Capítulo

9

Análisis de Covarianza

9.1. Introducción.

Hemos visto que el diseño de bloques es usado para eliminar el efecto de los factores de ruido que no son controlables, dicho procedimiento es valido cuando los factores que ocasionan el ruido son variables cualitativas. El análisis de la covarianza (ANCOVA) es un método que se utiliza para eliminar el efecto de ruido cuando al menos una de las variables que causan tal ruido es cuantitativa, dicha variable es conocida como covariable o variable concomitante.

Por lo tanto, el análisis de covarianza es una técnica estadística que permite conocer el efecto de una variable independiente categórica sobre una variable dependiente cuantitativa (variable respuesta), eliminando el efecto que tiene sobre esta última otra variable cuantitativa (variable concomitante). Según Montgomery (2008), el análisis de covarianza implica ajustar la variable respuesta observada para el efecto de la variable concomitante. Si no se hace este ajuste, la variable concomitante podría inflar el cuadrado medio del error y hacer que sean más difíciles de detectar las verdaderas diferencias en las respuestas debidas a los tratamientos. Por lo tanto, el análisis de covarianza es un método de ajuste para los efectos de una variable cuantitativa perturbadora no

controlable.

El análisis de la covarianza resulta ser una combinación entre el análisis de varianza (ANOVA) y el análisis de regresión.

9.2. Modelo unifactorial con una covariable

En esta sección vamos a desarrollar el análisis de covarianza más simple, cuando la variable dependiente es función de un sólo factor y una covariable. Suponiendo una relación lineal entre la variable dependiente y la covariable, la variable dependiente se puede modelar como:

$$y_{ij} = \mu + \alpha_j + \beta(x_{ij} - \bar{x}_{..}) + \varepsilon_{ij} \qquad \begin{matrix} i = 1, ..., a \\ j = 1, ..., n \end{matrix} \tag{9.1}$$

donde

y_{ij} es la j-ésima observación bajo el i-ésimo nivel del tratamiento

x_{ij} es la medida de la covariable que se hace para y_{ij}

$\bar{x}_{..}$ es la media de los valores de x_{ij}

μ es la media total.

α_i es el efecto del nivel i-ésimo del tratamiento

β es el coeficiente de regresión que relaciona y_{ij} con la covariable x_{ij}

ε_{ij} es el error aleatorio

Se asume que $\varepsilon_{ij} \sim N(0;\sigma^2)$ e independientes entre si, $\beta \neq 0$, $\sum_{i=1}^{n} \alpha_i = 0$ y que la covariable x no está afectada por los tratamientos.

Note de la ecuación (9.2.1), que el modelo de análisis de covarianza es una combinación de los modelos lineales empleados en el análisis de regresión y análisis de varianza. Es decir, se tienen efectos de los tratamientos $\{\alpha_i\}$, como en el análisis de varianza de un sólo factor, y un coeficiente de regresión β, como en una ecuación de regresión.

Para describir el análisis se introduce la siguiente notación

$$S_{yy} = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{..})^2 = \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - \frac{y_{..}^2}{an} \tag{9.2}$$

$$S_{xx} = \sum_{i=1}^{a}\sum_{j=1}^{n}(x_{ij}-\bar{x}_{..})^2 = \sum_{i=1}^{a}\sum_{j=1}^{n} x_{ij}^2 - \frac{x_{..}^2}{an} \tag{9.3}$$

$$S_{xy} = \sum_{i=1}^{a}\sum_{j=1}^{n}(x_{ij}-\bar{x}_{..})(y_{ij}-\bar{y}_{..}) = \sum_{i=1}^{a}\sum_{j=1}^{n} x_{ij}y_{ij} - \frac{(x_{..})(y_{..})}{an} \tag{9.4}$$

$$T_{yy} = n\sum_{i=1}^{a}(\bar{y}_{i.}-\bar{y}_{..})^2 = \frac{1}{n}\sum_{i=1}^{a} y_{i.}^2 - \frac{y_{..}^2}{an} \tag{9.5}$$

$$T_{xx} = n\sum_{i=1}^{a}(\bar{x}_{i.}-\bar{x}_{..})^2 = \frac{1}{n}\sum_{i=1}^{a} x_{i.}^2 - \frac{x_{..}^2}{an} \tag{9.6}$$

