Capítulo 12

ANÁLISIS DE INTERVENCIÓN

George C. Tiao (1933 -)

Estadístico chino. Realizó el doctorado en la Universidad de Wisconsin con George Box y ha sido profesor en las Universidades de Wisconsin y Chicago, donde es actualmente profesor emérito. Ha hecho contribuciones destacadas al estudio de la estacionalidad, los atípicos en series temporales, la incorporación de efectos deterministas, la predicción y los modelos para series temporales múltiples. Ha sido miembro del equipo científico interdisciplinario que descubrió los cambios en la capa de ozono de la tierra con técnicas de series temporales.

12.1 Introducción

Las series reales tienen con frecuencia efectos deterministas que deben tenerse en cuenta para modelarlas. Estos efectos pueden ser cualitativos, como una huelga, un año bisiesto, un cambio legal, un accidente o una festividad, o cuantitativos, como un cambio de base en un índice o un número distinto de días laborables en el mes.

La forma de modelar estos efectos deterministas es introducirlos como variables explicativas en el modelo. Por ejemplo, supongamos que en una serie de producción conocemos que se ha producido una huelga en un momento determinado. Podemos incluir este efecto construyendo una variable que tome el valor cero en todos los instantes de tiempo menos en el momento en que ha ocurrido la huelga, que toma el valor uno. A esta variable la llamaremos variable impulso, porque es diferente de cero únicamente en un punto. Si estudiamos la relación entre los valores de la serie y esta variable impulso podemos estimar el efecto que ha tenido en la serie. Como segundo ejemplo, supongamos en una serie de paro se produce un cambio legal a partir de un instante, ya que se modifica la definición de parado. Podemos representar este efecto mediante una variable que tome el valor cero antes del cambio legal y uno después del cambio. Esta variable se denomina variable escalón y veremos como modelar y estimar sus efectos.

En otras situaciones queremos tener en cuenta efectos que podemos representar mediante una serie conocida, pero que toma valores en todos los periodos. Por ejemplo, si el número de días laborables en el mes es una variable importante, como ocurre en muchas series mensuales económicas o técnicas, podríamos intentar tener en cuenta este efecto dividiendo las observaciones por el número de dias laborables dentro de cada mes y construir una serie de producción media diaria durante el mes. Una solución mejor es introducir una variable explicativa que incluyese los días laborables de cada mes dentro del modelo. Este segundo enfoque tiene la ventaja de que si los dias de la semana tienen distinta actividad, por ejemplo, no son iguales las ventas en un supermercado los lunes que los viernes, podemos investigar este efecto incluyendo siete variables explicativas que tengan en cuenta el número de lunes, martes, etc que existen en el mes. También, muchas series mensuales económicas y sociales están afectadas por festividades que ocurren de manera irregular, como la semana santa. Las predicciones de estas series son más precisas si introducimos este efecto determinista en el modelo, bien como una variable impulso, que toma el valor uno en el mes del año donde aparece la mayor parte de la semana santa y cero en el resto de los meses, o bien como una variable cuantitativa, que tome en cada mes del año el número de días de semana santa que hay en ese mes (esta variable será cero en todos los meses salvo en Marzo y Abril).

Box y Tiao (1975) denominaron análisis de intervención a la inclusión en un modelo de series temporales de efectos deterministas cualitativos mediante variables ficticias. Estudiaremos primero estas variables y luego comentaremos como incluir variables deterministas generales en el modelo de la serie.

12.2 Efectos cualitativos: variables impulso y escalón

Las variables ficticias más utilizadas para representar sucesos cualitativos que afectan a la serie son de dos tipos: *variables impulso* y *variables escalón.* Las variables impulso representan sucesos que ocurren únicamente en un instante, por ejemplo, un accidente, un error de medida o una huelga. La variables escalón representan acontecimientos que comienzan en un instante conocido y se mantienen a partir de ese instante, por ejemplo, una subida de precios, un cambio legal, un cambio de base en un índice, etc.

12.2.1 Variables impulso: función de respuesta a impulsos

Supongamos una variable y_t que sigue un modelo ARIMA que representaremos de forma simplificada por:

$$y_t = \psi(B)a_t. \tag{12.1}$$

Supongamos que esta serie esta afectada en un instante conocido, $t = h$, por un suceso concreto. Podemos representar el instante de ocurrencia de ese efecto mediante una variable impulso que se define por:

$$I_t^h = \begin{Bmatrix} 0 & t \neq h \\ 1 & t = h \end{Bmatrix}$$

La forma más simple de representar el efecto de este suceso sobre la serie es suponer que sólo la modifica en el instante en que ocurre. Llamando ω_0 al tamaño del efecto, el modelo para la serie observada, z_t, será:

$$z_t = \omega_0 I_t^h + \psi(B)a_t \tag{12.2}$$

Este modelo establece que la serie observada, z_t, sigue el modelo ARIMA $z_t = \psi(B)a_t$ en todos los instantes de tiempo en que la variable impulso, I_t^h, es cero. Sin embargo para $t = h$, el dato observado es:

$$z_h = \omega_0 + \psi(B)a_h = \omega_0 + y_h$$

de manera que, en ese instante, observamos la suma del valor de la serie que hubiese ocurrido de acuerdo con el modelo ARIMA, y_h, y que es desconocido, y del efecto determinista de tamaño ω_0.

