

Problemas que se plantean en la supresión de tendencias en series temporales^(*)

MIGUEL A. MARTÍNEZ ECHEVARRÍA Y ORTEGA
y JOSÉ MIGUEL PONCE NÚÑEZ

Departamento de Economía y Estadística
Universidad de Navarra

Un problema fundamental de la Estadística —en cuanto instrumento de observación de la realidad— es tratar de determinar la existencia de algún tipo de relación causal entre dos variables aleatorias. Es decir, poder establecer, desde un punto de vista empírico, que una variable es efecto de la otra.

Una primera etapa en la solución de este problema la constituye la simple determinación de relación entre dos variables, mediante el recurso al coeficiente estadístico de correlación. Esta conocida técnica estadística —si se aplica con las condiciones requeridas— establece un índice del grado de interdependencia entre esas dos variables, pero por su esencial simetría estructural, la información que suministra no es suficiente para atribuir el status causal a una de las variables estudiadas.

Una condición necesaria, aunque no suficiente —para no incurrir en la famosa falacia «post hoc ergo propter hoc»—, para que exista una relación causal entre dos variables aleatorias, es la presencia de un desfase temporal en la estructura de interdependencia. Es decir, que mientras que para valores contemporáneos de ambas variables la correlación sea nula o muy pequeña, para valores temporalmente retrasados, de una variable respecto de la otra, la correlación sea bastante elevada.

Esta manera de introducir el concepto de causalidad se fundamenta en la existencia de una asimetría temporal de la estructura de correlaciones de dos variables aleatorias.

(*) Conferencia impartida dentro del II Ciclo de Estadística Económica y Empresarial, Facultad de CC. Económicas y Empresariales, Universidad de Zaragoza, mayo, 1985.

El análisis de la matriz de correlaciones temporales se ha constituido así en etapa decisiva a la hora de investigar la naturaleza causal de la relación entre variables.

El trabajo de investigación que vamos a exponer se refiere a un aspecto muy concreto, pero importante, en el marco que acabamos de esbozar: se trata de determinar si el simple hecho de ordenar temporalmente las observaciones de una variable, introduce ya un tipo de causalidad implícita, difícilmente eliminable.

La primera aproximación a este tipo de problemas es la realizada por J. Durbin y G. S. Watson, y publicada en *Biometrika* en el año 1950, donde ponían de manifiesto las consecuencias de violar el supuesto de independencia-normalidad de las perturbaciones en el modelo de regresión MCO. Como es bien sabido, la estructura de las perturbaciones es complementaria a la estructura de las supuestas interdependencias, de tal manera que la violación del supuesto de independencia-normalidad es manifestación de una inapropiada definición del modelo de interdependencias.

Después de estudiar las aportaciones más importantes realizadas hasta ahora en el estudio de la estructura de autocorrelación en las perturbaciones, podemos afirmar que una fuente general de autocorrelación en los residuos proviene de que los residuos representan la influencia neta de las variables explicativas omitidas. Por sí misma, la exclusión de variables explicativas no introduce autocorrelación en los residuos, a no ser que las variables omitidas estén correlacionadas con las variables observadas. Y aún en este último caso podría suceder que la autocorrelación quedase compensada por interdependencias de efecto contrario. Una segunda fuente de autocorrelación en los residuos podría ser la inadecuada especificación del modelo de interdependencias. Una tercera fuente de autocorrelación residual es, finalmente, los errores de medición de la variable observada.

Los métodos seguidos para eliminar o reducir el efecto de autocorrelación en los residuos consisten, esencialmente, en restablecer la condición de independencia de los residuos, mediante la adecuada transformación de los residuos autocorrelacionados. Estos procedimientos son anteriores a la elaboración de tests para la detección de esta anomalía. El más conocido de ellos es el de Cochrane-Orcutt, descrito en *Application of least squares regressions to relationships containing autocorrelated error terms*, JASA, 1949. Posteriormente aparecieron los trabajos de Prais y Winsten (1954), Durbin (1960), Hildreth y Lu (1960), Griliches y Rao (1969), Lempers y Klock (1973), que, manteniendo la concepción básica de Cochrane-Orcutt, han añadido matizaciones y enfoques que representan un excelente medio para profundizar en la naturaleza de la estructura de la autocorrelación de los residuos.

La línea de investigación que hemos elegido viene a incidir una vez más sobre la autocorrelación de los residuos pero desde un planteamiento distinto.

El análisis de series de tiempo —en contraste con el análisis de regresión— adopta una postura agnóstica respecto de la naturaleza y el modelo de interdependencias, considerando los datos observados como realización estadística de un subyacente proceso generador de los datos observados.

El problema de correlación espúrea queda entonces planteado en el marco del análisis de series de tiempo.

Para modelar cualquier serie de tiempo, como un proceso ARMA, se requiere en primer lugar, suprimir cualquier «tendencia» que pudiera existir, es decir, hacerla estacionaria. Por otro lado es bien sabido que las series con «tendencia» presentan una fuerte correlación, lo que puede ser perturbador a la hora de establecer una relación causal.

Es costumbre generalizada atribuir la «tendencia» a una dependencia funcional del tiempo, y en consecuencia existe una lógica inclinación a suprimir esa «tendencia» regresando la serie no estacionaria observada, respecto del tiempo o de una función del tiempo.

Nuestra investigación se centra en el estudio de la supuesta naturaleza de la dependencia temporal de la tendencia observada en una serie de tiempo. Es decir, en realizar un análisis de la suposición de que la variable tiempo tiene status de variable explicativa, en lo que se refiere al comportamiento no estacionario de la serie.

Vamos a considerar un sencillo caso de tendencia lineal, tal como la siguiente:

$$y_t = \beta_0 + \beta_1 t + u_t \quad (1)$$

Esto implica que las observaciones y_t se dispersan alrededor de una línea recta: $\beta_0 + \beta_1 t$.

El método para eliminar la tendencia consiste en descomponer el modelo en dos partes —una determinística (no estacionaria) y otra estocástica (estacionaria)— mediante su ajuste MCO a la línea recta $\beta_0 + \beta_1 t$.

Planteamos ahora la siguiente situación: la verdadera subyacente estructura generadora de los datos observados se corresponde con un modelo tal como:

$$y_t = y_0 + \beta t + u_t \quad (2)$$

que tiene aparentemente la misma forma lineal que el modelo (1) aunque difiera radicalmente de él en su naturaleza estructural.

El término intersección no es un parámetro fijo, sino que depende del valor inicial y_0 .

Los residuos no son estacionarios, sino que representan una media móvil sobre los valores de la serie estacionaria $\{u_t\}$ de media cero y varianza constante σ_u^2 serialmente aleatorios.

El parámetro β representa el promedio fijo de las diferencias primeras, que se conoce con el nombre de «deriva» del proceso.

