MÓDULO DE ESTADÍSTICA SERIES DE TIEMPO

DIPLOMADO DE MATEMÁTICAS Y ESTADÍSTICA, MIDE

Ruslán Gómez Nesterkín

Noviembre 2022

Aviso: Los comentarios y opiniones expresados son solo del autor y no necesariamente reflejan a los del Banco de México.

ENCUESTAS

- A fin de tener una mayor interacción durante las presentaciones…
- Se realizarán algunas encuestas de opinión durante las sesiones…
- A través de la aplicación de celular: **Telegram**
- Favor de adherirse al grupo: **MIDE_DIP_MATS_2022**
- Liga directa: *https://t.me/MIDE_DIP_MATS_2022*

PLAN DE LA PRESENTACIÓN

PARTE 1

- **ANTECEDENTES:** Series determinísticas, variables aleatorias y estimación.
- **SERIES DE TIEMPO:** Series Determinísticas, Series de Tiempo, Suma de Variables Aleatorias.
- **DISTRIBUCIÓN DE SERIES DE TIEMPO:** Límite Central, Función Característica, Caminata Aleatoria.
- **CARACTERÍSTICAS DE LAS SERIES DE TIEMPO:** Autocorrelación, Estacionalidad.

PARTE 2

- **DESCOMPOSICIÓN:** Descomposición de Wold.
- **MODELOS DE SERIES DE TIEMPO:** AR y MA.
- **EJEMPLO INICIAL:** ¿Predicción al 75%?

PARTE 3

- **CARACTERÍSTICAS DE LAS SERIES DE TIEMPO:** Heteroskedasticidad, Causalidad de Granger y Cointegración.
- **MÁS MODELOS DE SERIES DE TIEMPO:** ARMA, Otras variantes.
- **PRONÓSTICOS**

PARTE 2
MODELOS

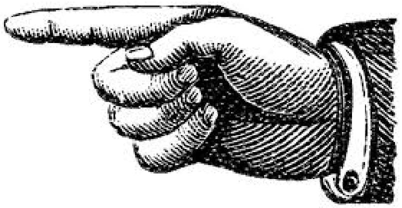
PLAN DE LA PRESENTACIÓN

PARTE 1

- **ANTECEDENTES**
- **SERIES DE TIEMPO**
- **DISTRIBUCIÓN DE SERIES DE TIEMPO**
- **CARACTERÍSTICAS DE LAS SERIES DE TIEMPO**

