

Tasas de potencia de dos enfoques robustos para analizar datos longitudinales

G. Vallejo^{*1}, P. Fernández* y M. Ato**

**Universidad de Oviedo, **Universidad de Murcia

El propósito del presente trabajo fue examinar la potencia de dos de las pruebas más novedosas, y quizás también más robustas, de cuantas se hallan disponibles actualmente para analizar datos longitudinales con matrices de dispersión arbitrarias. En concreto, se comparó la sensibilidad de la versión mejorada del procedimiento multivariado Brown-Forsythe con la potencia del enfoque del modelo mixto lineal con los grados de libertad corregidos mediante la técnica Satterthwaite de emparejar momentos en un diseño longitudinal con un factor de agrupamiento y otro de medidas repetidas. Por lo que respecta al efecto principal de las ocasiones de medida, el enfoque del modelo mixto resultó ser siempre más potente que el procedimiento Brown-Forsythe, sobre manera, cuando los tamaños de muestra eran pequeños. Cuando el interés se centró en los efectos de la interacción, se obtuvo un patrón de resultados similar; sin embargo, bajo esta condición las diferencias fueron menos acentuadas.

En un trabajo de reciente aparición, Vallejo, Fernández y Velarde (2001b), han examinado las características operantes del error de Tipo I de algunas de las técnicas más novedosas de cuantas se hallan actualmente disponibles para analizar datos obtenidos repetidamente a lo largo del tiempo bajo una misma condición experimental o bajo condiciones experimentales diferentes. En concreto, los citados autores compararon el comportamiento del enfoque del modelo mixto con los grados de libertad corregidos mediante la técnica de Satterthwaite de emparejar momentos con el de la versión multivariada del enfoque de Brown y Forsythe (BF, 1974) modificada de acuerdo con el trabajo que en el ámbito de los diseños de

¹ *Agradecimientos:* Los autores agradecen al editor y a los dos revisores anónimos sus valiosos comentarios a cerca del trabajo. Sus sugerencias nos han permitido mejorar sustancialmente la exposición y la calidad del mismo. Este trabajo ha sido financiado mediante el proyecto de investigación concedido por el MCT (Ref.: BOS-2000-0410). *Correspondencia:* Guillermo Vallejo. Facultad de Psicología. Plaza Feijoo S/n, 33003 Oviedo. E-mail: gvallejo@sci.cpd.uniovi.es

medidas repetidas doblemente multivariadas llevaron a cabo de Vallejo, Fidalgo y Fernández (2001a). Aunque ninguno de los dos enfoques requiere para su correcta aplicación que las matrices de dispersión sean homogéneas, el enfoque del modelo mixto permite a los investigadores modelar la estructura de covarianza de sus datos antes de probar los efectos del diseño. De acuerdo con los defensores de este enfoque, seleccionar estructuras de covarianza más parcias proporciona la base racional para obtener inferencias más exactas y eficientes de los parámetros del modelo y, por consiguiente, tests más potentes de los efectos de los tratamientos. Por su parte la eficiencia del enfoque BF puede ser más reducida, pues al no imponer restricciones sobre la forma de la matriz de dispersión exige estimar todas las varianzas y covarianzas entre las ocasiones de medida. No obstante, Vallejo *et al.* (2001a) utilizando un diseño de mediadas repetidas, pero con múltiples variables dependientes encontraron que el enfoque BF era menos liberal que la extensión multivariada que de los procedimientos de Welch (1951) y James (1951) desarrolló Johansen (1980) y en muchos casos igual de potente. Conviene resaltar que versión multivariada de Welch-James basada en medidas de tendencia central y de variabilidad robustas, en lugar del tradicional estimador de los mínimos cuadrados, está adquiriendo un destacado protagonismo en los últimos años (ver por ejemplo, Wilcox, 1997; 2001).

Mediante ambas técnicas se contrastaron las hipótesis concernientes tanto a los efectos principales de las ocasiones de medida, como a los efectos secundarios de la interacción de los grupos por las ocasiones de medida cuando los grupos estaban desequilibrados y las matrices de dispersión eran heterogéneas, pero proporcionales entre sí. En lo tocante al efecto principal de las ocasiones de medida, el estudio de simulación de Vallejo *et al.* (2001b) puso de relieve que el enfoque del modelo mixto con los grados de libertad corregidos mediante la técnica de Satterthwaite proporcionaba un notable control de las tasas de error bajo las diferentes condiciones manipuladas. Dicho control se constataba con que se cumpliera tan sólo el requisito de que el tamaño del grupo más pequeño excediera el número de medidas repetidas. Por su parte, el procedimiento BF modificado también ejercía un excelente control de las tasas de error cuando el diseño resultaba equilibrado o de no estarlo, la relación entre el tamaño de los grupos y el tamaño de las matrices de dispersión era positiva. Cuando la relación entre el tamaño de los grupos y el tamaño de las matrices de dispersión era negativa, el enfoque BF exhibía un comportamiento ligeramente conservador. Sin embargo, el conservadurismo de este procedimiento tendía a reducirse a medida que la relación entre el tamaño del grupo más pequeño y el número de ocasiones de medida se aproximaba a tres. En lo que a los efectos de la interacción se refiere, conviene resaltar

que el comportamiento de ambos procedimientos resultó ser bastante similar al descrito para las ocasiones de medida. No obstante, el enfoque del modelo mixto resultó ser algo más conservador bajo esta condición.