El efecto de la intervención definida por el impulso I_t^h puede ser más complejo y distribuirse en varios períodos. Por ejemplo, si tenemos una serie diaria de producción y se produce una avería grave en el día h es posible que este efecto se traslade a varios días posteriores y se tarde un cierto periodo en recobrar la normalidad. Por ejemplo, el efecto de una lluvia torrencial que ocurre en $t = h$ puede tener efectos en varios periodos posteriores. Un modelo más general para estos casos es:

$$z_t = w(B)I_t^h + \psi(B)a_t$$

donde $w(B) = (w_0 + w_1 B + ... + w_m B^m)$ es el efecto del impulso que es ahora un polinomio de orden m, lo que permite extender el efecto sobre los m periodos siguientes a la intervención. En efecto, los valores de la serie posteriores a h serán:

$$\begin{aligned} z_h &= w_0 + \psi(B)a_h \\ z_{h+1} &= w_1 + \psi(B)a_{h+1} \\ &\quad .... \\ z_{h+m} &= w_m + \psi(B)a_{h+m} \end{aligned}$$

y los m valores posteriores están afectados por la intervención. La figura 12.1 presenta un ejemplo de las cantidades w_i que se sumarán a los valores de la serie como consecuencia de la intervención. Esta representación se conoce como la *función de respuesta al impulso.*

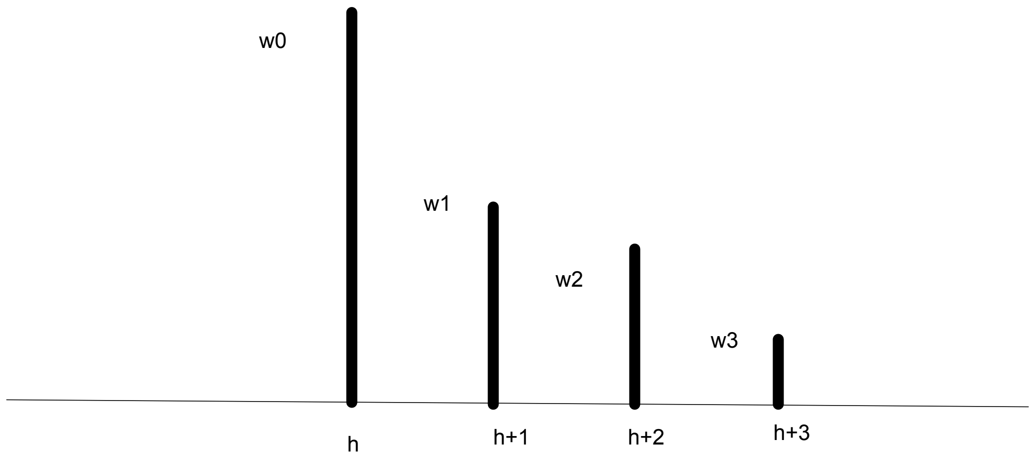

Figura 12.1: Representación de la respuesta de la serie a una intervención con función de respuesta $w_0 + w_1B + w_2B^2 + w_3B^3$

Si el número de periodos afectados por el impulso es largo, la representación anterior puede obligar a estimar muchos parámetros. Por ejemplo, en una serie horaria de consumo de energía electrica un avería en un punto puede afectar durante varias horas a la distribución pero con pesos decrecientes hasta que se establezca la normalidad. Un modelo para representar estos efectos es

$$z_t = \frac{w_0}{1-\delta B} I_t^h + \psi(B)a_t$$

donde $0 < \delta < 1$. En este modelo la función de respuesta a impulso es $w_0(1-\delta B)^{-1} = w_0(1+\delta B+\delta^2B^2+...) = (w_0+w_1B+...+w_mB^m+...)$ donde $w_k = w_0\delta^k$. El efecto de la intervención es ahora un polinomio indefinido, pero con pesos decrecientes que tienden a cero al aumentar el retardo. Los valores de la serie posteriores a h serán:

$$\begin{aligned} z_h &= w_0 + \psi(B)a_h \\ z_{h+1} &= w_0\delta + \psi(B)a_{h+1} \\ &\quad .... \\ z_{h+m} &= w_0\delta^m + \psi(B)a_{h+m} \end{aligned}$$

y todos los valores posteriores al momento de la intervención están afectados por ella, aunque con pesos decrecientes que tienden hacia cero.

12.2.2 Variables escalón: Ganancia

Estas ideas pueden extenderse para modelar intervenciones que tienen un efecto permanente sobre la serie a partir de su ocurrencia, como por ejemplo una subida de precios. Estas intervenciones se modelan con variables "escalón", que se definen mediante:

$$S_t^h = \begin{Bmatrix} 0 & t < h \\ 1 & t \geq h \end{Bmatrix}$$

El efecto de una variable escalón sobre la serie que sigue el modelo ARIMA dado por la ecuación (12.1) puede representarse mediante el modelo de intervención:

$$z_t = w(B)S_t^h + \psi(B)a_t$$

donde $w(B)$ tiene la misma interpretación anterior. En particular si $w(B) = w_0$ todos los valores posteriores al instante h están efectados por una cantidad constante, w_0. Este efecto equivale a un cambio de nivel en la serie a partir del instante h. En el caso general, la figura 12.2 presenta un ejemplo de los efectos que sobre la serie produce una intervención en forma de escalón. Esta función se conoce como *función de respuesta a un escalón*. Se observa que un escalón siempre afecta a todos los valores posteriores a su ocurrencia. Llamaremos *ganancia* de un escalón a su efecto final a largo plazo que es la suma de los efectos parciales y se calcula como

