

G*Power: importancia del tamaño muestral en análisis de series temporales

Marcel Elipe Miravet
al226029@uji.es

Patricia Flor Arasil
al228961@uji.es

Francisco Herrero Machancoses
herrerof@uji.es

Jesús Rosel Remírez
rosel@uji.es

Pilar Jara Jiménez
jara@uji.es

Resumen

Introducción: Los análisis de series temporales son especialmente útiles, sobre todo en disciplinas que requieren un seguimiento longitudinal minucioso. A pesar de su gran utilidad, son poco frecuentes en el campo de la psicología. Por este motivo, los autores de este trabajo utilizaron las series temporales para llevar a cabo un estudio sobre la adicción al tabaco. Los resultados mostraron que la conducta tabaquista seguía un modelo AR (2)(7)₈, es decir, la muestra tenía una memoria de 56 días. El objetivo del presente trabajo es comprobar la potencia estadística y el tamaño del efecto del modelo que encontraron.

Método: Dada la ausencia de información en la bibliografía previa, se realizó *a posteriori* un análisis de series temporales imitando los modelos encontrados en los estudios previos. Con la información obtenida, se calculó mediante el *software* G*Power si el tamaño muestral era suficientemente grande para tener un modelo con una buena potencia y un tamaño del efecto.

Resultados: El *output* indica que se necesita un mínimo de 17 sujetos con 63 datos diarios cada uno (1071 datos en total) para tener un modelo con buena potencia estadística y un tamaño del efecto digno.

Conclusión: Los análisis de series temporales tienen poca potencia, por lo que se necesitan registros con un número elevado de datos por sujeto. Además, la cantidad de sujetos para obtener una potencia y un tamaño del efecto adecuados debe ser verificado mediante estudios previos o, si no es posible, mediante análisis *a posteriori*.

Palabras clave: tabaco, series temporales agrupadas, conducta tabaquista, potencia estadística, G*Power.

Abstract

Introduction: Time series analysis is particularly useful, especially in disciplines that require close longitudinal monitoring. Despite their great usefulness, its use is not common in fields such as psychology. For this reason, the authors of this work used time series to carry out a study on tobacco addiction. The results showed that tobacco behaviour followed an AR (2)(7)₈ model, that is, the sample had a 56-day memory. The objective of the present work is to verify the statistical power and the effect size of the model that they found.

Method: Given the absence of information in the previous references, an analysis of time series was performed a posteriori imitating the models founded in the previous studies. It was calculated using G*Power software if our sample size is large enough to obtain a model with statistical power and a good effect size.

Results: The output indicates that a minimum of 17 subjects are needed, with 63 data each day (a total of 1071 data) to obtain a model with a good statistical power and effect size.

Conclusion: To sum up, we conclude with the affirmation that time series analysis has a poor statistical power, so samples for this type of analysis should be quite large. Furthermore, the ideal number of subjects to obtain an adequate statistical power and an effect size should be checked by a previous study or, if that is not possible, a posteriori analysis.

Keywords: tobacco, pooled time series, smoking behaviour, statistical power, G*Power.

Introducción

Puede parecer obvio indicar que la conducta llevada a cabo hoy por una persona es función de la conducta que ha tenido en días anteriores (ayer, anteayer, el mismo día de la semana pasada...); de modo que, si supiéramos qué ha hecho en esos días previos, podríamos anticipar qué hará esa persona el día de hoy. Pues bien, actualmente faltan estudios longitudinales referidos a conductas específicas en psicología y en ciencias de la salud, pese a que varios autores plantean la utilidad de este tipo de estudios en distintos ámbitos (Hamaker y Wichers 2017; Liu y West 2015; Schuurman, Ferrer, de Boer-Sonnenschein y Hamaker 2016; Weiner, Schinka y Velicer 2013).

