COMPARACION DE LA POTENCIA Y ROBUSTEZ DEL AMVAR DE MR FRENTE AL CORRESPONDIENTE MODELO MIXTO DEL AVAR CON DEPENDENCIA SERIAL EN EL ERROR

Guillermo VALLEJO y Paula FERNANDEZ
Facultad de Psicología - Universidad de Oviedo

RESUMEN

En el presente trabajo examinamos dos procedimientos de análisis apropiados a un diseño experimental aleatorio con varias medidas consecutivas a lo largo del tiempo para cada una de las unidades experimentales. Un enfoque para el análisis de tales datos consiste en aplicar el tradicional modelo mixto del AVAR con la estructura del error modelada a través de procesos ARMA. Alternativamente, el problema también puede ser considerado desde una perspectiva más general haciendo uso del enfoque multivariado de medidas repetidas. Datos simulados mediante procedimientos de Monte Carlo son empleados para invetigar el efecto que el incumplimiento de las asunciones subyacentes a los modelos tiene sobre el grado de sesgadez de los parámetros estimados, sobre la probabilidad empírica de cometer errores Tipo I y sobre la potencia de prueba de los test estadísticos para cada uno de los dos procedimientos.

Palabras clave: Medidas repetidas, errores con dependencia serial.

ABSTRACT

Comparison of the power and robustness of AMVAR of repeated measures versus the correspondent mixed model of AVAR with the error including serial dependence. In this research two procedures of analysis for a random experimental desing with several consecutive measures for each experimental unit are analized and composed. To analize this kind of data a traditional mixed model of AVAR with the error structure modeled through ARMA processes can be used. The problem can be worked out as well from a more general perspective, using a multivariate repeated measures approach. Both approaches are compared using simulated data obtaining with Monte Carlo procedures. Special attention was paid to a) the robustness of model parameter estimates when some assumtions are violated, b) the power of statistical test for both procedures.

Key words: Repeated measures. Errors with serial dependence

El objetivo primordial, de cualquier diseño experimental tiene que ver con el hecho de eliminar los sesgos sistemáticos y de minimizar la varianza del error a la hora de establecer relaciones causales entre las variables, o lo que es lo mismo, a la hora de intentar poner de manifiesto la posible acción de las variables experimentales o tratamientos sobre la conducta. Para lograr este objetivo, esencialmente dos tipos de diseños están disponibles para los investigadores. Por un lado, aquellos que emplean muestras independientes y, por otro lado, aquellos que emplean muestras correlacionadas. Obviamente, cada uno tiene sus propios méritos.

El método básico para lograr controlar los sesgos sistemáticos provinientes de los sujetos en los diseños que emplean muestras independientes se logra asignando aleatoriamente los sujetos a los diferentes niveles de tratamiento. Desafortunadamente, para que la aleatorización sea efectiva, o bien disponemos de sujetos muy homogéneos, o bien hacemos uso de tamaños de muestra elevados. El requerimiento de que el número de sujetos sea alto se hace aún más crítico cuando los niveles o combinaciones resultantes de éstos son numerosos. En estas circunstancias los diseños de medidas repetidas representan una alternativa viable a los diseños de corte transversal; pues, además de incrementar la validez de conclusión estadística y la validez externa, reducen drásticamente el tamaño de muestra. Por este motivo no es de extrañar que este tipo de diseños sean los más comunes en la investigación psicológica actual (Edgington, 1974; Keselman y Keselman, 1990).

En los últimos veinte años, los investigadores han enfocado su atención en determinar cual de las estrategias analíticas resulta más apropiada para este tipo de diseños (Andersen y otros, 1981; Barcikowski y Robey, 1984; Cole y Grizzle, 1966; Collier y otros, 1967; Davidson, 1972; Diggle, 1988; Hynh, 1978; Huynh y Feldt, 1976; Jones, 1987, Kenward, 1987; Keselman y Keslman, 1990; Lewis y van Knippenberg, 1984; Maxwell y Avey, 1982; O`Brien y Kaiser, 1985; Pantulla y Pollock, 1985; Rochon y Helms, 1989; Schaalje, Zhan, Pantulla y Pollock, 1991; Schluchter, 1988; Vallejo y Fernández, 1990;1992; Vallejo, 1991; Wallenstein y Fleiss, 1980). Esencialmente, podemos afirmar que los datos registrados bajo alguna de las variedades de los diseños de medidas repetidas pueden ser analizados bien mediante algún enfoque univariado o bien multivariadamente. Logicamente, la elección de una técnica analítica u otra va a depender en buena medida de la naturaleza de las asunciones que el investigador pueda efectuar acerca de sus datos; yendo más allá, podemos decir que nuestra elección deberá estar subordinada a la forma de la matriz de varianza-covarianza subyacente a nuestros datos.

