



Universidad Nacional Autónoma de México

Facultad de Ciencias

"COMPARACION DE METODOS PARA
ANALISIS DE DATOS EN TABLAS DE
DOBLE ENTRADA CON UNA OBSER-
VACION POR CELDA"

T E S I S

Que para obtener el título de

A C T U A R I O

p r e s e n t a :

José Antonio Flores Díaz

México, D. F.

1977



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A mi abuelo

Don Ezequiel Díaz M.

A mis padres

A mis hermanos

En verdad no debes nada a ningún hombre.

Debes todo a todos los hombres.

Gibrán Jalil Gibrán.

Deseo hacer patente mi reconocimiento, a quienes de una u otra forma colaboraron en la realización de este trabajo, especialmente:

Al Dr. Ignacio Méndez Ramírez por haber sugerido y dirigido la elaboración del mismo:

Al Actuario Manuel Francisco Román Enríquez por su constante apoyo y

Al Actuario Miguel Cervera Flores quien con sus cursos de Probabilidad y Estadística me motivó a realizar el trabajo que -- hoy presento.

J.A.F.D.

INDICE

	Página.
Introducción.	1
Capítulo I. Revisión de modelos lineales.	3
Capítulo II. Revisión de modelos para datos con dos criterios de clasificación.	28
Capítulo III. Análisis de la interacción en datos con dos criterios de clasificación y una observación por celda.	46
Capítulo IV. Comparación mediante simulación, de varios enfoques para el análisis de tablas de doble entrada.	62
Notas Encuadradas.	73
Apéndice A. Listado del programa.	77
Apéndice B. Gráficas de resultados.	101

INTRODUCCION.

Esta tesis es una comparación de los métodos estadísticos más importantes que para el análisis de tablas de doble entrada y una observación por celda existen.

Se espera que sea útil para todas aquellas personas, que se enfrenten con este tema, muy en particular, para los estudiantes que deseen ver un poco más "a fondo", el tema de la interacción en tablas de doble entrada.

El contenido del capítulo I es, una revisión general de los resultados más importantes que para modelos de rango completo e incompleto existe.

El capítulo II presenta la teoría de los modelos de rango incompleto para datos con dos criterios de clasificación con una y con varias observaciones por celda.

El capítulo III bosqueja la teoría relacionada con cada uno de los enfoques que se analizan en este trabajo.

Para finalizar, el capítulo IV trata los detalles más generales -- efectuados, al realizar la simulación que ahí mismo se comenta y -- concluye recomendando ciertas direcciones a seguir en futuros estudios.

CAPITULO I

REVISION DE MODELOS LINEALES^[1]

Parte 1.- Modelos de Rango Completo.

1.- Definiciones y suposiciones.

Consideremos la función de frecuencia $f(y; x_1, x_2, \dots, x_q; \beta_1, \beta_2, \dots, \beta_q)$ de una variable aleatoria y , la cual dependerá de q cantidades conocidas $x_1, x_2, x_3, \dots, x_q$ y de q parámetros desconocidos $\beta_1, \beta_2, \beta_3, \dots, \beta_q$.

Supondremos que y tiene una media igual a $\sum_{i=1}^q \beta_i x_i$ y una varianza de σ^2 también desconocida, la cual no dependen ni de β_i o x_i para $i = 1, 2, 3, \dots, q$.

Si se hace la transformación $e = y - \sum_{i=1}^q \beta_i x_i$ entonces e es una variable aleatoria con media cero y varianza σ^2 . Esta última ecuación puede ser escrita como

$$y = \sum_{i=1}^q \beta_i x_i + e$$

llamándola en este caso "ecuación de predicción".

Porque en general las β_i son desconocidas, tendremos que estimarlas tomando una muestra de tamaño n , denotándola por $p_1, p_2, p_3, \dots, p_n$, donde $p_j = (y_j; x_{j1}, x_{j2}, x_{j3}, \dots, x_{jq})$ representará una observación de $f(y)$, donde los valores correspondientes a las variables x_{ji} serán considerados constantes conocidas.

Por lo tanto el sistema de observaciones es:

$$y_j = \sum_{i=1}^q \beta_i x_{ji} + e_j \quad j = 1, 2, \dots, n$$

o en forma vectorial

$$\bar{y} = \bar{x}\bar{\beta} + \bar{e}$$

donde

$\bar{y} \equiv$ vector de n observaciones aleatorio.

$\bar{x} \equiv$ matriz de elementos conocidos fijos de rango q ($q < n$)

$\bar{\beta} \equiv$ vector de q parámetros desconocidos

$\bar{e} \equiv$ vector aleatorio de n componentes.

La consideración de las restantes secciones esta sujeta a la su posición que se haga en torno a la distribución del vector $\bar{e}$.

Las suposiciones son las siguientes:

Caso 1.- El vector $\bar{e}$ se distribuye como una $N(\bar{0}, \sigma^2 \bar{I})$, donde σ^2 es desconocida.

Caso 2.- $\bar{e}$ es un vector aleatorio tal que $E(\bar{e}) = \bar{0}$ y $cov(\bar{e}) = E(\bar{e}\bar{e}') = \sigma^2 \bar{I}$, donde σ^2 es desconocida.

Ambos casos serán considerados en el siguiente inciso, dejando el caso 1. o de la "teoría normal" para el resto.

II.- Estimación puntual de σ^2, β_i y funciones lineales de β_i .

Caso 1).-

De acuerdo con la suposición en turno del vector $\bar{e}$, estimaremos los parámetros $\beta_1, \beta_2, \beta_3, \dots, \beta_p$ y σ^2 , por el método de máxima-verosimilitud.

La ecuación de verosimilitud es

$$f(\bar{e}; \bar{\beta}, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp \left[-\frac{(\bar{y} - \bar{x}\bar{\beta})'(\bar{y} - \bar{x}\bar{\beta})}{2\sigma^2} \right]$$

usando logaritmos se obtiene

$$\log f(\bar{e}; \bar{\beta}, \sigma^2) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} (\bar{y} - \bar{x}\bar{\beta})'(\bar{y} - \bar{x}\bar{\beta})$$

siendo los estimadores de máxima-verosimilitud de $\bar{\beta}$ y σ^2 las soluciones a las ecuaciones.

$$\frac{\partial}{\partial \bar{\beta}} [\log f(\bar{e}; \bar{\beta}, \sigma^2)] = \bar{0} \quad \frac{\partial}{\partial \sigma^2} [\log f(\bar{e}; \bar{\beta}, \sigma^2)] = 0$$

siendo las soluciones las siguientes:

$$\bar{x}'\bar{x}\tilde{\bar{\beta}} = \bar{x}'\bar{y} \quad \text{y} \quad \tilde{\sigma}^2 = \frac{1}{n} (\bar{y} - \bar{x}\tilde{\bar{\beta}})'(\bar{y} - \bar{x}\tilde{\bar{\beta}})$$

La ecuación matricial $\bar{X}'\bar{X}\tilde{\beta} = \bar{X}'\bar{Y}$ es llamada o conocida también -- como las "ecuaciones normales", jugando un papel importante en -- toda la teoría.

Como $\bar{X}$ es de rango q esto implica que $\bar{X}'\bar{X}$ tenga inversa y -- por lo tanto que sea

$$\tilde{\beta} = (\bar{X}'\bar{X})^{-1}\bar{X}'\bar{Y}$$

Frecuentemente al producto matricial $\bar{X}'\bar{X}$ se le denotará por la matriz S .

Por ser estimadores de máxima-verosimilitud, son consistentes y eficientes pero sólo $\tilde{\beta}$ es insesgado, se puede obtener sin embargo el estimador insesgado de σ^2 que es:

$$\hat{\sigma}^2 = \frac{n}{n-q} \tilde{\sigma}^2 = \frac{(\bar{Y} - \bar{X}\hat{\beta})'(\bar{Y} - \bar{X}\hat{\beta})}{n-q}$$

En este momento se puede demostrar que $\hat{\sigma}^2$ y $\hat{\beta}$ son estimadores suficientes, completos y por lo tanto de varianza mínima y que -- además:

$$\begin{aligned} \hat{\beta} & \text{ se distribuye como } N(\bar{\beta}, \sigma^2 S^{-1}) \\ (n-q) \frac{\hat{\sigma}^2}{\sigma^2} & \text{ se distribuye como } \chi^2_{(n-q)} \\ \text{y } \hat{\beta} & \text{ y } (n-q) \frac{\hat{\sigma}^2}{\sigma^2} \text{ son independientes.} \end{aligned}$$

Caso 2).--

En este caso el supuesto no permite el empleo del método anteriormente usado por lo que se recurrirá al de "mínimos cuadrados".

Aquí también se trata de encontrar el valor de $\tilde{\beta}$ que minimice -- la suma de cuadrados siguiente $\sum_{i=1}^n e_i^2$.

La suma anterior vista vectorialmente da:

$$\sum_{i=1}^n e_i^2 = \bar{e}'\bar{e} = (\bar{Y} - \bar{X}\tilde{\beta})'(\bar{Y} - \bar{X}\tilde{\beta})$$

por lo tanto el valor de $\bar{\beta}$ que minimiza $\bar{e}'\bar{e}$ está dado por la solución a la ecuación:

$$\frac{\partial (\bar{e}'\bar{e})}{\partial \bar{\beta}} = 0$$

y ésto nos da que $\hat{\beta} = \bar{S}^{-1}\bar{X}'\bar{Y}$, el cual coincide con el estimador de máxima-verosimilitud bajo la "teoría normal".

La minimización de la suma de cuadrados $\bar{e}'\bar{e}$ no proporciona ningún estimador de σ^2 , sin embargo, el estimador insesgado de σ^2 basado en el obtenido para $\hat{\beta}$ en este mismo método está dado por

$$\hat{\sigma}^2 = \frac{(\bar{Y} - \bar{X}\hat{\beta})'(\bar{Y} - \bar{X}\hat{\beta})}{n - q}$$

En este caso, es de hacerse notar, que no es posible analizar la "bondad" del estimador $\hat{\beta}$ con respecto a todas las funciones. Sin embargo existe un teorema que demuestra que los estimadores de mínimos cuadrados en el modelo en cuestión, tienen una "mejor" bondad que cualquier otro estimador que esté en función específica de las observaciones $\bar{Y}$, este teorema es el de Gauss-Markoff, el cual simplemente se enunciará.

Teorema de Gauss-Markoff, 1.-

Si en la hipótesis del modelo lineal general de rango completo $\bar{Y} = \bar{X}\beta + \bar{e}$, el vector $\bar{e}$ se haya condicionado a:

- 1) $E(\bar{e}) = 0$
- 2) $E(\bar{e}\bar{e}') = \sigma^2 \bar{I}$

entonces, el mejor estimador lineal insesgado de β está dado por mínimos cuadrados.

Demostración: Ver referencia 1.

Hasta el momento se ha hablado únicamente de la estimación de las β_i , sin embargo, muy frecuentemente se requiere estimar ciertas funciones de las β_i . En el caso 1 no existe problema alguno debido a la "invarianza" de los estimadores, siendo algo directo el procedimiento de obtención de un estimador de cualquier transformación uno a uno de las β_i , esto mismo bajo el supuesto 2 no es tan simple; existe sin embargo un caso importante en el que la invarianza de los estimadores mínimos cuadrados se retiene, es en el caso de una transformación lineal de las β_i , lo cual se demostrará en el siguiente teorema.

Teorema 2.-

Bajo la hipótesis del modelo lineal general dada en el teorema anterior, el mejor estimador lineal insesgado de cualquier combinación lineal de las β_i , es la misma combinación lineal de los mejores estimadores lineales insesgados de las β_i .

Demostración.-

Sea $\bar{t} \neq \bar{0}$ un vector constante conocido de $q \times 1$ y $\bar{A}$ una matriz de $q \times n$, la cual tendremos que determinar, además sea $\beta^* = \bar{t}' \bar{A} \bar{Y}$ una función lineal general de $\bar{Y}$, la cual tomaremos como una estimación de $\bar{t}' \bar{\beta}$.

Sea $\bar{A}$ de tal forma que β^* sea el mejor estimador insesgado de $\bar{t}' \bar{\beta}$, en particular $\bar{A} = \bar{S}^{-1} \bar{X}' + \bar{B}$, como $\bar{S}^{-1} \bar{X}'$ es conocido, lo que nos hace falta es determinar $\bar{B}$.

Tomando la esperanza del estimador vemos que:

$$E(\beta^*) = E(\bar{t}' \bar{A} \bar{Y}) = \bar{t}' (\bar{S}^{-1} \bar{X}' + \bar{B}) \bar{X} \bar{\beta} = \bar{t}' \bar{\beta} + \bar{t}' \bar{B} \bar{X} \bar{\beta}$$

ahora, para que β^* sea insesgado tenemos que $\bar{t}' \bar{B} \bar{X} \bar{\beta} = 0$, en particular $\bar{B} \bar{X} = \bar{0}$.

Por la propiedad requerida de que sea "el mejor" estimador, debemos encontrar la matriz $\bar{A}$ que minimice la varianza, considerando la restricción $\bar{B} \bar{X} = \bar{0}$. Por lo que analizaremos la varianza:

$$\begin{aligned} \text{Var}(\beta^*) &= E[(\beta^* - \bar{t}' \bar{\beta})(\beta^* - \bar{t}' \bar{\beta})'] \\ &= E\left\{ [\bar{t}' (\bar{S}^{-1} \bar{X}' + \bar{B}) \bar{Y} - \bar{t}' \bar{\beta}] [\bar{t}' (\bar{S}^{-1} \bar{X}' + \bar{B}) \bar{Y} - \bar{t}' \bar{\beta}]' \right\} \end{aligned}$$

usando el hecho de que $\bar{B}\bar{X}=\bar{O}$ y sustituyendo $\bar{Y}$ por $\bar{X}\bar{\beta}+\bar{e}$ obtenemos que:

$$\begin{aligned} \text{Var}(\beta^*) &= E \{ \bar{t}' \bar{S}^{-1} \bar{X}' \bar{e} \bar{e}' \bar{X} \bar{S}^{-1} \bar{t} + \bar{t}' \bar{S}^{-1} \bar{X}' \bar{e} \bar{e}' \bar{B}' \bar{t} + \\ &\quad \bar{t}' \bar{B} \bar{e} \bar{e}' \bar{X} \bar{S}^{-1} \bar{t} + \bar{t}' \bar{B} \bar{e} \bar{e}' \bar{B}' \bar{t} \} = \\ &= \sigma^2 \{ \bar{t}' \bar{S}^{-1} \bar{t} + \bar{t}' \bar{B} \bar{B}' \bar{t} \} = \sigma^2 \{ \bar{t}' (\bar{S}^{-1} + \bar{B} \bar{B}^{-1}) \bar{t} \} \end{aligned}$$

Sea $\bar{B} \bar{B}' = \bar{G} = (g_{ij})$. Entonces tenemos que la $\text{Var}(\beta^*) = \sigma^2 \{ \bar{t}' (\bar{S}^{-1} + \bar{G}) \bar{t} \}$.

Observemos que si queremos minimizar la cantidad anterior, debemos minimizar $\bar{G}$ ya que los demás términos son constantes.

Por lo tanto, debemos encontrar una $\bar{G}$ tal que todo elemento de $\bar{G}$ sea mínimo.

Pero $\bar{G} = \bar{B} \bar{B}'$ es positiva semidefinida, de aquí que $g_{ii} \geq 0$ para $i, j = 1, 2, \dots, p$.

En particular observamos que para $i=j$ $g_{ii} = \sum_k b_{ik}^2$ y si $g_{ii} = 0$ para toda i , esto implica que $b_{ij} = 0$ para toda i e j , condición -- que también implica que $g_{ij} = \sum_k b_{ik} b_{jk} = 0$ y por lo tanto $\bar{B} = \bar{O}$.

La conclusión final ($\bar{B} = \bar{O}$) es compatible con la de que $\bar{B}\bar{X} = \bar{O}$ y por lo tanto $\bar{A} = \bar{S}^{-1}\bar{X}$ y $\beta^* = \bar{t}' \bar{S}^{-1} \bar{X}' \bar{Y}$.

III.- Prueba de las hipótesis de que $\beta_1 = \beta_1^*, \beta_2 = \beta_2^*, \dots, \beta_p = \beta_p^*$ donde las β_i^* son constantes dadas.

De aquí en adelante únicamente estamos considerando el caso 1 o de la teoría normal.

Probar que $\bar{\beta} = \bar{\beta}^*$ ($\bar{\beta}^*$ conocido) es lo mismo que probar que cada β_i es igual a una constante dada β_i^* .

Para probar la hipótesis $H_0: \bar{\beta} = \bar{\beta}^*$ se debe encontrar una función de prueba, digamos, $u = g(y_1, y_2, \dots, y_n, x_{11}, x_{12}, x_{13}, \dots, x_{np})$, tal que la distribución de u sea conocida cuando $\bar{\beta} = \bar{\beta}^*$. La distribución debe también ser conocida cuando la hipótesis alternativa $H_1: \bar{\beta} \neq \bar{\beta}^*$ es verdadera, para poder evaluar la potencia de la prueba.

Para hacer lo anterior emplearemos el cociente de verosimilitud L como la función de prueba. Definiremos en primer lugar la ecuación de verosimilitud como:

$$f(\bar{e}; \bar{\beta}, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{(\bar{Y} - \bar{X}\bar{\beta})'(\bar{Y} - \bar{X}\bar{\beta})}{2\sigma^2}\right]$$

y sea $L = L(\hat{\omega}) / L(\hat{\Omega})$ la función de prueba^[3], siendo

$$L(\hat{\omega}) = \frac{n^{n/2} e^{-n/2}}{(2\pi)^{n/2} [(\bar{Y} - \bar{X}\bar{\beta}^*)'(\bar{Y} - \bar{X}\bar{\beta}^*)]^{n/2}}$$

y

$$L(\hat{\Omega}) = \frac{n^{n/2} e^{-n/2}}{(2\pi)^{n/2} [(\bar{Y} - \bar{X}\hat{\beta})'(\bar{Y} - \bar{X}\hat{\beta})]^{n/2}}$$

por lo tanto tenemos que L es:

$$L = \left[\frac{(\bar{Y} - \bar{X}\hat{\beta})'(\bar{Y} - \bar{X}\hat{\beta})}{(\bar{Y} - \bar{X}\bar{\beta}^*)'(\bar{Y} - \bar{X}\bar{\beta}^*)} \right]^{n/2}$$

Analizando la ecuación de verosimilitud es fácilmente demostrable que L puede ser obtenida si se calculan las siguientes expresiones:

$$Q = (\bar{Y} - \bar{X}\bar{\beta}^*)'(\bar{Y} - \bar{X}\bar{\beta}^*) = \min [(\bar{Y} - \bar{X}\bar{\beta})'(\bar{Y} - \bar{X}\bar{\beta}) \text{ cuando } H_0: \bar{\beta} = \bar{\beta}^* \text{ es verdadera}]$$

$$Q_1 = (\bar{Y} - \bar{X}\hat{\beta})'(\bar{Y} - \bar{X}\hat{\beta}) = \min [(\bar{Y} - \bar{X}\bar{\beta})'(\bar{Y} - \bar{X}\bar{\beta}) \text{ sin ninguna restricción sobre } \bar{\beta}].$$

por lo tanto $L = (Q_1/Q)^{n/2}$.

El cálculo de las expresiones anteriores generalmente se presenta en una tabla denominada de análisis de varianza, la cual contiene los "grados de libertad" (GL), "suma de cuadrados" (SC) y "cuadrados medios" (CM) correspondientes a las diversas "fuentes de variación" (FV) descritas (Tabla 1).

Tabla de Análisis de Varianza para probar $\bar{\beta} = \bar{\beta}^*$.			
FV	GL	SC	CM
Total	n	$Q = (\bar{Y} - \bar{X}\bar{\beta}^*)'(\bar{Y} - \bar{X}\bar{\beta}^*)$	
Debido a $\bar{\beta}$	q	$Q_2 = (\bar{Y} - \bar{X}\bar{\beta}^*)'\bar{X}\bar{S}^{-1}\bar{X}'(\bar{Y} - \bar{X}\bar{\beta}^*)$	Q_2/q
Error	$n-q$	$Q_1 = \bar{Y}'(I - \bar{X}\bar{S}^{-1}\bar{X}')\bar{Y}$	$Q_1/(n-q)$

Tabla 1

Veamos ahora el tipo de distribución que la función de prueba - L nos determina cuando $H_0: \bar{\beta} = \bar{\beta}^*$ es verdadero y cuando H_0 es falso, es decir, $H_1: \bar{\beta} \neq \bar{\beta}^*$ es verdadera.

Sea

$$\bar{Z} = \bar{Y} - \bar{X}\bar{\beta}^*, \quad \bar{A}_1 = \bar{I} - \bar{X}\bar{S}^{-1}\bar{X}', \quad \text{y} \quad \bar{A}_2 = \bar{X}\bar{S}^{-1}\bar{X}'$$

si $\bar{Y}$ se distribuye como $N(\bar{X}\bar{\beta}, \sigma^2\bar{I})$, entonces $\bar{Z}$ se distribuye como $N(\bar{X}\bar{\beta} - \bar{X}\bar{\beta}^*, \sigma^2\bar{I})$ y por lo tanto en base al teorema 4.16 de la referencia 1 tenemos que:

$$\bar{Z}'\bar{A}_1\bar{Z}/\sigma^2 \text{ se distribuye como } \chi^2_{(n-q)}$$

$$\bar{Z}'\bar{A}_2\bar{Z}/\sigma^2 \text{ se distribuye como } \chi^2(q, \gamma) \text{ donde}$$

$$\gamma = \frac{(\bar{\beta} - \bar{\beta}^*)'\bar{X}'\bar{A}_2\bar{X}(\bar{\beta} - \bar{\beta}^*)}{2\sigma^2} = \frac{(\bar{\beta} - \bar{\beta}^*)'\bar{S}(\bar{\beta} - \bar{\beta}^*)}{2\sigma^2}$$

además $\bar{Z}'\bar{A}_1\bar{Z}/\sigma^2$ y $\bar{Z}'\bar{A}_2\bar{Z}/\sigma^2$ son independientes.

De lo anterior se observa que:

$u = (\bar{z}' \bar{A}_2 \bar{z} / \bar{z}' \bar{A}_2 \bar{z}) / ((n-q)/q)$ se distribuye como $F'(q, n-q, \lambda)$ y se reduce a una F de Snedecor si y solo si H_0 es verdadera.

También se puede observar que $v = q u / [(n-q) + q u] = \bar{z}' \bar{A}_2 \bar{z} / \bar{z}' \bar{z}$ tiene una distribución $F^2(q, n-q, \lambda)$.

Reconsiderando la igualdad que el cociente de máxima verosimilitud L nos da, encontramos que:

$$L = \left[\frac{(\bar{y} - \bar{x} \hat{\beta})' (\bar{y} - \bar{x} \hat{\beta})}{(\bar{y} - \bar{x} \hat{\beta}^*)' (\bar{y} - \bar{x} \hat{\beta}^*)} \right]^{n/2} = (1-v)^{n/2} = \left\{ \frac{1}{1 + [q/(n-q) u]} \right\}^{n/2}$$

observándose que L es una función monótonica tanto de v como de u . Siendo posible por este hecho que tanto u como v puedan ser usadas como funciones de prueba para la hipótesis -- $H_0: \beta = \beta^*$.

Todo lo anterior da las bases para la demostración del teorema que a continuación se enuncia.

Teorema 3.-

En la hipótesis del modelo lineal general dada en un principio donde $\bar{e}$ se distribuye como $N(\bar{0}, \sigma^2 \bar{I})$ la cantidad

$$u = \frac{n-q}{q} \cdot \frac{Q_2}{Q_1}$$

se distribuye como una $F'(q, n-q, \lambda)$, donde $Q_1 + Q_2$ es el mínimo con respecto a β de $(\bar{y} - \bar{x} \beta)' (\bar{y} - \bar{x} \beta)$ cuando H_0 es verdadero y Q_1 es el mínimo con respecto a β cuando no existe restricción sobre β y

$$\lambda = \frac{(\bar{\beta} - \bar{\beta}^*)' \bar{S} (\bar{\beta} - \bar{\beta}^*)}{2 \sigma^2}$$

Resumiendo, para probar la hipótesis $H_0: \beta = \beta^*$ se requiere evaluar la tabla de análisis de varianza correspondiente a los datos que se estén manejando, una vez hecho esto calcular la $F(q, n-q)$ con los cuadrados medios pertenecientes al vector de parámetros y al vector de error y compararla con una $F_{\alpha}(q, n-q)$ observada en tablas para un nivel de significancia α .

IV.- La prueba de hipótesis de que una función lineal $\bar{r}'\beta$ es igual a τ_0 , donde $\bar{r}$ es un vector conocido y τ_0 un escalar conocido.

En el inciso II se dijo que $\hat{\beta}$ se distribuía como una $N(\beta, \sigma^2 \bar{r}'\bar{r})$ y se puede demostrar que dada la función lineal de $\hat{\beta}$, $\bar{r}'\hat{\beta}$ con $\bar{r}$ conocida ésta se distribuye como una $N(\bar{r}'\beta, \sigma^2 \bar{r}'\bar{r})$ por lo que la variable u se distribuye como una $t(n-q)$ grados de libertad, donde u es

$$u = \frac{\bar{r}'\hat{\beta} - \bar{r}'\beta}{\sqrt{\bar{r}'\bar{r}\sigma^2}}$$

Por lo tanto para probar la hipótesis $H_0: \bar{r}'\beta = \tau_0$ simplemente se reemplaza $\bar{r}'\hat{\beta}$ por el escalar conocido τ_0 en la ecuación anterior, comparando el valor de u con el valor de tablas de la distribución t de Student con $n-q$ grados de libertad.

En la obtención de la fórmula anterior se observa que:

$$v = \frac{\bar{r}'\hat{\beta} - \tau_0}{\sqrt{\sigma^2 \bar{r}'\bar{r}}}$$

se distribuye como una $N(\mu, 1)$ donde $\mu = (\bar{r}'\beta - \tau_0) \cdot (\sigma^2 \bar{r}'\bar{r})^{-1/2}$, con lo cual v^2 se distribuye como una $\chi^2(1, 1)$ donde

$$\lambda = \frac{(\bar{r}'\beta - \tau_0)^2}{2\sigma^2 \bar{r}'\bar{r}}$$

y por lo tanto

$$z = \frac{v^2}{\sigma^2 / \sigma^2} = \frac{(\bar{r}'\hat{\beta} - \tau_0)^2}{\sigma^2 \bar{r}'\bar{r}}$$

se distribuye como una $F(1, n-q, 1)$, reduciéndose a una F de Snedecor si y sólo si $H_0: \bar{r}'\beta = \tau_0$ es verdadera.

V.- La prueba de hipótesis de que un subconjunto dado de β_i 's es igual a un conjunto de constantes conocidas, dejando las restantes β_i sin especificar.

Para poder llevar a cabo la prueba anterior, el modelo $\bar{y} = \bar{x}\bar{\beta} + \bar{e}$ se representa dividiendo a la matriz $\bar{X}$ y al vector $\bar{\beta}$ de la siguiente manera:

$$\bar{X} = (\bar{X}_1, \bar{X}_2) \quad \bar{\beta} = \begin{pmatrix} \bar{\beta}_1 \\ \bar{\beta}_2 \end{pmatrix}$$

con $\bar{X}_1$ de $n \times r$ dimensiones y $\bar{\beta}_1$ de $r \times 1$, entonces el modelo puede ser representado también de la siguiente manera,

$$\bar{y} = \bar{x}\bar{\beta} + \bar{e} = (\bar{X}_1, \bar{X}_2) \begin{pmatrix} \bar{\beta}_1 \\ \bar{\beta}_2 \end{pmatrix} + \bar{e}$$

$$\circ \quad \bar{y} = \bar{X}_1 \bar{\beta}_1 + \bar{X}_2 \bar{\beta}_2 + \bar{e}$$

En este caso, se trata de probar la hipótesis $H_0: \bar{\beta}_1 = \bar{\beta}_1^*$ ($\bar{\beta}_1^*$ conocido); sin estipular nada sobre $\bar{\beta}_2$, como $\bar{\beta}_1^* = (\beta_1^*, \beta_2^*, \dots, \beta_r^*)$ tenemos entonces que la hipótesis en cuestión, es equivalente a probar la hipótesis $\beta_1 = \beta_1^*, \beta_2 = \beta_2^*, \dots, \beta_r = \beta_r^*$ (con $r < q$) en el modelo $y_j = \sum_{i=1}^q x_{ji} \beta_i + e_j$.

El siguiente teorema establece la prueba de la hipótesis $H_0: \bar{\beta}_1 = \bar{\beta}_1^*$, además de su construcción.

Teorema 4.-

En el modelo definido en el inciso I, cuya representación puede ser

$$\bar{y} = \bar{X}_1 \bar{\beta}_1 + \bar{X}_2 \bar{\beta}_2 + \bar{e}$$

donde $\bar{\beta}_1$ es de dimensión $r \times 1$ y el vector $\bar{e}$ se distribuye como $N(0, \sigma^2 I)$, el procedimiento que hay que seguir para probar la hipótesis $H_0: \bar{\beta}_1 = \bar{\beta}_1^*$ por medio del cociente de verosimilitud, es el siguiente:

a) Obtener el valor que minimice $\bar{e}'\bar{e}$ con respecto al vector de parámetros completo $\bar{\beta}$ en el modelo $\bar{Y} = \bar{X}\bar{\beta} + \bar{e}$, denotando este valor por Q_0 .

b) Obtener el valor que minimice $\bar{e}'\bar{e}$ con respecto al vector de parámetros $\bar{\beta}_2$ en el modelo $\bar{Y} = \bar{X}_1\bar{\beta}_1^* + \bar{X}_2\bar{\beta}_2 + \bar{e}$, también conocido como el modelo restringido por la hipótesis. Denotando el valor correspondiente por $Q_0 + Q_1$.

c) Sea $Q = Q_0 + Q_1 + Q_2$ donde $Q = (\bar{Y} - \bar{X}_1\bar{\beta}_1^*)(\bar{Y} - \bar{X}_1\bar{\beta}_1^*)'$.

d) Evaluar la cantidad u , la cual está definida como

$$u = [(\eta - q)/r] \cdot [Q_1/Q_0]$$

y por lo tanto se distribuye como una $F'(\tau, \eta - q, \lambda)$, donde

$$\lambda = \frac{(\bar{J}_1 - \bar{J}_1^*)' \bar{B} (\bar{J}_1 - \bar{J}_1^*)}{2\sigma^2}$$

y $\bar{B}$ es igual a $\bar{B} = \bar{X}_1' \bar{X}_1 - \bar{X}_1' \bar{X}_2 (\bar{X}_2' \bar{X}_2)^{-1} \bar{X}_2' \bar{X}_1$; como $\bar{B}$ es una matriz positiva definida, entonces u se distribuye como una $F(\tau, \eta - q)$ si y solo si H_0 es verdadera, esto es, si y solo si $\bar{J}_1 = \bar{J}_1^*$.