Aunque los estudios longitudinales han aumentado en los últimos quince años (Liu y West 2015; Hamaker y Wichers 2017), son muy pocos los estudios de series temporales agrupadas que manifiestan la dependencia al tabaco en función de la memoria conductual autorregresiva (AR) de los fumadores. Esto es, el número de publicaciones desde la perspectiva longitudinal metodológica es mucho menor en comparación con otros ámbitos en cuanto a conducta tabaquista. Aun así, existen estudios como el de Leventhal y Cleary (1980) donde se plantean tres posibles modelos sobre la efectividad de la nicotina para mantener el hábito tabaquista. Por su parte, Velicer, Redding, Richmond, Greeley y Swift (1992) han revisado los tres modelos, asignándoles un modelo de series temporales a cada uno de ellos. El primer modelo sería el de *nicotine fixed-effect model*, en el que la nicotina estimularía centros específicos de recompensa del sistema nervioso, manteniendo relativamente constante un nivel de *arousal* en el fumador; a este modelo corresponderían coeficientes AR positivos. Un segundo modelo sería el de *nicotine regulation model*, donde la conducta tabaquista regularía el nivel de nicotina, siendo más afectada por factores externos, pero que provocaría una vuelta inmediata a los niveles anteriores de

nicotina y de conducta tabaquista. Esta conducta reflejaría una autocorrelación *ceró* y, por tanto, un modelo de conducta de *ruido blanco*. El tercer modelo sería el de *multiple regulation model*, donde habría una oscilación en los niveles de tabaquismo, de modo que un descenso en el consumo de tabaco produciría una mayor ansia de fumar, provocando un efecto oscilante en la conducta tabaquista del individuo. A este modelo se correspondería un modelo AR con coeficientes negativos.

Revisando la bibliografía existente, solamente dos estudios se asemejan a nuestra investigación. En primer lugar, Velicer, Colleen, Richmond, Greeley y Swift (1992) se plantearon qué modelo AR era más consistente con los datos de su estudio. Su muestra fue de 10 fumadores, recogiendo datos dos veces al día durante 62 días, es decir, 124 observaciones. En esta investigación encontraron que, en cuanto a la cantidad de cigarrillos fumados, el 70 % de la muestra mostraban un modelo AR de primer orden, o AR(1), es decir, lo que fuma una persona un día determinado es función de lo que fumó un día antes. El 20 % mostraba un modelo AR(1)₂ o, lo que es lo mismo, lo que fuma una persona un día determinado es función de lo que fumó un día antes y dos días antes. Finalmente, el 10 % mostró un modelo AR(0) o de ruido blanco, es decir, la cantidad de cigarrillos fumados un día determinado no estaba influenciada por los cigarrillos fumados en ningún día anterior. Por su parte, Rosel y Elósegui (1994), en una muestra de 29 participantes donde registraron durante 84 días seguidos el número de cigarrillos fumados, hallaron que el 48 % de su muestra seguía un modelo AR(1). El 21 % seguía un modelo AR(1)(1)₇, es decir, lo que fuman el día previo y siete días antes. El 17 % se ajustaba un modelo cíclico de 7 días. Finalmente, el 14 % mostró modelos distintos, donde la máxima memoria conductual seguía un modelo AR(2)(2)₇ o, lo que es lo mismo, su memoria conductual es de 1, 2, 7 y 14 días atrás.

Uno de los inconvenientes de estos estudios es que son de tipo univariado, es decir, solamente una persona por análisis, y, tal como establecen Makridakis, Wheelwright y McGee (1998) y Box y Jenkins (1976), tienen poca potencia estadística. Por lo tanto, si se analizaran todos los sujetos al mismo tiempo (multivariado), aumentaría la potencia estadística (Dielman 1989; Hoepfner, Goodwin, Velicer y Heltshe 2007; Johnson 1995; Moore, Osgood, Larzelere y Chamberlain 1994; Sayrs 1989; Soliday, Moore y Lande 2002; Sun y Li 2002). Este fue el motivo por el cual los autores decidieron realizar un estudio nuevo, aumentando la muestra de participantes para comprobar si un mayor número de sujetos, por una parte, y el análisis de los datos de todos los sujetos conjuntamente, por otra, daban como resultado un mayor número de retardos significativos y un modelo con una potencia estadística elevada.

