

Experimentación en agricultura



JUNTA DE ANDALUCÍA
CONSEJERÍA DE AGRICULTURA Y PESCA

EXPERIMENTACIÓN EN AGRICULTURA

R. FERNÁNDEZ ESCOBAR

A. TRAPERO

J. DOMÍNGUEZ

Sevilla, 2010



Fernández Escobar, Ricardo

Experimentación en agricultura / R. Fernández Escobar, A.
Trapero y J. Domínguez. -- Sevilla: Consejería de Agricultura y Pesca,
Servicio de Publicaciones y Divulgación, 2010
350 p. : gráf., tablas, diagr. ; 24 cm. -- (Agricultura: formación)
D.L. SE-1877-2010
ISBN 978-84-8474-281-4

Investigación. -- Experimentación. -- Agricultura
Trapero Casas, Antonio
Domínguez Giménez, Juan
Andalucía. Consejería de Agricultura y Pesca
631.001.4

EXPERIMENTACIÓN EN AGRICULTURA

© Edita: Junta de Andalucía.
Consejería de Agricultura y Pesca

Publica: Secretaría General Técnica
Servicio de Publicaciones y Divulgación

© de los textos: Autores

Serie (Agricultura: formación)

Depósito Legal: SE-1877-2010

ISBN: 978-84-8474-281-4

Diseño, Maquetación e Impresión: Ideas, Exclusivas y Publicidad. S.L.

RICARDO FERNÁNDEZ ESCOBAR

Catedrático de Pomología
Departamento de Agronomía
Escuela Técnica Superior de Ingenieros Agrónomos y de Montes
Universidad de Córdoba

ANTONIO TRAPERO CASAS

Catedrático de Patología Vegetal
Departamento de Agronomía
Escuela Técnica Superior de Ingenieros Agrónomos y de Montes
Universidad de Córdoba

JUAN DOMÍNGUEZ GIMÉNEZ

Investigador Coordinador
Área de Producción Agraria
IFAPA, Centro “Alameda del Obispo”, Córdoba

PRÓLOGO

Una de las mayores dificultades que se le presenta a un estudiante de segundo ciclo cuando ha de abordar un trabajo de investigación es, probablemente, el tratamiento estadístico de los datos. Esto suele ser así por dos motivos, el primero porque es consciente de su escasa formación en esta materia, y el segundo porque suele ser inconsciente de que el problema suele presentarse antes de comenzar el experimento, aunque es en el momento del análisis de los resultados cuando aparece la dificultad. El problema, no obstante, es general y no exclusivo de los estudiantes, pues muchos de los artículos que son rechazados en revistas científicas de impacto en el campo de la agronomía (y también en otros campos afines), lo son debido a un mal planteamiento de los experimentos o a un mal análisis estadístico de los datos. En el ámbito profesional el problema se agrava aún más, pues muchos técnicos con responsabilidad en la transferencia de tecnología, extensionistas agrarios o responsables de la dirección técnica de empresas productoras de insumos o de explotaciones agrarias, no suelen aplicar métodos estadísticos sencillos para determinar si una técnica nueva, un producto fitosanitario o una nueva variedad superan o no a lo habitual de la zona. El valor de la experiencia para juzgar la bondad de nuevas técnicas o las ventajas de un nuevo material, suele anteponerse al rigor de un análisis matemático, en muchas ocasiones por falta de pericia en el manejo de estos métodos.

Esta obra se ha concebido para aliviar las dificultades de investigadores, técnicos y estudiantes poniendo a su disposición una descripción y discusión de procedimientos para el diseño y establecimiento de experimentos, así como para facilitar el cálculo, el análisis y la interpretación de los datos de una forma sencilla y práctica sin grandes exigencias en conocimientos matemáticos. En definitiva, se ha tratado de manejar conceptos y métodos disponibles en la actualidad para un correcto planteamiento de un experimento o de una simple prueba comparativa y del análisis e interpretación de los datos. Para facilitar la comprensión de los procedimientos se expone el cálculo manual del análisis de datos y, para facilitar y agilizar el trabajo rutinario, se recurre a un programa estadístico, *Statistix Version 8.0* (Analytical Software, Tallahassee, FL, USA) que, en opinión de los autores, cumple el requisito de sencillez, resulta asequible, y su contenido abarca la práctica totalidad de las necesidades que requiere un experimentador agrícola.

En la actualidad existen magníficos libros sobre diseños de experimentos y análisis de datos, pero no suelen despejar las dificultades aludidas al principio, al menos en la experiencia que tienen los autores de esta publicación tras muchos años de dedicación a la docencia, a la investigación y a la formación de personal. A ello hay que añadir que esos libros, por lo general, solo tratan de aspectos estadísticos, fundamentales para el diseño de los experimentos y el análisis de datos, pero hay aspectos de la experimentación agrícola que escapan del ámbito matemático, como la definición de los objetivos, la selección de los tratamientos, la toma de datos o la interpretación y presentación de los resultados. Estos aspectos han sido incluidos en este texto como complemento al componente estadístico.

La estructura de la obra se ha basado en los más de 25 años de impartición de una asignatura en los cursos de doctorado, y actualmente en los másteres, en la Escuela Técnica Superior de Ingenieros Agrónomos y de Ingenieros de Montes de la Universidad de Córdoba, que en los últimos años se ha denominado "*Métodos Experimentales en Agronomía*". A lo largo de estos años se ha ido diseñando el perfil de esta materia y ajustándola a las necesidades reales de los alumnos, hasta llegar a la estructura que se recoge en esta publicación. En este momento, es justo mencionar la labor pionera del Profesor Luis Rallo Romero, que inició la impartición de esta asignatura con el afán altruista de ofrecer a sus alumnos una formación básica, de la que carecían, para la realización de sus trabajos experimentales. La perseverancia en el mantenimiento de la asignatura, acompañada por la buena aceptación por parte del alumnado y el consiguiente éxito en la matriculación, hizo que otros profesores nos integrásemos en las enseñanzas y mantuviéramos vigente el objetivo con el que el Profesor Rallo la concibió.

La obra se ha organizado en 24 capítulos y tres apéndices. Los dos primeros capítulos tratan sobre aspectos a considerar en el diseño y planteamiento de los experimentos; los tres siguientes abordan los procedimientos para la comparación de dos o más muestras, introduciendo al lector en el análisis de la varianza. A continuación se describen los diseños más comunes en la experimentación agrícola, que incluyen una introducción sobre el tipo de diseño, la aplicación práctica, el establecimiento de un experimento diseñado de esa forma y el análisis de los datos, tanto manual como con el programa *Statistix*. Le siguen capítulos dedicados al estudio de la correlación y las regresiones, así como al análisis de covarianza, de tanta utilidad al trabajar con plantas perennes. Un capítulo se dedica a los métodos no paramétricos, importantes en experimentos agrícolas y tradicionalmente ignorados por los investigadores. Para finalizar, se incluyen capítulos dedicados a la toma de datos en campo, las medidas en plantas y la interpretación y presentación de resultados tal como se exigen actualmente en las revistas de impacto. La obra termina con tres apéndices. El primero es un glosario que resulta de utilidad para aquellos menos familiarizados con la materia. El número de términos podría ampliarse considerablemente, pero se ha realizado un esfuerzo de síntesis evitando repeticiones de algunos conceptos claramente definidos en los capítulos. El segundo describe brevemente las distribuciones más comunes, para que el curioso pueda entender algo más sobre la distribución con la que trabaja. El tercero es una recopilación de tablas estadísticas, necesarias para el cálculo manual.

Tenemos la convicción de que esta publicación resultará de gran utilidad a estudiantes y profesionales, que requieren de estas herramientas para el desarrollo de sus trabajos. Con ese espíritu la hemos preparado y en la esperanza de su utilidad confiamos.

Primavera de 2009
LOS AUTORES

ÍNDICE

1. La experimentación en Agricultura	13
USOS Y ABUSOS DE LA ESTADÍSTICA	13
COMO REALIZAR UN BUEN EXPERIMENTO	15
EXPERIMENTOS EN EXPLOTACIONES COMERCIALES	17
2. Estructura y diseño de un experimento	19
DEFINICIÓN DE EXPERIMENTO: CLASES	19
UNIDAD EXPERIMENTAL Y TRATAMIENTOS	20
ERROR EXPERIMENTAL	21
CONTROL DEL ERROR EXPERIMENTAL	21
LAS REPETICIONES Y SUS FUNCIONES	24
FACTORES A TENER EN CUENTA PARA ELEGIR EL NÚMERO DE REPETICIONES	24
SORTEO	25
EMPLEO DE FILAS GUARDA	26
INFERENCIAS ESTADÍSTICAS	27
3. Comparación de dos muestras	29
MUESTRAS PAREADAS	29
MUESTRAS INDEPENDIENTES DE IGUAL TAMAÑO CON $\sigma_1=\sigma_2$	32
MUESTRAS INDEPENDIENTES DE DISTINTO TAMAÑO CON $\sigma_1=\sigma_2$	35
MUESTRAS INDEPENDIENTES CON $\sigma_1\neq\sigma_2$	38
CONSIDERACIONES FINALES	39
4. Análisis de varianza	41
MODELOS DE ANÁLISIS DE VARIANZA	41
PROCEDIMIENTO GENERAL DEL ANÁLISIS DE VARIANZA	43
Modelo I: efectos fijos	43
Modelo II: efectos aleatorios	47
SUPUESTOS DEL ANÁLISIS DE VARIANZA	48
Normalidad	48
Homogeneidad de las varianzas (Homoscedasticidad)	50
Independencia de medias y errores	52
Aditividad	53
TRANSFORMACIONES DE LOS DATOS	56
Logarítmica [$\log(Y)$]	56
Raíz cuadrada [\sqrt{Y}]	56
Inversa [$1/Y$]	57
Angular o Arcoseno $\arcsen \sqrt{Y/100}$	57
Escalas pretransformadas	57
Otras transformaciones	58
Datos periféricos o raros (“outliers”)	58
REALIZACIÓN DEL ANÁLISIS DE VARIANZA	58
CÁLCULO CON EL PROGRAMA <i>Statistix</i> (SX)	61

5. Separación de medias y contrastes	67
COMPARACIONES <i>A PRIORI</i>	68
Contrastes ortogonales	68
Contrates polinómicos	71
Comparaciones con un control	73
Comparaciones con el mejor	74
COMPARACIONES <i>A POSTERIORI</i>	75
Método de la Mínima Diferencia Significativa (<i>MDS</i> o " <i>LSD</i> ")	76
Método de <i>Tukey</i>	78
Procedimientos de <i>Bonferroni</i> y de <i>Sidak</i>	80
Procedimiento de <i>Scheffé</i>	82
Otros métodos de comparación múltiple	84
6. Diseño completamente aleatorio	85
EL MODELO LINEAL ADITIVO	85
DISEÑO Y ANÁLISIS DE VARIANZA	86
CÁLCULO CON EL PROGRAMA <i>Statistix (SX)</i>	89
7. Diseño en bloques al azar	93
MODELO LINEAL PARA UN DISEÑO EN BLOQUES COMPLETAMENTE AL AZAR	94
DISEÑO Y ANÁLISIS DE VARIANZA	94
CÁLCULO CON EL PROGRAMA <i>Statistix (SX)</i>	98
8. Cuadrado Latino	101
MODELO LINEAL PARA UN DISEÑO EN CUADRADO LATINO	101
DISEÑO Y ANÁLISIS DE VARIANZA	102
CÁLCULO CON EL PROGRAMA <i>Statistix (SX)</i>	107
9. Diseño aumentado	111
DISEÑO DEL EXPERIMENTO	111
ANÁLISIS	113
10. Experimentos factoriales	121
DISEÑOS EXPERIMENTALES Y MODELOS DE ANÁLISIS DE VARIANZA	122
REALIZACIÓN DEL ANÁLISIS DE VARIANZA FACTORIAL	127
CÁLCULO CON EL PROGRAMA <i>Statistix (SX)</i>	130
11. Diseño en parcelas divididas	135
FUNDAMENTOS DEL ANÁLISIS DE VARIANZA	136
REALIZACIÓN DEL ANÁLISIS DE VARIANZA	139
CÁLCULO CON EL PROGRAMA <i>Statistix (SX)</i>	142
12. Diseño en bloques divididos	149
FUNDAMENTOS DEL ANÁLISIS DE VARIANZA	150
REALIZACIÓN DEL ANÁLISIS DE VARIANZA	153
CÁLCULO CON EL PROGRAMA <i>Statistix (SX)</i>	157

13. Análisis de varianza combinado	167
OBSERVACIONES MÚLTIPLES	167
Análisis de muestreos	168
Análisis de mediciones temporales repetidas	172
SERIES DE EXPERIMENTOS	177
Análisis de diferentes épocas	177
Análisis de diferentes años	180
Análisis de diferentes localidades	184
ANÁLISIS DE EXPERIMENTOS DE LARGA DURACIÓN	190
14. Correlación lineal	193
CÁLCULO DEL COEFICIENTE DE CORRELACIÓN LINEAL	194
INTERPRETRACIÓN DE LOS VALORES DE r	195
CONSIDERACIONES FINALES	197
15. Regresión lineal	199
CÁLCULO DE LA REGRESIÓN LINEAL	200
RELACIONES CON LA CORRELACIÓN	203
CÁLCULO CON EL PROGRAMA <i>Statistix (SX)</i>	203
ESTIMACIONES DE LA REGRESIÓN POBLACIONAL	204
AJUSTE DE LA RECTA POR EL ORIGEN	207
COMPARACIÓN DE LÍNEAS DE REGRESIÓN	210
CONSIDERACIONES FINALES	214
16. Regresión múltiple	215
CORRELACIÓN PARCIAL Y MÚLTIPLE	215
REGRESIÓN CON DOS VARIABLES INDEPENDIENTES	218
REGRESIÓN CON MÁS DE DOS VARIABLES INDEPENDIENTES	226
CONSIDERACIONES SOBRE LA REGRESIÓN MÚLTIPLE	230
17. Regresión curvilínea	233
CURVAS DE TIPO LOGARÍTMICO Y EXPONENCIAL	233
Curvas de tipo logarítmico	234
Curvas de tipo exponencial	237
Otras curvas de tipo logarítmico o exponencial	239
CURVAS DE TIPO POLINÓMICO	240
CÁLCULO CON EL PROGRAMA <i>Statistix (SX)</i>	245
18. Análisis de covarianza	247
REALIZACIÓN DEL ANÁLISIS DE COVARIANZA	248
CÁLCULO CON EL PROGRAMA <i>Statistix (SX)</i>	253
COVARIANZA MÚLTIPLE	256
CÁLCULO CON EL PROGRAMA <i>Statistix (SX)</i>	261
CONSIDERACIONES SOBRE EL ANÁLISIS DE COVARIANZA	261

19. Métodos no paramétricos	263
PRUEBA DE LOS SIGNOS	263
PRUEBA DE RANGOS (CATEGORÍAS) CON SIGNOS DE WILCOXON	264
PRUEBA DE LA SUMA DE RANGOS (CATEGORÍAS) DE WILCOXON	266
PRUEBA DE KRUSKAL-WALLIS	270
PRUEBA DE FRIEDMAN	274
CORRELACIÓN DE RANGOS DE SPEARMAN	277
20. Toma de datos y medidas en campo	281
MÉTODOS DE MUESTREO	282
Muestreo aleatorio simple	283
Muestreo sistemático	285
ESTIMACIONES Y CATEGORÍAS	286
MEDIDAS INDIRECTAS	288
21. Medidas en plantas leñosas	289
EXPERIMENTACIÓN CON PLANTAS LEÑOSAS	289
Selección y calibración de plantas leñosas	290
La parcela elemental	290
MEDIDAS DEL CRECIMIENTO VEGETATIVO	291
Medidas del crecimiento nuevo	291
Medidas del tamaño del árbol	292
Otras medidas del crecimiento vegetativo	293
MEDIDAS DE LA PRODUCTIVIDAD	293
Medidas de la floración	294
Medidas del cuajado de frutos	295
Medidas de la producción	296
MEDIDAS DE CALIDAD DE LA COSECHA	296
Medidas del tamaño del fruto	296
Medidas del color del fruto	297
22. Medidas en plantas herbáceas	299
MEDIDAS DURANTE EL CICLO VEGETATIVO	299
MEDIDAS DE LA COSECHA	300
23. Interpretación y presentación de resultados	301
CONSIDERACIONES GENERALES	301
PRESENTACIÓN DE RESULTADOS	303
Experimento en bloques al azar con medida de una variable cualitativa	304
Experimento en bloques al azar con medida de una variable cuantitativa y un factor no cuantitativo	305
Experimento en bloques al azar con medida de una variable cuantitativa y un factor cuantitativo	306
Experimento factorial con interacción	307
Experimento factorial sin interacción	308
Presentación de regresiones	308
Presentación de datos por la media y el error estándar de la media	310

Apéndices	313
1.- Conceptos estadísticos	313
2.- Distribuciones más comunes	319
Distribución normal	319
Distribución de medias de muestras	320
Distribución t de Student	321
Distribución F	322
Distribución binomial	322
Distribución chi-cuadrado (χ^2)	323
3.- Tablas estadísticas.	325
Tabla A 1. Distribución de la t de Student.	325
Tabla A 2. Distribución F.	327
Tabla A 3. Distribución de χ^2 (Chi-cuadrado).	329
Tabla A 4. Coeficientes de correlación r.	330
Tabla A 5. Número de signos iguales requeridos para obtener Significación en la prueba de los signos.	332
Tabla A 6. Prueba de categorías con signos de Wilcoxon.	333
Tabla A 7. Prueba de la suma de categorías de Wilcoxon (Prueba de Mann-Whitney).	334
Tabla A 8. Coeficientes de correlación de Spearman r_s	336
Tabla A 9. Transformación $\arcsen\sqrt{\text{Porcentaje}/100}$ expresada en grados.	337
Tabla A 10. Coeficientes a_{n-i+1} para el test de Shapiro-Wilk.	338
Tabla A 11. Valores críticos (W) del test de Shapiro-Wilk.	340
Tabla A 12a. Valores críticos para el test de rachas (R_1).	342
Tabla A 12b. Valores críticos para el test de rachas (R_2).	343
Tabla A 13. Valores críticos (d) del test de Dunnett.	344
Tabla A 14. Valores críticos (Q) del test de Tukey.	346
Tabla A 15. Coeficientes para contrastes polinómicos ortogonales (con igual espaciamiento entre los niveles de la variable independiente)	348
 Bibliografía	 349

CAPÍTULO 1

LA EXPERIMENTACIÓN EN AGRICULTURA

La agricultura está siempre sujeta a continuos cambios. En la actualidad se está asistiendo a la transición desde una agricultura convencional, cuya productividad ha estado basada en la aplicación masiva de productos químicos para el control de la salud y la productividad de los cultivos, hacia una agricultura sostenible basada en la obtención de una producción de calidad competitiva respetando el medio ambiente y conservando los recursos naturales. Estos cambios se logran gracias al avance de los conocimientos generados por la investigación, que permiten desarrollar técnicas apropiadas para esos fines. No obstante, gran parte de la práctica agrícola presenta aún una base empírica.

El desarrollo de nuevas técnicas aplicables en la agricultura pasa necesariamente por la experimentación. No basta hoy en día con que destacados técnicos agrícolas establezcan sus impresiones sobre la bondad de una nueva técnica o la utilización de un determinado material, sino que las recomendaciones transferibles deben estar soportadas por datos sujetos a un riguroso examen. Como muchos ya conocen, es hoy en día difícil, si no imposible, que un trabajo de investigación sea aceptado en una revista agronómica de cierto impacto sin la utilización de procedimientos estadísticos adecuados para el diseño de los experimentos y el análisis de los datos.

El objetivo de la experimentación es obtener datos fiables que permitan establecer comparaciones entre tratamientos diferentes y apoyar o rechazar hipótesis de trabajo. El proceso experimental comprende diversas etapas hasta su conclusión. De forma breve, el proceso exige la definición del problema a resolver, el establecimiento de los objetivos, la selección correcta de los tratamientos a aplicar, del material vegetal a emplear en el experimento, el diseño experimental, la toma correcta de datos, su análisis y la interpretación y presentación correcta de los resultados. No basta, pues, con recopilar y presentar datos, sino que hay que obtenerlos de forma correcta y buscarle un sentido a los mismos. El componente estadístico representa tan solo una parte, aunque importante, de la experimentación; es la herramienta útil para el diseño correcto del experimento y el análisis de los datos.

USOS Y ABUSOS DE LA ESTADÍSTICA

El empleo de los métodos estadísticos resulta de interés y de importancia tanto para el investigador como para los técnicos que apliquen los conocimientos tecnológicos resultantes de las investigaciones. Hay que tener en consideración que en la agricultura, la complejidad existente en las relaciones entre el medio de cultivo y las plantas es de tal magnitud que escapan al control simple de una fórmula sencilla. Como indican Little and Hills (1991), por muy profundos conocimientos que se tengan sobre un cultivo, no es posible predecir con exactitud cual será la producción que se obtendría en determinadas condiciones de la misma

manera que lo sería calcular el área de un círculo conociendo su radio. Aún teniendo todo bajo control, cualquier variación aleatoria alteraría los resultados de la predicción. Es, por ello, por lo que resulta difícil contestar a las preguntas planteadas en una experimentación con absoluta seguridad, aún tras varios años de estudio, pues siempre se corre el riesgo de llegar a conclusiones incorrectas.

La estadística es una parte de las matemáticas que se fundamenta en la teoría de probabilidades, cuyos teoremas son tan precisos como otros teoremas matemáticos, pero cuyos resultados son probabilísticos y, en consecuencia, con el riesgo de aceptar o de rechazar una hipótesis incorrectamente. En estos términos hay que entenderse en agricultura, de ahí la importancia de que tanto investigadores como técnicos tengan un conocimiento básico de esas teorías. Si bien, como se ha especificado, es difícil la publicación de trabajos científicos sin un planteamiento estadístico correcto, es bien cierto que en el ámbito de la transferencia de tecnología y de la divulgación no sea frecuente el análisis estadístico de los datos, como si en esta fase final de aplicación del conocimiento el rigor de los trabajos y la interpretación de los resultados no fuera algo esencial. Esto hay que extenderlo también al ámbito de la empresa agraria, donde se requiere experimentar para la simple comprobación del efecto de una nueva técnica o del comportamiento de una nueva variedad. Se asiste, pues, hoy en día tanto a situaciones en las que el uso de la estadística es algo circunstancial, como también a aquellas en las que se trata de aplicar para solucionar problemas asociados a una mala planificación del experimento o, incluso, para adornar trabajos que rayan la mediocridad. En cualquiera de estos casos se está desaprovechando una técnica que, si no absolutamente precisa en la conclusión, los rígidos fundamentos matemáticos en los que se basa la hacen vital para la interpretación correcta de los datos.

Aún en los casos en los que la estadística trata de aplicarse racionalmente, los casos de mal uso de los procedimientos son más frecuentes de lo que podría imaginarse, sin que parezca que pueda ponerse freno a una práctica que conlleva una mala interpretación de los resultados obtenidos. Aunque este mal uso parece contagiado en muchos campos científicos, en el caso de la agronomía se vienen denunciando desde hace más de tres décadas casos de artículos científicos publicados en revistas internacionales de cierto impacto que presentan datos analizados estadísticamente, pero de forma tan incorrecta que pueden alterar las conclusiones obtenidas. Trabajos publicados en esas mismas revistas por Little (1978), Gates (1991) y Dyke (1997), por citar algunos a lo largo de casi dos décadas, ponen de manifiesto las deficiencias en el empleo de los métodos estadísticos en trabajos publicados. La lectura de estos artículos es recomendable a cualquier investigador. Sin tratar de resumir sus contenidos, se puede decir que los errores más frecuentemente encontrados son, sin que se establezca un orden de preferencia, la confusión entre el error experimental y el error de muestreo, lo que a veces lleva a diseñar experimentos sin repeticiones; el análisis incorrecto de experimentos factoriales, donde interviene más de un factor en estudio y se trata de observar si existe o no interacción entre ellos; el análisis de los datos de un diseño experimental que no se corresponde con el diseño establecido; el abuso de los procedimientos de comparación múltiple para la separación de medias, aplicados incluso cuando únicamente se comparan dos medias y también para comparar distintos niveles de un factor cuantitativo, que han de ser separados por regresión; y la ausencia de transformaciones de los datos cuando es un requisito requerido para el análisis de varianza.

En la actualidad el problema no se ha minimizado sino que, incluso, parece haberse complicado por el uso de paquetes informáticos. No cabe duda de la utilidad que representan estas técnicas para el manejo de datos, pues ahorran mucho tiempo y ganan precisión al evitar muchos errores de cálculo, pero hay que evitar la actitud de pensar que se trata de herramientas inteligentes que conocen lo que el investigador ha realizado por el simple hecho de introducirle los datos obtenidos. Con una misma hoja de datos cualquier paquete informático es capaz de analizarlos de muy diversas formas, dependiendo de la orden de cálculo que le demos, pero es evidente que la orden correcta es la que corresponde al diseño experimental del que proceden los datos. La falta de conocimientos sobre los procedimientos de cálculo, que hace años los investigadores solían tener porque tenían que realizarlo manualmente, conduce con frecuencia a la elección de procedimientos inadecuados. Si a esto se añade la cantidad de información que suelen generar los programas una vez analizados los datos, que a veces confunden al inexperto, es fácil imaginar que la expresión final de los resultados se aleje peligrosamente de la realidad en muchas ocasiones.

El rechazo de artículos científicos para su publicación en revistas especializadas tiene su origen, con frecuencia, en un uso inapropiado de los métodos estadísticos, y aún parece que son pocos los rechazados si se tiene en cuenta que muchos evaluadores carecen de conocimientos sólidos de esos procedimientos. Esto ha llevado a proponer a los Comités Editoriales de algunas revistas que dispongan de expertos en estadística para la evaluación de los trabajos presentados (Little, 1978) o, al menos, a exigir que se presenten más datos de los que suelen aportarse, aún a riesgo de aumentar la longitud de los manuscritos (Marini, 1999), de manera que el lector pueda interpretar los resultados por él mismo. Aunque sensatas, estas propuestas no parecen haber tenido aceptación en los Comités Editoriales de las revistas.

CÓMO REALIZAR UN BUEN EXPERIMENTO

El procedimiento para la investigación es el conocido *método científico*, más conocido a veces que comprendido. De forma breve, el método se basa en establecer hipótesis a partir de hechos observados, es decir, formular una idea de cómo se interpretan y se explican esos hechos. Para confirmar si la hipótesis establecida de esa manera es o no cierta, se diseña un experimento que permita probar su validez, y con los datos obtenidos, que aportan nuevos hechos a los ya conocidos, se interpreta si éstos apoyan, rechazan o alteran la hipótesis de partida. En este punto, por lo general, nos encontramos de nuevo al inicio, particularmente si la hipótesis se ha alterado, comenzando otra vez el proceso hasta poder llegar a una conclusión plausible. En el caso más simple de la experimentación de campo, que consiste en comparar una técnica usual con otra nueva, las hipótesis que pueden establecerse son dos, la que considera que ambas dan el mismo resultado y la que considera que ambas difieren en los resultados. Se denomina *hipótesis nula*, y se designa por H_0 , a aquélla que se formula en el sentido de que *no* hay diferencia entre las técnicas, es decir, que las diferencias que puedan observarse se deben a diferencias en el muestreo de la misma población. Esta suele ser la hipótesis de trabajo, en contraposición con la denominada *hipótesis alternativa* (H_1), que es la complementaria de la hipótesis nula, es decir, la que establece que

ambas técnicas difieren. El experimento puede complicarse al incluir diversos métodos de aplicación o al comparar varias técnicas simultáneamente, pero el procedimiento no varía.

Con independencia de lo que se pretende comparar, ya sea la producción, el tamaño del fruto o cualquier otra característica de dos árboles adyacentes, por ejemplo, los datos que se obtengan de cada árbol raramente van a coincidir, aún recibiendo ambos el mismo tratamiento. La diferencia es debida, fundamentalmente, a variaciones ambientales si ambos árboles pertenecen a la misma variedad y están injertados sobre un mismo patrón. Esa variabilidad suele escapar al control del investigador y representa el *error experimental*. En el supuesto de comparar dos técnicas, los resultados obtenidos son una mezcla del efecto de las técnicas y del error experimental, por lo que es necesario estimar éste para aislar el posible efecto de las técnicas. Los métodos estadísticos exigen la *repetición* de los tratamientos para estimar el error experimental, a la vez que *aleatoriedad*, es decir, que cada árbol del ejemplo tenga la misma probabilidad de recibir un determinado tratamiento. Esa manera de proceder asegura un procedimiento objetivo de evaluación de los datos, y al diseñar un experimento hay que pretender reducir en lo posible el error experimental para magnificar las posibles diferencias entre los tratamientos.

Al planificar un experimento hay que tener presente que las consideraciones estadísticas son importantes, pero no las únicas. El diseño del experimento debe hacer practicable los trabajos experimentales y no añadir variabilidad que aumente el error experimental. Por ejemplo, en una experimentación con árboles éstos deben disponerse a marcos adecuados si han de plantarse para el experimento, no a marcos más estrechos para ahorrar espacio o se provocará un crecimiento anormal de los mismos que alterará los datos experimentales. Salvado esto, el diseño debe ser correcto estadísticamente; en este sentido hay que evitar la actitud de pensar que cualquier diseño es válido porque cualquier experto en estadística o cualquier programa informático es capaz de resolverlo todo. La falta de repetición y de aleatoriedad es, con frecuencia, un aspecto de un mal diseño estadístico, lo que a veces se justifica por cuestiones prácticas.

La *simplicidad* es un aspecto que hay que considerar al planificar un experimento. Algunos investigadores se empeñan en emplear diseños complejos porque piensan que les proporcionarán mayor grado de información. Sin embargo, la información hay que obtenerla de la forma más sencilla y fácil posible y aumentar la complejidad en caso necesario, por requerimiento de la planificación del experimento o de los objetivos perseguidos, no para argumentar mayor precisión o mejor planteamiento del experimento. En adición a ello, el experimento debe tener *precisión* y *sensibilidad* para distinguir las posibles diferencias entre los tratamientos y evitar errores sistemáticos, es decir, que determinadas unidades experimentales que reciban un tratamiento difieran sistemáticamente de las que reciben otro tratamiento. Por último, hay que tener presente que la repetición de un experimento de campo en el tiempo y en el espacio permitirá aumentar la validez de las conclusiones que puedan obtenerse del mismo. En muy pocas ocasiones los resultados procedentes de un único experimento de campo pueden ser concluyentes.

Como conclusión, Pearce (1976) sugiere que un buen experimento debe responder a las siguientes cuestiones: ¿es practicable?, ¿es estadísticamente correcto?, ¿son los tratamientos

realmente lo que se proponen ser? Cualquier fallo en alguno de esos aspectos puede ser desastroso para la investigación que se pretende desarrollar.

EXPERIMENTOS EN EXPLOTACIONES COMERCIALES

Idealmente los experimentos deberían realizarse en fincas experimentales propias de centros de investigación y desarrollo. En estas explotaciones, el investigador puede controlar no solamente lo que sería la parte experimental, sino también todas las facetas relativas al manejo del cultivo. El personal que trabaja en esas fincas experimentales conoce lo que es un experimento, lo que puede representar un bloque, distingue el cultivo de una planta experimental de lo que sería el cultivo en una explotación comercial, está familiarizado con la toma de datos en cada parcela experimental y, en definitiva, forma parte de un equipo investigador.

En muchas ocasiones, sin embargo, los experimentos han de realizarse en explotaciones comerciales. Los motivos son variados, pero hay dos razones generales que hace obligado la experimentación en esas explotaciones. La primera es que en las fincas experimentales no existan las condiciones de medio necesarias para cubrir el objetivo de la experimentación. En muchos estudios sobre fertilización se requiere cultivar las plantas en suelos que sean pobres en un determinado elemento, por ejemplo, potasio; si no se dispone de una finca experimental con suelos deficientes en ese elemento, la única alternativa es buscar una explotación comercial con suelos de esas características.

Los experimentos de transferencia de tecnología constituyen la segunda y, a veces, la principal razón para realizar experimentos en explotaciones comerciales. En este tipo de experimentos se persigue, por lo general, evaluar y divulgar una nueva técnica que ha de ser comparada con la práctica habitual en la zona. Esto obliga a realizar los experimentos fuera de las fincas experimentales y repetirlos en varias localidades o zonas, con el objetivo de obtener mayor cobertura sobre la zona de trabajo y conseguir mayor divulgación de la técnica al establecer ensayos demostrativos con varios agricultores de la zona.

En cualquiera de los casos, existen diferencias sustanciales entre los experimentos realizados en fincas experimentales y los realizados en explotaciones comerciales. En primer lugar, éstos se desarrollan con la participación del agricultor, que es quien controla el manejo de la parcela experimental en lugar del investigador o de su equipo. Esto puede contribuir más a la variabilidad que la misma planta o el suelo en el que se cultiva, por lo que la selección del agricultor es esencial para que el experimento pueda finalizar satisfactoriamente. Han sido muchos los ensayos abandonados por el que escribe estos párrafos debido a la falta de interés del agricultor por el ensayo, o a su excesivo entusiasmo con el mismo, que le ha llevado, en ocasiones, a tratar también las plantas testigo antes de finalizar el experimento. En la selección del agricultor es importante que éste sea cooperativo con el experimento, que lo entienda y comprenda los objetivos, de manera que pueda identificarse con ellos y los vea de utilidad en su explotación. Asimismo, sería aconsejable que estuviera al corriente de los resultados, en la conciencia de que son parciales y no concluyentes hasta que se dé por finalizado el experimento, pues esto le haría sentirse, como de hecho lo es, parte del equipo investigador.

Otra diferencia sustancial de los experimentos en explotaciones comerciales, fundamentalmente los de transferencia de tecnología, es que por lo general el número de tratamientos es más limitado y la parcela experimental es de mayor tamaño, idealmente de un tamaño tal que permita realizar las operaciones habituales de cultivo. Parcelas de entre una y 15 ha, han sido utilizadas en ensayos de este tipo.

En los ensayos de transferencia de tecnología, normalmente se elige una zona objetivo y dentro de ella se seleccionan un número de explotaciones comerciales, cuyos criterios de selección variarán en función de los objetivos del experimento. Como todos los factores no están controlados, es conveniente aumentar en lo posible el número de explotaciones que participen en el experimento, lo que además permite soportar la posibilidad de anular una de ellas sin que peligre el experimento completo. Se ha sugerido (Petersen, 1994) que el número de explotaciones a seleccionar en un ensayo básico con dos tratamientos (nueva técnica frente a la habitual en la zona) sea tal que los grados de libertad del error en el análisis de varianza sea, al menos, de 10. Conforme aumente el número de tratamientos puede disminuir el número de explotaciones seleccionadas, que nunca deberían ser inferiores a cuatro. En cada explotación se realizará una única repetición de cada tratamiento y los datos experimentales se analizarán como un diseño en bloques al azar en el que cada explotación representa un bloque. De esta forma, las diferencias entre tipo de explotaciones, incluidas las debidas al manejo diferente de cada una, se acumularán en los bloques y no en los tratamientos. En este tipo de experimentos las parcelas elementales suelen ser grandes, como se ha indicado, deben ser iguales en tamaño y forma en todos los tratamientos de cada explotación y estar lo más cerca posible, aunque no hay necesidad de que sean adyacentes. Los tratamientos se sortean, como de costumbre, para asignarles una parcela a cada uno de ellos.

La técnica de disponer un bloque en cada explotación se pone en práctica en la asunción de que los resultados esperados sean similares en las distintas explotaciones, algo que puede extenderse a experimentos en distintas localidades. Si en éstos se esperase y se observase una interacción localidad x tratamiento, sería necesario establecer un experimento completo en cada localidad.

CAPÍTULO 2

ESTRUCTURA Y DISEÑO DE UN EXPERIMENTO

DEFINICIÓN DE EXPERIMENTO: CLASES

Existen diferentes definiciones de experimento pero, de un modo general, en ciencias biológicas y más, concretamente, en agronomía se puede definir un experimento como “aquel estudio en el que se manipulan deliberadamente una o más variables independientes (supuestas causas) para analizar las consecuencias que esa manipulación tiene sobre una o más variables dependientes (supuestos efectos), dentro de las condiciones controladas por el investigador”. En agronomía, los resultados de este estudio conducirán a tomar decisiones como la recomendación de una variedad, de una concentración de pesticida, de una dosis de abonado, etc.

Los experimentos pueden clasificarse según diferentes criterios; en nuestro caso los clasificaremos en:

- *Experimentos preliminares*: Son aquellos en los que se prueba un gran número de tratamientos con el objeto de obtener indicios para trabajos futuros.
- *Experimentos críticos o decisorios*: Son aquellos en los que el investigador compara las respuestas a diferentes tratamientos, utilizando un número suficiente de observaciones para tener la seguridad razonable de que detecta o no diferencias significativas.
- *Experimentos demostrativos*: Son aquellos en los que se compara uno o más tratamientos nuevos con un testigo, por lo general el tratamiento convencional.

En cualquiera de ellos será necesario definir la población a la que se aplicarán las inferencias, diseñar el experimento apropiado y realizar las medidas de las variables en estudio.

Cada experimento se establece para proporcionar respuestas a una o más preguntas y es el investigador quien debe decidir las comparaciones de tratamientos que proporcionarán las informaciones más relevantes. El experimento se ha de realizar para probar hipótesis relacionadas con diferencias entre tratamientos en condiciones comparables. Una vez hechas las mediciones y las observaciones sobre el material experimental, se responde a las preguntas planteadas al comienzo del experimento.

Así pues, un experimento utiliza un conjunto de reglas usadas para extraer una muestra de una población, por lo que la definición de la población es extremadamente importante. Ese conjunto de reglas es lo que se conoce como el procedimiento experimental o diseño del experimento.

Al diseñar un experimento se han de establecer claramente sus objetivos, en particular:

- Preguntas que han de responderse
- Hipótesis que se han de probar
- Efectos que se han de estimar

Es aconsejable clasificar los objetivos con arreglo a su importancia, ya que algunos diseños experimentales estiman con más precisión ciertas comparaciones entre tratamientos que otros.

UNIDAD EXPERIMENTAL Y TRATAMIENTOS

Una *unidad experimental* es la mínima unidad de material a la que se aplica un tratamiento; puede ser una parcela en el campo (Fig. 2.1), una maceta, una planta, un caldo de cultivo, una solución, media hora, etc. Se conoce también como *parcela elemental*. El *tratamiento* es el procedimiento cuyo efecto se mide y se compara con otros tratamientos, y puede ser una ración alimenticia, una variedad de semillas, un programa de aspersión, la concentración de un fármaco, una combinación temperatura/humedad, etc.

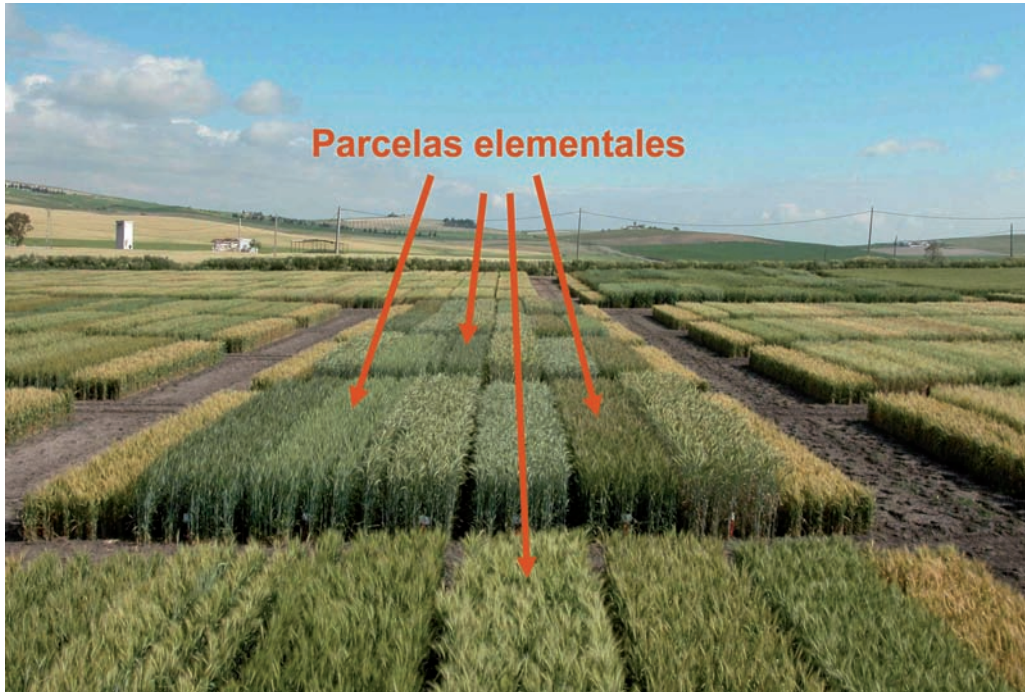


Fig. 2.1. Parcelas experimentales en un campo de ensayo.

Cuando se mide el efecto de un tratamiento se mide en una *unidad de muestreo*, que puede ser una fracción de la unidad experimental. Por lo tanto, la unidad de muestreo puede ser la unidad completa, como un animal sometido a una ración de tratamiento, o una fracción de la misma, como una muestra aleatoria de hojas de un árbol o la cosecha de 6 metros del surco central de una unidad experimental de tres líneas en una parcela en el campo.

Al seleccionar los tratamientos es importante definir cada uno cuidadosamente y considerarlo con respecto a los demás para asegurarse, en lo posible, que el conjunto del experimento proporcione respuestas eficientes relacionadas con los objetivos del mismo.

ERROR EXPERIMENTAL

El error experimental es una medida de la variación existente entre las observaciones realizadas en las unidades experimentales tratadas en forma similar. Esta definición es más sutil de lo que puede parecer a primera vista y se relaciona estrechamente con la definición de lo anterior. Por ejemplo, si se cultivan cinco plantas juntas en la misma maceta y se les aplica un mismo tratamiento, la unidad experimental consiste en las cinco plantas. Se necesitan otras macetas de cinco plantas cada una para poder medir la variación existente entre unidades experimentales tratadas de forma semejante. Esto es cierto aún si una medida, como la altura de la planta, se realiza individualmente. El problema está en que si se van a comparar dos tratamientos, cualquier diferencia observada será, en parte, atribuible a la diferencia entre macetas de cinco plantas y esto es probable que sea de mayor magnitud que las diferencias entre plantas de la misma maceta.

La variación proviene de dos fuentes principales: 1^a.- existe la variabilidad inherente al material experimental al cual se aplican los tratamientos y 2^a.- existe una variación resultante de cualquier falta de uniformidad en la realización física del experimento. En un experimento de nutrición con animales como material experimental, los individuos tendrán constitución genética diferente a menos que haya una alta consanguinidad; ésta es la variabilidad inherente al material experimental. Así pues, el error experimental incluye todos los tipos de variaciones extrañas a los tratamientos empleados.

CONTROL DEL ERROR EXPERIMENTAL

El error experimental puede ser controlado por el experimentador mediante:

1. *El diseño experimental.* El más idóneo es aquel que hace que la variación natural existente entre las unidades experimentales no afecte en nada a las diferencias entre las medias de los tratamientos. Una técnica efectiva para ello es agrupar las unidades experimentales en bloques o grupos de unidades homogéneas, de manera que los tratamientos se asignan al azar a cada unidad experimental dentro de cada bloque. En estas condiciones, en las que cada bloque contiene todos los tratamientos, la variación entre las unidades experimentales dentro de un mismo bloque es menor que entre las unidades de bloques distin-

tos, aumentando la precisión del experimento como resultado de un buen control del error experimental.

En la Fig. 2.2 se muestra la disposición de tres bloques de un experimento en función de la fertilidad del suelo. Si se aprecia un gradiente de fertilidad, la disposición de bloques en el sentido de esa variación permitiría extraer las diferencias entre bloques (en este caso entre diferente fertilidad) del error experimental, aumentando la precisión del experimento al comparar los tratamientos. La disposición contraria haría que dentro de cada bloque existiera una variación de fertilidad que afectaría de forma diferente a cada tratamiento, y que no se podría extraer del error experimental.

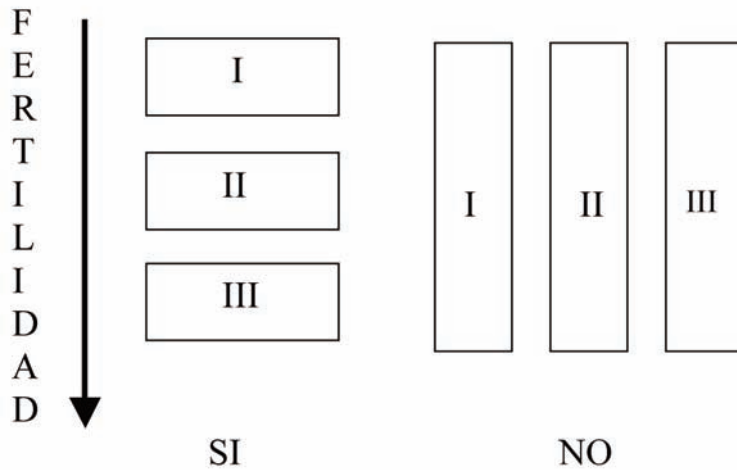


Fig. 2.2.- Disposición correcta e incorrecta de los bloques en un experimento.

Una gran variación entre bloques indica que su uso ha sido de gran ayuda, porque esta variación ha sido eliminada del error experimental y no contribuye a la diferencia entre tratamientos.

- 2. Uso de observaciones paralelas.** En algunos experimentos se puede aumentar la precisión haciendo uso de observaciones paralelas para realizar un análisis de la covarianza. El análisis de la covarianza se utiliza cuando las variaciones entre las unidades experimentales son, en parte, debidas a la variación en otros caracteres mensurables. Por ejemplo, el rendimiento depende del número de plantas por parcela, con independencia del tratamiento que se aplique.
- 3. Tamaño y forma de las unidades experimentales.** Como regla general, unidades experimentales grandes muestran menos variación que las pequeñas. Sin embargo, un aumento en el tamaño de la unidad experimental puede hacer disminuir el número de repeticiones que pueden ser controladas por el experimentador. La unidad experimental debe tener

unas dimensiones mínimas para poder realizar el experimento, pero más allá de esas dimensiones no se gana nada aumentando el tamaño y se obtiene mayor precisión aumentando el número de repeticiones. El tamaño de las unidades dependerá también del tipo de experimento, pues uno demostrativo normalmente exigirá mayores parcelas que uno preliminar. En general, en cultivos anuales suelen utilizarse unidades experimentales compuestas de una a seis filas de cultivo, o aún menos en programas de mejora, en no más de 50 m². Para el caso de especies perennes, véase el Capítulo 21.

En los experimentos de campo, la forma de la parcela experimental es importante en relación con la precisión. Los ensayos de uniformidad realizados por muchos investigadores utilizando varios cultivos, han demostrado que la parcela experimental relativamente larga y estrecha es más conveniente para una mayor precisión, pues cubren una amplitud mayor de condiciones posibles y aumentan la uniformidad dentro de cada bloque. No obstante, también es posible usar otras formas, a veces determinadas por las operaciones de cultivo. En el supuesto de la rectangularidad de las parcelas es conveniente, pues, que se orienten al contrario que lo referido anteriormente para los bloques, es decir, el lado mayor en la dirección de la mayor variabilidad.

En relación con la forma de los bloques, éstos han de tender a lo más cuadrado posible, aunque cualquier forma es correcta siempre que sea efectiva. Lo realmente importante en este caso es que exista la menor variabilidad posible dentro de ellos. No hay ninguna necesidad de que sean adyacentes, aunque si lo son facilita el manejo del experimento (Fig. 2.3).

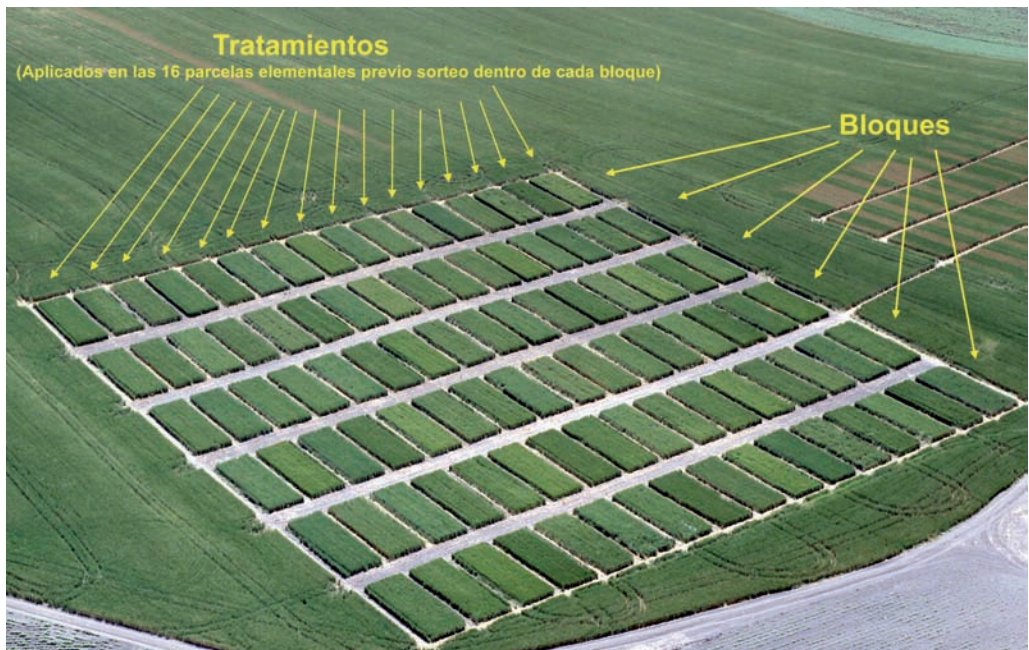


Fig. 2.3. Diseño experimental en bloques al azar con 6 bloques adyacentes y 16 tratamientos.

LAS REPETICIONES Y SUS FUNCIONES

Cuando un mismo tratamiento aparece más de una vez en un experimento se dice que está repetido. Las funciones de la repetición son: permitir una estimación del error experimental; mejorar la precisión de un experimento mediante la reducción del error estándar de una media de tratamiento; aumentar el alcance de la inferencia del experimento a través de la selección y del uso apropiado de unidades experimentales más variables; y ejercer control sobre la varianza del error.

Para realizar las pruebas de significación y para estimar el intervalo de confianza, es necesario obtener una estimación del error experimental. Un experimento en el cual cada tratamiento aparece sólo una vez, se dice que consiste en una repetición simple. De un experimento como éste no es posible estimar el error experimental, puesto que no es posible explicar si una diferencia observada se debe a una diferencia entre tratamientos o a una diferencia entre unidades experimentales. Por consiguiente, cuando no existe un método para estimar el error experimental, no hay manera de determinar si las diferencias observadas indican diferencias reales o si se deben a la variación inherente.

A medida que el número de repeticiones aumenta, las estimaciones de las medias poblacionales se hacen más precisas. Si se detecta una diferencia de cinco unidades usando cuatro repeticiones, un experimento de dieciséis repeticiones detectará la mitad de esa diferencia, o sea, 2,5 unidades, pues los errores estándar $\sigma/\sqrt{4}$ y $\sigma/\sqrt{16}$ están en proporción 2: 1. En definitiva, al aumentar el número de repeticiones disminuye el error estándar.

Aumentar el número de repeticiones puede significar el uso de material experimental menos homogéneo o una técnica menos cuidadosa, dando así una nueva población principal con un mayor error experimental. Sin embargo, el aumento en el número de repeticiones, por lo general, mejora la precisión, disminuyendo las amplitudes de los intervalos de confianza y aumentando el poder de las pruebas estadísticas.

FACTORES A TENER EN CUENTA PARA ELEGIR EL NÚMERO DE REPETICIONES

El número de repeticiones de un experimento depende de varios factores, de los cuales el más importante es el grado de precisión deseada. Cuanto más pequeña sea la discrepancia con respecto a la hipótesis nula que se ha de comprobar, mayor será el número de repeticiones requeridas. En cualquier experimento es, pues, muy importante tener en claro la magnitud correcta de la precisión deseada. No tiene sentido usar diez repeticiones para detectar una diferencia que se puede detectar con cuatro, ni tampoco realizar un experimento en el que el número de repeticiones sea insuficiente para detectar diferencias importantes, excepto ocasionalmente.

Cuando se comparan tratamientos que de antemano se sabe que van a ser muy similares (selecciones genéticas muy parecidas, por ejemplo), se debe aumentar el número de repeti-

ciones para detectar posibles diferencias; sin embargo, cuando se trata de comparar tratamientos que se prevén muy distintos, no será necesario establecer un número elevado de repeticiones (tratamientos de fungicidas y testigos no tratados, por ejemplo).

El número de tratamientos también afecta a la precisión de un experimento, así como al número de repeticiones necesarias para un grado de precisión determinado. Por ejemplo, si se aumenta el número de tratamientos y se mantiene constante el número de repeticiones para cada uno, entonces aumenta el tamaño del experimento y el número de grados de libertad para la estimación de la varianza del error. Si se mantiene constante el tamaño del experimento, entonces un mayor número de tratamientos implicará un menor número de repeticiones de cada uno de ellos y, por lo tanto, un menor número de grados de libertad para estimar la varianza del error. Como resultado se tiene una precisión menor, por lo que se debería aumentar el número de repeticiones para lograr una precisión prefijada. De cualquier manera, este razonamiento es más apropiado para experimentos pequeños, por ejemplo, con menos de 20 grados de libertad en el error.

El diseño experimental también afecta a la precisión de un experimento y al número de repeticiones necesarias. Cuando el número de tratamientos es grande y resulta obligado usar unidades experimentales más heterogéneas, aumenta el error experimental. Con diseños experimentales apropiados se puede controlar parte de esa variación.

Desafortunadamente, el número de repeticiones puede estar determinado, en parte, por los fondos y el tiempo disponible para el experimento, por lo que se ha de llegar a un equilibrio entre fondos, tiempo y precisión de modo que se minimicen los primeros y se maximice la última. No obstante, siempre hay que considerar que la falta de fondos o de tiempo no constituye una excusa para realizar un mal experimento.

SORTEO

La razón del sorteo no es más que asegurar la obtención de un error experimental que no esté sesgado. Esto se logra asegurando que un tratamiento en particular no estará consistentemente favorecido o discriminado en cada una de las repeticiones. En otras palabras, cada tratamiento deberá tener las mismas probabilidades de ser asignado a cualquier unidad experimental.

Sortear es algo análogo a una póliza de seguros, por la que se toman precauciones para una eventualidad que puede o no puede ocurrir. Aquellos diseños en donde los tratamientos se aplican de una forma prefijada y no al azar, a veces resultan en una estimación más alta o más baja que la adecuada del error experimental; asimismo, la precisión en la comparación entre medias se ve alterada.

Esto es particularmente importante en los ensayos de campo. Numerosos estudios han demostrado que las parcelas adyacentes tienden a tener rendimientos más parecidos que parcelas más distantes. Tales parcelas producen un “error correlacionado o residuo”. Como resultado de ello, si los tratamientos están dispuestos en el mismo orden sistemático en

cada repetición, puede haber considerables diferencias en la precisión de las comparaciones entre varios tratamientos. Normalmente resulta en una precisión mayor para los tratamientos más cercanos. El sorteo destruye la correlación entre los errores y hace válidos los tests de significación.

EMPLEO DE FILAS GUARDA

Las plantas que se encuentran en el exterior de las unidades experimentales suelen tener un comportamiento distinto al de aquellas que se encuentran en su interior, en el supuesto de que no haya otras plantas que rodeen a las del experimento. Esto es debido a varios factores, entre los que se encuentran una mejor exposición a la iluminación, a la humedad del suelo, los nutrientes, etc. Para evitar que ese efecto aumente la variabilidad, se suelen disponer filas guardas que rodeen las unidades experimentales, de manera que las plantas del experimento siempre se encuentren en el interior de las parcelas (Fig. 2.4). En el caso de la figura se han dispuesto tanto filas guarda externas, que rodean el conjunto del experimento, como internas, aislando cada unidad experimental.

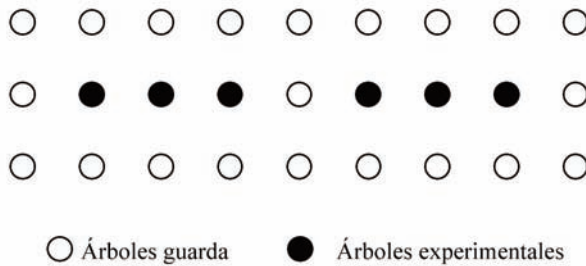


Fig. 2.4.- Disposición de árboles guarda en dos unidades experimentales de un experimento.

Las filas guardas dispuestas de esa forma también evitan el riesgo de afectar a plantas de tratamientos adyacentes, en particular cuando éstos consisten en la aplicación de productos que puedan derivar fácilmente por el aire o por el suelo. En estos casos, es corriente disponer de filas guardas dobles (Fig. 2.5). Si no es de temer ese efecto, basta en muchos casos con disponer filas guardas externas rodeando los bloques o el experimento si aquéllos se disponen de forma contigua.



Fig. 2.5.- Diseño experimental con cuatro bloques (indicado en diferentes colores), siete tratamientos, parcela elemental de cuatro árboles y filas guardas dobles internas y externas.

INFERENCIAS ESTADÍSTICAS

Como se ha visto, el objetivo de los experimentos es determinar si existen diferencias reales entre las medias de los tratamientos y estimar la magnitud de estas diferencias en caso de que existan. Inferir estadísticamente acerca de las diferencias obtenidas lleva consigo el asignar una magnitud de probabilidad a esa inferencia. Por ello, es necesario que sean introducidas la repetición y la aleatoriedad apropiadas al caso, ya que las repeticiones aseguran formas de calcular el error experimental y el sorteo asegura una medida válida de dicho error.

Escoger entre un experimento con el debido sorteo y uno con tratamientos sistemáticos, que aparentemente lleva consigo una mayor precisión, es como escoger entre un itinerario del que se conoce su estado y longitud y otro de la que lo único que se conoce es que es más corto.

CAPÍTULO 3

COMPARACIÓN DE DOS MUESTRAS

En la investigación es muy común evaluar las diferencias entre tratamientos distintos. En el caso sencillo de comparar únicamente dos tratamientos, se puede realizar la investigación de dos formas:

- 1.- *Muestras pareadas*. En este caso se seleccionan individuos o cosas de dos en dos, es decir, por pares, de forma que a un miembro de cada par se le aplica un tratamiento y al otro miembro el segundo tratamiento. En ocasiones, el mismo individuo proporciona los datos para dos muestras distintas que se pretenden comparar; por ejemplo, se puede medir el contenido de clorofila de una planta antes y después de una aplicación de un producto químico que se supone puede afectar al contenido en clorofila.
- 2.- *Muestras independientes*. Es un caso muy común en el que se desea comparar las medias de dos poblaciones y para ello se toman muestras de cada una de ellas independientemente. Debido a la naturaleza del muestreo y de las poblaciones, el tamaño de las muestras puede ser igual o distinto y las varianzas respectivas pueden, asimismo, ser iguales o diferentes.

MUESTRAS PAREADAS

La Tabla 3.1 recoge los datos de un experimento en el que se pretendía conocer la forma más efectiva de aplicación del nitrógeno para corregir deficiencias de este elemento en plantas jóvenes de olivo. Para ello, plantas autoenraizadas de olivo se trasplantaron a macetas de 2 litros de capacidad, que se colocaron en un invernadero durante varias semanas y se regaron con una solución nutritiva con deficiencia en nitrógeno. Transcurrido un tiempo de cultivo en esas condiciones, se seleccionaron plantas por pares, de manera que una planta de cada par recibió una aplicación adicional de nitrógeno al suelo para su absorción radical, y la otra recibió la misma cantidad de nitrógeno pero en aplicación foliar. La Tabla 3.1 muestra el contenido en nitrógeno de cada planta al finalizar el experimento.

Tabla 3.1. Contenido en nitrógeno (mg/planta) en plantas jóvenes de olivo tras la aplicación de nitrógeno vía foliar o radical.

Par	Foliar X ₁	Radical X ₂	Diferencia D = X ₁ - X ₂	Desviación d = D - \bar{D}	d ²
1	62,9	52,7	10,2	5,62	31,58
2	60,0	54,8	5,2	0,62	0,38
3	87,3	78,6	8,7	4,12	16,97
4	81,0	59,7	21,3	16,72	279,56
5	70,7	56,0	14,7	10,12	102,41
6	65,0	72,8	-7,8	-12,38	153,26
7	67,1	73,0	-5,9	-10,48	109,83
8	58,9	41,4	17,5	12,92	166,93
9	66,5	71,4	-4,9	-9,48	89,87
10	51,7	64,9	-13,2	-17,78	316,13
Total	671,1	625,3	45,8	0	1266,9
Media	67,11	62,53	4,58		

En el supuesto de que las desviaciones $D_i - \mu_D$ sean normal e independientemente distribuidas con la media poblacional de cero, la cantidad

$$t = \frac{\bar{D} - \mu_D}{s_{\bar{D}}}$$

sigue la distribución t de Student con (n-1) grados de libertad y puede utilizarse para probar la hipótesis nula de $\mu_D=0$ y calcular el intervalo de confianza de μ_D . Los cálculos son los siguientes:

$$s_D = \sqrt{\frac{\sum (D_i - \bar{D})^2}{n-1}} = \sqrt{\frac{1266,9}{9}} = 11,86$$

$$s_{\bar{D}} = \frac{s_D}{\sqrt{n}} = \frac{11,86}{\sqrt{10}} = 3,75$$

$$t = \frac{\bar{D}}{s_{\bar{D}}} = \frac{4,58}{3,75} = 1,22^{NS}$$

El valor 1,22 es menor de 2,262, valor en las tablas de $t_{0,05}$ para $(n-1)=9$ grados de libertad (Tabla A1), lo que no permite rechazar la hipótesis nula $\mu_D=0$. Los resultados indican, por consiguiente, que la forma de aplicación del nitrógeno no influye en la cantidad absorbida de este elemento en plantas jóvenes de olivo.

El intervalo de confianza al 95% para μ_D es:

$$\bar{D} \pm t_{0,05} s_{\bar{D}} = \bar{D} \pm (2,262)(3,75) = \bar{D} \pm 8,48$$

luego:

$$-3,9 \leq \mu_D \leq 13,06$$

Cálculo con el programa *Statistic (SX)*

Una vez introducidos los datos experimentales, seguir la secuencia:

Statistics>One, Two, Multi-Sample Tests>Paired T Test

y seleccionar las variables en estudio. En el ejemplo, se desea probar la hipótesis nula de que la $\mu_D=0$, esto es, que no existe diferencia entre las medias, siendo la hipótesis alternativa el que sean diferentes (*Not Equal*). Pinchar en OK y aparece la tabla:

Paired T Test for X1 - X2	
Null Hypothesis:	difference = 0
Alternative Hyp:	difference <> 0
Mean	4.5800
Std Error	3.7519
Mean - H0	4.5800
Lower 95% CI	-3.9075
Upper 95% CI	13.067
T	1.22
DF	9
P	0.2532
Cases Included	10
Missing Cases	0

En la tabla aparece el valor de la media de la diferencia (D), el error estándar de la media ($s_{\bar{D}}$), el intervalo de confianza, el valor de t, los grados de libertad y la significación (el valor de P).

El empleo de muestras pareadas en el curso de la experimentación viene impuesto muchas veces por las variaciones en el ambiente. Así, en condiciones de campo puede ser interesante poner los tratamientos uno al lado de otro para evitar variaciones respecto al suelo (aunque esto no siempre es posible). En condiciones de invernadero o de cámaras de crecimiento el método es de gran utilidad, pues si los tratamientos se aplican a macetas contiguas, o a grupos de macetas, se evitan diferencias ambientales, particularmente en temperatura y humedad, más que si las macetas se encuentran separadas, pues en esas condiciones de cultivo las variaciones ambientales pueden ser acusadas en espacios cortos.

MUESTRAS INDEPENDIENTES DE IGUAL TAMAÑO CON $\sigma_1 = \sigma_2$

Con dos muestras independientes normalmente distribuidas, con medias \bar{X}_1 y \bar{X}_2 , que estiman sus respectivas medias poblacionales μ_1 y μ_2 , las pruebas de significación están basadas en la distribución t de Student de la forma:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{s_{\bar{X}_1 - \bar{X}_2}}$$

siendo el denominador una estimada del error estándar de $(\bar{X}_1 - \bar{X}_2)$.

Se conoce que la varianza de una diferencia es la suma de las varianzas siempre que las variables estén distribuidas independientemente, de forma que:

$$\sigma_{\bar{X}_1 - \bar{X}_2}^2 = \sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2$$

Puede demostrarse que, aplicando esto a las dos medias \bar{X}_1 y \bar{X}_2 sacadas de poblaciones de varianza σ^2 , cada media tiene de varianza $\frac{\sigma^2}{n}$, siendo n el tamaño de la muestra. En consecuencia:

$$\sigma_{\bar{X}_1 - \bar{X}_2}^2 = \frac{2\sigma^2}{n}$$

Cada muestra proporciona una estimación de σ^2 , representadas por s_1^2 y s_2^2 . En muestras del mismo tamaño:

$$s^2 = \frac{s_1^2 + s_2^2}{2}$$

con $2(n-1)$ grados de libertad, es decir, la suma de los grados de libertad de s_1^2 y s_2^2 .

El error estándar de la diferencia entre las medias, como se indicó anteriormente, es:

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{2s^2}{n}}$$

y

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{s_{\bar{X}_1 - \bar{X}_2}}$$

El ejemplo siguiente recoge los datos de un experimento en el que se evaluó el contenido de $N-NO_3^-$ en un suelo cultivado con olivar en muestras tomadas de parcelas diferentes; en una de ellas no se aplicó nitrógeno y en la otra se aplicó 1 kg de nitrógeno por olivo durante varios años. Los resultados se muestran en la Tabla 3.2.

Tabla 3.2. Contenido de $N-NO_3^-$ en el suelo (mg N/kg suelo seco)

	X_1 (0 kg N/árbol)	X_2 (1 kg N/árbol)
	11.23	30.72
	14.55	62.15
	10.28	51.97
	33.83	51.13
	28.63	53.03
	8.31	28.42
	16.65	51.78
	13.13	52.03
	23.38	58.06
	34.99	58.28
	14.98	33.51
	8.39	69.58
	15.32	48.38
	17.31	40.31
	44.8	57.72
<i>n</i>	15	15
<i>Total</i>	295,78	747,07
\bar{X}	19,719	49,805
s^2	121,50	139,36
<i>gl</i>	14	14

$$s^2 = \frac{s_1^2 + s_2^2}{2} = \frac{121,50 + 139,36}{2} = 130,43 \quad \text{con} \quad \text{g.l.} = 28$$

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{2s^2}{n}} = \sqrt{\frac{2 * 130,43}{15}} = 4,17$$

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{s_{\bar{X}_1 - \bar{X}_2}} = \frac{19,719 - 49,805}{4,17} = -7,21^{***}$$

El alto valor de t, comparado con el de la Tabla A1 para 28 gl ($t_{0,001} = 3,674$), permite rechazar la hipótesis nula a un valor de P de, al menos, el 0,001%; en consecuencia, los resultados indican que las diferencias entre ambas muestras son altamente significativas.

Los límites de confianza al 95% para $(\mu_1 - \mu_2)$ son:

$$\bar{X}_1 - \bar{X}_2 \pm t_{0,05} s_{\bar{X}_1 - \bar{X}_2} = \bar{X}_1 - \bar{X}_2 \pm 2,048 * 4,17 = \bar{X}_1 - \bar{X}_2 \pm 8,54$$

luego:

$$-38,626 \leq \mu_1 - \mu_2 \leq -21,546$$

Cálculo con el programa *Statistic (SX)*.

Una vez introducidos los datos experimentales, seguir la secuencia:

Statistics>One, Two, Multi-Sample Tests>Two-Sample T Test

y seleccionar las variables en estudio. En el ejemplo, se desea probar la hipótesis nula de que $\mu_D=0$, esto es, que no existe diferencia entre las medias, siendo la hipótesis alternativa el que sean diferentes (*Not Equal*). Pinchar en OK y aparece la tabla:

Two-Sample T Tests for X1 vs X2

Variable	Mean	N	SD	SE
X1	19.719	15	11.023	2.8461
X2	49.805	15	11.805	3.0481
Difference	-30.086			

Null Hypothesis: difference = 0
 Alternative Hyp: difference <> 0

Assumption	T	DF	P	95% CI for Difference	
				Lower	Upper
Equal Variances	-7.21	28	0.0000	-38.628	-21.544
Unequal Variances	-7.21	27.9	0.0000	-38.630	-21.542

Test for Equality	F	DF	P
of Variances	1.15	14, 14	0.4006

Cases Included 30 Missing Cases 2

La tabla muestra los valores para ambas variables (media, número de datos, desviación típica y error estándar). El valor de t, con los grados de libertad, la significación (el valor de P) y los límites de confianza los muestra en los supuestos de igualdad o de desigualdad de las varianzas. En el ejemplo se suponen las varianzas iguales, y el programa aporta esta información al final de la tabla, donde aparece que el test para la igualdad de las varianzas no da significación, lo que indica que no se puede rechazar la hipótesis nula de igualdad de las varianzas.

MUESTRAS INDEPENDIENTES DE DISTINTO TAMAÑO CON $\sigma_1 = \sigma_2$

Aunque siempre es preferible comparar muestras de igual tamaño, en ocasiones esto no es posible por falta de datos o por pérdidas accidentales durante el proceso del experimento. En estos casos sería posible la comparación de las muestras siguiendo un análisis parecido al que se realiza cuando las muestras son iguales, pero con algunos cambios.

En el supuesto de que las varianzas sean iguales, tenemos dos muestras de tamaño n_1 y n_2 , con medias \bar{X}_1 y \bar{X}_2 que tienen varianzas $\frac{\sigma^2}{n_1}$ y $\frac{\sigma^2}{n_2}$. La varianza de la diferencia es:

$$\sigma_{\bar{X}_1 - \bar{X}_2}^2 = \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2} = \sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right) = \sigma^2 \left(\frac{n_1 + n_2}{n_1 n_2} \right)$$

La varianza global se obtiene ponderando por los grados de libertad las varianzas de cada muestra. Así, siendo x_1 y x_2 las correspondientes desviaciones respecto a su medias, esto es, $x_1 = X_1 - \bar{X}_1$ y $x_2 = X_2 - \bar{X}_2$,

$$s^2 = \frac{n_1 - 1}{n_1 + n_2 - 2} \left(\frac{\sum x_1^2}{n_1 - 1} \right) + \frac{n_2 - 1}{n_1 + n_2 - 2} \left(\frac{\sum x_2^2}{n_2 - 1} \right) = \frac{\sum x_1^2}{n_1 + n_2 - 2} + \frac{\sum x_2^2}{n_1 + n_2 - 2} = \frac{\sum x_1^2 + \sum x_2^2}{n_1 + n_2 - 2}$$

$$s^2 = \frac{\sum x_1^2 + \sum x_2^2}{(n_1 - 1) + (n_2 - 1)}$$

esto es, basta sumar las sumas de cuadrados de las desviaciones de cada muestra y dividir por la suma de sus grados de libertad.

Como ejemplo se utilizará el ejemplo anterior pero en el supuesto de que en la segunda muestra el tamaño sea de 8 en lugar de 15. Los datos se recogen en la Tabla 3.3.

Tabla 3.3. Contenido de N-NO₃ en el suelo (mg N/kg suelo seco)

	X ₁ (0 kg N/árbol)	X ₂ (1 kg N/árbol)
	11.23	30.72
	14.55	62.15
	10.28	51.97
	33.83	51.13
	28.63	53.03
	8.31	28.42
	16.65	51.78
	13.13	52.03
	23.38	
	34.99	
	14.98	
	8.39	
	15.32	
	17.31	
	44.8	
<i>n</i>	15	8
<i>Total</i>	295,78	381,23
\bar{X}	19,719	47,654
$\sum x^2$	1701,0	962,63
<i>gl</i>	14	7

$$s^2 = \frac{1701,0 + 962,63}{14 + 7} = 126,84 \quad \text{con g.l.}=21$$

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{s^2 \left(\frac{n_1 + n_2}{n_1 n_2} \right)} = \sqrt{126,84 \left(\frac{15 + 8}{15 * 8} \right)} = 4,93$$

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{s_{\bar{X}_1 - \bar{X}_2}} = \frac{19,719 - 47,654}{4,93} = -5,67***$$

que ha resultado, como en el ejemplo anterior, altamente significativo, pues 5,67 es mayor que lo reflejado en la Tabla A1 para 21 gl

Los límites de confianza al 95% para $(\mu_1 - \mu_2)$ son:

$$\bar{X}_1 - \bar{X}_2 \pm t_{0,05} s_{\bar{X}_1 - \bar{X}_2} = \bar{X}_1 - \bar{X}_2 \pm 2,080 * 4,93 = \bar{X}_1 - \bar{X}_2 \pm 10,25$$

luego:

$$-38,189 \leq \mu_1 - \mu_2 \leq -17,685$$

Cálculo con el programa *Statistic (SX)*.

En el programa *SX* se sigue exactamente la misma secuencia que si se tratase de muestras de igual tamaño. Por ello, una vez introducidos los datos experimentales se sigue la secuencia:

Statistics>One, Two, Multi-Sample Tests>Two-Sample T Test

y se seleccionan las variables en estudio. En el ejemplo, se desea probar la hipótesis nula de que $\mu_D=0$, esto es, que no existe diferencia entre las medias, siendo la hipótesis alternativa el que sean diferentes (*Not Equal*). Pinchar en OK y aparece la tabla:

Two-Sample T Tests for X1 vs X2

Variable	Mean	N	SD	SE
X1	19.719	15	11.023	2.8461
X2	47.654	8	11.727	4.1461
Difference	-27.935			

Null Hypothesis: difference = 0
 Alternative Hyp: difference <> 0

Assumption	T	DF	P	95% CI for Difference	
				Lower	Upper
Equal Variances	-5.67	21	0.0000	-38.189	-17.681
Unequal Variances	-5.55	13.6	0.0001	-38.748	-17.122

Test for Equality of Variances	F	DF	P
	1.13	7, 14	0.3977

Cases Included 23 Missing Cases 9

La tabla muestra los valores para ambas variables (media, número de datos, desviación típica y error estándar). El valor de t, con los grados de libertad, la significación (el valor de P) y los límites de confianza los muestra en los supuestos de igualdad o de desigualdad de las varianzas. En el ejemplo se suponen las varianzas iguales, y el programa aporta esta información al final de la tabla, donde aparece que el test para la igualdad de las varianzas no da significación, lo que indica que no se puede rechazar la hipótesis nula de igualdad de las varianzas.

MUESTRAS INDEPENDIENTES CON $\sigma_1 \neq \sigma_2$

Lo más común es presumir que las dos varianzas de población sean iguales, pero hay ocasiones en que no lo son pues las muestras pueden provenir de poblaciones diferentes, la σ puede tender a cambiar cuando cambia μ , o puede haberse muestreado una población sesgada. En estos casos la varianza de $(X_1 - X_2)$ no varía de lo recogido anteriormente:

$$\sigma_{\bar{X}_1 - \bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

y la t resulta:

$$t' = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

cantidad que no sigue la distribución t de Student cuando $\mu_1 = \mu_2$.

Se han desarrollado varias formas de la distribución t'. Siguiendo a Snedecor and Cochran (1974), en el caso de muestras de igual tamaño, $t' = t$, con lo que se calcula la t como se ha visto anteriormente pero se le da n-1 gl en lugar de 2(n-1). Si las muestras son de diferente tamaño, se calcula la t' y el nivel de significación viene dado por la expresión:

$$\frac{w_1 t_1 + w_2 t_2}{w_1 + w_2}$$

siendo t_1 y t_2 los niveles de significación de t para n_1-1 y n_2-1 gl, respectivamente, y $w_1 = \frac{s_1^2}{n_1}$

y $w_2 = \frac{s_2^2}{n_2}$. Si el valor de t' es mayor que el calculado de esta manera, se concluye que la diferencia es significativa.

El programa SX calcula, como ya se ha visto, los valores de t, la significación y los intervalos de confianza en los supuestos de igualdad y de desigualdad de las varianzas, así como la prueba de igualdad de las varianzas para no dudar de los valores a utilizar. No obstante, en caso de duda es preferible asumir que $\sigma_1 \neq \sigma_2$.

La prueba de igualdad de las varianzas se realiza mediante una prueba F de dos colas de forma que:

$$F = \frac{s_1^2}{s_2^2}$$

siendo s_1 la mayor de las dos. Si se conociera que $\sigma_1 > \sigma_2$, se puede utilizar la prueba de una cola, esto es, las mismas tablas de F que se utiliza en el análisis de varianza.

CONSIDERACIONES FINALES

Cuando se trata de comparar dos muestras, como se ha visto en este capítulo, hay que tener muy claro si las muestras son pareadas o independientes, pues los resultados del análisis pueden alterar las conclusiones si se usa el procedimiento inadecuado.

Las muestras pareadas tienen la ventaja de que sus pruebas t no exigen que $\sigma_1=\sigma_2$ y, en ocasiones, el tamaño de las muestras puede ser inferior que si se tratase de muestras independientes. Pero no siempre es posible hacer un emparejamiento efectivo al diseñar un experimento. Hay que tener en cuenta que puede perderse precisión si el criterio para establecer parejas no tiene relación con la variable en estudio, o si los miembros de un par están correlacionados negativamente.

Un problema adicional que se presenta, sean las muestras pareadas o independientes, es que se introduzca un factor aberrante al elegir los individuos o las muestras, que haga que $(\bar{X}_1 - \bar{X}_2)$ no sea una estimada de la diferencia media de población entre los dos tratamientos. Por ejemplo, que al disponer muestras pareadas de plantas en maceta en un invernadero las mayores pertenezcan a X_1 y las menores a X_2 . El investigador debe vigilar que estas situaciones no se produzcan. Una forma útil de evitar ese problema es hacer el muestreo al azar, de manera que en cada par o en cada muestra independiente los tratamientos se asignan por sorteo.

Una última cuestión es conocer cual sería el tamaño de la muestra que impidiera que una diferencia significativa aparezca como que no lo es porque el experimento fue demasiado pequeño. Se han desarrollado procedimientos para estimar el tamaño de la muestra en esas circunstancias (Snedecor and Cochran, 1974), pero a nivel práctico hay que considerar que si no se conoce nada acerca de la variabilidad de la variable en estudio, es preferible tomar al principio un tamaño de la muestra suficientemente grande para evitar ese problema.

CAPÍTULO 4 ANÁLISIS DE VARIANZA

El análisis de varianza o ANOVA (del inglés ANalysis Of VAriance) es una técnica paramétrica utilizada cuando hay más de dos grupos independientes. Se trata de un método para comparar medias, no varianzas como su nombre podría sugerir. Su hipótesis nula (H_0) establece la igualdad entre las medias de los a grupos o poblaciones ($\mu_1 = \mu_2 = \dots = \mu_a$), mientras que la hipótesis alternativa (H_1) establece que al menos una de las medias es distinta. El análisis se completa cuando se acepta la hipótesis H_0 , es decir, no hay diferencias entre grupos. En cambio, cuando se rechaza H_0 , se sabe que hay diferencias entre grupos, pero para conocer en concreto cuáles son esas diferencias es necesario continuar con los procedimientos de separación de medias y contrastes que se detallan en el Capítulo 5.

MODELOS DE ANÁLISIS DE VARIANZA

El modelo matemático que se aplica para el ANOVA en cualquier diseño experimental es el **modelo lineal aditivo**, que se formula en forma de suma:

$$Y_{ij\dots p} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \dots + \varepsilon_{ij\dots r}$$

$i = 1, 2, \dots, a$
 $j = 1, 2, \dots, b$
 $\dots\dots\dots$
 $r = 1, 2, \dots, n$

expresando que el valor de cualquier unidad experimental está compuesto por la media general (μ) más la suma de los efectos de los diferentes factores (α_i, β_j, \dots) y sus interacciones, $[(\alpha\beta)_{ij}, \dots]$, y la suma de los diferentes errores ($\varepsilon_{ij\dots r}$). Además de la aditividad, una característica fundamental de este modelo es que supone que los errores son independientes y se encuentran normalmente distribuidos con medias cero y varianzas iguales, es decir, $\varepsilon_{ij\dots r} \sim N(0, \sigma^2)$ para toda i, j, \dots, r .

La expresión concreta de este modelo varía con el diseño experimental y con el número y tipo de factores que intervienen. Cuando hay un solo factor para analizar se tiene el ANOVA unifactorial o de una vía (One Way ANOVA), mientras que si intervienen dos o más factores resulta el ANOVA multifactorial o factorial a secas. El número de factores debe ser limitado porque al aumentar éste se incrementan las posibles interacciones entre ellos, lo que complica el análisis y su interpretación.

Los factores o variables independientes pueden ser de dos clases, según sean sus efectos:

fijos o aleatorios. Ello da lugar a dos tipos de modelos de ANOVA: modelo de efectos fijos (modelo I) y modelo de efectos aleatorios (modelo II). Ambos modelos varían en las asunciones, en los propósitos y en la interpretación, aunque los cálculos y las pruebas de significación son idénticos.

En el **modelo de efectos fijos**, los diferentes niveles de los factores se seleccionan de antemano, por lo que se consideran como parámetros fijos tales que:

$$\sum_{i=1}^a \alpha_i = \sum_{j=1}^b \beta_j = \dots = 0$$
$$\sum_{i=1}^a (\alpha\beta)_{ij} = \sum_{j=1}^b (\alpha\beta)_{ij} = \dots = 0$$

La hipótesis nula de este modelo establece que todas las medias (factores e interacciones) son iguales, lo que puede escribirse en términos del modelo como:

$$H_0: \alpha_i = 0 \text{ para toda } i$$
$$H_0: \beta_j = 0 \text{ para toda } j$$

.....

$$H_0: (\alpha\beta)_{ij} = 0 \text{ para toda } i \text{ y } j$$

En este modelo I, las inferencias estadísticas respecto a los efectos de los tratamientos pertenecen, de forma exclusiva, a los niveles seleccionados, por lo que no se pueden extraer conclusiones para otros posibles niveles. Este es el caso más común de los experimentos agronómicos, donde generalmente estamos interesados en conocer el efecto de unos niveles concretos de una variable cualitativa o categórica. Cuando en este modelo se incluyen variables independientes cuantitativas (temperatura, tiempo, etc.), se supone que sólo estamos interesados en conocer el efecto de esos niveles concretos, pero no el de toda la población posible de niveles.

En el **modelo de efectos aleatorios**, los diferentes niveles de los factores se seleccionan al azar de una población mayor de posibles niveles. Por tanto, para cada factor (por ejemplo el factor T) de efectos aleatorios, $\tau_1, \tau_2, \dots, \tau_a$ son variables aleatorias independientes con $\tau_j \sim N(0, \sigma_\tau^2)$ para toda j . La hipótesis nula en este modelo consiste en la proposición de que la varianza entre los τ_j (o los efectos del tratamiento) es cero; es decir:

$$H_0 : \sigma_\tau^2 = 0$$

Así, al suponer independencia entre los errores y tratamientos aleatorios, tenemos que la varianza total es:

$$\text{Var}(Y_{ij}) = \sigma^2 + \sigma_\tau^2$$

En este modelo II, el interés recae en conocer cuánto de la varianza en las observaciones ($\text{Var}(Y_{ij})$) se debe a diferencias reales en las medias de los tratamientos (σ_τ^2) y cuánto se debe a errores aleatorios con respecto a estas medias (σ^2). Por ello, las inferencias estadísticas con respecto a los niveles del factor estudiado pertenecen a toda la población posible de niveles de ese factor. El modelo II es menos habitual en los experimentos agronómicos, pero debería considerarse sobre todo cuando se incluyen variables cuantitativas como variable independientes.

Un tercer modelo (modelo III), denominado **modelo mixto**, resulta cuando unos factores son de efectos fijos y otros de efectos aleatorios. Otra consideración sobre los modelos es si el diseño está equilibrado o no. Se dice que está equilibrado cuando cada combinación del modelo contiene el mismo número de casos o repeticiones ($n_1 = n_2 = \dots = n_j$). En cambio, se considera desequilibrado cuando todos los niveles de los tratamientos no contienen el mismo número de repeticiones. Estos modelos desequilibrados requieren cálculos más complicados y dificultan la construcción de los contrastes de hipótesis sobre cada factor, por lo que no suelen utilizarse en los experimentos agronómicos.

PROCEDIMIENTO GENERAL DEL ANÁLISIS DE VARIANZA

Modelo I: efectos fijos

Para explicar el desarrollo y los cálculos a realizar en el ANOVA se va a considerar el caso más simple: ANOVA de una vía para un diseño completamente aleatorio y el modelo de efecto fijos. Tenemos, pues, un solo factor o tratamiento (A) con a niveles ($i = 1, 2, \dots, a$) y varias observaciones (repeticiones) en cada nivel ($j = 1, 2, \dots, n_i$). El número de repeticiones en cada nivel de A no tiene por qué ser igual, pero es conveniente en algunos casos, sobre todo cuando se trabaja con muestras pequeñas (pocas observaciones). Para esos casos: $n_1 = n_2 = \dots = n_i = n$. El número total de casos (N) sería:

$$N = \sum_{j=1}^a n_j = n_1 + n_2 + \dots + n_{a-1} + n_a = n + \dots + n = a \times n$$

Los datos se podrían agrupar como se indica en la Tabla 4.1.

Tabla 4.1. Ordenación de los datos para un experimento con un solo factor (A) en un diseño completamente aleatorio con n repeticiones

Repeticiones	Niveles del factor A (tratamientos)									
	1	2	.	.	.	j	.	.	.	a
1	Y_{11}	Y_{12}	.	.	.	Y_{1j}	.	.	.	Y_{1a}
2	Y_{21}	Y_{22}	.	.	.	Y_{2j}	.	.	.	Y_{2a}
.
.
.
i	Y_{i1}	Y_{i2}	.	.	.	Y_{ij}	.	.	.	Y_{ia}
.
.
.
n	Y_{n1}	Y_{n2}	.	.	.	Y_{nj}	.	.	.	Y_{na}

El modelo lineal aditivo para este supuesto sería:

$$Y_{ij} = \mu + \alpha_j + \varepsilon_{ij}$$

donde, Y_{ij} es la i ésima observación del j ésimo tratamiento, μ es la media general sobre todas las α poblaciones, α_j es el efecto en la respuesta debido al j ésimo tratamiento y ε_{ij} es el error experimental para la i ésima observación bajo el j ésimo tratamiento.

Si $\mu_1, \mu_2, \dots, \mu_a$ son las medias de las α poblaciones y μ es la media de todas las poblaciones, el modelo puede escribirse como:

$$Y_{ij} = \mu + (\mu_j - \mu) + (Y_{ij} - \mu_j)$$

o bien:

$$Y_{ij} - \mu = (\mu_j - \mu) + (Y_{ij} - \mu_j)$$

Los parámetros $\mu_1, \mu_2, \dots, \mu_a$ y μ no son conocidos, pero pueden estimarse en base a las observaciones que tenemos de las α muestras aleatorias, según los estadísticos:

$$\bar{Y}_{\cdot j} = \sum_{i=1}^n \frac{Y_{ij}}{n_j} = \sum_{i=1}^n \frac{Y_{ij}}{n} \qquad \bar{Y}_{..} = \frac{\sum_{j=1}^a \sum_{i=1}^n Y_{ij}}{\sum_{j=1}^a n_j}$$

Al sustituir los parámetros μ_j y μ por sus estimadores $\bar{Y}_{\cdot j}$ y $\bar{Y}_{\cdot\cdot}$, tendremos la expresión:

$$Y_{ij} - \bar{Y}_{\cdot\cdot} = (\bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot}) + (Y_{ij} - \bar{Y}_{\cdot j})$$

que establece que la desviación de una observación (Y_{ij}) respecto al promedio de la muestra total ($\bar{Y}_{\cdot\cdot}$), se divide en dos componentes: la desviación de la media del tratamiento ($\bar{Y}_{\cdot j}$) respecto a la media global ($\bar{Y}_{\cdot\cdot}$), y la desviación de Y_{ij} respecto a la media del tratamiento ($\bar{Y}_{\cdot j}$).

Si en lugar de las desviaciones se toma el cuadrado de ellas, para eliminar el signo, y se suman sobre todos los i y j , se llega a la expresión:

$$\sum_{j=1}^a \sum_{i=1}^n (Y_{ij} - \bar{Y}_{\cdot\cdot})^2 = \sum_{j=1}^a \sum_{i=1}^n (\bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot})^2 + \sum_{j=1}^a \sum_{i=1}^n (Y_{ij} - \bar{Y}_{\cdot j})^2$$

Esta expresión se conoce como la ecuación fundamental del análisis de la varianza. Cada uno de sus términos se denomina suma de cuadrados (SC), pudiendo abreviarse en la expresión:

$$SC_{Total} = SC_{Tratamientos} + SC_{Error}$$

que indica que la variación total de un experimento es una suma de dos componentes: la variación entre grupos o tratamientos, más la variación dentro de cada grupo o error experimental. Si en lugar de un solo factor interviniesen varios (ANOVA factorial), el número de sumandos aumentaría incluyendo la suma de cuadrados de cada factor, de sus interacciones y del error o errores. La descomposición de la suma de cuadrados es, por tanto, aplicable a todos los diseños y modelos de ANOVA.

Igualmente, los grados de libertad (gl) correspondientes a este modelo se descomponen en sumandos según la expresión:

$$gl(SCT) = gl(SCA) + gl(SCE)$$

o bien:

$$N-1 = a-1 + N-a; \quad \text{ó} \quad (an-1) = (a-1) + a(n-1)$$

Dividiendo las sumas de cuadrados por sus respectivos grados de libertad se obtienen los cuadrados medios (CM), que serían por definición varianzas. Se ha demostrado que el cuadrado medio del error (CME) es un estimador no sesgado de la varianza residual del experimento (σ^2), mientras que el cuadrado medio de los tratamientos (CMA) es un estimador de σ^2 más la variación debida a los tratamientos:

$$CME = \frac{SC_{Error}}{a(n-1)} = \sigma^2$$

$$CMA = \frac{SC_{Tratamientos}}{a-1} = \sigma^2 + \frac{\sum_{j=1}^a n_j \times \alpha_j^2}{a-1} = \sigma^2 + \frac{n \sum_{j=1}^a \alpha_j^2}{a-1}$$

Si la hipótesis nula (H_0) fuera cierta, $\alpha_j = 0$ y $\sum n_j \alpha_j^2 = 0$,

entonces CMA sería también un estimador de σ^2 . En cambio, si H_0 no fuera cierta, CMA sería mayor que CME. El cociente CMA/CME tiene una distribución *F de Fisher* con $a-1$ y $N-a$ grados de libertad y es, por tanto, el estadístico apropiado para probar la hipótesis nula. Este cociente será próximo a 1 cuando las medias de los tratamientos sean similares y tanto mayor que 1 cuanto mayores sean las diferencias entre tratamientos.

