

METODOLOGIA DE LOS MODELOS LINEALES MIXTOS,  
UNA ALTERNATIVA AL DISEÑO EXPERIMENTAL  
CONVENCIONAL

SALOMON GONZALEZ DE LA TORRE

T E S I S

PRESENTADA COMO REQUISITO PARCIAL

PARA OBTENER EL GRADO DE:

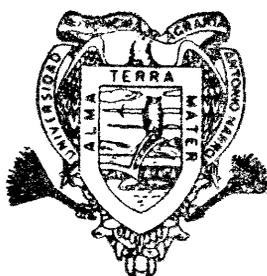
MAESTRO EN CIENCIAS

EN ESTADISTICA EXPERIMENTAL

Universidad Autónoma Agraria  
"ANTONIO NARRO"



BIBLIOTECA



Universidad Autónoma Agraria  
"Antonio Narro"

PROGRAMA DE GRADUADOS

Buenvista, Saltillo, Coah.

JUNIO DEL 2000

**UNIVERSIDAD AUTÓNOMA AGRARIA  
ANTONIO NARRO**

**SUBDIRECCIÓN DE POSTGRADO**

**METODOLOGÍA DE LOS MODELOS LINEALES MIXTOS, UNA  
ALTERNATIVA AL DISEÑO EXPERIMENTAL CONVENCIONAL**

**TESIS**

**Por**

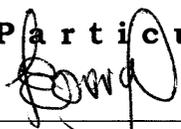
**SALOMÓN GONZÁLEZ DE LA TORRE**

Elaborada bajo la supervisión del comité particular de asesoría y  
aprobada como requisito parcial para obtener el grado de:

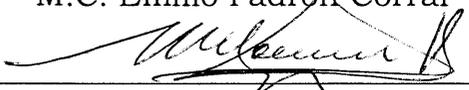
**MAESTRO EN CIENCIAS  
EN ESTADÍSTICA EXPERIMENTAL**

**Comité Particular**

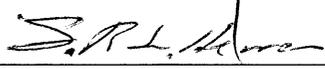
Asesor principal:

  
\_\_\_\_\_  
M.C. Emilio Padrón Corral

Asesor :

  
\_\_\_\_\_  
M.C. Regino Morones Reza

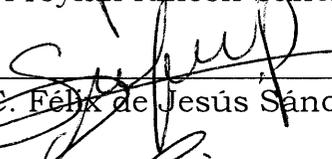
Asesor :

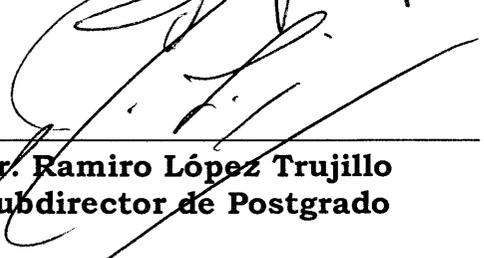
  
\_\_\_\_\_  
Dr. Sergio A. Rodríguez Herrera

Asesor :

  
\_\_\_\_\_  
Dr. Froylán Rincón Sánchez

Asesor :

  
\_\_\_\_\_  
M.C. Félix de Jesús Sánchez Pérez

  
\_\_\_\_\_  
**Dr. Ramiro López Trujillo**  
**Subdirector de Postgrado**

**Buenavista, Saltillo, Coah., México. Junio de 2000**

## **AGRADECIMIENTOS**

A la Universidad Autónoma Agraria Antonio Narro por brindarme la oportunidad para realizar este postgrado.

Al M.C. Emilio Padrón Corral asesor principal por guiar este trabajo y mantenerme continua la llama encendida de la motivación.

A los Dres. Sergio Rodríguez y Froylán Rincón S., quienes con su valiosa asesoría, ayudaron al desarrollo de este trabajo.

Al M.C. Regino Morones Reza, por sus insuperables cátedras y su ayuda incondicional.

Al M.C. Félix de Jesús, por iniciarme en esta maravillosa aventura de la Estadística Experimental.

A todos mis maestros de la Maestría de Estadística Experimental, quienes me dieron el conocimiento, tesoro inapreciable.

A todos mis compañeros de estudio, en quienes hallé fraternidad y comprensión.

A Ma. de Lourdes Villarreal, por su habilidad para reproducir mis manuscritos.

## **DEDICATORIA**

A mis padres ausentes: Salomón González y Consuelo de la Torre, quienes me dieron todo a cambio de nada.

A mi esposa: Irma Arellano de G., Crisol donde se amalgaman todas las virtudes.

A mis hijos: Salomón, Claudia, Edgar, Griselda y Virginia, como un presente por las infinitas satisfacciones que me han proporcionado.

A mis hermanos: Rubén, Lenin (ausente), Griselda, Edmundo, Hugo, Yolanda, Xóchitl, Ma de la Luz, Eduardo y Ma. de Jesús; todos profesionistas de quienes me siento orgulloso y procuro estar a su “altura”.

Por último, a mi querido pueblo: Jerez, Zac.

## COMPENDIO

### **Metodología de los Modelos Lineales Mixtos, Una Alternativa al Diseño Experimental Convencional**

**POR:**

**SALOMÓN GONZÁLEZ DE LA TORRE**

**MAESTRIA EN**

**ESTADISTICA EXPERIMENTAL**

**UNIVERSIDAD AUTONOMA AGRARIA ANTONIO NARRO**

**BUENAVISTA, SALTILLO, COAHUILA, JUNIO DE 2000**

**M.C. Emilio Padrón Corral –Asesor-**

**Palabras clave:** Algebra de matrices, inversas generalizadas, el modelo lineal general, el modelo lineal mixto, componentes de varianza.

Ante la creciente población mexicana y la escasez cada vez mayor de alimentos, se está ante una tarea urgente y prioritaria el hacer del campo un pilar donde se sustente la economía. Es importante tener presente que, dadas las características del territorio nacional, el país está afortunadamente destinado a ser un país agrícola, ganadero.

Compete por ello, a las instituciones agronómicas la noble tarea de hacer del campo una entidad eficiente y productiva.

El presente trabajo, que se inicia con el álgebra de matrices, luego presenta algo de la teoría del modelo lineal general, posteriormente aborda la metodología del modelo lineal mixto; presentando por último un ejemplo, comparativo con el diseño experimental convencional, así; intenta aportar una ayuda al avance de la investigación agropecuaria, ya que con el desarrollo exponencial de la computación, permite hacer de la teoría del modelo lineal mixto, una realidad factible de llevarse a la práctica.

## **ABSTRACT**

### **Methodology of the Mixed Linear Models, an Alternative to the Conventional Experimental Designs**

**BY**

**SALOMÓN GONZÁLEZ DE LA TORRE**

**MASTER OF SCIENCE**

**EXPERIMENTAL STATISTICS**

**UNIVERSIDAD AUTONOMA AGRARIA ANTONIO NARRO**

**BUENAVISTA, SALTILLO, COAHUILA, JUNE, 2000**

**M.C. Emilio Padrón Corral -Adviser-**

**Key words:** Matrix algebra, generalized inverse, the general linear model, the mixed linear model, variance components.

Due to the growing population of Mexico and the general scarcity of food, it has consider the country side a foundation which supports the economy, as an urgent task of high priority. It is important to remember that, given the characteristics of the national territory, the country is fortunately destined to be both agricultural and cattle dealer.

Therefore, it is the noble duty of the agronomic institutions to transform the country side into an efficiently productive entity.

The present work, which will begin with matrix algebra, will next present a part of the theory of the general linear model, in order to then touch the methodology of the mixed linear model. It will lastly present a comparative example with the conventional experimental design; thus, this thesis intends to provide help in the development of the agricultural research because the computation exponential advance allows us to make the mixed linear model theory a feasible reality.

## INDICE DE CONTENIDO

|   |           |
|---|-----------|
| <b>INDICE DE CUADROS</b> .....                                    | <i>xi</i> |
| <b>I. INTRODUCCIÓN</b> .....                                      | 1         |
| <b>II. EL MODELO LINEAL GENERAL</b> .....                         | 4         |
| Modelos Lineales .....  | 4         |
| Formas Cuadráticas.....   | 6         |
| Modelos de Rango Completo.....                                    | 10        |
| Modelos de Rango Incompleto.....                                  | 18        |
| Modelos Básicos del Diseño Experimental, ANVA<br>en una Vía.....  | 23        |
| <b>III. EL MODELO LINEAL MIXTO</b> .....                          | 31        |
| Generalidades.....  | 31        |
| Conceptos Básicos de la Metodología de los Modelos<br>Mixtos..... | 41        |
| Parámetros, Estimadores y Valores<br>Esperados .....              | 43        |
| El Mejor Estimador Lineal Inssegado (BLUE).....                   | 54        |
| El Mejor Predictor Lineal Inssegado (BLUP).....                   | 61        |
| Ecuaciones de los Modelos Mixtos (MME).....                       | 66        |
| Componentes de Varianza .....                                     | 72        |
| El Modelo Mixto como Alternativa de un Diseño de                  |           |

|  |    |
|--|----|
| Bloques Incompletos $\alpha$ -Látice Simple .....                                | 82 |
| <b>IV. EL MODELO DE BLOQUE ALEATORIO</b> .....                                   | 84 |
| Introducción .....   | 84 |
| Modelo de Bloque Aleatorizado, no Interacción,<br>una Observación por Celda..... | 86 |
| <b>V. LITERATURA CITADA</b> .....  | 90 |

## INDICE DE CUADROS

|   | <b>Página</b> |
|---|---------------|
| 3.1. Características del modelo de efectos fijos y del modelo de efectos aleatorios (clasificación: una vía)..... | 34            |
| 3.2. Comparaciones entre los modelos fijo y aleatorio.....  | 35            |
| 3.3. Pruebas F para modelos ANOVA de dos factores con repetición.....   | 36            |
| 3.4. Desarrollo histórico del modelo mixto.....   | 38            |
| 3.5. Información relevante de un Diseño de Látice Simple 16 x 16.....   | 83            |

## I. INTRODUCCIÓN

Es sabido que para los países latinoamericanos y con ellos México, la agricultura es prioritaria para su desarrollo económico y, en consecuencia; esta investigación agropecuaria debe incrementarse si se desea el progreso del país.

Este trabajo pretende aplicar una técnica de investigación: la metodología del modelo lineal mixto, la cual no siendo reciente y desarrollada con los trabajos de Henderson (1949), ha tomado notable actualidad debido en gran medida al desarrollo tecnológico de la computación. Se menciona, a manera de muestra y como justificación algunas investigaciones que han impulsado su avance:

El enfoque moderno (Steel y Torrie, 1980) de la notación matricial de la regresión múltiple, necesaria en la interpretación de resultados impresos obtenidos de la tecnología computacional, estimulando así al investigador para utilizar paquetes (programas) estadísticos computacionales.

La conveniente utilización de los modelos lineales mixtos (Henderson, 1975) en la mayoría de las aplicaciones de cría de animales. Los métodos para calcular BLUE (mejores estimadores lineales insesgados) de funciones estimables lineales de los elementos fijos del modelo y de los BLUP (mejores predictores lineales insesgados) de los elementos aleatorios deben ser aprovechados, puesto que los métodos usuales producen estimaciones y predicciones a menudo sesgados.

Investigadores como Johnson *et al.*, 1992 han obtenido en poblaciones de maíz, resultados satisfactorios, así como Núñez y Rodríguez (1995) y otros en el mejoramiento animal.

En el área computacional, diseñando programas (software) se tiene a Van Vleck (1994), Boldman (1993), Meyer (1988), Nelder y Mead (1965) quienes han diseñado e innovado paquetes estadísticos computacionales tales como DFREML, SPARSPAK, SPSS, SAS. MTDFREML, por citar algunos.

El desarrollo de computadoras tiende a reducir los problemas de cálculo asociados con el diseño de experimentos y es claro que los diseños requieren de matrices de inversión que pueden ahora utilizarse. Berenson y Levine (1992) afirman que los investigadores pueden ya utilizar las computadoras en la aplicación de métodos estadísticos, especialmente cuando se trabaja con bases de datos grandes o con

procedimientos de cálculo muy complejos, o bien con un número grande de variables.

Harville (1979) considera que estimadores GLS (estimadores por mínimos cuadrados generalizados) son más eficientes que los estimadores OLS (estimadores de mínimos cuadrados ordinarios) correspondientes. La más obligatoria ventaja (Stroup y Mc Lean, 1989) de la metodología del modelo mixto está en la comprensión de la teoría general del modelo lineal (modelo lineal general es un caso particular del modelo mixto).

En este trabajo se introducen aspectos de la teoría general de modelos lineales (Capítulo II) para luego abordar la metodología del modelo lineal mixto y concluir con un ejercicio ilustrativo donde se comparan los resultados obtenidos por el procedimiento de los modelos mixtos con el análisis convencional, además; se pretende disminuir la viciada tendencia del investigador al abuso de la utilización del modelo lineal fijo.

## II. EL MODELO LINEAL GENERAL

### Modelos Lineales

Un modelo es una representación simplificada de un objeto, evento o proceso. Esta representación deberá contener elementos esenciales que lo describan para obtener una buena interpretación y predicción (Priore *et al.*, 1995). Ningún modelo es una imagen exacta del mundo real; las aproximaciones son necesarias. La precisión de un modelo depende del objetivo al cual ha sido diseñado (Hadley, 1969).

Una primera clasificación de los modelos es la de ser o no ser matemáticos, de acuerdo a si se va a utilizar o no lenguaje matemático al modelar (Priore *et al.*, 1995). Para definir matemáticamente cualquier cosa del mundo real, es necesario diseñar un modelo matemático apropiado que relacione las variables que intervienen en él, y el cual consta de una o más ecuaciones o desigualdades. Méndez (1976) puntualiza que para las ciencias los modelos matemáticos son los importantes, que corresponden a representaciones abstractas donde se emplean símbolos para describir un fenómeno.

Podemos distinguir dos tipos de modelos matemáticos: los modelos determinísticos y los modelos estocásticos (probabilísticos o estadísticos). Un modelo determinístico es un postulado matemático que trata de describir exactamente las relaciones entre variables. Si el modelo matemático contiene un elemento aleatorio, entonces se tiene un modelo matemático estocástico. De lo anterior se desprende que la diferencia fundamental entre estos dos modelos estriba en el término aleatorio que aparece en el modelo probabilístico y no en el modelo determinístico.

Dentro de los modelos probabilísticos se realiza una clasificación en cualitativos y cuantitativos. En los modelos cualitativos interesan básicamente los modelos de diseño experimental y los de componentes de varianza (Graybill, 1976). En los modelos cuantitativos interesan sobre todo los de regresión lineal. Estos modelos (regresión lineal, diseño experimental y componentes de varianza) están relacionados y se consideran casos particulares del modelo lineal general.

El modelo estadístico, se caracteriza porque:

- Interviene un elemento aleatorio  $\varepsilon$ .
- Se aplica a un cierto número de eventos (promedio).
- A causa del elemento aleatorio, no se puede determinar con exactitud el valor de una variable en función de otra (s).

## Formas Cuadráticas

Para hacer contrastes cuando  $\sigma^2$  es desconocida y determinar sus propiedades se requiere calcular la distribución de

$$\hat{\mathbf{u}}' \hat{\mathbf{u}} = (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}})' (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}})$$

Esta magnitud es una componente básica de los diversos estimadores de  $\sigma^2$ . Se puede demostrar que la suma de los cuadrados de los residuos es una forma cuadrática en el proceso de los errores del GLM (Modelo Lineal General).

**Proposición 1.** Sea  $\mathbf{y} \sim N(\mathbf{0}, \mathbf{I})$ , donde  $\mathbf{y}$  es  $n \times 1$ , entonces

$$\mathbf{y}' \mathbf{y} \sim \chi^2_n$$

Es decir, una ji-cuadrada con un grado de libertad es la distribución del cuadrado de una variable aleatoria  $N(0,1)$ . La suma de  $n$  variables ji-cuadrada con un grado de libertad (e independientes) sigue una distribución ji-cuadrada con  $n$  grados de libertad.

**Proposición 2.** Sea  $\mathbf{y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ , donde  $\mathbf{y}$  es  $n \times 1$ , entonces

$$(\mathbf{y} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\mu}) \sim \chi^2_n$$

Las dos proposiciones anteriores proporcionan la forma canónica de una distribución ji-cuadrada. El siguiente resultado plantea

parcialmente el problema de dar un criterio que determine cuando una forma cuadrática se distribuye o no según una ji-cuadrada.

**Proposición 3.** Sea  $\mathbf{y} \sim N(\mathbf{0}, \mathbf{I})$ , donde  $\mathbf{y}$  es  $n \times 1$ , y sea  $\mathbf{A}$

$$\mathbf{y}' \mathbf{A} \mathbf{y} \sim \chi^2_r$$

si y sólo si  $\mathbf{A}$  es idempotente.

Puesto que  $\mathbf{A}$  es simétrica, sus valores propios son reales y se puede obtener una matriz de vectores propios ortogonales. Se puede escribir

$$\mathbf{A} = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}'$$

donde

$$\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n),$$

siendo  $\lambda_i$  los valores propios de  $\mathbf{A}$  y  $\mathbf{Q}$  la matriz de vectores propios.

Utilizando la simetría de  $\mathbf{A}$ , se puede representar

$$\mathbf{y}' \mathbf{A} \mathbf{y} = \mathbf{y}' \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}' \mathbf{y} = \sum_{i=1}^n \lambda_i \chi_i^2,$$

en consecuencia, las  $\chi_i^2$  son ji-cuadrado e independientes.

Se van ahora a dar condiciones para que dos formas cuadráticas con variables aleatorias normales, sean independientes.

**Proposición 4.** Sea  $\mathbf{y} \sim N(\mathbf{0}, \mathbf{I})$ , y sean  $\mathbf{A}$  y  $\mathbf{B}$  dos matrices simétricas  $n \times n$ , con  $\mathbf{AB} = \mathbf{0}$ ,  $\mathbf{y}'\mathbf{A}\mathbf{y}$  y  $\mathbf{y}'\mathbf{B}\mathbf{y}$ , son independientes. Entonces

$$\mathbf{AB} = \mathbf{BA}$$

Se observa que

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}' = \mathbf{BA}$$

ya que

$$\mathbf{AB} = \mathbf{0}$$

Aplicando la siguiente proposición (Diagonalización simultánea):

**Proposición 5.** Sean  $\mathbf{A}$  y  $\mathbf{B}$  dos matrices simétricas de orden  $m$ ; existe entonces una matriz ortogonal,  $\mathbf{P}$ , tal que

$$\mathbf{P}'\mathbf{A}\mathbf{P} = \mathbf{D}_1, \quad \mathbf{P}'\mathbf{B}\mathbf{P} = \mathbf{D}_2,$$

Siendo las  $\mathbf{D}_i$ ,  $i=1,2$ , matrices diagonales, si y sólo si

$$\mathbf{AB} = \mathbf{BA}$$

además, como

$$\mathbf{AB} = \mathbf{0}$$

se cumplirá también

$$\mathbf{D}_1\mathbf{D}_2 = \mathbf{0}$$

## Valores Esperados de Formas Cuadráticas

Sea el vector  $\mathbf{Y}$ , tal que  $E[\mathbf{Y}] = \boldsymbol{\mu}$ ,  $V[\mathbf{Y}] = \mathbf{V}$ , entonces  $E[\mathbf{Y}'\mathbf{A}\mathbf{Y}] = \text{tr}[\mathbf{A}\mathbf{V}] + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$

En mínimos cuadrados ordinarios  $E[\mathbf{Y}] = \mathbf{X}\boldsymbol{\beta}$  y  $V[\mathbf{Y}] = \sigma^2\mathbf{I}$ , por tanto  $E[\mathbf{Y}'\mathbf{A}\mathbf{Y}] = \sigma^2 \text{tr}[\mathbf{A}] + \boldsymbol{\beta}'\mathbf{X}'\mathbf{A}\mathbf{X}\boldsymbol{\beta}$ ; entonces, el valor esperado de la SC (modelo) es

$$E[\text{SC modelo}] = E[\mathbf{Y}'\mathbf{P}\mathbf{Y}] = \sigma^2 \text{tr}(\mathbf{P}) + \boldsymbol{\beta}'\mathbf{X}'\mathbf{P}\mathbf{X}\boldsymbol{\beta} = (p+1)\sigma^2 + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}\boldsymbol{\beta} = (p+1)\sigma^2 + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$$

$\boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$  es una forma cuadrática en  $\boldsymbol{\beta}$ . Como  $\mathbf{X}'\mathbf{X}$  es definida positiva, entonces  $\boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$  es definida positiva.

$$E[\text{SC regr}] = E\left[\mathbf{Y}'\left(\mathbf{P} - \frac{\mathbf{J}}{n}\right)\mathbf{Y}\right] = \sigma^2 \text{tr}(\mathbf{P} - \mathbf{J}/n) + \boldsymbol{\beta}'\mathbf{X}'(\mathbf{P} - \mathbf{J}/n)\mathbf{X}\boldsymbol{\beta} = p\sigma^2 + \boldsymbol{\beta}'(\mathbf{X}'\mathbf{P} - \mathbf{X}'\mathbf{J}/n)\mathbf{X}\boldsymbol{\beta} = p\sigma^2 + \boldsymbol{\beta}'\mathbf{X}'(\mathbf{I} - \mathbf{J}/n)\mathbf{X}\boldsymbol{\beta}$$

$\boldsymbol{\beta}'\mathbf{X}'(\mathbf{I} - \mathbf{J}/n)\mathbf{X}\boldsymbol{\beta}$  es una forma cuadrática en  $\boldsymbol{\beta}$  corregida.

$$E[\text{SC res}] = E[\mathbf{Y}'(\mathbf{I} - \mathbf{P})\mathbf{Y}] = \sigma^2 \text{tr}(\mathbf{I} - \mathbf{P}) + \boldsymbol{\beta}'\mathbf{X}'(\mathbf{I} - \mathbf{P})\mathbf{X}\boldsymbol{\beta} = \sigma^2[n - (p+1)] + \boldsymbol{\beta}'(\mathbf{X}'\mathbf{I} - \mathbf{X}'\mathbf{P})\mathbf{X}\boldsymbol{\beta}$$

Los valores esperados de los cuadrados medios son

$$E[\text{CM regr}] = \sigma^2 + [\boldsymbol{\beta}'\mathbf{X}'(\mathbf{I} - \mathbf{J}/n)\mathbf{X}\boldsymbol{\beta}]/p$$

$$E[\text{CM res}] = \sigma^2 \text{ (insesgado)}$$

## Modelos de Rango Completo

El Modelo Lineal General  $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$ , donde

$$\begin{array}{c} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \\ n \times 1 \end{array} = \begin{array}{c} \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1p} \\ 1 & X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix} \\ n \times (p+1) \end{array} \begin{array}{c} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} \\ (p+1) \times 1 \end{array} + \begin{array}{c} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \\ n \times 1 \end{array}$$

Con ( $n \geq p$ ). Se asume para  $\varepsilon$  una distribución normal multivariada con vector de medias cero y una matriz de varianza-covarianza  $\Sigma$ . Por ejemplo:

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}, \text{Var}[\varepsilon] = \begin{bmatrix} \sigma_{\varepsilon_1}^2 & \text{Cov}(\varepsilon_1, \varepsilon_2) & \text{Cov}(\varepsilon_1, \varepsilon_3) \\ \text{Cov}(\varepsilon_2, \varepsilon_1) & \sigma_{\varepsilon_2}^2 & \text{Cov}(\varepsilon_2, \varepsilon_3) \\ \text{Cov}(\varepsilon_3, \varepsilon_1) & \text{Cov}(\varepsilon_3, \varepsilon_2) & \sigma_{\varepsilon_3}^2 \end{bmatrix}$$

Cuando  $\varepsilon \sim N(0, \sigma^2 \mathbf{I}_n)$  se asume que los  $\varepsilon_i$  tienen varianza común  $\sigma^2$  y que los  $\varepsilon_i$  son estadísticamente independientes.