En esta situación, por un motivo que de momento no nos interesa, vamos a suponer que debido a la similitud formal de los modelos (1) y (2) —que llamaremos «modelo tendencia» y «modelo deriva», respectivamente— se produce una confusión entre ambos provocando que se intente —mediante una regresión MCO sobre el tiempo— eliminar una supuesta «tendencia», que en realidad es una «deriva».

El análisis y valoración de esta confusión proporciona a nuestro entender un interesante campo de investigación, que puede ayudar a esclarecer alguna de las cuestiones más importantes, que se insertan en el marco más general con el que hemos comenzado nuestra exposición.

Para llevar a cabo nuestra investigación es imprescindible recurrir a la simulación estadística, ya que no se dispone del suficiente número de datos reales que puedan soportar este tipo de investigación que requiere observar comportamientos aleatorios fiables de datos estadísticos MCO.

La condición imprescindible para que se puedan aplicar métodos Monte Carlo de simulación e investigación de las distribuciones de los datos supuestamente destendenciados y de los estadísticos deducidos, es que éstos no dependan de la inevitable arbitraria elección de los parámetros del modelo.

El análisis de los residuos y principales estadísticos MCO, tales como:

— las estimaciones MCO de β : $\hat{\beta} = (t't)^{-1}ty$

— la media cuadrática de residuos: $\frac{SRC}{n-1} = \frac{e'e}{n-1}$

— el coeficiente de correlación: $R^2 = \frac{y'P_0y}{y'y}$

— los estadísticos t : $t(\beta_i) = \frac{(\hat{\beta}_i - \beta) / \sqrt{\sigma^2 / t't}}{\sqrt{y'(I - P_0)y} / \sigma^2 (n - p - 1)}$

— los estadísticos de Durbin-Watson: $DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}$

— etc.

ponen de manifiesto que en nuestro caso se cumplen aquellas condiciones que justifican un empleo no perturbador de las técnicas Monte Carlo de simulación.

Para la realización concreta de nuestro estudio hemos elegido el «modelo-deriva» más sencillo, que es el constituido por un proceso «paso-aleatorio» tal como el representado por

$$y_t = y_{t-1} + \beta + \varepsilon_t$$

con «deriva» nula $\beta = 0$, con lo que queda:

$$y_t = y_{t-1} + \varepsilon_t$$

En esta situación, el verdadero valor del coeficiente de correlación R^2 será nulo, ya que hemos impuesto que la variable tiempo no ejerce ninguna influencia en la variación de los valores observados y_t .

El valor muestral del coeficiente de correlación R^2 está lógicamente sobresignificado (debido, como ya sabemos, a la presencia de autocorrelación en los residuos). Hemos juzgado conveniente realizar un estudio del valor muestral de R^2 para tratar de determinar las posibles fuentes de este error de sobresignificación.

De las dos posibles fuentes de error SEC y STC

$$R^2 = \frac{SEC}{STC} = 1 - \frac{SRC}{STC}$$

$$\frac{STC}{n-1} = \frac{\sum_{t=1}^n y_t^2}{n-1} - \frac{(\sum_{t=1}^n y_t)^2}{n(n-1)}$$

vemos que esta última es realmente responsable de una parte del error, al requerir una media muestral de y_t que es casi imposible que coincida con el verdadero valor nulo de esa media.

Hemos calculado que la magnitud promedio de la perturbación introducida en la STC al sustituir la verdadera media nula por una media muestral no nula es del orden de $1/3$.

Nelson y Kang (1984) sostienen que en la práctica esa perturbación es del orden de $2/3$. En nuestro caso, y quizá debido al mayor tamaño de las muestras empleadas, hemos obtenido valores empíricos de esa perturbación que se aproximan mucho al valor teórico de $1/3$.

La otra posible fuente de error, la SEC, o equivalentemente la SRC, aporta un error promedio que representa un 86 % de la variación total de los datos observados.

Para llegar a este resultado hemos utilizado la estimación de la función de autocovarianza de los residuos resultantes de la regresión MCO sobre un «paso aleatorio», obtenida por Chan, Hayya y Ord (1977), y mejorada por Nelson y Kang (1981), para, a partir de ahí, obtener la varianza de los residuos.

Alcanzamos así el siguiente importante resultado: Aún cuando no exista «tendencia», ni «deriva», el simple hecho de aplicar una regresión MCO sobre el tiempo, a un proceso «paso aleatorio», suprime el 86 % de la variación de los datos observados.

La siguiente fase de nuestro trabajo consiste ahora en realizar una comprobación empírica del resultado anterior.

Para ello hemos diseñado un algoritmo consistente en la simulación de cien valores de un proceso «paso aleatorio» a partir de los cuales se obtienen los parámetros, residuos y estadísticos, de la correspondiente regresión MCO sobre el tiempo.

El anterior algoritmo, escrito en lenguaje APL, ha dado lugar a un programa al que hemos designado con las siglas SIMULO, y que hemos iterado mil veces, lo cual nos ha permitido establecer los rasgos fundamentales del comportamiento probabilístico de las magnitudes MCO.

Como generador de valores pseudo-aleatorios hemos utilizado el algoritmo de Box-Muller, que permite obtener «pseudo ruido blanco gaussiano» a partir de una distribución uniforme, generada a su vez por el algoritmo en APL de Snorky.

El valor promedio simulado que hemos obtenido de la SRC 6.71 constituye una excelente aproximación al correspondiente valor teórico 6.73.

Como ya sabemos la varianza teórica total observada vale 50, tenemos entonces que la regresión sobre el tiempo reduce ficticiamente la varianza total observada en

$$50 - 6.71 = 43.29$$

que es el 86.58 % del total.

Luego los resultados empíricos simulados confirman nuestros resultados teóricos.

En lo que se refiere al coeficiente de correlación, el valor promedio simulado 0.452 se aparta algo del valor teórico 0.6. La causa de esta divergencia radica en que la situación de confusión entre modelos, que es la base de partida de nuestro experimento, conduce a admitir como cierto que el cociente de las esperanzas de sumas de cuadrados es igual a la esperanza del cociente de las sumas de cuadrados

$$\frac{E(\text{SRC})}{E(\text{STC})} \neq E\left(\frac{\text{SRC}}{\text{STC}}\right)$$

El valor de este perturbador coeficiente, próximo al 50 %, pone de manifiesto la conveniencia de ser muy cautos a la hora de establecer relaciones, causales o no, tomando como única fuente de evidencia empírica el valor de R^2 . La regresión sobre la variable tiempo suele ser la causa más frecuente de regresión espúrea.