PARTE 2

- **DESCOMPOSICIÓN**
- **MODELOS DE SERIES DE TIEMPO**
- **EJEMPLO**

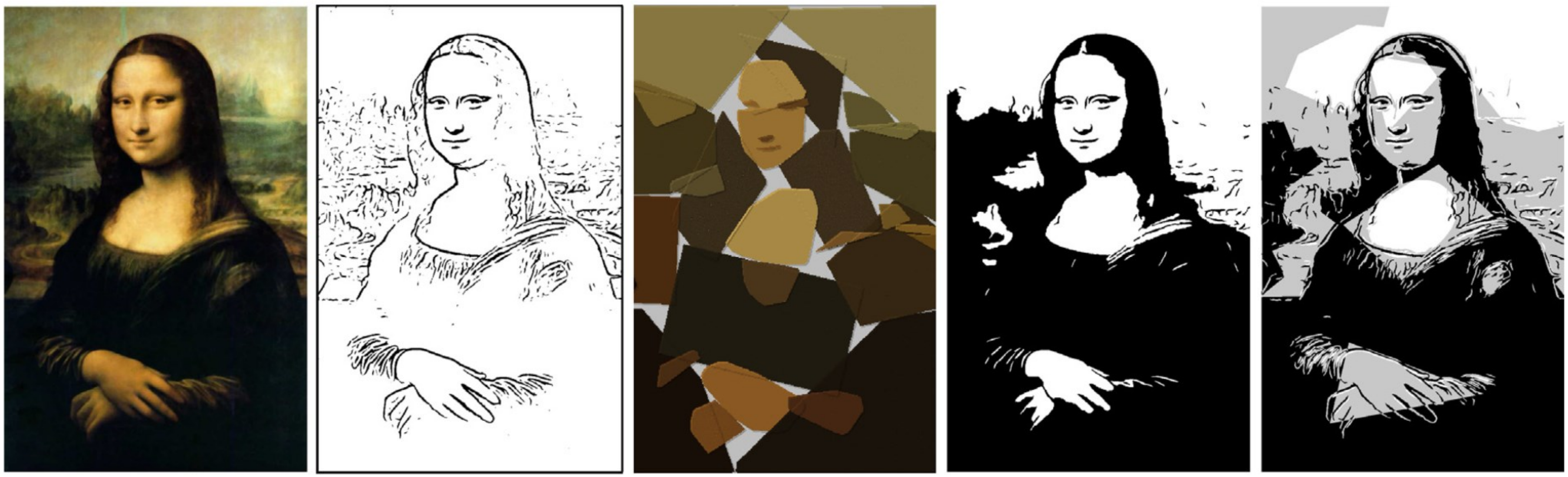


PARTE 3

- **MODELOS DE SERIES DE TIEMPO**
- **OTRAS CARACTERÍSTICAS**
- **PRONÓSTICOS**

DESCOMPOSICIÓN

Motivación ...



Retrato y otras representaciones del cuadro de Lisa Gherardini / La Gioconda / Mona Lisa. Reproducción original de Leonardo da Vinci a la izquierda, siglo XVI.
Fuente: P.L. Rosin, Y.-K. Lai; *"Artistic minimal rendering with lines and blocks"*; Graphical Models 75 (2013) pp. 208–229.

- Entenderemos por **modelo matemático** a aquella representación de una realidad o abstracción, dada a través de relaciones matemáticas.

- Dado un modelo, también puede considerarse la **descomposición** de éste en elementos principales, más útiles o fáciles de manejar.

DESCOMPOSICIÓN

Ejemplo 1/6

- **Datos:** Producto interno bruto de la República Federal de Alemania, datos trimestrales entre 1960 y 2011.