$$T_{xy} = n\sum_{i=1}^{a}(\bar{x}_{i.}-\bar{x}_{..})(\bar{y}_{i.}-\bar{y}_{..}) = \frac{1}{n}\sum_{i=1}^{a}(x_{i.})(y_{i.}) - \frac{(x_{..})(y_{..})}{an} \tag{9.7}$$

$$E_{yy} = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i.})^2 = S_{yy} - T_{yy} \tag{9.8}$$

$$E_{xx} = \sum_{i=1}^{a}\sum_{j=1}^{n}(x_{ij}-\bar{x}_{i.})^2 = S_{xx} - T_{xx} \tag{9.9}$$

$$E_{xy} = \sum_{i=1}^{a}\sum_{j=1}^{n}(x_{ij}-\bar{x}_{i.})(y_{ij}-\bar{y}_{i.}) = S_{xy} - T_{xy} \tag{9.10}$$

En general, $S = T + E$ donde los símbolos S, T y E son las sumas de cuadrados

y los dobles productos para el total, los tratamientos y el error respectivamente. A continuación se indica la forma en que el análisis de covarianza ajusta la variable respuesta para el efecto de la covariable. Los estimadores mínimos cuadrados de μ, α_i y β para el modelo de la ecuación (9.2.1) son, respectivamente,

$$\hat{\mu} = \bar{y}_{..}; \qquad \hat{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..} \qquad \text{y} \qquad \hat{\beta} = \frac{E_{xy}}{E_{xx}}$$

La suma de cuadrados del error en este modelo es

$$SC_E = E_{yy} - \frac{(E_{xy})^2}{E_{xx}} \tag{9.11}$$

el cual tiene $a(n-1)-1$ grados de libertad. El estimador de la varianza del error experimental esta dado por

$$CM_E = \frac{SC_E}{a(n-1)-1} \tag{9.12}$$

Ahora supongamos que no hay ningun efecto de los tratamientos, entonces para modelar la variable respuesta usamos el siguiente modelo

$$y_{ij} = \mu + \beta(x_{ij} - \bar{x}_{..}) + \varepsilon_{ij} \qquad \begin{matrix} i = 1, ..., a \\ j = 1, ..., n \end{matrix} \tag{9.13}$$

el cual es un modelo de regresión lineal simple. Los estimadores de mínimos cuadrados de μ y β para este modelo son, respectivamente,

$$\hat{\mu} = \bar{y}_{..}; \qquad \text{y} \qquad \hat{\beta} = \frac{E_{xy}}{E_{xx}}$$

La suma de cuadrada en este caso es

$$SC'_E = S_{yy} - \frac{(S_{xy})^2}{S_{xx}} \tag{9.14}$$

el cual tiene $an - 2$ grados de libertad. En la ecuación (9.14), la cantidad $\frac{(S_{xy})^2}{S_{xx}}$ es la reducción de la suma de cuadrados de y, obtenida por la regresión lineal de y sobre x. Además, $SC_E < SC'_E$ ya que el modelo de la ecuación (9.2.1) incluye los parámetros adicionales $\{\alpha_i\}$. Por lo tanto, la diferencia entre SC'_E y SC_E, es decir,$SC'_E - SC_E$ es una reducción en la suma de cuadrados debida a los términos $\{\alpha_i\}$, la cual tiene $a - 1$ grados de libertad, dicha cantidad es usada para probar la hiótesis de que no hay ningún efecto de los tratamientos luego de anular el efecto de la covariable sobre la respuesta. Es decir, para probar la hipótesis $H_0 : \alpha_i = 0$, se usa como estadístico de prueba

$$F_0 = \frac{\frac{SC'_E - SC_E}{a-1}}{\frac{SC_E}{a(n-1)-1}} \tag{9.15}$$

el cual, si la hipótesis nula es verdadera, se distribuye $F_{a-1,a(n-1)-1}$. La hipótesis nula es rechazada si $F_0 > F_{\alpha,a-1,a(n-1)-1}$. Todo este procedimiento se resume en la tabla (9.1), conocida como tabla de análisis de covarianza de un solo factor y una covariable. Los resultados presentados en esta tabla son útiles también para el cálculo de las medias de los tratamientos ajustadas, las cuales están dadas por

$$\bar{y}_{i.ajustada} = \bar{y}_{i.} - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..}) \qquad i = 1, 2, ..., a \tag{9.16}$$