El análisis de los valores simulados del estadístico t para la EMCO de los parámetros β_0 y β_1

media MC(100/1000) de $t(\beta_0) = -0.004$

varianza MC(100/1000) de $t(\beta_0) = 49.3$

desviación típica MC(100/1000) de $t(\beta_0) = 7.02$

frecuencia de rechazo (al 5 %) MC(100/1000) de $H_0(\beta_0 = 0) = 80 \%$

media MC (100/1000) de $t(\beta_1) = -0.15$
 varianza MC (100/1000) de $t(\beta_1) = 198.81$
 desviación típica MC (100/1000) de $t(\beta_1) = 14.1$
 frecuencia de rechazo (al 5 %) MC (100/1000) de $H_0(\beta_1 = 0) = 89 \%$

pone de manifiesto la ineficiencia de este test, con frecuencias de rechazo del 80 % y 89% para la hipótesis nula, al nivel de significación del 5 %, lo que es una prueba más de la naturaleza espúrea de esta regresión.

Nos ha parecido interesante realizar a continuación nuevos experimentos de simulación encaminados a poner de manifiesto la influencia del tamaño muestral de la serie observada sobre los anteriores resultados.

Para ello hemos realizado nueve mil simulaciones en nueve grupos, comenzando para una muestra de tamaño 90, bajando de 10 en 10 hasta una muestra de tamaño 10. Los resultados de este nuevo experimento se recogen en las siguientes tablas:

Sensibilidad al tamaño n de la muestra

Tamaño de la muestra	Media MC (n/1000) de R^2	Media MC (n/1000) de $\left(\frac{\text{SRC}}{n-1}\right)$	% de la dispersión inicial	$E\left(\frac{\text{SRC}}{n-1}\right)$	Var y_t
90	0.451	7.01	84.4 %	6.06	45
80	0.451	6.10	84.7 %	5.40	40
70	0.449	4.68	86.6 %	4.73	35
60	0.450	4.21	85.9 %	4.06	30
50	0.441	3.58	86.4 %	3.40	25
40	0.438	1.42	86.3 %	2.73	20
30	0.435	2.35	84.3 %	2.06	15
20	0.429	1.36	86.4 %	1.40	10
10	0.401	0.61	97.8 %	0.66	5

Tamaño n de la muestra	Desv. típica MC (n/1000) de $t(\beta_0)$	Frecuencia MC (n/1000) rechazo H_0 ($\beta_0 = 0$)	Desv. típica MC (n/1000) de $t(\beta_1)$	Frecuencia MC (n/1000) rechazo H_0 ($\beta_1 = 0$)
90	6.91	80.0 %	13.81	88.0 %
80	5.34	79.2 %	12.43	86.4 %
70	4.31	76.1 %	11.90	82.3 %
60	3.92	74.9 %	10.31	81.2 %
50	3.63	71.1 %	9.67	79.4 %
40	3.13	68.0 %	8.21	76.2 %
30	2.97	63.1 %	7.03	71.1 %
20	2.83	54.0 %	6.31	72.0 %
10	2.61	39.0 %	4.91	59.6 %

Un examen atento de estos resultados nos lleva a las siguientes conclusiones:

No hay casi influencia del tamaño de la muestra sobre el valor del coeficiente R^2 que se mantiene dentro del intervalo 0.43 ± 0.029 .

También se mantiene prácticamente constante el porcentaje de supuesta eliminación de la varianza total observada, que oscila alrededor de un porcentaje medio del 85.86 % para todos los tamaños de la muestra.

Sin embargo, como era de esperar, los estadísticos t MCO sí que son sensibles al tamaño de la muestra como recoge la última tabla, aunque, como se puede comprobar, las frecuencias de rechazo siguen siendo anormalmente altas.

Concluida esta fase de nuestro estudio, nos ha parecido oportuno dedicar especial atención a la naturaleza de los residuos MCO resultantes de la regresión, ya que son precisamente estos residuos el objetivo principal de la aparente supresión de la tendencia.

La simulación del estadístico Durbin-Watson:

media MC (100/1000) de DW = 0.196
varianza MC (100/1000) de DW = 0.009
desviación típica MC (100/1000) = 0.094

nos pone de manifiesto la necesidad de aceptar, en promedio, la presencia de una correlación serial entre esos residuos MCO.

Un procedimiento, ya clásico, de intentar eliminar esa correlación en los residuos es la transformación de Cochrane-Orcutt, basado en el supuesto de que la estructura de autocorrelaciones en los residuos se corresponde con la generada por un modelo AR (1). Aunque en un principio este supuesto puede parecer un tanto arbitrario, posteriormente, comprobaremos que es la elección más acertada. Con este supuesto la transformación del proceso vendría expresada por

$$(y_t - \rho y_{t-1}) = \beta_0 (1 - \rho) + \beta_1 [t - \rho (1 - t)] + (u_t - \rho u_{t-1})$$

Si no existiese la hipotética confusión de modelos, que es punto de partida de nuestro estudio, esta transformación sería efectivamente la adecuada, pues entonces sabríamos que el valor de $\rho = 1$ (caso de un proceso «paso aleatorio»), y la transformación anterior se convertiría en diferencias primeras.

Ahora bien, como imponemos la persistencia de la confusión, entonces, al ignorar el valor de ρ , se recurriría a su EMCO $\hat{\rho}$. La simulación Monte Carlo de ese parámetro nos ha dado:

media MC (100/1000) de $\hat{\rho} = 0.90 < 1$
varianza MC (100/1000) de $\hat{\rho} = 0.0031$
desviación típica MC (100/1000) de $\hat{\rho} = 0.055$

donde el valor medio de $\hat{\rho}$ es 0.90, claramente inferior a la unidad. Luego la transformación por diferencias queda por debajo de la que

sería adecuada para suprimir plenamente la correlación serial de los residuos.

La comprobación empírica, vía simulación Monte Carlo, de la transformación de Cochrane-Orcutt nos proporciona los siguientes valores del coeficiente de correlación R^2 y de los estadísticos t para β_0 y β_1 :

media MC (100/1000) de $R^2 = 0.156$
varianza MC (100/1000) de $R^2 = 0.079$
desviación típica MC (100/1000) de $R^2 = 0.281$

desviación típica MC (100/1000) de $t(\hat{\beta}_1) = 1.3$
frecuencia rechazo (5 %) MC (100/1000) de $H_0(\beta_0 = 0) = 45 \%$
desviación típica MC (100/1000) de $t(\hat{\beta}_0) = 2.63$
frecuencia rechazo (5 %) MC (100/1000) de $H_0(\beta_1 = 0) = 40.4 \%$

Las conclusiones que se siguen de estos resultados es que aunque se ha producido una reducción del 65 % en el valor del coeficiente de correlación, el comportamiento de los estadísticos t y de las respectivas frecuencias de rechazo, siguen sin ser información suficiente para aceptar la componente temporal como variable explicativa.

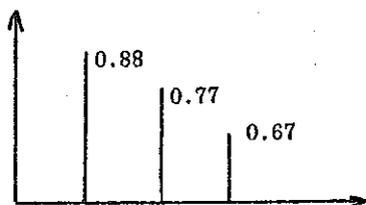
Para analizar las causas de esta regresión espúrea —todavía persistente—, nos ha parecido conveniente hacer un estudio más detallado de la estructura de autocorrelación de los residuos MCO, resultantes de la supuesta eliminación de la tendencia.