Casos especiales. El Modelo Lineal General tiene los casos:

1. La distribución de  $\varepsilon$  y por lo tanto de  $\mathbf{Y}$ .
2. La estructura de la matriz de varianza-covarianza  $\Sigma$ .
3. El rango y la estructura de  $\mathbf{X}$ .

**Caso 1.**  $\varepsilon \sim N(0, \sigma^2 \mathbf{I}_n)$

Si  $\text{Var} [\varepsilon] = \sigma^2 \mathbf{I}_n$ ,  $\Rightarrow$  los  $\varepsilon_i$  tienen varianza común y son no correlacionados ( $\text{Cov} [\varepsilon_i, \varepsilon_j]$  es cero).

**Caso 2.** Se desconoce la distribución, pero se sabe que está centrada en cero y que los  $\varepsilon_i$  tienen varianza común.

$$\varepsilon \sim ?(0, \sigma^2 \mathbf{I}_n), E[\varepsilon] = 0, \text{Cov} [\varepsilon_i, \varepsilon_j] = \sigma^2 \mathbf{I}$$

En ambos casos:

$$\begin{bmatrix} \sigma^2 & 0 & 0 \dots 0 \\ 0 & \sigma^2 & 0 \dots 0 \\ 0 & 0 & 0 \dots 0 \end{bmatrix}$$

**Caso 3.**  $\text{Var}[\varepsilon] \neq \sigma^2 \mathbf{I}$ , los  $\varepsilon_i$  tienen varianzas diferentes y no son independientes. Si los  $\varepsilon_i$  fueran independientes (si todos los elementos fuera de la diagonal de la matriz son cero) se usa el método de mínimos cuadrados ponderados. Si hay covarianza entre errores, (no son independientes) se usa el método de mínimos cuadrados generalizado.

Para estimar los parámetros  $(\beta, \sigma^2)$ , se utilizan los métodos de máxima verosimilitud o bien el de mínimos cuadrados.

Para el Caso 1, se puede demostrar utilizando el método de máxima verosimilitud que:

$$E[\hat{\sigma}^2] = \frac{n - (p + 1)}{n} \sigma^2, \text{ por lo tanto un estimador insesgado de } \sigma^2 \text{ es:}$$

$$\hat{\sigma}^2 = \frac{(\mathbf{Y} - \mathbf{X}\hat{\beta})'(\mathbf{Y} - \mathbf{X}\hat{\beta})}{n} \cdot \frac{n}{n - (p + 1)}$$

siendo la varianza muestral  $S^2$

$$S^2 = \frac{\text{SC res}}{n - (p + 1)}$$

La notación matricial de  $\mathbf{X}'\mathbf{X}$  es:

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} n & \Sigma X_{i1} & \Sigma X_{i2} \cdots \Sigma X_{ip} \\ \Sigma X_{i1} & \Sigma X_{i1}^2 & \Sigma X_{i1}X_{i2} \cdots \Sigma X_{i1}X_{ip} \\ \dots & \dots & \dots \\ \Sigma X_{ip} & \Sigma X_{i1}X_{ip} \cdots \cdots & \Sigma X_{ip}^2 \end{bmatrix}$$

En la diagonal principal están la suma de cuadrados y fuera de ella las sumas de productos.

$$\mathbf{X}'\mathbf{Y} = \begin{bmatrix} \Sigma Y_i \\ \Sigma X_{i1}Y_i \\ \Sigma X_{i2}Y_i \\ \vdots \\ \Sigma X_{ip}Y_i \end{bmatrix}$$

Para el Caso 2, se utiliza el método de mínimos cuadrados:

$$\varepsilon \sim (0, \sigma^2 \mathbf{I})$$

$$\varepsilon'\varepsilon = (\mathbf{Y} - \mathbf{X}\hat{\beta})'(\mathbf{Y} - \mathbf{X}\hat{\beta}) = (\mathbf{Y}' - \hat{\beta}'\mathbf{X}')(\mathbf{Y} - \mathbf{X}\hat{\beta})$$

$$= \mathbf{Y}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}\hat{\beta} - \hat{\beta}'\mathbf{X}'\mathbf{Y} + \hat{\beta}'\mathbf{X}'\mathbf{X}\hat{\beta}$$

$$= \mathbf{Y}'\mathbf{Y} - 2\mathbf{Y}'\mathbf{X}\hat{\beta} + \hat{\beta}'\mathbf{X}'\mathbf{X}\hat{\beta}$$

$$= \mathbf{Y}'\mathbf{Y} - 2(\mathbf{X}'\mathbf{Y})' \beta + \beta' \mathbf{X}'\mathbf{X} \beta$$

Derivado respecto a  $\beta$ , e igualando a cero se llega a las ecuaciones normales:

$(\mathbf{X}'\mathbf{X})\beta = \mathbf{X}'\mathbf{Y}$  de las que se pueden obtener los estimadores

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y}$$

El vector  $\mathbf{Y}$ :

Si  $\varepsilon \sim N(0, \sigma^2 \mathbf{I}_n)$ , como en el GLM  $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$ , entonces el vector  $\mathbf{Y}$  tiene distribución normal multivariada  $\mathbf{Y} \sim N(\mathbf{X}\beta, \sigma^2 \mathbf{I}_n)$  por lo que

$$E[\mathbf{Y}] = E[\mathbf{X}\beta + \varepsilon] = E[\mathbf{X}\beta] + E[\varepsilon] = \mathbf{X}\beta$$

$$\text{Var}[\mathbf{Y}] = \text{Var}[\mathbf{X}\beta + \varepsilon] = \text{Var}[\varepsilon] = \sigma^2 \mathbf{I}_n$$

Esto es así si se asume que el modelo es correcto, en caso contrario,  $E[\mathbf{Y}] \neq \mathbf{X}\beta$

### **Propiedades de $\hat{\beta}$**

1. Es insesgado, puesto que  $E[\hat{\beta}] = \beta$
2.  $\text{Var}[\hat{\beta}] = \text{Var}[(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y}] = (\mathbf{X}'\mathbf{X})^{-1} \sigma^2$

La matriz  $(\mathbf{X}'\mathbf{X})^{-1}$  es simétrica: los elementos de la diagonal principal contiene las varianzas de los estimadores de los parámetros y los otros elementos las covarianzas.

## **Teorema de Gauss-Markoff**

En el Modelo Lineal General  $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$ , donde  $\mathbf{X}$  es una matriz de rango completo  $n \times (p+1)$ ,  $\beta$  un vector  $(p+1) \times 1$  de parámetros desconocidos,  $\varepsilon$  un vector aleatorio  $n \times 1$  con media cero y varianza  $\sigma^2 \mathbf{I}$  el estimador de mínimos cuadrados  $\hat{\beta}$  es el mejor estimador lineal insesgado de (media).

## **Propiedades de los Estimadores y Combinaciones Lineales de los Estimadores**

Estimando funciones de  $\beta_0, \beta_1, \dots, \beta_p$ :

Para un vector  $\mathbf{t}'_{1 \times (p+1)}$  de números reales, el meli de  $(\mathbf{t}'\beta)$  es  $\mathbf{t}'\hat{\beta}$  donde  $\hat{\beta}$  son los estimadores de mínimos cuadrados de  $\beta$ .

$$\hat{\mu}_{Y, X} = \hat{\mathbf{Y}}_i \quad (\text{estimador de la media})$$

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta} = [\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Y} = \mathbf{P}\mathbf{Y}, \mathbf{P}_{n \times n} \text{ es simétrica e idempotente}$$

$$E[\hat{\mathbf{Y}}] = E[\mathbf{P}\mathbf{Y}] = \mathbf{P}[E(\mathbf{Y})] = \mathbf{P}\mathbf{X}\beta = \mathbf{X}\beta$$

$$\text{Var}[\hat{\mathbf{Y}}] = \text{Var}[\mathbf{X}\hat{\beta}] = \mathbf{X}[\text{Var}(\hat{\beta})]\mathbf{X}' = \mathbf{X}[(\mathbf{X}'\mathbf{X})^{-1}\sigma^2]\mathbf{X}' = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\sigma^2 = \mathbf{P}\sigma^2$$

$\hat{\mathbf{Y}}_i$  es un estimador insesgado de las medias de  $\mathbf{Y}$ .

## **Propiedades de los Residuales**

$$\varepsilon = \mathbf{Y} - \hat{\mathbf{Y}} = \mathbf{Y} - \mathbf{P}\mathbf{Y} = (\mathbf{I} - \mathbf{P})\mathbf{Y}, (\mathbf{I} - \mathbf{P}) \text{ simétrica e idempotente, } \mathbf{P}\mathbf{X} = \mathbf{X},$$

$$E[\varepsilon] = E[(\mathbf{I} - \mathbf{P})\mathbf{Y}] = (\mathbf{I} - \mathbf{P})E[\mathbf{Y}] = (\mathbf{I} - \mathbf{P})\mathbf{X}\beta = (\mathbf{X} - \mathbf{P}\mathbf{X})\beta = (\mathbf{X} - \mathbf{X})\beta = \mathbf{0}$$

$$\text{Var}[\varepsilon] = \text{Var}[(\mathbf{I} - \mathbf{P})\mathbf{Y}] = (\mathbf{I} - \mathbf{P})\text{Var}[\mathbf{Y}](\mathbf{I} - \mathbf{P})' = (\mathbf{I} - \mathbf{P})\mathbf{I}\sigma^2(\mathbf{I} - \mathbf{P})'$$

$$= (\mathbf{I} - \mathbf{P})(\mathbf{I} - \mathbf{P})'\sigma^2 = (\mathbf{I} - \mathbf{P})\sigma^2$$

Los vectores aleatorios  $\mathbf{Y}, \hat{\mathbf{Y}}, \hat{\boldsymbol{\beta}}, \mathbf{e}$  tienen distribución normal multivariada.

### Varianza de una Función Lineal

Sea la función  $U$  tal que  $U = \sum_{i=1}^n a_i Y_i$  donde  $Y_i$  son variables aleatorias y  $a_i$  son constantes, entonces:

$$\text{Var}[U] = \sum_{i=1}^n a_i^2 \text{Var}[Y_i] + \sum_i \sum_j a_i a_j \text{Cov}[Y_i, Y_j], \quad i \neq j$$

Si  $Y_i$  son independientes y  $\text{Var}[Y_i] = \sigma^2$ ,  $\Rightarrow \text{Var}[U] = (\sum a_i^2) \sigma^2$ .

En notación matricial:

Si  $\mathbf{U} = \mathbf{A}\mathbf{Y}$ , donde  $\mathbf{Y}$  es un vector aleatorio,  $\mathbf{A}$  es una matriz tal que cada hilera produce los coeficientes para una función lineal, entonces:

$$\text{Var}[\mathbf{U}] = \mathbf{A}[\text{Var}(\mathbf{Y})] \mathbf{A}'$$

$$\text{Si } \text{Var}[\mathbf{Y}] = \sigma^2 \mathbf{I}_n, \Rightarrow \text{Var}[\mathbf{U}] = \mathbf{A} \mathbf{A}' \sigma^2$$

### Prueba de Hipótesis

La forma general de las hipótesis lineales es  $H_0: \mathbf{K}\boldsymbol{\beta} = \mathbf{m}$ ,  $H_a: \mathbf{K}\boldsymbol{\beta} \neq \mathbf{m}$ , donde  $\mathbf{K}$  es una matriz de coeficientes que definen  $k$  funciones lineales de  $\boldsymbol{\beta}$  para probarse y  $\mathbf{m}$  es un vector  $k \times 1$  con elementos usualmente ceros.

En la hipótesis simple,  $\mathbf{K}$  tiene una fila, en la compuesta tiene más de una. Por ejemplo, sea  $H_0: \mathbf{K}\boldsymbol{\beta} = \mathbf{m}$

$$\mathbf{H}_0 : \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & -2 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \beta_1 - \beta_2 \\ \beta_1 + \beta_2 - 2\beta_3 \\ \beta_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Aquí, la hipótesis es compuesta,  $\mathbf{K}$  tiene más de una fila. Las  $k$  funciones lineales en  $H_0$  deben ser linealmente independientes (no necesariamente ortogonales), por lo que  $r(\mathbf{K}) = k$  y el número de funciones lineales en  $H_0$  no es mayor al número de parámetros, ya que  $\mathbf{K}$  no sería de rango completo.

Como  $\hat{\beta} \sim N(\beta, (\mathbf{X}'\mathbf{X})^{-1}\sigma^2)$ ,  $\mathbf{K}\hat{\beta} - \mathbf{m}$  son combinación lineal de  $\hat{\beta}$ ,

$$\Rightarrow (\mathbf{K}\hat{\beta} - \mathbf{m}) \sim N(\mathbf{K}\beta - \mathbf{m}, \mathbf{K}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K}'\sigma^2)$$

La suma de cuadrados de la hipótesis lineal, se calcula como forma cuadrática:

$$\mathbf{Q} = (\mathbf{K}\hat{\beta} - \mathbf{m})' [\mathbf{K}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K}']^{-1} (\mathbf{K}\hat{\beta} - \mathbf{m}),$$

$$\mathbf{Q} / \sigma^2 = (\mathbf{K}\hat{\beta} - \mathbf{m})' [\mathbf{K}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K}']^{-1} (\mathbf{K}\hat{\beta} - \mathbf{m}) / \sigma^2 \sim \chi^2_{(k, \lambda)}, \text{ donde}$$

$$\lambda = \frac{1}{2\sigma^2} (\mathbf{K}\beta - \mathbf{m})' [\mathbf{K}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K}']^{-1} (\mathbf{K}\beta - \mathbf{m}), \quad \Rightarrow \text{la } H_0 \text{ es cierta:}$$

$$\frac{\mathbf{Q} / r(\mathbf{K})\sigma^2}{\text{SC res} / [n - (p - 1)]\sigma^2} \sim F_{[k, n - (p - 1)]}$$

## Modelo Reducido

Una forma de calcular  $\mathbf{Q}$  es utilizando el modelo reducido. Se verá un ejemplo:

Sea  $H_0: \beta_1 = \beta_2$  y  $\beta_0 = 20$ , entonces

$$\mathbf{K} = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \mathbf{m} = \begin{bmatrix} 0 \\ 20 \end{bmatrix}, r(\mathbf{K}) = 2$$

$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i$ , con gl (res) =  $n-4$  (modelo completo)

$Y_i = 20 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i$ , (modelo reducido)

$Y_i - 20 = \beta_1(X_{i1} + X_{i2}) + \beta_3 X_{i3} + \varepsilon_i$

$Y^*_i = \beta_1 X^*_{i1} + \beta_3 X_{i3} + \varepsilon_i$

en forma general, el modelo reducido se puede expresar como

$\mathbf{Y}^* = \mathbf{X}^* \beta^* + \varepsilon$ , donde

$$\mathbf{Y}^* = \begin{bmatrix} Y_1 - 20 \\ Y_2 - 20 \\ \vdots \\ Y_n - 20 \end{bmatrix}, \mathbf{X}^* = \begin{bmatrix} (X_{11} + X_{12}) & X_{13} \\ (X_{21} + X_{22}) & X_{23} \\ \vdots & \vdots \\ (X_{n1} + X_{n2}) & X_{n3} \end{bmatrix} = \begin{bmatrix} X^*_{11} & X_{13} \\ X^*_{21} & X_{23} \\ \vdots & \vdots \\ X^*_{n1} & X_{n3} \end{bmatrix}, \beta^* = \begin{bmatrix} \beta^*_1 \\ \beta_3 \end{bmatrix}$$

$r(\mathbf{X}^*) = 2$ , SC (res red) tiene  $(n-2)$  gl, entonces:

$\mathbf{Q} = \text{SC (res red)} - \text{SC (res compl)}$

$gl = (n-2) - (n-4) = 2$

Si  $\beta_0$  no está incluido en  $H_0$ , una forma alternativa de obtener  $\mathbf{Q}$  es:

$$Q = \text{SC med compl} - \text{SC med red}$$

### Modelo de Rango Incompleto

Este modelo se representa igual que el modelo de rango completo, pero un número infinito de conjuntos de escalares  $\beta_0, \beta_1, \dots, \beta_p$  describen el sistema.  $\mathbf{X}$  es una matriz  $n \times (p+1)$ , de rango  $r < p$ ,  $\mathbf{X}'\mathbf{X}$  es singular y no tiene inversa ordinaria, por lo que el sistema  $(\mathbf{X}'\mathbf{X})\beta = \mathbf{X}'\mathbf{Y}$  tiene una infinidad de soluciones.

El modelo de rango incompleto, se puede escribir:

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad i = 1, 2, \dots, k; \quad j = 1, 2, \dots, n$$

En notación matricial:  $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$ , donde

$\mathbf{Y}$  es un vector de dimensiones  $\left( \sum_{i=1}^k n_i \right) \times 1$

$\beta$  es un vector de parámetros  $\beta' = [\mu \tau_1 \tau_2 \dots \tau_R]$

$\mathbf{X}$  es la matriz de diseño  $\left( \sum_{i=1}^k n_i \right) \times (k+1)$ ,  $r(\mathbf{X}) < (k+1)$ ; además  $r(\mathbf{X}) = r(\mathbf{X}'\mathbf{X})$ .

### Estimabilidad

Sea el modelo de rango incompleto  $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$ , donde  $r(\mathbf{X}) = r < (p+1)$ ,  $E[\varepsilon] = 0$ ,  $\text{Var}[\varepsilon] = \sigma^2 \mathbf{I}$ , entonces  $\mathbf{t}'\beta$  es estimable si existe un vector  $\mathbf{c}$  tal

que  $E[\mathbf{c}'\mathbf{Y}] = \mathbf{t}'\beta$ . Esto es, existe una combinación lineal de las observaciones cuya esperanza es  $\mathbf{t}'\beta$ . Estas especificaciones definen la estimabilidad.

**Definición 1.** Sea una matriz  $\mathbf{A}$  de rango incompleto,  $r(\mathbf{A}) = r$ , la inversa generalizada de  $\mathbf{A}$  es una matriz  $\mathbf{A}^G$  tal que  $\mathbf{A}\mathbf{A}^G\mathbf{A} = \mathbf{A}$ , la matriz  $\mathbf{A}^G$  no es única. Si  $\mathbf{A}$  tiene rango completo,  $\mathbf{A}^G = \mathbf{A}^{-1}$  es única.

**Definición 2.** Sea el modelo de rango incompleto  $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$ , donde  $r(\mathbf{X}) = r < (p+1)$ ,  $E[\varepsilon] = 0$ ,  $\text{Var}[\varepsilon] = \sigma^2\mathbf{I}$ ,  $\Rightarrow \mathbf{t}'\beta$  es estimable  $\Leftrightarrow \mathbf{t}'(\mathbf{X}'\mathbf{X})^G(\mathbf{X}'\mathbf{X}) = \mathbf{t}'$ .

### Teorema de Gauss-Markoff

Sea  $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$ , donde  $\mathbf{X}$  es  $n \times (p+1)$ , de rango  $r < (p+1)$ ,  $E[\varepsilon] = 0$ ,  $\text{Var}[\varepsilon] = \sigma^2\mathbf{I}$ , y  $\mathbf{t}'\beta$  sea estimable, entonces cualquier solución  $\beta$  de  $(\mathbf{X}'\mathbf{X})\beta = \mathbf{X}'\mathbf{Y}$  produce el mismo (único) estimador lineal insesgado de  $\mathbf{t}'\beta$  (meli).

