



PRECISIÓN EN LOS EXPERIMENTOS AGRÍCOLAS



1^{RA} EDICIÓN

PRECISIÓN EN LOS EXPERIMENTOS AGRÍCOLAS

Algunas Consideraciones y Alternativas



1^{RA} EDICIÓN

PRECISIÓN EN LOS EXPERIMENTOS AGRÍCOLAS

Algunas Consideraciones y Alternativas

Freddy Gavilánez Luna

Autor



1^{RA} EDICIÓN

PRECISIÓN EN LOS EXPERIMENTOS AGRÍCOLAS

Algunas Consideraciones y Alternativas

AUTOR

Freddy Carlos Gavilánez Luna

Magíster en Estadística Aplicada;

Magíster en Riego y Drenaje;

Doctor en Ciencias Ambientales;

Ingeniero Agrónomo;

Universidad Agraria del Ecuador;

Guayaquil, Ecuador;

 fgavilanez@uagraria.edu.ec

 <https://orcid.org/0000-0002-7861-514X>

1^{RA} EDICIÓN

PRECISIÓN EN LOS EXPERIMENTOS AGRÍCOLAS

Algunas Consideraciones y Alternativas

REVISORES

ACADÉMICOS

Aldo José Loqui Sánchez

Doctor en Educación;
Magíster en Riego y Drenaje;
Docente Facultad de Ciencias Agrarias;
Universidad de Guayaquil;
Guayaquil, Ecuador

✉ aldo_loqui@hotmail.com

🆔 <https://orcid.org/0000-0001-8953-5105>

Yhony Alfredo Valverde Lucio

Ingeniero Agropecuario;
Magíster en Gestión de Proyectos socio productivos;
Especialista en Metodología de investigación;
Doctorante de Biociencias agroalimentarias;
Universidad Estatal del Sur de Manabí;
Jipijapa, Ecuador;

✉ yhonny.valverde@unesum.edu.ec

🆔 <https://orcid.org/0000-0002-9792-9400>

CATALOGACIÓN BIBLIOGRÁFICA

AUTOR: Freddy Gavilánez Luna

Título: Precisión en los experimentos agrícolas

Subtítulo: Algunas consideraciones y alternativas

Descriptor: Agricultura; Investigación agrícola; Biotecnología

Código UNESCO: 31 Ciencias Agrarias

Clasificación Decimal Dewey/Cutter: 630/G244

Área: Ingeniería Agrícola

Edición: 1^{ra}

ISBN: 978-9942-654-06-9

Editorial: Mawil Publicaciones de Ecuador, 2024

Ciudad, País: Quito, Ecuador

Formato: 148 x 210 mm.

Páginas: 329

DOI: <https://doi.org/10.26820/978-9942-654-06-9>

URL: <https://mawil.us/repositorio/index.php/academico/catalog/book/105>

Texto para docentes y estudiantes universitarios

El proyecto didáctico: **Precisión en los experimentos agrícolas**, es una obra colectiva escrita por varios autores y publicada por MAWIL; publicación revisada bajo la modalidad de pares académicos y por el equipo profesional de la editorial siguiendo los lineamientos y estructuras establecidos por el departamento de publicaciones de MAWIL de New Jersey.

© Reservados todos los derechos. La reproducción parcial o total queda estrictamente prohibida, sin la autorización expresa de los autores, bajo sanciones establecidas en las leyes, por cualquier medio o procedimiento.



Usted es libre de:
Compartir — copiar y redistribuir el material en cualquier medio o formato.
Adaptar — remezclar, transformar y construir a partir del material para cualquier propósito, incluso comercialmente.

Director Académico: Lcdo. Alejandro Plúa Argoti

Dirección Central MAWIL: Office 18 Center Avenue Caldwell; New Jersey # 07006

Gerencia Editorial MAWIL-Ecuador: Mg. Vanessa Pamela Quishpe Morocho

Dirección de corrección: Mg. Yamara Galanton.

Editor de Arte y Diseño: Lic. Eduardo Flores, Arq. Alfredo Díaz

Corrector de estilo: Lic. Marcelo Acuña Cifuentes

1^{RA} EDICIÓN

PRECISIÓN EN LOS EXPERIMENTOS AGRÍCOLAS

Algunas Consideraciones y Alternativas

Dedicatoria



Dedicatoria

A mi imprescindible cimiento de antología: Betty, Samantha y Freddy...

A León y a la memoria de Carmen...

A mis queridos alumnos de la Universidad Agraria del Ecuador...

1^{RA} EDICIÓN

PRECISIÓN EN LOS EXPERIMENTOS AGRÍCOLAS

Algunas Consideraciones y Alternativas

Índices

Contenidos



Prólogo ----- 24

CAPÍTULO I

EXPERIMENTOS AGRÍCOLAS ----- 29
El diseño de experimentos ----- 30
Distribuciones experimentales ----- 32
Factores de estudio en la agronomía ----- 36
Experimentos de adaptabilidad de cultivos ----- 36
Experimentos sobre nutrición vegetal ----- 38
Experimentos relacionados con plaguicidas ----- 41
Experimentos relativos a ciertas prácticas culturales ----- 48

CAPÍTULO II

LA UNIDAD EXPERIMENTAL ----- 49
El tamaño de la parcela en los experimentos agrícolas ----- 50
Factores de influencia en la dimensión de parcelas experimentales ----- 51
Mapa de heterogeneidad de suelo ----- 54
Métodos analíticos aplicados en los ensayos en blanco
para establecer el tamaño de las unidades de muestreo ----- 56
Método de máxima curvatura ----- 56
Métodos multivariados en la definición de los tamaños de parcela ----- 59
Método de máxima curvatura bivariada ----- 60
Método de Hatheway ----- 61

CAPÍTULO III

TAMAÑO MUESTRAL ----- 64
Muestra estadística ----- 65
Consideraciones de la muestra estadística ----- 66
Tipos de muestra ----- 67
Muestras probabilísticas o aleatorias ----- 67
Muestras no probabilísticas o escogidas ----- 68
Tamaño de la muestra ----- 69
Factores que determinan el tamaño de una muestra ----- 69
Tamaño de muestra de poblaciones infinitas para estimar una media ----- 71
Tamaño de la muestra de poblaciones infinitas para estimar
un valor porcentual (proporción) ----- 73
Tamaño de la muestra de poblaciones finitas ----- 74
Tamaño de muestra para experimentos ----- 79
Tamaño de muestra para experimentos de dos tratamientos ----- 84
Tamaño de muestra para experimentos con más de dos tratamientos ----- 86
Tamaño de muestra por diferencia mínima significativa (intervalo de confianza) ----- 87
Recomendación empírica para definir el número de repeticiones ----- 89

Tamaño de muestra por el criterio empírico de los
grados de libertad del error experimental ----- 91

CAPÍTULO IV

VARIABLES Y ALGUNAS ALTERNATIVAS DE ANÁLISIS----- 95
Variables agronómicas de respuesta en los ensayos experimentales ----- 96
¿Qué es la unidad de muestreo?----- 107
Error experimental y el ajuste de datos----- 108
Importancia de los supuestos de normalidad, homocedasticidad
e independencia en la precisión experimental ----- 112
Ajuste o estabilización de la varianza intragrupos----- 122
Análisis de covarianza ----- 128
Análisis de varianza multivariado ----- 134
Análisis de varianza con submuestreo----- 145

CAPÍTULO V

DISTRIBUCIONES CON BLOQUES INCOMPLETOS ----- 150
Distribución en bloques incompletos balanceados ----- 151
Distribución en bloques incompletos parcialmente balanceados----- 159
Distribución en latice ----- 162
Distribución alfa - latice ----- 172
Diseños con doble bloqueo: alternativas a la distribución en cuadro latino----- 176
Distribuciones resolubles de bloques incompletos----- 178

CAPÍTULO VI

DISTRIBUCIONES CON BLOQUES ANIDADOS O MODIFICADOS ----- 183
Conjunto de experimentos replicados en el espacio (localidad) y en el
tiempo (temporada)----- 184
Experimentos replicados en el espacio (localidad) ----- 185
Experimentos replicados en el tiempo (temporada) ----- 192
Experimentos replicados en el espacio (localidad) y en el tiempo (temporada)----- 198
Distribución de bloques aumentados ----- 205
Distribución en cuadro latino con réplicas----- 208
Distribución en cuadro latino modificado ----- 216

CAPÍTULO VII

RECOMENDACIONES EN LOS EXPERIMENTOS AGRÍCOLAS ----- 222
Algunas consideraciones sobre el diseño de los experimentos agronómicos----- 223
Sobre el tipo de distribución experimental----- 223
Conveniencia de un experimento factorial sobre un ensayo de un solo factor por turno --- 225
Sobre los tratamientos----- 227
Sobre las variables de respuesta ----- 229
Sobre las herramientas estadísticas----- 230

Prueba de Kruskal - Wallis -----235
Prueba de Friedman -----239
Recomendaciones sobre las labores culturales en los experimentos-----242

CAPÍTULO VIII

ALGUNAS ALTERNATIVAS DE OPTIMIZACIÓN-----246
Arreglos experimentales para optimización -----247
Arreglos factoriales San Cristóbal-----249
Diseño central compuesto -----255
Experimento factorial 2^k con una sola repetición -----263
Experimentos 2^k con una sola repetición en la agronomía -----266
Experimentos factoriales fraccionados -----274
Factorial 2^k fraccionado a un medio-----276
Experimento factorial 2^k fraccionado a la cuarta parte-----284
Experimentos en bloques con efectos confundidos -----288
Experimento en franjas-----296

CAPÍTULO IX

UTILIDAD ECONÓMICA DE LOS TRATAMIENTOS EXPERIMENTALES -----305
Recomendaciones a partir de la rentabilidad de los tratamientos -----306
Definición de términos utilizados en la valoración económica de los tratamientos -----308
Tasa de retorno marginal -----310
Relación beneficio - costo-----313
Rendimiento óptimo económico -----315

ANEXOS-----317

BIBLIOGRAFÍA -----327

1^{RA} EDICIÓN

PRECISIÓN EN LOS EXPERIMENTOS AGRÍCOLAS

Algunas Consideraciones y Alternativas

Índices

Tablas



Tabla 1.1. Distribuciones experimentales básicas -----	33
Tabla 1.2. Escala cualitativa para valoración del control de malezas -----	42
Tabla 1.3. Escalas de severidad de la sigatoka negra en banano -----	45
Tabla 2.1. Coeficientes de variación para diferentes combinaciones de dimensiones parcela para un cultivo de yuca-----	57
Tabla 2.2. Definición del tamaño de parcela aplicando el test F de Fisher para diferencias significativas de $P < 0,05$ -----	59
Tabla 2.3. Dimensiones y tamaño de parcela (unidad de muestreo) para varios experimentos -----	62
Tabla 3.1. Tamaños de muestras para dos niveles de confianza y distintos errores admisibles considerando poblaciones infinitas -----	74
Tabla 3.2. Tamaños de muestra de poblaciones finitas para estimar proporciones considerando una varianza de $[0,5(1 - 0,5)]$ -----	75
Tabla 3.3. Esquema del análisis de varianza para un diseño completamente al azar. -----	93
Tabla 4.1. Datos de altura de planta (cm)-----	114
Tabla 4.2. Datos predichos o esperados -----	115
Tabla 4.3. Valores calculados para la elaboración de los gráficos de normalidad y homocedasticidad -----	116
Tabla 4.4. Análisis de varianza realizado en SPSS bajo un modelo de bloques completos al azar-----	119
Tabla 4.5. Análisis de normalidad realizado en SPSS -----	121
Tabla 4.6. Reporte del análisis de homocedasticidad de SPSS -----	122
Tabla 4.7. Variables con y sin ajuste a distribuciones normales -----	123
Tabla 4.8. Transformaciones sugeridas para estabilización de varianzas -----	124
Tabla 4.9. Datos de control de malezas a los 10dda (%)-----	126
Tabla 4.10. Análisis de varianza con los datos sin transformar -----	127
Tabla 4.11. Datos de control de malezas a los 10dda con transformación arcoseno -----	127
Tabla 4.12. Análisis de varianza con los datos transformados.-----	127
Tabla 4.13. Datos de producción en kg/racimo (Y) y diámetro del tallo (cm) como covariable (X) ³¹ -----	130
Tabla 4.14. Datos de longitud de guía (Y ₁) y diámetro del fruto (Y ₂)-----	135
Tabla 4.15. Estructura del análisis de varianza para una distribución completamente al azar con submuestreo-----	146
Tabla 4.16. Estructura de análisis de varianza para una distribución de bloques completos al azar con submuestreo -----	147
Tabla 4.17. Datos sobre la altura de planta (m) de banano con dos submuestras por parcela -----	148
Tabla 4.18. Análisis de varianza para una distribución de bloques completos al azar del ejemplo de aplicación 4.6 con submuestreo -----	148
Tabla 4.19. Análisis de varianza para una distribución de bloques completos al azar del ejemplo de aplicación 4.6 sin considerar el submuestreo-----	149

Tabla 5.1. Estructura del análisis de varianza para una distribución de bloques incompletos balanceados -----	155
Tabla 5.2. Datos de nitratos (ppm) de un ensayo bajo una DBIB del ejemplo de aplicación 5.1 -----	156
Tabla 5.3. Análisis de varianza para una DBIB del ejemplo de aplicación 5.1 -----	157
Tabla 5.4. Análisis de varianza obtenido con Minitab -----	158
Tabla 5.5. Medias ajustadas y error estándar de la media obtenidos con Minitab -----	158
Tabla 5.6. Aplicación de la prueba de Tukey ($p < 0,05$) realizada con Minitab -----	158
Tabla 5.7. Esquema de un experimento bajo una DBIPB -----	160
Tabla 5.8. Grupos de asociados en 3, 2 y 1 bloques -----	160
Tabla 5.9. Datos del rendimiento de ocho variedades de papa (kg/parcela) bajo una DBIPB -----	161
Tabla 5.10. Análisis de varianza de los datos del ejemplo de aplicación 5.2 sobre un experimento realizado bajo una DBIPB (salida de Minitab) -----	161
Tabla 5.11. Prueba de Tukey para las medias del ejemplo de aplicación 5.2 (salida de Minitab) -----	162
Tabla 5.12. Datos (kg/parcela) ordenados del ejemplo de aplicación 5.3 -----	166
Tabla 5.13. Análisis de varianza considerando una DBCA -----	166
Tabla 5.14. Análisis de varianza con los datos del ejemplo de aplicación 5.3 considerando una DLS -----	170
Tabla 5.15. Promedios corregidos del ejemplo de aplicación 5.3 -----	171
Tabla 5.16. Esquema de análisis de varianza para una distribución alfa – latice -----	172
Tabla 5.17. Análisis de varianza para el ejemplo de aplicación 5.4, considerando una distribución de bloques incompletos con arreglo alfa – latice -----	175
Tabla 5.18. Comparación de medias utilizando el test de Tukey ($p < 0,05$) para el ejemplo de aplicación 5.4 -----	175
Tabla 5.19. Datos de incremento de altura (cm) en plantas de banano de un experimento bajo un DR-C -----	176
Tabla 5.20. Análisis de varianza de los datos del ejemplo de aplicación 5.5 realizado bajo una distribución renglón – columna (cuadrado de Youden) -----	178
Tabla 5.21. Prueba de Tukey ($p < 0,05$) aplicado a las medias del ejemplo de aplicación 5.5 -----	178
Tabla 5.22. Análisis de varianza del ejemplo de aplicación 5.6 -----	181
Tabla 6.1. Esquema de análisis de varianza para un mismo experimento replicado en el espacio (localidad) bajo una DBCA -----	186

Tabla 6.2. Rendimientos de papa en experimento desarrollado en la Localidad 1 de la zona de estudio -----	187
Tabla 6.3. Rendimientos de papa en experimento desarrollado en la Localidad 2 de la zona de estudio -----	187
Tabla 6.4. Rendimientos de papa en experimento desarrollado en la Localidad 3 de la zona de estudio -----	187
Tabla 6.5. Sumatorias para análisis de localidad, según los bloques-----	188
Tabla 6.6. Sumatorias para análisis de localidad, según los tratamientos (variedades) -----	188
Tabla 6.7. Análisis de varianza con efecto de bloques anidados en localidades del ejemplo de aplicación 6.1 -----	189
Tabla 6.8. Resultados de los análisis de varianza por localidad del ejemplo de aplicación 6.1 -----	191
Tabla 6.9. Separación de medias por localidad con la aplicación de la prueba de Tukey-----	192
Tabla 6.10. Esquema de análisis de varianza para un mismo experimento replicado en el tiempo (temporada) bajo una DBCA-----	193
Tabla 6.11. Rendimientos de caña de azúcar (t/ha) obtenidos en la temporada 1 -----	194
Tabla 6.12. Rendimientos de caña de azúcar (t/ha) obtenidos en la temporada 2-----	194
Tabla 6.13. Rendimientos de caña de azúcar (t/ha) obtenidos en la temporada 3-----	194
Tabla 6.14. Sumatorias para obtención de los efectos individuales y de interacción entre tratamientos y bloques -----	194
Tabla 6.15. Sumatorias para obtención de los efectos individuales y de interacción entre temporadas y bloques -----	195
Tabla 6.16. Sumatorias para obtención de los efectos individuales y de interacción entre tratamientos y temporadas-----	195
Tabla 6.17. Análisis de varianza del ejemplo de aplicación 6.2-----	197
Tabla 6.18. Resultados de los análisis de varianza por temporada del ejemplo de aplicación 6.2 -----	197
Tabla 6.19. Separación de medias por temporada con la aplicación de la prueba de Tukey-----	197
Tabla 6.20. Esquema de análisis de varianza para un experimento replicado en las mismas localidades durante varias temporadas, utilizando una DBCA-----	198
Tabla 6.21. Esquema de análisis de varianza para un experimento replicado en diferentes localidades durante varias temporadas, utilizando una DBCA-----	199
Tabla 6.22. Rendimientos de maíz (kg/ha) en la localidad "Guamaní" en los años 1, 2 y 3 -----	200

Tabla 6.23. Rendimientos de maíz (kg/ha) en la localidad "Belén" en los años 1, 2 y 3-----	200
Tabla 6.24. Sumatorias para obtener la interacción de tratamientos por temporada-----	201
Tabla 6.25. Sumatorias para obtener la interacción de tratamientos por localidad-----	201
Tabla 6.26. Sumatorias para obtener la interacción de temporada por localidad-----	201
Tabla 6.27. Sumatorias para obtener la interacción de segundo orden entre tratamientos, localidades y temporadas-----	201
Tabla 6.28. Sumatorias para obtener la suma de cuadrados de bloques anidados en la interacción de localidad por temporada-----	202
Tabla 6.29. Análisis de varianza del ejemplo de aplicación 6.3-----	204
Tabla 6.30. Resultados de los análisis de varianza por localidad del ejemplo de aplicación 6.3-----	204
Tabla 6.31. Aplicación de la prueba de Tukey para la separación de medias por localidad del ejemplo de aplicación 6.3-----	205
Tabla 6.32. Detalle del cálculo del factor de ajuste de bloques (B_{aj})-----	207
Tabla 6.33. Rendimiento ajustado-----	207
Tabla 6.34. Análisis de varianza solo para obtención del cuadrado medio del error experimental-----	208
Tabla 6.35. Fórmulas para calcular el error estándar de medias para pruebas de comparación entre tratamientos-----	208
Tabla 6.36. Grados de libertad para diseños en cuadro latino replicados, según los casos A, B y C-----	211
Tabla 6.37. Formas de cálculo de las sumas de cuadrados para diseños en cuadro latino replicados, según los casos A, B y C-----	211
Tabla 6.38. Análisis de varianza del ejemplo de aplicación 6.5 sobre DCL con réplica bajo el caso C-----	214
Tabla 6.39. Tratamientos de un experimento bajo una DCLM-----	218
Tabla 6.40. Datos de peso de 100 semillas (g) de un experimento bajo una DCLM-----	219
Tabla 6.41. Análisis de varianza para el ejemplo de aplicación 6.6-----	221
Tabla 7.1. Magnitud de los grados de libertad del error experimental entre los análisis de varianza de una DCA y una DBCA-----	224
Tabla 7.2. Esquema del ANDEVA para el análisis de las tres dosis de NO ₃ K-----	225
Tabla 7.3. Esquema del ANDEVA para el análisis de las dosis de CO ₃ Ca-----	226
Tabla 7.4. Esquema del ANDEVA para el análisis del factorial A x B-----	226
Tabla 7.5. Fuentes de variación y modelos lineales de las distribuciones experimentales básicas-----	233

Tabla 7.6. Número de insectos muertos/planta -----	236
Tabla 7.7. Número de larvas de <i>S. frugiperda</i> muertas/planta -----	240
Tabla 8.1. Combinaciones factoriales en arreglo San Cristóbal para dos y tres factores -----	250
Tabla 8.2. Datos de rendimiento (kg/ha) de maíz para un arreglo San Cristóbal con fertilizantes -----	251
Tabla 8.3. Codificación de las dosis -----	252
Tabla 8.4. Análisis de varianza de los tratamientos (fórmulas) estudiados -----	252
Tabla 8.5. Descripción de la variables regresoras (efectos lineales, cuadráticos e interacción) y de la variable dependiente -----	252
Tabla 8.6. Estadísticas de la regresión -----	253
Tabla 8.7. Análisis de varianza para la regresión -----	253
Tabla 8.8. Coeficientes del modelo de regresión y las pruebas de significancia para cada uno de ellos -----	253
Tabla 8.9. Clasificación de promedios del rendimiento (kg/ha), según el test de Tukey -----	254
Tabla 8.10. Análisis de varianza con separación de factores -----	254
Tabla 8.11. Niveles bajo y alto de los fertilizantes -----	258
Tabla 8.12. Arreglo factorial bajo el DCC con sus niveles y sus variables codificadas -----	258
Tabla 8.13. Rendimiento (t/ha) bajo el arreglo DCC -----	259
Tabla 8.14. Combinaciones factoriales (tratamientos) e interacciones dobles del tipo 2^5 -----	265
Tabla 8.15. Número de tratamientos (combinaciones factoriales), efectos principales y efectos secundarios generados por experimentos factoriales del tipo 2^k , hasta $k = 5$ factores -----	267
Tabla 8.16. Contrastes, efectos y sumas de cuadrados para cada factor e interacción doble -----	269
Tabla 8.17. Análisis de varianza inicial -----	270
Tabla 8.18. Análisis de varianza definitivo -----	272
Tabla 8.19. Promedios del rendimiento (kg/Parcela) por cada nivel factorial -----	273
Tabla 8.20. Cuantificación de los efectos producidos por varios experimentos del tipo 2^k -----	275
Tabla 8.21. Fracción principal del experimento factorial 2^{4-1} -----	279
Tabla 8.22. Experimento factorial fraccionario 2^{5-1} con generador $I = +ABCDE$ y los rendimientos del cultivo en unidades de kg/parcela -----	279
Tabla 8.23. Contrastes y sumas de cuadrados para cada efecto principal e interacción doble -----	280
Tabla 8.24. Análisis de varianza considerando una sola réplica para un factorial fraccionado a la mitad del tipo 2^{5-1} con 8 efectos principales en el error experimental -----	282

Tabla 8.25. Análisis de varianza considerando una sola réplica para un factorial fraccionado a la mitad del tipo 2^{5-1} con nueve efectos principales en el error experimental -----	282
Tabla 8.26. Niveles para un experimento factorial 2^{5-2} -----	285
Tabla 8.27. Efectos aliados del factorial fraccionado 2^{5-2} -----	285
Tabla 8.28. Efectos aliados del factorial fraccionado 2^{5-2} hasta las interacciones dobles -----	286
Tabla 8.29. Matriz para un factorial fraccionado 2^{5-2} -----	286
Tabla 8.30. Contrastes, efectos y sumas de cuadrados (SC) para cada efecto principal de un factorial fraccionado 2^{5-2} -----	287
Tabla 8.31. Análisis de varianza considerando una sola réplica para un factorial fraccionado 2^{5-2} -----	288
Tabla 8.32. Arreglos ortogonales para un experimento factorial 2^2 -----	289
Tabla 8.33. Formación de bloque con efecto confundido para las interacciones ABC y ABCD en experimentos del tipo 23 y 24, respectivamente -----	291
Tabla 8.34. Bloques de confusión para un experimento 24, confundiendo primero la interacción ABCD y luego la interacción BCD -----	292
Tabla 8.35. Factores y niveles del ejemplo de aplicación 7.6 -----	292
Tabla 8.36. Tratamientos del ejemplo de aplicación 8.6 -----	292
Tabla 8.37. Arreglos ortogonales* para un experimento con tres factores a dos niveles cada uno -----	293
Tabla 8.38. Datos del rendimiento de arroz (kg/parcela) del ejemplo de aplicación 8.6. -----	294
Tabla 8.39. Sumas de cuadrados para los factores y las interacciones del ejemplo de aplicación 8.6 -----	295
Tabla 8.40. Análisis de varianza para los factores principales e interacciones dobles con la confusión del efecto de la interacción ABC -----	295
Tabla 8.41. Análisis de varianza sin la confusión del efecto de la interacción triple ABC, bajo una distribución de bloques completos al azar -----	296
Tabla 8.42. Niveles de los factores evaluados -----	298
Tabla 8.43. Matriz de datos organizados del ejemplo 8.7 -----	299
Tabla 8.44. Matriz de la suma de datos de cada nivel del factor A de acuerdo a cada repetición -----	299
Tabla 8.45. Matriz de la suma de datos de cada nivel del factor B de acuerdo a cada repetición -----	299
Tabla 8.46. Matriz de las interacciones de los factores A y B -----	300
Tabla 8.47. Análisis de varianza el ejemplo de aplicación 8.7 sobre un experimento en franjas -----	303
Tabla 8.48. Análisis de varianza del ejemplo de aplicación 8.7 realizado con el software Infostat -----	304

Tabla 9.1. Ejemplo de Presupuesto Parcial para un experimento sobre la valoración del efecto de 5 dosis de nitrógeno en arroz -----	311
Tabla 9.2. Tasa de retorno marginal-----	312
Tabla 9.3. Relación beneficio/costo de los tratamientos -----	313
Tabla 1A. Áreas de curva normal estándar con escala Z. -----	318
Tabla 2A. Valores críticos de la distribución t de student para distintos niveles de significancia α -----	319

1^{RA} EDICIÓN

PRECISIÓN EN LOS EXPERIMENTOS AGRÍCOLAS

Algunas Consideraciones y Alternativas

Índices

Figuras



Figura 2.1 Esquema de líneas de igual fertilidad en un mapa de heterogeneidad de suelo-----	55
Figura 2.2. Relación entre el coeficiente de variación y las dimensiones (ancho x largo) de las unidades de muestreo -----	57
Figura 2.3. Determinación de parcela experimental (unidad de muestreo) para ensayos de arroz -----	58
Figura 2.4. Descripción gráfica de la metodología de Hatheway -----	62
Figura 3.1. Esquema sobre la variabilidad del tamaño de una muestra n respecto del tamaño de una población N conocida-----	70
Figura 3.2. Error máximo admisible o distancia entre la media muestral y la media de su población infinita correspondiente-----	71
Figura 3.3. Ubicación del valor crítico de la variable y su correspondiente escala Z -----	82
Figura 3.4. Magnitud de la potencia -----	83
Figura 3.5. Incremento de la potencia de la prueba por el incremento del tamaño muestral -----	84
Figura 3.6. Esquema de la diferencia mínima entre dos promedios a partir de la cual se establece variación significativa entre dos tratamientos (DMS)-----	88
Figura 3.7. Esquema de la relación entre la varianza intragrupos (CM_E) y los grados de libertad del error experimental-----	92
Figura 4.1. Esquema de la medición de altura de plantas -----	99
Figura 4.2. Esquema de la formación de macollos en un cereal	
Figura 4.3. Detalle de la unidad de muestreo dentro de una unidad experimental (parcela)-----	101
Figura 4.4. Gráfica de probabilidad normal-----	116
Figura 4.5. Gráfica para verificar homocedasticidad-----	118
Figura 4.6. Secuencia en el software SPSS para realizar el análisis de varianza y la verificación gráfica de la normalidad y homocedasticidad-----	119
Figura 4.7. Gráfica de residuos reportado por el software SPSS para verificar normalidad y homocedasticidad-----	120
Figura 4.8. Secuencia en el software SPSS para realizar el análisis de normalidad con las pruebas de Kolmogorov – Smirnov y Shapiro – Wilk -----	121
Figura 4.9. Secuencia en el software SPSS para realizar el análisis de homocedasticidad mediante el test de Levene -----	122
Figura 4.10. Ingreso de datos en el Infostat -----	131
Figura 4.11. Secuencia del análisis en Infostat-----	131
Figura 4.12. Secuencia para la realización del análisis de varianza en Infostat -----	132
Figura 4.13. Secuencia para la realización del análisis de covarianza en Infostat -----	132
Figura 4.14. Proceso para la obtención de correlación entre las variables longitud de guía y diámetro del fruto-----	136

Figura 4.15. Secuencia para la obtención del análisis de varianza multivariado (ANDEVAM) -----	137
Figura 4.16. Clasificación de las variables dentro de la pestaña para la realización del ANDEVAM -----	137
Figura 4.17. Definición del modelo del ANDEVAM, considerando una distribución de bloques completos al azar -----	138
Figura 4.18. Selección del test de Tukey dentro de la pestaña <i>Post hoc</i> -----	140
Figura 4.19. Secuencia de pasos para verificar la normalidad mediante el test de Shapiro - Wilk en SPSS -----	142
Figura 4.20. Definición del modelo, considerando sólo el factor principal para verificar homocedasticidad -----	143
Figura 4.21. Selección de la pestaña para el análisis de homocedasticidad -----	144
Figura 5.1. Esquema de una distribución de bloques incompletos balanceados (DBIB) con cuatro tratamientos aleatorizados en tres unidades experimentales por bloque -----	153
Figura 5.2. Estructura de los datos de un experimento alfa – latice previo al análisis en un software estadístico -----	174
Figura 5.3. Escala lineal tipo hedónica -----	179
Figura 5.4. Esquema parcial de la matriz de datos para su análisis en un software -----	180
Figura 5.5. Análisis gráfico de residuos para los datos del ejemplo de aplicación 5.6 -----	181
Figura 6.1. Esquema de un experimento de seis tratamientos, bajo una distribución de bloques completos al azar y replicado en tres localidades (bloques anidados en localidad) -----	185
Figura 6.2. Esquema de la base de datos para la realización del análisis en un software estadístico -----	190
Figura 6.3. Salida del software Infostat, con el análisis de varianza y la prueba de Tukey -----	191
Figura 6.4. Esquema de la distribución de bloques aumentados del ejemplo de aplicación 6.4. -----	206
Figura 6.5. Esquema de un diseño en cuadro latino con cuatro tratamientos y dos réplicas, utilizando las mismas filas o las mismas columnas (a); o con diferentes filas y columnas (b) -----	210
Figura 6.6. Estructura de un diseño cruzado de tres tratamientos aplicados bajo una distribución de cuadro latino (3 x 3) con dos réplica -----	216
Figura 6.8. Aleatorización de los grupos por columna en el diseño DCLM -----	217
Figura 6.9. Croquis de ubicación de 12 tratamientos en una DCLM -----	218
Figura 6.10. Croquis de ubicación de los 18 tratamientos del ejemplo de aplicación 6.6. -----	219
Figura 7.1. Verificación gráfica de la falta de normalidad y de homocedasticidad -----	237
Figura 7.2. Ingreso de datos a Infostat para la realización del test de Kruskal – Wallis -----	238
Figura 7.3. Secuencia de pasos en el Infostat para la realización del test de Kruskal-Wallis -----	238

Figura 7.3. Estructura de la base de datos en Infostat para la realización del test de Friedman-----	241
Figura 7.4. Ubicación del test de Friedman en el software Infostat-----	241
Figura 8.1. Matriz de puntos experimentales en un arreglo San Cristóbal para dos factores-----	249
Figura 8.2. Matriz de puntos experimentales en un arreglo San Cristóbal para tres factores-----	250
Figura 8.3. Esquema de la matriz de puntos experimentales en un arreglo central compuesto para dos y tres factores-----	256
Figura 8.4. Superficie de contorno y de respuesta para los factores nitrógeno y fósforo-----	261
Figura 8.5. Superficie de contorno y de respuesta para los factores nitrógeno y potasio-----	261
Figura 8.6. Superficie de contorno y de respuesta para los factores fósforo y potasio-----	262
Figura 8.7. Identificación de niveles para los cuales el rendimiento es máximo-----	263
Figura 8.8. Indicación del postulado "80/20" de Pareto-----	272
Figura 8.9. Promedios registrados por cada nivel del factor B-----	273
Figura 8.10. Promedios registrados por la interacción de los factores A y C-----	273
Figura 8.11. Promedios registrados por la interacción de los factores A y D-----	274
Figura 8.12. Esquema de un diseño factorial 2^3 fraccionado a la mitad-----	276
Figura 8.13. Diagrama de Pareto para identificar los efectos de importancia-----	281
Figura 8.14. Efectos principales de los factores A, B, C y D-----	283
Figura 8.15. Efectos de las interacciones dobles AC y CE-----	284
Figura 8.16. Diagrama de Pareto para identificar los efectos principales en un factorial 2^{5-2} -----	287
Figura 8.17. Efectos principales por niveles-----	288
Figura 8.18. Esquema de un bloque (repetición) de un experimento factorial en franjas-----	297
Figura 8.18. Esquema de la estructura de datos y del modelo respectivo para un experimento en franjas en el software Infostat-----	303
Figura 9.1. Curva de dosis de nitrógeno versus beneficio neto-----	312
Figura 9.2. Curva de dosis de nitrógeno versus la relación B/C-----	314
Figura A1. Curva de operación característica para experimentos de dos tratamientos, como prueba de dos colas y 0,05 de significancia (Montgomery, 2005)-----	320
Figura A2. Curva de operación característica para el análisis de varianza, bajo el criterio de efectos fijos, con 1 grado de libertad en el numerador (V_1) y con 6 o más grados de libertad en el denominador (error experimental) (V_2) (Montgomery, 2005).-----	320

.....

Figura A3. Curva de operación característica para el análisis de varianza, bajo el criterio de efectos fijos, con 2 grados de libertad en el numerador (V_1) y con 6 o más grados de libertad en el denominador (error experimental) (V_2) (Montgomery, 2005).....	321
Figura A4. Curva de operación característica para el análisis de varianza, bajo el criterio de efectos fijos, con 3 grados de libertad en el numerador (V_1) y con 6 o más grados de libertad en el denominador (error experimental) (V_2) (Montgomery, 2005).....	321
Figura A5. Curva de operación característica para el análisis de varianza, bajo el criterio de efectos fijos, con 4 grados de libertad en el numerador (V_1) y con 6 o más grados de libertad en el denominador (error experimental) (V_2) (Montgomery, 2005).....	322
Figura A6. Curva de operación característica para el análisis de varianza, bajo el criterio de efectos fijos, con 5 grados de libertad en el numerador (V_1) y con 6 o más grados de libertad en el denominador (error experimental) (V_2) (Montgomery, 2005).....	322
Figura A7. Curva de operación característica para el análisis de varianza, bajo el criterio de efectos fijos, con 6 grados de libertad en el numerador (V_1) y con 6 o más grados de libertad en el denominador (error experimental) (V_2) (Montgomery, 2005).....	323
Figura A8. Curva de operación característica para el análisis de varianza, bajo el criterio de efectos fijos, con 7 grados de libertad en el numerador (V_1) y con 6 o más grados de libertad en el denominador (error experimental) (V_2) (Montgomery, 2005).....	323
Figura A9. Curva de operación característica para el análisis de varianza, bajo el criterio de efectos fijos, con 8 grados de libertad en el numerador (V_1) y con 6 o más grados de libertad en el denominador (error experimental) (V_2) (Montgomery, 2005).....	324
Figura A10. Estructuras en latice (cuadrados) balanceados para experimentos con 9, 16 y 25 tratamientos (Cochran y Cox, 1978).....	325
Figura A11. Estructuras en latice rectangulares no balanceados para experimentos con 12, 20 y 30 tratamientos (Cochran y Cox, 1978).....	326

1^{RA} EDICIÓN

PRECISIÓN EN LOS EXPERIMENTOS AGRÍCOLAS

Algunas Consideraciones y Alternativas

Prólogo



PRÓLOGO

La agricultura es una de las actividades humanas más diversas, la cual, por necesidad, debe cambiar continuamente para hacer producir los cultivos de la forma más eficientemente posible. En esta búsqueda por mejorar los rendimientos de las cosechas, toma contundencia a inicios del siglo XX, desde la perspectiva *fisheriana*, una de las herramientas del método científico fundamental, como es la experimentación. Los experimentos agrícolas, también llamados ensayos agrícolas, llevados a cabo mediante el debido protocolo, son los que han permitido establecer muchas alternativas aplicables a la agricultura y además han venido mejorándose continuamente. Por ello, el rigor científico de los experimentos en su planificación protocolaria es lo que los caracteriza dentro de la eficiencia y la objetividad, lo que a su vez está ligado a la precisión experimental; esta última, un componente que debe ser fundamental y subyacente de los experimentos para que éstos tengan la eficacia necesaria en la detección de los efectos de las alternativas de prueba.

La precisión de los experimentos es uno de los aspectos fundamentales que hace posible detectar diferencias entre los distintos tratamientos que se evalúan en un ensayo, cuya máxima ocurrencia depende básicamente de un oportuno diseño, de un adecuado manejo experimental y de un consecuente procedimiento para la medición de las variables de respuesta. En este sentido, un experimento preciso nos protege de caer en decisiones erradas como la de negar una alternativa que en un escenario determinado tiene efectos reales (error tipo II), sin embargo, por la falta de control de la varianza interna del ensayo son obviados a nivel experimental. Esta falta de precisión crea así un falso negativo que desechará, posiblemente, alguna oportunidad de mejorar la productividad de algún cultivo o, incluso, de afianzar o refutar una teoría fundada.

Un experimento preciso permite lograr eficiencia en los recursos que se destinan a la investigación, ya que al poder discernir claramente los efectos de las distintas alternativas que se prueban en un ensayo desde una primera experiencia, es posible acortar los tiempos e ir delineando relativamente rápido la vía de un proceso de investigación; evitando así el desarrollo de experimentos adicionales con fines de confirmación que, por lo general, hacen muy costoso el proceso. Al tener experimentos sensibles, con un control adecuado del error experimental, indirectamente es posible ampliar los programas de investigación a nivel institucional dentro de cualquier ciencia, pues quedan recursos que podrían invertirse en ensayos de otros procesos de indagación.

En pro de esa necesaria precisión experimental en los ensayos agrícolas se ha escrito este texto, explicando en sus nueve capítulos varios métodos y algunas consideraciones que el investigador podría tomar en cuenta, tanto en la planeación y desarrollo como en la toma, análisis e interpretación de sus datos. Es así como también se han incluido metodologías de la experimentación industrial, utilizados en optimización de procesos que, a pesar de tener fases no tan prácticas para aplicarse en la agricultura, se detallan para que se reconozcan componentes importantes de éstas. En este sentido se expone, por ejemplo, un caso de los diseños de segundo orden para generar superficies de respuesta, con el fin de que el investigador pueda adaptarlas parcialmente a la experimentación agrícola y así tomar decisiones más precisas.

Se inicia con el capítulo I haciendo una explicación sobre las distribuciones experimentales¹, sus fundamentos de uso y una descripción sobre los distintos tipos de experimentos que se desarrollan en la agronomía y que podrían aportar a una mayor precisión en los resultados. Con esta perspectiva, es preciso indicar que, si bien el manejo de un ensayo es de mucha

¹ Se prefiere utilizar en este texto la definición de DISTRIBUCIÓN EXPERIMENTAL haciendo referencia a lo que comúnmente se llama DISEÑO EXPERIMENTAL, dado que, bajo el criterio del autor, estos términos holísticamente son dos cosas distintas.

importancia, la selección de una distribución experimental que esté en consecuencia con un determinado ambiente en donde se desarrollará la experiencia puede ser incluso mucho más importante que dicho manejo. No hay que olvidar que, dentro del diseño o planeación de un experimento, la selección de la distribución se hace antes que el manejo.

En el capítulo II del texto se explica detalladamente sobre las diferentes metodologías para definir el tamaño de las unidades experimentales, pasando desde los mapas de heterogeneidad del suelo hasta la descripción del método de Hatheway; temáticas que, en la bibliografía a nivel regional, y especialmente a nivel local, son muy escasas, siendo de enorme importancia en la precisión de los experimentos. En el capítulo III se exponen las técnicas para seleccionar tanto los diferentes tipos de muestreo como la definición de los tamaños muestrales, incluyéndose métodos para investigaciones observacionales, así como para las de tipo experimental.

En el capítulo IV se señalan algunas consideraciones sobre las variables más comúnmente valoradas en los ensayos. También se precisan algunas alternativas para verificar normalidad y homocedasticidad de datos, además de las técnicas de ajustes para estabilizar las varianzas, utilizando de soporte algunos programas de estadística como *Infostat*, *IBM SPSS statistics* y *Minitab*. Igualmente, se detalla una alternativa adicional de ajuste como lo es el análisis de covarianza; así como el análisis de varianza multivariado para utilizarse cuando se tienen variables dependientes relacionadas que podrían involuntariamente incrementar, en la práctica, la probabilidad del error tipo I. Como parte final de este capítulo se incluye la descripción de un modelo de análisis de varianza que considera la variabilidad debido al submuestreo, presentado como otra alternativa para minimizar el error experimental al sustraer la variabilidad que ocurre al interior de las propias unidades experimentales.

Algunas de las distribuciones en bloques incompletos se tratan en el capítulo V. Este apartado describe varias alternativas de uso en cuanto a distribuciones que utilizan bloques que incluyen sólo una parte de un número elevado de tratamientos que se prevén evaluar en un ensayo, cuya característica principal es la de poder controlar de forma eficaz la heterogeneidad del suelo; a diferencia de los ensayos con bloques completos que pierden precisión al ampliarse extensamente en el campo. Se detallan temáticas relacionadas a distribuciones con bloques incompletos balanceados y parcialmente balanceados, distribuciones en látice, distribuciones en alfa - látice, algunas alternativas de distribución de cuadro latino y terminando con una explicación rápida sobre distribuciones resolubles de bloques incompletos. Complementariamente, en el capítulo VI se incluyen algunas distribuciones con bloques anidados por características de replicación en espacio y tiempo, experimentos exploratorios con bloques aumentados, además de cuadrados latinos replicados y modificados.

En el capítulo VII se especifican algunas consideraciones a tener en cuenta sobre el diseño de experimentos agrícolas, relacionadas al propio tipo de distribución experimental, a la conveniencia de los arreglos factoriales, a los tratamientos, a las variables de respuesta y a las herramientas estadísticas. Así también, se incluye información sobre las pruebas no paramétricas, análogas a los análisis de varianza para una distribución completamente al azar y de bloques completos al azar; concluyendo con algunas indicaciones sobre las labores culturales que se realizan en los experimentos.

Subsecuentemente, en el capítulo VIII se describen algunos arreglos experimentales (o diseño de tratamientos factoriales) para optimización, tales como los factoriales incompletos San Cristóbal y el Central Compuesto; este último utilizado en la metodología de Superficie de Respuesta. Igualmente, se hace una exposición sobre los factoriales 2^K con una sola réplica y fraccionados, los cuales son utilizados a nivel exploratorio y en procesos de optimización; siendo todos, arreglos factoriales que pueden ser adaptados para utilizarse en los ensayos agrícolas. También se incluye la temática de los factoriales que se realizan en bloques con

efecto de confusión (también con bloques incompletos), como alternativas de control de la heterogeneidad del suelo. Asimismo, se ha visto conveniente incluir como último tema de este capítulo lo referente a los experimentos en franjas, alternativa ideada para experiencias en las cuales se valoran dos factores que tienen cierta restricción al ser distribuidos dentro de las unidades experimentales, de allí que se ubiquen en franjas para un mejor control de los efectos de traslape entre tratamientos.

Finalmente, el capítulo IX se ha estructurado tomando en consideración el gran dilema entre lo que es significancia estadística y significancia práctica, definida esta última a través del beneficio económico que pueden ofrecer los tratamientos de un ensayo. En forma breve, se hace una descripción de la metodología de la tasa de retorno marginal como una opción para definir la recomendación de tratamientos de un ensayo en función de la utilidad económica de cada uno de ellos; además, se incluyen temas como la relación beneficio-costos y lo concerniente al rendimiento óptimo económico. La necesidad de incluir esta temática viene dada por su importancia en gran parte de los experimentos agrícolas, ya que éstos comúnmente se desarrollan en ambientes altamente volubles en espacio y tiempo, por consiguiente, podría ocurrir que alternativas con efectos estadísticamente significativos insoportablemente deben relegarse porque no presentan ninguna ventaja en términos de rentabilidad para el agricultor.

No es de menos indicar que este documento se ha propuesto para contrarrestar ciertos criterios arraigados que muchas veces restan precisión y/o exactitud a un ensayo, como el de creer que todo dato porcentual debe ser sometido a un ajuste angular, o el de medir el número de macollos por m² en siembras de arroz directa cuando no se lo ha voleado adecuadamente, o incluso el de seguir utilizando una distribución de bloques en un mismo campo cuando en experimentos previos no se han detectado diferencias significativas en la fuente de bloqueo. Estos son solo algunos de los elementos que han servido de fundamento para la redacción de este texto.

La obra se ha concebido como un complemento al diseño de experimentos agrícolas, aportando con descripciones conceptuales y de ejemplos en algunos casos, de tal forma que aquel que realiza investigación agrícola adquiera y/o amplíe el razonamiento integral en la definición de los ensayos, partiendo del supuesto de que el lector tiene un conocimiento previo del diseño experimental en esta ciencia. Se ha tratado de abarcar, al menos, las más importantes alternativas y metodologías a las que puede recurrir quien diseña un experimento agrícola, con el propósito de que los resultados experimentales tengan la mayor precisión posible. Queda a disposición esta humilde obra, esperando que sea un aporte importante en la planeación, desarrollo y éxito de sus experimentos.

Freddy Gavilánez Luna

Todo experimento eficiente es eficaz, pero no todo experimento eficaz es eficiente. La eficacia sólo depende de la precisión, la eficiencia de la precisión y de la exactitud. ¡Te allí la importancia de la base teórica, de la estructura del diseño y del manejo de un experimento! ¡Te allí la diferencia respecto de una aparente tautología!!

CAPÍTULO I

EXPERIMENTOS AGRÍCOLAS

El éxito en la precisión de un experimento se fundamenta en su adecuado diseño y en el rigor científico de su manejo, entendiéndose como éxito que dicha precisión sea absoluta y no relativa. Por lo tanto, también es fundamental la base teórica, empírica e intuitiva del investigador...



El diseño de experimentos

Cuando se ha identificado una problemática, sea ésta de aplicación práctica o de conocimiento y se decide buscar su solución evaluando una o varias alternativas de prueba, en esas circunstancias se está experimentando. Desde esta perspectiva, un experimento se realiza para identificar una posible solución que se establece al momento de evaluar los efectos de unos tratamientos, cuya objetividad y precisión se logran cuando se lo lleva a cabo con rigor científico; esto quiere decir que, tanto su planeación como su desarrollo, así como los resultados y conclusiones a los que se arribe, se fundamentan en un protocolo metodológico previamente establecido y bajo la concepción del método científico. Esta acepción contextual de experimento difiere con los de características netamente empíricas, que también permiten arribar a resultados, pero éstos en la mayoría de las veces no tienen una base científicamente explicativa que produzcan un conocimiento integral.

La agronomía tal vez sea una de las ciencias que más requiere de la realización de experimentos, debido al ambiente tan diverso sobre el que se desarrollan los cultivos y a la demanda constante de incrementar la producción de alimentos bajo términos de eficiencia. Sin embargo, y en concordancia con lo indicado en el párrafo anterior, un experimento requiere toda una parafernalia doctrinal para su diseño; partiendo desde la verificación del área experimental y su ambiente, la definición de cuáles y cuántos serán los tratamientos, sus repeticiones, el tipo de distribución, las variables a medirse, el manejo del cultivo durante la experiencia, hasta el tipo de herramienta estadística a utilizarse. De allí que cuando se habla de diseño experimental no sólo comprende el tipo de distribución a emplearse, sino el establecimiento y la definición de los límites de todos los elementos antes mencionados que permitan, a su vez, determinar claramente el efecto de los tratamientos de cuyas respuestas se tiene una idea hipotética.

Diseñar un experimento agrícola (o en cualquier otra ciencia) es contextualizar y conceptualizar una planificación a seguir, originado de una problemática para la cual se plantean unos objetivos fundamentados en consideraciones hipotéticas que vienen a constituirse en el impulso del investigador, a través de las cuales prevé conjugar la teoría y los hechos, de tal forma que se pueda solucionar dicha problemática. En esta sistematización del método científico se busca pasar del planteamiento deductivo al inductivo, realizando recomendaciones generales para los agricultores que les permitan mejorar productiva y rentablemente su actividad. Para este fin, el diseño de un experimento debe ser eficiente en el sentido de controlar el error experimental y eficaz para la comprobación de las hipótesis planteadas (Pantoja, 1989); lo que consecuentemente deberá permitirle al investigador extraer la mayor información posible de un ensayo con la mínima inversión económica, dado lo costoso que puede resultar una investigación y las limitaciones económicas que siempre existen.

El diseño de experimentos es una de las herramientas principales del método científico que se ha establecido con el propósito de producir conocimiento o de resolver problemas específicos, el cual, bajo la focalización de este método, se desarrolla dentro de límites lógicos y de objetividad. En este sentido, su fundamento reside en la relación de los métodos deductivo-inductivo establecidos por Galileo Galilei en el siglo XVI y cuya aplicación práctica, considerando elementos de control experimental como la aleatorización, empezó a difundirse a inicios del año 1930 por Ronald Fisher. El diseñar un experimento consiste en una definición o cambio deliberado de alternativas a probarse en un proceso productivo como el agrícola (o cualquier otro), bajo parámetros rigurosos que permitan tener la certeza de que el cambio que causen dichas alternativas de prueba no sea producto de efectos casuales. Dada la complejidad de fenómenos y ambientes sumamente variables que rodean a las ciencias biológicas, antes que en otras ciencias, la ortodoxia del método científico tal vez sea requerida con mayor intensidad en esta área si queremos alcanzar resultados contundentes.

Aparte del protocolo en el manejo experimental, para complementar el rigor científico de los ensayos en general, éstos deben desarrollarse bajo tres principios fundamentales. El primero de ellos es la *aleatorización* en la asignación de los tratamientos a las unidades experimentales (o viceversa), lo cual protege al investigador de caer fácilmente en el error I (α); el mismo que, según se vaya explicando en párrafos y capítulos posteriores, podría ser el más costoso dentro de la investigación agrícola. Además, la aleatorización es la que procura la independencia de los residuos, uno de los tres supuestos que demanda la aplicación de pruebas de la estadística paramétrica.

El segundo principio del diseño de experimentos es la *repetición*. Si los tratamientos de un experimento no tienen repeticiones se hace imposible valorar efectivamente el error experimental, de allí que sería muy riesgoso confiar en las respuestas que se obtengan en este tipo de experimento. Sólo las repeticiones, que no deben confundirse con réplicas de medición², son las que ofrecen precisión matemática a un experimento (la precisión empírica viene del manejo del ensayo), ya que permiten controlar el error II (β) y aumentar la potencia estadística de una prueba. No obstante, cuando se realizan experimentos de tipo exploratorios con más de un factor de estudio, utilizando sólo dos niveles por factor y que no sean tan comunes en la agronomía, puede optarse por realizar una sola repetición; en este caso, se utilizan los efectos de las interacciones triples o de mayor grado, que generalmente resultan no significativos y complicados de explicar, para formar la fuente del error experimental. Corrientemente, lo último puede realizarse cuando el ensayo contiene cuatro o más factores de exploración.

Finalmente, el tercer principio es la formación de *bloques* (bloqueo) en un experimento, también llamado en algunos casos como *control local*. Este principio específicamente está referido a distribuciones que se aplican en ambientes no controlados, tales como las de bloques completos e incompletos al azar o cuadrados latinos que se utilizan en ensayos de campo; y por su requerimiento de homogeneidad en las condiciones de suelo, no rige para las distribuciones completamente al azar. Un ejemplo de la necesidad de bloqueo experimental se da cuando se realiza un ensayo en un suelo con una pendiente, en donde puede existir una diferencia nutricional que influya de manera importante en los resultados; situación ante la cual una distribución como la de bloques al azar permitirá tener un mejor control de esta fuente de variación secundaria, reduciendo la influencia de dicho efecto al formar pequeños bloques homogéneos, de tal forma que esta variabilidad se vea reflejada entre dichos bloques y no entre los tratamientos.

Es oportuno indicar que, en la búsqueda de precisión en los experimentos, la planificación o diseño de éstos tomando en cuenta todas las previsiones del caso es fundamental. Por ello, su estructura debe ampararse bajo la percepción de gente que, por un lado, entienda sobre el fenómeno de estudio y, por otro, lo pueda describir. Los investigadores que estarán al frente de un ensayo podrán tener el conocimiento sobre lo que están indagando, pero muchas veces la estructura o el marco estadístico del experimento no lo definen claramente; o, en el peor de los casos, no consideran ciertos aspectos del diseño que al final lo pueden hacer muy impreciso. En este caso, la asistencia de un biometrista³ al momento de diseñar el experimento

² No debe confundirse la diferencia elemental entre la repetición de un tratamiento y lo que se llama réplica de medición. Una repetición viene a constituirse en una unidad experimental (parcela) sobre la que se aplica un tratamiento y son las que permiten estimar el error experimental en un ensayo cuando se tienen dos o más repeticiones; mientras que una réplica no es más que medir una variable dos o más veces dentro de una misma unidad experimental. Ejemplos prácticos de réplicas son las mediciones de varias plantas de una misma unidad experimental para obtener la altura de planta, cuya media (estimador de su valor) de dichas réplicas constituye el dato observado. Las repeticiones o unidades experimentales de los tratamientos son las que hacen posible valorar el error experimental; y las réplicas u observaciones al interior de las unidades experimentales sólo permiten valorar el error de muestreo.

³ Podríamos decir que lo más adecuado es que este profesional, además de tener un conocimiento estadístico, tenga la formación del área científica a la cual pertenece el experimento a desarrollar; esto es, el biometrista debe entender

siempre será fundamental para la efectividad de detección que tenga el ensayo; de allí la importancia de lo mencionado por Fisher, cuando dijo que “consultar a un estadístico después de que un experimento ha sido terminado, en general, puede no ser más que pedirle un examen *post mortem*; pues él puede quizás decir de qué murió el experimento”.

Distribuciones experimentales

Para el diseño de experimentos se tienen algunas alternativas de distribución básica que raudamente se mencionaron en uno de los párrafos del tema anterior y que deben ser seleccionadas, en el caso de la agronomía, de acuerdo al ambiente en donde se desarrollará la experiencia; de tal forma que se busque minimizar el error experimental. Estas distribuciones, debido a su capacidad de bloqueo de fuentes de variación secundaria⁴, pueden clasificarse de manera general en la forma indicada en la tabla 1.1.

Cabe aquí mencionar la importancia del principio de la aleatorización cuando no se detecta, en el ambiente experimental, alguna fuente de variación secundaria que produzca un efecto perturbador, pues sólo este principio minimiza dicho efecto. Así también es necesario indicar que, si en el contexto del experimento a desarrollar, el investigador identifica potenciales fuentes de variación secundaria, éstas podrían ser controladas con distribuciones que utilizan estructuras de bloques; de allí lo fundamental en el discernimiento indubitable sobre las circunstancias en que deberá utilizarse una distribución experimental dada. Existen también casos de fuentes de variación secundaria que no se pueden controlar con la distribución experimental, no obstante, pueden tratarse como covariables y aplicar un análisis de covarianza para nulificar el efecto perturbador de éstas. De esto último se realiza una explicación en el capítulo IV.

Una de las fuentes de variación secundarias principales (la más importante) que puede afectar los resultados de un experimento es el suelo, debido a su inexorable heterogeneidad. Por ello es que una distribución completamente al azar (DCA) se recomienda cuando se realizan experimentos en macetas, para reproducción de plantas o análisis de nutrientes, que generalmente se llevan a cabo en viveros o incluso en invernaderos; ensayos en los cuales no se evidencian fuentes con necesidad de bloqueo, ya que, por ejemplo, el suelo que se coloca en dichas macetas es homogenizado previamente. Una DCA es una distribución que también se utiliza a nivel de laboratorio, como en la reproducción de plantas *in vitro*, en donde también se tiene un ambiente prácticamente controlado. Su ventaja principal es la generación del mayor número de grados de libertad del error experimental, lo que comparativamente lo define como la distribución de mayor precisión estadística. Si bien su uso no se recomienda para experimentos que se realizan en campo, por su característica de no controlar ninguna fuente de variación secundaria, podrían utilizarse en este ambiente siempre que el suelo en donde se haya previsto el desarrollo del ensayo presente una relativa homogeneidad en cuanto a fertilidad, pendiente o cualquier otra característica física o edáfica del mismo.

el fenómeno para que su aporte en los resultados del ensayo tenga la eficacia necesaria. En la experimentación agrícola, no sólo basta con conocer sobre estadística, hay que tener los fundamentos para hacer las recomendaciones pertinentes en cuanto al diseño, desarrollo, medición de variables y análisis de los datos del experimento.

⁴ Las fuentes de variación secundarias son los factores que no son el objeto de estudio en los experimentos. Son efectos confirmados o probables debido a causas como, por ejemplo, la pendiente de un terreno, que el investigador identifica y bloquea con alguna distribución experimental como la de bloques completos o incompletos al azar o la de cuadro latino. Estas fuentes también se conocen como *variables de “ruido”*, debido a que pueden incluir en la variable respuesta una variación que no es atribuible a los tratamientos; por ello, al aplicar alternativas que bloqueen o nulifiquen su efecto, se denominan como *variables de ruido controladas*. En algunos textos, a estas variables secundarias, plenamente identificadas en un experimento, se las conocen también como variables “*NUISANCE*” (tontería), dado que su importancia, al ser variables de interés secundario dentro de un experimento, radica en que ellas sean controladas y no influyan en la respuesta del efecto de interés.

Tabla 1.1. Distribuciones experimentales básicas.

Ambiente controlado		Ambiente no controlado*	
Distribución completamente al azar (DCA)	Sin bloqueo de fuentes secundarias	Distribución de bloques completos al azar (DBCA)	Un solo bloqueo
		Distribución en cuadro latino (DCL)	Doble bloqueo
		Distribución en cuadro grecolatino (DCGL)	Triple opción de bloqueo

*Como distribución básica, en estos ambientes también se incluyen los de bloques incompletos.

La distribución de bloques completos al azar (DBCA) es la que más se utiliza en los experimentos de campo, en algunas veces no tanto por la presencia de fuentes de variación secundaria que sean evidentes en el suelo del área de ensayo, sino más bien como prevención de que existan los efectos de estas fuentes y puedan incidir en las respuestas de los tratamientos. Está indicada para contrarrestar sólo una fuente secundaria en experimentos con cultivos anuales de distancias de siembra reducidas, pues permite un solo bloqueo para afectaciones que pueden ser por pendiente del terreno, heterogeneidad en la fertilidad del suelo, cercanías a esteros o ríos o cualquier otra característica que se vea amenazante a la precisión en la determinación de los efectos de los tratamientos. Su mayor ventaja reside en el control que puede realizar de una fuente secundaria identificada en el área experimental, que es minimizada al formar bloques (que a su vez son las repeticiones) compactos, a manera de miniexperimentos, lo que consecuentemente evita una mayor afectación por la heterogeneidad del suelo y permite una comparación sin interferencia de efectos extraños entre los tratamientos; pues es de esta manera que hace posible reducir el error experimental. Su amplia utilización también se debe a cuestiones económicas, ya que el uso preventivo de bloques evita realizar experimentos en blanco o de uniformidad que duplican el tiempo y el costo de los ensayos.

A pesar de que el uso de una distribución experimental con bloqueo puede ser innecesario en algunos casos (si la fuente de bloqueo no tiene efectos significativos), ya que estadísticamente restan claridad a un ensayo al disminuir sus grados de libertad del error experimental⁵, su utilización permitirá definir en posteriores experimentos de campo la posibilidad de optar por la DCA; siendo esta última la distribución que comparativamente tiene una mayor precisión estadística. Asimismo, el uso de distribuciones con bloques se presenta como una gran alternativa cuando se tienen experimentos relativamente grandes, en donde algunas labores experimentales (riego, control de malezas, podas, medición de variables) no pueden realizarse en un solo día y, por lo tanto, una buena opción es llevarlas a cabo parcialmente, pero con bloques terminados; completando todo el experimento en el menor tiempo posible. El bloqueo también permite controlar fuentes de variación que pueden ser introducidas al ensayo cuando se utiliza distinta mano de obra para su manejo; en esta condición, la asignación de bloques completos a una misma persona permitirá controlar esta posible fuente de error.

La distribución en cuadro latino (DCL) se recomienda cuando el experimento involucra a cultivos de distancias de siembra amplias como mangos, aguacates, limones, cacao, palma africana, entre otros. Su mayor ventaja es la de poder controlar dos fuentes de variación secundarias (doble bloqueo), que pueden estar presentes en el suelo cuando el área de ensayo se extiende ampliamente y consecuentemente se incrementa la heterogeneidad en el suelo. Por cuestiones de precisión al generar escasos grados de libertad en el error experimental cuando

⁵ Los grados de libertad del error experimental se definen por diferencia de entre las distintas fuentes de variación que se hayan considerado en el modelo del análisis de varianza. Así, por ejemplo, en el caso de una distribución de bloques completos al azar, se deducen por: G. libertad del error experimental (GL_E) = G. libertad del total (GL_T) – G. libertad de tratamientos (GL_t) – G. libertad de repeticiones (GL_R).

se tienen 4 tratamientos o menos, es una distribución no recomendable; sin embargo, esta precisión puede mejorarse cuando esta distribución se repite dos o tres veces en el mismo lote (entiéndase que se repite dentro de un mismo espacio, pero ocupando distintas parcelas), modificando su nombre a cuadro latino replicado. Asimismo, cuando se tienen más de 7 tratamientos, la limitación más importante que se presenta en los cuadros latinos está referida a los costos de realización del ensayo, debido al incremento exponencial de unidades experimentales cada vez que aumenta un tratamiento. En el caso de la distribución en cuadro grecolatino (DCGL), que es una versión aumentada del DCL, no es de uso común en los experimentos agrícolas; no obstante, su ventaja principal es la de poder controlar tres fuentes de variación secundaria.

Las distribuciones experimentales indicadas son fundamentales en los ensayos que se realizan mayoritariamente para producción de conocimiento. En estos experimentos, su característica es la estrictez del método científico; a diferencia de otros ensayos que persiguen propósitos fundamentalmente de comercio, de allí que los ensayos experimentales con cultivos puedan distinguirse según los fines de la información que generan. Bajo esta distinción, los experimentos se pueden clasificar en dos grupos; los que se realizan con fines demostrativos y los que se realizan con fines netamente investigativos, a estos últimos, Steel & Torrie (1960), definen como experimentos críticos.

Un experimento que se realiza con fines demostrativos, en su contexto no es para producir conocimiento, más bien se realizan para exhibir resultados ya conocidos (De La Loma, 1966); en donde se comparan generalmente uno, dos o hasta tres tratamientos como máximo, junto a un tratamiento testigo. Los ensayos demostrativos comúnmente son realizados por productores de semillas y expendedores de agroquímicos, con el fin de convencer a los agricultores de la zona, en donde se ha llevado a cabo el ensayo, de que la nueva variedad o híbrido, o el nuevo producto, son mejores del que comúnmente utilizan y que pertenecen a la competencia.

En los ensayos demostrativos, en concordancia al propósito que persiguen, es necesario que el tratamiento que se quiere demostrar (vender) sea evidentemente mucho mejor que el tratamiento testigo; de tal forma que, con sólo ver la producción en el campo, sea suficiente para que el agricultor del sector decida optar por el cambio de la nueva propuesta. Son experimentos que, al igual que los de investigación, necesitan de un manejo riguroso en cuanto a las demás labores culturales, si lo que se pretende es que los resultados sean exclusivos efectos de los tratamientos. Generalmente se desarrollan con pocos tratamientos (máximo 4) y cada uno de ellos se aplica en parcelas únicas, relativamente grandes y homogéneas. Dada la gran diferencia que deberán tener los tratamientos que se evalúan en los ensayos demostrativos, la varianza interna de los mismos no se valora rigurosamente.

En los experimentos que se realizan con fines de investigación, comúnmente no se conoce lo que podría resultar al final de la experiencia; se realizan bajo un rigor científico, tal vez mucho más exigente que el rigor de los ensayos demostrativos. Pueden tener pocos o muchos tratamientos y son desarrollados generalmente por instituciones de investigación o por las universidades. Es fundamental el cumplimiento de los principios de aleatorización⁶, repetición y bloqueo; los cuales ineludiblemente deberán considerarse para poder verificar los efectos de los tratamientos y separarlos de los efectos de fuentes secundarias que pueden incrementar el error experimental. En estos tipos de experimentos es fundamental el cuidado de la varianza interna.

⁶ Este principio, junto con la selección intencional de los factores y/o niveles de estudio, definen básicamente lo que se conoce como *experimento verdadero* o *formal*; aparte de cumplir con los principios de bloqueo (cuando es necesario) y repetición, además del rigor en el cuidado de la uniformidad de las diferentes labores que se realizan en los experimentos agrícolas.

La aleatorización es el principio que protege a cualquier ensayo contra el cometimiento del error tipo I⁷. El sorteo en la asignación de los tratamientos a las unidades experimentales, o viceversa, disminuirá la probabilidad de rechazar una hipótesis nula que sea verdadera; esto evitará la facilidad de aceptar alternativas (variedades, dosis de fertilizantes o de otros agroquímicos, distancias de siembra, etc.) que, en la realidad, no mejoran la producción de algún cultivo. La repetición es el segundo principio que tiene la particularidad de darle precisión a un experimento, ya que mientras más repeticiones tiene un tratamiento, más se reduce el error experimental. Las suficientes repeticiones protegen a cualquier experimento de caer en el error tipo II, esto es, la aceptación de una hipótesis nula que es falsa; no obstante, si bien en la agronomía es menos importante que el error I, su cometimiento no permitirá detectar alguna alternativa (tratamiento) que en la realidad podría mejorar el rendimiento de un cultivo. Lo que no debe olvidar el investigador es que sólo la repetición permite estimar el error experimental aleatorio, fuente de variabilidad que siempre existirá en un ensayo. Las repeticiones son las que establecen si las diferencias encontradas entre tratamientos se deben al efecto real de éstos, o son diferencias casuales inherentes a las unidades experimentales.

Por último, el bloqueo es el principio aplicable a los experimentos de campo o que se realizan en condiciones no controladas (este principio no aplica en experimentos bajo ambientes controlados y/o de material experimental homogéneo), con el propósito de soslayar aparentes o comprobadas fuentes de variación secundarias que el investigador observa en el área de prueba y que podrían incrementar el error experimental. El bloqueo es una de las características de las distribuciones como la DBCA, la DCL y la DCGL, que controlan una, dos y hasta tres fuentes de variación secundarias, respectivamente. Asimismo, este último principio también permite disminuir la probabilidad de caer en el error II, ya que con él se controla de mejor manera el error experimental al dividir un experimento en bloques compactos, lo que hace que se limite más efectivamente la heterogeneidad del suelo y el efecto de ésta se evidencie mayoritariamente entre los bloques antes que tenga incidencia en la precisión del ensayo.

Si bien un experimento distribuido en bloques es eficaz en mejorar la precisión de un ensayo cuando existen fuentes de variación secundaria reales en un suelo con una variabilidad (heterogeneidad) en una dirección, su función puede quedar sin efecto cuando se tiene un alto número de tratamientos. Esta condición hace que los bloques se extiendan una longitud considerable en el terreno, permitiendo que la heterogeneidad del mismo actúe con mayor persistencia e incremente sustancialmente el error experimental. En esta situación, indicar un número máximo de tratamientos no tiene cabida, debido a que existen varias condicionantes para este propósito; este número máximo dependerá del factor (es) de estudio, del tipo de cultivo, del espacio de terreno experimental, de las variables a medirse y del tamaño de la unidad experimental.

Una alternativa para conservar o mejorar la precisión en un ensayo con bloqueo cuando éste tiene un importante número de tratamientos es el uso de distribuciones con bloques incompletos, siendo los más conocidos las distribuciones en latices. Este es un tipo de distribución experimental que generalmente se la utiliza con exclusividad en el área de la agronomía para investigar un número elevado de tratamientos, porque permite controlar, en el suelo o en el espacio experimental, la heterogeneidad que se acrecienta a medida que un experimento se extiende en el campo; situaciones en donde distribuciones como la DBCA pierden eficacia al tener bloques con mayor variación.

⁷ En los resultados de un experimento se tienen dos posibles sucesos de error. Uno de ellos se conoce como Error I y se comete cuando se rechaza una hipótesis nula verdadera; el otro se conoce como Error II y sucede cuando se acepta una hipótesis nula que es falsa. En la agronomía, como en la mayoría de las ciencias, se diseñan los ensayos con mayor cuidado hacia el error I, relegando al error II, debido a que, en términos de costos económicos, caer en el primero puede hacer que la rentabilidad de la actividad agrícola en vez de aumentar se reduzca.

En la distribución en latice, conocida también como distribución de bloques incompletos y en la cual los tratamientos también son asignados aleatoriamente, cada uno de los bloques contienen una parte del total de tratamientos que se investigan, lo cual constituye la característica principal de este arreglo experimental. Es una distribución que está indicada para estudios sobre mejoramiento de plantas, en donde se tiene una gran cantidad de variedades o líneas que son producto de cruces y de procesos de hibridación que requieren ser evaluados simultáneamente. Bajo la circunstancia mencionada, la distribución en latice es más eficaz que la DBCA por cuanto los bloques no contienen a todos los tratamientos, permitiendo consecuentemente tener grupos de parcelas (unidades experimentales) más homogéneos, debido a que la extensión de las repeticiones o bloques del experimento se reduce; además de ser una alternativa experimental de menor costo por cuanto generan un menor número de unidades experimentales, contrario a distribuciones como, por ejemplo, la DCL.

Factores de estudio en la agronomía

Llevar a cabo el diseño de un experimento implica el análisis de todas las circunstancias o la situación coyuntural que rodeará al mismo. Bajo esta óptica, los experimentos que se realizan dentro de la agronomía pueden definirse relativamente como los más complicados debido al ambiente muy variable que los rodean. Un experimento que se realiza a campo abierto tiene múltiples factores secundarios, menos controlables que incontrolables, que deben tenerse en cuenta si se quiere llegar a resultados contundentes y con el mínimo error. Esta situación es diferente para los experimentos industriales, en donde las condiciones de la experiencia prácticamente son de laboratorio; o los relacionados a la salud, si bien tienen también sus fuentes de variación secundaria las cuales son controladas por la propia unidad experimental (una persona), estas son menores con relación a los experimentos agronómicos.

La experimentación agrícola básicamente se fundamenta en el rendimiento de los cultivos, especialmente cuando se persiguen fines demostrativos. Estos rendimientos, a su vez, dependen de varios factores que los dominan. En forma general, considerando los aspectos que el investigador puede controlar, los experimentos agrícolas pueden clasificarse, según los factores de estudio, en la siguiente forma:

1. Experimentos de adaptabilidad de cultivos
2. Experimentos sobre nutrición vegetal
3. Experimentos relacionados con plaguicidas
4. Experimentos relativos a ciertas prácticas culturales

Los ensayos que se llevan a cabo para evaluar los factores antes mencionados están fundamentados, relativamente, en hipótesis de investigación de poca duración en el tiempo. Así, por ejemplo, un experimento sobre fertilizantes sintéticos permite tener resultados que pueden cambiar, incluso de una cosecha a otra, debido a que el nivel nutricional de un suelo varía en el tiempo; luego una dosis de fertilizantes que hoy es la adecuada, mañana no podría ser así. Además, los resultados de un experimento específico bajo condiciones edafoclimáticas determinadas, tienen aplicación exclusiva sólo para dichas condiciones; de allí que, si éstas cambian al replicar una experiencia, muy difícilmente se obtendrán los mismos resultados en experimentos similares. De allí que se diga que la investigación agrícola, por naturaleza, es localizada.

Experimentos de adaptabilidad de cultivos

Las condiciones ambientales son variables en el espacio, más que en el tiempo; de allí que las especies vegetales también tengan sus preferencias del medio que las rodea en donde su

expresión es la máxima. En este sentido, las diferentes variedades o híbridos de plantas de cultivo pueden diferir en cuanto a sus rendimientos de acuerdo a la influencia climática.

Los experimentos de adaptabilidad de cultivos, específicamente de variedades y de híbridos, no es más que seleccionar un determinado grupo de plantas de un mismo género para sembrarlas en una zona de condiciones agroclimáticas establecidas, de tal forma que se pueda verificar el material vegetativo que presente la mayor respuesta productiva. En estos experimentos, el factor de estudio lo constituyen las variedades sometidas a comparación; esto es, los tratamientos corresponden a cada una de estas variedades, en las cuales se evalúa su expresión productiva y fenotípica en un ambiente específico. Es un requisito muy importante estar seguro de que el material genético considerado para el estudio corresponde a una semilla certificada⁸; si este requerimiento no se cumple, los resultados que se obtengan no serán los correspondientes a un tipo de germoplasma determinado, cuyo error podrá hacer que se seleccione materiales no adecuados para una región específica.

Otro aspecto importante a tener en cuenta en esta clase de experimentos es lo relacionado al *efecto de borde* de las parcelas. Las hileras contiguas de distintas variedades pueden facilitar la polinización cruzada (alogamia); es decir, puede ocurrir que el polen de una variedad se deposite en el estigma de la flor de otra variedad, originando un material hibridado cuyo rendimiento no representa a las variedades evaluadas, constituyendo esto en una fuente error (De la Loma, 1966). Para soslayar este inconveniente, se hace necesario tener dentro de la parcela las suficientes hileras de plantas que sirvan como efecto de borde, de tal forma que las hileras que se encuentran dentro de la mal llamada área útil de parcela⁹ no tengan este efecto de alogamia.

El efecto de borde también puede servir para proteger sobre la influencia de la cobertura o competencia que unas variedades pueden hacer sobre otras. En estos experimentos se podrían tener especies de mayor altura y, por lo tanto, mayor cobertura foliar, que harán que especies de menor tamaño no puedan desarrollar su total expresión agronómica. Esta situación también obliga a tener hileras de plantas que reciban este efecto, dejando hileras libres en el área útil para realizar las mediciones respectivas.

El uso de materiales vegetativos empleados como testigos, constituye un elemento importante en estas experiencias, especialmente cuando las semillas y el mantenimiento del cultivo necesitan de una mayor inversión. Generalmente, tener un material de mayor rendimiento para una zona agrícola demanda también gastos extras, el cual, si a la postre no resulta rentable, no será una alternativa para el cambio de semilla.

En estos experimentos, dado que el objetivo es, además de evaluar el rendimiento, hacer una caracterización del material vegetal, se deben medir algunas variables complementarias que definen la expresión fenotípica de la especie en cuanto a la idoneidad para una zona determinada. En este sentido, también se evalúan características como resistencia a ciertas enfermedades imperantes por las condiciones ambientales propicias de la zona; precocidad para definir el periodo de siembra; y algunas otras características como altura de planta,

⁸ Cuando se trata de ensayos con rigor científico, este requisito es ineludible. El uso de una semilla no certificada incrementa la probabilidad de cometer el error tipo II de la estadística. En otras palabras, existe una mayor probabilidad de que el gasto en la investigación sea en vano, a pesar de que los tratamientos que se evalúen tengan efectos reales.

⁹ La costumbre de los biometristas agrícolas es la de llamar *parcela útil* al espacio dentro de cada unidad experimental de un ensayo en donde se evalúan las variables de respuesta; sin embargo, en un experimento de campo, aun los bordes que rodean a estas "parcelas útiles" también son útiles (en la realidad toda el área experimental es útil). En términos más acordes, es mejor definirlos como UNIDAD DE MUESTREO; de hecho, es así como en este texto se define a esta área en algunos párrafos.

rendimiento por planta, número de frutos por planta, características morfológicas de los frutos, etc.

En relación con el tipo de distribución requerido para llevar a cabo estos ensayos, generalmente se utiliza la distribución de bloques completos al azar; no obstante, cuando los materiales vegetativos a estudiar son numerosos, una alternativa que permite un mejor control sobre la variabilidad que introduce el suelo, es el uso de una distribución en latice simple o diseño de bloques incompletos que fueron creados con este propósito.

Los resultados de este tipo de experimentos permiten hacer las recomendaciones sobre las distintas variedades o híbridos que deberán cultivarse en zonas determinadas, cuyas condiciones hacen posible obtener la mayor producción; sin embargo, si no existe el debido acompañamiento en cuanto a un acertado manejo agronómico (con aplicación tecnológica comprobada), de nada servirá tener el mejor material genético para una zona dada. En consecuencia, habrá que realizar experimentos complementarios sobre el manejo cultural (como los que se detallan a continuación), con el fin de que el mejor material genético sea explotado a su máxima capacidad.

Experimentos sobre nutrición vegetal

Los experimentos agronómicos cuyos resultados de aplicación son de poca duración son los que se llevan a cabo con los fertilizantes, por ello se dicen que son de hipótesis rápidamente cambiables. Los resultados de estos experimentos tienen una fuerte influencia del estado edafológico del suelo, por consiguiente, tomar sus resultados de referencia después de unas dos a tres temporadas puede resultar inadecuado. Las condiciones nutricionales y de humedad de un suelo son tan móviles que inclusive podrían hacer que la fórmula fertilizante que resulte experimentalmente efectiva, en la subsecuente aplicación práctica a nivel de campo, presente efectos obsoletos.

Los ensayos sobre nutrición son experimentos con la mayor afectación de factores secundarios que aquellos en donde se realizan estudios de adaptabilidad. A la gran influencia que tiene el suelo sobre los resultados se une la potencial fuente de variación concerniente a la misma expresión genética de la semilla utilizada; por lo tanto, la recomendación respecto de esto es el uso de semilla certificada, de tal forma que no se confundan los efectos de mayor significancia con el mismo comportamiento agronómico intrínseco del cultivo, sino que sea producto de la aplicación de una determinada dosis de fertilizantes que se esté probando. A estas potenciales fuentes de variación también pueden unirse, y de hecho tienen mucha importancia, la frecuencia (épocas) y forma de aplicación de los fertilizantes; situación que está relacionada específicamente con la eficiencia de los diferentes tipos de nutrientes.

Cuando se evalúa más de un nutriente, la alternativa más adecuada para este tipo de experimento es el uso de un planteamiento factorial, en donde simultáneamente se valore el efecto de varios de estos nutrimentos. En este caso, si se realizan estudios de dosis equidistantemente espaciadas, son factibles los análisis regresionales, desde los cuales pueden establecerse las dosis de máximo rendimiento; o, en el mejor de los casos, los puntos de máximo rendimiento desde una superficie de respuesta generada a partir de una combinación de fertilizantes (experimento factorial). Además, utilizar un arreglo factorial en experimentos que involucran plantas cuando hay más de un factor actuando simultáneamente, permitirá valorar el efecto de la interacción de los factores que es muy importante, dado que está inmerso un organismo biológico que responde mejor a las interacciones antes que a causas individuales.

En relación al párrafo que antecede, es útil la indicación del procedimiento de análisis estadístico a seguir que se realiza cuando se evalúan dosis de nutrientes. Existe el criterio común de aplicar una valoración cualitativa de los efectos, utilizando en primera instancia el

análisis de varianza (ANDEVA) y luego las pruebas de comparación múltiple (pruebas *post hoc*) con las que definen al tratamiento (dosis) “ganador”. Sin embargo, en estos tipos de ensayos cabe la identificación de cuándo se tratan de factores cualitativos y cuándo representan a factores cuantitativos. Factores cualitativos en experimentos con fertilizantes son los casos en donde se comparan diferentes tipos (fuentes) de nutrientes, épocas o frecuencias de aplicación o incluso formas de aplicación; por lo tanto, el procedimiento más adecuado para establecer el tratamiento ganador es el indicado anteriormente (ANDEVA - prueba *post hoc*).

Cuando se trata de factores cuantitativos, esto es, cuando se evalúan dosis de fertilizantes y la variable de respuesta también es de tipo cuantitativa (por ejemplo, el rendimiento), lo más acorde a esta situación es la búsqueda de funciones continuas de respuesta, ante lo cual los análisis de regresión múltiple para definir funciones de producción se presentan como la alternativa estadística con mucha mayor explicación que un simple ANDEVA; cuyo panorama que le ofrece al investigador esta última herramienta estadística, es sólo de valoración cualitativa. Un modelo de regresión múltiple aplicado en estos casos, además de brindar información de los efectos de los factores de forma independiente y de interacción (que es lo único que hará un ANDEVA simple para un experimento factorial), también permitirá conocer cómo es la respuesta del cultivo de acuerdo a los incrementos subsecuentes de las dosis de nutrientes, en qué niveles el rendimiento es máximo o dónde éste se define como un óptimo económico. Adicionalmente, un modelo de regresión múltiple también hace posible incluir otros factores de producción, como ciertos fenómenos climáticos o edafológicos que tienen incidencia significativa y que pueden ser valorados en un modelo específico que, al final, definirá una dosis de nutrientes con mayor validez externa.

A sabiendas de que la respuesta fisiológica de las plantas a los nutrientes no es directamente proporcional, el planteamiento de evaluación estadística cuantitativa en los ensayos sobre dosis de fertilizantes demanda un conocimiento amplio de los requerimientos nutricionales de los cultivos; para lo cual el investigador deberá fundamentarse en una vasta revisión bibliográfica y empírica, de tal forma que le permita deducir los niveles de prueba que incluyan al rendimiento máximo y que sea posible definir el tipo de función de respuesta, a partir de la cual se establezcan dichos límites máximos. En este contexto, para poder verificar los efectos de la interacción entre los factores de interés siempre será beneficioso que se planifique bajo un arreglo factorial.

Por otro lado, si bien existe la suficiente ventaja de los arreglos factoriales sobre aquellos ensayos en donde se evalúa un solo factor a la vez, como así se lo demuestra en el capítulo 6, también debe tenerse presente que, tratar de desarrollar experimentos con múltiples factores y niveles en la agricultura, tiene sus límites en cuanto a la precisión y al componente económico que pueden hacerlo inviable. Ante esta situación, el investigador podría recurrir al uso de diseños con efectos confundidos que hacen posible disminuir sustancialmente el número de unidades experimentales; considerando para este propósito que las interacciones de tres o más factores en la práctica no contribuyen a un entendimiento claro de estas fuentes de variación en la respuesta de un cultivo y, por lo tanto, pueden mezclarse (confundirse) con el efecto de los bloques cuando se utiliza una DBCA. La segunda opción por la que puede optar el investigador es la selección arbitraria de algunos tratamientos (combinaciones factoriales) de interés para éste; sin embargo, esta opción podría causar que no se evalúen efectos importantes en el ensayo y se pierda información valiosa. Una tercera alternativa para contrarrestar las limitaciones del desarrollo de experimentos factoriales con todos sus niveles e interacciones es la utilización de ciertos arreglos factoriales incompletos, establecidos específicamente para procesos estadísticos de optimización; los cuales inicialmente se crearon para experimentos industriales, pero que, con cierta pericia y salvedad, podrían aplicarse en ensayos agronómicos. Algunos de estos arreglos factoriales incompletos para optimización se indican en el capítulo 7.

La necesidad de evaluar la rentabilidad inmediata de los tratamientos en esta clase de experimentos es fundamental, ya que sólo la verificación de que alguna dosis de nutrientes ofrece una mayor ganancia que otra se constituye en la razón para cambiar la dosificación que comúnmente se utiliza en la fertilización. Para este propósito, una técnica muy útil, difundida en su momento por el Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT), es el análisis de la *tasa de retorno marginal*, una alternativa viable que permite establecer las dosis de nutrientes que producirán los cambios más rentables. Para complementar esta técnica y tener una visión integral de la rentabilidad de los tratamientos, se puede optar por obtener también la relación beneficio-costo de cada uno de ellos cuando el experimento involucra a cultivos anuales¹⁰, con el propósito de definir el tratamiento de mayor utilidad económica. Adicionalmente, una alternativa de análisis, cuando se evalúan cuantitativamente fertilizantes, es la elaboración de funciones de producción; a partir de las cuales se establecen las dosis de máximo rendimiento y las dosis óptimas económicas. Lo ideal en la fase de aplicación práctica de los resultados de un experimento agrícola es que un tratamiento estadísticamente ganador deberá ser, también, rentablemente ganador.

Al igual que en el caso de los experimentos de adaptación de variedades o híbridos, las parcelas en las cuales se llevarán a cabo estos experimentos son relativamente las más pequeñas; por lo que, en cultivos como el maíz, cada unidad experimental podría contener un mínimo de tres hileras de plantas, en donde la hilera del medio servirá para la evaluación del efecto de los tratamientos. En el caso del arroz, bajo un sistema de cultivo tradicional en donde las plantas se mantienen en inundación durante todo su ciclo, hay que realizar muros de protección en cada unidad experimental para evitar el traslape de las diferentes dosis que se aplican en las parcelas adyacentes; además de tener un suficiente borde periférico que aisle el cruce de nutrientes a través del mismo suelo de forma interna.

En cuanto a la necesidad de utilizar tratamientos testigos, salvo que el suelo en donde se prevea desarrollar el ensayo esté recién incorporado a la actividad agrícola (llamado suelo virgen) o a su vez el cultivo no demande excesivas cantidades de nutrientes, no será necesario utilizar testigos absolutos; pues es esperable que en tierras mono-cultivadas por años, constituidos en lotes de tierra con grandes pérdidas de su fertilidad, el rendimiento de cultivos comerciales sea drásticamente menor. Por otro lado, en la mayoría de los casos, recurrir a los testigos convencionales (la dosis que comúnmente utiliza el agricultor), permitirá verificar la necesidad del cambio hacia una nueva fórmula de fertilizantes.

Tanto para pruebas de un solo tipo de nutriente a distintas dosis (experimento unifactorial) como para experimentos en donde se evalúen combinaciones de nutrientes (experimento factorial), la distribución experimental básica que comúnmente se utiliza es la de bloques completos (o también los incompletos) al azar. Para estos tipos de experimentos, una alternativa interesante de aplicar para evaluar al menos los macronutrientes N, P y K, pero no con el arreglo de un factorial clásico (totalmente cruzado), sino como uno incompleto, es el llamado arreglo *San Cristóbal* (Rojas, 1961); el mismo que fue ideado para experimentos con fertilizantes en caña de azúcar. Un factorial incompleto que también puede ser interesante utilizar, con muchas más ventajas que el San Cristóbal, es el *diseño central compuesto* que se utiliza en la metodología de superficie de respuesta; esta última, ampliamente aplicada en experimentos industriales.

¹⁰ En cultivos anuales, la definición de la utilidad económica de los mismos puede realizarse incluyendo todos los valores que forman el costo de producción de una especie determinada. Bajo este enfoque, será mejor identificar, en un experimento dado, cuánto es la ganancia por cada unidad monetaria invertida en cada uno de los tratamientos evaluados, lo que puede establecerse con una visión total al determinar la relación beneficio-costos de estos últimos. Esta visión no puede lograrse con metodologías que sólo valoran el costo del tratamiento, el cual es separado del ingreso económico total, dejando de lado el resto de costos de las actividades culturales (las que no son motivo de estudio) como preparación de suelos, riego, control de malezas, fertilización, etc.

En el caso de estudios de fertilizantes en cultivos definidos como perennes y con distanciamientos relativamente amplios (mango, palma africana, guayaba, limones, etc.), la recomendación es utilizar las distribuciones en cuadro latino; dado que el área que se determina para el experimento puede extenderse de manera importante y, por lo tanto, esto hará que también se incremente la variabilidad del suelo en sus dos direcciones. En estos ensayos, la unidad experimental generalmente corresponde a una planta, sobre la cual se evalúan las variables de respuesta. La propiedad de restringir dos fuentes de ruido (en doble dirección) de la distribución en cuadro latino la hacen muy indicada para estos tipos de experimentos (incluidos los de ciclo corto), pues controlan de mejor forma la heterogeneidad del suelo y, por ende, el error experimental que ésta causa.

Lo que hasta aquí se ha explicado hace referencia a fertilizantes de tipo edáficos o básicos; sin embargo, en los actuales momentos existen un sinnúmero de productos de aplicación foliar que contienen micronutrientes y/o estimulantes que parcialmente favorecen un mejor proceso metabólico y, por lo tanto, pueden favorecer incrementos en el rendimiento de los cultivos. A este respecto, existe un error muy común cuando se realizan experimentos utilizando bioestimulantes y/o fertilizantes de aplicación foliar bajo el criterio de complementariedad a la fertilización edáfica; no obstante, si verdaderamente existe un efecto de estos productos, la idea en la definición de los tratamientos deberá considerarse como un suplemento. De esto último se deduce que las dosis de fertilizantes que se definan como tratamientos, deberán establecerse en forma decreciente, si lo que se quiere es mejorar la rentabilidad.

Tanto para experimentos en donde se utilicen fertilizantes foliares (o algún estimulante) como en los casos en donde los tratamientos representan algún producto supuestamente nutricional de origen orgánico, siempre será necesario utilizar los testigos absolutos; ya que existen en el mercado una serie de estos productos, cuyos efectos apenas mejoran el rendimiento de los cultivos en algunos casos, o en otros solamente son líquidos inertes que no sirven para nada. En estos experimentos, sólo el testigo absoluto permitirá vislumbrar si es mejor no aplicar ninguno de estos productos.

Es necesario indicar que el requerimiento de borde en las parcelas, cuando el experimento consiste en aplicaciones de nutrientes (u otros agroquímicos) en forma foliar, deberá ser un poco más amplio que los bordes considerados en experimentos con aplicación edáfica. En esta circunstancia, por motivos de condiciones climáticas como el viento, o incluso una mala aplicación foliar, pueden hacer que los tratamientos tengan un mayor traslape entre parcelas contiguas, lo que podría confundir el efecto real de cada uno de ellos.

Experimentos relacionados con plaguicidas

Los experimentos que se realizan con plaguicidas son los que ofrecen resultados, comúnmente, con la mayor variabilidad. Tanto las malezas como los insectos, o las enfermedades, generalmente inciden negativamente de forma localizada dentro de un cultivo (heterogénea); de allí que en la recomendación sobre el tamaño de la parcela deberá considerarse esta característica, dándole la extensión suficiente para poder observar el efecto de algún plaguicida que se esté probando. No obstante, para la evaluación de plaguicidas siempre será necesario tener una infestación significativa en los cultivos que incida de manera importante en la producción de los mismos para poder observar el efecto de los agroquímicos. Si esta infestación es leve o nula y no se utilizan testigos absolutos, se podría llegar a la conclusión errónea de creer que dichos plaguicidas son eficaces cuando en la realidad no lo sean.

Para la evaluación de los herbicidas selectivos, en el caso de cultivos anuales y en condiciones de post-emergencia, siempre será necesario el uso de un tratamiento testigo en donde no se utilice el plaguicida y las malezas sean controladas de forma manual. Este tratamiento testigo

permitirá evaluar algún efecto de fitotoxicidad que manifieste el cultivo cuando reciba la aplicación del producto y que de alguna manera debe reflejarse en el rendimiento, y no únicamente a escasos días posteriores a la aplicación de los herbicidas, luego de lo cual se logra apreciar visualmente cierto efecto de resentimiento que muestra el cultivo y que luego desaparece.

Si bien las evaluaciones de control de los herbicidas se realizan de forma cualitativa y/o cuantitativa¹¹, las dos alternativas terminan reportando el resultado en forma porcentual, tanto de manera general como de forma específica por cada maleza presente. Una vez obtenidos los datos de cada unidad experimental, éstos generalmente se relacionan con una escala nominal estándar, establecida por la Asociación Latinoamericana de Malezas (ALAM), cuyos rangos se detallan en la tabla 1.2.

La utilización de un testigo absoluto en experimentos en donde se emplean herbicidas pre-emergentes (conocidos como mata semillas) es de suma importancia. Si no existe este testigo, jamás se podrá conocer qué clase de malezas pudo controlar el plaguicida. Para llevar a cabo estos experimentos de pre-emergencia, deben seleccionarse lotes en donde previamente se conozca que existe una proliferación masiva y pertinaz de malezas, de tal modo que sea factible la valoración del efecto herbicida con esta característica.

Tabla 1.2 Escala cualitativa para valoración del control de malezas

Nivel de control (%)	Valoración cualitativa
0 - 40	Ninguno o pobre
41 - 60	Regular
61 - 70	Suficiente
71 - 80	Bueno
81 - 90	Muy bueno
91 - 100	Excelente

Fuente: ALAM (1974).

En algunos casos, la evaluación del control de malezas realizado por los herbicidas en un estudio experimental se efectúa bajo un criterio netamente subjetivo, a partir de la observación que realiza el evaluador en cada unidad experimental; para lo cual debe tenerse la suficiente experticia y tratar la información con la mayor objetividad posible. Si esta forma de valoración se ejecuta cuando el área sobre la cual se aplican los productos de prueba tiene una cobertura de las malezas poco uniforme, el dato que se obtenga podría estar totalmente desubicado; por ello, para este tipo de experimentos deben utilizarse áreas en donde la presencia de estas arvenses tenga una cobertura total del suelo. Cuando existe esta cobertura total, se facilita la realización de evaluaciones más objetivas, utilizando para ello marcos cuadrados (100 cm x 100 cm) en cada unidad experimental y dentro los cuales solamente se hace una cuantificación proporcional de cada tipo de maleza presente en dicho marco.

Otra alternativa para cuantificar el efecto de herbicidas en forma general sobre las malezas, es determinando el peso de la masa foliar existente luego de aplicado el agroquímico, lo cual servirá para la comparación respectiva con el testigo absoluto.

¹¹ En realidad, las formas de evaluación de efectos de herbicidas en las malezas son de tipo semi-cuantitativa o cuantitativa. La primera es por estimación visual (cualitativa), cuya valoración al final se traduce en un valor porcentual (cuantitativa) mediante una escala. La segunda es mediante conteo de malezas dentro de marcos cuadrados o rectangulares al interior de las unidades experimentales, lo que al final se reporta también en porcentaje.

Las valoraciones económicas de los experimentos con herbicidas, generalmente se realizan basados únicamente en los costos de los tratamientos (costos variables), ya que éstos son experimentos relativamente rápidos y no se llega hasta la producción del cultivo. El procedimiento se fundamenta en los días a control de la maleza utilizando umbrales técnicos entre aplicaciones sucesivas del plaguicida; de allí que aquella dosis o producto que tenga la mayor cantidad de días a control, traducidos a costos, sea el tratamiento más económico. Cuando el experimento se realiza con cultivos perennes, o incluso con cultivos anuales en donde se evalúa el rendimiento, una metodología muy práctica para valorar económicamente los tratamientos es mediante la técnica de presupuesto parcial del CIMMYT.

Los datos obtenidos en experimentos con herbicidas, cuando se evalúa su efecto sobre las malezas, se expresan en forma porcentual, y por ende muchas veces no hay un ajuste normal de los errores; más bien éstos (en muchos casos) tienen distribución binomial. Dado que esta es una variable que puede tener una importante dispersión, a veces, para lograr homocedasticidad de los errores y verificar la eficacia del análisis de varianzas se recomienda el cambio de escala de los datos a una de tipo angular (ajuste arco-seno). Un error común de muchos biometristas relacionados con experimentos de malezas, es su uso sin restricción del ajuste angular sólo por el hecho de tener datos porcentuales; sin embargo, hay que tener presente que el análisis de varianzas tiene cierta robustez que hace que admita cierto grado de dispersión y, en la mayoría de situaciones, no es necesaria dicha transformación. Una transformación angular (o cualquier otra) realizada de manera indiscriminada a los datos, más bien puede favorecer la presencia del error tipo I de la estadística cuando el grado de significancia es estadístico y no práctico. Cabe mencionar que en las situaciones en donde se realizan estos u otros ajustes de varianzas, en las publicaciones debe hacerse constar las medias de los datos sin ajuste, realizando la explicación respectiva; ya que presentar un promedio ajustado puede confundir al lector.

La comprobación de normalidad y de igualdad de varianzas puede realizarse de forma gráfica planteando los residuos o errores versus la probabilidad (normalidad), o los residuos versus los valores predichos (homocedasticidad); o también mediante pruebas analíticas como la de Shapiro-Wilks o la de Levene, las cuales son pruebas estadísticas para verificar normalidad y homocedasticidad, respectivamente.

Cuando se realizan experimentos con insecticidas, los requerimientos de unidades experimentales o parcelas relativamente grandes son una necesidad, si es que realmente se quiere encontrar resultados. Una parcela grande tendrá, a su vez, una unidad de muestreo de mayor dimensión, en donde se podrán tener datos menos dispersos cuando se evalúe, por ejemplo, la presencia de insectos muertos después de la aplicación del plaguicida. Asimismo, a la cuantificación de insectos muertos, puede complementarse la evaluación de plantas afectadas dentro de cada unidad experimental, lo que en muchos casos es una variable más representativa cuando se requiere valorar la incidencia del insecto y éste es tan móvil que será complicado encontrarlo muerto en las plantas.

Cada parcela debe tener el suficiente borde perimetral que proteja al experimento de efectos de traslape de tratamientos, especialmente cuando se evalúa la presencia de insectos muertos. La toma de datos puede realizarse por plantas dentro de la unidad de muestreo, como por ejemplo la presencia de *Spodoptera frugiperda* (cogollero) en maíz; o a su vez, utilizando marcos cuadrados de 1 m de lado como en el caso de la evaluación de *Oebalus insularis* (chinche del arroz) en el cultivo de arroz sembrado al voleo. Incluso el efecto de algún insecticida puede evaluarse solamente en la parte de una planta (que se constituye a su vez en la unidad experimental) como, por ejemplo, la presencia del trips (*Trix* sp.) en banano y su daño conocido como la mancha roja; para lo cual las evaluaciones se realizan en los racimos.

Actualmente, dentro de los márgenes de una agricultura amigable con el ambiente, se están probando una serie de alternativas biológicas u orgánicas para prevenir y/o controlar el ataque de los insectos a los cultivos. Para poder determinar la eficacia de alguna alternativa ecológica en estos tipos de experimentos, dependiendo de la importancia del ataque de una plaga sobre el rendimiento de un cultivo, podría ser necesario la inclusión de hasta dos tratamientos testigo; uno de ellos correspondiente a la aplicación tradicional de un agroquímico, denominado como testigo convencional, muy utilizado para el control de una plaga específica; y otro que corresponda a las unidades experimentales en donde no se aplique ningún producto para el control de la plaga, al cual se lo definirá como testigo absoluto. El tratamiento convencional facilita comparar la eficacia de los tratamientos ecológicos, si es que estos últimos tienen algún efecto real; mientras que el testigo absoluto, permitirá establecer si la incidencia de la plaga sobre el cultivo es de importancia como para incurrir en los costos adicionales que demandarían su control.

Los datos provenientes de experimentos sobre insecticidas, en relación a las variables que miden la presencia del insecto en el cultivo (incidencia), presentan fuertes dispersiones que hacen imposible que sus residuos o errores presente normalidad u homocedasticidad. En este sentido, para poder realizar los análisis de varianza o para aplicar alguna otra prueba paramétrica, se necesitará la transformación del dato a una escala del tipo logarítmica, debido a la fuerte dispersión que generalmente presentan. Cuando el daño que causan los insectos se mide cualitativamente, estos pueden tomar una escala ordinal, para lo cual el uso de pruebas estadísticas no paramétricas como la de Kruskal -Wallis o la de Friedman se presentan como opciones análogas al análisis de varianza.

Uno de los casos que comúnmente requiere del uso de la transformación logarítmica es cuando se evalúan nematocidas. La población de nematodos asociados a cultivos se valora a través del número de estos fitoparásitos por cada 10 g de raíces o también en el suelo circundante al sistema radicular de las plantas por cada 100 cm³ de éste, variables que generalmente se presentan con una dispersión muy fuerte y que para ser analizadas con alguna prueba paramétrica requieren también un ajuste fuerte como el logarítmico.

La evaluación de productos para controlar enfermedades en los cultivos presenta variables cuya característica dispersa también es relativamente fuerte, de allí que las unidades experimentales para estudios sobre enfermedades requieran de parcelas grandes en donde se pueda valorar de manera más efectiva el efecto de algún fungicida. Los datos pueden expresarse en forma porcentual o utilizando escalas ordinales que, en la mayoría de los casos, deben transformarse para estabilizar las varianzas del error experimental de los grupos o tratamientos y poder de esta forma procesarlos a través de pruebas estadísticas paramétricas como el análisis de varianza.

El efecto principal que deberá medirse cuando se aplican fungicidas tiene que ver con la *incidencia* y la *severidad* de una enfermedad. Tanto la incidencia como la severidad son variables que se miden en campo a través de los síntomas de una enfermedad, por ello, se dice que es una forma indirecta de cuantificar el ataque de fitopatógenos. Este tipo de valoración demanda que en los ensayos sobre enfermedades de cultivos se corrobore la relación causal de la misma, definiendo al patógeno responsable de la enfermedad, ya sea a través de los signos de la misma cuando esto sea posible y/o mediante pruebas objetivas de laboratorio que demuestren la patogenicidad de la plaga aplicando el protocolo de los postulados de *Koch* (Koch, 1882). La incidencia es una variable cuantitativa que tiene por objeto valorar la velocidad y el patrón de avance de una enfermedad en el campo; mientras que la severidad es una variable, por lo general semicuantitativa, que complementa a la incidencia describiendo el nivel de daño que presenta el patógeno.

La incidencia es la afectación relativa de una determinada enfermedad sobre un cultivo, que se establece en las unidades de muestreo a partir de la selección aleatoria de un mínimo de 10 plantas o partes de ella (frutos, hojas, raíces, entre otras). La severidad es un tipo de variable más específica que la incidencia, en donde una vez que se identifican las plantas afectadas, se valora el nivel de daño en algunas de sus estructuras u órganos (hoja, tallo, raíces o frutos), utilizando para ello escalas estandarizadas o arbitrarias que toman en consideración el área o proporción de afectación en cada una de estas estructuras. Una de estas escalas muy utilizadas para valorar la severidad de la sigatoka negra (*Mycosphaerella fijiensis*) en hojas de banano es la conocida escala de *Stover*, modificada por Gauhl (1989), que se indica en la tabla 1.3. La incidencia generalmente se expresa en forma porcentual; mientras que la severidad se expresa en términos absolutos (grado de la escala) y que, para el procesamiento estadístico, debe al final reportarse en porcentaje según la correspondencia de la escala utilizada.

Dado el carácter cuantitativo de la incidencia, ésta se establece contabilizando las plantas u órganos vegetativos afectados entre el total muestreado. La severidad, cuya característica es ser semicuantitativa, se establece mediante una media ponderada entre el número de individuos afectados y clasificados de acuerdo a la magnitud de dicha afectación que es valorada con el apoyo de escalas como la antes mencionada, entre la totalidad de plantas u órganos muestreados en la unidad experimental. Bajo la apreciación indicada, las expresiones genéricas para cuantificar la incidencia y la severidad de una enfermedad son las siguientes:

$$I = \frac{100(N^\circ \text{ individuos afectados})}{N^\circ \text{ total de individuos muestreados}} \quad (1.1)$$

$$S = \frac{\sum n_i G_i}{n} \quad (1.2)$$

Como es deducible, en la expresión 1.1, la incidencia I está dada en porcentaje. En la expresión 1.2, la severidad S está en términos absolutos, dados por los grados de una escala; mientras que n_i son los elementos (plantas o partes de ella) seleccionados que se clasifican de acuerdo a un grado G_i de la escala y n es el total de elementos. La selección para las dos evaluaciones se realiza de forma aleatoria.

Tabla 1.3 Escalas de severidad de la sigatoka negra en banano.

Grado de severidad	Descripción de la severidad
0	Sin síntomas
1	Menos del 1% de área foliar afectada y/o hasta 10 estrías visibles
2	1 - 5% de área foliar afectada
3	6 - 15% de área foliar afectada
4	16 - 33% de área foliar afectada
5	34 - 50% de área foliar afectada
6	Más del 50% del área foliar afectada

Fuente: Stover, modificada por Gauhl (1989).

La distribución más comúnmente utilizada en los experimentos con plaguicidas, en el caso de los cultivos con pequeños distanciamientos de siembra, es el de bloques completos al azar; sin

embargo, cuando las parcelas incrementan sus dimensiones sustancialmente, una de las alternativas recomendadas es utilizar un diseño de doble bloqueo como el cuadrado latino.

Eficacia

Una de las variables importantes que debe ser valorada en los ensayos con plaguicidas es la eficacia de los tratamientos propuestos, definida ésta como la magnitud relativa (porcentual) de control de la plaga o enfermedad. Esta eficacia puede ser evaluada a nivel de infestación (individuos vivos) o de mortalidad de un insecto plaga, posterior a la aplicación de los pesticidas.