Para ello realizaron un análisis de series temporales agrupadas multivariado, con una muestra de 62 sujetos universitarios. Obtuvieron como resultado un modelo AR(2)(8)₇, con un porcentaje de varianza explicada del 83,7 %. En resumen, los días previos que influían en lo que fumaba una persona eran el día anterior, 2 días antes, 7, 14, 21... y 56 días atrás. A pesar de obtener un número tan elevado de retardos, no sabían si estos tenían una potencia estadística adecuada y desconocían el tamaño del efecto del modelo encontrado.

Por lo tanto, el principal objetivo de esta investigación es comprobar la potencia estadística y el tamaño del efecto obtenido en el modelo del estudio previo.

Existen dos grandes hipótesis en este trabajo. La primera de ellas es que el modelo autorregresivo obtenido tendría una potencia superior a 0,80, valor mínimo aceptado por la mayoría de manuales estadísticos. La segunda hipótesis del trabajo es que el tamaño del efecto del modelo será superior a 0,40, el cual es considerado como un tamaño del efecto medio según los criterios de la *d* de Cohen.

Método

Participantes

La muestra estaba formada por un total de 62 sujetos universitarios, los cuales dieron su consentimiento para registrar y analizar sus datos. Los participantes tenían que cumplir dos requisitos para poder formar parte en la investigación: 1) llevar fumando más de 2 años seguidos y 2) fumar de media más de 5 cigarrillos diarios.

De los 62 sujetos de la muestra, 36 eran hombres y 26 eran mujeres. La media de edad era de 20,44 años para los hombres (DT = 2,5), 20,37 para las mujeres (DT = 2,05) y 20,41 años para la muestra total (DT = 2,31).

Procedimiento

Cada participante realizó las anotaciones de los cigarrillos fumados mediante una tabla de autorregistro diario. Se les enseñó una serie de estrategias de anotación. Las anotaciones en la tabla de autorregistro se hacían antes de meterse en la cama por la noche. Los participantes recibían cada lunes una tabla de autorregistro diario, válida para una semana. Al siguiente lunes, después de transcurrir 7 días de anotaciones, cada participante entregaba la hoja rellena y se les daba una hoja de registro nueva. Los datos se recogieron durante 84 días seguidos, es decir, 12 semanas. Al finalizar los registros, se obtuvieron un total de 5159 datos registrados. Hubo 7 participantes que finalizaron los registros una semana antes que el resto de sujetos.

Análisis estadísticos

Los datos se almacenaron en el *software* estadístico SPSS versión 23. En este mismo *software* se realizaron los análisis descriptivos y se crearon las series temporales. Una vez organizados los datos mediante un sistema longitudinal-temporal, fueron trasladados y analizados en el *software* estadístico Mplus versión 8 (Muthén y Muthén 2017).

En este *software* se realizaron los análisis de series temporales agrupadas con los datos de los 62 sujetos. Primero, se analizaron todos los retardos para conocer cuántos eran significativos. Posteriormente, se realizó otro análisis de series temporales, esta vez siguiendo el mayor modelo AR encontrado en los estudios previos. Una vez obtenidos los resultados de los análisis, se calculó el r^2 parcial. El coeficiente de determinación parcial es la proporción de varianza que resulta explicada por la introducción de nuevas variables en un modelo explicativo. Una vez obtenidos todos los datos necesarios, se introdujeron en el *software* estadístico G*Power para estimar la muestra que sería necesaria para obtener un modelo con una potencia mínima de 0,80 y un tamaño del efecto aceptable.

Finalmente, y de forma adicional, se realizó un modelo mixto lineal del estudio inicial para verificar los resultados proporcionados por el G*Power y conocer de forma exacta la potencia de cada retardo y el tamaño del efecto del modelo completo.

Resultados

En la tabla 1 se pueden ver los resultados del análisis de series temporales agrupadas realizado *a posteriori* imitando el modelo hallado en la bibliografía previa, es decir, un modelo AR(2)(7)₂.

Tabla 1
Coeficientes del modelo de regresión realizado a posteriori.

