ELIMINACION DEL EFECTO DE LA COVARIABLE EN LOS DISEÑOS DE ANALISIS DE COVARIANZA

por FRANCISCO JAVIER TEJEDOR
Universidad de Santiago

1. *El análisis de covarianza*

El análisis de covarianza (Ancova) puede entenderse como una combinación de las técnicas de análisis de regresión y análisis de varianza.

En el análisis de varianza las variables independientes con las que se trabaja son de naturaleza cualitativa, nominal o categórica y sus modalidades (que se asocian con los distintos grupos del diseño) se denominan niveles, efectos o tratamientos. Las variables independientes incluidas en los diseños de análisis de varianza reciben el nombre de «factores».

Por su parte, en principio, en el análisis de regresión las variables independientes son continuas, de naturaleza estrictamente cuantitativa (la posibilidad de incorporar a la regresión variables nominales y la similitud de resultados que entonces se produce con los obtenidos en el análisis de varianza nos lleva a pensar en el modelo lineal general que subyace en gran parte de las técnicas estadísticas de investigación aplicada).

Planteadas así las cosas, el análisis de covarianza puede entenderse como un análisis de varianza en el que algunas de las variables independientes son estrictamente cuantitativas. Estas variables independientes cuantitativas incluidas en un diseño de análisis de varianza reciben el nombre de «covariables» o «concomitantes».

Algunos autores diferencian entre ambos términos reservando el término «covariable» para el caso en el que existe interacción entre los dos tipos de variables independientes incluidas en el diseño (los factores y las variables cuantitativas) y denominando «concomitantes» a las variables independientes cuantitativas cuando no se considera su interacción con los factores.

Nosotros, atendiendo expresamente a uno de los presupuestos estadísticos del análisis de covarianza que establece la exigencia de independencia entre la variable cuantitativa y los tratamientos, nos referiremos exclusivamente a esa situación y denominaremos «covariable» a toda variable cuantitativa incorporada a un diseño de análisis de covarianza, entendiendo que éstas desempeñan un papel distinto al que cabe atribuir a un factor en un diseño de análisis de varianza.

El modelo estadístico lineal que expresa esta «fusión», en los diseños con una sola variable independiente y una sola covariable, se formula como:

$$Y_{ij} = \mu + \alpha_i + \beta X_{ij} + \epsilon_{ij}$$

Donde:

Y_{ij} = Puntuación en la variable dependiente y del sujeto j para el tratamiento i.

μ = Efecto medio verdadero.

α_1 = Efecto verdadero del iésimo tratamiento o nivel del factor A (variable independiente incluida en el modelo).

$\beta\chi_{ij}$ = Efecto debido a la regresión; β es el valor de la pendiente; χ_{ij} es la puntuación de la covariable x del sujeto para el tratamiento i.

ε_{ij} = Efecto verdadero de la jotaésima unidad experimental (sujeto) en el iésimo tratamiento. Incluirá los efectos de variables extrañas. Es una variable aleatoria N (O,σ).

El investigador recurre a su utilización cuando ha decidido incluir en un diseño experimental alguna covariable para controlar su efecto sobre la variable dependiente, lo que le permitirá estudiar con más precisión el efecto que sobre ésta puedan tener las variables independientes que hemos incluido en nuestro estudio experimental (o cuasiexperimental).

La conocida dificultad para establecer en la investigación social los controles suficientes que nos permitan aproximarnos al establecimiento correcto de relaciones causales convierten al Ancova en una técnica específicamente idónea para la investigación educativa, ya que la inclusión de covariables en el diseño supone ampliar, por la vía estadística, las posibilidades de control de variables cuando no es posible hacerlo por medios físicos o estrictamente experimentales. Especialmente útil resulta, por ejemplo, para estudiar en el diseño «antes-después» la influencia del pre-test sobre los resultados del post-test (para ello se incorporarían al diseño los resultados del pre-test como covariable).

