

SOFTWARE, INSTRUMENTACIÓN Y METODOLOGÍA

ANÁLISIS DE INCREMENTOS Y DECREMENTOS: INCIDENCIA DE LA VIOLACION DE SUPUESTOS SOBRE LA TASA EMPIRICA DE ERROR TIPO I'

Antonio Solanas y Vicenta Sierra
Universidad de Barcelona

El análisis de incrementos-decrementos fue investigado en un amplio estudio de simulación Monte Carlo. Se demostró que la violación de supuestos produjo problemas en la inferencia, observándose una importante discrepancia entre las tasas exacta y empírica de error tipo I. La asimetría del término aleatorio del modelo autorregresivo de primer orden afecta dramáticamente la tasa empírica de error tipo I. Se encontró que los niveles de dependencia serial distintos de cero afectaron la tasa empírica de error tipo I, aunque la técnica mostró un nivel aceptable de robustez para valores de autocorrelación cercanos a cero. Extrema precaución es recomendable para utilizar la técnica en el análisis de datos de diseños A-B, donde no es extraño hallar autocorrelación en las series.

Increments-Decrements Analysis. Increments-decrements technique was investigated in an extensive Monte Carlo simulation study. It was demonstrated that violated assumptions yield problems of inference in the form of important differences between exact and empirical type I error rates. Skewness in random error of first order autoregressive model affects dramatically empirical type I error rate. It was found that the empirical type I error was affected by nonzero serial dependence, although this technique showed an acceptable robustness for levels of autocorrelation near zero. Extreme caution is recommended for use this strategy in A-B designs where autocorrelated series are not unusual.

El análisis de datos procedentes de diseños conductuales se fundamenta, generalmente, en la denominada *inferencia visual*. Esta estrategia se ha difundido extensamente y relegado a un segundo

plano el análisis estadístico, tanto en el ámbito básico como en el aplicado, aunque son numerosos los problemas en el estudio visual de diseños conductuales (Ballard, 1983; Jones, Weinrott & Vaughn, 1978; Wampold & Furlong, 1981). El acuerdo entre jueces, cuando se fundamenta el análisis en la inferencia visual, es escaso, agravándose con la existencia de dependencia serial en los datos; además, la estrategia de los evaluadores está fuertemente ligada a su formación

Correspondencia: Antonio Solanas
Dpto. de Metodología de las Ciencias del
Comportamiento
Facultad de Psicología
Universidad de Barcelona
C/Adolf Florensa, s/n
08028-Barcelona. Spain

y no es fácil sintetizarla en un conjunto de reglas.

Las anteriores consideraciones pretenden destacar aquellos inconvenientes más destacables, máxime cuando disponemos de numerosas técnicas estadísticas especialmente ideadas para el análisis de datos procedentes de diferentes tipos de diseños conductuales (Edgington, 1967; 1980a; 1980b; Revusky, 1967; White, 1974). Este último grupo de técnicas cobra importancia debido a la imposibilidad de recurrir a pruebas estadísticas clásicas (T-TEST, ANOVA, etc...), ya que los datos obtenidos longitudinalmente presentan dependencia serial, violándose los supuestos necesarios para su utilización. Es posible abordar la evaluación de la incidencia del tratamiento recurriendo al análisis de series temporales, pero el número de datos que son necesarios para una aceptable estimación de los diferentes parámetros supera el número de mediciones realizadas en la mayoría de investigaciones, tanto básicas como aplicadas.

En los estudios conductuales aplicados existen dificultades añadidas, pues no es posible controlar totalmente el amplio número de variables contaminantes y, en general, este motivo es suficiente para cuestionar el carácter experimental de un amplio grupo de investigaciones. Por otro lado, no es inusual la existencia de dificultades para obtener datos que, en casos extremos, limitan el número de mediciones en la fase A a tan solo un valor. La técnica estadística β_n (Solanas, Salafranca y Guàrdia, 1992) fue ideada para el análisis estructural de diseños A-B con un solo dato en la fase A. En lo sucesivo, nos referiremos a esta estrategia como *análisis de incrementos y decrementos*. Un antecedente de esta técnica se halla en Moore y Wallis (1943), trabajo en el cual se desarrollan un amplio grupo de técnicas para analizar la aleatoriedad de las series. El análisis de incrementos y decrementos

toma la idea de aleatoriedad para realizar un análisis estructural de la serie, frente al habitual estudio comparativo de niveles de conducta entre fases.

En el presente trabajo reconceptualizamos el análisis de incrementos y decrementos, estudiando el impacto de la violación de supuestos sobre la tasa empírica de error tipo I. En concreto, nos interesamos por la violación de la independencia serial y simetría del término error, evaluando su incidencia sobre la tasa empírica de error tipo I. Además de consideraciones de orden teórico, existen investigaciones previas (Crosbie, 1987; 1989), realizadas sobre otras técnicas estadísticas, donde se muestra el fuerte impacto sobre la tasa empírica de error tipo I ante la violación del supuesto de independencia. Igualmente, la violación del supuesto de simetría debe alterar la tasa empírica de error tipo I respecto a la tasa exacta. En ambos casos, nuestro objetivo consistió en determinar la magnitud de la alteración en el análisis de incrementos y decrementos.

Análisis de incrementos y decrementos: reconceptualización

En todo momento nos referiremos a un diseño conductual A-B con un único dato en la fase A. Modelizaremos el diseño anterior, bajo el supuesto de no incidencia del tratamiento, como un proceso aleatorio, donde para todo Δt se mantiene que

$$\text{Prob}\{x_{t+\Delta t} = x_t + \Delta x\} = \text{Prob}\{x_{t+\Delta t} = x_t - \Delta x\} = \frac{1}{2} \quad (1)$$

para cualquier Δx . De la expresión anterior se desprende que, aunque la función de distribución $F(\Delta x)$ se desconozca, debe aceptarse su simetría para todo momento t .