A nivel de infestación de insectos, una de las expresiones muy aceptada cuando la plaga se distribuye de forma dispersa o heterogénea en el ensayo, es la propuesta por Henderson y Tilton (1955). Esta fórmula valora la magnitud de la infestación a partir de la presencia de insectos dañinos, antes y después de aplicar los tratamientos, comparando este dato con la población de estos insectos en la parcela que se tiene como testigo absoluto. La expresión propuesta por estos autores es:

$$E = 100 \left[1 - \frac{I_{T_0A} \times I_{TD}}{I_{TA} \times I_{T_0D}} \right] \quad (1.3)$$

En la expresión 1.3, E es la eficacia (%), I_{T_0A} e I_{T_0D} es la infestación de la plaga en la parcela utilizada como testigo absoluto, antes y después de aplicar el tratamiento, respectivamente; I_{TA} e I_{TD} es la infestación en alguna parcela en donde se asigna uno de los tratamientos, antes y después de su aplicación, respectivamente. En ciertos casos, cuando existe un tiempo prudencial entre la aplicación del tratamiento y la evaluación posterior de su efecto, la expresión 1.3 también puede utilizarse para valorar la eficacia de alternativas para el control de enfermedades considerando para ello la incidencia.

Cuando la presencia de los insectos dañinos es relativamente homogénea, la eficacia se determina por la relación porcentual que es establecida dividiendo la diferencia entre la infestación de la parcela que se tiene como testigo absoluto (sin aplicación de tratamiento) (I_{T_0}) y la infestación en donde se aplica una de las alternativas o tratamientos propuestos (que puede ser un tipo de insecticida o una dosis de éste), para el primero de estos dos valores. La fórmula descrita fue originalmente propuesta por Abbott (1925). En este sentido, la eficacia E para plaguicidas, según este autor, puede también determinarse mediante la expresión:

$$E = \frac{100(I_{T_0} - I_T)}{I_T} \quad (1.4)$$

La indicación de que la expresión 1.4 se utilice cuando la plaga muestra cierta homogeneidad de su ataque va en el sentido de que fue propuesta para los casos de daños severos a los cultivos; por ejemplo, es recomendable su uso en ensayos que involucran a insectos como *Spodoptera frugiperda* en maíz o *Prodidiplosis longifila* en tomates. Cabe indicar que la expresión 1.4 también puede utilizarse para determinar la eficacia de alternativas para el control de enfermedades, considerando la incidencia que podría valorarse en forma transversal cuando se tiene un testigo absoluto, o de forma longitudinal cuando se toman lecturas iniciales (antes de la aplicación de los tratamientos) de la incidencia en la parcela testigo.

Cuando la eficacia de un insecticida se determina a partir de la mortalidad de una plaga, puede utilizarse la metodología propuesta por los autores Sun y Shepard (Püntener, 1981), considerando también lo que sucede en el testigo absoluto y la presencia de la plaga antes y después de la aplicación de los tratamientos. La expresión propuesta por estos autores es la siguiente:

$$E = 100 \left[\frac{\frac{100(I_{TA} - I_{TD})}{I_{TA}} \pm \frac{100(I_{TD} - I_{TA})}{I_{TA}}}{100 \pm \frac{100(I_{TD} - I_{TA})}{I_{TA}}} \right] \quad (1.5)$$

En la expresión 1.5, la razón $100(I_{TA} - I_{TD})/I_{TA}$ representa a la mortalidad de la plaga en valor porcentual.

Generalmente, la infestación de una plaga se valora de manera cuantitativa a través del número de insectos, ya sea por planta, por algún órgano de ésta (hojas, tallos o frutos) o por área foliar. Luego este valor, en algunos casos, se contrasta con los grados de infestación de alguna escala estándar o estructurada¹² para identificar la magnitud de dicha infestación. Para esto se seleccionan al menos 30 elementos (entiéndase como plantas, hojas, tallos, frutos, raíces, etc.) de forma aleatoria desde la unidad de muestreo y el dato se reporta como una media. Cuando se requiera reportar la *Intensidad de daño* de un insecto (valoración análoga a la severidad de una enfermedad) se utiliza la expresión 1.2, relacionando el número de órganos o plantas atacadas (n_i) en cada grado (G_i) de alguna escala para valorar el daño expresado en forma porcentual y el número total (n) de individuos seleccionados.

Otro aspecto a tener en cuenta en la definición de la eficacia de los plaguicidas está relacionado a la interrogante ¿qué evaluación debe considerarse para el reporte del dato de todo un grupo de evaluaciones subsecuentes que se hayan realizado? Como es conocido, cuando se valoran los efectos de los pesticidas, estas evaluaciones generalmente se realizan varias veces después de que se aplican los tratamientos; por tal motivo, la recomendación para establecer los datos que entrarán al análisis estadístico de la eficacia es la siguiente:

Insecticidas	Si lo que se evalúa es presencia o infestación de la plaga, los datos que se deberán considerar deben ser de la primera evaluación. Cuando se evalúa la intensidad de daño al cultivo, los datos que se deben considerar serán los de la última lectura, previo a la identificación de la estabilización de la variable que se visualiza en una gráfica de dispersión elaborada con los datos de las distintas evaluaciones realizadas durante el proceso experimental.
Fungicidas	Valorando la incidencia, los datos que deben considerarse para definir la eficacia deben ser los últimos de todas las evaluaciones, siempre que se tenga un testigo absoluto; caso contrario, se deberán contrastar entre valores finales e iniciales de una misma unidad experimental.
Herbicidas	Deberán considerarse los datos de la última evaluación, ya sea cualitativa o cuantitativamente.

¹² Una escala estructurada es aquella que el investigador la define en base a escalas con información de otros cultivos con características similares, debido a la ausencia de escalas estandarizadas que son las que se establecen a través de procesos experimentales que relacionan el rendimiento con diferentes grados de daño de una plaga bajo un rigor exhaustivo. En esta última se definen los umbrales económicos o de acción a partir de los cuales se procede con los métodos de control.

Experimentos relativos a ciertas prácticas culturales

Dentro de esta categoría se incluyen una serie de experimentos que tienen sus particularidades basados en ciertas labores culturales de prueba. Entre estos casos se tienen a los experimentos sobre distanciamientos de siembra, métodos de preparación de suelos, diferentes alternativas de poda, diferentes dosis de riego, dosis de encalado, frecuencias de corte de pastos, entre otras labores.

Cada una de estas alternativas experimentales tiene sus condiciones que deben ser previstas para mantener la homogeneidad del área experimental. Así, por ejemplo, en los experimentos sobre distancias de siembra es de esperar que las parcelas no tengan las mismas medidas, pero sí la misma cantidad de hileras de plantas; situación que podría afectar a los resultados, ya que habrá algunas parcelas que estarán expuestas a una mayor variabilidad por efecto del incremento de área de las mismas.

Los experimentos agrícolas tienen sus limitaciones que los hacen diferentes, como ya se lo mencionó al inicio, a los experimentos que se realizan en otras áreas. Sin embargo, cuando se tenga más de un factor por evaluar, indiscutiblemente siempre será mejor desarrollar un experimento factorial antes que realizar experimentos subsecuentes con un solo factor a la vez. Generalmente, los experimentos factoriales que se desarrollan en la agronomía no se ajustan a procesos de optimización factorial que sí se aplica en el campo industrial, ya que alternativas como la metodología de superficie de respuesta que definen puntos óptimos resultan muy caros de establecer; además, la aplicación práctica de los resultados experimentales de esta metodología, en el caso de la agronomía, tiene poca duración y el costo invertido no es recuperable.

CAPÍTULO II

LA UNIDAD EXPERIMENTAL

Gran parte del génesis del error sistemático que puede tener un experimento está en la inadecuada definición y selección de la unidad experimental. Si esta acción fundamental en el diseño de un experimento se la realiza a la ligera, el riesgo creciente puede no ser la aceptación de un falso negativo, sino de un falso positivo, que es más terrible...



El tamaño de la parcela en los experimentos agrícolas

La insensibilidad¹³ que un experimento pudiera tener está ligado al error tipo II de la estadística, en donde el riesgo que se corre en este ámbito de estudio es prácticamente nulo; empero, el mejoramiento productivo de los cultivos también depende del grado de sensibilidad que tenga el diseño experimental que se haya decidido implementar para la obtención de respuestas realmente significativas. Respuestas que pueden ser soluciones a una problemática y que por falta de sensibilidad experimental se obvian. De allí que el grado de sensibilidad sea inversamente proporcional a la magnitud del error experimental que se presenta en todo experimento; y consiguientemente, el éxito de tener un experimento sensible dependerá de cuánto control hagamos de dicho error experimental.

El error experimental, de acuerdo a lo manifestado en el capítulo I, representa a la variación no explicada que se produce en un experimento bien cuidado; sin embargo, en la práctica, se debe fundamentalmente a dos fuentes de variabilidad. Una de ellas corresponde al manejo no uniforme que se realiza en las unidades experimentales (error sistemático), mientras que la otra está referida básicamente a la heterogeneidad del suelo (Reyes, 1984). Si las condiciones de manejo se mantienen uniformes, como es de esperar al desarrollar un experimento, y si además se utilizan semillas de genética certificada, la heterogeneidad del suelo es el factor que más contribuirá dentro de la variación total correspondiente al error experimental.

El error experimental jamás podrá ser eliminado, pero sí se puede reducir su magnitud llevando a cabo ciertas acciones, que a su vez permitirán una mejor estimación de los efectos de los tratamientos que se estén evaluando. De allí que una de estas acciones que inciden sobre el control de esta fuente de variabilidad desconocida (siempre que exista un manejo uniforme del ensayo) sea el uso de tamaños y formas adecuadas de las unidades experimentales. Por lo tanto, debemos estar conscientes de que la información que hace posible cuantificar el error experimental de un ensayo proviene de las unidades experimentales; por consiguiente, el tamaño adecuado de estas últimas siempre será uno de los requerimientos importantes dentro del diseño experimental.

Definir la forma y el tamaño adecuado de las unidades experimentales, que en el caso de la agronomía incumben a las medidas que deberán tener las parcelas de los experimentos de campo, contribuyen, por un lado, a aumentar la sensibilidad de un experimento, y por otro, a incrementar la evidencia de las diferencias de tratamientos, de entre los cuales pueden estar algunas soluciones para mejorar la productividad agrícola. Un tamaño adecuado de la una unidad experimental puede hacer eficiente un experimento en el sentido del ahorro económico que pudiera lograrse al evitar utilizar parcelas grandes de forma innecesaria, que lo único que harán es magnificar el efecto distorsionador de la heterogeneidad del suelo. Sin embargo, uno de los parámetros muy difíciles de establecer *a priori* en la experimentación agrícola es el tamaño adecuado que deberá tener la parcela o unidad experimental en donde se evaluará el efecto de un tratamiento aplicado; debido a que tanto la forma como las dimensiones no sólo dependen de la variabilidad que presenta el suelo, sino de otras condiciones adicionales.

Por otro lado, existen evidencias en la literatura disponible indicando que en la precisión de un experimento incide con mayor contundencia el tamaño de la parcela antes que la forma de ésta. De hecho, en prácticamente todos los experimentos agrícolas en donde se han relacionado distintos tamaños de parcela con el coeficiente de variación respectivo, se ha podido establecer que este coeficiente disminuye a medida que las parcelas se hacen más grandes. Sin embargo, esta disminución de la variabilidad no tiene una relación inversamente proporcional (en forma lineal) respecto del tamaño de parcela. Siempre existe un límite a partir del cual, aumentos subsecuentes de tamaños de parcela no aportan disminuciones significativas del error

¹³ Insensibilidad debe entenderse aquí como la característica de no detección de cambios o diferencias (significativas) entre tratamientos en los experimentos. Lo contrario a esto sería sensibilidad.

experimental; más bien este último tiende a estabilizarse a partir de ciertos tamaños de parcela, ya que los cambios de variabilidad son insignificantes y lo único que se obtendría con tamaños que exceden estos límites es aumentar los costos del experimento, más no su precisión efectiva. Esta última relación es la que se aprovecha en el método de máxima curvatura que se indica más adelante, como alternativa para definir los tamaños de las unidades experimentales.

Si bien con mayores tamaños de parcela se puede apreciar una reducción en el error estándar o en el coeficiente de variabilidad cuando se realizan los llamados ensayos en blanco o de uniformidad (los cuales se especifican más adelante), es claro que en la práctica parcelas excesivamente grandes harán que distribuciones como la de bloques al azar pierdan eficacia; ya que dichos bloques se extenderán excesivamente, de tal forma que podrían ubicarse sobre condiciones de suelo muy heterogéneas (dentro de un mismo bloque), imposibles de separarlas por parte de esta distribución, lo que al final acrecentará la magnitud del error experimental.

Factores de influencia en la dimensión de parcelas experimentales

El investigador agrícola debe estar consciente de que en la práctica no existen medidas o tamaños estandarizados de las parcelas experimentales, ya que la selección de un tamaño definido requiere que se consideren varios aspectos coyunturales. De hecho, existen múltiples factores que pueden presentar ciertas restricciones respecto de la dimensión parcelaria, de entre los cuales están: el tipo de cultivo, tipo de experimento, las variables a medirse y fundamentalmente, la heterogeneidad del suelo en donde se llevará a cabo la experiencia.

El tipo de cultivo que se utilice en un estudio tiene una gran influencia en la dimensión de una parcela experimental. No es lo mismo realizar un ensayo con arroz y otro con un cultivo de distanciamiento amplio como la palma africana. En el primer caso, con áreas de unos 20 m² podría ser suficiente como unidad experimental; mientras que, en el segundo, para poder obtener un número suficiente de plantas en la evaluación, el tamaño de la unidad experimental deberá ser relativamente mucho más grande.

En el caso de unidades experimentales con plantas de distanciamientos amplios, si éstas son escasas dentro de la parcela, la variabilidad que puede existir de una a otra planta podría aumentar el error experimental de forma considerable. Por otro lado, una mayor cantidad de plantas por unidad experimental propiciará que los bloques, si es que se decide por una distribución con bloques al azar, se extiendan de manera importante dentro del campo, con lo cual el experimento se expone a una mayor variabilidad del medio ocasionada por una mayor heterogeneidad que pudiera presentar el suelo (pendiente, fertilidad, humedad, etc.). En este sentido, se podría utilizar una distribución con doble bloqueo como la de cuadrado latino, tal vez considerando sólo una planta como unidad experimental, pero teniendo en cuenta la cantidad de tratamientos y el número de réplicas que sean factibles de considerar. Incluso se podrá hacer uso de distribuciones con bloques incompletos que permitan un mejor control de la varianza del error experimental.

Otro de los elementos importantes que definen el tamaño de las parcelas experimentales tiene que ver con el tipo de ensayo que se pretende desarrollar, debido a la variabilidad que se pudieran esperar según sea la característica del experimento. Cuando se trata de experimentos sobre plagas, debido a lo estocástico que puede ser la infestación, la incidencia o el ataque de la misma, necesariamente deben utilizarse parcelas con un tamaño relativamente grande; aun si la experiencia involucra aplicaciones de tratamientos preventivos y la evaluación se realiza a través de variables indirectas como el rendimiento del cultivo. Cuando el estudio trata sobre nutrición de cultivos y los tratamientos representan sustancias inertes como los fertilizantes, el tamaño de la parcela es mucho más pequeño; dado que su efecto, si la aplicación se realiza con el suficiente cuidado, es más homogéneo. Experimentos de cruce genético, debido a la

gran cantidad de variedades que se utilizan y muchas veces a la escasez de semillas que se tienen en estos casos, utilizan parcelas de tamaño reducido. En otros casos podrían requerirse tamaños de parcelas grandes, no tanto por la heterogeneidad de la respuesta, sino por la forma como se aplican los tratamientos; bajo esta consideración se encuentran, por ejemplo, los métodos de preparación de suelos en donde se utilizan máquinas, las cuales evidentemente necesitan de áreas grandes para desarrollar su labor.

En el caso de las variables de respuesta que se miden en las parcelas experimentales, éstas a su vez deben tener una relación exclusiva con los tratamientos en estudio. Por ejemplo, si la única variable de respuesta es el rendimiento del cultivo, ésta podrá evaluarse en áreas de parcela relativamente pequeña; no obstante, para variables como el ataque de un hongo, deberán considerarse parcelas experimentales grandes, de tal forma que se pueda evaluar el nivel de incidencia. Esta última, al ser una variable que depende de la respuesta de un organismo vivo, hace que la presencia en el cultivo difícilmente tenga una distribución normal.

La heterogeneidad del suelo es el factor restrictivo, tal vez el más importante, que influye no solo en el tamaño de las parcelas sino también en el tipo de distribución experimental que debe implementarse. Esta heterogeneidad puede estar determinada por características físicas, químicas y hasta microbiológicas del suelo; las que, si no son tomadas en cuenta, pueden llevar al traste los resultados de un experimento y a su vez también incurrir en el error tipo I de la estadística. Además de estas características, la variabilidad que presenta un suelo puede estar influenciado por ciertas actividades agrícolas como el uso de agroquímicos, la quema de residuos de cosechas, la distribución de las malezas y los cambios en los niveles de la freática por las cercanías de los canales de riego y drenaje.

Procedimientos para establecer la heterogeneidad del suelo

La determinación del grado de heterogeneidad de los suelos puede realizarse mediante el empleo del coeficiente de correlación (r), que es una alternativa muy útil para analizar la variabilidad que presentan los mismos. Para este propósito se utilizan los *experimentos en blanco*, los cuales hacen referencia a ensayos en donde se siembra un cultivo anual (por ejemplo, maíz) de semilla certificada, cuyos cuidados en actividades comunes como el riego, control de malezas, fertilización y demás labores culturales se realizan de forma homogénea en toda la superficie del lote experimental; por lo cual también se los identifica como *experimentos de uniformidad*. Para esta práctica, una vez llegada la cosecha, se divide el campo, en donde se pretende evaluar la heterogeneidad, en un suficiente número de parcelas (llamadas unidades básicas o también subparcelas) del mismo tamaño y forma, en cada una de las cuales se evalúan los niveles de producción de la especie indicadora.

El uso del coeficiente de correlación se justifica en el sentido de que dentro de cualquier lote agrícola que se seleccione para experimentos en blanco, una heterogeneidad relativamente fuerte estará ligada al hecho de que parcelas adyacentes de un espacio determinado arrojen producciones altas; mientras que parcelas también adyacentes, pero alejadas de las primeras, presenten producciones menores. En este caso, el coeficiente de correlación será alto. Si la heterogeneidad es mínima dentro de todo el ensayo en blanco, presentando una variabilidad que obedezca exclusivamente a la casualidad, no debe existir una correspondencia entre las producciones de los distintos grupos de parcelas adyacentes; por lo tanto, la correlación que se obtenga en estas circunstancias debe ser un valor reducido.

Para establecer el coeficiente de correlación se enumeran las parcelas. Luego se deberán establecer dos grupos de parcelas contiguas; uno formado por las parcelas impares (1, 3, 5, 7, 9, ...) y el otro por las parcelas pares (2, 4, 6, 8, ...), de manera que las parcelas 1 y 2, 3 y 4, 5 y 6, etc., se correspondan. Como será deducible, el número de parcelas en que se ha dividido

el campo debe ser un número par para tener el mismo número de parejas correlacionadas. Una vez llegado el momento, se procede a la cosecha de cada parcela con cuyos datos se determina la existencia de correlación mediante la expresión indicada a continuación:

$$r = \frac{n \sum XY - \sum X \sum Y}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}} \quad (2.1)$$

En la expresión 2.1, X representa al rendimiento de las parcelas impares y Y al rendimiento de las parcelas pares; mientras que n es el número de parejas formadas.

Como se recordará, el valor de r puede variar en el rango de +1 a -1. Mientras este valor se acerca a la unidad, sea ésta positiva o negativa, la correlación es más importante; por otro lado, si este valor es o se aproxima a 0, la correlación es inexistente. La verificación de significancia del valor de r puede realizarse mediante una prueba de hipótesis, utilizando la función de densidad t de Student. En este caso, el estadístico de prueba está dado mediante la expresión siguiente:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad (2.2)$$

Para la cual, si la hipótesis nula (H_0) es verdadera, sigue la distribución t de Student, con $n - 2$ grados de libertad.

Finalmente, si el coeficiente de correlación resulta significativo (próximo a +1 o a -1), el suelo es heterogéneo. Caso contrario, cuando el coeficiente de correlación no registra significancia o éste tiende a 0, el suelo es relativamente homogéneo.

Otra de las alternativas para determinar la heterogeneidad de un suelo mucho más efectiva es la denominada ley de varianzas de Smith (Smith, 1938), a través de la cual se determina la existencia de una relación empírica entre el logaritmo de la varianza de las parcelas de un tamaño determinado y el logaritmo del tamaño de estas parcelas. Esta ley, según este autor, está definida como:

$$V_x = \frac{V_i}{X^b} \quad (2.3)$$

En donde V_x es la varianza del rendimiento entre parcelas conformadas por las diferentes combinaciones de unidades básicas (agrupaciones de tamaño X), V_i es la varianza entre unidades básicas, X es el área de parcelas en las diferentes combinaciones de unidades básicas y b se conoce como coeficiente de heterogeneidad de Smith.

La expresión 2.3 se linealiza aplicando logaritmos, con lo que se tiene la nueva función lineal:

$$\text{Log}(V_x) = \text{Log}(V_i) - b\text{Log}(X) \quad (2.4)$$

En esta última función, b es la pendiente que toma valores entre 0 y 1, que al igual que el coeficiente de correlación, para valores que tienden hacia el 0 las parcelas se definen como relativamente homogéneas, mientras que valores que tiendan hacia el 1 se definen como heterogéneas. Algo que debe tenerse presente es que, dada la relación inversa entre el tamaño

de la parcela y la variabilidad producida por la heterogeneidad del suelo, el valor de b comúnmente es negativo; no obstante, lo que siempre se considera es su valor absoluto. Este coeficiente de heterogeneidad se establece aplicando el método de los mínimos cuadrados ordinarios a la expresión 2.4.

En ciertos casos b puede presentar valores mayores a 1, lo que complica su interpretación. Por esta causa, varios autores recomiendan estimar el coeficiente de heterogeneidad del suelo ponderando los logaritmos de las varianzas por los grados de libertad asociados de cada combinación. Esto se hace de la siguiente forma:

$$b = \frac{\sum(G_L Q_i B_i) - \sum(G_L Q_i) \cdot \frac{\sum(G_L B_i)}{\sum G_L}}{\sum(G_L B_i^2) - \frac{\sum(G_L B_i)^2}{\sum G_L}} \quad (2.5)$$

Siendo G_L los grados de libertad asociados a la varianza de los distintos tamaños de parcela (combinaciones de unidades básicas), definido como $n - 1$; Q_i es el logaritmo de la varianza de rendimiento entre parcelas por unidad de área básica y B_i es el logaritmo del número de unidades básicas en cada tamaño de parcela.

El coeficiente de correlación y el coeficiente de heterogeneidad de Smith sólo son indicativos de la idoneidad de un campo para llevar a cabo los experimentos de interés, ya que permiten conocer si la heterogeneidad del suelo es importante, ya sea para comparar distintos suelos para experimentación (comparación que también puede realizarse a través del coeficiente de variación) o incluso como para decidir buscar otros campos experimentales menos heterogéneos. No obstante, una alternativa decisiva con la cual se podría también establecer el tamaño y la forma como deben distribuirse las parcelas, es la que presenta un mapa de heterogeneidad de suelo; una imagen del campo experimental que es complementario a los índices de heterogeneidad indicados y que son factibles de elaboración, ya que se realizan con los mismos resultados de producción de los experimentos en blanco.

Mapa de heterogeneidad de suelo

Los mapas de heterogeneidad se realizan para identificar espacialmente las distintas magnitudes de variabilidad que presenta un suelo que se ha seleccionado para la experimentación agrícola. La información que se utiliza para elaborar estos mapas también se obtiene a partir de los experimentos en blanco. Para esto se selecciona un cultivo, el mismo que es sembrado en el área de interés; el cual previamente se divide en un determinado número de parcelas. Todas las labores culturales (control de malezas, riego, fertilización, control de plagas, etc.) deben realizarse en igualdad de condiciones en cada una de las parcelas en que se haya dividido el terreno. Llegado el momento, se hace la cosecha y el rendimiento que se obtiene en cada una de estas unidades experimentales se toman de referencia (cotas), para interpolar y elaborar las curvas de igual valor, que en este caso corresponderá a las curvas de igual fertilidad. Los puntos que se unen con las isolíneas en el plano, son los centros geométricos de cada parcela que se hacen coincidir con los rendimientos obtenidos en cada una de ellas. Este es un trabajo análogo al de obtener curvas de igual nivel en altimetría.

En la figura 2.1 se expone un esquema de un mapa de heterogeneidad, indicando las respectivas isolíneas de fertilidad, obtenidas a partir de los datos de rendimiento cosechados en cada unidad básica de muestreo de un ejemplo hipotético. En este punto, es preciso indicar que todas las investigaciones que se realizan para definir el tamaño adecuado de parcela, en

realidad determinan el tamaño de la *unidad de muestreo* o lo que mal se conoce como *parcela útil*; lo que es diferente al tamaño de la unidad experimental, la cual es la parcela en la que también se considera el efecto de borde.

Cuando las isolíneas se muestran muy juntas, la variabilidad (pendiente) de la fertilidad es marcada; mientras que líneas más separadas indican una mayor homogeneidad, debido a que la pendiente de fertilidad es mínima. Bajo estas condiciones, de manera aproximada pueden establecerse las dimensiones de las parcelas.

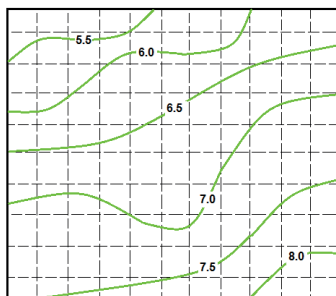


Figura 2.1 Esquema de líneas de igual fertilidad en un mapa de heterogeneidad de suelo.

Con los mapas de heterogeneidad también se facilita la identificación en el área experimental de zonas críticas, con reportes de rendimiento muy por arriba o por debajo de la media, que podrían ser causa de áreas localizadas de excesos de materia orgánica, niveles freáticos altos, fuertes concentraciones salinas, etc. Las delimitaciones de estas zonas críticas facilitan sus exclusiones al momento de implantar un experimento determinado.

Asimismo, otro de los aportes que brindan los mapas de heterogeneidad es cómo deberá realizarse el control local de un experimento, esto es, la distribución de los bloques. Para una situación dada, la poca pendiente que pueda observarse a través de las isolíneas, permitirá inclusive la selección de la distribución completamente al azar, la cual es una alternativa para utilizarse en ambientes homogéneos. La intensidad de la variabilidad de fertilidad que muestre el campo y la dirección que se evidencie, harán posible identificar la posición de cómo deberán colocarse los bloques para cuando sea necesario utilizar una distribución de bloques al azar; o incluso utilizar una distribución en cuadro latino, si el campo muestra dos direcciones ortogonales de fertilidad.

Otra opción de elaborar los mapas de heterogeneidad, y también de calcular los coeficientes de heterogeneidad (coeficiente de correlación), es mediante la medición de variables físicas o químicas del propio suelo en vez de realizar medidas de rendimientos de cultivos cuando se recurre a los experimentos en blanco. En estos casos el campo para ensayos también se divide en unidades básicas, desde las cuales se extraen muestras para evaluar la humedad, contenidos de materia orgánica o contenidos de nutrientes como nitrógeno, fósforo o potasio.

La identificación de la dispersión de la heterogeneidad en un campo experimental, por lo general, se mantiene en el tiempo; de allí que una vez identificada la dirección de la fertilidad, ésta se constituirá en una norma para ese suelo en cuanto al tipo de distribución que se adoptará y/o la dirección de cómo deberán colocarse los bloques de distribuciones como, por ejemplo, la DBCA.

Métodos analíticos aplicados en los ensayos en blanco para establecer el tamaño de las unidades de muestreo

Para determinar los tamaños de las unidades básicas de muestreo se recurre al uso del coeficiente de variación que presente el campo entre los datos recogidos en cada una de las unidades básicas en que se dividió el mismo y de los que se obtienen, subsecuentemente, según las combinaciones que se realicen de estas unidades básicas. Esto es, el tamaño de las unidades de muestreo está supeditado a la variabilidad que existe en el suelo, para lo cual debe haber una combinación de dimensiones (ancho y largo) en la que dicha variabilidad defina el tamaño óptimo de parcela.

La obtención del tamaño adecuado de las unidades de muestreo inicia en un experimento en blanco, y como ya se lo mencionó, va a depender del tipo de cultivo con los cuales se desarrollarán los experimentos subsiguientes y de las variables de respuesta que se evaluarán. Se siembra una semilla de pureza certificada en áreas que bien pueden ir desde los 20 m x 20 m hasta los 50 m x 50 m. Posteriormente, se inicia seleccionando como ancho de la unidad de muestreo una hilera del cultivo por 1 m de longitud, luego se realizan distintas combinaciones de dimensiones que podrían llegar hasta unidades de muestreo de 10 m x 10 m. Para definir el tamaño de las unidades de muestreo se cosechan todas las unidades básicas y se realizan distintas combinaciones para ir verificando, a partir del rendimiento, la variabilidad de cada dimensión que resulte de una combinación. Este procedimiento es básicamente el que se realiza en los distintos métodos que se indican en los párrafos siguientes.

Método de máxima curvatura

Una opción para establecer el tamaño de parcela óptimo es definiendo una gráfica que relacione la variabilidad y el rendimiento obtenido en las diferentes combinaciones de unidades básicas de muestreo. Para cada combinación de dimensiones de parcela se obtiene el respectivo coeficiente de variación, con los cuales puede realizarse una gráfica de dispersión con ajuste de alguna curva (generalmente de tipo potencial) y a partir de allí seleccionar el tamaño de parcela que registre la menor variación. Esta metodología, en donde se selecciona el tamaño de parcela con la variabilidad en el punto de mayor inflexión, es decir, en donde la variación empieza a ser mínima mientras se aumenta el tamaño de la parcela, se conoce como el de *máxima curvatura*.

En la figura 2.2, a manera de ejemplo, se presenta un esquema de la relación entre el coeficiente de variación obtenido para cada dimensión de parcela en que se ha considerado el área experimental a través de una curva ajustada. Según algunos estudios realizados por varios investigadores, se ha establecido que existe una disminución del valor de CV a medida que las parcelas tienen dimensiones relativamente mayores. Uno de estos resultados se detalla en la tabla 2.1 respecto de un experimento realizado sobre la determinación del tamaño óptimo de parcela experimental para el cultivo de yuca, realizado por Tineo y Villasmil (1988). En esta tabla puede notarse cómo varía ascendentemente el valor de CV cuando disminuye la cantidad de plantas por hilera para un mismo número de éstas; es decir, si el tamaño de parcela disminuye el CV se incrementa.

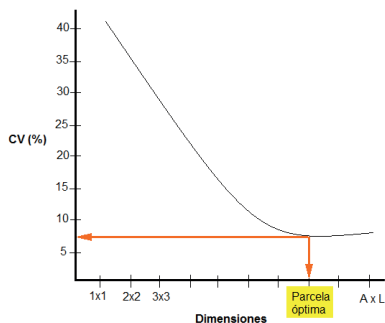


Figura 2.2. Relación entre el coeficiente de variación y las dimensiones (ancho x largo) de las unidades de muestreo.

Asimismo, según lo que se indica en la tabla 2.1, si se incrementa el ancho de la parcela (mayor número de hileras) para un mismo coeficiente de variación, se podrá disminuir la longitud de dichas hileras (menor cantidad de plantas por hilera). Por esta situación, existe la recomendación de utilizar parcelas cuadradas cuando no se tenga una definición del gradiente de variación de un campo; sin embargo, cuando este gradiente está identificado, la recomendación es utilizar parcelas rectangulares con el lado más largo paralelo a dicho gradiente.

Tabla 2.1. Coeficientes de variación para diferentes combinaciones de dimensiones parcela para un cultivo de yuca.

CV (%)	Ancho (N.º hileras)	Largo (plantas/hilera)	CV (%)	Ancho (N.º hileras)	Largo (plantas/hilera)
8	1	Sin solución	16	1	14
8	2	24	16	2	12
8	3	20	16	3	11
8	4	19	16	4	11
10	1	21	18	1	12
10	2	19	18	2	11
10	3	17	18	3	10
10	4	16	18	4	9
12	1	18	20	1	10
12	2	16	20	2	10
12	3	15	20	3	9
12	4	14	20	4	8
14	1	15			
14	2	14			
14	3	13			
14	4	12			

Fuente: Tineo J. y Villasmil J. (1988).

La figura 2.3 es uno de los resultados obtenidos de un experimento en blanco con el fin de deducir el tamaño adecuado de parcela experimental para el cultivo de arroz, realizado por el autor de esta obra. El experimento se desarrolló en un área de 1014,4 m², desde la cual se definieron 990 unidades básicas de 0,9 m² y con las cuales se formaron combinaciones de parcela desde 1 x 1 hasta 10 x 8. En cada una de estas combinaciones se obtuvo el CV relacionadas al área de cada una de ellas, lo cual corresponde a los puntos que se pueden

observar en la gráfica de dispersión de la figura 2.3. Estos puntos generaron un ajuste significativo ($r^2 = 0,9351$) para una función potencial que en la figura puede notarse. Al aplicar el criterio de la metodología de máxima curvatura, se logró establecer que el área adecuada (de mínima variación) de parcela estaba entre 5 a 10 m². Entiéndase este resultado como área de la 'parcela útil' (o de muestreo) de la unidad experimental.

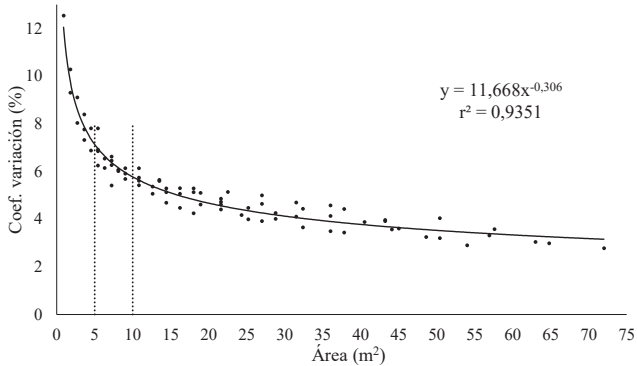


Figura 2.3. Determinación de parcela experimental (unidad de muestreo) para ensayos de arroz.

Para mejorar la objetividad del método de máxima curvatura puede aplicarse la prueba F de Fisher utilizando las varianzas de las distintas combinaciones de unidades básicas, con la cual se identifica el tamaño de parcela en donde las comparaciones empiezan a reportar variabilidad no significativa. En la tabla 2.2 se detalla esta forma de verificar el tamaño de parcela, en donde se puede observar la mínima variabilidad (varianza) cuando hay un cambio de 16 a 20 m² en el área de la unidad de muestreo en un experimento para determinar el tamaño adecuado de la unidad experimental en ensayos de arroz, publicado por Vargas y Navarro (2014), quienes utilizaron el método de curvatura máxima.

El método de máxima curvatura permite identificar el tamaño de parcela experimental adecuado en forma conjunta; es decir, es un método con el cual no es posible definir el ancho y el largo (forma de la parcela) de la unidad experimental de forma optimizada. Para este propósito se utilizan modelos multivariados como los de regresión múltiple, alternativa que se expone en los párrafos subsiguientes.

Tabla 2.2. Definición del tamaño de parcela aplicando el test *F* de Fisher para diferencias significativas de $P < 0,05$.

Tamaño de unidad experimental (m ²)	Varianza asociada a cada tamaño	Comparación de varianzas entre tamaños consecutivos	Valor de F entre varianzas de tamaños consecutivos	Probabilidad
1	9156,55			
2	21490,07	1 vs 2	2,35	0,000
4	53602,33	2 vs 4	2,49	0,000
5	75634,72	4 vs 5	1,41	0,006
8	124439,06	5 vs 8	1,65	0,003
10	187874,42	8 vs 10	1,51	0,014
16	331371,49	10 vs 16	1,76	0,020
20	461173,52	16 vs 20	1,39	0,176
25	371433,87	20 vs 25	1,24	0,330
40	877497,96	25 vs 40	2,36	0,053
50	926745,71	40 vs 50	1,06	0,426
80	1010271,80	50 vs 80	1,09	0,424
100	1172094,22	80 vs 100	1,16	0,422
200	573796,00	100 vs 200	2,04	0,178

Fuente: Vargas J. y Navarro J. (2014).

Métodos multivariados en la definición de los tamaños de parcela

Para obviar el uso de la gráfica indicada en la figura 2.2, una vez que se tienen los coeficientes de variación para distintas combinaciones de dimensión de parcela, se puede obtener un modelo de regresión múltiple como un polinomio de segundo orden que exprese una superficie de respuesta, en cuya concavidad se hace posible identificar el tamaño de parcela con la menor variación.

La expresión matemática del modelo de regresión que se aplica en estos casos se representa mediante la expresión 2.6, siendo su planteamiento el siguiente:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{12} X_1 X_2 + \varepsilon \quad (2.6)$$

En la expresión 2.6, *Y* es el coeficiente de variación, β_0 es el intercepto, β_1 y β_2 son los coeficientes lineales, β_{11} y β_{22} son los efectos cuadráticos o de curvatura del modelo, β_{12} es el efecto de interacción y ε es el error experimental. Las variables independientes están dadas por el ancho de la parcela (X_1) y la longitud de las mismas (X_2).

Este modelo (expresión 2.6) se utiliza para obtener las dimensiones de las unidades de muestreo desde dos perspectivas. La primera, permite identificar a la combinación (ancho y largo) que genera el punto de máxima curvatura que tiene un sentido económico; esto es, el punto a partir del cual cualquier incremento en las dimensiones de la parcela no justifica económicamente la mínima reducción que se logra en el coeficiente de variación. Este punto de máxima curvatura se obtiene aplicando derivadas parciales al modelo obtenido e igualando a -1 (este valor define al punto de máxima curvatura); de allí que a la izquierda de este punto la reducción del CV sea de importancia con respecto al incremento de las dimensiones de la

parcela, mientras que a su derecha la reducción del CV a un mayor aumento del tamaño de parcela sea insignificante.

El otro aspecto que establece el modelo como el de la expresión antes indicada (2.6), es el que está referido a determinar la dimensión de parcela para el mínimo coeficiente de variación; es decir, el tamaño de parcela en donde se tiene la influencia de la menor variabilidad del área de ensayo. Este punto también se obtiene de derivadas parciales del modelo, considerando una pendiente igual a cero.

De manera general, las derivadas parciales con las cuales se establecen estos puntos (óptimo económico y de mínima variabilidad) son:

$$\beta_1 + 2\beta_{11}X_1 + \beta_{12}X_2 = K \quad (2.7)$$

$$\beta_2 + 2\beta_{22}X_2 + \beta_{12}X_1 = K \quad (2.8)$$

Para resolver estos sistemas de ecuaciones, si se trata de hallar el punto de máxima curvatura para el óptimo económico, K deberá remplazarse por -1 ; y si lo que se busca es obtener el punto de mínima variabilidad, el valor a remplazar en K es 0. La resolución permitirá determinar el ancho (X_1) y el largo (X_2) de parcela en donde se optimiza (para $K = -1$) o minimiza (para $K = 0$) la variabilidad que presenta el suelo de un campo experimental.

Método de máxima curvatura bivariada

Otra de las metodologías para determinar el tamaño de parcela óptima es la propuesta por Barrientos (1981), conocida como *método de máxima curvatura bivariada*; el cual es una modificación al método de máxima curvatura, en donde sólo se relacionaban dos variables: el tamaño de la parcela y el coeficiente de variación. En este caso, la máxima curvatura bivariada relaciona el CV con el ancho y el largo de las parcelas, convirtiéndose en un modelo logarítmico múltiple, definido mediante la siguiente expresión:

$$Y = \beta_0 X_1^{\beta_1} X_2^{\beta_2} \varepsilon \quad (2.9)$$

siendo Y el coeficiente de variación; $\beta_0, \beta_1, \beta_2$ los parámetros de regresión desconocidos; X_1 y X_2 el ancho y el largo de parcela, respectivamente; y ε es el error experimental.

En el modelo 2.9, cuyos parámetros se obtienen aplicando logaritmos a las distintas dimensiones de ancho y largo de parcela, igual que en el caso de regresión múltiple (expresión 2.6), también se obtiene el tamaño de parcela óptimo cuando las derivadas parciales se igualan a -1 . Esto es, la pendiente -1 define la máxima curvatura a partir de la cual la variabilidad es insignificante comparada con el aumento del tamaño de parcela. Las dimensiones (ancho y largo) de parcela óptima, en función del modelo de máxima curvatura bivariada, se definen resolviendo el sistema de ecuaciones formado con las expresiones siguientes:

$$\beta_0 \beta_1 X_1^{\beta_1-1} X_2^{\beta_2} = -1 \quad (2.10)$$

$$\beta_0 \beta_2 X_1^{\beta_1} X_2^{\beta_2-1} = -1 \quad (2.11)$$

Una vez obtenidos los valores de X_1 y X_2 , se podrá estimar el valor de CV de parcela óptima para una situación específica, reemplazando estos valores en el modelo de la expresión 2.9.

Método de Hatheway

El método de Hatheway fue establecido para determinar el tamaño de parcela a partir del método de la Ley de varianzas de Smith y la expresión propuesta por Cochran y Cox para definir el número de repeticiones en un experimento. En su esencia, el método relaciona el tamaño de la parcela con el número de repeticiones y también con las diferencias que se espera encontrar entre los tratamientos que se evaluarán en un ensayo. Por esta causa, algunos autores la consideran como una de las metodologías más completas para este propósito. Tomando de base la fórmula propuesta por los autores Cochran y Cox para repeticiones, la expresión de esta metodología es:

$$X^b = \frac{2(t_1 + t_2)^2 CV^2}{Rd^2} \tag{2.12}$$

En donde X es el tamaño de la parcela expresada en número de unidades básicas, b es el coeficiente de heterogeneidad de Smith, t_1 es el valor de la distribución t de Student obtenido a partir de un nivel α pre-establecido y de los grados de libertad del error experimental; t_2 es también un valor inverso de la distribución t de Student para una probabilidad $\alpha = 2(1 - p)$, siendo p la probabilidad estimada de obtener una diferencia significativa; CV es el coeficiente de variación; R es el número de repeticiones y d es la diferencia que se pretende detectar, expresada como un porcentaje de la media muestral. Según el planteamiento de la expresión 2.12, se deduce que el tamaño de la parcela es directamente proporcional a la variación que existe en el suelo e inversamente proporcional al producto entre el número de repeticiones y la diferencia que se estime encontrar entre los tratamientos.

La descripción gráfica de la expresión 2.12 define una serie de curvas en donde evidentemente se relacionan el tamaño de la parcela (área), las repeticiones y el nivel de precisión requerido para el experimento. Un esquema de esta descripción se expone en la figura 2.4, en donde se observa que, al disminuir el área de muestreo, necesariamente debe incrementarse el número de repeticiones para mantener un mismo nivel de precisión; o viceversa, al disminuir el número de repeticiones, habrá que aumentar el área de muestreo de la parcela para poder tener igual precisión en las diferencias a detectar.

En la figura 2.4 se indica en línea roja el incremento del área de muestreo cuando se disminuye el número de repeticiones. Se puede observar que el área debe aumentar desde los 14,8 m² hasta los 24,4 m² para poder reducir el número de repeticiones de 5 a 3, manteniendo una precisión del 10%. Entiéndase como precisión a la diferencia que se quiere detectar entre tratamientos respecto del promedio; por ejemplo, si el promedio es de 2000 kg y se necesita detectar diferencias de unos 200 kg, la precisión deseada es de 10%.

Si bien el método de Hatheway es completo, éste tiene su falencia en la estimación de la precisión considerada. Esta precisión es una decisión subjetiva del investigador que, junto con los grados de libertad del error también establecidos de esta forma, hacen que el método reporte tamaños de parcela poco prácticos desde la perspectiva económica o estadística, lo que al final limita su aplicación en el campo. Además, es un procedimiento que sólo ofrece como resultado únicamente el tamaño de la parcela y no sus dimensiones, tal como lo hace el método de máxima curvatura.

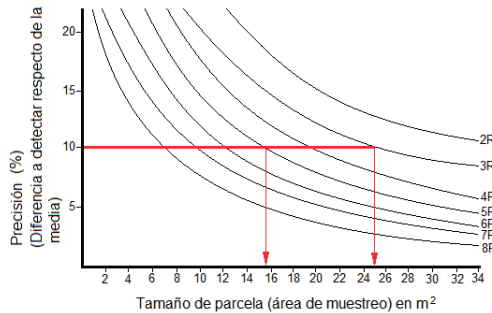


Figura 2.4. Descripción gráfica de la metodología de Hatheway.

La poca información que existe sobre experiencias para determinar las dimensiones de las unidades de muestreo, revela que el método de regresión múltiple es la alternativa más utilizada. Esta información es la que se ha tratado de compendiar en la tabla 2.3, en donde se detallan áreas y dimensiones optimizadas de parcela para algunos cultivos bajo el criterio de mínima variabilidad según varios autores. Son resultados referenciales, dado que la heterogeneidad de los suelos es una variable que no permite generalizar resultados para aplicarlos en experimentos ubicados en contextos edafo-climáticos diferentes.

Tabla 2.3. Dimensiones y tamaño de parcela (unidad de muestreo) para varios experimentos.

Cultivo	Ancho de parcela	Longitud de parcela	Área de parcela	Metodología	Autor (es)
Arroz	---	---	16 a 20 m ²	Máxima curvatura	Vargas J. y Navarro J. (2014)
Arroz	8,0 m	3,0 m	24,0 m ²	Regresión múltiple	Vargas J. y Navarro J. (2019)
Arroz	4,0 m	5,0 m	20 m ²	Máxima curvatura/Regresión múltiple	Gavilánez F., Andrade P. y Monserrate J. (2022)
Fréjol caupí	3,6 m (6 surcos a 0,6 m entre ellos)	5 m (a 0,10 m entre plantas)	18,0 m ²	Regresión múltiple	Escobar C., Buitrago I. y Jaramillo J. (1990)
Fréjol negro	1,2 m (2 surcos a 0,6 m entre ellos)	3,0 m	3,6 m ²	Regresión múltiple	Mamani L. (1971)
Papa	3,6 a 5,4 m (4 a 6 surcos a 0,9 m entre ellos)	3,3 a 4,8 m (a 0,3 m entre plantas)	11,8 a 25,9 m ²	Máxima curvatura bivariada/Regresión múltiple	Hernández A. (1982)
Sorgo	---	---	9,11 m ²	Máxima curvatura	Lechado I. y Rivera J. (1989)

Maíz	4,0 m	5,0 m	20,0 m ²	Regresión múltiple	Escobar C. (1981)
Maíz	---	---	12,0 m ²	Máxima curvatura	Vargas J. y Navarro J. (2017)
Maíz	---	---	4,5 m ²	Máxima curvatura	Avilés F. (1971)
Maíz	4,0 a 5,0 m (4 a 5 surcos)	5 a 8 m	20 a 40 m ²	Máxima curvatura bivariada/Regresión múltiple	Barrientos M. (1981)
Maíz	2,0 m	3,0 m	6,0 m ²	Máxima curvatura/Regresión múltiple	Gavilánez F. y García J. (2022)
Yuca	1,0 a 4,0 m (1 a 4 hileras)	12,0 a 9,6 m (a 0,8 m entre plantas)	12,0 a 38,4 m ²	Regresión múltiple	Tineo J. y Villasmil J. (1988)
Ajonjolí	3,0 a 4,0 m (4 a 5 surcos)	4,0 a 8,0 m	12 a 32 m ²	Máxima curvatura bivariada/Regresión múltiple	Barrientos M. (1981)
Tomate	3,6 a 4,5 m (4 a 5 surcos)	6,0 a 6,5 m	24 a 30 m ²	Máxima curvatura bivariada/Regresión múltiple	Bueso M. (1985)
Melón	5,4 m a 10,8 m (3 a 6 surcos)	5,0 a 7,0 m	27,0 a 75,6 m ²	Máxima curvatura bivariada/Regresión múltiple	Bueso M. (1985)

Fuente: Gavilánez *et al.* (2017) con modificaciones por el autor.

Una vez definido el tamaño y/o las dimensiones de la unidad de muestreo (parcela útil), el investigador debe analizar de manera holística el contexto del experimento a desarrollar para que determine el efecto de borde más adecuado, de tal forma que su ensayo sea económico y preciso. La adición de una, dos o más hileras de plantas en los lados laterales de la unidad de muestreo, así como las correspondientes plantas de cabecera, que (posiblemente) captarán el efecto de borde al contorno de la unidad de muestreo, en su conjunto, es lo que constituye la unidad experimental o también llamada parcela experimental. Si bien debe haber conciencia en el sentido de que el número plantas que formarán el efecto de borde no puede ser excesivo desde el punto de vista de control de la heterogeneidad del suelo, es una decisión que tiene una fuerte dependencia de la forma de aplicación de los tratamientos. Así, por ejemplo, si se trata de experimentos con aplicaciones manuales de productos, con una a tres hileras de plantas y las correspondientes en las cabeceras (tomando en consideración las distancias entre hileras) será suficiente; pero si esta aplicación es mediante equipos como las bombas de motor, la magnitud del efecto de borde necesariamente debe ser más grande.

CAPÍTULO III

TAMAÑO MUESTRAL

La base empírica del investigador es fundamental; ya que sí bien, en un experimento, el hecho de tener múltiples repeticiones disminuye el error experimental, este supuesto beneficio puede llevarnos al extremo de cometer el error tipo I de la estadística, desde la perspectiva de la aplicación práctica de resultados...



Muestra estadística

Uno de los pilares fundamentales de toda investigación es la muestra estadística a partir de la cual se obtiene ya sea una media o una proporción de alguna variable cuantitativa o cualitativa, respectivamente. Para que dicho valor estadístico sea un fiel reflejo de la población desde donde se extraen los datos, la muestra debe ser representativa de dicha población; esto es, la variable medida a través de cada uno de los individuos que conforman la población de estudio (definido como parámetro) no debe diferir sustancialmente con el valor estadístico obtenido en la muestra.

Lo que se ha indicado en el párrafo precedente permite colegir que una muestra es una parte de una población, cuya representatividad hará posible inferir el resultado de la muestra como un resultado que ocurre en la población de origen. La preferencia de optar por la muestra tiene justificativos contundentes. Se recurre al uso de muestras estadísticas en la mayoría de las investigaciones por cuestiones económicas, de tiempo o por la imposibilidad de realizar las medidas correspondientes de una variable en una población.

Estas poblaciones de estudio pueden ser tan pequeñas como los estudiantes de un curso de diseño experimental en una universidad, o inmensamente grande como los gránulos de arena del desierto del Sahara. En este sentido se ha realizado una clasificación de las poblaciones en *finitas* e *infinitas*. Una población es finita cuando todos los individuos que la integran (personas, animales, plantas, insectos o cosas) pueden fácilmente contabilizarse, como el primer ejemplo mencionado al inicio del párrafo; mientras que, una población es infinita, cuando el conteo total de los individuos que la integran es un proceso muy difícil de realizar debido a la inmensurable cantidad de los mismos¹⁴. Por ello se dice que una población infinita es desconocida.

En una población, dos son los aspectos importantes que el investigador no puede dejar de lado para que sus resultados, primero no se aparten de los objetivos del estudio, y segundo, no se realicen gastos innecesarios levantando información adicional que poco o nada contribuye al cumplimiento de dichos objetivos. Estos aspectos son la precisa definición de la unidad de muestreo y la consecuente delimitación de la población de estudio.

Las unidades de muestreo deben estar en concordancia con la problemática y los objetivos de una investigación, pudiendo ser éstas una persona, un animal, una planta, un fruto, un corral, una parcela (parcela útil), una rama de un árbol, una empresa o cualquier elemento sobre el que se realiza una medición de alguna variable. Definir adecuadamente la unidad de muestreo evita obtener una información errónea como, por ejemplo, levantar información sobre el nivel de instrucción académica de los productores de arroz de una localidad y que en su lugar se encueste al jornalero del productor, dejando de lado a la verdadera unidad de muestreo que es el productor.

Una vez que se ha definido plenamente cuál será la unidad de muestreo, el segundo aspecto a establecer son los límites de la población de estudio. Delimitar una población es definir sus características que son de interés para una investigación, de tal forma que las unidades de muestreo no sean seleccionadas de manera errónea. Una población se delimita en su contenido, de forma espacial y de forma temporal. Por ejemplo, digamos que se quiere conocer el tiempo de estudio de las matemáticas de un grupo de estudiantes, cuya característica es que sean del noveno y del décimo año de educación básica de colegios estatales, que sean de la ciudad de Milagro y durante los meses de mayo y junio del 2016. En este ejemplo, existe una delimitación de la población de estudio en cuanto al tipo de colegio (contenido), a la ciudad

¹⁴ En términos prácticos, se define a una población de estudio como infinita cuando está compuesta por más de 10000 elementos.

en donde se realizará la investigación (espacial) y al tiempo durante el cual se medirá la variable (temporal).

La magnitud de una muestra estadística representativa de una población depende de la variación externa a la que están expuestos los datos. No es lo mismo una muestra que se extrae para un estudio de característica descriptiva u observacional, en donde no se tiene ni un mínimo de control sobre el ambiente en donde se mide la variable; que una muestra que se establece a nivel de un experimento, en donde existe un importante control de la variabilidad interna y, por lo tanto, ésta puede ser relativamente pequeña.

Consideraciones de la muestra estadística

En la práctica, los valores estadísticos como las medias se obtienen a partir de una parte de una población que es conocida como la *muestra*. En la definición de esta última pueden o no existir criterios probabilísticos que definen el tipo de muestra, lo que consecuentemente establece el alcance de la inferencia, además de su tamaño. Estos criterios son características de importancia en los tamaños muestrales, de los cuales se tratarán en los siguientes párrafos.

Hay que resaltar la diferencia pertinente que existe entre los ámbitos de aplicación del muestreo. No es lo mismo lo que se pueda obtener de una muestra proveniente de una población en donde no se tiene idea de la variabilidad de la información que se busca, que aquello que es posible conseguir de una muestra en donde se conoce que la variabilidad de la información es reducida y los resultados que se obtendrían de forma aproximada. En el primer caso, estaremos tratando con tipos de *investigaciones netamente descriptivas*, en donde el efecto que se evalúa no ha sido manipulado arbitrariamente y el medio que rodea a la población en estudio no lo controla, ya sea parcial o totalmente, el investigador. En el segundo caso, estamos tratando con *investigaciones básicamente experimentales*, en las cuales la variabilidad desconocida (error experimental) que rodea al efecto de los tratamientos deberá ser prácticamente imperceptible, ya que el medio en donde la población se desarrolla está controlado por el investigador. Es decir, en investigaciones de tipo descriptivas necesitaremos un tamaño de muestra relativamente importante, que permita obtener conclusiones confiables acerca de una población; mientras que, en las investigaciones de carácter experimental, el tamaño de muestra será comparativamente menor. En el contexto de la investigación experimental, el tamaño de muestra se define como repeticiones.

La efectividad y magnitud de una muestra básicamente está subordinada a la representatividad de la misma, a la variabilidad dentro de la población y a los márgenes de error con que la muestra opere respecto de la población de origen. La representatividad de una muestra permitirá, en forma aproximada, que las características de la población no sean obviadas; es decir, lo que estadísticamente pueda concluirse con la muestra sea factible de extrapolación a la población estudiada. Asimismo, si hay razones que indiquen la existencia de una fuerte variabilidad en la información de una población, es lógico que la magnitud de la muestra vaya en aumento para contrarrestar dicha variabilidad (Teorema del límite central). Esto último, también es igual para el margen de error de una muestra, ya que mientras más nos distanciamos de un parámetro poblacional como la media, más grande deberá ser la muestra.

Tipos de muestra

La representatividad de una muestra no es más que la consideración del azar al momento de la medición (recolección) de una variable (característica), a partir de lo cual es posible tener dos tipos generales de muestras:

1. Muestras probabilísticas o aleatorias
2. Muestras no probabilísticas o escogidas

Muestras probabilísticas o aleatorias

Una muestra probabilística o aleatoria es aquella en la que todos los sujetos o unidades de observación de una población tienen la misma probabilidad de ser seleccionados. Son los tipos de muestras que tienen el mayor grado de certeza y, por lo tanto, permiten realizar inferencias respecto de una población con una mayor exactitud. Es el único tipo de muestra que relativamente tiene las características de la población de origen; así, por ejemplo, si una población tiene un 30% de una característica determinada (color, edad, educación, altura, etc.), existe una gran probabilidad de que esta misma proporción de dicha característica esté presente en la muestra. Hay varios tipos de muestreo probabilístico, de entre los cuales, los más utilizados son:

- ✓ **Muestreo aleatorio simple:** es un tipo de muestreo recomendado (no limitado) para poblaciones pequeñas o finitas. Básicamente consiste en identificar a cada una de las unidades de medida (numéricamente, mediante letras o cualquier otro símbolo) de la población, de tal manera que puedan someterse a una selección aleatoria mediante algún tipo de sorteo (tómbolas, tablas de números aleatorios, etc.).
- ✓ **Muestreo sistemático:** este tipo de muestreo también está indicado para poblaciones pequeñas y consiste en la selección de las unidades de observación mediante el uso de intervalos. Para realizar este tipo de muestreo primero se parte seleccionando al azar una primera unidad de observación, luego, a partir de ésta, el resto se seleccionan con un intervalo fijo. Por ejemplo, si asumimos un intervalo de 5, cada quinta unidad deberá ser seleccionada. Debe tenerse la debida precaución respecto de las unidades de medida seleccionadas de esta forma, ya que sólo la primera de éstas es seleccionada al azar, teniendo el resto una probabilidad relativa. Tiene cierta deficiencia respecto del muestreo aleatorio simple, ya que si el intervalo escogido coincide con alguna alteración de la población, esto se reflejará en la muestra y, en consecuencia, la información que se extraiga estará sesgada.
- ✓ **Muestreo estratificado:** es el tipo de muestreo más idóneo para poblaciones grandes o infinitas que poseen ciertas características heterogéneas, las cuales sugieren que dicha población se divida en estratos o subpoblaciones de mayor homogeneidad, de tal forma que la muestra constituida considere estas características y sea representativa de toda la población en estudio al momento de medir una variable. En forma concreta, lo que se pretende con este tipo de muestreo es asegurarse de que todos los estratos de importancia de una población estarán representados adecuadamente en la muestra. Bajo esta consideración, los estratos obedecen a características consideradas de importancia para un estudio (tales como edad, sexo, tiempo, etc.) y dentro de cada uno de ellos, en el proceso de muestreo, las unidades de observación son seleccionadas aleatoriamente.
- ✓ **Muestreo aleatorio por conglomerados:** en este tipo de muestreo las unidades de observación son escogidas desde grupos de elementos de la población que se forman

previamente y constituyen unidades muestrales a las cuales se las conoce como *conglomerados*. Dado que comúnmente se aplica sobre áreas geográficas extensas, se entiende su uso cuando el número de elementos de la población es muy grande y muy disperso, de allí que no se pueda estratificar la información. Para la toma de la muestra, la población se divide en conglomerados o subpoblaciones, dentro de los cuales se seleccionan las unidades de observación. Este tipo de muestreo puede realizarse en una, en dos o en varias etapas¹⁵; clasificándose, por tanto, entre *monoetápico*, *bietápico* y *polietápico*, respectivamente. Ejemplos de conglomerados pueden ser las universidades de un país, los hospitales, productos empacados, urnas de votación, etc. Es importante que los grupos sean heterogéneos, mutuamente excluyentes (sus unidades de observación) y exhaustivos.

Muestras no probabilísticas o escogidas

Muchas veces en los estudios de tipo descriptivo o en los exploratorios, el muestreo probabilístico como tal resulta excesivamente costoso porque los tamaños de muestra son relativamente grandes, especialmente cuando la dispersión es fuerte y/o se miden variables cuantitativas a nivel de laboratorio (analíticas). Para evitar estas limitaciones se acude a muestreos no probabilísticos que desafortunadamente tienen la desventaja de que no sirven para realizar generalizaciones (inferencia poblacional), ya que es muy difícil que la muestra extraída sea representativa debido a que no todos los sujetos de la población tienen la misma probabilidad de ser elegidos. Por esta particularidad que tiene este tipo de muestreo, el tamaño de la muestra generalmente es establecido bajo el criterio subjetivo del investigador y las conclusiones que se extraen sólo se circunscriben a reducidos espacios de una población objetivo, dentro de los cuales se han medido determinadas características. Entre los tipos de muestreo no probabilístico se tienen:

- ✓ **Muestreo intencional o de conveniencia:** es un tipo de muestreo que selecciona a las unidades de observación de acuerdo a su disponibilidad o por la representatividad de un grupo de individuos dentro de la población. Como ejemplos se pueden mencionar al muestreo de plantas de lotes de cultivo en predios de agricultores conocidos por el investigador o encuestas a personas con las cuales se tiene fácil acceso. La extrapolación a la población es específica para la condición evaluada la cual debe ser especificada.
- ✓ **Muestreo por cuotas:** viene a constituir una analogía al muestreo estratificado, pero se diferencia de éste por el hecho de que la selección de los datos no se hace en forma aleatoria. Es un tipo de muestra que, de forma predeterminada, se establecen grupos con ciertas características de interés para el investigador. Es una de las opciones de muestreo muy utilizada en encuestas de opinión.
- ✓ **Muestreo bola de nieve:** este tipo de muestreo ocurre cuando se utilizan unos pocos sujetos (unidades de medida), los cuales, a su vez, brindan información sobre otros sujetos hasta constituir una muestra relativamente adecuada. Es aplicable cuando la

¹⁵ El muestreo por conglomerado en una etapa (monoetápico) es aquel en donde sólo la selección de los grupos o conglomerados se hace de forma aleatoria, ya que, una vez seleccionados estos grupos, en el muestreo entran todas las unidades de observación que estos contengan. En el caso del muestreo por conglomerado bietápico se realizan dos sorteos: primero se seleccionan aleatoriamente los grupos y, subsecuentemente, también se hace una selección aleatoria de las unidades de observación dentro de los grupos escogidos aplicando el muestreo aleatorio simple (o también el muestreo sistemático). En el muestreo por conglomerados polietápico se realizan más de dos selecciones aleatorias y está recomendado para extensas poblaciones con mucha heterogeneidad, por lo cual se forman grupos, subgrupos, minigrupos, etc., definiendo al final los conglomerados desde donde se seleccionarán las unidades de observación también de forma aleatoria. Puede presentar una ventaja económica en cuanto al muestreo aleatorio simple, o sobre el estratificado, dado que se recurre a grupos concentrados en una misma unidad de muestreo, ahorrando costos de ubicación (movilidad) de las unidades de observación.

información no es tan común (evidente) o es de difícil acceso; ejemplos pueden ser los muestreos de enfermos de sida, de ciertos grupos delincuenciales, de plantas medicinales especiales, entre otros.

Tamaño de la muestra

En forma general y de acuerdo a lo mencionado al inicio de este capítulo, básicamente en las investigaciones se pueden obtener dos tipos de tamaños de muestra. Uno que está referido a investigaciones de carácter netamente observacional, en donde se miden variables cualitativas o cuantitativas y que, por lo general, está conformado por un número elevado de unidades de observación debido a que el investigador no tiene control de la varianza interna entre dichas unidades. El otro tipo de tamaño muestral es el que se realiza en investigaciones experimentales, generalmente compuesto de pocas unidades de observación, conocidas éstas en el contexto, como *unidades experimentales*. En este último caso, la magnitud del efecto que se valora se estima de antemano y el investigador puede tener un control parcial o total de la varianza interna, de allí que se tengan tamaños reducidos.

No obstante, una limitación muy importante que se presenta al momento de determinar el tamaño muestral en cualquier situación, es la necesidad del conocimiento previo de la varianza de la información (varianza no explicada o error experimental en el caso de investigaciones experimentales), por lo que se recurre a pruebas piloto o a investigaciones anteriores de las cuales se pueden obtener valores aproximados de este parámetro.

Factores que determinan el tamaño de una muestra

Bajo el supuesto de que la muestra sea la adecuada en cuanto a la representatividad de una población (definida bajo muestreo probabilístico), el tamaño ideal que contenga las características de importancia de esta última, de manera que se puedan extrapolar los resultados de la muestra a dicha población, depende fundamentalmente de tres factores: el *nivel de confianza*, el *error máximo admisible* y la *varianza*. La definición del valor de los dos primeros de estos términos depende del criterio subjetivo del investigador, el cual, en función de su experiencia, establecerá el más adecuado para su investigación; mientras que la varianza está establecida según la variabilidad de la población de estudio que es estimada de antemano.

En esta parte, es importante mencionar que el tamaño de una muestra dependerá de que la población tenga característica *finita* o *infinita*. En el caso de que la población esté definida como *finita*, es decir, la cantidad total de unidades de observación es cuantificable y, en ciertas ocasiones, fácilmente accesibles por su reducido tamaño, el tamaño de muestra se verá seriamente afectado por el tamaño de la población hasta cierto límite de esta última. Si la población es infinita no existe influencia de ésta en el tamaño de la muestra.

La relación entre el tamaño de una muestra (considerando constantes el nivel de confianza, la varianza y el error máximo admisible) y el tamaño de una población finita se describe gráficamente mediante la figura 3.1. Puede observarse que la relación no es proporcional y más bien hay un efecto asintótico en el cual se hace evidente el hecho de que a partir de cierto tamaño poblacional, el tamaño muestral toma un valor prácticamente constante. En un proceso de simulación, este tamaño de población en donde la muestra deja de diferir efectivamente para hacerse constante, se produce a partir de las 10000 unidades de observación. Por ello es que, a una población infinita, si bien se la ha definido como aquella cuyo tamaño se desconoce, también se la ha considerado como tal cuando se tienen más de 100.000 elementos de observación. A manera de conclusión, podríamos mencionar que el resultado práctico del

proceso de cálculo del tamaño de una muestra, teniendo población finita o infinita, es el mismo cuando una población tiene más de 10.000 unidades de observación.

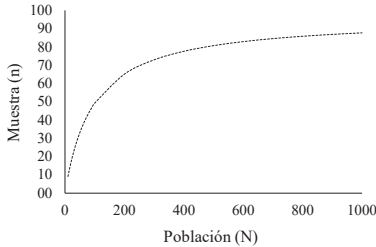


Figura 3.1. Esquema sobre la variabilidad del tamaño de una muestra n respecto del tamaño de una población N conocida.

La relación entre el error máximo admisible y el tamaño de una muestra no es directamente proporcional como pudiera colegirse. Es así como, para un error máximo admisible de 10%, manteniendo constantes tanto el nivel de confianza (95%) como la varianza (25% del parámetro), se requiere un tamaño de muestra de 95 unidades de observación; pero si dicho error es más conservador, por ejemplo 5%, el tamaño de muestra cambia a 370 unidades de observación. A nivel experimental esta relación es un tanto más factual, dado que no sirve de mucho incrementar el número de repeticiones, pues el error (que en ese contexto se llama error experimental) no se reduce sustantivamente si dichas repeticiones se incrementan desde cierta cantidad.

Relacionado a las muestras y para la aplicabilidad de las pruebas de la estadística paramétrica, un aspecto fundamental es el *teorema de límite central* (TLC). Bajo este teorema, las medias que se obtengan de muestras grandes siguen una *distribución normal*, esta última, fundamental en la aplicación de los procesos estadísticos inferenciales. El cumplimiento de este teorema no depende del tipo de distribución que sigan los datos, salvo que el tamaño de la muestra sea “pequeño”. En este sentido, las muestras estadísticas se han clasificado como muestras *grandes* aquellas que tienen más de 30 datos ($n > 30$) y *pequeñas* las que tiene 30 datos o menos ($n \leq 30$).

El procedimiento de cálculo del tamaño de una muestra depende del parámetro estadístico que se pretende estimar; por ello, en forma general, existen dos procesos. Uno de ellos se utiliza para estimar una media; siendo, por lo tanto, la variable de tipo cuantitativa-continua. El segundo proceso de cálculo muestral va a permitir estimar un valor porcentual o proporción, en donde la variable a medirse es discreta o categórica (cualitativa).

Aparte de los factores estadísticos indicados y dependiendo de lo que se vaya a investigar, según se ha venido mencionando, el factor determinante en la mayoría de estudios es el de tipo económico. Esta situación es más radical en investigaciones en las cuales se busca estimar promedios a base de distintos tipos de metodologías de laboratorio que demandan una importante inversión económica, lo que consecuentemente ocasiona que el tamaño de muestra básicamente esté subordinado al presupuesto de la investigación.

Tamaño de muestra de poblaciones infinitas para estimar una media

Una vez que se ha elegido el tipo de muestreo más idóneo para una población infinita con características determinadas, el tamaño de muestra (n) se establece a partir de la consideración de que la distribución de todas las medias posibles se ajusta normalmente. Por lo tanto, siempre existirá una diferencia entre la media de una muestra (\bar{x}) y la media de su población correspondiente (μ) dada por la expresión 3.1. Esta diferencia entre \bar{x} y μ es lo que se conoce como error máximo admisible (E), también llamado *error máximo de estimación* o *error de precisión*¹⁶ (también llamado *error de muestreo*), y representa a la distancia que hay entre \bar{x} y μ que se observa en la figura 3.2.

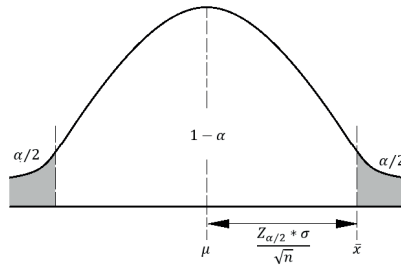


Figura 3.2. Error máximo admisible o distancia entre la media muestral y la media de su población infinita correspondiente.

Considerando este error máximo admisible, el cual podría ubicar a la media muestral por encima o por debajo de la media poblacional, está dado por:

$$E = Z_{\alpha/2} \frac{S}{\sqrt{n}} \quad (3.1)$$

El tamaño de una muestra (n) para poblaciones infinitas se calcula por:

$$n = \left(Z_{\alpha/2} \frac{S}{E} \right)^2 \quad (3.2)$$

Aunque parcialmente se ha indicado el significado de cada uno de los símbolos de la expresión 3.2, es preciso ampliar sus conceptualizaciones, lo cual se detalla en los siguientes párrafos.

En las expresiones indicadas (3.1 y 3.2), n define al tamaño de la muestra o número de unidades de observación que se deben seleccionar en la población de estudio.

El parámetro $Z_{\alpha/2}$ representa al grado o nivel de confianza que se tiene de que la muestra extraída contenga al parámetro estadístico que se pretende estimar. Corresponde al número de desviaciones estándar respecto de la media de una población cuya distribución muestral se supone sigue de una distribución normal, y dado que es un proceso general de conversión de valores a escala Z para cualquier tipo de variable continua, es común utilizar el término de *distribución normal estándar*.

¹⁶ Se prefiere tomarlo como error de precisión cuando E es calculado a partir de la información ya generada. No obstante, cuando es un valor asumido, es mejor llamarlo como error máximo admisible.

El nivel de confianza representado por Z es complementario a la probabilidad de error de que, la muestra extraída, no contenga al parámetro de estimación (p. e. una media o una proporción). Este error es de uso común en las pruebas de hipótesis, es representado por α , es conocido como *error tipo I* y convencionalmente toma el valor de 5% como nivel de significancia para estas pruebas. Por lo tanto, dado que α es de selección automática, la confianza establecida por $1 - \alpha$ es definida por decisión del investigador.

El valor de α es dividido por 2 ($\alpha/2$), según se indica como subíndice de Z , debido a que el error de máximo admisible E , en una distribución normal, tiene igual probabilidad de afectar a la media (\bar{x}) tanto en forma decreciente como en forma creciente ($\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$). En consecuencia, la media poblacional μ generalmente se la estima mediante un intervalo que es conocido, de acuerdo al valor de Z , como *intervalo de confianza*. Un nivel de significancia $\alpha/2$ dentro de las pruebas de hipótesis, originan las llamadas *pruebas bilaterales* o de *dos colas* porque la probabilidad de error se distribuye equitativamente a ambos lados (extremos) del área de curva normal estándar.

El valor de $Z_{\alpha/2}$ generalmente utilizado es el correspondiente a un nivel de confianza del 95% ($\alpha = 5\%$) y tiene una distancia estandarizada de 1,960. En ciertos casos, para tener un mayor grado de confianza, como por ejemplo del 99% ($\alpha = 1\%$), se utiliza el valor correspondiente de 2,575. Este último valor consecuentemente ocasiona que la muestra sea más grande (véase la fórmula 3.2).

El parámetro S es la desviación estándar obtenida, ya sea en una prueba piloto (preliminar) o por trabajos relacionados realizados previamente. Este valor es un estimador del parámetro poblacional σ , que es el que representa a la desviación estándar poblacional, el cual en poblaciones infinitas es imposible que se conozca; de allí el empleo de su estimador en la fórmula 3.2. Para estimar preliminarmente el valor de S en una población puede empezarse seleccionando una muestra inicial (n_0) de 25 unidades de observación; una vez establecido este valor, se lo emplea en expresiones como la 3.2 o cualquier otra fórmula para establecer el tamaño de la muestra definitiva, para a partir de allí, completar el resto de unidades de observación que faltarían.

Como ya se ha mencionado, E representa al error máximo admisible, el cual tiene las mismas unidades que la desviación estándar (por esto también se lo identifica como diferencia absoluta). Es la separación estadística que se estima existe entre la media poblacional μ y la media muestral \bar{x} al momento de extrapolar los resultados a la población original. En la práctica, el valor que le asignemos a E va a depender de cuánto error en la estimación de un parámetro (p. e. media o proporción) estamos dispuestos a aceptar en el tamaño de la muestra y, al igual que el grado de confianza, está fuertemente ligado a la disposición económica para llevar a cabo un estudio. El valor que se recomienda es que E no sea mayor a un 10% del parámetro que se quiere estimar; es decir, si una media es de 12,5 g, el error máximo de estimación deberá ser 1,25 g; o si una proporción es de 50%, este error será de 5%. Como es común no tener a la mano la media o la proporción que se quiere estimar, el valor de E también puede deducirse a través de los resultados de evaluaciones preliminares.

Una información importante al momento de reportar una media es conocer su desvío a una desviación estándar, término que se conoce como *error típico* o *estándar de la media* ($S_{\bar{x}}$), cuyo valor se establece por $S_{\bar{x}} = \frac{\sqrt{S^2/n}}{n} = S/\sqrt{n}$. Para una proporción (S_p), este error típico se define como $S_p = \sqrt{p(1-p)/n}$.

Tamaño de la muestra de poblaciones infinitas para estimar un valor porcentual (proporción)

Cuando la (o las) variable es de característica cualitativa no se tiene una media que se pueda cuantificar, en este caso se deberán establecer proporciones o porcentajes de individuos de la población. En estas investigaciones se tienen variables con un mínimo de dos respuestas (respuestas dicotómicas de sí o no, por ejemplo), típicas en investigaciones de carácter social, en las que se trabaja a través de encuestas y, por lo tanto, estas variables de medición son de tipo cualitativas. Un ejemplo podría ser la medición del acceso a internet en una población en donde habrá personas que sí tendrán acceso y otras que no, he allí una variable dicotómica.

En variables cualitativas de dos o más respuestas, se tendrá una opción que el investigador tomará de referencia como acierto y que en la población tendrá una proporción definida, de allí que si se toma como p a la proporción de éxito o el porcentaje de individuos que dan respuestas positivas a la opción de referencia y $1 - p$ (a veces a este término se los representa por q) a la proporción de falla o porcentaje de individuos que dan respuestas negativas a dicha opción de referencia, se podrá definir el *error estándar de la media* antes indicado y que, en este caso, se denota por $\sqrt{p(1 - p)/n}$. Al remplazar este último término en la expresión 3.1, el error máximo admisible se obtiene por:

$$E = Z_{\alpha/2} \sqrt{\frac{p(1 - p)}{n}} \quad (3.3)$$

Haciendo el despeje respectivo, el tamaño de la población estará dado por la siguiente expresión:

$$n = Z_{\alpha/2}^2 \frac{p(1 - p)}{E^2} \quad (3.4)$$

En la expresión 3.4, n , $Z_{\alpha/2}$ y E tienen las mismas definiciones antes indicadas. En el caso del parámetro p , éste representa el porcentaje de éxito o proporción de individuos que responderán positivamente al momento de medir una variable cualitativa, expresado en forma fraccionaria. Generalmente este valor es desconocido al momento de establecer el tamaño de la muestra. Si no se conoce de antemano, se recomienda utilizar $p = 0,5$, que es cuando se tiene la mayor varianza (condición crítica), y por lo tanto, no se correrá el riesgo de que el tamaño muestral sea insuficiente. De forma similar como en el proceso de cálculo del tamaño de muestra para una media, una alternativa viable para obtener un valor de p inicial es la realización de muestreos preliminares. En la agronomía, algunas variables que se expresan de forma proporcional están relacionadas al ataque de plagas; así, por ejemplo, la incidencia de una enfermedad se expresa en forma porcentual.

En forma general, como puede observarse en las expresiones 3.2 y 3.4, el tamaño de la muestra tiene una relación casi directa con la confianza que estemos dispuestos a aceptar y una relación inversa, pero no directamente proporcional, respecto del error máximo admisible. Esto es lo que se observa en la tabla 3.1. Por ejemplo, con un error máximo admisible del 5% y con una confianza del 95% ($Z_{0,05/2} = 1,960$), el tamaño de la muestra es de 384; si queremos para este mismo error una confianza del 99% ($Z_{0,01/2} = 2,575$), el tamaño de la muestra sube a 663. Si mantenemos la misma confianza, digamos 95%, con un error máximo admisible del 1%, el tamaño muestral es de 9604; mientras que, con un error máximo admisible del 2%, esta muestra se reduce hasta el valor de 2401.

Tabla 3.1. Tamaños de muestras para dos niveles de confianza y distintos errores admisibles considerando poblaciones infinitas.

Error máximo admisible (E)	$\alpha = 0,05$ $Z_{\alpha/2} = 1,960$	$\alpha = 0,01$ $Z_{\alpha/2} = 2,575$
0,05	384	663
0,04	600	1036
0,03	1067	1842
0,02	2401	4144
0,01	9604	16577

Tamaño de la muestra de poblaciones finitas

Cuando la población es pequeña, por ejemplo, que esté integrada por 100 individuos (en el caso de estudios sociales: un pequeño pueblo), utilizando la expresión 3.4 vamos a obtener valores salidos de contexto, como que el tamaño de nuestra muestra sea de 384 individuos (véase la tabla 3.1). Asimismo, tal como está planteada la expresión 3.4, si de 100 individuos de una población se tomaran 50 como muestra, no es lo mismo tener 100 individuos de una muestra representando a 10.000 que integran una población. En el primer caso estamos con el 50% de la población y en el segundo con apenas el 1% de la población. Es decir, en el primer caso tendremos una muestra con mayor representatividad que en el segundo caso. Para evitar estas situaciones extremas al momento de establecer el tamaño de una muestra perteneciente a una población finita (población de la que se conoce el total de unidades de observación), hay que utilizar un factor de corrección aplicado al error estándar de la media en la expresión establecida para el error máximo admisible. Es decir:

Factor de corrección:

$$\frac{N - n}{N - 1} \quad (3.5)$$

Realizando los despejes respectivos tenemos:

$$\frac{E}{Z_{\alpha/2}} = \frac{S}{\sqrt{n}} \quad (3.6)$$

Elevando al cuadrado la ecuación 3.6:

$$\frac{E^2}{Z_{\alpha/2}^2} = \frac{S^2}{n} \quad (3.7)$$

Aplicando el factor de corrección en la expresión 3.7:

$$\frac{E^2}{Z_{\alpha/2}^2} = \frac{S^2}{n} \left[\frac{N - n}{N - 1} \right] \quad (3.8)$$

Realizando el despeje de n :

$$n = \frac{Z_{\alpha/2}^2 S^2 N}{E^2 (N - 1) + Z_{\alpha/2}^2 S^2} \quad (3.9)$$

La fórmula 3.9 es la expresión para calcular el tamaño de una muestra de una población finita o conocida, en donde N representa al total de unidades de observación o de elementos que

conforman la población objetivo. El resto de símbolos tienen el mismo significado que ya fueron indicados para las expresiones en el caso de población infinita.

Cuando se trate de tamaños de muestra para estimar proporciones o porcentajes de poblaciones finitas (estudios sociales, principalmente), reemplazando el valor de la varianza: $S^2 = p(1 - p)$ en la expresión 3.9, se tiene:

$$n = \frac{Z_{\alpha/2}^2 p(1 - p)N}{E^2(N - 1) + Z_{\alpha/2}^2 p(1 - p)} \quad (3.10)$$

Las fórmulas para establecer el tamaño de muestra de poblaciones finitas tienen sus límites de operatividad, especialmente cuando se estima porcentajes en la medición de variables cualitativas, cuando las poblaciones desde donde se extraen son sustantivamente pequeñas ($N \leq 30$) o cuando éstas son relativamente grandes; siendo más incisivo el primer caso dado que se pierde la seguridad del tipo de distribución que siguen los datos, según el postulado establecido por el teorema del límite central. A este respecto, de acuerdo a lo indicado en la tabla 3.2, en donde se han calculado tamaños muestrales para proporciones con varianzas de 0,25 [0,5(1 - 0,5)], considerando diversas poblaciones y niveles de error más importantes, se pueden deducir las siguientes conclusiones:

- Cuando la población es relativamente grande, de unos 10.000 individuos o unidades de muestreo, es indiferente usar las expresiones para el cálculo de muestras de poblaciones finitas o infinitas (Figura 3.1). Por lo tanto, por su sencillez, es más práctico utilizar la expresión que está referida para poblaciones infinitas.
- Cuando la población es muy pequeña y si además se selecciona un error máximo admisible muy reducido, prácticamente la población constituye la muestra (censo).
- Si la población está compuesta con menos de 20 unidades de observación o de individuos, la muestra incluirá a toda la población. De 20 a 30 unidades de muestreo, la muestra sencillamente es equivalente a $N - 1$.

Como es deducible, el tamaño de una muestra no tiene una única solución, pues tiene una fuerte base intuitiva y subjetiva que reside en la selección del nivel de confianza y del error máximo admisible, además del aspecto económico; todo lo cual vuelve fundamental el carácter técnico-científico que el investigador debe hacer confrontar con todos estos aspectos mencionados. Dentro de estos elementos también se encuentra la heterogeneidad de la población de estudio, especialmente cuando éstos son de tipo observacional, haciendo que los tamaños de muestra sean más grandes mientras más grandes sean las varianzas.

Tabla 3.2. Tamaños de muestra de poblaciones finitas para estimar proporciones considerando una varianza de [0,5(1 - 0,5)].

Población	Niveles de error muestral para una confianza del 95%.				
	0,05	0,04	0,03	0,02	0,01
N = 5	5	5	5	5	5
N = 10	10	10	10	10	10
N = 15	14	15	15	15	15
N = 20	19	19	20	20	20
N = 25	24	24	24	25	25
N = 30	28	29	29	30	30
N = 35	32	33	34	35	35
N = 40	36	38	39	39	40
N = 45	40	42	43	44	45
N = 50	44	46	48	49	50
N = 60	52	55	57	59	60
N = 70	59	63	66	68	70

N = 80	66	71	74	77	79
N = 90	73	78	83	87	89
N = 100	80	86	92	96	99
N = 200	132	150	169	185	196
N = 300	169	200	234	267	291
N = 400	196	240	291	343	384
N = 500	217	273	341	414	475
N = 1000	278	375	516	706	906
N = 2000	322	462	696	1091	1655
N = 3000	341	500	787	1334	2286
N = 4000	351	522	843	1501	2824
N = 5000	357	536	880	1622	3288
N = 10000	370	566	964	1936	4899
N = 100000	383	597	1056	2345	8763
Tamaño de muestra en población infinita	384	600	1067	2401	9604

Otro aspecto que también debe analizarse al momento de establecer el tamaño muestral es el tipo de tratamiento estadístico que se realizará con las variables. Si éste es univariado, no existe mucha exigencia de importancia; sin embargo, para análisis bi-variados como los de correlación, al menos deberá contarse con 30 datos. Esta exigencia del tamaño muestral es mucho más fuerte en el caso de estudios multivariados como, por ejemplo, un análisis de regresión múltiple.

Ejemplo de aplicación 3.1

Ejercicio de aplicación sobre muestreo aleatorio simple para estimar una media

En una ciudad se ha decidido investigar la presencia de aluminio en el agua potable. La ciudad cuenta con aproximadamente 86.000 unidades de vivienda. En una prueba piloto se pudo establecer una desviación estándar de 0,082 mg/L de aluminio. Si se considera como error de precisión el valor de 0,015 mg/L, establecer el tamaño de la muestra con un 95% de confianza de ubicación del parámetro.

Solución:

Si se considera una población infinita, el tamaño de la muestra se podrá establecer con la expresión 3.2:

$$n = \left(Z_{\alpha/2} \frac{S}{E} \right)^2$$

Donde:

$$Z_{0,05/2} = 1,96$$

$$S = 0,082 \text{ mg/L}$$

$$E = 0,015 \text{ mg/L}$$

Por lo tanto, n tiene un valor de:

$$n = \left(\frac{1,96 \times 0,082}{0,015} \right)^2$$

$$n = 114,8 \approx 115$$

Si este tamaño de muestra se determina mediante la expresión 3.9, considerando el criterio de población finita, se tiene:

$$n = \frac{Z_{\alpha/2}^2 S^2 N}{E^2(N-1) + Z_{\alpha/2}^2 S^2}$$

$$n = \frac{1,96^2 \times 0,082^2 \times 85000}{0,015^2(85000 - 1) + 1,96^2 \times 0,082^2} = 114,7 \approx 115$$

Nótese la indiferencia en los resultados con las expresiones 3.2 y 3.9 cuando la población está por arriba de los 10.000 elementos.

Ejemplo de aplicación 3.2

Ejercicio de aplicación sobre muestreo aleatorio estratificado para estimar una proporción

Se requiere realizar un estudio socioeconómico de 345 familias de agricultores productores de maíz, ubicados en una amplia zona agrícola a lo largo de una cuenca hidrográfica. De este total, 89 están ubicados en la parte alta de la cuenca, 148 se ubican en la parte media de la cuenca y 108 están ubicados en la parte baja de dicha cuenca. ¿Cuál deberá ser el tamaño de la muestra?

Solución:

En este caso, el muestreo deberá ser estratificado. Además, la población se considera como finita. Sin embargo, dado que en poblaciones estratificadas es esperable que se tengan dispersiones muy dispares, será necesario considerar las varianzas de cada estrato (conocidas como *cuasi-varianzas*) S_i^2 si lo que se pretende es obtener un tamaño de muestra que represente a la población de estudio considerando la variabilidad de dichos estratos. En esta situación, la expresión para obtener una media, manteniendo fijo el error de muestreo E , es la siguiente:

$$n = \frac{(Z_{\alpha/2} N)^2 \sum N_i^2 S_i^2}{\sum N^2 W_i (E^2 N^2 + Z_{\alpha/2}^2 \sum N_i S_i^2)} \quad (3.11)$$

Siendo N el tamaño de la población, N_i el tamaño de la población de cada estrato, S_i^2 la varianza de cada estrato (cuasi-varianza) y W_i la proporción poblacional de cada estrato ($W_i = N_i/N$).

Para aplicar la expresión 3.11 será necesario detallar lo siguiente:

Estratos	N_i	W_i	N_i^2	S_i^2	$N_i^2 S_i^2$	$N^2 W_i$	$N_i S_i^2$
Cuenca alta	89	0,258	7921	0,25	1980,3	30705	22,3
Cuenca media	148	0,429	21904	0,25	5476,0	51060	37,0
Cuenca baja	108	0,313	11664	0,25	2916,0	37260	27,0
	345	1,000			10372,3	119025	86,3

Al final, considerando un valor de 0,05 como error de muestreo y una confianza del 95% dado por el valor de $Z_{0,025} = 1,96$, se tiene:

$$n = \frac{(1,96 \times 345)^2 \times 10372,3}{119025 \times (0,05^2 \times 345^2 + 1,96^2 \times 86,3)} = 63,3 \approx 64$$

Luego, el tamaño de la muestra es de 64 agricultores repartidos de la forma proporcional siguiente:

Estratos	W_i	n_i
Cuenca alta	0,258	17
Cuenca media	0,429	27
Cuenca baja	0,313	20
	1,0	64

La expresión 3.11 está establecida para definir un tamaño muestral que estime una media, sin embargo, en este ejercicio el parámetro de estimación es una proporción; de allí que se haya considerado como varianza $S^2 = p(1 - p) = 0,25$. En este caso se ha considerado el valor crítico para estimar p (de la población), esto es: $\hat{p} = 0,5$.

Es oportuno indicar que la expresión 3.11 no está diseñada para aplicarse en poblaciones pequeñas, en las cuales un estrato determinado (N_i) sea menor de 30. Para utilizar la expresión 3.11 en tamaños de poblaciones parciales de estratos < 30 , esta fórmula se debe modificar a la forma siguiente:

$$n = \frac{\frac{\sum N_i^3 p_i (1 - p_i)}{\sum N^2 W_i (N_i - 1)}}{\frac{E^2}{Z_{\alpha/2}^2} + \frac{\sum N_i^2 p_i (1 - p_i)}{\sum N^2 (N_i - 1)}} \quad (3.12)$$

En la cual p_i es la proporción parcial de aciertos de cada estrato.

Ejemplo de aplicación 3.3

Ejercicio de aplicación sobre muestreo aleatorio simple para estimar una proporción

Una fábrica pretende comercializar un nuevo producto de consumo, por lo cual necesita realizar una encuesta para validar su aceptación de tal forma que su difusión en el mercado esté asegurada. La población es mayor a los 10000 individuos y, por tanto, puede ser considerada como infinita. Si se requiere que la aceptación sea del 70% o más, ¿cuál será el mínimo tamaño muestral que deberá considerar en su encuesta?

Solución:

Para este cálculo se requiere la expresión 3.4:

$$n = Z_{\alpha/2}^2 \frac{p(1 - p)}{E^2}$$

Donde:

$$Z_{0,05/2} = 1,96$$

$$p = 0,7$$

$$E = 0,05$$

En consecuencia, el valor de n es:

$$n = \frac{1,96^2 \times 0,7 \times (1 - 0,7)}{0,05^2} = 322,7 \approx 323$$

Tamaño de muestra para experimentos

En los experimentos, la definición del tamaño muestral, que viene a constituirse en el número de repeticiones de cada tratamiento, tiene una relación determinante con la precisión de un ensayo; lo que a su vez le confiere “sensibilidad” para que un experimento detecte las diferencias entre los tratamientos de prueba que el investigador propone preliminarmente en su hipótesis. En este sentido, es necesario recordar los dos tipos de errores a los que está sujeta una decisión luego de realizado un experimento: el primero conocido como *error tipo I*, se hace presente cuando se rechaza una hipótesis nula que en la realidad es verdadera; y el segundo, identificado *error tipo II*, que ocurre cuando se acepta una hipótesis nula que en la realidad es falsa. Si bien el primero es aquel que reviste un mayor impacto, dado que involucra el cambio hacia una nueva alternativa (tratamiento); el segundo es el que está relacionado con la precisión γ , por tanto, con la capacidad de hacer detectables los efectos de los tratamientos.

En un experimento, una hipótesis nula (H_0) generalmente hace referencia a la ocurrencia de un fenómeno (cumplimiento de un parámetro de la población) o a la ausencia (o igualdad) de efectos, cuando se prueba un solo tratamiento o cuando se contrastan más de un tratamiento, respectivamente¹⁷. Cuando se rechaza la hipótesis nula se plantea una situación experimental alterna y sustantiva, conocida como hipótesis alternativa representada por H_a . La hipótesis alternativa puede proponerse indicando que el efecto de un tratamiento es menor, mayor o simplemente no es igual a una condición poblacional de un parámetro estadístico (media o proporción), o al efecto de otro u otros tratamientos que se contrastan en un experimento. Ejemplos de planteamientos hipotéticos son los siguientes:

- Un vendedor de semillas le asegura a un agricultor que tiene un híbrido de maíz que es capaz de rendir 8,0 t/ha. Este agricultor conoce, por amigos, que dicho híbrido no es capaz de llegar a tal rendimiento, aun aplicando toda la metodología de cultivo necesaria. En este sentido, previo a la compra, deciden ambos realizar un experimento de prueba, sembrando el material para poder evaluar el rendimiento y establecer si lo

¹⁷ Dos aspectos debemos resaltar en esta explicación. El primero tiene que ver con el hecho de considerar que un estudio sea experimental cuando se evalúa una sola muestra. Para este primer caso debemos recordar qué define a un suceso como experimental; el cual, en términos sustantivos, es establecer deliberadamente (o crear) las condiciones que en el contexto se definen como variables independientes (causas), para posteriormente verificar los efectos en el fenómeno evaluado a través de variables consecuentes o dependientes. Con este preámbulo conceptual de experimento, se puede decir que una investigación en la cual esté inmersa la toma de una sola muestra puede ser definida como experimental si se modifican las condiciones bajo las cuales se toma esa única muestra, que en este último caso puede llamarse tratamiento. Por ejemplo, la caracterización de una única semilla se constituiría en una investigación descriptiva si sólo se evaluaran sus caracteres con el propósito de conocer sus índices de producción; sin embargo, adicional a esto, puede dársele la definición de experimental cuando queremos valorar esos mismos índices bajo una condición agroclimática diferente.

El segundo aspecto está relacionado al formato de planteamiento de las hipótesis nulas. Si bien siempre se ha hecho mención que las hipótesis nulas se planteen considerando efectos neutrales, colocando el signo de igualdad con respecto a un parámetro o entre tratamientos, con lo que implícitamente se define que una prueba sea bilateral (o dos colas), es importante tener en cuenta que este proceso es relativo al contexto de lo que se pretende evaluar y que se explicita en la hipótesis alternativa. Es decir, si una hipótesis alternativa se plantea considerando que un efecto es superior o inferior a un parámetro u otro tratamiento, la hipótesis nula debe ser complementaria a este planteamiento. Por ejemplo, si la hipótesis alternativa se establece como $H_a > \mu$, la hipótesis nula deberá estar dada como $H_0 \leq \mu$.

indicado por el vendedor es posible. Para esta situación, las hipótesis planteadas serán las siguientes:

$H_0: \mu \geq 8,0 \text{ t/ha}$ Planteamiento del vendedor de semillas (Hipótesis nula)

$H_a: \mu < 8,0 \text{ t/ha}$ Planteamiento del agricultor (Hipótesis alternativa)

- Un agricultor cacaoero lleva a la industria su producto, del cual asegura tiene una humedad del 12%. El comprador, dada su experiencia, le dice que dichos granos contienen más del 12% de humedad. Para salir de la discusión, el comprador le plantea tomar 20 muestras al azar de los costales de venta del agricultor, el cual accede a la valoración de humedad. En este caso, las hipótesis se plantean de la siguiente forma:

$H_0: \mu \leq 12\%$ Planteamiento del agricultor (Hipótesis nula)

$H_a: \mu > 12\%$ Planteamiento del comprador (Hipótesis alternativa)

- Se desea evaluar el efecto de la labranza cero y la labranza mínima en caña de azúcar sobre el rendimiento de caña/hectárea. Debido a que existe información que avala las bondades, tanto a la ausencia de labranza como a la mínima que a su vez no permite presuponer qué metodología es mejor, se decide plantear las siguientes hipótesis:

$H_0: \mu_1 = \mu_2$ La hipótesis nula se plantea asumiendo que no existen diferencias entre los efectos de la labranza cero y la labranza mínima.

$H_a: \mu_1 \neq \mu_2$ La hipótesis alternativa se plantea asumiendo que existen diferencias de efectos entre la labranza cero y mínima.

El error tipo I se representa mediante el símbolo α y hace referencia al grado de significancia con que se rechaza H_0 . Convencionalmente se ha establecido como nivel crítico un valor del 5% (0,05), de allí que valores observados de la probabilidad de cometer el error I (llamados *p-valor* o simplemente *p*) menores a éste sea el indicativo del rechazo de H_0 . Por otro lado, el error tipo II se representa mediante el símbolo β y tiene una relación inversa con α ; esto es, si por un lado el investigador se impone valores estrictos de α , controla este error mientras que β tendrá una mayor probabilidad de ocurrencia. El error tipo I está relacionado con la aleatorización de los tratamientos en un experimento, por ello, la única forma de controlarlo o minimizar su probabilidad de ocurrencia es mediante el sorteo de cada uno de ellos al momento de su asignación a cada unidad experimental. Por otro lado, la alternativa para reducir (controlar) la probabilidad de ocurrencia del error tipo II consiste en establecer un número adecuado¹⁸ de repeticiones que permitan, a su vez, tener un ensayo con mayor precisión.

El error β está influido mayoritariamente por el grado de heterogeneidad que contenga el suelo en cuanto a humedad, fertilidad, condiciones texturales o cualquier otro componente edáfico que pueda incidir en la respuesta fisiológica o agroproductiva de un cultivo. En estas circunstancias, para poder controlar eficazmente la dispersión de los datos definida a través del error experimental y no caer en el error tipo II, el establecer un número suficiente de repeticiones es la opción de mayor beneficio. No obstante, existen otras alternativas como la selección de áreas experimentales visualmente homogéneas, manejo uniforme del ensayo,

¹⁸ Para un número adecuado de repeticiones debe entenderse que, si por un lado el aumento de *n* minimiza el error típico de las medias de tratamientos, eso no implica que las diferencias detectadas tengan una importancia en términos prácticos. Esto es, mientras más grande es *n* en un ensayo, dentro de la significancia práctica podríamos estar cayendo en el error tipo I y recomendando tratamientos que, a pesar de ser estadísticamente significativos, rentablemente no producirán un incremento real, sino más bien un decremento de la utilidad económica.

entre otras, que se indicaron en el capítulo anterior; opciones que de alguna manera pueden mejorar la precisión de un experimento y combatir la fácil ocurrencia del error tipo II.

En el campo de la agronomía, como en la mayoría de aplicaciones en otras ciencias, el único tipo de error que se considera como más importante en las pruebas estadísticas aplicadas a los experimentos es α , dejando relegado el error β . No obstante, si bien el cometimiento del error tipo I es el más costoso; el cometer el error tipo II puede mandar al traste alguna alternativa (tratamiento) que se prueba en un experimento y que, en ciencias como la medicina, podrían privar a la población de medicamentos que contrarresten alguna enfermedad relevante.

La forma de medir la falibilidad de una prueba estadística en el rechazo de una H_0 verdadera, es mediante la *potencia de la prueba*, la cual se establece a través del complemento del error tipo II, es decir $1 - \beta$. Generalmente se recomienda como máximo un $\beta = 0,2$, lo que consecuentemente ocasiona que la potencia de una prueba sea como mínimo 0,8; esto es, si existe un efecto, se lo podrá detectar en el 80% de las veces. Situaciones extremas es tener $1 - \beta = 0$ y $1 - \beta = 1$; en el primer caso, si existiera algún efecto, nunca se lo detectaría; mientras que en el segundo, el 100% de las veces será detectado.

La potencia de una prueba depende de tres elementos fundamentales: 1) del tamaño muestral, 2) del nivel de significancia, y 3) de la magnitud de las diferencias entre los tratamientos que se evalúan en un experimento. La relación de la potencia de una prueba tanto con el tamaño muestral como con el nivel de significancia es directa; es decir, si estos dos elementos se incrementan, también la potencia lo hace. Igual situación sucede con las diferencias entre los tratamientos evaluados, si esta diferencia es evidentemente importante, la potencia también se ve incrementada.

Cuando se aumenta el tamaño de la muestra se tiene una disminución matemática del error experimental, con ello una estimación estadística se hace más precisa y consiguientemente, cualquier prueba estadística tiene mayor potencia. Asimismo, dado que el error α es establecido de forma voluntaria por el investigador, para poder aumentar la potencia de una prueba se pueden seleccionar valores de α menos estrictos (más altos); situación que propiciaría un mayor incremento de la probabilidad de caer en el error tipo I, el cual según lo indicado anteriormente, es la opción más costosa.

La elección de la magnitud del error tipo I o nivel de significancia α del 5% ($\alpha = 0,05$), es un criterio adoptado convencionalmente e instaurado por Ronald Fisher (1919), quien fue el pionero del desarrollo de metodologías como el análisis de varianza utilizado en la experimentación. No obstante, adoptar este valor es una decisión de análisis por parte del investigador en cuanto a lo que tiene en juego y de la decisión que luego tendrá que tomar; de allí que se puedan utilizar valores tan estrictos como 0,01 o más permisivos como 0,10. Lo que sí se debe tener en cuenta en la experimentación es que si se quiere rechazar una hipótesis nula, o demostrar la existencia de un efecto de forma contundente, la opción es utilizar valores de α estrictos o reducidos; no obstante, esta decisión, de forma contraria, aumentará la probabilidad de cometer el error tipo II haciendo que la potencia de cualquier prueba estadística disminuya.

La potencia estadística de las pruebas de hipótesis utilizadas en los experimentos permite definir la capacidad de dichas pruebas para detectar diferencias significativas; por ello es que, si la potencia de una prueba es baja, se pone en duda la validez de unos resultados que reporten diferencias estadísticas. Si bien la consideración de un valor α menos estricto hace posible incrementar la potencia de una prueba con el consiguiente incremento de la probabilidad de caer en el error tipo I, la otra forma de aumentar la potencia de una prueba es a través de seleccionar un tamaño muestral adecuado.

Para entender cómo el tamaño muestral afecta la potencia de una prueba estadística, a continuación, se detalla un ejemplo.

Ejemplo de aplicación 3.4

Un agricultor está convencido de que cierta variedad de fréjol tiene un rendimiento máximo de 2,50 t/ha. El técnico de un agroservicio, para refutar la aseveración del agricultor y decir que dicha variedad es capaz de un mayor rendimiento, toma 10 muestras de un campo demostrativo desde donde obtiene una desviación estándar de 0,9 t/ha y una media de 2,90 t/ha. En este caso, la obtención de la potencia de la prueba estadística se detalla en los siguientes pasos:

Para este planteamiento, las hipótesis son:

$$H_0: \mu \leq 2,50 \text{ t/ha (Hipótesis nula)}$$

$$H_a: \mu > 2,50 \text{ t/ha (Hipótesis alternativa)}$$

Bajo la consideración de una prueba de una cola, asumiendo que los datos tienen un ajuste normal y además de considerar un error α del 5%, se tiene por valor crítico de la media el siguiente:

$$Z_\alpha = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

Desde donde se deduce:

$$\bar{X} = \frac{Z_\alpha S}{\sqrt{n}} + \mu$$

Sustituyendo valores en esta última expresión, el límite crítico del estimador de la media es:

$$\bar{X} = \frac{1,64 \times 0,9}{\sqrt{32}} + 2,50 = 2,76 \text{ t/ha}$$

La gráfica correspondiente para el error I será:

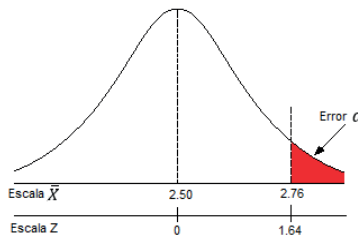


Figura 3.3. Ubicación del valor crítico de la variable y su correspondiente escala Z.

En la figura 3.3 se ha graficado el límite crítico de la variable \bar{X} , a partir de la cual todo valor es considerado significativo; esto es, a partir del cual se deberá rechazar la hipótesis nula. En esta situación, dado que la media tiene un valor de 2,90, se rechaza la hipótesis nula; es decir, la variedad evaluada es capaz de rendir más allá de los 2,50 t/ha de fréjol.

La potencia de la prueba (π) con la cual se está aceptando H_a se define en los siguientes cálculos:

$$Z_\beta = \frac{\bar{X} - \mu_a}{S/\sqrt{n}}$$

$$Z_\beta = \frac{2,76 - 2,90}{0,9/\sqrt{32}} = -0,88$$

La probabilidad del error β , según la distribución normal estándar (obtenida mediante la función *DISTR.NORM.ESTAND* de Excel), es:

$$\beta = 0,19$$

Por lo tanto, la potencia será:

$$\pi = 1 - \beta$$

$$\pi = 1 - 0,19 = 0,81$$

En la figura 3.4 se detalla la potencia de la prueba; de cuyo valor se deduce que existe la probabilidad de que, en 81 veces de 100, se detecten diferencias significativas que permitan establecer que la variedad de fréjol es capaz de rendir, en promedio, arriba de 2,50 t/ha.