### Funciones Estimables y no Estimables

$\beta_0, \beta_1, \dots, \beta_p$  no son estimables. Los elementos de  $\mathbf{X}\beta$  son estimables, ya que:  $E[\mathbf{Y}] = \mathbf{X}\beta$ . En particular, son funciones estimables del tipo  $\mu + \tau_i, \mu_i - \mu_j, \tau_i - \tau_j$ , para  $i \neq j$ . Toda combinación lineal de funciones estimables es estimable. Si  $\mathbf{t}_1\beta, \mathbf{t}_2\beta, \dots, \mathbf{t}_m\beta$ , son estimables, entonces  $\mathbf{z} = \mathbf{a}_1\mathbf{t}_1\beta + \mathbf{a}_2\mathbf{t}_2\beta + \dots + \mathbf{a}_m\mathbf{t}_m\beta$  es estimable

## Estimación de $\sigma^2$ en el Modelo de Rango Completo

En el modelo de rango completo el estimador insesgado de la varianza poblacional es

$$S^2 = \frac{SCres}{n - (p + 1)} \text{ donde } n \text{ es el tamaño de la muestra y } (p + 1) \text{ es } r(\mathbf{X}).$$

En el modelo de rango incompleto,  $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$ ,  $E[\varepsilon] = 0$ ,  $\text{Var}[\varepsilon] = \sigma^2 \mathbf{I}$ ,  $r(\mathbf{X}) = r < (p + 1)$ , por lo que el estimador insesgado de  $\sigma^2$  será  $S^2 = \frac{SCres}{n - r}$ .

## Reparametrización

El objeto de la reparametrización es pasar de un modelo de rango incompleto a uno de rango completo, redefiniendo los parámetros por medio de restricciones lineales impuestas a los estimadores. Se presentan tres tipos de reparametrizaciones.

**Caso 1.** El modelo sobreparametrizado  $Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$ ,  $i = 1, 2, \dots, p$ ;  $j = 1, 2, \dots, n$  con  $(p + 1)$  parámetros. Se define la reparametrización  $\mu_i = \mu + \tau_i$ , con lo que el modelo reparametrizado será  $Y_{ij} = \mu_i + \varepsilon_{ij}$ , y en forma matricial  $\mathbf{Y} = \mathbf{X}^* \beta^* + \varepsilon$ . Sea por ejemplo:

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \beta = \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \end{bmatrix}$$

En esta matriz  $\mathbf{X}$ , la suma de las últimas cuatro columnas producen la primera. El modelo reparametrizado tiene la siguiente matriz de diseño y vector de parámetros:

$$\mathbf{X}^* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \beta^* = \begin{bmatrix} \mu_1^* \\ \mu_2^* \\ \mu_3^* \\ \mu_4^* \end{bmatrix}$$

Con ecuaciones normales  $\mathbf{X}^* \mathbf{X}^* \beta^* = \mathbf{X}^* \mathbf{Y}$

**Caso 2.** Sea la restricción  $\sum \hat{\tau}_i = 0$  que implica  $\tau_p = -(\tau_1 + \dots + \tau_{p-1})$

El vector de parámetros y la matriz de diseño son:

$$\beta^* = \begin{bmatrix} \mu^* \\ \tau_1^* \\ \tau_2^* \\ \tau_3^* \end{bmatrix}, \quad \mathbf{X}^* = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{X}^* \mathbf{X}^* = \begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 4 & 2 & 2 \\ 0 & 2 & 4 & 2 \\ 0 & 2 & 2 & 4 \end{bmatrix}$$

$E[\beta^*] = [(\mathbf{X}^* \mathbf{X}^*)^{-1} \mathbf{X}^* \mathbf{X}] \beta$ , siendo

$$E[\beta^*] = \begin{bmatrix} 1 & 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 3/4 & -1/4 & -1/4 & -1/4 \\ 0 & -1/4 & 3/4 & -1/4 & -1/4 \\ 0 & -1/4 & -1/4 & 3/4 & -1/4 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_4 \end{bmatrix} = \begin{bmatrix} \mu + \tau \\ \tau_1 + \tau \\ \tau_2 + \tau \\ \tau_3 + \tau \end{bmatrix}$$

donde  $\bar{\tau} = \frac{\tau_1 + \tau_2 + \tau_3 + \tau_4}{4}$ ,  $\beta^* = (\mu^*, \tau_1^*, \tau_2^*, \tau_3^*)$  es un estimador insesgado de  $(\tau_4 - \bar{\tau})$

**Caso 3.** Sea la restricción  $\tau_4 = 0$ . Se genera igual resultado que utilizando la G inversa de SAS. En el ejemplo utilizado, el vector de parámetros y la matriz de diseño son:

$$\beta^* = \begin{bmatrix} \mu^* \\ \tau_1^* \\ \tau_2^* \\ \tau_3^* \end{bmatrix}, \quad X^* = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$$E[\beta^*] = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix} = \begin{bmatrix} \mu + \tau_4 \\ \tau_1 - \tau_4 \\ \tau_2 - \tau_4 \\ \tau_3 - \tau_4 \end{bmatrix}$$

Así que  $\beta^* = (\mu^*, \tau_1^*, \tau_2^*, \tau_3^*)$  estima a  $\begin{bmatrix} \mu + \tau_4 \\ \tau_1 - \tau_4 \\ \tau_2 - \tau_4 \\ \tau_3 - \tau_4 \end{bmatrix}$

## Modelos Básicos del Diseño Experimental, ANVA en una Vía

### Modelo Completo

El modelo completo  $Y_{ij} = \mu_i + \varepsilon_i$ ,  $\mathbf{Y} = \mathbf{Z}\alpha + \varepsilon$ ,  $\mathbf{Z}$  es  $n \times k$ ,  $\alpha$  un vector de parámetros  $k+1$ , tal que  $\alpha^1 = (\mu_1, \mu_2, \dots, \mu_k)$ , la suma de cuadrados del modelo completo es  $\mathbf{Y}'(\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}')\mathbf{Y}$ , donde

$$\mathbf{Z}'\mathbf{Z} = \begin{bmatrix} n_1 & 0 & \dots & 0 \\ 0 & n_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & n_k \end{bmatrix}, \quad (\mathbf{Z}'\mathbf{Z})^{-1} = \begin{bmatrix} \frac{1}{n_1} & 0 & \dots & 0 \\ 0 & \frac{1}{n_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{1}{n_k} \end{bmatrix}$$

entonces, SC (modelo completo) =  $\sum_{i=1}^k \frac{Y_i^2}{n_i}$

### Modelo Reducido

Se asume que  $H_0: \mu_1 = \mu_2 = \dots = \mu_k = \mu$  es verdadera, el modelo es  $Y_{ij} = \mu + \varepsilon_{ij}$ , en forma matricial  $\mathbf{Y} = \mathbf{Z}_2 \alpha_2 + \varepsilon$ , donde  $\mathbf{Z}'_2 = [1 \ 1 \ \dots \ 1]$  y  $\alpha_2 = \mu$ , la suma de cuadrados del modelo reducido es

$$\mathbf{Y}'[\mathbf{Z}_2(\mathbf{Z}'_2\mathbf{Z}_2)^{-1}\mathbf{Z}'_2]\mathbf{Y} = \frac{Y_{..}^2}{n},$$

Así la suma de cuadrados de la hipótesis es:

$$\text{SC (hipót)} = \text{SC (mod compl)} - \text{SC (mod red)}$$

$$= \mathbf{Y}'[\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}' - \mathbf{Z}_2(\mathbf{Z}'_2\mathbf{Z}_2)^{-1}\mathbf{Z}'_2]\mathbf{Y} = \sum_{i=1}^k \frac{Y_i^2}{n_i} - \frac{Y_{..}^2}{n}$$

Esta es la variabilidad en Y, que no es aleatoria (no explicada).

La  $H_0$  se prueba con:

$$\frac{[\text{SC(mod. compl)} - \text{SC(mod. red)}]/(k-1)}{\text{SC(res mod compl)}/(n-k)}$$

El Teorema de Cochran-Fisher afirma que estas sumas de cuadrados tienen distribución  $\chi^2$  y son independientes.

Cuando  $H_0: \mu_1 = \mu_2 = \dots = \mu_k = \mu$  es verdadera.

$$\lambda = \frac{1}{\partial \sigma^2} (\mathbf{Z}\alpha)^{-1} \left( \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{Z}' - \mathbf{Z}_2(\mathbf{Z}_2'\mathbf{Z}_2)^{-1} \mathbf{Z}_2' \right) (\mathbf{Z}\alpha) = 0$$

entonces 
$$\frac{\left[ \sum_{i=1}^k \frac{Y_i^2}{n_i} - \frac{Y_{..}^2}{n} \right] / (k-1)}{\text{SC(res modelo compl)}/(n-k)} \sim F_{[(k-1), (n-k)]}$$

La tabla del ANVA es

| FV                   | gl  | SC   |
|----------------------|-----|--|
| Modelo completo      | k   | $Y'[Z(Z'Z)^{-1}Z']Y$                         |
| Modelo reducido      | 1   | $Y'[Z_2(Z_2'Z_2)^{-1}Z_2']Y$                 |
| Hipótesis            | k-1 | $Y'[Z(Z'Z)^{-1}Z' - Z_2(Z_2'Z_2)^{-1}Z_2']Y$ |
| Residual (mod. Comp) | n-k | $Y'[I - Z(Z'Z)^{-1}Z']Y$                     |
| Total (no corr)      | n   | $Y'Y$  |

### Análisis de Dos Factores sin Interacción (Modelo Aditivo)

Sea el factor A con a niveles, el factor B con b niveles y se efectúa una observación en cada combinación ab.

El modelo propuesto es

$$Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}, \quad i = 1, 2, \dots, a; \quad j = 1, 2, \dots, b$$

donde  $\tau_i$  es el efecto debido al i-ésimo nivel del factor A,

$\beta_j$  es el efecto del j-ésimo nivel del factor B,

$\varepsilon_{ij}$  es el efecto del error aleatorio.

En forma matricial el modelo es

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}.$$

Se asume que  $\boldsymbol{\varepsilon} \sim N(0, \sigma^2\mathbf{I})$ . La matriz de diseño es

$$\mathbf{X} = \begin{bmatrix} \mu & \tau_1 & \tau_2 & \dots & \tau_a & \beta_1 & \beta_2 & \dots & \beta_b \\ 1 & 1 & 0 \dots 0 & & 0 & 1 & 0 & \dots & 0 \\ 1 & 1 & 0 \dots 0 & & 0 & 0 & 1 & \dots & 0 \\ \dots & \dots \\ 1 & 1 & 0 \dots 0 & & 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 1 \dots 0 & & 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 1 \dots 0 & & 0 & 1 & 0 & \dots & 0 \\ \dots & \dots \\ 1 & 0 & 1 \dots 0 & & 0 & 0 & 0 & \dots & 1 \\ \dots & \dots \\ 1 & 0 & 0 \dots 1 & & 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 \dots 1 & & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots \\ 1 & 0 & 0 \dots 1 & & 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

Tiene dos dependencias lineales,  $r(\mathbf{X}) = (a+b+1)-2 = a+b-1$  (rango incompleto).

### Hipótesis

Tienen la forma  $H_0: \mathbf{K}\beta = \mathbf{0}$ ,  $r(\mathbf{K}) \leq a+b-1$ . Las hipótesis más importantes son  $H_0: \tau_1 = \tau_2 = \dots = \tau_a$ ,  $H_0: \beta_1 = \beta_2 = \dots = \beta_b$ .

Éstas son contrastes entre  $\tau$ 's y  $\beta$ 's si son estimables; en este caso  $H_0: \tau_1 = \tau_2 = \dots = \tau_a$  equivale al siguiente sistema de contrastes ortogonales linealmente independiente.

$$\tau_1 - \tau_2 = 0$$

$$\tau_1 + \tau_2 - 2\tau_3 = 0$$

$$\tau_1 + \tau_2 + \tau_3 - 3\tau_4 = 0$$

·  
·  
·

$$\tau_1 + \tau_2 + \dots - (a-1)\tau_a = 0$$

y entonces

$$\mathbf{K} = \begin{bmatrix} \mu & \tau_1 & \tau_2 & \tau_3 & \dots & \tau_a & \beta_1 & \beta_2 & \dots & \beta_b \\ 0 & 1 & -1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 1 & -2 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots \\ 0 & 1 & 1 & 1 & \dots & -(a-1) & 0 & 0 & \dots & 0 \end{bmatrix}$$

Individualmente los parámetros no son estimables, pero sí son estimables sus diferencias.

## Análisis de Dos Factores con Interacción

Considerar el caso en que ambos factores se aplican a dos niveles en el experimento, y suponer los siguientes datos:

|          |   | Factor A |       |  |
|----------|---|----------|-------|--|
| Factor B | 1 | 2        | Total |  |
| 1        | 3 | 6        | 9     |  |
| 2        | 6 | 3        | 9     |  |
| Total    | 9 | 9        | 18    |  |

Hay interacción ya que hay un comportamiento diferencial de A a los distintos niveles de B. El modelo es:

$$Y_{ijk} = \mu + \tau_i + \beta_j + \delta_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, \dots, a; \quad j = 1, 2, \dots, b; \quad k = 1, 2, \dots, n$$

donde n es el número de observaciones en cada celda.

Las hipótesis en prueba son:

$H_0$ : no hay interacción

$H_0$ :  $\tau_1 = \tau_2 = \dots = \tau_a$ ,

$H_0$ :  $\beta_1 = \beta_2 = \dots = \beta_b$

Suponiendo dos observaciones por celda:

|          |           | Factor B  |  |
|----------|-----------|-----------|--|
| Factor A | 1         | 2         |  |
| 1        | $Y_{111}$ | $Y_{121}$ |  |
|          | $Y_{112}$ | $Y_{122}$ |  |
| 2        | $Y_{211}$ | $Y_{221}$ |  |
|          | $Y_{212}$ | $Y_{222}$ |  |

En forma matricial:

$$\begin{bmatrix} Y_{111} \\ Y_{112} \\ Y_{121} \\ Y_{122} \\ Y_{211} \\ Y_{212} \\ Y_{221} \\ Y_{222} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \beta \\ \beta_{12} \\ \delta_{11} \\ \delta_{12} \\ \delta_{21} \\ \delta_{22} \end{bmatrix} + \begin{bmatrix} \varepsilon_{111} \\ \varepsilon_{112} \\ \varepsilon_{121} \\ \varepsilon_{122} \\ \varepsilon_{211} \\ \varepsilon_{212} \\ \varepsilon_{221} \\ \varepsilon_{222} \end{bmatrix}$$

$r(\mathbf{X}) = 4$  cada columna de  $\mathbf{X}$  puede expresarse como una combinación lineal de  $\delta_{11}, \delta_{12}, \delta_{21}, \delta_{22}$ .

$$\mathbf{K} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \text{ la } H_0: \mathbf{K}\beta = 0 \text{ no es estimable.}$$

Cuando el número de niveles crece, el sistema de ecuaciones para expresar la idea de no interacción se hace compleja; por esto, conviene reparametrizar el modelo. Esta reparametrización es tal, que los términos de la interacción reparametrizada sean estimables.

$$Y_{ijk} = \mu + \tau_i + \beta_j + \delta_{ij} + \varepsilon_{ijk}, \quad i = 1, 2; j = 1, 2; k = 1, 2$$

$$\begin{aligned} \text{Sea} \quad \bar{\tau} &= \sum \frac{\tau_i}{2} & \bar{\beta} &= \sum \frac{\beta_j}{2} \\ \bar{\mu} &= \sum \sum \frac{\bar{\mu}_{ij}}{4} & \bar{\delta} &= \sum_{i=1}^2 \sum_{j=1}^2 \frac{\delta_{ij}}{4} \\ \bar{\delta}_i &= \sum_{j=1}^2 \frac{\delta_{ij}}{2} & \bar{\delta}_j &= \sum_{i=1}^2 \frac{\delta_{ij}}{2} \end{aligned}$$

Entonces:

$$\begin{aligned} \mu^* &= \mu + \bar{\tau} + \bar{\beta} + \bar{\delta} \\ \bar{\tau}_i^* &= \tau_i - \bar{\tau} + \bar{\delta}_i - \bar{\delta} \\ \beta^* &= \beta_j - \bar{\beta} + \bar{\delta}_j - \bar{\delta} \\ \bar{\delta}_{ij}^* &= \delta_{ij} - \bar{\delta}_i - \bar{\delta}_j + \bar{\delta} \end{aligned}$$

Así, el modelo puede expresarse como:

$$Y_{ijk} = \mu^* + \tau_i^* + \beta_j^* + \delta_{ij}^* + \varepsilon_{ijk}$$

Cada uno de estos parámetros es estimable y no hay interacción si  $\delta_{ij}^* = 0$  para cada  $i, j$ .

Definiendo  $\tau_i^*$ ,  $\beta_j^*$ ,  $\delta_{ij}^*$  como se ha hecho, implícitamente se han

definido las siguientes restricciones:

$$\begin{aligned} \sum_{j=1}^a \tau_i^* &= 0 & \sum_{i=1}^a \delta_{i.}^* &= 0 \\ \sum_{j=1}^b \beta_j^* &= 0 & \sum_{j=1}^b \delta_{.j}^* &= 0 \\ \sum_{i=1}^a \sum_{j=1}^b \delta_{ij}^* &= 0 \end{aligned}$$

que facilitan la solución de las ecuaciones normales del modelo reparametrizado.

### 3. EL MODELO LINEAL MIXTO

#### Generalidades

El modelo lineal aditivo más simple (Steel y Torrie, 1980) es

$$Y_i = \mu + \varepsilon_i$$

donde  $Y_i$  es la  $i$ -ésima observación de la media  $\mu$ , sujeta a un error de muestreo  $\varepsilon_i$ . Se supone que los  $\varepsilon_i$  pertenecen a la población de los  $\varepsilon$  con media cero, siendo independientes (no existe correlación). El modelo indica que existe una muestra finita de los  $\varepsilon$ , y estos llevan a un estimador de  $\sigma^2$ .

Para una adecuada evaluación de los datos experimentales y dentro del contexto del modelo lineal aditivo, debe establecerse el modelo de manera específica. Tres modelos lineales aditivos son: el modelo de efectos fijos o modelo I (Eisenhart, 1947), el modelo de efectos aleatorios o modelo II y el modelo de efectos mixtos o modelo III.

El modelo de efectos fijos. Para la clasificación de una vía, se tiene:

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad i = 1, 2, \dots, t \quad j = 1, 2, \dots, r_i$$

donde

Los  $\varepsilon_{ij}$  son componentes aleatorias, se deben hacer supuestos respecto a los  $\tau_i$ :

$$\text{Los } \tau_i \text{ son fijos y } \sum \tau_i = 0$$

forman una población finita y son parámetros de interés, además de  $\sigma^2$ . Para este modelo, una repetición del experimento generará el mismo conjunto de  $\tau$  en el nuevo experimento, donde se hacen inferencias acerca de los tratamientos particulares. Hacer  $\sum \tau_i = 0$ , es medir los efectos de los tratamientos como desviación respecto a una media general. Para el modelo fijo con efectos de interacción, estos serán fijos, ya que los efectos de tratamientos y bloques son fijos; se tratará de probar hipótesis sobre combinaciones lineales de los parámetros de interés, es decir, se tiene bajo estudio a toda la población de tratamientos. Aquí, los niveles de los factores no han sido seleccionados aleatoriamente de una población.

En el Modelo de Efectos Aleatorios, (Cuadro 3.1) los tratamientos son una muestra aleatoria de una población de  $\tau$  para la cual la media es cero y la varianza es  $\sigma^2_\tau$ , el parámetro de interés, además de  $\sigma^2$ . Para este modelo, una repetición producirá un nuevo conjunto de tratamientos de la misma población de  $\tau$  y, con el tiempo, la población completa de las  $\tau$  aparecerá en la posición  $i$ -ésima; interesa la variabilidad de las  $\tau$ . Para

un experimento al cual se le aplica el modelo de efectos aleatorios, la variabilidad del estimador de  $\sigma^2_{\tau}$ , proviene de la variabilidad en el estimador de  $\sigma^2$  y de las diferencias de los  $\tau$  de muestra a muestra.

Un ejemplo de un experimento en que se aplique un modelo de efectos fijos es la investigación de la eficiencia de los nutrientes N (nitrógeno), P (fósforo) y K (potasio) en producción de plantas, tal modelo podría ser:

$$y_{ij} = \mu + \alpha_i + e_{ij}$$

para clasificación de una vía. Y un ejemplo de un experimento típico donde halle aplicación el modelo de efectos aleatorios podría ser el estudio de la habilidad maternal de camadas de animales. El modelo apropiado con clasificación de una vía es:

$$y_{ij} = \mu + \beta_i + e_{ij}$$

con el término  $\beta_i$  que representan los datos relacionados con la habilidad maternal, variable sujeta a variaciones biológicas de animal a animal. En contraste con el modelo I, el modelo II, contiene factores en los cuales los niveles se seleccionan en forma aleatoria de entre una población.

