

GILBERTO PAEZ BOGARIN



Métodos de Investigación e PRODUCCION ANIMAL



INSTITUTO INTERAMERICANO DE CIENCIAS AGRICOLAS DE LA OEA
Centro de Enseñanza e Investigación
TURRIALBA, COSTA RICA

Disciplina de Zootecnia
1964



METODOS DE INVESTIGACION EN PRODUCCION ANIMAL

por

Gilberto Páez Bogarín

**Instituto Interamericano de Ciencias Agrícolas de la OEA
Centro de Enseñanza e Investigación
Departamento de Industria Animal
Turrialba, Costa Rica
1964**

00007018

~~007489~~

METODOS DE INVESTIGACION EN PRODUCCION ANIMAL

Apuntes del Curso del Ing. Gilberto Páez,

tomados por estudiantes graduados del

Departamento de Industria Animal

(Sin corregir)

Enero-Abril 1964

CIRCULACION PRIVADA

**Instituto Interamericano de Ciencias Agrícolas de la OEA
Centro de Enseñanza e Investigación
Turrialba, Costa Rica
1964**

ADVERTENCIA

No se tenía el propósito de poligrafiar estas notas sobre introducción a la técnica experimental, puesto que ellas posiblemente no contienen nada nuevo para la mayoría de las personas dedicadas a la investigación, quienes obviamente poseen un mayor cúmulo de conocimientos y experiencias que el autor, que les vinculan muy fuertemente con el Método Científico. Sin embargo se aceptó la reproducción de estos apuntes, accediendo a la petición de los profesores y estudiantes del Departamento de Industria Animal del IICA.

Probablemente se encuentren muchos errores, especialmente gramaticales, ya que no se ha hecho ninguna corrección a las notas originales preparadas para las clases; por esta razón las presentamos en forma preliminar y conciente de que en el futuro deben mejorarse y completarse.

Se notará asimismo que no se ha seguido una secuencia lógica en la presentación de los tópicos. Esto obedece a que las clases fueron dictadas atendiendo a las necesidades y problemas que presentaban los estudiantes en el aula. Por esa razón también se ha concedido mucho mayor énfasis al análisis e interpretación de resultados, y a algunos diseños que ellos consideraban de primordial interés para la investigación zootécnica.

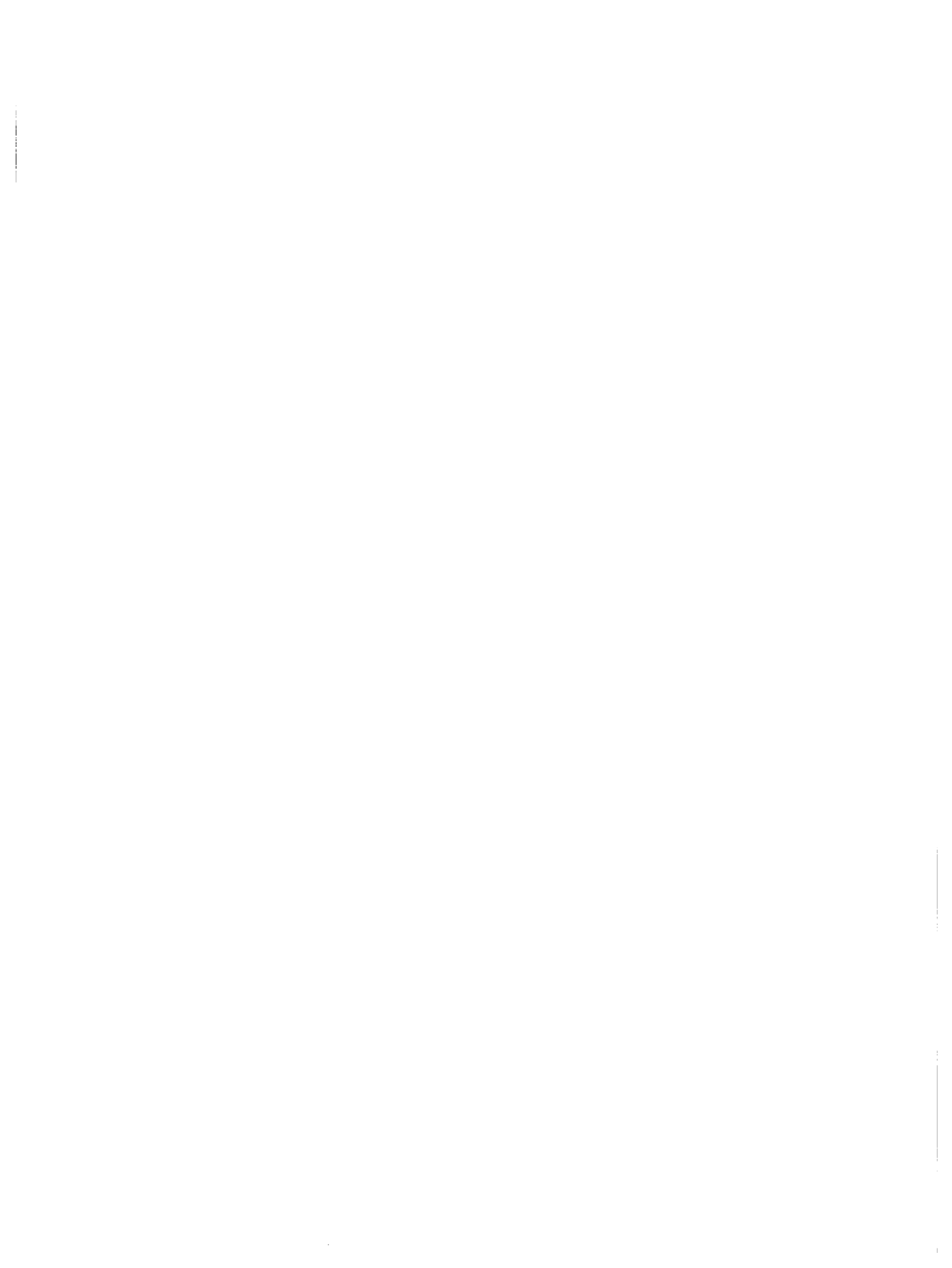
Las fuentes de información que sirvieron de base para las clases se indican en la parte final de estos apuntes.

Finalmente desea dejar constancia de su sincero reconocimiento al Dr. John V. Bateman, Jefe del Departamento de Industria Animal del IICA por su inestimable colaboración en el desarrollo de las clases, a los Ings. Héctor Muñoz y Alvaro Aguirre quienes participaron activamente en la planificación del curso, a todos los estudiantes graduados, quienes asistieron a estas clases con mucho interés y entusiasmo, a la Sra. Mireya de Vega y a la Srta. Emilia Fernández, nuestro reconocimiento especial.

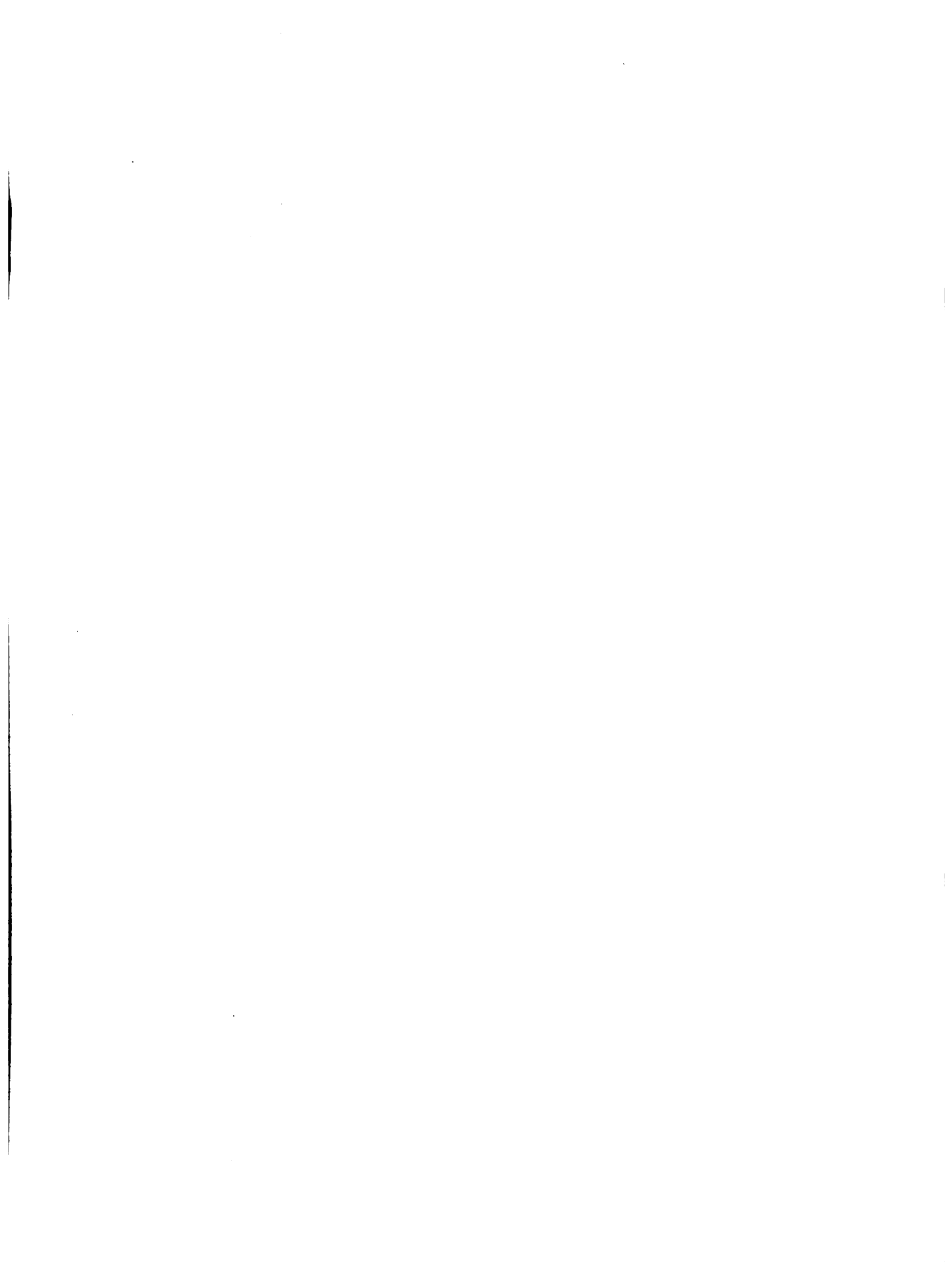
El autor

CONTENIDO

	Página
PARTE GENERAL - Método Científico	1
Experimentación e investigación	2
Característica general del método científico	3
Esquema de una investigación	4
Objetivos generales del diseño de experimentos	6
Postulados de la experimentación	7
Número de repeticiones	9
Determinación del número necesario de repeticiones	10
Sorteo	15
Tamaño y forma de la parcela	17
Métodos usualmente seguidos para determinar el tamaño y la forma óptima de la parcela	18
ANALISIS DE VARIANCIA	20
Propósito del análisis de variancia	20
Modelo matemático	21
Prueba de una hipótesis estadística	22
Análisis funcional de la variancia	25
Comparaciones de tendencias	38
Regresión polinomial aplicado a datos sin repetición ..	43
PRUEBA DE COMPARACION MULTIPLE	46
La prueba de diferencia mínima significativa o llamada también prueba múltiple de "t" o "l.s.d.'test"	47
Test de Student-Newman-Keuls	49
Prueba de rango múltiple de Duncan	50
Prueba de Tukey	51
Comentario general sobre las pruebas de rango múltiple.	52
Comentarios sobre la prueba de Duncan y de Tukey	52
Interpolación armónica	53
EJEMPLO DE APLICACION DE LAS PRUEBAS DE RANGO MULTIPLES	55
Prueba de lsd	56
Prueba de Student-Newman-Keuls	57
Prueba de Duncan	59
Prueba de Tukey	61
PARTE ESPECIAL	63
Análisis de variancia con números desiguales en la sub- clase	63
ANALISIS POR EL METODO DE MINIMOS CUADRADOS DE DATOS CON NUMERO DESIGUAL EN LA SUB-CLASE	97



Principios de mínimos cuadrados con datos de un criterio de clasificación	98
Un criterio de clasificación con regresión o covariancia	118
DATOS CON DOS CRITERIOS DE CLASIFICACION	135
Modelo matemático	136
Ecuaciones de mínimos cuadrados	137
Imposición de restricciones	137
Inversión de la matriz reducida y solución de las ecuaciones	137
Cálculos de las sumas de cuadrados para el análisis de variancia	138
Errores estandares, comparaciones ortogonales y otras comparaciones	139
Estimación de los componentes de variancia	139
Absorción de la ecuación $\mu + a_i$	141
DATOS CON DOS CRITERIOS DE CLASIFICACION CON INTERACCION	161
Restricciones impuestas	163
Análisis completo de mínimos cuadrados	164
Estimación de los componentes de variancia	173
MODELO DE ANALISIS USADO EN EXPERIMENTOS GENETICOS	179
Análisis de datos provenientes de cruzamientos	180
Ejemplo numérico	187
ANALISIS SIMULTANEO DE DATOS DE CRUZAS Y PUROS	197
Construcción y aplicación de las transformaciones matrices	202
ANALISIS SIMULTANEO DE DATOS DE CRUZAS Y PUROS	215
ESTIMACION DEL INDICE DE HERENCIA O HEREDABILIDAD	217
Métodos para estimar la heredabilidad	218
Método de regresión de descendientes sobre el promedio de los padres	220
Método de correlación	225
Regresión intra-toro de los descendientes sobre la madre	226
CONSANGUINIDAD	227
Parentesco	228
Cálculo de coeficiente de parentesco	229
Coeficiente de consanguinidad	231



	Página
Coeficiente de parentesco con antecesores consanguíneo	236
MODALIDAD DE LOS ENSAYOS CON ANIMALES	239
Experimento con animales menores	239
Experimentos con animales del tipo intermedio	240
Experimento con animales mayores	241
Experimento con vacas lecheras en períodos sucesivos .	243
Experimento de doble reversión o alternativo	245
ESTANDARIZACION DE DATOS	248
Método de regresión	249
Método por multiplicación	250
Corrección por adición	252
Otras correcciones	253
BIBLIOGRAFIA	253
APENDICE	257

PARTE GENERAL

METODO CIENTIFICO

Según O. Kempthorne, el método científico es la aplicación de la lógica y la objetividad al estudio de los fenómenos. Sin embargo debemos reconocer que es difícil dar una definición exacta de lo que es el método científico. Si por ejemplo nosotros nos preguntáramos qué es la ciencia? probablemente no estamos en condiciones de dar una definición clara y correcta. Esta dificultad surge del hecho, que el significado de la ciencia no es fijo sino es mas bien dinámico. El significado de la ciencia ha evolucionado en el transcurso del tiempo, en otras palabras ha ido tomando nuevo significado a medida que iba madurando (mayoría de edad de la ciencia).

Despojándonos de nuestras pretenciones de definir la ciencia, simplemente la consideraremos como un procedimiento para:

- 1) Contestar preguntas
- 2) Resolver problemas y
- 3) Desarrollar procedimientos más efectivos para contestar preguntas y resolver problemas.

Los procedimientos que caracterizan a la ciencia generalmente se refieren a herramientas, técnica y métodos científicos.

La herramienta científica constituye el instrumento físico o conceptual que se usa en la investigación científica.

La técnica científica es la manera de lograr un objetivo científico, un curso de acción científica o en otras palabras, es la manera de usar las herramientas científicas.

El método científico es el procedimiento técnico escogido para evaluar cursos alternativos de la acción científica.

El estudio del método científico frecuentemente se considera como metodología. El objetivo de la metodología es mejorar el procedimiento y criterio usado en la investigación científica. Por esta razón la metodología se considera como la lógica de la ciencia.

Sin pretensiones de ninguna otra laya nosotros nos limitaremos simplemente a mencionar los aspectos principales del método científico que son: a) observación y experimentación. Esta parte se refiere a la recolección de todas las informaciones posibles acerca de la investigación; b) análisis e interpretación de los datos y en base a ellos obtener conclusiones objetivas que respondan a las preguntas planteadas; c) postular leyes generales.

Unas de las principales palancas del método científico es la Estadística, que participa en una investigación científica por medio de sus dos ramas: a) diseño de experimentos, de muestras de censos, etc., esta parte de la estadística tiene por objeto la recolección de la información; b) análisis estadístico, esta parte tiene por objeto, resumir las informaciones, calcular ciertas constantes necesarias para describir la población (\bar{X} , s^2 , s_x , CV, etc.) y la inferencia estadística.

Experimentación e investigación

La palabra experimentación a veces se toma como sinónimo de investigación científica, sin embargo debe tenerse muy en cuenta que no todas las investigaciones científicas llevan envuelta la experimentación. La experimentación tal como se ha concebido en el siglo XIX envolvía el control físico de los objetos, eventos y sus

propiedades. Pero una investigación científica se puede hacer sin el control físico de los factores, por ejemplo lo relacionado con la astronomía, aunque la situación puede cambiar en el futuro, pero en la era en que vivimos los astrónomos no han podido aún controlar físicamente los objetos de sus estudios. El control se puede también obtener por manipulación conceptual o representaciones simbólicas (modelos) del fenómeno bajo estudio.

Por las consideraciones hechas podemos nosotros, desde el punto de vista práctico decir que la investigación científica se puede llevar a cabo por experimentación o por observación. En el primer caso se ejerce cierto, bastante o completo control sobre los factores en estudio y otros factores que pueden enmascarar los resultados. En el segundo caso como es del tipo observacional, a veces es imposible ejercer control físico sobre factores en estudio, lo mismo sobre los factores discrepantes.

Característica general del método científico

- Examen de las informaciones ya existentes. Esto se concreta a la revisión de literatura y comunicaciones personales.
- En base al conocimiento que se tiene del problema, derivar hipótesis por medio de la deducción.
- Someter la hipótesis a pruebas experimentales u observacionales.
- La verificación de la hipótesis no es absoluta, sino solo pone de manifiesto que los resultados son compatibles o nó con ella, en otras palabras, verificar una hipótesis no implica que la decisión sea cien por ciento cierta, ya que generalmente trabajamos

con números muy reducidos de la población; por esta razón en vez de decir prueba de hipótesis, nosotros vamos a preferir usar la palabra test de hipótesis. La palabra "test" significaría para nosotros una prueba no muy rigurosa; por otro lado si queremos referirnos a una prueba rigurosa, usaremos la palabra "proof" que implicaría pruebas de rigor.

Esquema de una investigación

Siguiendo los pasos del método científico, cuando a) surge un problema; b) se revisa la literatura existente al respecto, o consultar a personas que poseen ciertas experiencias en el asunto; c) en base a los conocimientos ya existentes formular una hipótesis; d) se plantean los objetivos generales y específicos del trabajo; e) se diseña un experimento o una muestra o un censo, etc. de acuerdo a los objetivos que se persiguen con la investigación. El diseño está sujeto a los objetivos del trabajo; f) se lleva a cabo el trabajo por el tiempo que el investigador cree necesario y desde luego en consonancia con los objetivos planteados; g) se recogen las informaciones (o datos); h) se analizan y se interpretan las informaciones; i) discusión y conclusiones de los resultados.

Un ejemplo práctico aclarará el proceso del método científico:

a) Problema. Existe el problema que preocupa a un grupo de personas, (jugadores) si la moneda (utilizada para el juego) tiene las dos caras (escudo y corona) balanceada o equilibrada o una de las caras tiene mayor peso y por lo tanto menor probabilidad de salir.

- b) Revisión de literatura. No se tiene referencia acerca del problema.
- c) Hipótesis. La probabilidad de que caiga escudo es igual a la probabilidad de que caiga corona (moneda equilibrada).
- d) Objetivo. El objetivo del trabajo consiste en averiguar si la moneda tiene las dos caras balanceadas.
- e) Diseño del experimento. Se tirará la moneda un número x de veces en condiciones completamente uniformes.
- f) Experimento. Se tiró la moneda 1000 veces.
- g) Resultados. Corona = 505 veces
Escudo = 495 veces
- h) Test de hipótesis. Para verificar la hipótesis (H_0) plantada en (b), usaremos en nuestro caso la prueba χ^2 (chi al cuadrado).

Alternativa	Observado (o)	Esperado e	o-e	(o-e) ²	$\frac{(o-e)^2}{e}$ = 2
Corona	505	500	5	25	$25/500 = 0.05$
Escudo	495	500	-5	25	$25/500 = 0.05$
Total	1000	1000	0		$\chi^2 = 0.10$

La χ^2 obtenida (0.10) es menor que la χ^2 tabular que es igual a 3.84 (χ^2 , P.05, GL 1 = 3.84), esto pone de manifiesto que el resultado obtenido es compatible con la hipótesis formulada, por eso aceptamos la hipótesis de que la moneda está balanceada.

- i) Conclusiones. La moneda está balanceada o equilibrada. Esto no es una prueba rigurosa. Para tener una prueba rigurosa se nece-

sitaría tirar la moneda infinito número de veces. Por qué se formula la hipótesis nula? En general resulta más fácil rechazar o aceptar una cosa que criticarla. Otra razón es que la hipótesis nula (H_0) es una sola y las hipótesis alternativas son varias. Por estas y otras razones se usa en la estadística la hipótesis de nulidad "null hypothesis".

Objetivos generales del diseño de experimentos

1. Recoger informaciones válidas o datos fidedignos
2. Eliminar al máximo posible los errores, pero como no es posible eliminarse totalmente:
3. Tratar de medir validamente el error experimental.

En un experimento aparecen una serie de factores incontrolables, tales como la heterogeneidad del suelo, diferencia genética entre plantas o animales, influencias climáticas, errores humanos, etc. Estos factores inciden en forma negativa sobre la uniformidad de la producción; como consecuencia introduce ciertos grados de incertidumbre en los resultados. Aquí pueden presentarse dos situaciones:

1. Puede suceder que una diferencia real entre dos tratamientos quede enmascarada por el error experimental, por ejemplo el promedio de un tratamiento A (μ_A) sea mayor que el promedio de otro tratamiento B (μ_B) pero en el experimento puede aparecer $\bar{X}_A = \bar{X}_B$ (debido a error experimental). Es decir que se obtiene un resultado contrario a lo real, en este caso aceptamos la hipótesis nula siendo ella falsa y cometemos error del Tipo II.

2. La otra situación es lo contrario, el promedio verdadero de dos tratamientos son iguales $\mu_A = \mu_B$ pero en el experimento se obtienen resultados diferentes:

$$\left. \begin{array}{l} A < B \\ A > B \end{array} \right\} \text{ debido a error experimental}$$

Basado en este resultado rechazamos la hipótesis nula. Siendo ella verdadera, cometemos error del Tipo I.

En las dos situaciones hemos cometido graves errores porque los resultados estaban enmascarados por el error experimental; entonces surge el diseño del experimento para controlar esos errores y dar validez a los resultados.

Algunos principios del diseño de experimentos se han aplicado desde hace mucho tiempo, por ser cuestiones de lógica. Sin embargo, el extraordinario progreso que alcanzó desde hace tres décadas más o menos, se debe a R. A. Fisher, quien también tiene el mérito de ser el primero que dio a conocer el desarrollo matemático del Análisis de Variancia, en 1923.

Postulados de la experimentación

1. El experimento no debe tener inclinación viciada "bias", sesgo. Es decir, que el experimentador no debe tener preferencia sobre ciertos tratamientos sino debe estimar los valores de ciertas constantes de la muestra $(\bar{X}, s^2, s_{\bar{x}})$ sin ningún vicio para que así estas estimaciones se aproximen a los valores de los parámetros $(\mu, \sigma^2, \sigma_{\bar{x}})$.

2. Debe existir una medida del error experimental.

Ocurre a veces que un mismo ensayo se lleva a cabo por dos personas "A" y "B". La persona "A" declara que los tratamientos son decididamente diferentes. La persona "B" dice que no existe diferencia real entre los tratamientos probados. A qué se debe esta discrepancia de opiniones? Posiblemente a error experimental, de ahí la importancia de tener una medida del error, para que las conclusiones sean objetivas. Para ayudar a esto se hace el "test" de significancia y para hacer esto se necesita del error experimental y para medir el error experimental se necesita: a) repetir los tratamientos de un experimento varias veces (Replication); y b) sortear la distribución de los tratamientos (Randomization). Hay casos que no necesitan una prueba formal, ("test" de significancia), por ejemplo, al comparar el efecto de dos herbicidas se observa que uno controló el 100%, mientras que el otro no tuvo efecto. En estos casos es evidente la superioridad del uno sobre el otro. Al hacer el "test" de significancia debe tenerse presente lo siguiente: a) Si nosotros aplicamos un criterio de aceptar solamente diferencias, muy evidentes o muy grandes y despreciar diferencias pequeñas, se debe tener en cuenta que la suma de muchas diferencias pequeñas dá diferencias grandes. En este caso se comete el error Tipo II; y b) si se quieren medir diferencias pequeñas se necesita hacer un buen experimento con muchas repeticiones, mediciones exactas de las variables, diseño y análisis adecuados. Si por el contrario se adopta el criterio de aceptar toda diferencia entre tratamiento, por pequeña que ella sea, se

comete error del Tipo I. Contra este error nos protege el nivel de significancia estadística (P.05, P.01).

3. Debe haber un objetivo bien definido. El objetivo es la base sobre la cual descansa el experimento; debe haber buenas razones para escoger los tratamientos, en qué área se aplicarán los resultados, por qué y para qué se va a conducir el experimento.
4. El experimento debe tener suficiente exactitud. La exactitud se puede alcanzar de varias maneras: a) eliminando errores de técnica; b) se puede aumentar a través del diseño y análisis; c) aumentando el número de repeticiones.
5. El experimento debe tener suficiente amplitud (scope). Si se conduce un experimento con la finalidad de determinar la mejor ración alimenticia para cerdos, y para el estudio se escogen cerdos jóvenes de cuatro meses, las conclusiones que se siguen de este experimento se referirán solo a cerdos de esa edad. Se puede aumentar la amplitud haciendo experimentos con cerdos de varias edades. A medida que aumenta la amplitud, las conclusiones son mejores.

Número de repeticiones

La experiencia ha demostrado que difícilmente se consiguen resultados satisfactorios con ensayos que tienen menos de 20 unidades experimentales. Este es el número que se toma en general como mínimo para experimentación agrícola o zootécnica. Así, si un experimento consta de dos tratamientos, por lo menos debe tener unas 10 repeticiones cada uno.

Otra advertencia útil que debe tenerse en cuenta es que el experimento por lo menos debe tener 10 GL para el error.

Estas dos restricciones en la práctica muchas veces no se cumplen (experimento de Física, Química, Laboratorio).

Determinación del número necesario de repeticiones

Antes de aplicar cualquier fórmula matemática para calcular el número de repeticiones requerido en un experimento, se necesita resolver dos dificultades de orden práctico.

1. Estimar la variabilidad debida a error experimental, que posiblemente se va a presentar en el experimento. Esto se puede conseguir mediante el análisis de experimentos previos.
2. El grado de precisión que se requiere en los resultados. Una forma de fijar la precisión deseada es contestando la siguiente pregunta: Cuál es la diferencia real entre dos tratamientos (δ) que se quieren detectar en el experimento?

$$\delta = \mu_A - \mu_B$$

La escogencia de δ depende hasta cierto punto del juicio personal del investigador. Hay, sin embargo, ciertas consideraciones que son de gran ayuda en la determinación de δ a saber:

1. El costo de los tratamientos. En experimentos planeados con un sentido económico, la δ debe ser mayor que el costo diferencial de los tratamientos.
2. El número de otros tratamientos promisorios que se van a ensayar o se pueden ensayar. No vale la pena hacer un experimento muy grande para detectar una δ muy pequeña, si eso

implica que faltaran recursos para otros experimentos en los cuales se espera que hayan algunos δ grandes. Cochran-Cox dan la siguiente fórmula para calcular el número de repeticiones.

$$r \geq 2 \left(\frac{\delta}{\sigma} \right)^2 (t_{\alpha} + t_{\gamma})^2$$

donde:

r = número de repeticiones en el experimento

σ = desviación estandar verdadera del error experimental

δ = diferencia, mínima verdadera que se quiere detectar

t = t de Student

α = nivel de probabilidad al que se hace la prueba

γ = probabilidad de obtener un resultado significativo, si la verdadera diferencia fuera igual a δ (γ se refiere a una cola de la distribución de t).

Como α es la probabilidad de cometer un error se escogen valores pequeños de α (generalmente 0.10, 0.05, 0.01); γ es la probabilidad de tener un acierto por lo que se escogen valores de γ alto ($\gamma = 0.50, 0.70, 0.80$).

La fórmula anterior hay que resolverla por aproximaciones sucesivas, por el hecho de que los grados de libertad de t son del error del futuro experimento, que dependen en parte de r (también depende del diseño y del número de tratamientos).

La t de Student (de una cola) para una probabilidad de 0.50 es igual a 0. De ahí tenemos para $\gamma = 0.50$, la anterior fórmula se simplifica en:

$$r \geq \left(\frac{\delta^2}{\sigma^2} \right)^2 t_{\alpha}^2$$

Un ejemplo práctico aclarará el concepto, si tenemos un experimento en bloques al azar con cinco tratamientos con los datos siguientes:

$$J = 2 \text{ fanegas de café}$$

$$\sigma = 1.6$$

$$\alpha = 0.05 \text{ y } \gamma = 80\% = 0.80$$

$$r \geq 2 \left(\frac{1.6}{2} \right)^2 \left(t_{0.05} + t_{0.80} \right)^2$$

la ecuación se resuelve por tanteo para varios valores de r

$$\text{Gl. error} = (5 - 1) (r - 1)$$

así Gl. 4, 5, 6, ---- n

$$\text{Gl. 11} = 4 \times (11-1) = 40$$

$$t_{0.05} = 2.02 \text{ tabla de dos colas}$$

$$t_{0.80} = 0.851 \text{ tabla de una cola}$$

$$r \geq \left(\frac{1.6}{2} \right)^2 (2.02 + 0.851)^2 = 10.6$$

$$\underline{r = 11}$$

Debido a que es difícil escoger el valor de J que se ha de sustituir en la fórmula, es más práctico despejar:

$$J \geq \frac{\sqrt{2} \sigma t_{\alpha}}{\sqrt{r}}$$

y calcular para varios valores de r las correspondientes J , ejemplo.

\sqrt{r}	$\sqrt{2}$	σ	t_{α}	GL error	$\frac{\sqrt{2} \sigma t_{\alpha}}{\sqrt{r}} = \int$ fanegas
2	1.4142	1.6	2.776	4	4.44
3	"	"	2.306	8	3.02
4	"	"	2.179	12	2.46
5	"	"	2.120	16	2.14
6	"	"	2.086	20	1.93
7	"	"	2.064	24	1.76
8	"	"	2.048	28	1.64
9	"	"	2.037	32	1.54
10	"	"	2.030	36	1.45
11	"	"	2.021	40	1.38
12	"	"	2.014	44	1.32

El despeje de \int también se puede hacer fácilmente con la fórmula dada por Cochran y Cox.

Hay otros dos puntos que hay que tener en cuenta al calcular r:

a) En los experimentos factoriales puede existir lo que se llama repetición oculta (hidden replication) lo que trae una reducción en el número de repetición. Por ejemplo, un 2^4 de NPKB donde la interacción de 3 orden es depreciable si $r = 8$; el número de repetición real que se necesita es $= \frac{8}{4} = 2$ bloques, en factoriales grandes se puede poner una sola repetición.

b) Una solución interesante y reciente del problema, puede ser obtenida por el uso de la prueba de Tukey (como veremos más adelante),

para esto se necesita conocer lo siguientes:

s_2 = desviación estandar del experimento previo

n_2 = GL del ensayo previo en condiciones análogas

\mathcal{J} = diferencia que se quiere detectar

q = amplitud total estudentizada (studentized range) para el experimento que se va a hacer.

F = valor tabular al nivel α con GL n_1 (GL del nuevo experimento) y n_2 está dado por la fórmula:

$$r = \frac{q^2 s_2^2 F}{\mathcal{J}^2}$$

Este número de repeticiones nos garantiza una probabilidad α de que el ensayo no compruebe la diferencia y una probabilidad de $1 - \alpha$ de que sea comprobada estadísticamente por la prueba de Tukey.

Como los valores de q y F dependen de r los valores serán obtenidos por tanteos y aproximación sucesiva.

Ejemplo: Un experimento con cinco variedades de caña de azúcar, que en el ensayo anterior tenía $s_2 = 7.4$ t./ha. con $n_2 = 60$. Se considera que el experimento a conducir debe detectar una diferencia de producción de 15 t./ha. o más. El ensayo se conducirá en bloques al azar; variedad; $5-1 = 4$ GL y $n_1 = 16 = (5-1)(5-1)$; al 5% la tabla de $q = 4.34$. F (con $n_1 = 16$ y $n_2 = 60$ GL) = 1.81 luego

$$r = \frac{(4.32)^2 (7.4)^2 (1.81)}{15^2} = 8.3 \text{ repeticiones}$$

El número de repeticiones a usar está entre el número usado inicialmente (5) y el obtenido 8.3.

Si tenemos $r = 7$; $q = 4.17$ y $F = 1.70$

$$r = \frac{(4.17)^2 (7.4)^2 (1.70)}{15^2} = 7.2 \text{ repeticiones}$$

El número de repeticiones debe ser de siete. Con este número a la probabilidad del 95% se puede detectar una diferencia de 15 t./ha.

Una solución interesante se obtiene también cuando se conoce el CV (coeficiente de variación) y la \mathcal{J} (diferencia en porcentaje).

$$r = \frac{q^2 (CV)^2 F}{2}$$

Sorteo "Randomization"

Existen varios métodos para distribuir los tratamientos al azar en las unidades experimentales (parcelas).

Fisher (1937) "The desing of Experiment") expone un método que consiste en tomar un mazo de tarjetas numeradas de 1 a 100, las que se disponen al azar barajándolas repetidas veces. Si 20 es el número de tratamientos a ensayarse con cinco repeticiones, se las numera del 1 al 20, luego se extrae una tarjeta del mazo; el número de la tarjeta extraída se divide entre 20, el residuo de la división corresponde al tratamiento a sembrarse en la primera parcela. Ejemplo, si la primera tarjeta extraída fuera 33; $\frac{33}{20} = 1$ más residuo 13; el tratamiento 13 irá en la primera parcela; si la tarjeta fuera 40 el residuo es 0. Se tomarán 20 y así sucesivamente hasta completar la primera repeticón, luego se baraja el mazo y se procede de la misma manera para el siguiente caso.

En este caso tenemos $r = 5$ y número de tratamiento = 20, entonces 100 es un número exactamente divisible por 20 y cada uno de los

tratamientos se halla representado cinco veces en el mazo.

Si el número de tratamiento fuera 19, con cinco repeticiones se usarán 95 cartas excluyendo el resto.

Otro método de hacer los sorteos es por medio de las tablas de número al azar de Tippett (1927) o de Fisher y Yates "Statistical Technique for Biological, Agricultural and Medical Research", en esas tablas se puede iniciar en cualquier punto escogido al azar "Random start" y seguir en cualquier sentido.

Si por ejemplo tenemos 30 macetas en las cuales queremos distribuir al azar dos tratamientos, se numeran las macetas del 1 al 30, y del 1 al 15 se aplica el tratamiento A; del 16 al 30 el tratamiento B. Como esto se hace al azar cada tratamiento tiene igual oportunidad de caer en cualquier parte o dicho de otro modo, toda maceta tiene igual oportunidad de recibir cualquier tratamiento.

Si hacemos el sorteo en la tabla del número al azar que tiene como principio, que todos los números de un dígito tienen igual probabilidad de salir; lo mismo los números de dos dígitos de 01, 02 - - - 98, 99 tienen igual probabilidad de salir. Todos los números de tres dígitos tienen igual probabilidad de salir, etc.

Para distribución al azar con 30 macetas (tomaremos en la tabla los valores) de 01 hasta 90.

01	=	31	=	61	=	tratamiento nº 1
02	=	32	=	62	=	" nº 2
.
.
.
30	=	60	=	90	=	" nº 30

También pueden emplearse las tablas que dan Cochran y Cox "Tables of random permutations".

Tamaño y forma de la parcela

Para cultivos anuales podemos considerar normalmente dos tamaños de parcelas. La pequeña parcela de almácigo para cultivo a mano y la grande para cultivo a máquina. En algunos casos esta distinción es arbitraria.

De manera general necesitamos parcela grande cuando estudiamos efecto de insecticida o fungicida en condiciones de campo. En el caso de insectos, el control se estudia en un área pequeña en el centro de la grande, para evitar la migración de insectos de las parcelas cercanas, por efectos del tratamiento adicionado.

Respecto a la forma de la parcela, puede decirse que la literatura cita a la forma rectangular, como mejor que la cuadrada, pero no siempre es mejor. Es claro, que si no se conoce la tendencia de la fertilidad, debería preferirse la rectangular.

En experimento de irrigación es más fácil irrigar parcelas grandes que pequeñas.

Federer resume los factores que afectan el tamaño de la parcela de la siguiente manera.

1. Consideraciones prácticas tales como facilidades y equipo con que se cuenta, material experimental y terreno, limitaciones de la mano de obra a pesarse.
2. Naturaleza del material experimental.

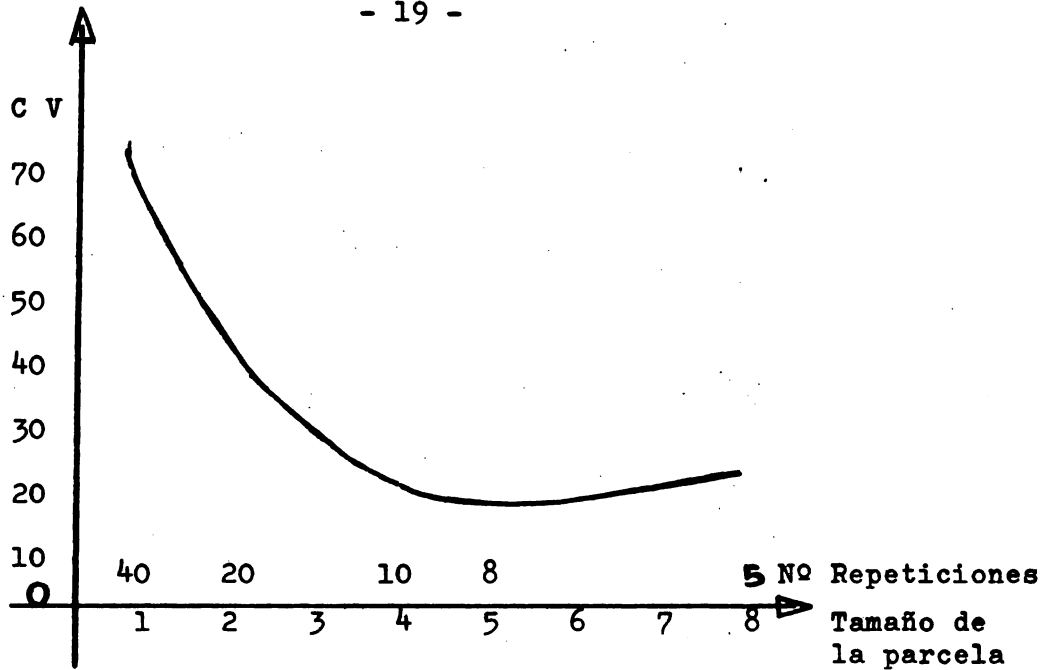
3. Número de tratamientos por bloque o bloques incompletos. A mayor número de tratamiento por bloque corresponderá una parcela de menor tamaño, para evitar que el bloque incluya un área más heterogénea.
4. Variabilidad entre individuo o unidad de muestreo dentro de la unidad experimental. La magnitud de la variabilidad tiene mucho que ver con el tamaño óptimo de la parcela.
5. Costo.

La orientación de la parcela tiene mucho que ver también, para juzgar si una forma es mejor que otra.

Hay casos en que se tiene necesidad de usar bordes; en estos casos las parcelas cuadradas desperdician menor cantidad de material experimental (bordes).

Métodos usualmente seguidos para determinar el tamaño y la forma óptima de la parcela

1) Máxima Curvatura de CV. Posiblemente este es el método más generalizado que consiste en calcular el CV para cada tamaño de parcela y plotear los valores obtenidos en un sistema de ejes coordenados, donde: el eje de la abscisa lleva los distintos tamaños de parcelas y la ordenada los CV, luego se unen los puntos y se obtiene una curva. Dicha curva presenta un punto de inflexión que corresponde a la Máxima Curvatura y ese punto se considera como el mejor tamaño de parcela.



En este ejemplo el mejor tamaño estaría entre cuatro y cinco plantas.

Este método fue objetado por Smith, quien anotó que la Máxima Curvatura depende de la escala de coordenadas.

2) Método de Keller. Este consiste en determinar el cuadrado medio de error para cada tamaño de parcela, luego se calcula el producto del cuadrado medio de error por el número de unidades que componen las parcelas. El producto que de menor valor se considera como óptimo por contrarrestar mejor la variabilidad. Generalmente la parcela unidad da el menor producto, pero considerando ciertos factores de orden práctico resulta que la parcela unidad no es la más recomendada. De ahí que se toma como óptimo, aquel de mayor número de unidades cuyo producto haya crecido en menor proporción.

3) Método de Smith. Smith demostró empíricamente la relación que existe entre el tamaño de la parcela y el cuadrado medio del error y expuso esta relación con la siguiente ecuación.

$$\sqrt{x} = \frac{V}{X^b} \quad \text{o} \quad \text{Log } \sqrt{x} = \text{Log } V - b \lg X \quad \text{o}$$
$$\text{Log } \sqrt{x} = \lg V + b \text{ colg } X$$

donde:

\sqrt{x} = CM error de la parcela de n unidades

V = CM error de la parcela de 1 unidad

b = Coeficiente de regresión que indica la relación entre unidades adyacentes.

Además Smith presenta una fórmula para estimar el costo del experimento.

$$X = \frac{b K_1}{1 - bK_2}$$

K_1 = Costo proporcional al número de parcelas por tratamiento en ensayos previos.

K_2 = Costo proporcional al área total en ensayos previos.

ANALISIS DE VARIANCIA

El análisis de variancia desarrollado por R. A. Fisher en 1923, facilita el análisis e interpretación de los datos obtenidos de las pruebas de campo, invernadero y laboratorio en investigaciones agrícolas y biológicas, también su uso se ha extendido a las ciencias sociales, físicas, ingeniería e industria, etc.

Propósito del análisis de variancia

1. Estimar correctamente ciertas diferencias entre tratamientos. Vale decir que la diferencia entre la estimación y el valor verdadero debe tener una variancia tan pequeña como

sea posible.

2. Obtener algunas ideas de la precisión de las estimaciones, calculando los errores estandares, límites de confianza, etc. Las estimaciones deben estar libres de inclinaciones viciadas (bias).
3. Efectuar las pruebas de significancia. Las más usadas son las pruebas de F y de "t" de la hipótesis de nulidad.

Modelo matemático

Todos los análisis de variancia de un experimento, presuponen un modelo matemático y ciertas hipótesis básicas. Si el experimento está diseñado en bloques al azar, el modelo matemático es el que sigue:

$$Y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$$

El valor Y_{ij} observado en la unidad experimental que recibió un tratamiento i en el bloque j ; μ es el promedio general, τ_i es el efecto del tratamiento i . β_j mide el efecto del bloque j y ϵ_{ij} es la parte de variación debida a factores no controlados.

Las presunciones básicas del análisis de variancia son:

1. Que los diversos efectos son aditivos como se ve en el modelo y no multiplicativo, por ejemplo, como sería el caso de $Y_{ij} = \mu \times \tau_i \times \beta_j \times \epsilon_{ij}$.
2. Que los errores ϵ_{ij} son independientes (no correlacionados).
3. Que los ϵ_{ij} tienen todos una misma variancia
4. Que los ϵ_{ij} tienen distribución normal.

Estas presunciones son muy restrictivas, pero en general no tienen mucha importancia que se verifiquen solo aproximadamente. La prueba de significancia (sea con t o F) no se altera mucho si una distribución fuera apenas aproximadamente normal.

Tal vez tiene mayor importancia la aditividad, pero en general se verifica. Si no se verificara nos obligaría a hacer transformaciones de las variables, aplicando por ejemplo logaritmos. El uso de las transformaciones también permite en general saldar las dificultades que surgen cuando la distribución de los errores se aparta mucho de la normalidad. Si ellos tuvieron distribución binomial, por ejemplo usaríamos una transformación.

$$Z_{ij} = \text{arc sen } \sqrt{y_{ij}} = \text{sen}^{-1} \sqrt{y_{ij}}$$

si tuviera una distribución de Poisson usaríamos una transformación

$$Z_{ij} = \sqrt{y_{ij}}$$

o cuando los valores de y_{ij} son muy bajos es preferible usar

$$Z_{ij} = \sqrt{y_{ij} + \frac{1}{2}}$$

Prueba de una hipótesis estadística

El análisis de variancia es un procedimiento estadístico que ayuda en la interpretación de los datos, pero en sí mismo no es un resumen del experimento. La interpretación deberá incluir un examen cuidadoso de los promedios para hacer una inferencia apropiada. Al hacer una inferencia estadística debe tenerse en cuenta que el promedio observado representa al valor de la población.

Un "test" de significancia es un procedimiento para determinar si cierta cantidad que está sujeta a fluctuaciones al azar, difiere de un valor teórico por una cantidad mayor que lo explicado por una variación al azar solamente. En un "test" de significancia el investigador debe decidir sobre el nivel de significancia apropiado para inferir acerca de los datos (por ejemplo $P 0.05$). Tal nivel indica la probabilidad de concluir erróneamente acerca de la hipótesis nula.

Los errores de aceptar o rechazar una hipótesis nula, se clasifican en dos grupos.

Error Tipo I. Consiste en rechazar la hipótesis nula siendo ella verdadera.

Error Tipo II. Consiste en aceptar la hipótesis nula siendo ella falsa.

La probabilidad de cometer Error Tipo I está bajo el control del investigador y se determina por el nivel de significancia especificada para la prueba de la hipótesis nula.

La probabilidad de cometer un Error Tipo II sin embargo no puede ser fijado precisamente a menos que una hipótesis alternativa pueda ser especificada. La frecuencia con que ocurre el error Tipo II depende de varios factores, principalmente de cual es la precisión con que las constantes estadísticas se calculan. Es obvio que un diseño apropiado y suficiente número de repeticiones podrán minimizar la probabilidad de cometer Error Tipo I y II. Aumentando la precisión de un experimento disminuye el riesgo de cometer ambos tipos de errores, pero particularmente el del Tipo II.

Test de F o \sqrt{v}

La prueba básica para el análisis de variancia es la prueba " Z " de R. A. Fisher. El Prof. Fisher tabuló la distribución en la forma $Z = \lg_e \sqrt{F}$. Fisher y Yates denominaron F a la razón de la variancia (cuadrados medios de tratamiento) / cuadrado medio de error. Hoy día la prueba de " Z " está sustituida por su equivalente " F " de (G. Snedecor, quien la denominó así en honor al Prof. Fisher). La prueba " \sqrt{v} " de F. G. Brieger que compara la desviación estandar del tratamiento con la desviación estandar del error.

Fórmulas:

$$Z = \lg_e \sqrt{F}$$

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

$$\sqrt{v} = \frac{s_1}{s} \quad \text{o} \quad \sqrt{v} = \sqrt{F}$$

Todas las tablas de F publicadas hasta el presente, fueron hechas suponiendo que $s_1^2 > s^2$ de ahí que los valores tabulares de F son mayores que 1; para \sqrt{v} hay tablas > 1 y < 1 .

Ejemplo:

FV	GL	SC	CM (s^2)	s
Tratamientos	3	32.64	10.88	3.30
Error	20	28.80	1.44	1.20
Total	23	61.44		

Estableciendo la $H_0 = \mu_A = \mu_B = \mu_C = \mu_D$ y que CM de la varian-
cia σ_e^2 es igual al CM del error.

$$F = \frac{10.88}{1.44} = 7.56^{**} \quad \left\{ \begin{array}{l} F \text{ tabular (0.01)} = 4.94 \\ \text{con 3 y 20 GL} \end{array} \right.$$

$$v = \frac{3.30}{1.20} = 2.75^{**} \quad \left\{ \begin{array}{l} v \text{ tabular (0.01)} = 2.22 \\ \text{con 3 y 20 GL} \end{array} \right.$$

El valor de v se consigue en la tabla de límite unilateral.

Cuando los valores de v son < 1 debe buscarse en la tabla de límite bilateral.

Análisis funcional de la variancia

El análisis funcional de la variancia se refiere a la prueba de significancia de las comparaciones individuales o de grupos de tratamientos. Estas comparaciones se pueden probar dividiendo los grados de libertad, para efecto de tratamientos en GL simples. Así mismo hacer las divisiones en sus correspondientes sumas de cuadrados.

Es digno destacar que la responsabilidad para decidir sobre las comparaciones apropiadas a hacer, es más del investigador que del estadístico, puesto que eso depende de la pregunta o preguntas a ser contestadas en el experimento.

El investigador generalmente está interesado en hacer dos tipos de comparaciones.

1. Comparaciones de clases
2. Comparaciones de tendencias

Comparaciones de clases

Las comparaciones de clases comprenden diferentes tipos de variables asignadas a las clases, según su naturaleza o el efecto que

produce. Las comparaciones a ser probadas deberán ser planeadas como parte del experimento.

Cuando dos o más comparaciones establecidas van a ser simultáneamente comparadas es deseable que cada una de ellas sean independientes de las otras (ortogonal) así los componentes correspondientes en el análisis de variancia van a ser también independientes. Es posible la división de la suma de cuadrados en muchas diferentes series ortogonales.

Las comparaciones no ortogonales o no independientes tienden a estar afectadas por cada uno de los otros y la separación de sus correspondientes componentes en el análisis resultan en probabilidades ambiguas respecto a la interpretación de los datos.

Comparaciones ortogonales

1) Tratamiento con igual número de repeticiones

Una función lineal de los totales de tratamientos (T_i) es $Z = l_1 T_1 + l_2 T_2 + l_3 T_3 + \dots + l_k T_k$. Si se cumple la condición de: $l_1 + l_2 + l_3 + \dots + l_k = 0$ ($\sum_{k=1}^n l_i = 0$) o sea que la suma de los coeficientes es igual a cero (0) Z se llama un contraste. Los l_i son ciertos coeficientes que uno escoge expreso según los tratamientos que desea comparar.

Regla 1 - Si Z es un contraste entre los totales de tratamiento la cantidad $\frac{Z^2}{r(l_1^2 + l_2^2 + \dots + l_k^2)}$ es un componente de la suma de cuadrados de tratamientos y tiene un grado de libertad.

Regla 2 - Dos contrastes Z y Z' son ortogonales si la suma de

producto de los coeficientes ($\sum l_i x_i$) es igual a cero por ejemplo:

$$Z = l_1 T_1 + l_2 T_2 + \dots + l_k T_k$$

$$Z' = l'_1 T_1 + l'_2 T_2 + \dots + l'_k T_k$$

$$l_1 \times l'_1 + l_2 \times l'_2 + \dots + l_k \times l'_k = 0$$

Regla 3 - Si los contrastes $Z_1, Z_2, Z_3, \dots, Z_{k-1}$ son mutuamente ortogonales, la SC de tratamientos puede dividirse en:

$$\frac{Z_1^2}{D_1} + \frac{Z_2^2}{D_2} + \dots + \frac{Z_{k-1}^2}{D_{k-1}} \text{ donde los } D \text{ son los denominadores}$$

$$D = (r \sum_{i=1}^n l_i)$$

2) Regla para diferentes números de repeticiones de los tratamientos

$Z = l_1 \times T_1 + l_2 T_2 + l_3 T_3 + \dots + l_k T_k$; Z es una función lineal de los totales de tratamientos y recibe el nombre de contraste si se cumple la condición de:

$$r_1 l_1 + r_2 l_2 + r_3 l_3 + \dots + r_k l_k = 0$$

Regla 4 - La cantidad $\frac{Z^2}{r_1 l_1^2 + r_2 l_2^2 + r_k l_k^2}$ es un componente de la SC de tratamientos y tiene un grado de libertad.

Regla 5 - Dos contrastes Z y Z' son ortogonales si se cumple la condición:

$$r_1 l_1 l'_1 + r_2 l_2 l'_2 + \dots + r_k l_k l'_k = 0$$

donde r_i son los números de repeticiones y l_i los coeficientes.

Regla 6 - Similar a la tercera regla para igual número de repeticiones.

Regla 7 - Si k tratamientos tiene el mismo número de repeticiones (r). El error estandar de la función lineal de los tratamientos.

$Z = l_1 \bar{Y}_1 + l_2 \bar{Y}_2 + \dots + l_k \bar{Y}_k$ de los promedios los trata-

mientos es $s_z = \sqrt{\frac{s^2}{r}(l_1^2 + l_2^2 + l_3^2 + \dots + l_k^2)}$

donde $s^2 = \text{CM error}$.

Regla 8 - Si el tratamiento i tiene r_i repeticiones, el error estandar es igual

$$s_z = \sqrt{s^2 \left(\frac{l_1^2}{r_1} + \frac{l_2^2}{r_2} + \dots + \frac{l_k^2}{r_k} \right)}$$

Ejemplos de comparaciones ortogonales

Ejemplo 1 - Un experimento planeado en un diseño de bloques al azar con cinco repeticiones en que se quiere comparar el efecto de diferentes fuentes de fertilizantes nitrogenados en el aumento del rendimiento de un cultivo.

- Las fuentes son:
1. $(\text{NH}_4)_2\text{SO}_4$
 2. $(\text{NH}_4)\text{NO}_3$
 3. Urea
 4. Testigo

Procedimiento

- 1) Se hace un análisis básico de variancia.

FV	GL
Repeticiones	4
Fuentes de N^1	3
Error	12
Total	

1 - Incluye al testigo

2) Podemos descomponer los GL de fuentes de N y la suma de cuadrados del mismo, en tres componentes.

Comparación	Fuentes de N				
	N ₀	Urea	(NH ₄) ₂ SO ₄	(NH ₄)NO ₃	Z
C ₁ = Presencia de N vs ausencia de N	+3	-1	-1	-1	Z ₁
C ₂ = N orgánico vs N Inorgánico	0	+2	-1	-1	Z ₂
C ₃ = (NH ₄) ₂ SO ₄ vs (NH ₄)NO ₃	0	0	+1	-1	Z ₃

$$SC_1 = \frac{z_1^2}{r(l_1^2 + l_2^2 + l_3^2 + \dots + l_n^2)} = \frac{z_1^2}{r \sum_{i=1}^n l_i^2}$$

$$SC_2 = \frac{z_2^2}{r \sum l_i^2}$$

$$SC_3 = \frac{z_3^2}{r \sum l_i^2}$$

3) Finalmente con los componentes calculados se hace la tabla final de variancia.

Forma Funcional	
FV	GL
Repeticiones	4
Fuente de N	3
Presencia de N vs ausencia de N	1
N orgánico vs N Inorgánico	1
(NH ₄) ₂ SO ₄ vs (NH ₄)NO ₃	1
Error	12
Total	19

El ejemplo 1 podemos expresarlo numéricamente en la forma siguiente. Considerando que los datos son de rendimiento de paste per parcelas, que están dados en kilogramos.

