Recibido, 20/02/2011 Aceptado, 21/07/2011 Artículo original

Estudio de métodos de prueba en el análisis de regresión lineal multivariado usando correcciones tipo Bartlett

Study of test methods in multivariate linear regression analysis using type corrections Bartlett

[1]Aouiqw Ascanio[a], Manuel Milla[a], Franklin Chacín[a], Margarita Cobo[a], Marisela Ascanio.[a]

RESUMEN

Para evaluar el efecto de las correcciones tipo Bartlett sobre los estadísticos de prueba de hipótesis en el análisis de regresión lineal multivariado, se remuestrearon 101 casos con cinco factores y tres variables, generándose 1000 muestras de diferentes tamaños, bajo modelos: completo y reducido. Con cada muestra se ejecutaron las pruebas de razón de verosimilitud (LR), Wald (W) y multiplicador de Lagrange (LM) y pruebas adicionales con los estadísticos ajustados. Se compararon las tasas de error tipo I y II y las potencias de pruebas. Los resultados fueron: 1) El método no ajustado que presentó la menor tasa de error tipo I fue LM (aunque muy superior al nivel de significación más alto involucrado), para los dos modelos y tamaños de muestra utilizados. La tasa de error tipo II fue baja y la potencia alta en las condiciones evaluadas; 2) El método ajustado que presentó la menor tasa de error tipo I fue W; aunque la respectiva corrección ejerció control sobre ésta sólo cuando el tamaño de muestra fue grande, ya que para conjuntos pequeños a moderados produjo un descenso drástico a los niveles de significación. No obstante, para los dos modelos, tamaños de muestra y niveles de significación evaluados, la potencia fue muy baja (tasa de error tipo II alta); 3) Los métodos ajustados no fueron efectivos bajo las condiciones de trabajo, ya que no reducen consistentemente la tasa de error tipo I, salvo con el ajuste implementado sobre el estadístico W, en cuyo caso redujo demasiado su potencia.

Palabra clave: Evaluación, correcciones tipo Bartlett, pruebas, regresión, multivariado.

ABSTRACT

In order to evaluate the effect of some Bartlett type correction on statistics to prove hypothesis in multivariate linear regression analysis, 101 cases of five factors and two variables was resampling, 1000 samples of each size: 5, 10, 20, 30, 40 and 50 under two different models: complete and reduced. With each sample, three hypothesis tests (likelihood ratio, Wald and Lagrange multiplier) were done, also three adjusted test. For every one used test error type I, II and power were evaluate and compared. Results allows to see: 1) The non-adjusted method multiplication of Lagrange outcome showed the lower error rate type I, even though present a significant lever more higher than the high level tested, for the two models and sample size used in the comparison. For all conditions tested, error rate type II was low and its power high; 2) The adjusted method submitting lower error rate type I was Wald method, although, the corresponding correction bring to control over error rate type I only if sample size is big enough. For sample size small to medium, correction falls drastic with all levels of significance. However, the two models studied in this approach, sample sizes and levels of significance, the test power corresponding was very low (error rate type II high); 3) The adjusted methods results non effective under work conditions used here, they do not reduces rate error type I, unless adjusted over non-adjusted Wald's statistics, in this case power test was so low that is consider unacceptable.

Keywords: Evaluation, Bartlett type corrections, tests, multivariate, regression.

[1] Facultad de Agronomía, Instituto de Ingeniería Agrícola, Universidad Central de Venezuela, [2]Inst. Universitario de Tecnología de Yaracuy, Dpto. Agrícola, Venezuela. [a] Ing. Agrícola

INTRODUCCIÓN

Los procedimientos de razón de verosimilitud, de Wald y del multiplicador de Lagrange, han sido utilizados en el análisis de regresión lineal multivariado, con el fin de probar la hipótesis que un conjunto de variables (dependientes) tienen una relación funcional definida en un grupo de factores (o variables independientes). Estos procedimientos utilizan estadísticos de prueba cuya distribución límite es ji-cuadrado con un número de grados de libertad dado por la cantidad de restricciones impuestas sobre la hipótesis a probar. Bajo estas circunstancias, para un determinado conjunto de datos, los tres métodos generan la misma región crítica; sin embargo, por sus expresiones de cálculo, es posible que los tres estadísticos sean numéricamente desiguales. Esta situación puede originar un serio conflicto al momento de decidir si se rechaza o no la hipótesis bajo consideración, al haber inconsistencia en las tres pruebas realizadas.