Searle *et al.* (1992) designa las siguientes características a estos dos modelos:

Cuadro 3.1. Características del modelo de efectos fijos y del modelo de efectos aleatorios (clasificación: una vía).

| Característica            | Modelo de efectos fijos  | Modelo de efectos aleatorios  |
|---------------------------|--|---|
| Ecuación del modelo       | $y_{ij} = \mu + \alpha_i + e_{ij}$   | $y_{ij} = \mu + \alpha_i + e_{ij}$  |
| Media de $y_{ij}$         | $E(y_{ij}) = \mu + \alpha_i$   | $E(y_{ij}   \alpha_i) = \mu + \alpha_i$<br>$E(y_{ij}) = \mu$  |
| El término $\alpha_i$     | Fijo (constante desconocida)   | $\alpha \sim \text{i.i.d.}(0, \sigma^2_\alpha)$   |
| El término $e_{ij}$       | $e_{ij} = y_{ij} - E(y_{ij})$<br>$= y_{ij} - (\mu + \alpha_i)$<br>$e_{ij} \sim \text{i.i.d.}(0, \sigma^2_e)$         | $e_{ij} = y_{ij} - E(y_{ij}   \alpha_i)$<br>$= y_{ij} - (\mu + \alpha_i)$<br>$e_{ij} \sim \text{i.i.d.}(0, \sigma^2_e)$   |
| $E(e_{ij}   \alpha_i)$    | $E(e_{ij}   \alpha_i) = \alpha_i E(e_{ij}) = 0$  | $E(e_{ij}   \alpha_i) = 0$  |
| Var ( $y_{ij}$ )          | Var ( $y_{ij}$ ) = $\sigma^2_e$  | Var ( $y_{ij}$ ) = $\sigma^2_\alpha + \sigma^2_e$   |
| Cov( $y_{ij}, y_{i'j'}$ ) | Cov( $y_{ij}, y_{i'j'}$ )<br>$= \begin{cases} \sigma^2_e & \text{para } i=i', j=j' \\ 0 & \text{d.o.f.} \end{cases}$ | Cov( $y_{ij}, y_{i'j'}$ )<br>$= \begin{cases} \sigma^2_\alpha + \sigma^2_e & \text{para } i=i', j=j' \\ \sigma^2_\alpha & \text{para } i=i', j \neq j' \\ 0 & \text{d.o.f} \end{cases}$ |

La decisión para modelos de un factor si los niveles de los tratamientos son fijos o aleatorios (Cuadro 3.2) afectará las suposiciones acerca del término  $\tau_j$  en el modelo

$$y_{ij} = \mu + \tau_j + e_{ij}$$

estas suposiciones difieren según el modelo. Hicks (1964) establece las siguientes comparaciones;

Cuadro 3.2. Comparaciones entre los modelos fijo y aleatorio.

| Modelo fijo   |        |  | Modelo Aleatorio  |        |                               |
|---|--------|--|---|--------|-------------------------------|
| 1. Suposiciones: $\tau_j$ son constantes fijas                        |        |  | Suposiciones: $\tau_j$ son variables aleatorias y están NID(0, $\sigma^2_\tau$ ). $\sigma^2_\tau$ representa la varianza entre las $\tau_j$ o entre las verdaderas medias de los tratamientos, $\mu_j$ 's. El promedio de las $\tau_j$ es cero para todos los niveles posibles, pero para los k niveles usualmente no promedian cero. |        |                               |
| $\sum_{j=1}^k \tau_j = \sum_{j=1}^k (\mu_j - \mu) = 0$                |        |  | Análisis: igual que para el modelo fijo   |        |                               |
| (sólo las medias son consideradas)                                    |        |  | ECM   |        |                               |
| 2. Análisis: los procedimientos son los usuales para el cálculo de SS |        |  |   |        |                               |
| 3. ECM  |        |  |   |        |                               |
| FV  | gl     | ECM  | FV  | gl     | ECM                           |
| $\tau_j$  | k-1    | $\sigma_e^2 + n \frac{\sum \tau_j^2}{k-1}$ | $\tau_j$  | k-1    | $\sigma_e^2 + n\sigma_\tau^2$ |
| $\varepsilon_{ij}$  | k(n-1) | $\sigma_e^2$                               | $\varepsilon_{ij}$  | k(n-1) | $\sigma_e^2$                  |
| 4. Prueba de hipótesis<br>$H_0: \tau_j = 0$ , para toda j             |        |  | Prueba de hipótesis<br>$H_0: \sigma_\tau^2 = 0$   |        |                               |

### Modelo Mixto

Este modelo (modelo III), es el que contiene ambos efectos: fijos y aleatorios (Cuadro 3.3). En rigor, cualquier modelo que contiene un efecto fijo  $\mu$  y un término aleatorio  $e$ , es un modelo mixto. En el modelo mixto, lo más frecuente es que los bloques suministren los efectos aleatorios. Los

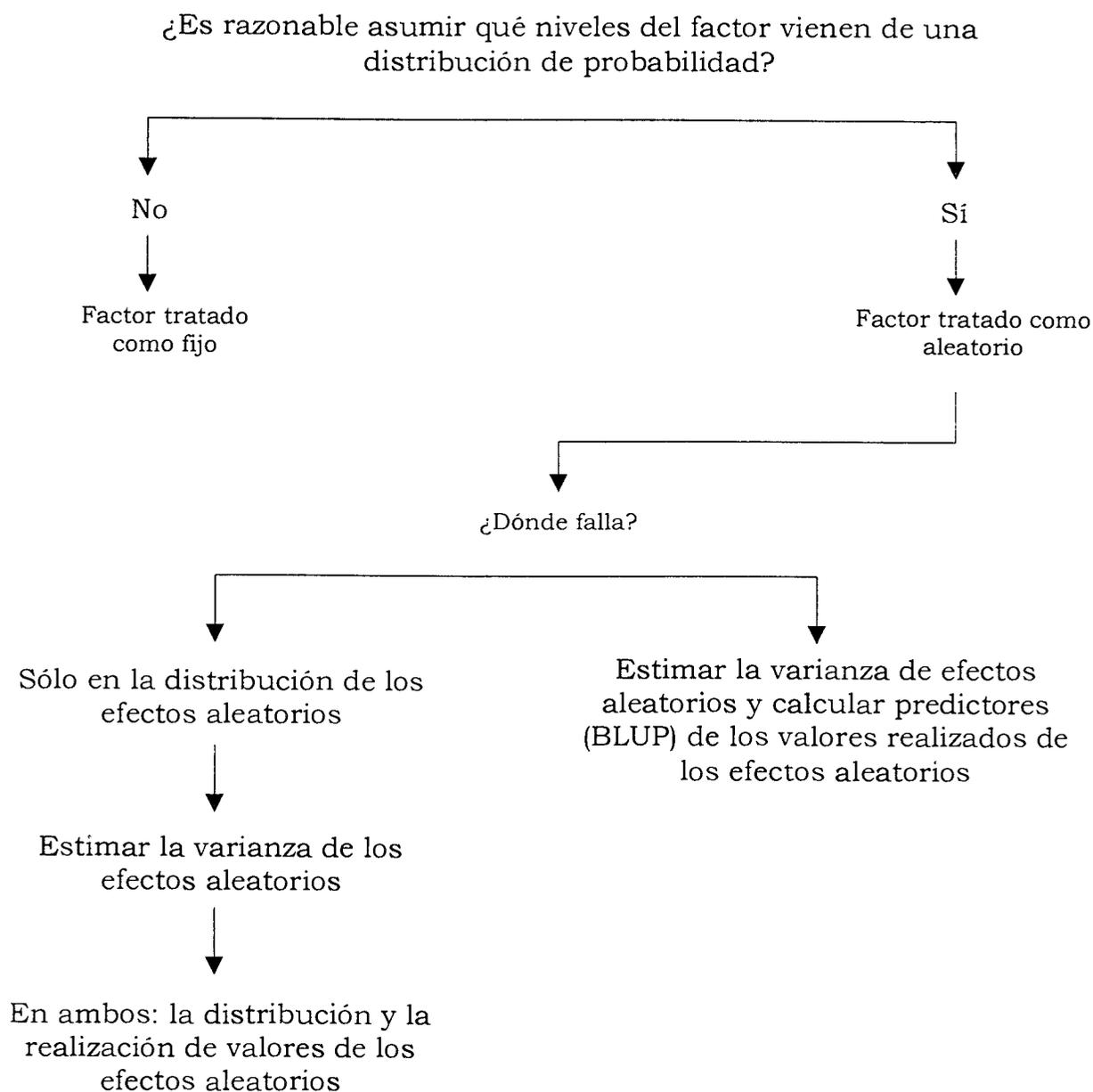
tratamientos se consideran fijos en experimentación repetida (Steel y Torrie, 1988).

Puesto que los compuestos de los modelos difieren en sus suposiciones, también conducen a diferentes pruebas F al evaluar la importancia de los efectos principales (factores A y B). Berenson y Levine (1992) resumen las pruebas F apropiadas para cada modelo:

Cuadro 3.3. Pruebas F para modelos ANOVA de dos factores con repetición.

| $H_0$          | Efectos fijos<br>(A y B) | Efecto aleatorios<br>(A y B) | Efectos mixtos<br>(A fijo, B aleat.) | Efectos mixtos<br>(A aleat., B fijo) |
|----------------|--------------------------|------------------------------|--------------------------------------|--------------------------------------|
| $\mu_{i.} = 0$ | $F = \frac{CM(FA)}{CME}$ | $F = \frac{CM(FA)}{CMAB}$    | $F = \frac{CM(FA)}{CMAB}$            | $F = \frac{CM(FA)}{CMFA}$            |
| $\mu_{.j} = 0$ | $F = \frac{CM(FB)}{CME}$ | $F = \frac{CM(FB)}{CME}$     | $F = \frac{CM(FB)}{CME}$             | $F = \frac{CM(FB)}{CMAB}$            |
| $AB_{ij} = 0$  | $F = \frac{CM(AB)}{CME}$ | $F = \frac{CM(AB)}{CME}$     | $F = \frac{CM(AB)}{CME}$             | $F = \frac{CM(AB)}{CME}$             |

Searle et al. (1992) enfatiza sobre el criterio de decisión de qué factores se deben considerar como fijos y cuáles factores como aleatorios. El siguiente esquema lo ilustra:



| Además, se cuestiona:   | Respuesta | Efectos    |
|---|-----------|------------|
| ¿Son los niveles del factor a considerar una muestra, aleatoria de una población de valores?        | Sí        | aleatorios |
|   | No        | fijos      |
| ¿Hay bastante información acerca de un factor para decidir que los datos son una muestra aleatoria? | Sí        | aleatorios |
|   | No        | fijos      |

Así, los efectos aleatorios ocurren en datos asumidos de poblaciones infinitas (o finitas muy grandes).

En el Cuadro 3.4 se da un breve resumen del desarrollo histórico del modelo mixto:

Cuadro 3.4. Desarrollo histórico del modelo mixto.

| Período | Autor     | Aportación   |
|---------|-----------|--|
| 1947    | Eisenhart | Denomina al modelo de efectos fijos, aleatorios y mixto, modelo I, II y III respectivamente.                             |
| 1949    | Henderson | Da a conocer las primeras bases teóricas del modelo mixto. Descubrimiento de las ecuaciones de los modelos mixtos (MME). |
| 1953    | Henderson | Presentó tres diferentes métodos (adaptación) tipo ANOVA para estimar componentes de varianza y ser utilizados en datos  |

desbalanceados. (método I, para modelos II, método II, para modelo I, con aplicación del método I y el método III para modelo III (mixto).

---

|      |                            |   |
|------|----------------------------|---|
| 1949 | Henderson<br><i>et al.</i> | Probaron que $\hat{\beta}$ de las MME son los mejores estimadores lineales insesgados (BLUE) de los efectos fijos, equivalentes a los estimadores lineales generales (GLS). |
|------|----------------------------|---|

---

|      |           |   |
|------|-----------|---|
| 1963 | Henderson | Probó que $\hat{u}$ de las MME son los mejores predictores lineales insesgados (BLUP) de los efectos aleatorios. Demostró que la $\text{var}(y) = V = \mathbf{ZGZ}' + \mathbf{R}$ . |
|------|-----------|---|

---

|      |                  |  |
|------|------------------|--|
| 1965 | Nelder<br>y Mead | Emplean el método simplex para actualizar $\mathbf{R}$ y $\mathbf{G}$ en forma iterativa, hasta lograr maximizar l (función de verosimilitud). |
|------|------------------|--|

---

|      |           |  |
|------|-----------|--|
| 1973 | Henderson | Describió las características de los siguientes métodos de predicción: mejor predictor lineal insesgado (BLUP) y sus aplicaciones en las evaluaciones genéticas. |
|------|-----------|--|

---

|      |           |  |
|------|-----------|--|
| 1975 | Henderson | Propuso reglas para obtener $(A^+)^{-1}$ donde $A^+$ es la matriz que incluye a todos los animales (con y sin registro) en un modelo mixto animal) |
|------|-----------|--|

---

|         |                      |   |
|---------|----------------------|---|
| 1977-79 | Harville<br>y Searle | Desarrollaron una forma equivalente del logaritmo de la función de verosimilitud restringida para una distribución normal multivariada, siendo $\Lambda = -.5 [\text{constante} + \log  \mathbf{R}  + \log  \mathbf{G}  + \log  \mathbf{C}  + \mathbf{y}'\mathbf{P}\mathbf{y}]$ . |
|---------|----------------------|---|

- 1981 Kackar y Harville Probaron que si las varianzas son estimadas con un método que cumpla con ciertas propiedades y estas estimaciones son sustituidas en las MME, los estimadores de los efectos fijos y los predictores de los efectos aleatorios son insesgados.
- 
- 1986-87 Smith y Graser *et al.* Propusieron un método para maximizar la forma equivalente de 1 de Harville y Searle son el empleo de derivadas o esperanzas.
- 
- 1993 Boldman *et al.* Popularizaron el procedimiento de libre derivación (DF-REML) entre genetistas animales
- 
- 1992 Searle *et al.* Demuestran que la matriz  $\mathbf{P}$  para simplificar las ecuaciones de ML es equivalente a  $\mathbf{K}(\mathbf{K}'\mathbf{V}\mathbf{K})^{-1}\mathbf{K}'$  que se tiene en las ecuaciones de REML.
- 
- 1992 Rodríguez-Herrera Estudió la metodología de modelos mixtos (MMM) con el objetivo de determinar aplicaciones potenciales en el mejoramiento genético. En una primera parte de su investigación encontró utilidad de la metodología de modelos mixtos al aplicarla en un diseño experimental de bloques incompletos desbalanceado donde fue posible obtener resultados igual que si el diseño fuera balanceado. En una segunda parte aplicó la metodología de modelos mixtos (MMM) a un problema de interacción genotipo-medio ambiente, encontrando que MMM provee información adicional a la obtenida por las

metodologías de parámetros de estabilidad de Eberhart y Russell, y ecovalencia de Wricke.

- 
- 1994 Madsen *et al.* Extendieron el procedimiento AI-REML a modelos multivariados y lo compararon con DF-REML.
- 
- 1995 Johnson y Thompson Para modelos univariados compararon el procedimiento AI-REML con dos procedimientos comúnmente usados (EM-REML y DF-REML).

### Conceptos Básicos de la Metodología de los Modelos Mixtos

La siguiente ecuación ilustra el modelo lineal mixto:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{ZU} + \varepsilon$$

donde

$\mathbf{y}$  es un vector  $n \times 1$  de medida de respuesta

$\mathbf{X}$  es una matriz  $n \times (p+1)$  conocida, con rango de  $\mathbf{X} < n, (p+1)$

$\beta$  es un vector  $(p+1) \times 1$  de efectos fijos desconocido

$\mathbf{Z}$  es una matriz  $n \times q$  incidente (concurrente)

$\mathbf{U}$  es un vector  $q \times 1$  de efectos aleatorios con  $E(\mathbf{U}) = \mathbf{0}$ ,

$\varepsilon$  es un vector  $n \times 1$  con  $E(\varepsilon) = \mathbf{0}$ , y

$$\text{Var} \begin{bmatrix} \mathbf{U} \\ \varepsilon \end{bmatrix} = \begin{bmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix}$$

Para este modelo (McLean, 1989) tenemos:

$$E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$$

$$\text{Var}(\mathbf{y}) = \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}$$

donde

$\mathbf{G}$  es una matriz simétrica de efectos aleatorios ( $E(\mathbf{u}\mathbf{u}')$ )

$\mathbf{R}$  es una matriz simétrica de residuales ( $E(\mathbf{e}\mathbf{e}')$ )

Henderson (1949) aplicando el método de mínimos cuadrados, obtuvo las ecuaciones de los modelos mixtos (MME):

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix}$$

Bajo la asunción de normalidad de los  $\varepsilon$  y  $\mathbf{u}$  y escribiendo la inversa generalizada, denotada con el superíndice  $-$ , de la matriz de coeficientes del modelo:

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix}^{-} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \mathbf{C}$$

Las ecuaciones de los modelos mixtos (MME) es uno de los descubrimientos de mayor importancia al desarrollo de estrategias de cómputo para la predicción; tales ecuaciones se cimentan en el procedimiento de mínimos cuadrados, así como ciertos algoritmos que permiten su resolución.

## Parámetros, Estimadores y Valores Esperados

**Definición 1.** Sea  $\{x_i: i=1,2,\dots, n\}$  una muestra y suponer que la densidad conjunta de la muestra viene dada por  $f(x; \theta)$  es un conjunto de parámetros desconocidos. De una función  $h(x)$  de los datos, que no dependa de  $\theta$ , se dice que es un estimador o un estadístico. El valor que toma esa función para una muestra concreta se denomina estimación.

Los parámetros (cantidades desconocidas de una población) más importantes para describir características continuas o cuantitativas son la media  $\mu$  y la varianza  $\sigma^2$ . En el mundo real los parámetros no se conocen con exactitud ya que el total de la población se desconoce, de ahí que los parámetros deberán estimarse con base en una muestra. Podrá justificarse el suponer parámetros conocidos cuando las muestras sean de tamaño muy grande.

La media. Si  $x_i(i=1,\dots,n)$  es la observación  $i$ -ésima para la característica  $X$ , entonces el estimador de  $\mu_x$  es  $\hat{\mu}_x$ . Un estimador sencillo es:

$$\hat{\mu}_x = \sum_{i=1}^n x_{i/n} = (X_1 + X_2 + \dots + X_n) / n$$

Si se conocieran todos los registros de la población,  $N$ , entonces:

$$\mu_x = \sum_{i=1}^n x_i / N$$

La varianza. El operador varianza se define como:

$$\sigma_x^2 = \sum_{i=1}^n (X_i - \mu)^2 p(X_i) = E[(X - \mu)^2], X \text{ discreta}$$

$$\text{Var}(X) = E(X - \mu)^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx = \sigma^2, X \text{ continua}$$

en palabras, la varianza es la esperanza del cuadrado de la desviación estándar para la característica X. Las varianzas son útiles para describir poblaciones y, más aún, para usarse junto con las covarianzas en el desarrollo de procedimientos para la predicción.

El operador esperanza (valor esperado) E, para una característica X, se define como:

$$E(X) = \int_{-\infty}^{+\infty} xf(x) dx = \mu$$

donde X es una variable aleatoria con función de densidad f(·).

Las definiciones de estos operadores, son para el contexto unidimensional. Se dan algunas definiciones para la obtención de la esperanza.

Sea c una constante;  $X_i$  una variable de alguna distribución con media  $\mu_x$  y varianza  $\sigma_x^2$ ,  $Y_i$  una variable de alguna distribución con media  $\mu_y$ , varianza  $\sigma_y^2$ , y covarianza  $\sigma_{xy}$ , entonces:

**Definición 2.**  $E(c) = C$

$$E(c^2) = C^2$$

**Definición 3.**  $E(X_i) = \mu_x$

**Definición 4.**  $E(cX_i) = cE(X_i) = c\mu_x$

**Definición 5.**  $E(X_i + Y_i) = E(X_i) + E(Y_i) = \mu_x + \mu_y$

**Definición 6.**  $E[(X_i - \mu_x)^2] = \sigma_x^2$

de esta definición se desprende que

$$E\left(X_i^2\right) = \sigma_x^2 + \mu_x^2$$

note que si  $\mu_x^2 = 0$ , entonces  $E\left(X_i^2\right) = \sigma_x^2$ .

**Definición 7.**  $E[(X_i - \mu_x)(Y_i - \mu_y)] = \sigma_{xy}$ , covarianza.

Usando esta última definición:

$$E(X_i Y_i) = \sigma_{xy} + \mu_x \mu_y$$

Note que

$E(X_i Y_i) = \sigma_{xy}$  cuando ambas  $\mu_{xy}$  y  $\mu_y$ , son cero.

**Definición 8.** Si  $a_i$ ,  $i=1,2,\dots,n$ , son constantes fijas y  $X_i$  son variables aleatorias con esperanzas (medias)  $\mu_i$ ,  $i=1,2,\dots,n$ , respectivamente, entonces

$$\text{Var}\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 \text{Var}(X_i) = \sum_{i=1}^n a_i^2 \sigma_i^2$$

Aplicando el caso multivariante:

**Definición 9.** Sea  $\mathbf{X}=(X_{ij})$  una matriz cuyos elementos,  $X_{ij}$ , son variables aleatorias. Se define la esperanza de  $\mathbf{X}$  como:

$$E(\mathbf{X})=[E(X_{ij})]$$

dicha esperanza no es más que la matriz cuyos elementos son las esperanzas de los elementos de la matriz  $\mathbf{X}$ . Una consecuencia de esta definición es la

**Proposición 1.** Sean  $\mathbf{A}$ ,  $\mathbf{B}$  y  $\mathbf{C}$ , tres matrices cuyos elementos son no aleatorios; sea  $\mathbf{X}$  una matriz aleatoria tal que están definidas las matrices  $\mathbf{AXB}$  y  $\mathbf{AX+C}$ , entonces

$$E(\mathbf{AXB})= \mathbf{AE(X)B}, E(\mathbf{AX+C})= \mathbf{AE(X)+C}$$

**Proposición 2.** Si  $\mathbf{v}$  es un vector de variables aleatorias, entonces;

(i)  $E[\mathbf{v}]=\mathbf{0}$ ,

(ii)  $E\{[\mathbf{v}-E(\mathbf{v})] [\mathbf{v}-E(\mathbf{v})]'\}=\text{Var}(\mathbf{v})= \mathbf{V}$

**Proposición 3.** Si  $\mathbf{v}'\mathbf{A}\mathbf{v}$  es una forma cuadrática, entonces  $E[\mathbf{v}'\mathbf{A}\mathbf{v}]=\text{tr}[\mathbf{AE}(\mathbf{v}\mathbf{v}')] = \text{tr}(\mathbf{A}\mathbf{V}) = \sum \sum a_{ij}v_{ij}$

**Proposición 4.** Sea  $\mathbf{y}=\mathbf{Xb}+\mathbf{e}$  un modelo lineal, entonces

$$E[\mathbf{y}]=\mathbf{Xb}; E[\hat{\mathbf{b}}]=E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}]=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E[\mathbf{y}]=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Xb} = \mathbf{b}$$

**Proposición 5.** Sea  $\mathbf{y}=\mathbf{Xb}+\mathbf{Zu}+\mathbf{e}$  un modelo lineal con  $\text{Var}(\mathbf{u})=\mathbf{G}$  y  $\text{Var}(\mathbf{e})=\mathbf{R}$ , entonces

$$(i) \quad E\{[\mathbf{y}-E(\mathbf{y})][\mathbf{y}-E(\mathbf{y})]'\}=\mathbf{ZGZ}'+\mathbf{R},$$

$$(ii) \quad E\{[\mathbf{y}-E(\mathbf{y})][\mathbf{e}-E(\mathbf{e})]'\}=\mathbf{R},$$

$$(iii) \quad E\{[\mathbf{y}-E(\mathbf{y})][\mathbf{u}-E(\mathbf{u})]'\}=\mathbf{ZG}$$

**Definición 10.** Sean  $\{X_1, X_2, \dots, X_n\}$  un conjunto de variables aleatorias y sean  $f(\cdot, \cdot, \dots, \cdot)$  su función de densidad conjunta. Se define la covarianza entre  $X_i$  y  $X_j$ ,  $i, j=1, 2, \dots, n$ , como

$$\text{Cov}(X_i, X_j)=E[(X_i-\mu_i)(X_j-\mu_j)],$$

donde  $\mu_i=E(X_i)$ ,  $i=1, 2, \dots, n$

Una representación de las propiedades ligadas a los momentos de segundo orden de una distribución multivariante viene dada en función de la matriz de covarianzas.