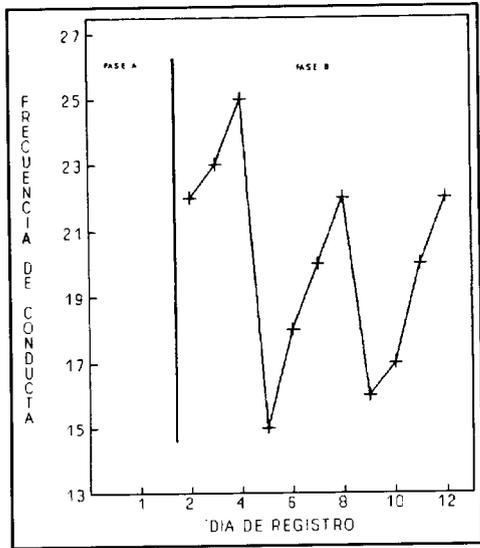
Sobre Δx , que denotaremos mediante ε , imponemos determinadas condiciones: $E_t(\varepsilon_t) = 0$, $\text{Var}_t(\varepsilon_t) = \sigma^2$ y $\text{Cov}(\varepsilon_t, \varepsilon_{t+k}) = 0$,

en todo momento t y para cualquier valor k . Respecto al término error ϵ , debe advertirse que no se exige que la varianza sea constante en todo momento (no se requiere que el proceso sea estacionario en varianza), aunque ésta debe ser finita. Este hecho añade generalidad, por lo cual no tiene consecuencias en la técnica expuesta la más que probable inconstancia de la variabilidad de la conducta a lo largo del tiempo.

En la investigación conductual el interés puede consistir en determinar si el tratamiento incrementa o decreta la tasa o el nivel de conducta. Tradicionalmente, el enfoque estadístico se ha centrado en determinar si entre fases adyacentes se ha producido un cambio de nivel, independientemente de que éste se produzca mediante un cambio abrupto o progresivo. De forma alternativa, puede analizarse la incidencia del tratamiento a partir de la estructura incremental o decremental, según sea el caso, de la serie. Esta última estrategia supone que los cambios son progresivos; por tanto, el análisis de las diferencias de primer orden de la serie deben reflejar si el comportamiento muestra cambios, ya sean en sentido incremental o decremental.

La estrategia de análisis de la eficacia o no de intervención aquí expuesta no es incompatible con aquellas fundamentadas en la comparación de niveles de conducta, sino complementaria. Pueden existir situaciones en las cuales los cambios de la conducta sean abruptos, donde estadísticos de estructura no son adecuados; pero es posible que la modificación producida por la intervención no se refleje en cambios de nivel, sino en el patrón de incrementos-decrementos de la serie, casos en los cuales los estadísticos de nivel no son pertinentes (la Figura 1 muestra un ejemplo en el cual el análisis de incrementos y decrementos detecta una estructura incremental en la serie, aunque el proceso parece estable en media).

Figura 1.
Un ejemplo de estructura incremental



Para realizar la formalización estadística de las anteriores ideas, considérese la serie $x_{1A}, x_{2B}, \dots, x_{nB}$ como datos de un diseño A-B con un solo dato en la fase A. Puesto que los tratamientos pueden incidir incrementando o decrementando la conducta, se diferenciará entre ambos casos. Nos referimos al *análisis incremental* cuando el cambio esperado por la acción del tratamiento se traduce en el aumento de la conducta. La hipótesis nula, suponiendo la aleatoriedad de la serie, es

$$H_0: \text{Prob}(x_{i+1} \leq x_i) = \text{Prob}(x_{i+1} > x_i) = 0.5 \tag{2}$$

Definamos β_n como,

$$\beta_n = \sum_{i=1}^{n-1} I_i \tag{3}$$

donde I_i denota la función indicadora, que, en el análisis incremental, toma valores según el siguiente criterio

$$I_i = 1, \text{ si } x_{i,t} > x_i \\ I_i = 0, \text{ en cualquier otro caso} \quad (4)$$

La razón $\beta_n / (n-1)$ estima la probabilidad de incremento, dadas dos observaciones adyacentes. El contraste de la hipótesis nula puede realizarse sobre la mencionada proporción o sobre el valor β_n . En ambos casos la distribución binomial es el modelo de probabilidad adecuado para obtener la probabilidad asociada al estadístico. Sólo porque requiere menos cómputo (aunque la diferencia es despreciable) proponemos realizar el contraste sobre β_n , considerando el modelo de probabilidad $Bi(n-1, 0.5)$, siendo $E(\beta_n) = (n-1)/2$ y su varianza $Var(\beta_n) = (n-1)/4$, siempre bajo el supuesto de hipótesis nula. La siguiente expresión proporciona la probabilidad asociada

$$2^{1-n} \sum_{i=\beta_n}^{n-1} \binom{n-1}{i} \quad (5)$$

Cuando el cambio esperado por la acción del tratamiento sobre la conducta sea la disminución de la misma, nos referiremos al *análisis decremental*. La hipótesis nula es

$$H_0: \text{Prob}(x_{i,t} \geq x_i) = \text{Prob}(x_{i,t} < x_i) = 0.5 \quad (6)$$

Para β_n obtenerla función indicadora toma valores, en el caso del análisis decremental, según la regla

$$I_i = 1, \text{ si } x_{i,t} < x_i \\ I_i = 0, \text{ en cualquier otro caso} \quad (7)$$

El cálculo mediante la expresión (5) puede resultar pesado, aunque se puede realizar mediante alguno de los programas de aplicación estadística comercializados.