$$Ganancia = w(1) = w_0 + w_1 + ... + w_m.$$

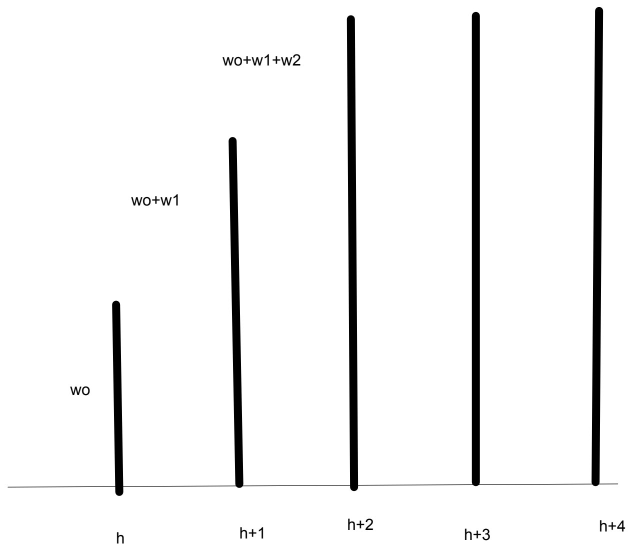

Figura 12.2: Respuesta ante un escalón de la función $(w_0 + w_1 B + w_2 B^2)$

La diferencia básica entre una variable impulso y una escalón es que el efecto de la primera se reduce a un intervalo de tiempo, mientras que el de la segunda se extiende hasta el final del período observado. Suele decirse que el efecto de un impulso en la serie es *transitorio* mientras que el de un escalón es *permanente*.

En determinados casos no se conoce a priori si un efecto es permanente o transitorio. Por ejemplo, un aumento de precios en h puede en algunos productos disminuir las ventas a medio y largo plazo (efecto permanente) mientras que en otros provoca únicamente una leve disminución de ventas en los periodos posteriores a la subida que desaparece en pocos periodos. Para diferenciar ambas respuestas, tendremos en cuenta que un impulso es la primera diferencia de un escalón:

$$I_t^h = S_t^h - S_{t-1}^h = \nabla S_t^h \tag{12.3}$$

Por lo tanto, podemos permitir que los datos nos indiquen si el efecto debe modelarse mediante un impulso o un escalón. Si estimamos el modelo:

$$z_t = (\omega_0 - \omega_1 B) S_t^h + \psi(B) a_t$$

y obtenemos que $\widehat{\omega}_0 \simeq \widehat{\omega}_1$, el modelo se reduce a:

$$z_t = \widehat{\omega}_0 \nabla S_t^h + \psi(B) a_t \tag{12.4}$$

que puede escribirse:

$$z_t = \widehat{\omega}_0 I_t^h + \psi(B) a_t$$

indicando que hay únicamente un efecto transitorio. Por el contrario, si obtenemos $\widehat{\omega}_1 = 0$ o un valor de signo opuesto al de $\widehat{\omega}_1$, esto implica que existe un efecto permanente. Otra manera equivalente de decidir entre un efecto permanente o transitorio es partir de que el efecto es transitorio y estimar el modelo

$$z_t = \frac{\omega_0}{1-\delta B} I_t + \psi(B) a_t$$

Si el parámetro δ se estima como próximo a la unidad, tenemos que, aproximadamente el modelo es

$$z_t = \omega_0 \frac{1}{\nabla} I_t + \psi(B) a_t$$

y utilizando la relación (12.3), el modelo se reduce a

$$z_t = \widehat{\omega}_0 S_t^h + \psi(B) a_t,$$

que representa un efecto permanente.

En resumen, cuando tengamos dudas respecto al tipo de efecto que produce la intervención podemos estimar un modelo que permita ambas posibilidades según los valores de algún parámetro y permitir que los datos nos indiquen el tipo de efecto. Por supuesto, también podemos estimar los dos modelos y seleccionar entre ambos con un criterio de selección.

Ejemplo 12.1

Consideremos las series semanales de cuota de mercado de los dentríficos Colgate y Crest (fichero crescolgate.dat). En Agosto de 1960 la Asociación Dental Americana (America Dental Association, ADA) hizo público que sus estudios sobre la influencia del fluor introducido por Crest habían demostrado un efecto positivo del fluor en la reduccion de caries. Esta declaración corresponde a la observación 135 en la figura 12.3, donde se observa un incremento en el nivel de la serie despúes de esta declaración. Wichern y Jones (1977) estudiaron esta serie y encontraron que esta declaración tuvo efectos permanentes sobre el aumento de la cuota de mercado de Crest.

Para estudiar este efecto vamos permitir que la intervención tenga efectos retardados. Definiremos la variable escalón

$$S_t^{135} = \begin{Bmatrix} 0 & t < 135 \\ 1 & t \geq 135 \end{Bmatrix}$$

y estimaremos el modelo:

$$z_t = (\underset{(.045)}{.065} + \underset{(.045)}{.112B}) S_t^{135} + \frac{(1-.77B)}{\nabla} a_t$$

Este modelo indica que en la semana en que ocurrió la intervención la cuota de mercado de Crest aumento en 6.5 puntos y en la semana siguiente en 11.2. La ganancia o efecto total es la suma de ambos efectos y concluimos que la declaración de la ADA supuso un aumento de la cuota de mercado total para Crest de 17.7 puntos.

Observemos que el modelo estimado equivale a introducir dos variables escalón. La primera, S_t^{135}, corresponde a la semana del anuncio y la segunda , S_t^{136}, definida por:

$$S_t^{136} = \begin{Bmatrix} 0 & t < 136 \\ 1 & t \geq 136 \end{Bmatrix}$$

corresponde a la semana siguiente. El modelo puede también escribirse

$$z_t = .065 S_t^{135} + .112 S_t^{136} + \frac{(1-.77B)}{\nabla} a_t$$

La figura 12.3 presenta el gráfico de la serie y se observa claramente el cambio de nivel que experimenta alrededor del periodo 135.