Bajo el enfoque univariado las diferentes medidas repetidas representan diferentes niveles de la variable independiente. Por tanto, en cada nivel una observación de la misma variable dependiente es tomada, de este modo la escala de medida es la misma a lo largo de todos los niveles. Lo que cambia es la magnitud de la variable dependiente, no la escala de medida.

Así pues, dado que en este procedimiento todas las medidas son registradas bajo la misma escala, la covarianza muestral a lo largo de los q niveles puede ser promediada, y de hecho así ocurre a la hora de decidir sobre las hipótesis nulas asociadas con las medidas repetidas. Ahora bien, dado que esto es lo que sucede con el AVAR, para aplicar este modelo debemos de hacer algunas asunciones acerca de la forma de las matrices de varianza-covarianza de la población.

Geisser y Greenhouse (1958) y Greenhouse y Geisser (1959) al extender el trabajo inicialmente desarrollado por Box (1954) con diseño de medidas repetidas de una sóla muestra, a diseños de muestras divididas, pusieron de manifiesto que si la matriz de varianza-covarianza tiene una forma arbitraria (se asume que las matrices grupales constituyen muestras al azar de una misma población) los distintos estadísticos asociados con las medidas repetidas se siguen distribuyendo de acuerdo con la F ordinaria, pero con los grados de libertad corregidos en función de la desviación de la matriz de varianza-covarianza del patrón de uniformidad requerido. Si esta matriz tiene estructura de simetría combinada, entonces la razón F empírica se distribuye como la F tabulada y, por tanto, el factor de corrección ε es igual a uno. Este hecho es de capital importancia tenerle presente, ya que conforme nos desviamos de la condición de simetría las razones F se vuelven cada vez más liberales incrementándose de este modo, las posibilidades de capitalizar sobre el azar.

Es decir, cuando la asunción de uniformidad no se cumple, la prueba F está positivamente sesgada, lo cual conlleva a rechazar más veces de las requeridas hipótesis nulas de acuerdo con el nivel de significación precisado. Sin embargo, la verdadera distribución del estadístico F con sus respectivos grados de libertad, podrá ser aproximada mediante el usual estadístico F con los grados de libertad ajustados mediante el parámetro ε, el cual como hemos dicho depende del grado de heterogeneidad que presenta la matriz de varianza-covarianza.

Si nos detenemos a examinar un poco más sustantivamente el contenido de la matriz de varianza-covarianza vemos, por un lado, que ésta al igual que cualquier otra matriz subyacente a un modelo mixto del AVAR incorpora dentro de sí las desviaciones resultantes de sustraer las puntuaciones de cada uno de los sujetos de las respectivas medias en los diferentes niveles de tratamiento; desviaciones que incluyen tanto la variación submuestral dentro de las unidades experimentales (errores de medida asociados con las distintas observaciones), como la variación aleatoria en el promedio de las respuestas de los diferentes sujetos asignados a los p niveles de tratamiento (errores aleatorios asociados con las unidades experimentales). Y, por otro lado, que para cada sujeto se asume que los errores son independientes a lo largo de los distintos puntos de observación. Este hecho se deriva directamente del cumplimiento de la condición de simetría combinada, lo cual como es sabido implica que las variables aleatorias del modelo están igualmente correlacionadas para todos los pares de observaciones de un mismo sujeto y tienen varianza constante.

Durante bastantes años se pensó que la única matriz que satisfacia el requisito de ε =1 era la de simetría combinada, sin embargo, Huynh y Feldt (1970) y Rouanet y Lépine (1970) han demostrado independientemente que la asunción de uniformidad es tan sólo una condición suficiente, no una condición necesaria para la validez de la razón F univariada en los diseños de medidas repetidas.

Por ejemplo, Huynh y Feldt (1970) han demostrado que la razón F también se distribuye como F si las diferencias entre las varianzas de cualquier par de tratamientos son iguales. Cuando el supuesto de simetría está presente, es obvio que lo anterior se cumple, dado que la matriz de simetría es un caso especial de una condición más general que presentan las matrices en las cuales se mantiene la asunción de igualdad en las diferencias de las varianzas entre dos cualesquiera tratamientos. De este modo, la igualdad de las varianzas y de las covarianzas no es requerido para que el estadístico F se comporte conforme a la distribución F, la condición necesaria y suficiente para la validez de la razón F es que la matriz de varianzas-covarianzas sea circular o esférica (esto es, que los vectores de diferencias entre dos cualesquiera pares de tratamiento tengan la misma varianza).