A pesar de que ambos procedimientos ofrecían un control bastante satisfactorio de las tasas de error de Tipo I en la mayor parte de las condiciones experimentales investigadas, Vallejo *et al.* (2001b) recomiendan utilizar provisionalmente el procedimiento de BF, ya que bajo este enfoque el investigador aplica directamente la prueba multivariada sin la necesidad de conocer la correcta estructura de la matriz de dispersión de sus datos. Con todo, decantarse taxativamente por uno u otro procedimiento tomando como punto de referencia la superioridad de los mismos a la hora de controlar las tasas de error de Tipo I, se nos antoja una tarea condenada al fracaso, más si tenemos en cuenta que las tasas de error estimadas no favorecieron universalmente a ninguno de los dos enfoques, pues la ligera superioridad exhibida por el enfoque del modelo mixto a la hora de controlar las tasas de error referidas a los efectos principales, se vio contrarrestada por la ligera superioridad exhibida por la versión mejorada del enfoque de BF a la hora de controlar las tasas de error correspondientes a los efectos secundarios o de la interacción. En cualquier caso, conviene resaltar dos propiedades nada desdeñables de los mismos. Por un lado, que ninguno de los dos enfoques exhibió comportamiento liberal alguno bajo las múltiples condiciones examinadas y, por otro lado, que los valores críticos empíricos registrados para cada enfoque se aproximaban estrechamente a los teóricos. Aunque estos aspectos puedan parecer poco enjundiosos, no carecen de interés, dado que permiten satisfacer la exigencia de los investigadores que abogan por utilizar únicamente procedimientos cuyas tasas de error sean siempre menores o a lo sumo iguales al valor nominal (Mehta y Srinivasan, 1970), al tiempo que facilitarían la interpretación de los resultados obtenidos bajo hipótesis alternativa por ser éstos comparados sobre la misma base.

A raíz de lo expuesto, pensamos que la elección entre uno de los dos procedimientos puede depender de la sensibilidad de los mismos para detectar los efectos implicados en el diseño. Pues, como acabamos de exponer en los apartados anteriores, los dos enfoques en litigio proporcionaban un control más que aceptable de las tasas de error de Tipo I de los efectos del diseño cuando las matrices de dispersión eran heterogéneas y los grupos no estaban convenientemente equilibrados. Hasta la fecha, en el ámbito de los diseños de medidas repetidas univariadas, la potencia del enfoque BF nunca ha sido evaluada, ni tampoco comparada con la potencia del modelo mixto lineal. Por consiguiente, el objetivo fundamental del presente trabajo fue comparar la potencia de la versión mejorada del procedimiento multivariado de Brown-Forsythe con la

potencia del enfoque del modelo mixto con los grados de libertad corregidos mediante la técnica Satterthwaite de emparejar momentos para detectar los efectos distintos de cero en un diseño longitudinal con un factor de agrupamiento y otro de medidas repetidas. Una definición detallada de los procedimientos estadísticos examinados en la presente investigación se encuentra en el trabajo de Vallejo *et al.*(2001b).

MÉTODO

Un estudio de simulación Monte Carlo fue llevado a cabo para evaluar la potencia estadística del modelo mixto en relación con la del enfoque BF para detectar la presencia de efectos principales e interacciones. Estos efectos fueron probados en un diseño univariado de medidas repetidas con un factor de agrupamiento. El factor de agrupamiento tenía tres niveles y la variable ocasiones de medida cuatro. Las cinco variables que siguen fueron manipuladas en el estudio: El tamaño de muestra, la relación entre el tamaño de los grupos y el de las matrices de dispersión, el coeficiente de variación de los diferentes tamaños de los grupos, el tipo de matrices de dispersión y la permutación del vector de medias.

Tamaño de muestra. Las hipótesis a comparar son las referidas a diseños de medidas parcialmente repetidas con vectores de observaciones del mismo tamaño ($n_1=n_2=n_3=10$ y $n_1=n_2=n_3=20$) y vectores de observaciones de tamaño distinto ($n_1=8, n_2=10, n_3=12$; $n_1=6, n_2=10, n_3=14$; $n_1=16, n_2=20, n_3=24$; $n_1=12, n_2=20, n_3=28$; $n_1=12, n_2=10, n_3=8$; $n_1=14, n_2=10, n_3=6$; $n_1=24, n_2=20, n_3=16$; $n_1=28, n_2=20, n_3=12$). Estos tamaños de muestra son representativos de los empleados con mayor frecuencia en las investigaciones psicológicas y educacionales que utilizan diseños similares al nuestro (Kowalchuk, Lix y Keselman, 1996).

Relación entre el tamaño de los grupos y el de las matrices de dispersión. Se investigó el comportamiento de los dos enfoques referidos cuando la naturaleza de la relación entre el tamaño de los grupos y el tamaño de las matrices de dispersión era tanto positiva como negativa. Una relación positiva implica que el grupo de menor tamaño se asocia con la matriz de dispersión menor, mientras que una relación negativa implica que el grupo de menor tamaño se asocia con la matriz de dispersión mayor. Dependiendo del emparejamiento descrito, muchas pruebas estadísticas se comportan de forma liberal y/o conservadora. El grado de heterogeneidad de las matrices de dispersión incluida en el presente trabajo fue $\Sigma_1 = \frac{1}{3}\Sigma_2$ y

$\Sigma_3 = \frac{5}{3}\Sigma_2$. Este tipo de relación se ha manifestado perjudicial en otros estudios y por tanto también puede afectar a la sensibilidad de los procedimientos evaluados en la presente investigación.

Relación entre el tamaño de los grupos y el tamaño de las matrices. Se manipuló la desigualdad del tamaño de los grupos de una manera leve y moderada. En concreto, el coeficiente de variación de los diferentes tamaños de muestra (Δ) fue de .16 y de .33, donde $\Delta = \frac{1}{\bar{n}}[\sum_j (n_j - \bar{n})^2 / p]^{1/2}$ siendo \bar{n} el tamaño promedio de los grupos. Cuando los tamaños de muestra son iguales Δ vale 0, mientras que el valor del coeficiente se incrementa cuánto más difiere el tamaño de los grupos entre sí.