Demostración.- Ver referencia 1.

La tabla que sigue es la de análisis de varianza, que corresponde a la prueba de la hipótesis en cuestión. (Tabla 2)

Tabla de Análisis de Varianza para probar la subhipótesis $\bar{J}_2 = \bar{J}_1^*$.

F.V.	G.L.	S.C.	F
Total	η	$\bar{Y}'\bar{Y}$	
Debida a $\bar{\beta}$	q	$\hat{\beta}'\bar{X}'\bar{Y}$	
" " $\bar{J}_2$ (no ajustada)	$q - r$	$\bar{J}_2' (\bar{J}_1^*)' \bar{X}_2' (\bar{Y} - \bar{X}_1 \bar{J}_1^*)$	
" " $\bar{J}_1$ (ajustada)	r	$Q_1 = \bar{T}' \bar{X} \bar{S}^{-1} \bar{X}' \bar{T} - \hat{\bar{J}}_2' (\bar{J}_1^*) \bar{X}_2' \bar{T}$	$\frac{\eta - q}{r} \frac{Q_1}{Q_0}$
Error	$\eta - q$	$Q_0 = \bar{Y}'\bar{Y} - \hat{\beta}'\bar{X}'\bar{Y}$	

⁽¹⁾ $\bar{T}$ se define como $\bar{T} = \bar{Y} - \bar{X}_1 \bar{J}_1^*$.

Tabla 2

VI.- Prueba de las hipótesis $\bar{\lambda}_1\bar{\beta} = \bar{\lambda}_2\bar{\beta} = \dots = \bar{\lambda}_k\bar{\beta} = 0$, donde $\bar{\lambda}_1\bar{\beta}, \bar{\lambda}_2\bar{\beta}, \dots, \bar{\lambda}_k\bar{\beta}$ son funciones lineales de $\bar{\beta}$, y $\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_k$ son -- vectores linealmente independientes.

Para realizar esta prueba se "asociarán" los vectores $\bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3, \dots, \bar{\lambda}_k$ en una matriz de dimensiones $k \times q$, en este caso sea $\bar{H}' = (\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_k)$ la cual tiene rango k . Sea $\bar{H}$ una matriz de dimensiones $q \times q$ de rango q tal que $\bar{H}' = (\bar{H}'_1, \bar{H}'_2)$.

Y sea $\bar{H}^{-1}\bar{\Lambda} = (\bar{\Lambda}_1, \bar{\Lambda}_2)$, donde la matriz $\bar{\Lambda}_2$ tiene dimensiones $q \times k$, como $\bar{H}$ es una matriz conocida, se deduce que $\bar{H}^{-1} = \bar{\Lambda}$ también es conocida.

Definamos los siguientes productos:

$$\bar{H}\bar{H}^{-1}\bar{\Lambda} = (\bar{H}\bar{\Lambda}_1, \bar{H}\bar{\Lambda}_2) = (\bar{\bar{\Lambda}}_1, \bar{\bar{\Lambda}}_2) = \bar{\bar{\Lambda}} \quad \text{y sea}$$

$$\bar{H}\bar{\beta} = \begin{pmatrix} \bar{H}_1\bar{\beta} \\ \bar{H}_2\bar{\beta} \end{pmatrix} = \begin{pmatrix} \bar{\alpha}_1 \\ \bar{\alpha}_2 \end{pmatrix} = \bar{\alpha}$$

por lo tanto el modelo que se ha analizado hasta ahora $\bar{Y} = \bar{X}\bar{\beta} + \bar{e}$ puede ser expresado como, $\bar{Y} = \bar{X}\bar{H}^{-1}\bar{H}\bar{\beta} + \bar{e}$ o sea como $\bar{Y} = \bar{\bar{X}}\bar{\alpha} + \bar{e}$, o alternativamente como

$$\bar{Y} = \bar{\bar{X}}_1\bar{\alpha}_1 + \bar{\bar{X}}_2\bar{\alpha}_2 + \bar{e}.$$

Es evidente que probar la hipótesis $\bar{\alpha}_1 = 0$, es equivalente a probar $\bar{H}_1\bar{\beta} = 0$ y por lo tanto es de hacerse notar, que el inciso anterior (V) es el que hay que seguir para probar la hipótesis en cuestión, por lo que se presenta la tabla de análisis de varianza correspondiente, que no es más que una forma abreviada de representar dicho procedimiento. (Tabla 3).

Tabla de Análisis de Varianza para probar $\hat{\gamma}_1\beta = \hat{\gamma}_2\beta = \dots = \hat{\gamma}_K\beta = 0^*$			
F.V.	G.L.	S.C.	F
Total	n	$\bar{Y}'\bar{Y}$	
Debida a β	q	$\hat{\beta}'\bar{X}'\bar{Y}$	
" $\bar{\alpha}_2$ (no ajustada)	$q-h$	$\bar{\alpha}_2'\bar{Z}_2'\bar{Y}$	
" $\bar{\alpha}_1$ (ajustada)	h	$Q_1 = \hat{\beta}'\bar{X}'\bar{Y} - \bar{\alpha}_2'\bar{Z}_2'\bar{Y}$	$\frac{n-q}{h} \frac{Q_1}{Q_0}$
Error	$n-q$	$Q_0 = \bar{Y}'\bar{Y} - \hat{\beta}'\bar{X}'\bar{Y}$	

Tabla 3

* Nótese que en este caso, el estimador de verosimilitud de la segunda parte del vector de parámetros, difiere del de la tabla 2 porque en este caso se encontró bajo el modelo $\bar{Y} = \bar{Z}_2\bar{\alpha}_2 + \bar{e}$ y en la tabla 2 se hizo bajo el modelo $\bar{Y} = \bar{X}_1\bar{\alpha}_1 + \bar{X}_2\bar{\alpha}_2 + \bar{e}$.

Parte 2.- Modelos de Rango Incompleto.

1.- Introducción.- Hasta ahora se ha revizado escuetamente la teoría relacionada con los modelos lineales de rango completo y como se vió, se suponía que los datos captados en un momento dado podían ser representados por uno o más parámetros reales, sin embargo el mayor interés estribaba en determinar la dependencia funcional del significado de la información en base a los parámetros escogidos.

En esta parte analizaremos escuetamente los modelos de diseños de experimentos, también conocidos como modelos de clasificación y cuyo objetivo es determinar las diferencias que implica la consideración de distintos factores.^[47]

Debe hacerse notar que en este caso las variables son indicadoras de la existencia de un "nivel" del factor o factores en estudio.

El modelo se define como

$$y = \sum_{i=1}^p \beta_i x_i + e$$

donde cada elemento se identifica de la siguiente manera:

$y \equiv$ variable aleatoria observable

$\beta_i \equiv$ parámetro desconocido ($i=1, 2, \dots, p$)

$x_i \equiv$ cantidades conocidas que sólo toman el valor 0 o 1
($i=1, 2, \dots, p$)

$e \equiv$ variable aleatoria no observable.

Matricialmente el modelo se representa como:

$$\bar{y} = \bar{x}\bar{\beta} + \bar{e}$$

modelo que difiere del visto en la parte 1, porque la matriz $\bar{x}$ de $n \times p$ es de rango $k < p$, además de que cada $(x)_{ij} = 0$ o 1 . Este modelo frecuentemente se identifica también como la hipótesis del modelo lineal general de rango incompleto.

Planteamientos donde se observen más claramente las características o formas de estos modelos se pueden ver en:

a) Referencia 1, páginas 223 a 225,

b) Referencia 2, página 556,

c) Referencia 3, páginas 360 y 361.

En este caso el vector $\bar{e}$ también se sujetará a las suposiciones planteadas en la sección 1, parte 1.

II.- Estimación puntual de σ^2 y ciertas funciones lineales de las β_i .

Caso 2).-

Analizaremos inicialmente la estimabilidad de ciertas funciones lineales de las β_i , suponiendo que $E(\bar{e}) = 0$ y $E(\bar{e}\bar{e}') = \sigma^2 I$ o sea considerando el caso 2.

Si empleamos los mínimos cuadrados, ya hemos visto que la derivada parcial respecto a β del producto $\tilde{e}'\tilde{e}$ da $2\tilde{x}'\tilde{x}\tilde{\beta} - 2\tilde{x}'\tilde{y}$, e igualando a cero la anterior cantidad, se obtienen las ecuaciones normales, que son $\tilde{x}'\tilde{x}\tilde{\beta} = \tilde{x}'\tilde{y}$.

Si la matriz $\tilde{x}$ es de rango completo, $\tilde{x}'\tilde{x}$ tiene una inversa y la solución para $\tilde{\beta}$ es única, siendo $\tilde{\beta}$ un estimador insesgado de β .

En el caso de que $\tilde{x}$ no sea de rango completo, $\tilde{x}'\tilde{x}$ no tiene inversa y el sistema de ecuaciones normales debe ser analizado, para ver si existe una solución,

Procediendo de acuerdo con lo anterior, notamos lo siguiente:

1) Las matrices asociadas con el sistema de ecuaciones normales son $A = \tilde{x}'\tilde{x}$, y la matriz aumentada $B = (\tilde{x}'\tilde{x} | \tilde{x}'\tilde{y})$, la cual ----- también se puede expresar como $\tilde{x}'(\tilde{x} | \tilde{y})$.

2) Analizando los rangos de las matrices asociadas observamos -- que:

- a) Dado que el rango de $\tilde{x}$ es k , el rango de la matriz A es k también.
- b) Como el rango de $\tilde{x}'$ es k y el rango de la matriz $(\tilde{x} | \tilde{y})$ no puede ser más que mayor o igual que k , y como B es el producto de las dos matrices anteriores se concluye, que el rango de B es k también.

Por lo tanto de a) y b) se concluye que el sistema de ecuaciones es consistente.

3) Por último, como la dimensión de $\tilde{x}'\tilde{x}$ es $q \times q$ y el rango es $k < q$, la teoría de algebra de matrices indica que en este caso existe un número infinito de soluciones.

Todo lo anterior nos proyecta a una situación no muy deseable, - puesto que dos investigadores con un mismo modelo e iguales observaciones podrían obtener diferentes estimaciones de β .

Como cualquier solución a las ecuaciones normales para $\hat{\beta}$ estará en función lineal del vector de observaciones $\tilde{y}$, es importante hacer notar que no existe combinación lineal alguna de $\tilde{y}$ que dé un estimador insesgado de β o β_1 .

Demostremos lo enunciado.

Supongamos que existe al menos una combinación lineal de $\bar{Y}$, y sea $\bar{A}$ la matriz de $q \times n$, constante e independiente de $\bar{\beta}$, que da la combinación lineal que proporciona el estimador insesgado de $\bar{\beta}$, tomando la esperanza se observa que:

$$E(\bar{A}\bar{Y}) = E(\bar{A}(\bar{X}\bar{\beta} + \bar{e})) = \bar{A}\bar{X}\bar{\beta}$$

por lo tanto si $\bar{A}\bar{Y}$ es estimador insesgado de $\bar{\beta}$, tenemos que -- $\bar{A}\bar{X}$ debe ser igual a $\bar{I}$, pero tal cosa es imposible ya que $\bar{A}\bar{X}$ es de rango k a lo más, con $k < q$, lo cual implica que no existe por lo tanto estimadores insesgados de $\bar{\beta}$.

Por lo tanto existen estimadores insesgados únicamente de combinaciones lineales de $\bar{\beta}$ ($\bar{\lambda}\bar{\beta}$), donde $\bar{\lambda}$ es un vector de constantes conocidas de $q \times 1$, se puede observar que lo anterior no es cierto para cualquier vector $\bar{\lambda}$; como sería el caso para cualquier vector canónico de R^q , ya que ello implicaría que $\bar{\beta}$ tuviera un estimador insesgado y se ha demostrado que eso no es posible.

Una vez que se ha visto ya, que cierto tipo de combinaciones lineales son estimables, veremos un teorema que ayudará a determinar si ésto es cierto para la combinación lineal que se intente estimar.

Teorema 1.-

En el modelo en cuestión, caso 2, la combinación lineal $\bar{\lambda}\bar{\beta}$ ($\bar{\lambda}$ vector de constantes conocidas) es estimable si y sólo si existe una solución para $\bar{\tau}$ en las ecuaciones $\bar{X}'\bar{X}\bar{\tau} = \bar{\lambda}$.

Demostración.- Ver referencia 1.

Es importante hacer notar que para el modelo que se estudia en esta parte, existe un teorema análogo al de Gauss-Markoff (Teorema 1, parte 1), el cual se ve a continuación.

Teorema 2.-

En el modelo que se está tratando, caso 2, el mejor estimador lineal insesgado para cualquier función estimable $\bar{\lambda}\bar{\beta}$ es $\bar{\tau}'\bar{X}'\bar{Y}$ donde $\bar{\tau}$ satisface la ecuación $\bar{X}'\bar{X}\bar{\tau} = \bar{\lambda}$.

Demostración.- Ver referencia 1.

Porque como se vió, $\bar{S}$ no es de rango completo, observamos que esto implicaba la existencia de una cantidad infinita de soluciones al sistema $\bar{S}\bar{\tau}=\bar{\lambda}$ si $\bar{\lambda}'\bar{\beta}$ es una combinación lineal estimable, sin embargo el siguiente teorema estipula que éste no es el caso.

Teorema 3.-

Si $\bar{\lambda}'\bar{\beta}$ es una función estimable, entonces cualquier vector $\bar{\tau}$ -- que satisface el sistema de ecuaciones $\bar{S}\bar{\tau}=\bar{\lambda}$ proporciona el mismo estimador de $\bar{\lambda}'\bar{\beta}$.

Demostración.- Ver referencia 1.

Hasta este punto podríamos suponer que existe un conjunto infinito de $\bar{\lambda}$'s, tal que para cada $\bar{\lambda}$ contenida en este conjunto tenemos que $\bar{\lambda}'\bar{\beta}$ es una combinación lineal estimable. A continuación vemos un teorema que demuestra que ése no es el caso.

Teorema 4.-

Existen exactamente k funciones estimables linealmente independientes, donde k es el rango de $\bar{X}$.

Demostración.- Ver referencia 1.

Del teorema anterior se puede observar que las funciones matriciales $\bar{X}\bar{\beta}$ y $\bar{X}'\bar{X}\bar{\beta}$ son estimables, lo cual se demostrará a continuación.

$\bar{X}\bar{\beta}$ es estimable.

Demostración.-

Partamos la matriz $\bar{X}$ en la forma

$$\bar{X} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_n \end{pmatrix}$$

en donde cada $\bar{x}_i$ es un vector de $1 \times q$ y que forma la i -ésima fila de $\bar{X}$. Se intentará demostrar que $\bar{x}_i\bar{\beta}$ es estimable para toda $i=1, 2, \dots, n$, si esto es cierto, se habrá demostrado que $\bar{X}\bar{\beta}$ es estimable.

Ahora bien, sabemos que $\bar{x}_i\beta$ es estimable si existe una $\bar{r}_i$ tal - que $\bar{X}'\bar{X}\bar{r}_i = \bar{x}_i$ y por el teorema 4 parte 2 se ve que cualquier $\bar{x}_i\beta$ es estimable, por lo tanto $\bar{x}\beta$ es estimable.

De lo anterior se concluye que $E(\hat{y}) = \bar{x}\beta$ es estimable, en consecuencia, aplicando un razonamiento semejante, se puede demostrar que $\bar{X}'\bar{X}\beta$ es estimable también. Emplee en este caso $\bar{x}_i^*\bar{x}\beta$ donde $\bar{x}_i^*$ es un renglón de la matriz $\bar{X}'$.

Teorema 5.-

Si $\bar{x}_1\beta, \bar{x}_2\beta, \bar{x}_3\beta, \dots, \bar{x}_s\beta$ son funciones estimables, cualquier combinación lineal de estas funciones estimables, es estimable.

Demostración.- Ver referencia 1.

Teorema 6.-

El mejor estimador lineal insesgado de una combinación lineal de funciones estimables, está dado por la misma combinación lineal de los mejores estimadores lineales insesgados de las funciones estimables.

Demostración.-

Sea $\bar{x}_1\beta, \bar{x}_2\beta, \bar{x}_3\beta, \dots, \bar{x}_s\beta$ un conjunto de funciones estimables, por - el teorema 2 parte 2, se sabe que existen los siguientes productos $\bar{r}_i'\bar{X}'\bar{Y}$ desde $i=1, 2, 3, \dots, s$, que proporcionan correspondientemente los mejores estimadores lineales insesgados de $\bar{x}_i\beta$; ahora sean $a_1, a_2, a_3, \dots, a_s$ constantes escalares, si determinamos la combinación lineal $\bar{r}_0' = \sum a_i \bar{r}_i'$, se demuestra por el teorema 5 -- parte 2, que $\bar{r}_0'\beta$ es una función estimable; y en este caso la mejor estimación lineal insesgada de la combinación lineal está dada por:

$$\sum a_i \bar{r}_i' \bar{X}' \bar{Y} = \bar{r}_0' \bar{X}' \bar{Y}$$

donde

$$\bar{r}_0' = \sum_i a_i \bar{r}_i'$$

Caso 1).-

Consideración del caso normal para la estimación de una cierta - combinación lineal de β_i 's.

Si resulta que el vector $\bar{e}$ se distribuye como $N(\bar{0}, \sigma^2 \bar{I})$, la maximización de la ecuación de verosimilitud lleva a los mismos resultados que en el caso anterior.

Sin embargo, es posible reparametrizar, el modelo de rango incompleto a uno de rango completo, y en este caso poder aplicar todo lo observado en dicho modelo.

Por lo que se aconseja revisar todo lo relacionado con reparametrización, en la referencia 1, a partir de la página 235.

Estimación de la varianza (σ^2).

Por lo visto hasta este momento se sabe que si $\bar{X}\beta$ es una función estimable, la mejor estimación lineal insesgada está dada por $\bar{X}\hat{\beta}$, donde $\hat{\beta}$ es cualquier solución a las ecuaciones normales; vamos a demostrar tal hecho formalmente en el siguiente teorema para posteriormente demostrar en otro teorema que esto mismo se observa en la estimación de σ^2 .

Teorema 7.-

Sea $\bar{X}\beta$ una función estimable en el modelo en cuestión, y sean $\hat{\beta}$ y $\hat{\alpha}$ dos vectores $q \times 1$ que satisfacen las ecuaciones normales; es decir $\bar{X}'\bar{X}\hat{\beta} = \bar{X}'Y$ y $\bar{X}'\bar{X}\hat{\alpha} = \bar{X}'Y$, entonces $\bar{X}\hat{\beta} = \bar{X}\hat{\alpha}$ y $E(\bar{X}\hat{\beta}) = E(\bar{X}\hat{\alpha}) = \bar{X}\beta$.

Demostración.-

Por hipótesis, se tiene que $\hat{\beta}$ y $\hat{\alpha}$ son dos vectores que satisfacen las ecuaciones normales o sea:

$$(a) \quad \bar{X}'\bar{X}\hat{\beta} = \bar{X}'Y \quad y \quad \bar{X}'\bar{X}\hat{\alpha} = \bar{X}'Y$$

(b) asimismo hemos establecido que $\bar{X}\beta$ es estimable y por lo tanto considerando el teorema 2 parte 2, se observa que existe un vector $\bar{\tau}$ tal que $\bar{\tau}'\bar{X}\bar{X} = \bar{X}'$.

(c) Premultiplicando por $\bar{\tau}'$ ambas ecuaciones de (a), se obtiene:

$$\bar{\tau}'\bar{X}'\bar{X}\hat{\beta} = \bar{\tau}'\bar{X}'Y \quad ; \quad \bar{\tau}'\bar{X}'\bar{X}\hat{\alpha} = \bar{\tau}'\bar{X}'Y$$

$$\text{o sea} \quad \bar{\tau}'\hat{\beta} = \bar{\tau}'\bar{X}'Y \quad y \quad \bar{\tau}'\hat{\alpha} = \bar{\tau}'\bar{X}'Y$$

$$\text{lo cual implica que} \quad \bar{\tau}'\hat{\beta} = \bar{\tau}'\hat{\alpha}$$

y por lo tanto se observa que:

$$E(\bar{\tau}'\hat{\beta}) = E(\bar{\tau}'\hat{\alpha}) = E(\bar{\tau}'\bar{X}'Y) = \bar{\tau}'\bar{X}'E(Y) = \bar{\tau}'\bar{X}'\bar{X}\beta = \bar{\tau}'\bar{X}\beta.$$

Teorema 8.-

En el modelo que se está analizando, la cantidad

$$\hat{\sigma}^2 = \frac{1}{n-k} (\bar{y} - \bar{x}\hat{\beta})'(\bar{y} - \bar{x}\hat{\beta})$$

es invariante para cualquier $\hat{\beta}$, que sea solución para las ecuaciones normales $\bar{x}'\bar{x}\hat{\beta} = \bar{x}'\bar{y}$.

Entendiéndose por invariante, el hecho de que $\hat{\sigma}^2$ sea la misma, - para cualquier solución que se emplee de las ecuaciones normales.

Demostración.-

Sean $\hat{\alpha}$ y $\hat{\beta}$ dos soluciones cualesquiera de las ecuaciones normales, anteriormente se demostró que $\bar{x}\hat{\beta}$ es estimable, y por el teorema anterior se tiene que $\bar{x}\hat{\alpha}$ y $\bar{x}\hat{\beta}$ son iguales, es decir, $\bar{x}\hat{\beta}$ es la misma para cualquier vector $\hat{\beta}$ que satisfaga a las ecuaciones normales, por lo tanto $\bar{y} - \bar{x}\hat{\beta}$ resulta ser la misma, independientemente de la solución de las ecuaciones normales que se utilice y de ahí que $\hat{\sigma}^2$ sea también invariante.

III.- Prueba de la hipótesis de que una cierta función estimable de β es igual a una constante conocida.

Para llevar a cabo una prueba de hipótesis ($H_0: \bar{\lambda}'\beta = \alpha_0^*$, α_0^* como α_0) de una función estimable $(\bar{\lambda}'\beta)$, debemos en primer lugar suponer que $\bar{e}$ se distribuye como $N(0, \sigma^2 I)$, en segundo lugar reparametrizar el modelo $\bar{y} = \bar{x}\beta + \bar{e}$ con $\bar{x}$ de rango k y dimensiones $n \times q$ (recuérdese $k < q \leq n$), al modelo $\bar{y} = \bar{z}\bar{\alpha} + \bar{e}$ con $\bar{z}$ de $n \times k$ y rango k , donde un elemento cualquiera de $\bar{\alpha}$, es decir $\alpha_i = \bar{\lambda}_i'\beta$.

En tercer lugar y considerando que una vez reparametrizado un modelo de rango incompleto, éste tiene todas las características - vistas en la parte 1 de este capítulo, lo que queda por hacer en realidad es la prueba de la hipótesis siguiente: $H_0: \alpha_i = \alpha_0^*$ donde el modelo considerado sería $(\bar{y} = \bar{z}\bar{\alpha} + \bar{e}$ con $\bar{e}$ distribuyéndose como $N(0, \sigma^2 I)$ y $\bar{z}$ de $n \times k$ y rango k), el que se hubiera obtenido de la reparametrización realizada anteriormente.

Por todo lo anterior y resumiendo se puede comprobar que:

- a) $(\bar{x}'\bar{x})\hat{\alpha} = \bar{x}'\bar{y}$ son las ecuaciones normales del modelo reparametrizado,
- b) $\hat{\alpha}_i$ obviamente puede ser obtenida,
- c) $\hat{\alpha}_i$ se distribuye como $N(\alpha_i, c_{ii}\sigma^2)$ donde c_{ii} es el i -ésimo elemento de la diagonal de $(\bar{x}'\bar{x})^{-1}$
- d) $(n-k)(\hat{\sigma}^2/\sigma^2)$ se distribuye como $\chi^2(n-k)$
- e) Que $\hat{\alpha}_i$ y $\hat{\sigma}^2$ son independientes y por lo tanto $v = (\hat{\alpha}_i - \alpha_i)/(c_{ii}\hat{\sigma}^2)^{1/2}$ se distribuye como una t de Student con $n-k$ grados de libertad.

Por lo tanto si α_i se substituye por α_0 y si la hipótesis ($H_0: \alpha_i = \alpha_0^*$) es cierta, tenemos que la cantidad $v^2 = (\hat{\alpha}_i - \alpha_0^*)^2 / (c_{ii}\hat{\sigma}^2)$ se distribuye como una $F(1, n-k, \lambda)$ donde λ es igual a

$$\lambda = \frac{1}{2c_{ii}\sigma^2} (\bar{x}'\bar{\beta} - \alpha_0^*)^2$$

y naturalmente λ es igual a cero si y sólo si $\bar{x}'\bar{\beta}$ es igual a α_0^* .

IV.- Prueba de la hipótesis $\beta_1 = \beta_2 = \dots = \beta_l$ ($l \leq q$).

Naturalmente, sólo se discutirán las pruebas de hipótesis estimables, esto es, se considerará la prueba de la hipótesis $\beta_1 = \beta_2 = \dots = \beta_l$ si y sólo si existe un conjunto de funciones estimables, linealmente independientes $\bar{x}_1\bar{\beta}, \bar{x}_2\bar{\beta}, \dots, \bar{x}_m\bar{\beta}$ tal que H_0 sea verdadera, si y sólo si $\bar{x}_1\bar{\beta} = \bar{x}_2\bar{\beta} = \dots = \bar{x}_m\bar{\beta} = 0$.

Definición.- Una hipótesis H_0 será llamada estimable si existe un conjunto de funciones estimables linealmente independiente $\bar{x}_1\bar{\beta}, \bar{x}_2\bar{\beta}, \bar{x}_3\bar{\beta}, \dots, \bar{x}_m\bar{\beta}$ tal que H_0 es verdadera si y sólo si $\bar{x}_1\bar{\beta} = \bar{x}_2\bar{\beta} = \dots = \bar{x}_m\bar{\beta} = 0$.

Teorema 9.-

En el modelo de rango incompleto, para probar la hipótesis de que $\beta_1 = \beta_2 = \beta_3 = \dots = \beta_l$ ($l \leq k$), la cual suponemos es equivalente a probar que las funciones estimables linealmente independientes $\bar{x}_1\bar{\beta} = \bar{x}_2\bar{\beta} = \bar{x}_3\bar{\beta} = \dots = \bar{x}_m\bar{\beta} = 0$, lo que tenemos que hacer, es lo siguiente:

- 1) Obtener cualquier solución a las ecuaciones normales $\bar{x}'\bar{x}\bar{\beta} = \bar{x}'\bar{y}$ y formar $Q_0 = (\bar{y} - \bar{x}\hat{\beta})'(\bar{y} - \bar{x}\hat{\beta})$

2) Obtener cualquier solución a las ecuaciones normales $\bar{z}_2' \bar{z}_2 \hat{\beta} = \bar{z}_2' \bar{y}$, las cuales se obtienen del modelo $\bar{y} = \bar{x}\beta + \bar{e}$, bajo las condiciones $\beta_1 = \beta_2 = \dots = \beta_q$ (denotando al modelo reducido por $\bar{y} = \bar{z}_2 \beta + \bar{e}$) y formar $Q_0 + Q_1 = (\bar{y} - \bar{z}_2 \hat{\beta})'(\bar{y} - \bar{z}_2 \hat{\beta})$ y

3) Calcular

$$v = \frac{n-q}{m} \cdot \frac{Q_1}{Q_0} = \frac{\hat{\beta}' \bar{x}' \bar{y} - \hat{\beta}' \bar{z}_2' \bar{y}}{\bar{y}' \bar{y} - \hat{\beta}' \bar{x}' \bar{y}} \cdot \frac{n-q}{m}$$

donde v se distribuye como $F'(m, n-q, \lambda^{\text{II}})$.

Demostración.-

Suponga que $\bar{\alpha}$ es un vector constituido de la siguiente manera $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_2, \dots, \bar{\alpha}_m)'$, además considere, las siguientes funciones también estimables $\bar{x}_{m+1}\beta, \dots, \bar{x}_k\beta$, $k-m$ funciones, tales que con las del vector $\bar{\alpha}$ forman un conjunto de k funciones estimables linealmente independiente, además, sea $\hat{\beta}' = (\bar{x}_{m+1}\beta, \bar{x}_{m+2}\beta, \dots, \bar{x}_k\beta)$.

Si hacemos $\bar{\delta} = \begin{pmatrix} \bar{\alpha} \\ \hat{\beta} \end{pmatrix}$, entonces considerando que podemos reparametrizar el modelo en cuestión, a uno de rango completo, observamos que este último lo podemos representar como

$$\bar{y} = \bar{z}_1 \bar{\alpha} + \bar{z}_2 \hat{\beta} + \bar{e}$$

y por lo tanto H_0 es verdadera si y sólo si $\bar{\alpha} = \bar{0}$, y para probar la hipótesis podemos emplear el inciso V de la parte 1.