Esta situación donde la varianza no depende del intervalo temporal entre las observaciones, puesto que todas las posibles diferencias entre dos cualesquiera tratamientos tienen la misma varianza, si bien es menos restrictiva que la situación de simetría no por ello deja de ser rebuscada a la hora de modelar estructuralmente las respuestas de los sujetos. Sobre todo, en aquellas situaciones en las cuales la secuencia de administración de los niveles de la variable tratamiento a los np bloques que conforman la variable entre grupos no se contrabalancean, bien sea porque los propios objetivos de investigación así lo requieran o, bien sea porque la propia naturaleza de la variable imposibilite al investigador para manipularla intencionalmente.

Así pues, satisfacer la condición de simetría nos parece irreal y cumplir la condición de circularidad es difícil de lograr bajo muchas condiciones, como ocurre por ejemplo, con los datos obtenidos en investigaciones de cáracter longitudinal en los campos clínicos, educativos e industriales donde la dependencia serial entre las observaciones tomadas desde una misma unidad de análisis en distintos momentos o intervalos temporales puede llegar a ser de considerable importancia. Obviamente, en estas situaciones un investigador no debe esperar que sus datos conformen el tipo de matriz estructural esférica requerida por el usual modelo del AVAR con valores esperados mixtos definido con anterioridad. Por supuesto, nada más apartado de la realidad, que esperar que las correlaciones entre todos los pares de medida repetida sean uniformes. Claramente, como apuntan Kogan y otros (1979) uno esperaría que las observaciones llevadas a cabo en tiempos sucesivos o contiguos estén más altamente correlacionadas que las de los registros que no son adyacentes; es decir, que la correlación entre las puntuaciones decrece a medida que incrementamos los intervalos temporales.

Tradicionalmente, para hacer frente al incumplimiento de la asunción de circularidad multimuestral los procedimientos de F conservadora y de F ajustada propuestos por Geisser y Greenhouse (1958) y por Greenhouse y Geisser (1959) son utilizados. Sin embargo, el problema con estas pruebas es que son excesivamente conservadoras, pues fueron desarrolladas para satisfacer bandas seguras más aproximadas (Jones, 1987). Alternativamente, Cole y Grizzle (1966) partiendo del hecho de que las observaciones obtenidas con diseños de parcela dividida están correlacionadas y, por tanto, son esencialmente de naturaleza multivariada han sugerido que el procedimiento adecuado para analizar tales diseños es el multivariado de medidas repetidas. Y, si bien es cierto que el enfoque multivariado permite a la matriz de varianzas-covarianzas tener cualquier estructura a la hora de contrastar los efectos principales, interacciones, efectos simples y subefectos, no lo es menos que este procedimiento, además de requerir que el número de sujetos sea mayor que el número de observaciones (para hacer inferencias acerca de los parámetros desconocidos, la matriz, además de arbitraria, debe ser simétrica y definida positiva), está sobreparametrizado (ya que utiliza todos los elementos contenidos en la matriz de varianzas-covarianzas), es menos parsimonioso y tiene dificultades a la hora de incorporar observaciones perdidas (Kenward, 1987). Además, tampoco conviene perder de vista los resultados de las investigaciones efectuadas por Davidson (1972) y Collier y otros (1967). Pues, mientras que en el trabajo de Collier y otros se comprobó como el valor de ε es un potente determinante del sesgo potencial que puede acaecer en el nivel de significación cuando calculamos la prueba de F ordinaria, F ajustada y F conservadora, en el de Davidson (1972) se constató la influencia decisiva que sobre la potencia de la T^2 de Hotelling tenía la diferencia entre el número de sujetos (n) y el número de niveles de la variable de tratamiento (q). En concreto, Vitalino (1982) señala que la potencia del procedimiento multivariado es superior a la del correspondiente univariado de medidas repetidas tan sólo cuando $n\text{-}q > 20$ y ε se encuentra por debajo de 0,5 ($1/q\text{-}1 \leq \varepsilon \leq 0.45$).

Por tanto, consideramos de capital importancia tener presente algún procedimiento que modele la correlación serial de los datos, de modo que si ésta resultase significativa pueda ser removida de las respuestas efectuadas por los sujetos. Si bien en la actualidad existen abundantes modelos a la hora de describir la estructura correlacional de la matriz (ver Box-Jenkins, 1976), nosotros en la presente investigación nos centraremos en situaciones donde los individuos son observados sucesivamente sobre el tiempo y consideraremos el procedimiento de estimación mínimo cuadrático generalizado con la estructura del error modelada a través de procesos autorregresivos de primer orden. Este enfoque se puede generalizar a una amplia variedad de modelos a la hora de describir la dependencia serial de la matriz de varianza-covarianza entre las medidas repetidas y depende de un reducido número de parámetros; por lo general no más de cuatro, si bien lo más frecuente es que la matriz sea una función del componente de varianza asociado con la kth medida del ith sujeto, del componente de varianza asociado con el factor aleatorio sujeto y del coeficiente del modelo de dependencia serial implicado.