Tratamientos	Repeticiones					Total
	I	II	III	IV	V	
T	3	4	2	3	4	16
(NH ₄)NO ₃	5	6	8	4	10	36
(NH ₄) ₂ SO ₄	6	5	7	6	5	29
Urea	8	9	10	10	12	49
	22	24	27	26	31	130

$$FC = \frac{130^2}{20} = 845$$

$$SC \text{ total} = 988 - FC = 143$$

$$SC \text{ tratamiento} = \frac{4794}{5} = FC = 959 - FC = 114$$

$$SC \text{ repeticiones} = \frac{3426}{4} - FC = 11$$

$$\text{Error} = 18$$

Forma básica

FV	GL	SC	CM	F
Repeticiones	4	11	2.75	1.83
Tratamientos	3	114	38.00	25.33
Error	12	18	1.50	
Total	19	143		

Ahora vamos a descomponer la suma de cuadrados de tratamientos en tres contrastes, aplicando las reglas antes mencionadas.

	(T ₁)	(T ₂)	(T ₃)	(T ₄)			
	Testigo	NH ₄ NO ₃	(NH ₄) ₂ SO ₄	Urea	Z	Z ²	Divisores
	16	36	29	49			
Testigo vs Tratados	+3	-1	-1	-1	(Z ₁) -66	4356	5x12 = 60
Inorgánico vs Orgánico	0	+1	+1	-2	(Z ₂) -33	1089	5x 6 = 30
NO ₃ vs SO ₄	0	+1	-1	0	(Z ₃) + 7	49	5x 2 = 10

Estos contrastes son mutuamente ortogonales porque se cumplen las condiciones siguientes:

Condición 1

Condición 2

en Z ₁ :	l _i = 0	C ₁ x C ₂ = 3 x 0 + 2 (-1) + 1 (-1) + (1 x 2) = 0
en Z ₂ :	l _i = 0	C ₁ x C ₃ = 3 x 0 + (-1 x 1) + (-1 x -1) + (-2 x 0) = 0
en Z ₃ :	l _i = 0	C ₂ x C ₃ = 0 x 0 + (1 x 1) + (1 x -1) + (-2 x 0) = 0

y las tres sumas de cuadrados siguientes son componentes de la SC de tratamientos.

$$SC_1 = \frac{Z_1^2}{D} = 72.6 \approx 73$$

$$SC_2 = \frac{Z_2^2}{D} = 36.3 \approx 36$$

$$SC_3 = \frac{Z_3^2}{D} = 4.9 \approx 5$$

Así podemos comprobar que

$$SC \text{ Tratamiento} = SC_1 + SC_2 + SC_3$$

$$114 = 73 + 36 + 5$$

$$114 = 114$$

Hay otros procedimientos para descomponer la SC de tratamientos en componentes individuales, así como el siguiente que es un método directo.

$$SC_1 = \frac{16^2}{5} + \frac{(36 + 29 + 49)^2}{15} - FC_1 = 51.2 + 866.4 - \frac{130^2}{20} = \underline{72.6}$$

$$SC_2 = \frac{(36+29)^2}{10} + \frac{49^2}{5} - FC_2 = 422.5 + 480.2 = 902.7 - \frac{(114)^2}{15} = \underline{36.3}$$

$$SC_3 = \frac{36^2 + 29^2}{5} - FC_3 = \frac{36^2 + 29^2}{5} - \frac{65^2}{10} = 4.9$$

Forma funcional

FV	GL	SC	CM	F
Repeticiones	4	11	2.75	.1.83 NS
Tratamientos	3	114	38	25.33 **
Ausencia de N vs presencia de N	$\left\{ \begin{array}{l} 1 \\ 1 \\ 1 \end{array} \right.$	$\left\{ \begin{array}{l} 73 \\ 36 \\ 5 \end{array} \right.$	$\left\{ \begin{array}{l} 73 \\ 36 \\ 5 \end{array} \right.$	48.67 **
N orgánico vs N inorgánico				24.00 **
SO ₄ vs NO ₃				3.33 NS
Error	12	18	1.50	
Total	19	143		

Ejemplo 2

Un método ligeramente diferente de la división de los grados de libertad es útil en experimentos factoriales. Si por ejemplo se ensayan dos vermicidas a dos concentraciones y su efectividad para combatir vermes:

Característica del ensayo

Vermicida = 2 = A y B

Dosis = 2 = a_1, a_2 y b_1, b_2

10 animales por niveles

Total animales = 40

Tratamiento	Repetición										
	I	II	III	IV	V	VI	VII	VIII	IX	X	
A	d_1	Y_{11}	Y_{12}	-	-	-	-	-	-	-	$Y_{1r} = T_1$
	d_2	Y_{21}	-	-	-	-	-	-	-	-	$Y_{2r} = T_2$
B	d_1	Y_{31}	-	-	-	-	-	-	-	-	$Y_{3r} = T_3$
	d_2	Y_{41}	-	-	-	-	-	-	-	-	$Y_{4r} = T_4$
		Y_{k_1}	Y_{k_2}							Y_{kr}	

Plan básico de análisis

FV	GL
Tratamiento	3
Error	36
Total	39

Comparaciones entre Tratamientos

	Vermicida A		Vermicida B		Z
	d ₁	d ₂	d ₁	d ₂	
C ₁ = Entre vermicida	-1	-1	+1	+1	Z ₁
C ₂ = Entre dosis	-1	+1	-1	+1	Z ₂
C ₃ = Interacción	+1	-1	-1	+1	Z ₃

Ejemplo 3

Si por ejemplo se estudia la acción de seis fungicidas en cinco repeticiones. Tres de ellos son compuesto cupricos y tres azufre. El análisis puede hacerse de la siguiente manera:

Forma básica	
FV	GL
Repeticiones	4
Fungicidas	5
Error	20
Total	29

Forma funcional	
FV	GL
Repeticiones	4
Fungicida	5
a) Entre grupos	{ 1
b) Dentro de grupos	
i Entre cu	{ 2
Entre S	
Error	20
Total	29

Ejemplo 4

Si desea conocer la concentración de Fe en una solución por 10 analistas diferentes, cinco mujeres y cinco hombres, cada uno hace cinco determinaciones de la solución Fe. El modelo de análisis a seguir podría ser como sigue:

Forma básica	
FV	GL
Entre analista	9
Entre muestras	40

Forma funcional	
FV	GL
Entre analista	9
(a) hombre vs mujer	} 1 8
(b) dentro de grupos	
(i) entre hombres	} 4 4
(ii) entre mujeres	
Entre muestras	40
Total	49

Ejemplo 5

Se quiere probar o comparar la eficiencia de cuatro tipos de ordeñadoras mecánicas. Las repeticiones son representadas por los días en que la prueba se repite con el grupo de cuatro máquinas.

Máquina A - corriente

- " B - modelo experimental desarrollado para reemplazar a A
- " C - modelo experimental desarrollado para reemplazar a
- " D - modelo experimental desarrollado para reemplazar a

	Ordeñadoras					Z
	A	B	C	D	Z	
$C_1 = A$ vs otras	+3	-1	-1	-1		Z_1
$C_2 = B$ y C vs D	0	+1	+1	-2		Z_2
$C_3 = B$ vs C	0	+1	-1	0		Z_3

Comparaciones no ortogonales

Frecuentemente es deseable diseñar un experimento para hacer ciertas comparaciones que no sean ortogonales. En tales situaciones la probabilidad establecida en el nivel de significancia (PO.01 PO.05) no es exactamente discernible. En estas situaciones debe hacerse una interpretación cuidadosa de los datos.

Debido a la naturaleza no ortogonal de estos tipos de análisis, la suma de cuadrado de las comparaciones individuales sumada no son iguales a la suma de cuadrados de tratamientos de la forma básica. Es decir que no se cumple el teorema de la aditividad.

Ejemplo 1. (Nutrición)

Se trata de la ganancia diaria de peso de oveja, alimentada con cuatro raciones diferentes. La ración una es una ración estandar de proteína que es lo que se quiere comparar con las tres restantes.

La comparación (no ortogonal) de cada ración con el testigo podría ser simbólicamente representada como sigue:

Comparación	Raciones				Z
	1	2	3	4	
$C_1 = 1 \text{ vs } 2$	+1	-1	0	0	Z_1
$C_2 = 1 \text{ vs } 3$	+1	0	-1	0	Z_2
$C_3 = 1 \text{ vs } 4$	+1	0	0	-1	Z_3

Del presente cuadro vemos que la suma de los coeficientes de cada una de las comparaciones es igual a cero, pero la suma de productos en cada par de comparaciones es mayor que cero. Ejemplo:

$$C_1 \times C_2 = (+1) + (+1) + (-1)(0) + (0)(-1) + (0)(0) = 1$$

Esto nos indica que estas comparaciones no son independientes o sea no ortogonales, pero éllas dan la respuesta a las preguntas para cuyo propósito el experimento fue diseñado.

Otro ejemplo: Se estudia la evolución del CO_2 proveniente de hongos del suelo y se incorporaron ciertos aditivos en el sustrato sobre el cual crecen.

Los tratamientos son:

- 1 - testigo
- 2 - harina de alfalfa
- 3 - harina de tusa de maíz
- 4 - harina de tusa de maíz + $MgSO_4$

Comparación	Tratamiento				Z
	1	2	3	4	
$C_1 = 1 \text{ vs } 2$	+1	-1	0	0	Z_1
$C_2 = 2 \text{ vs } 3$	0	+1	-1	0	Z_2
$C_3 = 3 \text{ vs } 4$	0	0	+1	-1	Z_3

Comparaciones de tendencias

Es muy frecuente que el investigador esté interesado en estudiar una variable a varios niveles, tales como niveles de temperatura, concentraciones de productos químicos, niveles de alimento en una ración, niveles de un nutrimento aplicado al suelo; en tales casos el investigador estaría interesado en la naturaleza de la tendencia de respuesta, a la variación de niveles. La tendencia de la respuesta (variable dependiente) con el aumento de niveles (variable independiente) puede ser lineal o curvilínea. Estimar la tendencia de la relación entre causa y efecto, es más importante conocer, que hacer una prueba de significancia entre los promedios de las variables dependientes, en muchos casos de la investigación.

Debe recordarse que como sea posible los niveles de las variables independientes deberían ser igualmente espaciados o de igual intervalo. Los espacios iguales son aplicables a la escala natural o logarítmica o a algunas otras transformaciones simples. Ejemplo de igual espaciamiento en escala natural y logarítmica.

Escala		Niveles de la variable independiente							Razón (r)
Natural	\div	0,	2	4,	6,	8,	10,	etc.	2
Logarítmica	$\div\div$	1,	2,	4,	8,	16,	32,	etc.	2

Un procedimiento fácil seguido para estimar la tendencia de la respuesta, es por medio del uso de los polinomios ortogonales. Este procedimiento es especialmente válido si la tendencia deseada es de grado superior al polinomio de segundo grado (cuadrático).

En este método de análisis de regresión por medio de polinomios ortogonales, si la variable independiente está igualmente espaciada (sea en el tiempo o en el espacio) el método conveniente de ajuste de la curva por polinomios ortogonales puede ser usado.

La ventaja del método de regresión por medio de polinomios ortogonales sobre el método usual de ajuste por polinomios no ortogonales se desprende del hecho de que los polinomios ortogonales son contruidos en tal forma que cualquier término del polinomio es independiente de cualquier otro. Esta propiedad de independencia permite la comparación de cada coeficiente de regresión independiente de los otros y también facilita la prueba de significancia de cada uno de ellos.

En la práctica el método (ξ) es de uso corriente, por ser fácil, rápido y expeditivo, si se desea la prueba de significancia de los términos sucesivos. Este método requiere tener una tabla de los valores de ξ (ver tabla de los coeficientes de polinomios ortogonales).

Ejemplo 1. El experimento se refiere a abonamiento de maíz con P_2O_5 a niveles de 0, 25, 50, 75, 100 kg./ha. de P_2O_5 llevado a cabo en bloques al azar con cuatro repeticiones.

Repeticiones	0	25	50	75	100	Total
I	3.38	7.15	10.07	9.55	9.14	39.29
II	5.77	9.78	9.73	8.95	10.17	44.40
III	4.90	9.99	7.92	10.24	9.75	42.80
IV	4.54	10.10	9.48	8.66	9.50	42.28
Total	18.59	37.02	37.20	37.40	38.56	168.77

Análisis preliminar

FV	GL	SC	CM	F
Repetición	3	2.73		
Tratamiento	4	72.22	18.055	19.84**
Error	12	10.92	0.910	
Total	19	85.87		

Ahora podemos descomponer los cuatro GL del tratamiento en cuatro componentes: 1° - componente lineal o sea si la producción es función lineal de los niveles de P_2O_5 ; 2° - función cuadrática; 3° - función cúbica y 4° - función cuarta o bicuadrática. Haremos uso de los coeficientes de polinomios ortogonales (de la tabla).

Tratamientos

	18.59	37.02	37.20	37.40	38.56	Z	D	$\frac{Z^2}{D}$
F. Lineal	-2	-1	0	+1	+2	40.32	4x10	40.64
F. Cuadrática	+2	-1	-2	-1	+2	-34.52	4x14	21.28
F. Cúbica	-1	+2	0	-2	+1	19.21	4x10	9.23
F. Cuarta	+1	-4	+6	-4	+1	-17.33	4x70	1.07

Análisis completo

FV	GL	SC	CM	F
Repeticiones	3	2.73	0.910	1.00 NS
Tratamiento	4	72.22		
F. Lineal	1	40.64	40.64	44.66**
F. Cuadrático	1	21.28	21.28	23.38**
F. Cúbica	1	9.23	9.23	10.14**
F. Cuarta	1	1.07	1.07	1.18 NS
Error	12	10.92	0.910	
Total	19			

Interpretación. Los componentes de 1º, 2º y 3º grado son significativos. Para decir cual es la ecuación de regresión que corresponde a nuestros datos, tenemos que calcular los coeficientes de cada una de las funciones que resultaron significativas, lo mismo que en este intervalo haya algún componente no significativo.

$$\text{F. Lineal} \quad - - B_1 = \frac{\sum_{l=1}^n Z}{\sum_{l=1}^n l^2} = \frac{40.32}{4 \times 10} = 1.01$$

$$\text{F. Cuadrática} \quad - - B_2 = \frac{-34.52}{4 \times 14} = -0.616$$

$$\text{F. Cúbica} \quad - - B_3 = \frac{19.21}{4 \times 10} = 0.480$$

Luego calculamos la media de X, y de Y

$$\bar{y} = \frac{\sum y_i}{n} = \frac{168.77}{20} = 8.44$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{250}{5} = 50$$

La ecuación de regresión, corresponde a la forma

$$\hat{y} - \bar{y} = B_1 M_1 P_1 + B_2 M_2 P_2 + B_3 M_3 P_3$$

Donde $P_1 P_2 - - P_3$ indican los polinomios dados por la tabla, $M_1 M_2$

- - M_3 son valores constantes también de la tabla; entonces tenemos

$$\hat{y} - 8.44 = 1.01 \times 1 \times P_1 - 0.616 \times 1 \times P_2 + 0.480 \left(\frac{5}{6}\right) P_3$$

Donde : $P_1 = X$

$$P_2 = X^2 - \frac{n^2 - 1}{12} = X^2 - 2$$

$$P_3 = X^3 - \frac{3n^2 - 7}{20}(X) = X^3 - \frac{17}{5}X$$

Si tuviéramos P_4 y P_5 usaríamos:

$$P_4 = X^4 - \frac{3n^2 - 13}{14}(X^2) + \frac{3(n^2 - 1)(n^2 - 9)}{560}$$

$$P_5 = X^5 - \frac{5(n^2 - 7)}{18}(X^3) + \frac{15n^4 - 230n^2 + 407}{1.008}(X)$$

n = número de niveles.

Entonces,

$$\hat{y} = 8.44 + 1.01X - 0.616 (X^2 - 2) + 0.400 (X^3 - \frac{17}{5}X)$$

$$\hat{y} = 9.672 - 0.350X - 0.616X^2 + 0.400X^3$$

La variable auxiliar X está dada por la ecuación

$$X = \frac{x - \bar{x}}{q}$$

Donde q es el intervalo o sea espacio de dosis a dosis, en nuestro caso $q = 25$

$$X = \frac{x - 50}{25} = \frac{x}{25} - 2$$

Luego reemplazando en la fórmula tenemos

$$\hat{y} = 9.672 - 0.350 \left(\frac{x-2}{25}\right) - 0.616 \left(\frac{x-2}{25}\right)^2 + 0.400 \left(\frac{x-2}{25}\right)^3$$

Por simplificación nos da la ecuación polinomial siguiente

$$\hat{y} = 4.708 + 0.277x - 0.00483x^2 + 0.0000256x^3$$

Regresión polinomial aplicado a datos sin repetición

En muchos experimentos no hay posibilidad de recoger con repeticiones los datos: por ejemplo datos meteorológicos de una determinada localidad.

Si desean analizar estos datos sin posibilidad de una estimación correcta del error. Esto se puede hacer con justificación aceptable utilizando como residuo el cuadrado medio de la desviación de regresión.

Por ejemplo tenemos datos de temperatura media en varios años.

Años (x)	Temperatura (y)
1941	24.5
1942	22.0
1943	25.8
1944	24.4
1945	24.6
1946	24.0
1947	23.0
1948	26.1
1949	25.7
1950	25.5
1951	24.5
1952	25.5
1953	25.0
1954	27.3
1955	25.7

Podemos plantear el siguiente análisis.

N	Temperatura Y	Años X	Coefficientes C ₁ lineal	Coefficientes C ₂ analítico
1	24.5	1	-7	+91
2	22.0	2	-6	+52
3	25.8	3	-5	+19
4	24.4	4	-4	-8
5	24.6	5	-3	-29
6	24.0	6	-2	-44
7	23.0	7	-1	-53
8	26.1	8	0	-56
9	25.7	9	+1	-53
10	25.5	10	+2	-44
11	24.5	11	+3	-29
12	25.5	12	+4	-8
13	25.0	13	+5	+19
14	27.3	14	+6	+52
15	25.7	15	+7	+91

$$SC \text{ total } (\sum Y_i^2 - FC) = \underline{23,510}$$

$$SCF \text{ lineal} = \frac{\sum Z_1^2}{\sum l_i^2} = 7.560$$

$$SCF \text{ cuadrática} = \frac{\sum Z_2^2}{\sum l_i^2} = 0.076$$

FV	GL	SC	CM	F
Regresión lineal	1	7.560	7.560	7.05*
Regresión cuadrática	1	0.076	0.076	0.071 NS
Desviación de la regresión	12	15.874	1.073	
Total	14	23.510		

Esto nos indica que existe una relación lineal entre años y temperatura. Ahora vamos a calcular los coeficientes respectivos.

$$B_1 = \frac{Z_1}{\sum 1i^2} = \frac{46.0}{280} = 0.164$$

La ecuación de regresiones

$$\hat{Y} = \bar{Y} + B_1 x$$

$$\hat{Y} = 24.907 + 0.164 x$$

$$x = \frac{X-8}{1}$$

$$Y = 24.907 + 0.164 (X-8)$$

$$Y = 23.595 + 0.164 X$$

Esto nos indica que la temperatura media aumenta casi linealmente a razón de 0.0164 grados por año.

PRUEBA DE COMPARACION MULTIPLE

El análisis de variancia (F - test) indica si existe o no diferencias significativas entre un grupo de promedios. Sin embargo una no significancia no implica que los efectos verdaderos no existen. De la misma manera un valor significativo de F no indica que un

promedio "A" difiera significativamente de algún otro promedio.

En algunos experimentos no hay mucha razón de partir los grados de libertad dentro de la clase o hacer comparaciones de tendencias, pero el investigador puede estar interesado en conocer cual promedio debe ser considerado diferente de otros.

Un gran número de procedimientos ha sido desarrollado para probar la diferencia entre varios promedios. Se puede hacer una clasificación en lo siguiente:

1. Pruebas de rango múltiples (multiple range tests)
 - a) Prueba de diferencia mínima significativa o prueba de "t" múltiple.
 - b) Prueba de Student - Newman-Keuls.
 - c) Prueba de rango múltiple de Duncan.
 - d) Prueba de Tukey.
2. Pruebas múltiples de F (multiple F - tests)
 - a) Prueba de Scheffe basadas en contrastes.
 - b) Pruebas de comparaciones múltiples de Duncan.

La prueba de diferencia mínima significativa (least significant difference) o llamada también prueba múltiple de "t" o "l.s.d.'test"

Esta prueba se usa comúnmente para comparar cada uno de los promedios con un tratamiento estandar. Este valor es igual a $t_{0.05} \times$

$\sqrt{\frac{2s^2}{r}}$ a veces se usa en vez de $t_{0.05}$ el valor de $t_{0.01}$, al valor

$t_{1\%} \sqrt{\frac{2s^2}{r}}$ se le llama a veces m.sd (most significant difference).

La prueba de l.s.d. es apropiada si las comparaciones a hacer, se seleccionan antes de llevar a cabo el experimento. El caso más extremo es la comparación del promedio más alto con el más bajo entre "v" tratamientos, en este caso Pearson y Hartley han demostrado que el error Tipo I de los v tratamientos no es 5% para v = 2 sino es mayor. Se ha demostrado que el valor observado de "t" para tres tratamientos excede al valor tabular al 5% cerca del 13% de veces; con seis tratamientos alrededor del 40%, con 10 cerca del 60% y con 20 cerca del 90%. Así que cuando el investigador piensa que está tomando t al 5% para tres tratamientos, realmente está probando al 13% (ver ejemplo).

La magnitud del Error Tipo I de los tratamientos es igual a $1 - \frac{1}{f} P_n(Q)$ donde $\frac{1}{f} P_n(Q)$ es igual a $\frac{1}{f} P_n(Q) = P_n(Q) + \frac{1}{f} a_n Q + \frac{1}{f^2} b_n(Q)$ donde $P_n(Q)$, $a_n(Q)$ y $b_n(Q)$ se obtienen de la tabla de Pearson y Hartley.

n = v = número de tratamientos

f = grado de libertad del error

$$Q = \sqrt{2} \cdot t_{\alpha}$$

Muy frecuentemente y especialmente con datos de animales, se aplica la prueba de l.s.d a la diferencia entre promedios que provienen de números desiguales de observaciones. En estos casos se puede aplicar la fórmula

$$\text{l.s.d.} = t_{.05} \sqrt{\frac{s^2}{r_1} + \frac{s^2}{r_2}}$$

Test de Student-Newman-Keuls

Cada una de las tres personas nombradas contribuyeron a crear un procedimiento para la prueba de rango múltiple. Este método es relativamente fácil de aplicar y consiste en:

1. Calcular el error estandar de promedio $s\bar{x} = \sqrt{\frac{s^2}{r}}$
2. Escoger el nivel de significancia " α " (usualmente es PO.05 o PO.01).
3. Escoger el valor tabular (Studentized ranges) de $q_2, q_3 \dots q_n$ según el número de tratamiento que se desea comparar.
4. Calcular el valor W 's = $q'_s \times s\bar{x}$; o sea $W_n = q_{\alpha, n} s\bar{x}$, $W_{n-1} = q_{\alpha, n-1} s\bar{x}$ etc.
5. Comparar el rango de n tratamientos $\bar{X}_n = \bar{X}_1$ con el valor calculado de W_n . Si $\bar{X}_n - \bar{X}_1$ es menor que W_n se suspende el proceso y se acepta que los promedios pertenecen a grupos no diferentes. Si $\bar{X}_n - \bar{X}_1 > W_n$ se subdividen los promedios en dos grupos de $n - 1$ promedios cada uno de \bar{X}_n a \bar{X}_2 y \bar{X}_{n-1} a \bar{X}_1 y se observa si \bar{X}_n es diferente de \bar{X}_1 . Entonces comparar los rangos $\bar{X}_n - \bar{X}_2$ y $\bar{X}_{n-1} - \bar{X}_1$ con W_{n-1} . Si cualquiera de los rangos es menor que W_{n-1} se dice que los promedios en el grupo pertenecen a un grupo simple. Si ambos rangos exceden a W_{n-1} de $n-1$ promedios se divide en dos grupos de $n-2$ promedios cada uno y compararlo con W_{n-2} . El proceso continúa así hasta un paso en que el promedio no excede al valor calculado de W_i (ver ejemplo).

Si dos promedios que deseamos comparar tuvieran diferentes números de repeticiones, se puede usar la fórmula aproximada de

$$W = qx \sqrt{\frac{1}{2} \left(\frac{s^2}{r_1} + \frac{s^2}{r_2} \right)}$$

Prueba de rango múltiple de Duncan (Duncan's multiple range test)

La prueba de Duncan combina la simplicidad del método de Student-Newman-Keuls y las ventajas de las comparaciones múltiples. Este método es mejor que los dos propuestos anteriormente y permite al investigador cometer menor "Error tipo II" pero mayor "Error tipo I" que los métodos explicados.

El método de Duncan es esencialmente lo mismo que el de Student-Newman-Keuls excepto que el valor tabular se obtiene de otra tabla de amplitud total estudentizada para prueba de Duncan.

Los pasos a seguir son los siguientes:

1. = Calcular $s\bar{x} = \sqrt{\frac{s^2}{r}}$
2. = Obtener los valores de "z" de la tabla, entrando con los GL del error y n, n-1, n-2 - - - - 2; promedios ordenados de acuerdo a su magnitud.
3. = Calcular los valores de $D = z \times s\bar{x}$
4. = Comparar $\left\{ \begin{array}{l} D_n \text{ con } \bar{X}_n - \bar{X}_1 \\ D_{n-1} \text{ con } \bar{X}_n - \bar{X}_2 \\ D_{n-1} \text{ con } \bar{X}_{n-1} - \bar{X}_2 \text{ etc. y seguir así sucesivamente formando grupos de igual amplitud.} \end{array} \right.$

Si en un grupo de promedios el mayor no difiere significativamente

de el menor, no admite diferencia significativa entre promedios intermedios. Más detalles se encontrarán en el ejemplo numérico.

Si el número de repeticiones de los promedios que se desea comparar no son las mismas se puede usar la fórmula aproximada del Error estandar del promedio.

$$s\bar{x} = \sqrt{\frac{1}{2} \left(\frac{s^2}{r_1} + \frac{s^2}{r_2} \right)}$$

Prueba de Tukey (Tukey's test based on allowances or w-procedure)

Este método comúnmente llamado hsd (honestly significant difference) es similar en su aplicación a la prueba de lsd. El método puede ser usado para comparar todos o cualquier contraste entre los promedios de tratamientos. La Prueba de Tukey es muy exigente, fija una probabilidad de α -por ciento de tomar como significativa una diferencia que realmente es nula entre todos los promedios de los tratamientos. Es decir que el Error Tipo I de los v - tratamientos se fija al nivel de α por ciento.

Los pasos a seguir en este método son:

1. Calcular $s\bar{x} = \sqrt{\frac{s^2}{r}}$
2. Obtener el valor de "q" de la tabla de amplitud total estudentizada con grado de libertad del Error y número de tratamiento.
3. Calcular el valor $\Delta = q \times s\bar{x}$
4. Comparar el valor de Δ con cualquier diferencia entre dos promedios (Ver ejemplo numérico).

En caso que los promedios a comparar tengan desigual número de repeticiones se puede estimar el Error estandar del promedio con la fórmula

$$s\bar{x} = \sqrt{\frac{1}{2} \left(\frac{s^2}{r_1} + \frac{s^2}{r_2} \right)}$$

Comentario general sobre las pruebas de rango múltiple

El uso de la prueba l_{sd} conduce a un error Tipo I más grande que $\alpha = 5\%$ cuando el promedio más grande se compara con el más pequeño, si el número de tratamiento es mayor de 2. Para saldar este defecto se han propuesto las otras pruebas que hemos discutido anteriormente y otras que no las discutimos por falta de tiempo.

La escogencia del tipo de prueba a usar de entre las que vimos, depende del tamaño del Error Tipo I deseado. La prueba de Duncan tiene mayor Error Tipo I pero menor Error Tipo II que las otras dos. La prueba de Student-Newman-Keuls tiene mayor Error Tipo I y menor Error Tipo II que la de Tukey.

Comentarios sobre la prueba de Duncan y de Tukey

Las dos pruebas tienen fundamentos muy semejantes, la de Duncan es menos exigente y menos conservadora, esto es que da diferencias significativas con mayor facilidad. El criterio del autor, es que toma el nivel de significancia del 5% de probabilidad en un contraste de dos promedios. Esto da una probabilidad del 95% (0.95) de no tomar como significativa una diferencia realmente nula, ya que para tres promedios la probabilidad será $0.9025 = (0.95)^2$ o 90.25%, para

cuatro promedios = $0.8574 = (0.95)^3 = 85.74\%$. En general para n promedios la probabilidad será $(0.95)^{n-1}$. La prueba de Tukey es más exigente, teniendo una probabilidad de 95% de no tomar como significativa una diferencia realmente nula, entre todos los promedios de tratamientos, no es de sorprenderse pues, que con $n > 2$ la prueba de Duncan de resultados significativos, en casos en que esto no ocurra con la de Tukey, y entonces la prueba de Duncan nos lleva a afirmaciones erradas con mayor frecuencia; lo mismo acontece y con mayor peligro con la prueba de LSD aplicada indiscriminadamente y que hoy ya no se acepta.

La prueba de Duncan establece pues un término medio entre el rigor un tanto excesivo de la de Tukey y la poca exigencia de la prueba de LSD usada sin precaución.

La prueba Sheffe es aun más rigurosa y no se aconseja para una comparación de dos promedios, pero funciona muy bien para probar contrastes más complicados, y es el método indicado para estos casos.

Interpolación armónica

A veces los valores buscados no se encuentran en la tabla (t , v , F , o Z), los valores se pueden calcular por interpolación armónica, en que se usan las recíprocas de los números de GL para formar una regla de tres.

Ejemplo 1. Cual es el límite al 1% de t , correspondiente a 36 GL.

En la tabla de t encontramos los siguientes valores

para 30 GL $t = 2.75$

para 40 GL $t = 2.70$

Cálculo

$$\frac{1}{30} - \frac{1}{40} = 2.75 - 2.70; \frac{1}{120} = 0.05$$

$$\frac{1}{30} - \frac{1}{36} = 2.75 - x ; \frac{1}{180} = 2.75 - x$$

$$2.75 - x = \frac{\frac{1}{180} (0.05)}{\frac{1}{120}}$$

$$x = 2.75 - 0.03 = 2.72$$

$$t \text{ para GL } 36 = 2.72$$

Ejemplo 2. Calcular el valor de v para $n_1 = 12$ GL y $n_2 = 200$ GL al nivel del 1% de probabilidad.

Cálculo

$$n_1 = 12 \text{ y } n_2 = 120 \text{ ————— } 1.53$$

$$n_1 = 12 \text{ y } n_2 = \infty \text{ ————— } 1.48$$

$$\frac{1}{120} - \frac{1}{\infty} = 1.53 - 1.48; \frac{1}{120} \text{ ————— } 0.05$$

$$\frac{1}{120} - \frac{1}{200} = 1.53 - x ; \frac{1}{300} \text{ ————— } 1.53 - x$$

$$1.53 - x = \frac{\frac{1}{300} (0.05)}{120}$$

$$x = 1.53 - 0.02 = 1.51$$

$$v \text{ cu GL } 12 \text{ y } 200 = 1.51$$

EJEMPLO DE APLICACION DE LAS PRUEBAS DE RANGO MULTIPLES

Masa Nº	Gramos de grasa absorbidos por la masa de 24 buñuelos								
	TRATAMIENTOS								
	1	2	3	4	5	6	7	8	
I	164	173	177	178	163	175	178	155	
II	172	161	183	191	165	193	146	166	
III	168	190	197	197	144	178	141	149	
IV	177	180	169	182	177	171	150	164	
V	156	197	179	185	165	163	169	170	
VI	195	167	187	177	176	176	182	168	
Total	1032	1068	1092	1110	990	1056	966	972	8286
\bar{X}	172	178	182	185	165	176	161	162	

Análisis de Variancia

FV	GL	SC	CM	F
Entre Grasas	7	3.527	503.9	3.56**
Dentro de Grasas	40	5.666	141.6	
Total	47	9.193		

El análisis estadístico nos indica que existe marcada diferencia en la cantidad de grasa absorbida por la masa, dependiendo este de la clase de grasa. Ahora estamos interesados en averiguar que

diferencia promedio existe entre las diferentes grasas. A continuación ilustramos los cuatro métodos para comparar promedios, que hemos discutido anteriormente.

Prueba de lsd (least significant difference)

Pasos a seguir:

1. $s\bar{d} = \sqrt{\frac{2 s^2}{r}} = \sqrt{\frac{2 \times 141.6}{6}} = 6.8702$
2. $t_{0.05}$ con GL = 40 = 2.021
3. $lsd = t_{0.05} \times s\bar{d} = 2.021 \times 6.8702 = 13.885 \approx \underline{13.9}$
4. Se calculan las diferencias entre promedios de tratamientos y se comparan con 13.9

		4	3	2	6	1	5	8
$\bar{X} \backslash \bar{X}$		185	182	178	176	172	165	162
7	161	24	21	17	15	11	4	1
8	162	23	20	16	14	10	3	x
5	165	20	17	13	11	7	x	x
1	172	13	10	6	4	x	x	x
6	176	9	6	2	x	x	x	x
2	178	7	4	x	x	x	x	x
3	182	3	x	x	x	x	x	x

Interpretación Gráfica
Promedios (ordenado de menor a mayor)

Tratamiento	7	8	5	1	6	2	3	4
\bar{Y}	161	162	165	172	176	178	182	185

Prueba de Student-Newman-Keuls

Pasos a seguir

1. $s\bar{x} = \sqrt{\frac{s^2}{r}} = \sqrt{\frac{141.6}{6}} = \underline{4.86}$

2. Obtener los valores de q 0.05 o q 0.01 con grados de libertad del error y n -promedios de tratamientos ordenados de mayor a menor o viceversa. Se calcularán tanto valores de q como $n-1$ tratamientos haya. Estos valores se consiguen en la tabla de amplitud total estudentizada (Studentized range).

En nuestro caso los valores tabulares de q 0.05 con 40 GL y $n=8$

---- $n=2$ son: $q_8 = 4.52$; $q_7 = 4.39$; $q_6 = 4.23$; $q_5 = 4.04$;
 $q_4 = 3.79$; $q_3 = 3.44$; $q_2 = 2.86$.

3. Cálculo de los comparadores $W_n = q_n \times s\bar{x}$

$W_8 = 4.52 (4.86) = 22.0$

$W_7 = 4.39 (4.86) = 21.3$

$W_6 = 4.23 (4.86) = 20.6$

$W_5 = 4.04 (4.86) = 19.6$

$W_4 = 3.79 (4.86) = 18.4$

$W_3 = 3.44 (4.86) = 16.7$

$W_2 = 2.86 (4.86) = 13.9$

4. Cálculo de la diferencia entre los promedios de tratamientos

		4	3	2	6	1	5	8
$\bar{X} \backslash \bar{X}$		185	182	178	176	172	165	162
7	161	24	21	17	15	11	4	1
8	162	23	20	16	14	10	3	x
5	165	20	17	13	11	7	x	x
1	172	13	10	6	4	x	x	x
6	176	9	6	2	x	x	x	x
2	178	7	4	x	x	x	x	x
3	182	3	x	x	x	x	x	x

Interpretación del cuadro

$$\begin{cases} \bar{X}_4 - \bar{X}_7 = 185 - 161 = 24 > W_8 = 22.0 \\ \bar{X}_4 - \bar{X}_8 = 185 - 162 = 23 > W_7 = 21.3 \\ \bar{X}_3 - \bar{X}_7 = 182 - 161 = 21 < W_7 = 21.3 \\ \bar{X}_4 - \bar{X}_5 = 185 - 165 = 20 < W_6 = 20.64 \text{ El proceso termina aquí} \\ \bar{X}_3 - \bar{X}_8 = 182 - 162 = 20 < W_6 = 20.6 \\ \bar{X}_2 - \bar{X}_7 = 178 - 161 = 17 < W_6 = 20.6 \end{cases}$$

Interpretación Gráfica de los resultados

Promedios (ordenados de menor a mayor)

Trata- miento	7	8	5	1	6	2	3	4
\bar{X}	161	162	165	172	176	178	182	185

Prueba de Duncan

Pasos a seguir:

1. $s\bar{x} = \sqrt{\frac{141.6}{6}} = 4.86$

2. Se obtienen los valores tabulares de la tabla (new multiple range test) con n_2 GL del error y p número de promedios.

Z con (40:8) = 3.30

Z con (40:7) = 3.27

Z con (40:6) = 3.22

Z con (40:5) = 3.17

Z con (40:4) = 3.10

Z con (40:3) = 3.01

Z con (40:2) = 2.86

3. Cálculo de los comparadores

$D_8 = 3.30 \times 4.86 = 16.0$

$D_7 = 3.27 \times 4.86 = 15.9$

$D_6 = 3.22 \times 4.86 = 15.6$

$D_5 = 3.17 \times 4.86 = 15.4$

$D_4 = 3.10 \times 4.86 = 15.1$

$D_3 = 3.01 \times 4.86 = 14.6$

$D_2 = 2.86 \times 4.86 = 13.9$

4. Cálculo de la diferencia entre los promedios de tratamientos

		4	3	2	6	1	5	8
$\bar{X} \backslash \bar{X}$		185	182	178	176	172	165	162
7	161	24	21	17	15	11	4	1
8	162	23	20	16	14	10	3	x
5	165	20	17	13	11	7	x	x
1	172	13	10	6	4	x	x	x
6	176	9	6	2	x	x	x	x
2	178	7	4	x	x	x	x	x
3	182	3	x	x	x	x	x	x

Interpretación del cuadro

$$\bar{X}_4 - \bar{X}_7 = 185 - 161 = 24 > D_8 = 16.0$$

$$\bar{X}_4 - \bar{X}_8 = 165 - 162 = 23 > D_7 = 15.9$$

$$\bar{X}_3 - \bar{X}_7 = 182 - 161 = 21 > D_7 = 15.9$$

$$\bar{X}_4 - \bar{X}_5 = 185 - 165 = 20 > D_6 = 15.6$$

$$\bar{X}_3 - \bar{X}_8 = 182 - 162 = 20 > D_6 = 15.6$$

$$\bar{X}_3 - \bar{X}_8 = 182 - 162 = 20 > D_6 = 15.6$$

$$\bar{X}_2 - \bar{X}_7 = 178 - 161 = 17 > D_6 = 15.6$$

$$\bar{X}_4 - \bar{X}_1 = 183 - 172 = 13 < D_5 = 15.4$$

$$\bar{X}_3 - \bar{X}_5 = 182 - 165 = 17 > D_5 = 15.4$$

$$\bar{X}_3 - \bar{X}_5 = 182 - 165 = 17 > D_5 = 15.4$$

$$\bar{X}_2 - \bar{X}_8 = 178 - 162 = 16 > D_5 = 15.4$$

$$\bar{X}_2 - \bar{X}_8 = 178 - 162 = 16 > D_5 = 15.4$$

$$\bar{X}_6 - \bar{X}_7 = 176 - 161 = 15 < D_5 = 15.4$$

$$\bar{X}_2 - \bar{X}_5 = 178 - 165 = 13 < D_4 = 15.1$$

Interpretación Gráfica de los resultados

Promedios (ordenado de menor a mayor)

Tratamiento	7	8	5	1	6	2	3	4
\bar{X}	161	162	165	172	176	178	182	185

Prueba de Tukey

Pasos a seguir:

1. $s\bar{x} = \sqrt{\frac{141.6}{6}} = 4.86$

2. $q_8 = 4.52$ (obtenido de la tabla de amplitud total estudentizada) con GL = 40 y n = 8

3. $W = s\bar{x} q_n = 4.52 \times 4.86 = 22.0$

hsd = 22.0

4. El valor de hsd se compara con la diferencia de los promedios de cualquiera de los dos tratamientos, si este excede al valor de hsd se dice que son significativas, y a la inversa se dice que no son significativas. En nuestro ejemplo solo son significativas.

$$\bar{X}_4 - \bar{X}_7 = 185 - 161 = 24 > hsd = 22.0 \text{ y también}$$

$$\bar{X}_4 - \bar{X}_8 = 185 - 162 = 23 > hsd = 22.0$$

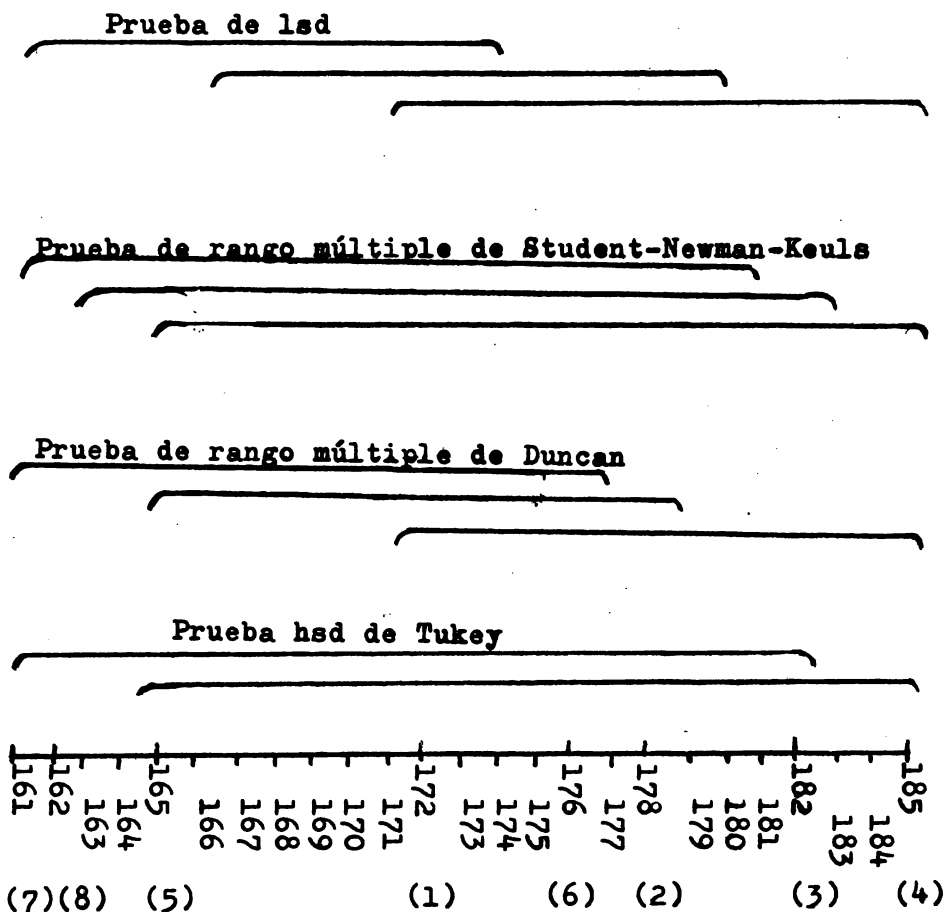
Como se podrá observar en el cuadro anterior de diferencia entre promedios, ninguna excepto las dos que se mencionan aquí fueron significativas.

Interpretación gráfica

Promedios (ordenado de menor a mayor)

Tratamiento	7	8	5	1	6	2	3	4
\bar{X}	161	162	165	172	176	178	182	185

Resumen de los resultados obtenidos por los cuatro métodos.



PARTE ESPECIAL

Análisis de variancia con números desiguales en la sub-clase

En ciertos campos de la investigación, no siempre es posible tener un número igual de observaciones en todas las combinaciones de tratamientos; especialmente en producción animal, es conveniente prestar atención al problema que plantea la desigualdad del número de observaciones en la sub-clase.

El análisis estadístico de la variancia corrientemente usado, pierde validez en muchos casos y necesita ser modificado para tener una estimación correcta de los componentes de variación.

Los datos con números desiguales de observaciones se pueden clasificar en tres grupos (A, B, C).

A - En este grupo se incluyen los experimentos donde se estudia un solo factor con desigual número de observaciones en cada tratamiento. Generalmente se les llama a estos datos con un criterio de clasificación (one-way classification).

B - Dentro de este grupo caen los experimentos donde se estudian dos factores (factorial) con número de observaciones en la sub-clase, desiguales pero proporcionales. A estos datos se les llaman con dos criterios de clasificación (two-way classification).

C - En este grupo se encuentran los experimentos con más de un factor en estudio (factorial), con número desigual y desproporcionales en la sub-clase.

A continuación presentamos el método o los métodos de análisis empleados según el caso.

1. Un solo factor en estudio con números desiguales en la sub-clase

Los datos de una sola clasificación no presentan ninguna dificultad para su análisis. Los componentes de variación se pueden estimar de la siguiente manera.

$$FC = \frac{(\sum_{ij} y_{ij})^2}{n_i}$$

$$SC \text{ total} = \sum_i y_{ij}^2 - FC$$

$$SC \text{ trat} = \sum_{\lambda} \frac{y_{\lambda}^2}{n_{\lambda}} - FC = \frac{y_1^2}{n_1} + \dots + \frac{y_k^2}{n_k} - FC$$

$$SC \text{ error} = \sum_{\lambda, j} y_{\lambda j}^2 - \sum_{\lambda} \frac{y_{\lambda}^2}{n_{\lambda}} = SC \text{ total} - SC \text{ tratamiento}$$

Ejemplo numérico

Pesos al nacer de 8 camadas de cerdos

Nº de animales	C a m a d a								Total
	1	2	3	4	5	6	7	8	
1	2.0	3.5	3.3	3.2	2.6	3.1	2.6	2.5	
2	2.8	2.8	3.6	3.3	2.6	2.9	2.3	2.4	
3	3.3	3.2	2.6	3.2	2.9	3.1	2.3	3.0	
4	3.2	3.5	3.1	2.9	2.5	2.5	2.5	1.5	
5	4.4	2.3	3.2	3.3	2.0		1.2		
6	3.6	2.4	3.3	2.5	2.1		1.2		
7	1.9	2.0	2.9	2.6					
8	3.3	1.6	3.4	2.8					
9	2.8		3.2						
10	1.1								
\bar{y}_i	28.4	21.3	31.8	23.8	14.2	11.6	11.9	9.4	152.4
n_i	10	8	10	8	6	4	6	4	56
\bar{y}_i	2.84	2.66	3.18	2.98	2.37	2.90	1.98	2.35	

Análisis

$$FC = \frac{(152.4)^2}{56} = 414.75$$

$$SC \text{ total} = (2.0)^2 + \text{-----} + (1.5)^2 - FC = 24.65$$

$$SC \text{ trat} = \frac{(28.4)^2}{10} + \text{-----} + \frac{(9.4)^2}{4} - FC = 7.48$$

$$SC \text{ error} = 24.65 - 7.48 = 17.17$$

Tabla de análisis de variancia

FV	GL	SC	CM	F
Entre camadas	7	7.48	1.07	2.97*
Dentro de Camadas (Error)	48	17.17	0.36	
Total	55	24.65		

En ciertos casos se necesita calcular el número promedio de animales por tratamiento (como se verá al calcular componentes de variancia para estimar el Índice de herencia). En vez de usar el promedio simple de "ni" se prefiere la fórmula siguiente:

$$k_o = \frac{1}{n-1} \left(\sum k + \frac{k^2}{\sum k} \right) \text{ donde:}$$

(en nuestro caso esto significa lo siguiente)

k_o = número promedio de animales por camada

n = número de camada.

Cuál es el número promedio de animales por camada?

$$k^2 = 10^2 + 8^2 + 10^2 + 8^2 + 6^2 + 4^2 + 6^2 + 4^2 = 432$$

$$k_o = \frac{1}{8-1} \left(56 - \frac{432}{56} \right) = 6.90 \text{ animales}$$

2. Los factores en estudio con números desiguales en la sub-clase pero proporcionales

El análisis de variancia no presenta dificultades particulares, para este tipo de datos. La suma de cuadrados para efectos principales se calcula de los totales marginales correspondientes a cada factor, como si fuera un solo factor con números desiguales. La interacción se calcula por diferencia entre la suma de cuadrados entre sub-clase menos los efectos principales. El procedimiento que se sigue es idéntico al anterior. Un ejemplo numérico.

Porcentaje de faenamiento (disminuida en 70%) de 93 cerdos clasificados por raza y por sexo

Número	R A Z A									
	A		B		C		D		E	
	Macho	Hembra	Macho	Hembra	Macho	Hembra	Macho	Hembra	Macho	Hembra
1	13.3	18.2	10.9	14.6	13.6	12.9	11.6	13.8	10.3	12.8
2	12.6	11.3	3.3	15.3	13.1	14.4	13.2	14.4	10.3	8.4
3	11.5	14.2	10.5	11.8	4.1		12.6	4.9	10.1	10.6
4	15.4	15.9	11.6	11.0	10.8		15.2	.	6.9	13.9
5	12.7	12.9	15.4	10.9			14.2		13.2	10.0
6	15.7	15.1	14.4	10.5			12.4		11.0	
7	13.2		11.6	12.9					12.2	
8	15.0		14.4	12.5					13.3	
9	14.3		7.5	13.0					12.9	
10	16.5		10.8	7.6					9.9	
11	15.0		10.5	12.9						

Continuación

Número	R A Z A									
	A		B		C		D		E	
	Macho	Hembra	Macho	Hembra	Macho	Hembra	Macho	Hembra	Macho	Hembra
12	13.7		14.5	12.4						
13			10.9	12.8						
14			13.0	12.9						
15			15.9	10.9						
16			12.8	13.9						
17			14.0							
18			11.1							
19			12.1							
20			14.7							
21			12.7							
22			13.1							
23			10.8							
24			11.9							
25			10.7							
26			14.4							
27			13.3							
28			13.0							
29			12.7							
30			12.6							
\bar{y}_{ij}	168.9	87.6	362.7	182.7	41.6	27.3	79.7	33.1	110.1	55.7
n_{ij}	12	6	30	15	4	2	6	3	10	5

El número de animales en la sub-clase sigue la siguiente proporción:

$$\text{Raza A} = \text{Macho : Hembra } 12:6 = 2$$

$$\text{Raza B} = 30:15 = 2$$

$$\text{Raza C} = 4:2 = 2$$

$$\text{Raza D} = 6:3 = 2$$

$$\text{Raza E} = 10:5 = 2$$

$$12:6 = 30:15 = 4:2 = 6:3 = 10:5$$

$$12:30:4:6:10 = 6:15:2:3:5$$

Análisis

$$FC = \frac{(1149.4)^2}{93} = 14,205.60$$

$$SC \text{ total} = (13.3)^2 + \text{-----} + (10.0)^2 - FC = 580.02$$

$$SC \text{ sub-clase} = \frac{(168.9)^2}{12} + \text{---} + \frac{(55.7)^2}{5} - FC = 122.83$$

$$SC \text{ dentro de sub-clase (error)} = 580.02 - 122.83 = 457.19$$

$$SC \text{ Razas} = \frac{(256.5)^2}{18} + \text{---} + \frac{(165.8)^2}{15} - FC = 97.38$$

$$SC \text{ sexo} = \frac{(763.0)^2}{62} + \text{---} + \frac{(386.4)^2}{31} - FC = 0.52$$

$$SC \text{ interacción} = 122.83 - (97.38 + 0.52) = 24.93$$

Cuadro de análisis de variancia

FV	GL	SC	CM	F
Raza	4	97.38	24.34	4.42 ^{**}
Sexo	1	0.52	0.52	1
Interacción R x S	4	24.93	6.23	1.13
Error	83	457.19	5.51	
Total	92	580.02		

3. Número de sub-clases desproporcionales clasificado en más de una dirección

Este es el caso más general ya que es difícil encontrar proporcionalidad en el número de la sub-clase. La dificultad se presenta, en que teniendo por ejemplo dos factores en estudio (A y B), no podemos calcular la variancia entre los promedios ponderados de A, libre del efecto de B. La imposibilidad de obtener una suma de cuadrados para A o para B que esté libre del otro factor, recibe el nombre de no ortogonalidad de los datos. En otras palabras no podemos aplicar la propiedad de la aditividad de la suma de cuadrados, en consecuencia requiere un método especial de análisis.

El método básico, es el método de mínimos cuadrados (least-squares method). En general se puede decir que la estimación de los promedios en las sub-clases son estimaciones no viciadas de los correspondientes promedios de la población y la variancia dentro de la sub-clase, es una estimación también no viciada de la variancia común en la población. El problema radica en conseguir estimaciones no

no viciadas de los efectos principales (main effects) y la interacción y las variancias atribuibles a ellos.

A continuación vamos a describir algunos métodos corrientemente usados para estimar los componentes de variancia, en datos con sub-clases desproporcionales; sin embargo antes de estudiar cada método, vamos a presentar un ejemplo numérico para aclarar las dificultades que se presentan en el análisis de estos datos.

En el siguiente cuadro se presenta la ganancia promedio en pesos (gramos) de 149 ratas, en cuatro generaciones sucesivas.

Generación	M A C H O		H E M B R A		T O T A L	
	Nº de ratas	Ganancia promedio	Nº de ratas	Ganancia promedio	Nº de ratas	Ganancia promedio
1	21	176.952	27	109.518	48	139.021
2	15	161.467	25	114.080	40	131.850
3	12	155.667	23	108.522	35	124.686
4	7	171.000	19	106.790	26	125.077
T O T A L	55	167.327	94		149	131.121

De los resultados que se presentan en este cuadro de doble entrada se pueden derivar los componentes de efectos principales, sexo y generación y también la interacción entre los dos factores. El análisis de variancia es el método más conveniente y efectivo para probar la significación estadística de los varios efectos. Si el número de individuos dentro las sub-clases fueran iguales no se

tendría ninguna dificultad para calcular todos los componentes de variación. Tampoco se tendría dificultad si los números de individuos dentro de la sub-clase fueran desiguales, pero proporcionales (como se indicó en B).

La verdadera dificultad se presentan cuando el número de individuos en las sub-clases son desproporcionados como los datos que estamos considerando. Cabe destacar en el cuadro que la poca ganancia de peso observada en la 4ª generación puede ser enteramente, debido a la escasez de macho, que generalmente tienen mayor peso.

No hay dificultad para hacer un análisis preliminar de la variancia. Calculando los componentes grupo de Sexo-Generación e interacción (entre sub-clases) y error experimental. De los datos originales de ganancia de pesos individuales, se pueden calcular los siguientes:

Análisis preliminar

FV	GL	SC	CM
Sub-clase (sexo y generación e interacción)	7	119,141	17,020
Error	141	57,695	409
Total	148	176,836	

El error experimental es una estimación correcta de la variación al azar. Ahora si nos proponemos descomponer la suma de cuadrados de las sub-clases en sus componentes: sexo, generación y la interacción sexo x generación nos encontramos con dificultades.

1. El teorema de la adición de la suma de cuadrados no se cumple cuando el número de individuos en las sub-clases no es proporcional, así en nuestro caso al calcular la suma de cuadrados para sexo nos da = 114.287 y la de generación 5.756. La suma de $114.287 + 5.756 = 120.043$, este valor es mayor que la suma de cuadrados total entre la sub-clase ($120.043 > 119.141$).

Este es un ejemplo bien claro de que el teorema de la aditividad no se cumple.

