

SOFTWARE, INSTRUMENTACIÓN Y METODOLOGÍA

COMPARACIONES MÚLTIPLES EN DISEÑOS TRANSVERSALES CON DATOS DEPENDIENTES

Guillermo Vallejo e Ignacio A. Menéndez
Universidad de Oviedo

El supuesto de independencia de las observaciones parece una asunción razonable cuando se hace uso del diseño experimental de grupos al azar con el propósito de examinar datos de corte transversal. Sin embargo, en la práctica dicho supuesto raramente es verificado. Por este motivo, pretendemos determinar mediante el conocido método de Monte Carlo las tasas de error Tipo I y II cometidas al utilizar diversos procedimientos de comparación múltiple en presencia de correlación dentro de los grupos.

Multiple comparisons in cross-sectional designs with dependent data. The assumption of independence of observations may seem like a reasonable assumption in examining data cross-sectional using experimental designs of single-factor. The assumption of independence, however is seldom verified by the investigators. By this reason we pretend to examine the relationship between true Types I and II error probabilities under the effects of departures from independence assumptions on hypothesis testing in six multiple comparison procedures.

Son muchos los experimentos diseñados con el propósito de poder determinar si algún efecto de los tratamientos está presente. Por regla general, en el caso de que una prueba de significación conduzca al rechazo de una hipótesis nula, la atención del investigador se dirige a explorar los datos en orden a descubrir tales efectos. Para tal fin diversos métodos han sido propuestos. Con todo, la elección de un procedimiento para comparar múltiples pares de medias no es precisamente

una tarea sencilla, pues en la actualidad existen diversos procedimientos, y en opinión de Kirk (1990) la lista de opciones disponibles continúa incrementándose. Afortunadamente, durante los últimos veinte años varios investigadores (Petronovich y Hardyck, 1969; Carmer y Swanson, 1973; Einot y Gabriel, 1975; Ramsey, 1978; Martin, Toothaker y Nixon, 1989; Seeman, Levin y Serlin, 1991) han llevado a cabo una serie de trabajos de carácter empírico en los que se han examinado las tasas de error Tipo I y Tipo II cometidos al utilizar los procedimientos de comparaciones múltiples más corrientes (Fisher, Fisher-Hayter, Dunn-Bonferroni, Holm, Shaffer, Tukey, Newman-

Correspondencia: Guillermo Vallejo
Departamento de Psicología
Universidad de Oviedo. Spain

Keuls, Ryan-Welsch, Peritz, Scheffé, etc). En la mayor parte de los trabajos reseñados, además de manipularse el tamaño de la muestra, el número de grupos y en ocasiones la forma de las distribuciones, se examinaron una gran cantidad de configuraciones de medias (rango máximo, rango mínimo, e igualmente espaciadas), diversas conceptualizaciones de la tasa de error (tasa de error por comparación, α_{pc} ; tasa de error por experimento, α_{pe} , y tasa de error por experimento concreto, $\alpha_{i,w}$) y varias operativizaciones de la potencia de las pruebas (potencia de cualquier par de comparaciones, potencia de todos los pares y potencia por comparación). A continuación, con el fin de enmarcar adecuadamente el problema que va a constituir nuestro objeto de estudio, comentaremos algunos de los descubrimientos que a nuestro juicio resultan más notorios.

Los resultados de estos experimentos ponen de relieve que cuando la unidad conceptual de error adoptada por los investigadores es la tasa de error por comparación y la hipótesis nula es globalmente verdadera la usual prueba de t controla la tasa de error al nivel del umbral estipulado. El resto de los procedimientos seleccionados para efectuar comparaciones múltiples manifiestan un control de la tasa de error por debajo del α estipulado. A su vez, cuando la hipótesis nula es globalmente verdadera, pero la unidad conceptual para la tasa de error adoptada por el investigador es la α_{cw} la usual prueba t arroja estimaciones empíricas de α que oscilan entre 0.25 cuando el número de grupos es igual a 5 ($p=5$) y 0.85 ($p=20$); resultados similares, aunque menos acentuados, proporciona la prueba de Duncan, en este caso α oscila entre 0.18 ($p=5$) y 0.60 ($p=20$). Por su parte el método *LSD* de Fisher mantiene la tasa de error controlada al nivel estipulado para la prueba F global. Los métodos de *HSD* de Tukey y de Newman-Keuls (en

adelante *NK*) controlan la tasa de error al nivel estipulado. Únicamente el procedimiento de Scheffé manifiesta un control de la tasa de error por debajo del umbral elegido. Por último, cuando la hipótesis nula es parcialmente verdadera, la unidad conceptual que se adopta para la tasa de error es la α_{cw} y el valor de p es igual a tres ($p=3$) todos los procedimientos post-hoc controlan aceptablemente la tasa de error, además los métodos *LSD* y *NK* resultan ser muy poderosos. Sin embargo, cuando $p>3$ los procedimientos *LSD* y *NK* son inaceptables en lo que al control de la α_{cw} se refiere, pues arrojan tasas de error empíricas que oscilan entre 0.11 ($p=5$) y 0.46 ($p=20$). En los restantes procedimientos el valor empírico de α se aproxima al α teórico, o bien resulta ser algo más pequeño dependiendo de la configuración de las medias manipulada por el investigador.

Por lo que a la potencia se refiere (solamente nos vamos a referir a la potencia por comparación), decir que cuando la hipótesis nula es parcialmente verdadera y la configuración de las medias es de rango mínimo las modificaciones efectuadas en el procedimiento *LSD* por Hayter (1986) y en el procedimiento *NK* por Ryan (1960), Einot y Gabriel (1975) y Welsch (1977), conlleva que éstos se muestren muy poderosos. El procedimiento de Tukey se encuentra en una posición intermedia y el de Scheffé en el extremo opuesto. Resultados similares aparecen cuando las medias están espaciadas igualmente. Sin embargo, cuando las medias tienen una configuración de rango máximo, el procedimiento de Tukey se manifiesta como un método enormemente poderoso. Por su parte, el de Scheffé continua siendo poco poderoso. En términos generales, podemos decir que el procedimiento de Fisher-Hayter y el de Tukey son alrededor de un 2% y 6%, respectivamente, menos poderosos

que el procedimiento más poderoso, de acuerdo con estos estudios el honor de ser la prueba más poderosa le corresponde al método de Peritz, pero con la ventaja añadida de su enorme simplicidad computacional. (Un programa escrito en Fortrán para ejecutar comparaciones múltiples haciendo uso del método Q de Peritz aparece en el n° 21 de la revista *Behavior Research Methods, Instruments and Computers*).

