

Ahora que se tienen bases más sólidas para comprender cómo utilizar el modelo de regresión múltiple para las aplicaciones de corte transversal, podemos estudiar el análisis econométrico de los datos de series de tiempo. Dado que nos basaremos mucho en el método de mínimos cuadrados ordinarios, ya se ha realizado buena parte del trabajo concerniente a su mecánica e inferencia. No obstante, como se señaló en el capítulo 1, los datos de series de tiempo cuentan con ciertas características que no poseen aquellos de corte transversal y exigen atención especial al aplicar MCO.

El capítulo 10 cubre el análisis de regresión básico y presta atención a los problemas exclusivos de los datos de series de tiempo. Se proporciona un conjunto de supuestos de Gauss-Markov y del modelo lineal clásico para las aplicaciones de series de tiempo. También se estudian problemas sobre la forma funcional, las variables binarias, las tendencias y la estacionalidad.

En virtud de que ciertos modelos de series de tiempo violan por fuerza los supuestos de Gauss-Markov, en el capítulo 11 se describen la naturaleza de estas violaciones y se presentan las propiedades de mínimos cuadrados ordinarios para muestras grandes. Como ya no es posible suponer un muestreo aleatorio, se deben cubrir las condiciones que limitan la correlación temporal en una serie de tiempo para asegurar que el análisis asintótico usual sea válido.

En el capítulo 12 se plantea un problema nuevo e importante: la correlación serial de los términos de error en las regresiones con series de tiempo. Se analizarán las consecuencias, las formas de hacer pruebas y los métodos para tratar con la correlación serial. El capítulo 12 también contiene una explicación de cómo surge la heterocedasticidad en los modelos de series de tiempo.

10.1 *Naturaleza de los datos de series de tiempo*

Una característica obvia de los datos de series de tiempo que los distingue de aquellos de corte transversal es que **tienen un orden temporal**. Por ejemplo, en el capítulo 1 se analizó de forma breve una base de datos de series de tiempo sobre el empleo, el salario mínimo y otras variables económicas de Puerto Rico. En dicha base de datos se debe saber que los datos de 1970 preceden de inmediato a los de 1971. Para analizar los datos de series de tiempo en las ciencias sociales, es necesario reconocer que el pasado influye en el futuro, pero no a la inversa (a diferencia del universo de “Viaje a las estrellas”). Para destacar el orden adecuado de los datos de series de tiempo, en la tabla 10.1 aparece una lista parcial de datos sobre la inflación y la tasa de desempleo en Estados Unidos tomados de varias ediciones del informe *Economic Report of the President*, incluido el informe de 2004 (tablas B-42 y B-64).

Existe otra diferencia más sutil entre los datos de corte transversal y los datos de series de tiempo. En los capítulos 3 y 4 se estudiaron las propiedades estadísticas de los esti-

¹ Tomado en forma parcial de Jeffrey M. Wooldridge - *Introducción a la econometría. Un enfoque moderno*. 4a. edición, 2009, Capítulo 10. Se ha simplificado la exposición, y se han dejado de lado los tecnicismos, pero se mantuvo la numeración de secciones, tablas y ecuaciones y gráficos. También hemos saltado otras cuestiones que ya están incluidas en otras lecturas.

madores de MCO con base en la noción de que las muestras se extrajeron de forma aleatoria de la población adecuada. Entender por qué los datos de corte transversal deben verse como resultados aleatorios es muy sencillo: cada muestra diferente extraída de la población, por lo general, producirá valores distintos de las variables dependientes e independientes (como la educación, la experiencia, el salario, etc.). Por tanto, las estimaciones por MCO calculadas a partir de diferentes muestras aleatorias por lo general serán diferentes, razón por la cual se considera que los estimadores de MCO son variables aleatorias. ¿Cómo se debe considerar la aleatoriedad en los datos de series de tiempo? Desde luego, las series de tiempo económicas satisfacen los requisitos intuitivos de ser resultados de variables aleatorias. Por ejemplo, hoy no se sabe cómo cerrará el promedio industrial Dow Jones al final de las operaciones de mañana. Ni tampoco cuál será el crecimiento anual de la producción en Canadá el año próximo. **Dado que los resultados de estas variables no se conocen por anticipado, deben verse sin lugar a dudas como variables aleatorias.**

| Año | Inflación | Desempleo |
|------|-----------|-----------|
| 1948 | 8.1 | 3.8 |
| 1949 | -1.2 | 5.9 |
| 1950 | 1.3 | 5.3 |
| 1951 | 7.9 | 3.3 |
| ⋮ | ⋮ | ⋮ |
| 1998 | 1.6 | 4.5 |
| 1999 | 2.2 | 4.2 |
| 2000 | 3.4 | 4.0 |
| 2001 | 2.8 | 4.7 |
| 2002 | 1.6 | 5.8 |
| 2003 | 2.3 | 6.0 |

Tabla 10.1

En términos formales, a una secuencia de variables aleatorias indexadas en el tiempo se le llama **proceso estocástico** o **proceso de series de tiempo** (“estocástico” es sinónimo de aleatorio). Cuando se conforma una base de datos de series de tiempo, se obtiene un

resultado posible, o realización, del proceso estocástico. Únicamente se puede ver una sola realización, ya que no es posible retroceder en el tiempo y empezar de nuevo el proceso. (Esto es análogo al análisis de corte transversal en el que únicamente se puede reunir una sola muestra aleatoria.) No obstante, si ciertas condiciones históricas fueran distintas, por lo general se obtendría una realización diferente para el proceso estocástico y es por ello que los datos de series de tiempo se consideran como el resultado de variables aleatorias. El conjunto de todas las realizaciones posibles de un proceso de series de tiempo desempeña el papel de la población en el análisis de corte transversal. El tamaño de



Una realización simulada por computadora de un proceso de movimiento de Wiener o Browniano en la superficie de una esfera. El proceso de Wiener es considerado el proceso estocástico más estudiado y central en la teoría de la probabilidad. (Wikipedia)

muestra para una base de datos de series de tiempo es el número de periodos durante los cuales se observan las variables de interés.

10.2 Ejemplos de modelos de regresión con series de tiempo

En esta sección se estudian dos ejemplos de modelos de series de tiempo que han sido útiles en el análisis empírico y que se estiman con facilidad mediante mínimos cuadrados ordinarios. En el capítulo 11 se verán otros modelos.

En esta sección se estudian dos ejemplos de modelos de series de tiempo que han sido útiles en el análisis empírico y que se estiman con facilidad mediante mínimos cuadrados ordinarios. En el capítulo 11 se verán otros modelos.

Modelos estáticos

Imaginen que tienen a su disposición datos de series de tiempo sobre dos variables, y y z , en las cuales y_t y z_t son contemporáneas. Un modelo estático que relaciona a y con z es

$$[10.1] \quad y_t = \beta_0 + \beta_1 z_t + u_t, \quad t = 1, 2, \dots, n.$$

El nombre de “modelo estático” proviene del hecho de que se está representando una relación contemporánea entre y y z . Por lo común, un modelo de esta índole se postula cuando se considera que un cambio en z en el momento t ejerce un efecto inmediato sobre y : $\Delta y_t = \beta_1 \Delta z_t$, cuando $\Delta u_t = 0$. Los modelos de regresión estáticos también se usan cuando se tiene interés en conocer el trade-off entre y y z .

Un ejemplo de modelo estático es la curva estática de Phillips, dada por

$$[10.2] \quad \ln f_t = \beta_0 + \beta_1 \ln em_t + u_t,$$

donde inf_t es la tasa de inflación anual y $unem_t$ es la tasa de desempleo. Esta forma de la curva de Phillips supone una tasa natural de desempleo constante y expectativas inflacionarias constantes, y se emplea para estudiar el trade-off contemporáneo entre la inflación y el desempleo.²

Como es natural, podemos tener diversas variables explicativas en un modelo de regresión estático. Sean $mrdрте_t$ los homicidios por cada 10,000 personas en una determinada ciudad durante el año t , $convрте_t$ la tasa de condena por homicidio, $unem_t$ la tasa de desempleo local y $yngmle_t$ la fracción de la población que se compone de hombres entre 18 y 25 años de edad. De esta manera, un modelo estático de regresión múltiple que explica las tasas de homicidio es

$$[10.3] \quad mrdрте_t = \beta_0 + \beta_1 convрте_t + \beta_2 unem_t + \beta_3 yngmle_t + u_t.$$

Con ayuda de un modelo como éste esperamos estimar, por ejemplo, el efecto *cæteris paribus* de un incremento en la tasa de condenas para una actividad delictiva en particular.

Modelos de rezagos distribuidos finitos

En un modelo de rezagos distribuidos finitos (RDF) se permite que una o más variables influyan sobre y en forma rezagada. Por ejemplo, para las observaciones anuales, consideren el modelo

$$[10.4] \quad gfr_t = \alpha_0 + \delta_0 pe_t + \delta_1 pe_{t-1} + \delta_2 pe_{t-2} + u_t,$$

donde gfr_t es la tasa de fertilidad general (niños nacidos por cada 1,000 mujeres en edad de concebir) y pe_t es el valor en dólares reales de la exención personal de impuestos. La idea es ver si, en conjunto, la decisión de tener hijos se asocia con el valor impositivo de tenerlos. La ecuación (10.4) reconoce que, tanto por motivos biológicos como conductuales, las decisiones de procrear no serían un resultado inmediato de los cambios en la exención personal.

La ecuación (10.4) es un ejemplo del modelo

$$[10.5] \quad y_t = \alpha_0 + \delta_0 z_t + \delta_1 z_{t-1} + \delta_2 z_{t-2} + u_t,$$

el cual es un modelo de RDF de orden dos. Para interpretar los coeficientes en la ecuación (10.5), suponga que z es una constante, igual a c , en todos los periodos antes del periodo t . En t , z aumenta una unidad a $c + 1$, y después vuelve a su nivel anterior en el periodo $t + 1$. (Es decir, el incremento en z es temporal.) De manera más concreta,

$$\dots, z_{t-2} = c, z_{t-1} = c, z_t = c + 1, z_{t+1} = c, z_{t+2} = c, \dots$$

Para concentrarse en el efecto *cæteris paribus* que z tiene sobre y , se fija el término de error de cada periodo en cero. Por consiguiente,

² Ver por ejemplo la sección 11.2 en Laurence Ball y N. Gregory Mankiw, *A Sticky-Price Manifesto*, 1994.

$$\begin{aligned}
 y_{t-1} &= \alpha_0 + \delta_0 c + \delta_1 c + \delta_2 c, \\
 y_t &= \alpha_0 + \delta_0 (c + 1) + \delta_1 c + \delta_2 c, \\
 y_{t+1} &= \alpha_0 + \delta_0 c + \delta_1 (c+1) + \delta_2 c, \\
 y_{t+2} &= \alpha_0 + \delta_0 c + \delta_1 c + \delta_2 (c + 1), \\
 y_{t+3} &= \alpha_0 + \delta_0 c + \delta_1 c + \delta_2 c,
 \end{aligned}$$

y así sucesivamente. De las dos primeras ecuaciones se obtiene, $y_t - y_{t-1} = \delta_0$, que demuestra que δ_0 es el cambio inmediato en y debido al aumento de una unidad en z en el periodo t . Por lo común, a δ_0 se lo conoce como *propensión de impacto* o **multiplicador de impacto**.

Asimismo, $\delta_1 = y_{t+1} - y_{t-1}$ es el cambio en y un periodo después de la modificación temporal, y $\delta_2 = y_{t+2} - y_{t-1}$ es el cambio en y dos periodos después de la modificación. En el momento $t + 3$, y se ha revertido a su nivel inicial: $y_{t+3} = y_{t-1}$. Esto se debe a que se ha supuesto que sólo dos rezagos de z aparecen en la ecuación (10.5). Cuando se traza la gráfica de δ_j como una función de j , se obtiene la distribución de rezagos, la cual resume el efecto dinámico que un aumento temporal en z tiene sobre y . En la figura 10.1 se aprecia una posible distribución de rezagos para el modelo de RDF de orden dos. (Desde luego, nunca se conocerán los parámetros de δ_j ; así que se estimará δ_j y luego se graficará la distribución de rezagos estimada.)

