

## **PUBLICACIONES DE 4<sup>er</sup> CURSO**

**Curso: 4º**

**Grado: Economía**

**Asignatura: ECONOMETRÍA III**

**TRANSPARENCIAS PARTE 3: MODELOS DE SRIES TEMPORALES**

**TEMA 6: ANÁLISIS UNIVARIANTE**

**Profesores: Antonio Aznar y M<sup>a</sup> Isabel Ayuda**

**Departamento de ANÁLISIS ECONÓMICO**

**Curso Académico 2015/16**



**Facultad de  
Economía y Empresa  
Universidad Zaragoza**

## **PARTE 3: MODELOS DE SERIES TEMPORALES**

### **Tema 6. Análisis Univariante**

#### **Indice**

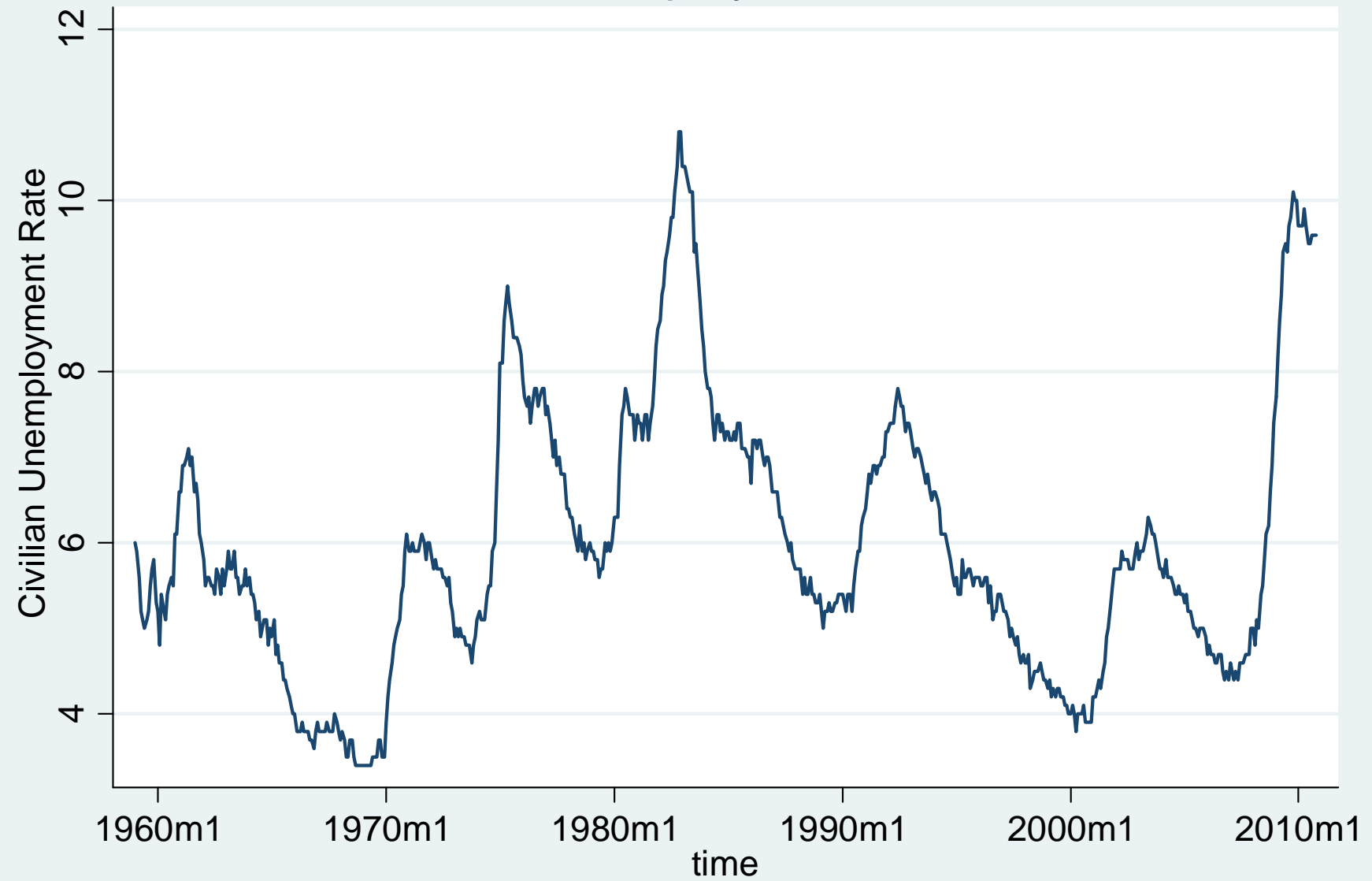
1. Datos de Series Temporales: ¿Qué es diferente?
2. Tendencias Deterministas y Estocásticas.
3. Contraste de Raíz Unitaria: Dickey-Fuller
4. Enfoque Box-Jenkins

# **1. Datos de Series Temporales: ¿Qué es diferente?**

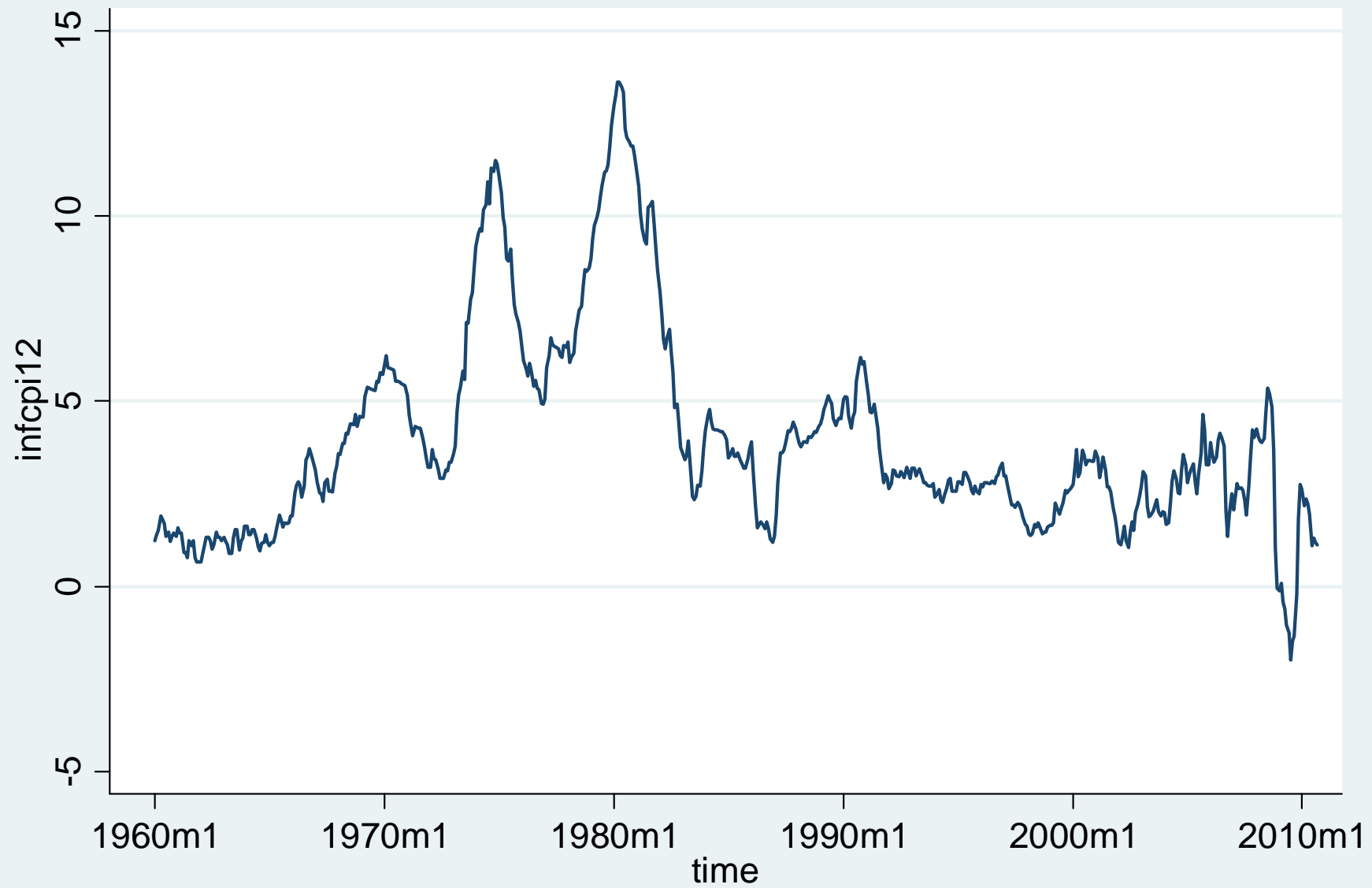
**Los datos de series temporales** son datos recogidos para una unidad observacional en diferentes periodos temporales.

- Consumo agregado y PNB para un país (Por ejemplo, 20 años de observaciones trimestrales = 80 observaciones)
- Yen/\$, libra/\$ and Euro/\$ tipos de cambio (datos diarios para un año = 365 observaciones)
- Consumo de cigarros per cápita en Aragón, para una serie de años.

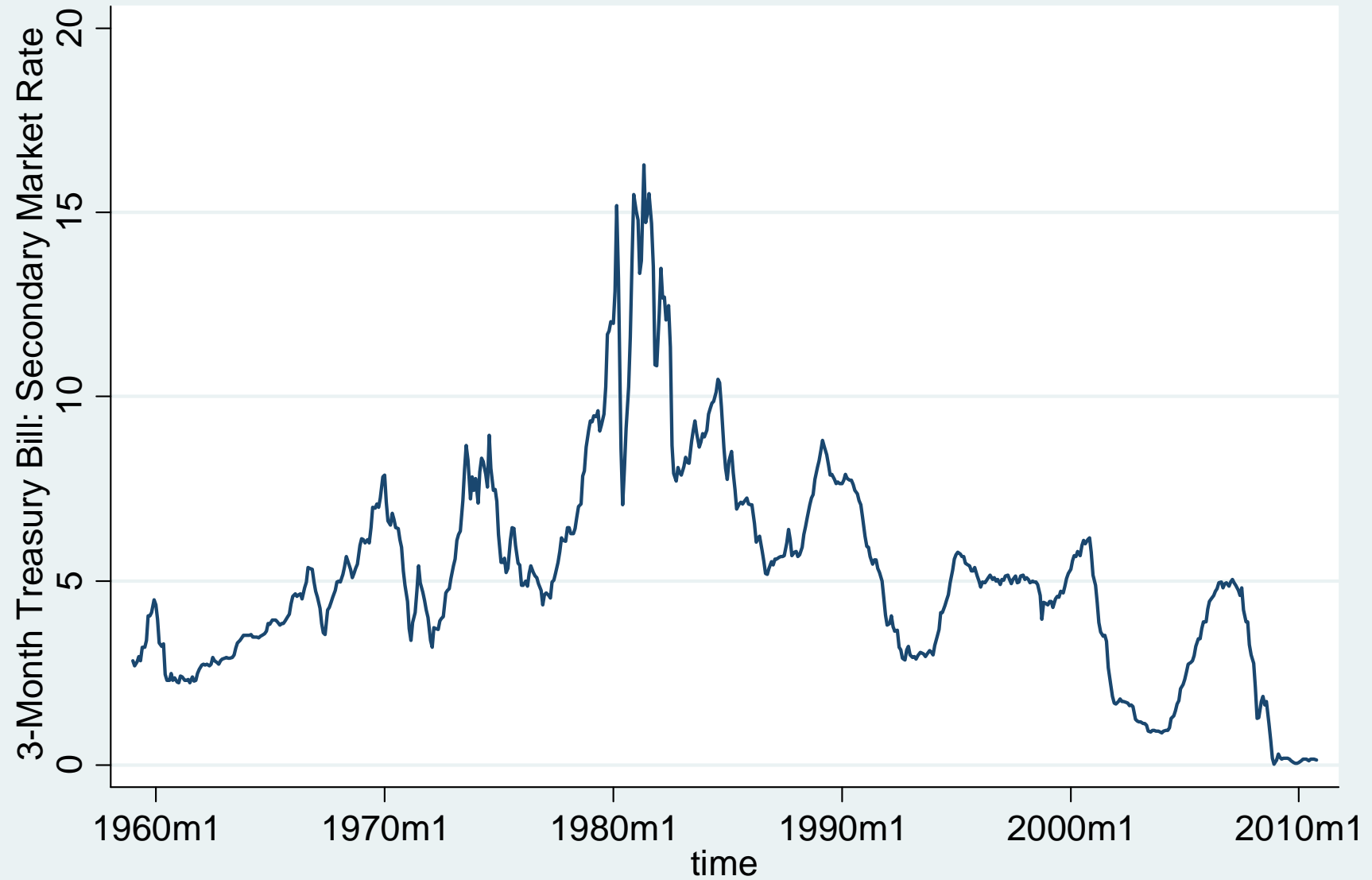
## Civilian unemployment rate, U.S.