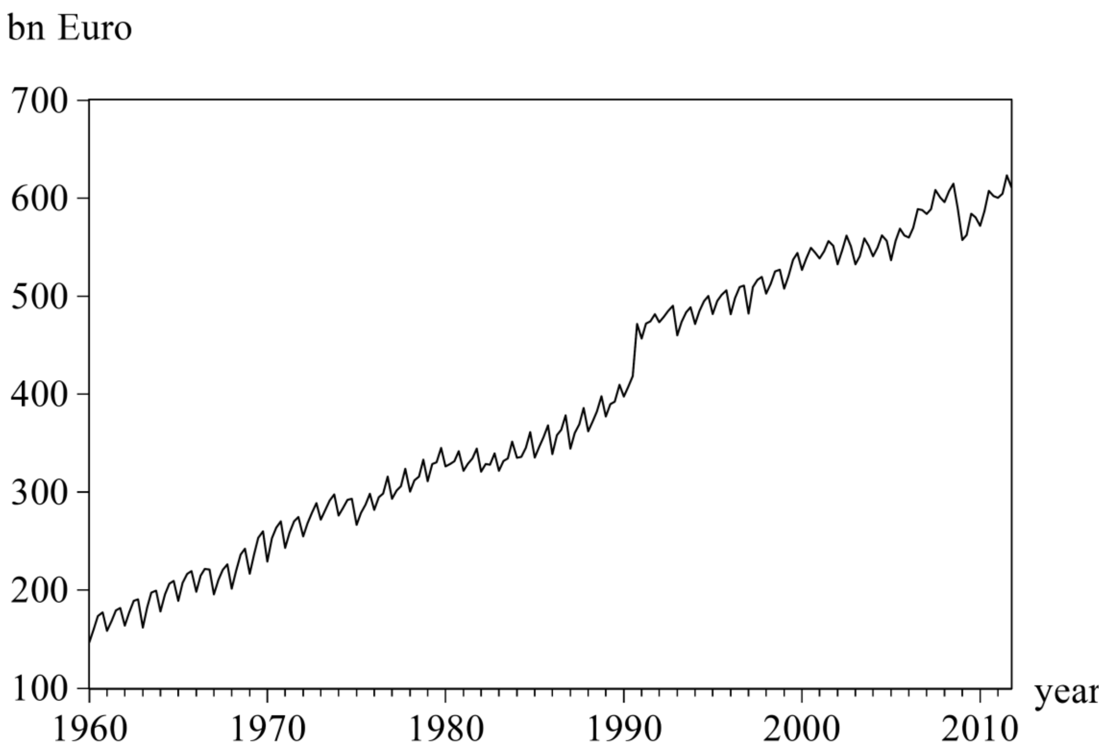


Figure 1.1: Real Gross Domestic Product of the Federal Republic of Germany in billions of Euro, 1960 – 2011

- Veamos distintas maneras de **transformar** dicha serie de tiempo…

DESCOMPOSICIÓN

Ejemplo 2/6

- **Datos:** Producto interno bruto de la Rep. Federal de Alemania, datos trimestrales entre 1960 y 2011:

$$X = \{X_1, X_2, \dots, X_n\}$$

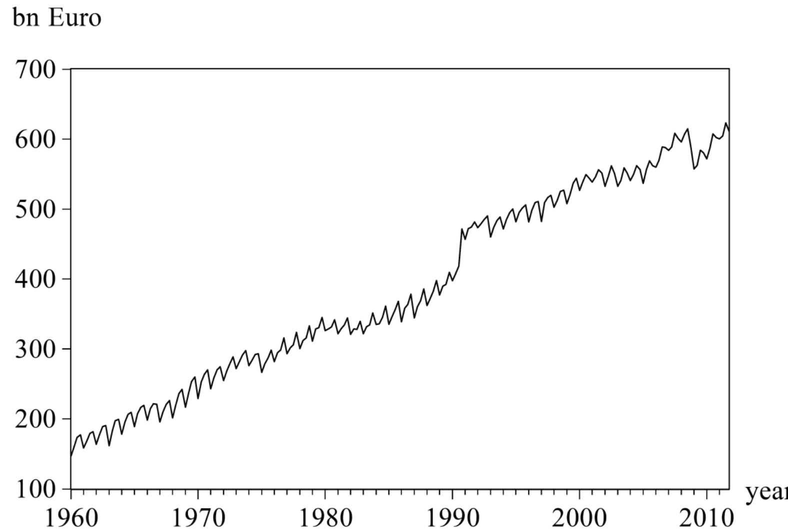


Figure 1.1: Real Gross Domestic Product of the Federal Republic of Germany in billions of Euro, 1960 – 2011

- **Transformación:** Eliminación de la tendencia

$$\Delta X_n = X_n - X_{n-1}$$

- **Comentario:** Se elimina la tendencia, permaneciendo la periodicidad propia del proceso original X. La amplitud es esencialmente la misma.

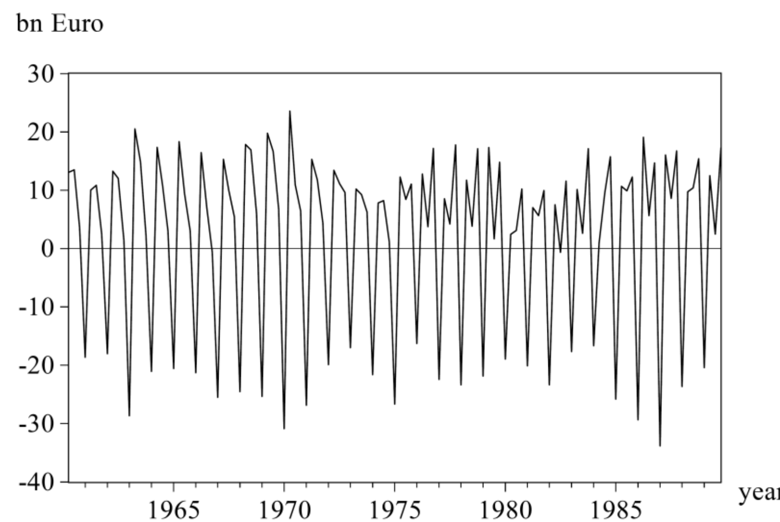


Figure 1.2: Quarterly Changes of the Real Gross Domestic Product (ΔGDP) of the Federal Republic of Germany, 1960 – 1989

