

## ANÁLISIS DE LA POTENCIA DEL ESTADÍSTICO $C$ MEDIANTE SIMULACIÓN

Roser Bono y Jaume Arnau  
Universidad de Barcelona

El estadístico  $C$  (Young, 1941) constituye una alternativa de análisis adecuada a las series temporales interrumpidas breves o cortas, propias de diseños conductuales. Por este motivo, se lleva a cabo un estudio de simulación de Monte Carlo, con objeto de hallar la potencia estadística de la prueba  $C$  para cada una de las estrategias de uso propuestas por Tryon (1982). Las series generadas permiten estudiar las limitaciones que presenta el estadístico  $C$  y comparar, para distintos patrones de cambio, tres alternativas de análisis posibles cuando la fase de línea base presenta una tendencia (comparación de series utilizando la técnica de la regresión, comparación de series con datos directos y análisis del conjunto de las fases).

*C statistic power analysis through simulation.* The  $C$  statistic (Young, 1941) constitutes a suitable analysis alternative for interrupted brief or short-time series used in behavioral designs. For this reason, a Monte Carlo simulation study is carried out to determine the statistical power of the  $C$  statistic for each of the usage strategies proposed by Tryon (1982). The series generated allow to evaluate the limitations of the  $C$  statistic and to compare the three possible analysis alternatives when the baseline phase shows a trend, for different change patterns (i.e. comparison of series using the regression technique, comparison of series with direct data, and analysis of the set

Un problema que, a partir de la década de los setenta, se plantea a la metodología de investigación comportamental es el de la inferencia del efecto de los tratamientos. En especial cuando se trata de analizar datos de diseños de sujeto único. El inconveniente principal es la reducida cantidad de registros. Ante este hecho, se ha demostrado que no es adecuado el enfoque propuesto por Box & Jenkins (1970). Por

este motivo, se ha desarrollado una serie de procedimientos alternativos que permiten inferir el efecto de la intervención en diseños de series temporales cortas (Algina & Olejnik, 1982; Algina & Swaminathan, 1977, 1979; Edgington, 1975, 1980, 1992; Gorsuch, 1983; Kazdin, 1976; Levin, Marascuilo & Hubert, 1978; Marascuilo & Busk, 1988; Simonton, 1977; Swaminathan & Algina, 1977; Tryon, 1982; Velicer & McDonald, 1984; Wampold & Worsham, 1986).

El objetivo de la presente investigación es estudiar el estadístico  $C$ . Dicho estadístico es una prueba de la aleatoriedad que

---

Correspondencia: Roser Bono  
Departamento de Metodología de las Ciencias del Comportamiento  
Facultad de Psicología. Universidad de Barcelona  
08035 Barcelona (Spain)

fue desarrollada por Young (1941) y recogida, posteriormente, por Tryon (1982) para el análisis de diseños AB con pocas observaciones por fase. Con esta finalidad, Tryon (1982) propone tres estrategias de aplicación. En primer lugar, en la estrategia 1, se usa el estadístico  $C$  para constatar si hay o no tendencia dentro de la fase de línea base. Cuando, según la estrategia 1, la línea base no presenta tendencia, entonces en la estrategia 2 se utiliza el estadístico  $C$  para verificar si hay una tendencia en los datos al considerar conjuntamente la fase de línea base y la de tratamiento. En el caso de que la prueba  $C$  sea estadísticamente significativa, se concluye que el tratamiento es efectivo. Si, de acuerdo con la estrategia 1, la línea base presenta una tendencia significativa, entonces en la estrategia 3 se usa el estadístico  $C$  con objeto de probar si esta tendencia se prolonga en la fase de intervención, o bien se produce un cambio. En caso de que haya un cambio de tendencia, se concluye que el tratamiento es efectivo. Las mayores dificultades aparecen cuando la línea base es no estacionaria. Para estos casos, Tryon (1982) propone dos aplicaciones del estadístico  $C$  basadas en la comparación de series. La más apropiada consiste en ajustar una recta de regresión a los datos de línea base para, luego, realizar una comparación de series sustrayendo los valores obtenidos de la línea ajustada a la fase de pre-intervención de los de la fase de tratamiento (Hayes, 1981). La otra alternativa es más simple, puesto que la comparación de series se lleva a cabo con las puntuaciones directas. En ambas aplicaciones, la existencia de una tendencia en los datos obtenidos de la comparación de series se prueba a través del estadístico  $C$ . Si éste es significativo, se infiere la presencia de un cambio de tendencia. Estos dos métodos de comparación de series tienen una limitación: el estadístico  $C$  no es significativo