En lo que a la robustez se refiere, comentar que la investigación empírica tan sólo se ha dirigido al incumplimiento de las suposiciones de normalidad y homogeneidad. Por lo que respecta a la ausencia de normalidad los trabajos de Petrinovich y Hardick, ya citados, Keselman y Rogan (1978), Dunnett (1982) y Ringland (1983) ponen de relieve, como cuando el número de unidades experimentales dentro de cada grupo es el mismo y las varianzas son homogéneas la forma de la distribución tiene poco que ver en la protección que estas pruebas nos ofrecen frente a la probabilidad de cometer uno o más errores de tipo I, tan sólo la potencia se incrementa ligeramente en el caso de que p sea elevado ($p > 10$). Únicamente, si la distribución presenta un apuntamiento excesivo o varias puntuaciones extremas los procedimientos de Tukey y de Scheffé se vuelven ligeramente conservadores (en los trabajos de los autores citados y en lo referido a la robustez de los *PCM* los investigadores se han centrado principalmente en los procedimientos de Dunn, Bechhofer-Dunnett, Tukey y Scheffé). A su vez, por lo que respecta al incumplimiento del supuesto de homogeneidad resaltar que los distintos *PCM* también se comportan honestamente, sobre todo, cuando el número de unidades experimentales dentro de cada grupo es el mismo y sólo se efectúan comparaciones entre pares de medias. Más aún, en aquellas situaciones

en que la heterogeneidad es muy acentuada (1: 50) la mayoría de las pruebas ofrecen un adecuado control del error tipo I. A destacar entre las mencionadas la de Scheffé, la cual raramente ofrece un valor que va más allá del $\alpha = 0.06$; en estos casos la menos robusta resulta ser la Q de Peritz.

Como se puede apreciar de lo dicho, sorprende que no se haya realizado ninguna investigación de carácter empírico destinada a analizar el comportamiento de los procedimientos más comúnmente utilizados para efectuar comparaciones múltiples cuando se incumple el supuesto de independencia. Más aún si cabe, si se tiene presente que derivaciones analíticas recientes (Pavur, 1988, Scariano y Davenport, 1987) han puesto de relieve la incidencia que el incumplimiento de dicho supuesto tiene sobre las tasas de error de la mayor parte de las pruebas de significación existentes, sobre todo, de las clásicas pruebas de t y F . Obviamente, éste aspecto contrasta con la aceptable robustez de estas pruebas cuando los supuestos de normalidad y/o homogeneidad no se satisfacen; sobre manera, cuando el diseño se encuentra perfectamente balanceado y tan sólo se practican comparaciones entre pares de medias.

En nuestra opinión el que esto sea así se debe a que el supuesto de independencia de las observaciones en datos de sección cruzada parece un supuesto razonable y, por ende, raramente es verificado. Para comprobar lo dicho bastaría con echar una ojeada a la mayoría de los textos que sobre diseño existen. En estas obras todos los autores están de acuerdo que el incumplimiento de la suposición de independencia es la peor enfermedad que pueden presentar los datos, sin embargo, sólo en textos muy específicos se presentan pruebas para comprobar este supuesto. Por el contrario, no deja de ser paradójico que al tiempo que se rela-

tiviza la importancia del incumplimiento de las suposiciones de normalidad y/o de homogeneidad, se presenten numerosas pruebas para su verificación e inclusive pruebas denominadas robustas que no requieren ni tan siquiera verificarlas. En relación con el tópico que nos ocupa, esto es, el incumplimiento del supuesto de independencia, cabe suponer que los investigadores confían que el acto físico de la aleatorización asegure la independencia de las observaciones. Con todo, al margen de la propia forma de conducir el experimento (por ejemplo, trabajar todos juntos en grupos, medir a los sujetos más de una vez, etc), bien puede ocurrir que el proceso de aleatorización no sea todo lo correcto que debiera por no contar con un adecuado tamaño de muestra, e inclusive que éste no se llegue a efectuar.

Objetivos e hipótesis

A raíz de lo expuesto en el apartado anterior, resulta obvio que se necesitan nuevas investigaciones con el propósito de poder determinar de la manera más clara que nos sea posible los tres objetivos que siguen:

1.—Comprobar empíricamente el comportamiento de los procedimientos *LSD* de Fisher (1935), *FH* de Fisher (1935) y Hayter (1986), *NK* de Newman (1939) y Keuls (1952), *REGW* de Ryan (1960), Einot y Gabriel (1975) y Welsch (1977), *HSD* de Tukey (1953) y *S* de Scheffé (1959) a la hora de llevar a cabo comparaciones de medias bajo el incumplimiento del supuesto de independencia con datos de sección cruzada. Pues, aunque usualmente este supuesto se da por sentado, puede ocurrir que no se satisfaga; bien sea por deficiencias en el propio proceso de aleatorización o bien por la propia forma de conducir el experimento.

2.—Verificar si los datos se comportan conforme a lo que cabría esperar a partir

de las derivaciones teóricas efectuadas desde los trabajos de Pavur (1988) y de Scariano y Davenport (1987), entre otros.

3.—Comprobar hasta que punto, una vez que las puntuaciones están infectadas por la presencia de correlación, el método que nos proponemos desarrollar para corregir la falta de independencia proporciona resultados coincidentes con los resultados obtenidos por otros investigadores (Games y Serlin, 1991, Klockars y Hancock, 1994, y Seaman y otros, 1991) cuando se cumplen los supuestos del modelo; haciendo uso, claro está, de los mismos procedimientos de comparaciones múltiples y manipulando variables similares.

De este modo, consideramos la posibilidad de aportar un modelo de análisis para los datos obtenidos desde diseños experimentales de corte transversal en presencia de dependencia dentro de los grupos y/o a través de los grupos. Nosotros en el presente trabajo tan sólo nos centraremos en la primera, no obstante, el modelo matemático que vamos a desarrollar se generaliza fácilmente a diseños con datos correlacionados tanto dentro como a través de los grupos. Con ello, pensamos; que además de contribuir al desarrollo de nuevas técnicas de análisis más apropiadas para esta clase de datos (que sin ningún género de dudas hasta el momento han sido insuficientemente probados), también resolvemos, en parte, uno de los principales problemas que afectan a la investigación, cómo es decidir ¿cuál de los diferentes procedimientos propuestos para llevar a cabo comparaciones múltiples es más adecuado?, tanto cuando se da correlación en los datos, como cuando dicha correlación ha sido corregida. Todo ello a la luz de la tasa de error tipo I y de la potencia de prueba para cada uno de los dos procedimientos de análisis.

A su vez, por lo que respecta a las hipótesis, en la tabla 1 presentamos un

cuadro resumen derivado analíticamente por nosotros para esta ocasión con las probabilidades de error tipo I teóricas que esperamos encontrar con la prueba de Scheffé en el caso de que H_0 sea globalmente verdadera. Dicha tabla esta configurada en función de la correlación existente (r) del número de unidades experimentales dentro de cada grupo (n), del número de niveles de la variable (p) y de los valores c_1 y c_2 que definiremos en el próximo apartado; y a través de ella estamos en condiciones de poder predecir una serie de acontecimientos que esperamos que ocurran en la investigación empírica. En concreto, a partir de la tabla 1 podemos observar como a medida que la correlación se incrementa la probabilidad de cometer, al menos un error de tipo I, crece dramáticamente. Además, contrariamente a lo que cabría esperar, la tasa de error guarda una

Corrección de la correlación

Considérese el siguiente modelo de diseño experimental de un sólo factor

$$y = x\beta + \epsilon$$

donde

y es el vector de observaciones de orden $N \times 1$

X es la matriz de diseño de orden $N \times (p+1)$ la cual adopta la forma que sigue:

$$[X = (1_p I_p) \otimes 1_n]$$

b es el vector de parámetros del modelo de orden $(p+1) \times 1$.