Las cantidades importantes según la notación empleada hasta el momento, son definidas como sigue:

a) $(\bar{y} - \bar{z}_1 \hat{\delta})'(\bar{y} - \bar{z}_1 \hat{\delta})$ o sea la suma de cuadrados debida al error donde $\hat{\delta}$ es la solución a las ecuaciones normales del modelo reparametrizado.

b) $\hat{\delta}' \bar{x}' \bar{y} - \hat{\beta}' \bar{z}_2' \bar{y}$, también conocida como la suma de cuadrados debida a $\bar{\alpha}$ ajustada para $\hat{\beta}$, donde $\hat{\beta}$ es la solución a las ecuaciones normales $\bar{z}_2' \bar{z}_2 \hat{\beta} = \bar{z}_2' \bar{y}$. Estas ecuaciones normales son derivadas del modelo $\bar{y} = \bar{z}_2 \beta + \bar{e}$.

De acuerdo con a) y b), observamos que siempre que queramos probar la hipótesis en cuestión, tendremos necesidad de calcular la matriz $\bar{Z}$, lo cual sería completamente tedioso, sin embargo lo anterior afortunadamente no es necesario, ya que podemos emplear el teorema 8 parte 2 para demostrar que $(Y - \bar{Z}\hat{\beta})'(Y - \bar{Z}\hat{\beta}) = (Y - X\hat{\beta})'(Y - X\hat{\beta})$ lo cual evita tal molestia. Evidentemente por la misma razón tenemos que el término $(Y - \bar{Z}\hat{\beta})'(Y - \bar{Z}\hat{\beta})$ puede ser obtenido de las ecuaciones normales que son derivadas del modelo $\bar{Y} = X\beta + \varepsilon$ con la condición de que $\beta_1 = \beta_2 = \beta_3 = \dots = \beta_k$.

Por lo tanto con lo anterior se justifica el uso de la notación empleada en el teorema 9 parte 2.

Referencias.

- 1.- Graybill Franklin A., An Introduction to Linear Statistical Models, Volume I, New York: McGraw-Hill Book Co. Inc., 1961 (First Edition).
- 2.- Lindgren Bernard W., Statistical Theory, New York. Macmillan Publishing, 1976 (Third Edition).
- 3.- Searle S. R., Linear Models, New York: John Wiley & Sons, Inc. 1971 (First Edition).

CAPITULO 11

Revisión de modelos para datos con dos criterios de clasificación. (Modelos de rango incompleto).

I.- Introducción.- En el capítulo anterior se vió de una forma general, la teoría relacionada con los modelos lineales de rango -- completo e incompleto respectivamente, analizando o revisando --- siempre resultados para el caso más general, o sea suponiendo que cualquiera que fuese el modelo, este implicaba el mayor número de factores en la composición "analítica" del mismo.

Sin embargo, cualquiera que fuera el modelo o cualesquiera que -- fueran las situaciones, nunca se propuso considerar el nada remoto fenómeno de la interacción entre los factores¹⁾ considerados en el estudio en cuestión.

Este hecho que es de suma importancia y que forma la parte medular del presente trabajo se planteará en este capítulo, haciendo notar y como el título del mismo lo indica, que se considerarán - los modelos de diseños de experimentos con dos factores de clasificación o sea, se analizará el modelo de diseños de experimentos más sencillo que permite la consideración de la interacción.

Esta decisión no se debe de ninguna manera a la intención de minimizar trabajo, sino que responde a una tendencia general de la -- gente que desarrolló esta área de investigación y que consiste en representar a los fenómenos en estudio con la consideración del - menor número de factores posible, la cual dió como resultado que hoy en día existan un mayor número de métodos para modelos con me - nos factores, que para modelos con muchos factores.

Por último es importante mencionar, que muy frecuentemente y por diversas razones la determinación de un modelo no permite la realización de varias observaciones²⁾ para un conjunto de valores determinados para las variables X_i , esto implica la necesidad de - considerar modelos que serán determinados con más de una observación por celda³⁾ y modelos que se determinarán con una observación por celda.

Para hacer más notorio el fenómeno de la interacción, se analizarán los modelos con más de una y una observación por celda, para los casos de no-interacción e interacción respectivamente.

II.- Modelos sin interacción.

a) Con una observación por celda. .

Antes de seguir adelante con el caso en cuestión, definiremos ciertos elementos que serán comunes a cualquiera de los modelos que se analizarán:

μ será un parámetro desconocido, denominado frecuentemente como la media general:

$\{\tau_i\}$ será un conjunto de parámetros asociados con uno de los criterios o factores en consideración también desconocidos;

$\{\delta_j\}$ será otro conjunto de parámetros asociados con el criterio o factor restante, obviamente desconocidos.

Los datos que se analizan de acuerdo a dos criterios de clasificación y una observación por celda tienen el siguiente modelo:

$$y_{ij} = \mu + \tau_i + \delta_j + e_{ij}$$

$$i = 1, 2, \dots, I \quad j = 1, 2, \dots, J$$

donde y_{ij} es la observación de la i j -ésima celda, e_{ij} es la variable aleatoria asociada a dicha observación, la cual no tiene correlación alguna con otra variable aleatoria y además tiene media cero y varianza σ^2 ; los demás elementos son como se definieron previamente.

Es claro que para realizar pruebas de hipótesis y/o calcular intervalos de confianza se supondrá que los errores se distribuyen normalmente.

Generalmente la finalidad de este modelo será estimar los contrastes $\sum e_i \tau_i$ y $\sum \delta_j \delta_j$, además de probar las hipótesis $\tau_1 = \tau_2 = \dots = \tau_I$ y $\delta_1 = \delta_2 = \dots = \delta_J$.

Estimación Puntual.- Como se recordará, en el capítulo anterior -- parte dos, se planteó el hecho de que para saber si una combinación lineal, digamos $\bar{x}'\bar{\beta}$, era estimable, era necesario encontrar un vector $\bar{\tau}$ que cumpliera la condición $\bar{x}'\bar{\tau} = \bar{\beta}$; todo lo cual provoca en sí un procedimiento laborioso sobretodo si las dimensiones de la matriz $\bar{x}'\bar{x}$ son de un tamaño respetable.

Sin embargo esta situación poco agradable se puede evitar, si en lugar de considerar lo anterior empleamos el hecho (también planteado en el capítulo anterior) de que $\bar{x}\bar{\beta}$ y $\bar{x}'\bar{x}\bar{\beta}$ son estimables, en cuyo caso algunas veces se podrá decir casi inmediatamente si -

una cierta función de los parámetros es estimable, esto es, si es función lineal de los renglones de $\bar{x}\beta$ o de $\bar{x}'\beta$.

Una vez que se ha escogido la última alternativa, el problema que queda por resolver es el de la obtención del mejor estimador de la combinación lineal correspondiente (Nótese que según lo expuesto en el capítulo anterior, el mejor estimador estaría dado por $\bar{r}'\bar{x}'\bar{y}$; sin embargo nosotros en este momento hemos decidido ya no encontrar $\bar{r}$ por lo laborioso que representa ser.) lo cual se llevará a cabo empleando las ecuaciones normales correspondientes y el siguiente corolario:

Corolario.- Si $\bar{x}\beta$ es estimable, existe un vector δ tal que $\bar{x}' = \delta'\bar{x}'\bar{x}$, y el mejor estimador lineal insesgado de $\bar{x}\beta$ es $\delta'\bar{x}'\bar{y}$.

Obsérvese por lo tanto que de las ecuaciones normales ($\bar{x}'\bar{x}\beta = \bar{x}'\bar{y}$) y lo dicho anteriormente, cualquier combinación lineal de los renglones del lado izquierdo de las ecuaciones normales es estimable (eliminando los sombreros) y por el corolario se tiene que el mejor estimador lineal insesgado está dado por la misma combinación lineal de renglones del lado derecho de las ecuaciones normales.

Por lo tanto, procederemos a obtener las ecuaciones normales del caso en cuestión;

$$\begin{aligned} \mu: & I\hat{\mu} + \hat{\tau}_1 + \dots + \hat{\tau}_I + I\hat{\eta}_1 + \dots + I\hat{\eta}_J = Y.. \\ \tau_p: & \hat{\mu} + \hat{\tau}_p + \hat{\eta}_1 + \dots + \hat{\eta}_J = Y_{p.} \\ \eta_q: & I\hat{\mu} + \hat{\tau}_1 + \dots + \hat{\tau}_I + \hat{\eta}_q = Y_{.q} \end{aligned}$$

$J, p=1, 2, \dots, J$
 $q=1, 2, \dots, J$

De las ecuaciones anteriores se puede observar que efectivamente las únicas combinaciones de las τ_i o de las η_j que son estimables son los contrastes.

Prueba de Hipótesis.- Para llevar a cabo la prueba de hipótesis - $\tau_1 = \tau_2 = \dots = \tau_I$ o la de $\eta_1 = \eta_2 = \dots = \eta_J$ se debe encontrar en primer lugar:

1.- La reducción debida al vector de parámetros β , la cual se suele denotar como $R(\beta) = R(\mu, \tau, \mu) = \beta' \bar{X} \bar{Y}$.

2.- La reducción debida al conjunto de parámetros no incluidos bajo la hipótesis, considerando a los incluidos como nulos. Cantidad que se denota como $R(\alpha_2) = R(\mu, \mu) (\sigma = R(\mu, \tau)) = \bar{\alpha}_2' \bar{X}_2' \bar{Y}$.

3.- La reducción debida al conjunto de parámetros incluidos bajo la hipótesis ajustada para el conjunto de parámetros no incluidos. Esta cantidad se denota como

$$R(\beta) - R(\alpha_2) = R(\mu, \tau, \mu) - R(\mu, \mu) (\sigma = R(\mu, \tau, \mu) - R(\mu, \tau)) \quad \text{y}$$

4.- La suma de errores al cuadrado, la cual se denota por

$$\sum_{i,j} y_{ij}^2 - R(\mu, \tau, \mu)$$

Para calcular uno, observamos que necesitamos encontrar $\hat{\beta}$ ya que $\bar{X}'\bar{Y}$ lo tenemos del lado derecho de las ecuaciones normales, o sea tenemos que

$$R(\mu, \tau, \mu) = \hat{\mu} Y_{..} + \sum \hat{\tau}_j Y_{.j} + \sum \hat{\mu}_p Y_{.p}$$

ahora una forma sencilla de encontrar un valor de $\hat{\beta}$ es considerar las combinaciones lineales no estimables $\sum \hat{\tau}_j = 0$ y $\sum \hat{\mu}_p = 0$ que junto con las ecuaciones normales nos da los siguientes resultados:

$$\hat{\mu} = \frac{Y_{..}}{IJ} \quad , \quad \hat{\tau}_j = \frac{Y_{.j}}{J} - \frac{Y_{..}}{IJ} \quad \text{y} \quad \hat{\mu}_p = \frac{Y_{.p}}{I} - \frac{Y_{..}}{IJ}$$

De lo anterior se obtiene que:

$$R(\mu, \tau, \mu) = \frac{Y_{..}^2}{IJ} + \left(\frac{\sum Y_{.j}^2}{J} - \frac{Y_{..}^2}{IJ} \right) + \left(\frac{\sum Y_{.p}^2}{I} - \frac{Y_{..}^2}{IJ} \right)$$

La cantidad dos se calcula considerando al modelo en este caso, - como

$$y_{ij} = (\mu + \tau) + \mu_j + e_{ij}$$

o $y_{ij} = \mu^* + \mu_j + e_{ij} \quad \text{con} \quad \mu^* = \mu + \tau$

el cual tiene como ecuaciones normales a

$$\mu^*: IJ \hat{\mu}^* + I \hat{y}_1 + \dots + I \hat{y}_J = Y_{..}$$

$$y_p: I \hat{\mu}^* + I \hat{y}_p = Y_{.p}$$

empleando la combinación lineal no estimable $\sum \hat{y}_p = 0$ obtenemos los siguientes resultados para el vector de parámetros $(\mu^*, y_1, y_2, y_3, \dots, y_J)$

$$\mu^* = \frac{Y_{..}}{IJ} \quad y \quad \hat{y}_p = \frac{Y_{.p}}{I} - \frac{Y_{..}}{IJ}$$

y como $\bar{Y}'\bar{X} = (\bar{X}'\bar{Y})' = (Y_{..}, Y_{.1}, Y_{.2}, Y_{.3}, \dots, Y_{.J})$ se tiene

$$que \quad R(\mu, y) = \frac{Y_{..}^2}{IJ} + \left(\frac{\sum Y_{.p}^2}{I} - \frac{Y_{..}^2}{IJ} \right)$$

$$\left[\sigma R(\mu, \tau) = \frac{Y_{..}^2}{IJ} + \left(\frac{\sum Y_{.p}^2}{I} - \frac{Y_{..}^2}{IJ} \right) \right] \quad \text{[5]}$$

La cantidad del punto tres es igual a la diferencia (1 - 2) de los dos puntos anteriores, lo que da

$$\begin{aligned} & R(\mu, \tau, y) - R(\mu, y) \quad [\sigma R(\mu, \tau, y) - R(\mu, \tau)] \\ &= \left(\frac{\sum Y_{.p}^2}{J} - \frac{Y_{..}^2}{IJ} \right) = \sum_{p,q} (y_{p.} - y_{..})^2 \left[\sigma \sum_{p,q} (y_{.q} - y_{..})^2 \right] \end{aligned}$$

La cantidad del punto cuatro no es más que

$$\sum_{p,q} y_{p.q}^2 - R(\mu, \tau, y)$$

Todo lo anterior se puede exponer en una tabla de análisis de - varianza como la que sigue a continuación.

Tabla de Análisis de Varianza para un Modelo con dos criterios de clasificación.

F.V.	G.L.	S.C.	C.M.	F
Total	IJ	$\sum_{ij} y_{ij}^2$		
Media	1	$\frac{Y..^2}{IJ}$		
γ (ajust.)	(J-1)	$SC\gamma = \frac{\sum Y_{.j}^2}{J} - \frac{Y..^2}{IJ}$	$CM\gamma = \frac{SC\gamma}{J-1}$	$\frac{CM\gamma}{CME}$
τ (ajust.)	(I-1)	$SC\tau = \frac{\sum Y_{i.}^2}{I} - \frac{Y..^2}{IJ}$	$CM\tau = \frac{SC\tau}{I-1}$	$\frac{CM\tau}{CME}$
Error	(I-1)(J-1)	$SCE = \sum_{ij} y_{ij}^2 - \frac{Y..^2}{IJ} - SC\gamma - SC\tau$	$CME = \frac{SCE}{(I-1)(J-1)}$	

b) Con más de una observación por celda.

Los datos que se analizan de acuerdo a las características especificadas para el caso en cuestión tienen el siguiente modelo:

$$y_{ijm} = \mu + \tau_i + \gamma_j + e_{ijm}$$

$$i = 1, 2, \dots, I; j = 1, 2, \dots, J; m = 1, 2, \dots, M$$

donde y_{ijm} es la m-ésima observación de la ij-ésima celda, e_{ijm} es la variable aleatoria asociada a la observación realizada, la cual se supone no tiene correlación alguna con las demás, que tiene media cero y varianza σ^2 .

No se olvide las definiciones preestablecidas.

Al igual que en el caso anterior la realización de pruebas de hipótesis y/o cálculo de intervalos de confianza se basarán en el supuesto de que los errores se distribuyen normalmente.

Del mismo modo, la finalidad de este modelo será estimar los contrastes $\sum e_i \tau_i$ y $\sum d_j \gamma_j$, además de probar las hipótesis $\tau_1 = \tau_2 = \dots = \tau_I = 0$ y $\gamma_1 = \gamma_2 = \dots = \gamma_J = 0$.

Estimación puntual.- Empleando todo lo expuesto en el punto a), obtenemos las ecuaciones normales del caso en cuestión.

$$\begin{aligned} \mu: IJM\hat{\mu} + JM\hat{\tau}_1 + \dots + JM\hat{\tau}_I + IM\hat{\gamma}_1 + \dots + IM\hat{\gamma}_J &= Y.. \\ \tau_p: JM\hat{\mu} + JM\hat{\tau}_p + M\hat{\gamma}_1 + \dots + M\hat{\gamma}_J &= Y_{p.} \end{aligned}$$

$p = 1, 2, \dots, I$

$$\mu_q: IM\hat{\mu} + M\hat{\tau}_1 + \dots + M\hat{\tau}_I + IM\hat{\beta}_q = Y_{.q}$$

$q = 1, 2, \dots, J$

Se puede comprobar de nueva cuenta que las únicas combinaciones - de las τ y β_q que son estimables son los contrastes.

Prueba de Hipótesis.- Aquí también, como en el modelo anterior, - para probar cualquiera de las hipótesis que se planteen, se deben calcular las cuatro cantidades ya expuestas, que son:

1.-

$$R(\mu, \tau, \beta) = \frac{Y_{...}^2}{IJM} + \left(\frac{\sum Y_{p..}^2}{JM} - \frac{Y_{...}^2}{IJM} \right) + \left(\frac{\sum Y_{.q.}^2}{IM} - \frac{Y_{...}^2}{IJM} \right)$$

2.- En este caso el modelo será:

$$y_{ijm} = (\mu + \tau) + \beta_j + e_{ijm}$$

o $y_{ijm} = \mu^* + \beta_j + e_{ijm}$ con $\mu^* = \mu + \tau$

el cual tiene como ecuaciones normales a:

$$\mu^*: IJM\hat{\mu}^* + IM\hat{\beta}_1 + \dots + IM\hat{\beta}_J = Y_{...}$$

$$\beta_q: IM\hat{\mu}^* + IM\hat{\beta}_q = Y_{.q.}$$

Empleando la combinación lineal no estimable $\sum \hat{\beta}_q = 0$, la cual da que

$$\hat{\mu}^* = \frac{Y_{...}}{IJM} \quad \text{y} \quad \hat{\beta}_q = \frac{Y_{.q.}}{IM} - \frac{Y_{...}}{IJM}$$

como $\bar{Y}'\bar{Y} = (\bar{X}'\bar{Y})' = (Y_{...}, Y_{.2.}, Y_{.2.}, \dots, Y_{.J.})'$ se tiene que

$$R(\mu, \beta) = \frac{Y_{...}^2}{IJM} + \left(\frac{\sum Y_{.q.}^2}{IM} - \frac{Y_{...}^2}{IJM} \right)$$

$$\left[\sigma R(\mu, \tau) = \frac{Y_{...}^2}{IJM} + \left(\frac{\sum Y_{p..}^2}{JM} - \frac{Y_{...}^2}{IJM} \right) \right]^{(8)}$$

3.-

$$R(\mu, \tau, \gamma) - R(\mu, \gamma) \left[\sigma R(\mu, \tau, \gamma) - R(\mu, \tau) \right] \\ = \left(\frac{\sum Y_{p\gamma}^2}{JM} - \frac{Y_{...}^2}{IJM} \right) = \sum_{p, \gamma, \tau} (y_{p..} - y_{...})^2 \left[\sigma \sum_{p, \gamma, \tau} (y_{p\gamma} - y_{...})^2 \right]$$

4.-

$$\sum_{p, \gamma, \tau} y_{p\gamma\tau}^2 - R(\mu, \tau, \gamma)$$

Todo lo cual se resume en la siguiente tabla de análisis de varian-
za.

Tabla de Análisis de Varianza para un Modelo con dos criterios de
clasificación y más de una observación por celda.

F.V.	G.L.	S.C.	C.M.	F
Total	IJM	$\sum y_{ijm}^2$		
Media	1	$Y_{...}^2 / IJM$		
γ (ajust.)	J-1	$SC\gamma = \frac{\sum Y_{j.}^2}{IM} - \frac{Y_{...}^2}{IJM}$	$CM\gamma = SC\gamma / (J-1)$	$CM\gamma / CME$
τ (ajust.)	I-1	$SC\tau = \frac{\sum Y_{. \tau}^2}{JM} - \frac{Y_{...}^2}{IJM}$	$CM\tau = SC\tau / (I-1)$	$CM\tau / CME$
Error	IJM-I-J+1	$SCE = \sum y_{ijm}^2 - \frac{Y_{...}^2}{IJM} - SC\gamma - SC\tau$	$CME = SCE / (IJM-I-J+1)$	

III.- Modelos con interacción.- Antes de que procedamos a estudiar
los modelos que consideran ya el fenómeno de la interacción, se ha
ce necesario ver que se entiende por interacción o por no-interac-
ción.

Considérese en primer lugar, que tenemos una función de dos variables, sea $f(x_1, x_2)$ y fórmúlese la siguiente definición,

Definición 1.- $f(x_1, x_2)$ se definirá como una función sin interacción si y sólo si existen funciones $J_1(x_1)$ y $J_2(x_2)$ tal -- que

$$f(x_1, x_2) = J_1(x_1) + J_2(x_2)$$

Ahora bien si el modelo para datos con dos criterios de clasificación se define como

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

donde μ_{ij} se reconoce como la suma del i-ésimo efecto de A, el cual es τ_i , más el j-ésimo efecto de B, el cual es β_j , más μ , tenemos entonces que $\mu_{ij} = \mu + \tau_i + \beta_j$. Esta es una función de τ_i y β_j sin interacción.

En una forma general se observa que

$$(\mu_{ij} - \mu_{i'j}) - (\mu_{ij'} - \mu_{i'j'}) = (\tau_i - \tau_{i'}) - (\tau_i - \tau_{i'}) = 0$$

para toda i, i', j y j' , nótese sin embargo que reanreglando los términos tenemos que:

$$(\mu_{ij} - \mu_{ij'}) - (\mu_{i'j} - \mu_{i'j'}) = (\beta_j - \beta_{j'}) - (\beta_j - \beta_{j'}) = 0$$

por lo tanto tenemos la siguiente definición:

Definición 2.- El modelo para datos con dos criterios de clasificación $y_{ij} = \mu_{ij} + e_{ij}$ será un modelo aditivo o en otras palabras sin interacción, si μ_{ij} es tal que $(\mu_{ij} - \mu_{ij'}) - (\mu_{i'j} - \mu_{i'j'}) = 0$ para toda i, i', j y j' . En caso contrario se dirá que es un modelo no aditivo o con interacción.

a) Con una observación por celda.

El modelo correspondiente para datos analizados de acuerdo con las características en cuestión se define como:

$$y_{ij} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + e_{ij}$$

con $i = 1, 2, 3, \dots, I$; $j = 1, 2, 3, \dots, J$ el cual difiere del planteado en II a) en el término $(\tau\beta)_{ij}$, que no representa el producto de los niveles i e j -ésimo de los factores A y B respectivamente (o viceversa); sino que representa la cantidad de interacción que el factor A tiene sobre el factor B (o viceversa) en los niveles i e j -ésimo respectivamente.

Nótese que además de lo planteado en otros casos, aquí se intentará probar la hipótesis de que $(\tau\beta)_{pq} = 0$ para todo p y q .

Estimación Puntual.- Empleando nuevamente lo expuesto en los casos anteriores, se obtienen las ecuaciones normales abreviadas - las cuales son:

$$\begin{aligned} \mu: & IJ\hat{\mu} + J(\sum \hat{\tau}_i) + I(\sum \hat{\beta}_j) + \sum (\hat{\tau\beta})_{ij} = Y.. \\ \tau_p: & J\hat{\mu} + J\hat{\tau}_p + \sum \hat{\beta}_j + \sum_j (\hat{\tau\beta})_{pj} = Y_{p.} \\ \beta_q: & I\hat{\mu} + \sum \hat{\tau}_i + I\hat{\beta}_q + \sum_i (\hat{\tau\beta})_{iq} = Y_{.q} \\ (\tau\beta)_{pq}: & \hat{\mu} + \hat{\tau}_p + \hat{\beta}_q + (\hat{\tau\beta})_{pq} = Y_{pq} \end{aligned}$$

para $p = 1, 2, \dots, I$ e $q = 1, 2, \dots, J$.

No es difícil comprobar que en este caso los contrastes de las τ_i o de las β_j no son estimables, debido a que los parámetros $(\tau\beta)_{p.}$ o $(\tau\beta)_{.q}$ no pueden ser separados de τ_p o β_q respectivamente (Hágase para los contrastes sencillos como son $\tau_i - \tau_{i'}$ y $\beta_j - \beta_{j'}$, con $i \neq i'$ y $j \neq j'$).

Sin embargo el modelo propuesto puede ser escrito reparametrizado en la siguiente forma

$$y_{ij} = [\mu + \bar{\tau} + \bar{\beta} + (\tau\bar{\beta})_{..}] + [\tau_i - \bar{\tau} + (\tau\bar{\beta})_{i.} - (\tau\bar{\beta})_{..}] + [\beta_j - \bar{\beta} + (\tau\bar{\beta})_{.j} - (\tau\bar{\beta})_{..}] + [(\tau\beta)_{ij} - (\tau\bar{\beta})_{i.} - (\tau\bar{\beta})_{.j} + (\tau\bar{\beta})_{..}] + e_{ij} \quad (1)$$

en donde

$$\begin{aligned} \bar{x}_i &= \frac{1}{I} \sum x_{ij} ; \bar{y}_j = \frac{1}{J} \sum y_{ij} ; (\bar{xy})_{ij} = \frac{1}{J} \sum (xy)_{ij} \\ (\bar{xy})_{ij} &= \frac{1}{I} \sum (xy)_{ij} \quad y \quad (\bar{xy})_{..} = \frac{1}{IJ} \sum (xy)_{ij} \end{aligned} \quad (2)$$

en este caso el modelo se puede representar como

$$y_{ij} = \mu^* + \tau_i^* + \beta_j^* + (xy)_{ij}^* + e_{ij}$$

Las ecuaciones normales del último modelo son:

$$\begin{aligned} \mu^* : & \quad IJ \hat{\mu}^* &= Y_{..} \\ \tau_p^* : & \quad J \hat{\mu}^* + J \hat{\tau}_p^* &= Y_{p.} \\ \beta_q^* : & \quad I \hat{\mu}^* + I \hat{\beta}_q^* &= Y_{.q} \\ (xy)_{pq} : & \quad \hat{\mu}^* + \hat{\tau}_p^* + \hat{\beta}_q^* + (\hat{xy})_{pq}^* &= Y_{pq} \end{aligned}$$

y se puede observar que los parámetros τ_i^* o β_j^* son estimables y además que las ecuaciones en realidad siguen siendo las mismas - que se obtuvieron en un principio.

Prueba de Hipótesis. - La realización de pruebas de hipótesis requiere del cálculo de las cantidades explicadas en la parte II a).

Sin embargo dichas pruebas de hipótesis no se podrán hacer por - lo que se verá más adelante.

Obtención de la cantidad del punto uno.

Para obtener esta cantidad se hace necesario tener una solución de $\hat{\beta}$, por lo que se emplearán las siguientes combinaciones no estimables, $\sum \hat{\tau}_i = 0$; $\sum \hat{\rho}_j = 0$; $(\hat{\tau}\hat{\rho})_{ij} = 0$ para $i = 1, 2, \dots, I$ e $(\hat{\tau}\hat{\rho})_{ij} = 0$ para $j = 1, 2, \dots, J$.

La solución para $\hat{\beta}$ sería:

$$\hat{\mu} = y_{..}$$

$$\hat{\tau}_p = y_{p.} - y_{..}$$

$$\hat{\rho}_q = y_{.q} - y_{..}$$

$$(\hat{\tau}\hat{\rho})_{pq} = y_{pq} - y_{p.} - y_{.q} + y_{..}$$

para $p = 1, 2, \dots, I$ e $q = 1, 2, \dots, J$,

y entonces tenemos que la cantidad es:

$$R[\mu, \tau, \rho, (\tau\rho)] = \frac{Y_{..}^2}{IJ} + \left(\frac{\sum Y_{i.}^2}{J} - \frac{Y_{..}^2}{IJ} \right) + \left(\frac{\sum Y_{.j}^2}{I} - \frac{Y_{..}^2}{IJ} \right) + \sum_{ij} y_{ij} \left(y_{ij} - \frac{Y_{i.}}{J} - \frac{Y_{.j}}{I} + \frac{Y_{..}}{IJ} \right) = \sum_{ij} y_{ij}^2$$

Antes de seguir con cualquier otro cálculo se puede ver que el punto cuatro que es la suma de errores al cuadrado menos $R(\mu, \tau, \rho, (\tau\rho))$ va a resultar ser cero y por lo tanto no habrá un estimador de la varianza que permita la realización de pruebas de hipótesis o cálculo de intervalos de confianza.

Es de notarse, que para obtener un estimador de la varianza --- sólo se necesita una observación más para al menos una de las celdas.

Por lo tanto se procederá con el modelo que considerando la interacción tiene más de una observación por celda.

b) Con más de una observación por celda.

Supongamos ahora, el mismo modelo que en el caso anterior, nada más con la diferencia como ya se dijo, de que aquí tendremos M ($M > 1$) observaciones por celda.

Notamos entonces que el modelo en cuestión es

$$y_{ijm} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + e_{ijm}$$

con $i = 1, 2, \dots, I$; $j = 1, 2, \dots, J$ y $m = 1, 2, \dots, M$.

Estimación Puntual.- Las ecuaciones normales para este caso son:

$$\mu: IJM\hat{\mu} + JM(\sum \hat{\tau}_i) + IM(\sum \hat{\beta}_j) + M(\sum (\hat{\tau}\beta)_{ij}) = Y_{...}$$

$$\tau_p: JM\hat{\mu} + JM\hat{\tau}_p + M(\sum \hat{\beta}_j) + M(\sum (\hat{\tau}\beta)_{pj}) = Y_{p..}$$

$$\beta_q: IM\hat{\mu} + M(\sum \hat{\tau}_i) + IM\hat{\beta}_q + M(\sum (\hat{\tau}\beta)_{iq}) = Y_{.q.}$$

$$(\tau\beta)_{pq}: M\hat{\mu} + M\hat{\tau}_p + M\hat{\beta}_q + M(\hat{\tau}\beta)_{pq} = Y_{pq.}$$

Empleando las condiciones no estimables $\hat{\tau}_i = \hat{\beta}_j = (\hat{\tau}\beta)_{ij} = 0$ para $i = 1, 2, \dots, I$ e $j = 1, 2, \dots, J$ se obtiene la siguiente solución:

$$\hat{\mu} = y_{...}$$

$$\hat{\tau}_p = y_{p..} - y_{...}$$

$$\hat{\beta}_q = y_{.q.} - y_{...}$$

$$(\hat{\tau}\beta)_{pq} = y_{pq.} - y_{p..} - y_{.q.} + y_{...}$$

con $p = 1, 2, \dots, I$ e $q = 1, 2, \dots, J$.