cuando ambas fases -línea base y tratamiento- tienen la misma pendiente y hay un cambio abrupto de nivel. Como solución al problema planteado, Tryon (1984) propone llevar a cabo un análisis del conjunto de las fases. De este modo, el estadístico  $C$  será sensible tanto a un cambio de tendencia como de nivel. No obstante, este tipo de análisis no está exento de dificultades, puesto que también será significativo en el caso de que la tendencia de la línea base se prolongue en la fase de tratamiento y no haya ningún cambio de nivel.

En la presente investigación se llevan a cabo tres estudios de simulación de Monte Carlo, con objeto de hallar la potencia estadística de la prueba  $C$  para distintos patrones de cambio.

### Estudio 1

#### *Método*

A fin de evaluar la primera estrategia de Tryon (1982) se ha diseñado un experimento de simulación de Monte Carlo. La generación de los datos, así como todos los cálculos posteriores, se ha llevado a cabo a través de la hoja de cálculo *LOTUS 1-2-3* (versión 3.1+).

Mediante simulación se han generado fases de línea base con longitudes desde  $n_1=8$  a  $n_1=15$  para tres modelos distintos: línea base estable, línea base con variabilidad (baja, media y alta) y línea base con pendiente programada, cuyos valores van de -0.1 a -1 disminuyendo de décima en décima. En total se han generado 22400 fases de pre-intervención, 200 para cada una de las combinaciones.

Cada serie de línea base con ausencia de tendencia ha sido generada de acuerdo a la siguiente fórmula:

$$Y_t = b_0 + e_t \quad (1)$$

donde  $Y_t$  es la observación obtenida en el punto  $t$ ;  $b_0$  es la constante, y  $e_t$  es el término de error.

Las medidas de variabilidad presentes en las fases se han fijado en términos de amplitudes totales. Para ello se establece, arbitrariamente, una media de 10 y unas puntuaciones que se extienden de 8 a 12 (amplitud 4) para una variabilidad baja; de 5 a 15 (amplitud 10) para una variabilidad media, y de 0 a 20 (amplitud 20) para una variabilidad alta. Así, se tiene, para cada tamaño muestral, tres distribuciones de frecuencias con la misma tendencia central, pero de muy distinta variabilidad.

En períodos de no intervención con tendencia, los datos se han generado a partir de la siguiente función matemática:

$$Y_t = b_0 + b_1T + e_t \quad (2)$$

donde  $b_1$  simboliza la pendiente en la fase de pre-intervención y  $T$  es una unidad de tiempo particular en la secuencia observacional. El error aleatorio que se añade a las diferentes series temporales, tanto en este estudio como en los siguientes, ha sido generado de forma que sea normal e independientemente distribuido con media cero y variancia uno.

**Resultados y discusión**

En este estudio, siguiendo la afirmación de Tryon (1982) según la cual una serie es estable cuando el estadístico  $C$  es no significativo, se muestran las tasas de error Tipo I cuando la línea base es estable (tabla 1), y de error Tipo II cuando la fase de pre-intervención presenta distintas amplitudes de variabilidad aleatoria no sistemática (tabla 2).

Como se puede apreciar en la tabla 1, los valores empíricos de  $\alpha$  son estimaciones de los correspondientes valores nominales. Los resultados hallados indican que cuando las series son estables, el estadístico

$C$  es no significativo. Por otro lado, en la tabla 2 se observa que, si se utiliza el concepto de estabilidad en sentido estricto, la aplicación de la prueba  $C$  produce tasas de error Tipo II elevadas cuando en la serie hay variabilidad. En consecuencia, el estadístico  $C$  no discrimina entre una fase estable y una variable. Tan sólo es posible sostener que un conjunto de datos ordenados se comporta de acuerdo con una expectativa al azar.