Esta media de los tratamientos ajustada es el estimador de mínimos cuadrados de $\mu + \alpha_i, i = 1, 2, ..., a$, en el modelo (9.2.1). El error estándar de cualquier media ajustada

Tabla 9.1: Anlisis de covarianza de un experimento de un solo factor con una covariable

Fuente de Variación	GL	Sumas de Cuadrados			Ajustados por la regresión			F
		x	xy	y	y	GL	CM	
Tratamiento	$a-1$	T_{xx}	T_{xy}	T_{yy}			$CM_E = \frac{SC_E}{a(n-1)-1}$	$F_0 = \frac{\frac{SC'_E - SC_E}{a-1}}{CM_E}$
Error	$a(n-1)$	E_{xx}	E_{xy}	T_{yy}	$SC_E = E_{yy} - \frac{(E_{xy})^2}{E_{xx}}$	$a(n-1)-1$		
Total	$an-1$	S_{xx}	S_{xy}	S_{yy}	$SC'_E = S_{yy} - \frac{(S_{xy})^2}{S_{xx}}$	$an-2$		
Tratamiento					$SC'_E - SC_E$	$a-1$	$\frac{SC'_E - SC_E}{a-1}$	

de los tratamientos es

$$S_{\bar{y}_{i.ajustada}} = \left[CM_E \left(\frac{1}{n} + \frac{(\bar{x}_{i.} - \bar{x}_{..})^2}{E_{xx}} \right) \right]^{1/2} \tag{9.17}$$

Otra hipótesis de interés a probar está relacionada con el parámetro de regresión, es decir, $H_0 : \beta = 0$. El no rechazo de dicha hipótesis indica que la covariable puede omitirse del estudio. E estadístico de prueba en este caso es

$$F_1 = \frac{(E_{xy})^2/E_{xx}}{CM_E} \tag{9.18}$$

que bajo la hipótesis nula se distribuye $F_{1,a(n-1)-1}$. Por lo tanto, $H_0 : \beta = 0$ se rechaza si $F_0 > F_{\alpha,1,a(n-1)-1}$.

Ejemplo 9.2.1 (Montgomery) *Considere un estudio realizado para determinar si existe diferencia en la resistencia de una fibra de monofilamento producida por tres máquinas diferentes. Se sospecha que, la resistencia de la fibra también se afecta por su grosor; por consiguiente, una fibra más gruesa será por lo general más resistente que una delgada. Los datos de este experimento se muestran en la tabla (9.2). Es evidente*

Tabla 9.2: Datos de la resistencia a la ruptura

Máquina 1		Máquina 2		Máquina 3	
y	x	y	x	y	x
36	20	40	22	35	21
41	25	48	28	37	23
39	24	39	22	42	26
42	25	45	30	34	21
49	32	44	28	32	15
207	126	216	130	180	106

que para resolver el problema debemos realizar un análisis de covarianza con el objeto

de eliminar el efecto del grosor (x) *sobre la resistencia* (y)*. Suponiendo que la relación lineal entre la resistencia a la ruptura y el diámetro es apropiada, el modelo es*

$$y_{ij} = \mu + \alpha_j + \beta(x_{ij} - \bar{x}_{..}) + \varepsilon_{ij} \qquad \begin{matrix} i = 1, ..., a \\ j = 1, ..., n \end{matrix}$$