**Definición 11.** Sea  $\{X_1, X_2, \dots, X_n\}$  un conjunto de variables aleatorias. Entonces la matriz de covarianzas (de su distribución conjunta) es la matriz

$$\Sigma=(\sigma_{ij}), \quad i, j=1, 2, \dots, n,$$

donde

$$\sigma_{ij} = \text{Cov}(X_i, X_j)$$

De esta última definición se deduce la siguiente propiedad de una matriz de covarianzas.

**Proposición 6.** Sea  $\{X_1, X_2, \dots, X_n\}$  un conjunto de variables aleatorias; sea  $\Sigma$  la matriz de covarianzas de su distribución conjunta.

Entonces:

(a)  $\Sigma$  es simétrica

(b)  $\Sigma$  es definida positiva, salvo que las variables aleatorias sean linealmente dependientes.

**Definición 12.** Sea  $X = \{X_1, X_2, \dots, X_n\}$  un conjunto de variables aleatorias y sea  $\Sigma = (\sigma_{ij})$  su matriz de covarianzas. El coeficiente de correlación (simple) entre  $X_i$  y  $X_j$ ,  $i, j = 1, 2, \dots, n$ , se define como

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}$$

La matriz de correlación es, pues,

$$\mathbf{R} = (\rho_{ij}) = \mathbf{S}\Sigma\mathbf{S}$$

donde

$$\mathbf{S} = \text{diag}(\sigma^{-1/2_{11}}, \sigma^{-1/2_{22}}, \dots, \sigma^{-1/2_{nn}})$$

siendo  $\mathbf{R}$  simétrica y definida positiva, salvo que los elementos de  $\mathbf{X}$  sean linealmente dependientes.

**Proposición 7.** Sea  $\{X_i, i=1,2,\dots,n\}$  un conjunto de variables aleatorias independientes e idénticamente distribuidas, tales que

$$X_i \sim N(0,1), i=1,2,\dots,n,$$

considerar

$$\mathbf{y} = \mathbf{A}\mathbf{X} + \boldsymbol{\mu},$$

donde  $\mathbf{A}$  es una matriz  $n \times n$  no aleatoria y no singular y  $\boldsymbol{\mu}$  un vector de  $n$  elementos no aleatorios. La distribución conjunta de los elementos de  $\mathbf{y}$  es:

$$\Phi(\mathbf{y}) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp\{-1/2(\mathbf{y}-\boldsymbol{\mu})' \Sigma^{-1}(\mathbf{y}-\boldsymbol{\mu})\}$$

**Definición 13.** Sea  $\mathbf{y}$  un vector aleatorio de  $n$  elementos con distribución conjunta  $\Phi(\mathbf{y})$ . Decimos que  $\mathbf{y}$  tiene una distribución normal multivariante, de vector de medias  $\boldsymbol{\mu}$  y matriz de covarianzas  $\Sigma$ . Esto se escribe

$$\mathbf{y} \sim N(\boldsymbol{\mu}, \Sigma)$$

Una vez definida la normal multivariante, veamos algunas propiedades:

**Proposición 8.** Sea  $\mathbf{y} \sim N(\mu, \Sigma)$ ; suponiendo que  $\mathbf{y}$  es  $n \times 1$  y sea  $\mathbf{B}$  una matriz  $n \times n$  no singular y no aleatoria. Entonces

$$\mathbf{Z} = \mathbf{B}\mathbf{y}$$

tiene la distribución

$$\mathbf{Z} \sim N(\mathbf{B}\mu, \mathbf{B}\Sigma\mathbf{B}')$$

**Proposición 9.** Sea  $\mathbf{y} \sim N(\mu, \Sigma)$ , donde  $\Sigma$  es definida positiva e  $\mathbf{y}$  es  $n \times 1$ ; se particiona

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}$$

de forma que  $\mathbf{y}_1$  tenga  $k$  elementos,  $\mathbf{y}_2$  tenga  $n-k$  elementos.

Se particiona

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix},$$

Coherentemente con  $\mathbf{y}$ , se hace

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix},$$

de forma que  $\Sigma_{11}$  sea  $k \times k$ ,  $\Sigma_{22}$  sea  $(n-k) \times (n-k)$  y  $\Sigma_{12}$  sea  $k \times (n-k)$ .

Entonces

$$\mathbf{y}_1 \sim N(\mu_1, \Sigma_1)$$

Las distribuciones condicionadas asociadas a la normal multivariante se obtienen como sigue:

**Proposición 10.** Sea  $\mathbf{y} \sim N(\mu, \Sigma)$  y particionándola  $\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}$

entonces la densidad de  $\mathbf{y}_1$  condicionada por  $\mathbf{y}_2$  es

$$N[\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{y}_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}]$$

A menudo la matriz  $\Sigma_{12}\Sigma_{22}^{-1}$  se llama matriz de los coeficientes de regresión de  $\mathbf{y}_1$  sobre  $\mathbf{y}_2$ . De los resultados de esta proposición, se tiene

**Proposición 11.** Sea  $\mathbf{y} \sim N(\mu, \Sigma)$  y particionando  $\mathbf{y}$ ,  $\mu$  y  $\Sigma$ :

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} \quad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

donde  $\mathbf{y}_1$  es  $k \times 1$ ,  $\mathbf{y}_2$  es  $(n-k) \times 1$ ,  $\Sigma_{11}$  es  $k \times k$ ,  $\Sigma_{22}$  es  $(n-k) \times (n-k)$ .

**Definición 14.** Sea  $\mathbf{y} \sim N(\mu, \Sigma)$  y particionando

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} \quad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

Se define el coeficiente de correlación parcial entre  $y_i$  e  $y_j$  dados  $y_{k+1}, \dots, y_n$  como el coeficiente de correlación simple entre  $y_i$  e  $y_j$  en el contexto de la distribución de  $\mathbf{y}_1$  condicionada por  $\mathbf{y}_2$ ; se denota como  $\rho_{ij.k+1,k+2,\dots,n}$ .

**Proposición 12.** Sea  $\mathbf{A}$  una matriz no aleatoria  $s \times n$ , con  $s < n$  y  $r(\mathbf{A})=s$ , y si se supone que  $\mathbf{X} \sim N(\mu, \Sigma)$ , con  $\Sigma$  no singular. Entonces

$$\mathbf{y} \sim N(\mathbf{A}\mu, \mathbf{A}\Sigma\mathbf{A}'),$$

donde

$$\mathbf{y} = \mathbf{A}\mathbf{X}$$

**Proposición 13.** Sea  $\mathbf{X}$  una variable aleatoria de  $n$  elementos, con media  $\mu$  y matriz de covarianzas  $\Sigma$ ; si toda combinación lineal  $\alpha\mathbf{X}$  se distribuye normalmente, entonces

$$\mathbf{X} \sim N(\mu, \Sigma)$$

**Proposición 14.** Sea  $\mathbf{X} \sim N(0, \mathbf{I})$ , donde  $\mathbf{X}$  es  $n \times 1$ . Entonces

$$\mathbf{X}'\mathbf{X} \sim \chi^2_n$$

**Proposición 15.** Sea  $\mathbf{X} \sim N(\mu, \Sigma)$ , donde  $\mathbf{X}$  es  $n \times 1$ . Entonces

$$(\mathbf{X} - \mu)' \Sigma^{-1} (\mathbf{X} - \mu) \sim \chi^2_n$$

**Proposición 16.** Sea  $\mathbf{X} \sim N(0, \mathbf{I})$ , donde  $\mathbf{X}$  es  $n \times 1$ , y sea  $\mathbf{A}$  una matriz simétrica de rango  $r$ . Entonces

$$\mathbf{X}'\mathbf{A}\mathbf{X} \sim \chi^2_r$$

sí y sólo sí  $\mathbf{A}$  es idempotente.

**Proposición 17.** Sea  $\mathbf{X} \sim N(0, \mathbf{I})$ , y sean  $\mathbf{A}$  y  $\mathbf{B}$  dos matrices  $n \times n$ . Si  $\mathbf{A}\mathbf{B} = \mathbf{0}$ ,  $\mathbf{X}'\mathbf{A}\mathbf{X}$  y  $\mathbf{X}'\mathbf{B}\mathbf{X}$  son idempotentes.

**Definición 15.** Sea  $\{x_i: 1, 2, \dots, n\}$  una muestra y  $f(x; \theta)$  su función de densidad conjunta; sea  $\hat{\theta} = h(x)$  un estimador de  $\theta$ . Se dice que el estimador es insesgado si

$$E(\hat{\theta}) = \theta$$

Se dice que el estimador es sesgado si  $E(\hat{\theta}) \neq \theta$

**Definición 16.** Sea  $\hat{\theta}_n$  un estimador de un parámetro (escalar)  $\theta$ , basado en una muestra de tamaño  $n$ . Se dice que es un estimador consistente de  $\theta$  si, dados  $\epsilon$  y  $\delta$  positivos, existe un tamaño muestral,  $n^*$ , tal que para todo  $n > n^*$

$$\Pr[|\hat{\theta}_n - \theta| > \delta] <$$

La consistencia también se llama convergencia en probabilidad.

**Proposición 18.** (Desigualdad de Chebycheff). Sea  $X$  una variable aleatoria de media  $\mu$  y varianza  $\sigma^2$ , y sea  $k > 0$  un número real.

Entonces

$$\Pr[|X - \mu| > k] \leq \frac{\sigma^2}{k^2}$$

**Proposición 19.** (Teorema de Khinchine). Sea  $\{X_t:t=1,2,\dots\}$  una sucesión de variables aleatorias i.i.d. tales que existe  $E(X_t)=\mu$ . Entonces

$$p \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T X_t = \mu$$

**Proposición 20.** Sea  $C$  una clase de estimadores de un vector de parámetros,  $\theta$ , que tienen ciertas propiedades; sean  $\hat{\theta}_1$  y  $\hat{\theta}_2 \in C$  tales que existen sus matrices de covarianzas. Se dice que  $\hat{\theta}_1$  es más eficiente que  $\hat{\theta}_2$  si

$$Cov(\hat{\theta}_2) - Cov(\hat{\theta}_1)$$

es semidefinida positiva. Si para todo  $\hat{\theta}_i \in C$ ,

$$Cov(\hat{\theta}_i) - Cov(\hat{\theta}_1)$$

es semidefinida positiva, se dice que  $\hat{\theta}_1$  es el mejor dentro de la clase  $C$ .

### El Mejor Estimador Lineal Insesgado (BLUE)

**Proposición 21.** El estimador OLS del parámetro  $\beta$  del modelo  $\mathbf{y} = \mathbf{X}\beta + \mathbf{e}$ , es el vector que minimiza

$$S(\beta) = (\mathbf{y} - \mathbf{X}\beta)' (\mathbf{y} - \mathbf{X}\beta)$$

con respecto a  $\beta$ , y viene dado por

$$\hat{\beta} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{y}$$

que tiene las siguientes propiedades:

(a) es insesgado

(b) es consistente

(c) dentro de la clase de estimadores lineales insesgados de  $\beta$ , es eficiente (en el sentido de que si  $\hat{\beta}$  es cualquier otro estimador lineal e insesgado de  $\beta$ , entonces  $Cov(\beta) - Cov(\hat{\beta})$  es semidefinida positiva).

Como una muestra es un conjunto de variables aleatorias y un estimador es una función de la muestra, un estimador será también una variable aleatoria. Como tal, tendrá función de densidad y de distribución. Que sea insesgado significa que la media de la distribución es el parámetro que se está estimando.

Para que el estimador OLS de  $\beta$  esté definido de forma única, debe existir la inversa de  $\mathbf{X}'\mathbf{X}$ . Al proceso de obtener el estimador  $\hat{\beta}$  se denomina regresión de  $y$  sobre  $X$ .

Si se quieren obtener las ecuaciones normales de mínimos cuadrados, se tiene:

$$\mathbf{y} = \mathbf{X} \boldsymbol{\beta} + \mathbf{e}$$

$n \times 1 \quad n \times k \quad k \times 1 \quad n \times 1$

$$E(\mathbf{e}) = \mathbf{0}$$

$$\text{Cov}(\mathbf{e}) = \sigma^2 \mathbf{I}$$

$\mathbf{X}$  es una matriz fija, con  $r(\mathbf{X}) = k < n$

donde  $\mathbf{I}$  es la matriz identidad de  $n \times n$ .

Además

$$E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta} + E(\mathbf{e}) = \mathbf{X}\boldsymbol{\beta}$$

donde la matriz  $\mathbf{X}$  y el vector  $\boldsymbol{\beta}$  son estocásticos.

Se define el vector de residuos en la  $i$ -ésima observación como

$$\hat{\mathbf{e}}_i = y_i - \mathbf{x}_i \hat{\boldsymbol{\beta}}, \quad i=1,2,\dots,n$$

donde  $\hat{\boldsymbol{\beta}}$  es el vector de coeficientes estimados. La suma de cuadrados de los residuos resulta entonces

$$S = \sum_{i=1}^n \hat{\mathbf{e}}_i^2 = \sum_{i=1}^n (y_i - \mathbf{x}_i \hat{\boldsymbol{\beta}})^2$$

la cual, de acuerdo con la técnica de mínimos cuadrados, tiene que ser minimizada. En notación matricial

$$\hat{\mathbf{e}} = \mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}}$$

$n \times 1 \quad n \times 1 \quad n \times k \quad k \times 1$

es el vector de residuos y la suma de cuadrados puede escribirse

$$S = \hat{\mathbf{e}}' \hat{\mathbf{e}} = (\mathbf{y} - \mathbf{x} \hat{\beta})' (\mathbf{y} - \mathbf{x} \hat{\beta})$$

efectuando la transposición y multiplicación vectorial

$$S = \mathbf{y}' \mathbf{y} - 2 \hat{\beta}' \mathbf{x}' \mathbf{y} + \hat{\beta}' \mathbf{x}' \mathbf{x} \hat{\beta}$$

Los estimadores mínimos cuadráticos son entonces la solución de  $\hat{\beta}$ .

$$\min_{\hat{\beta}} S = S(\hat{\beta}) = \mathbf{y}' \mathbf{y} - 2 \hat{\beta}' \mathbf{x}' \mathbf{y} + \hat{\beta}' \mathbf{x}' \mathbf{x} \hat{\beta}$$

Las condiciones necesarias y suficientes para minimizar  $S$  son que todas las derivadas parciales de primer orden sean cero, es decir

$$\frac{\partial S}{\partial \hat{\beta}} = -2 \mathbf{x}' \mathbf{y} + 2 \mathbf{x}' \mathbf{x} \hat{\beta} = \mathbf{0}$$

Esta condición puede escribirse

$$\mathbf{x}' \mathbf{x} \hat{\beta} = \mathbf{x}' \mathbf{y}$$

siendo las ecuaciones normales de mínimos cuadrados, donde  $\hat{\beta}$  es el estimador mínimo cuadrático de  $\beta$ . Premultiplicando ambos lados por  $(\mathbf{x}' \mathbf{x})^{-1}$  produce la solución única para  $\hat{\beta}$ :

$$\hat{\beta} = (\mathbf{x}' \mathbf{x})^{-1} \mathbf{x}' \mathbf{y}$$

que es el estimador MCO de  $\beta$ .

El estimador  $\hat{\beta}$  es lineal en  $\mathbf{y}$ , pero  $\mathbf{y}$  contiene un término estocástico aditivo y, por lo tanto, es estocástico.

Para analizar las varianzas y covarianzas de los BLUE se requiere del cálculo de la matriz de covarianzas definida como

$$\text{Cov}(\hat{\beta}) = E[(\hat{\beta} - E(\hat{\beta}))(\hat{\beta} - E(\hat{\beta}))']$$

pero por ser  $\hat{\beta}$  insesgado

$$E(\hat{\beta}) = \beta$$

entonces

$$\hat{\beta} - \beta = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{e}$$

así que

$$\text{Cov}(\hat{\beta}) = E[(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{e}\mathbf{e}'\mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}]$$

Utilizando los supuestos estocásticos de la covarianza de  $\mathbf{e}$

$$\text{Cov}(\mathbf{e}) = E[(\mathbf{e} - E(\mathbf{e}))(\mathbf{e} - E(\mathbf{e}))'] = E(\mathbf{e}\mathbf{e}') = \sigma^2\mathbf{I}$$

se sigue que

$$\text{Cov}(\hat{\beta}) = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\text{Cov}(\mathbf{e})\mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}$$

al combinar  $\text{Cov}(\mathbf{e}) = \sigma^2\mathbf{I}$  y cancelar  $(\mathbf{x}'\mathbf{x})^{-1}(\mathbf{x}'\mathbf{x})$  produce

Existen varias opciones para obtener soluciones al sistema de ecuaciones normales; algunas son:

- a) Aplicar restricciones sobre las soluciones. El número de restricciones debe ser igual al orden de  $\mathbf{X}'\mathbf{X}-r(\mathbf{X}'\mathbf{X})$ . Las restricciones no se deben aplicar sobre funciones estimables.
- b) Usar un procedimiento iterativo (evaluaciones cíclicas). En general, la  $i$ -ésima solución de las  $N$  ecuaciones, para el  $n$ -ésimo ciclo de iteración, se obtiene de la siguiente fórmula:

$$S_i^{(n)} = S_i^{(n-1)} + (1 / c_{ii})[r_i - \sum_{j=1}^{i-1} c_{ij}S_j^{(n)} - c_{ii}S_i^{(n-1)} - \sum_{j=i+1}^{i-1} c_{ij}S_j^{(n-1)}]$$

la iteración se suspende hasta llegar a la convergencia (la expresión entre paréntesis rectangulares es cero). Este procedimiento se le denomina Gauss-Seidel; existe también el procedimiento iterativo de Jacobi, el cual involucra las soluciones del ciclo previo.

- c) Uso de la inversa generalizada. Una forma de obtener la  $g$ -inversa de una matriz, es identificar una submatriz de rango completo (i.e. no singular), invertir la submatriz y agregar ceros fuera de la submatriz. Otra forma es aplicar programas computacionales de matrices (SAS, MATLAB, etc.) o bien existen algoritmos para su cálculo (Searle, 1971 y Graybill, 1976 los describen).

d) Uso del modelo para las medias de cada subclase. Este modelo se escribe

$$\mu_{ij} = \mu_i + e_{ij}$$

donde

$\mu_i = \mu + G_i$ , siendo  $G_i$  los estimadores de las funciones.

**Definición 17.** Se dice que una función paramétrica lineal,  $\lambda'\beta$ , es linealmente estimable si existe una función lineal de las observaciones,  $\mathbf{a}'\mathbf{y}$ , tal que su valor esperado sea igual a  $\lambda'\beta$  independientemente de los valores verdaderos de los parámetros desconocidos. Si tal función,  $\mathbf{a}'\mathbf{y}$ , no existe entonces se dice que  $\lambda'\beta$  no es estimable linealmente.

Si  $\lambda'\beta$  es estimable entonces por definición existe un vector columna  $\mathbf{a}$  tal que  $\lambda'\beta = E(\mathbf{a}'\mathbf{y})$ .

**Proposición 22.** Sea el Modelo Lineal General

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e},$$

donde

$$r(\mathbf{X}) < n+1$$

entonces:

- (i)  $\mathbf{G}\beta$  es una función linealmente estimable si  $\mathbf{G}$  pertenece al subespacio engendrado por las filas de  $\mathbf{X}$ , o sea, si existe una matriz  $\mathbf{C}$  tal que

$$\mathbf{G}=\mathbf{C}\mathbf{X}$$

- (ii)  $\sigma^2$  es un parámetro estimable

**Corolario 1.** Si se supone que los elementos del proceso de los errores,  $\{e_t:t=1,2,\dots\}$ , se distribuyen normalmente, entonces

$$\mathbf{G}\beta \sim N(\mathbf{G}\beta, \sigma^2 \mathbf{G}\mathbf{X}_g \mathbf{X}_g' \mathbf{G}')$$

Note que si  $\mathbf{G}=\mathbf{X}$ ,

$$\mathbf{X}\beta \sim N(\mathbf{X}\beta, \sigma^2 \mathbf{X}\mathbf{X}_g \mathbf{X}_g' \mathbf{X}')$$

Ahora bien,

$$\mathbf{X}\mathbf{X}_g \mathbf{X}_g' \mathbf{X}' = \mathbf{X}\mathbf{X}_g \mathbf{X}\mathbf{X}_g = \mathbf{X}\mathbf{X}_g$$

y sabemos que  $\mathbf{X}\mathbf{X}_g$  es una matriz  $T \times T$  de rango  $r$ , luego

$$\mathbf{X}\beta \sim N(\mathbf{X}\beta, \sigma^2 \mathbf{X}\mathbf{X}_g)$$

### Mejor Predictor Lineal Insesgado (BLUP)

varios métodos de predicción se han desarrollado. Henderson (1973) describió las principales características de los siguientes métodos de predicción:

- a) Mejor predictor (BP). Sea  $\{y_n\}$  un conjunto de registros, y sea  $\mathbf{u}$  un vector no observable el cual se desea predecir y está conjuntamente distribuido con  $\mathbf{y}$ , entonces el predictor del  $i$ -ésimo elemento de  $\mathbf{u}$  es:

$$\mathbf{u}_i = f(\mathbf{y}),$$

Se desea minimizar el error de predicción  $(\hat{\mathbf{u}}_i - \mathbf{u}_i)$ , medido como  $E(\hat{\mathbf{u}}_i - \mathbf{u}_i)^2$ , o sea, el BP minimiza el cuadrado medio del error.