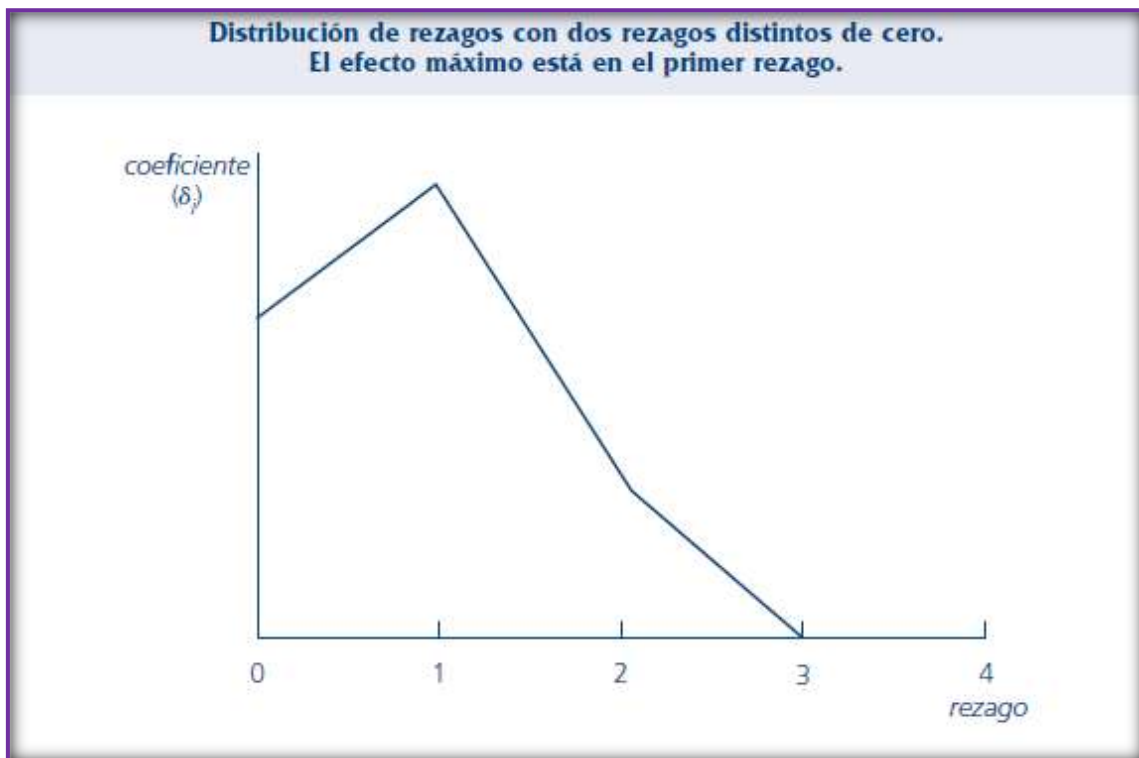


Figura 10.1

La distribución de rezagos de la figura 10.1 implica que el efecto más grande está en el primer rezago. Dicha distribución tiene una interpretación útil. Si se estandariza el valor inicial de y en $y_{t-1} = 0$, la distribución de rezagos determina todos los valores subsiguientes de y debido a un incremento temporal de una unidad en z .

También nos interesa el cambio en y debido a un incremento *permanente* en z . Antes del periodo t , z es igual a la constante c . En t , z aumenta de manera permanente a $c + 1$: $z_s = c$, $s < t$ y $z_s = c + 1$, $s \geq t$. Una vez más, al igualar los errores a cero, se tiene

$$\begin{aligned}y_{t-1} &= \alpha_0 + \delta_0 c + \delta_1 c + \delta_2 c, \\y_t &= \alpha_0 + \delta_0 (c + 1) + \delta_1 c + \delta_2 c, \\y_{t+1} &= \alpha_0 + \delta_0 (c + 1) + \delta_1 (c + 1) + \delta_2 c, \\y_{t+2} &= \alpha_0 + \delta_0 (c + 1) + \delta_1 (c + 1) + \delta_2 (c + 1),\end{aligned}$$

y así sucesivamente. Con el aumento permanente en z , luego de un periodo, y ha aumentado en $\delta_0 + \delta_1$, y después de dos periodos, y se ha incrementado en $\delta_0 + \delta_1 + \delta_2$. No hay modificaciones adicionales en y al cabo de dos periodos. Esto muestra que la suma de los coeficientes de la z actual y las z rezagadas, $\delta_0 + \delta_1 + \delta_2$, es el cambio a largo plazo en y dado un incremento permanente en z y se denomina *propensión de largo plazo* (PLP) **o multiplicador de largo plazo**. Con frecuencia, la PLP es de interés en los modelos de rezagos distribuidos.

A manera de ejemplo, en la ecuación (10.4), δ_0 mide el cambio inmediato en la fertilidad debido a un aumento de un dólar en pe . Como se mencionó antes, hay razones para pensar que δ_0 es pequeña, si no es que cero. Sin embargo δ_1 o δ_2 , o ambas, podrían ser positivas. Si pe aumenta un dólar de manera permanente, entonces al cabo de dos años, gfr habrá cambiado en $\delta_0 + \delta_1 + \delta_2$. Este modelo supone que no hay más modificaciones luego de dos años. Si esto es de este modo o no, se trata de un asunto empírico.

Un modelo de rezagos distribuidos finitos de orden q se escribe como

$$[10.6] \quad y_t = \alpha_0 + \delta_0 z_t + \delta_1 z_{t-1} + \dots + \delta_q z_{t-q} + u_t.$$

Éste contiene el modelo estático como un caso especial, al establecer que $\delta_1, \delta_2, \dots, \delta_q$ son iguales a cero. En ocasiones, el propósito primordial de la estimación de un modelo de rezagos distribuidos es probar si z ejerce un efecto rezagado sobre y . El multiplicador de impacto siempre es el coeficiente de la z actual, δ_0 . En ocasiones, se omite la z_t en la ecuación (10.6), en cuyo caso el multiplicador de impacto es cero. La distribución de rezagos es de nuevo la δ_j graficada como función de j . El multiplicador de largo plazo es la suma de todos los coeficientes de las variables z_{t-j} :

$$[10.7] \quad PLP = \delta_0 + \delta_1 + \dots + \delta_q.$$

Por la correlación, a menudo sustancial, entre z y sus diferentes rezagos, es decir, debido a la multicolinealidad en (10.6), resulta difícil obtener estimadores precisos de cada δ_j individual. Curiosamente, aun cuando la δ_j no pueda estimarse con precisión, a menudo se pueden obtener buenas estimaciones de la PLP. Más adelante se verá un ejemplo.

Es posible tener más de una variable explicativa que aparezca con rezagos o agregar variables contemporáneas a un modelo de RDF. Por ejemplo, el nivel de educación promedio para las mujeres en edad fértil se podría sumar a (10.4), lo que permite dar cuenta de los cambios en los grados de escolaridad de las mujeres.

Una convención sobre el índice de tiempo

Cuando los modelos cuentan con variables explicativas rezagadas (y, como se verá en el siguiente capítulo, con y rezagadas), puede surgir confusión respecto al tratamiento de las observaciones iniciales. Por ejemplo, si se supone que la ecuación (10.5) se cumple comenzando en $t = 1$, entonces las variables explicativas del primer periodo son z_t, z_0 y z_{-1} . Por convención, éstos son los valores iniciales de nuestra muestra, de modo que siempre comienza el índice temporal en $t = 1$. En la práctica, esto no es muy importante, ya que los paquetes de regresión dan seguimiento de forma automática a las observaciones disponibles para la estimación de modelos con rezagos. Sin embargo, para éste y algunos de los capítulos siguientes, se necesitan algunas convenciones sobre el primer periodo que puede ser representado por la ecuación de la regresión.

10.3 Propiedades en muestras finitas de MCO bajo los supuestos clásicos

En esta sección se proporciona una lista completa de las propiedades en muestras finitas o muestras pequeñas de los MCO de acuerdo con los supuestos estándar. Se presta particular atención a la forma en que deben modificarse las premisas de nuestro análisis de corte transversal con el fin de cubrir las regresiones con series de tiempo.

Insesgamiento de MCO

El primer supuesto plantea sencillamente que el proceso de series de tiempo sigue un modelo que es lineal en sus parámetros.

Supuesto ST.1 Lineal en los parámetros

El proceso estocástico $\{(x_{t1}, x_{t2}, \dots, x_{tk}, y_t) : t = 1, 2, \dots, n\}$ sigue el modelo lineal

[10.8]
$$y_t = \beta_0 + \beta_1 x_{t1} + \dots + \beta_k x_{tk} + u_t,$$

donde $\{u_t : t = 1, 2, \dots, n\}$ es una secuencia de errores o perturbaciones. Aquí, n es el número de observaciones (periodos).

En la notación x_{ij} , t denota el periodo y j es, como de costumbre, una etiqueta para señalar una de las k variables explicativas. En este punto se aplica la terminología empleada en la regresión de corte transversal: y_t es la variable dependiente, la variable explicada o el regresando; las x_{ij} son las variables independientes, variables explicativas o los regresores.

Debemos pensar en el supuesto ST.1 como si fuera idéntico en esencia a la premisa RLM.1 (el primer supuesto del corte transversal), sólo que ahora se está especificando un modelo lineal para datos de series de tiempo. Los ejemplos tratados en la sección 10.2 se pueden expresar en la forma de la ecuación (10.8) al definir x_{ij} de manera adecuada. Por ejemplo, la ecuación (10.5) se obtiene al establecer $x_{t1} = z_t, x_{t2} = z_{t-1}$ y $x_{t3} = z_{t-2}$.

Para plantear y analizar varios de los supuestos restantes, supongan que $x_t = (x_{t1}, x_{t2}, \dots, x_{tk})$ denota el conjunto de todas las variables independientes en la ecuación en el momento t . Además, \mathbf{X} denota la colección de todas las variables independientes para todos los periodos. Es útil considerar a \mathbf{X} como una matriz, con n filas y k columnas, lo

que refleja la manera en que se almacenan los datos de series de tiempo en el software de econometría: la t -ésima fila de \mathbf{X} es x_t , que consta de todas las variables independientes en el periodo t . Por tanto, la primera fila de \mathbf{X} corresponde a $t = 1$, la segunda fila a $t = 2$ y la última a $t = n$. En la tabla 10.2 se proporciona un ejemplo que utiliza $n=8$ y las variables explicativas de la ecuación (10.3).

Como es natural, al igual que con la regresión de corte transversal, es necesario descartar la colinealidad perfecta entre los regresores.

Supuesto ST.2 No hay colinealidad perfecta

En la muestra (y, por ende, en los procesos de series de tiempo subyacentes) no hay variables independientes que sean constantes ni que sean una combinación lineal perfecta de las otras.

| t | $convrte$ | $unem$ | $yngmle$ |
|-----|-----------|--------|----------|
| 1 | .46 | .074 | .12 |
| 2 | .42 | .071 | .12 |
| 3 | .42 | .063 | .11 |
| 4 | .47 | .062 | .09 |
| 5 | .48 | .060 | .10 |
| 6 | .50 | .059 | .11 |
| 7 | .55 | .058 | .12 |
| 8 | .56 | .059 | .13 |

Tabla 10.2

Este supuesto se estudió con detenimiento en el contexto de los datos de corte transversal en el capítulo 3. Los problemas son en esencia los mismos con los datos de series de tiempo. Recuerden, el supuesto ST.2 permite que las variables explicativas se correlacionen, pero **descarta la correlación perfecta en la muestra**.

El supuesto final requerido para el insesgamiento de MCO es el equivalente en series de tiempo a la premisa RLM.4, y también obvia la necesidad de un muestreo aleatorio de la premisa RLM.2.

Supuesto ST.3 Media condicional cero

Para cada t , dadas las variables explicativas para todos los periodos, el valor esperado del error u_t es cero. Matemáticamente,

$$[10.9] \quad E(u_t | \mathbf{X}) = 0, \quad t = 1, 2, \dots, n.$$

Este supuesto es crucial y es necesario tener una idea de su significado. Como en el caso del corte transversal, resulta mucho más sencillo ver este supuesto en función de la ausencia de correlación: el supuesto ST.3 implica que el error en el periodo t , u_t , no se correlaciona con ninguna variable explicativa en **cada** uno de los periodos. El hecho de que esto se plantee en términos de la esperanza condicional implica que se debe especificar de forma adecuada la relación funcional entre y_t y las variables explicativas. Si u_t es independiente de \mathbf{X} y $E(u_t) = 0$, el supuesto ST.3 es válido de manera automática.

Dado el análisis de corte transversal del capítulo 3, no es de sorprender que se necesite que u_t no esté correlacionada con las variables explicativas fechadas en el mismo periodo t : en términos de la media condicional,

$$[10.10] \quad E(u_t | x_{t1}, \dots, x_{tk}) = E(u_t | \mathbf{x}_t) = 0.$$

Cuando la ecuación (10.10) es válida, se dice que las x_{tj} son **contemporáneamente exógenas**. La ecuación (10.10) implica que u_t y las variables explicativas no se correlacionan de manera contemporánea: $\text{Corr}(x_{tj}, u_t) = 0$, para toda j .

