Capítulo 13

Valores Atípicos

Carl Harald Cramér (1893-1985)
Matemático y estadístico sueco. Fué el primer catedrático de Estadística Matemática y Ciencia Actuarial en la Universidad de Estocolmo en 1929 y durante cerca de cuarenta años trabajó como consultor para varias compañías de seguros suecas. Hizó contribuciones fundamentales a la teoría de los procesos estacásticos estacionarios y publicó en 1945 un libro de Estadística Matemática que ha sido una de las referencias básicas durante más de 30 años. Fue presidente de la Universidad de Estocolmo desde 1950 hasta su retiro en 1961.

13.1 Introducción

Hemos visto en el capítulo anterior como tener en cuenta efectos deterministas conocidos sobre una serie. Sin embargo, con mucha frecuencia ocurren en las series reales hecho puntuales que desconocemos. Por ejemplo, la serie puede haber estado sometida a intervenciones desconocidas como huelgas, cambios de base, errores de medición etc. Las observaciones afectadas por estas intervenciones pueden presentar una estructura distinta de las demás y aparecer como datos atípicos, es decir, aparentemente no generados igual que las demás. Por ejemplo, si se comete un error de medida o de transcripción en un dato, este valor no mantendrá la estructura de dependencia del resto de la serie. Es importante ser capaz de identificar estas situaciones desconocidas porque:

1. Si sus efectos son grandes pueden sesgar la estimación de los parámetros, lo que producirá a malas predicciones futuras.

2. Si el suceso ha ocurrido en la última parte de la serie y alguna observación afectada se utiliza para generar predicciónes estas no seán buenas, incluso aunque los parámetros estén bien estimados.

3. Si estos sucesos atípicos pueden volver a aparecer en el futuro y los identificamos y estimamos sus efectos, podemos incoporar esta información en las predicciones y obtener intervalos de predicción más realistas.

A continuación estudiaremos como identificar la ocurrencia de estos sucesos atípicos por las huellas que han dejado en la serie.

13.2 Atípicos aditivos

Diremos que ha ocurrido un atípico aditivo (AO) sobre una serie temporal en el instante h si el valor de la serie se genera en ese instante de manera distinta al resto. Por ejemplo, si existe un error de medida apreciable en el instante h que no aparece en las otras observaciones, el dato z_h será un atípico aditivo en la serie z_t. En general, si el dato z_h tiene propiedades distintas del resto por un cambio en las condiciones externas o del sistema de medición, diremos que este data es atípico. El modelo que seguirá la serie observada, z_t si ha sido afectada por un AO en t será:

$$z_t = \begin{cases} y_t & t \neq h \\ y_t + \omega_A & t = h. \end{cases}$$

donde y_t sigue un modelo ARIMA

$$y_t = \psi(B)a_t$$

Entonces, el modelo que sigue la serie observada, z_t, es

$$z_t = \omega_A I_t^{(h)} + \psi(B)a_t \tag{13.1}$$

donde $I_t^{(h)} = 0,\ t \neq h;\ I_h^{(h)} = 1$. Observemos la similitud entre este modelo y el básico de análisis de intervención. Una forma equivalente de escribir el modelo ARIMA para la serie y_t es

$$\pi(B)y_t = a_t$$

donde se verifica que $\psi(B)\pi(B) = 1$, entonces el modelo de un AO es:

$$\pi(B)(z_t - \omega_A I_t^{(h)}) = a_t. \tag{13.2}$$

Las ecuaciones (13.1) y (13.2) son intercambiables y cualquiera de las dos nos sirve para definir el AO.

13.2.1 Efectos en los residuos

La huella que un atípico aditivo deja en la serie es una alteración del valor en un punto. Cuando desconozcamos su presencia y construyamos un modelo ARIMA, su presencia puede notarse en los residuos del modelo. Para ver estos efectos, vamos a suponer que conocemos los verdaderos parámetros del proceso. Para simplificar supondremos primero que el modelos es un AR(1). Entonces, si ignoramos la aparición de un atípico en el instante h y calculamos los residuos suponiendo que todas las observaciones han sido generadas por el mismo modelo, los residuos se calculan con

$$e_t = z_t - \phi z_{t-1}$$

y como el modelo para el AO es, según (13.2),

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

con lo que deducimos que la relación entre los residuos calculados sin saber que hay atípicos y las verdaderas innovaciones es:

$$e_t = \omega_A I_t^{(h)} - \phi\omega_A I_{t-1}^{(h)} + a_t.$$

Podemos concluir que antes de la ocurrencia del suceso atípico en $t = h$ los residuos e_t serán iguales a las verdaderas innovaciones a_t. Sin embargo, pero a partir del instante h la relación entre ambas variables es

$$e_h = \omega_A + a_h$$

$$e_{h+1} = -\phi\omega_A + a_{h+1}$$

y

$$e_{h+j} = a_{h+j} \qquad j \geq 2.$$

Podemos concluir que un AR(1) afectado por un AO tendrá sólo dos residuos afectados: el más afectado es el del instante en que ocurre el atípico, donde el residuo viene modificado por el efecto de la intervención. El residuo posterior también estará afectado, pero en sentido contrario y por una magnitud que es el producto del tamaño del atípico y el parámetro $|\phi < 1|$.

Vamos a generalizar este resultado para un AR(p). Entonces el modelo es

$$\phi(B)z_t = \phi(B)\omega_A I_t^{(h)} + a_t$$

donde $\phi(B)$ es el operador AR de orden p. La relación entre los residuos y las innovaciones será:

$$e_t = \omega_A I_t^{(h)} - \phi_1\omega_A I_{t-1}^{(h)} - \ldots - \phi_p\omega_A I_{t-p}^{(h)} + a_t$$

y los p residuos posteriores a h estarán afectados según la relación

$$e_{h+j} = a_{h+j} - \phi_j \omega_A, \qquad j \geq 0$$

con $\phi_0 = -1$. Ahora es posible que el mayor valor del efecto ocurra en un instante posterior a la ocurrencia del atípico. Por ejemplo en el AR(2) estacionario con operador $(1 - 1.7B + .72B^2)$ el efecto de un atípico aditivo en h sobre las innovaciones será ω_A en h pero $-1.7\omega_A$. es decir un 70% más fuerte, en el instante $h+1$. Concluimos que en un AR(p) los p residuos +posteriores estarán afectados de manera compleja, que depende de los parámetros AR y que la suma de todos los efectos sobre los residuos es $\omega_A(1-\phi_1-...-\phi_p) = \omega_A\phi_p(1)$. Como en un proceso no estacionario $\phi_p(1) = 0$ (por tener una raiz unitaria) en todo proceso no estacionario la suma de los efectos sobre los residuos debe ser cero, y los efectos positivos deben compensarse con los negativos.

Estos resultados se geenralizan sin dificultad para procesos ARIMA, o $\pi(B)y_t = a_t$. Los residuos serán:

$$e_t = \pi(B)z_t$$

y utilizando la relación (13.2), la relación entre los residuos calculados en la serie contaminada, e_t, y las verdaderas inovaciones, a_t, es:

$$e_t = \pi(B)\omega_A I_t^{(h)} + a_t. \tag{13.3}$$

Esta ecuación indica que tendremos un número de residuos afectados por el atípico igual al orden del polinomio $\pi(B)$. En efecto, particularizando esta ecuación para $t = h + j$, para $j = 0, 1, ...$, tenemos:

$$e_{h+j} = -\pi_j \omega_A + a_{h+j} \qquad j \geq 0$$

donde $\pi_0 = -1$, y si todos los π_j son no nulos todos los residuos posteriores al instante de ocurrencia del atípico estarán afectados por el atípico ocurrido en el instante h.