El valor F obtenido se contrastará con el valor de la distribución teórica $F_{a-1, N-a}$, consultando la Tabla A2 y se obtendrá el valor de P , o probabilidad de encontrar por azar diferencias iguales o mayores a las observadas en la muestra si la hipótesis nula fuera cierta. El nivel de significación (α) o probabilidad con la que se está dispuesto a rechazar la hipótesis nula cuando es cierta, es decir, de cometer un error de tipo I, se ha fijado de forma arbitraria en el 0.05 (5%), aunque también se han utilizado otros niveles: 0.01 (1%), 0.001 (1‰) en función de los objetivos del experimento. Por tanto, si se obtiene un valor de $P < 0.05$, se puede rechazar la hipótesis nula y decir que existen diferencias significativas entre las medias de los tratamientos. En cambio, si $P > 0.05$, no se tienen motivos fundados para rechazar la hipótesis nula y el ANOVA concluye señalando que no existen diferencias significativas entre los tratamientos.

Los resultados del ANOVA se suelen presentar generalmente en forma de tabla, como se indica a continuación para el caso analizado, en el que se ha supuesto el mismo número de repeticiones para todos los niveles del tratamiento ($r_1 = r_2 = \dots = r_j = r$):

Fuente de Variación	gl	SC ⁽¹⁾	CM	F	P ⁽²⁾
Factor A	$a-1$	$\sum_{j=1}^a \sum_{i=1}^r (\bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot})^2$	$\frac{SCA}{(a-1)}$	CMA/CME	$\left. \begin{array}{l} >0.05 \text{ NS} \\ <0.05 \text{ *} \\ <0.01 \text{ **} \\ <0.001 \text{ ***} \end{array} \right\}$
Error	$a(r-1)$	$\sum_{j=1}^a \sum_{i=1}^r (Y_{ij} - \bar{Y}_{\cdot j})^2$	$\frac{SCE}{a(r-1)}$		
Total	$ar-1$	$\sum_{j=1}^a \sum_{i=1}^r (Y_{ij} - \bar{Y}_{\cdot\cdot})^2$			

¹Para el facilitar el cálculo manual, las sumas de cuadrados se calculan mediante una expresión equivalente:

$$\begin{aligned}
 SC_{Total} &= \sum_{j=1}^a \sum_{i=1}^r (Y_{ij} - \bar{Y}_{..})^2 = \sum_{j=1}^a \sum_{i=1}^r Y_{ij}^2 - \frac{(\sum_{j=1}^a \sum_{i=1}^r Y_{ij})^2}{ar} = \sum_{j=1}^a \sum_{i=1}^r Y_{ij}^2 - \frac{Y_{..}^2}{ar} \\
 SC_{Tratamientos} &= \sum_{j=1}^a \sum_{i=1}^r (\bar{Y}_{.j} - \bar{Y}_{..})^2 = \frac{1}{r} \sum_{j=1}^a Y_{.j}^2 - \frac{Y_{..}^2}{ar} \\
 SC_{Error} &= SC_{Total} - SC_{Tratamientos} = \sum_{j=1}^a \sum_{i=1}^r Y_{ij}^2 - \frac{1}{r} \sum_{j=1}^a Y_{.j}^2
 \end{aligned}$$

Al término común $\frac{Y_{..}^2}{ar}$ se le denomina factor de corrección y se abrevia como C .

²En algunos trabajos todavía no se indican los valores de P y en su lugar aparece, después del valor de F , la designación NS, *, **, ***, refiriéndose a diferencias no significativas ($P > 0.05$), significativas ($P < 0.05$) o altamente significativas ($P < 0.01$; $P < 0.001$), respectivamente. Esta es una práctica que debe evitarse actualmente, ya que todos los programas estadísticos ofrecen la posibilidad de calcular el valor exacto de P .

Modelo II: efectos aleatorios

La diferencia con el modelo I de efectos fijos es que los niveles empleados en el experimento no son fijos, sino que fueron seleccionados de forma aleatoria de una población de posibles niveles. Se supone, por tanto, que $\tau_j \sim N(0, \sigma_\tau^2)$, en la que σ_τ^2 es la varianza de los niveles o tratamientos aleatorios.

La descomposición de la suma de cuadrados y los cálculos para el ANOVA son iguales al modelo de efectos fijos, pero en este caso los valores esperados para el cuadrado medio del tratamiento y del error, suponiendo muestras de igual tamaño (n), serían:

$$\begin{aligned}
 E(CME) &= \sigma^2 \\
 E(CMTr) &= \sigma^2 + n\sigma_\tau^2
 \end{aligned}$$

de donde se obtiene que $\sigma_\tau^2 = \frac{CMTr - CME}{n}$

El cociente $CMTr / CME$ tiene el mismo significado que en el modelo de efectos fijos, ya que un valor grande sugiere que debe rechazarse la hipótesis nula ($H_0: \sigma_\tau^2 = 0$). Si $P > 0.05$, concluimos que la variabilidad observada entre tratamientos es debida a errores aleatorios ($\sigma_\tau^2 = \sigma^2$), mientras que si $P < 0.05$, podemos decir que existe variación debida al factor aleatorio. Por tanto, se puede calcular la proporción de la variabilidad total, $Var(Y_{ij})$, debida a los tratamientos aleatorios como:

$$\frac{\sigma_{\tau}^2}{\text{Var}(Y_{ij})} = \frac{\sigma_{\tau}^2}{\sigma^2 + \sigma_{\tau}^2}$$

Con esto concluiría el análisis, ya que las inferencias sobre los tratamientos se extienden a la población total de posibles tratamientos y no interesa realizar comparaciones entre niveles concretos, como ocurre en el modelo de efectos fijos.

SUPUESTOS DEL ANÁLISIS DE VARIANZA

En cualquier ANOVA tenemos siempre dos tipos de variables: la variable medida que se quiere comparar, o variable dependiente (peso de cosecha, altura, tiempo, etc.) y la variable o variables de agrupación, categóricas o independientes (tipo de fertilización o de riego, tratamientos fitosanitarios, etc.).

La variable dependiente debe ser cuantitativa (numérica), bien discreta o continua. En cambio, la variable independiente es únicamente el factor de clasificación de las observaciones en diferentes grupos y suele ser cualitativa (nominal u ordinal). A veces se incluyen variables cuantitativas como independientes (por ejemplo temperatura, concentración de inóculo, dosis de riego o de productos fitosanitarios, etc.). Esto sólo debería utilizarse cuando los niveles ensayados de estas variables sean bajos (≤ 4). En otros casos, el análisis más apropiado para tales datos sería la regresión lineal o curvilínea.

Para que los resultados de un ANOVA puedan considerarse válidos hay que comprobar que se cumplen unas condiciones generales de aplicación. Estas condiciones o supuestos tienen que ver fundamentalmente con la necesidad de que los valores residuales o errores en la variable medida (variable dependiente) sean independientes entre sí y se ajusten a una distribución normal con varianza uniforme, además del efecto aditivo de los factores analizados y errores. No obstante, se acepta generalmente que el análisis de la varianza es una técnica "robusta" porque sus resultados se alteran muy poco si se producen transgresiones de los supuestos en los que se basa. A continuación se comentan cada una de estas condiciones.

Normalidad

Aunque suele decirse que el requisito es que los valores de la variable dependiente presenten una distribución normal o de Gauss, lo realmente importante es que la distribución de los residuales se aproxime a una normal. Los residuales son la diferencia entre cada valor y la media de su grupo, es decir, cada uno de los elementos que se suman para construir la suma de cuadrados residual o del error (SCE).

Afortunadamente, las desviaciones del supuesto de normalidad no afectan muy seriamente la validez del ANOVA. No obstante, existen pruebas de normalidad que deben considerarse, sobre todo cuando el número de muestras es elevado.

El *test de normalidad de Shapiro-Wilk* prueba si una muestra aleatoria simple de tamaño n (X_1, X_2, \dots, X_n) procede de una población con distribución normal. El test utiliza el estadístico W que se define como:

$$W = \frac{b^2}{s^2}$$

siendo $b = \sum_{i=1}^{n/2} (X_{(n-i+1)} - X_{(i)}) a_{n-i+1}$ $s^2 = \sum_{i=1}^n (X_i - \bar{X})^2$

donde $X_{(i)}$ denota la i -ésima variable de menor valor una vez ordenados los valores en orden creciente ($X_{(1)}, X_{(2)}, \dots, X_{(n)}$) y a_{n-i+1} son los coeficientes de Shapiro-Wilk que se encuentran en la Tabla A10. La hipótesis nula H_0 establece que la muestra procede de una población normal, y se rechaza si $W < W_{n,\alpha}$, donde $W_{n,\alpha}$ está tabulada (Tabla A11).

Tomemos como ejemplo el de la Tabla 4.2. El procedimiento es el siguiente:

1. Se ordenan los 18 valores de menor a mayor:

70,2; 48,5; 20,3; 16,8; 10,4; 5,9; 5,6; 4,2; 2,8; 2,4; 2,3; 1,6; 1,6; 1,1; 0,8; 0,5; 0,4; 0,2

2. Se calcula el valor de b :

$X_{(n-i+1)} - X_{(i)}$	a_{n-i+1}
70,2 - 0,2 = 70,0	0,4886
48,5 - 0,4 = 48,1	0,3253
20,3 - 0,5 = 19,8	0,2553
16,8 - 0,8 = 16,0	0,2027
10,4 - 1,1 = 9,3	0,1587
5,9 - 1,6 = 4,3	0,1197
5,6 - 1,6 = 4,0	0,0837
4,2 - 2,3 = 1,9	0,0496
2,8 - 2,4 = 0,4	0,0163

A la observación mayor se le subtrae la menor y se multiplica la diferencia por el valor tabulado específico de a . Se hace lo mismo con la segunda observación y la penúltima y así sucesivamente. Los nueve coeficientes a pueden encontrarse en la Tabla A10 para $n=18$.

Luego:

$$b = \sum_{i=1}^9 (X_{(n-i+1)} - X_{(i)}) a_{n-i+1} = 60,573$$

3. Se calcula s^2 :

$$s^2 = \sum_{i=1}^{18} (X_i - \bar{X})^2 = (70,2 - 10,87)^2 + \dots + (0,5 - 10,87)^2 = 6067,38$$

4. Se calcula W :

$$W = \frac{b^2}{s^2} = 0,6047 < W_{18,0,01} = 0,858$$

El valor $W=0,858$ puede encontrarse en la Tabla A11. Como es mayor que el W calculado, se rechaza la hipótesis nula concluyendo que la población no es normal.

En el programa *Statistix*, introducir la variable en forma Categórica y seguir la siguiente secuencia:

Statistic>Randomness/Normality Test>Shapiro-Wilk Test

seleccionar la variable y se obtiene el valor de W y de P .

Otra forma más sencilla de probar la normalidad es representar los datos en el eje de abscisas y la frecuencia acumulada en ordenadas. La línea recta indica normalidad (véase más adelante el cálculo con *Statistix*).

Homogeneidad de las varianzas (Homoscedasticidad)

Una condición necesaria para poder comparar las medias de varios grupos es que las varianzas de cada grupo deben ser iguales, es decir, las dispersiones de los datos no deben diferir entre los grupos a comparar. Promediar las varianzas mayores y menores podría arrojar resultados engañosos, como considerar significativa la diferencia entre dos tratamientos con varianzas grandes cuando en realidad pudo haber ocurrido por casualidad (falsos positivos), o declarar no significativa la diferencia entre dos tratamientos con varianzas pequeñas cuando, de hecho, lo era (falsos negativos).

El *test de Bartlett* es una de las pruebas que más se utilizan para la detección de esta grave transgresión de los supuestos del ANOVA. Nos da la probabilidad "P" de tener diferencias mayores o iguales a las observadas en la muestra si las varianzas fuesen homogéneas. Si $P < 0,05$, se acepta la hipótesis alternativa de que las varianzas no son homogéneas. Para ver el procedimiento general, supongamos que tenemos a medias, cada una con su varianza (s_i^2) y sus grados de libertad (γ_i). Si $\gamma_1 = \gamma_2 = \dots = \gamma_i = \gamma$, y la varianza global es

$$s^2 = \frac{\sum_{i=1}^a s_i^2}{a} \text{ se calcula:}$$

$$M = \gamma \left[a \times \ln(s^2) - \sum_{i=1}^a \ln(s_i^2) \right] \quad y \quad C = 1 + \frac{a+1}{3 \times a \times \gamma}$$

El cociente M/C sigue una distribución χ^2 con $(a-1)$ grados de libertad, luego se calcularía el valor de la χ^2 experimental y se compararía con el valor $\chi^2_{0,05;a-1}$ teórico. Los cálculos para el ejemplo de la Tabla 4.2 son:

$$s_1^2 \text{ (No solarizado)} = \frac{\sum (Y_i - \bar{Y})^2}{r-1} = \frac{(70,2 - 28,68)^2 + \dots + (5,9 - 28,68)^2}{6-1} = \frac{3184,05}{5} = 636,81$$

$$s_2^2 \text{ (Solarizado 1)} = \frac{\sum (Y_i - \bar{Y})^2}{r-1} = \frac{(4,2 - 3,15)^2 + \dots + (2,3 - 3,15)^2}{6-1} = \frac{10,90}{5} = 2,18$$

$$s_3^2 \text{ (Solarizado 2)} = \frac{\sum (Y_i - \bar{Y})^2}{r-1} = \frac{(70,2 - 28,68)^2 + \dots + (5,9 - 28,68)^2}{6-1} = \frac{1,33}{5} = 0,27$$

$$s^2 = \frac{\sum_{i=1}^a s_i^2}{a} = \frac{636,81 + 2,18 + 0,27}{3} = 213,087$$

$$M = 5[3 \times \ln(213,087) - 5,9265] = 50,793$$

$$C = 1 + \frac{3+1}{3 \times 3 \times 5} = 1 + \frac{4}{45} = 1,089$$

$$\chi^2 = \frac{M}{C} = \frac{50,793}{1,089} = 46,64$$

En la Tabla A3 se observa que χ^2 con 2 gl, $(a-1 = 2)$ y $P = 0,05$ es $\chi^2_{0,05;2} = 5,99 < 46,64$, por lo que se concluye que las varianzas no son homogéneas.

La heterogeneidad de las varianzas es una de las transgresiones más comunes e importantes en los ANOVA. Para su corrección, se recomienda realizar transformaciones de los datos cuando dicha heterogeneidad es de tipo regular o no siguen la distribución normal. Si la heterogeneidad es del tipo irregular, se recomienda partir el error en grupos separados con varianzas homogéneas y realizar análisis independientes para cada grupo. Véase más adelante la transformación realizada a los datos y la aplicación del test de Bartlett a los datos transformados.

Independencia de medias y errores

Los residuales o errores, además de ser uniformes (varianzas homogéneas), deben distribuirse aleatoriamente. Un caso frecuente de falta de aleatorización es cuando existe una relación definida entre las medias de las muestras y sus varianzas. Este es un caso especial y la causa más común de heterogeneidad de la varianza.

El *test de rachas* se utiliza para determinar si los datos procedentes de una muestra de observaciones pueden ser aleatorios. En este apartado el test prueba la independencia de medias y varianzas o, expresado de otra forma, si hay falta de aleatoriedad en los errores. En una muestra de tamaño n que ha sido dividida en dos categorías, por ejemplo + y - (que pueden corresponder a machos y hembras; frutos sanos y dañados; etc.) con n_1 y n_2 observaciones cada una, se denomina racha a una sucesión de valores de la misma categoría. El número de rachas R es una variable aleatoria cuya distribución está tabulada para $n_1 < 20$ y $n_2 < 20$ (Tabla A12). La aleatoriedad se rechaza cuando $R \leq R_{n_1, n_2, \alpha/2}$ (Tabla A12a) ó $R \geq R_{n_1, n_2, 1-\alpha/2}$ (Tabla A12b). En otras palabras, hay dos valores críticos de R ; si el valor obtenido se encuentra entre los dos, se acepta la hipótesis nula de que la muestra es aleatoria.

En el ejemplo de la Tabla 4.2, supongamos que se asigna el signo + cuando un valor es mayor que el anterior, y - cuando es menor. Se tiene, pues:

	<i>No solarización</i>			<i>Solarización 1</i>			<i>Solarización 2</i>			
	+ - -	+ - -		- - -	+ - -		- + +	- + -	+ -	
Rachas:	1	2	3	4	5	6	7	8	9	10

Nº de rachas $R = 10$
 n_1 (signos +) = 6
 n_2 (signos -) = 12

Los valores críticos son 4 (Tabla A12a) y 13 (Tabla A12b). Como el valor de R está comprendido entre ambos, no se rechaza la hipótesis nula.

En el programa *Statistix*, los signos + y - se asignan según el dato esté por encima o por debajo de la mediana o de la media. Si consideramos la mediana, cuyo valor es 2,60 tenemos:

	<i>No solarización</i>					<i>Solarización 1</i>				<i>Solarización 2</i>			
	+ + +	+ + +		+ - -	+ + -		- - -	- - -	- - -				
Rachas:	1					2	3	4					

Nº de rachas $R = 4$
 n_1 (signos +) = 9
 n_2 (signos -) = 9

Los valores críticos son 5 y 15, que rechazan la hipótesis nula de que la muestra es aleatoria. Las conclusiones difieren de las obtenidas anteriormente con los mismos datos. En este caso se tienen menos rachas pero más largas, lo que altera el resultado.

En el programa *Statistix*, introducir la variable en forma Categórica y seguir la siguiente secuencia:

Statistic>Randomness/Normality Test>Runs Test

Seleccionar la variable, elegir entre la media o la mediana, y aparece:

Runs Test for ...		
Median		2.6000
Values Above the Median		9
Values below the Median		9
Values Tied with the Median		0
Runs Above the Median		2
Runs Below the Median		2
Total Number of Runs		4
Expected Number of Runs		10.0
P-Value, Two-Tailed Test		0.0060
Probability of getting 4 or fewer runs		0.0030
Probability of getting 4 or more runs		0.9996
Cases Included	18	Missing Cases
		1

Otra forma de detectar esta transgresión de los supuestos del ANOVA es realizar un examen visual de la distribución de los residuos estandarizados sobre los valores ajustados por el modelo ANOVA en un gráfico. Si los residuos no se distribuyen aleatoriamente por encima y debajo de la línea y existe algún patrón (lineal, curvilíneo, triangular, etc.) en dicha distribución, hay evidencia de una falta de independencia de los errores. Véase más adelante la obtención de este gráfico en el ejemplo.

Para corregir este fallo de los supuestos del ANOVA se recomienda realizar una buena aleatorización de las unidades experimentales, o realizar un análisis de covarianza si los datos guardan relación con otra variable que podría ser responsable de esta anomalía.

Aditividad

El modelo matemático que se aplica para el ANOVA en cualquier diseño experimental es el modelo lineal aditivo, como se ha indicado anteriormente, expresando que el valor de cual-

quier unidad experimental está compuesto por la media general más la suma de los efectos de los diferentes factores y sus interacciones y la suma de los diferentes errores.

Cuando el efecto de un factor no es el mismo sobre todos los niveles de un factor aleatorio (por ejemplo: bloques) tenemos un efecto no aditivo, ya que unos bloques responderán de una forma y otros de manera diferente. Otro caso más evidente es cuando el efecto de un factor sobre la variable medida es aumentar o disminuir ésta en un cierto porcentaje o proporción. Este es el caso de un efecto multiplicativo del factor.

Existen también varios test para detectar la falta de aditividad de los factores y errores. Uno de ellos es el *test de no aditividad de Tukey*, que da la probabilidad “P” de tener diferencias mayores o iguales a las observadas suponiendo el modelo aditivo. Si $P < 0.05$, se acepta la hipótesis alternativa de que los efectos no son aditivos.

El método trata de extraer una parte de la suma de cuadrados del error para probar la aditividad con un grado de libertad. Resulta muy útil para saber si es necesaria una transformación de los datos y sugiere la transformación apropiada de la forma $Y=X^p$, donde X está en la escala original y se busca el valor de p para que los efectos sean aditivos. Así, tenemos:

- $p \approx 1$ la transformación no es necesaria
- $p \approx -1$ transformar X en $1/X$
- $p \approx 0,5$ transformación \sqrt{X}
- $p \approx 0$ transformación logarítmica

El método es aplicable a clasificaciones de dos vías, por lo que siguiendo con el ejemplo de la Tabla 4.2, supongamos para este caso que se trata de un experimento en bloques al azar con tres tratamientos y seis bloques. Los pasos a seguir están basados en Snedecor y Cochran (1974) y se detallan a continuación.

Tratamientos	Bloques						X_{ij}	$\bar{X}_{i.}$	d_i	w_i
	1	2	3	4	5	6				
No Solarización	70,2	20,3	16,8	48,5	10,4	5,9	172,1	28,68	17,81	1094,93
Solarización1	4,2	2,4	1,6	5,6	2,8	2,3	18,9	3,15	-7,72	50,113
Solarización2	0,4	0,8	1,6	0,2	1,1	0,5	4,6	0,77	-10,10	-12,777
$X_{.j}$	74,8	23,5	20,0	54,3	14,3	8,7				
$\bar{X}_{.j}$	24,93	7,83	6,67	18,1	4,77	2,9	$\bar{X}_{..} = 10,87$			
d_j	14,06	-3,04	-4,2	7,23	-6,1	-7,97	$\sum d_i = \sum d_j = 0$			

$$d_i = \bar{X}_{i.} - \bar{X}_{..}$$

$$d_j = \bar{X}_{.j} - \bar{X}_{..}$$

$$w_i = \sum X_{ij} d_j$$

$$w_1 = (70,2)(14,06) + (20,3)(-3,04) + \dots + (5,9)(-7,97) = 1094,93$$

$$w_2 = (4,2)(14,06) + (2,4)(-3,04) + \dots + (2,3)(-7,97) = 50,113$$

$$w_3 = (0,4)(14,06) + (0,8)(-3,04) + \dots + (0,5)(-7,97) = -12,777$$

$$N = \sum w_i d_i = (1094,93)(17,81) + (50,113)(-7,72) + (-12,777)(-10,10) = 19226,31$$

$$\sum d_i^2 = (17,81)^2 + (-7,72)^2 + (-10,10)^2 = 478,80$$

$$\sum d_j^2 = (14,06)^2 + \dots + (-7,97)^2 = 377,57$$

$$D = (\sum d_i^2) (\sum d_j^2) = (478,80)(377,57) = 180780,52$$

La suma de cuadrados por no aditividad viene dada por:

$$SC = \frac{N^2}{D} = \frac{19226,31^2}{180780,52} = 2044,75$$

Se puede construir el siguiente análisis de varianza:

Fuente de variación	gl	SC	CM	F
Tratamientos	2	2873,94		
Bloques	5	1133,13		
Error	10	2060,30		
No aditividad	1	2044,75	2044,75	1181,94***
Remanente	9	15,55	1,73	

Como puede apreciarse, se ha restado de la suma de cuadrados del error el valor calculado de la suma de cuadrados por no aditividad, con un grado de libertad, quedando 15,55 restante como remanente cuyo cuadrado medio sirve para calcular la F.

Con F altamente significativa se concluye que el modelo no es aditivo, lo que aconsejaría transformar los datos para que lo sea. Para ello hay que seguir los siguientes pasos:

$$B = \frac{N}{D} = \frac{19226,31}{180780,52} = 0,1$$

$$p = 1 - B\bar{X}_{..} = 1 - (0,1)(10,87) = -0,087$$

El valor de p es próximo a cero, lo que sugiere una transformación logarítmica.

El programa *Statistix* presenta directamente el test de no aditividad en la salida del análisis de varianza de un diseño en bloques al azar y en cuadrado latino.

TRANSFORMACIONES DE LOS DATOS

Una de las soluciones más habituales para lograr que los datos (residuales) se ajusten a una distribución normal es realizar una transformación de la variable dependiente. Esto es particularmente importante cuando la heterogeneidad del error (heterocedasticidad) es de tipo regular, es decir, las varianzas están relacionadas de alguna manera con las medias.

Con frecuencia la transformación logra el doble objetivo de normalidad y aditividad de los datos. En estos casos, los análisis realizados con los datos transformados son perfectamente válidos. Cuando se hace una transformación, todas las comparaciones o las estimaciones de los intervalos de confianza deben hacerse en la escala transformada. Para la presentación de los resultados, las medias deben transformarse volviendo a la escala original, pero no es apropiado hacer lo mismo con los errores estándar o varianzas.

Si existen dudas sobre la transformación más adecuada es necesario examinar la relación entre varianzas y medias y elegir la transformación para la cual esta relación sea mínima. No obstante, cuando no es posible hallar una transformación que normalice los datos, deben usarse otros métodos de análisis, como la ponderación de las medias de acuerdo con sus varianzas o utilizar test no paramétricos.

A continuación se comentan las transformaciones más comunes.

Logarítmica [$\log(Y)$]

El uso de esta transformación, o bien de su recíproca [$-1/\log(Y)$], está indicado cuando los datos presenten una distribución muy sesgada a la derecha o a la izquierda; o bien, cuando las desviaciones típicas (no las varianzas) son proporcionales a las medias, o cuando el efecto de los factores es multiplicativo, en lugar de aditivo. Si existen valores pequeños, menores que 10, y especialmente ceros, es más adecuada la transformación $\log(Y+1)$.

Raíz cuadrada [\sqrt{Y}]

Esta transformación, o su recíproca [$-1/\sqrt{Y}$], es aplicable cuando las varianzas son proporcionales a las medias, lo que ocurre a menudo cuando estamos tratando con conteos o recuentos de acontecimientos poco comunes, cuyos datos tienden a seguir una distribución especial, denominada distribución de *Poisson*. En este caso, los datos presentan una distribución moderadamente sesgada a la derecha, o a la izquierda, respecto a la curva normal. Cuando existen valores por debajo de 10, y especialmente cuando hay ceros, es mejor utilizar la transformación [$\sqrt{Y + 0.5}$].

Inversa [$1/Y$]

Se utiliza para casos poco comunes en los que los datos presentan una elevada variabilidad y las varianzas son proporcionales a las medias elevadas a 4. Los datos presentan una distribución fuertemente sesgada a la derecha (*J*invertida) o a la izquierda (en forma de *J*). Esta transformación también precisa que sumemos una cantidad a cada valor, por ejemplo 1, si existen ceros.

Angular o Arcoseno [$\arcsen \sqrt{Y/100}$]

Cuando los datos son proporciones o porcentajes de la muestra total, tienen una distribución binomial en vez de una distribución normal. En los datos binomiales, las varianzas tienden a ser pequeñas en los dos extremos de los intervalos de valores (ceranos a 0 y a 100%), pero mayores en el medio (alrededor del 50%). La transformación apropiada para este tipo de datos, especialmente cuando las proporciones o los porcentajes cubren un intervalo amplio de valores, es $\text{seno}^{-1}\sqrt{Y}$ o $\text{arcoseno}\sqrt{Y}$. Cuando el intervalo de porcentajes está entre 0 y 20, o bien entre 80 y 100, pero no ambos, se recomienda la transformación raíz cuadrada. En este caso, los porcentajes entre 80 y 100 deberán restarse de 100 antes de hacer la transformación. Si el intervalo de porcentajes se sitúa entre 30 y 70, es dudosa la necesidad de transformación.

Escalas pretransformadas

En numerosos experimentos agronómicos, la variable dependiente es una variable cuantitativa continua que se podría expresar como porcentaje (incidencia de plantas enfermas, tejido afectado, etc.), pero que resulta muy difícil y engorroso su determinación exacta. En estos casos, es una práctica habitual para realizar un mayor número de mediciones en un tiempo dado, hacer estimaciones visuales aproximadas de los porcentajes, en vez de efectuar mediciones precisas. Estas estimaciones se hacen utilizando una escala de 0 a *n*, donde el cero representa ausencia de enfermedad o de control, y *n* indica el 100%. Las escalas varían en función de los objetivos del trabajo, desde simples, que incluyen 3 o 4 valores, hasta escalas más sofisticadas con más de 10 valores. Si los peldaños de dicha escala representan incrementos iguales de porcentajes, los datos deben ser transformados mediante la transformación angular, como se ha indicado para los porcentajes.

Una alternativa a lo indicado sería pretransformar la escala, es decir, escoger escalones de porcentajes tales que al aplicarles la transformación angular resultaran en números enteros de la escala. Otras veces, los valores enteros de la escala se refieren a porcentajes transformados, mediante $\log(Y)$, $\log(Y/(1-Y))$, u otras funciones similares, que asumen el hecho de que es más fácil detectar pequeñas diferencias en las inmediaciones de 0 y 100% que alrededor del 50%. Al analizar los datos basados en este tipo de escalas, éstos no deberían ser transformados, aunque habría que comprobar que se cumplen las características de una distribución normal.

Con las diferentes escalas se convierte una variable cuantitativa continua (porcentaje) en una variable cualitativa ordinal, ya que los datos sólo pueden tener unos cuantos valores enteros, correspondientes a los diferentes peldaños de la escala. Se viola, por tanto, uno de los fundamentos del ANOVA, el que establece que la variable medida debe ser cuantitativa (continua o discreta). Ello ha llevado a algunos autores a proponer otras alternativas de análisis en estos casos, como los métodos no paramétricos. No obstante, el uso de las escalas ha sido muy popular en algunas disciplinas, por lo que han sido y son todavía ampliamente utilizadas en los análisis paramétricos. El gran desarrollo que han experimentado recientemente los métodos no paramétricos, unido a la disponibilidad de software que facilita su aplicación incluso en diseños complicados, posibilitará que en un futuro próximo los datos medidos con escalas se analicen mayoritariamente utilizando tests no paramétricos.

Otras transformaciones

Las transformaciones expuestas anteriormente son las más usadas para los datos de experimentos agronómicos. No obstante, otras transformaciones pueden deducirse tras un cuidadoso examen de los datos. Así, si la distribución de los datos no es simétrica y hay una mayor concentración en el lado derecho y la cola en la izquierda, se puede utilizar la transformación Y^2 , que comprime la escala para valores pequeños y la expande para valores altos.

Datos periféricos o raros (“outliers”)

Cuando se encuentran valores inusualmente grandes o pequeños respecto al conjunto de los valores observados, la opción más correcta es averiguar si se trata de un error y corregirlo. Si hay motivos fundados para sospechar que se trata de un error, pero resulta imposible corregirlo, debería tratarse como dato perdido y realizar el ANOVA con el valor(es) estimado(es) con la consiguiente pérdida de uno o varios grados de libertad. Cuando el “outlier” no se debe a ningún error, existen métodos para determinar si debería o no incluirse en el análisis. En caso de duda, se realizarían los análisis con y sin los “outlier”, con los datos transformados y sin transformar, para ver el efecto de dicho cambio en los resultados. Los programas estadísticos actuales disponen de opciones para visualizar los datos que permiten detectar con facilidad la existencia de estos “outliers” y que deberían consultarse siempre antes de realizar el análisis.

REALIZACIÓN DEL ANÁLISIS DE VARIANZA

Supongamos un experimento sobre el efecto de la solarización del suelo sobre la población de propágulos viables del hongo fitopatógeno *Verticillium dahliae*. En un suelo con una infestación elevada y uniforme del patógeno, se aplicaron tres tratamientos (sin solarizar, solarizado una vez y solarizado dos veces). A los 2 meses de la solarización se tomaron seis muestras de suelo de cada tratamiento (6 repeticiones) y se analizaron por separado, contabilizándose el número de propágulos viables por g de suelo (variable dependiente). Los datos se incluyen en la Tabla 4.2.

Tabla 4.2. Número de propágulos de *V. dahliae* por g de suelo.

Repetición	Tratamiento			Total
	No solarizado	Solarizado 1	Solarizado 2	
1	70,2	4,2	0,4	
2	20,3	2,4	0,8	
3	16,8	1,6	1,6	
4	48,5	5,6	0,2	
5	10,4	2,8	1,1	
6	5,9	2,3	0,5	
$Y_{.j} = \sum_{i=1}^6 Y_{ij}$	172,1	18,9	4,6	195,6
$Y_{.j}^2 = \sum_{i=1}^6 Y_{ij}^2$	8117,59	70,45	4,86	8192,90
$\bar{Y}_{.j} = \sum Y_{ij} / 6$	28,68	3,15	0,77	$\bar{Y}_{..} = 10,87$

Antes de proceder con el análisis, algo que destaca en los datos es su heterogeneidad: los valores correspondientes a las muestras no solarizadas son muy superiores a los de las muestras solarizadas y, además, presentan una variación mayor. Por tanto, hay que comprobar la homogeneidad de las varianzas. La aplicación del test de Bartlett realizado en un apartado anterior, mostró que las varianzas no eran homogéneas. Los datos del ejemplo indican un cierto efecto multiplicativo de los tratamientos. Por ello, si se aplica la transformación logarítmica se obtiene la Tabla 4.3 con los datos transformados, a los que se puede aplicar de nuevo el test de Bartlett para comprobar si ha tenido efecto la transformación.

Tabla 4.3. Número de propágulos de *V. dahliae* por g de suelo (datos transformados)*

Repetición	Tratamiento			Total
	No solarizado	Solarizado 1	Solarizado 2	
1	3,85	2,62	1,60	
2	3,31	2,38	1,90	
3	3,22	2,20	2,20	
4	3,69	2,75	1,30	
5	3,02	2,45	2,04	
6	2,77	2,36	1,70	
$Y_{.j} = \sum_{i=1}^6 Y_{ij}$	19,86	14,76	10,74	195,6
$Y_{.j}^2 = \sum_{i=1}^6 Y_{ij}^2$	66,56	36,50	19,75	122,81
$\bar{Y}_{.j} = \sum Y_{ij} / 6$	3,31	2,46	1,79	$\bar{Y}_{..} = 10,87$

*A los datos originales (Y propágulos/g suelo) se aplicó la transformación $Y = \log(100 \times Y)$.

Las varianzas de los datos transformados son:

$$s_1^2 \text{ (No solarizado)} = \frac{\sum (Y_i - \bar{Y})^2}{n-1} = \frac{(3,85 - 3,31)^2 + \dots + (2,77 - 3,31)^2}{6-1} = \frac{0,820}{5} = 0,164$$

$$s_2^2 \text{ (Solarizado 1)} = \frac{\sum (Y_i - \bar{Y})^2}{n-1} = \frac{(2,62 - 2,46)^2 + \dots + (2,36 - 2,46)^2}{6-1} = \frac{0,1938}{5} = 0,039$$

$$s_3^2 \text{ (Solarizado 2)} = \frac{\sum (Y_i - \bar{Y})^2}{n-1} = \frac{(1,60 - 1,79)^2 + \dots + (1,70 - 1,79)^2}{6-1} = \frac{0,527}{5} = 0,105$$

$$s^2 = \frac{0,164 + 0,039 + 0,105}{3} = 0,103$$

$$M = 5[3 \times \ln(0,103) + 7,306] = 2,435$$

$$C = 1 + \frac{3+1}{3 \times 3 \times 5} = 1 + \frac{4}{45} = 1,089$$

$$\chi^2 = \frac{M}{C} = \frac{2,435}{1,089} = 2,24$$

De la Tabla A3 se obtiene que $\chi^2_{0,05;2} = 5,99 > 2,24$ y se concluye que las varianzas son ahora homogéneas.

Comprobados los datos, procedemos con el análisis. Los cálculos necesarios para construir la tabla ANOVA, utilizando el procedimiento más simplificado, son:

$$\text{Factor de corrección: } C = \frac{Y_{..}^2}{ar} = \frac{(45,36)^2}{3 \times 6} = \frac{2057,53}{18} = 114,31$$

$$SC_{Total} = \sum_{\uparrow} \sum_{\uparrow} Y_{ij}^2 - C = (3,85^2 + \dots + 1,70^2) - 114,31 = 122,81 - 114,31 = 8,50$$

$$SC_{Tratamientos} = \frac{1}{6} \sum_{\uparrow} Y_{.j}^2 - C = \frac{19,86^2 + 14,76^2 + 10,74^2}{6} - 114,31 = 121,27 - 114,31 = 6,96$$

$$SC_{Error} = SC_{Total} - SC_{Tratamientos} = 8,50 - 6,96 = 1,54$$

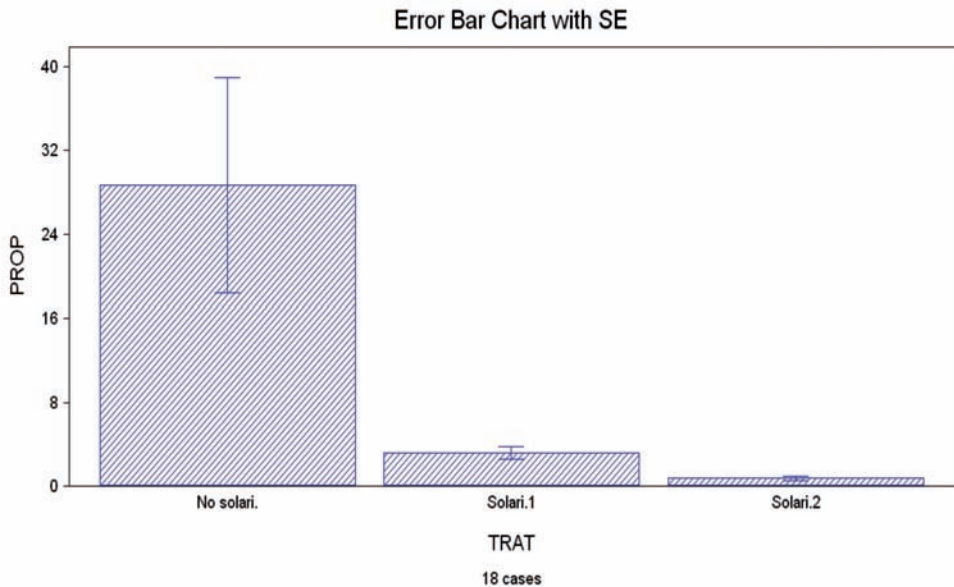
La tabla del ANOVA quedaría, según se ha explicado anteriormente, como sigue:

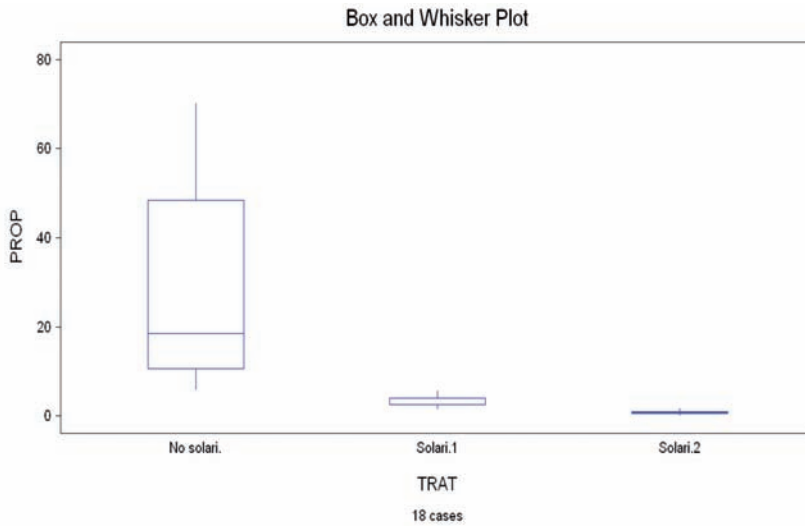
Fuente de Variación	gl	SC	CM	F	P
Solarización	2	6,96	3,48	33,89	0,00000
Error	15	1,54	0,1027		
Total	17	8,50			

Como $P < 0.05$ (Tabla A2), se rechaza la hipótesis nula y, por tanto, concluimos que existen diferencias significativas entre tratamientos. Para comprobar la naturaleza de dichas diferencias tendríamos que realizar algún test adicional de comparación de medias o contrastes como se indica en el Capítulo 5.

CÁLCULO CON EL PROGRAMA *Statistix* (SX)

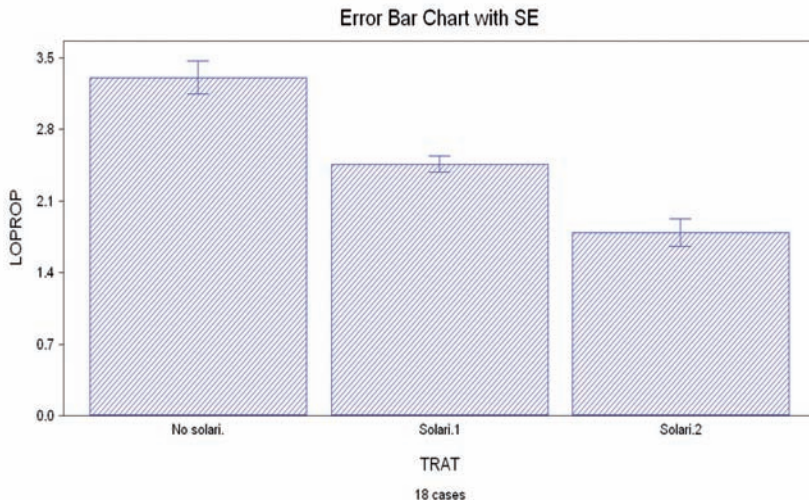
Para realizar el ANOVA utilizando el programa *Statistix*, lo primero, después de introducir los datos, sería visualizarlos. Para ello existen diversas opciones, pero las más comunes son “*Error Bar Chart*” y “*Box and Whisker Plot*” del menú “*Summary Statistics*”:

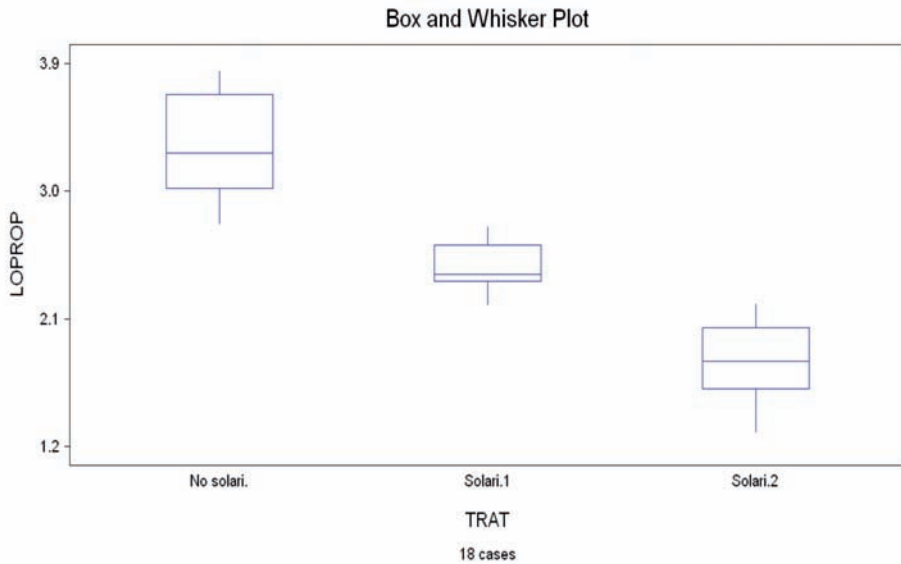




Como se puede apreciar, ambos gráficos muestran una amplia variabilidad de los datos. En el gráfico de arriba aparecen las medias de los tratamientos con su error estándar. Existen también las opciones alternativas de representar la desviación estándar, en lugar del error estándar, y una línea en vez del gráfico de barras. El gráfico de abajo ofrece la mediana (línea horizontal), el intervalo intercuartílico (cuartil 3º - cuartil 1º) representado por la altura de la caja, y las prolongaciones o líneas verticales superior e inferior ("bigotes"). El bigote superior termina en el valor absoluto igual o inferior al tercer cuartil más 1,5 veces el intervalo intercuartílico, mientras que el bigote inferior termina en el valor absoluto que sea igual o superior al primer cuartil menos 1,5 veces el intervalo intercuartílico. Este gráfico señala también todos los datos que no se incluyen en la caja o bigotes como los raros o "outlier". En el ejemplo no hay ningún valor "outlier".

En el programa *Statistix*, el *test de Bartlett* está incluido dentro del análisis de la varianza. Por ello, se comenta más tarde y ahora se visualizan los datos transformados:





A la vista de los gráficos de los datos transformados, es evidente que la transformación ha igualado las varianzas. Tampoco existe ningún dato raro, luego se procede con el análisis siguiendo la secuencia:

Statistics>Linear Models>Analysis of Variance>Completely Randomized Design

A efectos comparativos introducimos dos variables dependientes: la original PROP (propá-gulos/g de suelo) y la transformada LOPROP [$\log(100 \times PROP)$]. Se introduce también la variable categórica TRAT (tratamientos: No solari., Solari. 1, Solari. 2) y se pulsa OK. Los resultados del análisis son:

Completely Randomized AOV for PROP

Source	DF	SS	MS	F	P
TRAT	2	2873.94	1436.97	6.75	0.0081
Error	15	3193.44	212.90		
Total	17	6067.38			

Grand Mean	10.867	CV	134.27			
				Chi-Sq	DF	P
Bartlett's Test of Equal Variances				46.7	2	0.0000
Cochran's Q			0.9962			
Largest Var / Smallest Var			2385.9			

Component of variance for between groups	204.013
Effective cell size	6.0

```

TRAT           Mean
No solari.     28.683
Solari.1       3.150
Solari.2       0.767
Observations per Mean           6
Standard Error of a Mean       5.9567
Std Error (Diff of 2 Means)    8.4241
    
```

Completely Randomized AOV for LOPROP

Source	DF	SS	MS	F	P
TRAT	2	6.93608	3.46804	33.9	0.0000
Error	15	1.53405	0.10227		
Total	17	8.47012			

Grand Mean 2.5204 CV 12.69

	Chi-Sq	DF	P
Bartlett's Test of Equal Variances	2.22	2	0.3294
Cochran's Q	0.5297		
Largest Var / Smallest Var	4.2460		

Component of variance for between groups 0.56096
 Effective cell size 6.0

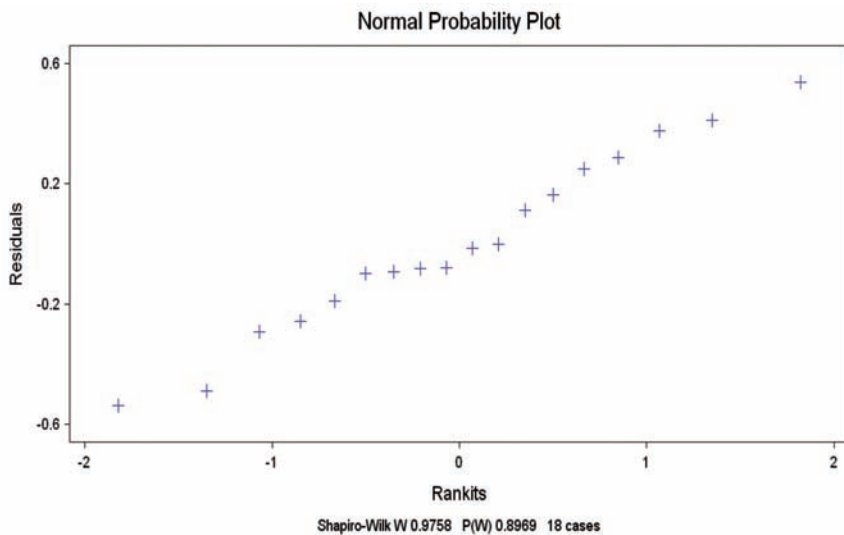
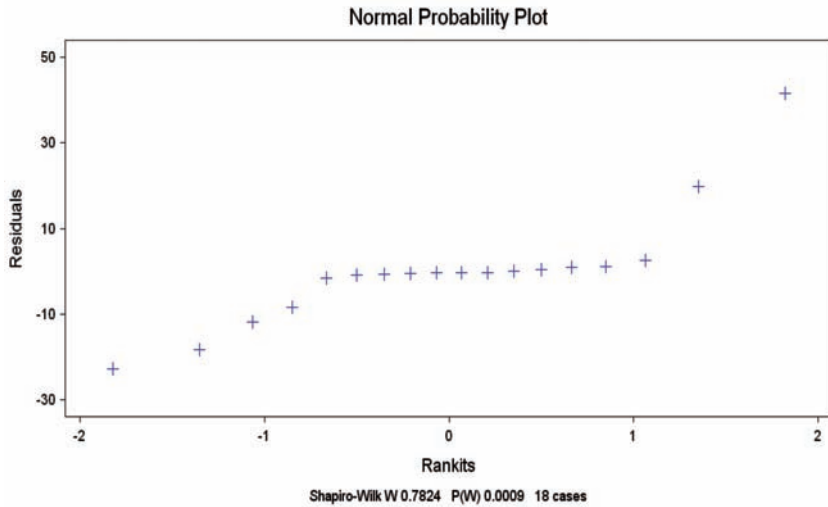
```

TRAT           Mean
No solari.     3.3088
Solari.1       2.4608
Solari.2       1.7918
Observations per Mean           6
Standard Error of a Mean       0.1306
Std Error (Diff of 2 Means)    0.1846
    
```

Además de la tabla de ANOVA, *Statistix* incluye varios estadísticos, como la media general del experimento, el coeficiente de variación (CV), errores estándar, y tres test de homogeneidad de las varianzas: *Bartlett*, *Cochran* y *razón Varianza mayor/Varianza menor*. En el *test de Bartlett*, valores de $P < 0.05$ indican que podemos rechazar la hipótesis nula de igualdad de las varianzas, luego las varianzas de los datos originales son heterogéneas ($P < 0.0001$), mientras que los datos transformados presentan varianzas homogéneas ($P = 0,3294 > 0,05$). Obsérvese también el alto valor del CV en los datos no transformados.

Se puede comprobar si se cumplen otros preceptos del ANOVA. La normalidad de los datos se comprueba eligiendo la siguiente secuencia en la pantalla de resultados:

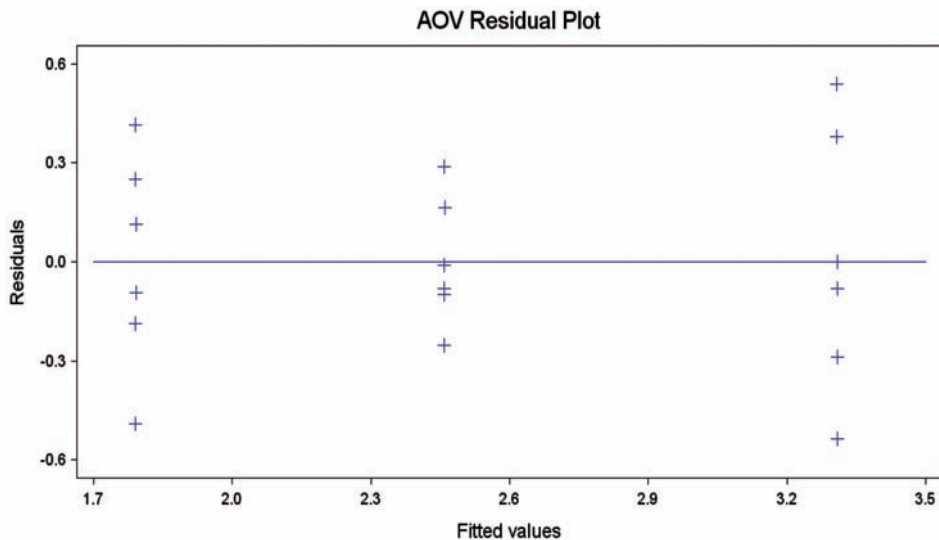
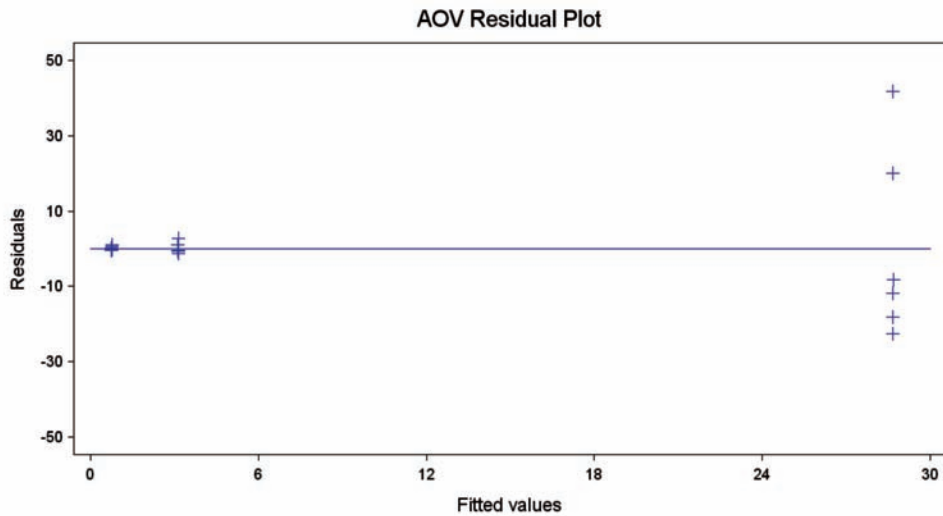
Results>Plots>Normal Probability Plot



El test *Shapiro-Wilk* indica que los datos originales (arriba) no se ajustan a una normal ($P < 0.05$), pero los transformados (abajo) sí ($P > 0.05$), ajustándose los datos a una línea recta.

De igual forma se puede ver la distribución que presentan los residuos en la misma pantalla de resultados siguiendo la secuencia:

Results>Plots>Resids By Fitted Values



Los gráficos de distribución de residuos también indican grandes diferencias entre los valores originales (arriba) y los transformados (abajo). La distribución de residuos de los datos originales muestra un claro patrón en cuña, inaceptable, mientras que los residuos de los transformados se distribuyen aleatoriamente, indicando su independencia.

Todas las comprobaciones realizadas demuestran claramente que los datos originales no son adecuados para el ANOVA, mientras que los datos transformados se adecuan perfectamente a dicho análisis. El ANOVA realizado se completaría con la comparación de los tratamientos que se desarrolla en el Capítulo siguiente.

CAPÍTULO 5

SEPARACIÓN DE MEDIAS Y CONTRASTES

Cuando el ANOVA resulta estadísticamente significativo el paso siguiente consiste en averiguar entre qué grupos o tratamientos hay diferencias. A los procedimientos para comparar medias después del ANOVA se les denomina contrastes y se pueden clasificar en dos categorías principales: contrastes *a priori* y contrastes *a posteriori*. Como su nombre indica, los primeros son los que se establecen antes de iniciar el análisis de los datos; los segundos, en cambio, son todas las posibles comparaciones de medias dos a dos, es decir, $a(a-1)/2$, si a es el número de medias a comparar.

Los numerosos métodos utilizados para realizar comparaciones entre las medias de los tratamientos varían principalmente en función del nivel de protección frente a los dos tipos de error asociados a cualquier contraste de hipótesis, según se indica en la Tabla 5.1.

Tabla 5.1. Principales tipos de error asociados con un contraste de hipótesis

		Verdad (Realidad)	
		H_0	H_1
Decisión	H_0	Acierto Nivel de confianza $(1-\alpha)$	Error de Tipo II Riesgo β (falsos-negativos)
	H_1	Error de Tipo I Riesgo α (falsos-positivos)	Acierto Potencia $(1-\beta)$

Los dos tipos de error están inversamente relacionados, luego los métodos que den más protección frente al error tipo I ofrecerán menor protección frente al error tipo II y viceversa. La elección de unos métodos u otros dependerá del planteamiento y expectativas de la investigación. Una forma de disminuir la probabilidad de aceptar H_0 cuando es falsa, es decir, protegernos frente al error tipo II, es aumentando el número de observaciones o repeticiones (r).

A continuación se comentan los métodos más usados para comparar medias según el tipo de comparación: *a priori* o planificadas y *a posteriori*.

COMPARACIONES A PRIORI

Se utilizan en experimentos que tienen una estructura de tratamientos establecida de antemano, es decir, las comparaciones a realizar son hipótesis especificadas antes de realizar el experimento. Esta es la situación deseable siempre porque reduce el número de comparaciones, dejando sólo las esenciales y eliminando las irrelevantes, reduciendo así el riesgo de errores tipo I y II.

El método más utilizado de este grupo es el de los contrastes ortogonales, aunque también son empleados los contrastes polinómicos cuando la variable independiente es cuantitativa.

Contrastes ortogonales

La estrategia de este método consiste en repartir la suma de cuadrados de los tratamientos entre las diversas hipótesis o contrastes. La suma de cuadrados asociada a cada contraste se usa como numerador y el cuadrado medio del error como denominador para calcular razones F y comprobar su significación (valor P). Cada contraste consume un grado de libertad, por lo que tendremos $a-1$ contrastes posibles.

El procedimiento a seguir para calcular contrastes ortogonales se resume a continuación:

a) Calcular los contrastes planificados (C_i):

$$C_1 = \sum w_{1j} \times T_j = w_{11} \times T_1 + w_{12} \times T_2 + \dots w_{1a} \times T_a$$
$$C_2 = \sum w_{2j} \times T_j = w_{21} \times T_1 + w_{22} \times T_2 + \dots w_{2a} \times T_a \dots$$

donde T_j son los valores totales de cada grupo de tratamientos y w_i son pesos o coeficientes asignados a cada grupo. Los coeficientes son números enteros arbitrarios (-3, -2, -1, 0, +1, +2, +3) que indican el peso relativo de cada medida, cuyos signos positivos y negativos son convencionales, sólo sirven para identificar los grupos que están a cada lado de la comparación. Por tanto, la suma de todos los coeficientes de cada contraste debe ser 0. Los coeficientes de valor 0 identifican los grupos que no intervienen en una determinada comparación.

Los contrastes de interés deben establecerse de forma que no haya superposiciones, es decir, que no utilicen la misma fuente de varianza, por eso se denominan ortogonales. Las condiciones que deben cumplir los contrastes para ser ortogonales son:

$$\sum (w_i \times w_j) = 0$$

donde, i se refiere a los coeficientes de un contraste y j a los del otro.

b) Calcular la suma de cuadrados de cada contraste (SCC_i), según la expresión:

$$SCC_i = \frac{(\sum_j w_{ij} \times T_j)^2}{r \sum_j w_{ij}^2}$$

donde w_i son los coeficientes de cada contraste (j), T_j los totales de los tratamientos y r el número de repeticiones incluidas en cada total. A cada contraste le corresponde un grado de libertad. Si C_i son mutuamente ortogonales, entonces se cumple:

$$SC_{\text{Tratamientos}} = \sum_i SCC_i$$

c) Calcular los cocientes F . Con el valor de cada SCC_i y el cuadrado medio del error (CME) se calcula la F de Fisher con 1 gl del numerador y tantos gl del denominador como tenía la varianza residual: $(N - a) = a \times r - a = a \times (r - 1)$.

d) Ver la significación estadística (P) de la F calculada. De la tabla A2 o de algún programa estadístico, se obtiene el valor de P. Si $P < 0.05$, se rechaza la hipótesis nula y el contraste es significativo.

Para explicar los contrastes ortogonales vamos a utilizar el ejemplo sobre tratamientos fungicidas para el control de una enfermedad en el olivar, que luego se desarrolla en el Capítulo 12. Tenemos un factor, tratamientos fungicidas, con cuatro niveles o tratamientos y 10 observaciones por cada nivel. Los cuatro tratamientos son: sin tratar (ST), tratamiento de primavera (P), tratamiento de otoño (O), tratamiento de primavera y de otoño (P+O), en un diseño en bloques al azar con 10 repeticiones. Los valores totales de la incidencia de la enfermedad para estos tratamientos son 653,3; 239,0; 467,7 y 138,5; respectivamente (Tabla 12.4). Como se tienen cuatro tratamientos, hay $4-1=3$ grados de libertad y, en consecuencia, 3 contrastes posibles. Un posible conjunto de contrastes ortogonales sería:

Contraste	ST	Coeficientes			C_i	SCC_i
		P	O	P+O		
Sin tratar vs. Tratados	3	-1	-1	-1	1114,7	10354,63
Trat. indiv. vs. Trat. doble	0	1	1	-2	429,7	3077,37
P vs. O (individual)	0	1	-1	0	-228,7	2615,18
$SC_{\text{Tratamientos}} = 16047,18$						

Se puede comprobar que la suma de los coeficientes para cada contraste, es decir, para cada línea de la tabla anterior, es cero; por ejemplo, para el contraste primero: $3 + (-1) + (-1) + (-1) = 0$. Asimismo, la suma de los productos entre los coeficientes de cada par de contrastes también es cero; para los dos primeros contrastes, por ejemplo, se tiene que $(3 \times 0) + (-1 \times 1) + (-1 \times 1) + (-1 \times 2) = 0$. Se trata, por tanto, de contrastes mutuamente ortogonales. En las dos columnas de la derecha se indican el valor de cada contraste (C_i) y de la suma de cuadrados correspondiente (SCC_i). Su cálculo detallado para cada contraste sería:

$$C_1 = 3 \times 653,3 - 239,0 - 467,7 - 138,5 = 1114,7$$

$$SCC_1 = \frac{(C_1)^2}{r \times [3^2 + (-1)^2 + (-1)^2 + (-1)^2]} = \frac{(1114,7)^2}{10 \times 12} = 10354,63$$

$$C_2 = 239,0 + 467,7 - (2 \times 138,5) = 429,7$$

$$SCC_2 = \frac{(C_2)^2}{r \times [1^2 + 1^2 + (-2)^2]} = \frac{429,7^2}{10 \times 6} = 3077,37$$

$$C_3 = 239,0 - 467,7 = -228,7$$

$$SCC_3 = \frac{(C_3)^2}{r \times [1^2 + (-1)^2]} = \frac{(-228,7)^2}{10 \times 2} = 2615,18$$

Como se tiene 1 gl por cada contraste, los valores de la suma de cuadrados coinciden con el cuadrado medio ($SCC_i = CMC_i$). La significación de estos contrastes mediante el test F , teniendo en cuenta que el cuadrado medio del error (CME) para este experimento es 54,1 con 12 gl (véase tabla del ANOVA del Capítulo 12), sería:

$$F_1 = \frac{CMC_1}{CME} = \frac{10354,63}{54,1} = 191,39 \Rightarrow P < 0,0001$$

$$F_2 = \frac{CMC_2}{CME} = \frac{3077,37}{54,1} = 56,88 \Rightarrow P < 0,0001$$

$$F_3 = \frac{CMC_3}{CME} = \frac{2615,18}{54,1} = 48,34 \Rightarrow P < 0,0001$$

Los tres contrastes resultan altamente significativos, por lo que se concluye que los cuatro tratamientos son diferentes. Es decir, los olivos tratados presentan una incidencia menor de la enfermedad que los olivos no tratados, el doble tratamiento también produce una reducción de la enfermedad frente a los tratamientos individuales y el tratamiento individual de primavera es más eficaz en el control de la enfermedad que el tratamiento de otoño.

Para el mismo ejemplo podríamos haber elegido otro grupo de contrastes ortogonales, como se indica a continuación, obteniéndose los resultados siguientes:

Contraste	ST	Coeficientes			C _i	SCC _i
		P	O	P+O		
P (Primavera vs. no primavera)	-1	1	-1	1	-743,5	13819,80
O (Otoño vs. no otoño)	-1	-1	1	1	-286,1	2046,33
P*O	1	-1	-1	1	85,1	181,05
SC _{Tratamientos} = 16047,18						

Al ser estos contrastes también ortogonales, se obtiene la misma suma de cuadrados total, que coincide con la suma de cuadrados del factor Tratamientos (véase tabla del ANOVA del Capítulo 12). La significación de estos nuevos contrastes sería:

$$F_1 = \frac{CMC_1}{CME} = \frac{13819,80}{54,1} = 255,45 \Rightarrow P < 0,0001$$

$$F_2 = \frac{CMC_2}{CME} = \frac{2046,33}{54,1} = 37,82 \Rightarrow P < 0,0001$$

$$F_3 = \frac{CMC_3}{CME} = \frac{181,05}{54,1} = 3,35 \Rightarrow P = 0,0913$$

Utilizando estos nuevos contrastes llegaríamos a conclusiones parecidas, pero no iguales, ya que no incluyen las mismas comparaciones. Así, se puede decir que el efecto más importante es el del tratamiento de primavera. El tratamiento de otoño, aunque con menor efecto, también resultó muy significativo, pero no hubo interacción entre ambos tratamientos, por lo que su efecto fue aditivo y no hubo sinergia ni antagonismo.

Para utilizar los contrastes ortogonales con el programa *Statistix*, una vez realizado el ANOVA correspondiente, en la pantalla de resultados se elige la siguiente secuencia:

Results>Contrasts...

En la pantalla resultante se introducen los coeficientes con sus signos, de acuerdo con las comparaciones que deseamos realizar, y se pulsa OK. Para cada uno de los contrastes ortogonales elegidos, con un grado de libertad, aparecerá el valor de P correspondiente al estadístico *t* del *test de Student* y a la *F* del *test de Scheffé* (veáanse las comparaciones *a posteriori*). Los resultados correspondientes al ejemplo utilizado en este apartado, que obviamente coinciden con el cálculo manual, se indican en el Capítulo 12.

Contrates polinómicos

Cuando la variable independiente es un factor cuantitativo (temperatura, tiempo, densidad de inóculo, etc.), el concepto de contrastes ortogonales se puede extender a los diferentes ni-

veles de este factor para determinar si existe alguna tendencia (lineal, cuadrática, cúbica) en los datos. En este caso, los procedimientos de comparación múltiple de medias son claramente inapropiados. El contraste polinómico tiene su principal aplicación cuando los niveles del factor cuantitativo son pocos (<5). Sin embargo, cuando los niveles de ese factor son mayores, el análisis de regresión es más apropiado.

Al igual que en los contrastes ortogonales, la suma de cuadrados del factor se divide en diferentes contrastes (lineal, cuadrático, cúbico, etc.), con un grado de libertad cada uno. El contraste lineal determina si existen diferencias significativas entre los niveles bajos y los niveles altos del factor cuantitativo. El cuadrático detecta diferencias entre los niveles intermedios y los extremos. El cúbico compara entre niveles alternos. La hipótesis nula (H_0) es que no existe ningún efecto lineal, cuadrático, cúbico, etc., debido a los diferentes niveles del factor cuantitativo. Los valores de los coeficientes para realizar estos contrastes están tabulados en la Tabla A15, considerando igual espaciamiento entre los niveles de la variable independiente, lo que facilita el cálculo manual. Algunos paquetes estadísticos, incluido *Statistix*, permiten los cálculos para un espaciamiento desigual de la variable.

Para ilustrar el procedimiento a seguir con los contrastes polinómicos vamos a utilizar un ejemplo sobre el rendimiento de un cultivo sometido a cuatro tratamientos de fertilización nitrogenada en un diseño en bloques completos al azar con 4 repeticiones. Tenemos, por tanto, un factor Nitrógeno con cuatro niveles: 0, 40, 80 y 120 kg de N/ha. Los totales de estos tratamientos son: 5961,2, 9538,0, 11922,4 y 12160,8, respectivamente. El cálculo de los contrastes sería idéntico a al realizado con los contrastes ortogonales. En este caso serían:

Contraste	Coeficientes				C_i	SCC_i
	0	40	80	120		
1. Lineal	-3	-1	+1	+3	20983,2	5503683,5
2. Cuadrático	+1	-1	-1	+1	3338,4	696557,2
3. Cúbico	-1	+3	-3	+1	953,6	11366,9
$SC_{\text{Tratamientos}} = 6211607,6$						

Como tenemos 1 gl por cada contraste, los valores de la suma de cuadrados coinciden con el cuadrado medio ($SCC_i = CMC_i$). La significación de estos contrastes mediante el test F , teniendo en cuenta que el cuadrado medio del error (CME) para este experimento es 98736,5 con 9 gl, sería:

$$F_1 = \frac{CMC_1}{CME} = \frac{5503683,5}{98736,5} = 55,74 \Rightarrow P < 0,0001$$

$$F_2 = \frac{CMC_2}{CME} = \frac{696557,2}{98736,5} = 7,05 \Rightarrow P = 0,0263$$

$$F_3 = \frac{CMC_3}{CME} = \frac{11366,9}{98736,5} = 0,12 \Rightarrow P = 0,7370$$

Luego podemos concluir que la mayor parte de la variación se debe al efecto lineal, aunque el efecto cuadrático también ha resultado significativo, pero no existe efecto cúbico. La tendencia del incremento de la cosecha con la dosis de nitrógeno no es sólo lineal sino que sigue una curva parabólica creciente, similar a una curva de saturación, donde existe una tasa de crecimiento más pronunciada para dosis bajas, la cual va reduciéndose a medida que aumenta la dosis de nitrógeno.

Para utilizar los contrastes polinómicos con el programa *Statistix*, una vez realizado el ANOVA correspondiente, en la pantalla de resultados, se elige la siguiente secuencia:

Results>Polynomial Contrasts...

En la pantalla resultante se introduce el grado del polinomio (1, 2, 3, ...) y se pulsa OK. Para cada tendencia (lineal, cuadrática, cúbica...), con un grado de libertad, aparecerá el valor de P correspondiente al estadístico *t* del *test de Student* y a la *F* del *test de Scheffé* (veáanse las comparaciones *a posteriori*).

Comparaciones con un control

Cuando se pretende comparar un conjunto de medias, una a una, frente a una única media, la del control, el procedimiento más usado es el *test de Dunnett*. Este test establece un valor crítico para la comparación de dos medias, a partir del cual podemos declarar que dichas medias difieren significativamente. El valor crítico (C) viene dado por:

$$C = d_{(\alpha_e, a-1, \gamma)} \times S_{\bar{D}}$$

donde $S_{\bar{D}}$ es el error estándar de la diferencia entre dos medias y d es el estadístico de *Dunnett* para γ grados de libertad del error, a medias y α_e la tasa del error experimental, es decir, el riesgo de error tipo I.

Se trata, por tanto, de un método conservador que protege principalmente contra el error tipo I (falsos positivos). No obstante, si se compara con los métodos conservadores de comparaciones múltiples es más liberal, o potente, que el más liberal de éstos, el de *Tukey*, ya que al haber menos comparaciones disminuye el riesgo de error tipo II (falsos negativos).

Para ilustrar este procedimiento tomemos como ejemplo el de la Tabla 18.1, que recoge la producción obtenida en un experimento donde se trataba de comparar tres épocas de aclareo de la nectarina 'Armking' frente a un control sin aclarar. Como se muestra en el capítulo 18, el error estándar de la diferencia entre dos medias fue $S_{\bar{D}} = 0,94$ y los grados de libertad del error 11. Para un nivel de significación $\alpha = 0,05$, se obtiene en la Tabla A13 el valor $d = 2,72$. El valor crítico será, pues:

$$C = d \times s_{\bar{D}} = 2,72 * 0,94 = 2,5568$$

Las comparaciones de las medias con el control muestran que todos los tratamientos difieren del mismo, como se indica a continuación:

Comparaciones	Diferencia	Valor crítico	Significación
18,816 – 13,624	5,19	2,5568	*
18,816 – 12,106	6,71	2,5568	*
18,816 – 9,184	9,63	2,5568	*

Para utilizar las comparaciones con un control en el programa *Statistix*, una vez realizado el ANOVA correspondiente, en la pantalla de resultados, se elige la siguiente secuencia:

Results>Multiple Comparisons>Comparisons with a Control

En la pantalla que aparece se introduce el tratamiento control, el nivel de significación α (generalmente $\alpha = 0.05$), y la alternativa considerada (diferente del control, menor que el control, o mayor que el control), se pulsa OK y se obtiene la siguiente salida:

```

Two-sided Dunnett's Multiple Comparisons with a Control of PRO
Control: TRA=1

Simultaneous 95% confidence intervals of treatment mean -
control mean

  TRA      Mean      Lower      Difference      Upper
   1  18.816      -7.750      -5.192*        -2.635
   2  13.624      -9.223      -6.710*        -4.197
   3  12.106     -12.190     -9.632*       -7.075
   4   9.184

Alpha                0.05
Critical D Value     2.717
Error term used: BLK*TRA, 11 DF
    
```

que muestra el valor de las medias a comparar con el control (TRA=1), el intervalo de confianza de la diferencia de cada media con el control, el nivel de significación α , el valor de d en tablas y los grados de libertad del error.

Comparaciones con el mejor

Cuando en lugar de comparar con la media del control se pretende identificar los tratamientos que producen el mejor resultado (media mayor o menor), se utiliza el test de *Hsu* (1996).

Al igual que el test de *Dunnett*, el de *Hsu* es un método conservador pero, al reducir el número de comparaciones, aumenta su protección contra el error tipo II siendo, por tanto, más potente o liberal que el test de *Tukey*.

Para utilizar las comparaciones con el Mejor en el programa *Statistix*, una vez realizado el ANOVA correspondiente, en la pantalla de resultados se elige la siguiente secuencia:

Results>Multiple Comparisons>Comparisons with the Best

En la pantalla resultante se introduce el nivel de significación α (generalmente $\alpha = 0,05$) y la alternativa considerada (el valor mayor o el valor menor) y se pulsa OK. El resultado del análisis presenta varios estadísticos, entre los que destaca el valor crítico para comparación según el test de *Hsu*, y las medias con su intervalo de confianza, destacando con un asterisco aquellas que difieren significativamente del valor mayor o del valor menor, dependiendo de la alternativa elegida.

COMPARACIONES A POSTERIORI

Se utilizan en experimentos que no tienen una estructura *a priori* de los tratamientos y permiten realizar todas las comparaciones posibles por parejas [$a \times (a - 1) / 2$] entre las medias. Al ser los procedimientos más empleados tradicionalmente, se han denunciado numerosos abusos derivados de su utilización para comparar medias de un factor cuantitativo o de experimentos con una estructura claramente definida de los tratamientos.

Los métodos de comparaciones múltiples son numerosos y se pueden agrupar en dos categorías en función de la protección que ofrecen contra los errores de tipo I y de tipo II. Los que presentan una mayor potencia contra el error tipo II (falsos negativos) y, por tanto, ofrecen una menor protección contra el error tipo I (falsos positivos), se consideran “liberales”, mientras que los que protegen especialmente contra el error tipo I y presentan una menor potencia contra el error tipo II, serían los “conservadores”. En el primer grupo se incluyen la mínima diferencia significativa (MDS o “least significant difference, *LSD*”), o mínima diferencia significativa protegida de Fisher (*FLSD*), y la *LSD* de *Waller-Duncan-Bayes* (*BLSD*). En el segundo grupo se incluyen los tests de *Duncan*, *Student-Newman-Keuls* (*SNK*), *Tukey*, *Sidak*, *Bonferroni* y *Scheffé*. El orden de estos métodos respecto a la protección frente al error de tipo I sería:

FLSD < *BLSD* < *Duncan* < *SNK* < *Tukey* < *Sidak* < *Bonferroni* < *Scheffé*

La estrategia general de estos métodos se basa en establecer un valor crítico de comparación (*C*) y si el valor absoluto de la diferencia entre las medias que comparamos supera ese nivel crítico se declara que dichas medias son diferentes significativamente. El nivel de significación (α) para el cálculo del valor crítico suele ser 0.05, aunque puede utilizarse otro. Los intervalos de confianza para las comparaciones se calculan como $\bar{X}_1 - \bar{X}_2 \pm C$. Respecto a la utilización de uno u otro de estos métodos no existe un acuerdo unánime entre los esta-

dísticos, por lo que tendrá que ser una decisión particular en función de los objetivos del experimento y la protección deseada frente a los errores tipo I y tipo II. Por eso, muchos programas estadísticos incluyen varios de ellos.

Para utilizar las comparaciones múltiples en el programa *Statistix*, una vez realizado el ANOVA correspondiente, en la pantalla de resultados se elige la siguiente secuencia:

Results>Multiple Comparisons>All-pairwise Comparisons

En la pantalla resultante se introduce el método de comparación (*Tukey, LSD, Scheffé, Multiplicative Sidak, Bonferroni*), el nivel de significación α (generalmente $\alpha = 0.05$), y el formato de presentación de los resultados (grupos homogéneos o matriz triangular; generalmente se utiliza el primero) y se pulsa OK. El resultado del análisis presenta varios estadísticos, entre los que destaca el valor crítico para comparación, según el test seleccionado. En la opción grupos homogéneos, las medias aparecen seguidas de una o varias letras, cada letra señala a un grupo homogéneo de medias, de forma que medias seguidas por alguna letra común no difieren significativamente entre sí para el nivel α elegido. En la otra opción de formato de resultados, las medias aparecen en una matriz triangular, destacándose con un asterisco todos los pares de medias que difieren significativamente.

A continuación se indican las características más sobresalientes de los métodos de comparación múltiple más utilizados en los experimentos agronómicos.

Método de la Mínima Diferencia Significativa (MDS o “LSD”)

El método de la mínima diferencia significativa (*MDS*), o “*Least Significant Difference*” (*LSD*), es el que suele dar el valor crítico más bajo y, por tanto, el menos conservador, o más liberal, de los métodos de comparación múltiple. El valor crítico para una comparación es:

$$MDS(LSD) = t_{(\alpha, \gamma)} \times \sqrt{CME \times \left(\frac{1}{r_1} + \frac{1}{r_2} \right)} = t \times \sqrt{\frac{2CME}{r}} = t \times S_{\bar{D}}$$

donde t es la t de Student para el nivel de significación α , y γ grados de libertad del error; CME es el cuadrado medio del error y r_1 y r_2 son el número de observaciones de cada media. Cuando éstas son iguales $r_1 = r_2 = r$.

El problema de este método es que el error tipo I aumenta al aumentar el número de comparaciones. Por ello, este método se ha recomendado sólo después de que el ANOVA haya resultado significativo, denominándose como el *test LSD protegido de Fisher (FLSD)*. En cualquier caso, este test es el que ofrece menor protección contra el error tipo I, por lo que debería usarse sólo en contrastes establecidos *a priori* y cuando el número de comparaciones es pequeño.

Tomemos como ejemplo los datos de la Tabla 7.1, que recoge los datos de un experimento en bloques al azar con cinco tratamientos (variedades) y cuatro bloques. El $CME = 1,775$ con

12 grados de libertad. Para un nivel de significación $\alpha=0,05$ el valor de t en la Tabla A1 es 2,179, luego:

$$MDS = t_{0,05} \sqrt{\frac{2CME}{R}} = 2,179 \sqrt{\frac{2 \times 1,775}{4}} = 2,179 \times 0,9421 = 2,05$$

La hipótesis nula se rechaza cuando la diferencia entre dos medias es superior al valor crítico MDS. Las comparaciones entre medias y su significación se muestran a continuación:

Comparaciones	Diferencia	Valor crítico	Significación
12,0-10,5	1,5	2.05	NS
12,0-8,0	4	2.05	*
12,0-9,25	2,75	2.05	*
12,0-10,25	1,75	2.05	NS
10,5-8,0	2,5	2.05	*
10,5-9,25	1,25	2.05	NS
10,5-10,25	0,25	2.05	NS
8,0-9,25	-1,25	2.05	NS
8,0-10,25	-2,25	2.05	*
9,25-10,25	-1,0	2.05	NS

Se han encontrado diferencias significativas entre las variedades 1 y 3; 1 y 4; 2 y 3; 5 y 3, no difiriendo significativamente las demás comparaciones.

En el programa *Statistic*, siguiendo la secuencia:

Results >Multiple comparisons >All-pairwise comparisons

se elije LSD para $\alpha = 0,05$ y se obtiene la siguiente salida:

```

LSD All-Pairwise Comparisons Test of Rendimiento for Variedades

Variedad      Mean  Homogeneous Groups
1              12.000      A
2              10.500      AB
5              10.250      AB
4               9.250      BC
3               8.000      C

Alpha                0.05      Standard Error for Comparison  0.9421
Critical T Value     2.179      Critical Value for Comparison  2.0526
Error term used: V002*V001, 12 DF
There are 3 groups (A, B, etc.) in which the means
are not significantly different from one another.
    
```

Como es lógico los resultados coinciden con los obtenidos anteriormente. El programa muestra el nivel de significación α , el valor de t para ese nivel de significación, el valor de $S_{\bar{D}}$ y el valor crítico para la comparación (MDS).

Método de Tukey

Este método, conocido también como “*Honestly Significant Difference (HSD)*”, hace uso del estadístico Q (“amplitud estudentizada”) y es aplicable a comparaciones por pares de medias, por lo que puede dar diferencias significativas aunque el valor de F en el ANOVA no resulte significativo. El valor crítico se calcula teniendo en cuenta el número de tratamientos, lo que protege contra el error tipo I. Aunque es, por tanto, un método conservador, resulta el más liberal de ellos. Su principal desventaja, aparte de su sensibilidad al error tipo II, es que requiere el mismo número de observaciones o repeticiones en las medias. Este inconveniente se supera con una variante del método conocida como el *test de Tukey-Kramer*.

El estadístico Q viene dado por la expresión:

$$Q = \frac{\bar{X}_i - \bar{X}_j}{\sqrt{\frac{CME}{n}}}$$

siendo CME el cuadrado medio del error y n el número de repeticiones. Los valores críticos del test están tabulados en la Tabla A14. La hipótesis nula es rechazada si el valor de Q calculado es mayor que el valor crítico de la tabla.

Tomemos como ejemplo los datos de la Tabla 7.1, que recoge los datos de un experimento en bloques al azar con cinco tratamientos (variedades) y cuatro bloques. El CME = 1,775 con 12 grados de libertad. Para un nivel de significación $\alpha=0,05$ y cinco tratamientos, el valor crítico mostrado en la Tabla A14 es 4,51. Los valores de Q para las diferentes comparaciones de todas las medias entre sí serían:

$$Q = \frac{\bar{X}_i - \bar{X}_j}{\sqrt{\frac{1,775}{4}}} = \frac{\bar{X}_i - \bar{X}_j}{0,6661}$$

Las comparaciones se muestran a continuación:

Comparaciones	Q	Valor crítico	Significación
12,0-10,5	$1,5 / 0,6661 = 2,25$	4,51	NS
12,0-8,0	$4 / 0,6661 = 6,00$	4,51	*
12,0-9,25	$2,75 / 0,6661 = 4,12$	4,51	NS
12,0-10,25	$1,75 / 0,6661 = 2,62$	4,51	NS
10,5-8,0	$2,5 / 0,6661 = 3,75$	4,51	NS
10,5-9,25	$1,25 / 0,6661 = 1,87$	4,51	NS
10,5-10,25	$0,25 / 0,6661 = 0,37$	4,51	NS
8,0-9,25	$-1,25 / 0,6661 = -1,87$	4,51	NS
8,0-10,25	$-2,25 / 0,6661 = -3,37$	4,51	NS
9,25-10,25	$-1,0 / 0,6661 = -1,50$	4,51	NS

Se concluye que la única diferencia significativa la muestran las variedades 1 y 3, y el resto de las comparaciones no muestran diferencias significativas entre sí. Los resultados son algo diferentes a los obtenidos con la mínima diferencia significativa, donde se aprecian tres grupos de significación. Puede comprobarse en este ejemplo que el test de Tukey se muestra más conservador que el de la MDS.

En el programa *Statistic*, siguiendo la secuencia:

Results >Multiple comparisons >All-pairwise comparisons

se elije Tukey para $\alpha = 0,05$ y se obtiene la siguiente salida:

```

Tukey HSD All-Pairwise Comparisons Test of Rendimien for Variedade

Variedade      Mean  Homogeneous Groups
1  12.000  A
2  10.500  AB
5  10.250  AB
4   9.250  AB
3   8.000  B

Alpha                0.05      Standard Error for Comparison  0.9421
Critical Q Value    4.515      Critical Value for Comparison  3.0078
Error term used: Bloques*Variedade, 12 DF
There are 2 groups (A and B) in which the means
are not significantly different from one another.
    
```

El programa muestra las medias, la asociación en grupos de significación, el valor de α , el valor crítico de Q que se recoge en las tablas, y el valor crítico para la comparación, que indica la diferencia mínima que debe existir entre dos medias para que su diferencia resulte significativa, valor que solo se supera al comparar las variedades 1 y 3.

Procedimientos de *Bonferroni* y de *Sidak*

Ambos métodos arrojan resultados muy parecidos y siguen la misma estrategia para protección del error tipo I. Ésta consiste en penalizar el valor de significación P . En el *test de Student*, que utiliza el método *LSD*, el valor de significación es la constante α . En cambio, en estos métodos el nivel de significación ajustado (p) es una función decreciente de m , siendo m el número de comparaciones. El método de *Bonferroni* emplea la función $p = \alpha / m$, mientras que el de *Sidak* usa la función $p = 1 - (1 - \alpha)^{\frac{1}{m}}$. Por tanto, en cualquiera de ellos la tasa del error tipo I es inferior a α , siendo generalmente el método de *Bonferroni* más conservador que el de *Sidak*. El problema de estos procedimientos es que resultan demasiado conservadores cuando el número de comparaciones (m) aumenta, por lo que sólo son útiles cuando el número de comparaciones es bajo.

La popularidad de ambos métodos, sobre todo el de *Bonferroni*, también se debe a que al ser tan conservadores presentan la ventaja de darle más credibilidad a las conclusiones del investigador. Además, estos procedimientos se pueden usar en muchas otras situaciones y contextos distintos del ANOVA, incluso asociados con tests no paramétricos.

Para ambos métodos, el valor crítico se obtiene igual que con la mínima diferencia significativa, pero utilizando el valor de p :

$$\text{Valor crítico} = t_{(p; \gamma_{\text{error}})} \cdot S_{\bar{D}} = t \sqrt{\frac{2CME}{n}}$$

La hipótesis nula se rechaza cuando la diferencia entre dos medias es superior o igual al valor crítico. Un problema para el cálculo manual es encontrar tablas *t de Student* para cualquier valor de α , algo superado en los programas informáticos.

Vamos a ilustrar el procedimiento de cálculo de uno de estos métodos, el de *Bonferroni*, con el mismo ejemplo anterior de la Tabla 7.1, que recoge los datos de un experimento en bloques al azar con cinco tratamientos (variedades) y cuatro bloques. El $CME = 1,775$ con 12 grados de libertad. Para un nivel de significación $\alpha=0.05$, el valor de $p = \frac{0,05}{10} = 0,005$, y el valor crítico resulta:

$$\text{Valor crítico} = t_{(0,005;12)} \sqrt{\frac{2 * 1,775}{4}} = 3,428 * 0,9421 = 3,23$$

El valor de t no aparece en la Tabla A1 que no muestra los valores para un nivel de significación de 0,005, aunque se puede intuir el valor aproximado en esa tabla.

Las comparaciones se muestran a continuación:

Comparaciones	Diferencia	Valor crítico	Significación
12,0-10,5	1,5	3,23	NS
12,0-8,0	4	3,23	*
12,0-9,25	2,75	3,23	NS
12,0-10,25	1,75	3,23	NS
10,5-8,0	2,5	3,23	NS
10,5-9,25	1,25	3,23	NS
10,5-10,25	0,25	3,23	NS
8,0-9,25	-1,25	3,23	NS
8,0-10,25	-2,25	3,23	NS
9,25-10,25	-1,0	3,23	NS

Los resultados son los mismos que los obtenidos con el test de Tukey.

En el programa *Statistic*, siguiendo la secuencia:

Results >Multiple comparisons >All-pairwise comparisons

se elije el test de Bonferroni y se obtiene la siguiente salida:

Bonferroni All-Pairwise Comparisons Test of Rendimien for Variedade

Variedade	Mean	Homogeneous Groups
1	12.000	A
2	10.500	AB
5	10.250	AB
4	9.250	AB
3	8.000	B

Alpha 0.05 Standard Error for Comparison 0.9421

Critical T Value 3.428 Critical Value for Comparison 3.2298

Error term used: Bloques*Variedade, 12 DF

There are 2 groups (A and B) in which the means are not significantly different from one another.

El programa muestra las medias, la asociación en grupos de significación, el valor de α , el valor de t para la prueba de dos colas, y el valor crítico para la comparación, que es el calculado anteriormente.

Procedimiento de Scheffé

Es el método que generalmente suele dar un valor crítico más elevado y, por tanto, el más conservador de los tests de comparación múltiple. El valor crítico para la comparación es una función de la F general del ANOVA, por lo que las comparaciones con este método nunca son significativas si la F no es significativa. Este método se recomienda cuando la comisión de un error tipo I pueda tener consecuencias muy graves. También es útil para establecer comparaciones sugeridas por los datos. El programa *Statistix* lo incluye, junto a la t de *Student*, en los contrastes ortogonales, pero esta opción sólo deberá tenerse en cuenta cuando el contraste no se haya establecido *a priori*, sino en función de los datos obtenidos, o cuando se realicen todas las comparaciones ortogonales posibles con un número elevado de medias.

Este test puede utilizarse aún cuando el número de repeticiones de cada media sea diferente. El valor crítico para la comparación es:

$$\text{Valor crítico} = \left[(t-1)F_{0,05} \right]^{1/2} \left[\left(\frac{1}{n_i} + \frac{1}{n_j} \right) \cdot CME \right]^{1/2}$$

siendo:

t el número de tratamientos

$F_{0,05}$ el valor de F en la Tabla A2

n el número de repeticiones de cada muestra

CME el cuadrado medio del error.

La hipótesis nula se rechaza cuando la diferencia entre dos medias es superior al valor crítico.

Tomando el mismo ejemplo anterior de la Tabla 7.1, que recoge los datos de un experimento en bloques al azar con cinco tratamientos (variedades) y cuatro bloques, se tiene que el $CME = 1,775$ con 12 grados de libertad. Los grados de libertad de los tratamientos son 4, por lo que el valor de F en la Tabla A2 resulta 3,26. El valor crítico es:

$$\text{Valor crítico} = \left[(5-1)3,26 \right]^{1/2} \left[\left(\frac{1}{4} + \frac{1}{4} \right) 1,775 \right]^{1/2} = 3,40$$

Las comparaciones entre medias se muestran a continuación:

Comparaciones	Diferencia	Valor crítico	Significación
12,0-10,5	1,5	3,40	NS
12,0-8,0	4	3,40	*
12,0-9,25	2,75	3,40	NS
12,0-10,25	1,75	3,40	NS
10,5-8,0	2,5	3,40	NS
10,5-9,25	1,25	3,40	NS
10,5-10,25	0,25	3,40	NS
8,0-9,25	-1,25	3,40	NS
8,0-10,25	-2,25	3,40	NS
9,25-10,25	-1,0	3,40	NS

Los resultados son los mismos que los obtenidos con el test de Tukey y con el de Bonferroni, pero nótese que el valor crítico para la comparación es el mayor de todos.

En el programa *Statistic*, siguiendo la secuencia:

Results >Multiple comparisons >All-pairwise comparisons

se elije el test de Scheffé y se obtiene la siguiente salida:

```

Scheffe All-Pairwise Comparisons Test of Rendimien for Variedade

Variedade      Mean  Homogeneous Groups
    1  12.000  A
    2  10.500  AB
    5  10.250  AB
    4   9.250  AB
    3   8.000  B

Alpha                0.05      Standard Error for Comparison  0.9421
Critical F Value    3.259      Critical Value for Comparison  3.4015
Error term used: Bloques*Variedade, 12 DF
There are 2 groups (A and B) in which the means
are not significantly different from one another.
    
```

El programa muestra las medias, la asociación en grupos de significación, el valor de α , el valor de F y el valor crítico para la comparación, que es el calculado anteriormente.