Además de lo planteado, persiste el problema que las pruebas producen un número grande de rechazos errados, lo cual significa que la tasa de error tipo I es mayor que el nivel de significación correspondiente. En la literatura revisada, se ha propuesto algunos *ajustes* en los estadísticos de prueba, que dependen del tamaño de las muestras y el número de variables y factores. Estas correcciones, que suelen llamarse *tipo Bartlett*, han logrado reducir de manera notable, el problema de los rechazos errados exagerados de la hipótesis, en conformidad con los resultados de Cribari-Neto y Zarkos (1995). Por otro lado, Dufour y Khalaf (2000), quienes trabajaron con muestras aleatorias simuladas, señalan que en sistemas complejos, en términos del número de variables y factores involucrados, la corrección tipo Bartlett no es suficientemente efectiva en el control del rechazo exagerado de la hipótesis nula.

Ante tal situación, surge la necesidad de medir cuánto es el beneficio, en términos de la reducción de la tasa de error tipo I, que se obtiene con la aplicación de estas correcciones y de evaluar el comportamiento de la tasa de error tipo II, bajo diferentes tamaños de muestra y por consiguiente, de la potencia de las pruebas.

MATERIALES Y MÉTODOS

El modelo clásico de regresión lineal multivariado de acuerdo a Berndt y Savin (1977), suele plantearse como:

$$y(t) = \Pi x(t) + \varepsilon(t); 1 \leq t \leq T$$

donde: $y'(t) = [y_1(t) \; \cdots \; y_m(t) \; \ldots \; y_M(t)]$ es un vector $1xM$ de variables, $\Pi = [\pi_{km}]$ es una matriz KxM de parámetros desconocidos, $x'(t) = [x_1(t) \; \ldots \; x_K(t)]$ es un vector $1xK$ de factores y $\varepsilon'(t) = [\varepsilon_1(t) \; \ldots \; \varepsilon_m(t) \; \ldots \; \varepsilon_M(t)]$ es un vector $1xM$ de errores. Las T ecuaciones (i) pueden escribirse de forma general matricial como:

$$Y = X\Pi + \varepsilon$$

Las suposiciones básicas que se establecen con el modelo son: X es una matriz no estocástica de rango completo (rango de $X = K<T$) y los $\varepsilon(t)$ son independientes, normalmente distribuidos con media vector cero y matriz de covarianzas Σ desconocida. $\hat{\Pi} = \left[\hat{\pi}_{km}\right]$ es una matriz KxM de parámetros estimados. Con base en las estimaciones, la matriz de residuales es: $\hat{\varepsilon} = Y - X\hat{\Pi} = \left[\hat{\varepsilon}_m(t)\right]$ y la matriz de sumas de cuadrados y productos cruzados de residuales estará dada por $W = \left(Y - X\hat{\Pi}\right)'\left(Y - X\hat{\Pi}\right)$

Por lo general, se elige a $P = (X'X)^{-1}(X'X)$ como expresión de cálculo de $\hat{\Pi}$. Se denota la m-ésima columna de P por p_m y la m-ésima columna de Y por y_m. Luego, $p_m = (X'X)^{-1} X' y_m$ es obtenida por regresión múltiple (no multivariada) de y_m en X, $m = 1, \ldots M$. Los residuales producidos así, son no correlacionados y denotados por $\hat{E} = Y - XP = [e_m(t)]$, donde la m-ésima columna de $\hat{E}$ es $e_m = y_m - Xp_m$, $m = 1,\ldots,M$. La matriz de sumas de cuadrados y productos cruzados de residuales no correlacionados, vendrá dada por:

$$S = \hat{E}'\hat{E} = (Y - XP)'(Y - XP)$$

donde S es definida positiva. En muchos casos, una hipótesis es formulada en términos de restricciones lineales en los coeficientes de regresión poblacionales. En tales casos, tomarán la forma especial de restricciones lineales uniformes (RLU), las cuales se utilizan en caso que se realicen transformaciones idénticas (individuales o cruzadas) sobre los coeficientes de regresión o se plantee la hipótesis que un elemento simple de Π es igual a cero: $F\Pi G = D$

donde F es una matriz JxK (de rango J, $\leq JK$), G es una matriz MxN (de rango N, $\leq NM$) y D es una matriz JxN, todas conocidas y de constantes. F y G asignan valores particulares a los coeficientes de regresión relativos a los factores y las variables, en orden respectivo.