Prueba de Hipótesis.- Antes de proceder con la prueba más importante que se puede hacer en el modelo en cuestión, se justificará su validez.

Usemos nuevamente la otra representación que al modelo que considera la interacción, se le puede dar (ver ecuaciones 1 y 2), el modelo en este caso se representa como:

$$y_{ijm} = \mu^* + \tau_i^* + \beta_j^* + (\tau\beta)_{ij}^* + \rho_{ijm}$$

con $i = 1, 2, \dots, I$; $j = 1, 2, \dots, J$ y $m = 1, 2, \dots, M$.

Se verá que no existirá interacción en el modelo si y sólo si $(\tau\beta)_{ij} - (\tau\bar{\beta})_{i.} - (\tau\bar{\tau})_{.j} + (\tau\bar{\tau})_{..}$ es igual a cero para toda i, j , - pero esto por la definición 2 sería lo mismo que si demostráramos que $(\mu_{ij} - \mu_{i'j}) - (\mu_{ij'} - \mu_{i'j'}) = 0$ para toda i, i', j y j' donde $\mu_{ij} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij}$.

Demostración.- Si tenemos que

$$(\mu_{ij} - \mu_{i'j}) - (\mu_{ij'} - \mu_{i'j'}) = 0$$

para toda i, i', j y j' observamos que

$$\begin{aligned} & \frac{1}{IJ} \sum_{i,j} [(\mu_{ij} - \mu_{i'j}) - (\mu_{ij'} - \mu_{i'j'})] \\ &= \mu_{.j} - \bar{\mu}_{i'.} - \bar{\mu}_{i.} + \bar{\mu}_{..} = 0 \end{aligned}$$

para toda i, i', j .

Más esta última expresión en realidad es igual a

$$(\tau\beta)_{ij} - (\tau\bar{\beta})_{.j} - (\tau\bar{\tau})_{i.} + (\tau\bar{\tau})_{..}$$

Ahora bien supongamos que $(\tau\mu)_{ij}^* = 0$ para toda $i \in j$, entonces tenemos que

$$(\tau\mu)_{ij}^* - (\tau\mu)_{i'j}^* - (\tau\mu)_{ij'}^* + (\tau\mu)_{i'j'}^*$$

debe ser igual a cero para toda i, i', j y j' , pero sustituyendo cada $(\tau\mu)_{ij}^*$ por $\mu_{ij} - \mu - \tau_i - \mu_j$ obtenemos la siguiente igualdad

$$\begin{aligned} 0 &= (\tau\mu)_{ij}^* - (\tau\mu)_{i'j}^* - (\tau\mu)_{ij'}^* + (\tau\mu)_{i'j'}^* \\ &= (\mu_{ij} - \mu_{i'j}) - (\mu_{ij} - \mu_{i'j'}) \end{aligned}$$

con lo cual, lo supuesto queda demostrado.

Obsérvese que la notación en si no cambia las ecuaciones y por lo tanto, emplear la solución que se obtenga para el sistema de ecuaciones con asterisco en los elementos, es lo mismo que si se emplea la solución obtenida en la estimación puntual.

Procedamos pues a calcular los puntos 1 a 4 para la realización de la prueba.

1.-

$$\begin{aligned} R(\mu, \tau, \mu, [\tau\mu]) &= \frac{Y_{...}^2}{IJM} + \left(\frac{\sum Y_{0..}^2}{JM} - \frac{Y_{...}^2}{IJM} \right) + \left(\frac{\sum Y_{.0.}^2}{IM} - \frac{Y_{...}^2}{IJM} \right) \\ &+ \sum_{i,j,m} (y_{ij.} - y_{i..} - y_{.j.} + y_{...}) y_{ijm} = \sum_{i,j} \frac{y_{ij.}^2}{M} \end{aligned}$$

2.- En este caso el modelo más importante a suponer es el de -- que

$$y_{ijm} = \mu + \tau_i + \mu_j + e_{ijm}$$

o sea el modelo que resulta de suponer en el modelo en cuestión que no existe interacción, sin dejar de considerar los modelos que para el caso de sin interacción se prueban en este mismo -- punto, los cuales son iguales.

Antes de seguir adelante hagamos una importante observación.

Según la notación empleada y de acuerdo con los fundamentos con los cuales se calculan los modelos en este punto, también sería posible considerar los modelos $(\mu, \beta, [\tau\beta I])$, $(\mu, \tau, [\tau\beta I])$ o $(\mu, [\tau\beta I])$. Sin embargo resulta que cada uno de estos posibles modelos, en la reducción en suma de cuadrados correspondientes, da una cantidad igual a la del modelo inicialmente considerado $(\mu, \tau, \beta, [\tau\beta I])$ y por lo tanto las cantidades que se calculan en el punto 3 resultarían ser cero, cualquiera que fuera el modelo que se considerará y por consiguiente siempre -resultaría que $F_c < F_r$ con las consabidas consecuencias.⁹

Debe de observarse también, que razonablemente los modelos mencionados en realidad no tienen ningún sentido, pues en general sería ilógico pensar en un modelo estadístico, que considerara la presencia de términos de interacción entre dos factores, si en si mismo el modelo no tiene términos asociados con dos factores distintos.

El cálculo correspondiente, para la suposición hecha es el siguiente:

$$\begin{aligned} R(\mu, \tau, \beta) &= \frac{Y_{...}^2}{IJM} + \left(\frac{\sum Y_{p..}^2}{JM} - \frac{Y_{...}^2}{IJM} \right) + \left(\frac{\sum Y_{.q.}^2}{IM} - \frac{Y_{...}^2}{IJM} \right) \\ &= \left(\frac{\sum Y_{p..}^2}{JM} + \frac{\sum Y_{.q.}^2}{IM} - \frac{Y_{...}^2}{IJM} \right) \end{aligned}$$

$$\begin{aligned} 3.- \quad R(\mu, \tau, \beta, [\tau\beta I]) - R(\mu, \tau, \beta) \\ &= \sum_{ij} \frac{Y_{ij.}^2}{M} - \frac{\sum Y_{p..}^2}{JM} - \frac{\sum Y_{.q.}^2}{IM} + \frac{Y_{...}^2}{IJM} \\ &= \sum_{ij,m} (y_{ij.} - y_{i..} - y_{.j.} + y_{...})^2 \end{aligned}$$

4.- Por último tenemos que

$$\sum_{ij,m} y_{ijm}^2 - R(\mu, \tau, \beta, [\tau\beta I]) = \sum_{ij,m} (y_{ijm} - y_{ij.})^2$$

Los cuatro puntos anteriores, como en los otros casos, también se presentan en una forma abreviada, empleando una tabla, como la que a continuación se da.

Tabla de Análisis de Varianza para la prueba de la interacción en un modelo con dos criterios de clasificación y más de una observación por celda.

F.V.	G.L.	S.C.	C.M.	F
Total	IJM	$\sum_{ijm} y_{ijm}^2$		
(τ_j) (ajustada)	$(I-1)(J-1)$	$SCT_j = \sum_{ijm} (y_{ij.} - y_{i..} - y_{.j.} + y_{...})^2$	$CM_{\tau_j} = \frac{SCT_j}{(I-1)(J-1)}$	CM_{τ_j}/CHE
Error	$IJ(M-1)$	$SCE = \sum_{ijm} (y_{ijm} - y_{ij.})^2$	$CHE = \frac{SCE}{IJ(M-1)}$	

Referencias.

- 1.- Graybill Franklin A., An Introduction to Linear Statistical Models, Volume I, New York McGraw-Hill Book Co. Inc., 1961 (First Edition).
- 2.- Lindgren Bernard W., Statistical Theory, New York. Macmillan Publishing, 1976 (Third Edition).
- 3.- Searle S.R., Linear Models, New York: John Wiley & Sons, Inc. 1971 (First Edition).

CAPITULO III

Análisis de la interacción en datos con dos criterios de clasificación y una observación por celda.

I.- Introducción.- Del capítulo anterior se pudo notar que la consideración de la "interacción" para el análisis de datos, es de una importancia elemental, pues aún cuando el modelo lineal general aditivo, puede ser "completo", pudiera ser que la consideración de la interacción ayudara a describir más correctamente los datos en cuestión, por lo que siempre se hace recomendable probar un modelo más completo o que considere la interacción.

Otra cosa que se pudo observar, y que forma la base de existencia del presente capítulo, fué la imposibilidad de realizar el análisis de varianza para el modelo que considera la interacción y que tiene una observación por celda, caso muy particular, el cual ha sido estudiado con diversos puntos de vista por Tukey, Mandel, -- Graybill, Milliken, Scheffé, Méndez y Bustos, entre otros, cuyos enfoques o métodos se bosquejarán a continuación.

II.- Enfoque de Tukey.

Tukey fué uno de los primeros que realizó estudios acerca de la interacción en tablas de doble entrada y una observación por celda.

El modelo que él propuso fué:

$$(1) \quad y_{ij} = \mu + \tau_i + \beta_j + \lambda \tau_i \beta_j + e_{ij}$$

con $i = 1, 2, \dots, I$ e $j = 1, 2, \dots, J$.

suponiendo que las e_{ij} se distribuyen normalmente con media cero y varianza σ^2 y las condiciones laterales $\sum \tau_i = \sum \beta_j = 0$, se cumplen.

La prueba de más interés en este caso, como se dijo en el capítulo anterior, consiste en considerar que la interacción es nula en cada "celda", y que para el modelo en cuestión sería probar la hipótesis $H_0: \lambda = 0$.

A continuación se bosquejan los fundamentos en que se basa la prueba de la hipótesis anterior.

Como se recordará, la pérdida de la linealidad en la esperanza de las observaciones y_{ij} , impide la estimación directa de los paráme-

tros μ , $\{\tau_i\}$, $\{\mu_j\}$ y λ . Sin embargo supóngase que los parámetros -- $\{\tau_i\}$ y $\{\mu_j\}$ son conocidos y minimícese la siguiente expresión por el método de mínimos cuadrados:

$$(2) \sum_i \sum_j e_{ij}^2 = \sum_i \sum_j (y_{ij} - \mu - \tau_i - \mu_j - \lambda \tau_i \mu_j)^2$$

como podrá comprobarse, se encuentra que el estimador de λ es:

$$(3) \quad \hat{\lambda} = \frac{\sum_i \sum_j \tau_i \mu_j y_{ij}}{\sum_i \tau_i^2 \sum_j \mu_j^2}$$

ahora, si se considera una suma de cuadrados para la interacción en (1), tenemos que ésta, se haya dada por

$$(4) \quad \lambda^2 \sum_i \tau_i^2 \sum_j \mu_j^2$$

la cual es igual a

$$(5) \quad \frac{[\sum_i \sum_j \tau_i \mu_j y_{ij}]^2}{\sum_i \tau_i^2 \sum_j \mu_j^2}$$

si reemplazamos en (5) las τ_i y μ_j por sus estimadores $\hat{\tau}_i$ e $\hat{\mu}_j$ obtendremos que

$$SC\lambda = \frac{[\sum_i \sum_j \hat{\tau}_i \hat{\mu}_j y_{ij}]^2}{\sum_i \hat{\tau}_i^2 \sum_j \hat{\mu}_j^2}$$

donde los estimadores $\{\hat{\tau}_i, \hat{\mu}_j\}$ serían estimados del modelo

$$y_{ij} = \mu + \tau_i + \mu_j + e_{ij}$$

con e_{ij} distribuyéndose como $N(0, \sigma^2)$ y $\sum \tau_i = \sum \mu_j = 0$ para toda i e j .

En este momento se intenta emplear $SC\lambda$, con la intención de rechazar la $H_0: \lambda=0$ si resulta que la $SC\lambda$ es "suficientemente grande", lo cual se presenta en el siguiente teorema que nada más se enuncia.

Teorema.-

Suponga que tenemos el siguiente modelo

$$y_{ij} = \mu + \tau_i + \mu_j + e_{ij}$$

con e_{ij} distribuyéndose como $N(0, \sigma^2)$ y las condiciones laterales $\sum \tau_i = \sum \beta_j = 0$, y que definimos

$$SC\lambda = \frac{[\sum_i \sum_j \tau_i \beta_j y_{ij}]^2}{\sum_i \tau_i^2 \sum_j \beta_j^2}$$

$$SCRXC = \sum_i \sum_j (y_{ij} - \bar{y} - \tau_i - \beta_j)^2$$

$$y \quad SCE_{M.N.A.} = SCRXC - SC\lambda$$

donde $\tau_i = y_{i.} - y_{..}$, $\beta_j = y_{.j} - y_{..}$.

Entonces bajo el modelo propuesto, tenemos que, $SC\lambda/\sigma^2$ y $SCE_{M.N.A.}/\sigma^2$ se distribuyen independientemente y tienen una distribución χ^2 con uno e $IJ - I - J$ grados de libertad respectivamente.

Demostración.- Ver referencia 14.

La tabla de análisis de varianza que se genera en este caso es la siguiente:

Tabla de Análisis de Varianza generada por el
Modelo de Tukey.

F.V.	G.L.	S. de C.	C.M.	F
Total	IJ	$\sum \sum y_{ij}^2$		
Media	1	$y_{..}^2/IJ$		
τ (ajustada)	$I-1$	$SC\tau = \frac{\sum y_{i.}^2}{I} - \frac{y_{..}^2}{IJ}$	$CM\tau = \frac{SC\tau}{I-1}$	$CM\tau/CM_{E_{M.N.A.}}$
β (ajustada)	$J-1$	$SC\beta = \frac{\sum y_{.j}^2}{J} - \frac{y_{..}^2}{IJ}$	$CM\beta = \frac{SC\beta}{J-1}$	$CM\beta/CM_{E_{M.N.A.}}$
ERROR $M.A.$ *	$(I-1)(J-1)$	$SCE_{M.A.} = \sum \sum y_{ij}^2 - \frac{y_{..}^2}{IJ} - SC\tau - SC\beta$	$CM_{E_{M.A.}} = \frac{SCE_{M.A.}}{(I-1)(J-1)}$	
No-aditividad	1	$SC\lambda = \frac{[\sum_i (y_{i.} - \bar{y})(y_{.j} - \bar{y}) y_{ij}]^2}{I(\sum_i (y_{i.} - \bar{y})^2) J(\sum_j (y_{.j} - \bar{y})^2)}$	$CM\lambda = SC\lambda$	$F\lambda = \frac{CM\lambda}{CM_{E_{M.N.A.}}}$
Error $M.N.A.$ **	$IJ - I - J$	$SCE_{M.N.A.} = SCE_{M.A.} - SC\lambda$	$CM_{E_{M.N.A.}} = \frac{SCE_{M.N.A.}}{IJ - I - J}$	

* M.A. = Modelo aditivo.

** M.N.A. = Modelo no aditivo.

Donde si F_{λ} es mayor que una $F_{I, IJ-1-J}$ de tablas con un nivel de significancia α , la hipótesis $H_0: \lambda=0$ se rechaza.

Por último es importante hacer notar que esta prueba de aditividad de Tukey, será en general un dispositivo para indicar cierto tipo de transformación en la información, para producir aditividad en los datos.

III.- Enfoque de Graybill y Milliken.^[1]

Este enfoque tiene como propósito demostrar que la teoría del modelo lineal general, puede ser usada para probar la hipótesis --- $H_0: \alpha = \bar{0}$, en el siguiente modelo

$$(1) \quad \bar{Y} = \bar{X}\bar{\beta} + \bar{F}\alpha + \bar{e} \quad [2]$$

donde $\bar{Y}$ es un vector de observaciones de $n \times 1$, $\bar{X}$, es una matriz de constantes conocidas de $n \times q$ y rango $k (k < q)$, $\bar{\beta}$ es un vector de parámetros desconocidos, $\bar{F}$ es una matriz de $n \times l$ cuya forma funcional es conocida y es función de las funciones estimables de $\bar{\beta}$ ($\bar{F} = [f_{ij}(\bar{x}\bar{\beta})]$), α un vector desconocido y $\bar{e}$ un vector de errores aleatorios no observable.

Antes de proceder con la teoría, es importante mencionar que el modelo (1) es una ligera generalización del modelo propuesto por Scheffé^[2] el cual a su vez es una generalización del modelo propuesto por Tukey.

El modelo (1) ya planteado se puede representar también como:

$$\bar{Y} = [\bar{X}, \bar{F}] \begin{bmatrix} \bar{\beta} \\ \alpha \end{bmatrix} + \bar{e}$$

donde $[\bar{X}, \bar{F}]$ es de rango r con $k < r < n$, y la hipótesis que se intenta probar es $H_0: \alpha = \bar{0}$ contra la alternativa $H_0: \alpha \neq \bar{0}$, por el momento supóngase que $\bar{F}$ es una matriz de números conocidos^[3] y observemos que la hipótesis planteada es muy semejante a la que se realiza en el teorema 4 de la parte 1 del primer capítulo, por lo tanto, es posible desarrollar un procedimiento similar al empleado en dicho teorema.

La función de prueba será:

$$v = \frac{Q_1 \cdot (n-r)}{Q_0 \cdot (r-l)}$$

donde $Q_1 = \bar{Y}' [(\bar{I} - \bar{X}\bar{X}') \bar{F}] [(\bar{I} - \bar{X}\bar{X}') \bar{F}]' \bar{Y} \quad [4]$

$$Q_0 = \bar{Y}' [\bar{I} - \bar{X}\bar{X}'] \bar{Y} - Q_1$$

y donde el símbolo A^- denota a la inversa generalizada de Moore-Penrose de una matriz A .

Bajo la hipótesis a probar se obtienen las siguientes propiedades distribucionales:

A- Q_0/σ^2 se distribuye como $\chi^2(n-r)$

B- Q_1/σ^2 se distribuye como $\chi^2(r-l)$

C- Q_0 y Q_1 son independientes, y

D- $v = \frac{Q_1 \cdot (n-r)}{Q_0 \cdot (r-l)}$ se distribuye como $F(r-l, n-r)$

"Por lo tanto v podría ser usada como una estadística de prueba para H_0 excepto que no puede calcularse de los datos observados por que la matriz F es desconocida", ahora si tenemos en cuenta -- que F es una función de las cantidades desconocidas $\bar{x}\beta$; y además que la forma funcional de cada elemento de F es conocida, observamos que podemos modificar v de tal manera que pueda ser calculada de los datos y aún usada para probar H_0 , si reemplazamos F en Q_0 y Q_1 por $\hat{F}$ donde se emplea $\bar{x}\bar{x}^{-}\bar{y}$ para $\bar{x}\beta$ en $f_{ij}(\bar{x}\beta)$ o sea $\bar{x}\hat{\beta}$ por $\bar{x}\beta$, donde $\bar{x}\hat{\beta}$ es el estimador de mínimos cuadrados de $\bar{x}\beta$ en el modelo bajo la hipótesis H_0 . Por lo tanto tenemos que $\hat{F} = [f_{ij}(\bar{x}\bar{x}^{-}\bar{y})]$ y observamos que $[\bar{x}, \hat{F}]$ tiene rango r con probabilidad 1.

Por lo tanto después de la sustitución tenemos que:

$$\hat{Q}_1 = \bar{y}'[(\bar{I} - \bar{x}\bar{x}^{-})\hat{F}][(\bar{I} - \bar{x}\bar{x}^{-})\hat{F}]^{-}\bar{y}$$

$$\hat{Q}_0 = \bar{y}'[\bar{I} - \bar{x}\bar{x}^{-}]\bar{y} - \hat{Q}_1$$

$$\hat{v} = \frac{\hat{Q}_1(n-r)}{\hat{Q}_0(r-l)}$$

Pero $\bar{y}'(\bar{I} - \bar{x}\bar{x}^{-})\bar{y}$ es independiente de $\bar{x}\hat{\beta}$ y de aquí también se -- tiene que es independiente de $\hat{F}$.

Por lo tanto, bajo el modelo dado por la hipótesis H_0 , la distribución condicional de $\hat{Q}_1$ dada $\hat{F} = \hat{F}^*$ (una matriz constante) no -- depende de $\hat{F}^*$, lo mismo sucede para $\hat{Q}_0$, por lo tanto también $\hat{Q}_1$ y $\hat{Q}_0$ son independientes. Luego entonces, cuando H_0 es verdadera, te

nemos que $\hat{v}$ tiene la misma distribución que v en (D) y por lo -- tanto $\hat{v}$ puede ser calculada de los datos observados $\bar{X}$ y $\bar{Y}$. De todo lo anterior, concluimos que $\hat{v}$ es una estadística de prueba -- de la hipótesis $\bar{\alpha} = \bar{\delta}$.

Nótese también que v es el cociente de dos formas cuadráticas en las X_{ij} y que $\hat{v}$ no lo es, y por último, que cuando $\bar{\alpha} \neq \bar{\delta}$, v es una F no central y que $\hat{v}$ no es una F no central e incluso poco se conoce de $\hat{v}$ cuando $\bar{\alpha} \neq \bar{\delta}$.

IV.- Enfoque de Mandel.

Antes de proceder con el enfoque en cuestión es importante hacer notar que el presente enfoque, ataca el problema de la interac--- ción de una manera completamente sistemática, dado que se parte la interacción (η_{ij}) en tantos términos individuales como los mismos datos lo requieran, además nótese que no se hace ninguna suposi--- ción previa sobre linealidad.

Este método o enfoque se deriva del de componentes principales, -- usado para el análisis de factores, en análisis estadístico multi variado, cuyo estudio se debe principalmente a Harold Hotelling.

Mientras que los resultados matemáticos en los que se apoya el -- presente método son conocidos, el punto de vista es realmente no- vedoso.

Otra cosa importante de mencionar es el hecho de que el presente enfoque, trata indistintamente a los renglones y a las columnas, quedando la asignación del papel de renglón o columna sujeto a -- una elección subjetiva de los criterios que se han mencionado con anterioridad.

Por último, debe observarse que el enfoque en cuestión podrá ser usado para probar la aditividad de efectos de los criterios, para lo cual será preciso formular ciertas suposiciones en torno de la forma particular en que se expresa el efecto de interacción en el modelo en cuestión.

El modelo que considera Mandel es

$$(1) \quad y_{ij} = \mu + \tau_i + \beta_j + \eta_{ij}$$

con $i = 1, 2, \dots, I$ e $j = 1, 2, \dots, J$.

donde μ es una literal asignada para la media general, las τ_i y las β_j son las literales asignadas para los efectos principales por renglón y por columna respectivamente, y las literales η_{ij} representan la interacción entre el renglón i -ésimo y la columna j -ésima; obviamente las condiciones laterales acostumbradas se tienen, recuerdense que son:

$$(2) \quad \sum \tau_i = 0 \quad \sum \beta_j = 0 \quad \sum_i \eta_{ij} = \sum_j \eta_{ij} = 0$$

las η_{ij} muy frecuentemente se consideran como una variable aleatoria con media cero y varianza σ^2 , en este caso, al modelo (1) se le conoce como aditivo, sin embargo, esto último no siempre es válido, en cuya situación se recomienda considerar alguna otra "estructura" de la η_{ij} .

Mandel supone que la estructura de la η_{ij} es la siguiente:

$$(3) \quad \eta_{ij} = \theta u_i v_j + \theta' u_i' v_j' + \theta'' u_i'' v_j'' + \dots + e_{ij}$$

donde e_{ij} es tal que $E(e_{ij}) = 0$, $V(e_{ij}) = \sigma^2$ y $E(e_{ij} e_{i'j'}) = 0$ para $i \neq i'$ e $j \neq j'$, cumpliéndose además las siguientes restricciones

$$(4) \quad \sum u_i = \sum u_i' = \sum u_i'' = \dots = \sum v_j = \sum v_j' = \dots = 0$$

$$(5) \quad \sum u_i^2 = \sum u_i'^2 = \sum u_i''^2 = \dots = \sum v_j^2 = \sum v_j'^2 = \dots = 1$$

De lo anterior observamos que el problema que tenemos ahora, es encontrar estimadores para los parámetros θ , θ' , θ'' , ..., u_i , u_i' , u_i'' , ..., v_j , v_j' , ..., y para la varianza del error σ^2 .

La estimación se realizará empleando el método de mínimos cuadrados.

Un estimador de mínimos cuadrados del efecto de interacción lo forma el residuo d_{ij} que se da por la siguiente relación

$$\begin{aligned} d_{ij} &= Y_{ij} - \hat{\mu} - \hat{\tau}_i - \hat{\beta}_j \\ \text{donde} \quad \hat{\mu} &= y_{..} = (1/ij) \sum \sum Y_{ij} \\ \hat{\tau}_i &= y_{i.} - y_{..} = (1/j) \sum Y_{ij} - y_{..} \\ \hat{\beta}_j &= y_{.j} - y_{..} = (1/i) \sum Y_{ij} - y_{..} \end{aligned}$$

Con el objeto de ir ajustando la interacción, suponga primero que ésta, está dada por la siguiente relación

$$\eta_{ij} = \theta u_i v_j + e_{ij}$$

donde u_i y v_j se hallan sujetas a las condiciones (4) y (5), por lo que los estimadores de mínimos cuadrados se obtienen minimizando la siguiente expresión:

$$H = \sum_i \sum_j (d_{ij} - \theta u_i v_j)^2 - \lambda_1 (\sum u_i^2 - 1) - \lambda_2 (\sum v_j^2 - 1) - 2\mu_1 (\sum u_i) - 2\mu_2 (\sum v_j)$$

donde $\lambda_1, \lambda_2, \mu_1$ y μ_2 son multiplicadores de Lagrange.

Después de efectuar los cálculos correspondientes se obtienen las siguientes relaciones:

$$(6) \quad u_k = \theta^{-1} \sum_j d_{kj} v_j$$

$$(7) \quad v_k = \theta^{-1} \sum_i d_{ik} u_i$$

sustituyendo (7) en (6) y si se hace

$$(8) \quad S_{it} = \sum_j d_{ij} d_{tj}$$

se obtiene que:

$$(9) \quad \theta^2 u_t = \sum_i u_i S_{it}$$

en notación matricial tenemos que, la ecuación (8) llega a ser:

$$(10) \quad \bar{S} = \bar{D} \cdot \bar{D}'$$

con $\bar{S} \in M_{1 \times 1}$, $\bar{D} \in M_{1 \times r}$, donde $\bar{D}$ representa a la matriz de residuos y el apóstrofe es para la matriz transpuesta. Entonces (9) se convierte en:

$$(11) \quad \theta^2 \bar{u} = \bar{S} \bar{u}$$

con $\tilde{u} \in \mathbb{R}^I$; de (11) se observa que θ^2 es un eigenvalor de la matriz $\tilde{S}$ y $\tilde{u}$ el eigenvector asociado.

En una forma similar se puede demostrar que

$$(12) \quad \theta^2 \tilde{v} = (\tilde{D}'\tilde{D})\tilde{v}$$

con $\tilde{v} \in \mathbb{R}^J$; y por lo tanto la minimización de H , se logrará -- considerando como θ^2 el eigenvalor más grande, ya sea de la matriz $\tilde{S}$ o de la matriz $(\tilde{D}'\tilde{D})$. Además puede probarse el hecho de -- que habiendo escogido como θ^2 al eigenvalor más grande de la matriz $\tilde{S}$, si se deseara minimizar la siguiente expresión

$$(13) \quad \sum_i \sum_j ((d_{ij} - \theta u_i v_j) - \theta' u_i' v_j')^2$$

con respecto a $\theta' u_i'$ y v_j' , con u_i' y v_j' sujetos a las restricciones (4) y (5), se obtendrá que el valor mínimo se logra tomando a θ'^2 como el segundo eigenvalor más grande de la matriz $\tilde{S}$ original, con $\tilde{u}'$ y $\tilde{v}'$ como los vectores asociados en sus respectivos casos.

Por lo tanto, obteniendo los eigenvalores de cualquiera de las -- dos matrices y los dos conjuntos de vectores asociados en sus respectivos casos, es suficiente para tener los estimadores de mínimos cuadrados para todos los parámetros de (3).

Por otro lado, ya que los residuos se obtienen a partir de los estimadores que se calculan por el método de mínimos cuadrados, se tiene que la matriz $\tilde{D}\tilde{D}' = \tilde{S}$ tiene a lo más rango igual a $\{\min(I, J) - 1\}$; es decir existirán cuando mucho $\{\min(I, J) - 1\}$ eigenvalores distintos de cero y por lo tanto, se tendrán como máximo $\{\min(I, J) - 1\}$ términos del tipo $\theta u_i v_j$.

La manera en que Mandel determina el número de términos de la forma $\theta u_i v_j$ que se han de incluir en el modelo es interesante e ilustrativa, en cierta forma es una generalización del criterio usado en el método de componentes principales, y es explicada por Mandel de la siguiente manera:

"Supóngase que se calculan todos los términos del tipo $\theta u_i v_j$ de modo que las e_{ij} son hechas igual a cero".

Entonces se obtiene lo siguiente:

$$(14) \sum_i \sum_j d_{ij}^2 = \hat{\theta}^2 + \hat{\theta}'^2 + \hat{\theta}''^2 + \dots$$

Resultado obtenido a partir de la condición de ortogonalidad de los eigenvectores y de la relación (3).

"La ecuación (14) corresponde a una partición de la suma de cuadrados de la interacción, por lo que uno se ve tentado a realizar un tratamiento del tipo de análisis de varianza. Sin embargo, las θ^2 no son formas cuadráticas de las observaciones originales y_{ij} ."

A pesar de lo cual es posible formular el problema en un lenguaje de análisis de varianza en virtud de las siguientes consideraciones".

Suponga en primer lugar que las y_{ij} forman una muestra aleatoria de una población $N(0,1)$. Entonces las cantidades $\theta^2, \theta'^2, \theta''^2, \dots$, formarán cada una poblaciones estadísticas particulares.

Sean

$$M_1 = E_N(\theta^2), M_2 = E_N(\theta'^2), \dots$$

"Donde E es el operador esperanza y N indica una distribución normal estandar ($N(0,1)$)".

Si la población normal tuviera una varianza igual a σ^2 en vez de igual a 1, entonces las cantidades $M_1, M_2, \dots$, estarían simplemente multiplicadas por σ^2 .