Otros métodos de comparación múltiple

Una variante del test *LSD* es el método de *Waller-Duncan*, también conocido como el *test LSD bayesiano (BLSD)*. Este test determina un valor crítico similar al método *LSD*, pero en este caso el investigador tiene la posibilidad de establecer una estimación subjetiva de la importancia relativa de los errores tipo I y tipo II. El programa *Statistix* no lo incluye entre los procedimientos de comparación múltiple, pero las tablas de los valores *t* de riesgo-promedio-mínimo aparecen en diversos textos de estadística, como en Steel y Torrie (1985).

Dos métodos utilizados habitualmente en experimentos agronómicos y en otras áreas de la biología son el de *amplitud múltiple de Duncan* y el de *Student-Neuman-Keuls (SNK)*. Ambos tienen en común que usan amplitudes múltiples para comparar. Es decir, no calculan un único valor crítico, sino varios, en función del número de medias y de la distancia o separación entre ellas. Son métodos menos conservadores que el de *Tukey*, pero algo más que el *LSD*. La desventaja principal de estos métodos es que son más complejos de utilizar y, sobre todo, que no permiten establecer intervalos de confianza. Por ello, no son muy utilizados actualmente y algunos programas, como *Statistix*, no los incluyen. Se han desarrollado algunas modificaciones del test *SNK* que eliminan algunos de sus inconvenientes y mejoran su potencia (menor riesgo de error tipo II), como el *test Shaffer-Ryan*, pero todavía son poco utilizados y no se incluyen en la mayoría de los programas estadísticos.

Otro método recomendado para comparaciones múltiples en experimentos con numerosas medias en los que es deseable obtener grupos de medias que no se solapen (por ejemplo, comparación de cultivares) es el de *Scott-Knott* basado en el análisis "cluster". No obstante, experimentos de simulación han demostrado que este método protege menos contra el error tipo I que el método *FLSD* (Willavize et al., 1980), por lo que no ha tenido una amplia difusión.

Como conclusión sobre los diferentes contrastes y procedimientos de comparación múltiple hay que destacar que no existe un único método y que la elección del más adecuado depende del experimento a analizar. Lo más recomendable es planificar las comparaciones *a priori* y aplicar el método de los contrastes ortogonales, o los contrastes polinómicos para un factor cuantitativo con pocos niveles, o bien la comparación con un testigo de referencia o con el mejor. Cuando no es posible planificar las comparaciones *a priori*, se debería limitar el número de comparaciones y hacer sólo las que realmente son relevantes, aplicando métodos conservadores como los de *Bonferroni* o *Sidak*. Finalmente, cuando interese realizar todas las comparaciones posibles de pares de medias, el método más recomendado por diversos autores es el de *Tukey* (Gilligan, 1986; Hsu, 1996; Analytical Software, 2003).

CAPÍTULO 6

DISEÑO COMPLETAMENTE ALEATORIO

Es el diseño más simple que resulta de utilidad cuando las unidades experimentales son esencialmente homogéneas, es decir, cuando la variación entre ellas es pequeña y agruparlas en bloques sería poco más que un proceso aleatorio. Este es el caso en muchos tipos de experimentos de laboratorio, en los que una cantidad de material está completamente mezclado y luego se divide en porciones pequeñas para formar unidades experimentales a las cuales se les asigna los tratamientos de forma aleatoria. Casos similares son los experimentos con animales o plantas con condiciones ambientales muy parecidas.

Los tratamientos se asignan a las unidades experimentales de forma completamente aleatoria, sin ninguna restricción. El diseño completamente aleatorio es flexible en cuanto a que el número de tratamientos y de repeticiones sólo está limitado por el número de unidades experimentales disponibles. El número de repeticiones puede variar de un tratamiento a otro, aunque lo ideal sería tener un número igual por tratamiento. Como la aleatorización no tiene restricciones, el error experimental contiene toda la variación entre unidades experimentales excepto la debida a los tratamientos, por lo que aporta el máximo de grados de libertad para el error.

El principal inconveniente del diseño completamente aleatorio es su poca precisión si las unidades experimentales no son uniformes.

EL MODELO LINEAL ADITIVO

El modelo lineal aditivo para clasificaciones de una vía, es decir, para un diseño completamente al azar, viene dado por la siguiente ecuación:

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

donde Y_{ij} es la observación en la unidad experimental j (repetición) a la que se le ha aplicado el tratamiento i , μ es la media general del experimento, α es el efecto del tratamiento i y ε_{ij} es el efecto del error de la observación, un elemento aleatorio de una población normal de media cero y desviación típica σ . En otras palabras, cualquier valor observado es la suma de esos tres componentes.

Los valores de las esperanzas medias para un diseño completamente al azar son:

Fuente de variación	Grados de libertad	Cuadrado medio	
		Modelo I (Efectos fijos)	Modelo II (Efectos aleatorios)
Tratamientos	t-1	$\sigma^2 + r \sum \alpha_t^2 / (t-1)$	$\sigma^2 + r \sigma_t^2$
Error	t(r-1)	σ^2	σ^2
Total	tr-1		

DISEÑO Y ANÁLISIS DE VARIANZA

Supongamos un experimento con cuatro tratamientos (T1, T2, T3 y T4) que queremos repetir cuatro veces. Un esquema del diseño podría ser:

T3	T4	T3	T1	T2	T1	T1	T4	T2	T3	T1	T3	T4	T4	T2	T2
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

Puede observarse que hay una aleatorización total, sin ninguna restricción, pudiendo encontrarse dos tratamientos seguidos.

El análisis de varianza se realizaría de la siguiente manera:

Fuente de Variación	gl	SC	CM	F
Tratamientos	t-1	SCTR	CMTR	CMTR/CME
Error	t(r-1)	SCE	CME	
Total	tr-1	SCT		

siendo:

t = número de tratamientos

r = número de repeticiones

El factor de corrección:
$$C = \frac{\left(\sum_{i=1}^n x\right)^2}{n} = \frac{(\text{total.de.observaciones})^2}{\text{número.total.observaciones}}$$

$$SCT = \sum_{i,j} x_{ij}^2 - C$$

$$SCTR = \sum_r T_i^2 - C = \frac{T_1^2 + \dots + T_t^2}{r} - C$$

$$SCE = SCT - SCTR$$

$$CMTR = \frac{SCTR}{t-1}$$

$$CME = \frac{SCE}{t(r-1)}$$

En la Tabla 6.1 se recogen los datos de un experimento en el que se ensayaron cuatro variedades de trigo para determinar su comportamiento en un terreno. Las variedades se han aleatorizado totalmente cuatro veces.

Tabla 6.1. Rendimiento (t/ha) de cuatro variedades de trigo en cada una de las cuatro parcelas elementales.

	VAR 1	VAR 2	VAR 3	VAR 4	TOTAL
	4,58	3,57	5,68	4,68	
	4,33	3,21	5,43	4,21	
	4,60	3,59	5,10	4,36	
	3,26	2,35	4,57	3,79	
T_i	16,77	12,72	20,78	17,04	67,31
$\sum_j x_{ij}^2$	71,51	41,46	108,64	73,00	294,61

Así pues, tenemos que:

$$C = \frac{67,31^2}{16} = 283,16$$

$$SCT = 294,61 - 283,16 = 11,45$$

$$SCTR = \frac{16,77^2 + \dots + 17,04^2}{4} - 283,16 = 291,30 - 283,16 = 8,14$$

$$SCE = 11,45 - 8,14 = 3,31$$

$$CMTR = \frac{8,14}{3} = 2,71$$

$$CME = \frac{3,31}{12} = 0,27$$

$$F = \frac{2,71}{0,27} = 10,04$$

Si todos estos valores los agrupamos en el correspondiente cuadro tenemos:

Fuente de Variación	gl	SC	CM	F
Tratamientos	3	8,14	2,71	10,04 **
Error	12	3,31	0,27	
Total	15	11,45		

Si vamos a la Tabla A2 de los valores de F, se comprueba que el valor de F para 3 grados de libertad del numerador y 12 del denominador es de 3,49 para $P = 0.05$ y de 5,95 para $P = 0,01$, valores inferiores al calculado de 10,04, luego se rechaza la hipótesis nula de que no había diferencia entre los tratamientos y se acepta la hipótesis alternativa de que los tratamientos aplicados en el experimento cambian las medias de las poblaciones a las que se les aplican, declarando que el experimento es significativo (lo que se denota con asteriscos detrás del valor de F).

Ahora se puede estar interesados en conocer entre qué tratamientos existen realmente diferencias significativas; para conocer este hecho hay que calcular un estadístico que, en base al cuadrado medio del error del experimento, que representa la mejor estimación de la varianza poblacional, indique qué diferencias entre las medias de los tratamientos son las estadísticamente significativas. Existen varios estadísticos para discernir qué tratamientos son diferentes de otros, como se ha recogido en el Capítulo 5. Si aplicamos la mínima diferencia significativa, tenemos que:

$$MDS_{0,05} = t_{0,05} * S_{\bar{D}}$$

siendo

$t_{0,05}$ = valor tabulado de t para los grados de libertad del error (Tabla A1).

$S_{\bar{D}} = \sqrt{\frac{2s^2}{r}}$, el error estándar de la diferencia entre dos medias.

s^2 = varianza del error (CME).

r = número de observaciones por media (en este caso coincide con el número de repeticiones).

En el ejemplo:

$t_{0,05}$ para 12 grados de libertad = 2,179

$$s_{\bar{D}} = \sqrt{\frac{2 \times 0,27}{4}} = 0,37$$

Luego

$$MDS = 2,179 \times 0,37 = 0,81$$

Si se ordenan las medias de los cuatros tratamientos de mayor a menor:

T	Media
3	5.1950
4	4.2600
1	4.1925
2	3.1800

se observa que la diferencia entre el tratamiento 3 y el 4 y el 1 es mayor que la MDS y, a su vez, las diferencias entre los tratamientos 4 y 1 con el tratamiento 2 son también mayores que 0,81, por lo que estas diferencias las calificamos de estadísticamente significativas.

CÁLCULO CON EL PROGRAMA *Statistix (SX)*

En el programa SX introducir los datos experimentales definiendo las variables “Variedades” y “Rendimiento”, obteniendo la siguiente tabla de valores:

Variedades	Rendimiento
1	4.58
1	4.33
1	4.60
1	3.26
2	3.57
2	3.21
2	3.59
2	2.35
3	5.68
3	5.43
3	5.10
3	4.57
4	4.68
4	4.21
4	4.36
4	3.79

Una vez introducidos los datos experimentales, seguir la siguiente secuencia:

Statistic>Linear Models>Analysis of Variance>Completely Randomized Design

Una vez dentro de esta opción, en el cajetín correspondiente escoger como variable dependiente "Rendimiento" y como variable de tratamientos "Variedades". Pulsar OK y se obtiene el siguiente cuadro de análisis de la varianza:

Completely Randomized AOV for Rendimien						
Source	DF	SS	MS	F	P	
Variedade	3	8.1356	2.71186	9.82	0.0015	
Error	12	3.3146	0.27621			
Total	15	11.4501				
Grand Mean		4.2069	CV	12.49		
				Chi-Sq	DF	P
Bartlett's Test of Equal Variances				0.83	3	0.8423
Cochran's Q				0.3634		
Largest Var / Smallest Var				2.9397		
Component of variance for between groups				0.60891		
Effective cell size				4.0		
Variedade	Mean					
1	4.1925					
2	3.1800					
3	5.1950					
4	4.2600					
Observations per Mean		4				
Standard Error of a Mean		0.2628				
Std Error (Diff of 2 Means)		0.3716				

La salida del programa incluye la tabla del análisis de varianza, cuyos datos son sensiblemente similares a los obtenidos manualmente por el efecto del redondeo de los decimales, la media general (4,2069), el coeficiente de variación (CV = 12,49 %), el test de Bartlett para la igualdad de la varianzas (véase el Capítulo 4), las medias de cada tratamiento y los valores de s y de S_D .

Con objeto de detectar los tratamientos que difieren significativamente entre sí, seguir la secuencia:

Results>Multiple comparisons>All-pairwise comparisons

y escoger alguno de los métodos ofrecidos para calcular la significación de las diferencias.

Si se escoge la LSD (mínimas diferencias significativas) se obtiene el siguiente cuadro:

LSD All-Pairwise Comparisons Test of Rendimiento by Variedades

Variedad	Mean	Homogeneous Groups
3	5.1950	A
4	4.2600	B
1	4.1925	B
2	3.1800	C

Alpha	0.05	Standard Error for Comparison	0.3716
Critical T Value	2.179	Critical Value for Comparison	0.8097

There are 3 groups (A, B, etc.) in which the means are not significantly different from one another.

Se muestran claramente cuales son los valores tomados para t , el valor de $S_{\bar{D}}$ y el valor de LSD (MDS), que separa las medias en tres grupos dentro de los cuales no hay diferencias significativas y entre los cuales si las hay. En definitiva, se puede afirmar que la variedad 3 produce significativamente el mayor rendimiento, que no existen diferencias significativas entre las variedades 4 y 1 y que la variedad 2 es la de menor rendimiento.

CAPÍTULO 7

DISEÑO EN BLOQUES AL AZAR

Este diseño es el más común en la experimentación agrícola y se utiliza cuando las unidades experimentales pueden agruparse de una forma sustancialmente lógica. El número de unidades experimentales en cada grupo, o bloque, ha de ser igual al número de tratamientos.

El objeto del agrupamiento en bloques es tener las unidades experimentales de forma tan uniforme como sea posible, de manera que las diferencias observadas sean fundamentalmente debidas a los tratamientos. En otras palabras, la variabilidad entre unidades experimentales que se encuentran en bloques diferentes debe de ser mayor que la que exista entre unidades que están en el mismo bloque si no se aplica ningún tratamiento.

La variabilidad entre bloques no afecta a las diferencias entre las medias de los tratamientos, puesto que cada tratamiento aparece en cada bloque. En el Capítulo 2 se recogen algunas consideraciones en relación a los bloques.

Durante el curso del experimento todas las unidades del mismo bloque deben ser tratadas tan uniformemente como sea posible en todos los aspectos. Cualquier cambio en alguna de las técnicas o condiciones que pueda afectar a los resultados se deberá hacer en el bloque completo. Así, por ejemplo, si la recolección se ha de prolongar durante varios días, todas las parcelas del mismo bloque deberán recolectarse el mismo día. En la toma de datos, lo ideal es que el mismo individuo tomara todos los datos del experimento, pero si ello no es posible, al menos los datos de cada bloque deberían ser tomados por la misma persona. Estas prácticas ayudan a controlar la variación entre bloques y, por lo tanto, el error experimental, y al mismo tiempo no contribuyen en nada a las diferencias entre tratamientos. La variación entre bloques es aritméticamente extraída del error experimental.

Este diseño es totalmente equilibrado, pues cada tratamiento aparece un número igual de veces en cada bloque y cada bloque contiene todos los tratamientos. Se dice que los bloques y tratamientos son ortogonales unos con respecto al otro.

Como se ha indicado anteriormente, este diseño se utiliza con más frecuencia que cualquier otro y si da una precisión satisfactoria, no tiene sentido utilizar otro. El mayor inconveniente es que cuando la variación entre unidades experimentales dentro del mismo bloque es grande, la varianza debida al error es grande. Esto ocurre, a veces, cuando hay muchos tratamientos y no es posible asegurar una agrupación uniforme en bloques. En tales situaciones es necesario utilizar otros diseños.

MODELO LINEAL PARA UN DISEÑO EN BLOQUES COMPLETAMENTE AL AZAR

El modelo lineal para un diseño en bloques al azar viene dado por la siguiente ecuación:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

siendo

- Y_{ij} = una observación del tratamiento i en el bloque j
- μ = la media general del experimento
- α_i = el efecto de los tratamientos
- β_j = el efecto de los bloques
- ε_{ij} = el efecto del error

La esperanza de los cuadrados medios para un diseño de bloques al azar es:

Fuente de variación	Grados de libertad	Modelo I (Efectos fijos)	Modelo II (Efectos aleatorios)
Bloques	$r-1$	$\sigma_\varepsilon^2 + t \sum \beta_j^2 (r-1)$	$\sigma_\varepsilon^2 + t \sigma_\beta^2$
Tratamientos	$t-1$	$\sigma_\varepsilon^2 + r \sum \alpha_i^2 (t-1)$	$\sigma_\varepsilon^2 + r \sigma_\alpha^2$
Error	$(r-1)(t-1)$	σ_ε^2	

DISEÑO Y ANÁLISIS DE VARIANZA

En el supuesto de un experimento con cuatro tratamientos y tres bloques, el esquema del diseño podría ser el siguiente:

Bloque 1	3	1	2	4
Bloque 2	1	3	4	2
Bloque 3	4	1	2	3

Dentro de cada bloque, que contiene cada uno de los cuatro tratamientos, éstos se asignan al azar a las unidades experimentales. Los bloques no tienen que ser adyacentes, e incluso cada bloque puede estar en una localidad distinta.

El análisis de la varianza de un diseño de este tipo se efectúa de forma similar al de un diseño completamente al azar añadiéndole una fuente de variación nueva, la de los bloques. Así, pues, tenemos:

Fuente de Variación	gl	SC	CM	F
Tratamientos	t-1	SCTR	CMTR	CMTR/CME
Bloques	r-1	SCB	CMB	CMB/CME
Error	(r-1)(t-1)	SCE	CME	
Total	tr-1	SCT		

siendo:

t = número de tratamientos

r = número de bloques

El factor de corrección:

$$C = \frac{\left(\sum_{i=1}^n x\right)^2}{n} = \frac{(\text{total.de.observaciones})^2}{\text{número.total.observaciones}}$$

$$SCT = \sum_{i,j} x_{ij}^2 - C$$

$$SCTR = \frac{\sum T_i^2}{r} - C$$

$$SCB = \frac{\sum B_j^2}{t} - C$$

$$SCE = SCT - SCTR - SCB$$

$$CMTR = \frac{SCTR}{t-1}$$

$$CMB = \frac{SCB}{r-1}$$

$$CME = \frac{SCE}{(r-1)(t-1)}$$

Supongamos un experimento de 5 variedades de maíz que se han ensayado en cuatro bloques al azar y cuyos rendimientos se recogen en la Tabla 7.1.

Tabla 7.1. Rendimiento (t/ha) de cinco variedades de maíz en cada uno de los cuatro bloques.

	VAR 1	VAR 2	VAR 3	VAR 4	VAR 5	B_j	B_j^2
B1	16	14	9	12	13	64	4096
B2	12	12	7	9	12	52	2704
B3	10	9	8	9	8	44	1936
B4	10	7	8	7	8	40	1600
T_i	48	42	32	37	41	200	
$\sum_j x_{ij}^2$	600	470	258	355	441	2124	
T_i^2	2304	1764	1024	1369	1681		

Tenemos que:

$$C = \frac{200^2}{20} = 2000$$

$$SCT = 2124 - 2000 = 124$$

$$SCTR = \frac{2304 + \dots + 1681}{4} - 2000 = 35,5$$

$$SCB = \frac{4096 + \dots + 1600}{5} - 2000 = 67,2$$

$$SCE = 124 - 35,5 - 67,2 = 21,3$$

$$CMTR = \frac{35,5}{4} = 8,87$$

$$CMB = \frac{67,2}{3} = 22,4$$

$$CME = \frac{21,3}{12} = 1,77$$

Si toda esta información la agrupamos en una tabla de análisis de la varianza, tenemos:

Fuente de variación	gl	SC	CM	F
Tratamientos	4	35,5	8,87	5,01 *
Bloques	3	67,2	22,4	12,65 **
Error	12	21,3	1,77	
Total	19	124,0		

En la Tabla A2 de los valores de F se comprueba que el valor de F para 4 grados de libertad del numerador y 12 del denominador, que corresponde a los tratamientos, es de 3,26 para $P=0,05$ y de 5,41 para $P=0,01$, el primero inferior al valor calculado y el segundo superior, luego se rechaza la hipótesis nula de que no había diferencia entre las variedades y se asume que los rendimientos de éstas representan medias poblacionales diferentes a un nivel de probabilidad de $P \leq 0,05$, lo que se representa por un asterisco. En el caso de los bloques, se comprueba que existen diferencias significativas entre ellos, al menos a un nivel de $P \leq 0,01$. Esto indica que la diferencia que existe entre los bloques ha sido extraída del error, disminuyendo CME y, en consecuencia, aumentando la precisión del experimento.

Ahora puede interesar conocer entre qué variedades existen realmente esas diferencias significativas; para ello se debe calcular algún estadístico que, en base al cuadrado medio del error del experimento, indique qué diferencias entre las medias de los tratamientos son las estadísticamente significativas. Si se utiliza la mínima diferencia significativa (MDS), tenemos:

$$MDS = t_{0,05} \cdot \sqrt{\frac{2s^2}{r}} = 2,18 \cdot \sqrt{\frac{2 \times 1,77}{4}} = 2,18 \times 0,94 = 2,05$$

Si se ordenan las medias de los rendimientos de las cinco variedades por orden decreciente, tenemos que:

Variedad	Rendimiento
1	12,000
2	10,500
5	10,250
4	9,250
3	8,000

Las diferencias entre los rendimientos de las 3 primeras variedades (1, 2 y 5) son menores que 2,05, luego no existen diferencias estadísticamente significativas entre el rendimiento de

esas tres variedades; sin embargo, si las hay entre la variedad 1 y la variedad 4 y entre la variedad 1 y la variedad 3, así como entre las variedades 2 y 5 y la variedad 3.

CÁLCULO CON EL PROGRAMA *Statistix (SX)*

En el programa SX introducir los datos experimentales definiendo las variables “Variedades”, “Bloques” y “Rendimiento” y se obtiene la siguiente tabla:

Bloques	Variedades	Rendimiento
1	1	16
1	2	14
1	3	9
1	4	12
1	5	13
2	1	12
2	2	12
2	3	7
2	4	9
2	5	12
3	1	10
3	2	9
3	3	8
3	4	9
3	5	8
4	1	10
4	2	7
4	3	8
4	4	7
4	5	8

Una vez introducidos los datos experimentales, seguir la siguiente secuencia:

Statistic >Linear Models >Analysis of Variance >Randomized Complete Block.

Una vez dentro de esta opción, en el cajetín correspondiente introducir como variable dependiente “Rendimiento”, como variable de bloques “Bloques” y como variable de tratamientos, “Variedades”. Pulsar OK y se obtiene el siguiente cuadro de análisis de la varianza:

Randomized Complete Block AOV Table for Rendimien

Source	DF	SS	MS	F	P
Bloques	3	67.200	22.4000		
Variedade	4	35.500	8.8750	5.00	0.0132
Error	12	21.300	1.7750		
Total	19	124.000			

Grand Mean 10.000 CV 13.32

Tukey's 1 Degree of Freedom Test for Nonadditivity

Source	DF	SS	MS	F	P
Nonadditivity	1	10.0366	10.0366	9.80	0.0096
Remainder	11	11.2634	1.0239		

Relative Efficiency, RCB 2.76

Means of Rendimien for Variedade

Variedade	Mean
1	12.000
2	10.500
3	8.000
4	9.250
5	10.250

Observations per Mean 4
 Standard Error of a Mean 0.6661
 Std Error (Diff of 2 Means) 0.9421

Los datos son sensiblemente similares a los obtenidos manualmente, con el efecto del redondeo por los decimales. En este caso se obtiene directamente la probabilidad ($P = 0,0132$) de obtener un valor igual o superior al de la $F(5,00)$, por lo que efectivamente, no se acepta la hipótesis nula de no diferencia entre Variedades. La salida aporta también la media general (10,0) el coeficiente de variación (13,32 %) y el test de Tukey para no aditividad. En este caso se aprecia que ese efecto existe, por lo que podría recomendarse la realización de alguna transformación de los datos (véase el Capítulo 4). La eficiencia relativa indica el grado de precisión que se consigue con el establecimiento de bloques comparada con un diseño totalmente aleatorio; en este caso el valor es muy superior a la unidad, lo que indica que se ha ganado mucha precisión al establecer los bloques. También se muestran las medias de cada tratamiento y los valores de s y de $s_{\bar{D}}$.

Con objeto de detectar qué diferencias entre tratamientos son significativas, seguir la siguiente secuencia:

Results >Multiple comparisons >All-pairwise comparisons.

y escoger alguno de los métodos ofrecidos para calcular la significación de las diferencias.

Si se elige la LSD (mínima diferencia significativa) se obtiene el siguiente cuadro:

LSD All-Pairwise Comparisons Test of Rendimiento for Variedades			
Variedad	Mean	Homogeneous	Groups
1	12.000	A	
2	10.500	AB	
5	10.250	AB	
4	9.250	BC	
3	8.000	C	

Alpha	0.05	Standard Error for Comparison	0.9421
Critical T Value	2.179	Critical Value for Comparison	2.0526
Error term used: V002*V001, 12 DF			
There are 3 groups (A, B, etc.) in which the means are not significantly different from one another.			

Como puede observarse, se explica claramente cuales son los valores tomados para t , el valor de $s_{\bar{D}}$ y el valor de la LSD, que separa las medias en 3 grupos dentro de los cuales no hay diferencias significativas y entre los cuales si las hay. Las variedades 1, 2 y 5 producen similarmente y las tres significativamente más que la 3; la 2 y la 5 no se diferencian significativamente de la 4, aunque sí de la 3, mientras que la 4 y la 3 rinden de forma similar.

CAPÍTULO 8

CUADRADO LATINO

En el diseño en Cuadrado Latino el número de parcelas elementales es igual al cuadrado del número de tratamientos, y éstos se disponen de dos maneras diferentes, por filas y por columnas. Cada tratamiento se presenta una y sólo una vez en cada fila y en cada columna, representando cada fila y cada columna un bloque completo. Mediante un análisis adecuado es posible eliminar del error la variabilidad debida a diferencias tanto en filas como en columnas.

Este diseño se puede utilizar en aquellos experimentos donde hay claramente dos fuentes de variación. Concretamente, en experimentos de campo permite la eliminación de la variación proveniente de las diferencias de fertilidad en el suelo en dos direcciones, o bien la variación de la fertilidad del suelo en una dirección y el gradiente de humedad generado al dar los riegos en otra dirección determinada. También puede ser útil en experimentos de invernadero en donde la iluminación y la temperatura pueden ser dos fuentes de variación claras que deberían extraerse del error experimental.

Es necesario indicar que los bloques, tanto en este diseño como en el de bloques al azar, no necesariamente tienen que ser físicos; por ejemplo, al probar distintas mezclas de combustible en varios tractores conducidos por varios tractoristas, la variación debida a ambas variables podría ser extraída del error experimental utilizando un diseño en cuadrado latino.

El principal inconveniente del Cuadrado Latino es que el número de filas, columnas y tratamientos debe ser el mismo. Así pues, si hay muchos tratamientos, el número de unidades experimentales resulta excesivo y poco práctico. A estos efectos, los cuadrados latinos suelen disponer entre cuatro y ocho tratamientos. El análisis también se complica si se pierden unidades experimentales en el transcurso del experimento.

MODELO LINEAL PARA UN DISEÑO EN CUADRADO LATINO

El modelo lineal para un diseño en cuadrado latino viene dado por la siguiente ecuación:

$$Y_{ij(t)} = \mu + \alpha_{(t)} + \beta_i + k_j + \varepsilon_{ij}$$

siendo

$Y_{ij(t)}$ = una observación de la fila i en la columna j para el tratamiento t .

μ = la media general del experimento.

$\alpha_{(t)}$ = el efecto de los tratamientos.

β_i = el efecto de las filas.

k_j = el efecto de las columnas.

ε_{ij} = el efecto del error.

Las esperanzas de los cuadrados medios para un Cuadrado Latino resultan:

Esperanzas de los cuadrados medios				
Fuente de variación	Grados de libertad	Modelo I (Efectos fijos)	Modelo II (Efectos aleatorios)	
Filas	r-1	$\sigma^2 + r(\sum \beta_i^2)/(r-1)$	$\sigma^2 + r\sigma_\beta^2$	
Columnas	r-1	$\sigma^2 + r(\sum k_j^2)/(r-1)$	$\sigma^2 + r\sigma_k^2$	
Tratamientos	r-1	$\sigma^2 + r(\sum \alpha_i^2)/(r-1)$	$\sigma^2 + r\sigma_t^2$	
Error	(r-1)(r-2)	σ^2	σ^2	

DISEÑO Y ANÁLISIS DE VARIANZA

Una disposición en cuadrado latino con 4 tratamientos podría ser la siguiente:

A	D	C	B
B	C	A	D
D	A	B	C
C	B	D	A

donde cada tratamiento es atribuido aleatoriamente a cada parcela elemental dentro de cada fila y de cada columna. Esto puede hacerse de varias maneras; por ejemplo, pueden sortearse los tratamientos dentro de cada fila y luego sortear las columnas.

El análisis de varianza en el Cuadrado latino sería:

Fuente de Variación	gl	SC	CM	F
Tratamientos	r-1	SCTR	CMTR	CMTR/CME
Filas	r-1	SCF	CMF	CMF/CME
Columnas	r-1	SCC	CMC	CMC/CME
Error	(r-1)(r-2)	SCE	CME	
Total	r ² -1	SCT		

siendo:

r = número de filas = n° de columnas = n° de tratamientos

El factor de corrección:

$$C = \frac{\left(\sum_{i=1}^n x\right)^2}{r^2}$$

$$SCT = \sum_{i,j} x_{ij}^2 - C$$

$$SCTR = \frac{\sum T_i^2}{r} - C$$

$$SCF = \frac{\sum F_i^2}{r} - C$$

$$SCC = \frac{\sum C_j^2}{r} - C$$

$$SCE = SCT - SCF - SCC - SCTR$$

$$CMTR = \frac{SCTR}{r-1}$$

$$CMF = \frac{SCF}{r-1}$$

$$CMC = \frac{SCC}{r-1}$$

$$CME = \frac{SCE}{(r-1)(r-2)}$$

Como ejemplo, se va a analizar un experimento realizado con cuatro variedades de maíz en el que se sospechaba que había un gradiente de fertilidad en un sentido y otro gradiente producido por la forma en la que se tenía que regar la parcela. Los datos correspondientes al rendimiento de las variedades de maíz se presentan en la Tabla 8.1.

Tabla 8.1.- Rendimiento (t/ha) de cuatro variedades de maíz (A, B, C y D) en cada parcela elemental de un Cuadrado Latino. Las filas se han dispuesto siguiendo el gradiente de fertilidad y las columnas siguiendo el gradiente de riego.

		Columnas			
		C 10,5	D 7,7	B 12,0	A 13,2
Filas	B	11,1	A 12,0	C 10,3	D 7,5
	D	5,8	C 12,2	A 11,2	B 13,7
	A	11,6	B 12,3	D 5,9	C 10,2

La resolución de este experimento de una forma convencional se haría exactamente igual que la de un bloques al azar, con la particularidad de que habría que añadir un sumando más a la descomposición de la suma de cuadrados total, pues en este caso se disponen de dos bloques: filas y columnas. Se disponen los datos en filas y en columnas como se presenta en la Tabla 8.2.

Tabla 8.2.- Tabla de datos en filas y columnas del Cuadrado Latino.

Fila	Columna				Suma (F _i)
	1	2	3	4	
1	10,5	7,7	12,0	13,2	43,4
2	11,1	12,0	10,3	7,5	40,9
3	5,8	12,2	11,2	13,7	42,9
4	11,6	12,3	5,9	10,2	40,0
Suma (C_j)	39,0	44,2	39,4	44,6	167,2

A continuación se disponen los datos como en la Tabla 8.3 para recoger las sumas y medias de tratamientos.

Tabla 8.3.- Totales y medias de tratamientos del Cuadrado Latino.

Tratamientos (Variedades)	A	B	C	D	Suma
Total (T_{\uparrow})	48,0	49,1	43,2	26,9	167,2
Media	12,0	12,27	10,8	6,72	10,45

Los cálculos serían:

$$C = \frac{167,2^2}{4^2} = 1747,24$$

$$SCT = (10,5^2 + \dots + 10,2^2) - 1747,24 = 1837,64 - 1747,24 = 90,4$$

$$SCF = \frac{(43,4^2 + \dots + 40,0^2)}{4} - C = \frac{6996,78}{4} - 1747,24 = 1,955$$

$$SCC = \frac{(39,0^2 + \dots + 44,6^2)}{4} - C = \frac{7016,16}{4} - 1747,24 = 6,8$$

$$SCTR = \frac{(48,0^2 + \dots + 26,9^2)}{4} - C = \frac{7304,66}{4} - 1747,24 = 78,925$$

$$SCE = 90,4 - 1,955 - 6,8 - 78,925 = 2,72$$

$$CMF = \frac{1,955}{3} = 0,65$$

$$CMC = \frac{6,8}{3} = 2,27$$

$$CMTR = \frac{78,925}{3} = 26,308$$

$$CME = \frac{2,72}{6} = 0,453$$

La tabla del análisis de varianza quedaría, pues:

Fuente de variación	gl	SC	CM	F
Variedades	3	78,925	26,308	58,07**
Filas	3	1,955	0,65	1,43 ^{NS}
Columnas	3	6,8	2,27	5,01*
Error	6	2,72	0,453	
Total	15	90,4		

En la Tabla A2 de los valores de F se observa que los valores tabulados para 3 grados de libertad del numerador y 6 del denominador son 4,76 para un nivel de $P=0,05$ y 9,78 para $P=0,01$. El valor de F de los tratamientos (58,07) supera ampliamente estos valores, lo que permite rechazar la hipótesis nula y asumir que las diferencias entre variedades son altamente significativas. El valor de F correspondiente a filas es inferior a los valores tabulados, lo que indica que la fertilidad del suelo, en este caso, no parece contribuir realmente al error experimental, pero sí lo hace la forma de riego, pues la F de columnas es mayor que la F tabulada al nivel de probabilidad del 5%.

Si se realiza la separación de medias por la mínima diferencia significativa (MDS), tenemos que:

$$MDS_{0,05} = t_{0,05} * s_{\bar{D}}$$

siendo

$t_{0,05}$ = valor tabulado de t para los grados de libertad del error (Tabla A1).

$s_{\bar{D}} = \sqrt{\frac{2s^2}{r}}$, el error estándar de la diferencia entre dos medias.

s^2 = varianza del error (CME).

r = número de observaciones por media.

En el ejemplo tenemos:

$$s_{\bar{D}} = \sqrt{\frac{2 * 0,453}{4}} = 0,4759$$

$$MDS = 2,447 * 0,4759 = 1,1645$$

Si se vuelve a la Tabla 8.3, que recoge la media de los tratamientos, se observa que, de acuerdo con la MDS, las variedades A y B no difieren estadísticamente en sus resultados, pero

sí lo hacen de las variedades C y D y, a la vez, estas últimas también muestran diferencias significativas entre sí.

CÁLCULO CON EL PROGRAMA *Statistix* (SX)

En el programa SX introducir los datos experimentales definiendo las variables “Filas”, “Columnas”, “Variedades”, y “Rendimiento” teniendo en cuenta que el programa solo admite caracteres numéricos en la hoja de datos, por lo que a cada variedad hay que asignarle un número (A=1; B=2; C=3; D=4). La tabla de datos quedaría, pues, de la siguiente forma:

Fila	Columna	Variedad	Rendimiento
1	1	3	10.5
1	2	4	7.7
1	3	2	12
1	4	1	13.2
2	1	2	11.1
2	2	1	12
2	3	3	10.3
2	4	4	7.5
3	1	4	5.8
3	2	3	12.2
3	3	1	11.2
3	4	2	13.7
4	1	1	11.6
4	2	2	12.3
4	3	4	5.9
4	4	3	10.2

Una vez introducidos los datos experimentales, seguir la siguiente secuencia:

Statistic >Linear Models >Analysis of Variance >Latin Square Design

Una vez dentro de esta opción, en el cajetín correspondiente indicar como variable dependiente “Rendimiento”, como variable de filas (row) “Filas”, como variable de columnas (columns) “Columna” y como variable de tratamientos, “Variedad”. Pulsar OK y se obtiene el siguiente cuadro de análisis de la varianza:

Latin Square AOV Table for Rendimien

Source	DF	SS	MS	F	P
Fila	3	1.9550	0.6517		
Columna	3	6.8000	2.2667		
Variedad	3	78.9250	26.3083	58.03	0.0001
Error	6	2.7200	0.4533		
Total	15	90.4000			

Grand Mean 10.450 CV 6.44

Tukey's 1 Degree of Freedom Test for Nonadditivity

Source	DF	SS	MS	F	P
Nonadditivity	1	0.01412	0.01412	0.03	0.8780
Remainder	5	2.70588	0.54118		

Relative Efficiency

Completely Randomized Design	1.89
Randomized Complete Block, Fila	1.87
Randomized Complete Block, Columna	1.04

Means of Rendimien for Variedad

Variedad	Mean
1	12.000
2	12.275
3	10.800
4	6.725
Observations per Mean	4
Standard Error of a Mean	0.3367
Std Error (Diff of 2 Means)	0.4761

En este caso se obtiene directamente la probabilidad ($P = 0,0001$) de obtener un valor igual o superior al de la $F(58,03)$, por lo que efectivamente, no se acepta la hipótesis nula de no diferencia entre variedades. Obsérvese que la salida no incluye los valores de F para filas y columnas, que han de calcularse a mano si se quiere esa información. Como en otros diseños, la salida incluye los valores de la media general (10,45), del coeficiente de variación (6,44 %) y el test de Tukey para la no aditividad, que en este caso no existe.

El programa calcula la eficiencia del cuadrado latino sobre un diseño completamente al azar o bien sobre diseños en bloque utilizando únicamente las filas como bloques o bien las columnas. En el ejemplo, el diseño en cuadrado latino es bastante eficiente en comparación con un diseño completamente al azar y también con un diseño en bloques al azar si se hubieran escogido la filas como bloques, consiguiendo una mejora sobre esos diseños del 89 % y del 87%, respectivamente, pero tan solo de un 4% comparado con un diseño en bloques al azar si se hubieran escogido las columnas como bloques. Por último, también se muestran las medias de cada tratamiento y los valores de s y de $S_{\bar{D}}$.

Para realizar la separación de medias, seguir la secuencia:

Results >Multiple comparisons >All-pairwise comparisons.

y elegir uno de los métodos ofrecidos para calcular la significación de las diferencias. Si se elige LSD (mínima diferencia significativa) se obtiene:

LSD All-Pairwise Comparisons Test of REND for VAR			
VAR	Mean	Homogeneous Groups	
B	12.275	A	
A	12.000	A	
C	10.800	B	
D	6.725	C	
Alpha	0.05	Standard Error for Comparison	0.4761
Critical T Value	2.447	Critical Value for Comparison	1.1650
Error term used: FILAS*COLUMNAS*VAR, 6 DF			
There are 3 groups (A, B, etc.) in which the means are not significantly different from one another.			

Aparecen los valores tomados para t , el valor de S_D y el valor de la LSD (MDS), que separa las medias en tres grupos dentro de los cuales no hay diferencia significativas y entre los cuales si hay, como se había interpretado anteriormente.

CAPÍTULO 9

DISEÑO AUMENTADO

En ciertos programas de mejora genética de plantas, es necesario cribar una gran cantidad de selecciones nuevas de las que se dispone de pequeñas cantidades de semilla. Como parte de ese proceso de cribado, estas selecciones deben ser evaluadas por rendimiento en experimentos preliminares. Los métodos más clásicos para llevar a cabo estas evaluaciones consisten en sembrar estas nuevas selecciones en líneas o surcos individuales junto con surcos de uno o varios testigos, que son situados sistemáticamente dentro del experimento. Las nuevas selecciones son evaluadas subjetivamente, comparando su rendimiento con el del testigo más cercano. Sin embargo, como las nuevas selecciones no se han repetido, no es posible realizar un análisis estadístico de los rendimientos.

El diseño aumentado, desarrollado por Federer (1956) y descrito por Federer y Ragavarao (1975), trata de poner el análisis del rendimiento de estos experimentos preliminares dentro de una base estadística más convincente; su objetivo es, pues, la evaluación de un gran número de selecciones incluyendo análisis estadísticos.

El plan básico del diseño consiste en dividir el área experimental en una serie de bloques que, a su vez, contienen las unidades experimentales o parcelas. Tres o más testigos se asignan al azar a las unidades experimentales de cada bloque y en las unidades restantes se siembran las selecciones nuevas que se quieren probar, de manera que en cada bloque se dispone un grupo de las nuevas selecciones. Aunque los testigos están repetidos, las nuevas selecciones no lo están, pues se asignan al azar dentro de las unidades experimentales que quedan libres una vez que se han sorteado los testigos. El rendimiento de las selecciones nuevas es ajustado por las diferencias de bloques, que son estimadas por el rendimiento de los testigos que se repiten en cada uno de ellos.

Los bloques no tienen que contener necesariamente el mismo número de unidades experimentales, pero el ensayo es mucho más eficiente si el tamaño de los bloques es idéntico.

DISEÑO DEL EXPERIMENTO

El tamaño de los bloques viene determinado por el número de éstos, **b**, el número de testigos, **c**, y el número de selecciones nuevas, **v**. Si el tamaño de los bloques es constante (mismo número de unidades experimentales en cada bloque) se puede definir lo siguiente:

c = número de testigos por bloque

v = número de nuevas selecciones

b = número de bloques

n = v/b = número de selecciones nuevas por bloque

p = $c + n$ = número de unidades experimentales por bloque

N = $bc + v = b(c + n)$ = número total de unidades experimentales

El número total de bloques viene determinado por la necesidad de que el error experimental del ensayo tenga al menos 10 grados de libertad. Esto, a su vez, viene determinado por el número de variedades testigos c utilizadas en el experimento. En el análisis de la varianza para las variedades testigo, el error experimental tiene $(b-1)(c-1)$ grados de libertad, por lo tanto el número de bloques debe ser tal que se cumpla:

$$b > \frac{10}{c-1} + 1$$

que para cuatro testigos, por ejemplo, sería:

$$b > \frac{10}{4-1} + 1 = 4,33$$

El número mínimo de bloques será 5 y cada bloque contendrá 5 o más unidades experimentales, dependiendo del número de nuevas selecciones.

Así pues, cuando se establezca el experimento, los testigos han de ser asignados al azar a unidades experimentales dentro de cada bloque. En caso de que se asigne sistemáticamente un testigo a una unidad experimental determinada (por ejemplo la primera), se pierde muy poca eficiencia, siempre que los restantes testigos ($c-1$) se asignen al azar dentro de las unidades experimentales de cada bloque. Posteriormente, las v nuevas selecciones son asignadas al resto de las unidades experimentales en el ensayo.

Supongamos que se quieren evaluar 24 selecciones nuevas y utilizar tres testigos, con uno asignado sistemáticamente a la primera unidad experimental de cada bloque. Se requiere, entonces, un mínimo de:

$$b > \frac{10}{3-1} + 1 = 6 \text{ bloques}$$

Se tiene, pues, que:

$$c = 3: A, B, C$$

$$v = 24: 1, 2, \dots, 24$$

$$b = 6$$

$$n = v/b = 24/6 = 4$$

$$p = c + n = 3 + 4 = 7$$

$$N = bc + v = (6)(3) + 24 = 42$$

El experimento en el campo podría distribuirse de la siguiente manera:

BLOQUES					
I	II	III	IV	V	VI
A	A	A	A	A	A
13	17	21	2	B	19
8	9	C	B	12	C
B	C	15	10	5	20
C	24	B	C	16	B
18	B	1	3	6	4
7	11	23	22	C	14

ANÁLISIS

El primer paso a llevar a cabo en el análisis del experimento es la construcción de una tabla de dos entradas con los rendimientos de los testigos, los totales y las medias.

TESTIGOS	BLOQUES					TOTAL	MEDIAS
	1	2	3	b		
1	X ₁₁	X ₁₂	X ₁₃	X _{1b}	C ₁	\bar{x}_1
2	X ₂₁	X ₂₂	X ₂₃		X _{2b}	C ₂	\bar{x}_2
.
.
.
c	X _{c1}	X _{c2}	X _{c3}	X _{cb}	C _c	\bar{x}_c
TOTAL	B ₁	B ₂	B ₃		B _b	G	M

Donde:

x_{ij} = rendimiento del testigo *i* en el bloque *j*

$B_j = \sum_i x_{ij}$ suma del rendimiento de todos los testigos en el bloque *j*

$C_i = \sum_j x_{ij}$ suma de los rendimientos del testigo *i*

$G = \sum_j B_j = \sum_i C_i$ = suma total de todos los rendimientos de los testigos

$\bar{x}_i = C_i / b$ = rendimiento medio del testigo *i*

$M = \sum_i \bar{x}_i = G / b$ = suma de las medias de todos los testigos

El siguiente paso es el cálculo del factor de ajuste r_j para cada bloque, que se calcula de la siguiente forma:

$$r_j = \frac{B_j - M}{c}$$

A continuación se construye una tabla con los rendimientos reales y los ajustados por el efecto del bloque de las selecciones ensayadas:

SELECCION	RENDIMIENTO	
	OBSERVADO	AJUSTADO
1	y_{1j}	\hat{y}_{1j}
2	y_{2j}	\hat{y}_{2j}
·	·	·
·	·	·
·	·	·
v	y_{vj}	\hat{y}_{vj}

Donde:

y_{ij} = rendimiento de la selección i en el bloque j

$\hat{y}_{ij} = y_{ij} - r_j$ = rendimiento ajustado por el efecto del bloque de la selección i

Una estimación del error experimental que puede ser utilizado para calcular errores estándar y mínimas diferencias significativas (MDS) se puede obtener fácilmente realizando el análisis de varianza de los testigos, que sería:

Fuente de Variación	gl	SC	CM
Testigos	c-1	SCTR	
Bloques	b-1	SCB	
Error	(b-1)(c-1)	SCE	CME
Total	bc-1	SCT	

Se trata del análisis de varianza de un diseño en bloques al azar, cuyo desarrollo y cálculo se ha visto en el Capítulo 7. En este caso, los tratamientos serían las variedades testigos.

Ahora hay que calcular una serie de diferencias, específicas de este diseño, que son las siguientes junto con sus correspondientes varianzas:

- Diferencia entre las medias de dos testigos:

$$\text{Varianza: } s_c^2 = \frac{2CME}{b}$$

- Diferencia entre los rendimientos ajustados de dos selecciones en el mismo bloque:

$$\text{Varianza: } s_b^2 = 2CME$$

- Diferencia entre rendimientos ajustados de dos selecciones en distintos bloques:

$$\text{Varianza: } s_v^2 = \frac{2CME(c+1)}{c}$$

- Diferencia entre el rendimiento ajustado de una selección y la media de un testigo:

$$\text{Varianza: } s_{vc}^2 = \frac{CME(b+1)(c+1)}{bc}$$

- Diferencia entre rendimientos ajustados de dos selecciones cualquiera:

$$\text{Varianza: } s_a^2 = \frac{CME(2c+1)}{c}$$

Las mínimas diferencias significativas (MDS) pueden calcularse utilizando esas varianzas de la siguiente forma:

$$MDS = t_a \sqrt{s^2}$$

en donde:

t_a = el valor de t en la prueba de dos colas para $(b-1)(c-1)$ grados de libertad (grados de libertad del error).

s^2 = la varianza de la diferencia para la que se quiera calcular la mínima diferencia significativa.

El ejemplo práctico utilizado para ilustrar este análisis ha sido tomado de Petersen (1985). En el mismo se enuncia que un mejorador de cereales quiere realizar un ensayo preeliminar con 30 selecciones de trigo duro. Particularmente está interesado en comparar estas nue-

vas selecciones con tres variedades típicas de la zona: ‘Someto’ (ST), ‘Don Pedro’ (DP) y ‘Yavaros’(YV). Tiene semilla suficiente como para sembrar una sola línea de 2,5 m de cada una de las selecciones, así que decide utilizar un diseño aumentado para realizar el experimento. Puesto que va a utilizar 3 variedades testigo, requerirá al menos:

$$b > \frac{10}{c-1} + 1 = \frac{10}{3-1} + 1 = 6 \text{ bloques}$$

para tener suficientes grados de libertad para estimar el error experimental. Así pues, utilizando estos seis bloques, el diseño tendrá las siguientes características:

1. Numero de testigos: $c = 3$; ST, DP, YV
2. Número de selecciones nuevas: $v = 30$: 1,2,.....,30.
3. Número de bloques: $b = 6$
4. Número de nuevas selecciones por bloque: $n = v/b = 30/6 = 5$
5. Número de unidades experimentales por bloque: $p = c + n = 3 + 5 = 8$
6. Número total de unidades experimentales, $N = bc + v = (3)(6) + 30 = 48$

Supongamos que el diseño en el campo y los rendimientos de cada una de las unidades experimentales son las que figuran en la Tabla 9.1.

Tabla 9.1.- Rendimiento (kg/ha) de diferentes variedades y nuevas selecciones de trigo duro en cada una de las parcelas elementales de los seis bloques que forman un diseño aumentado.

Bloque	I		II		III	
	Selección	Rendimiento	Selección	Rendimiento	Selección	Rendimiento
	14	2405	DP	3023	18	2603
	26	2855	4	3018	ST	2260
	DP	2592	15	2477	27	2857
	17	2572	30	2955	DP	2918
	YV	2608	YV	2477	25	2825
	ST	2972	3	3055	28	1903
	22	2705	ST	3122	5	2065
	13	2391	24	2783	YV	3107
Bloque	IV		V		VI	
	9	2268	2	1055	29	2915
	6	2148	21	1688	7	3265
	DP	2940	ST	1315	DP	3483
	YV	2850	YV	1625	1	3013
	20	2670	DP	1398	YV	3400
	11	3380	10	1293	12	2385
	23	2770	8	1253	ST	3538
	ST	3348	16	1495	19	3643

Para iniciar el análisis se construye una tabla con los rendimientos de los testigos, sus valores medios, los totales por bloque y el total general. Con los datos del ejemplo, se construye la Tabla 9.2.

Tabla 9.2.- Rendimientos, totales y medias de las variedades testigo en cada bloque.

Variedad	BLOQUES						TOTAL	MEDIA
	I	II	III	IV	V	VI		
ST	2972	3122	2260	3348	1315	3538	16555	2759,17
DP	2592	3023	2918	2940	1398	3483	16354	2725,67
YV	2608	2477	3107	2850	1625	3400	16067	2677,83
TOTAL	8172	8622	8285	9138	4338	10421	48976	2720,83
								8162,67

Los factores de corrección r_j para cada bloque serán, con $c = 3$ y $M = 8162,67$, los recogidos en la Tabla 9.3.

Tabla 9.3.- Factores de corrección para cada bloque del experimento.

Bloque	I	II	III	IV	V	VI	Σ
r_j	3,11	153,11	40,78	325,11	-1274,89	752,78	0,00

A continuación se confecciona una tabla con los valores observados y ajustados de las nuevas selecciones (Tabla 9.4).

Tabla 9.4.- Valores observados y valores ajustados de las nuevas selecciones.

SELECCIÓN	BLOQUE	VALOR OBSERVADO	FACTOR CORRECCIÓN BLOQUE	VALOR AJUSTADO
		y_{ij}	r_j	\hat{y}_{ij}
1	6	3013	752,78	2260,22
2	5	1055	-1274,89	2329,89
3	2	3055	153,11	2901,89
4	2	3018	153,11	2864,89
5	3	2065	40,78	2024,22
6	4	2148	325,11	1822,89
7	6	3265	752,78	2512,22
8	5	1253	-1274,89	2527,89

SELECCIÓN	BLOQUE	VALOR OBSERVADO y_{ij}	FACTOR CORRECCIÓN BLOQUE r_j	VALOR AJUSTADO \hat{y}_{ij}
9	4	2268	325,11	1942,89
10	5	1293	-1274,89	2567,89
11	4	3380	325,11	3054,89
12	6	2385	752,78	1632,22
13	1	2391	3,11	2387,89
14	1	2405	3,11	2401,89
15	2	2477	153,11	2323,89
16	5	1495	-1274,89	2769,89
17	1	2572	3,11	2568,89
18	6	2603	752,78	1850,22
19	3	3643	40,78	3602,22
20	4	2670	325,11	2344,89
21	5	1688	-1274,89	2962,89
22	1	2705	3,11	2701,89
23	4	2770	325,11	2444,89
24	2	2783	153,11	2629,89
25	3	2825	40,78	2784,22
26	1	2855	3,11	2851,89
27	3	2857	40,78	2816,22
28	3	1903	40,78	1862,22
29	6	2915	752,78	2162,22
30	2	2915	153,11	2761,89

El análisis de la varianza de los testigos se realiza de la forma descrita en el Capítulo 7. Si utilizamos el programa *Statistix (SX)*, tenemos:

Source	DF	SS	MS	F	P
Bloque	5	6968486	1393697		
Testigo	2	20051	10025	0.11	0.8969
Error	10	911027	91103		
Total	17	7899564			

Grand Mean 2720.9 CV 11.09

Tukey's 1 Degree of Freedom Test for Nonadditivity					
Source	DF	SS	MS	F	P
Nonadditivity	1	115840	115840	1.31	0.2817
Remainder	9	795187	88354		

Relative Efficiency, RCB 4.96

Means of Rendimien for Testigo

Testigo	Mean
1	2759.2
2	2725.7
3	2677.8
Observations per Mean	6
Standard Error of a Mean	123.22
Std Error (Diff of 2 Means)	174.26

Realmente de este análisis lo único que interesa es el valor del cuadrado medio del error (CME = 91103) para utilizarlo en la estimación de las varianzas que van a formar parte del cálculo de las mínimas diferencias significativas (MDS) para cada una de las diferencias establecidas. En el ejemplo puede interesar fundamentalmente las siguientes:

1. Diferencia entre el rendimiento ajustado de las selecciones y un testigo. La varianza de esta diferencia será:

$$s_{vc}^2 = \frac{CME(b+1)(c+1)}{bc} = \frac{(91103)(7)(4)}{(6)(3)} = 141716$$

y la mínima diferencia significativa para esta comparación será:

$$MDS_{(0,05)} = t_{(0,05)} \sqrt{s_{vc}^2} = 2,228 \sqrt{141716} = 838,7 \text{ kg.}$$

2. Diferencia entre los rendimientos ajustados de dos selecciones cualesquiera. La varianza será:

$$s_a^2 = \frac{CME(2c+1)}{c} = \frac{91103[(2)(3)+1]}{3} = 212574$$

y la mínima diferencia significativa para esta comparación será:

$$MDS_{(0,05)} = t_{(0,05)} \sqrt{s_a^2} = 2,228 \sqrt{212574} = 1027,2 \text{ kg.}$$

Comparando los rendimientos ajustados de las selecciones (Tabla 9.4) con los rendimientos medios de los testigos (Tabla 9.2) y conociendo las MDS calculadas, pueden obtenerse conclusiones pertinentes para realizar las selecciones y avanzar en el programa de mejora. Por ejemplo, aunque hay varias selecciones que superan el rendimiento medio de la variedad testigo más productiva, 'Someto', ninguna de ellas mostró aumentos significativos del rendimiento.

CAPÍTULO 10

EXPERIMENTOS FACTORIALES

Cuando hay que estudiar de forma simultánea los efectos que tienen varios factores sobre una respuesta, se utilizan experimentos factoriales en los que todos los niveles de un factor se combinan con todos los niveles de cualquier otro factor para formar los tratamientos. Así, en un experimento factorial de tres factores A, B y C, con a, b y c niveles respectivamente, el número de tratamientos es $a \times b \times c$.

El análisis de varianza de este tipo de experimentos recibe el nombre de ANOVA multifactorial, o simplemente ANOVA factorial. En él se analizan no sólo los efectos individuales de los factores (*efectos principales*), sino también el efecto causado por sus *interacciones*. Cuando el efecto de un factor sobre la respuesta no depende de los niveles del otro factor, se dice que ambos factores son independientes. En cambio, si el efecto de un factor varía para los distintos niveles del otro factor, entonces los dos factores no son independientes y se dice que interaccionan entre sí.

La posibilidad de establecer inferencias sobre el efecto de un factor para los diferentes niveles de otros factores, y sobre las relaciones entre ellos, es lo que diferencia principalmente a un experimento factorial de varios experimentos de factor único en los que se incluyan los diferentes tratamientos analizados. Comparativamente, el experimento factorial presenta una mayor precisión y, por tanto, un considerable ahorro de tiempo y material dedicado a los experimentos.

Los experimentos factoriales son ampliamente utilizados en investigaciones agronómicas cuando se sabe poco de los niveles más adecuados de los factores, o ni siquiera cuáles son importantes, y cuando se desea obtener recomendaciones que deben aplicarse en una gran variedad de condiciones. Asimismo, muchas veces el investigador está interesado en la interacción entre factores, más que en el efecto individual de éstos. La presencia de interacción significativa hace que los efectos principales de los factores que interaccionan resulten irrelevantes, por lo que debemos centrar nuestra atención en el análisis de la interacción.

La limitación principal de los experimentos factoriales es que al aumentar los factores se incrementa el número de interacciones y su complejidad, lo que dificulta la interpretación de los resultados. Si los diferentes factores interaccionan de una forma confusa, ello es indicativo de que los efectos son complejos, pero no implica que la experimentación sea defectuosa. En estos casos, será necesario un estudio profundo de los resultados y una experimentación más amplia para comprender enteramente los hechos.

DISEÑOS EXPERIMENTALES Y MODELOS DE ANÁLISIS DE VARIANZA

Los experimentos factoriales admiten los diferentes diseños que se han descrito para un único factor: completamente aleatorio, bloques completos al azar, cuadrado latino, etc. Sin embargo, cuando el número de factores y niveles de éstos aumentan considerablemente, algunos diseños, como el cuadrado latino, necesitan una cantidad de repeticiones que lo hacen impracticable. Otras veces, resulta difícil lograr repeticiones homogéneas en un diseño en bloques completos al azar, por lo que el error experimental tiende a aumentar. Por ello, cuando el número de tratamientos es elevado, se utilizan otros diseños, como el de bloques incompletos que sacrifica las interacciones de alto orden, o el de parcelas divididas o subdivididas que sacrifican un factor o un grupo de factores y sus interacciones, o bien, otros diseños en los que la pérdida de precisión se distribuye uniformemente entre los factores y sus interacciones, como látices balanceados, bloques incompletos balanceados y látices cuadrados (Cochran y Cox, 1978).

Para explicar el fundamento del ANOVA factorial se va a considerar un caso sencillo: dos factores de efectos fijos A y B, con a y b niveles respectivamente, en un diseño en bloques al azar con r bloques. El modelo lineal para este caso es:

$$Y_{ijr} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \rho_r + \varepsilon_{ijr}$$

que puede expresarse en términos de desviaciones como:

$$Y_{ijr} - \mu = \alpha_i + \beta_j + (\alpha\beta)_{ij} + \rho_r + \varepsilon_{ijr}$$

estableciendo que la desviación de una observación (Y_{ijr}) respecto a la media global (μ), está formada por cinco componentes: las desviaciones causadas por el efecto de los factores principales (α_i y β_j), por la interacción entre ellos ($(\alpha\beta)_{ij}$), por los bloques (ρ_r) y por el error aleatorio (ε_{ijr}).

Las hipótesis nulas de interés en este caso son tres:

1. $H_0: (\alpha\beta)_{ij} = 0$ para toda i y j
2. $H_0: \alpha_i = 0$ para toda i
3. $H_0: \beta_j = 0$ para toda j

teniendo en cuenta que si la primera hipótesis se rechaza, es decir, la interacción entre A y B resultara significativa, las otras hipótesis carecen de interés. En el caso de un modelo II (efectos aleatorios), estas hipótesis vendrían formuladas como $\sigma_{\alpha\beta} = 0$, $\sigma_{\alpha} = 0$, y $\sigma_{\beta} = 0$, respectivamente.

La ecuación fundamental de descomposición de la suma de cuadrados (SC) para el ANOVA en este caso sería:

$$SCT = SCA + SCB + SCAB + SCR + SCE$$

que establece que la suma de cuadrados total (*SCT*) está compuesta de cinco sumandos: la suma de cuadrados del factor (*SCA*), la del factor (*SCB*), la de la interacción (*SCAB*), la de los bloques (*SCR*) y la del error aleatorio (*SCE*).

Igualmente, la descomposición del número de grados de libertad (*gl*) sería:

$$gl(SCT) = gl(SCA) + gl(SCB) + gl(SCAB) + gl(SCR) + gl(SCE)$$

o bien:

$$(rab - 1) = (a - 1) + (b - 1) + (a - 1) \times (b - 1) + (n - 1) + (n - 1) \times (ab - 1)$$

Los estadísticos de prueba para las tres hipótesis nulas (interacción y dos efectos principales) son los cocientes entre los cuadrados medios (*CM*) correspondientes y el cuadrado medio del error (*CME*), que tienen una distribución *F*. Los valores esperados de los cuadrados medios para el caso de efectos fijos son:

$$E(CME) = \sigma^2$$

$$E(CMA) = \sigma^2 + nb \frac{\sum \alpha_i^2}{a - 1}$$

$$E(CMB) = \sigma^2 + na \frac{\sum \beta_i^2}{b - 1}$$

$$E(CMAB) = \sigma^2 + n \frac{\sum \sum (\alpha\beta)_{ij}^2}{(a - 1) \times (b - 1)}$$

La tabla del ANOVA para este caso sería:

Fuente de Variación	gl	SC	CM	F
Bloques	r-1	SCR	$\frac{SCR}{(r-1)}$	
Factor A	a-1	SCA	$\frac{SCA}{(a-1)}$	CMA/CME
Factor B	b-1	SCB	$\frac{SCB}{(b-1)}$	CMB/CME
AxB	(a-1)(b-1)	SCAB	$\frac{SCAB}{(a-1)(b-1)}$	CMAB/CME
Error	(r-1)(ab-1)	SCE	$\frac{SCE}{(r-1)(ab-1)}$	
Total	rab-1	SCT		

$$SCR = \frac{1}{ab} \sum_r Y_{..r}^2 - \frac{Y_{...}^2}{rab}$$

$$SCA = \frac{1}{rb} \sum_i Y_{i..}^2 - \frac{Y_{...}^2}{rab}$$

$$SCB = \frac{1}{ra} \sum_j Y_{.j.}^2 - \frac{Y_{...}^2}{rab}$$

$$SCAB = \frac{1}{r} \sum_i \sum_j Y_{ij.}^2 - \frac{1}{ra} \sum_i Y_{i..}^2 - \frac{1}{rb} \sum_j Y_{.j.}^2 + \frac{Y_{...}^2}{rab}$$

$$SCE = \sum_i \sum_j \sum_r (Y_{ijr} - \bar{Y}_{ij.})^2$$

$$SCT = \sum_i \sum_j \sum_r Y_{ijr}^2 - \frac{Y_{...}^2}{rab}$$

Si los cocientes CMA/CME , CMB/CME y $CMAB/CME$ dan un valor inferior o próximo a 1, obtendríamos un valor P no significativo ($P > 0.05$) y aceptaríamos las tres hipótesis nulas de que $\alpha_i = 0$, $\beta_j = 0$ y $(\alpha\beta)_{ij} = 0$. Por tanto, los tres cocientes serían estimadores de la varianza residual común σ^2 . En cambio, si alguno de estos cocientes es significativamente mayor que 1 ($P < 0.05$), rechazaríamos la hipótesis nula para ese caso, concluyendo que el efecto de ese factor principal o de la interacción es estadísticamente significativo. Con tres cocientes (CMA/CME , CMB/CME , $CMAB/CME$) y dos posibilidades de significación del efecto (significativo, no significativo) tendríamos $2^3 = 8$ posibles resultados. No obstante, cuando la interacción es significativa, la significación de los efectos principales de los factores implicados carece de interés ya que éstos no son independientes. En este caso, la interpretación de los resultados se debe basar exclusivamente en el análisis de la interacción. En las Figuras 10.1 y 10.2 se ilustran los ocho posibles resultados de un experimento factorial con dos factores.

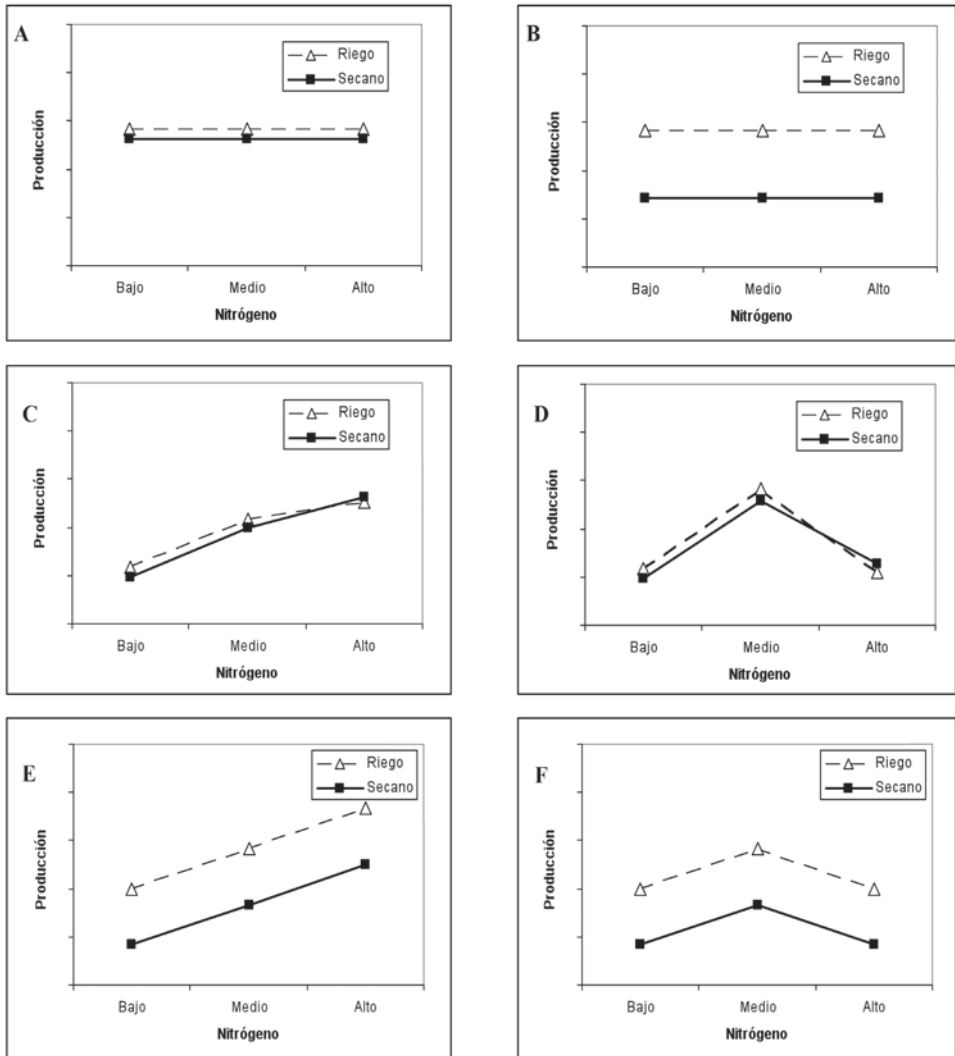


Fig. 10.1.- Ilustración del efecto de dos factores (riego y nitrógeno) en un ANOVA factorial con interacción no significativa. A) Riego y Nitrógeno no significativos. B) Riego significativo, Nitrógeno no. C y D) Nitrógeno significativo, riego no. E y F) Riego y Nitrógeno significativos

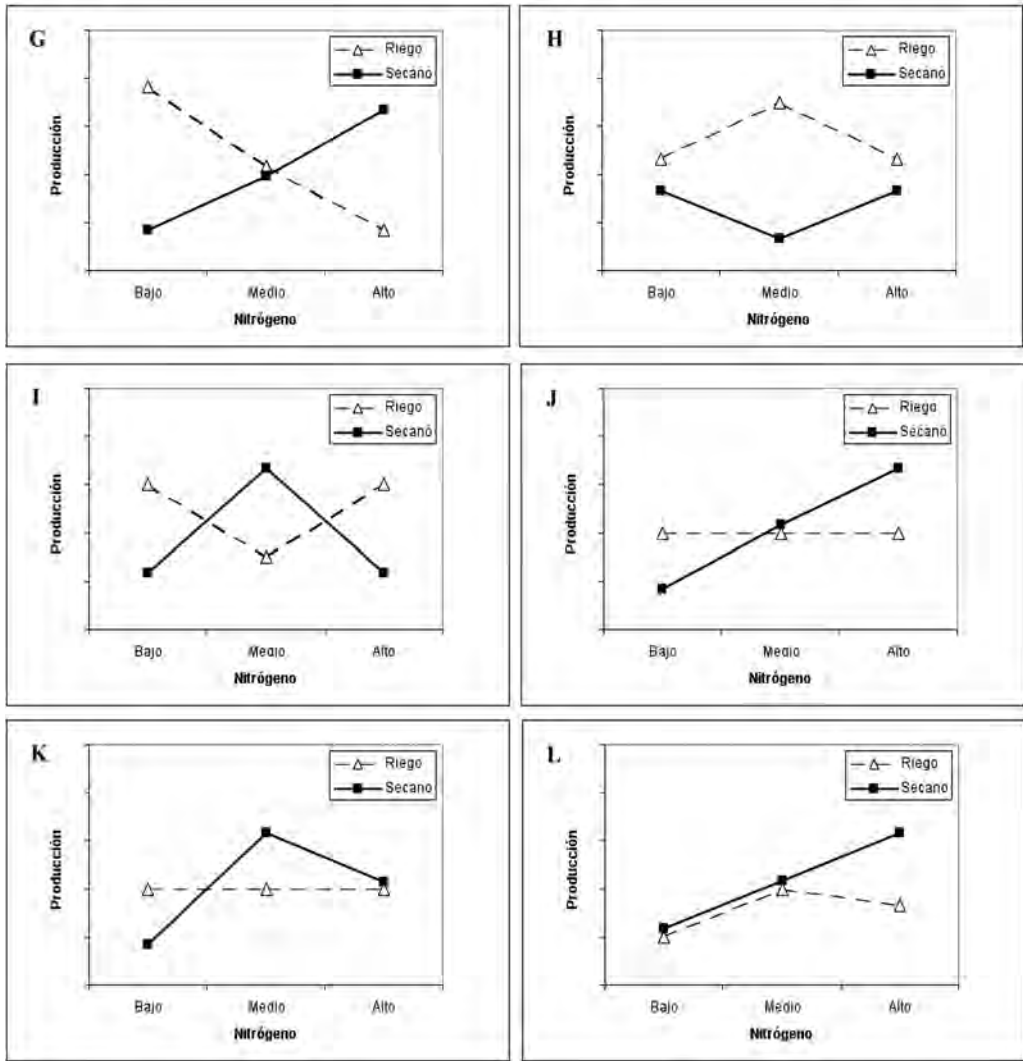


Fig. 10.2.- Ilustración del efecto de dos factores (riego y nitrógeno) en un ANOVA factorial con interacción significativa. G) Riego y Nitrógeno no significativos. H, I) Riego significativo, Nitrógeno no. J, K) Nitrógeno significativo, Riego no. L) Riego y Nitrógeno significativos.

El supuesto analizado corresponde a un ANOVA factorial con dos factores de efectos fijos (modelo I). Cuando los dos factores son de efectos aleatorios (modelo II), o bien, si uno de ellos es de efectos fijos y el otro es de efectos aleatorios (modelo III o mixto), los valores esperados de los cuadrados medios varían y también varía el estadístico de prueba F . En la Tabla 10.1 se indican las esperanzas de los cuadrados medios para estos modelos.

Tabla 10.1. Esperanzas de los cuadrados medios (ECM) para un ANOVA factorial con dos factores: modelos de efectos aleatorios o de efectos mixtos*

Efectos aleatorios (A y B aleatorios)			Efectos mixtos (A fijo, B aleatorio)	
Fuente	ECM	Estadístico F	ECM	Estadístico F
A	$\sigma^2 + n\sigma_{\alpha\beta}^2 + nb\sigma_\alpha^2$	CMA/CMAB	$\sigma^2 + n\sigma_{\alpha\beta}^2 + nb \frac{\sum \alpha_i^2}{(a-1)}$	CMA/CMAB
B	$\sigma^2 + n\sigma_{\alpha\beta}^2 + na\sigma_\beta^2$	CMB/CMAB	$\sigma^2 + na\sigma_\beta^2$	CMB/CME
AB	$\sigma^2 + n\sigma_{\alpha\beta}^2$	CMAB/CME	$\sigma^2 + n\sigma_{\alpha\beta}^2$	CMAB/CME
Error	σ^2		σ^2	

*Fuente: Canavos, 1988.

Conviene destacar que en el modelo de efectos fijos el CME es el denominador común de los cocientes de cuadrados medios para todos los efectos principales e interacciones, pero esto no ocurre así para los otros modelos (Tabla 10.1). Por ello, la elección de los niveles del factor en la fase de diseño de un experimento factorial es muy importante, ya que tiene una influencia directa en el análisis (Canavos, 1988).

REALIZACIÓN DEL ANÁLISIS DE VARIANZA FACTORIAL

La Tabla 10.2 recoge los resultados obtenidos en un experimento sobre el efecto del riego y de la fertilización potásica en el crecimiento de plantones de olivo cultivados en macetas. Se ensayaron dos niveles de riego (Bajo: riego cada 5 días; Alto: riego diario) y dos dosis de fertilizante potásico (K Bajo, K Alto). La unidad experimental fue una maceta con una planta y se dispusieron 6 macetas de cada combinación experimental en un diseño en bloques completos al azar. Los datos recogen el peso total (g de materia seca) de los brotes que crecieron durante el experimento. Tenemos, pues, un experimento factorial de dos factores, dos niveles de cada factor y 6 repeticiones (plantones) en bloques al azar.

Tabla 10.2. Peso de los brotes (g de materia seca) de olivos sometidos a dos tratamientos de riego y de fertilización potásica (K).

Bloques (r)	Riego Bajo		Riego Alto		Total
	K Bajo	K Alto	K Bajo	K Alto	
1	11	14	17	21	63
2	13	13	19	19	64
3	12	10	20	22	64
4	14	15	17	21	64
5	15	14	18	23	70
6	13	11	16	20	60
$Y_{ij\cdot} = \sum_{r=1}^6 Y_{ijr}$	78	77	107	126	388
$Y_{ij\cdot}^2 = \sum_{r=1}^6 Y_{ijr}^2$	1024	1007	1919	2656	6606
$\bar{Y}_{ij\cdot} = \sum_{r=1}^6 Y_{ijr} / 6$	13,00	12,83	17,83	21,00	16,17

El primer paso consiste en calcular las sumas de cuadrados. Utilizando la fórmula abreviada calculamos el factor de corrección (C) y las diferentes sumas de cuadrados: Total (SCT), Factor A (SCA), Factor B (SCB), Interacción A×B (SCAB), Bloques (SCR) y Error (SCE).

El factor de corrección es: $C = \frac{Y_{ijr}^2}{rab} = \frac{388^2}{6 \times 2 \times 2} = \frac{150544}{24} = 6272,667$

$$SCT = \sum_i \sum_j \sum_r Y_{ijr}^2 - C = (11^2 + 13^2 + \dots + 23^2 + 20^2) - C = 6606 - 6272,667 = 333,333$$

$$SCA = \frac{1}{rb} \sum_r Y_{i\cdot}^2 - C = \frac{(78 + 77)^2}{12} + \frac{(107 + 126)^2}{12} - 6272,667 = \frac{155^2}{12} + \frac{233^2}{12} - 6272,667 = 253,5$$

$$SCB = \frac{1}{ra} \sum_r Y_{\cdot j}^2 - C = \frac{(78 + 107)^2}{12} + \frac{(77 + 126)^2}{12} - 6272,667 = \frac{185^2}{12} + \frac{203^2}{12} - 6272,667 = 13,5$$

$$SCR = \frac{1}{ab} \sum_r Y_{\cdot\cdot r}^2 - C = \left(\frac{63^2}{4} + \frac{64^2}{4} + \dots + \frac{70^2}{4} + \frac{60^2}{4} \right) - 6272,667 = 14,833$$

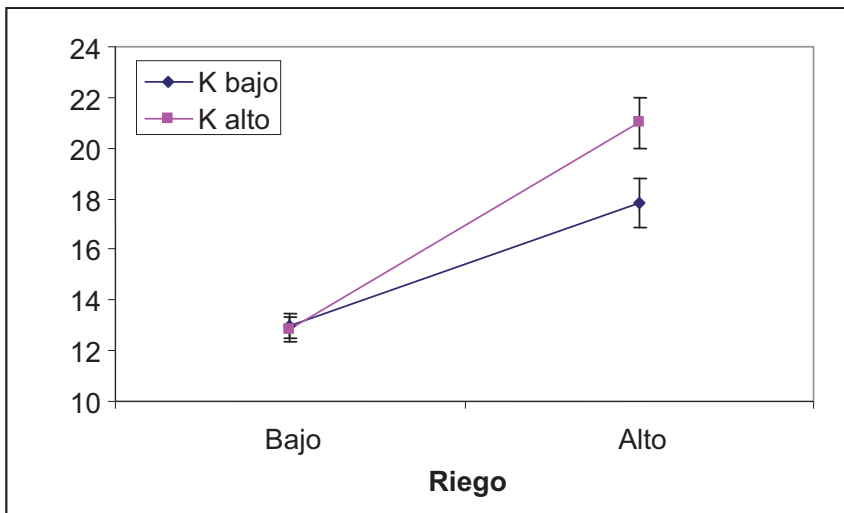
$$SCAB = \frac{1}{r} \sum_i \sum_j Y_{ij\cdot}^2 - \frac{1}{rb} \sum_r Y_{i\cdot}^2 - \frac{1}{ra} \sum_j Y_{\cdot j}^2 + C = \left(\frac{78^2}{6} + \frac{77^2}{6} + \frac{107^2}{6} + \frac{126^2}{6} \right) - \left(\frac{155^2}{12} + \frac{233^2}{12} \right) - \left(\frac{185^2}{12} + \frac{203^2}{12} \right) + 6272,667 = 16,667$$

$$SCE = SCT - SCA - SCB - SCAB - SCR = 34,833$$

La Tabla ANOVA para este ejemplo sería:

Fuente de Variación	gl	SC	CM	F	P
Bloques	5	14,833	2,967	1,28	0,3242
Riego (A)	1	253,500	253,500	109,16	0,0000
Potasio(B)	1	13,500	13,500	5,81	0,0292
AxB	1	16,667	16,670	7,18	0,0172
Error	15	34,833	2,322		
Total	23	333,333			

Los resultados del ANOVA indican que no hay efecto significativo de los bloques ($P > 0.05$), pero sí del riego, del potasio y de su interacción ($P < 0.05$). Al ser la interacción significativa, los efectos de los factores individuales carecen de interés y centramos nuestra atención en dicha interacción. Como sólo existen dos niveles de cada factor, el análisis de la interacción es sencillo. En el gráfico adjunto se ilustra la interacción:



Como se deduce del ANOVA y se aprecia en el gráfico, el crecimiento de los plántones de olivo aumentó con la mayor aportación de agua de riego, pero la respuesta a la fertilización potásica dependió de la dosis de riego aplicada. Así, para aportaciones bajas de riego no hubo diferencias de crecimiento entre los plántones fertilizados con dosis bajas y altas de potasio. En cambio, para aportaciones mayores de riego, los plántones fertilizados con la dosis alta de potasio crecieron más que los que recibieron una dosis más baja.

El valor de la Mínima Diferencia Significativa (MDS o LSD) sería:

$$MDS = t_{0,05,15} \times \sqrt{\frac{2 \times CME}{r}} = 2,131 \times \sqrt{\frac{2 \times 2,322}{6}} = 1,875$$

CÁLCULO CON EL PROGRAMA *Statistix (SX)*

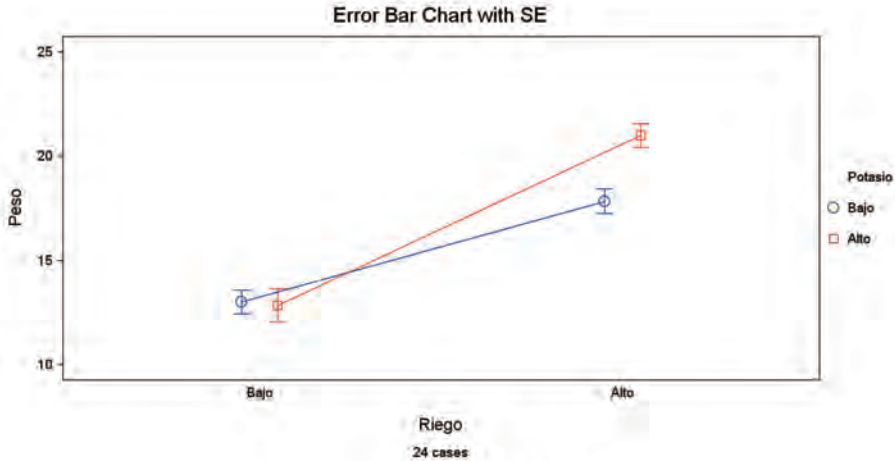
Se introducen los datos del experimento denominando a los dos factores a estudiar o variables categóricas como Riego (con dos niveles 2-Alto y 1-Bajo) y Potasio (con dos niveles 2-Alto y 1-Bajo), a los bloques como Rep (con seis valores) y a la variable dependiente como Peso (peso seco de los brotes en gramos). La matriz de los datos sería:

CASE	Riego	Potasio	Rep	Peso
1	1	1	1	11
2	1	1	2	13
3	1	1	3	12
4	1	1	4	14
5	1	1	5	15
6	1	1	6	13
7	1	2	1	14
8	1	2	2	13
9	1	2	3	10
10	1	2	4	15
11	1	2	5	14
12	1	2	6	11
13	2	1	1	17
14	2	1	2	19
15	2	1	3	20
16	2	1	4	17
17	2	1	5	18
18	2	1	6	16
19	2	2	1	21
20	2	2	2	19
21	2	2	3	22
22	2	2	4	21
23	2	2	5	23
24	2	2	6	20

Una vez introducidos los datos es recomendable su visualización. Para ello, se elige la siguiente secuencia:

Statistics>Summary Statistics>Error Bar Chart

También se podría elegir la opción “*Box and Whisker Plot*” en *Summary Statistics* para comprobar la variabilidad de los datos y ver si existen datos raros. En la pantalla resultante se introduce la variable dependiente y las dos variables categóricas, se seleccionan las opciones línea (“line”) y error estándar (“Std Error”) y se pulsa OK. Se obtiene el gráfico siguiente:



Se obtiene así una visualización bastante clara de los datos: las varianzas parecen homogéneas y también parece que existe una clara interacción entre los dos factores riego y potasio. Si se hubiese elegido la opción “*Box and Whisker Plot*” se apreciaría que las varianzas parecen homogéneas y que no existen “outliers”. Se realiza al análisis eligiendo la siguiente secuencia:

Satistics>Linear Models>Analysis of Variance>Factorial Design...

En la pantalla que se abre se introduce el nombre de la variable dependiente, el factor bloque, las dos variables independientes y su interacción, y se pulsa OK. Se obtiene así la Tabla ANOVA:

Analysis of Variance Table for Peso					
Source	DF	SS	MS	F	P
Rep	5	14.833	2.967		
Riego	1	253.500	253.500	109.16	0.0000
Potasio	1	13.500	13.500	5.81	0.0292
Riego*Potasio	1	16.667	16.667	7.18	0.0172
Error	15	34.833	2.322		
Total	23	333.333			
Grand Mean	16.167	CV 9.43			

Los resultados del análisis coinciden totalmente con el cálculo manual realizado anteriormente. Se puede comprobar si se cumplen algunos preceptos del ANOVA. En este análisis no se incluye ningún test de homogeneidad de varianzas. Para disponer de ellos, habría que realizar el ANOVA como si fuera unifactorial en un diseño completamente aleatorio. Los resultados para el riego y potasio por separado, serían:

Completely Randomized AOV for Peso (Riego)

Source	DF	SS	MS	F	P
Riego	1	253.500	253.500	69.9	0.0000
Error	22	79.833	3.629		
Total	23	333.333			

Grand Mean 16.167 CV 11.78

	Chi-Sq	DF	P
Bartlett's Test of Equal Variances	0.83	1	0.3620
Cochran's Q	0.6378		
Largest Var / Smallest Var	1.7608		

Component of variance for between groups 20.8226
Effective cell size 12.0

Riego Mean

Bajo	12.917
Alto	19.417
Observations per Mean	12
Standard Error of a Mean	0.5499
Std Error (Diff of 2 Means)	0.7777

Completely Randomized AOV for Peso (Potasio)

Source	DF	SS	MS	F	P
Potasio	1	13.500	13.5000	0.93	0.3457
Error	22	319.833	14.5379		
Total	23	333.333			

Grand Mean 16.167 CV 23.58

	Chi-Sq	DF	P
Bartlett's Test of Equal Variances	2.17	1	0.1410
Cochran's Q	0.7157		
Largest Var / Smallest Var	2.5179		

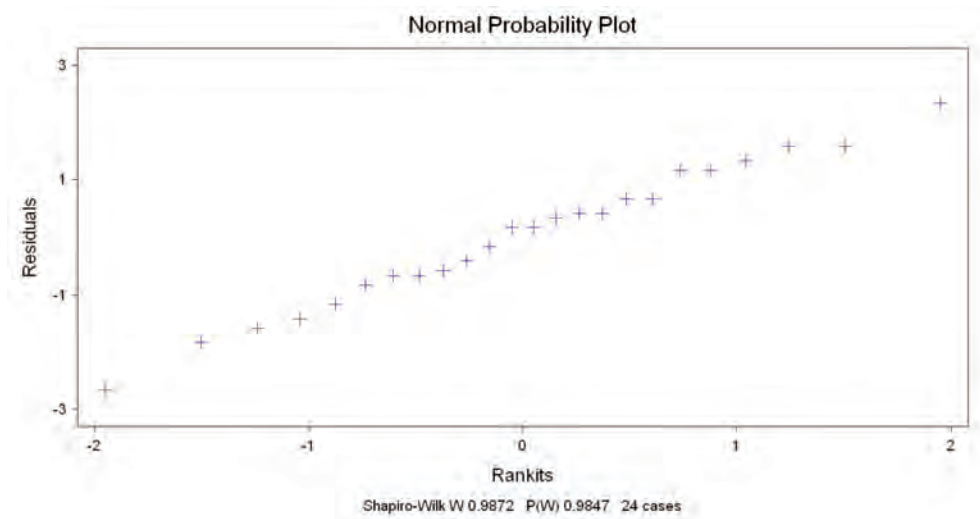
Component of variance for between groups -0.08649
Effective cell size 12.0

Potasio Mean

Bajo	15.417
Alto	16.917
Observations per Mean	12
Standard Error of a Mean	1.1007
Std Error (Diff of 2 Means)	1.5566

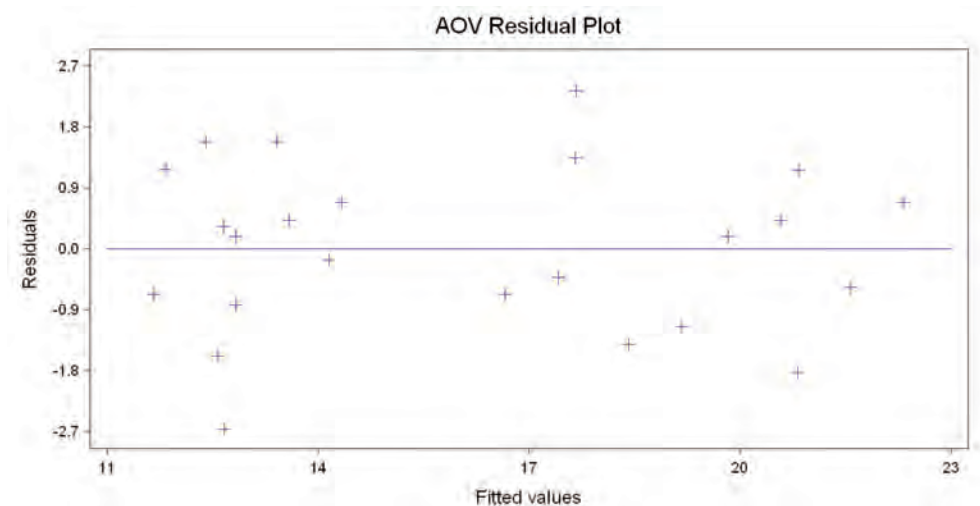
El *test de Bartlett*, el *test de Cochran* y la razón *Varianza mayor/Varianza menor* confirman que las varianzas son homogéneas. La normalidad de los datos se puede comprobar eligiendo la siguiente secuencia en la pantalla de resultados:

Results>Plots>Normal Probability Plot



El *test Shapiro-Wilk* indica que los datos se ajustan bien a una normal ($P > 0.05$). Asimismo, se puede ver la distribución que presentan los residuos en la misma pantalla de resultados siguiendo la secuencia:

Results>Plots>Resids By Fitted Values



La distribución observada en el gráfico es un ejemplo de distribución aleatoria sin ningún patrón específico, lo que confirma la independencia de los errores.

Finalmente, en la misma pantalla de resultados, se puede realizar la comparación de medias, teniendo en cuenta que la interacción es significativa. Para ello, se sigue la secuencia:

Results>Multiple Comparisons>All-pairwise Comparisons

En la pantalla resultante se seleccionan la interacción Riego*Potasio, el método *LSD*, $\alpha = 0.05$ y la opción grupos homogéneos y se obtiene:

LSD All-Pairwise Comparisons Test of Peso for Riego*Potasio

Riego	Potasio	Mean	Homogeneous Groups
Alto	Alto	21.000	A
Alto	Bajo	17.833	B
Bajo	Bajo	13.000	C
Bajo	Alto	12.833	C

Alpha 0.05 Standard Error for Comparison 0.8798
Critical T Value 2.131 Critical Value for Comparison 1.8753
Error term used: Rep*Riego*Potasio, 15 DF
There are 3 groups (A, B, etc.) in which the means are not significantly different from one another.

Como cabía esperar, se llega a la misma conclusión que con el cálculo manual. Para el nivel bajo de riego no existen diferencias significativas entre los tratamientos de potasio, mientras que para el nivel alto de riego sí existen diferencias significativas entre los tratamientos de potasio, siendo mayor el peso de los brotes de los plantones fertilizados con la dosis más alta de potasio.

El bajo número de factores y de los niveles de cada uno de ellos ha facilitado notablemente el análisis de la interacción en este ejemplo, utilizando la opción más sencilla de comparaciones múltiples. No obstante, cuando el número de factores y sus niveles son mayores, el análisis de la interacción es complejo y, generalmente, requiere realizar alguna partición del experimento y/o restringir el número de comparaciones utilizando otros tests de comparación de medias como los contrastes ortogonales.

CAPÍTULO 11

DISEÑO EN PARCELAS DIVIDIDAS

Los diseños en parcelas divididas o “split-plot” constituyen un caso particular de los experimentos factoriales, en el que los factores que intervienen no se combinan aleatoriamente entre sí, sino que están subordinados unos a otros. Las unidades o parcelas principales, a las que se les aplican los niveles de uno o más factores principales, se dividen en subunidades o subparcelas, a las cuales se les aplican los niveles de uno o más factores subordinados o secundarios. De este modo, cada unidad principal completa se convierte en un bloque para los tratamientos de las subunidades.