Se necesita probar la hipótesis $F\Pi G = D$ contra $F\Pi G \neq D$. Considérese la estimación de Π sujeta a $F\Pi G=D$. El método de estimación por máxima verosimilitud permite elegir un $\hat{\Pi}$ que minimice $|W|$ sujeto a $F\Pi G=D$, mientras que el método de la distancia mínima restringida de Zellner, permite la selección de un $\hat{\Pi}$ que minimice $tr\,(S^{-1}W)$ sujeto a $F\Pi G=D$. Bajo la condición de uniformidad, ambos métodos coinciden en la escogencia de $\hat{\Pi}$, la cual es llamada:

$$P_0=P-(X'X)^{-1}F'\left[F(X'X)^{-1}F'\right]^{-1}(FPG-D)[G'SG]^{-1}G'S$$

Cuando se elige P_0, W resulta:

$$W_0 = S + SG(G'SG)^{-1}(FPG-D)'\left[F(X'X)^{-1}F'\right]^{-1}(FPG-D)(G'SG)^{-1}G'S$$

El estadístico lambda de Wilks (utilizado en la prueba de razón de verosimilitud), se denota Λ, y se define como:

$$\Lambda = \frac{|S|}{|W_0|} = \frac{|S|}{|S+H|}$$

Si $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_m \geq 0$ son las raíces de $|H-\lambda S|=0$ donde $H=W_0-S$ y

$$A = (FPG-D)'\left[F(X'X)^{-1}F'\right]^{-1}(FPG-D)$$

$$B = (G'SG)$$

se tiene que las raíces no nulas de $S^{-1}H$ son las mismas que las raíces no nulas de $B^{-1}A$, dado que las raíces del producto de dos matrices son independientes del orden de multiplicación. De esta manera, el criterio de prueba Λ(estadístico de razón de verosimilitud), puede escribirse como:

$$\Lambda = \frac{|B|}{|B+A|} = \prod_{q=1}^{Q}(1+\lambda_q)^{-1}$$

donde $Q=\min,(NJ)$ es el número de raíces no nulas de $|A-\lambda B|=0$. Cuando la hipótesis $F\Pi G=D$ es cierta, la variable Λ se denota $\Lambda\ (N,J,T-K)$. La región crítica para la prueba de $F\Pi G=D$ contra $F\Pi G\neq D$ a un nivel de significación α es $\Lambda<\Lambda_\alpha(N,J,T-K)$, donde $\Lambda_\alpha(N,J,T-K)$ es el cuantil asociado a una probabilidad α de la distribución $\Lambda\ (N,J,T-K)$. Cuando la hipótesis $F\Pi G=D$ es cierta, $-T\ln\Lambda = T\sum_{q=1}^{Q}\ln(1+\lambda_q)$ tiene distribución límite χ_2 con NJ grados de libertad. La región crítica asintótica para la prueba, con un nivel de significación α, es $-T\ln\Lambda \geq \chi^2_{\alpha(NJ)}$

donde $\chi^2_{\alpha(NJ)}$ es el cuantil asociado a una probabilidad α de la distribución $\chi^2_{(NJ)}$

El estadístico de Wald se define como:

$$tr\left(S^{-1}W_0\right)-tr\left(S^{-1}S\right)=tr\left(S^{-1}H\right)=tr\left(AB^{-1}\right)=\sum_{q=1}^{Q}\lambda_q$$

$U(N,J,T-K)$ es la variable $tr\ (B^{-1}A)$ cuando la hipótesis $F\Pi G=D$ es cierta. $TU(N,J,T-K)$ tiene distribución límite $\chi^2_{(NJ)}$. La región crítica asintótica de la prueba de Wald es $T\,tr\left(S^{-1}H\right)\geq \chi^2_{\alpha(NJ)}$