El supuesto ST.3 exige algo más que exogeneidad contemporánea: u_t no debe correlacionarse con x_{sj} , ni siquiera cuando $s \neq t$. **Este es el sentido estricto en el que las variables explicativas deben ser exógenas, y cuando ST.3 es válido decimos que las variables explicativas son estrictamente exógenas**. En el capítulo 11 se mostrará que la ecuación (10.10) es suficiente para probar la consistencia de los estimadores de MCO. Pero para demostrar que los estimadores de MCO son insesgados, se necesita el supuesto de exogeneidad estricta.

En el caso del corte transversal, no se estipula de manera explícita la forma en que el término de error, por ejemplo para la persona i , u_i , se correlaciona con las variables explicativas para **otras** personas de la muestra. Esto no fue necesario debido a que con el muestreo aleatorio (premisa RLM.2), u_i automáticamente es independiente de las variables explicativas en las observaciones distintas de i . **En un contexto de series de tiempo, el muestreo aleatorio casi nunca es adecuado, de manera que se debe suponer de forma explícita que el valor esperado de u_t no se relaciona con las variables explicativas en ningún periodo**.

Es importante tomar en cuenta que el supuesto ST.3 no pone restricciones a la correlación de las variables independientes o en u_t a lo largo del tiempo. El supuesto ST.3 sólo indica que el valor esperado de u_t no se relaciona con las variables independientes en todos los periodos.

Todo lo que provoque que los factores inobservables en el periodo t se correlacionen con alguna de las variables explicativas en cualquier periodo invalida el supuesto ST.3. Dos candidatos para esta invalidación son las variables omitidas y el error de medición en alguno de los regresores. Sin embargo, el supuesto de exogeneidad estricta también puede invalidarse por otras razones menos obvias. En el modelo de regresión estática simple

$$y_t = \beta_0 + \beta_1 z_t + u_t,$$

El supuesto ST.3 requiere no sólo que u_t y z_t no estén correlacionadas, sino que u_t tampoco se correlacione con valores pasados ni futuros de z . Esto tiene dos implicancias. **En primer lugar, z no puede tener un efecto rezagado sobre y .** Si z tiene un efecto rezagado sobre y , entonces se debe estimar un modelo de rezagos distribuidos. **Un punto más sutil es que la exogeneidad estricta excluye la posibilidad de que cambios en el término de error de hoy ocasionen modificaciones futuras en z .** Esto descarta en forma eficaz la retroalimentación de y para valores futuros de z . Por ejemplo, consideren un modelo estático simple para explicar la tasa de homicidios de una ciudad en términos de oficiales de policía per cápita:

$$mrd_rte_t = \beta_0 + \beta_1 polpc_t + u_t.$$

Tal vez sea razonable suponer que u_t no se correlaciona con $polpc_t$ e incluso con valores pasados de $polpc_t$; supongan que lo anterior es cierto, pero que la ciudad ajusta el tamaño de su fuerza policiaca con base en los valores pasados de la tasa de homicidios. Esto significa que, $polpc_{t+1}$ se correlacionaría con u_t (ya que una variable u_t mayor conduce a una mrd_rte_t mayor). Si así fuera el caso, en general se violaría el supuesto ST.3.

Hay consideraciones semejantes en los modelos de rezagos distribuidos. Por lo general, no preocupa que u_t se correlacione con la z anterior debido a que estamos controlando las z anteriores en el modelo. **Pero la retroalimentación de u sobre las z futuras siempre representa un problema.**

Las variables explicativas que son estrictamente exógenas no reaccionan a lo que ha sucedido con y en el pasado. Un factor como la cantidad de precipitación pluvial en una función de producción agrícola satisface este requisito: la precipitación en cualquier año en el futuro no se ve influida por la producción del año en curso o de los años anteriores. Sin embargo, una variable como la cantidad de trabajo podría no ser estrictamente exógena, ya que el agricultor es quien la elige y éste puede ajustarla con base en la producción del último año. Las variables de política, como el crecimiento en la masa monetaria, los gastos de prestaciones sociales y los límites de velocidad en las autopistas, a menudo se ven influidos por lo que ha sucedido en el pasado. **En las ciencias sociales, muchas variables explicativas bien pueden infringir la premisa de exogeneidad estricta.**

Aun cuando el supuesto ST.3 sea poco realista, partimos de éste para llegar a la conclusión de que los estimadores de MCO son insesgados. La mayor parte de los tratamientos de los modelos estáticos y de los rezagos distribuidos dan por hecho ST.3 al considerar el supuesto más estricto de que las variables explicativas no son aleatorias o son fijas en muestras repetidas. El supuesto de no aleatoriedad desde luego es falso para las observaciones de series de tiempo; el supuesto ST.3 cuenta con la ventaja de ser más realista respecto a la naturaleza aleatoria de las x_{jt} , mientras que aísla la premisa necesaria respecto a cómo se relacionan u_t y las variables explicativas para que los estimadores de MCO sean insesgados.

Teorema 10.1 Insesgamiento de los estimadores de MCO

Bajo los supuestos ST.1, ST.2 y ST.3, los estimadores de MCO son insesgados condicionales sobre \mathbf{X} , y por tanto también incondicionalmente: $E(\hat{\beta}_j) = \beta_j, j = 0, 1, \dots, k$.

Es necesario agregar dos supuestos para complementar las premisas de Gauss-Markov para las regresiones con series de tiempo. La primera resulta familiar gracias al análisis de corte transversal.

Supuesto ST.4 Homocedasticidad

La varianza de u_t condicional en \mathbf{X} , es la misma para cualquier t : $Var(u_t | \mathbf{X}) = Var(u_t) = \sigma^2, t = 1, 2, \dots, n$.

Este supuesto significa que $Var(u_t | \mathbf{X})$ no puede depender de \mathbf{X} —es suficiente con que u_t y \mathbf{X} sean independientes— además que $Var(u_t)$ debe ser constante en el tiempo. Cuando el supuesto ST.4 no es válido, se dice que los errores son **heterocedásticos**, como en el caso del corte transversal. Por ejemplo, consideren una ecuación para determinar las tasas de las letras del Tesoro estadounidense a tres meses ($i3_t$) con base en la tasa de inflación (inf_t) y en el déficit federal como porcentaje del producto interno bruto (def_t):

$$[10.11] \quad i3_t = \beta_0 + \beta_1 inf_t + \beta_2 def_t + u_t.$$

Entre otras cosas, el supuesto ST.4 exige que los factores inobservables que influyen sobre las tasas de interés tengan una varianza constante respecto al tiempo. Como se sabe que los cambios en el régimen de políticas influyen sobre la variabilidad de las tasas de interés, este supuesto bien puede ser falso. Además, tal vez la variabilidad de las tasas de interés dependa del nivel de la inflación o del tamaño relativo del déficit. Esto también violaría la premisa de la homocedasticidad.

Cuando $Var(u_t | \mathbf{X})$ depende de \mathbf{X} , a menudo está subordinada a las variables explicativas en el periodo t , x_t . En el capítulo 12 se verá que las pruebas de heterocedasticidad del capítulo 8 también se emplean para las regresiones con series de tiempo, por lo menos bajo ciertos supuestos.

El supuesto final de Gauss-Markov para el análisis de series de tiempo es nuevo.

Supuesto ST.5 No hay correlación serial

Los errores, condicionales sobre \mathbf{X} , en dos periodos distintos, no están correlacionados: $Corr(u_t, u_s | \mathbf{X}) = 0$, para cualquier $t \neq s$.

La forma más sencilla de considerar este supuesto es ignorar el condicionamiento sobre \mathbf{X} . Así, el supuesto ST.5 es sencillamente

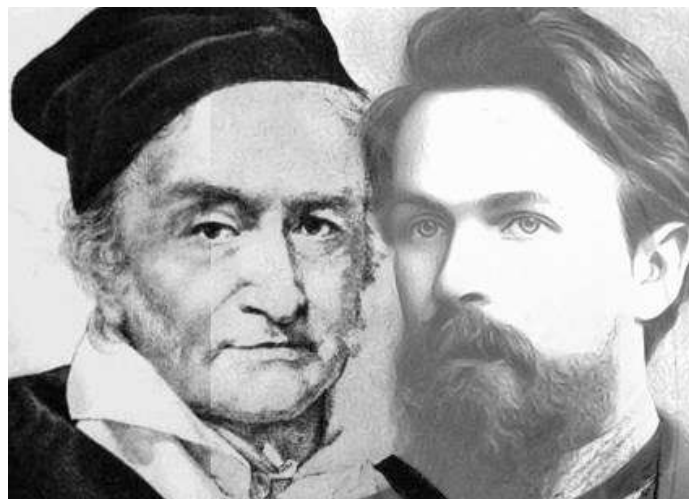
$$[10.12] \quad Corr(u_t, u_s) = 0, \text{ para cualquier } t \neq s.$$

(Así es como se plantea el supuesto de no correlación serial cuando \mathbf{X} se trata como no aleatoria.) Cuando se considera que el supuesto ST.5 tiene posibilidades de ser válido, nos concentramos en la ecuación (10.12) debido a que es fácil interpretarla.

Cuando (10.12) es falsa, se dice que los errores en (10.8) tienen correlación serial o autocorrelación, debido a que se correlacionan a lo largo del tiempo. Consideren el caso de los errores de periodos adyacentes. Supongan que cuando $u_{t-1} > 0$ entonces, en promedio, el error en el siguiente periodo, u_t , también es positivo. Por tanto, $\text{Corr}(u_t, u_{t-1}) > 0$, y los errores sufren de correlación serial. Esto quiere decir, en la ecuación (10.11), que si las tasas de interés aumentan de manera abrupta en este periodo, es probable que en el siguiente estén por encima del promedio (para los niveles dados de inflación y déficit). Esta caracterización resulta ser razonable para los términos de error en muchas aplicaciones de series de tiempo, las cuales se verán en el capítulo 12. Por ahora, damos por sentado el supuesto ST.5.

Es importante destacar que el ST.5 no supone nada sobre la correlación temporal de las variables independientes. Por ejemplo, en la ecuación (10.11) es casi seguro que inf_t esté correlacionada en el tiempo. Pero esto no tiene nada que ver con la validez del ST.5.

Una pregunta natural es: ¿por qué en los capítulos 3 y 4 no supusimos que los errores para diferentes observaciones de corte transversal



Carl Friedrich Gauss y Andrey Markov
Quinlan Lee - [OLS, BLUE and the Gauss Markov Theorem](#)

no están correlacionados? La respuesta proviene del supuesto de muestreo aleatorio: en el muestreo aleatorio, u_i y u_h son independientes para cualesquiera dos observaciones i y h . También se demuestra que bajo el muestreo aleatorio los errores para las diferentes observaciones, condicionales sobre las variables explicativas de la muestra, son independientes. Así, para nuestros propósitos, consideramos que la correlación serial sólo es un problema en las regresiones con series de tiempo. (En los capítulos 13 y 14, el problema de la correlación serial se presentará en el análisis de datos de panel.)

Los supuestos ST.1 a ST.5 son los adecuados de Gauss-Markov para las aplicaciones de series de tiempo, aunque también tienen otros usos. A veces los supuestos ST.1 a ST.5 se satisfacen en las aplicaciones de corte transversal, aun cuando el muestreo aleatorio no sea un supuesto razonable, como cuando las unidades de corte transversal son grandes en relación con la población. Imaginen que tiene una base de datos de corte transversal en el ámbito ciudadano. Podría suceder que exista una correlación en algunas de las variables explicativas entre las ciudades de un mismo estado, como el impuesto predial o los pagos de seguridad social per cápita. La correlación de las variables explicativas entre las observaciones no genera problemas para verificar los supuestos de Gauss-Markov, siempre y cuando los errores no se correlacionen entre estas ciudades. Sin embargo, en este capítulo nuestro principal interés es la aplicación de los supuestos de Gauss-Markov a los modelos de regresión con datos de series de tiempo.

Teorema 10-2 Varianzas de muestreo de los estimadores de MCO

Con base en los supuestos ST.1 a ST.5 de Gauss-Markov para las series de tiempo, la varianza de β_j° , condicional sobre \mathbf{X} , es

$$[10.13] \quad \text{Var}(\beta_j^\circ | \mathbf{X}) = \sigma^2 / [STC_j (1 - R_j^2)], j = 1, \dots, k,$$

donde STC_j es la suma total de cuadrados de x_{jt} y R_j^2 es la R-cuadrada de la regresión de x_{jt} sobre las otras variables independientes.