Utilizando las ecuaciones (9.3) a (9.10), pueden calcularse

$$\begin{aligned}
S_{yy} &= \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - \frac{y_{..}^2}{an} = 36^2 + 41^2 + ... + 32^2 - \frac{(603)^2}{(3)(5)} = 346{,}40 \\
S_{xx} &= \sum_{i=1}^{a}\sum_{j=1}^{n} x_{ij}^2 - \frac{x_{..}^2}{an} = 20^2 + 25^2 + ... + 15^2 - \frac{(362)^2}{(3)(5)} = 261{,}73 \\
S_{xy} &= \sum_{i=1}^{a}\sum_{j=1}^{n} x_{ij}y_{ij} - \frac{(x_{..})(y_{..})}{an} = (20)(36) + (25)(41) + ... + (15)(32) \\
&= -\frac{(362)(603)}{(3)(5)} = 282{,}60 \\
T_{yy} &= \frac{1}{n}\sum_{i=1}^{a} y_{i.}^2 - \frac{y_{..}^2}{an} = \frac{1}{5}(207^2 + 216^2 + 180^2) - \frac{(603)^2}{(3)(5)} = 140{,}40 \\
T_{xx} &= \frac{1}{n}\sum_{i=1}^{a} x_{i.}^2 - \frac{x_{..}^2}{an} = \frac{1}{5}(126^2 + 130^2 + 106^2) - \frac{(362)^2}{(3)(5)} = 66{,}13 \\
T_{xy} &= \frac{1}{n}\sum_{i=1}^{a}(x_{i.})(y_{i.}) - \frac{(x_{..})(y_{..})}{an} = \frac{1}{5}[(126)(207) + (130)(216) + (106)(184)] \\
&= -\frac{(362)(603)}{(3)(5)} = 96{,}00 \\
E_{yy} &= S_{yy} - T_{yy} = 346{,}40 - 140{,}40 = 206{,}00 \\
E_{xx} &= S_{xx} - T_{xx} = 261{,}73 - 66{,}13 = 195{,}60 \\
E_{xy} &= S_{xy} - T_{xy} = 282{,}60 - 96{,}00 = 186{,}60
\end{aligned}$$

Por la ecuación (9.14) se encuentra

$$SC'_E = S_{yy} - \frac{(S_{xy})^2}{S_{xx}} = 346{,}40 - \frac{186{,}60^2}{261{,}73} = 41{,}27$$

con $an - 2 = (3)(5) - 2 = 13$ *grados de libertad; y por la ecuación (9.11)*

$$SC_E = E_{yy} - \frac{(E_{xy})^2}{E_{xx}} = 206{,}00 - \frac{186{,}60^2}{195{,}60} = 41{,}27 = 27{,}99$$

con $a(n-1) - 1 = 3(5-1) - 1 = 11$ *grados de libertad.*

La suma de cuadrados para probar $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0$ *es*

$$SC'_E - SC_E = 41{,}27 - 27{,}99 = 13{,}28$$

con $a - 1 = 3 - 1 = 2$ *grados de libertad. Estos datos se resumen en la tabla 9.3 Para*

Tabla 9.3: Anlisis de covarianza de los datos de la resistencia a la ruptura

Fuente de Variación	GL	Sumas de Cuadrados			Ajustados por la regresión			F
		x	xy	y	y	GL	CM	
Tratamiento	2	66.13	96.00	140.40				
Error	12	195.60	186.60	206.00	27.99	11	2.54	
Total	14	261.73	282.60	346.40	41.27	13		
Tratamiento					13.28	2	6.64	2.61

probar la hipótesis de que las máquinas difieren en la resistencia a la ruptura de la fibra producida, es decir, $H_0 : \alpha_i = 0$*, por la ecuación (9.19) el estadístico de prueba se calcula como*

$$\begin{aligned} F_0 &= \frac{\frac{SC'_E - SC_E}{a-1}}{\frac{SC_E}{a(n-1)-1}} \\ &= \frac{13{,}28/2}{27{,}99/11} = 2{,}91 \end{aligned}$$

Al comparar este valor con $F_{0,10,2,11} = 2{,}86$*, se encuentra que no puede rechazarse la hipótesis nula. Por lo tanto, no hay evidencia sólida de que las fibras producidas por las tres máquinas difieran en la resistencia a la ruptura.*

La estimación del coeficiente de regresión se calcula con

$$\hat{\beta} = \frac{E_{xy}}{E_{xx}} = \frac{186{,}60}{195{,}60} = 0{,}9540$$

La hipótesis $H_0 : \beta = 0$ *puede probarse usando la ecuación (9.18. El estadístico de prueba es*

$$F_1 = \frac{(E_{xy})^2/E_{xx}}{CM_E} = \frac{(186{,}60)^2/195{,}60}{2{,}54} = 70{,}08 \tag{9.19}$$

y puesto que $F_{0,01,1,11} = 9{,}65$*, se rechaza la hipótesis de que* $\beta = 0$*. Por lo tanto, existe una relación lineal entre la resistencia a la ruptura y el diámetro, y el ajuste proporcionado por el análisis de covarianza fue necesario.*