El estadístico multiplicador de Lagrange, se define como:

$$tr\left(W_0^{-1}W_0\right)-tr\left(W_0^{-1}S\right)=tr\left(W_0^{-1}H\right)=tr\left[A(A+B)^{-1}\right]=\sum_{q=1}^{Q}\lambda_q\Big/(1+\lambda_q)=\sum_{q=1}^{Q}\theta_q$$

La variable aleatoria $tr\ (W_0^{-1}H)$ cuando la hipótesis $F\Pi G=D$ es cierta, es denotada por $V(N,J,T-K)$. La hipótesis puede ser probada comparando el valor de $tr\ (W_0^{-1}H)$ con $V\alpha(N,J,T-K)$, el cuantil asociado a una probabilidad α en la distribución de $V(N,J,T-K)$.

La distribución límite de $TV(N,J,T-K)$ es $\chi^2_{(NJ)}$. La región crítica asintótica para la prueba de multiplicador de Lagrange es $T\,tr\left(W_0^{-1}H\right)\geq \chi^2_{\alpha(NJ)}$

Las versiones corregidas de los estadísticos por el tamaño de las muestras, número de factores y cantidad de restricciones impuestas sobre la hipótesis nula, generan las llamadas *pruebas ajustadas* o *métodos ajustados*, los cuales fueron sugeridos por Cribari-Neto y Zarkos (1995). Los estadísticos corregidos son los siguientes:

$$LR^* = \underline{LR}\left[1-\frac{N-J+1}{2(T-K)}\right]$$

$$W^* = \underline{W}\left[1-\left(\frac{N-J+1}{2(T-K)}+\frac{N+J+1}{2(T-K)(JN+2)}\underline{W}\right)\right]$$

$$LM^* = \underline{LM}\left[1-\left(\frac{N-J+1}{2(T-K)}-\frac{N+J+1}{2(T-K)(JN+2)}\underline{LM}\right)\right]$$

donde $\underline{LR}$, $\underline{W}$ y $\underline{LM}$ son los estadísticos no ajustados de razón de verosimilitud, de Wald y multiplicador de Lagrange, respectivamente. De las tres expresiones, la primera es conocida como *corrección de Bartlett*, mientras que las dos últimas son *correcciones tipo Bartlett*. Los estadísticos de prueba ajustados, tienen distribución $\chi^2_{(NJ)}$ bajo la hipótesis $F\Pi G=D$. Las regiones críticas son las mismas que para las pruebas realizadas con los estadísticos no ajustados.

Se dispuso de una matriz de datos provenientes de 101 mediciones registradas en el banco de germoplasma de café (población) de la Estación Experimental Jaime Henao Jaramillo de la Facultad de Agronomía, Universidad Central de Venezuela. Esta matriz cuenta con cinco factores (K=5), tres variables (M=3) y 101 casos (T=101). Los factores y las variables son las siguientes: X_1: número de flores por nudo; X_2: longitud de los entrenudos (en centímetros); X_3: altura de la planta (en metros); X_4: ángulo de la rama con el tallo (en grados); X_5: diámetro de la planta (en metros); Y_1: rendimiento (en Kilogramos por planta); Y_2: rendimiento en beneficio; Y_3: porcentaje de granos vanos.

Los valores determinados de Π, calculados por $\Pi = (X'X)^{-1}(X'Y)$, son obtenidos a partir de la matriz original de datos, utilizando el programa Statistix (Analytical Software) versión 8.0 y los valores falsos son obtenidos a partir del siguiente criterio: Se construye con los casos de la población, un intervalo al 95% de confianza para cada coeficiente de regresión o cada elemento de Π. La distancia absoluta que se establece entre los límites del intervalo y su centro, para cada caso, es utilizada para localizar un punto, sustrayendo esta cantidad al límite inferior, donde está ubicado el valor falso del parámetro. Estos valores se han obtenido por aplicación de las fórmulas correspondientes al procedimiento descrito sobre la matriz de parámetros determinados, en una hoja de cálculo preparada en Microsoft Office Excel 2003.

Se prepararon 1000 muestras aleatorias cada una de tamaño n dentro de la matriz original de datos, a través del método *bootstrap*. La selección de las muestras fue ejecutada por el programa Resampling Stats for Excel (2006 Resampling Stats, Inc) versión W3.20 (Resampling Stats, Inc 2006).