*Tabla 1*  
Tasas de error Tipo I obtenidas mediante el estadístico  $C$ : series de línea base estable

$n_1$	$\alpha = 0.05$	$\alpha = 0.01$
8	0.0500	0.0075
9	0.0575	0.0100
10	0.0275	0.0050
11	0.0525	0.0075
12	0.0425	0.0100
13	0.0325	0.0025
14	0.0675	0.0125
15	0.0375	0.0075

*Tabla 2*  
Tasas de error Tipo II obtenidas mediante el estadístico  $C$ : series de línea base variable

$n_1$	Variabilidad baja (amplitud 4)		Variabilidad media (amplitud 10)		Variabilidad alta (amplitud 20)	
	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
8	0.9350	0.9950	0.9375	0.9850	0.9475	0.9850
9	0.9425	0.9925	0.9475	0.9975	0.9500	0.9925
10	0.9575	0.9950	0.9400	0.9775	0.9550	0.9950
11	0.9600	0.9925	0.9450	0.9775	0.9600	0.9925
12	0.9325	0.9825	0.9350	0.9825	0.9600	0.9875
13	0.9450	0.9925	0.9600	0.9825	0.9600	0.9950
14	0.9325	0.9925	0.9375	0.9850	0.9775	0.9900
15	0.9375	0.9775	0.9425	0.9925	0.9425	0.9950

Con objeto de hallar en qué medida el estadístico  $C$  es capaz de detectar una tendencia en la serie, se ha construido la tabla 3. En ella se muestra (con  $\alpha=0.05$ ) el límite -línea discontinua- a partir del cual la

potencia del estadístico  $C$  es superior o igual a 0.80, cifra recomendada por la mayoría de metodólogos (Cohen, 1992). Lo que básicamente aporta esta tabla es el número de registros necesarios para una potencia dada, en función de la pendiente y el nivel de significación. Se constata que a mayor número de observaciones, mayor potencia, y que al aumentar el tamaño del efecto, es decir, la pendiente, también aumenta la potencia. Los resultados obtenidos no son alentadores para efectos pequeños. Por ejemplo, se tiene que para pendientes de  $-0.1$  y  $-0.2$  no existe ni un 60% de posibilidades de detectar tal pendiente, cualquiera que sea la muestra. Otro aspecto importante a destacar es que cuando la potencia estadística es baja, la probabilidad de rechazar una hipótesis nula ( $H_0$ ) verdadera puede ser tan sólo un poco más pequeña que la probabilidad de rechazar la  $H_0$  cuando la hipótesis alternativa es verdadera. Así, en la tabla 3 se observa que para una pendiente de  $-0.1$  y  $n_1=10$ , la potencia es de 0.085. Esto significa que la tasa de error Tipo I con respecto a la potencia es muy grande y, en consecuencia, una proporción sustancial de resultados significativos se puede deber a falsos rechazos de la  $H_0$ . Por este motivo, la baja potencia estadística pone en duda tanto los resultados estadísticamente significativos como los no significativos.

A fin de verificar la relación existente entre las variables implicadas se ha llevado a cabo, en este estudio y en los posteriores, un AVAR factorial de bloques de acuerdo con el modelo aditivo. La variable de bloques por la que se ha optado ha sido el tamaño muestral, puesto que el aspecto más interesante de la presente investigación es el estudio del efecto de la pendiente sobre la potencia estadística de la prueba  $C$ . Por otro lado, con objeto de simplificar los cálculos se han seleccionado, dentro de cada bloque, la mitad de los niveles de la variable pendiente.

*Tabla 3*  
Potencia del estadístico  $C$ : series de línea base con tendencia y  $\alpha=0.05$

		Pendiente									
$n_1$	-0.1	-0.2	-0.3	-0.4	-0.5	-0.6	-0.7	-0.8	-0.9	-1	
8	0.050	0.130	0.200	0.395	0.580	0.720	0.800	0.875	0.930	0.925	
9	0.055	0.175	0.290	0.495	0.690	0.830	0.895	0.945	0.970	1.000	
10	0.085	0.165	0.430	0.615	0.770	0.915	0.960	0.990	0.985	1.000	
11	0.100	0.245	0.500	0.730	0.880	0.950	1.000	1.000	0.995	1.000	
12	0.115	0.295	0.640	0.885	0.935	0.985	0.990	1.000	1.000	1.000	
13	0.140	0.335	0.650	0.935	0.975	0.980	1.000	1.000	1.000	1.000	
14	0.150	0.405	0.755	0.920	0.990	0.995	1.000	1.000	1.000	1.000	
15	0.155	0.560	0.835	0.975	1.000	1.000	1.000	1.000	1.000	1.000	