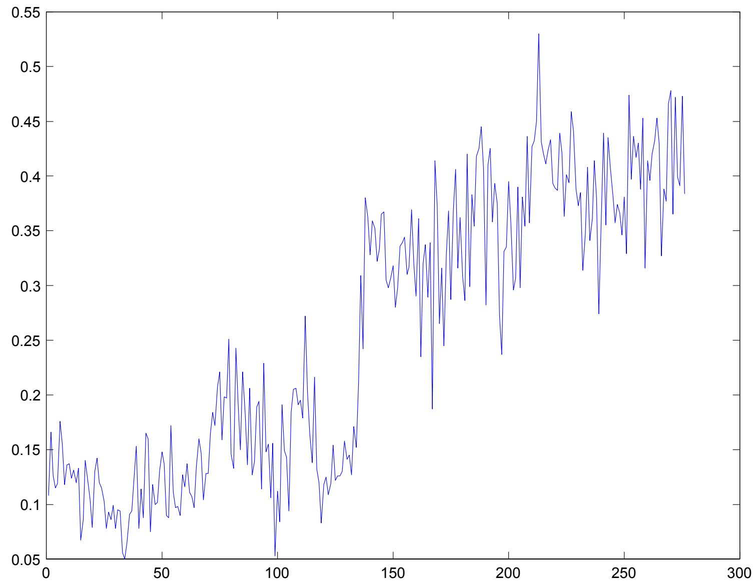

Figura 12.3: Serie semanal de cuota de mercado del dentífrico Crest en el periodo 1958-1963.

12.3 Efectos deterministas generales

Además de efectos impulso y escalón podemos incluir en el modelo un efecto determinista mediante una serie general, x_t, que recoge los efectos que queremos modelar. Por ejemplo, una serie de ventas puede venir afectada por la serie de días laborables con relación a los festivos de cada mes. Como segundo ejemplo, una serie de demanda de servicios pueden estar afectada por los precios que, si son regulados, como ocurre con algunos servicios públicos, formarán una serie determinista de precios.

El efecto de una serie de variables deterministas, x_t, sobre la serie de interés puede representarse por:

$$z_t = f(x_t) + y_t$$

donde z_t es la serie observada que es la suma de dos componentes: y_t la serie limpia del efecto de la variable determinista x_t que sigue el modelo ARIMA (12.1), y $f(x_t)$ el efecto de x_t que se suma a los valores no observados de la variable y_t. Suponiendo que el efecto x_t es lineal podemos escribir :

$$f(x_t) = v_0 x_t + v_1 x_{t-1} + v_2 x_{t-2} + ... = (v_0 + v_1 B + v_2 B^2 + ...)x_t \tag{12.5}$$

donde B es el operador de retardo, definido por $Bx_t = x_{t-1}$.

Una forma de interpretar esta función es analizar la respuesta que obtendríamos si pudiésemos hacer variar a la serie x_t de una manera controlada simple. Por ejemplo, supongamos que x_t es una variable impulso que es cero hasta x_{t-1}, toma el valor 1 en x_t y vuelve después al valor cero, de manera que $x_{t+k} = 0$,para $k \geq 1$. El efecto sobre la variable respuesta de este aumento unitario en x_t será:

$$\begin{aligned}
f(x_{t-1}) &= v_0 x_{t-1} + v_1 x_{t-2} + ... = 0 \\
f(x_t) &= v_0 x_t + v_1 x_{t-1} + ... = v_0 \\
f(x_{t+1}) &= v_0 x_{t+1} + v_1 x_t + ... = v_1 \\
f(x_{t+2}) &= v_0 x_{t+2} + v_1 x_{t+1} + ... = v_2 \\
&\quad .... \\
f(x_{t+k}) &= v_0 x_{t+k} + ... + v_k x_t + ... = v_k
\end{aligned}$$

En consecuencia, cada coeficiente de la función de transferencia, v_i, representa la respuesta en $t+i$ cuando la serie x_t sufre un impulso unitario en el instante t, y la representación de v_k en función del retardo k proporciona la *función de respuesta a impulsos (fri).* Otra forma de caracterizar esta respuesta es estudiar cómo respondería el sistema suponiendo que x_t es una variable escalón que es cero hasta x_{t-1}, y toma el valor 1 en x_t y en todos los valores posteriores. Se obtiene entonces la respuesta

$$
\begin{aligned}
f(x_{t-1}) &= v_0 x_{t-1} + v_1 x_{t-2} + ... = 0 \\
f(x_t) &= v_0 x_t + v_1 x_{t-1} + ... = v_0 \\
f(x_{t+1}) &= v_0 x_{t+1} + v_1 x_t + ... = v_0 + v_1 \\
f(x_{t+2}) &= v_0 x_{t+2} + v_1 x_{t+1} + ... = v_0 + v_1 + v_2 \\
& \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
f(x_{t+k}) &= v_0 x_{t+k} + ... + v_k x_t + ... = \sum_{j=1}^{k} v_j
\end{aligned}
$$

que es la función de respuesta a escalones. Su valor límite es la Ganancia, dada por $\sum_{j=1}^{\infty} v_j$.

La representación general de $v(B)$ puede contener infinitos parámetros y no es operativa. Encontramos este mismo problema en la representación del proceso lineal general cuando estudiamos como aproximar $\psi(B)$ mediante un cociente de dos operadores finitos. En particular, cuando el operador del denominador era la unidad y sólo teníamos el operador finito del numerador, $\theta(B)$, se obtenían los procesos MA, mientras que la combinación de ambos operadores daba lugar a los procesos ARMA. Extendiendo esta idea, podríamos pensar en aproximar $v(B)$ mediante una representación finita del tipo:

$$w_0 + w_1 B + ... + w_m B^m = w_m(B)$$

utilizando las letras w para indicar un número finito de coeficientes. Otra posibilidad es permitir un número infinito de ellos, pero con estructura simple de amortiguamiento, por ejemplo:

$$\frac{w_0}{(1-\delta B)} = w_0 \left(1 + \delta B + \delta^2 B^2 + ...\right)$$

donde suponemos $|\delta| < 1$ de manera que la serie sea convergente. Por analogía con lo estudiado en los modelos ARMA diremos que $w(B)$ representa una aproximación de «media móvil»a $v(B)$, mientras que $w_0(1-\delta B)^{-1}$ es del tipo «autorregresiva». Una representación general es suponer:

$$v(B) = \frac{w_m(B)}{\delta_a(B)}$$

donde

$$w_m(B) = w_0 + w_1 B + ... + w_m B^m; \;\; \delta_a(B) = (1 - \delta_1 B - ... - \delta_a B^a)$$

y llamaremos m y a a los órdenes de los operadores de media móvil y autorregresivo. Observemos que el polinomio $w_m(B)$ no comienza con 1, sino con w_0, y que los coeficientes se toman con signo positivo, como es habitual en regresión.

Para que el sistema sea estable, es decir, para que un incremento finito en x_t no conduzca a un incremento infinito de y_t, las raíces del polinomio $\delta_a(B) = 0$ tienen que estar fuera del círculo unidad. (Esta condición es equivalente a la del operador autorregresivo de un proceso ARMA para asegurar la estacionaridad.) No existen sin embargo restricciones sobre las raíces de $w_m(B)$.

Finalmente, muchos sistemas responden a un impulso en t después de cierto período de retardo que llamaremos b, es decir, los primeros b coeficientes v_i son cero. Tendremos en cuenta este efecto introduciendo un tercer parámetro b y escribiendo:

$$v(B) = \frac{w_m(B)}{\delta_a(B)} B^b \tag{12.6}$$

donde si $b = 0$ la respuesta es instantánea y si $b \neq 0$ comienza con b períodos de retardo. El filtro lineal $v(B)$ queda caracterizado por los órdenes (m, a, b).

Diremos que la relación definida for $v(B)$ es estable si un aumento finito en x_t produce un efecto finito en y_t. En consecuencia, el sistema será estable si la ganancia es finita, $\sum_0^\infty v_i < \infty$, lo que exige que la serie $v_0 + v_1 B + v_2 B^2 + ...$ sea convergente para $|B| \leq 1$. Observemos que al sustituir $B = 1$ en la expresión de $v(B)$ se obtiene precisamente la ganancia. Este resultado es útil para calcular la ganancia de cualquier función de transferencia (m, a, b), ya que:

$$G = v(1) = \frac{w_m(1)}{\delta_a(1)} (1)^b \tag{12.7}$$

es decir

$$G = \frac{w_0 + w_1 + ... + w_m}{1 - \delta_1 - ... - \delta_a}$$

la ganancia será finita si las raíces de $\delta(B) = 0$ están fuera del círculo unidad.

Ejemplo 12.2

Consideremos la serie $z_t = \ln A_t$, donde A_t es la serie de accidentes laborales. Vamos a estimar dos efectos deterministas sobre esta serie. El primero es el número de días laborables del mes, que será una variable determinista D_t. Esta variable la introducimos porque si un mes tiene menos dias laborables sólo por ese efecto debería tener menos accidentes. El segundo es el efecto de Semana Santa, que modelaremos como una variable impulso que toma el valor 1 si la Semana Santa ocurre en ese mes y cero en caso contrario. Por tanto, esta variable sólo toma valores distintos de cero en Marzo y/o Abril. El modelo a estimar será

$$z_t = \omega_1 D_t + \omega_2 I_t + \psi(B) a_t$$

donde $\psi(B)a_t$ es el modelo para la serie. El resultado es

$$\ln A_t = \underset{(12.14)}{0.00936} D_t - \underset{(-6.19)}{.0752}\, I_t + \frac{\underset{(-64.3)}{(1 - 0.34B)}\underset{(-7.3)}{\left(1 - 0.67B^{12}\right)}}{\nabla\nabla_{12}} a_t$$

Este modelo puede escribirse como

$$\ln A_t^C = \ln A_t - \underset{(12.14)}{0.00936} D_t + \underset{(-6.19)}{.0752}\, I_t$$

donde $\ln A_t^C$ es la serie corregida de estos dos efectos. Esta serie sigue el modelo ARIMA :

$$\nabla\nabla_{12} \ln A_t^C = \underset{(-64.3)}{(1 - 0.34B)}\underset{(-7.3)}{\left(1 - 0.67B^{12}\right)} a_t$$

donde hemos puesto bajo paréntesis los valores de la t para el parámetro estimado.

La estimación de este modelo se ha hecho con el programa TSW, donde la variable D_t se genera de forma automática y tiene en cuenta el número de días laborables en el mes. Este programa la define como número de dias laborables (contando de lunes a viernes) menos número de sábados y domingos multiplicados por 5/ 2. De esta manera, la variable D_t será cero si hay exactamente un número fijo de semanas completas en el mes, como en un mes con 28 días que no tiene ninguna fiesta: ese mes tiene 20 días laborables y 8 festivos y $D_t = 20 - 8 \times 5/2 = 0$. La variable será positiva cuando haya relativamente más días laborables en el mes, como en un mes de 31 días que comienza en lunes y tiene 23 días laborables y 8 festivos, $D_t = 23 - 8 \times 5/2 = 3$ y negativa en un mes de 30 días que comienza en sábado y $D_t = 20 - 10 \times 5/2 = -5$. La interpretación del coeficiente de esta variable es que un dia extra laborable sobre la relación 20 y 8 produce un aumento de los accidentes en un .9%. Observemos que un aumento de 1 dia sobre 20 es un aumento de .05, por lo que esos días laborables extras tienen un efecto más que proporcional sobre los accidentes laborales.