En los diseños de análisis de varianza se supone que el influjo sobre la variable dependiente de las posibles variables intervinientes se controla a través de métodos de control experimental: doble proceso de aleatorización (en la asignación de las unidades experimentales a los grupos y en la asignación de los tratamientos a los grupos una vez constituidos), constancia en las condiciones de aplicación de los tratamientos... En muchos casos no es posible incorporar al diseño todas las estrategias deseables de control experimental ya que, por ejemplo, el investigador no puede constituir aleatoriamente los grupos de trabajo sino que ha de contar con los grupos ya establecidos (alumnos de una clase, trabajadores de una empresa...), lo que va a incidir en el diseño incrementando el término de error experimental al no haber eliminado por aleatorización la fuente de error que supone la «individualidad» de las unidades experimentales.

Si tenemos en cuenta que la repercusión del incremento del término de error (Tejedor, 1979) supone una disminución del valor del estadístico de contraste que pone a prueba el efecto de los tratamientos, comprenderemos fácilmente la razón de ser de la incorporación de la covariable en los diseños de investigación experimental: aumentar la precisión del diseño de análisis de varianza eliminando de la variable dependiente (por medio de técnicas de regresión) aquella parte que es debida al efecto de la covariable. Se obtendrán así nuevas puntuaciones, denominadas puntuaciones «ajustadas por regresión», que nos proporcionan nuevas sumas de cuadrados, denominadas «sumas de cuadrados ajustadas por regresión»; el análisis de varianza que realicemos a partir de estas sumas de cuadrados nos permitirá interpretar la hipótesis básica del estudio (el efecto de los tratamientos) en condiciones de mayor control estadístico; es decir, en condiciones de mayor credibilidad estadística.

La gama de diseños de análisis de covarianza es muy variada, tanto como la de diseños de análisis de varianza. Así podemos hablar de diseños de uno o más factores, con una o más covariables, con aleatorización completa o por bloques, con efectos fijos o aleatorios, con medidas repetidas o no...

No es nuestra intención proceder a la presentación formalizada de los presupuestos estadísticos y metodológicos del Ancova ni detenernos en los procesos de cálculo implicados en las diversas modalidades que pueden adoptar estos diseños (Huitema, 1980; Arnau, 1984; Tejedor, 1984; Amón, 1985; Sanmartín y Pardo, 1989; Bisquerra, 1989; Etxeberria, Joaristi y Lizasoain, 1990).

Vamos a presentar dos procedimientos estadísticos para eliminar el efecto de la covariable en los diseños de Ancova de un factor. El primero de ellos responde a las pautas conceptuales implícitas en el Ancova: estudio de las diferencias provocadas en la variable dependiente (Y) por los distintos niveles de la variable independiente (factor A) una vez que hemos eliminado en la variable dependiente el efecto de la covariable X (mediante el estudio de la regresión de Y sobre X). El segundo, un procedimiento «abreviado» que realiza el ajuste de la variable Y sin la preocupación de seguir «literalmente» la definición conceptual del Ancova; este segundo procedimiento es el normalmente presentado en los libros de texto y, por tanto, el que suele utilizarse para la resolución de los diseños de Ancova; también es el seguido por los algoritmos de los programas estadísticos informáticos.

Desde nuestro punto de vista, el carácter exclusivamente operativo del procedimiento abreviado no facilita al investigador la comprensión de los fundamentos estadísticos subyacentes en el Ancova. Contribuir a esa comprensión es lo que pretendemos al presentar el «primer procedimiento».

2. *Ejemplo de aplicación*

Veamos con detalle sobre un ejemplo ficticio cómo se procedería en la aplicación de cada uno de estos procedimientos para la eliminación del efecto de la covariable en un diseño de Ancova.

Supongamos que estamos investigando el efecto que el método de lectura (variable independiente o factor A, con cuatro modalidades o niveles: a_1, a_2, a_3 y a_4) tiene sobre la variable «comprensión lectora» (variable dependiente) en alumnos del primer ciclo de EGB. En el estudio se han controlado determinadas variables consideradas de interés en el marco teórico que orienta la realización del trabajo: tiempo dedicado a la actividad instructiva, motivación e instrucción previa del alumno, preparación del profesor... Nuestro conocimiento de la historia académica de los alumnos y su pertenencia a un mismo centro facilita nuestra tarea de control de variables. La asignación aleatoria de los alumnos a los grupos (para evitar la influencia de otras posibles variables intervinientes) y la asignación aleatoria de los métodos a los grupos ya constituidos añade al estudio un control aleatorio que contribuirá a garantizar en mayor medida la obtención de inferencias precisas.