En todo lo anterior hemos supuesto que hay un único AO en la serie. Cuando hay varios, sus efectos se superponen y pueden afectar mucho a la estructura de la serie. Por ejemplo, un proceso AR(5) que sufre un 5% de atípicos aditivos puede llegar a tener el 25% de los residuos afectados.

13.2.2 Efectos en la estimación de los parámetros

El efecto de un AO es sesgar todos los coeficientes de autocorrelación hacia cero, sesgando también la estimación de los parámetros del proceso hacia cero. Ilustraremos este efecto en el caso de un proceso AR(1) de media cero donde ocurre un AO en el instante h. La estimación del parámetro será:

$$\widehat{\phi} = \frac{\sum(z_t - \overline{z})z_{t-1}}{\sum(z_t - \overline{z})^2}$$

utilizando que $z_h = y_h + \omega_A$, y $z_t = y_t$, $t \neq h$, entonces $\overline{z} = \overline{y} + \frac{1}{T}\omega_A$. Suponiendo que T es grande de manera que podemos suponer $\overline{z} \simeq 0$, podemos escribir,

$$\widehat{\phi} = \frac{\sum y_t y_{t-1} + \omega_A(y_{h-1} + y_{h+1})}{\sum y_t^2 + \omega_A^2 + 2\omega_A y_h}$$

y dividiendo ambos miembros por $\sum y_t^2 = Ts_y^2$ y llamando $\widehat{\phi}_0 = \sum y_t y_{t-1} / \sum y_t^2$ a la estimación obtenida con la serie sin atípicos, tenemos que

$$\widehat{\phi} = \frac{\widehat{\phi}_0 + T^{-1}\widetilde{\omega}_A(\widetilde{y}_{h-1} + \widetilde{y}_{h+1})}{1 + T^{-1}\widetilde{\omega}_A(2\widetilde{y}_h + \widetilde{\omega}_A)}$$

donde hemos llamado $\widetilde{\omega}_A = \omega_A/s_y$ y $\widetilde{y}_t = y_t/s_y$ a los valores estadarizados por la desviación típica de la serie. Esta expresión muestra que si el atípico es grande con relación a la variabilidad de la serie el coeficiente estimado irá a cero, ya que $\widetilde{\omega}_A$ aparece linealmente en el numerador y al cuadrado en el denominador, por lo que $\widehat{\phi} \to 0$ si $\widetilde{\omega}_A \to \infty$.

En general el efecto de un atípico aditivo depende mucho del tamaño muestral y para tamaños muestrales muy grandes un AO de tamaño moderado puede tener un efecto pequeño.

Ejemplo 13.1

La figura 13.1 presenta una muestra de 100 observaciones de un proceso AR(1) contaminado por un AO en el instante 60 de seis desviaciones típicas y la figura 13.2 la función de autocorrelación del proceso.

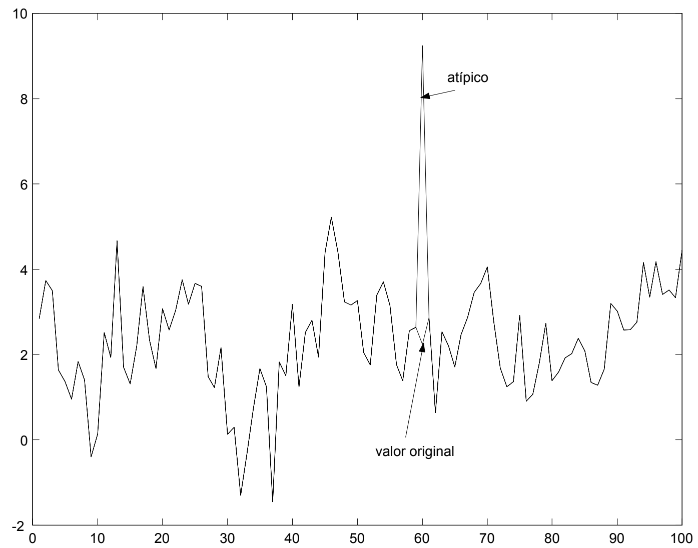

Figura 13.1: Efecto de un AO sobre un AR(1).

Se observa que todos los coeficientes de autocorrelación de los primeros retardos tienen un sesgo hacia cero. Como el parámetro AR es .6 y $.6^6 = .04$, para retardos altos los coeficientes reales son prácticamente cero, pero includo en estos valores la figura muestra que las estimaciones del proceso contaminado se aproximan más a cero en valor absoluto que los del proceso original.

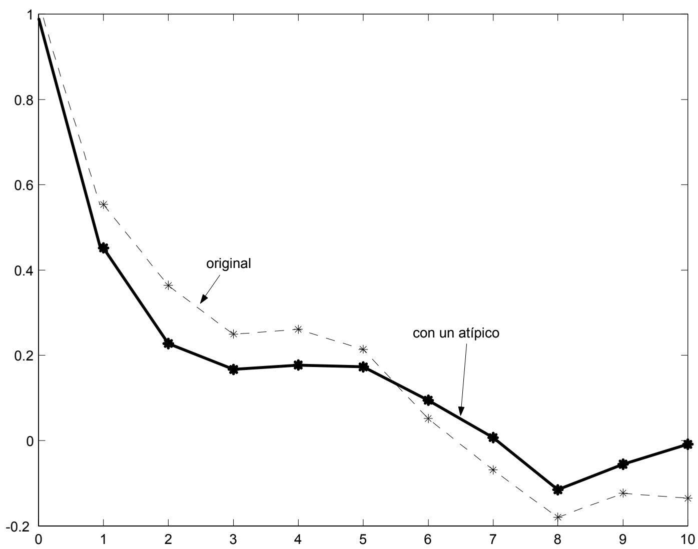

Figura 13.2: Función de autocorrelación del proceso AR(1) contaminado por un atípico

13.3 Atipicos innovativos (IO)

Diremos que ha ocurrido un atípico innovativo (IO) en una serie temporal en $t = h$ cuando cuando la innovación en ese punto esta directamente afectada por una cantidad desconocida debida a un suceso atípico. Recordemos que la innovación de una serie temporal representa el efecto agregado de todas las variables que influyen sobre la serie observada de forma aleatoria. Un atípico sobre la innovación puede interpretarse como un cambio imprevisible en las variables que afectan a la evolución de la serie temporal. Otra forma de verlo es recordar que la innovación representa el error de previsión a un paso. Un IO es un incremento del error de previsión de la serie en un punto debido a un cambio en las condiciones externas.

El modelo para la serie que un sufre un atípico innovativo de magnitud ω_I en el instante h es

$$z_t = \psi(B)(\omega_I I_t^{(h)} + a_t) \tag{13.4}$$

o también, llamando $e_t = \pi(B)z_t$ como en la sección anterior a los residuos calculados suponiendo que la serie esta libre de atípicos y que conocemos los parámetros del modelo:

$$e_t = \omega_I I_t^{(h)} + a_t, \tag{13.5}$$

que podemos escribir como:

$$e_h = a_h + \omega_I, \tag{13.6}$$

y para $j > h$

$$e_{h+j} = a_{h+j}$$

Por tanto, cuando ocurren un IO en $t = h$ los residuos estimados conociendo los parámetros del proceso serán igual a las innovaciones en todos los puntos menos en el de la ocurrencia del IO. De esta definición se concluye que un AO y IO son idénticos sobre una serie que sea ruido blanco. En otro caso, el efecto de un IO sobre la serie es muy distinto que el de un AO ya que este último produce un efecto fijo, alteración de una observación, mientras que el efecto de un IO sobre la serie depende del modelo, como vemos a continuación

13.3.1 Efecto de un IO sobre la serie

Supongamos que la serie es estacionaria. Utilizando la definición (13.4) y escribiendo

$$z_t = y_t + \psi(B)\omega_I I_t^{(h)}$$

donde $y_t = \psi(B)a_t$ es la serie sin contaminar que sigue el modelo ARMA. Según esta ecuación la relación entre las observaciones contaminadas, z_t, y las originales, y_t, es:

$$z_t = \begin{cases} y_t & t < h \\ y_t + \omega_I \psi_j & t = h + j, \quad j \geq 0. \end{cases} \tag{13.7}$$

Esta expresión indica que

(1) Para t=h el valor de la serie es siempre el valor original más la magnitud del atípico

(2) Si la serie es estacionaria y sigue un modelo MA(q) los q valores siguientes de la serie estarán afectados por una magnitud que depende de los coeficientes de la media móvil. En particular, si uno de estos coeficietnes es mayor que la unidad, lo que es posible si $q > 1$, el valor más afectado de la serie ocurrirá en ese punto.