En cualquier caso, si el número de datos n es suficientemente elevado, es posible simplificar el cómputo mediante la aproximación normal de la distribución binomial, obteniéndose la significación mediante,

$$\text{Prob}\left(z \geq \frac{2\beta_n - n + 1}{(n-1)^{1/2}}\right) \quad (8)$$

aunque el número de datos habituales en las investigaciones no aconsejan utilizar la aproximación propuesta, pues, para un número reducido de observaciones, existe excesiva diferencia entre la probabilidad asociada al estadístico obtenida con (8) y el valor exacto. Aunque desaconsejamos la aproximación, si nos atenemos a los requisitos tradicionales, puede utilizarse (8) cuando se cumple $\beta_n \geq 5$ y $n - \beta_n - 1 \geq 5$.

Estudio de la violación de supuestos: simulación Monte Carlo

Nos referiremos a los distintos supuestos implícitos en el análisis de incrementos y decrementos, en extremo importantes para garantizar que la tasa de error tipo I empírica se mantenga dentro de los límites admisibles y se corresponda con la tasa exacta. Como ya se ha explicitado, se requieren los supuestos de simetría de $F(\epsilon)$ y la no existencia de dependencia entre observaciones consecutivas. La violación de estos supuestos, desde una perspectiva teórica, deben alterar la probabilidad asociada al estadístico y, por tanto, la tasa empírica de error tipo I. En la presente investigación analizamos el efecto de la asimetría y la autocorrelación sobre la tasa empírica de error tipo I, estimando, mediante simulación Monte Carlo, la magnitud del efecto de esta fuente de error estadístico.

Procedimiento

Las series se generaron mediante simulación Monte Carlo, con longitud, dependiendo del subexperimento, de 6, 11 o 21 datos, que incluyen los datos de ambas fases. Estos valores fueron seleccionados por dos motivos: a) $n = 6$ es el número mínimo de observaciones necesarias para obtener niveles de significación inferiores a 0.05 para el estadístico; b) los tres tamaños de serie establecidos recogen la gama de observaciones típicas en diseños conductuales A-B en investigaciones aplicadas. Se establecieron diferentes niveles de ϕ_i (—0.9 a 0.9, incremento de 0.1), a fin de evaluar el efecto de la dependencia serial sobre la tasa empírica de error tipo I. Cada dato se obtuvo según la siguiente expresión

$$x_{i+1} - x_i = \phi_i x_i + \varepsilon_{i+1}$$

donde $E(\varepsilon) = 0$ y $\text{Var}(\varepsilon) = \sigma^2$ son constantes para todo momento, especificándose tres condiciones para $F(\varepsilon)$: uniforme, normal y exponencial. Para asemejar estas condiciones, establecida la restricción $E(\varepsilon) = 0$, se determinaron los parámetros de las distribuciones para que $\text{Var}(\varepsilon) = 1$.

Al existir dos análisis posibles (incremental o decremental), las anteriores condiciones experimentales se probaron para ambos casos. Se realizaron $3 \times 19 \times 3 \times 2 = 342$ subexperimentos y, en cada uno de éstos se generaron 1000 diseños A-B con un solo dato en la fase A. Para cada serie de observaciones (para todo diseño), los primeros 50 datos no fueron seleccionados, a fin de reducir los efectos artificiales (Greenwood y Matyas, 1990): o sea, mitigar en lo posible la incidencia de valores iniciales (semillas) anómalos del término pseudoaleatorio y estabilizar la serie.

La generación del error aleatorio ε se realizó mediante la función RND del lenguaje de programación BASIC (se utilizó el TURBO-BASIC), que permite obtener

números pseudoaleatorios uniformemente distribuidos en (0,1). Se trata de un generador lineal congruencial que, previamente fue probado respecto a diferentes condiciones de aleatoriedad (prueba de frecuencia mediante el estadístico χ^2 , prueba de la autocorrelación, test de rachas, ajuste a una distribución uniforme mediante la técnica de Kolmogorov-Smirnov y prueba de la mediana), con resultados satisfactorios. En cualquier caso, el principal inconveniente reside en que se trata de un generador con ciclo no excesivamente elevado, y, para mitigar este problema, la semilla se variaba con gran frecuencia en base a lecturas del reloj interno del ordenador. Estos números pseudoaleatorios fueron transformados a variables con distribución uniforme, exponencial y normal mediante la técnica de la inversión, los dos primeros, mientras la última se obtuvo en base a una aproximación racional.

Análisis de datos

Para cada una de las series de datos se calculó el número de incrementos o decrementos, según el caso. La tasa nominal (α_N) de error tipo I se fijó en 0.05, que corresponde a los puntos de corte 5, 9 y 15 en series de longitud 6, 11 y 21, respectivamente. Se contabilizó el número de veces que fue superior o igual a los puntos de corte correspondientes a la tasa nominal establecida. Este último dato, dividido por el total de muestras Monte Carlo, permitió estimar la tasa empírica de error tipo I. A partir de la tasa empírica $\hat{\alpha}_i$ estimada se pudo analizar la correspondencia entre valor estimado y esperado, y conocer la magnitud de la incidencia de la violación de los supuestos de simetría e independencia. Dada la distribución discreta del estadístico, es fundamental diferenciar entre las tasas nominal y exacta de error tipo I, valor, este último, al cual debe ajustarse la tasa empírica. Las tasas

exactas α_i son 0.03125, 0.010742 y 0.020694, respectivamente, para series de $n = 6$, $n = 11$ y $n = 21$. Por otro lado, para determinar si la tasa empírica se corresponde con la tasa exacta, considerando la fluctuación aleatoria, se determinaron los intervalos,

$$\alpha_i \pm Z_{\alpha/2} \left(\frac{\alpha_i - \alpha_i^2}{MC} \right)^{1/2}$$

donde MC denota el número de muestras Monte Carlo. Se fijó, arbitrariamente, $\alpha = 0.05$, obteniéndose los intervalos

$$\begin{aligned} n = 6, & 0.02046 \div 0.04203 \\ n = 11, & 0.00435 \div 0.01713 \\ n = 21, & 0.01187 \div 0.02951 \end{aligned}$$