Aún así, y debido a la influencia que sobre la comprensión lectora (variable dependiente) puede tener, en la edad estudiada, la variable «aptitud verbal» de los alumnos, decidimos incluir como covariable la puntuación obtenida por los alumnos en el factor verbal de un test de inteligencia. Nuestra intención es eliminar de la puntuación en la variable dependiente, mediante la regresión, la parte que pudiera deberse a la covariable de forma que realicemos la comparación de medias entre los grupos (contrastando mediante el análisis de varianza una hipótesis nula general) en condiciones más precisas al haber reducido una posible fuente de contaminación en los resultados. En definitiva, podremos obtener una información más exacta del efecto de los tratamientos (métodos) sobre la variable dependiente (comprensión lectora); la obtención de esa información, no olvidemos, es el objetivo básico del estudio realizado.

Podemos entonces decir que en el Ancova el análisis de regresión contribuye a preparar los datos para realizar después el análisis de varianza en mejores condiciones metodológicas de control. Si la covariable no se elige adecuadamente (lo que probaremos mediante el contraste de la hipótesis nula Ho: $\beta = O$ o contraste del efecto de covariación entre X e Y) no se modificarán los datos y no habrá diferencia entre la realización de un Ancova y un análisis de varianza ordinario. A medida que la covariable tenga una mayor influencia sobre la variable dependiente habrá más diferencias producidas por aplicación de la regresión y el Ancova producirá un estadístico de contraste para los datos que puede llegar a modificar la interppretación de la hipótesis básica planteada (la hipótesis nula general sobre el efecto de los tratamientos, Ho: $\alpha_i = = O$); es decir, la presencia de la covariable puede llegar a producir una interpretación diferente del efecto de los tratamientos sobre la variable dependiente.

Los datos sobre los que vamos a trabajar se presentan en la tabla 1. Es un diseño de un único factor o variable independiente, con $n = 6$ unidades experimentales para cada uno de los $t = 4$ tratamientos.

TABLA 1

a_1		a_2		a_3		a_4	
X	Y	X	Y	X	Y	X	Y
2	7	22	8	21	14	14	7
20	11	25	12	35	17	13	9
9	9	30	12	37	19	16	11
3	5	21	10	22	18	10	13
20	7	14	11	25	19	14	10
16	8	10	8	28	15	17	8

¿Qué queremos decir exactamente con la expresión «eliminar el influjo» de la covariable (X) en la variable dependiente (Y) a partir de la regresión de Y sobre X? Como es sabido, en el estudio de la regresión de Y (variable criterio) sobre X (variable predictora) generamos dos nuevas variables:

a) Y, o variable de puntuaciones pronosticadas a partir del conocimiento de X y del valor de los parámetros a y b de la recta de regresión (en la regresión lineal simple) una vez que hemos convertido la relación estadística entre X e Y en una relación funcional.

b) (Y—Y'), o variable de puntuaciones de error de pronóstico; son los residuos de la regresión; es decir, lo que queda de la variable Y una vez que hemos eliminado, mediante regresión, la parte de la puntuación debida a su relación con X.

Pues bien, en el Ancova, eliminar de las puntuaciones Y el influjo de la covariable X mediante la regresión equivale a transformar las puntuaciones Y en unas nuevas puntuaciones residuales consecuencia de la consideración sucesiva de dos procesos de regresión implicados en los datos y que posteriormente nos van a proporcionar las sumas de cuadrados de dos de las fuentes de variación del análisis de varianza:

a) Un proceso relacionado con un valor b (pendiente de regresión) único para todos los sujetos del estudio; es decir, considerando a todos los sujetos como integrantes de un grupo único; llamaremos a este valor b_{total}. Las puntuaciones residuales que genera serán denominadas $(Y_{ij} — Y'_{ij})$.

b) Un proceso relacionado con un valor *b* que actuará de forma específica para los sujetos de cada uno de los grupos; llamaremos a este valor b_{intra}. Las puntuaciones residuales generadas serán denominadas $(Y_{ij} — Y'_{ij\text{-}g\text{-}})$.