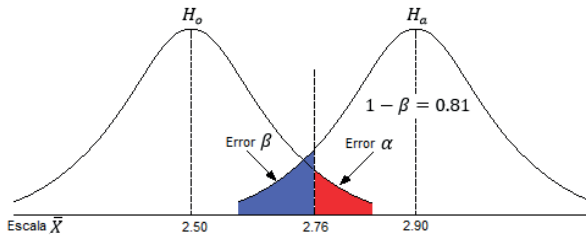


Figura 3.4. Magnitud de la potencia de la prueba.

Ahora, con los datos planteados, veamos qué sucede con la potencia al incrementar la muestra a 80 datos. Para esto, mantengamos los mismos valores de la media y de la desviación estándar; los cuales, debido a que se considera que se distribuyen normalmente, no deben cambiar de forma importante.

Con esta consideración, el nuevo valor crítico del estimador es:

$$\bar{X} = \frac{Z_\alpha S}{\sqrt{n}} + \mu = \frac{1,64 \times 0,9}{\sqrt{80}} + 2,50 = 2,67 \text{ t/ha}$$

La nueva potencia de la prueba será:

$$Z_\beta = \frac{\bar{X} - \mu_a}{S/\sqrt{n}} = \frac{2,76 - 2,90}{0,9/\sqrt{80}} = -1,39$$

Mediante la función *DISTR.NORM.ESTAND* de Excel, se tiene:

$$\beta = 0,08$$

$$\pi = 1 - \beta = 1 - 0,08 = 0,92$$

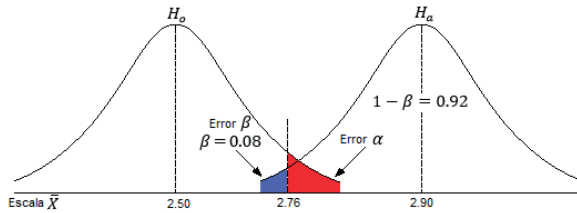


Figura 3.5. Incremento de la potencia de la prueba por el incremento del tamaño muestral.

En este caso, al aumentar la muestra de 32 a 80 se ha incrementado la potencia de 0,81 a 0,92, con lo cual los resultados gozan de una mayor confiabilidad.

En este ejemplo se ha expuesto la forma de cómo incrementar la potencia de una prueba estadística. Si bien las actividades que tienen que ver con el manejo de un experimento son fundamentales en la precisión de los resultados, las cuales son factibles y dependientes del cuidado del investigador, la forma más objetiva de que la validez experimental sea contundente y no sea cuestionable es aumentando el tamaño de la muestra estadística hasta un número adecuado.

Tamaño de muestra para experimentos de dos tratamientos

Como se ha visto en los párrafos precedentes, la precisión del tamaño de una muestra para revelar un efecto está ligado a la potencia de la prueba estadística que se esté utilizando. Esta potencia es dependiente del grado de significancia, de la diferencia entre los tratamientos que se comparan y del tamaño muestral; a su vez, también es el complemento del error β . Estos factores están incluidos en una gráfica conocida como *curva de operación característica* (COC), la misma que se elabora coordinando en las abscisas las diferencias entre las medias de dos tratamientos (δ) para diferentes valores de β bajo un tamaño de muestra determinado. Para más de dos tratamientos, las curvas de operación característica se elaboran utilizando el parámetro Φ , el cual está relacionado con la diferencia entre las medias de dos tratamientos δ .

En los anexos del texto se incluyen varias curvas de operación característica. La figura A1 es una COC para establecer el tamaño muestral de un experimento en donde se compara sólo dos tratamientos. Las figuras A2 hasta la A9 son curvas de operación característica para experimentos con más de dos tratamientos, con las cuáles se pueden definir tamaños muestrales en ensayos de hasta ocho tratamientos.

La COC de la figura A1 del anexo es una curva para la prueba t de Student de dos colas, asumiendo varianzas iguales y considerando una significancia de $\alpha = 0,05$; además de considerar que el experimento es balanceado ($n = n_1 = n_2$). En ésta como en las demás curvas indicadas en los anexos, puede notarse lo siguiente:

- Para un tamaño de muestra específico y un valor de α establecido, mientras más grande sea la diferencia entre las medias de dos tratamientos ($\delta = \mu_1 - \mu_2$), menor es la probabilidad de cometer el error tipo II. En esta situación y para condiciones de un

tamaño muestral y de α establecidos, una prueba estadística detectará con mayor facilidad las diferencias grandes antes que las pequeñas.

- Para una determinada diferencia entre medias de tratamientos y un valor de α específico, a medida que se incrementa el tamaño muestral, la probabilidad de cometer el error tipo II se reduce. De allí que para detectar una diferencia δ establecida, la opción para incrementar el poder o potencia de una prueba se logra incrementando el tamaño muestral.

Ejemplo de aplicación 3.5

Como ejemplo de aplicación consideremos el caso de un experimento en donde se quiere ver las bondades de un *biol* (mezcla de desechos orgánicos fermentados y utilizados como abono) en el rendimiento del maíz. Según la experiencia, el incremento productivo que puede originar dicho fertilizante orgánico es de 1,1 t/ha. Asimismo, por información sobre dicha variedad de maíz se conoce que la desviación estándar del rendimiento, en un cultivo más o menos cuidado, aproximadamente es de 0,7 t/ha. Con estos datos procedemos a determinar el tamaño muestral.

Para utilizar la curva de operación característica de la figura A1 de los anexos, primero debe establecerse el parámetro d . Este parámetro se define mediante la expresión siguiente:

$$d = \frac{|\mu_1 - \mu_2|}{2\sigma} = \frac{|\delta|}{2\sigma} \quad (3.13)$$

El conocimiento de la desviación estándar σ de la expresión 3.13 es la medida más complicada de establecer en la definición del tamaño de muestras, ya que si no se cuenta con información de primera mano establecida por experiencia propia, habrá que recurrir a información generada en otros estudios, los cuales rara vez existen. Esta situación induce a realizar estimaciones de la variabilidad de los datos que en muchos casos llevan a tamaños muestrales insuficientes o excesivos.

Para este ejemplo, el parámetro d es:

$$d = \frac{|\mu_1 - \mu_2|}{2\sigma} = \frac{1,1}{2 \times 0,7} = 0,8$$

Nótese que este resultado es un valor adimensional, lo que facilita el uso de la expresión 3.13 para distintas curvas haciéndola independiente para σ y δ .

Una vez obtenido el valor de d , se puede utilizar la figura A1 del anexo que corresponde a la curva de operación característica para un experimento de dos tratamientos. Otro parámetro requerido para poder utilizar esta curva es el valor del error β . En este caso, para tener una potencia de prueba de un 90% puede seleccionarse un valor de $\beta = 0,10$, lo que permite tener un valor de $n' = 19$. Con este último valor ya se puede definir el tamaño muestral por tratamiento utilizando la relación siguiente:

$$n = \frac{n' + 1}{2} = \frac{19 + 1}{2} = 10$$

Para lograr una precisión del 90% en la decisión que se tome sobre la hipótesis nula del experimento a realizar, cada tratamiento (grupo) deberá tener un mínimo de 10 unidades muestrales o repeticiones ($n = n_1 = n_2 = 10$).

Tamaño de muestra para experimentos con más de dos tratamientos

Para determinar el tamaño de muestra en experimentos de más de dos tratamientos, esto es, el número de repeticiones que deberá tener cada tratamiento, puede hacerse uso de las curvas de operación característica de las figuras A2 hasta A9 que se indican en los anexos. Estas curvas, que en su estructura se elaboran para el error II considerando un tamaño de muestra determinado, para una diferencia entre tratamientos especificada y un nivel de significación dado, se han elaborado ubicando en las abscisas el parámetro Φ bajo el enfoque de efectos fijos y asumiendo un experimento balanceado (igual número de repeticiones por tratamiento). Este parámetro tiene una relación directa con la diferencia o variabilidad entre tratamientos, debido a que en las curvas α y n ya están dados. Su determinación puede realizarse mediante la expresión:

$$\Phi = \frac{D}{\sigma} \sqrt{\frac{n}{2a}} \tag{3.14}$$

siendo D la diferencia entre las medias de dos tratamientos cualesquiera (efecto) y a (no alfa) el número tratamientos en estudio. En este caso, el cuadrado de la desviación estándar σ , que corresponde a la varianza, es equivalente al cuadrado medio del error experimental CM_E ($\sigma^2 = CM_E$) de todos los tratamientos. En realidad D deberá corresponder con la diferencia máxima entre tratamientos que se espera suceda en el experimento, de allí que el tamaño muestral que se defina sea el mínimo para una potencia especificada. Para la aplicación de este procedimiento, veamos el siguiente ejemplo:

Ejemplo de aplicación 3.6

Se ha previsto realizar un experimento sobre riego en una plantación de tomates. El sistema es localizado (goteo) y se propone determinar el efecto de cinco láminas de riego a partir de la evapotranspiración (ETC) del cultivo, definiéndose como tratamientos los siguientes: 0,7ETC; 0,8ETC; 0,9ETC; 1,0ETC y 1,1ETC. El investigador estima que la diferencia máxima entre dos tratamientos cualesquiera es 4,0 kg tomates/planta. Además, a partir de estudios previos, el investigador supone que la variabilidad de los datos (varianza) es: $CM_E = 2,0$. Determinar el tamaño muestral (número de repeticiones por tratamiento), de tal forma que el rechazo de la hipótesis nula acertadamente tenga una probabilidad de 90% o más (potencia de la prueba).

Según la expresión 3.14, se tiene lo siguiente:

$$\Phi = D \sqrt{\frac{n}{2CM_E a}} = 4 \times \sqrt{\frac{n}{2 \times 2,0 \times 5}}$$

$$\Phi = 0,89\sqrt{n}$$

Número de repeticiones n	Parámetro Φ	Grados de libertad del error experimental $a(n - 1)$	Error tipo II β	Potencia de la prueba $(1 - \beta)$
3	1,54	5(3-1)=10	0,45	0,55
4	1,78	15	0,22	0,78
5	1,99	20	0,10	0,90

Según lo resuelto en la tabla antes indicada, para que la prueba tenga una potencia de 0,90 hacia arriba, se requieren como mínimo 5 repeticiones por tratamiento.

Nótese que, para utilizar las curvas de operación característica, el parámetro Φ es fundamental. Sin embargo, su deducción se hace compleja debido a los valores de varianza y a la diferencia entre tratamientos que se requieren, los mismos que, salvo experiencias previas o datos de literatura existente, son difíciles de tenerlos a mano. El enfoque del parámetro Φ puede ser más complicado aún si se establece mediante la expresión siguiente:

$$\Phi = \frac{n \sum \tau_i^2}{\alpha \sigma^2} \quad (3.15)$$

En donde:

$$\tau_i^2 = \sum (\mu_i - \bar{\mu})^2$$

Siendo μ_i la media de un tratamiento cualquiera y $\bar{\mu}$ la media de medias de todos los tratamientos a evaluar.

Como se podrá advertir, en esta expresión si bien no se requiere estimar una diferencia entre tratamientos (D), de manera aproximada se hace necesario conocer cuáles serán las medias de cada uno de los tratamientos que se hayan considerado en el estudio, condición que hace de este enfoque una alternativa mucho más complicada que la propuesta por la expresión 3.14.

Tamaño de muestra por diferencia mínima significativa (intervalo de confianza)

Una alternativa un poco más subjetiva para establecer el tamaño muestral o número de repeticiones de cada tratamiento de un estudio es a partir de la diferencia entre dos cualesquiera de ellos que el investigador toma como límite para definir que dicha diferencia es significativa. Esta diferencia, en términos de no significancia, corresponde a un intervalo de confianza cuyo incremento a partir de la media es justamente la diferencia establecida subjetivamente por el investigador. La diferencia, a partir de la cual existe un efecto de importancia estadística, se corresponde con una prueba de comparación de medias conocida como diferencia mínima significativa (DMS).

La prueba DMS es una de las alternativas que se utilizan para comparar promedios de tratamientos y segregar los mejores, luego de que el análisis de varianza revela diferencias significativas entre ellos. La DMS entre dos promedios se representa mediante el esquema de la figura 3.6 y su valor se obtiene mediante la expresión 3.16.

$$DMS = t_{\alpha/2; N-t_i} \sqrt{\frac{2CM_E}{n}} \quad (3.16)$$

Despejando desde la expresión 3.16, el número de repeticiones se determina por:

$$n = \frac{2(t_{\alpha/2; N-t_i})^2 CM_E}{(DMS)^2} \quad (3.17)$$

En esta expresión, el valor de $t_{\alpha/2; N_o - t_i}$ corresponde al estadístico de prueba bajo la distribución t de Student, que es seleccionado de acuerdo a un nivel de significancia ($\alpha/2$) (prueba de dos colas) y a los grados de libertad del error experimental ($N_o - t_i$). Los tratamientos que previamente son establecidos en un experimento están representados por t_i , mientras que N_o corresponde al número de total de unidades experimentales que tendrá el ensayo (valor de prueba) y lo cual también implica la selección de un tamaño muestral de prueba (n_o). El valor de CM_E es el cuadrado medio del error experimental que se calcula en los análisis de varianza y es igual a la varianza que existe dentro de los tratamientos o grupos ($CM_E = \sigma^2$).

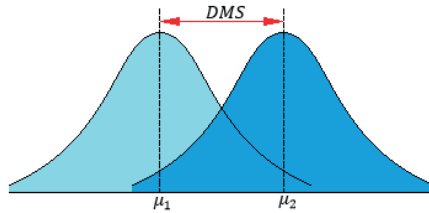


Figura 3.6. Esquema de la diferencia mínima entre dos promedios a partir de la cual se establece variación significativa entre dos tratamientos (DMS).

Como la DMS corresponde a la diferencia mínima significativa entre dos promedios de tratamientos, esta diferencia puede expresarse como $DMS = d$. Por lo tanto, replanteando la expresión 3.17, se tiene:

$$n = \frac{2(t_{\alpha/2; N_o - t_i})^2 \sigma^2}{d^2} \quad (3.18)$$

Ejemplo de aplicación 3.7

Ejercicio de aplicación sobre el tamaño de muestra para experimentos agrícolas

Se ha previsto realizar un experimento en donde se van a probar cinco dosis de nitrógeno en maíz bajo condiciones no controladas (campo abierto). Se pretende utilizar para tal propósito 4 repeticiones por cada tratamiento. Según lo indicado en estudios preliminares y de acuerdo a la magnitud de las dosis, se estima que la varianza dentro de los tratamientos es de 225 qq/ha y que la diferencia entre dos tratamientos, a partir de la cual se considera un efecto significativo, es de 20 qq/ha. Establecer el número de repeticiones necesarias que deberá tener cada tratamiento.

Datos:

$t_i = 5$	Número de tratamientos
$R_o = 4$	Número de repeticiones de prueba
$N_o = 5 \times 4 = 20$	Número de unidades experimentales de prueba
$\sigma^2 = 225 \text{ qq/ha}$	Varianza dentro de los grupos (cuadrado medio del error experimental)
$d = 20 \text{ qq/ha}$	Diferencia estimada entre dos tratamientos cualesquiera

Valor del estadístico t de Student se puede obtener de la tabla 2A del anexo en función de $\alpha/2 = 0,05/2 = 0,025$ y de acuerdo al número de grados de libertad en el error experimental que tendrá el experimento, esto es: $N_0 - t_i = 20 - 5 = 15$. Por lo tanto:

$$t_{0,025;15} = 2,131$$

El número de repeticiones más adecuado será:

$$R = \frac{2 \times (2,131)^2 \times 225}{20^2}$$

$$R = 5,1$$

De acuerdo a este resultado y considerando que la diferencia entre tratamientos establecida por el investigador es un límite mínimo a partir de la cual esta es significativa, el experimento deberá tener un tamaño muestral de seis repeticiones por tratamiento.

Recomendación empírica para definir el número de repeticiones

Hasta aquí se ha indicado que el tamaño muestral, aparte de factores como la confianza, el error máximo admisible o la varianza, también tiene una dependencia según el tipo de investigación. Cuando ésta es de característica no experimental, los datos de la (s) variable (s) dependiente (s) evaluada (s) podría tener un comportamiento totalmente estocástico, debido a que sobre la unidad de muestreo o de observación¹⁹, desde la cual se extraen, no hay influencia de la acción del investigador y por ello es que se requieren tamaños muestrales relativamente grandes.

En el caso de investigaciones experimentales (también en las cuasi-experimentales), la variabilidad presente en las unidades de muestreo desde donde se extraen los datos, en cierta magnitud, es controlada por el investigador porque quiere conocer el efecto de un tratamiento sin la influencia de agentes externos e intrínsecos que pudieran tergiversar la respuesta. En este sentido, el tamaño muestral de un experimento es reducido. Sin embargo, muy a pesar de valores estadísticos como la varianza, un factor elemental de los tamaños muestrales tanto para investigaciones no experimentales como en las que sí lo son, es el económico. En estas últimas, los costos que se originan por el desarrollo de todo el ensayo y la medición de las variables son los limitantes principales del tamaño de una muestra. Por esta causa, en los experimentos se deja de lado la potencia estadística de una prueba que, en términos prácticos, podrían hacer que estas investigaciones sean excesivamente caras.

Como hasta esta parte se ha visto, para estimar el tamaño muestral en los experimentos existen algunas alternativas estadísticas que se pueden encontrar en los textos de diseño experimental; empero, no todas son de fácil aplicación, debido a que se requiere un conocimiento previo de la variabilidad interna (varianza) a la que están expuestos los experimentos, además debe conocerse aproximadamente las diferencias entre las respuestas de los tratamientos. Esta información generalmente no se tiene a la mano y se termina estimando condiciones que bien pueden llevarnos a los extremos de caer en los errores I o II.

¹⁹ Es importante distinguir que no es lo mismo “unidad de muestreo” que “unidad de observación”. Si bien en algunos estudios pueden significar lo mismo, en otros hacen referencia a conceptos distintos. Una unidad de observación (o también llamada “unidad de análisis”) es un elemento de la unidad de muestreo. Por ejemplo, en los muestreos por conglomerados, cada conglomerado es una unidad de muestreo, pues dentro de éstas es desde donde se obtendrán las unidades de observación. En el muestreo aleatorio simple, la unidad de muestreo es la unidad de observación.

Según se ha visto en las fórmulas de cálculo de tamaños muestrales, dichos tamaños tienen una relación directamente exponencial con la variabilidad (desviación estándar) de los datos, requiriéndose, mientras el resto de términos se mantengan constantes, un mayor tamaño muestral a medida que dicha variabilidad sea mayor; situación que se contrapone a la disponibilidad económica de una investigación. Con estos antecedentes, para definir el número de repeticiones en un experimento habrá que considerar lo siguiente:

- Si se estima que las diferencias que se podrían encontrar entre los tratamientos son relativamente grandes, se necesitarán pocas repeticiones por tratamiento.
- Si se estima que las diferencias que se vayan a encontrar entre tratamientos son mínimas, habrá la necesidad de considerar un importante número de repeticiones, de tal manera que se puedan detectar diferencias estadísticas entre ellos.
- Si se estima que en el experimento se tendrá un considerable error experimental (especialmente en variables que traten sobre mediciones de plagas), o lo que es lo mismo, que al interior de cada tratamiento se produzca una gran variabilidad, se hará necesario un mayor número de repeticiones por tratamiento; lo que consecuentemente permita reducir dichas divergencias y hagan posible una mayor precisión en los resultados del ensayo.
Considerando lo indicado en este punto, es importante considerar que en el caso de ensayos en donde se valora más de una variable, el investigador deberá seleccionar la que más dispersión pudiera presentar; siempre que esté directamente relacionada a los objetivos de la experiencia.
- Si el experimento va a estar constituido por muchos tratamientos, de tal manera que se tengan suficientes grados de libertad en el error experimental, con pocas repeticiones se podrá tener conclusiones válidas sobre el experimento.

Por otro lado, en los experimentos agrícolas existen dos situaciones muy diferentes que, en cierta medida, pueden incidir en el tamaño muestral. La primera tiene que ver cuando el experimento se lleva a cabo bajo situaciones muy controladas u homogéneas, lo cual ocurre cuando se realizan ensayos en invernaderos, en viveros, en macetas, en jaulas; o en ambientes de laboratorio como, por ejemplo, la reproducción de plantas *in vitro*. En estas condiciones se pueden manejar relativamente un número reducido de repeticiones ya que la única fuente de variación permitida será la de los tratamientos bajo estudio. Bajo esta situación, el diseño experimental que se utiliza es el completamente aleatorizado, el cual no puede controlar ninguna fuente de variación secundaria; sin embargo, es el diseño que genera el mayor número de grados de libertad en el error experimental, lo que consecuentemente favorece una mayor precisión al minimizar más intensamente la varianza dentro de los tratamientos.

La segunda situación está relacionada a los experimentos agropecuarios que se realizan a campo abierto, en donde las condiciones en que se desarrollan los sujetos de estudio (cultivos o animales) no son controladas por el experimentador. En esta condición existen muchos factores que podrían incidir en los resultados, tal es el caso del clima, las propiedades del suelo, la forma del riego, el pasto que comerá un animal, etc., lo que ocasiona que cada unidad experimental estará en contacto con un ambiente muy heterogéneo. Esta heterogeneidad del medio fundamentalmente está influenciada por las características físicas, químicas y biológicas del suelo, que permiten prever que si se aumentan repeticiones el experimento se expondrá a una mayor variabilidad del medio; incrementando consecuentemente la probabilidad de que las conclusiones que extraigamos del ensayo nieguen algún efecto real de los tratamientos (error tipo II).

Por otro lado, realizar un experimento con insuficientes repeticiones permitirá que el error experimental de un ensayo sea lo suficientemente grande como para impedir la detección de diferencias entre tratamientos, ocasionando que la prueba contenga un déficit de precisión y que incurramos en el error tipo II. Sin embargo, al observar detenidamente el valor del error

estándar de la media de un tratamiento (S/\sqrt{n}), es fácil deducir que si queremos minimizar este error, tendremos que aumentar el número de repeticiones. Teniendo en cuenta las dos contraposiciones indicadas y de acuerdo a la experiencia del autor y de muchos investigadores, es recomendable que un experimento (agrícola, especialmente) tenga como mínimo 4 repeticiones, pudiendo ampliarse a un máximo de 10, según el número de tratamientos del ensayo. Esta última recomendación permitirá asegurar, tal como veremos a continuación, los suficientes grados de libertad del error experimental que logren una mayor potencia en las pruebas estadísticas como en el caso del análisis de varianza.

Tamaño de muestra por el criterio empírico de los grados de libertad del error experimental

Una consideración establecida bajo una base empírica para definir la cantidad de repeticiones que deberá tener cada tratamiento de un ensayo se ha venido fortaleciendo al considerar el número de grados de libertad del error experimental que generan los modelos de análisis de varianza (ANDEVA), cuyos esquemas pueden establecerse al inicio de la experiencia. Dada la gran variabilidad del ambiente y del propio suelo bajo los cuales se deben necesariamente llevar a cabo los experimentos agrícolas, es esperable que las diferencias entre los tratamientos sean contundentes y de aplicación práctica en el campo, de allí que el número de repeticiones no sea insuficiente como para no detectar nada y no tan exagerado como para declarar diferencias que, dentro de la estadística, podrían ser significativas, pero de conveniencia limitada desde el punto de vista económico.

Los extremos en los grados de libertad del error experimental (GL_E) de un ensayo (muy altos o muy bajos) pueden llevarnos a cometer los errores I y II, respectivamente. Con valores mínimos de GL_E , o lo que es lo mismo con escasas repeticiones, el experimento carecerá de precisión y las pruebas tendrán una mínima potencia ($1 - \beta$); con lo que fácilmente podríamos cometer el error β . Por otro lado, cuando se tiene un elevado número de grados de libertad en el error experimental se protege al experimento del error II, pero esto puede llevarnos a incurrir en el error I²⁰ al tomar como efecto real una mínima diferencia que bien podría ser considerada como un efecto que ocurre por la casualidad. Un excesivo número de grados de libertad en el error experimental implica un número importante de repeticiones que tenderán a minimizar el error estándar de la media de un tratamiento; esto hará que la varianza o cuadrado medio del error experimental sea tan pequeño que efectos nimios sean declarados significativos.

Incrementar sustancialmente los grados de libertad del error experimental, lo cual se logra aumentando el tamaño muestral, no favorece de forma correspondiente al incremento de la precisión de un experimento; esta precisión tiene un límite práctico, a partir del cual un mayor número de repeticiones sólo encarecerá la investigación. Este contexto se explica en el esquema de la figura 3.7, la cual se ha realizado considerando varios ensayos con igual número de tratamientos, pero con distintas repeticiones, en donde el cuadrado medio del error experimental (varianza intragrupos) no disminuye linealmente al aumentar los GL_E ; más bien se observa que existe una inflexión en la curva desde donde el aumento de repeticiones no contribuye de manera sustancial (o éste es nulo) a una mayor precisión (menor varianza).

²⁰ Entiéndase que el cometimiento del error tipo I en un experimento que tiene un elevado número de repeticiones no va por el lado estadístico, sino por el criterio de aplicación práctica. Un gran número de repeticiones “minimiza” el error experimental en términos estadísticos, a tal punto de permitir detectar diferencias significativas entre tratamientos que, llevados a la realidad de la aplicación en el campo, podría inducirnos a seleccionar tratamientos que no ofrezcan rentabilidad, aunque hayan resultado con la mayor media que la estadística define como “significativa”. Si bien un tratamiento debe ser un “ganador” en términos estadísticos, también lo debe ser en términos de rentabilidad.

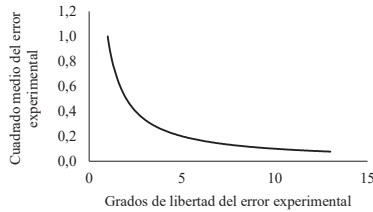


Figura 3.7. Esquema de la relación entre la varianza intragrupos (CM_E) y los grados de libertad del error experimental.

Un número suficiente de grados de libertad del error experimental permitirá observar diferencias significativas entre tratamientos con una precisión de importancia y evitar, consecuentemente, la aceptación de una hipótesis nula cuando es falsa en el contexto estadístico. Lo indicado es deducible al observar el esquema de un análisis de varianza para un diseño completamente al azar (tabla 3.3), en el que se conoce que el cuadrado medio del error experimental (CM_E) tiene una relación inversa con los grados de libertad del error experimental (GL_E); es decir, aumentando convenientemente los GL_E se logra minimizar la varianza entre unidades experimentales de un mismo grupo (tratamiento) y, por ende, obtener un estadístico calculado de Fisher (F_{cal}) de mayor contundencia. Esto último evitará que estemos aceptando una hipótesis nula falsa que podría suceder cuando los GL_E son insuficientes.

En este escenario, un criterio muy práctico en la determinación del tamaño muestral de los experimentos agrícolas es la consideración de que los grados de libertad del error experimental en los modelos de ANDEVA deben ser al menos 12 ($GL_E \geq 12$), teniendo cuidado de no extendernos demasiado por cuanto, si bien con un valor de GL_E de 12 estamos asegurados con una gran probabilidad de no caer en el error tipo II (β), aumentando extensiva e innecesariamente los GL_E a través de más repeticiones, podríamos estar cayendo virtualmente en el error tipo I (α) al establecer efectos significativos de poca importancia; o lo que significa esto último, adoptar un tratamiento que no ofrezca una mayor rentabilidad aunque estadísticamente sea el ganador. A este respecto, no olvidemos que los ambientes experimentales de los ensayos agrícolas son tan heterogéneos que, para evitar caer en el error α (el más costoso comparado con el error β), las diferencias que se detecten entre los tratamientos deben ser sustanciales como para hacer una recomendación general de la alternativa ganadora. Además, tener valores mucho más allá de los 12 grados de libertad en el error experimental no incidirá en un cambio sustancial del valor del estadístico crítico (tabular) que se toma de referencia en los análisis de varianza.

Tabla 3.3. Esquema del análisis de varianza para un diseño completamente al azar.

Fuentes de variación	GL	SC	CM	F_{cal}	F_{tab}
Total	GL_T	SC_T			
Tratamientos	GL_t	SC_t	$CM_t = \frac{SC_t}{GL_t}$	$F = \frac{CM_t}{CM_E}$	F_α
Error experimental	GL_E	SC_E	$CM_E = \frac{SC_E}{GL_E}$		

El manejo de un excesivo número de grados de libertad en el error experimental por el hecho de tener un número importante de repeticiones, en el caso de ensayos de laboratorio o de ambientes controlados, podría resultar en una ineficiencia económica de la investigación en el sentido de que los efectos significativos entre tratamientos con pocas repeticiones pueden hacerse fácilmente evidentes; situación que se debe a la alta precisión que tendría la experiencia al mantener reducida la varianza intragrupos. No obstante, queda bajo el criterio del investigador analizar lo que se ha indicado en esta parte del texto al momento de establecer el número de repeticiones que tendrá su experimento. En todo caso, siempre existirán dos situaciones contrapuestas, el aspecto económico y la precisión del ensayo que en alguna medida deberán equilibrarse para que la experiencia resulte económica y sus resultados definan claramente los efectos reales (y prácticos) del factor de estudio.

Teniendo en cuenta los modelos de análisis de varianza respecto de los grados de libertad del error experimental, tanto para la distribución completamente al azar como para la de bloques completos al azar, el número de repeticiones de cada tratamiento se define con las expresiones que a continuación se describen:

Considerando un diseño completamente al azar (DCA), se tendrá:

$$GL_E = nt - t \quad (3.19)$$

de donde se tiene:

$$n = \frac{GL_E + t}{t} \quad (3.20)$$

siendo n las repeticiones y t el número de tratamientos a evaluar. El valor de GL_E deberá cuidarse que siempre sea igual o mayor a 12.

Para el caso de diseños con bloques completos al azar (DBCA) la expresión puede modificarse como:

$$GL_E = nt - t - n + 1 \quad (3.21)$$

Desde donde también se tiene:

$$n = \frac{GL_E}{t - 1} + 1 \quad (3.22)$$

Considerando a los bloques como las repeticiones n de cada tratamiento y que $GL_E \geq 12$.

Retomando el dato de las 5 dosis de nitrógeno en maíz del ejemplo de aplicación 3.7, además de considerar un diseño completamente al azar, el número de repeticiones para este experimento según este criterio deberá ser:

$$n = \frac{GL_E + t}{t} = \frac{12 + 5}{5} = 3,4$$

Por el lado de la seguridad, el valor calculado se deja en 4 repeticiones.

CAPÍTULO IV

VARIABLES Y ALGUNAS ALTERNATIVAS DE ANÁLISIS

Aunque comúnmente el rendimiento es la variable más importante en los ensayos agronómicos, el efecto de los tratamientos no debe supeditarse exclusivamente al valor que toma esta variable en los cultivos. Se necesita verificar otras variables concomitantes para tener certeza del aparente efecto y, como beneficio científico colateral, la posibilidad de fundar teoría agronómica...



Variables agronómicas de respuesta en los ensayos experimentales

La razón de ser de los experimentos son las respuestas que de ellos se obtienen. Estas respuestas, al igual que el efecto de cualquier fenómeno en la naturaleza y a pesar del contexto controlado bajo el que se miden, siempre son diferentes, debido a los múltiples factores que influyen en el comportamiento divergente de las mismas. Este conjunto de factores causantes y de respuestas se definen como variables que le sirven a cualquier investigador para describir y explicar el suceso que investiga. Si bien los experimentos se ajustan a un rigor de condiciones que son total o parcialmente controladas, la relación causa-efecto también se hace presente; por ello, estos caben también dentro del tipo de investigación explicativa.

En forma general, en los ensayos experimentales se tienen dos tipos de variables. La *variable independiente de estudio*, o también llamada factor, que es la que manipula el investigador, la misma que se define a través de distintos niveles y a los cuales se los denomina como tratamientos²¹; y la variable dependiente, que vienen a constituirse como la respuesta de un efecto producido por la variable independiente de estudio. Una definición de estas variables en los ensayos agrícolas puede ser la relación funcional del nitrógeno con la altura de planta. El nutriente mencionado viene a constituirse como la variable independiente de estudio, que se valora a uno o varios niveles (una o varias dosis), y la variable dependiente es la respuesta de una planta medida a través de su altura.

En un experimento también pueden considerarse dos tipos de variables adicionales a la variable de estudio que también actúan de forma independiente, pero como no son de interés para el investigador, se acostumbra a definir las como *variables independientes secundarias*²² o de ruido. En este sentido, se pueden tener variables independientes controladas, lo cual se logra con la propia distribución experimental y/o con el manejo del ensayo. Un ejemplo de esta variable es la pendiente unidireccional del suelo que puede ser controlada con una distribución de bloques completos al azar, ubicando los bloques en forma perpendicular a dicha pendiente. El segundo tipo de variable adicional independiente se conoce como no controlada o de "ruido"²³; debido a que, o bien se presentan de forma imprevista, o está tan implícita en el ambiente experimental que no se puede controlar. Este último tipo de variable influye el resultado de un experimento de forma general, confundiendo el efecto real del tratamiento; como, por ejemplo, una lluvia imprevista en un ensayo que trate sobre dosis de riego.

Las variables de respuesta tienen una relación directa con los objetivos de una investigación y deben ser consecuentes con el factor o variable independiente de estudio; de hecho, si los objetivos de un ensayo están claramente definidos, las variables de respuesta son fáciles de establecer. Así, por ejemplo, si un experimento se ha diseñado para probar varias dosis de

²¹ En experimentos unifactoriales, los tratamientos son los niveles del factor de estudio; mientras que en los experimentos que incluyen a más de un factor de estudio, los tratamientos son las combinaciones de los niveles de dichos factores. Con lo último se generan los arreglos factoriales como resultado de la combinación de niveles de dos o más factores de estudio.

²² Una variable independiente aquí definida como secundaria, debe entenderse como aquella fuente de variabilidad adicional a los tratamientos (la variable primaria de interés) que toma su importancia en el sentido de que ésta puede disminuir la precisión de un experimento, de allí que el investigador debe buscar la combinación adecuada entre distribución con bloqueo (el diseño experimental) y manejo del ensayo para su control efectivo y, se convierta así, en una variable de ruido controlada o factor *nuisance*.

²³ Una variable independiente secundaria, no controlada, afectará la exactitud de una respuesta cuando su incidencia en un ensayo es general, pero si es parcial, su efecto mayoritario será en la precisión. Cuando esta variable independiente afecta a todo el ensayo, haciendo que la respuesta se aleje sustancialmente del verdadero valor (restando exactitud), se la conoce como *variable de "ruido"*. Aquí debe entenderse como "ruido" al efecto que sesgará a la (s) variable (s) de respuesta de un experimento, ocasionado por variables que son difíciles de controlar, como es el caso del clima. Las variables independientes secundarias que pueden controlarse con la distribución experimental, la aleatorización y/o manejo del ensayo, si se les permite que actúen sin estos controles, incidirán directamente en la precisión y, en algunos casos, en el cometimiento del error tipo I de la estadística.

encalado con el fin de recuperar un suelo ácido, es lógico que la variable de medición debe ser el pH del suelo. Si se evalúan insecticidas, es lógico que las variables directas deben ser las poblaciones o el nivel de daño de la plaga al cultivo; si se evalúan nematocidas, las variables deben estar relacionadas a la población de nematodos a nivel radicular o el nivel de daño al sistema radicular; si se estudian métodos de polinización artificial, la variable directa deberá ser la proporción de fructificación de un cultivo. Sin embargo, para contribuir a la contundencia de un resultado, o en algunos casos verificar el alcance del efecto de un tratamiento, se miden otras variables complementarias; siendo una de estas variables, casi infaltable en los ensayos, el rendimiento de un cultivo que es lo que al final le interesa al productor.

Las variables dependientes de un experimento pueden ser de medición relativamente inmediata a la aplicación de los tratamientos, o su medición puede esperar todo el tiempo que dure el cultivo. Para el primer caso, como ejemplo están los ensayos con insecticidas, cuyos efectos se miden a las pocas horas de haber aplicado los tratamientos; mientras que, para el segundo caso, están los estudios sobre distancias de siembra y cuya verificación del efecto en el rendimiento habrá de esperar todo el ciclo del cultivo.

Según la forma como se midan, las variables dependientes pueden clasificarse como *cualitativas* o como *cuantitativas*. Las variables cualitativas, también conocidas como *categorías*, definen la magnitud de una variable mediante una cualidad o característica que no es numérica. Las variables cualitativas se clasifican a su vez, en variables *nominales* u *ordinales*. Las variables clasificadas como nominales delimitan una característica mediante palabras que no obedecen a ningún orden, por lo tanto, no pueden ser analizadas estadísticamente mediante pruebas paramétricas. Las variables nominales se pueden encontrar en estudios descriptivos; o generalmente como variables independientes de los experimentos agrícolas como, por ejemplo, variedades de semillas en un ensayo de adaptación.

Las variables cualitativas ordinales también se definen mediante palabras, pero obedecen a un orden. Generalmente se utilizan escalas para hacerlas numéricas y poder aplicar alguna prueba de la estadística paramétrica. Como ejemplo de este tipo de variables podría ser la severidad de una enfermedad en un cultivo que, al final, toma un valor numérico luego de la traducción cualitativa-cuantitativa que se hace con una escala que puede ser estándar (probada científicamente) o de tipo arbitraria y adaptada por el investigador. Si bien en estas variables es posible el uso de escalas, la herramienta de valoración tiene en algunos casos un fuerte predominio subjetivo que toma de base los sentidos humanos como los ojos (generalmente en agricultura), de allí que quienes realicen estas valoraciones deberán poseer una gran experiencia y ser objetivos en sus decisiones.

Finalmente, se tienen las variables cuantitativas que, a su vez, se dividen en *continuas* y *discretas*. En el primer caso están las variables que admiten cualquier valor real dentro de un intervalo, tales como el peso de frutos o la altura de plantas; por lo tanto, en este tipo de variables se admiten valores con decimales. Las variables discretas son valores aislados en un intervalo expresado mediante números enteros, tales como el número de frutos por planta o el número de plantas por unidad de muestreo; motivo por el cual también se las conoce como *variables de conteo*. En esta parte, es preciso indicar que las variables de tipo ordinal, en su traducción numérica a través de una escala, pueden ser procesadas como variables discretas.

Otra de las formas de clasificar las variables es la que toma de referencia a los niveles de medición de las mismas. En este caso se han establecido cuatro niveles: *nominal*, *ordinal*, *de intervalo* y *de razón*. Una medida nominal, de acuerdo a lo indicado en párrafos precedentes, corresponde al nivel más bajo de medición, cuyo proceso no va más allá de una clasificación de los datos y en la cual queda restringida la aplicación de la mayoría de pruebas estadísticas. Asimismo, el nivel ordinal puede traducirse desde su concepción cualitativa a cuantitativa a

través de escalas que manifiestan un orden y con cuyos datos podrían aplicarse determinadas pruebas estadísticas, especialmente de tipo no paramétrica. Finalmente, los niveles de intervalo²⁴ y de razón se encasillan dentro de las variables cuantitativas continuas y discretas, con sus datos es posible toda prueba correspondiente que la estadística paramétrica ofrece.

En la agronomía existen una serie de variables de respuesta, medibles experimentalmente y cuya mayoría son de tipo cuantitativa. Estas variables deben estar en consonancia con los objetivos que se persiguen en un ensayo y con la (s) hipótesis propuesta, de tal forma que permitan valorar el verdadero efecto de los tratamientos y no se incluyan otras que más bien tiendan a confundir dicho efecto. Es así como también existen casos en que se abunda con variables supuestamente dependientes y que, en términos agronómicos de una especie, no son una respuesta (o efecto) a los tratamientos aplicados. De esto último se tiene, por ejemplo, la evaluación de la precocidad de un cultivo mediante los días a la cosecha en un ensayo sobre fertilizantes; siendo que, en esta situación, la duración de todo este periodo vegetativo no está dado por los fertilizantes, sino por las condiciones ambientales y la expresión genética de las plantas.

En los párrafos subsiguientes se hace una descripción sobre algunas variables dependientes de característica agronómica que se evalúan en los experimentos agrícolas, con el propósito de tener una clara concepción del proceso de valoración; pues en ciertos casos se omiten criterios importantes que deben tomarse en cuenta al momento de medirse y que, en muchos casos, están relacionados con la precisión experimental.

Altura de planta

La altura de planta es una variable cuantitativa continua que se mide en cultivos anuales como el arroz, maíz, frejol, maní, soya, pimiento, tabaco, siempre que su desarrollo sea erecto y sean plantas que no necesiten de tutorado. Esta variable debe medirse en plantas monocotiledóneas desde el nivel del suelo hasta el ápice central; y en plantas dicotiledóneas, su medición deberá realizarse hasta el ápice del brote más alto que muestren cada una de ellas, tal como se indica en la figura 4.1.

Generalmente es una variable complementaria, que permite tener una idea del desarrollo de la masa foliar de las plantas y deben medirse por una única vez en la etapa de máximo crecimiento (etapa de floración); salvo en estudios de dinámica de crecimiento, en donde se realizan varias mediciones con el propósito de generar un modelo del comportamiento de esta variable en el tiempo o para correlación con otras variables.

²⁴ Una variable medida en un nivel de intervalo contiene un cero relativo y con sus valores se pueden realizar operaciones aritméticas de suma y resta, y admite la multiplicación y la división únicamente con las diferencias entre dos valores de una escala. La temperatura en grados Celsius (°C) es un ejemplo clásico de este tipo de variables, cuyo valor cero no significa ausencia de temperatura.



Figura 4.1. Esquema de la medición de altura de plantas.

Se demanda su medición en ensayos de nutrición, de fitohormonas, de adaptación de variedades o híbridos, en estudios de densidades de siembra, de preparación de suelos, de dosis de riego, hasta en estudios sobre ciertas plagas que pueden atrofiar el crecimiento de las plantas como en el caso de *Spodoptera frugiperda* en maíz.

Para su medición se utilizan los flexómetros o reglas graduadas y la unidad de medida es en centímetros para plantas de porte pequeño o en metros para plantas de porte alto, utilizando un decimal para la primera unidad y dos decimales para la segunda, todos redondeados. Es una variable que comúnmente se ajusta a una distribución normal siempre que el experimento se haya manejado homogéneamente, considerando como coeficiente de variación un máximo del 20%.

En ciertos cultivos como la sandía, el pepino o el melón, cuyo desarrollo es de tipo rastrero, las mediciones deben hacerse en la guía principal, siempre que no se realicen podas de formación y/o fructificación que puedan incidir en el efecto de los tratamientos en esta variable. La definición acorde en estos casos es *longitud de guía*.

Número de frutos/vainas por planta

Esta es una variable cuantitativa discreta que se mide en experimentos de nutrición, de distanciamientos de siembra, de inducción floral y polinización, de adaptación de variedades, de control de plagas y de preparación de suelos. Pueden medirse tanto en cultivos anuales como el tomate o el fréjol; o en los perennes, tales como el cacao, mango o aguacates. Para su medición, deben considerarse las mismas plantas seleccionadas aleatoriamente al inicio de la etapa de cosecha cuando esta variable se reporta en función de más de una cosecha.

A veces en este tipo de variable se hace una distinción de los frutos que son comerciales de los no comerciales, especialmente cuando se

realizan evaluaciones sobre cruce genético o de técnicas de poda; o también la diferencia entre los frutos fisiológicamente maduros de los no maduros. También puede realizarse una distinción de frutos sanos por planta cuando se realizan estudios sobre fungicidas, como en el caso de los ensayos sobre *moniliasis* en cacao.

Diámetro y longitud de frutos

Estas son variables cuantitativas continuas que se miden en ensayos de nutrición, de técnicas de poda, de distanciamientos de siembra, de asociación de plantas, de bioestimulantes hormonales, de adaptabilidad y caracterización de variedades e híbridos, de mejoramiento genético, de riego, hasta en experimentos sobre ciertas plagas. Son variables que se miden generalmente en frutales y en hortalizas cosechadas, utilizando calibradores en unidades de milímetros para frutos relativamente pequeños y en centímetros para frutos grandes, como las sandías, redondeando a un decimal en los dos casos.

Debe evitarse el uso de cinta métrica para la determinación de estas variables, dado que pueden reportarse valores engañosos (falta de exactitud) al medir las longitudes o por el cometimiento de errores sistemáticos al transformar un valor de circunferencia a diámetro. Es preferible elaborar un calibrador de madera (o cualquier otro material) que facilite la medición directamente en el caso de frutos relativamente grandes como las sandías.

Longitud y ancho de hojas

Estas variables se miden en plantas cuyo producto comercial son justamente sus hojas, tales como el tabaco, el nabo o la acelga. Son variables cuantitativas continuas, que se las considera como respuesta a ensayos de nutrientes, estimulantes, distancias de siembra, estudios de adaptabilidad y de caracterización de variedades. Salvo para estudios en donde se evalúe el comportamiento del crecimiento foliar en el tiempo, las mediciones de estas variables deben realizarse en el momento de la cosecha.

Para su determinación se utilizan flexómetros, reportando los resultados en centímetros, hasta en un decimal redondeado. En cada unidad de muestreo pueden seleccionarse aleatoriamente al menos unas 30 hojas en un mínimo de 10 plantas, definiendo y considerando estrictamente la zona (parte alta, media o baja) de la planta desde donde se extraerán las hojas.

Macollamiento

El macollamiento es una variable que se mide en cereales, tales como el arroz o el trigo, es de tipo cuantitativa discreta y se la emplea en estudios sobre adaptabilidad de variedades y/o híbridos, de distanciamientos de siembra, de nutrición, de mejoramiento genético

y en experimentos sobre láminas de riego. Consiste en cuantificar la cantidad de macollos que se forman en una planta durante su desarrollo que, en el caso del arroz, cesa hasta la aparición de los primordios florales. Esta formación de macollos se indica en la figura 4.2, en donde se verifica el incremento de éstos durante el desarrollo de una planta de trigo.

Para el reporte del dato, la recomendación es seleccionar aleatoriamente dentro de la unidad de muestreo un mínimo de 30 plantas en la etapa de floración, determinándose el promedio de macollos por planta; forma de medición que, en el caso del arroz, es aplicable cuando la siembra es directa (al voleo). Bajo este tipo de siembra, cabe mencionar el error que se comete en muchos estudios al reportar esta variable como macollos por m^2 , por cuanto no es posible en la práctica tener una distribución homogénea de las semillas por unidad de área cuando éstas son distribuidas en el área de siembra.

En la siembra directa o al voleo, hay una importante probabilidad de tener un experimento cargado de un elevado error experimental al evaluar, por ejemplo, la variable más importante, como lo es el rendimiento. Como esta variable se define a partir de la cosecha que se obtiene en la unidad de muestreo, este dato podría presentar una fuerte variabilidad, aun en un mismo tratamiento, debido a lo disperejo que pudiéramos distribuir la semilla al volear la misma.

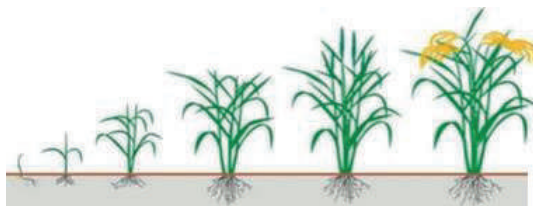


Figura 4.2. Esquema de la formación de macollos en un cereal como el trigo.

Para contrarrestar estos efectos introducidos hay que ser muy cuidadosos al momento de la siembra, distribuyendo las semillas en las parcelas lo más homogéneamente posible y utilizando las mismas cantidades (siempre que el factor de estudio no sea densidad de siembra) en cada unidad experimental. También será recomendable obviar la determinación de la variable rendimiento a partir de pocas unidades cuadradas de área dentro de la unidad de muestreo (por ejemplo, $1 m^2$), siendo preferible reportar el dato a partir de la cosecha obtenida sobre toda la superficie de dicha unidad de muestreo.

Cuando la siembra de las gramíneas es indirecta (por trasplante) deberá tenerse el cuidado suficiente al momento de trasplantar el cultivo al sitio definitivo, haciendo que ésta sea homogénea; es decir,

se coloquen en cada sitio de siembra la misma cantidad de plántulas. En este sentido, otra alternativa para reportar esta variable puede ser como número de macollos por sitio de siembra; o también reportar este dato por unidad de área, pero considerando que las dimensiones de dicha área sean proporcionales a las distancias de siembra, de tal forma que en cada unidad de muestreo se encierre la misma cantidad de sitios o puntos de trasplante.

Número de panículas por planta

El número de panículas por planta es una variable cuantitativa-discreta que se evalúa en algunos cereales como el arroz. Junto con la cuantificación de macollos por planta, permite deducir el número de macollos efectivos (macollos con panículas) por planta. Es una variable que se mide en ensayos de cruce genético, de adaptación de variedades y/o híbridos, de distanciamientos de siembra o de densidad por unidad de área y en ensayos de láminas de riego.

El procedimiento para valorar esta variable es similar al de macollamiento. Deben seleccionarse aleatoriamente un mínimo de 30 plantas dentro de cada unidad de muestreo y el dato se lo reportará como un promedio, redondeado a un decimal. En siembras por trasplante esta variable deberá reportarse como el número de panículas por sitio de siembra, para lo cual también deberán seleccionarse al azar un mínimo de 10 sitios dentro de la unidad de muestreo, siempre que, al momento de la siembra (trasplante), se coloque la misma cantidad de plántulas en cada sitio.

Longitud y peso de raíces

La longitud y el peso radicular de un cultivo son variables que, consecuentemente, deben medirse en experimentos sobre productos que favorecen el crecimiento radicular, sobre ciertos fertilizantes como el fósforo o en ensayos sobre métodos de preparación de suelos.

Son variables de tipo cuantitativa continua y para su determinación se pueden utilizar flexómetros hasta una décima de centímetro (longitud) y balanzas de precisión hasta la décima de gramo (masa). El dato de reporte es el promedio de un mínimo de 10 plantas seleccionadas aleatoriamente de cada unidad de muestreo cuando se haya completado el desarrollo fisiológico del cultivo (entre floración y cosecha).

En este grupo también se incluyen los ensayos con ciertas fitohormonas que se utilizan para reproducción asexual de plantas, utilizando esquejes o por injerto; en donde la variable que se mide es la eficacia de los tratamientos, determinando la proporción de plantas enraizadas.

Nótese que estas variables son de índole destructiva. Es decir, deberá extraerse la planta para poder realizar su medición; por ello, debe planificarse adecuadamente el momento y/o sitio de la unidad

experimental desde donde se extraerán, de tal forma que no se afecte la medición de otras variables como, por ejemplo, el rendimiento.

Fecundación

Los ensayos para medir la fecundación involucran experimentos sobre métodos de polinización, de inductores florales y de técnicas de poda. Es una variable cuantitativa discreta, que se expresa en forma porcentual y se ajusta a la distribución binomial.

La fecundación es una variable que se mide en experimentos relacionados generalmente a plantas frutales, como la guanábana, el maracuyá, el mango, el cacao, entre otros. Por esta causa, la distribución que comúnmente se utiliza en el ensayo es el cuadro latino, en donde cada unidad experimental es una planta.

Para la determinación de la fecundación previamente se realiza una selección aleatoria de flores en el árbol, o seleccionando tramos de ramas o tronco en donde se verifica cantidades importantes de yemas florales, luego de lo cual se marcan y se aplican los tratamientos. El experimento puede tener una duración que va desde el estado floral hasta la cosecha del fruto o hasta cierto tiempo de formación de éste, de tal forma que se haya producido la fecundación. La proporción de la fecundación se establece con la expresión:

$$F = \frac{100F_F}{F_I} \quad (4.1)$$

donde F es la fecundación (%), F_I es la cantidad de flores seleccionadas en cada unidad experimental al inicio del ensayo y a las cuales se les aplicarán los tratamientos; y F_F es la cantidad de flores, de las seleccionadas, que fecundaron o terminaron en fruto.

Vigor vegetativo

El vigor vegetativo es una variable cualitativa de característica ordinal (categórica) que se mide en estudios sobre fertilizantes foliares o de fitohormonas, tanto en plántulas de vivero como en campo. Complementariamente, también se determina en ensayos sobre productos que favorecen el crecimiento radicular o en experimentos sobre el control de nematodos.

Es una variable que depende del sujeto, el cual utiliza el sentido de la vista para hacer la valoración. En esta determinación se tiene de apoyo escalas estandarizadas o arbitrarias (algunas son gráficas-visuales) que traducen a la forma cuantitativa la valoración cualitativa del vigor.

Días a la floración/días a la cosecha

Para determinar la precocidad de un cultivo se utiliza la variable de los días que tarde éste en llegar a la etapa de floración o de cosecha desde la siembra. Esta es una variable de tipo cuantitativa discreta y se la utiliza principalmente en estudios de cruce genético, de adaptabilidad de variedades y/o híbridos, en técnicas de poda y en estudios sobre bioestimulantes con contenidos hormonales.

Es una variable que requiere mucho cuidado si lo que se quiere es objetividad, ya que la verificación es diaria cuando el cultivo está próximo a estas etapas. Para la contabilización de los días hasta que el cultivo llega a la floración o cosecha, deberá considerarse que el 50% de las plantas de cada unidad de muestreo tengan flores totalmente abiertas o tengan sus frutos listos para la cosecha, respectivamente.

Número de granos por panícula

Es una variable cuantitativa-discreta, medible en cereales, cuyo producto de interés se forma en panículas. Se requiere en estudios sobre cruce genético, adaptación de variedades y/o híbridos y de distanciamientos de siembra. Para su determinación deberán seleccionarse al azar un mínimo de 10 panículas, reportándose el dato como un promedio, redondeado a un decimal.

Peso de 100/1000 semillas

Esta variable es de tipo cuantitativa continua y se recomienda su valoración en estudios sobre adaptación de variedades y/o de híbridos, estudios sobre cruce genético, ensayos sobre nutrición o en estudios sobre plagas, como complemento en la valoración de semillas vanas. Para su determinación se utilizan balanzas con precisiones de hasta una décima de gramo. Se consideran 100 semillas cuando éstas son relativamente grandes como, por ejemplo, las de fréjol o de maíz; y 1000 semillas en granos pequeños como los de arroz o cebada.

Para una mayor certeza, es recomendable que esta variable se mida al menos por triplicado en cada unidad de muestreo. Su medición podría ser tediosa en granos pequeños como el arroz, debido a que se deben contar aleatoriamente 1000 granos en cada evaluación; no obstante, para acelerar el proceso es preferible tarar envases plásticos de volúmenes reducidos, de tal forma que se facilite la rápida medición con sólo llenar hasta la señal de tara.

Debido a que el peso se lo determina en estado aparentemente seco, siempre será necesario realizar la corrección respectiva por los distintos contenidos de humedad que se pueden encontrar en las muestras, ajustándolo a contenidos de humedad uniforme, que pueden ser de un 12 o de un 14%. Una vez que se haya determinado la humedad de cada muestra, mediante equipos para medir la humedad en granos, el ajuste se realizará empleando la expresión siguiente:

$$P_{aj} = \frac{P_m(100 - H_0)}{100 - H_1} \quad (4.2)$$

donde P_{aj} es el peso ajustado de los 100 o 1000 granos expresado en gramos, P_m es el peso de la muestra (g), H_0 es la humedad real o instantánea de la muestra y H_1 es la humedad de ajuste o deseada (12 o 14%). Como es lógico suponer, las mediciones del peso y del contenido de humedad deberán determinarse en un mismo instante, con un máximo de diferencia de tiempo de 24 horas entre muestras.

Granos/vainas vanas

La variable de granos vanos se mide en cultivos como el arroz y otros cereales, dado que se expresa en forma porcentual se la considera cuantitativa-continua. Se valora en estudios de distanciamientos de siembra o de densidad, en cruces genéticos, en los de adaptación de variedades y/o híbridos, en ensayos de fertilización y en estudios sobre plagas. Su determinación es un poco tediosa por cuanto, en el caso de cereales, deben seleccionarse aleatoriamente 1000 semillas en cada unidad de muestreo, a partir de las cuáles habrá que separar las semillas vanas y reportar el resultado proporcionalmente.

La variable de vainas vanas se valora también en cultivos leguminosos como los de fréjol, arvejas, habas o los de maní. En estos casos, se seleccionan al azar un mínimo de 100 vainas para la verificación de las que están vanas.

Índices o ratios

Relación grano/tusa

Esta variable se la mide en el cultivo de maíz y corresponde a la relación del peso de los granos de mazorcas “secas” con respecto al peso de las tusas de las mismas mazorcas. Es una variable cuantitativa-continua y adimensional, que se evalúa especialmente en estudios sobre cruces genéticos, de adaptación de variedades y/o híbridos, en caracterización de semillas y a veces en algunos estudios sobre plagas que atacan las mazorcas.

El valor de esta relación permite conocer la proporción de los granos con respecto a la de las tusas. Estos valores pueden oscilar entre un mínimo de 3 a un máximo de 6, obteniéndose aproximaciones a este último valor en los híbridos de maíz más promisorios.

Índice de mazorca en cacao

Este índice corresponde a la relación del número de mazorcas de cacao necesarias para obtener 1.0 kg de grano seco. Es una variable cuantitativa continua, utilizada especialmente para valoración de cruces genéticos, de adaptabilidad de variedades o híbridos, de dosis de nutrientes y de caracterización.

Cajas/racimo

Esta es una variable de tipo cuantitativa continua, utilizada para expresar la producción comercial del cultivo de banano. Se contabiliza el número cajas (expresado en fracción) de 43,0 libras en peso neto que se obtiene por racimo, con características comerciales de que cada banano tenga una longitud máxima de 18,0 cm y un calibre (diámetro) de entre 39 a 46 mm. Es un tipo de variable medida, principalmente, en estudios de nutrición, de control de plagas, de riego y de manejo del cultivo.

Rendimiento

El rendimiento es una variable de tipo cuantitativa continua que se expresa en unidades de peso por unidad de área (a veces por planta como en algunos frutales). Ya sea como variable principal o complementaria, es la más importante para todos los ensayos que involucran producción. Las unidades de peso que se consideran pueden estar dadas en kilogramos (kg) o en toneladas métricas (t), referidas a un área de 1 ha. La unidad de kg/ha se utiliza en cultivos cuyo rendimiento máximo es de unos 10000 kg/ha, mientras que en cultivos con producciones arriba de este valor, se sugiere la unidad de t/ha. Para la primera unidad, deberá utilizarse un decimal redondeado; mientras que, para la segunda, se recomienda entre 2 a 3 decimales redondeados.

Para el reporte de esta variable, en cultivos de ciclo corto o de distanciamientos de siembra reducidos, se recomienda que se coseche toda el área de la unidad de muestreo de cada unidad experimental, establecida en la delimitación del ensayo; luego este dato (peso de la cosecha) deberá extrapolarse a la unidad de 1 ha.

En ciertos casos no es una evaluación adecuada la obtención del rendimiento a partir de cierto número de plantas que se seleccionan aleatoriamente dentro de cada unidad de muestreo. Por ejemplo, en experimentos sobre plagas pueden no existir algunas plantas debido al ataque de éstas. En este caso, al no considerarse toda el área de la unidad de muestreo, no se reflejará en el dato de rendimiento la falta de algunas plantas o de la parte cosechable que es afectada por la plaga; lo que viene a ser el indicativo de que un tratamiento determinado no ha funcionado o su efecto fue parcialmente eficaz.

En otras situaciones experimentales como, por ejemplo, los ensayos con fertilizantes, es recomendable obtener el rendimiento a partir de un determinado número de plantas (entre 10 a 30) que se seleccionan aleatoriamente dentro de la unidad de muestreo de la parcela. Con esta forma de evaluación se evitará (o se corregirá) el efecto de la falta de plantas dentro de la unidad experimental debido a una mala siembra o un mal manejo, o incluso cierta deficiencia germinativa de la semilla; lo que, en el caso de experiencias con nutrientes, pasarían a constituir una variable de ruido en la respuesta de los tratamientos. Además, en

este tipo de experiencias lo que se busca es valorar la respuesta de una planta, de allí la necesidad de que ella esté presente.

Otra de las consideraciones que debe tenerse en cuenta al momento de valorar el rendimiento es que, en todos los cultivos de granos que se cosechan “secos”, sus datos deberán ser ajustados o corregidos por el contenido de humedad que tengan al momento de su recolección. Esta corrección por humedad, al igual como se explicó en el peso de 100 o 1000 semillas, se realiza con la expresión 4.2.

Para los cultivos en los cuales se tienen múltiples cosechas o en algunos perennes, como el cacao, en donde permanentemente se hace la recolección, para el reporte deberá considerarse el total de al menos tres cosechas consecutivas, de tal forma que el dato tenga consistencia y el efecto de los tratamientos sea mejor evaluado.

En esta variable también se incluyen las de rendimiento efectivo que consideran el producto comercial del total cosechado. En estos casos pueden mencionarse la producción de cajas/ha en banano o la producción de azúcar obtenida de las toneladas de caña cosechada en una hectárea.

Tal como se ha dado a entender, en los experimentos agrícolas por lo general se tienen una o más variables que se relacionan directamente con el efecto de los tratamientos; y una o más variables que se miden complementariamente, que respaldan a las primeras, de tal forma que se pueda tener una mayor certidumbre en la conclusión que se deduzca de la experiencia. No obstante, es criterio del investigador identificar plenamente las variables que considera pertinente medirlas en su experimento, teniendo en cuenta lo que es comparable y coherentemente medible, que al mismo tiempo permitan cumplir con los objetivos de la investigación y corroborar la (s) hipótesis planteada; todo a la luz de una base teórica-empírica consistente.

¿Qué es la unidad de muestreo?

En la descripción de algunas de las variables de respuesta que se miden en los experimentos agrícolas de los párrafos anteriores se ha mencionado recurrentemente el término *unidad de muestreo*, haciendo alusión al espacio desde donde se seleccionan las unidades de observación para las mediciones correspondientes. En este sentido, en lo que sigue se hace una distinción con respecto al criterio “consuetudinario” de *parcela útil*, ampliamente utilizado en los textos de diseño experimental agrícola.

En los experimentos de campo, en los cuales se utilizan distribuciones como la de bloques completos al azar, uno de los elementos de precaución muy importante es el llamado *efecto de borde*. Este efecto es definido de esa manera debido al traslape que podría suceder entre tratamientos que se aplican en parcelas contiguas; dado que, en el sentido de la conveniencia, estas parcelas dentro de cada bloque no se separan con espacios intermedios, resultando algún efecto de cruce entre los tratamientos en los lados laterales de las mismas. Asimismo, en las cabeceras de dichas parcelas también podría reportarse algún efecto de confusión del tratamiento debido a que las plantas, si bien están separadas en bloques diferentes, compiten con otro tipo de plantas (generalmente malezas), que podrían interactuar con el cultivo de estudio.

Para evitar estos efectos se dejan hileras de plantas a los costados y en las cabeceras de cada parcela al momento de medir las variables, separando en la parte central de estas unidades

experimentales un área con cierto número de plantas a partir de las cuales se tomarán los datos y que muchos autores conocen como *área útil de parcela* o simplemente como *parcela útil*. Un término que no es una definición adecuada, dado que cada unidad cuadrada de superficie de todo un ensayo (incluidos los espacios que separan bloques) es útil; por lo tanto, debido a que desde este espacio dentro de cada unidad experimental se extraerán los datos de todas las variables que se hayan previsto en la evaluación, su definición más acorde sería *unidad de muestreo*. Al respecto, en la figura 4.3 se ha indicado en el recuadro pintado de celeste la unidad de muestreo de una parcela que contiene 4 hileras de plantas; para lo cual el efecto de borde está representado por una hilera lateral a cada lado de la unidad experimental y dos plantas en sus cabeceras.

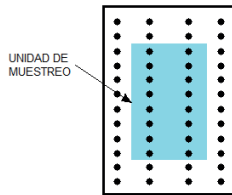


Figura 4.3. Detalle de la unidad de muestreo dentro de una unidad experimental (parcela).

También es necesaria la distinción entre unidad de muestreo y *unidad de observación*. Esta última es la planta o elemento sobre la cual se mide la variable de interés, de varias que se ubican dentro de la unidad de muestreo. Es decir, al final, el dato reportado por unidad experimental viene a ser una media de las unidades de observación que se consideraron en una unidad de muestreo. En algunos casos, especialmente cuando se trata de cultivos perennes o frutales, de distanciamientos amplios como el mango, la unidad de observación es igual a la unidad de muestreo; lo cual implica que una planta es una unidad experimental.

Error experimental y el ajuste de datos

A pesar de que los experimentos se llevan a cabo bajo condiciones controladas, siempre existirá una variabilidad que no se pueda explicar entre las unidades experimentales de cada tratamiento (variación dentro de los grupos); es decir, siempre se tendrán diferencias entre los datos que provienen de las parcelas en donde se aplica un mismo tratamiento. Esta variación es conocida como *error aleatorio* y se supone que es invariable en cada unidad experimental; estadísticamente definido bajo distribución normal, con media cero y varianza constante $N(0, \sigma^2)$. Sin embargo, cuando el experimento no es bien diseñado y/o manejado, el investigador adiciona fuentes de variación secundarias que causan otro tipo de error que ya no obedecen a un comportamiento aleatorio. En este sentido, todo experimento contiene un error experimental que inherentemente corresponderá al error aleatorio, al cual podrá adicionarse un error no aleatorio (ocasionado por el propio investigador, conocido también como error de manejo o sistemático). Es decir, en un ensayo con diseño y/o manejo defectuoso, se tendrá que:

$$\text{Error experimental} = \text{error aleatorio} + \text{error no aleatorio}$$

El éxito de un experimento es que el error no aleatorio sea cero, lo cual se logra con rigor científico obtenido de un conocimiento teórico, aplicando un extremo cuidado en el diseño y

en el desarrollo del ensayo; todo lo cual, consecuentemente, conferirá la debida precisión²⁵ en la determinación de efectos entre tratamientos. En un experimento adecuadamente planeado y manejado, es decir, amparado bajo el método científico, el error aleatorio afectará sólo a la precisión del ensayo; mientras que los errores sistemáticos o no aleatorios, no sólo afectarán a la precisión, sino también a la exactitud, lo cual invalida los resultados de cualquier prueba.

Fundamentalmente y siempre que se utilice una semilla certificada, el error experimental dependerá exclusivamente de dos grandes factores: 1) el ambiente bajo el cual se desarrolla el ensayo y 2) el manejo del experimento. Por ello, si se realiza una debida selección de un material genético homogéneo y escogiendo una adecuada distribución experimental en campo, es lógico suponer que si los datos de una variable presentan un elevado error experimental (una gran variación no explicada), será por un mal manejo que se hizo del ensayo, entendiéndose también dentro del mal manejo alguna forma errónea en el proceso de medición de las variables de respuesta. Esta última indicación debe cumplirse en experimentos en donde se evalúan factores inermes como, por ejemplo, los fertilizantes. Cuando un ensayo busca solucionar problemas de plagas (ataques de insectos, nematodos, enfermedades y otros), la información generalmente presenta una importante variabilidad atribuida al error experimental, aun teniendo un riguroso cuidado del manejo del experimento.

La utilidad del coeficiente de determinación y del coeficiente de variación en la cuantificación del error experimental

La verificación cuantitativa de cómo la influencia del ambiente y la expresividad genética de una semilla (variación genotipo x ambiente), así como un mal manejo del ensayo y/o una mala selección de los niveles (especialmente los cuantitativos) de los tratamientos, afectan a un experimento, puede observarse mediante el coeficiente de determinación (r^2). Este coeficiente cuantifica la variabilidad explicada o conocida generada por los tratamientos o por las fuentes de variación que se incluyen en los modelos de ANDEVA de un ensayo; es decir, relaciona la variabilidad explicada por el modelo, respecto de la variación total de los datos. En este caso, utiliza las sumas de cuadrados del total de datos (SC_T) y del error experimental (SC_E) para plantear la expresión 4.3 indicada a continuación:

$$r^2 = \frac{SC_T - SC_E}{SC_T} \quad (4.3)$$

En términos numéricos, los límites de r^2 están dados entre 0 y 1, siendo aceptables (también confiables) para un experimento valores $\geq 0,6$. Según la expresión 4.3, la variabilidad del modelo, comprendida por $SC_T - SC_E$, incluyen a todas las fuentes de variación que se desglosan en el ANDEVA. Así, por ejemplo, en una distribución completamente al azar (DCA), $SC_T - SC_E$ es la variabilidad que producen los tratamientos expresada a través de su suma de cuadrados.

El límite (inferior) de 0,6 para r^2 como valoración de aceptabilidad de un experimento debe verse como un elemento de juicio de su precisión; pues en el análisis de modelos lineales múltiples, cuando se trata del ajuste de éstos y se eliminan variables independientes que no aportan significativamente a un modelo dado, dentro de los que se obtienen con carácter exploratorio, se exige que este valor sea $\geq 0,7$. En este caso, debe obtenerse el coeficiente de

²⁵ La precisión experimental es distinta de la exactitud. La primera se controla con el cuidado en el desarrollo del ensayo (manejo) y está ligada inversamente a la varianza desconocida; mientras la segunda depende de la planificación o diseño del experimento, especialmente, del diseño de tratamientos. Lo último permite coleccionar que la exactitud queda establecida previo al inicio de la fase experimental, aunque podría darse el caso de la presencia de variables de ruido cuyos efectos ocurran en la etapa de desarrollo del experimento. Por lo mencionado, se puede decir que el error experimental sólo juzga la precisión más no la exactitud.

determinación ajustado (r_{aj}^2), similar a la expresión 4.3, pero utilizando los cuadrados medios (varianzas) del total de los datos (CM_T) y del error experimental (CM_E).

La valoración que hace r^2 de la bondad de un experimento la podemos observar, por ejemplo, en los ensayos sobre fertilizantes, ya que, si un investigador selecciona niveles de algún macronutriente poco equidistantes, es de esperarse que entre ellos no exista variación significativa ($p < 0,05$); de allí que el coeficiente de determinación tienda a cero, indicándonos que los tratamientos no han sido bien seleccionados. En este caso la variabilidad que presentará el experimento estará referido mayoritariamente al error experimental ocasionado por un mal diseño del experimento, debido a una falta de base teórica y/o empírica (experiencia de un experto o resultado de experimento preliminar) que debió tomarse en cuenta al inicio de la experiencia.

Si bien el coeficiente de determinación nos ofrece una visión más o menos integral de las posibles causas de falta de precisión de un ensayo, gran parte de los biometristas, especialmente los del área agrícola, basan la bondad de sus ensayos en el coeficiente de variación (CV). Este coeficiente nos ofrece una visión general del manejo o cuidado que se ha realizado en un ensayo, valorado a través de la proporción de la raíz cuadrada de la varianza del error experimental (CM_E) respecto de la media general de los datos (\bar{Y}), lo cual se plantea mediante la expresión 4.4. El CV es una medida relativa de la dispersión de toda la información de una variable determinada.

$$CV = \frac{100\sqrt{CM_E}}{\bar{Y}} \quad (4.4)$$

Aunque los dos coeficientes valoran la magnitud del error experimental y ofrecen conjuntamente un panorama sobre la bondad de un experimento; en ciertos casos, el r^2 puede ser más efectivo²⁶ que el CV para calificar dicha bondad. Se afirma esta comparación en el sentido de que el r^2 considera los desvíos totales sin ser afectados por los grados de libertad del error experimental; situación que es distinta para el CV , con el cual puede enmascarse la variabilidad del error experimental dada por su suma de cuadrados que incluso puede ser mucho más grande que las otras fuentes conocidas en el modelo de ANDEVA. Este engaño puede ser palpable en experimentos que tienen un número elevado de repeticiones como, por ejemplo, en una prueba sensorial. En una prueba sensorial de un alimento de característica hedónica, las repeticiones, representadas por los jueces, deben estar arriba de 30 personas; quienes, por su evaluación subjetiva, ocasionarán que el error experimental se incremente sustancialmente, no obstante, será minimizado por la importante cantidad de grados de libertad que en estos casos se generan (el postulado del teorema del límite central). En todo caso, es obligación del biometrista hacer una verificación intuitiva de lo que realmente está ocasionando la falta precisión en un experimento determinado.

El investigador debe estar consciente de lo que representa la suma de cuadrados en los modelos de ANDEVA: la variabilidad “original” de cada una de las fuentes del modelo. Es decir, observando la magnitud de estas variaciones podemos darnos cuenta, en el caso de

²⁶ Esta efectividad del r^2 ocurre por el engaño que puede encubrirse en el CV cuando un experimento tiene un gran número de grados de libertad en el error experimental. En esta situación la varianza intragrupos, o también llamada cuadrado medio del error experimental, disminuirá excesivamente, lo que podría ocasionar una falsa valoración haciéndonos creer que la variabilidad que presenta el error experimental es menor que la variabilidad que producen los tratamientos. En estas circunstancias, es posible que se tenga un CV con un valor aparente por debajo de los límites aceptables (menos del 30% en experimentos de campo o menos del 10% en experimentos de laboratorio), dado que considera al cuadrado medio del error; mientras que el r^2 indica una posición contraria, presentando un valor no aceptable para los experimentos (menos de 0,6). Un caso como este podría favorecer (en términos de aplicación práctica), incluso, el cometimiento del error tipo I, ya que existiría una gran probabilidad latente de que se estaría declarando efectos significativos en donde realmente (en la práctica) no existen.

experimentos mal diseñados y/o mal manejados, quién tiene los mayores efectos (tratamientos, bloques o el error experimental). El aumento de repeticiones no va a disminuir el error experimental de un ensayo leído como suma de cuadrados, lo que realmente hacen es aumentar la precisión matemática (no la precisión práctica) de las estimaciones; pues, como puede verse a través del error estándar de las medias, éste se reduce de forma proporcional a la raíz cuadrada del tamaño muestral (σ/\sqrt{n}). Bajo esta apreciación, considerando que el r^2 se determina con las sumas de cuadrados y no con los cuadrados medios como lo hace el CV, se ha manifestado cierta supremacía del primero de los nombrados para verificar la bondad de un experimento.

Por lo que ya se ha manifestado sobre el error experimental, cuando éste es exclusivo del error aleatorio, no debe interpretarse como lo que sugiere su denominación; es decir como un error propiamente dicho, sino como lo que ya se ha indicado: hace referencia a toda variación externa al experimento que no es explicada por los tratamientos. Lo ideal en los ensayos es que el error experimental sólo sea producto del error aleatorio; de hecho, para los cálculos matemáticos del análisis de varianza obligadamente debe ser mayor a cero, caso contrario no será posible determinar el valor del estadígrafo de contraste F de Fisher. En cualquier experimento es imposible eliminar totalmente el error experimental, lo máximo que podemos hacer es reducir su efecto hasta que éste sólo corresponda a la casualidad, de manera que el resultado y diferencia o efecto de los tratamientos evaluados sean evidenciados claramente; de allí que el propósito del diseño experimental sea el de controlar o reducir este error. Se puede eliminar, o cuando menos minimizar, la variación correspondiente al error no aleatorio en un ensayo si se consideran las siguientes indicaciones:

- Delimitación homogénea del área experimental (igual tamaño y forma de parcela, selección de suelo en condiciones más o menos uniformes de pendientes y libres de variaciones importantes de fertilidad).
- Manejo de las unidades experimentales lo más homogéneamente posible (aplicación del riego, aplicación de fitosanitarios, aplicación de los tratamientos, etc.).
- El uso de borde de parcela o de una adecuada unidad de muestreo que evite el traslape parcial de los tratamientos o de especies. Esto es, un adecuado tamaño de parcela o unidad experimental.
- Uso y ubicación correcta de una distribución experimental, de tal modo que se separen las potenciales fuentes de variación secundarias que podrían afectar al experimento.
- Utilizar un apropiado e igual número de repeticiones para cada tratamiento.
- Siempre aleatorizar la asignación de los tratamientos a las unidades experimentales (o viceversa), de tal forma que se reduzca la probabilidad de caer en el error tipo I.
- Aplicación homogénea de los tratamientos.
- Utilización de las pruebas estadísticas que puedan separar las potenciales causas de variación identificadas.
- Toma de datos bajo un criterio uniforme. Cuando se utilice más de un equipo, es imprescindible que éstos estén debidamente calibrados y reporten un mismo resultado.
- Durante la toma de datos en ensayos relativamente grandes, cuando la disponibilidad de tiempo, personal y/o recursos no permite culminar la medición en un mismo día, no se deben dejar bloques inconclusos.
- Cuidado permanente de las cosechas. Existen casos, especialmente en los experimentos que se realizan con frutales, en donde se requiere guardianía; ya que una vez que sean visibles los frutos, éstos pueden ser sustraídos.
- Verificar posibles errores en la tabulación de datos. Es decir, suspicacia en datos que muestren una dispersión importante dentro de los grupos o tratamientos.

La magnitud del error experimental permite una visión integral sobre el éxito de un experimento, el cual como se habrá deducido de los párrafos que anteceden, tiene una

dependencia mayoritaria en el investigador más que en el propio experimento *per se*. En forma general, lo que debe tener siempre presente todo investigador es que, cuanto menor sea el error experimental, más certero serán los resultados y más fácil será determinar diferencias reales entre los tratamientos evaluados, especialmente si dichas diferencias son pequeñas. Es decir, un menor error experimental, favorecido por un control eficaz de esta fuente, hará que se consiga una mayor precisión en los ensayos.

Importancia de los supuestos de normalidad, homocedasticidad e independencia en la precisión experimental

Otro problema que podemos encontrar por la falta de precisión en los experimentos, desde el punto de vista estadístico y como hecho colateral, es la situación en donde los residuales de los datos de las distintas variables evaluadas no cumplen con las condiciones de que se ajusten a una distribución normal, de que sean homocedásticos (esto es, los tratamientos deben tener varianzas estadísticamente constantes) y que sean independientes. El cumplimiento de estos supuestos son los fundamentos en los que reside la efectividad de las pruebas paramétricas como el análisis de varianza, una de las herramientas estadísticas clásicas utilizadas para determinar la existencia de diferencias significativas entre tratamientos. En palabras de Montgomery (2005), “no es prudente confiar en una herramienta estadística como el ANDEVA sin antes haber verificado estos supuestos”; ya que, tanto el grado de significancia como la potencia estadística de una prueba paramétrica se ven afectados por el no cumplimiento de estos tres supuestos. En un experimento, por ejemplo, la falta normalidad en los residuos comúnmente causará una pérdida de potencia del ANDEVA; esto es, una falta de precisión del ensayo que restará “sensibilidad” para detectar diferencias significativas entre los tratamientos.

Aun con todo el cuidado en el manejo de un ensayo y cumpliendo los tres principios básicos de los diseños experimentales, como son la aleatorización, el bloqueo y la repetición, en los ensayos agronómicos existen algunas variables que comúnmente no se ajustan a los supuestos de normalidad y de homocedasticidad. Generalmente esto sucede en variables relacionadas a organismos vivos como los insectos o los hongos, evaluación de daño en cultivos por estos organismos, en variables de conteo como el de frutos por planta o en evaluaciones de malezas; las cuales más bien tienden a ajustarse a distribuciones como la binomial o la de Poisson en muchos de los casos.

Para comprobar la normalidad, la homocedasticidad y la independencia se utilizan los errores o residuales que representan a la variabilidad aleatoria de cada dato y que, por tanto, son los valores más indicados para verificar el cumplimiento de estos supuestos; especialmente en herramientas estadísticas como el ANDEVA. Esta verificación se puede realizar mediante una descripción gráfica o alguna prueba analítica utilizando los residuales. De estas dos opciones de comprobación, la más práctica es la descripción gráfica, la misma que a pesar de ser subjetiva, se la utiliza con mayor frecuencia. Una prueba analítica requiere un conjunto de datos relativamente grande para que tenga la suficiente potencia; mientras que una gráfica, con pocos datos ya permite visualizar el comportamiento de la información (cuando éstos muestran un ajuste inobjetable), de allí que esta sea una de las ventajas que ofrece la comprobación gráfica.

Una ventaja importante que ofrece el análisis gráfico de la normalidad de los residuos es poder observar *valores atípicos*²⁷, conocidos también como *puntuaciones outliers*. Estos valores,

²⁷ Una forma de identificar los puntos o valores atípicos (*outliers*) en un conjunto de datos es estandarizando el o los residuales que se manifiestan ser marcadamente distintos del resto de residuales. Esta estandarización se puede realizar con la expresión: $e_s = e_{ij} / \sqrt{CM_E}$, siendo e_s el residual estandarizado y CM_E el cuadrado medio del error

como se explicará más adelante, se presentan distanciados del resto de puntos de forma evidente. Es importante observar las puntuaciones *outliers*, dado que, con sólo tener uno de ellos entre los datos en ensayos pequeños, es posible que se produzca una distorsión grave en los resultados de un ANDEVA. La presencia de estos valores atípicos deberá ser seriamente investigada, pues su origen podría no deberse a un error sistemático en el experimento.

En cada uno de los supuestos se elabora una gráfica respectiva; es decir, se elabora una para la normalidad, una para la homogeneidad de las varianzas y una para representar la independencia de los valores residuales. Los residuos se obtienen de las diferencias entre el dato observado (Y_{ij}) y la media respectiva de dichos datos en cada grupo o tratamiento (\bar{Y}_i) correspondiente, según sea el tipo de distribución experimental utilizado. Estos valores, que también se conocen como errores (\hat{e}_{ij}), deben ser estimados considerando el modelo estadístico que se esté utilizando; así por ejemplo, si la distribución es completamente al azar, los residuales están dados por: $\hat{e}_{ij} = Y_{ij} - \bar{Y}_i$. La media \bar{Y}_i de un tratamiento i para un modelo dado de ANDEVA (en este caso de un DCA) es en realidad el estimador de un valor predicho de una función lineal (\hat{Y}_i); por ello es que en modelos de regresión lineal, estos errores se calculen como: $\hat{e}_{ij} = Y_{ij} - \hat{Y}_i$ (el circunflejo en la letra Y indica que son medidas muestrales).

Para la comprobación de los supuestos de normalidad, homocedasticidad e independencia mediante el uso de pruebas analíticas, existen varias alternativas. Así se tienen, por ejemplo, la prueba Chi cuadrado, la de Smirnov-Kolmogorov o la de Shapiro-Wilks, para el análisis de normalidad de los datos. Para comprobar la homogeneidad de las varianzas se puede utilizar la prueba de Barlett, o también la prueba de Levene. Para analizar la independencia de los residuales puede utilizarse la prueba de Durbin-Watson.

De los tres supuestos mencionados, a nivel agrícola, la falta de cumplimiento de la independencia de los residuos es el más problemático; ya que, por la misma heterogeneidad del suelo, parcelas contiguas tenderán a correlacionarse positivamente cuando existan gradientes de fertilidad, humedad u otras potenciales fuentes de ruido que podrían afectar a las respuestas de los tratamientos, produciendo una subestimación del error experimental y un mayor riesgo de cometer el error tipo I. No obstante, con el cumplimiento del principio de aleatorización y acciones necesarias como el uso de bordes en las unidades experimentales, queda sin efecto la incidencia de la falta de independencia de los residuales. En consecuencia, la verificación debe realizarse con mayor atención hacia el cumplimiento de normalidad y homocedasticidad; sin embargo, este rasero no debe ser excesivamente estricto dado que el ANDEVA, o más precisamente el estadístico F de Fisher, es apenas afectado por la falta de normalidad o de homocedasticidad cuando se trata de modelos fijos y experimentos balanceados (cada tratamiento con igual número de repeticiones); debido a ello es que se dice que el ANDEVA es una herramienta estadística *robusta*.

Como una forma ilustrativa se expone a continuación un ejemplo en donde se verifican los supuestos de normalidad y homocedasticidad (igualdad de varianzas) en forma gráfica y analítica, utilizando Excel y el programa *IBM SPSS Statistics 20.0*.

experimental. Por lo tanto, bajo el teorema del límite central en donde los e_{ij} deben tener media cero y varianza unitaria $N(0, \sigma^2)$, el 68% de los e_{ij} deberá estar dentro de los límites de ± 1 , el 95% de los e_{ij} deberá ubicarse entre el límite ± 2 y prácticamente todos los e_{ij} deberán caer en los límites ± 3 . Si existen residuales e_{ij} mayores a 3, éstos serán considerados como valores atípicos.

Ejemplo de aplicación 4.1

En un experimento con plantas de fréjol se probaron cuatro dosis de un biol, con un testigo de referencia. Cada uno de los tratamientos se valoró con cuatro repeticiones, utilizando una distribución de bloques completos aleatorizados. La variable evaluada fue la altura de plantas (cm), cuyos datos son los que se indican en la tabla 4.1. Previo a la aplicación del análisis de varianza, la verificación del cumplimiento de los datos respecto de los supuestos de normalidad y homocedasticidad es lo que se detalla a continuación:

Tabla 4.1. Datos de altura de planta (cm).

N°	Tratamientos	Repeticiones				Total tratamientos
		I	II	III	IV	($\sum Y_j$)
1	Biol 1,0 l/ha	20,8	26,8	28,2	25,2	101,0
2	Biol 2,0 l/ha	21,1	25,4	22,8	20,8	90,1
3	Biol 3,0 l/ha	22,4	21,8	24,4	23,2	91,8
4	Biol 4,0 l/ha	26,4	26,8	25,6	27,4	106,2
5	Testigo	17,3	18,5	19,2	20,2	75,2
Total repeticiones ($\sum Y_i$)		108,0	119,3	120,2	116,8	464,3

Comprobación gráfica de la normalidad

El problema de la falta de normalidad, así como de la ausencia de homocedasticidad de un grupo de datos, es la causa para que las inferencias que se hacen con pruebas de la estadística paramétrica pierdan precisión al tener una menor potencia, así como un ligero incremento en la probabilidad de caer en el error tipo I (alfa). No obstante y de acuerdo a lo ya indicado, el análisis de varianza es un test robusto ante una ligera falta de normalidad, ya que la prueba *F* que se utiliza es afectada levemente por el incumplimiento moderado de este supuesto, de allí que su afectación sobre la inferencia respecto de las medias es relativamente pequeño. Esta última situación puede ser diferente con el cumplimiento de la homocedasticidad de los errores, ya que ante este principio existe una mayor “sensibilidad” en la precisión de este estadígrafo, especialmente cuando el experimento es desbalanceado. En todo caso, se vuelve a recalcar que, para prevenir estos problemas, es fundamental que se cumplan adecuadamente los tres principios básicos que rigen para todo experimento: la repetición, la aleatorización y el bloque.

Las gráficas que permiten visualizar si una información estadística se ajusta a una distribución normal se pueden desarrollar utilizando un papel con las dos escalas, tanto las *X* como las *Y*, en distancias aritméticas o constantes; es decir, un papel común y corriente. También es factible utilizar un papel que tiene una de las escalas, y generalmente las ordenadas, en distancias logarítmicas, el cual se conoce como papel normal y se encuentra en cualquier librería o se aplica directamente a los datos mediante Microsoft Excel.

Para graficar los datos que provienen de un experimento se puede utilizar el papel con escalas aritméticas, en donde el eje de las *X* se usa para ubicar los errores o residuos en forma creciente y el eje de las *Y* para ubicar la probabilidad acumulada de ocurrencia del residuo²⁸. Si los puntos siguen o sugieren explícitamente la descripción de una línea recta, esta es una evidencia suficiente para corroborar que los datos tienen aproximadamente una distribución normal. La precisión del análisis de varianza puede verse afectado cuando uno o más residuos se

²⁸ Esta es una de las formas de realizar estas gráficas. Existen otras como la que se elabora con los cuantiles estandarizados (valores *Z*), obtenidos de la probabilidad de ocurrencia acumulada de los datos ordenados ascendientemente utilizando la expresión $P_{X \leq X_T} = (n_i - 0,5)/n$. La gráfica se elabora colocando los cuantiles en el eje *X* y los datos observados (ordenados ascendientemente) en el eje *Y*.

encuentran distanciados de manera importante de la alineación del resto en una gráfica de normalidad. Cuando se verifica uno de estos datos atípicos u *outliers*²⁹ el investigador deberá dedicar la atención suficiente, ya que muchas veces y especialmente cuando la variable debe ser de un comportamiento normal, existen errores de tabulación; si este no es el caso, deberá tener la reserva necesaria incluso como para desechar este dato y realizar el análisis con el resto.

Los pasos para la elaboración de la gráfica de normalidad que se siguen en este texto son los siguientes:

1. Calcular las probabilidades iguales o menores a cada dato u observación ($P_{(X \leq X_T)}$) con la expresión:

$$P_{(X \leq X_T)} = 100[n_i / (n + 1)] \quad (4.5)$$

2. Obtener los residuos, restando de los datos observados los promedios de cada tratamiento (los predichos) de manera correspondiente siempre que el modelo de ANDEVA sea de una distribución completamente al azar. En el caso de una distribución de bloques completos al azar, los valores predichos (P) se obtienen con la siguiente expresión:

$$P = \frac{\sum Y_i \sum Y_j}{\sum Y_{ij}} \quad (4.6)$$

Donde $\sum Y_i$ e $\sum Y_j$ son las sumatorias de los datos que se encuentran en la columna i y en la fila j , respectivamente; e $\sum Y_{ij}$ es la sumatoria total.

3. Ordenar los residuos ascendentemente.
4. Proceder a ubicar los puntos con los residuos ordenados en el eje de las X y las probabilidades respectivas en el eje de las Y .

Los valores predichos con referencia a los datos de la tabla 4.1, obtenidos con la expresión 4.6, son los siguientes:

Para el primer y último valores se tiene:

$$P_1 = \frac{108 \times 101}{464,3} = 23,5$$

$$P_{20} = \frac{116,8 \times 75,2}{464,3} = 18,9$$

Tabla 4.2. Datos predichos o esperados.

N°	Tratamientos	Repeticiones			
		I	II	III	IV
1	Biol 1,0 l/ha	23,5	26,0	26,1	25,4
2	Biol 2,0 l/ha	21,0	23,2	23,3	22,7
3	Biol 3,0 l/ha	21,4	23,6	23,8	23,1
4	Biol 4,0 l/ha	24,7	27,3	27,5	26,7
5	Testigo	17,5	19,3	19,5	18,9

²⁹ Una herramienta gráfica y clásica para verificar (entre sus aplicaciones) los *outliers* es el diagrama de cajas (*box-plot*).

Considerando los pasos antes indicados, en la tabla 4.3 se tienen las probabilidades y los residuales ordenados ascendentemente:

Tabla 4.3. Valores calculados para la elaboración de los gráficos de normalidad y homocedasticidad.

N.º	Tratamientos	Valores observados	Valores predichos	Probabilidad $P_{(X \leq x_T)}$	Residuales \hat{e}_i	Residuales \hat{e}_i ordenados ascendentemente
1	Biol 1,0 L/ha	20,8	23,5	4,76	-2,7	-2,7
2		26,8	26,0	9,52	0,8	-1,9
3		28,2	26,1	14,29	2,1	-1,9
4		25,2	25,4	19,05	-0,2	-1,8
5	Biol 2,0 L/ha	21,1	21,0	23,81	0,1	-0,8
6		25,4	23,2	28,57	2,2	-0,5
7		22,8	23,3	33,33	-0,5	-0,5
8		20,8	22,7	38,10	-1,9	-0,3
9	Biol 3,0 L/ha	22,4	21,4	42,86	1,0	-0,2
10		21,8	23,6	47,62	-1,8	-0,2
11		24,4	23,8	52,38	0,6	0,1
12		23,2	23,1	57,14	0,1	0,1
13	Biol 4,0 L/ha	26,4	24,7	61,90	1,7	0,6
14		26,8	27,3	66,67	-0,5	0,7
15		25,6	27,5	71,43	-1,9	0,8
16		27,4	26,7	76,19	0,7	1,0
17	Testigo	17,3	17,5	80,95	-0,2	1,3
18		18,5	19,3	85,71	-0,8	1,7
19		19,2	19,5	90,48	-0,3	2,1
20		20,2	18,9	95,24	1,3	2,2

Tomando como eje X los residuales y como el eje Y a la probabilidad $P_{(X \leq x_T)}$, el gráfico para verificar la normalidad de los datos es el que se indica en la figura 4.4. En esta figura, puede notarse que la distribución de los residuales tiende a describir un ajuste lineal de los datos, sin mostrar alguna dispersión fuerte de algún punto. Por lo tanto, puede asumirse que los datos de altura de planta dados en la tabla 4.1 cumplen con el supuesto de normalidad.

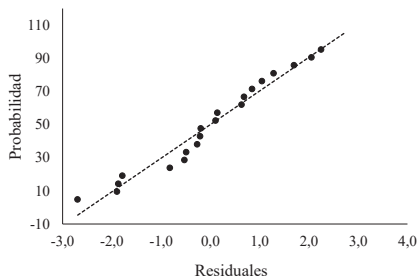


Figura 4.4. Gráfica de probabilidad normal.

Comprobación gráfica de la homocedasticidad

La potencia del análisis de varianza y de las demás pruebas estadísticas paramétricas dependen, en gran magnitud, de la homogeneidad de las varianzas de las muestras que representan a los diferentes conjuntos de tratamientos de un experimento. Esta situación es específica en el caso de los ensayos agronómicos, en donde es característica de algunas variables presentar heterogeneidad en las varianzas de los grupos o tratamientos, especialmente de las que están relacionadas a organismos como insectos, fitopatógenos y demás microorganismos inmersos en la producción de cultivos. Si existe el rigor científico necesario en un experimento, el supuesto de la homocedasticidad pasa a ser la principal propiedad que el investigador no debe dejar de verificar, si lo que pretende es que el modelo de análisis de varianza que está utilizando como herramienta estadística tenga la precisión necesaria.

La falta de igualdad o de similitud entre las varianzas de los tratamientos dentro de un estudio con plantas, fundamentalmente se debe a dos situaciones. La primera tiene que ver con el material experimental; si éste es altamente heterogéneo, la dispersión de los datos en cada grupo es de tal magnitud que incluso puede llevarnos a cometer el error tipo I. Un caso de material heterogéneo en un ensayo puede ser el uso de semilla que no es certificada, en donde la mezcla de genotipos podría confundir el efecto real de los tratamientos. La segunda situación está relacionada a la falta de cuidado en el manejo del ensayo, que puede hacer que variables que no tienen razón para presentar dispersiones importantes, se presenten como heterocedásticas. En estas condiciones, un factor no menos importante es la selección adecuada de la distribución experimental, que permita bloquear potenciales fuentes secundarias que también pueden influir para que las varianzas no sean estadísticamente homogéneas.

La precisión de los experimentos agrícolas es altamente “sensible” a la variabilidad del ambiente y del manejo de los mismos, tal vez con una intensidad mayor al de los experimentos de otras áreas; la cual puede ser verificada mediante el coeficiente de variación (CV), cuyo valor podría darnos una idea de la falta de igualdad en las varianzas de los tratamientos. En este sentido, el CV es uno de los estadísticos que se utiliza en los ensayos agronómicos como herramienta para calificar su precisión; no obstante que, debido a los múltiples factores secundarios que inciden en el proceso experimental, es esperable que las variables de respuesta presenten cierta magnitud en sus varianzas que puedan impedir que se mantengan, en términos estadísticos, como homocedásticas.

Dado que el CV de un experimento es la relación porcentual del cuadrado medio del error experimental (varianza) respecto del promedio de los datos, es un buen indicador general de la dispersión que muestran éstos, que a su vez nos alertan de la falta de varianzas homogéneas que tienen los tratamientos y que, por consiguiente, se constituye en un elemento que explica la falta de eficacia del ANDEVA. Sin embargo, según lo antes indicado, en un experimento con un elevado número de repeticiones se prorratea la variabilidad de los datos de un tratamiento dado al dividirse por los grados de libertad del error experimental, presentando un valor engañoso respecto de la falta de homocedasticidad; ante lo cual, el coeficiente de determinación (r^2) permite un mejor panorama.

La comprobación del supuesto de homocedasticidad, al igual que en la comprobación de la normalidad, también se puede realizar mediante gráficas o mediante fórmulas analíticas. El método gráfico de verificación del supuesto de igualdad de varianzas consiste, generalmente, en plotear los predichos, ubicándolos en el eje X ; contra los residuos o errores, que van en el eje Y (\hat{Y}_{ij} vs \hat{e}_{ij}). Una vez obtenido el diagrama de dispersión, si los puntos describen aproximadamente una franja de éstos en forma horizontal, aleatoria y sin ningún patrón definido (sin estructura), es la condición necesaria para concluir que los tratamientos cumplen

con el supuesto de igualdad de varianzas. En contraposición, si la dispersión de los residuos es fuerte mostrando una clara distorsión de la franja hacia una forma de “embudo”, entonces es un indicativo de que la información no se ajusta al supuesto de homocedasticidad. Esta forma de embudo se debe a que la media tiene una relación directa con la varianza, en donde el aumento de la una afecta a la otra; o también, mediante una relación inversa: el aumento de una produce una reducción de la otra.

En el ejemplo que se ha considerado, los valores predichos y los residuos están calculados en la tabla 4.3. En la gráfica correspondiente de la figura 4.5 se observa la dispersión de los residuos, los cuales no muestran un patrón definido; de allí que los datos de altura de planta que están tabulados en la tabla 4.1 pueden considerarse como homocedásticos.

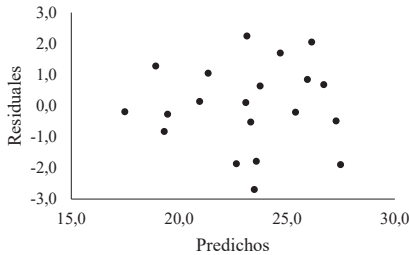


Figura 4.5. Gráfica para verificar homocedasticidad.

Al utilizar el programa SPSS, se puede verificar el cumplimiento de normalidad y de homocedasticidad al momento de realizar el análisis de varianza de los datos. Para esta observación la secuencia es la que se indica a continuación:

Analizar → Modelo lineal general → Univariante

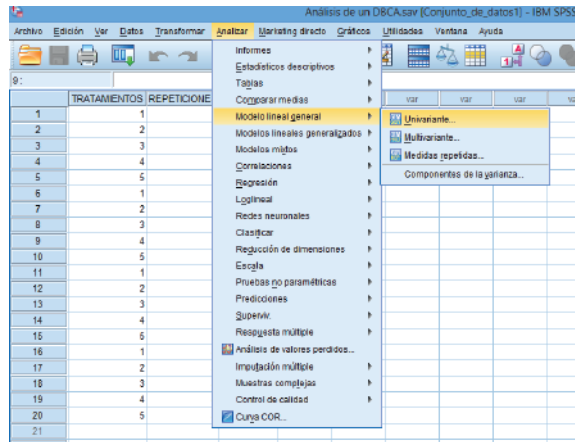


Figura 4.6. Secuencia en el software SPSS para realizar el análisis de varianza y la verificación gráfica de la normalidad y homocedasticidad.

Definiendo el modelo de una distribución de bloques completos al azar y seleccionando en el submenú la opción de gráfico de residuos, se reporta el análisis de varianza (tabla 4.4) junto con las gráficas correspondientes en donde se puede observar el cumplimiento de los supuestos de normalidad y homocedasticidad (figura 4.7).

Tabla 4.4. Análisis de varianza realizado en SPSS bajo un modelo de bloques completos al azar.

Variable dependiente: Altura de planta (cm)

Origen	Suma de cuadrados tipo III	GL	Media cuadrática	F	Sig.
Modelo corregido	159,837 ^a	7	22,834	7,699	0,001
Intersección	10778,725	1	10778,725	3634,503	0,000
Tratamientos	141,208	4	35,302	11,904	0,000
Repeticiones	18,629	3	6,210	2,094	0,155
Error experimental	35,588	12	2,966		
Total	10974,150	20			
Total corregida	195,425	19			

a. r cuadrado = 0,818 (r cuadrado corregido = 0,712)

Según puede observarse en la figura 4.7, al relacionar los datos observados con los de pronóstico (valores predichos), se destaca cierta tendencia hacia una línea recta; por lo cual se pueda afirmar que, aproximadamente, existe un comportamiento normal de la variable. Para verificar la homocedasticidad, en esta figura también se puede notar la relación entre los valores de pronóstico y los residuales, en donde al igual que lo sucedido en la figura 4.5, no

se nota un patrón definido; consiguientemente, puede decirse que los datos presentan varianzas homogéneas.

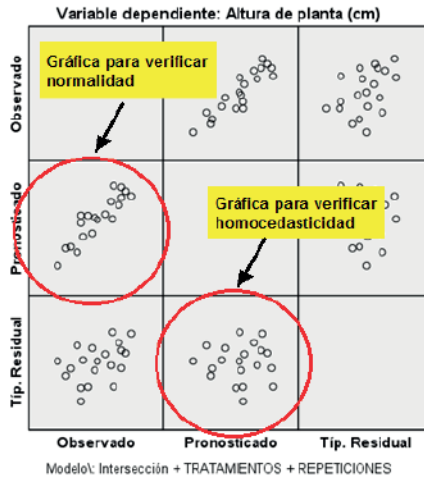


Figura 4.7. Gráfica de residuos reportado por el software SPSS para verificar normalidad y homocedasticidad.

Analíticamente, mediante el software SPSS también es posible verificar la normalidad y la homocedasticidad de datos. Para comprobar la normalidad la secuencia en el programa es:

Analizar → Estadísticos descriptivos → Explorar

En la figura 4.8 se presenta la secuencia para realizar el análisis de normalidad. Para este análisis, el programa utiliza el test de Smirnov – Kolmogorov indicado para tamaños muestrales arriba de los 50 datos; y el test de Shapiro – Wilk, que se utiliza en tamaños muestrales menores. En este ejemplo, considerando la última prueba mencionada, el software reporta que cada grupo (tratamiento) de datos tiene un comportamiento normal³⁰ (tabla 4.5) dada la significancia establecida, cuyo valor es mayor al 5% ($p > 0,05$). Esta significancia permite aceptar la hipótesis nula de que los datos se ajustan a una distribución normal.

³⁰ Debe entenderse que los análisis para verificar los supuestos no necesitan hacerse para los grupos (tratamientos), sino de forma general para todos los datos; lo cual es factible debido a que se realizan estos procesos a partir de los residuos y no de los datos originales. En este ejemplo, se realizó el análisis por grupo (tratamiento), únicamente con el sentido de verificar en cuál de ellos se hubiese ubicado la falta de normalidad.

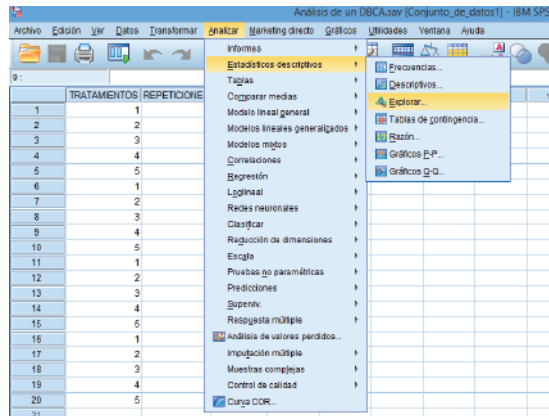


Figura 4.8. Secuencia en el software SPSS para realizar el análisis de normalidad con las pruebas de Kolmogorov – Smirnov y Shapiro – Wilk.

Tabla 4.5. Análisis de normalidad realizado en SPSS.

		Pruebas de normalidad					
		Kolmogorov - Smirnov ^a			Shapiro - Wilk		
	Tratamientos	Estadístico	GL	Sig.	Estadístico	GL	Sig.
Altura de planta (cm)	Biol 1,0 L/ha	0,244	4	.	0,927	4	0,575
	Biol 2,0 L/ha	0,250	4	.	0,889	4	0,377
	Biol 3,0 L/ha	0,188	4	.	0,973	4	0,858
	Biol 4,0 L/ha	0,171	4	.	0,994	4	0,976
	Testigo	0,153	4	.	0,998	4	0,993

a. Corrección de la significación de Lilliefors.

Para la igualdad de varianzas, la hipótesis nula a comprobar es que los datos del ejemplo en cuestión son homocedásticos. Para este caso, la secuencia en el software es:

Analizar → Comparar medias → ANOVA de un factor

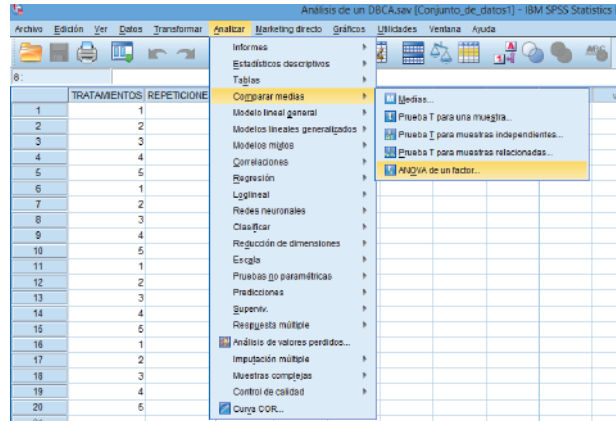


Figura 4.9. Secuencia en el software SPSS para realizar el análisis de homocedasticidad mediante el test de Levene.

Una vez que se cumple la secuencia indicada en la figura 4.9, se deberá seleccionar en el submenú opciones la prueba de homogeneidad de las varianzas. En este caso, el software utiliza el test de Levene para realizar este análisis; una prueba que es robusta a la falta de normalidad, muy indicada para utilizarse a nivel experimental. El reporte para el ejemplo que se ha venido explicando es el siguiente:

Tabla 4.6. Reporte del análisis de homocedasticidad de SPSS.