A partir de estos intervalos es posible determinar si la tasa empírica se mueve dentro de los valores esperados para el valor α fijado.

Complementariamente al análisis descrito, se halló para cada una de las combinaciones de tipo de hipótesis, tamaño de la serie y tipo de distribución del término error (un total de 18 condiciones) un modelo de regresión que permita determinar la tasa empírica de error tipo I en función del valor del término autorregresivo. La inspección de gráficas bivariantes mostró que los datos podían ser representados mediante un modelo de crecimiento exponencial

$$\bar{\alpha}_E = \beta_0 \exp(\beta_1 (\Phi_1 - \gamma))$$

fijándose de forma arbitraria $\gamma = 0$. El análisis de la pertinencia de los modelos mostró una aparente adecuación, estimándose todos los parámetros y determinándose R^2 , además de la razón F. Debemos notar que, en la medida que la tasa empírica de error tipo I está acotada superior

e inferiormente, teóricamente, otros modelos son adecuados; pero en los valores de ϕ_1 estudiados, no existen indicios de punto de inflexión, por lo cual se optó por la modelización expuesta, únicamente aceptable para los valores $-0.9 \leq \phi_1 \leq 0.9$.

Resultados

Respecto a la incidencia de la autocorrelación sobre la tasa empírica de error tipo I, se aprecia (Tablas 1, 2 y 3) un incremento monótono de la misma, resultado constante para cualquier tamaño de la serie, tipo de hipótesis y distribución del error. Cuando la autocorrelación es negativa se nota la infraestimación de la tasa exacta de error tipo I, mientras para valores positivos se obtienen tasas empíricas de error tipo I infladas. Los niveles moderados de autocorrelación muestran incidencia sobre la tasa empírica de error tipo I, incluso en aquellos casos en que la dependencia serial no dista en exceso de $\phi_1 = 0$; sólo en el rango de valores entre -0.2 y 0.1 , considerando únicamente distribuciones del término error simétricas, la autocorrelación no incide en exceso sobre la tasa empírica de error tipo I.

Aunque el tamaño de la serie no cambia notablemente el patrón de resultados, sí es destacable que, cuando $n = 11$, se observa una mayor proximidad entre la tasa empírica de error tipo I y la tasa exacta. Las Tablas 1 y 2 muestran que, cuando el error posee distribución simétrica, las tasas empíricas de error tipo I se ubican dentro del intervalo de probabilidad que recoge las fluctuaciones aleatorias, sin ser excesivamente rigurosos, para valores de autocorrelación comprendidos entre -0.3 y 0.2 para $n=11$. En general, la tasa empírica de error tipo I es menor para $n = 11$ comparativamente a los otros tamaños de la serie considerados en este estudio, aunque este resultado no siempre supone considerar óptima una longitud de serie $n = 11$.

Tabla 1

Tasa empírica de error tipo I ($\hat{\alpha}_1$) estimada con simulación Monte Carlo (MC = 1000). $\epsilon \sim U(-3^{1/2}, 3^{1/2})$, $\gamma_1 = 0$. a y b denotan α_1 infraestimada y α_1 sobreestimada, respectivamente

ϕ_1	n = 6		n = 11		n = 21	
	Increment.	Decrem.	Increment.	Decrem.	Increment.	Decrem.
-0.9	0.003 ^a	0.003 ^a	0.001 ^a	0.000 ^a	0.000 ^a	0.001 ^a
-0.8	0.006 ^a	0.007 ^a	0.000 ^b	0.001 ^a	0.003 ^a	0.001 ^a
-0.7	0.010 ^a	0.007 ^a	0.005	0.001 ^a	0.002 ^a	0.002 ^a
-0.6	0.008 ^a	0.007 ^a	0.000 ^b	0.001 ^a	0.002 ^a	0.004 ^a
-0.5	0.011 ^a	0.012 ^a	0.004 ^a	0.000 ^b	0.004 ^a	0.008 ^a
-0.4	0.012 ^a	0.014 ^a	0.005	0.004 ^a	0.005 ^a	0.011 ^a
-0.3	0.016 ^a	0.015 ^a	0.007	0.002 ^a	0.009 ^a	0.008 ^a
-0.2	0.024	0.017 ^a	0.005	0.011	0.022	0.014
-0.1	0.028	0.027	0.010	0.009	0.008 ^a	0.010 ^b
0.0	0.027	0.036	0.010	0.015	0.012	0.023
+0.1	0.035	0.028	0.015	0.016	0.020	0.021
+0.2	0.058 ^b	0.045 ^b	0.018 ^b	0.024 ^b	0.037 ^b	0.021
+0.3	0.060 ^b	0.047 ^b	0.029 ^b	0.018 ^b	0.045 ^b	0.050 ^b
+0.4	0.071 ^b	0.079 ^b	0.041 ^b	0.042 ^b	0.052 ^b	0.051 ^b
+0.5	0.087 ^b	0.103 ^b	0.049 ^b	0.060 ^b	0.095 ^b	0.078 ^b
+0.6	0.108 ^b	0.134 ^b	0.090 ^b	0.075 ^b	0.111 ^b	0.123 ^b
+0.7	0.174 ^b	0.162 ^b	0.127 ^b	0.124 ^b	0.158 ^b	0.149 ^b
+0.8	0.225 ^b	0.220 ^b	0.188 ^b	0.161 ^b	0.238 ^b	0.213 ^b
+0.9	0.290 ^b	0.270 ^b	0.260 ^b	0.267 ^b	0.280 ^b	0.292 ^b