ϵ es el vector de errores aleatorios de orden $N \times 1$ con $E(\epsilon) = 0$ y $V(\epsilon) = \sigma_\epsilon^2 U$

σ_ϵ^2 es la varianza de cada componente del vector de observaciones y U es una matriz de correlaciones.

Tabla 1

Verdaderas probabilidades de error tipo I para la prueba de Scheffé bajo H_0

p	n	ρ					
		0.00	0.05	0.10	0.15	0.20	0.30
3	10	0.050	0.130	0.223	0.312	0.386	0.537
	15	0.050	0.177	0.309	0.421	0.518	0.651
	19	0.050	0.214	0.367	0.487	0.579	0.708

* El nivel nominal elegido fue $\alpha = 0.05$

relación directa con el tamaño de la muestra y con el número de grupos. Similarmente, a partir de esta derivación teórica y, en base al trabajo de Pavur también cabe esperar que los procedimientos más conservadores no serán en el caso que nos ocupa los que nos ofrecen mayor protección frente a la probabilidad de cometer errores tipo I.

Si asumimos que $N=pn$, entonces el número de observaciones dentro de cada uno de los grupos que componen el diseño será el mismo. $b' = [m, b_1, b_2, \dots, b_p]$, donde m es la media general y b_1, b_2, \dots, b_p son los efectos de los p grupos bajo investigación. Con el fin de poder estimar los parámetros desconocidos del modelo asumimos que

$$\sum_{j=1}^p b_j = 0$$

La estructura de correlación que vamos a considerar entre las puntuaciones del vector $y' = [y_{11}, y_{12}, \dots, y_{1n}, y_{21}, y_{22}, \dots, y_{p1}, y_{p2}, \dots, y_{pn}]$ es una en la cual la correlación dentro de los grupos $r \neq 0$ y entre los grupos $r = 0$. De esta manera U es una matriz diagonal de bloques de orden $N \times N$ cuya forma general puede ser representada como sigue

$$U = \begin{bmatrix} A & 0 & \dots & 0 \\ 0 & A & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A \end{bmatrix}$$

donde 0 es una matriz de ceros de orden $n \times n$ y A es una matriz que presenta simetría combinada de orden $n \times n$. Por consiguiente, la matriz de varianzas-covarianza V puede expresarse como

$$V = \sigma^2 U = \sigma^2 I \otimes A$$

Si siguiendo el trabajo de Pavur y Lewis (1983), si definimos

$$M_n = \frac{1}{n} 1_n 1_n', \quad M_p = \frac{1}{p} 1_p 1_p', \quad M_{np} = M_n \otimes M_p, \quad C_n = I_n - M_n \text{ y } C_p = I_p - M_p.$$

la matriz de correlación U puede escribirse como

$$U = c_1 C_p \otimes M_n + c_2 I_p \otimes C_n + c_3 M_{np}$$

donde c_1, c_2 y c_3 son constantes positivas que guardan la siguiente relación

$$1 = (c_1 - c_2) \frac{1}{n} + c_2 + (c_3 - c_1) \left(\frac{1}{pn} \right) \\ \rho = (c_1 - c_2) \frac{1}{n} + (c_3 - c_1) \left(\frac{1}{pn} \right)$$

Por tanto, si r ($0 < r < 1$), $c_1 = 1 + p(n-1)$, $c_2 = (1 - r)$ y $c_3 = c_1$. Como nos señalan Scariano y Davenport (1987), si $c_1 = c_2 = c_3$ entonces $U = I_N$.

Establecidas las asunciones sobre el modo de alcanzar las observaciones en el modelo, debemos pasar a obtener el vec-

tor de coeficientes b . Para ello, es bien conocido que el estimador mínimo cuadrático generalizado de Aitken puede ser utilizado

$$\hat{\beta} = [X'V\rho^{-1}X]^{-1} X'V(\rho)^{-1}y$$

Obtener este estimador puede resultar bastante trabajoso dado que V es una matriz de orden $N \times N$. Sin embargo, si una matriz de transformación P puede ser descubierta, de modo que la matriz de varianzas-covarianzas de los residuales filtrados esté libre de correlación, el procedimiento se simplifica enormemente, pues el estimador b puede lograrse utilizando el método de los mínimos cuadrados ordinarios con la matriz y vector de variables X e y transformados por P .

Para la construcción de la matriz de transformación existen varios procedimientos, nosotros en una publicación relativamente reciente (Vallejo y Fernández, 1990), desarrollamos dos de estos métodos. Sin embargo, en el trabajo que nos ocupa, una vez que tenemos construida la matriz U conforme al procedimiento descrito por Pavur y Lewis, la forma más rápida de descubrir P de modo que se cumpla

$$P'P = U^{-1} \\ PUP' = I$$

es el ofrecido por Pavur (1988). De acuerdo con este investigador P viene dada por

$$P = \left(\frac{1}{\sqrt{c_1}} \right) C_p \otimes M_n + \left(\frac{1}{\sqrt{c_2}} \right) I_p \otimes C_n \left(\frac{1}{\sqrt{c_3}} \right) M_{np}$$

En consecuencia, al multiplicar la ecuación $y = Xb + e$ por P se obtiene

$$Py = PXb + Pe$$

haciendo $y = Py$, $X = PX$ y $e = Pe$ tenemos $y = Xb + e$. Donde e están distribuidos $N(0, s^2I)$, y el estimador b que cumple

con las propiedades del teorema Gauss-Markov es

$$\begin{aligned} \tilde{\beta} &= [X'V\rho^{-1}X]^{-1}X^{-1}V(\rho)^{-1}y \\ \tilde{\beta} &= (XPPX)^{-1}XPPy \\ \tilde{\beta} &= (\tilde{X}\tilde{X})^{-1}\tilde{X}\tilde{y} \end{aligned}$$

La prueba estadística para llevar a cabo las comparaciones múltiples la podemos expresar como sigue

$$t = \frac{\tilde{\psi}_i}{\tilde{\sigma}_{\psi_i}} = \frac{\bar{Y}_j^* - \bar{Y}_j^*}{\sqrt{\frac{2CM_c^*}{n}}}$$

donde $CM_c^* \psi_i$ son

$$\begin{aligned} CM_c^* &= [\tilde{y}'(I_p \otimes C_n)\tilde{y}] / v \\ \tilde{\psi}_i &= (c_{ij} \otimes I_n) \tilde{y} \end{aligned}$$

C_n e y -ya han sido definidos y c'_{ij} es un vector de orden $p \times 1$ cuyos componentes son todos ceros excepto los que ocupan las posiciones ith y jth que valen respectivamente $1/n$ y $-1/n$.

Método

Si bien es cierto que algunos resultados pueden derivarse analíticamente, no es menos cierto, que la complejidad y el excesivo número de variables hacen esta tarea prácticamente prohibitiva. Por tanto, en orden a evaluar los objetivos que nos hemos marcado en el apartado anterior diseñamos un experimento de simulación Monte Carlo en base a cuatro áreas de interés: Tipo de hipótesis (hipótesis nula global y parcialmente verdadera), configuración de las medias (configuración de rango mínimo, máximo e igualmente espaciadas), número de unidades experimentales dentro de cada grupo ($n= 10, 15, 19$) y grado de correlación dentro de los grupos ($r = 0.00, 0.05, 0.10, 0.15, 0.20$ y 0.30).