El BP de  $\mathbf{u}_i$  equivale a la media condicional de  $\mathbf{u}_i$ , o sea,  $E(\mathbf{u}_i | \mathbf{y})$ . Este método supone: 1) la distribución conjunta de  $\mathbf{y}$  y  $\mathbf{u}$ ; 2) se debe derivar la distribución condicional; y 3) se deben conocer los valores de los parámetros de la distribución condicional.

Algunas propiedades del BP son:

i) El BP es un predictor insesgado:  $E(\hat{\mathbf{u}}_i) = E(\mathbf{u}_i)$

ii)  $\text{Var}(\hat{\mathbf{u}} - \mathbf{u}) = \text{Var}(\mathbf{u} | \mathbf{y})$

$$\text{Var}(\hat{\mathbf{u}}) = \text{Cov}(\hat{\mathbf{u}}, \mathbf{u})$$

iii) El BP maximiza el promedio de los valores verdaderos del grupo seleccionado (selección truncada).

- b) Mejor predictor lineal (BLP). El BLP del vector  $\mathbf{u}$ , que minimiza  $E(\hat{\mathbf{u}} - \mathbf{u})^2$ , puede describirse como:

$$\hat{\mathbf{u}}_{BLP} = \mathbf{a} + \mathbf{b}'\mathbf{y}$$

Minimizando  $E[(\mathbf{a} + \mathbf{b}'\mathbf{y}) - \mathbf{u}]^2$  y derivando con respecto a  $\mathbf{a}$  y  $\mathbf{b}$ , se tiene:

$$\mathbf{a} = \mathbf{k}'\boldsymbol{\beta} - \mathbf{C}'\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\beta}, \text{ y } \mathbf{b} = \mathbf{V}^{-1}\mathbf{C}, \text{ donde}$$

$\mathbf{V} = \text{Var}(\mathbf{y})$ ,  $\mathbf{C} = \text{Cov}(\mathbf{y}, \mathbf{u})$ . Entonces

$$\hat{\mathbf{u}}_{BLP} = \mathbf{k}'\boldsymbol{\beta} + \mathbf{C}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

El BLP supone que se conoce: 1) el primer momento de la distribución (medias); 2) el segundo momento de la distribución de  $(\mathbf{y}, \mathbf{u})$  (varianzas). El BLP es más exigente que el BP. Las propiedades del BLP son similares a las de BP:

- (i) El BLP de  $\mathbf{u}$  es insesgado,
- (ii)  $\text{Var}(\hat{\mathbf{u}}) = \mathbf{C}'\mathbf{V}^{-1}\mathbf{C} = \mathbf{b}'\mathbf{V}\mathbf{b}$ ,
- (iii)  $\text{Cov}(\mathbf{u}, \hat{\mathbf{u}}) = \text{Var}(\hat{\mathbf{u}})$ ,
- (iv) Minimiza los errores de predicción, donde  $\text{Var}(\hat{\mathbf{u}} - \mathbf{u}) = \text{Var}(\mathbf{u}) - \text{Var}(\hat{\mathbf{u}})$ ,
- (v) Sí  $\mathbf{y}$  y  $\mathbf{u}$  se distribuyen normalmente, se maximiza el valor promedio de los valores verdaderos del grupo seleccionado (selección truncada).

Existen algunas dificultades con el BLP tales como: 1) a veces no hay precisión en las medias de los efectos fijos; 2) ocasionalmente las varianzas y covarianzas se desconocen; 3) puede generarse un gran número de subclases y con datos extremadamente desbalanceados, siendo el BLP inadecuado.

c) Mejor predictor lineal insesgado (BLUP). El BLUP es menos restrictivo que el BLP, se trata de derivar un predictor de  $\mathbf{u}$  de manera que sea insesgado,  $E(\mathbf{u}) = \mathbf{k}'\beta$ ,  $E(\hat{\mathbf{u}} - \mathbf{u})^2$  es mínimo.

Considerar el predictor

$$\hat{\mathbf{u}} = \mathbf{b}'\mathbf{y}$$

$$E(\hat{\mathbf{u}}) = \mathbf{k}'\beta, \text{ pero}$$

$$E(\mathbf{b}'\mathbf{y}) = \mathbf{b}'\mathbf{X}\beta, \text{ entonces}$$

$$\mathbf{k}' = \mathbf{b}'\mathbf{X}, \text{ o también, } \mathbf{X}'\mathbf{b} = \mathbf{k}$$

En este procedimiento, primero minimiza  $\text{Var}(\mathbf{b}'\mathbf{y} - \mathbf{u}) = \mathbf{b}'\mathbf{V}\mathbf{b} - 2\mathbf{b}'\mathbf{C} + \text{Var}(\mathbf{u})$ , tomando derivadas parciales con respecto a  $\mathbf{b}$ , e igualando a cero, obteniendo  $\mathbf{V}\mathbf{b} = \mathbf{C}$ , y luego aplicar la restricción de que  $\mathbf{X}'\mathbf{b} = \mathbf{k}$ . Para esto último, se aplica el multiplicador de Lagrange,  $\lambda$ , quedando las ecuaciones:

$$\begin{bmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X}' & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{C} \\ \mathbf{k} \end{bmatrix}$$

donde

$\mathbf{b} = \mathbf{V}^{-1}(\mathbf{C} - \mathbf{X}\lambda)$ . Sustituyendo  $\mathbf{b}$  y despejando  $\lambda$ , se tiene:

$\lambda = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{C} - (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{k}$ . Sustituyendo nuevamente  $\lambda$  en la primera ecuación, se tiene,

$$\mathbf{b} = \mathbf{V}^{-1}\mathbf{C} - \mathbf{V}^{-1}\mathbf{X}[(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{C} - (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{k}],$$

si  $\hat{\mathbf{u}} = \mathbf{b}'\mathbf{y}$ , luego

$$\hat{\mathbf{u}} = \mathbf{C}'\mathbf{V}^{-1}\mathbf{y} - \mathbf{C}'\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} + \mathbf{k}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y},$$

pero,  $\beta^{\circ} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$ , es el estimador de mínimos cuadrados generalizados (GLS) de  $\beta$ , así

$$\hat{\mathbf{u}} = \mathbf{C}'\mathbf{V}^{-1}\mathbf{y} - \mathbf{C}'\mathbf{V}^{-1}\mathbf{X}\beta^{\circ} + \mathbf{k}'\beta^{\circ}$$

$$\hat{\mathbf{u}} = \mathbf{k}'\beta^{\circ} + \mathbf{C}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\beta^{\circ})$$

el cual es BLUP de  $\mathbf{u}$ , similar al BLP.

Algunas propiedades importantes de la solución de BLUP son:

- En la clase de predictores lineales insesgados, BLUP maximiza la correlación entre  $U$  y  $u$ .
- $\mathbf{K}'\mathbf{b}$  es BLUE del conjunto de funciones lineales estimables,  $\mathbf{K}'\mathbf{B}$ .

- $E(U|u)=u$
- $u$  es única.
- $K'b+m'u$  es BLUP de  $K'B+m'U$  que proviene de  $K'B$  estimable.
- $\text{Var}(K'b)=K'C_{11}K$ .
- $\text{Var}(K'b+m'u)= K'C_{11}K+m'(G-C_{22})m$
- $\text{Var}(K'b+m'u-K'B-m'U)=(K'm')C(K'm)'$
- $\text{Cov}(K'b,u)=0$
- $\text{Cov}(K'b,U)=-K'C_{12}$
- $\text{Cov}(K'b,u-U)=K'C_{12}$
- $\text{Var}(u)=\text{Cov}(u,U)=G-C_{22}$
- $\text{Var}(u-U)= C_{22}$

### **Ecuaciones de los Modelos Mixtos**

La dificultad principal de este predictor, es que requiere la inversa de la matriz de varianzas- covarianzas de los registros cuyo tamaño es igual al número de registros. Las alternativas para soluciones este problema las dio Henderson con las ecuaciones de los modelos mixtos.

Henderson (1949) propuso las ecuaciones de los modelos mixtos, las cuales simplifican el cálculo del estimador de cuadrados mínimos generalizados ( $\hat{\beta}$ ) y el mejor predictor lineal insesgado ( $\mathbf{u}$ ). De forma general, para un conjunto de  $n$  registros, el modelo lineal mixto se puede escribir:

$$\mathbf{y}=\mathbf{X}\beta+\mathbf{Z}\mathbf{u}+\mathbf{e}$$

donde

$\mathbf{y}$  es el vector  $n \times 1$  de registros de comportamiento,

$\mathbf{X}$  es una matriz  $n \times p$  que asocia  $\mathbf{y}$  con  $\beta$ ,

$\beta$  es un vector  $p \times 1$  de efectos fijos ( $\mu$ ),

$\mathbf{Z}$  es una matriz  $n \times q$  que asocia  $\mathbf{y}$  con  $\mathbf{u}$ ,

$\mathbf{u}$  es un vector  $q \times 1$  de efectos aleatorios,

$\mathbf{e}$  es el vector  $n \times 1$  de residuales peculiares de cada registro.

Las medias y varianzas de los efectos en el modelo son:

$$E(\mathbf{y})=\mathbf{X}\beta, \text{ esto es, } E(\mathbf{u})=E(\mathbf{e})=\mathbf{0},$$

$$\text{Var}(\mathbf{u})=E(\mathbf{u}\mathbf{u}')=\mathbf{G}, \text{ Var}(\mathbf{e})=E(\mathbf{e}\mathbf{e}')=\mathbf{R},$$

$$\text{Cov}(\mathbf{u},\mathbf{e}')=\mathbf{0}, \text{ Cov}(\mathbf{u},\mathbf{y}')=\mathbf{G}\mathbf{Z}',$$

$$\text{Var}(\mathbf{y}) = \mathbf{E}\{[(\mathbf{y} - \mathbf{E}(\mathbf{y}))][(\mathbf{y} - \mathbf{E}(\mathbf{y}))']\} = \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}$$

Comúnmente para el caso de una variable dependiente ( $\mathbf{y}$ ),  $\mathbf{R} = \mathbf{I}\sigma_e^2$ , y  $\mathbf{G} = \mathbf{A}\sigma_u^2$ , donde  $\mathbf{A}$  para el caso en que  $\mathbf{u}$  son efectos genéticos aditivos, es la matriz de relaciones genéticas aditivas. Si  $\mathbf{A} = \mathbf{I}$ , entonces no se consideran parentescos y los efectos genéticos con  $\mathbf{u}$  son independientes.

La forma general de las ecuaciones de los modelos mixtos es:

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \beta \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix}$$

Las ecuaciones de los modelos mixtos son importantes ya que  $\beta^\circ$  y  $\hat{\mathbf{u}}$  se obtienen sin necesitarse el cálculo de  $\mathbf{V}^{-1}$ , ( $\text{NXN}$ ), y aunque  $\mathbf{R}$  es también  $\text{NXN}$ , ésta es generalmente diagonal para análisis univariados y diagonal en bloques para análisis multivariados. Henderson (1959) probaron que  $\beta^\circ$  de estas ecuaciones son BLUE de los efectos fijos, equivalentes a estimadores de mínimos cuadrados generalizados GLS. Henderson (1963) también probó que  $\hat{\mathbf{u}}$  de estas ecuaciones son BLUP de los efectos aleatorios.

Dado que generalmente  $\mathbf{R}^{-1} = \mathbf{I}(1/\sigma_e^2)$  y  $\mathbf{G}^{-1} = \mathbf{A}^{-1}(1/\sigma_u^2)$ , entonces las ecuaciones de los modelos mixtos pueden multiplicarse por  $\sigma_e^2$ , resultando:

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}\lambda \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}$$

Este sistema de ecuaciones se puede representar como  $\mathbf{C}\hat{\mathbf{s}}=\mathbf{r}$ ,  
donde

$\mathbf{C}$  es la matriz de coeficientes,

$\hat{\mathbf{s}}$  es el vector de soluciones, y

$\mathbf{r}$  es el vector de elementos del lado derecho de las ecuaciones.

La matriz  $\mathbf{C}$  y su inversa se pueden representar como:

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{XX} & \mathbf{C}_{XZ} \\ \mathbf{C}_{ZX} & \mathbf{C}_{ZZ} \end{bmatrix}, \quad \mathbf{C}^{-1} = \begin{bmatrix} \mathbf{C}^{XX} & \mathbf{C}^{XZ} \\ \mathbf{C}^{ZX} & \mathbf{C}^{ZZ} \end{bmatrix}$$

### Varianzas y Valores Esperados de las Soluciones a las MME

Sea  $\mathbf{k}'$  un vector de coeficientes para los efectos fijos y  $\mathbf{m}'$  un vector de coeficientes para los efectos aleatorios, entonces la función de predicción es  $\mathbf{k}'\hat{\beta} + \mathbf{m}'\hat{\mathbf{u}}$ , con varianza,

$$\text{Var}(\mathbf{k}'\hat{\beta} + \mathbf{m}'\hat{\mathbf{u}}) = \begin{bmatrix} \mathbf{k}' \\ \mathbf{m}' \end{bmatrix} \mathbf{C}^{-1} [\mathbf{k} \ \mathbf{m}] \sigma_e^2$$

donde se tiene

a)  $\text{Var}(\hat{\beta}) = \mathbf{C}^{XX} \sigma_e^2$

$$b) \text{Var}(\mathbf{k}'\hat{\beta}) = \mathbf{k}'\mathbf{C}^{XX}\mathbf{k}\sigma_e^2$$

$$c) \text{Var}(\hat{\mathbf{u}}) = \text{Cov}(\hat{\mathbf{u}}, \hat{\mathbf{u}}) = \mathbf{G} - \mathbf{C}^{ZZ}\sigma_e^2$$

$$d) \text{Var}(\hat{\mathbf{u}} - \mathbf{u}) = \mathbf{C}^{ZZ}\sigma_e^2$$

Esto indica que conforme el número de registros (información) aumenta, la  $\text{Var}(\hat{\mathbf{u}})$  tiende a  $\mathbf{G}$ , y  $\text{Var}(\hat{\mathbf{u}} - \mathbf{u})$  tiende a cero.

Los valores esperados de las soluciones,  $E(\hat{\mathbf{s}})$ , se obtienen así:

$$E(\hat{\mathbf{s}}) = \mathbf{C}^{-1} \begin{bmatrix} \mathbf{X}'\mathbf{X} \\ \mathbf{Z}'\mathbf{X} \end{bmatrix} \beta$$

Si las ecuaciones de los modelos mixtos son de rango completo,  $E(\hat{\beta}) = \beta$ . En el caso de los efectos aleatorios no existe problema de estimabilidad, entonces  $E(\hat{\mathbf{u}}) = \mathbf{u}$ , esto es, siempre los predictores de los valores genéticos son insesgados.

Para poder escribir las ecuaciones de los modelos mixtos es necesario que se conozcan las varianzas y covarianzas de los efectos aleatorios. Cuando esto ocurre, las propiedades de las soluciones son:

- Las soluciones de los efectos fijos son BLUE de las funciones estimables de los efectos fijos ( $\mathbf{k}'\hat{\beta}$ ). El estimador de efectos

fijos,  $\hat{\beta}$  (estimador de mínimos cuadrados generalizados, GLS) se obtiene como sigue:

$$\beta^{\circ} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}), \text{ donde } \mathbf{V} = \text{Var}(\mathbf{y}) = \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}$$

- Las soluciones de los efectos aleatorios son BLUP de los verdaderos efectos aleatorios. Estos predictores se obtienen como sigue:

$$\hat{\mathbf{g}} = \mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\beta})$$

El producto  $\mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}$  se puede interpretar como una regresión  $[\text{Cov}(\mathbf{g}, \mathbf{y})/\text{Var}(\mathbf{y})]$  de  $\mathbf{g}$  sobre  $\mathbf{y}$  después de ajustar  $\mathbf{y}$  por las medias del grupo contemporáneo.

Considérese el modelo mixto,  $\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{u} + \mathbf{e}$ . Henderson (1963) demostró que si  $\text{Var}(\mathbf{y}) = \mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}$ , entonces:

$$\mathbf{V}^{-1} = \mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}$$

Sea  $\mathbf{Q} = (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})$ , luego,  $\mathbf{V}^{-1} = \mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z}\mathbf{Q}^{-1}\mathbf{Z}'\mathbf{R}^{-1}$

Considerando que  $\mathbf{V}^{-1}\mathbf{V} = \mathbf{I}$ . Así:

$$[\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z}\mathbf{Q}^{-1}\mathbf{Z}'\mathbf{R}^{-1}][\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}] = \mathbf{R}^{-1}\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{I} - \mathbf{R}^{-1}\mathbf{Z}\mathbf{Q}^{-1}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{G}\mathbf{Z}' - \mathbf{R}^{-1}\mathbf{Z}\mathbf{Q}^{-1}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{R}$$

para que esta expresión sea igual a  $\mathbf{I}$ , los términos 1 y 3 deben ser iguales, pero de signo contrario al término 4. Sumando los términos 1 y 3 se tiene,

$$\mathbf{R}^{-1}\mathbf{Z}[\mathbf{I}-\mathbf{Q}^{-1}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}]\mathbf{G}\mathbf{Z}'$$

y multiplicando “ $-\mathbf{Q}^{-1}$ ” por  $\mathbf{G}^{-1}$  y por  $-\mathbf{G}^{-1}$ , no altera la expresión, quedando

$$\mathbf{R}^{-1}\mathbf{Z}[\mathbf{I}-\mathbf{Q}^{-1}\mathbf{Q}+\mathbf{Q}^{-1}\mathbf{G}^{-1}]\mathbf{G}\mathbf{Z}' = \mathbf{R}^{-1}\mathbf{Z}\mathbf{Q}^{-1}\mathbf{Z}'$$

lo cual es el término 4 con signo contrario.

### **Componentes de Varianza**

Uno de los elementos de las ecuaciones de los modelos mixtos son las varianzas de los efectos aleatorios, las cuales habrá que estimar. Kackar y Harville (1981) probaron que si las varianzas son estimadas bajo ciertas suposiciones, tanto los estimadores de los efectos fijos (BLUE) como los predictores de los efectos aleatorios (BLUP) son insesgados. Los supuestos para obtener estos resultados son:

- Las observaciones se distribuyen simétricamente.

- Los estimadores de las varianzas sean de traslación invariante (no dependen de los efectos fijos).
- Los estimadores de las varianzas sean funciones de las observaciones.

Existen diversos métodos para la estimación de componentes de varianza en los modelos mixtos, brevemente se describirán los más usuales:

Métodos tipo ANOVA. Consisten en igualar los cuadrados medios del análisis de varianza con sus esperanzas. Henderson (1973) presentó tres diferentes métodos: el método I se usa para modelos con efectos aleatorios, el método II, se ajustan los datos por los efectos fijos en el modelo y luego se aplica el método II; el método III, se usa para modelos mixtos y se basa en la suma de cuadrados que se obtienen al ajustar un modelo lineal y sus submodelos. Se analizarán los métodos I y III.

Método I. Consiste en calcular la suma de cuadrados para cada factor en el modelo e igualarlas con sus esperanzas, sólo es aplicable para efectos aleatorios en el modelo. Los pasos para estimar los componentes de varianza son:

- Calcular suma de cuadrados a todos los factores del modelo,
- Igualar los cuadrados medios con sus esperanzas;

- Al resolver el conjunto de ecuaciones considerar los parámetros como sus estimadores (colocar la notación “^”).