Tabla 2

Tasa empírica de error tipo I ($\hat{\alpha}_1$) estimada con simulación Monte Carlo (MC = 1000). $\epsilon \sim N(0,1)$, $\gamma_1 = 0$. a , b y c denotan α_1 infraestimada, α_1 sobreestimada e insuficiente número de simulaciones para decidir, respectivamente

ϕ_1	n = 6		n = 11		n = 21	
	Increment.	Decrem.	Increment.	Decrem.	Increment.	Decrem.
-0.9	0.000 ^a					
-0.8	0.001 ^a	0.003 ^a	0.001 ^a	0.000 ^b	0.000 ^a	0.000 ^a
-0.7	0.005 ^a	0.004 ^a	0.000 ^a	0.000 ^b	0.003 ^a	0.001 ^a
-0.6	0.004 ^a	0.007 ^a	0.000 ^a	0.000 ^b	0.003 ^a	0.000 ^a
-0.5	0.007 ^a	0.005 ^a	0.003 ^a	0.004 ^a	0.004 ^a	0.001 ^a
-0.4	0.016 ^a	0.010 ^a	0.003 ^a	0.001 ^a	0.006 ^a	0.005 ^a
-0.3	0.011 ^a	0.010 ^a	0.005	0.003 ^a	0.005 ^a	0.009 ^a
-0.2	0.020 ^a	0.018 ^a	0.006	0.005	0.010 ^a	0.011 ^c
-0.1	0.024	0.018 ^a	0.008	0.011	0.015	0.007 ^a
0.0	0.025	0.042	0.007	0.009	0.019	0.021
+0.1	0.041	0.043 ^b	0.025 ^b	0.014	0.014	0.022
+0.2	0.051 ^b	0.047 ^b	0.020 ^b	0.016	0.029 ^b	0.033 ^b
+0.3	0.064 ^b	0.060 ^b	0.039 ^b	0.033 ^b	0.044 ^b	0.045 ^b
+0.4	0.089 ^b	0.090 ^b	0.043 ^b	0.041 ^b	0.080 ^b	0.051 ^b
+0.5	0.110 ^b	0.129 ^b	0.066 ^b	0.058 ^b	0.097 ^b	0.088 ^b
+0.6	0.133 ^b	0.139 ^b	0.096 ^b	0.097 ^b	0.106 ^b	0.125 ^b
+0.7	0.196 ^b	0.184 ^b	0.133 ^b	0.135 ^b	0.157 ^b	0.161 ^b
+0.8	0.237 ^b	0.241 ^b	0.205 ^b	0.171 ^b	0.204 ^b	0.215 ^b
+0.9	0.307 ^b	0.308 ^b	0.270 ^b	0.291 ^b	0.299 ^b	0.294 ^b

Tabla 3

Tasa empírica de error tipo I ($\hat{\alpha}_i$) estimada con simulación Monte Carlo (MC = 1000). $\varepsilon_i \sim 1 - \text{EXP}(1)$, $\gamma_i = 2$. *a*, *b* y *c* denotan α_i infraestimada, α_i sobreestimada e insuficiente número de simulaciones para decidir, respectivamente

ϕ_i	n = 6		n = 11		n = 21	
	Increment.	Decrem.	Increment.	Decrem.	Increment.	Decrem.
-0.9	0.001 ^a	0.006 ^a	0.000 ^a	0.000 ^a	0.000 ^a	0.001 ^a
-0.8	0.000 ^a	0.017 ^a	0.000 ^b	0.001 ^a	0.000 ^b	0.008 ^a
-0.7	0.001 ^a	0.033	0.000 ^a	0.005	0.000 ^a	0.015
-0.6	0.000 ^b	0.034	0.000 ^b	0.025 ^b	0.000 ^a	0.031 ^b
-0.5	0.002 ^a	0.031	0.000 ^b	0.021 ^b	0.000 ^a	0.048 ^b
-0.4	0.002 ^a	0.056 ^b	0.000 ^b	0.028 ^b	0.000 ^b	0.101 ^b
-0.3	0.000 ^b	0.066 ^b	0.000 ^b	0.031 ^b	0.000 ^a	0.116 ^b
-0.2	0.002 ^a	0.066 ^b	0.000 ^a	0.049 ^b	0.001 ^a	0.153 ^b
-0.1	0.003 ^a	0.085 ^b	0.000 ^b	0.043 ^b	0.000 ^a	0.197 ^b
0.0	0.007 ^a	0.100 ^b	0.001 ^a	0.076 ^b	0.001 ^a	0.191 ^b
+0.1	0.014 ^a	0.107 ^b	0.001 ^a	0.080 ^b	0.002 ^a	0.205 ^b
+0.2	0.023	0.134 ^b	0.004 ^c	0.092 ^b	0.001 ^a	0.202 ^b
+0.3	0.034	0.164 ^b	0.010	0.106 ^b	0.015	0.225 ^b
+0.4	0.039	0.167 ^b	0.013	0.127 ^b	0.013	0.195 ^b
+0.5	0.053 ^b	0.169 ^b	0.034 ^b	0.134 ^b	0.037 ^b	0.210 ^b
+0.6	0.098 ^b	0.220 ^b	0.043 ^b	0.162 ^b	0.077 ^b	0.227 ^b
+0.7	0.139 ^b	0.268 ^b	0.087 ^b	0.197 ^b	0.089 ^b	0.285 ^b
+0.8	0.196 ^b	0.294 ^b	0.168 ^b	0.232 ^b	0.145 ^b	0.314 ^b
+0.9	0.273 ^b	0.342 ^b	0.238 ^b	0.286 ^b	0.269 ^b	0.336 ^b