La idea de incorporar la correlación serial dentro del modelo no es original, pues de una forma u otra ha estado presente con anterioridad en los trabajos clásicos de Box (1954), Guttman (1955) y Danford y otros (1960); sin embargo, hasta fechas recientes no se ha efectuado un modelamiento explícito de la estructura correlacional de la matriz tal y como ponen de manifiesto los trabajos de Diggle (1988); Pantulla y Pollock (1985) y Röchom y Helms (1989). A nuestro modo de ver esta forma de proceder constituye un ajuste razonable entre la situación en exceso restrictiva donde se asume que la matriz tiene una estructura correlacional que cumple con la condición de esfericidad y la situación en exceso liberal donde se permite a la matriz tener cualquier estructura.

Así pues, a raíz de lo expuesto hasta aquí un problema que nos surge de inmediato, y que nos parece de capital importancia, es investigar el comportamiento del enfoque multivariado de medidas repetidas frente al correspondiente univariado, pero con la estructura del error modelada mediante algún proceso ARMA. En relación con la probabilidad de cometer errores Tipo I y en relación con la potencia de prueba de los tests estadísticos para cada uno de los dos procedimientos de análisis.

METODO

En orden a evaluar el objetivo expuesto, las propiedades de un modelo mixto del AVAR con errores correlacionados y las propiedades de un AMVAR de medidas repetidas serán investigadas por medio de datos simulados bajo diferentes situaciones. Las hipótesis a comparar son las referidas a un diseño de medidas parcialmente repetidas (2x8) con ocho vectores de observaciones en cada uno de los dos diferentes grupos. Para alcanzar este objetivo, desde distribuciones multivariadas normales extrajimos múltiples conjuntos de vectores pseudoaleatorios $\mathbf{y}'_{ij}$ $[y_{ij1}, y_{ij2}, ...y_{ijk}]$ con vector de medias μ'_j. $[\mu_{j1}, \mu_{j2} ..., \mu_{jk}]$ y matriz de varianza-covarianza Σ.

El procedimiento seguido en la obtención de los vectores pesudoaleatorios μ'_{ij} se efectuó en tres fases. En la primera haciendo uso del método congruencial multiplicativo descrito por Naylor y otros (1966) generamos vectores de variables aleatorias independientes y uniformemente distribuidas entre cero y uno. A continuación haciendo uso del método de Teichroew descrito por Knuth (1969) las variadas uniformes fueron convertidas en vectores de variadas normales $\mathbf{z}'_{ij}$ $[z_{ij1}, z_{ij2},..., z_{ijk}]$. Finalmente, los vectores de variadas $\mathbf{z}_{ij}$. fueron transformados en vectores de observaciones $\mathbf{y}_{ij}$ por medio de la ecuación $\mathbf{y}'_{ij} = \mathbf{T}\mathbf{z}_{ij} + \mu$, donde $\mathbf{T}$ es la factorización de Cholesky de Σ y e $y_{ij} \sim N(\mu, \Sigma)$. Tanto la precisión del procedimiento de normalización, como el ajuste de las matrices de varianzas-covarianzas muestrales a la poblacional haciendo uso de la prueba Chi-cuadrado descrita por Anderson (1958, pp. 264-267) resultaron completamente satisfactorios.

En la tabla 1 presentamos las matrices de varianza-covarianza utilizadas en este trabajo. La primera matriz sigue un proceso Markov, en concreto ha sido construida desde,

$$\Sigma = \sigma_{\varepsilon}^{2} (\Omega - 1/ q\, j) + (\sigma_{\varepsilon}^{2} + q\sigma_{\pi}^{2})$$
$$1/ q\, j(\Phi = 0.8;\ \sigma_{e}^{2} = 50;\ \sigma_{v}^{2} = 72)$$

Aquí se asume que el proceso AR es estacionario e invertible. Como consecuencia inmediata de la estacionaridad la covarianza entre dos cualesquiera términos de error depende únicamente de la distancia temporal entre ellos. De este modo todos los elementos a lo largo de cualquier diagonal de la matriz Σ toman los mismos valores, los cuales configuran una matriz conocida con el nombre Toeplitz. Así mismo, en la Tabla 1 también presentamos una matriz en la cual, además de darse heterocedasticidad sobre el tiempo, los valores situados en las diagonales secundarias también difieren sensiblemente. Los datos generados a partir de estas matrices nos van a permitir comparar las dos técnicas analíticas aludidas, así como el comportamiento del AVAR de medidas repetidas con la estructura del error modelada mediante algún proceso ARMA cuando la propiedad de igualdad de las diferentes diagonales de la matriz Σ es parcial o totalmente incumplida.