Si lo que tenemos dado ahora es una matriz de observaciones

$$\bar{Y} = (y_{ij})$$

donde los términos de interacción no son sino errores normales -- (con distribución de probabilidad normal), entonces los cocientes de los valores θ^2 obtenidos a partir de estos datos, entre los correspondientes valores de M , obtenidos de una matriz de observaciones $N(0,1)$ de las mismas dimensiones, son simplemente estimadores de σ^2 .

Continuando un poco más con este argumento, y haciendo un razonamiento de tipo heurístico, podemos esperar que si el modelo real contiene p términos del tipo $\theta_{ij} y_j$, entonces los correspondientes p valores de θ^2 serán inflados por los efectos sistemáticos de estos términos mientras que los términos restantes $(\theta^{(p)})^2, (\theta^{(p+1)})^2, \dots$, sólo serán estimadores de:

$$M_{p+1}\sigma^2, M_{p+2}\sigma^2, M_{p+3}\sigma^2, \dots$$

y por lo tanto

$$\frac{(\theta^{(p1)})^2}{M_{p+1}}, \frac{(\theta^{(p11)})^2}{M_{p+2}}, \frac{(\theta^{(p12)})^2}{M_{p+3}}, \dots$$

serán estimadores de σ^2 .

Es de esta forma que uno puede realizar un juicio sobre el número de términos $\theta u_i v_j$ que han de considerarse en el modelo.

Realizando un estudio Monte-Carlo se determinaron los valores de $M_1, M_2, M_3, \dots$ para distintos valores de I y J . Valiéndose de estos valores se planteó una analogía de las tablas de análisis de varianza, la cual se presenta a continuación:

Tabla de Análisis de Varianza generada por el Modelo de Mandel.

F.V.	G.L.	S. de C.	C.M.
Total	IJ	$\sum \sum y_{ij}^2$	
τ_i	$I-1$	$\sum y_i^2/I - Y^2/IJ$	
γ_j	$J-1$	$\sum y_j^2/I - Y^2/IJ$	
η_{ij}	$(I-1)(J-1)$	$\sum \sum d_{ij}^2$	$\sum \sum d_{ij}^2 / (I-1)(J-1)$
$\theta u_i v_j$	M_1	$\hat{\sigma}^2$	$\hat{\sigma}^2 / M_1$
$\theta' u_i v_j'$	M_2	$\hat{\sigma}'^2$	$\hat{\sigma}'^2 / M_2$
$\theta'' u_i v_j''$	M_3	$\hat{\sigma}''^2$	$\hat{\sigma}''^2 / M_3$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

"Ahora se examinan los cuadrados medios correspondientes a la partición de η_{ij} en la suma de términos $\theta u_i v_j$. En este momento no se tiene una distribución teórica para estos cuadrados medios, y las pruebas exactas de significancia no se pueden realizar. Sin embargo, aún la interpretación intuitiva de los cuadrados medios nos llevará a determinar fácilmente el número de términos multiplicativos que han de ser considerados en el modelo.

Así, si ninguno de los cuadrados medios relacionados con términos -- $\theta_{u_i v_j}$ resulta ser drásticamente mayor que los demás, por lo expuesto antes, ningún término multiplicativo será considerado en el modelo que se ha de ajustar a los datos que se estudien. Es decir, el modelo es un modelo estrictamente aditivo:

$$y_{ij} = \mu + \tau_i + \beta_j + e_{ij}$$

Esta puede ser una alternativa a la prueba de aditividad de Tukey; -- sin embargo como se notó anteriormente, se desconoce la distribución de θ^2 y en consecuencia el nivel de significancia real de la prueba.

Si acaso resulta ser, que solamente el primer cuadrado medio, es decir el correspondiente a $\theta_{u_i v_j}$ es significativamente mayor que los correspondientes a los términos semejantes posteriores, el modelo en este caso sería [2]

$$y_{ij} = \mu + \tau_i + \beta_j + \theta_{u_i v_j} + e_{ij}$$

el cual incluye como caso particular la suposición de Tukey para su prueba de no aditividad

$$y_{ij} = \mu + \tau_i + \beta_j + \lambda \tau_i \beta_j + e_{ij}$$

Así podría continuarse usando los valores de la tabla de análisis de varianza".

V.- Enfoque de Méndez y Bustos.

En términos generales se puede decir que este enfoque, es el mismo -- que emplea Mandel, con la diferencia de que el criterio de determinación del número de términos $\theta_{u_i v_j}$ en el modelo, está sujeto a las -- propiedades de los residuos.

Por lo tanto se explicará únicamente los fundamentos en que se basa el criterio de elección para su aplicación.

El desarrollo del criterio anterior consiste en el ajuste sucesivo -- de modelos como el siguiente:

$$(1) \quad y_{ij} = \mu + \tau_i + \beta_j + \theta_{u_i v_j} + \theta' u_i' v_j' + \theta'' u_i'' v_j'' + \dots + e_{ij}$$

$$y_{ij} = \mu + \tau_i + \eta_j + \sum_{k=0}^p \theta^{(k)} u_i^{(k)} v_j^{(k)} + e_{ij} \quad (9)$$

con $p = 0, 1, 2, \dots$ en el que cada vez se hacen pruebas para la normalidad, independencia y homocedasticidad.

"El criterio en sí para la determinación de p (número de términos del tipo $\theta u_i v_j$), tiene una generación básicamente heurística.

Esto es, la consideración de que para datos reales, las siguientes premisas se cumplen:

- 1.- Los datos reales generados por un proceso físico, químico o biológico reciben la influencia de un número muy grande de factores.
- 2.- El modelo estadístico lineal incluye los factores más importantes.
- 3.- El término de error, además de los errores de medida se genera por un gran número de efectos de poca importancia cada uno de ellos, y
- 4.- Los efectos que generan el error, por ser de poca importancia actúan en forma aditiva".

De lo anterior se concluye, que "los errores deben tener distribución normal, con base en el teorema central del límite. Además -- tendrán independencia y homogeneidad de varianzas".

Por todo lo anterior, se puede esperar que en un modelo como (1), agregando términos de interacción (del tipo ya descrito), los residuos alcancen las propiedades deseadas en un momento dado, esto será una clara indicación de que han quedado incluidos en el modelo en forma explícita los efectos de más importancia, por un lado y por otro, que la "explicación" de la interacción de los efectos a logrado su mejor forma en los datos en cuestión.

Es de hacerse notar, que para datos generados artificialmente, -- los supuestos mencionados con anterioridad no se cumplen, por lo que se espera que este criterio para determinar el número de términos $(\theta u_i v_j)$ en (1) no funcione satisfactoriamente.

Referencias.

- 1.- Bustos Alfredo y Méndez R. Ignacio, Especificación de la interacción en tablas de doble entrada, Facultad de Ciencias y -- CIMAS, UNAM.
- 2.- Graybill Franklin A., An Introduction to Linear Statistical Models, Volume I, New York: McGraw-Hill Book Co. Inc., 1961 --- (First Edition).
- 3.- Johnson, N.L. "A method for separation of residual and interaction effects in cross-classifications" Institute of Statistics Mimeo Series N° 842 Department of Statistics, University of North Carolina at Chapel Hill 1972.
- 4.- Johnson, N.L. and Leone, F.C. (1968). "Statistics and Experimental Design in Engineering and the Physical Sciences". Volume II., New York: John Wiley and Sons.
- 5.- Lilliefors, H.W. (1967). "On the Kolmogorov-Smirnov Test -- for Normality with Mean and Variance Unkown", Journ. Am. Stat. Assoc. 62.
- 6.- Mandel, J. A method for fitting empirical surfaces to physical or chemical data. Technometrics 11,411-429 (1969).
- 7.- Mandel, J. (1969) The partitioning of interaction in analysis of variance. Journal of Research of the Natl. Bureau of --- Standards, Sec. B., N° 4 Vol. 73.
- 8.- Mandel, J., (1971), A new analysis of variance model for -- Non-additive data. Technometrics, Natl. Bureau of Standars, Vol. 13, N° 1.
- 9.- Mandel, J., Structure analysis in two way tables of measurement data., National Bureau of Standards, Washington, D.C., U.S.A. (39th Session of the International Statistical Institute, -- Vienna, Austria, August. 20-30, 1973) Vol. II.
- 10.- Méndez R. Ignacio, Introducción a la metodologfa estadfstica, Universidad Autónoma de Chapingo; Chapingo, México. (1976).
- 11.- Méndez R. Ignacio y Márquez Landa Julio A., Las hipótesis que se prueban en modelos lineales con dos criterios de clasificación. I.I.M.A.S. (U.N.A.M.) Comunicaciones Técnicas (1976).

12.- Milliken George A. and Graybill, F.A. (1970) Extensions of the general linear hypothesis model., Jour. Am. Stat. Assoc. 65, 797-807, Vol. 65.

13.- Rao C. Radhakrishna, Linear statistical inference and its applications; New York: John Wiley & Sons, Inc. 1973 (Second Edition).

14.- Scheffé Henry, The analysis of variance, New York: John -- Wiley & Sons, Inc. (1959).

15.- Scheffé Henry, Testing linear hypothesis against a certain class of nonlinear alternatives: A generalization of Tukey's -- " 1d.f. for non-additivity.

16.- Searle S.R., Linear Models., New York: John Wiley & Sons, Inc. 1971 (First Edition).

17.- Simon Gary, Fitting models to $R \times C$ tables, Technical Report 34, Series 2, Department of Statistics Princeton University (1973).

18.- Snedecor George W. and Cochran William G., Statistical methods., Ames Iowa, U.S.A. (Sixth Edition).

19.- Tukey, J.W. One degree of freedom for non-additivity, Biometrics 5, 232-242 (1949).

20.- Tukey, J.W. The future of data analysis, Anns. Math. Stat. 33, 1-67 (1962).

21.- Yamane Taro, Statistics, an introductory analysis, (1967) - New York. Harper and Row.

CAPITULO IV

COMPARACION MEDIANTE SIMULACION, DE VARIOS ENFOQUES PARA EL ANALISIS DE TABLAS DE DOBLE ENTRADA.

I.- Introducción.- La comparación de los enfoques o métodos planteados en el capítulo anterior, implican la realización de una simulación que permita, más o menos, darse una "buena" idea de los pros y contras que el empleo de los mismos puede acarrear.

En este capítulo se presentan, los detalles más importantes de la simulación que se efectuó, en la computadora Burroughs B6700 del I.I.M.A.S. (U.N.A.M.), mediante un programa en FORTRAN (Programa que se incluye en el apéndice A) y el uso de unas rutinas del paquete I.M.S.L. (International Mathematical and Statistical Libraries).

II.- Descripción de las consideraciones más importantes de los enfoques a comparar.

a) Enfoque de Tukey.

La aplicación de este método, de acuerdo a lo estipulado en el capítulo anterior fue directa, la $H_0: \lambda=0$ se probó siempre a tres niveles de significancia (0.05, 0.025 y 0.01).

b) Enfoque de Graybill y Milliken.

En este caso, considerando lo general del método, se realizaron los siguientes dos supuestos;

1) La matriz $\tilde{F}$ de dimensiones $n \times l$, se escogió con $l=1$ siempre, por lo que las interacciones estuvieron dadas en términos de un vector.

De lo anterior, se concluye también, que el vector $\tilde{\alpha}$ se redujo a un elemento nada más.

2) Se consideraron cuatro formas funcionales conocidas, como son:

$$i) f(\tau_i, \mu_j) = \tau_i * \mu_j$$

$$ii) f(\tau_i, \mu_j) = \tau_i^3 * \mu_j$$

$$iii) f(\tau_i, \mu_j) = (e^{\tau_i}) * \mu_j$$

$$iv) f(\tau_i, \theta_j) = \begin{cases} (ln |\tau_i|) * \theta_j & \text{para } \tau_i \neq 0 \\ -10 ** (60 * RI) & \text{para } \tau_i = 0 \end{cases}$$

con $RI \in [75, 1)$

Al igual que en a), la $H_0: \alpha = 0$ correspondiente se hizo para tres niveles de significancia.

c) Enfoque de Mandel.

En este caso, la aplicación del método fue directa, con la única variación de que el criterio muy "subjetivo" de determinación del número de términos $\theta_{u,v}$, fué substituido por el criterio que considera, que un término del tipo $\theta_{u,v}$ debe estar incluido dentro del modelo, si su cuadrado medio correspondiente es mayor que la mitad del cuadrado medio correspondiente al término anterior.

En particular, en la tabla de análisis de varianza generada por el modelo Mandel, del capítulo anterior, tenemos que si σ^2/μ_1 es mayor (menor o igual) que la mitad del cuadrado medio correspondiente a la fuente de variación η_{ij} , entonces el término $\theta_{u,v}$ (no) debe aparecer en la descripción del modelo en cuestión, etc.

d) Enfoque de Méndez y Bustos.

En este caso y como ya se dijo anteriormente, se realizaron --- pruebas de hipótesis estadísticas, sobre los residuos correspondientes a cada uno de los ajustes sucesivos.

Las pruebas que se realizaron fueron:

1.- Normalidad.- Se aplicó la prueba de Kolmogorov-Smirnov ---- empleando las tablas de Lilliefors (1967). Los niveles que se probaron fueron 0.2, 0.15, 0.10, 0.05 y 0.01.

2.- Simetría.- El parámetro $\xi_1 = E(x-\mu)^3/\sigma^3$ se estima por $m_3/m_2^{3/2} = g_1$, donde $m_k = (\sum (x_i - \bar{x})^k)/n$, se obtuvo g_1 , y se probó la hipótesis $\xi_1 = 0$ al 0.01 y 0.05 usando las tablas de Snedecor y Cochran (1967).

3.- Kurtosis.- El parámetro $\xi_2 = E(x-\mu)^4/\sigma^4$ que es igual a 3 en el caso normal, se estimó con $g_2 = m_4/m_2^2$. La hipótesis $\xi_2 = 3$, se probó usando las tablas de Snedecor y Cochran (1967).

4.- Autocorrelación.- Se empleó la prueba de Von Newman con $\omega = S^2 / S^2$ donde $S^2 = (\sum_{t=2}^n (e_t - e_{t-1})^2) / (n-1)$ y $S^2 = (\sum_{t=1}^n (e_t - \bar{e})^2) / n$.

Se probó la hipótesis de independencia usando las tablas de -- Yamane (1967).

5.- Homocedasticidad.- Se usó la prueba de Cochran a hileras y columnas de los residuos. Las tablas empleadas fueron las de -- Johnson y Leone (1968).

Aquí cabe hacer notar, que la aplicación exacta, que se propone en la referencia 1, resultaba en un número de veces bastante -- considerable con la aceptación de la hipótesis $H_0: \rho=0$ o sea -- que no se rechazaba la hipótesis de que el modelo era aditivo, debido principalmente a que por lo menos algunas pruebas consideraban dos niveles de significancia para aceptárcele como verdadera, se optó por tomar el nivel más pequeño en consideración, lo cual en particular "relajo" el criterio inicial de prueba sin duda alguna, pero sin embargo se pudo comprobar que el número -- de veces en que se llegaría a la conclusión de que el modelo -- era no aditivo, cuando en realidad los datos creados presentaban una clara interacción aumento, por lo que se esperó que funcionara, mejor, aún para aceptar la hipótesis de que el modelo era aditivo.

III.- Descripción de la simulación.

La obtención de datos u observaciones se realizó, fijando diferentes niveles de los factores A y B, además se generaron números pseudo-aleatorios con distribución $N(0, \sigma^2)$, empleando la -- función GGNOR del paquete IMSL (para $\sigma^2=2$).

Se consideraron además tres tipos de interacción, a la hora de generar los datos, como son:

- a) $\tau_i * \beta_j$
- b) $\tau_i^3 * \beta_j$
- c) $e^{\tau_i * \beta_j}$

Por lo tanto, considerando los siguientes valores $\mu=0.0$, $\tau_i = -6.0$, $\beta_j = -8.0$ y $e_i = -1.56243879$, tenemos que la observación y_{ij} estaría dada para cada tipo de interacción de la siguiente manera:

- a) $Y_{11} = \mu + \tau_1 + \beta_1 + \tau_1 \beta_1 + e_{11} = 0 + (-6) + (-8) + (-6)*(-8) + (-1.56243879) = 32.43756121$
- b) $Y_{11} = \mu + \tau_1 + \beta_1 + \tau_1^3 \beta_1 + e_{11} = 0 + (-6) + (-8) + (-6)^3(-8) + (-1.56243879) = 1712.43756121$
- c) $Y_{11} = \mu + \tau_1 + \beta_1 + (EXP(\tau_1)) * \beta_1 + e_{11} = 0 + (-6) + (-8) + (0.002478752)*(-8) + (-1.56243879) = -15.58226881$

Analogamente para las demas observaciones.

Una vez generadas las observaciones para un tipo de interacción, se aplicaron los cuatro enfoques para obtener las tablas de análisis de varianza correspondientes.

En el enfoque de Tukey y en el de Graybill y Milliken se emplearon valores de F tabulados a tres diferentes niveles de significancia (0.05, 0.025 y 0.01), con sus apropiados grados de libertad.

En el enfoque de Méndez y Bustos y en el de Mandel, nada más se aplicaron los criterios correspondientes tal cual.

Con la finalidad de construir la potencia de las pruebas de hipótesis, se hicieron variar los valores de las τ_i y β_j cuantificando el grado de dispersión de la interacción, por una suma de cuadrados, según se indica.

$$SCAB = \sum (\delta_{ij} - \bar{\delta}_{..})^2$$

Para cada combinación de dispersiones de las τ_i 's y de las β_j 's se generaron cien tablas de observaciones y se analizó cada conjunto por los cuatro métodos descritos.

En esta forma se obtuvieron las potencias empíricas de las pruebas, con el programa que se desarrolló. Los puntos que dieron lugar a las gráficas de las curvas de potencia empírica, se obtuvieron, usando como abscisa, las diferentes sumas de cuadrados de los parámetros al fijarse los diferentes valores de los mismos y como ordenada la potencia empírica de la prueba la cual se obtuvo del cociente $P.E. = R/T$, donde R es el número de veces en que se rechaza

zó la hipótesis de no interacción en cada caso y T es el número total de pruebas que se hicieron ($T=100$).

El factor a estudiar en las gráficas de potencia empírica fué -- uno: la interacción $A \cdot B$; las funciones de interacción consideradas en la generación de datos fueron tres; tres niveles de -- significancia (0.05, 0.025 y 0.01), aunque no para todos los en -- foques; un valor de σ^2 (2). Todo esto hace que se tenga $1 \times 3 \times 3 \times 1 = 9$ gráficas de potencia empírica.

Además, si para cada punto generado en la suma de cuadrados para la interacción, graficamos la media de los cuadrados medios del error, podremos analizar como se comportan en general los -- enfoques para estimar la varianza del error. En este caso, como se tienen tres tipos de interacción, son tres las gráficas que se generan.

IV.- Resultados y Discusión.-

En esta parte, se analizan seis gráficas que se presentan en el apéndice B. Las primeras tres corresponden a los patrones observados en la potencia empírica para cada tipo de interacción y -- al nivel de significancia del 0.05, ya que para 0.025 y 0.01 los resultados fueron muy semejantes, las tres restantes corresponden al comportamiento estimativo de la varianza del error según sea la interacción. El análisis se efectuara por interacción -- considerada.

a) Interacción ($A \cdot B$).-

En términos generales se observa que la gráfica de potencia -- empírica (P.E.), cuando la suma de cuadrados para la interacción (SCAB) es cero empieza en el nivel correspondiente de significancia, con excepción de los enfoques de Mandel y de Méndez y -- Bustos, los cuales para este punto presentan una potencia muy -- grande, lo cual implica que están cometiendo error tipo I con -- probabilidades (0.324 para Mandel y 0.635 para el de Méndez y Bustos) mucho mayores que el nivel de significancia deseado. Aquí se observa también que el enfoque de Tukey o su equivalente el de Graybill y Milliken opción uno presentan las características más apropiadas de un método estadístico "bueno", puesto que, -- conforme la SCAB va creciendo la P.E. aumenta monótonamente llegando a 1 en el punto igual a 189.84 manteniéndose así posteriormente.

También se observa que las opciones dos y tres del enfoque de Graybill y Milliken se comportan más o menos como la opción uno del mismo enfoque o sea como el enfoque de Tukey e incluso estas dos opciones presentan un comportamiento alternante, pues algunas veces es mejor la opción dos y otras veces la opción tres.

Lo que sí, verdaderamente es grave, es la opción cuatro de Graybill y Milliken puesto que aún cuando en 0 (para SCAB) comienza aproximadamente bien, conforme la SCAB va creciendo, hasta antes de 150, la P.E. crece muy poco, tan poco que aún en el nivel más alto de significancia, nunca alcanza el valor de .2 y si en cambio a partir de 150 en adelante va decreciendo hasta que la potencia es cero, cosa que significa que se está cometiendo el error tipo II con probabilidad 1 o sea, se están aceptando hipótesis falsas, en este caso, se está aceptando la hipótesis de aditividad, cuando en general existe una gran interacción en la información.

Por otra parte, la gráfica de medias de los cuadrados medios del error (CME) con respecto a SCAB, nos muestra que las curvas correspondientes para los enfoques de Mandel, Méndez y Bustos y de Graybill y Milliken opción cuatro, tienden a crecer desmesuradamente conforme la SCAB aumenta, produciendo sobreestimaciones muy grandes de la varianza del error, no así los restantes enfoques, al menos cuando la SCAB es menor que 700, porque posteriormente siguen la misma tendencia, aunque no tan marcadamente.

b) Interacción ($A^3 * B$).-

Aquí, igual que en el caso anterior, observamos que los enfoques de Mandel y de Méndez y Bustos cuando la SCAB es 0, tienen una potencia muy grande (error tipo I con alta probabilidad).

También se puede observar, que a partir de que la SCAB es igual a 7.6171 la potencia de preferencia sería la del enfoque de Mandel y/o la del enfoque de Méndez y Bustos, ya que crecen más rápidamente, conforme la SCAB aumenta y sobre todo, porque permanece constantemente en 1 una vez que alcanzado este valor, no así los restantes métodos que en particular observan dos bajas de potencia con su correspondiente regreso a 1, y una de esas bajas va más allá del .5, cosa muy delicada. Puesto que, de acuerdo a la teoría, SCAB es parte del parámetro de no centralidad de la F para probar interacción, esperando por tal hecho que si aumenta la SCAB, aumenta la potencia; es decir la potencia es función monótona creciente de SCAB. Es raro el hecho de que la P.E. no resulte monótona creciente.

El enfoque de Graybill y Milliken opción cuatro continúa con las características del caso anterior, nadamás que más acentuadas.

Por otro lado, la gráfica de medias de los CME con respecto a SCAB muestra una tendencia tremendamente creciente para todos los enfoques, siendo la más "suave" la que corresponde al enfoque de Graybill y Milliken opción dos, que en particular concuerda con el tipo de interacción empleado para generar los datos, en este caso.

c) Interacción (EXP(A)*B)

En este caso, observamos que en 0 (para SCAB) casi todos los enfoques excepto los mismos que en los otros casos, observan valores - más o menos apropiados al nivel de significancia correspondiente, pero conforme el valor de SCAB aumenta, todos los métodos entran - en una situación oscilante hasta aproximadamente 6201.63 punto a - partir del cual, se sostiene la potencia empírica de 1, para casi todos los métodos menos la opción cuatro del enfoque de Graybill y Milliken.

Lo expuesto anteriormente, es muy grave en general para todos los métodos, por lo discutido en el caso (b).

También se puede notar que la opción tres del enfoque de Graybill y Milliken es mejor que la opción uno del mismo enfoque, debiendo se esto a que la clase de datos generados en este caso y la caracterización de la interacción en este enfoque coinciden.

Finalmente, analizando la gráfica de medias de los CME contra la - SCAB, observamos que en términos generales, el comportamiento que siguen las "curvas" en este caso, es aproximadamente el mismo que presentan las "curvas" de potencia ya analizadas para el caso en - cuestión, notando únicamente que la curva que en este caso corresponde al enfoque de Graybill y Milliken opción tres es más "suave" que las demás, pero con la misma tendencia a la larga, de producir sobreestimaciones muy grandes cuando la interacción es grande.

V.- Conclusiones y Recomendaciones.-

De todo lo visto anteriormente se puede concluir lo siguiente:

I.- Que el enfoque de Méndez y Bustos es un enfoque demasiado "ri- guroso" o "exigente".

Además que no detecta muy bien la aditividad o sea que tiene una - potencia muy grande cuando la interacción es nula, resultados todos

estos, los cuales más o menos se esperaban, debido a las consideraciones que se hicieron en sí del enfoque.

2.- Que los altibajos que se observan en las curvas de potencia -- para la interacción $(EXP(A)*B)$ son bastantes irregulares, puesto -- que conforme la interacción aumenta, debería esperarse que la potencia de los métodos fuera 1 y no descendiera a niveles muy bajos, ya que eso está indicando que no existe gran interacción en los datos, lo cual definitivamente no es cierto.

Con respecto a este punto es de hacerse notar que se revisó una -- vez más el programa hecho y no se encontró falla alguna, por lo -- que en este caso no se sabe exactamente que suceda.

3.- Que las opciones del enfoque Graybill y Milliken que corresponden a el tipo de interacción, considerado en cada caso para la generación de datos, son las que mejor estiman la varianza del error supuesto, aún cuando se observe que conforme la SCAB aumenta, la -- estimación de la varianza aumenta también. Además es importante notar que la opción dos de Graybill y Milliken en su caso, es la que más mal estima la varianza del error en comparación con las otras opciones del enfoque Graybill y Milliken en sus casos correspon---dientes.

4.- Que los enfoques de Tukey y la opción uno del de Graybill y Milliken (que son idénticos) son los que mejor tratan la información, para cualquiera de los tres tipos de interacción considerados.

5.- Que la consideración del enfoque de Graybill y Milliken opción cuatro, resulta pésima para cualquier tipo de interacción que se -- considere.

De todo lo anterior se recomienda:

1.- Realizar estudios comparativos en datos reales para poder analizar más a fondo el enfoque de Méndez y Bustos.

2.- Suponer otros valores de varianza para los errores, a la hora de generar datos.

3.- Modificar el criterio de determinación del número de términos del tipo $\theta_{u,v}$ en el enfoque de Mandel, con respecto al empleado en este trabajo. O sea, incluir un término en el modelo, solo si -- su cuadrado medio correspondiente es mayor que la parte proporcional que se escoja, del cuadrado medio correspondiente al término --

anterior. E incluso realizar un estudio, respecto a este punto -- para un cierto conjunto de partes proporcionales.

4.- Eliminar de futuros estudios, la consideración del enfoque de Graybill y Milliken opción cuatro, pues según parece, el tipo de función que es en si el logaritmo natural es lo que provoca que el enfoque no sea "bueno".

5.- Considerar algunos otros tipos de interacción tanto en la generación de datos como en la aplicación del enfoque de Graybill y Milliken.

Referencias.

- 1.- Bustos Alfredo y Méndez R. Ignacio, Especificación de la interacción en tablas de doble entrada, Facultad de Ciencias y CIMAS, UNAM.
- 2.- Graybill Franklin A., An Introduction to Linear Statistical Models, Volume I, New York: McGraw-Hill Book Co. Inc., 1961 (First Edition).
- 3.- Johnson, N.L. "A method for separation of residual and interaction effects in cross-classifications" Institute of Statistics Mimeo Series N° 842 Department of Statistics, University of North Carolina at Chapel Hill 1972.
- 4.- Johnson, N.L. and Leone, F.C. (1968). "Statistics and Experimental Design in Engineering and the Physical Sciences". Volume - II., New York: John Wiley and Sons.
- 5.- Lilliefors, H.W. (1967). "On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown", Journ. Am. Stat. Assoc. 62.
- 6.- Mandel, J. A method for fitting empirical surfaces to physical or chemical data. Technometrics 11,411-429 (1969).
- 7.- Mandel, J. (1969) The partitioning of interaction in analysis of variance. Journal of Research of the Natl. Bureau of Standards, Sec. B., N° 4 Vol. 73.
- 8.- Mandel, J., (1971), A new analysis of variance model for Non-additive data. Technometrics, Natl. Bureau of Standards, Vol. 13, N° 1.

- 9.- Mandel, J., Structure analysis in two way tables of measurement data., National Bureau of Standards, Washington, D.C. U.S.A. (39th Session of the International Statistical Institute, Vienna, Austria, August. 20-30, 1973) Vol. II.
- 10.- Méndez R. Ignacio, Introducción a la metodología estadística, Universidad Autónoma de Chapingo; Chapingo, México. (1976).
- 11.- Méndez R. Ignacio y Márquez Landa Julio A., Las hipótesis - que se prueban en modelos lineales con dos criterios de clasificación. I.I.M.A.S. (U.N.A.M.) Comunicaciones Técnicas (1976).
- 12.- Milliken George A. and Graybill, F.A. (1970) Extensions of the general linear hypothesis model., Jour. Am. Stat. Assoc. 65. 797-807, Vol. 65.
- 13.- Rao C. Radhakrishna, Linear statistical inference and its - applications; New York: John Wiley & Sons, Inc. 1973 (Second Edition).
- 14.- Scheffé Henry, The analysis of variance, New York: John Wiley & Sons, Inc. (1959).
- 15.- Scheffé Henry, Testing linear hypothesis against a certain class of nonlinear alternatives: A generalization of Tukey's - " Id.f. for non-additivity.
- 16.- Searle S.R., Linear Models., New York: John Wiley & Sons, - Inc. 1971 (First Edition).
- 17.- Simon Gary, Fitting models to $R \times C$ tables, Technical Report 34, Series 2, Department of Statistics Princeton University (1973).
- 18.- Snedecor George W. and Cochran William G., Statistical methods., Ames Iowa, U.S.A. (Sixth Edition).
- 19.- Tukey, J.W. One degree of freedom for non-additivity, Biometrics 5, 232-242 (1949).
- 20.- Tukey, J.W. The future of data analysis, Anns. Math. Stat., 33, 1-67 (1962).
- 21.- Yamane Taro, Statistics, an introductory analysis, (1967) - New York. Harper and Row.