Para comprenderlo mejor, consideremos un experimento sencillo en el que intervienen dos factores: uno asignado a la parcela principal (A) con dos niveles (a_1 y a_2) y otro a la subparcela (B) con tres niveles (b_1 , b_2 y b_3). Si el diseño experimental fuese en tres bloques completos al azar, tendríamos el siguiente esquema del experimento:

Bloque I		Bloque II		Bloque III	
a_1	a_2	a_2	a_1	a_1	a_2
b_2	b_1	b_3	b_1	b_2	b_3
b_1	b_3	b_2	b_2	b_3	b_1
b_3	b_2	b_1	b_3	b_1	b_2

Hay que destacar que la aleatorización para distribuir los niveles de los factores se realiza en dos etapas: primero se sortean los niveles del factor A en cada bloque y después se hace lo propio con el factor subordinado (B) para cada nivel del factor A en cada bloque.

Este diseño sacrifica la precisión en la estimación de los efectos del factor principal, aunque suele incrementar la precisión para comparar los efectos del factor subordinado y los de la interacción de éste con el principal. Ello se debe a que al haber dos tamaños de parcelas existen dos errores experimentales, uno para cada tamaño de parcela. Generalmente, el error de las subparcelas es menor que el de las parcelas principales. Por ello, es muy importante la asignación de los factores como principal y subordinado.

Para la elección de la jerarquía de los factores en los experimentos agronómicos hay que considerar varios aspectos, como el grado relativo de precisión que se desee, el tamaño relativo de las parcelas y las posibilidades de aplicación práctica de los tratamientos o niveles de los factores. Como se ha indicado, el factor del que se desee estimar sus efectos con más precisión se deberá asignar a las subparcelas. Por el contrario, el factor que tenga un efecto mayor y más fácilmente detectable debería asignarse a las parcelas principales. Asimismo,

si un factor requiere mayor cantidad de material experimental o mayor tamaño de parcela para facilitar su manejo o aplicación, dicho factor debería asignarse a la parcela principal.

El diseño en parcelas divididas está especialmente indicado cuando se desee incorporar en un experimento un factor adicional para aumentar su alcance. Por ejemplo, supongamos que el objetivo principal de un experimento es comparar la eficacia de varios fungicidas contra una enfermedad. Para aumentar el alcance del experimento, se incluyen varios cultivares que difieren en su resistencia a la enfermedad. En este caso, los cultivares se organizarían en parcelas principales y los tratamientos fungicidas como subparcelas. Igualmente, este diseño es apropiado para incorporar un nuevo factor a un experimento ya establecido. Aquí los niveles del nuevo factor se incorporarían como subparcelas dentro de la parcela experimental ya establecida, que actuaría como parcela principal.

El número de factores jerarquizados que interviene en estos diseños puede ser superior a dos, lo que da lugar a otras variantes del mismo diseño. Así, cuando se tienen tres factores se denomina parcelas subdivididas o "split-split-plot". En este caso, tendríamos tres tamaños de parcelas (principal, subparcela y sub-subparcela) y, por tanto, tres errores experimentales. No obstante, el aumento del número de factores jerarquizados complica notablemente las interacciones, los cálculos y la interpretación de los resultados, por lo que son menos utilizados.

FUNDAMENTOS DEL ANÁLISIS DE VARIANZA

Para explicar el fundamento del ANOVA en un diseño de parcelas divididas vamos a considerar un caso sencillo: un factor principal (A), un factor subordinado (B), ambos de efectos fijos y con a y b niveles, respectivamente, en un diseño en bloques al azar con n bloques. El modelo lineal para este caso, expresado en términos de desviaciones, sería:

$$Y_{ijr} - \mu = \rho_r + \alpha_i + \varepsilon_{ir} + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijr}$$

estableciendo que la desviación de una observación (Y_{ijr}) respecto a la media global (μ) está formada por seis componentes: las desviaciones causadas por los bloques (ρ_r), por el efecto del factor principal (α_i), por el factor subordinado (β_j) y su interacción con el principal ($(\alpha\beta)_{ij}$), y por dos errores aleatorios (ε_{ir} y ε_{ijr}). Así pues, la diferencia principal con el diseño factorial ordinario en bloques al azar es la existencia de dos errores:

Error(a), de las parcelas principales, con $(n-1) \times (a-1)$ grados de libertad

Error (ab), de las subparcelas, con $a \times (n-1) \times (b-1)$ grados de libertad

Como en un factorial ordinario en el que intervienen dos factores, las hipótesis nulas de interés son tres:

1. $H_0: (\alpha\beta)_{ij} = 0$ para toda i y j
2. $H_0: \alpha_i = 0$ para toda i
3. $H_0: \beta_j = 0$ para toda j

teniendo en cuenta que si la primera hipótesis se rechaza, es decir, la interacción entre A y B resultara significativa, las otras hipótesis carecen de interés. En el caso de un modelo II (efectos aleatorios), estas hipótesis vendrían formuladas como $\sigma_{\alpha\beta} = 0$, $\sigma_{\alpha} = 0$, y $\sigma_{\beta} = 0$, respectivamente.

Las sumas de cuadrados, grados de libertad y cuadrados medios se calcularían igual que para un factorial ordinario, aunque resulta algo más complicado. El cálculo de las sumas de cuadrados, utilizando la fórmula abreviada con el factor de corrección C, sería:

Factor de corrección: $C = \frac{Y_{...}^2}{nab}$

$$SCT (\text{Total}) = \sum_{i=1}^a \sum_{j=1}^b \sum_{r=1}^n Y_{ijr}^2 - C$$

$$SCR (\text{Bloques}) = \frac{1}{ab} \sum_{r=1}^n Y_{..r}^2 - C$$

$$SCA (\text{Factor A}) = \frac{1}{nb} \sum_{i=1}^a Y_{i..}^2 - C$$

$$SCEa [\text{Error (a)}] = \frac{1}{b} \sum_{i=1}^a \sum_{r=1}^n Y_{i,r}^2 - C - SCR - SCA$$

$$SCB (\text{Factor B}) = \frac{1}{na} \sum_{j=1}^b Y_{.j.}^2 - C$$

$$SCAB (\text{Interacción A} \times \text{B}) = \frac{1}{n} \sum_{i=1}^a \sum_{j=1}^b Y_{ij.}^2 - C - SCA - SCB$$

$$SCEab [\text{error (b)}] = SCT - SCR - SCA - SCEa - SCB - SCAB$$

La tabla del ANOVA para este caso sería:

Fuente de Variación	gl	SC	CM	F
Bloque	n-1	SCR		
Factor A	a-1	SCA	SCA / (a-1)	CMA/CMEa
Error (a)	(n-1)(a-1)	SCEa	SCEa / (n-1)(a-1)	
Factor B	b-1	SCB	SCB / (b-1)	CMB/CMEab
A×B	(a-1)(b-1)	SCAB	SCAB / (a-1)(b-1)	CMAB/CMEab
Error (ab)	a(n-1)(b-1)	SCEab	SCEab / a(n-1)(b-1)	
Total	nab-1	SCT		

La interpretación del análisis es similar a lo indicado para los experimentos factoriales (Capítulo 10). Si los tres valores de F calculados fueran próximos a 1, obtendríamos un valor P no significativo ($P > 0.05$) y aceptaríamos las tres hipótesis nulas de que no hay efecto de los factores A, B y de la interacción entre ellos. En cambio, si alguno de los valores de F fuera significativamente mayor que 1 ($P < 0.05$), rechazaríamos la hipótesis nula para ese caso, concluyendo que el efecto de ese factor o de la interacción es estadísticamente significativo. Conviene recordar que cuando la interacción es significativa, la significación de los efectos principales de los factores implicados carece de interés ya que éstos no son independientes. En este caso, la interpretación de los resultados se debe basar exclusivamente en el análisis de la interacción.

Un aspecto que cambia en el diseño en parcelas divididas respecto a un factorial ordinario es el cálculo de los errores estándares de las medias y de las diferencias entre medias, ya que hay que considerar los dos errores que intervienen en las parcelas divididas. Las diferentes medias y errores estándares que tenemos son:

$$\text{Factor A: } \bar{Y}_{i..} = \frac{\sum_{j=1}^b Y_{i.j.}}{nb}; \quad s_{\bar{A}} = \sqrt{\frac{CMEa}{nb}}$$

$$\text{Factor B: } \bar{Y}_{.j.} = \frac{\sum_{i=1}^a Y_{i.j.}}{na}; \quad s_{\bar{B}} = \sqrt{\frac{CMEab}{na}}$$

$$\text{Interacción A*B: } \bar{Y}_{ij.} = \frac{\sum_{i=1}^a \sum_{j=1}^b Y_{ij.}}{n}; \quad s_{\bar{Y}} = \sqrt{\frac{CMEab}{n}}$$

Los errores estándares de la diferencia entre dos medias dependerán de las medias a comparar y se calculan como sigue:

$$\text{Dos medias de A: } s_{\bar{dA}}(\bar{Y}_{2..} - \bar{Y}_{1..}) = \sqrt{\frac{2CMEa}{nb}}$$

$$\text{Dos medias de B: } s_{\bar{dB}}(\bar{Y}_{.2.} - \bar{Y}_{.1.}) = \sqrt{\frac{2CMEab}{na}}$$

$$\text{Dos medias de B para el mismo nivel de A: } s_{\bar{dAB}}(\bar{Y}_{12.} - \bar{Y}_{11.}) = \sqrt{\frac{2CMEab}{n}}$$

Dos medias de A para el mismo o diferente nivel de B:

$$s_{\bar{d}}[(\bar{Y}_{21.} - \bar{Y}_{11.}); (\bar{Y}_{22.} - \bar{Y}_{11.})] = \sqrt{\frac{2(b-1)CMEa + CMEa}{nb}}$$

REALIZACIÓN DEL ANÁLISIS DE VARIANZA

Supongamos un experimento de campo en el que se desea comparar la eficacia de dos tratamientos fungicidas (F_1 y F_2) entre sí y frente a un control no tratado (F_0). Para realzar el alcance de los tratamientos se ha incluido otro factor (Variedad) con tres niveles o variedades (V_1, V_2, V_3), que representan tres grados diferentes de susceptibilidad a una enfermedad. El diseño empleado es en parcelas divididas con 4 bloques completamente al azar, siendo el factor de las parcelas principales la variedad y el de las subparcelas el tratamiento fungicida. Al final del experimento se recolectaron todas las parcelas ($4 \times 3 \times 3 = 36$) y se determinó el peso de la cosecha (kg/parcela). En la Tabla 11.1 se muestran los pesos obtenidos, redondeados a valores enteros para facilitar los cálculos manuales.

Tabla 11.1.- Rendimiento de un cultivo en un experimento diseñado en parcelas divididas para estudiar el efecto de tres variedades y tres tratamientos fungicidas.

Fungicida	Rendimiento del cultivo (kg/parcela)			
	Bloque I	Bloque II	Bloque III	Bloque IV
Variedad V₁ (Susceptible)				
F ₀ (Control)	5	18	7	10
F ₁ (Fungicida 1)	22	13	6	16
F ₂ (Fungicida 2)	29	35	40	24
Variedad V₂ (Moderadamente susceptible)				
F ₀ (Control)	33	43	28	30
F ₁ (Fungicida 1)	38	47	51	57
F ₂ (Fungicida 2)	62	52	54	44
Variedad V₃ (Resistente)				
F ₀ (Control)	47	43	51	60
F ₁ (Fungicida 1)	52	56	49	38
F ₂ (Fungicida 2)	54	34	53	48

Para facilitar el cálculo se suelen utilizar dos tablas de totales, la tabla factor A × Bloque (Tabla 11.2) y la tabla factor A × factor B (Tabla 11.3).

Tabla 11.2. Totales de la interacción Variedad × Bloque

Variedad	Bloque				$Y_{i..} = \sum_{r=1}^n Y_{i.r}$
	Bloque I	Bloque II	Bloque III	Bloque IV	
V ₁	56	66	53	50	225
V ₂	133	142	133	131	539
V ₃	153	133	153	146	585
$Y_{..r} = \sum_{i=1}^a Y_{i.r}$	342	341	339	327	1349
$Y_{..r}^2 = \sum_{i=1}^a Y_{i.r}^2$	44234	42209	43907	40977	171327
$\bar{Y}_{..r} = \frac{Y_{..r}}{ab}$	38	40	35,6	36,3	$\bar{Y}_{...} = \frac{Y_{...}}{abn} = 37,47$

Tabla 11.3. Totales de la interacción Variedad × Fungicida

Variedad	Fungicida			$Y_{i..} = \sum_{j=1}^b Y_{i.j}$
	F ₀	F ₁	F ₂	
V ₁	40	57	128	225
V ₂	134	193	212	539
V ₃	201	195	189	585
$Y_{.j.} = \sum_{i=1}^a Y_{i.j}$	375	445	529	1349
$Y_{.j.}^2 = \sum_{i=1}^a Y_{i.j}^2$	59957	78523	97049	235529
$\bar{Y}_{.j.} = \frac{Y_{.j.}}{na}$	31,25	37,08	44,08	$\bar{Y}_{...} = 37,47$

Se calcula el factor de corrección (C) y las sumas de cuadrados:

$$C = \frac{Y_{...}^2}{nab} = \frac{(1349)^2}{4 \times 3 \times 3} = \frac{1819801}{36} = 50550,03$$

$$SCT(\text{Total}) = \sum_{i=1}^a \sum_{j=1}^b \sum_{r=1}^n Y_{ijr}^2 - C = (5^2 + 22^2 + \dots + 38^2 + 48^2) - C = 60339 - 50550,03 = 9788,97$$

$$SCR \text{ (Bloques)} = \frac{1}{a \times b} \sum_{r=1}^n Y_{..r}^2 - C = \frac{342^2 + 341^2 + 339^2 + 327^2}{3 \times 3} - C = 50566,11 - 50550,03 = 16,08$$

$$SCA \text{ (Factor A)} = \frac{1}{n \times b} \sum_{i=1}^a Y_{i..}^2 - C = \frac{225^2 + 539^2 + 585^2}{4 \times 3} - C = 56947,58 - 50550,03 = 6397,55$$

$$SCEa \text{ [Error (a)]} = \frac{1}{b} \sum_{i=1}^a \sum_{r=1}^n Y_{i.r}^2 - C - SCR - SCA = \frac{171327}{3} - 50550,03 - 16,08 - 6397,55 = 145,33$$

$$SCB \text{ (Factor B)} = \frac{1}{a \times n} \sum_{j=1}^b Y_{.j.}^2 - C = \frac{375^2 + 445^2 + 529^2}{3 \times 4} - C = 51540,92 - 50550,03 = 990,89$$

$$SCAB \text{ (Interacción A} \times \text{B)} = \frac{1}{n} \sum_{i=1}^a \sum_{j=1}^b Y_{ij.}^2 - C - SCA - SCB = \frac{235529}{4} - 50550,03 - 6397,55 - 990,89 = 943,78$$

$$SCEab \text{ [error (b)]} = SCT - SCR - SCA - SCEa - SCB - SCAB = 9788,97 - 16,08 - 6397,55 - 145,33 - 990,89 - 943,78 = 1295,34$$

Se obtiene la siguiente tabla ANOVA:

Fuente de Variación	gl	SC	CM	F	P
Bloque	3	16,08	5,36		
Factor A	2	6397,55	3198,78	132,06	0,0000
Error (a)	6	145,33	24,22		
Factor B	2	990,89	495,44	6,88	0,0060
AxB	4	943,78	235,94	3,28	0,0348
Error (ab)	18	1295,34	71,96		
Total	35	9788,97			

Los resultados del ANOVA indican que los dos factores, variedad y fungicida, tienen un efecto significativo ($P < 0.05$) en la cosecha, pero también la interacción variedad \times fungicida, lo que indica que el efecto de un factor varía con los niveles del otro factor. Este hecho es muy patente si se consideran los datos separadamente por variedades. En la variedad resistente a la enfermedad (V_3) no existen diferencias entre fungicidas, mientras que en las variedades moderadamente susceptible (V_2) y susceptible (V_1), estas diferencias son acusadas, destacando por su mayor eficacia el fungicida F_2 . La comparación de medias para este ejemplo se realizaría conforme se indica en el Capítulo 5 y se especifica en el apartado siguiente.

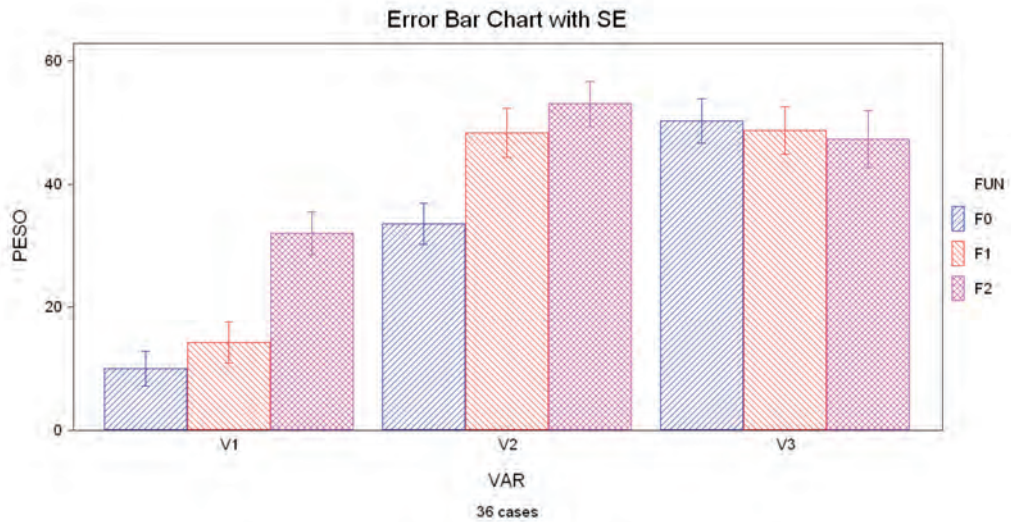
CÁLCULO CON EL PROGRAMA *Statistix (SX)*

Se introducen los datos del experimento denominando al factor principal como Variedad (V_1 , V_2 , V_3) y al subordinado como Fungicida (F_1 , F_2 , F_3), a los bloques como Rep (con cuatro valores) y a la variable dependiente como Peso (peso de la cosecha en kg/parcela). Una vez introducidos los datos se obtiene la siguiente matriz:

CASE	VAR	FUN	REP	PESO
1	1	1	1	5
2	1	1	2	18
3	1	1	3	7
4	1	1	4	10
5	1	2	1	22
6	1	2	2	13
7	1	2	3	6
8	1	2	4	16
9	1	3	1	29
10	1	3	2	35
11	1	3	3	40
12	1	3	4	24
13	2	1	1	33
14	2	1	2	43
15	2	1	3	28
16	2	1	4	30
17	2	2	1	38
18	2	2	2	47
19	2	2	3	51
20	2	2	4	57
21	2	3	1	62
22	2	3	2	52
23	2	3	3	54
24	2	3	4	44
25	3	1	1	47
26	3	1	2	43
27	3	1	3	51
28	3	1	4	60
29	3	2	1	52
30	3	2	2	56
31	3	2	3	49
32	3	2	4	38
33	3	3	1	54
34	3	3	2	34
35	3	3	3	53
36	3	3	4	48

Antes de proceder con el ANOVA se recomienda visualizar los datos. Para ello, se elige la siguiente secuencia:

Statistics>Summary Statistics>Error Bar Chart



Se obtiene así una visualización bastante clara de los datos: las varianzas parecen homogéneas y también parece que existe una clara interacción entre los dos factores variedad y fungicida. Entrando en la opción “Box and Whisker Plot” del mismo menú (*Summary Statistics*) se comprueba que las varianzas son homogéneas y que no existen “outliers”. Se procede al análisis eligiendo la siguiente secuencia:

Statistics>Linear Models>Analysis of Variance>Split-Plot Design...

En la pantalla resultante se introduce el nombre de la variable dependiente (Peso), las repeticiones o bloques (Rep), el factor de la parcela principal (VAR), el factor de la subparcela (FUN) y se pulsa OK. Se obtiene así la tabla del ANOVA:

<u>Analysis of Variance Table for PESO</u>					
Source	DF	SS	MS	F	P
REP	3	16.08	5.36		
VAR	2	6397.56	3198.78	132.06	0.0000
Error REP*VAR	6	145.33	24.22		
FUN	2	990.89	495.44	6.88	0.0060
VAR*FUN	4	943.78	235.94	3.28	0.0348
Error REP*VAR*FUN	18	1295.33	71.96		
Total	35	9788.97			
Grand Mean 37.472					
CV (REP*VAR) 13.13					
CV (REP*VAR*FUN) 22.64					

Los resultados del análisis coinciden totalmente con el cálculo manual realizado anteriormente. Obsérvese que debajo de la tabla se indica la media general del experimento ($\bar{Y}...=37,472$) y dos coeficientes de variación (CV), pues hay dos errores. El primer CV da el grado de precisión asociado al factor de la parcela principal y el segundo se refiere al factor de la subparcela y su interacción con el factor de la parcela principal.

Como en otros ejemplos hay que comprobar si se cumplen algunos preceptos del ANOVA. En este análisis no se incluye ningún test de homogeneidad de varianzas. Para disponer de ellos, habría que realizar el ANOVA como si fuera unifactorial en un diseño completamente aleatorio. Los resultados para la Variedad y el Fungicida por separado serían:

Completely Randomized AOV for PESO (VARIEDAD)

Source	DF	SS	MS	F	P
VAR	2	6397.56	3198.78	31.1	0.0000
Error	33	3391.42	102.77		
Total	35	9788.97			

Grand Mean 37.472 CV 27.05

	Chi-Sq	DF	P
Bartlett's Test of Equal Variances	2.17	2	0.3378
Cochran's Q	0.4323		
Largest Var / Smallest Var	2.4027		

Component of variance for between groups 258.001
Effective cell size 12.0

VAR Mean

V1 18.750

V2 44.917

V3 48.750

Observations per Mean 12

Standard Error of a Mean 2.9265

Std Error (Diff of 2 Means) 4.1386

Completely Randomized AOV for PESO (FUNGICIDA)

Source	DF	SS	MS	F	P
FUN	2	990.89	495.444	1.86	0.1719
Error	33	8798.08	266.609		
Total	35	9788.97			

Grand Mean 37.472 CV 43.57

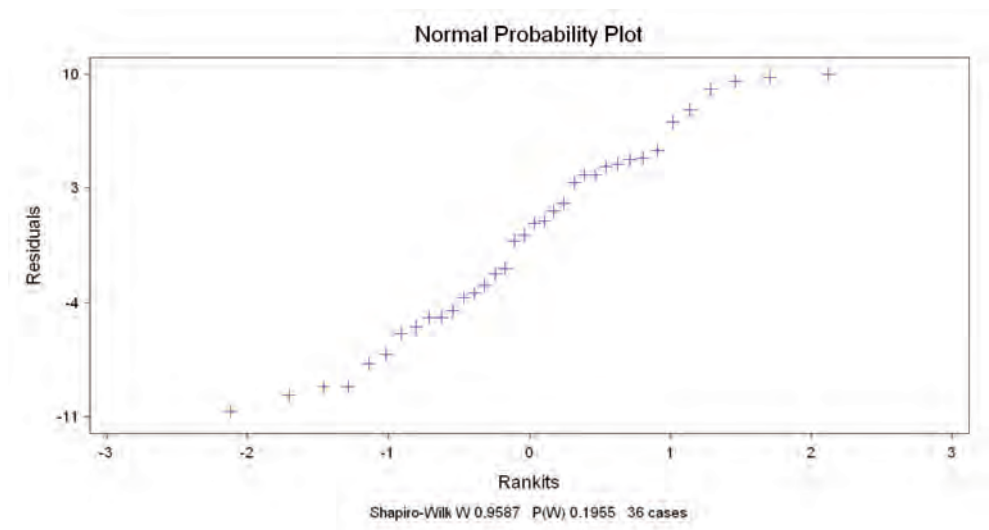
	Chi-Sq	DF	P
Bartlett's Test of Equal Variances	2.44	2	0.2954
Cochran's Q	0.4160		
Largest Var / Smallest Var	2.4290		

Component of variance for between groups 19.0697
 Effective cell size 12.0

FUN	Mean
F0	31.250
F1	37.083
F2	44.083
Observations per Mean	12
Standard Error of a Mean	4.7135
Std Error (Diff of 2 Means)	6.6659

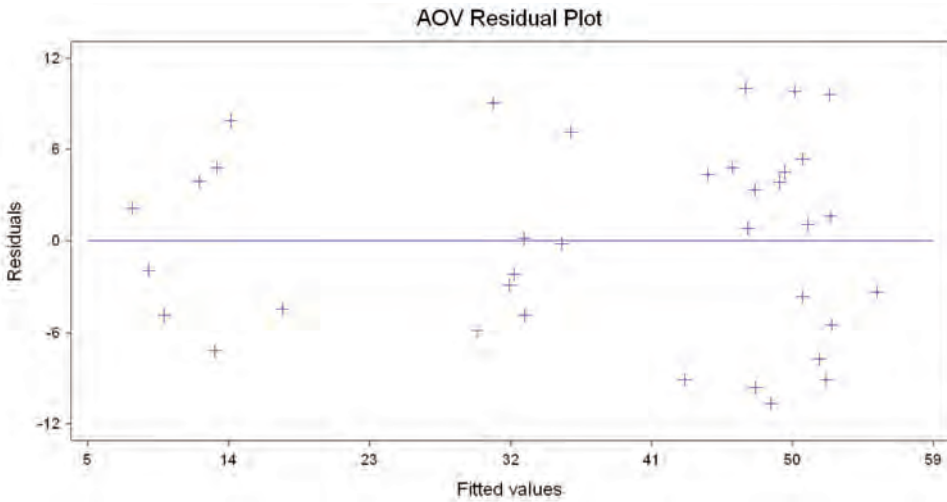
El *test de Bartlett*, el *test de Cochran* y la razón *Varianza mayor/Varianza menor* confirman que las varianzas son homogéneas. La normalidad de los datos se comprueba eligiendo la siguiente secuencia en la pantalla de resultados:

Results>Plots>Normal Probability Plot



El *test Shapiro-Wilk* muestra que los datos se ajustan bien a una normal ($P > 0.05$). Asimismo, se puede ver la distribución que presentan los residuos en la misma pantalla de resultados siguiendo la secuencia:

Results>Plots>Resids By Fitted Values



La distribución observada en el gráfico es un ejemplo de distribución aleatoria sin ningún patrón específico, lo que confirma la independencia de los errores.

Finalmente, en la misma pantalla de resultados se puede realizar la comparación de medias, teniendo en cuenta que la interacción es significativa. Como en este ejemplo el objetivo *a priori* era comparar dos fungicidas entre sí (F1 vs. F2) y frente al testigo sin tratar (F0) en tres situaciones diferentes representadas por tres variedades con diferente grado de resistencia a una enfermedad, optaríamos por los contrastes ortogonales. Para ello, se elige la opción “Contrast...” del menú de resultados. En la pantalla que aparece se seleccionan la interacción VAR*FUN, el factor VAR, y los coeficientes del contraste 2 -1 -1 y 0 1 -1. En el primer contraste se comparan los dos fungicidas frente al testigo sin tratar, ya que el orden de los fungicidas en la matriz de datos del experimento es F0, F1 y F2. En el segundo contraste no se considera el testigo sin tratar, al que se asigna el coeficiente 0, y se compara sólo entre los dos fungicidas. Como se ha elegido la variable VAR, eso significa que los contrastes se realizarán para cada nivel de esa variable. De esta forma se obtiene:

AOV Contrasts of PESO by FUN for VAR

Contrast Number 1

Contrast Coefficients: 2 -1 -1

FOR VAR = V1

Contrast	-26.250	SS (Contrast)	459.38
----------	---------	---------------	--------

FOR VAR = V2

Contrast	-34.250	SS (Contrast)	782.04
----------	---------	---------------	--------

FOR VAR = V3

Contrast	4.5000	SS (Contrast)	13.500
----------	--------	---------------	--------

```

Contrast Number 2
Contrast Coefficients: 0 1 -1

FOR VAR = V1
Contrast          -17.750          SS (Contrast)          630.13

FOR VAR = V2
Contrast          -4.7500          SS (Contrast)          45.125

FOR VAR = V3
Contrast           1.5000          SS (Contrast)          4.5000
    
```

El programa da los valores de la suma de cuadrados correspondiente a cada contraste. En este caso no da la significación del contraste porque, como se ha indicado anteriormente, hay dos errores experimentales y, por tanto, diferentes errores estándares para comparación de medias. Según la comparación deseada, se elige su error estándar correspondiente y se realiza el test de la *t de Student*, que para dos medias es equivalente a la *F de Fisher*. En nuestro ejemplo, como los contrastes establecidos *a priori* implican la comparación de medias entre tratamientos fungicidas dentro de cada variedad, habría un único error para la comparación: Error (ab). Considerando que cada contraste consume un grado de libertad y que el cuadrado medio del error es $CME_{ab} = 71,96$, con 18 gl, obtendríamos la siguiente tabla de significación de los contrastes:

Variedad	Contraste	SC = CM	F	P
V ₁	F ₀ vs. (F ₁ + F ₂)	459,38	6,38	0,0214
V ₁	F ₁ vs. F ₂	630,13	8,76	0,0084
V ₂	F ₀ vs. (F ₁ + F ₂)	782,04	10,87	0,0040
V ₂	F ₁ vs. F ₂	45,125	0,63	0,4377
V ₃	F ₀ vs. (F ₁ + F ₂)	13,50	0,19	0,6681
V ₃	F ₁ vs. F ₂	4,50	0,062	0,8062

Los resultados obtenidos con los contrastes permiten clarificar la naturaleza de la interacción Variedad*Fungicida, demostrando que para cada variedad se alcanza una conclusión diferente. Así, para la variedad susceptible (V₁) los dos fungicidas difieren del testigo y entre sí; para la variedad moderadamente susceptible (V₂) los fungicidas difieren del testigo, pero no hay diferencias entre ellos; y para la variedad resistente (V₃) no hay diferencias entre los tres tratamientos.

Aunque faltan los datos de la gravedad de los ataques de la enfermedad en las distintas parcelas, suponiendo que las diferencias en el rendimiento del cultivo se deben exclusivamente

a la enfermedad y que la variedad resistente no fue afectada, se podría decir que los tratamientos fungicidas redujeron la enfermedad y que el fungicida F_2 fue más eficaz que el F_1 . Este hecho destacó especialmente en la variedad susceptible V_1 , la más afectada por la enfermedad. En la variedad resistente los fungicidas no tuvieron ningún efecto, ya que no fue afectada por la enfermedad.

CAPÍTULO 12

DISEÑO EN BLOQUES DIVIDIDOS

Los diseños en bloques divididos, “strip-plot”, “split-block” o “criss-cross”, constituyen un caso particular de los diseños “split-plot”, que a su vez forman parte de los experimentos factoriales, en el que los factores que intervienen no se combinan aleatoriamente entre sí, sino que están subordinados unos a otros. En este caso los niveles de un primer factor A se asignan a franjas de parcelas a lo largo de los bloques en una dirección, mientras que los niveles del segundo factor B se asignan a franjas de parcelas orientadas perpendicularmente a las del primero. Debido a la orientación perpendicular de los niveles de los dos factores, éstos suelen denominarse factor horizontal (A) y factor vertical (B). Para cada bloque se realiza una aleatorización independiente de los niveles de los dos factores.

Para comprenderlo mejor, veamos un experimento sencillo en el que intervienen dos factores: uno horizontal (A) con tres niveles (a_1 , a_2 y a_3) y otro vertical (B) con cuatro niveles (b_1 , b_2 , b_3 y b_4). Si el diseño experimental fuese en bloques completos al azar, tendríamos el siguiente esquema del experimento para dos de los bloques:

	Bloque I				Bloque II				
	b_2	b_4	b_1	b_3	b_1	b_3	b_2	b_4	
a_3	a_3b_2	a_3b_4	a_3b_1	a_3b_3	a_2	a_2b_1	a_2b_3	a_2b_2	a_2b_4
a_1	a_1b_2	a_1b_4	a_1b_1	a_1b_3	a_3	a_3b_1	a_3b_3	a_3b_2	a_3b_4
a_2	a_2b_2	a_2b_4	a_2b_1	a_2b_3	a_1	a_1b_1	a_1b_3	a_1b_2	a_1b_4

En el diseño en bloque divididos hay tres tamaños de parcelas: la franja horizontal de cuatro parcelas en las que se aplica el factor A, la franja vertical de tres parcelas en las que se aplica el factor B y las pequeñas parcelas de la intersección que contienen las combinaciones $a_i b_j$. Existen, por tanto, tres errores experimentales, uno por cada tamaño de parcela. Generalmente, el error de las parcelas pequeñas es menor que el de las parcelas principales.

La principal ventaja del este diseño, comparando con el diseño de parcelas divididas, es que facilita la aplicación de factores que son difíciles de aplicar en parcelas pequeñas. Por ejemplo, en un ensayo de comparación de variedades puede interesar que existan filas completas de cada una de ellas por diversas razones, como que forma parte de un experimento demostrativo y se pretenda observar el hábito de crecimiento en visitas específicas; si a la vez se pretende aprovechar el experimento para comparar sistemas de mantenimiento del suelo, que obliga a aplicar tratamientos en parcelas relativamente grandes, el diseño en bloques divididos resulta de gran utilidad. Al igual que el de parcelas divididas, presenta una diferente precisión en la estimación de los efectos principales y de la interacción, siendo

generalmente esta precisión mayor para la interacción que para los efectos principales. También este diseño es favorable para incorporar nuevos factores en un experimento en curso.

Como se indicó para el “split-plot”, el número de factores jerarquizados que interviene en estos diseños factoriales puede ser superior a dos, lo que da lugar a otras variantes del mismo diseño. Así, cuando tenemos tres factores se denomina parcelas subdivididas o “split-split-plot”, donde existen tres tamaños de parcelas (principal, subparcela y sub-subparcela) y, por tanto, tres errores experimentales. Una variante del diseño “strip-plot” se denomina “strip-split-plot”, donde las parcelas pequeñas de intersección de los dos factores principales se dividen en subparcelas para incluir a un tercer factor. No obstante, conviene recordar que el aumento del número de factores jerarquizados complica notablemente las interacciones, los cálculos y la interpretación de los resultados, por lo que son menos utilizados.

FUNDAMENTOS DEL ANÁLISIS DE VARIANZA

Para explicar el fundamento del ANOVA en un diseño de bloque divididos vamos a considerar un caso sencillo: un factor horizontal (A), un factor vertical (B), ambos de efectos fijos y con a y b niveles, respectivamente, en un diseño en bloques al azar con n bloques. El modelo lineal para este caso, expresado en términos de desviaciones, sería:

$$Y_{ijr} - \mu = \rho_r + \alpha_i + \varepsilon_{ir} + \beta_j + \varepsilon_{jr} + (\alpha\beta)_{ij} + \varepsilon_{ijr}$$

estableciendo que la desviación de una observación (Y_{ijr}) respecto a la media global (μ) está formada por seis componentes: las desviaciones causadas por los bloques (ρ_r), por el efecto del factor horizontal (α_i) por el factor vertical (β_j) y la interacción entre ambos ($\alpha\beta$) $_{ij}$, y por tres errores aleatorios (ε_{ir} , ε_{jr} y ε_{ijr}). Así pues, la diferencia principal con el diseño factorial ordinario y con el “split plot” en bloques al azar es la existencia de tres errores:

Error(a), de las franjas horizontales, con $(n-1) \times (a-1)$ grados de libertad.

Error (b), de las franjas verticales, con $(n-1) \times (b-1)$ grados de libertad.

Error (ab), de las parcelas de intersección (interacción A*B), con $(n-1) \times (a-1) \times (b-1)$ grados de libertad.

Como en un factorial ordinario en el que intervienen dos factores, las hipótesis nulas de interés son tres:

1. $H_0: (\alpha\beta)_{ij} = 0$ para toda i y j
2. $H_0: \alpha_i = 0$ para toda i
3. $H_0: \beta_j = 0$ para toda j

teniendo en cuenta que si la primera hipótesis se rechaza, es decir la interacción entre A y B resultara significativa, las otras hipótesis carecen de interés. En el caso de un modelo II (efectos aleatorios), estas hipótesis vendrían formuladas como $\sigma_{\alpha\beta} = 0$, $\sigma_\alpha = 0$, y $\sigma_\beta = 0$, respectivamente.

Las sumas de cuadrados, grados de libertad y cuadrados medios se calculan igual que para un factorial ordinario, aunque resulta algo más complicado. El cálculo de las sumas de cuadrados, utilizando la fórmula abreviada con el factor de corrección C , sería:

Factor de corrección: $C = \frac{Y_{\dots}^2}{nab}$

$$SCT \text{ (Total)} = \sum_{i=1}^a \sum_{j=1}^b \sum_{r=1}^n Y_{ijr}^2 - C$$

$$SCR \text{ (Bloques)} = \frac{1}{ab} \sum_{r=1}^n Y_{\dots r}^2 - C$$

$$SCA \text{ (Factor A)} = \frac{1}{nb} \sum_{i=1}^a Y_{i\dots}^2 - C$$

$$SCEa \text{ [Error (a)]} = \frac{1}{b} \sum_{i=1}^a \sum_{r=1}^n Y_{i\dots r}^2 - C - SCR - SCA$$

$$SCB \text{ (Factor B)} = \frac{1}{na} \sum_{j=1}^b Y_{\dots j}^2 - C$$

$$SCEb \text{ [Error (b)]} = \frac{1}{a} \sum_{j=1}^b \sum_{r=1}^n Y_{\dots jr}^2 - C - SCR - SCB$$

$$SCAB \text{ (Interacción A} \times \text{B)} = \frac{1}{n} \sum_{i=1}^a \sum_{j=1}^b Y_{ij\dots}^2 - C - SCA - SCB$$

$$SCEab \text{ [error (b)]} = SCT - SCR - SCA - SCEa - SCB - SCEb - SCAB$$

La tabla del ANOVA para este caso sería:

Fuente de Variación	gl	SC	CM	F
Bloque	n-1	SCR		
Factor A	a-1	SCA	SCA / (a-1)	CMA/CMEa
Error (a)	(n-1)(a-1)	SCEa	SCEa / (n-1)(a-1)	
Factor B	b-1	SCB	SCB / (b-1)	CMB/CMEb
Error (b)	(n-1)(b-1)	SCEb	SCEb/(n-1)(b-1)	
A×B	(a-1)(b-1)	SCAB	SCAB / (a-1)(b-1)	CMAB/CMEab
Error (ab)	(n-1)(a-1)(b-1)	SCEab	SCEab / (n-1)(a-1)(b-1)	
Total	nab-1	SCT		

La interpretación del análisis es similar a lo indicado para los experimentos factoriales (Capítulo 10). Si los tres valores de F calculados fueran próximos a 1, obtendríamos un valor P no significativo ($P > 0.05$) y aceptaríamos las tres hipótesis nulas de que no hay efecto de los

factores A, B y de la interacción entre ellos. En cambio, si alguno de los valores de F fuera significativamente mayor que 1 ($P < 0.05$), rechazaríamos la hipótesis nula para ese caso, concluyendo que el efecto de ese factor o de la interacción es estadísticamente significativo. Conviene recordar que cuando la interacción es significativa, la significación de los efectos principales de los factores implicados carece de interés ya que éstos no son independientes. En este caso, la interpretación de los resultados se debe basar exclusivamente en el análisis de la interacción.

Un aspecto que cambia en el diseño en bloques divididos respecto a un factorial ordinario, o a las parcelas divididas, es el cálculo de los errores estándares de las medias y de las diferencias entre medias, ya que hay que considerar los tres errores que intervienen en los bloques divididos. Las diferentes medias y errores estándares son:

$$\text{Factor A: } \bar{Y}_{i..} = \frac{\sum_{j=1}^b Y_{i.j.}}{nb} ; \quad s_{\bar{A}} = \sqrt{\frac{CMEa}{nb}}$$

$$\text{Factor B: } \bar{Y}_{.j.} = \frac{\sum_{i=1}^a Y_{i.j.}}{na} ; \quad s_{\bar{B}} = \sqrt{\frac{CMEb}{na}}$$

$$\text{Interacción A*B: } \bar{Y}_{ij.} = \frac{\sum_{i=1}^a \sum_{j=1}^b Y_{ij.}}{n} ; \quad s_{\bar{Y}} = \sqrt{\frac{CMEab}{n}}$$

Los errores estándares de la diferencia entre dos medias dependerán de las medias a comparar y se calculan como sigue:

$$\text{Dos medias de A: } s_{\bar{dA}}(\bar{Y}_{2..} - \bar{Y}_{1..}) = \sqrt{\frac{2CMEa}{nb}}$$

$$\text{Dos medias de B: } s_{\bar{dB}}(\bar{Y}_{.2.} - \bar{Y}_{.1.}) = \sqrt{\frac{2CMEb}{na}}$$

$$\text{Dos medias de A para el mismo nivel de B: } s_{\bar{dAB}}(\bar{Y}_{21.} - \bar{Y}_{11.}) = \sqrt{\frac{2[(b-1)CMEab + CMEa]}{nb}}$$

$$\text{Dos medias de B para el mismo nivel de A: } s_{\bar{dBA}}(\bar{Y}_{12.} - \bar{Y}_{11.}) = \sqrt{\frac{2[(a-1)CMEab + CMEb]}{na}}$$

Dos medias de A para diferente nivel de B o dos medias de B para diferente nivel de A:

$$s_{\bar{d}}(\bar{Y}_{22.} - \bar{Y}_{11.}) = \sqrt{\frac{2[(ab - a - b)CMEab + aCMEa + bCMEb]}{nab}}$$

REALIZACIÓN DEL ANÁLISIS DE VARIANZA

Supongamos un experimento de campo en el que se desea comparar el efecto de dos sistemas de manejo del suelo sobre el desarrollo de una enfermedad foliar del olivar pero, al mismo tiempo, se desea conocer el efecto sobre la enfermedad de varios momentos de aplicación de un fungicida protector. Tenemos dos factores, el sistema de manejo del suelo (SUELO) con dos niveles (suelo desnudo y cubierta vegetal) y los momentos de los tratamientos (TRATAMIENTO) con cuatro niveles (sin tratar, aplicación en primavera, aplicación en otoño y aplicación en primavera y otoño). Por la dificultad de aplicación de los dos factores en el campo, se optó por un diseño en “strip-plot”, con 5 bloques completos al azar. El experimento se inició en marzo en una finca que presentaba un nivel de infección elevado y uniforme. En febrero del año siguiente se tomó al azar una muestra de 2000 hojas jóvenes (desarrolladas en la primavera anterior) de cada una de las 40 parcelas experimentales (5×2×4) y se determinó el porcentaje de hojas afectadas por la enfermedad. En la Tabla 12.1 se recogen los datos obtenidos.

Tabla 12.1. Porcentaje de hojas de olivo afectadas en un experimento diseñado en bloques divididos para estudiar el efecto sobre una enfermedad de dos manejos del suelo y cuatro tratamientos fungicidas.

Tratamiento	Hojas enfermas (%)				
	Bloque I	Bloque II	Bloque III	Bloque IV	Bloque V
Suelo con cubierta vegetal					
Sin tratar	85,5	87,8	81,4	74,3	94,5
Primavera	16,7	12,0	38,1	22,2	18,0
Otoño	65,5	60,7	38,2	78,7	56,9
Primavera+Otoño	6,2	7,7	5,4	3,6	15,0
Suelo desnudo					
Sin tratar	89,3	76,8	70,0	78,5	81,4
Primavera	6,9	24,8	10,0	14,1	9,2
Otoño	52,3	68,7	41,5	28,1	38,6
Primavera+Otoño	0,7	6,3	11,2	2,9	4,5

Al tratarse de datos porcentuales que cubren un amplio intervalo (1–95%), la primera recomendación es realizar la transformación angular ($\arcseno \sqrt{Y/100}$) para ajustarlos a una distribución normal (véase la Tabla A9). Por ello, se van a realizar todos los cálculos con los datos transformados suponiendo que éstos no incumplen los supuestos del ANOVA. Los datos transformados aparecen en la Tabla 12.2.

Tabla 12.2. Valores transformados mediante $Y_T = \arcseno \sqrt{Y/100}$ de los datos de la Tabla 12.1.

Tratamiento	Hojas enfermas (%)				
	Bloque I	Bloque II	Bloque III	Bloque IV	Bloque V
Suelo con cubierta vegetal					
Sin tratar	67,6	69,6	64,4	59,5	76,4
Primavera	24,1	20,3	38,1	28,1	25,1
Otoño	54,0	51,2	38,2	62,5	49,0
Primavera+Otoño	14,4	16,1	13,4	10,9	22,8
Suelo desnudo					
Sin tratar	70,9	61,2	56,8	62,4	64,5
Primavera	15,2	29,9	18,4	22,1	17,7
Otoño	46,3	56,0	40,1	32,0	38,4
Primavera+Otoño	4,8	14,5	19,6	9,8	12,2

En el diseño en bloque divididos, para facilitar el cálculo se suelen utilizar tres tablas de totales, la tabla factor A*Bloque (Tabla 12.3), la tabla factor B*Bloque (Tabla 12.4) y la tabla Factor A*Factor B (Tabla 12.5).

Tabla 12.3. Totales de la interacción Suelo*Bloque

Suelo	Bloque					$Y_{i..} = \sum_{r=1}^n Y_{i.r}$
	I	II	III	IV	V	
Cubierta vegetal	160,1	157,2	154,1	161,0	173,3	805,7
Desnudo	137,2	161,6	134,9	126,3	132,8	692,8
$Y_{..r} = \sum_{i=1}^a Y_{i.r}$	297,3	318,8	289,0	287,3	306,1	1498,5
$Y_{..r}^2 = \sum_{i=1}^a Y_{i.r}^2$	44455,85	50826,4	41944,82	41872,69	47668,73	226768,49
$\bar{Y}_{..r} = \frac{Y_{..r}}{ab}$	37,16	39,85	36,13	35,91	38,26	$\bar{Y}_{...} = \frac{Y_{...}}{nab} = 37,46$

Tabla 12.4. Totales de la interacción Tratamiento*Bloque

Tratamiento	Bloque					$Y_{.j.} = \sum_{r=1}^n Y_{.jr}$
	I	II	III	IV	V	
Sin tratar	138,5	130,8	121,2	121,9	140,9	653,3
Primavera	39,3	50,2	56,5	50,2	42,8	239,0
Otoño	100,3	107,2	78,3	94,5	87,4	467,7
Primavera+Otoño	19,2	30,6	33,0	20,7	35,0	138,5
$Y_{..r} = \sum_{j=1}^b Y_{.jr}$	297,3	318,8	289,0	287,3	306,1	1498,5
$Y_{..r}^2 = \sum_{j=1}^b Y_{.jr}^2$	31155,47	32056,88	25101,58	26738,39	30548,41	145600,73
$\bar{Y}_{..r} = \frac{Y_{..r}}{ab}$	37,16	39,85	36,13	35,91	38,26	$\bar{Y}_{...} = \frac{Y_{...}}{nab} = 37,46$

Tabla 12.5. Totales de la interacción Suelo*Tratamiento

Suelo	Tratamiento				
	Sin tratar	Primavera	Otoño	Primavera+Otoño	
Cubierta vegetal	337,5	135,7	254,9	77,6	805,7
Desnudo	315,8	103,3	212,8	60,9	692,8
$Y_{.j.} = \sum_{i=1}^a Y_{ij.}$	653,3	239,0	467,7	138,5	1498,5
$Y_{.j.}^2 = \sum_{i=1}^a Y_{ij.}^2$	213635,89	29085,38	110257,85	9730,57	362709,69
$\bar{Y}_{.j.} = \frac{Y_{.j.}}{na}$	65,33	23,90	46,77	13,85	$\bar{Y}_{...} = \frac{Y_{...}}{nab} = 37,46$

Se calcula el factor de corrección (C) y las sumas de cuadrados:

$$C = \frac{Y_{...}^2}{nab} = \frac{(1498,5)^2}{5 \times 2 \times 4} = \frac{2245502,25}{40} = 56137,56$$

$$SCT \text{ (Total)} = \sum_{i=1}^a \sum_{j=1}^b \sum_{r=1}^n Y_{ijr}^2 - C = (67,6^2 + 69,6^2 + \dots + 9,8^2 + 12,2^2) - C = 73957,73 - 56137,56 = 17820,17$$

$$SCR \text{ (Bloques)} = \frac{1}{a \times b} \sum_{r=1}^n Y_{..r}^2 - C = \frac{297,3^2 + \dots + 306,1^2}{2 \times 4} - C = 56222,53 - 56137,56 = 84,97$$

$$SCA \text{ (Factor A)} = \frac{1}{n \times b} \sum_{i=1}^a Y_{i..}^2 - C = \frac{805,7^2 + 692,8^2}{5 \times 4} - C = 56456,22 - 56137,56 = 318,66$$

$$SCEa \text{ [Error (a)]} = \frac{1}{b} \sum_{i=1}^a \sum_{r=1}^n Y_{i.r}^2 - C - SCR - SCA = \frac{226768,49}{4} - 56137,56 - 84,97 - 318,66 = 150,93$$

$$SCB \text{ (Factor B)} = \frac{1}{a \times n} \sum_{j=1}^b Y_{.j.}^2 - C = \frac{653,3^2 + 239,0^2 + 467,7^2 + 138,5^2}{2 \times 5} - C = 72184,74 - 56137,56 = 16047,18$$

$$SCEb \text{ [Error (b)]} = \frac{1}{a} \sum_{j=1}^b \sum_{r=1}^n Y_{.jr}^2 - C - SCR - SCB = \frac{145600,73}{2} - 56137,56 - 84,97 - 16047,18 = 530,65$$

$$SCAB \text{ (Interacción AxB)} = \frac{1}{n} \sum_{i=1}^a \sum_{j=1}^b Y_{ij.}^2 - C - SCA - SCB = \frac{362709,69}{5} - 56137,56 - 318,66 - 16047,18 = 38,54$$

$$SCEab \text{ [Error (ab)]} = SCT - SCR - SCA - SCEa - SCB - SCEb - SCAB = 17820,17 - 84,97 - 318,66 - 150,93 - 16047,18 - 530,65 - 38,54 = 649,24$$

Se obtiene la siguiente tabla ANOVA:

Fuente de Variación	gl	SC	CM	F	P
Bloque	4	84,97	21,24		
Factor A	1	318,66	318,66	8,45	0,0438
Error (a)	4	150,93	37,73		
Factor B	3	16047,18	5349,06	120,96	0,0000
Error (b)	12	530,65	44,22		
AxB	3	38,54	12,85	0,24	0,8668
Error (ab)	12	649,24	54,10		
Total	39	17820,17			

Los resultados del ANOVA indican que los dos factores, sistema de manejo del suelo y tratamientos fungicidas, tienen un efecto significativo ($P < 0.05$) en el porcentaje de hojas afectadas por la enfermedad. En cambio, la interacción suelo \times tratamiento fungicida no tiene efecto significativo ($P > 0.05$), lo que indica que los dos factores tienen efectos independientes. En este caso, para completar el análisis, bastaría con realizar las comparaciones de medias para los dos factores por separado. En el caso del manejo del suelo, al haber sólo dos niveles, concluiríamos que la cubierta vegetal ha resultado más favorable que el suelo des-

nudo para el desarrollo de la enfermedad. En el caso de los tratamientos fungicidas, habría que realizar algún test de comparación de medias, conforme se indica en el Capítulo 5, para establecer conclusiones sobre los diferentes tratamientos. Esta comparación y la comprobación de los supuestos del ANOVA se realizarán en el siguiente apartado utilizando el programa *Statistix*.

CÁLCULO CON EL PROGRAMA *Statistix* (SX)

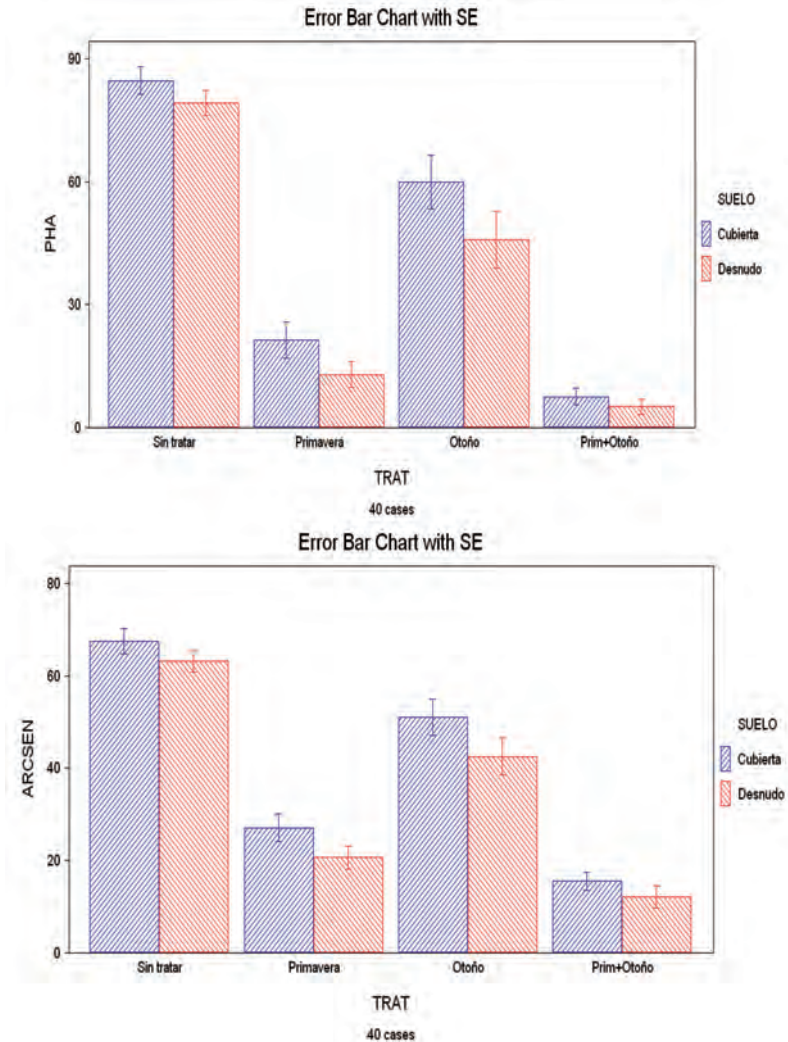
Se introducen los datos del experimento denominando al factor sistema de manejo del suelo como SUELO, con dos niveles: 1-Cubierta y 2-Desnudo. Al segundo factor, considerado como vertical, lo denominamos TRAT, con cuatro niveles: 1-Sin tratar, 2-Primavera, 3-Otoño y 4-Primavera+Otoño. Los bloques se designan como Rep (con cinco valores) y la variable dependiente (porcentaje de hojas afectadas) como PHA. Como los datos originales (Y) están expresados en porcentaje, realizamos también una transformación de los mismos mediante la función angular ($Y' = \arcseno \sqrt{Y/100}$), utilizando la opción "Transformations.." del menú "Data" (véase la Tabla A9). A la variable transformada la denominamos ARCSEN, para realizar un análisis comparativo con la variable original PHA. Se obtiene la siguiente matriz de datos:

CASE	SUELO	TRAT	REP	PHA	ARCSEN
1	1	1	1	85.5	67.617457
2	1	1	2	87.8	69.556412
3	1	1	3	81.4	64.451359
4	1	1	4	74.3	59.539005
5	1	1	5	94.5	76.436623
6	2	1	1	89.3	70.906630
7	2	1	2	76.8	61.205877
8	2	1	3	70	56.789089
9	2	1	4	78.5	62.375112
10	2	1	5	81.4	64.451359
11	1	2	1	16.7	24.120455
12	1	2	2	12.0	20.267901
13	1	2	3	38.1	38.115735
14	1	2	4	22.2	28.11019
15	1	2	5	18.0	25.104090
16	2	2	1	6.9	15.229055
17	2	2	2	24.8	29.867503
18	2	2	3	10.0	18.434948
19	2	2	4	14.1	22.055198
20	2	2	5	9.2	17.65682
21	1	3	1	65.5	54.029615
22	1	3	2	60.7	51.178433
23	1	3	3	38.2	38.174711
24	1	3	4	78.7	62.514815
25	1	3	5	56.9	48.966065
26	2	3	1	52.3	46.318268
27	2	3	2	68.7	55.981260
28	2	3	3	41.5	40.106090
29	2	3	4	28.1	32.011828
30	2	3	5	38.6	38.410324

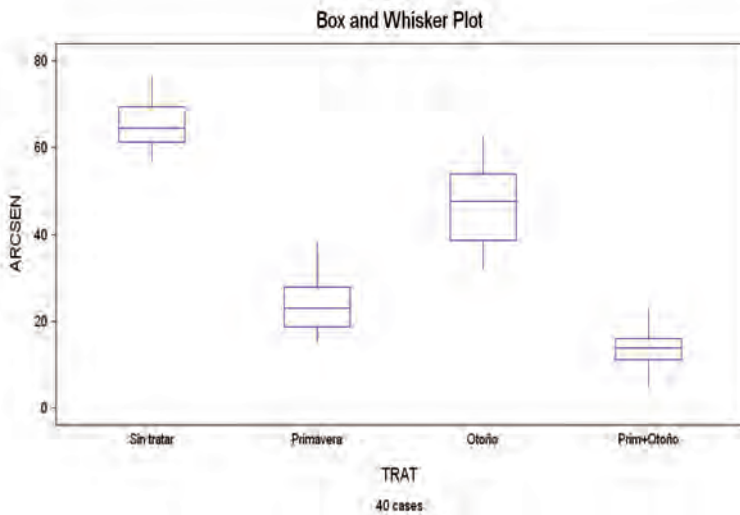
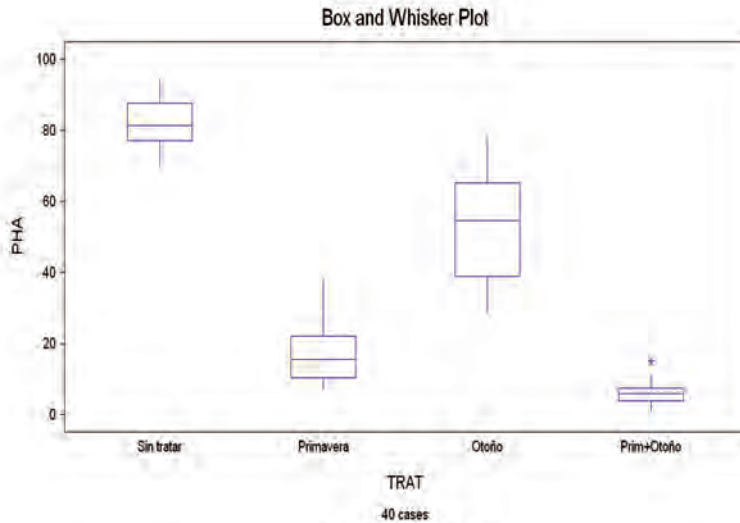
31	1	4	1	6.2	14.418226
32	1	4	2	7.7	16.110381
33	1	4	3	5.4	13.437174
34	1	4	4	3.6	10.937416
35	1	4	5	15	22.786498
36	2	4	1	0.7	4.7993192
37	2	4	2	6.3	14.536577
38	2	4	3	11.2	19.552107
39	2	4	4	2.9	9.8049053
40	2	4	5	4.5	12.247324

Una vez introducidos los datos se visualizan. Para ello, se elige la siguiente secuencia:

Statistics>Summary Statistics>Error Bar Chart



Se obtiene así una visualización bastante clara de los datos: las varianzas parecen más homogéneas en los datos transformados y también parece que no existe interacción entre los dos factores SUELO y TRAT. Para ver mejor la variabilidad de los datos, se puede elegir la opción "Box and Whisker Plot" del mismo menú (*Summary Statistics*) y se obtienen los siguientes gráficos:



Se comprueba visualmente que las varianzas de la variable TRAT son más homogéneas para los datos transformados. Además, existe un "outlier" en el gráfico de datos originales que no se presenta con los datos transformados. Los mismos gráficos se pueden obtener para las otras variables, SUELO y REP, pero en este caso no hay grandes diferencias y las varianzas de estas medias parecen homogéneas.

Para comprobar la homogeneidad de las varianzas, se utiliza el procedimiento habitual mediante un ANOVA completamente aleatorio de la variable que se quiere comprobar. En este caso, se hace con la variable TRAT, siguiendo la secuencia:

Statistics>Linear Models>Analysis of Variance>Completely Randomized Design...

y obtenemos una tabla ANOVA con los test de homogeneidad de la varianza:

Completely Randomized AOV for PHA

Source	DF	SS	MS	F	P
TRAT	3	35782.7	11927.6	114	0.0000
Error	36	3750.5	104.2		
Total	39	39533.1			

Grand Mean 39.605 CV 25.77

	Chi-Sq	DF	P
Bartlett's Test of Equal Variances	14.7	3	0.0021
Cochran's Q	0.6177		
Largest Var / Smallest Var	14.860		

Component of variance for between groups 1182.34
Effective cell size 10.0

TRAT	Mean
Sin tratar	81.950
Primavera	17.200
Otoño	52.920
Prim+Otoño	6.350
Observations per Mean	10
Standard Error of a Mean	3.2277
Std Error (Diff of 2 Means)	4.5646

Completely Randomized AOV for ARCSIN

Source	DF	SS	MS	F	P
TRAT	3	16043.5	5347.85	109	0.0000
Error	36	1772.5	49.24		
Total	39	17816.1			

Grand Mean 37.465 CV 18.73

	Chi-Sq	DF	P
Bartlett's Test of Equal Variances	4.02	3	0.2591
Cochran's Q	0.4603		
Largest Var / Smallest Var	3.5819		

Component of variance for between groups 529.861
Effective cell size 10.0

TRAT	Mean
Sin tratar	65.333
Primavera	23.896
Otoño	46.769
Prim+Otoño	13.863
Observations per Mean	10
Standard Error of a Mean	2.2189
Std Error (Diff of 2 Means)	3.1381

El *test de Bartlett* indica que las varianzas de los datos originales (PHA) no son homogéneas ($P < 0.05$). En cambio, los datos transformados (ARCSEN) sí presentan varianzas homogéneas ($P > 0.05$), luego es necesario realizar el análisis con los datos transformados que no violan este importante supuesto del ANOVA. No obstante, a efectos de comparación, se va a continuar el análisis con los dos tipos de datos. Para ello, se elige la siguiente secuencia:

Statistics > Linear Models > Analysis of Variance > Strip-Plot Design...

En la pantalla resultante introducimos el nombre de las variables dependientes (PHA, ARCSEN), las repeticiones o bloques (REP), el factor horizontal (SUELO), el factor vertical (TRAT) y se pulsa OK. Se obtiene así la tabla del ANOVA:

<u>Analysis of Variance Table for PHA</u>					
Source	DF	SS	MS	F	P
REP	4	184.2	46.1		
SUELO	1	582.2	582.2	7.26	0.0544
Error REP*SUELO	4	320.6	80.2		
TRAT	3	35782.7	11927.6	132.26	0.0000
Error REP*TRAT	12	1082.2	90.2		
SUELO*TRAT	3	186.2	62.1	0.53	0.6676
Error REP*SUELO*TRAT	12	1395.0	116.3		
Total	39	39533.1			
Grand Mean 39.605					
CV (REP*SUELO) 22.61					
CV (REP*TRAT) 23.98					
CV (REP*SUELO*TRAT) 27.22					
<u>Analysis of Variance Table for ARCSEN</u>					
Source	DF	SS	MS	F	P
REP	4	84.1	21.01		
SUELO	1	320.0	319.98	8.42	0.0440
Error REP*SUELO	4	152.0	37.99		
TRAT	3	16043.5	5347.85	120.80	0.0000
Error REP*TRAT	12	531.2	44.27		
SUELO*TRAT	3	38.1	12.69	0.24	0.8700
Error REP*SUELO*TRAT	12	647.2	53.93		
Total	39	17816.1			
Grand Mean 37.465					
CV (REP*SUELO) 16.45					
CV (REP*TRAT) 17.76					
CV (REP*SUELO*TRAT) 19.60					

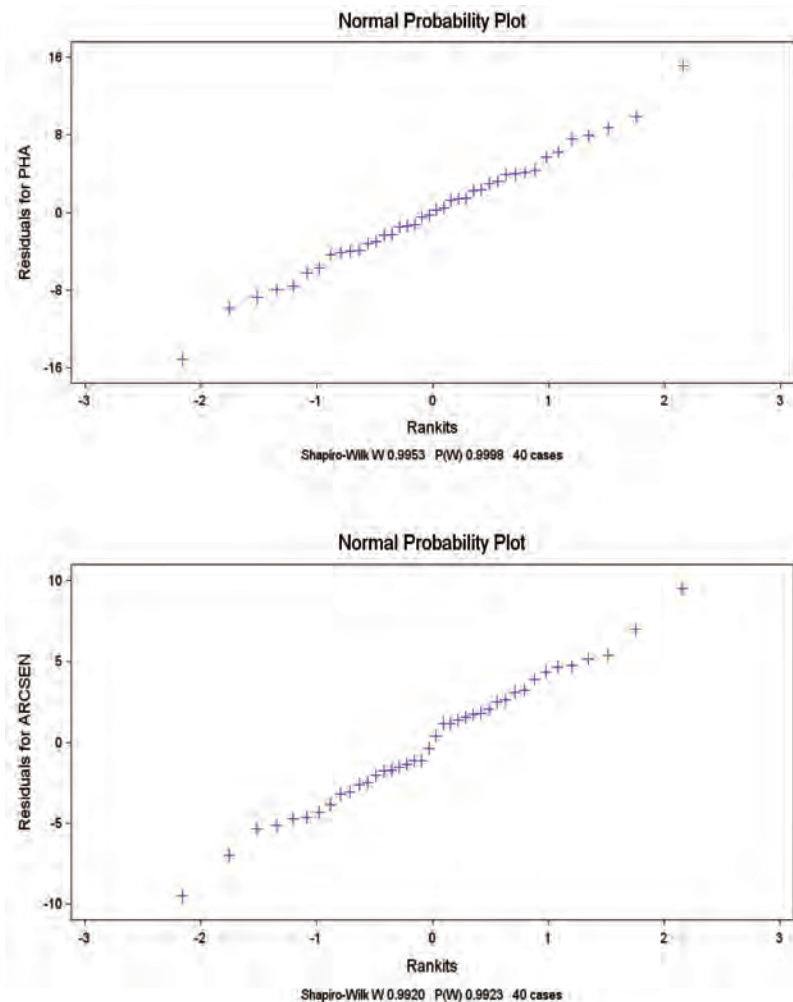
Los resultados del análisis de los datos transformados (ARCSEN) coinciden casi totalmente con el cálculo manual realizado anteriormente, llegándose, por tanto, a las mismas conclusiones, aunque existen pequeñas diferencias numéricas que se deben a errores de redondeo. Obsérvese que debajo de la tabla se indica la media general del experimento ($\bar{Y}_{...} = 37.46$) y tres coeficientes de variación (CV), correspondientes a los tres errores. El primer CV da el

grado de precisión asociado al factor horizontal (SUELO), el segundo se refiere al factor vertical (TRAT) y el tercero a la interacción SUELO*TRAT. Compruébese que el CV es menor con los datos transformados que con los originales, lo que indica mayor precisión al transformar los datos.

Comparando con el análisis de los datos originales se observa una gran diferencia, ya que con los datos originales no se obtiene un efecto significativo ($P=0.0544$) para el factor SUELO, mientras que con los datos transformados sí se obtiene ($P=0.0440$). Se llegaría, pues, a una conclusión completamente distinta si se hubiesen utilizado los datos sin transformar, o sin comprobar la homogeneidad de las varianzas.

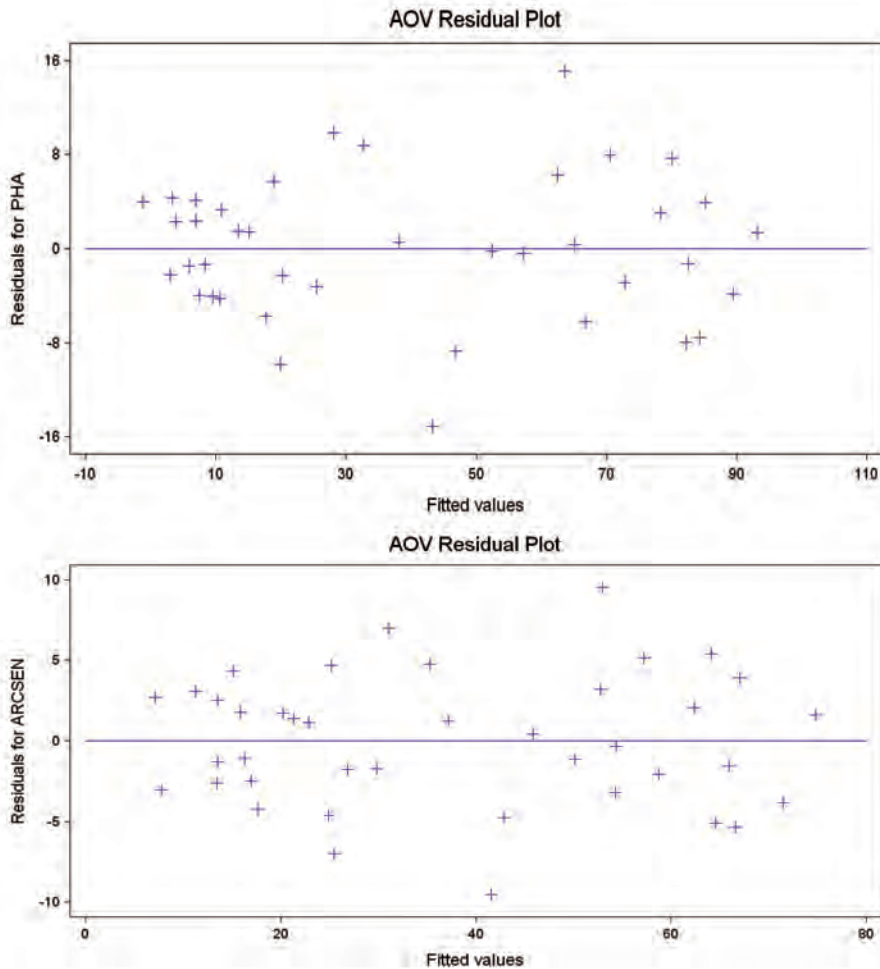
Se continúa el análisis comparativo comprobando si se cumplen otros preceptos del ANOVA. La normalidad de los datos se comprueba eligiendo la siguiente secuencia en la pantalla de resultados:

Results>Plots>Normal Probability Plot



El *test Shapiro-Wilk* muestra que, en ambos casos (originales y transformados), los datos se ajustan bien a una normal ($P > 0.05$). Asimismo, se puede ver la distribución que presentan los residuos en la misma pantalla de resultados siguiendo la secuencia:

Results>Plots>Resids By Fitted Values



Tanto en los datos originales como en los transformados, la distribución de residuos es un ejemplo de distribución aleatoria sin ningún patrón específico, lo que confirma la independencia de los errores.

Para concluir el análisis habría que realizar la comparación de medias del factor TRAT, ya que el otro factor significativo (SUELO) tiene sólo dos niveles y la comparación entre ellos, obviamente, se concluye con el ANOVA: la enfermedad tuvo una mayor incidencia en la cubierta

vegetal (43.4%) que en el suelo desnudo (35.8%). Para comparar las medias del factor TRAT se utilizará el test de comparación múltiple de *Tukey*, eligiendo la opción "*Multiple Comparisons>All-pairwise Comparisons*" del menú de resultados, y obtendríamos:

Tukey HSD All-Pairwise Comparisons Test of PHA for TRAT

TRAT	Mean	Homogeneous Groups
Sin tratar	81.950	A
Otoño	52.920	B
Primavera	17.200	C
Prim+Otoño	6.350	C

Alpha 0.05 Standard Error for Comparison 4.2469
Critical Q Value 4.199 Critical Value for Comparison 12.611
Error term used: REP*TRAT, 12 DF
There are 3 groups (A, B, etc.) in which the means are not significantly different from one another.

Tukey HSD All-Pairwise Comparisons Test of ARCSIN for TRAT

TRAT	Mean	Homogeneous Groups
Sin tratar	65.333	A
Otoño	46.769	B
Primavera	23.896	C
Prim+Otoño	13.863	D

Alpha 0.05 Standard Error for Comparison 2.9755
Critical Q Value 4.199 Critical Value for Comparison 8.8356
Error term used: REP*TRAT, 12 DF
All 4 means are significantly different from one another.

También con este test se llega a una conclusión diferente según se utilicen los datos originales o transformados. Con los datos originales sólo hay tres grupos de medias, no existiendo diferencias significativas entre el tratamiento de Primavera y el de Primavera + Otoño, mientras que con los datos transformados los cuatro tratamientos difieren entre sí. Obviamente la conclusión válida es la segunda, la que corresponde a los datos transformados.

Otro test de comparación válido para este ejemplo podría ser el de contrastes ortogonales. Se supone que el interés principal del investigador fue comparar los tratamientos fungicidas frente al testigo sin tratar y, además, los tratamientos individuales (primavera u otoño) frente al doble tratamiento (primavera + otoño), así como los tratamientos individuales entre sí (primavera frente a otoño). Para realizar estas comparaciones ortogonales, se elige la opción "*Contrast...*" del menú de resultados. En la pantalla que aparece se selecciona el factor TRAT, y los coeficientes del contraste $3 -1 -1 -1$; $0 1 1 -2$; y $0 1 -1 0$. En el primer contraste se compara el testigo sin tratar frente a los tres tratamientos, ya que el orden de los tratamientos en la matriz de datos del experimento es: Sin tratar, Primavera, Otoño y Primavera + Otoño. En el segundo contraste no se considera el testigo sin tratar, al que se asigna el coeficiente 0, y se comparan los tratamientos individuales frente al doble tratamiento. En el tercer contraste se elimina (coeficiente 0) el testigo sin tratar y el tratamiento doble y se compara Primavera frente a Otoño. De esta forma se obtiene:

AOV Contrasts of PHA by TRAT

Contrast Number 1

Contrast Coefficients: 3 -1 -1 -1

Contrast	169.38	SS (Contrast)	23908
Scheffe's F	88.37	P (Scheffe's F)	0.0000
T-Statistic	16.28	P (T-Statistic)	0.0000
SE (Contrast)	10.403		

Contrast Number 2

Contrast Coefficients: 0 1 1 -2

Contrast	57.420	SS (Contrast)	5495.1
Scheffe's F	20.31	P (Scheffe's F)	0.0001
T-Statistic	7.81	P (T-Statistic)	0.0000
SE (Contrast)	7.3558		

Contrast Number 3

Contrast Coefficients: 0 1 -1 0

Contrast	-35.720	SS (Contrast)	6379.6
Scheffe's F	23.58	P (Scheffe's F)	0.0000
T-Statistic	-8.41	P (T-Statistic)	0.0000
SE (Contrast)	4.2469		

Error term used: REP*TRAT, 12 DF

AOV Contrasts of ARCSEN by TRAT

Contrast Number 1

Contrast Coefficients: 3 -1 -1 -1

Contrast	111.47	SS (Contrast)	10355
Scheffe's F	77.97	P (Scheffe's F)	0.0000
T-Statistic	15.29	P (T-Statistic)	0.0000
SE (Contrast)	7.2885		

Contrast Number 2

Contrast Coefficients: 0 1 1 -2

Contrast	42.939	SS (Contrast)	3073.0
Scheffe's F	23.14	P (Scheffe's F)	0.0000
T-Statistic	8.33	P (T-Statistic)	0.0000
SE (Contrast)	5.1538		

Contrast Number 3

Contrast Coefficients: 0 1 -1 0

Contrast	-22.873	SS (Contrast)	2615.9
Scheffe's F	19.70	P (Scheffe's F)	0.0001
T-Statistic	-7.69	P (T-Statistic)	0.0000
SE (Contrast)	2.9755		

Error term used: REP*TRAT, 12 DF

Los tres contrastes resultan significativos, con lo que se llega prácticamente a la misma conclusión que en la comparación múltiple. Obsérvese que en este caso los resultados para los dos tipos de datos, originales o transformados, son similares para el test *t de Student* (T-Statistic), que es el recomendado para las comparaciones *a priori*. Si los contrastes se hubieran establecido a la vista de los resultados se debería utilizar los resultados del otro test (Scheffe's F), que también resulta significativo en todos los contrastes. En estas comparaciones, sin embargo, no existen diferencias notables entre los datos originales y los datos transformados.

Por último, conviene recordar que cuando se usan transformaciones de los datos, los análisis se realizan con los datos transformados, pero la presentación de los mismos hay que realizarla con los datos sin transformar. En la Tabla 12.6 se incluye, a modo de ejemplo, una forma de presentar los datos de este experimento (véase también el Capítulo 23).

Tabla 12.6. Efecto del sistema de manejo del suelo y de tratamientos con un fungicida protector sobre el porcentaje de hojas de olivo afectadas por una enfermedad.

Tratamiento	Manejo del suelo		Media
	Cubierta vegetal	Desnudo	
Sin tratar	84.7	79.2	82.0a
Primavera	21.4	13.0	17.2c
Otoño	60.0	45.8	52.9b
Primavera + Otoño	7.6	5.1	6.4d
Media	43.4A	35.8B	39.6

Como se puede observar en la Tabla 12.6, sólo se comparan las medias generales de los dos factores (Manejo del suelo y Tratamiento), ya que la interacción entre ellos no resultó significativa. Además, se utilizan letras distintas para la comparación de los niveles de cada factor, ya que se trata de comparaciones diferentes.

CAPÍTULO 13

ANÁLISIS DE LA VARIANZA COMBINADO

Hasta ahora todos los ANOVA realizados se han referido a un único experimento, pero frecuentemente el investigador trabaja con experimentos relacionados o en serie. Por ejemplo, un experimento que se repite varias veces en distintas localidades, tipos de suelo, condiciones ambientales, años, etc. Obviamente, el análisis combinado de estos datos ofrecerá una información más completa y, por supuesto, con mayor valor predictivo que el análisis por separado de los diferentes experimentos.

La limitación principal para poder realizar estos análisis combinados es que se deben cumplir todos los supuestos de un ANOVA y, sobre todo, la homogeneidad de las varianzas. Por ello, antes de proceder a realizar un análisis combinado es necesario comprobar que las varianzas de los errores de los diferentes experimentos que se quieren combinar son homogéneas. Si no ocurre así, habrá que realizar las transformaciones o particiones oportunas de los datos. Un factor añadido de estos análisis combinados es que los cálculos son más complejos, pero esto no representa una limitación con los programas estadísticos disponibles actualmente.

El procedimiento general del análisis en estos casos sigue la estructura general de los modelos jerarquizados, como el de parcelas divididas o el de bloques divididos, para la partición de la suma de cuadrados y de los grados de libertad. Las diferentes repeticiones del experimento serán los niveles del nuevo factor, de efectos fijos o aleatorios, que se introducirá en el análisis combinado.

Las dos situaciones más comunes de estos análisis combinados son los experimentos que implican varias observaciones repetidas sobre la misma unidad experimental (muestreos, mediciones temporales) y los experimentos repetidos en varias localidades, estaciones o años. A continuación se desarrollan varios ejemplos de análisis combinado para ambas situaciones, aunque para simplificar la exposición se ha prescindido de los cálculos manuales y de los pormenores del análisis que ya se han indicado en los ANOVA individuales. Una información más detallada de los ANOVA combinados y de su cálculo manual se puede encontrar en los textos de Gomez y Gomez (1984) y Petersen (1994).

OBSERVACIONES MÚLTIPLES

Cuando la variable dependiente se mide más de una vez en cada unidad experimental se habla de observaciones o datos múltiples. Existen dos clases principales de datos múltiples:

- Datos procedentes de un número determinado (s) de muestras tomadas en cada unidad experimental, tales como número, peso o contenido en nutrientes de las plantas en cada una de las muestras.

- Datos de mediciones temporales en las que un mismo carácter se mide periódicamente en indiferentes momentos o estados de desarrollo, como altura de las plantas, gravedad de infecciones, población de insectos o patógenos, etc.

El ANOVA estándar, que hemos visto hasta ahora, sería aplicable sólo para la media de todas las muestras o mediciones temporales realizadas en cada unidad experimental.

Análisis de muestreos

En el análisis de los datos procedentes de muestreos tenemos una fuente de variación adicional, la debida a la variación entre muestras, que es referida como error de muestreo. El formato del ANOVA es similar a los experimentos sin muestreos, por lo que se pueden utilizar cualquiera de los diseños experimentales explicados, con la diferencia de que al final de la tabla ANOVA se incluyen los valores correspondientes al error de muestreo. Así, para un experimento con dos factores (A, B) en un diseño en “split-plot” con r bloques al azar y s muestras de cada unidad experimental, la Tabla ANOVA correspondiente sería:

Fuente de Variación	gl	SC	CM	F
Bloque	$r-1$	SCR		
Factor A	$a-1$	SCA	$SCA / (a-1)$	CMA/CME_A
Error (a)	$(r-1)(a-1)$	SCE_A	$SCE_A / (r-1)(a-1)$	
Factor B	$b-1$	SCB	$SCB / (b-1)$	CMB/CME_E
A×B	$(a-1)(b-1)$	SCAB	$SCAB / (a-1)(b-1)$	$CMAB/CME_E$
Error (ab)	$(r-1)(a-1)(b-1)$	SCE_E	$SCE_E / (r-1)(a-1)(b-1)$	CME_E/CME_S
Error de muestreo	$rab(s-1)$	SCE_S	$SCE_S / rab(s-1)$	
Total	$rabs-1$	SCT		

El ANOVA y el cálculo de las sumas de cuadrados sería similar a lo indicado para el diseño en parcelas divididas (véase el Capítulo 11), pero se ha añadido un error adicional, el error de muestreo. La suma de cuadrados de este nuevo error (SCE_S) se calcularía como:

$$SCE_S = SCT \text{ (Total)} - (SCR + SCA + SCE_A + SCB + SCAB + SCE_E)$$

El cuadrado medio del error de muestreo (CME_S) es un estimador de σ^2 , mientras que el cuadrado medio del error experimental (ab) (CME_E) es un estimador de $\sigma^2 + s\sigma_{E'}^2$, siendo s el número de muestras tomadas. Por ello, podemos calcular la varianza del error de muestreo (s_s^2) y la del error experimental (s_E^2) y sus correspondientes coeficientes de variación (CV) como:

$$s_S^2 = CME_S; \quad CV(S) = \frac{\sqrt{s_S^2}}{\bar{Y}} \times 100$$

$$s_E^2 = \frac{CME_E - CME_S}{s}; \quad CV(E) = \frac{\sqrt{s_E^2}}{\bar{Y}} \times 100$$

con lo que tendríamos una indicación de la importancia relativa del error de muestreo respecto al error experimental. No obstante, si el cociente de varianzas CME_E/CME_S resultara en un valor de F no significativo ($P > 0,05$), concluiríamos que la variación entre unidades o parcelas experimentales no difiere de la que existe entre las muestras tomadas en una misma parcela, por lo que pueden combinarse los dos errores para obtener una nueva estimación de la varianza apropiada para probar las parcelas experimentales. La nueva varianza del único error (s^2) sería:

$$s^2 = \frac{SCE_E + SCE_S}{GL_E + GL_S}$$

Ejemplo de un ANOVA con muestreo

Se trata de un experimento de campo sobre efecto de la solarización del suelo sobre el número de propágulos del hongo fitopatógeno *Verticillium dahliae*. Se realizaron tres tratamientos del suelo: sin solarizar, solarizado una vez y solarizado dos veces, en un diseño en bloques al azar con 4 repeticiones. Después de la segunda solarización, se tomaron 6 muestras de suelo de cada parcela elemental, que se analizaron en el laboratorio calculándose el número de propágulos por gramos de suelo. Los datos obtenidos se presentan en la Tabla 13.1.

Tabla 13.1. Número de propágulos de *Verticillium dahliae* por g de suelo (año 1)

TRAT	BLOQUE (REP)	MUESTRA						Media
		1	2	3	4	5	6	
Sin solarizar	1	18,4	19,8	23,6	60,2	34,5	15,7	28,7
	2	28,3	65,4	23,1	34,5	9,6	16,4	29,6
	3	13,4	12,0	7,1	15,0	9,6	45,3	17,1
	4	16,5	8,1	12,7	15,8	6,6	33,2	15,5
Solarizado 1	1	0,6	1,5	1,1	6,2	0,5	0,2	1,7
	2	0,7	2,7	1,2	4,4	7,3	0,5	2,8
	3	7,0	3,6	0,8	0,5	1,2	0,9	2,3
	4	2,3	0,8	3,7	1,9	6,5	1,5	2,8
Solarizado 2	1	2,8	1,2	0,6	0,9	0	0,3	1,0
	2	0	0,2	0,4	2,6	0,3	0,5	0,7
	3	1,0	2,7	1,8	0,5	0	0,8	1,1
	4	0,5	1,6	0,4	0,2	0,2	0	0,5

El programa *Statistix* no incluye un apartado específico para el ANOVA de experimentos con muestreos. Para los casos como éste, *Statistix* dispone de la opción “*General AOV/AOCV...*” en el menú de “*Analysis of Variance*” que permite cualquier formato de ANOVA si se indican los factores, interacciones y errores que deben ser considerados. Para realizar el análisis seguimos la siguiente secuencia:

Statistics>Linear Models>Analysis of Variance>General AOV/AOCV...

y aparece una pantalla donde introducimos la variable o variables dependientes, el modelo ANOVA y las covariables (para análisis de covarianza). El modelo ANOVA se define indicando ordenadamente los factores, las interacciones y el error o errores. En el manual del programa se explica la forma de definir el modelo para diferentes diseños experimentales. Para este ejemplo, denominando TRAT a los tratamientos de solarización, REP a los bloques, MUES a las muestras, e indicando con (E) el término o términos de error, el modelo es:

REP TRAT REP*TRAT(E),

o bien, REP TRAT REP*TRAT(E) REP*TRAT*MUES(E).

Introduciendo cualquiera de estos dos modelos y pulsando OK se obtiene el siguiente ANOVA para los datos normalizados mediante la transformación logarítmica (LOPROP):

Analysis of Variance Table for LOPROP

Source	DF	SS	MS	F	P
TRAT	2	15.4419	7.72097	92.08	0.0000
REP	3	0.0727	0.02422	0.29	0.8322
Error TRAT*REP	6	0.5031	0.08385		
Error TRAT*REP*MUES	60	3.3539	0.05590		
Total	71	19.3716			

Grand Mean 2.6547

CV (TRAT*REP) 10.91

CV (TRAT*REP*MUES) 8.91

La tabla muestra dos errores, el primero (TRAT*REP) es el error experimental, con un coeficiente de variación del 10.91%, mientras que el segundo (TRAT*REP*MUES) es el error de muestreo, con su coeficiente de variación del 8.91%. El error experimental es el que se usa para ver la significación de los tratamientos (TRAT) y los bloques (REP), resultando muy significativos los tratamientos y no significativos los bloques. El error de muestreo sirve para estimar la variación debida a las muestras, siendo ésta (8.91%) muy parecida a la ocasionada conjuntamente por los bloques y el muestreo (10.91%). Como el cuadrado medio del error de muestreo (CME_S) es un estimador de σ^2 y el cuadrado medio del error experimental (CME_E) lo es de $\sigma^2 + s\sigma_E^2$, se puede calcular s_E^2 y s_S^2 y sus coeficientes de variación (CV) como:

$$s_E^2 = \frac{CME_E - CME_S}{s} = \frac{0,08385 - 0,05590}{6} = 0,00466; CV = \frac{\sqrt{0,00466}}{2,6547} = 2,6\%$$

$$s_S^2 = 0,05590; CV = \frac{\sqrt{0,0559}}{2,6547} = 8,9\%$$

Estos resultados indican un error debido a los bloques relativamente pequeño en comparación con el error de muestreo. Es decir, para aumentar la precisión del experimento se debería incrementar el número de muestras o modificar la técnica de análisis del suelo para hacerla más eficiente, ya que es la fuente de variación más importante. Una forma común de reducir el error de muestreo sin necesidad de cambiar la técnica de análisis es eliminar el valor más alto y el más bajo de las diferentes muestras tomadas en cada parcela experimental. Obviamente, para ello es necesario disponer de un número elevado de muestras. Aplicando la prueba *F* al cociente entre los cuadrados medios de los dos errores se obtiene:

$$\frac{CME_E}{CME_S} = \frac{0,08385}{0,0559} = 1,5 \quad \text{con 6 y 60 grados de libertad, se obtiene } P = 0,19368$$

El valor de *F* no es significativo, luego se puede admitir que las dos varianzas no difieren entre sí, o que la variación debida a los bloques es insignificante comparada con la debida al muestreo, y combinar las dos varianzas en una varianza común (*s*²):

$$s^2 = CME = \frac{SCE_E + SCE_S}{GL_E + GL_S} = \frac{0,5031 + 3,3539}{6 + 60} = \frac{3,857}{66} = 0,05844$$

Se tiene un único error, cuyo cuadrado medio (*CME*) servirá para todas las comparaciones posibles de medias, entre tratamientos o entre muestras. El nuevo ANOVA, resultante del modelo en bloques al azar TRAT REP, sería:

<u>Analysis of Variance Table for LOPROP</u>						
Source	DF	SS	MS	F	P	
TRAT	2	15.4419	7.72097	132.12	0.0000	
REP	3	0.0727	0.02422	0.41	0.7432	
Error	66	3.8570	0.05844			
Total	71	19.3716				
Grand Mean		2.6547	CV		9.11	

En este caso, se llega a la misma conclusión que en el ANOVA anterior con los dos errores: los tratamientos son muy significativos, pero los bloques no lo son, aunque hay un único error con 66 grados de libertad y un coeficiente de variación del 9,11%.

Para la comparación de las medias de los tratamientos se pueden aplicar los contrastes ortogonales. Utilizando los mismos coeficientes que en el ejemplo similar del Capítulo 10, se obtiene:

AOV Contrasts of LOPROP by TRAT

Contrast Number 1

Contrast Coefficients: 2 -1 -1 (No solarizado vs. Solarizado 1 + Solarizado 2)

Contrast	1.9232	SS (Contrast)	14.795
Scheffe's F	126.59	P (Scheffe's F)	0.0000
T-Statistic	15.91	P (T-Statistic)	0.0000
SE (Contrast)	0.1209		

Contrast Number 2

Contrast Coefficients: 0 1 -1 (Solarizado 1 vs. Solarizado 2)

Contrast	0.2321	SS (Contrast)	0.6466
Scheffe's F	5.53	P (Scheffe's F)	0.0060
T-Statistic	3.33	P (T-Statistic)	0.0014
SE (Contrast)	0.0698		

Error term used: Error, 66 DF

Los resultados indican que los dos tratamientos de solarización reducen significativamente los propágulos del patógeno en el suelo respecto al testigo no solarizado ($\bar{X}_0 = 22,70$) y que la doble solarización ($\bar{X}_2 = 0,77$) reduce significativamente la población del patógeno respecto al suelo solarizado una sola vez ($\bar{X}_1 = 2,40$).