## 12-month inflation rate, CPI



## Interest rate on 90-day Treasury bills

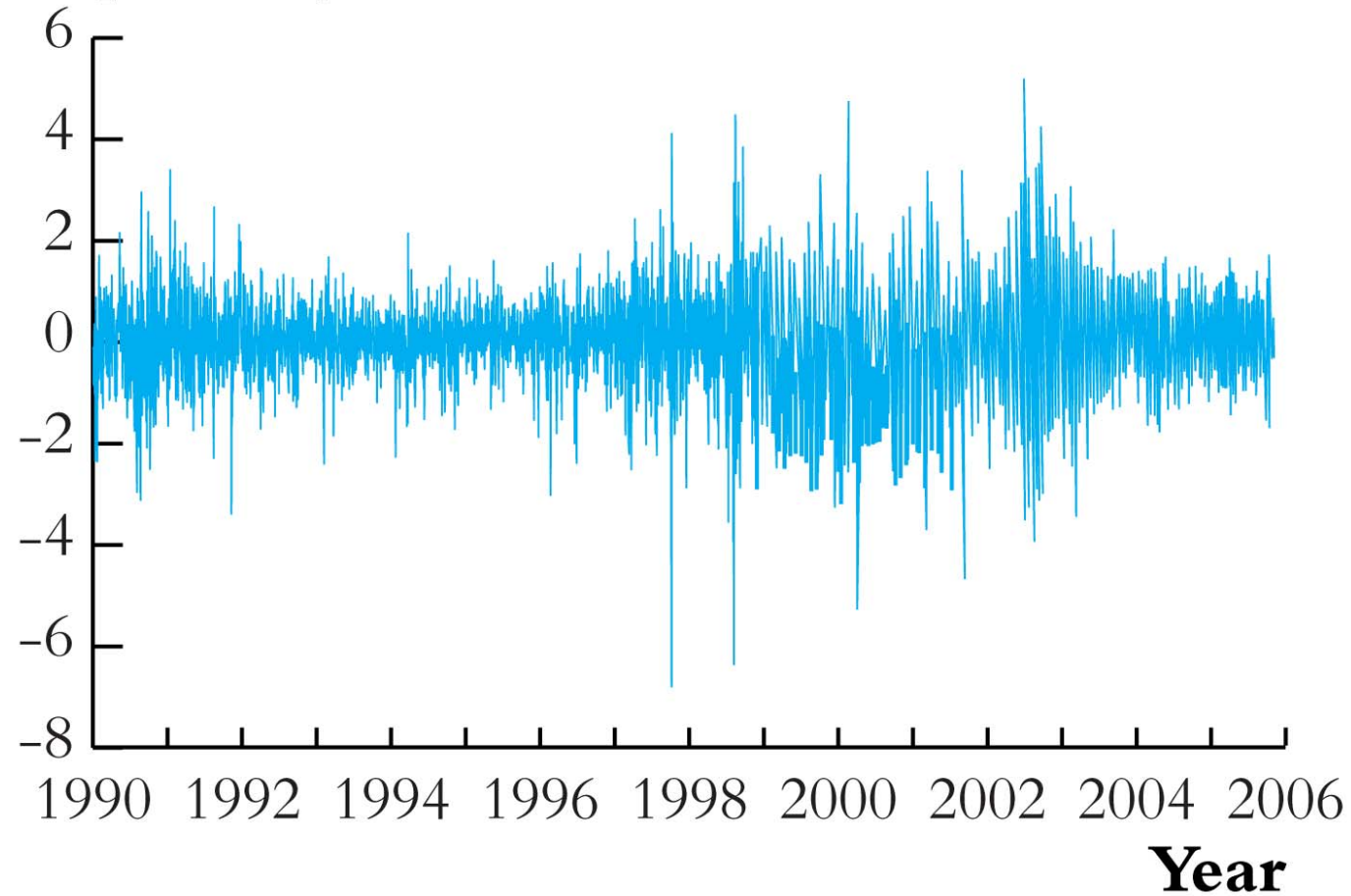


## Yield on 10-year Treasury bonds



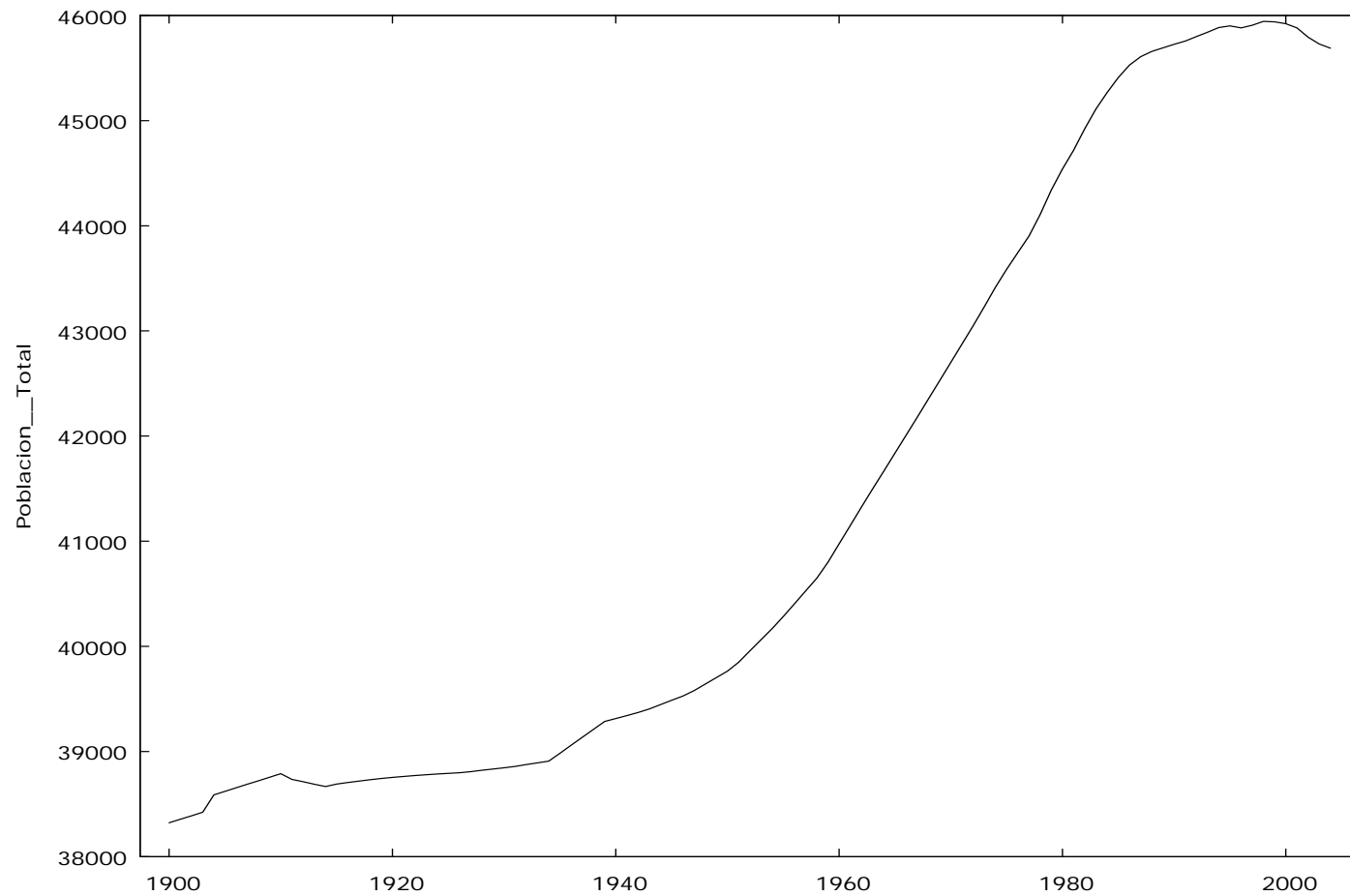
## Serie finanziaria diaria de USA

Percent per Day

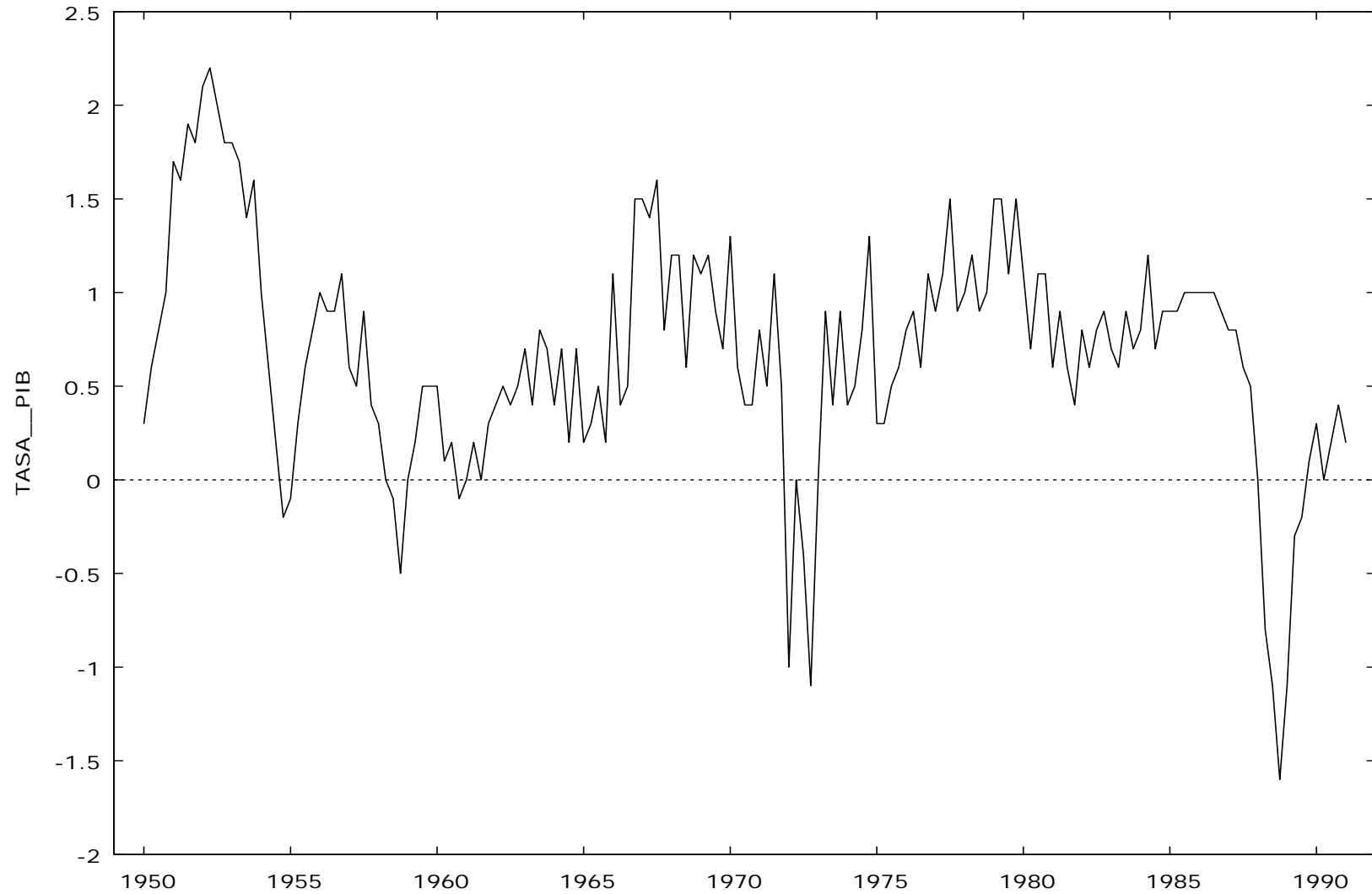


**(d)** Percentage Changes in Daily Values of the NYSE Composite Stock Index





## Total población española(1900/2004)



## Tasa de variación del PIB español.

## Los datos de Series Temporales generan nuevas demandas técnicas

- Retardos Temporales
- Correlación en el tiempo (correlación *serial* , *autocorrelación*)
- Cálculo de los errores estándar cuando los errores están serialmente correlados.

# La Correlación Serial

## A. Notación

- $Y_t$  = valor de  $Y$  en el periodo  $t$ .
- Conjunto de datos:  $\{Y_1, \dots, Y_T\}$  son  $T$  observaciones temporales de la variable  $Y$
- Consideramos observaciones consecutivas, igualmente espaciadas (por ejemplo, mensuales de 1960 a 1999 sin que falte ningún mes)

## B. Retardos, Primeras diferencias y tasas de crecimiento

- El primer retardo de una serie temporal  $Y_t$  es  $Y_{t-1}$
- La primera diferencia de una serie es  $\Delta Y_t = Y_t - Y_{t-1}$
- La primera diferencia del logaritmo es  
$$\Delta \ln(Y_t) = \ln(Y_t) - \ln(Y_{t-1})$$
- El cambio porcentual de una serie  $Y_t$  entre los periodos  $t-1$  y  $t$  es aproximadamente  $100\Delta \ln(Y_t)$  en donde la aproximación es más exacta cuando el cambio porcentual es pequeño.

Ejemplo: Tasa de inflación trimestral (U.S.A.)

*CPI = Consumer Price Index (Indice de precios del consumo)*

- CPI en el primer trimestre de 2004 (2004:I) = 186.57
- CPI en el Segundo trimestre de 2004 (2004:II) = 188.60
- Cambio porcentual en el CPI, 2004:I to 2004:II

$$= 100 \times \left( \frac{188.60 - 186.57}{186.57} \right) = 100 \times \left( \frac{2.03}{186.57} \right) = 1.088\%$$

- Cambio porcentual en CPI, 2004:I to 2004:II, anualizado =  
 $4 \times 1.088 = 4.359\% \approx 4.4\%$  (cambio porcentual anual)
- Como los tipos de interés, las tasas de inflación siempre se dan anualizadas. Simple convención.
- Utilizando la aproximación logarítmica tenemos  
 $4 \times 100 \times [\log(188.60) - \log(186.57)] = 4.329\%$

## C. Autocorrelacion o correlación serial

La correlación de una serie con su propio retardo se llama serial o correlación.

- La primera **autocovarianza** de  $Y_t$  es  $\text{cov}(Y_t, Y_{t-1})$
- La primera **autocorrelacion** de  $Y_t$  es  $\text{corr}(Y_t, Y_{t-1})$
- Así

$$\text{corr}(Y_t, Y_{t-1}) = \frac{\text{cov}(Y_t, Y_{t-1})}{\sqrt{\text{var}(Y_t) \text{var}(Y_{t-1})}} = \rho_1$$

- Estas son correlaciones poblacionales – describen la distribución conjunta poblacional de  $(Y_t, Y_{t-1})$

## Autocorrelaciones muestrales

La  $j$ -ésima autocorrelación muestral es la estimación de la correspondiente autocorrelación poblacional:

$$\hat{\rho}_j = \frac{\widehat{\text{cov}}(Y_t, Y_{t-j})}{\widehat{\text{var}}(Y_t)}$$

donde

$$\widehat{\text{cov}}(Y_t, Y_{t-j}) = \frac{1}{T} \sum_{t=j+1}^T (Y_t - \bar{Y}_{j+1,T})(Y_{t-j} - \bar{Y}_{1,T-j})$$

donde  $\bar{Y}_{j+1,T}$  es la media muestral de  $Y_t$  calculada para las observaciones  $t = j+1, \dots, T$ . NOTAR:

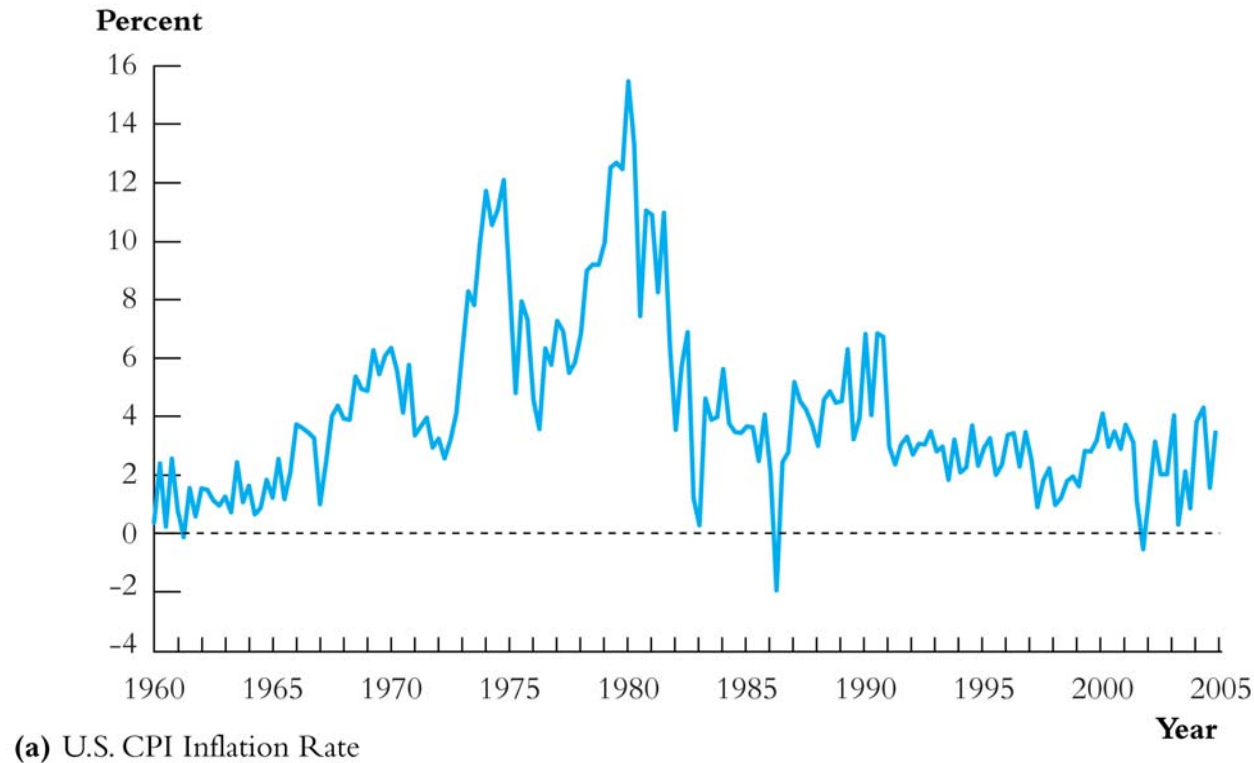
- La suma es de  $j+1$  a  $T$ . (¿por qué?)
- El divisor es  $T$  y no  $T-j$ . Es una convención



## Ejemplo: Inflación USA

**Tabla 4.2 Primeras 4 autocorrelaciones muestrales de la tasa de inflación y de su variación, 1960:I-2004:IV**

retardos	Tasa de inflación	variación
1	0,84	-0,26
2	0,76	-0,25
3	0,76	0,29
4	0,67	-0,06



- La tasa de inflación está serialmente muy correlacionada ( $\rho_1 = .84$ ). El trimestre anterior influye mucho sobre el siguiente.
- En el gráfico ondas a largo plazo. Pero no excluida la sorpresa.

## **Estacionariedad**

La estacionariedad dice que la historia es relevante. La estacionariedad es un requisito clave para la validez externa de la regresión de series temporales.

### **CC 14.5 Estacionariedad**

Una serie temporal es estacionaria si su distribución de probabilidad no cambia en el tiempo; es decir, si la distribución conjunta de  $T$  elementos del proceso que comienzan en el periodo  $s$  no depende de  $s$ . La definición se extiende a dos o más series temporales, diciendo que son conjuntamente estacionarias. La estacionariedad requiere que el futuro sea como el pasado, al menos en probabilidad.

Por lo tanto, una variable es estacionaria cuando la pauta de comportamiento temporal que sigue no está influida por el paso del tiempo. Hay dos sentidos de estacionariedad: en sentido fuerte y en sentido débil. Decimos que una variable es estacionaria en **sentido fuerte** cuando la distribución de probabilidad de cualquier conjunto del mismo número de elementos sucesivos del proceso no depende del periodo al que corresponde el primer elemento del conjunto. Si  $f(\cdot)$  denota la función de probabilidad entonces la estacionariedad en sentido fuerte implica que,

$$\begin{aligned} f(y_1) &= f(y_2) = \dots = f(y_t) \\ f(y_1, y_2) &= f(y_4, y_5) = \dots = f(y_t, y_{t+1}) \\ f(y_1, y_2, \dots, y_T) &= f(y_{1+s}, y_{2+s}, \dots, y_{T+s}) \end{aligned}$$

Decimos que una variable es estacionaria en **sentido débil** cuando se cumple lo siguiente

$$E y_t = \mu \quad \forall t$$

$$Var(y_t) = \sigma^2 \quad \forall t$$

$$Cov(y_t, y_s) = \gamma_{t-s} \quad \forall t, s$$

Lo que significa esta caracterización de la estacionariedad es que todos los elementos del proceso, con independencia del periodo temporal en que nos encontremos, giran en torno a un mismo valor, con una variación en torno a ese valor medio que no cambia con el tiempo y con una dependencia temporal que tampoco cambia.

La implicación de las propiedades asociadas con la estacionariedad débil es que cualquier elemento del proceso tiene momentos de primer y de segundo orden que no dependen del tiempo. Podemos decir que  $y_t$  es  $O_p(1)$  para indicar que cualquier elemento del proceso tiene una distribución de probabilidad bien definida y no precisa de ninguna normalización. Otras implicaciones se derivan para combinaciones de los elementos del proceso. Por ejemplo, el

sumatorio de  $T$  elementos del proceso no tiene un límite bien definido pero si se divide por  $\sqrt{T}$  entonces, por el teorema central del límite, sabemos que la expresión normalizada converge a una distribución de probabilidad bien definida. Podemos decir que  $\sum_1^T y_t$  es  $O_p(T^{\frac{1}{2}})$  o, alternativamente,

$$\frac{1}{T^{1/2}} \sum_1^T y_t \xrightarrow[T \rightarrow \infty]{d} N(0, \sigma_y^2) \quad (1)$$

El mensaje importante es que algunas expresiones necesitan ser normalizadas para que cuando la evidencia empírica sea grande se obtenga una distribución con momentos finitos.

Otra expresión que tiene interés es la suma de los cuadrados

de los elementos del proceso,  $\sum_1^T y_t^2$ . Claramente se ve que

este sumatorio se va a infinito cuando el tamaño muestral crece por lo que es necesaria la normalización. Es fácil demostrar que si esta suma se divide por  $T$  entonces si que

converge en probabilidad, en este caso no en distribución, a una cantidad finita. Podemos decir que  $\sum_1^T y_t^2$  es  $O(T)$ , o bien que

$$\frac{1}{T} \sum_1^T y_t^2 \xrightarrow[p]{T \rightarrow \infty} q_{yy} \quad (2)$$

Siendo  $q_{yy}$  una constante.

Una característica importante de las variables estacionarias es que los momentos muestrales convergen en probabilidad a los correspondientes momentos poblacionales y esta propiedad es muy importante para derivar las propiedades de estimadores y contrastes. Así, podemos escribir,

$$\frac{1}{T} \sum_1^T y_t \xrightarrow[p]{T \rightarrow \infty} \mu$$

$$\frac{1}{T} \sum_1^T (y_t - \bar{y})^2 \xrightarrow[p]{T \rightarrow \infty} \sigma_y^2$$



## 2. Tendencias deterministas y estocásticas

### No estacionariedad I: Tendencias (Trends)

Hasta el momento hemos asumido que los datos son estacionarios, es decir, que la distribución de  $(Y_{s+1}, \dots, Y_{s+T})$  no depende de  $s$ .

Si la estacionariedad no se mantiene, decimos que la serie es no estacionaria.

Dos tipos importantes de no estacionariedad son:

- Tendencias (SW Section 14.6)
- Rupturas estructurales (modelo inestable) (SW Section 14.7)

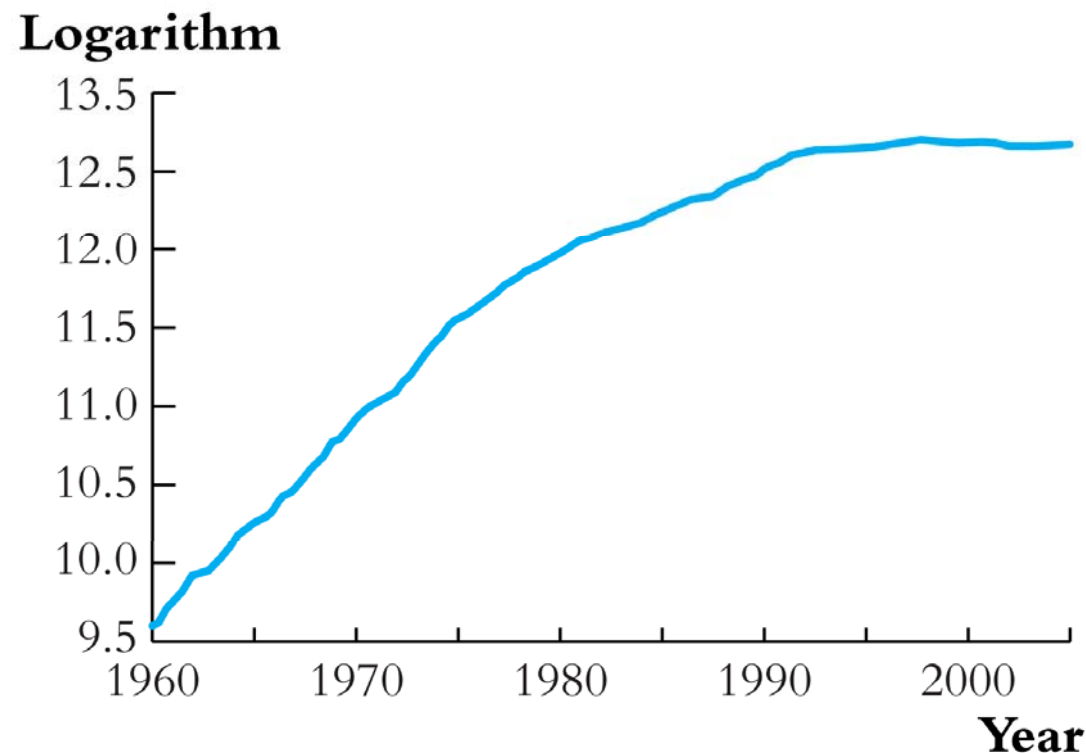
## **Nociones generales sobre las tendencias:**

- A. ¿Qué es una tendencia?
- B. Tendencias deterministas y estocásticas (aleatorias)
- C. ¿Qué problemas son causados por las tendencias?
- D. ¿Cómo detectar las tendencias estocásticas (contrastes estadísticos)?
- E. ¿Cómo solucionar los problemas surgidos con las tendencias?

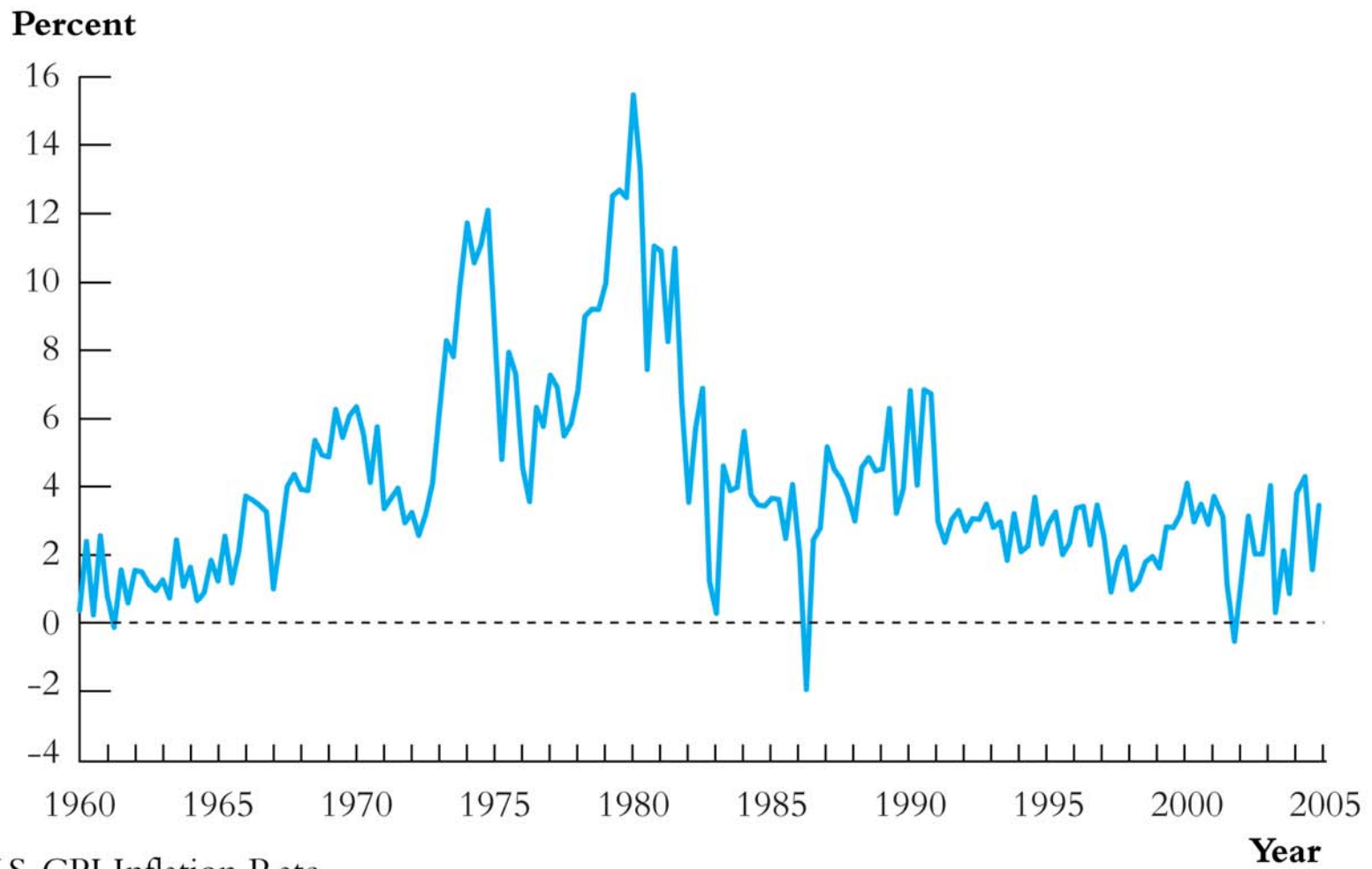
## A. ¿Qué es una tendencia?

Una tendencia es un movimiento persistente, a largo plazo en los datos. Las tendencias no necesariamente son líneas rectas!