Los resultados del AVAR revelan que tanto el nivel de significación como la pendiente son estadísticamente significativos:  $F(1,63)=32.07$ ,  $p < 0.001$ ;  $F(4,63)=262.30$ ,  $p < 0.001$ , respectivamente. También es significativa la interacción entre estas dos variables:  $F(4,63)=2.52$ ,  $p < 0.05$ , observándose que cuando la pendiente es pequeña ( $-0.1$ ), la potencia es casi nula, y cuando está presente una pendiente de gran tamaño ( $-0.7$  y  $-0.9$ ), la potencia se acerca considerablemente a la unidad, tanto para  $\alpha=0.05$  como para  $\alpha=0.01$ . En cambio, con pendientes medias ( $-0.3$  y  $-0.5$ ), la potencia es significativamente mayor con un valor  $\alpha=0.05$ .

## Estudio 2

### *Método*

Siguiendo el mismo proceso de simulación descrito en el estudio 1, se pone a prueba el estadístico  $C$  con objeto de evaluar la segunda estrategia de Tryon (1982). Las series temporales generadas tienen una longitud que oscila entre  $n_1=8$  y  $n_2=8$  hasta  $n_1=15$  y  $n_2=15$ , línea base estable, pendiente negativa en la fase de tratamiento que va desde  $-0.1$  a  $-1$ , con decrementos de una décima, y cambios de nivel de cero y una desviación estándar ( $\sigma=0$  y  $\sigma=1$ ).

Bajo estas condiciones se han generado un total de 32000 series de tiempo (200 por combinación).

El modelo estadístico seguido, para aquellas situaciones con línea base estable y tendencia en la fase de intervención, es el siguiente:

$$Y_t = b_0 + b_2 X_t T + e_t \quad (3)$$

donde  $b_0$  estima el punto de corte con el eje  $Y_t$  antes de la intervención;  $b_2$  simboliza la pendiente de la fase de post-intervención;  $X_t$  es una variable *dummy* codificada dicotómicamente con 0 antes de la intervención y 1 después de la misma, y  $T$  define la unidad temporal concreta dentro de la secuencia de observaciones. Cuando además de la tendencia en la fase de tratamiento existe un cambio de nivel, el modelo queda representado por la siguiente ecuación:

$$Y_t = b_0 + b_2 X_t T + b_3 X_t + e_t \quad (4)$$

donde  $b_3$  simboliza el cambio de nivel.

**Resultados y discusión**

Las tablas de potencia de la prueba  $C$  ponen de relieve que el porcentaje de identificaciones correctas de un efecto de tratamiento aumenta a medida que incrementa el número de observaciones, así como la pendiente de la fase de intervención (Bono, 1994). En series de línea base estable, pendiente negativa en la fase de tratamiento, sin cambio de nivel y  $\alpha=0.05$ , la prueba  $C$  tiene baja potencia con pendientes de -0.1 y -0.2, cualquiera que sea el tamaño muestral. En cambio, el estadístico  $C$  empieza a ser bastante potente con ocho observaciones por fase y una pendiente de -0.5 (se tiene el 78.5% de confianza de que se ha adoptado la decisión correcta).

Cuando, además, se añade un cambio de nivel de una desviación estándar, la po-

tencia aumenta notablemente, siendo para  $\alpha=0.05$  superior a 0.80, cualquiera que sea su tamaño de muestra y pendiente en la fase de tratamiento (Bono, 1994).

Los resultados del AVAR indican que las variables nivel de significación, pendiente en la fase de tratamiento y cambio de nivel son significativas, con una probabilidad de error menor de 0.001. La interacción más importante es la que se da entre pendiente y cambio de nivel, resultando ser significativa [F(4,133)=106.4,  $p<0.001$ ]. De las representaciones gráficas de esta interacción, con  $\alpha=0.05$  y  $\alpha=0.01$  (figura 1), se concluye que, para aquellas condiciones en que hay una fase de línea base estable y una pendiente negativa baja (-0.1 y -0.3) en la fase de tratamiento, el estadístico  $C$  es sensible a los cambios de nivel, es decir, al aumentar el tamaño del efecto ( $\sigma=1$ ), la potencia incrementa notablemente.