Variables	β	S.E.	t	p-value
Lag 1	0,273	0,014	19,004	<0,001
Lag 2	0,140	0,013	10,565	<0,001
Lag 7	0,307	0,014	21,563	<0,001
Lag 14	0,248	0,014	17,744	<0,001

Nota: $r^2 = 0,820$.

Todos los retardos del modelo anterior son significativos. Por su parte, el coeficiente de determinación del modelo es de 0,820.

Una vez conocido el r^2 de los dos modelos, se calculó el r^2 parcial. En nuestro caso, con respecto a los modelos de los estudios previos, se han introducido 6 retardos de componente cíclico semanal, es decir, los retardos 21, 28, 35, 42, 49 y 56; además del día de la semana, introducida como variable *dummy*.

El coeficiente de determinación parcial se obtuvo restando el r^2 del modelo inicial menos el r^2 del modelo calculado *a posteriori*. Como el r^2 de nuestro estudio inicial era de 0,837, el r^2 parcial será 0,017 (0,837-0,820).

Una vez introducidos todos los estadísticos requeridos por el G*Power, se calculó el tamaño muestral necesario para mejorar la varianza explicada en 0,017. En la figura 1 puede verse el *output* de salida de los resultados.

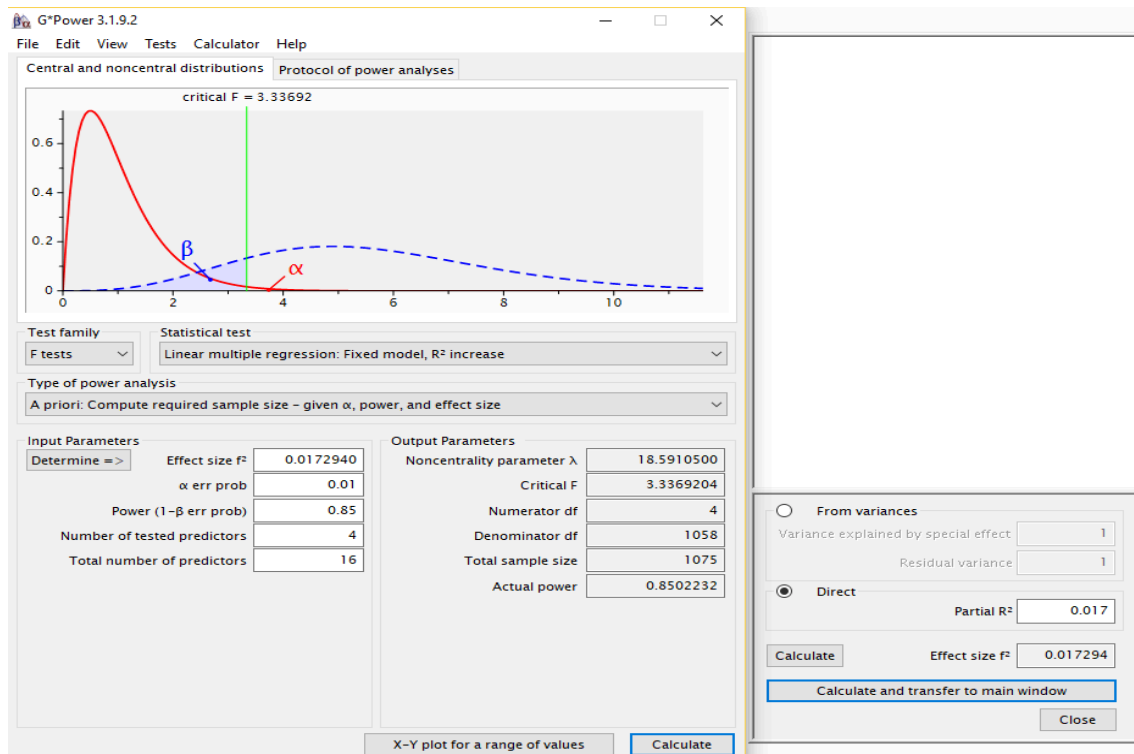


Figura 1. Parámetros de salida del G*Power.