DESCOMPOSICIÓN

Ejemplo 3/6

- **Datos:** Producto interno bruto de la Rep. Federal de Alemania, datos trimestrales entre 1960 y 2011:

$$X = \{X_1, X_2, \dots, X_n\}$$

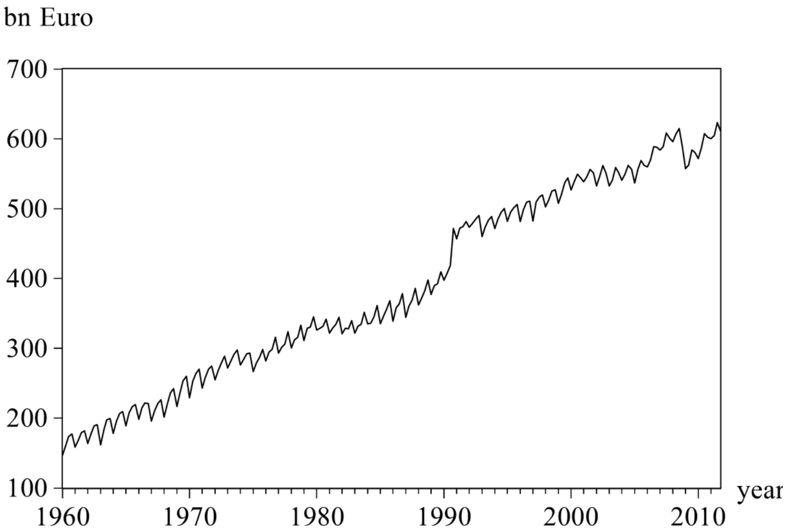


Figure 1.1: Real Gross Domestic Product of the Federal Republic of Germany in billions of Euro, 1960 – 2011

- **Transformación:** **Cambio porcentual**

$$gqr_n = 100 \cdot \ln\left(\frac{X_n}{X_{n-1}}\right)$$

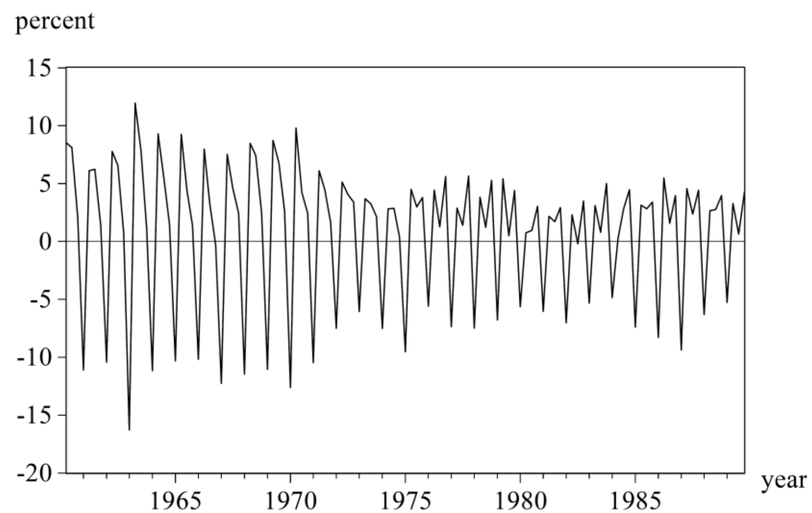


Figure 1.3: Quarterly Growth Rates of the Real Gross Domestic Product (qgr) of the Federal Republic of Germany, 1960 – 1989

- **Comentario:** Se observa un patrón de periodicidad que sufre una disrupción a partir de 1975. Aquí la amplitud es importante y se muestra decreciente en el tiempo.