Partiendo de la estimación de la función de autocorrelación que para esos residuos hicieron, como ya hemos citado, Chan, Hayya y Ord (1977), posteriormente corregida por Nelson y Kang (1981), es fácil establecer que esa función de autocorrelación:

- a) debe ser fuertemente decreciente para los primeros retrasos $\tau = 1, 2, 3,$
- b) y que su forma es independiente del tamaño de la muestra.

Las simulaciones Monte Carlo realizadas que se recogen en el correlograma:

media MC (100/1000) de $\hat{\rho}(\tau = 1) = 0.88$
media MC (100/1000) de $\hat{\rho}(\tau = 2) = 0.77$
media MC (100/1000) de $\hat{\rho}(\tau = 3) = 0.67$



confirman plenamente estas previsiones teóricas, a la vez que sugieren un modelo AR(1) como el generador de esos residuos.

Sugerencia que se ve confirmada por el comportamiento de las simulaciones de las autocorrelaciones parciales:

media MC (100/1000) de $\hat{\Pi}(\tau = 2) = -0.039$

media MC (100/1000) de $\hat{\Pi}(\tau = 3) = -0.018$

banda de significación $1/\sqrt{n} = 0.1$

confirmación de estructura según modelo AR (1).

Resultado que elimina la opinión de arbitrariedad sobre la elección de la estructura AR (1) como generadora de los residuos, que hicimos con anterioridad.

Persiguiendo la exhaustividad de nuestro análisis volvamos de nuevo a examinar si persiste o no la correlación serial en los residuos MCO resultantes de la transformación de Cochrane-Orcutt.

Las simulaciones realizadas para el estadístico de Durbin-Watson y la correspondiente autocorrelación para el retraso 1:

media MC (100/1000) de DW = 1.876

media MC (100/1000) de $\hat{\rho}(\tau = 1) = 0.050$

nos permite afirmar que aunque más reducida persiste una autocorrelación en los residuos.

A estas alturas de la investigación cabría plantearse la duda de que los anteriores resultados están condicionados por la imposición de que $\beta_1 = 0$ en el modelo, pero ¿qué sucedería si β_1 fuese distinto de cero?

Un análisis de la estructura del coeficiente R^2 , para este supuesto, nos lleva al resultado teórico de que el valor de ese coeficiente se aproxima muy rápidamente al valor unitario, para tamaños de la muestra suficientemente grandes.

Resultado teórico que se ve confirmado por la simulación realizada. Hemos utilizado diez valores de la muestra para el caso de estricta dependencia temporal ($\beta_1 = 1$). Los resultados se exponen en la tabla adjunta:

Tamaño de la muestra	Media MC (100/1000) de R^2	Valor teórico de R^2
n = 100	0.989	0.992
n = 90	0.983	0.991
n = 80	0.979	0.990
n = 70	0.970	0.988
n = 60	0.963	0.987
n = 50	0.954	0.984
n = 40	0.947	0.980
n = 30	0.932	0.975
n = 20	0.930	0.963
n = 10	0.921	0.933

Luego, es evidente que la presencia de deriva no nula, no mitiga en absoluto el fenómeno de la regresión espúrea.

Una vez que hemos estudiado, bastante exhaustivamente, el modelo inicialmente planteado, nos pareció oportuno intentar una generalización del modelo introduciendo una variable explicativa x_t no estacionaria. El modelo será ahora:

$$y_t = \beta_0 + \beta_1 t + \beta_2 x_t + u_t$$

Luego la serie

$$y_t - \beta_2 x_t = \beta_0 + \beta_1 t + u_t$$

dará comportamiento estacionario, una vez que se suprima la supuesta «tendencia», ya que se mantiene la confusión de modelos, con la correspondiente confusión sobre la estacionaridad de las perturbaciones.

Si ante tal supuesto se intentase realizar una regresión MCO sobre las observaciones originales, aparecería un tipo de regresión espúrea como el ya estudiado por Granger y Newbold (1974), con sobreestimación de los coeficientes y estadísticos.

Un posible procedimiento para tratar de eliminar esa regresión espúrea es regresar previamente cada una de las series x_t y y_t sobre el tiempo, tratando de eliminar la supuesta influencia temporal, y, a continuación, regresar la serie x_t sobre la serie y_t . Pero ya Lovell había demostrado en 1963 que eso era equivalente a la regresión MCO entre los datos originales.

Un procedimiento alternativo consiste en considerar la «estacionarización» de una serie como una adecuada transformación matricial:

$$y_t^{dt} = T y_t$$

$$x_t^{dt} = T x_t$$

$$u_t^{dt} = T u_t$$

siendo T la matriz de transformación.

Con lo que el problema se reduce a estimar el parámetro β_2 en

$$y_t^{dt} = \beta_2 x_t^{dt} + u_t^{dt}$$

para una regresión MCO del «paso aleatorio» y_t^{dt} sobre el «paso aleatorio» x_t^{dt} . Es decir, una regresión en la que tanto la variable explicativa x_t^{dt} como la perturbación u_t^{dt} tienen la misma función de autocorrelación.

Empleando el factor establecido por Wold (1953) para medir el efecto multiplicador de la presencia de autocorrelación en las perturbaciones, sobre la varianza de las estimaciones MCO

$$[1 + \sum \hat{\rho}_x(\tau = i) \hat{\rho}_u(\tau = i)]$$

donde $\hat{\rho}_x(\tau = i)$ es la autocorrelación de la variable explicativa al retraso i y $\hat{\rho}_u(\tau = i)$ es la autocorrelación de la perturbación al retraso i como en nuestro caso se verifica la identidad de las autocorrelaciones, entonces el coeficiente de Wold queda

$$\{1 + \sum [\hat{\rho}_u(\tau = i)]^2\}$$

es decir, la estimación MCO de β_2 , aumentará en una cuantía todavía mayor que la prevista.

Factor que además aumentará con el tamaño de la muestra, ya que como vimos, las autocorrelaciones al retraso i se incrementan con el tamaño de la muestra.

Las simulaciones Monte Carlo realizadas confirman todas estas previsiones teóricas.

Para estas simulaciones diseñamos un nuevo programa llamado SIMULI, que recoge las características del nuevo modelo. Al que en nuestro caso le hemos impuesto las condiciones no distorsionantes $\beta_1 = 0$ y $\beta_2 = 0$.

El examen de los datos obtenidos:

media MC (100/1000) de $\hat{\beta}_2 = -0.010$
 varianza MC (100/1000) de $\hat{\beta}_2 = 0.163$
 desviación típica MC (100/1000) de $\hat{\beta}_2 = 0.404$
 media MC (100/1000) de $(\beta_2 - \hat{\beta}_2)^2 = 0.101$

en concreto, la comparación de la desviación típica simulada de la EMCO de $\hat{\beta}_2$, con el error típico simulado de $\beta_2 - \hat{\beta}_2$, nos sugiere la presencia de una sobreestimación de orden de magnitud 1/4.