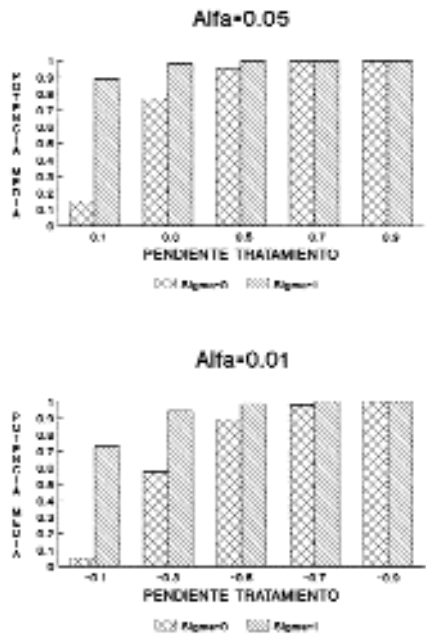


Figura 1. Interacción entre  $\sigma$  y pendiente en la fase de tratamiento.

## Estudio 3

*Método*

Con objeto de evaluar la tercera estrategia de Tryon (1982) se ha seguido el mismo proceso de simulación aplicado en los estudios precedentes. Mediante simulación de Monte Carlo, se han generado series temporales interrumpidas con longitudes desde  $n_1=8$  y  $n_2=8$  a  $n_1=15$  y  $n_2=15$ , a las que se ha aplicado los tres métodos de análisis propuestos por Tryon (1982, 1984): comparación de series utilizando la técnica de la regresión, comparación de series con datos directos y análisis del conjunto de las fases. Las combinaciones llevadas a cabo (200 réplicas para cada una de ellas) son las siguientes: a) longitud de la serie (ocho niveles), tendencia en la línea base y en la fase de tratamiento con pendiente igual  $(-0.1(-0.1)-1)$ , y cambio de nivel en el punto de intervención ( $\sigma=0$  y  $\sigma=1$ ); b) longitud de la serie (ocho niveles), pendiente en la línea base (0.2, 0.4, 0.6, y 0.8), cambio de tendencia en la fase de tratamiento  $(-0.1(-0.1)-1)$ , y cambio de nivel ( $\sigma=0$  y  $\sigma=1$ ), y c) longitud de la serie (ocho niveles), pendiente en la fase de pre-intervención de  $-0.1$ , cambio de tendencia en la fase de tratamiento  $(-0.2(-0.1)-1)$ , y sin cambio de nivel.

Del primer grupo de combinaciones experimentales se han obtenido 96000 simulaciones; 384000 del segundo grupo, y un total de 43200 del último.

Los modelos lineales de la regresión seleccionados quedan representados a través de la siguiente ecuación general:

$$Y_t = b_0 + b_1T + b_2X_tT + b_3X_t + e_t \quad (5)$$

donde  $T$  simboliza la unidad temporal de cada observación;  $X_t$  es una variable ficticia dicotómica con valores 0 para las observaciones anteriores a la intervención

y 1 para las observaciones posteriores;  $b_0$  estima el punto de corte con el eje  $Y_t$  antes de la intervención;  $b_1$  representa la pendiente de la fase de pre-intervención;  $b_2$  es un parámetro que simboliza el cambio de tendencia durante el período post-intervención, y  $b_3$  representa el cambio de nivel.

*Resultados y discusión*

En este estudio se han analizado tres patrones de cambio que se ajustan a la tercera estrategia de Tryon. En la descripción de los resultados se seguirá el orden establecido en el método. Del total de tablas que recogen el error de Tipo I y la potencia estadística de la prueba  $C$  (Bono, 1994), tan sólo se analizan las que corresponden a un valor  $\alpha=0.05$ . El motivo de ello es, por un lado, simplificar los análisis estadísticos y, por otro, que la mayoría de veces es más prudente utilizar un nivel de significación del 5%, a fin de mejorar la potencia.