Las sumas de cuadrados que se utilizan para la estimación de componentes de varianza, se pueden expresar en forma matricial por medio de formas cuadráticas:

$$\mathbf{y}'\mathbf{Q}\mathbf{y}$$

donde  $\mathbf{Q}$  es simétrica, de no serlo se puede hacer con:

$$\mathbf{y}'\mathbf{Q}\mathbf{y} = \frac{1}{2} \mathbf{y}'\mathbf{Q}\mathbf{y} + \frac{1}{2} \mathbf{y}'\mathbf{Q}'\mathbf{y} = \mathbf{y}'(\frac{1}{2} \mathbf{Q} + \frac{1}{2} \mathbf{Q}')\mathbf{y}$$

Algunas propiedades de las formas cuadráticas. Para  $\mathbf{y} \sim N(\boldsymbol{\mu}, \mathbf{V})$ :

- (i)  $E(\mathbf{y}'\mathbf{Q}\mathbf{y}) = \text{tr}(\mathbf{Q}\mathbf{V}) + \boldsymbol{\mu}'\mathbf{Q}\boldsymbol{\mu}$
- (ii)  $\text{Var}(\mathbf{y}'\mathbf{Q}\mathbf{y}) = 2\text{tr}(\mathbf{Q}\mathbf{V}\mathbf{Q}\mathbf{V}) + 4\boldsymbol{\mu}'\mathbf{Q}\mathbf{V}\mathbf{Q}\boldsymbol{\mu}$
- (iii)  $\text{Cov}(\mathbf{y}'\mathbf{Q}\mathbf{y}, \mathbf{y}'\mathbf{C}\mathbf{y}) = 2\text{tr}(\mathbf{Q}\mathbf{V}\mathbf{C}\mathbf{V}) + 4\boldsymbol{\mu}'\mathbf{Q}\mathbf{V}\mathbf{C}\boldsymbol{\mu}$

Si  $\mathbf{Q}\mathbf{V}\mathbf{C} = \mathbf{0}$ :

- (iva)  $\mathbf{y}'\mathbf{Q}\mathbf{y}$  y  $\mathbf{y}'\mathbf{C}\mathbf{y}$  son independientes
- (ivb)  $\mathbf{Q}\mathbf{y}$  y  $\mathbf{C}\mathbf{y}$  son independientes
- (ivc)  $\mathbf{Q}\mathbf{y}$  y  $\mathbf{y}'\mathbf{C}\mathbf{y}$  son independientes

Si  $\mathbf{Q}\mathbf{V}$  es idempotente

(V)  $\mathbf{y}'\mathbf{Q}\mathbf{y} \sim \chi^2_{(r(\mathbf{Q}), \frac{1}{2}\boldsymbol{\mu}'\mathbf{Q}\boldsymbol{\mu})}$ , donde

$r(\mathbf{Q})$  es el rango de la matriz  $\mathbf{Q}$

$\frac{1}{2}\boldsymbol{\mu}'\mathbf{Q}\boldsymbol{\mu} = \lambda =$  parámetro de no centralidad

(Vi) Si  $\mathbf{Q}\boldsymbol{\mu} = \mathbf{0}$ , entonces  $\mathbf{y}'\mathbf{Q}\mathbf{y} \sim \chi^2_{r(\mathbf{Q})}$

(Vii) Si  $\mathbf{Q}$  es idempotente;  $\mathbf{Q}^2 = \mathbf{Q}$ , el rango de  $\mathbf{Q} = \text{tr}(\mathbf{Q})$

(Viii) Si  $\mathbf{V}$  es  $n \times n$  y su rango es  $n$ , entonces

$$r(\mathbf{QV}) = r(\mathbf{Q})$$

Método III. Se basa en el cálculo de la suma de cuadrados reducidos (SCR) para los diferentes efectos en el modelo, suponiendo un modelo fijo. Las sumas de cuadrados reducidos se igualan a sus esperanzas, tomando en cuenta también los efectos aleatorios; las ecuaciones resultantes se resuelven para estimar los componentes de varianza.

Las SCR se obtienen a partir de las diferencias de las sumas de cuadrados del error (SCE) o de las diferencias de las sumas de cuadrados del modelo (SCM). En general, el cálculo de las SCR no tienen que ser en un cierto orden, una alternativa es seguir las estrategias del SAS. Otra forma de obtener las SCR es mediante el proceso de absorción.

Las soluciones para las ecuaciones de los mínimos cuadrados  $(\hat{\beta}, \hat{\mu}_1, \dots, \hat{\mu}_1)$  se pueden obtener a través del cálculo de la inversa de la matriz de coeficientes, una vez que se hayan aplicado las restricciones correspondientes o con la factorización de Choleski; o bien, con algún algoritmo iterativo. Las SCE se obtiene como la diferencia entre la suma de cuadrados total (SCT) y la suma de cuadrados del modelo (SCM).

### **Máxima verosimilitud (Maximum Likelihood, ML)**

Este método usado para la estimación de parámetros estadísticos, se basa en la selección de una distribución que describa adecuadamente el problema. La distribución más usada es la normal y en ella nos basaremos.

Sea el modelo lineal

$$Y_i = \mu + e_i,$$

donde

$Y_i$ : es la  $i$ -ésima observación

$\mu$ : es una constante común a todas las observaciones

$e_i$ : es la desviación aleatoria de la  $i$ -ésima observación con respecto a la constante  $\mu$ .

$$e_i \sim \text{NID}(0, \sigma^2)$$

Para la variable aleatoria  $Y_i$ , la función de densidad de probabilidades se puede expresar como

$$f(Y_i; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-1/2\sigma^2(Y_i-\mu)^2}$$

y si se tienen  $n$  variables aleatorias independientes  $Y_1, Y_2, \dots, Y_n$ ; la función conjunta de densidad de probabilidades es:

$$f(Y_1, Y_2, \dots, Y_n; \mu, \sigma^2) = f(Y_1; \mu, \sigma^2) \cdot f(Y_2; \mu, \sigma^2) \cdot \dots \cdot f(Y_n; \mu, \sigma^2)$$

$$= \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right)^{n/2} e^{-1/2\sigma^2 \sum_{i=1}^n (Y_i-\mu)^2}$$

La función de verosimilitud  $L(\mu, \sigma^2; Y_1, Y_2, \dots, Y_n)$  es algebraicamente igual que  $f(Y_1, Y_2, \dots, Y_n; \mu, \sigma^2)$ .

A todos los posibles valores que pueden tomar los parámetros  $\mu$  y  $\sigma^2$  se le conoce como espacio paramétrico. Los estimadores de ML de  $\mu$ ,  $\sigma^2$  son aquellos valores  $\hat{\mu}$  y  $\hat{\sigma}^2$  que maximizan la función de verosimilitud, esto es:

$L(\hat{\mu}, \hat{\sigma}^2; Y_1, Y_2, \dots, Y_n) \geq L(\mu, \sigma^2; Y_1, Y_2, \dots, Y_n)$ , para todas las  $\mu$  y  $\sigma^2$  en el espacio paramétrico.

Similarmente,  $\hat{\mu}$  y  $\hat{\sigma}^2$  son los valores que maximizan el logaritmo de la función de verosimilitud:

$$l((\mu, \sigma^2; Y_1, Y_2, \dots, Y_n) = \log L(\mu, \sigma^2; Y_1, Y_2, \dots, Y_n)$$

Para hallar los valores de  $\hat{\mu}$  y  $\hat{\sigma}^2$  que maximicen el logaritmo de la función de verosimilitud, se siguen los siguientes pasos:

- a) Desarrollar el logaritmo de la función de verosimilitud ( $l$ )
- b) Definir el espacio de los parámetros
- c) Hallar las derivadas parciales de  $l$  con respecto a los parámetros
- d) Igualar a cero y resolver la ecuación para los parámetros
- e) Revisar que las soluciones pertenezcan al espacio paramétrico; de ser así, verificar que correspondan a un máximo global de la función de verosimilitud y no a un máximo local.

Para el ejemplo del modelo  $Y_i = \mu + e_i$ :

$$l(\mu, \sigma^2; \mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n) = -\frac{\mathbf{n}}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{\mathbf{n}} (\mathbf{Y}_i - \mu)^2$$

$$\frac{\delta l}{\delta \mu} = \frac{\delta \left( \frac{1}{2\sigma^2} \left[ \sum_{i=1}^{\mathbf{n}} \mathbf{Y}_i^2 - 2 \sum_{i=1}^{\mathbf{n}} \mathbf{Y}_i \mu + \sum_{i=1}^{\mathbf{n}} \mu^2 \right] \right)}{\delta \mu} = \frac{1}{2\sigma^2} \left[ -2 \sum_{i=1}^{\mathbf{n}} \mathbf{Y}_i + 2\mathbf{n}\mu \right]$$

Igualando a cero y resolviendo la ecuación resultante para  $\mu$ , se tiene

$$\hat{\mu} = \frac{\sum_{i=1}^{\mathbf{n}} \mathbf{Y}_i}{\mathbf{n}} = \bar{y}$$

Por otra parte

$$\begin{aligned} \frac{\delta l}{\delta \sigma^2} &= \frac{\delta \left( -\frac{\mathbf{n}}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \left[ \sum_{i=1}^{\mathbf{n}} \mathbf{Y}_i^2 - 2 \sum_{i=1}^{\mathbf{n}} \mathbf{Y}_i \mu + \sum_{i=1}^{\mathbf{n}} \mu^2 \right] \right)}{\delta \sigma^2} \\ &= -\frac{\mathbf{n}}{2\sigma^2} + \frac{2 \sum_{i=1}^{\mathbf{n}} \mathbf{Y}_i^2}{4\sigma^4} - \frac{\sum_{i=1}^{\mathbf{n}} \mathbf{Y}_i \mu}{\sigma^4} + \frac{2 \sum_{i=1}^{\mathbf{n}} \mu^2}{4\sigma^4} \\ &= \frac{-2n\sigma^2 + 2 \sum_{i=1}^{\mathbf{n}} \mathbf{Y}_i^2 - 4 \sum_{i=1}^{\mathbf{n}} \mathbf{Y}_i \mu + 2 \sum_{i=1}^{\mathbf{n}} \mu^2}{4\sigma^4} \end{aligned}$$

Igualando a cero y resolviendo para  $\sigma^2$ , se tiene:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{\mathbf{n}} \mathbf{Y}_i^2 - 2 \sum_{i=1}^{\mathbf{n}} \mathbf{Y}_i \bar{y} + \mathbf{n} \bar{y}^2}{\mathbf{n}} = \frac{\sum_{i=1}^{\mathbf{n}} (\mathbf{Y}_i - \bar{y})^2}{\mathbf{n}}$$

El procedimiento ML, tiene la propiedad de que al estimar componentes de varianza, no considera los gl, que se pierden al estimar los efectos fijos.

### Máxima Verosimilitud Restringida (REML)

El procedimiento de ML consiste en maximizar el logaritmo de la función de verosimilitud de las observaciones ( $\mathbf{y}$ ). Una alternativa es operar con el logaritmo de la función de verosimilitud de los residuales calculados al sustraer los efectos fijos de las observaciones, constituyendo así el procedimiento de máxima verosimilitud restringida (Restricted Maximum Likelihood o Residual Maximum Likelihood, REML). Aquí, en vez de usar el vector de observaciones  $\mathbf{y}$  se seleccionan combinaciones lineales de los elementos en  $\mathbf{y}$  de tal manera que estas combinaciones no contengan ninguno de los efectos fijos. En general, el conjunto de las combinaciones lineales de las observaciones puede representarse como  $\mathbf{K}'\mathbf{y}$ , donde  $\mathbf{K}$  debe ser tal que  $E(\mathbf{K}'\mathbf{y}) = \mathbf{0}$ . De esta forma,  $\mathbf{K}'\mathbf{y} \sim N(\mathbf{0}, \mathbf{K}'\mathbf{V}\mathbf{K})$ . Para nuestro modelo  $\mathbf{Y}_i = \mu + \mathbf{e}_i$ , donde  $\mathbf{Y} = [Y_1, Y_2, \dots, Y_n] \sim N(\mu\mathbf{1}, \sigma^2\mathbf{I})$ , se puede usar el conjunto de combinaciones lineales  $Y_i - Y_n$ , para  $i = 1, 2, \dots, n-1$ . Matricialmente queda:

$$(\mathbf{y}^* - \mathbf{1}y_n) = \mathbf{K}'\mathbf{y}$$

donde:

$$\mathbf{y}^* = [Y_1, Y_2, \dots, Y_n],$$

$$\mathbf{K}'\mathbf{y} \sim N(\mathbf{0}, \mathbf{K}'\mathbf{K}\sigma^2)$$

$\mathbf{K}'$  es una matriz  $(n-1) \times n$  y de la forma

$$\mathbf{K}' = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & -1 \\ 0 & 1 & 0 & \dots & 0 & -1 \\ 0 & 0 & 1 & \dots & 0 & -1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & -1 \end{bmatrix}$$

El logaritmo de la función de verosimilitud de los residuales es:

$l(\sigma^2 | \mathbf{y}^* | \mathbf{y}_n) = -1/2 (n-1) \log 2\pi - 1/2 [\mathbf{K}'\mathbf{K}\sigma^2]^{-1} / 2\mathbf{Y}'\mathbf{K}(\mathbf{K}'\mathbf{K}\sigma^2)^{-1}\mathbf{K}'\mathbf{y}$  y la derivada de  $l$  con respecto a  $\sigma^2$  resulta ser:

$$\frac{\delta l}{\delta \sigma^2} = -1/2 \operatorname{tr} \left[ \mathbf{I} \frac{1}{\sigma^2} - \mathbf{J} \left( \frac{1}{n\sigma^2} \right) \right] + 1/2 \mathbf{y}' \left[ \mathbf{I} \frac{1}{\sigma^2} - \mathbf{J} \left( \frac{1}{n\sigma^2} \right) \right] \left[ \mathbf{I} \frac{1}{\sigma^2} - \mathbf{J} \left( \frac{1}{n\sigma^2} \right) \right] \mathbf{y}$$

donde  $\mathbf{J}$  es una matriz cuadrada de tamaño  $n$ , con todos sus elementos 1 (matriz unidad).

$$\begin{aligned} &= -\frac{n-1}{2\sigma^2} + \frac{\sum_{i=1}^n y_i^2 - \frac{\left(\sum_{i=1}^n Y_i\right)^2}{n}}{2\sigma^4} \\ &= \frac{-(n-1)\sigma^2 + \sum_{i=1}^n y_i^2 - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n}}{2\sigma^4} \end{aligned}$$

Igualando a cero, se tiene:

$$\sigma^2 = \frac{\sum_{i=1}^n y_i^2 - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n}}{n-1} = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1},$$

mostrando que el estimador REML de  $\sigma^2$  sí considera la pérdida de 1 gl al estimar  $\mu$  con  $\hat{y}$ .

Los métodos ML y REML resultan adecuados para estimar componentes de varianza para modelos simples o modelos para conjuntos de datos provenientes de diseños balanceados y con estructuras de (co) varianzas sencillas. Hay otros métodos (iterativos, análisis numéricos) para maximizar o minimizar funciones no lineales para obtener estimaciones de componentes de varianza de máxima verosimilitud.

### **Método Libre de Derivadas (DF-REML)**

El logaritmo de la función de verosimilitud restringida para una distribución normal multivariada es:

$$l = -0.5 (N - \text{rango}(\mathbf{X})) \log(2\pi) + \log |\mathbf{K}' \mathbf{V} \mathbf{K}| + \mathbf{Y}' \mathbf{K} [\mathbf{K}' \mathbf{V} \mathbf{K}]^{-1} \mathbf{K}' \mathbf{y}$$

Harville (1977) y Searle (1979) desarrollaron una forma equivalente de  $l$ :

$$l = -0.5 [\text{constante} + \log |\mathbf{R}| + \log |\mathbf{G}| + \log |\mathbf{C}'| + \mathbf{Y}' \mathbf{P} \mathbf{y}]$$

donde

$\mathbf{C}'$  es la matriz de coeficientes de las MME con rango completo.

$\mathbf{Y}' \mathbf{P} \mathbf{y}$  con  $\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1} \mathbf{X} (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{V}^{-1}$  es la suma de cuadrados residual generalizada.

Smith y Graser (1986) propusieron el método llamado “método libre de derivadas” que consiste en asignar diferentes valores a  $\mathbf{R}$  y  $\mathbf{G}$  (matrices de covarianzas de  $\mathbf{e}$  y  $\mathbf{U}$ ) hasta hallar la combinación que maximice  $l$ .

Existen procedimientos como el método simplex de Nelder y Mead (1965) que permiten sistemáticamente actualizar  $\mathbf{R}$  y  $\mathbf{G}$  iterativamente hasta maximizar  $l$ . La evaluación de  $l$  no es tan simple, ya que involucra el cálculo de determinantes de matrices con estructuras complejas. Van Vleck (1994) describe algunos algoritmos computacionales que permiten evaluar  $l$  eficientemente.

### **El Modelo Mixto como Alternativa de un Diseño de Bloque Incompleto ( $\alpha$ -Látice)**

A continuación, se ilustra con un ejemplo utilizando el paquete estadístico computacional SAS, un diseño de  $\alpha$ -Látice 16 x 16 como un Diseño de Bloques Completamente al Azar, como Látice Convencional, como Modelo Mixto y como  $\alpha$ -Látice (tomado del curso de Experimentación Agrícola Avanzada\*). Tomando la información relevante, se tiene el Cuadro 3.5 siguiente:

---

\* Por cortesía del Dr. Froylán Rincón Sánchez: FIT 605: Notas 14)

Cuadro 3.5. Información relevante de un Diseño de Látice Simple 16 x 16.

| Modelo                     | Ls mean                        | CME    | E.E              | Análisis          |
|----------------------------|--------------------------------|--------|------------------|-------------------|
| DBCA                       | 3.314                          | .5259  | .5128            | GLM               |
| Látice efectos fijos       | 3.628                          | .3248  | .4259            | GLM               |
| Látice (Mixto)             | 3.577                          | .3248. | .4302            | Mixto             |
| $\alpha$ -Látice           | 3.577                          | .3250  | .4031            | $\alpha$ - Látice |
| Diferencia de Tratamientos |                                |        |                  |                   |
| Modelo                     | T <sub>1</sub> -T <sub>2</sub> | E.E.   | Análisis         |                   |
| DBCA                       | -. 657                         | .725   | GLM              |                   |
| Látice efectos fijos       | -. 713                         | .604   | GLM              |                   |
| Látice (Mixto)             | -. 704                         | .599   | Mixto            |                   |
| $\alpha$ -Látice           | -----                          | .725   | $\alpha$ -Látice |                   |

De la información relevante del Cuadro 3.1, se desprenden las siguientes conclusiones:

- (1)- El error estándar (E.E.) del contraste de los tratamientos T<sub>1</sub> y T<sub>2</sub> es menor en el modelo mixto, como se observa en el Cuadro 3.1.
- (2)- El CME del Látice es menor en el modelo mixto que en el modelo convencional.
- (3)- El modelo mixto se apega más al mundo real, es decir representa y describe mejor la realidad de un fenómeno, en base al E.E.

Por lo tanto, se puede resumir que en base a las referencias que se presentan, la mayoría de los autores hacen las mismas recomendaciones que las encontradas en este trabajo, es decir, que los errores estándar de los contrastes entre tratamientos, es mejor por el método del modelo mixto.

Lattice 16x16 analizado como DBCA  
 Modelo fijo

General Linear Models Procedure  
 Class Level Information

| Class | Levels | Values  |
|-------|--------|---|
| REP   | 2      | 1 2   |
| ENT   | 256    | 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22<br>23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41<br>42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60<br>61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79<br>80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98<br>99 100 101 102 103 104 105 106 107 108 109 110 111 112<br>113 114 115 116 117 118 119 120 121 122 123 124 125 126<br>127 128 129 130 131 132 133 134 135 136 137 138 139 140<br>141 142 143 144 145 146 147 148 149 150 151 152 153 154<br>155 156 157 158 159 160 161 162 163 164 165 166 167 168<br>169 170 171 172 173 174 175 176 177 178 179 180 181 182<br>183 184 185 186 187 188 189 190 191 192 193 194 195 196<br>197 198 199 200 201 202 203 204 205 206 207 208 209 210<br>211 212 213 214 215 216 217 218 219 220 221 222 223 224<br>225 226 227 228 229 230 231 232 233 234 235 236 237 238<br>239 240 241 242 243 244 245 246 247 248 249 250 251 252<br>253 254 255 256 |

Number of observations in data set = 512

Lattice 16x16 analizado como DBCA  
 Modelo fijo

General Linear Models Procedure

Dependent Variable: REND

| Source          | DF  | Sum of Squares | Mean Square | F Value | Pr > F |
|-----------------|-----|----------------|-------------|---------|--------|
| Model           | 256 | 676.6684407    | 2.6432361   | 5.03    | 0.0001 |
| Error           | 255 | 134.1069540    | 0.5259096   |         |        |
| Corrected Total | 511 | 810.7753947    |             |         |        |

| R-Square | C.V.     | Root MSE | REND Mean |
|----------|----------|----------|-----------|
| 0.834594 | 21.03853 | 0.725196 | 3.446991  |

| Source | DF  | Type I SS   | Mean Square | F Value | Pr > F |
|--------|-----|-------------|-------------|---------|--------|
| REP    | 1   | 182.0182631 | 182.0182631 | 346.10  | 0.0001 |
| ENT    | 255 | 494.6501777 | 1.9398046   | 3.69    | 0.0001 |

| Source | DF  | Type III SS | Mean Square | F Value | Pr > F |
|--------|-----|-------------|-------------|---------|--------|
| REP    | 1   | 182.0182631 | 182.0182631 | 346.10  | 0.0001 |
| ENT    | 255 | 494.6501777 | 1.9398046   | 3.69    | 0.0001 |

Lattice 16x16 analizado como DBCA  
 Modelo fijo

General Linear Models Procedure  
 Least Squares Means

| ENT | REND<br>LSMEAN |
|-----|----------------|
| 1   | 3.31423500     |
| 2   | 3.97152000     |
| 3   | 4.13082000     |
| 4   | 6.03839500     |
| 5   | 5.81978000     |
| 6   | 8.45157500     |
| 7   | 5.95619000     |
| 8   | 3.29682000     |
| 9   | 3.36343000     |
| 10  | 2.70734000     |
| 11  | 3.75762500     |
| 12  | 3.08470500     |
| 13  | 4.13470000     |
| 14  | 1.82236500     |
| 15  | 2.12911500     |

Lattice 16x16 analizado como DBCA  
 Modelo fijo

General Linear Models Procedure

Dependent Variable: REND

| Parameter    | Estimate    | T for H0:<br>Parameter=0 | Pr >  T | Std Error of<br>Estimate |
|--------------|-------------|--------------------------|---------|--------------------------|
| MEAN TRT1    | 3.31423500  | 6.46                     | 0.0001  | 0.51279120               |
| MEAN TRT2    | 3.97152000  | 7.74                     | 0.0001  | 0.51279120               |
| TRT1 vs TRT2 | -0.65728500 | -0.91                    | 0.3656  | 0.72519627               |
| TRT1 vs TRT3 | -0.81658500 | -1.13                    | 0.2612  | 0.72519627               |