Por tanto, en el Ancova las puntuaciones Y se transformarán, mediante dos procesos diferenciados de regresión, en dos tipos de puntuaciones residuales:

— $(Y_{ij} — Y'_{ij})$, que nos van a permitir calcular la suma de cuadrados total a utilizar en la tabla de análisis de varianza.

— $(Y_{ij} — Y'_{ij\text{-}g\text{-}})$, que nos permitirán calcular la suma de cuadrados del error (o suma de cuadrados intragrupo) para la tabla de análisis de varianza.

A estos planteamientos teóricos es a los que responde el primer procedimiento, que exponemos con referencia al ejemplo presentado.

2.1. *Primer procedimiento de resolución*

Calculamos en primer lugar el «coeficiente de regresión total» asociado a los datos de la tabla 1, considerando las puntuaciones como integrantes de un único grupo. Nos resulta:

$$b_{total} = \frac{\text{Cov } XY}{\text{Var } X} = \frac{\Sigma xy/n}{\Sigma x^2/n} = \frac{\Sigma xy}{\Sigma x^2} = \frac{600}{1816} = 0{,}33$$

A continuación obtenemos las puntuaciones pronosticadas para cada uno de los sujetos, Y'_{ij}, dadas por:

$Y'_{ij} = b_{total}\ (X_{ij} — X_{..}) + Y_{..}$

Los valores X.. e Y.. son, respectivamente, las medias de X e Y para el conjunto de sujetos del estudio.

Estas puntuaciones se presentan en la tabla 2.

TABLA 2.—*Puntuaciones pronosticadas Y'*

a_1	a_2	a_3	a_4
5'72	12'32	11'99	9'68
11'66	13'31	16'62	9'35
8'03	14'97	17'28	10'34
6'04	11'99	12'32	8'36
11'66	9'68	13'31	8'68
10'34	8'36	14'31	10'67

Las puntuaciones residuales (Y_{ij}—Y'_{ij}) se presentan en la tabla 3.

TABLA 3.—*Puntuaciones residuales (Y_{ij} - Y'_{ij})*

a_1	a_2	a_3	a_4
1'28	—4'32	2'01	—2'68
—0'66	—[illegible]1	0'38	—0'35
0'97	—2'97	1'72	0'66
—1'04	—1'99	5'68	4'64
—4'66	1'32	5'69	0'32
—2'34	—0'36	0'69	—2'67

La suma de cuadrados asociada a esta tabla es la «suma de cuadrados total ajustada» o «suma de cudrados total corregido el efecto de la regresión», que utilizaremos en el análisis de varianza; podemos calcularla a partir de:

$$SC_{total(ajus)} = \Sigma(Y_{ij} - Y'_{ij})^2 - \frac{[\Sigma(Y_{ij} - Y'_{ij})]^2}{n \cdot t} = \Sigma(Y_{ij} - Y'_{ij})^2 = 175{,}10$$

(hemos tenido en cuenta que la suma de los residuos, $\Sigma(Y_{ij}—Y'_{ij})$, es igual a cero).

Pasamos ahora a obtener el denominado «coeficiente de regresión intragrupo», que viene dado por:

$$b_{intra} = \frac{Exy}{Exx}$$

Donde los valores Exy y Exx vienen dados, respectivamente, por:

$$Exy = Sxy - Txy$$
$$Exx = Sxx - Txx$$

Siendo:

Sxy = Suma de productos total para X e Y =

$$= \sum^{t}\sum^{n} X_{ij}\, Y_{ij} - \frac{(\sum^{t}\sum^{n} X_{ij})\,(\sum^{t}\sum^{n} Y_{ij})}{t\, n}$$

Txy = Suma de productos de los tratamientos para X e Y =

$$= \sum^{t} \left[\frac{(\sum^{n} X_{ij})\,(\sum^{n} Y_{ij})}{n}\right] - \frac{(\sum^{t}\sum^{n} X_{ij})\,(\sum^{t}\sum^{n} Y_{ij})}{t\, n}$$

Sxx = Suma de cuadrados total para X =

$$= \sum^{t}\sum^{n} X_{ij}^{2} - \frac{(\sum^{t}\sum^{n} X_{ij})^{2}}{t\, n}$$

Txx = Suma de cuadrados de los tratamientos para x =

$$= \sum^{t} \left[\frac{(\sum X_{ij})^{2}}{n}\right] - \frac{(\sum^{t}\sum^{n} X_{ij})^{2}}{t\, n}$$

En nuestro ejemplo (ver tabla 7), tenemos:

Exy = 600 — 498'67 = 101'33.