## A) IGUAL PENDIENTE ENTRE FASES, CON Y SIN CAMBIO DE NIVEL

En el primer subestudio se analiza la principal limitación del estadístico  $C$ . Se trata de la dificultad de identificar un cambio de nivel cuando se aplica un método de comparación de series y tanto la fase de línea base como la de tratamiento presentan la misma pendiente.

La figura 2a recoge los porcentajes de significaciones positivas del estadístico  $C$  para la interacción método de análisis x cambio de nivel. Estos porcentajes corresponden a errores de Tipo I cuando  $\sigma=0$ , y a la potencia del estadístico  $C$  cuando  $\sigma=1$ . El efecto de la interacción es significativo [ $F(2,203)=87$ ,  $p<0.001$ ]. Se observa que los dos métodos de análisis de comparación de series, independientemente del valor de  $\sigma$ , son similares, es decir, tanto la potencia como el error de Tipo I tienen va-

lores muy bajos cuando se aplica alguna técnica de comparación de series. En cambio, el error de Tipo I aumenta considerablemente con un análisis del conjunto de las fases y, en mayor medida, la potencia.

A partir de estos resultados queda claro que es preciso tener mucha cautela al inferir el efecto de un tratamiento cuando la fase de intervención sigue la misma pendiente iniciada en la fase de línea base, puesto que dependerá de si el método de análisis adoptado es el correcto.

**B) TENDENCIAS OPUESTAS ENTRE FASES, CON Y SIN CAMBIO DE NIVEL**

En términos generales, a partir de las tablas de potencia (Bono, 1994), se observa que la potencia estadística es mayor

cuando los efectos de la intervención (cambio de nivel y cambio de tendencia) son grandes y que las posibilidades de detectar efectos pequeños son escasas.

El efecto de la interacción método de análisis x cambio de nivel es significativo [ $F(2,413)=20.91, p<0.001$ ]. La figura 2b pone de manifiesto que los niveles de potencia son mayores para el método de la regresión. Los dos métodos de comparación de series no difieren significativamente en función de los valores de  $\alpha$ ; únicamente en el análisis del conjunto de las fases, la potencia aumenta al presentarse un cambio de nivel de una desviación estándar.

Por lo que respecta a la interacción entre método de análisis y pendiente en la línea base, ésta es significativa [ $F(2,413)=$