Para determinar cómo afecta a la aceptación o rechazo incorrecto de las hipótesis referidas a las medidas repetidas cuando se incumplen tanto los supuestos referidos a la forma de distribución, como a la forma de la matriz Σ en la Tabla 2 presentamos los vectores de medias μ utilizadas.

El diseño fue analizado por cada uno de los dos procedimientos analíticos bajo cada una de las siguientes condiciones:

a) Homocedasticidad más correlación en la estructura del error.
b) Heterocedasticidad y correlación arbitraria.

Bajo cada una de estas dos condiciones efectuaremos comparaciones en las siguientes tres áreas

1 Tasas de error Tipo I ,
2 Tasas de error Tipo II ,
3 Precisión de las estimaciones efectuadas.

Medidas empíricas de la probabilidad de cometer errores Tipo I serán obtenidas tabulando el número de veces que cada estadístico excede su valor crítico cuando las diferencias en el vector medias es nula y dividiendo por el número de pruebas efectuadas. La probabilidad de cometer errores Tipo II son derivadas registrando el número de veces que las hipótesis nulas son indebidamente aceptadas al nivel α especificando y dividiendo por el número de contrastes realizados. Para estimar el grado de sesgo presente en los efectos del diseño, el valor medio de los parámetros estimados es computado. Si el estimador es insesgado, entonces el valor promedio se aproxima estrechamente al verdadero valor del parámetro poblacional. Similarmente, conclusiones relativas a la eficacia de los diferentes procedimientos de estimación pueden ser extraídas comparando los respectivos errores estándar.

Para finalizar este apartado apuntar tan sólo lo siguiente: dado que los datos han sido simulados desde poblaciones conocidas y que cada conjunto de observaciones será analizado mediante los dos procedimientos, contamos con una situación óptima para determinar el grado de robustez de ambos enfoques. De este modo, si un procedimiento es robusto con respecto a sus asunciones los resultados esperados se obtendrán aunque éstas no se hayan cumplido. Por el contrario, si los supuestos son críticos y el enfoque no tolera desviarse de ellas los resultados diferirán de las conclusiones esperadas.

RESULTADOS

Estimaciones empíricas de la probabilidad de cometer errores Tipo I y Tipo II para cada uno de los dos procedimientos de análi-

TABLA 1.- Matrices de varianza covarianza utilizadas en nuestro estudio de simulación junto con sus correspondientes descomposiciones de Cholesky

$$\Sigma_U = \begin{bmatrix} \mathbf{122} & 112 & 104 & 98 & 93 & 88 & 85 & 82 \\ & \mathbf{122} & 112 & 104 & 98 & 83 & 88 & 85 \\ & & \mathbf{122} & 112 & 104 & 98 & 83 & 88 \\ & & & \mathbf{122} & 112 & 104 & 98 & 93 \\ & & & & \mathbf{122} & 112 & 104 & 98 \\ & & & & & \mathbf{122} & 112 & 104 \\ & & & & & & \mathbf{122} & 112 \\ & & & & & & & \mathbf{122} \end{bmatrix}$$

$$\Sigma_M = \begin{bmatrix} \mathbf{120} & 103 & 95 & 91 & 110 & 92 & 110 & 84 \\ & \mathbf{114} & 97 & 100 & 116 & 102 & 90 & 80 \\ & & \mathbf{110} & 104 & 111 & 100 & 101 & 98 \\ & & & \mathbf{125} & 120 & 110 & 96 & 103 \\ & & & & \mathbf{129} & 112 & 107 & 104 \\ & & & & & \mathbf{125} & 75 & 79 \\ & & & & & & \mathbf{131} & 96 \\ & & & & & & & \mathbf{122} \end{bmatrix}$$

MATRIZ UNIVARIADA

11.0454	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10.1400	4.37954	0.0	0.0	0.0	0.0	0.0	0.0
9.415712	3.77314	4.37123	0.0	0.0	0.0	0.0	0.0
8.8725	3.20417	3.74476	4.35761	0.0	0.0	0.0	0.0
8.41982	2.88226	3.16756	3.71719	4.35295	0.0	0.0	0.0
7.96715	2.78868	2.85081	3.14404	3.71317	4.35295	0.0	0.0
7.69554	2.27586	2.73466	2.79704	3.12109	3.71305	4.33236	0.0
7.42391	2.21198	2.23101	2.68243	2.77475	3.12134	3.68871	4.33255