Como se esperaba, la violación del supuesto de simetría en la distribución del error (Tabla 3) muestra graves consecuencias sobre la tasa empírica de error tipo I, pudiéndose calificar de aberrantes los resultados obtenidos. Destaca la presencia de desajustes entre las tasas empírica y exacta de error tipo I, especialmente cuando los datos fueron generados por un proceso en el cual $\phi_i = 0$. Por otro lado, el comportamiento de la incidencia de la autocorrelación produce tasas aceptables para niveles positivos en el caso incremental, apareciendo un patrón opuesto cuando se realiza el análisis decremental. La longitud de la serie no muestra importantes diferencias en el patrón de resultados expuesto.

En la Tabla 4 se muestra la información correspondiente a la modelización de crecimiento exponencial realizada. En general, las estimaciones de los parámetros

para los casos en los que el término error posee distribución simétrica es similar, sólo dependiendo de la longitud de la serie. Nótese que, aunque en la serie de longitud 11 se obtiene la mayor velocidad de crecimiento, sistemáticamente, la estimación del parámetro β_0 (en algún sentido, un parámetro de magnitud) es inferior. Por otro lado, siempre para distribuciones simétricas del término error, existe una elevada correspondencia entre los valores de β_0 estimados y las tasas exactas de error tipo I, como era previsible, si el modelo de crecimiento exponencial era una elección adecuada. Los valores F y R^2 , además de otros datos complementarios (Ver Tabla 4), proporcionan información sobre la modelización, aunque debe advertirse que R^2 es un indicador excesivamente optimizado. Por un lado, existen sólo 19 pares de datos para estimar los parámetros de la función; por otro, no

Tabla 4

Estimación de parámetros del modelo, error estándar, razón F y coeficiente de determinación correspondientes a las diferentes funciones exponenciales propuestas para relacionar ϕ_1 con \bar{a}_1 .

Distribución	β_0	S.E.	β_1	S.E.	F	R ²
UNIFORME						
<i>Incremental</i>						
n = 60	02580	00192	67950	09781545	20	9907
n = 11	00095	00006	36936	00814	45403	09973
n = 21	00188	00018	30512	01217	13590	09906
<i>Decremental</i>						
n = 6	00285	00014	25124	00643	32122	09954
n = 11	00079	00009	38682	01399	16761	09929
n = 21	00165	00011	31919	00864	29233	09956
NORMAL						
<i>Incremental</i>						
n = 6	00303	00011	25784	00492	58266	09976
n = 11	00119	00008	34854	00827	39013	09969
n = 21	00182	00015	30864	01055	18343	09930
<i>Decremental</i>						
n = 6	00317	00015	25274	00642	32611	09956
n = 11	00087	00009	38591	01336	18468	09936
n = 21	00179	00009	31161	00629	52571	09976
EXPONENCIAL						
<i>Incremental</i>						
n = 6	00101	00007	36737	00861	40323	09971
n = 11	00025	00005	50871	02234	12394	09915
n = 21	00025	00005	51415	02291	12074	09912
<i>Decremental</i>						
n = 6	00927	00032	14567	00504	19386	09883
n = 11	00603	00026	17104	00605	17547	09892
n = 21	01421	00111	09973	01232	1999	08551

posee el mismo valor informativo que en regresión lineal, cuando se obtiene para modelos no lineales. Los gráficos donde se muestran los datos empíricos y la función de crecimiento exponencial obtenida, que han sido omitidos por motivos de extensión, muestran una más que adecuada correspondencia entre ambos. También el análisis de residuos mostró la adecuación del modelo.

Cuando el término de error posee una distribución con fuerte asimetría, el modelo de crecimiento exponencial también es

adecuado, excepto para la hipótesis decremental con $n = 21$. En este caso, los datos empíricos muestran un comportamiento más complejo, especialmente entre $-0.3 \leq \phi_1 \leq 0.3$. Es probable que un problema en la ejecución de esa simulación explicara estos resultados, pero difícilmente compatible con los otros datos disponibles; por lo tanto, creemos que la situación se hace mucho más compleja para otros valores no incluidos en este trabajo. Si el término de error posee distribución asimétrica, no se observa un mismo patrón de resultados

para ambos tipos de hipótesis. Respecto al caso incremental, notamos mayor valor de la velocidad de crecimiento, con estimaciones de β_0 menores que la tasa exacta; por su parte, en el contraste decremental se observa un patrón opuesto. Esta parece ser la consecuencia más destacable de la asimetría del término error sobre el tipo de contraste realizado. Por otro lado, comprendemos mejor cómo las tasas de error tipo I empíricas tienden a equilibrarse si, para $\phi_1 = 0$, muestran un elevado desajuste, hecho que recoge la estimación de β_0 . La razón se halla en las velocidades de crecimiento diferenciales.