NOTAS ENCUADRADAS.

Capítulo I

1.- Por un modelo lineal se entenderá una ecuación o igualdad matemática que comprenda variables matemáticas, variables aleatorias y parámetros siendo lineal en los elementos de los dos últimos tipos. Con la posibilidad de que las variables aleatorias sean definidas como pertenecientes a una cierta función de distribución.

2.- En esta sección se empleará la notación $\sim$ para denotar a los estimadores de máxima verosimilitud y $\wedge$ para los estimadores insesgados.

3.- Ω es el espacio $q+1$ dimensional, definido por las desigualdades $0 < \sigma^2 < \infty$, $-\infty < \beta_i < \infty$ ($i = 1, 2, 3, \dots, q$) y donde $\hat{\Omega}$ es el punto de dicho espacio donde la función de verosimilitud se maximiza.

ω es el espacio definido por la desigualdad $0 < \sigma^2 < \infty$ y por $\beta_1 = \beta_1^*, \beta_2 = \beta_2^*, \beta_3 = \beta_3^*, \dots, \beta_q = \beta_q^*$, en este caso el espacio es de una dimensión y $\hat{\omega}$ es el punto de dicho espacio donde la función de verosimilitud se maximiza.

4.- Forma en que denotaremos de ahora en adelante a las variables que se consideren en un modelo de rango incompleto, cuya característica principal suele ser la no "medibilidad".

5.- En este caso el valor de λ se halla determinado por el siguiente teorema;

Teorema.- Bajo las condiciones del teorema 9 el parámetro de no-centralidad es:

$$\lambda = \frac{E(\hat{\beta}' \bar{X}' \bar{Y} - \hat{\beta}' \bar{X}_2' \bar{Y})}{2 \sigma^2} - \frac{S}{2}$$

Capítulo II

1.- Recuerdese la nota número 4 del capítulo anterior.

2.- En este trabajo únicamente consideramos los modelos, que en todas y cada una de las celdas (vease la nota número 3) tienen el mismo número de observaciones, en particular M (con $M \geq 1$), [Balanceados].

- 3.- Se llamará celda a la "intersección" muestreable, del i -ésimo nivel del factor τ , con el j -ésimo nivel del factor μ .
- 4.- En el modelo en cuestión, una combinación lineal de parámetros $\sum a_i \mu_i$ es llamada contraste si se tiene que $\sum a_i = 0$.
- 5.- Cantidades que también se suelen denotar como:

$$R(\tau/\mu, \mu) [\text{o } R(\mu/\mu, \tau)]$$
- 6.- Los resultados obtenidos en este punto se podrán lograr, si se eliminan las columnas (de la matriz $\bar{X}$) asociadas con el factor que se supone es constante y además del vector de parámetros, se eliminan los elementos asociados con dicho factor.
- 7.- Recuerdese la nota número 6.
- 8.- También se pudo considerar $R(\mu)$.
- 9.- O sea, siempre se aceptaría la hipótesis correspondiente al modelo en cuestión.

Capítulo III

- 1.- Este método tiene la particularidad de que puede aplicarse -- tanto a modelos de diseños de experimentos, como de covarianza y en ciertos modelos de regresión.
- 2.- Nótese que si la hipótesis $H_0: \alpha = \bar{0}$ no se rechaza, el modelo (1) llega a ser el planteado en el caso de diseños de experimentos o sea $Y = X\beta + \bar{e}$.
- 3.- Ver la referencia 14 problema 4.19.
- 4.- Esta suposición también se efectúa en el modelo de Tukey para $\{\tau_i\}$ y $\{\mu_j\}$, además nótese que en el caso de Tukey la f_{ij} $(\bar{X}\bar{\beta}) = \bar{X}_i \cdot \bar{\beta}_j$, siendo un vector de IJ elementos.
- 5.- Nótese que si la hipótesis H_0 no es cierta, v se distribuye como una F no central.
- 6.- Considere las siguientes relaciones:
 $\theta = \theta^{(0)}, \theta' = \theta^{(1)}, \theta'' = \theta^{(2)}, \dots$ para comprender mejor la notación.

7.- Nótese que la determinación de un término multiplicativo en el modelo se hace de manera subjetiva.

8.- Adviertase que podemos denotar, a $\bar{E}X'$ como $\bar{X}^-$, por lo que el producto que tenemos, que es $\bar{X}(\bar{X}\bar{X})^-X'$ con $\bar{E} = (\bar{X}\bar{X})^-$ se puede -- simplificar a $\bar{X}\bar{X}^- = \bar{X}\bar{E}X' = X(\bar{X}'\bar{X})^-X'$.

9.- Vea la nota **Nº 6** y considere además las siguientes relaciones:

$$u_i = u_i^{(0)}, u_i' = u_i^{(1)}, u_i'' = u_i^{(2)}, \dots, \dots, v_j = v_j^{(0)}, v_j' = v_j^{(1)}, v_j'' = v_j^{(2)}, \dots.$$

Capítulo IV

1.- Los valores empleados para los factores A y B fueron los que se dan a continuación.

Factor A:

- i) 0.0, 0.0, 0.0, 0.0, 0.0;
- ii) -0.125, -0.0625, 0.0, 0.0625, 0.125;
- iii) -0.5, -0.25, 0.0, 0.25, 0.5;
- iv) -1.5, -0.75, 0.0, 0.75, 1.5;
- v) -4.0, -2.0, 0.0, 2.0, 4.0.

Factor B:

- i) 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0;
- ii) -0.25, -0.1875, -0.125, -0.0625, 0.0625, 0.125, 0.1875, 0.25;
- iii) -1.0, -0.75, -0.5, -0.25, 0.25, 0.5, 0.75, 1.0;
- iv) -3.0, -2.25, -1.5, -0.75, 0.75, 1.5, 2.25, 3.0;
- v) -8.0, -6.0, -4.0, -2.0, 2.0, 4.0, 6.0, 8.0.

El valor de la media fue 0.0.

APENDICE A

```

BEGIN
:WAIT('PONTAP INSL',OK);
:COMPILE 'FTEST FORTAL';
:DATA
$SET AUTORIND
$RIND = FROM SERVICE/INSL/2=1
$RINDR SET LIST TIME
FILE 6=NONREARCH,UNIT=PRINTER
FILE 1=INTERAC1,UNIT=DISK,RECORD=5
FILE 2=INTERAC2,UNIT=DISK,RECORD=5
FILE 3=INTERAC3,UNIT=DISK,RECORD=5
FILE 4=TARLASMEN,UNIT=DISK,RECORD=2
FILE 7=TEMPINT5,UNIT=DISK,RECORD=5
FILE 8=TEMPINT2,UNIT=DISK,RECORD=5
FILE 9=TEMPINT1,UNIT=DISK,RECORD=5
DIMENSION YESIJ(5,8),FACRES(5),FACRES(8),RESI(5,8),FACREN(5),
FACCOL(8),FACRPI(5),VEC(75),EIVA(5),EIVE(5,5)
IMTGFEE SEED /258439173/
COMMON /NUMBER/ M,N,MINI,MXN,SQRMXN,DESEST,ARR(2,5),GL(UJV(4)
COMMON /CINCO/ IL,CMEFAN(100),CMEFUX(100),CMEGM(400),CMEHBU(100),
ICME(100)
CALL DAVALS
DO 4 I=1,5
CALL DAALFA(FACREN,I)
DO 3 K=1,3
CALL CRUCE(FACREN,FACRPI,K)
CALL CUADRI(SUM1,FACRPI)
DO 2 J=1,5
CALL DARETA(FACCOL,J)
CALL CUADRI2(SUM2,FACCOL)
CALL FACIN2(SUM1,SUM2,K,I,J,VEC)
CALL CEFNES(100)
DO 1 L=1,100
IL=I
CALL GENERA(YFSTJ,FACREN,FACCOL,FACRPI,SEED)
CALL MANDEL(YFSTJ,FACRES,FACRES,RESI,EIVA,EIVE)
CALL MENBUS(RESI,EIVA,EIVE)
CALL TUKEY(YFSTJ,FACRES,FACRES)
CALL GRAMIL(YFSTJ,FACRES,FACRES)
1 CONTINUE
CALL FSCRIB(I,K,J)
CALL CHAMEN
2 CONTINUE
3 CONTINUE
4 CONTINUE
CALL TARLON(I-1,J-1,K-1)
CALL PLOTS(VEC,I-1,J-1,K-1)
END

```

```

SUBROUTINE DAVALS
DIMENSION ERR(2,5)
COMMON /NUMBER/ M,N,MINI,MXN,SQRMXN,DESEST,ARR(2,5),GL(UJV(4)
DATA ERR / .736,20.,.768,15.,.805,10.,.886,5.,.1031,1. /
M=5
N=8
N(N)=(M-1)*(N-1)
MXN=M*N
SQRMXN=SQRT(FLD(4,MXN))
DESEST=SQRT(2.)
DO 1 I=1,M
ARR(1,I)=ERR(1,I)/SQRMXN
ARR(2,I)=ERR(2,I)
1 CONTINUE
GL(UJV(1)=.5367*101

```

KOLMOG

```

GLDIV(2)=.2813*M*MINI
GLDIV(3)=.1341*M*MINI
GLDIV(4)=MINI-GLDIV(1)-GLDIV(2)-GLDIV(3)
RETURN
END

```

```

SUBROUTINE DAALFA(VFCT0),IPART)
  DIMENSION SUSALF(25),VECT01(5)
  COMMON /NUMBER/ M,N,MINI,MXN
  DATA SUSALF / 5*0.,
1  -.125, -.0625, 0., .0625, .125,
3  -.5, -.25, 0., .25, .5,
5  -1.5, -.75, 0., .75, 1.5,
7  -4., -2., 0., 2., 4.,
*/
  K1=M*(IPART-1)
  DO 1 I=1,M
    VECT01(I)=SUSALF(K1+I)
1  CONTINUE
  RETURN
END

```

A 1 SET
A 2 SET
A 4 SET
A 6 SET
A 8 SET
A

```

SUBROUTINE CRUCE(VECT01,VECT11,IJOB)
  DIMENSION VECT01(5),VECT11(5)
  COMMON /NUMBER/ M,N,MINI,MXN
  GO TO(1,3,5,7),IJOB
1  DO 2 I=1,M
    VECT11(I)=VECT01(I)
2  CONTINUE
  RETURN
3  DO 4 I=1,M
    VECT11(I)=VECT01(I)*VECT01(I)*VECT01(I)
4  CONTINUE
  RETURN
5  DO 6 I=1,M
    VECT11(I)=EXP(VECT01(I))
6  CONTINUE
  RETURN
7  DO 8 I=1,M
    IF(VECT01(I).EQ.0.) GO TO 9
    VECT11(I)=ALOG(ABS(VECT01(I)))
    GO TO 8
9  RT=RANDOM(DUMMY)
    IF(RT.GE..75) GO TO 11
    GO TO 9
11 VECT11(I)=-10*(60*RT)
8  CONTINUE
  RETURN
END

```

```

SUBROUTINE CIADRI(SUM1,FACRPI)
  DIMENSION FACRPI(5)
  COMMON /NUMBER/ M,N,MINI,MXN
  SUM1=0.
  DO 1 I=1,M
    SUM1=SUM1+FACRPI(I)*FACRPI(I)
1  CONTINUE
  RETURN
END

```

```

SUBROUTINE DAREIA(VECT0),IPART)
  DIMENSION SUSSET(40),VECT01(4)
  COMMON /NUMBER/ M,N,MINI,MXN
  DATA SUSSET / 4*0.,

```

A 1 SET

1	-.25.	-.1875.	-.125.	-.0625.	.0625.	.125.	.1875.	.25.	H 2 SET
3	-1.	-.75.	-.5.	-.25.	.25.	.5.	.75.	1.	H 4 SET
5	-3.	-2.25.	-1.5.	-.75.	.75.	1.5.	2.25.	3.	H 6 SET
7	-8.	-6.	-4.	-2.	2.	4.	6.	8.	H 8 SET

```

1/
K1=N*(IPART-1)
DO 1 I=1,N
VFCOL(I)=SUSRET(K1+I)
1 CONTINUE
RETURN
END

```

```

SUBROUTINE CUADP2(SUM2,FACCOL)
DIMENSION FACCOL(8)
COMMON /NUMBER/ M,N,MIN1,MXN
SUM2=0.
DO 1 I=1,N
SUM2=SUM2+FACCOL(I)*FACCOL(I)
1 CONTINUE
RETURN
END

```

```

SUBROUTINE FACIN2(SUM1,SUM2,K,I,J,VEC)
DIMENSION VEC(1)
L=25*(K-1)+5*(J-1)+1
VEC(L)=SUM1*SUM2
RETURN
END

```

```

SUBROUTINE CIENES(NTRIAL)
COMMON /UNO/ NHR5(3),NHR2P5(3),NHR1(3)
COMMON /DOS/ NHR6M5(4),NHR6M2(4),NHR6M1(4)
COMMON /TRES/ NHR15,NHR12P,NHR11,NHRM
DO 1 I=1,3
NHR5(I)=NTRIAL
NHR2P5(I)=NTRIAL
NHR1(I) = NTRIAL
NHR6M5(I) = NTRIAL
NHR6M2(I) = NTRIAL
NHR6M1(I) = NTRIAL
1 CONTINUE
NHR15 = NTRIAL
NHR12P = NTRIAL
NHR11 = NTRIAL
NHRM = NTRIAL
NHR6M5(4) = NTRIAL
NHR6M2(4) = NTRIAL
NHR6M1(4) = NTRIAL
RETURN
END

```

1MH23RC

```

SUBROUTINE ESCPIR(I,K,J)
DIMENSION IGRAF(3)
COMMON /UNO/ NHR5(3),NHR2P5(3),NHR1(3)
COMMON /DOS/ NHR6M5(4),NHR6M2(4),NHR6M1(4)
COMMON /TRES/ NHR15,NHR12P,NHR11,NHRM
DATA IGRAF / 7,9,9 /
100 FORMAT(2I2,6I3)
101 FORMAT(12,13,7I3)
102 FORMAT(4I3)
WRITE(6,100) 1,3,1000(2),NHR2P5(2),NHR1(2),NHR5(3),NHR2P5(3),
1 NHR1(3)
IJC=5*(J-1)+1
WRITE(6,101) IGRAF(1),IJC,NHR6M5(1),NHR6M2(1),NHR6M1(1),NHR15

```

1MH23RC

```

WRITE (K,101) IGRA= (2),1,JC,FEUPO (2),FEU3,FEU5 (1),FEU12P
WRITE (K,101) IGRA= (3),1,JC,FEUPO (1),FEU3,FEU5 (1),FEU11
WRITE (K,102) 100-FEU5 (1),FEU-FEU2P (1),FEU-FEU1 (1),100-FEU3
RETURN
END

```

```

SUBROUTINE CUAMEO
DIMENSION X(100)
COMMON /GINCO/ IL,CMEFAN(100),CMEFUK(100),CMEGN(400),CMEPRU(100),
1CME(100)
IG=0
DO 1 I=1,100
X(I)=CMEFAN(I)
1 CONTINUE
IGRA=1
GO TO 10
2 DO 3 I=1,100
X(I)=CMEFUK(I)
3 CONTINUE
IGRA=2
GO TO 10
4 IG=IG+1
IPAR=100*(IG-1)
DO 5 I=1,100
X(I)=CMEGN(IPAR+I)
5 CONTINUE
IGRA=IGRA+1
GO TO 10
6 DO 7 I=1,100
X(I)=CMEPRU(I)
7 CONTINUE
IGRA=IGRA+1
GO TO 10
8 DO 9 I=1,100
X(I)=CME(I)
9 CONTINUE
IGRA=IGRA+1
10 CALL USUMXX(X,100,1,XMIN,XMAX)
CALL MUVAR(X,RMEDIA,VAR)
CALL TITULO(IGRA,IG)
PRINT /," LA MEDIA Y LA VARIANZA DE LOS CME FUE ",RMEDIA,VAR
PRINT /," EL CME MAS GRANDE FUE ",XMAX
PRINT /," EL CME MAS CHICO FUE ",XMIN
GO TO (2,4,4,4,4,5,8,11),IGRA
11 PRINT 200
200 FORMAT(3(/))
RETURN
END

```

```

SUBROUTINE PLOTS(VEC,II,JJ,KK)
DIMENSION VEC(1),NVEC(75),VECP(75)
N1=1+JJ*KK
PRINT /," VECTOR TOTAL DE PUNTOS SIN TRATAR "
DO 4 IN=1,75,25
IN2=IN+24
PRINT /," ",(VEC(IN1),IN1)=(IN,IN2)
4 CONTINUE
DO 1 I=1,N1
VECP(I)=VEC(I)
1 CONTINUE
K=0
2 K1=K+1
K2=25*(K1-1)+1
CALL TRATA(VECP(K2),VEC(I2),VECP(K2),1+JJ)

```

```

      IF (K1 .LT. 3) GO TO 2
3     K=K+1
      CALL ABCD1(K)
      CALL GWA1C(I1,JJ,K,7)
      K2=25*(K-1)+1
      CALL PLOTAB(VECP(K2),MVEC(K2),118,1,7)
      IF (K .EQ. 3) RETURN
      GO TO 3
      END

```

```

SUBROUTINE TRATA(VECP,VEC,MVEC,N1)
  DIMENSION VECP(1),VEC(1),MVEC(1),VECP1(75)
  CALL ORD(VECP,N1)
  CALL CHANGE(VECP,VECP1,N1)
  DO 1 I=1,N1
    MVEC(I)=0
1   CONTINUE
  K=0
  DO 4 I=1,N1
    IF (I .GT. N1) GO TO 4
    K=K+1
    DO 2 IP=1,M1
      IF (VECP1(I) .NE. VEC(IP)) GO TO 2
      MVEC(IP)=K
2   CONTINUE
3   IF (I .EQ. N1) GO TO 4
      IF (VECP1(I) .NE. VECP1(I+1)) GO TO 4
      I=I+1
      GO TO 3
4   CONTINUE
  DO 5 I=1,N1
    VECP(I)=VECP1(I)
5   CONTINUE
  PRINT /,"  VECP  ",(VECP(I),I=1,N1)
  PRINT /,"  VEC   ",(VEC(I),I=1,M1)
  PRINT /,"  MVEC  ",(MVEC(I),I=1,M1)
  RETURN
  END

```

```

SUBROUTINE ORD(R,K)
  DIMENSION R(1)
  IF (R(1) .GE. R(2)) GO TO 3
  CALL CAMBIA(R(1),R(2))
3   CALL CAMBIA(C(1),C(2))
4   MXX1=K-1
  DO 5 I=2,MXX1
    IF (R(I) .GE. R(I+1)) GO TO 5
    CALL CAMBIA(R(I),R(I+1))
    CALL CAMBIA(C(I),C(I+1))
    I=I+1
  DO 4 J=1,I1
    JJ=I+1-J
    IF (R(JJ) .LE. R(JJ-1)) GO TO 5
    CALL CAMBIA(R(JJ),R(JJ-1))
    CALL CAMBIA(C(JJ),C(JJ-1))
5   CONTINUE
  RETURN
  END

```

```

SUBROUTINE CORDA(P1,P2,GG,P2)
  IF (P1 .GT. P2)
    K(P1)=GG*P2
    GG=P2=118

```

```

      RETURN
      END
*****
      SUBROUTINE CHANGE(VEC1,VEC2,I)
      DIMENSION VEC1(1),VEC2(1)
      DO I=1,N
      N=N+1-I
      VEC2(I)=VEC1(I)
1    CONTINUE
      RETURN
      END

```

```

*****
      SUBROUTINE APGHT(K)
      DIMENSION NVEC1(3),NVEC2(3),KOSISTV(7)
100  FORMAT(2I2,6I3)
101  FORMAT(I2,I3,7I3)
200  FORMAT(20X,2I2,6I3)
201  FORMAT(20X,I2,I3,7I3)
      IPEG = 0
      PRINT 202,K
202  FORMAT(5X,'INFORMATION DEL ARCHIVO NOMELOM,I2)
      PRINT *,K
1    READ(K,100,END=3) I,J,NVEC1,NVEC2
      IPEG = IPEG+4
      PRINT 200,I,J,NVEC1,NVEC2
2    READ(K,101) I6PAF,I6PBT,KOSISTV
      KI = KI+1
      PRINT 201,I6PAF,I6PBT,KOSISTV
      IF(KI .LT. 3) GO TO 2
      KI = 0
      GO TO 1
3    PRINT 203,IPEG
203  FORMAT(10X,'NUMBER OF RECORDS',I4)
      RETURN
      END

```

```

*****
      SUBROUTINE GRAFIC(I1,JJ,F,J)
      DIMENSION KOSISTV(7),IPA(10),IPB(10),IPG5A(10),IPG2A(10),IPG1A(10),
      IPG5B(10),IPG2B(10),IPG1B(10),PRGR(10),NVEC1(3),NVEC2(3),
      PDALFA(10),PDBETA(10)
      DATA PDALFA / 0.,
1      .0390625,
3      .625,
5      5.625,
7      40.,
8      500. /
      DATA PDBETA / 0.,
1      .234375,
3      7.75,
5      33.75,
7      240.,
8      500. /
100  FORMAT(2I2,6I3)
101  FORMAT(I2,I3,7I3)
      IJ=I1*JJ
      JK=0
      DO I I=1,10
      IPA(I) = 0
      IPB(I) = 0
      IPG5A(I) = 0
      IPG2A(I) = 0
      IPG1A(I) = 0
      IPG5B(I) = 0
      IPG2B(I) = 0

```

```

      A 1 SET
      A 2 SET
      A 4 SET
      A 6 SET
      A 8 SET
      B 1 SET
      B 2 SET
      B 4 SET
      B 6 SET
      B 8 SET

```

```

      IPA(I) = 0
1    CONTINUE
      REWRITE
2    READ(5,100) (I,J,VECI),I=1,2
3    READ(K,101) IGRAT,IPRAT,COSTIV
      QJ(I,J,101) = IGRAT*IPRAT*COSTIV
      JI=J+1
      IF(J .EQ. 10) GO TO 4
      GO TO 3
4    JI=7
      IPA(I)=IPA(I)+1
      IPGSA(I)=IPGSA(I)+VECI(1)
      IPGZA(I)=IPGZA(I)+VECI(2)
      IPGIA(I)=IPGIA(I)+VECI(3)
      JPH(J)=JPH(J)+1
      IPGSR(J)=IPGSR(J)+VECI(1)
      IPGZR(J)=IPGZR(J)+VECI(2)
      IPGIR(J)=IPGIR(J)+VECI(3)
      JK=JK+1
      IF(JK .LT. 1) GO TO 2
      GO TO 5
      CALL PROPR(PROP,IPGSA,IPA,1)
      CALL GRAF(PDATA,PROF,1)
      CALL PROPR(PROP,IPGZA,IPA,1)
      CALL GRAF(PDATA,PROF,1)
      CALL PROPR(PROP,IPGIA,IPA,1)
      CALL GRAF(PDATA,PROF,1)
      CALL PROPR(PROP,IPGSR,IPH,J)
      CALL GRAF(PDATA,PROF,J)
      CALL PROPR(PROP,IPGZR,IPH,J)
      CALL GRAF(PDATA,PROF,J)
      CALL PROPR(PROP,IPGIR,IPH,J)
      CALL GRAF(PDATA,PROF,J)
5    CONTINUE
      REWRITE
      RETURN
      END
*****
SUBROUTINE PROPR(PROP,IPGFI,IARF2,K)
  DIMENSION PROP(10),IARF1(10),IARF2(10)
  PRINT /," IARF1 ",(IARF1(I),I=1,K)
  PRINT /," IARF2 ",(IARF2(I),I=1,K)
  DO 1 I=1,K
    PROP(I)=I*AT(IARF1(I))/(100.*I*CAF(IARF2(I)))
1  CONTINUE
  PRINT /," PROP ",(PROP(I),I=1,K)
  RETURN
*****
SUBROUTINE GRAF(X,Y,K)
  DIMENSION A(100),IMAGG(SI),X(10),Y(10)
100  FORMAT(A0A1)
200  FORMAT(IH1)
201  FORMAT(20X,A0A1,/)
  READ(5,100) (A(I),I=1,80)
  READ(5,100) (A(I),I=81,160)
C    CALL DSPH(X,Y,N,1,1,N,A,IMAGG,IER)
  PRINT 200
  PRINT 201,A
  PRINT /," VAL OFS OF X ",X
  PRINT /," VAL OFS OF Y ",Y
  RETURN
      END
*****

```

```

SUBROUTINE PLOTAF (NVEC, NVFC, N1, K)
DIMENSION VFCP(1), NVFC(1), NVFCP(75), X(25), Y(25,7), IMA64(515),
1A(160), NSIRVE(7)
100 FORMAT(12,13,7I3)
101 FORMAT(A6A1)
200 FORMAT(1H1)
DO 1 I=1,N1
  NVFCP(I)=0
1 CONTINUE
DO 2 I=1,N1
  NVFCP(NVFC(I))=NVFCP(NVFC(I))+1
2 CONTINUE
PRINT /," NVFCP ",(NVFCP(I),I=1,N1)
3 DO 5 I=1,N1
  DO 4 J=1,7
    Y(I,J)=0.
4 CONTINUE
5 CONTINUE
IJ=0
REWIND K
6 READ(K,100) IGRAF, IPUNTO, NSIRVE
  IJ=IJ+1
  DO 7 I=1,7
    Y(NVFC(IPUNTO),I)=Y(NVFC(IPUNTO),I)+NSIRVE(I)
7 CONTINUE
  IF(IJ.EQ. N1) GO TO 6
  GO TO 6
8 IF(K.EQ. 7) GO TO 10
  K1=1
  X(K1)=VFCP(1)
  DO 9 I=2,N1
    IF(X(K1).EQ. VFCP(I)) GO TO 9
    K1=K1+1
    X(K1)=VFCP(I)
9 CONTINUE
10 DO 12 I=1,K1
  DO 11 J=1,7
    PRINT /," I ",I," J ",J," Y(I,J) ",Y(I,J)," NVFCP(I)
1 " ,NVFCP(I)
    Y(I,J)=Y(I,J)/(100.*NVFCP(I))
11 CONTINUE
12 CONTINUE
  PFA0(5,101) (A(I),I=1,80)
  READ(5,101) (A(I),I=81,160)
  PRINT 201,A
201 FORMAT(20X,A6A1,/)
  K2=K1
C14 CALL DISPLH(X,Y,K2,7,1,K2,A,16A64,IFR)
  K2=K2-1
C IF(K2.GT. 3) GO TO 14
  REWIND K
  PRINT 200
  PRINT /," VALORES DE X ",X
  PRINT /," VALORES DE Y "
  DO 13 I=1,25
    PRINT /," ",(Y(I,J),J=1,7)
13 CONTINUE
  IF(K.EQ. 9) RETURN
  K=K+1
  GO TO 3
END

```

```

*****
SUBROUTINE CONTRA (OUVCN1,OUVCN2,OUVCN3,ERRCM,IP)
COMMON /TRFS/ NHRI5,NHRT2P,NHRI1,ERR0

```

```

COMMON /CUATRO/ SCMPR(
IF (OUVCM1 .GT. (SCMPR(2/2.)) GO TO 4
IP=0
GO TO 3
4 IF (OUVCM2 .GT. (OUVCM1/2.)) GO TO 1
IP=1
GO TO 3
1 IF (OUVCM3 .GT. (OUVCM2/2.)) GO TO 2
IP=2
GO TO 3
2 IF (ERFCM .LE. (OUVCM3/2.)) GO TO 5
C SI 11FGAS AQUÍ EL MODELO ES ADITIVO
IP=4
NHRM=NHRM-1
RETURN
5 IP=3
3 IF (IP .EQ. 0) NHRM=NHRM-1
RETURN
END)
*****
SUBROUTINE MUVAR (VECT, RMU, VAR)
DIMENSION VECT(100)
SUM1=0.
SUM2=0.
DO 1 I=1,100
SUM1=SUM1+VECT(I)
SUM2=SUM2+VECT(I)*VECT(I)
1 CONTINUE
RMU=SUM1/100.
VAR=(SUM2-100.*RMU*RMU)/99.
RETURN
END
*****
SUBROUTINE TITULO (IGRA, IC)
GO TO (1,2,3,3,3,3,4,5), IGRA
1 PRINT 200
200 FORMAT(20X,"MODELO MANDEL")
RETURN
2 PRINT 201
201 FORMAT(20X,"MODELO TURKEY")
RETURN
3 PRINT 202, IG
202 FORMAT(20X,"MODELO GRAYBILL-MILLIKEN ", IG)
RETURN
4 PRINT 203
203 FORMAT(20X,"MODELO MENDEZ-MUSTOS")
RETURN
5 PRINT 204
204 FORMAT(20X,"SUMA DE F(1..J) AL CUADRADO")
RETURN
END
*****
SUBROUTINE TABLON (IJ, JJ, KK)
DIMENSION NWORD(4)
100 FORMAT(4I3)
200 FORMAT(1H1,///,22X,"MODELO MENDEZ-MUSTOS",15X,"MODELO MANDEL",/,10X
1,"MOD. ADITIVOS*EST. LA VAR.*RES1. APROPIADOS MOD. ADITIVOS" )
201 FORMAT(15X,13,11X,11,12X,11,13X,13,/)
IJ=IJ*JJ*KK
RETURN
4 PRINT 200
1 READ(4,100) NWORD
1 JK=1,JK+1
PRINT 201,NWORD

```

```

IF (IJK .EQ. 1J) RETURN
GO TO 1
END

```

```

SUBROUTINE GENFRA (ARREG1, VECT01, VECT02, VECT11, SEED)
DIMENSION ARREG1 (5,3), VECT01 (5), VECT02 (8), VECT11 (5), VNOR (1)
INTEGER SEED
COMMON /NUMBER/ M,N,MIN1,MXN, SQRMXN, DESEST, ARK (2,5), GLOUV (4)
COMMON /CINCO/ IL, CMEMAN (100), CMETUK (100), CMEGM (400), CMEMHU (100),
1 CMF (100)
DATA AMUGR/0./
GO TO 3
CALL MERFI (P,Y,IER)
3 CONTINUE
SUM1=0.
DO 2 I=1,M
DO 1 J=1,N
CALL GGNOR (SEED,1,VNOR)
TEMP=VNOR (1)*DESEST
ARREG1 (I,J)=AMUGR+VECT01 (I)+VECT02 (J)+VECT11 (I)*VECT02 (J)+TEMP
SUM1=SUM1+TEMP*TEMP
1 CONTINUE
2 CONTINUE
CMF (IL)=SUM1/MXN
RETURN
END