Prueba de homogeneidad de varianzas			
Altura de planta (cm)			
Estadístico de Levene	G _L 1	G _L 2	Sig.
1,684	4	15	0,206

Según lo indicado en la tabla 4.6, la significancia es mayor al 5% ($p > 0,05$), por lo tanto debe aceptarse la hipótesis nula; es decir, los datos son homocedásticos y existe confianza en lo reportado por el análisis de varianza.

Ajuste o estabilización de la varianza intragrupos

En la agronomía existen variables cuyo comportamiento no se ajusta a distribuciones normales y tienen por naturaleza ser heterocedásticas, lo que precisamente no se debe a que la selección del material experimental y/o el manejo del ensayo no hayan sido los adecuados. Son variables

que intrínsecamente tienen tendencia a la dispersión y que generalmente se utilizan para medir efectos relacionados con organismos móviles como los insectos u obedecen a procesos estocásticos como el ataque de hongos, las cuales pueden ajustarse a distribuciones de Poisson, binomiales o extremas como la exponencial. En estos casos, una de las soluciones para poder seguir utilizando una herramienta estadística paramétrica importante como el análisis de varianza es ajustar (cambio de escala) la variable, de manera que la dispersión se reduzca y se cumplan aproximadamente los supuestos de normalidad y homocedasticidad. Si luego de este ajuste la dispersión persiste, el investigador puede optar por utilizar pruebas estadísticas equivalentes, pero de tipo no paramétricas, las cuáles no requieren del cumplimiento de los supuestos mencionados; o utilizar modelos lineales generalizados que tampoco tienen la obligación de cumplir con el criterio de normalidad u homocedasticidad, debido a que se basan en otros tipos de distribuciones diferentes a la normal y con supuestos menos restrictivos. También podría utilizarse el test de Welch para la comparación de medias, dado que es una prueba para grupos heterocedásticos, aunque requiere normalidad en los residuos.

Si bien el *CV* cuantifica de forma general la dispersión con que puede presentarse una variable, también sirve para calificar el manejo del ensayo, por cuanto una falta de normalidad y/o de varianzas homogéneas también es causada por la falta de cuidado en el desarrollo del experimento. El investigador que está relacionado con experimentos agrícolas debe tener una idea de cuáles son las variables que no deben presentar dispersiones importantes y cuáles sí podrían hacerlo; de tal forma que no realice el ajuste de una variable en donde la dispersión relativamente fuerte (error experimental de importancia) no es naturaleza de la misma, sino debido a un mal manejo del ensayo, lo cual sería una causal para no confiar en los resultados y no hacer válido el ensayo. A este respecto, en la tabla 4.7 se detallan algunas variables que deben ajustarse a distribuciones normales y otras que generalmente no lo hacen.

Tabla 4.7. Variables con y sin ajuste a distribuciones normales.

Variables con ajuste a una distribución normal	Variables que comúnmente no cumplen el supuesto de normalidad
Altura de planta	Nematodos por 100 g de suelo
Peso de materia seca	Nematodos por 100 g de raíces
Diámetro del tallo	Incidencia de una enfermedad
Longitud del fruto	Presencia de insectos
Diámetro de fruto	Ataque de un insecto
Peso de 1000 semillas	Control de malezas*
Peso de 100 semillas	
Días a floración	
Días a la cosecha	
Número de frutos por planta*	
Longitud de raíces	
Macollamiento*	
Número de panículas por planta	
Granos/vainas vanas	
Rendimiento	

*Variables cuyos errores pueden o no ajustarse a la distribución normal.

Debido a que el valor del *CV* de un experimento depende, además del manejo, también del ambiente en donde se desarrolla, y dado que el análisis de varianza es una prueba relativamente robusta, pueden admitirse ciertos límites. Si el ensayo se desarrolla en condiciones de laboratorio, es decir en medios y ambientes controlados, y la variable cuya expresión endógena debe ajustarse a una distribución normal, se recomienda como máximo un *CV* del 10%. En estas circunstancias pueden ubicarse los experimentos de reproducción de plantas *in vitro*, los cuáles se realizan en laboratorios apropiados; los que se realizan con macetas, en donde el

suelo es homogeneizado; o estudios que se realizan en viveros e invernaderos. En el caso de los experimentos que se realizan a nivel de campo y siempre que la variable tenga por naturaleza expresar normalidad, podrían permitirse *CV*s de hasta un 30%; sin embargo, en el caso de variables como las indicadas en el lado derecho de la tabla 4.7, las mismas que generalmente no muestran comportamientos de normalidad y de homocedasticidad, no se puede definir un límite de aceptación del resultado de dichas variables debido a su naturaleza intrínseca de presentar dispersiones de importancia.

Muy aparte de lo que pueda admitir el investigador respecto de la heterocedasticidad que presente una variable a través de su *CV*, lo que si debe considerarse de manera estricta es la afectación de la precisión del análisis de varianza cuando los datos registran coeficientes de variación relativamente altos. Si bien un ANDEVA tiene cierto grado de robustez, el test *F* que utiliza esta herramienta estadística incrementa su falta de precisión cuando el error experimental de un ensayo se vuelve trascendental. En este sentido, si el *CV* de una variable rebasa el 30%, la falta de normalidad y de homocedasticidad se hace evidente, y mucho más si el tamaño muestral es reducido. En estas circunstancias y considerando que la naturaleza inherente de la variable sea la de presentar dispersión, la alternativa más práctica para poder seguir confiando en la precisión del ANDEVA es realizar la transformación o ajuste de los datos de la variable.

La transformación de datos de una variable de respuesta, también llamado *método de estabilización de varianzas*, se realiza con el propósito de que la dispersión que causa la heterocedasticidad se minimice hasta lograr que los tratamientos muestren varianzas homogéneas; consecuentemente esto ocasionará que los valores se ajusten más a una distribución normal y que exista una mayor correspondencia con el modelo estadístico. En este caso, el ajuste es una forma de controlar virtualmente el error experimental; ya que permite que, herramientas como el análisis de varianza, tengan una mayor precisión y se evite caer fácilmente en el error β .

Para optar por alguna alternativa de transformación es necesario observar el tipo de relación que tiene la varianza con la media de una variable, lo que se define proporcionalmente de acuerdo al crecimiento exponencial de esta última. De allí que este ajuste puede ser más o menos restrictivo (más o menos fuerte), dependiendo de que los datos muestren una mayor o menor dispersión, que puede apoyarse al observar el *CV* de los mismos. Una de las gráficas que también permite definir el tipo de ajuste es la que se realiza para verificar la homocedasticidad, en donde se utilizan los residuales y los predichos (figura 4.5). Si en la dispersión que describen los residuales en una gráfica de comprobación de varianzas homogéneas es claramente visible una forma de “corneta”, el ajuste debe ser relativamente más fuerte. De acuerdo al tipo de relación entre la media y la varianza, en la tabla 4.8 se sugieren algunos tipos de transformaciones.

Tabla 4.8. Transformaciones sugeridas para estabilización de varianzas.

Relación proporcional entre la varianza y la media	Tipo de transformación
σ^2 constante	No se requiere transformación.
$\sigma^2 \propto \mu q$	Angular o arcoseno (siendo q la probabilidad de falla).
$\sigma^2 \propto \mu$	Raíz cuadrado de los datos.
$\sigma^2 \propto \mu^2$	Logaritmo de los datos.
$\sigma^2 \propto \mu^3$	Recíproco de la raíz cuadrada de los datos.
$\sigma^2 \propto \mu^4$	Recíproco de los datos

Entre las transformaciones más comunes en el análisis de datos de experimentos agrícolas están las de raíz cuadrada, la logarítmica y la angular; indicada esta última para valores porcentuales.

Transformación a raíz cuadrada

Esta transformación está indicada cuando la información comprende números enteros (pequeños y positivos) relacionada con variables de conteo que se miden en el espacio o en el tiempo como, por ejemplo, el número de panículas por m² en arroz de siembra al voleo. En este caso los datos describen aproximadamente una distribución de Poisson y existe una relación proporcional entre las varianzas y las medias de los tratamientos, debido a lo cual puede ocurrir una falta de normalidad y homocedasticidad de la variable que incide en el grado de precisión de herramientas de la estadística paramétrica como el análisis de varianza. Esta anomalía se puede corregir extrayendo la raíz cuadrada de cada uno de los datos antes de someterlos al ANDEVA. Para este propósito tenemos dos casos:

$Y' = \sqrt{Y}$ Siempre que la matriz de datos contenga valores iguales o mayores a 10.

$Y' = \sqrt{Y + 0,5}$ Si la matriz contiene valores menores a 10 e incluyendo el cero.

Esta es una transformación que también puede aplicarse en valores dados en porcentaje, siempre que tengan una variación entre 0 a 20% o de 80 a 100%. En el caso del segundo intervalo, para poder aplicar la transformación habrá que restar de 100 cada uno de los datos.

Transformación logarítmica

Esta alternativa es aplicable cuando las varianzas son proporcionales al cuadrado de las medias de los tratamientos, produciéndose una falta de independencia de los datos al relacionarse medias mayores con varianzas también mayores. En otras palabras, debe aplicarse en rangos de datos positivos muy amplios. Un ejemplo de esta variable es la velocidad de infiltración de un suelo, cuyo rango a través del tiempo, hará que se tengan valores desde menos de 1 hasta valores mayores a 100. Otro ejemplo puede ser el conteo de nematodos en raíces, situación que podría involucrar desde unos pocos e incluso cero, hasta miles de individuos en cada muestra. Aquí también se pueden incluir variables relacionadas con insectos y de microorganismos, variables cuya dispersión puede ajustarse a una distribución extrema como la exponencial. Para la transformación de los datos en estos casos, tenemos las alternativas:

$Y' = \log Y$ Para matrices de datos que no contienen ceros.

$Y' = \log(Y + 1)$ Para matrices de datos que contienen ceros.

$Y' = \log(Y + 10)$ Para matrices de datos que contienen ceros y la dispersión es extrema.

Transformación angular o arcoseno

Este tipo de transformación se recomienda cuando los datos describen una distribución binomial, están expresados en fracciones o porcentajes, con una dispersión de relativa importancia. Está indicada cuando la variabilidad de los datos puede ir desde 0 hasta 30%, o desde 70 a 100%. Un ejemplo para este caso es la medición del control de malezas producto

de la aplicación de varias dosis de algún herbicida, en los cuales se tendrán valores porcentuales bajos y altos. Las opciones de transformación al respecto son las siguientes:

$$Y' = \text{Sen}^{-1}\sqrt{Y/100} \quad Y = \% \quad \text{Transformación a grados.}$$

$$Y' = \frac{\pi \text{Sen}^{-1}\sqrt{Y/100}}{90} \quad Y = \% \quad \text{Transformación a radianes.}$$

Cuando se realiza un análisis de varianza con los datos transformados es importante que se mencione que la información ha sido modificada, especialmente cuando el ajuste permite detectar un cambio en el *p*-valor del estadístico *F* que establezca una significancia entre los tratamientos menor al 5%. Esto hará que quien utilice estos resultados tenga las reservas del caso al momento de aplicar el tratamiento ganador.

En todo caso hay que estar conscientes que las transformaciones no eliminan la falta de normalidad que la información inherentemente o por error inadvertido tiene. Lo que sí se consigue es una mejor estimación de los efectos de los tratamientos al lograr una mayor potencia en el análisis de varianza. En todo caso, se vuelve a recalcar que es potestad del investigador aplicar alguna transformación para estabilizar las varianzas, pues es él quien se supone conoce la naturaleza del material con el que está experimentando y no muchas veces el estadístico; de allí que, variables agrícolas como el rendimiento de un cultivo no necesitan de estas transformaciones porque insoslayablemente deben presentar normalidad, caso contrario, es un mal manejo del ensayo que no justifica ningún ajuste de datos. Al final, previo a la interpretación de los resultados de un análisis estadístico con datos transformados, los estimadores (medias, límites de intervalos u otros) deben ser reconvertidos, pues sus valores estarán sesgados sino se realiza este proceso. A continuación, se detalla un ejemplo de aplicación respecto de datos transformados.

Ejemplo de aplicación 4.2

En un ensayo sobre *pyrazosulfuron ethyl* se han probado cuatro dosis en el control de malezas de hoja angosta, cuya variable se ha medido a los 10 días después de la aplicación (dda). Para su ejecución se empleó una distribución de bloques completos al azar, en donde cada tratamiento se evaluó con cinco repeticiones. Los datos son los que se detallan en la tabla 4.9.

Tabla 4.9. Datos de control de malezas a los 10dda (%).

N°	Tratamientos	Repeticiones					Promedios
		I	II	III	IV	V	
1	15 gramos i.a./ha	82	76	69	35	77	67,8
2	20 gramos i.a./ha	76	81	73	74	79	76,6
3	25 gramos i.a./ha	92	88	94	82	81	87,4
4	30 gramos i.a./ha	32	93	91	83	78	75,4
Promedio general							76,8

Al realizar el análisis de varianza con estos datos se tiene lo siguiente:

Tabla 4.10. Análisis de varianza con los datos sin transformar.

Fuentes de variación	GL	SC	CM	F	P-valor
Total	19	5089,20			
Tratamientos	3	976,80	325,60	1,17	0,3603
Repeticiones	4	784,70	196,17	0,71	0,6021
Error experimental	12	3327,70	277,31		

Utilizando la varianza del error experimental (cuadrado medio del error experimental) obtenido en el ejemplo, según la expresión 4.4, el coeficiente de variación de los datos sin ningún ajuste es el siguiente:

$$CV = \frac{100 \times \sqrt{277,31}}{76,8} = 21,68\%$$

De acuerdo al tipo de dato y a la variabilidad presentada, medido a través del coeficiente de variación ($CV = 21,68\%$), el ajuste más indicado en este caso es el angular o arcoseno. Para aplicar esta transformación puede recurrirse a algún texto sobre diseño de experimentos, en donde se registren estos valores transformados en forma tabular. Sin embargo, dada la gran cantidad de programas informáticos sobre estadística, también puede recurrirse a uno de ellos para realizar este ajuste. Asimismo, con Excel también es factible esta transformación, considerando el conversor de $180/\pi$ por cuanto el resultado que se obtiene con esta herramienta es en radianes.

En este ejemplo, utilizando Excel, la función que hace la transformación es:

$$= \text{ASENO}[\text{RAIZ}(\text{CELDA}/100)] * 180/\pi$$

Aplicando esta función, los datos ajustados son:

Tabla 4.11. Datos de control de malezas a los 10dda con transformación arcoseno.

N°	Tratamientos	Repeticiones					Promedios
		I	II	III	IV	V	
1	15 gramos i.a./ha	64,9	60,7	56,2	36,3	61,3	55,9
2	20 gramos i.a./ha	60,7	64,2	58,7	59,3	62,7	61,1
3	25 gramos i.a./ha	73,6	69,7	75,8	64,9	64,2	69,6
4	30 gramos i.a./ha	34,4	74,7	72,5	65,6	62,0	61,9
Promedio general							62,1

Con los datos ajustados mediante la transformación arcoseno de la tabla 4.11, el nuevo análisis de varianza es el siguiente:

Tabla 4.12. Análisis de varianza con los datos transformados.

Fuentes de variación	GL	SC	CM	F	P-valor
Total	19	2172,68			
Tratamientos	3	483,18	161,06	1,43	0,2812
Repeticiones	4	342,65	85,66	0,76	0,5690
Error experimental	12	1346,86	112,24		

Con el nuevo análisis de varianza realizado con los datos ajustados, el coeficiente de variación resultante es:

$$CV = \frac{100 \times \sqrt{112,23}}{62,1} = 17,06\%$$

Con este valor puede observarse como la variabilidad de los datos se ha reducido, pasando de un 21,68% a 17,06%; esto quiere decir que la dispersión de los errores se ha reducido, haciendo que el análisis de varianza gane precisión al obtenerse inicialmente un valor de probabilidad (*P-valor*) de 0,36 cuando los datos no tienen ajuste angular, y luego de éste, esta probabilidad es de 0,28.

Otra alternativa de precisión en los experimentos cuando una variable revela un alto *CV* podría ser la realización de un análisis estadístico considerando alguna variable independiente secundaria o de ruido, adicional a la variable independiente de estudio y la dependiente (respuesta), que deben ser caracteres medibles en los cultivos y que no se controlan directamente con el diseño experimental. En estas circunstancias se tiene una covariable o variable concomitante, que obliga a realizar lo que se conoce como *análisis de covarianza*, de la cual se trata en el siguiente tema.

Análisis de covarianza

El análisis de covarianza (ANCOVA) es un procedimiento estadístico que se realiza con el propósito de corregir los efectos de los tratamientos de un experimento, debido a la presencia de una o más variables de ruido que pueden tergiversar la respuesta real de los mismos. El investigador deberá tener la capacidad para identificar estas variables, conocidas comúnmente como *covariables*, y que podrían tener una influencia sobre la variable de respuesta, ya que sus efectos son diferentes al causado por los tratamientos y no son parte del error experimental aleatorio. En esta identificación es importante la experiencia del investigador, pues las covariables no son susceptibles de control ya sea con el manejo o con la distribución experimental, sino sólo son medibles.

El ANCOVA permite mejorar la precisión de un ensayo y hace más efectivo al análisis de varianza, por cuanto minimiza el error experimental a su condición de aleatoriedad; lo que permite a su vez también realizar un ajuste de los promedios de los tratamientos, por cuanto el efecto de estos últimos en alguna medida está relacionado con el efecto de las covariables. Esta prueba es un análisis que tiene su importancia cuando se ha identificado plenamente la influencia de variables secundarias (o covariables) que actúan en forma concomitante junto a la variable independiente de estudio, sobre la respuesta de un experimento; no obstante, es una herramienta que, por razones de aparente complejidad, generalmente no se la aplica. Sin embargo, debido al actual desarrollo de programas computacionales, el procedimiento para realizar un ANCOVA es menos engorroso.

En un ANCOVA se consideran dos o más variables independientes y una dependiente, en donde el análisis facilita verificar, además del efecto de la variable independiente de estudio (tratamientos), si alguna(s) variable de ruido, que también deberá ser medida, tiene una influencia significativa sobre la variable dependiente (respuesta). Si esta influencia es significativa, el investigador deberá hacer la corrección de los promedios de los tratamientos separando mediante un modelo de regresión lineal, el efecto de la variable de ruido que viene a constituirse en la covariable del experimento; de allí que un ANCOVA en la práctica,

involucra al análisis de varianza y a la regresión lineal. Sin embargo, aun cuando no exista efecto significativo de la variable concomitante sobre la dependiente, siempre será mejor tener promedios ajustados; por ello la consideración inicial de una covariable no es una previsión que nos ha hecho perder tiempo si es que no existieran efectos “reales” en términos estadísticos.

Ejemplos sobre incidencia de covariables en el ámbito de los ensayos agrícolas hay muchos. Uno de estos ejemplos puede ser un experimento en donde se evalúen varias dosis de nitrógeno (variable independiente de estudio) en arroz sembrado al voleo y la variable de respuesta sea el rendimiento. En este caso puede existir una influencia de distintos números de macollos por m^2 que podrían ocurrir en algunas unidades experimentales debido a la no uniformidad con que se distribuyó la semilla en cada parcela al momento de la siembra, constituyéndose el macollamiento como una covariable. Otro ejemplo puede ser un ensayo sobre inductores florales de aplicación foliar (variable independiente) en cacao, donde la variable de respuesta sea el número de frutos por planta; teniéndose como covariable al número de ramas productivas por árbol. Otro ejemplo es la valoración de crecimiento en plántulas establecidas que han recibido un tratamiento a base de bioestimulantes, en el cual el ensayo se inicia con plantas que tienen deferentes alturas, siendo en este caso la altura inicial la covariable.

Los supuestos bajo los cuales es eficaz el ANCOVA son los mismos que se consideran para el análisis de varianza; tales como la normalidad, la homocedasticidad y la independencia de sus residuales. El modelo estadístico es aditivo y no debe existir interacción entre una covariable y los tratamientos de estudio, además la variable concomitante debe tener una relación lineal con la variable de respuesta.

Para ampliar un poco más esta técnica de ajuste de datos, seguidamente se expone un ejemplo utilizando el software *Infostat*.

Ejemplo de aplicación 4.3

Se ha realizado un ensayo sobre el efecto de cuatro dosis de un fertilizante de aplicación foliar con contenido hormonal en la etapa de floración del plátano, además de un tratamiento testigo (sin aplicación); para lo cual se ha utilizado una distribución de bloques completos al azar. Como variable de respuesta (Y) se ha medido la producción en $kg/racimo$. Asimismo, dado el momento de aplicación de los tratamientos y las condiciones del fuste de las plantas, se ha considerado medir el diámetro (cm) del tallo como variable concomitante (covariable). Los tratamientos y los datos son los que se indican en la tabla 4.13.

Dado que uno de los objetivos principales de un análisis de covarianza es de ajustar los promedios, además de verificar si la covarianza es significativa, las hipótesis estadísticas que se plantean son:

- | | |
|---|---|
| $H_0: t_{1(a_j)} = t_{2(a_j)} = \dots = t_{5(a_j)}$ | Hipótesis nula, planteada como la igualdad entre tratamientos con sus medias ajustadas. |
| $H_0: COV(X) = 0$ | Hipótesis nula, indicando que no existe covarianza entre las variables; es decir, a la variable Y no le afecta la variable concomitante X . |

Tabla 4.13. Datos de producción en kg/racimo (Y) y diámetro del tallo (cm) como covariable (X)³¹.

N°	Tratamientos	Variables	Repeticiones				Total	Promedios
			I	II	III	IV		
1	0,4 L/ha	Y	5,4	6,1	4,3	5,2	21,0	5,3
		X	20,3	22,1	17,9	17,1	77,4	19,4
2	0,8 L/ha	Y	4,5	3,8	5,1	5,7	19,1	4,8
		X	21,7	18,4	25,3	21,8	87,2	21,8
3	1,2 L/ha	Y	7,7	6,9	6,2	8,3	29,1	7,3
		X	23,6	25,8	24,9	29,0	103,3	25,8
4	1,6 L/ha	Y	7,6	7,9	8,2	5,9	29,6	7,4
		X	26,7	28,1	27,2	23,6	105,6	26,4
5	Testigo	Y	4,5	3,2	4,1	3,8	15,6	3,9
		X	18,5	19,3	20,0	19,2	77,0	19,3

El proceso de cálculo consiste en primer lugar en realizar el análisis de covarianza y luego la obtención de los promedios ajustados. Estos cálculos involucran una serie de pasos que pueden resultar un tanto tediosos, cuya relativa complejidad depende del tipo de distribución y arreglo experimental que se desarrolle en el ensayo. Para este ejemplo se tiene una distribución de bloques completos al azar y bajo este diseño se lo ha analizado en el software estadístico Infostat.

Para demostrar la reducción del error experimental se han realizado dos análisis, un ANDEVA sólo con la variable dependiente y el análisis de covarianza incluyendo a la variable concomitante (diámetro delseudotallo). Los pasos para estos análisis son los siguientes:

³¹ Quizás el lector estadístico se esté preguntando qué tiene que ver el diámetro del tallo de una planta con su producción en este ejemplo para haber aplicado un ANCOVA, cuando es deducible que la covariable debe tener relación con la respuesta medida. No obstante, el agrónomo advertirá que una planta, especialmente como el banano, producirá mejor mientras el diámetro del tallo es más grande, debido a un mayor flujo de agua y nutrientes; de allí que se haya tomado en cuenta dicha covariable como ejemplo. Se hace esta distinción en el sentido de que, si el investigador decide aplicar un ANCOVA, tendrá que tener una fundamentación teórica de la relación que tenga la variable que designa como covariable.

1. Ingreso de datos

Caso	TRATAMIENTOS	REPETICIONES	Producción (kg/racimo)	Diámetro de pseudotallo (cm)
1	T1: 0.4 L/ha	1	5,4	20,3
2	T2: 0.8 L/ha	1	4,5	21,7
3	T3: 1.2 L/ha	1	7,7	23,6
4	T4: 1.6 L/ha	1	7,6	26,7
5	T5: Testigo	1	4,5	18,5
6	T1: 0.4 L/ha	2	6,1	22,1
7	T2: 0.8 L/ha	2	3,8	18,4
8	T3: 1.2 L/ha	2	6,9	25,8
9	T4: 1.6 L/ha	2	7,9	28,1
10	T5: Testigo	2	3,2	19,3
11	T1: 0.4 L/ha	3	4,3	17,9
12	T2: 0.8 L/ha	3	5,1	25,3
13	T3: 1.2 L/ha	3	6,2	24,9
14	T4: 1.6 L/ha	3	8,2	27,2
15	T5: Testigo	3	4,1	20,0
16	T1: 0.4 L/ha	4	5,2	17,1
17	T2: 0.8 L/ha	4	5,7	21,8
18	T3: 1.2 L/ha	4	8,3	29,0
19	T4: 1.6 L/ha	4	6,9	23,6
20	T5: Testigo	4	3,8	19,2

Figura 4.10. Ingreso de datos en el Infostat.

2. Aplicación de las opciones para el análisis

Caso	TRATAMIENTOS	REPE	Producción (kg/racimo)	Diámetro de pseudotallo (cm)
1	T1: 0.4 L/ha			20,3
2	T2: 0.8 L/ha			21,7
3	T3: 1.2 L/ha			23,6
4	T4: 1.6 L/ha			26,7
5	T5: Testigo			18,5
6	T1: 0.4 L/ha			22,1
7	T2: 0.8 L/ha			18,4
8	T3: 1.2 L/ha			25,8
9	T4: 1.6 L/ha			28,1
10	T5: Testigo			19,3
11	T1: 0.4 L/ha			17,9
12	T2: 0.8 L/ha			25,3
13	T3: 1.2 L/ha			24,9
14	T4: 1.6 L/ha			27,2
15	T5: Testigo			20,0
16	T1: 0.4 L/ha			17,1
17	T2: 0.8 L/ha			21,8
18	T3: 1.2 L/ha			29,0
19	T4: 1.6 L/ha			23,6
20	T5: Testigo			19,2

Figura 4.11. Secuencia del análisis en Infostat.

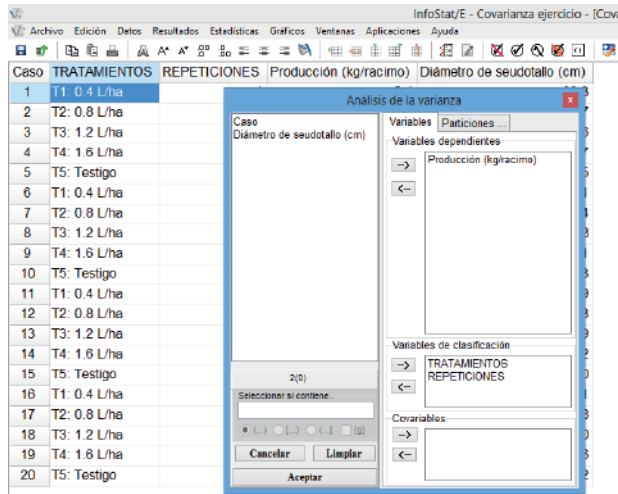


Figura 4.12. Secuencia para la realización del análisis de varianza en Infostat.

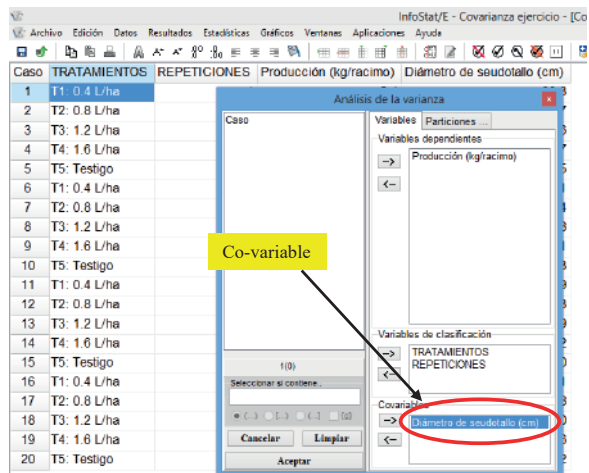


Figura 4.13. Secuencia para la realización del análisis de covarianza en Infostat.

En el primer cuadro de reporte del análisis de varianza 1 puede observarse que los tratamientos presentan diferencias significativas ($P = 0,0001$), pero con un coeficiente de variación (CV) del 13,49%.

En el segundo cuadro, que es otro análisis de varianza con inclusión de la covariable diámetro delseudotallo (Análisis de varianza 2), puede apreciarse que los tratamientos vuelven a presentar diferencia significativa ($P = 0,0076$), sin embargo el coeficiente de variación se ha reducido al 9,66%; lo que consecuentemente se ha cumplido con uno de los objetivos del análisis de covarianza al tener un ANDEVA de mayor precisión. Asimismo, también se observa que la covariable tiene un efecto significativo ($P = 0,0048$); es decir, la variable de respuesta está influenciada de forma significativa por el diámetro delseudotallo, por lo tanto, se hace necesario realizar el ajuste de los promedios.

Análisis de varianza 1

Variable	N	R ²	R ² Aj	CV
Producción (kg/racimo)	20	0,86	0,77	13,49

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo.	42,95	7	6,14	10,13	0,0003
TRATAMIENTOS	42,23	4	10,56	17,43	0,0001
REPETICIONES	0,73	3	0,24	0,40	0,7559
Error	7,27	12	0,61		
Total	50,22	19			

Análisis de varianza 2 (Covarianza)

Variable	N	R ²	R ² Aj	CV
Producción (kg/racimo)	20	0,93	0,88	9,66

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor	Coef
Modelo.	46,81	8	5,85	18,84	<0,0001	
TRATAMIENTOS	7,62	4	1,90	6,13	0,0076	
REPETICIONES	1,53	3	0,51	1,65	0,2352	
Diámetro deseudotallo (cm..)	3,85	1	3,85	12,41	0,0048	0,24
Error	3,42	11	0,31			
Total	50,22	19				

El ajuste de promedios se realiza mediante el modelo de regresión lineal indicado en la expresión 4.8. En este modelo, \hat{Y}_i es el promedio ajustado de cada tratamiento, \bar{Y}_i es la media de cada tratamiento sin ajuste, b es el coeficiente de regresión, \bar{X}_i es la media de la covariable en cada tratamiento y \bar{X} es el promedio general de la covariable.

$$\hat{Y}_i = \bar{Y}_i - b(\bar{X}_i - \bar{X}) \quad (4.8)$$

Los promedios con y sin ajuste, obtenidos también en Infostat, son los siguientes:

Promedios sin ajuste

TRATAMIENTOS	Media
T1: 0.4 L/ha	5,25
T2: 0.8 L/ha	4,78
T3: 1.2 L/ha	7,28
T4: 1.6 L/ha	7,65
T5: Testigo	3,90

Promedios con ajuste

TRATAMIENTOS	Media
T1: 0.4 L/ha	6,02
T2: 0.8 L/ha	4,95
T3: 1.2 L/ha	6,47
T4: 1.6 L/ha	6,71
T5: Testigo	4,70

Una vez aplicado el modelo de regresión lineal de la expresión 4.8, los promedios se han ajustado. Por ejemplo, para el tratamiento 1 este valor sufrió un cambio desde 5,25 a 6,02 kg/racimo.

Las explicaciones que se han realizado respecto de contar con alternativas y acciones que mejoren la precisión de un experimento sólo han considerado el hecho de tener una experiencia con una sola variable de respuesta. No obstante, en el caso de los ensayos agrícolas no solo se mide una variable de respuesta sino por lo general se evalúa más de una variable, debido a que el ambiente en el cual se desarrollan estos experimentos es de mucha heterogeneidad y el experimentador está obligado a valorar más de un efecto para poder establecer si los tratamientos evaluados han incidido en las respuestas valoradas. Además, es lógico esperar que ciertas respuestas tengan relación entre ellas o estén correlacionadas ya sea positiva o negativamente; de allí que valorarlas estadísticamente de forma independiente a través de sus respectivos análisis de varianza, es posible que se termine declarando efectos no significativos cuando en realidad si los haya. En este contexto, el análisis deberá realizarse o cotarse haciendo una comparación de forma simultánea con todas las variables dependientes evaluadas mediante lo que se conoce como *análisis de varianza multivariado*, herramienta que se explica en el siguiente tema de este capítulo.

Análisis de varianza multivariado

En los experimentos en donde se miden los efectos a través de más de una variable de respuesta, sus resultados pueden estar correlacionados con expresiones genéticas de los cultivos y que no tienen nada que ver con los tratamientos que se estén probando; por lo tanto, ante esta situación se podrían estar enmascarando efectos que no pueden detectarse al utilizar herramientas clásicas como el análisis de varianza, ya que este último valora las variables de respuesta de forma independiente (análisis univariado). En este escenario, el ANDEVA también puede llevarnos al extremo más peligroso de cometer el error tipo I, ya que al analizar un grupo de variables dependientes que están correlacionadas no mantiene estable el valor α ; más bien éste se incrementa, haciendo que efectos no significativos al tradicional 5% se presenten como significativos.

En las circunstancias mencionadas en el párrafo anterior, cuando se tienen más de una variable de respuesta y éstas están correlacionadas, para evitarnos un incremento de la probabilidad de caer en los errores estadísticos, se puede realizar un análisis simultáneo de todas estas variables mediante el análisis de varianza multivariado (ANDEVAM) o también conocido como MANOVA (*Multivariate Analysis of Variance*). El ANDEVAM es una generalización del ANDEVA, diferenciándose de este último en el cual, la detección de diferencias significativas entre tratamientos lo hace a través de una sola variable de respuesta a la vez; mientras que, en el ANDEVAM, estas diferencias se establecen mediante dos o más variables de respuesta. De hecho, el análisis de varianza multivariado puede detectar diferencias analizando las variables de respuesta en conjunto cuando están correlacionadas, cosa que no es posible en ciertos casos cuando se realizan los análisis de varianza por separado.

El análisis conjunto del ANDEVAM genera un mayor número de grados de libertad en el error experimental, estableciendo un estadístico F general para todas las variables dependientes de manera indirecta a partir de otro estadígrafo conocido como Λ (lambda). Para definir este último estadígrafo existen varios test (cuatro de los cuales se indicarán más adelante), los mismos que se basan en una matriz de suma de cuadrados y en otra de productos cruzados. A continuación, se detalla un ejemplo de aplicación de esta herramienta estadística empleando para ello el software estadístico SPSS.

Ejemplo de aplicación 4.4

En un experimento se han evaluado cinco distancias de siembra de una variedad de sandía, utilizando una distribución de bloques completos al azar con 4 repeticiones. Realizar el análisis de varianza multivariado a partir de las variables evaluadas: longitud de guías (m) y el diámetro del fruto (cm). Identificando como Y_1 a la primera variable y como Y_2 a la segunda, la información al respecto es la siguiente:

Tabla 4.14. Datos de longitud de guía (Y_1) y diámetro del fruto (Y_2).

N°	Tratamientos (t_i)	Repeticiones (R_j)							
		I		II		III		IV	
		Y_1	Y_2	Y_1	Y_2	Y_1	Y_2	Y_1	Y_2
1	0,60 x 1,50	0,95	28,3	1,12	32,1	0,99	30,2	1,01	29,3
2	0,80 x 1,50	1,35	21,6	1,33	23,6	1,28	23,1	1,41	25,1
3	1,00 x 1,50	1,88	22,9	1,95	20,8	2,02	19,6	2,23	23,6
4	0,60 x 2,00	1,22	24,5	1,15	25,5	1,34	28,3	1,32	25,6
5	0,80 x 2,00	1,45	28,0	1,37	26,9	1,50	24,9	1,49	24,8
Total ΣR_j		6,85	125,3	6,92	128,9	7,13	126,1	7,46	128,4

Tratamientos	1		2		3		4		5	
Variabes	Y_1	Y_2	Y_1	Y_2	Y_1	Y_2	Y_1	Y_2	Y_1	Y_2
Total	4,07	119,9	5,37	93,4	8,08	86,9	5,03	103,9	5,81	104,6
Promedios	1,02	30,0	1,34	23,4	2,02	21,7	1,26	26,0	1,45	26,2

Variable	Total	N° de datos (n)	Promedio general (\bar{Y})
Y_1	28,36	20	1,42
Y_2	508,7	20	25,4

Las hipótesis estadísticas, de acuerdo a las variables y a los tratamientos son las siguientes:

$H_0: \mu_{11} = \mu_{12} = \mu_{13} = \mu_{14} = \mu_{15}$ No existe variación en conjunto entre los tratamientos estudiados.

$H_0: \mu_{21} = \mu_{22} = \mu_{23} = \mu_{24} = \mu_{25}$

$H_a: \mu_{11} \neq \mu_{12} \neq \mu_{13} \neq \mu_{14} \neq \mu_{15}$ Existe variación en conjunto entre los tratamientos estudiados.

$H_a: \mu_{21} \neq \mu_{22} \neq \mu_{23} \neq \mu_{24} \neq \mu_{25}$

Para iniciar el análisis, primero habrá que verificar la correlación entre las variables dependientes. La secuencia en el software es la siguiente:

Analizar → *C orrelaciones* → *Bivariadas*

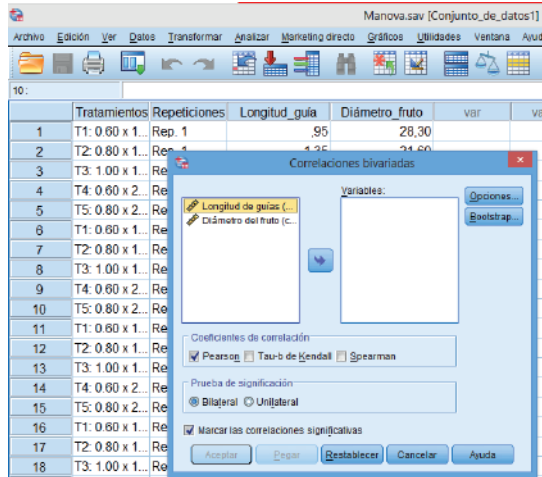


Figura 4.14. Proceso para la obtención de correlación entre las variables longitud de guía y diámetro del fruto.

Una vez que las variables son seleccionadas y se define el coeficiente de correlación (Pearson para datos normalizados), el reporte es el que se indica a continuación:

Correlaciones		Longitud de guías (m)	Diámetro del fruto (cm)
Longitud de guías (m)	Correlación de Pearson	1	-0,695**
	Sig. (bilateral)		0,001
	N	20	20
Diámetro del fruto (cm)	Correlación de Pearson	-0,695**	1
	Sig. (bilateral)	0,001	
	N	20	20

** . La correlación es significativa al nivel 0,01 (bilateral).

Según este análisis, las variables se encuentran correlacionadas (sig. = 0,001) y, por tanto, es procedente el desarrollo del ANDEVAM en esta situación. La secuencia de pasos para desarrollar en el software este análisis es el siguiente:

Analizar → Modelo lineal general → Multivariante

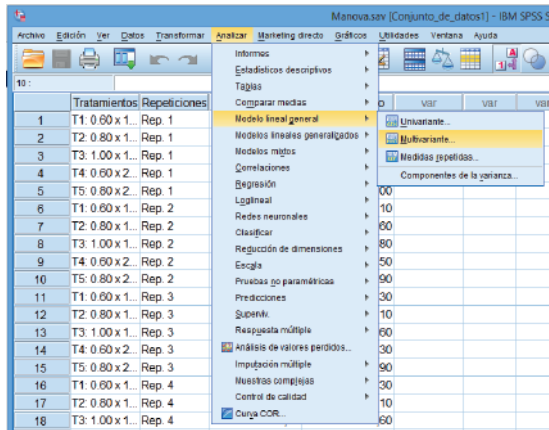


Figura 4.15. Secuencia para la obtención del análisis de varianza multivariado (ANDEVAM).

La secuencia indicada permite llegar a la siguiente ventana, en donde se clasifica en casilleros las variables independientes (Factores fijos) y las dependientes:

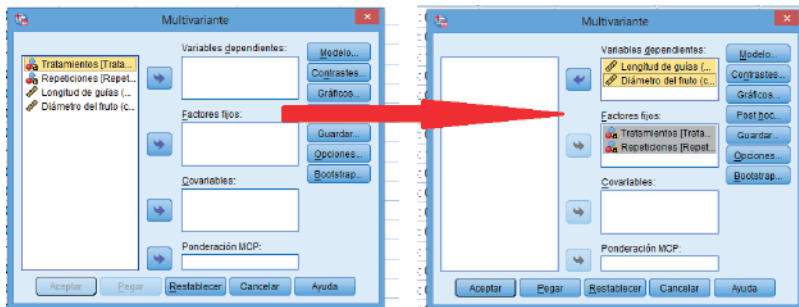


Figura 4.16. Clasificación de las variables dentro de la pestaña para la realización del ANDEVAM.

Una vez clasificadas las variables, lo primero que debe estructurarse es el modelo. Debido a que el experimento se realizó bajo una distribución de bloques completos al azar, el modelo a definir en el software es:

$$\text{Tratamientos} + \text{Repeticiones}$$

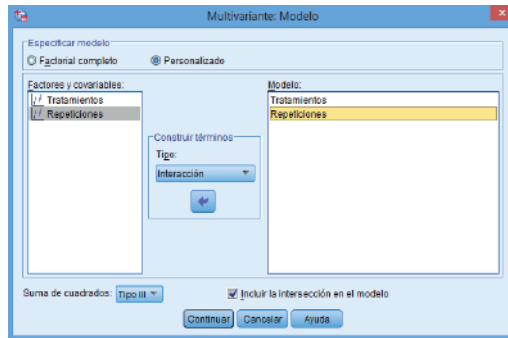


Figura 4.17. Definición del modelo del ANDEVAM, considerando una distribución de bloques completos al azar.

En el análisis que realiza el software se presenta primero, como salida en forma general, el grado de significancia entre todas las variables dependientes. Esto lo hace mediante cuatro estadígrafos (la mayoría de paquetes estadísticos también lo hacen) conocidos como: *Traza de Pillai*, *Lambda de Wilks*, *Traza de Hotelling* y *Raíz máxima de Roy*. De estos, la de mayor contundencia es el estadístico *Lambda de Wilks* (Λ), cuyo valor varía entre 0 y 1; siendo mayor la significancia si el valor se acerca a 0 y menor, si éste se acerca a 1 (este valor no tiene nada que ver con el valor P de significancia).

Para el ejemplo, según el reporte del software, existen diferencias significativas ($p < 0,05$) entre los distintos tratamientos. Esto lo registran los cuatro estadígrafos; que en el caso del test de *Lambda de Wilks*, se ha alcanzado un valor de 0,012 para Λ . Este último valor, al transformarse al estadígrafo *F* de Fisher, equivale a 22,812.

Contrastes multivariados^a

Efecto		Valor	F	Gl de la hipótesis	Gl del error	Sig.
Intersección	Traza de Pillai	0,998	3272,166 ^b	2	11	0,000
	Lambda de Wilks	0,002	3272,166 ^b	2	11	0,000
	Traza de Hotelling	594,939	3272,166 ^b	2	11	0,000
	Raíz mayor de Roy	594,939	3272,166 ^b	2	11	0,000
Tratamientos	Traza de Pillai	1,440	7,708	8	24	0,000
	Lambda de Wilks	0,012	22,812 ^b	8	22	0,000
	Traza de Hotelling	46,416	58,020	8	20	0,000
	Raíz mayor de Roy	45,560	136,680 ^c	4	12	0,000
Repeticiones	Traza de Pillai	0,421	1,065	6	24	0,410
	Lambda de Wilks	0,595	1,085 ^b	6	22	0,401
	Traza de Hotelling	0,653	1,088	6	20	0,403
	Raíz mayor de Roy	0,609	2,435 ^c	3	12	0,115

a. Diseño: Intersección + Variedades + Repeticiones

b. Estadístico exacto

c. El estadístico es un límite superior para la F el cual ofrece un límite inferior para el nivel de significación.

El análisis de varianza multivariado, en términos generales, comprende dos fases. Primero debe realizarse el contraste para verificar la existencia de diferencia significativa entre todas las variables dependientes, lo que se cumple con el estadígrafo *Lambda de Wilks*. Si en este primer contraste se detecta significancia, lo segundo que debe realizarse es un análisis univariado, esto es, realizar un ANDEVA con cada una de las variables dependientes para definir qué variables son las que están reportando los efectos de los tratamientos. Este último análisis es el que a continuación se detalla, según el reporte del software:

Pruebas de los efectos inter-sujetos

Origen	Variante dependiente	SC tipo III	GL	Media cuadrática	F	Sig.
Modelo corregido	Longitud de guías (m)	2,267 ^a	7	0,32	47,22	0,000
	Diámetro del fruto (cm)	159,933 ^b	7	22,85	7,25	0,002
Intersección	Longitud de guías (m)	40,214	1	40,21	5863,59	0,000
	Diámetro del fruto (cm)	12938,79	1	12938,79	4103,97	0,000
Tratamientos	Longitud de guías (m)	2,222	4	0,56	80,99	0,000
	Diámetro del fruto (cm)	158,103	4	39,53	12,54	0,000
Repeticiones	Longitud de guías (m)	0,045	3	0,015	2,19	0,142
	Diámetro del fruto (cm)	1,830	3	0,61	0,19	0,899
Error	Longitud de guías (m)	0,082	12	0,007		
	Diámetro del fruto (cm)	37,833	12	3,153		
Total	Longitud de guías (m)	42,564	20			
	Diámetro del fruto (cm)	13136,55	20			
Total corregida	Longitud de guías (m)	2,349	19			
	Diámetro del fruto (cm)	197,766	19			

a. R cuadrado = 0,965 (R cuadrado corregida = 0,945)

b. R cuadrado = 0,809 (R cuadrado corregida = 0,697)

En los análisis de varianza de las dos variables, longitud de guía y diámetro de fruto, se han podido detectar diferencias significativas ($p < 0,05$). Esto permite establecer que la variabilidad reportada por el ANDEVAM es debida a las dos variables analizadas. Lo que sigue es determinar qué tratamientos, en las dos variables, definen las diferencias significativas. Para obtener este análisis, en la ventana Multivariante (figura 4.16) habrá que seleccionar la pestaña *Post hoc* (Pruebas a posteriori), desde donde se podrá seleccionar una de las varias pruebas de comparación múltiple que ofrece el software. En este caso, seleccionaremos el test de Tukey como prueba aplicable a experimentos de campo. Esta ventana es la que se indica en la figura 4.18.

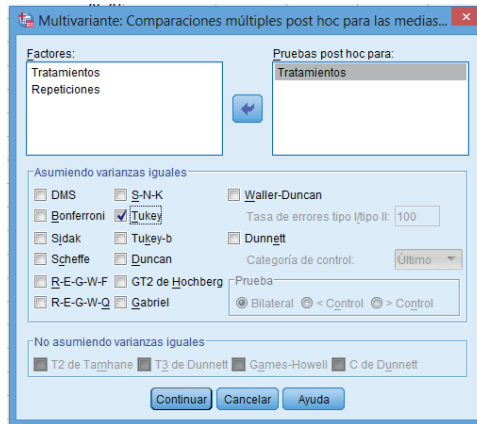


Figura 4.18. Selección del test de Tukey dentro de la pestaña *Post hoc*.

Según el reporte dado por el software y de acuerdo al test de Tukey ($p < 0,05$) para la longitud de guía, el tratamiento con el mayor promedio estadístico es el de distanciamiento 1,00 x 1,50 (T3), con una media de 2,02 m. En el caso del diámetro del fruto, los tratamientos con las mayores medias son los de distanciamientos 0,80 x 2,00 (T5) y 0,60 x 1,50 (T1), con valores de 26,15 cm y 30,00 cm, respectivamente. Cabe indicar que el tratamiento T5, si bien pertenece al subconjunto de las medias más altas (3), también se ubica en el rango del subconjunto 2.

Longitud de guías (m)

DHS de Tukey^{a,b,c}

Tratamientos	N	Subconjunto			
		1	2	3	4
T1: 0,60 x 1,50	4	1,02			
T4: 0,60 x 2,00	4		1,26		
T2: 0,80 x 1,50	4		1,34	1,34	
T5: 0,80 x 2,00	4			1,45	
T3: 1,00 x 1,50	4				2,02
Sig.		1,000	0,609	0,378	1,000

Se muestran las medias de los grupos de subconjuntos homogéneos.

Basadas en las medias observadas.

El término de error es la media cuadrática(Error) = 0,007.

a. Usa el tamaño muestral de la media armónica = 4,000

b. Los tamaños de los grupos son distintos. Se empleará la media armónica de los tamaños de los grupos. No se garantizan los niveles de error tipo I.

c. Alfa = 0,05.

Diámetro del fruto (cm)

DHS de Tukey^{a,b,c}

Tratamientos	N	Subconjunto		
		1	2	3
T3: 1,00 x 1,50	4	21,73		
T2: 0,80 x 1,50	4	23,35	23,35	
T4: 0,60 x 2,00	4		25,98	25,98
T5: 0,80 x 2,00	4		26,15	26,15
T1: 0,60 x 1,50	4			29,98
Sig.		0,700	0,233	0,050

Se muestran las medias de los grupos de subconjuntos homogéneos.

Basadas en las medias observadas.

El término de error es la media cuadrática(Error) = 3,153.

a. Usa el tamaño muestral de la media armónica = 4,000

b. Los tamaños de los grupos son distintos. Se empleará la media armónica de los tamaños de los grupos. No se garantizan los niveles de error tipo I.

c. Alfa = 0,05.

Como el ANDEVAM es una extensión del ANDEVA, para su precisión también debe cumplir con los supuestos de normalidad y de homocedasticidad de los datos. En el SPSS estos supuestos pueden establecerse calculando el test de Shapiro – Wilk para la normalidad y la prueba de Levene para la igualdad de varianzas de los grupos.

Para verificar la normalidad de los datos, la secuencia de pasos en el software es el siguiente:

Analizar → Estadísticos descriptivos → Explorar

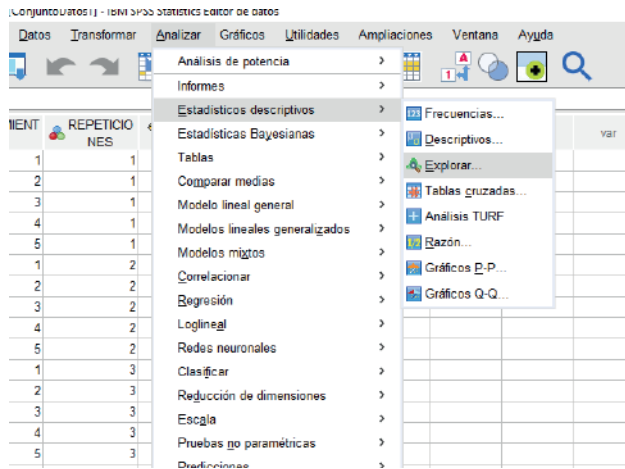


Figura 4.19. Secuencia de pasos para verificar la normalidad mediante el test de Shapiro - Wilk en SPSS.

Para establecer que los datos son normales, la hipótesis nula a comprobar es que la información se ajusta a una distribución normal; para ello, la significancia debe ser mayor al 5%. En este caso, el software presenta una salida no solo aplicando el test de Shapiro – Wilk, sino también el de Kolmogorov – Smirnov (con corrección por Lilliefors), tal como puede observarse en la tabla que sigue a continuación. Considerando el primer test como el recomendado para menos de 50 datos, los residuales de la longitud de guías no presentan normalidad, según el P-valor observado (0,047); lo cual es diferente para el diámetro del fruto, en donde se cumple el supuesto de normalidad de acuerdo al P-valor observado (0,970). Cuando hay falta de normalidad como en la primera variable, habrá que observar también la falta de homocedasticidad de la misma para decidir no aplicar directamente el ANDEVAM, ya que, como el ANDEVA, estas son herramientas estadísticas relativamente robustas.

Pruebas de normalidad

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Long. de guías (m)	0,208	20	0,024	0,903	20	0,047
Diám. del fruto (cm)	0,130	20	0,200*	0,983	20	0,970

*. Esto es un límite inferior de la significación verdadera.

a. Corrección de significación de Lilliefors

Asimismo, tal como se mencionó antes, para verificar la homocedasticidad de los datos de las variables, el test que utiliza el SPSS es el de Levene. Esta prueba se aplica desde la ventana

multivariante (figura 4.16), seleccionando la ventana *Opciones*, desde donde se deberá escoger la pestaña *Pruebas de homogeneidad* (figura 4.21).

Cabe indicar que el análisis de homocedasticidad se realiza considerando solo el factor principal que se evalúa en un ensayo (bajo el modelo de un DCA); por lo tanto, la fuente secundaria de bloqueo no debe ingresar en el modelo. Esta indicación se detalla en la figura 4.20.

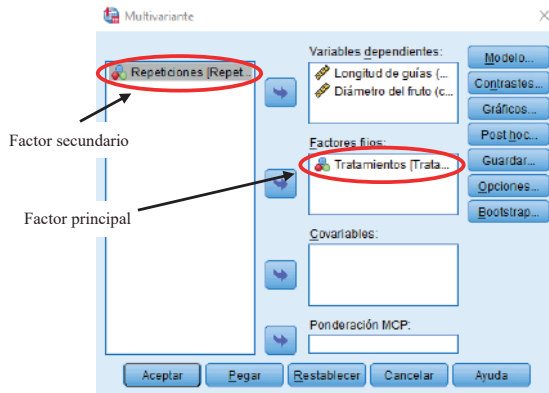


Figura 4.20. Definición del modelo, considerando sólo el factor principal para verificar homocedasticidad.

En el reporte del software para la homogeneidad de las varianzas no sólo se obtiene el resultado del test de Levene, adicionalmente se puede verificar la homogeneidad de las matrices de covarianzas mediante la prueba de Box. Esta igualdad también debe cumplirse para el caso del ANDEVAM. El resultado de este análisis es el que se indica a continuación, primero con la selección de *Pruebas de homogeneidad* (figura 4.21) y luego el reporte de estas pruebas.

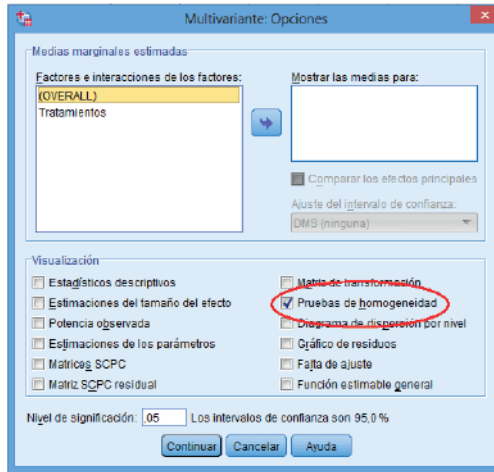


Figura 4.21. Selección de la pestaña para el análisis de homocedasticidad.

La hipótesis nula a comprobar es que, tanto las matrices de covarianza como las varianzas (entre tratamientos) son homogéneas, para lo cual se debe cumplir que $p > 0,05$. Puede observarse en los resultados que, según el test de Box, las matrices de covarianzas son homogéneas; así como también las varianzas de los tratamientos, según el test de Levene. Los valores 0,684 del test de Box y de 0,359 y 0,896 de Levene, para las variables longitud de guías y diámetro del fruto, respectivamente, corroboran la homocedasticidad de los residuales.

Prueba de Box sobre la igualdad de las matrices de covarianzas^a

M de Box	13,100
F	0,768
gl1	12
gl2	1667,647
Sig.	0,684

Contrasta la hipótesis nula de que las matrices de covarianza observadas de las variables dependientes son iguales en todos los grupos.

a. Diseño: Intersección + Tratamientos

Contraste de Levene sobre la igualdad de las varianzas error^a

	F	gl1	gl2	Sig.
Longitud de guías (m)	1,180	4	15	0,359
Diámetro del fruto (cm)	0,264	4	15	0,897

Contrasta la hipótesis nula de que la varianza error de la variable dependiente es igual a la largo de todos los grupos.

a. Diseño: Intersección + Tratamientos

Entre los elementos aplicables al análisis de datos es importante también valorar la variabilidad existente entre los individuos de una misma unidad experimental que, a la postre y siendo de efecto significativo, sino se separa del error experimental terminará engrosándolo y haciendo que se pierda precisión. En este contexto, la varianza que se produce entre individuos de una misma unidad experimental (diferente a unidades experimentales de un mismo tratamiento) se define como *error de muestreo* y es el tema que se trata en el siguiente apartado como otra alternativa de análisis que busque mejorar la precisión de los experimentos.

Análisis de varianza con submuestreo

A nivel de experimentos de campo, cuando la unidad experimental corresponde a una parcela de un cultivo de ciclo corto, la cantidad de individuos que se seleccionan dentro del área de muestreo (“área útil” de la parcela) puede corresponder a grupos de más de 10 plantas, a partir de las cuales se extrae una media que viene a constituirse en el dato³² de la parcela. En este caso la variabilidad de las mediciones de la variable puede minimizarse al seleccionar una cantidad importante de unidades de observación; sin embargo, cuando el número de individuos (o unidades de evaluación) dentro de la unidad experimental es escaso (menos de 10), el promedio o dato puede verse afectado por la variabilidad que pudiera existir entre los pocos individuos o plantas que conforman la unidad experimental.

Cuando existe una importante dispersión entre los pocos individuos que conforman una unidad experimental, de tal forma que inciden sustancialmente en el incremento del error experimental, lo que puede hacerse es una valoración de dicha variabilidad separándola de toda la variación “inexplicada” que debería circundar el antes mencionado error experimental. En estas circunstancias, la varianza intra-grupo en realidad estará compuesta de la varianza que ocurre entre los individuos de una misma unidad experimental, que se identifican como submuestras, y la varianza entre las unidades experimentales de un mismo tratamiento. Por lo tanto, bajo el modelo de una distribución completamente al azar, la variabilidad que no es parte de la variación entre los tratamientos estará compuesta por el cuadrado medio del error experimental (varianza dentro de los tratamientos) y el cuadrado medio de un error de muestreo³³. Este modelo es el siguiente:

³² En un experimento, un dato es un valor de una variable determinada que representa a una unidad experimental; de allí que el número de datos de alguna variable específica esté definido en función del número de tratamientos y repeticiones que tuviera el ensayo. Este dato, puede establecerse a partir de un solo individuo cuando la unidad experimental es solamente ese individuo; o un grupo de individuos, a partir de los cuales se extrae una media. Para el primer caso, un solo árbol de mango puede representar a la unidad experimental; y en el segundo, un grupo de 100 plantas de maíz podrían constituirse como tal unidad experimental, definiendo en esta última la mal llamada “parcela útil” (en realidad, es un área de muestreo de la parcela) desde donde se seleccionará un grupo de plantas para la obtención del dato, el mismo que será un promedio de las observaciones que se realicen en este grupo de plantas.

³³ Al respecto de este error de muestreo, se hace preciso distinguir lo que corresponde relativamente al error experimental. Este último es exclusivamente la variabilidad que sucede entre unidades experimentales en las cuales se aplica un mismo tratamiento, mientras que el error de muestreo (que es un error aleatorio), tal como se lo ha

$$Y_{ijk} = \mu + \tau_i + \varepsilon_{ij} + \delta_{ijk} \quad (4.7)$$

En el modelo 4.7, Y_{ijk} es una observación cualquiera dentro del experimento (un dato), μ es la respuesta constante del material experimental, τ_i es el efecto del tratamiento, ε_{ij} es el efecto del error experimental y δ_{ijk} es el efecto del error de muestreo. Las sumas de cuadrados respectivas para este mismo modelo, se definen como:

$$SC_{total} = SC_{tratamiento} + SC_{error\ exp.} + SC_{error\ muestreo}$$

Los cálculos respectivos del análisis de varianza, considerando un modelo para una distribución completamente al azar, se indican en la tabla 4.15. Cómo podrá observarse, el estadístico F se calcula utilizando como divisor el cuadrado medio del error experimental; considerando que este último representa al efecto de la varianza de causas desconocidas (que es la concepción de la metodología del análisis de varianza), una vez que se le ha retirado la variabilidad debida al error de muestreo.

Tabla 4.15. Estructura del análisis de varianza para una distribución completamente al azar con submuestreo.

Fuentes de variación	GL	SC	CM	F
Total	$trs - 1$	$\sum Y_{ijk}^2 - \frac{(\sum Y_{ijk})^2}{trs}$		
Tratamientos	$t - 1$	$\sum \frac{(\sum Y_i)^2}{rs} - \frac{(\sum Y_{ijk})^2}{trs}$	$\frac{SC_t}{t - 1}$	$\frac{CM_t}{CM_E}$
Error experimental	$t(r - 1)$	$\frac{\sum Y_{ij}^2}{s} - \sum \frac{(\sum Y_i)^2}{rs}$	$\frac{SC_E}{t(r - 1)}$	
Error de muestreo	$tr(s - 1)$	$\sum Y_{ijk}^2 - \frac{\sum Y_{ij}^2}{s}$		

En la tabla 4.15, t es el número de tratamientos, r es el número de repeticiones, s es el número de submuestras de cada unidad experimental, SC_t es la suma de cuadrados de tratamientos y SC_E es la suma de cuadrados del error experimental. Nótese que el planteamiento indicado en esta tabla se ha establecido considerando un experimento balanceado (tratamientos con igual número de repeticiones). En forma análoga, la estructura del análisis de varianza definida bajo una distribución de bloques completos al azar también con el desglose del submuestreo se detalla en la tabla 4.16.

indicado en esta parte, es la variabilidad de los individuos que pertenecen a una misma unidad experimental. Por ello es que algunos autores mencionan que, si bien el error experimental es una variación aleatoria, no toda variación aleatoria es error experimental. En experimentos mal manejados, la variación aleatoria es sólo una parte del error experimental.

Tabla 4.16. Estructura de análisis de varianza para una distribución de bloques completos al azar con submuestreo.

Fuentes de variación	GL	SC	CM	F
Total	$trs - 1$	$\sum Y_{ijk}^2 - \frac{(\sum Y_{ijk})^2}{trs}$		
Tratamientos	$t - 1$	$\sum \frac{(\sum Y_i)^2}{rs} - \frac{(\sum Y_{ijk})^2}{trs}$	$\frac{SC_t}{t - 1}$	$\frac{CM_t}{CM_{EE}}$
Repeticiones	$r - 1$	$\sum \frac{(\sum Y_j)^2}{ts} - \frac{(\sum Y_{ijk})^2}{trs}$	$\frac{SC_r}{r - 1}$	$\frac{CM_r}{CM_{EE}}$
Error experimental	$(t - 1)(r - 1)$	$\frac{\sum Y_{ij}^2}{s} - \frac{(\sum Y_{ijk})^2}{trs} - SC_t - SC_r$	$\frac{SC_{EE}}{(t - 1)(r - 1)}$	
Error de muestreo	$tr(s - 1)$	$\sum Y_{ijk}^2 - \frac{\sum Y_{ij}^2}{s}$		

Siendo r , en el caso de la tabla 4.16, el número de bloques o repeticiones del experimento.

En la práctica, existen diferentes tipos de ensayos en donde puede realizarse la valoración estadística del submuestreo con el propósito de mejorar la precisión de éstos. Hay ejemplos como los ensayos de sandía o melones, en donde la variable de respuesta es el número de frutos por planta o de longitud de guía, en los cuales podría valorarse la varianza al interior de las parcelas; dado que generalmente son pocas plantas las que conforman la unidad experimental y mucho menos las que se ubican dentro del área de muestreo de la parcela. Otros ejemplos pueden ser experimentos con animales, en donde la unidad experimental está representada por corrales conteniendo menos de 10 individuos, en el cual es común encontrar una dispersión importante entre los individuos de una misma unidad experimental. A continuación, se detalla la aplicación de este procedimiento en el análisis de varianza de un ensayo sobre dosis de potasio en plántulas de banano.

Ejemplo de aplicación 4.5

Se prueban cinco dosis de potasio en banano en estado de plantilla. Para esto se utiliza una distribución de bloques completos al azar, empleando para ello 4 repeticiones. La parcela se definió con tres hileras de plantas, considerando en cada una de ellas un total de 4 plantas, circunscribiendo en total 12 plantas por unidad experimental. En cada parcela se consideraron dos plantas (las centrales) para el muestreo respectivo, evaluándose en cada una de ellas la altura de las mismas. Los datos son los siguientes:

Tabla 4.17. Datos sobre la altura de planta (m) de banano con dos submuestras por parcela.

N°	Tratamientos	Repeticiones				Total por muestra	Total por tratamiento	Promedios
		I	II	III	IV			
1	50 kg/ha	2,87	1,85	1,95	1,12	7,79	16,11	2,01
		3,12	1,82	1,26	2,12	8,32		
		5,99	3,67	3,21	3,24			
2	100 kg/ha	2,99	3,23	3,12	3,34	12,68	23,86	2,98
		3,15	2,86	2,12	3,05	11,18		
		6,14	6,09	5,24	6,39			
3	150 kg/ha	1,26	2,25	3,25	2,95	9,71	23,32	2,92
		3,35	3,26	3,55	3,45	13,61		
		4,61	5,51	6,80	6,40			
4	200 kg/ha	2,96	3,35	1,54	3,05	10,90	22,86	2,86
		3,25	2,85	3,96	1,90	11,96		
		6,21	6,20	5,50	4,95			
5	250 kg/ha	2,67	3,54	1,75	3,30	11,26	23,52	2,94
		3,35	2,95	3,25	2,71	12,26		
		6,02	6,49	5,00	6,01			
Total repeticiones		28,97	27,96	25,75	26,99			
Total general							109,67	

Siguiendo lo indicado en la tabla 4.16, las sumas de cuadrados del total y de las diferentes fuentes de variación son las siguientes:

Suma de cuadrados del total	$SC_T = 2,87^2 + 1,85^2 + \dots + 2,71^2 - 109,67^2/40 = 20,67$
Suma de cuadrados de tratamientos	$SC_t = \frac{16,11^2 + 23,86^2 + \dots + 23,52^2}{4 \times 2} - 109,67^2/40 = 5,36$
Suma de cuadrados de repeticiones	$SC_r = \frac{28,97^2 + 27,96^2 + \dots + 26,99^2}{5 \times 2} - 109,67^2/40 = 0,57$
Suma de cuadrados de las unidades experimentales	$SC_{UE} = \frac{5,99^2 + 3,67^2 + \dots + 6,01^2}{2} - 109,67^2/40 = 10,95$
Suma de cuadrados del error experimental	$SC_E = 10,95 - 5,36 - 0,57 = 5,02$
Suma de cuadrados del error de muestreo	$SC_{EM} = 20,67 - 10,95 = 9,72$

El análisis de varianza con el submuestreo realizado se detalla en la tabla 4.18. Puede notarse como la variabilidad representada por la suma de cuadrados en el caso de la fuente tratamientos (5,36) es prácticamente equivalente a la del error experimental (5,02), separando lo que corresponde al error de muestreo (9,72). Sin embargo, si el análisis de varianza se realiza sin considerar el submuestreo (tabla 4.19); esto es, considerando las dos submuestras como repeticiones, la suma de cuadrados del error experimental es la fuente con mucha más variabilidad (14,74).

Tabla 4.18. Análisis de varianza para una distribución de bloques completos al azar del ejemplo de aplicación 4.6 con submuestreo.

Fuentes de variación	GL	SC	CM	F	P-valor
Total	39	20,67			
Tratamientos	4	5,36	1,34	3,19	0,052
Repeticiones	3	0,57	0,19	0,45	0,721
Error experimental	12	5,02	0,42		
Error de muestreo	20	9,72			

Tabla 4.19. Análisis de varianza para una distribución de bloques completos al azar del ejemplo de aplicación 4.6 sin considerar el submuestreo.

Fuentes de variación	GL	SC	CM	F	P
Total	39	20,67			
Tratamientos	4	5,36	1,34	1,10	0,491
Repeticiones	3	0,57	0,19	0,16	0,925
Error experimental	32	14,74	0,46		

La ganancia en precisión al realizar un ANDEVA considerando submuestreo se ve reflejada en el coeficiente de variación. Para el análisis realizado en la tabla 4.18, el CV tiene un valor de 23,6%; mientras que, el ANDEVA sin submuestreo de la tabla 4.19, reporta un CV de 24,7%.

CAPÍTULO V

DISTRIBUCIONES CON BLOQUES INCOMPLETOS

El investigador debe entender que en la precisión de un ensayo agrícola es fundamental la adecuada estructura de la distribución de las unidades experimentales (o tratamientos), incluso, con mayor importancia que el propio manejo; dado que éste último no sirve de nada, aun procurando homogeneidad, si la distribución experimental adoptada no controla fuentes de variación secundaria en ambientes heterogéneos...



Distribución en bloques incompletos balanceados

Siguiendo con las diversas alternativas metodológicas que podrían incrementar la precisión de los ensayos, en el presente capítulo se detallan algunas opciones de distribuciones experimentales que se llevan a cabo mediante bloques incompletos, es decir, bloques con un número menor de unidades experimentales del que sería necesario para contener a todos los tratamientos de un determinado experimento. Este procedimiento resulta con una eficacia creciente en cuanto a precisión mientras el número de tratamientos sea más elevado, ya que bloques con una menor extensión sobre el campo estarán expuestos a una menor incidencia de la heterogeneidad del suelo y, por ende, es esperable que el error experimental presente una menor varianza intra-bloques.

En ensayos en donde se requieren unidades experimentales relativamente grandes, la distribución de bloques completos al azar (DBCA) podría perder eficacia dado que las extensiones (amplitud transversal) de cada uno de dichos bloques harían que se expongan a la heterogeneidad del suelo en el sentido paralelo y no en el perpendicular que es cuando este diseño cumple con su razón de ser. En esta situación el error experimental se incrementará y hará que la distribución pierda precisión; o incluso, si dicha heterogeneidad corresponde a una alta distinción de fertilidad en el suelo, se podría llegar a cometer el error tipo I al rechazar una hipótesis nula verdadera y terminar recomendando el tratamiento erróneo. Ejemplos de ensayos para estos casos pueden ser los que involucran control de plagas, evaluaciones de láminas de riego o métodos de preparación de suelo, los cuales demandan tamaños de parcelas más grandes que las que comúnmente se utilizan.

Otra situación experimental en donde podría limitarse el uso de bloques completos es cuando la disponibilidad de terreno es escasa o tiene cierta forma irregular (especialmente cuando el área es relativamente angosta) que complica tener la extensión total de dichos bloques y que cada uno incluya a todos los tratamientos que se hayan previsto en la investigación. En este caso se tendrían que formar obligadamente bloques con sólo una parte del número de tratamientos que se hayan definido para el estudio.

Ante lo expuesto en los dos párrafos anteriores habría la necesidad de correr el experimento utilizando bloques incompletos, esto es, una distribución en donde la totalidad de tratamientos no está presente en cada uno de los bloques sino sólo una parte de ellos. Esta clase de distribución se conoce como *diseño de bloques incompletos*, pudiendo ser éstos balanceados o desbalanceados. El hecho de definir que una distribución de bloques incompletos tenga que ser balanceada o desbalanceada, teniendo en cuenta que esto último no ocurra por pérdida de unidades experimentales, está relacionado a la importancia que tenga cada uno de los tratamientos a evaluarse. Si éstos son considerados con la misma importancia por el investigador al momento de su comparación, deberá pensarse en que la distribución sea balanceada (igual número de repeticiones por tratamiento).

Una situación adicional que tal vez el lector haya advertido, es la de los ensayos con un elevado número de tratamientos que también limitan el uso de la tradicional distribución DBCA, debido a la extensibilidad en la amplitud de los bloques y su consecuente pérdida potencial de precisión cuando la heterogeneidad del suelo se acrecienta. Frente a esta problemática que también justifica el uso de las distribuciones de bloques incompletos, se han ideado ciertas estructuras encasilladas dentro de los llamados *diseños resolubles*, algunos de los cuales se tratan más adelante, tal como la clásica distribución en *latice*.

Entre los experimentos de agronomía que se realizan bajo condiciones no controladas (campo), en forma general, las distribuciones en bloques incompletos permiten reducir la varianza del error experimental; es decir, realizar una comparación de tratamientos con una mayor precisión y en igualdad de condiciones cuando estos ensayos son balanceados. Sin embargo, el hecho de que un ensayo desarrollado en bloques incompletos sea desbalanceado no es una

condición para que éstos tengan menor precisión que los balanceados; esa precisión, más que por la diferencia del número de valores que se utilicen en una estimación, está relacionada con mayor intensidad a la heterogeneidad del suelo, de allí que sea importante la selección de lotes más o menos homogéneos especialmente para los bloques incompletos.

En lo que respecta a este texto, en esta parte se tratará de forma general y única sobre el *diseño de bloques incompletos balanceados* (DBIB). El DBIB fue desarrollado por Yates en 1930 y consiste en una distribución experimental en donde todos los tratamientos se comparan en igualdad de condiciones, es decir, tienen el mismo número de repeticiones y cada pareja de tratamientos se presenta en una misma cantidad de bloques dentro de todo el experimento³⁴ (de allí el término balanceado y que se diga que todos los tratamientos tienen el mismo grado de importancia). Según Kuehl (2001), el balance obtenido en este diseño con igual número de ocurrencias de todos los pares de tratamientos en el mismo bloque tiene como resultado una precisión igual en todas las comparaciones entre los pares de medias de tratamientos. Este es un diseño equi-replicado, con un esquema binario tal que la diferencia entre dos efectos de tratamiento se estima con la misma varianza (Hinkelmann y Kempthorne, 2005).

La concepción primigenia de este tipo de distribución es la de considerar que el número de unidades experimentales k dentro de cada bloque es menor que el número de tratamientos t de un experimento dado ($k < t$); no obstante, para lograr el balanceo de las repeticiones r dentro del ensayo se tendrá que generar la combinatoria $\binom{t}{k}$ de tratamientos en bloques. Esta situación puede ser vista como una cortapisa en la aplicación de estas distribuciones en los experimentos que se realizan en campo cuando éstos tienen un elevado número de tratamientos y escasas unidades experimentales establecidas por la poca cantidad de terreno disponible y/o material experimental; dado que, en contraposición, buscar el balanceo del ensayo generará un elevado número de combinatorias imprácticas en términos económicos y que, además, podrían limitar el control del error experimental. Por ejemplo, en un ensayo de ocho tratamientos para bloques de sólo 4 de ellos, tendríamos que estructurar en campo un total de 70 bloques.

Por otro lado, es destacable el hecho de que este tipo de distribución podría presentar una ventaja en cuanto a la eficacia del control del error experimental, pero dicha eficacia está subordinada al planteamiento de un experimento que de pronto no vaya más allá de seis tratamientos con combinaciones en bloques de al menos unos tres de ellos. Desde esta perspectiva se ha visto la necesidad de incluir el detalle de esta metodología en este texto, en función del objetivo que se busca con su difusión: conocer alternativas que permitan hacer precisos, y por tanto eficaces³⁵, los resultados de la aplicación de los diferentes diseños experimentales en los estudios agrícolas. Asimismo, debe conocerse que existen alternativas de los DBIB que pueden plantearse con menos $\binom{t}{k}$ bloques, tal como la distribución de *bloques incompletos resolubles* en donde la fuente de control local realizado por los bloques se anida en otra fuente general que en el contexto del diseño se conoce como réplica o repetición.

En la figura 5.1 se detalla el esquema de un experimento de cuatro tratamientos bajo una DBIB, considerando sólo tres de ellos (tres unidades experimentales) en cada bloque. Nótese el balanceo del experimento ya que cada par de tratamientos en un mismo bloque se repite dos veces y cada uno de ellos tiene tres repeticiones en todo el ensayo. Cuando el número de tratamientos t es igual al número de bloques generados, se dice que el diseño es simétrico.

El número de bloques en que las parejas de tratamientos aparecen juntos dentro del experimento (λ) se deduce al considerar que existen $\lambda(t - 1)$ pares de un determinado

³⁴ En la práctica, estos dos requerimientos son los que definen que un ensayo tenga una DBIB.

³⁵ Se recalca que, la eficacia a la que se hace referencia es a la de detección de efectos y que involucra específicamente a la precisión experimental; diferente a la eficiencia, la cual incluye, además de la precisión, a la exactitud. Ésta última, fuertemente dependiente de la base teórica del experimento que el investigador debe conocer.

tratamiento dentro del experimento que es igual al producto de las r repeticiones en que aparece dicho tratamiento por $k - 1$, es decir: $\lambda(t - 1) = r(k - 1)$. De allí que el número de bloques en que las parejas de tratamientos se presenten juntos en la DBIB se establezca a través de la expresión 5.1.

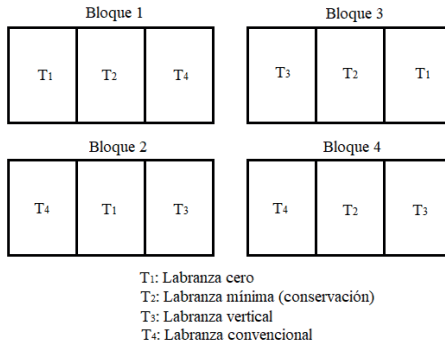


Figura 5.1. Esquema de una distribución de bloques incompletos balanceados (DBIB) con cuatro tratamientos aleatorizados en tres unidades experimentales por bloque.

$$\lambda = \frac{r(k - 1)}{(t - 1)} \quad (5.1)$$

Esquemas tabulares sobre estos tipos de distribuciones se pueden encontrar en Cochran y Cox (1978). Lo que si debe tenerse en cuenta es que, en la práctica, no existe una forma estándar de estructurar una distribución de bloques incompletos; de hecho, hay alternativas como el antes mencionado diseño de bloques incompletos resolubles que pueden formarse de forma intuitiva o también aquellas distribuciones de bloques incompletos que se definen de manera cíclica.

Aleatorización de los tratamientos en la DBIB

Algo especial, tal vez con cierta complejidad, podría ser el proceso de aleatorización de los tratamientos en las DBIB. Sin embargo, esta asignación aleatoria podría llevarse a cabo de manera parcialmente análoga a la que se realiza en la distribución en cuadro latino. Este proceso se describe mediante los siguientes pasos:

1. Se generan los grupos de tratamientos de cada bloque incompleto considerando los $\binom{t}{k}$ bloques y desde allí se sortean éstos últimos, obteniéndose así una permutación determinada.
2. Una vez estructurados los bloques y realizada la permutación respectiva, complementariamente se sortea la ubicación de los tratamientos dentro de cada bloque para minimizar el efecto de heterogeneidades definidas en el suelo.

Al respecto de esto, veamos el siguiente ejemplo:

Retomando como ejemplo lo detallado por la figura 5.1 sobre un ensayo para evaluar métodos de preparación de suelos, las permutaciones iniciales en cada uno de los bloques de tres tratamientos son los siguientes:

Bloque 1	1	2	3
Bloque 2	1	2	4
Bloque 3	1	3	4
Bloque 4	2	3	4

Sorteando el lugar de ocupación de los bloques para definir sus nuevas posiciones, se tiene:

Bloque 3	1	3	4
Bloque 1	1	2	3
Bloque 4	2	3	4
Bloque 2	1	2	4

Finalmente se sorte la ubicación de cada tratamiento dentro de los bloques, quedándonos definitivamente la distribución que se indica a continuación:

Bloque 3	3	1	4
Bloque 1	2	3	1
Bloque 4	4	2	3
Bloque 2	1	4	2

Correcciones por falta de ortogonalidad en la DBIB

La variabilidad general del modelo estadístico de la DBIB tiene los mismos componentes de la distribución de bloques completos al azar³⁶, esto es: $Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$. Igualmente para los residuales (ε_{ij}) se asume que son independientes y normalmente distribuidos [$NID(0, \sigma^2)$]. Sin embargo, dado que todos los tratamientos no se repiten completamente en los bloques, la distribución pierde ortogonalidad estadística y, por tanto, la suma de cuadrados de tratamientos deberá ser ajustada por las condiciones en que aparecen en los bloques. Con este ajuste se trata de separar el efecto de los tratamientos del efecto de los bloques, dado que los primeros no se comparan en igualdad de condiciones dentro de cada bloque. En este caso, la variabilidad total, vista como suma de cuadrados (SC), está compuesta de la siguiente forma:

$$SC_{total} = SC_{trat.(ajustados)} + SC_{bloques} + SC_{error\ exp.}$$

La suma de cuadrados de tratamientos ajustados [$SC_{t(aj)}$] se define con la expresión 5.2, indicada a continuación:

$$SC_{t(aj)} = \frac{k \sum Q_i^2}{\lambda t} \quad (5.2)$$

³⁶ Para recordar: en una distribución de bloques completos al azar (DBCA) de un experimento uni-factorial (o uni-variado), una respuesta Y_{ij} está integrada de un efecto constante μ , el efecto del tratamiento τ_i , el efecto de los bloques β_j y el efecto del error experimental ε_{ij} . Por ello, $Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$.

en donde Q_i representa el total de cada uno de los tratamientos ajustados de acuerdo a los datos existentes por bloques. Este valor se determina con la expresión 5.3 siguiente:

$$Q_i = \sum y_{i.} - \frac{\sum n_{ij}y_{.j}}{k} \quad (5.3)$$

En la expresión 5.3, $\sum n_{ij}y_{.j}$ corresponde a los totales de todos los bloques que incluyen al i -ésimo tratamiento, considerando que $n_{ij} = 1$ si el tratamiento i aparece en el bloque j , o $n_{ij} = 0$ cuando no aparece. En la tabla 5.1 se detallan los componentes del modelo de análisis de varianza para una DBIB, en el cual b representa al número de bloques de la distribución.

Tabla 5.1. Estructura del análisis de varianza para una distribución de bloques incompletos balanceados.

Fuentes de variación	GL	SC	CM	F
Total	$kb - 1$	$SC_T = \sum Y_{ij}^2 - \frac{(\sum Y_{ij})^2}{kb}$		
Tratamientos (ajustados)	$t - 1$	$SC_{t(a)} = \frac{k \sum Q_i^2}{\lambda t}$	$\frac{SC_{t(a)}}{t - 1}$	$\frac{CM_t}{CM_E}$
Bloques	$b - 1$	$SC_b = \frac{\sum (\sum Y_{.j})^2}{k} - \frac{(\sum Y_{ij})^2}{kb}$	$\frac{SC_b}{b - 1}$	
Error experimental	$kb - t - b + 1$	$SC_E = SC_T - SC_{t(a)} - SC_b$	$\frac{SC_E}{kb - t - b + 1}$	

Nótese en la tabla 5.1 que, al tener bloques completos ($k = t$), el procedimiento de cálculo de los grados de libertad del error experimental se reduce a la forma de deducción en las DBCA; esto es: $(t - 1)(b - 1)$.

Ejemplo de aplicación 5.1

Se efectuó un experimento para valorar la residualidad del nitrato (NO_3) de cuatro tipos de fertilizantes (tratamientos) comerciales en el agua subterránea producto de la percolación de piscinas de arroz. Se estableció un ancho de 10 m para cada parcela y se consideró una DBIB con tres tratamientos por bloque, debido a que la extensión del terreno no permitía colocar todos los tratamientos dentro de un mismo bloque tal como en una DBCA. Se evaluó la presencia de NO_3 en ppm en el agua percolada. Los datos se describen en la tabla 5.2.

Tabla 5.2. Datos de nitratos (ppm) de un ensayo bajo una DBIB del ejemplo de aplicación 5.1.

Nº	Tratamientos (fertilizantes)	Bloques				Total de tratamientos
		I	II	III	IV	
1	Urea	58	62		59	179
2	Nitrato de amonio		42	38	39	119
3	Sulfato de amonio	33	29	35		97
4	Nitrato de calcio	48		54	43	145
Total de bloques		139	133	127	141	540

En este ejemplo, los parámetros de la DBIB son los siguientes: $k = 3$ (unidades experimentales/bloque), $\lambda = 2$ (número de bloques en que un par de tratamientos aparecen juntos) $b = 4$ (número de bloques del ensayo) y $r = 3$ (número de repeticiones de los tratamientos).

La suma de cuadrados del total está dada por:

$$SC_T = \sum Y_{ij}^2 - \frac{(\sum Y_{ij})^2}{kb} = (58^2 + 33^2 + \dots + 43^2) - \frac{540^2}{3 \times 4} = 1342$$

Para obtener la suma de cuadrados de tratamientos, previamente deben ajustarse los totales de cada tratamiento utilizando la expresión 5.3. Es decir:

$$Q_1 = 179 - \frac{139 + 133 + 141}{3} = 41,33$$

$$Q_2 = 119 - \frac{133 + 127 + 141}{3} = -14,67$$

$$Q_3 = 97 - \frac{139 + 133 + 127}{3} = -36$$

$$Q_4 = 145 - \frac{139 + 127 + 141}{3} = 9,33$$

Utilizando estos valores ajustados y la expresión 5.2, la suma de cuadrados de tratamientos ajustados es:

$$SC_{t(a_j)} = \frac{k \sum Q_i^2}{\lambda t} = \frac{3[41,33^2 + (-14,67)^2 + (-36)^2 + 9,33^2]}{2 \times 4} = 1240$$

La suma de cuadrados de bloques es:

$$SC_b = \frac{\sum (\sum Y_j)^2}{k} - \frac{(\sum Y_{ij})^2}{tb} = \frac{139^2 + 133^2 + 127^2 + 141^2}{3} - \frac{540^2}{3 \times 4} = 40$$

Y la suma de cuadrados del error es:

$$SC_{EE} = SC_T - SC_{t(a_j)} - SC_b = 1342 - 1240 - 40 = 62$$

Tabla 5.3. Análisis de varianza para una DBIB del ejemplo de aplicación 5.1.

Fuentes de variación	GL	SC	CM	F	P-valor
Total	11	1342			
Tratamientos (ajustados)	3	1240	413,33	33,33	0,001
Bloques (sin ajuste)	3	40	13,33	1,08	0,439
Error experimental	5	62	12,40		

En este caso, el coeficiente de determinación del modelo de análisis de varianza se obtiene por:

$$r^2 = \frac{SC_{t(a_j)} + SC_b}{SC_T}$$

Para el ejercicio, este valor es:

$$r^2 = \frac{1240 + 40}{1342} = 0,954$$

Asimismo, el error estándar de las medias (S_Y), según este tipo de diseño, se determina mediante la expresión:

$$S_Y = \sqrt{\frac{CM_E}{rt} \left[1 + \frac{kr(t-1)}{\lambda t} \right]} \quad (5.4)$$

El ajuste por mínimos cuadrados para las medias de tratamientos se define mediante la expresión 5.5 indicada a continuación:

$$\bar{Y}_i = \frac{\sum Y}{n} + \frac{kQ_i}{\lambda t} \quad (5.5)$$

Para los cuatro tratamientos de este ejemplo, los promedios ajustados son:

$$\begin{aligned} \bar{Y}_1 &= \frac{540}{12} + \frac{3 \times 41,33}{2 \times 4} = 60,50 \\ \bar{Y}_2 &= \frac{540}{12} + \frac{3 \times (-14,67)}{2 \times 4} = 39,50 \\ \bar{Y}_3 &= \frac{540}{12} + \frac{3 \times (-36,0)}{2 \times 4} = 31,50 \\ \bar{Y}_4 &= \frac{540}{12} + \frac{3 \times 9,33}{2 \times 4} = 48,50 \end{aligned}$$

Cuando se requiera realizar un ajuste de la suma de cuadrados de los bloques (cuando éste es de interés principal), deberá procederse de forma análoga a la obtención de la suma de cuadrados ajustados para tratamientos, considerando que la suma de cuadrados del total se desagregue en la forma siguiente:

$$SC_{total} = SC_{trat.} + SC_{bloques(ajustados)} + SC_{error exp.}$$

En donde la suma de cuadrados de los bloques ajustados se obtiene mediante la expresión 5.6 expuesta a continuación:

$$SC_{b(a_j)} = \frac{r \sum Q_j^2}{\lambda b} \quad (5.6)$$

El valor de Q_j se determina mediante la expresión 5.7 que se indica seguidamente:

$$Q_j = \sum y_{.j} - \frac{\sum n_{ij}y_{it}}{r} \quad (5.7)$$

En las tablas 5.4 hasta la 5.6 se detallan el análisis de varianza, las medias ajustadas con su respectivo error estándar y la prueba de Tukey, respectivamente, obtenidos mediante el programa Minitab. Nótese en la tabla 5.4 el ajuste realizado también a los bloques, cuya suma de cuadrados difiere del indicado en la tabla 5.3 en donde el efecto de los bloques no está ajustado. Nótese también que, al ajustar las sumas de cuadrados de tratamientos y de bloques, la suma de cuadrados del total (1342,0) no coincide con la sumatoria: $SC_{t(a_j)} + SC_{b(a_j)} + SC_{EE}$, cuyo valor es 1336, 67. Esto último sucede debido a la falta de ortogonalidad que tiene este tipo de distribución experimental.

Tabla 5.4. Análisis de varianza obtenido con Minitab.

Fuentes de variación	GL	SC Ajust.	MC Ajust.	F	P-valor
Tratamientos	3	1240,00	413,33	33,33	0,001
Bloques	3	34,67	11,56	0,93	0,490
Error exp.	5	62,00	12,40		
Total	11	1342,00			

Tabla 5.5. Medias ajustadas y error estándar de la media obtenidos con Minitab.

Término	Media ajustada	Error estándar de la media
Tratamientos		
T1: Urea	60,50	2,13
T2: Nitrato de amonio	39,50	2,13
T3: Sulfato de amonio	31,50	2,13
T4: Nitrato de calcio	48,50	2,13

Tabla 5.6. Aplicación de la prueba de Tukey (p<0,05) realizada con Minitab.

Tratamientos	N	Media	Agrupación
T1: Urea	3	60,5	a
T4: Nitrato de calcio	3	48,5	b
T2: Nitrato de amonio	3	39,5	b c
T3: Sulfato de amonio	3	31,5	c

Las medias que no comparten una letra son significativamente diferentes.

La eficiencia relativa de la DBIB respecto de las distribuciones de bloques completos aleatorizados (DBCA) se obtiene a partir de la relación de varianzas del error experimental (CM_E) de estas dos estructuras de ensayo. No obstante, establecer el valor de esta relación es imposible si se pretendiera obtener el CM_E de una DBCA a partir de los datos de una DBIB,

salvo que se desarrolle el ensayo bajo una DBCA de forma alterna para realizar esta comparación. Sin embargo, dado el propósito de la DBIB, es esperable que el CM_E de esta distribución sea más pequeño que el CM_E de una DBCA con el mismo número de tratamientos y repeticiones; con lo que se tendría una mayor precisión en la detección de diferencias entre los tratamientos que se estén evaluando en el ensayo. En este sentido, lo único que podríamos establecer al realizar un experimento bajo una DBIB es la relación potencial entre los cuadrados medios de ésta distribución con la DBCA, a partir de la cual se declare la primera como de mayor precisión. Esta relación es:

$$\frac{CM_{EE(DBIB)}}{CM_{EE(DBCA)}} < \frac{\lambda t}{rk} \quad (5.8)$$

Es decir, si el error experimental no tiene una “carga extra” debido a un mal manejo del ensayo, un experimento llevado a cabo bajo una DBIB es más preciso que si éste se realizara bajo una DBCA (considerando el mismo número de tratamientos y repeticiones), siempre que se cumpla la desigualdad dada por la expresión 5.8.

Distribución en bloques incompletos parcialmente balanceados

Tal como se ha mencionado al inicio del tema sobre las distribuciones de bloques incompletos balanceados, estas distribuciones tienen sus límites en el sentido de que, al buscar el balanceo en los tratamientos, se generará una cantidad elevada de bloques que limitarán su aplicación en términos económicos; y que, además, harán que este tipo de ensayo pierda el propósito de su utilidad: disminuir la varianza del error experimental al permitir evaluar bloques más pequeños y consecuentemente hacer que el diseño experimental tenga mayor precisión. A esta limitación que podría ser real en una condición dada, se une otra desventaja en el sentido de que las DBIB se estructuran sólo para determinados parámetros que deben converger para un valor λ (parejas de tratamientos en bloques) el cual debe ser un entero.

Ante la situación planteada en el párrafo anterior y dada la necesidad de controlar o disminuir la varianza del error experimental en experimentos con un número importante de tratamientos, una alternativa a la DBIB es la distribución de bloques incompletos parcialmente balanceados (DBIPB). Si bien esta distribución comparativamente podría presentar una menor precisión que la DBIB, esta diferencia en precisión podría también no ser tan importante como para desmerecer la DBIPB, a sabiendas que ésta última ofrece la gran ventaja de correr un experimento con un número menor de repeticiones (o menor cantidad de bloques). Este número menor de bloques ocurre porque el parámetro λ , que es constante en la DBIB, en la DBIPB es variable por el mismo requerimiento de reducir las repeticiones r del ensayo, de tal manera que se genere un número de unidades experimentales que sea manejable en la práctica. En la DBIPB pueden existir λ_1 hasta λ_n que indicarán diferentes parejas de tratamientos en los bloques del ensayo.

La “gran” ventaja de las distribuciones de bloques incompletos parcialmente balanceados

La gran ventaja de las DBIPB sobre las DBIB es la de poder correr un ensayo con menos cantidades de bloques. Esta característica de este tipo de distribución es lo que la define como parcialmente balanceado, debido a que habrá diferentes pares de tratamientos que aparecerán en uno, en dos, en tres o en más bloques, que al final causarán una comparación injusta de los tratamientos. En la tabla 5.7 se ha detallado el esquema de un experimento bajo una DBIPB con los siguientes parámetros: $t = 8$, $r = 3$, $b = 6$ y $k = 4$. En este ejemplo, la pareja de tratamientos (4, 7) aparece en tres bloques del ensayo, hay nueve pares que aparecen en dos

bloques y el resto de parejas sólo aparecen en un solo bloque; es decir, las comparaciones de tratamientos que se realicen tendrán distintas precisiones, siendo en este caso el par (4, 7) el que se valore con mayor consistencia.

Tabla 5.7. Esquema de un experimento bajo una DBIPB.

Bloques	Tratamientos			
1	5	6	8	2
2	5	7	3	4
3	1	4	7	8
4	1	3	2	6
5	7	4	2	8
6	1	6	3	5

Un aspecto importante para las DBIPB son las clases asociadas de tratamientos en el arreglo; resultando que, cada tratamiento, podría pertenecer a dos o más clases asociadas. Entiéndase como clase asociada al grupo de tratamientos en el que cada par asociado ocurre en λ_i bloques (Kuehl, 2001). En la tabla 5.8 se han desglosado los grupos asociados según el número de bloques en que aparecen (λ_i), así se tienen los primeros tratamientos asociados (4, 7), los cuales aparecen en tres bloques; así también se indican los que están asociados en dos bloques (λ_2) y los asociados que sólo aparecen en un solo bloque (λ_3).

Tabla 5.8. Grupos de asociados en 3, 2 y 1 bloques.

Tratamientos	$\lambda_1 = 3$	$\lambda_2 = 2$	$\lambda_3 = 1$
1		3, 6	2, 4, 5, 7, 8
2		6, 8	1, 3, 4, 5, 7
3		1, 5, 6	2, 4, 7, 8
4	7	8	1, 2, 3, 5
5		6	1, 2, 4, 7, 8
6		1	8
7		8	1, 2, 3, 5
8		7	1, 3, 5, 6

Sumas de cuadrados en la DBIPB

Igual como en la DBIB, la falta de ortogonalidad de los tratamientos en los bloques exige que las sumas de cuadrados de tratamientos se ajusten. De manera general, la suma de cuadrados de tratamientos ajustada se deduce por diferenciación entre la suma de cuadrados del modelo completo ($Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$) y la que se obtiene obviando el efecto de tratamientos en el modelo, llamado éste como modelo reducido ($Y_j = \mu + \beta_j + \varepsilon_j$). Los cálculos presentan cierto proceso engorroso dado que, si bien los bloques no son ortogonales, también habrá que considerar la diversidad de clases asociadas a cada tratamiento, con lo cual se tendrán una serie de cálculos diferenciados para cada experimento realizado bajo una DBIPB. Para facilitar esta situación actualmente existen programas que, utilizando procesos de modelos lineales, facilitan la obtención de los análisis de varianza, así como las medias de los tratamientos ajustadas y los errores estándar correspondientes.

Ejemplo de aplicación 5.2

En una zona determinada se requirió la evaluación de ocho variedades de papa en un experimento. Para este propósito se utilizó una distribución de bloques incompletos parcialmente balanceados debido al poco espacio de terreno disponible que no permitía extender bloques completos, además de que se generaban una cantidad importante de bloques si se buscaba el balanceo del ensayo. Con los datos que se detallan en la tabla 5.9, realizar el análisis de varianza y la prueba de comparación correspondiente.

Tabla 5.9. Datos del rendimiento de ocho variedades de papa (kg/parcela) bajo una DBIPB.

Tratamientos (Variedades)	Bloques						Total
	1	2	3	4	5	6	
1	14,2	10,9	13,7				38,8
2	15,1	12,7		15,8			43,6
3	20,7		22,0		24,1		66,8
4	17,8			15,4	13,9		47,1
5		26,8	27,3			24,8	78,9
6		22,4		17,2		19,9	59,5
7			25,9		23,6	21,8	71,3
8				12,1	9,8	10,1	32,0
Total	67,8	72,8	88,9	60,5	71,4	76,6	438,0

Parámetros del ensayo: $t = 8, b = 6, r = 3, k = 4$

La tabla 5.10 es el análisis de varianza obtenido mediante Minitab con el ajuste de las sumas de cuadrados de tratamientos en base a los bloques y de los bloques en base a los tratamientos. Nótese que el total de las sumas de cuadrados (713,74) no coincide con la suma de los valores registrados de las sumas de cuadrados de tratamientos, bloques y error experimental (559,49+12,50+41,34), esto debido a la falta de ortogonalidad del diseño. En este caso, los tratamientos presentan diferencias altamente significativas ($p < 0,01$), mientras que los bloques muestran efectos homogéneos. Esta falta de variabilidad significativa en los bloques, de acuerdo a la magnitud registrada del estadístico de prueba ($F = 0,67$), faculta para que el análisis se realice bajo la consideración de una distribución completamente al azar, la misma que queda como tarea para que el lector la realice.

Tabla 5.10. Análisis de varianza de los datos del ejemplo de aplicación 5.2 sobre un experimento realizado bajo una DBIPB (salida de Minitab).

Fuentes de variación	GL	SC Ajust.	MC Ajust.	F	P-valor
Tratamientos	7	559,49	79,927	21,27	0,000
Bloques	5	12,50	2,500	0,67	0,658
Error exp.	11	41,34	3,758		
Total	23	713,74			

Resumen del modelo de análisis de varianza
(salida de Minitab).

Error estándar	r-cuad.	r-cuad. (ajustado)	r-cuad. (pred)
1,93856	94,21%	87,89%	72,43%

En la tabla 5.11 se detalla la clasificación de medias realizada con la prueba de Tukey, en el que se observan cuatro grupos definidos (a, b, c y d), ubicando a los tratamientos 5, 7, 3 y 6 en el primer grupo con los valores estadísticamente significativos ($p < 0,05$) respecto del resto de tratamientos. Es preciso destacar también que las medias indicadas son valores ajustados dada la falta de ortogonalidad del diseño. También se resalta la bondad del modelo, cuyo coeficiente de determinación es de 94,21%.

Tabla 5.11. Prueba de Tukey para las medias del ejemplo de aplicación 5.2 (salida de Minitab).

Tratamientos	N	Media	Agrupación
T5	3	26,4417	a
T7	3	23,9333	a
T3	3	21,6417	a b
T6	3	20,4583	a b c
T4	3	15,5583	b c d
T2	3	14,3667	c d
T1	3	12,2833	d
T8	3	11,3167	d

Distribución en Lattice

La distribución en lattice (DL) es un tipo de asignación experimental de tratamientos que se creó para utilizarse en el área de la investigación agrícola, específicamente en ensayos sobre cruces genéticos de plantas y comparación de variedades, aunque su uso no es tan común. Es una distribución que está indicada para experiencias con un número elevado de tratamientos porque tiene la característica de controlar de mejor forma la variabilidad o heterogeneidad del suelo que crece a medida que un experimento se extiende en el campo. Presenta una mayor precisión que la distribución clásica utilizada extensivamente como lo es la distribución de bloques completos al azar (DBCA), dado que en esta distribución los bloques no contienen a todos los tratamientos, de allí que estos últimos tengan una menor extensión lateral y, por tanto, un mejor control de la varianza del error experimental.

Como la característica principal de una DL es la de constituirse en bloques con un número de k unidades experimentales por bloque menores a los t tratamientos que se prevén evaluar en un ensayo, es una estructura experimental que se conoce también como *distribución de bloques incompletos*. Por esta característica de la distribución en lattice de formar bloques relativamente reducidos con $k < t$, se hace posible tener áreas más pequeñas (en cuanto a los bloques) con grupos de parcelas uniformes o de menor heterogeneidad respecto del suelo. En este sentido, la DL permite evaluar los efectos de los tratamientos con una mayor precisión que aquella que se obtiene cuando se utilizan bloques demasiado extendidos al utilizar una distribución como la DBCA. Debe entenderse que los bloques incompletos aquí mencionados son todos del

mismo tamaño; es decir, para un experimento específico, todos los bloques tiene la misma cantidad de k unidades experimentales.

La repartición aleatoria de las unidades experimentales en el campo de una distribución en latice, o lo que se conoce como el croquis de campo del diseño, requiere de un proceso muy parecido a la asignación de tratamientos que se realiza en una distribución en cuadro latino. Al respecto de esto, para el uso de este tipo de distribución es necesario que la cantidad de tratamientos sean cuadrados perfectos, tales como 16, 25, 36, 49, 64, etc., con el propósito de que el número de b bloques de cada repetición sea equivalente a las k unidades experimentales por bloque. Esta última condición define la igualdad de $t = k^2$.

Para la asignación de los tratamientos en el campo partimos de dos permutaciones fundamentales clasificadas como grupos básicos X e Y . Estos grupos representan repeticiones completas del ensayo, formadas de bloques en donde se distribuirá equitativamente una parte del grupo de tratamientos. La cantidad de tratamientos que se ubican en cada bloque (k unidades experimentales) debe ser equivalente a \sqrt{t} (siendo t el número total de tratamientos), por lo que cada repetición contará con un número de bloques también equivalente a \sqrt{t} .

Los grupos básicos X e Y a que se hace referencia en las distribuciones en latice son permutaciones horizontales y verticales de la distribución de los tratamientos dentro de estos grupos, respectivamente. También puede existir un tercer grupo básico que se representa con Z , que no es más que una permutación sistemática obtenida entre el montaje del primer grupo básico X y un tercer grupo en que se considera una distribución en cuadro latino (con letras latinas representado a los tratamientos) que tiene una permutación vertical, definiendo los bloques en orden a los que coinciden con las letras A, B, C, etc., según sea el número de tratamientos. Según este orden, si la distribución contiene sólo a los grupos X e Y se la denomina como *distribución en latice simple* (DLS); y si contiene los tres grupos básicos mencionados, se denomina *distribución en latice triple* (DLT). Asimismo, se hace preciso indicar que, a diferencia de las DBCA, repetición y bloque son dos términos distintos para este tipo de distribución.

Es importante indicar que las distribuciones en latice son estructuras que están incluidas entre las distribuciones de *bloques incompletos resolubles* que se indican más adelante, propuestos principalmente por Yates (1936). Para lograr un balance estricto en estas distribuciones es necesario que se tengan $r = k + 1$ repeticiones y $b = k(k + 1)$ bloques, logrando con esto un balanceo de $\lambda = 1$ (un par único de tratamientos por bloque); además de que las k unidades experimentales por bloque deberán ser números primos o potencias de números primos (Kuehl, 2001). En ciertas situaciones el carácter de resoluble de estas distribuciones le confiere una ventaja en el sentido de que un experimento puede realizarse por etapas (repeticiones) cuando no pueda desarrollarse completamente al mismo tiempo, debido a restricciones económicas o de limitaciones en el material experimental.

Un aspecto que podría verse como restrictivo en el caso de las DL, especialmente cuando el material experimental es escaso (por ejemplo, pocas semillas), es el sentido del balanceo, pues como se ha mencionado en el párrafo anterior, éste es posible cuando se tienen $k + 1$ repeticiones de la distribución. Sin embargo, dado que se trata de distribuciones resolubles, es posible aminorar el número de repeticiones, obteniéndose así las *distribuciones en latice parcialmente balanceadas*, de la cual se hace énfasis en este texto y se expone un ejemplo más adelante.

Como se ha visto, las distribuciones en latice son estructuras cuadráticas de ubicación de los tratamientos o unidades experimentales, que son equivalentes al número s de bloques de una repetición del ensayo ($k = s$). Sin embargo, también existen distribuciones en latice que tienen forma rectangular, que se establecen considerando $k = s - 1$ y en donde el número de

t tratamientos está dado como $t = s(s - 1)$. Al respecto, varias estructuras con distribuciones para latices cuadrados - balanceados y latices rectangulares se incluyen en las figuras A10 y A11 del anexo de este texto, tomadas de Cochran y Cox (1978).

¿Cómo se aleatorizan los tratamientos en una distribución en latices?