*¿Cuál de estas series tiene una tendencia?*

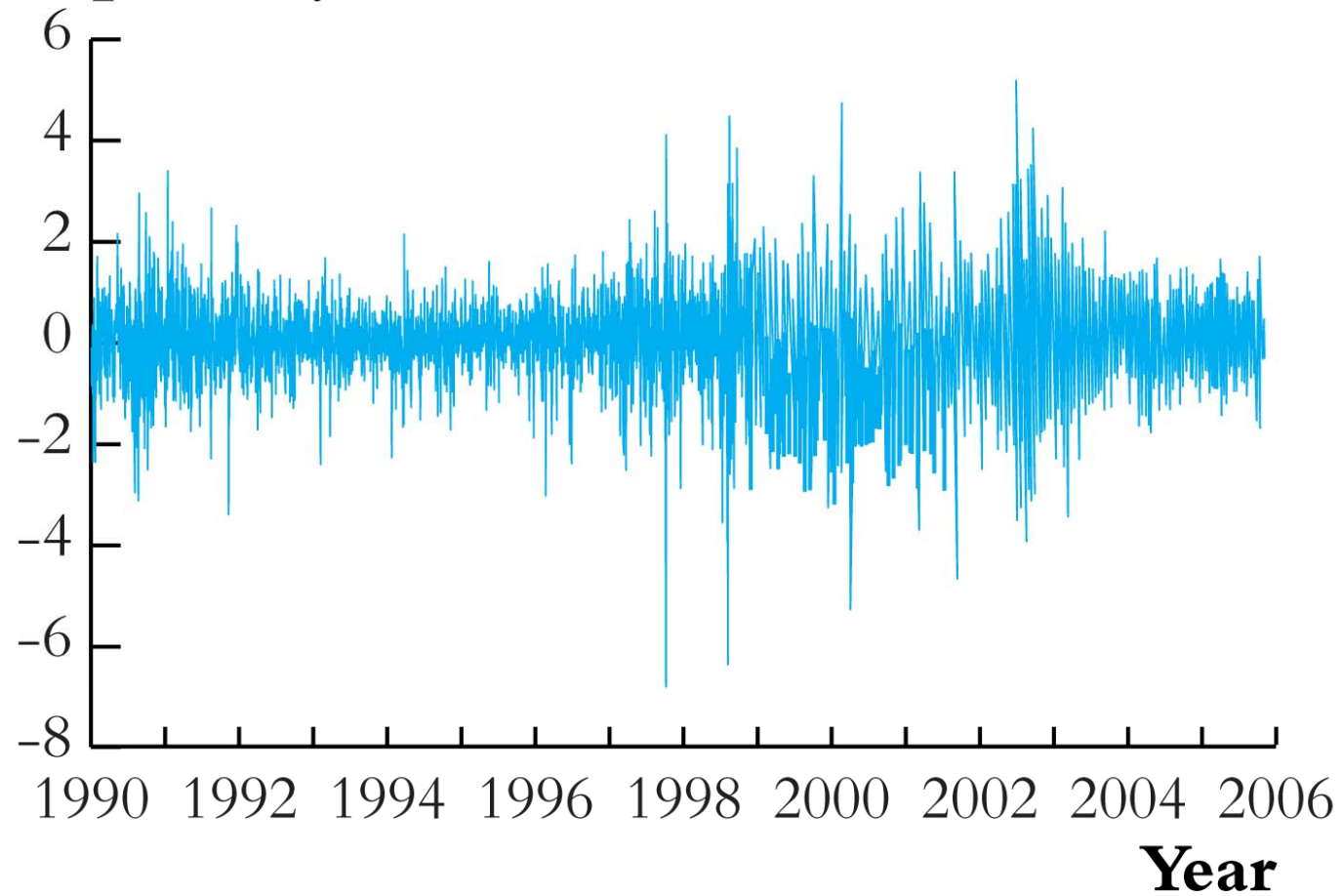


(c) Logarithm of GDP in Japan



**(a)** U.S. CPI Inflation Rate

## Percent per Day



**(d)** Percentage Changes in Daily Values of the NYSE Composite Stock Index

# Concepto de tendencia

Las tres series:

- El log del PNB de Japón claramente tiene una tendencia a largo plazo – no necesariamente una línea recta, sino una tendencia lentamente creciente –rápido crecimiento entre 1960s y 1970s, menor durante los 1980s, estancamiento en los 1990s/2000s.
- La inflación tiene largas oscilaciones, periodos en los que persistentemente está en niveles altos ('70s/primeros'80s) y periodos consecutivos en los que permanece en niveles bajos. Puede que tenga una tendencia pero el diagnóstico no es fácil.
- Los cambios diarios del NYSE no parecen tener una tendencia. Hay periodos con alta volatilidad pero eso no es una tendencia.

## B. Tendencias determinísticas y Estocásticas

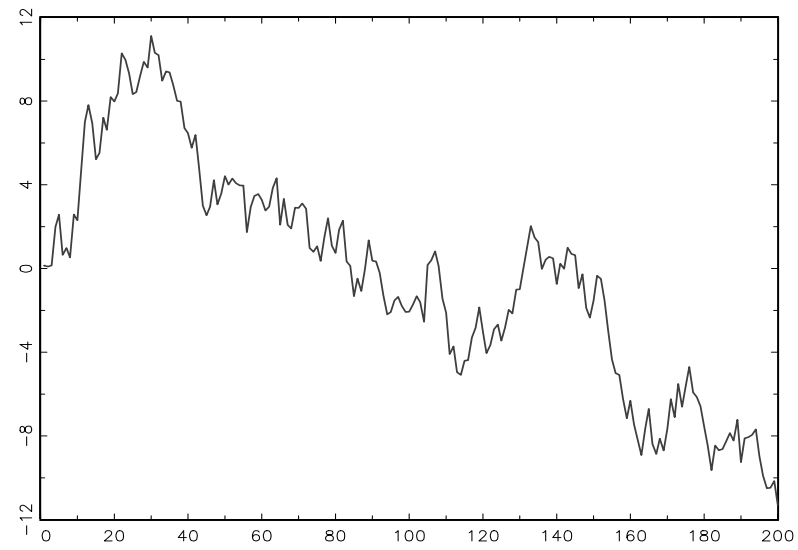
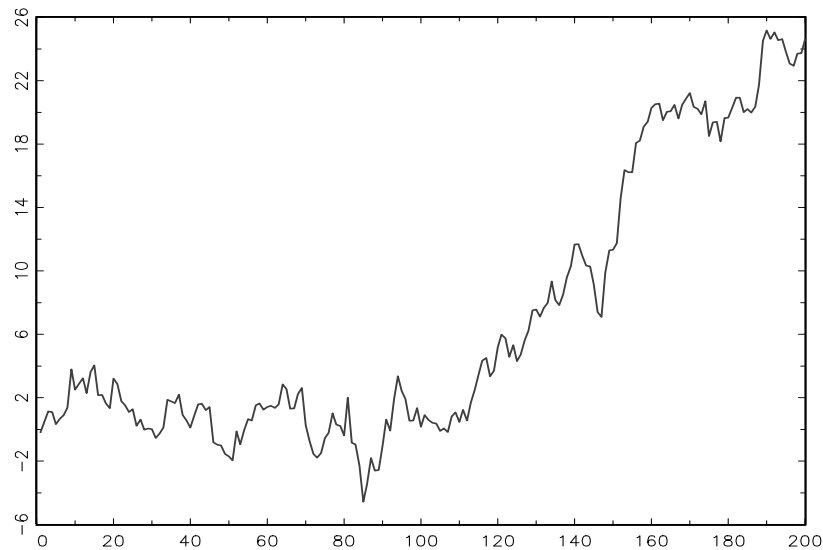
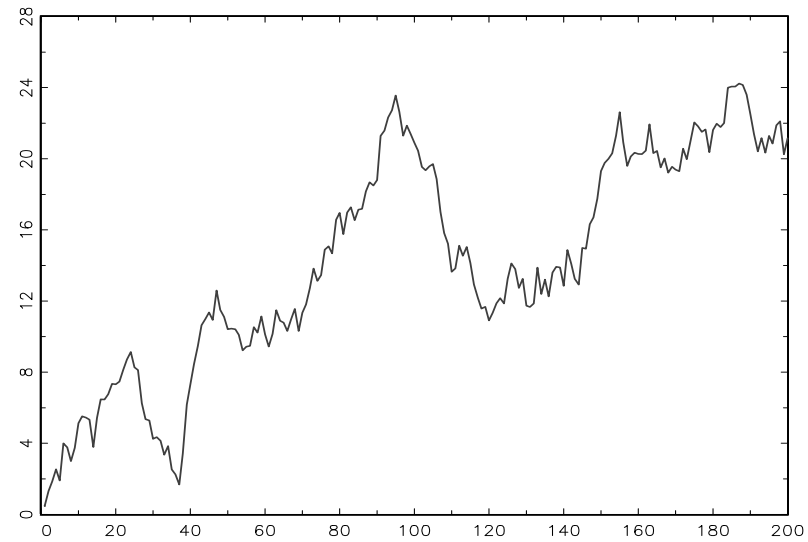
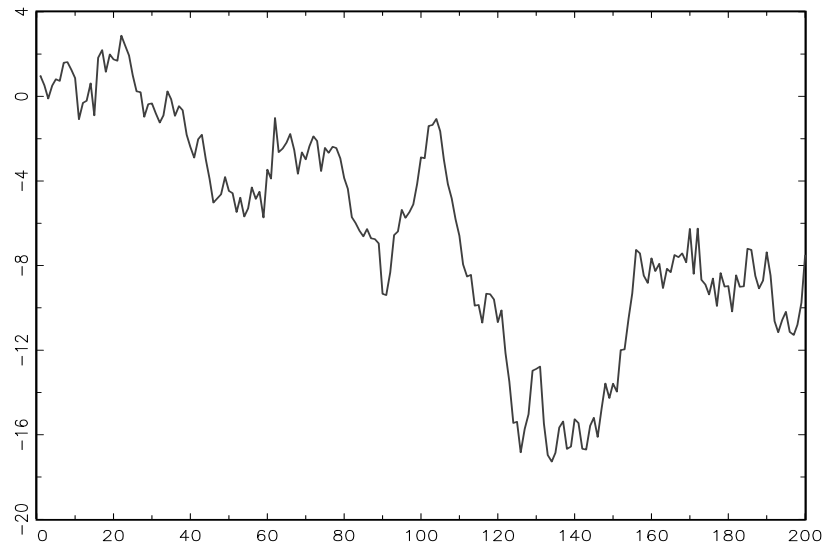
Hemos dicho que una tendencia es un movimiento a largo plazo en los datos.

- Una **tendencia determinista** es una función no aleatoria del tiempo (e.g.  $y_t = t$ , or  $y_t = t^2$ ).
- Una **tendencia estocástica** es aleatoria y cambia con el tiempo.
- Un ejemplo importante de tendencia estocástica es un **paseo aleatorio**:

$Y_t = Y_{t-1} + u_t$ , en donde  $u_t$  no está seriamente correlacionado

Si  $Y_t$  sigue un paseo aleatorio, entonces el valor de  $Y$  de mañana es el valor de  $Y$  de hoy, más una perturbación que no se puede predecir.

Cuatro paseos aleatorios generados artificialmente,  $T = 200$ :



*¿Cómo generas un paseo aleatorio en el ordenador?*



## No estacionariedad

Decimos que una variable es no estacionaria cuando la pauta de comportamiento temporal que sigue depende de la variable tiempo. Esta dependencia puede afectar solo a la media, solo a la varianza o a las dos. Podemos distinguir tres tipos de variables no estacionarias.

### Variable solo con tendencia determinista

El modelo sería

$$y_t = \gamma_0 + \gamma_1 t + u_t$$

Se supone que  $u_t$  es una variable estacionaria con media cero. Los momentos de esta variable son los siguientes,

$$E y_t = \gamma_0 + \gamma_1 t$$

$$Var(y_t) = Var(u_t) = \sigma^2$$

$$Cov(y_t, y_s) = Cov(u_t, u_s) = \gamma_{t-s}$$

Se ve como, para esta variable, la dependencia temporal se manifiesta solo en la media.

## Variable sólo con tendencia estocástica

En este caso el modelo es

$$y_t = y_{t-1} + u_t = y_0 + \sum_1^t u_i$$

Por lo que los momentos serán

$$Ey_t = y_0$$

$$Var(y_t) = t\sigma^2$$

$$Cov(y_t, y_s) = s\sigma^2$$

En este caso, la dependencia temporal se manifiesta en la varianza y en la covarianza.

## Variable con las dos tendencias

El modelo es,

$$y_t = \delta + y_{t-1} + u_t = \delta t + y_0 + \sum_1^t u_i$$

Los momentos serán

$$E y_t = \gamma_0 + \gamma_1 t$$

$$Var(y_t) = t\sigma^2$$

$$Cov(y_t, y_s) = s\sigma^2$$

La dependencia temporal afecta tanto a la media como a la varianza y a las covarianzas.

Para analizar las implicaciones de la no estacionariedad consideremos la variable que solo tiene tendencia estocástica.

Lo primero que observamos es que, conforme el tamaño muestral se hace grande,  $T \rightarrow \infty$ , entonces  $y_\infty$  no tiene una distribución bien definida con momentos finitos porque, como ya hemos visto,

$$Var(y_t) = t\sigma^2 \xrightarrow{T \rightarrow \infty} \infty$$

Tenemos que normalizar para neutralizar la influencia del tiempo. Primero, tener en cuenta que  $t=rT$  con  $0 \leq r \leq 1$ , de forma que la varianza puede escribirse como

$$Var(y_t) = rT\sigma^2$$

A partir de esta expresión se puede pensar en una normalización dividiendo la variable por la raíz cuadrada del tamaño muestral de tal forma que,

$$Var(\frac{y_t}{T^{1/2}}) = r\sigma^2$$

Hemos neutralizado el efecto del tiempo pero no hemos llegado exactamente a un marco estacionario porque cada elemento del proceso normalizado tiene una varianza diferente, aunque finita. Recordar que, en el marco estacionario, todos los elementos tenían la misma varianza. Por eso, las convergencias que se van a derivar para el proceso normalizado van a ser diferentes de las obtenidas en el caso estacionario. En concreto, para la variable no estacionaria en varianza normalizada la convergencia es

$$\frac{y_t}{T^{1/2}} \xrightarrow[T \rightarrow \infty]{d} N(0, r\sigma^2) \cong \sigma W(r)$$

$W(r)$  es una  $N(0,r)$  que se llama variable Browniana. En lo que respecta al sumatorio, hay que tener en cuenta que

$$\sum y_t = T^{1/2} T \frac{1}{T} \sum \frac{y_t}{T^{1/2}}$$

Por lo tanto, podemos decir que el sumatorio de los elementos del proceso es  $O_p(T^{3/2})$ . ¿Tendrá este sumatorio el mismo límite que el obtenido en el caso estacionario? No, porque estamos sumando variables con diferente varianza. El límite al que se llega es una suma de variables en el espacio continuo. El resultado es,

$$\frac{1}{T^{3/2}} \sum y_t \xrightarrow[T \rightarrow \infty]{d} \int_0^1 W(r) dr$$

- De forma similar puede derivarse la distribución a la que tienden otras expresiones funciones del proceso.