Análisis de mediciones temporales repetidas

Algunos experimentos requieren observaciones repetidas de un mismo carácter en el tiempo. En estos casos es importante determinar el efecto de la interacción entre un tratamiento y el momento de la observación, lo que no puede hacerse si el ANOVA se realiza separadamente para cada observación. El ANOVA combinado se lleva a cabo considerando el tiempo de observación como un factor adicional y tratándolo como una subparcela o la unidad experimental más pequeña. Por tanto, el formato de un ANOVA combinado de medidas repetidas en el tiempo para un diseño en bloques completos al azar es idéntico a un diseño en "split-plot" con los tratamientos como parcela principal y el tiempo de observación como subparcela. Asimismo, el ANOVA combinado de medidas repetidas en el tiempo para un diseño en "split-plot" es similar al que corresponde a un diseño en "split-split-plot" con el tiempo de observación como sub-subparcela.

Ejemplo de ANOVA de observaciones temporales repetidas

Para ilustrar este tipo de análisis se utilizará el mismo ejemplo del apartado anterior, en el que vamos a añadir una segunda observación un año después. Los datos correspondientes a esta segunda observación se indican a continuación:

Tabla 13.2. Número de propágulos de *Verticillium dahliae* por g de suelo (año 2)

TRAT	BLOQUE (REP)	MUESTRA						Media
		1	2	3	4	5	6	
Sin solarizar	1	12,7	9,5	25,2	58,4	33,9	17,3	26,2
	2	13,3	33,6	14,7	16,9	8,2	7,5	15,7
	3	14,5	22,8	7,8	16,1	30,2	59,5	25,1
	4	17,6	8,7	31,8	16,9	7,1	15,5	16,3
Solarizado 1	1	0,6	1,3	1,4	5,9	2,8	0,2	2,0
	2	2,9	3,4	1,5	5,5	7,1	0,6	3,5
	3	7,7	2,0	1,0	0,6	1,5	1,1	2,3
	4	2,8	1,0	4,6	2,9	8,1	1,9	3,5
Solarizado 2	1	7,1	2,4	2,2	3,1	0	1,0	2,6
	2	0,3	0,7	1,4	5,3	1,0	1,8	1,7
	3	3,4	9,2	6,2	1,7	0	2,7	3,9
	4	1,7	6,4	1,3	2,1	0,7	0	2,0

En primer lugar se analizan separadamente las dos observaciones utilizando los datos normalizados mediante la transformación logarítmica. Las tablas ANOVA correspondientes a estos dos análisis se indican a continuación:

<u>Analysis of Variance Table for LOPROP (T=1)</u>					
Source	DF	SS	MS	F	P
TRAT	2	15.4419	7.72097	92.08	0.0000
REP	3	0.0727	0.02422	0.29	0.8322
Error TRAT*REP	6	0.5031	0.08385		
Error	60	3.3539	0.05590		
Total	71	19.3716			
Grand Mean 2.6547					
CV (TRAT*REP) 10.91					
CV (Error) 8.91					

Analysis of Variance Table for LOPROP (T=2)

Source	DF	SS	MS	F	P
TRAT	2	9.75527	4.87763	63.09	0.0001
REP	3	0.05174	0.01725	0.22	0.8771
Error TRAT*REP	6	0.46390	0.07732		
Error	60	4.37250	0.07288		
Total	71				

Grand Mean 2.7454
 CV (TRAT*REP) 10.13
 CV (Error) 9.83

Los dos análisis individuales indican resultados similares: existen diferencias significativas entre tratamientos, pero no entre bloques. Para comprobar si las varianzas son homogéneas realizamos el test *F* para los cuadrados medios del error experimental:

$$F = \frac{0,08385}{0,07732} = 1,08 \text{ con 6 y 6 grados de libertad, } P = 0,4640$$

El test *F* no es significativo, luego las varianzas son homogéneas y podemos realizar el ANOVA combinado de los dos experimentos. Además, si aplicamos el test *F* a las varianzas del error de muestreo, obtenemos:

$$F = \frac{0,07288}{0,05590} = 1,30 \text{ con 60 y 60 grados de libertad, } P = 0,1561$$

En este caso, el test *F* tampoco es significativo, luego las varianzas del error de muestreo no difieren entre tiempos (años). Realizando el ANOVA combinado con el modelo:

TRAT REP TRAT*REP(E) TIEMPO TRAT*TIEMPO TRAT*REP*TIEMPO(E)

se obtiene:

Analysis of Variance Table for LOPROP

Source	DF	SS	MS	F	P
TRAT	2	24.7384	12.3692	96.10	0.0000
REP	3	0.0356	0.0119	0.09	0.9616
Error TRAT*REP	6	0.7723	0.1287		
TIEMPO	1	0.2966	0.2966	9.41	0.0134
TRAT*TIEMPO	2	0.4588	0.2294	7.28	0.0132
Error TRAT*REP*TIEMPO	9	0.2835	0.0315		
Error TRAT*REP*TIEMPO*MUES	120	7.7264	0.0644		
Total	143	34.3115			

Grand Mean 2.7001
 CV (TRAT*REP) 13.29
 CV (TRAT*REP*TIEMPO) 6.57
 CV (TRAT*REP*TIEMPO*MUES) 9.40

El test F para los cuadrados medios del error experimental y del error de muestreo ($F = \frac{0,0315}{0,0644} = 0,49$) tampoco resulta significativo ($P = 0,879$), por lo que se realiza el ANOVA con el error global, utilizando el modelo: TRAT TRAT*REP(E) TIEMPO TRAT*TIEMPO. Se obtiene:

Source	DF	SS	MS	F	P
TRAT	2	24.7384	12.3692	96.10	0.0000
REP	3	0.0356	0.0119	0.09	0.9616
Error TRAT*REP	6	0.7723	0.1287		
TIEMPO	1	0.2966	0.2966	4.78	0.0307
TRAT*TIEMPO	2	0.4588	0.2294	3.69	0.0275
Error	129	8.0099	0.0621		
Total	143	34.3115			

Grand Mean 2.7001
 CV (TRAT*REP) 13.29
 CV (Error) 9.23

El resultado del análisis indica que, además de los tratamientos, el tiempo y la interacción tratamiento×tiempo resultan significativos. Tras comprobar que se cumplen los supuestos del ANOVA, pasamos a la comparación de medias mediante contrastes ortogonales y considerando la interacción significativa. Utilizando los mismos coeficientes que en el ejemplo anterior y considerando que cada contraste tiene un grado de libertad y que el cuadrado medio del error es $CME = 0,0621$, con 129 gl, obtendríamos dos conjuntos de contrastes ortogonales, cuya significación se indica en las dos tablas siguientes (véase el Capítulo 5):

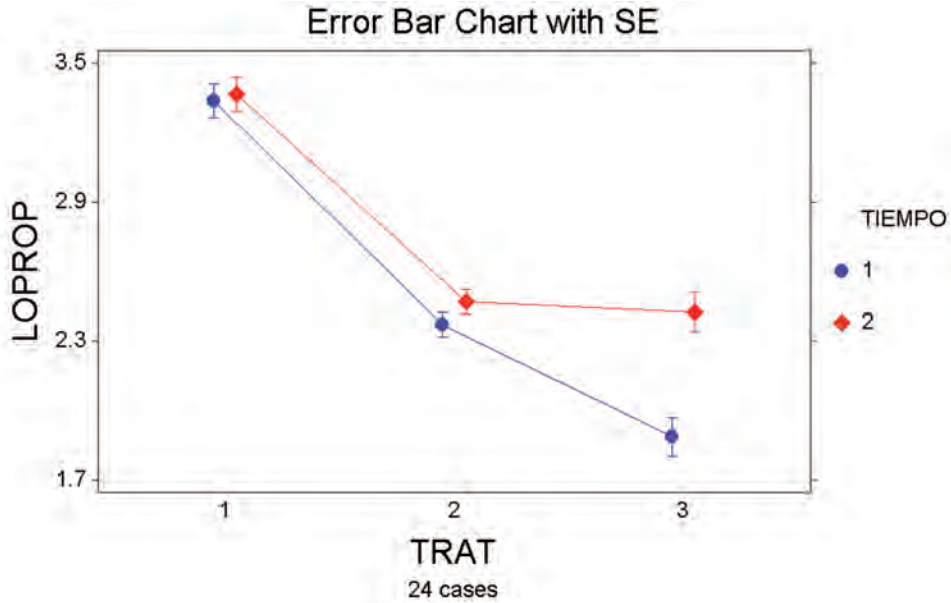
Tiempo	Contraste*	SC = CM	F	P
1	S_0 vs. ($S_1 + S_2$)	14,795	238,24	0,0000
1	S_1 vs. S_2	0,6466	10,41	0,0016
2	S_0 vs. ($S_1 + S_2$)	9,7221	156,56	0,0000
2	S_1 vs. S_2	0,0331	0,53	0,4679

* S_0 = No solarizado, S_1 = solarizado una vez, S_2 = solarizado dos veces.

Tratamiento	Contraste*	SC = CM	F	P
S_0	T_1 vs. T_2	0,0113	0,53	0,6721
S_1	T_1 vs. T_2	0,0456	0,73	0,3945
S_2	T_1 vs. T_2	0,6985	11,25	0,0011

* T_1 = Tiempo 1, T_2 = Tiempo 2.

Como indicó el ANOVA, las diferencias entre tratamientos dependían del tiempo de evaluación. Así, mientras que en la primera evaluación los tres tratamientos diferían entre sí, en la segunda evaluación no hubo diferencias entre los suelos solarizados una vez y los solarizados dos veces, debido a un incremento de la población del patógeno en estos últimos. Los testigos no solarizados tuvieron una población mayor de propágulos que los suelos solarizados, aunque su densidad de propágulos no varió significativamente entre los dos tiempos de evaluación. Estas conclusiones se aprecian claramente en el gráfico de los datos obtenido con el programa *Statistix*:



Otra forma de analizar la interacción, sin necesidad de partir el experimento por los tiempos o los tratamientos, sería descomponer la interacción (2 gl) en dos contrastes, cada uno de ellos con 1 gl:

Fuente de variación	gl	SC	CM	F	P
Tiempo (T) × TRAT (S)	2	0,4588	0,2294	3,69	0,0275
T × [S ₀ vs. (S ₁ + S ₂)]	(1)	0,2654	0,2654	4,27	0,0408
T × (S ₁ vs. S ₂)	(1)	0,1934	0,1934	3,11	0,0802

Con este nuevo análisis llegaríamos a una conclusión similar: la diferencia significativa entre los dos tratamientos solarizados y el testigo sin solarizar se mantiene durante los dos tiempos de muestreo, mientras que la diferencia entre los dos tratamientos de solarización no es similar para los dos tiempos de muestreo, es decir, en el primer tiempo la diferencia entre

los dos tratamientos es significativa, mientras que en la segunda evaluación esta diferencia no es significativa.

SERIES DE EXPERIMENTOS

Con frecuencia un mismo experimento se repite varias veces en diferentes épocas, lugares o años para estudiar la adaptación de una nueva tecnología respecto a diferentes condiciones ambientales, diferentes localidades, o su efecto en la productividad a lo largo del tiempo, o bien, para comprobar la repetitividad de los resultados obtenidos. En cualquier caso, se trata de combinar datos de una serie de experimentos para analizarlos conjuntamente. A continuación se consideran varias situaciones que incluyen diferentes épocas, años, localidades y experimentos de larga duración.

Análisis de diferentes épocas

Este tipo de análisis es aplicable para las especies vegetales que pueden cultivarse en diferentes estaciones o épocas, o en diferentes condiciones climáticas, dentro del mismo año. La fecha o época de siembra dentro del año se puede considerar como una variable de efectos fijos. El objetivo principal del análisis combinado es examinar la interacción entre época y tratamiento para determinar si es necesario recomendar una tecnología diferente para cada época de siembra. El ANOVA combinado para este caso sería:

Fuente de variación	gl	SC	CM	F
Época de cultivo (<i>A</i>)	$a-1$	SCA	$SCA / (a-1)$	CMA / CME_A
Reps. dentro de época	$a(r-1)$	SCE_A	$SCE_A / a(r-1)$	
Tratamiento (<i>B</i>)	$b-1$	SCB	$SCB / (b-1)$	CMB / CME_B
<i>A</i> × <i>B</i>	$(a-1)(b-1)$	$SC_{A \times B}$	$SC_{A \times B} / (a-1)(b-1)$	$CM_{A \times B} / CME_{AB}$
Error global	$a(r-1)(b-1)$	SCE_E	$SCE_E / a(r-1)(b-1)$	
Total	$abr-1$	SCT		

Las sumas de cuadrados se calcularían según lo indicado en el Capítulo 11 para el diseño en parcelas divididas.

El ejemplo que vamos a utilizar para este supuesto es el de un cultivo en el que se quiere comparar la adaptación de tres variedades (V_1, V_2 y V_3) en dos épocas de siembra (invierno y primavera) que dan lugar a dos tipos de cultivo con condiciones ambientales muy diferentes. La siembra de invierno es muy favorable para el desarrollo de una enfermedad, mientras que la de primavera es poco favorable. Asimismo, las tres variedades difieren por su resistencia genética a la enfermedad: V_1 es muy susceptible, V_2 moderadamente susceptible y V_3 resis-

tente. Las variedades V_1 y V_2 son de origen local, mientras que V_3 es una nueva variedad procedente de un programa de mejora genética. El objetivo del ensayo es determinar la época más adecuada para el cultivo, teniendo en cuenta la limitación de la enfermedad y la adaptación de las variedades a los dos sistemas de cultivo. Se empleó un diseño experimental en bloques completos al azar con 4 repeticiones y 50 m² de parcela elemental. En la Tabla 13.3 se incluyen los datos de producción de las parcelas.

Tabla 13.3. Rendimiento (kg/ha) de tres variedades en dos épocas de siembra

Variedad	Bloque (REP)				Media
	I	II	III	IV	
Invierno					
V_1	0	214	425	40	169,75
V_2	1252	627	716	1068	915,75
V_3	2163	2714	2521	2240	2409,50
Primavera					
V_1	1036	697	849	1258	960,00
V_2	1524	1861	2220	1744	1837,30
V_3	1312	874	695	1133	1003,50

El procedimiento a seguir para el ANOVA combinado comienza por el análisis individual de cada época de cultivo según el diseño en bloques al azar utilizado. Un resumen de estos análisis individuales se indica a continuación:

Fuente de variación	gl	CM	F	P
Invierno				
Bloque	3	6611		
Variedad	2	5202857	57,06	0,0001
Error	6	91182		
Primavera				
Bloque	3	28236		
Variedad	2	977733	10,31	0,0115
Error	6	94855		

Para comprobar si las varianzas son homogéneas, se aplica el test F , ya que sólo tenemos dos varianzas:

$$F = \frac{CME_mayor}{CME_menor} = \frac{94855}{91182} = 1,04; P = 0,4816$$

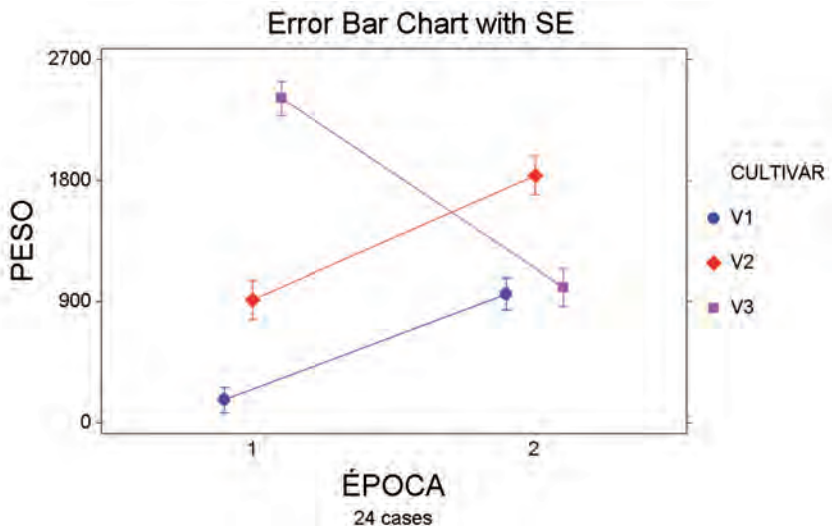
luego las varianzas son homogéneas para las diferentes épocas de siembra y se puede realizar un ANOVA combinado global. Si las varianzas fueran heterogéneas también se podría realizar un ANOVA combinado, pero en este caso habría que hacer una partición del error global (Gomez y Gomez, 1984). El ANOVA combinado para varianzas homogéneas, que se obtiene con el modelo:

EPOCA EPOCA*REP(E) VARIEDAD VARIEDAD*EPOCA

sería:

Fuente de variación	gl	SC	CM	F	P
Época (A)	1	62322	62322	3,58	0,1075
Reps. dentro época	6	104542	17424		
Variedad (V)	2	5522514	2761257	29,69	0,0000
A × V	2	6838665	3419332	36,76	0,0000
Error global	12	1116223	93019		
Total	23	1.36E+07			

Tanto el factor Variedad como la interacción Variedad×Época han resultado muy significativos. La interacción se aprecia claramente en el gráfico de datos que ofrece el programa *Statistix*:



El gráfico muestra que el orden de producción de las variedades varía drásticamente con la época de siembra. En la siembra de invierno (época 1) el orden sería $V_3 > V_2 > V_1$, mientras que en la siembra de primavera (época 2) dicho orden sería $V_2 > V_3 = V_1$.

Para analizar la interacción, se ha realizado una descomposición de las sumas de cuadrados de ambos factores en contrastes ortogonales con un grado de libertad. Para ello, se ha comparado la nueva variedad resistente a la enfermedad (V_3) frente a las dos variedades locales (V_3 vs. V_1+V_2) y éstas entre sí (V_1 vs. V_2). La nueva tabla ANOVA que resulta de esta descomposición sería:

Fuente de variación	gl	SC	CM	F	P
Época (A)	1	62322	62322	3,58	0,1075
Reps. dentro época	6	104542	17424		
Variedad (V)	2	5522514	2761257	29,69	0,0000
V_3 vs. V_1+V_2	(1)	2887573	2887573	31,04	0,0001
V_1 vs. V_2	(1)	2634941	2634941	28,33	0,0002
$A \times V$	2	6838665	3419332	36,76	0,0000
$A \times (V_3$ vs. $V_1+V_2)$	(1)	6821439	6821439	73,33	0,0000
$A \times (V_1$ vs. $V_2)$	(1)	17226	17226	0,19	0,6707
Error global	12	1116223	93019		
Total	23	1.36E+07			

Puesto que la diferencia fundamental entre las dos épocas de cultivo se debió al ataque de una enfermedad, que resultó muy grave en el cultivo de invierno y de escasa importancia en primavera, las conclusiones del ANOVA son que la variedad resistente es más productiva que las locales en invierno, pero no en primavera, donde fue superada por una de ellas (V_2). De las variedades locales V_2 , que es moderadamente susceptible a la enfermedad, resultó más productiva que la variedad susceptible (V_1) tanto en invierno como en primavera. La variedad susceptible tuvo unas producciones muy bajas, sobre todo en invierno, por lo que bajo el criterio de productividad no se recomienda su siembra. Habría que considerar otros criterios, como calidad de la semilla, para ver si interesa su cultivo, aunque para esta variedad se recomendaría únicamente el cultivo de primavera.

Análisis de diferentes años

En los estudios de adaptación de una nueva tecnología es habitual repetir el experimento en varios años para comprobar si el efecto de la nueva tecnología es significativo y estable en el tiempo. A diferencia del caso anterior, en el que las estaciones u épocas se pueden definir por unas características determinadas y predecibles, la variabilidad entre años es generalmente impredecible, por lo que los años se consideran como una variable aleatoria. La

interacción entre tratamiento y año no tiene un claro significado agronómico, por lo que es menos importante que la interacción entre tratamiento y época.

El procedimiento a seguir es parecido al del apartado anterior, con la diferencia de que la variable época se consideraba fija y el término de error era el *CME*, mientras que el año se considera una variable aleatoria y el término de error es la interacción *año × tratamiento*. La Tabla del ANOVA combinado de *a* años (*A*) para un diseño en bloques completos al azar con *t* tratamientos sería:

Fuente de variación	gl	SC*	CM	F
Año (<i>A</i>)	$a-1$	SCA	SCA / ($a-1$)	CMA / CME_A
Reps. dentro de año	$a(r-1)$	SCE_A	$SCE_A / a(r-1)$	
Tratamiento (<i>T</i>)	$t-1$	SCB	SCB / ($t-1$)	CMB / CME_E
<i>A × T</i>	$(a-1)(t-1)$	$SC_{A \times T}$	$SC_{A \times T} / (a-1)(t-1)$	$CM_{A \times T} / CM_E$
Error global	$a(r-1)(t-1)$	SCE_E	$SCE_E / a(r-1)(t-1)$	
Total	$art-1$	SCT		

* Las sumas de cuadrados se calcularían según lo indicado en el Capítulo 11 para el diseño en parcelas divididas.

Se va a ilustrar el procedimiento para combinar los datos sobre años con un experimento de tratamientos fungicidas en el cultivo del trigo. Tenemos cuatro tratamientos o momentos de aplicación de un fungicida protector (sin tratar, aplicación en invierno, aplicación en primavera y aplicación en invierno y en primavera) en un diseño en bloques completos al azar con 5 repeticiones. La variable dependiente es el porcentaje de tejido foliar afectado por una enfermedad medido al final del ciclo del cultivo. Como el intervalo de valores es muy amplio, se ha realizado la transformación angular para homogeneizar las varianzas. En la Tabla 13.4 se presentan los datos transformados de este experimento para dos años (año 1 y año 2).

Tabla 13.4. Porcentaje de tejido foliar afectado por una enfermedad del trigo según varios tratamientos y años*

Tratamiento	Bloque (REP)				
	I	II	III	IV	V
	Año 1				
Sin tratar	61,9	55,9	64,5	57,7	50,9
Invierno	20,3	16,5	15,0	8,5	23,8
Primavera	54,0	60,4	49,7	70,4	49,0
Primavera+Otoño	14,4	16,1	13,4	10,9	22,8

Tratamiento	Bloque (REP)				
	I	II	III	IV	V
Año 2					
Sin tratar	70,9	61,2	56,8	62,4	64,5
Invierno	15,2	29,9	18,4	22,1	17,7
Primavera	46,3	32,0	40,1	56,0	38,4
Invierno+Primavera	4,8	14,5	19,6	9,8	12,2

* Valores transformados mediante $Y_T = \arcsen \sqrt{Y/100}$, siendo Y el porcentaje de hojas afectadas por la enfermedad (veáse la Tabla A9).

Lo primero es realizar los análisis individuales para cada año y comprobar si las varianzas de los errores de los dos experimentos son homogéneas. Un resumen de estos análisis individuales se indica a continuación:

Fuente de variación	gl	CM	F	P
Año 1				
Bloque	4	8,84		
Tratamiento	3	2838,44	53,83	0,0000
Error	12	52,73		
Año 2				
Bloque	4	11,60		
Tratamiento	3	2625,38	49,28	0,0000
Error	12	53,27		

Para comprobar si las varianzas son homogéneas, aplicamos el test F , ya que sólo tenemos dos varianzas:

$$F = \frac{CME_{mayor}}{CME_{menor}} = \frac{53,27}{52,73} = 1,01; P = 0,4221$$

luego las varianzas son homogéneas para los dos años y podemos realizar un ANOVA combinado con un único error global. Si las varianzas fueran heterogéneas también podríamos realizar un ANOVA combinado, pero en este caso habría que realizar una partición del error global (Gomez y Gomez, 1984). El ANOVA combinado para los dos años, considerando las varianzas homogéneas, sería:

Fuente de variación	gl	SC	CM	F	P
Año (A)	1	47,4	47,38	6,86	0,0307
Reps. dentro año	8	55,2	6,90		
Tratamiento (T)	3	15813,8	5271,28	27,38	0,0111
A × T	3	577,6	192,54	3,63	0,0272
Error global	24	1272,0	53,00		
Total	39	17766,1			

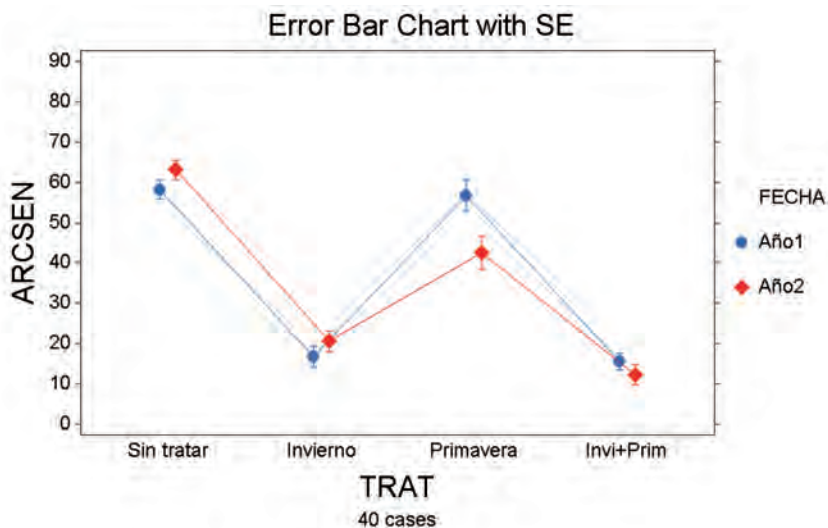
Puesto que la interacción año × tratamiento ha resultado significativa ($CM(A \times T) / CME = 3,63$), podemos calcular el valor de la varianza debida a dicha interacción (σ_{at}^2) como:

$$\sigma_{at}^2 = \frac{CMA \times T - CME}{r} = \frac{192,54 - 53,00}{5} = 27,91$$

luego, la varianza debida a la interacción año × tratamiento respecto a la varianza total sería:

$$\frac{\sigma_{at}^2}{\sigma_{at}^2 + \sigma^2} \times 100 = \frac{27,91}{27,91 + 53,00} \times 100 = 34,5\%$$

Esta proporción es relativamente grande, por lo que cabe esperar que el efecto de los tratamientos no sea muy estable y varíe con los años. La interacción se puede observar en el gráfico de datos del programa *Statistix*:



Experimentación en agricultura

Al ser significativa la interacción, ésta es el término de error para determinar la significación de los tratamientos. Para examinar más cuidadosamente el efecto de los tratamientos y su variabilidad interanual, se puede descomponer la suma de cuadrados de los tratamientos en tres contrastes ortogonales, según los coeficientes siguientes:

Contraste	gl	Tratamiento			
		Sin tratar	Invierno	Primavera	Invierno+Primavera
I	1	-1	1	-1	1
P	1	-1	-1	1	1
I × P	1	1	-1	1	-1

La nueva Tabla ANOVA con la descomposición de los tratamientos sería:

Fuente de variación	gl	SC	CM	F	P
Año (<i>A</i>)	1	47,4	47,38	6,86	0,0307
Reps. dentro año	8	55,2	6,90		
Tratamiento (<i>T</i>)	3	15813,8	5271,28	27,38	0,0111
Invierno (<i>I</i>)	(1)	15086,0	15086,0	78,35	0,0024
Primavera (<i>P</i>)	(1)	633,3	633,3	3,29	0,1771
I × P	(1)	94,5	94,5	0,49	0,7136
<i>A</i> × <i>T</i>	3	577,6	192,54	3,63	0,0272
Error global	24	1272,0	53,00		
Total	39	17766,1			

Utilizando la interacción $A \times T$ como término de error para los contrastes, se obtiene que el único contraste significativo es el efecto del invierno. Así pues, la variación interanual afectaría especialmente al tratamiento primaveral, que resultaría efectivo unos años y no efectivo en otros. En cambio, el tratamiento de invierno parece más efectivo y estable, no siendo afectado por la variación interanual. No obstante, esta conclusión debería confirmarse con un mayor número de años.

Análisis de diferentes localidades

Los experimentos de adaptación de una nueva tecnología (variedad, fertilización, tratamiento fitosanitario, etc.) frecuentemente requieren su realización en varios lugares, condiciones climáticas o localidades. El objetivo básico es conocer el intervalo de adaptabilidad de la nueva tecnología en diferentes áreas geográficas, suelos o ambientes. Para ello, se eligen varios sitios en diferentes estaciones experimentales o fincas de agricultores particulares, por lo que

estos sitios o localidades no son seleccionados al azar. La localidad no se puede considerar, por tanto, una variable aleatoria. Los tratamientos suelen consistir en alguna nueva tecnología que ha resultado prometedora en, al menos, un ambiente determinado. Por ello, al menos uno de los tratamientos incluirá un control, que representará la ausencia de dicha nueva tecnología o la tecnología utilizada tradicionalmente.

Este tipo de experimentos en serie requiere un número considerable de sitios o localidades, por lo que el tamaño de cada ensayo es generalmente pequeño y el diseño experimental simple. Si se emplean experimentos factoriales, no se suelen incluir más de dos factores. Por ello, los diseños más comunes son el de bloques completos al azar y el split-plot. Para el análisis de los datos es fundamental examinar el efecto de la interacción tratamiento*localidad y los efectos de los tratamientos sobre localidades homogéneas. Estos efectos constituyen la base para identificar las mejores tecnologías y su rango de adaptabilidad.

El ejemplo que vamos a utilizar para explicar este tipo de análisis combinado es el de la Tabla 13.5, suponiendo que tenemos tres variedades (V₁, V₂ y V₃) y dos épocas de cultivo (Invierno y Primavera), en un diseño en split-plot con 4 bloques al azar, siendo la parcela principal la época de cultivo. El experimento se repitió en 4 localidades (L₁, L₂, L₃, L₄) representativas del área geográfica de estudio, aunque la localidad 4 se eligió en una zona aislada donde la enfermedad no ocurre.

Tabla 13.5. Rendimiento (kg/ha) de tres variedades en dos épocas de siembra (Invierno y Primavera) y cuatro localidades (L₁, L₂, L₃, L₄)

Variedad	Bloque (REP)				Media
	I	II	III	IV	
L₁, Invierno					
V ₁	0	214	425	40	169,75
V ₂	1252	627	716	1068	915,75
V ₃	2163	2714	2521	2240	2409,50
L₁, Primavera					
V ₁	1036	697	849	1258	960,00
V ₂	1524	1861	2220	1744	1837,30
V ₃	1312	874	695	1133	1003,50
L₂, Invierno					
V ₁	26	125	0	358	127,25
V ₂	813	986	657	432	722
V ₃	2475	2671	3044	2850	2760
L₂, Primavera					
V ₁	488	369	714	983	638,50
V ₂	1815	1662	1233	1486	1549
V ₃	1958	1604	1760	1325	1661,8

Variedad	Bloque (REP)				Media
	I	II	III	IV	
L₃, Invierno					
V ₁	257	64	756	452	382,25
V ₂	789	491	608	1066	738,50
V ₃	2801	3125	2977	3453	3089
L₃, Primavera					
V ₁	412	1118	833	621	746
V ₂	1325	1950	1507	1773	1638,8
V ₃	1582	1236	1844	1371	1508,3
L₄, Invierno					
V ₁	1060	1383	889	723	1013,8
V ₂	816	1125	1463	1277	1170,3
V ₃	892	1112	1521	1366	1222,8
L₄, Primavera					
V ₁	960	474	899	763	774
V ₂	725	583	1179	914	850,25
V ₃	897	985	544	718	786

El ANOVA combinado se inicia con los ANOVA individuales para cada localidad, que están indicados en la siguiente Tabla:

Fuente de variación	gl	L ₁		L ₂		L ₃		L ₄	
		CM	F	CM	F	CM	F	CM	F
Bloque	3	8191		1025		81717		39163	
Época (S)	1	62322	2,34ns	38400	0,97 ns	66887	0,69 ns	662008	11,24*
Error (a)	3	26657		39543		96996		58915	
Variedad(V)	2	2761257	29,69**	6752613	87,76**	6174121	86,38**	34384	0,48 ns
S × V	2	3419332	36,76**	2132259	27,71**	3408091	47,68**	19627	0,28 ns
Error (b)	12	93019		76947		71476		71235	
Total	23								
CV (a)%		13,43		16,00		23,06		25,04	
CV (b)%		25,08		22,31		19,80		27,53	

* = significativo al nivel 5%, ** = significativo al 1%, ns = no significativo ($P > 0.05$)

Para comprobar la homogeneidad de las varianzas, se aplica el *test de Bartlett* (véase Capítulo 4) a los dos errores en las cuatro localidades y se obtiene:

$$\text{Error (a): } \chi^2 = 1.21; \chi^2_{0.05,3} = 7.81 > 1.21 \text{ (Tabla A3)}$$

$$\text{Error (b): } \chi^2 = 0.07; \chi^2_{0.05,3} = 7.81 > 0.07 \text{ (Tabla A3)}$$

Se concluye que las varianzas de los dos errores son homogéneas, por lo que se puede proceder al análisis combinado, según el esquema general para el diseño en “split-plot” que se indica a continuación (véase Capítulo 11):

Fuente de variación	gl ^a	SC	CM	F
Localidad (<i>L</i>)	m-1	SCL	SCL / (m-1)	CML / CMEr
Reps. dentro de localidad	m(r-1)	SCEr	SCEr / m(r-1)	
Factor principal (<i>A</i>)	a-1	SCA	SCA / (a-1)	CMA / CMEa
<i>L</i> × <i>A</i>	(m-1)(a-1)	SCL × <i>A</i>	SCL × <i>A</i> / (m-1)(a-1)	CML × <i>A</i> / CMEa
Error global (a)	m(r-1)(a-1)	SCEa	SCEa / m(r-1)(a-1)	
Factor subparcela (<i>B</i>)	b-1	SCB	SCB / (b-1)	CMB / CMEb
<i>L</i> × <i>B</i>	(m-1)(b-1)	SCL × <i>B</i>	SCL × <i>B</i> / (m-1)(b-1)	CML × <i>B</i> / CMEb
<i>A</i> × <i>B</i>	(a-1)(b-1)	SCA × <i>B</i>	SCA × <i>B</i> / (a-1)(b-1)	CMA × <i>B</i> / CMEb
<i>L</i> × <i>A</i> × <i>B</i>	(m-1)(a-1)(r-1)	SCL × <i>A</i> × <i>B</i>	SCL × <i>A</i> × <i>B</i> / (m-1)(a-1)(r-1)	CML × <i>A</i> × <i>B</i> / CMEb
Error global (b)	ma(r-1)(b-1)	SCEb	SCEb / ma(r-1)(b-1)	
Total	mrab-1			

^am = número de localidades, r = número de repeticiones, a = número de niveles del factor principal, b = número de niveles del factor subparcela.

El ANOVA combinado para los datos del ejemplo sería:

Fuente de variación	gl	SC	CM	F	P
Localidad (<i>L</i>)	3	1866445	622148	19,13	0,0001
Reps. dentro localidad	12	390285	32523,8		
Época (<i>S</i>)	1	98176,0	98176,0	1,77	0,2083
<i>L</i> × <i>S</i>	3	731441	243814	4,39	0,0264
Error global (a)	12	666331	55527,6		
Variedad (<i>V</i>)	2	23194537	11597268	148,36	0,0000
<i>L</i> × <i>V</i>	6	8250212	1375035	17,59	0,0000
<i>S</i> × <i>V</i>	2	13852498	6926249	88,61	0,0000
<i>L</i> × <i>S</i> × <i>V</i>	6	4106118	684353	8,75	0,0000
Error global (b)	48	3752123	78169,2		
Total	95	56908166			

Como todas las interacciones resultan significativas, se procede a su descomposición según contrastes ortogonales. Para ello, se va a considerar un grupo de contrastes ortogonales para la variedad y otro para la localidad. Los contrastes elegidos para el efecto de la variedad son los ya utilizados anteriormente: V_3 vs. $V_{1,2}$ y V_1 vs. V_2 , siendo V_3 una variedad resistente a la enfermedad, V_1 susceptible y V_2 moderadamente susceptible. Para el efecto de la localidad, se han elegido las comparaciones L_4 vs. $L_{1,2,3}$, L_1 vs. $L_{2,3}$ y L_2 vs. L_3 , ya que L_4 es la única localidad donde no se desarrolló la enfermedad y L_1 , la que tuvo el ataque más grave. La nueva Tabla ANOVA con el descomposición de los contrastes correspondientes se indica a continuación:

Fuente de variación	gl	SC	CM	F	P
Localidad (L)	3	1866445	622148	19,13	0,0001
Reps. dentro localidad	12	390285	32523,8		
Época (S)	1	98176,0	98176,0	1,77	0,2083
$L \times S$	3	731441	243814	4,39	0,0264
$S \times (L_4 \text{ vs. } L_{1,2,3})$	(1)	575486	575486	10,36	0,0074
$S \times (L_1 \text{ vs. } L_{2,3})$	(1)	52632	52632	0,95	0,3490
$S \times (L_2 \text{ vs. } L_3)$	(1)	103323	103323	1,86	0,1977
Error global (a)	12	666331	55527,6		
Variedad (V)	2	23194537	11597268	148,36	0,0000
$L \times V$	6	8250212	1375035	17,59	0,0000
$L \times (V_3 \text{ vs. } V_{1,2})$	(3)	7048921	2349640	30,06	0,0000
$L \times (V_1 \text{ vs. } V_2)$	(3)	1201291	400430	5,12	0,0037
$S \times V$	2	13852498	6926249	88,61	0,0000
$S \times (V_3 \text{ vs. } V_{1,2})$	(1)	13648533	13648533	174,60	0,0000
$S \times (V_1 \text{ vs. } V_2)$	(1)	203965	203965	2,61	0,1128
$L \times S \times V$	6	4106118	684353	8,75	0,0000
Error global (b)	48	3752123	78169,2		
Total	95	56908166			

Como las interacciones $L \times V$ y $L \times S \times V$ son significativas y contienen varios grados de libertad, vamos a proceder a su descomposición en contrastes ortogonales con 1 grado de libertad. Para ello, multiplicamos los tres contrastes en que se ha dividido el factor Localidad (L) por los dos contrastes del factor Variedad (V), obteniendo seis contrastes cada uno con un grado de libertad. Estos contrastes y sus coeficientes serían:

Contraste	gl	L ₁			L ₂			L ₃			L ₄		
		V ₁	V ₂	V ₃	V ₁	V ₂	V ₃	V ₁	V ₂	V ₃	V ₁	V ₂	V ₃
$(L_4 \text{ vs. } L_{1,2,3}) \times (V_3 \text{ vs. } V_{1,2})$	1	-1	-1	2	-1	-1	2	-1	-1	2	3	3	-6
$(L_1 \text{ vs. } L_{2,3}) \times (V_3 \text{ vs. } V_{1,2})$	1	2	2	-4	-1	-1	2	-1	-1	2	0	0	0
$(L_2 \text{ vs. } L_3) \times (V_3 \text{ vs. } V_{1,2})$	1	0	0	0	1	1	-2	-1	-1	2	0	0	0
$(L_4 \text{ vs. } L_{1,2,3}) \times (V_1 \text{ vs. } V_2)$	1	-1	1	0	-1	1	0	-1	1	0	3	-3	0
$(L_1 \text{ vs. } L_{2,3}) \times (V_1 \text{ vs. } V_2)$	1	2	-2	0	-1	1	0	-1	1	0	0	0	0
$(L_2 \text{ vs. } L_3) \times (V_1 \text{ vs. } V_2)$	1	0	0	0	1	-1	0	-1	1	0	0	0	0

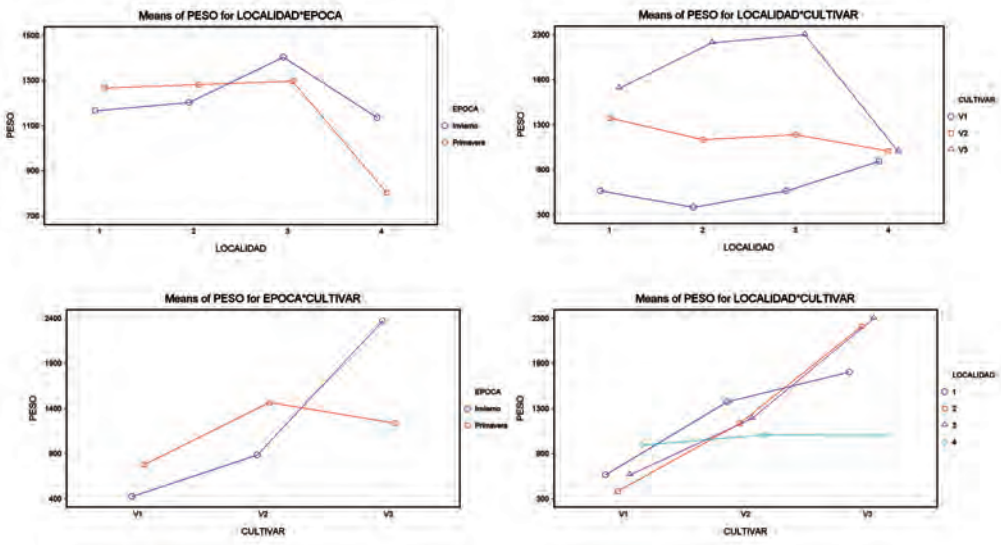
Utilizando estos contrastes se llega a la descomposición de las interacciones que se indica en la siguiente Tabla ANOVA:

Fuente de variación	gl	SC	CM	F	P
$L \times V$	6	8250212	1375035	17,59	0,0000
$(L_4 \text{ vs. } L_{1,2,3}) \times (V_3 \text{ vs. } V_{1,2})$	(1)	5298628	5298628	67,778	0,0000
$(L_1 \text{ vs. } L_{2,3}) \times (V_3 \text{ vs. } V_{1,2})$	(1)	1747982	1747982	22,36	0,0000
$(L_2 \text{ vs. } L_3) \times (V_3 \text{ vs. } V_{1,2})$	(1)	2311	2311	0,03	0,8632
$(L_4 \text{ vs. } L_{1,2,3}) \times (V_1 \text{ vs. } V_2)$	(1)	1128074	1128074	14,43	0,0004
$(L_1 \text{ vs. } L_{2,3}) \times (V_1 \text{ vs. } V_2)$	(1)	40385	40385	0,52	0,4743
$(L_2 \text{ vs. } L_3) \times (V_1 \text{ vs. } V_2)$	(1)	32832	32832	0,42	0,5200
$S \times L \times V$	6	4106118	684353	8,75	0,0000
$S \times (L_4 \text{ vs. } L_{1,2,3}) \times (V_3 \text{ vs. } V_{1,2})$	(1)	3700975	3700975	47,35	0,0000
$S \times (L_1 \text{ vs. } L_{2,3}) \times (V_3 \text{ vs. } V_{1,2})$	(1)	65673	65673	0,84	0,3640
$S \times (L_2 \text{ vs. } L_3) \times (V_3 \text{ vs. } V_{1,2})$	(1)	132239	132239	1,69	0,1998
$S \times (L_4 \text{ vs. } L_{1,2,3}) \times (V_1 \text{ vs. } V_2)$	(1)	124899	124899	1,60	0,2120
$S \times (L_1 \text{ vs. } L_{2,3}) \times (V_1 \text{ vs. } V_2)$	(1)	57967	57967	0,74	0,3939
$S \times (L_2 \text{ vs. } L_3) \times (V_1 \text{ vs. } V_2)$	(1)	24365	24365	0,31	0,5803
Error global (b)	48	3752123	78169,2		

El ANOVA realizado pone de manifiesto que todas las interacciones entre los factores analizados son significativas, lo que dificulta notablemente la interpretación de los resultados. No obstante, la descomposición de las interacciones mediante contrastes ortogonales permite aclarar la naturaleza de dichas interacciones. Así, se puede concluir que la localidad L₄, donde la enfermedad está ausente, es muy diferente a las demás localidades. En esta localidad, no existen diferencias entre cultivares, pero sí entre épocas de siembra, siendo más productivo el cultivo de invierno. En las localidades restantes, el orden de productividad de los cultivares es $V_3 > V_2 > V_1$ en invierno, y $V_2 \geq V_3 \geq V_1$ en primavera, dependiendo del nivel de ataque de la enfermedad en cada localidad. Se puede concluir, por tanto, que la nueva

tecnología de siembra invernal permite un incremento significativo de la producción del cultivo, pero sólo si no hay ataques de la enfermedad o si la variedad es resistente. Los beneficios de la siembra invernal en cuanto al incremento de la productividad están limitados por la enfermedad, que se ve muy favorecida por estas condiciones de cultivo. Así pues, la siembra invernal es recomendable, pero sólo si se usan variedades resistentes a la enfermedad. Las variedades susceptibles o moderadamente susceptibles se deben sembrar en primavera, donde la incidencia de la enfermedad es menor, siendo más recomendable la variedad moderadamente susceptible desde el punto de vista de la productividad.

Los gráficos de las diferentes interacciones, obtenidos del programa *Statistix*, se indican a continuación:



Estos gráficos permiten aclarar la naturaleza de las interacciones entre los tres factores, que llevaría a las conclusiones expuestas anteriormente.

ANÁLISIS DE EXPERIMENTOS DE LARGA DURACIÓN

El efecto de algunas nuevas prácticas de cultivo sobre la productividad debe ser evaluado con experimentos repetidos durante un número considerable de años para poder sacar conclusiones válidas. En estos experimentos, denominados de larga duración, el diseño experimental y los tratamientos se establecen el primer año y permanecen constantes durante todo el tiempo en el mismo campo experimental. La variación temporal de los efectos de los tratamientos es el carácter más importante. Por ello, el rendimiento del cultivo o alguna ca-

racterística del suelo son los indicadores utilizados más habitualmente, aunque generalmente se emplean varios de ellos. Algunos ejemplos de este tipo de experimentos son ensayos de fertilización, de rotación de cultivos, o de control de plagas, enfermedades y malas hierbas.

Para ilustrar la realización del ANOVA en estos casos, vamos a considerar un experimento sencillo sobre control de una enfermedad foliar del olivar. A lo largo de 12 años se han comparado dos manejos diferentes de los tratamientos fungicidas: el tratamiento convencional a calendario fijo (TC) y un nuevo tratamiento o tratamiento dirigido (TD), en el que el momento de aplicación está basado en un sistema de predicción de riesgos de infección. El diseño experimental fue en bloques completos al azar con 5 repeticiones y, entre otros índices, se evaluó el porcentaje de hojas afectadas (PHA) por la enfermedad. Los valores medios del PHA se presentan en la tabla siguiente:

Tabla 13.6. Porcentaje de hojas de olivo afectadas por una enfermedad.

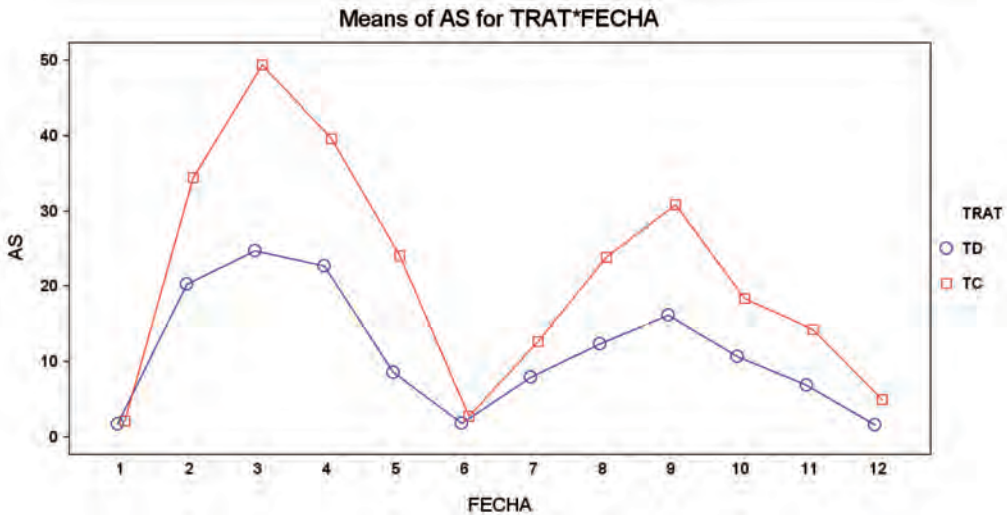
Tratamiento	Año											
	1	2	3	4	5	6	7	8	9	10	11	12
TC	0,20	32,6	57,3	40,8	16,9	0,40	5,1	16,7	26,9	10,2	6,3	1,00
TD	0,16	12,4	18,0	15,2	2,5	0,20	2,3	4,8	8,1	3,6	1,8	0,16

El ANOVA para este caso sigue el procedimiento general para un diseño “split-plot” estándar con los tratamientos como parcela principal y los años como subparcela. Algunos estadísticos dudan de la adecuación de este procedimiento en estos casos (véase para mayor información el apartado sobre experimentación en plantas leñosas del Capítulo 21). No obstante, a efectos del ejercicio se ha considerado su validez. La Tabla correspondiente a este análisis, utilizando el programa *Statistix* y los datos ajustados a la normalidad mediante la transformación angular, sería:

<u>Analysis of Variance Table for AS</u> $[AS = \text{Arcsen}(\sqrt{PHA/100})]$					
Source	DF	SS	MS	F	P
REP	4	597.9	149.48	0.65	0.6550
TRAT	1	3103.5	3103.48	13.56	0.0212
Error REP*TRAT	4	915.3	228.82		
FECHA	11	14710.9	1337.36	187.44	0.0000
TRAT*FECHA	11	1477.8	134.35	18.83	0.0000
Error REP*TRAT*FECHA	88	627.9	7.14		
Total	119	21433.4			

Grand Mean 16.335
 CV (REP*TRAT) 92.60
 CV (REP*TRAT*FECHA) 16.35

Tanto los tratamientos, como los años y su interacción han resultado significativos. Al ser la interacción significativa, el procedimiento general a seguir sería descomponer los tratamientos en sumas de cuadrados de 1 gl mediante contrastes ortogonales y estudiar la interacción de los años con dichos contrastes. Como en el ejemplo tenemos sólo dos tratamientos, la partición no procede, y sólo nos queda examinar la naturaleza de la interacción a lo largo del tiempo. Con el programa *Statistix*, después de comprobar que se cumplen los supuestos de normalidad y de distribución aleatoria de los residuos, entraríamos en la opción "Plots" del menú de resultados y elegimos "Means Plot...". Introducimos las variables año (FECHA) y tratamiento (TRAT), pedimos un gráfico de líneas, y obtenemos:



En el gráfico obtenido se puede comprobar claramente que la interacción TRAT*FECHA se debe a que en los años de menor ataque de la enfermedad no hay diferencias significativas entre los dos tratamientos, mientras que éstas son muy acusadas en los años de mayor incidencia de la enfermedad. Así pues, el tratamiento dirigido (TD), aplicando los fungicidas según la predicción de riesgos de infección, resulta más eficaz para el control de la enfermedad que el tratamiento convencional (TC), aunque los dos tratamientos se ven afectados por las oscilaciones derivadas de la irregularidad anual de los ataques de la enfermedad. Habría que considerar el coste de los tratamientos en uno y otro caso, los residuos de fungicidas y el rendimiento del cultivo, para poder recomendar de forma generalizada el nuevo tratamiento en lugar del convencional.

El estudio de la relación entre los años y las medias de los tratamientos conforme se ha realizado en el gráfico anterior, evitando el análisis en "split-plot", es la alternativa sugerida para el estudio de estos casos, que puede completarse a veces con el ajuste de curvas.

CAPÍTULO 14

CORRELACIÓN LINEAL

Por correlación se entiende la tendencia de dos o más variables a estar relacionadas de una manera definida. La medida de la relación entre dos variables aleatorias se expresa por el coeficiente de correlación lineal (r). Para conocer si la relación es real o accidental, hay que referirse a las tablas de probabilidad.

El coeficiente de correlación lineal r tiene dos propiedades: a) es un número sin dimensiones, y b) sus valores están comprendidos entre -1 y $+1$. Si los pares de valores correspondientes a dos variables se expresan en una gráfica, los puntos de la misma tienden a disponerse en una banda que suele tomar la forma de una elipse, cuya amplitud depende del valor de r . La Fig. 14.1 muestra gráficamente las relaciones entre dos variables en función del valor de r .

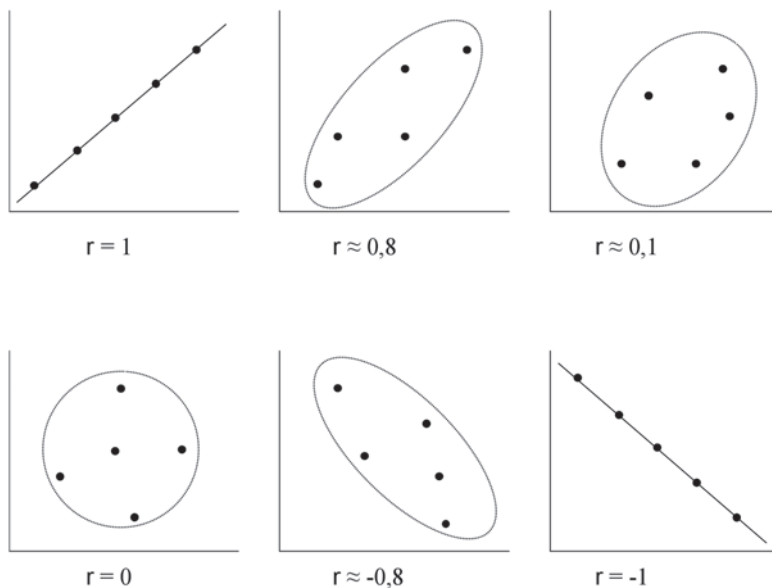


Fig. 14.1. Diagramas de dispersión para diferentes valores de r .

La correlación perfecta acontece para valores de $r = 1$ ó $r = -1$, algo que difícilmente ocurre con datos biológicos. El valor $r = 0$ indica que las dos variables no están relacionadas. Para realizar la prueba de hipótesis nula $\rho = 0$, los valores de r deben contrastarse con las tablas de probabilidad para $n-2$ grados de libertad.

CÁLCULO DEL COEFICIENTE DE CORRELACIÓN LINEAL

El ejemplo que se muestra en la Tabla 14.1 recoge los datos correspondientes a un ensayo con plantas jóvenes de olivo sometidas a diferentes tratamientos con distintos fertilizantes nitrogenados. Se establecieron ocho grupos de plantas, cada uno compuesto de cinco plantas de olivo, que recibieron un tratamiento diferente. La tabla muestra el contenido total de nitrógeno en las plantas de olivo al final del experimento (X_1) y la concentración de nitrógeno en las hojas de esas mismas plantas (X_2). Interesa conocer si ambas variables están correlacionadas.

Tabla 14.1. Contenido de nitrógeno en plantas jóvenes de olivo y concentración de nitrógeno en hojas en ocho grupos de plantas sometidas a diferentes tratamientos.

Grupo	X_1 (mg N/planta)	X_2 (% N en hoja)
1	390,6	1,79
2	376,7	1,58
3	524,9	2,20
4	502,8	2,52
5	371,4	1,60
6	494,8	2,61
7	1042,6	4,24
8	1168,9	5,12
Media	609,09	2,71

$$\sum x_1^2 = \sum (X_1 - \bar{X}_1)^2 = 691005,41$$

$$\sum x_2^2 = \sum (X_2 - \bar{X}_2)^2 = 11,81$$

$$\sum x_1 x_2 = \sum (X_1 - \bar{X}_1) (X_2 - \bar{X}_2) = 2814,42$$

$$r = \frac{\sum x_1 x_2}{\sqrt{(\sum x_1^2) (\sum x_2^2)}} = \frac{2814,42}{\sqrt{(691005,41)(11,81)}} = 0,9852$$

Los valores de r en la tabla de probabilidad (Tabla A4) para $8-2 = 6$ grados de libertad son $r = 0,7067$ para $P = 0,05$; $r = 0,8343$ para $P = 0,01$; y $r = 0,9249$ para $P = 0,001$; en todos los casos inferiores al valor obtenido de $r = 0,9852$. Esto permite rechazar la hipótesis nula $\rho = 0$, es decir, r ha resultado ser altamente significativo e indica que las dos variables estudiadas están significativamente relacionadas entre sí con un nivel de significación del 1%.

Nótese que si se expresa la covarianza de X_1 y X_2 como $s_{X_1X_2}$ y las correspondientes varianzas de X_1 y de X_2 como $s_{X_1}^2$ y $s_{X_2}^2$,

$$s_{X_1X_2} = \frac{\sum x_1x_2}{n} \quad s_{X_1} = \sqrt{\frac{\sum x_1^2}{n}} \quad s_{X_2} = \sqrt{\frac{\sum x_2^2}{n}}$$

el coeficiente de correlación puede escribirse como:

$$r = \frac{s_{X_1X_2}}{s_{X_1} s_{X_2}}$$

es decir, el cociente entre la covarianza y el producto de las desviaciones típicas de ambas variables.

CÁLCULO CON EL PROGRAMA *Statistix* (SX)

En el programa SX, una vez introducidos los datos experimentales, seguir la siguiente secuencia:

Statistics > Linear Models > Correlations (Pearson)

Al llegar a este punto, indicar las variables a correlacionar, marcar la casilla *Compute P-Values* para que incluya el nivel de significación de r y pulsar OK. Aparece, en el caso del ejemplo, la siguiente tabla:

Correlations (Pearson)	
	NHOJA
NPLANTA	0.9852
P-VALUE	0.0000
Cases Included	8
Missing Cases	0

INTERPRETRACIÓN DE LOS VALORES DE r

El cálculo del coeficiente de correlación es uno de los más sencillos en estadística y, a veces, está sujeto a interpretaciones incorrectas. Para un uso correcto de la interpretación de los valores de r , tenemos que tener en cuenta las siguientes cautelas:

1. *Tamaño de la muestra.* En muestras pequeñas las correlaciones muestrales son muy variables. En el caso del ejemplo anterior, si cambiamos el valor de X_1 en el grupo 8 por 25, el valor de r pasa de 0,9852 a 0,0214, mostrando una clara falta de significación en la relación entre ambas variables. La obtención de una medida anormal puede darse con alguna frecuencia en la experimentación, por lo que hay que ser cautelosos al interpretar los valores de r si la muestra es pequeña.

En muestras muy grandes el problema anterior se minimiza y un error en la medida de algún dato no altera la interpretación final de los resultados. Pero los valores de r en muestras grandes pueden ser bajos y resultar significativos, pues en las tablas de probabilidad los valores disminuyen conforme aumentan los grados de libertad. En estos casos cabe preguntarse acerca de lo que puede significar un bajo valor de r , aún mostrando significación. Las figuras anteriores pueden clarificar esto, pues una elipse cercana a la forma de una circunferencia indica una dispersión de los puntos en la gráfica tal que no resulta factible establecer una relación clara entre las variables. Esto es un factor que con frecuencia se olvida, pudiendo llegar a conclusiones alejadas de la realidad.

2. *Relación de causa-efecto.* Aún obteniendo un valor alto y significativo de r no se puede interpretar que un cambio en una de las variables cause un cambio en la otra, pues otras variables no estudiadas pueden afectar al proceso. Incluso se obtienen con frecuencia correlaciones absurdas y sin sentido pero altamente significativas. Así, por ejemplo, existe una correlación altamente significativa ($r = -0,99^{***}$) entre la acumulación de horas-frío (horas bajo $7\text{ }^{\circ}\text{C}$) en Córdoba durante el otoño-invierno de 1984 y la puntuación obtenida por los equipos de fútbol en la Liga Española en la temporada 2002-03. Probablemente nadie dude de la falta de relación entre ambas variables, aunque el valor de r sea real. En otros casos la falta de relación no resulta tan evidente; por ejemplo, ¿existe relación causal entre la incidencia de una enfermedad de las plantas y el abonado? Solo el investigador puede establecer el sentido de la relación en base al conocimiento que tiene sobre las variables.

3. *Linealidad.* A veces se olvida que r es el coeficiente de correlación lineal y solo tiene validez si la relación entre las variables estudiadas es de ese tipo. Por ejemplo, si se calcula r para el siguiente par de valores:

X:	0	1	2	3	4	5	6
Y:	1	13	25	30	25	13	1

se obtiene que $r = 0$, lo que indicaría que no existe relación alguna entre las variables, pero si se representa la gráfica de los datos observamos que se ajustan perfectamente a una parábola cuya ecuación (véase el capítulo 18) es $Y = -0,05 + 18,6 X - 3,09 X^2$, $R^2=0,98^{***}$. No podemos decir que no existe relación entre las variables, pero sí que esa relación no es lineal.

CONSIDERACIONES FINALES

El cálculo del coeficiente de correlación r requiere que la población siga una distribución normal bivalente, esto es, una distribución normal de dos variables. Si esto no es así, convendría calcular el coeficiente de correlación de Spearman conforme se indica en el capítulo 19.

La tabla de probabilidad de r (Tabla A4) puede utilizarse, como se ha hecho en el ejemplo, para probar únicamente la hipótesis nula $H_0: \rho = 0$, pero no otras hipótesis nulas, como $H_0: \rho = 0,5$ ó $H_0: \rho_1 = \rho_2$, o para obtener intervalos de confianza de ρ . Para realizar estos cálculos hay que transformar r en una cantidad z distribuida casi normalmente. Estos procedimientos no se han considerado en este capítulo pues, por lo general, cuando un investigador necesita realizar estudios más profundos de las dos variables consideradas recurre a un estudio de regresión como se describe en los próximos capítulos. No obstante, una descripción detallada de esos métodos puede encontrarse en Snedecor and Cochran (1974) y en Steel and Torrie (1985).

CAPÍTULO 15

REGRESIÓN LINEAL

Desde un punto de vista experimental, se entiende por regresión la magnitud del cambio de una variable asociada con una unidad de cambio de otra. Se distingue de la correlación en que mientras ésta se refiere al hecho de que dos variables están relacionadas, la regresión se refiere a la naturaleza de esa relación.

En la regresión existe relación de dependencia entre las variables en estudio, y suele designarse por **Y** a la variable dependiente y por **X** a la independiente. Por ejemplo, en un estudio sobre la influencia de las dosis de riego en la producción, la variable dependiente Y sería la producción mientras que la independiente X sería las dosis de riego.

La regresión lineal viene expresada por la ecuación de una recta,

$$Y = a + bX$$

Si se representa en una gráfica (Fig. 15.1) la línea recta es la regresión de Y sobre X, y su posición está fijada por:

1. La recta pasa por el punto (\bar{X}, \bar{Y}) .
2. La intersección con el eje de ordenadas es el coeficiente a .
3. Su pendiente es el coeficiente de regresión b .

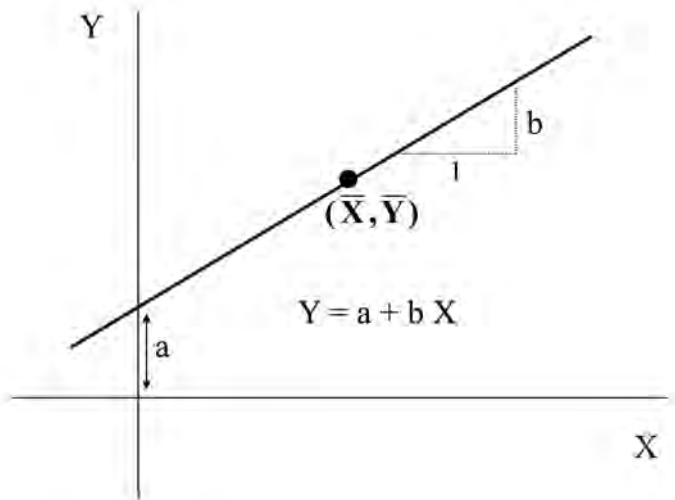


Fig. 15.1. Recta de regresión de Y sobre X.

La regresión suele emplearse para conocer si Y depende de X, para predecir Y conociendo X y, en general, para probar hipótesis sobre causa-efecto. Como se ha indicado en el capítulo 5, la separación de medias que sigue al análisis de la varianza ha de hacerse por regresión cuando la variable en estudio es un factor cuantitativo, como la dosis de aplicación de un producto químico, fechas de siembra, etc. En el ejemplo anterior, donde se trataría de estudiar el efecto en la producción de las distintas dosis de agua aplicada, la separación de medias habría que realizarse por regresión.

CÁLCULO DE LA REGRESIÓN LINEAL

El ejemplo siguiente recoge datos de un estudio en el que se pretende conocer si la concentración de nitrógeno en hojas de olivo depende o no de la cantidad de nitrógeno aportado en la fertilización. Los datos experimentales correspondientes a ambas variables, sus desviaciones respecto a las medias y las sumas de cuadrados y de productos se recogen en la Tabla 15.1.

Tabla 15.1. Concentración de nitrógeno en hojas de olivo en función de la fertilización nitrogenada recibida.

	Fertilización (kg N/árbol)	N en hoja (%)	Desviaciones de la media		Cuadrados		Productos
	X	Y	$x = X - \bar{X}$	$y = Y - \bar{Y}$	x^2	y^2	xy
	0	1,41	-0,37	-0,29	0,137	0,084	0,1073
	0,12	1,62	-0,25	-0,08	0,062	0,0064	0,020
	0,25	1,74	-0,12	0,04	0,015	0,0016	-0,005
	0,50	1,78	0,13	0,08	0,017	0,0064	0,010
	1	1,93	0,63	0,23	0,397	0,053	0,1449
Suma	1,87	8,48	0	0	0,628	0,1514	0,2772
Media	0,37	1,70					

El coeficiente de regresión b se determina según la expresión:

$$b = \frac{\sum xy}{\sum x^2} = \frac{0,2772}{0,628} = 0,44$$

y la ecuación de regresión se escribe como:

$$\hat{Y} - \bar{Y} = b(X - \bar{X}), \text{ de donde:}$$

$$\hat{Y} = \bar{Y} + b(X - \bar{X}) = 1,70 + 0,44(X - 0,37) = 1,54 + 0,44 X$$

siendo \hat{Y} el valor estimado de Y en la regresión, esto es, el valor en la recta para cada valor de X.

Desde un punto de vista matemático, la regresión lineal se define como la recta que hace mínima la suma de cuadrados de las desviaciones entre los puntos observados y los estimados. Para determinar entonces la bondad del ajuste de la línea de datos, se calculan las desviaciones de regresión $d_{y,x}$, es decir, las diferencias entre las Y observadas y las Y estimadas por la ecuación de regresión (Fig. 15.2). Por ejemplo, para X = 0,12 la $\hat{Y} = 1,59$ y la observada resultó 1,62 luego $d_2=0,03$.

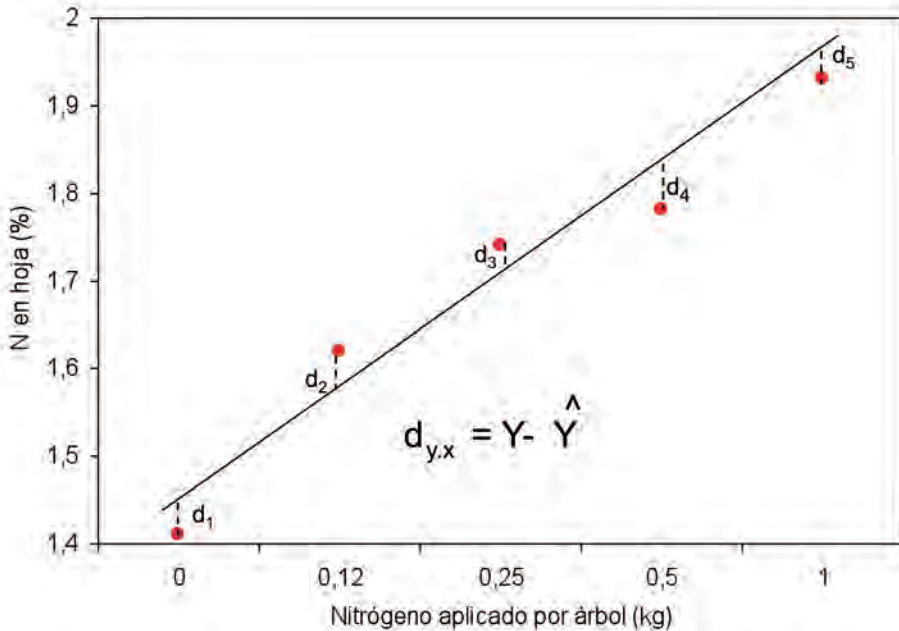


Fig. 15.2. Desviaciones de regresión.

La suma de cuadrados de las desviaciones $\sum d_{y,x}^2$ es, pues, la base para estimar la bondad del ajuste y puede calcularse sumando cada $d_{y,x}$ ($d_1^2 + d_2^2 + \dots + d_n^2$) o bien directamente por la fórmula:

$$\sum d_{y,x}^2 = \sum y^2 - \frac{(\sum xy)^2}{\sum x^2} = 0,1514 - \frac{0,2772^2}{0,628} = 0,029$$

La desviación típica de regresión es:

$$s_{y.x} = \sqrt{\frac{\sum d_{y.x}^2}{n-2}} = \sqrt{\frac{0,029}{3}} = 0,098$$

El error estándar de b es:

$$s_b = \frac{s_{y.x}}{\sqrt{\sum x^2}} = \frac{0,098}{0,79} = 0,12 \quad \text{con } n-2 = 3 \text{ gl}$$

La prueba de significación de b está dada por la t de Student de la siguiente forma:

$$t = \frac{b}{s_b} = \frac{0,44}{0,12} = 3,67$$

El valor de $t = 3,67 > t_{0,05} = 3,182$ en la Tabla A1 para 3 gl, luego t ha resultado significativa al nivel del 5%.

La prueba de significación también puede hacerse mediante un análisis de varianza. Para la regresión queda como sigue:

Fuente de Variación	gl	SC	CM	F
Total	n-1	SC _T		
Regresión lineal	1	SC _R	SC _R /1	CM _R /CM _D
Desviaciones de Regresión	n-2	SC _D	SC _D /n-2	

$$n = 5$$

$$SC_T = \sum y^2$$

$$SC_R = \sum y^2 - \sum d_{y,x}^2$$

$$SC_D = \sum d_{y,x}^2$$

En el ejemplo, los cálculos quedan de la siguiente forma:

Fuente de Variación	gl	SC	CM	F
Total	4	0,1514		
Regresión lineal	1	0,1224	0,1224	12,62*
Desviaciones de Regresión	3	0,029	0,0097	

Contrastando en las tablas F (Tabla A2) para 1 y 3 grados de libertad, se obtiene que $12,62 > 10,13$ lo que indica que la regresión ha resultado significativa al nivel del 5% ($P = 0,05$). Obsérvese que el cuadrado medio de las desviaciones de regresión CM_D es $s^2_{y,x}$.

RELACIONES CON LA CORRELACIÓN

Cuando las variables están muy relacionadas hay coincidencias entre la correlación y la regresión. El coeficiente de determinación r^2 indica la proporción de la suma de cuadrados total que es atribuible a la variable independiente, y suele utilizarse en conexión con el análisis de regresión. El coeficiente de determinación es el producto de los coeficientes de regresión de y sobre x y de x sobre y:

$$r^2 = b_{yx} \cdot b_{xy}$$

y algebraicamente resulta ser el cuadrado del coeficiente de correlación.

En el ejemplo, $r^2 = 0,81$ e indica que el 81% de la variación en la concentración de nitrógeno en hojas de olivo es debida a la fertilización nitrogenada.

CÁLCULO CON EL PROGRAMA *Statistix* (SX)

En el programa SX, una vez introducidos los datos experimentales, seguir la siguiente secuencia:

Statistics >Linear Models> Linear Regresión

Al llegar a este punto, indicar cuál es la variable dependiente y cuál la independiente y luego pulsar OK, teniendo la precaución de que quede señalada la opción *Fit constant*, que indica el cálculo del coeficiente a . Aparece, en el caso del ejemplo, la siguiente tabla:

Unweighted Least Squares Linear Regression of NHOJA

Predictor Variables	Coefficient	Std Error	T	P
Constant	1.53038	0.06313	24.24	0.0002
NFERTI	0.44282	0.12254	3.61	0.0364

R-Squared	0.8132	Resid. Mean Square (MSE)	0.00942
Adjusted R-Squared	0.7509	Standard Deviation	0.09707

Source	DF	SS	MS	F	P
Regression	1	0.12305	0.12305	13.06	0.0364
Residual	3	0.02827	0.00942		
Total	4	0.15132			

Cases Included 5 Missing Cases 0

El valor de los coeficientes ($a = 1,53$ y $b = 0,44$) permite escribir la ecuación de regresión. El programa realiza la prueba de significación basado tanto en la t de Student como en la F , conforme se ha realizado anteriormente. Los valores de $s_b = 0,12254$, $r^2 = 0,81$ y $s_{y,x} = 0,097$ aparecen con claridad en la tabla. Podrá apreciarse que los valores calculados anteriormente en el ejemplo y los que emite el programa no siempre coinciden exactamente, lo que es debido a las aproximaciones en los cálculos con decimales.

ESTIMACIONES DE LA REGRESIÓN POBLACIONAL

La ecuación de la recta $Y = a + bX$ se expresa para la población como $\mu = \alpha + \beta x$, donde μ es la media de la población de los valores de Y correspondientes a una X determinada, cuya desviación típica es $\sigma_{y,x}$. Para cada X , pues, existe una distribución normal de Y de media μ y desviación típica $\sigma_{y,x}$. La Fig. 15.3 muestra esa relación.

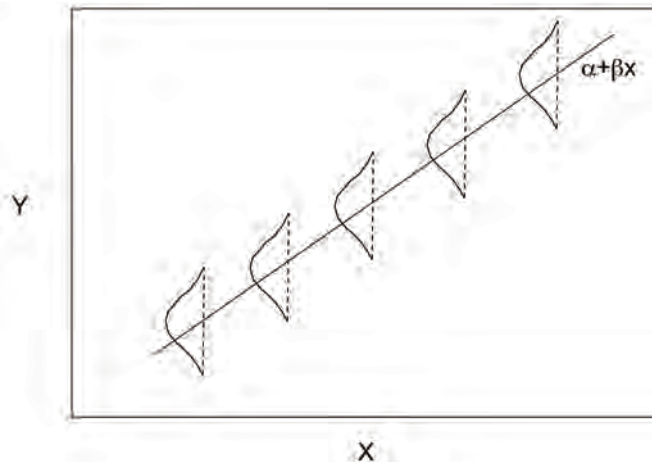


Fig. 15.3. Modelo de regresión lineal de Y sobre X.

El modelo matemático de la regresión lineal queda expresado por la ecuación:

$$Y = \alpha + \beta X + \varepsilon$$

donde ε es una variable aleatoria tomada de la distribución normal de la población de media cero y desviación típica $\sigma_{y,x}$. Para cada X la diferencia entre la Y estimada y la media poblacional es:

$$\hat{Y} - \mu = (\bar{Y} - \alpha) + (b - \beta)X$$

El primer sumando de la ecuación representa la diferencia en las elevaciones de la recta y el segundo las diferencias entre las pendientes, ambas debidas a ε .

Las estimaciones de μ pueden hacerse en una X no muestreada pero que se encuentra entre dos X que sí lo fueron. Por ejemplo, para 0,3 kg N/árbol. Aunque esta dosis no se aplicó, su valor debe estar situado en la recta de regresión. Sin embargo, si pretendemos extrapolar y estimar μ para una X cuyo valor excede al experimentado, por ejemplo para $X = 2$ kg N/árbol, no tenemos certeza de que los resultados puedan ser correctos, pues la relación entre las variables puede que no sea lineal más allá de los valores muestreados.

Los intervalos de confianza de μ y de β resultan de interés en el estudio de la regresión. Para los β el intervalo de confianza al 95%, siendo $t_{0,05}$ la t de Student a ese nivel de significación para $n-2$ gl, es:

$$b - t_{0,05} s_b \leq \beta \leq b + t_{0,05} s_b$$

que en el ejemplo resulta:

$$0,44 - 3,182 * 0,12 \leq \beta \leq 0,44 + 3,182 * 0,12$$

$$0,058 \leq \beta \leq 0,82$$

Análogamente, el intervalo de confianza de las μ es:

$$\hat{Y} - t_{0,05} s_{\hat{y}} \leq \mu \leq \hat{Y} + t_{0,05} s_{\hat{y}}$$

siendo

$$s_{\hat{y}} = s_{y,x} \sqrt{\frac{1}{n} + \frac{x^2}{\sum x^2}}$$

el error estándar de \hat{Y} .

Para cada valor de X puede obtenerse de ese modo el intervalo de confianza de μ . Si se hacen los cálculos para todos los valores de X y se expresan en una gráfica, se puede observar que los límites de confianza están representados por las ramas de una hipérbola (Fig. 15.4). Obsérvese el riesgo de predecir en valores de X muy alejados de la media.

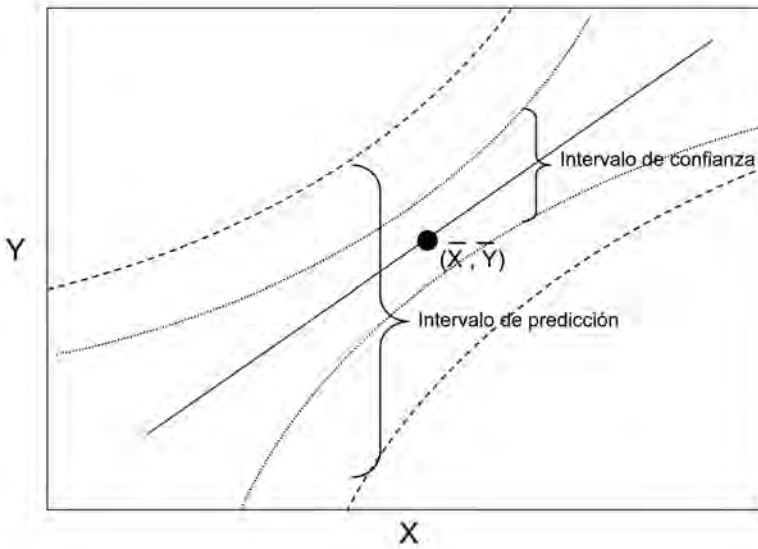


Fig. 15.4. Intervalos de confianza y de predicción en la regresión lineal.

En el ejemplo, para $X = 0,25$ tenemos que:

$$s_{\hat{y}} = 0,098 \sqrt{\frac{1}{5} + \frac{0,015}{0,628}} = 0,046$$

y el intervalo de confianza para ese valor de X es:

$$1,65 - 3,182 * 0,046 \leq \mu \leq 1,65 + 3,182 * 0,046$$

$$1,50 \leq \mu \leq 1,796$$

A veces interesa predecir la Y en lugar de μ , esto es, el valor individual de Y para un nuevo miembro de la población para el que se midió la X. En el caso anterior se trataba de estimar la media de la respuesta a la aplicación de 0,25 kg N/árbol, y ahora se trata de predecir cuál sería la respuesta a una nueva aplicación de esa cantidad de nitrógeno. Los cálculos son idénticos, pero

$$s_{\hat{y}} = s_{y,x} \sqrt{1 + \frac{1}{n} + \frac{x^2}{\sum x^2}}$$

Para el caso anterior tenemos que ahora $s_y = 0,108$ y el intervalo de predicción de Y resulta:

$$1,65 - 3,182*0,108 \leq Y \leq 1,65 + 3,182*0,108$$

$$1,30 \leq Y \leq 1,99$$

Estos cálculos pueden ser muy interesantes cuando en el estudio de una regresión aparece un valor individual de Y que se desvía mucho de la recta, esto es, presenta una $d_{y,x}$ muy grande. Interesaría conocer en este caso si el valor medido está dentro del intervalo de predicción de Y en la regresión. Para determinarlo basta con omitir el punto, calcular la nueva recta de regresión y determinar el intervalo de predicción de Y como si fuera un nuevo punto. Si el valor medido está dentro del intervalo no existe ningún problema y seguimos considerando el valor medido. Por el contrario, si se encuentra fuera del intervalo es necesario determinar la causa. Por lo general debemos considerar lo siguiente:

1. Si hubo un error en la medida. En ocasiones es posible hacerla de nuevo; por ejemplo, si se conserva parte de un extracto vegetal es posible hacer de nuevo determinaciones químicas que permitan corregir el valor de ese punto extraño.
2. Si hubo influencia de un factor aberrante. En experimentos de campo es conveniente llevar un diario donde se anoten todas las eventualidades que se observen durante el desarrollo del experimento. Una lluvia ocasional, por ejemplo, puede alterar la dosis de riego en un experimento y provocar una respuesta no esperada. Si se conoce ese hecho, el valor obtenido debería omitirse para no alterar las conclusiones del experimento.
3. Si no se encuentra explicación para ese valor la situación es difícil. Lo más correcto en este caso es examinar las conclusiones obtenidas con ese punto y sin él; si difieren, lo aconsejable sería repetir el experimento o probar otros ajustes diferentes de la regresión lineal.

Los intervalos de confianza y de predicción se obtienen directamente en el programa *Statistix* pulsando la siguiente secuencia en el menú que aparece junto a la tabla de resultados:

Results > Plots > Simple Regression Plot

AJUSTE DE LA RECTA POR EL ORIGEN

En algunos casos es posible establecer que cuando $X = 0$ el valor de Y debería ser igualmente cero, por lo que sería posible forzar el paso de la recta por el origen, esto es, eliminar la intersección de la línea con el eje Y. La ecuación de la recta quedaría de la forma

$$Y = bX$$

donde se ha eliminado el término independiente.

Para ajustar una recta por el origen, los cálculos son los siguientes:

$$b = \frac{\sum XY}{\sum X^2} \quad s_{y \cdot x}^2 = \frac{\sum Y^2 - \frac{(\sum XY)^2}{n-1}}{n-1} \quad s_b = \frac{s_{y \cdot x}}{\sqrt{\sum X^2}}$$

En el ejemplo anterior tenemos que:

$$\sum X^2 = 1,3269 \quad \sum XY = 3,4494 \quad \sum Y^2 = 14,5334$$

con lo que calculamos:

$$b = \frac{3,4494}{1,3269} = 2,599$$

y la ecuación de la recta sería $\hat{Y} = 2,599X$. Para determinar la bondad del ajuste calculamos:

$$s_{y \cdot x}^2 = \frac{14,5334 - \frac{3,4494^2}{5-1}}{5-1} = 1,3916 \quad \text{luego } s_{y \cdot x} = 1,1796$$

$$s_b = \frac{1,1796}{\sqrt{1,3269}} = 1,024$$

$$t = \frac{b}{s_b} = \frac{2,599}{1,024} = 2,538$$

que con $n-1 = 4$ gl resulta no significativo.

El forzado por el origen tiene interés para ganar precisión en la predicción de Y, pero debe realizarse únicamente cuando tenga sentido. En el caso del ejemplo no lo tiene, pues la falta de abonado no implica que el árbol no tenga nitrógeno, ya que absorbe el disponible en la solución del suelo, luego el supuesto para $X = 0, Y = 0$ carece de sentido. La regresión forzada por el origen ha resultado, además, no significativa, pero esto no es debido a su falta de sentido, sino que no resulta lineal como puede comprobarse si se representa la figura correspondiente.

Por otra parte, si los valores de X se encuentran lejos del origen (del valor 0) puede ocurrir que la relación entre las variables sea curvilínea en la población, pero lineal o ligeramente curva en el intervalo estudiado, por lo que al forzarla a pasar por el punto (0,0) la regresión lineal no resulta significativa. La representación gráfica de los datos siempre es de utilidad antes de iniciar los cálculos de la regresión.

Los cálculos anteriores están indicados por su precisión cuando la varianza del residual ε es constante. Sin embargo no siempre lo es, sino que aumenta conforme aumenta X , lo que se puede apreciar gráficamente pues los puntos caen cerca de la línea cuando X es pequeña y tienden a alejarse de ella cuando X aumenta. Si se han realizado varias medidas de Y para cada X , se puede apreciar en una gráfica que la amplitud de los valores para cada grupo de Y aumenta al hacerlo X (Fig. 15.5).

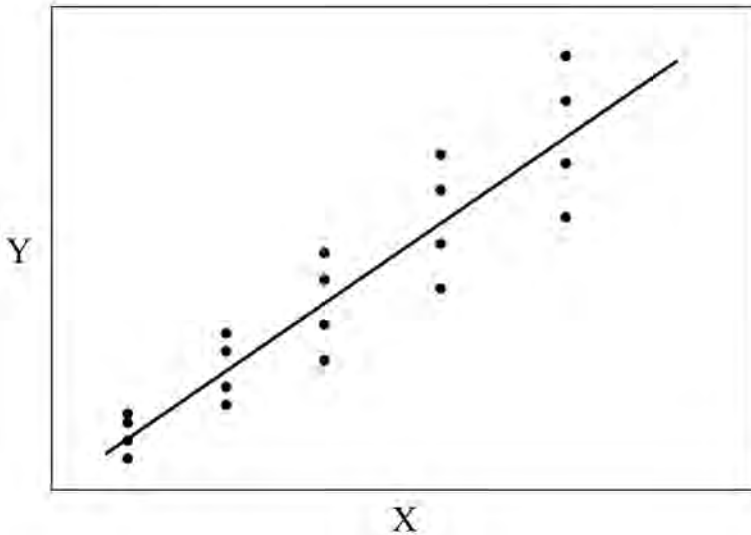


Fig. 15.5. Regresión de Y sobre X con la varianza del residual que aumenta al aumentar X .

En este caso, cuando todas las X son positivas, los cálculos siguientes resultan de mayor precisión:

$$b = \frac{\sum Y}{\sum X}$$

$$s_{y.x}^2 = \frac{\sum \frac{Y^2}{X} - \frac{(\sum Y)^2}{\sum X}}{n-1}$$

$$s_b = \frac{s_{y.x}}{\sqrt{\sum X}}$$

CÁLCULO CON EL PROGRAMA *Statistix* (SX)

El programa permite de una forma sencilla forzar el paso de la recta por el origen en el primer supuesto, cuando la varianza de ε es constante. Para ello, basta con quitar la señal en la casilla *Fit constant* antes de pulsar OK para generar la salida de datos, una vez que se ha indicado cuál es la variable dependiente y cuál la independiente. Se obtiene, de esta manera, la siguiente tabla:

Unweighted Least Squares Linear Regression of NHOJA					
NOTE: Model Forced through origin					
Predictor					
Variables	Coefficient	Std Error	T	P	
NFERTI	2.59959	1.02409	2.54	0.0641	
R-Squared	0.6170	Resid. Mean Square (MSE)	1.39159		
Adjusted R-Squared	0.5212	Standard Deviation	1.17966		
Centered R-Squared	-35.7854				
Source	DF	SS	MS	F	P
Regression	1	8.9670	8.96704	6.44	0.0641
Residual	4	5.5664	1.39159		
Total	5	14.5334			
Cases Included	5	Missing Cases	0		

Como en casos anteriores, el programa realiza las pruebas t y F de bondad del ajuste.

COMPARACIÓN DE LÍNEAS DE REGRESIÓN

La relación entre X e Y se estudia muchas veces en muestras diferentes tomadas en distintas localidades, en diferentes épocas o por distintos investigadores. Cuando se dispone de datos de muestras distintas, cabe preguntarse si las líneas de regresión son o no idénticas, pues en caso de serlo podrían combinarse todos los datos de todas las muestras y ganar precisión en el estudio de la relación entre ambas variables. A veces tan solo se pretende conocer si son o no iguales las rectas de regresión obtenidas de muestras diferentes.

Dos rectas de regresión pueden diferir en las varianzas residuales, en la pendiente o en la elevación, esto es, en la intersección de la recta con el eje Y , por lo que habrá que comparar esos valores para conocer la igualdad de las rectas. Siguiendo con el ejemplo del principio, se trataría de comprobar si las rectas obtenidas con datos de otro año son iguales o no a las del ejemplo. La nueva muestra aporta los datos recogidos en la Tabla 15.2.

Tabla 15.2. Concentración de nitrógeno en hojas de olivo en función de la fertilización nitrogenada recibida.

Fertilización (kg N/árbol)	N en hoja (%)
X	Y
0	1,45
0,12	1,60
0,25	1,56
0,50	1,70
1	1,75

Siguiendo el procedimiento normal de cálculo obtenemos que:

$$\sum x^2 = 0,628 \qquad \sum xy = 0,168 \qquad \sum y^2 = 0,056$$

y la ecuación de regresión se escribe como:

$$\hat{Y} = 1,51 + 0,26 X ; \quad r^2 = 0,80^*$$

Resulta una relación significativa, luego se está en condiciones de poder comparar con la recta del ejemplo primero para conocer si son o no idénticas. Se comparan primero las varianzas residuales, si resultan que no difieren significativamente se comparan las pendientes de las rectas, esto es, los coeficientes de regresión, y si tampoco difieren, se comparan por último las elevaciones de las rectas. Los cálculos se realizan como se indica en la siguiente tabla:

					Desviaciones de regresión			
	gl	$\sum x^2$	$\sum xy$	$\sum y^2$	gl	SC	CM	F
Muestra 1	4	0,628	0,277	0,151	3	0,029	0,0097	2,62 ^{NS}
Muestra 2	4	0,628	0,168	0,056	3	0,011	0,0037	
		Suma de gl y de SC			6	0,040	0,0067	
Muestras 1+2	8	1,256	0,445	0,207	7	0,049	0,007	
		Diferencia entre pendientes			1	0,009	0,009	1,34 ^{NS}
Global	9	1,255	0,446	0,225	8	0,066		
		Diferencia entre elevaciones			1	0,017	0,017	2,43 ^{NS}

En las dos primeras líneas de la tabla se ponen los datos calculados para ambas muestras. Las sumas de cuadrados de las desviaciones ($\sum d^2_{y,x}$), con $n-2$ grados de libertad, se calculan como se indicó al principio para cada muestra, así como los cuadrados medios. La prueba F de las varianzas residuales da un valor de:

$$F = 0,0097/0,0037 = 2,62 \text{ (gl} = 3,3) \text{ NS}$$

Al no mostrar diferencias significativas se asume que las varianzas residuales son homogéneas, por lo que sigue el proceso comparando las pendientes de las rectas. Para ello, se suman los gl y las SC de las desviaciones de regresión individuales, y se obtiene un cuadrado medio residual de 0,0067. A continuación se suman los gl y las sumas de cuadrados y de productos de ambas muestras, obteniéndose una SC de las desviaciones de 0,049 y un CM de 0,007. La diferencia $0,049-0,040=0,009$, con 1 gl, indica la contribución de la diferencia entre los coeficientes de regresión a la SC de las desviaciones. Se obtiene así el valor de F para la comparación de las pendientes, que resulta:

$$F = 0,009/0,0067 = 1,34 \text{ (gl} = 1,6) \text{ NS}$$

La falta de significación de las pendientes permite continuar el proceso comparando las elevaciones de la recta. Para ello se calculan las sumas de cuadrados y de productos considerando ambas muestras conjuntas formando una única recta, en el caso del ejemplo, con $n = 10$. Se obtiene así una SC de las desviaciones de 0,066 y la diferencia con 0,049 da un valor de 0,017, que sirve para calcular la F que permite comparar la elevación de la recta. En el ejemplo,

$$F = 0,017/0,007 = 2,43 \text{ (gl} = 1,7) \text{ NS}$$

No existe tampoco diferencia entre elevaciones, lo que permite concluir que ambas rectas no difieren.