En los parámetros de salida se observa que sería necesario un mínimo de 1075 datos para obtener un modelo estadísticamente significativo, el cual tendría una potencia estimada de 0,85 con un $\alpha = 0,01$. 1071 datos, a razón de 63 datos por persona (con el fin de poder captar un retardo de 56 días, pues se ha de coger algún dato más), haría necesaria una muestra de 17 personas.

De forma adicional, se realizó un modelo mixto lineal del estudio inicial para comprobar si los resultados anteriores son ciertos. En la tabla 2 se pueden ver los resultados con la potencia de cada retardo y el tamaño del efecto de nuestro modelo inicial.

Tabla 2
Coeficientes y potencia del modelo de regresión del estudio inicial.

VARIABLES	β	S.E.	t	p-value	Power
Intercept	0,2618	0,078	3,348	0,001	
Lag 1	0,2468	0,023	10,624	<0,001	1,000
Lag 2	0,0983	0,022	4,501	<0,001	0,944
Lag 7	0,1934	0,023	8,494	<0,001	0,990
Lag 14	0,0867	0,024	3,662	<0,001	0,830
Lag 21	0,0716	0,023	3,065	0,001	0,725
Lag 28	0,0653	0,016	4,179	<0,001	0,757
Lag 35	0,0652	0,016	4,173	<0,001	0,919
Lag 42	0,0523	0,011	4,792	<0,001	0,806
Lag 49	0,0522	0,011	4,783	<0,001	0,783
Lag 56	0,0521	0,011	4,778	<0,001	0,917
Day of the Week (dummy)				<0,001	

Nota: $\eta^2 = 0,873$; $r^2 = 0,837$.

Como se observa en la tabla anterior, todos los retardos tienen una potencia estadística por encima del mínimo aceptado por los manuales en análisis de series temporales. Además, el tamaño del efecto es de 0,873. Finalmente, el coeficiente de determinación es muy elevado (0,837).

Discusión y conclusiones

Los resultados obtenidos en esta investigación confirman las hipótesis que se habían planteado inicialmente. Por una parte, todos los retardos del estudio inicial son estadísticamente significativos, además de tener una potencia estadística muy elevada, por encima del mínimo aceptado (Little 2013). Por otra parte, nuestro modelo inicial cuenta con un tamaño del efecto muy grande. Es decir, la fuerza que tiene la conducta tabaquista previa para predecir lo que una persona puede llegar a fumar en un día determinado es muy grande.

Como se indica en la introducción, los participantes de las investigaciones previas seguían modelos AR con un máximo de 14 días. Sin embargo, en el presente estudio, además de confirmarse que los participantes siguen un modelo AR de 2 semanas ($AR(2)(2)_7$), se ha de añadir un componente cíclico de 7 días con 8 ciclos, es decir, de 56 días ($AR(2)(7)_8$). Este hallazgo demuestra que la cantidad de cigarrillos que fuma una persona un día está influida por el día de la semana y por el número de cigarrillos que fumó 1 día antes, 2 días, 7 días, 14, 21... y hasta 56 días antes, es decir, un total de 8 semanas, lo cual supone una memoria implícita de 56 días atrás. Además, todos los coeficientes AR son positivos, por lo que todos los participantes fuman de acuerdo con el *nicotine fixed-effect model* (Velicer, Redding, Richmond, Greeley y Swift 1992).

El hecho de que las personas fumen con un patrón semanal regular indica que sus hábitos semanales están ligados a normas no escritas de comportamiento, o bien a normas por demandas sociales. Del mismo modo, cuando una persona fuma, implica distintos aspectos de su conducta en la vida cotidiana: conducta psicológica de supuesta disminución de estrés por medio del tabaco, hábitos asociados al tabaco (horas de tabaquismo, otros consumos simultáneos como café o alcohol, fumar después de determinadas comidas) o incluso hábitos de tipo social (fumar en compañía de otros compañeros o amigos).