Exx = 1816 — 963'33 = 852'67.

Por tanto:

$$b_{intra} = \frac{101'33}{852'67} = 0'12$$

Pasamos a obtener las puntuaciones pronosticadas para los sujetos en cada uno de los grupos, dadas por:

$$Y'_{ij(a)} = b_{intra}\,(X_{ij} - \bar{X}_{i}) + \bar{Y}_{i}$$

Los valores X_i e Y_i son, respectivamente, las medias de X e Y en cada uno de los grupos del estudio.

TABLA 4.—*Puntuaciones pronosticadas $Y'_{ij\cdot g\cdot}$*

a_1	a_2	a_3	a_4
6'67	10'37	16'16	9'67
8'83	10'73	17'84	9'55
7'51	11'33	18'08	9'91
6'79	10'25	16'28	9'19
8'83	9'41	16'64	9'67
8'35	8'93	17'00	10'03

Teniendo en cuenta los valores de las medias correspondientes en el ejemplo que seguimos, calculamos las puntuaciones $Y'_{ij\text{-}g\text{-}}$ (tabla 4). En la tabla 5 se incluyen los valores (Y_{ij} — $Y'_{ij\text{-}g\text{-}}$) o puntuaciones residuales resultantes del nuevo proceso de regresión considerado.

TABLA 5.—*Puntuaciones residuales (Y_{ij} - $Y'_{ij\cdot g\cdot}$)*

a_1	a_2	a_3	a_4
0'33	—2'37	—2'16	—2'67
2'17	1'27	—0'84	—0'55
1'49	0'67	0'92	1'09
—1'79	—0'25	1'72	3'81
—1'83	1'59	2'36	0'33
—0'35	—0'93	—2'00	—2'03

La suma de cuadrados asociada a esta tabla es la «suma de cuadrados de error ajustada» o «suma de cuadrados de error corregido el efecto de la regresión», que utilizaremos igualmente en el análisis de varianza; se calcula a partir de la fórmula:

$$SC_{error(ajus)} = \Sigma\ [Y_{ij} - Y'_{ij(g)}]^2 = 70{,}96$$

La «suma de cuadrados de los tratamientos ajustada por regresión» (suma de cuadrados de los tratamientos una vez corregido el efecto de la regresión), viene dada por:

Con los valores calculados de las sumas de cuadrados de todas las fuentes de variación incluidas en el diseño adoptado, podemos especificar la tabla correspondiente de análisis de varianza, que presentamos sin explicaciones añadidas (tabla 6).

TABLA 6

Fuentes de variación	Sumas de cuadrados	Grados de ...	Cuadrados medios	F
Tratamientos	104'14	3	34'71	9'29
Error experimental	70'96	19	3'73	
Total	175'10	22		

2.2. *Segundo procedimiento*

Exponemos el segundo procedimiento exclusivamente para facilitar la lector la comparación con el anterior, ya que se trata del procedimiento abreviado que presentan la práctica totalidad de los manuales al uso (Arnau, 1984; Tejedor, 1984; Amón, 1985; Sanmartín y Pedro, 1989).

Se procede calculando las sumas de cuadrados para X e Y y las sumas de productos para XY en las tres fuentes de variación (total, tratamientos y error):

— Sumas de cuadrados para X: Sxx, Txx y Exx.

— Sumas de cuadrados para Y: Syy, Tyy y Eyy.

— Sumas de productos para XY: Sxy, Txy y Exy.

Los resultados de la aplicación de las fórmulas correspondientes (Tejedor, 1984, pp. 197-199) para los datos del ejemplo que venimos siguiendo se presentan en la tabla 7).