FV	GL	SC	CM
Sub-clase (sexo-generación e interacción)	7	119,141	
Sexo	1	114,287	114,287
Generación	3	5,756	1,919
Interacción	3	---	---
Error	141	57,695	409

2. La segunda dificultad se encuentra en la descripción de la población de la cual se ha tomado la muestra, ya que el análisis depende de la característica de la población. Las hipótesis que deben ser planteadas acerca de la población son:

a) En las sub-clases de la población los individuos pueden ocurrir con frecuencia igual o proporcional y que la desproporción se debe a accidente de muestreo.

- b) La interacción en la sub-clase de la población puede no existir.
3. Una tercera dificultad ocasionada por el número de sub-clase no proporcional es la estimación de los efectos principales, ya que:
- a) La diferencia entre promedios de sexo como en el cuadro anterior, no es una función simple de la diferencia entre sexo en las varias generaciones.
- b) Si examinamos en el mismo cuadro la ganancia de peso de los machos en la generación 1 y 4, la ganancia promedio en la generación 1 es sólo 6 gramos más que en la generación 4 y la diferencia para hembra es aún menor. A continuación se exponen algunos métodos que pueden ser usados según la circunstancia.

Método de número esperado en la sub-clase (The method of expected subclass numbers)

Este método parte de la suposición de que de la población de la cual se extrae la muestra tiene números proporcionales en la sub-clase y que la desproporción en el número en la muestra se atribuye a accidente de muestreo. El método es apropiado solamente si toda la sub-clase contiene por lo menos un valor observado.

El procedimiento general será ilustrado por medio de los datos (promedio) del cuadro anterior.

Números observados y esperados (ganancia de pesos
disminuídos en 100 para facilitar los cálculos)

Gene- ración	M A C H O		H E M B R A		T O T A L	
	Obs- vado	Esperado	Obs- vado	Esperado	Obs- vado	Esperado
1	Nº 21	17.7181	27	30.2819	48	48.0000
		76.952	9.518		39.021	34.41
	Suma	1363.45		288.24	1873	1651.69
2	Nº 15	14.7651	25	25.2349	40	40.0000
		61.467	14.080		31.850	31.57
	Suma	907.56		355.31	1274	1262.87
3	Nº 12	12.9195	23	22.0805	35	35.0000
		55.667	8.522		24.686	25.92
	Suma	719.19		188.16	864	907.35
4	Nº 7	9.5973	19	16.4027	26	26.0000
		71.000	6.790		24.077	30.49
	Suma	681.41		111.37	626	792.77
Total	Nº 55	55.000	94	94.000	149	149.0000
		67.327	9.936	10.03	31.121	30.97
	Suma	3703	3661.61	934	943.08	4637

Los números esperados se calcularon multiplicando el total mar-
ginal hilera por total marginal columna dividido por total general.
Por ejemplo, número esperado para macho.

$$\text{Generación 1} = \frac{55 \times 48}{149} = \underline{17.7181}$$

Pasos a seguir en el análisis

Primer paso. Consiste en hacer una prueba cuantitativa de la validez de la presunción del número de sub-clase proporcional. Esto se hace por medio de la prueba de chi cuadrado.

1. Se calculan los números esperados en cada sub-clase.
2. Comparar los números esperados con los observados.
3. Calcular χ^2 para cada sub-clase y sumar los χ^2 para obtener el total y probar su significación.

$$\chi^2 = \sum \frac{(N_{\text{ob}} - N_{\text{esp}})^2}{N_{\text{esperado}}}$$

$$\chi^2 = \frac{(21 - 17.7181)^2}{17.7181} + \dots + \frac{(19 - 16.4027)^2}{16.4027} = 2.19$$

$$GL = (-1)(h - 1) = 3 \times 1 = 3 \quad P = .54$$

No es significativa, de ahí que el postulado del número proporcional estadísticamente se justifica. Esta es una garantía cuantitativa para adoptar el método del número de clase esperado.

Segundo paso. Se calcula la suma esperada para cada sub-clase multiplicando el promedio verdadero por el número esperado, así para macho.

$$\text{Generación 1 se tiene } (76.952)(17.7181) = \underline{1363.45}$$

Tercer paso. Se calculan las sumas y el promedio esperado en las columnas e hileras de los totales en la forma siguiente:

$$\text{Hilera. Generación 1} = 1363.45 + 288.24 = \underline{1651.69}$$

$$\bar{y} = \frac{1651.69}{48} = \underline{34.41}$$

$$\begin{aligned} \text{Columna. Macho} &= 1363.45 + 907.56 + 719.19 + 681.41 = \underline{\underline{366.61}} \\ &= \frac{3671.61}{55} = \underline{\underline{66.76}} \end{aligned}$$

Cuarto paso. Con los valores esperados se hace un análisis de variancia para estimar los componentes sexo, generación e interacción sexo x generación. Para el error se utiliza el mismo valor de la suma de cuadrados dentro de sub-clases deducido de los datos originales.

$$FC = \frac{(4614.68)^2}{149} = \underline{\underline{142,921}}$$

$$SC \text{ sub-clase} = \frac{1363.45^2}{17.7181} + \dots + \frac{111.37^2}{16.4027} - FC = \underline{\underline{116,305}}$$

$$SC \text{ sexo} = \frac{3681.61^2}{55} + \frac{943.08^2}{94} - FC = 111.644$$

$$SC \text{ generaciones} = \frac{1651.69^2}{48} + \dots + \frac{792.77^2}{26} - FC = \underline{\underline{1,480}}$$

$$SC \text{ interacción} = SC \text{ sub-clase} - (SC \text{ sexo} + SC \text{ Gen.}) = \underline{\underline{3,180}}$$

FV	GL	SC	CM	F
Sub-clase	7	116,305	-	
Sexos	1	111,644	111,644	272.97 **
Generación	3	1,480	493	1.21 NS
Interacción	3	3,181	1,060	2.60 *
Error	141	57,695	409	

Aquí se demuestra claramente el error que se tuvo con el primer análisis. En este análisis se pudo separar el componente Interacción y es significativo al 5%.

El componente de Variación, Generaciones, en este análisis no es significativo, lo que en el cuadro anterior si lo era.

El método que se ha descrito aquí se ha ensayado bajo diversas condiciones, durante varios años y ha dado buen resultado cuando los números en la sub-clase no son muy diferentes.

Método de los promedios no ponderados (The method of unweighted means)

Este método propuesto por Yates como un método aproximado, y su uso se justifica solo si los números en las sub-clase son aproximadamente iguales. Este método tiene algunas ventajas; tal como la manera familiar de cálculo. Si la interacción está presente ningún cálculo posterior se necesita hacer. Si la interacción es claramente despreciable los efectos principales pueden estar obviamente presente o ausente y ya no es necesario aplicar los métodos exactos, que veremos más adelante. Este método aproximado permite el uso de comparaciones planeadas o de todas las comparaciones que se desearan hacer.

A continuación presentamos un ejemplo numérico y los pasos a seguir en el análisis.

Ganancia promedio mensual de pesos de novillos
Holstein en tres edades

Mes de aumento	E D A D E N M E S E S						Total	Nº
	13		18		24			
	Nº (k)	Gan. (\bar{x})	Nº (k)	Gan. (\bar{x})	Nº (k)	Gan. (\bar{x})		
Agosto	5	22.2	7	21.4	3	40.7	84.3	15
Setiembre	8	24.4	3	30.3	1	39.0	93.7	12
Octubre	4	45.0	3	30.3	3	48.3	123.6	10
Noviembre	4	18.0	3	67.0	6	49.0	134.0	13
		109.6	149.0		177.0			

El objetivo de la investigación fue determinar si los meses afectaban los aumentos de peso.

Pasos a seguir

1. Hacer un análisis de variancia corriente con los datos originales.

a) Calcular la SC total = 44,931 con GL = 49

b) Calcular la SC entre sub-clases = 9,522 con GL = 11

c) Calcular el error (original) = SC dentro de sub-clases = SC total - SC entre sub-clases = 35,409 con 38 GL.

2. Con los promedios dados en el cuadro calcular:

$$FC = \frac{(435.6)^2}{12}$$

a) SC total sub-clases (\bar{x}) = $22.2^2 + \dots + 49.0^2 -$

$$\frac{(435.6)^2}{12} = \underline{2,319}$$

b) SC entre promedio de edades = $\frac{109.6^2 + \dots + 177.0^2}{4}$ -

FC = 573

c) SC entre promedio de estaciones = $\frac{84.3^2 + \dots + 134.0^2}{3}$ -

FC = 561

d) SC interacción = 2.319 - (573 + 561) = 1185

3. Ajustar el error en la siguiente forma.

a) Calcular el cuadrado medio dentro de sub-clases =

$$\frac{35.409}{38} = 931.8$$

b) Calcular un promedio del número de sub-clases k_o , que se calcula usando el promedio armónico de los 12 números en las sub-clases.

$$\frac{1}{k_o} = \frac{1}{12} \left(\frac{1}{5} + \frac{1}{8} + \dots + \frac{1}{3} + \frac{1}{6} \right) = 0.3168$$

c) Dividir el cuadrado medio dentro de sub-clases por (931.8) (0.3168) = 295.2 este es el error experimental apropiado a ser usado con todos los cuadrados medios calculados.

4. Cuadro de análisis de variancia completo y prueba de significancia.

FV		SC	CM	F
Entre meses	3	361	187	< 1
Entre edades	2	573	287	< 1
Interacción	6	1185	198	< 1
Error	38		295	--

Si algunos de los valores de F hubieran estado cerca de los niveles de significancia se contaría con uno de los métodos exactos que explicaremos enseguida. Este método proporciona buena información acerca de la interacción, para orientarle al investigador qué método debe seguir.

El método de los cuadrados ponderados de los promedios (The method of weighted squares of means)

Este método es aplicable a los análisis de datos con dos criterios de clasificación donde la interacción está presente, en este caso las estimaciones y las pruebas de significancia de los efectos principales no están viciadas. No provee información acerca de la interacción, a menos que uno de los factores contenga solo dos categorías.

Si no existe interacción, la estimación y prueba de significancia son ineficientes pero no viciadas. Este método exige que todas las sub-clases contengan por lo menos una observación. A continuación se presenta un ejemplo

Los pasos a seguir en este método de análisis se exponen a continuación:

1. Con los datos originales se hace un análisis de variancia corriente.

FV	GL	SC	CM
Entre sub-clases	9	678.62	75.40
Dentro de sub-clases	374	8,003.13	21.40
	383	8,681.75	

2. Calcular los valores presentados en el cuadro, para hileras y columnas, así como se detalla en el cuadro siguiente.
3. Cálculo de los efectos principales con interacción presente.
 - a) La variancia apropiada para probar la hipótesis nula de la diferencia entre los promedios de las hileras no ponderadas es una variancia ponderada con suma de cuadrados igual.

$$SC A = c^2 \sum w_i \left[\frac{(\bar{y}_{i1} + \bar{y}_{i2})}{2} - \bar{y}_w \right]^2 = c^2 \left\{ \sum w_i \left(\frac{\bar{y}_{i1} + \bar{y}_{i2}}{2} \right)^2 - \frac{\left[\sum w_i \left(\frac{y_{i1} + y_{i2}}{2} \right) \right]^2}{\sum w_i} \right\}$$

donde c = es el número de columna y

w_i = son las recíprocas de los coeficientes de las variancias correspondientes, se usa la recíproca debido a que una observación con una variancia pequeña tiene alta precisión así es que recibiría un mayor peso. Se ha usado w_i y v_j para las ponderaciones para no tener que desarrollar notaciones adicionales para los promedios no ponderados.

Largo promedio de gestación (días) de acuerdo al sexo del becerro (B) y de acuerdo a la secuencia de parición (A) (observaciones individuales disminuidas en 270)

Secuencia de parición (A)	Sexo (B)		$\sum \left(\frac{1}{n_{ij}} \right)$	$w_i = \frac{1}{\sum \left(\frac{1}{n_{ij}} \right)}$	Promedio simple $\frac{\bar{y}_{i1} + \bar{y}_{i2}}{2}$	Diferencia de promedio simple $\bar{y}_{i1} - \bar{y}_{i2} = \bar{d}_i$	$w_i \bar{d}_i$
	Macho b_1	Hembra b_2					
1 a ₁	54 0.018518 8.722	52 0.019231 6.462	0.037750	26.49057	7.592	2.260	59.869
2 a ₂	37 0.027027 9.649	34 0.029412 6.500	0.056439	16.71831	8.074	3.149	53.795
3 a ₃	37 0.027027 9.622	37 0.027027 8.973	0.054054	18.50000	9.297	0.649	12.007
4 a ₄	28 0.035714 9.679	28 0.035714 5.750	0.071428	14.00000	7.714	3.929	55.006
5 a ₅	46 0.021739 8.261	31 0.032258 8.581	0.053997	18.51948	8.421	-0.320	-5.926
$\sum \frac{1}{n_{ij}}$	0.130026	0.143642	0.053997	95.22836 = w_i	780.118		$\sum w_i \bar{d}_i = 176.751$
$v_j = \frac{1}{\sum \left(\frac{1}{n_{ij}} \right)}$	7.69077	6.96175	14.65252 = $\sum v_j$				$\sum (\bar{y}_{i1} + \bar{y}_{i2})/2$
Promedio simple $\sum \bar{y}_{ij}/5$	9.187	7.253	121.149 = $\sum v_j \frac{\sum \bar{y}_{ij}}{5}$	8.220			
$v_j \sum \frac{\bar{y}_{ij}}{5}$	70.655	50.494					

$$SC \text{ secuencia de parto} = 2^2 \left(6,427.322 - \frac{(770.118)^2}{95.22836} \right) = 146.144$$

b) La suma de cuadrados entre sexo (factor B) se calcula en la misma forma que para el factor A.

$$SC B = r^2 \left[\sum_{\lambda} v_j (\sum \bar{y}_{ij})^2 - \frac{\sum v_j (\sum y_{ij}/5)^2}{\sum v_j} \right]$$

donde r = número de hileras

$$SC \text{ entre sexos} = 5^2 \left[(1,015.340 - \frac{(121.149)^2}{14.65252}) \right] = 341.60$$

4. Cálculo de la interacción en este caso particular. La interacción se puede estimar, como explicamos más arriba, porque tenemos un factor con solo dos categorías. Si tuviéramos un r x c no sería eficiente la estimación de la interacción.

$$SC (A B) = \frac{(\sum w_i \bar{d}_i)^2}{\sum w_i}$$

$$SC \text{ interacción sexo x secuencia} = 536.810 - \frac{(176.751)^2}{95.22836} = 208.747$$

Análisis de variancia completo

FV	GL	SC	CM ajustado
Entre secuencias de Parto	4	146.14	36.54
Entre sexos	1	341.60	341.60 **
Interacción S de P x S	4	208.75	52.19 *
Error	374	8003.13	21.40

Aquí se puede notar que la interacción es fuerte. El procedimiento para estimar efectos principales se aplica para cualquier tabla $r \times c$. Si la interacción no existe, las estimaciones son todavía válidas pero ineficientes, por eso no se recomienda este método para el caso de interacción ausente.

El método de constantes de ajuste (The method of fitting constants)

El método de constantes de ajuste es general; es el único método que se puede aplicar cuando algunas observaciones en la sub-clase no contienen ninguna información. Así cuando en una tabla de $r \times c$ se tiene alguna sub-clase vacante.

El análisis de datos con número no proporcional clasificado en una tabla de $r \times c$ y cuando la interacción está presente, está dado por el modelo matemático.

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

Las restricciones que se imponen a este método son:

- Que la suma de las constantes a_i (estimador de α_i) sea igual a cero; en símbolo $\sum_k a_i = 0$; $a_1 + a_2 + \dots + a_n = 0$
- Que la suma de las constantes b_j (estimador de β_j) sea igual a cero; simbólicamente es igual a $\sum b_j = 0$; $b_1 + b_2 + \dots + b_n = 0$
- Que $\sum_k (ab)_{ij} = 0$ para cada j y que $\sum_k (ab)_{ij} = 0$ para cada i

El método de constantes de ajuste exige la estimación de μ de α 's de β 's y $(\alpha\beta)$'s, así mismo exige el cálculo de la reducción en el total de suma de cuadrados debido a ellos (generalmente se le llama

reducción debida a constantes de ajustes).

- Las estimaciones deben ser tal que la suma de cuadrados residual sea un mínimo.

La descripción del modelo cuando la interacción no existe es

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \xi_{ijk} \quad \text{con las siguientes restricciones}$$

$$\sum_i a_i = 0 = \sum_j b_j$$

Puesto que el método de constante de ajuste requiere mucho más trabajo si la interacción está presente, es deseable tener primero conocimiento de ella. Cuando la interacción está presente, la suma de cuadrados de los efectos principales es solo una aproximación, acentuándose más esto, cuando la desproporción es mayor. Sin embargo la estimación de la interacción es correcta.

Si no existe interacción la suma de cuadrados de los efectos principales no están viciados, pero ineficientemente estimados.

Un ejemplo numérico donde la interacción no está presente.

Intervalo en días entre el parto y el primer calor
en vacas Holstein-Friesian

Estación del año (factor A)	Partos (factor B)						Totales	
	Segundo	Tercero	Cuarto	Tercero	Cuarto	Tercero		
Invierno	26	987	26	1024	27	962	79	2,973
Primavera	30	997	22	676	28	873	80	2,546
Verano	17	463	18	463	18	558	53	1,484
Otoño	15	321	11	331	14	401	40	1,053
Totales	88	2768	77	2494	87	2794	252	8,056

Necesitamos hacer los siguientes cálculos.

1. Hacer un análisis de variancia preliminar con los datos originales en la forma siguiente:

- a) Calcular la suma de cuadrados total
- b) Calcular la suma de cuadrados entre sub-clases
- c) Calcular la SC para factor (A) ignorando el factor (B)
- d) Calcular la SC para factor B ignorando el factor A
- e) Calcular la SC dentro de las sub-clases

En nuestro ejemplo esto es igual

$$FC = \frac{(8,056)^2}{252} =$$

$$a) \text{ SC total} = \sum Y_{ij}^2 - FC = \underline{64,980.65}$$

$$b) \text{ SC entre sub-clases} = \frac{987^2}{26} + \dots + \frac{401^2}{14} - FC = \underline{5,794.11}$$

$$c) \text{ SC factor A (ignorando B)} = \frac{2,973^2}{79} + \dots + \frac{(1,053)^2}{40} - FC = \underline{4,645.07}$$

$$d) \text{ SC factor B (ignorando A)} = \frac{2,768^2}{88} + \dots + \frac{2,794^2}{87} - FC = \underline{38.77}$$

$$e) \text{ SC dentro de sub-clases} = \text{SC total} - \text{SC entre sub-clases} \\ = 64,980.65 - 5,794.11 = \underline{59,186.64}$$

Análisis de variancia preliminar

FV	GL	SC
Entre sub-clases	11	5,794.11
Estación (ignorando partos)	3	4,645.07
Partos (ignorando estación)	2	37.77
Dentro de sub-clases	240	59,186.64
Error		
Total	251	64,980.65

2. Plantear una serie de ecuaciones simultáneas de acuerdo al número de parámetros que se van a estimar. Dichos parámetros son μ (promedio general); α_i constantes para hileras (estaciones) y β_j constantes para columnas (partos). Nosotros usaremos los símbolos $\hat{\mu}$ como estimador de μ , a_i como estimador de α y b_j como estimador de β

Las ecuaciones normales son:

Parámetros	$\hat{\mu}$	a_1	a_2	a_3	a_4	b_1	b_2	b_3
μ :	$n \cdot \hat{\mu}$	$+ n_1 a_1$	$+ n_2 a_2$	$+ n_3 a_3$	$+ n_4 a_4$	$+ n_{.1} b_1$	$+ n_{.2} b_2$	$+ n_{.3} b_3 = Y_{..}$
α_1 :	$n_{1.} \hat{\mu}$	$+ n_{1.} a_1$				$+ n_{11} b_1$	$+ n_{12} b_2$	$+ n_{13} b_3 = Y_{1.}$
α_2 :	$n_{2.} \hat{\mu}$		$+ n_{2.} a_2$			$+ n_{21} b_1$	$+ n_{22} b_2$	$+ n_{23} b_3 = Y_{2.}$
α_3 :	$n_{3.} \hat{\mu}$			$+ n_{3.} a_3$		$+ n_{31} b_1$	$+ n_{32} b_2$	$+ n_{33} b_3 = Y_{3.}$
α_4 :	$n_{4.} \hat{\mu}$				$+ n_{4.} a_4$	$+ n_{41} b_1$	$+ n_{42} b_2$	$+ n_{43} b_3 = Y_{4.}$
β_1 :	$n_{.1} \hat{\mu}$	$+ n_{.1} a_1$	$+ n_{.2} a_2$	$+ n_{.3} a_3$	$+ n_{.4} a_4$	$+ n_{.1} b_1$		$= Y_{.1}$
β_2 :	$n_{.2} \hat{\mu}$	$+ n_{.2} a_1$	$+ n_{.2} a_2$	$+ n_{.3} a_3$	$+ n_{.4} a_4$		$+ n_{.2} b_2$	$= Y_{.2}$
β_3 :	$n_{.3} \hat{\mu}$	$+ n_{.3} a_1$	$+ n_{.3} a_2$	$+ n_{.3} a_3$	$+ n_{.4} a_4$			$+ n_{.3} b_3 = Y_{.3}$

Para resolver estas ecuaciones nosotros imponemos las siguientes restricciones:

$$\sum_i a_i = 0; a_1 + a_2 + \dots + a_n = 0$$

$$\sum_j b_j = 0; b_1 + b_2 + \dots + b_n = 0$$

Ahora podemos resolver fácilmente las ecuaciones por eliminación del juego de constantes que tiene mayor número de incógnitas. En

nuestro ejemplo resolveremos para las b's, imponiendo la condición

$\sum_j b_j = 0$, luego sustituir estos resultados en las a's ecuaciones y resolver para los $(\hat{\mu} + a_i)$'s. Puesto que $\sum (\hat{\mu} + a_i) = \sum \hat{\mu} + \sum a_i$ siendo $\sum \hat{\mu} = n\hat{\mu}$ y $\sum a_i = 0$, tendremos $\sum (\hat{\mu} + a_i) = n\hat{\mu}$, luego se obtienen fácilmente los valores de a_i's sustituyendo $\hat{\mu}$ en la ecuación.

3. Cálculo de la suma de cuadrados de reducción debido a constante de ajuste. Esto se obtiene fácilmente por medio de la ecuación siguiente.

$$\hat{\mu} \sum y_i + \sum a_i y_i + \sum b_j y_j = \sum (\hat{\mu} + a_i) y_i + \sum b_j y_j$$

A continuación presentamos el cálculo completo del ejemplo propuesto más arriba.

Datos arreglados por el método de constantes de ajuste

	n_{i1}	n_{i1}/n_i	n_{i2}	n_{i2}/n_i	n_{i3}	n_{i3}/n_i	$n_{i.}$	y_i
n_{1j} y $n_{1j}/n_{1.}$	26	0.329114	26	0.329114	27	0.341772	79	2,973
n_{2j} y $n_{2j}/n_{2.}$	30	0.375000	22	0.275000	28	0.350000	80	2,546
n_{3j} y $n_{3j}/n_{3.}$	17	0.370755	18	0.339623	18	0.339623	53	1,484
n_{4j} y $n_{4j}/n_{4.}$	15	0.375000	11	0.275000	14	0.355000	40	1,053
$n_{.j}$	88		77		87		252	
y_j	2,768		2,494		2,994			

Los pasos a seguir son los siguientes (después del análisis preliminar)

1. Obtener los n_{ij}/n_i valores, por ejemplo: $n_{11}/n_1 = \frac{26}{79} = 0.329114$

$$n_{21}/n_1 = \frac{28}{80} = 0.350000$$

Chequear que la $\sum n_{ij}/n_i = 1$; por ejemplo para $i = 3$

$$0.320755 + 0.33923 + 0.339623 = 1.000001 \approx 1$$

2. Calcular los coeficientes de las ecuaciones, c_i y c_{ij} en la siguiente forma y plantear las ecuaciones

$$a) \quad c_1 = \sum_i (n_{i1}/n_i) Y_i - Y_1$$

$$c_2 = \sum_i (n_{i2}/n_i) Y_i - Y_2$$

$$c_3 = \sum_i (n_{i3}/n_i) Y_i - Y_3$$

$$b) \quad c_{11} = \sum_i (n_{i1}/n_i) n_{i1} - n_{.1}$$

$$c_{22} = \sum_i (n_{i2}/n_i) n_{i2} - n_{.2}$$

$$c_{33} = \sum_i (n_{i3}/n_i) n_{i3} - n_{.3}$$

$$c) \quad c_{12} = \sum_i (n_{i1}/n_i) n_{i2}$$

$$c_{13} = \sum_i (n_{i1}/n_i) n_{i3}$$

$$c_{23} = \sum_i (n_{i2}/n_i) n_{i3}$$

Aplicando esta fórmula a nuestro ejemplo tenemos

$$a) \quad c_1 = 0.329114(2,973) + 0.375000(2,546) + 0.320755(1,484) \\ + 0.375000(1,053) - 2,768 = \underline{36.0813}$$

$$c_2 = 0.329114(2,973) + 0.275000(2,546) + 0.339623(1,484) \\ + 0.27500(1,053) - 2,494 = - \underline{21.8185}$$

$$c_3 = 0.341772(2,973) + 0.350000(2,546) + 0.339623(1,484) \\ + 0.350000(1,053) - 2,794 = - \underline{14.2613}$$

Se pueden chequear, si las operaciones están bien la suma de los c_i , deben ser igual a cero ($\sum_i c_i = 0$)

$$c_i = 36.0813 - 21.8185 - 14.2613 = 0.0015 \doteq 0$$

$$b) \quad C_{11} = 0.329114(26) + 0.375000(30) + 0.32175(17) + \\ 0.375000(15) - 88 = - \underline{57.1152}$$

$$C_{22} = 0.329114(26) + 0.275000(2) + 0.339623(18) + \\ 0.275000(11) - 77 = - \underline{53.2548}$$

$$C_{33} = 0.341772(27) + 0.350000(28) + 0.339623(18) + \\ 0.350000(14) - 87 = - \underline{56.9589}$$

$$c) \quad C_{12} = 0.329114(26) + 0.375000(22) + 0.320755(18) + \\ 0.375000(11) = \underline{26.7055}$$

$$C_{13} = 0.329114(27) + 0.375000(28) + 0.320755(18) + \\ 0.375000(14) = \underline{30.4097}$$

$$C_{23} = 0.329114(27) + 0.275000(28) + 0.339623(18) + \\ 0.275000(14) = \underline{26.5493}$$

ahora podemos plantear las siguientes ecuaciones simultáneas.

$$C_{11}b_1 + C_{12}b_2 + C_{13}b_3 = c_1$$

$$C_{12}b_1 + C_{22}b_2 + C_{23}b_3 = c_2$$

$$C_{13}b_1 + C_{23}b_2 + C_{33}b_3 = c_3$$

antes de proceder a resolver las ecuaciones se debe hacer el siguiente chequeo.

$$C_{11} + C_{12} + C_{13} = 0$$

$$C_{12} + C_{22} + C_{23} = 0$$

$$C_{13} + C_{23} + C_{33} = 0$$

En nuestro ejemplo tenemos las siguientes ecuaciones

$$-57.1152b_1 + 26.7055b_2 + 30.4097b_3 = 36.0813 \quad (1)$$

$$26.7055b_1 - 53.2548b_2 + 26.5493b_3 = 21.8185 \quad (2)$$

$$30.4097b_1 + 26.5493b_2 - 56.9589b_3 = 14.2613 \quad (3)$$

Estas ecuaciones se pueden resolver por cualquier método, sin embargo nosotros vamos a usar el método de determinantes y matrices invertida. El método de determinantes lo usaremos cuando la ecuación consta de 3 o menos incógnitas, y para cualquier número de incógnitas usaremos matrices invertidas (repaso sobre matrices y determinantes, se encuentra en la parte final de este apunte).

Las ecuaciones que hemos planteado las vamos a resolver por el método de determinantes, para que sea más fácil eliminaremos una de las incógnitas, valiéndonos de la restricción $\sum_j b_j = 0 = b_1 + b_2 = -b_3$ donde $b_3 = -(b_1 + b_2)$ reemplazando b_3 por su valor en las ecuaciones (1) y (2) tendremos:

$$(C_{11} - C_{13}) b_1 + (C_{12} - C_{13}) b_2 = C_1$$

$$(C_{12} - C_{23}) b_1 + (C_{22} - C_{23}) b_2 = C_2$$

$$(-57.1152 - 30.4097)b_1 + (26.7055 - 30.4097)b_2 = 36.0813 \quad (4)$$

$$(26.7055 - 26.5493) b_1 + (53.2548 - 26.5493) b_2 = -21.8185 \quad (5)$$

reduciendo estas ecuaciones tenemos:

$$-87.5249b_1 - 3.7042b_2 = 36.0813 \quad (4)$$

$$0.1562b_1 - 79.8041b_2 = -21.8185 \quad (5)$$

ahora podemos resolver fácilmente estas dos ecuaciones de primer grado con dos incógnitas b_1 y b_2 por el método de determinantes.

$$|A| = \begin{vmatrix} -87.5249 & -3.7042 \\ 0.1562 & -79.8041 \end{vmatrix} = (-87.5249)(-79.8041) - (0.1562)(-3.7042) = 6,985.4245$$

$$Ab_1 = \begin{vmatrix} 36.0813 & -3.7042 \\ -21.8185 & -79.8041 \end{vmatrix} = (36.0813)(-79.8041) - (-21.8185)(-3.7042) = -2,960.2558$$

$$Ab_2 = \begin{vmatrix} -87.5249 & 36.0813 \\ 0.1562 & -21.8185 \end{vmatrix} = (-87.5249)(-21.8185) - (0.1562)(36.0813) = 1,904.0261$$

$$b_1 = \frac{Ab_1}{|A|} = \frac{-2,960.2558}{6,985.4245} = -0.423776$$

$$b_2 = \frac{Ab_2}{|A|} = \frac{1,904.0261}{6,985.4245} = 0.272571$$

$$b_3 = -(b_1 + b_2) = -(-0.423776 + 0.272571) = 0.151205$$

Verificación. Vamos a reemplazar estos valores de b_1 , b_2 y b_3 en la ecuación 3 (original).

$$30.4097(-0.423776) + 26.5493(0.272571) - 56.9589(0.151205) = -14.2613 - 14.2628 = -14.2613$$

Se ha comprobado que los valores de b 's están correctos $\sum_{j=1}^3 b_j = 0$

3. Resolver las ecuaciones de Q_i para los valores de $\hat{\mu} + Q_i$ en la siguiente forma:

$$\hat{\mu} + a_1 = \frac{Y_1}{n_1} - \frac{n_{11}b_1}{n_1} - \frac{n_{12}b_2}{n_1} - \frac{n_{13}b_3}{n_1}$$

$$\hat{\mu} + a_2 = \frac{Y_2}{n_2} - \frac{n_{21}b_1}{n_2} - \frac{n_{22}b_2}{n_2} - \frac{n_{23}b_3}{n_2}$$

$$\hat{\mu} + a_3 = \frac{Y_3}{n_3} - \frac{n_{31}b_1}{n_3} - \frac{n_{32}b_2}{n_3} - \frac{n_{33}b_3}{n_3}$$

$$\hat{\mu} + a_4 = \frac{Y_4}{n_4} - \frac{n_{41}b_1}{n_4} - \frac{n_{42}}{n_4} - \frac{n_{43}}{n_4}$$

$$\begin{aligned}\hat{\mu} + a_1 &= \frac{2,973}{79} = 0.329114 (-0.423776) - 0.329114(0.272571) \\ &\quad - 0.341772 (0.151205) = \underline{37.6310}\end{aligned}$$

$$\begin{aligned}\hat{\mu} + a_2 &= \frac{2,546}{80} - 0.375000(-0.423776) - 0.275000(0.272571) \\ &\quad - 0.350000(0.151205) = \underline{31.8560}\end{aligned}$$

$$\begin{aligned}\hat{\mu} + a_3 &= \frac{1,484}{53} - 0.370755(-0.423776) - 0.339623(0.272571) \\ &\quad - 0.339623(0.151205) = \underline{27.9920}\end{aligned}$$

$$\begin{aligned}\hat{\mu} + a_4 &= \frac{1,053}{40} - 0.37000(-0.423776) - 0.275000(0.272571) \\ &\quad - 0.355000(0.151205) = \underline{26.3560}\end{aligned}$$

sumando miembro a miembro tenemos

$$\begin{aligned}4\hat{\mu} + \sum_1^4 a_i &= 123.835 \text{ siendo } \sum_1^4 a_i = 0 \\ \hat{\mu} &= \frac{123.835}{4} = 30.9588\end{aligned}$$

Para obtener los valores de a_1 , a_2 , a_3 , y a_4 , bastaría reemplazar $\hat{\mu}$ por su valor en las cuatro ecuaciones, vale decir que:

$$a_i = (\hat{\mu} + a_i) - \hat{\mu} = \hat{\mu} + a_i - \hat{\mu}$$

de donde:

$$a_1 = 37.6310 - 30.9588 = \underline{6.6722}$$

$$a_2 = 31.8560 - 30.9588 = \underline{0.8972}$$

$$a_3 = 27.9920 - 30.9588 = \underline{-2.9668}$$

$$a_4 = 26.3560 - 30.9588 = \underline{-4.6028}$$

$$\text{Lógicamente } \sum_1^4 a_i = 0$$

4. Cálculo de la suma de cuadrados de reducción debido a constante de ajuste, por la fórmula dada anteriormente

$$\sum_1^A (\hat{\mu} + a_i) Y_i + \sum_1^3 b_j Y_j - \frac{Y_{1j}^2}{n}$$

Estaciones	Partos			Y _{1..}	$\hat{\mu} + a_i$
	Primero Y ₁₁	Tercero Y ₁₂	Cuarto Y ₁₃		
Y _{1j}	Y ₁₁	Y ₁₂	Y ₁₃	2,973	37.6310
Y _{2j}	Y ₂₁	Y ₂₂	Y ₂₃	2,546	31.8560
Y _{3j}	Y ₃₁	Y ₃₂	Y ₃₃	1,484	27.9920
Y _{4j}	Y ₄₁	Y ₄₂	Y ₄₃	1,053	26.3560
	2,768	2,494	2,794		$\hat{\mu} = 30.9588$
	-0.423776	+0.272571	+0.151205		

$$\begin{aligned} \text{SC de constantes de ajuste} &= 37.6310(2,973) + 31.8560(2,546) + 27.9920 \\ &\quad (1,484) + 26.3560 (1,053) \\ &\quad - 0.423776 (2,768) + 0.272571 (2,494) + \\ &\quad 0.151205 (2,794) - \frac{(8,056)^2}{252} \end{aligned}$$

$$\text{SC de constantes de ajuste} = 262,204.59 - 257,536.25 = \underline{4,668.34}$$

5. Cálculo de SC de la interacción y prueba de significancia de la interacción.

$$\begin{aligned} \text{SC Interacción A x B} &= \text{SC sub-clases (original)} - \text{SC d. a} \\ &\quad \text{constante de ajuste.} \end{aligned}$$

$$= 5,794.11 - 4,668.34 = 1,125.77$$

$$\text{CM Interacción A x B} = 1,125.77/6 = 187.63$$

$$F = \frac{187.63}{246.61} < 1$$

Esto nos da una garantía cuantitativa de que la interacción no existe, entonces podemos seguir adelante.

6. Cálculo del factor de ajuste para los efectos principales.

Es necesario ajustar los cuadrados medios de los efectos principales para probar su significancia estadística; para el efecto procedemos de las siguientes maneras:

a) Factor de ajuste = SC factor A + SC factor B - SC d, a constante de ajuste.

$$= 4,645.07 + 37.77 - 4,668.34 = \underline{14.50}$$

b) SC de estaciones ajustadas = SC original - factor de ajuste.

$$4,645.07 - 14.50 = \underline{4,630.57}$$

c) SC de partos ajustado = SC original - factor de ajuste

$$37.77 - 14.50 = \underline{23.27}$$

7. Análisis de variancia completo.

FV	GL	SC ajustado	CM ajustado	F	
Entre estaciones (A)	3	4,630.57	1,543.52	6.26	**
Entre partos (B)	2	23.27	11.64	< 1	N S
Interacción A x B	6	1,125.77	187.63	< 1	N S
Error	240	59,186.64	246.61		

Promedios ajustados de los tratamientos (interacción despreciable)

Estación	Constantes	PARTOS				Promedio ajustado de estación	
		Segundo $b_1 = -0.4238$ observado ajuste	Tercero $b_2 = +0.2726$ observado ajuste	Cuarto $b_3 = +0.1511$ observado ajuste			
Invierno	$\bar{C}_1 = 6.6722$	37.9615	39.3846	37.9036	35.6296	37.7821	37.6310
Primavera	$\bar{C}_2 = 0.8972$	33.2333	30.7273	32.1289	31.1786	32.0071	31.8560
Verano	$\bar{C}_3 = 2.9668$	27.2353	25.7222	28.2646	31.0000	28.1431	27.9920
Otoño	$\bar{C}_4 = 4.6028$	21.4000	30.0909	26.6286	28.6429	26.5071	26.3560
Promedio ajustado de partos		30.5350	31.2316	31.1100			30.9588 = $\bar{\mu}$

9. Ajuste de promedios. La ausencia de interacción nos indica que los a's y b's son estimadores no viciados de los correspondientes parámetros de la población α 's y β 's. Ahora vamos a ajustar los promedios de las sub-clases y efectos principales, aplicando la fórmula, $\bar{y}_{ij} = \hat{\mu} + \alpha_i + \beta_j$

Cuando la interacción es significativa las pruebas de los efectos principales pierde validez. Para probar una hipótesis nula para un efecto principal es necesario ajustar la constante por el otro efecto principal y por la interacción con el doble juego de restricciones. La diferencia así obtenida entre la reducción y los efectos principales más la interacción calculada, como entre promedios de sub-clases es apropiada para probar la hipótesis nula acerca del efecto principal. Por esta razón el método de constante de ajuste generalmente no se usa para probar efectos principales cuando la interacción está presente.

ANALISIS POR EL METODO DE MINIMOS CUADRADOS DE DATOS CON NUMERO DESIGUAL EN LA SUB-CLASE

Ya hemos hablado de las múltiples dificultades que se hallan asociadas al análisis de los datos no ortogonales. Vamos a ilustrar con ciertos detalles el procedimiento general para estimar las constantes estadísticas, con datos que tienen número desigual en la sub-clase.

Principios de mínimos cuadrados con datos de un criterio de clasificación

Se ha indicado anteriormente que los datos con un criterio de clasificación no presenta ningún problema para su análisis, y que el método ordinario de análisis de variancia es eficiente para estimar los componentes de variancia, sin embargo vamos a demostrar aquí que el análisis por el método de constante de ajuste en el cual se hace uso de las "matrices" da el mismo resultado que el método corriente, pero para aclarar el concepto de mínimos cuadrados vamos a trabajar con este tipo de datos por ahora, para luego generalizar con otros ejemplos.

Modelo matemático

Los datos con un criterio de clasificación siguen el modelo matemático siguiente:

$$Y_{ij} = \mu + a_i + e_{ij}$$

$$i = 1, 2 - - - , p$$

$$j = 1, 2 - - - , n_i$$

donde:

Y_{ij} = El valor de la observación (Y_a sea edad, toros, años, tratamientos etc.)

μ = Promedio general de la población cuando existe igual frecuencia entre las clases de A

a_i = Efecto de tratamientos expresado como desviación del promedio general.

e_{ij} = Errores al azar.

Los términos de la ecuación representan los parámetros de la población. Como estimadores de estos parámetros por la facilidad de representar simbólicamente, usaremos las mismas letras con un pequeño ángulo arriba ($\hat{}$) así $\hat{\mu}$ es estimador de μ

Ecuaciones de mínimos cuadrados

Las ecuaciones normales resultan del uso de un principio del cálculo diferencial, sin embargo resulta realmente innecesario operar con los cálculos diferenciales para escribir las ecuaciones. Lo principal que debe recordarse es que por cada constante a ser estimada debe haber una ecuación. Las ecuaciones para datos de un criterio de clasificación son como sigue:

	$\hat{\mu}$	\hat{a}_1	\hat{a}_2	-----	\hat{a}_p	
$\mu:$	$n \cdot \hat{\mu}$	$+ n_1 \hat{a}_1$	$+ n_2 \hat{a}_2$	-----	$+ n_p \hat{a}_p$	$= Y_1$
$a_1:$	$n_1 \hat{\mu}$	$+ n_1 \hat{a}_1$				$= Y_1$
$a_2:$	$n_2 \hat{\mu}$		$+ n_2 \hat{a}_2$			$= Y_2$
$a_p:$	$n_p \hat{\mu}$				$+ n_p \hat{a}_p$	$= Y_p$

En las ecuaciones $n = \sum_i n_i$ donde n_i es el número de observaciones en la clase i^{th} . La columna de valores en el lado derecho de las ecuaciones se refiere a los miembros derecho de la ecuación (raight hand members = RHM). Muy amenudo la misma matriz variancia-covariancia se usa con varios RHM's.

Forma tabular de las ecuaciones de mínimos cuadrados

	$\hat{\mu}$	\hat{a}_1	\hat{a}_2 - - - - -	\hat{a}_p	RHM
μ :	n .	n_1	n_2 - - - - -	n_p	Y .
a_1 :	n_1	n_1	0 - - - - -	0	Y_1
a_2 :	n_2	0	n_2 - - - - -	0	Y_2
.	.				
.	.				
.	.				
.	.				
.	.				
a_p :	n_p	0	0 - - - - -	n_p	Y_p

Cuatro aspectos de las ecuaciones de mínimos cuadrados se pueden notar en el presente cuadro.

1. Los elementos de la matriz variancia - covariancia (los miembros izquierdos de las ecuaciones) forman una matriz simétrica alrededor de la diagonal principal (la parte superior a la diagonal es igual a los elementos de la parte inferior).
2. Los elementos de la ecuación μ , como los coeficientes para \hat{a}_i son los mismos para los elementos de la diagonal en las ecuaciones a_i .
3. Los elementos fuera de la diagonal en la sección de hileras y columnas para las \hat{a}_i son cero.
4. La suma de los coeficientes para los \hat{a}_i en la ecuación μ es igual a los coeficientes para $\hat{\mu}$ en la misma ecuación. Lo

mismo la suma de los RHM's para las ecuaciones a_i es igual al RHM para la ecuación μ

Restricciones impuestas

Un método directo de obtener los estimadores de a_i como desviaciones de μ consiste en imponer la restricción en las ecuaciones de mínimos cuadrados que $\sum_L \hat{a}_i = 0$. Si esto se cumple los coeficientes de uno de los \hat{a}_i digamos \hat{a}_p se puede sustraer de los coeficientes de los otros, los elementos que resultaron en las a_p ecuaciones son entonces sustraídos de los elementos correspondientes en la a_i ecuaciones para mantener simetría en las ecuaciones. Además el RHM para la ecuación a_p se sustrae de los RHM's para las otras ecuaciones de a_i . Cuando estas subtracciones se completan la columna e hilera para la ecuación a_p desaparece y el resultado de la serie simétrica de ecuaciones se resuelve para obtener los estimadores de μ y de a_i directamente.

Utilización de la matriz inversa

Existen varios métodos para resolver ecuaciones simultáneas y para obtener la inversa de una matriz, nosotros no vamos a discutir ninguno, sino simplemente nos limitaremos a la aplicación de los elementos inversos en los análisis de datos.

Cálculo de las constantes

Aunque los estimadores de las constantes se pueden obtener directamente resolviendo las ecuaciones, a menudo es más conveniente obtenerlo de los elementos inversos y los RHM's de las ecuaciones, puesto que

$$\sum_j c^{ij} y_j = \hat{c}_i$$

donde:

C^{ij} = elemento inverso para la hilera i^{th} y la columna j^{th} de la matriz inversa.

Y_{ij} = RHM para las hileras j^{th}

C_i = constantes estimadas i^{th}

por ejemplo:

$$\hat{\mu} = C^{11} Y. + C^{12}(Y_1 - Y_p) + C^{13}(Y_2 - Y_p) + \dots + C^{1p}(Y_{p-1} - Y_p)$$

$$a_1 = C^{21} Y + C^{22}(Y_1 - Y_p) + C^{23}(Y_2 - Y_p) + \dots + C^{2p}(Y_{p-1} - Y_p)$$

etc. etc.

Cálculo de los errores estandares de las constantes estimadas o funciones lineales de las constantes estimadas

Tales como $\hat{\mu}$ o \hat{a}_i es igual a

$$S_{\hat{c}_i} = \sqrt{C^{ii} \delta^2}$$

donde:

C^{ii} = es el correspondiente elemento inverso de la diagonal para la constante considerada.

$\delta_e^2 = \frac{1}{n-p} \left[\sum_k \sum_j Y_{ij}^2 - R(\mu, a_i) \right]$, esta es la fórmula para datos con un solo criterio de clasificación, donde $\sum_k \sum_j Y_{ij}^2$ = suma de cuadrado total (sin corregir) y $R(\mu, a_i)$ es la reducción debido a todas las constantes de ajuste, $R(\mu, a_i)$, puede ser calculado de todas las constantes estimadas y los RHM's original en la forma siguiente:

$$R(\mu, a_i) = \hat{\mu} Y + \hat{a}_1 Y_1 + \hat{a}_2 Y_2 + \dots + \hat{a}_p Y_p$$

o también puede calcularse de las constantes y las RHM's reducida

$$R(\mu, a_i) = \hat{\mu} Y + a_1 (Y_1 - Y_p) + a_2 (Y_2 - Y_p) + \dots + a_{p-1} (Y_{p-1} - Y_p)$$

Error estandar de la diferencia entre dos constantes estimadas

El error estandar de la diferencia entre dos constantes estimadas, está dada por la fórmula

$$S_{\hat{a}_i - \hat{a}_j} = \sqrt{(c^{ii} + c^{jj} - 2c^{ij})\hat{\sigma}^2}$$

Cuando se impone la restricción que $\sum_I \hat{a}_i = 0$, los elementos inversos para la columna e hilera \hat{a}_p se calcula de las dependencias que tienen los elementos inversos dentro de su juego para sumar cero por hileras y columnas, por ejemplo:

$$c^{1(p+1)} = - (c^{12} + c^{13} + c^{14} + \dots + c^{1p})$$

$$c^{2(p+1)} = - (c^{22} + c^{23} + c^{24} + \dots + c^{2p})$$

.

.

.

$$c^{(p+1)(p+1)} = - (c^{1(p+1)} + c^{2(p+1)} + c^{3(p+1)} + \dots + c^{p(p+1)})$$

El error estandar del promedio de mínimos cuadrados se obtiene en la forma siguiente:

$$S_{\hat{\mu} + \hat{a}_i} = \sqrt{(c^{11} + c^{ii} + 2c^{1i})\hat{\sigma}^2}$$

Cálculo de las sumas de cuadrados para el análisis de variancia

La suma de cuadrados para A se puede obtener fácilmente por el procedimiento corriente para datos con un criterio de clasificación. Sin embargo esta misma suma de cuadrados se puede obtener de una manera más general como se demuestra a continuación.

$$SC = B' Z^{-1} B$$

donde B' es una hilera vector de las constantes estimadas para un juego dado (tal como \hat{a}_i); Z^{-1} es la inversa del segmento de la

inversa de la matriz variancia-covariancia correspondiente por hileras y columnas a este juego de constantes. La SC obtenida así es igual a la SC de reducción, debido a todas las constantes de ajuste. La SC de reducción debida a todas las constantes, excepto el juego que se está considerando, es decir $R(\mu, a_i) - R(\mu)$

Ejemplo numérico.

Vamos a suponer que los datos que presentamos en el cuadro siguiente representan las ganancias de pesos de cerdos alimentados con tres raciones. (Los cerdos fueron asignados al azar a las raciones)

	R A C I O N			Total
	1	2	3	
	3	5	7	
	5	6	6	
	6	2	4	
	2	7	3	
		8	6	
		3	4	
		9		
		8		
Y_i	16	48	30	94
\bar{Y}_i	4	6	5	
n_i	4	8	6	

1. Modelo matemático

$$Y_{ij} = \mu + r_i + e_{ij}$$

donde:

Y_{ij} = ganancia de peso de los cerdos j^{th} con la ración i^{th}

μ = promedio de la población con igual número

r_i = efecto de la ración

e_{ij} = error al azar que se asume independiente y normalmente distribuido NID $(0, \sigma^2)$

2. Ecuaciones de mínimos cuadrados.

	$\hat{\mu}$	\hat{r}_1	\hat{r}_2	\hat{r}_3	RHM
μ :	18	4	8	6	94
r_1 :	4	4	0	0	16
r_2 :	8	0	8	0	48
r_3 :	6	0	0	6	30

3. Restricción impuesta $\sum \hat{r}_1 = 0$. Con esta restricción impuesta podemos proceder a reducir las ecuaciones.

a) Reducción de hileras (substrayendo r_3 de r_1 y r_2)

	$\hat{\mu}$	\hat{r}_1	\hat{r}_2	RHM
μ :	18	-2	2	94
r_1 :	4	4	0	16
r_2 :	8	0	8	48
r_3 :	6	-6	-6	30

b) Reducción de columnas y RHM's

	$\hat{\mu}$	\hat{r}_1	\hat{r}_2	RHM
$\mu:$	18	-2	2	94
$r_1:$	-2	10	6	-14
$r_2:$	2	6	14	18

A estas ecuaciones se les llaman ecuaciones reducidas de mínimos cuadrados.

4. Inversión de la matriz variancia-covariancia.

a) Cálculo del determinante $|A|$. Esto se obtiene fácilmente de las ecuaciones reducidas, operando de la siguiente manera.

$$|A| = \begin{vmatrix} 18 & -2 & 2 \\ -2 & 10 & 6 \\ 2 & 6 & 14 \end{vmatrix} = 18 \begin{vmatrix} 10 & 6 \\ 6 & 14 \end{vmatrix} + 2 \begin{vmatrix} -2 & 6 \\ 2 & 14 \end{vmatrix} + 2 \begin{vmatrix} -2 & 10 \\ 2 & 6 \end{vmatrix} = 1,728$$

o lo mismo se puede tomar la segunda hilera

$$|A| = \begin{vmatrix} 18 & -2 & 2 \\ -2 & 10 & 6 \\ 2 & 6 & 14 \end{vmatrix} = 2 \begin{vmatrix} -2 & 2 \\ 6 & 14 \end{vmatrix} + 10 \begin{vmatrix} 18 & 2 \\ 2 & 14 \end{vmatrix} - 6 \begin{vmatrix} 18 & -2 \\ 2 & 6 \end{vmatrix} = 1,728$$

tomando la tercera hilera

$$|A| = \begin{vmatrix} 18 & -2 & 2 \\ -2 & 10 & 6 \\ 2 & 6 & 14 \end{vmatrix} = 2 \begin{vmatrix} -2 & 2 \\ 10 & 6 \end{vmatrix} - 6 \begin{vmatrix} 18 & 2 \\ -2 & 6 \end{vmatrix} + 14 \begin{vmatrix} 18 & -2 \\ -2 & 10 \end{vmatrix} = 1,728$$

o lo mismo se puede tomar cualquier columna

$$|A| = 1,728$$

b) Cálculo de los cofactores de la matriz $A^{ij} = A^{ji}$

$$A^{11} = (-1)^{1-1} \begin{bmatrix} 10 & 6 \\ 6 & 14 \end{bmatrix} = 104$$

$$A^{12} = (-1)^{1+2} \begin{bmatrix} -2 & 6 \\ 2 & 14 \end{bmatrix} = 40$$

$$A^{13} = (-1)^{1+3} \begin{bmatrix} -2 & 10 \\ 2 & 6 \end{bmatrix} = -32$$

$$A^{21} = (-1)^{2+1} \begin{bmatrix} -2 & 2 \\ 6 & 14 \end{bmatrix} = 40$$

$$A^{22} = (-1)^{2+2} \begin{bmatrix} 18 & 2 \\ 2 & 14 \end{bmatrix} = 248$$

$$A^{23} = (-1)^{2+3} \begin{bmatrix} 18 & -2 \\ 2 & 6 \end{bmatrix} = -112$$

$$A^{31} = (-1)^{3+1} \begin{bmatrix} -2 & 2 \\ 10 & 6 \end{bmatrix} = -32$$

$$A^{32} = (-1)^{3+2} \begin{bmatrix} 18 & 2 \\ -2 & 6 \end{bmatrix} = -112$$

$$A^{33} = (-1)^{3+3} \begin{bmatrix} 18 & -2 \\ -2 & 10 \end{bmatrix} = 176$$

$$A_{ij} = \begin{bmatrix} A^{11} & A^{12} & A^{13} \\ A^{21} & A^{22} & A^{23} \\ A^{31} & A^{32} & A^{33} \end{bmatrix} = \begin{bmatrix} 104 & 40 & -32 \\ 40 & 248 & -112 \\ -32 & 112 & 176 \end{bmatrix}$$

c) La matriz inversa de A, (A^{-1}) se obtiene dividiendo cada elemento de A^{ji} por $|A|$ el resultado de cada división $\frac{A^{ji}}{|A|} = C^{ij}$ que son los coeficientes o elementos de la matriz variancia-covariancia.