### 3. Contraste de Raíz Unitaria: Dickey-Fuller

Decimos que un proceso es integrado de orden cero y lo denotamos como  $I(0)$ , si, después de quitar los elementos deterministas, es un proceso estacionario e invertible.

Decimos que un proceso es integrado de orden 1 y lo denotamos como  $I(1)$  si, después de quitar los elementos deterministas, la primera diferencia del mismo es un proceso  $I(0)$ . También decimos que el proceso tiene una raíz unitaria o, equivalentemente, que tiene una tendencia estocástica.

Decimos que un proceso es integrado de orden  $d$  y lo denotamos como  $I(d)$  si, después de quitar los elementos deterministas, la diferencia  $d$ -ésima del mismo es un proceso  $I(0)$ . También decimos que el proceso tiene  $d$  raíces unitarias.

Se verá posteriormente la importancia que tiene el conocer el orden de integración de las variables con las que se pretende especificar un modelo econométrico. Ahora vamos a describir un proceso de contraste para determinar el orden de

integración de una variable. El punto de partida son las  $T$  observaciones y el gráfico con la variable tiempo en el eje de abscisas y, en ordenadas, los valores de la variable que se analiza. Con carácter general podemos encontrarnos con tres gráficos:

**Gráfico 1.** Los valores de la variable giran en torno a cero, que podemos considerar es la media de la variable. El grado de reversión hacia la media puede ser variable.

**Gráfico 2.** Los valores de la variable giran en torno a una constante diferente de cero. Tampoco en este caso el grado de reversión está definido.

**Gráfico 3.** Los valores de la serie van en torno a una pauta creciente o decreciente que, suponemos, es una función lineal del tiempo. Tampoco en este caso la reversión está definida.

Veamos ahora como diseñar el contraste del orden de integración de una serie según sea el gráfico observado.

**Si observamos el Gráfico 1** entonces el modelo de contraste que utilizamos es

$$y_t = \rho y_{t-1} + u_t \quad y_0 = 0$$

En principio, suponemos que  $u_t$  es un ruido blanco. La hipótesis nula del contraste es que la variable es integrada de orden 1, es decir que es I(1). En el marco del modelo de contraste la hipótesis nula se puede escribir como,

$$H_0 : \rho = 1$$

La hipótesis alternativa es que la serie es integrada de orden cero, es decir, es I(0) que se puede formular como

$$H_1 : \rho < 1$$

Es una hipótesis compuesta unilateral.

El estadístico de contraste es

$$t = \frac{\hat{\rho} - 1}{\hat{\sigma}_{\hat{\rho}}}$$



Con  $\hat{\rho} = \frac{\sum y_{t-1}y_t}{\sum y_{t-1}^2}$  y  $\hat{\sigma}_{\hat{\rho}} = \sqrt{\frac{\hat{\sigma}^2}{\sum y_{t-1}^2}}$

Siendo  $\hat{\sigma}^2 = \frac{\sum \hat{u}_t^2}{T} = \frac{\sum (y_t - \hat{\rho}y_{t-1})^2}{T}$

La región crítica del contraste, es decir, los valores de  $t$  para los que se rechaza la hipótesis nula, viene dada por

$$t < t_{\varepsilon}$$

En donde  $t_{\varepsilon}$  es el punto crítico que cumple

$$Prob.\{t < t_{\varepsilon} / H_0\} = \varepsilon$$

Siendo  $\varepsilon$  el nivel de significación elegido a priori. Los puntos críticos correspondientes a este caso pueden verse en la Tabla 1 en la columna encabezada por Gráfico 1. En esta columna aparecen varios puntos críticos dependiendo del tamaño muestral.

Normalmente el modelo de contraste que se utiliza no es el original sino otro equivalente que se obtiene restando a ambos lados de ese modelo el valor retardado del proceso, obteniéndose

$$\Delta y_t = \rho^* y_{t-1} + u_t$$

Con  $\rho^* = \rho - 1$ . En este caso, la hipótesis nula se formula como  $H_0 : \rho^* = 0$ , siendo el estadístico de contraste

$$t^* = \frac{\hat{\rho}^*}{\hat{\sigma}_{\hat{\rho}^*}}$$

$$\text{Con } \hat{\rho}^* = \frac{\sum y_{t-1} \Delta y_t}{\sum y_{t-1}^2} \quad \text{y} \quad \hat{\sigma}_{\hat{\rho}^*} = \sqrt{\frac{\hat{\sigma}^{*2}}{\sum y_{t-1}^2}}$$

Los puntos críticos son los mismos comentados previamente. Por ejemplo, si  $T=100$ , entonces la hipótesis nula se rechaza si

$$t^* < -1,95$$

**Si observamos el gráfico 2** entonces el modelo de contraste es

$$\Delta y_t = \delta_0 + \rho^* y_{t-1} + u_t \quad \text{con} \quad y_0 \neq 0$$

La hipótesis nula es la misma, aunque en este caso se puede demostrar que esa hipótesis nula implica también que  $\delta_0 = 0$ . El estadístico de contraste es el mismo pero definiendo los estimadores a partir del modelo con constante,

$$\begin{bmatrix} \hat{\delta}_0 \\ \hat{\rho}^* \end{bmatrix} = \begin{bmatrix} T & \sum y_{t-1} \\ \sum y_{t-1} & \sum y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum \Delta y_t \\ \sum y_{t-1} \Delta y_t \end{bmatrix}$$

Los puntos críticos para definir la región crítica están recogidos en la Tabla 1 en la columna encabezada por Gráfico 2. Por ejemplo, si  $T=100$ , con un nivel de significación del 5% la hipótesis nula es rechazada si

$$t^* < -2,89$$

**Tabla 1. Puntos Críticos. Nivel de Significación: 5%.**

<b>Tamaño Muestral</b>	<b>Gráfico 1</b>	<b>Gráfico 2</b>	<b>Gráfico 3</b>
50	-1,95	-2,93	-3,50
100	-1,95	-2,89	-3,45
250	-1,95	-2,88	-3,43

**Si observamos el gráfico 3** entonces el modelo de contraste es

$$\Delta y_t = \delta_0 + \delta_1 t + \rho^* y_{t-1} + u_t$$

La hipótesis nula es la misma, aunque en este caso se puede demostrar que esa hipótesis nula implica también que  $\delta_1 = 0$ . El estadístico de contraste es el mismo pero definiendo los

estimadores a partir del modelo con constante y tendencia

$$\text{lineal} \quad \begin{bmatrix} \hat{\delta}_0 \\ \hat{\delta}_1 \\ \hat{\rho}^* \end{bmatrix} = \begin{bmatrix} T & \sum t & \sum y_{t-1} \\ \sum t & \sum t^2 & \sum ty_{t-1} \\ \sum y_{t-1} & \sum ty_{t-1} & \sum y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum \Delta y_t \\ \sum t \Delta y_t \\ \sum y_{t-1} \Delta y_t \end{bmatrix}$$

Los puntos críticos para definir la región crítica están recogidos en la Tabla 1 en la columna encabezada por Gráfico 3. Por ejemplo, si  $T=100$ , con un nivel de significación del 5% la hipótesis nula es rechazada si

$$t^* < -3,45$$

Hasta ahora hemos supuesto que la perturbación de los tres modelos es un ruido blanco. Pero si la perturbación  $u_t$  esté afectada por **autocorrelación** serial entonces la estrategia tiene que cambiar porque la distribución de probabilidad del estadístico de contraste,  $t^*$ , ya no es la misma. Para evitar este problema lo que se hace es estimar los modelos tal como

los hemos escrito y contrastar la presencia de autocorrelación utilizando los residuos. Si el resultado del contraste es que hay autocorrelación entonces se añade un retardo de la variable endógena como nuevo regresor estimándose el siguiente modelo, en el caso de observar el Gráfico 3

$$\Delta y_t = \delta_0 + \delta_1 t + \rho^* y_{t-1} + \gamma_1 \Delta y_{t-1} + u_t$$

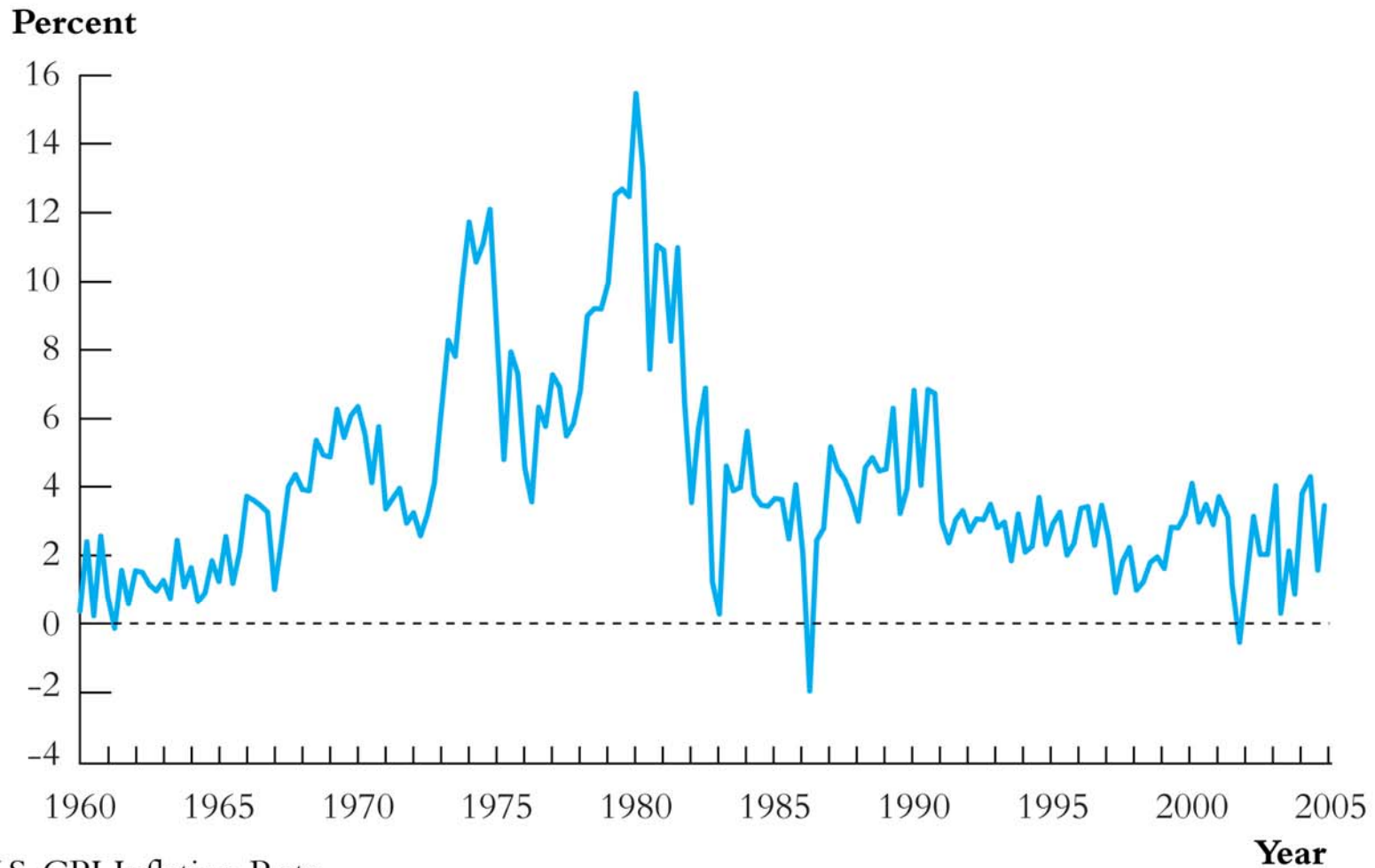
Se vuelve a contrastar la autocorrelación y si se rechaza la hipótesis nula se añade un segundo retardo y el proceso sigue hasta que se llega a un modelo en el que la hipótesis nula de no autocorrelación no es rechazada.

El proceso de contraste que acabamos de describir se suele denominar **contraste de Dickey-Fuller** por ser estos dos autores los que lo desarrollaron.

## ¿Qué problemas causan las tendencias?

1. Las estimaciones de los coeficientes AR son muy sesgadas hacia cero y esto acaba en malas predicciones.
2. Algunos estadísticos-t no siguen la distribución Normal estándar, ni siquiera en muestras grandes.
3. Si las dos  $Y$  y  $X$  siguen un paseo aleatorio entonces puede parecer que están relacionadas cuando realmente no lo están – Es lo que se llama “correlación espúrea.” He aquí un ejemplo...

***Ejemplo: ¿Tiene la tasa de inflación de USA una raíz unitaria?***





La alternativa es que la inflación es estacionaria en torno a una constante.

## *Ejemplo: Salida GRETL*

DF test for a unit root in U.S. inflation – using  $p = 4$  lags

**Modelo 3: MCO, usando las observaciones 1962:1-2004:4 (T = 172)**

**Variable dependiente: d\_Inft**

	Coeficiente	Desv. Típica	Estadístico t	Valor p	
-----					
const	0.506816	0.214181	2.366	0.0191	
Inft_1	-0.113417	0.0422344	-2.685	0.0080	*
d_Inft_1	-0.186443	0.0805144	-2.316	0.0218	*
d_Inft_2	-0.256388	0.0814630	-3.147	0.0020	*
d_Inft_3	0.199049	0.0793514	2.508	0.0131	*
d_Inft_4	0.00999942	0.0779921	0.1282	0.8981	
Media de la vble. dep.	0.017016	D.T. de la vble. dep.	1.707934		
Suma de cuad. residuos	380.6090	D.T. de la regresión	1.514208		
R-cuadrado	0.236972	R-cuadrado corregido	0.213989		
F(5, 166)	10.31083	Valor p (de F)	1.27e-08		
Log-verosimilitud	-312.3653	Criterio de Akaike	636.7307		
Criterio de Schwarz	655.6157	Crit. de Hannan-Quinn	644.3928		
rho	0.002177	Durbin-Watson	1.992076		

**Estadístico DF = -2.69**

No compara este con -1.645 – utilizar la tabla de Dickey-Fuller !

*Nota:*

- *El número de retardos en la regresión DF puede determinarse con el BIC, AIC o t-student. Utilizaremos el AIC.*
- La rutina en Gretl es **MODELO/CONTRASTES DE RAIZ UNITARIA/CONTRASTE AUMENTADO DE DICKY-FULLER**

## ¿Cómo solucionar los problemas que causan las tendencias?

Si  $Y_t$  tiene una raíz unitaria (un paseo aleatorio), la forma más fácil de evitar los problemas es modelizar tomando primeras diferencias de  $Y_t$ .