En el caso de un experimento con 25 tratamientos, las permutaciones horizontales y verticales de los grupos básicos X e Y son las que se indican en los siguientes cuadros:

Repetición I – Grupo básico X

Bloques	a	1	2	3	4	5
	b	6	7	8	9	10
	c	11	12	13	14	15
	d	16	17	18	19	20
	e	21	22	23	24	25

Repetición II – Grupo básico Y

Bloques	a	1	6	11	16	21
	b	2	7	12	17	22
	c	3	8	13	18	23
	d	4	9	14	19	24
	e	5	10	15	20	25

En cada grupo básico se sortean, tanto los bloques dentro de cada repetición como los tratamientos dentro de cada bloque. Por lo tanto, una vez realizado el respectivo sorteo de bloques y de tratamientos, una opción de distribución podría ser la siguiente:

Repetición I - Grupo básico X (Sorteado)

Bloques	c	12	13	11	15	14
	e	24	23	21	25	22
	b	9	10	7	6	8
	a	4	1	5	3	2
	d	18	20	16	17	19

Repetición II - Grupo básico Y (Sorteado)

Bloques	c	18	13	3	23	8
	a	6	11	1	16	21
	e	20	5	15	25	10
	b	17	12	7	22	2
	d	9	4	14	24	19

Los grupos básicos (X e Y) se repiten el mismo número de veces, según se haya previsto desde el contexto estadístico o en función de las limitaciones reales del ensayo. Algo que debe tenerse en cuenta es que una repetición (grupo básico) debe establecerse completa dentro de

un mismo lote, de tal forma que no exista una pérdida de eficacia del experimento y aprovechar esta ventaja que es una de las que ofrece este tipo de distribución experimental.

Ejemplo de aplicación 5.3

Se ha evaluado el rendimiento de 16 híbridos de maíz utilizando una distribución en latice simple parcialmente balanceada con dos grupos básicos de 4 x 4 y con 4 repeticiones, es decir, dos repeticiones por cada grupo. La variable fue medida en unidades de kg/parcela. Con la información respectiva dada en los siguientes cuadros, realizar el análisis de varianza.

Considérese en los cuadros que los valores superiores de cada celda representan al número de identificación de los tratamientos y los valores inferiores a la variable medida.

Grupo X - Repetición I					
Bloques					Total
1	11	9	12	10	18,1
	5,5	4,3	3,2	5,1	
2	3	4	1	2	13,8
	2,7	3,7	4,2	3,2	
3	14	16	15	13	15,9
	4,1	3,4	4,7	3,7	
4	6	7	5	8	15,8
	4,9	3,6	4,5	2,8	
Total	17,2	15,0	16,6	14,8	63,6

Grupo Y - Repetición II					
Bloques					Total
1	14	10	2	6	17,4
	3,9	4,9	3,5	5,1	
2	16	8	4	12	13,6
	3,6	3,1	3,3	3,6	
3	13	5	1	9	16,4
	3,4	4,4	4,7	3,9	
4	7	3	15	11	11,7
	3,2	2,6	4,1	5,0	
Total	14,1	15,0	15,6	17,6	62,3

Grupo X - Repetición III					
Bloques					Total
1	13	16	15	14	13,0
	2,8	3,2	4,1	2,9	
2	2	4	1	3	14,1
	3,8	4,1	3,1	3,1	
3	9	11	12	10	17,2
	4,6	4,9	2,9	4,8	
4	6	5	8	7	14,9
	4,3	4,1	3,4	3,1	
Total	15,5	16,3	13,5	13,9	59,2

Grupo Y - Repetición IV

Bloques					Total
1	6	14	10	2	17,4
	4,8	3,8	5,1	3,7	
2	5	13	9	1	16,2
	4,6	3,7	3,7	4,2	
3	8	16	12	4	12,5
	2,9	3,2	3,2	3,2	
4	7	15	11	3	14,9
	3,5	3,9	5,3	2,2	
Total	15,8	14,6	17,3	13,3	61,0

La información ordenada por tratamiento y repetición se detalla en la tabla 5.12.

Tabla 5.12. Datos (kg/parcela) ordenados del ejemplo de aplicación 5.3.

Nº Trat.	Repeticiones				Total	Promedios
	I	II	III	IV		
1	4,2	4,7	3,1	4,2	16,2	4,05
2	3,2	3,5	3,8	3,7	14,2	3,55
3	2,7	2,6	3,1	2,2	10,6	2,65
4	3,7	3,3	4,1	3,2	14,3	3,58
5	4,5	4,4	4,1	4,6	17,6	4,40
6	4,9	5,1	4,3	4,8	19,1	4,78
7	3,6	3,2	3,1	3,5	13,4	3,35
8	2,8	3,1	3,4	2,9	12,2	3,05
9	4,3	3,9	4,6	3,7	16,5	4,13
10	5,1	4,9	4,8	5,1	19,9	4,98
11	5,5	5,0	4,9	5,3	20,7	5,18
12	3,2	3,6	2,9	3,2	12,9	3,23
13	3,7	3,4	2,8	3,7	13,6	3,40
14	4,1	3,9	2,9	3,8	14,7	3,68
15	4,7	4,1	4,1	3,9	16,8	4,20
16	3,4	3,6	3,2	3,2	13,4	3,35
Total	63,6	62,3	59,2	61,0	246,1	

El análisis de varianza, considerando los datos ordenados de la tabla 5.12 y una distribución en bloques completos al azar, es el siguiente:

Tabla 5.13. Análisis de varianza considerando una DBCA.

Fuentes de variación	GL	SC	CM	F	P-valor
Total	63	37,40			
Tratamientos	15	31,15	2,08	17,33	< 0,01
Repeticiones	3	0,66	0,22	1,83	0,16
Error experimental	45	5,59	0,12		

Coefficiente de variación (CV): 9,12%

Al realizar el ANDEVA asumiendo una distribución de bloques completos al azar, el coeficiente de variación revela que el ensayo merece un nivel alto de confianza o, dicho de otra manera, ha habido un control exhaustivo del error experimental. Sin embargo, la precisión de este análisis puede aumentar si se incluye en el proceso la variación de los bloques incompletos, entre los cuales pudiera existir una fuente adicional de variación que no se ha considerado en el análisis de la tabla 5.13. Si se descuenta la variación respecto de los bloques

incompletos, la varianza del error experimental debe reducirse y, consiguientemente, el coeficiente de variación.

El análisis de una DLS separa la variación generada por los bloques incompletos, por tanto, habrá que realizar el análisis para la variación de éstos dentro de un mismo grupo básico identificado como *Componente A*. Así también se tiene que analizar la variación de los bloques incompletos, pero entre distinto grupo básico, efecto que se conoce como *Componente B*. La variación conjunta de estos dos componentes forma la variación atribuible a los *bloques incompletos*, que es la fuente que también debe considerar en el análisis de estas distribuciones. En lo que sigue se exponen los cálculos respectivos de cada una de estas fuentes de variación adicionales, representadas por sus respectivas sumas de cuadrados.

Cálculo de la suma de cuadrados del componente A

Para el cálculo de esta fuente de variación se deben obtener las diferencias entre las sumas de los bloques incompletos que contengan a los mismos tratamientos, dentro de un mismo grupo básico; es decir:

Grupos X

Repetición I		Repetición III		Diferencias (D)
Nº Bloque	Total de bloque	Nº Bloque	Total de bloque	
1	18,1	3	17,2	0,9
2	13,8	2	14,1	- 0,3
3	15,9	1	13,0	2,9
4	15,8	4	14,9	0,9
Total	63,6	Total	59,2	4,4

Grupos Y

Repetición II		Repetición IV		Diferencias (D)
Nº Bloque	Total de bloque	Nº Bloque	Total de bloque	
1	17,4	1	17,4	0,0
2	13,6	3	12,5	1,1
3	16,4	2	16,2	0,2
4	11,7	4	14,9	- 3,2
Total	59,1	Total	61,0	- 1,9

La suma de cuadrados para la componente A se obtiene con la siguiente expresión:

$$SC_{com-A} = \frac{\sum D^2}{R_G \cdot t_G} - \frac{(\sum D_X)^2 + (\sum D_Y)^2}{R_G \cdot t_G^2} \quad (5.9)$$

Siendo SC_{com-A} la suma de cuadrados para el componente A, D es la diferencia entre el total de bloques similares de un mismo grupo (X e Y), R_G es el número de repeticiones de cada grupo básico, t_G es el número de tratamientos en cada bloque incompleto, $\sum D_X$ es la sumatoria

algebraica de las diferencias del grupo básico X y $\sum D_Y$ es la sumatoria algebraica de las diferencias del grupo básico Y . Para este ejemplo:

$$SC_{com-A} = \frac{0,9^2 + (-0,3)^2 + \dots + 0,2^2 + (-3,2)^2}{2 \times 4} - \frac{4,4^2 + (-1,9)^2}{2 \times 4^2}$$

$$SC_{com-A} = 1,98$$

Cálculo de la suma de cuadrados del componente B

En el cálculo de la suma de cuadrados de la componente B es necesario obtener las diferencias entre las columnas del grupo básico Y con las sumatorias de los mismos tratamientos, y las filas del grupo básico X , también con las sumatorias de los mismos tratamientos.

La suma de los mismos tratamientos del grupo Y (tratamientos ordenados por columna) son los que se detallan en el cuadro siguiente, considerando los valores correspondientes a la repetición II y IV. Por ejemplo, el valor del casillero 1 es la suma de los valores correspondientes del tratamiento 1, es decir de 4,7 y 4,2.

	Tratamientos / Grupos Y				Total filas
	1	2	3	4	
1	1 8,9	5 9,0	9 7,6	13 7,1	32,6
2	2 7,2	6 9,9	10 10,0	14 7,7	34,8
3	3 4,8	7 6,7	11 10,3	15 8,0	29,8
4	4 6,5	8 6,0	12 6,8	16 6,8	26,1
Total columnas	27,4	31,6	34,7	29,6	123,3

	Tratamientos / Grupos X				Total filas
	1	2	3	4	
1	1 7,3	2 7,0	3 5,8	4 7,8	27,9
2	5 8,6	6 9,2	7 6,7	8 6,2	30,7
3	9 8,9	10 9,9	11 10,4	12 6,1	35,3
4	13 6,5	14 7,0	15 8,8	16 6,6	28,9
Total columnas	31,3	33,1	31,7	26,7	122,8

Las diferencias (D) se obtienen de la resta del total de columna menos el total de fila (o viceversa) de los grupos X e Y .

Grupo Y		Grupo X		Diferencias [$D_{C(Y)F(X)}$]
Columnas	Total columnas	Filas	Total filas	
1	27,4	1	27,9	-0,5
2	31,6	2	30,7	0,9
3	34,7	3	35,3	-0,6
4	29,6	4	28,9	0,7
Total	123,3	Total	122,8	0,5

Grupo X		Grupo Y		Diferencias [$D_{C(X)F(Y)}$]
Columnas	Total columnas	Filas	Total filas	
1	31,3	1	32,6	-1,3
2	33,1	2	34,8	-1,7
3	31,7	3	29,8	1,9
4	26,7	4	26,1	0,6
Total	122,8	Total	123,3	-0,5

Asimismo, la suma de cuadrados para la componente B se obtiene mediante la expresión:

$$SC_{com-B} = \frac{\sum D_{C(Y)F(X)}^2 + \sum D_{C(X)F(Y)}^2}{R \cdot t_G} - \frac{(\sum D_{C(Y)F(X)})^2 + (\sum D_{C(X)F(Y)})^2}{R \cdot t_G^2} \quad (5.10)$$

En la cual SC_{com-B} es la suma de cuadrados para el componente B; D son las diferencias entre el total de tratamientos similares del grupo X y el total de tratamientos del grupo Y , considerando el orden de columnas menos las filas; R es el número de repeticiones del experimento; t_G es el número de tratamientos en cada bloque incompleto; $\sum D_{C(Y)F(X)}$ es la sumatoria algebraica de las diferencias de los grupos básicos Y y X , considerando las columnas menos las filas, respectivamente; y $\sum D_{C(X)F(Y)}$ es la sumatoria algebraica de las diferencias de los grupos básicos X e Y , considerando las columnas menos las filas, respectivamente. En consecuencia, la suma de cuadrados es:

$$SC_{com-B} = \frac{(-0,5)^2 + 0,9^2 + \dots + 1,9^2 + 0,6^2}{4 \times 4} - \frac{(0,5)^2 + (-0,5)^2}{4 \times 4^2}$$

$$SC_{com-B} = 0,65$$

Los grados de libertad para cada componente se determinan considerando los grupos básicos (en este caso son 2) y los tratamientos que forman un bloque incompleto dentro de cada grupo, es decir:

$$GL_{com-A/B} = G(t_G - 1)$$

Siendo G el número de grupos básicos.

$$GL_{com-A} = 2(4 - 1) = 6$$

$$GL_{com-B} = 2(4 - 1) = 6$$

Finalmente, el análisis de varianza, considerando las componentes A y B respecto de la variabilidad atribuida a los bloques incompletos, propia para una distribución en latice simple, es:

Tabla 5.14. Análisis de varianza con los datos del ejemplo de aplicación 5.3 considerando una DLS.

Fuentes de variación	GL	SC	CM	F	P-valor
Total	63	37,40			
Tratamientos	15	31,15	2,08	23,11	< 0,01
Repeticiones	3	0,66	0,22	2,44	0,082
Bloques incompletos	12	2,63	0,22	2,44	0,021
Componente A	6	1,98			
Componente B	6	0,65			
Error experimental	33	2,96	0,09		

Coefficiente de variación (CV): 7,89%

Al comparar los análisis de varianza, el realizado inicialmente considerando una distribución en bloques completos al azar (tabla 5.13) con el de la distribución en latice simple (tabla 5.14), se puede notar una disminución de la varianza desconocida (cuadrado medio del error experimental); resultado que se revela en la disminución del coeficiente de variación, el cual pasó de 9,12% a 7,89%. Bajo esta apreciación, es deducible el aumento de la precisión en la distribución en latice simple con relación a distribuciones como la DBCA, cuya magnitud de dicha precisión vendrá dada por la extensión que tendrían los bloques, los cuales se definirán en función del número de tratamientos.

Ajuste de los promedios en una distribución en latice

Debido a la variación entre los bloques incompletos, así como la variación dentro de cada uno de ellos, es necesario realizar una corrección a los promedios de los tratamientos evaluados. El ajuste puede realizarse mediante la expresión 5.11 indicada a continuación:

$$\bar{X}_{i(a_j)} = \bar{X}_i + C_X + C_Y \quad (5.11)$$

siendo $\bar{X}_{i(a_j)}$ el promedio ajustado; mientras que C_X y C_Y los coeficientes de ajuste, los mismos que se determinan mediante las expresiones siguientes:

$$C_X = H \cdot D_{C(Y)F(X)} \quad (5.12)$$

$$C_Y = H \cdot D_{C(X)F(Y)} \quad (5.13)$$

donde:

$$H = \frac{4CM_B - 7CM_E}{rk(4CM_B - CM_E)}$$

en la cual, CM_B es el cuadrado medio de bloques incompletos; CM_E el cuadrado medio del error experimental; r el número de repeticiones; k el número de tratamientos por bloque incompleto; $D_{C(Y)F(X)}$ las diferencias de los grupos básicos Y e X , considerando las columnas menos las filas, respectivamente; y $D_{C(X)F(Y)}$ las diferencias de los grupos básicos X e Y , considerando las columnas menos las filas, respectivamente.

Para los datos del ejemplo, los coeficientes de ajuste según las expresiones 5.12 y 5.13 son:

$$H = \frac{4CM_B - 7CM_E}{RK(4CM_B - CM_E)} = \frac{4 \times 0,22 - 7 \times 0,09}{4 \times 4 \times (4 \times 0,22 - 0,09)} = 0,0198$$

$D_{C(Y)F(X)}$	$D_{C(X)F(Y)}$	H	C_X	C_Y
-0,5	-1,3	0,0198	-0,0099	-0,0257
0,9	-1,7	0,0198	0,0178	-0,0337
-0,6	1,9	0,0198	-0,0119	0,0376
0,7	0,6	0,0198	0,0139	0,0119

Utilizando la expresión 5.11 y los coeficientes de ajuste, los promedios corregidos de los tratamientos son los descritos en la tabla 5.15.

Tabla 5.15. Promedios corregidos del ejemplo de aplicación 5.3.

Nº Trat.	Promedios sin ajustar	C_X	C_Y	Promedios ajustados
1	4,05	-0,0099	-0,0257	4,0144
2	3,55	-0,0099	-0,0337	3,5064
3	2,65	-0,0099	0,0376	2,6777
4	3,58	-0,0099	0,0119	3,5770
5	4,40	0,0178	-0,0257	4,3921
6	4,78	0,0178	-0,0337	4,7591
7	3,35	0,0178	0,0376	3,4054
8	3,05	0,0178	0,0119	3,0797
9	4,13	-0,0119	-0,0257	4,0874
10	4,98	-0,0119	-0,0337	4,9294
11	5,18	-0,0119	0,0376	5,2007
12	3,23	-0,0119	0,0119	3,2250
13	3,40	0,0139	-0,0257	3,3882
14	3,68	0,0139	-0,0337	3,6552
15	4,20	0,0139	0,0376	4,2515
16	3,35	0,0139	0,0119	3,3758

Nótese en la tabla 5.15 que los valores de C_X son los mismos para cada grupo de 4 tratamientos ordenados; sin embargo, los valores de C_Y son distintos para cada grupo de 4 tratamientos. Esto se debe a la forma como se calcularon los valores de los coeficientes de ajuste (las

columnas del grupo básico Y menos las filas del grupo básico X). Para la aplicación de las pruebas de comparación múltiple se utilizarán los promedios ajustados.

El lector también puede consultar sobre las estructuras de latice rectangulares que se contraponen a la limitación de las distribuciones en latice cuadradas (k^2) explicadas en este texto. Estas distribuciones pueden revisarse en la obra “*Experimental Designs*” de Cochran y Cox (1957) para estructuras de 12, 20, 30, 42, 56, 72 y 90 tratamientos. El propósito es exponer algunas alternativas de distribución que pueden ofrecer un incremento en la precisión de un ensayo y hacerlo eficaz en la detección de efectos de tratamientos, lo que está en correspondencia con el objetivo de este libro.

Distribución Alfa - Latice

Si bien dentro de las distribuciones de bloques incompletos resolubles se encuentran las estructuras en latice para ensayos con un importante número de tratamientos, éstas tienen la “ortodoxia inquebrantable” de que deben ser cuadrados o cubos perfectos o; en el caso de latices rectangulares, estructuras $s(s - 1)$, es decir, bloques s compuestos irremediablemente con $k = s - 1$ unidades experimentales. Para contrarrestar esta limitación en estos tipos de distribución se idearon las distribuciones *alfa - latice*, difundidos por Paterson y Williams (1976). La distribución alfa - latice ofrece una mayor versatilidad en el manejo de múltiples alternativas de prueba (básicamente para evaluaciones de semillas) y fue establecido en función de una reglamentación en Inglaterra en donde se estipulaba la realización de los experimentos con un número fijo de tratamientos, dejando fuera de la subordinación de quien diseñaba el experimento la decisión de la cantidad de tratamientos que debían probarse.

Como nada en la vida es perfecto, estas distribuciones tienen la exigencia de que el número de tratamientos t deben ser múltiplos del tamaño de los bloques k ; es decir, $t = sk$. De esta forma se logran bloques de un mismo tamaño, lo que permite que estas distribuciones sigan conservando el carácter de resolubles, además de ser parcialmente balanceados. Sus autores definieron estructuras para 2, 3 y 4 repeticiones, con un número de hasta 100 tratamientos, con tamaños de bloque entre 4 y 16 unidades experimentales.

Considerando un efecto de anidamiento de los bloques en las repeticiones de la distribución, los grados de libertad y las sumas de cuadrados pueden calcularse de acuerdo al esquema de análisis de varianza detallado en la tabla 5.16.

Tabla 5.16. Esquema de análisis de varianza para una distribución alfa - latice.

Fuentes de variación	Grados de libertad	Sumas de cuadrados
Total	$tr - 1$	$\sum Y_{ij(k)}^2 - FC$
Tratamientos (sin ajustar)	$t - 1$	$\frac{\sum (\sum t_i)^2}{r} - FC$
Repeticiones	$r - 1$	$\frac{\sum (\sum r_j)^2}{t} - FC$
Bloques en repeticiones	$r(s - 1) = p - r$	$\frac{\sum (\sum p_{J(k)})^2}{k} - \frac{\sum (\sum r_j)^2}{t}$
Error experimental	$tr - t - rs + 1$	Por sustracción

s representa el número de bloques por repetición y p es el número total de bloques entre todas las repeticiones del ensayo.

Considérese que el factor de corrección (FC) de las sumas de cuadrados viene dado por $FC = (\sum Y_{ij(k)})^2 / tr$, siendo tr el total de observaciones del experimento.

El lector debe tener presente que el esquema de ANDEVA de la tabla 5.16 corresponde a un experimento cuyas repeticiones se valoran transversalmente; es decir, desde un mismo momento y dentro de un mismo lote. No obstante, dado que este tipo de distribución fue ideado para evaluar fundamentalmente la adaptabilidad genotípica de semillas, es práctica común que las repeticiones se lleven a cabo en distintas zonas agroclimáticas con el propósito de valorar el efecto de la interacción genotipo x ambiente, el mismo que debe incluirse en el modelo de análisis de varianza. Esto último facilitará la identificación de genotipos con mayor adaptación que otros en una determinada zona agroclimática.

Ejemplo de aplicación 5.4

Se evaluaron 20 genotipos de soya compuestos entre híbridos (8) y variedades (12), dentro de una distribución en alfa – latice de 4 x 5. La variable evaluada fue el rendimiento (t/ha) de grano seco, cuyos valores se indican en los cuadros adjuntos. Realizar el análisis de varianza y la comparación de medias.

Repetición I					Repetición II				
Bloques					Bloques				
1	2	3	4	5	6	7	8	9	10
T ₁₈	T ₃	T ₂	T ₁₀	T ₁₉	T ₃	T ₁₁	T ₁₀	T ₁₆	T ₁
2,41	2,05	1,89	3,45	1,86	2,21	2,25	2,94	2,05	1,62
T ₁₄	T ₁₆	T ₁₅	T ₇	T ₁₁	T ₁₈	T ₂	T ₉	T ₅	T ₆
3,40	2,56	1,95	3,15	2,54	2,39	1,51	1,96	2,69	2,01
T ₁₇	T ₉	T ₁₂	T ₂₀	T ₅	T ₈	T ₁₇	T ₁₄	T ₁₉	T ₂₀
3,15	1,78	3,95	2,89	2,78	3,22	3,95	3,86	1,99	3,05
T ₈	T ₁₃	T ₁	T ₄	T ₆	T ₁₃	T ₇	T ₄	T ₁₅	T ₁₂
3,16	2,47	1,72	2,57	1,90	1,98	2,85	2,39	1,83	4,12

Repetición III				
Bloques				
11	12	13	14	15
T ₉	T ₂	T ₃	T ₁₃	T ₈
1,35	1,39	1,48	1,24	2,24
T ₁	T ₆	T ₁₄	T ₇	T ₁₇
1,41	1,85	2,12	2,40	3,05
T ₁₀	T ₂₀	T ₁₈	T ₁₆	T ₄
2,06	2,09	1,87	1,65	1,74
T ₅	T ₁₂	T ₁₉	T ₁₁	T ₁₅
1,92	3,12	1,53	1,88	1,20

Un aspecto importante que debe observarse cuando se trata de analizar experimentos con un arreglo de tratamientos bajo un alfa – latice, especialmente en el área agrícola, es el anidamiento de los bloques incompletos dentro de cada repetición de la estructura experimental. Es decir, cada bloque incompleto debe tratarse como un espacio de terreno

disímil; por ello, cada repetición del experimento anida un grupo distinto de bloques. Esto se ha tratado de explicitar en los cuadros de datos de este ejercicio, definiendo el anidamiento de los bloques del 1 al 5, del 6 al 10 y del 11 al 15, en la primera, segunda y tercera repetición, respectivamente. Este anidamiento no podrá soslayarse aun a pesar de que las repeticiones se valoren longitudinalmente (de un momento a otro) ocupando exactamente el mismo sitio, ya que, aparte de tener diferentes condiciones agroclimáticas, las condiciones agroquímicas del suelo también diferirán de un momento a otro; de allí que se prevea también valorar la varianza intra-bloques considerando el anidamiento del efecto.

La forma de ordenar los datos para analizarlos en cualquier software estadístico es el que se indica parcialmente en la figura 5.2. Nótese la numeración de los bloques incompletos, según la repetición de la estructura experimental.

ID	Tratamientos	Repetición	Bloque_Inc	Rendimiento
1	T1	1	3	1,72
2	T2	1	3	1,89
3	T3	1	2	2,05
4	T4	1	4	2,57
5	T5	1	5	2,78
6	T6	1	5	1,90
7	T7	1	4	3,15
8	T8	1	1	3,16
9	T9	1	2	1,78
10	T10	1	4	3,45
11	T11	1	5	2,54
12	T12	1	3	3,95
13	T13	1	2	2,47
14	T14	1	1	3,40

Figura 5.2. Estructura de los datos de un experimento alfa – latice previo al análisis en un software estadístico.

El análisis de varianza respectivo, tomando en consideración las indicaciones de la tabla 5.16, es el que se detalla en la tabla 5.17. En este análisis se observan diferencias significativas entre los tratamientos (genotipos) y entre las repeticiones de la distribución; diferente a las diferencias entre bloques incompletos dentro de las repeticiones que, según su p-valor (0,705), no fueron significativas. Con este resultado podría volverse y realizar un nuevo ANDEVA eliminando la fuente de *Bloques en repeticiones* (tiene un valor F menor a 1), tal como si se tratase de una distribución de bloques completos al azar, pues no se perdería precisión y más bien se lograría incrementarla debido a que el error experimental ya no tendría 26 sino 38 grados de libertad. Esto último se le deja al lector para que lo compruebe. Asimismo, el experimento goza de credibilidad de acuerdo al CV (11,25%) y a la importante variabilidad conocida del modelo, reportada por el r^2 (87%).

Tabla 5.17. Análisis de varianza para el ejemplo de aplicación 5.4, considerando una distribución de bloques incompletos con arreglo alfa – latice.

Fuentes de variación	GL	SC	CM	F	P-valor
Total	59	30,98			
Tratamientos (genotipos)	19	22,35	1,176	17,05	0,000
Repeticiones	2	6,23	3,117	45,20	0,000
Bloques en repeticiones (efecto anidado)	12	0,61	0,051	0,74	0,705
Error experimental	26	1,79	0,069		

$CV = 11,25\%$; $r^2(a_j) = 0,87$

En la tabla 5.18 se indica el resultado de la comparación de medias, aplicando el test de Tukey al 5% de probabilidad. Los resultados representan el compendio de la salida de Minitab, realizándose el ajuste de las medias, según el error estándar correspondiente, cuyos valores difieren por el efecto de los bloques incompletos. En la tabla se identifican 6 tratamientos (genotipos) con las respuestas estadísticamente más altas, reunidas en el grupo a.

Tabla 5.18. Comparación de medias utilizando el test de Tukey ($p < 0,05$) para el ejemplo de aplicación 5.4.

Tratamientos	EE de la media	Media ajustada	Grupo
T1	0,185	1,51	h
T2	0,184	1,56	gh
T3	0,188	1,97	efgh
T4	0,179	2,10	defgh
T5	0,178	2,53	bcdefg
T6	0,189	1,86	efgh
T7	0,178	2,75	abcde
T8	0,186	3,01	abcd
T9	0,179	1,58	gh
T10	0,181	2,63	abcdef
T11	0,178	2,31	cdefgh
T12	0,197	3,61	a
T13	0,185	1,91	efgh
T14	0,183	3,20	abc
T15	0,175	1,73	fgh
T16	0,18	2,12	defg
T17	0,179	3,53	ab
T18	0,191	2,39	cdefgh
T19	0,178	1,95	efgh
T20	0,189	2,48	bcdefg

Medias con letras iguales no difieren significativamente.

El investigador tiene en las distribuciones en latice el medio para lograr precisión cuando un experimento tiene un elevado número de tratamientos. Si aún todavía existen ciertas cantidades de tratamientos que no pueden ajustarse a las estructuras de latices³⁷ cuadrados, cúbicos, rectangulares, o dentro de los mismos alfa – latices, éstas son muy escasas y el

³⁷ Un ejemplo de experimento que no calza exactamente en una estructura en latice es aquel que tenga 23 tratamientos para evaluación. Esta cantidad no se ajusta, de acuerdo a sus fundamentos, a un latice cuadrado, cúbico o rectangular, ni tampoco en un alfa – latice. Sin embargo, en términos prácticos, podría estructurarse en campo la última distribución mencionada, considerando un arreglo de 6 x 4, dejando uno de los bloques incompletos desbalanceado debido a que se tendrá una unidad experimental faltante. Esto puede reflejarse en las sumas de cuadrados, la misma que deberá considerar el respectivo bloque incompleto desbalanceado.

biometrista deberá escoger la distribución que mayor ventaja le ofrezca desde la perspectiva del modelo de ANDEVA que aplique.

Diseños con doble bloqueo: alternativas a la distribución en cuadro latino

Diseño renglón – columna: cuadrados de Youden

En algunas situaciones de experimentación agrícola se busca la precisión de un ensayo utilizando la distribución en cuadro latino, de tal manera que se puedan controlar dos fuentes de variabilidad o de ruido (fuentes secundarias de variación) al mismo tiempo; esto especialmente cuando se involucran cultivos frutales que demandan tamaños de parcela (o distancias entre las unidades experimentales) relativamente amplios, con lo que la afectación de la heterogeneidad del suelo contribuiría a aumentar la varianza del error experimental. En estos contextos caben los cuadrados latinos. Sin embargo, la principal desventaja de estas distribuciones es el requerimiento de que el número de unidades experimentales sea equivalente al cuadrado del número de tratamientos ($n = t^2$), lo cual restringe su uso en experimentos estructurados para un número relativamente alto de tratamientos (en la práctica, 7 o más tratamientos) debido a factores netamente económicos y de control de fuentes de ruido.

Alternativamente, si un experimento necesita bloquear dos fuentes de variabilidad, es práctico utilizar los llamados *diseño renglón – columna* (DR-C). Los DR-C pueden tener los renglones, las columnas o ambos como bloques incompletos, lo cual permite reducir el tamaño del experimento y mantener la precisión requerida. Para lograr un mayor entendimiento de este tipo de distribución, analicemos el siguiente ejemplo de aplicación.

Ejemplo de aplicación 5.5

Se desarrollan siete formulaciones de biol (mezcla orgánica fermentada para nutrición de plantas) para aplicarse, vía *drench*, a plantas de banano en crecimiento de un estado desarrollo determinado. Debido a la complicación de tener plantas en un mismo estado de forma continua, se decide bloquear el experimento en los dos sentidos con el fin de separar fuentes de variabilidad (de ruido) que pudieran afectar los resultados. Sin embargo, dado el espacio del área experimental que complica utilizar una distribución en cuadro latino, se decide utilizar un diseño renglón – columna (DR-C) como el que se indica en la tabla 5.19. La variable evaluada fue el incremento de altura de las plantas en cm.

Tabla 5.19. Datos de incremento de altura (cm) en plantas de banano de un experimento bajo un DR-C.

Renglones	Columnas						
	1	2	3	4	5	6	7
1	Form. 7	Form. 1	Form. 2	Form. 3	Form. 4	Form. 5	Form. 6
	22,7	22,5	18,9	12,5	28,9	25,3	24,7
2	Form. 6	Form. 7	Form. 1	Form. 2	Form. 3	Form. 4	Form. 5
	32,4	28,9	32,5	24,8	18,7	32,8	30,2
3	Form. 5	Form. 6	Form. 7	Form. 1	Form. 2	Form. 3	Form. 4
	23,1	22,8	23,5	21,9	14,7	10,5	24,7
4	Form. 4	Form. 5	Form. 6	Form. 7	Form. 1	Form. 2	Form. 3
	21,1	24,8	28,1	25,4	20,5	19,7	13,1

Form.: Formulación

Nótese que el arreglo de la tabla 5.19 es un diseño de bloques completos para los renglones e incompleto para las columnas, aunque estas últimas están balaceadas. Es decir, esta distribución es ortogonal para los renglones debido a que todos los tratamientos están presentes en cada uno de dichos renglones. Esta distribución se circunscribe a los cuadrados latinos incompletos, en los que se omiten dos o más renglones, y son conocidos también como *cuadrados de Youden* en honor a quien los propuso inicialmente. Los parámetros de estas distribuciones se definen por: $t = b$, $r = k$ y $\lambda = k(k - 1)/(t - 1)$. En textos como el de Kuehl (2001) o el de Cochran y Cox (1957) pueden consultarse algunos arreglos bajo la distribución de cuadrados de Youden.

Para el ejemplo, se tiene $t = 7$, $b = 7$, $r = k = 4$ (bloques incompletos) y $\lambda = 2$ bloques incompletos en que se tienen pares de tratamientos. Los tratamientos son aleatorizados mediante un proceso de permutación horizontal similar al que se realiza en las distribuciones de cuadro latino.

Análisis de varianza en una distribución renglón - columna

El modelo estadístico para este tipo de distribución es el mismo que se aplica a distribuciones en cuadro latino, definido como: $Y_{ijk} = \mu + \tau_i + \beta_j + \gamma_k + \varepsilon_{ijk}$. En este caso, la suma de cuadrados de tratamientos debe ajustarse debido a los bloques incompletos; no así en el caso de los renglones, los cuales son ortogonales y no requieren de ajuste por esta causa. Esta suma de cuadrados se obtiene con la expresión 5.2 antes indicada:

$$SC_{t(af)} = \frac{k \sum Q_i^2}{\lambda t}$$

Definiéndose Q_i para esta expresión de la forma siguiente:

$$Q_i = y_{i..} - \frac{\sum B_j}{k}$$

siendo B_j los totales de bloques que incluyen el tratamiento i .

En la tabla 5.20 se presenta el análisis de varianza para los datos del ejemplo de aplicación 5.5 realizado con Minitab. En este caso, las sumas de cuadrados se han corregido debido a los bloques (columnas) incompletos; así también la variabilidad de las columnas. La falta de ortogonalidad de los tratamientos respecto de las columnas ha hecho que las sumas de cuadrados de tratamientos, renglones, columnas y error experimental, no se corresponda con el valor de 935,57 que es la variabilidad total; las sumas de cuadrados de estos componentes tienen un valor menor, equivalente a 872,78. El análisis revela diferencias significativas ($p < 0,05$) en los tratamientos y en los renglones, mientras que, entre columnas, este efecto es nulo. En esta situación podría probarse un análisis bajo una distribución de bloques completos al azar, dado que en forma horizontal los bloques no muestran variabilidad.

Tabla 5.20. Análisis de varianza de los datos del ejemplo de aplicación 5.5 realizado bajo una distribución renglón – columna (cuadrado de Youden).

Fuentes de variación	GL	SC Ajust.	MC Ajust.	Valor F	P-valor
Tratamientos (aj.)	6	512,21	85,368	25,20	0,000
Renglones	3	289,65	96,550	28,50	0,000
Columnas (aj.)	6	30,26	5,043	1,49	0,262
Error exp.	12	40,66	3,388		
Total	27	935,57			

Resumen del modelo de análisis de varianza.

Error estándar	R-cuad.	R-cuad. (ajustado)	R-cuad. (pred)
1,84068	95,65%	90,22%	75,61%

El modelo de ANDEVA es aceptable según el coeficiente de determinación (95,65%). Asimismo, según el test de Tukey ($p < 0,05$) cuyos resultados se indican en la tabla 5.21, las formulaciones 1, 4, 5, 6 y 7 pueden aplicarse al cultivo de acuerdo a los efectos más altos presentados.

Tabla 5.21. Prueba de Tukey ($p < 0,05$) aplicado a las medias del ejemplo de aplicación 5.5.

Tratamientos	N	Media	Agrupación
Form. 4	4	27,3399	a
Form. 6	4	27,3002	a
Form. 5	4	26,5094	a
Form. 7	4	25,0941	a
Form. 1	4	23,9302	a
Form. 2	4	18,7281	b
Form. 3	4	13,5231	c

Distribuciones resolubles de bloques incompletos

Otra alternativa de desarrollo de experimentos con múltiples tratamientos y con k unidades experimentales por bloque que obligan a que $k < b$, ya sea por términos económicos, por control complicado de fuentes de ruido, por falta de material experimental y/o por que alguna técnica así lo establece, es la de las *distribuciones resolubles de bloques incompletos* (DRBI). Una DRBI es una estructura en grupos de bloques incompletos, cada uno constituyendo una repetición del experimento, lo cual le confiere eficacia al mismo en el sentido de realizar un mejor control del error experimental y que al final redunda en una mayor precisión del ensayo.

La DRBI es una estructura experimental que se puede llevar a cabo por etapas, pudiendo cada etapa comprender la corrida de una o más repeticiones del ensayo; de allí su definición de resoluble. Es muy útil cuando un experimento con todas sus repeticiones completas no puede desarrollarse al mismo tiempo (Kuehl, 2001), siempre y cuando se cumpla con la obligación

de reproducir exactamente las condiciones experimentales en el desarrollo de cada repetición. No obstante, sino pudieran cumplirse las condiciones de desarrollo de cada repetición o éstas son parciales, podrían evaluarse como fuentes de ruido en los modelos de análisis de varianza incluyéndolas como fuente de anidamiento de las repeticiones.

Las DRBI son distribuciones que se establecen con r grupos de repeticiones de s bloques con k unidades experimentales por bloque. Para que la DRBI sea balanceada es requerimiento que el número de t tratamientos sea un múltiplo del producto sk , es decir: $t = sk$; además, debe satisfacerse que, para el número total de bloques (b), se cumpla con la relación: $b = rs \geq t + r - 1$. Debe haberse advertido que el tema sobre distribuciones en latice antes vista es un ejemplo clásico de las distribuciones resolubles, no obstante, para una mayor captación de la aplicabilidad de este tipo de distribución, veamos el siguiente ejemplo aplicado a una evaluación sensorial de alimentos.

Ejemplo de aplicación 5.6

Un experimento sobre evaluación sensorial se ha realizado con 15 formulaciones (tratamientos) de una jalea. Se evaluó el sabor mediante un criterio hedónico utilizando una escala lineal (figura 5.3) y un panel de jueces tipo consumidor integrado de 30 personas. Un problema técnico de una evaluación sensorial es que un catador no debería probar más de cuatro muestras (formulaciones) para detectar alguna diferenciación entre ellas con mayor precisión (menor error experimental), además de que la disponibilidad de los panelistas para el proceso de catación muchas veces se limita por cuestiones personales y económicas del propio ensayo. En este contexto, se diseñó este experimento bajo una distribución resoluble de bloques incompletos, cuya estructura se detalla a continuación:

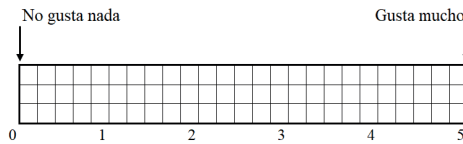


Figura 5.3. Escala lineal tipo hedónica.

Repetición 1				Repetición 2			
Jueces				Jueces			
1	8	7	9	6	12	4	9
	1,6	4,4	2,2		2,6	3,2	2,4
2	14	13	15	7	14	5	7
	3,4	2,8	3,2		3,2	1,8	4,4
3	2	1	3	8	13	1	8
	3,6	2,2	4,8		1,2	2,8	1,0
4	4	6	5	9	15	3	6
	2,8	2,4	1,6		2,8	4,6	1,8
5	11	10	12	10	11	2	10
	2,0	5,0	3,2		4,0	0,6	4,6

Repetición 3				Repetición 4			
Jueces				Jueces			
11	14	13	15	16	8	3	13
	2,6	3,4	2,2		2,8	4,8	3,2
12	8	7	9	17	10	5	15
	0,8	4,6	1,6		4,2	5,0	4,0
13	4	6	5	18	6	1	11
	2,0	2,8	1,8		0,8	1,6	4,6
14	11	10	12	19	9	4	14
	3,2	5,0	1,2		2,8	3,6	2,6
15	2	1	3	20	7	2	12
	1,0	2,4	4,4		4,8	3,2	1,2

Como podrá notarse, la distribución de este ensayo no tiene balanceo y el efecto de los bloques está anidado en las repeticiones; por lo tanto, el modelo estadístico para el análisis de los datos en este ejemplo es:

$$Y_{ijk} = \mu + \tau_i + \alpha_j + \beta_{k(j)} + \varepsilon_{ijk}$$

En el modelo, μ es el efecto constante, τ_i es el efecto de los i tratamientos, α_j representa al efecto de las j repeticiones (grupo de tratamientos y bloques), $\beta_{k(j)}$ es el efecto de los $k(j)$ bloques (en el ejemplo, los jueces) dentro de cada repetición y ε_{ijk} representa al error experimental. Asimismo, la suma total de cuadrados estará compuesta de la suma de cuadrados de tratamientos, de repeticiones, de bloques incompletos anidados en repeticiones y del error experimental. Lo último, de forma respectiva corresponde a: $SC_T = SC_t + SC_r + SC_{b(r)} + SC_E$.

En este caso el análisis de los datos se realiza con Minitab aplicando la sentencia: Estadísticas → ANOVA → Modelo lineal general → Ajustar modelo lineal general. La matriz de datos en su parte inicial puede observarse en la figura 5.4.

Formulaciones (tratamientos)	Repeticiones	Jueces (bloques)	Sabor
8	1	1	1,6
14	1	2	3,4
2	1	3	3,6
4	1	4	2,8
11	1	5	2,0
7	1	1	4,4
13	1	2	2,8
1	1	3	2,2
6	1	4	2,4
10	1	5	5,0
9	1	1	2,2
15	1	2	3,2
3	1	3	4,8

Figura 5.4. Esquema parcial de la matriz de datos para su análisis en un software.

El análisis de varianza de este ejemplo se detalla en la tabla 5.22. Según este procedimiento, las formulaciones presentan diferencias significativas ($p < 0,05$) y deben someterse a un análisis de comparación de medias para establecer la mejor (s) formulación. Es importante indicar que la suma de cuadrados de las formulaciones (tratamientos) están ajustadas por el efecto de los bloques incompletos. Por otro lado, las repeticiones y los bloques incompletos anidados en

repeticiones no han presentado efectos significativos, lo que permite entrever la homogeneidad en el criterio de los jueces evaluadores.

El modelo de ANDEVA tiene una importante aceptación dado por el coeficiente de determinación, cuyo valor es de 73,39%. Así también se respalda el análisis con la verificación gráfica detallada en la figura 5.5 en cuanto a la normalidad, la homocedasticidad e independencia de los residuos.

Tabla 5.22. Análisis de varianza del ejemplo de aplicación 5.6.

Fuentes de variación	GL	SC Ajust.	MC Ajust.	F	P-valor
Formulaciones (Aj.)	14	47,743	3,4102	3,58	0,002
Repeticiones	3	4,123	1,3742	6,46	0,128 *
Jueces(Repeticiones)	16	6,367	0,3979	0,42	0,964
Error experimental	26	24,791	0,9535		
Total	59	93,157			

*No es una prueba F exacta.

Resumen del modelo de análisis de varianza del ejemplo de aplicación 5.7.

Error Estándar	R-cuad.	R-cuad. (ajustado)	R-cuad. (pred)
0,976465	73,39%	39,61%	0,00%

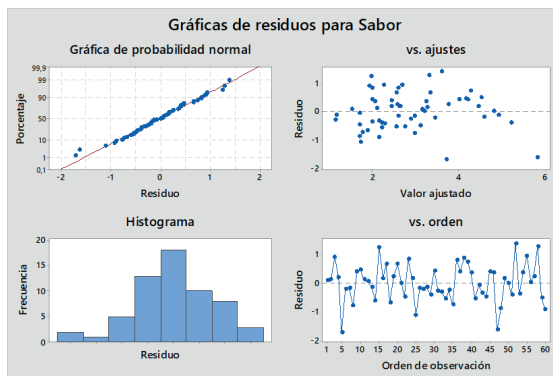


Figura 5.5. Análisis gráfico de residuos para los datos del ejemplo de aplicación 5.6.

Es importante indicar que en este capítulo, a través de algunos de los ejemplos expuestos (como este último), se ha considerado dentro de los análisis de varianza el efecto de anidamiento. La anidación de factores en un experimento es un aspecto que a veces, sea por negligencia del experimentador o por su inobservancia, es obviada y cargada, al final, al error experimental. De allí que en ensayos que son probados simultáneamente en diferentes zonas, no se logre la precisión necesaria para detectar los efectos de los tratamientos de prueba. En el

siguiente capítulo se explica el efecto anidado en las réplicas de experimentos con mayor detalle.

CAPÍTULO VI

DISTRIBUCIONES CON BLOQUES ANIDADOS O MODIFICADOS

Si bien la precisión es la que regula la validez interna de un experimento, para los dominios de recomendación a grupos de agricultores no solo basta que un ensayo logre esa necesaria precisión puntual, se requiere que la misma se extienda a toda una zona agrícola mediante réplicas experimentales; de tal forma que en los resultados de una prueba se incluya también a la importante validez externa...



Conjunto de experimentos replicados en el espacio (localidad) y en el tiempo (temporada)

Dentro de los programas de investigación agrícola que tienen básicamente propósitos de inducción o de generalización de resultados, que buscan especialmente plasmar una fundamentación a las recomendaciones de aplicación que se les brinda a los agricultores, se encuentran los experimentos realizados *en serie y/o en paralelo*, los cuales involucran ensayos replicados ya sea en el tiempo y/o en el espacio, respectivamente. Las pruebas a las que nos referimos bajo las condiciones mencionadas, por causa del ambiente tan variable que es el que rodea a estos tipos de experimentos cuyo objeto de estudio son las plantas, generalmente son réplicas similares en todos sus aspectos; es decir, tienen un igual número y clase de tratamientos, contienen la misma cantidad de bloques (repeticiones), tamaños de parcela semejantes, igual manejo experimental, las mismas variables a medirse y demás actividades exactamente idénticas. El investigador debe advertir que la similitud del grupo de experimentos que prevé desarrollar en las diferentes localidades de una zona y/o temporadas le evitará introducir complicaciones en el ajuste de los datos como, por ejemplo, recurrir a un análisis de covarianza cuando tiene parcelas de diferente tamaño.

Las réplicas de un experimento en el espacio y/o tiempo permiten afianzar la confiabilidad en los resultados; pues en esas circunstancias el investigador somete a un mismo grupo de tratamientos ante factores exógenos que tiene influencia en las respuestas, con el fin de darle validez externa a alguna alternativa de prueba que se haya presentado con diferencias significativas dentro de un único experimento respecto al resto de tratamientos. Esto es particularmente importante a nivel de los experimentos agrícolas realizados a campo abierto, en donde los factores edáficos y ambientales tienen una fuerte influencia, ocasionando efectos de interacción que deben ser separadas del error experimental para mejorar la precisión experimental. La validación experimental espacio-temporal de los resultados de un ensayo al interior de una zona agrícola hace posible definir los llamados *dominios de recomendación*, esto es, el alcance de los resultados para aplicarlos exclusivamente a un grupo de fincas que se encuentran circunscritas a dicho dominio.

En estos experimentos que se prueban, ya sea en serie (longitudinal) y/o en paralelo (transversal), un efecto que muchas veces no se advierte es el de *anidamiento*³⁸. Este efecto está presente entre los bloques de las distribuciones utilizadas (DBCA o DCL), sin embargo, son variaciones localizadas del sitio y/o de la época en el cual se lleva a cabo el ensayo; es decir, los bloques están anidados en la localidad y/o periodo dentro del cual se realizaron las pruebas experimentales. Esta consideración no es tomada en cuenta por muchos investigadores, con lo cual, le están restando precisión a sus experimentos al dejar fuera la variabilidad localizada de los bloques. Nótese en esta parte que la anidación a la cual se hace referencia es sobre una fuente de variación secundaria (fuente bloqueada), sin embargo, también los factores primarios pueden ser anidados. En este último caso, cuando se trata de dos o más factores principales con anidamiento, diferentes niveles de un factor son probados a distintos niveles del otro factor y su interacción no tiene razón para valorarse³⁹. En este libro solo nos remitiremos a la anidación de los bloques, salvo las necesarias explicaciones que se hacen cuando un ensayo se prueba en el espacio y también en el tiempo.

³⁸ Sobre esta temática existen capítulos enteros de diseños anidados en textos como el de Montgomery (2005), conocidos también como diseños jerárquicos y que el lector puede revisar.

³⁹ Un error común cometido en los análisis de experimentos con factores anidados en otros, es pensar que los tratamientos están definidos por las combinaciones (sin restricción) de los niveles de los factores; lo cual, al final, puede inducir a realizar recomendaciones erróneas.

Experimentos replicados en el espacio (localidad)

En este caso se trata de experimentos similares en todos los aspectos, que se desarrollan a campo abierto (en medios no controlados) de forma simultánea y en distintas localidades dentro de una zona agrícola. Este tipo de estudio se denomina de característica transversal, ya que todas las réplicas del ensayo se desarrollan dentro de un mismo periodo que es seleccionado según las condiciones climáticas de mayor producción para un cultivo determinado, en donde se toman en cuenta aspectos como contenidos de humedad residual del suelo, época de menor ataque de plagas, entre otros. Los tratamientos o factores de interés que se prueban pueden corresponder a estudios sobre variedades e híbridos, dosificaciones de plaguicidas, distanciamientos de siembra, además de tipos y dosificaciones de sustancias nutricionales.

Un esquema de experimento bajo una DBCA, replicado en tres localidades, se describe en la figura 6.1. Como podrá observarse en la figura, en esta clase de ensayos todos los tratamientos se repiten en cada localidad, por lo tanto, se tiene una interacción completa entre ellos (tratamiento x localidad). Un aspecto que debe resaltarse en este planteamiento es la consideración de que los tratamientos y las localidades son factores principales. Por otro lado, lo que es distinto en cada sitio, por la misma heterogeneidad del suelo, son los bloques; de allí que su efecto se encuentre anidado a la localidad respectiva. Cuantificar la variabilidad entre bloques dentro de cada localidad siempre será favorable al incremento de la precisión de los experimentos, aun a pesar de que el efecto de estos resulte no significativo y aparentemente le reste grados de libertad al error experimental, siendo esto último imperceptible en el análisis por cuanto las réplicas permiten tener un valor relativamente alto en esta fuente de variación.

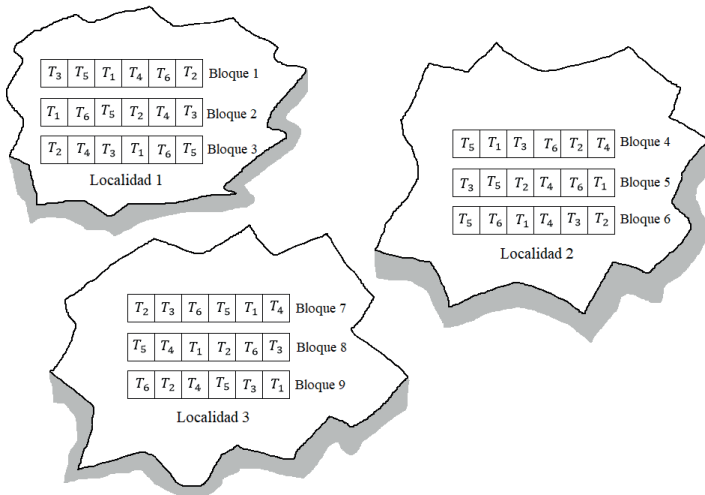


Figura 6.1. Esquema de un experimento de seis tratamientos, bajo una distribución de bloques completos al azar y replicado en tres localidades (bloques anidados en localidad).

La estructura de ANDEVA para estos ensayos se establece considerando un modelo de efectos absolutamente fijos, dado que tanto los tratamientos como las localidades no se seleccionan de forma aleatoria sino en base a un análisis teórico-práctico que debe ser realizado intencionalmente por el investigador de acuerdo a los intereses de éste. Los componentes de varianza están determinados por los tratamientos (sean estos niveles de un solo factor o combinaciones de más de un factor) (t), la localidad (L), la interacción de tratamientos y localidad ($t \times L$), los bloques anidados en las localidades [$b(L)$] y el error experimental. Un esquema de ANDEVA para un experimento bajo una DBCA replicado en el espacio es el que se detalla en la tabla 6.1.

Tabla 6.1. Esquema de análisis de varianza para un mismo experimento replicado en el espacio (localidad) bajo una DBCA.

Fuentes de variación	Grados de libertad	Sumas de cuadrados
Total	$tLb - 1$	$\sum Y_{ij[k(j)]}^2 - FC$
Tratamientos	$t - 1$	$\frac{\sum(\sum t_i)^2}{b \times L} - FC$
Localidad	$L - 1$	$\frac{\sum(\sum L_j)^2}{b \times t} - FC$
Interacción de tratamientos por localidad	$(t - 1)(L - 1)$	$\frac{\sum(\sum tL_{ij})^2}{b} - \frac{\sum(\sum t_i)^2}{b \times L} - \frac{\sum(\sum L_j)^2}{b \times t} + FC$
Bloques (Localidad)	$L(b - 1)$	$\frac{\sum(\sum b_{[k(j)]})^2}{t} - \frac{\sum(\sum L_j)^2}{b \times t}$
Error experimental	$L(t - 1)(b - 1)$	Por sustracción

En las sumas de cuadrados indicadas en la tabla 6.1, FC es el factor de corrección y se define por la expresión:

$$FC = \left(\sum Y_{ij[k(j)]} \right)^2 / tLb$$

Las sumas de cuadrados de los bloques están definidas por su anidamiento en las localidades. Es decir, la expresión $\sum(\sum b_{[k(j)]})^2$ establece la suma de cuadrados de todos los bloques en cada una de las localidades. El modelo estadístico que toma en cuenta el efecto de anidamiento de los bloques (repeticiones) en las localidades, además de contemplar una asignación de tratamientos en una DBCA, es el que se indica a continuación:

$$Y_{ijk(j)} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \gamma_{k(j)} + \varepsilon_{ijk(j)}$$

en donde $Y_{ijk(j)}$ es la respuesta de la especie, μ es el efecto constante, τ_i es el efecto de tratamientos, β_j es el efecto de la localidad (espacio), $(\tau\beta)_{ij}$ es el efecto de la interacción entre tratamientos y localidad, $\gamma_{k(j)}$ es el efecto de los bloques anidados en localidad y $\varepsilon_{ijk(j)}$ es el efecto del error experimental.

Ejemplo de aplicación 6.1

Se realiza un experimento en el cual se comparan cuatro variedades de papa, replicado simultáneamente en tres localidades de una zona. Los ensayos se desarrollaron bajo una distribución de bloques completos al azar, utilizando cinco repeticiones (bloques) en cada uno de ellos y considerando las mismas características dimensionales del área experimental. Los datos que se indican en las tablas 6.2, 6.3 y 6.4 corresponden al rendimiento reportado en t/ha.

Tabla 6.2. Rendimientos de papa en experimento desarrollado en la Localidad 1 de la zona de estudio.

N°	Tratamientos	Bloques					Total	Promedios
		I	II	III	IV	V		
1	Chaucha	82,47	89,94	76,82	70,14	75,86	395,23	79,05
2	Sta. Rosa	62,30	55,71	74,10	48,96	55,21	296,28	59,26
3	Colorada	75,68	69,98	81,32	80,25	76,45	383,68	76,74
4	Chola	92,37	88,65	90,54	84,12	96,35	452,03	90,41
	Total	312,82	304,28	322,78	283,47	303,87	1527,22	

Tabla 6.3. Rendimientos de papa en experimento desarrollado en la Localidad 2 de la zona de estudio.

N°	Tratamientos	Bloques					Total	Promedios
		I	II	III	IV	V		
1	Chaucha	74,12	80,25	69,23	78,95	82,36	384,91	76,98
2	Sta. Rosa	42,25	40,13	49,09	39,80	45,23	216,5	43,30
3	Colorada	95,12	75,89	80,23	86,20	75,41	412,85	82,57
4	Chola	87,42	79,56	82,36	67,80	74,58	312,16	62,43
	Total	298,91	196,27	280,91	272,75	277,58	1326,42	

Tabla 6.4. Rendimientos de papa en experimento desarrollado en la Localidad 3 de la zona de estudio.

N°	Tratamientos	Bloques					Total	Promedios
		I	II	III	IV	V		
1	Chaucha	79,98	71,20	80,14	83,05	75,32	389,69	77,94
2	Sta. Rosa	56,31	61,25	49,89	58,47	62,34	288,26	57,65
3	Colorada	80,01	77,08	84,25	80,14	86,54	408,02	81,60
4	Chola	95,12	103,52	89,04	95,60	90,03	473,31	94,66
	Total	311,42	313,05	303,32	317,26	314,23	1559,28	

A partir de las tablas 6.2, 6.3 y 6.4 se obtienen las tablas 6.5 y 6.6 correspondientes a las sumatorias de localidad por bloque y localidad por tratamiento. Estos valores servirán para obtener los efectos de los factores y de la interacción de los mismos.

Tabla 6.5. Sumatorias para análisis de localidad, según los bloques.

Bloques	Localidad			Total
	1	2	3	
I	312,82	298,91	311,42	923,15
II	304,28	275,83	313,05	893,16
III	322,78	280,91	303,32	907,01
IV	283,47	272,75	317,26	873,48
V	303,87	277,58	314,23	895,68
Total	1527,22	1405,98	1559,28	4492,48

Tabla 6.6. Sumatorias para análisis de localidad, según los tratamientos (variedades).

Variedades	Localidad			Total
	1	2	3	
Chaucha	395,23	384,91	389,69	1169,83
Sta. Rosa	296,28	216,5	288,26	801,04
Colorada	383,68	412,85	408,02	1204,55
Chola	452,03	391,72	473,31	1317,06
Total	1527,22	1405,98	1559,28	4492,48

El factor de corrección es:

$$FC = \frac{(\sum Y_{ij[k(j)]})^2}{n} = \frac{4492,48^2}{60} = 336372,94$$

Cálculos de las sumas de cuadrados:

- SC del total de los datos:

$$SC_{Tot} = \sum Y_{ij[k(j)]}^2 - FC$$

$$SC_{Tot} = 82,47^2 + 62,30^2 \dots + 90,03^2 - 336372,94 = 13402,91$$

- SC de variedades (tratamientos):

$$SC_V = \frac{\sum (\sum t_i)^2}{b \times L} - FC$$

$$SC_V = \frac{1169,83^2 + \dots + 1317,06^2}{5 \times 3} - 336372,94 = 10010,73$$

- SC de localidades:

$$SC_L = \frac{\sum (\sum L_j)^2}{b \times t} - FC$$

$$SC_L = \frac{1527,22^2 + \dots + 1559,28^2}{5 \times 4} - 336372,94 = 653,80$$

- SC de la interacción de variedad por localidad:

$$SC_{VL} = \frac{\sum (\sum tL_{ij})^2}{b} - FC - SC_V - SC_L$$

$$SC_{VL} = \frac{395,23^2 + \dots + 473,31^2}{5} - 336372,94 - 10010,73 - 653,80 = 943,03$$

- SC de bloques anidados en localidades:

$$SC_{b(L)} = \frac{\sum(\sum b_{(k(j))})^2}{t} - \frac{\sum(\sum L_j)^2}{b \times t}$$

$$SC_{b(L)} = \frac{312,82^2 + \dots + 314,23^2}{4} - \frac{1527,22^2 + \dots + 1559,28^2}{5 \times 4} = 344,51$$

- SC del error experimental:

$$SC_E = SC_T - SC_V - SC_L - SC_{VL} - SC_{b(L)}$$

$$SC_E = 13402,91 - 10010,73 - 653,80 - 943,03 - 344,51 = 1450,85$$

Una vez obtenidas las sumas de cuadrados de cada fuente de variación, el análisis de varianza se detalla en la tabla 6.7. Según este análisis, las variedades, las localidades y la interacción de estos dos factores presentan diferencias significativas ($p < 0,05$). En el caso de los bloques, éstos son relativamente homogéneos dentro de cada localidad, dado que su efecto no presenta significancia estadística. Los resultados del experimento, en esta variable, revelan una relativa confianza según el coeficiente de variación (8,48%) y el coeficiente de determinación (0,89) establecidos.

Tabla 6.7. Análisis de varianza con efecto de bloques anidados en localidades del ejemplo de aplicación 6.1.

Fuentes de variación	GL	SC	CM	F	P
Total	59	13402,91			
Variedades (V)	3	10010,73	3336,91	82,80	0,000
Localidades (L)	2	653,80	326,899	8,11	0,001
Interacción V x L	6	943,03	157,172	3,90	0,004
Bloques (Localidades)	12	344,51	28,7088	0,71	0,730
Error experimental	36	1450,85	40,3013		
CV = 8,48%					
$r^2 = 0,89$					

Es importante volver a indicar que el ANDEVA detallado en la tabla 6.7 se ha realizado considerando efectos fijos para los tratamientos, localidades y bloques. Sin embargo, cuando no existen restricciones de ningún tipo para el investigador al momento de replicar un experimento dentro de una zona, las localidades deberán procesarse bajo la consideración de efectos aleatorios porque se seleccionarán con este criterio. Así, por ejemplo, si en el ANDEVA de la tabla 6.7, únicamente las localidades se hubiesen seleccionado de forma aleatoria (efecto aleatorio), los estadísticos de prueba (F) para tratamientos (F_t), localidades (F_L) y su interacción ($F_{t \times L}$) se calcularán de la forma siguiente:

$$F_t = \frac{CM_t}{CM_{t \times L}} \quad F_L = \frac{CM_L}{CM_E} \quad F_{t \times L} = \frac{CM_{t \times L}}{CM_E}$$

siendo $CM_{t \times L}$ el cuadrado medio para la interacción de tratamientos y localidades. Si los bloques se consideran definidos bajo selección aleatoria, el valor del estadístico F para el

planteamiento de ANDEVA descrito en la tabla 6.7 de esta fuente de variación no se modificará y seguirá manteniendo como denominador al cuadrado medio del error experimental (CM_E).

En la figura 6.2 se expone parcialmente la forma de elaboración de la base de datos para el análisis de este tipo de experimento. La salida de la figura 6.3 es el análisis realizado con Infostat, en donde, además de observarse el ANDEVA, se exhibe el resultado de la prueba de comparación de medias realizada con el test de Tukey, desde donde se puede apreciar que el mayor rendimiento posible se logra con la variedad de papa “Chola”. El modelo definido en el software es el siguiente:

Variedades
 Localidad
 Variedades*Localidad
 Localidad>Bloques

Variedades	Bloques	Localidad	Rendimiento
Chaucha	I	1	82,47
Sta. Rosa	I	1	62,30
Colorada	I	1	75,68
Chola	I	1	92,37
Chaucha	II	1	89,94
Sta. Rosa	II	1	55,71
Colorada	II	1	69,98
Chola	II	1	88,65
Chaucha	III	1	76,82
Sta. Rosa	III	1	74,10
Colorada	III	1	81,32
Chola	III	1	90,54
Chaucha	IV	1	70,14
Sta. Rosa	IV	1	48,96
Colorada	IV	1	80,25
Chola	IV	1	84,12
Chaucha	V	1	75,86
Sta. Rosa	V	1	55,21
Colorada	V	1	76,45
Chola	V	1	96,35

Figura 6.2. Esquema de la base de datos para la realización del análisis en un software estadístico.

Cuadro de Análisis de la Varianza (SC tipo III)						
F.V.	SC	gl	CM	F	p-valor	
Modelo	11952,06	23	519,65	12,89	<0,0001	
Variedades	10010,73	3	3336,91	82,80	<0,0001	
Localidad	653,80	2	326,90	8,11	0,0012	
Variedades*Localidad	943,03	6	157,17	3,90	0,0042	
Localidad>Bloques	344,51	12	28,71	0,71	0,7295	
Error	1450,85	36	40,30			
Total	13402,91	59				

Test: Tukey Alfa=0,05 DMS=6,24312			
Error: 40,3013 gl: 36			
Variedades	Medias	n	E.E.
Chola	87,80	15	1,64 A
Colorada	80,30	15	1,64 B
Chaucha	77,99	15	1,64 B
Sta. Rosa	53,40	15	1,64 C

Medias con una letra común no son significativamente diferentes (p > 0,05)

Figura 6.3. Salida del software Infostat, con el análisis de varianza y la prueba de Tukey.

En la tabla 6.8 se compendian los análisis de varianza realizados por cada localidad bajo una DBCA. Se puede observar que las variedades reportan diferencias significativas en cada Localidad, esto es, existe una o algunas variedades que deben tener mejor respuesta que el resto y que muy posiblemente sea la variedad “Chola”. El manejo de los experimentos presenta consistencia y esmero, según los coeficientes de variación (CV) y de determinación (r^2) que se indican para cada localidad, con valores muy por debajo del 30% y arriba del 0,70, respectivamente.

Tabla 6.8. Resultados de los análisis de varianza por localidad del ejemplo de aplicación 6.1.

Fuentes de variación	GL	Valor P por Localidad		
		Loc. 1	Loc. 2	Loc. 3
Total	19			
Variedades	3	0,0001	0,0001	0,0001
Bloques	4	0,3733	0,6902	0,9124
Error experimental	12			
CV:		8,79%	9,74%	6,90%
r^2 :		0,83	0,90	0,91

El resultado de la aplicación de la prueba de Tukey que se detalla en la figura 6.3 representa al efecto combinado (o de interacción) de variedades por localidad, en donde se ha destacado la variedad “Chola” como la mejor; no obstante, este resultado debe confirmarse al realizar estas comparaciones por cada localidad, tal como lo que se indica en la tabla 6.9. Los resultados de esta última tabla confirman la mayor potencialidad productiva de la variedad “Chola”, ya que ésta se mantiene en las tres localidades.

Tabla 6.9. Separación de medias por localidad con la aplicación de la prueba de Tukey.

Tratamientos	Temporadas		
	Loc. 1	Loc. 2	Loc. 3
Chaucha	79,05 ab	76,98 a	77,94 b
Sta. Rosa	59,26 c	43,30 b	57,65 c
Colorada	76,74 b	82,57 a	81,60 b
Chola	90,41 a	78,34 a	94,66 a
Medias con letras iguales no difieren significativamente ($p > 0,05$).			

Experimentos replicados en el tiempo (temporada)

Los experimentos que se replican en el tiempo, conocidos como *experimentos en serie*, llevados a cabo en estudios longitudinales, buscan confirmar el efecto de algún tratamiento ganador en las primeras pruebas, midiendo su capacidad en ensayos subsecuentes para, de esta forma, valorar su respuesta bajo distintas modificaciones de un mismo clima; a pesar de que también habría condiciones fisicoquímicas del suelo que podrían variar de un año a otro. Un aspecto que es favorable para el desarrollo de estos ensayos es el hecho de que los periodos o temporadas para la mayoría de cultivos de ciclo corto están circunscritos a unos meses determinados de cada año. Así, por ejemplo, en el caso de Ecuador, cultivos como el maíz o la soya, tienen definidas sus temporadas entre la época lluviosa o la época seca; con lo que un experimento que se replica en varios años ajustándose a una misma temporada, solo permitirá corroborar la respuesta positiva de uno o más tratamientos que hayan presentado las mejores medias en un primer experimento. En cultivos en donde sus producciones no tienen temporadas definidas, como por ejemplo el banano, los ensayos replicados en el tiempo, aparte de verificar la consistencia de un tratamiento ganador obtenido en una prueba preliminar, pueden servir también para identificar los periodos dentro de un año en donde dicho tratamiento se presente con la mejor respuesta.

Un experimento que se repite en distintas temporadas ocupando un mismo espacio (en el mismo terreno), estará sometido a una menor heterogeneidad del suelo que aquellos que se repiten en distintos lotes de terreno. En esta perspectiva es deducible que en aquellos ensayos que involucran cultivos perennes y siempre que sean las mismas plantas, invariablemente se estará utilizando los mismos espacios de terreno para cada experiencia. Sin embargo, cuando en cada prueba subsecuente se utilizase un lote distinto (dentro del mismo predio), se estaría tratando con plantaciones diferentes en cada periodo y, por tanto, se deberá incluir una fuente de variación adicional representada por la variabilidad intrínseca (expresividad genética) de las especies. Esta fuente de variación estará representada por la unidad experimental y en los análisis podría llamarse como error de muestreo.

Son variados los experimentos que pueden valorarse a través del tiempo. Dentro de estos casos están los ensayos de variedades, fórmulas fertilizantes, plaguicidas, métodos de preparación de suelos, entre otros. Estas pruebas pueden realizarse tanto con cultivos de ciclo corto como también con los perennes. Dependiendo de la cantidad de ensayos que se realicen durante varios años, estas experiencias servirán para certificar la bondad de una alternativa de mejoramiento en cualquier temporada y/o identificar en qué periodos es más recomendable su aplicación.

El efecto del clima en los experimentos que se replican durante varias temporadas podría hacerse presente en el efecto de los bloques (repeticiones) cuando se trata de distribuciones como la DBCA o la DCL y en la respuesta de los mismos tratamientos. Además, es razonable

que el factor “temporada” (o tiempo si se prefiere) no puede ser aleatorizado como si se tratara de un ensayo factorial sin restricciones, no obstante que la interacción de interés para verificar si la respuesta de los tratamientos es generalizada entre temporadas es la de estas dos fuentes de variación. En esta situación el análisis se realiza de forma análoga a la de un arreglo factorial de *Parcelas divididas* (o de *bloques divididos*), en el cual el error experimental estará compuesto de tres tipos de errores:

1. Interacción doble de tratamientos y bloques.
2. Interacción doble de temporadas y bloques.
3. Interacción triple de tratamientos, temporadas y bloques.

El modelo estadístico, considerando que el experimento se desarrolla bajo una DBCA, replicado integralmente en un mismo sitio durante varias temporadas, es el siguiente:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_a + \gamma_k + \varepsilon_b + (\alpha\gamma)_{ik} + \varepsilon_c$$

En este modelo cada observación (Y_{ijk}) estará influenciada por el efecto constante μ , el efecto de tratamientos (α_i), el efecto de los bloques (β_j), el efecto del error a (ε_a), el efecto de las temporadas (γ_k), el efecto del error b (ε_b), el efecto de la interacción entre los tratamientos y las temporadas $[(\alpha\gamma)_{ik}]$ y el efecto del error c (ε_c). Los errores a , b y c corresponden a las interacciones que no son de interés primario en el experimento, pero son requeridas para realizar una comparación estadística en correspondencia a las fuentes de variación que tendrá el modelo de ANDEVA: tratamientos, temporadas y bloques. La interacción triple (Error c) es la que al final toma en cuenta el efecto de anidamiento de los tratamientos versus los bloques. El esquema de ANDEVA para un planteamiento como este es el que se describe en la tabla 6.10.

Tabla 6.10. Esquema de análisis de varianza para un mismo experimento replicado en el tiempo (temporada) bajo una DBCA.

Fuentes de variación	Grados de libertad	Sumas de cuadrados
Total	$tbT - 1$	$\sum Y_{ijk}^2 - FC$
Tratamientos (t)	$t - 1$	$\frac{\sum (\sum t_i)^2}{b \times T} - FC$
Bloques (B)	$b - 1$	$\frac{\sum (\sum b_j)^2}{t \times T} - FC$
Error a	$(t - 1)(b - 1)$	$\left[\frac{\sum (\sum tb_{ij})^2}{T} - FC \right] - SC_t - SC_b$
Temporada (T)	$T - 1$	$\frac{\sum (\sum T_k)^2}{t \times b} - FC$
Error b	$(T - 1)(b - 1)$	$\frac{\sum (\sum Tb_{jk})^2}{t} - FC - SC_{Temp} - SC_b$
Interacción de tratamientos por temporada	$(t - 1)(T - 1)$	$\frac{\sum (\sum tT_{ik})^2}{b} - FC - SC_t - SC_{Temp}$
Error c	$(t - 1)(T - 1)(b - 1)$	$SC_{Tot} - SC_t - SC_b - SC_{E(a)} - SC_{Temp} - SC_{E(b)} - SC_{t \times T}$

t : tratamientos, b : bloques, T : temporada, SC_{Tot} : suma de cuadrados del total de los datos, SC_t : suma de cuadrados de tratamientos, SC_b : suma de cuadrados de bloques, SC_{Temp} : suma de cuadrados de temporada, $SC_{E(a)}$: suma de cuadrados del error a (SC de la interacción temporada por bloques), $SC_{E(b)}$: suma de cuadrados del error b (SC de tratamiento por temporada), $SC_{t \times T}$: suma de cuadrados de la interacción de tratamientos por temporada.

Ejemplo de aplicación 6.2

Se prueban cuatro distanciamientos de siembra en caña de azúcar en un experimento llevado a cabo bajo una distribución de bloques completos al azar. El ensayo se repite durante tres temporadas consecutivas (tres años), conservando la misma estructura experimental y dentro de la misma superficie. Los datos, expresados en t/ha, se detallan en las tablas 6.11, 6.12 y 6.13.

Tabla 6.11. Rendimientos de caña de azúcar (t/ha) obtenidos en la temporada 1.

N°	Tratamientos	Repeticiones					Total	Promedios
		I	II	III	IV	V		
1	Dist. 1	89,0	81,9	78,2	85,3	94,1	428,5	85,7
2	Dist. 2	112,5	98,2	118,2	100,6	110,7	540,2	108,0
3	Dist. 3	87,3	85,2	79,1	75,9	74,2	401,7	80,3
4	Dist. 4	98,1	90,2	81,8	99,1	88,4	457,6	91,5
Total		386,9	355,5	357,3	360,9	367,4	1828,0	

Tabla 6.12. Rendimientos de caña de azúcar (t/ha) obtenidos en la temporada 2.

N°	Tratamientos	Repeticiones					Total	Promedios
		I	II	III	IV	V		
1	Dist. 1	100,1	94,8	89,9	99,1	101,3	485,2	97,0
2	Dist. 2	122,4	109,6	113,4	108,9	125,4	579,7	115,9
3	Dist. 3	65,3	55,1	71,2	81,5	66,1	339,2	67,8
4	Dist. 4	96,2	89,7	94,6	79,1	94,0	453,6	90,7
Total		384,0	349,2	369,1	368,6	386,8	1857,7	

Tabla 6.13. Rendimientos de caña de azúcar (t/ha) obtenidos en la temporada 3.

N°	Tratamientos	Repeticiones					Total	Promedios
		I	II	III	IV	V		
1	Dist. 1	105,4	96,8	109,4	100,0	112,7	524,3	104,86
2	Dist. 2	117,3	124,1	108,6	114,7	107,0	571,7	114,34
3	Dist. 3	45,7	66,1	54,7	49,8	50,0	266,3	53,26
4	Dist. 4	102,4	91,8	87,2	96,1	101,6	479,1	95,82
Total		370,8	378,8	359,9	360,6	371,3	1841,4	

Desde las tablas 6.11, 6.12 y 6.13 se obtienen las tablas 6.14, 6.15 y 6.16 que servirán para definir las sumas de cuadrados de los factores y las interacciones respectivas. En este caso, las tablas de doble entrada se formarán entre los tratamientos y los bloques, las temporadas y los bloques, y los tratamientos con las temporadas.

Tabla 6.14. Sumatorias para obtención de los efectos individuales y de interacción entre tratamientos y bloques.

Bloques	Tratamientos				Total
	Dist. 1	Dist. 2	Dist. 3	Dist. 4	
I	294,5	352,2	198,3	296,7	1141,7
II	273,5	331,9	206,4	271,7	1083,5
III	277,5	340,2	205,0	263,6	1086,3
IV	284,4	324,2	207,2	274,3	1090,1
V	308,1	343,1	190,3	284,0	1125,5
Total	1438,0	1691,6	1007,2	1390,3	5527,1

Tabla 6.15. Sumatorias para obtención de los efectos individuales y de interacción entre temporadas y bloques.

Bloques	Temporadas			Total
	Temp. 1	Temp. 2	Temp. 3	
I	386,9	384,0	370,8	1141,7
II	355,5	349,2	378,8	1083,5
III	357,3	369,1	359,9	1086,3
IV	360,9	368,6	360,6	1090,1
V	367,4	386,8	371,3	1125,5
Total	1828,0	1857,7	1841,4	5527,1

Tabla 6.16. Sumatorias para obtención de los efectos individuales y de interacción entre tratamientos y temporadas.

Tratamientos	Temporadas			Total
	Temp. 1	Temp. 2	Temp. 3	
Dist. 1	428,5	485,2	524,3	1438,0
Dist. 2	540,2	579,7	571,7	1691,6
Dist. 3	401,7	339,2	266,3	1007,2
Dist. 4	457,6	453,6	479,1	1390,3
Total	1828,0	1857,7	1841,4	5527,1

El factor de corrección es:

$$FC = \frac{(\sum Y_{ijk})^2}{n} = \frac{(1828,0 + 1857,7 + 1841,4)^2}{60} = 509147,2$$

Cálculos de las sumas de cuadrados:

- SC del total de los datos:

$$SC_{Tot} = \sum Y_{ijk}^2 - FC$$

$$SC_{Tot} = [89,0^2 + 112,5^2 + \dots + 50,0^2 + 101,6^2] - 509147,2 = 21408,5$$

- SC de tratamientos:

$$SC_t = \frac{\sum (\sum t_i)^2}{b \times T} - FC$$

$$SC_t = \frac{1438,0^2 + 1691,6^2 + 1007,2^2 + 1390,3^2}{5 \times 3} - 509147,2 = 15968,8$$

- SC de bloques:

$$SC_b = \frac{\sum (\sum b_j)^2}{t \times T} - FC$$

$$SC_b = \frac{1141,7^2 + 1083,5^2 + 1086,3^2 + 1090,1^2 + 1125,5^2}{4 \times 3} - 509147,2 = 233,4$$

- SC de la interacción tratamientos por bloques (t x B) o error a:

$$SC_{E(a)} = \frac{\sum(\sum tb_{ij})^2}{T} - FC - SC_t - SC_b$$

$$SC_{E(a)} = \frac{294,5^2 + 273,5^2 + \dots + 284,0^2}{3} - 509147,2 - 15968,8 - 233,4 = 463,1$$

- SC de temporada:

$$SC_{Temp} = \frac{\sum(\sum T_k)^2}{t \times b} - FC$$

$$SC_{Temp} = \frac{1828,0^2 + 1857,7^2 + 1841,4^2}{4 \times 5} - 509147,2 = 22,2$$

- SC de la interacción temporada por bloques (T x B) o error b:

$$SC_{E(b)} = \frac{\sum(\sum T b_{jk})^2}{t} - FC - SC_{Temp} - SC_b$$

$$SC_{E(b)} = \frac{386,9^2 + 355,5^2 + \dots + 371,3^2}{4} - 509147,2 - 22,2 - 233,4 = 218,3$$

- SC de la interacción tratamiento por temporada (t x T):

$$SC_{t \times T} = \frac{\sum(\sum t T_{ik})^2}{b} - FC - SC_t - SC_{Temp}$$

$$SC_{t \times T} = \frac{428,5^2 + 540,2^2 + \dots + 479,1^2}{5} - 509147,2 - 15968,8 - 22,2 = 2992,5$$

- SC de la interacción de tratamientos, temporada y bloques (t x T x B) o error c:

$$SC_{E(c)} = SC_{Tot} - SC_t - SC_b - SC_{E(a)} - SC_{Temp} - SC_{E(b)} - SC_{t \times T}$$

$$SC_{E(c)} = 21408,5 - 15968,8 - 233,4 - 463,1 - 22,2 - 218,3 - 2992,5 = 1510,2$$

El análisis de varianza totalmente desarrollado de este ejemplo se describe en la tabla 6.17. Como puede observarse, se registran diferencias significativas ($p < 0,05$) entre los tratamientos evaluados; sin embargo, este efecto, según la interacción de tratamientos por temporada, depende de esta última fuente de variación. Es decir, los distanciamientos de siembra tienen respuestas diferentes según la temporada en que se realizó el ensayo. En este caso, habrá que hacer un análisis individual por temporada para poder observar, en el caso de que así ocurra, si es que el tratamiento con la mayor media se mantiene en cada una de ellas. Si esto último no se comprueba se tendrá que seguir probando nuevos distanciamientos, de tal forma que se pueda identificar uno que sea constante en el tiempo. El efecto no significativo de las temporadas ($p \geq 0,05$) revela que su acción independiente no incide en los rendimientos de caña de azúcar, distinto a su acción en combinación con los tratamientos.

Tabla 6.17. Análisis de varianza del ejemplo de aplicación 6.2.

Fuentes de variación	GL	SC	CM	F	P
Total	59	21408,5			
Tratamientos (t)	3	15968,8	5322,9	137,90	0,000
Bloques (B)	4	233,4	58,4	1,51	0,260
Error a (t x B)	12	463,1	38,6		
Temporadas (T)	2	22,2	11,1	0,41	0,680
Error b (T x B)	8	218,3	27,3		
Interacción t x T	6	2992,5	498,8	7,93	0,000
Error c (t x B x T)	24	1510,2	62,9		
CV = 7,66%					
$r^2 = 0,90$					

Nótese que el estadígrafo F de prueba en el ANDEVA de la tabla 6.17 se define según el tipo de error relacionado a las fuentes de variación: tratamientos, bloques, temporadas y la interacción t x T. Todas estas fuentes se analizan bajo la consideración de efectos fijos, dado que ninguna toma como denominador al cuadrado medio del error experimental total.

En las tablas 6.18 y 6.19 se resumen los análisis de varianza y la aplicación de la prueba de comparación de medias de Tukey por cada temporada. Los tratamientos presentan diferencias significativas ($p < 0,05$) en cada una de estas temporadas, destacándose de forma constante el distanciamiento de siembra 2 con las medias más altas y diferentes, estadísticamente, al resto de medias (tabla 6.19). Esto es, hay un tratamiento que mantiene la mejor respuesta a través del tiempo y, por tanto, puede ser recomendado. También puede observarse que hay consistencia en los resultados según los coeficientes de variación (CV bastante menores al 30%) y los coeficientes de determinación (r^2 arriba del 0,70) que se han incluido en la tabla 6.18.

Tabla 6.18. Resultados de los análisis de varianza por temporada del ejemplo de aplicación 6.2.

Fuentes de variación	GL	Valor P por temporada		
		Temp. 1	Temp. 2	Temp. 3
Total	19			
Distanciamientos	3	0,0003	0,0000	0,0000
Bloques	4	0,5484	0,4273	0,8923
Error experimental	12			
CV:		7,80%	7,93%	8,37%
r^2 :		0,79	0,90	0,94

Tabla 6.19. Separación de medias por temporada con la aplicación de la prueba de Tukey.

Tratamientos	Temporadas		
	Temp. 1	Temp. 2	Temp. 3
Dist. 1	85,7 b	97,0 b	104,9 ab
Dist. 2	108,0 a	115,9 a	114,3 a
Dist. 3	80,3 b	67,8 c	53,3 c
Dist. 4	91,5 b	90,7 b	95,8 b
Medias con letras iguales no difieren significativamente ($p > 0,05$).			

Experimentos replicados en el espacio (localidad) y en el tiempo (temporada)

Para una generalización mucho más extensa, en los programas de investigación se puede prever que un experimento se repita simultáneamente en varias localidades de una región y durante varias temporadas o años. En esta situación se tienen dos condiciones que definen la existencia de efectos de interacción y/o efectos anidados, tales como:

1. Un experimento que se repite en unas mismas localidades durante varias temporadas.
2. Un experimento que se repite en diferentes localidades en cada una de las temporadas subsiguientes que dura el proceso de investigación.

Como podrá colegirse en el primer caso, las localidades se valorarán en cada una de las temporadas, generando así un efecto de interacción o de cruzamiento sin restricción; así también puede observarse que los bloques estarán anidados en la interacción de localidad por temporada, ya que las condiciones de estas dos fuentes de variabilidad serán distintas tanto en el espacio como en el tiempo aún a pesar de que se empleen los mismos bloques en cada localidad.

En la segunda condición, las localidades serán distintas en cada temporada, estableciéndose, por consiguiente, un efecto anidado (de localidad en temporada). Otro efecto con anidación doble identificable en esta segunda condición, es el de bloques en las localidades y, estas últimas, en las temporadas. El doble anidamiento se produce al desarrollar el experimento en distintos sitios y en distintas temporadas, de allí que es esperable que el efecto de los bloques sea diferente en cada uno de estos factores. Nótese que los bloques no están anidados en la interacción de localidades por temporadas como en el primer caso, debido a que dicha interacción no sería posible. Los esquemas de ANDEVA para cada caso se describen en las tablas 6.20 y 6.21.

Tabla 6.20. Esquema de análisis de varianza para un experimento replicado en las mismas localidades durante varias temporadas, utilizando una DBCA.

Fuentes de variación	Grados de libertad	Sumas de cuadrados
Total	$tLtb - 1$	$\sum Y_{ijkl}^2 - FC$
Tratamientos (t)	$t - 1$	$\frac{\sum (\sum t_i)^2}{L \times T \times b} - FC$
Localidad (L)	$L - 1$	$\frac{\sum (\sum L_j)^2}{t \times T \times b} - FC$
Temporada (T)	$T - 1$	$\frac{\sum (\sum T_k)^2}{t \times L \times b} - FC$
Bloques (Localidad x Temporada)	$LT(b - 1)$	$\frac{\sum (\sum (LT)_{jkl})^2}{t} - FC - SC_L - SC_T - SC_{L \times T}$
Interacción t x L	$(t - 1)(L - 1)$	$\left[\frac{\sum (\sum (tL)_{ij})^2}{T \times b} - FC \right] - SC_t - SC_L$
Interacción t x T	$(t - 1)(T - 1)$	$\left[\frac{\sum (\sum (tT)_{ik})^2}{L \times b} - FC \right] - SC_t - SC_T$
Interacción L x T	$(L - 1)(T - 1)$	$\left[\frac{\sum (\sum (LT)_{jk})^2}{t \times b} - FC \right] - SC_L - SC_T$

Interacción t x L x T	$(t - 1)(L - 1)(T - 1)$	$\left[\frac{\sum(\sum(tLT)_{ijk})^2}{b} - FC \right] - SC_t - SC_L$ $- SC_T - SC_{t \times L}$ $- SC_{t \times T} - SC_{L \times T}$
Error experimental	$(tLT - 1)(b - 1)$	Por sustracción

t: tratamientos, *L*: localidades, *T*: temporada *b*: bloques, SC_t : suma de cuadrados de tratamientos, SC_L : suma de cuadrados de localidades, SC_T : suma de cuadrados de temporada SC_b : suma de cuadrados de bloques, $SC_{t \times L}$: suma de cuadrados de tratamientos por localidades, $SC_{t \times T}$: suma de cuadrados de tratamientos por temporada, $SC_{L \times T}$: suma de cuadrados de localidades por temporada.
NOTA: el efecto de los bloques está anidado en la interacción de localidades por temporada.

Tabla 6.21. Esquema de análisis de varianza para un experimento replicado en diferentes localidades durante varias temporadas, utilizando una DBCA.

Fuentes de variación	Grados de libertad	Sumas de cuadrados
Total	$tTlb - 1$	$\sum Y_{ijkl}^2 - FC$
Tratamientos (t)	$t - 1$	$\frac{\sum(\sum t_i)^2}{L \times T \times b} - FC$
Temporada (T)	$T - 1$	$\frac{\sum(\sum T_k)^2}{t \times L \times b} - FC$
Localidad (Temporada)	$T(L - 1)$	$\frac{\sum(\sum L_j)^2}{t \times T \times b} - \frac{\sum(\sum T_k)^2}{t \times L \times b}$
Bloques [Localidad(Temporada)]	$TL(b - 2)$	$\frac{\sum(\sum b_l)^2}{t \times L \times T} - \frac{\sum(\sum L_j)^2}{t \times T \times b} - \frac{\sum(\sum T_k)^2}{t \times L \times b} - FC$
Interacción t x L	$(t - 1)(L - 1)$	$\left[\frac{\sum(\sum tL_{ij})^2}{T \times b} - FC \right] - SC_t - SC_L$
Interacción t x T	$(t - 1)(T - 1)$	$\left[\frac{\sum(\sum tT_{ik})^2}{L \times b} - FC \right] - SC_t - SC_T$
Error experimental	$(tLT - 1)(b - 1)$	Por sustracción

Para explicitar un poco más el caso de experimentos replicados en el espacio y en el tiempo, a continuación, se expone un ejemplo considerando las mismas localidades en temporadas subsiguientes para hacer uso del planteamiento de ANDEVA detallado en la tabla 6.20.

Ejemplo de aplicación 6.3

En dos localidades (Guamaní y Belén) se evalúa el rendimiento (kg/ha) de cuatro híbridos de maíz durante tres años consecutivos y en la época seca anual. Para esto se ha utilizado la misma distribución experimental basada en bloques completos al azar, con cinco repeticiones. Los datos se indican en las tablas 6.22 y 6.23, según la localidad.

Tabla 6.22. Rendimientos de maíz (kg/ha) en la localidad “Guamani” en los años 1, 2 y 3.

	Tratamientos	Bloques					Total
		I	II	III	IV	V	
AÑO 1	Híbrido 01	4424,8	3854,9	4024,9	3628,5	3575,5	19508,6
	Híbrido 02	5115,7	4983,9	4982,7	5456,1	5148,3	25686,7
	Híbrido 03	5826,2	6563,5	8346,6	7374,6	6242,3	34353,2
	Híbrido 04	4985,2	5451,9	7114,2	5059,0	5635,7	28246,0
		20351,9	20854,2	24468,4	21518,2	20601,8	107794,5
AÑO 2	Híbrido 01	3993,6	3658,9	2591,0	6916,0	2241,4	19400,9
	Híbrido 02	4989,9	5140,7	4684,7	5257,1	4904,3	24976,7
	Híbrido 03	5040,0	5424,8	7436,2	7075,1	8781,2	33757,3
	Híbrido 04	4717,3	5551,6	1593,5	2669,7	4685,9	19218,0
		18740,8	19776,0	16305,4	21917,9	20612,8	97352,9
AÑO 3	Híbrido 01	5191,2	5102,7	2086,7	5867,0	6075,4	24323,0
	Híbrido 02	5062,1	4987,1	4843,7	5062,0	4781,2	24736,1
	Híbrido 03	9356,8	7211,4	5358,9	6394,3	7637,7	35959,1
	Híbrido 04	5298,3	6451,7	4046,1	6079,2	2455,3	24330,6
		24908,4	23752,9	16335,4	23402,5	20949,6	109348,8

Tabla 6.23. Rendimientos de maíz (kg/ha) en la localidad “Belén” en los años 1, 2 y 3.

	Tratamientos	Bloques					Total
		I	II	III	IV	V	
AÑO 1	Híbrido 01	3551,5	2862,6	1400,3	4706,2	5071,6	17592,2
	Híbrido 02	8027,8	7995,5	7875,1	7632,4	8048,8	39579,6
	Híbrido 03	5739,5	3296,5	9554,6	6344,4	4469,4	29404,4
	Híbrido 04	3779,8	5918,9	4191,7	5786,5	5550,9	25227,8
		21098,6	20073,5	23021,7	24469,5	23140,7	111804,0
AÑO 2	Híbrido 01	3904,1	3012,4	3813,1	5057,7	5110,5	20897,8
	Híbrido 02	8044,6	8307,6	8053,4	8023,6	8504,1	40933,3
	Híbrido 03	4886,3	7828,3	4140,3	6606,1	7127,9	30588,9
	Híbrido 04	5911,6	1838,5	3619,4	4337,7	4391,3	20098,5
		22746,6	20986,8	19626,2	24025,1	25133,8	112518,5
AÑO 3	Híbrido 01	3485,8	5113,6	3682,6	4964,0	1266,8	18512,8
	Híbrido 02	7702,6	7766,2	7650,9	8264,3	7926,9	39310,9
	Híbrido 03	7665,2	6998,8	6090,2	8321,4	6163,7	35239,3
	Híbrido 04	6377,8	3425,0	6437,4	3857,7	5575,1	25673,0
		25231,4	23303,6	23861,1	25407,4	20932,5	118736,0

A partir de las tablas 6.22 y 6.23 se elaboran las tablas 6.24, 6.25, 6.26, 6.27 y 6.28 para la obtención de las sumas de cuadrados de los factores, las interacciones y los valores de anidamiento.

Tabla 6.24. Sumatorias para obtener la interacción de tratamientos por temporada.

Tratamientos	Temporada			Total
	1	2	3	
Híbrido 01	37100,8	40298,7	42835,8	120235,3
Híbrido 02	65266,3	65910,0	64047,0	195223,3
Híbrido 03	63757,6	64346,2	71198,4	199302,2
Híbrido 04	53473,8	39316,5	50003,6	142793,9
	219598,5	209871,4	228084,8	657554,7

Tabla 6.25. Sumatorias para obtener la interacción de tratamientos por localidad.

Tratamientos	Localidad		Total
	Guamaní	Belén	
Híbrido 01	63232,5	57002,8	120235,3
Híbrido 02	75399,5	119823,8	195223,3
Híbrido 03	104069,6	95232,6	199302,2
Híbrido 04	71794,6	70999,3	142793,9
	314496,2	343058,5	657554,7

Tabla 6.26. Sumatorias para obtener la interacción de temporada por localidad.

Temporadas	Localidad		Total
	Guamaní	Belén	
1	107794,5	111804,0	219598,5
2	97352,9	112518,5	209871,4
3	109348,8	118736,0	228084,8
	314496,2	343058,5	657554,7

Tabla 6.27. Sumatorias para obtener la interacción de segundo orden entre tratamientos, localidades y temporadas.

Tratamientos	Guamaní			Belén		
	1	2	3	1	2	3
Híbrido 01	19508,6	19400,9	24323,0	17592,2	20897,8	18512,8
Híbrido 02	25686,7	24976,7	24736,1	39579,6	40933,3	39310,9
Híbrido 03	34353,2	33757,3	35959,1	29404,4	30588,9	35239,3
Híbrido 04	28246,0	19218,0	24330,6	25227,8	20098,5	25673,0

Tabla 6.28. Sumatorias para obtener la suma de cuadrados de bloques anidados en la interacción de localidad por temporada.

Loc.	Temp.	Bloques				
		I	II	III	IV	V
Loc. Guamani	1	20351,9	20854,2	24468,4	21518,2	20601,8
	2	18740,8	19776,0	16305,4	21917,9	20612,8
	3	24908,4	23752,9	16335,4	23402,5	20949,6
Loc. Belén	1	21098,6	20073,5	23021,7	24469,5	23140,7
	2	22746,6	20986,8	19626,2	24025,1	25133,8
	3	25231,4	23303,6	23861,1	25407,4	20932,5
		133077,7	128747,0	123618,2	140740,6	131371,2

El cálculo de las sumas de cuadrados, según lo indicado en la tabla 6.20, se detalla en los siguientes pasos:

El factor de corrección, tomando las sumas totales de las tablas 6.22 y 6.23, es:

$$FC = \frac{(\sum Y_{ijkl})^2}{n} = \frac{(107794,5 + 97352,9 + \dots + 118736,0)^2}{120} = 3603151529,10$$

Cálculos de las sumas de cuadrados (SC):

- SC del total de los datos (tablas 6.22 y 6.23):

$$SC_{Tot} = \sum Y_{ijkl}^2 - FC$$

$$SC_{Tot} = [4424,8^2 + 5115,7^2 + \dots + 6163,7^2 + 5575,1^2] - 3603151529,10 = 393639738,39$$

- SC de tratamientos (tabla 6.24 o 6.25):

$$SC_t = \frac{\sum (\sum t_i)^2}{L \times T \times b} - FC$$

$$SC_t = \frac{120235,3^2 + 195223,3^2 + 199302,2^2 + 142793,9^2}{2 \times 3 \times 5} - 3603151529,10 = 152852771,93$$

- SC de localidad (tabla 6.25 o 6.26):

$$SC_L = \frac{\sum (\sum L_j)^2}{t \times T \times b} - FC$$

$$SC_L = \frac{314196,2^2 + 343058,5^2}{4 \times 3 \times 5} - 3603151529,10 = 6798374,84$$

- SC de temporada (tabla 6.24 o 6.26):

$$SC_T = \frac{\sum (\sum T_k)^2}{t \times L \times b} - FC$$

$$SC_T = \frac{219598,5^2 + 209871,4^2 + 228084,8^2}{4 \times 2 \times 5} - 3603151529,10 = 4153014,18$$

- SC de la interacción de localidad por temporada (tabla 6.26):

$$SC_{L \times T} = \left[\frac{\sum(\sum(LT)_{jk})^2}{t \times b} - FC \right] - SC_L - SC_T$$

$$SC_{L \times T} = \left[\frac{107794,5^2 + 97352,9^2 + \dots + 118736,0^2}{4 \times 5} - 3603151529,10 \right] - 6798374,84$$

$$- 4153014,18 = 1556401,09$$

- SC de bloques anidados en la interacción de localidad por temporada (tabla 6.28):

$$SC_{B(L \times T)} = \frac{\sum(\sum(LTb)_{jkl})^2}{t} - FC - SC_L - SC_T - SC_{L \times T}$$

$$SC_{B(L \times T)} = \frac{20351,9^2 + 18740,8^2 + \dots + 20932,5^2}{4} - 3603151529,10 - 6798374,84$$

$$- 4153014,18 - 1556401,09 = 30289824,92$$

- SC de la interacción de tratamiento por localidad (tabla 6.25):

$$SC_{t \times L} = \left[\frac{\sum(\sum(tL)_{ij})^2}{T \times b} - FC \right] - SC_t - SC_L$$

$$SC_{t \times L} = \left[\frac{63232,5^2 + 75399,5^2 + \dots + 70999,3^2}{3 \times 5} - 3603151529,10 \right] - 152852771,93$$

$$- 6798374,84 = 62903380,61$$

- SC de la interacción de tratamientos por temporada (tabla 6.24):

$$SC_{t \times T} = \left[\frac{\sum(\sum(tT)_{ik})^2}{L \times b} - FC \right] - SC_t - SC_T$$

$$SC_{t \times T} = \left[\frac{37100,8^2 + 65266,3^2 + \dots + 50003,6^2}{2 \times 5} - 3603151529,10 \right] - 152852771,93$$

$$- 4153014,18 = 11989506,29$$

- SC de la interacción de tratamientos, localidad y temporada (tabla 6.27):

$$SC_{t \times L \times T} = \left[\frac{\sum(\sum(tLT)_{ijk})^2}{b} - FC \right] - SC_t - SC_L - SC_T - SC_{t \times L} - SC_{t \times T} - SC_{L \times T}$$

$$SC_{t \times L \times T} = \left[\frac{19508,6^2 + 25686,7^2 + \dots + 25673,0^2}{5} - 3603151529,10 \right] - 152852771,93$$

$$- 6798374,84 - 4153014,18 - 62903380,61 - 11989506,29$$

$$- 1556401,09 = 3387504,74$$

- SC del error experimental:

$$SC_E = SC_{Tot} - \text{demás } SC = 119708959,77$$

En la tabla 6.29 se detalla el análisis de varianza completo del ejemplo de aplicación 6.3. En este pueden observarse efectos significativos ($p < 0,05$) entre tratamientos, localidad y en la

interacción de tratamientos por localidad. Aparentemente los híbridos responden de forma diferente según la localidad en la que se siembran, lo cual demanda que se realice un ANDEVA por cada una de estas localidades y aplicar un test de comparación de medias para verificar cuál es la semilla con los rendimientos más prometedores. En cuanto a las temporadas, éstas no tienen incidencia en los rendimientos de los híbridos, mientras que los lotes seleccionados para los ensayos han revelado una relativa homogeneidad, según los efectos no significativos ($p > 0,05$) para estas dos fuentes de variación analizadas. Además, de acuerdo al coeficiente de variación ($CV \leq 30\%$) y el de determinación ($0 \leq r^2 \leq 0,60$), lo resuelto en este ensayo revela un manejo adecuado del mismo y puede admitírsele cierta confiabilidad.

Tabla 6.29. Análisis de varianza del ejemplo de aplicación 6.3.

Fuentes de variación	GL	SC	CM	F	P
Total	119	393639738,39			
Tratamientos (t)	3	152852771,93	50950923,98	30,645	0,0000
Localidad (L)	1	6798374,84	6798374,84	4,089	0,0469
Temporada (T)	2	4153014,18	2076507,09	1,249	0,2929
Bloques (Localidad x Temporada)	24	30289824,92	1262076,04	0,759	0,7725
Interacción t x L	3	62903380,61	20967793,54	12,611	0,0000
Interacción t x T	6	11989506,29	1998251,05	1,202	0,3153
Interacción L x T	2	1556401,092	778200,55	0,468	0,6281
Interacción t x L x T	6	3387504,743	564584,12	0,340	0,9137
Error experimental	72	119708959,77	1662624,44		

CV: 23,53%

r^2 : 0,70

Según el análisis de varianza de la tabla 6.30, los híbridos evaluados presentan diferencias significativas en las dos localidades. No obstante, en estas últimas, distintos híbridos tienen las medias más altas. En este caso, de acuerdo a la prueba de Tukey de la tabla 6.31, las mejores respuestas dependen de la zona, pudiéndose recomendar al híbrido 03 para la localidad “Guamaní” y al híbrido 02 para la localidad “Belén”.

Tabla 6.30. Resultados de los análisis de varianza por localidad del ejemplo de aplicación 6.3.

Fuentes de variación	GL	Valor P por temporada	
		Guamaní	Belén
Total	19		
Distanciamientos	3	0,0000	0,0000
Bloques	4	0,5786	0,6199
Error experimental	52		

CV: 23,54%

r^2 : 0,46

22,13%

0,65

Tabla 6.31. Aplicación de la prueba de Tukey para la separación de medias por localidad del ejemplo de aplicación 6.3.

Tratamientos	Temporadas	
	Guamáni	Belén
Híbrido 01	4215,5 b	3800,2 c
Híbrido 02	5026,6 b	7988,3 a
Híbrido 03	6938,0 a	6348,8 b
Híbrido 04	4786,3 b	4733,3 c

Medias con letras iguales no difieren significativamente ($p > 0,05$).

Distribución de bloques aumentados

La distribución experimental conocida como *bloques aumentados*⁴⁰ es aplicable en los casos en donde se tiene un gran número de tratamientos, como en aquellos que tratan sobre la selección de semillas que se valoran en investigaciones preliminares de fitomejoramiento. Son ensayos que se utilizan cuando el material experimental (tratamientos) es escaso y no se puede evaluar más de una repetición, o por limitaciones de los terrenos sobre los cuales se desarrollarán estas experiencias; de esta forma, también pueden ser vistos como alternativas que contrarrestan parcialmente el efecto de la heterogeneidad de los suelos y permiten cierta eficiencia económica en un programa de investigación por el hecho de que se utiliza menos material y área experimental. Es importante que se entienda que esta distribución fue ideada para un proceso inicial de cribado de tratamientos, circunscritos a experimentos exploratorios, a partir de los cuales se pretende seleccionar aquellas alternativas de prueba que tengan efectos claramente evidentes para utilizarlos posteriormente en ensayos con mayor fundamento estadístico⁴¹.

La característica de los bloques aumentados consiste en distribuir un grupo importante de tratamientos sin repeticiones (los que aumentan los bloques), junto a otros tratamientos que son tomados como testigos o referentes (tratamientos estándar). Los tratamientos testigos pueden ser asignados bajo distribuciones en bloques al azar, cuadro latino o latice, dado que estos sí se valoran a través de varias repeticiones; siendo los que finalmente definen la estructura del diseño en el campo. En este sentido, se cumple lo manifestado por el autor de esta distribución (Federer, 1961) al indicarnos que: “el diseño experimental seleccionado debe ajustarse a las condiciones experimentales, en lugar de que los tratamientos y las unidades experimentales se ajusten a los requerimientos del diseño experimental”.

En una distribución con bloques aumentados, son los tratamientos testigos los que permiten estimar el error experimental debido a que son valorados a través de varias repeticiones; con lo cual se hace posible calcular los errores estándar en las diferentes comparaciones que se realicen entre medias de tratamientos (aunque en los tratamientos de prueba no se tienen

⁴⁰ Esta modificación de las distribuciones con bloqueo es aplicable dentro del criterio de análisis de las distribuciones de bloques completos al azar, cuadro latino o latice. Fue ideada por Federer allá por los años 50 del siglo pasado.

⁴¹ Se hace esta mención en el sentido de que, como en su momento lo dijo Fisher, para estimar el error experimental se requiere que los tratamientos de un experimento tengan repeticiones; lo cual no es posible cuando un tratamiento se valora con una sola repetición, como aquellos que se evalúan bajo una distribución en bloques aumentados. El investigador debe tener presente que, los resultados de un experimento con bloques aumentados, no sirven para tomar decisiones en procesos de productividad económica; solo son resultados preliminares para la búsqueda de alternativas cuya efectividad deberá valorarse posteriormente a través de experimentos con un mayor fundamento estadístico de cara a su posible recomendación.

medias porque son datos únicos), utilizando pruebas como la DMS, la de Duncan o la de Tukey. Las repeticiones de los tratamientos testigos también hacen posible estimar el efecto de los bloques y, con ello, ajustar las medias de los tratamientos asignados a los bloques sin repeticiones.