En nuestro ejemplo la matriz inversa variancia-covariancia es el siguiente:

$$A^{-1} = \begin{bmatrix} \frac{104}{1,728} & \frac{40}{1,728} & \frac{-32}{1,728} \\ \frac{40}{1,728} & \frac{248}{1,728} & \frac{-112}{1,728} \\ \frac{-32}{1,728} & \frac{-112}{1,728} & \frac{176}{1,728} \end{bmatrix} = \begin{bmatrix} \underline{0.060185} & 0.023148 & -0.018519 \\ 0.023148 & \underline{0.143519} & -0.064815 \\ -0.018519 & -0.064815 & \underline{0.101852} \end{bmatrix}$$

5. Cálculo de las constantes

$$\begin{aligned} \hat{\mu} &= \sum_j (C^{1j} Y_j) = (0.060185)(94) + (0.023148)(-14) + (-0.018519)(18) = 5 \\ \hat{r}_1 &= \sum_j (C^{2j} Y_{1j}) = (0.023148)(94) + (0.143519)(-14) + (-0.064815)(18) = -1 \\ \hat{r}_2 &= \sum_j (C^{3j} Y_{2j}) = (-0.018519)(94) + (-0.064815)(-14) + (0.101852)(18) = 1 \\ \hat{r}_3 &= - (r_1 + r_2) = - (-1+1) = 0 \end{aligned}$$

6. Análisis de variancia

a) SC de reducción debida a todas las constantes de ajustes $R(\mu, r_1), R(\mu, r_1) = 5(94) - 1(-14) + 1(18) = 502$

b) SC error = $\sum_i \sum_j Y_{ij}^2 - R(\mu, r_1) = 568 - 502 = 66$

c) SC ración = $B' Z^{-1} B =$

$$\begin{array}{c} \overline{\quad} \\ -1 \quad 1 \\ \overline{\quad} \end{array} \begin{bmatrix} 0.143519 & -0.064815 \\ -0.064815 & 0.101852 \end{bmatrix} \begin{array}{c} -1 \\ 1 \end{array}$$

$$\begin{array}{c} \overline{\quad} \\ -1 \quad 1 \\ \overline{\quad} \end{array} \begin{bmatrix} \frac{-0.101852}{0.0104167} & \frac{0.064815}{0.0104167} \\ \frac{0.064815}{0.0104167} & \frac{0.143519}{0.0104167} \end{bmatrix} \begin{array}{c} -1 \\ 1 \end{array}$$

$$\begin{array}{c} \overline{\quad} \\ -1 \quad 1 \\ \overline{\quad} \end{array} \begin{bmatrix} 9.7778 & 6.2222 \\ 6.2222 & 13.7778 \end{bmatrix} \begin{array}{c} -1 \\ 1 \end{array}$$

$$= \begin{bmatrix} -9.7778 + 6.2222 & -6.2222 + 13.7778 \end{bmatrix} \begin{array}{c} -1 \\ 1 \end{array}$$

$$= \begin{bmatrix} -3.5556 & 7.5556 \end{bmatrix} \begin{array}{c} -1 \\ 1 \end{array}$$

$$= 3.5556 + 7.5556$$

SC ración = 11.1112

FV	GL	SC	CM	F
Raciones	2	11.1112	5.5556	1.263 N S
Error	15	66.0000	4.4000	
Total	17	77.1112		

7. Errores estandares. El error estandar de un efecto, tal como \hat{r}_1 tiene poco significado por si mismo, puesto que no puede existir sin el promedio, así calcularemos el error estandar para el promedio de mínimos cuadrados $\hat{\mu}$ y el error estandar de $\hat{\mu} + \hat{r}_1$

$$S_{\hat{\mu}} = \sqrt{(0.060185)(4.4)} = 0.51$$

$$S_{\hat{\mu} + \hat{r}_1} = \sqrt{[(0.060185 + 0.143519 + 2 \times 0.023148)] (4.4)} = 1.05$$

$$S_{\hat{\mu} + \hat{r}_2} = \sqrt{[0.060185 + 0.101852 + 2(-0.018519)] 4.4} = 0.77$$

$$S_{\hat{\mu} + \hat{r}_3} = \sqrt{[0.060185 + 0.11574 + 2(-0.004629)] 4.4} = 0.86$$

Los elementos inversos para los coeficientes en la ecuación de mínimos cuadrados que fueron sustraídos más arriba son necesarios para obtener el error estandar de $\hat{\mu} + \hat{r}_3$. Esto se puede obtener en la siguiente manera:

$$C^{44} = 0.143519 + 0.101852 + (2)(-0.064815) = 0.115741$$

$$C^{14} = -(0.023148 - 0.018519) = -0.004629$$

8. Comparaciones individuales entre los r_i .

La conocida prueba de "t" se puede emplear para comparar los efectos de r_i de dos en dos, pero este método se vuelve poco satisfactorio cuando el número de tratamientos aumenta, la fórmula es la siguiente:

$$t = \frac{\hat{r}_i - \hat{r}_j}{S_{\hat{r}_i - \hat{r}_j}}$$

a) r_1 vs r_2

$$t = \frac{\frac{-1-1}{1.285}}{\sqrt{\frac{[0.143519 + 0.101852 - (2)(-0.064815)] 4.4}{(0.375) 4.4}}} = \frac{-2}{\sqrt{(0.375) 4.4}} =$$

$t = - 1.56$ NS

b) r_2 vs r_3

$$t = \frac{\frac{-1-0}{1.354}}{\sqrt{\frac{[0.143519 + 0.115741 - (2)(-0.078704)] 4.4}{0.416668 \times 4.4}}} = \frac{1}{\sqrt{0.416668 \times 4.4}} =$$

$t = - 0.74$ NS

c) r_2 vs r_3

$$t = \frac{\frac{1-0}{1.133}}{\sqrt{\frac{[0.101852 + 0.115741 - (2)(-0.037037)] 4.4}{(0.291667)(4.4)}}} = \frac{1}{\sqrt{(0.291667)(4.4)}} =$$

$t = 0.88$ NS

La prueba de rango múltiple de Duncan modificada por Kramer se puede emplear para hacer todas las comparaciones deseadas, para esto se hace uso de los elementos inversos y la desviación estandar del error. Si los valores de $\bar{Y}_i - \bar{Y}_j \sqrt{\frac{2}{C^{ii} + C^{jj} - 2C^{ij}}}$ son mayores que $\hat{\sigma}_e Z_p$, n_2 la diferencia es significativa; Z_p , n_2 es el valor de la amplitud total estudentizada en la tabla de Duncan (P.05 o P.01), p es el número de promedios en el rango escogido y n_2 el grado de libertad del error.

Vamos a ilustrar el método con el ejemplo a mano.

- i) Cálculo de la diferencia promedia entre los tratamientos a comparar.

Comparación	$\bar{Y}_i - \bar{Y}_j$	$d\bar{Y}_i - \bar{Y}_j$
r_2 vs r_3	$\bar{Y}_2 - \bar{Y}_3 = 6 - 5$	1
r_2 vs r_1	$\bar{Y}_2 - \bar{Y}_1 = 6 - 4$	2
r_3 vs r_1	$\bar{Y}_3 - \bar{Y}_1 = 5 - 4$	1

- ii) Cálculo del factor para la diferencia entre promedios, por la fórmula siguiente:

$$\sqrt{\frac{2}{c^{ii} + c^{jj} - 2c^{ij}}}$$

Comparación	$\sqrt{\frac{2}{c^{ii} + c^{jj} - 2c^{ij}}}$	Factor
r_2 vs r_3	$\sqrt{0.101852 + 0.115741 - (2)(-0.037037)}$	= 2.619
r_2 vs r_1	$\sqrt{0.143519 + 0.101852 - (2)(-0.64815)}$	= 2.309
r_3 vs r_2	$\sqrt{0.143519 + 0.115741 - (2)(-0.078704)}$	= 2.191

iii) Cálculo de la desviación estandar = $\sqrt{CM \text{ error}} =$
 $= \sqrt{4.4000} = 2.097659$

- iv) Obtener de la tabla de amplitud total estudentizada para pruebas de Duncan los valores de Z_p , n_2 con grado de

libertad del error n_1 y el número de promedio en el rango escogido n_2 , en nuestro ejemplo $n_1 = 15$ y $n_2 = 2$ y 3

para $n_1 = 15$ y $n_2 = 2$ el valor de $Z_p = 3.01$

para $n_1 = 15$ y $n_2 = 3$ el valor de $Z_p = 3.16$

Ahora podemos resumir en una tabla las comparaciones.

Comparación	$\bar{Y}_i - \bar{Y}_j$	$\sqrt{\frac{2}{C^{ii} + C^{jj} - 2C^{ij}}}$	Factor x Diferencia	$\hat{C}_e Z_{p, n_2}$
r_2 vs r_3	1	2.619	2.62	6.31
r_2 vs r_1	2	2.309	4.62	6.63
r_3 vs r_1	1	2.191	2.19	6.31

Todas las diferencias por supuesto no son significativas para este ejemplo, hemos presentado solamente para ilustrar el procedimiento.

9. Contrastes ortogonales

Las sumas de cuadrados requeridos para las pruebas de significancia por contrastes ortogonales entre un juego de constantes se obtienen por transformación de los segmentos matriz del elemento inverso correspondiente a dicho juego. Esta transformación es equivalente al cálculo de los elementos inversos correspondientes a la variancia-covariancia de los contrastes ortogonales escogidos. Estos elementos son más fácilmente calculados por medio de un simple procedimiento de multiplicación de matrices. Si T es el segmento transformado de la inversa, K es la transformación matriz y Z_A el segmento

cuadrado de la inversa variancia-covariancia con la columna e hileras adicional sumada, entonces se puede demostrar que $T = KZAK'$.

La transformación matriz K se obtiene de los coeficientes ortogonales que define el juego de contrastes ortogonales deseados.

Suponiendo que queremos hacer las siguientes comparaciones ortogonales entre las tres raciones de nuestro ejemplo.

$$\begin{array}{ccc} \underline{r_1} & \underline{r_2} & \underline{r_3} \\ 2 & -1 & -1 \\ 0 & 1 & -1 \end{array}$$

Entonces la transformación matriz

$$K = \frac{1}{4} \begin{bmatrix} 2 & -1 & -1 \\ 0 & 2 & -2 \end{bmatrix}$$

Se puede notar que la suma de los coeficientes por hileras en la transformación matriz es igual a cero.

La transformación del segmento completo de los elementos inversos para las \hat{r}_i y el cálculo de la suma de cuadrados para cada contraste se dan a continuación.

$$\begin{aligned} KZA &= \frac{1}{4} \begin{bmatrix} 2 & -1 & -1 \\ 0 & 2 & -2 \end{bmatrix} \begin{bmatrix} \underline{-0.143519} & -0.04815 & -0.078704 \\ -0.064815 & \underline{0.101852} & -0.037037 \\ -0.078704 & -0.037037 & \underline{0.115741} \end{bmatrix} = \\ &= \frac{1}{4} \begin{bmatrix} 0.430557 & -0.194445 & -0.236112 \\ 0.027778 & 0.277778 & -0.305556 \end{bmatrix} \end{aligned}$$

$$T = KZ_A K' = \frac{1}{4} \begin{bmatrix} 0.430557 & -0.194445 & -0.236112 \\ 0.027778 & 0.277778 & -0.305556 \end{bmatrix} \frac{1}{4} \begin{bmatrix} 2 & 0 \\ -1 & 2 \\ -1 & -2 \end{bmatrix}$$

$$= \frac{1}{16} \begin{bmatrix} \underline{1.291671} & 0.083334 \\ 0.083334 & \underline{1.166668} \end{bmatrix}$$

Un método más fácil de calcular T es por sustracción de los coeficientes en la última columna de la transformación matriz K de los coeficientes en las otras primeras columnas, a la matriz producto. En este caso, la última columna e hilera para Z no necesita ser sumada.

$$T = K_A Z K'_A = K Z_A K'$$

Las constantes para los dos contrastes ortogonales se obtienen de la transformación matriz y los $\hat{\tau}_i$ constantes, en la siguiente forma.

$$\frac{1}{4} \begin{bmatrix} 2 & -1 & -1 \\ 0 & 2 & -2 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \hat{c}_1 \\ \hat{c}_2 \end{bmatrix} ; \hat{c}_1 = \frac{-3}{4} \text{ y } \hat{c}_2 = \frac{1}{2}$$

Las dos constantes \hat{c}_1 y \hat{c}_2 son los estimadores de las constantes del contraste.

Las sumas de cuadrados para los dos contrastes se obtienen de $B'Z^{-1}B$ en la forma usual.

$$\text{SC ración 1 vs raciones 2 y 3} = \frac{\left(\frac{-3}{4}\right)^2 (16)}{1.291671} = 6.9677$$

$$\text{SC ración 2 vs ración 3} = \frac{\left(\frac{1}{2}\right)^2 (16)}{1.166668} = 3.4286$$

La suma de estos dos componentes de la SC de raciones no es igual a la SC de raciones (11.1112) que hemos presentado en el cuadro preliminar de análisis de variancia. Esto se debe a la desigualdad del número de observaciones entre las raciones.

10. Un método corto para calcular la matriz inversa variancia-covariancia y las constantes estimadas es el siguiente.

Cuando las constantes se ajustan para todos los grados de libertad, entre un juego de clases o sub-clases la matriz inversa variancia-covariancia y las constantes estimadas se pueden obtener más fácilmente con la ayuda de una transformación matriz. Si llamamos D a la diagonal de la matriz de coeficiente de mínimos cuadrados para las clases o sub-clases y K es la transformación matriz, entonces la matriz inversa variancia-covariancia C se calcula por la fórmula $KD^{-1}K'$. Las constantes estimadas se obtienen de $KS = B$ donde S es una columna vector de los promedios de clase o sub-clase y B es una columna vector de las constantes. Aplicando este método a nuestro ejemplo, tenemos que la diagonal de la matriz de coeficientes de mínimos cuadrados es:

$$D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

Estos son los miembros izquierdos de las ecuaciones de mínimos cuadrados.

La transformación matriz es:

$$K = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ -1 & 2 & -1 \end{bmatrix}$$

Puesto que esta es la relación funcional de $\mu + r_i$ el promedio de raciones y las dos clases de contrastes μ y r_i es el siguiente:

$$\hat{\mu} = \frac{\hat{s}_1 + \hat{s}_2 + \hat{s}_3}{3}$$

$$\hat{r}_1 = \hat{s}_1 - \frac{\hat{s}_1 + \hat{s}_2 + \hat{s}_3}{3} = \frac{2}{3} \hat{s}_1 - \frac{1}{3} \hat{s}_2 - \frac{1}{3} \hat{s}_3$$

$$\hat{r}_2 = \hat{s}_2 - \frac{\hat{s}_1 + \hat{s}_2 + \hat{s}_3}{3} = -\frac{1}{3} \hat{s}_1 + \frac{2}{3} \hat{s}_2 - \frac{1}{3} \hat{s}_3$$

Si D es una diagonal de la matriz, su inversa consiste solamente en la recíproca de los elementos diagonales. De ahí que la matriz inversa variancia-covariancia y las constantes estimadas se pueden calcular individualmente en la siguiente forma.

	μ	r_1	r_2	Promedio de raciones	D^{-1}	Constantes
s_1	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{3}$	4	0.250000	5
s_2	$\frac{1}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$	6	0.125000	-1
s_3	$\frac{1}{3}$	$\frac{1}{3}$	$-\frac{1}{3}$	5	0.166667	1

Las constantes (5, -1, 1) se calcularon por multiplicación de cada una de las columnas en el transpuesto de la matriz de transformación por el promedio de raciones correspondientes.

$$\hat{\mu} = \frac{1}{3} (4 + 6 + 5) = 5$$

$$\hat{r}_1 = \frac{1}{3} [(2)(4) - 6 - 5] = -1$$

$$\hat{r}_2 = \frac{1}{3} [-4 + (2)(6) - 5] = 1$$

Si $\hat{r}_1 + \hat{r}_2 + \hat{r}_3 = 0$, de ahí se puede fácilmente calcular \hat{r}_3 .

Los elementos inversos de la matriz variancia-covariancia se obtienen por multiplicación de cada una de las columnas del transpuesto de la matriz de transformación por sí misma y con cada una de las otras columnas por vez, luego multiplicar este producto por la columna D^{-1} , por ejemplo:

$$c^{11} = \frac{1}{9} (0.250000 + 0.125000 + 0.166667) = 0.060185$$

$$c^{12} = \frac{1}{9} \quad 2 (0.250000) - 0.125000 - 0.166667 \quad = 0.023148$$

$$c^{13} = \frac{1}{9} \quad (-0.250000) + 0.125000 (2) - 0.166667 \quad = -0.018519$$

etc., seguir así sucesivamente hasta completar el cálculo de todos los elementos inversos.

$$C = \begin{bmatrix} \underline{0.069185} & 0.023148 & -0.018519 \\ 0.023148 & \underline{0.143519} & -0.064815 \\ -0.018519 & -0.064815 & \underline{0.101852} \end{bmatrix}$$

Se puede notar que estos elementos concuerdan con los dados anteriormente.

Un criterio de clasificación con regresión o covariancia

El análisis de covariancia de los datos con un criterio de clasificación se puede lograr por el método estandar, aunque existan números desiguales de clase a clase. Sin embargo resultados idénticos se pueden obtener por el procedimiento de mínimos cuadrados que implica el uso de la matriz. Este procedimiento es aplicable a una amplia variedad de problemas.

Modelo matemático

El modelo matemático para datos con un criterio de clasificación y con una variable continua independiente "x" es $Y_{ij} = \mu + a_i + b(x_{ij} - \bar{X}) + e_{ij}$ donde:

Y_{ij} = j^{th} observaciones en la i^{th} clase A

μ = promedio general de Y_{ij} con igual frecuencia en cada una de las clases de

a_i = efecto i^{th} de la clase A

b = regresión parcial de Y sobre x

X_{ij} = variable continua independiente

\bar{X} = promedio aritmético de X_{ij}

e_{ij} = error al azar

o lo mismo, se puede usar el siguiente modelo.

$$Y_{ij} = \alpha + a_i + bX_{ij} + e_{ij}$$

El símbolo μ es el promedio de la población cuando $X = 0$ o sea

$$\hat{\mu} = \hat{\alpha} + b\bar{X}$$

Ecuaciones de mínimos cuadrados

Las ecuaciones de mínimos cuadrados para datos de un criterio de clasificación con una variable independiente continua es el siguiente.

	$\hat{\alpha}$	\hat{a}_i	\hat{b}	RHM
α :	n	n_i	X	Y
a_i :	n_i	$\sum n_i^0$	X_i	Y_i
b :	X	X_i	$\sum \sum X_{ij}^2$	$\sum \sum X_{ij} Y_{ij}$

Si X_{ij} se expresa como desviaciones de \bar{X} entonces la ecuación para b sería la siguiente:

$$\sum_l \sum_j (X_{ij} - \bar{X}) \hat{a}_l + \sum_l \sum_j (X_{ij} - \bar{X})^2 \hat{b} = \sum_l \sum_j (X_{ij} - \bar{X}) Y_{ij}$$

y la estimación de μ se hace directamente en vez de ∞

Restricciones impuestas

Las restricciones impuestas son las mismas ya vistas en el párrafo anterior para un criterio de clasificación, sin análisis de regresión es decir $\sum_l \hat{a}_l = 0$.

Inversión de la matriz reducida y solución de las ecuaciones

Cuando las constantes para todos los grados de libertad entre una serie de clases o sub-clases están siendo ajustadas, además para la regresión (o regresiones), se dispone de un procedimiento alternativo para obtener la matriz inversa variancia-covariancia. En este caso la regresión o regresiones están siendo ajustadas sobre una base dentro de las clases o sub-clases. Por ejemplo en nuestro caso el estimador de b es \hat{b} .

$$\hat{b} = \frac{\sum_l \sum_j X_{ij} Y_{ij} - \frac{\sum_l X_l Y_l}{n_l}}{\sum_l \sum_j X_{ij}^2 - \sum_l \frac{X_l^2}{n_l}}$$

Los elementos inversos de la diagonal de la matriz variancia-covariancia correspondiente a \hat{b} es:

$$\frac{1}{\sum_l \sum_j X_{ij}^2 - \sum_l \frac{X_l^2}{n_l}}$$

El problema está en ajustar los elementos inversos del coeficiente de la diagonal de la matriz de $\infty + a_1$ y entonces transformar esta matriz inversa ajustada, con la transformación matriz apropiada. Para obtener los elementos inversos de la sección $\infty + a_1$, primero es necesario obtener los elementos inversos de la matriz variancia-covariancia para la b - columna o hilera. Estos elementos se refieren a lo mismo de la sección G de la inversa variancia-covariancia y se obtiene de la siguiente manera.

$$G^{ib} = G^{bi} = - \frac{X_i}{n_1 \left[\sum_I \sum_J X_{ij}^2 - \sum \frac{X_i^2}{n_1} \right]}$$

Los elementos inversos de la diagonal para la sección $\infty + a_1$ de A se calculan como sigue:

$$A^{ii} = \frac{1}{n_1} (1 - X_i G^{bi})$$

y los elementos inversos fuera de la diagonal para $\infty + a_1$ están dados por:

$$A^{ij} = - \frac{1}{n_1} (X_i G^{bj})$$

$$A^{ij} = - \frac{1}{n_j} (X_j G^{bi})$$

La matriz inversa variancia-covariancia para ∞ y a_1 se puede obtener separadamente del producto de las matrices $K A K'$, donde K es la matriz de relación funcional entre las clases ($\infty + a_1$) y las constantes individuales (∞ y a_1) y A es la matriz inversa para $\infty + a_1$. Con datos de un criterio de clasificación, como se demostró en

el ejemplo numérico del párrafo anterior, K es igual a:

$$K = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ \frac{p-1}{p} & -\frac{1}{p} & \frac{1}{p} & \dots & -\frac{1}{p} \\ -\frac{1}{p} & \frac{p-1}{p} & \frac{1}{p} & \dots & -\frac{1}{p} \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ -\frac{1}{p} & -\frac{1}{p} & -\frac{1}{p} & -\frac{p-1}{p} & -\frac{1}{p} \end{bmatrix}$$

Los estimadores de las constantes, $\hat{\alpha}$ y \hat{a}_1 , también se pueden obtener de dos maneras. Multiplicando la matriz inversa para las $\alpha + a_1$ y b por el apropiado RHM's, por ejemplo Y_1 y $\sum_l \sum_j X_{lj} Y_{lj}$, una hilera o columna por vez, para obtener $\hat{\alpha} + \hat{a}_1$. Entonces imponiendo la restricción que $\sum_l a_l = 0$, los estimadores de α se obtienen por la fórmula:

$$\hat{\alpha} = \frac{\sum_l (\hat{\alpha} + \hat{a}_l)}{p}$$

y los \hat{a}_1 son entonces calculados en forma usual.

El segundo método para calcular $\hat{\alpha}$ y los \hat{a}_1 proviene de la matriz inversa completa de variancia-covariancia para α , a_1 y b y también los RHM's para la matriz reducida de mínimos cuadrados. Los elementos inversos obtenidos arriba (los G's) para los b-hileras y columnas pueden ser transformados ya que ellos corresponderían a la matriz variancia-covariancia para α , a_1 y b mas bien que para $\alpha +$

+a_i y b. Estos se obtienen fácilmente de KG, donde G es la columna vector de los G^{ib}.

Cálculo de las sumas de cuadrados y errores estandares

La suma de cuadrados de error es igual a:

$$\sum_l \sum_j Y_{lj}^2 - R(\infty, a_i, b)$$

La reducción debido a todas las constantes de ajustes R(∞, a_i, b) es igual a R(μ, a_i, b) y se obtiene de la siguiente manera.

$$\hat{\alpha}Y. + \sum_l \hat{a}_l Y_l + \hat{b} \sum_l \sum_j X_{lj} Y_{lj}$$

o más fácilmente.

$$\hat{\alpha}Y. + \sum_{l=1}^{p-1} \hat{a}_l (Y_l - Y_p) + \hat{b} \sum_l \sum_j X_{lj} Y_{lj}$$

Si las X_{lj} han sido expresadas como desviaciones de X, entonces μ̂ se obtiene directamente en vez de ∞ y

$$\begin{aligned} R(\mu, a_i, b) &= R(\infty, a_i, b) = \hat{\mu}Y + \sum_l a_l Y_l + \hat{b} \sum_l \sum_j (X_{lj} - \bar{X}) Y_{lj} \\ &= \hat{\alpha}Y + \sum_{l=1}^{p-1} \hat{a}_l (Y_l - Y_p) + \hat{b} \sum_l \sum_j (X_{lj} - \bar{X}) Y_{lj} \end{aligned}$$

Con datos de un criterio de clasificación con covariancia, la suma de cuadrados de error se puede obtener más directamente, por supuesto de:

$$\sum_l \sum_j Y_{lj}^2 - \sum_l \frac{F_l^2}{n_l} - b^2 \left(\sum_l \sum_j X_{lj}^2 - \sum_l \frac{X_l^2}{n_l} \right)$$

desde que R(μ, a_i, b) es igual a los dos últimos términos en esta fórmula. La suma de cuadrados entre las clases A, ajustadas por las variaciones en el promedio de x, se puede obtener de B'Z⁻¹B como se ha explicado en las secciones previas.

La suma de cuadrados para la prueba de significancias del coeficiente de regresión se obtiene en la misma manera de $B'Z^{-1}B$.

$$\hat{b}^2 \left(\sum_k \sum_j X_{ij}^2 - \sum_k \frac{X_i^2}{n_i} \right)$$

El error estandar de los promedios de mínimos cuadrados para las clases A, $\hat{\mu} + \hat{a}_i$ se puede obtener de la matriz inversa variancia-covariancia para ∞ , ai y b, por la fórmula siguiente:

$$S_{\hat{\mu} + \hat{a}_i} = \sqrt{(C^{aa} + \bar{X} C^{bb} + C^{ai ai} = 2\bar{X} C^{\infty b} + 2C^{\infty ai} + 2\bar{X} C^{bai}) \hat{\sigma}^2}$$

donde:

$$\hat{\sigma}^2 = \text{CM del error}$$

$$C'_g = \text{elementos inversos}$$

Los elementos inversos que hubieran sido obtenidos para μ - hileras (o columnas). Si las X'_g hubieran sido tomadas como desviaciones de \bar{X} . Se puede calcular si se desea en la forma siguiente:

$$C^{\mu\mu} = C^{\infty\infty} + 2\bar{X} C^{\infty b} + \bar{X}^{-2} C^{bb}$$

$$C^{\mu ai} = C^{\infty ai} + \bar{X} C^{bai}$$

Si se dispone de esto, los errores estandares de $\hat{\mu} + \hat{a}_i$ pueden ser fácilmente calculados por la fórmula

$$S_{\hat{\mu} + \hat{a}_i} = \sqrt{(C^{\mu\mu} + C^{ai ai} + 2 C^{\mu ai}) \hat{\sigma}^2}$$

El cálculo de los errores estandares de los promedios no necesita de los elementos inversos de μ a una hilera (o columna), puesto que:

$$S(\hat{\mu} + \hat{a}_i) - (\hat{\mu} + \hat{a}_i) = S_{a_i} - a_j = \sqrt{(C^{ai ai} + C^{aj aj} - 2C^{ai aj}) \hat{\sigma}^2}$$

Ejemplo numérico. Vamos a usar las mismas series de datos del ejemplo anterior, con la adición de una variable calibradora (W) que suponemos sea el peso inicial de cada cerdo.

	R a c i ó n					
	1		2		3	
	W	Y	W	Y	W	Y
	5	3	4	5	8	7
	9	5	7	6	7	6
	11	6	0	2	3	4
	3	2	8	7	2	3
			10	8	8	6
			2	3	2	4
			12	9		
			5	8		
Total	28	16	48	48	30	30
n_i	4		8		6	
\bar{X}	7	4	6	6	5	5

1. Modelo matemático.

$$Y_{ij} = \alpha C + r_i + b w_{ij} + e_{ij}$$

2. Ecuaciones de mínimos cuadrados.

	$\hat{\alpha}$	\hat{r}_1	\hat{r}_2	\hat{r}_3	\hat{b}	RHM
α :	18	4	8	6	106	94
r_1 :	4	4	0	0	28	16
r_2 :	8	0	8	0	48	48
r_3 :	6	0	0	6	30	30
b :	106	28	48	30	832	656

3. Imponiendo la restricción $\sum_i \hat{r}_i = 0$

a) Reducción de hileras.

	$\hat{\alpha}$	\hat{r}_1	\hat{r}_2	\hat{b}	RHM
α :	18	-2	2	106	94
r_1 :	4	4	0	28	16
r_2 :	8	0	8	48	48
r_3 :	6	-6	-6	30	30
b :	106	-2	18	832	656

b) Reducción de columnas. Efectuando esta reducción se tienen las ecuaciones reducidas de mínimos cuadrados.

	$\hat{\alpha}$	\hat{r}_1	\hat{r}_2	\hat{b}	RHM
α :	18	-2	2	106	94
r_1 :	-2	10	6	-2	-14
r_2 :	2	6	14	18	18
b :	106	-2	18	832	656

4) Inversa de las ecuaciones reducidas de mínimos cuadrados.

a) Cálculo del determinante de la matriz.

$$\begin{aligned}
 |A| &= 18 \begin{vmatrix} 10 & 6 & -2 \\ 6 & 14 & 18 \\ -2 & 18 & 832 \end{vmatrix} + 2 \begin{vmatrix} -2 & 6 & -2 \\ 2 & 14 & 18 \\ 106 & 18 & 832 \end{vmatrix} + 2 \begin{vmatrix} -2 & 10 & -2 \\ 2 & 6 & 18 \\ 106 & -2 & 832 \end{vmatrix} - 106 \begin{vmatrix} -2 & 10 & 6 \\ 2 & 6 & 14 \\ 106 & -2 & 18 \end{vmatrix} \\
 &= 18 \left(\begin{vmatrix} 14 & 18 \\ 18 & 832 \end{vmatrix} - 6 \begin{vmatrix} 6 & 18 \\ -2 & 832 \end{vmatrix} - 2 \begin{vmatrix} 6 & 14 \\ -2 & 18 \end{vmatrix} \right) - 18 \left(113,240 - 30,168 - 272 \right) = 1,490,400 \\
 &+ 2 \left(\begin{vmatrix} 14 & 18 \\ 18 & 832 \end{vmatrix} - 6 \begin{vmatrix} 2 & 18 \\ 106 & 832 \end{vmatrix} - 2 \begin{vmatrix} 2 & 14 \\ 106 & 18 \end{vmatrix} \right) = 2 \left(-22,648 + 1,464 + 2,896 \right) = -36,296 \\
 &+ 2 \left(\begin{vmatrix} 6 & 18 \\ -2 & 832 \end{vmatrix} - 10 \begin{vmatrix} 2 & 18 \\ 106 & 832 \end{vmatrix} - 2 \begin{vmatrix} 2 & 6 \\ 106 & -2 \end{vmatrix} \right) = 2 \left(-10,096 + 2,440 + 1,880 \right) = -12,672 \\
 &- 106 \left(\begin{vmatrix} 6 & 14 \\ -2 & 18 \end{vmatrix} - 10 \begin{vmatrix} 2 & 14 \\ 106 & 18 \end{vmatrix} + 6 \begin{vmatrix} 2 & 6 \\ 106 & -2 \end{vmatrix} \right) = -106 \left(-272 + 14,480 - 3,840 \right) = 1,099,008 \\
 |A| &= 1,490,400 - 36,296 - 12,672 - 1,099,008 = + 342,144
 \end{aligned}$$

b) Cálculo de los cofactores.

$$\begin{aligned}
 A_{11} &= (-1)^{1+1} \begin{vmatrix} 6 & 14 & 18 \\ -2 & 18 & 832 \end{vmatrix} = 1 \left(\begin{vmatrix} 14 & 18 \\ 18 & 832 \end{vmatrix} - 6 \begin{vmatrix} 6 & 18 \\ -2 & 832 \end{vmatrix} - 2 \begin{vmatrix} 6 & 14 \\ -2 & 18 \end{vmatrix} \right) = 82,800 \\
 A_{12} &= (-1)^{1+2} \begin{vmatrix} -2 & 6 & -2 \\ 106 & 18 & 832 \end{vmatrix} = -1 \left(\begin{vmatrix} 14 & 18 \\ 18 & 832 \end{vmatrix} - 6 \begin{vmatrix} 2 & 18 \\ 106 & 832 \end{vmatrix} - 2 \begin{vmatrix} 2 & 14 \\ 106 & 18 \end{vmatrix} \right) = -18,288 \\
 A_{13} &= (-1)^{1+3} \begin{vmatrix} -2 & 10 & -2 \\ 2 & 6 & 18 \\ 106 & -2 & 832 \end{vmatrix} = 1 \left(\begin{vmatrix} 6 & 18 \\ -2 & 832 \end{vmatrix} - 10 \begin{vmatrix} 2 & 18 \\ 106 & 832 \end{vmatrix} - 2 \begin{vmatrix} 2 & 6 \\ 106 & 2 \end{vmatrix} \right) = -6,336 \\
 A_{14} &= (-1)^{1+4} \begin{vmatrix} -2 & 10 & 6 \\ 2 & 6 & 14 \\ 106 & -2 & 18 \end{vmatrix} = -1 \left(\begin{vmatrix} 6 & 14 \\ -2 & 18 \end{vmatrix} - 10 \begin{vmatrix} 2 & 14 \\ 106 & 18 \end{vmatrix} + 6 \begin{vmatrix} 2 & 6 \\ 106 & -2 \end{vmatrix} \right) = -10,368 \\
 A_{21} &= (-1)^{2+1} \begin{vmatrix} -2 & 2 & 106 \\ 6 & 14 & 18 \\ -2 & 18 & 832 \end{vmatrix} = -1 \left(\begin{vmatrix} 14 & 18 \\ 18 & 832 \end{vmatrix} - 2 \begin{vmatrix} 6 & 18 \\ -2 & 832 \end{vmatrix} + 106 \begin{vmatrix} 6 & 14 \\ -2 & 18 \end{vmatrix} \right) = 18,288 \\
 A_{22} &= (-1)^{2+2} \begin{vmatrix} 18 & 2 & 106 \\ 2 & 14 & 18 \\ 106 & 18 & 832 \end{vmatrix} = 1 \left(\begin{vmatrix} 14 & 18 \\ 18 & 832 \end{vmatrix} - 2 \begin{vmatrix} 2 & 18 \\ 106 & 832 \end{vmatrix} + 106 \begin{vmatrix} 2 & 14 \\ 106 & 18 \end{vmatrix} \right) = +50,832 \\
 A_{23} &= (-1)^{2+3} \begin{vmatrix} 18 & -2 & 106 \\ 2 & 6 & 18 \\ 106 & -2 & 832 \end{vmatrix} = -1 \left(\begin{vmatrix} 6 & 18 \\ -2 & 832 \end{vmatrix} + 2 \begin{vmatrix} 2 & 18 \\ 106 & 832 \end{vmatrix} + 106 \begin{vmatrix} 2 & 6 \\ 106 & -2 \end{vmatrix} \right) = -22,176 \\
 A_{24} &= (-1)^{2+4} \begin{vmatrix} 18 & -2 & 2 \\ 2 & 6 & 14 \\ 106 & -2 & 18 \end{vmatrix} = 1 \left(\begin{vmatrix} 18 & 14 \\ 18 & -2 \end{vmatrix} + 2 \begin{vmatrix} 2 & 14 \\ 106 & 18 \end{vmatrix} + 2 \begin{vmatrix} 2 & 6 \\ 106 & -2 \end{vmatrix} \right) = -1,728 \\
 A_{31} &= (-1)^{3+1} \begin{vmatrix} -2 & 2 & 106 \\ 10 & 6 & -2 \\ -2 & 18 & 832 \end{vmatrix} = 1 \left(\begin{vmatrix} 6 & -2 \\ 18 & 832 \end{vmatrix} - 2 \begin{vmatrix} 10 & -2 \\ -2 & 832 \end{vmatrix} + 106 \begin{vmatrix} 10 & 6 \\ -2 & 18 \end{vmatrix} \right) = -6,336 \\
 A_{32} &= (-1)^{3+2} \begin{vmatrix} 18 & 2 & 106 \\ -2 & 6 & -2 \\ 106 & 18 & 832 \end{vmatrix} = -1 \left(\begin{vmatrix} 6 & -2 \\ 18 & 832 \end{vmatrix} - 2 \begin{vmatrix} -2 & -2 \\ 106 & 832 \end{vmatrix} + 106 \begin{vmatrix} -2 & 6 \\ 106 & 18 \end{vmatrix} \right) = -22,176 \\
 A_{33} &= (-1)^{3+3} \begin{vmatrix} 18 & -2 & 106 \\ -2 & 10 & -2 \\ 106 & -2 & 832 \end{vmatrix} = 1 \left(\begin{vmatrix} 10 & -2 \\ -2 & 832 \end{vmatrix} + 2 \begin{vmatrix} -2 & -2 \\ 106 & 832 \end{vmatrix} + 106 \begin{vmatrix} -2 & 10 \\ 106 & -2 \end{vmatrix} \right) = 34,848 \\
 A_{34} &= (-1)^{3+4} \begin{vmatrix} 18 & -2 & 2 \\ -2 & 10 & 6 \\ 106 & -2 & 18 \end{vmatrix} = -1 \left(\begin{vmatrix} 10 & 6 \\ -2 & 18 \end{vmatrix} + 2 \begin{vmatrix} -2 & 6 \\ 106 & 18 \end{vmatrix} + 2 \begin{vmatrix} -2 & 10 \\ 106 & -2 \end{vmatrix} \right) = 0,000 \\
 A_{41} &= (-1)^{4+1} \begin{vmatrix} -2 & 2 & 106 \\ 10 & 6 & -2 \\ 6 & 14 & 18 \end{vmatrix} = -1 \left(\begin{vmatrix} 6 & -2 \\ 14 & 18 \end{vmatrix} - 2 \begin{vmatrix} 10 & -2 \\ 6 & 18 \end{vmatrix} + 106 \begin{vmatrix} 10 & 2 \\ 6 & 14 \end{vmatrix} \right) = -10,368 \\
 A_{42} &= (-1)^{4+2} \begin{vmatrix} 18 & 2 & 106 \\ -2 & 6 & -2 \\ 2 & 14 & 18 \end{vmatrix} = 1 \left(\begin{vmatrix} 6 & -2 \\ 14 & 18 \end{vmatrix} - 2 \begin{vmatrix} -2 & -2 \\ 2 & 18 \end{vmatrix} + 106 \begin{vmatrix} -2 & 6 \\ 2 & 14 \end{vmatrix} \right) = -1,728 \\
 A_{43} &= (-1)^{4+3} \begin{vmatrix} 18 & -2 & 106 \\ -2 & 10 & -2 \\ 2 & 6 & 18 \end{vmatrix} = -1 \left(\begin{vmatrix} 10 & -2 \\ 6 & 18 \end{vmatrix} + 2 \begin{vmatrix} -2 & -2 \\ 2 & 18 \end{vmatrix} + 106 \begin{vmatrix} -2 & 10 \\ 2 & 6 \end{vmatrix} \right) = 0,000 \\
 A_{44} &= (-1)^{4+4} \begin{vmatrix} 18 & -2 & 2 \\ -2 & 10 & 6 \\ 2 & 6 & 14 \end{vmatrix} = 1 \left(\begin{vmatrix} 10 & 6 \\ 6 & 14 \end{vmatrix} + 2 \begin{vmatrix} -2 & 6 \\ 2 & 14 \end{vmatrix} + 2 \begin{vmatrix} -2 & 10 \\ 2 & 6 \end{vmatrix} \right) = 1,728
 \end{aligned}$$

5. Cálculo de la matriz inversa.

$$A^{-1} = \begin{bmatrix} \frac{A_{11}}{A} & \frac{A_{12}}{A} & \frac{A_{13}}{A} & \frac{A_{14}}{A} \\ \frac{A_{21}}{A} & \frac{A_{22}}{A} & \frac{A_{23}}{A} & \frac{A_{24}}{A} \\ \frac{A_{31}}{A} & \frac{A_{32}}{A} & \frac{A_{33}}{A} & \frac{A_{34}}{A} \\ \frac{A_{41}}{A} & \frac{A_{42}}{A} & \frac{A_{43}}{A} & \frac{A_{44}}{A} \end{bmatrix} = \begin{bmatrix} c^{11} & c^{12} & c^{13} & c^{14} \\ c^{21} & c^{22} & c^{23} & c^{24} \\ c^{31} & c^{32} & c^{33} & c^{34} \\ c^{41} & c^{42} & c^{43} & c^{44} \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} \frac{82,800}{342,144} & \frac{18,288}{342,144} & \frac{-6,336}{342,144} & \frac{-10,368}{342,144} \\ \frac{18,288}{342,144} & \frac{50,832}{342,144} & \frac{-22,176}{342,144} & \frac{-1,728}{342,144} \\ \frac{-6,336}{342,144} & \frac{-22,176}{342,144} & \frac{34,848}{342,144} & \frac{0,000}{342,144} \\ \frac{-10,368}{342,144} & \frac{-1,728}{342,144} & \frac{0,000}{342,144} & \frac{1,728}{342,144} \end{bmatrix}$$

Matris inversa

	\hat{c}	\hat{r}_1	\hat{r}_2	b
c_3 :	0.242003	0.053451	-0.018519	-0.030303
r_1 :	0.053451	0.048569	-0.064815	-0.005051
r_2 :	-0.018519	-0.064815	0.101852	0.000000
r_3 :	-0.030303	-0.005051	0.000000	0.005051

6. Estimación de las constantes de ajustes.

$$\hat{\alpha} = 0.242003 \times 94 + 0.053451(-14) - 0.018519 \times 18 - 0.030303 \times 656 = \underline{1.7879}$$

$$\hat{r}_1 = 0.053451 \times 94 + 0.148569(-14) - 0.064815 \times 18 - 0.005051 \times 656 = \underline{-1.5354}$$

$$\hat{r}_2 = -0.018519 \times 94 - 0.064815(-14) + 0.101852 \times 18 + 0.000000 \times 656 = \underline{1.0000}$$

$$\hat{b} = -0.030303 \times 94 - 0.005051(-14) + 0.000000 \times 18 + 0.005051 \times 656 = \underline{0.5354}$$

$$\hat{r}_3 = -(\hat{r}_1 + \hat{r}_2) = -(-1.5354 + 1.0000) = \underline{0.5354}$$

$$\hat{\mu} = \hat{\alpha} + \hat{b} \bar{w} = 1.7879 + (0.5354)(5.8889) = \underline{4.9408}$$

7. Elementos inversos de la diagonal para $\hat{\mu} + \hat{r}_1$ pues que

$\hat{\mu} + \hat{r}_1 = \hat{\alpha} + \hat{b}\bar{w} + \hat{r}_1$ los elementos inversos se pueden calcular de la siguiente manera:

$$C^{\mu+r_1} = C^{\alpha\alpha} + \frac{2}{\bar{w}} C^{bb} + C^{r_1 r_1} + 2 \bar{w} C^{\alpha b} + 2 C^{\alpha r_1} + 2 \bar{w} C^{b r_1}$$

Otro método para calcular el elemento inverso de la diagonal, que es necesario, si el error estandar de $\hat{\mu} + \hat{r}_1$ se desea, es posible obtenerlo de:

$$C^{\mu\mu} + C^{r_1 r_1} + 2 C^{\mu r_1}$$

Sin embargo los elementos inversos que contienen μ no son directamente utilizables a partir de la matriz inversa, a menos que los W_{ij} hayan sido expresados como desviaciones del promedio general. Estos elementos inversos para μ -hilera y columnas se pueden calcular de los elementos inversos disponibles, en la forma siguiente.

En nuestro caso

$$C^{\mu\mu} = 0.242003 + (2)(5.88889)(-0.030303) + (5.88889)^2(0.005051) = \underline{0.060265}$$

$$C^{\mu r_1} = 0.053451 + (5.88889)(-0.005051) = \underline{0.023706}$$

$$C^{\mu r_2} = -0.018519 + (5.88889)(0.000000) = \underline{-0.018519}$$

Un procedimiento alternativo para calcular la matriz inversa variancia-covariancia más fácilmente es el siguiente, que se puede aplicar cuando las constantes para todos los grados de libertad entre las clases o sub-clases están siendo ajustadas.

Los elementos inversos para b , \hat{b} de la matriz inversa es igual a:

$$C^{bb} = \frac{1}{\sum_I \sum_J W_{ij} - \sum_I \frac{W_i}{n_i}} = \frac{1}{198} = \underline{0.00505051}$$

Este valor concuerda con el obtenido para este elemento en la inversión simultánea de la matriz completa. Los elementos inversos para las b -columna e hilera de $\alpha + r_i$ se obtiene en la siguiente manera:

$$G^{1b} = G^{b1} = -\frac{W_1}{n_1 \left[\sum_I \sum_J W_{ij} - \sum_I \frac{W_i}{n_i} \right]} = -(0.00505051) \left(\frac{28}{4} \right) = -\underline{0.03535357}$$

$$G^{2b} = G^{b2} = -(0.00505051) \left(\frac{48}{8} \right) = -\underline{0.03030306}$$

$$G^{3b} = G^{b3} = -(0.00505051) \left(\frac{30}{6} \right) = -\underline{0.02525255}$$

Los elementos inversos de la diagonal, ajustada por $\hat{a}_i + \hat{r}_i$ se calcula de las siguientes maneras:

$$A^{11} = \frac{1}{n_1} (1 - W_1 G^{b1}) = \frac{1}{4} \left[1 - (28)(-0.03434357) \right] = \underline{0.49747499}$$

$$A^{21} = \frac{1}{n_2} (1 - W_2 G^{b2}) = \frac{1}{8} \left[1 - (48)(-0.03030306) \right] = \underline{0.30681836}$$

$$A^{33} = \frac{1}{n_3} (1 - W_3 G^{b3}) = \frac{1}{6} \left[1 - (30)(-0.02525255) \right] = \underline{0.29292942}$$

Los elementos inversos fuera de la diagonal, ajustado por los

$\infty + r_i$ se calculan de la siguiente manera:

$$A^{12} = A^{21} = -\frac{1}{n_1} (w_1) G^{b2} = -\frac{1}{4}(28)(-0.03030306) = \underline{0.21212142}$$

$$A^{13} = A^{31} = -\frac{1}{n_1} (w_1) G^{b3} = -\frac{1}{4}(28)(-0.02525255) = \underline{0.17676785}$$

$$A^{23} = A^{32} = -\frac{1}{n_2} (w_2) G^{b3} = -\frac{1}{8}(48)(-0.02525255) = \underline{0.15151530}$$

La transformación del segmento $\infty + r_i$ de la matriz inversa se obtiene en la forma siguiente:

$$KAK' = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ -1 & 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} 0.49747499 & 0.21212142 & 0.17676785 \\ 0.21212142 & 0.30681836 & 0.15151530 \\ 0.17676785 & 0.15151530 & 0.29292942 \end{bmatrix} K'$$

$$= \frac{1}{3} \begin{bmatrix} 0.88636426 & 0.67045508 & 0.62121257 \\ 0.60606071 & -0.03409082 & -0.09090902 \\ -0.25000000 & 0.25000000 & -0.16666667 \end{bmatrix} \frac{1}{3} \begin{bmatrix} 1 & 2 & -1 \\ 1 & -1 & 2 \\ 1 & -1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} \underline{0.24200355} & 0.05345121 & -0.01851852 \\ 0.05345121 & \underline{0.14856903} & -0.06481481 \\ -0.01851852 & -0.06481481 & \underline{0.10185185} \end{bmatrix}$$

Estos elementos inversos coinciden con los obtenidos directamente por el procedimiento anterior de inversión de matriz.

Los elementos inversos para la b-columna e hilera se obtienen:

$$K G = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ -1 & 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} -0.03535357 \\ -0.03030306 \\ -0.02525255 \end{bmatrix} = \begin{bmatrix} -0.03030306 \\ -0.00505051 \\ 0.00000000 \end{bmatrix}$$

que también coincide con los valores para estos elementos que fueron obtenidos directamente.

8. El análisis de variancia.

a) SC de reducción debido a todas las constantes de ajuste. $R(\mu, r_i, b) = 1.7879(94) + (-1.5354)(-14) + (1.0000)(18) + (0.5354)(656) = \underline{558.7806}$

b) SC de error = $\sum_k \sum_j Y_{ij}^2 - R(\mu, r_i, b) = 568 - 558.7806 = \underline{9.2194}$.

c) SC de razones = $B' Z^{-1} B$

$$= \begin{bmatrix} -1.5354 & 1.0000 \\ -1.5353 & 1.0000 \end{bmatrix} \begin{bmatrix} 0.148569 & -0.064815 \\ -0.064815 & 0.101852 \\ 9.317664 & 5.929431 \\ 5.929431 & 13.591447 \end{bmatrix}^{-1} \begin{bmatrix} 1.5354 \\ 1.0000 \\ 1.5354 \\ 1.0000 \end{bmatrix}$$

$$= \begin{bmatrix} -8.376910 & 4.487399 \end{bmatrix} \begin{bmatrix} -1.5354 \\ 1.0000 \end{bmatrix}$$

SC razones = 17.3493

d) SC debido a regresión = $B' Z^{-1} B$

$$\frac{(0.5354)^2}{0.005051} = (0.5354)^2 (198) = \underline{56.7573}$$

El análisis de variancia es como sigue

	GL	SC	CM	F
Raciones	2	17.3493	8.6746	13.17**
Regresión	1	56.7573	56.7573	86.19**
Error	14	9.2194	0.6585	

Las diferencias entre raciones son ahora altamente significativas, lo que ignorando el peso inicial (sin covariancia), no fue significativa en el párrafo anterior. El error se redujo de 4.400 a 0.6585.

9. Errores estandares y comparaciones individuales.

a) Errores estandar del promedio.

$$S_{\hat{\mu}} = \sqrt{(0.060265)(0.6585)} = \underline{0.199}$$

$$S_{\hat{\mu} + \hat{r}_1} = \sqrt{[0.60265 + 0.148569 + 2(0.023706)] (0.6585)} = \underline{0.411}$$

$$S_{\hat{\mu} + \hat{r}_2} = \sqrt{[0.60265 + 0.101852 + (2)(-0.018519)] (0.6585)} = \underline{0.287}$$

$$S_{\hat{\mu} + \hat{r}_3} = \sqrt{[0.60265 + 0.120791 + (2)(-0.005187)] (0.6585)} = \underline{0.335}$$

b) Comparaciones individuales entre las \hat{r}_i por la prueba de "t". Esto se verifica en tres pasos.

(i) Cálculo de los errores estandares de las diferencias entre promedios.

$$S_{\hat{r}_1 - \hat{r}_2} = \sqrt{[0.148519 + 0.101852 - (2)(-0.064815)] (0.6585)} = \underline{0.5003}$$

$$S_{\hat{r}_1 - \hat{r}_3} = \sqrt{[0.148519 + 0.120791 - (2)(-0.083754)] (0.6585)} = \underline{0.5364}$$

$$S_{\hat{r}_2 - \hat{r}_3} = \sqrt{[0.101852 + 0.120791 - (2)(-0.037037)] (0.6585)} = 0.4420$$

(ii) Cálculo de los promedios ajustados.

	\bar{Y}_i	\bar{W}	\bar{W}	$(\bar{W} - \bar{W})$	b	$b(\bar{W} - \bar{W})$	$Y_i - b(\bar{W} - \bar{W}) =$ Promedio ajustado
r_1	4	7	6	1	0.5354	0.5354	$4 - 0.5354 = 3.4646$
r_2	6	6	6	0	0.5354	0.0000	$6 - 0.0000 = 6.0000$
r_3	5	5	6	-1	0.5354	-0.5354	$5 + 0.5354 = 5.5354$

(iii) Cálculo de "t"

$$r_1 \text{ vs } r_2 \quad t = \frac{3.4646 - 6.0000}{0.5003} = 5.068^{**}$$

$$r_1 \text{ vs } r_3 \quad t = \frac{3.4646 - 5.5354}{0.5364} = 3.861^{**}$$

$$r_2 \text{ vs } r_3 \quad t = \frac{6.0000 - 5.5354}{0.4420} = 1.051 \text{ N S}$$

c) Comparación entre promedios por la prueba de rango múltiple de Duncan (P.05). El mismo procedimiento y criterio de la sección precedente se aplica a este caso. A continuación se presenta el resumen de las comparaciones.

Comparación	$Y_i - Y_j$	$\sqrt{\frac{2}{C^{ii} + C^{jj} - 2C^{ij}}}$	Producto	$\hat{C} Z_p, n_2$
$r_2 \text{ vs } r_3$	0.4646	2.294	1.07	2.46
$r_2 \text{ vs } r_1$	2.5354	2.140	5.43*	2.58
$r_3 \text{ vs } r_1$	2.0708	2.596	5.28*	2.46

d) Contrastes ortogonales.

Suponiendo que las raciones sean función de una variable nos interesaría probar la significancia de los efectos lineales y cuadráticos entre los \hat{r}_i . En este caso la transformación matriz para el segmento r_i de la matriz inversa (incluyendo la tercera hilera y columna) es:

$$K = \frac{1}{4} \begin{bmatrix} 2 & 0 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

$$\begin{aligned}
 T = KZK' &= \frac{1}{4} \begin{bmatrix} 2 & 0 & -2 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 0.148569 & -0.064815 & -0.083654 \\ -0.064815 & 0.101852 & -0.037237 \\ -0.083754 & -0.037037 & 0.120791 \end{bmatrix} K' \\
 &= \frac{1}{4} \begin{bmatrix} 0.464646 & -0.055556 & -0.409090 \\ 0.194445 & -0.305556 & 0.111111 \end{bmatrix} \frac{1}{4} \begin{bmatrix} 2 & 1 \\ 0 & -2 \\ -2 & 1 \end{bmatrix} \\
 &= \frac{1}{16} \begin{bmatrix} 1.747472 & 0.166668 \\ 0.166668 & 0.916668 \end{bmatrix} \\
 &= \begin{bmatrix} \underline{0.109217} & 0.010417 \\ 0.010417 & \underline{0.057292} \end{bmatrix}
 \end{aligned}$$

$$\text{SCF linear} = \frac{\frac{1}{4}(2)(-1.5354) - \frac{1}{4}(2)(0.5353)^2}{0.109217} = \frac{(1.0354)^2}{0.109217} = 9.816^{**}$$

$$\text{SCF cuadrática} = \frac{\frac{1}{4}(-1.5354) - \frac{1}{4}(2)(1.0000) + \frac{1}{4}(5354)^2}{0.057292} = \frac{(-0.7500)^2}{0.057292} = 9.818^{**}.$$

Otra vez se nota que la suma de cuadrados de tratamientos no es igual a la SC para contrastes ortogonales.

DATOS CON DOS CRITERIOS DE CLASIFICACION

Los casos más simples de este problema ya los hemos tratado en los capítulos precedentes, ahora solamente nos referiremos al método de constantes de ajuste (mínimos cuadrados) para los casos generales de datos con dos criterios de clasificación sin interacción.

Modelo matemático

$$Y_{ijk} = \mu + a_i + b_j + e_{ij}$$

donde:

Y_{ijk} = la observación K^{th} en la j^{th} clase B y i^{th} clase A

μ = promedio general con número de sub-clases iguales

a_i = efecto de la i^{th} clase A

b_j = efecto de la j^{th} clase B

e_{ijk} = errores al azar, asumiendo que sea NID $(0, \hat{\sigma}^2)$

Los a_i y b_j pueden ser de efectos fijos o al azar. Si los efectos de a_i como los de b_j son fijos, el modelo se llama "fijo", o Modelo I de Eisenhart. Si ambos a_i y b_j se obtienen al azar de ciertas poblaciones infinitas tales como vacas, toros, etc. el modelo se llama al azar o Modelo II de Eisenhart. Si uno de los efectos se fija y el otro se obtiene al azar, el modelo se llama "mixto".

Si una clase de efectos es fijo entonces el investigador está interesado en los promedios de mínimos cuadrados, errores estandares, prueba de significancia, comparaciones de promedios, comparaciones ortogonales, etc.

Por otro lado si una clase de efectos es al azar, el investigador estaría principalmente interesado en el componente de variancia de esa fuente. El procedimiento general para las dos clases de efectos se presentará a continuación, por procedimiento de absorción que reduce los cálculos cuando los efectos al azar tienen un gran número de grados de libertad.

Ecuaciones de mínimos cuadrados

	$\hat{\mu}$	\hat{a}_i	\hat{b}_j	RHM
μ :	n..	ni.	n _j	Y..
a_i :	ni	$\circ ni^{\circ}$	n _{ij}	Y _i
b_j :	n.j	n _{ij}	$\circ nj^{\circ}$	Y _j

Imposición de restricciones

Dos clases de restricciones se pueden imponer a estas ecuaciones: una sería $a_p = 0 = b_p$ y despreciar las ecuaciones para a_p y b_p . Sin embargo es preferible imponer las restricciones de $\sum_i \hat{a}_i = 0 = \sum_j \hat{b}_j$ tal como lo veníamos haciendo, luego proceder a hacer las sustracciones necesarias dentro de cada serie de constantes a ajustar los mismos para la columna RHM, antes de invertir la matriz.