La segunda variable, I_t, es una variable impulso que toma el valor uno si la semana santa ocurre en ese mes y cero en caso contrario. Se observa que la aparición de la semana santa supone una disminución de los accidentes ese mes del 7%. Como la semana santa supone dos días adicionales de vacaciones esto supone 2/31=.0667 dias de no trabajo en el mes y el valor obtenido es consistente con estos datos.

La desviación típica residual es 0.05, que indica que el error de predicción de la serie corregida por los efectos deterministas es el 5%. Como comparación, este valor para el mismo modelo sin corrección de efectos deterministas es del 6.6%.

12.4 Construcción de modelos de intervención

La estimación de un modelo de intervención se realiza por máxima verosimilitud de manera similar a la estudiada en el Capítulo 10. La estimación incluye ahora los parámetros del modelo ARIMA y los de la intervención y puede interpretarse como una estimación en etapas. Para ilustrar el procedimiento de estimación consideremos el modelo simple

$$z_t = \omega_0 I_t^h + \frac{1}{1-\phi B} a_t. \tag{12.8}$$

El proceso de estimación por máxima verosimilitud se realiza construyendo la función de verosimilutud y evaluandola con el filtro de Kalman como vimos en el capítulo 10. Este proceso equivale a las tres etapas siguientes:

(1) Partiendo de una estimación inicial del efecto de la intervención, $\widehat{\omega}_0$, se construye la serie corregida del efecto de la intervención:

$$y_t = z_t - \widehat{\omega}_0 I_t^h.$$

(2) Se estiman los parámetros del modelo ARIMA en la serie corregida, y_t de la forma habitual maximizando la verosimilitud. En este caso el problema se reduce a estimará el prámetro ϕ en el modelo AR(1) $y_t = \phi y_{t-1} + u_t$.

(3) Se calculan los residuos del modelo con los parámetros estimados, y se utilizan estos residuos para estimar el efecto de la intervención escribiendo los residuos como función de estos efectos. Para ello se multiplica toda la ecuación por la estructura del modelo ARIMA estimado. En este ejemplo, multiplicando por $1-\widehat{\phi}B$ el modelo (12.8), donde $\widehat{\phi}$ es la estimación de la etapa 2, se obtiene que, llamando $\widehat{e}_t = z_t - \widehat{\phi} z_{t-1}$ a los residuos estimados sobre la serie observada:

$$\widehat{e}_t = \omega_0 (1-\widehat{\phi}B) I_t^h + a_t$$

De esta manera se obtiene una ecuación de regresión donde los residuos $\widehat{e}_t$ son la variable dependiente, la dependiente es la variable $\xi_t = (1-\widehat{\phi}B) I_t^h$, que es conocida. En efecto en este ejemplo $\xi_t = I_t^h - \widehat{\phi}\ I_{t-1}^h$, es una variable que toma siempre el valor cero menos en $\xi_h = 1$ y $\xi_{h+1} = -\widehat{\phi}$. Por tanto tenemos una ecuación de regresión y el coeficiente de regresión es el parámetro a estimar. Escribiendo la ecuación como:

$$\widehat{e}_t = \omega_0 \xi_t + a_t$$

la estimación será $\widehat{\omega}_0 = \sum \widehat{e}_t \xi_t / \sum \xi_t^2 = (\widehat{e}_h - \widehat{\phi} \widehat{e}_{h+1})/(1+\widehat{\phi}^2)$. Con este nuevo valor de ω_0 volveremos a (1) iterando estas tres etapas (1) a (3) hasta obtener convergencia.

La diagnosis de un modelo de intervención se realiza de la misma forma estudiada para los modelos ARMA.

12.5 Estimación de valores ausentes

En algunas series los valores en determinados periodos de tiempo no se conocen. Decimos entonces que la serie tiene valores ausentes. Supondremos que estos huecos ocurren aleatoriamente, de manera independiente del valor de la serie, (es decir, que no faltan precisamente los valores más altos, más bajos u otros con una pauta fija) de manera que la aparición de un valor ausente no proporciona información sobre su magnitud. En estas condiciones el análisis de intervención puede utilizarse para estos valores ausentes. Recordemos que la interpretación de una intervención es que cuando ocurre no observamos el valor del proceso ARIMA, sino una cantidad que incluye el efecto de la intervención. Intuitivamente esto sugiere que cuando exista un valor ausente podemos suponer que ha ocurrido un suceso puntual en esa fecha, suponer un modelo de intervención y estimar el valor de la serie no observado.

Supongamos que la observación h esta ausente. Llamaremos $\mathbf{Z}_{(h)} = (z_1, ..., z_{h-1}, z_{h+1}, ..., z_T)$ a los datos disponibles, donde el subíndice (h) indica que el valor h está ausente. Estimar el valor ausente supone obtener su esperanza condicionada a la información disponible

$$\widehat{z}_h = E(z_h|\mathbf{Z}_{(h)})$$

Si conocemos los parámetros del modelo podemos calcular esta esperanza. Sin embargo, en la práctica los parámetros debemos estimarlos con la información disponible. Vamos a comprobar que podemos obtener simultaneamente esta esperanza y estimar los valores de los parámetros mediante el método siguiente:

(1) Insertar en la posición vacía un valor arbitrario z_h^*, por ejemplo cero, para tener una serie completa, z_t;

(2) Estimar un modelo de intervención para la serie completada, $\mathbf{Z}^* = (z_1, ..., z_{h-1}, z_h^*, z_{h+1}, ..., z_T)$ incluyendo un impulso en la posición de la observación ausente.