(3) Si la serie es estacionaria y tiene componente AR, todos los valores posteriores al instante en que ocurre el IO estarán afectados, aunque con peso globalmente decrecientes, ya que los coeficientes ψ_j van a cero en una serie estacionaria al aumentar el retardo. Sin embargo, como ψ_1 puede ser mayor que uno, es posible, como en el caso del MA, que el valor más afectado de la serie ocurra después del instante de aparición del atípico.

(4) Cuando la serie es no estacionaria el efecto de un IO puede ser muy complejo. Supongamos que la serie es I(1), entonces el modelo es

$$\phi(B)\nabla z_t = \theta(B)(\omega_I I_t^{(h)} + a_t)$$

que puede escribirse, teniendo en cuanta que $\nabla^{-1} I_t^{(h)} = S_t^{(h)}$, como

$$z_t = \phi^{-1}(B)\theta(B)\omega_I S_t^{(h)} + y_t$$

donde y_t sigue el proceso ARIMA. Este efecto supone un cambio de nivel complejo sobre la serie, como veremos en la sección siguiente. En particular si el modelo es un paseo aleatorio tendremos un cambio de nivel puro. Este resultado se obtiene también con (13.7) ya que entonces, $\psi_j = 1$ para $j \geq 1$ y todos los valores posteriores a h están afectados en la misma magnitud, con lo que los valores posteriores al IO sufren un escalón que les sitúa a un nivel ω_I unidades más altas que los anteriores.

(6) Si la serie es I(2), por ejemplo del tipo $\nabla^2 z_t = a_t$, igualando coeficientes en $\psi(B)\nabla^2 = 1$ obtenemos que los coeficientes ψ_j son $\psi_1 = 2$, $\psi_2 = 3, \psi_j = j+1$ y la serie z_t estará afectada por lo que se denomina un efecto rampa, que estudiaremos en las secciones siguientes. La serie sufrirá un crecimiento lineal determinista con pendiente ω_I desde el instante en que aparece el atípico. Si además la serie es estacional, como por ejemplo la serie $\nabla\nabla_{12} z_t = a_t$ el efecto de un atípico IO es un conjunto de cambios de nivel de la misma magnitud.

Concluimos que los atípicos innovativos pueden producir estructuras muy complejas sobre la serie y con series no estacionarias pueden confundirse facilmente con otros efectos: cambio de nivel, para series I(1) y cambio de tendencia, para series I(2).

Ejemplo 13.2

Obtener el efecto teórico de un IO sobre las observaciones de una serie que sigue el modelo de los pasajeros de avión, $\nabla\nabla_{12} z_t = (1-\theta B)(1-\Theta B^{12})a_t$.

Para ver el efecto sobre la serie necesitamos la representación de media móvil, que se obtiene igualando potencias de B en

$$\psi(B)\nabla\nabla_{12} = (1-\theta B)(1-\Theta B^{12})$$

Escribiendo en esta ecuación las inversas de los operadores diferencia como operadores de suma indefinida,:

$$\psi(B) = (1 + B + B^2 + ...)(1 + B^{12} + B^{24} + ...)(1-\theta B)(1-\Theta B^{12})$$

lo que resulta, multiplicando la parte regular y la estacional, en

$$\psi(B) = (1 + (1-\theta)B + (1-\theta)B^2 + ...)(1 + (1-\Theta)B^{12} + (1-\Theta)B^{24} + ...)$$

y operando

$$\psi(B) = 1 + (1-\theta)B + (1-\theta)B^2 + ...((1-\Theta) + (1-\theta))B^{12} + ((1-\Theta)(1-\theta) + (1-\theta))B^{13} + ...$$

El efecto sobre la serie de un IO de tamaño 1 se representa en la figura 13.3. En el instante en que ocurre el efecto es igual al tamaño del IO, pero en los periodos siguientes el efecto cae a $(1-\theta)$ y se convierte en un cambio de nivel de magnitud $(1-\theta)$. Al llegar al periodo 12 se superpone al escalón un valor de magnitud $(1-\Theta)$ y, a continuación, todas las observaciones de la 13 a la 23 sufren un escalón de magnitud $(1-\Theta)(1-\theta)+(1-\theta)$. En el periodo 24 se superpone un valor de magnitud $(1-\Theta)$ y se repite la estructura como indica la figura **??**.

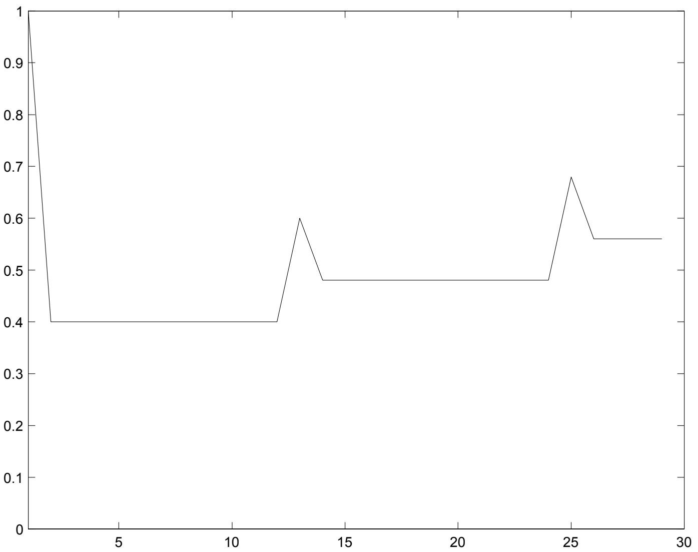

Figura 13.3: Efectos sobre los siguientes 30 valores posteriores a su ocurrencia de un atípico innovativo en una serie que sigue el modelo de pasajeros de avión

Por ejemplo, si Θ=.8 y $\theta = .6$, y el IO ocurre cuando quedan 36 observaciones para el final se la serie el efecto es: 1 en el instante de ocurrencia, después tendrá un escalón de tamaño .4. En la posición 12 tendrá un valor de .6 y entre la 13 y la 23 un escalón de tamaño .4+.2×.4 =.48. Entre la 25 y la 35 un escalón de tamaño .56. Observemos que este efecto tan complejo puede ser aproximado por un AO en la posición 1, seguido de un cambio de nivel de alrededor de .5