DESCOMPOSICIÓN

Ejemplo 4/6

- **Datos:** Producto interno bruto de la Rep. Federal de Alemania, datos trimestrales entre 1960 y 2011:

$$X = \{X_1, X_2, \dots, X_n\}$$

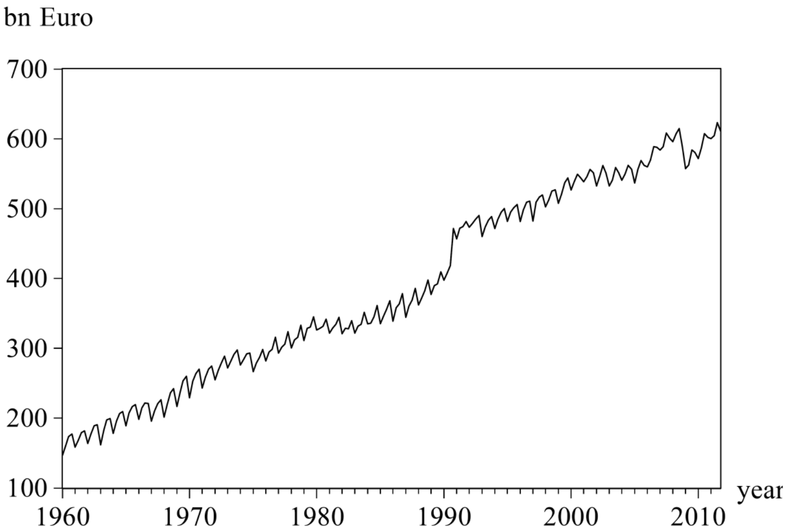


Figure 1.1: Real Gross Domestic Product of the Federal Republic of Germany in billions of Euro, 1960 – 2011

- **Transformación:** **Eliminación de periodicidad**

$$\Delta_4 X_n = X_n - X_{n-4}$$

- **Comentario:** Se usa la referencia del cambio respecto al año anterior y no respecto al trimestre anterior. Se elimina así la variación estacional. Se observa **negativo solamente en épocas de recesión** (1967, 1975, 1981 y 1982), las últimas por shocks del petróleo.

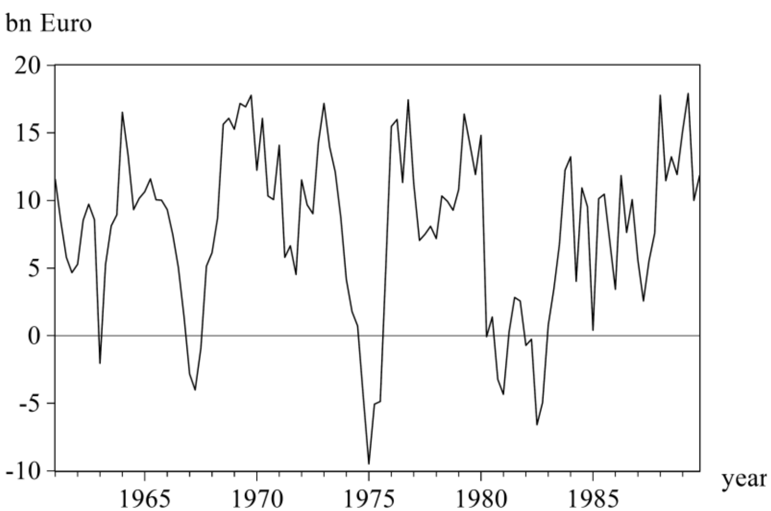


Figure 1.4: Annual Changes of the Real Gross Domestic Product (Δ_4GDP) of the Federal Republic of Germany, 1961 – 1989

DESCOMPOSICIÓN

Ejemplo 5/6

- **Datos:** Producto interno bruto de la Rep. Federal de Alemania, datos trimestrales entre 1960 y 2011:

$$X = \{X_1, X_2, \dots, X_n\}$$

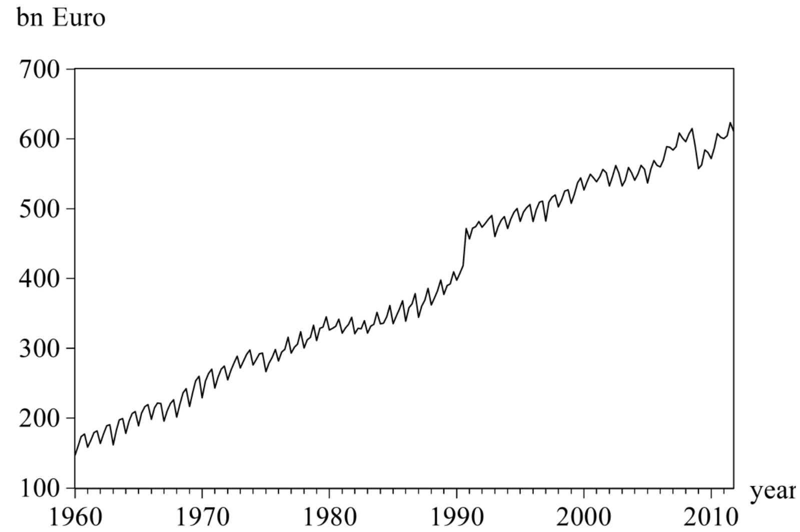


Figure 1.1: Real Gross Domestic Product of the Federal Republic of Germany in billions of Euro, 1960 – 2011

- **Transformación: Cambios anuales porcentuales**

$$agr_n = 100 \cdot \ln\left(\frac{X_n}{X_{n-4}}\right)$$

- **Comentario:** La tasa de crecimiento en los años 60 y 70 fluctúa entre **-3.5%** y **10%**. Se distingue la recesión de 1968, 1975 y de los 80's. Posteriormente a 1975, el crecimiento ha sido cada vez menor (entre **0%** y **5%**).

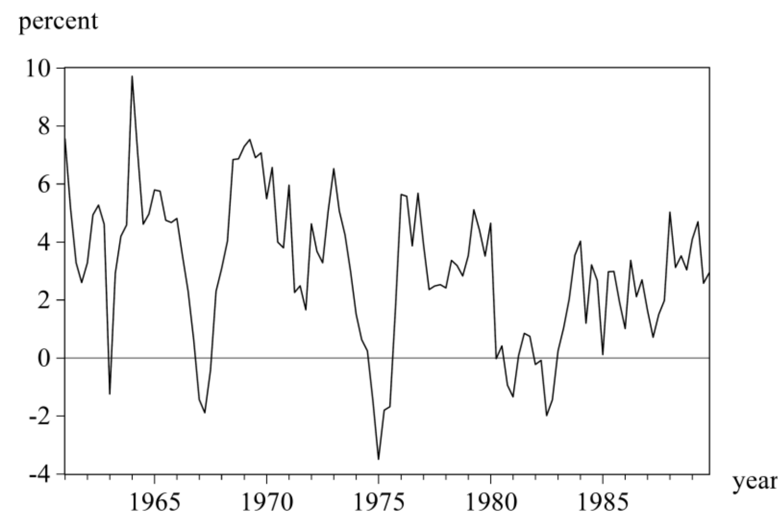


Figure 1.5: Annual Growth Rates of Real Gross Domestic Product (agr) of the Federal Republic of Germany, 1960 – 1989

DESCOMPOSICIÓN

Ejemplo 6/6

- **Transformación:** **Eliminación de estacionalidad sin afectar la tendencia**

$$GDPS_n = \frac{1}{4}(X_n + X_{n-1} + X_{n-2} + X_{n-3})$$

- **Comentario:** Este es un promedio móvil de grado 4. Tiene un rezago de 1 año.