Patrón de las matrices de dispersión. Las estructuras de covarianza manipuladas para generar los datos simulados fueron las tres que siguen: Simetría compuesta homogénea (CS), simetría compuesta heterogénea (CSH) y autorregresiva de primer orden heterogénea (ARH(1)). Una descripción formal de las mismas se ofrece a continuación:

$$(a) CS = \begin{bmatrix} \sigma^2 & \sigma^2\rho & \sigma^2\rho & \sigma^2\rho \\ & \sigma^2 & \sigma^2\rho & \sigma^2\rho \\ & & \sigma^2 & \sigma^2\rho \\ & & & \sigma^2 \end{bmatrix}, (b) CSH = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho & \sigma_1\sigma_4\rho \\ & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho \\ & & \sigma_3^2 & \sigma_3\sigma_4\rho \\ & & & \sigma_4^2 \end{bmatrix}, (c) ARH(1) = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho^2 & \sigma_1\sigma_4\rho^3 \\ & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho^2 \\ & & \sigma_3^2 & \sigma_3\sigma_4\rho \\ & & & \sigma_4^2 \end{bmatrix}$$

El patrón expuesto en primer lugar representa una estructura de covarianza homogénea. Las matrices que satisfacen dicho patrón se las denomina de simetría compuesta o combinada, pues las observaciones tomadas desde una misma unidad experimental, además de mantener entre sí la misma correlación, tienen varianza constante. Por el contrario, los patrones expuestos en segundo y tercer lugar, representan estructuras de covarianza heterogéneas. Este tipo de heterogeneidad dentro de las ocasiones de medida no es más que una mera generalización de la matriz de simetría compuesta que acabamos de describir y de la matriz autorregresiva de primer orden. Pues, mientras que en la matriz CSH la correlación permanece constante, en la matriz ARH(1) decae exponencialmente. En cualquier caso, ambas matrices requieren estimar el mismo número de parámetros. Para el trabajo que nos ocupa, los valores de los parámetros de

la primera matriz fueron $\sigma^2 = 10$ y $\rho = .73$, mientras que los parámetros de las otras dos matrices adoptaron los valores que siguen: $\sigma_1^2 = 8$, $\sigma_2^2 = 10$, $\sigma_3^2 = 10$, $\sigma_4^2 = 12$ y $\rho = .73$. El índice de esfericidad de la matriz CS es uno y el de las matrices CSH y ARH(1) se fijó en 0.75. Repare el lector que la manipulación conjunta de la variable tipo de matrices de dispersión y relación entre el tamaño de los grupos y de las matrices de covarianza da lugar a que la heterogeneidad esté presente, tanto entre los grupos como a través de los mismos.

Permutación de la configuración del patrón del vector de medias.

Sobre la base del trabajo de Algina y Keselman (1998) la configuración seleccionada fue la de rango máximo. De acuerdo con Ramsey (1978), esta configuración implica que la primera media del vector toma el valor más pequeño, la última media toma el valor mayor y las dos medias intermedias el promedio de las anteriores. Para detectar la sensibilidad de las ocasiones de medida, dentro de cada uno de los tres grupos del diseño, se incluyeron las permutaciones que siguen: $(-\mu, 0, 0, \mu)$, $(-\mu, 0, \mu, 0)$, $(-\mu, \mu, 0, 0)$, $(0, -\mu, 0, \mu)$, $(0, -\mu, \mu, 0)$ y $(0, 0, -\mu, \mu)$. Cuando el interés se centró en comparar la potencia de ambas pruebas para detectar los efectos de la interacción, únicamente el tercer grupo tenía un vector de medias idéntico al descrito; para los otros dos grupos, el vector de medias fue el vector nulo. Para cada uno de los dos tamaños de muestra manipulados y para cada uno de los efectos implicados en el diseño, se seleccionó un valor de μ que proporcionara un valor objetivo de potencia *a priori* de 0.80. No hay razón para pensar que el patrón de resultados se viese alterado de haber utilizado algún otro valor, por ejemplo $1-\beta = .40$, más en consonancia con los obtenidos en las investigaciones psicológicas publicadas (Sedlmeier y Gigerenzer, 1989). El valor de potencia objetivo se obtuvo para el tradicional análisis de la varianza propuesto por Scheffé (1956) cuando los datos satisfacían el supuesto de esfericidad multimuestral. Para el efecto principal de las ocasiones de medida, los valores de los parámetros de no centralidad requeridos para satisfacer la potencia objetivo fueron: $\delta = 11.440$ para $N = 30$ y $\delta = 11.154$ para $N = 60$. A su vez, para la interacción de los grupos por las ocasiones de medida, los valores de los parámetros de no centralidad requeridos para satisfacer la potencia objetivo fueron: $\delta = 14.714$ para $N = 30$ y $\delta = 14.134$ para $N = 60$.

Las tasas de potencia estimadas fueron obtenidas simulando datos desde distribuciones normales multivariadas. Para ello, en un principio, para cada nivel de la variable entre sujetos se generaron vectores pseudoaleatorios de variables normales mediante la subrutina RNDN del

programa GAUSS (Aptech Systems, 1997). Dicha rutina crea una matriz de números pseudoaleatorios sin requerir especificar explícitamente ninguna semilla como uno de sus argumentos. Por defecto GAUSS genera una semilla inicial que se actualiza automáticamente cada vez que el programa es invocado. A continuación, las correspondientes observaciones multivariadas se consiguieron siguiendo el procedimiento descrito por Vallejo *et al.* (2001b), de acuerdo con el método propuesto por Schauer y Stoller (1966).