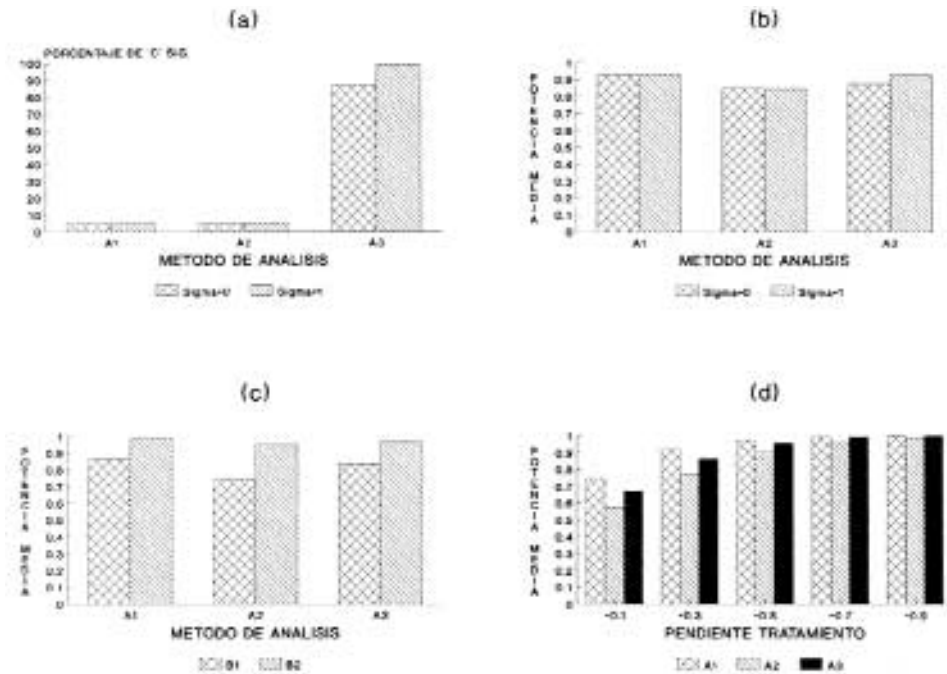


Figura 2. Series con igual pendiente entre fases: (a) interacción entre método de análisis y  $\sigma$ . Series con tendencias opuestas entre fases: interacción entre (b) método de análisis y  $\sigma$ ; (c) método de análisis y pendiente en la línea base, y (d) método de análisis y pendiente en la fase de tratamiento (A<sub>1</sub>= Comparación de series utilizando la técnica de la regresión; A<sub>2</sub>= Comparación de series con datos directos; A<sub>3</sub>= Análisis del conjunto de las bases; B<sub>1</sub>= Pendiente línea base de 0.2; B<sub>2</sub>= Pendiente línea base de 0.6).

12.49,  $p < 0.001$ ]. En la figura 2c queda corroborada la mayor efectividad de la técnica de la regresión, y se observa que al incrementar el tamaño del efecto (en este caso pendiente en la línea base de 0.6), la potencia aumenta. También cabe destacar la menor efectividad del método de comparación de series con datos directos cuando la pendiente de la fase de pre-intervención es de 0.2.

En cuanto al efecto de la interacción método de análisis x pendiente en la fase de tratamiento también es significativo [ $F(8,413)=3.57$ ,  $p < 0.001$ ]. Al analizar dicha interacción, se comprueba que la comparación de series utilizando la técnica de la regresión es más eficaz que el análisis del conjunto de las fases, y este último presenta una mayor potencia que la comparación de series con datos directos (figura 2d).

Al considerar conjuntamente las figuras 2c y 2d, se observa que las diferencias entre los tres métodos de análisis se van diluyendo a medida que aumenta el tamaño del efecto (sea la pendiente en la línea base o en la fase de tratamiento, respectivamente).

C) TENDENCIAS IGUALES ENTRE FASES (MÁS PRONUNCIADAS EN LA FASE DE TRATAMIENTO) Y SIN CAMBIO DE NIVEL

Por último se estudia un patrón de cambio donde la línea base presenta una moderada pendiente que pasa a ser más acusada después de la intervención. La figura 3 recoge la interacción método de análisis x pendiente en la fase de tratamiento. Se observa que con una pendiente de línea base de -0.1 y pendientes mayores en la fase de tratamiento, el orden en que aumenta la potencia de la prueba C, según los distintos métodos de análisis, es el siguiente: comparación de series con datos directos, comparación de series usando la técnica de la regresión, y análisis del con-

junto de las fases. Además, la potencia media incrementa conforme aumenta la pendiente en la fase de intervención. La interacción que se muestra es significativa [ $F(8,98)= 19.83$ ,  $p < 0.001$ ]. La mayor diferencia se produce entre los dos métodos de comparación de series y el análisis del conjunto de las fases, y ésta va disminuyendo a medida que incrementa la pendiente en la fase de tratamiento.

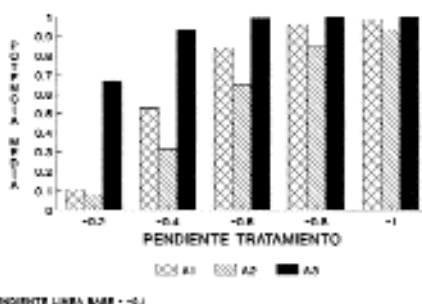


Figura 3. Interacción entre método de análisis y pendiente en la fase de tratamiento, para una pendiente en la línea base de -0.1 (A<sub>1</sub>= Comparación de series utilizando la técnica de la regresión; A<sub>2</sub>= Comparación de series con datos directos; A<sub>3</sub>= Análisis del conjunto de las fases).