El proceso práctico de análisis de una distribución con bloques aumentados comprende la realización de un ANDEVA normal, ajustado a un modelo de bloques completos al azar, cuadro latino o laticé, según sea la distribución de base de los tratamientos testigos. Este ANDEVA permitirá obtener la varianza del error experimental (cuadrado medio del error experimental), con lo que se podrá obtener el error estándar respectivo y realizar las diferentes comparaciones de medias entre tratamientos testigos, tratamientos de prueba o entre un tratamiento testigo y uno de prueba. Así también, la estimación del efecto de bloques obtenida sólo con los tratamientos testigos permitirá realizar una corrección de los datos de cada tratamiento, previo al proceso de comparación de medias. Para ampliar la explicación de esta temática, veamos un ejemplo aplicado a una distribución de bloques al azar.

Ejemplo de aplicación 6.4

Se tiene un experimento de característica exploratoria, en el cual se ha evaluado el efecto de tres fórmulas tradicionales de fertilización orgánica (A, B, C), tipo *Biol*, en el rendimiento de maíz y bajo una distribución de bloques al azar. Adicionalmente se valoraron 15 formulaciones (F_i) (de prueba) de este tipo de fertilización en el mismo experimento, no obstante, dada las escasas cantidades de estas nuevas formulaciones que se tuvieron, solo fue posible tener una única repetición por cada una de ellas. La estructura del diseño en campo con los datos respectivos (kg/parcela) se detalla en la figura 6.4 (croquis de campo).

Bloq. I	A 3,14	F ₁ 2,13	B 2,25	F ₂ 2,96	C 3,32	F ₃ 2,45
Bloq. II	F ₄ 2,65	A 2,99	F ₅ 3,14	B 2,55	F ₆ 2,54	C 3,12
Bloq. III	C 2,58	F ₇ 2,67	A 3,45	F ₈ 3,33	B 1,96	F ₉ 2,14
Bloq. IV	F ₁₀ 2,89	C 3,79	F ₁₁ 2,15	A 3,10	F ₁₂ 1,98	B 2,68
Bloq. V	B 2,05	F ₁₃ 3,11	C 3,87	F ₁₄ 2,95	A 2,85	F ₁₅ 3,05

Figura 6.4. Esquema de la distribución de bloques aumentados del ejemplo de aplicación 6.4.

Como un primer paso se elabora la tabla 6.32 sólo con los datos de los tratamientos testigos (A, B y C). Aquí se obtienen los totales y los promedios, tanto de tratamientos como de bloques. Con los promedios de estos últimos se calcula el factor de ajuste de bloques (B_{aj}), obteniendo la diferencia entre el promedio de un bloque determinado ($\bar{Y}_{b(i)}$) y el promedio de todos los datos (\bar{Y}). Esto es:

$$B_{aj} = \bar{Y}_{b(i)} - \bar{Y}$$

Nótese que para obtener el factor de ajuste de bloques solo se toman en cuenta los datos de los tratamientos testigos.

Tabla 6.32. Detalle del cálculo del factor de ajuste de bloques (B_{aj})

Bloques	A	B	C	Total de bloque	Promedio de bloque	Factor de ajuste (B_{aj})
I	3,14	2,25	3,32	8,71	2,90	-0,01
II	2,99	2,55	3,12	8,66	2,89	-0,03
III	3,45	1,96	2,58	7,99	2,66	-0,25
IV	3,10	2,68	3,79	9,57	3,19	0,28
V	2,85	2,05	3,87	8,77	2,92	0,01
Total	15,53	11,49	16,68	43,70	2,91	
Promedios	3,11	2,30	3,34			

Una vez definidos los factores de ajuste de bloques, se corrigen los datos de los tratamientos de prueba (nuevas formulaciones), tal como se indica en la tabla 6.33.

Se realiza un ANDEVA de forma parcial hasta obtener el cuadrado medio del error experimental (CM_E). Es importante tener en cuenta que este análisis no se realiza con el propósito de verificar diferencias significativas entre los tratamientos, sino para la obtención del CM_E ; dado que, con esta varianza, es posible calcular los diferentes errores estándar que se necesitan para la aplicación de alguna prueba de comparación de medias.

Tabla 6.33. Rendimiento ajustado.

Tratamientos		Rendimiento	Factor de ajuste	Rendimiento ajustado
Fórmulas tradicionales	A	3,11	-	3,11
	B	2,30	-	2,30
	C	3,34	-	3,34
Fórmulas de prueba (Tratamientos aumentados)	F ₁	2,13	-0,01	2,14
	F ₂	2,96	-0,01	2,97
	F ₃	2,45	-0,01	2,46
	F ₄	2,65	-0,03	2,68
	F ₅	3,14	-0,03	3,17
	F ₆	2,54	-0,03	2,57
	F ₇	2,67	-0,25	2,92
	F ₈	3,33	-0,25	3,58
	F ₉	2,14	-0,25	2,39
	F ₁₀	2,89	0,28	2,61
	F ₁₁	2,15	0,28	1,87
	F ₁₂	1,98	0,28	1,70
	F ₁₃	3,11	0,01	3,10
	F ₁₄	2,95	0,01	2,94
	F ₁₅	3,05	0,01	3,04

Adviértase que el ANDEVA descrita en la tabla 6.34 se ha obtenido bajo un modelo de una distribución en bloques completos al azar únicamente con los tratamientos testigos, en donde el total de grados de libertad del modelo se deduce como: $GL_T = 2 + 4 + 8 = 14$.

Tabla 6.34. Análisis de varianza solo para obtención del cuadrado medio del error experimental.

Fuentes de variación	GL	SC	CM
Total	14	4,67	
Tratamientos	2	2,97	
Bloques	4	0,42	
Error experimental	8	1,28	0,159

Conociendo el valor del CM_E , los errores estándar de las medias ($S_{\bar{y}}$), según los tratamientos que se vayan a comparar, se establecen con las expresiones que se detallan en la tabla 6.35. Con estas fórmulas el investigador podrá aplicar alguna prueba de comparación de medias para ir definiendo los mejores tratamientos.

De forma axiomática, nótese que este tipo de experimento requiere un sumo cuidado tanto en el diseño como en el manejo del mismo, ya que toma el cuadrado medio del error experimental que se define a partir de los tratamientos testigos (los tratamientos que tienen repeticiones) para hacer comparaciones con y entre los tratamientos de prueba. De esto puede deducirse que, a un mayor número de tratamientos testigos y/o más repeticiones de éstos, se tendrá mucha mayor precisión en las comparaciones siempre que las posibles fuentes de ruido se mantengan controladas.

Tabla 6.35. Fórmulas para calcular el error estándar de medias para pruebas de comparación entre tratamientos.

Para comparar dos tratamientos testigos:	$S_{\bar{y}} = \sqrt{2CM_E/b}$
Para la comparación de dos tratamientos de prueba en un mismo bloque:	$S_{\bar{y}} = \sqrt{2CM_E}$
Para la comparación de dos tratamientos de prueba en bloques distintos:	$S_{\bar{y}} = \sqrt{2CM_E + 2CM_E/t}$
Para la comparación entre un tratamiento testigo y uno de prueba:	$S_{\bar{y}} = \sqrt{2CM_E \left(\frac{bt + t + b - 1}{bt} \right)}$

t : es el número de tratamientos testigos.

b : es el número de bloques del ensayo.

Distribución en cuadro latino con réplicas

Los temas tratados en los apartados que anteceden en este capítulo hacen alusión a experimentos con elevados números de tratamientos; sin embargo, sobre una superficie dada de terreno que espacialmente sea muy heterogénea en cuanto a pendientes, fertilidad, cambios texturales, etc., podríamos estar interesados en desarrollar ensayos de unos pocos tratamientos. Estas circunstancias de heterogeneidad podrían afectar más drásticamente cuando se trata de experimentos que involucran a cultivos perennes (generalmente frutales), que se siembran sobre superficies topográficamente irregulares, con amplios distanciamientos de siembra, muy a pesar de que el ensayo incluya unos pocos tratamientos.

Para situaciones como las descritas en el párrafo anterior, posibles en el ámbito agrícola, una alternativa que estructuralmente ofrece una ventaja en el sentido de poder controlar simultáneamente dos fuentes de variabilidad secundarias, plenamente identificadas, es la distribución en cuadro latino (DCL). Esta distribución genera un doble bloqueo para el control de dos fuentes de variación secundaria que, en el espacio de terreno dentro del cual se desarrolla la experiencia, actúan de forma perpendicular. Es una de las distribuciones muy recomendadas para experimentos de campo con fertilizantes, debido a que permite un mayor control del error experimental que una distribución como la de bloques completos al azar. Sin embargo, desde el punto de vista de la potencia estadística para evitar cometer el error II, una de sus desventajas importantes es la de generar pocos grados de libertad en el error experimental cuando se llevan a cabo ensayos con un reducido número de tratamientos. Así, por ejemplo, si quisiéramos emplear una DCL con apenas tres tratamientos, tendríamos solo dos grados de libertad en el error experimental; si son cuatro tratamientos, se tendrían seis grados de libertad en dicha fuente de variabilidad. Experimentos con estos pocos grados de libertad para la varianza desconocida no sirven en la práctica para tomar una decisión, debido a la escasa confiabilidad que ofrecen.

Una alternativa para soslayar el inconveniente mencionado en el párrafo anterior, es hacer réplicas de todo el ensayo estructurado bajo una DCL, dando así origen a la *distribución de cuadro latino con réplicas*. Las réplicas permiten incrementar los grados de libertad del error experimental, logrando así incrementar la precisión cuantitativa de un ensayo.

La replicación de una DCL puede abarcar repeticiones del experimento completo en cuanto a distintos sitios, distintas épocas, distintas condiciones ambientales, etc. El hecho de poder realizar réplicas en distintas épocas se constituye en una ventaja respecto al factor económico cuando éste tiene influencia en la investigación, pues al correr las réplicas durante un lapso de tiempo hay la oportunidad de prorratear los costos del proyecto. Según como se establezcan las réplicas del diseño, tomando en consideración la característica longitudinal o transversal de la investigación, es posible tener los siguientes casos:

Caso A: las unidades experimentales de las réplicas ocupan el mismo espacio o sitio, es decir, se consideran las mismas filas y columnas.

Caso B: las unidades experimentales ocupan sitios distintos en cada réplica, pero dentro del espacio de terreno, pueden ubicarse a la altura de las mismas filas o en la dirección de las mismas columnas.

Caso C: las unidades experimentales se ubican en distintos sitios en cada réplica, pero en posiciones de filas y columnas distintas.

El caso A se aplica para ensayos inmersos en *estudios longitudinales*, para cuyo desarrollo se utiliza el mismo lote experimental con el fin de hacer evaluaciones durante dos o más temporadas de cosecha. Es decir, un mismo tratamiento ocupa la misma unidad experimental (el mismo sitio, el mismo árbol, el mismo animal, etc.) en cada una de las réplicas que se hayan considerado para el ensayo. Es importante hacer una distinción cuando se trata de experimentos en ámbitos diferentes al agrícola, en los cuales la investigación puede desarrollarse desde un mismo instante sin necesidad de que las réplicas se realicen en distintos tiempos. Un ejemplo de lo último es lo que se puede dar en el ámbito agroindustrial cuando los lotes de materia prima, la mano de obra, los instrumentos, etc., son suficientes para la realización de todas las réplicas del DCL en un mismo momento.

Los casos B y C se aplican para los *estudios de característica transversal* en cuanto a las réplicas, dado que éstas se valoran a un mismo tiempo. En estas condiciones, un mismo tratamiento ocupa un sitio distinto en cada réplica que es evaluada simultáneamente. Dentro de las consideraciones de estos casos, puede suceder que, en el lote de terreno en donde se

desarrollará la investigación, las réplicas compartan las mismas filas o las mismas columnas (caso B, indicados en la figura 6.5a); o que, tanto filas como columnas sean distintas para cada réplica (caso C, indicado en la figura 6.5b). A los casos detallados en la figura 6.5a, por las formas que generan, ocasionalmente se los identifican como *rectángulos latinos*.

Las distintas formas de réplicas de cuadros latinos también generan diferentes resultados en los análisis de varianza que se utilizan en cada caso, producto de los distintos procedimientos que se aplican para definir las sumas de cuadrados y los grados de libertad de las fuentes bloqueadas (filas y columnas) y del error experimental. Estas formalidades son las que se describen en las tablas 6.36 y 6.37.

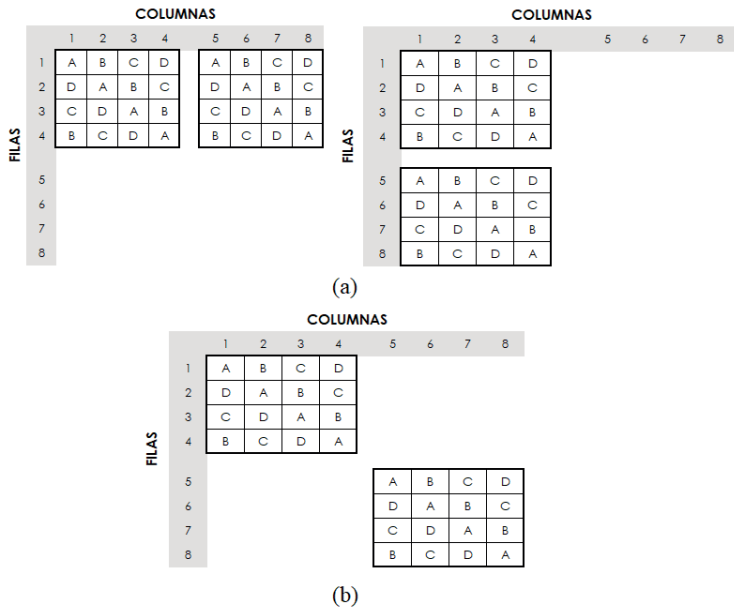


Figura 6.5. Esquema de un diseño en cuadro latino con cuatro tratamientos y dos réplicas, utilizando las mismas filas o las mismas columnas (a); o con diferentes filas y columnas (b).

Tabla 6.36. Grados de libertad para diseños en cuadro latino replicados, según los casos A, B y C.

Fuente de variación	Caso A (las mismas filas y columnas)	Caso B (cuando las filas son las mismas)	Caso B (cuando las columnas son las mismas)	Caso C (cuando filas y columnas son distintas)
Total	$Rt^2 - 1$	$Rt^2 - 1$	$Rt^2 - 1$	$Rt^2 - 1$
Tratamientos	$t - 1$	$t - 1$	$t - 1$	$t - 1$
Filas	$F - 1$	$F - 1$	$R(F - 1)$	$R(F - 1)$
Columnas	$C - 1$	$R(C - 1)$	$C - 1$	$R(C - 1)$
Réplicas	$R - 1$	$R - 1$	$R - 1$	$R - 1$
Error experimental	$Rt^2 - 3t - R + 3$	$(t - 1)(Rt - 2)$	$(t - 1)(Rt - 2)$	$(t - 1)(Rt - R - 1)$

Nótese que $t = F = C$.

Tabla 6.37. Formas de cálculo de las sumas de cuadrados para diseños en cuadro latino replicados, según los casos A, B y C.

Fuente de variación	Caso A (las mismas filas y columnas)	Caso B (cuando las filas son las mismas)	Caso B (cuando las columnas son las mismas)	Caso C (cuando filas y columnas son distintas)
Total	$\sum Y_{ijkl}^2 - FC$			
Tratamientos	$\frac{\sum(\sum t_i^2)}{Rt} - FC$			
Filas	$\frac{\sum(\sum F_{j+j'}^2)}{Rt} - FC$	$\frac{\sum(\sum F_{j+j'}^2)}{Rt} - FC$	$\frac{\sum(\sum F_j^2)}{t} - \frac{\sum(\sum R_i^2)}{t^2}$	$\frac{\sum(\sum F_j^2)}{t} - \frac{\sum(\sum R_i^2)}{t^2}$
Columnas	$\frac{\sum(\sum C_{k+k'}^2)}{Rt} - FC$	$\frac{\sum(\sum C_k^2)}{t} - \frac{\sum(\sum R_i^2)}{t^2}$	$\frac{\sum(\sum C_{k+k'}^2)}{Rt} - FC$	$\frac{\sum(\sum C_k^2)}{t} - \frac{\sum(\sum R_i^2)}{t^2}$
Réplicas	$\frac{\sum(\sum R_i^2)}{t^2} - FC$			
Error experimental	Por sustracción			

Ejemplo de aplicación 6.5

Se ha desarrollado un experimento en el cual se han valorado 4 tratamientos (tres dosis de nitrato de calcio y un testigo) en una plantación de palma africana al interior de un espacio de terreno que posee diferentes pendientes, tanto de este a oeste como de norte a sur. Bajo esta circunstancia y, buscando una mayor precisión, se estableció utilizar una distribución en cuadro latino. Por otro lado, debido al grado de confianza que tendría dicho experimento respecto de los grados de libertad del error experimental, éste se llevó a cabo con 2 réplicas dentro del mismo lote y bajo el caso C. La información del experimento y el análisis respectivo se detallan a continuación:

Suma total T_1 (Réplica 1)	Suma total T_2 (Réplica 2)	Nº total de datos n		Promedios generales	
		n_1	n_2	\bar{Y}_1	\bar{Y}_2
3595	4100	16	16	224,7	256,3
$T = 3595 + 4100 = 7695$		$n = 16 + 16 = 32$		$\bar{Y} = 240,5$	

Los grados de libertad de cada componente del análisis de varianza, considerando el caso C (tabla 5.30) se indican a continuación:

$$\begin{aligned}
 GL_T &= Rt^2 - 1 & GL_T &= 2 \times 4^2 - 1 = 31 \\
 GL_t &= t - 1 & GL_t &= 4 - 1 = 3 \\
 GL_F &= R(F - 1) & GL_F &= 2(4 - 1) = 6 \\
 GL_C &= R(C - 1) & GL_C &= 2(4 - 1) = 6 \\
 GL_R &= R - 1 & GL_R &= 2 - 1 = 1 \\
 GL_E &= (t - 1)(Rt - R - 1) & GL_E &= (4 - 1)(2 \times 4 - 2 - 1) = 15
 \end{aligned}$$

El factor de corrección de las sumas de cuadrados está dado por el valor:

$$FC = \frac{T^2}{Rt^2} = \frac{7695^2}{2 \times 4^2} = 1850407,03$$

Cálculo de las sumas de cuadrados:

- Suma de cuadrados para el total de los datos:

$$SC_T = \sum Y_{ijkl}^2 - FC$$

$$SC_T = (214^2 + 185^2 + \dots + 316^2 + 305^2) - 1850407,03 = 231942,0$$

- Suma de cuadrados para los tratamientos:

$$SC_t = \frac{\sum (\sum t_i)^2}{Rt} - FC$$

$$\begin{aligned}
 SC_t &= \frac{(647 + 776)^2 + (802 + 878)^2 + (1246 + 1350)^2 + (900 + 1096)^2}{2 \times 4} \\
 &\quad - 1850407,03 = 95913,1
 \end{aligned}$$

- Suma de cuadrados para las filas:

$$SC_F = \frac{\sum (\sum F_j^2)}{t} - \frac{\sum (\sum R_i^2)}{t^2}$$

$$SC_F = \frac{1050^2 + 1057^2 + \dots + 1084^2}{4} - \frac{3595^2 + 4100^2}{4^2} = 44544,2$$

- Suma de cuadrados para columnas:

$$SC_C = \frac{\sum (\sum C_k^2)}{t} - \frac{\sum (\sum R_i^2)}{t^2}$$

$$SC_C = \frac{937^2 + 851^2 + \dots + 1117^2}{4} - \frac{3595^2 + 4100^2}{4^2} = 51062,7$$

- Suma de cuadrados para las réplicas:

$$SC_R = \frac{\sum(\sum R_i^2)}{t^2} - FC$$

$$SC_R = \frac{3595^2 + 4100^2}{4^2} - 1850407,03 = 7969,5$$

- Suma de cuadrados del error experimental:

$$SC_E = SC_T - SC_t - SC_F - SC_C - SC_R$$

$$SC_E = 231942,0 - 95913,1 - 44544,2 - 51062,7 - 7969,5 = 32452,5$$

El análisis de varianza correspondiente es el que se detalla en la tabla 6.38. En esta tabla se observan diferencias significativas ($p < 0,05$) para los tratamientos, para las filas y para las columnas; mientras que las réplicas no presentan significancia en sus diferencias al nivel del 5% de probabilidad. El experimento presenta una confianza aceptable, según los valores de los coeficientes de variación ($CV = 19,34\%$) y de determinación ($r^2 = 0,86$).

Tabla 6.38. Análisis de varianza del ejemplo de aplicación 6.5 sobre DCL con réplica bajo el caso C.

Fuentes de variación	GL	SC	CM	F	P-valor
Total	31	231942,0			
Tratamientos	3	95913,1	31971,0	14,78	0,000
Filas	6	44544,2	7424,0	3,43	0,025
Columnas	6	51062,7	8510,5	3,93	0,015
Réplicas	1	7969,5	7969,5	3,68	0,074
Error experimental	15	32452,5	2163,5		
CV: 19,34%					
$r^2: 0,86$					

Los errores estándar de una media y de la diferencia entre dos medias pueden estimarse con $\sqrt{CME/Rt}$ y $\sqrt{2CME/Rt}$, respectivamente. La última expresión se empleará para las expresiones de comparación de medias.

Cuadros latinos con réplicas en experimentos cruzados

Una de las aplicaciones de los cuadros latinos con réplicas, utilizados para estudios de tipo longitudinal, es la formación de los llamados *diseños cruzados* o diseños “*crossover*”. Estos tipos de distribuciones experimentales consisten en aplicar todos los tratamientos, previamente establecidos, en cada una de las unidades experimentales. Esto es, en una misma unidad experimental, se aplica un tratamiento y se evalúa su efecto en un periodo, luego se aplica el siguiente tratamiento y se evalúa su efecto en un segundo periodo y así sucesivamente, hasta completar todos los tratamientos en estudio. Dado su carácter secuencial es que se dice que son aplicables en procesos de investigación longitudinal, a pesar de que también tienen carácter transversal en el sentido de que utilizan réplicas de cuadros latinos de forma simultánea.

Las ventajas más importantes de los diseños cruzados son la de implementarse relativamente con menores costos de investigación y de poder lograr una mayor precisión. Se logra reducir los costos al utilizar unas mismas unidades experimentales para evaluar varios tratamientos, de allí que pueda seleccionarse un número significativamente menor de elementos que harán más manejable todo el ensayo. Consecuentemente, gran parte de la varianza entre unidades experimentales que reciben un mismo tratamiento queda eliminada cuando los efectos de todos los tratamientos se miden sobre una misma unidad experimental, con lo que se logra reducir el error experimental y mejorar la precisión del experimento.

Por la naturaleza de los diseños cruzados, en la práctica son aplicables a experimentos con animales o con humanos; quedando casi impracticables en el ámbito de la agronomía, aunque pudiera pensarse su utilidad para cultivos perennes como los frutales. No obstante, la situación en contra sería el tiempo que se requiere para poder observar los resultados de un ensayo completo, especialmente cuando se involucra al rendimiento de la especie. Empero, su mención va en el sentido del tipo de distribución que utiliza para el cruce de tratamientos en cada unidad experimental. Comúnmente estos experimentos hacen uso de los cuadros latinos para formar las diferentes estructuras experimentales, dando el carácter de fuentes de variación secundaria a los periodos (fases de la investigación) y a las unidades experimentales que se utilizan en las pruebas como filas y columnas, respectivamente.

Un aspecto importante de resaltar en cuanto a la razón de ser de los ensayos cruzados, es que éstos no confunden los efectos de los periodos con los efectos de los tratamientos. Así, por ejemplo, si una unidad experimental recibe la secuencia de tratamientos $A \rightarrow B$, habrá otra que reciba la secuencia $B \rightarrow A$ dentro de un mismo periodo, de tal forma que la comparación de tratamientos es independiente de los periodos. La estructura de un experimento cruzado que se describe en la figura 6.6 es un ejemplo clásico sobre estos tipos de distribuciones, en las cuales se toma de base las distribuciones en cuadro latino; asumiendo como filas a los periodos y como columnas a las unidades experimentales, mientras que los tratamientos están representados por las letras latinas. En esta figura son notorias las dos réplicas del DCL; además de observarse que todos los tratamientos son aplicados en cada periodo con la misma frecuencia, con el fin de evitar la confusión del efecto de los periodos y mantener un balanceo de las secuencias que facilitan estimar posibles *perturbaciones por acarreo*⁴² desde un tratamiento al siguiente en la aplicación secuencial. Debido a las perturbaciones por acarreo es deducible que los experimentos cruzados son adecuados cuando los efectos de tratamientos son reversibles y de corta duración.

Los análisis de un experimento cruzado, cuando se considera la no existencia de perturbaciones por acarreo, se formulan con los grados de libertad y sumas de cuadrados para distribuciones en cuadro latino con réplicas que tienen diferentes filas y columnas (caso C), indicado en el tema anterior y formulados en las tablas 6.36 y 6.37. En función del propósito del texto, se ha visto no pertinente ampliar más esta temática, no sin antes indicarle al lector que, para un mayor detalle al respecto, puede revisar el texto de Jones y Kenward (1989) mencionado en la bibliografía.

⁴² Una de las desventajas importantes en los experimentos cruzados es la posibilidad latente de que el efecto de un tratamiento, valorado en un periodo dado, se traslapa con el efecto del tratamiento inmediatamente posterior que se aplica en el siguiente periodo dentro de la misma unidad experimental. Estos efectos residuales son conocidos como *perturbaciones de acarreo* y su potencial influencia debe ser considerada en el análisis de varianza. La no consideración de un efecto de acarreo implica que se tenga en cuenta un intervalo de tiempo que, en cierta medida, permita la "eliminación total" del efecto residual del tratamiento aplicado antes de proceder con el siguiente tratamiento (siguiente periodo).

Período	Secuencia					
	1	2	3	4	5	6
I	A	B	C	A	B	C
	↓	↓	↓	↓	↓	↓
II	B	C	A	C	A	B
	↓	↓	↓	↓	↓	↓
III	C	A	B	B	C	A

Figura 6.6. Estructura de un diseño cruzado de tres tratamientos aplicados bajo una distribución de cuadro latino (3 x 3) con dos réplicas.

Distribución en cuadro latino modificado

Apegado al objetivo de este texto de acercar conocimiento exegético y metodologías que ofrezcan precisión y ventajas económicas a los experimentos agrícolas, se detalla a continuación las características y aplicación de la distribución en *cuadro latino modificado* (DCLM). La distribución en cuadro latino modificado es una alternativa para llevar a cabo experimentos que contienen un número elevado de tratamientos y que, a diferencia de la distribución en cuadro latino normal (DCL), es una opción que permite la reducción de costos, ya que el número de unidades experimentales que se manejan en esta distribución es relativamente menor. Además, esta modalidad conserva la capacidad de la distribución experimental en cuanto a la precisión que ofrece una DCL clásica cuando se trata de controlar dos fuentes adicionales de ruido, siempre y cuando la distribución sea adecuadamente estructurada e implantada. En la práctica, esta distribución es una especie de mezcla entre la distribución en cuadro latino y la de bloques completos al azar (DBCA).

La idea de la distribución consiste en que los tratamientos se los agrupe en un número equivalente al número de repeticiones que se hayan previsto utilizar, por ello los tratamientos deben ser múltiplos de las repeticiones. El total de los grupos se aleatorizan en los dos sentidos, de tal forma que los mismos no se repitan en las filas ni en las columnas. Dentro de cada grupo los tratamientos también deben ser aleatorizados.

Esta distribución tiene una ventaja sobre el DBCA. Si un experimento tiene un alto número de repeticiones y se bloquea en una sola dirección utilizando para ello un DBCA, las repeticiones o bloques se extenderán de manera importante en el suelo, lo que causará que la varianza del error experimental se incremente debido a que aumentará la heterogeneidad en el suelo. Por otro lado, si en este caso se prevé utilizar una distribución como el DCLM, puede lograrse un mayor control debido a que la variabilidad a lo ancho de un bloque es separada del error experimental en la fuente que para el modelo de análisis de varianza de esta distribución se conoce como "Filas".

Ubicación de los tratamientos en una distribución en cuadro latino modificado

Para explicar la asignación de los tratamientos en una DCLM consideremos un experimento que tiene 12 tratamientos y que se prevea correrlo con tres repeticiones. Tomando en cuenta

que en esta distribución el número de tratamientos debe ser múltiplo del número de repeticiones, para este ejemplo se deberán estructurar tres grupos de tratamientos. Esto es:

Grupo 1				Grupo 2				Grupo 3			
G ₁				G ₂				G ₃			
T ₁	T ₂	T ₃	T ₄	T ₅	T ₆	T ₇	T ₈	T ₉	T ₁₀	T ₁₁	T ₁₂

La aleatorización de los grupos puede iniciarse desde una permutación horizontal, es decir:

G₁ G₂ G₃
 G₃ G₁ G₂
 G₂ G₃ G₁

Haciendo la aleatorización de filas, se tendrá:

Filas				Filas			
1	G ₁	G ₂	G ₃	2	G ₃	G ₁	G ₂
2	G ₃	G ₁	G ₂	3	G ₂	G ₃	G ₁
3	G ₂	G ₃	G ₁	1	G ₁	G ₂	G ₃

Figura 6.7. Aleatorización de los grupos por fila en el diseño DCLM.

La aleatorización de las columnas es la forma como se distribuirán los grupos en el campo. Esto es lo que se indica en la figura 6.8.

Filas	1	2	3	Filas	2	3	1
2	G ₃	G ₁	G ₂	2	G ₁	G ₂	G ₃
3	G ₂	G ₃	G ₁	3	G ₃	G ₁	G ₂
1	G ₁	G ₂	G ₃	1	G ₂	G ₃	G ₁

Dist. para campo

Figura 6.8. Aleatorización de los grupos por columna en el diseño DCLM.

Al final, luego que las columnas se han aleatorizado, se ubican los tratamientos de acuerdo a los grupos establecidos, pero teniendo en cuenta que los tratamientos de cada grupo deben ser sorteados también. Es lo que puede observarse en la figura 6.9.

Filas (Repeticiones)	Columnas											
	1				2				3			
1	Grupo 1				Grupo 2				Grupo 3			
	G ₁				G ₂				G ₃			
	T ₃	T ₄	T ₁	T ₂	T ₇	T ₈	T ₅	T ₆	T ₁₀	T ₉	T ₁₂	T ₁₁
2	Grupo 3				Grupo 1				Grupo 2			
	G ₃				G ₁				G ₂			
	T ₁₁	T ₁₂	T ₁₀	T ₉	T ₁	T ₂	T ₄	T ₃	T ₈	T ₅	T ₇	T ₆
3	Grupo 2				Grupo 3				Grupo 1			
	G ₂				G ₃				G ₁			
	T ₅	T ₆	T ₈	T ₇	T ₁₂	T ₁₁	T ₉	T ₁₀	T ₂	T ₄	T ₃	T ₁

Figura 6.9. Croquis de ubicación de 12 tratamientos en una DCLM.

Ejemplo de aplicación 6.6⁴³

Se ha evaluado la adaptación de 18 variedades de soya bajo la distribución de un DCLM con tres repeticiones. La variable medida fue el peso de 100 semillas (g). La información del experimento es la siguiente:

Tabla 6.39. Tratamientos de un experimento bajo una DCLM.

Tratamientos			
Nº	Variedades	Nº	Variedades
1	10108	10	10003
2	RVBV-08-14(V14)	11	10564
3	JOSEFINA-2	12	J-865
4	10734	13	IJ-112-111
5	10010	14	10489
6	10032	15	J-854
7	10781	16	IJ-112-97
8	PADRE	17	INIAP-307
9	IJ-112-62	18	INIAP-308

La ubicación de cada unidad experimental en el campo se detalla en la figura 6.10. Nótese la estructura combinada de una DBCA con una DCL. También obsérvese que la estructura del cuadrado latino es la de un 3 x 3.

⁴³ Ejemplo tomado del texto “Diseños y análisis estadísticos para experimentos agrícolas” de este autor.

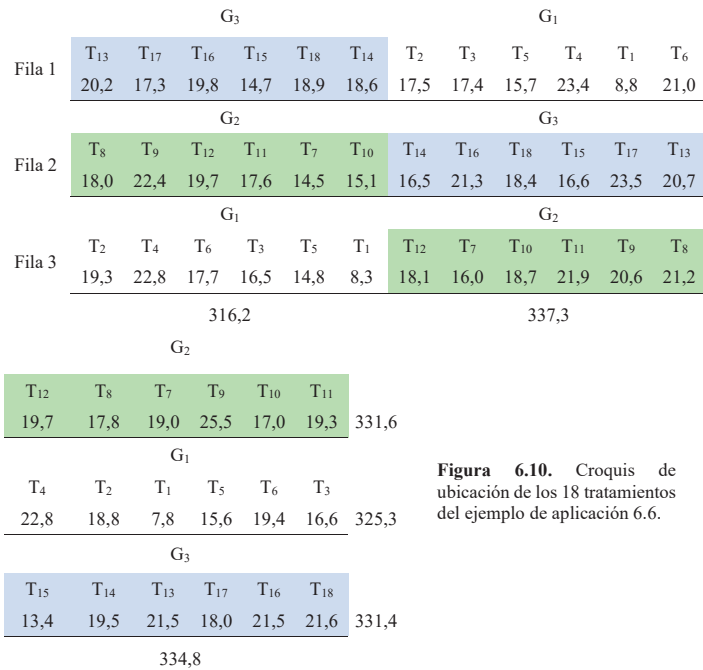


Figura 6.10. Croquis de ubicación de los 18 tratamientos del ejemplo de aplicación 6.6.

Tabla 6.40. Datos de peso de 100 semillas (g) de un experimento bajo una DCLM.

Grupos	Nº	Tratamientos (variedades)	Filas*			Total tratamientos	Promedios
			1	2	3		
G ₁	1	10108	8,8	7,8	8,3	24,9	8,3
	2	RVBV-08-14(V14)	17,5	18,8	19,3	55,6	18,5
	3	JOSEFINA-2	17,4	16,6	16,5	50,5	16,8
	4	10734	23,4	22,8	22,8	69,0	23,0
	5	10010	15,7	15,6	14,8	46,1	15,4
	6	10032	21,0	19,4	17,7	58,1	19,4
G ₂	7	10781	19,0	14,5	16,0	49,5	16,5
	8	PADRE	17,8	18,0	21,2	57,0	19,0
	9	IJ-112-62	25,5	22,4	20,6	68,5	22,8
	10	10003	17,0	15,1	18,7	50,8	16,9
	11	10564	19,3	17,6	21,9	58,8	19,6
	12	J-865	19,7	19,7	18,1	57,5	19,2
	13	IJ-112-111	20,2	20,7	21,5	62,4	20,8

	14	10489	18,6	16,5	19,5	54,6	18,2
	15	J-854	14,7	16,6	13,4	44,7	14,9
G ₃	16	IJ-112-97	19,8	21,3	21,5	62,6	20,9
	17	INIAP-307	17,3	23,5	18,0	58,8	19,6
	18	INIAP-308	18,9	18,4	21,6	58,9	19,6
Total de repeticiones			331,6	325,3	331,4	988,3	

*En este experimento, las filas son las repeticiones de cada tratamiento.

El modelo de análisis de varianza para este ensayo es el mismo que se utiliza en una DCL clásica, definido como:

$$Y_{ijk} = \mu + \tau_i + \beta_j + \gamma_k + \varepsilon_{ijk}$$

En el cual μ es el efecto constante, τ_i es el efecto de tratamientos, β_j es el efecto de filas, γ_k es el efecto de columnas y ε_{ijk} es el efecto del error experimental. La diferencia con respecto al análisis clásico de una DCL es que, en este caso, los grupos definen las filas y columnas; por ello será importante, al momento de establecer el número de tratamientos que integrarán cada grupo, que estos últimos queden ubicados en lotes más o menos homogéneos.

La suma total de cuadrados se define como:

$$SC_T = SC_t + SC_F(\text{suma grupos/filas}) + SC_C(\text{suma grupos/columnas}) + SC_E$$

$$SC_T = \sum Y_{ijk}^2 - \frac{\sum Y_{ijk}^2}{n}$$

$$SC_T = (8,8^2 + 17,5^2 + \dots + 21,6^2) - \frac{988,3^2}{54} = 669,53$$

$$SC_t = \frac{\sum (\sum \tau_i)^2}{r} - \frac{\sum Y_{ijk}^2}{n}$$

$$SC_t = \frac{24,9^2 + 55,6^2 + \dots + 58,9^2}{3} - \frac{988,3^2}{54} = 571,51$$

$$SC_F = \frac{\sum (\sum F_j)^2}{t} - \frac{\sum Y_{ijk}^2}{n}$$

$$SC_F = \frac{331,6^2 + 325,3^2 + 331,4^2}{18} - \frac{988,3^2}{54} = 1,42$$

$$SC_C = \frac{\sum (\sum C_k)^2}{t} - \frac{\sum Y_{ijk}^2}{n}$$

$$SC_C = \frac{316,2^2 + 337,3^2 + 334,8^2}{18} - \frac{988,3^2}{54} = 14,77$$

Tabla 6.41. Análisis de varianza para el ejemplo de aplicación 6.6.

Fuentes de variación	GL	SC	CM	F	P-valor
Total	53	669,53			
Tratamientos	17	571,51	33,62	13,15	0,000
Filas (por grupos)	2	1,42	0,71	0,28	0,759
Columnas (por grupos)	2	14,77	7,38	2,89	0,070
Error experimental	32	81,83	2,56		

De acuerdo al análisis de varianza detallado en la tabla 6.41, los tratamientos (las variedades) presentan diferencias significativas ($p < 0,05$) entre sus efectos que habrá que analizar con alguna prueba de comparación de medias. Asimismo, si bien entre filas el suelo es homogéneo, entre las columnas hay cierta variabilidad que, aunque no es significativa al 5% de probabilidad ($p = 0,07$), deberá tomarse en cuenta para posteriores experimentos sobre el mismo terreno.

CAPÍTULO VII

RECOMENDACIONES EN LOS EXPERIMENTOS AGRÍCOLAS

A veces hay que preguntarse si realmente es más beneficioso utilizar una distribución de bloques al azar, la que se usa sin reparo, antes que una distribución completamente al azar; a sabiendas de que esta última es estadísticamente más precisa...



Algunas consideraciones sobre el diseño de los experimentos agronómicos

Si bien el propósito de esta obra ha sido el de describir algunas alternativas metodológicas y estadísticas que deben tenerse en cuenta para hacer que un experimento sea lo más preciso posible, no obstante, existen también algunas consideraciones que se indican a continuación que todo investigador deberá tener presente durante la planeación, desarrollo y recolección de los datos del ensayo, con miras a favorecer dicha precisión.

Sobre el tipo de distribución experimental

Los experimentos agrícolas que se realizan a nivel campo, tradicionalmente antes que científicamente, en la gran mayoría hacen uso de las distribuciones de bloques completos al azar (DBCA). Es una situación que ocurre aun a pesar de que el lote de terreno seleccionado muestre condiciones evidentemente homogéneas y no presente factores de variación secundaria que induzcan a utilizar una distribución como la DBCA con el fin de bloquearlos. Esta decisión que se ha perennizado en la concepción de muchos biometristas, incluso en instituciones de investigación o en sitios en donde se conoce de antemano que la fuente de bloqueo en los modelos de análisis de varianza resulta con efecto no significativo, se sigue implementando la DBCA en experimentos subsecuentes dentro de los mismos espacios en donde se han realizado experimentos anteriores.

Todo investigador debe conocer que entre la distribución completamente al azar (DCA), la de bloques completos al azar y la de cuadro latino (DCL), las cuales son las estructuras básicas que se utilizan en prácticamente todos los experimentos agrícolas⁴⁴, la primera es la que estadísticamente tiene la mayor precisión al generar un ANDEVA de mayor potencia de entre las que se realizarían en las otras dos distribuciones. De todas estas distribuciones, la DCA⁴⁵ goza de la gran ventaja de contener la mayor cantidad de grados de libertad en el error experimental (siempre que sea la misma cantidad de tratamientos respecto de las otras distribuciones), lo que consecuentemente hace que el análisis de varianza tenga relativamente una mayor precisión para detectar efectos de los tratamientos al minimizar la varianza dentro de ellos, disminuyendo la probabilidad de caer en el error tipo II.

La precisión indicada en el párrafo anterior es la que se demuestra en la tabla 7.1 en donde se incluyen los análisis de varianzas realizados con unos mismos datos. En la parte superior de esta tabla se detalla el ANDEVA de una DCA y en la parte inferior el correspondiente al de una DBCA. En el primer caso, la varianza o cuadrado medio del error es 20,70, mientras que al considerar una distribución de bloques completos al azar este valor aumenta a 22,96. Consecuentemente, esta disminución de los grados de libertad del error experimental en la DBCA es lo que hace que las varianzas dentro de los grupos hayan aumentado, haciendo también que el valor (cuantil) de la probabilidad α (P – valor observado) aumente de 0,2044 en la DCA a 0,2562 en la DBCA.

⁴⁴ Esta afirmación se hace tomando en cuenta las distribuciones de base que se utilizan en los experimentos. Esto no debe confundirse con los arreglos o diseño de tratamientos que son los que utilizan estas distribuciones experimentales de base.

⁴⁵ Un aspecto práctico que debe tener presente el investigador es la reversibilidad en los análisis que puede aplicarse en los experimentos que inicialmente utilizan distribuciones con bloqueo, pero que en los modelos de ANDEVA dichas fuentes bloqueadas (fuentes de variación secundaria) no presentan efectos significativos. En estos casos, sea que se haya utilizado una distribución como la DBCA o la DCL y sus modelos de ANDEVA correspondientes no detectan significancia en las fuentes bloqueadas, el biometrista puede re-hacer los cálculos y analizar el experimento como si se tratase de una DCA. Con ello puede lograr un incremento en la precisión estadística al tener una mayor cantidad de grados de libertad en el error experimental.

Lo que se ha mencionado es con el propósito de que, al momento de seleccionar una distribución experimental, el investigador deberá ser muy observador del espacio de suelo en donde desarrollará su experiencia. Lotes de terrenos sin pendientes, que no tengan canales, esteros o ríos muy próximos al área experimental y/o bosques o cultivos intensivos que limiten con el experimento, a primera vista no presentan ningún efecto secundario que se haga necesario bloquearlo; de allí que la distribución completamente al azar podría ser la más indicada. Sin embargo, la decisión de adoptar una distribución experimental *a priori* también involucra otras consideraciones de importancia como el tipo de cultivo que se utilizará en la prueba. Experimentos que se realizan con cultivos de distancias de siembra amplias (más de 6 m) no se desarrollan en parcelas pequeñas como los que se llevan a cabo con cultivos como el arroz, el maíz o el trigo. El hecho de requerir mayores áreas, hace que se complique cada vez más el control sobre la heterogeneidad del suelo; por lo tanto, en estas circunstancias utilizar una DCA podría ser la opción menos precisa.

Tabla 7.1. Magnitud de los grados de libertad del error experimental entre los análisis de varianza de una DCA y una DBCA.

Distribución completamente al azar

Variable	N	R ²	R ² Aj	CV
Datos	25	0,25	0,10	15,01

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo.	135,44	4	33,86	1,64	0,2044
TRATAMIENTOS	135,44	4	33,86	1,64	0,2044
Error	414,00	20	20,70		
Total	549,44	24			

Distribución de bloques completos al azar

Variable	N	R ²	R ² Aj	CV
Datos	25	0,33	0,00	15,80

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo.	182,08	8	22,76	0,99	0,4781
TRATAMIENTOS	135,44	4	33,86	1,47	0,2562
REPETICIONES	46,64	4	11,66	0,51	0,7308
Error	367,36	16	22,96		
Total	549,44	24			

Para identificar no sólo visualmente alguna fuente de variabilidad secundaria en un experimento, sino también cuantitativamente y tener una mayor precisión en los resultados al seleccionar la distribución más adecuada, los estudios de heterogeneidad de suelo indicados en el capítulo 2 son fundamentales. Cuando existen patrones de comportamiento en el suelo, como por ejemplo el de fertilidad, éste puede identificarse al desarrollar los *experimentos de uniformidad* (llamados también ensayos en blanco), los cuales permitirán decidir si es necesario utilizar una DBCA, de tal forma que se pueda neutralizar la acción de dicho efecto. Lo que se debe tener bien definido es que, aparte del manejo de un ensayo, la precisión se ve afectada por el área de la unidad experimental y el número de repeticiones que tendrá cada tratamiento; a este respecto existen varias metodologías que pueden aplicarse, en donde los experimentos de uniformidad son esenciales. Sin embargo, debido al costo y tiempo que se

debe invertir en estas investigaciones previas, son alternativas que ni siquiera en los institutos de investigación agrícola se hace énfasis de su uso.

La orientación de los bloques, cuando se utiliza una distribución como la DBCA, es otro aspecto muy importante para la búsqueda de precisión en los ensayos. Muy aparte de observar la dirección del viento o de la radiación solar, el factor determinante de cómo deberán ubicarse los bloques en el campo es la heterogeneidad del suelo. Para que éstos cumplan su función, siempre deberán colocarse en forma perpendicular a alguna dirección de variabilidad que muestre el suelo.

Otro aspecto importante para tener experimentos eficaces es la observancia de la conveniencia de ensayos con más de un factor de estudio antes que los que se realizan con solo un factor a la vez. Los primeros tienen muchas ventajas, tanto en precisión como en el aspecto económico, de lo cual se trata en el siguiente tema.

Conveniencia de un experimento factorial sobre un ensayo de un solo factor por turno

Desde el punto de vista del diseño de tratamientos se define el tipo de experimento; por ello, la clasificación más general es la de unifactorial (un factor a la vez) cuando se evalúa un único factor, o factorial cuando se evalúan dos o más factores. En este sentido, hay que tener en cuenta la ventaja indiscutible que se tiene cuando se desarrolla un experimento factorial antes que un unifactorial. Cuando se tienen dos o más factores identificados para un estudio, el peor error que puede cometer un investigador es desarrollar experimentos subsiguientes moviendo un solo factor a la vez. El uso de experimentos factoriales tiene su primacía sobre el de un solo factor en todo aspecto, tanto en el de precisión como en el económico; además, en estos experimentos se evalúa un efecto muy importante que jamás se podrá evaluar en los unifactoriales subsiguientes como lo es la interacción de los factores, la cual ocurre con mayor frecuencia en los ensayos biológicos como es el caso de los agronómicos. Para constatar la conveniencia de un experimento factorial sobre los que evalúan un solo factor a la vez, veamos lo que se explica sobre el ejemplo que se describe en el siguiente párrafo.

Se ha propuesto la evaluación experimental de tres dosis de nitrato de potasio (NO₃K) en condiciones de campo aplicadas a un cultivo determinado, utilizando para ello una distribución de bloques completos al azar con siete repeticiones para cada tratamiento. Luego de que se haya establecido la dosis de NO₃K que ofrezca la mayor producción, se planea evaluar cuatro dosis de carbonato de calcio (CO₃Ca) por problemas de acidez en el suelo, utilizando también una distribución de bloques con cinco repeticiones por tratamiento. En el caso de evaluar estos dos factores simultáneamente como un experimento factorial, tendríamos una DBCA con arreglo factorial 3 x 4, utilizando un mínimo de tres repeticiones por cada combinación. Los esquemas de ANDEVA correspondientes son los detallados en las tablas 7.2, 7.3 y 7.4.

- El esquema de ANDEVA del primer experimento para el primer factor identificado como nitrato de potasio se indica en la tabla 7.2.

Tabla 7.2. Esquema del ANDEVA para el análisis de las tres dosis de NO₃K.

Fuentes de variación	Grados de libertad
Total	20
Tratamientos (Dosis NO ₃ K)	2
Repeticiones	6
Error experimental	12

- El esquema de ANDEVA del segundo experimento para el segundo factor identificado como carbonato de calcio es el que se indica en la tabla 7.3. Cabe indicar que este segundo experimento se realiza una vez que el primero concluye y se conoce cuál es la mejor dosis de NO_3K , la misma que es aplicada con este segundo ensayo de forma general.

Tabla 7.3. Esquema del ANDEVA para el análisis de las dosis de CO_3Ca .

Fuentes de variación	Grados de libertad
Total	19
Tratamientos (Dosis de CO_3Ca)	3
Repeticiones	4
Error experimental	12

- El esquema de ANDEVA para un único experimento en donde se evalúan los dos factores (dosis de NO_3K y CO_3Ca) en forma simultánea es el que se describe en la tabla 7.4.

Tabla 7.4. Esquema del ANDEVA para el análisis del factorial $A \times B$.

Fuentes de variación	Grados de libertad
Total	35
Factor A (NO_3K)	2
Factor B (CO_3Ca)	3
Interacción AB	6
Repeticiones	2
Error experimental	22

Evaluando los dos factores en forma independiente, primero al factor NO_3K y luego el factor CO_3Ca , se necesitan 41 unidades experimentales (21 para el primer experimento y 20 para el segundo); mientras que haciéndolo simultáneamente mediante un factorial, se necesitarían 36 unidades experimentales. Es decir, en menos parcelas podemos estudiar estos dos factores a la vez, ahorrando espacio, costos y tiempo; logrando así una mayor eficiencia económica en el experimento factorial antes que en el estudio independiente de cada factor. Asimismo, y por razones obvias, un experimento factorial permite reducir el tiempo de estudio, lo que consecuentemente trae consigo un gran ahorro de recursos y la oportunidad de tener resultados a la mitad del tiempo que al hacer dos experimentos unifactoriales subsecuentes, de un factor a la vez.

En cuanto a la precisión estadística, el experimento factorial es mucho más preciso respecto de los experimentos unifactoriales. Esto es evidente al observar los grados de libertad del error experimental que ofrece el ANDEVA de la tabla 7.4 con respecto a las tablas 7.2 y 7.3, en donde se consideran a los experimentos en forma independiente. Para el esquema del ensayo factorial tenemos 22 grados de libertad, mientras que, en los esquemas independientes, sólo se tienen 12 grados de libertad en cada uno de los experimentos. Por lo tanto, puede decirse que el factorial de este ejemplo es aproximadamente dos veces más preciso que los ensayos de un solo factor.

Es imperativo también indicar que al realizar un experimento para determinar primero la dosis de NO_3K que resulte con la mayor producción, y luego esta dosis probarla para las cuatro proporciones de CO_2/Ca , se eliminan combinaciones que podrían resultar mucho más rentables que el uso de la dosis que resultó con los mayores rendidos. Esto puede hacerse notorio debido a que únicamente el experimento factorial (con todas sus combinaciones) permitirá determinar si existe alguna interacción entre los factores de estudio que propicien un mayor rendimiento; situación que es de mucha importancia cuando se evalúan productos agrícolas que pueden actuar con sinergia, o incluso en forma antagonica, en donde sólo un experimento factorial podrá verificar este efecto. Además, dada la dependencia múltiple de las plantas hacia diversos elementos, los efectos de interacción por lo general son más comunes en los experimentos biológicos.

Aunque la optimización como tal no es tan aplicable en el ámbito agrícola, la ubicación de rangos de respuestas máximas puede tener utilidad práctica en los experimentos de esta área; sin embargo, evaluando un factor a la vez se corre el riesgo de seleccionar niveles factoriales que se encuentran muy lejos de la región experimental dentro de la cual se ubican los mejores efectos. En este sentido, las combinaciones factoriales, cuyos niveles hayan sido adecuadamente seleccionados, permiten identificar espacios experimentales en donde se encuentran las respuestas máximas o mínimas, según sea el propósito del experimento.

Otras ventajas que ofrecen los experimentos factoriales se refieren al uso de un menor número de repeticiones, al aumentar la cantidad de tratamientos. Permiten lograr un análisis más compacto, debido a que se estudian los efectos de un factor con la influencia del efecto de otro (u otros) factor (s), lo que resulta en conclusiones más valederas y eficaces. Además, en la realización de estos experimentos se aprovecha las condiciones genotípicas de las semillas y ambientes climáticos más homogéneamente que al realizar experimentos subsecuentes, en donde, desde un experimento a otro pueden cambiar estas condiciones, lo que al final se traduce en un aumento del error experimental.

La bondad de los experimentos factoriales, en su momento, fue expresada por el propio precursor y referente del diseño de experimentos, el insigne Ronald Fisher (1890 – 1962). Este científico, a través de su obra, *Contribution to mathematical statistics*, mencionó que: “las experiencias amplias y complejas son mucho más eficaces que las experiencias sencillas. Ningún aforismo es más frecuentemente repetido, en relación con la experimentación agrícola, que aquel que dice que a la naturaleza debemos plantearle pocas cuestiones simultáneamente, e incluso en rigor una sola. El autor está convencido de que se trata de un completo error. La naturaleza responderá de buena gana a un cuestionario lógico cuidadosamente preparado, e incluso, si le planteamos una sola cuestión, rehusará el responder antes de que otra esté resuelta”.

Sobre los tratamientos

Los tratamientos constituyen la variable independiente de un experimento. Pueden ser directamente los niveles de un factor cuando el experimento es unifactorial o una combinación de niveles cuando el experimento es un factorial (más de un factor). La selección de éstos se fundamenta en el método deductivo, ya que su definición no puede realizarse a partir de ideas vagas sino con la ayuda de información bibliográfica (basada en hechos empíricos) pertinente que haga posible escoger niveles que no estén fuera de contexto. Esto es, para definir los tratamientos, el investigador deberá basarse en el mayor marco teórico posible, de tal forma que lo guíe en esta selección. Además, la ayuda de técnicos relacionados al tema de la investigación en esta selección, con vasta experiencia, siempre será un gran aporte para lograr un experimento exitoso. Si no existiera esta información de respaldo respecto de la definición

de los tratamientos de estudio, la única alternativa que le puede ayudar al investigador es realizar experimentos preliminares; de tal manera que pueda identificar, más o menos, el espacio experimental sobre el cual los efectos posiblemente sean los más sobresalientes y, a partir de aquello, establecer los tratamientos que evaluará en el experimento definitivo. Casos como este último, se enmarcan dentro de los llamados *experimentos exploratorios*.

Los tratamientos pueden ser nominales como, por ejemplo, varias técnicas de podar; o también cuantitativos, como el caso de la evaluación de varias dosis de un fertilizante. Cuando el experimento contiene tratamientos cuantitativos, los niveles que se establecen para el estudio deben tener la suficiente amplitud para poder detectar efectos entre ellos; ya que, cuando éstos son muy reducidos, se complica descubrir algún efecto de alguno de los niveles de prueba. Otra recomendación que debe tenerse en cuenta cuando los tratamientos son cuantitativos y se quiere investigar más allá de encontrar diferencias estadísticas, es utilizar al menos tres niveles (de preferencia, equidistantemente espaciados), que permita deducir si el factor de estudio tiene una incidencia de curvatura, que a la vez haga posible establecer la existencia de un efecto máximo a partir del cual un tratamiento deja de ser productivo, o a su vez un rango en donde dichas respuestas máximas se presenten. Debido a la bondad de poder modelar con los tratamientos cuantitativos a partir de ecuaciones de regresión, dichos tratamientos son muy utilizados en la metodología de superficie de respuesta, a partir de la cual se establecen niveles experimentales óptimos; con la salvedad de que esta última debe aplicarse adecuadamente, debido a que los ensayos agronómicos son los que más expuestos están a la inestabilidad espacial y temporal del ambiente que los rodea.

Otro aspecto que hay que tener muy en cuenta es el número de tratamientos, especialmente aquellos que provienen de mezclas factoriales que generan una importante cantidad de combinaciones. En este sentido, no se puede dejar de lado el factor heterogéneo del suelo, el cual crece conforme aumentamos el área experimental con un gran número de tratamientos que pueden hacer que se confundan los efectos, además de lo costoso que resulta mantener un experimento con muchas unidades experimentales. En esta situación el investigador debe ser muy observador del área en donde llevará a cabo su ensayo, pues necesita definir la distribución que utilizará para contrarrestar el requerimiento de un área relativamente grande; de allí que existan alternativas como los experimentos en látice, los de bloques incompletos fraccionados o desbalanceados, en bloques con efecto confundido o los factoriales fraccionados, los cuales permiten un mejor control de la heterogeneidad del suelo.

Otra observación que debe tenerse presente es la forma de aplicación de los tratamientos. Se requieren unidades experimentales relativamente grandes cuando los tratamientos son de aplicación foliar, lo que es diferente cuando los tratamientos son de aplicación localizada como el caso de los fertilizantes edáficos. Asimismo, en ciertos cultivos como en el arroz, las unidades experimentales deben estar protegidas con muros para que los tratamientos de aplicación al suelo no se traslapen entre ellos, especialmente los que se aplican al suelo. En estos casos, el uso de bordes suficientemente amplios que definan adecuadamente la unidad de muestreo también ayudan a controlar los efectos de traslape de los tratamientos.

En ciertas investigaciones no se puede dejar de lado a los tratamientos testigos. Dentro de los ensayos, estas parcelas referenciales (que en todo experimento también se definen como tratamientos) pueden ser de dos clases: aquellos que se conocen como testigos convencionales y aquellos que se definen como testigos absolutos. En algunos casos se requiere solamente el testigo convencional como, por ejemplo, en los experimentos en donde se evalúan distintas clases de fertilizantes y el testigo convencional hace referencia a un fertilizante de uso común en la zona de estudio. En otras circunstancias se requerirá de ambos testigos; así, por ejemplo, cuando se evalúan plaguicidas de formulación orgánica, el testigo convencional puede ser un plaguicida químico, mientras que el testigo absoluto corresponderá a aquella unidad experimental en donde no se realice la aplicación de ningún plaguicida. Los tratamientos

testigos, ya sean convencionales o absolutos, son los que permiten hacer las comparaciones de los tratamientos de prueba en un ensayo en términos económicos; de allí que, en la mayoría de las circunstancias, sólo los testigos permitirán tomar la decisión más acertada sobre la aplicación de un tratamiento o seguir sin ningún cambio.

Sobre las variables de respuesta

Las variables de respuesta constituyen (holísticamente), ya sea directa o indirectamente, los objetivos específicos de un experimento. Estas variables deben estar directamente relacionadas con lo que se quiere probar en un ensayo; no obstante, existen casos en donde el efecto de los tratamientos se complica medirlo directamente y se opta por evaluar variables que valoran los efectos en forma indirecta. Un ejemplo puede ser un estudio de diferentes cebos envenenados para disminuir la población de ratas que afectan a la caña de azúcar. En este ejemplo, medir el número de ratas muertas por cada uno de los cebos como variable de respuesta resultaría muy complicado por la gran movilidad de la plaga y el efecto gradual de los productos; sin embargo, puede optarse por variables indirectas como el daño en el cultivo y la cantidad de cebo consumido en 24 horas.

Una falta que se comete comúnmente en los ensayos agrícolas es la selección de múltiples variables de respuesta, que muchas veces ni siquiera tienen relación con los tratamientos. Seleccionar un montón de variables no hace mejor a un experimento que aquellos en los que apenas se tienen unas pocas variables de respuesta, siempre que estén directa y/o indirectamente relacionadas al factor (s) de estudio, a las hipótesis y a los objetivos planteados. No obstante, la medición de más de una variable en un experimento, que estén vinculadas al efecto de los tratamientos, es una forma concomitante de confirmar un resultado cuando una de las variables declara efectos no tan convenientes respecto de uno o más tratamientos.

Las labores culturales, tales como el riego, control de malezas, control de plagas y otras que se realicen en un experimento pueden llevarse a cabo en bloques cuando el ensayo es extenso y no existen los recursos ni la mano de obra disponible. La medición de variables no puede realizarse en partes dentro de un experimento, de preferencia ni si quiera por bloques. Esta medición debe ejecutarse en un mismo instante y no puede dejarse para sucesivos días, pues ésta podría ser una fuente de variación que voluntariamente el investigador estaría introduciendo en el experimento. El investigador debe prever toda esta situación al momento de planear el experimento, incluso deberá pensar en utilizar guardianes para el cuidado del experimento en época de cosecha, especialmente en aquellos ensayos en donde se experimenta con cultivos frutales como las sandías o los melones, cuyas pérdidas por el robo de los frutos pueden mandar al traste toda una investigación.

Otro aspecto que debe prestarse mucha atención en la medición de variables está relacionado con la parte humana de la investigación (imaginación, suspicacia y subjetivismo). En la obtención de los datos debe procederse con sumo cuidado, si existen valores salidos de un margen de variación normal, deberán en lo posible, volver a medirse para descartar algún error de medición. Asimismo, mientras se utilicen criterios y herramientas objetivas de medición de variables, las conclusiones tendrán una mayor validez científica. Si la medición insoslayablemente debe realizarse de forma cualitativa, es un menester que quién la realice tenga la experiencia y seriedad necesaria para evitar, en la medida de lo posible, resultados extremadamente subjetivos y sesgados.

Sobre las herramientas estadísticas

En un experimento bien planeado y desarrollado bajo el método científico, las pruebas o test estadísticos que se utilizan, en algunos casos sólo para reafirmar lo evidente, vienen a constituirse en las herramientas que, en alguna medida, dan fe de la objetividad⁴⁶ de los resultados. De hecho, mencionar la prueba estadística utilizada en una investigación experimental es parte fundamental de la metodología para poder publicar un artículo sobre un ensayo determinado en alguna revista científica. Sin embargo, dado el aporte abstracto y la aparente perogrullada que ofrecen las pruebas estadísticas en un ensayo bien planeado y manejado, con efectos claramente contundentes, las aplicaciones prácticas de los resultados experimentales están cargadas más de criterios subjetivos antes que la objetividad que pudieran ofrecer dichas pruebas.

La herramienta estadística que se utilice depende del tipo de distribución experimental y no al revés; esto trae como consecuencia que lo importante en la investigación experimental es el ensayo y no las pruebas estadísticas, ya que las conclusiones de un experimento realizado con rigor científico y con efectos contundentes, aun con el más elemental test estadístico, fácilmente pueden ser deducidas. A su vez también, ni el uso de sofisticadas pruebas estadísticas podrían remediar resultados imprecisos de experimentos mal diseñados y/o mal manejados.

Dentro de los componentes de la experimentación, tales como el diseño, el manejo del ensayo y las herramientas estadísticas, éstas últimas no deben tener supremacía en las decisiones que de un experimento se tomen. En este contexto, las diferencias definidas por un test estadístico como significativas pueden no representar efectos que permitan lograr condiciones mejoradas reales y aplicables en la práctica para el agricultor. Hay que tener claro que la importancia de los datos no está en las técnicas de análisis estadístico, sino en la incidencia de ellos sobre el conocimiento que se quiere obtener del objeto de estudio; de allí que un efecto real de los tratamientos no sólo puede ser declarado a partir de una de estas técnicas sin considerar un análisis agrobiológico y productivo del factor de estudio, esto es, los métodos estadísticos no deben sustituir el sentido común y pragmático. Además, en estudios en donde las diferencias entre tratamientos son notoriamente evidentes (con un reducido error experimental), un test estadístico viene a constituirse en un requisito para poder difundir lo obtenido en un experimento en las famosas revistas científicas, pero que es irrelevante en la toma de decisiones; de hecho, un tratamiento estadísticamente ganador puede relegarse al no ser también rentablemente el mejor.

Las pruebas clásicas de la estadística paramétrica en los ensayos son el test t de Student y el análisis de varianza (ANDEVA). El primero para comparar, ya sea un solo tratamiento contra una media poblacional o para comparar dos tratamientos; mientras que el segundo se utiliza para definir diferencias significativas entre más de dos tratamientos. Dado que ambas son pruebas paramétricas, los requerimientos de normalidad subyacente y de homocedasticidad residual los datos, dando por hecho que la independencia se cumple con la aleatorización, son importantes para que estas pruebas tengan la precisión necesaria.

Al igual que la prueba de t, el ANDEVA también puede utilizarse para detectar diferencias significativas entre dos tratamientos. Sin embargo, la comparación de más de dos tratamientos realizándola de dos en dos a través de la prueba t de Student no es lo más adecuado, ya que por este método, el grado de significancia α pre-establecido para el estudio se ve incrementado. Esto último significa que se estará estableciendo diferencias significativas a un valor de α más alto, con el consiguiente aumento del riesgo de caer en el error I. Así por ejemplo, aplicando el test t de Student para un experimento de tres tratamientos, si el valor de α establecido es de

⁴⁶ Muchos investigadores, especialmente los de formación en estadística pura, dicen que la objetividad de una investigación se define por las pruebas estadísticas. Sin embargo, esta "objetividad" solo reside en el análisis de los números; lo que suceda en un experimento en cuanto a su diseño, manejo, así como en la toma y tabulación de los datos, cuando el investigador es honesto, es lo que al final le confiere una base objetiva sólida y confiable.

0,05, en la realidad se estará rechazando la hipótesis nula a un valor de 0,0975. Este incremento de α en la prueba de t para el análisis de más de dos tratamientos se explica al considerar que la significancia real viene dada por los rangos (m) en que se amplían las comparaciones y puede obtenerse por:

$$Sig. = 1 - (1 - \alpha)^m$$

Al comparar los tratamientos 1 y 2 no existe inconveniente con el valor establecido de 0,05:

$$Sig. = 1 - (1 - 0,05)^1 = 0,05$$

Sin embargo, el problema con la prueba se presenta cuando entran más de un rango en la comparación, esto es, entre los tratamientos 1 y 3:

$$Sig. = 1 - (1 - 0,05)^2 = 0,0975$$

Al final, lo último quiere decir que en la comparación de los tratamientos 1 y 3, el nivel de significancia real con el cual se está realizando la comparación es al 9,75% y no del 5% como inicialmente se consideró.

El modelo de análisis de varianza bajo el criterio de efectos fijos⁴⁷ es considerado como la prueba estadística clásica de los experimentos que se plantean en función del diseño experimental. Es una herramienta que desglosa la variabilidad que presentan los datos en sus correspondientes fuentes, establecidas por el tipo de distribución utilizada, cuya identificación previa fue realizada por el investigador. Los distintos modelos estadísticos del ANDEVA consideran tanto a la distribución como al tipo de experimento que debe plantearse adecuadamente, con el fin de poder observar efectos tan importantes como la interacción de factores en los ensayos en donde se evalúan más de una variable de interés. Esta descomposición queda expuesta si partimos por la variabilidad total que presentan los datos que se han recolectado en un experimento bajo una distribución completamente al azar.

Cualquier dato bajo una DCA estará representado por la siguiente suma algebraica de efectos (modelo indicado en la tabla 7.5):

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad (7.1)$$

Es decir, cualquier dato tendrá un componente constante (μ), uno del efecto del tratamiento (τ_i) más el otro del efecto del error experimental (ε_{ij}). Visto desde otro punto, la variabilidad total de este modelo está definida como:

⁴⁷ La definición de los niveles de los factores de estudio de un experimento puede realizarse de acuerdo a la necesidad del investigador o en función de una selección al azar de los mismos. En el primer caso, los niveles los establece el investigador en función de lo que le interesa a éste para solucionar un problema, de allí que se diga que el experimento obedece a un modelo de *efectos fijos*, cuyo ANDEVA correspondiente deberá ajustarse a este modelo. Los experimentos que se realizan seleccionando al azar los niveles de un factor se denominan de *efectos aleatorios*, en los cuales ya no tiene sentido identificar la mejor media sino el tratamiento de menor variación; por ello de manera general, estos experimentos se llevan a cabo a nivel industrial para la calibración de equipos respecto de su precisión. La mayoría (por no decir todos) de los experimentos agrícolas se realizan bajo el criterio de efectos fijos, debido a que el ambiente en el cual se desarrollan es tan cambiante (varianza externa) que hace innecesaria la realización de experimentos bajo el criterio de efectos aleatorios. Las conclusiones de un experimento de efectos fijos sólo son válidas para los niveles seleccionados de un factor; mientras que, en el caso de los experimentos de efectos aleatorios, lo que resulte puede extrapolarse a la población del factor de estudio. El ANDEVA que se elabora cuando el experimento es unifactorial, pero de efectos aleatorios, es idéntico al que se realiza cuando los niveles del factor son fijos; no obstante, cuando se tiene más de un factor, la diferencia ocurre en el cálculo del estadístico F . También pueden existir experimentos de dos factores, uno de ellos de efectos fijos y el otro de efectos aleatorios, a los cuales se los conoce como de *modelo mixto* (fijo y aleatorio). Ejemplos en la agronomía de experimentos de efectos mixtos son los de efectos anidados y los de parcelas divididas.

$$Y_{ij} - \mu = \tau_i + \varepsilon_{ij} \quad (7.1a)$$

Al hacer la consideración de que $\mu = \bar{Y}$ (la media general de los datos de un experimento), es deducible que el efecto de un tratamiento i es igual a la diferencia entre el efecto medio de ese tratamiento (\bar{Y}_i) y la media general de los datos de todo el ensayo. Esto es:

$$\tau_i = \bar{Y}_i - \bar{Y}$$

Asimismo, el error experimental corresponderá a la variabilidad residual entre lo que expresa cualquier dato (Y_{ij}) y la media de cualquier tratamiento (\bar{Y}_i). Es decir:

$$\varepsilon_{ij} = Y_{ij} - \bar{Y}_i$$

Al realizar el remplazo respectivo en la expresión 6.1a, ésta última se replantea en la forma que se indica a continuación:

$$Y_{ij} - \bar{Y} = (\bar{Y}_i - \bar{Y}) + (Y_{ij} - \bar{Y}_i) \quad (7.1b)$$

Si se eleva al cuadrado la expresión 7.1b y se resuelve el miembro derecho de esta ecuación, se tendrá:

$$\begin{aligned} (Y_{ij} - \bar{Y})^2 &= [(\bar{Y}_i - \bar{Y}) + (Y_{ij} - \bar{Y}_i)]^2 \\ (Y_{ij} - \bar{Y})^2 &= (\bar{Y}_i - \bar{Y})^2 + 2(\bar{Y}_i - \bar{Y})(Y_{ij} - \bar{Y}_i) + (Y_{ij} - \bar{Y}_i)^2 \end{aligned}$$

Si se suman respectivamente todos estos desvíos, tendremos:

$$\sum (Y_{ij} - \bar{Y})^2 = \sum (\bar{Y}_i - \bar{Y})^2 + 2 \sum (\bar{Y}_i - \bar{Y})(Y_{ij} - \bar{Y}_i) + \sum (Y_{ij} - \bar{Y}_i)^2$$

Dado que la suma de los desvíos es igual a cero [$\sum (\bar{Y}_i - \bar{Y})(Y_{ij} - \bar{Y}_i) = 0$], la última expresión se reduce a la siguiente forma:

$$\sum (Y_{ij} - \bar{Y})^2 = \sum (\bar{Y}_i - \bar{Y})^2 + \sum (Y_{ij} - \bar{Y}_i)^2 \quad (7.1c)$$

La expresión 7.1c corresponde a la suma de cuadrados de la variabilidad del total de los datos (el término del primer miembro de la ecuación), la que a su vez está compuesta por la suma de cuadrados de la variabilidad entre los diferentes grupos y la suma de cuadrados de la variabilidad dentro de estos grupos; es decir, la variabilidad que hay entre tratamientos y la variabilidad del error experimental según el primero y segundo términos del segundo miembro de esta expresión, respectivamente.

Trasladando la expresión 7.1c a expresiones de varianzas bajo la forma abreviada, esta fórmula se modifica planteándose en la estructura rutinaria de cálculo que se desarrolla en el ANDEVA. Esto es:

$$\frac{\sum Y_{ij}^2 - \frac{(\sum Y_{ij})^2}{n}}{n-1} = \frac{\sum \bar{Y}_i^2 - \frac{(\sum Y_{ij})^2}{n}}{t-1} + \frac{\sum Y_{ij}^2 - \frac{\sum \bar{Y}_i^2}{r}}{n-t} \quad (7.1d)$$

En la fórmula 7.1d, el primer miembro representa a la varianza total de los datos; el primer término del segundo miembro es la varianza de los tratamientos (varianza entre grupos) y el segundo término de este último miembro es la varianza del error experimental (varianza dentro

de grupos). Nótese que las sumas de cuadrados de las desviaciones están corregidas por el término: $(\sum Y_{ij})^2/n$. En un experimento, considérese a $n = tr$.

En una distribución completamente al azar, el modelo de ANDEVA separa la variabilidad total de los datos en dos fuentes: la que corresponde a los tratamientos y la que representa al error experimental (según se expuso entre las expresiones 7.1b a 7.1d). Mientras que, en distribuciones como la de bloques completos al azar, la varianza total se descompone en tres fuentes: las mencionadas anteriormente y una adicional que representa a un factor secundario identificado como bloques o repeticiones. El modelo de ANDEVA para una distribución en cuadro latino genera cuatro fuentes, las dos mencionadas en la distribución completamente al azar, y dos representadas por el doble bloqueo que se analiza en esta distribución, conocidas generalmente como “filas” y “columnas”. Bajo esta apreciación, los modelos de análisis de varianza se plantean de acuerdo a la distribución experimental utilizada. En la tabla 7.5 se indica el número de fuentes de variación de las distribuciones básicas mencionadas, así como sus modelos de ANDEVA respectivos.

Tabla 7.5. Fuentes de variación y modelos lineales de las distribuciones experimentales básicas.

Distribución	Fuentes de variación en el modelo de ANDEVA	Modelo
Completamente al azar	2	$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$
Bloques completos al azar	3	$Y_{ij} = \mu + \tau_i + \gamma_j + \varepsilon_{ij}$
Cuadro latino	4	$Y_{ijk} = \mu + \tau_i + \varphi_j + \delta_k + \varepsilon_{ijk}$

Y: respuesta; μ efecto promedio; τ_i : efecto del tratamiento; γ_j : efecto de repeticiones; φ_j : efecto de filas; δ_k : efecto de columnas. El efecto del error experimental se representa por ε en sus distintos componentes.

Para establecer la existencia de diferencias significativas entre tratamientos de un ensayo, el ANDEVA, en la situación más simple, compara la variabilidad o varianza entre las medias de los tratamientos con la varianza de los datos de cada tratamiento (varianza del error experimental). A la primera se la conoce como varianza entre grupos y a la segunda como varianza dentro de los grupos. Si la varianza entre grupos es mayor que la varianza dentro de los grupos, se dice que existe variación significativa entre los tratamientos; esto en términos más explícitos quiere decir que existe una mayor variabilidad entre las medias de los tratamientos que aquella variabilidad aleatoria (error experimental) que ocurre dentro de cada tratamiento. A la relación de estas variaciones se le conoce como *F* de Fisher⁴⁸ (estadístico de prueba u observado), cuyos resultados pueden ser menores, iguales o mayores a 1. Si $F \leq 1$, esto determina automáticamente que no existen diferencias significativas entre tratamientos; y cuando $F > 1$, se declara que existen diferencias significativas siempre que el nivel o cuantil de probabilidad del error I (α) sea como máximo del 5%.

Como se ha visto en los párrafos anteriores, el ANDEVA permite determinar diferencias entre tratamientos utilizando una relación de varianzas, donde la correspondiente a tratamientos se ubica en el numerador y la del error experimental en el denominador; de allí la gran importancia del cuidado de un ensayo para que la varianza del error experimental se mantenga controlada y se origine únicamente del error aleatorio, con el propósito de tener un análisis de

⁴⁸ A este estadígrafo también se le conoce como *F* de Snedecor, ya que fue George Snedecor quien la definió así en honor a Ronald Fisher.

mayor precisión⁴⁹. Si no existe una planeación y manejo del ensayo con rigor científico, que involucre también la eliminación de preferencias o subjetividades, se tendrá un error experimental cargado de un error no aleatorio (varianza externa) que, aun habiendo efectos reales de los tratamientos, hará impreciso al ensayo e incapaz de detectar las diferencias de dichos efectos.

Como se ha mencionado antes, el ANDEVA es una herramienta estadística relativamente robusta, esto es, puede admitir cierto grado de dispersión en los datos; no obstante, no se puede abusar de su uso en cualquier situación en donde el error experimental es comparativamente alto. En estos casos hay que realizar una evaluación crítica del tipo de variable que se somete al análisis de varianza, pues en el ámbito agrícola existen características medibles que no deben presentar dispersiones o errores de magnitud dentro de un mismo tratamiento; mientras que otras, como aquellas que están relacionadas con las plagas, es su naturaleza variar de forma sustancial intra-grupalmente (véase la tabla 4.7).

Un experimento que presenta una variable cuya naturaleza no debería comportarse como significativamente dispersa no puede ser avalado en ningún caso, por lo tanto, no deben buscarse alternativas que lo “validen”. En estas circunstancias, antes de desechar todo el trabajo experimental, el investigador previamente deberá revisar sus datos en la búsqueda de algún error de transcripción y corregir aquellos valores (si es que los hubiera) que están ocasionando la falta de normalidad y/u homocedasticidad. Si el problema no es de escritura, éste se debe a un mal manejo del ensayo; en consecuencia, es una experiencia sin precisión y que no debe ser aceptada.

Si la variable tiene por característica ser dispersa, tales como las que están relacionadas a insectos, malezas o enfermedades, la primera acción que deberá realizar el investigador es ver la necesidad del ajuste previo de sus datos antes de someterlos al ANDEVA; pues la distribución de sus errores podría ser de tipo binomial, Poisson o Log-normales en algunos casos. Este cambio de escala de la variable estabiliza las varianzas de los tratamientos indicados en los datos y permite que esta herramienta estadística sea más precisa para detectar diferencias significativas entre tratamientos.

Si el ajuste no alcanza para estabilizar la varianza de los datos, una segunda alternativa es utilizar pruebas estadísticas no paramétricas, llamadas de distribución libre por cuanto no requieren del cumplimiento de normalidad, homocedasticidad ni el de independencia. En este caso, como pruebas análogas al ANDEVA para distribuciones completamente al azar y de bloques completos al azar, pueden utilizarse los test de Kruskal – Wallis para la primera y la de Friedman para la segunda, dos pruebas estadísticas que se explican más adelante. Una tercera opción es el uso de los modelos lineales generalizados (MLG) con los cuales tampoco se precisa normalidad y homocedasticidad de los residuos, ya que se pueden observar efectos generando modelos lineales bajo otras distribuciones de probabilidad (binominal, Poisson, etc.).

La decisión del nivel de error α en los ensayos debe ser establecida juiciosamente. Si bien las pruebas convencionalmente se realizan al 5%, este valor puede resultar peligroso en ensayos a partir de los cuales se tomarán decisiones que involucren una importante inversión

⁴⁹ Si bien la comparación que se realiza en el ANDEVA es respecto de las medias de los tratamientos, tal como se describe en la hipótesis nula de prueba ($H_0: \mu_1 = \mu_2 = \dots = \mu_n$) que generalmente se plantea, esto debe ser visto sólo de forma convencional. Lo que realmente se comparan son las varianzas: la que corresponde a los tratamientos y la que representa al error experimental. Por ello, aun a pesar de que la hipótesis alternativa ($H_a: \mu_1 \neq \mu_2 \neq \dots \neq \mu_n$) que se establece sugiere una prueba bilateral (dos colas), bajo la distribución de probabilidad F de Fisher, este test siempre ha sido de forma lateral por derecha (de una cola). Este criterio se fundamenta en el sentido que de que, en los experimentos planeados, se supone que el investigador tiene control de las fuentes de variación secundaria, de tal forma que él esperará que la varianza entre los tratamientos sea mayor a la varianza que hay dentro de éstos.

económica. Dado que la agronomía, y en específico la agricultura, es una actividad con fines de lucro, decidir un cambio en la forma de producir va mucho más allá de tener un valor de significancia apenas inferior o escasamente superior al “famoso” 0,05; límite crítico que, de paso, no tiene un fundamento científico y ha sido adoptado tradicionalmente antes que por ciencia. El investigador deberá protegerse contra efectos errados (o de caer en el error I), primero diseñando adecuadamente su experimento (con las suficientes repeticiones) y manejándolo homogéneamente, y luego adoptando valores de alfa con cierto nivel de exigencia.

Asimismo, dada a la gran variabilidad que presenta el medio que rodea a los experimentos agrícolas, los cuales necesariamente deben desarrollarse en las condiciones en que se realiza un cultivo comercial para poder tener resultados más acordes a la realidad, siempre será necesario utilizar pruebas estadísticas estrictas. En este sentido, una vez que el ANDEVA ha detectado diferencias significativas, una de las alternativas entre las pruebas de comparación múltiple para definir el o los tratamientos de mejor respuesta es el test de Tukey. Esta última es una prueba ampliamente utilizada por presentar valores críticos estrictos y por tener la particularidad de mantener fijo el valor de α pre-establecido, con lo cual se protege una decisión contra el error I.

Prueba de Kruskal – Wallis

En los experimentos agronómicos también pueden existir casos de variables que ningún tipo de ajuste logra que sus residuales tengan cierto comportamiento de normalidad y homocedasticidad, porque tienen una fuerte dispersión y sus varianzas son totalmente distintas. En estas circunstancias, una de las alternativas que podría utilizar el investigador es remitir a las pruebas no paramétricas, las cuales no se ajustan a ningún tipo de distribución, de allí que se las defina también como pruebas de distribución libre. Por esta característica son pruebas que no están obligadas a cumplir con los supuestos de normalidad, homocedasticidad ni de independencia.

Una de estas pruebas es el estadístico propuesto por Kruskal y Wallis (1952). Prueba no paramétrica que se la ha diseñado para analizar la variabilidad de más de dos niveles (más de dos tratamientos) dentro de un factor de estudio, cuya información es muy dispersa y se comporta en una forma ordinal. Es una prueba análoga a la técnica paramétrica del análisis de varianza en experimentos unifactoriales (analiza un solo factor de estudio) sin bloqueo, o llamado también, análisis de varianza en un sentido; es decir, la prueba es susceptible de aplicarse cuando el experimento ha utilizado una distribución completamente al azar (DCA). La diferencia de esta prueba respecto del ANDEVA, es que ésta no requiere que los datos se ajusten a una distribución normal ni cumplan con el criterio de varianzas iguales.

No obstante, hay que tener una visión crítica sobre la factibilidad de descartar el uso del ANDEVA en una situación específica, ya que esta poderosa prueba paramétrica no se restringe estrictamente al cumplimiento de las condiciones de normalidad y homocedasticidad; ésta admite hasta cierto punto que la información se presente de manera relativamente dispersa o variada. Es decir, la robustez del ANDEVA se extiende un poco más allá de la perfecta uniformidad teórica que pudieran presentar un conjunto de datos, admitiendo cierto grado de variabilidad de la información.

La prueba de Kruskal-Wallis, también llamada prueba H , al igual que las pruebas paramétricas, se la utiliza para probar si un conjunto de varias muestras aleatorias e independientes, producto de la aplicación de dos o más tratamientos, provienen de una misma población (H_0); o en su defecto, provienen de poblaciones distintas (H_a). Si las muestras provienen de una misma población, la prueba permitirá demostrar que los datos aproximadamente se ajustan a una

distribución *Chi – cuadrada* (χ^2) con $k - 1$ grados de libertad, siempre que dichas muestras contengan cinco o más datos.

Para la aplicación de este test, lo primero que se debe realizar es ordenar la información en forma ascendente de todos los grupos o tratamientos en forma conjunta, como si se tratara de un sólo grupo, definiendo distintos rangos. Luego que se han establecido los rangos de toda la información, se puede obtener el estadístico calculado H mediante la siguiente expresión:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1) \quad (7.2)$$

siendo N la suma de las muestras de cada grupo o tratamiento, esto es: $N = n_1 + n_2 + \dots + n_k$; R_i son los rangos asignados a cada uno de los grupos y n_i el número de datos de cada tratamiento.

La expresión 7.2 es aplicable cuando en el conjunto de datos, una vez ordenados ascendentemente, no tiene valores repetidos (conocidos como empates), ya sea dentro de una misma muestra o entre muestras; sin embargo, si el conjunto de datos ordenados contiene valores repetidos es mejor utilizar la expresión 6.3 siguiente:

$$H = \frac{\frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1)}{1 - \frac{\sum(I^3 - I)}{N^3 - N}} \quad (7.3)$$

donde I representa al número de empates dentro de cada conjunto de valores repetidos.

Por cuanto la condición de la hipótesis nula (H_o) es que los datos se ajusten aproximadamente a la distribución Chi cuadrado, el rechazo de H_o se produce cuando el estadístico H es mayor a χ_{α}^2 con $k - 1$ grados de libertad.

Ejemplo de aplicación 7.1

Los datos que se indican en la tabla 7.6 son los correspondientes al número de moscas blancas muertas por planta, como respuesta en un experimento en donde se han probado tres dosis de un insecticida aplicado en soya, utilizando la distribución completamente al azar. Para desarrollar este ejemplo de aplicación se ha utilizado el software Infostat en la versión estudiantil.

Tabla 7.6. Número de insectos muertos/planta.

Dosis 1	Dosis 2	Dosis 3
8	43	59
19	22	72
25	35	3
6	57	15
0	3	1
20	58	47

6	72	40
12	98	125

La hipótesis nula a comprobar es la siguiente:

H_0 : Las dosis tienen efectos semejantes.

Aunque el investigador por intuición y aplicando pruebas preliminares podría verificar la falta de normalidad y de varianzas homogéneas⁵⁰, antes de proceder a la aplicación del test de Kruskal – Wallis, con los datos de la tabla 7.6 se ha procedido a realizar dos gráficas indicadas mediante la figura 7.1. En la gráfica de la izquierda se han relacionado los residuos con sus correspondientes probabilidades para poder observar si los datos cumplen con el supuesto de normalidad; mientras que en la gráfica de la derecha se han planteando las medias versus sus residuales. En el caso de normalidad, los datos no se alinean contundentemente, por lo tanto, es previsible una aparente falta de este supuesto; mientras que la falta de homocedasticidad es contundente, ya que los residuales no se distribuyen formando una franja homogénea.

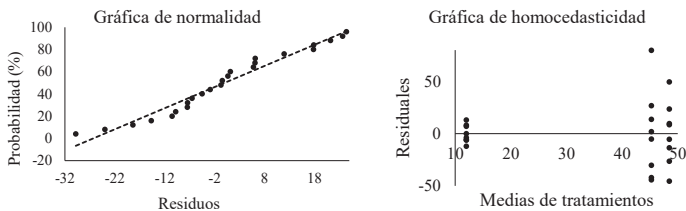


Figura 7.1. Verificación gráfica de la falta de normalidad y de homocedasticidad.

La base de datos estructurada en el software para el análisis correspondiente es la siguiente:

⁵⁰ Una forma preliminar de verificar si un grupo de datos se ajustan a una distribución normal es comparando la media con la mediana y la moda de dichos datos. Estos estadísticos descriptivos deben ser cercanamente similares (en una distribución normal ideal deben ser iguales) para saber que la información presenta normalidad de forma aproximada; sin embargo, para apoyar a esta forma de verificación podría obtenerse el coeficiente de asimetría y de curtosis (estadígrafos de forma), considerando que unos datos presentan normalidad cuando los valores de estos dos coeficientes se ubican en el rango de $\pm 0,5$. Para evaluar la homocedasticidad de los tratamientos de un ensayo, un criterio práctico es obtener la varianza de cada grupo, para lo cual se relacionan la mayor y la menor de todas de entre los grupos y se obtiene así un estadístico de prueba conocido como test de Hartley o prueba F_{MAX} . Para un experimento adecuado, con un mínimo de 12 grados de libertad en el error experimental, tener un valor de F_{MAX} mayor a 10 (con este valor se supera los valores críticos) puede ser suficiente para decidir que no existe homogeneidad de las varianzas.

Caso	TRATAMIENTOS	Insectos muertos/planta
1	Dosis 1	8,00
2	Dosis 1	19,00
3	Dosis 1	25,00
4	Dosis 1	6,00
5	Dosis 1	0,00
6	Dosis 1	20,00
7	Dosis 1	6,00
8	Dosis 1	12,00
9	Dosis 2	43,00
10	Dosis 2	22,00
11	Dosis 2	35,00
12	Dosis 2	57,00
13	Dosis 2	3,00
14	Dosis 2	58,00

Figura 7.2. Ingreso de datos a *InfoStat* para la realización del test de Kruskal – Wallis.

Una vez que se ha estructurado la base datos en el software Infostat, la secuencia para el análisis mediante la prueba no paramétrica de Kruskal-Wallis es:

Estadísticas → *Análisis de varianza no paramétrica* → *Kruskal Wallis*

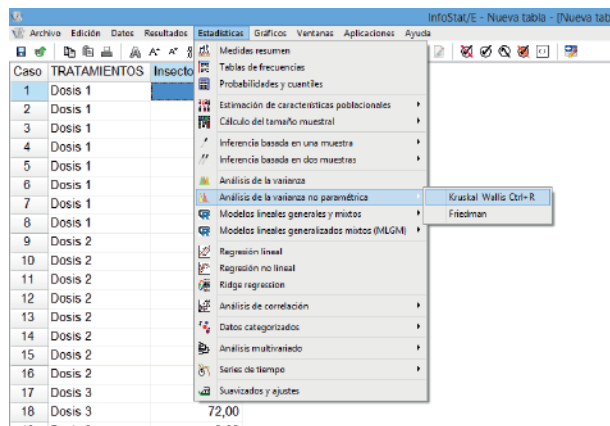


Figura 7.3. Secuencia de pasos en el *InfoStat* para la realización del test de Kruskal-Wallis.

Al realizar la clasificación respectiva en la ventana que inmediatamente se abre y seleccionando los estadígrafos pertinentes, el resultado del análisis es el siguiente:

Prueba de Kruskal Wallis

Variable	TRATAMIENTOS	N	Medias	D.E.	Medianas	H	p
Insectos muertos/planta	Dosis 1	8	12,00	8,57	10,00	5,99	0,0500
Insectos muertos/planta	Dosis 2	8	48,50	29,64	50,00		
Insectos muertos/planta	Dosis 3	8	45,25	41,32	43,50		

Trat.	Ranks
Dosis 1	7,63 A
Dosis 3	14,00 A B
Dosis 2	15,88 B

Medias con una letra común no son significativamente diferentes ($p > 0,05$)

El estadístico H calculado es 5,99, mientras que el P -valor observado es de 0,05, significancia que coincide con el límite máximo convencionalmente adoptado, haciendo que la declaración de diferencias significativas sea un aspecto facultativo. Es decir, con esta probabilidad, tomar la decisión de que existen efectos reales entre tratamientos será criterio del investigador. En este análisis también se adjunta el test de comparaciones realizado con el software, en donde se puede notar que la dosis 2 y 3 fueron las que estadísticamente presentaron los efectos más altos indicados a través de sus rangos.

Prueba de Friedman

La prueba de Friedman (1940) es también un test no paramétrico análogo al análisis de varianza en dos sentidos, es decir permite valorar estadísticamente a los tratamientos considerando las repeticiones de éstos. Desde esta perspectiva, este test hace las veces del ANDEVA para una distribución de bloques completos al azar (DBCA). Al igual que la prueba de Kruskal-Wallis, es una de las alternativas cuando ningún ajuste de una variable estabiliza su varianza. Es una prueba de distribución libre que no requiere el cumplimiento de los supuestos de normalidad, homocedasticidad e independencia y es aplicada en datos que al menos tienen escala ordinal.

Se la conoce también como prueba de análisis de varias muestras relacionadas debido a que permite establecer si los n elementos de un grupo k , o el grupo al que se le aplicó un tratamiento, pertenecen a una misma población. Dentro de esta concepción y haciendo analogía al análisis estadístico de un DBCA, con la prueba podemos establecer si existen diferencias entre los bloques o repeticiones a los cuales se les aplicó un mismo tratamiento. En forma general, las hipótesis estadísticas que se valoran en esta prueba son las siguientes:

H_0 : Los elementos que pertenecen a un mismo grupo k son homogéneos. En un experimento, esta hipótesis dice que las unidades experimentales a las que se les aplicó distinto tratamiento, llamada bloque o repetición en los DBCA, son homogéneas o pertenecen a una misma población.

H_a : Los elementos pertenecientes a un mismo grupo k son heterogéneos. En los experimentos se dirá que las unidades experimentales, llamadas bloques o repeticiones, a las que se les aplicó distinto tratamiento, producen efectos diferentes; es decir, provienen de distintas poblaciones.

La prueba se realiza asignando un rango u ordenando los datos en forma horizontal o por filas (bloques). Luego estos rangos, pertenecientes a cada grupo k , se suman con el propósito de obtener el estadístico de prueba de Friedman (χ^2_f), a través de la siguiente expresión:

$$\chi_r^2 = \frac{12 \sum R_i^2}{bk(k+1)} - 3b(k+1) \quad (7.4)$$

siendo χ_r^2 es el estadístico de prueba de Friedman, R_i es el rango en la columna i , b son las filas o bloques y k son los grupos o tratamientos analizados.

El estadístico de prueba χ_r^2 sigue una distribución Chi cuadrado con $k - 1$ grados de libertad, de tal forma que se rechaza la hipótesis nula si el estadístico calculado es mayor al estadístico crítico; es decir: $\chi_r^2 > \chi_{\alpha}^2$.

Para la aplicación de esta prueba, veamos el siguiente ejemplo en el cual se utiliza también el software Infostat en la versión estudiantil.

Ejemplo de aplicación 7.2

Para el detalle de la aplicación del test de Friedman se considera un experimento en donde se han probado 4 insecticidas para controlar el insecto “cogollero” (*Spodoptera frugiperda*) del maíz. Para su desarrollo se ha implementado una distribución en bloques completos al azar, en donde cada tratamiento se corrió en cinco repeticiones o bloques. La variable evaluada fue el número de larvas muertas por planta.

Tabla 7.7. Número de larvas de *S. frugiperda* muertas/planta.

Repeticiones	Tratamientos			
	Insecticida A	Insecticida B	Insecticida C	Insecticida D
I	0	2	0	1
II	4	5	1	7
III	2	1	1	2
IV	3	2	2	4
V	5	1	3	0

La hipótesis nula a probarse es:

H_0 : Los tratamientos tienen efectos semejantes.

Al ingresar los datos al software, la base de datos se forma de columnas que dependerán del número de tratamientos del experimento, en este caso se tiene lo siguiente:

Caso	Insecticida A	Insecticida B	Insecticida C	Insecticida D
1	0	2	0	1
2	4	5	1	7
3	2	1	1	2
4	3	2	2	4
5	5	1	3	0

Figura 7.3. Estructura de la base de datos en Infostat para la realización del test de Friedman.

Luego de estructurada la base datos en el Infostat, la secuencia para el análisis mediante la prueba no paramétrica de Friedman es:

Estadísticas → Análisis de varianza no paramétrica → Friedman

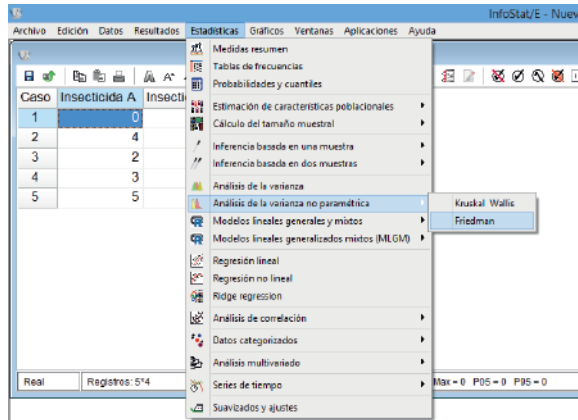


Figura 7.4. Ubicación del test de Friedman en el software Infostat.

Ingresando las variables a la ventana que se despliega y escogiendo la significancia (5% en este ejemplo), el análisis es el siguiente:

Prueba de Friedman

Insecticida A	Insecticida B	Insecticida C	Insecticida D	T ²	p
2,80	2,40	1,70	3,10	1,26	0,3328

Minima diferencia significativa entre suma de rangos = 8,320

Tratamiento	Suma (Ranks)	Media (Ranks)	n
Insecticida C	8,50	1,70	5 A
Insecticida B	12,00	2,40	5 A
Insecticida A	14,00	2,80	5 A
Insecticida D	15,50	3,10	5 A

Medias con una letra común no son significativamente diferentes (p > 0,050)

En el análisis se ha obtenido un P-valor de 0,3328; por lo tanto, según esta prueba no existen diferencias significativas entre los tratamientos y se acepta la hipótesis nula.

Recomendaciones sobre las labores culturales en los experimentos

Como se ha indicado antes en este documento, el manejo del ensayo es de suma importancia si lo que se quiere es tener un experimento con la debida precisión y que se ajuste al rigor científico que demandan estas investigaciones. Las labores que deben realizarse en las parcelas de un ensayo deberán tener el mayor cuidado, de tal forma que el manejo de éstas se realice dentro de términos de total homogeneidad. Entre las labores que se llevan a cabo en los experimentos a nivel de campo, que en algunos casos pueden diferir, están la preparación de suelos, la siembra, podas en algunos casos, el riego, la fertilización, el control de malezas y el control de plagas. Son labores que necesariamente deberán realizarse de manera uniforme, salvo que alguna de ellas sea motivo de estudio y constituyan los tratamientos a evaluarse.

Preparación de suelos

Dado que en la gran mayoría de los ensayos agronómicos realizados en campo se trata de experimentos de aplicación práctica, la preparación de suelos deberá ser idéntica a la que tradicionalmente se realiza en la zona, de tal forma que los resultados se den para una condición real del agricultor. Allí puede existir una gama de labranzas, en donde incluso se involucra a aquella en la que no se realiza preparación (labranza cero); sin embargo, si lo que se pretende es realizar un ensayo de comparación y/o adaptación de semillas, cuyas demandas de labores tecnificadas no pueden eludirse para que exista una expresión genética máxima del germoplasma, siempre será un menester realizar una adecuada preparación del suelo.

Siembra

Lo primero que deberá considerar el investigador al momento de sembrar sus plantas en el ensayo es su calidad. Tanto para experimentos en donde la semilla no es motivo de estudio como en los que sí lo son, el germoplasma a utilizar deberá ser certificado. Bajo ningún concepto y para cumplir con el rigor científico, no se deben utilizar semillas recicladas o materiales del cual no se tiene certeza de su pureza; el uso de estas últimas puede incrementar el error experimental, haciendo que el ensayo pierda precisión, inclusive también llevarlos al extremo de caer en el error I (la decisión errada más costosa de los ensayos).

Otro elemento que debe tenerse en cuenta en el caso de los ensayos en donde se efectúan siembras directas, es que no se deben realizar resiembras. Para evitar esta situación deberán hacerse las debidas pruebas de germinación y depositar más de una semilla por sitio de

siembra, de tal forma que una vez que las plantas tengan unas tres hojas verdaderas o hasta unos 20 cm de altura se proceda al raleo correspondiente.

Es una labor que debe desarrollarse de manera homogénea, depositando las semillas a una misma profundidad y conservando siempre el distanciamiento que fue establecido en el protocolo o proyecto de investigación, considerando el espaciado común de la zona; salvo que el experimento consista en la evaluación de distintos distanciamientos de siembra, ante lo cual se tendrán varios espaciados de prueba. Cuando se trate de siembras por trasplante, se deberán seleccionar las plántulas que visualmente presenten una misma condición; además tanto en esta forma de siembra como en la directa, esta labor debe efectuarse en un mismo día.

En siembras directas que se realizan al voleo, como en el caso del arroz o de la soya, es de sumo cuidado que al volear la semilla ésta se distribuya de manera uniforme en el área experimental. Quien realice esta labor debe tener la suficiente experiencia y estar debidamente consciente de que, si no se volean homogéneamente, los datos que se generen propiciarán a que el experimento, desde el inicio, adolezca de precisión. La recomendación es que se separe la cantidad de semilla en unidades de peso para cada unidad experimental o parcela y ésta se distribuya totalmente de acuerdo a su cantidad. Es preciso tener en cuenta la gran importancia que tiene el voleo uniforme en ensayos de siembra directa; si ésta no se realiza de esta forma, la información que se extraiga puede no ser válida, especialmente en variables como el rendimiento, en donde se evalúan unidades de muestreo que abarcan una determinada área (1 o más metros cuadrados) de la unidad experimental. Un voleo deficiente en siembras directas presenta una fuerte variabilidad poblacional en las parcelas.

Riego

Al igual que las otras labores, el riego debe llevarse a cabo de manera homogénea en las parcelas experimentales y de forma cautelosa, especialmente cuando se riega por gravedad. Aplicaciones con caudales fuertes podrían lavar los fertilizantes y borrar el efecto de los tratamientos cuando el factor de estudio corresponde a enmiendas edáficas o aplicaciones de nutrientes sintéticos u orgánicos. Cuando se trata de ensayos de adaptación o de comparación de semillas, siempre será recomendable manejar el riego de forma optimizada; esto es, con tinas de evaporación para determinar la evapotranspiración, tensiómetros para medir los contenidos de humedad del suelo y el uso de dosificadores o medidores de caudal, de manera que se obtengan dosis de riego adecuadas a la capacidad productiva de la semilla.

Cuando se utilizan sistemas de riego presurizados, tales como el de aspersión, el de microaspersión o el de goteo, además de la contabilización de la evapotranspiración del cultivo y de los contenidos de humedad en el suelo, es necesario que se evalúe el sistema de riego para definir la uniformidad con el propósito de realizar las correcciones necesarias y que el agua se distribuya homogéneamente.

En cultivos como el arroz, en donde tradicionalmente se inunda el suelo y los tratamientos deben aplicarse de forma edáfica, habrá que realizar muros de protección en cada unidad experimental para evitar el traslape de las alternativas de prueba. Asimismo, en esta situación siempre será recomendable utilizar cabeceras e hileras de borde relativamente más amplias para evitar el cruce interno de la humedad y que se pueda producir un traslape entre los efectos de los tratamientos.

Fertilización

En los ensayos agrícolas, el manejo de la fertilización debe efectuarse con mucho cuidado, especialmente en la distribución de los mismos dentro de cada unidad experimental. En estos casos son muy importantes las balanzas de precisión que permitan establecer las dosis exactas

en cada parcela. En los experimentos con rigor, el establecimiento de las dosis deberá realizarse considerando el estado de fertilidad del suelo; esto es, se definen a partir de un análisis de suelo de laboratorios certificados. Si el experimento tiene un fin demostrativo, los cuales están dirigidos a los agricultores y quienes en la gran mayoría no realizan evaluaciones de fertilidad de sus suelos, es aconsejable de que la metodología o dosis que se pretende demostrar para la aceptación de los productores se defina más por la demanda del cultivo que por la condición nutricional del suelo descrita a partir de un análisis del mismo.

Definir la fertilización en experimentos demostrativos a partir de la demanda de los cultivos permite proteger los resultados de alguna metodología expuesta a nivel experimental, debido a que la condición de fertilidad de los suelos es muy cambiante en el espacio y en el tiempo; de allí que se diga que los ensayos sobre fertilizantes y/o dosis de éstos tengan una fuerte restricción tanto espacial como temporal.

Se comete un error al considerar la no aplicación de fertilizantes sintéticos cuando se están probando ciertos productos nutricionales que se emplean de forma generalmente foliar, o incluso otros que también se dosifican de manera edáfica. Hay que tener mucho criterio para suprimir la fertilización sintética básica que se aplica al suelo, especialmente en cultivos comerciales que tienen fuertes demandas de éstos. Lo adecuado, dada la necesidad de hacer inclusiva la agricultura orgánica, es realizar fertilizaciones en déficits para suplirlas con la tecnología orgánica que se esté probando. El uso de enmiendas orgánicas no implica retirar totalmente la fertilización tradicional, especialmente en los actuales momentos en donde se demandan exigentes eficiencias debido la globalización de los mercados y a que se tienen suelos degradados en su gran mayoría. El uso de productos nutricionales que se aplican vía foliar y que constituyen los factores de estudio, tales como los oligoelementos, deben ser considerados como complementarios a la fertilización básica con el propósito de nutrir completamente a los cultivos.

El investigador debe estar consciente de lo que pretende lograr con su ensayo. La realización de cualquier experimento agrícola debe tener como fin ineludible buscar soluciones prácticas para una agricultura comercial que satisfaga a una población que crece de forma exponencial. Como se ha mencionado, la supresión total de la fertilización sintética (y de otras alternativas agroquímicas) productivamente no es posible, ni aun con las macro aplicaciones de enmiendas a base de estiércol que no igualan a las producciones que se obtienen con los fertilizantes sintéticos; situación que coadyuva a una baja sustancial de la rentabilidad, o incluso pérdidas que harían que el proceso de producción de cualquier cultivo sea una actividad inviable. Si bien la agricultura comercial, a más de alimentar a la población, tiene una visión de lucro, lo que no se puede obviar y sobre lo cual se tiene que trabajar, es en el manejo eficiente de los recursos no sólo en el aspecto económico, sino también en la parte ambiental. La agricultura desarrollada sobre pilares de sostenibilidad no sólo implica el uso exclusivo de medios orgánicos, mayoritariamente también depende de alternativas inorgánicas que, bien manejadas, deben satisfacer la exigencia de la tríada entre sociedad, productor y ambiente.