Por último, mediante sendos programas escritos en GAUSS se efectuó el análisis del conjunto de datos simulados con cada uno de los enfoques expuestos. En la presente investigación cada uno de los procedimientos fue aplicado a muestras pseudoaleatorias diferentes, por tanto, se considera que el diseño de simulación empleado implica grupos independientes. Para cada condición manipulada se efectuaron 10.000 replicaciones adoptando el nivel de significación $\alpha = 0.05$. La razón para efectuar el estudio de potencia utilizando los valores críticos nominales en lugar de los valores críticos empíricos obtenidos en el estudio de robustez por Vallejo *et al.* (2001b), se debió a que el margen de discrepancias entre ambos valores resultó muy reducido para los dos procedimientos examinados. Por tanto, esta forma de proceder garantiza que las tasas de potencia registradas bajo cada uno de los enfoques sean comparadas sobre la misma base. A pesar de lo costoso del asunto, de existir un dilema interpretativo por ser una prueba liberal y otra conservadora, tal vez lo más sensato hubiese sido abordar el estudio de potencia utilizando tanto los valores críticos empíricos de algún estudio de robustez previo, como valores críticos teóricos. El primer planteamiento permitiría que las comparaciones fuesen efectuadas como si las verdaderas regiones críticas resultasen conocidas, mientras que el segundo permitiría conocer la probabilidad de rechazar la hipótesis nula falsa para un nivel de significación prefijado de antemano.

RESULTADOS

1. Tasas de potencia empíricas para las ocasiones de medida.

En la Tabla 1 aparecen recogidas las tasas de potencia empírica correspondientes al efecto principal de las ocasiones de medida para cada una de las 180 condiciones manipuladas. Con el propósito de comprobar si las tasas obtenidas bajo cada uno de los procedimientos utilizados eran estadísticamente significativas se aplicó la prueba Fisher-Irwin para muestras grandes (Conover, 1999). El enfoque del modelo mixto lineal fue

elegido como procedimiento de referencia porque probablemente sea más potente, aunque las tasas de error de Tipo I obtenidas con el mismo en el estudio de Vallejo *et al.* (2001b) fuesen comparables a las obtenidas con el procedimiento BF.

Tabla 1. Tasas de potencia estimadas para las ocasiones de medida. Potencia simulada =.8, N = 30 y N = 60.

Tabla 1. Tasas de potencia estimadas para las ocasiones de medida. Potencia simulada =.8, N = 30 y N = 60

E.C	N	-μ 0 0 μ		-μ 0 μ 0		-μ μ 0 0		0 -μ 0 μ		0 -μ μ 0		0 0 -μ μ	
		Mixto	MBF										
CS	10-10-10	.8112	.7388	.8145	.7322	.8026	.7320	.8152	.7358	.8138	.7389	.8157	.7374
CS	08-10-12	.8421	.7663	.8432	.7790	.8362	.7761	.8415	.7728	.8449	.7754	.8467	.7767
CS	06-10-14	.8428	.7864	.8397	.7921	.8443	.7888	.8482	.7885	.8441	.7880	.8505	.7890
CS	12-10-08	.7661	.6342	.7669	.6330	.7716	.6258	.7655	.6265	.7637	.6320	.7667	.6272
CS	14-10-06	.6870	.4443	.6853	.4468	.6805	.4488	.6752	.4451	.6733	.4438	.6778	.4397
CSH	10-10-10	.7985	.7169	.8543	.7827	.8593	.7790	.7641	.6840	.8187	.7347	.7684	.6888
CSH	08-10-12	.8160	.7565	.8718	.8198	.8755	.8061	.7982	.7233	.8412	.7753	.7979	.7397
CSH	06-10-14	.8281	.7763	.8761	.8276	.8740	.8277	.7998	.7495	.8432	.7851	.8088	.7472
CSH	08-10-12	.7437	.6099	.7862	.6723	.8062	.6721	.7157	.5803	.7626	.6272	.7252	.5850
CSH	06-10-14	.6675	.4295	.7163	.4799	.7247	.4850	.6295	.4016	.6750	.4444	.6352	.4001
ARH	10-10-10	.5550	.4808	.6902	.6134	.8814	.8147	.6267	.5409	.9103	.8477	.8349	.7695
ARH	08-10-12	.5772	.5170	.7180	.6511	.9046	.8527	.6448	.5839	.9303	.8892	.8651	.8017
ARH	06-10-14	.5760	.5239	.7182	.6689	.9013	.8678	.6595	.5993	.9326	.8952	.8703	.8169
ARH	08-10-12	.5029	.3821	.6341	.5055	.8402	.7153	.5717	.4426	.8683	.7514	.7924	.6616
ARH	06-10-14	.4268	.2555	.5538	.3506	.7710	.5178	.4864	.2973	.8003	.5577	.7122	.4649
CS	20-20-20	.8351	.7653	.8364	.7736	.8370	.7625	.8378	.7731	.8392	.7611	.8401	.7738
CS	16-20-24	.8599	.8041	.8653	.8102	.8632	.8106	.8627	.8899	.8623	.8068	.8622	.8106
CS	12-20-28	.8722	.8223	.8730	.8141	.8744	.8175	.8748	.8167	.8680	.8106	.8678	.8178
CS	24-20-16	.7868	.7032	.7864	.7013	.7829	.6966	.7838	.7045	.7856	.7003	.7888	.6930
CS	28-20-12	.7027	.5724	.7029	.5730	.6961	.5661	.7113	.5689	.6995	.5807	.6965	.5787
CSH	20-20-20	.8242	.7496	.8715	.8058	.8663	.8158	.7973	.7245	.8337	.7667	.7898	.7184
CSH	16-20-24	.8394	.7911	.8921	.8496	.8992	.8379	.8262	.7580	.8688	.8138	.8146	.7571
CSH	12-20-28	.8493	.8016	.9047	.8611	.9032	.8503	.8313	.7726	.8738	.8161	.8296	.7773
CSH	24-20-16	.7664	.6783	.8224	.7469	.8310	.7408	.7349	.6506	.7854	.6968	.7442	.6523
CSH	28-20-12	.6874	.5662	.7525	.6263	.7395	.6230	.6562	.5283	.7080	.5632	.6434	.5388
ARH	20-20-20	.5755	.5054	.7065	.6481	.9026	.8407	.6443	.5796	.9340	.8733	.8645	.8020
ARH	16-20-24	.6015	.5480	.7435	.6775	.9288	.8774	.6786	.6175	.9505	.9057	.8882	.8290
ARH	12-20-28	.6014	.5487	.7526	.6889	.9359	.8892	.6788	.6341	.9599	.9115	.8955	.8445
ARH	24-20-16	.5224	.4462	.6571	.5682	.8586	.7803	.5904	.5120	.8972	.8194	.8128	.7261
ARH	28-20-12	.4532	.3411	.5738	.4604	.7797	.6638	.5060	.4040	.8234	.6960	.7366	.6056