MATRIZ MULTIVARIADA

10.9545	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9.40257	5.05882	0.0	0.0	0.0	0.0	0.0	0.0
8.467227	3.05572	5.045222	0.0	0.0	0.0	0.0	0.0
8.30712	4.32742	3.71340	4.84517	0.0	0.0	0.0	0.0
10.0416	4.26648	2.15640	2.08721	0.978439	0.0	0.0	0.0
8.39841	4.55318	2.62679	2.22388	-2.11110	4.175236	0.0	0.0
10.0416	-0.873063	3.28717	0.857526	1.03560	-3.28467	2.4499	0.0
7.66812	1.56163	5.29771	2.65618	3.44361	-1.21298	-2.80798	2.10375

Tabla 1: Matrices de varianza covarianza utilizadas en nuestro estudio de simulación junto con sus correspondientes descomposiciones de Cholesky

Casos	μ'_j	q_1	q_2	q_3	q_4	q_5	q_6	q_7	q_8				Potencia de Prueba: Sin Correg. Esferic.	Corrigien. Esferic.	Multiv.
A	**Valores de los vectores de medias de la población bajo H_o**														
	μ'_1	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	$SC_A=0$		—	—	
	μ'_2	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	$SC_{S(A)}=24.555$				
											$SC_B=0$		—	—	
		4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	$SC_{AB}=0$		—	—	
											$SC_{BXS(A)}=4725$				
B	**Valores de los vectores de medias de la población bajo H_I Modelo aditivo**														
	μ'_1	1	2	3	4	5	6	7	8	4.5	$SC_A=0$		0.65	—	
	μ'_2	1	2	3	4	5	6	7	8	4.5	$SC_{S(A)}=24.555$				
											$SC_B=1344$	$\alpha=0.05$	0.99	0.66	0.60
		1	2	3	4	5	6	7	8	4.5	$SC_{AB}=0$	$\alpha=0.10$	0.99	0.91	0.89
											$SC_{BXS(A)}=4725$				

Tabla 2: Valores numéricos de los vectores de parámetros de medias poblacionales y de las potencias de prueba

sis (univariado con la correlación corregida y multivariado de medidas repetidas) a través de ocho poblaciones simuladas son expuestas en las Tablas 3, 4 y 5. Los errores estándar reportados para las estimaciones empíricas de α y β y son calculadas desde,

$$\sqrt{pxq/m}$$

donde p es la probabilidad teórica del error Tipo I o Tipo q es el complementario de p y m es el número de experimentos efectuados.

Cuando las asunciones del modelo univariado con la correlación corregida son efectuadas (Caso I-A) y no hay efectos ni en la parte entre (Tratamiento A) ni en la parte intra (Tratamiento B e Interacción AxB) el modelo multivariado de medidas repetidas (en adelante AMVAR de MR) es ligeramente más liberal que el modelo mixto del AVAR al nivel $\alpha = 0.05$. Sin embargo, cuando el nivel de significación es el $\alpha = 0.01$ los dos procedimientos analíticos manifiestan una tasa de rechazo similar. A su vez, cuando Σ es consistente con las asunciones subyacentes al AMVAR de MR existen sólamente ligeras diferencias empíricas entre los dos modelos tanto al $\alpha = 0.05$, como al $\alpha = 0.01$.

Por otro lado, cuando los efectos de las medidas repetidas están presentes, en nuestro caso sólamente el efecto del tratamiento B (modelo aditivo), Casos I-B y II-B, el modelo mixto del AVAR con la correlación corregida es siempre más poderoso a la hora de detectar los efectos del tratamiento que el AMVAR de MR (ver Tabla 5). Sin embargo, en este punto debemos de hacer una consideración, cuando los datos han sido generados desde una matriz que presenta correlación (Caso I-B), el modelo mixto es sustancialmente más poderoso que el AMVAR de MR tanto al $\alpha = 0.05$ como al $\alpha = 0.10$; ahora bien, cuando los datos han sido generados desde una matriz que satisface las asunciones del modelo multivariado, las diferencias se dan aislamente al $\alpha = 0.05$.

Las estimaciones empíricas de la potencia pueden ser comparadas con los valores teóricos expuestos en la Tabla 2. Para el caso I-B la potencia empírica de ambos modelos es menor de la predicha, siendo ésta mucho más reducida en el caso del AMVAR de MR. Para el caso II-B la potencia empírica se aproxima extrechamente al valor teórico.

Finalmente, queremos reseñar que cuando las estimaciones empíricas son comparadas con los valores conocidos de la población, ambos procedimientos cuentan con idénticas estimaciones de las medidas repetidas. Lo cual, obviamente, no es de extrañar dado que tanto el modelo mixto del AVAR como del AMVAR de MR son procedimientos de estimación mínimo cuadráticos para estimar los efectos de las medidas repetidas y prueban las hipótesis por medio de la razón F. Sin embargo, la formación de la razón F es diferente. Para el modelo mixto del AVAR, la razón F es calculada desde las apropiadas sumas de cuadrados y grados de libertad descritos con anterioridad en la Introducción. Por su parte, en el AMVAR de MR las sumas de cuadrados son remplazadas por las matrices sumas de cuadrados y productos cruzados correspondientes a la hipótesis y al error, las cuales son utilizadas para calcular la razón F multivariada.