Inversión de la matriz reducida y solución de las ecuaciones

La inversión de la matriz para datos con dos criterios de clasificación no presentan ninguna dificultad, y se puede hacer en la forma usual y resolver las ecuaciones en la misma forma que acostumbramos resolverlas para estimar la constante μ , a_i y b_j . Las constantes a_p y b_p se obtienen en la forma siguiente:

$$\hat{a}_p = - \sum_i a_i$$

$$\hat{b}_p = - \sum_j b_j$$

Cálculos de las sumas de cuadrados para el análisis de variancia

a) SC de reducción debido a las constantes de ajustes = $R(\mu, ai, bj)$, $R(\mu, ai, bj) = \mu Y_{..} + \sum_{i=1}^{q-1} \hat{a}_i (Y_{i.} - Y_{..}) + \sum_{j=1}^{q-1} \hat{b}_j (Y_{.j} - Y_{..})$

b) SC error (cuando la interacción está ausente)

$$= \sum_i \sum_j \sum_k Y_{ijk}^2 - R(\mu, ai, bj)$$

c) SC ajustado de A. Cuando no existe interacción se puede calcular de dos maneras:

$$\begin{aligned} \text{SC ajustado de A} &= R(\mu, ai, bj) - R(\mu, bj) \\ &= R(\mu, ai, bj) - \sum_j \frac{Y_{.j}^2}{n_j} \end{aligned}$$

donde $\sum_j \frac{Y_{.j}^2}{n_j}$ es la SC no corregida de B

d) SC ajustado de B = $R(\mu, ai, bj) - R(\mu, ai)$

$$= R(\mu, ai, bj) - \sum_i \frac{Y_{i.}^2}{n_i}$$

e) Si existiera la interacción, la suma de cuadrados de ella se puede calcular indirectamente, aún cuando las constantes de interacción no estén ajustadas.

$$\text{SC interacción} = R[(\mu, ai, bj, (ab)_{ij})] - R(\mu, ai, bj)$$

$$= \sum_i \sum_j \frac{Y_{ij}^2}{n_{ij}} - R(\mu, ai, bj)$$

donde $\sum_i \sum_j \frac{Y_{ij}^2}{n_{ij}}$ es la suma de cuadrados no corregidas de las sub-clases AB. El mismo procedimiento es aplicable para modelos más complejos. Esto es aplicable a las constantes para todos los efectos entre

las sub-clases que son ajustadas, excepto la interacción de mayor orden. La SC para esta interacción se puede obtener indirectamente por diferencia de la SC entre sub-clase menos SC reducción para todas las constantes de ajustes. Un procedimiento indirecto, similar se puede usar para calcular la interacción cuando al mismo tiempo se ajustan por regresiones parciales para variables continuas.

La suma de cuadrados del error para probar la significancia de la interacción es:

$$\sum_i \sum_j \sum_k Y_{ijk}^2 - \sum_i \sum_j \frac{Y_{ij}^2}{n_{ij}}$$

que es la SC dentro de las sub-clases.

Si la interacción es significativa, las SC de los efectos principales están viciados y estos necesitan ser recalculados por el método de los cuadrados ponderados de los promedios con el cual estamos ya familiarizados.

Errores estandares, comparaciones ortogonales y otras comparaciones

El procedimiento para calcular lo expuesto arriba es el mismo descrito para los datos con un criterio de clasificación.

Estimación de los componentes de variancia

Cuando \hat{a}_i y \hat{b}_j son muestras al azar de poblaciones, los estimadores de los componentes de variancia asociados con estos efectos son de interés primordial. Con los datos de dos criterios de clasificación en ausencia de interacción el análisis de variancia es como sigue:

FV	GL	SC	CM	Componente estimado
A	p-1	$R(\mu, a_{ij}) - R(\mu, b_j)$	$\frac{SCA}{p-1}$	$\sigma_a^2 + k_2 \sigma_a^2$
B	q-1	$R(\mu, a_{ij}) - R(\mu, a_i)$	$\frac{SCB}{q-1}$	$\sigma_b^2 + k_1 \sigma_b^2$
Error	n-p-q+1	$\sum_{i,j,k}^2 Y_{ijk} - R(\mu, a_{ij})$	$\frac{SC \text{ error}}{n-p-q+1}$	σ_e^2

Los coeficientes de los componentes de variancia estimados se pueden calcular por dos métodos, uno directo y otro indirecto.

Método directo para calcular K's

Si las restricciones fueron impuestas que $\sum_i a_i = 0 = \sum_j b_j$ en las ecuaciones originales y las sumas de cuadrados se calcularon por medio de la fórmula $B'Z^{-1}B$. Los coeficientes de los componentes de variancia, K_1 y K_2 se calculan fácilmente de la siguiente manera:

$$K_1 = \frac{1}{q} \left[\left(\sum_I Z_B^{ii} - \frac{1}{q-1} \left(\sum \sum Z_B^{ij} \right) \right) \right]$$

$$K_2 = \frac{1}{p} \left[\left(\sum_I Z_A^{ii} - \frac{1}{p-1} \left(\sum_I \sum_J Z_A^{ij} \right) \right) \right]$$

El valor de Z_A^{ii} y Z_B^{ii} identifica los elementos en la matriz inversa para el segmento simétrico cuadrado de la matriz inversa variancia-covariancia. Para un diseño más complejo la fórmula general para el cálculo de los coeficientes es: $K = \frac{1}{m} \left(\sum Z^{ii} - \frac{1}{GL} \sum_I \sum_J Z^{ij} \right)$ donde m es el número de clases o sub-clases y GL grado de libertad para esa línea en el análisis de variancia.

Método indirecto para calcular K's

El método indirecto para calcular los coeficientes para σ_a^2 y σ_b^2 se reduce a los siguientes:

$$K_1 = \frac{1}{q-1} \left(n - \sum_i \frac{\sum_j n_{ij}^2}{n_i} \right)$$

$$K_2 = \frac{1}{p-1} \left(n - \sum_j \frac{\sum_i n_{ij}^2}{n_j} \right)$$

Hay otro método indirecto o de Henderson que veremos más adelante.

Absorción de la ecuación $\mu + ai$

Aunque los coeficientes de la matriz inversa variancia-covariancia; se requieren por el modelo completo, a menudo se puede obtener más fácilmente por un procedimiento de partición, que de la inversión directa de la serie de ecuaciones reducidas de mínimos cuadrados. La absorción de las ecuaciones $\mu + ai$ dentro de las ecuaciones b_j , es sólo un paso en este procedimiento de partición.

Se demostró con los datos de un criterio de clasificación que las ecuaciones para $\mu + ai$ son idénticas a las ecuaciones para ai . Cuando μ se combina con la ai es innecesario imponer la restricción sobre ai puesto que hay p grados de libertad asociado con $\mu + ai$ pero solo $p-1$ GL asociado con ai solo. Esto siendo verdadero para $\mu + ai = \frac{1}{n_i} (Y_i - \sum_j n_{ij} b_j)$ y como es fácil absorber las ecuaciones $\mu + ai$ por las b_j , se efectuará la absorción. Si los nuevos coeficientes para los b_j después de la absorción de los $\mu + ai$ son $C(b_j b_j)$ y $C(b_j b_j')$ y los nuevos RHM's son $S(b_j)$ las ecuaciones reducidas son como sigue.

	\hat{b}_1	\hat{b}_2	\hat{b}_3	\hat{b}_q	RHM
$b_1:$	$C(b_1 b_1)$	$C(b_1 b_2)$	$C(b_1 b_3)$	$\dots C(b_1 b_q)$	$S(b_1)$
$b_2:$	$C(b_2 b_1)$	$C(b_2 b_2)$	$C(b_2 b_3)$	$\dots C(b_2 b_q)$	$S(b_2)$
$b_3:$	$C(b_3 b_1)$	$C(b_3 b_2)$	$C(b_3 b_3)$	$\dots C(b_3 b_q)$	$S(b_3)$
.
.
.
.
$b_q:$	$C(b_q b_1)$	$C(b_q b_2)$	$C(b_q b_3)$	$\dots C(b_q b_q)$	$S(b_q)$

Los nuevos coeficientes y RHM's se calculan en la forma siguiente:

$$C(b_j b_j) = n_j - \sum_i \frac{n_{ij}^2}{n_i}$$

$$C(b_j b_{j'}) = C(b_j b_j) = - \sum_i \frac{n_{ij} n_{ij'}}{n_i} \text{ donde } j \neq j'$$

$$S(b_j) = Y_j - \sum_i \frac{n_{ij} Y_i}{n_i}$$

Ninguna restricción se necesita imponer a las ecuaciones de mínimos cuadrados hasta después de haber completado la absorción. Después de la absorción se notará que la suma de los nuevos coeficientes para b_j suman cero por hileras y columnas y que la suma de RHM's por columna es igual a cero. Esto provee un chequeo del proceso de absorción y también demuestra qué dependencias existen en las series de ecuaciones reducidas. Las dependencias se remueven

por imposición de la restricción que $\sum_j \hat{b}_j = 0$, donde los coeficientes \hat{b}_q se sustraen de los coeficientes de los otros \hat{b}_j y los coeficientes resultantes en las b ecuaciones, entonces se sustraen de los coeficientes resultantes en las otras b_j ecuaciones. La $S(b_q)$ también se sustrae de la otra $S(b_q)$ en cada una de las columnas RHM.

La inversa de la serie de ecuaciones reducidas para la b_j , después de la absorción de los $\mu + ai$ y luego de imponer la restricción apropiada es exactamente lo mismo que los elementos inversos del segmento $b_j \hat{b}_j$ en la matriz inversa de la matriz completa variancia-covariancia. Las constantes obtenidas para b_j son también idénticas a b_j que se obtendría resolviendo las ecuaciones originales.

La reducción en SC obtenida por multiplicación de \hat{b}_j por $S(b_j) - S(b_q)$ es igual a la SC para B.

$$\begin{aligned} SCB &= \sum_{j=1}^{q-1} \hat{b}_j [S(b_j) - S(b_q)] \\ &= \sum_j b_j S(b_j) \\ &= R(\mu, ai, bj) - R(\mu, ai) \end{aligned}$$

Los estimadores de $\mu + ai$ y los promedios de mínimos cuadrados para la clase A se obtiene por simple solución de las ecuaciones $\mu + ai$ después de calculadas las constantes \hat{b}_j . Por ejemplo:

$$\begin{aligned} \hat{\mu} + \hat{ai} &= \frac{1}{n_1} (Y_1 - \sum_j n_{1j} \hat{b}_j) \\ \hat{\mu} + \hat{ai} &= \frac{1}{n_i} (Y_i - \sum_j n_{ij} \hat{b}_j) \end{aligned}$$

Los elementos inversos que quedan después de la absorción a la matriz completa variancia-covariancia, se puede obtener fácilmente si se desea, por el procedimiento de multiplicación de matrices. Si

llamamos C a la matriz inversa remanente, D^{-1} a la inversa de la diagonal de la matriz para las clases $(\mu + a_i)$ y N_M a la matriz modificada de los coeficientes fuera de la diagonal, en las ecuaciones originales de mínimos cuadrados, entonces los elementos inversos para la sección n_{ij} (G) se pueden obtener de la fórmula $D^{-1} N_M C$.

$$\begin{bmatrix} \frac{1}{n_1} & 0 & \dots & \dots & 0 \\ 0 & \frac{1}{n_2} & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \frac{1}{n_p} \end{bmatrix} \begin{bmatrix} n'_{11} & n'_{12} & \dots & n'_{1(q-1)} \\ n'_{21} & n'_{22} & \dots & n'_{2(q-1)} \\ \vdots & \vdots & \ddots & \vdots \\ n'_{p1} & n'_{p2} & \dots & n'_{p(q-1)} \end{bmatrix} \begin{bmatrix} C^{11} & C^{12} & \dots & C^{1(q-1)} \\ C^{21} & C^{22} & \dots & C^{2(q-1)} \\ \vdots & \vdots & \ddots & \vdots \\ C^{(q-1)1} & C^{(q-1)2} & \dots & C^{(q-1)(q-1)} \end{bmatrix}$$

donde n'_{ij} es el resultado de la sustracción de la última columna de los n_{ij} de cada uno de los otros n_{ij} columnas, por ejemplo: $n'_{ij} = n_{ij} - n_{iq}$. Si \hat{b}_q fuera fijado igual a cero para obtener C , entonces la última columna de los n_{ij} se pueden despreciar en vez de sustraer como se indicó arriba. Sin embargo en este caso los elementos intersos obtenidos no se aplican directamente a las constantes como se demostró en el modelo (3). También las constantes que se obtienen así directamente son funciones lineares de μ , a_i y b_j más bien que las constantes por si mismas.

Los elementos inversos para la $\mu + a_i$ segmento cuadrado (A) se obtienen de la siguiente manera:

$$A = D^{-1} (I - N_M G')$$

donde I es la matriz de identidad y G' el transpuesto de G .

Cuando los efectos de a_i son al azar a menudo se desea obtener la estimación de la máxima probabilidad para las clases A. Henderson ha demostrado que la estimación de la máxima probabilidad se puede obtener por regresión de los promedios de mínimos cuadrados. La cantidad para el cual los promedios de mínimos cuadrados deben ser regresados, es una función de los correspondientes elementos de la diagonal inversa.

$$\begin{aligned} \mu + \hat{a}_i &\doteq \hat{\mu} + \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + A^{ii}\hat{\sigma}_e^2} (\hat{\mu} + \hat{a}_i - \hat{\mu}) \\ &\doteq \mu + \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + A^{ii}\hat{\sigma}_e^2} (\hat{a}_i) \end{aligned}$$

donde $\mu + a_i$ es el promedio de mínimos cuadrados regresados, por ejemplo, la probabilidad máxima estimada y la A^{ii} son los elementos inversos de la diagonal para las $\hat{\mu} + \hat{a}_i$. Henderson demuestra que la máxima probabilidad estimada se reduce a $\hat{\mu} + \hat{a}_i \doteq \hat{\mu} + \frac{nr}{1 + (n-1)r}(\hat{a}_i)$

Para datos con un criterio de clasificación, donde n es el número de observaciones en la i^{th} clases A y r es la simple correlación intraclase o repetibilidad.

Ejemplo numérico.

A continuación presentamos un ejemplo numérico, usando los mismos datos de las secciones previas, con ciertos reareglos.

RACION	PADRES N°		
	1	2	3
	5	2	3
	6	3	
1		5	
		6	
		7	
Sub total	11	23	3
Nº	2	5	1
	2	8	4
	3	8	4
2		9	6
			6
			7
Sub total	5	25	27
Nº	2	3	5
Total	16	48	30

1. Modelo matemático.

$$Y_{ijk} = \mu + s_i + r_j + e_{ijk}$$

2. Ecuación de mínimos cuadrados.

	$\hat{\mu}$	\hat{s}_1	\hat{s}_2	\hat{s}_3	\hat{r}_1	\hat{r}_2	RHM
μ :	<u>18</u>	4	8	6	8	10	94
s_1 :	4	<u>4</u>	0	0	2	2	16
s_2 :	8	0	<u>8</u>	0	5	3	48
s_3 :	6	0	0	<u>6</u>	1	5	30
r_1 :	8	2	5	1	<u>8</u>	0	37
r_2 :	10	2	3	5	0	<u>10</u>	57

3. Imponiendo la restricción de $\sum_i \hat{s}_i = 0$ y $\sum_j \hat{r}_j = 0$
 Reducción de hileras.

	$\hat{\mu}$	\hat{s}_1	\hat{s}_2	\hat{r}_1	RHM
μ :	18	-2	2	-2	94
s_1 :	4	4	0	0	16
s_2 :	8	0	8	2	48
s_3 :	-6	-6	-6	-4	30
r_1 :	8	1	4	8	37
r_2 :	10	-3	-2	-10	57

Reducción de columnas

	$\hat{\mu}$	\hat{s}_1	\hat{s}_2	\hat{r}_1	RHM
μ :	<u>18</u>	-2	2	-2	94
s_1 :	-2	<u>10</u>	6	4	-14
s_2 :	2	6	<u>14</u>	6	18
r_1 :	-2	4	6	<u>18</u>	-20

De esta manera, con estas series de ecuaciones simultáneas hemos formado una matriz cuadrada.

$$|A| = \begin{bmatrix} \underline{18} & -2 & 2 & -2 \\ -2 & \underline{10} & 6 & 4 \\ 2 & 6 & \underline{14} & 6 \\ -2 & 4 & 6 & \underline{18} \end{bmatrix}$$

En primer término vamos a calcular el determinante de la matriz, para el efecto podemos tomar cualquier hilera o columnas y proceder en la forma ya explicada en las clases extraordinarias de álgebra de matrices (resumen apéndice).

Cálculo del determinante de la matriz

$$|A| = 18 \begin{bmatrix} 10 & 6 & 4 \\ 6 & 14 & 6 \\ 4 & 6 & 18 \end{bmatrix} + 2 \begin{bmatrix} -2 & 6 & 4 \\ 2 & 14 & 6 \\ -2 & 6 & 18 \end{bmatrix} + 2 \begin{bmatrix} -2 & 10 & 4 \\ 2 & 6 & 6 \\ -2 & 4 & 18 \end{bmatrix} + 2 \begin{bmatrix} -2 & 10 & 6 \\ 2 & 6 & 14 \\ -2 & 4 & 6 \end{bmatrix} =$$

$$= 18 \left\{ 10 \begin{bmatrix} 14 & 6 \\ 6 & 18 \end{bmatrix} - 6 \begin{bmatrix} 6 & 6 \\ 4 & 18 \end{bmatrix} + 4 \begin{bmatrix} 6 & 14 \\ 4 & 6 \end{bmatrix} \right\} = 18(2,160 - 504 - 80) = 28,368$$

$$+ 2 \left\{ -2 \begin{bmatrix} 14 & 6 \\ 6 & 18 \end{bmatrix} - 6 \begin{bmatrix} 2 & 6 \\ -2 & 18 \end{bmatrix} + 4 \begin{bmatrix} 2 & 14 \\ -2 & 6 \end{bmatrix} \right\} = 2(-432 - 288 + 160) = -1,120$$

$$+ 2 \left\{ -2 \begin{bmatrix} 6 & 6 \\ 4 & 18 \end{bmatrix} - 2 \begin{bmatrix} 10 & 4 \\ 4 & 18 \end{bmatrix} - 2 \begin{bmatrix} 10 & 4 \\ 6 & 6 \end{bmatrix} \right\} = 2(-168 - 1,480 + 80) = -1,136$$

$$+ 2 \left\{ -2 \begin{bmatrix} 6 & 14 \\ 4 & 6 \end{bmatrix} - 10 \begin{bmatrix} 2 & 14 \\ -2 & 6 \end{bmatrix} + 6 \begin{bmatrix} 2 & 6 \\ -2 & 4 \end{bmatrix} \right\} = 2(40 - 400 + 120) = -480$$

$$|A| = 28,368 - 1,120 - 1,136 - 480 = \underline{25,632}$$

o lo mismo tomando la segunda hilera, $|A| =$

$$|A| = +2 \begin{vmatrix} -2 & 2 & -2 \\ 6 & 14 & 6 \\ 4 & 6 & 18 \end{vmatrix} + 10 \begin{vmatrix} 18 & 2 & -2 \\ 2 & 14 & 6 \\ -2 & 6 & 18 \end{vmatrix} - 6 \begin{vmatrix} 18 & -2 & -2 \\ 2 & 6 & 6 \\ -2 & 4 & 18 \end{vmatrix} + 4 \begin{vmatrix} 18 & -2 & 2 \\ 2 & 6 & 14 \\ -2 & 4 & 6 \end{vmatrix} = 25.632$$

tomando la tercera hilera

$$= 2 \begin{vmatrix} -2 & 2 & -2 \\ 10 & 6 & 4 \\ 4 & 6 & 18 \end{vmatrix} - 6 \begin{vmatrix} 18 & 2 & -2 \\ -2 & 6 & 4 \\ -2 & 6 & 18 \end{vmatrix} + 14 \begin{vmatrix} 18 & -2 & -2 \\ -2 & 10 & 4 \\ -2 & 4 & 18 \end{vmatrix} - 6 \begin{vmatrix} 18 & -2 & 2 \\ -2 & 10 & 6 \\ -2 & 4 & 6 \end{vmatrix} = 25.632$$

tomando la última hilera

$$= 2 \begin{vmatrix} -2 & 2 & -2 \\ 10 & 6 & 4 \\ 6 & 14 & 6 \end{vmatrix} + 4 \begin{vmatrix} 18 & 2 & -2 \\ -2 & 6 & 4 \\ 2 & 14 & 6 \end{vmatrix} - 6 \begin{vmatrix} 18 & -2 & -2 \\ 2 & 10 & 4 \\ 2 & 6 & 6 \end{vmatrix} + 18 \begin{vmatrix} 18 & -2 & 2 \\ -2 & 10 & 6 \\ 2 & 6 & 14 \end{vmatrix} = 25.632$$

Cálculo de los cofactores A_{ij}

$$A_{11} = (-1)^{1+1} \begin{vmatrix} 10 & 6 & 4 \\ 6 & 14 & 6 \\ 4 & 6 & 18 \end{vmatrix} = 1 \left\{ 10 \begin{vmatrix} 14 & 6 \\ 6 & 18 \end{vmatrix} - 6 \begin{vmatrix} 6 & 4 \\ 6 & 18 \end{vmatrix} + 4 \begin{vmatrix} 6 & 4 \\ 14 & 6 \end{vmatrix} \right\} = +1(2.160 - 504 - 80) = 1.576$$

$$= (-1)^{1+2} \begin{vmatrix} -2 & 6 & 4 \\ 2 & 14 & 6 \\ -2 & 6 & 18 \end{vmatrix} = -1 \left\{ -2 \begin{vmatrix} 14 & 6 \\ 6 & 18 \end{vmatrix} - 2 \begin{vmatrix} 6 & 4 \\ 6 & 18 \end{vmatrix} - 2 \begin{vmatrix} 6 & 4 \\ 14 & 6 \end{vmatrix} \right\} = -1(-432 - 168 + 40) = +560$$

$$= (-1)^{1+3} \begin{vmatrix} -2 & 10 & 4 \\ 2 & 6 & 6 \\ -2 & 4 & 18 \end{vmatrix} = 1 \left\{ -2 \begin{vmatrix} 6 & 6 \\ 4 & 18 \end{vmatrix} - 2 \begin{vmatrix} 10 & 4 \\ 4 & 18 \end{vmatrix} - 2 \begin{vmatrix} 10 & 4 \\ 6 & 6 \end{vmatrix} \right\} = +1(168 - 328 - 72) = -568$$

$$= (-1)^{1+4} \begin{vmatrix} -2 & 10 & 6 \\ 2 & 6 & 14 \\ -2 & 4 & 6 \end{vmatrix} = -1 \left\{ -2 \begin{vmatrix} 6 & 14 \\ 4 & 6 \end{vmatrix} - 2 \begin{vmatrix} 10 & 6 \\ 4 & 6 \end{vmatrix} - 2 \begin{vmatrix} 10 & 6 \\ 6 & 14 \end{vmatrix} \right\} = -1(40 - 72 - 208) = +240$$

$$= (-1)^{2+1} \begin{vmatrix} -2 & 2 & -2 \\ 6 & 14 & 6 \\ 4 & 6 & 18 \end{vmatrix} = -1 \left\{ -2 \begin{vmatrix} 14 & 6 \\ 6 & 18 \end{vmatrix} - 6 \begin{vmatrix} 2 & -2 \\ 6 & 18 \end{vmatrix} + 4 \begin{vmatrix} 2 & -2 \\ 14 & 6 \end{vmatrix} \right\} = -1(432 - 288 + 160) = +560$$

$$= (-1)^{2+2} \begin{vmatrix} 18 & 2 & -2 \\ 2 & 14 & 6 \\ -2 & 6 & 18 \end{vmatrix} = 1 \left\{ 18 \begin{vmatrix} 14 & 6 \\ 6 & 18 \end{vmatrix} - 2 \begin{vmatrix} 2 & -2 \\ 6 & 18 \end{vmatrix} - 2 \begin{vmatrix} 2 & -2 \\ 14 & 6 \end{vmatrix} \right\} = +1(3.888 - 96 - 80) = +3712$$

$$= (-1)^{2+3} \begin{vmatrix} 18 & -2 & -2 \\ 2 & 6 & 6 \\ -2 & 4 & 18 \end{vmatrix} = -1 \left\{ 18 \begin{vmatrix} 6 & 6 \\ 4 & 18 \end{vmatrix} - 2 \begin{vmatrix} -2 & -2 \\ 4 & 18 \end{vmatrix} - 2 \begin{vmatrix} -2 & -2 \\ 6 & 6 \end{vmatrix} \right\} = -1(1512 + 56 + 0) = -1568$$

$$= (-1)^{2+4} \begin{vmatrix} 18 & -2 & 2 \\ 2 & 6 & 14 \\ -2 & 4 & 6 \end{vmatrix} = 1 \left\{ 18 \begin{vmatrix} 6 & 14 \\ 4 & 6 \end{vmatrix} - 2 \begin{vmatrix} -2 & 2 \\ 4 & 6 \end{vmatrix} - 2 \begin{vmatrix} -2 & 2 \\ 6 & 14 \end{vmatrix} \right\} = +1(-360 + 40 + 80) = -240$$

$$= (-1)^{3+1} \begin{vmatrix} -2 & 2 & -2 \\ 10 & 6 & 4 \\ 4 & 6 & 18 \end{vmatrix} = 1 \left\{ -2 \begin{vmatrix} 6 & 4 \\ 6 & 18 \end{vmatrix} - 10 \begin{vmatrix} 2 & -2 \\ 6 & 18 \end{vmatrix} + 4 \begin{vmatrix} 2 & -2 \\ 6 & 4 \end{vmatrix} \right\} = +1(-168 - 480 + 80) = -568$$

$$= (-1)^{3+2} \begin{vmatrix} 18 & 2 & -2 \\ -2 & 6 & 4 \\ -2 & 6 & 18 \end{vmatrix} = -1 \left\{ 18 \begin{vmatrix} 6 & 4 \\ 6 & 18 \end{vmatrix} + 2 \begin{vmatrix} 2 & -2 \\ 6 & 18 \end{vmatrix} - 2 \begin{vmatrix} 2 & -2 \\ 6 & 4 \end{vmatrix} \right\} = -1(1512 + 96 - 40) = -1568$$

$$= (-1)^{3+3} \begin{vmatrix} 18 & -2 & -2 \\ -2 & 10 & 4 \\ -2 & 4 & 18 \end{vmatrix} = 1 \left\{ 18 \begin{vmatrix} 10 & 4 \\ 4 & 18 \end{vmatrix} + 2 \begin{vmatrix} -2 & -2 \\ 4 & 18 \end{vmatrix} - 2 \begin{vmatrix} -2 & -2 \\ 10 & 4 \end{vmatrix} \right\} = 1(2952 - 56 - 24) = +2872$$

$$= (-1)^{3+4} \begin{vmatrix} 18 & -2 & 2 \\ -2 & 10 & 6 \\ -2 & 4 & 6 \end{vmatrix} = -1 \left\{ 18 \begin{vmatrix} 10 & 6 \\ 4 & 6 \end{vmatrix} + 2 \begin{vmatrix} -2 & 2 \\ 4 & 6 \end{vmatrix} - 2 \begin{vmatrix} -2 & 2 \\ 10 & 6 \end{vmatrix} \right\} = -1(648 + 40 - 64) = -672$$

$$= (-1)^{4+1} \begin{vmatrix} -2 & 2 & -2 \\ 10 & 6 & 4 \\ 6 & 14 & 6 \end{vmatrix} = -1 \left\{ -2 \begin{vmatrix} 6 & 4 \\ 14 & 6 \end{vmatrix} - 10 \begin{vmatrix} 2 & -2 \\ 14 & 6 \end{vmatrix} - 6 \begin{vmatrix} 2 & -2 \\ 6 & 4 \end{vmatrix} \right\} = -1(40 - 400 + 120) = +240$$

$$= (-1)^{4+2} \begin{vmatrix} 18 & 2 & -2 \\ -2 & 6 & 4 \\ 2 & 14 & 6 \end{vmatrix} = 1 \left\{ 18 \begin{vmatrix} 6 & 4 \\ 14 & 6 \end{vmatrix} + 2 \begin{vmatrix} 2 & -2 \\ 14 & 6 \end{vmatrix} + 2 \begin{vmatrix} 2 & -2 \\ 6 & 4 \end{vmatrix} \right\} = 1(-360 + 80 + 40) = -240$$

$$= (-1)^{4+3} \begin{vmatrix} 18 & -2 & -2 \\ -2 & 10 & 4 \\ 2 & 6 & 6 \end{vmatrix} = -1 \left\{ 18 \begin{vmatrix} 10 & 4 \\ 6 & 6 \end{vmatrix} + 2 \begin{vmatrix} -2 & -2 \\ 6 & 6 \end{vmatrix} + 2 \begin{vmatrix} -2 & -2 \\ 10 & 4 \end{vmatrix} \right\} = -1(648 + 0 + 24) = -672$$

$$= (-1)^{4+4} \begin{vmatrix} 18 & -2 & 2 \\ -2 & 10 & 6 \\ 2 & 6 & 14 \end{vmatrix} = 1 \left\{ 18 \begin{vmatrix} 10 & 6 \\ 6 & 14 \end{vmatrix} + 2 \begin{vmatrix} -2 & 2 \\ 6 & 14 \end{vmatrix} + 2 \begin{vmatrix} -2 & 2 \\ 10 & 6 \end{vmatrix} \right\} = 1(1872 - 80 - 64) = 1728$$

Ordenando los valores calculados de A_{ij} tendremos

$$A_{ij}=A_{ji} = \begin{bmatrix} \underline{A_{11}} & A_{12} & A_{13} & A_{14} \\ A_{21} & \underline{A_{22}} & A_{23} & A_{24} \\ A_{31} & A_{32} & \underline{A_{33}} & A_{34} \\ A_{41} & A_{42} & A_{43} & \underline{A_{44}} \end{bmatrix} = \begin{bmatrix} \underline{1,576} & 560 & -568 & 240 \\ 560 & \underline{3,712} & -1,568 & -240 \\ -568 & 1,568 & \underline{2,872} & -672 \\ 240 & -240 & -672 & \underline{1,728} \end{bmatrix}$$

5. Cálculo de la inversa de la matriz para la matriz de coeficientes reducido.

Los elementos a_{ij} de la matriz inversa A^{-1} de la matriz A se obtienen por división del cofactor A_{ji} por el determinante $|A|$ de la matriz.

Ejemplo: $a_{ij} = \frac{A_{ji}}{|A|}$

	\hat{u}	\hat{s}_1	\hat{s}_2	\hat{r}_1
$u:$	$\frac{1,576}{25,632}$	$\frac{560}{25,632}$	$\frac{-568}{25,632}$	$\frac{240}{25,632}$
$s_1:$	$\frac{560}{25,632}$	$\frac{3,712}{25,632}$	$\frac{-1,568}{25,632}$	$\frac{-240}{25,632}$
$s_2:$	$\frac{-568}{25,632}$	$\frac{-1,568}{25,632}$	$\frac{2,872}{25,632}$	$\frac{-672}{25,632}$
$r_1:$	$\frac{240}{25,632}$	$\frac{-240}{25,632}$	$\frac{-672}{25,632}$	$\frac{1,728}{25,632}$

$$A^{-1} = \begin{bmatrix} \underline{0.061486} & 0.021848 & -0.022160 & 0.009363 \\ 0.021848 & \underline{0.144819} & -0.061174 & -0.009363 \\ -0.022160 & -0.061174 & \underline{0.112047} & -0.026217 \\ 0.009363 & -0.009363 & -0.026217 & \underline{0.067416} \end{bmatrix}$$

6. Cálculos de las constantes.

Las constantes ($\hat{\mu}$, \hat{s}_1 , \hat{s}_2 , \hat{s}_3 , \hat{r}_1 , \hat{r}_2) se pueden estimar directamente por multiplicación de la matriz inversa por los RHM's de la matriz reducida.

$$\hat{\mu} = 0.061486 \times 94 + 0.021848 \times (-14) - 0.022160 \times 18 + 0.009363 \times (-20) = \underline{4.8876}$$

$$\hat{s}_1 = 0.021848 \times 94 + 0.144819 \times (-14) - 0.061174 \times 18 - 0.009363 \times (-20) = -\underline{0.8876}$$

$$\hat{s}_2 = -0.022160 \times 94 - 0.061174 \times (-14) + 0.112047 \times 18 - 0.026217 \times (-20) = \underline{1.3146}$$

$$\hat{s}_3 = ?$$

$$\hat{r}_1 = 0.09363 \times 94 - 0.009353 \times (-14) - 0.026217 \times 18 + 0.067416 \times (-20) = -\underline{0.8090}$$

$$\hat{r}_2 = ?$$

El valor de \hat{s}_3 y \hat{r}_2 podemos calcularlo fácilmente, recordando las restricciones impuestas mas arriba $\sum_k \hat{s}_i = 0 = \sum_j \hat{r}_j$ de donde:

$$\hat{s}_3 = -(\hat{s}_1 + \hat{s}_2) = -(-0.8876 + 1.3146) = -\underline{0.4270}$$

$$\hat{r}_2 = -\hat{r}_1 = -(-0.8090) = \underline{0.8090}$$

7. Cálculo de las sumas de cuadrados

a) SC de reducción $R(\mu, s_i, r_j)$

$$R(\mu, s_i, r_j) = (4.8876)(94) + (-0.8876)(-14) + (1.3146)(18) + (-0.8090)(-20) = 511.7036.$$

b) SC error = $\sum_l \sum_j \sum_k^2 Y_{ijk} - R(\mu, s_i, r_j)$

donde $\sum_l \sum_j \sum_k Y_{ijk} = \text{SC total sin corregir.}$

$$\text{SC error} = 568 - 511.7036 = \underline{56.2964} \text{ (sin interacción)}$$

c) SC entre padres (S) = $R(\mu, s_i, r_j) - \sum_j \frac{Y_{.j}^2}{n_{.j}}$

$$= 511.7036 - 496.0250 = \underline{15.6786}$$

$$d) \text{ SC raciones (R)} = R(\mu, s_i, r_j) - \sum_k \frac{Y_k^2}{n_k}$$

$$= 511.7036 - 502.000 = \underline{9.7036}$$

$$e) \text{ SC interacciones (S x R)} = \sum_i \sum_j \frac{Y_{ij}^2}{n_{ij}} - R(\mu, s_i, r_i)$$

$$541.9333 - 511.7036 = \underline{30.2297}$$

$$f) \text{ SC dentro de sub-clases S R} = 568 - 541.9333 = \underline{26.0667}$$

8. Análisis de variancia.

FV	GL	SC ¹	CM	F
Entre Padres (S)	2	15.6786	7.8393	
Raciones (R)	1	9.7036	9.7036	
Interacción S x R	2	30.2297	15.1148	6.96**
Dentro de sub-clases	12	26.0667	2.1722	

^{1/} Los SC para Padres y raciones no están ajustadas por la interacción.

A pesar de que la interacción está presente, vamos a considerarla como si no existiera, a fin de completar los cálculos, y la próxima sección se dedicará al análisis completo cuando la interacción está presente, entonces nuestro análisis quedaría así, si ignoramos la existencia de la interacción.

FV	GL	SC	CM	F
Entre Padres (S)	2	15.6786	7.8393	1.95 N S
Raciones	1	9.7036	9.7036	2.41 N S
Error	14	56.2964	4.0212	

9. Errores estandares de los promedios de mínimos cuadrados.

$$S_{\hat{\mu}} = \sqrt{(0.061486) 4.02112} = 0.50$$

$$S_{\hat{\mu} + \hat{s}_1} = \sqrt{[0.061486 + 0.144819 + 2(0.021848)] (4.0212)} = 1.00$$

$$S_{\hat{\mu} + \hat{s}_2} = 0.72$$

$$S_{\hat{\mu} + \hat{s}_3} = 0.89$$

$$S_{\hat{\mu} + \hat{r}_1} = 0.77$$

$$S_{\hat{\mu} + \hat{r}_2} = 0.67$$

10. Comparaciones individuales entre los \hat{s}_i

a) Prueba de rango múltiple de Duncan (0.05)

Comparaciones	$\bar{Y}_i - \bar{Y}_j$	$\sqrt{\frac{2}{c^{ii} + c^{jj} - 2c^{ij}}}$	Diferencias producto	$\hat{\sigma}_e z_p, n_2$
s_2 vs s_3	1.7416	2.396	4.17	6.08
s_2 vs s_1	2.2292	2.297	5.12	6.38
s_3 vs s_1	0.4606	2.116	0.97	6.08

b) Comparaciones ortogonales entre las \hat{r}_i .

Esto se puede llevar a cabo en la misma forma que en el capítulo precedente, por eso omitimos los cálculos.

11. Estimación de los componentes de variancia.

A pesar de que en nuestro ejemplo los GL para padres y raciones son muy pequeños pero como ilustración nos sirven muy bien; los efectos de padres y raciones serán consideradas como efectos al azar para estimar σ_s^2 y σ_r^2

El CM estimado para raciones es $\hat{\sigma}_e^2 + k_1 \hat{\sigma}_r^2$

El CM estimado para padres es $\hat{\sigma}_e^2 + k_2 \hat{\sigma}_s^2$ esto es cuando la interacción no existe.

K_1 y K_2 se pueden estimar por el método directo o indirecto.

a) Método directo de cálculo de K's

La matriz inversa de los segmentos Z de la matriz inversa completa es igual a Z_S^{-1} y Z_R^{-1}

$$Z_S^{-1} = \begin{bmatrix} 0.144819 & -0.061174 \\ -0.061174 & 0.112047 \end{bmatrix}^{-1} = \begin{bmatrix} 8.975050 & 4.900084 \\ 4.900084 & 11.600112 \end{bmatrix}$$

$$Z_R^{-1} = \begin{bmatrix} 0.067416 \end{bmatrix}^{-1} = 14.833274$$

Los coeficientes K_1 y K_2 se calculan de la siguiente manera:

$$K_1 = \frac{1}{2} (14.833274) = \underline{7.417}$$

$$K_2 = \frac{1}{3} \left[8.875050 + 11.600112 - \frac{1}{2} (4.900084 + 4.900084) \right]$$

$$= \frac{1}{3} (15.675078) = \underline{5.225}$$

Las sumas de cuadrados para padres y raciones se pueden verificar por medio del siguiente cálculo $B'Z^{-1}B$.

$$SC \text{ Padres} = \begin{bmatrix} -0.8876 & 1.3146 \end{bmatrix} Z_S^{-1}$$

$$= \begin{bmatrix} -1.5246 & 10.9002 \end{bmatrix} B = 15.683$$

$$SC \text{ Raciones} = \begin{bmatrix} -0.8090 \end{bmatrix} Z_R^{-1} B$$

$$\begin{bmatrix} -112.0001 \end{bmatrix} B = 9.708$$

b) Método indirecto para calcular K's

$$i) K_1 = \frac{1}{2-1} \left(18 - \frac{2^2 + 2^2}{4} - \frac{5^2 + 3^2}{8} - \frac{1^2 + 5^2}{6} \right) = \underline{7.417}$$

$$K_2 = \frac{1}{3-1} \left(18 - \frac{2^2 + 5^2 + 1^2}{8} - \frac{2^2 + 3^2 + 5^2}{10} \right) = \underline{5.225}$$

ii) El otro método indirecto o de Henderson para calcular K's es el que sigue.

Paso 1. Cálculo de las matrices inversas para las matrices reducidas.

Matriz inversa de coeficientes para $\hat{\mu} + \hat{s}_i$

$$\begin{array}{l} \mu + s_1 \\ \mu + s_2 \\ \mu + s_3 \end{array} \begin{bmatrix} \hat{\mu} + \hat{s}_1 & \hat{\mu} + \hat{s}_2 & \hat{\mu} + s_3 \\ \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{8} & 0 \\ 0 & 0 & \frac{1}{6} \end{bmatrix} = R^{ij}$$

Matriz inversa del coeficiente para $\hat{\mu} + \hat{r}_i$

$$\begin{array}{l} \mu + r_1 \\ \mu + r_2 \end{array} \begin{bmatrix} \hat{\mu} + \hat{r}_1 & \hat{\mu} + \hat{r}_2 \\ \frac{1}{8} & 0 \\ 0 & \frac{1}{10} \end{bmatrix} = R^{ij}$$

Paso 2. Cálculo de las sumas asociadas, matrices, NN'

$$N_{ij} = NN' = \begin{bmatrix} 2 & 2 \\ 5 & 3 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} 2 & 5 & 1 \\ 2 & 3 & 5 \end{bmatrix} = \begin{bmatrix} \underline{8} & 16 & 12 \\ 16 & \underline{34} & 20 \\ 12 & 20 & \underline{26} \end{bmatrix}$$

$$N_{ij} = NN' = \begin{bmatrix} 2 & 5 & 1 \\ 2 & 3 & 5 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 5 & 3 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} \underline{30} & 24 \\ 24 & \underline{38} \end{bmatrix}$$

Paso 3. Cálculo de la suma de productos de R^{ij} y N_{ij}

$$\text{Padres: } \sum_i \sum_j R^{ij} N^{ij} = 10.583$$

$$\text{Raciones: } \sum_i \sum_j R^{ij} N^{ij} = 7.550$$

Paso 4. Cálculo de los valores de K

$$K_1 = \frac{18 - 10.583}{1} = 7.417$$

$$K_2 = \frac{18 - 7.550}{2} = 5.225$$

Cálculo de los componentes de variancia estimados

FV	GL	SC	CM	Componentes estimados E(MS)
Padres (S)	2	15.6787	7.8393	$\sigma_e^2 + 5.225\sigma_s^2$
Raciones (R)	1	9.7036	9.7036	$\sigma_e^2 + 7.417\sigma_r^2$
Error	14	56.2964	4.0212	σ_e^2

$$\sigma_r^2 = \frac{9.7036 - 4.0212}{7.417} = 0.766$$

$$\sigma_s^2 = \frac{7.8393 - 4.0212}{5.225} = 0.731$$

12. Absorción de las ecuaciones $\mu + ai$

a) Cálculo de los nuevos coeficientes ajustados por \hat{r}_{ij}

Los cálculos envueltos en la absorción de las ecuaciones $\mu + ai$ se pueden obtener de la siguiente tabla.

	n_i	r_1	r_2	RHM
$\mu + s_1$	4	2(0.50000)	2(0.50000)	16(4.00000)
$\mu + s_2$	8	5(0.62500)	3(0.37500)	48(6.00000)
$\mu + s_3$	6	1(0.16667)	5(0.83333)	30(5.00000)

El cálculo de ajuste que se debe hacer en los coeficientes por \hat{r}_j y en los RHM's por las ecuaciones r_j que se completan en la siguiente forma:

$$C(r_1 r_1) = [8 - (2)(0.50000) - (5)(0.62500) - (1)(0.16667)] = \underline{3.70833}$$

$$C(r_2 r_1) = C(r_1 r_2) = - (2)(0.50000) - (5)(0.37500) - (1)(0.83333) \\ = - (0.50000)(2) - (3)(0.62500) - (5)(0.16667) = - \underline{3.70833}$$

$$C(r_2 r_2) = [10 - (2)(0.50000) - (3)(0.37500) - (5)(0.83333)] = \underline{3.70835}$$

$$s(r_1) = [37 - (2)(4.00000) - (5)(6.00000) - (1)(5.00000)] \\ = [37 - (0.50000)16 - (0.62500)(48) - (0.16667)(30)] = 6.00000$$

$$s(r_2) = 57 - (2)(4.00000) - (3)(6.00000) - (5)$$

Las ecuaciones reducidas después de la absorción de las ecuaciones $\mu + s_i$ son las que siguen:

	\hat{r}_1	\hat{r}_2	RHM
r_1 :	<u>3.70833</u>	-3.70833	-6.00000
r_2 :	-3.70833	<u>3.70835</u>	6.00000

- b) Imponiendo la restricción que $\sum_j r_j = 0$ se calcula la \hat{r}_j y la inversa de la matriz reducida. La ecuación resultante después de la sustracción por hileras y columnas es:

$$14.83334 \hat{r}_1 = -12.00000$$

$$r_1 = -\frac{12.00000}{14.83334} = -0.8090$$

$$r_2 = -r_1 = 0.8090$$

Estos valores coinciden con los obtenidos por el método directo. La inversa de la matriz reducida, de solo un orden en este caso es:

$$\frac{1}{14.83334} = 0.067416$$

- c) Solución para los $\hat{\mu} + \hat{s}_i$ constantes.

$$\begin{pmatrix} \hat{\mu} + \hat{s}_1 \\ \hat{\mu} + \hat{s}_2 \\ \hat{\mu} + \hat{s}_3 \end{pmatrix} = \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{8} & 0 \\ 0 & 0 & \frac{1}{6} \end{bmatrix} \left[\begin{pmatrix} 16 \\ 48 \\ 30 \end{pmatrix} - \begin{pmatrix} 2 & 2 \\ 5 & 3 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} -0.8090 \\ 0.8090 \end{pmatrix} \right]$$

$$= \begin{vmatrix} \frac{1}{4} & 0 & 0 & 16.0000 & 4.0000 \\ 0 & \frac{1}{8} & 0 & 49.6180 & 6.2022 \\ 0 & 0 & \frac{1}{6} & 26.7640 & 4.4607 \end{vmatrix}$$

- d) Cálculo de la matriz inversa variancia-covariancia y la máxima probabilidad estimadas para los promedios de padres.

$$G = \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{8} & 0 \\ 0 & 0 & \frac{1}{6} \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ -4 \end{bmatrix} \left[0.067416 \right] = \begin{bmatrix} 0 \\ \frac{2}{8} \\ \frac{4}{6} \end{bmatrix} \left[0.067416 \right]$$

$$G = \begin{bmatrix} 0 \\ \frac{-0.134832}{8} \\ \frac{0.269664}{6} \end{bmatrix} = \begin{bmatrix} 0 \\ -0.016854 \\ 0.044944 \end{bmatrix}$$

$$A = \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{8} & 0 \\ 0 & 0 & \frac{1}{6} \end{bmatrix} \left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 2 \\ -4 \end{pmatrix} (0.016854 \ 0.044944) \right]$$

$$= \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{8} & 0 \\ 0 & 0 & \frac{1}{6} \end{bmatrix} \begin{bmatrix} 1.000000 & 0.000000 & 0.000000 \\ 0.000000 & 1.033708 & -0.089888 \\ 0.000000 & -0.067416 & 1.179776 \end{bmatrix}$$

$$= \begin{bmatrix} 0.250000 & 0.000000 & 0.000000 \\ 0.000000 & 0.129214 & -0.011236 \\ 0.000000 & -0.011236 & 0.196629 \end{bmatrix}$$

La matriz inversa A se aplica directamente a los $\hat{\mu} + \hat{s}_1$. Esta matriz se puede transformar como se demuestra abajo para dar los elementos inversos que se aplican separadamente a $\hat{\mu}$ y los s_i .

$$\begin{aligned}
 KAK' &= \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ -1 & 2 & -1 \end{bmatrix} \begin{bmatrix} 0.250000 & 0.000000 & 0.000000 \\ 0.000000 & 0.129214 & -0.011236 \\ 0.000000 & -0.011236 & 0.196629 \end{bmatrix} K' \\
 &= \frac{1}{3} \begin{bmatrix} 0.250000 & 0.117978 & 0.185393 \\ 0.500000 & -0.117978 & -0.185393 \\ -0.250000 & 0.269664 & -0.219104 \end{bmatrix} \frac{1}{3} \begin{bmatrix} 1 & 2 & -1 \\ 1 & -1 & 2 \\ 1 & -1 & -1 \end{bmatrix} \\
 &= \begin{bmatrix} 0.061486 & 0.021848 & -0.022160 \\ 0.021848 & 0.144819 & -0.061174 \\ -0.022160 & -0.061174 & 0.112048 \end{bmatrix}
 \end{aligned}$$

Los elementos inversos transformados para la r_1 columna e hilera se obtienen de:

$$\begin{aligned}
 KG &= \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ -1 & 2 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ -0.016854 \\ -0.044944 \end{bmatrix} \\
 &= \begin{bmatrix} 0.009363 \\ -0.009363 \\ -0.026217 \end{bmatrix}
 \end{aligned}$$

La matriz inversa variancia-covariancia obtenida por esta forma indirecta está dada en el siguiente cuadro.

	$\hat{\mu}$	\hat{s}_1	\hat{s}_2	\hat{r}_1
μ :	<u>0.061486</u>	0.021848	-0.022160	0.009363
s_1 :	0.021848	<u>0.144819</u>	-0.061174	-0.009363
s_2 :	-0.022160	-0.061174	<u>0.112048</u>	-0.026217
r_1 :	0.009363	-0.009363	-0.026217	<u>0.067416</u>

Para calcular la máxima probabilidad estimada de los efectos de s_i , en nuestro ejemplo para $\mu + s_i$ es el siguiente:

$$\begin{aligned} \hat{\mu} + \hat{s}_1 &\doteq 4.8876 + \frac{0.731}{0.731 + (0.250000)(4.0212)} (-0.8876) \\ &\doteq 4.8876 + (0.4210)(-0.8876) \doteq 4.51 \end{aligned}$$

$$\begin{aligned} \hat{\mu} + \hat{s}_2 &\doteq 4.8876 + \frac{0.731}{0.731 + (0.129214)(4.0212)} \\ &\doteq 4.8876 + (0.5845)(1.3146) \doteq 5.65 \end{aligned}$$

$$\begin{aligned} \hat{\mu} + \hat{s}_3 &\doteq 4.8876 + \frac{0.731}{0.731 + (0.196629)(4.0212)} = (-0.4270) \\ &\doteq 4.8876 + (0.4802)(-0.4270) \doteq 4.68 \end{aligned}$$

DATOS CON DOS CRITERIOS DE CLASIFICACION CON INTERACCION

Ya hemos hablado del método de los cuadrados ponderados de los promedios como el indicado para analizar los datos con dos criterios de clasificación cuando la interacción está presente, esto da una estimación correcta de los efectos principales, con la condición de que las sub-classes tengan por lo menos una observación. A

continuación vamos a presentar el análisis de mínimos cuadrados para datos con dos criterios de clasificación cuando las constantes para los efectos de la interacción deben ser ajustados.

Modelo matemático

$$Y_{ijk} = \mu + a_i + b_j + (ab)_{ij} + e_{ijk}$$

donde: $(ab)_{ij}$ = efecto de la ij^{th} sub-clase, AB después de que los efectos de A y B hayan sido removidos.

Los demás términos de la ecuación tiene significado similar al anterior modelo.

El modelo matemático es el mismo si a_i y b_j son fijos o al azar. Si ambos son fijos el efecto de la interacción también está fijo. Cuando todos los efectos están fijos excepto el error el modelo se refiere al Modelo I de Eisenhart. Si todos los efectos están al azar excepto μ se refiere al Modelo II de Eisenhart. El modelo es mixto cuando uno de los factores está fijo y el otro al azar.

Ecuaciones de mínimos cuadrados

	$\hat{\mu}$	\hat{a}_i	\hat{b}_j	$(\hat{ab})_{ij}$	RHM
$\mu:$	$n..$	$ni.$	$n.j$	nij	$Y..$
$a_i:$	ni	$0ni^0$	nij	nij	$Y_i.$
$b_j:$	$n.j$	nij	$0n.j^0$	nij	$Y.j$
$(ab)_{ij}:$	nij	nij	nij	$0nij^0$	Yij

Como se puede ver la serie completa de ecuaciones son:

- una ecuación de μ
- una ecuación para cada una de las p clase de A

- una ecuación para cada una de las q clase de B
- una ecuación para cada una de las sub-classes A y B que tiene una o más observaciones.

Restricciones impuestas

Una solución a las ecuaciones de mínimos cuadrados no se puede obtener hasta que ellas están reducidas en número, a los números de grados de libertad. Las series de restricciones impuestas son, que la suma de las constantes para efectos principales sean igual a cero (0) dentro de una serie y que la suma de las constantes $(ab)_{ij}$ sea igual a cero en cada columna y cada hilera.

Los estimadores para las constantes son:

$$\hat{\mu}' = \hat{\mu} + \hat{a}_p + \hat{b}_q + \hat{a}b_{pq}$$

$$\hat{a}_i' = \hat{a}_i - \hat{a}_p + \hat{a}b_{iq} - \hat{a}b_{pq}$$

$$\hat{b}_j' = \hat{b}_j - \hat{b}_q + \hat{a}b_{pj} - \hat{a}b_{pq}$$

$$(\hat{a}b)'_{ij} = (\hat{a}b)_{ij} - (\hat{a}b)_{iq} - (\hat{a}b)_{pj} + (\hat{a}b)_{pq}$$

Además para la confusión de las constantes estimadas, cuando se usan posteriores restricciones la reducción total en suma de cuadrados se obtiene de:

$$\sum_{i=1}^{p-1} \hat{a}_i' Y_i + \sum_{j=1}^{q-1} \hat{b}_j' Y_{.j} + \sum_{i=1}^{p-1} \sum_{j=1}^{q-1} (\hat{a}b)'_{ij} Y_{ij}$$

Esto es incorrecto, siendo ya viciada más arriba, lo mismo ocurre con la SC para los efectos obtenidos de las matrices inversas.

Imponiendo las restricciones que $\sum_i \hat{a}_i = \sum_j \hat{b}_j = \sum_i (\hat{a}b)_{ij} =$

$(ab)_{ij} = 0$ sobre las primas (') estimadas de las constantes es posible calcular las \hat{a}_i , \hat{b}_j y $(\hat{a}b)_{ij}$. Además, la matriz inversa se

puede transformar para dar la inversa que se hubiera obtenido si hubieran sido impuestas las restricciones que $\sum_i \hat{a}_i = \sum_j \hat{b}_j = \sum_i (\hat{a}b)_{ij} = 0$ en las ecuaciones de mínimos cuadrados originales.

Cuando se imponen las restricciones $\sum_i \hat{a}_i = \sum_j \hat{b}_j = \sum_i (\hat{a}b)_{ij} = 0$ sobre las ecuaciones de mínimos cuadrados, es necesario hacer unas sustracciones y adiciones dentro de la matriz de coeficientes y los RHM's. Las sustracciones requeridas dentro de a_i y b_j , son las mismas para hileras y columnas como se explicó anteriormente. Dentro de la serie de coeficiente para las $(\hat{a}b)_{ij}$, las sustracciones y adiciones que pueden ser convenientemente escogidas para cada hilera son como sigue:

$$n_{ij} - n_{iq} - n_{pj} + n_{pq}$$

Después de que estos cambios hayan sido hechos por columnas el mismo procedimiento se sigue por hilera con los coeficientes modificados para los $(ab)_{ij}$ ecuaciones y para los RHM's.