Al completar la serie hemos transformado el problema de valores ausentes en un problema de intervención sobre una serie completa, $\mathbf{Z}^*$. La serie z_t no observada en su totalidad que sigue el modelo ARIMA esta relacionada con la observada mediante $z_t = z_t^*$ si $t \neq h$ y $z_h = z_h^* + \omega_0$, donde ω_0 es desconocida. Si estimamos ω_0 entonces conoceremos el valor de z_h. Con este objetivo, estimamos el modelo de intervención:

$$z_t^* = \omega_0 I_t^h + \psi(B)a_t = \omega_0 I_t^h + z_t \tag{12.9}$$

Tomando esperanzas en el instante $t = h$ condicionadas a los valores $\mathbf{Z}_{(h)} = \mathbf{Z}^*_{(h)}$, tenemos que

$$z_h^* = \omega_0 + E(z_h|\mathbf{Z}_{(h)})$$

ya que el valor de z_h es conocido, igual a z_h^*. La esperanza del valor del proceso ARIMA en el instante h será

$$\widehat{z}_h = E(z_h|\mathbf{Z}_{(h)}) = z_h^* - \omega_0 \tag{12.10}$$

Si obtenemos una estimación $\widehat{\omega}_0$ de ω_0 con el modelo (12.9) y la sustituimos en (12.10) tendremos la estimación del valor ausente. Puede demostrase (véase el Apendice 12.1) que al estimar el modelo (12.9) obtenemos simultaneamente una estimación de los parámetros que no depende de z_h^* y un estimador de ω_0 que tampoco depende de z_h^*, es decir la estimación de los parámetros del modelo y la predicción de los valores ausentes.

Como ilustración, supongamos que el modelo es un paseo aleatorio y que falta el valor z_h. Insertando un valor arbitrario z_h^* en esa posición. El modelo (12.9) será

$$z_t^* = \omega_0 I_t^h + \frac{1}{1-B} a_t$$

que equivale a

$$\nabla z_t^* = \omega_0 x_t + a \tag{12.11}$$

donde la variable $x_t = \nabla I_t^h$ es cero salvo para $x_h = 1$ y $x_{h+1} = -1$. El parámetro ω_0 en (12.11) se estimará mediante

$$\widehat{\omega}_0 = \frac{\sum_t x_t \nabla z_t^*}{\sum (x_t)^2} = \frac{\nabla z_h^* - \nabla z_{h+1}^*}{2} = \frac{z_h^* - z_{h-1} - z_{h+1} + z_h^*}{2} = z_h^* - \frac{z_{h-1} + z_{h+1}}{2}$$

y la estimación del valor ausente será, aplicando (12.10):

$$\widehat{z}_h = \frac{z_{h-1} + z_{h+1}}{2}$$

Es decir, el valor ausente se estima como la media de los valores anterior y posterior.

Este enfoque se generaliza sin dificultad si falta un bloque de k observaciones entre h y $h+k-1$. Sea $\mathbf{Z}_{(h,h+k-1)} = (z_1, ..., z_{h-1}, z_{h+k}, ..., z_T)$ la serie disponible. Introduciendo ceros en los k valores ausentes para construir la serie completa $\mathbf{Z}^*$, y definiendo k variables impulso I_t^{h+i-1} para $i = 1, ..., k$, estimaremos el modelo

$$z_t^* = \sum_{i=1}^{k} \omega_{0,i} I_t^{h+i-1} + z_t \tag{12.12}$$

y la estimación de los valores ausentes es obtiene de

$$\widehat{z}_{h+i} = E(z_{h+i}|\mathbf{z}_{(h,h+k-1)}) = z_{h+i}^* - \widehat{\omega}_{0,i}$$

Esta idea se generaliza para cualquier combinación de valores ausentes. Puede demostrarse, véase el apéndice 12.1, que si aplicamos análisis de intervención sobre la serie completada artificialmente rellenando los valores ausentes, los parámetros estimados son aproximadamente los que obtendríamos maximizando la verosimilitud de los datos observados y la estimación de los valores ausentes es aproximadamente su esperanza condicional

Ejemplo 12.3

Para ilustrar la estimación de valores ausentes vamos a suponer que los valores de la serie de accidentes 201 a 205 no se han observado. Podemos poner entonces cero en esas posiciones y estimar el modelo de intervención

$$\ln A_t = \sum_{i=1}^{5} w_i I_{it}^{h_i} + \frac{(1-0.34B)\left(1-0.67B^{12}\right)}{\nabla\nabla_{12}} a_t$$

donde las cinco variables $I_{it}^{h_i}$ para $i = 1, 2, 3, 4, 5$ toman el valor cero en todas las posiciones menos en la h_i, donde h_i =201,...,205. El resultado de la estimación se indica en la tabla siguiente. Puede verse que la estimación es razonablemente buena.

t	201	202	203	204	205
$-\ \widehat{w}$	11.396	11.543	11.530	11.407	11.367
real	11.366	11.465	11.540	11.335	11.397

Apéndice 12.1 Predicción de valores ausentes.

Vamos a demostrar para una serie AR(1) con un valor ausente que estimar los parámetros por máxima verosimilitud conduce, para tamaños muestrales grandes, a un estimador muy similar al obtenido estimando un modelo de intervención sobre una serie completa rellenada con un valor arbitrario. También comprobaremos que la estimación del valor ausente aproxima la esperanza condicionada.