```

```

SUBROUTINE MEDIAR (ARREG1, VECTOR)
DIMENSION ARREG1 (5,8), VECTOR (5)
COMMON /NUMBER/ M,N,MIN1,MXN
DO 1 I=1,M
VECTOR (I)=0.
1 CONTINUE
DO 3 I=1,M
DO 2 J=1,N
VECTOR (I)=VECTOR (I)+ARREG1 (I,J)
2 CONTINUE
3 CONTINUE
VECTOR (I)=VECTOR (I)/N
1 CONTINUE
RETURN
END

```

```

SUBROUTINE MEDIAG (VECTOR, VALOR)
DIMENSION VECTOR (5)
COMMON /NUMBER/ M,N,MIN1,MXN
VALOR=0.
DO 1 I=1,M
VALOR=VALOR+VECTOR (I)
1 CONTINUE
VALOR=VALOR/M
RETURN
END

```

```

SUBROUTINE CORREC (VECTOR, VALOR)
DIMENSION VECTOR (5)
COMMON /NUMBER/ M,N,MIN1,MXN
DO 1 I=1,M
VECTOR (I)=VECTOR (I)-VALOR
1 CONTINUE
RETURN
END

```

```

SUBROUTINE MEDIAC (ARREG1, VECTOR, VALOR)
DIMENSION ARREG1 (5,8), VECTOR (8)

```

```
COMMON /NUMBER/ M,N,M1,N1,MXN
```

```
DO 1 J=1,N  
VECTOR(I)=0.
```

```
1 CONTINUE
```

```
DO 3 J=1,N
```

```
DO 2 I=1,M
```

```
VECTOR(J)=VECTOR(J)+ARREG1(I,J)
```

```
2 CONTINUE
```

```
VECTOR(J)=VECTOR(J)/I-VALUE
```

```
3 CONTINUE
```

```
RETURN
```

```
END
```

```
*****
```

```
SUBROUTINE RESIDU(ARREG1,VECTOR1,VECTOR2,VALOR,ARREG2)
```

```
DIMENSION ARREG1(5,R),ARREG2(5,R),VECTOR1(5),VECTOR2(8)
```

```
COMMON /NUMBER/ M,N,M1,N1,MXN
```

```
DO 2 I=1,M
```

```
DO 1 J=1,N
```

```
ARREG2(I,J)=ARREG1(I,J)-VALOR-VECTOR1(I)-VECTOR2(J)
```

```
1 CONTINUE
```

```
2 CONTINUE
```

```
RETURN
```

```
END
```

```
*****
```

```
SUBROUTINE ANOVA(ARREG1,ARREG2,VECTOR1,VECTOR2,VALOR)
```

```
DIMENSION ARREG1(5,R),ARREG2(5,R),VECTOR1(5),VECTOR2(8)
```

```
COMMON /NUMBER/ M,N,M1,N1,MXN
```

```
COMMON /ANOVA/ SCPR,SCPR*,SCPC,SCPRC,SCTOT
```

```
SCPM=VALOR*VALOR*MXN
```

```
SCPR=0.
```

```
DO 1 J=1,M
```

```
SCPR=SCPR+VECTOR1(I)*VECTOR1(I)
```

```
1 CONTINUE
```

```
SCPR=SCPR*11
```

```
SCPC=0.
```

```
DO 2 J=1,N
```

```
SCPC=SCPC+VECTOR2(J)*VECTOR2(J)
```

```
2 CONTINUE
```

```
SCPC=SCPC*M
```

```
SCPRC=0.
```

```
DO 4 J=1,M
```

```
DO 3 I=1,N
```

```
SCPRC=SCPRC+ARREG2(I,J)*ARREG2(I,J)
```

```
3 CONTINUE
```

```
4 CONTINUE
```

```
SCTOT=0.
```

```
DO 6 J=1,M
```

```
DO 5 I=1,N
```

```
SCTOT=SCTOT+ARREG1(I,J)*ARREG1(I,J)
```

```
5 CONTINUE
```

```
6 CONTINUE
```

```
RETURN
```

```
END
```

```
*****
```

```
SUBROUTINE ANOVA1
```

```
DIMENSION FUEN(5),PH(2),EFES(9),FST(3)
```

```
COMMON /NUMBER/ M,N,M1,N1,MXN
```

```
COMMON /ANOVA/ SCPM,SCPR,SCPC,SCPRC,SCTOT
```

```
COMMON /DATA/ NHR5(3),NHR2P5(3),NHR1(3)
```

```
COMMON /CIATRO/ SCPRC
```

```
DATA FUEN,PH,PEL / SETOTAL,5H SCM,5H SCA,5H SCH,5H SCAB,2H S,
```

```
1 2HMS,1 /
```

```
DATA EFES / 4,2005,5,60723,7,6493,2,7185,3,29306,4,08473,2,36373, 0ANOVA1
```

```
1 2,7886,3,3683 /
```

```

200 FORMAT(51X,2PH TABLA DE ANALISIS DE VARIANZA/,132(1H*))
201 FORMAT(33X,6X,3HF V,4X,3HF L,6X,6HS DE C,10X,4HC 4,10X,1HF,6X,
15X,/,132(1H*))
202 FORMAT(33X,3X,2HX ,A5,3H X ,13,3H X ,2(f12,5,3H X ,)F7,2,3H X ,
1A2,2H X,7X)
203 FORMAT(10X,2PH ANALISIS DE VARIANZA ,A5,8H AL 5,% ,A2,9H AL 2,5% ,
1A2,8H AL 1,% ,A2)
204 FORMAT(33X,3X,2HX ,A5,3H X ,13,3H X ,F12,5,3H X ,12X,3H X ,7X,
13H X ,2X,3X)
205 FORMAT(132(1H*))
SCMPR=SCPR/(M-1)
SCMPC=SCPC/(N-1)
SCMPRC=SCPRC/M/N
EST(1)=SCPM/SCMPR
EST(2)=SCMPR/SCMPRC
EST(3)=SCMPC/SCMPRC
DO 5 I=1,3
K1=1
K2=1
K3=1
J1=3*(I-1)
DO 4 J=1,3
J2=J1+J
IF(EST(1) .IF. FFES(J2)) GO TO(1,2,3),J
GO TO 4
1 K1=2
IF(I .NE. 1) NHR5(I)=NHR5(I)-1
GO TO 4
2 K2=2
IF(I .NE. 1) NHR2P5(I)=NHR2P5(I)-1
GO TO 4
3 K3=2
IF(I .NE. 1) NHR1(I)=NHR1(I)-1
4 CONTINUE
I1=1+I
C PRINT 203,FUEN(I1),PH(K1),PH(K2),PH(K3)
5 CONTINUE
C9 PRINT 200
C PRINT 201
C PRINT 204,FUEN(1),M*N,SCTOT
C PRINT 202,FUEN(2),MFD,SCPR,SCPM,EST(1)
C PRINT 202,FUEN(3),M-1,SCPR,SCMPR,EST(2)
C PRINT 202,FUEN(4),N-1,SCPC,SCMPC,EST(3)
C PRINT 202,FUEN(5),M/N,SCPRC,SCMPRC
C PRINT 205
C RETURN
END

```

```

SUBROUTINE MATSIN(D1,D2,DDT,IJOB)
DIMENSION D1(5,8),D2(8,5),DDT(1)
COMMON /NUMBER/ M,N,M1N1,MXN
K1=M
K2=N
IF(IJOB .EQ. 2) GO TO 3
DO 2 I=1,M
K=(1*(I-1))/2
DO 1 J=1,1
K=K+1
CALL VIPREF(D1(I,1),D1(J,1),K2,K1,K1,DDT(K))
1 CONTINUE
2 CONTINUE
RETURN
3 DO 5 I=1,N
K=(1*(I-1))/2

```

```

      DO 4 J=1,1
      K=K+1
      CALL VPRFF(D2(1,1),D2(J,1),K1,K2,K2,D1(K))
4     CONTINUE
5     CONTINUE
      RETURN
      END
*****
      SUBROUTINE SUMCIA(ARREG2,ICONT,SCDA,PSC)
      DIMENSION ARREG2(5,8)
      COMMON /NUMBER/ M,N,M1N1,MXN
      SCDA=0.
      PRINT /," SUMA DE CUADRADOS TOTAL "PSC," SUMA DE CUADRADOS AN
      TERIOR "SCDA
      DO 2 I=1,M
      DO 1 J=1,N
      SCDA=SCDA+ARREG2(I,J)*ARREG2(I,J)
1     CONTINUE
2     CONTINUE
      PRINT /," LA SCI OBTENIDA ES "SCD
      IF(ICONT.LE. 0) GO TO 4
      SCDA=SCDA/PSC
      SCDA=SCDA/PSC
      IF(ABS(SCD-SCDA).LE. .1) GO TO 3
      DIF=SCD-SCDA
      PRINT /," LA SUMA DE LOS CUADRADOS DE LOS RESIDUOS ACTUALES DIFE
      RE DE LA DE LOS ANTERIORES EN"DIF," SIENDO LA ACTUAL S D C = "
      PSCD=PSC
      GO TO 4
      PRINT /," LA SUMA DE CUADRADOS DE LOS RESIDUOS CONVERGE PARA "
      ICONT," Y CONVERGE ALREDEDOR DEL VALOR "SCD*PSC
3     CONTINUE
4     IF(ICONT.GT. 0) GO TO 5
      SCDA=SCD
      GO TO 6
5     SCDA=SCD*PSC
6     RETURN
      END
*****
      SUBROUTINE MANDEL(ARREG1,VECTO1,VECTO2,ARREG2,FIVA,EIVE)
      DIMENSION ARREG1(5,8),ARREG2(5,8),VECTO1(5),VECTO2(8),EIVE(5,5),
      FIVA(5)
      LOGICAL MAL
      CALL MEDIAR(ARREG1,VECTO1)
      CALL MEDIAG(VECTO1,VALOR)
      CALL CORREC(VECTO1,VALOR)
      CALL MEDIAC(ARREG1,VECTO2,VALOR)
      CALL RESIDU(ARREG1,VECTO1,VECTO2,VALOR,ARREG2)
      CALL ANOVA(ARREG1,ARREG2,VECTO1,VECTO2,VALOR)
      CALL ANOVA1
      CALL FSTVII(ARREG2,FIVA,EIVE,MAL)
      IF(.NOT.MAL)CALL MANAOV(FIVA)
      RETURN
      END
*****
      SUBROUTINE RESIDU(ARREG2,FIVA,EIVE,EIVEC,ICONT)
      DIMENSION ARREG2(5,8),FIVA(5),EIVE(5,5),EIVEC(8,8)
      COMMON /NUMBER/ M,N,M1N1,MXN
      K=M-ICONT
      K1=N-ICONT
      FIVAK=SQRT(FIVA(K))
      DO 2 I=1,M
      DO 1 J=1,N
      ARREG2(I,J)=ARREG2(I,J)-FIVAK*EIVE(I,K1)*EIVEC(J,K1)

```

```

1  CONTINUE
2  CONTINUE
RETURN
END
*****
SUBROUTINE MENHUS (ARREG2,FIVA,EIVE)
DIMENSION ARREG2(5,8),VECT1(5),EIVA(5),FIVE(5,5),EIVEC(8,8),
1ARREG1(5,8),VECT01(5),VECT02(8)
LOGICAL HOMOCF,LOG1,SIMETR,LOG2,KURTOS,LOG3,AUTOCO,LOG4,KOLMOG,
1JICUAD,LOG5,MAL,YAPASE
COMMON /ADV/,SCPM,SCPR,SCPC,SCPRC,SCTOT
COMMON /NUMBER/ M,N,MIN1,MXN,SORMXN,DESEST,ARP(2,5),GLOUV(4)
COMMON /UNO/ NHR5(3),NHR2P5(3),NHR1(3)
COMMON /CUATRO/ SCMPRC
COMMON /CINCO/ IL,CNEMAN(100),CMETUK(100),CMEGM(400),CMEMRU(100)
200 FORMAT(20X,38H LA ESTIMACION DE U EN MENHUS ESTA MAL)
C CALL MEDIAR (ARREG1,VECT01)
C CALL MEDTAG (VECT01,VALOR)
C CALL COPREG (VECT01,VALOR)
C CALL MEDIAC (ARREG1,VECT02,VALOR)
C CALL RESIDU (ARREG1,VECT01,VECT02,VALOR,ARREG2)
C CALL ANOVA (ARREG1,ARREG2,VECT01,VECT02,VALOR)
C CALL ANOVA1
YAPASE=.FALSE.
ICONT=0
SCDA=SCPRC
SCTO=SCTOT
GO TO 2
1 CALL SIMCUA (ARREG2,ICONT,SCDA,SCTO)
2 LOG1=HOMOCF (ARREG2,VECT1)
LOG2=SIMETR (ARREG2,VECT1,S2,ICONT,S4)
LOG3=KURTOS (ARREG2,S2,ICONT,S4)
LOG4=AUTOCO (ARREG2,S2)
C IF (MXN .GT. 60) GO TO 3
LOG5=KOLMOG (ARREG2,SCDA)
GO TO 4
C3 LOG5=JICUAD(
4 IF (.NOT. (LOG1 .AND. LOG2 .AND. LOG3 .AND. LOG4 .AND. LOG5)) GO TO 5
C PRINT /." LOS RESIDUOS TIENE YA UNA DISTRIBUCION MUY PROXIMA A
C ILA NORMAL"
IF (ICONT .EQ. 0) NHR5(1)=NHR5(1)+1
IF (ICONT .NE. 4) NHR2P5(1)=NHR2P5(1)+1
NHR1(1)=NHR1(1)+1
IF (ICONT .EQ. 0 .OR. ICONT .EQ. 4) GO TO 7
SUM=0.
SUM1=0.
DO 8 I=ICONT,4
M1=M+1)-1
SUM=SUM+EIVA(M1)
SUM1=SUM1+GLOUV(1)
8 CONTINUE
CMEMRU(IL)=SUM/SUM1
RETURN
7 CMEMRU(IL)=SCMPRC
RETURN
5 ICONT=ICONT+1
IF (ICONT .EQ. 5) GO TO 9
IF (YAPASE) GO TO 6
C CALL ESTIVU (ARREG2,FIVA,FIVE,MAL)
C IF (MAL) PRINT 200
CALL ESTIVV (ARREG2,FIVE,FIVA,EIVEC)
YAPASE=.TRUE.
6 CALL GESTIO (ARREG2,FIVA,FIVE,EIVEC,ICONT-1)
GO TO 1

```

```

9      CMEMAN(IL)=SCMPRC
      RETURN
      END
*****
      SUBROUTINE MANADV(VADXT)
      DIMENSION FUEN(4),VADXT(5)
      COMMON /NUMBER/ M,N,M1,MXN,SQMXR,DESEST,ARR(2,5),GLOUV(4)
      COMMON /CUATRO/ SCMPRC
      COMMON /CINCO/ IL,CMEMAN(100),CMETUK(100),CMEGH(400),CMEMHU(100)
      DATA FUEN / 5H OUV,5H OUV,5H OUV,5H OUV,5H OUV /
200    FORMAT(132(1H*),/,56X,20HPLANTEAMIENTO MANDEL,/,132(1H*))
201    FORMAT(133X,3X,2HX,AS,3H X ,F4.1,3H X ,2(F12.5,3H X ),7X,3H X ,2X,
1      2H X,3X,/)
202    FORMAT(132(1H*))
      OUVCM1=VADXT(M)/GLOUV(1)
      OUVCM2=VADXT(M-1)/GLOUV(2)
      OUVCM3=VADXT(M-2)/GLOUV(3)
      EPRCM=VADXT(M-3)/GLOUV(4)
C      PRINT 200
C      PRINT 201,FUEN(1),GLOUV(1),VADXT(M),OUVCM1
C      PRINT 201,FUEN(2),GLOUV(2),VADXT(M-1),OUVCM2
C      PRINT 201,FUEN(3),GLOUV(3),VADXT(M-2),OUVCM3
C      PRINT 201,FUEN(4),GLOUV(4),VADXT(M-3),EPRCM
C      PRINT 202
C      PRINT /,"      MANDEL      ",SCMPRC,OUVCM1,OUVCM2,OUVCM3,EPRCM
      CALL COMMA(OUVCM1,OUVCM2,OUVCM3,EPRCM,IP)
      IF (IP.EQ. 0 .OR. IP.EQ. 4) GO TO 2
      SUM=0.
      SUM1=0.
      DO 1 J=IP,4
      M1=M+1-I
      SUM=SUM+VADXT(M1)
      SUM1=SUM1+GLOUV(I)
1      CONTINUE
      CMEMAN(IL)=SUM/SUM1
      RETURN
2      CMEMAN(IL)=SCMPRC
      RETURN
      END

```

```

*****
      LOGICAL FUNCTION HOMOCF(D,PFNM)
      DIMENSION D(5,R),PFNM(5),COLM(R),REN(5),COL(R),COCH(2)
      COMMON /NUMBER/ M,N,M1,MXN
      DATA COCH / 0.52H5,0.463 /
      IQ=0
      HOMOCF=.FALSE.
      DO 1 I=1,M
      PFNM(I)=0.
1      REN(I)=0.
      DO 2 J=1,N
      COLM(J)=0.
2      COL(J)=0.
      DO 4 I=1,M
      DO 3 J=1,N
      RENM(I)=RENM(I)+D(I,J)
3      COLM(J)=COLM(J)+D(I,J)/M
4      RENM(I)=RENM(I)/N
      SUM=0.
      SUM1=0.
      DO 6 J=1,M
      DO 5 J=1,N
      REN(I)=REN(I)+(D(I,J)-RENM(I))* (D(I,J)-RENM(I))
      COL(J)=COL(J)+(D(I,J)-COLM(J))* (D(I,J)-COLM(J))
5      CONTINUE
6      CONTINUE

```

HOMOCF

```

      SUM=SUM+REN(I)
6     CONTINUE
      DO 7 J=1,N
      SUM1=SUM1+COL(J)
7     CONTINUE
      M1=M-1
      DO 8 I=1,M1
      IF (REN(I) .LE. REN(I+1)) GO TO 8
      CALL CAMBIA(REN(I),REN(I+1))
8     CONTINUE
      N1=N-1
      DO 9 J=1,N1
      IF (COL(J) .LE. COL(J+1)) GO TO 9
      CALL CAMBIA(COL(J),COL(J+1))
9     CONTINUE
      DIV=REN(M)/SUM
      IF (DIV .LE. COCH(1)) GO TO 10
C     PRINT /," LA HOMOCEASTICIDAD SE RECHAZA: COCH= ",COCH(1)," LA E
C     ISTADISTICA ES ",DIV
      GO TO 11
10    CONTINUE
C10   PRINT /," LA HOMOCEASTICIDAD SE ACEPTA: COCH= ",COCH(1)," LA ES
C     ISTADISTICA ES ",DIV
      IQ=IQ+1
11    DIV=COL(N)/SUM1
      IF (DIV .LE. COCH(2)) GO TO 12
C     PRINT /," LA HOMOCEASTICIDAD SE RECHAZA: COCH= ",COCH(2)," LA E
C     ISTADISTICA ES ",DIV
      GO TO 13
12    CONTINUE
C12   PRINT /," LA HOMOCEASTICIDAD SE ACEPTA: COCH= ",COCH(2)," LA ES
C     ISTADISTICA ES ",DIV
      IQ=IQ+1
13    IF (IQ .EQ. 2) HOMOCF=.TRUE.
      RETURN
      END

```

```

      LOGICAL FUNCTION SIMETR(D,RENM,S2,ICONT,S4)
      DIMENSION D(5,8),RENM(5),CSIM(2),PSIM(2)
      COMMON /NUMBER/ M,N,M1,M2,MXN,SORMXN,DFSEST,ARR(2,5),GLDIV(4)
      DATA CSIM,PSIM / 0.587,0.870,5.,1. /
      SIMETR=.FALSE.
      IS=0
      CALL MFDIAG(RENM,VALMG)
      S1=0.
      S2=0.
      S4=0.
      DO 2 I=1,M
      DO 1 J=1,N
      TEMP=(D(I,J)-VALMG)*(D(I,J)-VALMG)
      S1=S1+TEMP*(D(I,J)-VALMG)
      S2=S2+TEMP
      S4=S4+TEMP*TEMP
1     CONTINUE
2     CONTINUE
      SKEW=(S1*MXN*SORMXN)/(MXN*S2*SORT(S2))
C     PRINT /," EL COEFICIENTE DE SIMETRIA PARA LA MATRIZ DE LA ITERAC
C     ION ",ICONT," ES ",SKEW
      ASKEW=ABS(SKEW)
C     DO 4 I=1,2
C     DO 4 I=2,2
      IF (ASKEW .LT. CSIM(1)) GO TO 3
C     PRINT /," DE DONDE LOS RESIDUOS NO SON SIMETRICOS AUN. EL VALOR
C     L DE LA COTA ES: ",CSIM(1)

```

SIMETR

```

GO TO 4
3  TS=TS+1
C  PRINT /," DE DONDE LOS RESIDUOS TIENEN UNA SIMEIRIA MARCADA AL "
C  1,PSIM(1),"% Y EL VALOR DE LA COTA ES: ",CSIM(1)
4  CONTINUE
C  IF (TS .EQ. 2) SIMETR=.TRUE.
  IF (TS .EQ. 1) SIMETR=.TRUE.
  RETURN
END
*****
LOGICAL FUNCTION KURTOS(D,S2,ICONT,S4)
DIMENSION D(5,8),CKU(4)
COMMON /NUMBER/ M,N,MIN1,MXN
DATA CKU / 1.95,4.88,2.15,3.99 /
KURTOS=.FALSE.
IK=0
CKURT=S4*MXN/(S2*S2)
C  PRINT /," EL COEFICIENTE DE KURTOSIS EN LA ITERACCION ",ICONT,
C  1" ES IGUAL A ",CKURT
ACKURT=ABS(CKURT)
DO 4 I=1,2
DO 4 I=2,2
  IF (ACKURT .GE. CKU(2*I-1) .AND. ACKURT .LE. CKU(2*I)) GO TO 3
C  PRINT /," POR LO QUE LOS DATOS AUN SE ENCUENTRAN LEJOS DE TENER
C  1UNA KURTOSIS NORMAL. LOS EXTREMOS SON ",CKU(2*I-1)," Y ",CKU(2*I)
GO TO 4
3  IK=IK+1
C  PRINT /," POR LO QUE LOS DATOS TIENEN UNA KURTOSIS NORMAL. LOS E
C  1XTREMOS SON ",CKU(2*I-1)," Y ",CKU(2*I)
4  CONTINUE
C  IF (IK .EQ. 2) KURTOS=.TRUE.
  IF (IK .EQ. 1) KURTOS=.TRUE.
  RETURN
END
*****
LOGICAL FUNCTION AUTOCO(P,EDEN)
DIMENSION D(5,8),H(40),PSIM(2),CNOR(4)
COMMON /NUMBER/ M,N,MIN1,MXN
DATA PSIM,CNOR / 10.,2.,1.3266,2.776,1.5304,2.5722 /
AUTOCO=.FALSE.
FNUM=0.
IQ=0
DO 2 I=1,M
  K1=(I-1)*N
  DO 1 J=1,N
    R(K1+J)=D(I,J)
1  CONTINUE
2  CONTINUE
DO 3 I=2,MXN
  FNUM=FNUM+(R(I)-R(I-1))*(R(I)-R(I-1))
3  CONTINUE
Q=MXN*FNUM/((MXN-1)*EDEN)
C  PRINT /," EL VALOR DE LA RAZON DE VON NEWMAN ES: ",Q
C  DO 5 I=1,2
DO 5 I=2,2
  IF (Q .GE. CNOR(2*I-1) .AND. Q .LE. CNOR(2*I)) GO TO 4
C  PRINT /," LA INDEPENDENCIA DE LOS RESIDUOS NO ES ACEPTADA AL "
C  1PSIM(1),"% SIENDO LOS EXTREMOS ",CNOR(2*I-1)," Y ",CNOR(2*I)
GO TO 5
4  IQ=IQ+1
C  PRINT /," LA INDEPENDENCIA DE LOS RESIDUOS SE ACEPTA AL ",PSIM(1)
C  1,"% SIENDO LOS EXTREMOS ",CNOR(2*I-1)," Y ",CNOR(2*I)
5  CONTINUE
C  IF (IQ .EQ. 2) AUTOCO=.TRUE.

```

```

      IF(IQ .EQ. 1) AUTOCN=.TRUE.
      RETURN
      END
*****
      LOGICAL FUNCTION KOLMOG(I,SCD)
      DIMENSION D(5,8),ORD(5,8),DIF(5,8),ODIF(5,8)
      COMMON /NUMBER/ M,N,MIN,MXN,SQRMXN,DESEST,ARR(2,5),GLOUV(4)
      KOLMOG=.FALSE.
      QP=0.
      IQ=0
      CALL ORDENA(D,ORD,PTF)
      K=0
      DO 5 I=1,M
        IT=M+1-I
        DO 4 J=1,N
          JJ=N+1-J
          IF(JJ .EQ. N .AND. IT .EQ. M) GO TO 2
          IF(JJ .NE. N)GO TO 1
          CALL INTEGR(6,ORD(IT+1,1),ORD(IT,JJ),RINT,SCD)
          GO TO 3
1        CALL INTEGR(6,ORD(IT,JJ+1),ORD(IT,JJ),RINT,SCD)
          GO TO 3
2        CALL INTEGR(6,ORD(M,N)-4*SCD,ORD(M,N),RINT,SCD)
3        QP=QP+RINT
          K=K+1
          DIF(I,J)=AMAX1(ABS((FLOAT(K)/FLOAT(MXN))-QP),ABS((FLOAT(K-1)/FLOAT
            (MXN))-QP))
4        CONTINUE
5        CONTINUE
      CALL ORDENA(DIF,ODIF,PTF)
      DO 7 I=1,5
        DO 7 J=1,5
          IF(ODIF(I,J) .GT. ARR(I,J)) GO TO 6
          IQ=IQ+1
      PRINT /," KOLMOGOROV-SMIRNOV LA NORMALIDAD SE ACEPTA CON ",ARR(
      12,1)," SIENDO LA ESTADISTICA: ",ODIF(1,1)
      GO TO 7
6        CONTINUE
      PRINT /," KOLMOGOROV-SMIRNOV LA NO NORMALIDAD ES SIGNIFICATIVA
      1AL",ARR(2,1)," SIENDO LA ESTADISTICA:",ODIF(1,1)
7        CONTINUE
      IF(IQ .EQ. 5)KOLMOG=.TRUE.
      IF(IQ .EQ. 1) KOLMOG=.TRUE.
      RETURN
      END
*****
      SUBROUTINE INTEGR(IP,A,B,RINT,SCD)
      COMMON /NUMBER/ M,N,MIN,MXN
      VAR=SCD/(MXN-1)
      CIF=10**(-5)
1      RINT1=RINT
      L=L+1
      IIP=2*IP
      IF(A .EQ. B) GO TO 2
      CALL INTEGA(IIP,A,B,RINT,VAR)
      GO TO 3
2      RINT=0.
3      IF(ABS(RINT-RINT1) .GT. CIF)GO TO 1
      IF(L .EQ. 15) RETURN
      RETURN
      END
*****
      SUBROUTINE INTEGA(IP,A,B,RINT,VAR)
      DIMENSION DES(50)

```

```

I=0
Q1=IT=0.
CTF=1/SQRT(6.2832*VAR)
CTF1=2*VAR
HAIP=(B-A)/IP
RJ=A
1 I=I+1
C UES(I)=(1/SQRT(6.2832*VAR))*EXP(-(R-IP*RJ)/(2*VAR))
UES(I)=CTF*EXP(-RJ*RJ/CTF)
RJ=RJ+HAIP
IF(RJ .LE. B) GO TO 1
IP=IP+1
DO 4 J=1,IP
IF(J .EQ. 1 .OR. J .EQ. IP) GO TO 3
RJ=J
IF(AMOD(RJ,2) .EQ. 0) GO TO 2
RINT=RINT+4*UES(J)
GO TO 4
2 RINT=RINT+2*UES(J)
GO TO 4
3 RINT=RINT+UES(J)
4 CONTINUE
RINT=RINT*(B-A)/(IP*3)
RETURN
END

```

```

*****
SUBROUTINE ESTVU(ARREG2,EIVAL,EIVEC,MAL)
DIMENSION ARREG2(5,8),DXDTMM(15),FIVAL(5),EIVEC(5,5),WK(20)
DIMENSION ARREG3(8,5),DXDT(15)
LOGICAL MAL,PAS,PAS2
COMMON /NUMBER/ M,N,M1,M2,MXN
200 FORMAT(5X,28H1 CALCULO DEL VECTOR NUMERO,12,15H NO SE REALIZO)
201 FORMAT(5X,49H1 PARAMETRO 1JON DE LA SUBROUTINA EIGRS ESTA MAL)
202 FORMAT(5X,53H1 EXISTE UNA INCONSISTENCIA ENTRE LOS PARAMETROS 1Z E 1
1 26H1JON DE LA SUBROUTINA EIGRS)
203 FORMAT(1X,43H1 CALCULO REALIZADO POR EIGRS FUE MUY MALO)
MAL=.FALSE.
PAS=.FALSE.
PAS2=.FALSE.
CALL MATSIN(ARREG2,ARREG3,DXDTMM,1)
M0=M*(M+1)/2
M1=M
M2=MXN
CALL ORDEN(DXDTMM,DXDT,M0)
IF(AMAX1(ABS(DXDT(1)),ABS(DXDT(M0))) .GT. 1.) GO TO 1
GO TO 3
1 IF(ABS(DXDT(1)) .GT. ABS(DXDT(M0))) GO TO 2
CALL DIVIDE(DXDTMM,DXDT(M0),M0)
PAS2=.TRUE.
GO TO 3
2 CALL DIVIDE(DXDTMM,DXDT(1),M0)
PAS=.TRUE.
3 CALL EIGRS(DXDTMM,M1,2,FIVAL,EIVEC,M),WK,IFW)
IF(WK(1) .LE. 100) GO TO 4
PRINT 203
MAL=.TRUE.
4 IF(IFW .GT. 128) GO TO 5
IF(IFW .EQ. 66) GO TO 6
IF(IFW .EQ. 67) GO TO 7
GO TO 8
5 IVFC=IFW-128
PRINT 200,IVFC
RETURN
6 PRINT 201