Conclusiones

Mediante un estudio de simulación de Monte Carlo, se demuestra que la prueba C no es significativa cuando se trata de series estables o variables. Por otra parte, se confirma que la potencia estadística de C está en función del tamaño de la muestra, y también se constata una relación funcional positiva entre aumento de la potencia e incremento de la pendiente y el nivel de significación. Además, los resultados demuestran que al definir un cambio de nivel de una desviación estándar, la potencia aumenta considerablemente, sobre todo cuando los niveles de pendiente son bajos. Con respecto a la limitación del estadístico



C (fases de línea base y tratamiento con la misma pendiente), ésta queda corroborada a partir del análisis del conjunto de las fases, puesto que la potencia es muy alta cuando  $\sigma=1$ , y la tasa de error Tipo I también es elevada cuando  $\sigma=0$ . Por último, los resultados de la simulación de Monte Carlo muestran que, para series con tendencias opuestas entre fases, el orden con que aumenta la potencia estadística de *C* se ajusta a la siguiente secuencia: compara-

ción de series con datos directos; análisis del conjunto de las fases, y comparación de series utilizando la técnica de la regresión. Cuando las series presentan una línea base con una moderada pendiente que pasa a ser más pronunciada en la fase de tratamiento, el orden en que aumenta la potencia del estadístico *C* es el siguiente: comparación de series con datos directos; comparación de series usando la técnica de la regresión, y análisis del conjunto de las fases.

### Referencias

- Algina, J. y Olejnik, S.F. (1982). Multiple group time-series design: An analysis of data. *Evaluation Review*, 6, 203-232.
- Algina, J. y Swaminathan, H.A. (1977). A procedure for the analysis of time series designs. *Journal of Experimental Education*, 45, 56-60.
- Algina, J. y Swaminathan, H.A. (1979). Alternatives to Simonton's analysis of the interrupted and multiple-group time series designs. *Psychological Bulletin*, 86, 919-926.
- Bono, R. (1994). *Diseños de series temporales interrumpidas: técnicas alternativas de análisis*. Tesis doctoral no publicada, Universidad de Barcelona.
- Box, G.E.P. y Jenkins, G.M. (1970). *Time-series analysis: Forecasting and control*. San Francisco, CA: Holden-Day.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- Edgington, E.S. (1975). Randomization tests for one-subject operant experiments. *Journal of Psychology*, 90, 57-68.
- Edgington, E.S. (1980). Random assignment and statistical tests for one-subject experiments. *Behavioral Assessment*, 2, 19-28.
- Edgington, E.S. (1992). Nonparametric test for single-case experiments. En T.R. Kratochwill y J.R. Levin (Eds.), *Single-case research design and analysis*. Hillsdale, NJ: Erlbaum.
- Gorsuch, R.L. (1983). Three methods for analyzing limited time-series (N of 1) data. *Behavioral Assessment*, 5, 141-154.
- Hayes, S.C. (1981). Single case experimental design and empirical clinical practice. *Journal of Consulting and Clinical Psychology*, 49, 193-211.
- Kazdin, A.E. (1976). Statistical analysis for single-case designs. En M. Hersen y D.H. Barlow (Eds.), *Single case experimental designs: Strategies for studying behavior change*. New York: Pergamon Press.
- Levin, J.R., Marascuilo, L.A. y Hubert, L.J. (1978). N=1 nonparametric randomization test. En T.R. Kratochwill (Ed.), *Single-subject research: Strategies for evaluating change*. New York: Academic Press.
- Marascuilo, L.A. y Busk, P.L. (1988). Combining statistics for multiple-baseline AB and replicated ABAB designs across subjects. *Behavioral Assessment*, 10, 1-28.
- Simonton, D.K. (1977). Cross-sectional time-series experiments: Some suggested statistical analysis. *Psychological Bulletin*, 84, 489-502.
- Swaminathan, H. y Algina, J. (1977). Analysis of quasi-experimental time-series designs. *Multivariate Behavioral Research*, 12, 111-131.
- Tryon, W.W. (1982). A simplified time-series analysis for evaluating treatment interventions. *Journal of Applied Behavior Analysis*, 15, 423-429.
- Tryon, W.W. (1984). "A simplified time-series analysis for evaluating treatment interventions": A rejoinder to Blumberg. *Journal of Applied Behavior Analysis*, 17, 543-544.

Velicer, W.F. y McDonald, R.P. (1984). Time series analysis without model identification. *Multivariate Behavioral Research*, 19, 33-47.

Wampold, B.E. y Worsham, N.L. (1986). Randomization tests for multiple-baseline. *Behavioral Assessment*, 8, 135-143.

Young, L.C. (1941). On randomness in ordered sequences. *The Annals of Mathematical Statistics*, 12, 293-300.

*Aceptado el 17 de noviembre de 1995*