Cálculo con el programa *Statistix (SX)*

Introducir los datos de ambas muestras creando una nueva variable denominada, por ejemplo, MUESTRA, que agrupa las variables en estudio y que en el ejemplo tomará los valores 1 y 2, que corresponden a los de la muestra del año 1 y a los del año 2. Los valores de la variable NFERTI se repiten para MUESTRA 1 y 2, pues no han variado, mientras que los de la variable NHOJA corresponden a MUESTRA 1 los de la Tabla 15.1 y a MUESTRA 2 los de la Tabla 15.2. Los datos quedarían de la siguiente forma:

MUESTRA	NFERTI	NHOJA
1	0	1.41
1	0.12	1.62
1	0.25	1.74
1	0.5	1.78
1	1	1.93
2	0	1.45
2	0.12	1.6
2	0.25	1.56
2	0.5	1.7
2	1	1.75

A continuación, seguir la secuencia usual:

Statistics >Linear Models> Linear Regresión

y provocar la salida de datos, que mostrará la regresión global considerando todos los datos para una misma recta.

```

Unweighted Least Squares Linear Regression of NHOJA

Predictor
Variables   Coefficient   Std Error       T           P
Constant      1.52126           0.04200           36.22         0.0000
NFERTI        0.35492           0.08153           4.35          0.0024

R-Squared           0.7031           Resid. Mean Square (MSE)   0.00834
Adjusted R-Squared 0.6660           Standard Deviation         0.09134

Source       DF           SS           MS           F           P
Regression       1           0.15810       0.15810       18.95         0.0024
Residual         8           0.06674       0.00834
Total            9           0.22484

Cases Included 10   Missing Cases 0
    
```

En el submenú *Results* seleccionar *Comparison of Regresión Lines*. En el cuadro que aparece indicar en *Group Variable* la variable que identifique los grupos a comparar, en este caso la variable MUESTRA, y pulsar OK. Aparece la siguiente salida:

Comparison of Regression Lines for NHOJA = NFERTI

MUESTRA	N	Intercept	Slope	MSE
1	5	1.53038	0.44282	0.00942
2	5	1.51213	0.26702	0.00371

	F	DF	P
Equality of Variances	2.54	3, 3	0.2322
Comparison of Slopes	1.48	1, 6	0.2700
Comparison of Elevations	2.51	1, 7	0.1568

En las primeras líneas aparecen las rectas de regresión correspondientes a cada muestra junto a los cuadrados medios de las desviaciones de regresión. A continuación aparecen las pruebas F para la homogeneidad de las varianzas residuales, la comparación de pendientes y la de elevaciones, junto con los grados de libertad correspondientes y la probabilidad. Las diferencias respecto a los cálculos expresados en la tabla manual se deben a las aproximaciones con los decimales.

CONSIDERACIONES FINALES

En los ejemplos que se han resuelto en este capítulo se ha supuesto un modelo en el que los valores de X eran fijos, pues son los tratamientos seleccionados en el experimento. Pero a veces el modelo puede ser aleatorio, de manera que se toman muestras de una población y luego se miden Y y X, de manera que no hay valores fijos de la variable independiente sino que estos varían de una muestra a otra. Por ejemplo, si en un muestreo de frutos se mide en cada uno el calibre y a la vez la infección causada por un determinado parásito, estamos en un caso de estos.

En estos supuestos, los métodos anteriormente expuestos son completamente válidos siempre que para cada X la correspondiente Y satisfaga las suposiciones a las que se ha hecho referencia anteriormente y, en particular, que exista una distribución normal de Y; que μ esté en la recta $\mu = \alpha + \beta x$; y que en cada población, $\sigma_{y,x}$ tenga el mismo valor.

CAPÍTULO 16

REGRESIÓN MÚLTIPLE

La regresión de Y en una sola variable independiente resulta, a veces, inadecuada porque puede haber otras variables que proporcionen más información sobre Y. En una muestra de n valores, la ecuación se escribe como:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_nX_n$$

donde b_1 es el coeficiente de regresión parcial de Y en X_1 y mide el cambio en Y cuando X_1 varía en una unidad, permaneciendo el resto de X_i inmutables.

Entre los usos principales de la regresión múltiple están: a) predecir mejor los valores de Y; b) elegir las X que mejor definen Y, de manera que se ahorran cálculos y medidas; y c) conocer las variables relacionadas con Y y ordenarlas según su importancia.

La regresión múltiple se complica conforme aumentan las variables en estudio, tanto en los cálculos, cada vez más complejos, como en la interpretación sobre la influencia de varias variables en la Y, cuando muchas de ellas podrían estar, a su vez, relacionadas entre sí. Se estudia en este capítulo la regresión con dos variables independientes.

CORRELACIÓN PARCIAL Y MÚLTIPLE

Si hay tres variables en estudio (Y, X_1 , X_2), tendremos siempre tres correlaciones simples entre ellas, cuyos coeficientes de correlación se calculan de la forma descrita en el capítulo 14:

$$r_{y x_1} = \frac{\sum y x_1}{\sqrt{(\sum y^2)(\sum x_1^2)}}$$

$$r_{y x_2} = \frac{\sum y x_2}{\sqrt{(\sum y^2)(\sum x_2^2)}}$$

$$r_{x_1 x_2} = \frac{\sum x_1 x_2}{\sqrt{(\sum x_1^2)(\sum x_2^2)}}$$

La correlación entre dos variables cuando la tercera la mantenemos constante, se denomina **correlación parcial**. El coeficiente de correlación parcial entre Y y X_1 manteniendo X_2 constante, es:

$$r_{y x_1, x_2} = \frac{r_{y x_1} - r_{y x_2} r_{x_1 x_2}}{\sqrt{(1 - r_{y x_2}^2)(1 - r_{x_1 x_2}^2)}}$$

y, análogamente, el coeficiente de correlación parcial entre Y y X₂ con X₁ constante, es:

$$r_{yX_2 \cdot X_1} = \frac{r_{yX_2} - r_{yX_1} r_{X_1X_2}}{\sqrt{(1 - r_{yX_1}^2)(1 - r_{X_1X_2}^2)}}$$

El coeficiente de correlación parcial se contrasta en las tablas con (n-3) grados de libertad. Estos coeficientes indican si la relación entre dos variables es o no dependiente de la otra. En el caso de r_{yX₁X₂}, por ejemplo, si resulta significativo quiere decir que parte de la relación entre Y y X₁ no depende de X₂.

El coeficiente de correlación múltiple R mide la relación de las tres variables entre sí, y se calcula de la siguiente forma:

$$R_{y \cdot X_1 X_2} = \sqrt{\frac{(r_{yX_1}^2 + r_{yX_2}^2 - 2r_{yX_1} r_{yX_2} r_{X_1X_2})}{(1 - r_{X_1X_2}^2)}}$$

El valor de R, al contrario que los anteriores, es siempre positivo y varía entre 0 y 1. Su valor es, al menos, tan grande como el mayor de los coeficientes simples o parciales.

Ejemplo

La Tabla 16.1 siguiente recoge datos de un ensayo realizado con la nectarina ‘Armking’ donde se trataba de estudiar la relación existente entre la producción de una rama principal, tomada como muestra del árbol completo, el número de ramos fructíferos de esa rama y su vigor, expresado como el perímetro de la rama.

Tabla 16.1. Producción, número de ramos fructíferos y vigor en once ramas principales de la nectarina ‘Armking’.

Rama N°	Producción (kg/rama)	Nº ramos fructíferos/rama	Vigor (perímetro de la rama, cm)
	Y	X ₁	X ₂
1	18,5	34	26
2	16,5	36	28
3	22,0	43	25
4	18,6	46	30
5	15,1	35	25
6	10,8	33	22
7	32,4	52	32
8	13,6	32	24
9	17,4	29	24
10	16,7	36	23
11	14,9	39	26
Suma	196,5	415	285
Media	17,86	37,73	25,91

Los coeficientes de correlación simple son:

$$r_{yx1} = 0,79^{**}$$

$$r_{yx2} = 0,75^{**}$$

$$r_{x1x2} = 0,81^{**}$$

Los coeficientes de correlación parcial:

$$r_{yx1.x2} = 0,48^{NS}$$

$$r_{yx2.x1} = 0,30^{NS}$$

El coeficiente de correlación múltiple :

$$R_{y.x1x2} = 0,81^{**}$$

Los resultados indican que todas las variables están correlacionadas significativamente ente sí, incluidas las dos variables independientes. La falta de significación de los coeficientes de correlación parciales indica que la relación que existe tanto entre Y y X₁ como entre Y y X₂, depende de la otra variable en estudio.

Cálculo con el programa *Statistix (SX)*

En el programa *Statistix* seguir la siguiente secuencia para el cálculo de los coeficientes de correlación parciales:

Statistics > Linear Models > Partial Correlations

En este punto indicar cuál es la variable dependiente (*Dependent Variable*), la variable con la que se quiere determinar la correlación (*Correlation Variable*) y aquella con la que se pretende contrastar la dependencia (*Adjust for Variables*). Pulsar OK para obtener el resultado. Para el caso de correlación parcial entre la producción y el número de ramos manteniendo el vigor como constante $r_{yx1.x2}$, el resultado aparece en la siguiente tabla:

Partial Correlations with PRO Controlled for VIGOR	
RAMOS	0.4787
Cases Included	11
Missing Cases	1

REGRESIÓN CON DOS VARIABLES INDEPENDIENTES

Si tenemos X_1 y X_2 dadas, los valores individuales de Y varían en torno a un plano de regresión en una distribución normal de media 0 y varianza σ^2 . En una muestra de n valores de Y , X_1 y X_2 , la ecuación es:

$$\hat{Y} = a + b_1X_1 + b_2X_2$$

El valor de a viene dado por la ecuación:

$$a = \bar{Y} - b_1\bar{X}_1 - b_2\bar{X}_2$$

Los b satisfacen las ecuaciones normales siguientes:

$$b_1 \sum X_1^2 + b_2 \sum X_1X_2 = \sum X_1Y$$

$$b_1 \sum X_1X_2 + b_2 \sum X_2^2 = \sum X_2Y$$

siendo x_1 , x_2 y y las desviaciones respecto de sus medias:

$$x_1 = X - \bar{X}_1; x_2 = X_2 - \bar{X}_2; y = Y - \bar{Y}.$$

Los sumatorios anteriores también pueden calcularse de la siguiente manera:

$$\sum x^2 = \frac{\sum X^2 - (\sum X)^2}{n}$$
$$\sum xy = \frac{\sum XY - (\sum X)(\sum Y)}{n}$$

En el ejemplo de la Tabla 16.1 tenemos:

$$\sum X_1^2 = (16117 - 415^2) / 11 = 460,18$$

$$\sum X_2^2 = (7475 - 285^2) / 11 = 90,91$$

$$\sum X_1X_2 = [10918 - (415)(285)]/11 = 165,73$$

$$\sum X_2Y = [5218,4 - (285)(196,5)] / 11 = 127,26$$

$$\sum X_1Y = [7716,4 - (415)(196,5)] / 11 = 302,99$$

$$\sum Y^2 = (3827,5 - 196,5^2) / 11 = 317,29$$

Sustituyendo estos valores en las ecuaciones anteriores, tenemos:

$$b_1 = 0,45 \quad b_2 = 0,58 \quad a = -14,14$$

La ecuación de regresión múltiple resulta:

$$\hat{Y} = -14,14 + 0,45 X_1 + 0,58 X_2$$

Los coeficientes de regresión parciales son próximos, lo que indica que para un aumento de una unidad en cualquiera de las variables independientes, el aumento esperado en la producción es parecido, aunque un cambio en el grosor de la rama parece aumentar la producción en mayor cantidad.

De la ecuación anterior pueden estimarse las \hat{Y} para cada rama y determinar las desviaciones $d = Y - \hat{Y}$. La suma de cuadrados de las desviaciones puede calcularse también de la forma:

$$\sum d^2 = \sum y^2 - \sum \hat{y}^2$$

con

$$\sum \hat{y}^2 = b_1 \sum x_{1,y} + b_2 \sum x_{2,y}$$

de manera que la varianza es:

$$s^2 = \frac{\sum d^2}{n - k}$$

siendo k el número de parámetros estimados en el ajuste, en nuestro caso a, b_1 y b_2 .

En el ejemplo:

$$\sum \hat{y}^2 = 0,45 * 302,99 + 0,58 * 127,26 = 210,15$$

$$\sum d^2 = 317,29 - 210,15 = 107,14$$

$$s^2 = 107,14 / (11-3) = 13,39 ; \quad y \quad s = 3,66$$

La prueba de hipótesis nula $\beta_1 = \beta_2 = 0$ se desarrolla como sigue:

$$s_{b_1} = \frac{s}{\sqrt{c_{11}}} \quad y \quad s_{b_2} = \frac{s}{\sqrt{c_{22}}}$$

donde c_{11} , c_{22} y $c_{12}=c_{21}$ son los multiplicadores de Gauss, cuyo cálculo es:

$$c_{11} = \frac{\sum x_2^2}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1x_2)^2} = 0,0063$$

$$c_{22} = \frac{\sum x_1^2}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1x_2)^2} = 0,032$$

$$c_{12} = c_{21} = -\frac{\sum x_1x_2}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1x_2)^2} = -0,0115$$

luego:

$$s_{b_1} = \frac{3,66}{\sqrt{0,0063}} = 0,29$$

$$s_{b_2} = \frac{3,66}{\sqrt{0,032}} = 0,65$$

La prueba t, con $n-k = 8$ gl, resulta:

$$t_1 = \frac{b_1}{s_{b_1}} = \frac{0,45}{0,29} = 1,55^{NS} \quad t_2 = \frac{b_2}{s_{b_2}} = \frac{0,58}{0,65} = 0,89^{NS}$$

El valor de $t_{0,05}$ en la Tabla A1 es 2,306 para 8 gl, superior a los t_1 y t_2 , por lo que no resultan significativos.

Realizando la prueba de significación mediante un análisis de varianza de la regresión, éste queda como sigue:

Fuente de Variación	gl	SC	CM	F
Total	n-1	SC _T		
Regresión múltiple	2	SC _R	SC _R /2	CM _R /CM _D
Desviaciones de Regresión	n-3	SC _D	SC _D /n-3	

$$n = 11$$

$$SC_T = \sum y^2$$

$$SC_R = \sum y^2 - \sum d^2$$

$$SC_D = \sum d^2$$

En el ejemplo, los cálculos quedan de la siguiente forma:

Fuente de Variación	gl	SC	CM	F
Total	10	317,29		
Regresión múltiple	2	210,15	105,07	7,85*
Desviaciones de Regresión	8	107,14	13,39	

El análisis de la varianza indica que existe una F significativa ($P \leq 0,05$) para la regresión múltiple, mientras que ninguno de los coeficientes de regresión parciales muestran significación. Esta situación puede deberse a la relación existente entre las variables, como se observó al calcular los coeficientes de correlación. También hay que tener en cuenta que aún queda una gran proporción de la suma de cuadrados en las desviaciones de regresión y, en consecuencia, no están explicados por la misma. Para despejar esta situación es conveniente descomponer la suma de cuadrados debida a la regresión múltiple en la debida a cada una de las variables, considerando el orden en el que se incluyen en el estudio, esto es, considerando en primer lugar el efecto de X_1 y, a continuación, el efecto adicional de X_2 , y luego al contrario en un análisis de la varianza distinto. Este procedimiento debería hacerse en todos los casos de estudio de la regresión múltiple, con independencia de cuales hayan sido los resultados anteriores.

Para el primer caso, el análisis de la varianza quedaría:

Fuente de Variación	gl	SC	CM	F
Total	$n-1$	SC_T		
Regresión debida a X_1	1	SC_{RX_1}	$\frac{SC_{RX_1}}{1}$	$\frac{CM_{RX_1}}{CM_{DX_1}}$
Desviaciones de la regresión simple	$n-2$	SC_{DX_1}	$\frac{SC_{DX_1}}{n-2}$	
Regresión adicional debida a X_2	1	SC_{RX_2}	$\frac{SC_{RX_2}}{1}$	$\frac{CM_{RX_2}}{CM_{DM}}$
Desviaciones de la regresión múltiple	$n-3$	SC_{DM}	$\frac{SC_{DM}}{n-3}$	

$$n = 11$$

$$SC_T = \sum y^2$$

$$SC_{RX_1} = r_{yx_1}^2 \left(\sum y^2 \right)$$

$$SC_{DX_1} = (1 - r_{yx_1}^2) \sum y^2$$

$$SC_{RX_2} = r_{yx_2 \cdot x_1}^2 (1 - r_{yx_1}^2) \sum y^2$$

$$SC_{DM} = \sum d^2 = (1 - R_{y \cdot x_1 x_2}^2) \sum y^2$$

En el ejemplo:

Fuente de Variación	gl	SC	CM	F
Total	10	317,29		
Regresión debida a X ₁	1	198,02	198,02	14,94**
Desviaciones de la regresión simple	9	119,27	13,25	
Regresión adicional debida a X ₂	1	10,73	10,73	0,80 ^{NS}
Desviaciones de la regresión múltiple	8	107,14	13,39	

Ha resultado significativa la regresión debida a X₁, y sin significación la regresión adicional de X₂, que representa una parte pequeña de la suma de cuadrados total. Nótese que $SC_{RX_1} + SC_{RX_2}$ (198,02+10,73=208,75) difiere del valor obtenido para la regresión múltiple en el análisis de varianza anterior (210,15). La diferencia es debida a la aproximación en los cálculos de tantos coeficientes.

Si consideramos ahora el efecto de X₂ en primer lugar y luego el efecto adicional de X₁, el análisis de varianza sería:

Fuente de Variación	gl	SC	CM	F
Total	n-1	SC _T		
Regresión debida a X ₂	1	SC _{RX₂}	$\frac{SC_{RX_2}}{1}$	$\frac{CM_{RX_2}}{CM_{DX_2}}$
Desviaciones de la regresión simple	n-2	SC _{DX₂}	$\frac{SC_{DX_2}}{n-2}$	
Regresión adicional debida a X ₁	1	SC _{RX₁}	$\frac{SC_{RX_1}}{1}$	$\frac{CM_{RX_1}}{CM_{DM}}$
Desviaciones de la regresión múltiple	n-3	SC _{DM}	$\frac{SC_{DM}}{n-3}$	

$$\begin{aligned}
 n &= 11 \\
 SC_T &= \sum y^2 \\
 SC_{RX_2} &= r_{yx_2}^2 \left(\sum y^2 \right) \\
 SC_{DX_2} &= (1 - r_{yx_2}^2) \sum y^2 \\
 SC_{RX_1} &= r_{yx_1 \cdot x_2}^2 (1 - r_{yx_2}^2) \sum y^2 \\
 SC_{DM} &= \sum d^2 = (1 - R_{y \cdot x_1 x_2}^2) \sum y^2
 \end{aligned}$$

En el ejemplo:

Fuente de Variación	gl	SC	CM	F
Total	10	317,29		
Regresión debida a X ₂	1	178,47	178,47	11,57**
Desviaciones de la regresión simple	9	138,81	15,42	
Regresión adicional debida a X ₁	1	31,98	31,98	2,39 ^{NS}
Desviaciones de la regresión múltiple	8	107,14	13,39	

Resulta ahora una F significativa para la regresión con X₂ y no significativa la regresión adicional de X₁. La mayor parte de la suma de cuadrados de la regresión se va con la variable considerada en primer lugar, lo que da idea de que el orden en el que se pongan las variables es de gran importancia y puede dar lugar a diferencias en las conclusiones.

Los resultados indican que ambas variables explican la Y, pero que están muy relacionadas, algo que ya conocíamos al calcular los coeficientes de correlación simple y parciales. Esto quiere decir que una sola variable basta para el estudio; X₁ (nº de ramos fructíferos) muestra algo más de correlación con Y, pero X₂ (perímetro de la rama) es más fácil de medir. Cualquiera de las dos sería elegible para explicar la Y, en cualquier caso.

Probando con X₁ exclusivamente, obtenemos la siguiente ecuación:

$$\hat{Y} = -6,97 + 0,66 X_1 ; r^2 = 0,63^{**}; t_1 = 3,9^{**}$$

y con X₂ la siguiente:

$$\hat{Y} = -18,4 + 1,40 X_2 ; r^2 = 0,56^{**}; t_2 = 3,39^{**}$$

Se observa que ambas regresiones se muestran significativas, tanto en el valor de las F como en las t, y que cualquiera de las dos variables puede servir para explicar la Y, como se ha comentado anteriormente. No obstante, con X₁ se obtiene un r² mayor, es decir, se explica un mayor porcentaje de la variabilidad de Y.

Los casos que se pueden presentar son muy variados y la interpretación correcta siempre es difícil y nunca ajena al conocimiento que el investigador posee sobre las variables estudiadas. Sin embargo, a modo de ejemplo, los resultados anteriores podrían haber sido de la siguiente manera:

	F
Regresión debida a X_1	14,82**
Regresión adicional de X_2	0,70 ^{NS}
Regresión debida a X_2	2,5 ^{NS}
Regresión adicional de X_1	10,53**

En este supuesto, la mayor suma de cuadrados va siempre asociada con la variable X_1 y muy poco con X_2 . Podría interpretarse con cierta facilidad que la variable X_2 aporta poco a la variación de Y , y podría eliminarse del estudio.

Otro caso frecuente puede darse como sigue:

	F
Regresión debida a X_1	15,43**
Regresión adicional de X_2	10,41**
Regresión debida a X_2	25,72**
Regresión adicional de X_1	3,70 ^{NS}

En el primer análisis de la varianza podría interpretarse que ambas variables independientes tienen peso en la explicación de Y , pero al colocar en primer lugar X_2 y estudiar luego la regresión adicional de X_1 , la mayor parte de la suma de cuadrados se va con X_2 , resultando que no existe regresión significativa con X_1 . Parece evidente que X_2 es la variable que está relacionada con la Y , y que la significación de la X_1 al ser considerada en primer lugar podría deberse a su relación con X_2 . Al colocarse en primer lugar, absorbe sumas de cuadrados debida a su relación con X_2 . El cálculo de los coeficientes de correlación ayudará a confirmar esa interpretación.

Cálculo con el programa *Statistix (SX)*

En el programa SX, una vez introducidos los datos experimentales, seguir la siguiente secuencia:

Statistics > Linear Models > Linear Regression

Al llegar a este punto, indicar cuál es la variable dependiente y cuales las independientes y luego pulsar OK. Aparece, en el caso del ejemplo, la siguiente tabla:

Unweighted Least Squares Linear Regression of PRO						
Predictor Variables	Coefficient	Std Error	T	P	VIF	
Constant	-14.1376	10.3861	-1.36	0.2105		
RAMOS	0.44913	0.29123	1.54	0.1616	2.9	
VIGOR	0.58114	0.65523	0.89	0.4010	2.9	
R-Squared	0.6620	Resid. Mean Square (MSE)			13.4057	
Adjusted R-Squared	0.5775	Standard Deviation			3.66139	
Source	DF	SS	MS	F	P	
Regression	2	210.040	105.020	7.83	0.0131	
Residual	8	107.246	13.406			
Total	10	317.285				
Cases Included 11		Missing Cases 1				

Los coeficientes permiten escribir la ecuación de regresión. Obsérvese la falta de significación de t_1 y t_2 , y la F significativa del análisis de varianza. $R^2 = 0,66^*$ es el coeficiente de determinación, que debe emplearse asociado a la regresión.

La descomposición de la suma de cuadrados de la regresión múltiple en la debida a cada variable según el orden en el que son consideradas, lo realiza el programa con el procedimiento *Stepwise*. Para ello seguir los pasos anteriores y cuando aparezca la tabla de resultados, en el submenú de la regresión, seguir la secuencia:

Results > Stepwise AOV Table

y aparece la tabla siguiente:

Stepwise Anlysis of Variance of PRO

Source	Individual SS	Cum DF	Cumulative SS	Cumulative MS	Adjusted R-Squared	Mallows' CP	P
Constant	3510.20						
RAMOS	199.494	1	199.494	199.494	0.5875	1.8	2
VIGOR	10.5455	2	210.040	105.020	0.5775	3.0	3
Residual	107.246	10	317.285	31.7285			
R-Squared		0.6620	Resid. Mean Square (MSE)		13.4057		
Adjusted R-Squared		0.5775	Standard Deviation		3.66139		

En la columna **Individual SS** aparece la descomposición de la suma de cuadrados cuando la variable Ramos (X_1) es considerada en primer lugar. Para obtener la descomposición al considerar primero la variable Vigor (X_2), repetir el proceso y al indicar cuales son las variables independientes, colocar Vigor en primer lugar. Obtenemos así la tabla siguiente:

Stepwise Anlysis of Variance of PRO

Source	Individual SS	Cum DF	Cumulative SS	Cumulative MS	Adjusted R-Squared	Mallows' CP	P
Constant	3510.20						
VIGOR	178.156	1	178.156	178.156	0.5128	3.4	2
RAMOS	31.8831	2	210.040	105.020	0.5775	3.0	3
Residual	107.246	10	317.285	31.7285			
R-Squared		0.6620	Resid. Mean Square (MSE)		13.4057		
Adjusted R-Squared		0.5775	Standard Deviation		3.66139		

El cálculo de las F correspondientes no las realiza el programa, pero se reduce a unas pocas divisiones.

REGRESIÓN CON MÁS DE DOS VARIABLES INDEPENDIENTES

Al aumentar el número de variables independientes los cálculos se hacen más complejos e inmanejables. Los métodos descritos pueden generalizarse para más de dos variables independientes, es decir, el álgebra no varía, pero la aritmética para el cálculo de coeficientes se complica. Para hacerse una idea, en la regresión lineal, con una sola variable independiente, se calcula un coeficiente de correlación; con dos variables independientes el número de coeficientes aumenta a siete (tres simples, tres parciales y uno múltiple); con cuatro variables independientes el número aumenta a 25; y con cinco variables independientes a 81. Si se considera ahora que con dos variables independientes había que realizar dos análisis de varianza, en función de la variable que se consideraba en primer lugar, con tres variables el número aumenta a seis y con n variables a $n!$ [$n(n-1)(n-2)(n-3)\dots \cdot 3 \cdot 2 \cdot 1$]. El problema que se plantea es cómo elegir el conjunto de variables independientes que mejor explique la ecuación. Esto es siempre difícil, pero para ayudar a los cálculos se requiere necesariamente el uso de programas informáticos que aproximen la solución al problema.

El programa *Statistix* incluye dos procedimientos que permiten seleccionar los modelos más adecuados para el estudio de la regresión múltiple. En otras palabras, todas las variables independientes que se puedan incluir en el estudio no tienen el mismo peso para explicar la variación de la variable dependiente, y algunas puede que no influyan en esa variabilidad. Lo que se trata es de elegir aquellas variables que influyan en la dependiente para realizar posteriormente el estudio de la regresión con las que tienen influencia y que no estén relacionadas entre sí.

Para ilustrar los procedimientos se incluye el ejemplo recogido en la Tabla 16.2. Se trata de un estudio en el que se pretendía conocer si el contenido de nitrógeno mineral en muestras de suelos cultivados de olivar, dependía de variables tales como el tratamiento de nitrógeno al suelo, la profundidad de la toma de muestras o el grosor de las partículas del suelo al realizar los análisis. Para ello, en parcelas tratadas y no tratadas con nitrógeno, se tomaron muestras de suelo a cinco profundidades distintas; una vez las muestras en el laboratorio, se trituraron y se tamizaron para conseguir tres tamaños distintos de partículas antes de proceder al análisis del nitrógeno.

Tabla 16.2. Contenido mineral de nitrógeno en el suelo, nitrógeno aplicado, profundidad del muestreo y tamaño de la partícula de suelo previo al análisis.

N mineral (mg N/kg suelo)	Tratamiento (kg N/árbol)	Grosor (mm)	Profundidad (cm)
26,34	0	0,5	20
28,25	0	0,5	30
26,24	0	0,5	50
37,28	0	0,5	70
42,04	0	0,5	90
21,15	0	1	20
34,38	0	1	30
26,67	0	1	50
28,39	0	1	70
40,5	0	1	90
35,53	0	2	20
22,3	0	2	30
20,75	0	2	50
20,35	0	2	70
47,32	0	2	90
44,75	1	0,5	20
74,16	1	0,5	30

N mineral (mg N/kg suelo)	Tratamiento (kg N/árbol)	Grosor (mm)	Profundidad (cm)
62,11	1	0,5	50
57,41	1	0,5	70
60,78	1	0,5	90
41,54	1	1	20
61,1	1	1	30
63,67	1	1	50
62,31	1	1	70
65,21	1	1	90
51,8	1	2	20
74,04	1	2	30
60,69	1	2	50
56,52	1	2	70
69,23	1	2	90

El **primer procedimiento** es el denominado *Best Subset Regression Models*, que permite, de una forma sencilla y rápida, elegir aquellas variables independientes que mejor explican la dependiente y, en definitiva, elegir el mejor modelo para el estudio. Para activar el procedimiento se sigue la siguiente secuencia:

Statistics > Linear Models > Best Subset Regressions

Indicar cuál es la variable dependiente y las demás incluirlas en la casilla *Non-forced Indep. Vars.*, a menos que se quiera que una de ellas forme parte del modelo, en cuyo caso se incluye dentro de *Forced Indep. Variables*. Pulsar OK y aparece la siguiente tabla:

Best Subset Regression Models for N_MINERAL

Unforced Independent Variables: (A)GROSOR (B)PROF (C)TRAT
 3 "best" models from each subset size listed.

P	CP	Adjusted		Resid SS	Model Variables
		R Square	R Square		
1	104.2	0.0000	0.0000	8892.87	Intercept Only
2	6.8	0.7429	0.7517	2207.81	C
2	99.4	0.0177	0.0516	8434.06	B
2	106.2	-0.0357	0.0000	8892.62	A
3	2.0	0.7888	0.8033	1749.01	B C
3	8.8	0.7334	0.7518	2207.56	A C
3	101.4	-0.0186	0.0516	8433.81	A B
4	4.0	0.7807	0.8034	1748.76	A B C

Cases Included 30 Missing Cases 0

La columna de la derecha indica los posibles modelos que pueden elaborarse con las variables introducidas; en el ejemplo se considera cada una de las variables por separado, esto es, las regresiones lineales con Grosor, Profundidad o Tratamiento, la regresión múltiple considerando las variables de dos en dos, o la regresión múltiple incluyendo las tres variables en el modelo. En las restantes columnas se muestra el número de parámetros que se incluyen en el modelo (p), el estadístico C_p de Mallows, los valores de R^2 y del R^2 ajustado y la suma de cuadrados de los residuales para cada modelo. El estadístico C_p y los R^2 sirven para elegir el mejor modelo. C_p es útil para eliminar variables que contribuyen poco al modelo y los mejores modelos presentan un valor de C_p cercano o menor que el número de parámetros en el modelo (p). El coeficiente de determinación R^2 sería un buen criterio para seleccionar modelos, pero R^2 suele aumentar al aumentar el número de variables incluidas, aunque contribuyan poco, de forma que el R^2 mayor siempre corresponderá al modelo que incluya más variables; en el ejemplo, el de la regresión múltiple con las tres variables independientes. Téngase en cuenta que R^2 indica el porcentaje de la suma de cuadrados que pueden explicar la variable dependiente que, lógicamente, aumenta al introducir más variables. El R^2 ajustado que se incluye en la columna tercera corrige ese problema y es un criterio indicado para elegir el modelo.

En el ejemplo, el R^2 ajustado mayor corresponde al modelo que incluye tan solo el tratamiento y la profundidad como variables independientes. Si se utiliza el estadístico C_p se llega a la misma conclusión, pues aporta un valor de $2,0 < 3$. Por consiguiente, se puede deducir que el mejor modelo es el que utilizaría una regresión múltiple con dos variables independientes, el tratamiento de nitrógeno al suelo y la profundidad de muestreo. El tamaño de la partícula parece no influir en los valores aportados sobre el contenido de nitrógeno en muestras de suelo.

El **segundo procedimiento** es el denominado *Stepwise Linear Regression*, que resulta de particular utilidad cuando el número de variables independientes a estudiar es alto. El problema es que no siempre el mejor modelo seleccionado es el que se elegiría al utilizar el R^2 ajustado y el C_p de Mallows. Por eso, si el número de variables a utilizar es bajo o moderado, estos dos valores dan un buen resultado. Para activar el procedimiento se sigue la siguiente secuencia:

Statistics > Linear Models > Stepwise Linear Regression

Indicar cuál es la variable dependiente y las demás incluirlas en la casilla *Non-forced Indep. Vars.*, a menos que se quiera que una de ellas forme parte del modelo, en cuyo caso se incluye dentro de *Forced Indep. Variables*. Pulsar OK y aparece la siguiente tabla:

```

Stepwise Linear Regression of N_MINERAL
Unforced Variables: GROSOR PROF TRAT
P to Enter 0.0500
P to Exit 0.0500

                                G
                                R
                                O P T
                                S R R
                                O O A
Step      R Sq      MSE      P      R F T
  1      0.0000      306.651      .      . . .
  2      0.7517      78.8505      0.0000 + . . C
  3      0.8033      64.7781      0.0129 + . B C

Resulting Stepwise Model
Variable      Coefficient      Std Error      T      P      VIF
Constant      22.5596      3.63579      6.20      0.0000
PROF      0.15269      0.05737      2.66      0.0129      1.0
TRAT      29.8553      2.93889      10.16      0.0000      1.0

Cases Included      30      R Squared      0.8033      MSE      64.7781
Missing Cases      0      Adjusted R Sq      0.7888      SD      8.04848

Variables Not in the Model
Correlations
Variable      Multiple      Partial      T      P
GROSOR      0.0000      0.0120      0.06      0.9517
    
```

Al principio la tabla muestra los valores de R^2 y la significación (valor de P) en los pasos considerados para la selección; en el ejemplo, para la variable Tratamiento únicamente y para la combinación de ésta con la Profundidad. A continuación, muestra el análisis de la regresión múltiple con esas dos variables, el modelo seleccionado por el procedimiento. Al final incluye las variables no consideradas con los coeficientes de correlación y el valor de t. El coeficiente de correlación parcial corresponde al de la variable dependiente N_MINERAL con el GROSOR, manteniendo constantes las otras dos variables. Con este procedimiento se ha llegado a la misma conclusión que con el anterior, pero ya aporta la ecuación de regresión y la significación de las t para el modelo seleccionado. No obstante, sería conveniente, una vez elegidas las variables, realizar el estudio de la regresión múltiple conforme se ha indicado en este capítulo.

CONSIDERACIONES SOBRE LA REGRESIÓN MÚLTIPLE

La regresión múltiple presenta importantes limitaciones sobre lo que el análisis puede proporcionar en estudios de observación. En primer lugar, nunca estamos seguros de que no existen otras variables relacionadas con Y, que no se han considerado por desconocimiento, por dificultad de la medida o porque han sido consideradas de menor importancia en base a conocimientos previos. Pero todas ellas pueden significar una proporción importante de

la variación de Y . Si esas variables están relacionadas con alguna de las incluidas en el estudio, mejorará la predicción de Y , pero si no lo están la predicción de Y puede ser parcial.

Otro problema que se presenta hace referencia a la precisión con la que un coeficiente de regresión muestral, por ejemplo b_1 , estima al poblacional β_1 . En la regresión múltiple el valor de cualquier coeficiente de regresión depende de las otras variables, lo que puede afectar a la estimación del poblacional. En estudios de campo es frecuente realizar experimentos en explotaciones particulares. Si tomamos muestras en distintas explotaciones encontraremos diferencias entre ellas en el manejo del cultivo, algo que puede afectar directamente a las variables que queremos estudiar. No es lo mismo, pues, realizar muestreos en explotaciones competentes que en otras que no lo son. Si las muestras se toman en un tipo u otro de explotación, los coeficientes b_i sobreestimarán o subestimarán, según el caso, a los poblacionales β_i .

En muchos estudios lo que realmente interesa es conocer la variación de Y ante cambios dados en las X . En concordancia con lo comentado anteriormente, la regresión múltiple no puede servirnos para alcanzar ese objetivo. El análisis de la regresión múltiple es útil para seleccionar las variables más interesantes relacionadas con Y que permitan posteriormente plantear ensayos específicos para estudiar los cambios en las variables y su efecto en Y ; en definitiva, seleccionar tratamientos cuyos efectos queremos medir. La cuestión es como seleccionar las variables más importantes que afectan a Y . No siempre hay una respuesta satisfactoria. Si las X son independientes, los coeficientes de regresión parcial pueden utilizarse para ordenar las variables. En cualquier caso, no debe olvidarse que el conocimiento previo de las variables por el investigador puede ayudar a seleccionarlas de una manera racional.

CAPÍTULO 17

REGRESIÓN CURVILÍNEA

La relación más sencilla entre dos variables es la regresión lineal y, por ello, es la de mayor utilidad. Pero no todas las variables se relacionan de esa manera tan simple y la forma en la que se explica mejor la relación entre ellas es mediante una curva. El análisis de la regresión curvilínea, o el ajuste de curvas, es el proceso matemático adecuado para estudiar la relación entre dos variables que no se ajusta a la linealidad.

Una regresión curvilínea puede interesar por varias razones. En primer lugar, la relación entre dos variables puede ser lineal dentro de un intervalo, pero no más allá de él, por lo que puede interesar estudiar la verdadera relación entre ellas. Por ejemplo, la producción está relacionada linealmente con la densidad de plantación, esto es, con el número de árboles por hectárea, hasta un límite a partir del cuál el aumento de árboles crea fenómenos de competencia que hacen disminuir la producción. El intervalo de estudio que suele interesar es el primero, hasta la aparición de la competencia, pero no muestra la verdadera relación entre producción y densidad de plantación.

En otros casos interesa probar si los datos se ajustan a curvas o modelos propuestos con anterioridad o, simplemente, encontrar un ajuste a los datos que elimine imprecisiones de la regresión lineal o bien encontrar una curva que muestre la relación natural de los datos. Pero, en general, y dada las complejas relaciones que pueden encontrarse al tratar con datos biológicos, el investigador trata de buscar con la regresión curvilínea una tendencia que explique la relación entre dos variables.

El número de curvas que pueden expresarse por ecuaciones matemáticas es casi ilimitado, por lo que la elección de la curva que se desea ajustar puede ser un problema complejo. Pero muchos de esos ajustes no resultan de utilidad en la práctica, pues suelen mostrar una relación tan compleja que resulta difícil extraer conclusiones. Sin embargo, en estudios agronómicos la elección puede resumirse en algunas curvas sencillas que permitan obtener alguna conclusión sobre la relación entre las variables.

CURVAS DE TIPO LOGARÍTMICO Y EXPONENCIAL

Las curvas de este tipo pueden transformarse en líneas rectas mediante la toma de logaritmos de X , de Y o de ambas variables. Con los datos transformados se realiza una regresión lineal en la forma indicada en el capítulo 15, y la ecuación de la recta se convierte en la forma original mediante la toma de antilogaritmos.

La elección del tipo de curva puede intentarse realizando la representación gráfica de los datos, representando en primer lugar los datos originales para observar la curvilinealidad de los mismos y luego los datos transformados de una u otra variable. Puede apreciarse, a

veces, que los datos originales se ajustan a una línea recta incluso con un r^2 alto, pero por las razones apuntadas anteriormente puede interesar intentar un ajuste más preciso con algunas de las curvas de este tipo.

Curvas de tipo logarítmico

La representación gráfica de estas curvas se muestra en la Fig.17.1, cuya tendencia varía según el valor de b .

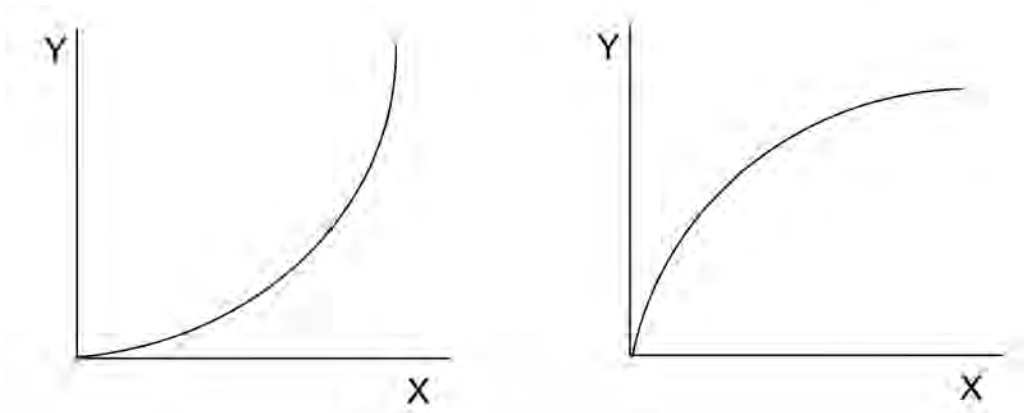


Fig 17.1. Curvas de tipo logarítmico.

La forma general de la ecuación de esas curvas, también denominadas de potencia, es:

$$Y = aX^b$$

Aplicando logaritmos tenemos:

$$\log Y = \log a + b \log X$$

y denominando $\log Y = Y'$, $\log a = a'$ y $\log X = X'$, tenemos la siguiente ecuación de una recta:

$$Y' = a' + b X'$$

En consecuencia, para analizar los datos se transforman los correspondientes a ambas variables en logaritmos, se realiza el ajuste de una regresión lineal como se indica en el Capítulo 15 y, posteriormente, se toman antilogaritmos para tener la ecuación de la curva de potencia. Como solo los números positivos tienen logaritmos, la ecuación no tiene sentido para valores negativos en ambas variables.

Ejemplo

Los datos de la Tabla 17.1 proceden de un ensayo cuyo objetivo era el estudio de la salida del reposo invernal de variedades de melocotonero. Para ello, se registraron el número de horas de frío (horas bajo 7° C) que se acumulaban a lo largo del invierno y, periódicamente, se tomaron ramos mixtos de los árboles, se colocaron en una cámara a 25° C para provocar el crecimiento de las yemas y, al cabo de un tiempo, se midió el peso seco de 10 yemas en cada tratamiento.

Tabla 17.1. Peso seco de yemas de melocotonero en relación con las horas-frío acumuladas en condiciones de campo y tras la exposición a 25° C durante un periodo de tiempo. Se incluyen las transformaciones logarítmicas de ambas variables.

Peso seco de 10 yemas (mg)	Horas-frío (horas<7°C)				
Y	X	Log Y	Log X	Ln Y	Ln X
3	50	0,48	1,70	1,10	3,91
4	100	0,60	2,0	1,39	4,60
6	200	0,78	2,30	1,79	5,30
8	300	0,90	2,48	2,08	5,70
14	350	1,15	2,54	2,64	5,86
30	400	1,48	2,60	3,40	5,99
71	450	1,85	2,65	4,26	6,11

La representación gráfica de los datos originales (Fig. 17.2) muestra una clara curvilinealidad, por lo que interesa un ajuste curvilíneo. No obstante, si se realiza una regresión lineal con los datos originales obtenemos la siguiente ecuación:

$$Y = -13,1 + 0,12 X \quad r^2 = 0,58^*$$

Hay una regresión lineal significativa al nivel del 5% con los datos originales. A veces esto es suficiente para el investigador que solo desea saber si hay relación entre las variables. Pero en este caso conviene probar otros ajustes.

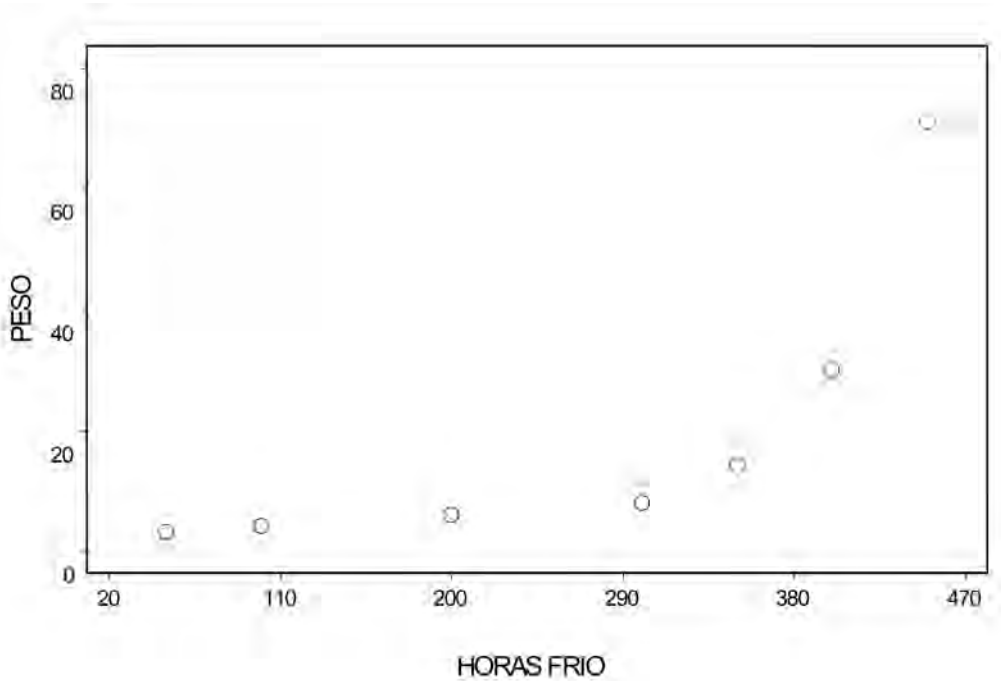


Fig 17.2. Relación entre el peso seco de yemas de melocotonero y las horas-frío acumuladas durante el invierno.

La regresión lineal

$$\log Y = \log a + b \log X$$

da la siguiente ecuación:

$$Y' = -1,71' + 1,18 X'$$

Tomando antilogaritmos, queda la ecuación de potencia original:

$$Y = 0,02 X^{1,18} \quad r^2 = 0,72^*$$

La ecuación de potencia ha resultado igualmente significativa, pero el aumento del coeficiente de determinación ha sido apreciable, indicando que esta ecuación representa mejor la relación entre las variables.

Curvas de tipo exponencial

La representación gráfica de estas curvas se muestra en la Fig. 17.3.

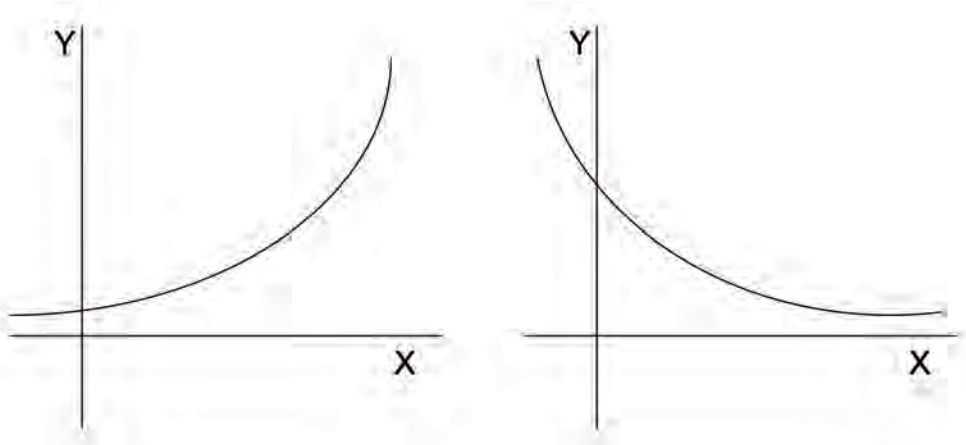


Fig. 17.3. Curvas de tipo exponencial.

La forma general de la ecuación de este tipo de curvas es:

$$Y = ab^X \quad \text{o bien} \quad Y = ae^{bX}$$

siendo e la base de los logaritmos neperianos. Las dos ecuaciones muestran la misma relación, bien expresada en logaritmos comunes de base 10 o naturales de base e . Tomando logaritmos en la primera ecuación, tenemos:

$$\log Y = \log a + X \log b$$

y haciendo $\log Y = Y'$, $\log a = a'$ y $\log b = b'$, tenemos la relación lineal:

$$Y' = a' + b' X$$

luego solo hay que transformar en logaritmos los datos de Y , quedando los originales de X .

La ecuación de la recta que se obtiene con los datos del ejemplo anterior es:

$$Y' = 0,22' + 0,003' X$$

Tomando antilogaritmos tenemos la ecuación:

$$Y = 1,66(1,007^X) \quad r^2 = 0,89^{**}$$

La curva exponencial parece ajustarse mucho mejor a los datos, pues ha aumentado tanto r^2 como el nivel de significación, que ha pasado del 5% al 1%. El investigador podía ser consciente de esto y haber probado directamente este ajuste, pues los fenómenos de crecimiento, como es el caso del ejemplo, suelen ajustarse a curvas de tipo exponencial. Es uno de los casos en los que el ajuste curvilíneo se realiza para comprobar que los datos obtenidos se ajustan a modelos conocidos o propuestos con anterioridad.

Si realizamos los cálculos tomando como base los logaritmos neperianos, tenemos que:

$$\ln Y = \ln a + bX \quad \ln e = \ln a + bX$$

$$Y' = a' + bX$$

Como en el caso de los logaritmos de base 10, solo hay que hacer transformaciones con los datos de Y. La ecuación de la recta que se obtiene es:

$$Y' = 0,52' + 0,007 X$$

Tomando antilogaritmos:

$$Y = 1,68 e^{0,007X} \quad r^2 = 0,89^{**}$$

Obsérvese que tanto r^2 como el nivel de significación son idénticos en ambos casos.

Otras curvas de tipo logarítmico o exponencial

La representación gráfica de este tipo de curvas se muestra en la Fig. 17.4.

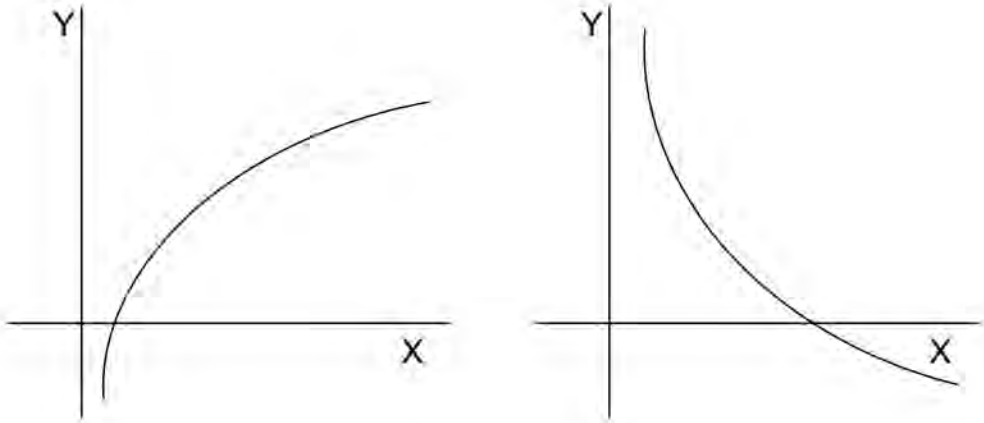


Fig. 17.4. Otras curvas de tipo logarítmico o exponencial.

La forma general de la ecuación es:

$$e^Y = a X^b$$

Procediendo como en los casos anteriores, tenemos:

$$Y \ln e = Y = \ln a + b \ln X$$

$$Y = a' + b X'$$

En este caso hay que transformar los datos originales de X y mantener intactos los de Y. La regresión lineal con los datos del ejemplo resulta:

$$Y = -81,0' + 18,7 X'$$

y tomando antilogaritmos tenemos la ecuación:

$$e^Y = (6,64 \cdot 10^{-36}) X^{18,7} \quad r^2 = 0,39^{NS}$$

La curva no se ajusta a los datos, pues la regresión ha resultado no significativa.

CURVAS DE TIPO POLINÓMICO

Este tipo de curvas tiene por ecuación general:

$$Y = a + bX + cX^2 + dX^3 + \dots + nX^n$$

en la que pueden incluirse tantos términos como se deseen, teniendo en cuenta que el grado del polinomio será, a lo sumo, uno menos que el número de pares de observaciones. La representación gráfica de estas curvas se muestra en la Fig. 17.5, y toman forma dependiendo del número de términos que se incluyan. Algunas curvas, entre las que se incluye la recta, toman nombre especiales muy conocidos.

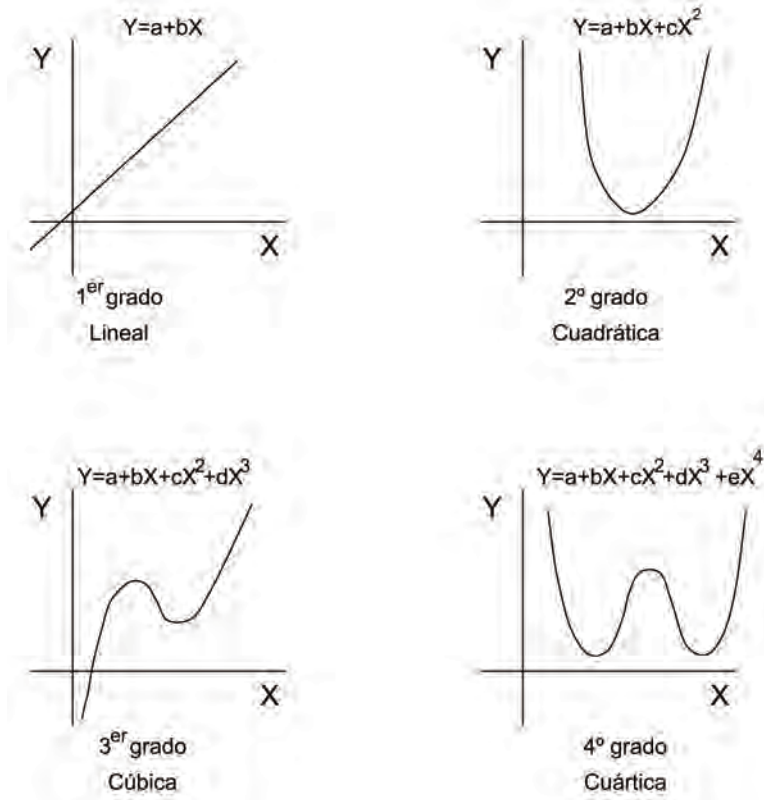


Fig. 17.5. Curvas de tipo polinómico.

Salvo que se desee ajustar unos datos a un modelo o tipo de curva determinado, como puede suceder con las logarítmicas o exponenciales, se puede decir que el polinomio es la expresión más utilizada para describir la relación entre dos variables, pues casi siempre es posible calcular una curva polinómica que se ajuste a los datos, si el número de pares de observaciones es elevado y se puedan añadir términos hasta lograr el ajuste. En la práctica, sin

embargo, rara vez resultan prácticas ecuaciones superiores al tercer o cuarto grado pues las curvas, como se aprecia en la Fig. 17.5, se muestran serpenteantes desde esos grados polinómicos, y suelen carecer de sentido para interpretar datos agronómicos.

El cálculo de la ecuación de regresión se basa en buscar la curva de grado tal que haga mínima la suma de cuadrados de las desviaciones, algo similar a lo realizado con anterioridad. Desde un punto de vista práctico, deseáramos la curva de menor grado que aporte el R^2 más elevado. Los coeficientes de la ecuación se calculan a partir de las ecuaciones siguientes:

$$\begin{aligned}
 a n + b \sum X + c \sum X^2 + d \sum X^3 + \dots &= \sum Y \\
 a \sum X + b \sum X^2 + c \sum X^3 + d \sum X^4 + \dots &= \sum XY \\
 a \sum X^2 + b \sum X^3 + c \sum X^4 + d \sum X^5 + \dots &= \sum X^2 Y \\
 a \sum X^3 + b \sum X^4 + c \sum X^5 + d \sum X^6 + \dots &= \sum X^3 Y \\
 &\dots\dots\dots
 \end{aligned}$$

siendo n el número de observaciones.

Para resolver la ecuación de una línea recta se necesitan solamente los dos primeros términos de las dos primeras ecuaciones, para la de una parábola los tres primeros de las tres primeras ecuaciones, y así sucesivamente.

El coeficiente de determinación puede calcularse de la siguiente forma:

$$R^2 = \frac{a \sum Y + b \sum XY + c \sum X^2 Y + d \sum X^3 Y + \dots - \frac{1}{n} (\sum Y)^2}{\sum Y^2 - \frac{1}{n} (\sum Y)^2}$$

Si \hat{Y}_L , \hat{Y}_Q e \hat{Y}_C representan los valores estimados de la Y en el caso de una regresión lineal, cuadrática o cúbica, respectivamente,

$$\sum d_L^2 = \sum (Y - \hat{Y}_L)^2; \quad \sum d_Q^2 = \sum (Y - \hat{Y}_Q)^2 \quad \text{y} \quad \sum d_C^2 = \sum (Y - \hat{Y}_C)^2$$

son las sumas de los cuadrados de las desviaciones, y aquella que resulte mínima corresponde a la ecuación buscada.

Expresando los datos en el análisis de varianza, tenemos:

Fuente de Variación	gl	SC	Porcentaje explicado de la SC
Total	n-1	SC _T	
Regresión lineal	1	SC _L	$\frac{SC_L}{SC_T} 100$
Desviaciones de la regresión lineal	n-2	SC _{DL}	
Regresión cuadrática	1	SC _Q	$\frac{SC_L + SC_Q}{SC_T} 100$
Desviaciones de la regresión cuadrática	n-3	SC _{DQ}	
Regresión cúbica	1	SC _C	$\frac{SC_L + SC_Q + SC_C}{SC_T} 100$
Desviaciones de la regresión cúbica	n-4	SC _{DC}	

$$SC_T = \sum y^2 = \sum (Y - \bar{Y})^2$$

$$SC_L = \sum y^2 - \sum d^2_L$$

$$SC_{DL} = \sum d^2_L$$

$$SC_Q = \sum d^2_L - \sum d^2_Q$$

$$SC_{DQ} = \sum d^2_Q$$

$$SC_C = \sum d^2_Q - \sum d^2_C$$

$$SC_{DC} = \sum d^2_C$$

El porcentaje de la suma de cuadrados totales explicado por cada regresión, esto es, el coeficiente de determinación, aumenta al aumentar el grado del polinomio, lo que podría indicar que la de mayor grado es siempre la mejor. Para evitar esto hay que probar la significación para cada curva mediante la prueba F.

Ejemplo

Los datos de la Tabla 17.2 proceden de un ensayo con variedades de melocotonero en el que se estudiaba el efecto del anillado del tronco en la producción y calidad de los melocotones. Como la anchura de la herida producida por el anillado puede afectar a la salud de los árboles si cicatriza mal, se midió el porcentaje de circunferencia del tronco en la que se apreciaba formación de callo y, en consecuencia, se producía una cicatrización satisfactoria. Para la variedad ‘Springtime’, que resultó una de las que mejor toleraban esta práctica, los resultados fueron los siguientes:

Tabla 17.2. Relación entre la anchura del anillado del tronco en la variedad de melocotonero ‘Springtime’ y la formación de callo.

Anchura del anillado (mm)	Formación de callo (% circunferencia cicatrizada)
X	Y
0	100
3	100
6	98
9	89
12	51

Resolviendo las ecuaciones anteriores para el cálculo de los coeficientes, obtenemos las siguientes ecuaciones y sus correspondientes R²:

Regresión lineal: $\hat{Y} = 109,4 - 3,63 X$ $r^2 = 0,676$

Regresión cuadrática: $\hat{Y} = 97,54 + 4,27 X - 0,66 X^2$ $R^2 = 0,956$

Regresión cúbica: $\hat{Y} = 100,24 - 2,18 X + 0,84 X^2 - 0,08 X^3$ $R^2 = 0,997$

El coeficiente de determinación aumenta al aumentar el grado del polinomio, pues la suma de cuadrados explicada respecto al total aumenta, como se demostró más arriba. Por ello hay que probar la significación para cada grado del polinomio y elegir, en consecuencia, la curva que mejor interesa para los datos. El análisis de varianza con los datos de la Tabla 17.2 resulta:

Fuente de Variación	gl	SC	Porcentaje de SC	Porcentaje explicado (R ²)
Total	4	1757,2		
Regresión lineal	1	1188,1	67,6	67,6
Desviaciones de la regresión lineal	3	569,1		
Regresión cuadrática	1	492,07	28,0	67,6+28,0 = 95,6
Desviaciones de la regresión cuadrática	2	77,03		
Regresión cúbica	1	72,9	4,1	67,6+28,0+4,1 = 99,7
Desviaciones de la regresión cúbica	1	4,13		

La regresión lineal acumula la mayor parte de la suma de cuadrados totales, pero solo explica el 67,6 % del total. Si calculamos la F para la regresión lineal, tenemos:

$$F_L = \frac{1188,1}{569,1/3} = 6,26^{NS}$$

que no resulta significativa. La adición de un término al polinomio explicaría el 95,6 % de la suma de cuadrados totales, y la prueba F para la regresión cuadrática resulta:

$$F_Q = \frac{(1188,1+492,07)/2}{77,03/2} = 21,81^*$$

que muestra significación al 5%. La adición de un término más, y último pues ya no quedan suficientes grados de libertad para seguir añadiendo términos, explicaría hasta el 99,7% de la suma de cuadrados totales. Si realizamos la prueba F para la regresión cúbica, tenemos:

$$F_C = \frac{(1188,1+492,07+72,9)/3}{4,13} = 141,49^{NS}$$

que no resulta significativa (aunque la F es muy elevada, solo dispone de 1 grado de libertad en el denominador). En realidad solo ha añadido un 4,1% al caso anterior y no ha resultado significativo. Aún habiéndolo sido, el investigador debe decidir si esa pequeña aportación en la explicación de Y tiene algún sentido como para elegir la curva de mayor grado. El cálculo

de la F del componente cúbico aislado ayudaría a ello; $F = \frac{72,9}{4,13} = 17,65$ resulta, lógicamente en este caso, también no significativo.

El cálculo de las F de los componentes aislados puede ayudar a elegir la curva en el caso de que todas las regresiones probadas, por ejemplo lineal, cuadrática y cúbica, resultasen significativas. Para ello se contrasta el cuadrado medio del componente con el de sus respectivas desviaciones. En el caso del ejemplo tenemos que para el componente lineal $F=6,26^{NS}$

y para la cuadrática $F = \frac{492,07}{77,03/2} = 12,78^{NS}$. En todos los casos la F ha resultado no significativa para los componentes aislados; solo la regresión cuadrática, que suma los efectos de los componentes lineal y cuadrático, ha dado significación. La relación entre las variables en estudio en este ejemplo puede explicarse, pues, mediante una relación cuadrática cuya ecuación se detalló anteriormente.

El cálculo de las F de los componentes aislados puede ayudar a elegir la curva en el caso de que todas las regresiones probadas, por ejemplo lineal, cuadrática y cúbica, resultasen significativas. Para ello se contrasta el cuadrado medio del componente con el de sus respectivas desviaciones. En el caso del ejemplo tenemos que para el componente lineal $F=6,26^{NS}$ y para la cuadrática $F = \frac{492,07}{77,03/2} = 12,78^{NS}$. En todos los casos la F ha resultado no significativa para los componentes aislados; solo la regresión cuadrática, que suma los efectos de los componentes lineal y cuadrático, ha dado significación. La relación entre las variables en estudio en este ejemplo puede explicarse, pues, mediante una relación cuadrática cuya ecuación se detalló anteriormente.

CÁLCULO CON EL PROGRAMA *Statistix* (SX)

En el programa SX, una vez introducidos los datos originales, hay que crear las “nuevas variables” X^2 y X^3 transformando los datos ($X^2 = X * X$ y $X^3 = X * X * X$). Luego se sigue el procedimiento explicado para la regresión múltiple señalando X, X^2 y X^3 como variables independientes, aunque en realidad no son tales, pero con la precaución de que el orden debe ser de acuerdo con el exponente de las X (no puede ir X^3 antes que X^2 , por ejemplo). Como el proceso es rápido puede calcularse la regresión lineal, la cuadrática y la cúbica y elegir de acuerdo con la significación de las F, o bien obtener la tabla *Stepwise AOV Table* para descomponer la suma de cuadrados en las debidas a la lineal, cuadrática y cúbica para seguir el proceso explicado más arriba. Esto es de interés cuando hay que calcular las F de los componentes aislados.

Calculemos a modo de ejemplo la regresión cúbica. Siguiendo la secuencia

Statistics > Linear Models > Linear Regresión

se indica cual es la variable dependiente y cuales las independientes (en este caso X, X^2 y X^3), se pincha en OK y aparece:

Unweighted Least Squares Linear Regression of CALLO					
Predictor Variables	Coefficient	Std Error	T	P	VIF
Constant	100.243	2.01732	49.69	0.0128	
ANILLADO	-2.17857	1.71062	-1.27	0.4238	63.8
ANILLADO2	0.84127	0.36203	2.32	0.2587	447.4
ANILLADO3	-0.08333	0.01983	-4.20	0.1487	200.7
R-Squared	0.9977	Resid. Mean Square (MSE)		4.12857	
Adjusted R-Squared	0.9906	Standard Deviation		2.03189	

Source	DF	SS	MS	F	P
Regression	3	1753.07	584.357	141.54	0.0617
Residual	1	4.13	4.129		
Total	4	1757.20			

Cases Included 5 Missing Cases 0

Se tiene la ecuación de regresión, que no resulta significativa. Para descomponer la suma de cuadrados debida a cada componente, se continúa con la secuencia

Results > Stepwise AOV Table

y aparece

Stepwise Anlaysia of Variance of CALLO

Source	Individual SS	Cum DF	Cumulative SS	Cumulative MS	Adjusted R-Squared	Mallows' CP	P
Constant	38368.8						
ANILLADO	1188.10	1	1188.10	1188.10	0.5682	136.8	2
ANILLADO2	492.071	2	1680.17	840.086	0.9123	19.7	3
ANILLADO3	72.9000	3	1753.07	584.357	0.9906	4.0	4
Residual	4.12857	4	1757.20	439.300			

R-Squared 0.9977 Resid. Mean Square (MSE) 4.12857
 Adjusted R-Squared 0.9906 Standard Deviation 2.03189

donde la suma de cuadrados de regresión de la tabla anterior (1753,07) se ha descompuesto en la suma de cuadrados individual de cada componente, indicado en la columna **Individual SS**. Con estos datos completamos la tabla del análisis de la varianza conforme se indicó más arriba y calculamos de forma manual las F de los componentes aislados como se ha especificado. Estas operaciones no la realiza el programa, pero se reducen a algunas restas y divisiones.

CAPÍTULO 18

ANÁLISIS DE COVARIANZA

El análisis de covarianza es una técnica que combina aspectos del análisis de varianza y de la regresión. Intervienen, por consiguiente, dos o más variables relacionadas de algún modo entre sí.

Uno de los usos más importantes del análisis de covarianza es para aumentar la precisión en experimentos realizados con plantas perennes, aunque no es exclusivo de estas plantas. En plantas perennes, la variación entre individuos influye más en el error experimental que la variación posicional en el experimento. Supongamos, por ejemplo, que se estudia la influencia de los tratamientos en la producción de una especie arbórea. Es muy posible que entre los árboles en ensayo unos sean más productivos que otros, o que unos presenten un tamaño mayor que otros, lo que supondrá una mayor producción por árbol. Si algunos tratamientos se aplican por azar a los árboles más productivos o a los mayores, se obtendrá una sobreestimación del efecto de esos tratamientos. Aunque esto no fuera así, y los tratamientos se distribuyeran por igual entre los diferentes árboles, esas características productivas de los mismos contribuirán de forma notable a aumentar la variabilidad. Si con anterioridad a la realización de los tratamientos se conoce la producción media de cada árbol experimental o se toma una medida del tamaño del árbol, se puede corregir por covarianza la parte de la producción debida al árbol, con lo cual se compara únicamente la producción debida a los tratamientos.

Como se desprende del ejemplo anterior, en el análisis de covarianza se supone que la variable en estudio está afectada por otra variable distribuida en las parcelas elementales y que no está relacionada con los tratamientos. La variable en estudio es la que denominamos variable dependiente (Y) y la que influye en ella, y no depende de los tratamientos, la variable independiente (X). En el análisis de covarianza se asume que:

1. La variable independiente X es independiente de los tratamientos y debería medirse antes de aplicarlos.
2. La variable dependiente Y da el mismo error para cada tratamiento, o debe ser transformada para que sea así.
3. La relación entre X e Y debe ser de la forma $Y = a + bX$, es decir, deben guardar una relación lineal. Si no fuera así, la covarianza no debe aplicarse para conseguir los objetivos.
4. El valor de **b** debe ser independiente de los tratamientos.

El modelo matemático del análisis de covarianza en un experimento en bloques al azar es:

$$Y_{ij} = \mu + \alpha_i + \rho_j + \beta(X_{ij} - \bar{X}_{..}) + \varepsilon_{ij}$$

donde:

μ = media global

α_i = efecto de los tratamientos

ρ_j = efecto de los bloques

β = coeficiente de regresión de Y sobre X

ε_{ij} = residuales

Si la ecuación anterior se pone de la forma:

$$Y_{ij} - \beta(X_{ij} - \bar{X}_{..}) = \mu + \alpha_i + \rho_j + \varepsilon_{ij}$$

queda el modelo de un análisis de varianza en el que los valores de Y han sido ajustados por regresión sobre la variable independiente.

REALIZACIÓN DEL ANÁLISIS DE COVARIANZA

La Tabla 18.1 recoge los resultados obtenidos en un ensayo sobre época de aclareo manual de frutos en la nectarina 'Arnking'. Los tratamientos fueron los siguientes:

1. Testigo sin aclarar
2. Aclareo realizado 33 días después de plena floración (PF + 33)
3. Aclareo en PF + 42
4. Aclareo en PF + 49

Cada tratamiento se aplicó en una rama principal, de manera que un árbol con cuatro ramas principales constituía un bloque. Los datos recogen la producción en kg de cada rama experimental (Y); como las ramas no eran todas idénticas podría ocurrir que unas fueran más productivas que otras en función del número de ramos fructíferos que poseyeran, por lo que antes de comenzar el experimento se determinó el número de ramos fructíferos en cada rama experimental (X), que guarda una relación lineal con la producción (véase ejemplo de la Tabla 16.1).

Tabla 18.1

Bloques	1		2		3		4		5		Total		Medias	
Tratam.	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	\bar{X}	\bar{Y}
1	34	18,55	36	16,55	43	22,0	46	18,65	35	15,15	194	90,9	38,8	18,18
2	19	7,5	54	17,25	50	17,2	37	11,25	55	18,1	215	71,3	43	14,26
3	39	11,65	34	10,45	55	14,65	33	9,45	33	11,15	194	57,35	38,8	11,47
4	27	5,95	43	10,6	37	6,25	53	12,3	55	14,0	215	49,1	43	9,82
Total	119	43,65	167	54,85	185	60,1	169	51,65	178	58,4	818	268,65	40,9	13,43

Antes de comenzar el análisis es conveniente revisar que se cumplen todos los supuestos del análisis de varianza (normalidad, homogeneidad de las varianzas, independencia de medias y varianzas, y aditividad conforme se ha indicado en el capítulo 4) y los específicos del análisis de covarianza que, como se ha indicado más arriba, son que las X sean independientes de los tratamientos y que la relación entre Y y X sea lineal. El primer caso puede comprobarse realizando un análisis de varianza para las X y comprobando que no hay significación entre los tratamientos. En este caso $F = 0,28^{NS}$ para la variable número de ramos fructíferos, cumpliéndose esa premisa. Para comprobar que la relación entre Y y X es lineal se realiza un análisis de regresión; en el ejemplo la regresión lineal aparece significativa, luego puede procederse a la realización del análisis de covarianza.

El primer paso consiste en calcular las sumas de cuadrados, o de productos en este caso, igual que se procedía en el análisis de varianza.

Términos correctores:

$$C_x = \frac{X_{..}^2}{rt} = \frac{818^2}{5 * 4} = 33.456,2$$

$$C_y = \frac{Y_{..}^2}{rt} = \frac{268,65^2}{5 * 4} = 3.608,64$$

$$C_{xy} = \frac{X_{.Y}}{rt} = \frac{818 * 268,65}{5 * 4} = 10.987,78$$

Suma de cuadrados o de productos totales:

$$SC_x = \sum x_{ij}^2 - C_x = 34^2 + 19^2 + 39^2 + \dots + 33^2 + 55^2 - 33.456,2 = 2.017,8$$

$$SC_y = \sum y_{ij}^2 - C_y = 18,55^2 + 7,5^2 + \dots + 11,15^2 + 14,0^2 - 3.608,64 = 381,82$$

$$SC_{xy} = \sum x_{ij} y_{ij} - C_{xy} = (34 * 18,55) + (19 * 7,5) + \dots + (55 * 14,0) - 10.987,78 = 483,41$$

Suma de cuadrados o de productos de bloques:

$$SC_x = \frac{\sum x_{.j}^2}{t} - C_x = \frac{119^2 + 167^2 + 185^2 + 169^2 + 178^2}{4} - 33.456,2 = 673,8$$

$$SC_y = \frac{\sum y_{.j}^2}{t} - C_y = \frac{43,65^2 + 54,85^2 + 60,1^2 + 51,65^2 + 58,4^2}{4} - 3.608,64 = 42,39$$

$$SC_{xy} = \frac{\sum x_{.j} y_{.j}}{t} - C_{xy} = \frac{(119 * 43,65) + \dots + (178 * 58,4)}{4} - 10.987,78 = 161,43$$

Suma de cuadrados o de productos de tratamientos:

$$SC_x = \frac{\sum x_{.i}^2}{r} - C_x = \frac{194^2 + 215^2 + 194^2 + 215^2}{5} - 33.456,2 = 88,2$$

$$SC_y = \frac{\sum y_{.i}^2}{r} - C_y = \frac{90,9^2 + 71,3^2 + 57,35^2 + 49,1^2}{5} - 3.608,64 = 200,62$$

$$SC_{xy} = \frac{\sum x_{.i} y_{.i}}{r} - C_{xy} = \frac{(194 * 90,9) + (215 * 71,3) + \dots + (215 * 49,1)}{5} - 10.987,78 = -58,48$$

Suma de cuadrados o de productos del error:

$$SC_{\text{error}} = SC_{\text{totales}} - SC_{\text{bloques}} - SC_{\text{tratamientos}}$$

Con los datos así calculados se completa el cuadro del análisis de covarianza como sigue:

Fuente de variación	gl	SC _x	SC _{xy}	SC _y	Desviaciones de regresión			
					gl	SC	CM	F
Total	19	2.017,8	483,41	381,82				
Bloques	4	673,8	161,43	42,39				
Tratamientos	3	88,2	-58,48	200,62				
Error	12	1.255,8	380,46	138,80	11	23,53	2,14	
Tratamientos + error	15	1.344,0	321,98	339,42	14	262,28		
Tratamientos ajustados (diferencia)					3	238,75	79,58	37,18***

Como se ha indicado anteriormente, el análisis de covarianza es esencialmente un análisis de varianza de la cantidad Y-bX. Para realizar la prueba F de hipótesis nula de las medias ajustadas hay que determinar las desviaciones de regresión o, lo que es igual, el ajuste de Y por las X. Si las sumas de cuadrados o productos del error se designan por:

$$E_{xx} = 1.255,8 \qquad E_{xy} = 380,46 \qquad E_{yy} = 138,80$$

la suma de cuadrados del error en las columnas de desviaciones de regresión es:

$$SC_{error} = E_{yy} - \frac{E_{xy}^2}{E_{xx}} = 138,80 - \frac{380,46^2}{1.255,8} = 23,53$$

con 12-1 = 11 grados de libertad, debido al consumo de 1 g.l. por la regresión.

Análogamente, para la fila tratamientos + error:

$$SC_{tratamientos+error} = 339,42 - \frac{321,98^2}{1.344,0} = 262,28$$

con 15-1 = 14 gl.

La línea de tratamientos ajustados se calcula por diferencia entre las dos filas anteriores, obteniéndose el valor de F:

$$F = \frac{79,58}{2,14} = 37,18***$$

En el ejemplo se puede afirmar que existen diferencias significativas entre los tratamientos en cuanto a producción se refiere, una vez corregidas las producciones por el número de ramos fructíferos iniciales de cada rama experimental.

Si no se hubiera utilizado la covarianza, el cuadrado medio del error sería:

$$CM_{error} = \frac{E_{yy}}{gl} = \frac{138,80}{12} = 11,56$$

y al ajustar las producciones ha resultado de 2,14 lo que ha provocado una notable disminución en ese valor y, consiguientemente, un aumento de la F de 5,78 a 37,18.

El siguiente paso consiste en calcular **b** y las medias ajustadas. Tenemos que:

$$b = \frac{E_{xy}}{E_{xx}} = \frac{380,46}{1.255,8} = 0,3$$

y las medias ajustadas:

$$\bar{Y}_{aj1} = \bar{Y}_1 - b(\bar{X}_1 - \bar{X}_{..}) = 18,18 - 0,3(38,8 - 40,9) = 18,81$$

$$\bar{Y}_{aj2} = \bar{Y}_2 - b(\bar{X}_2 - \bar{X}_{..}) = 14,26 - 0,3(43 - 40,9) = 13,63$$

$$\bar{Y}_{aj3} = \bar{Y}_3 - b(\bar{X}_3 - \bar{X}_{..}) = 11,47 - 0,3(38,8 - 40,9) = 12,1$$

$$\bar{Y}_{aj4} = \bar{Y}_4 - b(\bar{X}_4 - \bar{X}_{..}) = 9,82 - 0,3(43 - 40,9) = 9,19$$

El error estándar de la diferencia entre dos medias ajustadas es:

$$s_{\bar{D}} = s' \sqrt{\frac{2}{n}}$$

siendo n el tamaño de la muestra por tratamiento, esto es, el número de bloques.

$$s'^2 = s_{y,x}^2 \left(1 + \frac{t_{xx}}{E_{xx}} \right)$$

es la varianza efectiva del error, $s^2_{y,x} = 2,14$ el cuadrado medio del error con 11 gl y t_{xx} es el cuadrado medio de tratamientos para X. Luego:

$$t_{xx} = \frac{88,2}{3} = 29,4$$

$$s'^2 = 2,14 \left(1 + \frac{29,4}{1.255,8} \right) = 2,19$$

$$s' = 1,48$$

$$s_{\bar{D}} = 1,48 \sqrt{\frac{2}{5}} = 0,94$$

La mínima diferencia significativa puede calcularse ahora de la forma:

$$MDS = t_{0,05} * s_{\bar{D}} = 2,201 * 0,94 = 2,07$$

El coeficiente de variación es:

$$CV = \frac{s'}{\bar{Y}_{..}} 100 = \frac{1,48}{13,43} 100 = 11,02$$

Comparando los tratamientos de aclareo con el testigo por cualquiera de los procedimientos vistos en capítulos anteriores, puede deducirse que el aclareo siempre reduce la producción. Realizando un contraste polinómico para las fechas de aclareo, se aprecia que la reducción es mayor conforme el aclareo es más tardío.

CÁLCULO CON EL PROGRAMA *Statistix* (SX)

En el programa SX introducir los datos experimentales como de costumbre denominando, por ejemplo, como BLK los bloques, TRA los tratamientos, PRO la producción (variable Y) y RAMOS la variable X. Una vez introducidos los datos experimentales, seguir la siguiente secuencia:

Statistics>Linear Models> Analysis of Variance>General AOV/AOCV

Al llegar a este punto indicar cuál es la variable dependiente (PRO), el modelo (*AOV Model Statement*), que en este caso está compuesto por BLK y TRA, la covariable (RAMOS) y luego pulsar OK. Aparece la siguiente tabla:

Analysis of Variance Table for PRO

Source	DF	SS	MS	F	P
BLK	4	5.481	1.370	0.64	0.6446
TRA	3	238.758	79.586	37.21	0.0000
RAMOS	1	115.273	115.273	53.89	0.0000
Error	11	23.530	2.139		
Total	19				

Note: SS are marginal (type III) sums of squares

Grand Mean 13.432 CV 10.89

Covariate Summary Table

Covariate	Coefficient	Std Error	T	P
RAMOS	0.30297	0.04127	7.34	0.0000

En la tabla del análisis de varianza se presentan los datos para los tratamientos ajustados y para error conforme fueron calculados anteriormente. Se incluyen los de la covariable RAMOS, con un grado de libertad, que se refiere a la regresión lineal con la PRODUCCIÓN, cuya F muestra significación a un alto nivel. Esto es crítico, pues como se ha indicado anteriormente, en el análisis de covarianza se asume que las variables están relacionadas de forma lineal, luego hay que comprobar que la relación lineal es significativa o, en caso contrario, no sirve de nada la covarianza con esa variable. La suma de cuadrados de RAMOS, 115,273 es la reducción de la suma de cuadrados debida a la regresión, esto es, al haber utilizado la covariable. De no haberlo hecho, la suma de cuadrados del error hubiera sido $115,273+23,53 = 138,803$ (E_{yy}) y la F debida a tratamientos hubiera disminuido.

En la tabla se recoge también la media general ($\bar{Y} = 13,432$), el coeficiente de variación (10,89) y una tabla adicional para la covariable, donde calcula el coeficiente de regresión, $b = 0,30297$ y su significación por la prueba t.

En el submenú que aparece junto a la tabla anterior, puede seguirse la siguiente secuencia para obtener las medias ajustadas y la separación de medias correspondiente:

Results> Multiple Comparisons>Comparisons with a Control

En el caso concreto de este ejemplo, se comparan frente a un control tres épocas de aclareo. Lo correcto sería comparar en primer lugar las fechas de aclareo frente al control, y después averiguar si hay diferencias entre las fechas de aclareo. Para el primer caso la secuencia a seguir es la indicada anteriormente. Al llegar a ese punto, indicar el término del modelo que se usará para la comparación (*Term Selected for Mean Comparisons*), en este caso los tratamientos (TRA) y el tratamiento que es el control, en este caso el tratamiento 1. Puede elegirse el nivel de significación y la hipótesis alternativa que interese, en este caso que sea diferente del control. Se pulsa OK y aparece:

Two-sided Dunnett's Multiple Comparisons with a Control of PRO

Control: TRA=1

Simultaneous 95% confidence intervals of treatment mean - control mean

TRA	Mean	Lower Bound	Difference	Upper Bound
1	18.816			
2	13.624	-7.750	-5.192*	-2.635
3	12.106	-9.223	-6.710*	-4.197
4	9.184	-12.190	-9.632*	-7.075

Alpha 0.05
 Critical D Value 2.717
 Error term used: BLK*TRA, 11 DF

Los resultados indican que todos los tratamientos difieren significativamente del control. En la segunda columna aparecen las medias ajustadas para cada tratamiento; estas sirven para la comparación de medias, pero al presentar los resultados habría que incluir las medias originales y la separación obtenida con las ajustadas.

La comparación de medias entre las épocas de aclareo habría que hacerlas por regresión, pues se trata de un factor cuantitativo. Para ello habría que realizar de nuevo el análisis de covarianza pero omitiendo los datos del tratamiento 1 que es el control. A continuación, en el submenú, seguir la secuencia:

Results > Polynomial Contrasts

e indicar que el término de contraste son los tratamientos (TRA). Pulsar OK y aparece:

Polynomial Contrasts of PRO by TRA

Degree = 1, Linear Trend

Contrast	-3.1396	SS (Contrast)	49.284
Scheffe's F	27.67	P (Scheffe's F)	0.0005
T-Statistic	-7.44	P (T-Statistic)	0.0001
SE (Contrast)	0.4220		

Degree = 2, Quadratic Trend

Contrast	-0.5048	SS (Contrast)	1.2741
Scheffe's F	0.67	P (Scheffe's F)	0.5401
T-Statistic	-1.16	P (T-Statistic)	0.2838
SE (Contrast)	0.4349		

Error term used: BLK*TRA, 7 DF

La tabla muestra el grado 1 y 2 del polinómico, esto es, la tendencia lineal o cuadrática. Si nos fijamos en la prueba t se observa que los valores de la significación (valor de P) recogidos en la cuarta columna muestran significación para el caso de la tendencia lineal, lo que permite concluir que respecto a la época de aclareo, conforme más tardía es más, se reduce la producción, siguiendo una tendencia lineal.

Para más información acerca de la separación de medias en el análisis de varianza, véase el capítulo 5.

COVARIANZA MÚLTIPLE

Cuando hay más de una variable independiente los cálculos son más complicados al sumar más términos en el modelo, pero el concepto y el procedimiento no cambian. El modelo matemático para un experimento en bloques al azar con dos variables independientes es:

$$Y_{ij} = \mu + \alpha_i + \rho_j + \beta_1 (X_{1ij} - \overline{X_{1..}}) + \beta_2 (X_{2ij} - \overline{X_{2..}}) + \varepsilon_{ij}$$

La Tabla 18.2 muestra los datos del ejemplo anterior, pero se ha añadido una nueva variable medida con anterioridad, el vigor de la rama experimental expresado por su perímetro en la base de inserción, medido en cm. Esta variable es una medida del tamaño de la rama experimental, por lo que cabe suponer que al aumentar el tamaño de la rama aumente la producción de forma independiente de los tratamientos. También cabe suponer que aumente el número de ramos fructíferos, pero tras la poda de fructificación esa relación puede alterarse. En la Tabla 18.2, Y es el peso de la rama en kg, X₁ es el número de ramos fructíferos por rama y X₂ el perímetro de la rama en la base de inserción, en cm. Más información acerca de las relaciones entre estas tres variables puede encontrarse en el capítulo 16.

Tabla 18.2

Tratamiento	Variable	Bloque					Total	Medias
		1	2	3	4	5		
1	X ₁	34	36	43	46	35	194	38,8
	X ₂	26	28	25	30	25	134	26,8
	Y	18,55	16,55	22,0	18,65	15,15	90,9	18,18
2	X ₁	19	54	50	37	55	215	43
	X ₂	28	25	28	24	30	135	27
	Y	7,5	17,25	17,2	11,25	18,1	71,3	14,26

Trata- miento	Variable	Bloque					Total	Medias
		1	2	3	4	5		
3	X ₁	39	34	55	33	33	194	38,8
	X ₂	24	23	26	23	24	120	24
	Y	11,65	10,45	14,65	9,45	11,15	57,35	11,47
4	X ₁	27	43	37	53	55	215	43
	X ₂	24	26	20	27	31	128	25,6
	Y	5,95	10,6	6,25	12,3	14,0	49,1	9,82
Total	X ₁	119	167	185	169	178	818	40,9
	X ₂	102	102	99	104	110	517	25,85
	Y	43,65	54,85	60,1	51,65	58,4	268,65	13,43

El primer paso, una vez comprobados los supuestos del análisis de covarianza, consiste en calcular las sumas de cuadrados y de productos de la forma ya conocida y que da como resultado la tabla siguiente:

Fuente de variación	gl	SC _{x₁}	SC _{x₁x₂}	SC _{x₂}	SC _{x₁y}	SC _{x₂y}	SC _y
Total	19	2.017,8	237,7	142,55	483,41	120,5	381,82
Bloques	4	673,8	15,45	16,8	161,43	3,52	42,39
Tratamientos	3	88,2	18,9	28,55	-58,48	49,98	200,62
Error	12	1.255,8	203,35	97,2	380,46	67,0	138,80
Tratamientos + error	15	1.344,0	222,25	125,75	321,98	116,98	339,42

Se calculan a continuación los coeficientes de regresión, las reducciones de las sumas de cuadrados y las sumas de cuadrados de las desviaciones para las filas del error y de tratamientos + error. Las fórmulas son:

Cálculo de b_1 y b_2 :

$$(SC_{x_1})b_1 + (SC_{x_1x_2})b_2 = SC_{x_1y}$$

$$(SC_{x_1x_2})b_1 + (SC_{x_2})b_2 = SC_{x_2y}$$

Reducción en la suma de cuadrados:

$$\sum \hat{y}^2 = b_1(SC_{x_1,y}) + b_2(SC_{x_2,y})$$

Suma de cuadrados de las desviaciones:

$$\sum d^2 = \sum y^2 - \sum \hat{y}^2$$

Aplicando las fórmulas tenemos:

Para el error

$$1.255,8 b_1 + 203,35 b_2 = 380,46$$

$$203,35 b_1 + 97,2 b_2 = 67,0$$

luego $b_1 = 0,29$ y $b_2 = 0,081$

$$\sum \hat{y}^2 = 0,29 * 380,46 + 0,081 * 67,0 = 115,76$$

$$\sum d^2 = 138,8 - 115,76 = 23,04$$

Para tratamientos + error

$$1.344 b_1 + 222,25 b_2 = 321,98$$

$$222,25 b_1 + 125,75 b_2 = 116,98$$

luego $b_1 = 0,12$ y $b_2 = 0,72$

$$\sum \hat{y}^2 = 0,12 * 321,98 + 0,72 * 116,98 = 122,86$$

$$\sum d^2 = 339,42 - 122,86 = 216,56$$

Con estos datos se realiza la prueba F como se indica a continuación.

Fuente de Variación	gl	SC	CM	F
Tratamiento + error	13	216,56		
Error	10	23,04	2,304	
Diferencia	3	193,52	64,5	27,99***

Las medias ajustadas se calculan de la siguiente forma:

$$\bar{Y}_{aj i} = \bar{Y}_i - b_1(\bar{X}_{1i} - \bar{X}_1) - b_2(\bar{X}_{2i} - \bar{X}_2)$$

siendo b_1 y b_2 los coeficientes de regresión calculados para el error. Las medias ajustadas para los tratamientos resultan:

$$\bar{Y}_{aj1} = 18,18 - 0,29(38,8-40,9) - 0,081(26,8-25,85) = 18,71$$

$$\bar{Y}_{aj2} = 14,26 - 0,29(43-40,9) - 0,081(27-25,85) = 13,56$$

$$\bar{Y}_{aj3} = 11,47 - 0,29(38,8-40,9) - 0,081(24-25,85) = 12,23$$

$$\bar{Y}_{aj4} = 9,82 - 0,29(43-40,9) - 0,081(25,6-25,85) = 9,23$$

La varianza efectiva del error es:

$$s^2 = s^2 (1 + c_{11} t_{11} + 2 c_{12} t_{12} + c_{22} t_{22})$$

siendo los c_{ij} los multiplicadores inversos y t_{ij} los cuadrados medios de los tratamientos. Tenemos, pues, que:

$$t_{11} = \frac{88,2}{3} = 29,4$$

$$t_{12} = \frac{18,9}{3} = 6,3$$

$$t_{22} = \frac{28,55}{3} = 9,52$$

Si E_{ij} son las sumas de cuadrados del error, tenemos que:

$$c_{11} = \frac{E_{22}}{E_{11}E_{22} - E_{12}^2} = \frac{97,2}{(1.255,8)(97,2) - 203,35^2} = 0,0012$$

$$c_{12} = \frac{E_{12}}{E_{11}E_{22} - E_{12}^2} = \frac{203,35}{(1.255,8)(97,2) - 203,35^2} = 0,0025$$

$$c_{22} = \frac{E_{11}}{E_{11}E_{22} - E_{12}^2} = \frac{1.255,8}{(1.255,8)(97,2) - 203,35^2} = 0,015$$

con lo cual:

$$s^2 = 2,304 [1 + (0,0012*29,4) + (0,0025*6,3) + (0,015*9,52)] = 2,79$$

$$s' = 1,67$$

El error estándar de la diferencia entre dos medias ajustadas es:

$$s_{\bar{D}} = s' \sqrt{\frac{2}{n}} = 1,67 \sqrt{\frac{2}{5}} = 1,056$$

La mínima diferencia significativa al nivel del 5% resulta:

$$MDS = 2,228 * 1,056 = 2,35$$

y el coeficiente de variación:

$$CV = \frac{s'}{Y} 100 = \frac{1,67}{13,43} 100 = 12,43$$

Los resultados no varían de los obtenidos con la covarianza simple, lo que indica que la introducción de la variable X_2 (vigor) no ha mejorado la precisión. Como ya se conocía, ambas variables están correlacionadas. Pero además podría haber falta de significación en la regresión adicional con estos datos. Si se determina la significación de b_1 y b_2 obtenemos:

$$s_{b_1} = \sqrt{s^2 c_{11}} = \sqrt{(2,304)(0,0012)} = 0,0526$$

$$s_{b_2} = \sqrt{s^2 c_{22}} = \sqrt{(2,304)(0,015)} = 0,1859$$

Luego:

$$t_1 = \frac{b_1}{s_{b_1}} = \frac{0,29}{0,0526} = 5,51^{***}$$

$$t_2 = \frac{b_2}{s_{b_2}} = \frac{0,081}{0,1859} = 0,44^{NS}$$

La regresión adicional con X_2 no ha resultado significativa, lo que indica que la covarianza simple hubiera bastado para el análisis de estos datos. Véase el ejemplo del capítulo 16 para repasar la relación entre estas variables.