Los seis tamaños (n) de muestra son n_1=5, n_2=10, n_3=20, n_4=30, n_5=40 y n_6=50. Cada uno es aplicado para las dos especificaciones del modelo lineal multivariado que se describen ahora. El modelo lineal multivariado es, en ambos casos: $Y=X\Pi+\underline{\varepsilon}$

donde: Y es la matriz $nx3$ de variables; X es la matriz $nx5$de factores; Π es la matriz $5x3$ de parámetros determinados o falsos, según el tipo de error a medir y ε es la matriz $nx3$ de errores. La primera relación funcional que se plantea (modelo completo), es la siguiente: $Y_1, Y_2, Y_3 = f(X_1,X_2,X_3,X_4,X_5)$ y la segunda (modelo reducido) es: $Y_1, Y_2, = f(X_1,X_3,X_5)$

La hipótesis $F\Pi G=D$, en ambos casos, involucra a F y G como matrices de identidad JxK (donde $J=K$) y MxN (donde $M=N$), respectivamente. En los dos casos, el número de restricciones sobre los factores (J) es igual al número de ellos (K) y de restricciones sobre las variables (N) es igual al número de éstas (M); D es una matriz JxN (ó KxM bajo estas condiciones) que contiene los valores determinados o falsos, según el tipo de error a medir, de Π. Los niveles de significación (α) a los que son conducidas las pruebas son tres: α=0,1, α=0,05 y α=0,01.

Las hojas de cálculo para las expresiones de los estadísticos de prueba fueron creadas en Microsoft Office Excel 2003, con la adición del complemento ASAP Utilities (1999-2006 Bastien Mensink – eGate Internet Solutions) versión 3.10c.

DISCUSIÓN

Las tablas 1 y 2 muestran los resultados de las mediciones de la tasa de error tipo I y II y de la potencia de las pruebas para cada modelo, nivel de significación y tamaño de muestra.

En general, las pruebas realizadas con los estadísticos no ajustados producen números muy altos de rechazos de la hipótesis $F\Pi G=D$, planteada con los valores determinados o verdaderos y falsos de Π, lo cual implica que mientras la potencia de las correspondientes pruebas se mantiene muy elevada, también lo hace la tasa de error tipo I. Lo último sin embargo, parece controlable mientras se utilice los procedimientos con muestras de gran tamaño, mayor que 50 (tablas 1 y 2).

Comparación de las tasas de error tipo I

En todos los casos evaluados se mantuvo la relación de orden entre las pruebas realizadas con los estadísticos no ajustados, descrita por Savin (1976); Berndt y Savin (1977); Breusch (1979) y Evans y Savin (1982), según la cual se produce mayor número de rechazos equivocados de la hipótesis planteada con la prueba de Wald. Con muestras de tamaño cinco no fue posible la aplicación de correcciones sobre los estadísticos de prueba, ya que $T-K$=0 en las expresiones de cálculo. Como se observa en la tabla 1, la tasa de error tipo I es uno para las tres pruebas realizadas con los estadísticos no ajustados, con todos los niveles de significación.

Para el modelo completo y con muestras de tamaño 10, se sigue observando tasas iguales a uno de error tipo I para todas las pruebas realizadas con los estadísticos no ajustados, independientemente del nivel de significación utilizado. Sin embargo, se hace notar el efecto del factor de corrección aplicado sobre los estadísticos: Para todos los valores de α, se aprecia una disminución en los valores de la razón de rechazos equivocados de la hipótesis propuesta, sobre todo con el método ajustado de Wald, donde el valor es cero. En este caso, evidentemente, la corrección no resultó adecuada, pues todos los valores del estadístico fueron transformados a negativos y, por tanto jamás se produjo rechazo. La causa de este hecho es la siguiente: Con muestras de tamaño 10 la función de corrección correspondiente al estadístico de Wald tiene una raíz en 20,794.