13.3.2 Efectos en la estimación

El efecto de un IO en la estimación de los parámetros depende del modelo para la serie. Para series estacionarias los efectos de un IO sobre los parámetros estimados son menores que en el caso de los AO, pero para series no estacionarias pueden ser importantes. Para ilustrar los efectos en el caso estacionario supongamos un modelo AR(1) con media cero. Entonces el estimador del parámetro se calcula con

$$\widehat{\phi} = \frac{\sum z_t z_{t-1}}{\sum z_t^2}$$

y reemplazando para $t = h + j$, z_t por $y_t + \omega_I \psi_j$ este estimador puede escribirse

$$\widehat{\phi} = \frac{\widehat{\phi}_0 + T^{-1}\widetilde{\omega}_I S_{1+} T^{-1}\widetilde{\omega}_I^2(\phi/(1-\phi^2))}{1 + T^{-1}2\widetilde{\omega}_I S_2 + T^{-1}\widetilde{\omega}_I^2/(1-\phi^2)},$$

donde $\widehat{\phi}_0 = \sum y_t y_{t-1} / \sum y_t^2$ es la estimación obtenida de la serie sin atípicos y $S_1 = \sum_{j=0}(\widetilde{y}_{h-1+j} + \widetilde{y}_{h+1+j})\phi^j$ y $S_2 = \sum_{j=0} \widetilde{y}_{h+j}\phi^j$. Esta expresión indica que si $\widetilde{\omega}_I \to \infty$ entonces $\widehat{\phi} \to \phi$, y no existen sesgos asintóticos al estimar el parámetro.

Ejemplo 13.3

La figura 13.4 presenta una serie AR(1) simulada con parámetro .6 contaminada por un IO en la posición 60. Se observa que los valores posteriores al instante 60 de la serie también están afectados, aunque con peso decreciente. La figura 13.5 presenta la función de autocorrelación y se ve que está muy poco afectada por el atípico.

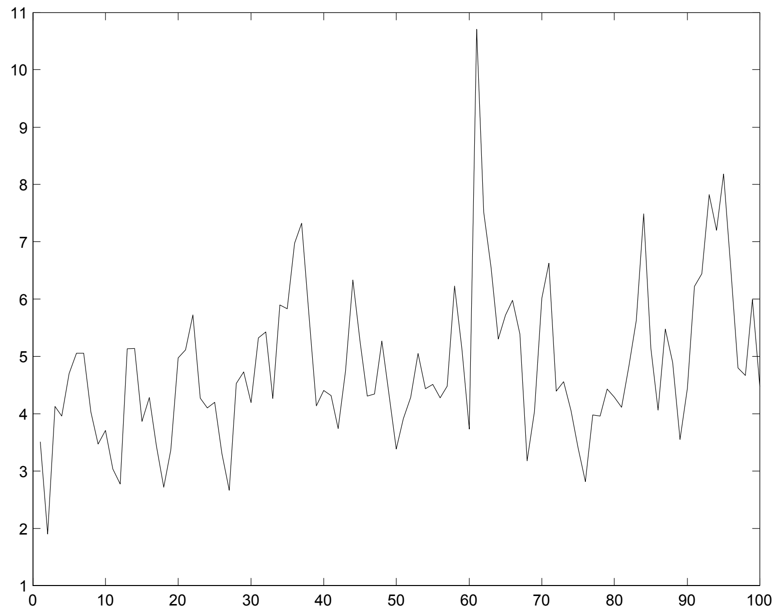

Figura 13.4: Un AR(1) con parámetro .6 y un IO en la posición 60 igual a seis desviaciones típicas de los residuos

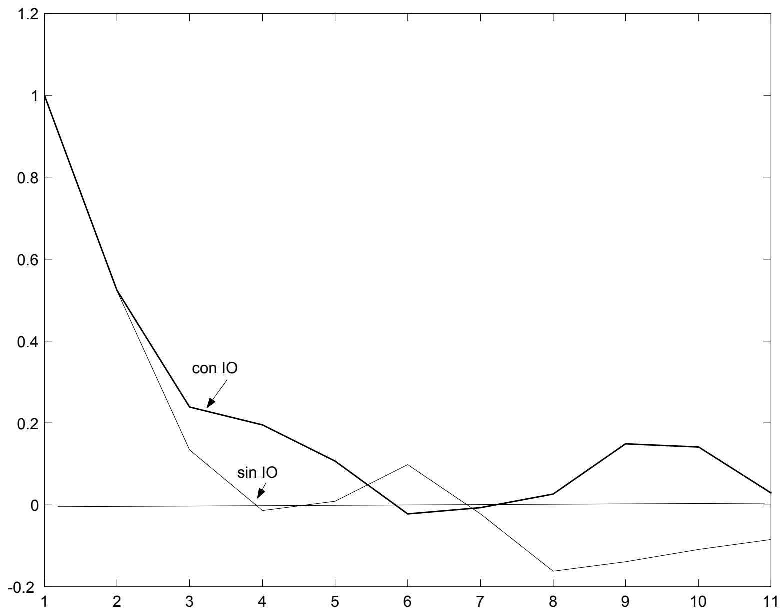

Figura 13.5: Dos realizaciones de una proceso AR(1), una con un IO y otra sin contaminar.

13.4 Cambio de nivel

Un efecto importante que puede aparecer sobre una serie temporal es el cambio de nivel. Diremos que una serie ha sufrido un cambio de nivel en el instante h si sigue el modelo:

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

donde $S_t^{(h)}$ es la variable escalón definida en el capítulo 12 por $S_t^{(h)} = 1$ si $t \geq h$ y cero en otro caso. Los valores de la serie observada estarán relacionados con la serie sin contaminar por el cambio de nivel mediante:

$$z_t = \begin{cases} y_t & t < h \\ y_t + \omega_L & t \geq h. \end{cases}$$

Si el proceso es estacionario un cambio de nivel convertirá a la serie en no estacionaria, ya que la esperanza de cada observación será μ antes del cambio de nivel y $\mu + \omega_L$ después del cambio. Si la serie es no estacionaria el cambio de nivel puede ser equivalente a un IO, como ya hemos visto. En efecto, una serie $I(1)$ con un cambio de nivel sigue el modelo

$$z_t = \omega_L S_t^{(h)} + \nabla^{-1} a_t$$

y tomando diferencias como al diferenciar un escalón se obtiene un impulso, el modelo anterior puede también escribirse como:

$$\nabla z_t = \omega I_t^{(h)} + a_t$$

que es el modelo de un IO para esta serie.

13.4.1 Efectos en los residuos

Escribiendo el cambio de nivel como

$$\pi(B)(z_t - \omega_L S_t^{(h)}) = a_t.$$

la relación entre los residuos calculados con los parámetros verdaderos y las innovaciones es

$$e_t = \omega_L \pi(B) S_t^{(h)} + a_t$$

y estas variables están relacionadas por:

$$e_t = \begin{cases} a_t & t < h \\ a_t + \omega_L l_j & t = h + j. \end{cases} \tag{13.8}$$

donde $l_j = 1 - \pi_1 - ... - \pi_j$ son los coeficientes de $l(B) = (1 + l_1 B + l_2 B^2 + ...) = \pi(B)/(1 - B)$. Concluimos que todos los residuos después del cambio están afectados, pero que el efecto depende: (1) del modelo, y el efecto será mayor para modelos estacionarios que para los no estacionarios; (2) de la distancia entre el momento de aparición, h, y el final.