La ecuación (10.13) es la misma varianza que se determinó en el capítulo 3 bajo los supuestos de Gauss-Markov para corte transversal. Como la demostración es muy parecida a la del teorema 3.2, aquí se omite. El planteamiento del capítulo 3 sobre los factores que generan grandes varianzas, incluida la multicolinealidad entre las variables explicativas, se aplica de inmediato al caso de las series de tiempo.

El estimador usual de la varianza del error también es insesgado bajo los supuestos ST.1 a ST.5 y el teorema de Gauss-Markov es válido.

Teorema 10-3 Estimación insesgada de σ^2

Bajo los supuestos ST.1 a ST.5, el estimador $\sigma^{\circ 2} = SRC/gl$ es un estimador insesgado de σ^2 , donde $gl = n - k - 1$ son los grados de libertad.

Teorema 10-4 Teorema de Gauss-Markov

Bajo los supuestos ST.1 a ST.5, los estimadores de MCO son los mejores estimadores lineales insesgados condicionales sobre \mathbf{X} .

Inferencia bajo los supuestos del modelo lineal clásico

Para utilizar los errores estándar usuales de MCO, los estadísticos t y F, se necesita agregar un último supuesto que es análogo a la premisa de normalidad empleada en el análisis de corte transversal.

Supuesto ST.6 Normalidad

Los errores u_t son independientes de \mathbf{X} son independientes y están idénticamente distribuidos como una Normal $(0, \sigma^2)$.

El supuesto ST.6 comprende los supuestos ST.3, ST.4 y ST.5, pero es más fuerte debido a los supuestos de independencia y normalidad.

Teorema 10-5 Distribuciones de muestreo normales

Bajo los supuestos ST.1 a ST.6, los supuestos del MCL para series de tiempo, los estimadores de MCO se distribuyen de forma normal, condicionales sobre \mathbf{X} . Además, bajo la hipótesis nula, cada estadístico t tiene una distribución \underline{t} y cada estadístico F tiene una distribución \underline{F} . También es válida la construcción usual de los intervalos de confianza.

Las consecuencias del teorema 10.5 son de suma importancia. Implican que, cuando los supuestos ST.1 a ST.6 son válidos, todo lo aprendido acerca de la estimación y la inferencia para las regresiones con cortes transversales se aplica de manera directa a las regresiones con series de tiempo. Así pues, los estadísticos t pueden emplearse para probar la significancia estadística de las variables explicativas individuales y los estadísticos F pueden utilizarse para probar la significancia conjunta.

Como en el caso del corte transversal, los usuales procedimientos de inferencia están supeditados a los supuestos subyacentes. Los supuestos del modelo lineal clásico para los datos de series de tiempo son mucho más restrictivos que aquellos de corte transversal —en particular, los supuestos, de exogeneidad estricta y no correlación serial pueden resultar poco realistas—. Con todo, el esquema del modelo lineal clásico (MLC) es un buen punto de partida para muchas aplicaciones.

10.4 *Forma funcional, variables binarias y números índice*

Todas las formas funcionales aprendidas en capítulos anteriores se utilizan en las regresiones con series de tiempo. La más importante es el *logaritmo natural*: en el trabajo aplicado con frecuencia aparecen regresiones con series de tiempo que tienen efectos porcentuales constantes.

Castillo-Freeman y Freeman utilizan datos anuales sobre la tasa de empleo en Puerto Rico, el salario mínimo y otras variables para estudiar los efectos que tiene el salario mínimo estadounidense en el empleo en Puerto Rico.³ Una versión simplificada de su modelo es

$$[10.16] \quad \log(\text{prepop}_t) = \beta_0 + \beta_1 \log(\text{mincov}_t) + \beta_2 \log(\text{usgnpt}_t) + u_t,$$

en donde *prepop* es la tasa de empleo en Puerto Rico durante el año t (proporción de las personas que trabajan respecto a la población total), *usgnpt* es el producto nacional bruto real de Estados Unidos (en miles de millones de dólares) y *mincov* mide la importancia del salario mínimo en relación con los salarios promedio. En particular, $\text{mincov} = (\text{avgmin}/\text{avgwage}) \cdot \text{avgcov}$, donde *avgmin* es el salario mínimo promedio, *avgwage* es el salario general promedio y *avgcov* es la tasa de cobertura promedio (la proporción real de trabajadores cubiertos bajo la ley del salario mínimo).

Usando la base de datos PRMINWGE.RAW para los años de 1950 a 1987 se obtiene

$$[10.17] \quad \log(\text{prepop}_t)^\circ = -1.05 - .154 \log(\text{mincov}_t) - .012 \log(\text{usgnpt}_t) \\ \phantom{\log(\text{prepop}_t)^\circ =} .077 .065 .089 \\ \phantom{\log(\text{prepop}_t)^\circ =} \phantom{-1.05 - .154 \log(\text{mincov}_t) - .012 \log(\text{usgnpt}_t)} n = 38, R^2 = 0.661, R^2_{aj} = 0.641.$$

La elasticidad estimada de *prepop* respecto a *mincov* es $-.154$, y es estadísticamente significativa con $t = -2.37$. Por consiguiente, un salario mínimo mayor disminuye la tasa de empleo, hecho que la economía clásica predice. La variable del PNB no resulta es-

³ Alida Castillo-Freeman y Richard B. Freeman, *When the Minimum Wage Really Bites: The Effect of the U.S.-Level Minimum on Puerto Rico*, Ch. 6 en *Immigration and the Workforce: Economic Consequences for the United States and Source Areas* (editado por George J. Borjas y Richard B. Freeman, 1992).

tadísticamente significativa, pero esto cambiará cuando se considere una tendencia en el tiempo en la siguiente sección.

Podemos emplear también formas logarítmicas funcionales en modelos de rezagos distribuidos. Por ejemplo, para los datos trimestrales, supongan que la demanda de dinero (M_t) se relaciona con el producto interno bruto (GDP_t) mediante

$$\log(M_t) = \alpha_0 + \delta_0 \log(GDP_t) + \delta_1 \log(GDP_{t-1}) + \delta_2 \log(GDP_{t-2}) + \delta_3 \log(GDP_{t-3}) + \delta_4 \log(GDP_{t-4}) + u_t.$$

La propensión de impacto en esta ecuación, δ_0 , también se conoce como la **elasticidad de corto plazo**: mide el cambio porcentual inmediato en la demanda de dinero cuando hay un aumento de 1% en el GDP. La propensión de largo plazo, $\delta_0 + \delta_1 + \dots + \delta_4$, con frecuencia se denomina la **elasticidad de largo plazo**: mide el incremento porcentual en la demanda de dinero luego de cuatro trimestres cuando hay un aumento permanente de 1% en el GDP.

Las variables binarias también son muy útiles en las aplicaciones de series de tiempo. Dado que la unidad de observación es el tiempo, una variable binaria representa si, en cada periodo, ha ocurrido un evento determinado. Por ejemplo, para los datos anuales podemos indicar en cada año si el presidente de Estados Unidos es un demócrata o un republicano con sólo definir una variable $democ_t$, igual a uno cuando el presidente sea demócrata e igual a cero cuando no lo sea. O bien, al considerar los efectos de la pena de muerte en las tasas de homicidio en Texas, podemos definir una variable binaria para cada año que sea igual a uno si se aplicó la pena de muerte en Texas durante ese año y a cero en caso contrario.

Con frecuencia, las variables binarias se emplean para aislar ciertos periodos que pueden ser sistemáticamente distintos de otros lapsos comprendidos en una base de datos.

Ejemplo 10.4 Efectos de la exención personal sobre las tasas de fertilidad

La tasa de fertilidad general (gfr) es el número de niños nacidos por cada 1,000 mujeres en edad de concebir. Para los años de 1913 a 1984, la ecuación,

$$gfr_t = \beta_0 + \beta_1 pe_t + \beta_2 ww2_t + \beta_3 pill_t + u_t,$$

explica gfr en términos del valor en dólares reales de la exención personal de impuestos promedio (pe) y dos variables binarias. La variable $ww2$ asume el valor de uno para los años de 1941 a 1945, cuando Estados Unidos se involucró en la Segunda Guerra Mundial. La variable $pill$ es uno de 1963 en adelante, cuando salió al mercado la píldora anticonceptiva para el control de la natalidad.

Usando los datos de FERTIL3.RAW, que se obtuvieron del artículo de Whittington, Alm y Peters,⁴ se obtiene

⁴ Whittington, Leslie A & Alm, James & Peters, H Elizabeth, 1990. *Fertility and the Personal Exemption: Implicit Pronatalist Policy in the United States*, American Economic Review, American Economic Association, vol. 80(3).

$$\begin{aligned}
 [10.18] \quad gfr_t = & 98.68 + .083 pe_t - 24.24 ww2_t - 31.59 pill_t \\
 & (3.21) \quad (.03) \quad (7.46) \quad (4.08) \\
 & n=72 \quad R^2 = .473, \quad R^2_{aj} = .450.
 \end{aligned}$$

Cada variable es estadísticamente significativa al nivel de 1% contra la alternativa de dos colas. Se ve que la tasa de fertilidad fue menor durante la Segunda Guerra Mundial: dada pe , hubo alrededor de 24 nacimientos menos por cada 1,000 mujeres en edad de procrear, lo cual es una gran reducción. (De 1913 a 1984, gfr osciló entre 65 y 127.) De igual modo, la tasa de fertilidad disminuyó de manera sustancial a partir de la introducción de la píldora anticonceptiva.

La variable de interés económico es pe . La pe media en este periodo es de \$100.40 dólares y oscila entre cero y \$243.83 dólares. El coeficiente de pe implica que un aumento de 12 dólares en pe provoca un incremento en gfr cercano a un nacimiento por cada 1,000 mujeres en edad de concebir. Este efecto no es poca cosa.

En la sección 10.2 se advirtió que la tasa de fertilidad reacciona a los cambios en pe en forma rezagada. La estimación de un modelo de rezagos distribuidos con dos rezagos nos da

$$\begin{aligned}
 [10.19] \quad gfr_t = & 95.87 + .073 pe_t - .0058 pe_{t-1} + .034 pe_{t-2} - 22.13 ww2_t - 31.30 pill_t \\
 & (3.28) \quad (.126) \quad (.1557) \quad (.126) \quad (10.73) \quad (3.98) \\
 & n = 70, \quad R^2 = .499, \quad R^2_{aj} = .459.
 \end{aligned}$$

En esta regresión sólo se cuenta con 70 observaciones, ya que se perdieron dos al rezagar pe dos veces. Los coeficientes en las variables pe se estiman de forma muy imprecisa y cada uno por separado es insignificante. Resulta que hay una alta correlación entre pe_t , pe_{t-1} y pe_{t-2} , y esta multicolinealidad dificulta la estimación del efecto de cada rezago. Sin embargo, pe_t , pe_{t-1} y pe_{t-2} en conjunto son significativas: el estadístico F tiene un $valor-p = .012$. Así, pe ejerce un efecto en gfr [como se vio en la ecuación (10.18)], pero no contamos con buenas estimaciones para determinar si es contemporáneo o tiene un rezago de uno o dos años (o algo de todo). En realidad, pe_{t-1} y pe_{t-2} son conjuntamente insignificantes en esta ecuación ($valor-p = .95$), de modo que en este momento se justificaría que utilizáramos el modelo estático. Pero para ilustrarlo mejor, obtengan un intervalo de confianza para la propensión de largo plazo de este modelo.

La PLP estimada en la ecuación (10.19) es $.073 - .0058 + .034 = .101$. Sin embargo, no contamos con información suficiente en (10.19) para obtener el error estándar de esta estimación. Con el fin de obtener el error estándar de la PLP estimada, se usa el truco propuesto en la sección 4.4. Sea $\theta_0 = \delta_0 + \delta_1 + \delta_2$ la PLP y escriban δ_0 en términos de θ_0 , δ_1 y δ_2 como $\delta_0 = \theta_0 - \delta_1 - \delta_2$. Ahora sustituyan δ_0 en el modelo

$$gfr_t = \alpha_0 + \delta_0 pe_t + \delta_1 pe_{t-1} + \delta_2 pe_{t-2} + \dots$$

para obtener

$$\begin{aligned}
 gfr_t = & \alpha_0 + (\theta_0 - \delta_1 - \delta_2) pe_t + \delta_1 pe_{t-1} + \delta_2 pe_{t-2} + \dots \\
 = & \alpha_0 + \theta_0 pe_t + \delta_1 (pe_{t-1} - pe_t) + \delta_2 (pe_{t-2} - pe_t) + \dots
 \end{aligned}$$

De esta última ecuación se obtiene θ^o y su error estándar al hacer la regresión de gfr_t sobre pe_t , $(pe_{t-1} - pe_t)$, $(pe_{t-2} - pe_t)$, $ww2_t$ y $pill_t$. El coeficiente y el error estándar asociado con pe_t son lo que necesitamos. Al realizar esta regresión se obtiene $\theta^o = .101$ como el coeficiente de pe_t (lo que ya se conocía) y $ee(\theta^o) = .030$ [que no podíamos calcular a partir de (10.19)]. Por tanto, el estadístico t para θ^o es cerca de 3.37, de manera que θ^o es estadísticamente distinto de cero a niveles de significancia pequeños. Aun cuando ninguna de las θ_j tiene significancia por separado, la PLP es muy significativa. El intervalo de confianza de 95% para la PLP es de entre .041 y .160 aproximadamente.