Control de malezas en los ensayos

A nivel experimental el control de las malezas puede realizarse, ya sea de forma manual o utilizando algún herbicida. La eliminación de las malezas de forma manual se realiza generalmente en experimentos con rigor, debido a la incidencia que podría tener el agroquímico en la genética del cultivo como, por ejemplo, en los ensayos de cruce genético o de hibridación. También el control manual de las malezas puede realizarse en los ensayos en donde se evalúan dosis y/o herbicidas en campos con cultivo, el mismo que deberá efectuarse en las parcelas testigo para valorar al final algún efecto fitotóxico del mismo, cuya mejor variable indicadora deberá ser el rendimiento.

El control de malezas no puede realizarse de forma parcial, con intervalos de tiempo más allá de 24 o 48 horas. Cuando menos, si el ensayo tiene múltiples unidades experimentales, el investigador tendrá que buscar mano de obra adicional para que pueda realizar esta labor en un mismo día. Dejar una parte del experimento, o aun cuando se realiza por bloques, en periodos largos de tiempo entre cada limpieza puede favorecer una interferencia de las arvenses con el desarrollo normal del cultivo, haciendo que se confunda el efecto de los tratamientos y se produzca una disminución de la precisión experimental.

Control de insectos y de enfermedades

La aplicación de los insecticidas y de los fungicidas, en algunos casos y siempre que éstos no sean el factor de estudio, a nivel experimental debe hacerse en forma preventiva para poder evaluar el efecto total de los tratamientos, especialmente cuando la respuesta involucra el rendimiento del cultivo. En otros casos, cuando se evalúan materiales vegetativos, se podría permitir el ataque de plagas con el propósito de identificar la susceptibilidad o resistencia de la especie. Si en forma localizada se identifica un ataque, el cual debe valorarse a la luz de un umbral económico, la aplicación de los plaguicidas deberá efectuarse de forma general dentro del ensayo para guardar el cuidado de homogeneidad en el experimento.

CAPÍTULO VIII

ALGUNAS ALTERNATIVAS DE OPTIMIZACIÓN

En un ambiente tan mudable como en el que se desarrolla el proceso de producción agrícola, hablar de una alternativa óptima en la experimentación agronómica podría ser un criterio desajustado del método inductivo...



Arreglos experimentales para optimización

Ubicar respuestas óptimas (máximas o mínimas) debido al efecto de factores que se evalúan en un experimento, en términos de utilidad práctica o real, va a depender del área científica en donde se aplique la metodología. A nivel industrial, por ejemplo, el punto óptimo de operación de un proceso establecido a través de varios experimentos secuenciales se convierte en un protocolo que deberá aplicarse siempre, si lo que se busca es mejorar rentablemente el proceso productivo. Esta situación es diferente para los experimentos que se realizan en el área de la agronomía con fines de optimización, en donde las fuentes de variación secundarias, móviles en el espacio y en el tiempo, no permiten tener resultados cuya aplicabilidad sea general y duradera.

Si bien existen métodos como el de superficie de respuesta (MSR), basados en técnicas estadísticas y matemáticas, que inician con arreglos factoriales exploratorios del tipo 2^K completos o fraccionados y mediante los cuales se hace posible arribar a respuestas óptimas utilizando modelos de regresión múltiple, en la agronomía su aplicación resulta costosa y lo que se obtiene como resultado tiene escasa vigencia de aplicación.

Para fundamentar el criterio antes descrito, tomemos de ejemplo la necesidad de optimizar una dosis de fertilizantes cuya fórmula contenga a los macronutrientes nitrógeno, fósforo y potasio aplicados en cualquier cultivo. Ahora imaginemos que en este ensayo se decide utilizar la MSR para determinar la dosis compuesta en la que se logre el máximo rendimiento, ante lo cual se tendrán las siguientes situaciones:

1. En la MSR se parte de experimentos del tipo 2^K que inicialmente tienen característica exploratoria, con el propósito de establecer los factores que realmente inciden en la respuesta de interés. Posteriormente se van incluyendo puntos de experimentación central y axial (nuevos tratamientos) para verificar si los efectos de los factores evaluados tienen comportamientos de curvatura, con los que se define un área de respuesta en donde se localiza el mayor rendimiento. Este proceso experimental involucra una serie de ensayos secuenciales que, en la agronomía y en algunos casos de otras ciencias, no es posible desarrollarlos por los altos costos que conlleva su realización y que en el menor de los tiempos hay que esperar un mínimo de 6 meses para poder tener resultados de un solo experimento de toda una obligada secuencia. De allí que llegar al punto óptimo de rendimiento de un cultivo hayan transcurrido, al menos un par de años.

Otro elemento en contra de la adopción plena de la MSR en la agronomía va en el sentido de ir adicionando corridas experimentales a las ya dadas, lo que a su vez demanda que en las posteriores corridas de tratamientos se tengan las mismas condiciones experimentales de las corridas precedentes. Esta situación es muy difícil o imposible de conseguir en los experimentos agrícolas que se realizan en campo, en donde las condiciones del suelo y del clima cambian en el tiempo.

2. En la agricultura, como en todo negocio, lo que define la adopción de una metodología no siempre es lo que dice la estadística, sino la rentabilidad de la misma. Con exclusividad para los experimentos que sirven para la toma de decisiones de inversión económica, la estadística es una herramienta secundaria que, bajo una probabilidad determinada para el error, permite establecer cuánto incide en el rendimiento y en el desarrollo fisiológico de un cultivo un tratamiento (que puede estar representando a uno o varios factores) de interés para el investigador; sin embargo, no siempre el tratamiento que estadísticamente registra el mayor rendimiento del cultivo es el más rentable. Hay que tener en cuenta que un mejoramiento sustancial de la producción agrícola involucra también una inversión importante en el manejo del cultivo; de tal

forma que al final, habrá que relacionar las ganancias y los costos para poder establecer si el tratamiento estadísticamente ganador también presenta una evidente rentabilidad que vaya mucho más allá de los costos de oportunidad del productor. En este sentido, la localización del punto óptimo mediante la MSR en la agronomía puede llegar a ser un gasto innecesario.

3. Los experimentos agronómicos son quizás los que más están expuestos a fuentes de variación secundaria; fuentes que, como el estado nutricional del suelo, varían en el espacio y en el tiempo. Una condición específica del suelo puede cambiar sustancialmente a escasos metros de distancia y variar de un periodo de cosecha a otro por factores como la lluvia, que hacen que la fertilidad del mismo no se mantenga estable. Por consiguiente, la duración de la validez de una dosis de fertilización compuesta, establecida mediante el MSR, puede llegar incluso a ser óptima sólo en el experimento; ya que posteriormente su aplicación comercial podría quedar obsoleta. Además, dichas dosis “óptimas” están dadas bajo unas condiciones de manejo en cuanto a preparación del suelo, aplicación del riego, control de malezas, control de plagas, entre otras, que muchas veces distan diametralmente opuestas a las que realiza el agricultor común.

También hay que considerar que el desarrollo de los experimentos con los arreglos factoriales del tipo 2^K , en los cuales se evalúan sólo dos niveles por factor debido a la fuerte demanda de tiempo y recursos económicos que requieren los ensayos agronómicos, pueden ser alternativas de investigación improductivas. Además, dado que la MSR requiere que los factores de estudio sean cuantitativos, éstos quedan relegados a ensayos en donde se traten con variables independientes de característica nutricional; que como se ha mencionado en los párrafos anteriores, las dosis óptimas podrían ser sólo experimentales y no comerciales.

A pesar de las limitantes mencionadas respecto del uso de la MSR en la agronomía, existen alternativas de arreglos factoriales incompletos (o algunos aumentados), que podrían aplicarse para más o menos definir en experimentos únicos, los mejores niveles factoriales en donde se obtienen los mayores rendimientos de un cultivo. Estos arreglos experimentales factoriales parciales, en ciertos casos, pueden tener ventaja sobre los arreglos factoriales completos en el sentido de que se pueden evaluar más de dos factores y hasta siete niveles de cada uno, con menos combinaciones factoriales (tratamientos). Así, por ejemplo, para un experimento de tres factores con cuatro niveles de cada uno, se tendrá un factorial completo de 64 tratamientos, cuyo desarrollo es complicado en el ámbito económico y en precisión experimental debido al incremento de la heterogeneidad del suelo y del ambiente en general que podrían aumentar el error experimental. Esta situación es diferente para un factorial incompleto como, por ejemplo, los llamados arreglos “Plan Puebla”, el mismo que, para los mismos factores y niveles del ejemplo indicado generarían sólo 14 tratamientos.

Además de los arreglos Plan Puebla, existen una serie de arreglos factoriales incompletos que pueden utilizarse en la búsqueda de niveles o tratamientos más o menos óptimos (o mejor, más o menos adecuados) en escala cuantitativa, como es el caso de los experimentos que se realizan con fertilizantes. Desde este punto de vista, dependiendo de la estructura del arreglo de tratamientos, un factorial incompleto puede ofrecernos la misma precisión que un factorial completo, pero con mucha más eficiencia en el uso de recursos. Uno de estos tipos de diseño de tratamientos son los arreglos *San Cristóbal* para dos y tres factores que se indican en los siguientes párrafos.

Arreglos factoriales *San Cristóbal*

Los factoriales incompletos *San Cristóbal*, tal como se ha indicado, requieren una escala cuantitativa de los niveles para poder obtener los modelos de regresión que al final permitan deducir los puntos óptimos. Estas escalas deben estar equidistantemente espaciadas y contener 4 niveles. Este requerimiento de tener niveles cuantitativos es lo que le dio origen a este arreglo, ya que inicialmente fue diseñado para experimentos sobre combinaciones de fertilizantes y es en donde comúnmente se lo utiliza.

Un experimento con dos factores a cuatro niveles genera un factorial compuesto de 16 tratamientos o combinaciones; no obstante, en un arreglo *San Cristóbal*, en este mismo experimento sólo se evaluarían siete combinaciones factoriales. Estas combinaciones representan a siete puntos estratégicamente seleccionados dentro del espacio experimental que definen las 16 combinaciones factoriales que tendría el experimento completo. Estos puntos estratégicos se indican en la figura 8.1, en la cual se hace evidente la definición de estos puntos partiendo de un eje coordenado, desde donde se escogen los cuatro puntos que forman un cuadrado (factorial completo del tipo 2^2), el punto central de este cuadrado y dos puntos externos o aristas.

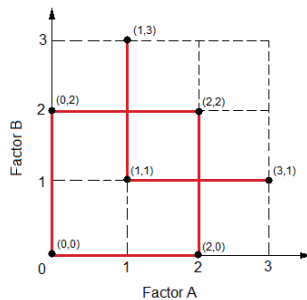


Figura 8.1. Matriz de puntos experimentales en un arreglo *San Cristóbal* para dos factores.

Para el caso de tener un experimento con tres factores, dado que la técnica *San Cristóbal* considera 4 niveles, se tendrían 64 combinaciones factoriales; los cuales, a nivel de campo, tendrían serias limitaciones para poder desarrollarlos en la práctica. Bajo el criterio del *San Cristóbal*, un experimento de esta magnitud sólo contendría 12 combinaciones factoriales, estratégicamente establecidos dentro del espacio experimental. Estos 12 tratamientos, de un total de 64, se detallan en la figura 8.2.

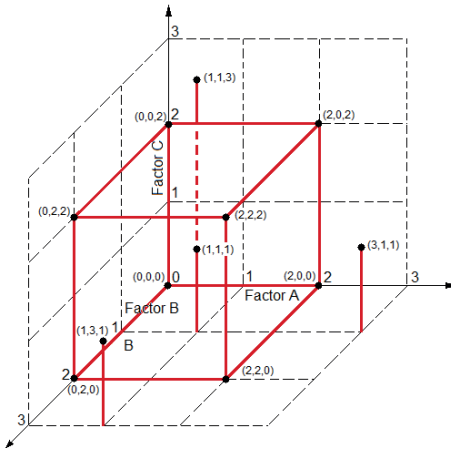


Figura 8.2. Matriz de puntos experimentales en un arreglo San Cristóbal para tres factores.

Puede notarse en la figura 8.2 como está integrado el arreglo: primero con un 2^k , que en este caso es un 2^3 , más un tratamiento central y tres combinaciones externas. Las combinaciones factoriales, según el arreglo San Cristóbal para dos y tres factores, se indican en la tabla 8.1.

Tabla 8.1. Combinaciones factoriales en arreglo San Cristóbal para dos y tres factores.

N°	Dos factores		N°	Tres factores		
	Factor A	Factor B		Factor A	Factor B	Factor C
1	0	0	1	0	0	0
2	2	0	2	2	0	0
3	0	2	3	2	0	2
4	2	2	4	0	0	2
5	1	1	5	2	2	0
6	3	1	6	2	2	2
7	1	3	7	0	2	0
			8	0	2	2
			9	1	1	1
			10	3	1	1
			11	1	3	1
			12	1	1	3

Los arreglos San Cristóbal son diseños de segundo orden, por cuanto la relación que tienen los factores dentro de una respuesta bien puede ser representada por un polinomio de segundo grado. Los modelos de regresión a los que se ajustan aproximadamente estos arreglos, con dos y con tres factores, en su orden, son los siguientes:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{12} X_1 X_2 + \varepsilon \quad (8.1)$$

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 + \beta_{23} X_2 X_3 + \varepsilon \quad (8.2)$$

En estos modelos matemáticos se representan los efectos lineales, los cuadráticos y las interacciones dobles; de allí que también a este tipo de arreglo se lo considere como un factorial fraccionado, que deja de lado las interacciones de mayor orden.

Los arreglos factoriales San Cristóbal se desarrollan utilizando un diseño experimental básico como el de bloques completos al azar; no obstante, si en el área experimental hay alguna evidencia fuerte que indique variabilidad en las parcelas en más de una fuente de variación secundaria, podría optarse por otro diseño que controle de mejor manera estas fuentes de error. El análisis de los datos se completa realizando los siguientes pasos:

1. Se realiza el análisis de varianza para detectar diferencias significativas entre los tratamientos valorados, es decir, entre las combinaciones factoriales que determinan el arreglo. Es preciso indicar que este análisis es básico y no desglosa la variabilidad de cada factor evaluado.
2. Se obtiene el modelo de regresión que se ajusta para cada caso, siempre que en el análisis de varianza se hayan detectado diferencias significativas en la fuente de tratamientos.
3. Se aplican las pruebas de hipótesis para evaluar cada uno de los parámetros del modelo de regresión obtenido en el paso 2. Este análisis permitirá ajustar dicho modelo con los parámetros que hayan presentado significancias.

Ejemplo de aplicación 8.1

Se ha realizado un experimento sobre nutrientes en el cultivo de maíz a base de los elementos nitrógeno, fósforo y potasio. Para su desarrollo se ha utilizado un diseño de bloques completos al azar con un factorial incompleto, bajo el arreglo de un San Cristóbal. Los tratamientos y su codificación, así como el registro de la variable de respuesta, que en este caso es el rendimiento del cultivo en kg/ha, se detallan en las siguientes tablas:

Tabla 8.2. Datos de rendimiento (kg/ha) de maíz para un arreglo San Cristóbal con fertilizantes.

Nº	Tratamientos			Repeticiones			Promedios
	N	P	K	I	II	III	
1	0	0	0	2350	2875	2645	2623,3
2	2	0	0	3035	3567	2984	3195,3
3	0	2	0	2945	2892	3007	2948,0
4	2	2	0	3956	4125	4232	4104,3
5	0	0	2	3654	4098	3985	3912,3
6	2	0	2	4355	4671	4870	4632,0
7	0	2	2	3054	3245	3150	3149,7
8	2	2	2	6129	6325	5987	6147,0
9	1	1	1	3176	3069	2956	3067,0
10	3	1	1	5874	5945	5345	5721,3
11	1	3	1	4468	4145	4765	4459,3
12	1	1	3	4058	3875	3950	3961,0

Tabla 8.3. Codificación de las dosis.

Nitrógeno	Fósforo	Potasio
0: 0 kg/ha	0: 0 kg/ha	0: 0 kg/ha
1: 50 kg/ha	1: 30 kg/ha	1: 60 kg/ha
2: 100 kg/ha	2: 60 kg/ha	2: 120 kg/ha
3: 150 kg/ha	3: 90 kg/ha	3: 180 kg/ha

1. Elaboración del análisis de varianza

Tabla 8.4. Análisis de varianza de los tratamientos (fórmulas) estudiados.

Fuentes de variación	GL	SC	CM	F	P-valor
Total	35	41494618,56			
Tratamientos	11	40336605,89	3666964,17	78,63	0,000
Repeticiones	2	131969,56	65984,78	1,41	
Error experimental	22	1026043,11	46638,32		

2. Una vez que el análisis de varianza ha detectado diferencias significativas entre los tratamientos, en este segundo paso se debe obtener el modelo de regresión múltiple para la variable de respuesta evaluada.

La codificación para los efectos lineales, cuadráticos y de interacción se indica en la tabla 8.5.

Tabla 8.5. Descripción de la variables regresoras (efectos lineales, cuadráticos e interacción) y de la variable dependiente.

N	P	K	N ²	P ²	K ²	NP	NK	PK	Y
									Promedios (kg/ha)
0	0	0	0	0	0	0	0	0	2623,3
2	0	0	4	0	0	0	0	0	3195,3
0	2	0	0	4	0	0	0	0	2948,0
2	2	0	4	4	0	4	0	0	4104,3
0	0	2	0	0	4	0	0	0	3912,3
2	0	2	4	0	4	0	4	0	4632,0
0	2	2	0	4	4	0	0	4	3149,7
2	2	2	4	4	4	4	4	4	6147,0
1	1	1	1	1	1	1	1	1	3067,0
3	1	1	9	1	1	3	3	1	5721,3
1	3	1	1	9	1	3	1	3	4459,3
1	1	3	1	1	9	1	3	3	3961,0

Resumen

Tabla 8.6.

Estadísticas de la regresión	
Coef. de correlación múltiple	0,979
Coef. de determinación r^2	0,958
r^2 ajustado	0,768
Error típico	531,967
Observaciones	12

Tabla 8.7. Análisis de varianza para la regresión.

Fuentes de variación	GL	SC	CM	F	P-valor
Regresión	9	12879556,79	1431061,87	5,057	0,176
Residuos	2	565978,50	282989,25		
Total	11	13445535,30			

Tabla 8.8. Coeficientes del modelo de regresión y las pruebas de significancia para cada uno de ellos.

Coefficientes	Error típico	Estadístico t	P-valor	
Intersección	2876,589	495,235	5,809	0,028
N	-655,344	586,774	-1,117	0,380
P	-580,428	586,774	-0,989	0,427
K	523,906	586,774	0,893	0,466
N^2	344,044	208,872	1,647	0,241
P^2	244,753	208,872	1,172	0,362
K^2	-66,331	208,872	-0,318	0,781
NP	357,750	188,079	1,902	0,197
NK	248,583	188,079	1,322	0,317
PK	-60,167	188,079	-0,320	0,779

La función de producción correspondiente, tomando los coeficientes que se indican en la tabla 8.8, es la siguiente:

$$\begin{aligned}
 Y = & 2876,59 - 655,34N - 580,43P + 523,91K + 344,04N^2 \\
 & + 244,75P^2 - 66,33K^2 + 357,75NP \\
 & + 248,58NK - 60,17PK
 \end{aligned}
 \tag{a}$$

La variabilidad que presenta este modelo tiene una explicación importante en los factores nitrógeno, fósforo y potasio, según puede observarse en la tabla 8.6, a través del coeficiente de determinación ajustado (0,768). Sin embargo, de acuerdo al análisis de varianza realizado para el modelo (tabla 8.7), éste no reporta significancia; esto es, no es muy adecuado para utilizarse como herramienta de pronóstico.

Por otro lado, si bien entre los tratamientos evaluados (fórmulas fertilizantes) se ha podido detectar diferencias significativas según el análisis de varianza indicado en la tabla 8.4; ninguno de los coeficientes, según el análisis que se detalla en la tabla 8.8, reportan

significancia estadística. En este sentido, en la función de producción dada por la expresión a , no es posible verificar la fórmula fertilizante para un máximo rendimiento del cultivo.

Una alternativa suplementaria que puede realizarse en este caso, en donde los tratamientos presentan diferencias significativas, mientras que los efectos lineales, cuadráticos y de interacción no reportan significancia dentro del modelo, es hacer una comparación de promedios utilizando la prueba de Tukey. Es decir, hacer una comparación de cada fórmula fertilizante con el testigo absoluto (0 - 0 - 0). Esta comparación es la que se detalla en la tabla 8.9, obtenida utilizando el software IBM SPSS 20.0.

Tabla 8.9. Clasificación de promedios del rendimiento (kg/ha), según el test de Tukey.

Tratamientos	n	Subconjunto			
		1	2	3	4
0 - 0 - 0	3	2623,3			
0 - 60 - 0	3	2948,0			
50 - 30 - 60	3	3067,0			
0 - 60 - 120	3	3149,7			
100 - 0 - 0	3	3195,3			
0 - 0 - 120	3		3912,3		
50 - 30 - 180	3		3961,0		
100 - 60 - 0	3		4104,3	4104,3	
50 - 90 - 60	3		4459,3	4459,3	
100 - 0 - 120	3			4632,0	
150 - 30 - 60	3				5721,3
100 - 60 - 120	3				6147,0
Sig.		0,110	0,144	0,176	0,436

Se muestran las medias de los grupos de subconjuntos homogéneos.

Basadas en las medias observadas.

El término de error es la media cuadrática (Error) = 46638,323.

a. Usa el tamaño muestral de la media armónica = 3,00

b. Alfa = 0,05.

Tabla 8.10. Análisis de varianza con separación de factores.

Cuadro de Análisis de la Varianza (SC tipo I)						
F.V.	SC	gl	CM	F	p-valor	
Modelo.	40468575,44	13	3112967,34	66,75	<0,0001	
Nitrógeno	20891693,67	3	6963897,89	149,32	<0,0001	
Fósforo	3266383,72	2	1633191,86	35,02	<0,0001	
Potasio	10461691,50	2	5230845,75	112,16	<0,0001	
Nitrógeno*Fósforo	3071641,50	1	3071641,50	65,86	<0,0001	
Nitrógeno*Potasio	1483048,17	1	1483048,17	31,80	<0,0001	
Fósforo*Potasio	86880,67	1	86880,67	1,86	0,1861	
Nitrógeno*Fósforo*Potasio	1075266,67	1	1075266,67	23,06	0,0001	
REPETICIONES	131969,56	2	65984,78	1,41	0,2642	
Error	1026043,11	22	46638,32			
Total	41494618,56	35				

Por cuanto el propósito de eficiencia experimental que se pretende en los ensayos que utilizan el arreglo San Cristóbal consistente en la determinación de los rendimientos máximos definidos por una dosis que se establece con la función de producción (como la expresión 8.1) a través de la aplicación de derivadas parciales, en este ejemplo no se ha logrado establecer la

dosis de mayor respuesta del cultivo; sin embargo, utilizando un test de comparación múltiple, se han identificado dos fórmulas de fertilizantes (150-30-60 y 100-60-120) con las cuales el rendimiento es estadísticamente superior al resto (tabla 8.9). La selección de estas fórmulas se respalda en el resultado del análisis de varianza con desglose por factores, detallado en la tabla 8.10. En esta última tabla puede observarse la existencia de interacción significativa al combinarse los tres factores (N, P y K).

Como se ha mencionado antes, hay varias alternativas de factoriales incompletos ideados para aplicarse en los experimentos agrícolas que el investigador puede seleccionar. No obstante, si bien tienen una aplicación práctica para los ensayos con plantas, ya que fueron creados para esta situación, tienen sus defectos en cuanto a la falta de rotabilidad y ortogonalidad⁵¹ que podrían adicionarle cierto grado de imprecisión. Además, debido al uso intensivo de la tierra agrícola, podría decirse que no existe la necesidad de utilizar tratamientos testigos (niveles cero) en factores cuantitativos como los fertilizantes; de allí que es recomendable que los intervalos inicien desde un valor mínimo que deberá ser respaldado en la bibliografía y experiencias previas. Ante esta situación hay una muy buena alternativa, también bajo la concepción de un factorial incompleto basado en el arreglo 2^K , como es el caso del diseño central compuesto (DCC) que se explica en el siguiente tema y que bien puede ser utilizado en la agronomía.

Diseño central compuesto

Uno de los arreglos factoriales de mayor uso en la MSR, como alternativa de un modelo de segundo orden, es la estructura de tratamientos que tiene como base un factorial 2^K aumentado, conocido como *diseño central compuesto* (DCC) y cuyo modelo de regresión, en el caso de tener tres factores, se define mediante la expresión 8.2. Este diseño es el que permite evaluar tanto los efectos lineales como los de interacción y de curvatura. Su característica importante es que su construcción se define en experimentos subsecuentes dentro de la MSR, esto es, se corre primero un experimento factorial del tipo 2^K ⁵² con puntos centrales para verificar curvatura; y si ésta existe, se corren tratamientos adicionales (puntos axiales) sin la necesidad de volver a correr los otros tratamientos. Esta forma de desarrollo es lo que le confiere flexibilidad a este diseño y por lo cual es el que más se utiliza en optimización; sin embargo, se requieren condiciones constantes en las corridas secuenciales de los tratamientos, situación que a nivel de la agronomía es muy difícil de conseguir.

Si bien la estructura completa del DCC, bajo la concepción del MSR, se logra en experimentos subsecuentes, este diseño de tratamientos puede ser aprovechado en el área agrícola si se lo aplica de forma simultánea en un mismo experimento, especialmente en las investigaciones sobre nutrientes; de allí que ese sea el propósito de su exposición en este texto. Este es un arreglo factorial que, con las previsiones teóricas que fundamentan la delimitación de espacios experimentales, tiene algunas ventajas que ofrecer a la investigación agrícola.

Entrando en su definición, el número total de tratamientos (N_T) que se generan en este diseño se define mediante la expresión:

⁵¹ La rotabilidad y la ortogonalidad son dos características que brindan mayor eficacia en la precisión de las distribuciones experimentales factoriales que permiten generar superficies de respuesta. La primera asegura modelos que predicen con la misma varianza dentro de un espacio experimental, desde el punto central a cualquier dirección; y la segunda, asegura que los efectos no estén correlacionados.

⁵² Los arreglos factoriales del tipo 2^K , en la MSR, son considerados diseños de primer orden, de tipo exploratorios, que sólo permiten observar efectos lineales. En el caso de notar la necesidad de verificar curvatura en el modelo, se añaden puntos centrales. Además, es el tipo de arreglo base para obtener algunos diseños de segundo orden como el central compuesto.

$$N_t = 2^K + 2K + n_{RC} \quad (8.3)$$

En función de lo mencionado y de lo que representa la expresión 8.3, el DCC está compuesto de una parte factorial (2^K), otra parte de puntos o tratamientos axiales ($2K$) y una última parte de puntos centrales (n_{RC}). En la figura 8.3 se detallan los puntos experimentales para un DCC con 2 y 3 factores; en los cuales, los puntos rojos representan a las estructuras factoriales 2^2 y 2^3 , respectivamente, mientras que las corridas centrales y las axiales están representadas por los puntos azules y verdes, en su orden. Nótese también en esta figura el número de tratamientos que se generan para dos y tres factores: nueve para el primero y 15 para el segundo.

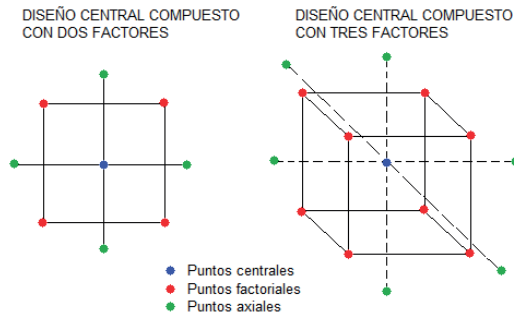


Figura 8.3. Esquema de la matriz de puntos experimentales en un arreglo central compuesto para dos y tres factores.

La precisión y eficiencia de este diseño pueden mejorarse si se hace que cumpla con condiciones de *rotabilidad* y *ortogonalidad*. Al hacer que un diseño de tratamientos sea rotable, eligiendo los puntos experimentales axiales adecuadamente, se logra que la varianza tenga dependencia exclusiva de la distancia del punto central del diseño y no de la dirección con respecto a este punto; propiedad que permite tener una igualdad de varianzas y comparaciones justas dentro del espacio experimental de estudio. Esta rotabilidad permite tener un arreglo eficiente en términos económicos, debido a que es posible estimar el error experimental sólo con unas cuantas réplicas del tratamiento central, dejando al resto de puntos experimentales sólo con una de ellas.

La falta de ortogonalidad en un arreglo de tratamientos puede ocasionar que ciertos efectos no estén bien claros, encontrándose confusos entre ellos; de allí que, un diseño con ausencia de esta propiedad pueda ser muy impreciso. Un diseño de tratamientos ortogonal evita que los coeficientes del modelo de regresión ajustado no estén correlacionados entre sí, de tal manera que el efecto de cada término del mismo se estime con mayor precisión (Gutiérrez y De la Vara, 2008).

Para lograr la rotabilidad de un diseño es necesario que la distancia de los puntos axiales (α) se defina a partir de la expresión siguiente:

$$\alpha = (2^K)^{1/4} \quad (8.4)$$

Por otro lado, para lograr que estos tipos de diseños sean ortogonales, la distancia de los puntos axiales α se debe determinar considerando todos los tratamientos o corridas experimentales (N) y la expresión siguiente:

$$\alpha = \left[\frac{(2^K N)^{1/2} - 2^K}{2} \right]^{1/2} \quad (8.5)$$

La ventaja económica que presenta el DCC sobre los factoriales completos se hace evidente a partir de tres factores a más de dos niveles. Así, por ejemplo, para un factorial del tipo 3^3 se tiene un total de 27 tratamientos; mientras que en un DCC se tendrían sólo 15 combinaciones factoriales. En el caso de utilizar un factorial completo de 4 factores a tres niveles, se tendrá un total de 81 corridas; lo que en un DCC se tendrían sólo 25 tratamientos. Otra de las ventajas de este tipo de diseño sobre los factoriales completos es notable en los niveles de prueba; para el primer ejemplo, cada factor tiene sólo tres niveles, mientras que en un DCC se tendrá un total de 5 niveles, lo que permitirá estimar de mejor manera los efectos de curvatura en la determinación de superficies que contienen a los puntos experimentales óptimos.

Para la aplicación de los DCC en la agronomía habrá que tener la suficiente información de base que permita elegir adecuadamente los niveles cuantitativos que se utilizarían en un experimento, de tal forma que el arreglo permita en un solo experimento, determinar los puntos de ensayo en donde la producción de un cultivo es máxima. Este diseño podría aplicarse, especialmente en ensayos en donde se valoran dosis de fertilizantes; también en algunos otros experimentos en donde se prueben estimulantes, nutrientes foliares o algún otro producto cuyos niveles necesariamente deben ser cuantitativos.

Ejemplo de aplicación 8.2

Se ha realizado un experimento sobre nutrientes en el cultivo de caña de azúcar, utilizando los elementos nitrógeno, fósforo y potasio. Para llevarlo a cabo se ha utilizado un diseño de bloques completos al azar, empleando el arreglo central compuesto como diseño de tratamientos, con el propósito de identificar las dosis en donde el rendimiento es máximo. Los niveles bajo y alto de los fertilizantes evaluados se detallan en la tabla 8.11. Asimismo, los tratamientos y su codificación, así como el registro de la variable de respuesta que en este caso fue el rendimiento del cultivo en t/ha, se detallan en las tablas 8.12 y 8.13.

En la tabla 8.12 se detallan los niveles evaluados por cada uno de los factores, considerando un DCC ajustado al modelo gráfico para tres factores de la figura 8.3. Adicionalmente, en esta tabla también se indican los valores codificados por cada nivel. Los primeros ocho tratamientos corresponden al factorial del tipo 2^3 , mientras que los tratamientos desde el 9 al 14 son los puntos axiales del diseño que se establecen considerando la rotabilidad del mismo mediante la expresión 8.13. La distancia α , desde el punto central a los puntos axiales, es:

$$\alpha = (2^3)^{1/4} = 1,682$$

Los niveles de los puntos axiales se establecen con la expresión siguiente:

$$X_i = \frac{\alpha \cdot \Delta X}{2} + \bar{X} \quad (8.6)$$

Donde X_i es el nivel original de uno de los factores; ΔX y \bar{X} son la diferencia y la media, respectivamente, entre el nivel superior e inferior del factor. Para el nitrógeno, este punto axial tiene los siguientes valores:

$$X_{-N} = \frac{-1,682 \times 80}{2} + 120 \quad X_{+N} = \frac{1,682 \times 80}{2} + 120$$

$$X_{-N} = 52,72 \text{ kg/ha} \quad X_{+N} = 187,28 \text{ kg/ha}$$

Por último, el tratamiento 15 representa al punto central del diseño (0,0,0).

Tabla 8.11. Niveles bajo y alto de los fertilizantes.

Factores	Niveles	
	Bajo (-1)	Alto (+1)
Nitrógeno N	80	160
Fósforo P	60	120
Potasio K	100	200

Tabla 8.12. Arreglo factorial bajo el DCC con sus niveles y sus variables codificadas.

Nº	Variables (factores)			Variables codificadas		
	N	P	K	X_1	X_2	X_3
1	80	60	100	-1	-1	-1
2	160	60	100	+1	-1	-1
3	80	120	100	-1	+1	-1
4	160	120	100	+1	+1	-1
5	80	60	200	-1	-1	+1
6	160	60	200	+1	-1	+1
7	80	120	200	-1	+1	+1
8	160	120	200	+1	+1	+1
9	52,72	90	150	-1,682	0	0
10	187,28	90	150	+1,682	0	0
11	120	39,54	150	0	-1,682	0
12	120	140,46	150	0	+1,682	0
13	120	90	65,9	0	0	-1,682
14	120	90	234,1	0	0	+1,682
15	120	90	150	0	0	0

Una vez definido el diseño de tratamientos, estos fueron corridos mediante tres réplicas en una distribución de bloques completos al azar. Los datos recolectados se indican en la tabla 8.13.

Tabla 8.13. Rendimiento (t/ha) bajo el arreglo DCC.

N°	Variables (factores)			Repeticiones			Promedios
	N	P	K	I	II	III	
1	80	60	100	70,3	69,4	73,5	71,1
2	160	60	100	79,3	83,2	80,2	80,9
3	80	120	100	77,5	79,2	74,3	77,0
4	160	120	100	86,5	88,6	80,2	85,1
5	80	60	200	72,5	74,1	76,8	74,5
6	160	60	200	84,3	77,4	80,2	80,6
7	80	120	200	82,3	85,5	83,1	83,6
8	160	120	200	70,4	80,3	75,4	75,4
9	52,72	90	150	72,5	70,4	74,3	72,4
10	187,28	90	150	69,5	66,2	70,3	68,7
11	120	39,54	150	73,8	70,2	66,9	70,3
12	120	140,46	150	66,3	65,5	67,1	66,3
13	120	90	65,9	66,9	68,9	67,4	67,7
14	120	90	234,1	67,8	64,8	68,4	67,0
15	120	90	150	93,4	98,1	95,6	95,7

La obtención del modelo y el análisis estadístico correspondiente, así como las superficies de respuesta para cada par de factores, se han determinado mediante el software estadístico Minitab. Cabe indicar que el arreglo de los tratamientos también está basado en la forma como lo diseña este software. El proceso de análisis se detalla en los siguientes pasos:

Características de la estructura del diseño central compuesto (DCC) determinado por Minitab:

```
Factores:          3      Réplicas:          3
Corridas base:    15     Total de corridas: 45
Bloques base:     1      Total de bloques:  1
```

Factorial de dos niveles: Factorial completo

```
Puntos del cubo:          24
Puntos centrales en el cubo: 3
Puntos axiales:          18
Puntos centrales en axial: 0
```

Alfa: 1,682

Análisis de regresión del modelo de superficie de respuesta:

Regresión de superficie de respuesta: Rend. (TM/ha vs. Nitrógeno. Fósforo. ...

El análisis se realizó utilizando unidades codificadas.

Coefficientes de regresión estimados de Rend. (TM/ha)

Término	Coef	SE Coef	T	P
Constante	93,5183	3,998	23,389	0,000
Nitrógeno	0,6995	1,088	0,643	0,525
Fósforo	0,5349	1,088	0,492	0,626
Potasio	-0,0879	1,088	-0,081	0,936
Nitrógeno*Nitrógeno	-5,8680	1,634	-3,590	0,001
Fósforo*Fósforo	-6,6574	1,634	-4,073	0,000
Potasio*Potasio	-6,9873	1,634	-4,275	0,000
Nitrógeno*Fósforo	-2,0208	1,422	-1,421	0,164
Nitrógeno*Potasio	-2,5042	1,422	-1,761	0,087
Fósforo*Potasio	-0,7792	1,422	-0,548	0,587

S = 6,96567 PRESS = 2754,62
R-cuad. = 44,37% R-cuad. (pred.) = 9,76% R-cuad. (ajustado) = 30,06%

Análisis de varianza de Rend. (TM/ha)

Fuente	GL	SC Sec.	SC Ajust.	MC Ajust.	F	P
Regresión	9	1354,31	1354,31	150,479	3,10	0,008
Lineal	3	32,09	32,09	10,696	0,22	0,882
cuadrado	3	1059,15	1059,15	353,048	7,28	0,001
interacción	3	263,08	263,08	87,694	1,81	0,164
Error residual	35	1698,22	1698,22	48,521		
Falta de ajuste	5	1479,07	1479,07	295,814	40,50	0,000
Error puro	30	219,15	219,15	7,305		
Total	44	3052,53				

Coefficientes de regresión estimados de Rend. (TM/ha) utilizando datos en unidades sin codificar

Término	Coef
Constante	-133,272
Nitrógeno	1,23706
Fósforo	1,62930
Potasio	1,03372
Nitrógeno*Nitrógeno	-0,00366748
Fósforo*Fósforo	-0,00739709
Potasio*Potasio	-0,00279491
Nitrógeno*Fósforo	-0,00168403
Nitrógeno*Potasio	-0,00125208
Fósforo*Potasio	-5,19444E-04

Según este análisis, el modelo de regresión en niveles de los factores sin codificar, es:

$$Y = -133,27 + 1,24N + 1,63P + 1,03K - 0,004N^2 - 0,007P^2 - 0,003K^2 - 0,001NP - 0,001NK - 0,0005PK$$

Sin embargo, este modelo apenas tiene, según el coeficiente de determinación, un ajuste del 30,06%; calificado desde esta perspectiva como medianamente aceptable para pronosticar. Este resultado también es corroborado por la falta de ajuste indicada en el análisis de varianza

del modelo, en donde puede notarse un valor significativo en esta fuente de variación ($P < 0,05$). Según el mismo análisis de varianza, también se puede observar que los factores tienen efectos cuadráticos y actúan de forma independiente, ya que no registran interacción significativa.

Obtener un punto (tratamiento) óptimo en la experimentación agrícola, en la práctica no tiene una utilidad por los motivos que se indicaron al inicio de este capítulo; sin embargo, ubicar unos rangos puede ser mucho mejor. Con este propósito se han realizado las gráficas de contorno y superficie de respuesta por cada par de factores evaluados.

Se observa en las figuras 8.4 y 8.5, de manera simultánea, que las dosis de nitrógeno para lograr un mayor rendimiento se encuentran entre los 100 y 150 kg/ha (color verde de mayor intensidad); ubicándose un valor máximo en aproximadamente 120 kg/ha. Para el caso del fósforo, según las figuras 8.4 y 8.6, el rango de las dosis en donde se puede obtener un mayor rendimiento se encuentra entre los 75 y 110 kg/ha, aproximadamente; con un máximo que podría estar alrededor de los 90 kg/ha. Finalmente, para lograr con el potasio un mayor rendimiento, de acuerdo a las figuras 8.5 y 8.6, éste debería aplicarse entre los 120 y 180 kg/ha; pudiéndose tener un máximo en aproximadamente los 150 kg/ha.

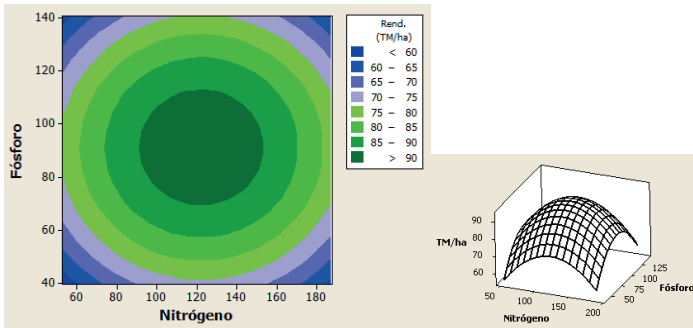


Figura 8.4. Superficie de contorno y de respuesta para los factores nitrógeno y fósforo.

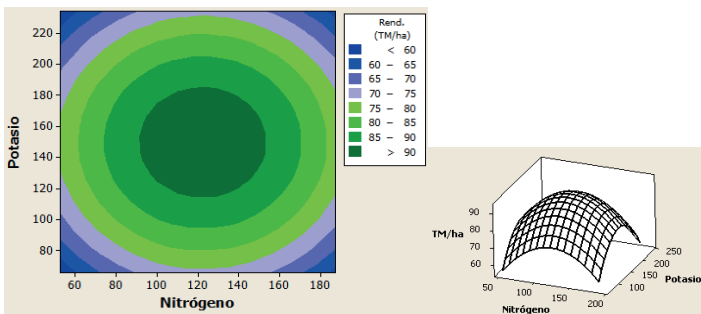


Figura 8.5. Superficie de contorno y de respuesta para los factores nitrógeno y potasio.

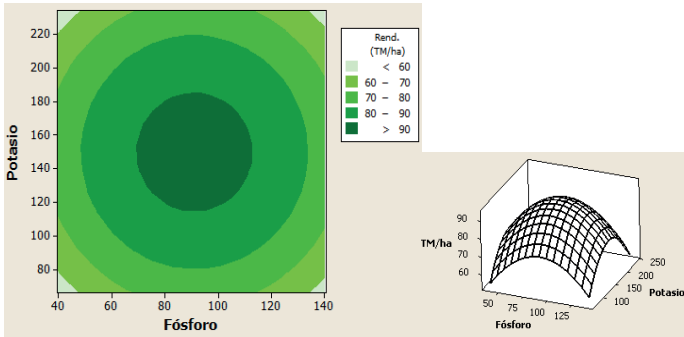


Figura 8.6. Superficie de contorno y de respuesta para los factores fósforo y potasio.

Al optimizar la respuesta para un rendimiento máximo, se puede identificar que la máxima cosecha que podría obtenerse es de 93,5 t/ha. Esto también puede observarse en el análisis realizado en Minitab, con su gráfica correspondiente indicada en la figura 8.7. Puede notarse que este rendimiento máximo ocurre cuando se adicionan 122,0 kg/ha de nitrógeno, 91,5 kg/ha de fósforo y 149,7 kg/ha de potasio; lo que en términos prácticos es la fórmula 120 – 90 – 150, representada por el tratamiento 15 (tabla 8.13).

Optimización de respuesta

Parámetros

	Meta	Inferior	Objetivo	Superior	Ponderación	Importar
Rend. (TM/ha)	Máximo	66,3	100	100	1	1

Solución global

Nitrógeno = 122,039
Fósforo = 91,5291
Potasio = 149,742

Respuestas pronosticadas

Rend. (TM/ha = 93,5490 , conveniencia = 0,808575

Conveniencia compuesta = 0,808575

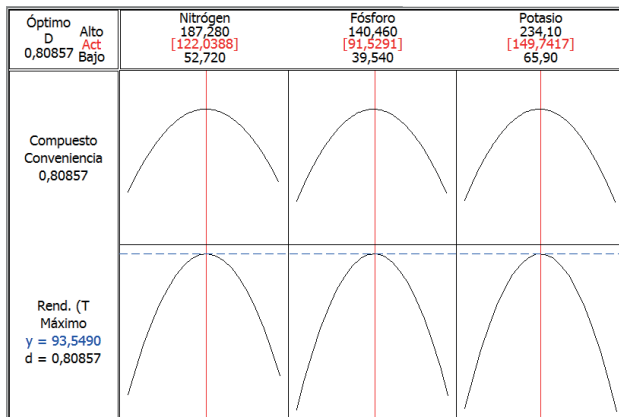


Figura 8.7. Identificación de niveles para los cuales el rendimiento es máximo.

Experimento factorial 2^K con una sola repetición

Cuando se inician estudios con objetivos de optimización, los primeros experimentos que se realizan empiezan seleccionando varios factores de estudio. Con este propósito se recurre a los experimentos factoriales del tipo 2^K ; es decir K factores a dos niveles cada uno (experimentos simétricos). En este contexto, las corridas experimentales se van incrementando exponencialmente conforme vayan aumentando los factores en estudio, lo cual hace que los costos experimentales también se incrementen de manera sustancial; además de ir incrementándose también los niveles de complejidad de los análisis estadísticos respectivos.

Los experimentos del tipo 2^2 y 2^3 producirán cuatro y ocho combinaciones factoriales (tratamientos), respectivamente, con los menores efectos (factores principales e interacciones de éstos) que necesariamente deben valorarse a través de dos o más repeticiones. En estos primeros arreglos, el número de corridas que deberá efectuarse le dan el carácter práctico al experimento en términos de los recursos empleados. Aún los experimentos con cuatro factores, producirán a su vez 16 combinaciones factoriales que, en la posibilidad de las circunstancias, todavía pueden ser prácticos en su aplicación si todo el ensayo se corre con al menos dos repeticiones. Sin embargo, en esta clase de experimentos lo más común es que las interacciones de tres o más factores no presenten efectos significativos, lo que se constituye en una oportunidad para desarrollar un experimento con sólo una réplica; es decir, utilizar estos efectos, que no deben ser significativos, para estimar aproximadamente la varianza del error experimental.

Para poder realizar un análisis de varianza, al menos se necesita que los tratamientos de un experimento se repitan dos veces, de tal forma que se pueda estimar el error experimental; ya que con una sola repetición no es posible su realización y, por lo tanto, tampoco se puede hacer la verificación del azar en los efectos de los tratamientos. Los experimentos con cuatro o más factores generan varios efectos relacionados a las interacciones de tres o más niveles factoriales que, como se ha mencionado en el párrafo anterior, generalmente no presentan significancias de importancia. Por consiguiente, estos efectos, potencialmente irrelevantes, pueden agruparse en el error experimental y obtener así una suma de cuadrados de dicha fuente compuesta de los efectos de todas las interacciones o, en ciertos, casos de los efectos principales; que no reportaron significancia estadística. Asimismo, la totalidad de los grados de libertad que tenga el error experimental se constituirá del total de efectos que, por su poca magnitud significativa, se hayan enviado a formar parte de la suma de cuadrados del error experimental.

La elaboración de un ANDEVA con una varianza del error experimental (cuadrado medio) obtenida a partir de la unión de efectos no significativos produce resultados que deben considerarse con cierta reserva, ya que en la realidad no se han repetido los tratamientos y, consecuentemente, dicha varianza es sólo una estimación aproximada del valor que se obtendría cuando la magnitud del error experimental se logra a partir de n efectos de un mismo tratamiento. Bajo esta concepción, es útil la verificación previa de los efectos que no presentan significancia en forma gráfica, utilizando el *diagrama de Pareto* o el papel para gráficas de normalidad. Lo último es el punto inicial para realizar un análisis de varianza considerando sólo las fuentes de variabilidad cuyos efectos no se deben al azar.

Debido a que la verificación gráfica de los efectos significativos es subjetiva al momento de decidir qué efectos deben ser parte del error experimental, la obtención del mejor ANDEVA también debe fundamentarse en ANDEVAs preliminares que indiquen si un efecto tiene cierta magnitud de significancia como para seguir considerándolo como efecto de importancia. En este sentido, podría considerarse que todos los efectos (léase suma de cuadrados) que tengan un $P - Valor < 0,2$ no deberían ser parte del error experimental; lo que, de forma complementaria, se ha establecido tomando en cuenta el postulado de Pareto a través de su regla "80/20": el 80% del problema es resuelto por el 20% de las causas esenciales. Asimismo, la observancia de los grados de libertad del error experimental es fundamental al momento de realizar un ANDEVA bajo estas circunstancias, para lo cual, es recomendable que al menos se cuente con un mínimo de 10 grados de libertad en esta fuente de variabilidad para que se tenga confianza del análisis.

Otro de los valores estadísticos que permiten tener una idea de los efectos que deberían enviarse al error experimental es el coeficiente de determinación ajustado (r_{aj}^2). Si un efecto se excluye del grupo de los que interesan en el análisis y el valor de r_{aj}^2 aumenta, esto es un indicativo de lo acertado de su exclusión; sin embargo, si este valor disminuye sustancialmente, dicha fuente que se había previsto excluir deberá mantenerse entre las fuentes principales del análisis aun no siendo significativa.

Uno de los riesgos que se corre en estos casos de experimentos es la probabilidad creciente de cometer el error tipo I (el error de mayor costo económico). Pues, como se verá en el ejemplo de aplicación detallado más adelante, el cuadrado medio del error experimental que se obtiene de las fuentes que se envían a formar parte de esta varianza, comparativamente es bastante menor a las fuentes de variabilidad que se analizan en el ANDEVA. Por esto, se corre el riesgo de detectar efectos que, si bien son estadísticamente significativos, no tienen una aplicación realmente práctica. En fin, si principalmente la disponibilidad de recursos económicos no es un problema, siempre será mejor que un ensayo factorial tenga las necesarias repeticiones; pues son sólo éstas las que hacen posible estimar correctamente la influencia del error experimental en cualquier ensayo.

De acuerdo a lo explicado en los párrafos que anteceden en este procedimiento, hay que resaltar primero el requerimiento de un material experimental (especialmente el suelo) que deberá ser, en lo posible, extremadamente homogéneo; si esto no es así, la probabilidad de caer en el error tipo I será inevitablemente alta. Este requerimiento también incluye el manejo uniforme de todas las actividades experimentales, pues si bien no se puede valorar la precisión de un ensayo debido a la ausencia de repeticiones de las alternativas de prueba, un mal manejo también nos llevará directamente a cometer el error tipo I. Como corolario, un experimento con una sola repetición sobre un suelo heterogéneo y/o mal manejado, tendrá como característica una inexactitud que, aunque el contexto de la investigación se haya situado teóricamente dentro del espacio experimental adecuado, nos llevará a una confusión total mucho más catastrófica que la falta de precisión. Nótese también que el requisito insoslayable de homogeneidad para un ensayo de una sola réplica, no permite utilizar una distribución que no sea la completamente al azar.

26	+1	-1	-1	+1	+1	-1	-1	+1	+1	+1	-1	-1	-1	-1	+1	<i>ade</i>
27	-1	+1	-1	+1	+1	-1	+1	-1	-1	-1	+1	+1	-1	-1	+1	<i>bde</i>
28	+1	+1	-1	+1	+1	-1	+1	+1	-1	-1	+1	+1	-1	-1	+1	<i>abde</i>
29	-1	-1	+1	+1	+1	+1	-1	-1	-1	-1	-1	-1	-1	+1	+1	<i>cde</i>
30	+1	-1	+1	+1	+1	-1	+1	+1	+1	-1	-1	-1	+1	+1	+1	<i>acde</i>
31	-1	+1	+1	+1	+1	-1	-1	-1	-1	+1	+1	+1	+1	+1	+1	<i>bcde</i>
32	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	<i>abcde</i>

En la expresión 8.7, Y_{ij} representa al tratamiento i en la repetición (bloque, si es una DBCA) j ; y n ($n = R \times 2^2$) es el total de datos. El número de repeticiones están representadas por R (expresiones 7.8 y 7.9). Nótese además que las tres fórmulas están diseñadas para aplicarse exclusivamente a combinaciones factoriales del tipo 2^5 . Igualmente, los efectos aproximados de factores e interacciones pueden calcularse con las siguientes expresiones:

$$EFECTO K = (CONTRASTE K)/(2^{K-1}R) \quad (8.10)$$

$$EFECTO K_1K_2 \dots K_n = (CONTRASTE K_1K_2 \dots K_n)/(2^{K-1}R) \quad (8.11)$$

Experimentos 2^K con una sola repetición en la agronomía

A los resultados poco prácticos que se pueden obtener para la agronomía de metodologías de optimización como la de superficie de respuesta, llevar a cabo estos ensayos para optimizar las actividades de la agricultura involucra también una cantidad de recursos económicos importantes que en muchos casos restringen su aplicación. Sin embargo, cuando se pretende evaluar varios factores para identificar exploratoriamente los que inciden de forma significativa en la producción de un cultivo, pueden desarrollarse experimentos con una sola réplica utilizando un factorial del tipo 2^K (aunque en la práctica también podrían emplearse más de dos niveles por factor). En forma general, puede decirse que todo proceso investigativo que busca eficiencia en las alternativas de prueba, las cuales están relacionadas directamente con los resultados productivos, se inicia con una selección de múltiples factores que luego se los va excluyendo hasta obtener los niveles idóneos que son determinantes en la producción.

Antes de seguir con el ejemplo que a continuación se explica, es preciso recordar la distinción entre lo que es una combinación factorial y el modelo que se genera en el ANDEVA. En las combinaciones factoriales, cada nivel de un factor se combina con cada nivel de otro (s) factor para formar lo que en el lenguaje del diseño experimental se conoce como tratamiento o corrida experimental. Esto quiere decir que, un experimento factorial del tipo 2^2 tendrá 4 combinaciones factoriales, o lo que es lo mismo, 4 tratamientos; consecuentemente, un experimento del tipo 2^3 tendrá 8 tratamientos; por lo que el número de tratamientos se incrementará exponencialmente conforme aumenten los factores.

Esta generación de tratamientos producirá una serie de efectos, principales y secundarios, que crecerán también conforme aumenten las combinaciones factoriales. Éstos siempre serán uno menos que las combinaciones factoriales que genere uno de estos experimentos de tipo simétricos⁵³. Así, por ejemplo, un factorial del tipo 2^2 tendrá tres efectos que deberán ser evaluados: los dos efectos principales (los dos factores) y una interacción doble. Un factorial del tipo 2^3 , tendrá tres efectos principales (tres factores), tres interacciones dobles y una interacción triple. Los efectos que se tendrán que evaluar en el primer caso serán 3, mientras

⁵³ Un experimento factorial es *simétrico* cuando tienen la misma cantidad de niveles en cada uno de los factores; para lo cual éstos se representan como una potencia, en donde la base representa a los niveles y el exponente se refiere a la cantidad de factores. Cuando se tienen distintos niveles, el factorial se conoce como *asimétrico* y su representación es detallarlo como un producto indicado.

que en el segundo caso serán 7. En forma análoga, todo esto se repite para factoriales con más niveles y/o factores. En la tabla 8.15 se describen los tratamientos, los efectos y las interacciones generadas para un experimento del tipo 2^K hasta un máximo de $K = 5$ factores.

Tabla 8.15. Número de tratamientos (combinaciones factoriales), efectos principales y efectos secundarios generados por experimentos factoriales del tipo 2^K , hasta $K = 5$ factores.

Factorial	Tratamientos generados (combinaciones)	Efectos principales (factores)	Efectos secundarios (Interacciones)			
			Doble	Triple	Cuádruples	Quíntuples
2^2	4	2: A y B	1: AB			
2^3	8	3: A, B y C	3: AB, AC y BC	1: ABC		
2^4	16	4: A, B, C y D	6: AB, AC, AD, BC, BD y CD	4: ABC, ABD, ACD y BCD	1: ABCD	
2^5	32	5: A, B, C, D y E	10: AB, AC, AD, AE, BC, BD, BE, CD, CE y DE	10: ABC, ABD, ABE, ACD, ACE, ADE, BCD, BCE, BDE y CDE	5: ABCD, ABCE, ABDE, ACDE y BCDE	1: ABCDE

Ejemplo de aplicación 8.3

Se ha realizado un experimento factorial del tipo 2^5 con el propósito de establecer los factores y niveles que son importantes dentro de la productividad del cultivo de arroz con aplicación de riego. El área experimental es relativamente plana y sin factores secundarios evidentes que podrían necesitar ser bloqueados, lo cual constituyó el fundamento para haber optado por una distribución completamente al azar; además de considerar una sola repetición por restricciones de área y de tipo económicas. Los factores y niveles evaluados son los que se detallan a continuación:

Factor	Descripción	Nivel 1	Nivel 2
A	Preparación de suelo	a_1 : Sin fangueo	a_2 : Con fangueo
B	Tipo de siembra	b_1 : Boleo	b_2 : Trasplante
C	Lámina de riego	c_1 : Sólo a capacidad de campo	c_2 : Con lámina de 10 cm
D	Pre-germinación	d_1 : Semilla pre-germinada	d_2 : Semilla sin pre-germinación
E	Encalado	e_1 : Sin encalado previo	e_2 : Con encalado previo

Los resultados experimentales del rendimiento del cultivo, en unidades de kg/parcela, son los que se describen en la columna 3 de la tabla 8.16, los cuales son el producto de sólo una

repetición. En esta misma tabla se han detallado los contrastes, los efectos aproximados y las sumas de cuadrados para cada factor e interacción doble.

Los contrastes de cada factor e interacción doble que se describen en la tabla 8.16 han sido obtenidos de la suma algebraica de los valores correspondientes a la notación de Yates que se detalla en la tabla 8.14, considerando los signos indicados en esta última tabla para cada caso específico. Así, por ejemplo, para el factor A, el contraste respectivo es:

$$CONTRAS.A = [-(1) + a - b + \dots + abcde] = (-4,5 + 5,6 - 14,3 + \dots + 33,2) = 40,7$$

Los efectos se han estimado mediante las expresiones 8.10 y 8.11. De la misma forma, las sumas de cuadrados para cada factor e interacción doble se han obtenido con las expresiones 8.8 y 8.9. Como ejemplo, para el caso del factor A, se tiene:

$$EFECTO A = (CONTRASTE A)/(2^{K-1}R) = 40,7/(2^4 \times 1) = 2,54$$

$$SC_A = (CONTRASTE A)^2/R2^5 = 40,7^2/(1 \times 32) = 51,77$$

El análisis de varianza que inicialmente se realiza, sólo tomará en consideración los efectos principales de los factores y sus respectivas interacciones dobles; ya que se asume que las interacciones de un mayor grado no presentan significancias de importancia. Esta consideración hace posible matemáticamente realizar este análisis; pues, de lo contrario, si en este tipo de combinación factorial (2^5) se asumen que todos los 31 efectos son importantes no habrá grados de libertad para el error experimental y, consecuentemente, no podrá realizarse el ANDEVA. Bajo esta concepción, todos los efectos que desde el ANDEVA inicial se los identifican como no significativos formarán parte del error experimental estimado. Por lo tanto, para llegar al ANDEVA que mejor interpretación estadística realice con los datos, el proceso puede circunscribirse a los siguientes tres pasos:

1.- Elaboración de un ANDEVA inicial, asumiendo que las interacciones triples y de mayor grado no son significativas, por lo que estas últimas serán las fuentes que inicialmente constituirán el error experimental estimado.

Tabla 8.16. Contrastes, efectos y sumas de cuadrados para cada factor e interacción doble.

Nº	Notación de Yates	Y _i	A	B	C	D	E	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE
1	(1)	15,7	-15,7	-15,7	-15,7	-15,7	-15,7	15,7	15,7	15,7	15,7	15,7	15,7	15,7	15,7	15,7	15,7
2	a	5,6	5,6	-5,6	-5,6	-5,6	-5,6	-5,6	-5,6	-5,6	-5,6	5,6	5,6	5,6	5,6	5,6	5,6
3	b	7,8	-7,8	7,8	-7,8	-7,8	-7,8	7,8	7,8	-7,8	-7,8	-7,8	-7,8	-7,8	-7,8	-7,8	-7,8
4	ab	12,6	12,6	12,6	-12,6	-12,6	12,6	12,6	-12,6	-12,6	-12,6	12,6	12,6	-12,6	-12,6	12,6	12,6
5	c	18,5	-18,5	18,5	-18,5	18,5	-18,5	18,5	18,5	18,5	18,5	-18,5	-18,5	18,5	-18,5	-18,5	18,5
6	ac	18,5	18,5	-18,5	-18,5	18,5	-18,5	-18,5	18,5	18,5	-18,5	-18,5	18,5	18,5	-18,5	-18,5	18,5
7	bc	13,5	-13,5	13,5	-13,5	13,5	-13,5	13,5	13,5	13,5	13,5	-13,5	-13,5	13,5	-13,5	-13,5	13,5
8	abc	35,7	35,7	35,7	-35,7	-35,7	35,7	35,7	-35,7	-35,7	-35,7	35,7	35,7	-35,7	-35,7	35,7	35,7
9	d	4,3	-4,3	-4,3	4,3	4,3	4,3	-4,3	-4,3	-4,3	-4,3	4,3	4,3	4,3	4,3	4,3	-4,3
10	ad	13,9	13,9	-13,9	-13,9	13,9	-13,9	-13,9	13,9	13,9	13,9	-13,9	-13,9	13,9	-13,9	13,9	-13,9
11	bd	22,4	-22,4	22,4	-22,4	22,4	-22,4	-22,4	22,4	-22,4	-22,4	22,4	22,4	-22,4	-22,4	22,4	-22,4
12	abd	25,7	25,7	-25,7	-25,7	25,7	25,7	-25,7	-25,7	25,7	25,7	-25,7	-25,7	25,7	-25,7	-25,7	25,7
13	cd	13,5	-13,5	13,5	-13,5	13,5	-13,5	13,5	13,5	13,5	13,5	-13,5	-13,5	13,5	-13,5	-13,5	13,5
14	acd	34,4	34,4	-34,4	34,4	-34,4	-34,4	34,4	34,4	34,4	-34,4	-34,4	34,4	34,4	-34,4	-34,4	34,4
15	bcd	25,7	-25,7	25,7	-25,7	25,7	-25,7	-25,7	25,7	-25,7	-25,7	25,7	25,7	-25,7	-25,7	25,7	-25,7
16	abcd	35,8	35,8	35,8	-35,8	-35,8	35,8	35,8	-35,8	-35,8	-35,8	35,8	35,8	-35,8	-35,8	35,8	-35,8
17	e	4,1	-4,1	-4,1	4,1	4,1	4,1	-4,1	-4,1	-4,1	-4,1	4,1	4,1	4,1	4,1	4,1	-4,1
18	ae	13,8	13,8	-13,8	-13,8	13,8	-13,8	-13,8	13,8	13,8	13,8	-13,8	-13,8	13,8	-13,8	-13,8	13,8
19	be	21,5	-21,5	21,5	-21,5	21,5	-21,5	21,5	21,5	21,5	-21,5	-21,5	21,5	21,5	-21,5	-21,5	21,5
20	abe	11,5	11,5	-11,5	-11,5	11,5	11,5	-11,5	-11,5	-11,5	-11,5	11,5	11,5	-11,5	-11,5	11,5	-11,5
21	ce	5,8	-5,8	5,8	-5,8	5,8	-5,8	5,8	5,8	5,8	5,8	-5,8	-5,8	5,8	-5,8	-5,8	5,8
22	ace	19,5	19,5	-19,5	19,5	-19,5	19,5	-19,5	19,5	19,5	-19,5	-19,5	19,5	19,5	-19,5	-19,5	19,5
23	bce	15,7	-15,7	15,7	-15,7	15,7	-15,7	-15,7	15,7	15,7	15,7	-15,7	-15,7	15,7	-15,7	-15,7	15,7
24	abce	33,5	33,5	33,5	-33,5	-33,5	33,5	33,5	-33,5	-33,5	-33,5	33,5	33,5	-33,5	-33,5	33,5	-33,5
25	de	5,8	-5,8	5,8	-5,8	5,8	-5,8	5,8	5,8	5,8	5,8	-5,8	-5,8	5,8	-5,8	-5,8	5,8
26	ade	13,6	13,6	-13,6	-13,6	13,6	-13,6	-13,6	13,6	13,6	13,6	-13,6	-13,6	13,6	-13,6	-13,6	13,6
27	bde	23,7	-23,7	23,7	-23,7	23,7	-23,7	-23,7	23,7	-23,7	-23,7	23,7	23,7	-23,7	-23,7	23,7	-23,7
28	abde	25,4	25,4	-25,4	25,4	-25,4	25,4	-25,4	-25,4	25,4	25,4	-25,4	-25,4	25,4	-25,4	-25,4	25,4
29	cde	15,7	-15,7	15,7	-15,7	15,7	-15,7	15,7	15,7	15,7	15,7	-15,7	-15,7	15,7	-15,7	-15,7	15,7
30	acde	35,8	35,8	-35,8	35,8	-35,8	35,8	-35,8	35,8	35,8	-35,8	-35,8	35,8	35,8	-35,8	-35,8	35,8
31	bcde	25,8	-25,8	25,8	-25,8	25,8	-25,8	-25,8	25,8	-25,8	-25,8	25,8	25,8	-25,8	-25,8	25,8	-25,8
32	abcde	37,9	37,9	37,9	-37,9	-37,9	37,9	37,9	-37,9	-37,9	-37,9	37,9	37,9	-37,9	-37,9	37,9	-37,9
Contrastes		133,7	135,7	157,9	106,1	5,5	9,7	127,1	74,5	12,1	-111,9	26,1	21,7	17,3	-17,3	10,5	
Efectos		8,36	8,48	9,87	6,63	0,34	-0,61	7,94	4,66	0,76	-0,74	2,19	1,63	1,36	-1,08	0,66	
Suma de cuadrados		558,62	575,45	779,14	351,79	0,95	2,94	504,83	173,45	4,58	4,43	38,50	21,29	14,72	9,35	3,45	
Total		612,7															

2.- Realizar un nuevo tamizado de los efectos que resulten significativos en el ANDEVA inicial mediante el análisis gráfico de Pareto y su regla del 80/20. En esta gráfica se podrá respaldar lo resuelto por el ANDEVA inicial y realizar una selección definitiva de las fuentes con efectos significativos y de las demás fuentes de variación que deberán ir al error experimental.

3.- Finalmente se realiza el ANDEVA definitivo, segregando las fuentes de variación significativas y estableciendo la magnitud del error experimental estimado.

Posteriormente se verifica cuánto influyen en los resultados las fuentes seleccionadas a través del coeficiente de determinación ajustado (r_{aj}^2). Asimismo, se podrá establecer un modelo explicativo de regresión con las fuentes de variación que resultaron importantes.

Paso 1

Para el ejemplo que se está describiendo, el proceso para llegar al ANDEVA inicial es el siguiente:

La suma de cuadrados del total, de acuerdo a la expresión 8.7 y a los datos de la columna 3 de la tabla 8.16, es la siguiente:

$$SC_T = (15,7^2 + 5,6^2 + 7,8^2 + \dots + 37,9^2) - 612,7^2/32$$

$$SC_T = 3229,10$$

El resto de sumas de cuadrados, tanto de los factores independientes como de las interacciones dobles, se detallan en la tabla 8.16. Los grados de libertad correspondientes para cada una de estas fuentes de variación, debido a que son integrantes de los grados de libertad de tratamientos, es uno para cada uno. Por lo tanto, el ANDEVA preliminar es el que se indica en la tabla 8.17.

Tabla 8.17. Análisis de varianza inicial.

Fuentes de variación	GL	SC	CM	F	P-valor
Total	31	3229,10			
Factor A (Preparación de suelo)	1	558,62	558,62	48,14	0,000
Factor B (Tipo de siembra)	1	575,45	575,45	49,60	0,000
Factor C (Lámina de riego)	1	779,14	779,14	67,15	0,000
Factor D (Pre-germinación)	1	351,79	351,79	30,32	0,000
Factor E (Encalado)	1	0,95	0,95	0,08	0,779
Interacción AB	1	2,94	2,94	0,25	0,622
Interacción AC	1	504,83	504,83	43,51	0,000
Interacción AD	1	173,45	173,45	14,95	0,001
Interacción AE	1	4,58	4,58	0,39	0,539
Interacción BC	1	4,43	4,43	0,38	0,546
Interacción BD	1	38,50	38,50	3,32	0,087
Interacción BE	1	21,29	21,29	1,83	0,194
Interacción CD	1	14,72	14,72	1,27	0,277
Interacción CE	1	9,35	9,35	0,81	0,383
Interacción DE	1	3,45	3,45	0,30	0,593
Error experimental	16	185,65	11,60		

Según la tabla 8.17, en el ANDEVA inicial puede notarse que la consideración de que las 16 fuentes de variación, compuestas de las interacciones triples y demás grados que no se han

detallado, forman parte de un error experimental estimado que es relativamente pequeño. Esto hace suponer que cualquier interacción inmersa en este efecto, como valor promedio, sería no significativo; todo lo cual constituye la condición necesaria, a partir de la cual, se estima que dicho ANDEVA es el adecuado. Asimismo, de acuerdo a este ANDEVA preliminar, existen cuatro efectos principales y dos interacciones dobles bien definidas que destacan en su significancia ($p < 0,05$) respecto del resto de fuentes de variación, situación que hace prever que el análisis más acertado tendrá sólo seis fuentes de variación. Este resultado está respaldado por la magnitud de los efectos que las mismas fuentes identificadas en el ANDEVA inicial muestran en la tabla 8.16.

Paso 2

Una vez que se ha realizado la verificación preliminar de las fuentes de variación principal y secundarias que definitivamente deberán integrar el error experimental, se hace necesario tener un respaldo adicional que indique si las fuentes que se han seleccionado en el ANDEVA preliminar como integrantes del error experimental son las que realmente deberán estar. Este respaldo puede estar en el postulado de Pareto, el cual manifiesta que el 80% de un problema es resuelto tratando el 20% de las causas consideradas esenciales. Bajo este postulado, podríamos decir que el 80% de la variabilidad que ha reportado el ANDEVA inicial se corresponde aproximadamente con el 20% de las fuentes de variación analizadas.

En la figura 8.8 se indica el postulado de Pareto gráficamente para este ejemplo. En este gráfico son notorios los efectos que muestran las mismas fuentes de variación que el ANDEVA inicial identificó como significativas. Esto es lo que se puede observar en el gráfico de Pareto al separar las seis fuentes de variabilidad (A, B, C, D, AC y AD) mediante líneas punteadas que permiten declararlas responsables del 80% de variación.

Paso 3

Finalmente, luego de realizar una identificación plena de los efectos que muestran una significancia contundente, ya se puede elaborar el ANDEVA definitivo o más adecuado a la variabilidad encontrada. Este análisis es el que se describe en la tabla 8.18, en el cual los efectos que resultaron no significativos en el primer paso se han adicionado al error experimental (sumas de cuadrados) así como sus respectivos grados de libertad.

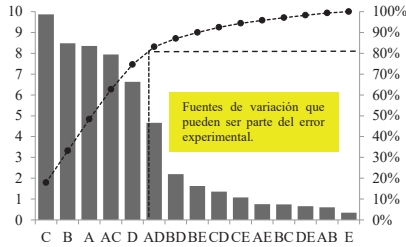


Figura 8.8. Indicación del postulado "80/20" de Pareto.

En el ANDEVA definitivo (tabla 8.18), la variación que explican las fuentes que finalmente se establecieron como significativas tienen un coeficiente de determinación ajustado muy importante ($r_{aj}^2 = 0,89$) que, a su vez, también respalda lo resuelto en este análisis.

Tabla 8.18. Análisis de varianza definitivo.

Fuentes de variación	GL	SC	CM	F	P-valor
Total	31	3229,10			
Factor A (Preparación de suelo)	1	558,62	558,62	48,85	0,000
Factor B (Tipo de siembra)	1	575,45	575,45	50,33	0,000
Factor C (Lámina de riego)	1	779,14	779,14	68,14	0,000
Factor D (Pre-germinación)	1	351,79	351,79	30,77	0,000
Interacción AC	1	504,83	504,83	44,15	0,000
Interacción AD	1	173,45	173,45	15,17	0,001
Error experimental	25	285,86	11,43		

$$r_{aj}^2 = 0,89$$

Selección de factores

El análisis de varianza final que se ha realizado ha permitido establecer los factores que tienen una influencia determinante en la productividad del cultivo, así como las interacciones que también inciden de manera significativa en los resultados. Lo que sigue es establecer los niveles de estos factores que son los que ofrecen las mejores respuestas. Hay que verificar el efecto independiente principal del factor B; y los niveles de los factores A, C y D deben verificarse tomando en consideración las interacciones dobles AC y AD.

En la tabla 8.19 se detallan los promedios de cada nivel factorial. En esta tabla puede notarse que, a excepción del factor E, los niveles codificados con +1 son los que han resultado con los rendimientos sustancialmente diferentes respecto de los codificados con -1. También se observa que, entre los niveles del factor E, no hay diferencias significativas; es decir, es evidente que el encalamiento no es necesario ya que el rendimiento obtenido es bastante similar al rendimiento logrado sin encalar el suelo.

Tabla 8.19. Promedios del rendimiento (kg/Parcela) por cada nivel factorial.

Factores	Promedios	
	- 1	+ 1
Factor A (Preparación de suelo)	15,0	23,3
Factor B (Tipo de siembra)	14,9	23,4
Factor C (Lámina de riego)	14,2	24,1
Factor D (Pre-germinación)	15,8	22,5
Factor E (Encalado)	19,0	19,3

Dado que el factor B (Tipo de siembra) fue el único que presentó un efecto significativo independiente, éste debe seleccionarse de acuerdo al nivel que mostró el mayor promedio en el rendimiento del cultivo. Esto es lo que se describe gráficamente en la figura 8.9, en la cual se muestra claramente que el mayor rendimiento se obtiene realizando la siembra por trasplante.

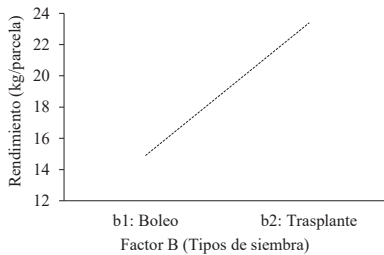


Figura 8.9. Promedios registrados por cada nivel del factor B.

En la figura 8.10 se describen gráficamente los promedios de los efectos producto de la interacción factorial AC. Es notorio que el mayor promedio se logra en la combinación de a_2c_2 ; es decir, realizando fangueo y manteniendo una lámina de agua de 10 cm. Asimismo, en la figura 8.11 se hace evidente que el mayor promedio se obtiene con la combinación factorial a_2d_2 ; o lo que es lo mismo, el mayor efecto se adquiere con fangueo y con semilla sin pre-germinación.

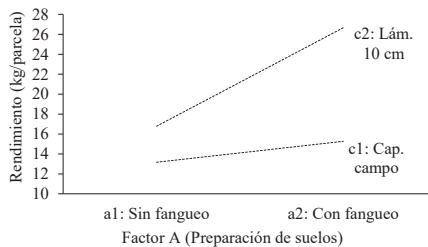


Figura 8.10. Promedios registrados por la interacción de los factores A y C.

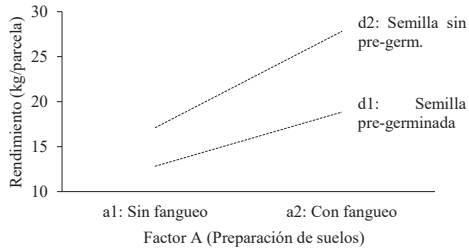


Figura 8.11. Promedios registrados por la interacción de los factores A y D.

Como una conclusión de este experimento y considerando los cinco factores evaluados, se puede decir que el mayor promedio en el rendimiento del cultivo de arroz se define con la combinación factorial $a_2b_2c_2d_2e_1$. En términos de los niveles, esto quiere decir que el mayor promedio se obtiene realizando fanguero, sembrando por trasplante, manteniendo una lámina de 10 cm, utilizando semilla sin pre-germinación y sin necesidad de encalar el suelo.

Experimentos factoriales fraccionados

Otra alternativa en la búsqueda exploratoria de los diseños experimentales agrícolas, cuando se requiere ir identificando los factores determinantes de un proceso productivo, es la realización de experimentos con varios factores de estudio, pero evaluándolos parcialmente. Es decir, sólo se estudia una fracción del experimento factorial debido a los motivos indicados en capítulos y párrafos precedentes sobre lo complicado que es llevar a cabo experimentos con múltiples factores en la agronomía, específicamente en aquellos en donde se planea tener más de tres factores de estudio. Estos ensayos comúnmente se conocen como factoriales fraccionados y se simbolizan como 2^{K-P} , siguiendo la base tradicional en donde el 2 son los niveles, K son los factores y P es el fraccionamiento que se ha realizado en el diseño de tratamientos. Este fraccionamiento puede ser a la mitad, a la cuarta parte, a la octava parte o en cualquier fraccionamiento en múltiplos de 2; todo lo cual depende del número de factores, de los efectos que se quieran evaluar y fundamentalmente de los costos que conlleva la experiencia.

Generalmente estos experimentos se emplean al inicio de todo un proceso investigativo, ya que después de éstos se siguen desarrollando experimentos subsiguientes con los factores que pudieron ser detectados como importantes en el factorial fraccionado; de allí que comúnmente se utilicen dos niveles por cada factor. También están indicados para empezar experimentos exploratorios que permitan identificar ciertos factores que son decisivos en un proceso de producción; para lo cual, generalmente se hace necesario valorar los efectos principales y las interacciones dobles, ya que se consideran de mayor importancia.

La dificultad que hace que se utilice un experimento factorial completo tiene que ver con el incremento sustancial de tratamientos que se tiene cuando se decide evaluar múltiples factores. Por ejemplo, al tener un experimento 2^K con $K = 5$ factores, se tendrán $2^5 = 32$ tratamientos; o el caso de $K = 6$ factores, se tendrán 64 tratamientos que harían muy complicado su

realización debido a los altos costos que demandarían y a la heterogeneidad del suelo que complicarían la precisión. Además, con múltiples factores se hace compleja la interpretación por el número importante de interacciones que se generan; las cuales, sin embargo, son innecesarias de evaluarlas por que comúnmente las interacciones de más de dos factores no influyen significativamente en los resultados.

La justificación de preferir la realización de un factorial del tipo 2^K fraccionado, tal como se lo ha indicado, reside en que los efectos tienen generalmente un grado descendente en importancia significativa. Esto es, los efectos de los factores principales generalmente son los más importantes; luego siguen las interacciones dobles, las triples y así sucesivamente. Esta situación que es de ocurrencia común en los experimentos factoriales, permite la realización experimental incompleta, ya que existen efectos que prácticamente pueden ignorarse por su escasa significancia; sin embargo, también están los efectos que no pueden dejar de tomarse en cuenta debido a su trascendental implicancia en una determinada respuesta.

Una consideración adicional e importante que de alguna manera justifica el uso de los experimentos fraccionados, tiene que ver con el principio de Pareto sobre el que se apoyan, dado que se ajustan muy bien a este principio cuando se tienen varios factores. Es decir, en un ensayo multifactorial, sólo unos pocos factores explican la mayor variabilidad de la respuesta, con lo cual se hace fácil la identificación de los mismos mediante un diagrama de Pareto.

Como ya se lo ha manifestado en el caso de experimentos con múltiples factores, las interacciones de tres o más de ellos pierden importancia por la poca significancia que muchas veces presentan; de allí que esta característica favorezca el desarrollo de experimentos factoriales en los cuales no se consideran la totalidad de las combinaciones. Los efectos que pueden obviarse en un experimento factorial del tipo 2^K tienen una relación proporcional con el incremento de factores; por ejemplo, experimentos como el 2^2 no presenta ningún efecto que pueda ignorarse, mientras que experimentos como el 2^5 tienen potencialmente 16 efectos que pueden despreciarse. En la tabla 8.20 se describen los efectos de los experimentos desde el 2^2 hasta el 2^7 , en donde se detallan tanto los que son de importancia y que obligadamente no pueden despreciarse, y los efectos que bien pueden ignorarse.

Tabla 8.20. Cuantificación de los efectos producidos por varios experimentos del tipo 2^K .

Experimento factorial	Efectos totales	Efectos principales y de las interacciones dobles (no despreciables)	Efectos de las interacciones de más de dos factores (despreciables)
2^2	3	3	0
2^3	7	6	1
2^4	15	10	5
2^5	31	15	16
2^6	63	21	42
2^7	127	28	99

Fuente: Gutiérrez y De La Vara (2005).

Conforme el número de factores se incrementan también los efectos de menor importancia lo hacen. Esta información es la que puede notarse en la tabla 8.20, en la cual se evidencia el hecho de que a partir del factorial 2^3 los efectos que podrían despreciarse son muchos más que los que no deben ignorarse. Por lo tanto, es deducible que experimentos con 5 o más factores puedan fraccionarse a la mitad o a una cantidad menor a esta parte sin que exista el riesgo de obviar información de importancia.

Factorial 2^K fraccionado a un medio

Los experimentos factoriales fraccionados a la mitad tienen la simbología de 2^{K-1} , ya que $2^K/2 = 2^{K-1}$. De estos, el primer factorial que permite un fraccionamiento de las corridas totales es el de tipo 2^3 , el cual fraccionado a la mitad y simbolizado como 2^{3-1} , permite evaluar sólo cuatro tratamientos de los ocho posibles que se pueden generar. El problema que surge en este tipo de arreglo experimental, dado que existen sólo 8 posibilidades (tratamientos), es la selección de los cuatro tratamientos que posibiliten detectar los efectos de mayor importancia.

Debido a que en un factorial fraccionado 2^3 se tiene sólo un efecto despreciable (tabla 8.20) correspondiente a la interacción triple ABC, las fracciones se generan en base a esta interacción; de aquí que sólo existen dos opciones de selección que permiten potencialmente identificar los efectos principales. Estas opciones se indican en la figura 8.12, en donde se hace la distinción en función de los signos de la interacción triple ABC; identificándose como **fracción principal** (color amarillo) cuando esta interacción lleva signos positivos (+) y **fracción complementaria** (color verde) cuando los signos son negativos (-).

COMBINACIONES		A	B	C	EFECTOS			
		A	B	C	AB	AC	BC	ABC
Fracción principal	a	+1	-1	-1	-1	-1	+1	+1
	b	-1	+1	-1	-1	+1	-1	+1
	c	-1	-1	+1	+1	-1	-1	+1
	abc	+1	+1	+1	+1	+1	+1	+1
Fracción complementaria	ab	+1	+1	-1	+1	-1	-1	-1
	ac	+1	-1	+1	-1	+1	-1	-1
	bc	-1	+1	+1	+1	-1	-1	-1
	(1)	-1	-1	-1	+1	+1	+1	+1

Figura 8.12. Esquema de un diseño factorial 2^3 fraccionado a la mitad.

En la fracción principal de la figura 8.12 se observa que la interacción ABC tiene únicamente signos positivos, mientras que en la fracción complementaria dicha interacción es negativa; por consiguiente, el efecto ABC no tiene contraste y más bien se confundirá entre todos los datos. Además, dado que el efecto ABC define las fracciones del factorial por sus signos, se lo conoce como generador de dichas fracciones.

Consideración del efecto alias en los experimentos factoriales fraccionados

La realización de experimentos factoriales fraccionados, si bien permiten un ahorro económico en su realización por cuanto en ellos no se corren todos los tratamientos, tienen la desventaja de no permitir detectar el efecto de las interacciones de forma clara. En el caso del factorial fraccionado 2^{3-1} se hace imposible observar el efecto de la interacción triple, mientras que los efectos de las interacciones dobles constituyen los mismos efectos principales de los factores. Esto último puede notarse en la figura 8.12, en donde el efecto de A se estima también a través del efecto de BC, es decir un mismo efecto es estimado por las fuentes A y BC; luego estos

efectos están aliados y se puede decir que A es alias de BC . Por lo tanto, desde la figura 8.12 se tienen los siguientes efectos alias:

$$\begin{aligned} A \text{ es alias de } BC: & \quad \text{Contraste } A = \text{Contraste } BC = a - b - c + abc \\ B \text{ es alias de } AC: & \quad \text{Contraste } B = \text{Contraste } AC = -a + b - c + abc \\ C \text{ es alias de } AB: & \quad \text{Contraste } C = \text{Contraste } AB = -a - b + c + abc \end{aligned}$$

Dado que un mismo efecto se estima con los efectos alias, puede decirse que dicho efecto es la suma de los efectos $A + BC$, de $B + AC$ y de $C + AB$ y no se puede conocer específicamente quien produce el efecto principal entre cada aliados o si los dos están actuando con la misma intensidad sobre el efecto. Esta situación es la que limita para que un experimento factorial del tipo 2^3 se lo desarrolle fraccionado a un medio, ya que sólo podrían conocerse (asumiéndolo que es así) los efectos principales y no las interacciones dobles, lo cual no es confiable experimentalmente.

Según la figura 8.12, la fracción principal está definida por $+ABC$, debido a ello se refiere su relación generadora como $I = +ABC$. Asimismo para la fracción complementaria, la relación generadora está dada por $I = -ABC$.

Si la decisión en las corridas experimentales es por la fracción complementaria, en donde la relación generadora está dada por $I = -ABC$, los efectos alias según puede observarse en la figura 8.12, están dados por: $A - BC$, $B - AC$ y $C - AB$. Esto permite deducir que, en un factorial fraccionado del tipo 2^{3-1} , un efecto principal de un factor está aliado con el efecto de interacciones dobles; por lo tanto, es indiferente la fracción que se seleccione ya que las dos estiman resultados similares.

Cómo se ha visto, en un factorial fraccionado a la mitad del tipo 2^{3-1} se complica poder identificar a quién pertenece un efecto importante entre los aliados de efecto principal (factor) e interacción doble. La única forma de separar sus efectos es corriendo completamente dicho experimento; esto es, si el investigador primero corrió la fracción principal (con generador $I = +ABC$), deberá correr también la fracción complementaria (con generador $I = -ABC$). Por ejemplo, corriendo las dos fracciones, los efectos independientes de A y de la interacción doble BC se define de la forma siguiente:

$$A = \frac{(A + BC) + (A - BC)}{2} \quad (8.12)$$

$$BC = \frac{(A + BC) - (A - BC)}{2} \quad (8.13)$$

En forma análoga también pueden determinarse independientemente los efectos de B , C , AC y AB .

Una vez corridas las dos fracciones, otra opción es analizar el experimento como si se hubiese realizado completamente desde el inicio; para ello, como es obvio, las condiciones en las que se desarrollen las dos fracciones del experimento deben ser homogéneas. Si esta condición no puede cumplirse y dado que el diseño básico es un completamente al azar, será necesario bloquear el experimento confundiendo los con el efecto de la interacción triple ABC . Esto permitirá tener los efectos independientes del factor principal y de las interacciones dobles por separado.

Como una regla general; se considera la interacción de mayor orden de un factorial, para definir el número de corridas, debido a que se supone que ésta tiene el menor efecto como relación generadora de la fracción experimental. Así, por ejemplo, un factorial del tipo 2^{5-1} se genera con la interacción $ABCDE$ y se realizará con $2^{5-1} = 16$ corridas experimentales.

Resolución de los experimentos factoriales fraccionados

De acuerdo a lo indicado, los efectos que son alias en un experimento factorial fraccionado no pueden separarse y, por ende, no se puede establecer qué efecto es el principal. Para este caso, una necesidad de antemano es la consideración de que un efecto que está aliado con otro es más importante; de allí que, lo ideal, es que un efecto importante sea un aliado de otro efecto de mucha menor importancia. En esta situación es destacable la *resolución* de un factorial fraccionado, siendo preferible que estos experimentos sean de *alta resolución*; entendiéndose como de alta resolución aquellos experimentos en donde los efectos de importancia no están aliados.

Un ejemplo de un factorial fraccionado de alta resolución es del tipo 2^{5-1} . En este experimento los efectos principales y las interacciones dobles tienen alias con las interacciones de tres factores en adelante; es decir, tanto los efectos de los factores principales como las interacciones dobles pueden deducirse fácilmente.

La terminología para identificar el tipo de resolución de un experimento se hace mediante numeración romana. En la práctica, los experimentos que más se desarrollan son los de resolución *III*, *IV* y *V*.

Un experimento es de resolución *III* cuando los efectos principales dados por los factores no son alias entre ellos, pero sus efectos se confunden con los efectos de las interacciones dobles. En este caso están los factoriales fraccionados 2^{3-1} . Los experimentos de resolución *IV* se caracterizan porque los efectos principales no se confunden entre ellos ni con las interacciones dobles; sin embargo, estas últimas si se confunden entre ellas. Un ejemplo de un factorial fraccionado de resolución *IV* es el de tipo 2^{4-1} . En los experimentos de resolución *V* los efectos principales y las interacciones dobles no tienen confusión entre ellos, pero si se alían con interacciones de más de dos factores. En este caso, tanto los efectos principales como las interacciones dobles se determinan claramente. Un ejemplo de un factorial fraccionado de resolución *V* es el de tipo 2^{5-1} .

Obtención de las fracciones 2^{K-1}

La alternativa para obtener las fracciones de experimentos factoriales del tipo 2^{K-1} puede realizarse aplicando los dos pasos siguientes:

1. Se detallan los niveles de $K - 1$ factores en el orden estándar.
2. Se obtiene la última columna en donde se ubica el último factor, multiplicando cada nivel con el signo correspondiente. Para la fracción complementaria, sólo será necesario cambiar de signo los coeficientes del último factor.

Como ejemplo de aplicación de estos pasos se considera un factorial fraccionado 2^{4-1} . Esta resolución se indica en la tabla 8.21, iniciando las tres primeras columnas en la forma estándar de $2^{4-1} = 2^3$; luego el factor D es obtenido del producto ABC.

Tabla 8.21. Fracción principal del experimento factorial 2^{4-1} .

A	B	C	D = ABC	Tratamientos
-1	-1	-1	-1	(1)
+1	-1	-1	+1	ad
-1	+1	-1	+1	bd
+1	+1	-1	-1	ab
-1	-1	+1	+1	cd
+1	-1	+1	-1	ac
-1	+1	+1	-1	bc
+1	+1	+1	+1	abcd

Ejemplo de aplicación 8.4

Para la explicación práctica sobre esta temática se considera lo planteado en el ejemplo de aplicación 8.3 que corresponde a un experimento factorial 2^K con una sola repetición. El experimento representa a un factorial del tipo 2^5 , de característica exploratoria y cuyo propósito es establecer los factores y niveles que son importantes dentro de la productividad del cultivo de arroz con aplicación de riego. En este caso su análisis se realiza considerando la mitad del factorial con el generador $I = +ABCDE$. Los factores y niveles evaluados son los que se detallan a continuación:

Factor	Descripción	Nivel 1 (-1)	Nivel 2 (+1)
A	Preparación de suelo	a_1 : Sin fangueo	a_2 : Con fangueo
B	Tipo de siembra	b_1 : Boleo	b_2 : Trasplante
C	Lámina de riego	c_1 : Sólo a capacidad de campo	c_2 : Con lámina de 10 cm
D	Pre-germinación	d_1 : Semilla pre-germinada	d_2 : Semilla sin pre-germinación
E	Encalado	e_1 : Sin encalado previo	e_2 : Con encalado previo

Los datos de rendimiento se indican en la última columna de la tabla 8.22, dado en unidades de kg/parcela.

Tabla 8.22. Experimento factorial fraccionario 2^{5-1} con generador $I = +ABCDE$ y los rendimientos del cultivo en unidades de kg/parcela.

Nº	A	B	C	D	E	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE	Comb.	Y
1	-1	-1	-1	-1	1	1	1	1	-1	1	1	-1	1	-1	-1	e	4,1
2	1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1	a	5,6
3	-1	1	-1	-1	-1	-1	1	1	1	-1	-1	-1	1	1	1	b	7,8
4	1	1	-1	-1	1	1	-1	-1	1	-1	-1	1	1	-1	-1	abe	11,5
5	-1	-1	1	-1	-1	-1	1	1	1	-1	1	1	-1	-1	1	c	18,5
6	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	-1	1	-1	ace	19,5
7	-1	1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	-1	bce	15,7
8	1	1	1	-1	-1	1	1	-1	-1	1	-1	-1	-1	-1	1	abc	35,7
9	-1	-1	-1	1	-1	1	1	-1	1	1	-1	1	-1	1	-1	d	4,3
10	1	-1	-1	1	1	-1	-1	1	1	1	-1	-1	-1	-1	1	ade	13,6
11	-1	1	-1	1	1	-1	1	-1	-1	-1	1	1	1	-1	-1	bde	23,7
12	1	1	-1	1	-1	1	-1	1	-1	-1	1	-1	-1	1	-1	abd	25,7
13	-1	-1	1	1	1	1	-1	-1	-1	-1	-1	-1	1	1	1	cde	15,7
14	1	-1	1	1	-1	-1	1	1	-1	-1	-1	-1	1	1	-1	acd	34,4
15	-1	1	1	1	-1	-1	-1	-1	1	1	1	-1	1	-1	-1	bcd	25,7
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	abcde	37,9

Solución:

Los contrastes para cada efecto se deducen desde la tabla 8.22 y están determinados por la suma algebraica de cada combinación, considerando los signos respectivos. Por ejemplo, para el factor A el contraste es:

$$\text{CONTRASTE } A = -4,1 + 5,6 - 7,8 + 11,5 - 18,5 + 19,5 - 15,7 + 35,7 - 4,3 + 13,6 - 23,7 + 25,7 - 15,7 + 34,4 - 25,7 + 37,9 = 68,4$$

La cuantificación de los efectos se realiza mediante la siguiente expresión:

$$A = \frac{2 \times \text{CONTRASTE } K}{r \times 2^{K-1}} \quad (8.14)$$

En la expresión 8.14, K es cualquier fuente a analizar (A, B, AB, etc.) y r representa a las repeticiones de cada tratamiento. Nótese que la fórmula está diseñada para utilizarse solamente con experimentos fraccionados del tipo 2^{K-1} . Para el factor A, el efecto respectivo es:

$$A = \frac{2 \times 68,4}{1 \times 2^{5-1}} = 8,6$$

En el caso de las sumas de cuadrados, éstas se definen por la siguiente expresión:

$$SC_K = \frac{(\text{CONTRASTE } K)^2}{R \times 2^{K-1}} \quad (8.15)$$

Para el factor A, la suma de cuadrados es:

$$SC_A = \frac{68,4^2}{1 \times 16} = 292,4$$

En la tabla 8.23 se detallan todos los contrastes y las sumas de cuadrados de cada factor en forma independiente y de las interacciones dobles.

Tabla 8.23. Contrastes y sumas de cuadrados para cada efecto principal e interacción doble.

Nº	A	B	C	D	E	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE
1	-4,1	-4,1	-4,1	-4,1	4,1	4,1	4,1	4,1	-4,1	4,1	4,1	-4,1	4,1	-4,1	-4,1
2	5,6	-5,6	-5,6	-5,6	-5,6	-5,6	-5,6	-5,6	-5,6	5,6	5,6	5,6	5,6	5,6	5,6
3	-7,8	7,8	-7,8	-7,8	-7,8	-7,8	7,8	7,8	7,8	-7,8	-7,8	-7,8	7,8	7,8	7,8
4	11,5	11,5	-11,5	-11,5	11,5	11,5	-11,5	-11,5	11,5	-11,5	-11,5	11,5	11,5	-11,5	-11,5
5	-18,5	-18,5	18,5	-18,5	-18,5	18,5	-18,5	18,5	18,5	-18,5	18,5	18,5	-18,5	-18,5	18,5
6	19,5	-19,5	19,5	-19,5	19,5	-19,5	19,5	-19,5	19,5	-19,5	19,5	-19,5	-19,5	19,5	-19,5
7	-15,7	15,7	15,7	-15,7	15,7	-15,7	-15,7	15,7	-15,7	15,7	-15,7	15,7	-15,7	15,7	-15,7
8	35,7	35,7	35,7	-35,7	-35,7	35,7	35,7	-35,7	-35,7	35,7	-35,7	-35,7	-35,7	-35,7	35,7
9	-4,3	-4,3	-4,3	4,3	-4,3	4,3	4,3	-4,3	4,3	4,3	-4,3	4,3	-4,3	4,3	-4,3
10	13,6	-13,6	-13,6	13,6	13,6	-13,6	-13,6	13,6	13,6	13,6	-13,6	-13,6	-13,6	-13,6	13,6
11	-23,7	23,7	-23,7	23,7	23,7	-23,7	23,7	-23,7	-23,7	-23,7	23,7	23,7	-23,7	-23,7	23,7
12	25,7	25,7	-25,7	25,7	-25,7	25,7	-25,7	25,7	-25,7	-25,7	25,7	-25,7	-25,7	25,7	-25,7
13	-15,7	-15,7	15,7	15,7	15,7	15,7	-15,7	-15,7	-15,7	-15,7	-15,7	-15,7	15,7	15,7	15,7
14	34,4	-34,4	34,4	34,4	-34,4	-34,4	34,4	34,4	-34,4	-34,4	-34,4	34,4	34,4	-34,4	-34,4

15	-25,7	25,7	25,7	25,7	-25,7	-25,7	-25,7	-25,7	25,7	25,7	25,7	-25,7	25,7	-25,7	-25,7
16	37,9	37,9	37,9	37,9	37,9	37,9	37,9	37,9	37,9	37,9	37,9	37,9	37,9	37,9	37,9
Contras.	68,4	68,0	106,8	62,6	-16,0	7,4	35,4	16,0	-21,8	-14,2	22,0	3,8	-14,0	-35,0	17,6
Efecto	8,6	8,5	13,4	7,8	-2,0	0,9	4,4	2,0	-2,7	-1,8	2,8	0,5	-1,8	-4,4	2,2
SC	292,4	289,0	712,9	244,9	16,0	3,4	78,3	16,0	29,7	12,6	30,3	0,9	12,3	76,6	19,4

De forma gráfica, utilizando el diagrama de Pareto que se detalla en la figura 8.13, es posible identificar los efectos de importancia. Según la regla del 80/20, puede notarse que los efectos que en este caso tienen las incidencias de mayor importancia son: A, B, C, D, AC, CE y BD. Para corroborar la significancia de estos resultados, la realización del análisis de varianza que se lleve a cabo con esta información tendrá que enviar las fuentes de variación que no contribuyen de manera importante al error experimental, dado que el experimento tiene una sola réplica. Estos efectos de menor importancia, según el diagrama de Pareto de la figura 8.13, corresponden a: AE, DE, E, AD, BC, CD, AB y BE.

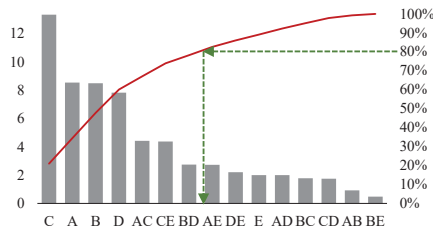


Figura 8.13. Diagrama de Pareto para identificar los efectos de importancia.

La suma de cuadrados del total de los datos es:

$$SC_T = (4,1^2 + 5,6^2 + \dots + 37,9^2) - 299,4^2/16$$

$$SC_T = 1834,6$$

El análisis de varianza respectivo sin considerar las interacciones de tres o más factores y cuyo error experimental está compuesto de los efectos de menor importancia es el que se detalla a continuación:

Tabla 8.24. Análisis de varianza considerando una sola réplica para un factorial fraccionado a la mitad del tipo 2^{5-1} con 8 efectos principales en el error experimental.

Fuentes de variación	GL	SC	CM	F	P-valor
Total	15	1834,60			
Factor A (Preparación de suelo)	1	292,41	292,41	21,22	0,002
Factor B (Tipo de siembra)	1	289,00	289,00	20,97	0,002
Factor C (Lámina de riego)	1	712,89	712,89	51,73	0,000
Factor D (Pre-germinación)	1	244,92	244,92	17,77	0,003
Interacción AC	1	78,32	78,32	5,68	0,044
Interacción BD	1	30,25	30,25	2,20	0,177
Interacción CE	1	76,56	76,56	5,56	0,046
Error experimental	8	110,25	13,78		

El análisis de varianza de la tabla 8.24 permite verificar efectos significativos para las fuentes A, B, C, D, AC y CE; mientras que la interacción doble BD no registra un efecto importante. Estos resultados sugieren la realización de un nuevo ANDEVA, enviando el efecto BD al error experimental, luego los grados de libertad de esta última fuente se incrementan a 9. El detalle de este nuevo ANDEVA se expone en la tabla 8.25.

Tabla 8.25. Análisis de varianza considerando una sola réplica para un factorial fraccionado a la mitad del tipo 2^{5-1} con nueve efectos principales en el error experimental.

Fuentes de variación	GL	SC	CM	F	P-valor
Total	15	1834,60			
Factor A (Preparación de suelo)	1	292,41	292,41	18,73	0,002
Factor B (Tipo de siembra)	1	289,00	289,00	18,51	0,002
Factor C (Lámina de riego)	1	712,89	712,89	45,67	0,000
Factor D (Pre-germinación)	1	244,92	244,92	15,69	0,003
Interacción AC	1	78,32	78,32	5,02	0,052
Interacción CE	1	76,56	76,56	4,90	0,054
Error experimental	9	140,49	15,61		

Con el nuevo análisis de varianza de la tabla 8.25 puede apreciarse que las significancias de los efectos principales A, B, C y D se mantienen; no obstante, los efectos de las interacciones AC y CE, si bien con unas centésimas exceden la regla convencional de tener un valor $p < 0,05$ para declarar significancias, por la magnitud de los valores también pueden ser considerados efectos de importancia. Esto último queda a decisión del investigador el declarar que estas interacciones son significativas.

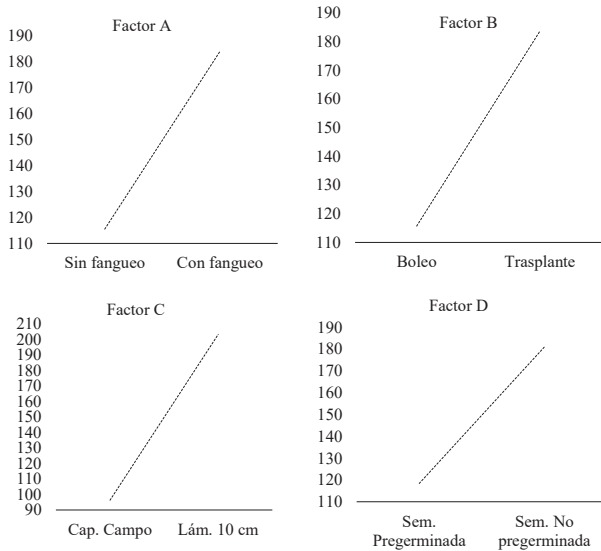


Figura 8.14. Efectos principales de los factores A, B, C y D.

En los análisis de varianza realizados con la información experimental es notoria que los efectos que inciden significativamente en el rendimiento del arroz sobre los factores evaluados son: la preparación del suelo (factor A), el tipo de siembra (factor B), el tipo de riego (factor C) y la pre-germinación (factor D). La combinación factorial de los niveles con los cuales es posible lograr el mayor rendimiento se deducen con la ayuda de las figuras 8.14 y 8.15. En la figura 8.14 se detalla el efecto de cada uno de los factores que resultaron significativos, mientras que las interacciones dobles significativas se indican en la figura 8.15. Dada la importancia que tiene una interacción sobre un efecto unifactorial, la combinación de mayor rendimiento es $a_2b_2c_2d_2e_1$.

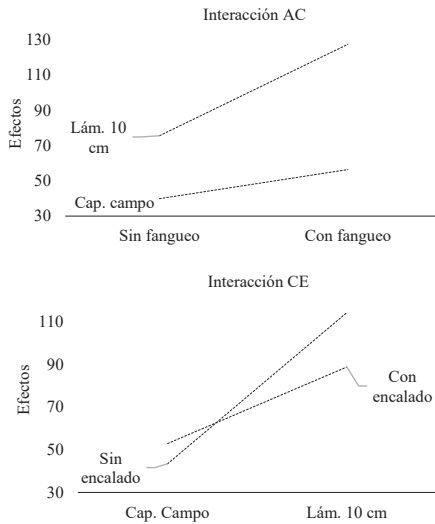


Figura 8.15. Efectos de las interacciones dobles AC y CE.

Nótese que el tratamiento resultante en este ejemplo coincide con el mismo que se obtuvo considerando el factorial completo del ejemplo de aplicación 8.3, de allí que esta sea una alternativa para optimizar costos experimentales sin pérdida de información importante.

Experimento factorial 2^K fraccionado a la cuarta parte

Si lo que se quiere es reducir aún más un experimento factorial por cuestiones económicas, una opción es realizar sólo la cuarta parte de las corridas experimentales. Así, el factorial incompleto que se desarrolla en este caso se conoce como factorial fraccionado a 1/4 y se lo representa como 2^{K-2} ; deducido de $2^K/4 = 2^K/2^2 = 2^{K-2}$. En consonancia con la obtención de niveles de este experimento, para el establecimiento de un factorial de este tipo, se requieren dos efectos generadores que se obtienen mediante las interacciones de mayor grado. Una condición para seleccionar estos dos generadores es que su producto también se corresponda con una interacción del mayor grado posible; de allí que estos experimentos tengan tres generadores: dos seleccionados convenientemente y un tercero, producto de estos dos.