Una de las ventajas de los resultados obtenidos es que concuerdan con investigaciones previas relacionadas con el abandono del tabaco. Existen estudios donde han hallado que, para que el abandono del hábito tabaquista sea más efectivo, los pacientes deben permanecer al menos 8 semanas bajo algún tipo de programa de tabaquismo, ya que, de lo contrario, existen más probabilidades de recaída a lo largo de los 12 meses siguientes desde el inicio del programa (McAfee et al. 2008; Barrueco et al. 2001). Por lo tanto, los profesionales de la salud deberían planificar terapias o tratamientos no inferiores a 56 días, para conseguir de esta forma mayores porcentajes de éxito en el cese del hábito tabaquista y evitar recaídas en los 12 meses siguientes.

En cuanto a datos se refiere, una de las ventajas de este estudio es la cantidad de observaciones por participante con las que se cuenta (84). Algunos manuales recogen que aquellos estudios con menos de 50 observaciones por participante precisan estimaciones erróneas (Ljung y Box 1978; Box y Pierce 1970; Glass, Willson y Gottman 1975).

Es digno de resaltar que los resultados de nuestro modelo han sido calculados con 1687 datos. Uno de los inconvenientes de las series temporales es que se pierden

tantos datos como retardos significativos se obtienen (Sayrs 1989). Dicho de otro modo, aunque la muestra de 62 sujetos contaba con un total de 5159 datos, había un total de 3472 datos perdidos. De esta forma, cuando en los resultados se ha especificado que se necesita un mínimo de 1075 datos, quiere decir que se requieren 17 sujetos aproximadamente ($1075/63$) para que todas las variables del modelo inicial tengan suficiente potencia estadística (es decir, en el caso de que haya efecto, para que este sea significativo y, si no lo hubiese, para que salga no significativo). En resumen, la muestra final de 62 sujetos supera con creces los 17 participantes mínimos requeridos para que el modelo sea estadísticamente adecuado. Con todo esto, concluimos con la afirmación que nuestra muestra es representativa en lo referido a análisis autorregresivos de series temporales.

Aun así, los hallazgos de esta investigación deben evaluarse en el contexto de varias limitaciones metodológicas. Una de ellas es la falta de generalizabilidad o validez externa, puesto que se ha tomado una muestra solamente compuesta por estudiantes universitarios. Otra limitación importante es que en el presente estudio no se han tenido en cuenta variables psicológicas como el estrés, la ansiedad o el estado de ánimo, las cuales podrían estar mediando en el consumo diario de cigarrillos. Creemos que en futuras investigaciones deberían tenerse en cuenta, pues el consumo de tabaco podría verse afectado por los distintos niveles en esas variables.

Una característica de este estudio es que cuenta con participantes que tienen su hábito tabaquista arraigado, pero no sabemos cuál sería el proceso temporal de las personas que se inician en el hábito. Tal vez presenten modelos de alternancia diaria en su frecuencia tabaquista con *multiple regulation model* y su modelo temporal sea AR con coeficientes negativos o bien dependan de estímulos externos, con lo que fumarían según el *nicotine regulation model*, ajustándose a un modelo de ruido blanco.

Por otra parte, el tabaco es una droga legal con una fuerte composición adictiva, pero cabe preguntarse cuál sería el modelo temporal de conducta adictiva de otras sustancias adictivas legales (alcohol, cafeína, azúcar...) o incluso ilegales (cocaína, heroína, LSD...).

Por todo ello, sería conveniente seguir investigando en este campo de estudio. De esta forma, podrían llegar a replicarse o ampliarse los hallazgos aquí encontrados, con la finalidad de poder aplicarlos en futuras terapias o tratamientos enfocados al abandono del tabaco, mejorando así la salud de la población.

Referencias bibliográficas

- Barrueco, Miguel, Miguel Torrecilla, José Ángel Maderuelo, Carlos Jiménez Ruiz, Miguel Ángel Hernández Mezquita y M. Dolores Plaza. 2001. «Two-month results predictors of outcome in smoking cessation programme». *Medicina Clínica* 116(7): 246-250.
- Box, George E. P., Gwilym M. Jenkins, Gregory C. Reinsel y Greta M. Ljung. 1976. *Time series analysis: Forecasting and control*. San Francisco: Holden-Day.
- Box, George E. P. y David A. Pierce. 1970. «Distribution of residual correlations in autoregressive-integrated moving average time series models». *Journal of the American Statistical Association* 65(332): 1509-1526.
- Dielman, Terry E. 1989. *Pooled cross-sectional and time series data analysis*. New York: Marcel Dekker.