- Esto es lo que hicimos en el primer tratamiento de la inflación –ya que la inspección visual del gráfico más los resultados del contraste DF hacían plausible aceptar la presencia de una raíz unitaria- eso nos llevó a estimar los AR utilizando  $\Delta Inf_t$

## **RESUMEN**

1. El modelo de paseo aleatorio es el caballo de batalla para las tendencias de los datos económicos de series temporales
2. Para determinar si  $Y_t$  tiene una tendencia estocástica, primero el gráfico de  $Y_t$ . Si una tendencia parece plausible, calcular el contraste DF (decidir que versión, constante o constante +tendencia)
3. Si el DF no rechaza, concluimos con que  $Y_t$  tiene una raíz unitaria
4. Si  $Y_t$  tiene una raíz unitaria, utilizar  $\Delta Y_t$  para el análisis de regresión y predicción. Si no hay raíz unitaria, usar  $Y_t$ .

## No estacionariedad II: Rupturas

El Segundo tipo de no estacionariedad que consideramos es la de que los coeficientes del modelo pueden no ser constantes a lo largo de toda la muestra. Esta posibilidad de inestabilidad del modelo plantea problemas cuando se va a predecir (Es una cuestión que afecta a la validación externa.)

Lo que haremos será:

- Considerar dos formas de detectar el cambio en los coeficientes: contrastes de ruptura, y predicciones pseudo fuera de la muestra.
- Trabajar con un ejemplo: la curva de Phillips en USA.

## Contrastes de ruptura (cambio) en los coeficientes de regresión

### Caso I: La fecha de la ruptura es conocida

Suponer que se conoce que la ruptura ha ocurrido en el periodo  $\tau$ . La estabilidad de los coeficientes puede ser contrastada estimando un modelo con variables ficticias. En el caso  $ARD(1,1)$ :

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \delta_1 X_{t-1} \\ + \gamma_0 D_t(\tau) + \gamma_1 [D_t(\tau) \times Y_{t-1}] + \gamma_2 [D_t(\tau) \times X_{t-1}] + u_t$$

donde  $D_t(\tau) = 1$  if  $t \geq \tau$ , y  $= 0$  en otro caso.

- Si  $\gamma_0 = \gamma_1 = \gamma_2 = 0$ , entonces los coeficientes son constantes a lo largo de toda la muestra.
- Si, al menos uno de  $\gamma_0$ ,  $\gamma_1$ , o  $\gamma_2$  son diferentes de cero, la regresión tiene una ruptura en  $\tau$ .

El contraste de Chow de una ruptura en  $\tau$  es el contraste F (robusto a la heterocedasticidad) que contrasta:

$$H_0: \gamma_0 = \gamma_1 = \gamma_2 = 0$$

vs.  $H_1$ : al menos uno de  $\gamma_0$ ,  $\gamma_1$ , or  $\gamma_2$  son diferentes de cero.

$$\text{El estadístico: } F = \frac{(R_A^2 - R_R^2) / q}{(1 - R_A^2)(n - k_A - 1)} \sim F_{(q, n - k_A - 1)}$$

(q es el número de restricciones)

- Destacar que se puede aplicar este contraste a un subconjunto de los coeficientes, por ejemplo al coeficiente de  $X_{t-1}$ .
- El problema es que, en general, no se conoce la fecha de la ruptura,  $\tau \dots$



## Caso II: La fecha de la ruptura no es conocida

¿Por qué consideramos este caso?

- Puedes sospechar que hay una ruptura pero no sabes cuando.
- Tú puedes desear contrastar la hipótesis nula de estabilidad frente a una alternativa general de que ha habido una ruptura en algún momento.
- Incluso si crees que conoces la fecha de ruptura después de ver el gráfico de la serie en realidad es como si se hubiera estimado esa fecha y eso puede cuestionar el uso del contraste de Chow.

# El Estadístico de la Razón de Verosimilitud de Quandt (The Quandt Likelihood Ratio (QLR) Statistic) (también llamado el estadístico supWald)

El estadístico  $QLR$  = el máximo del estadístico de Chow

- Sea  $F(\tau)$  = el contraste de Chow para la hipótesis nula de que hay una ruptura en  $\tau$ .
- El estadístico  $QLR$  es el **máximo** de todos los estadísticos  $F$  de Chow, en un rango de  $\tau$ ,  $\tau_0 \leq \tau \leq \tau_1$ :

$$QLR = \max[F(\tau_0), F(\tau_0+1), \dots, F(\tau_1-1), F(\tau_1)]$$

- Una práctica habitual ha consistido en determinar el rango de forma que  $\tau_0$  y  $\tau_1$  abarcan el 70% de la muestra(excluyen el primer y último 15%).
- ¿Se pueden usar los puntos críticos de  $F_{q,\infty}$  ?

## El contraste $QLR$ .

$$QLR = \max[F(\tau_0), F(\tau_0+1), \dots, F(\tau_1-1), F(\tau_1)]$$

- La distribución asintótica bajo la nula de  $F(\tau)$  para un  $\tau$  dado (fijo, no estimado) es  $F_{q,\infty}$
- Pero si calculas dos estadísticos Chow y eliges el mayor, entonces el valor crítico tiene que ser mayor que el que corresponde a un solo estadístico de Chow.
- Si calculas muchos estadísticos de Chow, por ejemplo, los correspondientes a todos los periodos del 70% central de la muestra, el valor crítico tiene que ser aún mayor!

**TABLE 14.6** Critical Values of the QLR Statistic with 15% Trimming

Number of Restrictions ( $q$ )	10%	5%	1%
1	7.12	8.68	12.16
2	5.00	5.86	7.78
3	4.09	4.71	6.02
4	3.59	4.09	5.12
5	3.26	3.66	4.53
6	3.02	3.37	4.12
7	2.84	3.15	3.82
8	2.69	2.98	3.57
9	2.58	2.84	3.38
10	2.48	2.71	3.23

Destacar que estos valores críticos son mayores que los de la distribución estándar  $F_{q,\infty}$ .

## Ejemplo: ¿Se ha mantenido estable la curva de Phillips?

Recordar el modelo estimado y presentado anteriormente para el periodo (1962 – 2004):

$$\widehat{\Delta Inf_t} = 1.30 - .42\Delta Inf_{t-1} - .37\Delta Inf_{t-2} + .06\Delta Inf_{t-3} - .04\Delta Inf_{t-4}$$

(.44) (.08)                      (.09)                      (.08)                      (.08)

$$- 2.64Unem_{t-1} + 3.04Unem_{t-2} - 0.38Unem_{t-3} + .25Unemp_{t-4}$$

(.46)                      (.86)                      (.89)                      (.45)

¿Se ha mantenido estable durante todo el periodo 1962-2004?

## Contrastes QLR de la estabilidad de la curva de Phillips en USA.

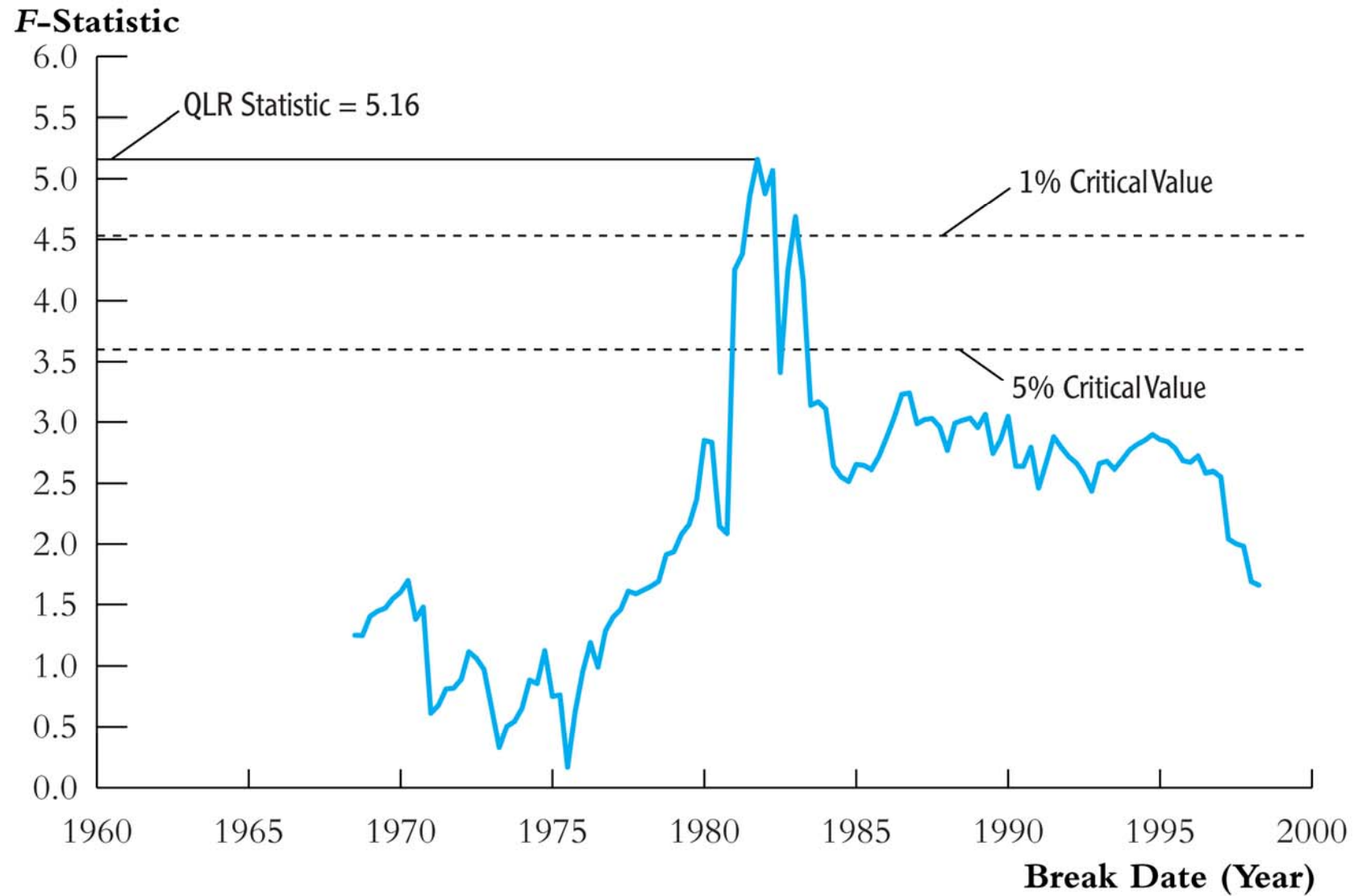
variable dependiente:  $\Delta Inf_t$

regresores: constante,  $\Delta Inf_{t-1}, \dots, \Delta Inf_{t-4}, Unemp_{t-1}, \dots, Unemp_{t-4}$

- Contraste de la estabilidad solo de la constante (se supone que los otros coeficientes son constantes):  $QLR = 2.865$  ( $q = 1$ ).
  - Valor crítico al 10% = 7.12  $\Rightarrow$  no rechazar al 10%
- Contraste de la estabilidad de la constante y de los coeficientes de  $Unemp_t, \dots, Unemp_{t-3}$  (suponiendo que los coeficientes de  $f_{t-1}, \dots, \Delta Inf_{t-4}$  son constantes):  $QLR = 5.158$  ( $q = 5$ )
  - 1% valor crítico = 4.53  $\Rightarrow$  rechazar al 1%

- Estimar la fecha de ruptura = máxima  $F$  que aparece en 1981:IV
- Concluir con la existencia de una ruptura en la relación tasa de inflación-desempleo, con fecha estimada de 1981:IV

**FIGURE 14.5** *F*-Statistics Testing for a Break in Equation (14.17) at Different Dates



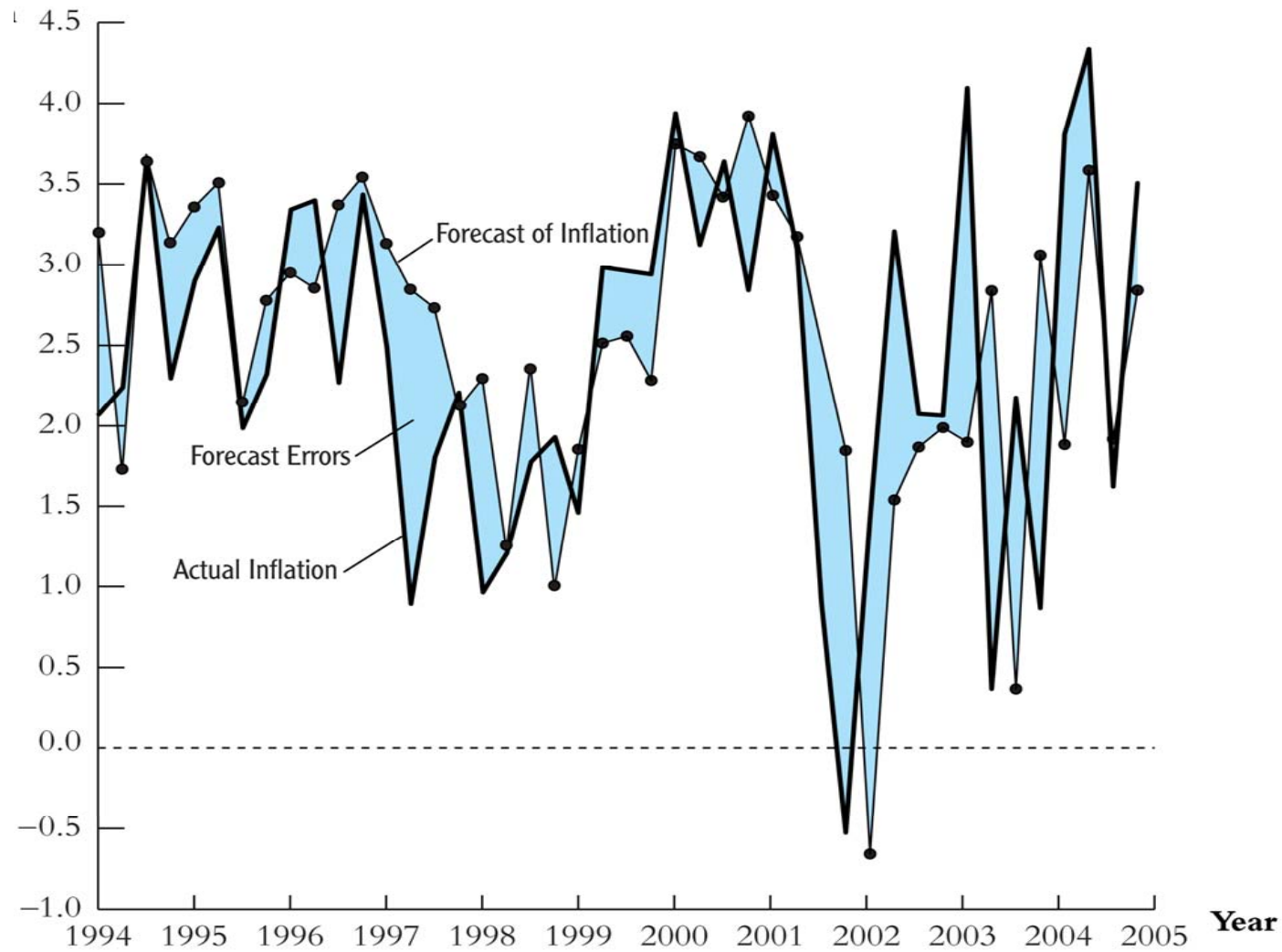


## Aplicación a la curva de Phillips :

*¿Era la curva de Phillips estable hacia el final de la muestra?*

- Hemos encontrado una ruptura en 1981:IV – por lo que para este análisis consideramos solamente los datos a partir de 1982:I – ignorando los datos anteriores del “viejo” modelo (“regimen”).
- Modelo de regresión:
  - variable dependiente:  $\Delta Inf_t$
  - regresores:  $1, \Delta Inf_{t-1}, \dots, \Delta Inf_{t-4}, Unemp_{t-1}, \dots, Unemp_{t-4}$
- Predicciones pseudo fuera de la muestra:
  - Calcular la regresión para el periodo  $t = 1982:I, \dots, P$
  - Calcular las predicciones  $\widehat{\Delta Inf}_{P+1|P}$ , y el error de predicción.
  - Repetirlo para  $P = 1994:I, \dots, 2005:I$

# Predicciones de $\Delta Inf$ utilizando el modelo ARD(4,4) con $Unemp$



Hay algunos errores que son grandes en torno a 2001 pero el modelo no se deteriora en periodos posteriores.