Respecto al coeficiente de correlación:

media MC (100/1000) de $R^2 = 0.516$
 varianza MC (100/1000) de $R^2 = 0.084$
 desviación típica MC (100/1000) de $R^2 = 0.289$

se produce un aumento de 0.452 a 0.516. Resultado que junto con el comportamiento de los estadísticos t y sus respectivas frecuencias de rechazo

media MC (100/1000) de $t(\hat{\beta}_2) = -0.049$
 varianza MC (100/1000) de $t(\hat{\beta}_2) = 19.243$
 desviación típica MC (100/1000) de $t(\hat{\beta}_2) = 4.387$

al 1 % la frecuencia de significación de $\hat{\beta}_2$ es 51 %

al 5 % la frecuencia de significación de $\hat{\beta}_2$ es 61 %

frecuencia de rechazo MC (100/1000) al 5 % de $H_0(\beta_2 = 0) = 83 %$

nos lleva a concluir, que la inclusión de una supuesta variable explicativa, no puede reducir la magnitud de la regresión espúrea. Es más, en el caso concreto que hemos estudiado, entre la variable tiempo y la variable «paso aleatorio», explican casi el 50 % de la variación del proceso «paso aleatorio» observada, cuando sabemos que en la realidad no depende ni de uno ni del otro.

Para mantener el rigor de nuestro estudio volvamos a analizar el comportamiento de los residuos MCO resultantes de la anterior regresión.

Las simulaciones del estadístico de Durbin-Watson, no nos permiten decidir sobre la existencia o no de correlación serial, al situarse en zona de ambigüedad:

media MC (100/1000) de DW = 0.25
 varianza MC (100/1000) de DW = 0.12
 desviación típica MC (100/1000) de DW = 0.34

Con el fin de adquirir más información, hemos analizado el comportamiento de la función de autocorrelación

media MC (100/1000) de $\hat{\rho}(\tau = 1) = 0.073$
 media MC (100/1000) de $\hat{\rho}(\tau = 2) = 0.065$
 media MC (100/1000) de $\hat{\rho}(\tau = 3) = 0.058$

que nos vuelve a sugerir un modelo AR(1) como generador de esos residuos.

Aceptado este modelo AR(1) hemos realizado la pertinente transformación

$$y_t = \beta_0 + \beta_1 t + \beta_2 x_t + u_t$$

$$y_t - \rho y_{t-1} = \beta_0 (1 - \rho) + \beta_1 [t - \rho(t - 1)] + \beta_2 (x_t - \rho x_{t-1}) + (u_t - \rho u_{t-1})$$

del modelo considerado. Volviendo entonces a establecer los valores del estadístico DW:

media MC (100/1000) de DW = 1.79
 varianza MC (100/1000) de DW = 0.143
 desviación típica MC (100/1000) de DW = 0.378

y del correlograma:

media MC (100/1000) de $\hat{\rho}(\tau = 1) = 0.81 < 1$
 media MC (100/1000) de $\hat{\rho}(\tau = 2) = 0.056$
 media MC (100/1000) de $\hat{\rho}(\tau = 3) = 0.051$

con $1/\sqrt{n} = 0.1$ raramente se rechazará H_0 (hay correlación serial) es necesario admitir la hipótesis de correlación serial en los residuos.

Los restantes efectos de la transformación de la serie observada se recogen en la siguiente tabla:

	Datos originales	Datos transformados
media de R^2	0.516	0.151
varianza de R^2	0.084	0.029
desviación típica de R^2	0.289	0.170
media $\hat{\beta}_1$	0.005	0.005
varianza $\hat{\beta}_1$	0.013	0.013
desviación típica de $\hat{\beta}_1$	0.115	0.115
media $\hat{\beta}_2$	- 0.010	0.003
varianza $\hat{\beta}_2$	0.163	0.035
desviación típica de $\hat{\beta}_2$	0.404	0.183
media $t(\hat{\beta}_1)$	0.325	0.141
varianza $t(\hat{\beta}_1)$	135.552	25.913
desviación típica $t(\hat{\beta}_1)$	11.557	5.09
media $t(\hat{\beta}_2)$	- 0.049	0.045
varianza $t(\hat{\beta}_2)$	19.243	1.732
desviación típica $t(\hat{\beta}_2)$	4.387	1.316

La lectura de esa tabla confirma la persistencia de la regresión espúrea.

De todo el trabajo teórico y experimental realizado y que hemos tratado de exponer lo más sucintamente posible, hemos obtenido las siguientes conclusiones:

1. Tratar de eliminar, mediante regresión MCO, una inexistente «tendencia» (relación lineal del tiempo) en un proceso «paso aleatorio», genera un valor del coeficiente de correlación R^2 , que induce al error de admitir la posibilidad de que efectivamente exista esa «tendencia».

2. En el supuesto de que el proceso considerado tenga «deriva» nula (no hay ninguna relación funcional del tiempo), los valores del coeficiente de correlación R^2 no dependen del tamaño de la muestra.

3. En el supuesto de que el proceso considerado no tenga «deriva» nula, los valores del coeficiente de correlación R^2 aumentan con el tamaño de la muestra, hasta alcanzar valores próximos a la unidad. Este resultado es independiente de la magnitud promedio de la «deriva», así como de su variabilidad.

4. La variación de los residuos resultantes de la regresión MCO sobre el tiempo de un proceso «paso aleatorio», es alrededor de un 15 % de la variación realmente observada. Luego el simple hecho de realizar esa regresión MCO conduce al error de admitir una supuesta «explicación temporal» de alrededor del 85 %.

5. El comportamiento de la varianza de los residuos es independiente de:

- el tamaño de la muestra,
- la magnitud media de la «deriva».

6. Si se realiza la falsa suposición de que los residuos de la regresión se configuran como una serie estacionaria, entonces, debido sin duda a que los procesos «paso aleatorio» tienen un buen ajuste «ex post» a la regresión, pero no «ex ante», la varianza residual queda muy por debajo del valor de la varianza observada: ($6.71 < 50$).

7. El clásico estadístico «t» para los EMCO de los coeficientes de regresión sobre el parámetro tiempo tienen un comportamiento muy insatisfactorio cuando se trata de detectar la «tendencia» de la serie observada. Según hemos comprobado rechazan la hipótesis nula, de no dependencia, en casi el 85 % de los casos, en muestras de tamaño 100, y coeficiente de significación de 0.05; cuando en realidad no existe dependencia funcional del tiempo.

8. Si por algún motivo se detectase la presencia de autocorrelación entre los residuos, y se quisiera tener en cuenta la presencia de esas correlaciones a la hora de realizar la regresión, las técnicas disponibles sólo son capaces de producir mejoras parciales y poco eficaces.