Lattice 16x16 analizado como LATICE  
 Modelo fijo

General Linear Models Procedure  
 Class Level Information

| Levels | Values  |
|--------|---|
| 2      | 1 2   |
| 16     | 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16  |
| 256    | 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22<br>23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41<br>42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60<br>61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79<br>80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98<br>99 100 101 102 103 104 105 106 107 108 109 110 111 112<br>113 114 115 116 117 118 119 120 121 122 123 124 125 126<br>127 128 129 130 131 132 133 134 135 136 137 138 139 140<br>141 142 143 144 145 146 147 148 149 150 151 152 153 154<br>155 156 157 158 159 160 161 162 163 164 165 166 167 168<br>169 170 171 172 173 174 175 176 177 178 179 180 181 182<br>183 184 185 186 187 188 189 190 191 192 193 194 195 196<br>197 198 199 200 201 202 203 204 205 206 207 208 209 210<br>211 212 213 214 215 216 217 218 219 220 221 222 223 224<br>225 226 227 228 229 230 231 232 233 234 235 236 237 238<br>239 240 241 242 243 244 245 246 247 248 249 250 251 252<br>253 254 255 256 |

Lattice 16x16 analizado como LATICE  
 Modelo fijo

General Linear Models Procedure

Variable: REND

|       | DF  | Sum of Squares | Mean Square | F Value | Pr > F |
|-------|-----|----------------|-------------|---------|--------|
|       | 286 | 737.6787894    | 2.5792965   | 7.94    | 0.0001 |
|       | 225 | 73.0966053     | 0.3248738   |         |        |
| Total | 511 | 810.7753947    |             |         |        |

| R-Square | C.V.     | Root MSE | REND Mean |
|----------|----------|----------|-----------|
| 0.909844 | 16.53550 | 0.569977 | 3.446991  |

| DF  | Type I SS   | Mean Square | F Value | Pr > F |
|-----|-------------|-------------|---------|--------|
| 1   | 182.0182631 | 182.0182631 | 560.27  | 0.0001 |
| 30  | 130.7911335 | 4.3597044   | 13.42   | 0.0001 |
| 255 | 424.8693929 | 1.6661545   | 5.13    | 0.0001 |

| DF  | Type III SS | Mean Square | F Value | Pr > F |
|-----|-------------|-------------|---------|--------|
| 1   | 182.0182631 | 182.0182631 | 560.27  | 0.0001 |
| 30  | 61.0103487  | 2.0336783   | 6.26    | 0.0001 |
| 255 | 424.8693929 | 1.6661545   | 5.13    | 0.0001 |

Latice 16x16 analizado como LATICE  
Modelo fijo

General Linear Models Procedure  
Least Squares Means

| ENT | REND<br>LSMEAN |
|-----|----------------|
| 1   | 3.62765094     |
| 2   | 4.34105000     |
| 3   | 3.87073094     |
| 4   | 5.70913250     |
| 5   | 6.30807031     |
| 6   | 8.12437031     |
| 7   | 5.73528906     |
| 8   | 3.11898812     |
| 9   | 3.26932656     |
| 10  | 2.51588594     |
| 11  | 3.13179469     |
| 12  | 2.76808938     |
| 13  | 4.78632313     |
| 14  | 1.38480531     |
| 15  | 2.52151844     |

Latice 16x16 analizado como LATICE  
Modelo fijo

General Linear Models Procedure

Dependent Variable: REND

| Parameter    | Estimate    | T for H0:<br>Parameter=0 | Pr >  T | Std Error of<br>Estimate |
|--------------|-------------|--------------------------|---------|--------------------------|
| MEAN TRT1    | 3.62765094  | 8.52                     | 0.0001  | 0.42599586               |
| MEAN TRT2    | 4.34105000  | 10.19                    | 0.0001  | 0.42599586               |
| TRT1 vs TRT2 | -0.71339906 | -1.18                    | 0.2392  | 0.60455192               |
| TRT1 vs TRT3 | -0.24308000 | -0.40                    | 0.6880  | 0.60455192               |

Lattice 16x16 analizado como LATICE  
 Modelo Mixto (BLK=ALEATORIO)

The MIXED Procedure  
 Class Level Information

| Class | Levels | Values   |
|-------|--------|--|
| REP   | 2      | 1 2  |
| BLK   | 16     | 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16   |
| ENT   | 256    | 1 2 3 4 5 6 7 8 9 10 11 12 13<br>14 15 16 17 18 19 20 21 22 23<br>243 244 245 246 247 248 249<br>250 251 252 253 254 255 256 |

| Cov Parm | Covariance Parameter Estimates (REML) |            |            |       |         |
|----------|---------------------------------------|------------|------------|-------|---------|
|          | Ratio                                 | Estimate   | Std Error  | Z     | Pr >  Z |
| BLK(REP) | 0.65748780                            | 0.21360056 | 0.06574826 | 3.25  | 0.0012  |
| Residual | 1.00000000                            | 0.32487380 | 0.03062940 | 10.61 | 0.0001  |

| Tests of Fixed Effects |     |     |            |        |  |
|------------------------|-----|-----|------------|--------|--|
| Source                 | NDF | DDF | Type III F | Pr > F |  |
| ENT                    | 255 | 225 | 5.15       | 0.0001 |  |
| REP                    | 1   | 30  | 48.64      | 0.0001 |  |

ESTIMATE Statement Results

| Parameter    | Estimate    | Std Error  | DDF | T     | Pr >  T |
|--------------|-------------|------------|-----|-------|---------|
| MEAN TRT1    | 3.57758371  | 0.43024019 | 225 | 8.32  | 0.0001  |
| MEAN TRT2    | 4.28201873  | 0.43024019 | 225 | 9.95  | 0.0001  |
| TRT1 vs TRT2 | -0.70443502 | 0.59916261 | 225 | -1.18 | 0.2410  |
| TRT1 vs TRT3 | -0.33469565 | 0.59916261 | 225 | -0.56 | 0.5770  |

| Least Squares Means |            |            |     |       |         |
|---------------------|------------|------------|-----|-------|---------|
| Level               | LSMEAN     | Std Error  | DDF | T     | Pr >  T |
| ENT 1               | 3.57758371 | 0.43024019 | 225 | 8.32  | 0.0001  |
| ENT 2               | 4.28201873 | 0.43024019 | 225 | 9.95  | 0.0001  |
| ENT 3               | 3.91227936 | 0.43024019 | 225 | 9.09  | 0.0001  |
| ENT 4               | 5.76173116 | 0.43024019 | 225 | 13.39 | 0.0001  |
| ENT 5               | 6.23006745 | 0.43024019 | 225 | 14.48 | 0.0001  |
| ENT 6               | 8.17664025 | 0.43024019 | 225 | 19.00 | 0.0001  |
| ENT 7               | 5.77057730 | 0.43024019 | 225 | 13.41 | 0.0001  |
| ENT 8               | 3.14739622 | 0.43024019 | 225 | 7.32  | 0.0001  |
| ENT 9               | 3.28435929 | 0.43024019 | 225 | 7.63  | 0.0001  |
| ENT 10              | 2.54647013 | 0.43024019 | 225 | 5.92  | 0.0001  |
| ENT 11              | 3.23176914 | 0.43024019 | 225 | 7.51  | 0.0001  |
| ENT 12              | 2.81866774 | 0.43024019 | 225 | 6.55  | 0.0001  |
| ENT 13              | 4.68222835 | 0.43024019 | 225 | 10.88 | 0.0001  |

#PLOTS= 512 #REPS= 2 # VARS= 256 # BLOCKS/REP= 16  
 EFFY= 0.9999 CRIT= 0.00 MISS. VAL. CODE=-99.

EFICIENCIA DE LATICE 16X16

ANALYSIS OF VARIANCE FOR VARIABLE REND

LATTICE ADJUSTED ANALYSIS

| SOURCE   | DF  | MS       |
|----------|-----|----------|
| REPS     | 1   | 182.0472 |
| BLK(ADJ) | 30  | 2.0328   |
| TRT      | 255 | 1.9398   |
| TRT(ADJ) | 255 | 1.6662   |
| RESID    | 225 | 0.3250   |
| TOTAL    | 511 | 1.5867   |

F (REPS)= 560.142 ( 1, 225 DF)  
 F (TRTS)= 5.127 ( 255, 225 DF)

AV. SE 0.4031 LSD(0.05) 1.1234 F SIG 1.0 DF 225. CV 16.5

RCBD ANALYSIS

| SOURCE  | D.F. | SUMS OF SQUARES | MEAN SQUARES | F VALUES |
|---------|------|-----------------|--------------|----------|
| REPS    | 1    | 182.04930000    | 182.04930000 | 346.138  |
| ENTRIES | 255  | 494.64890000    | 1.93980000   | 3.688    |
| ERROR   | 255  | 134.11570000    | 0.52594400   |          |
| TOTAL   | 511  | 810.81400000    |              |          |

REP MEANS  
 1 4.043  
 2 2.851

S.E. (DIFF) 0.72522 ( 255 ) D.F.  
 COEFF. OF VARIATION 21.04 OVERALL MEAN 3.4469

RELATIVE EFFICIENCY RCBD/LATT ADJ:  
 CALCULATED AS (SED(RCBD)\*\*2/SED(LATT.ADJ)\*\*2) = 1.6182

\*\*\*\*\*

1 UNADJUSTED ADJUSTED MEANS

| ENT  | REND  |       |
|------|-------|-------|
| 1    | 3.32  | 3.58  |
| 2    | 3.97  | 4.28  |
| 3    | 4.14  | 3.92  |
| 4    | 6.04  | 5.76  |
| 5    | 5.82  | 6.23  |
| 6    | 8.45  | 8.17  |
| 7    | 5.96  | 5.77  |
| 8    | 3.30  | 3.15  |
| 9    | 3.36  | 3.28  |
| 10   | 2.70  | 2.54  |
| .    | .     | .     |
| .    | .     | .     |
| .    | .     | .     |
| .    | .     | .     |
| 252  | 3.82  | 3.80  |
| 253  | 4.17  | 4.48  |
| 254  | 5.28  | 4.86  |
| 255  | 3.37  | 3.14  |
| 256  | 4.57  | 4.29  |
| MEAN | 3.45  | 3.45  |
| LSD  | 2.02  | 1.12  |
| CV   | 21.04 | 16.54 |
| FSIG | 1.00  | 1.00  |
| REFF |       | 1.62  |

## IV. EL MODELO DE BLOQUE ALEATORIZADO

### Introducción

Se mencionan algunos aspectos relevantes del modelo de bloque aleatorizado, siendo quizás el más útil de todos los modelos del diseño experimental y referido en la literatura como: "*two- way classification model*" o "*row by column model*" o "*two- way layout model*".

Existen muchos casos especiales de este modelo (Graybill, 1976) dependiendo de si presenta o no interacción y del número de observaciones disponibles (celdas) para cada combinación de los niveles de cada factor.

- a) En este modelo, interesa la comparación de los efectos fijos entre medias de los tratamientos para, probar hipótesis. El modelo del bloque aleatorizado, por esta razón, es clasificado como modelo fijo o modelo I.
- b) Se pretende minimizar la varianza de la media de los tratamientos ( $\bar{y}_i$ ), para luego aplicar la inferencia estadística, siendo necesario para ello obtener el error estándar y, finalmente, hacer predicciones.
- c) Este modelo consiste que los bloques son fijos, y la prueba F es:

$$F = \frac{CM(\text{blq})}{CME} \text{ (como efectos fijos)}$$

- d) El modelo es aditivo en los efectos de los bloques y de los tratamientos.
- e) No existe interacción entre bloque y tratamientos.
- f) El error experimental es aleatorio y se distribuye normalmente, como media cero y varianza  $\sigma^2$ . Aquí

$$\text{Var}(\bar{y}_i) = \frac{\sigma^2}{r}$$

En base a los aspectos considerados anteriormente y del concepto de aleatorización tan presente en el modelo aleatorio (modelo II) así como en el modelo mixto (modelo III), podemos sustentar (Rincón, 1998) que:

- a) En el modelo de efectos fijos, se utiliza la comparación entre medias de tratamientos para probar hipótesis.
- b) El modelo de efectos aleatorios, utiliza la comparación entre la diferencia de varianzas de medias.
- c) El modelo mixto, utiliza la comparación entre la diferencia de varianzas de medias, debido también a que el efecto de los bloques es fijo:
  - i) Las unidades experimentales son homogéneas dentro de los bloques, pero.
  - ii) Las unidades experimentales son heterogéneas entre los bloques.

Si el bloque es fijo,  $\text{Var}(\bar{y}_{i.}) = \frac{\sigma^2}{r}$  (modelo I)

Si el bloque es aleatorio  $\text{Var}(\bar{y}_{i.}) = \frac{\sigma_R^2 + \sigma^2}{b}$

con

$$\text{E.E.} = \frac{\sigma_R^2 + \sigma^2}{b}$$

### **Modelo de Bloque Aleatorizado, no Interacción, una Observación por Celda**

**Teorema 1.** Considerar el modelo de bloque aleatorizado con no interacción y una observación por celda, dada en la ecuación

$$Y_{ij} = \mu + \alpha_i + \tau_j + E_{ij} \quad i = 1, 2, \dots, a > 1, j = 1, 2, \dots, t > 1$$

La combinación lineal  $\sum c_i \alpha_i$  es estimable sí y sólo sí  $\sum c_i \alpha_i$  es un contraste de los  $\alpha_i$ 's. El mejor estimador insesgado es  $\sum c_i \hat{\alpha}_i = \sum c_i \bar{Y}_i$ .

Los mismos resultados aplicados a los  $\tau_j$ 's y el mejor estimador de una contraste  $\sum d_j \tau_j$  es  $\sum d_j \hat{\tau}_j = \sum d_j \bar{y}_{.j}$ . En particular, el mejor estimador insesgado de

$\alpha_i^* = \alpha_i - \bar{\alpha}$ . es  $\hat{\alpha}_i^* = \alpha_i - \bar{\alpha} = \bar{y}_{i.} - \bar{y}_{..}$  para cada  $i = 1, 2, \dots, a$  y de

$\tau_j^* = \tau_j - \bar{\tau}$ . es  $\hat{\tau}_j^* = \tau_j - \bar{\tau} = \bar{y}_{.j} - \bar{y}_{..}$  para cada  $j = 1, 2, \dots, t$

El mejor estimador insesgado de

$\mu + \alpha_i + \bar{\tau}$ . es  $\bar{y}_{i.}$  para cada  $i = 1, 2, \dots, a$  y de

$\mu + \bar{\alpha} + \tau_j$  es  $\bar{y}_{.j}$  para cada  $j = 1, 2, \dots, t$

En seguida, se establece un teorema acerca del estimador de  $\sigma^2$  y luego un teorema acerca de las propiedades distribucionales de los estimadores de  $\sum_{ci}\alpha_i$ ,  $\sum_{di}\tau_j$ , y  $\sigma^2$ .

**Teorema 2.** Considerar el modelo

$$Y_{ij} = \mu + \alpha_i + \tau_j + \varepsilon_{ij}, \quad i = 1, 2, \dots, a > 1, \quad j = 1, 2, \dots, t > j$$

Para el caso I:  $\varepsilon_{ijm} \sim \text{NID}(\varepsilon : 0, \sigma^2)$  para todo valor posible de  $i, j, m$ .

El estimador UMVU (estimador de  $\ell' \hat{\beta}$ ) de  $\sigma^2$  es

$$\hat{\sigma}^2 = \sum_{j=1}^t \sum_{i=1}^a \frac{(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2}{(a-1)(t-1)}$$

**Teorema 3.** Considerar el modelo

$$Y_{ij} = \mu + \alpha_i + \tau_j + \varepsilon_{ij}, \quad i = 1, 2, \dots, a > 1, \quad j = 1, 2, \dots, t > 1,$$

para el caso I:  $\varepsilon_{ijm} \sim \text{NID}(\varepsilon : 0, \sigma^2)$  para todo valor posible de  $i, j, m$ .

Los resultados son:

- $Z_1 = \sum c_i \hat{\alpha}_i = \sum c_i \bar{y}_i$  está distribuida  $N(Z_1: c_i \alpha_i, \sigma^2 \sum c_i^2 / t)$
- si  $\sum c_i = 0$ ;
- $Z_2 = \sum d_j \hat{\tau}_j = \sum d_j \bar{Y}_j$  está distribuida  $N(Z_2: \sum d_j \tau_j, \sigma^2 \sum d_j^2 / a)$
- si  $\sum d_j = 0$ ;
- $U = (a-1)(t-1) \hat{\sigma}^2 / \sigma^2$  está distribuida  $\chi^2 (u: (a-1)(t-1))$ ;
- $U, Z_1,$  y  $Z_2$  son mutuamente independientes;
- Con  $\left[ \sum c_i \hat{\delta}_i, \sum c^* i \hat{\delta}_i \right] = \sigma^2 \sum c_i c^* i / t$  si  $\sum c_i = \sum c^* i = 0$ ;
- Con  $\left[ \sum d_j \hat{\tau}_j, \sum d_j^* \hat{\tau}_j \right] = \sigma^2 \sum d_j d_j^* / a$  si  $\sum d_j = \sum d_j^* = 0$ ;

**Teorema 4.** Considera el modelo

$$Y_{ij} = \mu + \delta_i + \tau_j + \varepsilon_{ij}, \quad i=1,2,\dots, a>1, \quad j=1,2,\dots, t>1;$$

Para el caso:

$\varepsilon_{ijm} \sim \text{NID} (\varepsilon : 0, \sigma^2)$  para todo valor posible de  $i, j, m$ .

Los resultados se expresan:

- Uno a la vez  $1 - \delta$  intervalo de confianza en  $\sum_{d_j} \tau_j$ , donde  $\sum_{d_j} = 0$ , es

$$\sum_{d_j} \bar{Y}_{.j} \mp \tau_{\delta/2: (a-1)(t-1)} \hat{\sigma} \sqrt{\sum d_j^2 / a}$$

- Simultáneos  $1 - \delta$  intervalos de confianza en  $t_j - t_j'$  para  $j = 1, 2, \dots, t, j' = 1, 2, \dots, t, j \neq j'$  son

$$\bar{Y}_{.j} - \bar{Y}_{.j} \mp q \delta : t, (a-1)(t-1) \hat{\sigma} / \sqrt{a}$$

La hipótesis que interesa en este modelo es  $H_0 : \tau_1 = \tau_2 = \dots = \tau_t$  vs

$H_a$ : al menos uno igual es una desigualdad, es una hipótesis estimable.

- Una verosimilitud generalizada relativa de la prueba estadística para la hipótesis en (2) es  $W_T$ , donde

$$W_T = \frac{\sum \sum (\bar{Y}_{.j} - \bar{Y}_{..})^2}{\frac{t-1}{(a-1)(t-1)} \sum \sum (Y_{ij} - \bar{Y}_{.j} + \bar{Y}_{..})^2}$$

- Para un tamaño  $\delta$  la prueba de  $H_0$  vs  $H_a$  en (2), rechaza  $H_0$  si y sólo si

$$w_T \geq F_{\delta : t-1, (a-1)(t-1)};$$

$W_T$  es valor calculado del estadístico de prueba.

## LITERATURA CITADA

- Anderson, V.L. and Mc Lean, R.A. 1974. Design of Experiments, A Realistic Approach. Marcel Dekker, Inc. New York.
- Ayres, F.J. 1983. Teoría y Problemas de Matrices. McGraw Hill, Serie Schaum, México.
- Berenson M.L, Levine D.M. 1992. Estadística Básica en Administración. Prentice - Hall Hispanoamericana, S.A., México.
- Blaine J. Sergio Rdz, Walter Stroup., 1992. Recent Developments in Design and Analysis of corn Yield Trials. University of Nebraska, Lincoln, Nebraska, U.S.A.
- Dhrymes P.J. 1984. Econometría. Edit. AC, Madrid, España.
- Gray - bill, F.A. 1976, Theory and Application of the Linear Model. Wadsworth Publishing Company, Pasific Grove, California, U.S.A.
- Hadley, G. 1969. Linear Algebra. Fondo Educativo Interamericano, S.A., Bogotá.
- Henderson, C.R. 1949. Estimation of changes in herd environment. J. Dairy Sci. 32:706
- \_\_\_\_\_. 1975. Best linear unbiased prediction under a selection model. Biometrics 31:423
- \_\_\_\_\_. 1973. Sire evaluation and genetic trends. In : Proc. Symp. In honor Dr. Jay L. Lush , Amer. Soc. Animal Sci. And Amer. Dairy Assn., Champaign, 1 L.
- Hicks, Ch. R. 1964. Fundamental Concepts in the Design of Experiments. Holt Rinehart Winston, N.Y., U.S.A.
- Haffman. K y R. Kunze. 1992 Algebra Lineal. Mc Graw - Hill Serie Schaum, México.
- Méndez R., I. 1976. Modelos Estadísticos Lineales. Interpretación y Aplicaciones FOCCA VI/CONACYT, México.

- Morones R.E. 1996. Apuntes del Curso de Diseño Experimentales, Centro de Estadística y Cálculo, UAAAN, Saltillo, México.
- Nering, E.D. 1977. Algebra Lineal y Teoría de Matrices Editorial Limusa, México.
- Nuñez, R. y Rodríguez F. 1995. Curso Intensivo de introducción a la metodología de los modelos mixtos. Universidad Autónoma de Chihuahua, Chihuahua, México.
- Padrón C., E. 1996. Diseño Experimentales con Aplicación a la Agricultura y la Ganadería, Editoriales Trillas, México.
- Priore E. Franco J. Crossa J., Dartayete M. 1995. Curso de Modelos Lineales en la Agricultura. Universidad de la República, Facultad de Agronomía, Uruguay.
- Searle S.R. Casella G, Mc Culloch Ch. E. 1992. Variance Components. Wiley y Sons, Inc. New York.
- Searles. R. 1971. Linear Models. Wiley y Sons, Inc. New York.
- Steel R.G., Torrie J.H. 1980. Bioestadística. Mc Graw – Hill, México.
- Stroup. W.W., Mc Lean R.A. 1989 Applications of Mixed Models in Agriculture and Related Disciplines. U.S.A.
- Zyskind, G. 1975. Teoría General de las Hipótesis Lineales. Colegio de Postgraduados, Chapingo, México.