CÁLCULO CON EL PROGRAMA *Statistix* (SX)

En el programa SX el proceso es el mismo seguido para el análisis de covarianza simple, pero indicando ambas covariables RAMOS y VIGOR. La tabla del análisis es la siguiente:

Analysis of Variance Table for PRO					
Source	DF	SS	MS	F	P
BLK	4	5.010	1.2525	0.54	0.7084
TRA	3	193.574	64.5247	27.96	0.0000
RAMOS	1	69.545	69.5446	30.13	0.0003
VIGOR	1	0.452	0.4517	0.20	0.6676
Error	10	23.078	2.3078		
Total	19				

Note: SS are marginal (type III) sums of squares

Grand Mean 13.432 CV 11.31

Covariate Summary Table				
Covariate	Coefficient	Std Error	T	P
RAMOS	0.28940	0.05272	5.49	0.0003
VIGOR	0.08383	0.18949	0.44	0.6676

Nótese la falta de significación de la regresión para VIGOR, con $F = 0,20^{NS}$ y $t = 0,44^{NS}$ y, en consecuencia, la poca reducción adicional en la suma de cuadrados.

De haber resultado significativa la covarianza doble, se procedería a partir de ahora a obtener las medias ajustadas y la separación de medias conforme se ha visto para la covarianza simple.

CONSIDERACIONES SOBRE EL ANÁLISIS DE COVARIANZA

Como se ha destacado en el ejemplo anterior, las medias ajustadas estiman mejor el efecto de los tratamientos al haber eliminado una de las fuentes de error experimental, lo que ha provocado una notable reducción en la suma de cuadrados del error. La eficacia del análisis de covarianza depende, pues, de una elección acertada de la variable independiente. En este aspecto, recuérdese que una de las premisas del análisis de covarianza era que las X fueran independientes de los tratamientos, que se ha comprobado calculando la F para las X que ha resultado $F = 0,28^{NS}$, cumpliéndose esa premisa. La comprobación de relación lineal entre las variables dependiente e independiente, otra de las premisas del análisis, ha sido realizada con anterioridad; si la relación no fuera lineal habría que realizar transformaciones en las variables hasta conseguirlo, siempre que estén, lógicamente, correlacionadas.

En ocasiones, las X pueden mostrar diferencias entre tratamientos que han sido impuestas por la toma de datos. Esto debe evitarse, pero si no es posible los cálculos permanecerían sin cambios, pero al determinar la diferencia entre dos medias la s_D^2 se calcularía como sigue:

$$s_D^2 = s_{y.x}^2 \left(\frac{2}{n} + \frac{(\bar{X}_i - \bar{X}_j)^2}{E_{xx}} \right)$$

La s_D^2 , por consiguiente, no es constante y varía para cada dos medias. En otras palabras, para cada par de medias que se comparen hay que estimar s_D^2 . Si las X varían mucho entre tratamientos, s_D^2 aumenta porque el término $(\bar{X}_i - \bar{X}_j)^2$ es grande, lo que puede hacer que no aparezca significación donde realmente la hay.

Por otra parte, si las X están sujetas a errores sustanciales en su medición, el ajuste solo eliminaría parte de las diferencias de Y debida a las X, mientras que otra parte permanecería en las medias ajustadas. Es necesario insistir de nuevo que una buena elección de las X es crítica para que el análisis de covarianza sea eficiente.

CAPÍTULO 19

MÉTODOS NO PARAMÉTRICOS

En los capítulos anteriores se ha considerado que los datos se distribuyen de acuerdo con distribuciones conocidas, y por ello se han estimado parámetros y se han probado hipótesis relativas a ellos. Esta es una situación muy común, donde se manejan datos que, por lo general, se ajustan a una distribución normal o cercana a la normalidad, pues los métodos estadísticos descritos también resultan efectivos para poblaciones moderadamente anormales.

Sin embargo, en muchos datos que aporta la investigación no es fácil especificar la distribución original o ésta se aleja mucho de la normalidad. En estos casos se requieren métodos que no dependan de una distribución específica, esto es, que sean independientes de las distribuciones de la población y de los parámetros asociados. Se conocen como métodos no paramétricos y se aplican cuando no se conoce la naturaleza de la distribución en la que se fundamentan los datos; cuando se utilizan categorías, bien porque falta una escala de medida adecuada o porque se desean tomar datos con rapidez; y cuando se trata de datos no numéricos. Estos métodos suelen ser muy sencillos y fáciles de entender, y pueden aplicarse también en condiciones estándar, pero si la población es normal o cercana a la normalidad, o los datos pueden transformarse para que lo sean, los métodos no paramétricos dan menos información. Por lo general, la eficiencia de estos métodos es alta para muestras pequeñas, digamos que para $n \leq 10$, y disminuye conforme aumenta n .

PRUEBA DE LOS SIGNOS

Esta es una prueba no paramétrica para comparar **muestras pareadas**. De acuerdo con el método, las diferencias entre los miembros de un par se reemplazan por signos (+ ó -) ignorándose la magnitud de la diferencia. La prueba de los signos se utiliza para comparar dos muestras cuando las medias son categorías, aunque también se utiliza con datos continuos o discretos, bien porque se quiere evitar la suposición de normalidad o porque es un sustituto rápido de la prueba t, aunque es menos sensible. Cuando la variable es continua o discreta, la hipótesis nula es que la variable tiene la misma distribución bajo los dos tratamientos. Hasta para 20 pares de datos, la Tabla A5 muestra el número más pequeño de signos iguales que se requieren para una significación del 1% ó 5%.

Como ejemplo utilizaremos el mismo que se utilizó para comparar muestras pareadas, cuyos datos se recogen en la Tabla 3.1. La secuencia de signos para la diferencia entre X_1 y X_2 , asignando signo + cuando el valor de $X_1 > X_2$ y signo - cuando el valor de $X_1 < X_2$, es:

+ + + + + - - + - -

Para 10 pares de valores habría que obtener, según la Tabla A5, un número de 1 o menos de un signo y 9 o más del otro para alcanzar la significación al 5%. Como tenemos 4 nega-

tivos y 6 positivos, se concluye que no se rechaza la hipótesis nula, esto es, no hay diferencias significativas entre las dos muestras. Al mismo resultado se llegó utilizando la t de Student.

Cálculo con el programa *Statistic (SX)*

La secuencia a seguir es la siguiente:

Statistics>One, Two, Multi-Sample Tests>Sign Test

Indicar las variables a comparar, pinchar en OK y aparece:

Sign Test for X1 - X2

Number of Negative Differences	4
Number of Positive Differences	6
Number of Zero Differences (ignored)	0
Probability of a result as or more extreme than observed (one-tailed p-value)	0.3770

A value is counted as a zero if its absolute value is less than 0.00001

Cases Included 10 Missing Cases 0

El programa muestra el número de diferencias negativas y el de positivas y el valor de $P = 0,377$, que no permite rechazar la hipótesis nula.

PRUEBA DE RANGOS (CATEGORIAS) CON SIGNOS DE WILCOXON

Esta prueba, igual que la anterior, es una sustitución de la prueba t de Student para **muestras pareadas**, pero aporta mayor precisión que la prueba de los signos. Para la realización de la prueba se ordenan en orden ascendente los valores absolutos de las diferencias entre cada par, sin considerar el signo, y se asigna a la diferencia más pequeña la categoría (rango) 1, a la siguiente la 2, y así sucesivamente hasta llegar a la última, pero cada una con el signo de la diferencia.

Siguiendo con el ejemplo de las muestras pareadas de la Tabla 3.1, tenemos la siguiente ordenación y asignación de categorías:

Diferencias ordenadas:	-4,9	5,2	-5,9	-7,8	8,7	10,2	-13,2	14,7	17,5	21,3
Categoría con el signo:	-1	2	-3	-4	5	6	-7	8	9	10

Si dos o más diferencias son iguales, se asigna a cada una la categoría media de la que les correspondiera. Por ejemplo, si la segunda y tercera diferencia coincidieran en el valor (por ejemplo, la segunda sería 5,2 y la tercera -5,2), se asignaría a la segunda la categoría 2,5 y a la tercera -2,5, y la siguiente sería -4.

Una vez establecidas las categorías con el signo, se calcula la suma de categorías + (T_+) y la de - (T_-). En el ejemplo:

$$T_+ = 40 \quad \text{y} \quad T_- = -15$$

Obsérvese que $T_+ + T_- = \frac{n(n+1)}{2}$. Se elige el menor valor entre las dos y se le designa T .

En el ejemplo:

$$T = T = 15$$

Este valor, sin tener en cuenta el signo, se contrasta en la Tabla A6. Para 10 pares de muestras se requiere una suma de categorías ≤ 8 al nivel del 5%. Como $T = 15 > 8$, no permite rechazar la hipótesis nula, esto es, no existen diferencias significativas entre las dos muestras.

Para valores de n que sobrepasan los de la Tabla, se calcula el valor

$$z = \frac{|\mu - T| - 1/2}{\sigma}$$

donde:

$|\mu - T|$ es el valor absoluto de la diferencia

-1/2 es la corrección por continuidad

$$\mu = \frac{n(n+1)}{4}$$

T es el valor más pequeño de T_+ y T_-

$$\sigma = \sqrt{\frac{(2n+1)\mu}{6}}$$

Como es normal, $z > 1,96$ significa rechazar la hipótesis nula al 5%.

Cálculo con el programa *Statistic (SX)*

La secuencia a seguir es la siguiente:

Statistics>One, Two, Multi-Sample Tests>Wilcoxon Signed Rank Test

Indicar las variables a comparar, pinchar en OK y aparece:

Wilcoxon Signed Rank Test for X1 - X2	
Sum of Negative Ranks	-15.000
Sum of Positive Ranks	40.000
Exact probability of a result as or more extreme than the observed ranks (one-tailed p-value)	0.1162
Normal Approximation with Continuity Correction	1.223
Two-tailed P-value for Normal Approximation	0.2213
Total number of values that were tied	0
Number of zero differences dropped	0
Max. diff. allowed between ties	0.00001
Cases Included 10	Missing Cases 0

En la salida se encuentran los valores de T_+ y de T_- y la significación (el valor de P). Obsérvese que también incluye el valor de P (en tablas de dos colas) en el caso de realizar la comparación de muestras como en una distribución normal. El valor de P en una cola multiplicado por dos da aproximadamente el de dos colas, como sucede en el ejemplo. Quiere decir que realizando la comparación de muestras por los métodos paramétricos y por los no paramétricos, los resultados obtenidos son similares.

PRUEBA DE LA SUMA DE RANGOS (CATEGORÍAS) DE WILCOXON

Esta prueba también se conoce como la PRUEBA U DE MANN-WHITNEY que la ampliaron para muestras de tamaño distinto. Se trata de una prueba no paramétrica para comparar **muestras independientes**. Para la realización de la prueba se ordenan en orden ascendente los valores de ambas muestras juntos, pero marcando los de cada una para distinguirlos, y se asignan categorías a esos valores en la forma realizada en la prueba anterior.

Como ejemplo utilizaremos el mismo que se utilizó para comparar dos muestras independientes de igual tamaño, cuyos datos aparecen en la Tabla 3.2. Los valores ordenados, la categoría asignada y la muestra a la que pertenecen, se exponen a continuación:

Valores ordenados	Categoría	Muestra
8,31	1	X ₁
8,39	2	X ₁
10,28	3	X ₁
11,23	4	X ₁
13,13	5	X ₁
14,55	6	X ₁
14,98	7	X ₁
15,32	8	X ₁
16,65	9	X ₁
17,31	10	X ₁
23,38	11	X ₁
28,42	12	X ₂
28,63	13	X ₁
30,72	14	X ₂
33,51	15	X ₂
33,83	16	X ₁
34,99	17	X ₁
40,31	18	X ₂
44,80	19	X ₁
48,38	20	X ₂
51,13	21	X ₂
51,78	22	X ₂
51,97	23	X ₂
52,03	24	X ₂
53,03	25	X ₂
57,72	26	X ₂
58,06	27	X ₂
58,28	28	X ₂
62,15	29	X ₂
69,58	30	X ₂

Se obtiene a continuación la suma de categorías de la muestra que proporciona la suma más pequeña, en el ejemplo la muestra X₁:

$$T = 1+2+3+4+5+6+7+8+9+10+11+13+16+17+19 = 131$$

En la Tabla A7, para $n_1 = n_2 = 15$, $T_{0,05} = 185$ y $T_{0,01} = 171$. Como

$$T = 131 < T_{0,01} = 171$$

se rechaza la hipótesis nula al nivel del 1%, esto es, las diferencias entre las dos muestras son significativas.

Cálculo con el programa *Statistic (SX)*

La secuencia a seguir es la siguiente:

Statistics>One, Two, Multi-Sample Tests>Wilcoxon Rank Sum Test

Especificar la forma en la que se almacenaron las variables en la base de datos; si ambas entraron como dos variables, seleccionar *Table*; si entraron como una sola variable pero con una segunda que separa ambos grupos, seleccionar *Categorical*. Indicar las variables a comparar, pinchar en OK y aparece:

Wilcoxon Rank Sum Test for X1 VS X2				
Variable	Rank Sum	N	U Stat	Mean Rank
X1	131.00	15	11.000	8.7
X2	334.00	15	214.00	22.3
Total	465.00	30		
Normal Approximation with Corrections for Continuity and Ties				4.189
Two-tailed P-value for Normal Approximation				0.0000
Total number of values that were tied				0
Maximum difference allowed between ties				0.00001
Cases Included 30		Missing Cases 2		

Encontramos la suma de categorías para ambas variables y para el total, aunque la comparación se hace exclusivamente con la suma más pequeña. El valor exacto de P solo lo da para muestras pequeñas, y en este caso da el valor de P en caso de realizarse una comparación estándar, que puede utilizarse en este ejemplo para determinar la significación.

Cuando las **muestras** son **de tamaño distinto**, se estima la T para la muestra de menor tamaño, que se designa por n_1 , y se calcula

$$T' = n_1 (n_1 + n_2 + 1) - T$$

Tomando el ejemplo de muestras de distinto tamaño de la Tabla 3.3, y ordenando los datos en categorías, tenemos:

Valores ordenados	Categoría	Muestra
8,31	1	X ₁
8,39	2	X ₁
10,28	3	X ₁
11,23	4	X ₁
13,13	5	X ₁
14,55	6	X ₁
14,98	7	X ₁
15,32	8	X ₁
16,65	9	X ₁
17,31	10	X ₁
23,38	11	X ₁
28,42	12	X ₂
28,63	13	X ₁
30,72	14	X ₂
33,83	15	X ₁
34,99	16	X ₁
44,80	17	X ₁
51,13	18	X ₂
51,78	19	X ₂
51,97	20	X ₂
52,03	21	X ₂
53,03	22	X ₂
62,15	23	X ₂

$$T = 12+14+18+19+20+21+22+23 = 149$$

$$T' = 8(8+15+1)-149 = 43$$

En la Tabla A7, $T_{0,01} = 56$ para $n_1 = 8$ y $n_2 = 15$, valor superior que 43, por lo que se rechaza la hipótesis nula al nivel del 1%.

Cálculo con el programa *Statistic (SX)*

Sigue exactamente la misma secuencia que en el caso anterior. En el ejemplo, aparece la salida:

```
Wilcoxon Rank Sum Test for X1 VS X2

Variable      Rank Sum      N      U Stat      Mean Rank
X1            127.00       15     7.0000       8.5
X2            149.00        8     113.00      18.6
Total         276.00       23

Exact Permutation Test Two-tailed P-value      0.0326

Normal Approximation with Corrections for Continuity and Ties      3.389
Two-tailed P-value for Normal Approximation      0.0007

Total number of values that were tied      0
Maximum difference allowed between ties 0.00001

Cases Included 23      Missing Cases 9
```

En este caso incluye el valor de P, así como la aproximación a la normal.

Para valores de n_1 y n_2 fuera de la Tabla, se calcula

$$z = \frac{|\mu - T| - 1/2}{\sigma}$$

siendo:

$$\mu = \frac{n_1(n_1 + n_2 + 1)}{2} \quad \text{y} \quad \sigma = \sqrt{\frac{n_2 \mu}{6}}$$

PRUEBA DE KRUSKAL-WALLIS

Esta prueba es una generalización de la prueba de Wilcoxon-Mann-Whitney para k muestras, de modo que para $k = 2$ ambas pruebas son equivalentes, por lo que es preferible utilizar la de Wilcoxon-Mann-Whitney. En la práctica, la prueba de Kruskal-Wallis es un método no paramétrico para el **análisis de varianza de una vía**, que resulta de utilidad cuando los datos no reúnen los requisitos para dicho análisis, únicamente que las muestras sean independientes. La prueba se basa en la asignación de categorías (rangos) a los datos en la forma especificada en pruebas anteriores.

El procedimiento es similar al anterior, de forma que se asignan categorías a todas las observaciones juntas ordenadas en forma ascendente, se suman las categorías para cada muestra o tratamiento y se calcula el valor:

$$H = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(n+1)$$

siendo

n_i el número de observaciones de la i -ésima muestra ($i = 1 \dots k$)

$$n = \sum n_i$$

R_i la suma de categorías de la i -ésima muestra.

El valor H se distribuye como una χ^2 con $k-1$ grados de libertad, en el supuesto de que $n_1, n_2 \dots n_k$ sean al menos 5 todas ellas. Como en casos anteriores, si dos valores coinciden se les asigna una categoría intermedia.

En la Tabla 19.1 se recogen los datos de un experimento en el que plantas jóvenes de olivo cultivadas en macetas de 2 L de capacidad en condiciones de invernadero, se sometieron durante un período de tiempo a la aplicación de 0, 50 ó 100 ppm de nitrógeno en el agua de riego, junto a una solución con los demás elementos nutritivos. Transcurrido un tiempo el contenido en nitrógeno de las plantas varió según el tratamiento recibido y, cuando esto sucedió, todas las plantas recibieron una aplicación adicional de nitrógeno vía foliar al objeto de estudiar si la eficiencia del uso del nitrógeno depende o no del estado nutritivo de las plantas. Se utilizaron 6 plantas por tratamiento.

Tabla 19.1.- Contenido de nitrógeno en plantas jóvenes de olivo después de la aplicación adicional de nitrógeno vía foliar.

Concentración de N en el agua de riego	Contenido de N en las plantas (mg/planta)					
0 ppm	139,52	136,58	103,54	94,28	68,40	148,47
50 ppm	173,68	125,58	106,90	111,26	118,25	124,57
100 ppm	161,59	121,14	81,56	156,25	119,74	132,73

Siguiendo con el procedimiento descrito, se ordenan todos los datos de menor a mayor y se les asigna la categoría comenzando con el valor más bajo:

Valores ordenados	Categoría	Tratamiento
68,40	1	0 ppm
81,56	2	100 ppm
94,28	3	0 ppm
103,54	4	0 ppm
106,90	5	50 ppm
111,26	6	50 ppm
118,25	7	50 ppm
119,74	8	100 ppm
121,14	9	100 ppm
124,57	10	50 ppm
125,58	11	50 ppm
132,73	12	100 ppm
136,58	13	0 ppm
139,52	14	0 ppm
148,47	15	0 ppm
156,25	16	100 ppm
161,59	17	100 ppm
173,68	18	50 ppm

Los valores de R_i son:

$$R_0 = 50$$

$$R_{50} = 57$$

$$R_{100} = 64$$

con los que se calcula H :

$$H = \frac{12}{18(19)} \frac{50^2 + 57^2 + 64^2}{6} - 3(18 + 1) = 0,57 \quad \text{con } 3-1 = 2 \text{ g.l.}$$

El valor en la Tabla A3 para 2 g.l. es $\chi^2_{0,05} = 5,99 > H = 0,57$, luego no se rechaza la hipótesis nula, concluyendo que no existen diferencias significativas entre los tratamientos, esto es, las plantas deficientes en N hacen un uso más eficiente del mismo tras una aplicación foliar que las plantas bien nutridas en este elemento. A la misma conclusión se llega realizando un análisis de varianza estándar, el más correcto en este caso pues los datos cumplen los requisitos para utilizar un método paramétrico.

Cálculo con el programa *Statistic (SX)*

La secuencia a seguir es la siguiente:

Statistics>One, Two, Multi-Sample Tests>Kruskal-Wallis One-Way AOV

Especificar la forma en la que se almacenaron las variables en la base de datos; si entraron como tres variables, seleccionar *Table*; si entraron como una sola variable pero con una segunda que las separa, seleccionar *Categorical*. Indicar las variables a comparar, pinchar en OK y aparece:

```

Kruskal-Wallis One-Way Nonparametric AOV

Variable      Mean Rank   Sample Size
Cero          8.3         6
Cien          10.7        6
Cincuenta    9.5         6
Total        9.5         18

Kruskal-Wallis Statistic          0.5731
P-Value, Using Chi-Squared Approximation  0.7508

Parametric AOV Applied to Ranks
Source      DF      SS      MS      F      P
Between     2      16.333  8.1667  0.26  0.7732
Within     15     468.167 31.2111
Total      17     484.500

Total number of values that were tied      0
Max. diff. allowed between ties          0.00001

Cases Included 18      Missing Cases 3

```

El programa da la categoría media para cada tratamiento, en lugar de la suma total, el tamaño de la muestra o número de repeticiones, el valor de H y la significación (el valor de P). A continuación muestra el análisis de varianza paramétrico, que aporta los mismos resultados como ya se ha comentado.

En el caso de existir demasiadas coincidencias en los datos, esto es, demasiados datos con el mismo valor, particularmente cuando son de tratamientos diferentes, puede hacerse una corrección de H dividiendo su valor por:

$$1 - \frac{\sum (T^3 - T)}{n^3 - n}$$

siendo T el número de coincidencias correspondientes a cada observación, y el sumatorio corresponde a la suma de todas las observaciones. Si no hay coincidencias, $T = 0$, el divisor es 1 y no es necesaria la corrección. En la práctica, la corrección no cambia sustancialmente el valor de H. El programa SX calcula el valor de H corregido siempre que detecte una coincidencia en los datos.

PRUEBA DE FRIEDMAN

Se trata de una prueba no paramétrica para la realización del **análisis de varianza de dos vías**, concretamente para el más común de todos, el correspondiente a un diseño en bloques al azar con más de dos tratamientos. Como en el caso anterior, este método es aplicable particularmente cuando la asunción de normalidad es sospechosa.

El procedimiento consiste en asignar categorías a los datos correspondientes a los tratamientos dentro de cada bloque en orden ascendente, obtener la suma de categorías para cada tratamiento, y probar la hipótesis nula de que las poblaciones de un bloque son idénticas, con la hipótesis alternativa de que al menos un tratamiento no lo sea. Para ello se calcula:

$$\chi_R^2 = \frac{12}{bt(t+1)} \sum R_i^2 - 3b(t+1), \text{ con } t-1 \text{ grados de libertad.}$$

donde:

t es el número de tratamientos

b es el número de bloques

R_i es la suma de las categorías para el i-ésimo tratamiento.

χ_R^2 mide la homogeneidad de las t sumas y se distribuye aproximadamente como una χ^2 . La aproximación es diferente para valores pequeños de t y b, para los que Friedman elaboró tablas.

En la Tabla 19.2 se recogen datos correspondientes a un diseño en bloques al azar con tres tratamientos y seis bloques. Lo que significan los datos es irrelevante en este momento.

Tabla 19.2.- Datos correspondientes a un diseño en bloques al azar.

Bloque	Tratamiento		
	1	2	3
1	14,9	7,2	18,6
2	13,3	9,5	22,2
3	4,6	6,8	10,0
4	14,7	9,7	17,6
5	12,8	9,2	14,4
6	7,5	7,5	7,9

Asignando categorías a los tratamientos dentro de cada bloque, tenemos:

Bloque	Tratamiento		
	1	2	3
1	2	1	3
2	2	1	3
3	1	2	3
4	2	1	3
5	2	1	3
6	1,5	1,5	3
Ri:	10,5	7,5	18

En el bloque 6 había una coincidencia en los tratamientos 1 y 2, por lo que se ha asignado a ambos una categoría intermedia a la que les correspondería. Una vez obtenidas las Ri, se calcula

$$\chi_R^2 = \frac{12}{6(3)4} (10,5^2 + 7,5^2 + 18^2) - 3(6)4 = 9,75 \text{ con } 3-1 = 2 \text{ g.l.}$$

Como existe una coincidencia en los datos, se realiza una corrección por las mismas dividiendo el valor de χ_R^2 por:

$$1 - \frac{\sum_{i=1}^b T_i}{bt(t^2 - 1)}$$

donde

$$T_i = \sum_h t_{ih}^3 - \sum_h t_{ih}$$

t_{ih} es el número de observaciones coincidentes para una categoría dada en el i-ésimo bloque

h es el índice de la sumatoria para los conjuntos de coincidencias en el bloque.

En el ejemplo solo hay dos observaciones coincidentes, luego:

$$t_i = 2;$$

$$h = 1;$$

$$T_i = 2^3 - 2 = 6$$

luego el divisor resulta:

$$1 - \frac{6}{6(3)(3^2 - 1)} = 0,9583$$

y, en consecuencia,

$$\chi_R^2 = \frac{9,75}{0,9583} = 10,174$$

Este valor es mayor que 9,21 que es el valor en la Tabla A3 para la $\chi^2_{0,01}$, lo que permite rechazar la hipótesis nula y concluir que existen diferencias significativas entre los tratamientos al nivel del 1%.

Cálculo con el programa *Statistic (SX)*

La secuencia a seguir es la siguiente:

Statistics>One, Two, Multi-Sample Tests>Friedman Two-Way AOV

Especificar la forma en la que se almacenaron las variables en la base de datos. En este caso interesa introducir los datos de manera que pueda elegirse la opción *Categorical* para poder indicar cual es la variable dependiente (la variable en estudio). Para ello basta introducir tres variables: Tratamientos, Bloques y la Variable dependiente. Indicar el lugar correspondiente a estas variables, pinchar en OK y aparece:

Friedman Two-Way Nonparametric AOV for X = TRA BLK

TRA	Mean Rank	Sample Size
1	1.75	6
2	1.25	6
3	3.00	6

Friedman Statistic, Corrected for Ties 10.174
P-value, Chi-Squared Approximation 0.0062
Degrees of Freedom 2

BLK	Mean Rank	Sample Size
1	4.33	3
2	5.00	3
3	1.33	3
4	5.00	3
5	3.33	3
6	2.00	3

Friedman Statistic 10.429
P-value, Chi-Squared Approximation 0.0640
Degrees of Freedom 5

Max. diff. allowed between ties 0.00001

Cases Included 18 Missing Cases 0

El programa muestra el valor medio de cada categoría para cada tratamiento, el tamaño de la muestra (número de bloques), el valor de χ^2_R , la significación (el valor de P) y los grados de libertad. Nótese que los cálculos los ha realizado teniendo en cuenta la corrección por coincidencias. También aporta estos datos para los bloques, algo que no resulta de interés pues interesa la comparación de los tratamientos. En cualquier caso, los datos muestran que no existen diferencias significativas entre bloques.

CORRELACIÓN DE RANGOS DE SPEARMAN

La determinación del coeficiente de correlación r normalmente requiere que la población siga una distribución normal bivalente, algo que no siempre sucede. La correlación de rangos de Spearman es un método no paramétrico que trata de salvar esa situación y se aplica a datos en forma de categorías (rangos).

El procedimiento de cálculo comienza por asignar categorías a las observaciones de cada variable, teniendo en cuenta que si hay coincidencias en los valores de algunos datos se les asigna a cada uno la categoría media, como se ha hecho en situaciones anteriores. A continuación, se determina la diferencia de categorías en cada observación pareada y se calcula el coeficiente de correlación de rangos de la forma:

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

donde d es la diferencia entre las categorías asignadas a ambas variables, y n es el número de pares de datos. Igual que ocurre con r , los valores de r_s varían entre -1 y +1.

En la Tabla 19.3 se recogen los valores medios asignados por dos observadores distintos en la evaluación visual de síntomas de clorosis férrica que mostraban en un experimento árboles de olivo. Estos síntomas se caracterizan por la aparición de hojas de color amarillento, de manera que se utilizó una escala visual que variaba entre el 0 (color verde de las hojas, sin síntomas de clorosis) al 3 (clorosis muy pronunciada). En la tercera columna se incluye la cantidad de clorofila extraída de una muestra de hojas de los mismos árboles evaluados. Interesa, pues, conocer si los datos de los dos observadores están relacionados, esto es, si dan lugar a los mismos resultados, y también si existe relación entre la medida visual y la extracción de clorofila en las hojas, un método preciso pero que exige un trabajo extra en el laboratorio.

Tabla 19.3. Índice visual de clorosis asignado por dos observadores distintos y contenido en clorofila de hojas de olivo tomadas de los mismos árboles observados.

Muestra	Índice visual ¹ (Observador 1)	Índice visual ¹ (Observador 2)	Clorofila en hojas (mg/g)
1	2,5	2,0	0,60
2	0,1	0,5	1,25
3	1,4	1,0	1,04
4	1,7	1,2	0,66
5	0,2	0,7	1,32
6	1,9	2,0	0,62
7	1,2	1,2	1,06

¹Basado en una escala de 0 = no clorosis a 3 = clorosis muy pronunciada.

En el ejemplo se combinan datos de variables cualitativas (índice visual) y variables continuas (contenido en clorofila). El empleo de los métodos no paramétricos parece indicado. Veamos en primer lugar si existe relación entre los datos de observadores distintos. Para ello, se asignan categorías dentro de cada observador en la forma acostumbrada:

Muestra	Índice visual (Observador 1)	Índice visual (Observador 2)	Diferencia d	d ²
1	7	6,5	0,5	0,25
2	1	1	0	0
3	4	3	1	1
4	5	4,5	0,5	0,25
5	2	2	0	0
6	6	6,5	-0,5	0,25
7	3	4,5	-1,5	2,25
			∑d = 0	∑d ² = 4

El valor de r_s resulta:

$$r_s = 1 - \frac{6(4)}{7(49 - 1)} = 0,928^{**}$$

La significación se contrasta en la Tabla A8, cuyo valor para 7 pares de muestras resulta $0,893 < 0,928$ resultando una correlación significativa al 1%, esto es, ambos observadores han proporcionado los mismos resultados en su evaluación.

Del mismo modo se procedería para conocer si hay relación entre la medida visual y el contenido en clorofila. Comparemos, pues, los resultados del observador 1 y los del análisis de las muestras de hojas. La asignación de categorías queda como sigue:

Muestra	Índice visual (Observador 1)	Clorofila en hojas (mg/g)	Diferencia d	d ²
1	7	1	6	36
2	1	6	-5	25
3	4	4	0	0
4	5	3	2	4
5	2	7	-5	25
6	6	2	4	16
7	3	5	-2	4
			$\Sigma d = 0$	$\Sigma d^2 = 110$

Luego:

$$r_s = 1 - \frac{6(110)}{7(49 - 1)} = -0,964^{**}$$

Ha resultado una correlación negativa y altamente significativa. El resultado es lógico pues a mayor contenido en clorofila en las hojas mayor color verde muestran, esto es, menor grado de clorosis, de ahí la correlación negativa.

Cálculo con el programa *Statistic (SX)*

La secuencia a seguir es la siguiente:

Statistics>Association Test>Spearman Rank Correlations

Indicar las variables a correlacionar (pueden incluirse las tres del ejemplo), pedir que calcule el valor de P pinchando en *Compute P-Values*, y OK con lo que aparece:

Spearman Rank Correlations, Corrected for Ties

	Obs1	Obs2
Obs2	0.9274	
P-Value	0.0064	
Clorofila	-0.9643	-0.8911
	0.0064	0.0190
Maximum Difference Allowed Between Ties	0.00001	
Cases Included	7	Missing Cases 1

Obsérvese que indica que las categorías se han corregido por coincidencias (*Corrected for Ties*). El programa aporta los coeficientes r_s para las tres variables y la significación para cada uno (el valor de P).

CAPÍTULO 20

TOMA DE DATOS Y MEDIDAS EN CAMPO

Cuando se realiza un experimento cualquiera se pretende observar la respuesta de la planta a los tratamientos que han sido introducidos en el mismo. En muchas ocasiones la respuesta es tan clara, que basta una observación de las plantas para ver el efecto de los tratamientos. En un experimento realizado para determinar la influencia de la tiourea en la salida del reposo de yemas de melocotonero, se recogían periódicamente ramos fructíferos de los árboles experimentales durante los meses de invierno y se introducían en una cámara de crecimiento a 20 °C. En los primeros muestreos las yemas de los ramos procedentes de árboles tratados con tiourea brotaban a los siete días de ser introducidos en las cámaras, mientras que los ramos sin tratar permanecían intactos. Resultaba obvio que la aplicación de tiourea provocaba la salida del reposo de yemas de melocotonero cuando aún no se había acumulado frío suficiente para la brotación natural, y bastaba con observar los ramos al cabo de una semana para cerciorarse de ese efecto.

En otros casos, por el contrario, las diferencias entre tratamientos solo se detectan cuando la aplicación de una prueba de significación muestra que hay efectos diferenciales. En cualquiera de los casos, no obstante, es necesario tomar medidas de respuesta de las plantas. Como indica Pearce (1976), las plantas no se miden principalmente para proveer datos para los análisis estadísticos, sino para dejar constancia y para describir a otros lo que el investigador observa. Por ello, la relación que debe existir entre el investigador y las plantas experimentales no se sustituye por ningún sistema de medida, por preciso que pueda ser.

La toma de datos en experimentos de campo requiere mucho cuidado y precaución para evitar errores sustanciales que puedan falsear los resultados o, al menos, aumentar la variabilidad. Las medidas pueden realizarse en campo o en laboratorio y en ambos casos existen ventajas o dificultades. Si las medidas se realizan en campo se minimizan los errores que pueden cometerse por la falta de identificación de un tratamiento o de una parcela experimental, pero el trabajo suele ser más dificultoso y tedioso, agravarse con condiciones meteorológicas adversas, aumentar los errores por las prisas, particularmente si el experimento está lejos del centro de trabajo, o por las dificultades. No obstante, algunos datos han de tomarse necesariamente en campo, como la toma de un índice visual de cierta sintomatología o el tamaño del tronco de un árbol, pero siempre que se pueda suele ser preferible realizar las medidas en laboratorio. En este caso han de extremarse las precauciones para evitar un error fatal que confunda o dificulte la identificación de una muestra. Para ello hay que etiquetar todo lo que salga del campo experimental y vaya a medirse en el laboratorio, por ejemplo ramos, frutos, hojas, etc., indicando con claridad el tratamiento, la parcela, el bloque o la repetición que corresponda a la muestra y sin que haya posibilidad de error en la interpretación de la etiqueta. Hay veces que se invalida un experimento completo por no ser rigurosos en la identificación o confiar en anotaciones simples e imprecisas.

Es recomendable, en experimentos agronómicos, fotografiar todo aquello que no se mida o, aún midiéndose, resulte gráfico, como cualquier sintomatología, crecimiento diferencial de las plantas, etc. En algunos casos se utiliza la fotografía para medir el tamaño o las dimensiones de los árboles posteriormente en laboratorio, siempre que se tomen a la misma distancia, lo que no representa ninguna dificultad en una plantación regular.

Con independencia de las características de las medidas a tomar, éstas pueden clasificarse en tres tipos (Pearce, 1976):

1. Medidas *completas*, cuando se mide todo lo producido en un experimento (se cuentan todos los frutos, se miden todos los brotes, etc.), o basadas en un *muestreo*.
2. Medidas basadas en *mediciones, estimaciones* o *categorías*. En el primer caso es necesaria la lectura en un instrumento, como una balanza, un colorímetro, etc. En las estimaciones se toman valores numéricos, normalmente a ojo, sin que sean medidos, por ejemplo 90% de flores abiertas, 30% de hojas cloróticas, etc. Es preferible definir los límites de cada clase más que indicar su valor medio, por ejemplo entre 20-30% más que 25%. En las categorías, los individuos se colocan en grupos ordenados pero con los límites definidos arbitrariamente; por ejemplo, la incidencia de enfermedades puede clasificarse en ligera, moderada o fuerte.
3. Medidas *directas* o *indirectas*. En el primer caso se mide la variable que realmente se desea, mientras en el segundo se mide una variable, por lo general más fácil de medir, que esté relacionada con la que se desea medir. Por ejemplo, el peso de un árbol solo puede medirse antes de ser plantado, en el trasplante o al arrancarse, pero puede obtenerse una idea precisa del tamaño del árbol midiendo el perímetro del tronco o el volumen de copa sin necesidad de proceder a su arranque.

Las mejores medidas son, lógicamente, las completas y directas pero esto no siempre es posible. En cualquier caso hay que decidir qué datos se van a tomar durante la realización del experimento, qué medidas son necesarias y cómo pueden obtenerse. Merece la pena dedicar cierto tiempo a ello, pues en ocasiones se toman tantos datos que, desbordados por el trabajo, no se toman correctamente o no llegan a elaborarse ni, consiguientemente, a utilizarse, pero ha generado un gasto y un esfuerzo innecesario. En otros casos, por el contrario, se pretende simplificar tanto la toma de datos que, una vez finalizado el experimento, no se encuentran medidas que apoyen las hipótesis.

MÉTODOS DE MUESTREO

El objetivo que se pretende con el muestreo es conseguir un grado de precisión deseado con los mínimos recursos. Los métodos de muestreo son muy variados, pero en experimentación en campo hay dos que son de interés: el muestreo aleatorio simple y el muestreo sistemático. En cualquiera de ellos el mayor peligro es que aumente el error de muestreo, en particular porque el material seleccionado de cada tratamiento no sea representativo de la población.

Muestreo aleatorio simple

En este método los miembros de la muestra son tomados independientemente con igual probabilidad. El método resulta satisfactorio cuando la población no es muy variable y, en particular, cuando se estiman proporciones comprendidas entre el 20% y el 80%.

En todo experimento hay un error inherente al material en estudio, los árboles por ejemplo, (σ_i) que se tendría aún midiéndolos todos, y un error técnico de muestreo (σ_t) que, en general, aumenta al aumentar σ_i . Estos errores pueden estimarse para determinar la bondad del método de muestreo. Pearce (1976) sugiere para ello tomar, en un ensayo en curso, dos muestras de cada parcela elemental, sumar ambos valores, realizar un análisis de varianza con las sumas y el cuadrado medio del error será una estimación de:

$$CM_{error} = 2\sigma_t^2 + 4\sigma_i^2$$

A continuación, la suma de los cuadrados de las diferencias dividida por el número de parcelas elementales (p) es una estimación de:

$$\frac{\sum (X_{1ij} - X_{2ij})^2}{p} = 2\sigma_t^2$$

Con estas dos expresiones se pueden estimar ambos errores, pero hay que recordar que son estimaciones y, por consiguiente, sujetas a error, de manera que podría ocurrir el absurdo de que σ_t^2 resultase negativo. Esto indicaría que σ_t ha resultado muy superior a σ_i y que el método de muestreo no ha sido satisfactorio.

Lo ideal es que $\sigma_t < \frac{1}{3}\sigma_i$, es decir, que σ_t fuera tan pequeño que estuviera incluido en σ_i .

Con esos valores puede estimarse el efecto del muestreo. Así, si n es el número de repeticiones necesarias para mostrar diferencias en el caso de realizar una medida completa, con

el muestreo se necesitarían $n \left(1 + \frac{\sigma_t^2}{\sigma_i^2} \right)$

El problema real es determinar el tamaño de la muestra, ni tan pequeña que no resulte con la precisión deseada, ni tan grande que multiplique innecesariamente el trabajo. El primer paso consiste en decidir qué error es admisible en la estimación; por ejemplo, si se trata de tomar una muestra de frutos para obtener el calibre, hay que decidir cuantos mm del diámetro del fruto se pueden tolerar como error. Esto solo lo puede determinar la experiencia previa, pero la decisión ha de tomarse con cuidado pues el tamaño de la muestra puede variar ampliamente. El siguiente paso consiste en expresar el error admisible en términos de límites de confianza. Si L es el error admisible en la media de la muestra y se acepta un riesgo del 5% de que el error exceda a L , puede demostrarse que el tamaño de la muestra exigido es:

$$n = \frac{4\sigma^2}{L^2}$$

de manera que para conocer n se necesita una estimación de la varianza poblacional σ^2 . Esto no siempre es posible, por lo que muchos investigadores prefieren tomar una muestra amplia, por ejemplo los 2/3, que les da garantías de precisión y ahorran un tercio de trabajo.

Si se dispone de resultados de muestreos anteriores puede estimarse σ^2 . Por ejemplo, de una muestra de 150 frutos tomada de cuatro árboles de melocotonero se obtuvo una varianza $s^2 = 15,37$ para el calibre del fruto, expresado en mm de diámetro. Si en un ensayo posterior en las mismas condiciones se desea conocer el tamaño de la muestra de frutos para determinar el calibre en la misma variedad de melocotonero, con un error admisible de $L = 1$ mm de diámetro y con un límite de confianza del 5%, es decir, que en el 95% de los casos el error de la media no supere 1 mm de diámetro, tenemos que el tamaño de la muestra debe ser, al menos, de:

$$n = \frac{4\sigma^2}{L^2} = \frac{4(15,37)}{1^2} = 61,48 \approx 62 \text{ frutos}$$

Si no se conoce s^2 , el valor de σ^2 puede estimarse conociendo los valores extremos de la población, esto es, la amplitud. Si h es la diferencia entre el valor más alto y el más bajo, σ puede tomar valores de $0,21h$ para una distribución sesgada, $0,24h$ para una distribución simétrica ó $0,29h$ para una distribución uniforme.

Si la cantidad a estimar es una proporción binomial, el tamaño de la muestra es:

$$n = \frac{4pq}{L^2}$$

donde p , q y L pueden expresarse en proporciones siempre que todos se expresen en las mismas unidades.

Para aplicar la fórmula es necesaria una estimación de p . Por ejemplo, en un plantel de un vivero que contenía 700 plántones se estimó que alrededor del 10% estaban enfermos. Si se desea conocer el tamaño de la muestra necesaria para determinar el porcentaje de plántones enfermos, con un error admisible del 4%, tendríamos:

$$n = \frac{4pq}{L^2} = \frac{4(10)(90)}{4^2} = 225 \text{ plántones}$$

En las fórmulas anteriores no se ha tenido en cuenta la corrección de la población finita, ya que la población que se muestrea es de tamaño finito N . En general son válidas cuando n es menor del 10% de N , pero si el valor estimado es superior, puede obtenerse un n' que considera la corrección de la población finita.

$$n' = \frac{n}{1 + f}$$

siendo $f = n / N$ la fracción de muestreo. En el ejemplo anterior:

$$n' = \frac{225}{1 + 225/700} = 170,45 \approx 171 \text{ plantones}$$

Muestreo sistemático

Este tipo de muestreo es muy popular en experimentos de campo y resulta muchas veces más preciso que el anterior. Para entender la diferencia entre ambos, supongamos que se trata de coger cuatro frutos por caja de tres cajas de igual tamaño procedentes de la recolección. En el muestreo aleatorio simple deben seleccionarse en primer lugar las cajas al azar y elegir de entre todas tres y, posteriormente, los frutos dentro de cada caja también al azar. En el muestreo sistemático las cajas se eligen a intervalos prefijados, sorteando tan solo la primera de ellas. Así, por ejemplo, si tenemos un total de 15 cajas, se toma la 3ª (elegida al azar), la 8ª y la 13ª en el orden de recolección, esto es, se toma una de cada cinco. En este muestreo es más fácil conseguir la muestra y además se distribuye mejor sobre toda la población. A veces es el único método practicable. Si se pretende, por ejemplo, tomar una muestra de frutos en un árbol en un momento determinado de su desarrollo, en el muestreo sistemático se elegiría un ramo fructífero cada cuatro, diez, etc. según el volumen del árbol y se tomaría el fruto central de cada uno (o el que se considere representativo). En el muestreo aleatorio habría que estimar el número total de frutos del árbol y asignarle un número a cada uno para proceder al sorteo, algo que sería impracticable. En el muestreo sistemático, además, la muestra estaría distribuida por todo el árbol.

El mayor inconveniente del método es que el error técnico de muestreo es más difícil de calcular. Un método consiste en medir una muestra de cada parcela elemental (X_p) y luego medirla entera (X_c), con lo que se puede calcular el error de muestreo:

$$\sigma_t^2 = \frac{\sum (X_{c_{ij}} - X_{p_{ij}})^2}{p}$$

siendo p el número de parcelas elementales medidas. Realizando un análisis de varianza con las medidas completas, se estima:

$$\sigma_t^2 = CM_{error}$$

Aunque poco ortodoxo, se puede obtener una idea del tamaño de la muestra realizando los cálculos como si se tratase de un muestreo aleatorio simple y posteriormente revisar los valores de n .

Otro inconveniente del método es el riesgo de sesgo. Si la población contiene una variación periódica y el intervalo entre unidades coincide con ella, la muestra puede estar muy sesgada. Para estimar este riesgo puede hacerse lo de antes, medir una muestra en una parcela, medir luego toda la parcela completa y apreciar si hay o no sesgo. Pero a veces la medida

completa de una parcela no es fácil en la práctica. No obstante, hay que considerar que un sesgo constante no afectaría a la diferencia entre tratamientos.

ESTIMACIONES Y CATEGORÍAS

Este tipo de medidas son unas herramientas útiles y sencillas para la valoración de muchos resultados de un experimento, pero siempre que se utilicen con sentido crítico. En el mejor de los casos es un medio sencillo de mostrar diferencias entre tratamientos, y en el peor de ellos grandes diferencias pueden quedar enmascaradas. Todo depende, en parte, de la persona que utilice estas herramientas. Si el observador pone tal entusiasmo en la respuesta que espera de un determinado tratamiento que le altera la objetividad, el procedimiento puede ser desastroso.

Un punto que requiere especial atención es la elección de la escala, que debe permitir mostrar diferencias y a la vez ser de fácil comprensión por el observador, de forma que se minimice el riesgo de errores por confusión de la misma. Se han utilizado multitud de escalas que varían desde 3 a 15 grados. Pearce (1976) recomienda para estimaciones no pasar de diez grados y para categorías una escala de no más de cinco grados, para evitar el riesgo de confusión al aumentar el número y el de imprecisión al disminuirlo. Sobre esta base se han elaborado varias escalas; algunas de muy amplio uso en categorización se describen a continuación.

Escalas 1-5

En realidad son escalas de tres grados que utilizan otros dos para casos de emergencia o dudosos. Tienen la ventaja de la simplicidad y fácil categorización. La primera de ellas consiste en:

- 1.- Malo
- 3.- Medio
- 5.- Bueno

Los grados 2 y 4 se reservan para casos intermedios.

En una segunda escala, los valores serían:

- 2.- Individuos por debajo de la media
- 3.- Individuos que muestren un valor medio
- 4.- Individuos por encima de la media

Los valores 1 y 5 se reservarían para los individuos malísimos o los realmente excelentes, respectivamente.

Escala 0-3

Es una escala de cuatro grados también muy simple, que puede utilizarse con éxito para comparaciones con un testigo. Las categorías son, para el caso de observación de síntomas:

- 0.- Sin síntomas (o testigo, en su caso).
- 1.- Pocos síntomas.
- 2.- Moderados síntomas.
- 3.- Muchos síntomas.

Escala 1-9

En esta escala las categorías se establecen del siguiente modo:

- 1.- Extremadamente bajo (o muy malo)
- 3.- Bajo (o malo)
- 5.- Medio
- 7.- Alto (o bueno)
- 9.- Extremadamente alto (o excelente)

Obsérvese que esta escala es, en realidad, de cinco grados pero reserva los valores pares 2, 4 y 6 para los casos de duda en la asignación de un valor a un individuo, y deben asignarse únicamente en casos muy particulares. Muy utilizada en multitud de trabajos de observación.

Escala 0-5

Esta escala se utiliza con frecuencia para la evaluación de síntomas provocados por diversos agentes. Las categorías son las siguientes:

- 0.- Individuo sano, sin síntomas visibles.
- 1.- Individuo con algún síntoma asociado a la adversidad.
- 2.- Individuo con pocos a moderados síntomas.
- 3.- Individuo con muchos síntomas, indicando un estado avanzado de afectación.
- 4.- Individuo con síntomas graves.
- 5.- Individuo muerto.

Esta escala consta de 6 grados, pero los valores 0 y 5 no crean confusión quedando, en la práctica, cuatro grados para discriminar. Lo importante en esta escala es definir con precisión el valor 3, y asignar los demás por comparación con los síntomas mostrados por un individuo con ese valor.

El valor 0 en una escala ha sido cuestionado por algunos por su indefinición, ya que cero indica ausencia de algo; pero en estudios sobre incidencia de adversidades, por ejemplo, el valor 0 asignado a un individuo sano que no presenta ningún tipo de síntoma asociado a la adversidad, es algo preciso por lo que se incluye como valor categórico en algunas escalas.

En la interpretación de los datos hay que considerar que si los resultados son consistentes, pequeñas diferencias entre medias podrían indicar significación estadística. Pero para que realmente se puedan asignar diferencias significativas entre tratamientos, la diferencia entre sus medias debería ser, al menos, igual a la división más pequeña que pueda discernir un experto en la toma de datos. Por consiguiente, atribuir importancia a una diferencia inferior a 0,5 puntos probablemente no está justificada en la mayoría de los casos (Lipton, 1992). En cualquier caso, la atribución de diferencias entre tratamientos quedará, en último lugar, a juicio del investigador.

El mayor riesgo en la utilización de categorías para la toma de datos de observación es el sesgo de la persona que los toma. En ocasiones, una persona sin experiencia en el tema toma los datos más correctamente que otra especializada, si ésta incluye en la valoración de las observaciones apreciaciones personales u observaciones indirectas. No obstante es preferible una persona experimentada, siempre que actúe con sentido crítico y objetividad. A ello ayuda conocer con claridad qué se quiere medir y distinguir el objeto de la observación de otras características similares, y comprender la escala o elegir aquella que mejor se adapte al problema y a la comprensión del observador.

Para evitar el sesgo de cada observador, es conveniente también que si son varias las personas que deben tomar los datos, cada una los tome de un bloque completo, no de un tratamiento. De esta manera, el sesgo introducido es el mismo en cada tratamiento y la variabilidad entre observadores se añade a la varianza entre bloques, pero no entre tratamientos. Si el experimento no está constituido por bloques, por ejemplo, si se trata de determinar la época de floración en una colección de variedades frutales, los datos completos debe tomarlos una sola persona o, si lo hacen más, deben trabajar independientemente y luego realizar una media de sus resultados.

MEDIDAS INDIRECTAS

Este tipo de medidas consiste en medir una cantidad x por su sencillez, en lugar de y , la cantidad buscada, siempre que ambas estén altamente correlacionadas. Esto es algo factible, pero a la vez peligroso pues se puede correr el riesgo de introducir sesgos y distorsiones no deseados, ya que los tratamientos pueden alterar las relaciones existentes entre esas variables. Por ejemplo, un árbol que crece más suele producir más, luego una medida del crecimiento vegetativo podría utilizarse indirectamente como una medida de fructificación o de producción. Si se realiza un ensayo sobre intensidad de poda esa relación puede fallar, pues los árboles sometidos a una poda severa tendrán un crecimiento vegetativo mayor pero también una producción menor.

Si se plantea tomar este tipo de medidas, es necesario comprobar que ambas variables están altamente correlacionadas, lo que generalmente se hace, pero además que la naturaleza de la relación se mantiene en todos los tratamientos. Puede que las variables estén bien relacionadas en cada uno de los tratamientos, pero la naturaleza de esa relación cambie entre ellos. En definitiva, las rectas de regresión entre ambas variables deben ser idénticas en todos los tratamientos. Como esto no siempre es fácil de comprobar, es necesaria una investigación cuidadosa antes de plantear este tipo de medidas.

CAPÍTULO 21

MEDIDAS EN PLANTAS LEÑOSAS

EXPERIMENTACIÓN CON PLANTAS LEÑOSAS

Las plantas leñosas poseen unas características que las diferencian sustancialmente de las plantas herbáceas, hasta el punto que afectan a los métodos experimentales aplicables a ellas. Entre esas características destacan las siguientes:

- a) *Perennidad*. Las plantas leñosas son, por lo general, más longevas que las plantas herbáceas y los experimentos suelen realizarse con las mismas plantas en la misma localidad durante varios años. Esta continuidad con el mismo material vegetal hace que no puedan aplicarse los métodos estadísticos que combinan el análisis de experimentos realizados en varios años, como ocurre con plantas anuales. En éstas, los experimentos se inician y terminan en la misma campaña y si se requiere mayor información, se establece un nuevo experimento en la campaña siguiente, que incluye un nuevo orden aleatorio de los tratamientos y el empleo, lógicamente, de individuos distintos. Para estos casos se han desarrollado los experimentos en serie. Por otra parte, como la planta perenne permanece en el mismo lugar durante varios años, las observaciones que se hagan en un año pueden estar correlacionadas con las que se realicen en otro.

Para evitar ese problema, se han propuesto dos soluciones. Una de ellas es considerar el experimento que dura varios años como un split-plot, donde los tratamientos constituyen la parcela principal y los años las subparcelas. El principal problema asociado con esta propuesta es que persiste la falta de aleatoriedad en la variable años; en otras palabras, no se pueden sortear los años. Asimismo, persiste la correlación dentro de la parcela elemental entre las observaciones realizadas en varios años. La segunda propuesta consiste en ajustar una curva que describa la relación entre los años y la medida de interés en cada tratamiento, algo que no está sometido a error. En general, es preferible no buscar dificultades y si se trata de estudiar la producción o el crecimiento, por ejemplo, suele ser suficiente sumar los datos y obtener la producción total acumulada durante el período de estudio o el crecimiento total.

Otro problema asociado con la perennidad es la posibilidad de que pueda perderse alguna planta durante el experimento, lo que alteraría el análisis de los datos. Este hecho ha de tenerse en cuenta y deben tomarse precauciones para evitar o minimizar el efecto de las pérdidas. En experimentos a largo plazo es aconsejable, por consiguiente, aumentar el número de individuos en la parcela elemental, de manera que si ocurre la pérdida de uno, aún queden suficientes individuos que permitan obtener los datos de cada parcela.

- b) *Tamaño*. Las plantas leñosas suelen alcanzar mayor tamaño que las herbáceas, lo que hace que tengan mayor interés como individuos. Mientras que la producción de un árbol puede ser medida y considerada en los métodos estadísticos, la de una planta anual se

obtiene sobre un conjunto de varias plantas, de manera que si las hay grandes y pequeñas, los tamaños se compensan al final dejando los factores ambientales, como el suelo, los únicos para ser considerados. La concentración sobre el individuo en las plantas leñosas implica una fuente de variación, de manera que la variabilidad existente entre individuos contribuye más al error experimental que la posición que ocupe éste en el experimento (Pearce, 1976). Por ello, el uso de diseños experimentales complejos para controlar la variación posicional no tienen sentido en experimentos con plantas leñosas.

Selección y calibración de plantas leñosas

La consideración del individuo al experimentar con plantas leñosas hace que se ponga especial cuidado en la selección de las plantas experimentales. Como la variación entre individuos es fundamentalmente de carácter genético, el empleo de material clonal, siempre que sea posible, es preferible a otro material para controlar esa variabilidad. Si no existe esta posibilidad, el material elegido debe ser lo más homogéneo posible y proceder de la misma fuente.

Muchos experimentos se hacen con plantas jóvenes cultivadas en macetas. En estos casos es muy conveniente estandarizar el tamaño de las plantas antes de aplicar los tratamientos, salvo que en el experimento la variable sea la variedad. Para ello, hay que seleccionar previamente más plantas de las necesarias para la realización del experimento, dejarlas crecer y aclimatarse en el lugar donde se vaya a realizar el mismo y, antes de la aplicación de los tratamientos, elegir aquellas que sean lo más homogéneas posible. Grandes diferencias en el tamaño inicial de estas plantas aumentan de tal manera la variabilidad que pueden resultar sin significación diferencias entre tratamientos que realmente difieren. De ser posible, las plantas deberían proceder del mismo material parental.

El control de la variabilidad individual se realiza también mediante la calibración. Esto implica medidas de una o más variables en cada planta antes de iniciar el experimento, de forma que puedan usarse posteriormente como covariables en un análisis de covarianza. Este es uno de los métodos más eficaces para el control de la variabilidad entre individuos en plantas leñosas. Medidas clásicas empleadas para la calibración de los árboles han sido la producción media de años anteriores, el diámetro del tronco de los árboles, el volumen de copa o el número de órganos fructíferos por unidad experimental. En el caso de plantas jóvenes cultivadas en maceta, el tamaño inicial de cada planta, medido bien por el peso, la altura o el crecimiento total, puede utilizarse como covariable para un análisis de covarianza.

La parcela elemental

Debido al tamaño y demás características de las plantas leñosas, la parcela elemental en experimentos con estas plantas está compuesta por menos individuos que en experimentos con plantas herbáceas. El número diferirá en función del tipo de experimento, de las condiciones locales, de las posibilidades, etc., pero normalmente implicarán entre uno y 15 ó 20 árboles en el mejor de los casos. Aunque no hay regla general, a modo orientativo se puede indicar

que en experimentos en bloques al azar compuestos de entre cuatro y seis bloques, con material vegetal homogéneo, la parcela elemental suele estar compuesta de uno a cuatro árboles; si el experimento se plantea a largo plazo, es conveniente al menos cuatro árboles para asumir posibles pérdidas durante el transcurso del experimento. En ensayos comparativos de variedades serían necesarios, al menos, cinco árboles por parcela elemental.

En otros tipos de experimentos, por el contrario, la parcela elemental está compuesta de una rama principal o un sector del árbol, siendo éste el que constituye un bloque. En estos casos hay que tener la seguridad de que el efecto del tratamiento dado en una rama no se trasloca a otras partes del árbol y, por consiguiente, a otra parcela elemental. Ensayos sobre aclareo manual de frutos, algunos tipos de poda, etc., pueden realizarse de esta manera con un ahorro sustancial de árboles experimentales. En cualquier caso hay que tener en mente que, salvado el mínimo de parcela elemental, es preferible aumentar el número de bloques que el de árboles dentro de cada parcela elemental para ganar precisión en el experimento. No obstante, un incremento exagerado del número de bloques solo redundaría en un mayor esfuerzo; un número razonable está comprendido, en la mayoría de los casos, entre cuatro y ocho.

MEDIDAS DEL CRECIMIENTO VEGETATIVO

El crecimiento se define como el conjunto de procesos que provocan un aumento irreversible en el tamaño de los organismos, acompañado normalmente por un aumento en peso (SECH, 1999). En las plantas leñosas se puede diferenciar un crecimiento primario derivado del meristemo apical, que da lugar a la formación de nuevos brotes y a la extensión de los ya formados, y un crecimiento secundario derivado de la actividad de los meristemos laterales, que provoca el crecimiento en grosor con independencia del crecimiento en longitud. Ambos tipos de crecimiento pueden medirse para observar la respuesta de la planta a los tratamientos.

Medidas del crecimiento nuevo

La medida completa del crecimiento en longitud de todos los ramos al final de período de crecimiento es una medida laboriosa pero no impracticable, al menos en plantas de pequeño tamaño. Todo consiste en seguir un orden para evitar equivocarse: medir cada rama por separado y en cada una comenzar desde su inserción hasta el ápice, marcando la parte ya medida para cerciorarse de que no se hace dos veces.

En plantas de cierto tamaño la medida completa del crecimiento nuevo es, en realidad, impracticable, por lo que hay que realizar un muestreo para tener una medida de ese crecimiento. En este caso se puede medir una rama principal del árbol o bien un número de ramos distribuidos alrededor del mismo, eligiendo aquellos que sean típicos de la especie, que guarden una inclinación similar con la vertical y evitando ramos anormales bien por su excesivo vigor, como los chupones, o escaso vigor por encontrarse sombreados, demasiado inclinados, etc. Aunque esta medida no es una indicación del crecimiento total anual, sí lo es del

crecimiento medio del brote. Hay que mencionar que el crecimiento del brote es muy sensible a cualquier cambio o estrés, por lo que es una medida de gran utilidad para comparar la respuesta de la planta al efecto de distintos tratamientos. Una medida indirecta del crecimiento total lo puede dar el peso de la leña de poda.

Medidas del tamaño del árbol

Probablemente la mejor medida del tamaño de un árbol sea el peso, pero esto solo puede hacerse en la plantación, en el trasplante o al arrancarlo, por lo que hay que tomar otras medidas que puedan hacerse durante el desarrollo normal de un árbol. El grosor del tronco es una medida que está correlacionada con el peso del árbol. Puede obtenerse mediante la medida del diámetro, del perímetro o estimar el área de la sección transversal. Todas ellas son fáciles de realizar y suelen tomarse durante el período de reposo invernal. En experimentos de larga duración, la medida del grosor del tronco se toma anualmente, por lo que conviene señalar con pintura la zona donde se mide para realizarla en la misma todos los años. En árboles frutales suele tomarse a unos 30 cm del suelo, mientras que en árboles forestales, de tallo más alto, suele hacerse a la altura del pecho. Pearce (1976) indica que, a veces, es útil medir el grosor de las ramas y combinarlo en un solo valor que vendría dado por la suma de los grosores dividido por la raíz cuadrada del número de ramas medidas. Esto podría ser de gran utilidad para tener una medida del tamaño del árbol en aquellas especies que no presentan un solo tronco, sino varios bajo una misma copa.

Una medida común del tamaño del árbol lo es también el derivado de medir la altura y difusión de la copa. Con ambas, puede obtenerse el volumen de copa si ésta se asemeja a un esferoide y se aplican las fórmulas siguientes:

- Árboles de mayor altura que anchura (esferoide alargado): $V = \frac{4}{3} \pi a b^2$

- Árboles de mayor anchura que altura (esferoide aplastado): $V = \frac{4}{3} \pi a^2 b$

siendo a y b el semieje mayor y menor, respectivamente.

Si la proyección de la copa de los árboles no es aproximadamente circular debido al sistema de formación o a cualquier otra causa, entonces el volumen de copa podría asemejarse a un elipsoide, cuyo volumen es:

$$V = \frac{4}{3} \pi a b c$$

siendo a , b y c los semiejes del elipsoide.

Esta medida es de utilidad para árboles en los que el grosor del tronco y el peso del árbol no están relacionados, bien por no poseer un solo tronco o por estar sometidos a podas de renovación de la copa, como sucede en muchos casos con el olivo, que rompe la relación anterior.

Las medidas del tamaño del árbol se utilizan con asiduidad en experimentos con plantas leñosas pues, entre otras cosas, sirven para calibrar los árboles y se emplean como covariables en los análisis de covarianza.

Otras medidas del crecimiento vegetativo

La superficie foliar, tomada por un muestreo de hojas dentro de la parcela elemental, es una medida común en muchos experimentos con plantas leñosas. La medida se realiza con facilidad y precisión en un medidor de área foliar, y si se obtiene a la vez el peso seco de la muestra de hojas, se calculan dos parámetros de importancia: el *peso foliar específico*, que es el peso seco de las hojas por unidad de superficie foliar, expresado en g/cm^2 , y su inverso, el *área foliar específica*, que es la superficie foliar por unidad de peso seco de las hojas, expresado en cm^2/g .

El *índice de área foliar* (IAF) se define como la superficie foliar total de una planta por unidad de superficie de suelo que ocupa y el *índice de superficie de copa* como la superficie de la copa de un árbol por unidad de superficie de suelo que ocupa. Ambas medidas se utilizan en estudios sobre eficiencia fotosintética de las plantas, cobertura, etc. y están relacionados entre sí en algunas situaciones. El índice de superficie de copa es más fácil de estimar, por lo que suele utilizarse con frecuencia en estudios con árboles. Este índice suele medirse con instrumentos específicos, aunque también puede estimarse calculando la superficie de un esferoide, cuando la copa de un árbol pueda asemejarse a esta figura, aplicando la fórmula:

$$S = 4 \pi a b$$

La *densidad foliar* se define como la superficie foliar por unidad de volumen de copa, y es otra medida utilizada en diversos estudios para comparar la compacidad de las copas.

MEDIDAS DE LA PRODUCTIVIDAD

Este tipo de medidas en plantas leñosas presenta una complejidad asociada con la vecería, esto es, el fenómeno por el cual a una abundante cosecha le sigue otra de baja o nula producción. Se hace necesario, pues, la aportación de datos de varios años para que los resultados puedan ser comparables. Por otra parte, ya ha sido comentado en varias secciones que, en este caso en particular, la calibración de los árboles antes de comenzar el experimento es de particular importancia para corregir por covarianza los resultados finales.

En muchos experimentos, particularmente los de transferencia de tecnología, se suele estar interesado únicamente en la producción final. Aunque esta medida es interesante, hay otras

que reflejan toda la actividad reproductora del árbol que desembocarán en la producción. Quiere esto decir que la producción es una variable influida por muchas otras y que descomponer la actividad reproductora en medidas de otras variables relacionadas, puede ayudar en muchos casos a encontrar una explicación plausible de los resultados. Las medidas de mayor importancia en trabajos de campo se exponen a continuación.

Medidas de la floración

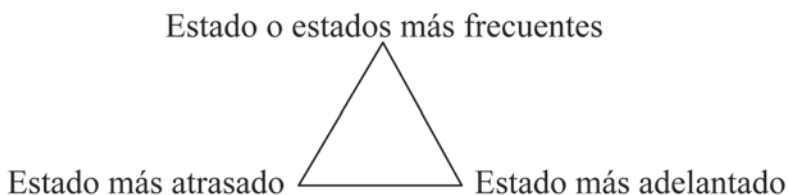
En floración suelen tomarse dos tipos de medidas generales: la cantidad de flor, que mide la potencialidad productiva del árbol en esa campaña, y la época de floración, una medida esencial para el estudio de la polinización y de la incidencia de adversidades meteorológicas, como las heladas, entre otros.

La *cantidad de flor* suele medirse mediante categorías en una escala del 0 (árbol sin flores) al 5 (el máximo de floración) como máximo. En algunos ensayos específicos puede interesar determinar el porcentaje de yemas de flor, que se realizaría por conteo sobre una muestra de ramos fructíferos.


La *época de floración* se establece haciendo visitas periódicas, entre 1 y 3 días según la evolución de la floración, y anotando el estado de los botones florales de acuerdo con los estados fenológicos definidos en cada especie. Por ejemplo, en especies cuyas yemas reproductivas dan lugar a una flor individual, los estados fenológicos suelen definirse como:


- A. Yema invernal
- B. Yema hinchada
- C. Se ve el cáliz
- D. Se ve la corola
- E. Se ven los estambres
- F. Flor abierta
- G. Caída de pétalos
- H. Fruto cuajado


Una forma usual de medir la época de floración es utilizando el método de Fleckinger (1954), con ligeras modificaciones. De acuerdo con este método, en cada fecha de observación y para cada árbol, los datos se anotan en un triángulo equilátero indicando en cada uno de los vértices los estados fenológicos observados en esa fecha de la siguiente forma:





Una vez finalizada la toma de datos en campo, se elaboran los gráficos de floración procediendo de la siguiente forma:

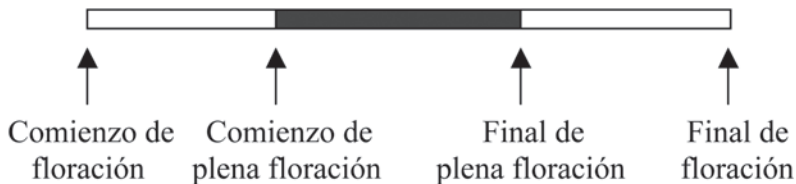
Comienzo de floración: El día en que aparezca por primera vez 

Comienzo de plena floración: El día en que aparezca por primera vez 

Fin de plena floración: El día que aparezca por última vez 

Fin de floración: { El día en que aparezca por última vez 
ó
El día en que aparezca por primera vez 

Esos datos se representan en un gráfico de la forma siguiente:



Medidas del cuajado de frutos

El cuajado suele expresarse como frutos por 100 flores o frutos por 100 inflorescencias, según el hábito de fructificación de la especie. Si la especie suele presentar un alto porcentaje de aborto ovárico, como es el caso del olivo, el cuajado se expresa sobre flores o inflorescencias perfectas o funcionales. En este tipo de medidas es conveniente marcar previamente los ramos sobre los que se realizará la medida, contar el número de flores o de inflorescencias y, posteriormente, en una segunda visita, una vez que hayan concluido los procesos de polinización y fecundación, contar el número de frutos cuajados.

El número de frutos cuajados por cm² de sección transversal de una rama o del tronco y el número de frutos por unidad de longitud del ramo fructífero (frutos/cm), según el hábito de fructificación de la especie, son medidas para estimar la productividad. Estas medidas sue-

len hacerse por muestreo sobre ramas o ramos fructíferos y se realizan en una sola visita a los árboles.

Medidas de la producción

La producción se mide por pesada directa de la cosecha de cada árbol o de cada parcela elemental. En ocasiones, si la parcela elemental es muy grande, se ha contado el número de cajas llenas de fruta que se obtiene en cada parcela y, por muestreo, se determina el peso medio de una caja llena, con lo cual se obtiene la producción total. Esta medida tiene algunos inconvenientes; si el peso medio de la caja se obtiene de una muestra de todo el experimento, hay que tener en cuenta que lo que realmente se obtiene es una medida del volumen de cosecha, pues si un tratamiento produce frutos de un peso específico diferente, o de distinta forma o tamaño que afectan al contenido de la caja, los resultados pueden estar sesgados. En este caso es mejor dejar la medida en número de cajas por parcela. Si el peso medio de la caja se estima dentro de cada tratamiento también presenta inconvenientes, pues si en uno de ellos se obtiene un valor alto se aplicará a todo el tratamiento pudiendo falsear los resultados. La alternativa posible a la pesada de la cosecha total por parcela es obtener el peso medio de la caja dentro de cada parcela elemental, pero esto puede ahorrar poco trabajo en algunos experimentos.

En experimentos con parcelas elementales que contienen muchos árboles, como puede suceder en algunos ensayos de transferencia de tecnología que se realizan en plantaciones privadas, una alternativa es obtener el peso medio de la cosecha por árbol determinado por muestreo sobre algunos árboles de la parcela elemental. Esto ahorra tiempo y, si el tamaño de la muestra es adecuado, se obtiene la misma precisión que con la medida completa.

MEDIDAS DE CALIDAD DE LA COSECHA

En la mayoría de los experimentos no debería bastar con las medidas de producción, sino que deben ir acompañadas de alguna medida de calidad de la cosecha. Entre ellas, cabe destacar por su carácter general las medidas del tamaño del fruto y las medidas del color del fruto. Otro tipo de medidas, como el contenido químico del fruto, la consistencia de la pulpa, las cualidades organolépticas, etc., son de interés en experimentos más específicos. En cualquier caso, la medida suele obtenerse por muestreo dentro de cada parcela elemental.

Medidas del tamaño del fruto

El tamaño del fruto se determina por muestreo dentro de cada parcela elemental, bien calibrando cada fruto individual u obteniendo el peso medio de 100 frutos. Ambas medidas son de utilidad y la decisión sobre una u otra dependerá del objetivo del ensayo y del tiempo disponible. El calibrado de cada fruto individual tiene la ventaja de que los frutos pueden agruparse por calibres comerciales, lo que permite obtener el volumen de cosecha producido dentro de cada uno además de una medida del tamaño del fruto, pero es una medida laboriosa.

En un ensayo realizado con frutos de melocotonero en el que se calibraron todos los frutos recogidos y se obtuvo además el peso medio de 100 frutos sobre el total producido, el coeficiente de variación obtenido fue menor cuando se calibraron los frutos ($CV = 1,1\%$) que cuando se obtuvo el peso de 100 frutos ($CV = 3,5\%$), aunque en ambos casos los coeficientes fueron bajos.

Medidas del color del fruto

Antes de tomar medidas sobre el color del fruto es necesario definir lo que se quiere medir, pues puede interesar la proporción de superficie coloreada (chapa), el color de la chapa o el color de fondo. En algunas variedades ambos colores son fácilmente discernibles, pero en otras el color de fondo puede estar enmascarado por la intensidad del color superficial y complica su medida. El color de la chapa suele estar afectado de una gran variabilidad, por lo que es menos utilizado.

La medida de cualquiera de las características debe hacerse por muestreo y establecer categorías. Nunca debe hacerse sobre una muestra de frutos ya embalados, pues en este caso los frutos suelen colocarse formando capas del mismo color, lo que podría falsear la medida. Si esto es así, hay que sacar los frutos de las cajas para tomar los datos.

Para la medida de la superficie coloreada, Pearce (1976) aconseja establecer las siguientes categorías:

1. hasta $\frac{1}{4}$ coloreado
2. de $\frac{1}{4}$ a $\frac{1}{2}$ coloreado
3. de $\frac{1}{2}$ a $\frac{3}{4}$ coloreado
4. $>\frac{3}{4}$ coloreado.

La medida del color de fondo, siempre que pueda realizarse sin interferencias con el color de la chapa, puede determinarse utilizando cartas de colores o estableciendo categorías en función del color de maduración del fruto. Por ejemplo, en un fruto como la aceituna, cuyo color durante el proceso de maduración varía desde el verde intenso al negro, se han establecido las siguientes categorías:

- 1.- Verde intenso
- 2.- Verde amarillento
- 3.- Verde con manchas rojizas o moradas (envero)
- 4.- Morado
- 5.- Negro

CAPÍTULO 22

MEDIDAS EN PLANTAS HERBÁCEAS

La metodología de la toma de datos en los experimentos con plantas herbáceas no difiere, básicamente, de la de otras plantas. Sin embargo, existen ciertos matices que se relacionan fundamentalmente con el tipo de cultivo y con la disposición de siembra en cada uno de ellos.

En relación a la siembra de la unidad experimental y a la toma de datos de un experimento, las plantas herbáceas se pueden dividir en:

- a) *Plantas de semilla pequeña*: Cereales de invierno, colza, algunas leguminosas, etc.
- b) *Plantas de escarda*: Girasol, maíz, algodón, remolacha, sorgo, etc.

Los experimentos realizados con plantas de semilla pequeña se suelen sembrar con espaciamientos entre líneas muy cortos (15-35 cm), mientras que los realizados con plantas de escarda se suelen sembrar con espaciamientos grandes (50-100 cm), con el objetivo de poder dar alguna labor entre líneas durante los primeros estadios del cultivo. Las medidas a realizar durante el ciclo de cultivo y al final de ciclo varían, en general, con el tipo de planta.

MEDIDAS DURANTE EL CICLO VEGETATIVO

Durante este periodo se suele determinar la altura de las plantas, su número, el número de hojas, la fecha de floración o de espigado, el tamaño de las inflorescencias (espigas, racimos, vainas, capítulos), etc. En función del tipo de plantas, tenemos:

- a) *Plantas de semilla pequeña*. Las medidas se suelen tomar en una o varias muestras de la parcela elemental. Estas muestras se seleccionan en un muestreo al azar, lanzando una cuadrícula de $\frac{1}{2}$ m² sobre la parcela y realizando las medidas sobre los individuos que se encuentran dentro de la cuadrícula.
- b) *Plantas de escarda*. Las medidas se suelen tomar en un número de plantas determinado que se escogen de acuerdo con un muestreo sistemático (por ejemplo la 2-4-6-8-10) dentro de cada surco, excluidos los surcos guarda o surcos borde.

Además de las medidas anteriores, durante este periodo se suelen tomar datos sobre incidencia de plagas y enfermedades. En este caso, las medidas se toman en la totalidad de la parcela, puesto que la distribución de los daños no tiene por qué ser aleatoria. Para la toma de datos de las infestaciones y/o incidencias de ciertas enfermedades, sobre todo aquellas enfermedades foliares como royas, mildius, oidios, etc., en las que se ha de valorar la superficie cubierta por pústulas, chancros, etc., existen una serie de tablas disponibles en al-

gunos manuales de patología vegetal y también en la *web*. En función de esto se establecen estimaciones o categorías para determinar el grado de infección.

Se ha de tener en cuenta que para la toma de ciertos datos en cada una de las unidades experimentales, el estado de crecimiento del cultivo ha de encontrarse en idéntica fase en todas y cada una de esas unidades experimentales. Para cerciorarse de este extremo, se han de conocer perfectamente los estados de crecimiento de la especie con la que se está experimentando. Por lo general, los estados de crecimiento se numeran en función del desarrollo alcanzado por el cultivo en un momento determinado. Suele distinguirse entre los estados vegetativos (V1, V2, V3, etc.) numerados según el número de hojas o de nudos, y los estados reproductivos (R1, R2, R3, etc.) numerados según el desarrollo de la flor, de la inflorescencia o del fruto. Estos estados se pueden encontrar en cualquier manual que trate sobre el cultivo de la especie correspondiente. Existen también páginas *web* en las que se puede obtener información sobre los estados de crecimiento de los cultivos herbáceos más comunes.

MEDIDAS DE LA COSECHA

a) Plantas de semilla pequeña. Lo normal es recoger la parcela elemental entera, pero si ésta es extremadamente grande se recogen una o dos pasadas de cosechadora experimental. En cualquier caso, se ha de medir cuidadosamente la superficie realmente cosechada, pues posteriormente se refiere el rendimiento de esa superficie a kg/ha. Una vez realizado el pesaje de la producción de la unidad experimental, se suelen tomar muestras de 100 a 1.000 g para realizar los análisis de las características cualitativas que, lógicamente, varían con el tipo de cultivo.

b) Plantas de escarda. Por lo general, las parcelas elementales se componen de 2 a 4 surcos rodeados de surcos guarda para evitar el efecto borde. Además, se suelen despreciar las primeras y últimas plantas de los surcos a recoger, todo ello con el objeto de evitar todos los efectos bordes posibles. Se ha de ser extremadamente cuidadoso en referir correctamente la cantidad de semilla recolectada a la superficie real sobre la que se ha recolectado dicha semilla. Téngase en cuenta que al ser las unidades experimentales relativamente pequeñas, la adición o la sustracción de la semilla de una cabeza, mazorca, etc, que no corresponda a la superficie real recolectada, puede sesgar significativamente los datos del experimento.

En estos cultivos la recolección de las unidades experimentales puede hacerse bien con una cosechadora de experimentos o bien a mano, llevando la producción de cada parcela a un sitio específico en donde se realiza la trilla y limpia de la semilla. A continuación se realiza la pesada de la producción, tomándose una muestra para la realización de los análisis de las características cualitativas de las semillas (contenido de aceite, proteínas, ácidos grasos, etc.).

CAPÍTULO 23

INTERPRETACIÓN Y PRESENTACIÓN DE RESULTADOS

CONSIDERACIONES GENERALES

La experimentación es la única vía para generar nuevos conocimientos en Agronomía. La complejidad de los fenómenos biológicos, cuyos mecanismos no son del todo conocidos, y la de las relaciones entre la planta y el medio, ha impedido hasta la fecha el desarrollo de métodos matemáticos que permitan predecir el comportamiento de un cultivo en unas condiciones determinadas o el efecto que podría producir una determinada técnica, por lo que el método de ensayo y error es la única herramienta válida para su aplicación a fenómenos biológicos.

El desarrollo de experimentos en Agronomía es un proceso caro, que consume tiempo, trabajo y dinero antes de poder obtener alguna conclusión aplicable en la práctica agrícola. El último paso de la experimentación es el análisis de los resultados obtenidos, en principio lo más fácil y, a la vez, lo más peligroso del proceso, pues un análisis equivocado de los datos o una mala interpretación de los mismos puede inutilizar todo el esfuerzo realizado durante la planificación y el desarrollo de los experimentos. Aunque esto parezca una obviedad, la realidad es que muchos de los experimentos que se realizan en Agronomía, entendida ésta con carácter general, adolecen de este problema. Basta hojear revistas de relevancia dentro de este campo publicadas en los últimos 30 años, para encontrar artículos que denuncian el mal uso de la estadística aplicada al diseño de experimentos (Little, 1978; Little, 1981; Gates, 1991; Dyke, 1997). El objetivo de estas publicaciones fue llamar la atención de los autores para que recapaciten sobre los aspectos estadísticos de su trabajo, desde la planificación y el diseño de los experimentos hasta el análisis, y a presentar los resultados de una forma clara y comprensible por el lector utilizando los métodos estadísticos apropiados.

El problema se complica en la actualidad por el empleo de programas informáticos. Antes de su disposición para el uso general, los autores necesitaban conocer los procedimientos estadísticos porque tenían que realizar los cálculos, algo que hoy hace el ordenador. Pero éste se limita a realizar los cálculos aritméticos, que representan una labor tediosa en su ausencia, pero realiza los cálculos que se le indique sin que pueda interpretar si se ajustan o no al diseño experimental seguido. El desconocimiento de los procedimientos estadísticos puede provocar errores a la hora de elegir el procedimiento adecuado, lo que llevaría a la obtención de resultados y conclusiones falsas. En definitiva, lo que representa una ventaja, como es el empleo de los programas estadísticos, puede convertirse en un agente peligroso si no se indica correctamente lo que debe calcular y no se hace un uso adecuado de los datos.

Antes de comenzar el análisis es aconsejable revisar los datos y observar si hay mucha variación dentro de los tratamientos. A veces se puede detectar que alguna medida fue mal realizada, hubo problemas que alteraron esa toma de datos, etc. En ocasiones es posible repetir una medida o detectar algún accidente que justifique esa variabilidad, con lo que se

puede corregir o, al menos, tener un criterio claro para eliminar el dato. Si todo está correcto, o se han tomado las medidas para que así lo sea, se procede al análisis de los datos en la forma indicada en los capítulos anteriores. En este punto, no debe quedar ninguna duda sobre lo que realmente se ha investigado y, en consecuencia, la respuesta que se persigue con el experimento. Tampoco se debe dudar sobre el método estadístico a aplicar a los datos, que será el que corresponde al diseño experimental seguido y no a otro que en ese momento parezca más oportuno, pues se pueden falsear los resultados. Es necesario mencionar aquí que la unidad de un experimento es la parcela elemental y, por consiguiente, los datos para el análisis deben expresarse por parcela elemental, no por individuos en el caso de que la parcela elemental esté compuesta por varios. De ser necesario, se procede a la transformación de los datos originales antes del análisis con el objetivo de que se cumplan los supuestos del análisis, algo que siempre hay que revisar.

Realizado el análisis de varianza o de covarianza oportuno, si se obtiene que la prueba F es significativa, se sigue el proceso normal de separación de medias. En este momento es necesario prestar una cuidadosa atención, pues es un punto donde se cometen muchos errores en el análisis de los datos que suelen llevar a interpretaciones erróneas o a enmascarar resultados que podrían aportar una información valiosa. Hay una tendencia general a la utilización de procedimientos de comparación múltiple, como el test de rango múltiple de Duncan, el test de Tukey, el de Scheffé y otros, aún cuando esos procedimientos no están indicados. En este sentido se utilizan con frecuencia para comparar tratamientos que son de naturaleza factorial o para comparar distintos niveles de un factor cuantitativo. En el primer caso es necesario descomponer la suma de cuadrados en efectos principales y en la interacción, y no incluir todo como si se tratara de un único factor. En el segundo caso, es necesario realizar un contraste polinómico, esto es, separar las medias por regresión; si ésta resultase significativa, quiere decir que todos los tratamientos, incluidos los que podrían haberse realizado dentro del intervalo estudiado, muestran unos efectos que son significativamente diferentes. Los procedimientos de comparación múltiple son apropiados cuando los tratamientos consisten en un conjunto de materias no relacionadas, como variedades o productos químicos. Aún así, hay que considerar si los tratamientos pueden formar grupos que puedan ser comparados, como variedades resistentes o susceptibles, grupos de familias de productos fungicidas, etc. Hay que mencionar aquí también, como ha sido indicado por algunos autores, el poco uso que se hace en la actualidad de la mínima diferencia significativa (MDS), un procedimiento útil pero que a muchos le parece menos sofisticado que los procedimientos de comparación múltiple. No es raro encontrar que se ha aplicado uno de estos procedimientos para la separación de tan solo dos medias!

Si la prueba F no resulta significativa, se termina ahí el proceso. En experimentos que incluyen más de un factor la suma de cuadrados hay que descomponerla en los efectos principales y en las interacciones entre ellos, como se ha indicado, por lo que una sola F no significativa no implica abandonar el proceso, sino estudiar el efecto de las respuestas significativas.

Es necesario, no obstante, reflexionar sobre el significado de una F no significativa. Cuando ésta aparece en el análisis puede ser que realmente no existan diferencias significativas en

la respuesta de la planta a los tratamientos, o que esa falta de significación sea exclusiva de nuestro experimento porque el diseño aplicado no haya sido lo suficientemente preciso para detectar esas diferencias. El investigador no debe conformarse con el resultado sin tratar de averiguar cuál puede ser la realidad. Esto no significa que si la F no resulta significativa haya que sacar conclusiones en el caso de que la media de un tratamiento sea superior a la de otro, sino que hay que dudar acerca de si la falta de significación es real o es la obtenida en nuestro experimento. Un procedimiento que ayuda a esclarecer los hechos es el cálculo del coeficiente de variación. Si resultase elevado, y no puede decirse qué es elevado en el coeficiente de variación sin conocer la variable en estudio, podría sospecharse que la variabilidad obtenida en el experimento haya enmascarado la realidad sobre el efecto de los tratamientos. Un método eficaz de corregir la alta variabilidad, además de poner atención en la precisión de las medidas, es aumentar el número de repeticiones, por lo que el planteamiento de un nuevo experimento con mayor número de repeticiones es la forma más correcta de actuar en tales circunstancias. No siempre pueden obtenerse conclusiones válidas con la realización de un simple experimento.

Pueden obtenerse, en ocasiones, resultados que no son esperados cuando se planteó el experimento. Si se estudian varios factores a la vez, esto puede ser normal, pero la interpretación de los resultados no siempre se puede hacer con claridad o se complica al manejar varios factores. Es necesario pensar de nuevo en cuales fueron los objetivos del experimento y los factores que realmente interesan estudiar para dar respuesta al mismo. En el caso que se expone, de un diseño factorial con varios factores, cabe preguntarse qué interpretación puede darse con datos agronómicos a una interacción de, por ejemplo, cuatro factores. Aplicando el sentido común, parecería lógico realizar nuevos experimentos en el que se limitase el número de factores a estudiar en cada uno, de manera que se pudieran obtener resultados que puedan interpretarse de una manera inteligente. Si los resultados inesperados se producen en otras circunstancias, la solución es similar; el investigador no puede conformarse con esos resultados, sino que ha de profundizar en futuros ensayos y tratar de averiguar la causa que los produjo.

PRESENTACIÓN DE RESULTADOS

La elaboración de tablas que contengan los datos analizados del experimento suele ser, con frecuencia, la mejor forma de presentar la mayoría de los resultados, aunque otros, por el contrario, se presentan en forma de figuras o gráficos. En cualquier caso, los datos hay que presentarlos de la forma más clara posible, y deben interpretarse sin que el lector necesite recurrir a la lectura del texto. Las instrucciones para autores que elaboran las editoriales de las revistas científicas o técnicas suelen dar algunas normas para la elaboración de tablas y figuras que, con ligeras diferencias de forma, suelen coincidir en el contenido. Lo importante es la claridad y la información que contienen. En este punto es necesario mencionar de nuevo las salidas de los programas informáticos a través del ordenador. En todos se genera una salida de datos interesante, pero no todos son útiles en todos los casos y, por supuesto, no siempre han de ser publicados. La salida generada en el ordenador debe servir para elegir los datos que se requieren para ese análisis y hacer un uso correcto de ellos.

A continuación se muestra, a modo de ejemplo, la presentación de datos en diversos supuestos que han sido publicados en revistas agronómicas.

Experimento en bloques al azar con medida de una variable cualitativa

El ejemplo está extraído de un experimento en el que se estudió la respuesta de árboles cloróticos de la variedad de olivo ‘Manzanilla’ a la inyección en el tronco de seis compuestos de hierro, comparados con un testigo sin inyectar. Se utilizaron cuatro bloques y se aplicaron siete tratamientos, incluido el testigo; un árbol individual constituyó la parcela elemental. Los compuestos utilizados no podían agruparse entre sí, pues se eligieron productos de composición química diferente. La variable que se midió fue el color de las hojas en varias épocas, obtenido mediante categorías utilizando una escala visual del 0 al 3. Los resultados se muestran en la Tabla 23.1.

Tabla 23.1.- Clorosis mostrada durante tres años por árboles de ‘Manzanilla’ inyectados en julio de 1988 con varios compuestos de hierro¹ (*HortScience* 1993, 28: 192-194).

Tratamiento ²	Fecha de observación				
	Julio 1988	Nov. 1988	Julio 1989	Nov. 1989	Julio 1990
Control	2,2	2,5	2,5	2,4	2,0
Sulfato ferroso	2,2	0	0,1	0,2	0,5
Fe-EDDHA	2,1	0,1	1,4	1,0	1,0
Fe-lignosulfonato	2,0	1,0	1,7	1,1	1,2
Fe-EDDHMA	2,0	0,1	0,2	0,6	0,7
Citrato férrico	2,2	1,0	1,9	1,9	2,0
Fe-heptagluconato	2,2	0,2	1,2	0,6	1,2
Significación ³	NS	**	**	**	*

¹ Clorosis medida en una escala visual del 0 = sin clorosis al 3 = clorosis muy pronunciada.

² Cada tratamiento contenía 0,5 g Fe, excepto el control.

³ NS, *, **: No significativo o significativo a $P \leq 0,05$ o $P \leq 0,01$, respectivamente, obtenida por el método no paramétrico de Friedman.

Al tratarse de una variable cualitativa el análisis correcto de los datos resulta de la aplicación de métodos no paramétricos, como se explica en el capítulo 19. En el ejemplo se ha realizado un análisis de varianza por el método de Friedman, que muestra que en el momento de la inyección férrica los árboles se presentan claramente cloróticos, sin que existan diferencias significativas entre los tratamientos. Cuatro meses después, el análisis muestra diferencias significativas entre tratamientos. Para la separación de medias, esto es, para saber

qué tratamientos difieren entre sí y cuales no, se puede considerar lo que se indicó en el capítulo 20 para las categorías, en el sentido de que diferencias inferiores a 0,5 puntos no son de importancia en la mayoría de los casos. Obsérvese el efecto persistente del sulfato ferroso y del Fe-EDDHA y el poco persistente del citrato férrico.