Si el estadístico no ajustado de Wald es menor que 20,794, el estadístico ajustado resulta positivo; mientras que si es mayor que dicho valor, el estadístico ajustado resulta negativo. En los casos señalados, todas las muestras produjeron un valor del estadístico no ajustado mucho mayor que este y por lo tanto, todos fueron cambiados de signo durante el ajuste. Es conveniente señalar que la prueba realizada sin ajuste correspondiente, había arrojado para todas las muestras analizadas, un valor del estadístico mucho mayor que el número crítico según su distribución.

Tabla 1: Errores tipo I y II cometidos y potencias de las seis pruebas, bajo cada nivel de significación y tamaño de muestra para el modelo completo.

		$\alpha = 0,1$						$\alpha = 0,05$						$\alpha = 0,01$					
n	Tasa	LR	W	LM	LR*	W*	LM*	LR	W	LM	LR*	W*	LM*	LR	W	LM	LR*	W*	LM*
5	Tipo I	1,00	1,00	1,00				1,00	1,00	1,00				1,00	1,00	1,00			
	Tipo II	0,00	0,00	0,00				0,00	0,00	0,00				0,00	0,00	0,00			
	Potencia	1,00	1,00	1,00				1,00	1,00	1,00				1,00	1,00	1,00			
10	Tipo I	1,00	1,00	1,00	0,91	0,00	0,91	1,00	1,00	1,00	0,91	0,00	0,91	1,00	1,00	1,00	0,91	0,00	0,91
	Tipo II	0,00	0,00	0,00	0,09	1,00	0,09	0,00	0,00	0,00	0,09	1,00	0,09	0,00	0,00	0,00	0,09	1,00	0,09
	Potencia	1,00	1,00	1,00	0,91	0,00	0,91	1,00	1,00	1,00	0,91	0,00	0,91	1,00	1,00	1,00	0,91	0,00	0,91
20	Tipo I	0,97	0,98	0,95	0,98	0,00	0,99	0,95	0,97	0,91	0,96	0,00	0,98	0,89	0,94	0,90	0,90	0,00	0,95
	Tipo II	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00
	Potencia	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00
30	Tipo I	0,86	0,90	0,77	0,86	0,44	0,89	0,79	0,86	0,69	0,81	0,00	0,84	0,65	0,75	0,47	0,67	0,00	0,71
	Tipo II	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00
	Potencia	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00
40	Tipo I	0,55	0,65	0,44	0,56	0,48	0,58	0,46	0,56	0,32	0,48	0,36	0,50	0,29	0,42	0,17	0,30	0,15	0,31
	Tipo II	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00
	Potencia	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00
50	Tipo I	0,35	0,44	0,25	0,36	0,33	0,36	0,27	0,36	0,15	0,27	0,24	0,27	0,13	0,22	0,05	0,13	0,10	0,12
	Tipo II	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00
	Potencia	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00

Tabla 2: Errores tipo I y II cometidos y potencias de las seis pruebas, bajo cada nivel de significación y tamaño de muestra para el modelo reducido.

		$\alpha = 0,1$						$\alpha = 0,05$						$\alpha = 0,01$					
n	Tasa	LR	W	LM	LR*	W*	LM*	LR	W	LM	LR*	W*	LM*	LR	W	LM	LR*	W*	LM*
5	Tipo I	1,00	1,00	1,00				1,00	1,00	1,00				1,00	1,00	1,00			
	Tipo II	0,00	0,00	0,00				0,00	0,00	0,00				0,00	0,00	0,00			
	Potencia	1,00	1,00	1,00				1,00	1,00	1,00				1,00	1,00	1,00			
10	Tipo I	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00
	Tipo II	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00
	Potencia	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00
20	Tipo I	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00
	Tipo II	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00
	Potencia	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00
30	Tipo I	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00
	Tipo II	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00
	Potencia	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00
40	Tipo I	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00
	Tipo II	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00
	Potencia	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00
50	Tipo I	0,23	0,28	0,19	0,23	0,23	0,23	0,16	0,20	0,12	0,16	0,16	0,16	0,06	0,10	0,03	0,06	0,05	0,05
	Tipo II	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00
	Potencia	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00

La función de corrección correspondiente al estadístico del multiplicador de Lagrange tiene dos raíces: una en -20,794 y otra en cero. Si el estadístico no ajustado está contenido en el intervalo [-20,794; 0], el ajustado adquiere valor negativo. Esto último no ocurrió, ya que el primero, por definición, es no negativo. La función de corrección correspondiente al estadístico de razón de verosimilitud, por ser sólo un factor de corrección y no una función con raíces, no realiza cambios de signo.