## Predicciones para la curva de *Phillips*

### Resumen:

- Media de los errores de predicción, 1999:I – 2004:IV = 0.11 ( $SE = 0.27$ )
  - No hay evidencia de que los errores sean sistemáticamente positivos o negativos.
- *RECOMP*, 1999:I – 2004:IV: 1.32
- *Error Estándar del modelo ajustado en* 1982:I – 1998:IV: 1.30
  - Por estar próximos el modelo se comporta de forma parecida dentro de la muestra que en las predicciones pseudo fuera.

Este análisis sugiere que no había un cambio sustancial de las predicciones generadas por el modelo ARD(4,4) en el último tramo de la muestra.

## 4. Enfoque Box-Jenkins

Proceso con cuatro etapas:

- Identificación
- Estimación
- Chequeo
- Predicción

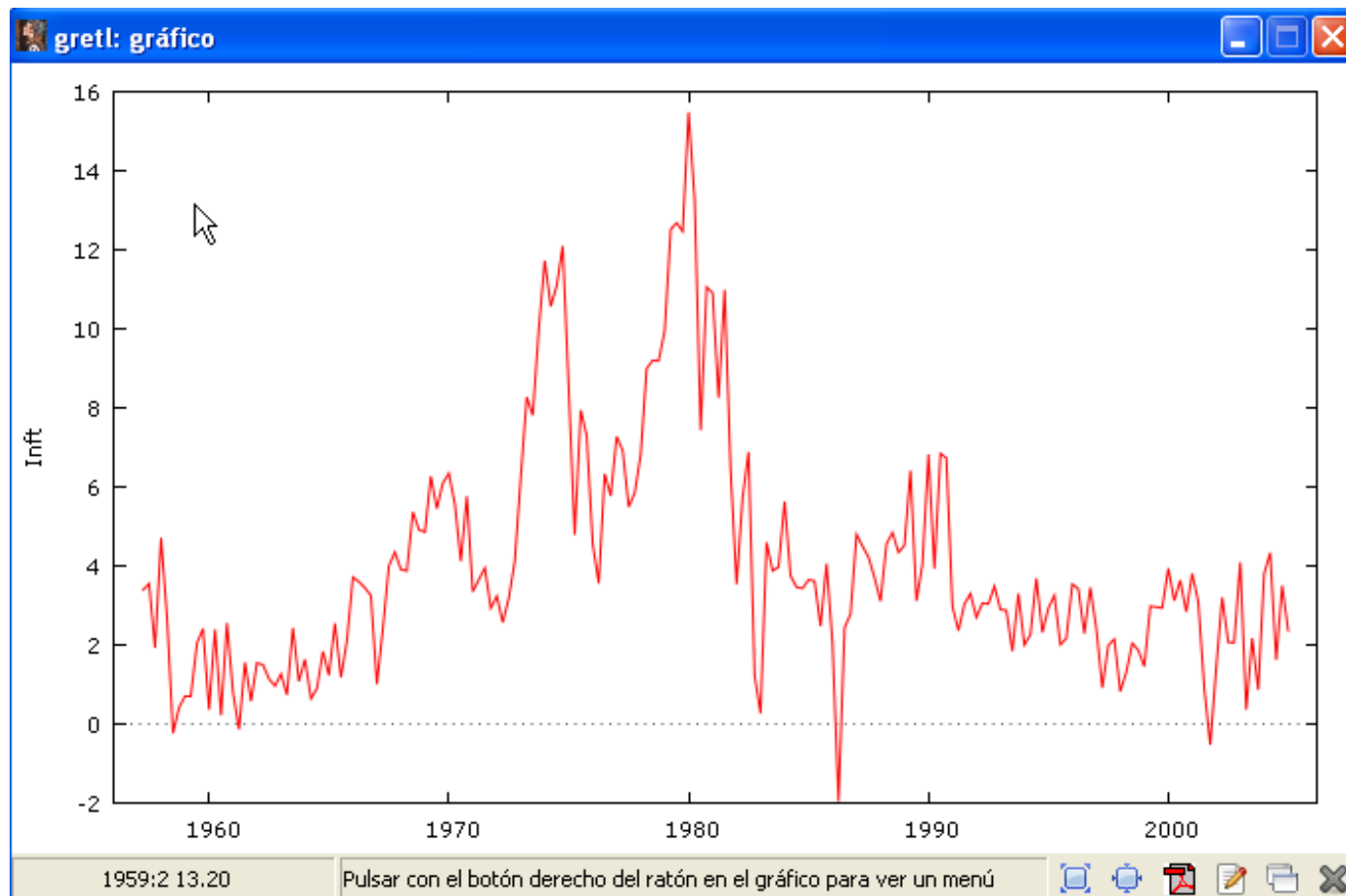
(Todo desarrollado en Econometría II) (ejemplo)

# **METODOLOGÍA BOX-JENKINS APLICADA A LA SERIE DE INFLACIÓN**

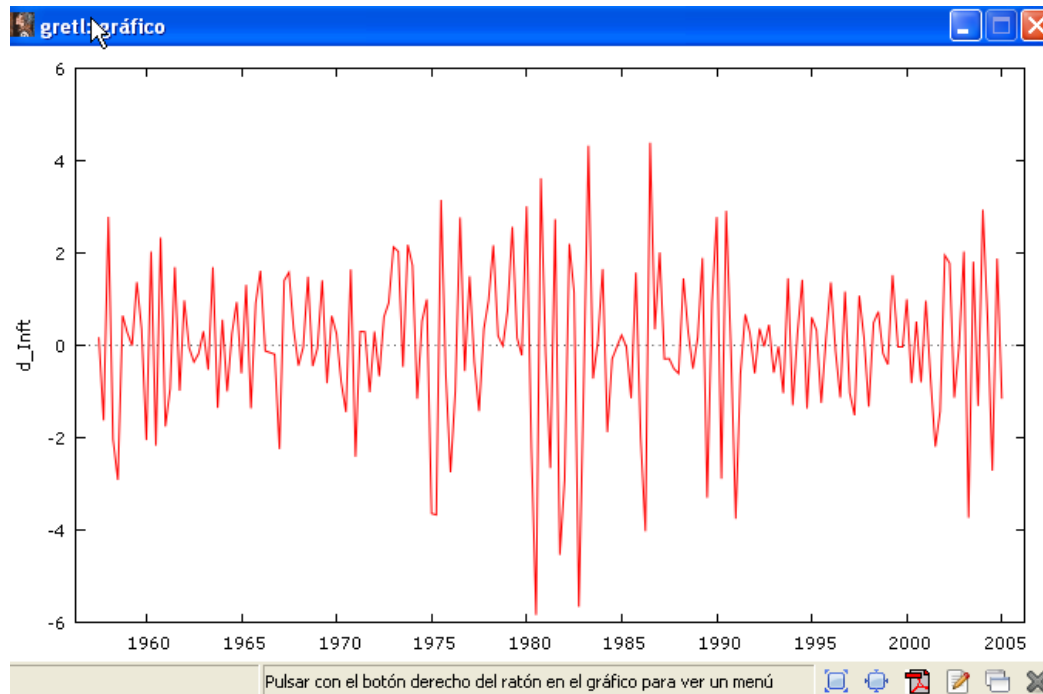
Utilizando la variable inflación. Estimando con la muestra de 1957:3-2004:2 y la predicción para 2004:3 a 2005:2.

## **1. IDENTIFICACIÓN**

- **Gráficos**



## Gráfico de la primera diferencia de la serie

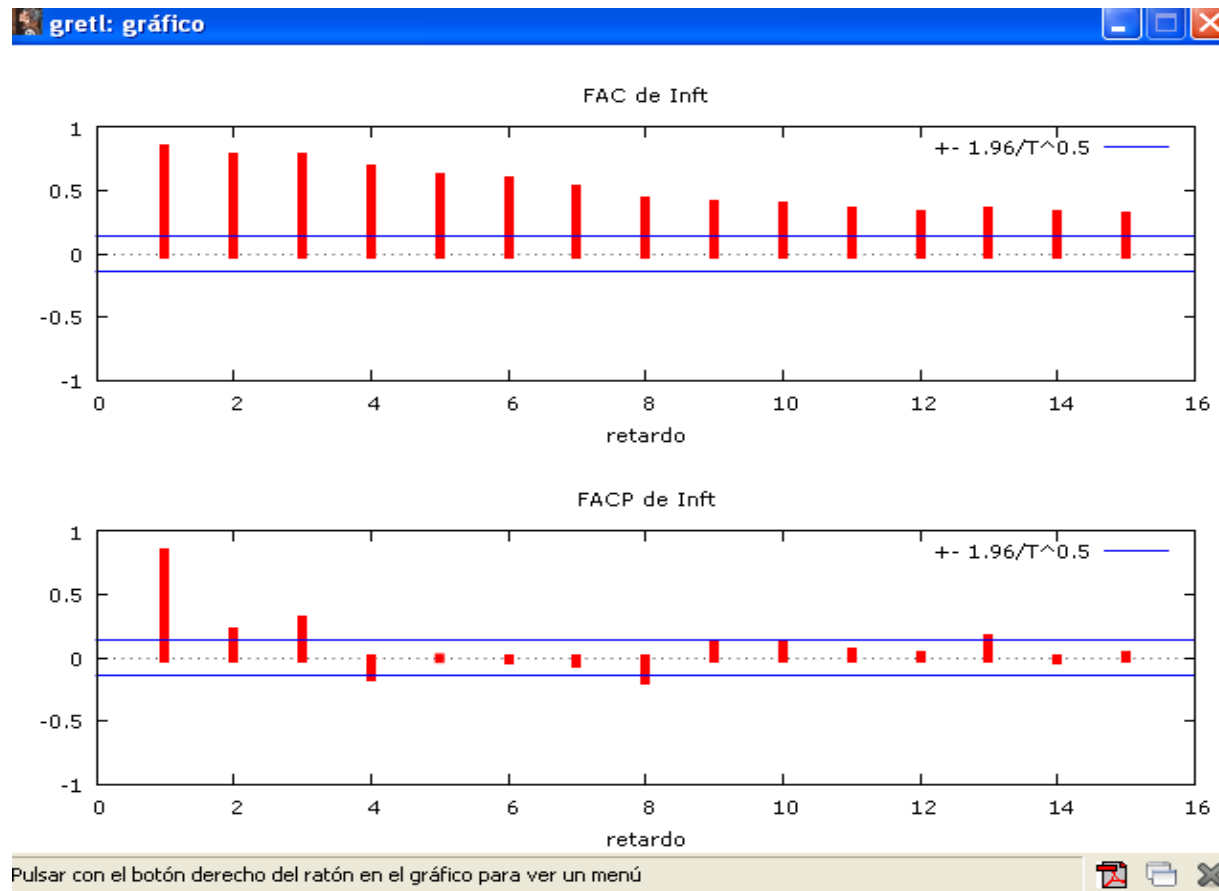


La serie diferenciada una vez ya es estacionaria en media, nos que damos con la serie diferenciada una vez,  $d = 1$ .