Los patrones de diferencias de medias que serán estudiados han sido tomados desde Seaman, Levin y Serlin (1991), la razón para ello es que, además de adaptarse perfectamente a las características de nuestra investigación, nos van a permitir ver el grado de coincidencia existente entre sus resultados y los nuestros con aquellos casos en los cuales $r = 0$, claro está, para aquellas pruebas que sean coincidentes en ambos estudios; en concreto, la de Fisher-Hayter, la de Tukey y la de Scheffé. El patrón de medias elegido para el estudio Monte Carlo es el que se muestra en la tabla 2.

Los valores especificados en la tabla 2 fueron elegidos de modo que cuando $n=10, n=15$ y $n=19$ la potencia de prueba fuese aproximadamente $1-\beta = 0.60, 1-\beta = 0.80$ y $1-\beta = 0.90$ al $\alpha = 0.05$ y $1-\beta = 0.32, 1-\beta = 0.56$ y $1-\beta = 0.74$ al $\alpha = 0.01$ siendo la varianza de cada población igual a la unidad.

Procedimiento

En orden a evaluar los objetivos expuestos, en una primera fase desde distribuciones normales, extraemos múltiples conjuntos de vectores pseudoaleatorios $y_i = [Y_{i1}, Y_{i2}, \dots, Y_{ip}]$ con vector de medias $m' = [m_1, m_2, \dots, m_p]$ y la matriz de varianza-covarianza V . Las observaciones son obtenidas mediante la descomposición triangular de V (frecuentemente referida como factorización de Cholesky o método de la raíz cuadrada, Harman, 1967), esto es,

$$y_i = \mu + Tz_i$$

donde T es una matriz triangular inferior que satisface la igualdad $V=TT'$ y z_i es un vector de puntuaciones normal e independientemente distribuidas con $m=0$ y $s^2=1$, generados de acuerdo al algoritmo propuesto por Kinderman y Ramage (1976).

Dos probabilidades han sido calculadas, la tasa de error por comparación y la potencia por comparación (probabilidad promedio de detectar una diferencia verdadera a lo largo de todas las comparaciones).

Para cada una de las 72 condiciones existentes (4 patrones de medias x 3 tamaños de muestras x 6 estructuras de correlación diferentes) se han efectuado 10.000 replicaciones adoptando los niveles de significación $\alpha=0.10, 0.05$ y 0.01 .

En una segunda fase procederemos a eliminar la dependencia que con anterioridad introdujimos en los datos que configuran nuestros grupos. Para ello, cada una de las series será transformada mediante el procedimiento descrito con anterioridad para corregir la correlación; de este modo, la matriz de varianza-covarianza del error será un escalar, satisfaciendo así las asunciones del modelo mínimo cuadrático. Al contar con series de puntuaciones en las cuales la correlación dentro de los grupos es nula ($r=0$) y

$r=0$ y entre aquellas otras en las cuales $r \neq 0$, pero ha sido corregido.

Por último, reseñar que para llevar a cabo los cálculos descritos anteriormente, hemos desarrollado un programa en Gauss (v. 2.0).

Resultados

Antes de comenzar a exponer los resultados hallados, queremos resaltar dos cuestiones. Por un lado, el análisis de las comparaciones múltiples ha sido realizado tanto cuando el análisis de la variancia resultó significativo como cuando no lo fue. Por otro lado, dado que el número de niveles de la variable tratamiento es de 3 ($p=3$), los resultados que obtendríamos por los procedimientos *LSD* y *FH* serían idénticos, lo mismo ocurriría con los resultados obtenidos mediante los procedimientos *NK* y *REGW*; de ahí que en las tablas que siguen tan sólo aparezcan los acronimos *LSD* y *NK*.

Tabla 2
Patrón de medias elegido para el estudio de simulación (adoptado de Seaman, Levin y Serlin, 1991)

m_1	m	m
0.0000	0.0000	0.0000
0.0000	0.0000	1.0210
0.0000	0.3859	1.1577
0.0000	0.5895	1.1790

series de datos en los cuales existe distinto grado de correlación ($r \neq 0$)

estamos en una disposición inmejorable para comprobar, no sólo cómo afectan las diferentes estructuras de correlación a las tasas de error, sino también para verificar si una vez corregida la correlación las probabilidades de cometer errores Tipo I y Tipo II se mantienen similares entre aquellas situaciones en las cuales

En el comentario de los resultados empezaremos detallando como afecta la dependencia de los datos al análisis de varianza, para posteriormente centrarnos en las comparaciones de medias.

Análisis de varianza

Cuando la hipótesis nula es globalmente verdadera encontramos que la tasa

de error Tipo I se dispara al ir aumentando la correlación, como se puede encontrar en la tabla 3. Este aumentó de

la tasa de error, una vez que hay dependencia en los datos, es directamente proporcional al tamaño de muestra y a.

Tabla 3
Proporciones de error tipo I y potencia del AVAR promediados a través de los diferentes patrones de medias

		10					
a		0.00	0.05	0.10	0.15	0.20	0.30
Error	.10	.102	.222	.323	.414	.499	.631
Tipo I		(.102)	(.106)	(.102)	(.099)	(.096)	(.097)
	.05	.050	.137	.228	.313	.398	.543
		(.050)	(.055)	(.051)	(.050)	(.048)	(.049)
	.01	.010	.047	.093	.155	.227	.363
		(.010)	(.011)	(.010)	(.010)	(.010)	(.009)

		10					
a		0.00	0.05	0.10	0.15	0.20	0.30
Potenc.	.10	.723	.736	.750	.767	.784	.820
		(.723)	(.574)	(.483)	(.413)	(.367)	(.307)
	.05	.596	.630	.663	.690	.715	.761
		(.596)	(.439)	(.346)	(.288)	(.249)	(.197)
	.01	.335	.405	.466	.514	.555	.631
		(.335)	(.202)	(.146)	(.109)	(.087)	(.065)

		15					
a		0.00	0.05	0.10	0.15	0.20	0.30
Error	.10	.096	.262	.409	.508	.600	.722
Tipo I		(.096)	(.097)	(.100)	(.101)	(.095)	(.096)
	.05	.046	.173	.309	.410	.512	.656
		(.046)	(.047)	(.053)	(.054)	(.047)	(.048)
	.01	.009	.064	.158	.247	.340	.508
		(.009)	(.009)	(.009)	(.010)	(.010)	(.008)

		15					
a		0.00	0.05	0.10	0.15	0.20	0.30
Potenc.	.10	.880	.858	.850	.855	.865	.880
		(.880)	(.688)	(.549)	(.461)	(.402)	(.321)
	.05	.798	.791	.792	.806	.819	.847
		(.798)	(.565)	(.417)	(.331)	(.279)	(.211)
	.01	.569	.618	.648	.679	.708	.758
		(.569)	(.309)	(.195)	(.140)	(.109)	(.076)

		19					
a		0.00	0.05	0.10	0.15	0.20	0.30
Error	.10	.106	.324	.471	.581	.663	.766
Tipo I		(.106)	(.101)	(.102)	(.098)	(.102)	(.104)
	.05	.052	.221	.375	.490	.578	.710
		(.052)	(.052)	(.051)	(.049)	(.050)	(.051)
	.01	.009	.092	.210	.323	.427	.578
		(.009)	(.010)	(.010)	(.010)	(.009)	(.010)

		19					
a		0.00	0.05	0.10	0.15	0.20	0.30
Potenc.	.10	.943	.910	.895	.891	.894	.906
		(.943)	(.745)	(.586)	(.488)	(.418)	(.333)
	.05	.894	.862	.850	.854	.857	.875
		(.894)	(.626)	(.454)	(.358)	(.295)	(.219)
	.01	.722	.727	.738	.756	.773	.803
		(.722)	(.370)	(.225)	(.157)	(.121)	(.077)

* Entre paréntesis proporciones de los datos corregidos

Con respecto a la potencia encontramos como esperabamos que para $\alpha=0.05$ ésta es aproximadamente 0.6, 0.8 y 0.9 para $n=10$, $n=15$ y $n=19$ respectivamente. Observamos que la potencia es mayor cuando aumenta la dependencia y cuando crece el tamaño de muestra. A lo apuntado en último lugar hay que hacer una salvedad, la potencia no aumenta e incluso disminuye ligeramente, aún en el caso de que se incremente r cuando la potencia es mayor de 0.85.