CONCLUSIONES

Los datos indican que con respecto a las tasas de error Tipo I y estimación de los efectos de medidas repetidas, ambos procedimientos arrojan resultados similares. Los resultados expuestos en la Tabla 3 ponen de manifiesto que los estimadores empíricos de α son más pequeños que los valores esperados; implicando que ambos procedimientos son conservadores bajo las asunciones especificadas. Con todo, el hecho de que las estimaciones empíricas de α se hallen compren-

Casos para Σ^*	Modelos de Análisis	*P* (ERROR TIPO I) Caso A		*F* AJUSTADA (tratamiento B)	
		α=0.05	α=0.01	α=0.05	α=0.01
I	AVAR	0.02	—*	—	
	MANOVA de MR	0.04	0.01		
		ES=0.0217	ES=0.0099		
II	AVAR	0.11*	0.03	0.05	0.02
	MANOVA de MR	0.06	0.01	0.06	0.01
		ES=0.0217	ES=0.0099	ES=0.02117	ES=0.0099

* Los datos de la matriz Σ son los definidos en la Tabla 1

N=100 Experimentos

Tabla 3: Estimación empírica de la tasa de error Tipo I en ausencia de efecto de tratamiento en la parte intra

Tabla 4. Estimación empírica de la Tasa de error Tipo I en ausencia de efectos significativos de la interacción

Casos para Σ *	Modelos de Análisis	P (ERROR TIPO I) Modelo Aditivo (interacción)		
		α=0.01	α=0.05	α=0.10
I	AVAR	0.008	0.044	0.078
	MANOVA de MR	0.002	0.040	0.106
		ES=0.004	ES=0.0099	ES=0.013
II	AVAR	0.02	0.07	0.10
	MANOVA de MR	0.02	0.06	0.10
		ES=0.004	ES=0.0099	ES=0.013

* Los datos de la matriz Σ son los definidos en la Tabla 1

N=500 Experimentos

Tabla 5. Estimación empírica de la potencia de prueba

Casos para Σ *	Modelos de Análisis	**Potencia de Prueba** **Modelo Aditivo (B efectos)**	
		α=0.05	α=0.10
I	AVAR	0.626	0.862
	MANOVA de MR	0.381*	0.598*
		ES=0.0211	ES=0.0134
II	AVAR	0.811	0.942*
	MANOVA de MR	0.702	0.886
		ES=0.0211	ES=0.0134

* Los datos de la matriz Σ son los definidos en la Tabla I

N=500 Experimentos

didas entre ± 2 se sugiere que la tendencia conservadora aludida se debe tan sólo a fluctuaciones del azar.

Por otra parte, cuando los efectos de los tratamientos están presentes, en nuestro caso el efecto de la variable B, el AVAR de medidas repetidas con la correlación corregida es más poderoso. Esto ocurre tanto en el caso I-B como en el caso II-B. Con todo, conviene reparar que para el caso II-B la diferencia es tan sólo un mero artefacto del procedimiento de análisis, pues al tener la matriz multivariada correlación cero y falta de esfericidad ($\alpha = 0$ y $\varepsilon \approx 0.47$), el modelo multivariado tiene en cuenta este hecho, mientras que el univariado se comporta como el usual modelo del AVAR sin corregir los grados de libertad por la falta de esfericidad.

Por último, apuntar que las generalizaciones de este estudio son limitadas por el diseño, valor de los parámetros utilizados (μ, Σ, α, β, ϕ y ε), tamaño de muestra, así como la forma de distribución subyacente que en este estudio es normal multivariada.

REFERENCIAS

Andersen, A.H.; Jensen, E.B. y Geart, S. (1981). Two-way analysis of with correlated errors. *International Statistical 49*, 153-169.

Anderson, T.W. (1958). *An Introduction Multivariate Statistical Analysis.* New York:Wiley and Sons, Inc...

Barcikowski, R.S y Robey, R.R. (1984). Decisions in single group repeated measures analysis: Statistical test and tree computer packages. *The American Statistician, 38*, 148-151.

Box, G. E. P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems II. Effects of inequality of variance and of correlation between errors in the two-way classification. *Annals of Mathematical Statistics, 25*, 484-498.

Box, G. E. P. y Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control, Edición revisada.* San Francisco. Holden-Day.

Cole, V. W. L. y Grizzle, J. E. (1966). Applications of multivariate analysis of variance to repeated measures experiments. *Biometrics 41*, 505-514.