13.4.2 Efecto en la estimación

Veámos el efecto de un cambio de nivel sobre un proceso estacionario. Supongamos que el proceso es un AR(1). Puede comprobarse que la estimación del parámetro puede escribirse como:

$$\widehat{\phi} = \frac{\widehat{\phi}_0 + T^{-1} S_3 + T^{-1} \widetilde{\omega}_L^2 (T - h)}{1 + T^{-1} 2 S_4 + T^{-1} \widetilde{\omega}_L^2 (T - h + 1)},$$

donde $\widetilde{\omega}_L = \omega_L / s_y$ y $\widetilde{y}_t = y_t / s_y$, $\widehat{\phi}_0 = \sum y_t y_{t-1} / \sum y_t^2$ es la estimación obtenida de la serie sin atípicos y $S_3 = \widetilde{\omega}_L (\widetilde{y}_{h-1} + \widetilde{y}_h + 2\widetilde{y}_{h+1} + ... + 2\widetilde{y}_T)$ y $S_4 = \widetilde{\omega}_L \sum_{j=0} \widetilde{y}_{h+j}$. Si $T - h$ no es muy pequeño,

$$\text{si} \quad \widetilde{\omega}_L \to \infty, \Longrightarrow \widehat{\phi} \to 1.$$

y el efecto del cambio de nivel es hacer que el parámetro tienda a uno y el proceso parecerá no estacionario. Cuando el proceso es no estacionario el efecto de un cambio de nivel sobre los parámetros puede ser más pequeño.

13.5 Cambios transitorios y efecto rampa.

Un efecto adicional que puede considerarse es el cambio transitorio (TC), definido por

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

vemos que si $\delta = 1$ el modelo es el del cambio de nivel, ya que $\nabla^{-1} I_t^{(h)} = S_t^{(h)}$, mientras que si $\delta = 0$ tenemos un AO. En la práctica este tipo de atípico se utiliza fijando el valor de δ de manera que no pueda confundirse con el IO o LS.

Otro tipo de efecto de interés es el efecto rampa, que se define por

$$z_t = \omega_R R_t^{(h)} + \psi(B)a_t,$$

donde la variable $R_t^{(h)}$ es una variable rampa definida por

$$R_t^{(h)} = \begin{cases} 0 & t < h \\ t+1-h & t \geq h. \end{cases}$$

y que introduce una tendencia determinista con pendiente ω_R en la serie a partir del instante h. La variable rampa, $R_t^{(h)}$ esta relacionda con las variables escalón e implulso, ya que $\nabla R_t^{(h)} = S_t^{(h)}$ y $\nabla^2 R_t^{(h)} = I_t^{(h)}$. Este efecto sólo es esperable en series no estacionarias y puede confundirse fácilmente con un IO en series con dos diferencias. En efecto, suponemos que se identifica sobre la serie estacionaria $\nabla^2 z_t$ un IO según el modelo:

$$\nabla^2 z_t = \omega I_t^{(h)} + a_t$$

Este modelo puede escribirse en la forma equivalente

$$\nabla z_t = \omega S_t^{(h)} + \frac{1}{\nabla} a_t$$

que indica que la serie ∇z_t tiene un cambio de nivel y un ruido no estacionario, o también

$$z_t = \omega R_t^{(h)} + \frac{a_t}{\nabla^2}$$

que indica que la serie sufre un efecto rampa. Los efectos rampa aparecen en serie I(2) confundidos con los IO. Muchas series mensuales requieren la diferenciación $\nabla\nabla_{12}$ para convertirse en estacionarias y si existe un efecto rampa

$$z_t = \omega_R R_t^{(h)} + \frac{\psi(B)}{\nabla\nabla_{12}} a_t, \tag{13.9}$$

donde el operador $\psi(B)$ recoge todos los operadores estacionarios regulares y estacionales. Como$\nabla\nabla_{12} = \nabla^2 S(B)$, donde $S(B) = 1 + B + ... + B^{11}$ es el operador suma y $\nabla^2 R_t^{(h)} = I_t^{(h)}$, podemos escribir

$$\nabla\nabla_{12} z_t = \omega_R S(B) I_t^{(h)} + \psi(B)a_t$$

y una rampa producirá sobre la serie estacionaria una racha de doce atípicos consecutivos de la misma magnitud. Vamos a comparar esta rampa pura con un IO sobre las serie. Entonces la serie estacionaria sigue el modelo

$$\nabla\nabla_{12} z_t = \psi(B)\omega_R I_t^{(h)} + \psi(B)a_t \tag{13.10}$$

y sobre la serie estacionaria veremos una secuencia de atípicos cuya posición y tamaño dependen de los coeficientes $\psi(B)$. Como estos efectos pueden confundirse, cuando aparezca un IO sobre una serie I(2) conviene siempre estimar el modelo (13.9) y compararlo con el (13.10) para ver si se trata realmente de un IO o de un efecto rampa puro.

13.6 Procedimientos de estimación de atípicos

En la práctica la posición y la naturaleza de los atípicos que pueden aparecer sobre una serie es desconocida y se precisa un procedimiento para identificarlos y detectar sus efectos. En concreto se requiere, para cada dato atípico:

(1) Detectar el momento de aparición.

(2) Identificar el tipo de atípico.

(3) Estimar su tamaño.

Una vez conocidos estos aspectos es posible eliminar su efecto sobre la serie.

El procedimiento que comentamos a continuación está diseñado para encontrar atípicos del tipo aditivo, AO, innovativo, IO, cambio transitorio, TC y cambio de nivel, LS, y fué desarrollado por Chang y Tiao (1983), Tsay (1986), Chang, Tiao and Chen (1988) y Chen y Liu (1993). Este método está implantado en los programas SCA y TSW, aunque este último no incluye los IO. En los cambios transitorios se supone que el denominador es fijo, igual a .7, con lo que el modelo para este efecto es $\omega/(1-.7B)$, y en todos los cuatro casos el atípico queda definido por un único parámetros ω. Vamos a comentar como realizar los tres pasos anteriores. En primer lugar estudiaremos como estimar el tamaño de un atípico cuando su posición y tipo es conocida. A continuación nos basaremos en este resultado para presentar el procedimiento general.

13.6.1 Estimación del tamaño del atípico

Vamos a analizar cómo estimar el tamaño del atípico cuando suponemos conocido los parámetros del proceso, el tipo de atípico y el momento de ocurrencia. Utilizando la relación entre los residuos estimados cuando existe un atípico y las innovaciones verdaderas de un proceso, tenemos que:

1. Si la serie ha sufrido un AO, la relación entre los residuos y las innovaciones que vimos en (13.3) es:

$$e_t = \omega_i x_t + a_t, \tag{13.11}$$

donde $\omega_i = \omega_A$ y x_t es la variable :

$$x_t = x_t^{AO} = \pi(B) I_t^{(h)} = (0, 0, ...1, -\pi_1, -\pi_2, ..., -\pi_{T-h}). \tag{13.12}$$

que depende de los parámetros de la representación AR(∞) del proceso. La ecuación (13.11) indica que, como vimos en la sección 13.1, los residuos a partir del instante h pueden contener información sobre el tamaño del atípico y esto ocurre cuando la variable x_t es distinta de cero. Por ejemplo, si la serie es una AR(1), los residuos h y $h+1$ contienen información sobre el AO y esto se traduce en que la variable x_t toma sólo dos valores distintos de cero, que son $x_h = 1$ y $x_{h+1} = -\phi$, los tamaños del efecto del AO sobre los residuos.

2.Si la serie ha sufrido un IO, según la ecuación (13.5) sólo un residuo esta efectado, la variable $x_t^{IO} = I_t^{(h)} = (0, 0, ...1, 0, 0, ..., 0)$ sólo toma un valor distinto de cero y la ecuación (13.11) sigue siendo cierta, pero tomando ahora $\omega_i = \omega_I$.