Whittington, Alm y Peters permiten más rezagos, pero limitan los coeficientes para ayudar a paliar el problema de multicolinealidad que dificulta la estimación de las δ_j individuales. (El problema 10.6 proporciona un ejemplo de cómo llevarlo a cabo.) Para estimar la PLP, lo cual parecería ser de interés primordial aquí, tales restricciones son innecesarias. Whittington, Alm y Peters también controlan variables adicionales, como el salario promedio de las mujeres y la tasa de desempleo.

Variables binarias y estudios de eventos

Las variables binarias explicativas son el componente clave en lo que se denomina estudio de evento. En un estudio de este tipo, el objetivo es ver si un determinado evento influye en algún resultado. Los economistas que estudian organización industrial han investigado los efectos de ciertos eventos sobre los precios de las acciones de las empresas. Por ejemplo, Rose estudió los efectos que tienen las nuevas reglamentaciones del transporte por carretera sobre los precios de las acciones de las compañías del ramo.⁵

Una versión sencilla de una ecuación utilizada para tales estudios de eventos es

$$R^f_t = \beta_0 + \beta_1 R^m_t + \beta_2 d_t + u_t,$$

donde R^f_t es el rendimiento de las acciones de la empresa f durante el periodo t (por lo común una semana o un mes), R^m_t es el rendimiento del mercado (que por lo general se calcula para un amplio índice bursátil), y d_t es una variable binaria que indica cuándo ocurrió el evento. Por ejemplo, si la empresa es una aerolínea, d_t podría denotar si la empresa sufrió un accidente públicamente conocido, o estuvo próxima a uno, durante la semana t . Incluir R^m_t en la ecuación controla la posibilidad de que amplios movimientos del mercado pudieran coincidir con los accidentes de la aerolínea. A veces se emplean múltiples variables binarias. Por ejemplo, si el evento es la imposición de una nueva reglamentación que pudiese influir sobre una empresa determinada, se incluiría una variable binaria que abarcara unas cuantas semanas antes del anuncio público de la reglamentación y una segunda variable para unas semanas después de dicho anuncio. La primera variable podría detectar la presencia de información interna.

Antes de dar un ejemplo de un estudio de evento, es necesario formular la noción de un número índice y la diferencia entre las variables económicas nominales y las reales. Un número índice por lo general agrega un gran volumen de en una sola cantidad. Las cifras de esta índole se emplean por lo general en el análisis de series de tiempo, en especial en aplicaciones macroeconómicas. Un ejemplo de número índice es el índice de la

⁵ Nancy L. Rose, *The incidence of regulatory rents in the motor carrier industry*, Rand Journal of Economics, Vol. 16, No. 3, Autumn 1985.

producción industrial (IPI), calculado cada mes por el consejo directivo de la Reserva Federal. El IPI es una medida de producción de un amplio espectro de industrias y, como tal, su magnitud en un determinado año no tiene un sentido cuantitativo. Para interpretar la magnitud del IPI se debe conocer el **periodo base** y el **valor base**. En el informe *1997 Economic Report of the President* (ERP) de 1997, el año base es 1987 y el valor base es 100. (Fijar el valor del IPI en 100 en el periodo base es sólo una convención; tiene tanto sentido como establecer IPI= 1 en 1987, y algunos índices se definen con un valor base de 1.) Como el IPI fue de 107.7 en 1992, se puede decir que la producción industrial fue 7.7% mayor en 1992 que en 1987. Se puede utilizar el IPI de cualquier bienio para calcular la diferencia porcentual en la producción industrial durante ese lapso. Por ejemplo, como IPI= 61.4 en 1970 e IPI= 85.7 en 1979, la producción industrial creció cerca de 39.6% durante la década de los setenta.

Es fácil modificar el periodo base de cualquier número índice y, a veces, se debe hacer para dar un año base común a los números índices reportados con distintos años base. Por ejemplo, si se desea cambiar el año base del IPI de 1987 a 1982, sólo se divide el IPI para cada año por el valor de 1982 y luego se multiplica por 100 para asignar el valor de 100 al periodo base. En general, la fórmula es

$$[10.20] \quad \text{nuevoíndice}_t = 100 (\text{viejoíndice}_t / \text{viejoíndice}_{\text{nuevabase}}),$$

donde $\text{viejoíndice}_{\text{nuevabase}}$ es el valor original del índice en el año base nuevo. Por ejemplo, con el año base 1987, el IPI en 1992 es de 107.7; si se modifica el año base a 1982, el IPI en 1992 se vuelve 100 ($107.7/81.9$) = 131.5 (ya que el IPI en 1982 fue 81.9).

Otro ejemplo importante de un número índice es el *índice de precios*, como el Índice de Precios al Consumidor (IPC). Ya se ha usado el IPC para calcular las tasas de inflación anuales en el ejemplo 10.1. Al igual que con el índice de producción industrial, el IPC sólo es importante cuando se compara a lo largo de varios años (o meses, si se tienen datos mensuales). En el ERP de 1997, se tiene $IPC = 38.8$ en 1970 y en 1990 $IPC = 130.7$. De esta manera, el nivel general de precios creció cerca de 237% en este periodo de 20 años. (En el ERP de 1997, el IPC se define de tal manera que su promedio en 1982, 1983 y 1984 es igual a 100; por tanto, el periodo base figura como 1982-1984.)

Además de utilizarse para calcular las tasas de inflación, los índices de precios son necesarios para convertir una serie de tiempo medida *en dólares nominales* (o dólares *corrientes*) en *dólares reales (o dólares constantes)*. Se supone que la mayor parte del comportamiento económico se ve influido por las variables reales y no por las nominales. Por ejemplo, la economía laboral clásica supone que la oferta de trabajo se basa en el salario real por hora y no en el nominal. Obtener el salario real a partir del nominal resulta sencillo si contamos con un índice de precios como el IPC. Se debe procurar dividir primero el IPC por 100, para que el valor en el año base sea 1. Así, si w denota el salario medio por hora en dólares nominales y $p = IPC/100$, *el salario real es simplemente w/p* . Este salario se mide **en dólares del periodo base del IPC**.

Por ejemplo, en la tabla B-45 del ERP de 1997, los ingresos por hora promedio se reportan en términos nominales y en dólares de 1982 (lo que significa que el IPC utilizado para calcular el salario real tuvo como año base 1982). Esta tabla informa que el salario nominal por hora en 1960 fue de \$2.09, pero medido en dólares de 1982, el salario fue

de \$6.79. El salario por hora real alcanzó su punto máximo en 1973, en \$8.55 de 1982, y disminuyó a \$7.40 en 1995. Así pues, hubo una disminución importante en los salarios reales durante esos 20 años. (Si se comparan los salarios nominales de 1973 y 1995, se obtiene un panorama muy engañoso: \$3.94 en 1973 y \$11.44 en 1995. En virtud de que el salario real bajó, el aumento en el salario nominal se debe por completo a la inflación.)

Las medidas estándar de las variables económicas se proporcionan en términos reales. La más importante de estas medidas es el **producto interno bruto o GDP**. Cuando la prensa popular informa el crecimiento del GDP, casi siempre es un crecimiento del GDP real. En el ERP de 1997 (tabla B-9), se reporta el GDP en miles de millones de dólares de 1992. En el ejemplo 10.3 se usó una medida similar de producción, el producto nacional bruto real.

Sucedan cosas interesantes cuando las variables en dólares reales se utilizan en combinación con los **logaritmos naturales**. Supongan, por ejemplo, que el promedio de horas semanales trabajadas se relaciona con el salario real mediante

$$\log(\text{hours}) = \beta_0 + \beta_1 \log(w/p) + u.$$

Considerando el hecho de que $\log(w/p) = \log(w) - \log(p)$, se puede escribir esto como

$$[10.21] \quad \log(\text{hours}) = \beta_0 + \beta_1 \log(w) - \beta_2 \log(p) + u,$$

con la restricción de que $\beta_2 = -\beta_1$. Por tanto, el supuesto de que sólo el salario real influye sobre la oferta de trabajo impone una restricción en los parámetros del modelo (10.21). Si $\beta_2 \neq -\beta_1$, entonces el nivel de precios ejercerá un efecto sobre la oferta de trabajo, algo que puede suceder si los trabajadores no entienden muy bien la distinción entre salarios reales y nominales.

Hay muchos aspectos prácticos para el cálculo de los números índice, pero llevaría mucho explicarlos aquí. La mayor parte de los libros de macroeconomía intermedia, como el de Mankiw (1994, capítulo 2),⁶ contienen análisis detallados sobre los índices de precios. Lo que aquí importa es saber aplicar los números índice en el análisis de regresión. Como se mencionó, dado que las magnitudes de los números índice no son en especial informativas, a menudo aparecen en forma logarítmica, de modo que los coeficientes de regresión se interpretan en términos de cambio porcentual.

Ahora se dará un ejemplo de un estudio de evento en el que también se aplican números índice.

Ejemplo 10.5 Demandas antidumping e importaciones químicas

Krupp y Pollard analizaron los efectos de las demandas antidumping de las industrias químicas estadounidenses sobre la importación de diversas sustancias químicas.⁷ Aquí se considera en particular una sustancia química industrial, el cloruro de bario, un agente limpiador utilizado en diversos procesos químicos y en la producción de gasoli-

⁶ N. Gregory Mankiw, *Macroeconomics*, 6th ed., 1994.

⁷ Corinne M. Krupp y Patricia Pollard, *Market Responses to Antidumping Laws: Some Evidence from the U.S. Chemical Industry*, Canadian Journal of Economics, 1996.

na. Los datos provienen del archivo BARIUM.RAW. A principios de la década de los ochenta, los productores estadounidenses de cloruro de bario consideraron que China estaba ofreciendo sus exportaciones a Estados Unidos a un precio injustamente bajo (acción conocida como *dumping*) y la industria del cloruro de bario interpuso una demanda ante la Comisión Estadounidense de Comercio Internacional (CCI) en octubre de 1983. La CCI falló a favor de la industria del cloruro de bario estadounidense en octubre de 1984. Hay varios interrogantes que resultan interesantes en este caso, pero sólo se mencionarán algunos de ellos. En primer lugar, ¿las importaciones son insólitamente altas en el periodo inmediato anterior a la demanda inicial? En segundo lugar, ¿las importaciones cambian de manera notable después de una demanda antidumping? Y por último, ¿cuál es la reducción en las importaciones luego del fallo a favor de la industria estadounidense?

Para responder a estas preguntas, seguimos a Krupp y Pollard y se definen tres variables binarias: *befile6* es igual a 1 durante los seis meses anteriores a la demanda, *affile6* indica los seis meses posteriores a la misma, y *afdec6* denota los seis meses que siguieron al fallo positivo. La variable dependiente es el volumen de importaciones de cloruro de bario de China, *chnimp*, que se emplea de forma logarítmica. Se incluyen como variables explicativas, todas en forma logarítmica, un índice de la producción química, *chempi* (para controlar la demanda general de cloruro de bario), el volumen de la producción de nafta o gasolina, *gas* (otra variable de demanda) y un índice del tipo de cambio, *rtwex*, que mide la solidez del dólar en comparación con otras divisas. El índice de producción química se definió en 100 en junio de 1977. Aquí el análisis difiere ligeramente del de Krupp y Pollard en cuanto a que aquí se utilizan logaritmos naturales de todas las variables (salvo las binarias, desde luego) y se incluyen las tres variables binarias en la misma regresión.