Comparaciones de medias

Es importante resaltar que la realización o no del análisis de varianza tan

sólo modifica los resultados para los procedimientos *LSD* y *FH*; en el resto de los procedimientos de comparaciones de medias, se obtienen los mismos resultados tanto si se realiza el *AVAR* como si no.

Como se puede comprobar en las tablas 4 y 5 podemos ordenar los procedimientos según sea su tasa de error por comparación de menor a mayor: *S*, *HSD*, *NK* y *REGW*, y *LSD* y *FH*. Por otra parte también podemos ordenarlos según su potencia de mayor a menor: *LSD* y *FH*, *NK* y *REGW*, *HSD* y *S* (estos resultados se encuentran en las tablas 6 y 7).

Tabla 4

Proporciones de error tipo I promediados a través de los diferentes patrones de medias

		10					
Prueba		0.00	0.05	0.10	0.15	0.20	0.30
0.10	LSD	.076	.147	.219	.274	.336	.436
	S	.032	.081	.139	.185	.240	.338
	HSD	.039	.092	.152	.201	.258	.355
	NK	.075	.145	.216	.271	.333	.433
0.10	LSD	.072	.184	.278	.346	.417	.519
	S	.031	.105	.181	.249	.319	.425
	HSD	.037	.119	.198	.266	.337	.444
	NK	.071	.182	.275	.342	.414	.517
0.10	LSD	.079	.214	.319	.400	.467	.558
	S	.033	.129	.221	.296	.367	.466
	HSD	.038	.144	.240	.320	.385	.484
	NK	.078	.212	.317	.397	.464	.556
		15					
Prueba		0.00	0.05	0.10	0.15	0.20	0.30
0.05	LSD	.036	.086	.145	.194	.252	.352
	S	.015	.046	.088	.130	.176	.268
	HSD	.018	.054	.100	.143	.193	.286
	NK	.034	.084	.142	.191	.248	.348
0.05	LSD	.034	.116	.195	.263	.336	.446
	S	.013	.065	.126	.187	.251	.361
	HSD	.017	.074	.139	.204	.269	.380
	NK	.033	.114	.192	.260	.332	.443
		19					
Prueba		0.00	0.05	0.10	0.15	0.20	0.30
0.05	LSD	.037	.170	.238	.314	.386	.487
	S	.016	.085	.163	.234	.305	.408
	HSD	.018	.094	.175	.247	.320	.423
	NK	.036	.138	.234	.310	.383	.484
0.01	LSD	.006	.026	.054	.089	.130	.214
	S	.002	.012	.030	.056	.087	.158
	HSD	.003	.016	.036	.064	.098	.173
	NK	.006	.025	.053	.085	.127	.210
0.01	LSD	.006	.041	.089	.143	.202	.312
	S	.002	.021	.055	.098	.149	.250
	HSD	.003	.025	.064	.110	.163	.266
	NK	.006	.040	.087	.140	.199	.309
0.01	LSD	.006	.055	.120	.184	.256	.360
	S	.002	.029	.076	.131	.193	.297
	HSD	.003	.034	.088	.145	.209	.314
	NK	.006	.053	.117	.180	.252	.358

Tabla 5
 Proporciones de error tipo I promediados a través de los
 diferentes patrones de medias (sin realizar AVAR)

		10								19							
Prueba		0.00	0.05	0.10	0.15	0.20	0.30			Prueba		0.00	0.05	0.10	0.15	0.20	0.30
0.10	LSD	.098	.180	.256	.319	.368	.467	0.05	LSD	.047	.163	.259	.340	.404	.508		
	S	.032	.081	.136	.193	.240	.339		S	.014	.084	.162	.238	.302	.411		
	HSD	.040	.094	.152	.212	.258	.358		HSD	.018	.095	.175	.254	.317	.427		
	NK	.076	.150	.224	.285	.336	.439		NK	.036	.141	.235	.315	.380	.488		
0.10	LSD	.100	.214	.308	.384	.445	.535										
	S	.032	.105	.182	.252	.318	.418										
	HSD	.040	.119	.200	.274	.340	.438										
	NK	.078	.185	.278	.356	.419	.514										
0.10	LSD	.099	.243	.348	.427	.486	.582										
	S	.031	.128	.217	.298	.363	.469										
	HSD	.038	.143	.237	.318	.381	.488										
	NK	.077	.214	.319	.400	.462	.563										
0.05	LSD	.049	.109	.172	.232	.280	.381										
	S	.015	.045	.087	.134	.179	.270										
	HSD	.019	.054	.099	.149	.197	.289										
	NK	.037	.087	.146	.203	.251	.354										
0.05	LSD	.049	.137	.222	.296	.362	.459										
	S	.015	.063	.128	.192	.251	.354										
	HSD	.019	.075	.144	.211	.271	.373										
	NK	.037	.115	.196	.269	.335	.437										

Tabla 6
 Proporciones de potencia por comparación promediados a través de los
 diferentes patrones de medias

		10								19							
Prueba		0.00	0.05	0.10	0.15	0.20	0.30			Prueba		0.00	0.05	0.10	0.15	0.20	0.30
0.10	LSD	.523	.538	.553	.571	.588	.610	0.10	LSD	.753	.725	.710	.713	.716	.733		
	S	.405	.433	.458	.481	.501	.550		S	.640	.631	.628	.636	.644	.669		
	HSD	.430	.454	.479	.498	.518	.583		HSD	.662	.650	.645	.651	.658	.682		
	NK	.517	.533	.549	.568	.585	.600		NK	.751	.723	.709	.711	.714	.732		
0.10	LSD	.677	.661	.657	.665	.676	.699										
	S	.553	.561	.568	.581	.598	.628										
	HSD	.579	.583	.588	.598	.614	.641										
	NK	.673	.658	.654	.662	.671	.695										