Nota: Las tasas en negrita indican diferencias significativas del modelo mixto ($\alpha=.05$).

Los resultados presentados en la Tabla 1 revelan que el enfoque del modelo mixto con los grados de libertad corregidos mediante la técnica de Satterthwaite de emparejar momentos fue siempre más potente que la correspondiente versión mejorada del procedimiento multivariado de BF. Las tasas de potencia fueron estadísticamente significativas en las 180 condiciones examinadas. No obstante, es importante resaltar que las tasas de potencia registradas bajo el enfoque del modelo mixto se incrementaron levemente al aumentar el tamaño de muestra, mientras que las tasas de

potencia registradas bajo el procedimiento de BF se incrementaron moderadamente. Como consecuencia, las diferencias entre los patrones de potencia de ambos procedimientos, eran sustanciales cuando $N = 30$ y moderadas cuando $N = 60$. Especialmente, cuando $\Delta = 0$ y la relación entre el tamaño de los grupos y el tamaño de las matrices de dispersión era positiva. Como se puede apreciar en la Figura 1, esta pauta de conducta se mantuvo inalterable con independencia de las permutaciones efectuadas con la configuración del vector de medias. Aunque no se muestra en la Figura 1, dicha pauta de conducta tampoco dependía de la forma que tuviese la matriz de covarianza manipulada. La Figura 1 también pone de relieve las tasas de potencia promediadas a través de las variables estructuras de covarianza y tipo de apareamiento estaban afectadas por el tipo de permutación del vector de medias. En particular, las tasas de potencia mayores se registraron bajo la permutación $(-\mu, \mu, 0, 0)$ y las menores bajo la permutación $(-\mu, 0, 0, \mu)$.

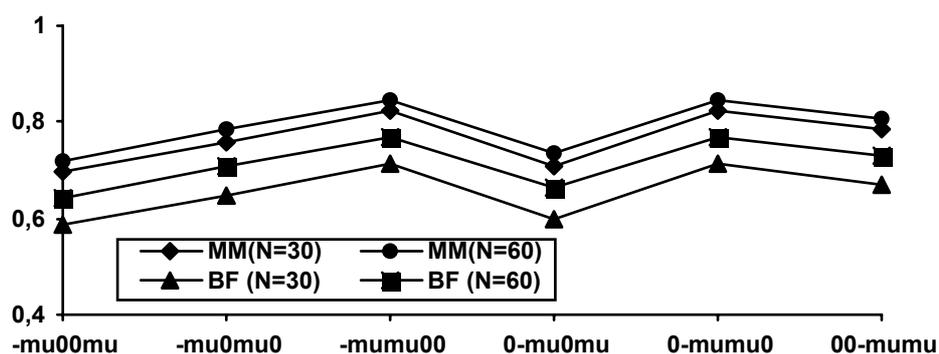


Figura 1. Tasas de potencia promediadas a través de las matrices de dispersión y condiciones de apareamiento para las ocasiones de medida con $N = 30$ y $N=60$.

Un examen más pormenorizado de los resultados presentados en la Tabla 1, pone de manifiesto otra serie de aspectos que resultaron comunes a los dos procedimientos y que también requieren ser descritos. En primer lugar, cuando el índice de esfericidad era igual a uno, como ocurría en el caso de la matriz CS, las tasas de potencia estimadas se veían afectadas por la relación existente entre el tamaño de los grupos y el tamaño de las matrices de dispersión; sin embargo, no guardaban ninguna relación con las permutaciones realizadas en la configuración del vector de medias. En

segundo lugar, cuando el índice de esfericidad era distinto de uno, como ocurría en el caso de las matrices CSH y ARH(1) donde dicho índice se fijó en 0.75, las tasas de potencia estimadas dependían tanto de la relación existente entre el tamaño de los grupos y el tamaño de las matrices de dispersión, como de las permutaciones realizadas con la configuración del vector de medias. En tercer lugar, independientemente del tamaño de muestra, de la permutación del vector de medias y de la estructura que tuviesen las matrices de dispersión, las tasas de potencia más elevadas se registraban cuando la relación existente entre el tamaño de los grupos y el tamaño de las matrices de dispersión era positiva y las tasas de potencia menos elevadas cuando la relación era negativa. Cuando la relación era nula, las tasas de potencia ocupaban un lugar intermedio. Resaltar, por último, que la forma de las matrices dispersión también afectaba a las tasas de potencia estimadas. La mayor estabilidad en las tasas de potencia acontecía cuando las matrices tenían simetría combinada y la menor cuando la estructura de las matrices era heterogénea y autorregresiva de primer orden.

2. Tasas de potencia empíricas para la interacción grupos x ocasiones de medida

En la Tabla 2 se muestran las tasas de potencia estimadas correspondientes a la interacción de los grupos x las ocasiones de medida para cada una de las condiciones manipuladas.