Un comportamiento contrario al esperado para las pruebas ajustadas de razón de verosimilitud y multiplicador de Lagrange, se aprecia con muestras de mayor tamaño, hasta 50, donde las correcciones tipo Bartlett incrementaron los valores de los correspondientes estadísticos y se vio aumentada esta tasa de error. Sin embargo, tuvieron el efecto deseado sobre el estadístico de Wald con muestras de tamaño 40 y 50.

Aún en los casos en que la corrección resultara favorable, no se pudo llevar la tasa de error tipo I, bajo ninguna circunstancia, a valores menores o iguales al nivel de significación correspondiente, lo cual sugiere que los ajustes sugeridos por Cribari-Neto y Zarkos (1995) no son suficientemente efectivos en el control del número exagerado de rechazos errados de la hipótesis bajo consideración; además, tampoco contribuye a reducir la posibilidad de conflictos generados al momento de tomar la decisión en torno al rechazo o no de ésta, por resultar aún más diferentes las tasas de error tipo I una vez aplicados los ajustes, en ciertas ocasiones. Este hallazgo apoya lo expresado por Dufour y Khalaf (2000), sobre la insatisfacción que produce la implementación de las correcciones.

Para el modelo reducido, la tasa de error tipo I es uno para las tres pruebas sin ajustes realizadas con muestras hasta de tamaño 40 (ver tabla 2). Los ajustes sobre los estadísticos fueron inútiles, tomando en cuenta que la única disminución observada sobre la razón de rechazos errados corresponde a la prueba ajustada de Wald, con la cual se lleva a cero el valor como consecuencia de una inapropiada transformación aplicada. Para el modelo reducido y con muestras de tamaño 50, las tasas de error tipo I resultan las más bajas obtenidas en este trabajo, pero aún son altas si se comparan con el nivel de significación respectivo. En este caso, sólo se logra mejoras con el método ajustado de Wald. Con el procedimiento ajustado del multiplicador de Lagrange, en cambio, se obtiene tasas más elevadas, en relación a su correspondiente equivalente no ajustado. Una vez más, no resultan satisfactorios los efectos de las correcciones tipo Bartlett. Sin embargo, aunque esto guarda concordancia con lo expresado por Dufour y Khalaf (2000), se ha encontrado que con un sistema más complejo de variables y factores, en términos del número de éstos, se logra mayores efectos. Con muestras de tamaño 50, bajo el modelo completo se obtiene una reducción del 25%, para 0,1a=, con la aplicación del método ajustado de Wald, respecto al equivalente no ajustado. Bajo el modelo reducido, con el mismo tamaño de muestra, esta reducción sólo es de 18%. De manera que las correcciones serían precisamente convenientes cuando los sistemas son complejos, con gran número de variables y factores, para controlar el error tipo I.

Comparación de las tasas de error tipo II y potencias

Para el modelo completo, salvo por aquellos casos en que la tasa de error tipo I fue cero en la prueba ajustada de Wald, y por el control ejercido sobre la razón de rechazos errados de la hipótesis bajo consideración por las pruebas ajustadas de razón de verosimilitud y multiplicador de Lagrange con muestras de tamaño 10 y todos los niveles de significación, la tasa de error tipo II es cero (ver tabla 1). De igual manera para el modelo reducido, con la excepción de aquellos casos en que la tasa de error tipo I fue cero en la prueba ajustada de Wald, la razón de no rechazos equivocados de la hipótesis planteada es cero (ver tabla 2).

Para el modelo completo, los métodos no ajustados son muy potentes; la potencia es uno, en todos los casos, bajo la especificación de la hipótesis alternativa hecha en este trabajo, lo cual significa que las pruebas de razón de verosimilitud, de Wald y del multiplicador de Lagrange tienen una alta capacidad de producir rechazos acertados de la hipótesis nula planteada, en base al modelo elegido (ver tabla 1).