15								19							
	Prueba	0.00	0.05	0.10	0.15	0.20	0.30		Prueba	0.00	0.05	0.10	0.15	0.20	0.30
0.05	LSD	.402	.433	.463	.486	.509	.555	0.01	LSD	.196	.245	.291	.327	.358	.420
	S	.300	.339	.375	.415	.429	.479		S	.138	.182	.224	.261	.293	.357
	HSD	.345	.362	.396	.424	.447	.496		HSD	.155	.201	.244	.278	.311	.374
	NK	.394	.427	.457	.481	.504	.552		NK	.188	.237	.284	.320	.352	.415
0.05	LSD	.570	.575	.582	.596	.612	.642	0.01	LSD	.350	.398	.427	.455	.483	.530
	S	.453	.479	.495	.515	.538	.574		S	.263	.316	.351	.385	.414	.469
	HSD	.481	.502	.517	.533	.554	.558		HSD	.288	.340	.374	.405	.434	.486
	NK	.563	.570	.579	.592	.609	.639		NK	.340	.390	.421	.449	.478	.527
0.05	LSD	.664	.648	.644	.651	.658	.682	0.01	LSD	.464	.488	.505	.527	.547	.583
	S	.554	.561	.566	.580	.592	.623		S	.363	.400	.430	.458	.483	.524
	HSD	.575	.578	.582	.594	.605	.635		HSD	.393	.425	.452	.478	.501	.541
	NK	.660	.645	.641	.649	.656	.681		NK	.455	.481	.500	.523	.543	.581

Tabla 7
 Proporciones de potencia por comparación promediados a través de los
 difentes patrones de medias (sin realizar AVAR)

10								19								
	Prueba	0.00	0.05	0.10	0.15	0.20	0.30		Prueba	0.00	0.05	0.10	0.15	0.20	0.30	
0.10	LSD	.580	.587	.596	.608	.620	.649	0.05	LSD	.614	.610	.613	.619	.631	.657	
	S	.405	.433	.458	.482	.502	.543		S	.458	.478	.497	.516	.534	.574	
	HSD	.434	.459	.480	.503	.521	.561		HSD	.490	.505	.520	.536	.554	.591	
	NK	.528	.543	.557	.576	.591	.626		NK	.576	.577	.585	.595	.610	.641	
0.10	LSD	.708	.693	.684	.686	.693	.712	0.05	LSD	.690	.672	.664	.662	.675	.696	
	S	.558	.563	.571	.580	.595	.626		S	.557	.557	.567	.574	.591	.620	
	HSD	.586	.587	.592	.599	.613	.641		HSD	.580	.577	.583	.588	.605	.636	
	NK	.681	.667	.662	.665	.674	.697		NK	.666	.648	.643	.644	.658	.683	
0.10	LSD	.772	.744	.729	.722	.731	.747	0.01	LSD	.253	.295	.332	.365	.393	.446	
	S	.641	.629	.616	.628	.643	.669		S	.137	.184	.223	.261	.294	.354	
	HSD	.665	.651	.646	.645	.658	.682		HSD	.160	.207	.247	.284	.317	.376	
	NK	.756	.726	.711	.706	.717	.736		NK	.194	.242	.286	.324	.358	.417	
0.05	15								0.01	LSD	.410	.437	.461	.483	.506	.546
	LSD	.465	.486	.505	.526	.543	.580	S		.266	.314	.351	.385	.416	.467	
	S	.302	.340	.373	.405	.430	.478	HSD		.301	.344	.379	.408	.438	.486	
	HSD	.332	.368	.398	.428	.451	.497	NK		.354	.394	.425	.453	.481	.526	
0.01	LSD	.404	.436	.463	.490	.510	.554	0.01	LSD	.505	.516	.531	.541	.562	.598	
	S	.361	.397	.431	.452	.479	.526		S	.361	.397	.431	.452	.479	.526	
	HSD	.397	.427	.457	.474	.500	.543		HSD	.397	.427	.457	.474	.500	.543	
	NK	.460	.482	.504	.518	.542	.581		NK	.460	.482	.504	.518	.542	.581	

Los resultados obtenidos cuando la hipótesis nula es verdadera, nos indican que la tasa de error crece peligrosamente al incrementarse la dependencia. Volvemos a encontrar que la tasa de error aumenta (en presencia de correlación) al aumentar el tamaño de muestra y a La única diferencia encontrada entre realizar o no el AVAR es como se comentó con anterioridad para los procedimientos LSD y FH; si no realizamos el AVAR su tasa de error Tipo I se mantiene al nivel nominal elegido cuando $r=0$, mientras que si se hace un AVAR la tasa de error es aproximadamente la mitad.

Si analizamos los resultados hallados con respecto a la potencia de prueba, encontramos que los procedimientos que presentan una mayor potencia per se (cuando no hay dependencia) son los que menos la incrementan, porcentualmente, al aumentar la dependencia entre las puntuaciones. Las tasas más altas de potencia por comparación para todos los PCM se encuentran en la configuración de rango mínimo, encontrándose unos resultados muy similares entre la de rango máximo y la de medias igualmente espaciadas. En todos los casos se encuentra que la potencia aumenta al incrementarse el tamaño de muestra. Al igual que ocurría para el error Tipo I, en el caso de la potencia también se encuentran dife-

rencias entre los procedimientos LSD y FH si se realiza el AVAR o no, ya que ambos procedimientos presentan mayor potencia caso de no realizarse el AVAR.

Corrección de la dependencia

Hemos podido comprobar que la violación de la suposición de independencia es la peor enfermedad que pueden presentar los datos. Como indicamos anteriormente, Pavur (1988) presenta un procedimiento de corrección de la dependencia, que es el seguido en este trabajo. Los resultados hallados, se presentan respectivamente, en las tablas 8 y 9 para los errores Tipo I por comparación, y en las tablas 10 y 11 para la potencia por comparación.

Cuando la hipótesis nula es globalmente verdadera, comprobamos como la corrección de la dependencia del error, mantiene la tasa de error tipo I controlada al nivel nominal elegido para todos los PCM evaluados. No obstante, cuando la configuración de las medias es de rango mínimo, únicamente los procedimientos desarrollados por Scheffé y Tukey logran controlar adecuadamente la tasa de error. El resto de los procedimientos no logran controlar la tasa de error indistintamente de la configuración que presenten las medias.

Tabla 8

Proporciones de error tipo I promediados a través de los diferentes patrones de medias una vez corregida la dependencia

		10													
Prueba		0.00	0.05	0.10	0.15	0.20	0.30	0.10	LSD	.072	.075	.071	.069	.068	.066
0.10	LSD	.076	.073	.073	.069	.067	.062		S	.031	.033	.033	.034	.031	.032
	S	.032	.034	.033	.035	.032	.032		HSD	.037	.040	.038	.039	.038	.037
	HSD	.039	.040	.039	.040	.038	.038		NK	.071	.073	.069	.067	.065	.064
	NK	.075	.071	.070	.066	.065	.060	0.10	LSD	.079	.077	.074	.068	.070	.065
									S	.033	.032	.032	.031	.034	.031
									HSD	.038	.039	.039	.037	.040	.037
									NK	.078	.076	.072	.066	.068	.063