Puesto que los RHM's reducidos para los $(ab)_{ij}$ que queda son

$$Y_{ij} - Y_{iq} - Y_{pj} + Y_{pq}$$

Análisis completo de mínimos cuadrados

Los elementos inversos para columna e hilera de a_p o b_q se pueden obtener sumando los elementos inversos para las a_i columnas o hileras y cambiar el signo de la suma, como se explicó en la primera sección. El mismo procedimiento se usa para obtener los elementos inversos de las columnas e hileras de la interacción que fueron eliminadas.

$$\begin{aligned}
 C^{\mu abiq} &= C^{abiq\mu} = - \sum_{j=1}^{q-1} C^{\mu abij} \\
 C^{aiabiq} &= C^{abiq ai} = - \sum_{j=1}^{q-1} C^{aiabij} \\
 C^{bjabiq} &= C^{abiqbj} = - \sum_{j=1}^{q-1} C^{bjabij} \\
 C^{\mu abpj} &= C^{abpj\mu} = - \sum_{i=1}^{p-1} C^{\mu abij} \\
 C^{\mu abpq} &= C^{abpq\mu} = - \sum_{i=1}^{p-1} C^{\mu abiq} = \sum_{j=1}^{q-1} C^{\mu abpj}
 \end{aligned}$$

La suma de cuadrados para error en el análisis de variancia se calcula de:

$$\sum_i \sum_j \sum_k Y_{ijk}^2 - R[\mu, ai, bj, (ab)_{ij}]$$

$$\text{donde: } R[\mu, ai, bj, (ab)_{ij}] = \hat{\mu} Y. + \sum_{i=1}^{p-1} \hat{a}_i (Y_{i1} - Y_{ip}) + \sum_{j=1}^{q-1} \hat{b}_j (Y_{j1} - Y_{jq}) + \sum_{i=1}^{p-1} \sum_{j=1}^{q-1} (\hat{ab})_{ij} (Y_{ij} - Y_{iq} - Y_{pj} + Y_{pq})$$

La reducción total en S de C, $R[\mu, ai, bj, (ab)_{ij}]$ se puede calcular también de $\sum_i \sum_j \frac{Y_{ij}^2}{n_{ij}}$ en este caso, puesto que toda la variabilidad entre las AB sub-clases se dan por las constantes que están siendo ajustadas.

La SC para A, B y AB se puede calcular por el procedimiento directo $B'Z^{-1}B$ que envuelve los segmentos de la matriz inversa de la matriz variancia-covariancia y las constantes estimadas o por el procedimiento indirecto que envuelve las diferencias en varias reducciones en la SC. Por ejemplo la SC para A usando el método indirecto es $R[\mu, ai, bj, (ab)_{ij}] - R[\mu, bj, (ab)_{ij}]$

Cuando los GL de la interacción es grande se puede obtener más

fácilmente la interacción por la fórmula $R[\mu, a_i, b_j, (ab)_{ij}]$ - $R(\mu, a_i, b_j)$ que de $B'Z^{-1}B$. Cuando las constantes para interacción deben ser ajustadas solo con los efectos principales; las sumas de cuadrados para los efectos principales pueden ser calculados directamente de $B'Z^{-1}B$.

Los promedios de mínimos cuadrados para las clases de A y de B son $\hat{\mu} + \hat{a}_i$ y $\hat{\mu} + \hat{b}_j$ respectivamente. Los errores estandares para $\hat{\mu}$, $\hat{\mu} + \hat{a}_i$ y $\hat{\mu} + \hat{b}_j$ se calculan en la misma forma ya vista anteriormente. Sin embargo cuando los efectos de la interacción es significativa el investigador estaría más interesado en los promedios de las sub-clases AB que los promedios de clases, ya que el promedio de la sub-clase de mínimos cuadrados es:

$$\hat{s}_{ij} = \hat{\mu} + \hat{a}_i + \hat{b}_j + (\hat{ab})_{ij}$$

El error estandar puede ser calculado de la matriz inversa y en la siguiente forma:

$$S\hat{s}_{ij} = \sqrt{\frac{(C^{\mu\mu} + C^{a_ia_i} + C^{b_jb_j} + C^{ab_{ij}} + 2C^{\mu a_i} + 2C^{\mu b_j} + 2C^{\mu ab_{ij}} + 2C^{a_ib_j} + 2C^{a_iab_{ij}} + 2C^{b_iab_{ij}}) \hat{\sigma}_e^2}{}}$$

Con datos de dos criterios de clasificación cuando todas las sub-clases tienen observaciones y las constantes de interacción se ajustan, el error estandar del promedio de sub-clase, \hat{s}_{ij} se reduce a:

$$\sqrt{\frac{1}{n_{ij}} \hat{\sigma}_e^2}$$

El procedimiento para la comparación de promedios para \hat{a}_i y \hat{b}_j es igual a lo visto en las secciones previas, lo mismo para las comparaciones ortogonales entre \hat{a}_i o entre \hat{b}_j . El mismo procedimiento se puede seguir para $(\hat{a}\hat{b})_{ij}$ descomponiendo en grados de libertad simple para contrastes ortogonales, por ejemplo: suponiendo que hay dos clases de A y tres de B y los siguientes grados de libertad de los contrastes ortogonales se desean entre los promedios de sub-clases.

	s_{11}	s_{12}	s_{13}	s_{21}	s_{22}	s_{23}
μ	1	1	1	1	1	1
A	1	1	1	-1	-1	-1
B	2	-1	-1	2	-1	-1
	0	1	-1	0	1	-1
AB	2	-1	-1	-2	1	1
	0	1	-1	0	-1	1

La matriz de transformación K, es:

$$K = \frac{1}{24} \begin{bmatrix} 4 & 4 & 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & -4 & -4 & -4 \\ 6 & -3 & -3 & 6 & -3 & -3 \\ 0 & 6 & -6 & 0 & 6 & -6 \\ 6 & -3 & -3 & -6 & 3 & 3 \\ 0 & 6 & -6 & 0 & -6 & 6 \end{bmatrix}$$

La matriz transformada T se calcula de la siguiente manera:

$$T = K D^{-1} K'$$

y las constantes para los contrastes ortogonales, $\hat{\epsilon}_i$ de $KB = \hat{\epsilon}_i$ donde B es una columna vector de \hat{s}_{ij} .

El cuadrado medio dentro de sub-classes es el error apropiado para probar la significancia de la interacción si el modelo es fijo mixto o al azar, lo mismo para probar la significancia de los efectos principales, si el modelo es fijo.

La expectación de los cuadrados medios para el modelo al azar es el siguiente:

	E(MS)
A	$\sigma_e^2 + K_4 \sigma_{ab}^2 + K_5 \sigma_a^2$
B	$\sigma_e^2 + K_2 \sigma_{ab}^2 + K_3 \sigma_b^2$
AB	$\sigma_e^2 + K_1 \sigma_{ab}^2$
Error	σ_e^2

Los coeficientes K_1 , K_3 y K_5 se pueden calcular por el método directo y K_2 , K_4 se pueden calcular por el procedimiento indirecto.

La expectación de los cuadrados medios para el modelo mixto con los efectos de ai al azar y bj fijo es el siguiente.

	E(MS)
A	$\sigma_e^2 + K_5 \sigma_a^2$
B	$\sigma_e^2 + K_2 \sigma_{ab}^2 + K_3 \sigma_b^2$
AB	$\sigma_e^2 + K_1 \sigma_{ab}^2$
Error	σ_e^2

La variancia debida a a_i , σ_a^2 ahora se estima sobre una serie fija de los efectos de B.

Ejemplo numérico

El mismo ejemplo usado en las secciones previas. Vamos a considerar ahora bajo dos criterios de clasificación con interacción presente.

Modelo matemático

$$Y_{ijk} = \mu + s_i + r_j + (sr)_{ij} + e_{ijk}$$

El significado de cada término de la ecuación ya lo hemos dado en las secciones anteriores y es el mismo, excepto que en este modelo consideramos la interacción $(sr)_{ij}$.

Ecuaciones de mínimos cuadrados

	$\hat{\mu}$	\hat{s}_1	\hat{s}_2	\hat{s}_3	\hat{r}_1	\hat{r}_2	$(\hat{sr})_{11}$	$(\hat{sr})_{12}$	$(\hat{sr})_{21}$	$(\hat{sr})_{22}$	$(\hat{sr})_{31}$	$(\hat{sr})_{32}$	RHM
μ :	<u>18</u>	4	8	6	8	10	2	2	5	3	1	5	94
s_1 :	4	<u>4</u>	0	0	2	2	2	2	0	0	0	0	16
s_2 :	8	0	<u>8</u>	0	5	3	0	0	5	3	0	0	48
s_3 :	6	0	0	<u>6</u>	1	5	0	0	0	0	1	5	30
r_1 :	8	2	5	1	<u>8</u>	0	2	0	5	0	1	0	37
r_2 :	10	2	3	5	0	<u>10</u>	0	2	0	3	0	5	57
$(sr)_{11}$:	2	2	0	0	2	0	<u>2</u>	0	0	0	0	0	11
$(sr)_{12}$:	2	2	0	0	0	2	0	<u>2</u>	0	0	0	0	5
$(sr)_{21}$:	5	0	5	0	5	0	0	0	<u>5</u>	0	0	0	23
$(sr)_{22}$:	3	0	3	0	0	3	0	0	0	<u>3</u>	0	0	25
$(sr)_{31}$:	1	0	0	1	1	0	0	0	0	0	<u>1</u>	0	3
$(sr)_{32}$:	5	0	0	5	0	5	0	0	0	0	0	<u>5</u>	27

Imponiendo las restricciones que $\sum_i \hat{s}_i = \sum_j \hat{f}_j = \sum_k (\hat{sr})_{kj} = \sum_j (sr)_{ij} = 0$ tenemos las ecuaciones reducidas de mínimos cuadrados.

	$\hat{\mu}$	\hat{s}_1	\hat{s}_2	\hat{r}_1	$(\hat{sr})_{11}$	$(\hat{sr})_{12}$	RHM
μ :	<u>18</u>	-2	2	-2	4	6	94
s_1 :	-2	<u>10</u>	6	4	-4	-4	-14
s_2 :	2	6	<u>14</u>	6	-4	-2	18
r_1 :	-2	4	6	<u>18</u>	-2	2	-20
$(sr)_{11}$:	4	-4	-4	-2	<u>10</u>	6	30
$(sr)_{12}$:	6	-4	-2	2	6	<u>14</u>	22

Matriz inversa de la matriz variancia-covariancia

	$\hat{\mu}$	\hat{s}_1	\hat{s}_2	\hat{r}_1	$(\hat{sr})_{11}$	$(\hat{sr})_{12}$
μ :	<u>0.075926</u>	0.007407	-0.031481	0.018519	-0.018519	-0.029630
s_1 :		<u>0.159259</u>	0.051852	-0.018519	0.018519	0.029630
s_2 :			<u>0.120370</u>	-0.029630	0.029630	0.007407
r_1 :				<u>0.075926</u>	0.007407	-0.031481
$(sr)_{11}$:					<u>0.159259</u>	-0.051852
$(sr)_{12}$:						<u>0.120370</u>

El cálculo de las constantes se calcula por el procedimiento usual.

$$\begin{aligned} \hat{\mu} &= \underline{4.8889} & \hat{r}_1 &= -\underline{0.5222} \\ \hat{s}_1 &= -\underline{0.8889} & (\hat{sr})_{11} &= \underline{2.0222} \\ \hat{s}_2 &= \underline{1.5778} & (\hat{sr})_{21} &= -\underline{1.3444} \end{aligned}$$

Las constantes que hemos reducido en las ecuaciones originales se calculan de la siguiente manera:

$$\hat{S}_3 = -(-0.8889 + 1.5778) = \underline{0.6889}$$

$$\hat{r}_1 = -(-0.5222) = \underline{0.5222}$$

$$(\hat{sr})_{12} = -(2.0222) = \underline{-2.0282}$$

$$(\hat{sr})_{22} = -(1.3444) = \underline{1.3444}$$

$$(\hat{sr})_{31} = -(2.0222 - 1.3444) = \underline{0.6778}$$

$$(\hat{sr})_{32} = -(-2.0222 + 1.3444) = \underline{0.67778}$$

Cálculos de las sumas de cuadrados.

$$\begin{aligned} \text{a) SC de error} &= \sum_i \sum_j \sum_k Y_{ijk}^2 - R[\mu, s_i, r_j, (sr)_{ij}] \\ &= 568 - (4.8889)(94) - (-0.8889)(-14) - (1.5778)(18) - (-0.5222) \\ &\quad (-20) - (2.0222)(30) - (-1.3444)(22) = 26.0652 \end{aligned}$$

También se puede calcular por la siguiente fórmula:

$$\sum_i \sum_j \sum_k Y_{ijk}^2 - \sum_{nij} \frac{Y_{ij}^2}{n_{ij}} = 568 - 541.9333 = \underline{26.0667}$$

$$\text{b) SC entre padres (s)} = B'Z^{-1}B.$$

$$\text{SC(s)} = \begin{bmatrix} -0.8889 & 1.5778 \end{bmatrix} \begin{bmatrix} 0.159259 & -0.051852 \\ -0.051852 & 0.120370 \end{bmatrix}^{-1} \begin{bmatrix} -0.8889 \\ 1.5778 \end{bmatrix} = \underline{21.0015}$$

$$\text{c) SC raciones (r)} = \frac{(-0.5222)^2}{0.075926} = \underline{3.5916}$$

$$\text{d) SC Interacción (S x R)} = B'Z^{-1}B.$$

$$\begin{aligned} \text{SC (SxR)} &= \begin{bmatrix} 2.0222 & -1.34444 \end{bmatrix} \begin{bmatrix} 0.159259 & -0.051852 \\ -0.051852 & 0.120370 \end{bmatrix}^{-1} \begin{bmatrix} 2.0222 \\ -1.3444 \end{bmatrix} \\ &= 30.2245 \end{aligned}$$

Análisis de variancia

FV	GL	SC	CM	F
S	2	21.0015	10.5008	4.83 *
R	1	3.5916	3.5916	1.65 N S
S x R	2	30.2245	15.1122	6.96 *
Error	12	26.0667	2.1721	

Promedios de mínimos cuadrados, errores estandares y comparaciones individuales.

a) Promedios: \circ

$$\hat{\mu} = 4.9$$

$$\hat{\mu} + \hat{s}_1 = 4.0$$

$$\hat{\mu} + \hat{s}_2 = 6.5$$

$$\hat{\mu} + \hat{s}_3 = 4.2$$

$$\hat{\mu} + \hat{r}_1 = 4.4$$

$$\hat{\mu} + \hat{r}_2 = 5.4$$

$$\hat{\mu} + \hat{s}_1 + \hat{r}_1 + (\hat{sr})_{11} = 5.5$$

$$\hat{\mu} + \hat{s}_1 + \hat{r}_2 + (\hat{sr})_{12} = 2.5$$

$$\hat{\mu} + \hat{s}_2 + \hat{r}_1 + (\hat{sr})_{21} = 4.6$$

$$\hat{\mu} + \hat{s}_2 + \hat{r}_2 + (\hat{sr})_{22} = 8.3$$

$$\hat{\mu} + \hat{s}_3 + \hat{r}_1 + (\hat{sr})_{31} = 3.0$$

$$\hat{\mu} + \hat{s}_3 + \hat{r}_2 + (\hat{sr})_{32} = 5.4$$

b) Los errores estandares de estos promedios se calculan en la forma usual, a partir de los elementos de la matriz inversa y el estimador de σ_e^2

- c) Las comparaciones individuales también se pueden hacer en la forma conocida.

Estimación de los componentes de variancia

Vamos a considerar los efectos de padres y raciones como si fueran ambos, variables al azar.

FV	GL	Componentes estimados	E(MS)
S	p-1	$\sigma_e^2 + K_4 \sigma_{sr}^2 + K_5 \sigma_s^2$	
R	q-1	$\sigma_e^2 + K_2 \sigma_{sr}^2 + K_3 \sigma_r^2$	
SR	r-p-q+1	$\sigma_e^2 + K_1 \sigma_{sr}^2$	
Error	n. - r	σ_e^2	

Los coeficientes K_1 , K_3 y K_5 se pueden calcular por el método directo de las matrices Z^{-1} que fueron usados para calcular las sumas de cuadrados.

$$K_1 = \frac{7.303395 + 9.662967 - \left(-\frac{2}{2}\right) (3.146096)}{6} = \frac{13.820266}{6} = 2.303$$

$$K_3 = \frac{1}{2} \frac{0.075926}{2} = \frac{13.17072}{2} = 6.585$$

$$K_5 = \frac{7.303395 + 0.662967 - \left(-\frac{2}{2}\right) (3.146096)}{3} = \frac{13.820266}{3} = 4.607$$

Los coeficientes de σ_{sr}^2 en los componentes estimados de R y S se calculan por el método indirecto.

$$K_2 = \frac{1}{1} (18 - \sum_l \sum_j R^{lj} N^{lj})$$

$$R = \begin{bmatrix} \underline{18} & -2 & 2 & 4 & 6 \\ -2 & \underline{10} & 6 & -4 & -4 \\ 2 & 6 & \underline{14} & -4 & -2 \\ 4 & -4 & -4 & \underline{10} & 6 \\ 6 & -4 & -2 & 6 & \underline{14} \end{bmatrix}$$

$$R^{-1} = \begin{bmatrix} \underline{0.071409} & 0.011924 & -0.024254 & -0.020326 & -0.021952 \\ & \underline{0.154742} & -0.059079 & 0.020326 & 0.021952 \\ & & \underline{0.108807} & 0.032521 & -0.004878 \\ & & & \underline{0.158536} & 0.048781 \\ & & & & \underline{0.107317} \end{bmatrix}$$

Los elementos que están a la izquierda de la diagonal se omitieron.

$$N^{ij} = NN' = \begin{bmatrix} 2 & 2 & 5 & 3 & 1 & 5 \\ 2 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix} N'$$

μ	s_1	s_2	s_3	$(sr)_{11}$	$(sr)_{12}$	$(sr)_{21}$	$(sr)_{22}$	$(sr)_{31}$	$(sr)_{32}$
<u>68</u>	8	34	26	4	4	25	9	1	25
	<u>8</u>	0	0	4	4	0	0	0	0
		<u>34</u>	0	0	0	25	9	0	0
			<u>26</u>	0	0	0	0	1	25
				<u>4</u>	0	0	0	0	0
					<u>4</u>	0	0	0	0
						<u>25</u>	0	0	0
							<u>9</u>	0	0
								<u>1</u>	0
									<u>25</u>

La matriz N^{ij} reducida es

<u>68</u>	-18	8	24	40
	<u>34</u>	26	-24	-24
		<u>60</u>	-24	-8
			<u>34</u>	26
				<u>60</u>

$$K_2 = 18 - (68)(0.071409) - (-18)(0.011924) - - - (60)(0.107317)$$

$$K_2 = 18 - 15.804544 = 2.195$$

$$K_4 = \frac{1}{2} (18 - \sum_i \sum_j R^{ij} N^{ij})$$

La matriz reducida N_{ij}

$$\begin{bmatrix} 68 & -8 & 24 & 40 \\ & \underline{68} & -18 & 8 \\ & & \underline{34} & 26 \\ & & & \underline{60} \end{bmatrix}$$

$$K_4 = \frac{1}{2} (18 - 13.393274) = \frac{1}{2} (4.606726) = 2.303$$

Cálculo de los componentes de variancia

FV	GL	CM	Componentes estimados E(MS)
S	2	10.5008	$\sigma_e^2 + 2.303\sigma_{sr}^2 + 4.607\sigma_s^2$
R	1	3.5916	$\sigma_e^2 + 2.195\sigma_{sr}^2 + 6.585\sigma_r^2$
SR	2	15.1122	$\sigma_e^2 + 2.303\sigma_{sr}^2$
Error	12	2.1721	σ_e^2

$$\hat{\sigma}_e^2 = 2.172$$

$$\hat{\sigma}_{sr}^2 = 5.618$$

$$\hat{\sigma}_r^2 = -1.657$$

$$\hat{\sigma}_s^2 = -1.001$$

Análisis por el método de los cuadrados ponderados de los promedios. Ya hemos descrito la técnica de este método en una sección previa. Los pasos a seguir son los siguientes:

- 1) Cálculo de las constantes estimadas.

$$\hat{\mu} + \hat{s}_i = \frac{\sum \hat{s}_{ij}}{q}$$

$$\hat{\mu} + \hat{s}_1 = \frac{5.5 + 2.5}{2} = 4.0000$$

$$\hat{\mu} + \hat{s}_2 = \frac{4.60000 + 8.3333}{2} = 6.4667$$

$$\hat{\mu} + \hat{s}_3 = \frac{3.0 + 5.4}{2} = 4.2000$$

$$\hat{\mu} + \hat{r}_j = \frac{\sum \hat{s}_{ij}}{p}$$

$$\hat{\mu} + \hat{r}_1 = \frac{5.5 + 4.6 + 3.0}{3} = 4.3667$$

$$\hat{\mu} + \hat{r}_2 = \frac{2.5000 + 8.3333 + 5.4000}{3} = 5.4111$$

$$\hat{\mu} = \frac{4.0000 + 6.46667 + 4.2000}{3} = \frac{4.3667 + 5.4111}{2} = 4.8889$$

$$\hat{s}_1 = 4.0000 - 4.8889 = -0.8889$$

$$\hat{s}_2 = 6.4667 - 4.8889 = 1.5778$$

$$\hat{s}_3 = 4.2000 - 4.8889 = -0.6889$$

$$\hat{r}_1 = 4.3667 - 4.8889 = -0.5222$$

$$\hat{r}_2 = 5.4111 - 4.8889 = 0.5222$$

2) Cálculo de las ponderaciones.

$$w_i = \frac{q^2}{\sum_j \frac{1}{n_{ij}}}$$

$$v_j = \frac{p^2}{\sum_i \frac{1}{n_{ij}}}$$

$$w_1 = \frac{4}{\frac{1}{2} + \frac{1}{2}} = 4.000.$$

$$v_1 = \frac{9}{\frac{1}{2} + \frac{1}{5} + 1} = 5.2941$$

$$w_2 = \frac{4}{\frac{1}{5} + \frac{1}{3}} = 7.5000$$

$$v_2 = \frac{9}{\frac{1}{2} + \frac{1}{3} + \frac{1}{5}} = 8.7097$$

$$w_3 = \frac{4}{1 + \frac{1}{5}} = 3.3333$$

3) Cálculo de las SC.

$$SC \text{ error} = SC \text{ dentro de sub-clases} = \underline{26.0067}$$

$$SC(S \times R) = \underline{30.2245}$$

$$SC(R) = (4.0000)(-0.8889)^2 + (7.5000)(1.5778)^2 + (3.3333)(-0.6889)^2 \\ - \left[\frac{(4.0000)(-0.8889) + (7.5000)(1.5778) + (3.3333)(-0.6889)}{4.0000 + 7.5000 + 3.3333} \right] \\ = \underline{21.0013}$$

$$SC(S) = (5.2941)(-0.5222)^2 + (8.7097)(0.5222)^2 \\ - \left[\frac{(5.2941)(-0.5222) + (8.7097)(0.5222)}{5.2941 + 8.7097} \right] = 3.5916$$

4) Cálculo de K's

$$K_2 = \frac{1}{(3)(1)} \left[(5.2941 + 8.7097) - \frac{(5.2941)^2 + (8.7097)^2}{5.2941 + 8.7097} \right] = 2.195$$

$$K_3 = \frac{1}{1} (6.5854) = 6.585$$

$$K_4 = \frac{1}{(2)(2)} \left[14.8333 - \frac{(4.0000)^2 + (7.5000)^2 + (3.3333)^2}{14.8333} \right] = 2.303$$

$$K_5 = \frac{1}{2} (9.2135) = 4.607$$

Estos resultados concuerdan con los obtenidos por el otro método.

Además hay otros métodos para el análisis de variancia, como el procedimiento corto para calcular la matriz inversa $C^{-1} = KD^{-1}K'$. El otro método es el I de Henderson.

MODELO DE ANALISIS USADO EN EXPERIMENTOS GENETICOS

Hay tres casos principales que deben tenerse en cuenta para el análisis de datos de cruzamientos.

1. Análisis de datos obtenidos de todos los cruzamientos posibles entre líneas, sin incluir las endocriadas (puros).
2. Análisis de todos los datos provenientes de cruza y endocrias (puros) producidos juntos, en donde los efectos de las líneas puras se estiman independientemente de la habilidad combinatoria general y los efectos de la habilidad combinatoria materna.
3. Análisis de todos los datos provenientes de endocrias (puros) y cruza producidas juntas, donde los efectos de los puros se estiman simultáneamente con los efectos de la habilidad combinatoria general y materna.

Análisis de datos provenientes de cruzamientos

Si no se consideran las líneas puras (endocriadas) el modelo sería el siguiente.

1) Modelo matemático (I)

$$Y_{ijk} = \mu + g_i + g_j + m_j + c_{ij} + r_{ij} + e_{ijk}$$

donde:

Y_{ijk} = observación K^{th} en la progenie proveniente de la línea padre i^{th} y madre j^{th} .

μ = promedio general

$g_i(g_j)$ = efecto de la habilidad combinatoria general de la línea $i^{th}(j^{th})$.

m_j = efecto materno para la línea j^{th} de madre.

c_{ij} = efecto de la habilidad combinatoria específica.

r_{ij} = efecto ligado al sexo o recíproco.

e_{ijk} = error al azar que se asume NID $(0, \sigma^2)$

El valor $g_i(g_j)$ se asume igual a $\frac{1}{2}$ del valor genético aditivo (valor del breeding) de la i^{th} (j^{th}) línea y el valor está expresado como desviación de μ . La m_j mide la habilidad materna pre y post-natal de una línea y cada una es más bien una función del genotipo de una línea que de los genes transmitidos a la progenie hembra de la línea.

Los c_{ij} son efectos, además de los efectos de g_i y m_j y se miden sobre la progenie de la línea padre i^{th} con la línea madre j^{th} y la progenie de la línea padre j^{th} con la línea madre i^{th} .

Los c_{ij} miden si cuantas veces es mejor o peor que el promedio, los de los apareamientos recíprocos, que se esperarían sobre la base de un conocimiento exacto de los valores de los efectos genéticos aditivos, y valores maternos de las líneas. Los r_{ij} son efectos que existen además de la genética aditiva, maternal y efectos específicos y se miden por diferencia entre cruzas recíprocas después de haber sido estimada la diferencia en habilidad combinatoria materna entre las líneas i^{th} y j^{th} .

2) Ecuaciones de mínimos cuadrados.

	$\hat{\mu}$	\hat{g}_i	\hat{m}_j	\hat{c}_{ij}	\hat{r}_{ij}	BHM
μ	$n_{..}$	$n_{i.} + n_{.i}$	$n_{.j}$	$n_{ij} + n_{ji}$	n_{ij}	$Y_{..}$
g_i	$n_{i.} + n_{.i}$	$n_{i.} + n_{.i}$ $n_{ij} + n_{ji}$	n_{ij}	$n_{ij} + n_{ji}$	n_{ij}	$Y_{i.} + Y_{.i}$
m_j	$n_{.j}$	n_{ij}	n_{ij}^0	n_{ij}	n_{ij}^0	$Y_{.j}$
c_{ij}	$n_{ij} + n_{ji}$	$n_{ij} + n_{ji}$	n_{ij}	$n_{ij} + n_{ji}^0$	n_{ij}	$Y_{ij} + Y_{ji}$
r_{ij}	n_{ij}	n_{ij}	n_{ij}	n_{ij}	n_{ij}^0	Y_{ij}

Estas ecuaciones difieren algo de las ecuaciones de mínimos cuadrados presentados anteriormente, en que los elementos fuera de la diagonal en el segmento de la matriz de coeficientes concerniente a los efectos de la habilidad combinatoria general no son cero. Esto se debe al hecho que cada uno del grupo de cruza se usa en la estimación del efecto de la habilidad combinatoria general de dos líneas. Por ejemplo, una cruce entre un toro de la línea 1 y una vaca de la línea 2 se usa para estimar la habilidad combinatoria general de ambas líneas 1 y 2. Así, la suma de los coeficientes para la \hat{g}_i en la ecuación μ es igual a dos veces el coeficiente para μ . Sin embargo, las sumas de los coeficientes para los efectos de \hat{m}_j , \hat{c}_{ij} y \hat{r}_{ij} en la ecuación μ son iguales para uno y otro y también igual al coeficiente para $\hat{\mu}$. Además el total de RHM's para todas las ecuaciones excepto para g_i igual al gran total Y.. El total de RHM's para las ecuaciones g_i es igual a 2 Y. Antes de que la matriz de coeficientes sea invertida para obtener los elementos inversos se debe imponer ciertas restricciones.

3) La posición de restricciones sobre las constantes a estimar.

$$\sum_i \hat{g}_i = \sum_j \hat{m}_j = \sum_k \hat{c}_{ij} = \sum_l \hat{c}_{ij} = \sum_{k=1}^{p-1} \sum_j \hat{c}_{ij} = \sum_k \hat{r}_{ij} = \sum_j \hat{r}_{ij} = \hat{r}_{ij} + \hat{r}_{ij} = 0$$

Las restricciones que $\sum_i \hat{g}_i = \sum_j \hat{m}_j = 0$ se puede imponer por sustracciones de hileras y columnas como se explicó anteriormente. Las sustracciones y adiciones requeridas para imponer las restricciones sobre las constantes a estimar para los efectos específicos, $\sum_k \hat{c}_{ij} = \sum_j \hat{c}_{ij} = \sum_{k=1}^{p-1} \sum_j \hat{c}_{ij} = 0$ será dado para el diseño específico de todas las cruza posibles entre 4 líneas.

Si el número en las sub-classes consideradas con la medida de los

efectos específicos son representados por:

$$x \quad n_{12} \quad n_{13} \quad n_{14}$$

$$n_{21} \quad x \quad n_{23} \quad n_{24}$$

$$n_{31} \quad n_{32} \quad x \quad n_{34}$$

$$n_{41} \quad n_{42} \quad n_{43} \quad x$$

entonces de las sustracciones de la última columna en cada hilera de todos los otros coeficientes resultó en:

$$x \quad n_{12}^{-n_{14}} \quad n_{13}^{-n_{14}} \quad n_{14}^{-n_{14}}$$

$$n_{21}^{-n_{24}} \quad x \quad n_{23}^{-n_{24}} \quad n_{24}^{-n_{24}}$$

$$n_{31}^{-n_{34}} \quad n_{32}^{-n_{34}} \quad x \quad n_{34}^{-n_{34}}$$

$$n_{41}^{-n_{43}} \quad n_{42}^{-n_{43}} \quad n_{43}^{-n_{44}} \quad x$$

Las sustracciones por hileras entonces producen:

$$x \quad n_{12}^{-n_{14}} - n_{42}^{+n_{43}} \quad n_{13}^{-n_{14}}$$

$$n_{21}^{-n_{24}} - n_{41}^{+n_{43}} \quad x \quad n_{23}^{-n_{24}}$$

$$n_{31}^{-n_{34}} - n_{41}^{+n_{43}} \quad n_{32}^{-n_{34}} - n_{42}^{+n_{43}} \quad x$$

Puesto que el efecto específico para una cruz particular se mide sobre las dos cruza recíprocas estos valores se pueden combinar en la siguiente forma:

$$n_{12}^{-n_{14}} - n_{41}^{-n_{42}} + n_{43}^{+n_{21}} - n_{24}^{-n_{41}} + n_{43}^{+n_{23}} \quad n_{13}^{-n_{14}} + n_{31}^{-n_{34}} - n_{41}^{-n_{43}} + n_{43}^{+n_{32}}$$

$n_{23}^{-n_{24}} + n_{32}^{-n_{34}} - n_{42}^{-n_{43}} + n_{43}^{+n_{31}}$ o lo mismo se puede expresar en la siguiente manera.

$$n'_{12} - n'_{14} - n'_{24} + 2n_{43}$$

$$n'_{13} - n'_{14} - n_{34} + n_{43}$$

$$n'_{23} - n'_{24} - n_{34} + n_{43}$$

donde

$$n'_{12} = n_{12} + n_{21}, n'_{14} = n_{14} + n_{41} \text{ etc.}$$

La tercera restricción es $\sum_i \sum_j \hat{c}_{ij} = 0$ resulta en las ecuaciones siguientes:

$$\begin{aligned} n'_{12} - n'_{14} - n'_{24} + 2 n_{43} - n'_{23} + n'_{24} + n_{34} - n_{43} &= n'_{12} - n'_{14} - n'_{23} + \\ n_{34} \text{ y } n'_{13} - n'_{14} - n_{34} + n_{43} - n'_{23} + n'_{24} + n_{34} - n_{43} &= n'_{13} - n'_{14} - \\ n'_{23} + n'_{24}. \end{aligned}$$

La restricción sobre las ecuaciones para las cruza recíprocas que $\sum_i \hat{r}_{ij} + \sum_j \hat{r}_{ji} = 0$ resulta en la resta por hilera y columna, siguiendo el mismo procedimiento para los efectos específicos.

La restricción que los efectos recíprocos para cada cruza debe sumar a cero ($\hat{r}_{ij} + \hat{r}_{ji} = 0$) produce las siguientes ecuaciones $n'_{12} - n'_{14} - n_{42} + n_{43} - n_{21} + n_{24} + n_{41} - n_{43} = n'_{12} - n'_{14} - n_{42} - n_{21} + n_{24} + n_{41} \dots n'_{13} - n'_{14} - n_{31} + n_{34} + n_{41} - n_{43} = n'_{13} - n'_{14} - n_{31} + n_{34} + n_{41} - n_{43} \dots n'_{23} - n'_{24} - n_{32} + n_{34} + n_{42} - n_{43} = n'_{23} - n'_{24} - n_{32} + n_{34} + n_{42} - n_{43} \dots$ hasta aquí hemos completado las ecuaciones correspondientes a la habilidad combinatoria específica y para los efectos ligados al sexo, para todas las otras ecuaciones se practican las mismas operaciones (resta y suma), así mismo debe completarse los RHM's.

Después de que la matriz de coeficientes hayan sido reducidos quedará una matriz simétrica de orden $p(p-1)$ donde p es el número de líneas incluídas en el estudio.

4) Inversión de la matriz reducida.

Si la matriz es muy grande, es muy laborioso y hasta casi

prohibitivo hacerlo con calculadora de escritorio, es mejor consultar a instituciones de Estados Unidos u otras partes donde ya disponen de los programas para cálculos de esta naturaleza y los equipos electrónicos necesarios.

5) Cálculo de los promedios de mínimos cuadrados y errores estandares.

Esto no presenta ninguna dificultad, se realiza en la forma familiar para nosotros.

6) Cálculos de las sumas de cuadrados.

Las sumas de cuadrados para g_i , m_j , c_{ij} , y r_{ij} se obtienen directamente por la fórmula $B'Z^{-1}B$.

Las sumas de cuadrados también pueden ser calculadas indirectamente, por diferencias en la siguiente forma:

- a) SC habilidad combinatoria general = $R(\mu, g, m, c, r) - R(\mu, m, c, r)$
 - b) SC habilidad materna = $R(\mu, g, m, c, r) - R(\mu, g, e, r)$
 - c) SC habilidad combinatoria específica = $R(\mu, g, m, c, r) - R(\mu, g, m, r)$
 - d) SC efecto ligado al sexo = $R(\mu, g, m, c, r) - R(\mu, g, m, c)$
 - e) SC error = $\sum_i \sum_j \sum_k^2 Y_{ijk} - R(\mu, g, m, c, r)$
- 7) Grados de libertad.

FV	GL
Habilidad combinatoria general	p-1
Habilidad materna	p-1
Habilidad combinatoria específica	$\frac{p(p-3)}{2}$
Efectos ligados al sexo	$\frac{p(p-3)}{2} + 1$
Error	n - p (p-1)
Total	n - 1

De esto se desprende que cuando el estudio incluye solo tres líneas, no se puede estimar la habilidad combinatoria específica y solo se tiene un grado de libertad para efectos ligados al sexo.

Cuando algunos de los efectos no son significativos, las ecuaciones de estos efectos se pueden despreciar del modelo a fin de estimar más eficientemente los efectos que quedan, después de haber eliminado ciertos efectos, la inversa de la matriz quedaría como C_A^{-1} que es igual a:

$$C_A^{-1} = C_R^{-1} - RZ^{-1}R'$$

donde Z^{-1} es el segmento eliminado por hilera y columna $R(R')$ es el segmento de la matriz inversa que asocia a estos efectos con los que quedan y C_R^{-1} es el segmento correspondiente por hileras y columnas.

Las nuevas constantes estimadas de los efectos que quedan, B_A se pueden obtener por medio de la fórmula: $B_A = B_2 - RZ^{-1}B_1$. donde B_2 representa las constantes estimadas para los efectos que quedan antes del ajuste y B_1 las constantes estimadas de los efectos eliminados. Otra manera fácil de estimar las constantes de los efectos que quedan, es multiplicando los RHM's por cada hilera de la nueva matriz.

8) Estimación de los componentes de variancias.

Si se asume que todos los efectos están presentes, como efectos al azar se puede estimar los componentes de variancia en la siguiente forma.

FV	Componentes estimados E(MS)
Habilidad combinatoria general	$\sigma_e^2 + K_6 \sigma_r^2 + K_7 \sigma_c^2 + K_8 \sigma_g^2$
Habilidad materna	$\sigma_e^2 + K_4 \sigma_r^2 + K_5 \sigma_m^2$
Habilidad combinatoria específica	$\sigma_e^2 + K_2 \sigma_r^2 + K_3 \sigma_c^2$
Efectos ligados al sexo	$\sigma_e^2 + K_1 \sigma_r^2$
Error	σ_e^2

Los valores K_1 , K_3 , K_5 y K_8 se pueden estimar por el procedimiento directo ya explicado; K_2 , K_4 , K_6 y K_7 se pueden calcular por el procedimiento indirecto también ya explicado.

Ejemplo numérico

A continuación se presentan los datos correspondientes a 208 ratas (a los 28 días), resultado de un experimento llevado a cabo por Kidwell et al .

Línea padre	L i n e a m a d r e				Suma	
	1	2	3	4		
1	Número	22	12	10	12	56
	Total	1231.0	767.2	621.0	787.5	3406.7
	Promedio	55.955	63.933	62.100	65.625	
2	Número	12	26	10	10	58
	Total	712.0	1725.0	597.4	610.5	3644.9
	Promedio	59.333	66.346	59.740	61.050	
3	Número	12	12	12	12	48
	Total	771.3	773.9	645.5	799.9	2990.6
	Promedio	64.275	64.492	53.792	66.658	
4	Número	14	8	12	12	46
	Total	860.8	459.2	605.4	731.6	2657.0
	Promedio	61.486	57.400	50.450	60.967	
Suma	Número	60	58	44	46	208
	Total	3575.1	3725.3	2469.3	2929.5	12699.2

1) Ecuaciones de mínimos cuadrados bajo el modelo

$$y_{ijk} = \mu + \xi_i + \xi_j + \xi_k + \eta_j + \eta_k + \epsilon_{ijk} + \tau_{ij} + \tau_{ik} + \tau_{jk} + e_{ijk}$$

	$\hat{\mu}$	$\hat{\xi}_1$	$\hat{\xi}_2$	$\hat{\xi}_3$	$\hat{\xi}_4$	$\hat{\eta}_1$	$\hat{\eta}_2$	$\hat{\eta}_3$	$\hat{\eta}_4$	$\hat{\epsilon}_{12}$	$\hat{\epsilon}_{13}$	$\hat{\epsilon}_{14}$	$\hat{\epsilon}_{23}$	$\hat{\epsilon}_{24}$	$\hat{\epsilon}_{34}$	$\hat{\tau}_{12}$	$\hat{\tau}_{13}$	$\hat{\tau}_{14}$	$\hat{\tau}_{21}$	$\hat{\tau}_{23}$	$\hat{\tau}_{24}$	$\hat{\tau}_{31}$	$\hat{\tau}_{32}$	$\hat{\tau}_{34}$	$\hat{\tau}_{41}$	$\hat{\tau}_{42}$	$\hat{\tau}_{43}$	RHM	
μ :	136	72	64	68	38	32	32	32	34	24	22	26	22	18	24	12	10	12	12	12	10	10	12	12	12	14	8	12	8366.1
ξ_1 :	72	72	24	26	38	12	10	12	24	24	22	26	22	18	24	12	10	12	12	12	10	10	12	12	14	8	12	14	4519.8
ξ_2 :	64	24	64	22	18	12	32	10	10	24	22	26	22	18	24	12	10	12	12	12	10	10	12	12	14	8	12	14	3920.2
ξ_3 :	68	22	22	68	24	12	12	32	12	12	22	26	22	18	24	10	10	12	12	12	10	10	12	12	14	8	12	14	4168.9
ξ_4 :	68	26	18	24	68	14	8	12	34	12	12	14	22	18	24	12	10	12	12	12	10	10	12	12	14	8	12	14	4123.3
η_1 :	38	38	12	12	14	38				12	12	14	12	8	24	12	10	12	12	12	10	10	12	12	14	8	12	14	2344.1
η_2 :	32	12	32	12	8	32	32	32		12	10	12	12	8	24	12	10	12	12	12	10	10	12	12	14	8	12	14	2000.3
η_3 :	32	10	10	32	12	32	32	32		10	10	12	12	10	24	12	10	12	12	12	10	10	12	12	14	8	12	14	1823.8
η_4 :	34	12	10	12	34				34	24	22	26	22	10	24	12	10	12	12	12	10	10	12	12	14	8	12	14	2197.9
ϵ_{12} :	24	24	24			12	12	12	12	24	22	26	22	10	24	12	10	12	12	12	10	10	12	12	14	8	12	14	1479.2
ϵ_{13} :	22	22	22	22		12	12	10	12	24	22	26	22	10	24	12	10	12	12	12	10	10	12	12	14	8	12	14	1392.3
ϵ_{14} :	26	26		26	14	12	14	12	12		22	26	22	10	24	12	10	12	12	12	10	10	12	12	14	8	12	14	1648.3
ϵ_{23} :	22	22	22	22		12	12	12	12		22	26	22	18	24	12	10	12	12	12	10	10	12	12	14	8	12	14	1371.3
ϵ_{24} :	18	18	18	18		8	8	8	10		18	22	22	18	24	12	10	12	12	12	10	10	12	12	14	8	12	14	1069.7
ϵ_{34} :	24	24	24	24		12	12	12	12		24	26	22	10	24	12	10	12	12	12	10	10	12	12	14	8	12	14	1405.3
τ_{12} :	12	12	12	12		12	12	12	12		12	14	12	10	24	12	10	12	12	12	10	10	12	12	14	8	12	14	767.2
τ_{13} :	10	10	10	10		10	10	10	10		10	12	12	10	24	12	10	12	12	12	10	10	12	12	14	8	12	14	621.0
τ_{14} :	12	12	12	12		12	12	12	12		12	14	12	10	24	12	10	12	12	12	10	10	12	12	14	8	12	14	787.5
τ_{21} :	12	12	12	12		12	12	12	12		12	14	12	10	24	12	10	12	12	12	10	10	12	12	14	8	12	14	712.0
τ_{23} :	10	10	10	10		10	10	10	10		10	12	12	10	24	12	10	12	12	12	10	10	12	12	14	8	12	14	597.4
τ_{24} :	10	10	10	10		10	10	10	10		10	12	12	10	24	12	10	12	12	12	10	10	12	12	14	8	12	14	610.5
τ_{31} :	12	12	12	12		12	12	12	12		12	14	12	10	24	12	10	12	12	12	10	10	12	12	14	8	12	14	771.3
τ_{32} :	12	12	12	12		12	12	12	12		12	14	12	10	24	12	10	12	12	12	10	10	12	12	14	8	12	14	773.9
τ_{34} :	12	12	12	12		12	12	12	12		12	14	12	10	24	12	10	12	12	12	10	10	12	12	14	8	12	14	779.9
τ_{41} :	14	14	14	14		14	14	14	14		14	16	14	12	24	12	10	12	12	12	10	10	12	12	14	8	12	14	860.8
τ_{42} :	8	8	8	8		8	8	8	8		8	10	8	12	24	12	10	12	12	12	10	10	12	12	14	8	12	14	459.2
τ_{43} :	12	12	12	12		12	12	12	12		12	14	12	10	24	12	10	12	12	12	10	10	12	12	14	8	12	14	605.4

En este cuadro se puede observar que el segmento correspondiente por hileras y columnas a la habilidad combinatoria general están llenadas en forma completa; considerando los segmentos correspondientes por hilera y columna a los otros efectos que tienen entradas solo en las diagonales, además los RHM's para la habilidad combinatoria general no suma igual a la ecuación

2) Imponiendo las restricciones $\sum_i \hat{g}_i = \sum_j \hat{m}_j = \sum_j \hat{c}_{ij} = \sum_j \hat{c}_{ij} = \sum_i \hat{r}_{ij} = \sum_j \hat{r}_{ij} = \hat{r}_{ij} = \hat{r}_{ji} = 0$ se procedió a la reducción de la matriz original, dando como resultado el cuadro que sigue.

Ecuaciones reducidas de mínimos cuadrados y RHM's

	$\hat{\mu}$	\hat{g}_1	\hat{g}_2	\hat{g}_3	\hat{m}_1	\hat{m}_2	\hat{m}_3	\hat{c}_{12}	\hat{c}_{13}	\hat{r}_{12}	\hat{r}_{13}	\hat{r}_{23}	RHM
μ :	136	4	-4	0	4	-2	-2	0	-8	4	0	-4	8366.1
g_1 :	4	<u>88</u>	48	40	46	26	20	0	4	-2	-2	2	396.5
g_2 :	-4	48	<u>96</u>	48	22	48	22	4	4	-2	-2	-2	-203.1
g_3 :	0	40	48	<u>88</u>	20	26	42	4	8	-4	-4	0	45.6
m_1 :	4	46	22	20	<u>72</u>	34	34	-2	0	4	2	-2	146.2
m_2 :	-2	26	48	26	34	<u>66</u>	34	0	-2	6	0	-6	-197.6
m_3 :	-2	20	22	42	34	34	<u>66</u>	2	2	2	-2	-4	-374.1
c_{12} :	0	0	4	4	-2	0	2	<u>96</u>	48	-2	-2	2	-135.1
c_{13} :	-8	4	4	8	0	-2	2	48	<u>88</u>	0	-4	0	-557.6
r_{12} :	4	-2	-2	-4	4	6	2	-2	0	<u>68</u>	26	-18	279.8
r_{13} :	0	-2	-2	-4	2	0	-2	-2	-4	26	<u>72</u>	24	117.5
r_{23} :	-4	2	-2	0	-2	-6	-4	2	0	-18	24	<u>64</u>	-133.3

Matriz inversa variancia-covariancia ($\times 10^{-6}$)

	$\hat{\mu}$	$\hat{\epsilon}_1$	$\hat{\epsilon}_2$	$\hat{\epsilon}_3$	\hat{m}_1	\hat{m}_2	\hat{m}_3	\hat{c}_{12}	\hat{c}_{13}	\hat{r}_{12}	\hat{r}_{13}	\hat{r}_{23}	Constant Estimates
μ	: 7498	-434	645	-484	-595	174	694	-554	1009	-384	50	434	61.378
ϵ_1	: -434	25651	-9375	-6901	-16667	4948	5729	434	-347	-391	1302	-1450	2.941
ϵ_2	: 645	-9375	25595	-8519	6101	-17708	5729	-645	-124	1488	-856	781	-1.294
ϵ_3	: -484	-6901	-8519	24163	5580	4427	-15625	484	-1339	577	186	-391	2.753
m_1	: -595	-16667	6101	5580	31548	-10417	-11458	595	-446	149	-1414	521	-1.641
m_2	: 174	4948	-17708	4427	-10417	35937	-11458	-174	1389	-2344	521	781	1.426
m_3	: 694	5729	5729	-15625	-11458	-11458	33333	-694	347	0	521	521	-5.784
c_{12}	: -554	434	-645	484	595	174	-694	14443	-7953	384	-50	-434	-1.285
c_{13}	: 1009	-347	-124	-1339	-446	1389	347	-7953	16005	-918	992	-174	-1.172
r_{12}	: -384	-391	1488	577	149	-2344	0	384	-918	22080	-11384	10286	.767
r_{13}	: .50	1302	-856	186	-1414	521	521	-50	992	-11384	21949	-11458	.984
r_{23}	: 434	-1450	781	-391	521	781	521	-434	-174	10286	-11458	23047	1.229

3) Inversión de la matriz reducida y cálculo de constantes.

Una vez invertida la matriz, se pueden calcular las constantes en la forma usual $\sum C^{ij} Y_j = \hat{c}_i$. (Ver cuadro página siguiente).

4) Cálculo de las sumas de cuadrados.

a) SC de reducción

$$R(\mu, g, m, c, r) = (61.378)(8366.1) + \dots + (1.229)(-133.3) = 517,127$$

b) SC del error = $\sum_i \sum_j \sum_k y_{ijk}^2 - R(\mu, g, m, c, r)$.

Sin embargo, por lo explicado más arriba, en este análisis el efecto ligado al sexo y sexo por sub-clase se englobaría en el error. Por esta razón la SC del error se obtuvo por sustracción de la suma de cuadrados corregida por efecto ligado al sexo y sexo por sub-clase de la SC total sin corregir y se encontró que la SC del error con GL (136 - 24), 122 era igual a 5,786.3.

c) SC del efecto ligado al sexo = $B' Z^{-1} B$.

$$\begin{bmatrix} 0.767 & 0.984 & 1.229 \end{bmatrix} (10^{-6}) \begin{bmatrix} \underline{22080} & -11384 & 10286 \\ -11384 & \underline{21949} & -11458 \\ 10286 & -11458 & \underline{23047} \end{bmatrix}^{-1} \begin{bmatrix} 0.767 \\ 0.984 \\ 1.229 \end{bmatrix}$$

$$= \begin{bmatrix} 0.767 & 0.984 & 1229 \end{bmatrix} \begin{bmatrix} -66.3559 & 25.6001 & -16.8877 \\ 25.6001 & 71.4053 & 24.0742 \\ -16.8877 & 24.0742 & 62.8954 \end{bmatrix} \begin{bmatrix} 0.767 \\ 0.984 \\ 1.229 \end{bmatrix}$$

$$= 268.21$$

Prueba de significancia del efecto ligado al sexo

FV	GL	SC	CM	F
Efectos ligados al sexo	3	268.2	89.4	1.89 NS
Error	122	5,786.3	47.4	

Puesto que las diferencias entre \hat{r}_{ij} no son significativas el efecto ligado al sexo será eliminado del análisis para tener mayor eficiencia en la estimación de μ , g_i , m_j y s_{ij} . Los elementos inversos del cuadro presentado en esta sección es decir el cuadro correspondiente a las ecuaciones reducidas, de mínimos cuadrados serán ajustadas por el uso de la fórmula: $C_A^{-1} = C_R^{-1} - R Z^{-1} R'$, hecha la operación indicada resulta el siguiente cuadro.

La matriz $R Z^{-1} R'$ usada para remover r_{ij} ($\times 10^6$)

	$\hat{\mu}$	\hat{g}_1	\hat{g}_2	\hat{g}_3	\hat{m}_1	\hat{m}_2	\hat{m}_3	\hat{c}_{12}	\hat{c}_{13}
μ :	<u>28</u>	-33	-23	-31	8	103	20	-28	27
g_1 :	-33	<u>127</u>	-35	48	-91	-91	-3	33	28
g_2 :	-23	-35	<u>101</u>	35	17	-149	0	23	-63
g_3 :	-31	48	35	<u>44</u>	-35	-130	6	31	-22
m_1 :	8	-91	17	-35	<u>112</u>	35	-46	-8	-51
m_2 :	103	-91	-149	-130	35	<u>441</u>	51	-103	121
m_3 :	20	-3	0	6	-46	51	<u>50</u>	-20	37
c_{12} :	-28	33	23	31	-8	-103	-20	<u>28</u>	-27
c_{13} :	27	28	-63	-22	-51	121	37	-27	<u>68</u>

Estos valores fueron sustraídos elemento por elemento del segmento correspondiente a los efectos que quedaron.

La nueva matriz inversa que resulta se presenta en el siguiente cuadro.

Matriz inversa (C_A^{-1}) de la matriz variancia-covariancia descartando los efectos rij. ($\times 10^6$)

	$\hat{\mu}$	\hat{g}_1	\hat{g}_2	\hat{g}_3	\hat{m}_1	\hat{m}_2	\hat{m}_3	\hat{c}_{12}	\hat{c}_{13}
μ :	<u>7470</u>	-401	668	-453	-603	71	674	-526	982
g_1 :	-401	<u>25524</u>	-9340	-6949	-16576	5039	5732	401	-375
g_2 :	668	-9340	<u>25494</u>	-8554	6084	-17559	5729	-668	-61
g_3 :	-453	-6949	-8554	<u>24119</u>	5615	4557	-15619	453	-1317
m_1 :	-603	-16576	6084	5615	<u>31436</u>	-10452	-11412	603	-395
m_2 :	71	5039	-17539	4557	-10452	<u>35496</u>	-11509	277	1268
m_3 :	674	5732	5729	-15619	-11412	-11509	<u>33283</u>	-674	310
c_{12} :	-526	401	-668	453	603	277	-674	<u>14415</u>	-7926
c_{13} :	982	-375	61	-1317	-395	1268	310	-7926	<u>15937</u>

5) Estimación de las nuevas constantes.

Puesto que los rij se asumieron igual a cero los RHM's se omiten en el cálculo de las nuevas constantes y se procede al cálculo en la forma ya conocida.

$$\hat{\mu} = 10^{-6} [(7470)(8366.1) + \dots + (982)(-557.6)] = 61.35$$

$$\hat{g}_1 = 2.94 \qquad \hat{m}_1 = -1.53 \qquad \hat{c}_{12} = -1.33$$

$$\hat{m}_2 = 1.38 \qquad \hat{c}_{13} = -0.22$$

$$\hat{g}_2 = -1.34 \qquad \hat{m}_3 = -5.89$$

$$\hat{g}_3 = 2.73$$

Las demás constantes se pueden obtener por diferencia

$$\hat{g}_3 = -4.33 \qquad \hat{m}_4 = 6.04$$

y los \hat{c}_{ij} también se calculan por diferencia sabiendo que \hat{c}_{ij} suman cero por hilera y columna y por $\sum_l \sum_j \hat{c}_{ij}$.