Discusión

Parece incuestionable desaconsejar, tanto por razones de orden teórico como por los resultados obtenidos en la simulación, el análisis de incrementos y decrementos, cuando existan dudas sobre la simetría del término error. El estadístico carece de robustez ante la violación del supuesto de simetría. Desde una perspectiva aplicada, debe evitarse recurrir a esta técnica analítica cuando, dada la existencia de *techo y/o suelo* en la variable respuesta, el nivel conductual se aproxima, en algún o distintos momentos (especialmente en este último caso) del período de observación, a los máximos y/o mínimos niveles de respuesta.

Respecto al efecto diferencial de la asimetría del término error sobre la tasa empírica de error tipo I, dependiendo de la hipótesis contrastada, los resultados deben ser función del sentido e intensidad de la asimetría, como argumentaremos. En una de las condiciones experimentales del presente trabajo se fijó una distribución exponencial para el término error, que, para cualquier valor del parámetro determinante de la función, posee asimetría positiva ($\gamma_1 = 2$). Este hecho se traduce en una mayor probabilidad de obtener valores negativos en el término error; por tanto,

cuando se contrasta la hipótesis incremental, las series tienden a presentar menor número de incrementos que el esperado, siempre bajo el supuesto de hipótesis nula, existiendo un punto en el cual los valores de autocorrelación positiva compensan este efecto. Si se prueba la hipótesis decremental, la situación se invierte, pues el número de decrementos supera al esperado por azar, dado el supuesto de hipótesis nula, existiendo un punto en el cual los valores de autocorrelación negativa neutralizan este efecto. Por este motivo, el rango de valores de autocorrelación en los cuales se produce un ajuste entre las tasas empírica y exacta de error tipo I se halla desplazado respecto al valor esperado $\phi_1 = 0$, siempre que la distribución del error sea asimétrica. En síntesis, las distribuciones asimétricas positivas del término error conllevarán los resultados obtenidos en la presente investigación, más acentuados cuanto mayor sea la intensidad de la asimetría y, por otro lado, amortiguados si la asimetría es inferior a $\gamma_1 = 2$; por su parte, cuando el término de error presente distribución con asimetría negativa, cabe esperar un patrón de resultados opuesto al hallado en la presente investigación, compensándose la asimetría con valores de autocorrelación negativos para el contraste incremental, mientras la dependencia serial positiva neutraliza el efecto de la asimetría en la prueba decremental.

Aunque esta investigación no fue planificada para estudiar la potencia de la técnica, los resultados que muestran cómo la tasa empírica de error tipo I es infraestimada cuando existe autocorrelación negativa, indican posible pérdida de potencia en estas condiciones. No existe discrepancia, en este punto, con los resultados hallados en otras investigaciones sobre distintas técnicas utilizadas para el análisis de datos de diseños conductuales (Crosbie, 1986; 1987). Por su parte, la presencia de un componente autorregresivo positivo

incrementa la tasa empírica de error tipo I, aunque esta tasa se sitúa en niveles más aceptables que los hallados en investigaciones realizadas sobre otras técnicas, como *T-test* (Crosbie, 1986), *Split-middle* (Crosbie, 1987) y *Estadístico C* (Crosbie, 1989). Estos resultados no son sorprendentes, puesto que, si la significación del estadístico se fundamenta en la distribución binomial, la presencia de un componente autorregresivo vulnera el supuesto de independencia y, debido a la actuación del mismo sobre la serie, el número de series básicamente incrementales o decrementales será mayor al esperado bajo el supuesto de independencia.

El modelo exponencial parece recoger adecuadamente la relación entre el parámetro autorregresivo y la tasa empírica de error tipo I, proporcionando información complementaria sobre el efecto del tamaño de la serie. Considerando sólo los casos con distribución simétrica del error, los parámetros estimados recogen la similitud observada entre ambos tipos de hipótesis, además de la invarianza respecto a la distribución del error. Para el término error con distribución asimétrica, el patrón es distinto según el tipo de hipótesis considerada; desajuste a la baja entre las tasas exacta y empírica, para hipótesis incrementales, mostrando inversión de la situación, cuando se contrasta la hipótesis decremental.

El modelo fijado en este estudio para establecer la relación entre las observaciones es arbitrario, pues no conocemos el patrón subyacente en el comportamiento humano. Hemos supuesto que la interdependencia existía entre el incremento y el nivel de conducta actual, pero son posibles otras estructuras no estudiadas en este trabajo. Recientemente, a partir de los comentarios realizados por uno de los revisores, probamos el modelo,

$$x_{i,t} - x_i = \phi_1 (x_i - x_{i,t}) + \varepsilon_{i,t}$$

que corresponde a un ARIMA(1,1,0), donde establecimos $\delta = 0$ (parámetro de nivel). Si el comportamiento es modelizado adecuadamente por la anterior expresión, la incidencia de la autocorrelación sobre la tasa empírica de error tipo I es más notable, como mostró una nueva simulación. También estos resultados aconsejan no utilizar el análisis de incrementos-decrementos cuando exista evidencia o sospecha de dependencia serial.

El análisis de incrementos-decrementos aplicado a diseños A-B con un solo dato en la fase A presenta un problema insalvable, consistente en la imposibilidad de conocer la existencia de tendencia en la fase de línea base. Sólo si la tendencia en la fase A es opuesta a la dirección del cambio esperado por la acción del tratamiento o la línea base es estacionaria, está justificado utilizar esta herramienta de análisis de datos. Difícilmente es posible determinar el patrón de conducta en la fase de pre-tratamiento, si tan solo se posee una observación; por tanto, únicamente es factible recurrir a la búsqueda de algún tipo de información retrospectiva que permita decidir de forma aproximada sobre el patrón de conducta en la fase A.