Con los métodos ajustados de razón de verosimilitud y del multiplicador de Lagrange y muestras de tamaño 10, se reduce un poco el valor de las potencias de las pruebas, lo cual es un costo por un ajuste inútil sobre los correspondientes estadísticos, que reduce la tasa de error tipo I hasta 0,91, valor que excede por mucho al más alto de los niveles de significación utilizados. Se aceptaría una disminución en la potencia, a cambio de una disminución justa de la razón de rechazos errados o error tipo I. Conviene llamar la atención en el sentido que la tasa de error tipo I se reduce, conforme aumenta el tamaño de las muestras, de acuerdo con los resultados observados. La potencia se mantiene alta dentro del rango evaluado, lo que implica que en estos términos, los estadísticos utilizados se ajustan adecuadamente a la distribución límite descrita. De igual manera para el modelo reducido, la potencia en todos los casos es uno, salvo por la correspondiente a la prueba ajustada de Wald, con todos los tamaños de muestra y los niveles de significación involucrados (ver tabla 2).

Tomando en cuenta los dos aspectos, error tipo I y potencia, la prueba más conveniente bajo las condiciones expuestas aquí, es el método del multiplicador de Lagrange, ya que presenta la menor tasa de error tipo I (dentro de los no ajustados), mientras que mantiene su potencia alta, independientemente del tamaño de muestra y nivel de significación. Ahora bien, para ambos modelos (completo y reducido), el mínimo tamaño de muestra ($5n=$) es adecuado en términos de potencia, para ejecutar el método no ajustado del multiplicador de Lagrange. Sin embargo, la mínima tasa de error tipo I hallada en estos análisis para esta prueba, fue aún mayor que el correspondiente nivel de significación.

CONCLUSIONES

1. Las correcciones tipo Bartlett no mostraron el comportamiento esperado, ya que no resulta satisfactorio el control ejercido sobre la tasa de error tipo I por las razones siguientes: 1) el control fue ejercido prácticamente sólo con la prueba ajustada de Wald, con serias consecuencias, ya que se reduce drásticamente. Aunque su efecto mejoraba con el aumento del tamaño de muestra; 2) no hubo control consistente con los otros métodos y 3) cuando hubo control, éste no fue suficiente, ya que no colocó la tasa de error tipo I debajo del nivel de significación correspondiente.

2. El efecto de las correcciones aplicadas a los estadísticos de prueba sobre la tasa de error tipo II y, consecuentemente sobre la potencia, es negativa porque: 1) aplicadas sobre los estadísticos de razón de verosimilitud y del multiplicador de Lagrange con muestras pequeñas, disminuyen la potencia de las pruebas, a cambio de un efecto no satisfactorio en la reducción de la tasa de error tipo I y 2) utilizadas sobre el estadístico de Wald, produjo una reducción drástica en la potencia.

3. Es necesario estudiar diferentes condiciones teóricas y prácticas que permitan predecir su funcionalidad. Mientras tanto, se recomienda sólo la implementación de los métodos no ajustados, en particular el del multiplicador de Lagrange que, de acuerdo con los resultados de este trabajo, es el más conveniente.

REFERENCIAS BIBLIOGRÁFICAS

Berndt, E. R. y N. E. Savin 1977, Conflict among criteria for testing hypotheses in the multivariate linear regression model, Econométrica 45(5), 1263-1277.

Breusch, T. 1979, Conflict among criteria for testing hypotheses: Extensions y comments, Econometrica, 47(1), 203-207.

Cribari-Neto, F y S. Zarkos 1995, Improved test statistics for multivariate regression, Economics Letters, 49(2), 113-120.

Dufour, J-M y L. Khalaf 2000, Simulation based finite y large sample tests in multivariate regressions, CIRANO: Scientific Series. Montreal, 26 p.

Evans, G.B.A y N.E. Savin 1982, Conflict among the criteria revisited: The W, LR y LM tests, Econometrica, 50(3), 737-748.

Resampling Stats, Inc. 2006, Resampling Stats add-in for Excel, User's Guide Version 3, Resampling Stats, Inc., 177 p.

Savin, N.E. 1976, Conflict among testing procedures in a linear regression model with autoregressive disturbances, Econometrica, 44(6), 1303-1315.

Correspondencia:
Aouiqw Ascanio
aascanioe@gmail.com,
memilla22@yahoo.com.mx
Venezuela