```

```

      RETURN
7     PRINT 202
      RETURN
A     IF (PAS1) CALL DIVIDE (FIVAL,1/DXDT(1),M1)
      IF (PAS2) CALL DIVIDE (FIVAL,1/DXDT(M0),M1)
      RETURN
      END
*****
      SUBROUTINE DIVIDE (VECTOR,DENUM,K)
      DIMENSION VECTOR(1)
      DO 1 I=1,K
      VECTOR(I)=VECTOR(I)/DENOM
1     CONTINUE
      RETURN
      END
*****
      SUBROUTINE ESTIVV (ARREG2,EIVE,FIVAL,EIVFC)
      DIMENSION ARREG2(5,R),EIVE(5,5),FIVAL(5),EIVFC(A,H)
      COMMON /NUMBER/ M,N,M1N1
      N1=N-M+2
      DO 2 J=N1,N
      DO 1 I=1,M
      EIVFC(I,J)=0.
1     CONTINUE
2     CONTINUE
      M1=N-1
      DO 5 K=1,M1
      IK=M+1-K
      IC=N+1-K
      RSOE1=SQRT(FIVAL(IK))
      DO 4 J=1,N
      J1=N+1-J
      DO 3 I=1,M
      EIVFC(J,IC)=EIVFC(J1,IC)*ARREG2(I,J1)*EIVE(I,IK)
3     CONTINUE
      EIVFC(J1,IC)=EIVFC(J1,IC)/RSOE1
4     CONTINUE
5     CONTINUE
      RETURN
      END
*****
      SUBROUTINE ORDEN (H,C,K)
      DIMENSION H(15),C(15)
      DO 1 I=1,K
      C(I)=H(I)
1     CONTINUE
      CALL ORD (C,K)
      RETURN
      END
*****
      SUBROUTINE ORDENA (A,AOR,PTE)
      DIMENSION A(5,R),AOR(5,R),H(40)
      COMMON /NUMBER/ M,N,M1N1,MXN
      DO 2 I=1,M
      K1=N*(I-1)
      DO 1 J=1,N
      H(K1+J)=A(I,J)
C     C(K1+J)=J
1     CONTINUE
2     CONTINUE
      K=MXN
      CALL ORD (H,K)
      DO 7 I=1,M
      K1=N*(I-1)

```

```

      DO 6 J=1,N
      AOR(I,J)=R(K1,J)
6     CONTINUE
7     CONTINUE
      PTF=C(I)
      RETURN
      END
*****
      SUBROUTINE GRAMIL(YESI,J,FACRES,FACCES)
      DIMENSION YESI(5,B),FIJ(5,B),RESIF(5,B),FACKES(5),FACRCE(5),
      1VECRF1(5),FACCES(B),VECCF2(B),ESTVVV(4),
      2RESI(5,B)
      COMMON /ADV/ SCPH,SCPH,SCPC,SCPRC,SCTOT
      COMMON /NUMBER/ M,N,MINI,MXN
      COMMON /CINCO/ IL,CMEAN(100),CMEIUK(100),CMEGM(400),CMEMBU(100)
      CALL MEDIAR(YESI,J,FACRES)
      CALL MEDIAG(FACRES,VALOR)
      CALL CORREC(FACRES,VALOR)
      CALL MEDIAC(YESI,J,FACCES,VALOR)
      CALL RESIDU(YESI,J,FACRES,FACCES,VALOR,RESI)
      CALL ANOVA(YESI,J,RESI,FACRES,FACCES,VALOR)
      CALL ANOVA1
      K1=MXN
      DO 1 I=1,4
      CALL CRUCE(FACRES,FACRCE,I)
      CALL GENERO(FIJ,FACRCF,FACCES)
      CALL MEDIAR(FIJ,VECRF1)
      CALL MEDIAG(VECRF1,VALMGF)
      CALL CORREC(VECRF1,VALMGF)
      CALL MEDIAC(FIJ,VECCF2,VALMGF)
      CALL RESIDU(FIJ,VECRF1,VECCF2,VALMGF,RESIF)
      CALL VIPRF(YESI,J,RESIF,K1,1,1,RNUMQ1)
      CALL VIPRF(RESIF,RESIF,K1,1,1,DENQ1)
      QIES=(RNUMQ1*RNUMQ1)/DENQ1
      QOES=SCPRC-QIES
      PRINT /,"      QIES =",QIES,"      QOES =",QOES
      ESTVVV(I)=(QIES*(MINI-1))/QOES
      IL=100*(I-1)+1
      CMEGM(IL)=QOES/(MINI-1)
1     CONTINUE
      CALL ESTADI(ESTVVV)
      RETURN
      END
*****
      SUBROUTINE GENERO(ARREG1,VECTO1,VECTO2)
      DIMENSION ARREG1(5,B),VECTO1(5),VECTO2(B)
      COMMON /NUMBER/ M,N,MINI,MXN
      DO 2 I=1,M
      DO 1 J=1,N
      ARREG1(I,J)=VECTO1(I)*VECTO2(J)
1     CONTINUE
2     CONTINUE
      RETURN
      END
*****
      SUBROUTINE ESTADI(VECTO1)
      DIMENSION VECTO1(4),VECTO2(3),PH(2)
      COMMON /DOS/ NHRGM5(4),NHRGM2(4),NHRGM1(4)
      DATA PH,VECTO2 / 2H 5.2PHMS,4.2153,5.6421,7.6927 /
      200 FORMAT(10X,19HGPAYRIL=-MILLIKEN (,11,9H) AL 5.2 ,42,9H AL 2.53 ,A2
      1,8H AL 1.8 ,A2)
      DO 5 I=1,4
      K1=1
      K2=1

```

DOOMY

1M23RC

```

K3=1
DO 4 J=1,3
IF(VECT01(I).LE.VECT02(J)) GO TO (1,2,3),J
GO TO 4
1 K1=2
NHRGM5(I)=NHRGM5(I)-1
GO TO 4
2 K2=2
NHRGM2(I)=NHRGM2(I)-1
GO TO 4
3 K3=2
NHRGM1(I)=NHRGM1(I)-1
4 CONTINUE
C PRINT 200,I,PH(K1),PH(K2),PH(K3)
5 CONTINUE
C PRINT /,"      EFFE CALCULADAS GRAY-MILL",VECT01
RETURN
END
*****
SUBROUTINE PAS02(ARREG1,VECT01,VECT02)
DIMENSION ARREG1(5,8),VECT01(5),VECT02(8),FUEN(2),F1(3),PH(2),
IWENCOL(5)
COMMON /NUMBER/ M,N,M1N1
COMMON /ADV/SCPM,SCPR,SCPC,SCPRC,SCITOT
COMMON /TRES/ NHRT5,NHRT2P,NHRT1,NHRM
COMMON /CINCO/ IL,CMEAN(100),CMETUK(100),CMEGM(400),CMEMRU(100)
DATA F1,PH / 4.2153,5.6421,7.6927,2H 5,2HNS /
DATA FUEN / 5HNDIT,5HRLAN /
DATA NNDIT / 1 /
200 FORMAT(132(1H*),/,57X,19HPLANTEAMIENTO TUKEY /,132(1H*))
201 FORMAT(33X,3X,2HX ,A5,3H X ,13,3H X ,2(12,5,3H X ) ,F7.2,3H X ,
1 A2,2H X,3X)
202 FORMAT(20X,"TUKEY AL 5.% ",A2," AL 2.5% ",A2," AL 1.% ",A2)
DO 1 I=1,M
WENCOL(I)=0.
1 CONTINUE
DO 3 I=1,M
DO 2 J=1,N
WENCOL(I)=WENCOL(I)+VECT02(J)*ARREG1(I,J)
2 CONTINUE
3 CONTINUE
SUM=0.
DO 4 I=1,M
SUM=SUM+WENCOL(I)*VECT01(I)
4 CONTINUE
SDPRC=0.
DO 5 I=1,M
SDPRC=SDPRC+VECT01(I)*VECT01(I)
5 CONTINUE
SDPCC=0.
DO 6 J=1,N
SDPCC=SDPCC+VECT02(J)*VECT02(J)
6 CONTINUE
ADITNO=SUM*(SUM/(SDPCC*SDPRC))
BALANC=SCPRC-ADITNO
CMRAL=BALANC/(M1N1-1)
CMETUK(IL)=CMRAL
FNADIT=ADITNO/CMRAL
K1=1
K2=1
K3=1
DO 10 J=1,3
IF(FNADIT.LE.F1(J)) GO TO (7,8,9),J
GO TO 10

```

```

/      K1=2
      NHRT5=NHRT5-1
      GO TO 10
H      K2=2
      NHRT2P=NHRT2P-1
      GO TO 10
Q      K3=2
      NHRT1=NHRT1-1
10     CONTINUE
C      PRINT 202,PH(K1),PH(K2),PH(K3)
C      PRINT 200
C      PRINT 201,FUEN(1),NOADIT,ADITNO,ADITNO,FNADIT
C      PRINT 201,FUEN(2),MIN1-1,BALANC,CMBAL
      RETURN
      END

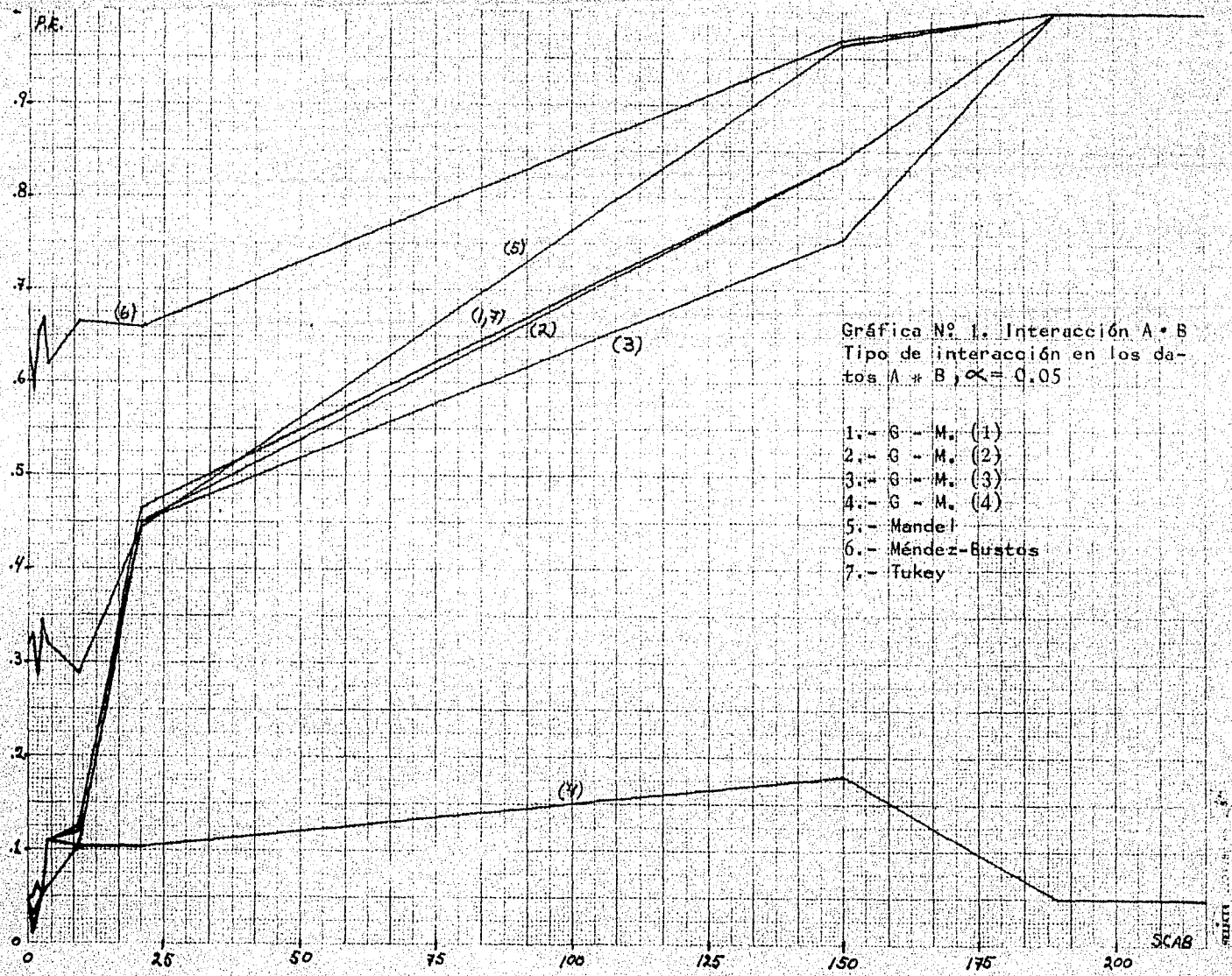
```

```

*****
SUBROUTINE TUKEY (ARREG1,VECT01,VECT02)
DIMENSION ARREG1(5,A),VECT01(5),VECT02(8),ARREG2(5,B)
C      CALL MEDIAR (ARREG1,VECT01)
C      CALL MEDIAG (VECT01,VALOR)
C      CALL CORREC (VECT01,VALOR)
C      CALL MEDIAC (ARREG1,VECT02,VALOR)
C      CALL RESIDU (ARREG1,VECT01,VECT02,VALOR,ARREG2)
C      CALL ANOVA (ARREG1,ARREG2,VECT01,VECT02,VALOR)
C      CALL ANOVA1
      CALL PASO2 (ARREG1,VECT01,VECT02)
      RETURN
      END

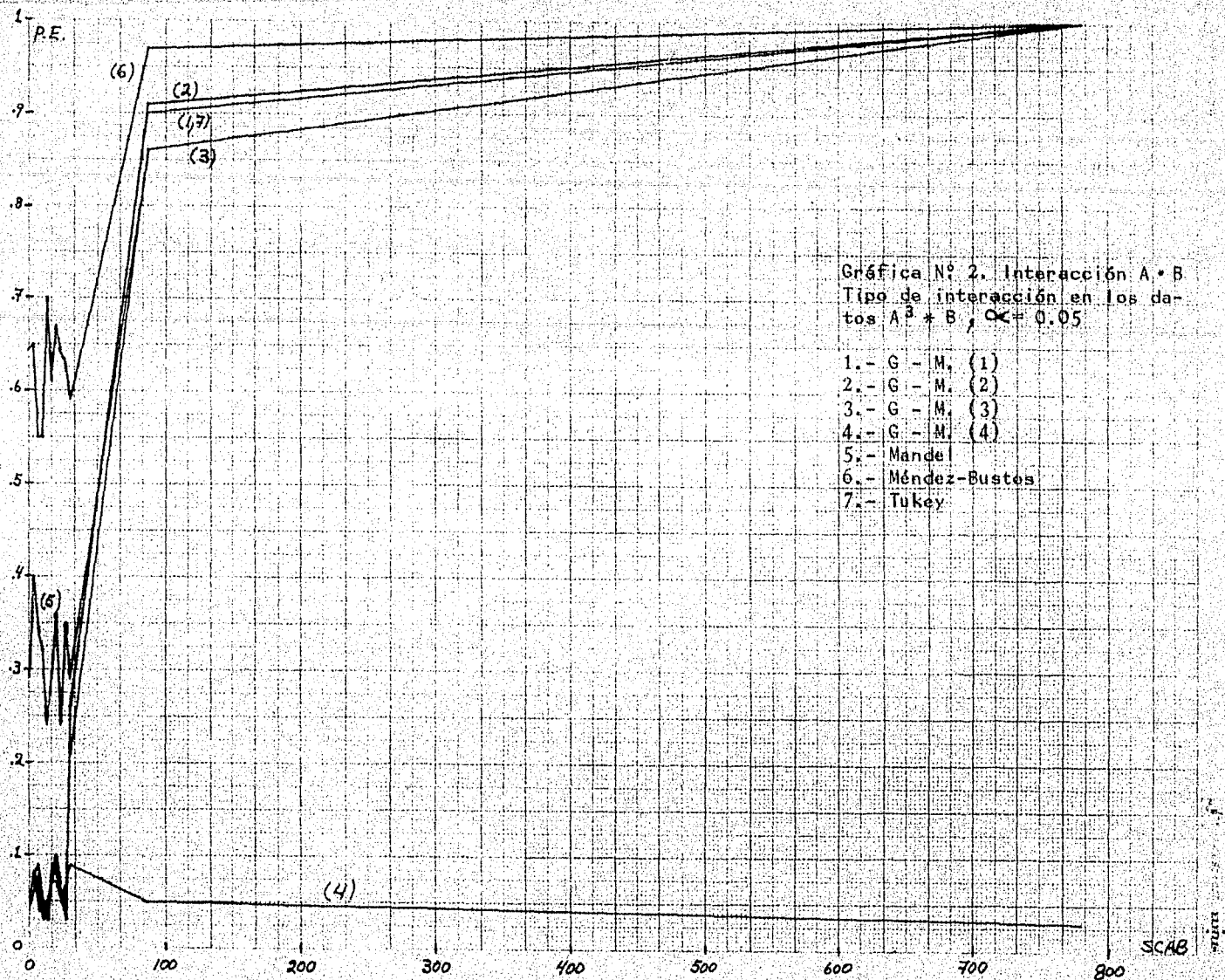
```

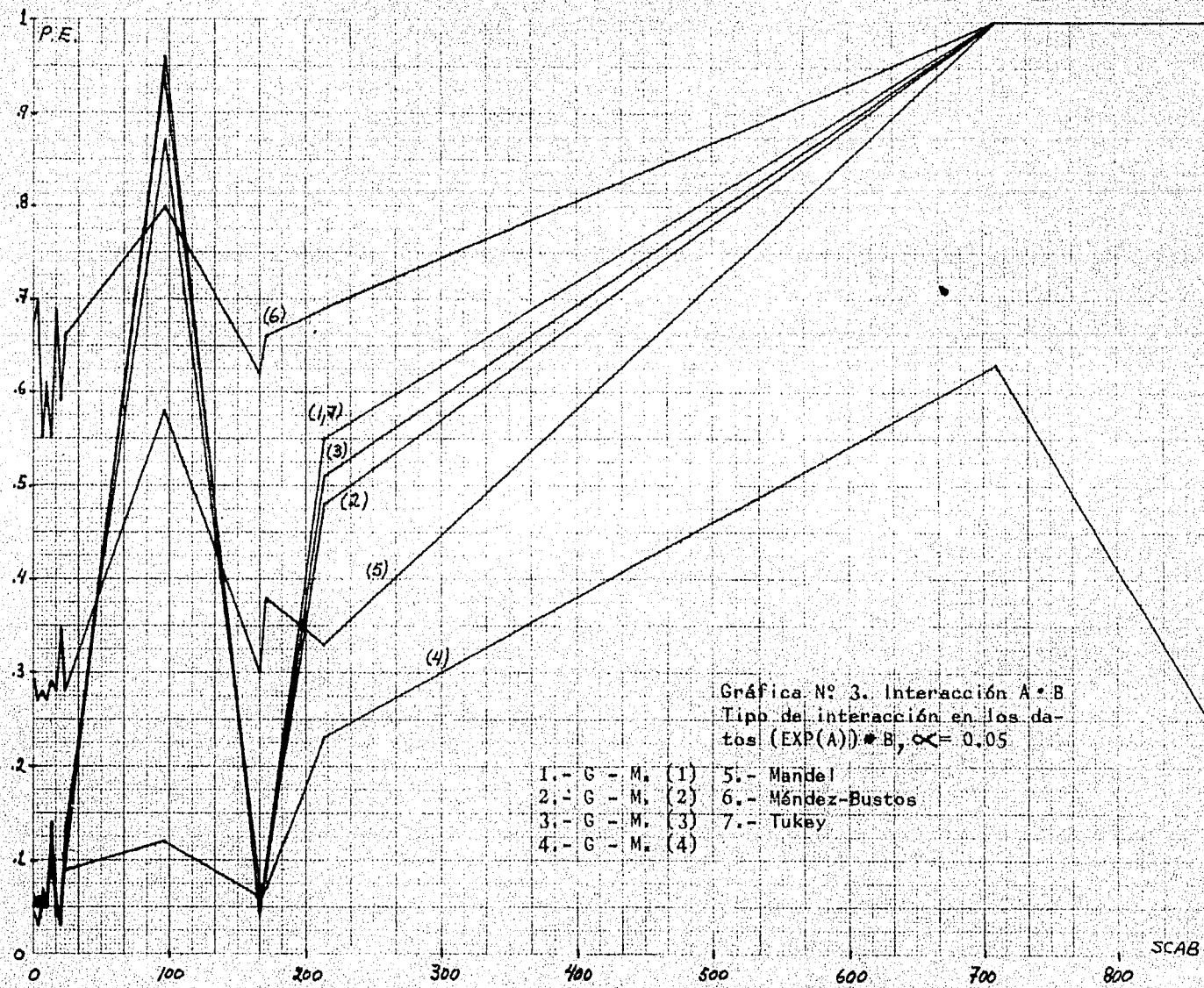
APENDICE B



Gráfica N° 1. Interacción A * B
 Tipo de interacción en los da-
 tos A * B, $\alpha = 0.05$

- 1.- G - M. (1)
- 2.- G - M. (2)
- 3.- G - M. (3)
- 4.- G - M. (4)
- 5.- Mandel
- 6.- Méndez-Bustos
- 7.- Tukey

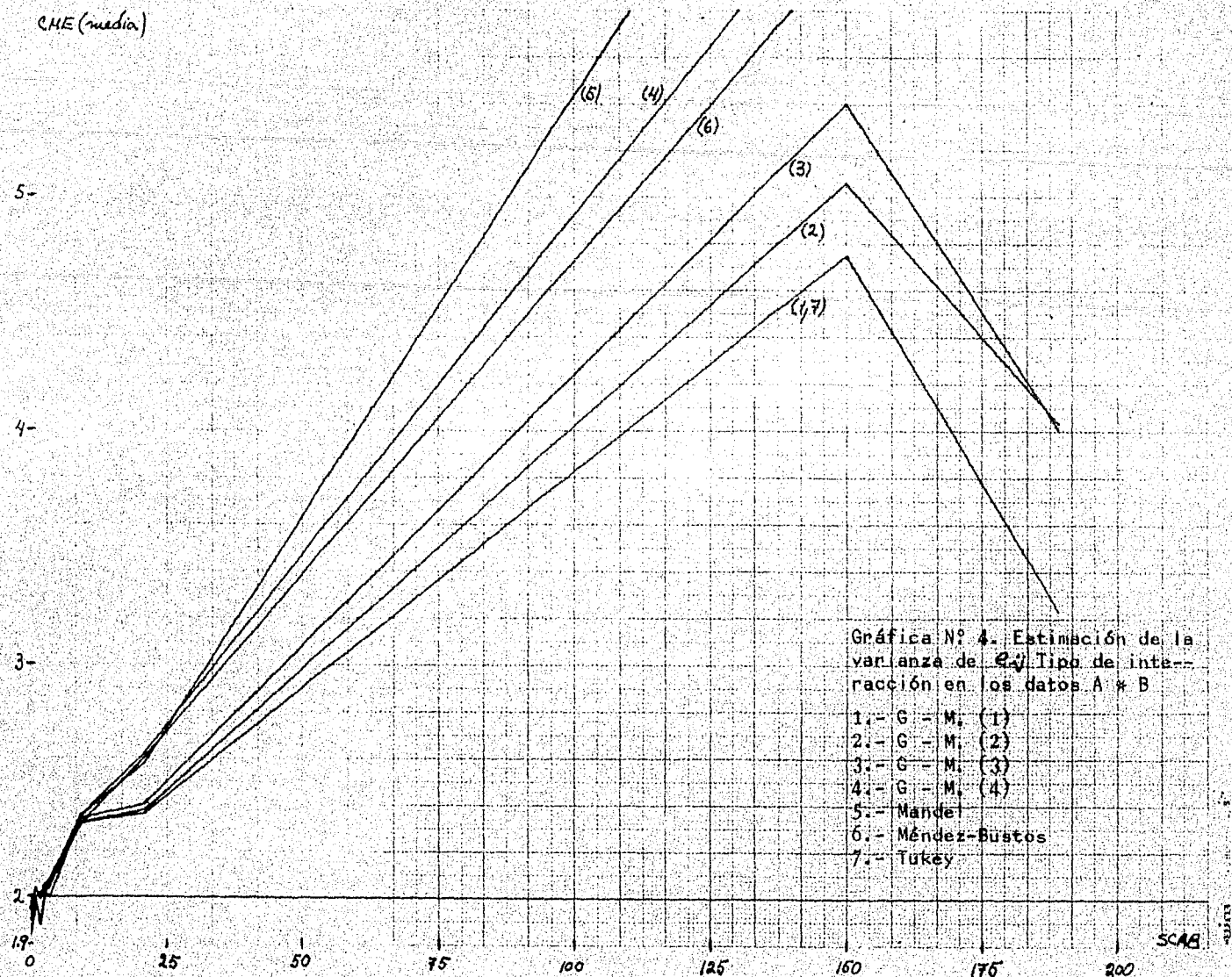


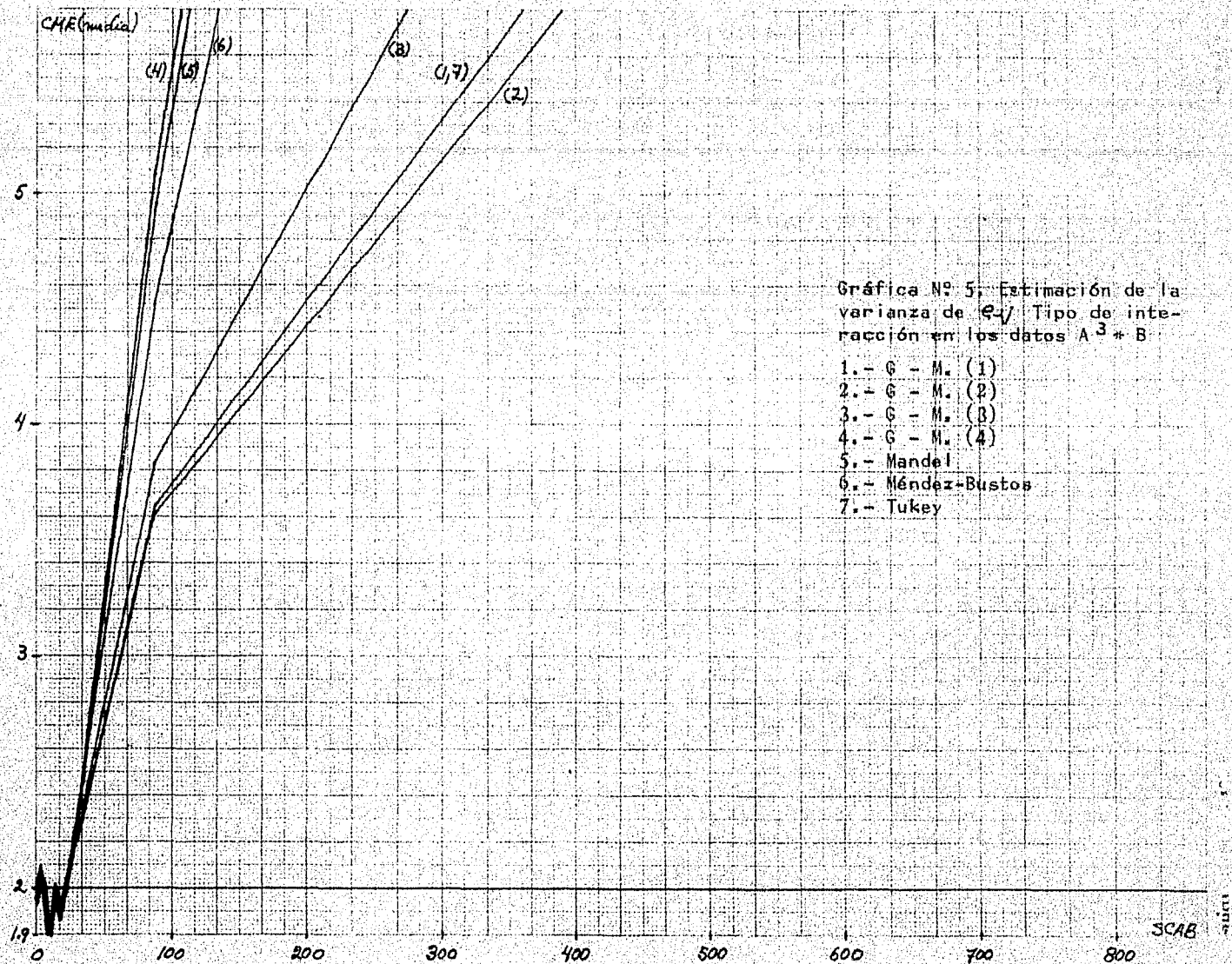


Gráfica N° 3. Interacción A*B
Tipo de interacción en los da-
tos (EXP(A)) * B, $\alpha \leq 0.05$

- | | |
|----------------|-------------------|
| 1.- G - M. (1) | 5.- Mandel |
| 2.- G - M. (2) | 6.- Mández-Bustos |
| 3.- G - M. (3) | 7.- Tukey |
| 4.- G - M. (4) | |

CME (media)





CHE(media)

Gráfica N° 6. Estimación de la
varianza de Q_y Tipo de inte-
racción en los datos (EXP(A))*

B.

- 1.- G - M. (1)
- 2.- G - M. (2)
- 3.- G - M. (3)
- 4.- G - M. (4)
- 5.- Mandel
- 6.- Méndez-Bustos
- 7.- Tukey