Experimento en bloques al azar con medida de una variable cuantitativa y un factor no cuantitativo

En este supuesto la separación de medias que sigue al análisis de la varianza que muestra una F significativa, se realiza mediante la aplicación de la mínima diferencia significativa o la de un procedimiento de comparación múltiple. La presentación de la tabla de resultados es similar a la anterior, pero las diferencias entre tratamientos se indican mediante letras que siguen a las medias, de manera que medias seguidas con la misma letra no muestran diferencias significativas al nivel que se indique, por lo general al 5% o al 1%.

En el ejemplo que se expone el objetivo fue conocer la respuesta de encinas afectadas por *Phytophthora cinnamomi* al tratamiento mediante inyección al tronco de dos productos fungicidas, que se compararon con árboles testigo sin inyectar. Se diseñó un experimento en bloques al azar con tres tratamientos y cuatro bloques; puesto que se esperaba una alta variabilidad debido tanto al carácter de la enfermedad como al hecho de tratarse de encinas adultas procedentes de semilla, la parcela experimental la constituyó una superficie de dehesa de 10 ha, donde se marcaron 20 árboles sobre los que se realizaron las medidas. La amplitud de la parcela experimental se aprovechó también para estudiar costes de aplicación y demostración de la técnica. Se realizó un único tratamiento en la primavera de 1993. La Tabla 23.2 muestra el crecimiento vegetativo de los árboles en los tres años siguientes al tratamiento.

Tabla 23.2.- Efecto de los tratamientos de inyección con carbendazima y fosfonato potásico sobre el crecimiento vegetativo de las encinas (*European Journal of Forest Pathology* 1999, 29: 29-38).

Tratamiento ²	Longitud del brote (cm.) ¹		
	Febrero 94	Febrero 95	Febrero 96
Control	5,5 a	7,3 a	7,5 a
Carbendazima	5,5 a	7,6 a	8,2 ab
Fosfonato K	7,1 a	9,3 b	12,1 b
CV (%) ³	18,0	7,1	20,2

¹Cada dato representa la media de 320 brotes.

²Dentro de cada columna, letras diferentes a continuación de las medias indican diferencias significativas a $P \leq 0,05$ según el test de Tukey.

³Coefficiente de variación.

La inclusión de los coeficientes de variación es de gran utilidad para interpretar los resultados, pues da una idea de la sensibilidad del análisis. En el ejemplo, los coeficientes obtenidos no parecen elevados, dado el carácter del material en estudio. El efecto de los tratamientos durante el primer año de crecimiento no muestra diferencias significativas pero cabría preguntarse, revisando las medias obtenidas, si la variación estimada por el CV podría enmascarar el efecto del fosfonato potásico. En el segundo año disminuye claramente el CV y aparecen diferencias significativas entre los tratamientos. En el último año se obtuvieron diferencias significativas a pesar del mayor valor del CV.

El ejemplo ilustra también una de las particularidades de las plantas leñosas en relación con el diseño de experimentos. Es evidente que los árboles experimentales son los mismos durante los tres años de estudio, y que las observaciones realizadas en un año guardan relación con la de años anteriores; de hecho, se observa que al avanzar los años, los árboles tratados con fosfonatos crecen más, lo que puede interpretarse como un signo de recuperación paulatina de la enfermedad. Como se indicó en el capítulo 21, el análisis de una serie de experimentos realizados en años distintos no es aplicable a este caso debido a esas particularidades. El crecimiento acumulado durante los tres años podría servir para comparar los tratamientos, pero la descomposición por años muestra una imagen de evolución que ayuda a interpretar los datos.

Experimento en bloques al azar con medida de una variable cuantitativa y un factor cuantitativo

En este caso la separación de medias se realiza mediante un contraste polinómico, pues se trata de comparar distintos niveles de un factor cuantitativo. El ejemplo trata de un experimento en el que pretendía estudiar la respuesta del olivo a la aplicación foliar de varias dosis de un extracto de leonardita que contenía sustancias húmicas. La Tabla 23.3 corresponde a un ensayo realizado con árboles adultos de la variedad ‘Hojiblanca’ cultivados en secano. El diseño fue en bloques al azar con cuatro bloques y tres tratamientos, y una parcela elemental compuesta de cinco árboles. Los tratamientos consistieron en la aplicación foliar del extracto de leonardita a las concentraciones del 0,5% y del 1% en primavera antes de la floración, y en un testigo sin tratar.

Tabla 23.3.- Efecto de la aplicación foliar de extractos de leonardita sobre el crecimiento vegetativo y la producción del cultivar de olivo ‘Hojiblanca’ (*Scientia Horticulturae*, 1996, 66: 191-200).

Concentración del extracto (%)	Longitud del brote (cm.)	Cuajado (Frutos/cm.)	Producción (kg/árbol)
0	6,64	0,42	24,61
0,5	8,08	0,55	29,08
1	8,03	0,44	25,49
Significación ¹	L*	Q*	NS
CV (%) ²	9,8	12,2	14,2

¹L = lineal; Q = cuadrática; NS = No significativo; * P ≤ 0,05.
²Coefficiente de variación.

La significación en este caso no se representa mediante letras sino indicando si la relación encontrada ha sido lineal (L), cuadrática (Q) o cúbica (C), en su caso, y el nivel de probabilidad, representado por asteriscos (*, **, *** a $P \leq 0,05$; 0,01 ó 0,001 respectivamente). De nuevo conviene indicar la utilidad de incluir el coeficiente de variación obtenido.

Experimento factorial con interacción

Cuando se combinan en un mismo experimento más de un factor, lo que se pretende conocer es si existe o no interacción entre ellos, esto es, si los factores actúan independientemente o no. En el caso de existir una interacción significativa, los factores no actúan de forma independiente, por lo que los efectos principales dejan de tener sentido por sí mismos.

Una forma correcta y práctica de presentar las interacciones es de forma gráfica, pues se puede apreciar a simple vista la convergencia o la intersección de las tendencias. En el ejemplo ilustrativo se estudió la salida del reposo invernal de yemas de melocotonero tras la aplicación de productos químicos. Se diseñó un experimento factorial con dos factores, aplicación de tiourea a niveles del 0%, 0,25% y 1,5% y aplicación de nitrato potásico también a tres niveles, 0%, 1% y 5%, y cuatro repeticiones. Estas concentraciones no trataban de responder a series aritméticas, sino que fueron elegidas en función de los resultados recogidos en la literatura sobre el tema. Los tratamientos se aplicaron a ramos fructíferos recogidos en campo a intervalos regulares durante el invierno, que se introdujeron en cámaras a 20 °C para provocar la brotación. La parcela elemental estuvo compuesta por cinco ramos fructíferos, donde se realizaron las mediciones. Los resultados, tras la acumulación de 328 horas-frío, se recogen en la Fig. 23.1.

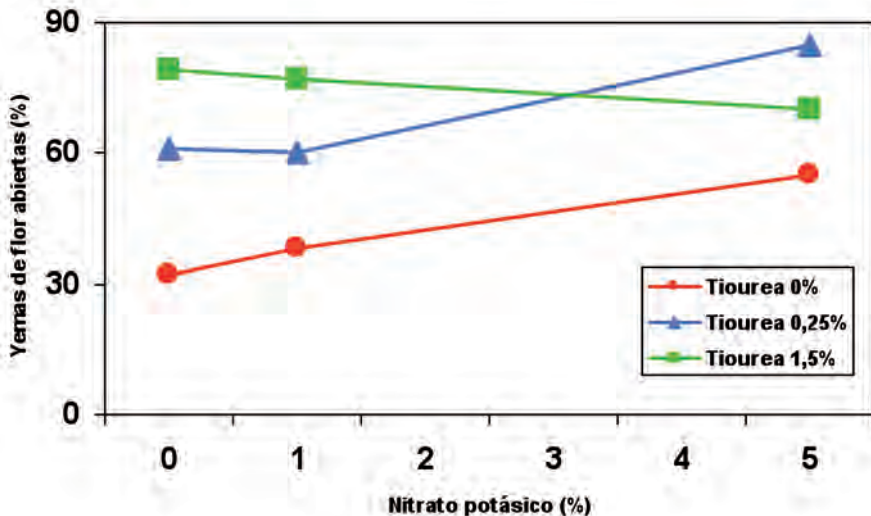


Fig. 23.1.- Interacción entre tiourea y nitrato potásico en la apertura de las yemas de flor de melocotonero 'Springcrest' tras la acumulación de 328 horas-frío (Journal of Horticultural Science 1987, 62: 457-461).

En la figura se aprecia que al aumentar la concentración de tiourea se favorece la apertura de las yemas de flor, pero a bajas concentraciones de nitrato potásico, pues cuando ésta aumenta se produce la interacción, expresada por la intersección de las curvas. El efecto de la tiourea depende, pues, de la concentración de nitrato potásico.

Los datos pueden expresarse cambiando la posición de los factores, como se muestra en la Fig. 23.2, y la interacción se aprecia de igual forma por la intersección de las curvas, pero quizá la interpretación de los resultados resulta más fácil.

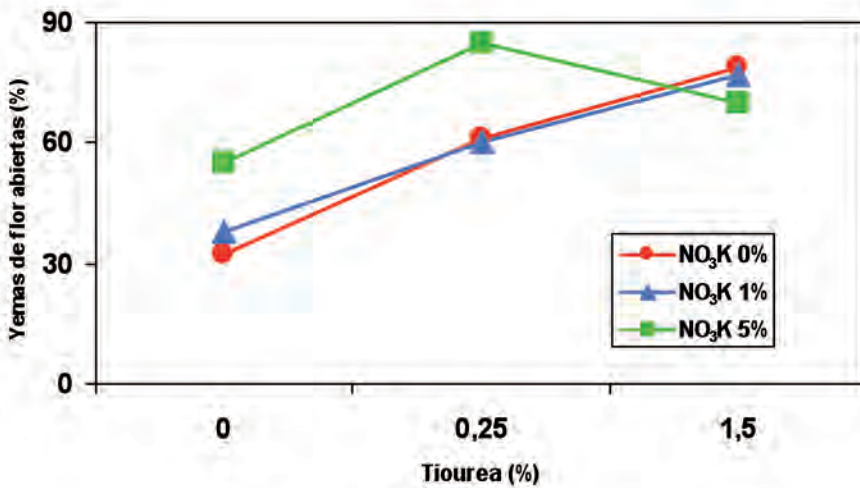


Fig. 23.2.- Interacción entre tiourea y nitrato potásico en la apertura de las yemas de flor de melocotonero 'Springcrest' tras la acumulación de 328 horas-frío.

Experimento factorial sin interacción

Si la interacción no resulta significativa indica que los factores actúan independientemente, por lo que hay que prestar atención a los efectos principales que muestren significación. Se agrupan los datos correspondientes a cada factor estudiado y, en el que haya resultado una F significativa, se separan las medias de la forma indicada en los ejemplos anteriores.

Presentación de regresiones

Cuando interesa el estudio de la regresión entre dos variables, la representación gráfica resulta de gran utilidad. La forma más apropiada es presentar las líneas de tendencias y los datos reales, representados éstos por puntos, pues el lector puede hacerse una idea aproximada sobre la magnitud de las desviaciones de regresión. En el ejemplo se recoge un estudio sobre anillado en melocotoneros y nectarinas. La Fig. 23.3 recoge la tolerancia de dos

variedades al anillado, que fue evaluada midiendo la formación de callo sobre el descortezado. La medida se realizó al final del periodo de crecimiento, antes de la poda invernal, y se expresó como el porcentaje de circunferencia en la que el callo cerró la herida.

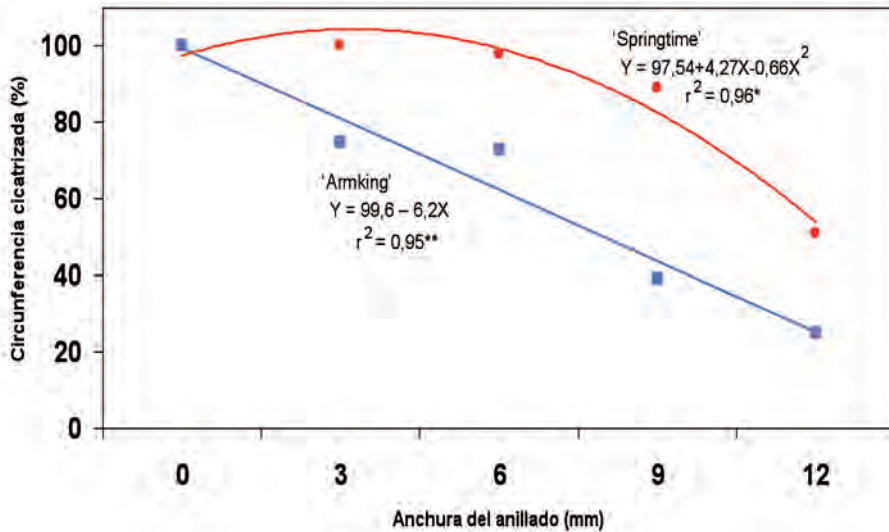


Fig. 23.3.- Cicatrización de la herida provocada por el anillado a diferentes anchuras en 'Springtime' y 'Armking' (Journal of Horticultural Science 1987, 62: 463-468).

Las líneas de regresión deben identificarse en la figura, en el ejemplo con indicación de las variedades a las que corresponden, y la ecuación de regresión debe escribirse junto a ellas incluyendo el coeficiente de determinación y la significación de la regresión, expresada por asteriscos junto al valor de r^2 . Hay que indicar que es el valor de r^2 , el coeficiente de determinación, y no el de r , el coeficiente de correlación, el que debe incluirse junto a la ecuación. Aunque algebraicamente el primero es el cuadrado de r , lo que expresa es la proporción de la variabilidad de la variable dependiente que es explicado por la independiente.

En ocasiones las ecuaciones de regresión y los r^2 se escriben en el pie de la figura, indicando que los datos se ajustaron a esas funciones. La disposición en uno u otro lugar es algo a decidir por el sentido común, pues si la información que se incluye dentro de la figura impide al lector su interpretación por la densidad de información incluida, es preferible ponerla al pie para facilitar su lectura.

En el ejemplo se aprecia claramente que la respuesta de la nectarina 'Armking' al tratamiento es lineal, mientras que en el melocotonero 'Springtime' es cuadrática, lo que indica que esta última variedad es más tolerante al anillado pues se consigue una cicatrización casi total siempre que la anchura del anillado no exceda de 6 mm. En 'Armking', por el contrario, el anillado que se practique a cualquier anchura dejará parte de la herida sin cicatrizar, y la proporción de herida cicatrizada disminuirá conforme la anchura del anillado sea mayor.

Presentación de datos por la media y el error estándar de la media

Esta es la forma más sencilla de presentar los resultados y, a veces, la más eficaz. Muchas figuras suelen presentarse de esta forma, excepto en los casos tratados en el apartado anterior donde se realiza un estudio de la regresión. Hay que considerar, no obstante, que el error estándar de la media da información sobre la variación de cada media y su representación es sencilla y de gran utilidad en muchos casos, pero no es útil para comparar medias. La presentación de resultados de esta forma debería limitarse, pues, a aquellos casos en los que el autor desea mostrar la variación alrededor de las medias o cuando las varianzas no sean homogéneas.

Un ejemplo del primer caso, probablemente el más común, es el representado en la Fig. 23.4. Se estudiaba en ese trabajo la movilización del boro en hojas de olivo en un período próximo a antesis. Se tomaron, para ello, muestras periódicas de hojas jóvenes y de hojas maduras de varios brotes fructíferos procedentes de cuatro árboles en carga. La figura muestra los cambios de la concentración de boro en ambos tipos de hojas durante el periodo estudiado.

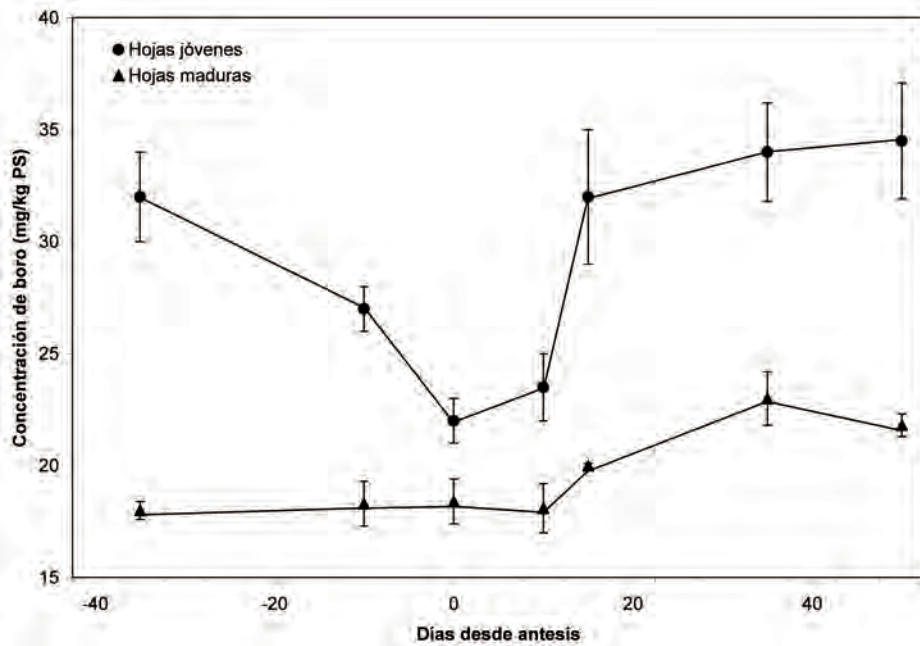


Fig. 23.4. Cambios en la concentración de boro en hojas jóvenes y maduras de olivo en fechas próximas a antesis (*HortScience* 1991, 26: 867-870).

La regresión entre la concentración de boro y los días desde antesis podría haberse realizado, pero no tiene sentido, pues lo que los autores indican es la diferente concentración de boro

en ambos tipos de hojas y el hecho de que en hojas jóvenes encuentran un mínimo en la concentración en antesis que no aparece en las hojas maduras. Interpretan que el boro de las hojas jóvenes se moviliza hacia las flores para atender los procesos de floración, mientras que el de las hojas maduras permanece inmóvil pues forma parte de la estructura de las mismas. La representación de la figura con la media y su error estándar parece en este caso la más simple y adecuada.

La homogeneidad de las varianzas es uno de los requisitos previos al análisis de varianza. Si el test de Bartlett indica que existe heterogeneidad, el análisis no podría practicarse a menos que mediante una transformación de los datos se consiga homogeneizar las varianzas. De no conseguirse, la presentación de resultados de la forma indicada anteriormente no procedería. La solución es presentar la media de los tratamientos y su error estándar y, para un nuevo experimento, corregir los problemas que hubieran causado esa heterogeneidad. Un ejemplo de este caso podría ser el de un ejemplo anterior donde se recogía un estudio sobre el tratamiento de encinas enfermas mediante inyecciones al tronco de productos fungicidas. Como la parcela elemental era grande, se realizaron medidas sobre el rendimiento de la operación en función de la densidad del sotobosque. Los resultados se muestran en la Tabla 23.4.

Tabla 23.4.- Efecto de la densidad del sotobosque sobre el número de árboles inyectados por día¹ (*European Journal of Forest Pathology* 1999, 29: 29-38).

Sotobosque	Inyecciones colocadas por día	Árboles inyectados por día
Ausente	660 ± 13	189 ± 4
Medio	585 ± 16	167 ± 5
Denso	420 ± 20	120 ± 6

¹Valores expresados por la media ± error estándar.

Es evidente que la densidad del sotobosque no representaba ningún tratamiento, sino que se identificaron parcelas con diferente densidad dentro de cada bloque sobre las que se realizaron las mediciones. El tamaño de esas parcelas no era, lógicamente, homogéneo ni cabía sospechar que las varianzas fueran homogéneas. Los resultados presentados de esa forma parecen indicar que el número de árboles inyectados disminuye conforme aumenta la densidad del sotobosque, algo que parece ajustado a la lógica.

APÉNDICE 1

CONCEPTOS ESTADÍSTICOS

Amplitud (*intervalo de variación, recorrido*). Índice de dispersión que consiste en la diferencia entre el valor mayor y el valor menor de un conjunto de datos. La amplitud es la estimada más sencilla de la desviación típica poblacional (σ), pues basta multiplicar la amplitud por un coeficiente que depende del tamaño de la muestra para obtener σ . Estos coeficientes están

tabulados, pero para muestras de tamaño $n \leq 10$ σ puede estimarse como $\sigma = \frac{\text{amplitud}}{\sqrt{n}}$.

La eficiencia relativa de la estimación de σ comienza a ser poco satisfactoria para $n > 10$.

Coefficiente de variación. Medida de dispersión relativa que se define como el cociente entre la desviación típica y la media, y se expresa normalmente en porcentaje:

$$CV = \frac{\sigma}{\mu} 100 \quad \text{y en muestras} \quad CV = \frac{s}{\bar{X}} 100$$

El CV es útil para conocer el éxito de un experimento y si resulta más elevado de lo esperado habría que buscar las causas para aumentar la precisión. Sin embargo, no es posible saber si un CV es muy grande o muy pequeño sin tener experiencia con datos similares. Como el CV es la razón de dos promedios que tienen la misma unidad de medida, es independiente de la unidad que se utilice, lo que lo hace útil para comparar distribuciones con unidades diferentes.

Contraste de hipótesis. Procedimiento que permite aceptar o rechazar si determinadas afirmaciones son ciertas o falsas en función de los datos observados de una muestra. Existe el riesgo de que la conclusión a la que se llegue sea falsa, por lo que hay que tener un cierto grado de precisión, que se denomina nivel de confianza.

Covarianza. La media de los productos entre las desviaciones respecto a su media de los valores de dos variables. Indica el grado de variación conjunta de dos variables, y se expresa por:

$$\sigma_{XY} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{n}$$

Cuantiles. Valores que dividen el conjunto de datos ordenados por magnitud en partes iguales. Los más usados son los cuantiles, los deciles y los percentiles.

Cuartiles. Valores que dividen el conjunto de datos ordenados por magnitud en cuatro partes iguales. Se designan Q_1 , Q_2 y Q_3 y se llaman primer, segundo y tercer cuartil, respectivamente. El Q_2 coincide con la mediana. $Q_1 = 15$ indica que el 25% de los datos es inferior o igual a 15.

Deciles. Valores que dividen el conjunto de datos ordenados por magnitud en 10 partes iguales. Se designan D_1 , D_2 , D_3 , ..., D_9 . El D_5 coincide con la mediana.

Desviación media. Medida de dispersión igual a la media aritmética de los valores absolutos de las desviaciones de los datos respecto a su media.

$$DM = \frac{\sum |X_i - \bar{X}|}{n}$$

Desviación típica. Medida de dispersión igual a la raíz cuadrada de la media de los cuadrados de las desviaciones respecto a la media. Es la raíz cuadrada de la varianza. En una población:

$$\sigma = \sqrt{\frac{\sum (X_i - \mu)^2}{n}}$$

Desviación típica de la muestra. Idéntica a la desviación típica de la población, pero usando de divisor $(n-1)$:

$$s = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}}$$

Error de tipo I. El que se produce cuando se rechaza la hipótesis nula H_0 cuando es verdadera. La probabilidad es igual al nivel de significación α .

Error de tipo II. El que se produce cuando se acepta la hipótesis nula H_0 siendo falsa. Es lo mismo que rechazar H_1 siendo verdadera, cuya probabilidad se representa por β .

Error estándar de la media. Desviación típica de una población de medias obtenida de todas las posibles muestras del mismo tamaño de una población. Es un indicador de la precisión lograda en la estimación de la media mediante muestras. Se expresa por:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Error estándar de la diferencia entre dos medias. Es la desviación típica de la diferencia entre dos medias \bar{X}_1 y \bar{X}_2 sacadas de poblaciones de varianza σ_1^2 y σ_2^2 . Como la varianza de una diferencia es la suma de las varianzas, la desviación típica de esa diferencia se expresa por:

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Estadístico. Es una característica de una muestra, como la media o la desviación típica, que suele utilizarse para estimar un parámetro. Se representa con letras latinas (\bar{X} , s , ...). Véase también parámetro.

Grados de libertad. El número de observaciones independientes en la muestra (n), es decir, el tamaño de la muestra, menos el número k de parámetros de la población que deben ser estimados a partir de observaciones de la muestra. **2.** Número de relaciones lineales independientes que se pueden establecer entre los datos de una muestra.

Hipótesis alternativa (H_1). Es el suceso complementario a H_0 . En el caso siguiente:

$$H_1 : t_1 \neq t_2$$

$$H_1 : t_1 < t_2$$

$$H_1 : t_1 > t_2$$

Hipótesis nula (H_0). Indica la afirmación sobre los parámetros poblacionales que se van a contrastar. Por ejemplo, si deseamos probar si un tratamiento es mejor que otro, se formula la hipótesis nula de que *no* hay diferencia entre los tratamientos, es decir, las diferencias observadas se deben a fluctuaciones en el muestreo de la misma población.

$$H_0 : t_1 = t_2$$

Individuo. Cada uno de los elementos sobre las que se realiza el estudio.

Inferencia estadística. La que trata de obtener propiedades de la población a partir de las propiedades de la muestra. Los resultados que se obtienen se basan en teoremas del Cálculo de Probabilidades, tan exactos como otros teoremas matemáticos, pero hay que tener en cuenta que son probabilísticos con un nivel de significación que se ha de especificar.

Intervalo de confianza. Espacio comprendido entre los límites formados por un parámetro de la muestra y el error correspondiente, dentro del cual debe encontrarse el verdadero valor del parámetro con una probabilidad prefijada.

Intervalo de variación. Véase amplitud.

Logaritmo. Es el número que expresa la potencia a la que hay que elevar la base para obtener un número positivo. Suelen utilizarse los logaritmos comunes o de base 10, y los naturales o neperianos, de base $e = 2,71828$. Si N es un número positivo, puede expresarse como $N = 10^p$; se dice entonces que p es el logaritmo de N en base 10 y se escribe $p = \log_{10} N$ o bien $p = \log N$, y que N es el antilogaritmo de p . Se emplean para simplificar operaciones. Los cálculos comunes con logaritmos son:

$$\log MN = \log M + \log N ; \quad \log \frac{M}{N} = \log M - \log N ; \quad \log M^p = p \log M$$

Media aritmética. La suma de observaciones o de valores dividida por su número.

$$\bar{X} = \frac{\sum X_i}{n}$$

La suma algebraica de las desviaciones del conjunto de valores respecto de su media es cero.

Mediana. Valor central, o la media de los dos centrales, que divide un conjunto de datos ordenados por magnitud en dos partes iguales.

Métodos no paramétricos. Aquellos que se aplican cuando no se conoce la distribución de la población y se realizan hipótesis sobre ésta. Son, pues, métodos independientes de las distribuciones de población y de los parámetros asociados.

Métodos paramétricos. Los que se aplican cuando los datos provienen de una distribución conocida cuyos parámetros se quieren estimar.

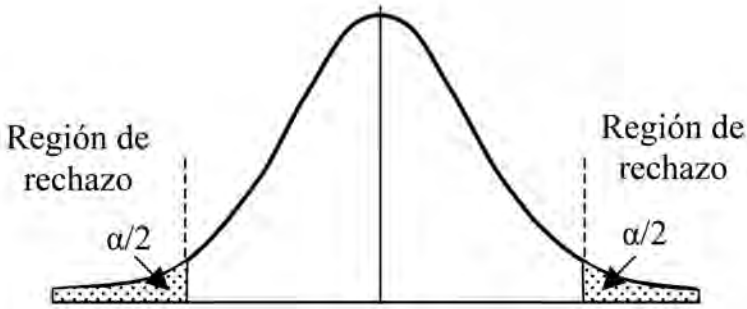
Moda. En un conjunto de datos, el valor más frecuente. Puede no existir o no ser única en caso de existir.

Muestra. Parte de una población sobre la que se realiza el estudio para obtener conclusiones para toda la población.

Nivel de confianza (1- α). Probabilidad de aceptar la hipótesis nula cuando es cierta.



Nivel de significación (α). Probabilidad con la que se está dispuesto a rechazar la hipótesis nula cuando es cierta, es decir, de cometer un error de tipo I. En la práctica, los niveles de significación que suelen emplearse se han fijado de forma arbitraria en el 0,05 (5%), en el 0,01 (1%) y en el 0,001 (1‰), que se representan por asteriscos (*, **, ***, respectivamente).



Parámetro. Una característica de la población, como la media o la desviación típica. Se estima a partir de las muestras y se representa con letras griegas (μ , σ ,...). Véase también **estadístico**.

Percentiles. Valores que dividen el conjunto de datos ordenados por magnitud en 100 partes iguales. Se designan P_1 , P_2 , P_3 , ..., P_{99} . El P_{50} coincide con la mediana. Por ejemplo, $P_{30} = 25$ indica que el 30% de los datos es inferior o igual a 25.

Población. Conjunto sobre el que se realiza un estudio. Puede ser finita o infinita.

Prueba de una y de dos colas. Prueba estadística en la que se utiliza para el contraste uno o los dos extremos de la curva de distribución muestral. Por ejemplo, si se quiere contrastar la hipótesis de que un tratamiento es superior a otro, solo habría que considerar el extremo derecho de la distribución muestral, y se estaría en una prueba de una cola. Por el contrario, si se quiere contrastar si un tratamiento es mejor o peor que otro, habría que considerar ambos extremos, y se trataría de una prueba de dos colas.

Recorrido. Véase **amplitud**.

Valor P. Probabilidad de obtener un resultado como el observado o más extremo, suponiendo que la hipótesis nula es cierta. H_0 se rechaza si $P \leq \alpha$ (el nivel de significación).

Variable cualitativa. La que no es medible numéricamente, como el sexo, el color de los pétalos, etc.

Variable cuantitativa. La que puede medirse numéricamente. Puede ser **discreta**, cuando los posibles valores que puede tomar son aislados, como el número de pétalos de una flor,

el número de semillas producidas, etc., o **continua**, cuando los posibles valores que puede tomar son todos los de un intervalo y, en consecuencia, infinitos, como el peso de la cosecha, el tamaño, etc. Una variable discreta puede considerarse continua para estudios posteriores cuando pueda tomar un gran número de valores, de manera que puedan agruparse los datos en intervalos. Por ejemplo, la edad de un individuo que podría agruparse en intervalos de tres, cuatro o más años.

Variable tipificada. Variable transformada al dividir la desviación respecto a la media por la desviación típica. Es adimensional y viene dada por:

$$z = \frac{X - \bar{X}}{s}$$

Cuando se expresa de esta manera, se dice expresada en unidades estándar y es de gran utilidad para comparar distribuciones. Véase también la distribución normal estándar.

Varianza. El cuadrado de la desviación típica, esto es, la media de los cuadrados de las desviaciones respecto a la media. En una población:

$$\sigma^2 = \frac{\sum (X_i - \mu)^2}{n}$$

APÉNDICE 2

DISTRIBUCIONES MÁS COMUNES

DISTRIBUCIÓN NORMAL

Es aquella que en su forma se aproxima a la de la curva normal, curva en forma de campana definida por la ecuación:

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-(X-\mu)^2 / 2\sigma^2}$$

donde μ es la media, σ la desviación típica, $\pi = 3,1416$ y $e = 2,71828$.

La distribución normal está determinada por dos parámetros, la media (μ) que se localiza en el centro de la distribución, y la desviación típica (σ) que es una medida de dispersión. Su representación gráfica se muestra en la Fig. A1.

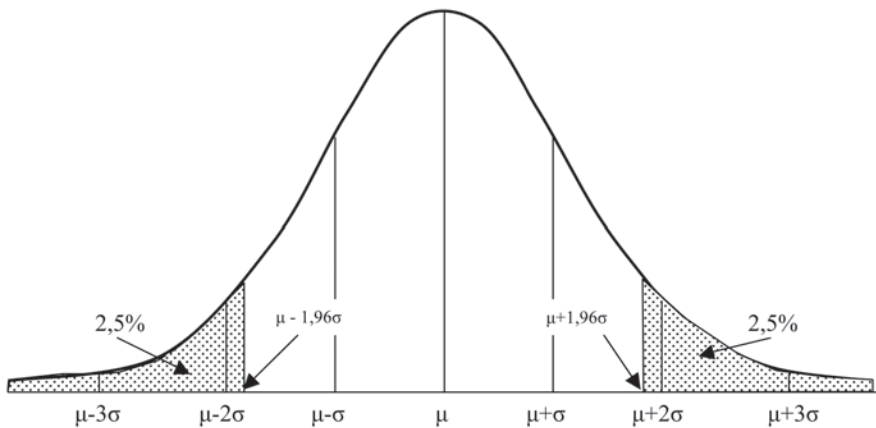


Fig. A1.- Distribución normal con media μ y desviación típica σ . Las áreas sombreadas comprenden el 5% del total.

Las distribuciones normales varían entre sí con relación a la media y a la desviación típica. Las áreas bajo las curvas limitadas por cualquier intervalo de valores dado sobre el eje X comprenden al porcentaje de elementos de la población que caen dentro de ese intervalo. Así, por ejemplo, el intervalo de $\mu \pm 1,96\sigma$ contiene el 95% de los elementos, y el intervalo $\mu \pm 2,58\sigma$ contiene el 99%.

Las áreas bajo las curvas se han tabulado para el caso de $\mu = 0$ y $\sigma = 1$, que se conoce como *distribución normal estándar* y cualquier curva normal puede convertirse en una curva normal estándar cambiando el eje Y en eje de probabilidades ($P = f/n$, siendo f la frecuencia y n el número total de observaciones) y cambiando el eje X en eje de las verdaderas unidades de medida para la desviación típica, denominada z . El valor de z de cualquier valor de X es:

$$z = \frac{X - \mu}{\sigma}$$

El área total bajo la curva es igual a 1 y el área bajo la curva entre dos valores de z es el porcentaje de población que se encuentra en ese intervalo.

La distribución normal constituye una piedra angular de la estadística, por lo que ha sido tabulada con gran precisión y se dispone de tablas para su uso. Resulta de particular interés en Agronomía, pues muchos de los fenómenos biológicos muestran distribuciones de sus datos en forma cercana a la curva normal; pero si no se ajustasen, esto es, si los datos siguen una distribución alejada de la normal, pueden ser transformados para convertirlos en normales. Es más, aún disponiendo de datos no normales, muchos de los resultados son también válidos aplicando la distribución normal. Estas razones justifican el uso extendido de esta distribución.

DISTRIBUCIÓN DE MEDIAS DE MUESTRAS

Si se toman muestras aleatorias repetidas de tamaño n de una población, no necesariamente normal, de media μ y desviación típica σ , la distribución de frecuencias de las medias de las muestras \bar{X} tiene de media $\mu_{\bar{X}}$ y de desviación típica $\sigma_{\bar{X}}$, cuyos valores son:

$$\mu_{\bar{X}} = \mu \quad \text{y} \quad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

La desviación típica de las medias $\sigma_{\bar{X}}$ se conoce como *error estándar de la media* o simplemente *error estándar*.

Para valores grandes de n ($n \geq 30$) la distribución de medias de muestras es aproximadamente normal de media $\mu_{\bar{X}}$ y de desviación típica $\sigma_{\bar{X}}$, independientemente de la población, es decir, cualquiera que sea la forma de la distribución original de X , pues la distribución de \bar{X} como se ha definido antes tiende a hacerse normal conforme aumenta n . Si la población está normalmente distribuida, la distribución de \bar{X} también lo está para $n < 30$. En estas condiciones, se puede convertir una población de medias en unidades estándar de la forma:

$$z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

Los límites de confianza dentro de los cuales estará la media verdadera de la población están definidos por:

$$\mu = \bar{X} \pm z\sigma_{\bar{X}}$$

que para un 95% de confianza resulta:

$$\bar{X} - 1,96 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + 1,96 \frac{\sigma}{\sqrt{n}}$$

y para un 99%:

$$\bar{X} - 2,58 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + 2,58 \frac{\sigma}{\sqrt{n}}$$

DISTRIBUCIÓN *t* de Student

En la distribución de medias de muestras, la conversión de una población de medias en medidas estándar se expresaba por:

$$z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

En el caso de pequeñas muestras (digamos $n < 30$) el uso de la desviación típica de la muestra (s) en lugar de la desviación típica de la población (σ) no resultaba satisfactorio. Se definió el estadístico:

$$t = \frac{\bar{X} - \mu}{s_{\bar{X}}} = \frac{\bar{X} - \mu}{s / \sqrt{n}}$$

conocido como la *t* de Student. $s_{\bar{X}}$ es una estimación de $\sigma_{\bar{X}}$ calculada para cada muestra, luego la variación de *t* depende tanto de \bar{X} como de $s_{\bar{X}}$ en cada muestra, mientras que $\sigma_{\bar{X}}$ es una constante.

t sigue una distribución no normal pero, como ella, tiene forma de campana y es simétrica respecto a la media, pero con un mayor número de valores hacia los extremos de la distribución. Conforme aumenta el tamaño de la muestra, *t* se aproxima al valor *z* normalmente distribuido, y en muestras grandes (digamos $n > 60$) la distribución *t* es prácticamente normal con $\mu = 0$ y $\sigma = 1$. En la investigación agrícola por lo general se emplean muestras de tamaño inferior a 60, por lo que los valores de *z* rara vez se utilizan, pero sí es corriente el uso de la *t*.

Los límites de confianza de μ basados en la distribución t se expresan por:

$$\bar{X} \pm t s_{\bar{X}}$$

que para un 95% de confianza resulta:

$$\bar{X} - t_{0,05} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{0,05} \frac{s}{\sqrt{n}}$$

siendo $t_{0,05}$ el valor en la Tabla A1 para $n-1$ grados de libertad.

DISTRIBUCIÓN F

La distribución F es una distribución de una variable aleatoria continua formada por el estadístico:

$$F = \frac{s_1^2}{s_2^2}$$

donde s_1 y s_2 son dos estimaciones independientes de la varianza de una población normal obtenida a partir de n_1 y n_2 observaciones.

La prueba F es, pues, una razón entre dos varianzas y se utiliza en el análisis de varianza para determinar si dos estimaciones independientes de la varianza pueden ser admitidas como estimaciones de la misma varianza (σ^2). Si es así, F tomará un valor próximo a 1, por lo que el interés es conocer la probabilidad de que F sea mayor que la unidad. Los valores de F están tabulados y se recogen en la Tabla A2.

DISTRIBUCIÓN BINOMIAL

La distribución binomial es una distribución de variables aleatorias discretas que solo admiten dos alternativas: sí o no, representadas por p y q , siendo $q = 1 - p$. Así, la probabilidad de que un suceso ocurra exactamente x veces en n intentos es:

$$y = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

que describe la distribución binomial. En esta distribución,

$$\mu = np \quad \text{y} \quad \sigma = \sqrt{npq}$$

Si en lugar del número de éxitos se está interesado en la proporción de éxitos (x/n), entonces:

$$\mu = p \quad \text{y} \quad \sigma = \sqrt{\frac{pq}{n}}$$

Si n es grande y p y q no están próximos a cero, la distribución binomial se aproxima a la normal, con z :

$$z = \frac{X - np}{\sqrt{npq}}$$

Conforme aumenta n aumenta la aproximación a la normal y en el límite es exacta. En la práctica, la aproximación es buena si np y nq son mayores de 5.

DISTRIBUCIÓN CHI-CUADRADO (χ^2).

La distribución χ^2 es una distribución continua que se define como la suma de los cuadrados de variables aleatorias independientes normalmente distribuidas con medias 0 y varianzas 1. Por lo tanto:

$$\chi^2 = \sum_i z_i^2 = \sum \left(\frac{X_i - \mu_i}{\sigma_i} \right)^2$$

La distribución χ^2 es una distribución asimétrica que se acerca a la normal al aumentar el número de variables independientes, pero más lentamente que la distribución t . Depende del número de desviaciones independientes, es decir, de los grados de libertad, de manera que para cada número de grados de libertad hay una distribución χ^2 . La media y la varianza coinciden con los grados de libertad y con dos veces los grados de libertad, respectivamente. Al ser una suma de cuadrados, la distribución χ^2 no puede ser negativa.

La distribución χ^2 tiene múltiples aplicaciones en estadística, pero se suele asociar más con datos de conteos en los que normalmente interviene una variable discreta. En general, se trata de número de individuos pertenecientes a clases bien definidas, como por ejemplo, machos y hembras; individuos sanos y enfermos; etc. Para estos casos, la prueba χ^2 es:

$$\chi^2 = \sum \frac{(\text{Observados} - \text{Esperados})^2}{\text{Esperados}}$$

tratándose de los resultados observados u obtenidos en el muestreo y de los resultados esperados teóricamente de acuerdo con las leyes de las probabilidades. Ambos resultados

difícilmente coinciden, por lo que la prueba χ^2 trata de averiguar si los valores observados difieren significativamente de los esperados, para lo cual se recurre a contrastar los resultados con los valores de las tablas tabuladas de χ^2 con los grados de libertad apropiados (Tabla A3). El número de grados de libertad no es siempre obvio; por lo general, son uno menos que el número de clases, y si los individuos pueden clasificarse de dos formas distintas, por ejemplo, estado de salud y aplicación o no de un tratamiento, entonces los grados de libertad son el producto de los grados de libertad correspondientes a cada forma de clasificación. Valores altos de χ^2 son significativos.

APÉNDICE 3 TABLAS

Tabla A 1. Distribución de la t de Student (Prueba de dos colas).

Grados de libertad	Probabilidad de un valor más alto de t, sin tener en cuenta el signo		
	0,05	0,01	0,001
1	12,706	63,657	
2	4,303	9,925	31,598
3	3,182	5,841	12,941
4	2,776	4,604	8,610
5	2,571	4,032	6,859
6	2,447	3,707	5,959
7	2,365	3,499	5,405
8	2,306	3,355	5,041
9	2,262	3,250	4,781
10	2,228	3,169	4,587
11	2,201	3,106	4,437
12	2,179	3,055	4,318
13	2,160	3,012	4,221
14	2,145	2,977	4,140
15	2,131	2,947	4,073
16	2,120	2,921	4,015
17	2,110	2,898	3,965
18	2,101	2,878	3,922
19	2,093	2,861	3,883
20	2,086	2,845	3,850
21	2,080	2,831	3,819
22	2,074	2,819	3,792
23	2,069	2,807	3,767
24	2,064	2,797	3,745
25	2,060	2,787	3,725
26	2,056	2,779	3,707
27	2,052	2,771	3,690
28	2,048	2,763	3,674
29	2,045	2,756	3,659

Tabla A 1. (cont.)

Grados de libertad	Probabilidad de un valor más alto de t, sin tener en cuenta el signo		
	0,05	0,01	0,001
30	2,042	2,750	3,646
35	2,030	2,724	3,591
40	2,021	2,704	3,551
45	2,014	2,690	3,520
50	2,008	2,678	3,496
55	2,004	2,669	3,476
60	2,000	2,660	3,460
70	1,994	2,648	3,435
80	1,989	2,638	3,416
90	1,986	2,631	3,402
100	1,982	2,625	3,390
120	1,980	2,617	3,373
∞	1,960	2,5758	3,2905

Fuente: Esta tabla es un resumen de la Tabla A 4 de Snedecor and Cochran (1974).

Tabla A 2. Distribución F

gl del denominador	P	gl del numerador													
		1	2	3	4	5	6	7	8	9	10	14	20	40	∞
1	0,05	161	200	216	225	230	234	237	239	241	242	245	248	251	254
	0,01	4052	4999	5403	5625	5764	5859	5928	5981	6022	6056	6142	6208	6286	6366
2	0,05	18,51	19,00	19,16	19,25	19,30	19,33	19,36	19,37	19,38	19,39	19,42	19,44	19,47	19,50
	0,01	98,49	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,43	99,45	99,48	99,50
3	0,05	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,71	8,66	8,60	8,53
	0,01	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	26,92	26,69	26,41	26,12
4	0,05	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,87	5,80	5,71	5,63
	0,01	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,54	14,24	14,02	13,74	13,46
5	0,05	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,64	4,56	4,46	4,36
	0,01	16,26	13,27	12,06	11,39	10,97	10,67	10,45	10,29	10,15	10,05	9,77	9,55	9,29	9,02
6	0,05	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,96	3,87	3,77	3,67
	0,01	13,74	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,60	7,39	7,14	6,88
7	0,05	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,52	3,44	3,34	3,23
	0,01	12,25	9,55	8,45	7,85	7,46	7,19	7,00	6,84	6,71	6,62	6,35	6,15	5,90	5,65
8	0,05	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,23	3,15	3,05	2,93
	0,01	11,26	8,65	7,59	7,01	6,63	6,37	6,19	6,03	5,91	5,82	5,56	5,36	5,11	4,86
9	0,05	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,02	2,93	2,82	2,71
	0,01	10,56	8,02	6,99	6,42	6,06	5,80	5,62	5,47	5,35	5,26	5,00	4,80	4,56	4,31
10	0,05	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,86	2,77	2,67	2,54
	0,01	10,04	7,56	6,55	5,99	5,64	5,39	5,21	5,06	4,95	4,85	4,60	4,41	4,17	3,91
11	0,05	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,74	2,65	2,53	2,40
	0,01	9,65	7,20	6,22	5,67	5,32	5,07	4,88	4,74	4,63	4,54	4,29	4,10	3,86	3,60
12	0,05	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,64	2,54	2,42	2,30
	0,01	9,33	6,93	5,95	5,41	5,06	4,82	4,65	4,50	4,39	4,30	4,05	3,86	3,61	3,36
13	0,05	4,67	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,55	2,46	2,34	2,21
	0,01	9,07	6,70	5,74	5,20	4,86	4,62	4,44	4,30	4,19	4,10	3,85	3,67	3,42	3,16
14	0,05	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,48	2,39	2,27	2,13
	0,01	8,86	6,51	5,56	5,03	4,69	4,46	4,28	4,14	4,03	3,94	3,70	3,51	3,26	3,00
15	0,05	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,43	2,33	2,21	2,07
	0,01	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,56	3,36	3,12	2,87
16	0,05	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,37	2,28	2,16	2,01
	0,01	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,45	3,25	3,01	2,75
17	0,05	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,33	2,23	2,11	1,96
	0,01	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,35	3,16	2,92	2,65
18	0,05	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,29	2,19	2,07	1,92
	0,01	8,28	6,01	5,09	4,58	4,25	4,01	3,85	3,71	3,60	3,51	3,27	3,07	2,83	2,57
19	0,05	4,38	3,52	3,13	2,90	2,74	2,63	2,55	2,48	2,43	2,38	2,26	2,15	2,02	1,88
	0,01	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,19	3,00	2,76	2,49
20	0,05	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40	2,35	2,23	2,12	1,99	1,84
	0,01	8,10	5,85	4,94	4,43	4,10	3,87	3,71	3,56	3,45	3,37	3,13	2,94	2,69	2,42
21	0,05	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,20	2,09	1,96	1,81
	0,01	8,02	5,78	4,87	4,37	4,04	3,81	3,65	3,51	3,40	3,31	3,07	2,88	2,63	2,36
22	0,05	4,30	3,44	3,05	2,82	2,66	2,55	2,47	2,40	2,35	2,30	2,18	2,07	1,93	1,78
	0,01	7,94	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,02	2,83	2,58	2,31
23	0,05	4,28	3,42	3,03	2,80	2,64	2,53	2,45	2,38	2,32	2,28	2,14	2,04	1,91	1,76
	0,01	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	2,97	2,78	2,53	2,26
24	0,05	4,26	3,40	3,01	2,78	2,62	2,51	2,43	2,36	2,30	2,26	2,13	2,02	1,89	1,73
	0,01	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,25	3,17	2,93	2,74	2,49	2,21

Tabla A 2. (cont.)

gl del denominador	P	gl del numerador														
		1	2	3	4	5	6	7	8	9	10	14	20	40	∞	
25	0,05	4,24	3,38	2,99	2,76	2,60	2,49	2,41	2,34	2,28	2,24	2,11	2,00	1,87	1,71	
	0,01	7,77	5,57	4,68	4,18	3,86	3,63	3,46	3,32	3,21	3,13	2,89	2,70	2,45	2,17	
26	0,05	4,22	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,10	1,99	1,85	1,69	
	0,01	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,17	3,09	2,86	2,66	2,41	2,13	
27	0,05	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,30	2,25	2,20	2,08	1,97	1,84	1,67	
	0,01	7,68	5,49	4,60	4,11	3,79	3,56	3,39	3,26	3,14	3,06	2,83	2,63	2,38	2,10	
28	0,05	4,20	3,34	2,95	2,71	2,56	2,44	2,36	2,29	2,24	2,19	2,06	1,96	1,81	1,65	
	0,01	7,64	5,45	4,57	4,07	3,76	3,53	3,36	3,23	3,11	3,03	2,80	2,60	2,35	2,06	
29	0,05	4,18	3,33	2,93	2,70	2,54	2,43	2,35	2,28	2,22	2,18	2,05	1,94	1,80	1,64	
	0,01	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,08	3,00	2,77	2,57	2,32	2,03	
30	0,05	4,17	3,32	2,92	2,69	2,53	2,42	2,34	2,27	2,21	2,16	2,04	1,93	1,79	1,62	
	0,01	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,06	2,98	2,74	2,55	2,29	2,01	
32	0,05	4,15	3,30	2,90	2,67	2,51	2,40	2,32	2,25	2,19	2,14	2,02	1,91	1,76	1,59	
	0,01	7,50	5,34	4,46	3,97	3,66	3,42	3,25	3,12	3,01	2,94	2,70	2,51	2,25	1,96	
34	0,05	4,13	3,28	2,88	2,65	2,49	2,38	2,30	2,23	2,17	2,12	2,00	1,89	1,74	1,57	
	0,01	7,44	5,29	4,42	3,93	3,61	3,38	3,21	3,08	2,97	2,89	2,66	2,47	2,21	1,91	
36	0,05	4,11	3,26	2,86	2,63	2,48	2,36	2,28	2,21	2,15	2,10	1,98	1,87	1,72	1,55	
	0,01	7,39	5,25	4,38	3,89	3,58	3,35	3,18	3,04	2,94	2,86	2,62	2,43	2,17	1,87	
38	0,05	4,10	3,25	2,85	2,62	2,46	2,35	2,26	2,19	2,14	2,09	1,96	1,85	1,71	1,53	
	0,01	7,35	5,21	4,34	3,86	3,54	3,32	3,15	3,02	2,91	2,82	2,59	2,40	2,14	1,84	
40	0,05	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,07	1,95	1,84	1,69	1,51	
	0,01	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,88	2,80	2,56	2,37	2,11	1,81	
42	0,05	4,07	3,22	2,83	2,59	2,44	2,32	2,24	2,17	2,11	2,06	1,94	1,82	1,68	1,49	
	0,01	7,27	5,15	4,29	3,80	3,49	3,26	3,10	2,96	2,86	2,77	2,54	2,35	2,08	1,78	
44	0,05	4,06	3,21	2,82	2,58	2,43	2,31	2,23	2,16	2,10	2,05	1,92	1,81	1,66	1,48	
	0,01	7,24	5,12	4,26	3,78	3,46	3,24	3,07	2,94	2,84	2,75	2,52	2,32	2,06	1,75	
46	0,05	4,05	3,20	2,81	2,57	2,42	2,30	2,22	2,14	2,09	2,04	1,91	1,80	1,65	1,46	
	0,01	7,21	5,10	4,24	3,76	3,44	3,22	3,05	2,92	2,82	2,73	2,50	2,30	2,04	1,72	
48	0,05	4,04	3,19	2,80	2,56	2,41	2,30	2,21	2,14	2,08	2,03	1,90	1,79	1,64	1,45	
	0,01	7,19	5,08	4,22	3,74	3,42	3,20	3,04	2,90	2,80	2,71	2,48	2,28	2,02	1,70	
50	0,05	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,02	1,90	1,78	1,63	1,44	
	0,01	7,17	5,06	4,20	3,72	3,41	3,18	3,02	2,88	2,78	2,70	2,46	2,26	2,00	1,68	
60	0,05	4,00	3,15	2,76	2,52	2,37	2,25	2,17	2,10	2,04	1,99	1,86	1,75	1,59	1,39	
	0,01	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,40	2,20	1,93	1,60	
70	0,05	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,01	1,97	1,84	1,72	1,56	1,35	
	0,01	7,01	4,92	4,08	3,60	3,29	3,07	2,91	2,77	2,67	2,59	2,35	2,15	1,88	1,53	
80	0,05	3,96	3,11	2,72	2,48	2,33	2,21	2,12	2,05	1,99	1,95	1,82	1,70	1,54	1,32	
	0,01	6,96	4,88	4,04	3,56	3,25	3,04	2,87	2,74	2,64	2,55	2,32	2,11	1,84	1,49	
100	0,05	3,94	3,09	2,70	2,46	2,30	2,19	2,10	2,03	1,97	1,92	1,79	1,68	1,51	1,28	
	0,01	6,90	4,82	3,98	3,51	3,20	2,99	2,82	2,69	2,59	2,51	2,26	2,06	1,79	1,43	
150	0,05	3,91	3,06	2,67	2,43	2,27	2,16	2,07	2,00	1,94	1,89	1,76	1,64	1,47	1,22	
	0,01	6,81	4,75	3,91	3,44	3,14	2,92	2,76	2,62	2,53	2,44	2,20	2,00	1,72	1,33	
∞	0,05	3,84	2,99	2,60	2,37	2,21	2,09	2,01	1,94	1,88	1,83	1,69	1,57	1,40	1,00	
	0,01	6,64	4,60	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,07	1,87	1,59	1,00	

Fuente: Esta tabla es un resumen de la Tabla A 14 de Snedecor and Cochran (1974).

Tabla A 3. Distribución de χ^2 (Chi-cuadrado)

Grados de libertad	Probabilidad de obtener un valor igual o mayor		
	0,05	0,01	0,001
1	3,84	6,63	10,83
2	5,99	9,21	13,81
3	7,81	11,34	16,27
4	9,49	13,28	18,46
5	11,07	15,09	20,52
6	12,59	16,81	22,46
7	14,07	18,48	24,32
8	15,51	20,09	26,12
9	16,92	21,67	27,88
10	18,31	23,21	29,59
11	19,68	24,72	31,26
12	21,03	26,22	32,91
13	22,36	27,69	34,53
14	23,68	29,14	36,12
15	25,00	30,58	37,70
16	26,30	32,00	39,25
17	27,59	33,41	40,79
18	28,87	34,81	42,31
19	30,14	36,19	43,82
20	31,41	37,57	45,31
21	32,67	38,93	46,80
22	33,92	40,29	48,27
23	35,17	41,64	49,73
24	36,42	42,98	51,18
25	37,65	44,31	52,62
26	38,89	45,64	54,05
27	40,11	46,96	55,48
28	41,34	48,28	56,89
29	42,56	49,59	58,30
30	43,77	50,89	59,70
40	55,76	63,69	
50	67,50	76,15	
60	79,08	88,38	
70	90,53	100,42	
80	101,88	112,33	
90	113,14	124,12	
100	124,34	135,81	

Fuente: Esta tabla ha sido elaborada a partir de la Tabla A 5 de Snedecor and Cochran (1974) y de la Tabla A 6 de Little and Hills (1991).

Tabla A 4. Coeficientes de correlación r a diferentes niveles de significación

Grados de libertad	Probabilidad de obtener un valor igual o mayor de r		
	0,05	0,01	0,001
1	0,997	1,000	1,000
2	0,950	0,990	0,999
3	0,878	0,959	0,991
4	0,811	0,917	0,974
5	0,754	0,874	0,951
6	0,707	0,834	0,925
7	0,666	0,798	0,898
8	0,632	0,765	0,872
9	0,602	0,735	0,847
10	0,576	0,708	0,823
11	0,553	0,684	0,801
12	0,532	0,661	0,780
13	0,514	0,641	0,760
14	0,497	0,623	0,742
15	0,482	0,606	0,725
16	0,468	0,590	0,708
17	0,456	0,575	0,693
18	0,444	0,561	0,679
19	0,433	0,549	0,665
20	0,423	0,537	0,652
21	0,413	0,526	
22	0,404	0,515	
23	0,396	0,505	
24	0,388	0,496	
25	0,381	0,487	0,597
26	0,374	0,478	
27	0,367	0,470	
28	0,361	0,463	
29	0,355	0,456	
30	0,349	0,449	0,554
35	0,325	0,418	0,519
40	0,304	0,393	0,490
45	0,288	0,372	0,468
50	0,273	0,354	0,443

Tabla A 4. (cont.)

Grados de libertad	Probabilidad de obtener un valor igual o mayor de r		
	0,05	0,01	0,001
60	0,250	0,325	0,408
70	0,232	0,302	0,380
80	0,217	0,283	0,357
90	0,205	0,267	0,337
100	0,195	0,254	0,321
125	0,174	0,228	
150	0,159	0,208	
200	0,138	0,181	
300	0,113	0,148	
400	0,098	0,128	
500	0,088	0,115	

Fuente: Esta tabla ha sido elaborada a partir de la Tabla A 11 de Snedecor and Cochran (1974) y de la Tabla A 7 de Little and Hills (1991).

Tabla A 5. Número de signos iguales requeridos para obtener significación en la prueba de los signos (Prueba de dos colas).

Número de pares	Probabilidad	
	0.05	0,01
≤ 8	0	0
9	1	0
10	1	0
11	1	0
12	2	1
13	2	1
14	2	1
15	3	2
16	3	2
17	4	2
18	4	3
19	4	3
20	5	3

Fuente: Esta tabla es un resumen de la Tabla A 8 de Snedecor and Cochran (1974).

Tabla A 6. Prueba de categorías con signos de Wilcoxon (Prueba de dos colas)¹.

Número de pares	Probabilidad	
	0,05	0,01
6	0	–
7	2	–
8	4	0
9	6	2
10	8	3
11	11	5
12	14	7
13	17	10
14	21	13
15	25	16
16	30	20
17	35	23
18	40	28
19	46	32
20	52	38
21	59	43
22	66	49
23	73	55
24	81	61
25	89	68

¹Valores iguales o menores indican rechazo.

Fuente: Esta tabla es un resumen de la Tabla A 18 de Steel and Torrie (1985), que la tomaron, a su vez, de Wilcoxon (1949).

Tabla A 7. Prueba de la suma de categorías de Wilcoxon (Prueba de Mann-Whitney) Prueba de dos colas^{1, 2}.

n ₂	P	n ₁													
		2	3	4	5	6	7	8	9	10	11	12	13	14	15
4	0,05			10											
	0,01			-											
5	0,05		6	11	17										
	0,01		-	-	15										
6	0,05		7	12	18	26									
	0,01		-	10	16	23									
7	0,05		7	13	20	27	36								
	0,01		-	10	17	24	32								
8	0,05	3	8	14	21	29	38	49							
	0,01	-	-	11	17	25	34	43							
9	0,05	3	8	15	22	31	40	51	63						
	0,01	-	6	11	18	26	35	45	56						
10	0,05	3	9	15	23	32	42	53	65	78					
	0,01	-	6	12	19	27	37	47	58	71					
11	0,05	4	9	16	24	34	44	55	68	81	96				
	0,01	-	6	12	20	28	38	49	61	74	87				
12	0,05	4	10	17	26	35	46	58	71	85	99	115			
	0,01	-	7	13	21	30	40	51	63	76	90	106			
13	0,05	4	10	18	27	37	48	60	73	88	103	119	137		
	0,01	-	7	14	22	31	41	53	65	79	93	109	125		
14	0,05	4	11	19	28	38	50	63	76	91	106	123	141	160	
	0,01	-	7	14	22	32	43	54	67	81	96	112	129	147	
15	0,05	4	11	20	29	40	52	65	79	94	110	127	145	164	185
	0,01	-	8	15	23	33	44	56	70	84	99	115	133	151	171
16	0,05	4	12	21	31	42	54	67	82	97	114	131	150	169	
	0,01	-	8	15	24	34	46	58	72	86	102	119	137	155	

Tabla A 7. (cont.)

n ₂	P	n ₁													
		2	3	4	5	6	7	8	9	10	11	12	13	14	15
17	0,05	5	12	21	32	43	56	70	84	100	117	135	154		
	0,01	–	8	16	25	36	47	60	74	89	105	122	140		
18	0,05	5	13	22	33	45	58	72	87	103	121	139			
	0,01	–	8	16	26	37	49	62	76	92	108	125			
19	0,05	5	13	23	34	46	60	74	90	107	124				
	0,01	3	9	17	27	38	50	64	78	94	111				
20	0,05	5	14	24	35	48	62	77	93	110					
	0,01	3	9	18	28	39	52	66	81	97					
21	0,05	6	14	25	37	50	64	79	95						
	0,01	3	9	18	29	40	53	68	83						
22	0,05	6	15	26	38	51	66	82							
	0,01	3	10	19	29	42	55	70							
23	0,05	6	15	27	39	53	68								
	0,01	3	10	19	30	43	57								
24	0,05	6	16	28	40	55									
	0,01	3	10	20	31	44									
25	0,05	6	16	28	42										
	0,01	3	11	20	32										
26	0,05	7	17	29											
	0,01	3	11	21											
27	0,05	7	17												
	0,01	4	11												
28	0,05	7													
	0,01	4													

¹Valores iguales o menores indican rechazo.

² $n_1 \leq n_2$, es decir, si los grupos son de distinto tamaño n_1 es el más pequeño.

Fuente: Esta tabla ha sido reproducida de la Tabla A 19 de Steel and Torrie (1985), que la tomaron, a su vez, de White (1950)

Tabla A 8. Coeficientes de correlación de Spearman r_s a diferentes niveles de significación

Tamaño de muestra	Probabilidad de obtener un valor igual o mayor de r_s	
	0,05	0,01
≤ 4	–	–
5	1,000	–
6	0,886	1,000
7	0,750	0,893
8	0,714	0,857
9	0,683	0,833
10	0,648	0,794
≥ 11	Utilícese la Tabla A 4.	

Fuente: Esta tabla es una reproducción de la Tabla 7.11.2 de Snedecor and Cochran (1974).

Tabla A 9. Transformación $\arcsen \sqrt{\text{Porcentaje}/100}$ expresada en grados¹.

%	Ángulo	%	Ángulo	%	Ángulo
0	0	34	35,67	68	55,55
1	5,74	35	36,27	69	56,17
2	8,13	36	36,87	70	56,79
3	9,97	37	37,46	71	57,42
4	11,54	38	38,06	72	58,05
5	12,92	39	38,64	73	58,69
6	14,18	40	39,23	74	59,34
7	15,34	41	39,81	75	60,00
8	16,43	42	40,40	76	60,67
9	17,46	43	40,98	77	61,34
10	18,43	44	41,55	78	62,03
11	19,37	45	42,13	79	62,73
12	20,27	46	42,71	80	63,43
13	21,13	47	43,28	81	64,16
14	21,97	48	43,85	82	64,90
15	22,79	49	44,43	83	65,65
16	23,58	50	45,00	84	66,42
17	24,35	51	45,57	85	67,21
18	25,10	52	46,15	86	68,03
19	25,84	53	46,72	87	68,87
20	26,56	54	47,29	88	69,73
21	27,27	55	47,87	89	70,63
22	27,97	56	48,45	90	71,56
23	28,66	57	49,02	91	72,54
24	29,33	58	49,60	92	73,57
25	30,00	59	50,18	93	74,66
26	30,66	60	50,77	94	75,82
27	31,31	61	51,35	95	77,08
28	31,95	62	51,94	96	78,46
29	32,58	63	52,53	97	80,02
30	33,21	64	53,13	98	81,87
31	33,83	65	53,73	99	84,26
32	34,45	66	54,33	100	90
33	35,06	67	54,94		

¹Para realizar la transformación en el *Programa SX*, seguir la siguiente secuencia:

Data>Transformations

Escribir la siguiente expresión, siendo p el porcentaje:

$$\text{Ángulo} = \text{Arcsin}((p/100)^{0.5}) * 360 / (2 * \text{Pi})$$

pinchar *Go* y aparecerán los valores transformados y expresados en grados.

Tabla A 10. Coeficientes a_{n-i+1} para el test de Shapiro-Wilk

$i \backslash n$	2	3	4	5	6	7	8	9	10	
1	0,7071	0,7071	0,6872	0,6646	0,6431	0,6233	0,6052	0,5888	0,5739	
2		0,0000	0,1667	0,2413	0,2806	0,3031	0,3164	0,3244	0,3291	
3			0,0000	0,0875	0,1401	0,1743	0,1976	0,2141		
4						0,0000	0,0561	0,0947	0,1224	
5								0,0000	0,0399	
$i \backslash n$	11	12	13	14	15	16	17	18	19	20
1	0,5601	0,5475	0,5359	0,5251	0,5150	0,5056	0,4968	0,4886	0,4808	0,4734
2	0,3315	0,3325	0,3325	0,3318	0,3306	0,3290	0,3273	0,3253	0,3232	0,3211
3	0,2260	0,2347	0,2412	0,2460	0,2495	0,2521	0,2540	0,2553	0,2561	0,2565
4	0,1429	0,1586	0,1707	0,1802	0,1878	0,1939	0,1988	0,2027	0,2059	0,2085
5	0,0695	0,0922	0,1099	0,1240	0,1353	0,1447	0,1524	0,1587	0,1641	0,1686
6	0,0000	0,0303	0,0539	0,0727	0,0880	0,1005	0,1109	0,1197	0,1271	0,1334
7			0,0000	0,0240	0,0433	0,0593	0,0725	0,0837	0,0932	0,1033
8					0,0000	0,0196	0,0359	0,0496	0,0612	0,0711
9							0,0000	0,0163	0,0303	0,0422
10									0,0000	0,0144
$i \backslash n$	21	22	23	24	25	26	27	28	29	30
1	0,4643	0,4590	0,4542	0,4493	0,4450	0,4407	0,4366	0,4328	0,4291	0,4254
2	0,3185	0,3156	0,3126	0,3098	0,3069	0,3043	0,3018	0,2992	0,2968	0,2944
3	0,2578	0,2571	0,2563	0,2554	0,2543	0,2533	0,2522	0,2510	0,2499	0,2487
4	0,2119	0,2131	0,2139	0,2145	0,2148	0,2151	0,2152	0,2151	0,2150	0,2148
5	0,1736	0,1764	0,1787	0,1807	0,1822	0,1836	0,1848	0,1857	0,1864	0,1870
6	0,1399	0,1443	0,1480	0,1512	0,1539	0,1563	0,1584	0,1601	0,1616	0,1630
7	0,1092	0,1150	0,1201	0,1245	0,1283	0,1316	0,1346	0,1372	0,1395	0,1415
8	0,0804	0,0878	0,0941	0,0997	0,1046	0,1089	0,1128	0,1162	0,1192	0,1219
9	0,0530	0,0618	0,0696	0,0764	0,0823	0,0876	0,0923	0,0965	0,1002	0,1036
10	0,0263	0,0368	0,0459	0,0539	0,0610	0,0672	0,0728	0,0779	0,0822	0,0862
11	0,0000	0,0122	0,0228	0,0321	0,0403	0,0476	0,0540	0,0598	0,0650	0,0697
12			0,0000	0,0107	0,0200	0,0284	0,0358	0,0424	0,0483	0,0537
13					0,0000	0,0094	0,0178	0,0253	0,0320	0,0381
14							0,0000	0,0084	0,0159	0,0227
15									0,0000	0,0076

Tabla A10. (cont.)

i \ n	31	32	33	34	35	36	37	38	39	40
1	0,4220	0,4188	0,4156	0,4127	0,4096	0,4068	0,4040	0,4015	0,3989	0,3964
2	0,2921	0,2898	0,2876	0,2854	0,2834	0,2813	0,2794	0,2774	0,2755	0,2737
3	0,2475	0,2462	0,2451	0,2439	0,2427	0,2415	0,2403	0,2391	0,2380	0,2368
4	0,2145	0,2141	0,2137	0,2132	0,2127	0,2111	0,2116	0,2110	0,2104	0,2098
5	0,1874	0,1878	0,1880	0,1882	0,1883	0,1883	0,1883	0,1881	0,1880	0,1878
6	0,1641	0,1651	0,1660	0,1667	0,1673	0,1678	0,1683	0,1686	0,1689	0,1691
7	0,1433	0,1449	0,1463	0,1475	0,1487	0,1496	0,1505	0,1513	0,1520	0,1526
8	0,1243	0,1265	0,1294	0,1301	0,1317	0,1331	0,1344	0,1356	0,1366	0,1376
9	0,1066	0,1093	0,1118	0,1140	0,1160	0,1179	0,1196	0,1211	0,1225	0,1237
10	0,0899	0,0931	0,0961	0,0988	0,1013	0,1036	0,1056	0,1075	0,1092	0,1108
11	0,0739	0,0777	0,0812	0,0844	0,0873	0,0900	0,0924	0,0947	0,0967	0,0986
12	0,0585	0,0629	0,0669	0,0706	0,0739	0,0770	0,0798	0,0824	0,0848	0,0870
13	0,0435	0,0485	0,0530	0,0572	0,0610	0,0645	0,0677	0,0706	0,0733	0,0759
14	0,0289	0,0349	0,0395	0,0441	0,0484	0,0523	0,0559	0,0592	0,0622	0,0611
15	0,0144	0,0206	0,0262	0,0314	0,0361	0,0404	0,0444	0,0481	0,0515	0,0546
16	0,0000	0,0068	0,0131	0,0187	0,0239	0,0287	0,0331	0,0372	0,0409	0,0444
17			0,0000	0,0062	0,0119	0,0172	0,0220	0,0264	0,0305	0,0343
18					0,0000	0,0057	0,0110	0,0158	0,0203	0,0244
19							0,0000	0,0053	0,0101	0,0146
20									0,0000	0,0049
i \ n	41	42	43	44	45	46	47	48	49	50
1	0,3940	0,3917	0,3894	0,3872	0,3850	0,3830	0,3808	0,3789	0,3770	0,3751
2	0,2719	0,2701	0,2684	0,2667	0,2651	0,2635	0,2620	0,2604	0,2589	0,2574
3	0,2357	0,2345	0,2334	0,2323	0,2313	0,2302	0,2291	0,2281	0,2271	0,2260
4	0,2091	0,2085	0,2078	0,2072	0,2065	0,2058	0,2052	0,2045	0,2038	0,2032
5	0,1876	0,1874	0,1871	0,1868	0,1865	0,1862	0,1859	0,1855	0,1851	0,1847
6	0,1693	0,1694	0,1695	0,1695	0,1695	0,1695	0,1695	0,1693	0,1692	0,1691
7	0,1531	0,1535	0,1539	0,1542	0,1545	0,1548	0,1550	0,1551	0,1553	0,1554
8	0,1384	0,1392	0,1398	0,1405	0,1410	0,1415	0,1420	0,1423	0,1427	0,1430
9	0,1249	0,1259	0,1269	0,1278	0,1286	0,1293	0,1300	0,1306	0,1312	0,1317
10	0,1123	0,1136	0,1149	0,1160	0,1170	0,1180	0,1189	0,1197	0,1205	0,1212
11	0,1004	0,1020	0,1035	0,1049	0,1062	0,1073	0,1085	0,1095	0,1105	0,1113
12	0,0891	0,0909	0,0927	0,0943	0,0959	0,0972	0,0986	0,0998	0,1010	0,1020
13	0,0782	0,0804	0,0824	0,0842	0,0860	0,0876	0,0892	0,0906	0,0919	0,0932
14	0,0677	0,0701	0,0724	0,0745	0,0765	0,0783	0,0801	0,0817	0,0832	0,0846
15	0,0575	0,0602	0,0628	0,0651	0,0673	0,0694	0,0713	0,0731	0,0748	0,0764
16	0,0476	0,0506	0,0534	0,0560	0,0584	0,0607	0,0628	0,0648	0,0667	0,0685
17	0,0379	0,0411	0,0442	0,0471	0,0497	0,0522	0,0546	0,0568	0,0588	0,0608
18	0,0283	0,0318	0,0352	0,0383	0,0412	0,0439	0,0465	0,0489	0,0511	0,0532
19	0,0188	0,0227	0,0263	0,0296	0,0328	0,0357	0,0385	0,0411	0,0436	0,0459
20	0,0094	0,0136	0,0175	0,0211	0,0245	0,0277	0,0307	0,0335	0,0361	0,0386
21	0,0000	0,0045	0,0087	0,0126	0,0163	0,0197	0,0229	0,0259	0,0288	0,0314
22			0,0000	0,0042	0,0081	0,0118	0,0153	0,0185	0,0215	0,0244
23					0,0000	0,0039	0,0076	0,0111	0,0143	0,0174
24							0,0000	0,0037	0,0071	0,0104
25									0,0000	0,0035

La primera columna (i) dentro de cada bloque indica el orden de las diferencias entre dos medias dispuestas de mayor a menor valor. La primera fila de cada bloque (n) indica el tamaño de la muestra.

Tabla A 11. Valores críticos (W) del test de Shapiro-Wilk.

n	Probabilidad	
	0,05	0,01
3	0,767	0,753
4	0,748	0,687
5	0,762	0,686
6	0,788	0,713
7	0,803	0,730
8	0,818	0,749
9	0,829	0,764
10	0,842	0,781
11	0,850	0,792
12	0,859	0,805
13	0,866	0,814
14	0,874	0,825
15	0,881	0,835
16	0,887	0,844
17	0,892	0,851
18	0,897	0,858
19	0,901	0,863
20	0,905	0,868
21	0,908	0,873
22	0,911	0,878
23	0,914	0,881
24	0,916	0,884
25	0,918	0,888
26	0,920	0,891
27	0,923	0,894
28	0,924	0,896
29	0,926	0,898
30	0,927	0,900

Tabla A11. (cont.)

n	Probabilidad	
	0,05	0,01
31	0,929	0,902
32	0,930	0,904
33	0,931	0,906
34	0,933	0,908
35	0,934	0,910
36	0,935	0,912
37	0,936	0,914
38	0,938	0,916
39	0,939	0,917
40	0,940	0,919
41	0,941	0,920
42	0,942	0,922
43	0,943	0,923
44	0,944	0,924
45	0,945	0,926
46	0,945	0,927
47	0,946	0,928
48	0,947	0,929
49	0,947	0,929
50	0,947	0,930

Tabla A 12a. Valores críticos para el test de rachas (R_1)

n_1	n_2																			
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
2											2	2	2	2	2	2	2	2	2	2
3					2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3
4				2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4
5			2	2	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5	5
6		2	2	3	3	3	3	4	4	4	4	5	5	5	5	5	5	6	6	6
7		2	2	3	3	3	4	4	5	5	5	5	5	6	6	6	6	6	6	6
8		2	3	3	3	4	4	5	5	5	6	6	6	6	6	7	7	7	7	7
9		2	3	3	4	4	5	5	5	6	6	6	7	7	7	7	8	8	8	8
10		2	3	3	4	5	5	5	6	6	7	7	7	7	8	8	8	8	9	9
11		2	3	4	4	5	5	6	6	7	7	7	8	8	8	9	9	9	9	9
12	2	2	3	4	4	5	6	6	7	7	7	8	8	8	9	9	9	10	10	10
13	2	2	3	4	5	5	6	6	7	7	8	8	9	9	9	10	10	10	10	10
14	2	2	3	4	5	5	6	7	7	8	8	9	9	9	10	10	10	11	11	11
15	2	3	3	4	5	6	6	7	7	8	8	9	9	10	10	11	11	11	12	12
16	2	3	4	4	5	6	6	7	8	8	9	9	10	10	11	11	11	12	12	12
17	2	3	4	4	5	6	7	7	8	9	9	10	10	11	11	11	12	12	13	13
18	2	3	4	5	5	6	7	8	8	9	9	10	10	11	11	12	12	13	13	13
19	2	3	4	5	6	6	7	8	8	9	10	10	11	11	12	12	13	13	13	13
20	2	3	4	5	6	6	7	8	9	9	10	10	11	12	12	13	13	13	14	14

Fuente: Esta tabla ha sido adaptada de Sweed, Frieda, S. and Eisenhart, C. 1943. Tables for testing randomness of grouping in a sequence of alternatives. *Ann. Math. Statist.*, 14:83-86.

Tabla A 12b. Valores críticos para el test de rachas (R_2)

n_1	n_2																			
	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20		
2																				
3																				
4			9	9																
5		9	10	10	11	11														
6		9	10	11	12	12	13	13	13	13										
7			11	12	13	13	14	14	14	14	15	15	15							
8			11	12	13	14	14	15	15	16	16	16	16	17	17	17	17	17	17	
9				13	14	14	15	16	16	16	17	17	18	18	18	18	18	18	18	
10				13	14	15	16	16	17	17	18	18	18	19	19	19	20	20	20	
11				13	14	15	16	17	17	18	19	19	19	20	20	20	21	21	21	
12				13	14	16	16	17	18	19	19	20	20	21	21	21	22	22	22	
13					15	16	17	18	19	19	20	20	21	21	22	22	23	23	23	
14					15	16	17	18	19	20	20	21	22	22	23	23	23	24	24	
15					15	16	18	18	19	20	21	22	22	23	23	24	24	25	25	
16						17	18	19	20	21	21	22	23	23	24	25	25	25	25	
17						17	18	19	20	21	22	23	23	24	25	25	26	26	26	
18						17	18	19	20	21	22	23	24	25	25	26	26	27	27	
19						17	18	20	21	22	23	23	24	25	26	26	27	27	27	
20						17	18	20	21	22	23	24	25	25	26	27	27	28	28	

Fuente: Esta tabla ha sido adaptada de Sweed, Frieda, S. and Eisenhart, C. 1943. Tables for testing randomness of grouping in a sequence of alternatives. *Ann. Math. Statist.*, 14:83-86.

Tabla A 13. Valores críticos (d) del test de Dunnett

gl del error	P	Número de grupos, incluido el control								
		2	3	4	5	6	7	8	9	10
5	0,05	2,57	3,03	3,29	3,48	3,62	3,73	3,82	3,90	3,97
	0,01	4,03	4,63	4,98	5,22	5,41	5,56	5,69	5,80	5,89
6	0,05	2,45	2,86	3,10	3,26	3,39	3,49	3,57	3,64	3,71
	0,01	3,71	4,21	4,51	4,71	4,87	5,00	5,10	5,20	5,28
7	0,05	2,36	2,75	2,97	3,12	3,24	3,33	3,41	3,47	3,53
	0,01	3,50	3,95	4,21	4,39	4,53	4,64	4,74	4,82	4,89
8	0,05	2,31	2,67	2,88	3,02	3,13	3,22	3,29	3,35	3,41
	0,01	3,36	3,77	4,00	4,17	4,29	4,40	4,48	4,56	4,62
9	0,05	2,26	2,61	2,81	2,95	3,05	3,14	3,20	3,26	3,32
	0,01	3,25	3,63	3,85	4,01	4,12	4,22	4,30	4,37	4,43
10	0,05	2,23	2,57	2,76	2,89	2,99	3,07	3,14	3,19	3,24
	0,01	3,17	3,53	3,74	3,88	3,99	4,08	4,16	4,22	4,28
11	0,05	2,20	2,53	2,72	2,84	2,94	3,02	3,08	3,14	3,19
	0,01	3,11	3,45	3,65	3,79	3,89	3,98	4,05	4,11	4,16
12	0,05	2,18	2,50	2,68	2,81	2,90	2,98	3,04	3,09	3,14
	0,01	3,05	3,39	3,58	3,71	3,81	3,89	3,96	4,02	4,07
13	0,05	2,16	2,48	2,65	2,78	2,87	2,94	3,00	3,06	3,10
	0,01	3,01	3,33	3,52	3,65	3,74	3,82	3,89	3,94	3,99
14	0,05	2,14	2,46	2,63	2,75	2,84	2,91	2,97	3,02	3,07
	0,01	2,98	3,29	3,47	3,59	3,69	3,76	3,83	3,88	3,93
15	0,05	2,13	2,44	2,61	2,73	2,82	2,89	2,95	3,00	3,04
	0,01	2,95	3,25	3,43	3,55	3,64	3,71	3,78	3,83	3,88
16	0,05	2,12	2,42	2,59	2,71	2,80	2,87	2,92	2,97	3,02
	0,01	2,92	3,22	3,39	3,51	3,60	3,67	3,73	3,78	3,83
17	0,05	2,11	2,41	2,58	2,69	2,78	2,85	2,90	2,95	3,00
	0,01	2,90	3,19	3,36	3,47	3,56	3,63	3,69	3,74	3,79
18	0,05	2,10	2,40	2,56	2,68	2,76	2,83	2,89	2,94	2,98
	0,01	2,88	3,17	3,33	3,44	3,53	3,60	3,66	3,71	3,75
19	0,05	2,09	2,39	2,55	2,66	2,75	2,81	2,87	2,92	2,96
	0,01	2,86	3,15	3,31	3,42	3,50	3,57	3,63	3,68	3,72

Tabla A13. (cont.)

gl del error	P	Número de grupos, incluido el control								
		2	3	4	5	6	7	8	9	10
20	0,05	2,09	2,38	2,54	2,65	2,73	2,80	2,86	2,90	2,95
	0,01	2,85	3,13	3,29	3,40	3,48	3,55	3,60	3,65	3,69
24	0,05	2,06	2,35	2,51	2,61	2,70	2,76	2,81	2,86	2,90
	0,01	2,80	3,07	3,22	3,32	3,40	3,47	3,52	3,57	3,61
30	0,05	2,04	2,32	2,47	2,58	2,66	2,72	2,77	2,82	2,86
	0,01	2,75	3,01	3,15	3,25	3,33	3,39	3,44	3,49	3,52
40	0,05	2,02	2,29	2,44	2,54	2,62	2,68	2,73	2,77	2,81
	0,01	2,70	2,95	3,09	3,19	3,26	3,32	3,37	3,41	3,44
60	0,05	2,00	2,27	2,41	2,51	2,58	2,64	2,69	2,73	2,77
	0,01	2,66	2,90	3,03	3,12	3,19	3,25	3,29	3,33	3,37

Fuente: Esta tabla es un resumen de Dunnett, C.W. (1964). New tables for multiple comparisons with a control. *Biometrics*, 482-491.

Tabla A 14. Valores críticos (Q) del test de Tukey

gl del error	P	Número de muestras a comparar									
		2	3	4	5	6	7	8	9	10	11
5	0,05	3,64	4,60	5,22	5,67	6,03	6,33	6,58	6,80	6,99	7,17
	0,01	5,70	6,98	7,80	8,42	8,91	9,32	9,67	9,97	10,24	10,48
6	0,05	3,46	4,34	4,90	5,30	5,63	5,90	6,12	6,32	6,49	6,65
	0,01	5,24	6,33	7,03	7,56	7,97	8,32	8,61	8,87	9,10	9,30
7	0,05	3,34	4,16	4,68	5,06	5,36	5,61	5,82	6,00	6,16	6,30
	0,01	4,95	5,92	6,54	7,01	7,37	7,68	7,94	8,17	8,37	8,55
8	0,05	3,26	4,04	4,53	4,89	5,17	5,40	5,60	5,77	5,92	6,05
	0,01	4,75	5,64	6,20	6,62	6,96	7,24	7,47	7,68	7,86	8,03
9	0,05	3,20	3,95	4,41	4,76	5,02	5,24	5,43	5,59	5,74	5,87
	0,01	4,60	5,43	5,96	6,35	6,66	6,91	7,13	7,33	7,49	7,65
10	0,05	3,15	3,88	4,33	4,65	4,91	5,12	5,30	5,46	5,60	5,72
	0,01	4,48	5,27	5,77	6,14	6,43	6,67	6,87	7,05	7,21	7,36
11	0,05	3,11	3,82	4,26	4,57	4,82	5,03	5,20	5,35	5,49	5,61
	0,01	4,39	5,15	5,62	5,97	6,25	6,48	6,67	6,84	6,99	7,13
12	0,05	3,08	3,77	4,20	4,51	4,75	4,95	5,12	5,27	5,39	5,51
	0,01	4,32	5,05	5,50	5,84	6,10	6,32	6,51	6,67	6,81	6,94
13	0,05	3,06	3,73	4,15	4,45	4,69	4,88	5,05	5,19	5,32	5,43
	0,01	4,26	4,96	5,40	5,73	5,98	6,19	6,37	6,53	6,67	6,79
14	0,05	3,03	3,70	4,11	4,41	4,64	4,83	4,99	5,13	5,25	5,36
	0,01	4,21	4,89	5,32	5,63	5,88	6,08	6,26	6,41	6,54	6,66
15	0,05	3,01	3,67	4,08	4,37	4,59	4,78	4,94	5,08	5,20	5,31
	0,01	4,17	4,84	5,25	5,56	5,80	5,99	6,16	6,31	6,44	6,55
16	0,05	3,00	3,65	4,05	4,33	4,56	4,74	4,90	5,03	5,15	5,26
	0,01	4,13	4,79	5,19	5,49	5,72	5,92	6,08	6,22	6,35	6,46
17	0,05	2,98	3,63	4,02	4,30	4,52	4,70	4,86	4,99	5,11	5,21
	0,01	4,10	4,74	5,14	5,43	5,66	5,85	6,01	6,15	6,27	6,38
18	0,05	2,97	3,61	4,00	4,28	4,49	4,67	4,82	4,96	5,07	5,17
	0,01	4,07	4,70	5,09	5,38	5,60	5,79	5,94	6,08	6,20	6,31

Tabla A14. (cont.)

gl del error	P	Número de muestras a comparar									
		2	3	4	5	6	7	8	9	10	11
19	0,05	2,96	3,59	3,98	4,25	4,47	4,65	4,79	4,92	5,04	5,14
	0,01	4,05	4,67	5,05	5,33	5,55	5,73	5,89	6,02	6,14	6,25
20	0,05	2,95	3,58	3,96	4,23	4,45	4,62	4,77	4,90	5,01	5,11
	0,01	4,02	4,64	5,02	5,29	5,51	5,69	5,84	5,97	6,09	6,19
24	0,05	2,92	3,53	3,90	4,17	4,37	4,54	4,68	4,81	4,92	5,01
	0,01	3,96	4,55	4,91	5,17	5,37	5,54	5,69	5,81	5,92	6,02
30	0,05	2,89	3,49	3,85	4,10	4,30	4,46	4,60	4,72	4,82	4,92
	0,01	3,89	4,45	4,80	5,05	5,24	5,40	5,54	5,65	5,76	5,85
40	0,05	2,86	3,44	3,79	4,04	4,23	4,39	4,52	4,63	4,73	4,82
	0,01	3,82	4,37	4,70	4,93	5,11	5,26	5,39	5,50	5,60	5,69
60	0,05	2,83	3,40	3,74	3,98	4,16	4,31	4,44	4,55	4,65	4,73
	0,01	3,76	4,28	4,59	4,82	4,99	5,13	5,25	5,36	5,45	5,53
∞	0,05	2,77	3,31	3,63	3,86	4,03	4,17	4,29	4,39	4,47	4,55
	0,01	3,64	4,12	4,40	4,60	4,76	4,88	4,99	5,08	5,16	5,23

Fuente: Esta tabla es un resumen de la Tabla A8 de Steel and Torrie (1985).

Tabla A 15. Coeficientes para contrastes polinómicos ortogonales (con igual espaciamento entre los niveles de la variable independiente)

Número de tratamientos	Grado del polinomio	Tratamientos					
		T1	T2	T3	T4	T5	T6
	2	1 - Lineal	-1	+1			
3	1 - Lineal	-1	0	+1			
	2 - Cuadrático	+1	-2	+1			
4	1 - Lineal	-3	-1	+1	+3		
	2 - Cuadrático	+1	-1	-1	+1		
	3 - Cúbico	-1	+3	-3	+1		
5	1 - Lineal	-2	-1	0	+1	+2	
	2 - Cuadrático	+2	-1	-2	-1	+2	
	3 - Cúbico	-1	+2	0	-2	+1	
	4 - Cuártico	+1	-4	+6	-4	+1	
6	1 - Lineal	-5	-3	-1	+1	+3	+5
	2 - Cuadrático	+5	-1	-4	-4	-1	+5
	3 - Cúbico	-5	+7	+4	-4	-7	+5
	4 - Cuártico	+1	-3	+2	+2	-3	+1
	5- Quintico	-1	+5	-10	+10	-5	+1

BIBLIOGRAFÍA

- Canavos, G.C. 1988. *Probabilidad y Estadística. Aplicaciones y métodos*. McGraw-Hill, México, 651 pp.
- Cochran, W.G., Cox, G.M. 1978. *Diseños experimentales*. Editorial Trillas, México, 661 pp.
- Dyke, G., 1997. "How to avoid bad statistics". *Field Crops Research*, 51:165-187.
- Federer, W.T., 1961. "Augmented design with one way elimination of heterogeneity". *Biometrics*, 17:447-473.
- Federer, W.T. and Ragavarao, D., 1975. "On augmented designs". *Biometrics*, 31: 29-35.
- Fleckinger, J., 1954. "Observations recentes sur l'ecologie du pommier a cidre". *VIII Congres Inter. Botan.*, 10 pp.
- Gates, C.E., 1991. "A user's guide to misanalyzing planned experiments". *HortScience*, 26(10):1262-1265.
- Gilligan, C.A., 1986. "Use and misuse of the analysis of variance in plant pathology". *Advances in Plant Pathology*, 5: 225-261.
- Gomez, K.A. and Gomez, A.A., 1984. *Statistical procedures for agricultural research*. John Wiley & Sons, New York, 680 pp.
- Hsu, J.C., 1996. *Multiple comparisons: Theory and Methods*. Chapman and Hall, New York.
- Lipton, W.J., 1992. "How does your rating system rate? (Part III)". *ASHS Newsletter* 8(1):9.
- Little, T.M. and Hills, F.J., 1991. *Métodos estadísticos para la investigación en agricultura*, 2ª Ed. Trillas, México, 270 pp.
- Little, T.M., 1978. "If Galileo published in HortScience". *HortScience*, 13(5):504-506.
- Little, T.M., 1981. "Interpretation and presentation of results". *HortScience*, 16(5):637-640.

- Marini, R.P., 1999. "Are nonsignificant differences really not significant?" *HortScience*, 34(5):761-762.
- Martínez-González, M.A., Sánchez-Villegas, A. and Faulin Fajardo, J., 2006. *Bioestadística amigable*. 2ª ed. Díaz de Santos, Madrid, 919 pp.
- Pearce, S.C., 1976. *Field experimentation with fruit trees and other perennial plants. Technical Communication No. 23*, 2nd Ed. Commonwealth Agricultural Bureaux, England, 182 pp.
- Petersen, R.G., 1985. "Augmented designs for preliminary yield trials (Revised)". *Rachis*, 4(1): 27-32.
- Petersen, R.G., 1994. *Agricultural Field Experiments. Design and Analysis*. Marcel Dekker, Inc., New York, 409 pp.
- Snedecor, G.W. and Cochran, W.G., 1974. *Métodos estadísticos*. C.E.C.S.A., México, 703 pp.
- Sociedad Española de Ciencias Hortícolas (SECH), 1999. *Diccionario de Ciencias Hortícolas*. Mundi-Prensa, Madrid, 605 pp.
- Steel, R.G.D. and Torrie, J.H., 1985. *Bioestadística: Principios y procedimientos*. McGraw-Hill, Bogotá, 622 pp.
- Willavize, S.A., Carmer, S.G. and Walker, W.M., 1980. "Evaluation of cluster analysis for comparing treatment means". *Agronomy Journal*, 72: 317-320.



AGRICULTURA



FORMACIÓN



GANADERÍA



PESCA Y ACUICULTURA



JUNTA DE ANDALUCÍA
CONSEJERÍA DE AGRICULTURA Y PESCA