3. Para un cambio de nivel todos los residuos posteriores están afectados, de acuerdo (13.8). En la ecuación(13.11) ahora $\omega_i = \omega_L$ y

$$x_t^{LS} = \pi(B) S_t^{(h)} = (0, 0, ...1, 1-\pi_1, 1-\pi_1-\pi_2, ..., 1-\pi_1 - ... - \pi_{T-h}).$$

Es decir, los coeficientes no nulos de la variable x_t son ahora $l_i = 1 - \pi_1 - ... - \pi_i$ y podemos escribir $x_t^{LS} = (0, 0, ...1, l_1, l_2, ..., l_{T-h})$, donde $l_i = 1 - \pi_1 - ... - \pi_i$.

4. Para el cambio transitorio, según (**??**), la variable $x_t^{TC} = (0, 0, ...1, g_1, g_2, ..., g_{T-h})$ donde los g_i son los coeficientes de $(1-.7B)^{-1}\pi(B)$.

En resumen, en los cuatro casos la ecuación (13.11) permite estimar el tamaño del atípico conocido su clase y momento de ocurrencia. Esta es una ecuación de regresión simple, y la estimación del parámetro desconocido, que es el tamaño del atípico, se realiza de la forma habitual por mínimos cuadrados:

$$\tilde{\omega}_i = \frac{\sum e_t x_t}{\sum x_t^2} \tag{13.13}$$

y la varianza de esta estimación es

$$var(\tilde{\omega}_i) = \frac{\sigma_a^2}{\sum x_t^2} \tag{13.14}$$

El resultado de aplicar esta ecuación a los distintos tipos de atípicos es la siguiente:

(1) Para el AO la estimación es

$$\tilde{\omega}_A = (e_h - \pi_1 e_{h+1} - .... - \pi_{T-h} e_T)/\rho_A^2 \tag{13.15}$$

donde $\rho_A^2 = (1 + \pi_1^2 + \cdots + \pi_{T-h}^2)$. El estimador es una combinación lineal de los residuos poseriores a la ocurrencia del atípico. Esta ecuación tiene la siguiente interpretación intuitiva. Cada residuo después de h contiene información sobre el efecto del atípico ya que por (13.11) y (13.15)

$$e_{h+j} = -\omega_A \pi_j + a_{h+j}$$

con $\pi_0 = -1$. De esta ecuación obtenemos que el estimador para el parámetro ω_A que se obtiene con el residuo e_{h+j}, que llamaremos $\widehat{\omega}_A^{(j)}$ se obtiene con $\widehat{\omega}_A^{(j)} = -e_{h+j}/\pi_j$. Estos estimadores son insesgados e independientes, ya que sus errores dependen de las variables a_{h+j} que son independientes. Cada uno de ellos tiene varianza $\sigma^2/\ \pi_j^2$, y el estimador global será una media ponderada de todos ellos con ponderaciones iguales a las precisiones (inversas de las varianzas) de los estimadores, π_j^2/σ^2. Podemos escribir:

$$\widehat{\omega}_A = \sum_{j=1}^{T-h} \frac{\pi_j^2/\sigma^2}{\sum \pi_j^2/\sigma^2} \widehat{\omega}_A^{(j)}$$

que es equivalente a (13.15). La varianza es $Var(\widehat{\omega}_A) = \sigma_a^2/\rho_A^2$

(2) Para IO el estimador es simplemente el valor del residuo en ese punto:

$$\widehat{\omega}_I = e_h.$$

esta estimación tendra varianza $Var(\widehat{\omega}_I) = \sigma_a^2$.

(3) Para el LS, como $x_t^{LS} = (0, 0, ...1, l_1, l_2, ..., l_{T-h})$, tenemos que el estimador es:

$$\widehat{\omega}_L = (e_h + l_1 e_{h+1} + ... + l_{T-h} e_T)/\rho_L^2 \tag{13.16}$$

donde $\rho_L^2 = (1 + l_1^2 + l_2^2 + \cdots + l_{T-h}^2)$. El estimador se interpreta de forma análoga al caso del AO. Un estimador insesgado del tamaño utilizando la información de e_{h+j} es e_{h+j}/l_j y tiene varianza σ^2/l_j^2. El estimador (13.16) es una combinación lineal de estos estimadores con pesos iguales a su precisión relativa. Por otro lado, la interpretación de esta estimación es muy simple: compara el nivel de la serie antes y después del cambio.

En el apéndice se demuestra que el estimador del salto tiene un sesgo que puede ser importante cuando no conocemos los parámetros y los estimamos con los datos. Se demuestra también que cuando el modelo es ARIMA(0,1,1) es la diferencia entre dos medias ponderadas de las observaciones antes y después del cambio.

(4) Para el TC el resultado tiene una interpretación similar.

13.6.2 El procedimiento general

Dada una serie temporal que sigue un modelo estimado, $\widehat{\pi}(B) z_t = \widehat{e}_t$, la deteccion de los posibles valores atípicos se realiza detectando el instante y tipo más probable y contrastando si es significativo. Si lo es, se limpia de la serie y se repite el proceso hasta que el más probable no sea significativo, con lo que el procedimiento termina. Para implantarlo se procede como sigue:

1. Se calculan los residuos $\widehat{e}_t = \widehat{\pi}(B) z_t$, utilizando el modelo estimado para la serie.

2. Se comienza investigando la presencia de cada tipo de atípico considerado. Si comenzamos con el AO, se estima el tamaño de un posible AO, $\widehat{\omega}_t^{A0}$, en todos los instantes, $t = 1, ..., T$ utilizando los residuos $\widehat{e}_t$ y suponiendo que la relación entre ellos y las innovaciones es igual a la teórica presentada en la sección anterior y utilizando los parámetros estimados $\widehat{\pi}$ como si fuesen los verdaderos en (13.13). Se obtiene el

error típico de estimación, $s(\widehat{\omega}_t^{A0})$, con (13.14), donde se sustituyen los parámetros estimados por lo teóricos y como estimación de σ se toma un estimador robusto, como la mediana de las desviaciones absolutas, la meda, dividida por .675 o la desviación típica recortada, donde se eliminan del cálculo las $\alpha\%$ observaciones más extremas. Se compara la estimación del tamaño del atípico con la desviación típica de la estimación, formando el estadístico:

$$\lambda_t^{AO} = \frac{\widehat{\omega}_t^{A0}}{s(\widehat{\omega}_t^{A0})}$$

A continuación se toma el máximo de este estadístico en todos los puntos muestrales. Sea $\lambda_{AO} = \max(\lambda_t^{AO})$ este valor máximo y supongamos que corresponde al instante $t = t_{AO}$.

3. Se repite el proceso anterior para los otros tipos de atípicos. Por ejemplo, suponiendo un IO, por (13.6) la estimación del atípico será directamente el residuo y su desviación típica es σ_a, que se sustituye por un estimador robusto, $\widehat{\sigma}_a$, como comentamos en el paso anterior. El estadístico anterior se reduce a

$$\lambda_t^{IO} = \frac{\widehat{\omega}_t^{I0}}{s(\widehat{\omega}_t^{I0})} = \frac{\widehat{e}_t}{\widehat{\sigma}_a}.$$

Para los otros dos tipos de atípicos considerados, cambio de nivel y cambio transitorio y se construyen los mismos estadísticos utilizando lo mismos principios (parámetros estimados por verdaderos y estimación robusta de σ_a):

$$\lambda_t^{LS} = \frac{\widehat{\omega}_t^{LS}}{s(\widehat{\omega}_t^{LS})}, \qquad \lambda_t^{CT} = \frac{\widehat{\omega}_t^{CT}}{s(\widehat{\omega}_t^{CT})}$$

y de nuevo se toma el valor máximo de estos estadísticos en la muestra, que llamaremos $\lambda_{IO} = \max(\lambda_t^{IO})$, $\lambda_{LS} = \max(\lambda_t^{LS})$ y $\lambda_{TC} = \max(\lambda_t^{LS})$. Llamaremos t_{IO}, t_{LS} y t_{TC} a los instantes en que ocurren estos máximos.