9. La única manera eficiente de suprimir las autocorrelaciones en los residuos es tomando diferencias primeras.

10. La regresión de un «paso aleatorio» sobre otro «paso aleatorio», que incluye el tiempo, con el fin de eliminar la «tendencia», está fuertemente sometida a la generación de fenómenos de regresión espúrea. Es decir, los clásicos «test t» sugieren la existencia de relaciones entre las variables, que realmente no existen.

11. Cuando se intenta eliminar las regresiones espúreas corrigiendo la correlación serial, pero manteniendo el supuesto, de que las perturbaciones son estacionarias AR(1), se lograrán correcciones parciales muy poco eficientes. Así, por ejemplo, la frecuencia de rechazo de la verdadera hipótesis se reducirá en apenas un 10 %.

12. Las diferencias primeras constituyen el único método eficiente para eliminar las regresiones espúreas.

13. Como conclusión práctica opinamos que la actitud generalizada de explicar la no-estacionaridad de una serie a partir de un modelo que supone estacionaridad alrededor de una función del tiempo, debe ser considerada como una tentativa inicial orientativa, pero que su mantenimiento no suficientemente justificado puede conducir a resultados erróneos. Por eso recomendamos que antes de realizar una regresión sobre una serie que puede no seguir el modelo de estacionaridad alrededor de una función del tiempo es muy recomendable aplicar algún

contraste sobre la estacionaridad de los residuos. En cualquier caso, recomendamos aplicar previamente diferencias primeras, ya que, en último caso, la aplicación innecesaria de diferencias primeras a series que conllevan subseries estacionarias produce menos distorsiones y más fácilmente detectables que la injustificada aplicación de métodos de regresión.

BIBLIOGRAFIA

- CHAN, K. H., J. C. HAYYA and J. K. ORD (1977): *A note trend removal methods: the case of polynomial regression us variate differencing*. *Econometrica* 45, pp. 737-744.
- COCHRANE, D. and G. H. ORCUTT (1949): *Application of least squares regressions to relationships containg autocorrelated error terms*. *Journal of the American Statistical Association*, vol. 44, pp. 32-61.
- GRANGER, C. W. J. and NEWBOLD (1974): *Spurious regressions in econometrics*. *Jour. Econometrics* 2, 111-120.
- LOVELL, M. C. (1963): *Seasonal adjustment of Economic time series*. *Journal of the American Statistical Association* 58, pp. 93-101.
- NELSON, CH. R. and KANG, H. (1981): *Spurious priodicity inappropriately detrend time series*. *Econometrica* 34, 235-238.
- NELSON, CH. R. and KANG, H. (1984): *Pitfalls in the use of time as an explanatory variable in regression*. *Journal of Business and Economic Statistics*, vol. 2, n.º 1, 73-82.
- PONCE NÚÑEZ, José Miguel (1985): *Ausencia de estacionaridad en las perturbaciones de una serie temporal y su influjo en la regresión sobre el tiempo*. Tesis doctoral. Universidad de Granada.
- REUVEN, Y. R. (1981): *Simulation and the Monte Carlo Method*. John Wiley, New York.
- WOLD, H. (1949): *On least squares regression with autocorrelated error terms*. *Bulletin of the International Statistical Institute*, 32.

Estudio del sistema de distribuciones de probabilidad bivariantes discretas del tipo Pearson-Ord^(*)

R. HERRERÍAS PLEGUEZUELO

H. CALVETE FERNÁNDEZ

Departamento de Estadística

Facultad de Ciencias Económicas y Empresariales

Universidad de Zaragoza

1. INTRODUCCION

Sea la familia de distribuciones discretas bivariantes solución del sistema de ecuaciones en diferentes finitas parciales:

$$\frac{\Delta_1 f_{r-1,s}}{f_{r-1,s}} = \frac{a_0 + a_1 r + a_2 s}{g_0 + g_1 r + g_2 s + g_3 r(r-1) + g_4 s(s-1) + g_5 rs} \quad (1.1)$$

$$\frac{\Delta_2 f_{r,s-1}}{f_{r,s-1}} = \frac{b_0 + b_1 r + b_2 s}{h_0 + h_1 r + h_2 s + h_3 r(r-1) + h_4 s(s-1) + h_5 rs}$$

donde

$$\Delta_1 f_{r-1,s} = f_{r,s} - f_{r-1,s}$$

$$\Delta_2 f_{r,s-1} = f_{r,s} - f_{r,s-1}$$

Estas ecuaciones son análogas, en la forma, a la ecuación de Ord (1967), que sirve para definir el sistema de Pearson univariante discreto y están en la línea de las de Van Uven (1947), utilizadas para definir el sistema de Pearson bivalente continuo.

(*) Trabajo presentado en la XV Reunión Nacional de la S.E.I.O. (Gijón, 1985).

Escribiendo el sistema (1.1) en la forma

$$f_{r,s} = A(r,s) f_{r-1,s}$$

$$f_{r,s} = B(r,s) f_{r,s-1}$$

Herrerías y Cobos (1984) obtienen la expresión de la solución general de dicho sistema:

$$f_{r,s} = f_{00} \frac{\prod_{i=1}^s f_{0i}}{\prod_{i=0}^{s-1} f_{ri}} \frac{\prod_{i=1}^{s+1} B(0,i)}{\prod_{i=1}^r A(j,i)}$$

donde f_{00} juega el papel de constante normalizadora y se determina haciendo:

$$\sum_{r,s} f_{rs} = 1$$

En el presente trabajo se obtienen las relaciones de recurrencia entre los momentos factoriales descendentes, verificándose su coincidencia con la presentada por Ord en el caso univariante (para $r = \text{cte}$ ó $s = \text{cte}$). Se dan, asimismo, ejemplos de familias de distribuciones que pertenecen al sistema, comprobándose, para ellas, las relaciones entre los momentos obtenidas.

2. RELACION ENTRE LOS MOMENTOS FACTORIALES DESCENDENTES

Denotando

$$r^{(j)} = r(r-1)\dots(r-j+1)$$

$$s^{(k)} = s(s-1)\dots(s-k+1)$$

y multiplicando ambos miembros de cada una de las ecuaciones del sistema (1.1) por $r^{(j)} s^{(k)}$ puede escribirse:

$$f_{r-1,s}(r-1)^{(j)}s^{(k)}[(a_0 + g_0) + (a_1 + g_1)r + (a_2 + g_2)s + g_3r(r-1) + g_4s(s-1) + g_5rs] + jf_{r-1,s}(r-1)^{(j-1)}s^{(k)}[(a_0 + g_0) + (a_1 + g_1)r + (a_2 + g_2)s + g_3r(r-1) + g_4s(s-1) + g_5rs] - f_{r,s}r^{(j)}s^{(k)}[g_0 + g_1r + g_2s + g_3r(r-1) + g_4s(s-1) + g_5rs] = 0$$

$$f_{r,s-1}r^{(j)}(s-1)^{(k)}[(b_0 + h_0) + (b_1 + h_1)r + (b_2 + h_2)s + h_3r(r-1) + h_4s(s-1) + h_5rs] + kf_{r,s-1}r^{(j)}(s-1)^{(k-1)}[(b_0 + h_0) + (b_1 + h_1)r + (b_2 + h_2)s + h_3r(r-1) + h_4s(s-1) + h_5rs] - f_{r,s}r^{(j)}s^{(k)}[h_0 + h_1r + h_2s + h_3r(r-1) + h_4s(s-1) + h_5rs] = 0$$