- Glass, Gene V., Víctor L. Willson y John M. Gottman. 1975. *Design and analysis of time-series experiments*. Boulder, CO: Colorado Associated University Press.
- Hamaker, Ellen L. y Marieke Wichers. 2017. «No time like the present». *Current Directions in Psychological Science* 26(1): 10-15.
- Hoeppner, Bettina, Matthew S. Goodwin, Wayne Velicer y James F. Heltshe. 2007. «An applied example of pooled time series analysis: Cardiovascular reactivity to stressors in children with autism». *Multivariate Behavioral Research* 42(4): 707-727.
- Johnson, David R. 1995. «Alternative methods for the quantitative analysis of panel data in family research: Pooled time-series models». *Journal of Marriage and Family* 57(4): 1065-1077.
- Leventhal, Howard y Paul D. Cleary. 1980. «The smoking problem: A review of the research and theory in behavioral risk modification». *Psychological Bulletin* 88(2): 370-405.
- Little, Todd. 2013. *Longitudinal structural equation modeling*. New York: Guilford.
- Liu, Yu y Stephen West. 2015. «Weekly cycles in daily report data: An overlooked issue». *Journal of Personality* 84(5): 560-579.
- Ljung, Greta M. y George E. P. Box. 1978. «On a measure of a lack of fit in time series models». *Biometrika* 65(2): 297-303.
- Makridakis, Spyros, Stephen C. Wheelwright y Víctor E. McGee. 1998. *Forecasting: methods and applications*. New York: John Wiley & Sons.
- McAfee, Timothy, Terry Bush, T. Mona Deprey, Lisa D. Mahoney, Susan M. Zbikowski, Jeffrey L. Fellows y Jennifer B. McClure. 2008. «Nicotine patches and uninsured quitline callers». *American Journal of Preventive Medicine* 35(2): 103-110.
- Moore, Kevin, D. Wayne Osgood, Robert E. Larzelere y Patricia Chamberlain. 1994. «Use of pooled time series in the study of naturally occurring clinical events and problem behavior in a foster care setting». *Journal of Consulting and Clinical Psychology* 62(4): 718-728.
- Muthén, Linda K. y Bengt Muthén. 2017. *MplusUser'sGuide (Versión 8)*. Los Angeles, CA: Muthén&Muthén.
- Rosel, Jesús y Eduardo Elósegui. 1994. «Daily and weekly smoking habits: A Box-Jenkins analysis». *Psychological Reports* 75(3): 1639-1648.
- Sayrs, Lois W. 1989. *Pooled Time series analysis*. Newbury Park, CA: Sage.
- Schuurman, Noémi K., Emilio Ferrer, Mieke de Boer-Sonnenschein y Ellen L. Hamaker. 2016. «How to compare cross-lagged associations in a multilevel autoregressive model». *Psychological Methods* 21(2): 206-221.
- Soliday, Elizabeth, Kevin J. Moore y Marc B. Lande. 2002. «Daily reports and pooled time series analysis: Pediatric psychology applications». *Journal of Pediatric Psychology* 27(1): 67-76.
- Sun, Yongmin y Yuanzhang Li. 2002. «Children's well-being during parents' marital disruption process: A pooled time-series analysis». *Journal of Marriage and Family* 64(2): 472-488.
- Velicer, Wayne F., Colleen A. Redding, Robyn Richmond, Janet Greeley y Wendy Swift. 1992. «A time series investigation of three nicotine regulation models». *Addictive Behaviors* 17(4): 325-345.
- Weiner, Irvin, John A. Schinka y Wayne F. Velicer. 2013. *Handbook of psychology*. Hoboken, NJ: Wiley.