Para diseñar un factorial fraccionado 2^{K-2} lo primero que debe hacerse es considerar los $K - 2$ factores; luego los factores restantes (2) se obtienen del producto de cualquiera de los $K - 2$ factores por los generadores correspondientes. Por ejemplo, para establecer el factorial fraccionado 2^{5-2} , se detallan las combinaciones de los tres primeros factores como un factorial completo 2^3 . Una vez que se han establecido las combinaciones de los tres factores, el siguiente paso es la consideración de los generadores que darán origen a los niveles de los factores D y E.

Si se procede seleccionado los generadores:

$$I = ABD$$

$$I = ACE$$

El producto correspondiente de éstos es:

$$ABD \times ACE = BCDE$$

Por lo tanto, la relación definidora del experimento es:

$$I = ABD = ACE = BCDE$$

En consecuencia, se tiene que:

$$D = I \times D = ABD \times D = ABD^2 = AB$$

$$E = I \times E = ACE \times E = ACE^2 = AC$$

Esto último quiere decir que los niveles de D y E se establecen entre los productos de niveles de los factores A y B en el primer caso, y de A y C en el segundo. Esto es lo que se detalla en la tabla 8.26, en donde los tres primeros factores tienen el color amarillo, mientras que los niveles de D y E se indican con un color turquesa.

Tabla 8.26. Niveles para un experimento factorial 2^{5-2} .

A	B	C	D = AB	E = AC
-1	-1	-1	+1	+1
+1	-1	-1	-1	-1
-1	+1	-1	-1	+1
+1	+1	-1	+1	-1
-1	-1	+1	+1	-1
+1	-1	+1	-1	+1
-1	+1	+1	-1	-1
+1	+1	+1	+1	+1

Los efectos aliados se obtienen multiplicando cada efecto principal por la relación definidora en sus tres generadores. Estos efectos de alias se describen en la tabla 8.27, en la cual se puede apreciar que los efectos principales empiezan confundándose con las interacciones dobles. Por lo tanto, este es un factorial de resolución III, según la interpretación de resolución dada al inicio de estos temas.

Tabla 8.27. Efectos aliados del factorial fraccionado 2^{5-2} .

$$A = BD = CE = ABCDE$$

$$B = AD = ABCE = CDE$$

$$C = ABCD = AE = BDE$$

$$D = AB = ACDE = BCE$$

$$E = ABDE = AC = BCD$$

$$BC = ACD = ABE = DE$$

$$BE = ADE = ABC = CD$$

$$(1) = ABD = ACE = BCDE$$

En la tabla 8.27 se describen la totalidad de los efectos aliados, sin embargo, según se lo ha manifestado varias veces, evaluar los efectos de las interacciones de más de dos factores resulta innecesario; por lo tanto, la matriz de un factorial fraccionado 2^{5-2} se establece considerando sólo hasta las interacciones dobles. Bajo esta concepción es que los efectos alias sólo se consideran hasta los efectos de interacciones de segundo orden (tabla 8.28).

Tabla 8.28. Efectos aliados del factorial fraccionado 2^{5-2} hasta las interacciones dobles.

A = BD = CE
B = AD
C = AE
D = AB
E = AC
BC = DE
BE = CD

Ejemplo de aplicación 8.5

Nuevamente para la explicación práctica de un factorial fraccionado 2^{5-2} se ha considerado lo planteado en el ejercicio de aplicación 8.3 que es el ejemplo para el caso de un experimento factorial 2^K con una sola repetición sobre arroz. El experimento correspondía a un factorial del tipo 2^5 , en el cual los factores y sus niveles correspondientes son los que se indican a continuación:

Factor	Descripción	Nivel 1 (-1)	Nivel 2 (+1)
A	Preparación de suelo	a_1 : Sin fangueo	a_2 : Con fangueo
B	Tipo de siembra	b_1 : Boleo	b_2 : Trasplante
C	Lámina de riego	c_1 : Sólo a capacidad de campo	c_2 : Con lámina de 10 cm
D	Pre-germinación	d_1 : Semilla pre-germinada	d_2 : Semilla sin pre-germinación
E	Encalado	e_1 : Sin encalado previo	e_2 : Con encalado previo

Este experimento, como ya se lo indicó, contenía una sola réplica y la variable evaluada es el rendimiento, medido en unidades de kg/parcela. Estos datos de rendimiento se indican en la última columna de la tabla 8.29, en donde también se detallan los contrastes y las sumas de cuadrados para los efectos principales y las interacciones dobles BC y BE.

Tabla 8.29. Matriz para un factorial fraccionado 2^{5-2} .

Nº	A	B	C	D	E	BC	BE	COMB.	Y
1	-1	-1	-1	+1	+1	+1	-1	<i>de</i>	5,8
2	+1	-1	-1	-1	-1	+1	+1	<i>a</i>	5,6
3	-1	+1	-1	-1	+1	-1	+1	<i>be</i>	21,5
4	+1	+1	-1	+1	-1	-1	-1	<i>abd</i>	25,7
5	-1	-1	+1	+1	-1	-1	+1	<i>cd</i>	13,5
6	+1	-1	+1	-1	+1	-1	-1	<i>ace</i>	19,5
7	-1	+1	+1	-1	-1	+1	-1	<i>bc</i>	13,5
8	+1	+1	+1	+1	+1	+1	+1	<i>abcde</i>	37,9

Los efectos, contrastes y sumas de cuadrados de los factores y de las interacciones BC y BE se indican en la tabla 8.30. Nótese en esta tabla que las interacciones dobles están aliadas con los factores principales y sólo las interacciones BC y BE no tienen alias con estos factores (véase la tabla 8.28).

Tabla 8.30. Contrastes, efectos y sumas de cuadrados (SC) para cada efecto principal de un factorial fraccionado 2^{5-2} .

N°	A	B	C	D	E	BC	BE
1	-5,8	-5,8	-5,8	5,8	5,8	5,8	-5,8
2	5,6	-5,6	-5,6	-5,6	-5,6	5,6	5,6
3	-21,5	21,5	-21,5	-21,5	21,5	-21,5	21,5
4	25,7	25,7	-25,7	25,7	-25,7	-25,7	-25,7
5	-13,5	-13,5	13,5	13,5	-13,5	-13,5	13,5
6	19,5	-19,5	19,5	-19,5	19,5	-19,5	-19,5
7	-13,5	13,5	13,5	-13,5	-13,5	13,5	-13,5
8	37,9	37,9	37,9	37,9	37,9	37,9	37,9
Contrastes	34,4	54,2	25,8	22,8	26,4	-17,4	14,0
Efectos	8,6	13,6	6,5	5,7	6,6	-4,4	3,5
SC	147,9	367,2	83,2	65,0	87,1	37,8	24,5

A partir de los efectos de la tabla 8.30 se elabora un gráfico de Pareto, el cual se describe en la figura 8.16. Según la regla del 80/20 establecido por esta herramienta estadística, puede notarse que los efectos de importancia corresponden únicamente a los factores principales (A, B, C, D y E). Las interacciones dobles BC y BE quedan fuera de la regla porque no aportan significativamente a la respuesta productiva del cultivo.

Por otro lado, al realizar el análisis de varianza (tabla 8.31), el único factor que presenta alguna significancia (menos del 8%) es el factor B; situación que es atribuible a los insuficientes grados de libertad del error experimental.

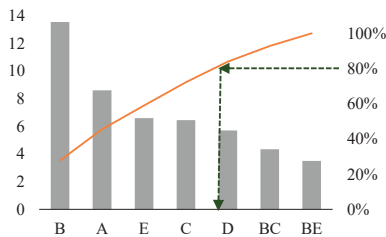


Figura 8.16. Diagrama de Pareto para identificar los efectos principales en un factorial 2^{5-2} .

Tabla 8.31. Análisis de varianza considerando una sola réplica para un factorial fraccionado 2^{5-2} .

Fuentes de variación	GL	SC	CM	F	P-valor
TOTAL	7	812,78			
Factor A (Preparación de suelo)	1	147,92	147,92	4,75	0,161
Factor B (Tipo de siembra)	1	367,21	367,21	11,78	0,075
Factor C (Lámina de riego)	1	83,21	83,21	2,67	0,244
Factor D (Pre-germinación)	1	64,98	64,98	2,08	0,286
Factor E (Encalado)	1	87,12	87,12	2,79	0,237
Error experimental	2	62,34	31,17		

La magnitud de los efectos por nivel, para cada uno de los factores, se describe en la figura 8.17. Este gráfico se ha elaborado considerando los efectos principales establecidos por el diagrama de Pareto de la figura 8.16. De acuerdo a los niveles que presentan los mayores efectos, la combinación factorial a partir de la cual se logra un mayor rendimiento del cultivo es: $a_2b_2c_2d_2e_2$.

Si se compara el resultado obtenido en este ejemplo con el resultado del anterior ejercicio en donde se consideró un factorial fraccionado 2^{5-1} , estos no concuerdan en los niveles del factor E; de allí que realizar sólo 1/4 de un factorial del tipo 2^5 no es muy adecuado.

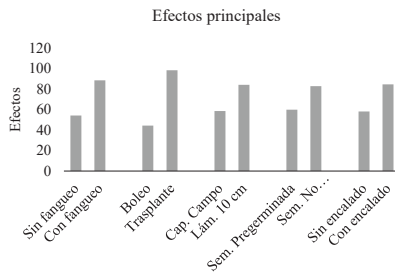


Figura 8.17. Efectos principales por niveles.

Experimentos en bloques con efectos confundidos

A pesar de la bondad en la realización de los experimentos factoriales respecto de aquellos en donde se evalúa un factor a la vez, en la agronomía se tiene una limitante en cuanto al número de factores y/o niveles que pueden valorarse al mismo tiempo en un experimento, debido al incremento de la heterogeneidad del suelo. Por ejemplo, en un experimento en donde se evalúen tres factores, cada uno con cuatro niveles, se tendrá un total de 64 tratamientos. Si este experimento se realiza a nivel de campo (en un ambiente no controlado) y se utiliza una distribución de bloques completos al azar, aun con el cultivo anual de distancia más corta, para

tener resultados distinguibles en la parcela, la amplitud de cada uno de estos bloques fácilmente rebasará los 100 m de longitud. De allí que el error experimental pueda incrementarse sustancialmente debido a que se tendría una superficie relativamente extensa y la presencia de una mayor heterogeneidad en el suelo.

En contraposición a la restricción indicada en el párrafo anterior sobre los experimentos factoriales, una alternativa para contrarrestar ese posible incremento del error experimental es el uso de bloques con efectos confundidos; de tal forma que se suprime alguna (o algunas) fuente de variación, propiciando a que se tengan bloques incompletos con solo una parte de los tratamientos, más pequeños y más homogéneos. En este sentido, voluntariamente se puede confundir (mezclar) el efecto de un factor o de una interacción con el efecto que presentan los bloques; no obstante, como es deducible, lo más lógico es que se confundan interacciones que generalmente no presentan efectos significativos o que de alguna manera son complicadas de explicarlas. Estas interacciones comúnmente son las de segundo orden en adelante (*ABC*) y que corresponden a experimentos con tres o más factores; sin embargo, existe un riesgo latente al que está expuesto el investigador, dado que el efecto que él haya decidido confundir previamente pudiera al final presentar una variabilidad significativa. Si esto último ocurre, en vez de mejorar la precisión de un ensayo, lo que se producirá es un incremento del error experimental y consiguientemente cierta disminución de la precisión.

El único objetivo del uso de arreglos confundidos en experimentos factoriales es el de minimizar el error experimental, a partir de grupos de bloques, aunque son incompletos por no contener a todas las combinaciones factoriales o tratamientos, pero mucho más homogéneos que si se analizara un factorial en forma ordinaria. Si bien existe la posibilidad de perder información que podría ser importante por el hecho de estar confundida; por otro lado, se gana precisión al poder reducir el error experimental por parcela. Este efecto confundido puede ser completo si uno de los componentes de variabilidad factorial se confunde con todas las repeticiones del experimento; sin embargo, en el caso de que la confusión de una fuente esté aplicada sólo a unas cuantas repeticiones del experimento, se dice que dicho arreglo confundido es parcial.

La forma más práctica de confundir un efecto y establecer los bloques incompletos es a través de los arreglos ortogonales (tablas de signos). Estos arreglos se definen considerando los efectos principales y de interacción, así como la notación de Yates representada por las sumatorias de cada uno de los tratamientos o combinaciones factoriales, tal como se indican en la última y segunda columnas de las tablas 8.14 y 8.16, respectivamente. Por ejemplo, si tuviéramos un experimento factorial 2^2 tendríamos la siguiente tabla de signos:

Tabla 8.32. Arreglos ortogonales para un experimento factorial 2^2 .

Fuentes de variación	(1)	<i>a</i>	<i>b</i>	<i>ab</i>
<i>A</i>	-1	+1	-1	+1
<i>B</i>	-1	-1	+1	+1
<i>AB</i>	+1	-1	-1	+1
	a_0b_0	a_1b_0	a_0b_1	a_1b_1

De acuerdo a la tabla 8.32, al confundir la interacción doble (*AB*) tendríamos los siguientes bloques básicos de confusión:

Bloque 1 (+)	
(1)	<i>ab</i>

Bloque 2 (-)	
<i>a</i>	<i>b</i>

Nótese que el bloque 1, según la fila de la interacción *AB*, está compuesto de los efectos de los tratamientos que en la tabla 8.32 se representan con el signo +; mientras que el bloque 2 se conforma con los efectos de los tratamientos que tienen signo -. Asimismo, puede observarse como el efecto de la interacción *AB* sólo se hace presente en el bloque 1, por ello este efecto está confundido con el efecto del bloque y no puede ser valorado de forma independiente. No obstante, el efecto de los factores en sus niveles bajos (a_0, b_0) y altos (a_1, b_1) son ortogonales, no están confundidos y pueden ser valorados independientemente. De manera aproximada podría estimarse el efecto de la interacción *AB* al obtener la diferencia entre los efectos del bloque 1 y 2.

Lo importante que hay que destacar en este ejemplo básico es la magnitud del error experimental que se ve disminuido a la mitad al tener bloques de sólo dos tratamientos antes que, de cuatro, si estuviéramos utilizando una distribución de bloques completos al azar a nivel de campo. Para verificar esta bondad de los experimentos con bloques con efecto confundido, consideramos que este ejemplo se realiza a nivel de campo con cinco repeticiones, cuyo croquis se indica a continuación:

Croquis de campo

I	(1)	<i>ab</i>	<i>a</i>	<i>b</i>
II	<i>a</i>	<i>b</i>	(1)	<i>ab</i>
III	(1)	<i>ab</i>	<i>a</i>	<i>b</i>
IV	<i>a</i>	<i>b</i>	(1)	<i>ab</i>
V	(1)	<i>ab</i>	<i>a</i>	<i>b</i>

Ahora analicemos los modelos de ANDEVA que se generarían, asumiendo que el experimento se realiza dentro de una distribución de bloques completos al azar y también en bloques con efecto confundido:

ANDEVA con todas las fuentes de variación del factorial 2^2 , considerando una distribución de bloques completos al azar.

Fuentes de variación	Grados de libertad
Total	19
A	1
B	1
AB	1
Repeticiones	4
Error experimental	12

ANDEVA del factorial 2^2 con el efecto de la interacción *AB* confundido.

Fuentes de variación	Grados de libertad
Total	19
A	1
B	1
Bloques	9
Error experimental	8

Estos dos modelos de análisis de varianza permiten entrever una aparente pérdida de precisión (matemática) del experimento al pasar de 12 grados de libertad en el error experimental si el ensayo no tuviera ningún efecto confundido y utilizara una distribución de bloques completos al azar, a 8 grados de libertad en dicha fuente al tener confundido el efecto de la interacción

AB. Por otro lado, hay una ganancia práctica en precisión respecto de los bloques; pues en el primer caso se tienen 4 grados de libertad, mientras que en el modelo de efecto confundido la fuente de bloques registra 9 grados de libertad. Esto es, los bloques se analizan considerando una menor área, por ello el error típico por parcela se ve disminuido, permitiendo valorar los efectos de los factores *A* y *B* con mayor contundencia.

La técnica de la confusión es una alternativa para mejorar la precisión, que está indicada para experimentos agrícolas en los cuales las unidades experimentales están representadas por parcelas; es decir, se recomienda su uso cuando el ensayo se lleva a cabo en ambientes no controlados, cuya heterogeneidad del suelo necesita ser agrupada en bloques para controlarla. Dado que su aplicación toma importancia cuando se tiene un número elevado de tratamientos (entiéndase como combinaciones factoriales), es una opción cuando se trata de investigaciones exploratorias, en donde se evalúan varios factores a dos niveles cada uno; por ello es que se trata de confundir efectos de interacciones de segundo orden o más, que en la práctica y por lo general no aportan una información significativa. Así, por ejemplo, para el caso de experimentos como los de tipo 2^3 o los de 2^4 , los grupos de confusión serán:

Tabla 8.33. Formación de bloque con efecto confundido para las interacciones ABC y ABCD en experimentos del tipo 2^3 y 2^4 , respectivamente.

Con la interacción ABC confundida		Con la interacción ABCD confundida	
Bloque 1 (+)	Bloque 2 (-)	Bloque 1 (+)	Bloque 2 (-)
<i>a</i>	(1)	(1)	<i>a</i>
<i>b</i>	<i>ab</i>	<i>ab</i>	<i>b</i>
<i>c</i>	<i>ac</i>	<i>ac</i>	<i>c</i>
<i>abc</i>	<i>bc</i>	<i>bc</i>	<i>abc</i>
		<i>ad</i>	<i>d</i>
		<i>bd</i>	<i>abd</i>
		<i>cd</i>	<i>acd</i>
		<i>abcd</i>	<i>bcd</i>

En el caso del ensayo tipo 2^4 indicado en la tabla 8.33, cuyo efecto confundido es la interacción ABCD, también podría confundirse cualquiera de las cuatro interacciones triples; generándose cuatro bloques incompletos de cuatro tratamientos cada uno⁵⁴. En este caso, confundiendo la interacción triple BCD, se tendrían los siguientes bloques:

⁵⁴ El procedimiento aplicado para obtener cuatro bloques incompletos con efectos confundidos de un ensayo tipo 2^4 se inicia confundiendo la interacción cuádruple ABCD. Con este primer proceso se obtienen dos bloques de 8 tratamientos cada uno, agrupando a los positivos y negativos. Luego, para obtener los cuatro grupos, el investigador deberá definir qué interacción triple desea confundir (ABC, ACD, ABD y BCD) y, a partir de ésta, establecer dichos grupos. De estos cuatro bloques, el primero es positivo y el segundo es negativo, según los signos que tengan los 8 tratamientos del primer bloque en la confusión de la interacción ABCD. Asimismo, los signos del tercero y cuarto bloque serán también positivo y negativo, respectivamente, según los signos del segundo bloque de la interacción ABCD confundida inicialmente.

Tabla 8.34. Bloques de confusión para un experimento 2^4 , confundiendo primero la interacción ABCD y luego la interacción BCD.

Bloque 1 (+)	Bloque 2 (-)	Bloque 3 (+)	Bloque 4 (-)
<i>ab</i>	(1)	<i>b</i>	<i>a</i>
<i>ac</i>	<i>bc</i>	<i>c</i>	<i>abc</i>
<i>ad</i>	<i>bd</i>	<i>d</i>	<i>abd</i>
<i>abcd</i>	<i>cd</i>	<i>bcd</i>	<i>acd</i>

Para complementar el entendimiento de esta temática, a continuación se detalla un ejemplo del tipo 2^3 de factores cualitativos con efecto confundido entre bloques.

Ejemplo de aplicación 8.6

En un experimento con el cultivo de arroz se ha probado el efecto del riego (Factor A) en dos condiciones de humedad, junto con dos métodos de siembra (Factor B) y la aplicación de la fertilización nitrogenada en dos tipos de presentación (Factor C). La variable de respuesta fue el rendimiento (kg/parcela). Para el desarrollo del ensayo, se decidió confundir la interacción triple ABC, utilizando un diseño de bloques con arreglo confundido con 3 repeticiones. Los tratamientos y su identificación en el arreglo fueron los siguientes:

Tabla 8.35. Factores y niveles del ejemplo de aplicación 7.6.

Factores	Niveles
Riego (Factor A)	a_1 : humedad a saturación a_2 : humedad a capacidad de campo
Métodos de siembra (Factor B)	b_1 : siembra por trasplante b_2 : Siembra al voleo
Nitrógeno (Factor C)	c_1 : nitrógeno como urea c_2 : nitrógeno como nitrato

Las combinaciones factoriales (tratamientos) resultantes, de acuerdo a la notación de Yates, y considerando al subíndice 1 como nivel bajo y al subíndice 2 como nivel alto, son las que indican en la tabla 8.36.

Tabla 8.36. Tratamientos del ejemplo de aplicación 8.6.

Combinaciones factoriales (tratamientos)	Notación estándar (Notación de Yates)
$a_1b_1c_1$	(1)
$a_2b_1c_1$	<i>a</i>
$a_1b_2c_1$	<i>b</i>
$a_1b_1c_2$	<i>c</i>
$a_2b_2c_1$	<i>ab</i>
$a_2b_1c_2$	<i>ac</i>
$a_1b_2c_2$	<i>bc</i>
$a_2b_2c_2$	<i>abc</i>

Los arreglos ortogonales para los ensayos del tipo 2^3 de la agrupación de los bloques, de manera que se puedan observar los efectos que conformarán cada uno, son los que se detallan en la tabla 8.37.

Tabla 8.37. Arreglos ortogonales* para un experimento con tres factores a dos niveles cada uno.

Fuentes de variación	(1)	a	b	c	ab	ac	bc	abc
A	-1	+1	-1	-1	+1	+1	-1	+1
B	-1	-1	+1	-1	+1	-1	+1	+1
C	-1	-1	-1	+1	-1	+1	+1	+1
AB	+1	-1	-1	+1	+1	-1	-1	+1
AC	+1	-1	+1	-1	-1	+1	-1	+1
BC	+1	+1	-1	-1	-1	-1	+1	+1
ABC	-1	+1	+1	+1	-1	-1	-1	+1

*Los signos de la codificación se obtienen confrontando las filas con las columnas. En donde coinciden el mismo tipo de letra (por ejemplo A con a) el signo es positivo, caso contrario es negativo. En las interacciones los signos se obtiene del producto de los respectivos factores.

Los componentes de los bloques incompletos básicos, según la confusión de la interacción ABC, son:

$$\begin{array}{cccc}
 & \text{Bloque 1 (+)} & & \text{Bloque 2 (-)} \\
 a & b & c & abc \\
 (1) & ab & ac & bc
 \end{array}$$

Los datos de rendimiento (kg/parcela) por bloque se indican a continuación:

Croquis de campo

I	(1)	ab	ac	bc	a	b	c	abc
	12,1	10,5	17,3	13,5	12,4	12,1	9,8	18,8
Total bloque: 53,4				Total bloque: 53,1				
II	a	b	c	abc	(1)	ab	ac	bc
	18,7	10,8	7,3	20,2	8,5	11,1	16,9	14,8
Total bloque: 57,0				Total bloque: 51,3				
III	a	b	c	abc	(1)	ab	ac	bc
	8,5	10,2	11,0	17,9	13,6	9,8	18,3	21,2
Total bloque: 47,6				Total bloque: 62,9				

Los datos ordenados por tratamiento y según la notación de Yates, son los que se indican en la siguiente tabla:

Tabla 8.38. Datos del rendimiento de arroz (kg/parcela) del ejemplo de aplicación 8.6.

Notación de Yates	Tratamientos	Repeticiones			Total tratamientos
		I	II	III	
(1)	$a_1 b_1 c_1$	12,1	8,5	13,6	34,2
<i>a</i>	$a_2 b_1 c_1$	12,4	18,7	8,5	39,6
<i>b</i>	$a_1 b_2 c_1$	12,1	10,8	10,2	33,1
<i>c</i>	$a_1 b_1 c_2$	9,8	7,3	11,0	28,1
<i>ab</i>	$a_2 b_2 c_1$	10,5	11,1	9,8	31,4
<i>ac</i>	$a_2 b_1 c_2$	17,3	16,9	18,3	52,5
<i>bc</i>	$a_1 b_2 c_2$	13,5	14,8	21,1	49,5
<i>abc</i>	$a_2 b_2 c_2$	18,8	20,2	17,9	56,9
Total repeticiones		106,5	108,3	110,5	

Las sumas de cuadrados del total (SC_T) de datos y de los bloques incompletos (SC_B) se determinan mediante las siguientes expresiones:

$$SC_T = \sum Y^2 - \frac{(\sum Y)^2}{n}$$

$$SC_b = \frac{\sum B^2}{N_{p \times b}} - \frac{(\sum Y)^2}{n}$$

En las cuales $[(\sum Y)^2/n]$ se conoce como factor de corrección y es el término que minimiza los cuadrados. Cada dato está representado por Y y el número total de ellos por n . Los bloques incompletos se representan por b (bloques con efecto confundido) y $N_{p \times b}$ es el número de parcelas por bloques incompletos del ensayo. Estas sumas de cuadrados tienen los siguientes valores:

$$SC_T = 12,1^2 + 12,4^2 + \dots + 21,1^2 + 17,9^2 - \frac{325,2^2}{24} = 389,48$$

$$SC_b = \frac{53,4^2 + 57,0^2 + \dots + 62,9^2}{4} - \frac{325,2^2}{24} = 34,34$$

Las sumas de cuadrados, tanto de los factores como de sus interacciones dobles, se obtienen según las expresiones 8.8 o 8.9 indicadas en este capítulo y aplicadas en varios ejemplos. Para esto, se utilizan los valores codificados de los arreglos ortogonales dados en la tabla 8.37. Así, por ejemplo, para el caso de la suma de cuadrados del factor A, el valor obtenido según se indica en la tabla 8.39 se ha calculado de la forma siguiente:

$$SC_K = (CONTRASTE K)^2/n$$

$$SC_A = \frac{[(-1 \times 34,2) + (+1 \times 39,6) + \dots + (+1 \times 56,9)]^2}{24} = 52,51$$

Tabla 8.39. Sumas de cuadrados para los factores y las interacciones del ejemplo de aplicación 8.6.

Fuentes de variación	(1)	a	b	c	ab	ac	bc	abc	Efectos totales K	Suma de cuadrados
A	-1	+1	-1	-1	+1	+1	-1	+1	35,5	52,51
B	-1	-1	+1	-1	+1	-1	+1	+1	16,5	11,34
C	-1	-1	-1	+1	-1	+1	+1	+1	48,7	98,82
AB	+1	-1	-1	+1	+1	-1	-1	+1	-24,1	24,20
AC	+1	-1	+1	-1	-1	+1	-1	+1	28,1	32,90
BC	+1	+1	-1	-1	-1	-1	+1	+1	35,1	51,33
ABC	-1	+1	+1	+1	-1	-1	-1	+1	-9,9	4,08

El resumen del análisis de varianza con el efecto de la interacción ABC confundido es el que se detalla en la tabla 8.40.

Tabla 8.40. Análisis de varianza para los factores principales e interacciones dobles con la confusión del efecto de la interacción ABC.

Fuentes de variación	GL	SC	CM	F	P-valor
Total	23	389,48			
A	1	52,51	52,51	7,50	0,018
B	1	11,34	11,34	1,62	0,227
C	1	98,82	98,82	14,11	0,003
AB	1	24,20	24,20	3,46	0,088
AC	1	32,90	32,90	4,70	0,051
BC	1	51,33	51,33	7,33	0,019
Bloques incompletos	5	34,34	6,87	0,98	0,468
Error experimental	12	84,03	7,00		

El análisis de varianza de la tabla 8.40 indica efectos significativos ($p < 0,05$) para el factor A, para el factor C y la interacción doble BC; mientras el error experimental presenta una variabilidad dada por la suma de cuadrados de 84,03. Para poder observar el mejoramiento de la precisión, en la tabla 8.41 se presenta el análisis de varianza de este ejemplo sin la confusión del efecto; esto es, se presenta el análisis considerando una distribución de bloques completos al azar.

Tabla 8.41. Análisis de varianza sin la confusión del efecto de la interacción triple ABC, bajo una distribución de bloques completos al azar.

Fuentes de variación	GL	SC	CM	F	P-valor
Total	23	389,48			
A	1	52,51	52,51	6,49	0,023
B	1	11,34	11,34	1,40	0,256
C	1	98,82	98,82	12,21	0,004
AB	1	24,20	24,20	2,99	0,106
AC	1	32,90	32,90	4,07	0,063
BC	1	51,33	51,33	6,34	0,025
ABC	1	4,08	4,08	0,50	0,489
Repeticiones	2	1,00	0,50	0,06	0,940
Error experimental	14	113,28	8,09		

Nótese la disminución respecto de los valores de la suma de cuadrados del error experimental, entre el ANDEVA de la tabla 8.40 y el de la tabla 8.41. Este valor se reduce desde 113,28 en el análisis sin confusión a 84,03 en el análisis con el efecto de la interacción triple ABC confundido. Existe un incremento de la precisión debido a que la confusión permitió tener bloques de sólo 4 tratamientos y no 8, lo que a su vez hizo posible la separación de la variación relativamente alta que puede existir al interior de los bloques completos y que no es detectada por el análisis que se hace cuando se considera una distribución de bloques completos al azar.

Es de resaltar lo que se puede observar en el croquis de campo de este ejemplo. Si este ensayo hubiese sido desarrollado con sus bloques completos, aparentemente las variabilidades de esos bloques no presentan diferencias significativas; de hecho, las sumas de cuadrados de éstos apenas arrojan un valor de 1,00 (tabla 8.41). Sin embargo, en el croquis de campo, el tratamiento $a_2b_1c_1$ (que según notación de Yates es a) varía entre 18,7 en la segunda repetición a 8,5 en la tercera repetición; una variación de un poco más del 100% que no es detectada por el análisis de varianza indicado en la tabla 8.41. Esto se debe a que hay una compensación entre los datos, de tal forma que las sumatorias de repeticiones indicadas en la tabla 8.38 prácticamente presentan valores estadísticamente similares; de allí que sus sumas de cuadrados hayan reportado apenas un valor de 1,00. En este caso, el tener bloques incompletos con efecto confundido pudo separar esta variabilidad y con ello disminuir el error experimental.

Experimento en franjas

Las singularidades en la experimentación agrícola crean condiciones que son únicas para la agronomía. En estas circunstancias, hay factores que son inflexibles en cuanto a la maniobrabilidad de sus niveles y, en consecuencia, hay que buscar las condiciones para que sean aplicados. Uno de estos factores de estudio, por ejemplo, puede ser la preparación del suelo. Para resultados prácticos y consecuentes con la realidad, un experimento que valore métodos de preparación de suelo inevitablemente tiene que utilizar máquinas agrícolas; de allí que se complique la aplicación de dichos métodos en áreas escasamente reducidas como lo son las parcelas de los ensayos comunes. En estas circunstancias se requiere de áreas más grandes en donde la maquinaria pueda desarrollar completamente el método de prueba. Ante esto, una opción práctica son los experimentos en franjas.

Los experimentos en franjas en realidad son experimentos factoriales, cuyos niveles son difíciles de mover y requieren, por lo tanto, parcelas relativamente grandes. Comúnmente son

ensayos en donde se estudian dos factores y su fundamento es el de la distribución de bloques completos al azar, en el cual cada bloque es dividido en un número de franjas horizontales y verticales que estará en función de los niveles de cada uno de los factores previstos para la prueba.

Los tratamientos (combinaciones factoriales) se asignan a cada bloque por una única vez. Primero se hace una división del bloque en franjas horizontales, según los niveles de uno de los factores; los cuales, a su vez, son asignados aleatoriamente a cada franja horizontal. Este proceso se repite para las franjas verticales que también se establecen en el bloque, asignando al azar los niveles del segundo factor. Como ejemplo, consideremos que se tiene un ensayo en donde se evaluarán tres dosis de riego superficial: R_1 , R_2 y R_3 ; además también, se valorarán tres dosificaciones de nitrato de potasio (NO_3K): N_1 , N_2 y N_3 . Para este ejemplo, un bloque (repetición) del experimento, luego de haber realizado las divisiones de franjas horizontales y verticales y de haber sorteado los niveles respectivos, quedaría de la forma como se indica en la figura 8.18.

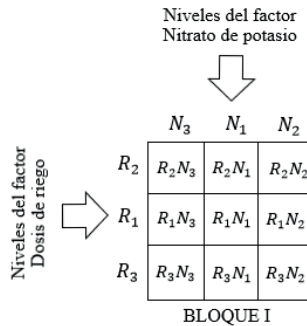


Figura 8.18. Esquema de un bloque (repetición) de un experimento factorial en franjas.

Este arreglo factorial, que en su base tiene una modificación de la distribución de bloques completos al azar, presenta la ventaja de que los dos factores son juzgados con igual número de repeticiones; a diferencia de un experimento factorial en parcelas divididas, en donde uno de los factores es analizado en desventaja en cuanto a precisión. La característica principal de un experimento factorial en franjas, dado que los dos factores utilizan parcelas “grandes”, es la de estimar tres tipos de errores experimentales: un error a , que está relacionado a los niveles del factor A con las franjas horizontales; un error b , relacionado a los niveles del factor B y las franjas verticales; y un error c , que relaciona a los tratamientos (combinaciones factoriales) con los bloques. El modelo estadístico de este arreglo factorial se define de la siguiente forma:

$$Y_{ijk} = \mu + \alpha_i + \delta_j + (\alpha\delta)_{ij} + \beta_k + (\beta\delta)_{jk} + (\alpha\beta)_{ik} + \varepsilon_{ijk}$$

En donde, cualquier dato (Y_{ijk}) representa a la suma de un efecto constante (μ), más el efecto del factor A (α_i) o del que se ubica en franjas horizontales, más el efecto de las repeticiones (δ_j), más el efecto del error a $[(\alpha\delta)_{ij}]$ (en realidad es la interacción entre el factor A y las repeticiones), más el efecto del factor B (β_k), más el efecto del error b $[(\beta\delta)_{jk}]$ (interacción

entre el factor B y las repeticiones), más la interacción de los factores $[(\alpha\beta)_{ik}]$ y más el efecto de un error $c (\varepsilon_{ijk})$ que está ligado a la interacción de los factores con las repeticiones.

Este es un tipo de distribución que, en cierta forma, ofrece facilidades y mayores ventajas en cuanto al control del error experimental, cuando los niveles de uno o de los dos factores requieren de algunas condiciones que hagan posible dicho control. Por ejemplo, si uno de los factores fuese el riego por surcos y sus niveles varias dosis de agua, la situación más adecuada (lo más práctico) sería que en cada franja se aplique una misma dosis (nivel). Esto es mucho más ventajoso que aplicar diferentes dosis en cada parcela de una misma franja; además, se evitaría un eventual traslape de las dosis y, por consiguiente, se controlaría de mejor manera el error experimental. Para una mayor comprensión de esta alternativa experimental, veamos el ejemplo siguiente.

Ejemplo de aplicación 8.7

Se ha evaluado el efecto combinado de tres tipos de labranza de suelos y la aplicación de tres láminas de riego en el cultivo de maíz. Para su desarrollo, considerando la facilidad en el manejo de estos factores de estudio, se recurrió al uso de un arreglo de tratamientos en franjas. La información al respecto es la siguiente:

Tabla 8.42. Niveles de los factores evaluados.

Tipos de labranza de suelo (factor A)		Láminas de riego (factor B)	
a_1 :	Labranza cero	b_1 :	3 mm/día
a_2 :	Labranza mínima	b_2 :	4 mm/día
a_3 :	Labranza convencional	b_3 :	5 mm/día

El experimento tuvo 4 repeticiones y los rendimientos⁵⁵ obtenidos (kg/parcela) son los que se indican a continuación:

Repetición I				Repetición II			
	a_2	a_1	a_3		a_3	a_2	a_1
b_3	46,8	55,3	43,6	b_1	32,7	22,8	39,4
b_2	45,5	33,6	26,6	b_3	44,7	48,5	46,9
b_1	34,6	37,8	28,6	b_2	37,5	49,3	30,6

Repetición III				Repetición IV			
	a_2	a_1	a_3		a_3	a_2	a_1
b_3	52,8	47,8	51,6	b_1	36,5	19,8	39,2
b_2	48,9	26,4	24,8	b_3	40,2	45,5	45,9
b_1	26,9	40,6	30,6	b_2	28,9	50,3	27,7

Las matrices de datos para el desarrollo del análisis de varianza son las siguientes:

⁵⁵ Datos tomados del texto “Diseños y análisis estadísticos para experimentos agrícolas” de este autor.

Tabla 8.43. Matriz de datos organizados del ejemplo 8.7.

Niveles de los tipos de labranza	Niveles de las láminas de riego	Repeticiones				Total	Promedio
		I	II	III	IV		
a_1	b_1	37,8	39,4	40,6	39,2	157,0	39,3
	b_2	33,6	30,6	26,4	27,7	118,3	29,6
	b_3	55,3	46,9	47,8	45,9	195,9	49,0
Subtotal		126,7	116,9	114,8	112,8		
a_2	b_1	34,6	22,8	26,9	19,8	104,1	26,0
	b_2	45,5	49,3	48,9	50,3	194,0	48,5
	b_3	46,8	48,5	52,8	45,5	193,6	48,4
Subtotal		126,9	120,6	128,6	115,6		
a_3	b	28,6	32,7	30,6	36,5	128,4	32,1
	b_2	26,6	37,5	24,8	28,9	117,8	29,5
	b_3	43,6	44,7	51,6	40,2	180,1	45,0
Subtotal		98,8	114,9	107,0	105,6		
Total repeticiones		352,4	352,4	350,4	334,0		

Total general ($\sum Y$)	Número de datos (n)	Promedio general (\bar{Y})
1389,2	36	38,6

Tabla 8.44. Matriz de la suma de datos de cada nivel del factor A de acuerdo a cada repetición.

Tipos de labranza (Niveles del factor A)	Repeticiones				Total
	I	II	III	IV	
a_1	126,7	116,9	114,8	112,8	471,2
a_2	126,9	120,6	128,6	115,6	491,7
a_3	98,8	114,9	107,0	105,6	426,3
Total	352,4	352,4	350,4	334,0	1389,2

Tabla 8.45. Matriz de la suma de datos de cada nivel del factor B de acuerdo a cada repetición.

Láminas de riego (Niveles del factor B)	Repeticiones				Total
	I	II	III	IV	
b_1	101,0	94,9	98,1	95,5	389,5
b_2	105,7	117,4	100,1	106,9	430,1
b_3	145,7	140,1	152,2	131,6	569,6
Total	352,4	352,4	350,4	334,0	1389,2

Tabla 8.46. Matriz de las interacciones de los factores A y B.

Tipos de labranza (Niveles del factor A)	Láminas de riego (Niveles del factor B)			Total
	b_1	b_2	b_3	
a_1	157,0	118,3	195,9	471,2
a_2	104,1	194,0	193,6	491,7
a_3	128,4	117,8	180,1	426,3
Total	389,5	430,1	569,6	1389,2

1. Las hipótesis a probarse en este ejemplo son:

Hipótesis nula	Hipótesis alternativa	Indicación
$H_0: a_1 = a_2 = a_3$	$H_a: a_1 \neq a_2 \neq a_3$	Para los niveles del factor A.
$H_0: b_1 = b_2 = b_3$	$H_a: b_1 \neq b_2 \neq b_3$	Para los niveles del factor B.
H_0 : No hay interacción factorial.	H_a : Si hay interacción factorial.	Para la interacción AB.

2. Grados de libertad de cada fuente de variabilidad del experimento:

$GL_T = n - 1$	$GL_T = 36 - 1 = 35$	Grados de libertad del total de los datos.
$GL_A = a - 1$	$GL_A = 3 - 1 = 2$	Grados de libertad para el factor A.
$GL_R = r - 1$	$GL_R = 4 - 1 = 3$	Grados de libertad para las repeticiones (r).
$GL_{E(a)} = (a - 1)(r - 1)$	$GL_{E(a)} = (3 - 1)(4 - 1) = 6$	Grados de libertad para error experimental referido a los niveles del factor A (error a).
$GL_B = b - 1$	$GL_B = 3 - 1 = 2$	Grados de libertad para el factor B.
$GL_{E(b)} = (b - 1)(r - 1)$	$GL_{E(b)} = (3 - 1)(4 - 1) = 6$	Grados de libertad para error experimental referido a los niveles del factor B (error b).
$GL_{AB} = (a - 1)(b - 1)$	$GL_{AB} = (3 - 1)(3 - 1) = 4$	Grados de libertad para la interacción de los factores (A x B).
$GL_{E(c)} = (a - 1)(b - 1)(r - 1)$	$GL_{E(c)} = (3 - 1)(3 - 1)(4 - 1) = 12$	Grados de libertad para error experimental referido a la interacción de los factores (A x B) (error c).

Nótese que, debido a la existencia de tres errores, los grados de libertad también se han calculado para estos parámetros. Asimismo, el error c también puede deducirse de la diferencia: $GL_{E(c)} = GL_T - (GL_R + GL_A + GL_{E(a)} + GL_B + GL_{E(b)} + GL_{AB})$.

3. Las sumas de cuadrados para cada una de las fuentes se detallan en los siguientes cálculos:

- Suma de cuadrados para el total de los datos:

$$SC_T = \Sigma Y^2 - \frac{(\Sigma Y)^2}{n}$$

$$SC_T = (37,8^2 + 39,4^2 + \dots + 40,2^2) - \frac{1389,2^2}{36}$$

$$SC_T = 56\,897,96 - 53\,607,68 = 3290,28$$

- Suma de cuadrados para los niveles del factor A:

$$SC_A = \frac{\Sigma(\Sigma a_i)^2}{B \times r} - \frac{(\Sigma Y)^2}{n}$$

$$SC_A = \frac{471,2^2 + 491,7^2 + 426,3^2}{3 \times 4} - \frac{1389,2^2}{36}$$

$$SC_A = 53\,794,17 - 53\,607,68 = 186,49$$

- Suma de cuadrados para las repeticiones:

$$SC_r = \frac{\Sigma(\Sigma r_j)^2}{A \times B} - \frac{(\Sigma Y)^2}{n}$$

$$SC_r = \frac{352,4^2 + 352,4^2 + 350,4^2 + 334,0^2}{3 \times 3} - \frac{1389,2^2}{36}$$

$$SC_r = 53\,634,19 - 53\,607,68 = 26,51$$

- Suma de cuadrados para el error a :

$$SC_{E(a)} = \frac{\Sigma(A \cdot r)_{ij}^2}{B} - SC_A - SC_r - \frac{(\Sigma Y)^2}{n}$$

El término $\Sigma(A \cdot r)_{ik}^2$ es la sumatoria al cuadrado de los totales de cada nivel del factor A por cada repetición (matriz de datos de la tabla 7.44).

$$SC_{E(a)} = \frac{126,7^2 + 116,9^2 + \dots + 105,6^2}{3} - 186,49 - 26,51 - \frac{1389,2^2}{36}$$

$$SC_{E(a)} = 53\,911,49 - 186,49 - 26,51 - 53\,607,68 = 90,81$$

- Suma de cuadrados para los niveles del factor B:

$$SC_B = \frac{\Sigma(\Sigma b_k)^2}{A \times r} - \frac{(\Sigma Y)^2}{n}$$

$$SC_B = \frac{389,5^2 + 430,1^2 + 569,6^2}{3 \times 4} - \frac{1389,2^2}{36} = 1487,36$$

- Suma de cuadrados para el error b :

$$SC_{E(b)} = \frac{\sum(B \cdot r)_{jk}^2}{A} - SC_B - SC_r - \frac{(\sum Y)^2}{n}$$

El término $\sum(B \cdot r)_{jk}^2$ es la sumatoria al cuadrado de los totales de cada nivel del factor B por cada repetición (matriz de datos de la tabla 8.45).

$$SC_{E(b)} = \frac{101,0^2 + 94,9^2 + \dots + 131,6^2}{3} - 1487,36 - 26,51 - \frac{1389,2^2}{36}$$

$$SC_{E(b)} = 55\,231,21 - 1487,36 - 26,51 - 53\,607,68 = 109,66$$

- Suma de cuadrados para la interacción de los factores ($A \times B$)

$$SC_{AB} = \frac{\sum(AB)_{ik}^2}{r} - SC_A - SC_B - \frac{(\sum Y)^2}{n}$$

$$SC_{AB} = \frac{157,0^2 + 118,3^2 + \dots + 180,1^2}{4} - 186,49 - 1487,36 - \frac{1389,2^2}{36} = 1161,94$$

- Suma de cuadrados para el error c

$$SC_{E(c)} = SC_T - [SC_A + SC_B + SC_{AB}] - SC_r - SC_{E(a)} - SC_{E(b)}$$

Nótese que la suma de los valores encerrados entre corchetes corresponde a lo que en un experimento unifactorial es la fuente "Tratamientos".

$$SC_{E(c)} = 3290,28 - [186,49 + 1487,36 + 1161,94] - 26,51 - 90,81 - 109,66 = 227,51$$

4. Tabla del análisis de varianza

En la tabla 8.47 se describe todo el análisis de varianza de este ejemplo. El estadígrafo F se ha calculado para las fuentes de interés (Factor A, Factor B e Interacción AB) de acuerdo a los cuadrados medios de los diferentes errores. Para obtener el valor de F en el factor A y las repeticiones, se toma de denominador al cuadrado medio del error a ; para el caso del factor B, el denominador es el cuadrado medio del error b ; y para la interacción AB, se toma como denominador al cuadrado medio del error c .

Tabla 8.47. Análisis de varianza el ejemplo de aplicación 8.7 sobre un experimento en franjas.

Fuentes de variación	GL	SC	CM	F	P-valor
Total	35	3290,28			
Factor A (Tipos de labranza)	2	186,49	93,25	6,16	0,0351
Repeticiones	3	26,51	8,84	0,47	0,7087
Error <i>a</i>	6	90,81	15,14		
Factor B (Láminas de riego)	2	1487,36	743,68	40,68	0,0003
Error b	6	109,66	18,28		
Interacción AB	4	1161,94	290,49	15,32	0,0001
Error <i>c</i>	12	227,51	18,96		

Según el ANDEVA de la tabla 8.47, tanto los factores como la interacción factorial presentan efectos significativos ($p < 0,05$). En este caso, alguna prueba de comparación de medias deberá aplicarse a las combinaciones factoriales.

Si este análisis se realiza en Infostat, el modelo que deberá declararse para este caso es el que se indica en la figura 8.18. Nótese también la estructura de los datos en el programa. La salida con el reporte del análisis de varianza realizado con la ayuda de este software se detalla en la tabla 8.48, cuyos resultados son los mismos que se describen en la tabla 8.47 realizado de forma manual.

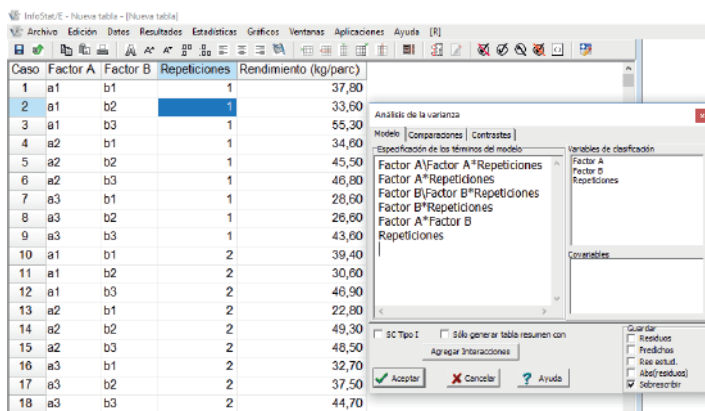


Figura 8.18. Esquema de la estructura de datos y del modelo respectivo para un experimento en franjas en el software Infostat.

Tabla 8.48. Análisis de varianza del ejemplo de aplicación 8.7 realizado con el software Infostat.

Análisis de la varianza

Variable	N	R ²	R ² Aj	CV
Rend. (kg/parc)	36	0,93	0,80	11,28

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor	(Error)
Modelo	3062,79	23	133,16	7,02	0,0006	
Factor A	186,48	2	93,24	6,16	0,0351	(Factor A*Repeticiones)
Factor A*Repeticiones	90,82	6	15,14	0,80	0,5891	
Factor B	1487,35	2	743,68	40,68	0,0003	(Factor B*Repeticiones)
Factor B*Repeticiones	109,68	6	18,28	0,96	0,4881	
Factor A*Factor B	1161,95	4	290,49	15,32	0,0001	
Repeticiones	26,50	3	8,83	0,47	0,7114	
Error	227,49	12	18,96			
Total	3290,28	35				

CAPÍTULO IX

UTILIDAD ECONÓMICA DE LOS TRATAMIENTOS EXPERIMENTALES

Para concluir recomendando una alternativa ganadora de un experimento agrícola realizado con estricto rigor científico y de manejo, dos aspectos son fundamentales: la repetitividad del ensayo para que adquiera validez externa y una clara rentabilidad que vaya más allá de los costos de oportunidad del productor...



Recomendaciones a partir de la rentabilidad de los tratamientos

El uso de la estadística en la investigación científica es fundamental desde una perspectiva de objetividad en el reporte de un resultado. No podemos publicar una información proveniente de un ensayo experimental sino se menciona bajo qué prueba estadística se han determinado diferencias significativas o igualdades estadísticas. Sin embargo, cuando se realiza un ensayo bajo un estricto rigor de manejo y sus resultados no son para publicarlos porque interesan a un sólo productor, además existe una contundente diferencia entre tratamientos cuyos datos son relativamente homocedásticos, las herramientas de la inferencia estadística pueden ser relegadas; de allí que a quien interese dicho resultado, con sólo hacer una comparación intuitiva de las medias será suficiente para que tome una decisión. Es por esta causa que se afirma que la estadística es una herramienta secundaria, o incluso abstracta, menos importante que el diseño y el manejo de un ensayo agronómico.

Por cuanto la agricultura es una actividad comercial, el fin último que se persigue al realizar un experimento es que una nueva alternativa o tratamiento de estudio se presente como una opción, que le permita al agricultor tener una mayor utilidad económica de su actividad. En este contexto, podría pensarse que, más que considerar unas diferencias llamadas “significativas”, mientras no se rebase el famoso y consuetudinario 5% de probabilidad de error tipo I (alfa), la estadística es un elemento secundario al momento de decidir un cambio en el manejo de un cultivo; no obstante, esto no es del todo cierto. Si bien, lo que interesa principalmente es incrementar el rendimiento de un cultivo con el menor costo, la estadística sirve al biometrista, botánico, genetista o investigador agrónomo, para definir si la respuesta fisiológica de las plantas al efecto de un tratamiento es real y no algo que sucede por la casualidad.

Dentro de la agronomía, la estadística es más importante en experimentos que conciernen a investigaciones básicas, ya que en estos ensayos el interés se circunscribe a lo que sucede biológicamente con las plantas. Esto puede ser diferente en experimentos de investigaciones aplicadas y/o demostrativas, de las cuales debe salir una recomendación que adoptará el agricultor debido a que le ofrece una mayor rentabilidad en su proceso productivo. Por ello, en la práctica, la decisión final de seleccionar un tratamiento lo establece la utilidad económica que presenta antes que su “diferencia significativa” respecto del resto de alternativas evaluadas.

Una vez que los resultados de un experimento aplicado se han analizados estadísticamente, lo siguiente deberá ser la valoración económica de los tratamientos con el propósito de establecer cuál de ellos ofrece la mayor rentabilidad, ya que no siempre el tratamiento con el mayor promedio estadístico de producción significa que sea el más rentable. En este sentido, consideremos que tenemos un experimento con dos tratamientos, consistentes en dos dosis de nitrógeno aplicadas en un cultivo. Para una primera dosis de 120 kg/ha de nitrógeno, el rendimiento reportado es de 6800 kg/ha; mientras que para la segunda dosis de 180 kg/ha, el rendimiento es de 8100 kg/ha. Posteriormente, luego de aplicar una prueba estadística, se llega a la conclusión que entre las dos dosis de nitrógeno existen diferencias significativas. Este conocimiento es el límite del criterio estadístico, porque la interrogante que deberá contestarse posteriormente es si la dosis de 180 kg/ha de nitrógeno es más rentable que la dosis de 120 kg/ha como para decidir su aplicación.

Para valorar la utilidad económica que los tratamientos de un experimento ofrecen existen algunas metodologías que pueden adoptarse, dependiendo si los factores de estudio son de naturaleza cuantitativa o cualitativa. En el primer caso se ubican las funciones matemáticas de producción, las mismas que mediante técnicas de superficie de respuesta como las indicadas en el capítulo VIII y/o mediante los modelos de ajuste respectivos, pueden obtenerse soluciones de máximo rendimiento y de rentabilidad óptima. Para factores de característica

.....

cuantitativa se pueden utilizar metodologías como la de *Presupuesto parcial*, la *Tasa de retorno marginal* o la *Curva de beneficios netos*, las cuales son metodologías propuestas por el Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT). Dada las condiciones de la experimentación agrícola, las alternativas del CIMMYT también podrían ser utilizadas en estudios con factores cuantitativos debido a la alta variabilidad del ambiente en donde se desarrollan los experimentos de campo, lo cual le quita pragmatismo al criterio de optimización que matemáticamente sólo podría ser ideal en el papel.

La existencia de resultados con diferencias inobjetables traerá consigo que el investigador termine recomendando la alternativa que en su ensayo se haya revelado como el tratamiento de mayor rendimiento (o al menos uno de los mayores rendimientos), pero con la condición de que dicha superioridad también esté presente en su rentabilidad. Al respecto de esto, no hay que olvidar que, tanto la superficie como las condiciones en las cuales se realizan los ensayos agrícolas distan sustancialmente del área y de las condiciones en las que se desarrollan comercialmente los mismos cultivos que se los utiliza experimentalmente.

Asimismo, no puede obviarse un aspecto fundamental que es la repetitividad de un resultado experimental a través de los ensayos replicados en el espacio y/o tiempo como los descritos en el capítulo 6, de tal forma que se pueda generalizar dentro de una superficie o zona de producción importante; especialmente cuando estos resultados revelan diferencias o efectos no tan convincentes. En este sentido, para realizar las recomendaciones, los experimentos deberán replicarse al menos unas 20 veces (Perrin et al., 1983) y ser valorados por un tiempo mínimo de dos años. Este número de ensayos no es una regla general, ya que su definición más acertada estará en función de la variabilidad que se encuentre dentro de la zona de estudio⁵⁶ y hacia quienes estará dirigida la recomendación; además no debe olvidarse que los resultados de un experimento agronómico son muy inestables en el tiempo y/o en el espacio.

Como será lógico esperar, las recomendaciones emanadas de los ensayos realizados en una zona de estudio deberán incluir fincas con condiciones semejantes de climas y manejos de cultivos, aunque la heterogeneidad del suelo puede contrarrestarse con la realización de estos ensayos en todos los tipos existentes dentro de dicha zona. Los ejemplos clásicos de la agronomía son los experimentos que se realizan con fertilizantes, en donde el efecto de éstos puede confundirse o quedar inexistente cuando el suelo contiene una alta fertilidad; por ello es que estos ensayos deben desarrollarse comúnmente en suelos intensamente cultivados y su dosificación debe realizarse en función de los requerimientos nutricionales de los cultivos.

En los casos en donde no existen diferencias significativas entre los tratamientos, el análisis económico podrá sólo considerar los costos de producción del cultivo en cada una de las alternativas, de tal forma que se pueda escoger a la menos costosa. Cuando el ensayo incluye más de un factor de estudio se realizará el análisis de todos los tratamientos (todas las combinaciones) sólo cuando las interacciones y los promedios sean significativos. No obstante, cuando sólo entre los niveles de uno de los factores existen efectos significativos, el análisis podrá realizarse sólo para éstos niveles; considerando el nivel del otro factor de efecto no significativo que resulte con el menor costo.

Por otro lado, existen algunos ensayos agronómicos que necesariamente no involucran a la variable rendimiento; por lo tanto, la valoración económica de los tratamientos sólo se hace posible hasta un análisis de costos. En estos casos se encuentran, por ejemplo, los ensayos de herbicidas en cultivos perennes o bianuales, en los cuales sólo se valora el efecto de los mismos

⁵⁶ Debe entenderse como zona de estudio a un conglomerado de productores que comparten las mismas condiciones climáticas y de manejo de cultivos, que el CIMMYT define como DOMINIO DE RECOMENDACIÓN. Así, por ejemplo, si se ha previsto realizar las recomendaciones en función de las respuestas de experimentos con cultivos que se realizan con riego, estas recomendaciones no podrían ser las adecuadas para los productores que cultivan sólo en época de secano.

sobre las malezas, y algunas veces, el efecto de toxicidad sobre el cultivo en forma cualitativa mediante algún tipo de escala. Una opción para valorar económicamente estos tipos de ensayos es considerando su efecto residual, contabilizando los días transcurridos hasta el próximo control, resultando ganador el producto que permita tener un mayor intervalo de tiempo entre controles sucesivos, lo que a su vez lo define como la alternativa de menor costo.

En este capítulo, debido al fundamento práctico de los métodos y al interés del agricultor común sobre los réditos de su inversión económica en los cultivos, se realiza una explicación más o menos detallada de la metodología de la Tasa de retorno marginal (*TRM*), compaginándola con la elaboración de curvas de beneficios netos, relación beneficio – costo y una breve descripción sobre el rendimiento óptimo económico. Empero, para entender estas temáticas de valoración se hace preciso explicar ciertas definiciones que se utilizan en estas metodologías antes de entrar a exponer los ejemplos respectivos.

Definición de términos utilizados en la valoración económica de los tratamientos

La concepción adecuada de los parámetros que definen el análisis económico de los tratamientos de un experimento es fundamental en el sentido de tener valores de costos y ganancias muy cercanos o exactos con relación a la situación real de los agricultores. Por ello, a continuación, se hace una explicación de cada uno de estos componentes, tomando en consideración las indicaciones dadas por el CIMMYT en su momento.

Costos fijos (*CF*): en un experimento, los costos fijos hacen referencia al costo de producción del cultivo, excluyendo el costo que involucra a los tratamientos. Dentro del análisis son valores constantes e independientes del costo de los tratamientos, que conciernen a actividades culturales tales como la preparación de suelos, compra de semilla, agroquímicos, riego, etc.

Costos variables (*CV_{ar}*): los costos variables de un experimento son los valores monetarios que demanda cada tratamiento, desde el costo del insumo objeto de estudio (fertilizante, herbicida, fungicida, etc.) en el almacén, su transporte y su aplicación en la finca. Por ejemplo, si los tratamientos son dosis de fósforo, los costos variables corresponderán al precio del fertilizante que lo contiene, los valores por el transporte y los concernientes a la aplicación del mismo. En estos costos variables deben incluirse los *Costos de oportunidad*, los cuales no tiene una cancelación física monetaria, pero son los costos que representan a la otra acción o labor que se dejó de hacer por escoger a la que se realizó en su lugar. Por ejemplo, si un agricultor realiza la poda de sus parcelas, existen unos jornales que se hubiesen tenido que cancelar si dicha labor la hubiese realizado una persona particular. En este caso, existe un costo de oportunidad para el agricultor que tendría el mismo valor que se le hubiese tenido que cancelar a la persona que hubiera realizado la poda.

La definición de los costos variables puede resultar un poco complicado en algunas circunstancias. Por ejemplo, en el caso de que los tratamientos se refieran a distancias de siembra, gran parte del costo total de producción corresponderá al costo variable; ya que existirá todo un grupo de labores culturales que diferirán de un tratamiento a otro. Habrá distintas cantidades de semilla, una mayor cantidad de surcos y más jornales para desarrollar el riego, mayor mano de obra mientras mayor es la densidad de plantas en la cosecha, entre otras circunstancias que tendrán que ser valoradas. Por lo tanto, en este caso los tratamientos involucrarían al costo de producción total o al menos gran parte de éste.

Costo Total (CT): el costo total corresponde a la sumatoria del costo fijo con el costo variable; es decir, es el costo total de producción del cultivo incluyendo al costo de los tratamientos.

Debido a las condiciones de la mayoría de agricultores, los cuales realizan sus ventas a intermediarios en sus mismos campos y a la condición reacia hacia los cambios mientras no existan beneficios contundentes, es preciso indicar que los costos totales deben considerar los costos en el mismo sitio de trabajo del agricultor; es decir en el campo. Por ejemplo, los fertilizantes tienen un costo en el almacén que es distinto cuando dichos fertilizantes están en el campo, debido al transporte, carga y descarga del producto. Los precios que deben considerarse son los precios de los insumos puestos en el campo listos para aplicarse, de allí que se hable de precios en el campo.

Ingresos totales (IT): los ingresos totales se obtienen para cada tratamiento del experimento, multiplicando los rendimientos medios ajustados de cada uno de ellos por el precio del producto cancelado en el campo. A este precio tendrán que disminuirse los costos por la cosecha si el cultivo es vendido en pie, embalaje y hasta los descuentos aplicados por falta de humedad óptima, ya que el intermediario generalmente compra el producto listo para transportarlo.

Rendimiento medio ajustado (RMA): por los cuidados extremos que se deben aplicar en un experimento, generalmente los rendimientos que se obtienen en ellos son más altos que los que se obtienen posteriormente a nivel comercial. Esta es la causa por la que se deben realizar ajustes hacia abajo en los rendimientos experimentales, con el propósito de simular los rendimientos reales que el agricultor obtendría cuando aplique las alternativas en forma masiva en sus campos de producción.

El ajuste del rendimiento va a depender básicamente de la información previa que tenga el investigador sobre los rendimientos comerciales obtenidos en la localidad en donde se realizó el experimento y a la cual se le hará la recomendación del tratamiento más rentable o eficiente. Por ello, una gran ayuda es la de poder comparar los rendimientos reportados por los tratamientos testigos con los resultados reales que obtienen los agricultores de la zona de estudio. Si no existe ninguna información previa, el CIMMYT hace la recomendación de reducir los rendimientos experimentales entre un 5 a un 30% para establecer el RMA. Puesto que hasta el testigo relativo a la práctica del agricultor puede resultar con rendimientos mayores al cultivo comercial, para la realización de este ajuste habrá que tener en cuenta los siguientes puntos:

- **Precisión y cuidado:** en un experimento cada labor se realiza en el momento preciso y con la debida homogeneidad. Por ejemplo, el riego (si este no es el factor de estudio) se realiza de forma uniforme, tratando de evitar una aplicación variable que pueda tergiversar (crear ruido) los resultados posteriores. En los campos definitivos, la uniformidad de las labores es muy difícil y costosa de lograr, aun en la misma aplicación del tratamiento recomendado. Para el agricultor común, muchas veces el factor económico es una restricción para que el empleo de un insumo agrícola sea realizado en el momento oportuno y en la cantidad recomendada. Además, en un experimento el tipo de diseño experimental seleccionado permitirá controlar alguna fuente de variación, lo cual no es posible en un cultivo comercial.
- **Dimensión del área:** en un experimento y dependiendo del tipo de cultivo, las superficies que se utilizan generalmente no van más allá de 1000 m², y si a esto se le une la consideración del área de la unidad de muestreo, en donde no se toman datos de los espacios que hacen de efecto de borde de una unidad experimental, esta área

disminuye aún más; de allí el hecho de tener un mayor control del desarrollo del cultivo y, por consiguiente, unos rendimientos más altos. No es lo mismo un rendimiento obtenido en unos pocos metros cuadrados, con sumo cuidado y extrapolado a una hectárea, que los rendimientos reales obtenidos en extensas áreas de cultivo.

- **Método de cosecha:** existe una gran diferencia entre la cosecha manual y la cosecha mecanizada de un cultivo. Mientras en un experimento comúnmente se hace uso de la primera, en las grandes extensiones de cultivo insoslayablemente deben utilizarse máquinas cosechadoras que producen alguna pérdida del producto.

Tasa de retorno marginal

La alternativa metodológica conocida como la tasa de retorno marginal (TRM), utilizada como un medio de valoración económica de los tratamientos experimentales, es una extensión del método de presupuesto parcial y difundido por el CIMMYT. Está diseñada para establecer la magnitud de la rentabilidad que ofrece el cambio de una alternativa a otra; pues se basa en realizar comparaciones en serie, respecto de los *beneficios netos parciales* obtenidos cada vez que se incrementa ascendentemente la inversión en el siguiente tratamiento. Es decir, esta forma de valoración económica lo que hace es establecer cuál será el *beneficio neto marginal* que se obtendría si se cambia de un tratamiento (metodología, dosis, técnica, etc.) a otro.

La TRM se obtiene a partir de otra metodología propuesta por el CIMMYT, que se define previamente y que es conocida como *PRESUPUESTO PARCIAL*. Esta metodología, llamada así porque sólo se consideran los costos relacionados a los tratamientos, no es más que la diferencia entre los ingresos totales y los costos variables; esto es, entre el producto del rendimiento medio ajustado y los precios de campo de la cosecha, menos los costos de los tratamientos. Esta diferencia es el beneficio neto parcial (*BNP*); el cual no debe confundirse con la rentabilidad de un tratamiento, ya que en la elaboración del presupuesto parcial no se consideran los costos fijos (costos de producción del cultivo, a excepción del costo de los tratamientos). Todos estos cálculos se realizan en unidades de área de 1 hectárea. La TRM se obtiene con la expresión siguiente:

$$TRM = \frac{100 \times \Delta_{BNP}}{\Delta CV_{ar}} \quad (9.1)$$

donde la TRM está dada en porcentaje. El símbolo Δ_{BNP} representa a la diferencia entre los beneficios netos parciales y ΔCV_{ar} es la diferencia de los costos variables; en ambos casos, obtenidos de dos tratamientos adyacentes y ordenados ascendentemente en relación a los costos variables.

El o los tratamientos que en la ordenación ascendente muestren un beneficio neto marginal menor al precedente, cuyos costos variables son más altos que aquel, se llaman *tratamientos dominados*; los cuales tienen una rentabilidad mínima, o en algunos casos, contrariamente reportan pérdidas. Los tratamientos dominados se excluyen de este tipo de análisis económico, por tanto, la TRM no debe registrar valores negativos.

Un elemento importante a considerar en esta forma de valoración es la *Tasa de retorno mínima (TRm)* aceptable para el productor. Esta tasa estará en función de los costos de capital en los que hubiere incurrido un agricultor cuando realiza un préstamo para hacer posible su cosecha, o únicamente para modificar un testigo por un tratamiento nuevo. En términos

comparativos, para realizar un cambio hacia un nuevo tratamiento (por ejemplo, ir de una labranza cero a una labranza completa), la tasa de retorno marginal debe ser mayor a la tasa de retorno mínima; esta condición permitirá, además de pagar la deuda o cubrir el costo del capital, tener una ganancia extra que sustente el trabajo del productor.

Para establecer la TRm, una alternativa es recurrir a criterios empíricos al considerar tasas que estén comprendidas entre el 50% y el 100% (CIMMYT, 1988), seleccionándose valores cercanos al menor cuando el tratamiento requiere de cambios que demandan mínimos costos y valores cercanos al mayor cuando los tratamientos demandan una mayor sofisticación como, por ejemplo, el empleo de equipos.

Otra opción de establecer la TRm es considerar los intereses causados por los préstamos realizados (uso de capital), ya sea en mercados formales (instituciones bancarias privadas u estatales) o en informales (prestamistas). En estas tasas debe considerarse el tiempo que dura el cultivo desde la siembra hasta su cosecha.

Ejemplo de aplicación 9.1

En un ensayo de arroz se han evaluado 5 dosis de nitrógeno, junto a un tratamiento testigo en donde no se aplicó este nutriente. Los rendimientos medios obtenidos son los que se indican en la tabla 9.1. En esta tabla también se describe el detalle de la metodología de presupuesto parcial.

Tabla 9.1. Ejemplo de Presupuesto Parcial para un experimento sobre la valoración del efecto de 5 dosis de nitrógeno en arroz.

Descripción	Tratamientos					
	1 0 kg N/ha	2 30 kg N/ha	3 60 kg N/ha	4 90 kg N/ha	5 120 kg N/ha	6 150 kg N/ha
Rendimiento (kg/ha)	1550d	2150c	3810b	4720ab	5220a	5240a
Rendimiento ajust. (kg/ha)	1395	1935	3429	4248	4698	4716
Ingresos totales	488,3	677,3	1200,2	1486,8	1644,3	1650,6
Costos del fertilizante	0,0	52,2	104,4	156,6	208,8	261,0
Costo de mano de obra	0,0	15,0	30,0	45,0	60,0	75,0
Total costo variable	0,0	67,2	134,4	201,6	268,8	336,0
Beneficio neto parcial	488,3	610,1	1065,8	1285,2	1375,5	1314,6

Letras iguales en los promedios de rendimiento no difieren significativamente (p<0,05).

Según lo descrito en la tabla 9.1, el rendimiento se ajustó considerando una afectación (reducción) del 10%. Los ingresos totales se calcularon con un precio de campo del arroz en cáscara de \$0,35/kg, disminuyendo \$0,05/kg debido a su cosecha. El costo del fertilizante se estableció en un precio de \$1,74 por kg de nitrógeno a través del producto comercial urea. Cada jornal se ha considerado bajo un precio de \$15,0.

De acuerdo al beneficio neto parcial, la mayor utilidad se registra en el tratamiento 5, con una dosis de 120 kg N/ha. En este caso, puede observarse desde ya la existencia de un tratamiento dominado, con una menor utilidad y mayor costo variable, representado por el tratamiento 6, en el cual se produce una reducción de \$60,9 respecto del tratamiento 5.

En la figura 9.1 se define la relación entre los beneficios netos parciales y el costo variable de cada tratamiento. Su relación se ajusta de forma contundente ($r^2 = 0,9819$) a un modelo polinómico de tercer grado.

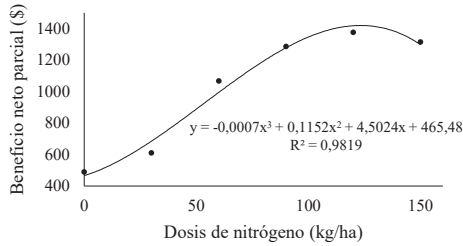


Figura 9.1. Curva de dosis de nitrógeno versus beneficio neto parcial.

Puede definirse al modelo $Y = 465,48 + 4,5024X + 0,1152X^2 - 0,0007X^3$ como una función de rentabilidad parcial, cuyo máximo se establece cuando su primera derivada tiene pendiente cero; esto es:

$$\frac{dY}{dX} = 4,5024 + 0,2304X - 0,0021X^2 = 0$$

Función cuadrática cuya solución positiva es:

$$X = 126,6 \text{ kg/ha}$$

Esto quiere decir que el máximo beneficio neto parcial se obtiene cuando se aplican 126,6 kg de nitrógeno por hectárea.

En la tabla 9.2 se describe la tasa de retorno marginal para el cambio de un tratamiento al siguiente, obtenida con la expresión 9.1 y excluyéndose al tratamiento dominado (tratamiento 6). Asimismo, se detallan los costos variables y beneficios netos marginales.

Tabla 9.2. Tasa de retorno marginal.

N°	Tratamientos	Total costo variable (\$)	Beneficio neto parcial (\$)	Costos variables marginales (\$)	Beneficios netos marginales Δ_{BNP} (\$)	Tasa de retorno marginal ΔCV_{ar} (%)
1	0 kg N/ha	0,0	488,3			
2	30 kg N/ha	67,2	610,1	67,2	121,8	181,3
3	60 kg N/ha	134,4	1065,8	67,2	455,7	678,1
4	90 kg N/ha	201,6	1285,2	67,2	219,5	326,6
5	120 kg N/ha	268,8	1344,0	67,2	58,8	87,5

De acuerdo al análisis marginal, cambiarse del tratamiento 1 al 2 es una buena alternativa, ya que se tendría una TRM de 181,3%; sin embargo, pasar del tratamiento 2 al 3 se tendrá una

mejor TRM (el máximo retorno); y aun ir del tratamiento 3 al 4 también se presenta como una aceptable alternativa. Sin embargo, cambiar el tratamiento 4 por el 5, a pesar de existir ganancias, el 87,5% de retorno puede resultar una tasa no tan atractiva para el productor si los costos del capital apenas son cubiertos con estas ganancias. Sólo en el caso de que el costo del tratamiento 4 al 5 permita tener unos ingresos que hagan posible cumplir con la tasa de retorno mínima, y además generar una utilidad (ganancia extra), podría aceptarse como el mejor tratamiento.

Asumiendo que el productor necesita obtener unos ingresos que sean superiores a una TRM de 100%, la mejor opción es el tratamiento 4; es decir, la aplicación de 90 kg N/ha. El empleo de esta dosis, considerando sólo el costo variable (fertilizante, transporte y mano de obra) del tratamiento, le permitirá al productor tener una ganancia de \$3,27 por cada unidad monetaria invertida. Es importante indicar que los \$3,27 no debe ser visto como utilidad para el productor, ya que la TRM sólo considera los costos variables (costo de los tratamientos) y no suma los costos fijos (costo de producción del cultivo, excepto los costos de los tratamientos).

Relación beneficio – costo

La relación beneficio – costo, cuyo símbolo es RB/C, es un indicador de eficiencia económica de los proyectos productivos que, en el caso de los ensayos, permiten una visión completa de la utilidad económica que recibirá el productor por la aplicación de las diferentes alternativas (tratamientos) evaluadas. Se define a partir de la relación entre los ingresos netos o beneficio (ingreso total menos el costo total); que se obtiene por la aplicación de un tratamiento; y el costo total de dicha alternativa, siendo éste último la suma entre el costo fijo más el costo variable.

La interpretación de la RB/C toma en consideración al cero; valores mayores a éste, indican la utilidad que se obtiene por cada unidad monetaria invertida, mientras que valores menores indican pérdidas. Para su obtención es necesario considerar el costo del capital invertido (intereses) o los costos de oportunidad, dado que los beneficios se obtienen después de cierto tiempo que dependerá del desarrollo de los cultivos.

Para lo resuelto en el ejemplo de aplicación 9.1, al obtener la utilidad real para el productor a través de la relación beneficio/costo (RB/C) de cada uno de ellos, se genera una discrepancia respecto de la mejor alternativa. Si bien, el tratamiento ganador bajo la TRM fue el 4 (90 kg N/ha); no obstante, quien presenta la mayor RB/C es el tratamiento siguiente (T5), con 120 kg N/ha según se observa en la tabla 9.3.

Tabla 9.3. Relación beneficio/costo de los tratamientos.

Nº	Tratamientos	Costos fijo (\$)	Costos variables (\$)	Costo total (\$)	Ingresos totales (\$)	Beneficios (\$)	Relación B/C
1	0 kg N/ha	1100	0,0	1100,0	488,3	-611,8	-0,56
2	30 kg N/ha	1100	67,2	1167,2	677,3	-490,0	-0,42
3	60 kg N/ha	1100	134,4	1234,4	1200,2	-34,3	-0,03
4	90 kg N/ha	1100	201,6	1301,6	1486,8	185,2	0,14
5	120 kg N/ha	1100	268,8	1368,8	1644,3	275,5	0,20
6	150 kg N/ha	1100	336,0	1436,0	1650,6	214,6	0,15

Para la obtención de la RB/C se ha considerado un costo fijo de producción del cultivo de \$1100 por hectárea, lo que sumado a los costos variables respectivos se tienen los costos totales. Estos costos, restados de los ingresos totales producto de los rendimientos medios ajustados por el precio del kg de arroz en cáscara, se obtienen los beneficios. Estos beneficios

al ser relacionados con los costos totales permiten establecer los valores de la RB/C de cada tratamiento. De los 6 tratamientos evaluados, según lo que se detalla en la tabla 9.3, el tratamiento de mayor rentabilidad es el 5, generando una utilidad de \$0,20 por cada unidad monetaria invertida.

Cabe recalcar que en la deducción del costo fijo de un cultivo deberá considerarse, ya sea el costo de oportunidad cuando se utiliza mano de obra familiar o el costo del capital cuando se recurre a los préstamos. Esta valoración en la RB/C permitirá conocer las ganancias reales (posibles) para el productor.

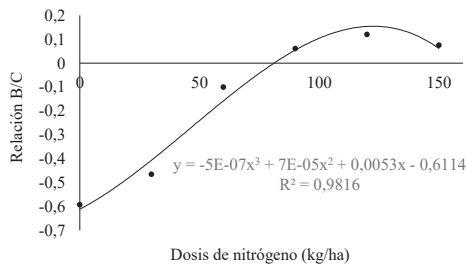


Figura 9.2. Curva de dosis de nitrógeno versus la relación B/C.

Al generar la función de rentabilidad utilizando la relación B/C junto con las dosis de nitrógeno (tratamientos), se ha obtenido el modelo indicado en la figura 9.2. Si desde este modelo se obtiene la máxima RB/C considerando su primera derivada igual a cero y la solución positiva de la expresión de segundo grado definida con los datos, la dosis óptima para obtener este valor es la que se indica a continuación:

$$X = 122,2 \text{ kg}$$

Esta dosis de máxima rentabilidad permitirá tener una RB/C siguiente:

$$Y = -0,6114 + 0,0053(122,2) + 7 \times 10^{-5}(122,2)^2 - 5 \times 10^{-7}(122,2)^3$$

$$Y = 0,17$$

Pensar en optimización en el campo de la agronomía, especialmente en las labores que están relacionadas al suelo agrícola, puede ser una pérdida de tiempo debido a la alta variabilidad que espacial y temporalmente sucede en el ambiente de un cultivo; consecuentemente y tal como ya se lo ha indicado en el capítulo 8, mejor sería aplicar el criterio de un intervalo. Bajo este enfoque, en este ejemplo, las dosis más adecuadas de la fertilización nitrogenada estarían entre los 90 y 120 kg/ha para tener el mayor rendimiento y la mejor rentabilidad; sin embargo, una recomendación adicional es seguir con ensayos de ajuste en el rango de las dosis indicadas para tener una mejor certeza en la recomendación para los productores.

Un punto importante que debe indicarse al respecto de este ejercicio en la comparación entre la TRM y la RB/C, es el conocimiento hasta qué nivel de un factor en escala cuantitativa se

producen las ganancias netas de forma efectiva. Si se obtuviera una TRM promedio entre el tratamiento testigo (0 kg N/ha) y el tratamiento de 120 kg N/ha, ésta tasa sería la siguiente:

$$TRM = \frac{100 \times (1344,0 - 488,3)}{(268,8 - 0)} = 318,3 \%$$

Una tasa aparentemente muy buena; ya que, asumiendo que el productor tenga que cumplir la obligación de un costo de capital del 100% (TRm), estaría obteniendo ganancias muy por encima de dicha obligación. No obstante, en este proceder se está ocultando el hecho de que la mayoría de ganancias se producen a niveles de menor inversión; esto es, se produce la mayor rentabilidad a menos de los 120 kg N/ha (hasta los 90 kg N/ha, según la tabla 9.2). Al considerar una tasa de retorno mínima, esta apreciación permite entrever que el tratamiento recomendable no es el que tiene la mayor TRM o la más alta RB/C, ni mucho menos el de mayor rendimiento; sino aquel cuyas ganancias netas hagan posible cubrir el costo del capital y generar ingresos extras.

Rendimiento óptimo económico

De acuerdo a lo indicado al inicio de este capítulo, el rendimiento máximo que se tiene para un determinado nivel de un factor cuantitativo o de ciertos niveles cuando se trata de experimentos factoriales, no significa que también en esos mismos niveles se obtengan las máximas ganancias. Esta situación se puede evidenciar con el ejemplo que se ha indicado sobre las cinco dosis de nitrógeno en arroz; en el cual, a partir de los datos del rendimiento ajustado y de las dosis correspondientes, es posible obtener la función de producción de mejor ajuste. Esta función está definida por una expresión polinómica de tipo cuadrática indicada a continuación:

$$Y = 1329,90 + 19,26X + 0,33X^2 - 0,0021X^3 \quad (b)$$

Desde este planteamiento, el máximo rendimiento estará dado cuando su primera derivada sea igual a cero. Por lo tanto, la cantidad de nitrógeno para lograr este valor máximo se deduce de la siguiente igualdad:

$$dY/dX = 19,26 + 0,66X - 0,0063X^2 \quad (c)$$

Cuya solución positiva es:

$$X = 128,5 \text{ kg}$$

Dosis que sustituida en la expresión *b*, corresponde al rendimiento de:

$$Y = 4798 \text{ kg/ha}$$

Por otro lado, la utilidad que tendrá un productor en cualquiera de las alternativas estará dada por la expresión:

$$U = YP_Y - XC_X - C_F \quad (9.2)$$

donde U es la utilidad por unidad de cosecha obtenida, P_Y es el precio unitario pagado en campo por la cosecha, C_X es el costo variable unitario por unidad de producto utilizado (en el ejemplo, kg de fertilizante) y C_F es el costo fijo en que se incurre para obtener una unidad de cosecha.

Asimismo, la máxima utilidad se podrá obtener al igualar a cero la primera derivada respecto de X en la función 9.2; esto es:

$$\frac{dU}{dX} = \frac{dY}{dX}P_Y - C_X = 0$$
$$dY/dX = C_X/P_Y \quad (9.3)$$

De esta relación puede deducirse que la máxima rentabilidad puede lograrse cuando se iguala la primera derivada de la función de producción con la relación entre el costo unitario del tratamiento (costo variable unitario) y el precio unitario en campo de la cosecha del cultivo.

Considerando las expresiones c y la 9.3, respecto del ejemplo de aplicación 9.1, se tiene la función siguiente:

$$19,26 + 0,66X - 0,0063X^2 = C_X/P_Y$$

Si de acuerdo a lo indicado en la tabla 9.1 el costo de cada kg de nitrógeno es de \$1,74 y si a esto también se le adiciona el costo de aplicación de \$0,50/kg, se tiene un costo variable unitario (C_X) de \$2,24. El precio unitario de cosecha vendido en campo (P_Y) es de \$0,35. Con estos valores, la mayor utilidad se obtiene cuando se aplica la siguiente cantidad de nitrógeno por hectárea:

$$19,26 + 0,66X - 0,0063X^2 = 2,24/0,35$$
$$X = 121,6 \text{ kg}$$

Valor que es apenas diferente del nitrógeno necesario para obtener el máximo rendimiento del cultivo (128,5 kg/ha).

1^{RA} EDICIÓN

PRECISIÓN EN LOS EXPERIMENTOS AGRÍCOLAS

Algunas Consideraciones y Alternativas

Anexos



Anexos

Tabla 1A. Áreas de curva normal estándar con escala Z.

Z	0,0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990

Tabla 2A. Valores críticos de la distribución *t* de student para distintos niveles de significancia α .

<i>G.L.</i> <i>error</i>	Nivel de significancia								
	0,25	0,20	0,15	0,10	0,05	0,025	0,01	0,005	0,0005
1	1,000	1,376	1,963	3,078	6,314	12,70	31,821	63,65	636,619
2	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,598
3	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,941
4	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,405
8	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,700	0,879	1,093	1,371	1,812	2,228	2,764	3,169	4,587
11	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,833
20	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
60	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,674	0,842	1,936	1,282	1,645	1,960	2,326	2,576	3,291

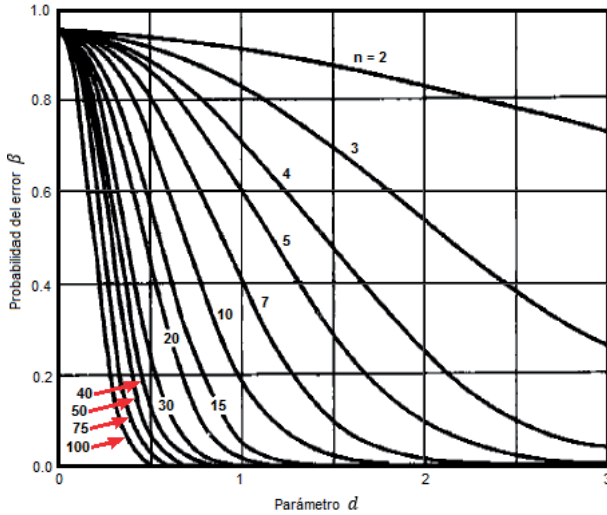


Figura A1. Curva de operación característica para experimentos de dos tratamientos, como prueba de dos colas y 0,05 de significancia (Montgomery, 2005).

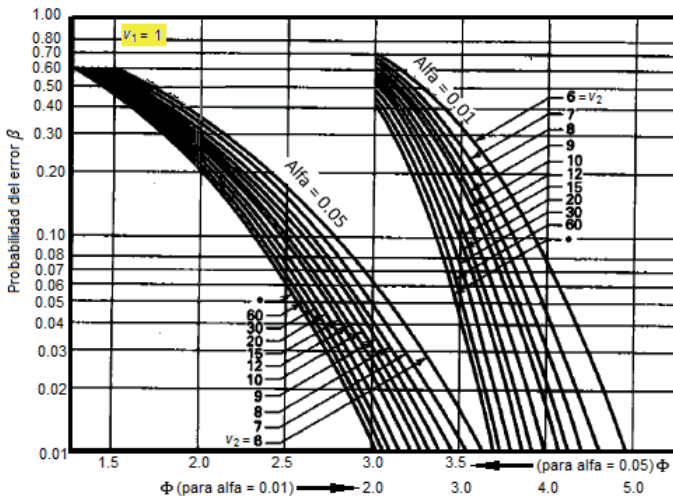


Figura A2. Curva de operación característica para el análisis de varianza, bajo el criterio de efectos fijos, con 1 grado de libertad en el numerador (V_1) y con 6 o más grados de libertad en el denominador (error experimental) (V_2) (Montgomery, 2005).

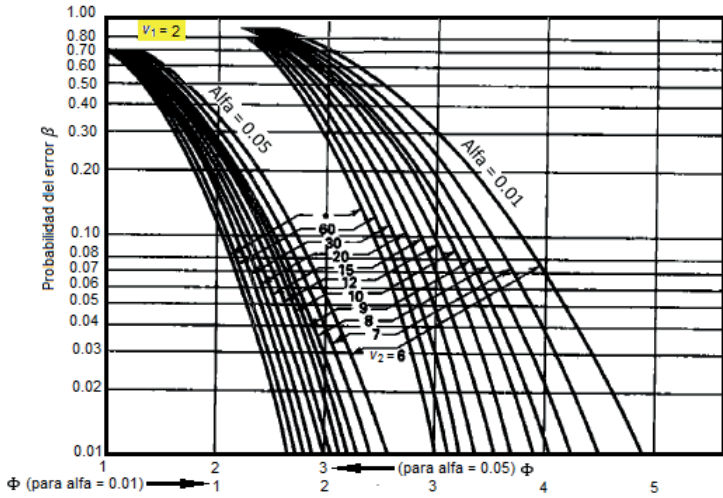


Figura A3. Curva de operación característica para el análisis de varianza, bajo el criterio de efectos fijos, con 2 grados de libertad en el numerador (V_1) y con 6 o más grados de libertad en el denominador (error experimental) (V_2) (Montgomery, 2005).

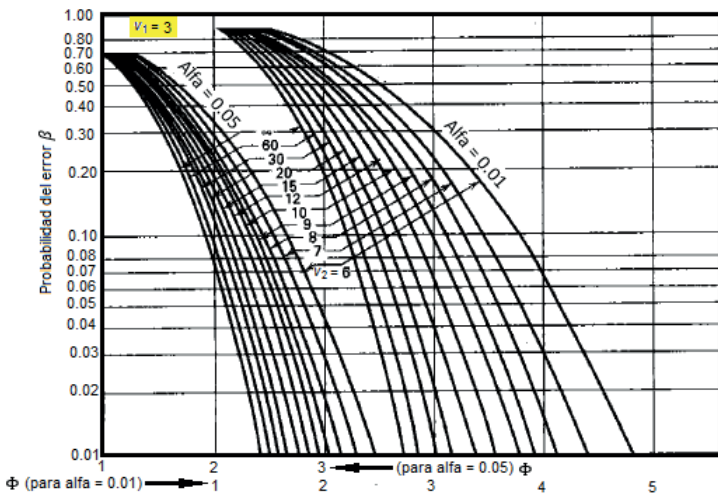


Figura A4. Curva de operación característica para el análisis de varianza, bajo el criterio de efectos fijos, con 3 grados de libertad en el numerador (V_1) y con 6 o más grados de libertad en el denominador (error experimental) (V_2) (Montgomery, 2005).

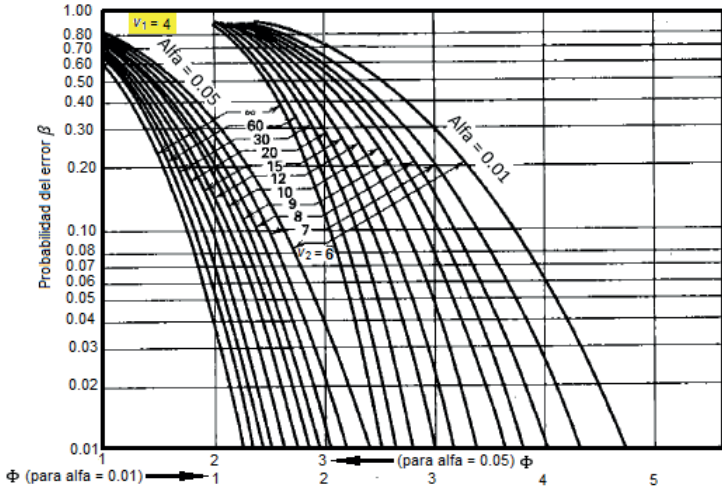


Figura A5. Curva de operación característica para el análisis de varianza, bajo el criterio de efectos fijos, con 4 grados de libertad en el numerador (V_1) y con 6 o más grados de libertad en el denominador (error experimental) (V_2) (Montgomery, 2005).

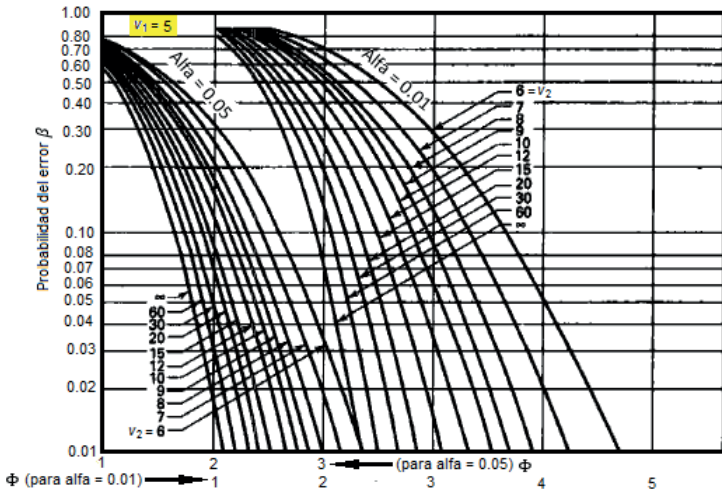


Figura A6. Curva de operación característica para el análisis de varianza, bajo el criterio de efectos fijos, con 5 grados de libertad en el numerador (V_1) y con 6 o más grados de libertad en el denominador (error experimental) (V_2) (Montgomery, 2005).

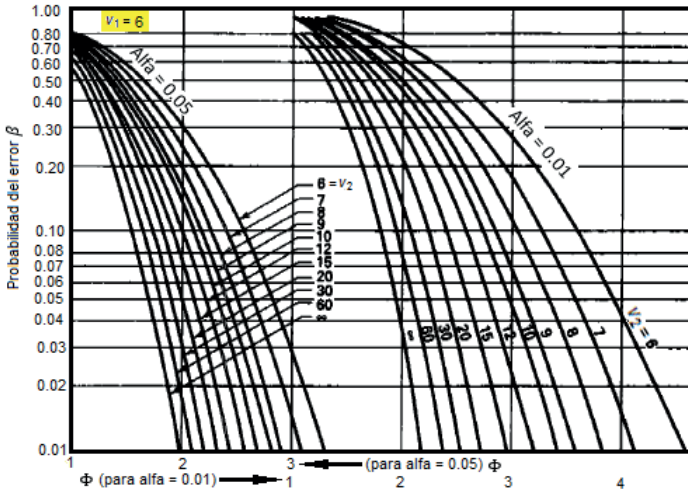


Figura A7. Curva de operación característica para el análisis de varianza, bajo el criterio de efectos fijos, con 6 grados de libertad en el numerador (V_1) y con 6 o más grados de libertad en el denominador (error experimental) (V_2) (Montgomery, 2005).

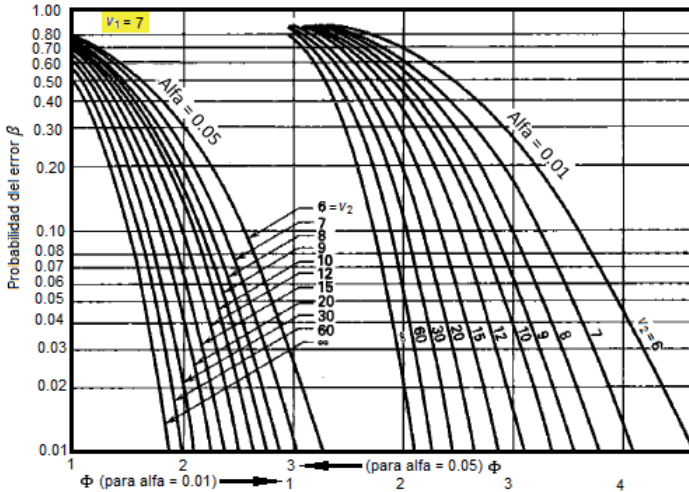


Figura A8. Curva de operación característica para el análisis de varianza, bajo el criterio de efectos fijos, con 7 grados de libertad en el numerador (V_1) y con 6 o más grados de libertad en el denominador (error experimental) (V_2) (Montgomery, 2005).

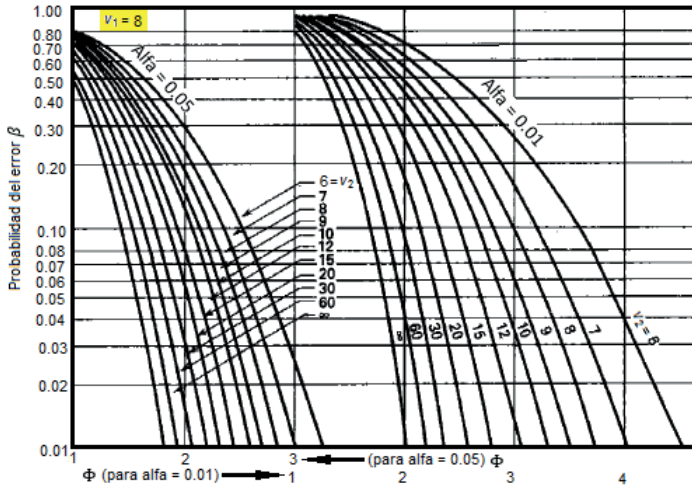


Figura A9. Curva de operación característica para el análisis de varianza, bajo el criterio de efectos fijos, con 8 grados de libertad en el numerador (V_1) y con 6 o más grados de libertad en el denominador (error experimental) (V_2) (Montgomery, 2005).

Latice balanceado 3 x 3

$$t = 9, k = 3, r = 4, b = 12, \lambda = 1$$

Bloq.	Rep. I			Bloq.	Rep. II			Bloq.	Rep. III			Bloq.	Rep. IV		
(1)	1	2	3	(4)	1	4	7	(7)	1	5	9	(10)	1	8	6
(2)	4	5	6	(5)	2	5	8	(8)	7	2	6	(11)	4	2	9
(3)	7	8	9	(6)	3	6	9	(9)	4	8	3	(12)	7	5	3

Latice balanceado 4 x 4

$$t = 16, k = 4, r = 5, b = 20, \lambda = 1$$

Bloq.	Rep. I				Bloq.	Rep. II				Bloq.	Rep. III			
(1)	1	2	3	4	(5)	1	5	9	13	(9)	1	6	11	16
(2)	5	6	7	8	(6)	2	6	10	14	(10)	5	2	15	12
(3)	9	10	11	12	(7)	3	7	11	15	(11)	9	14	3	8
(4)	13	14	15	16	(8)	4	8	12	16	(12)	13	10	7	4

Bloq.	Rep. IV				Bloq.	Rep. V			
(13)	1	14	7	12	(17)	1	10	15	8
(14)	13	2	11	8	(18)	9	2	7	16
(15)	5	10	3	16	(19)	13	6	3	12
(16)	9	6	15	4	(20)	5	14	11	4

Latice balanceado 5 x 5

$$t = 25, k = 5, r = 6, b = 30, \lambda = 1$$

Bloq.	Rep. I					Bloq.	Rep. II					Bloq.	Rep. III				
(1)	1	2	3	4	5	(6)	1	6	11	16	21	(11)	1	7	13	19	25
(2)	6	7	8	9	10	(7)	2	7	12	17	22	(12)	21	2	8	14	20
(3)	11	12	13	14	15	(8)	3	8	13	18	23	(13)	16	22	3	9	15
(4)	16	17	18	19	20	(9)	4	9	14	19	24	(14)	11	17	23	4	10
(5)	21	22	23	24	25	(10)	5	10	15	20	25	(15)	6	12	18	24	5

Bloq.	Rep. IV					Bloq.	Rep. V					Bloq.	Rep. V				
(16)	1	12	23	9	20	(21)	1	17	8	24	15	(26)	1	22	18	14	10
(17)	16	2	13	24	10	(22)	11	2	18	9	25	(27)	6	2	23	19	15
(18)	6	17	3	14	25	(23)	21	12	3	19	10	(28)	11	7	3	24	20
(19)	21	7	18	4	15	(24)	6	22	13	4	20	(29)	16	12	8	4	25
(20)	11	22	8	19	5	(25)	16	7	23	14	5	(30)	21	17	13	9	5

Figura A10. Estructuras en latice (cuadrados) balanceados para experimentos con 9, 16 y 25 tratamientos (Cochran y Cox, 1978).

Lattice rectangular 3 x 4

Bloq.	Grupo X			Bloq.	Grupo Y			Bloq.	Grupo Z		
(1)	1	2	3	(1)	4	7	10	(1)	6	8	12
(2)	4	5	6	(2)	1	8	11	(2)	2	9	10
(3)	7	8	9	(3)	2	5	12	(3)	3	4	11
(4)	10	11	12	(4)	3	6	9	(4)	1	5	7

Lattice rectangular 4 x 5

Bloq.	Grupo X				Bloq.	Grupo Y			
(1)	1	2	3	4	(1)	5	9	13	17
(2)	5	6	7	8	(2)	1	10	14	18
(3)	9	10	11	12	(3)	2	6	15	19
(4)	13	14	15	16	(4)	3	7	11	20
(5)	17	18	19	20	(5)	4	8	12	16

Bloq.	Grupo Z			
(1)	8	11	15	18
(2)	2	9	16	20
(3)	4	7	14	17
(4)	1	5	12	19
(5)	3	6	10	13

Lattice rectangular 5 x 6

Bloq.	Grupo X					Bloq.	Grupo Y				
(1)	1	2	3	4	5	(1)	6	11	16	21	26
(2)	6	7	8	9	10	(2)	1	12	17	22	27
(3)	11	12	13	14	15	(3)	2	7	18	23	28
(4)	16	17	18	19	20	(4)	3	8	13	24	29
(5)	21	22	23	24	25	(5)	4	9	14	19	30
(6)	26	27	28	29	30	(6)	5	10	15	20	25

Bloq.	Grupo Z				
(1)	7	13	19	25	27
(2)	5	14	16	23	29
(3)	1	8	20	21	30
(4)	2	9	15	22	26
(5)	3	10	11	17	28
(6)	4	6	12	18	24

Figura A11. Estructuras en lattice rectangulares no balanceados para experimentos con 12, 20 y 30 tratamientos (Cochran y Cox, 1978).

1^{RA} EDICIÓN

PRECISIÓN EN LOS EXPERIMENTOS AGRÍCOLAS

Algunas Consideraciones y Alternativas

Bibliografía



- Box G., Hunter J. y Hunter W. (2008). *Estadística para investigadores: diseño, innovación y descubrimiento* (2da. Ed.). Barcelona, España: Reverté.
- Cochran G. y Cox G. (1978). *Diseños experimentales*. México: Trillas.
- De La Loma J. (1966). *Experimentación agrícola* (2da. Ed.). México: Unión Tipográfica Editorial Hispano Americana.
- Escobar J. (1981). *Elementos estructurales de un experimento agrícola*. Colombia: Centro Internacional de Agricultura Tropical CIAT.
- (1982). *La heterogeneidad del suelo y los ensayos de uniformidad*. Colombia: Centro Internacional de Agricultura Tropical CIAT.
- Federer, W. (1961). Augmented design with one-way elimination of heterogeneity. *Biometrics*, 17(1): 447 - 473.
- Fisher R., (1950). *Contributions to mathematical statistics*. Nueva York: John Wiley & Sons, Inc.
- García R. (2004). *Inferencia estadística y diseño de experimentos*. Argentina: Editorial Universitaria de Buenos Aires.
- Gavilánez et. al. (2017). Tamaño de la parcela de los experimentos agrícolas. *El Misionero del agro*, 12(1): 04 – 12.
- Gavilánez F. (2015). *Diseños y análisis estadísticos para experimentos agrícolas*. Madrid, España: Editorial Díaz de Santos.
- González G. (1979). *Métodos estadísticos y principios de diseño experimental*. Quito, Ecuador: Universidad Central.
- Gutiérrez H. y De la Vara R. (2008). *Análisis y diseño de experimentos* (2da. Ed.). México: McGraw – Hill Interamericana.
- Hinkelmann K. y Kempthorne O. (2005). *Design and analysis of experiments: Advanced experimental design*. EE. UU.: Wiley – Interscience.
- Jones B. y Kenward. M. (1989). *Design and analysis of cross-over trials* (3era. Ed.). Reino Unido: Chapman and Hall.
- Kuehl R. (2001). *Diseño de experimentos* (2da. Ed.). México: International Thomson Editores S. A.
- Little T. y Jackson F. (1989). *Métodos estadísticos para la investigación en la agricultura*. México: Trillas.
- Mayor V., Blair M. y Muñoz J. (2012). Metodología para determinar el coeficiente de heterogeneidad del suelo, el número de repeticiones y el tamaño de parcela en investigaciones con frijol (*Phaseolus vulgaris* L.). *Acta Agronómica*, 61(1): 32 – 39.
- Montgomery D, Peck E. y Vining G. (2006). *Introducción al análisis de regresión lineal*. México: Compañía Editorial Continental.
- Montgomery D. (2005). *Diseño y análisis de experimentos*. México: Limusa Wiley.

- Osorio F. e Ibarra E. (1987). *Métodos estadísticos aplicados a la investigación agrícola*. Honduras: Instituto Interamericano de Cooperación para la Agricultura IICA.
- Pantoja C. (1989). Algunos criterios y procedimientos estadísticos en la experimentación en fincas de agricultores. Ramakrishna B, *II curso corto. Pruebas en fincas* (pp. 218 – 236). Quito: Programa Cooperativo de Investigación Agrícola para la Subregión Andina – PROCIANDINO.
- Paterson, H. D. & Williams, E. R. (1976). A new class of resolvable incomplete block designs. *Biometrika*, 63(1): 83 – 92.
- Pedroza H., (1993). *Fundamentos de experimentación agrícola*. Nicaragua: Centro de estudios de ecodesarrollo para el Trópico.
- Perrín R., Winkelmann D., Moscardi E. y Anderson J. (1983). *Formulación de recomendaciones a partir de datos agronómicos: Manual metodológico de evaluación económica*. México: Centro Internacional de Mejoramiento de Maíz y Trigo – CIMMYT.
- Quiroga V. (1976). *Manual práctico para el análisis de experimentos de campo*. Costa Rica: Instituto Interamericano de Ciencias Agrícolas IICA.
- Reyes P. (1984). *Diseño de experimentos aplicados* (2da. Ed.). México: Editorial Trillas.
- Roselló J. y Fernández M. (1993). *Guía técnica para ensayos de variedades de campo*. Italia: FAO.
- Smith H. (1938). An empirical law describing heterogeneity in the yield of agricultural crops. *J. Agric. Sci.* 28: 1 – 23.
- Steel R. & Torrie J. (1960). *Principles and procedures of statistics: with special reference to the biological sciences*. Nueva York: McGraw – Hill.
- Tineo J. y Villasmil J. (1988). Determinación del tamaño óptimo de parcela experimental en yuca (*Manihot esculenta* Crantz). *Revista de la Facultad de Agronomía*, 7(2): 116 – 126.
- Vargas J. y Navarro J. (2014). Determinación de un tamaño adecuado de unidad experimental, utilizando el método de curvatura máxima, para ensayos de arroz (*Oryza sativa*), en Bagaces, Guanacaste. *Revista de las Sedes Regionales*, 15(31): 128 – 144.
- Yates F. (1936). Incomplete randomized blocks. *Annals of Eugenics*, 7: 121 – 140.

1^{RA} EDICIÓN

PRECISIÓN EN LOS EXPERIMENTOS AGRÍCOLAS

Algunas Consideraciones y Alternativas

Autor

Freddy Gavilánez Luna

ISBN: 978-9942-654-06-9



Usted es libre de:

Compartir — copiar y redistribuir el material en cualquier medio o formato.

Adaptar — remezclar, transformar y construir a partir del material para cualquier propósito, incluso comercialmente.



PRECISIÓN EN LOS EXPERIMENTOS AGRÍCOLAS