Evidentemente, la técnica aquí analizada se trata de un caso extremo, pues es inusual disponer de una sola medición en la fase A. En ámbitos de investigación básica es improbable obtener un sólo dato para la fase de línea base, además de poco pertinente optar por un diseño A-B, máxime cuando estas estrategias no son las más adecuadas para determinar el efecto de las intervenciones. Únicamente en aquellos casos de la investigación aplicada donde es recomendable una intervención rápida (pensemos en conductas con consecuencias lesivas para el individuo, por ejemplo), no existencia de registros previos e imposibilidad de realizar múltiples mediciones válidas (como la evaluación de intervenciones psicológicas realizadas

sobre individuos con comportamiento antisocial sometidos, mediante mandato judicial, a regímenes de custodia abierta) o, cuando, por otros motivos, el investigador no pueda disponer de más información, podría considerarse analizar los datos mediante la estrategia aquí expuesta.

El punto más destacable en favor de esta estrategia analítica reside en su forma de evaluar la incidencia del tratamiento, pues no se centra en los cambios de nivel, como es habitual en la mayoría de las técnicas aplicadas al análisis de datos procedentes de diseños conductuales. Aunque tampoco es una alternativa a estas técnicas, sí es un complemento, pues el efecto del tratamiento pudiera no producir un cambio de nivel y, por contra, generar unos patrones de estructura interna en la serie de datos. También es cierto que la técnica tal como se ha presentado, en esta versión extrema, únicamente detecta estructuras en las cuales, para dos datos adyacentes, domina el crecimiento o decrecimiento; pero no es posible detectar estructuras de cambio caracterizadas, por ejemplo, por un incremento o decremento en forma de sierra, ni, obviamente, estructuras más complejas en las series. Estos problemas y la generalización a diseños A-B con varios datos en la fase A, además de su extensión a diseños A-B-A, entre otros, han sido recientemente estudiados (Solanas y Sierra, 1993); pero, en base a los resultados del presente trabajo, el análisis de incrementos-decrementos parece poco adecuado en el ámbito de la investigación

psicológica, excepto que sea posible suponer inexistencia de dependencia serial.

A modo de síntesis, ¿qué puede decirse sobre la robustez de la técnica? Si nos referimos al efecto de la dependencia serial, puede considerarse una técnica robusta, al menos en el sentido tradicional de relativa insensibilidad ante leves violaciones de los supuestos (Huber, 1981). En un rango de valores del componente autorregresivo cercanos a la independencia, no hemos hallado una alteración significativa de la tasa empírica de error tipo I; pero, considerando el ámbito en el cual se utilizaría esta estrategia, donde los niveles de dependencia serial pueden ser elevados, parece aconsejable referirnos a una *robustez débil*. La expresión anterior es un intento de reflejar que, aunque la técnica es robusta, quizá no lo suficiente, en general, para analizar datos procedentes de diseños conductuales. Por otra parte, la asimetría del término error sí incide de forma severa, careciendo la técnica de robustez frente a la violación de este supuesto.

Los autores desean agradecer los comentarios y sugerencias realizadas por los revisores del manuscrito.

Notas

(1) Trabajo realizado con la ayuda de una beca predoctoral de formación de profesorado en áreas deficitarias, otorgada por el Departament d'Ensenyament de la Generalitat de Catalunya a Vicenta Sierra.

Referencias

- Ballard, K.D. (1983). The visual analysis of time series data: Issues affecting the assessment of behavioral interventions. *New Zealand Journal*, 12, 69-73.
- Crosbie, J. (1986). *The inability of t Test to control Type I error with typical applied behavioral data*. Manuscrito no publicado.

- Crosbie, J. (1987). The inability of the binomial test to control Type I error with single-subject data. *Behavioral Assessment*, 9, 141-150.
- Crosbie, J. (1989). The inappropriateness of the C statistic for assessing stability or treatment effects with single-subject data. *Behavioral Assessment*, 11, 315-325.
- Edgington, E.S. (1967). Statistical inference $N = 1$ experiments. *The Journal of Psychology*, 65, 195-199.
- Edgington, E.S. (1980a). Random assignment and statistical test for one-subject experiments. *Behavioral Assessment*, 2, 19-28.
- Edgington, E.S. (1980b). *Randomization test*. New York: Marcel Dekker.
- Greenwood, K.M. & Matyas, T.A. (1990). Problems with application of interrupted time series analysis for brief single-subject data. *Behavioral Assessment*, 12, 355-370.
- Huber, P.J. (1981). *Robust statistics*. John Wiley & Sons: New York.
- Jones, R.R.; Weinrott, M.R. & Vaught, R.S. (1978). Effects of serial dependency between visual and statistical inference. *Journal of Applied Behavior Analysis*, 11, 277-283.
- Moore, G.H. & Wallis, W.A. (1943). Time series significance test based on signs of differences. *Journal of the American Statistical Association*, 38, 153-164.
- Revusky, S.H. (1967). Some statistical treatments compatible with individual organism methodology. *Journal of Experimental Analysis of Behavior*, 19, 319-330.
- Solanas, A.; Salafranca, L.I. y Guàrdia, J. (1992). Análisis estadístico β_n de diseños conductuales: Estadístico. *Psicothema*, 1, 253-259.
- Solanas, A. y Sierra, V. (1993). *Estadístico β_n de k-ésimo orden en diseños conductuales A-B*. Comunicación presentada en el III Simposium de Metodología de las Ciencias Sociales y del Comportamiento, Santiago de Compostela.
- White, O.R. (1974). *The «split middle»: a «quickie» method for trend estimation*. Experimental Education Unit, Child Development and Mental Retardation Center, University of Washington.

Aceptado el 16 de mayo de 1994