15								19							
	Prueba	0.00	0.05	0.10	0.15	0.20	0.30		Prueba	0.00	0.05	0.10	0.15	0.20	0.30
0.05	LSD	.036	.036	.034	.034	.031	.030	0.01	LSD	.006	.006	.006	.006	.006	.005
	S	.015	.015	.014	.016	.015	.014		S	.002	.003	.002	.002	.002	.002
	HSD	.018	.019	.018	.020	.018	.017		HSD	.003	.003	.003	.003	.003	.002
	NK	.034	.034	.033	.032	.030	.028		NK	.006	.006	.006	.006	.005	.005
0.05	LSD	.034	.037	.034	.034	.032	.030	0.01	LSD	.006	.006	.005	.006	.005	.006
	S	.013	.015	.015	.016	.014	.014		S	.002	.002	.002	.003	.002	.002
	HSD	.017	.019	.018	.020	.017	.017		HSD	.003	.003	.003	.003	.003	.003
	NK	.033	.036	.033	.033	.031	.029		NK	.006	.006	.005	.006	.005	.005
0.05	LSD	.037	.037	.034	.033	.033	.029	0.01	LSD	.006	.007	.007	.006	.006	.005
	S	.016	.015	.015	.015	.016	.014		S	.002	.002	.003	.002	.002	.002
	HSD	.018	.018	.017	.018	.019	.016		HSD	.003	.003	.003	.003	.003	.002
	NK	.036	.036	.032	.031	.031	.028		NK	.006	.006	.006	.006	.005	.004

Tabla 9

Proporciones de error tipo I promediados a través de los diferentes patrones de medias una vez corregida la dependencia (sin realizar AVAR)

10								19							
	Prueba	0.00	0.05	0.10	0.15	0.20	0.30		Prueba	0.00	0.05	0.10	0.15	0.20	0.30
0.10	LSD	.098	.100	.100	.101	.100	.099	0.05	LSD	.049	.047	.050	.050	.051	.048
	S	.032	.033	.035	.034	.033	.032		S	.015	.014	.015	.014	.014	.014
	HSD	.040	.041	.042	.041	.041	.039		HSD	.019	.018	.019	.017	.019	.018
	NK	.076	.075	.074	.073	.070	.066		NK	.037	.034	.035	.034	.034	.030
0.10	LSD	.100	.098	.099	.099	.099	.099	0.05	LSD	.047	.050	.049	.051	.047	.049
	S	.032	.030	.033	.032	.032	.031		S	.014	.016	.016	.017	.015	.015
	HSD	.040	.038	.041	.041	.040	.038		HSD	.018	.020	.019	.020	.018	.019
	NK	.078	.074	.075	.074	.072	.068		NK	.036	.033	.035	.035	.031	.031
0.10	LSD	.099	.102	.099	.102	.099	.098	19							
	S	.031	.033	.031	.034	.032	.032	0.01	LSD	.009	.010	.010	.010	.009	.009
	HSD	.038	.040	.039	.042	.040	.040	S	.002	.002	.002	.003	.002	.002	
	NK	.077	.080	.074	.075	.071	.068	HSD	.003	.004	.003	.004	.003	.003	
0.05	15								NK	.006	.007	.006	.006	.005	.005
	LSD	.049	.049	.052	.051	.050	.047	0.01	LSD	.010	.009	.010	.009	.010	.009
	S	.015	.015	.015	.015	.015	.014		S	.002	.002	.002	.002	.002	.002
	HSD	.019	.019	.019	.019	.020	.018		HSD	.003	.003	.003	.003	.003	.003
NK	.037	.036	.036	.034	.033	.030	NK		.007	.005	.006	.006	.006	.004	
0.01	LSD	.009	.010	.010	.010	.009	.010	0.01	LSD	.009	.010	.010	.010	.009	.010
	S	.002	.002	.002	.002	.002	.002		S	.002	.002	.002	.002	.002	.002
	HSD	.003	.003	.003	.003	.003	.003		HSD	.003	.003	.003	.004	.003	.003
	NK	.006	.006	.006	.006	.005	.005		NK	.006	.006	.006	.006	.005	.005

Sin embargo, no podemos afirmar lo mismo con respecto a la potencia de prueba, ya que la corrección de la depen-

dencia la infraestima notablemente sea cual sea la configuración de nuestras medias.

Tabla 10

Proporciones de potencia por comparación promediados a través de los difentes patrones de medias una vez corregida la dependencia

		10						0.05						0.01																										
Prueba		0.00	0.05	0.10	0.15	0.20	0.30	LSD	.570	.374	.266	.208	.169	.126	S	.453	.276	.191	.145	.115	.084	HSD	.481	.301	.211	.161	.130	.095	NK	.563	.365	.258	.201	.163	.121					
0.10	LSD	.523	.398	.327	.275	.240	.197	LSD	.664	.419	.291	.224	.182	.132	S	.554	.321	.214	.161	.128	.092	HSD	.575	.342	.230	.174	.140	.101	NK	.660	.412	.282	.217	.176	.127					
	S	.405	.295	.235	.195	.167	.135																																	
	HSD	.430	.318	.255	.213	.184	.149																																	
	NK	.517	.391	.319	.268	.234	.191																																	
0.10	LSD	.677	.490	.376	.308	.262	.204																																	
	S	.553	.375	.277	.221	.184	.140																																	
	HSD	.579	.400	.300	.243	.202	.156																																	
	NK	.673	.483	.369	.301	.256	.198																																	
0.10	LSD	.753	.539	.405	.328	.278	.213																																	
	S	.640	.418	.300	.236	.196	.147																																	
	HSD	.662	.444	.323	.257	.214	.163																																	
	NK	.751	.533	.398	.321	.271	.207																																	
		15						19																																
Prueba		0.00	0.05	0.10	0.15	0.20	0.30	Prueba	0.00	0.05	0.10	0.15	0.20	0.30																										
0.05	LSD	.402	.283	.219	.179	.152	.119	0.01	LSD	.196	.114	.080	.059	.047	.035																									
	S	.300	.204	.152	.123	.102	.079		S	.138	.078	.054	.038	.030	.022																									
	HSD	.324	.223	.169	.138	.115	.089		HSD	.155	.089	.062	.045	.035	.026																									
	NK	.394	.276	.212	.173	.146	.114		NK	.188	.108	.076	.055	.043	.032																									
0.05	LSD	.350	.178	.110	.077	.058	.040	0.01	LSD	.464	.215	.126	.086	.065	.041																									
	S	.263	.123	.074	.050	.038	.025		S	.363	.152	.085	.057	.042	.025																									
	HSD	.288	.140	.085	.058	.045	.030		HSD	.393	.172	.098	.065	.049	.030																									
	NK	.340	.169	.104	.072	.054	.037		NK	.455	.206	.118	.080	.060	.037																									
0.01	LSD	.464	.215	.126	.086	.065	.041																																	
	S	.363	.152	.085	.057	.042	.025																																	
	HSD	.393	.172	.098	.065	.049	.030																																	
	NK	.455	.206	.118	.080	.060	.037																																	

Tabla 11

Proporciones de potencia por comparación promediados a través de los difentes patrones de medias una vez corregida la dependencia (sin realizar AVAR)

		10						0.10																											
Prueba		0.00	0.05	0.10	0.15	0.20	0.30	LSD	.708	.552	.450	.385	.341	.280	S	.558	.374	.277	.222	.187	.142	HSD	.586	.405	.305	.247	.212	.163	NK	.681	.496	.383	.316	.272	.212
0.10	LSD	.580	.471	.400	.352	.316	.267	0.10	LSD	.772	.590	.478	.398	.353	.289																				
	S	.405	.296	.234	.195	.169	.134		S	.641	.414	.300	.231	.194	.146																				
	HSD	.434	.324	.258	.217	.189	.152		HSD	.665	.445	.329	.258	.218	.167																				
	NK	.528	.405	.329	.280	.245	.200		NK	.756	.539	.413	.329	.281	.219																				