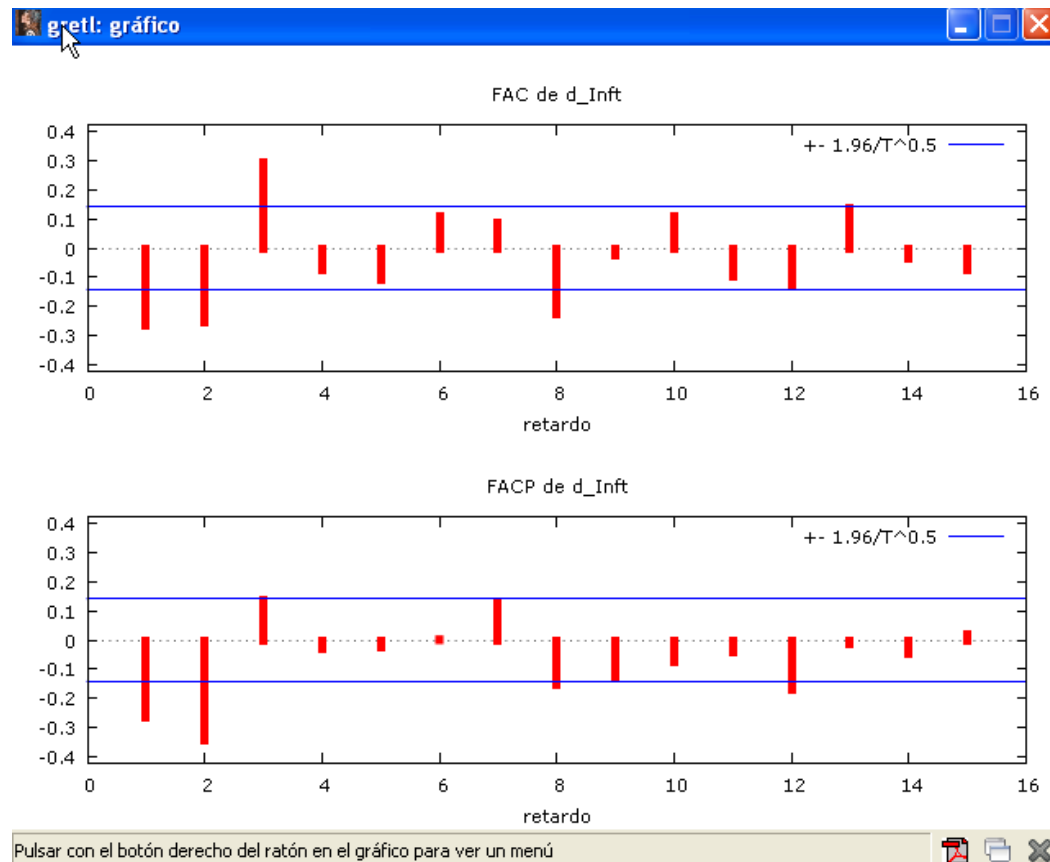


- Correlogramas

- Correlograma de la serie original



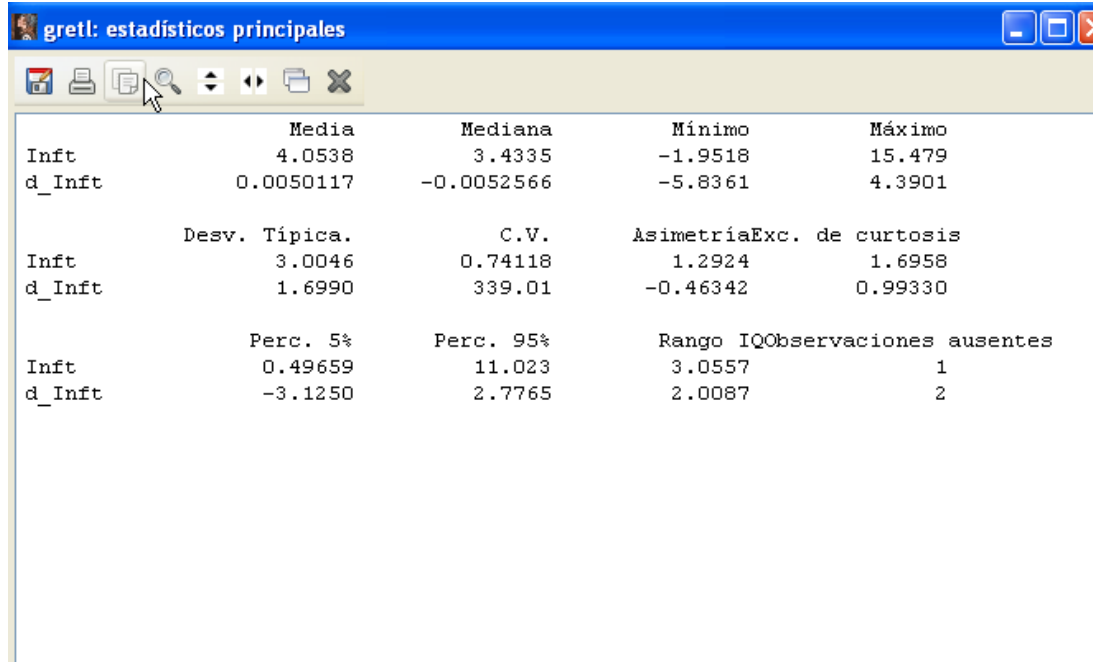
# Correlogramas de la primera diferencia de la serie



Conclusión:

Según los correlogramas diríamos que  $d = 1$ .

- **Desviación típica**



The screenshot shows the 'gretl: estadísticos principales' window. It displays statistical results for two variables: 'Inft' and 'd\_Inft'. The window has a blue title bar and a toolbar with icons for file operations and navigation. The data is presented in three sections: summary statistics, descriptive statistics, and percentiles.

	Media	Mediana	Mínimo	Máximo
Inft	4.0538	3.4335	-1.9518	15.479
d_Inft	0.0050117	-0.0052566	-5.8361	4.3901

	Desv. Típica.	C.V.	Asimetría	Exc. de curtosis
Inft	3.0046	0.74118	1.2924	1.6958
d_Inft	1.6990	339.01	-0.46342	0.99330

	Perc. 5%	Perc. 95%	Rango IQ	Observaciones ausentes
Inft	0.49659	11.023	3.0557	1
d_Inft	-3.1250	2.7765	2.0087	2

Conclusión:

Según todos los criterios  $d = 1$ .

- **Contraste de Dickey-Fuller:**

**$H_0: I(1)$**

**$H_1: I(0)$**

Según el gráfico de la serie en niveles deberíamos de utilizar el modelo con constante.

```

contraste aumentado de Dickey-Fuller para Inft
incluyendo 2 retardos de (1-L) Inft
el máximo fue 5, el criterio AIC modificado)
año muestral 186
hipótesis nula de raíz unitaria: a = 1

contraste con constante
modelo: (1-L)y = b0 + (a-1)*y(-1) + ... + e
Coef. de autocorrelación de primer orden de e: 0.061
diferencias retardadas: F(2, 182) = 13.128 [0.0000]
valor estimado de (a - 1): -0.0899355
Estadístico de contraste: tau_c(1) = -2.24691
valor p asintótico 0.1898

regresión aumentada de Dickey-Fuller
AIC, usando las observaciones 1958:1-2004:2 (T = 186)
variable dependiente: d_Inft

      Coeficiente   Desv. Típica   Estadístico t   Valor p
-----
const      0.374155      0.197993         1.890         0.0604   *
Inft_1     -0.0899355      0.0400262        -2.247         0.1898
d_Inft_1   -0.288566      0.0727929        -3.964         0.0001   ***
d_Inft_2   -0.307696      0.0710921        -4.328         2.48e-05   ***

AIC: 692.416   BIC: 705.319   HQC: 697.644

```

DF= -2.2469      p-valor = 0.1898

No rechazamos la hipótesis de que al menos la serie es  $I(1)$ .

**$H_0: I(2)$**

**$H_1: I(1)$**

Modelo para calcular el estadístico de Dickey-Fuller →

Gráfico de la primera diferencia de la serie (sin constante)

```

Contraste aumentado de Dickey-Fuller para d_Inft
incluyendo 2 retardos de (1-L)d_Inft
(el máximo fue 5, el criterio AIC modificado)
tamaño muestral 185
hipótesis nula de raíz unitaria: a = 1

contraste sin constante
modelo: (1-L)y = (a-1)*y(-1) + ... + e
Coef. de autocorrelación de primer orden de e: 0.013
diferencias retardadas: F(2, 182) = 14.229 [0.0000]
valor estimado de (a - 1): -1.43863
Estadístico de contraste: tau_nc(1) = -8.72064
valor p asintótico 5.705e-016

```

```

Regresión aumentada de Dickey-Fuller
MCO, usando las observaciones 1958:2-2004:2 (T = 185)
Variable dependiente: d_d_Inft

```

	Coeficiente	Desv. Típica	Estadístico t	Valor p	
d_Inft_1	-1.43863	0.164968	-8.721	5.71e-016	***
d_d_Inft_1	0.151306	0.120427	1.256	0.2106	
d_d_Inft_2	-0.142736	0.0734622	-1.943	0.0536	*

AIC: 686.806    BIC: 696.467    HQC: 690.722

DF = -8.7206

p-valor = 0.0000

Conclusión del contraste de Dickey-Fuller:

Se rechaza la hipótesis nula de que es I (2) por lo que concluimos que es I(1), por lo tanto  $d = 1$ .

Del correlograma de la serie diferenciada una vez, se puede identificar un  $AR(2)$  o  $AR(3)$ , vamos a estimar estos modelos, los anidados en él,  $AR(1)$  y uno más amplio  $AR(4)$ .



## 2. ESTIMACIÓN, CHEQUEO Y PREDICCIÓN

	AR(1)	AR(2)	AR(3)	AR(4)
<b>cte</b>	0.0043 (0.96)	0.0021 (0.97)	0.004 (0.96)	0.0035 (0.96)
$\phi_1$	-0.2533 (0.00)	-0.3418 (0.00)	-0.29 (0.00)	-0.29 (0.00)
$\phi_2$		-0.3423 (0.00)	-0.29 (0.00)	-0.30 (0.00)
$\phi_3$			0.14 (0.07)	0.13 (0.07)
$\phi_4$				-0.03 (0.72)
<b>LB(1)</b>	1.44 (0.23)	0.39 (0.53)	0.00 (0.99)	0.00 (0.99)
<b>LB(2)</b>	16.19 (0.00)	0.51 (0.78)	0.00 (0.99)	0.01 (0.99)

<b>LB(3)</b>	27.41 (0.00)	2.23 (0.13)	0.11 (0.99)	0.06 (0.99)
<b>LB(4)</b>	27.56 (0.00)	2.83 (0.24)	0.32 (0.57)	0.11 (0.99)
<b>ARCH(4)</b>	11.62 (0.02)	7.68 (0.10)	5.73 (0.21)	5.93 (0.20)
<b>J-B</b>	36.55 (0.00)	23.33 (0.00)	20.82 (0.00)	20.63 (0.00)
<b>SBIC</b>	734.98	717.12	718.55	723.66
<b>AIC</b>	725.27	704.18	702.37	704.24
<b>EAPM</b>	86.53	51.55	56.60	53.39

Entre paréntesis p-valor

Tenemos un ARIMA(2,1,0) o un ARIMA(3,1,0)

## PREDECIR CON UN ARIMA(2,1,0)

El modelo estimado es:

$$\widehat{\Delta Inf}_t = \hat{\phi}_0 + \hat{\phi}_1 \Delta Inf_{t-1} + \hat{\phi}_1 \Delta Inf_{t-2}$$

Vamos a distinguir entre la predicción dinámica y estática y del incremento de la inflación o de la inflación misma.

A) **Predicción dinámica.** Incremento de la inflación:

$$\widehat{\Delta Inf}_{T+1/T} = \hat{\phi}_0 + \hat{\phi}_1 \Delta Inf_T + \hat{\phi}_1 \Delta Inf_{T-1}$$

$$\widehat{\Delta Inf}_{T+2/T} = \hat{\phi}_0 + \hat{\phi}_1 \widehat{\Delta Inf}_{T+1/T} + \hat{\phi}_1 \Delta Inf_T$$

$$\widehat{\Delta Inf}_{T+3/T} = \hat{\phi}_0 + \hat{\phi}_1 \widehat{\Delta Inf}_{T+2/T} + \hat{\phi}_1 \widehat{\Delta Inf}_{T+1/T}$$

**B) Predicción dinámica.** Tasa de inflación:

En niveles el modelo estimado puede escribirse como:

$$Inf_t = \hat{\phi}_0 + (\hat{\phi}_1 + 1) Inf_{t-1} + (\hat{\phi}_2 - \hat{\phi}_1) Inf_{t-2} - \hat{\phi}_2 Inf_{t-3}$$

Por lo que las predicciones serán:

$$\widehat{Inf}_{T+1/T} = \hat{\phi}_0 + (\hat{\phi}_1 + 1) Inf_T + (\hat{\phi}_2 - \hat{\phi}_1) Inf_{T-1} - \hat{\phi}_2 Inf_{T-2}$$

$$\widehat{Inf}_{T+2/T} = \hat{\phi}_0 + (\hat{\phi}_1 + 1) \widehat{Inf}_{T+1/T} + (\hat{\phi}_2 - \hat{\phi}_1) Inf_T - \hat{\phi}_2 Inf_{T-1}$$

$$\widehat{Inf}_{T+3/T} = \hat{\phi}_0 + (\hat{\phi}_1 + 1) \widehat{Inf}_{T+2/T} + (\hat{\phi}_2 - \hat{\phi}_1) \widehat{Inf}_{T+1/T} - \hat{\phi}_2 Inf_T$$

**C) Predicción Estática: Diferencia de la tasa de inflación**

$$\widehat{\Delta Inf}_{T+1/T} = \hat{\phi}_0 + \hat{\phi}_1 \Delta Inf_T + \hat{\phi}_1 \Delta Inf_{T-1}$$

$$\widehat{\Delta Inf}_{T+2/T} = \hat{\phi}_0 + \hat{\phi}_1 \Delta Inf_{T+1} + \hat{\phi}_1 \Delta Inf_T$$

$$\widehat{\Delta Inf}_{T+3/T} = \hat{\phi}_0 + \hat{\phi}_1 \Delta Inf_{T+2} + \hat{\phi}_1 \Delta Inf_{T+1}$$

**D) Predicción Estática: Tasa de inflación**

$$\widehat{Inf}_{T+1/T} = \hat{\phi}_0 + (\hat{\phi}_1 + 1) Inf_T + (\hat{\phi}_2 - \hat{\phi}_1) Inf_{T-1} - \hat{\phi}_2 Inf_{T-2}$$

$$\widehat{Inf}_{T+2/T} = \hat{\phi}_0 + (\hat{\phi}_1 + 1) Inf_{T+1} + (\hat{\phi}_2 - \hat{\phi}_1) Inf_T - \hat{\phi}_2 Inf_{T-1}$$

$$\widehat{Inf}_{T+3/T} = \hat{\phi}_0 + (\hat{\phi}_1 + 1) Inf_{T+2} + (\hat{\phi}_2 - \hat{\phi}_1) Inf_{T+1} - \hat{\phi}_2 Inf_T$$

## Indicadores para medir la calidad de las predicciones

Estimaremos el modelo con las primeras  $T-H$  observaciones.

Cuantificaremos los errores cometidos al efectuar, en base a él, las predicciones para los  $H$  periodos reservados:

- Importante ver el gráfico de los errores de predicción.

A) Error Cuadrático Medio (ECM):

$$ECM(H) = \frac{1}{H} \sum_{l=1}^H \left( y_{T-H+l} - \hat{y}_{T-H+l/T-H} \right)^2$$

B) Raíz del Error Cuadrático Medio (RECM):

$$RECM(H) = \sqrt{\frac{1}{H} \sum_{l=1}^H \left( y_{T-H+l} - \hat{y}_{T-H+l/T-H} \right)^2}$$

C) Error Absoluto Medio (EAM):

$$EAM(H) = \frac{1}{H} \sum_{l=1}^H \left| y_{T-H+l} - \hat{y}_{T-H+l/T-H} \right|$$

D) Error Absoluto Medio Porcentual (EAMP):

$$EAMP(H) = \frac{1}{H} \sum_{l=1}^H \frac{\left| y_{T-H+l} - \hat{y}_{T-H+l/T-H} \right|}{y_{T-H+l}} \times 100$$