Los datos mensuales que se utilizan de febrero de 1978 a diciembre de 1988 proporcionan lo siguiente:

$$\begin{aligned}
 [10.22] \quad \log(\text{chnimp})^o = & -17.80 + 3.12 + (\text{chempi}) + .196 \log(\text{gas}) + \\
 & (21.05) \quad (.48) \quad \quad \quad (.907) \\
 & + .983 \log(\text{rtwex}) + .060 \text{befile6} - .032 \text{affile6} - .565 \text{afdec6} \\
 & (.400) \quad \quad \quad (.261) \quad \quad \quad (.264) \quad \quad \quad (.286)
 \end{aligned}$$

$$n = 131, R^2 = .305, R^2_{aj} = .271$$

La ecuación muestra que *befile6* es insignificante en términos estadísticos, de modo que no hay evidencia de que las importaciones chinas fueran de manera inusitada altas en los seis meses anteriores a la interposición de la demanda. Además, aun cuando el estimador del coeficiente de *affile6* es negativo, es pequeño (lo que indica una disminución de alrededor de 3.2% en las importaciones chinas) y es insignificante en términos estadísticos. El coeficiente de *afdec6* muestra una disminución sustancial en las importaciones chinas de cloruro de bario luego del fallo a favor de la industria estadounidense, que no es sorprendente. Como el efecto es tan grande, se calcula el cambio porcentual exacto: $100 [\exp(-.565) - 1] \approx -43.2\%$. El coeficiente es estadísticamente significativo a un nivel de 5% en contra de una alternativa de dos colas.

Los signos de los coeficientes de las variables de control son los esperados: un incremento en la producción general de químicos aumenta la demanda del agente limpiador. La producción de gasolina no afecta de forma significativa las importaciones chinas. El coeficiente sobre $\log(rtwex)$ muestra que un incremento en el valor del dólar respecto a otras monedas incrementa la demanda de las importaciones chinas, como lo predice la teoría económica. (De hecho, la elasticidad no es estadísticamente diferente de uno. ¿Por qué?)

Las *interacciones entre las variables cualitativas y cuantitativas* se emplean también en el análisis de series de tiempo. A continuación se da un ejemplo de importancia práctica.

Fair resume su trabajo explicando los resultados de la elección presidencial en términos del desempeño económico.⁸ Explica la proporción del voto que en una elección bipartidista obtiene el candidato demócrata con datos de 1916 a 1992 (cada cuatro años), para un total de 20 observaciones. Se estima una versión simplificada del modelo de Fair (utilizando nombres de variables que son más descriptivos que los suyos):

$$demvote = \beta_0 + \beta_1 partyWH + \beta_2 incum + \beta_3 partyWH \cdot gnews + \beta_4 partyWH \cdot inf + u,$$

donde *demvote* es la proporción del voto que corresponde al candidato demócrata en una elección bipartidista. La variable explicativa *partyWH* es similar a una variable binaria, pero adopta el valor 1 si un demócrata está en la Casa Blanca y -1 si está un republicano. Fair utiliza esta variable para imponer la restricción de que el efecto de un republicano en la Casa Blanca tiene la misma magnitud, pero con signo opuesto, que el de un demócrata en la Casa Blanca. Se trata de una restricción natural ya que, por definición, las participaciones de los partidos deben sumar uno. También conserva dos grados de libertad, lo que es importante con tan pocas observaciones. De igual modo, la variable *incum* se define como 1 si el mandatario titular que está buscando la reelección es un demócrata, -1 si es republicano, y 0 en otros casos. La variable *gnews* es el número de trimestres durante los primeros 15 (de un total de 16) en la administración actual, en los que el crecimiento trimestral de la producción real per cápita estuvo por encima de 2.9% (a una tasa anual), e *inf* es la tasa de inflación anual promedio de los primeros 15 trimestres de la administración. Vean Fair (1996) para obtener definiciones precisas.

Los economistas están más interesados en los términos de interacción *partyWH·gnews* y *partyWH·inf*. Como *partyWH* es igual a uno cuando un demócrata está en la Casa Blanca, β_3 mide el efecto de las buenas noticias económicas sobre el partido en el poder; se espera que $\beta_3 > 0$. Asimismo, β_4 mide el efecto que la inflación tiene sobre el partido en el poder. Como la inflación durante una administración se considera una mala noticia, esperamos que $\beta_4 < 0$.

La ecuación estimada con ayuda de los datos de FAIR.RAW es

$$[10.23] \quad demvote = .481 - .0435 partyWH + .0544 incum + .0108 partyWH \cdot gnews - \\ \quad \quad \quad (.012) \quad \quad (.0405) \quad \quad (.0234) \quad \quad (.0041)$$

⁸ Ray C. Fair, *Econometrics and Presidential Elections*, Journal of Economic Perspectives—Volume 10, Num 3, Summer 1996.

$$.0077 \text{ partyWH} \cdot \text{inf}$$

$$(.0033)$$

$$n = 20; R^2 = .663; R^2_{aj} = .573$$

Todos los coeficientes, salvo el de *partyWH*, son estadísticamente significativos a un nivel de 5%. Buscar la reelección vale alrededor de 5.4 puntos porcentuales en la participación del voto (recuerden que *demvote* se mide como proporción). Además, la variable de las noticias económicas ejerce un efecto positivo: un trimestre más de buenas noticias vale alrededor de 1.1 puntos porcentuales. La inflación, como se esperaba, tiene un efecto negativo: si la inflación anual promedio es, por decir, dos puntos porcentuales mayor, el partido en el poder pierde cerca de 1.5 puntos porcentuales de los votos en la elección bipartidista.

Se pudo haber utilizado esta ecuación para predecir el resultado de la elección presidencial de 1996 entre Bill Clinton, demócrata, y Bob Dole, republicano (el candidato independiente, Ross Perot se excluye porque la ecuación de Fair es sólo para una elección bipartidista). Dado que Clinton buscaba la reelección, *partyWH* = 1 e *incum* = 1. Para pronosticar el resultado de la elección se necesitan las variables *gnews* e *inf*. Durante los primeros 15 trimestres del mandato de Clinton, el PIB real per cápita excedió 2.9% en tres ocasiones, de modo que *gnews* = 3. Además, utilizando el deflactor del PIB reportado en la tabla B-4 del ERP, de 1997, la tasa de inflación media anual (calculada usando la fórmula de Fair) del cuarto trimestre de 1991 al tercer trimestre de 1996 fue de 3.019. Al sustituir estos valores en la ecuación (10.23) se obtiene

$$\text{demvote}^o = .481 - .0435 + .0544 + .0108(3) - .0077(3.019) = .5011.$$

Por tanto, en base a la información conocida antes de la elección de noviembre, se pronosticaba que Clinton recibiría una mayoría muy escasa del voto bipartidista: cerca de 50.1%. De hecho, Clinton ganó con más facilidad: su participación en el voto bipartidista fue de 54.65%.

10.5 *Tendencias y Estacionalidad*

Caracterización de la tendencia en las series de tiempo

Muchas series de tiempo económicas tienen una tendencia común de crecimiento a lo largo del tiempo. ***Se debe reconocer que ciertas series contienen una tendencia en el tiempo para hacer inferencias causales con los datos de las series de tiempo.*** Ignorar el hecho de que dos secuencias tienen tendencias en la misma dirección o en direcciones opuestas nos conduce a la conclusión falsa de que los cambios en una variable en realidad son ocasionados por modificaciones en otra variable. En muchos casos, parece que dos procesos de series de tiempo están correlacionados sólo porque ambos tienen una tendencia en el tiempo por razones relacionadas con otros factores inobservables.

La figura 10.2 contiene una gráfica de productividad laboral (producción por hora de trabajo) para los años de 1947 a 1987 en Estados Unidos. Esta serie muestra una clara tendencia hacia arriba, que refleja el hecho de que los trabajadores se han vuelto más productivos con el tiempo.

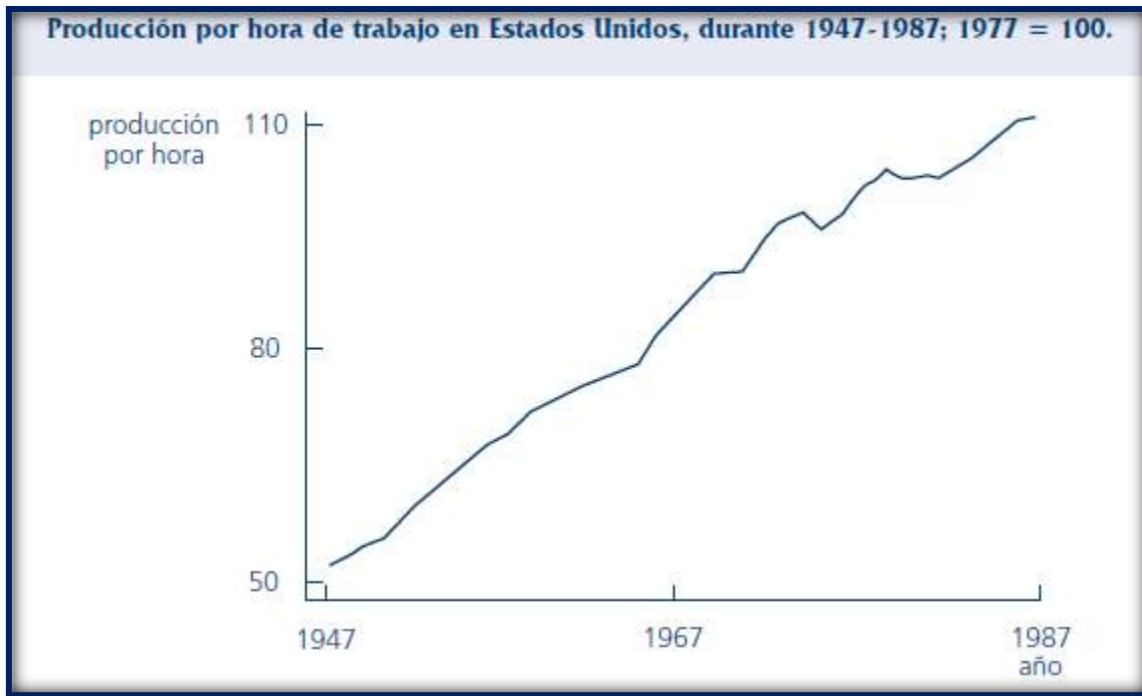


Figura 10.2

Otras series, al menos sobre ciertos periodos, tienen claras tendencias hacia abajo. Como las tendencias positivas son más comunes, en ellas se centrará el siguiente análisis.

¿Qué clase de modelos estadísticos representan de manera adecuada el comportamiento de la tendencia? Una formulación popular es escribir la serie $\{y_t\}$ como

$$[10.24] \quad y_t = \alpha_0 + \alpha_1 t + e_t, \quad t = 1, 2, \dots,$$

en donde, en el caso más sencillo, $\{e_t\}$ es una secuencia independiente e idénticamente distribuida (i.i.d.) con $E(e_t) = 0$ y $Var(e_t) = \sigma_e^2$. Adviertan cómo el parámetro α_1 multiplica al tiempo, t , lo que da como resultado una tendencia lineal en el tiempo. La interpretación de α_1 en la ecuación (10.24) es sencilla: si se mantienen fijos todos los demás factores (los de e_t), α_1 mide el cambio en y_t de un periodo al siguiente debido al transcurso del tiempo: cuando $\Delta e_t = 0$,

$$\Delta y_t = y_t - y_{t-1} = \alpha_1.$$

Otra manera de considerar una secuencia que tiene una **tendencia lineal** en el tiempo es que su valor promedio es una función lineal del tiempo:

$$[10.25] \quad E(y_t) = \alpha_0 + \alpha_1 t.$$

Si $\alpha_1 > 0$, entonces, en promedio, y_t está creciendo en el tiempo y por consiguiente tiene una tendencia hacia arriba. Si $\alpha_1 < 0$, entonces y_t tiene una tendencia hacia abajo. Los valores de y_t no caen de manera exacta sobre la línea en la ecuación (10.25) debido a la aleatoriedad, pero los valores esperados sí están sobre ella. A diferencia de la media, la varianza de y_t es constante en el tiempo: $Var(y_t) = Var(e_t) = \sigma_e^2$.

Si $\{e_t\}$ es una secuencia i.i.d., entonces $\{y_t\}$ es una secuencia independiente, pero no es idénticamente distribuida. Una caracterización más realista de la tendencia en las se-

ries de tiempo permite que $\{e_t\}$ esté correlacionada a través del tiempo, pero esto no modifica la cualidad de una tendencia lineal en el tiempo. De hecho, lo que es importante para el análisis de regresión, bajo los supuestos del modelo lineal clásico, es que $E(y_t)$ sea lineal en t . Cuando estudien las propiedades de muestras grandes de MCO en el capítulo 11, tendrán que plantearse cuánta correlación temporal se permite en $\{e_t\}$.

Muchas series de tiempo económicas se aproximan mejor por medio de una **tendencia exponencial**, la cual se sigue cuando una serie tiene la misma tasa de crecimiento promedio de un periodo a otro. En la figura 10.3 se trazan los datos de las importaciones nominales anuales de Estados Unidos para los años de 1948 a 1995 (ERP de 1997, tabla B-101).