Un examen de los resultados presentados en la Tabla 2 revela que cuando $N=30$, el enfoque del modelo mixto con los grados de libertad corregidos mediante la técnica de Satterthwaite de emparejar momentos fue casi siempre más potente que la correspondiente versión mejorada del procedimiento multivariado de BF. En concreto, de acuerdo con la prueba Fisher-Irwin en 83 de las 90 condiciones manipuladas las tasas de potencia fueron estadísticamente significativas. Cuando $N=60$, el patrón de potencia observado también favorecía al enfoque del modelo mixto, no obstante, las diferencias eran más pequeñas que en la condición anterior. En concreto, las tasas de potencia promediadas a lo largo de las 180 condiciones experimentales manipuladas en la investigación resultaron ser de 0.664 para el enfoque del modelo mixto y de 0.643 para el enfoque de BF. Con todo, en 68 de las 90 condiciones manipuladas las tasas de potencia continuaron siendo estadísticamente significativas.

Al igual que ocurría en el caso de las ocasiones de medida, es importante resaltar que mientras las tasas de potencia registradas bajo el procedimiento de BF se incrementaban sustancialmente al aumentar el tamaño de muestra, dicho incremento afectó moderadamente a las tasas de

potencia registradas bajo el enfoque del modelo mixto. Sin embargo, como se puede apreciar en la Figura 2, a diferencia de lo que sucedía con el efecto principal de las ocasiones de medida, las diferencias entre los patrones de potencia de ambos procedimientos, fueron más modestas cuando N = 30. En este caso, los resultados también se mantuvieron inalterables independientemente de la forma de las matrices de dispersión y de las permutaciones efectuadas con la configuración del vector de medias. Desde la Figura 2 también es fácil de constatar que las diferencias entre ambos procedimientos tendían a estrecharse a medida que se incrementaba el tamaño de muestra.

Tabla 2. Tasas de potencia para la interacción de los grupos y las ocasiones de medida. Potencia simulada = .8, N = 30 y N = 60.

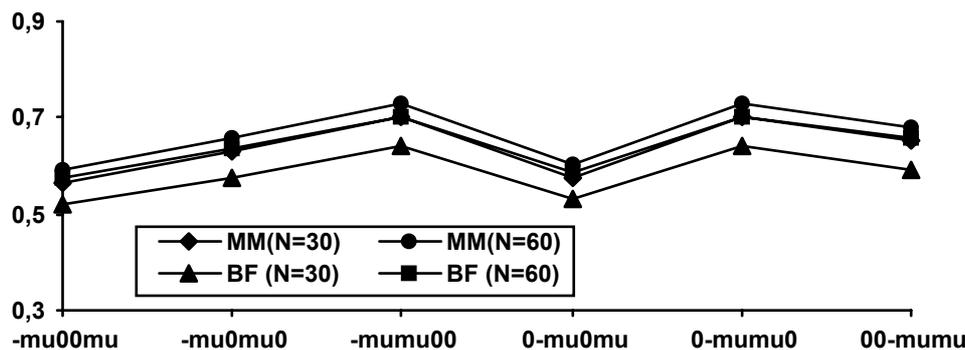
Tabla 2. Tasas de potencia para la interacción de los grupos y las ocasiones de medida. Potencia simulada = .8, N = 30 y N = 60

E.C	N	-μ 0 0 μ		-μ 0 μ 0		-μ μ 0 0		0 -μ 0 μ		0 -μ μ 0		0 0 -μ μ	
		Mixto	MBF										
CS	10-10-10	.6922	.6536	.6836	.6383	.6918	.6360	.6584	.6394	.6935	.6464	.6912	.6450
CS	08-10-12	.7496	.7219	.7594	.7253	.7574	.7218	.7583	.7239	.7581	.7257	.7524	.7251
CS	06-10-14	.7930	.7766	.7917	.7715	.7913	.7689	.7923	.7705	.7968	.7683	.7919	.7594
CS	12-10-08	.5996	.5126	.5962	.5163	.5966	.5196	.5939	.5230	.6024	.5251	.5926	.5209
CS	14-10-06	.4694	.3410	.4777	.3491	.4759	.3421	.4721	.3476	.4708	.3428	.4792	.3382
CSH	10-10-10	.6700	.6288	.7388	.6833	.7365	.6818	.6389	.5941	.6803	.5961	.6375	.5938
CSH	08-10-12	.7260	.7068	.7928	.7676	.8006	.7722	.7040	.6725	.7461	.7300	.7093	.6749
CSH	06-10-14	.7665	.7451	.8295	.8132	.8329	.8103	.7394	.7246	.7895	.7677	.7326	.7219
CSH	08-10-12	.5806	.5053	.6369	.5579	.6353	.5569	.5433	.4716	.5929	.5992	.5558	.4728
CSH	06-10-14	.4595	.3353	.5071	.3758	.5098	.3764	.4339	.3074	.4698	.3506	.4278	.3109
ARH	10-10-10	.4097	.3953	.5471	.5093	.7730	.7296	.4815	.4584	.8182	.7728	.7258	.6690
ARH	08-10-12	.4632	.4487	.6091	.5789	.8391	.8110	.5402	.5380	.8702	.8415	.7811	.7546
ARH	06-10-14	.4948	.4985	.6484	.6416	.8639	.8482	.5807	.5725	.8966	.8789	.8160	.8000
ARH	08-10-12	.3478	.3091	.4649	.4030	.6911	.6045	.4092	.3595	.7290	.6529	.6230	.5567
ARH	06-10-14	.2703	.2156	.3575	.2635	.5526	.4204	.3202	.2383	.6030	.4453	.4999	.3728
CS	20-20-20	.7248	.6958	.7280	.6985	.7225	.7058	.7181	.6995	.7281	.6932	.7207	.6942
CS	16-20-24	.7932	.7630	.7804	.7613	.7832	.7579	.7782	.7589	.7833	.7635	.7821	.7616
CS	12-20-28	.8279	.7988	.8215	.7994	.8214	.8019	.8211	.7999	.8182	.7961	.8213	.8028
CS	24-20-16	.6313	.5962	.6379	.6046	.6291	.5938	.6152	.5955	.6388	.5848	.6355	.5962
CS	28-20-12	.5008	.4541	.4993	.4604	.5000	.4581	.4982	.4578	.5070	.4629	.5095	.4503
CSH	20-20-20	.7057	.6834	.7604	.7359	.7612	.7370	.6645	.6372	.7068	.7005	.6352	.6458
CSH	16-20-24	.7612	.7379	.8157	.8061	.8202	.7941	.7252	.7141	.7795	.7471	.7299	.7159
CSH	12-20-28	.7920	.7789	.8380	.8339	.8479	.8350	.7572	.7492	.8126	.7991	.7495	.7529
CSH	24-20-16	.6044	.5850	.6730	.6424	.6855	.6469	.5750	.5464	.6207	.5887	.5762	.5459
CSH	28-20-12	.4866	.4410	.5477	.4982	.5445	.4952	.4610	.4147	.5037	.4576	.4571	.4197
ARH	20-20-20	.4381	.4368	.5762	.5627	.8037	.7829	.5050	.5010	.8486	.8183	.7422	.7231
ARH	16-20-24	.4538	.4935	.6361	.6388	.8606	.8395	.5633	.5645	.9044	.8725	.8097	.7992
ARH	12-20-28	.5190	.5239	.6734	.6743	.8901	.8777	.5891	.5959	.9275	.9038	.8394	.8279
ARH	24-20-16	.3687	.3600	.4986	.4739	.7170	.6878	.4251	.4131	.7620	.7193	.6674	.6331
ARH	28-20-12	.2956	.2684	.3842	.3518	.5854	.5426	.3392	.3052	.6232	.5813	.5435	.4896