Las medias de los tratamientos ajustadas pueden calcularse con la ecuación (9.17). Estas medias ajustadas son

$$\begin{aligned}
\bar{y}_{1.ajustada} &= \bar{y}_{1.} - \hat{\beta}(\bar{x}_{1.} - \bar{x}_{..} \\
&= 41{,}40 - (0{,}9540)(25{,}20 - 24{,}13) = 40{,}38 \\
\bar{y}_{2.ajustada} &= \bar{y}_{2.} - \hat{\beta}(\bar{x}_{2.} - \bar{x}_{..} \\
&= 43{,}20 - (0{,}9540)(26{,}00 - 24{,}13) = 41{,}42 \\
\bar{y}_{3.ajustada} &= \bar{y}_{3.} - \hat{\beta}(\bar{x}_{3.} - \bar{x}_{..} \\
&= 36{,}00 - (0{,}9540)(21{,}20 - 24{,}13) = 38{,}80
\end{aligned}$$

Al comparar las medias ajustadas con las medias no ajustadas de los tratamientos (las $\bar{y}_{i.}$*), se observa que las medias ajustadas se encuentran mucho más próximas entre sí,*

una indicación más de que el análisis de covarianza fue necesario.

Un supuesto básico en el análisis de covarianza es que los tratamientos no influyen en la covariable x, ya que la técnica elimina el efecto de las variaciones en las $\bar{x}_{i.}$. Sin embargo, si la variabilidad en la $\bar{x}_{i.}$ se debe en parte a los tratamientos, entonces el análisis de covarianza elimina parte del efecto de los tratamientos. Por lo tanto, deberá tenerse una seguridad razonable de que los tratamientos no afectan los valores de x_{ij}. En algunos experimentos esto puede ser obvio a partir de la naturaleza de la covariable, mientras que en otros puede ser más dudoso. En el ejemplo tratado aquí puede haber una diferencia en el diámetro de la fibra (x_{ij}) entre las tres máquinas. En tales casos, Cochran y Cox sugieren la posible utilidad de un análisis de varianza de los valores x_{ij} para determinar la validez de este supuesto. Para el problema tratado aquí, con este procedimiento se obtiene

$$F_0 = \frac{66{,}13/2}{195{,}60/12} = 2{,}03$$

que es menor que $F_{0,10,2,12} = 2{,}81$, por lo que no hay razón para creer que las máquinas producen fibras con diámetros diferentes.

Una manera alternativa de presentar el desarrollo del análisis de covarianza se resume en la tabla (9.4). En ella el análisis de covarianza se presenta como una análisis de varianza "ajustado". En la columna de la fuente de variación, la variabilidad total se mide por S_{yy}, con $an-1$ grados de libertad. La fuente de variación "regresión" tiene la suma de cuadrados $\frac{(S_{xy})^2}{S_{xx}}$ con un grado de libertad. Si no hubiera ninguna variable concomitante, se tendría $S_{xy} = S_{xx} = E_{xy} = E_{xx} = 0$. Entonces la suma de cuadrados del error sería simplemente E_{yy} y la suma de cuadrados de los tratamientos sería $S_{yy} - E_{yy} = T_{yy}$. Sin embargo, debido a la presencia de la variable concomitante, S_{yy} y E_{yy} deben "ajustarse" para la regresión de y sobre x, como se muestra en la tabla

(9.4). La suma de cuadrados del error ajustada tiene $a(n-1)-1$ grados de libertad en lugar de $a(n-1)$ grados de libertad debido a que se ajusta un parámetro adicional (la pendiente β) a los datos.