6) Cálculo de las sumas de cuadrados.

a) SC reducción $R(\mu, g, m, c)$

$$R(\mu, g, m, c) = (61,35)(8366.1) + \dots + (0.22)(-557.6) = 516,472.7$$

b) SC error = 5,786.3

c) SC habilidad combinatoria general $B' Z^{-1} B$

$$SCHCG = \begin{bmatrix} 2.94 & -1.34 & 2.73 \end{bmatrix} \begin{bmatrix} 0.025524 & -0.009340 & -0.006949 \\ -0.009340 & 0.025494 & -0.008554 \\ -0.006949 & -0.008554 & 0.024119 \end{bmatrix}^{-1} \begin{bmatrix} 2.94 \\ -1.34 \\ 2.73 \end{bmatrix}$$

$$= \begin{bmatrix} 2.94 & -1.34 & 2.73 \end{bmatrix} \begin{bmatrix} 57.7447 & 30.3487 & 27.4003 \\ 30.3487 & 60.4735 & 30.1920 \\ 27.4003 & 30.1920 & 60.0631 \end{bmatrix} \begin{bmatrix} 2.94 \\ -1.34 \\ 2.73 \end{bmatrix}$$

$$= \begin{bmatrix} 203.9050 & 90.6148 & 204.0219 \end{bmatrix} \begin{bmatrix} 2.94 \\ -1.34 \\ 2.73 \end{bmatrix} = 1035.2$$

d) SC Habilidad materna

$$= \begin{bmatrix} -1.53 & 1.38 & -5.89 \end{bmatrix} \begin{bmatrix} .031436 & -.010452 & -.011412 \\ -.010452 & .035496 & -.011509 \\ -.011412 & -.011509 & .033283 \end{bmatrix}^{-1} \begin{bmatrix} -1.53 \\ 1.38 \\ -5.89 \end{bmatrix}$$

$$= \begin{bmatrix} -1.53 & 1.38 & -5.89 \end{bmatrix} \begin{bmatrix} 47.7436 & 21.8113 & 23.9123 \\ 21.8113 & 41.6942 & 21.8964 \\ 23.9123 & 21.8964 & 45.8169 \end{bmatrix} \begin{bmatrix} -1.53 \\ 1.38 \\ -5.89 \end{bmatrix}$$

$$= \begin{bmatrix} -183.7916 & -104.8031 & -276.2250 \end{bmatrix} \begin{bmatrix} -1.53 \\ 1.38 \\ -5.89 \end{bmatrix}$$

$$= 1763.5$$

e) SC Habilidad combinatoria especifica

$$= \begin{bmatrix} -1.33 & -.22 \end{bmatrix} \begin{bmatrix} .014415 & -.007926 \\ -.007926 & .015937 \end{bmatrix}^{-1} \begin{bmatrix} -1.33 \\ -.22 \end{bmatrix}$$

$$= \begin{bmatrix} -1.33 & -.22 \end{bmatrix} \begin{bmatrix} 95.4826 & 47.4867 \\ 47.4867 & 86.3639 \end{bmatrix} \begin{bmatrix} -1.33 \\ -.22 \end{bmatrix}$$

$$= \begin{bmatrix} 137.4389 & 82.1574 \end{bmatrix} \begin{bmatrix} -1.33 \\ -.22 \end{bmatrix}$$

= 200.9

7) Análisis de variancia completa

FV	GL	SC	CM	F
Habilidad combinatoria general (G)	3	1035.2	345.1	7.28 **
Habilidad materna (M)	3	1763.5	587.8	12.40 **
Habilidad combinatoria especifica (C)	2	200.9	100.4	2.18 N S
Error	122	5786.3	47.4	

El efecto de la habilidad combinatoria especifica no es significativa, lo cual indica que hubiéramos podido despreciar del modelo con la misma forma que los efectos ligados al sexo.

8) Estimación de los componentes de variancia.

FV	GL	Componente estimado E(MS)
Habilidad combinatoria general	3	$\sigma_e^2 + K_3 \sigma_c^2 + K_4 \sigma_g^2$
Habilidad materna	3	$\sigma_e^2 + K_2 \sigma_m^2$
Habilidad combinatoria especifica	2	$\sigma_e^2 + K_1 \sigma_c^2$
Error	122	σ_e^2

Cálculo de los coeficientes.

Los coeficientes K's para los componentes principales se calculan por la fórmula general ya vista

$$K = \frac{\sum z^{ii} - \frac{1}{GL} \sum \sum z^{ii}}{m}$$

$$K_1 = \frac{181.8465 - 47.4867}{6} = 22.39$$

$$K_2 = \frac{135.2538 - 45.0800}{4} = 22.54$$

$$K_4 = \frac{178.2813 - 58.6273}{4} = 29.91$$

El coeficiente K_3 que aparece con $\hat{\sigma}_c^2$ asociado al componente habilidad combinatoria general se puede calcular cuando todas las celdas en las sub-clases están completas, aplicando la fórmula siguiente:

$$K' = \frac{\sum z^{ii} - \frac{1}{GL} \sum \sum z^{ij}}{m(m-2)}$$

$$K_3 = \frac{178.2813 - 58.6273}{4(4-2)} = 14.96$$

Cálculo de los componentes de variancia.

$$\hat{\sigma}_c^2 = \frac{100.4 - 47.4}{22.39} = 2.4$$

$$\hat{\sigma}_{m^2} = \frac{587.8 - 47.4}{22.54} = 24.0$$

$$\hat{\sigma}_g^2 = \frac{345.1 - 47.4 - (14.96)(2.4)}{29.91} = 8.8$$

9) Comparaciones de promedios por la prueba de Duncan.

Vamos a considerar la habilidad combinatoria general de los

distintas líneas, también si se quiere aplicar para cualquier otro efecto se puede seguir en la misma forma.

Los cálculos son familiares para nosotros así omitimos la explicación.

Comparación de promedios para el efecto de la habilidad combinatoria general (P.05)

Comparación	$\bar{Y}_i - \bar{Y}_j$	$\sqrt{\frac{2}{c^{ii} + c^{jj} - 2c^{ij}}}$	Diferencias Producto	$G_{\alpha} Z_{p, n_2}$
ϵ_1 vs ϵ_3	0.21	5.610	1.18	19.28
ϵ_1 vs ϵ_2	4.28 [*]	5.357	22.93	20.31
ϵ_1 vs ϵ_4	7.27 [*]	5.366	39.01	21.00
ϵ_3 vs ϵ_2	4.07 [*]	5.475	22.28	19.28
ϵ_3 vs ϵ_4	7.06 [*]	5.472	38.63	20.31
ϵ_2 vs ϵ_4	2.99	5.499	16.44	19.28

ANALISIS SIMULTANEO DE DATOS DE CRUZAS Y PUROS

Cuando es posible incluir las endocrías o puros en el diseño de un experimento de crusa simple entonces se dispone de un estimador de la heterosis. El efecto del puro se puede estimar independientemente o simultaneamente con la habilidad combinatoria general y efecto de la habilidad materna. El análisis que se va a considerar en esta sección incluye los puros en el experimento, pero su efecto se va a estimar independientemente de la habilidad combinatoria general y el efecto materno.

1) Modelo matemático (II)

$$Y_{hijk} = \mu + a_h + P_{1ii} + G_{2i} + S_{2j} + M_{2j} + C_{2ij} + r_{2ij} + e_{hijk}$$

donde:

Y_{hijk} = la observación K^{th} en la progenie proveniente de apareamiento de una vaca de la línea j^{th} con un toro de la línea i^{th} en el tipo h^{th} breeding (puro o crusa).

μ = Promedio general (con igual frecuencia en la sub-clase.

a_h = Efecto común a todas las progenies del tipo de breeding h^{th} (puros o cruza), la diferencia entre estos efectos constituye un estimador de la heterosis.

P_{1ii} = Efecto común a todas las progenies de un apareamiento entre una vaca de la línea i^{th} y un toro de la línea i^{th} .

Los demás símbolos representados en el modelo, tienen el mismo significado que en el modelo presentado anteriormente.

2) Ecuaciones de mínimos cuadrados.

Las ecuaciones de mínimos cuadrados para este modelo se presentan en el cuadro siguiente. Como se puede notar en este cuadro solo la suma de los coeficientes para heterosis es igual al coeficiente μ y ninguno de los otros efectos.

3) Imposición de restricciones.

Las restricciones impuestas a las constantes en este modelo de análisis son similares a las anteriores excepto que los efectos tipos-de-breeding y los de los puros se deben tomar dentro de su grupo.

$$\sum_h \hat{a}_h = \sum_i \hat{P}_{1ii} = \sum_i \hat{G}_{2i} = \sum_j \hat{M}_{2j} = \sum_i \hat{C}_{ij} = \sum_j \hat{C}_{ji} = \sum_l \sum_j \hat{C}_{lj} = \sum_l \hat{r}_{lj} = \sum_j \hat{r}_{ij} = \hat{r}_{ij} + \hat{r}_{ji} = 0.$$

Ecuaciones de mínimos cuadrados

	$\hat{\mu}$	$\hat{\sigma}_h$	\hat{p}_{11}	$\hat{\epsilon}_{21}$	\hat{m}_{2j}	\hat{c}_{21j}	\hat{r}_{21j}	RHM
μ	$n_{...}$	$n_{h..}$	n_{111}	$n_{21.} + n_{2.1}$	$n_{2..j}$	$n_{21j} + n_{2j1}$	n_{21j}	$Y_{...}$
σ_h	$n_{h..}$	$0^{n_j..0}$	n_{111}	$n_{21.} + n_{2.1}$	$n_{2..j}$	$n_{21j} + n_{2j1}$	n_{21j}	$Y_{h..}$
p_{11}	$n_{1..}$	$n_{1..}$	$0^{n_{1..}0}$	-	-	-	-	Y_{111}
ϵ_{q1}	$n_{21.} + n_{2.1}$	$n_{21.} + n_{2.1}$	-	$n_{21j} + n_{2j1}$	$n_{2..j}$	$n_{21j} + n_{2j1}$	n_{21j}	$Y_{21.} + Y_{2.1}$
m_{2j}	$n_{2..j}$	$n_{2..j}$	-	$n_{2..j}$	$0^{n_{2..j}0}$	n_{21j}	n_{21j}	$Y_{2..j}$
$c_{2k,j}$	$n_{21j} + n_{2j1}$	$n_{21j} + n_{2j1}$	-	$n_{21j} + n_{2j1}$	n_{21j}	$0^{n_{21j} + n_{2j1}0}$	n_{21j}	$Y_{21j} + Y_{2j1}$
r_{21j}	n_{21j}	n_{21j}	-	n_{21j}	n_{21j}	n_{21j}	$0^{n_{21j}0}$	Y_{21j}

Estas restricciones exigen a las constantes estimadas para efecto tipo-de-breeding sumar a cero respecto al promedio $\hat{\mu}$. A las constantes estimadas para puros a sumar a cero respecto a $\hat{\mu} + \hat{a}_1$ y a las constantes estimadas que quedan a sumar cero respecto a $\hat{\mu} + \hat{a}_2$. Las restricciones a las constantes estimadas para las \hat{a}_h y \hat{p}_{1ii} se efectúan por sustracción, por hileras y columnas como se discutió en la primera sección. Las restricciones que aun quedan se imponen por el procedimiento bosquejado para el modelo anterior.

4) Inversión de la matriz reducida.

Después de las restricciones discutidas quedará una matriz de orden p^2 . Invertir este tipo es por supuesto muy laborioso sin una calculadora de alta velocidad.

5) Cálculo de los promedios de mínimos cuadrados y errores estándares.

Para los cálculos se sigue el mismo procedimiento ya explicado.

Para los promedios de mínimos cuadrados para los puros están dado por $\hat{\mu} + \hat{a}_1$ y los promedios para cruzas $\hat{\mu} + \hat{a}_2$, para los promedios por grupo de puros es $\hat{\mu} + \hat{a}_1 + \hat{p}_{1ii}$.

Los promedios para el grupo de cruzas está dado por $\hat{\mu} + \hat{a}_2 + \hat{g}_{21} + \hat{g}_{2j} + \hat{m}_{2j} + \hat{c}_{2ij} + \hat{r}_{2ij}$. Los errores estándares para estos promedios se determinan como sigue.

$$s_{\hat{\mu}} = \sqrt{c^{\mu\mu}} \hat{\sigma}_e$$

$$s_{(\hat{\mu} + \hat{a}_1)} = \sqrt{c^{\mu\mu} + c^{a_1a_1} + 2 c^{\mu a_1}} \hat{\sigma}_e$$

$$s_{(\hat{\mu} + \hat{a}_2)} = \sqrt{c^{\mu\mu} + c^{a_2a_2} + 2 c^{\mu a_2}} \hat{\sigma}_e$$

$$s(\hat{\mu} + \hat{a}_1 + \hat{p}_{11i}) = \sqrt{c^{\mu\mu} + c^{a_1 a_1} + c^{p_{11i} p_{11i}} + 2c^{\mu a_1} + 2c^{\mu p_{11i}} + 2c^{a_1 p_{11i}}} \quad \hat{\sigma}_e$$

$$s(\hat{\mu} + \hat{a}_2 + \hat{g}_{21} + \hat{g}_{2j} + \hat{m}_{2j} + \hat{c}_{21j} + \hat{r}_{21j}) = \sqrt{(c^{\mu\mu} + c^{a_i a_i} + c^{g_{21} g_{21}} + c^{g_{2j} g_{2j}} + c^{m_{2j} m_{2j}} + c^{c_{21j} c_{21j}} + c^{r_{21j} r_{21j}} + 2c^{\mu a_2} + 2c^{\mu g_{21}} + 2c^{\mu g_{2j}} + 2c^{\mu m_{2j}} + 2c^{\mu c_{21j}} + 2c^{\mu r_{21j}} + 2c^{a_2 g_{21}} + 2c^{a_2 g_{2j}} + 2c^{a_2 m_{2j}} + 2c^{a_2 c_{21j}} + 2c^{a_2 r_{21j}} + 2c^{g_{21} g_{2j}} + 2c^{g_{21} m_{2j}} + 2c^{g_{21} c_{1j}} + 2c^{g_{21} r_{21j}} + 2c^{g_{2j} m_{2j}} + 2c^{g_{2j} c_{21j}} + 2c^{g_{2j} r_{21j}} + 2c^{m_{2j} c_{21j}} + 2c^{m_{2j} r_{21j}} + 2c^{c_{21j} r_{21j}})^{\frac{1}{2}} \hat{\sigma}_e$$

$$+ c^{c_{21j} c_{21j}} + c^{r_{21j} r_{21j}} + 2c^{\mu a_2} + 2c^{\mu g_{21}} + 2c^{\mu g_{2j}} + 2c^{\mu m_{2j}} + 2c^{\mu c_{21j}} + 2c^{\mu r_{21j}} + 2c^{a_2 g_{21}} + 2c^{a_2 g_{2j}} + 2c^{a_2 m_{2j}} + 2c^{a_2 c_{21j}} + 2c^{a_2 r_{21j}} \hat{\sigma}_e^2$$

$$= 2c^{g_{21} g_{2j}} + 2c^{g_{21} m_{2j}} + 2c^{g_{21} c_{1j}} + 2c^{g_{21} r_{21j}} + 2c^{g_{2j} m_{2j}} + 2c^{g_{2j} c_{21j}} + 2c^{g_{2j} r_{21j}} + 2c^{m_{2j} c_{21j}} + 2c^{m_{2j} r_{21j}} + 2c^{c_{21j} r_{21j}})^{\frac{1}{2}} \hat{\sigma}_e$$

La variancia del error $\hat{\sigma}_e^2$ normalmente se obtiene por la fórmula:

$$\hat{\sigma}_e^2 = \frac{\sum_i \sum_j \sum_k \sum_l Y_{hijkl} - R(\mu, a, g, m, c, r)}{GL \text{ del error}}$$

6) Cálculo de las sumas de cuadrados.

Este se obtiene en la forma usual aplicando la fórmula $B'Z^{-1}B$

o se puede obtener por diferencia entre las SC de reducciones.

Análisis de variancia

FV	GL
Heterosis	1
Entre puros	p-1
Habilidad combinatoria general	p-1
Habilidad materna	p-1
Habilidad combinatoria específica	$\frac{p(p-3)}{2}$
Efecto ligado al sexo	$\frac{p(p-3)}{2} + 1$
Error	n - p ²
Total	n - 1

Construcción y aplicación de las transformaciones matriz

El análisis de los datos descritos se puede facilitar grandemente por el uso de una transformación matriz para el modelo en cuestión, en general para este modelo y que puede ser seguido por otros análisis de mínimos cuadrados.

Para simplificar vamos a considerar $p=4$ con una sola observación en cada sub-clase; la matriz de coeficientes de las ecuaciones reducidas de mínimos cuadrados que resulta se presente a continuación.

Matriz de coeficientes de mínimos cuadrados reducidos con una observación por sub-clase para el modelo

$$Y_{hijk} = \mu + a_h + P_{111} + g_{21} + g_{2j} + m_{2j} + c_{21j} + r_{21j} + c_{nijK}$$

	$\hat{\mu}$	\hat{a}_1	\hat{P}_{111}	\hat{P}_{122}	\hat{P}_{133}	\hat{g}_{21}	\hat{g}_{22}	\hat{g}_{23}	\hat{m}_{21}	\hat{m}_{22}	\hat{m}_{23}	\hat{c}_{212}	\hat{c}_{213}	\hat{r}_{212}	\hat{r}_{213}	\hat{r}_{223}
μ	16	-8	0	0	0	0	0	0	0	0	0	0	0	0	0	0
a_1	-8	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P_{111}	0	0	2	1	1	0	0	0	0	0	0	0	0	0	0	0
P_{122}	0	0	1	2	1	0	0	0	0	0	0	0	0	0	0	0
P_{133}	0	0	1	1	2	0	0	0	0	0	0	0	0	0	0	0
g_{21}	0	0	0	0	0	8	4	4	4	2	2	0	0	0	0	0
g_{22}	0	0	0	0	0	4	8	4	2	4	2	0	0	0	0	0
g_{23}	0	0	0	0	0	4	4	8	2	2	4	0	0	0	0	0
m_{21}	0	0	0	0	0	4	2	2	6	3	3	0	0	0	0	0
m_{22}	0	0	0	0	0	2	4	2	3	6	3	0	0	0	0	0
m_{23}	0	0	0	0	0	2	2	4	3	3	6	0	0	0	0	0
c_{12}	0	0	0	0	0	0	0	0	0	0	0	8	4	0	0	0
c_{13}	0	0	0	0	0	0	0	0	0	0	0	4	8	0	0	0
r_{12}	0	0	0	0	0	0	0	0	0	0	0	0	0	6	2	-2
r_{13}	0	0	0	0	0	0	0	0	0	0	0	0	0	2	6	2
	0	0	0	0	0	0	0	0	0	0	0	0	0	-2	2	6

Esta matriz se invierte fácilmente, puesto que cada uno de los cinco segmentos se pueden invertir separadamente, para obtener la inversa.

El próximo paso consiste en escribir algebraicamente las subclases que se usan en la medida de cada uno de los efectos que se incluyen en el modelo. Imponiendo las restricciones apropiadas sobre estos valores, los siguientes RHM's son obtenidos:

$$\mu : \sum_{\lambda} \sum_j s_{1j}$$

$$a_1 : \sum_{\lambda} \sum_{\substack{j \\ i=j}} s_{1j} - \sum_{\lambda} \sum_{\substack{j \\ i \neq j}} s_{1j}$$

$$P_{111} : s_{11} - s_{44}$$

$$P_{122} : s_{22} - s_{44}$$

$$P_{133} : s_{33} - s_{44}$$

$$s_{21} : \left(\sum_j s_{1j} + \sum_{\lambda} s_{i1} \right) - \left(\sum_j s_{4j} + \sum_i s_{i4} \right)$$

$$s_{22} : \left(\sum_j s_{2j} + \sum_{\lambda} s_{i2} \right) - \left(\sum_j s_{4j} + \sum_i s_{i4} \right)$$

$$s_{23} : \left(\sum_j s_{3j} + \sum_{\lambda} s_{i3} \right) - \left(\sum_j s_{4j} + \sum_i s_{i4} \right)$$

$$m_{21} : \sum_{\lambda} s_{i1} - \sum_i s_{i4}$$

$$m_{22} : \sum_{\lambda} s_{i2} - \sum_i s_{i4}$$

$$m_{23} : \sum_{\lambda} s_{i3} - \sum_i s_{i4}$$

Matriz Inversa

	\hat{u}	\hat{a}_1	\hat{p}_{111}	\hat{p}_{122}	\hat{p}_{133}	\hat{g}_{21}	\hat{g}_{22}	\hat{g}_{23}	\hat{m}_{21}	\hat{m}_{22}	\hat{m}_{23}	\hat{c}_{212}	\hat{c}_{213}	\hat{f}_{212}	\hat{f}_{213}	\hat{f}_{223}
u	8	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
a_1	4	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0
p_{111}	0	0	72	-24	-24	0	0	0	0	0	0	0	0	0	0	0
p_{122}	0	0	-24	72	-24	0	0	0	0	0	0	0	0	0	0	0
p_{133}	0	0	-24	-24	72	0	0	0	0	0	0	0	0	0	0	0
g_{21}	0	0	0	0	0	27	-9	-9	-18	6	6	0	0	0	0	0
g_{22}	0	0	0	0	0	-9	27	-9	6	-18	6	0	0	0	0	0
g_{23}	0	0	0	0	0	-9	-9	27	6	6	-18	0	0	0	0	0
m_{21}	0	0	0	0	0	-18	6	6	36	-12	-12	0	0	0	0	0
m_{22}	0	0	0	0	0	6	-18	6	-12	36	-12	0	0	0	0	0
m_{23}	0	0	0	0	0	6	6	-18	-12	-12	36	0	0	0	0	0
c_{12}	0	0	0	0	0	0	0	0	0	0	0	16	-8	0	0	0
c_{13}	0	0	0	0	0	0	0	0	0	0	0	-8	16	0	0	0
r_{12}	0	0	0	0	0	0	0	0	0	0	0	0	0	24	-12	+12
r_{13}	0	0	0	0	0	0	0	0	0	0	0	0	0	-12	24	-12
r_{23}	0	0	0	0	0	0	0	0	0	0	0	0	0	+12	-12	24

$\frac{1}{96}$

$$c_{212}: s_{12} + s_{21} - s_{14} - s_{41} - s_{23} - s_{32} + s_{34} + s_{43}$$

$$c_{213}: s_{13} + s_{31} - s_{14} - s_{41} - s_{23} - s_{32} + s_{24} + s_{42}$$

$$r_{212}: s_{12} - s_{14} - s_{42} - s_{21} + s_{24} + s_{41}$$

$$r_{213}: s_{13} - s_{14} - s_{31} + s_{34} + s_{41} - s_{43}$$

$$r_{223}: s_{23} - s_{24} - s_{32} + s_{34} + s_{42} - s_{43}$$

donde s_{ij} se refiere a la sub-clase para la progenie que resulta del apareamiento de un toro de la línea i^{th} con una vaca de la línea j^{th} .

El tercero y último paso consiste en la multiplicación de los EHM's dado arriba por los coeficientes (matriz inversa)

$$\mu: \frac{1}{12} \sum_{ij} s_{ij} + \frac{1}{24} \left(\sum_{i=j} s_{ij} - \sum_{i \neq j} s_{ij} \right) - \frac{3}{24} \sum_{i=j} s_{ij} + \frac{1}{24} \sum_{i \neq j} s_{ij}$$

$$a_1: \frac{1}{24} \sum_{ij} s_{ij} + \frac{1}{12} \left(\sum_{i=j} s_{ij} - \sum_{i \neq j} s_{ij} \right) = \frac{3}{24} \sum_{i=j} s_{ij} - \frac{1}{24} \sum_{i \neq j} s_{ij}$$

$$p_{111}: \frac{3}{4}(s_{11}-s_{44}) - \frac{1}{4}(s_{22}-s_{44}) - \frac{1}{4}(s_{33}-s_{44}) = \frac{3}{4} s_{11} - \frac{1}{4} s_{22} - \frac{1}{4} s_{33} - \frac{1}{4} s_{44}$$

·
·
·

$$r_{223}: \frac{1}{8}(s_{12}-s_{14}-s_{42}-s_{21}+s_{24}+s_{41}) - \frac{1}{8}(s_{13}-s_{14}-s_{31}+s_{34}+s_{41}-s_{43})$$

$$+ \frac{1}{4}(s_{23}-s_{24}-s_{32}+s_{34}+s_{42}-s_{43}) = \frac{1}{8} s_{12} - \frac{1}{8} s_{13} - \frac{1}{8} s_{21} + \frac{1}{4} s_{23}$$

$$- \frac{1}{8} s_{24} + \frac{1}{8} s_{31} - \frac{1}{4} s_{32} + \frac{1}{8} s_{34} + \frac{1}{8} s_{42} - \frac{1}{8} s_{43}$$

Los valores que resultan de las multiplicaciones arriba indicadas, son los coeficientes de la matriz de transformación K. Si D es la matriz de coeficiente para las sub-classes, entonces la matriz inversa C^{-1} se obtiene por la fórmula $K D^{-1} K'$.

Las constantes estimadas se obtienen por $KS = B$ donde S es una columna vector de los promedios de sub-classes y B es la columna vector de las constantes estimadas.

Matriz de transformación para el modelo $Y_{hijk} = \mu + a_n + P_{lii} + S_{2j} + m_{2j} + c_{2ij} + r_{2ij} + e_{hijk}$. Efectos de puros estimados independientemente de la habilidad combinatoria general y habilidad materna

	$\hat{\mu}$	\hat{s}_{11}	\hat{s}_{12}	\hat{s}_{13}	\hat{s}_{14}	\hat{s}_{21}	\hat{s}_{22}	\hat{s}_{23}	\hat{s}_{24}	\hat{s}_{31}	\hat{s}_{32}	\hat{s}_{33}	\hat{s}_{34}	\hat{s}_{41}	\hat{s}_{42}	\hat{s}_{43}	\hat{s}_{44}
μ :	3	1	1	1	1	3	1	1	1	1	3	1	1	1	1	3	
a_1 :	3	-1	-1	-1	-1	3	-1	-1	-1	-1	3	-1	-1	-1	-1	3	
p_{111} :	18	0	0	0	0	-6	0	0	0	0	-6	0	0	0	0	-6	
p_{122} :	-6	0	0	0	0	18	0	0	0	0	-6	0	0	0	0	-6	
p_{133} :	-6	0	0	0	0	-6	0	0	0	0	18	0	0	0	0	-6	
s_{21} :	0	6	6	6	0	0	-3	-3	0	-3	0	-3	0	-3	-3	0	
s_{22} :	0	0	-3	-3	6	0	6	6	-3	0	0	-3	-3	0	-3	0	
$s_{23}^{\frac{1}{24}}$:	0	-3	0	-3	-3	0	0	-3	6	6	0	6	-3	-3	0	0	
m_{21} :	0	-6	-6	-6	6	0	0	0	6	0	0	0	6	0	0	0	
m_{22} :	0	6	0	0	-6	0	-6	-6	0	6	0	0	0	6	0	0	
m_{23} :	0	0	6	0	0	0	6	0	-6	-6	0	-6	0	0	6	0	
c_{212} :	0	4	-2	-2	4	0	-2	-2	-2	-2	0	4	-2	-2	4	0	
c_{213} :	0	-2	4	-2	-2	0	-2	4	4	-2	0	-2	-2	4	-2	0	
r_{212} :	0	6	-3	-3	-6	0	3	3	3	-3	0	0	3	-3	0	0	
r_{213} :	0	-3	6	-3	3	0	-3	0	-6	3	0	3	3	0	-3	0	
r_{223} :	0	3	-3	0	-3	0	6	-3	3	-6	0	3	0	3	-3	0	

Ejemplo numerico

Los datos usados anteriormente los vamos a usar para ilustrar este caso, incluyendo los puros.

En las ecuaciones de mínimos cuadrados se puede ver que solo el segmento correspondiente por hilera y columna a los efectos de la habilidad combinatoria general contiene entradas alrededor de la diagonal, todos los otros segmentos solo tienen entradas en la diagonal. Se puede notar también que la suma de los coeficientes para los efectos de la habilidad combinatoria general es igual a dos veces al coeficiente para μ , excepto en la ecuación μ debido a la inclusión de puros en el análisis.

Después de imponer las restricciones necesarias se procede a invertir la matriz reducida, y calcular las constantes, errores estándares, sumas de cuadrados, etc.

Las constantes que resultan son las siguientes:

$$\begin{array}{lll} \hat{\mu} = 60.322 & \hat{\epsilon}_{22} = -1.293 & \hat{c}_{13} = -0.172 \\ \hat{a}_1 = -1.057 & \hat{\epsilon}_{23} = 2.753 & \hat{r}_{12} = 0.766 \\ \hat{p}_{111} = -3.310 & \hat{m}_{21} = -1.641 & \hat{r}_{13} = 0.982 \\ \hat{p}_{122} = 7.081 & \hat{m}_{22} = 1.426 & \hat{r}_{23} = 1.229 \\ \hat{p}_{133} = -5.473 & \hat{m}_{23} = -5.784 & \\ \hat{\epsilon}_{21} = 2.941 & \hat{c}_{12} = -1.285 & \end{array}$$

Sumas de cuadrados

La reducción total en SC = $R(\mu, a, p, g, m, c, r) = \underline{779,783.44}$

SC error = $779,783.44 - 779,783.44 = \underline{13,328.45}$ y así los demás

SC se calculan por el procedimiento conocido.

Análisis de variancia

FV	GL	SC	CM
Heterosis (A)	1	192.93	192.93
Puros (P)	3	1878.50	626.17
Habilidad combinatoria general (G)	3	1041.44	347.15
Habilidad materna (M)	3	1725.26	575.09
Habilidad combinatoria específica (C)	2	180.73	90.36
Efectos ligados al sexo (R)	3	267.73	89.24
Error	192	13328.45	69.42

En vista de que el CM para efectos ligados al sexo no es significativo, se podría eliminar del modelo y estimar nuevamente los efectos que quedan. Después de la nueva estimación, si la habilidad combinatoria específica no es significativa, también se puede eliminar y reestimar los efectos restantes.

Los errores estandares se calculan por el procedimiento usual.

$$s_{\hat{\mu}} = \sqrt{.00578994} \sqrt{69.42} = (.0761)(8.33) = .63$$

$$s_{(\hat{\mu} + \hat{a}_1)} = \sqrt{.00578994 + .00578994 + 2(.00204077)} (8.33) = (.125)(8.33) = 1.04$$

$$s_{(\hat{\mu} + \hat{a}_1 + \hat{p}_{111})} = \sqrt{.00578994 + .00578994 + .03838870 + 2(.00204077) + 2(-.00214889) + 2(-.00214889)} (8.33) = (.213)(8.33) = 1.77$$

El error estandar para un grupo de puras cruza se puede obtener facilmente considerando la diagonal de la matriz y el número en la sub-clase. Cada una de estas sub-clases contiene el efecto de $\hat{\mu}$ más el de la sub-clase, así el elemento inverso de la diagonal correspondiente a una sub-clase de s_{ij} . El error estandar de esto sería $S_{\hat{\mu}} + s_{ij}$. Por ejemplo, para calcular el error estandar de la crua entre un toro de la línea 1 y una vaca de la línea 2, el número en la sub-clase es 12 y su inversa $\frac{1}{12}$, de ahí que el error estandar para esta crua sería.

$$s_{12} = \sqrt{\frac{1}{12}} \sqrt{69.42} = \sqrt{5.7850} = 2.41$$

Los datos que se han usado en este análisis se podría analizar también, usando el modelo siguiente:

$$Y_{ijk} = \mu + s_i + d_j + (sd)_{ij} + e_{ijk}$$

Si bien es cierto que el análisis que sigue este modelo no da una información detallada como la que vimos, pero da buena idea del performance de las líneas padres y madres. Si se dispone de una matriz de transformación esto ahorra considerable trabajo, para el caso de 4 líneas y bajo este modelo que se presenta a continuación.

Matriz de transformación para el modelo

$$Y_{ijk} = \mu + s_i + d_j + (sd)_{ij} + e_{ijk}$$

	s ₁₁	s ₁₂	s ₁₃	s ₁₄	s ₂₁	s ₂₂	s ₂₃	s ₂₄	s ₃₁	s ₃₂	s ₃₃	s ₃₄	s ₄₁	s ₄₂	s ₄₃	s ₄₄
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
s ₁	3	3	3	3	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
s ₂	-1	-1	-1	-1	3	3	3	3	-1	-1	-1	-1	-1	-1	-1	-1
s ₃	-1	-1	-1	-1	-1	-1	-1	1	3	3	3	3	-1	-1	-1	-1
d ₁	3	-1	-1	-1	3	-1	-1	-1	3	-1	-1	-1	3	-1	-1	-1
d ₂	-1	3	-1	-1	-1	3	-1	-1	-1	3	-1	-1	-1	3	-1	-1
d ₃	-1	-1	3	-1	-1	-1	3	-1	-1	-1	3	-1	-1	-1	3	-1
sd ₁₁ $\frac{1}{16}$	9	-3	-3	-3	-3	1	1	1	-3	1	1	1	-3	1	1	1
sd ₁₂	-3	9	-3	-3	1	-3	1	1	1	-3	1	1	1	-3	1	1
sd ₁₃	-3	-3	9	-3	1	1	-3	1	1	1	-3	1	1	1	-3	1
sd ₂₁	-3	1	1	1	9	-3	-3	-3	-3	1	1	1	-3	1	1	1
sd ₂₂	1	-3	1	1	-3	9	-3	-3	1	-3	1	1	1	-3	1	1
sd ₂₃	1	1	-3	1	-3	-3	9	-3	1	1	-3	1	1	1	-3	1
sd ₃₁	-3	1	1	1	-3	1	1	1	9	-3	-3	-3	-3	1	1	1
sd ₃₂	1	-3	1	1	1	-3	1	1	-3	9	-3	-3	1	-3	1	1
sd ₃₃	1	1	-3	1	1	1	-3	1	-3	-3	9	-3	1	1	-3	1

Analizando bajo este modelo nuestros datos, obtenemos las siguientes constantes.

$$\begin{aligned} \hat{\mu} &= 60.850 & \hat{d}_3 &= -4.330 & (\hat{sd})_{23} &= 2.452 \\ \hat{s}_1 &= 1.053 & (\hat{sd})_{11} &= -5.361 & (\hat{sd})_{31} &= 2.559 \\ \hat{s}_2 &= .767 & (\hat{sd})_{12} &= -.163 & (\hat{sd})_{32} &= -.005 \end{aligned}$$

$$\begin{aligned}\hat{s}_3 &= 1.454 & (\hat{sd})_{13} &= 4.526 & (\hat{sd})_{33} &= -4.183 \\ \hat{d}_1 &= - .588 & (\hat{sd})_{21} &= 1.696 \\ \hat{d}_2 &= 2.193 & (\hat{sd})_{22} &= 2.536\end{aligned}$$

La obtención de las sumas de cuadrados sigue el procedimiento usual.

Análisis de variancia

FV	GL	SC	CM
Padres (P)	3	658.49	219.50
Madres (M)	3	1409.62	469.87
Padres x Madres	9	1812.72	201.41
Error	191	13327.55	69.78

Este modelo de análisis sólo da información del valor de la línea madre y padre y ninguna información acerca del tipo de genes que actúa (aditivo y no aditivo).

Sin embargo bajo el modelo anterior se tiene una información mucho más completa de los tipos y efectos de genes. El uso de una simple matriz de transformación hace posible estimar el efecto de la heterosis en cada una de las cruzas y también probar la significancia entre las heterosis estimadas. La matriz de transformación y la transformada se demuestran a continuación.

Matriz de transformación usada para obtener los estimadores de las constantes de efectos de heterosis para las cruzas por separado y la matriz inversa transformada de los números en las subclases

	s ₁₁	s ₁₂	s ₁₃	s ₁₄	s ₂₁	s ₂₂	s ₂₃	s ₂₄	s ₃₁	s ₃₂	s ₃₃	s ₃₄	s ₄₁	s ₄₂	s ₄₃	s ₄₄
h ₁₂ :	-.5	.5	0	0	.5	-.5	0	0	0	0	0	0	0	0	0	0
h ₁₃ :	-.5	0	.5	0	0	0	0	0	.5	0	-.5	0	0	0	0	0
h ₁₄ :	-.5	0	0	.5	0	0	0	0	0	0	0	0	.5	0	0	-.5
h ₂₃ :	0	0	0	0	0	-.5	.5	0	0	.5	-.5	0	0	0	0	0
h ₂₄ :	0	0	0	0	0	-.5	0	.5	0	0	0	0	0	.5	0	-.5
h ₃₄ :	0	0	0	0	0	0	0	0	0	0	-.5	.5	0	0	.5	-.5

	h ₁₂	h ₁₃	h ₁₄	h ₂₃	h ₂₄	h ₃₄
h ₁₂	<u>.06264569</u>	.01136364	.01136364	.00961538	.00961538	.00000000
h ₁₃	.01136364	<u>.07803030</u>	.01136364	.02083333	.00000000	.02083333
h ₁₄	.01136364	.01136364	<u>.07088744</u>	.00000000	.02083333	.02083333
h ₂₃	.00961538	.02083333	.00000000	<u>.07628205</u>	.00961538	.02083333
h ₂₄	.00961538	.00000000	.02083333	.00961538	<u>.08669872</u>	.02083333
h ₃₄	.00000000	.02083333	.02083333	.02083333	.02083333	<u>.08333333</u>

Se puede ver de la matriz de transformación, que las constantes estimadas para heterosis en cada craza, es un contraste entre el promedio de los puros y el promedio de las cruzas recíprocas. Las constantes estimadas y las pruebas de significancia se demuestran a continuación.

Cruzas	Constantes estimados	Pruebas de significancia entre cruzas
13	8.314*	
14	5.095*	
23	2.047	
34	1.175	
12	0.483	
24	-4.431	

Las pruebas de significancia de la heterosis estimadas para cada craza se hicieron por el análisis de variancia, con solo un GL para cada constante estimada. La SC para cada contraste utiliza la constante estimada y el elemento inverso de la matriz transformada, correspondiente por hilera y columna a la craza considerada. Por ejemplo la SC para el estimador de la heterosis que resulta de la craza línea 1 x línea 3 es:

$$SC_{h_{13}} = \frac{(8.314)^2}{0.078030030} = 885.84$$

El análisis de variancia se demuestra a continuación.

	FV	GL	SC	CM
Heterosis				
Línea 1 y línea 2		1	3.72	3.72
Línea 1 y línea 3		1	885.93	885.93 **
Línea 1 y línea 4		1	366.16	366.16 *
Línea 2 y línea 3		1	54.93	54.93
Línea 2 y línea 4		1	226.51	226.51
Línea 3 y línea 4		1	16.57	16.57
Error		191	13327.55	69.78

Las pruebas de significancia entre las constantes estimadas, fueron hechas para cada cruz a por separado, por la prueba de rango múltiple de Duncan.

ANALISIS SIMULTANEO DE DATOS DE CRUZAS Y PUROS

En el análisis anterior hemos estimado los efectos del puro independientemente del efecto de la habilidad combinatoria general, sin embargo podemos estimar simultáneamente considerando a los efectos de la habilidad combinatoria general y la habilidad materna incluidos en las ecuaciones de puros.

El modelo matemático que describe el tercer método de análisis es el siguiente.

Modelo matemático (III)

$Y_{hijk} = \mu + a_h + P_{iiii} + S_i + S_j + C_{2ij} + r_{2ij} + e_{hijk}$. Los símbolos tienen el mismo significado para el modelo II.

Los sub-índices indican una consideración simultánea de la habilidad combinatoria general, materna y los efectos de puros. Las ecuaciones de mínimos cuadrados para este modelo se presentan a continuación.

Las restricciones impuestas a los estimadores de las constantes son las mismas para el modelo II. Los demás procesos de cálculos son similares a los del método I y II, por eso no presentamos ningún ejemplo numérico.

Ecuaciones de mínimos cuadrados para el modelo III

$\hat{\mu}$	$\hat{\sigma}_h$	\hat{P}_{111}	$\hat{\sigma}_{.1}$	$\hat{m}_{.j}$	\hat{c}_{21j}	\hat{r}_{21j}	RHM
μ	$n_{h..}$	n_{111}	$n_{.1.} + n_{..1}$	$n_{..j}$	$n_{21j} + n_{2j1}$	n_{21j}	$Y_{...}$
σ_h	$n_{h..}^0$	n_{111}	$n_{h1.} + n_{h.1}$	$n_{h..j}$	$n_{21j} + n_{2j1}$	n_{21j}	$Y_{h..}$
P_{111}	n_{111}	n_{111}^0	n_{111}	n_{111}	—	—	Y_{111}
$\sigma_{.1}$	$n_{.1.} + n_{..1}$	n_{111}	$n_{.1.} + n_{..1}$	$n_{..j}$	$n_{21j} + n_{2j1}$	n_{21j}	$Y_{.1.} + Y_{.1}$
$m_{.j}$	$n_{h..j}$	n_{111}	$n_{..j}$	$n_{..j}^0$	$n_{2.j}$	n_{21j}	$Y_{..j}$
c_{21j}	$n_{21j} + n_{2j1}$	—	$n_{21j} + n_{2j1}$	$n_{2.j}$	$n_{21j} + n_{2j1}^0$	n_{21j}	$Y_{21j} + Y_{2j1}$
r_{21j}	n_{21j}	—	n_{21j}	n_{21j}	n_{21j}	n_{21j}^0	Y_{21j}

ESTIMACION DEL INDICE DE HERENCIA O HEREDABILIDAD

El concepto de la heredabilidad está asociado con la importancia relativa de la herencia y el ambiente como factores influyentes en la variación de los caracteres.

La heredabilidad es la razón entre la variancia genética aditiva y la fenotípica. La expresión matemática es la que se da a continuación.

$$h^2 = \frac{V_A}{V_p}$$

donde

h^2 = heredabilidad o índice de herencia.

V_A = variancia genética aditiva

V_p = variancia fenotípica

Un significado equivalente de la heredabilidad está dado por la regresión del valor del breeding sobre el valor fenotípico. La expresión matemática es

$$h^2 = b_{AP}$$

donde b_{AP} representa el coeficiente de regresión del valor del breeding sobre el valor fenotípico.

Realmente estas dos fórmulas son equivalentes, puesto que si nosotros dividiéramos el valor del fenotipo en valor del breeding mas un recíduo (R) que contiene el efecto medio ambiental, más la dominancia, más las desviaciones de la interacción, entonces tendríamos $P = A + R$, donde A y R no están correlacionados, de ahí que $Cov = V_A$ y $b_{AP} = \frac{V_A}{V_p}$.

También se puede notar que la correlación entre el valor del breeding y el valor fenotipo (r_{Ap}) es igual a la raíz cuadrada de la heredabilidad.

$$r_{Ap} = b_{Ap} \frac{\sigma_P}{\sigma_A} = h^2 \times \frac{1}{h} = h$$

La heredabilidad como regresión del valor del breeding sobre el valor fenotípico nos muestra que la mejor estimación del valor del breeding de un individuo es el producto de su valor fenotípico y la heredabilidad

$$A \text{ esperado} = h^2 P$$

Esto nos indica que la heredabilidad expresa la realidad del valor fenotípico como una guía del valor del breeding.

El valor de la heredabilidad depende de la magnitud de todos los componentes de variancia. Todos los componentes genéticos están influenciados por la frecuencia de genes y por eso difieren de una población a otra, de acuerdo a la historia pasada de la población. Las poblaciones pequeñas mantenidas por largo tiempo "cerrado" (hato cerrado) demuestra poca heredabilidad. La variancia medio ambiental depende de las condiciones del manejo, cuanto más variable es este componente se reduce la heredabilidad y a la inversa, cuanto más uniforme es esta.

Métodos para estimar la heredabilidad

Los métodos para estimar la heredabilidad dependen de los datos de que se dispone y la precisión que uno desee en la estimación. Nosotros los dividiremos en general en los siguientes:

- 1) Regresión de los descendientes sobre los padres.
 - a) Regresión de los descendientes sobre uno de los padres.
 - b) Regresión de los descendientes sobre el promedio de los padres.
- 2) Correlación intraclase (Análisis de componentes de variancia)
 - a) Análisis de medios hermanos.
 - b) Análisis de hermanos completos

Naturalmente como los materiales disponibles son diferentes, el valor de la variancia aditiva será diferente en cada caso. A continuación resumiremos los valores.

Parientes	Covariancia	Regresión(b)	Correlación(t)
Descendientes y uno de los padres	$\frac{1}{2} V_A$	$b = \frac{1}{2} h^2$	-
Descendientes y promedios de los padres	$\frac{1}{2} V_A$	$b = h^2$	-
Medios hermanos	$\frac{1}{4} V_A$	-	$t = \frac{1}{4} h^2$
Hermanos completos	$\frac{1}{2} V_A + \frac{1}{4} V_D + V_{Ec}$	-	$t > \frac{1}{2} h^2$

Generalmente se dice que la correlación entre medios hermanos y la regresión de hijos sobre el padre dan mejores estimaciones de la heredabilidad, ya que la regresión de los hijos sobre la madre puede estar viciada por el efecto materno y es menos real todavía la correlación entre hermanos completos, ya que el componente debido al medio ambiente común a menudo está presente y es difícil corregir por el diseño, y la covariancia entre hermanos completos se ve aumentada por

la variancia dominante (sobre-estimación).

Nosotros suponemos que los datos con los cuales trabajamos provienen de apareamiento al azar en las sucesivas generaciones, sin embargo a menudo los grupos de padres se escogen consecuentemente, la variancia fenotípica entre los padres escogidos es menor que la de la población original y menor también a la de los descendientes. Sin embargo la regresión de hijos sobre padres no está afectado por la selección de los padres, debido a que la covariancia se reduce en la misma magnitud que la variancia de los padres, así que la pendiente de la línea de regresión no se altera, de ahí que la regresión de hijos sobre unos de los padres es una medida válida de $\frac{1}{2} h^2$ y la regresión de descendientes sobre el promedio de los padres es una estimación válida de h^2 , pero la covariancia no es una medida válida de V_A .

Si el apareamiento se hace por semejanza de los fenotipos entonces existe una correlación entre el valor del fenotipo de las parejas apareados; esto trae como consecuencia un incremento del valor de la variancia del promedio de los padres y consecuentemente la covariancia entre los hermanos completos. Sin embargo la regresión de descendientes sobre el promedio de los padres es muy poco afectado y se puede tomar como una medida válida de h^2 .

Método de regresión de descendientes sobre el promedio de los padres

Los datos obtenidos de los padres (x) se promedian siguiendo el procedimiento usual ($\frac{x^{\sigma} + x^{\varphi}}{2}$). Luego se calcula el promedio de los valores de sus descendientes. Estos valores promedios deben ser

ponderados de acuerdo al número de descendientes por familia.

Los factores de ponderación apropiados a usar los discutiremos a continuación, ya que el cálculo del coeficiente de regresión, es cosa de rutina.

El factor de ponderación

El cálculo de regresión con número variable de descendientes por padres tiene el problema de cómo calcular el factor de ponderación debido a la desigualdad de los números de descendientes. Dos son los procedimientos corrientemente seguidos: 1) repetir el record del progenitor con cada uno de los records de los descendientes; 2) promediar los datos de todos los descendientes de un progenitor y regresar tal promedio sobre el record de su progenitor. La primera práctica sería válida si la correlación entre los descendientes de un padre fuera cero, mientras la otra sería válida si la correlación entre los miembros de un grupo de progenie fuera uno. Obviamente la situación es intermedia al primero y al segundo, aunque más cerca al primero. Para encontrar qué factor de ponderación se debe aplicar para obtener una estimación no viciada de la regresión que haga mínima la variancia de la muestra. Para ello se siguen los siguientes pasos (según Kempthorne y Tandon).

1) Suponer que el valor de $T = \frac{\rho}{1-\rho}$ y que este sea un estimador de T (corrientemente se usa $T = 0.04$)

2) Estimar β por la fórmula $\hat{\beta} = \frac{\sum W_i (X_i - \bar{X}) Y_i}{\sum W_i (X_i - \bar{X})^2}$

donde

$$W_i = \frac{n_i}{1 + n_i T}$$

T es el estimador de T

$$\bar{X} = \frac{\sum W_i X_i}{\sum W_i}$$

- 3) Estimar ρ_1 usando los cuadrados medios dentro y entre grupos de X . El cuadrado medio esperado dentro de grupos es

$$\sigma_p^2 (1 - \rho_1) \text{ y el cuadrado medio entre grupos es } \sigma_p^2 (1 - \rho_1) + \frac{1}{K-1} \left(\sum n_i - \frac{\sum n_i^2}{\sum n_i} \right) \sigma_p^2 \rho_1$$

donde K es el número de grupos de X y así se puede estimar ρ_1 igualando la ecuación de cuadrados medios observados a esperados.

- 4) Usando los estimadores de $\hat{\rho}_1$ y $\hat{\beta}$ obtenemos T

$$T = \frac{\hat{\rho}_1 - \hat{\beta}^2}{1 - \hat{\rho}_1}$$

- 5) La variancia estimada de β es el cuadrado medio entre Y 's dentro X grupos.

$$\sum n_i \frac{1 + n_i \hat{T}}{(1 + n_i T)^2} \frac{(X_i - \bar{X})^2}{\left[\sum_i w_i (X_i - \bar{X})^2 \right]^2}$$

Sin embargo Reeve (1955) objeta en parte este procedimiento y sugiere un procedimiento más simple, que es el siguiente.

Suponiendo que tenemos familias de n hermanos, es decir que las familias no producen igual número de descendientes, llamando P al valor promedio de los padres, o al valor de la progenie (individualmente) y M al promedio de la progenie para cada familia. Con estos valores se establecen las siguientes correlaciones.

- a) r_{op} = coeficiente de correlación entre promedio de padres y el valor de cada individuo descendiente.
- b) r_{MP} = coeficiente de correlación entre promedio de padres y promedio de sus descendientes.
- c) r_{oo} = correlación intraclase entre los hermanos.

Entonces se utiliza la fórmula ya dada para calcular el coeficiente de regresión b_{MP} ponderado por W_n , donde

$$W_n = \frac{n}{1 + nT}$$

$$T = \frac{r_{oo} - r_{op}^2}{1 - r_{oo}}$$

y la variancia del coeficiente de regresión b_{MP} es

$$V(b_{MP}) = \frac{60(1 - r_{oo})^2}{W_n \sum (P - \bar{P})^2}$$

Solo el denominador cambia con el tamaño n de la familia, así es que la ponderación de un coeficiente de regresión basado en el tamaño n de la familia es proporcional a

$$W_n = W_n \sum (P - \bar{P})^2$$

Para obtener el coeficiente de regresión combinado con una variancia mínima para la variable n , primero debemos escoger un estimador común del promedio de los padres \bar{P} y luego obtener el promedio ponderado de los coeficientes de regresión para cada valor de n , usando el factor de ponderación W_n . El estimador de \bar{P} es $\bar{P} =$

$\left(\frac{\sum W_i P_i}{\sum W_i} \right)$ de ahí que el coeficiente de regresión promedio es

$$\bar{B} = \frac{\sum W_n b_n}{\sum W_n} = \frac{W_i M_i (P_i - \bar{P})}{W_i (P_i - \bar{P})^2}$$

Las mismas fórmulas se aplican cuando se dispone de datos de medios hermanos correlacionados con su ascendiente común P, es decir $r_{op} = \beta$ y $r_{oo} = \rho_1$. Ahora podemos ver que la ecuación que estamos discutiendo es simplemente el resultado de escoger el promedio de los padres comunes \bar{P} y ponderar los coeficientes de regresión para cada valor de n en proporción a las recíprocas de sus variancias.

Hay dos casos a considerar:

- 1) Familias de medios hermanos correlacionados con su padre común.

$$r_{op} = \frac{1}{2} h^2; r_{oo} = \frac{1}{4} h^2; \beta = \frac{1}{2} h^2$$

- 2) Familias de hermanos completos correlacionados con el promedio de sus padres.

$$r_{op} = \frac{h^2}{\sqrt{2}}; r_{oo} = \frac{1}{2} h^2, \beta = h^2$$

donde β es en cada caso la regresión de O sobre P, de esto se sigue que para el caso (1) tenemos:

$$T = \frac{h^2(1-h^2)}{4-h^2} = \frac{\beta(1-2\beta)}{2-\beta}$$

y para el caso (2)

$$T = \frac{h^2(1-h^2)}{2-h^2} = \frac{\beta(1-\beta)}{2-\beta}$$

Estos dos fórmulas para T caso (1) y caso (2) se pueden usar muy satisfactoriamente para calcular el factor de ponderación W_n , y el valor de β para el caso (1) es el coeficiente de regresión no ponderado de progenie sobre el padre y para el caso (2) de progenie sobre el promedio de los padres.

Método de correlación (Análisis de hermanos)

Corrientemente los datos que se obtienen con animales de fincas provienen del apareamiento de un grupo de toros, donde cada uno se aparee con varias vacas y se tiene un número m de descendientes provenientes de cada hembra, es decir que tenemos una población de medios hermanos y hermanos completos.

Con estos datos se hace un análisis de variancia, y se estiman los componentes de variación σ_s^2 (entre toros), σ_D^2 (entre vacas dentro de toros), y σ_w^2 (dentro de progenies), es decir que el análisis quedaría en la siguiente forma.

FV	GL	CM	Componentes estimados
Entre toros	$s-1$	CM_S	$\sigma_w^2 + K_1 \sigma_D^2 + K_2 \sigma_s^2$
Entre vacas dentro de toros	$s(d-1)$	CM_D	$\sigma_w^2 + K_1 \sigma_D^2$
Dentro de progenies	$sd(K-1)$	CM_w	σ_w^2

s = número de toros

d = número de vacas por toro

K = número de descendientes por vaca

Si el número de d y K son diferentes por toro y descendientes por vaca, se calcula el promedio de las mismas por las fórmulas indirectas que vimos en secciones anteriores.

Los componentes estimados nos indican lo siguiente: el componente entre toros nos mide $\frac{1}{4}$ de la variancia aditiva ($\sigma_s^2 = \frac{1}{4} V_A$) y la variancia fenotípica $V_p = \sigma_T^2 = \sigma_s^2 + \sigma_D^2 + \sigma_w^2$.

Entonces el coeficiente de correlación intraclase entre medios hermanos (t); $t_{HS} = \frac{\sigma_s^2}{\sigma_T^2}$

Regresión intra-toro de los descendientes sobre la madre

Los datos disponibles para este caso se supone que provienen del apareamiento de cada padre con varias madres, la regresión de descendientes sobre el promedio de los padres es inapropiada, ya que usualmente hay pocos padres machos. En el presente caso la heredabilidad se puede estimar satisfactoriamente de la regresión promedia de descendientes sobre la madre, calculado separadamente para cada serie de vacas apareadas con un toro, luego combinar los varios coeficientes calculados, y estimar su promedio ponderado. La regresión intra-toro de descendientes sobre la madre estima la mitad de la heredabilidad. La validez de esta estimación depende de la influencia materna.

Si existe heterogeneidad de la variancia dentro de machos y dentro de hembras, el coeficiente de regresión debe ser ajustada, multiplicando b por la razón de la desviación estandar para hembra sobre la desviación estandar para macho.

Por lo demás no requiere mayores explicaciones ya que lo hemos tratado con bastante detalle en el caso anterior y los mismos conceptos se pueden aplicar a este caso.

CONSANGUINIDAD (INBREEDING)

La consanguinidad es otro método de cría del cual puede valerse el hombre para modificar la herencia. La maximización de este método es la autofecundación, muy empleado en plantas, pero imposible de utilizar en animales.

La consanguinidad llevada a un extremo produce un estado de máxima homocigosis, pero la proporción de genes dominantes y recesivos permanece invariable mientras no se efectúe selección a favor o en contra de alguno de ellos.

Influencia de distintos métodos de reproducción sobre el coeficiente de consanguinidad.

Generación	Autofec.	Hermanos completos	1 solo macho	2 machos	Entre primos
1	50	25	12,5	6,6	6,2
2	75	37.5	21,8	12,8	9,4
3	87.5	50.0	30,4	18,5	12,2
4	93.8	59.4	38,0	23,9	14,6
5	96.9	67.2	44,8	28,9	16,8

En cría animal existen dos tipos de consanguinidad:

- a) Estrecha o simple
- b) Colateral o familiar

Puede observarse en el cuadro anterior que el cruzamiento entre hermanos completos es el método más estrecho que se aplica en

animales, y lleva a una homocigosis muy inferior a la obtenida por autofecundación. Para obtener la misma homocigosis que es lograda a través de 6 generaciones de autofecundación se necesitan aproximadamente 17 generaciones de consanguinidad entre hermanos completos.