4. Se considera el máximo de estos cuatro estadísticos:

$$\lambda_{\max} = max(\lambda_{AO}, \lambda_{IO}, \lambda_{LS}, \lambda_{TC}).$$

Si $\lambda_{\max} > c_1$, donde c_1 es un valor crítico que se obtiene por simulación, concluimos que existe en la serie un cambio de nivel del tipo y en el instante que corresponde a $\lambda_{\max}$. En otro caso, suponemos que no hay atípicos en la serie. El valor c_1 suele tomarse como 3.25 o 3.5

5. Si se detecta un atípico en la etapa anterior se corrige en la serie, si es AO, LS, o TC, o en los residuos si es IO. La corrección de la serie por un AO detectado en h es $z_h^c = z_h - \widehat{\omega}_{AO}$, por un LS a partir de h, $z_{h+j}^c = z_{h+j} - \widehat{\omega}_{LS}$ para todas las observaciones posteriores a h, por un TC en h se aplican correcciones $z_{h+j}^c = z_{h+j} - \widehat{\omega}_{TC}7^j$ .Con esta serie corregida se vuelve al paso 1. Para un IO se corrige simplemente el residuo en ese punto como $\widehat{e}_h^c = \widehat{e}_h - \widehat{\omega}_{IO}$.y se vuelve al paso 2. El proceso se itera hasta que no se detecten mas atípicos.

5. Cuando el proceso de detección ha terminado se estiman conjuntamente los efectos de los atípicos y los parámetros del modelo. Supongamos que se han detectado atípicos en los instantes $t_1, t_2, \ldots, t_k$. El modelo a estimar es

$$z_t = \sum_{j=1}^{k} \omega_j v_{t_i}(B) I_t^{(t_i)} + \psi(B) a_t,$$

donde $v_{t_i}(B) = 1$ si AO, $v_{h_j}(B) = \psi(B)$ si IO, $v_{h_j}(B) = \nabla^{-1}$ si LS y $v_{h_j}(B) = (1 - .7)^{-1}$ si es TC.

Ejemplo 13.4

Como ejemplo de este procedimiento vamos a aplicarselo a la series de cuota de mercado de Crest y Colgate. Los modelos estimados inicialmente para estas series son:

Tabla 13.1 Modelos iniciales para Crest y Colgate

	d	MA	st	t	$\widehat{\sigma}$
Crest	1	.7661	.0392	19.54	0.047 .
Colgate	1	.6648	0448	14.85	0.0455

Con estos parámetros se ha hecho la búsqueda de atípicos con los resultados indicados en la tabla con el programa SCA

Tabla 13.2 Atipicos detectados para Crest y Colgate

Crest						Colgate			
t	99	136	167	196	213	43	102	136	196
tamaño	-0.102	0.167	-0.142	-0.131	0.121	-0.130	-0.161	-0.100	0.142
est-t	-3.07	7.04	-3.69	-3.97	3.14	-3.77	-3.82	-4.49	4.10
tipo	TC	LS	AO	TC	AO	TC	AO	LS	TC

Se observa que ambas series tienen algunos atípicos comunes. En el instante 136 ambas series sufren un cambio de nivel de dirección opuesta. Una explicación de este hecho es que en Agosto de 1960, que corresponde a la observación 136, la American Dental Association declaró que la inclusión de fluor en la pasta dentrífica hecha por Crest se había demostrado eficaz en la prevención de las caries dentales. Según la estimación del cambio nivel, esto supuso un aumento de la cuota de mercado de Crest de 16.7 puntos porcentuales y un decrecimiento de Colgate de 10.0 puntos. Se observa que en el instante 196 se produjo un movimiento de sentido opuesto, aunque de caracter transitorio, correpondiente a un esfuerzo comercial de Colgate de incluir también fluor y bajar sus precios.

La tabla presenta los modelos estimados cuando se han tenido en cuenta estos valores atípicos. Se observa que ha aumentado el coeficiente de la media móvil como corresponde a un modelo con tendencia más regular.

Tabla 13.3 Modelos finales para Crest y Colgate

	d	MA	st	t	$\widehat{\sigma}$
Crest	1	..8104	.0375	21.61	0.040 .
Colgate	1	.8607	.0326	26.41	0.0435

13.7 Comentarios al procedimiento de detección de atípicos

Un aspecto a tener en cuenta del procedimiento anterior es que, como hemos comentado, si las series son no estacionarias los IO pueden confundirse facilmente con otros efectos. Por ejemplo, con cambios de nivel en series I(1) y con efectos rampa, en series I(2). Conviene siempre antes de aceptar un IO en una serie integrada asegurarse que no esta encubriendo otro tipo de efecto de más fácil interpretación.

El método presentado funciona bastante bien para detectar AO y IO en series con bajo nivel de contaminación, pero tiene dos problemas principales en series con muchos atípicos (véase Sánchez y Peña, 2003). El primero es que no identifica bien cambios de nivel que pueden confundirse con otros atípicos o pasar inadvertidos. El segundo es que la estimación inicial de los parámetros puede ser muy mala, con lo que el método puede fallar.

Para ilustrar estos problemas consideremos la figura 13.6. Esta serie es de datos incorrelados y sufre un cambio de nivel de 20 unidades en la posición 30 y la media de la serie pasa de ser 10 a ser 30. Al aplicar el método estudiado el primer paso es calcular los residuos de la serie. La media de todo el periodo será 20 y por tanto la estimación de la magnitud del cambio que obtenemos de los residuos es, como se demuestra en el apéndice, la mitad del cambio real. Esto hace más dificil la identificación del cambio de nivel. Sin embargo, si detectamos su posición la estimación del efecto si me proprocionará la magnitud del salto. Otro problema con los LS es que podemos fácilmente confundirlos con IO en series estacionarias. En un amplio estudio de simulación, Sánchez y Peña (2003) encontraron que el percentil 95% del estadístico para LS cuando no hay atípicos está entre 2.5 y 3 para procesos estacionarios, pero está entre 3.5 y 3.8 para procesos no estacionarios. Sin embargo, el percentil 95% para los efectos transitorios AO, IO y TC es similar, próximo a 3.5. Este resultado indica que no adecuado comparar el máximo del estadístico del contraste para LS con el maximo de los otros efectos, porque la distribución es distinta. La razón es fácil de entender: como se demuestra en el apéndice, el estadístico para LS en series estacionarias es la diferencia entre dos medias de los datos. Esto implica que en el centro de la serie el estadístico para el instante h y para el instante $h+1$ estarán muy correlados, ya que las medias que se comparán son sustancialmente idénticas, al diferir sólo en una observación. El máximo del estadístico es el máximo de variables muy correladas, que siempre será menor que el máximo de variables independientes. Sin embargo, para un IO la razón de verosimilitudes en h depende sólo del residuo en ese punto y estará incorrelada con la razón en h+1 que depende del residuo en ese punto. Por tanto, el estadístico para IO será mayor que el estadístico para LS y si existe un LS, es frecuente que el estadístico para IO en ese punto aparezca como mayor que el estadístico para LS, ya que no tienen la misma distribución. Por tanto podemos identificar erróneamente un IO en lugar de un LS.