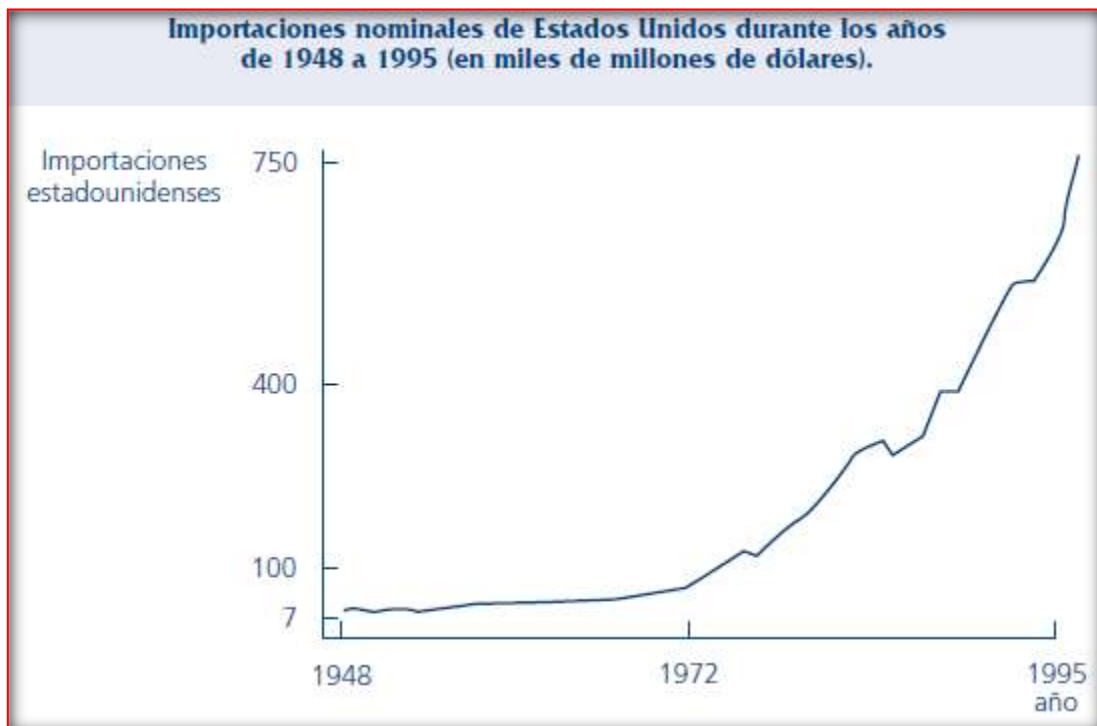


Figura 10.3

En los primeros años, se ve que el cambio en las importaciones anuales es relativamente pequeño, pero aumenta con el transcurso del tiempo. Esto es consistente con una **tasa de crecimiento promedio constante**: el cambio porcentual es casi el mismo en cada periodo.

En la práctica, una tendencia exponencial en una serie de tiempo se consigue modelando el **logaritmo natural de la serie como una tendencia lineal** (suponiendo que $y_t > 0$):

$$[10.26] \quad \log(y_t) = \beta_0 + \beta_1 t + e_t, \quad t = 1, 2, \dots$$

La exponenciación muestra que y_t por sí misma tiene una tendencia exponencial: $y_t = \exp(\beta_0 + \beta_1 t + e_t)$. Como se deseará utilizar series de tiempo de tendencia exponencial en los modelos de regresión lineal, la ecuación (10.26) resulta ser la forma más conveniente para representar estas series.

¿Cómo se interpreta β_t en la ecuación (10.26)? Recuerden que, para cambios pequeños, $\Delta \log(y_t) = \log(y_t) - \log(y_{t-1})$ es aproximadamente el cambio proporcional en y_t :

$$[10.27] \quad \Delta \log(y_t) \approx (y_t - y_{t-1})/y_{t-1}.$$

Al lado derecho de la ecuación (10.27) también se lo llama **tasa de crecimiento de y del periodo $t - 1$ al periodo t** . Para convertir la tasa de crecimiento en un porcentaje, sólo se multiplica por 100. Si y_t cumple con (10.26), entonces, tomando cambios y estableciendo $\Delta e_t = 0$,

$$[10.28] \quad \Delta \log(y_t) = \beta_t, \text{ para toda } t.$$

En otras palabras, β_t es aproximadamente la tasa de crecimiento promedio por periodo de y_t . Por ejemplo, si t denota un año y $\beta_t = .027$, entonces y_t crece alrededor de 2.7% al año en promedio.

Aunque las tendencias lineales y exponenciales son las más comunes, las tendencias en el tiempo pueden ser más complicadas. Por ejemplo, en vez del modelo de tendencia lineal de la ecuación (10.24), se podría tener una **tendencia cuadrática** en el tiempo:

$$[10.29] \quad y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + e_t.$$

Si α_1 y α_2 son positivas, entonces la pendiente de la tendencia está aumentando, como se ve con facilidad al calcular la pendiente aproximada (manteniendo a e_t fijo):

$$[10.30] \quad (\Delta y_t / \Delta t) \approx \alpha_1 + \alpha_2 t.$$

[Si están ustedes familiarizados con el cálculo, reconocerán el lado derecho de la ecuación (10.30) como la derivada de $\alpha_0 + \alpha_1 t + \alpha_2 t^2$ respecto a t .] Si $\alpha_1 > 0$, pero $\alpha_2 < 0$, la tendencia tiene una forma de U invertida. Tal vez ésta no sea una muy buena descripción de ciertas series que muestran tendencia debido a que requieren que una tendencia decreciente siga después de una tendencia creciente. Sin embargo, durante un periodo dado, puede ser una manera flexible de modelar series de tiempo que tienen tendencias más complicadas que las de las ecuaciones (10.24) o (10.26).

Uso de variables con tendencia en el análisis de regresión

En el análisis de regresión resulta muy fácil llevar la cuenta de las variables explicadas o explicativas que muestran tendencia. En primer lugar, nada respecto a las variables con tendencia viola necesariamente los supuestos ST.1 a ST.6 del modelo lineal clásico. Sin embargo, debe procurarse tener en cuenta el hecho de que factores inobservables con tendencia que afectan a y_t podrían estar correlacionados con las variables explicativas. Si se ignora esta posibilidad, podemos encontrar una relación falsa entre y_t y una o más variables explicativas. **El fenómeno de encontrar una relación falsa entre dos o más variables con tendencia, sencillamente porque cada una está creciendo con el tiempo, es un problema de **regresión espuria****. Por fortuna, la adición de una tendencia en el tiempo elimina este problema.

En concreto, consideren un modelo donde dos factores inobservables, x_{t1} y x_{t2} , afectan a y_t . Además, hay factores inobservables que están creciendo o decreciendo de manera sistemática con el tiempo. Un modelo que captura esto es

$$[10.31] \quad y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \beta_3 t + u_t.$$

Este modelo encaja en el esquema de la regresión lineal múltiple con $x_{t3} = t$. Tomar en cuenta la tendencia en esta ecuación reconoce de manera explícita que y_t puede estar creciendo ($\beta_3 > 0$) o decreciendo ($\beta_3 < 0$) en el tiempo por motivos que en esencia no están relacionados con x_{t1} o x_{t2} . Si la ecuación (10.31) satisface los supuestos ST.1, ST.2 y ST.3, entonces al omitir t y hacer la regresión de y_t sobre x_{t1} , x_{t2} por lo común generará estimadores sesgados de β_1 y β_2 : se ha omitido, en efecto, una variable importante, t , de la regresión. Esto es válido en particular si x_{t1} y x_{t2} poseen una tendencia, porque entonces pueden estar muy correlacionadas con t . El siguiente ejemplo muestra cómo la omisión de una tendencia en el tiempo puede dar como resultado una regresión espuria.

Ejemplo 10.7 Inversión y precios de la vivienda

Los datos de HSEINV.RAW son observaciones anuales sobre la inversión y el índice de precios de la vivienda en Estados Unidos de 1947 a 1988. Sea *invpc* la inversión real en vivienda per cápita (en miles de dólares) y *price* un índice del precio de vivienda (igual a 1 en 1982). Una regresión simple en la forma de elasticidad constante, que puede pensarse como una ecuación de oferta para la cantidad de viviendas disponibles, da

$$[10.32] \quad \log(\text{invpc})^o = -.550 + 1.241 \log(\text{price}) \\ (.043) \quad (.382) \\ n = 42, R^2 = .208, R^2_{aj} = .189.$$

La elasticidad de la inversión per cápita respecto al precio es muy grande y significativa; estadísticamente no es diferente de 1. **Deben ser cuidadosos aquí.** Tanto *invpc* como *price* muestran sesgos hacia arriba. En particular, si hacen la regresión de $\log(\text{invpc})$ se obtiene un coeficiente sobre la tendencia igual a .0081 (error estándar = .0018); la regresión de $\log(\text{price})$ sobre t arroja un coeficiente de tendencia igual a .0044 (error estándar .0004). **Aunque los errores estándar en los coeficientes de tendencia no son necesariamente confiables, ya que estas regresiones tienden a contener una correlación serial sustancial, las estimaciones de los coeficientes revelan sesgos hacia arriba.**

Para representar el comportamiento de la tendencia de las variables, se añade una tendencia en el tiempo:

$$[10.33] \quad \log(\text{invpc})^o = -.913 - .381 \log(\text{price}) + .0098 t \\ (.136) \quad (.679) \quad (.0035) \\ n = 42, R^2 = .341, R^2_{aj} = .307.$$

La historia es muy distinta ahora: la elasticidad estimada del precio es negativa y en términos estadísticos no es diferente de cero. La tendencia en el tiempo es estadísticamente significativa y su coeficiente implica un incremento aproximado de 1% anual en *invpc*, en promedio. A partir de este análisis, no se puede concluir que la inversión real per cápita en vivienda esté afectada en lo absoluto por el precio. Hay otros factores,

capturados en la tendencia en el tiempo, que influyen sobre $invpc$, pero no los hemos modelado. **Los resultados en la ecuación (10.32) muestran una relación espuria entre $invpc$ y $price$ debido al hecho de que el precio también tiene una tendencia ascendente en el tiempo.**

En algunos casos, la adición de una tendencia en el tiempo hace que una variable explicativa clave resulte *más* significativa. Esto puede suceder si las variables dependiente e independiente tienen tendencias distintas (por ejemplo, una hacia arriba y otra hacia abajo), pero el movimiento en la variable independiente *alrededor* de su línea de tendencia provoca un movimiento de la variable dependiente alejándose de su línea de tendencia.

Ejemplo 10.8 Ecuación de fertilidad

Si se añade una tendencia lineal en el tiempo a la ecuación (10.18), se obtiene

$$[10.34] \quad gfr^o_t = 111.77 + .279 pe_t - 35.59 ww2_t + .997 pill_t - 1.15 t$$

$$(3.36) (.040) \quad (6.30) \quad (6.626) \quad (.19)$$

$$n = 72, R^2 = .662, R^2_{aj} = .642.$$

El coeficiente de pe es más de tres veces la estimación de la ecuación (10.18) y es mucho más significativo en términos estadísticos. Resulta interesante que $pill$ no sea significativa una vez que se permite una tendencia lineal. Como se ve por la estimación, gfr estaba descendiendo, en promedio, durante este periodo, manteniendo constantes los demás factores.

Como la tasa general de fertilidad mostró tendencias tanto hacia arriba como hacia abajo durante el periodo de 1913 a 1984, se puede ver la intensidad que tiene el efecto estimado de pe cuando se utiliza una tendencia cuadrática:

$$[10.35] \quad gfr^o_t = 124.09 + .348 pe_t - 35.88 ww2_t + -10.12 pill_t - 2.53 t + .0196 t^2$$

$$(4.36) (.040) \quad (5.71) \quad (6.34) \quad (.39) \quad (.0050)$$

$$n = 72, R^2 = .727, R^2_{aj} = .706.$$

El coeficiente de pe es incluso más grande y más significativo en términos estadísticos. Ahora, $pill$ tiene el efecto negativo esperado y es poco significativa, y ambos términos de tendencia son estadísticamente significativos. La tendencia cuadrática es una forma flexible de representar el comportamiento inusual de la tendencia de gfr .

Tal vez se pregunten respecto al ejemplo 10.8: ¿por qué detenerse en una tendencia cuadrática? Nada nos impide agregar, por ejemplo, t^3 como variable independiente y, de hecho, esto podría justificarse (vean el ejercicio para computadora C10.6). Pero debemos procurar no entusiasmarnos cuando se incluyan términos de tendencia en un modelo. Son recomendables las tendencias más o menos simples que registran los amplios movimientos en la variable dependiente que no se explican por medio de las variables independientes en el modelo. **Si se incluyen los suficientes términos polinomiales en t , entonces se puede rastrear muy bien cualquier serie. Pero esto nos ofrece poca ayuda para determinar qué variables explicativas influyen sobre y_t** [y agrego, es posible que se enfrente un severo problema de multicolinealidad.]