Nota: Las tasas en negrita indican diferencias significativas del modelo mixto ($\alpha=.05$).

El patrón de resultados restante coincide sustancialmente con el informado anteriormente para las ocasiones de medida, de ahí que no nos detengamos en repetir lo mismo. No obstante, no queremos terminar la descripción de los resultados, sin destacar que, si bien desde el punto de vista cualitativo el patrón de potencia obtenido para la interacción era

prácticamente idéntico al registrado para el efecto principal de las ocasiones de medida, desde el punto de vista cuantitativo no ocurría lo mismo. En concreto, las tasas de potencia resultaron ser moderadamente más elevadas cuando el interés se centró en los efectos principales, la razón pudo deberse



a que en el caso de la interacción el vector de medias no nulo se asoció con el grupo que tenía la mayor matriz de dispersión; esto es, con el tercer grupo. Recuérdese que para los grupos uno y dos un el vector de medias manipulado era nulo. Al margen de la asociación utilizada en la presente investigación, aún son posibles otras cinco asociaciones, por ejemplo, asociar el vector de medias no nulo con el grupo que tenía la matriz de dispersión más pequeña.

Figura 2. Tasas de potencia promediadas a través de las matrices de dispersión y condiciones de apareamiento para la interacción grupos x ocasiones con N = 30 y N=60.

DISCUSIÓN Y CONCLUSIONES

El propósito de la presente investigación se ha centrado en evaluar la sensibilidad de dos de los procedimientos más novedosos, y quizás más robustos de cuantos existen actualmente para analizar diseños de medidas repetidas en ausencia de esfericidad multimuestral. En concreto, hemos comparado la potencia del procedimiento multivariado de Brown-Forsythe modificado de acuerdo con el trabajo de Vallejo *et al.*(2001a) con la potencia del enfoque del modelo mixto con los grados de libertad corregidos mediante la técnica Satterthwaite de emparejar momentos.

Con independencia de la fuente de variación implicada, los resultados de nuestro estudio indican que cuando la estructura de dispersión de los datos está correctamente identificada el enfoque del modelo mixto era

siempre más potente que la versión mejorada del enfoque BF. No obstante, a medida que el tamaño de muestra se incrementa, nuestros resultados también sugieren que no se produciría una pérdida apreciable de potencia de utilizarse el procedimiento BF, en especial, para el caso de la interacción. No hay razón para pensar que los resultados se viesan sustancialmente alterados por utilizar otras potencias objetivo o inclusive datos obtenidos desde distribuciones no normales, a juzgar por estudios de potencia realizados por otros investigadores (Algina y Keselman, 1998; Keselman, Algina, Kowalchuk y Wolfinger (1999). Tampoco es probable que los valores potencia estimados difiriesen sustancialmente de los aquí expuestos de haber emparejado los resultados de ambos métodos y, por ende, haber contrastado la significación estadística de las potencias estimadas mediante alguno de los procedimientos existentes para probar proporciones correlacionadas (por ejemplo, la prueba de McNemar). Sin embargo, no estamos seguros de poder afirmar que el patrón de resultados resultase similar al descrito de haber utilizado matrices de dispersión que no fuesen múltiples unas de otras, variables de medida con un mayor número de niveles y estructuras de covarianza diferentes. Núñez-Antón y Zimmerman (2001) describen detalladamente diversas matrices de dispersión para modelar el comportamiento de datos longitudinales. Además, en el caso de la interacción también sería interesante verificar otras posibles asociaciones entre el grupo que presenta el vector de medias no nulo y la desigualdad de las matrices de dispersión.