Tabla 9.4: El Anlisis de covarianza como un anlisis de varianza ajustado

Fuente de Variación	Suma de Cuadrados	Grados de Libertad	Cuadrado Medio	F
Regresión	$(S_{xy})^2/S_{xx}$	1		
Tratamientos	$SC'_E - SC_E$	$a-1$	$\frac{SC'_E - SC_E}{a-1}$	$\frac{(SC'_E - SC_E)/(a-1)}{CM_E}$
Error	SC_E	$a(n-1)-1$	$CM_E = \frac{SC_E}{a(n-1)-1}$	
Total	S_{yy}	$an-1$		

9.3. Ejercicios

1. Un distribuidor de bebidas gaseosas está estudiando la efectividad de los métodos de descarga. Se han desarrollado tres tipos diferentes de carretillas, y se lleva a cabo un experimento en el laboratorio de ingeniería de métodos de la compañía. La variable de interés es el tiempo de descarga en minutos (y): sin embargo, el tiempo de descarga también guarda una estrecha relación con el volumen de las cajas guardadas (x). Cada carretilla se usó cuatro veces y se obtuvieron los siguientes datos. Analizar estos datos y sacar las conclusiones apropiadas.

Carretilla 1		Carretilla 2		Carretilla 3	
y	x	y	x	y	x
27	24	25	26	40	38
44	40	35	32	22	26
31	34	46	42	53	50
41	40	26	25	18	20

2. Calcular las medias ajustadas de los tratamientos y los errores estándar de éstas para los datos del problema anterior.

3. A continuación se presentan las sumas de cuadrados y los productos de un análisis de covarianza de un sólo factor. Terminar el análisis y sacar las conclusiones apropiadas. Utilizar $\alpha = 0{,}05$.

Fuente de Variación	Grados de libertad	Sumas de cuadrados y productos		
		x	xy	y
Tratamiento	3	1500	1000	650
Error	12	6000	1200	550
Total	15	7500	2200	1200

4. Se están probando cuatro formulaciones diferentes de un adhesivo industrial. La resistencia a la tensión del adhesivo cuando se aplica para unir piezas se relaciona con el espesor de la aplicación. Se obtienen cinco observaciones de la resistencia (y) en libras y del espesor (x) en 0.01 pulgadas para cada formulación. Los datos se muestran en la siguiente tabla. Analizar estos datos y sacar las conclusiones apropiadas.

Formulación del adhesivo							
1		2		3		4	
y	x	y	x	y	x	y	x
46.5	13	48.7	12	46.3	15	44.7	16
45.9	14	49.0	10	47.1	14	43.0	15
49.8	12	50.1	11	48.9	11	51.0	10
46.1	12	48.5	12	48.2	11	48.1	12
44.3	14	45.2	14	50.3	10	48.6	11

5. Calcular las medias ajustadas de los tratamientos y los errores estándar de éstas para los datos del problema anterior.

6. Un ingeniero estudia el efecto de la rapidez de corte sobre el índice de metal eliminado en una operación de maquinado. Sin embargo, el índice de metal eliminado se relaciona también con la dureza del ejemplar de prueba. Se hacen cinco observaciones de cada rapide de corte. La cantidad de metal eliminado (y) y la dureza del ejemplar (x) se muestran en la siguiente tabla. Analizar los datos usando un análisis de covarianza. Utilizar $\alpha = 0{,}05$.

Rapidez de corte (rpm)					
1000		1200		1400	
y	x	y	x	y	x
68	120	112	165	118	175
90	140	94	140	82	132
98	150	65	120	73	124
77	125	74	125	92	141
88	136	85	133	80	130

7. Demostrar que en un análisis de covarianza de un solo factor con una sola covariable, un intervalo de confianza de $100(1-\alpha)$ por ciento para la media ajustada del tratamiento $i-$ésimo es

$$\bar{y}_{i.} - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..}) \pm t_{\alpha/2,a(n-1)-1} \left[CM_E \left(\frac{1}{n} + \frac{(\bar{x}_{i.} - \bar{x}_{..})^2}{E_{xx}} \right) \right]^{1/2}$$

Usando esta fórmula, calcular un intervalo de confianza de 95 % para la media ajustada para la rapidez de corte de 1000 del ejercicio anterior.

8. Demostrar que en un análisis de covarianza de un solo factor con una sola covariable, el error estándar de la diferencia entre dos medias ajustadas de los

tratamientos cualesquiera es

$$S_{\bar{y}_{i.ajustada}-\bar{y}_{j.ajustada}} = \left[CM_E \left(\frac{2}{n} + \frac{(\bar{x}_{i.} - \bar{x}_{j.})^2}{E_{xx}} \right) \right]^{1/2}$$