Interpretación de las regresiones con tendencia en el tiempo mediante la eliminación de la tendencia

Incluir una tendencia en el tiempo en un modelo de regresión permite hacer una interpretación atractiva en términos de la eliminación de la tendencia en las series de datos originales, antes de utilizarlas en el análisis de regresión. Para concretar, el análisis se concentrará en el modelo (10.31), pero las conclusiones son mucho más generales.

Cuando se hace la regresión de y_t sobre x_{t1} , x_{t2} y t , se obtiene la ecuación ajustada

$$[10.36] \quad y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \beta_3 t.$$

Podemos ampliar la interpretación de MCO como descuento de efectos parciales que se cubrieron en el capítulo 3, para mostrar que β_1 y β_2 se obtienen como sigue.

i) Hagan la regresión de cada una de las variables y_t , x_{t1} y x_{t2} sobre una constante y la tendencia en el tiempo t y guarden los residuales, digamos, y'_t , x'_{t1} , x'_{t2} , $t = 1, 2, \dots, n$. Por ejemplo,

$$y'_t = y_t - \alpha_0 - \alpha_1 t.$$

Así, puede pensarse en y'_t como el resultado de haber eliminado la tendencia lineal. Para eliminar la tendencia de y_t , se ha estimado por MCO el modelo

$$y_t = \alpha_0 + \alpha_1 t + e_t$$

los residuales de esta regresión, $e_t = y'_t$, tienen la tendencia en el tiempo eliminada (al menos en la muestra). Una interpretación similar es válida para x'_{t1} y x'_{t2} .

ii) Realicen la regresión de

$$[10.37] \quad y'_t \text{ sobre } x'_{t1} \text{ y } x'_{t2}.$$

(El intercepto no es necesario, pero si se incluye no afecta en nada; se estimará como cero.) Esta regresión produce exactamente β_1 y β_2 de la ecuación (10.36).

Esto significa que las estimaciones de interés primordial, β_1 y β_2 , pueden interpretarse como si provinieran de una regresión sin una tendencia en el tiempo, pero en la que se eliminó primero la tendencia de la variable dependiente y de todas las independientes. La misma conclusión es válida para cualquier número de variables independientes y si la tendencia es cuadrática o de algún otro grado polinomial.

Si t se omite de (10.36), entonces no ocurre la eliminación de la tendencia y parecería que y_t se relaciona con una o más de las x_{tj} sencillamente porque cada una contiene una tendencia; lo vimos en el ejemplo 10.7. Si el término de tendencia es estadísticamente significativo, y los resultados se modifican de manera importante cuando se agrega una tendencia en el tiempo a una regresión, entonces los resultados iniciales sin una tendencia deben tratarse con reserva.

La interpretación de β_1 y β_2 muestra que es una buena idea incluir una tendencia en la regresión si alguna variable independiente tiene tendencia, incluso si y_t no la tiene. Si y_t

no cuenta con una tendencia observable, pero, digamos, x_{ti} está creciendo con el tiempo, entonces la exclusión de una tendencia de la regresión haría parecer como si x_{ti} no tuviera ningún efecto sobre y_t , aun cuando los movimientos de x_{ti} en torno a su tendencia sí influyan en y_t . Esto se captará si se incluye en la regresión.

Ejemplo 10.9 Empleo en Puerto Rico

Cuando se añadió una tendencia lineal a la ecuación (10.17), las estimaciones pasaron a ser

$$\begin{aligned}
 [10.38] \quad \log(\text{prepop}_t) = & -8.70 - .169 \log(\text{mincov}_t) + 1.06 \log(\text{usgnp}_t) - .032 t \\
 & (1.30) \quad (.044) \quad (0.18) \quad (.005) \\
 & n = 38, R^2 = .847, R^2_{aj} = .834.
 \end{aligned}$$

El coeficiente de $\log(\text{usgnp})$ ha cambiado en forma drástica: de $-.012$ a 1.06 , y de ser insignificante a ser muy significativo. El coeficiente del salario mínimo se ha modificado sólo un poco, aunque el error estándar es en forma notable más pequeño, volviendo a $\log(\text{mincov})$ más significativo que antes. La variable prepop_t no muestra una tendencia clara hacia arriba o hacia abajo, pero $\log(\text{usgnp})$ posee una tendencia lineal ascendente. [Una regresión de $\log(\text{usgnp})$ sobre t da una estimación de alrededor de $.03$, de modo que usgnp está creciendo cerca de 3% al año durante el periodo.] Puede considerarse la estimación 1.06 como sigue: cuando usgnp aumenta 1% por encima de su tendencia a largo plazo, prepop se incrementa alrededor de 1.06 por ciento.

Estacionalidad

Si se observa una serie de tiempo a intervalos mensuales o trimestrales (incluso semanales o diarios), ésta puede manifestar **estacionalidad**. Por ejemplo, la venta mensual de viviendas en la región central de Estados Unidos se ve influida sobremanera por el clima. Si bien los patrones climáticos en cierto modo son aleatorios, es seguro que el clima durante enero por lo general, es más inclemente que en junio y, por tanto, la construcción de viviendas nuevas es mayor en junio que en enero. Una manera de modelar este fenómeno es permitir que el valor esperado de la serie, y_t , sea distinto cada mes. En otro ejemplo, las ventas al menudeo en el cuarto trimestre en general son más altas que en los tres trimestres anteriores debido a las vacaciones de Navidad. Una vez más, esto puede ser captado si se permite que las ventas al menudeo sean diferentes en el transcurso de un año. Esto además de permitir la posibilidad de considerar una media con tendencia. Por ejemplo, las ventas al menudeo en el primer trimestre de este año fueron mayores que las del cuarto trimestre de hace 30 años, ya que han crecido a un ritmo constante. No obstante, si se comparan las ventas medias dentro de un año cualquiera, el factor del periodo vacacional suele hacer que las ventas resulten mayores durante el cuarto trimestre.

Aun cuando muchas series de datos mensuales y trimestrales muestran patrones estacionales, no todas lo hacen. Por ejemplo, no existe un patrón estacional observable en los intereses mensuales o en las tasas de inflación. Asimismo, las series que manifiestan patrones estacionales a menudo se ajustan estacionalmente antes de reportarlas para su uso público. Una serie ajustada estacionalmente es aquella a la que, en principio, se le han eliminado los factores estacionales. El ajuste estacional puede realizarse de una gran variedad de maneras, pero un análisis minucioso rebasa el alcance de este libro.

[Puede consultarse Harvey⁹ e Hylleberg¹⁰ para obtener una idea de los tratamientos pormenorizados.]

El ajuste estacional se ha vuelto tan común que en muchos casos no es posible obtener datos no ajustados estacionalmente. El PIB trimestral estadounidense es un ejemplo destacado. En el informe anual *Economic Report of the President*, muchas bases de datos macroeconómicas reportadas mensualmente (al menos en los años más recientes) y aquellas que muestran patrones estacionales se ajustan todas estacionalmente. Las principales fuentes para series de tiempo macroeconómicas, incluida Citibase, también ajustan estacionalmente a muchas de las series. Por esta razón, el alcance para usar nuestro propio ajuste estacional con frecuencia se ve limitado.

En ocasiones, se trabaja con datos que no están ajustados estacionalmente y resulta útil saber con cuáles métodos sencillos se cuenta para tratar la estacionalidad en los modelos de regresión. En general, se puede incluir un conjunto de **variables binarias estacionales** para representar la estacionalidad en la variable dependiente, las independientes o ambas.

El método es sencillo. Imaginen que cuentan con datos mensuales y piensan que los patrones estacionales de un año son más o menos constantes en el tiempo. Por ejemplo, como la Navidad siempre se celebra en la misma época del año, podemos esperar que las ventas al menudeo sean más altas, en promedio, en los últimos meses del año que en los primeros. O bien, como los patrones climáticos son parecidos, en términos generales, a lo largo de los años, la construcción de viviendas nuevas en la región central de Estados Unidos en promedio será mayor durante el verano que en el invierno. Un modelo general para datos mensuales que representa este fenómeno es

$$[10.41] \quad y_t = \beta_0 + \delta_1 feb_t + \delta_2 mar_t + \delta_3 apr_t + \dots + \delta_{11} dec_t + \beta_1 x_{1t} + \dots + \beta_k x_{kt} + u_t,$$

donde feb_t , mar_t , ..., dec_t son variables binarias que indican si el periodo t corresponde al mes adecuado. En esta formulación, enero es el mes base y β_0 es el intercepto de enero. Si no hay estacionalidad en y_t , una vez que las x_{ij} se han controlado, entonces δ_1 a δ_{11} son todas cero. Esto se verifica con facilidad por medio de una prueba F.

Ejemplo 10.11 Efectos de las demandas antidumping

En el ejemplo 10.5, se usaron datos mensuales que no se ajustaron estacionalmente. Por tanto, deben agregarse variables binarias estacionales para asegurarse que ninguna de las condiciones importantes cambie. Podría suceder que los meses previos a la demanda sean meses en que las importaciones son mayores o menores, en promedio, que en los demás meses. Cuando se agregan las 11 variables binarias mensuales como en la ecuación (10.41) y se prueba su significancia conjunta, se obtiene el *valor-p* = .59, y de este modo se prueba que las variables binarias estacionales son conjuntamente insignificantes. Además, no hay un cambio importante en las estimaciones una vez que se toma en cuenta la significancia estadística. Krupp y Pollard (1996) en realidad utilizaron

⁹ Harvey, A. (1990), *The Econometric Analysis of Economic Time Series*. 2d ed. Cambridge, MA: MIT Press.

¹⁰ Hylleberg, S. (1992), *Modelling Seasonality*. Oxford: Oxford University Press.

tres variables binarias para las estaciones (otoño, primavera y verano, con el invierno como estación base), en lugar de un conjunto completo de binarias mensuales; el resultado es esencialmente el mismo.

Si los datos son trimestrales, entonces se incluirían variables binarias para tres de los cuatro trimestres, y la categoría omitida sería el trimestre base. A veces resulta útil que las binarias estacionales interactúen con algunas de las x_{tj} para permitir que el efecto de x_{tj} sobre y_t difiera a lo largo del año.

De la misma manera que incluir una tendencia en el tiempo en una regresión se interpreta como la eliminación de la tendencia inicial de los datos, incluir binarias estacionales en una regresión puede interpretarse como la eliminación de la estacionalidad de los datos. En concreto, consideren la ecuación (10.41) con $k=2$. Los coeficientes de pendiente de MCO β_1^o y β_2^o correspondientes a x_1 y x_2 se obtienen como sigue:

i) Determinen las regresiones de y_t , x_{t1} y x_{t2} sobre una constante y las binarias mensuales, feb_t , mar_t, \dots, dec_t , y guarden los residuales, digamos, y'_t , x'_{t1} y x'_{t2} , para toda $t = 1, 2, \dots, n$. Por ejemplo,

$$y'_t = y_t - \alpha^o - \alpha^o_1 feb_t - \alpha^o_2 mar_t - \dots - \alpha^o_{11} dec_t.$$

Este es un método de eliminación de la estacionalidad de una serie de tiempo mensual. La misma interpretación es válida para x'_{t1} y x'_{t2} .

ii) Realicen la regresión, sin las binarias mensuales, de y'_t sobre x'_{t1} y x'_{t2} como en la ecuación (10.37). Esto da como resultado β'_1 y β'_2 .

En algunos casos, si y_t tiene una estacionalidad pronunciada, una mejor medida de la bondad de ajuste es una R-cuadrada basada en la y_t sin estacionalidad. Esto compensa cualquier efecto estacional que no explique la x_{tj} .¹¹

Las series de tiempo que muestran patrones estacionales también pueden tener tendencia, en cuyo caso debe estimarse un modelo de regresión con una tendencia en el tiempo y variables binarias estacionales. Las regresiones pueden entonces interpretarse como regresiones que utilizan series sin tendencia ni estacionalidad. Los estadísticos de la bondad de ajuste se estudian en el documento mencionado en la nota previa: en particular, se eliminan la tendencia y la estacionalidad de y_t al hacer la regresión sobre una tendencia en el tiempo y binarias estacionales, antes de calcular la R-cuadrada.

¹¹ Los ajustes de grados de libertad específicos se estudian en Wooldridge, J. M. (1991), *A Note on Computing R-Squared and Adjusted R-Squared for Trending and Seasonal Data*, *Economics Letters* 36.