En el estudio de consanguinidad es necesario distinguir entre:

- a) Parentesco o coeficiente de parentesco
- b) Consanguinidad

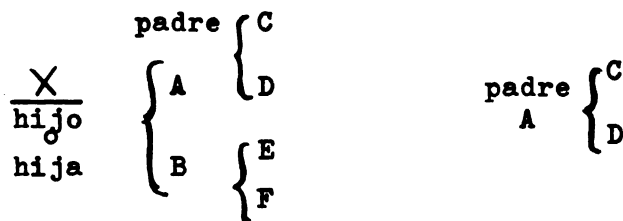
Parentesco

Se dice que dos animales son parientes cuando poseen mayor porcentaje de genes comunes que el que tendrían dos animales promedios de la raza. También se denominan parientes a aquellos individuos que tienen un antecesor común en la 4ª a 6ª generaciones anteriores.

Parentesco directo existe cuando hay una relación directa entre los animales.

Ejemplo:

Hijo o hija con respecto a su padre o madre

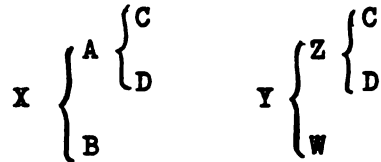


Estos animales están emparentados porque poseen genes en común que han heredado de sus antecesores comunes C y D. El animal A los recibió directamente y X a través de este.

Parentesco colateral es aquel que existe entre dos individuos que tienen uno o más antecesores comunes, pero que no caen dentro del primer grupo.

Ejemplo:

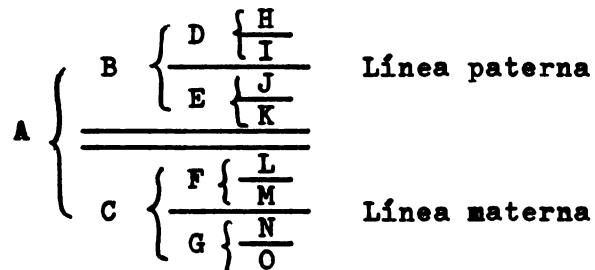
Primos hermanos dobles



Los animales X e Y tienen genes comunes que han heredado de sus abuelos paternos D y C y los han recibido a través del padre.

Nota: En todo pedigrée los antecesores paternos se escriben arriba y los maternos abajo.

Ejemplo



Cálculo de coeficiente de parentesco

$$R_{xy} = \sum \frac{1}{2}^{n+n'}$$

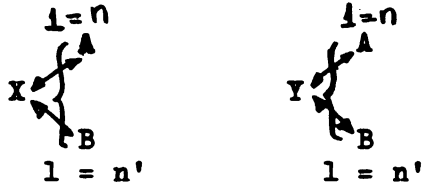
$\frac{1}{2}$ indica el proceso de partición propio de la herencia en cada generación.

n = indica el número de generaciones que existe desde el antecesor común a uno de los animales considerados.

n' = indica el número de generaciones que existe desde el mismo antecesor común al otro animal considerado.

Ejemplo 1:

Hermanos completos



Para A: $n = 1$ } $R_x = (1/2)^{1+1} = 25\%$
 $n' = 1$ }

Para B: $n = 1$ } $R_y = (1/2)^{1+1} = 25\%$
 $n' = 1$ }

50% (coeficiente de parentesco)

Ejemplo 2:

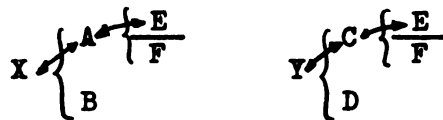
Medios hermanos



$R_{xy} = (1/2)^{1+1} = 25\%$ (coeficiente de parentesco)

Ejemplo 3:

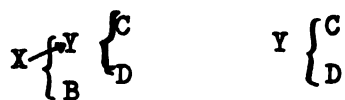
Primos hermanos simples



$R_{xy} = (1/2)^{2+2} = 6,25\%$ (coeficiente de parentesco)

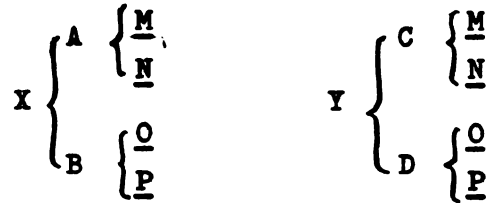
Ejemplo 4:

Padre e hijo



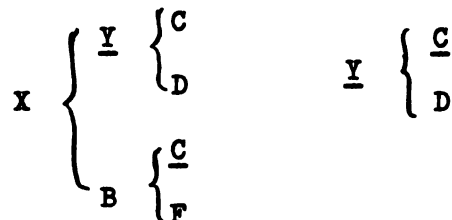
$$R_{xy} = (1/2)^{1+0} = 50\% \text{ (coeficiente de parentesco)}$$

Ejemplo 5:



$$\begin{array}{l} R_{xy}: \text{ Para M} = (1/2)^{2+2} = 6,25\% \\ \text{ Para N} = (1/2)^{2+2} = 6,25\% \\ \text{ Para O} = (1/2)^{2+2} = 6,25\% \\ \text{ Para P} = (1/2)^{2+2} = 6,25\% \\ \hline 25,00\% \text{ (coeficiente de parentesco)} \end{array}$$

Ejemplo 6:



En este caso Y es padre de X (parentesco directo) C es abuelo materno de X y padre de Y (parentesco colateral)

Parentesco directo

$$R_{xy} = (1/2)^{1+0} = 50\%$$

Parentesco colateral

$$R_{xy} = (1/2)^{2+1} = \frac{12,25\%}{62,25\%} \text{ (parentesco total)}$$

Coefficiente de consanguinidad

La consanguinidad deriva del hecho de aparear dos individuos

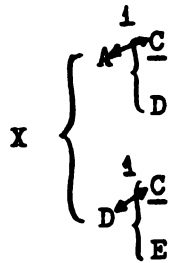
que son parientes, por lo tanto la cría estará alejada en una generación del antecesor común, y tendrá menos genes en común con aquellos que los que tenían sus padres.

Esta es la razón por la cual la fórmula para calcular consanguinidad se le agrega 1 al exponente.

$$F_x = \sum (1/2)^{n+n'+1}$$

El coeficiente de consanguinidad mide el aumento de homocigosis o disminución de heterocigosis con respecto a los primeros animales considerados en el pedegree.

Ejemplo 1:



El antecesor común en la línea paterna y materna es c, por tanto

$$F_x = (1/2)^{1+1+1} = 12,25\%$$

n = indica el número de generaciones desde el antecesor común en la línea paterna al padre del animal considerado.

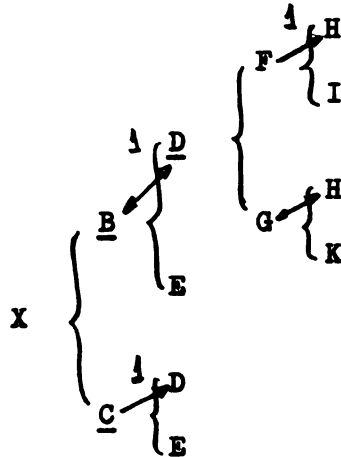
n' = indica el número de generaciones desde el mismo antecesor en la línea materna a la madre del animal considerado.

Cuando existe ascendencia consanguínea la fórmula toma una nueva forma, ya que los descendientes de animales consanguíneos tienen menores probabilidades de recibir genes diferentes de sus antecesores.

$$F_x = \sum (1/2)^{n+n'+1} (1 + F_A)$$

Fa representa el coeficiente de consanguinidad del antecesor común (tanto por uno).

Ejemplo 2:



En este caso el animal X es consanguíneo ya que resulta del cruzamiento de 2 hermanos completos B y C, pero D es consanguíneo a su vez porque proviene del apareamiento de 2 medios hermanos. Por tanto X tendrá mayor número de genes al estado homocigota.

$$F_D = (1/2)^{1+1+1} = 0.125$$

$$F_X \text{ a través de D} = (1/2)^{1+1+1} = 12,5$$

$$F_X \text{ a través de E} = (1/2)^{1+1+1} = \frac{12,5}{25,0\%}$$

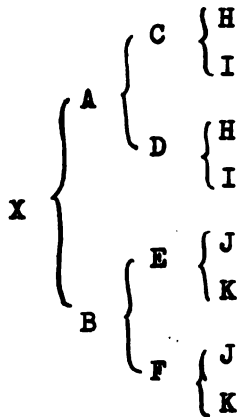
Consanguinidad total

$$F_X = 25 (1 + 0,125)$$

$$F_X = 25 \times 1,125 = 28,125\%$$

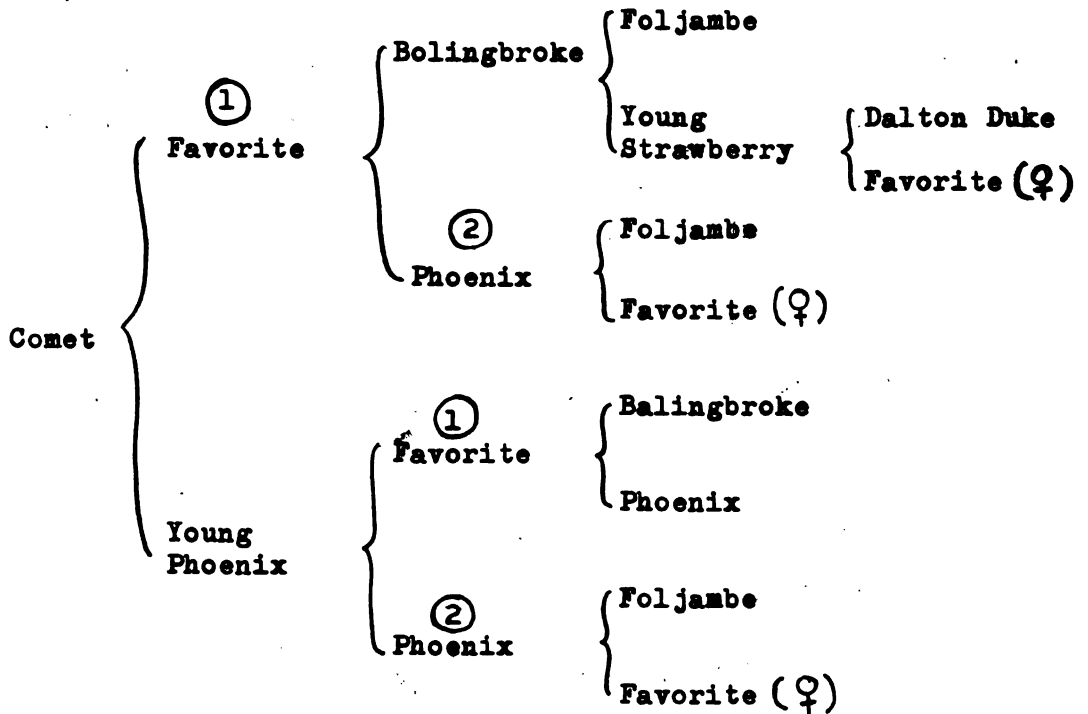
El individuo x tendrá 28,125% más homocigosis que el promedio de los individuos no consanguíneos que forman esa población.

Ejemplo 3:



La consanguinidad de X es 0 ya que no tiene antecesores comunes representa dos en la línea paterna y materna, a pesar de que A y B lo son y en alto grado.

Ejemplo 4:



Primer paso. Determinar los individuos representados en la línea materna y paterna. En este caso son:

- a) Favorite macho
- b) Phoenix
- c) Foljambe
- d) Favorite hembra

Segundo paso. Determinar cuales de los individuos que están produciendo la consanguinidad de comet son consanguíneos a su vez. En este caso solo es:

- a) Favorite macho (a través de Foljambe y Favorite Q Young Phoenix también es consanguíneo, pero no se considera en el cálculo ya que solo está representado en la línea materna.

Tercer paso. Calcular la consanguinidad.

- a) Consanguinidad de Favorite

$$F_{\text{Favorite}} \text{ a través de Foljambe} = (1/2)^{1+1+1} = .1250$$

$$F_{\text{Favorite}} \text{ a través de Favorite} = (1/2)^{2+1+1} = .0625$$

.1875

- b) Consanguinidad de comet a través de Favorite ♂

Favorite es padre y abuelo materno

$$F_{\text{comet}} \text{ a través de Favorite} = (1/2)^{0+1+1} = 25\%$$

Como Favorite es consanguíneo a su vez esta cifra deberá ser corregida.

$$F_{\text{comet}} \text{ a través de Favorite ♂} = 25(1+.1875) = \underline{29.69\%}$$

c) Consanguinidad de comet a través de Phoenix

Phoenix es abuela materna y paterna

$$F_{\text{comet}} \text{ a través de Phoenix} = (1/2)^{1+1+1} = \underline{12,5\%}$$

d) Consanguinidad de comet a través de Foljambe

En este caso se considera a Foljambe como abuelo de Favorite a través de Bolingbroke y como abuelo de Young Phoenix a través de Phoenix. Se toman estos individuos ya que no deben considerarse los padres o antecesores comunes de un individuo que aparecerá en ambas ramas de pedegree.

$$F_{\text{comet}} \text{ a través de Foljambe} = (1/2)^{2+2+1} = \underline{3,12\%}$$

e) Consanguinidad de comet a través de Favorite Q

La situación aquí es semejante a la anterior, considerándosele entonces a través de Young Strawberry, madre de Bolingbroke, en la rama paterna y como madre de Phoenix en la rama materna.

$$F_{\text{comet}} \text{ a través de Favorite} = (1/2)^{3+2+1} = \underline{1,56\%}$$

Cuarto paso. Cálculo de la consanguinidad total de comet

$$\text{Coeficiente consanguinidad total} = 29,69 + 12,50 + 3,12 + 1,56 = 46,87\%$$

Coeficiente de parentesco con antecesores consanguíneos

En este caso existirán dos tipos de correcciones:

1. Cuando se tiene un antecesor común consanguíneo existirán mayores posibilidades de que este transmita genes idénticos a su descendencia, porque la consanguinidad lo ha hecho más homocigoto. Por ello el numerador de la fórmula del

coeficiente de parentesco debe ir multiplicado por el factor $(1+F_A)$.

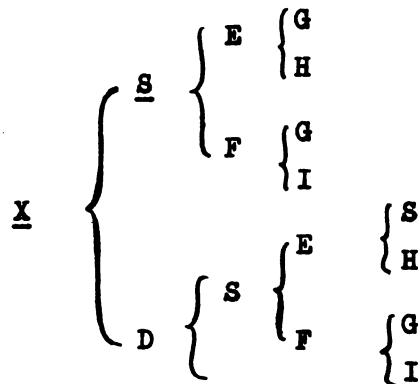
2. Para corregir el hecho que la consanguinidad tiende a formar familias o strains diferentes, el denominador de la fórmula en el coeficiente de parentesco debe incluir la consanguinidad que posee cada uno de los individuos a los cuales se les está determinando el parentesco. Esto proviene del hecho de que la consanguinidad aumenta la variación entre las diferentes líneas, aunque la disminuye dentro de la línea considerada en especial. Un individuo puede ser homocigota para igual locus, pero no para el mismo gene.

La fórmula será entonces:

$$R_{xy} = \frac{(1/2)^{n+n'} (1+F_A)}{(1+F_x) (1+F_y)}$$

Ejemplo 1

Apareamiento de un padre consanguíneo con su hija



Primer paso. Calcular la consanguinidad de S

$$F_S = (1/2)^{1+1+1} = .125$$

Segundo paso. Calcular la consanguinidad de X

$$F_X = (1/2)^{0+1+1} = 25$$

Como el antecesor común de X es consanguíneo a su vez deberá corregirse.

$$F_X = (1/2)^{n+n'+1} (1+F_A) = 25 \times 1,125 = \underline{28,125\%}$$

Tercer paso. Cálculo del coeficiente de parentesco R_{SX} .

a) Parentesco directo (S como padre de X)

$$(1/2)^{1+0} \times 1,125 = \underline{56,25}$$

b) Parentesco colateral (S como abuelo de X)

$$(1/2)^{2+0} \times 1,125 = \underline{28,125}$$

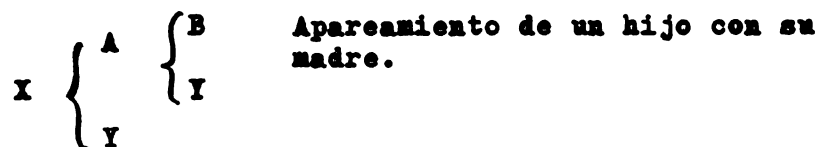
Se corrige ya que S es consanguíneo.

La suma de los dos valores obtenidos anteriormente 84,375% representa por tanto el parentesco corriente corregido por la consanguinidad que presenta el antecesor común.

Para calcular el coeficiente de parentesco real entre R y S deberá corregirse por las consanguinidades, presentadas individualmente por el antecesor y el descendiente.

$$R_{XS} = \frac{84,375}{\sqrt{(1+0,28125)(1+0,125)}} = \underline{70,28\%}$$

Ejemplo 2:



En este caso no habrá corrección en el número donde la fórmula, porque los antecesores comunes no son consanguíneos.

$$\text{Parentesco X, Y a través de A} = (1/2)^{2+0} = 25\%$$

$$\text{Parentesco X, Y a través de B} = (1/2)^{1+0} = 50\%$$

$$75\%$$

Consanguinidad de X

$$F_X = (1/2)^{1+0+1} = 0.25$$

consanguinidad a través de Y no existe por tanto:

$$R_{XY} = \frac{75}{\sqrt{1 + 0,25}} = \frac{75}{1,11} = \underline{67,5\%}$$

MODALIDAD DE LOS ENSAYOS CON ANIMALES

Experimente con animales menores

Dentro de este grupo incluimos, las aves, los conejos, ratas, etc. La modalidad de los experimentos con estos tipos de animales son bastante simples, ya que generalmente son pequeños, de un valor relativamente bajo que permite usar varios animales por unidad experimental, esto es una ventaja ya que tiende a homogenizar la información (datos). Por otra parte se puede usar bastante número de repeticiones por tratamientos, en un espacio relativamente pequeño y

con un costo relativamente bajo.

Se puede usar cualquier tipo de diseños, de acuerdo a los factores que se presume obran como fuentes de variación, sin embargo los más usados son: el irrestrictamente al azar y el bloques al azar.

Las unidades experimentales deben considerarse como un todo. No interesa, por ejemplo conocer el peso individual, sino por unidad de observación (parcela).

No es raro que al final del experimento se tengan unos animales muertos por enfermedades o accidentes, si el porcentaje de mortalidad no es alta digamos alrededor del 10%, se pueden considerar las medidas provenientes de las unidades experimentales, sin ningún ajuste, es decir que la diferencia en el número de animales se consideran como variación del azar.

Sin embargo, una mortalidad elevada puede ser debido al efecto de los mismos tratamientos, este aspecto debe ser verificado cuidadosamente analizando el número de animales vivos y muertos, casi siempre los análisis de este tipo, requieren una transformación de n_{ij} digamos $\sqrt{n_{ij}}$. Si la desproporción es muy grande se puede usar el método de constantes de ajustes para ponderar los Y_{ij} , también cabe hacer acá ajustes por sexo, o cualquier otro ajuste (ver última sección).

Experimentos con animales del tipo intermedio

En este grupo incluimos, suinos, ovinos, caprinos etc. En cuanto a la modalidad de los experimentos con estas clases de animales, tiene fundamento bastante similar al anterior aunque un poco más complicado. Según el objetivo del experimento se puede usar 1, 2, 3 o 4 animales

por observación (unidad experimental), aunque muchas veces es más importante usar unidades experimentales pequeñas con bastantes repeticiones, aunque esto es más costoso. Con este tipo de animales se puede aplicar cualquier diseño. Los ajustes de datos se hacen en la forma usual tomando en cuenta el tipo de ajuste, sea aditivo, multiplicativo, etc., según el factor que se quiere ajustar.

Experimento con animales mayores

En este grupo incluiremos los animales grandes equinos y bovinos. Especial interés presentan estos últimos grupos por su valor económico bien alto, con los cuales se han llevado a cabo miles de experimentos. Para mejor ilustración los dividiremos en dos grupos:

- a) bovinos de carne
- b) bovinos de leche

Los experimentos con ganados de carne o cotre tienen menos problemas que los de leche, por esta razón consideraremos a los bovinos de leche como representativo del grupo y lo comentaremos con ciertos detalles.

Los ensayos con ganados lecheros son los más complejos que existen, ya que las vacas lecheras son animales de valor bastante alto, generalmente se dispone de muy pocos números para el estudio y muy heterogéneo. La heterogeneidad del material experimental nos obliga a usar numerosas restricciones, que hacen necesariamente complicado los ensayos, a fin de aumentar la precisión de las estimaciones.

Es muy corriente el uso de diseño de bloques al azar con los ganados lecheros, donde los bloques están formados por el nivel de la

producción de cada vaca en un período pre experimental, digamos producción alta, intermedia y baja o lo mismo podría ser edad avanzada intermedia y jóvenes. Entonces dentro de cada bloque formado por n vacas se aplican todos los tratamientos. Se puede usar una vaca como unidad experimental.

Un ejemplo aclarará el concepto. Supongamos que queremos estudiar el efecto de 5 fórmulas de raciones balanceadas sobre la producción de leche. Disponemos de 20 animales, 5 de producción muy alta, 5 alta, 5 intermedia y 5 baja; en este caso podemos hacer los bloques con los niveles de producción. Esto lo podemos hacer en la siguiente forma.

Raciones	Repeticiones			
	I (muy alta)	II (alta)	III (mediana)	IV (baja)
r_1 :	1 (vaca)	6	11	16
r_2 :	2	7	12	17
r_3 :	3	8	13	18
r_4 :	4	9	14	19
r_5 :	5	10	15	20

Estas informaciones las analizaríamos con un bloque al azar o sea.

FV	GL
Repeticiones	$r - 1$
Raciones	$k - 1$
Error	$(k-1)(r-1)$
Total	$n - 1$

Experimento con vacas lecheras en períodos sucesivos

Un caso muy interesante de ensayos con vacas lecheras consiste en tomar cada animal como bloque, y la producción de leche en períodos sucesivos se considera como unidad de observación (parcela), la unidad de observación puede ser la producción semanal o tomada cada dos semanas, etc., según el criterio del investigador y el objetivo del trabajo. Los experimentos de esta naturaleza, en inglés recibe el nombre de "change-over designs" o "Swith-over designs", nosotros le podemos llamar ensayos con reversión.

Un defecto de este método reside en la posibilidad de efecto residual de un tratamiento sobre el subsiguiente, esto se puede corregir despreciando la producción del animal en los primeros días o primera semana después del cambio de tratamiento o recurrir a procedimientos matemáticos propuestos por Cochran y otros. Otro defecto que se anota en este tipo de experimento, es que cada animal se usa como bloque y su producción es alta en los dos o tres primeros meses luego queda estacionaria para finalmente decaer. Esta fuente de error se puede corregir utilizando cuadrados latinos, tomando los animales como columnas, y los períodos como hileras.

Ejemplo. Un ensayo con tres tratamientos (A, B y C) se puede diseñar en la siguiente forma.

Hileras	Columnas		
	Vaca A	Vaca 2	Vaca 2
Primer período	A	B	C
Segundo período	C	A	B
Tercer período	B	C	A

Entonces el análisis de variancia se haría en la siguiente forma.

FV	GL
Columnas (vacas)	c - 1
Hileras (períodos)	h - 1
Tratamientos	K - 1
Error	(k-1)(K-2)
Total	n - 1

Cada tratamiento cae tres veces en los tres períodos, para que esto sea eficiente es necesario que los animales tengan producción similar, una desigualdad en la producción de la vaca impide una buena compensación de los efectos de los períodos.

Como con pocos tratamientos, caso de cuadrado latino 3 x 3, el grado de libertad del error es muy pequeño. Se necesita usar varios cuadrados latinos, 3 o 4, y entre estos cuadrados latinos puede variar la producción de las vacas, pero dentro de cada cuadrado se necesita que sean bien uniformes.

El análisis con tres tratamientos y tres cuadrados latinos sería el siguiente.

FV	GL
Grupos de vacas (cuadrados latinos)	2
Vacas dentro de grupos	6
Períodos (hileras) dentro de grupos	6
Tratamientos	2
Interacción tratamiento x cuadrados latinos	4
Error	6
Total	26

Experimento de doble reversión o alternativo

En inglés, este tipo de delineamiento recibe el nombre de "switch-back". Este diseño se usa especialmente con vacas lecheras o casos análogos (gallinas ponedoras).

Cada animal o grupo de animal se utiliza en 3 períodos experimentales sucesivos. Cada uno de alrededor de tres semanas, para evitar los efectos residuales de un período sobre el siguiente y permitir que los animales se habitúen al nuevo tratamiento, se debe despreñar la producción de leche de la semana inicial de cada período.

El fundamento del método está basado en que la curva de producción de leche se mantiene más o menos estacionaria después de un breve ascenso en el inicio de la producción, por eso este tipo de diseño se debe aplicar a animales que ya hayan sobrepasado el pico de la lactación, o sea después de dos meses de lactación aproximadamente y

que el experimento finalice antes de pasar la mitad del período de gestación subsiguiente.

Siendo p el número de tratamientos cada ensayo debe tener $P(p-1)$ animales o múltiplo de este número.

Para dos tratamientos (A, B) debemos usar dos animales o múltiplo de este número. Los animales de cada par recibirán por sorteo los tratamientos, ejemplo.

	Animal 1	Animal 2
Primer período	A	B
Segundo período	B	A
Tercer período	A	B

Para el caso de tres tratamientos el número de animales es seis.

A	B	C	B	C	A
B	C	A	A	B	C
A	B	C	B	C	A

Si tuviéramos se puede comenzar el ensayo simultáneamente con todos o si es necesario podemos comenzar con tres animales en un bloque, dejando los otros animales para otro bloque.

Bloque 1			Bloque 2		
A	B	C	B	C	A
B	C	A	A	B	C
A	B	C	B	C	A

Para cada animal en cualquier caso el valor $D = Y_1 - 2Y_2 + Y_3$ que es la base de todo el análisis estadístico.

$$FC = \frac{\sum D^2}{3np(p-1)}$$

$n = 2r$, r es el número de vacas usadas para cada una de las secuencias de tratamientos.

$$SC \text{ total} = \left(\frac{1}{6}\right) \sum D^2 - FC$$

Cálculo de Q_i correspondiente a cada uno de los tratamientos.

$$Q_i = \sum 'D - \sum '' D$$

$\sum 'D$ es la suma de los valores D de las vacas que reciben el tratamiento en cuestión, en el primero y tercer período y $\sum ''D$ es la suma de los valores de D de las vacas que lo reciben en el segundo período.

$$SC \text{ tratamientos} = \left(\frac{1}{6} np\right) \sum Q^2$$

En el caso de que no haya bloques, el análisis de variancia es el siguiente.

FV	GL	SC
Tratamiento	$p - 1$	$(1/6^{np}) \sum Q^2$
Error	por diferencia	por diferencia
Total	$1/2 np (p-1) - 1$	$1/6 \sum D^2 - FC$

en el caso de que haya bloque

$$SC \text{ bloque} = (1/6) \sum (1/m_u) B_u^2 - FC$$

B_u = total de bloque

y m_u = número de animales en el mismo; Análisis es como sigue.

FV	GL	SC
Bloques	$b - 1$	$(1/6) \sum (1/\mu) B_{\mu}^2 - FC$
Tratamientos	$p - 1$	$(1/6 np) \sum Q^2$
Error	por diferencia	por diferencia
	$(1/2)np(p-1)-1$	$(1/6) \sum D^2 - FC$

En los dos casos se hace la prueba de F corriente. Las medias de los tratamientos se calculan por la fórmula $m_i = \bar{Y} + (1/2 np) Q_i$

En el caso de que p sea impar y mayor que 4 se puede usar un delineamiento reducido con $(\frac{1}{2}) p (p-1)$ secuencia de tratamientos y se toman $n = r$.

ESTANDARIZACION DE DATOS

Dos dificultades principales se presentan con la manipulación de los datos provenientes de animales: El primero se debe a la no ortogonalidad y el segundo al efecto de los factores discrepantes que puede enmascarar la información obtenida.

El primer punto (no ortogonalidad) lo hemos discutido ampliamente y hemos dados varios métodos que se pueden usar según el caso para alcanzar una interpretación correcta de los datos.

El segundo punto lo vamos a discutir a continuación, vale decir el tipo de correcciones que deben aplicarse para estandarizar los datos. La estandarización tiene por objeto eliminar el efecto de algunos factores que pueden alterar la veracidad de la información,

o que no tiene razón de ser incluido en el estudio. De esta manera se puede aumentar la precisión con que se estiman el efecto de los factores en estudio.

Sin embargo debe hacerse una advertencia importante que la estandarización que se aplica en forma desmedida puede viciar más la información en vez de aumentar la precisión y la confianza en ella.

Para decidir que tipo de correcciones se debe aplicar conviene consultar la experiencia que se tiene sobre ella, esto se consigue por medio de la revisión de literatura u otros medios. Nosotros tomaremos la estandarización como un procedimiento para ajustar los datos a una base común.

El criterio más común aplicado para apreciar la bondad de un factor de ajuste es la similitud que guardan los promedios, variancias o la estabilización de los coeficientes de variación.

De modo general se pueden clasificar los métodos empleados para la estandarización de los datos en cuatro tipos principales:

- a) correcciones por regresión
- b) correcciones por multiplicación (o porcentaje)
- c) correcciones por adicción
- d) otros.

Método de regresión

Este método se usa siempre y cuando se disponga de una (o más) variable concomitante, es decir además de la variable en estudio y (dependiente) se tiene información de una variable independiente (x) que se utiliza para calibrar aquella y de este modo conseguir

informaciones más válidas, libres del efecto de la variable o las variables que están asociadas a ella.

Ya hicimos una advertencia acerca del criterio a usar para la aplicación de los métodos de ajustes. Con el método de regresión debe tenerse mayor cuidado para su aplicación. Como ejemplo citaremos la aplicación del método de regresión para estandarizar los datos de lactancia según la edad de la vaca. En este caso claro está que la ecuación de regresión incluye no solo el efecto de la edad sino también el efecto de ciertos cambios medioambientales que ocurre en la vida del animal.

Método por multiplicación

Un ajuste multiplicativo es apropiado si los coeficientes de variación entre los factores que se quieren ajustar a una base común son iguales. En este caso la razón entre los valores promedios es el factor de ajuste adecuado. Los valores de los promedios se deben calcular en forma ponderada.

Este tipo de corrección corrientemente se aplica para estandarizar el sexo. A continuación presentamos un modelo que aclarará el concepto. Si tenemos dos factores en estudio Padres y años en forma esquemática podemos representar así.

		Padres							
Años		P ₁		P ₂		P ₃		P _n	
		M	H	M	H	M	H	M	H
A ₁									
A ₂									
...									
A _n									

Cada celda formada por A_i x P_i constituye una sub-clase donde se conoce el número de machos (M) y el número de hembras (H). Los promedios se obtienen ponderando cada promedio de sub-clase Padre-año, para cada sexo, por el factor $\frac{M \cdot H}{M+H}$; donde

M = número de machos en cada sub-clase padre-año

H = número de hembras en cada sub-clase padre - año

Este método se usa para remover los factores genéticos y ambientales que pueden haber diferido entre grupo.

Para calcular el factor de ajuste (multiplicativo) se dividen los promedios ponderados (razón de promedios) en el caso de que se tenga macho y hembra se podría ajustar al sexo que tiene mayor número. Por ejemplo estandarizar al sexo macho. Se usaría como factor $\frac{\bar{M}}{\bar{H}}$ para multiplicar por cada observación de la hembra, o lo mismo se puede usar el factor $\frac{\bar{H}}{\bar{M}}$ para estandarizar macho a hembra.

Si se tuviera macho, hembra y capones se podría operar en la misma forma ya indicada.

Corrección por adición

La estandarización por este método tiene fundamento similar al anterior, pero en vez de usar un factor multiplicativo, se usa un valor aditivo. Esto deja ver claramente que entre los niveles de la variable que se quiere estandarizar tiene que existir una relación de aumento lineal.

Si pretendiéramos aplicar a nuestro ejemplo anterior de estandarización de sexo, el método aditivo procederíamos en la misma forma que el anterior, pero para calcular el estandarizador (valor aditivo) se estimaría la diferencia promedio ($\bar{M} - \bar{H}$), estos promedios también son ponderados.

El ajuste aditivo es apropiado si las desviaciones estandares entre las observaciones que se quiere estandarizar sean iguales. En nuestro ejemplo anterior para ajustar a una base común de sexo digamos a macho, se tomaría la diferencia ($\bar{M} - \bar{H}$) para sumarle a cada observación H.

Es muy corriente también observar la estandarización de la edad de los animales lecheros sobre la base de una lactancia determinada digamos tercera para este ajuste corrientemente se usa un valor aditivo. Debe tenerse en cuenta que las vacas que comienzan con una producción inicial baja tienen un aumento proporcionalmente más alto en la producción con la edad, que las vacas que comienzan con una producción inicial alta. Si en las correcciones usáramos el criterio de

aplicar aumentos porcentuales (corrección en porcentaje) esta tendería a sub-corregir el merito de los animales que comienzan con una producción baja y a sobre-corregir a aquellas que inician con una producción alta.

Las correcciones aditivas dan resultados insatisfactorios con animales con producción consistentemente alta o consistentemente baja.

Otras correcciones

Dentro de este tipo de correcciones caen una diversidad de fórmulas específicas aplicadas, según el caso. Simplemente hacemos una llamada acerca de estas fórmulas, que no las vamos a discutir ni presentar. Sin embargo puede decirse que todas las fórmulas tienen fundamento en uno u otro de los tres discutidos o combinación de los criterios usados en cada una de ellas.

BIBLIOGRAFIA

1. ACKOFF, R. L., GUPTA, S. K. & SAYER MINAS, J. Scientific method. Willey. New York. 1962. 464 p.
2. ANDERSON, R. L. & BANCROFT, T. A. Statistical theory in research. McGraw-Hill. New York. 1952. 399 p.
3. BELLMAN, R. Introduction to matrix analysis. McGraw-Hill. 1960. 328 p.
4. BRINKS, J. S. et al. Adjusting birth weight, weaning weight and preweaning gain for sex of calf in range Hereford cattle. Journal of Animal Science 20(2):363-367. 1961.
5. BROWN, C. J. Heritability of weight and certain body dimensions of beef calves at weaning. Arkansas Agricultural Experiment Station. Bulletin No 597. 29 p.

6. COCHRAN, W. G. & COX, G. M. Experimental designs. Wiley. New York. 1950. 454 p.
7. FALCONER, O. S. Introduction to quantitative genetics. Oliver & Boyd. Edinburgh. 1960. 365 p.
8. FEDERER, W. T. Experimental designs: Theory and application. MacMillan. New York. 1955. pp. 544-547.
9. GOMEZ, F. P. Curso de Estadística experimental. Escola Superior de Agricultura "Luiz de Queiros". Sao Paulo. 2nd. ed. (Portugués). 1963. 384 p.
10. GOULDEN, C. H. Methods of Statistical analysis. Wiley. New York. 1952. 467 p.
11. HENDERSON, C. R. Estimation of variance and covariance components. Biometrics 9(2):226-252. 1953.
12. _____. Estimation of general, specific and maternal combining abilities in crosses among inbred of swine. Unpublished Ph.D. Thesis. Iowa State College Library. 199 p.
13. JOHANSSON, I. The heritability of milk and butterfat yield. Animal Breed. Abst. 18:1-12. 1950.
14. KASSAB, S. A. & KAROM, H. A. I. Effects of some environmental factors on body and fleece weights of Barki sheep. Journal of Animal Production. 1(2):149-167. 1961.
15. KRAMER, C. Y. Extension of multiple range tests to group correlated adjusted means. Biometrics 13(1):13-18. 1957.
16. KEMOTHRNE, O. The design and analysis of experiments. Wiley. New York. 1952. 631 p.
17. _____ & TANDON, O. B. The estimation of heretability by regression of offspring on parent. Biometric 9(1):90-100. 1953.
18. KIDWELL, J. F., et al. Heterosis in crosses of inbred lines of rats. Genetics 45(2):225-231. 1960.
19. KOCH, R. M. et al. Evaluating the influence of sex on birth weight an preweaning gain in beef cattle. Journal of Animal Science 18(2):738-744. 1959.
20. LUSH, J. L. & SHRODE, R. R. Changes in milk production with age and milking frequency. Journal Dairy Science 33(5):338-357. 1950.

21. MAHADEVAN, P. The effect of environment and heredity on lactation. I. Milk yield. *Journal of Agricultural Science* 41(1-2):80-88. 1951.
22. MARLOWE, T. J. & GAINES, J. A. The influence of age, sex, and season of birth of calf, and age of dam on preweaning growth rate and type score of beef calves. *Journal of Animal Science* 17(3):706-713. 1958.
23. PAEZ, G. Estudios sobre tamaño y forma de parcela para ensayos en café. Tesis Mag. Agr. Turrialba, Costa Rica. Instituto Interamericano de Ciencias Agrícolas. 1962. 114 p. (Mimeografiada)
24. REEVE, E. C. R. The variance of the genetic correlation coefficient. *Biometrics* 11(3):357-374. 1955.
25. _____ & ROBERTSON, F. W. Studies in quantitative inheritance. II. Analysis of a strain of *Drosophila melanogaster* selected for long wings. *J. Genetics* 51:276-316. 1953.
26. SANDERS, H. G. The variations in milk yields caused by season of the year, service, age, and dry period and their elimination. Part I. *Journal of Agricultural Science* 17(3):339-379. 1927.
27. _____. The variations in milk yields caused by season of the year, service, age, and dry period and their elimination. Part II. *Journal of Agricultural Science* 17(4):502-523. 1927.
28. SNEDECOR, G. W. & COX, G. M. Disproportionate subclass number in table of multiple classification. *Iowa State College Agricultural Experiment Station* 15(180):235-272. 1935.
29. SNEDECOR, G. W. *Statistical methods applied to experiments in agriculture and biology*. 5th. ed. Ames, Iowa State College Press. 1956. 534.
30. STEEL, R. G. D. & TORRIE, J. H. *Principles and procedures of statistics, with special reference to biological sciences*. McGraw-Hill. New York. 1960. 481 p.
31. UMAÑA, R. Número de repeticiones necesario en varios ensayos. SEDETE. Doc. 8. Instituto Interamericano de Ciencias Agrícolas. 1962.
32. _____. *Diseño y análisis de experimentos*. Comunicación personal. 1961.

33. UNITED STATES DEPARTMENT OF AGRICULTURE. Least-squares analysis of data with unequal subclass numbers. USDA Agricultural Research Service. USA. 1960. 157 p.
34. _____. Mean separation by the functional analysis of variance and multiple comparisons. USDA. Agricultural Research Service. USA. 1957. 33 p.
35. WILSON, E. B. An introduction to scientific research. McGraw-Hill. New York. 1952. 375 p.
36. WORTHING, A. G. & GEFNER, J. Treatment of experimental data. Wiley. New York. 1943. 342 p.

APENDICE

Determinantes

Con el fin de aclarar el significado de la palabra determinante comenzaremos con un ejemplo, si consideramos dos ecuaciones lineales homogéneas:

$$a_1X + b_1Y = 0 \quad (1)$$

$$a_2X + b_2Y = 0 \quad (2)$$

Si multiplicamos la primera ecuación por b_2 y la segunda por b_1

$$a_1b_2X + b_1b_2Y = 0$$

restando miembro a miembro

$$-a_2b_1X + b_1b_2Y = 0$$

$$a_1b_2X - a_2b_1X = 0$$

dividiendo por X tenemos

$$a_1b_2 - a_2b_1 = 0$$

Este resultado se puede escribir así:

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = 0$$

La expresión resultante de este producto cruzado y sus diferencias reciben el nombre de determinante de 2º orden. Las letras a_1 , a_2 , b_1 y b_2 reciben el nombre de elementos del determinante.

Propiedad 1. El valor del determinante no se altera por el cambio de hileras en columnas, y las columnas en hileras.

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1 = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$$

Propiedad 2. Si intercambiamos dos hileras o dos columnas del determinante, se obtiene un determinante que difiere del primero solamente en signo.

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = - \begin{vmatrix} b_1 & a_1 \\ b_2 & a_2 \end{vmatrix} \quad \text{y} \quad \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = - \begin{vmatrix} a_2 & b_2 \\ a_1 & b_1 \end{vmatrix}$$

Si ahora consideramos tres ecuaciones lineares homogéneas.

$$a_1X + b_1Y + c_1Z = 0$$

$$a_2X + b_2Y + c_2Z = 0$$

$$a_3X + b_3Y + c_3Z = 0$$

Por eliminación de X, Y, Z obtenemos

$$a_1(b_2c_3 - b_3c_2) + b_1(c_2a_3 - c_3a_2) + c_1(a_2b_3 - a_3b_2) = 0$$

o sea

$$a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} + b_1 \begin{vmatrix} c_2 & a_2 \\ c_3 & a_3 \end{vmatrix} + c_1 \begin{vmatrix} c_2 & a_2 \\ c_3 & a_3 \end{vmatrix} = 0$$

Generalmente esto se escribe así:

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0 \quad \text{Este es el valor del determinante de tercer orden.}$$

Por arreglo de los términos en la forma arriba explicada de determinante puede escribirse así:

$$a_1(b_2c_3 - b_3c_2) + a_2(b_3c_1 - b_1c_3) + a_3(b_1c_2 - b_2c_1)$$

o sea

$$a_1 \begin{bmatrix} b_2 & b_3 \\ c_2 & c_3 \end{bmatrix} + a_2 \begin{bmatrix} b_3 & b_1 \\ c_2 & c_1 \end{bmatrix} + a_3 \begin{bmatrix} b_1 & b_2 \\ c_1 & c_3 \end{bmatrix}$$

$$\text{Puesto que } \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_3 & c_3 \end{bmatrix}$$

Recordamos acá la propiedad. (El valor de los determinantes no se altera por el cambio de hilera en columna o columna en hileras).

De lo anterior se desprende que:

$$\begin{aligned} \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} &= a_1 \begin{bmatrix} b_2 & c_2 \\ b_3 & c_3 \end{bmatrix} + a_2 \begin{bmatrix} b_3 & c_3 \\ b_1 & c_1 \end{bmatrix} + a_3 \begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix} \\ &= a_1 \begin{bmatrix} b_2 & c_2 \\ b_3 & c_3 \end{bmatrix} - a_2 \begin{bmatrix} b_1 & c_1 \\ b_3 & c_3 \end{bmatrix} + a_3 \begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix} \quad (1) \end{aligned}$$

o también

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = a_1 \begin{bmatrix} b_2 & c_2 \\ b_3 & c_3 \end{bmatrix} - b_1 \begin{bmatrix} a_2 & c_2 \\ a_3 & c_3 \end{bmatrix} + c_3 \begin{bmatrix} a_2 & b_2 \\ a_3 & b_3 \end{bmatrix}$$

De esto podemos escribir una sencilla regla para desarrollar el determinante de tercer orden.

El coeficiente de cualquiera de los componentes a_1 a_2 a_3 es aquel determinante de segundo orden que se obtiene omitiendo la hilera y la columna en que este aparece. Estas determinantes son "Minor" del determinante original y el "left-hand side" de la ecuación se puede escribir (1). $a_1 A_1 - a_2 A_2 + a_3 A_3$

donde A_1 , A_2 y A_3 son los "minors" de a_1 , a_2 y a_3 respectivamente.

La ecuación (2) se puede escribir así:

$$a_1 A_1 - b_1 B_1 + c_1 C_1$$

El determinante $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1(b_2c_3 - b_3c_2) + b_1(c_2a_3 - c_3a_2) + c_1(a_2b_3 - a_3b_2)$
 $= -b_1(a_2c_3 - a_3c_2) - a_1(a_2b_3 - a_3b_2) - c_1(b_2a_3 - b_3a_2)$

Puesto que

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = - \begin{vmatrix} b_1 & a_1 & c_1 \\ b_2 & a_2 & c_2 \\ b_3 & a_3 & c_3 \end{vmatrix} \quad \text{(de acuerdo con la propiedad (2))}$$

Propiedad 3. Si dos hileras o dos columnas del determinante son idénticas el valor del determinante es nulo.

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \quad \text{El determinante} = a_1b_2c_3 - (b_1a_2c_3) = - (a_1b_2c_3) = (a_1b_2c_3)$$

Si llamamos D al valor determinante, por intercambio de dos hileras y dos columnas obtenemos un determinante cuyo valor es - D pero

el determinante se altera puesto que $D = -$ (menos) D que es $D=0$, así tenemos las siguientes ecuaciones.

$$a_1A_1 - a_2A_2 + a_3A_3 = D$$

$$b_1A_1 - b_2A_2 + b_3A_3 = 0$$

$$c_1A_1 - c_2A_2 + c_3A_3 = 0$$

Propiedad 4. Si cada componente en cualquier hilera o columnas se multiplica por el mismo factor, entonces el determinante queda multiplicado por ese factor.

$$\begin{vmatrix} ma_1 & b_1 & c_1 \\ ma_2 & b_2 & c_2 \\ ma_3 & b_3 & c_3 \end{vmatrix} = ma_1A_1 - ma_2A_2 + ma_3A_3 = m(a_1A_1 - a_2A_2 + a_3A_3)$$

Propiedad 5. Si cada uno de los componentes en cualquier hilera o columna consta de dos términos, entonces el determinante se puede expresar como la suma de dos otros determinantes.

$$\begin{vmatrix} a_1 + \alpha_1 & b_1 & c_1 \\ a_2 + \alpha_2 & b_2 & c_2 \\ a_3 + \alpha_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} \alpha_1 & b_1 & c_1 \\ \alpha_2 & b_2 & c_2 \\ \alpha_3 & b_3 & c_3 \end{vmatrix} = (a_1 + \alpha_1)A_1 - (a_2 + \alpha_2)A_2 + (a_3 + \alpha_3)A_3 = a_1A_1 - a_2A_2 + a_3A_3 + (\alpha_1A_1 - \alpha_2A_2 + \alpha_3A_3)$$

De la misma manera, si cada componente en cualquier hilera o columna consiste de m términos, el determinante se puede expresar como la suma de otras determinantes.

$$\begin{bmatrix} a_1 + \alpha_1 & b_1 + \beta_1 & c_1 \\ a_2 + \alpha_2 & b_2 + \beta_2 & c_2 \\ a_3 + \alpha_3 & b_3 + \beta_3 & c_3 \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} + \begin{bmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{bmatrix} + \begin{bmatrix} \alpha_1 & b_1 & c_1 \\ \alpha_2 & b_2 & c_2 \\ \alpha_3 & b_3 & c_3 \end{bmatrix} + \begin{bmatrix} \alpha_1 & \beta_1 & c_1 \\ \alpha_2 & \beta_2 & c_2 \\ \alpha_3 & \beta_3 & c_3 \end{bmatrix}$$

Matrices

Una matriz es un arreglo de cantidades en hileras y columnas, como por ejemplo:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{1m} & a_{m2} & & a_{mn} \end{bmatrix}$$

Las cantidades a_{ij} reciben el nombre de elemento de la matriz. Esta es una matriz de m hilera y n columnas. Si $m = n$ la matriz se llama matriz cuadrada. El número de hileras o columnas de una matriz cuadrada se llama orden de la matriz. Si a_{ij} es igual a_{ji} la matriz se llama matriz simétrica. Dos matrices son iguales si los elementos correspondientes son idénticos.

Para la mayoría de nuestros casos nos referiremos a la matriz cuadrada y simétrica.

Suma de matrices. La suma de dos matrices se calcula de la manera siguiente:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}$$

Resta de matrices. Se define en la misma forma que la suma.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} - \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11} - b_{11} & a_{12} - b_{12} \\ a_{21} - b_{21} & a_{22} - b_{22} \end{bmatrix}$$

Multiplicación de dos matrices. En la multiplicación de dos matrices, los elementos del producto-Matriz se obtienen multiplicando los elementos de una hilera de la primera matriz por el correspondiente elemento de una columna de la segunda matriz, luego sumar los términos productos. Por ejemplo:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

Si nosotros llamamos A stand a la primera matriz de arriba y B a la segunda es obvio que A . B no necesariamente debe ser igual a B . A.

División de una matriz por otra. La división se define como una operación inversa a la multiplicación.

$$A \cdot B = G$$

Si multiplicamos los dos miembros de esta ecuación por A^{-1} que es la inversa de A, tenemos

$$A^{-1} \cdot A \cdot B = A^{-1} G$$

El valor de A^{-1} se define como una matriz tal que

$$A^{-1} \cdot A = A \cdot A^{-1} = I$$

donde I recibe el nombre de matriz de identidad.

La matriz idéntica I en término de sus elementos es:

$$I = \begin{bmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & \dots & 1 \end{bmatrix}$$

Volviendo a la ecuación $A \cdot B = G$ se puede ver que B es igual a GA^{-1} , entonces la definición de la operación de división de la matriz G dividido por la matriz A sería como: El producto de la matriz G por la Inversa de la matriz A ($B = G.A^{-1}$).

Para explicar el método general de obtener la inversa de una matriz se necesita recordar las propiedades básicas de los determinantes.

Los elementos de una matriz cuadrada determina el determinante de la matriz. Un determinante es una función de los elementos de un arreglo cuadrado, se podría expresar como un polinomio, desarrollando los determinantes por "minors" de acuerdo a la regla de Cramer.

El orden de un determinante es su número de hileras o columnas. Un "minor" de un elemento de un determinante es el determinante de un orden inferior, encontrado por resolución de hileras y columnas que contienen el elemento particular.

El cofactor A_{ij} del elemento a_{ij} se obtiene por multiplicación de $(-1)^{i+j}$ por el menor de a_{ij} . La regla de Cramer nos permite evaluar un determinante por multiplicación de los elementos de cualquier hilera o columna por el correspondiente cofactor y luego sumar los productos.

Los elementos a_{ij} de la matriz inversa A^{-1} de la matriz A se obtiene por división del cofactor A_{ij} por el determinante A de la matriz que es:

$$a_{ij} = \frac{A_{ij}}{A}$$

Un ejemplo numérico. Para encontrar la matriz inversa de A que es en el ejemplo siguiente se procede como se explicó más arriba.

$$|A| = \begin{vmatrix} 0 & 3 & 2 \\ 1 & 2 & 4 \\ 3 & 0 & 2 \end{vmatrix}$$

1º) Cálculo del determinante de la matriz

$$|A| = \begin{vmatrix} 0 & 3 & 2 \\ 1 & 2 & 4 \\ 3 & 0 & 2 \end{vmatrix} = 0 \begin{vmatrix} 2 & 4 \\ 0 & 2 \end{vmatrix} - 3 \begin{vmatrix} 1 & 4 \\ 3 & 2 \end{vmatrix} + 2 \begin{vmatrix} 1 & 2 \\ 3 & 0 \end{vmatrix} = 0 - 3(2-12) + 2(0-6)$$

$$|A| = \underline{\underline{18}}$$

2º) Cálculo de los cofactores:

$$A_{11} = (-1)^{1+1} \begin{vmatrix} 2 & 4 \\ 0 & 2 \end{vmatrix} = +4$$

$$A_{12} = (-1)^{1+2} \begin{vmatrix} 1 & 4 \\ 3 & 2 \end{vmatrix} = +10$$

$$A_{13} = (-1)^{1+3} \begin{vmatrix} 1 & 2 \\ 3 & 0 \end{vmatrix} = -6$$

$$A_{21} = (-1)^{2+1} \begin{vmatrix} 3 & 2 \\ 0 & 2 \end{vmatrix} = -6$$

$$A_{22} = (-1)^{2+2} \begin{bmatrix} 0 & 2 \\ 3 & 2 \end{bmatrix} = -6$$

$$A_{32} = (-1)^{2+3} \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix} = +9$$

$$A_{23} = (-1)^{3+1} \begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix} = +8$$

$$A_{32} = (-1)^{3+2} \begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix} = +2$$

$$A_{33} = (-1)^{3+3} \begin{bmatrix} 0 & 3 \\ 1 & 2 \end{bmatrix} = -3$$

Resumiendo tenemos

$$A_{ij} = \begin{bmatrix} 4 & 10 & -6 \\ -6 & -6 & 9 \\ 8 & 2 & -3 \end{bmatrix}$$

3º) Intercambiando hileras y columnas de A_{ij} tenemos:

$$A_{ji} = \begin{bmatrix} 4 & -6 & 8 \\ 10 & -6 & 2 \\ -6 & 9 & -3 \end{bmatrix}$$

4º) Dividiendo cada elemento de A_{ji} por $|A|$ tenemos la inversa de la matriz A que es A^{-1} .

$$A^{-1} = \begin{bmatrix} \frac{4}{18} & \frac{-6}{18} & \frac{8}{18} \\ \frac{10}{18} & \frac{-6}{18} & \frac{2}{18} \\ \frac{-6}{18} & \frac{9}{18} & \frac{-3}{18} \end{bmatrix} \quad A^{-1} = \begin{bmatrix} 0.222222 & -0.333333 & 0.444444 \\ 0.555555 & -0.333333 & 0.111111 \\ -0.333333 & 0.500000 & -0.166667 \end{bmatrix}$$

52) Podemos verificar numericamente que

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Los elementos de la matriz I se obtienen en la siguiente forma:

1^{er} elemento $0 \left(\frac{4}{18}\right) + 3 \left(\frac{10}{18}\right) + 2 \left(\frac{-6}{18}\right) = 1$

luego: $0 \left(\frac{-6}{18}\right) + 3 \left(\frac{-6}{18}\right) + 2 \left(\frac{9}{18}\right) = 0$

$0 \left(\frac{8}{18}\right) + 3 \left(\frac{2}{18}\right) + 2 \left(\frac{-3}{18}\right) = 0$

$1 \left(\frac{4}{18}\right) + 2 \left(\frac{10}{18}\right) + 4 \left(\frac{-6}{18}\right) = 0$

$1 \left(\frac{-6}{18}\right) + 2 \left(\frac{-6}{18}\right) + 4 \left(\frac{9}{18}\right) = 1$

$1 \left(\frac{8}{18}\right) + 2 \left(\frac{2}{18}\right) + 4 \left(\frac{-3}{18}\right) = 0$

$3 \left(\frac{4}{18}\right) + 0 \left(\frac{10}{18}\right) + 2 \left(\frac{-6}{18}\right) = 0$

$3 \left(\frac{-6}{18}\right) + 0 \left(\frac{-6}{18}\right) + 2 \left(\frac{9}{18}\right) = 0$

$3 \left(\frac{8}{18}\right) + 0 \left(\frac{2}{18}\right) + 2 \left(\frac{-3}{18}\right) = 1$

IICA
LOO
32
Autor

METODOS DE INVESTIGACION
EN PRODUCCION ANIMAL

Título

Nombre del solicitante

Fecha Devolución

8 SET 1984

Jorge Caro