TABLA 7

Fuentes de variación	Sumas de cuadrados de X	Sumas de producto de XY	Sumas de cuadrados de Y
Tratamientos	Txx = 963'33	Txy = 498'67	Tyy = 290'33
Error	Exx = 852'67	Exy = 101'33	Eyy = 83
Total	Sxx = 1816	Sxy = 600	Syy = 373'33

Las sumas de cuadrados ajustadas (una vez corregido el efecto de la regresión), obtenidas ahora a partir de los datos de la tabla 7, vienen dadas por:

$$SC_{total(ajus)} = Syy - \frac{Sxy^2}{Sxx} = 373,33 - \frac{600^2}{1816} = 175,10$$

$$SC_{error(ajus)} = Eyy - \frac{Exy^2}{Exx} = 83 - \frac{101,33^2}{852,67} = 70,96$$

$$SC_{trat(ajus)} = SC_{total(ajus)} - SC_{error(ajus)} = 175,10 - 70,96 = 104,14$$

Valores que coinciden con los obtenidos en el «primer procedimiento», que nos llevan a la tabla de análisis de varianza ya presentada (tabla 6) y que tendrán un tratamiento e interpretación semejante.

3. *Conclusión*

Hemos presentado dos procedimientos para corregir, a partir del análisis de regresión, el efecto de la covariable en un diseño de Ancova. Nuestra intención ha sido mostrarlos simultáneamente referidos a unos mismos datos para poder conectar el proceso de resolución seguido en el «procedimiento abreviado» con el proceso que sigue las pautas de los fundamentos estadísticos y metodológicos del Ancova.

Una vez eliminado el efecto de la regresión nos encontramos con la tabla de análisis de varianza (tabla 6), que habrá de ser tratada en la forma conocida. Calculado el valor F interpretaremos la hipótesis nula referida al efecto de los tratamientos (Ho: $\alpha_i = O$). En el ejemplo el valor F = 9'29 se interpretará con 3 y 19 grados de libertad. Como el valor tabular para $F_{3\,19\,0'05}$ = 3'13, rechazaremos la hipótesis nula y admitiremos que existen diferencias significativas entre los tratamientos; es decir, en las condiciones en que se ha realizado el estudio, los diferentes métodos de lectura seguidos generan diferencias en el comportamiento lector medio de los grupos. Una vez rechazada la hipótesis nula habría que proceder a determinar con exactitud entre qué grupos se producen diferencias significativas. Estos contrastes posteriores al rechazo de la hipótesis nula implican peculiaridades que debemos tener en cuenta a fin de que también las medidas que comparamos respondan al criterio de «medias ajustadas por regresión» (Tejedor, 1984, pp. 201, 211, 230).

¿Qué hubiera ocurrido en el estudio ficticio que estamos analizando si no se hubiera incorporado la covariable? Está claro que prescindiendo de las puntuaciones X en la tabla 1 nos encontraríamos ante un diseño de análisis de varianza de un factor, con cuatro tratamientos y seis unidades experimentales por tratamiento.

Las sumas de cuadrados correspondientes a ese diseño, únicamente referidas ahora a la variable dependiente Y, serían los valores Syy, Tyy y Eyy de la tabla 7. El contraste de la hipótesis referida al efecto de los tratamientos en la tabla de análisis de varianza correspondiente nos proporcionaría el valor F = 23'32 (tabla 8).

TABLA 8

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Cuadrados medios	F
Tratamientos	290'33	3	96'78	23'32
Error experimental	83	20	4'15	
Total	373'33	23		

Las diferencias entre los valores de la tabla 8 y los ofrecidos en la tabla 6 se deben al efecto de la covariable. Podemos preguntarnos si esas diferencias son importantes. La respuesta viene dada por el contraste de la hipótesis nula referida al coeficiente de regresión b_{intra} (Ho: β_{intra} = = O). Esta hipótesis se contrasta mediante el estadístico F, que se interpreta con $v_1 = 1$ y v_2 = nt-t-1 grados de libertad, dado por:

$$F = \frac{Exy^2\ (nt\text{-}t\text{-}1)}{(Exx\ Eyy — Exy^2)}$$

A medida que aumente el efecto de la covariable habrá mayor diferencia entre las sumas de cuadrados ajustadas y no ajustadas y, por tanto, en la interpretación de la hipótesis referida al efecto de los tratamiento. En el ejempplo que hemos expuesto el valor F, pasa de 23'32 a 9'29 al incluir la covariable; aunque los dos valores nos llevan a la misma interpretación de la hipótesis, es evidente que la influencia es notoria, lo que se constataría estadísticamente al comprpobar el rechazo de la hipótesis Ho: β_{intra} = O.