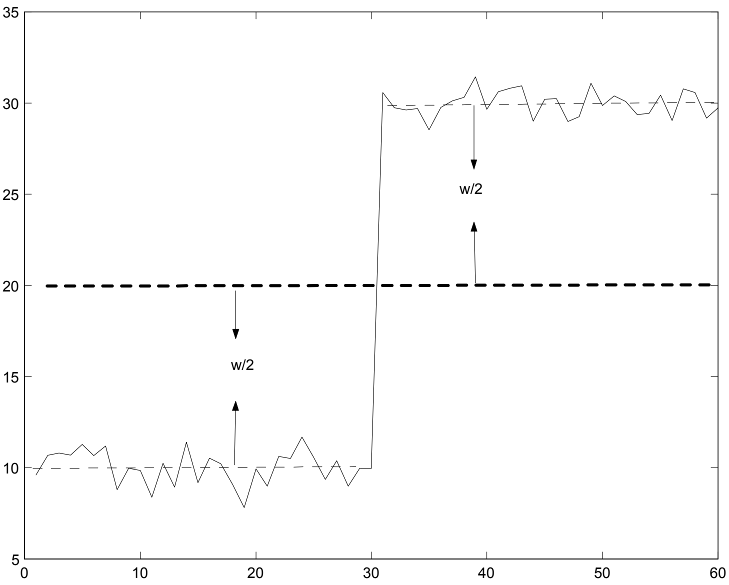

Figura 13.6: Una serie estacionaria con un cambio de nivel en la mitad del periodo de observación

Una alternativa mejor es separar los cambios de nivel e identificarlos separadamente. Además existe considerable evidencia, véase Carnero et al (2003) y Galeano y Peña (2003), que métodos básados en sumas de residuos acumuladas son más efectivos que el contraste de verosimilitudes para identificar cambios de nivel.

Para resolver el problema de la mala estimación inicial de los parámetros podemos identificar las observaciones que afectan más a la estimación de los parámetros, eliminarlas suponiendo que son valores ausentes, y estimar los parámetros inciales en esta serie limpia de observaciones influyentes. Véase Peña(1990, 1991) y Sánchez y Peña (2003) para mas detalles de este enfoque.

Ejercicios

Apéndice 13 Estimación de los cambios de nivel

Supongamos para simplificar que z_t es ruido blanco, de manera que $\pi(B) = 1$ y $l(B) = (1 - B)^{-1}$ con $l_i = 1$. Entonces el estimador del cambio de nivel es:

$$\widehat{\omega}_{LS} = \frac{(e_h + e_{h+1} + ... + e_T)}{(T - h + 1)} = \frac{\sum_{t=h}^{T}(z_t - \overline{z})}{(T - h + 1)} = \overline{z}_2 - \overline{z}$$

y sustituyendo en esta ecuación la expresión de la media

$$\overline{z} = \frac{(h-1)\overline{z}_1 + (T - h + 1)\overline{z}_2}{T}$$

resulta finalmente:

$$\widehat{\omega}_{LS} = \frac{h-1}{T}(\overline{z}_2 - \overline{z}_1)$$

Este estimador tiene sesgo como consecuencia de no conocer los parámetros. Por ejemplo, supongamos que la serie tiene un cambio de nivel de magnitud ω_{LS}. Entonces $(\overline{z}_2 - \overline{z}_1)$ si estima en promedio la magnitud del cambio pero si, por ejemplo, $h = T/2$ el valor estimado por $\widehat{\omega}_{LS}$ será la mitad del cambio real.

Sin embargo, cuando identifiquemos el momento de cambio y estimemos conjuntamente su tamaño y os parámetros del modelo el problema desaparece. El modelo para estimar el cambio de nivel es:

$$z_t = \mu + \omega_{LS} S_t^{(h)} + a_t, \tag{13.17}$$

La estimación de este modelo de regresión por mínimos cuadrados conduce a un valor de la constante de

$$\widehat{\mu} = \overline{z} - \widehat{\omega}_{LS}\overline{S}^{(h)}$$

donde $\overline{z} = \sum_{t=1}^{T} z_t/T$ y $\overline{S}^{(h)} = \sum_{t=1}^{T} S_t^{(h)}/T = (T-h+1)/T$. La pendiente se estima por:

$$\widehat{\omega}_{LS} = \frac{\sum_{t=1}^{T}(z_t - \overline{z})S_t^{(h)}}{\sum_{t=1}^{T}(S_t^{(h)} - (T-h+1)/T)^2} = \frac{\sum_{t=h}^{T}(z_t - \overline{z})}{(T-h+1)(h-1)/T}$$

que puede escribirse

$$\widehat{\omega}_{LS} = \frac{T}{(h-1)}(\overline{z}_2 - \overline{z})$$

y sustituyendo en esta ecuación la expresión de la media

$$\overline{z} = \frac{(h-1)\overline{z}_1 + (T-h+1)\overline{z}_2}{T}$$

resulta finalmente:

$$\widehat{\omega}_{LS} = \overline{z}_2 - \overline{z}_1$$

y esta ecuación estima bien el cambio de nivel, como diferencia entre las medias antes y después del instante en que suponemos aparece el cambio. La varianza de esta estimación se obtiene es

$$var(\widehat{\omega}_{LS}) = \frac{\sigma_a^2}{\sum(S_t^{(h)} - \overline{S}^{(h)})^2} = \frac{\sigma_a^2}{(T-h+1)(h-1)/T}$$

que coincide con la expresión habitual de la varianza de la diferencia de medias en dos muestras independientes:

$$var(\widehat{\omega}_L) = \sigma_a^2(\frac{1}{(T-h+1)} + \frac{1}{h-1})$$

En esta expresión σ_a^2 se estima mediante la varianza residual de la regresión

$$\widehat{\sigma}_a^2 = (\sum_{t=1}^{h-1}(z_t - \overline{z}_1)^2 + \sum_{t=h}^{T}(z_t - \overline{z}_2)^2)/T$$

que equivale a la de la regresión (13.17).

Estos resultados pueden generalizarse para otros modelos. Supongamos un modelo un poco más complejo, como el

$$\nabla z_t = (1-\theta B)a_t$$

de manera que el cambio de nivel es

$$z_t = c + \omega S_t^{(h)} + \frac{(1-\theta B)}{\nabla}a_t$$

entonces, multiplicando por $\nabla(1-\theta B)^{-1}$, la ecuación a estimar es

$$y_t = \omega x_t + a_t$$

donde $y_t = (1-\theta B)^{-1}\nabla z_t$ and $x_t = (1-\theta B)^{-1}I_t^{(h)} = (0, 0, ..., 1, \theta, \theta^2, ..)$. Suponiendo que el tamaño muestral es grande y que h no esta situada ni al principio ni al final de la serie, tenemos que, como

$$\widehat{\omega} = \frac{\sum y_t x_t}{\sum x_t^2} = \frac{y_T + \theta y_{T+1} + \theta^2 y_{T+2} + ..}{1 + \theta^2 + \theta^4 + ..} = (1-\theta^2)(1-\theta F)^{-1}y_T$$

donde $F = B^{-1}$ es el operador de adelanto y la suma de la progresión geométrica indefinida con razón θ^2 del denominador es $(1-\theta^2)^{-1}$. Entonces

$$\widehat{\omega} = (1-\theta^2)(1-\theta F)^{-1}(1-\theta B)^{-1}\nabla z_t$$

y expresando $(1-\theta F)^{-1} = (1+\theta F+\theta^2 F^2+..)$ y $(1-\theta B)^{-1} = (1+\theta B+\theta^2 B^2+..)$ y operando y agrupando términos se obtiene que

$$\widehat{\omega} = Z(antes) - Z(después)$$

donde $Z(antes)$ es una media ponderada de las observaciones antes del cambio de nivel

$$Z(antes) = (1-\theta)(z_{T-1} + \theta z_{T-2} + ...)$$

y $Z(después)$ es una media ponderada de las observaciones a partir del cambio de nivel

$$Z(después) = (1-\theta)(z_T + \theta z_{T+1} + ...)$$