Para concluir, resaltar, que si bien los resultados del estudio de robustez realizado por Vallejo *et al.* (2001b) no permitía decantarse taxativamente por ninguno de los dos enfoques, dado que las tasas de error estimadas no favorecieron universalmente a ninguno de los dos enfoques, los resultados del estudio de potencia que acabamos de comentar nos deberían llevar a recomendar el enfoque del modelo mixto, especialmente cuando el tamaño de muestra disponible fuese reducido. No obstante, conviene tener presente que mediante el enfoque del modelo mixto los investigadores no utilizan directamente pruebas univariadas o multivariadas para analizar los efectos del diseño, sino que deben seleccionar previamente la estructura de covarianza más adecuada mediante alguno de los criterios existentes para elegir entre modelos. Por ejemplo, una estructura de covarianza se puede seleccionar comparando los valores obtenidos tras ajustar varias estructuras de covarianza potenciales mediante el criterio de información de Akaike (1974). Ahora bien, como señala Galecki (1994), la identificación del modelo constituye la principal dificultad en el análisis paramétrico de datos longitudinales. En consecuencia, si el tamaño de muestra es relativamente grande, no se dispone de covariantes y se cuenta con el mismo número de observaciones para todas las unidades, nosotros

entendemos que una solución alternativa al enfoque del modelo mixto, es la versión mejorada del procedimiento multivariado de BF. Dicha recomendación se basa en la notable robustez exhibida por el enfoque en las diferentes situaciones que ha sido examinado, en su aceptable sensibilidad estadística y por no exigir a los investigadores tener que modelar las matrices de dispersión subyacentes a sus datos.

ABSTRACT

Power rates of two robust approaches for the analysis of longitudinal data. The purpose of this paper was to compare the power of two of the newest, and perhaps more robust, approaches among current available procedures for analyzing longitudinal data, when dispersion matrices were unequal across groups. In particular, the sensibility of the improved multivariate Brown-Forsythe approach was compared to the linear mixed model procedure, with the residual degrees of freedom based on Satterthwaite's technique for matching moments in a longitudinal design involving one grouping factor and one repeated-measures factor. The results indicate that the power value of the mixed-model approach was higher than the one observed in the Brown-Forsythe procedure for the repeated-measures main effects, particularly when the sample size was small. When interaction effects were the focus of interest, a similar pattern of results was obtained; however, being the differences less marked under this condition.

REFERENCIAS

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transaction on Automatic Control*, AC-19, 716-723.
- Algina, J. & Keselman, H. J. (1998). A power comparison of the Welch-James and Improved General Approximation tests in the split-plot design. *Journal of Educational and Behavioral Statistics*, 23, 152-169.
- Aptech Systems. (1997). *The Gauss System* (Vers. 3.2.32). Kent-Kangley Road Mapley Valley, Washington: Aptech Systems, Inc.
- Brown, M. B., & Forsythe, A. B. (1974). The small sample behavior of some statistics which test the equality of several means. *Technometrics*, 16, 129-132.
- Conover, W. J. (1999). *Practical Nonparametric Statistics*. Third Edition New York: John Wiley & Sons, Inc.
- Galecki, A.T. (1994). General class of covariance structures for two or more repeated factors in longitudinal data analysis. *Communications in Statistics-Theory and Methods*, 23, 3105-3119.
- James, G.S. (1951). The comparison of several groups of observations when the ratios of the population variances are unknown. *Biometrika*, 38, 324-329.
- Johansen, S. (1980). The Welch-James approximation of the distribution of the residual sum of squares in weighted linear regression. *Biometrika*, 67, 85-92.
- Keselman, H. J., Algina, J., Kowalchuk, R. K., & Wolfinger, R. D. (1999). The analysis of repeated measurements: A comparison of mixed-model Satterthwaite F tests and a nonpooled adjusted degrees of freedom multivariate test. *Communications in Statistics-Theory and Methods*, 28, 2967-2999.
- Kowalchuk, J. C., Lix, L. M., & Keselman, H. J. (1996). The analysis of repeated measures designs. Paper presented at the Annual Meeting of the Psychometric Society, Banff, Alberta.
- Mehta, J. S. & Srinivasan, R. (1970). On the Behrens-Fisher problem. *Biometrika*, 57, 649-655.
- Núñez-Antón, V. & Zimmerman, D. L. (2001). Modelización de datos longitudinales con estructuras de covarianza no estacionarias: Modelos de coeficientes aleatorios frente

- a modelos alternativos. *Qüestió. Quaderns d'Estadística i Investigació Operativa*, 25, 225-262.
- Ramsey, P. H. (1978). Power differences between pairwise multiple comparisons. *Journal of the American Statistical Association*, 73, 479-485.
- Scheffé, H. (1956). A mixed model for the analysis of variance. *Annals of Mathematical Statistics*, 27, 23-36.
- Schuer, E. M. & Stoller, D. S. (1966). On the generation of normal random vectors. *Technometrics*, 4, 278-280.
- Sedlmeier, P. & Gigerenzer, G. (1989). Do studies of statistical power have an effect on the power of studies?. *Psychological Bulletin*, 105, 309-316.
- Vallejo, G., Fidalgo, A. M., & Fernández, P. (2001a). Effects of covariance heterogeneity on three procedures for analysing multivariate repeated measures designs. *Multivariate Behavioral Research*, 36, 1-27.
- Vallejo, G., Fernández, P. & Velarde, H. (2001b). Un estudio comparativo de pruebas robustas para el análisis de datos longitudinales. *Metodología de las Ciencias del Comportamiento*, 3, 35-52.
- Welch, B.L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika*, 38, 330-336.
- Wilcox, R. R. (1997). *Introduction to Robust Estimation and Hypothesis Testing*. New York: Academic Press.
- Wilcox, R. R. (2001). *Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy*. New York: Springer.

(Manuscrito recibido: 13/11/01; aceptado: 18/2/02)