		15					
	Prueba	0.00	0.05	0.10	0.15	0.20	0.30
0.05	LSD	.465	.353	.287	.243	.214	.173
	S	.302	.204	.153	.122	.103	.079
	HSD	.332	.230	.175	.141	.121	.094
	NK	.404	.286	.220	.180	.155	.121
0.05	LSD	.614	.370	.334	.275	.237	.184
	S	.458	.275	.190	.145	.119	.086
	HSD	.490	.305	.216	.168	.138	.101
	NK	.576	.373	.269	.211	.176	.130
0.05	LSD	.690	.477	.360	.287	.244	.189
	S	.557	.319	.215	.157	.127	.090
	HSD	.580	.344	.237	.176	.144	.104
	NK	.666	.418	.294	.221	.182	.133

		19					
	Prueba	0.00	0.05	0.10	0.15	0.20	0.30
0.01	LSD	.253	.164	.119	.093	.080	.058
	S	.137	.080	.053	.040	.032	.023
	HSD	.160	.096	.065	.049	.040	.028
	NK	.194	.116	.079	.060	.050	.035
0.01	LSD	.410	.233	.154	.115	.092	.065
	S	.266	.124	.073	.050	.038	.025
	HSD	.301	.147	.089	.063	.049	.032
	NK	.354	.178	.108	.076	.059	.039
0.01	LSD	.505	.202	.173	.124	.097	.067
	S	.361	.151	.085	.056	.041	.026
	HSD	.397	.176	.103	.069	.051	.034
	NK	.460	.210	.124	.083	.062	.041

Discusión

Para terminar, y de una manera muy breve vamos a proceder a realizar una comparación entre los resultados que presentamos en este trabajo y los hallados por otros autores, obviamente, en aquellas situaciones en que se hayan utilizado diseños y procedimientos similares.

Los resultados presentados por Scariano y Davenport (1987) acerca de los efectos del incumplimiento del supuesto de independencia en el AVAR, coinciden con los reseñados por nosotros, claro está, cuando las variables implicadas son las mismas.

En otra publicación, Seaman, Levin y Serlin (1991) muestran la diferencia entre diversos procedimientos de comparación múltiple. Nuevamente, encontramos una coincidencia con nuestros resultados cuando las variables manejadas (patrones de medias, tamaños de muestra y $r=0$)

por estos autores son las mismas que las nuestras; en concreto, para el caso de potencia por comparación en ausencia de dependencia para los procedimientos *HSD* y *LSD*.

Con respecto a la tabla 1, donde se presentaban las verdaderas probabilidades de error tipo I de la prueba de Scheffé bajo H_0 , hemos de decir que los resultados producto de la simulación presentan una tasa de error menor que los de la derivación analítica.

Por último, queremos resaltar que la presencia de dependencia en los datos es lo peor que le puede ocurrir a un investigador, ya que si bien bajo hipótesis nula verdadera el procedimiento desarrollado para corregir la correlación parece comportarse honradamente, no podemos afirmar lo mismo cuando la hipótesis alternativa es la verdadera; esto es, cuando las medias presentan configuración de rango mínimo, máximo e igualmente espaciadas

Referencias

- Carmer, S.G. y Swanson, M.R. (1973). An evaluation of ten multiple comparison procedures by Monte Carlo methods. *Journal of the American Statistical Association*, 68, 66-74.
- Dunnett, C.W. (1982). Robust multiple comparisons. *Communications in Statistics, Simulation and Computation*, 11, 2611-2624.
- Einot, I. y Gabriel, K.R. (1975). A study of the power of several methods of multiple comparisons. *Journal of the American Statistical Association*, 70, 574-583.
- Fisher, R.A. (1935). *The Design of Experiments*. Edinburgh and London: Oliver and Boyd.
- Gauss (1988). Washington: Aptech Systems, Inc.
- Hayter, A.J. (1986). The maximum familywise error rate of Fisher's least significant difference test. *Journal of the American Statistical Association*, 81, 1000-1004.
- Hochberg, Y. y Tamhane, A.C. (1987). *Multiple Comparison Procedures*. New York, NY: John Wiley.
- Keselman, H.J. y Rogan J.C. (1978). A comparison of modified-Tukey and Scheffé methods of multiple comparisons for pairwise contrasts. *Journal of the American Statistical Association*, 73, 47-51.
- Keuls, M. (1952). The use of the 'Studentized range' in connection with and analysis of variance. *Euphytica*, 1, 112-122.
- Kinderman, A.J. y Ramage, J.G. (1976). Computer generation of normal random numbers. *Journal of the American Statistical Association*, 71, 893-896.
- Kirk, S.A. (1990). Multiple comparisons: From Fisher to the future. *Paper presented at the annual meeting of the American Educational Research Association*. Boston.
- Klockars, A.J. y Hancock, G.R. (1994). Per-experiment error rates: The hidden costs of several multiple comparison procedures. *Educational and Psychological Measurement*, 54, 292-298.
- Newman, D. (1939). The distribution of the range in samples from a normal population expressed in terms of an independent estimate of standard deviation. *Biometrika*, 31, 20-30.
- Pavur, R. (1988). Type I error rates for multiple comparison procedures with dependent data. *American Statistician*, 42, 171-174.
- Pavur, R. y Lewis, T.D. (1983). Unbiased *F* test for factorial experiments for correlated data. *Communications in Statistics, Theory and Methods*, 12, 829-840.
- Petrinovich, L.F. y Hardyck, C.D. (1969). Error rates for multiple comparison methods. *Psychological Bulletin*, 71, 43-54.
- Ramsey, P.H. (1978a). Power differences between pairwise multiple comparisons. *Journal of the American Statistical Association*, 73, 479-485.
- Ramsey, P.H. (1978b). Rejoinder to comment on power differences between pairwise multiple comparisons. *Journal of the American Statistical Association*, 73, 487.
- Ringland, J.T. (1983). Robust multiple comparisons. *Journal of the American Statistical Association*, 78, 145-151.
- Ryan, T.A. (1959). Multiple comparisons in psychological research. *Psychological Bulletin*, 56, 26-47.
- Scariano, S.M. y Davenport, J.M. (1987). The effects of violations of independence assumptions in the one-way ANOVA. *American Statistician*, 41, 123-129.
- Scheffé, H. (1959). *The Analysis of Variance*. New York: John Wiley and Sons.
- Seaman, M.A., Levin, J.R. y Serlin, R.C. (1991). New developments in pairwise multiple comparisons: Some powerful and practicable procedures. *Psychological Bulletin*, 110, 577-586.
- Toothaker, L.E. (1991). *Multiple Comparisons for Researchers*. Newbury Park, CA: Sage Publications.
- Tukey, V.W. (1953). *The Problem of Multiple Comparisons*. Mimeographed Monograph, Princeton University.
- Vallejo, G. y Fernández, P. (1990). Diseños de medidas repetidas con errores autocorrelacionados. *Psicothema*, 2, 189-209.
- Welsch, R.E. (1977). Stepwise multiple comparison procedures. *Journal of the American Statistical Association*, 72, 566-575.