Sumando en $[u,v]$ y $[w,z]$, recorridos, respectivamente, de las variables r y s , y haciendo uso, entre otras, de las siguientes relaciones entre los factoriales descendentes:

$$\begin{aligned} r(r-1)(r-1)^{(j)} &= (r-1)^{(j+2)} + 2(j+1)(r-1)^{(j+1)} + j(j+1)(r-1)^{(j)} \\ r(r-1)r^{(j)} &= r^{(j+2)} + 2jr^{(j+1)} + j(j-1)r^{(j)} \\ r(r-1)^{(j)} &= (r-1)^{(j+1)} + (j+1)(r-1)^{(j)} \\ rr^{(j)} &= r^{(j+1)} + jr^{(j)} \end{aligned}$$

se obtiene el siguiente sistema de ecuaciones:

$$\begin{aligned} &\mu_{(j+1,k)}[a_1 + (j+2)g_3] + \mu_{(j,k+1)}[a_2 + (j+1)g_5] + \mu_{(j-1,k+2)}jg_4 + \\ &+ \mu_{(j,k)}[a_0 + (2j+1)a_1 + ka_2 + (j+1)g_1 + 2j(j+1)g_3 + k(j+1)g_5] + \\ &+ \mu_{(j-1,k+1)}[(a_2 + g_2)j + 2kjg_4 + j^2g_5] + \mu_{(j-1,k)}[(a_0 + g_0)j + (a_1 + g_1)j^2 + \\ &+ (a_2 + g_2)kj + j^2(j-1)g_3 + kj(k-1)g_4 + j^2kg_5] + E_{(j+1,k)} = 0 \quad (2.1) \end{aligned}$$

$$\begin{aligned} &\mu_{(j+1,k)}[b_1 + (k+1)h_5] + \mu_{(j,k+1)}[b_2 + (k+2)h_4] + \mu_{(j+2,k-1)}kh_3 + \\ &+ \mu_{(j,k)}[b_0 + jb_1 + (2k+1)b_2 + (k+1)h_2 + 2k(k+1)h_4 + j(k+1)h_5] + \\ &+ \mu_{(j+1,k-1)}[(b_1 + h_1)k + 2jkh_3 + k^2h_5] + \mu_{(j,k-1)}[(b_0 + h_0)k + (b_1 + h_1)jk + \\ &+ (b_2 + h_2)k^2 + j(j-1)kh_3 + k^2(k-1)h_4 + jk^2h_5] + E_{(j,k+1)} = 0 \quad (2.2) \end{aligned}$$

donde $\mu_{(j,k)}$ representa el momento factorial descendente bidimensional para $j, k = 1, 2, 3, \dots$ y $\mu_{(0,0)} = 1$; siendo $E_{(j+1,k)}$ y $E_{(j,k+1)}$ términos complementarios que vienen dados por:

$$\begin{aligned} E_{(j+1,k)} &= u^{(j)} \sum_s [g_0 + g_1u + g_2s + g_3u(u-1) + g_4s(s-1) + \\ &+ g_5us] s^{(k)} f_{u,s} - (v+1)^{(j)} \sum_s [(a_0 + g_0) + (a_1 + g_1)(v+1) + \\ &+ (a_2 + g_2)s + g_3(v+1)v + g_4s(s-1) + g_5(v+1)s] s^{(k)} f_{v,s} \quad (2.3) \end{aligned}$$

$$\begin{aligned} E_{(j,k+1)} &= w^{(k)} \sum_r [h_0 + h_1r + h_2w + h_3r(r-1) + h_4w(w-1) + \\ &+ h_5rw] r^{(j)} f_{r,w} - (z+1)^{(k)} \sum_r [(b_0 + h_0) + (b_1 + h_1)r + (b_2 + h_2)(z+1) + \\ &+ h_3r(r-1) + h_4(z+1)z + h_5r(z+1)] r^{(j)} f_{r,z} \quad (2.4) \end{aligned}$$

Si el rango de ambas variables es $(-\infty, \infty)$ o $[0, \infty)$ ambos términos complementarios son iguales a cero para todo j, k .

Es inmediato observar que si hacemos $s = \text{cte.}$ (equivalentemente $r = \text{cte.}$) la relación (2.1) se transforma en

$$\begin{aligned} &\mu_{(j+1)}[a_1 + (j+2)g_3] + \mu_{(j)}[(a_0 + a_2s) + (j+1)(g_1 + g_5s) + \\ &+ (2j+1)a_1 + 2j(j+1)g_3] + j\mu_{(j-1)}[(a_0 + a_2s) + (g_0 + g_2s + \\ &+ g_4s(s-1)) + ja_1 + j(j-1)g_3 + j(g_1 + g_5s)] + E_{(j+1)} = 0 \end{aligned}$$

siendo el término complementario

$$E_{j+1} = u^{(j)}[(g_0 + g_2s + g_4s(s-1)) + (g_1 + g_3s)u + g_3u(u-1)]f_{u,s} - \\ - (v+1)^{(j)}[(a_0 + a_2s) + a_1(v+1) + (g_0 + g_2s + g_4s(s-1)) + \\ + (g_1 + g_3s)(v+1) + g_3(v+1)v]f_{v,s}$$

que coincide con la obtenida por Ord en el caso univariante.

3. ALGUNAS DISTRIBUCIONES CONTENIDAS EN EL SISTEMA

Consideraremos en este apartado algunas de las distribuciones discretas bivariantes más importantes, que son caso particular de la familia de distribuciones presentada, y que pueden encontrarse en Johnson y Kotz (1969).

3.1. Distribución binomial negativa bivalente

Se obtiene dando a los parámetros los siguientes valores:

$$a_0 = (N-1)P_1; \quad a_1 = P_1 - Q; \quad a_2 = P_1 \\ b_0 = (N-1)P_2; \quad b_1 = P_2; \quad b_2 = P_2 - Q \\ g_0 = g_2 = g_3 = g_4 = g_5 = h_0 = h_1 = h_3 = h_4 = h_5 = 0 \\ g_1 = h_2 = Q$$

Las relaciones (2.1) y (2.2) toman la forma (para $j = 1 = k$)

$$\mu_{(2,1)}(P_1 - Q) + \mu_{(1,2)}P_1 + \mu_{(1,1)}[(N+3)P_1 - Q] + \mu_{(0,2)}P_1 + \\ + \mu_{(0,1)}(N+1)P_1 + E_{(2,1)} \\ \mu_{(2,1)}P_2 + \mu_{(1,2)}(P_2 - Q) + \mu_{(1,1)}[(N+3)P_2 - Q] + \mu_{(2,0)}P_2 + \\ + \mu_{(1,0)}(N+1)P_2 + E_{(1,2)}$$

Sustituyendo

$$\mu_{(j,k)} = N^{(j+k)}P_1^jP_2^k$$

donde $N^{(j)}$ denota una factorial ascendente, y como ambos términos complementarios son cero se obtiene la igualdad a cero de ambas expresiones.