Supongamos que el valor z_h no se ha observado. Entonces la funcion de verosimilitud condicionada es:

$$f(z_2, ..., z_{h-1}, z_{h+1}, ..., z_T|z_1) = f(z_2|z_1)...f(z_{h+1}|z_{h-1})f(z_{h+2}|z_{h+1})...f(z_T|z_{T-1}).$$

Suponiendo normalidad, hemos visto en el capítulo 10 que para $t \geq 2$ y $t \neq h$ la densidad $f(z_h|z_{h-1})$ es $N(\phi z_{h-1}, \sigma^2)$. La distribución de $f(z_{h+1}|z_{h-1})$, para la observación después del hueco, es también normal pero con otros parámetros. Para obtenerla, escribamos

$$z_{h+1} = \phi(\phi z_{h-1} + a_h) + a_{h+1} = \phi^2 z_{h-1} + \phi a_h + a_{h+1}$$

y tomando esperanzas condicionadas a z_{h-1} la media es $\phi^2 z_{h-1}$. Pasando este término al otro lado y elevando al cuadrado para calcular la varianza condicionada se obtiene que esta es $\sigma^2(1+\phi^2)$. Por tanto, concluimos que $f(z_{h+1}|z_{h-1})$ es $N(\phi^2 z_{h-1}, \sigma^2(1+\phi^2))$. Sustituyendo estos valores la función de verosimilitud condicionada, se obtiene la verosimilitud de los datos cuando el valor z_h es ausente:

$$l(\phi, \sigma^2|z_1) = -\frac{(T-2)}{2}\ln\sigma^2 - \frac{1}{2}\ln(1+\phi^2) - \sum_{t\in A}\frac{(z_t - \phi z_{t-1})^2}{2\sigma^2} - \frac{(z_{h+1} - \phi^2 z_{h-1})^2}{2\sigma^2(1+\phi^2)} \tag{12.13}$$

donde $A = \{2, ..., h-1, h+2, ...T\}$.

Vamos a demostrar que con el modelo de intervención se obtiene una función de verosimilitud para los parámetros similar. Por tanto, ambos métodos llevarán a estimaciones de los parámetros similares.

Suponiendo que hemos rellenado la serie con un valor arbitrario para z_h el modelo de intervención que estimaremos es

$$z_t^* = \omega I_t^{(h)} + \frac{a_t}{1-\phi B}$$

que puede escribirse:

$$z_t^* = \phi z_{t-1}^* + \omega I_t^{(h)} - \phi\omega I_{t-1}^{(h)} + a_t$$

En este modelo para $t \neq h,\ h+1$, las variables $z_t = z_t^*$ son normales com media ϕz_{t-1} y varianza σ^2. La variable z_h^* tiene media $\phi z_{t-1} + \omega$ y varianza σ^2 y la z_{h+1} tiene media $\phi(z_h^* - \omega)$ y tambien varianza σ^2. La verosimilitud para los parámetros condicional la primer valor z_1 es entonces,

$$l(\omega,\phi,\sigma^2|z_1) = -\frac{(T-1)}{2}\ln\sigma^2 - \sum_{t\in A}\frac{(z_t-\phi z_{t-1})^2}{2\sigma^2} - \frac{(z_h^*-\phi z_{h-1}-\omega)^2}{2\sigma^2} - \frac{(z_{h+1}-\phi(z_h^*-\omega))^2}{2\sigma^2}. \quad (12.14)$$

Para maximizar esta función observemos que la estimación de ω sólo depende de los valores z_h^* y z_{h+1}. Es decir, sea cual sea el valor de ϕ al derivar la verosimiltud respecto a ω en la función de verosimilitud se obtiene que el estimador de ω debe verificar:

$$(z_h^* - \phi z_{h-1} - \omega) = (z_{h+1} - \phi(z_h^* - \omega))\phi$$

que resulta en:

$$\widehat{\omega} = z_h^* - \frac{\phi}{(1+\phi^2)}(z_{h+1} + z_{h-1}). \quad (12.15)$$

Este estimador puede interpretarse como la diferencia entre el valor arbitrario introducido, z_h^* y su predicción óptima con el resto de la muestra. En efecto, la predicción de z_h que podemos obtener desde sus valores previos es ϕz_{h-1} con varianza σ^2. Desde los valores futuros podemos obtener otra predicción mediante la relación $z_{h+1} = \phi z_h + a_{h+1}$, que implica:

$$z_h = \phi^{-1} z_{h+1} - \phi^{-1} a_{h+1}$$

y que conduce a la predicción $\phi^{-1} z_{h+1}$ con varianza $\phi^{-2}\sigma^2$. La predicción óptima combinara estas dos fuentes de información con peso proporcional a su precisión (inversa de la varianza) para obtener:

$$\widehat{z}_h = \frac{\sigma^{-2}}{\sigma^{-2}+\phi^2\sigma^{-2}}\phi z_{h-1} + \frac{\phi^2\sigma^{-2}}{\sigma^{-2}+\phi^2\sigma^{-2}}\phi^{-1} z_{h+1} = \frac{\phi}{(1+\phi^2)}(z_{h+1} + z_{h-1})$$

Sustituyendo el estimador (12.15) en la verosimilitud (12.14) se obtiene la verosimilitud concentrada que no depende ahora del parámetro ω

$$l(\phi,\sigma^2|z_1) = -\frac{(T-1)}{2}\ln\sigma^2 - \sum_{t\in A}\frac{(z_t-\phi z_{t-1})^2}{2\sigma^2} - \frac{(z_{h+1}-\phi^2 z_{h-1})^2}{2\sigma^2(1+\phi^2)} \quad (12.16)$$

y las funciones (12.14) y (12.16) serán basicamente equivalentes para tamaño muestral moderado. Observemos que estas funciónes sólo difieren en el factor $T-1$ o $T-2$ y en un término, $\frac{1}{2}\ln(1+\phi^2)$, que no depende del tamaño muestral y cuyo efecto está acotado. Para tamaños muestrales grandes su efecto será muy pequeño y ambos procedimientos llevarán a estimadores practicamente idénticos. Puede demostrarse que este resultado es general, véase Peña y Maravall (1991). Gómez, Maravall, y Peña (1999) estudian su implantación y comparan este método con un procedimiento exacto para resolver el problema que utiliza el filtro de Kalman.