Collier, R. O., Baker, F. B., Mandeville, G. K. y Hayes, T. F. (1967). Estimates of tests size for several test procedures based on conventional variance ratios in the repeated measures design. *Psychometrika, 32*, 339-353.

Danford, M. B., Hughes, H. M. y McNee, R. C. (1960). On the analysis of repeated measures experiments. *Biometrics, 16*, 547-565. ,

Davidson, M. L. (1972). Univariate versus multivariate test in repeated measures experiments. *Psychological Bulletin, 77*, 446-452. ,

Diggle, P. J. (1988). An approach to the analysis of repeated measurements. *Biometrics, 16*, 959-971.

Edgington, E.S. (1974). A new tabulation of statistical procedures used in APA Journals. *American Psychologist, 29*, 25-26.

Geisser, S. y Greenhouse, S. W. (1958). An extension of Box's result on the use of the F distribution in multivariate analyisis. *Annals of Mathematical Statistical, 29*, 885-891.

Greenhouse, S. W. y Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*, 95-102.

Guttman, L. G. (1955). An generalized simplex for factor analysis. *Psychometrika, 20*, 173-192.

Huynh, H. (1978). Some approximate test for repeated measurement designs. *Psychometrika, 43*, 161-175.

Huynh, H. y Feldt. (1976). Estimation of the Box correction for degrees of freedon from sample data in the randomized block an Split-plot designs. *Journal of Educational Statistics, 1*, 1582-1589.

Huynh, H. y Feldt, L. S. (1970). Conditions under which mean square ratios in repeated measurements designs have exact F-distributions. *Journal of the American Statistical Associations, 65*, 1582-1585.

Jones, R. H. (1987). Serial correlation in unbalanced mixed models. *Bulletin of the International Statistical Institute, 46*, 105-122.

Kenward, M. G. (1987).A method for comparing profiles of repeated measurements. *Applied Statistical, 36*, 296-308.

Keselman, J.C. y Keselman, H.J. (1990). Analysing unbalanced repeated measures designs. *British Journal of Mathematical and Statistical Psychology, 43*, 265-282.

Knuth, D.E. (1969). *Seminumerical Algorithms: The art of Computer Prograaming*. Massachusetts: Addison-Wesley.

Kogan, C. J., Keselman, H. J. y Mendoza, J. L. (1979). Analysis of repeated measurements. *British Journal of Mathematical and Statistical Psychology, 32*, 269-286.

Lewis, c. y van Knippnberg, C. (1984). Estimation and model comparisons for repeated measures data. *Psychological Bulletin, 96*, 182-194.

Maxwell, S.E. y Arvey, R.D. (1982). Small sample profile analysis with many variables. *Psychological Bulletin, 92*, 778-785.

Naylor, H.T.; Balintfy, J.L.; Burdick, D.S. y Chu, K. (1966). *Computer Simulation Techniques*. New York: John Wiley.

O'Brien, K.G. y Kaiser, M.K. (1985) MANOVA method for analyzing repeated measures designs: An extensive primer. *Psychological Bulletin 97*, 316-333.

Pantulla, S. G. y Pollock, K. M. (1985). Nested analysis of variance with autocorreleted errors. *Biometrics, 41*, 909-920.

Rochon, J. y Helms, R. W. (1989). Maximun Likelihood estimation for incomplete repeated measures experimental under on ARIMA covariance structure. *Biometrics, 45*, 207-218.

Rouanet, H. y Lepine, D. (1970). Comparison between treatments in a repeated-measurements design: ANOVA and multivariate methods. *British Journal of Mathematical on Statistical Psychology, 23*, 147-163.

Schaalje, B. Zhang,J.; Pantulla, S. y Pollock, K. (1991). Analysis of repeated measurements data from randomized block experiments. *Biometrics, 47*, 813-824.

Schluchter, M.D. (1988). BMDP5V- Unbalanced repeated measures: Models with structured covariance matrices. *Technical Report, 86 BMDP Statistical Software, Inc.*

Vallejo, G. (1991). *Análisis Univariado y Multivariado de los diseños de medidas repetidas de una sola muestra y de muestras divididas*. Barcelona: PPU.

Vallejo, G. y Fernández, P. (1990). Diseños de medidas repetidas con errores autocorrelacionados. *Psicothema, 2*, 189-209.

Vallejo, G. y Fernández, P.(1992). Diseños de medidas parcialmente repetidas con ausencia de independencia en los errores. En J. Arnau (Ed). *Diseños Longitudinales Aplicados a la Investigación Socio-Educativa*. México: Trillas (En prensa).

Vitalino,P.P. (1982). Parametric statistical analysis repeated measures experiments. *Psychoneuroendocrinology 1*, 3-13.