3.2. Distribución trinomial

Se obtiene haciendo:

$$a_0 = (N+1)P_1; \quad a_1 = -Q - P_1; \quad a_2 = -P_1 \\ b_0 = (N+1)P_2; \quad b_1 = -P_2; \quad b_2 = -P_2 - Q \\ g_0 = g_2 = g_3 = g_4 = g_5 = h_0 = h_1 = h_3 = h_4 = h_5 = 0 \\ g_1 = h_2 = Q$$

Las relaciones (2.1) y (2.2) son en este caso:

$$\begin{aligned} & (P_2 - 1)^{\mu_{(2,1)}} - P_1^{\mu_{(1,2)}} + [(N - 2)P_1 - 1 + P_2]^{\mu_{(1,1)}} - P_1^{\mu_{(0,2)}} + \\ & \quad + (N - 1)P_1^{\mu_{(0,1)}} + E_{(2,1)} \\ & - P_2^{\mu_{(2,1)}} + (P_1 - 1)^{\mu_{(1,2)}} + [(N - 2)P_2 - 1 + P_1]^{\mu_{(1,1)}} - P_2^{\mu_{(2,0)}} + \\ & \quad + (N - 1)P_2^{\mu_{(1,0)}} + E_{(1,2)} \end{aligned}$$

Sustituyendo

$$\mu_{(j,k)} = N^{(j+k)} P_1^j P_2^k$$

y teniendo en cuenta que por ser $u = w = 0$, $v = z = N$, los términos complementarios son:

$$\begin{aligned} E_{(2,1)} &= (N + 1)P_1 \sum_s s^2 f_{N,s} = 0 \\ E_{(1,2)} &= (N + 1)P_2 \sum_r r^2 f_{r,N} = 0 \end{aligned}$$

se obtiene la igualdad a cero de ambas expresiones.

3.3. Distribución hipergeométrica bivalente

Se obtiene dando a los parámetros los valores:

$$\begin{aligned} a_0 &= NM_1 + N + M_1 + 1; & a_1 &= M_2 - M - 2; & a_2 &= -M_1 - 1 \\ b_0 &= NM_2 + M_2 + N + 1; & b_1 &= -M_2 - 1; & b_2 &= M_1 - M - 2 \\ g_0 &= g_2 = g_4 = h_0 = h_1 = h_3 = 0 \\ g_1 &= h_2 = M - M_1 - M_2 - N + 1 \\ g_3 &= g_5 = h_4 = h_5 = 1 \end{aligned}$$

Las relaciones (2.1) y (2.2) se transforman en

$$\begin{aligned} & \mu_{(2,1)}(M_2 - M + 1) + \mu_{(1,2)}(1 - M_1) + \mu_{(1,1)}(NM_1 - N + M_2 - M - \\ & \quad - 2M_1 + 2) - \mu_{(0,2)}M_1 + \mu_{(0,1)}M_1(N - 1) + E_{(2,1)} \\ & \mu_{(2,1)}(1 - M_2) + \mu_{(1,2)}(M_1 - M + 1) + \mu_{(1,1)}(NM_2 - N + M_1 - M - \\ & \quad - 2M_2 + 2) - \mu_{(2,0)}M_2 + \mu_{(1,0)}M_2(N - 1) + E_{(1,2)} \end{aligned}$$

Sustituyendo

$$\mu_{(j,k)} = \frac{N^{(j+k)}}{M^{(j+k)}} M_1^{(j)} M_2^{(k)}$$

y teniendo en cuenta que, en este caso, $u = w = 0$, $v = M_1$ y $z = M_2$ y por tanto ambos términos complementarios son:

$$E_{(2,1)} = - (M_1 + 1) \sum_s [NM_1 + N + M_1 + 1 + (-M_1 - N - 1)(M_1 + 1) + (-M_1 - 1)s + (M_1 + 1)M_1 + (M_1 + 1)s] sf_{M_1, s} = 0$$

$$E_{(1,2)} = - (M_2 + 1) \sum_r [NM_2 + M_2 + N + 1 + (-M_2 - 1)r + (-M_2 - N - 1)(M_2 + 1) + (M_2 + 1)M_2 + r(M_2 + 1)] rf_{r, M_2} = 0$$

se comprueba la igualdad a cero de ambas ecuaciones.

3.4. Distribución serie logarítmica bivalente

Se obtiene haciendo:

$$a_0 = -\theta_1; \quad a_1 = \theta_1 - 1; \quad a_2 = \theta_1$$

$$b_0 = -\theta_2; \quad b_1 = \theta_2; \quad b_2 = \theta_2 - 1$$

$$g_0 = g_2 = g_3 = g_4 = g_5 = h_0 = h_1 = h_3 = h_4 = h_5 = 0$$

$$g_1 = h_2 = 1$$

En este caso (2.1) y (2.2) se convierten en

$$(\theta_1 - 1)\mu_{(2,1)} + \theta_1\mu_{(1,2)} + (3\theta_1 - 1)\mu_{(1,1)} + \theta_1\mu_{(0,2)} + \theta_1\mu_{(0,1)} + E_{(2,1)}$$

$$\theta_2\mu_{(2,1)} + (\theta_2 - 1)\mu_{(1,2)} + (3\theta_2 - 1)\mu_{(1,1)} + \theta_2\mu_{(2,0)} + \theta_2\mu_{(1,0)} + E_{(1,2)}$$

Los momentos factoriales son

$$\mu_{(j,k)} = \frac{(j+k-1)! \theta_1^j \theta_2^k}{[-\log(1-\theta_1-\theta_2)] (1-\theta_1-\theta_2)^{j+k}}$$

Sustituyendo esta expresión y teniendo en cuenta que ambos términos complementarios son iguales a cero, se demuestra que ambas expresiones son iguales a cero.

BIBLIOGRAFIA

- HERRERÍAS, R. y COBOS, J. (1984): «Solución general para un tipo de sistemas de distribución de probabilidad bivariantes discretas». *Cuadernos Aragoneses de Economía*, n.º 8, 133-137.
- JOHNSON, N. y KOTZ, S. (1969): «*Distributions in Statistics: Discrete distributions*». Wiley & Sons.
- ORD, J. K. (1967): «On a system of discrete distributions». *Biometrika*, n.º 54, 649-656.
- UVEN, M. J. van (1947): «Extensions of Pearson's probability distributions to two variables». *Proceeding of the Royal Acad. Sci. Amsterdam*, vol. 50, pp. 1063-1070, 1252-1254.