En el ejemplo que nosotros venimos exponiendo, y debido al reducido tamaño de las muestras, no podríamos rechazar las hipótesis Ho: β_{intra} = O a pesar de la importante diferencia encontrada entre las sumas de cuadrados corregidas y no corregidas y entre los correspondientes valores F. En efecto, vemos que por aplicación de la fórmula anteriormente expuesta:

$$F = \frac{101'33^2\ (24 — 4 — 1)}{(852'67 \times 83) — 101'33^2}$$

El valor tabular de $F_{1\,19\,0'05}$ = 4'38. Como el valor F obtenido (F = 3'22) es menor que el valor F tabular no podríamos rechazar la hipótesis nula y admitiríamos que el valor b_{intra} = 0'12 no es estadísticamente distinto de cero; es decir, la covariable no tiene una relación estadística significativa con la variable dependiente.

Cuando esa hipótesis pueda ser rechazada interpretaremos que la eliminación de la influencia de la covariable nos va a permitir analizar con más precisión los efectos que la variable independiente tiene sobre la dependiente; si se prefiere, podemos pensar que la inclusión de la covariable en el diseño contribuye a depurar el análisis de datos a realizar posibilitando la realización del contraste en mejores condiciones de control al eliminar la influencia de una variable interviniente (la covariable). Podríamos considerar, por tanto, en ese caso, un acierto la elección del análisis de covarianza frente a la alternativa del diseño correspondiente de análisis de varianza.

Dirección del autor: F. Javier Tejedor, Departamento de Métodos, Facultad de Filosofía y Ciencias de la Educación, Campus Universitario, 15706 Santiago de Compostela, La Coruña.

Fecha de recepción de la versión definitiva de este artículo: 8.III.1990.

BIBLIOGRAFIA

AMÓN, J. (1985) *Estadística para psicólogos* (Madrid, Pirámide).

ARNAU, J. (1984) *Diseños experimentales en psicología y educación* (México, Trillas).

BISQUERRA, R. (1989) *Introducción conceptual al análisis multivariable. Un enfoque informático con los paquetes SPSS-X, BMDP, LISREL y CPAD* (Barcelona, P.P.U.).

ETXEBERRÍA, J.; JOARISTI, L. y LIZASOAIN, L. (1990) *Programación y análisis estadísticos básicos con spss-pc(+)* (Madrid, Paraninfo).

HUITEMA, B. E. (1980) *The analysis of covariance and alternatives* (New York, Wiley).

MARTÍNEZ ARIAS, R. (1980) *Psicología Matemática II* (Madrid, UNED).

SANMARTÍN, R. y PARDO, A. (1989) *Psicoestadística. Contrastes paramétricos y no paramétricos* (Madrid, Pirámide).

TEJEDOR, F. J. (1979) El término de error experimental en los modelos estadísticos de análisis de varianza, *Revista Española de Pedagogía*, n. 145, pp. 77-113.

— (1984) *Análisis de varianza aplicada a la investigación en pedagogía y psicología* (Madrid, Anaya).

SUMMARY: REDUCTION OF THE COVARIATE EFFECT IN THE COVARIANCE ANALYSIS DESINGS.

In this article, we present two procedures followed in experimental research designs to eliminate the effect of the covariate before the hypothesis of mean difference is put to contrast between the differents treatments of the independent variable which are being subjected to analysis.

The first procedure follows the methodological process suggested by the statistical lay-outs underlying the Ancova. The second one is the short procedure which can generally be found in hand books. Its application does not permit to realize the progress of the implicit process. By comparing both procedures, the reader can understand much better the underlying assumptions in the Ancova.

KEY WORDS: Analysis of covariance. Analysis of variance. Covariate. Experimental designs. Statistics methods.