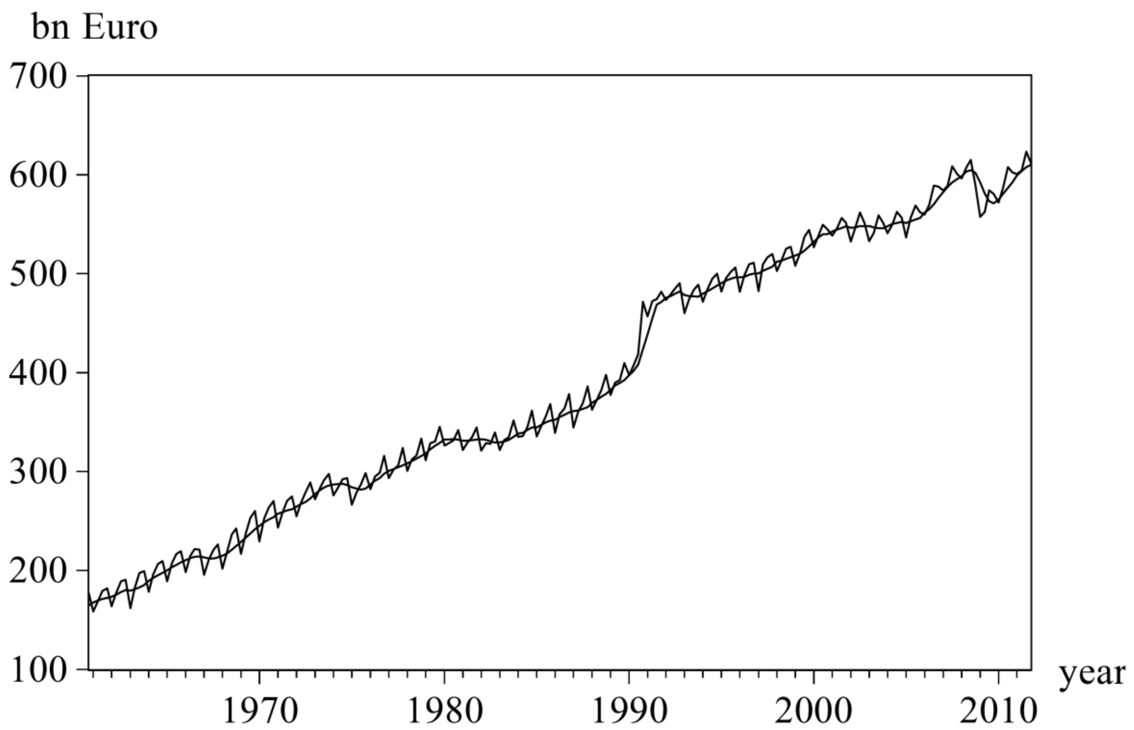


Figure 1.6: 'Smooth Component' and actual values of the Real Gross Domestic Product of the Federal Republic of Germany, 1961 – 2011

Descomposición de Wold (1/4)

- Consideremos una sucesión $\{X_0, X_1, \dots, X_t, \dots\}$ de variables aleatorias con varianza estacionaria:

$$\mathrm{Var}(X_t) = \sigma^2 \text{ para toda } t.$$

- Al substraerle la media, el proceso puede representarse por una combinación lineal de variables aleatorias no correlacionadas de media cero y varianza constante:

$$\boxed{X_t - \mu_t = \sum_{j=0}^{\infty} a_j\, u_{t-j}}$$

con **procesos estocásticos puros** u_t, los cuales cumplen:

$$\mathbb{E}[u_t] = 0 \quad \text{y} \quad \mathbb{E}[u_t u_s] = \begin{cases} \sigma^2, & \text{cuando } t = s \\ 0\ , & \text{cuando } t \neq s \end{cases}$$

y con coeficientes a_j constantes para $j = 1,2 \dots$, $a_0 = 1$ y $\sum_{j=1}^{\infty} a_j^2 < \infty$.

NOTA: Cuando $t = s$, se trata de la varianza de u_t:

$$\mathrm{Var}(u_t) \\ = \mathbb{E}[u_t^2] - \mathbb{E}[u_t]^2 \\ = \mathbb{E}[u_t^2] - 0 = \sigma^2 \\ = \mathrm{Var}(X_t)$$

Ejemplos…

Descomposición de Wold (2/4)

Valor Esperado:

- Asumiendo la descomposición de Wold para la sucesión $\{X_0, X_1, \dots, X_t, \dots\}$, se obtiene que la esperanza queda dada por:

$$\mathbb{E}[X_t - \mu_t] = \mathbb{E}\left[\sum_{j=0}^{\infty} a_j\, u_{t-j}\right] = \sum_{j=0}^{\infty} a_j\, \mathbb{E}[u_{t-j}] = 0$$

Por lo tanto

$$\boxed{\mathbb{E}[X_t] = \mu_t}$$

Conclusión: Si la serie X_t cumple la descomposición de Wold, tendría valor esperado constante (estacionario), siempre y cuando su descomposición de Wold cumpliera que $\mu_t =$ **constante** para todo tiempo t.

Descomposición de Wold (3/4)

Varianza:

- Asumiendo la descomposición de Wold para la sucesión $\{X_0, X_1, \dots, X_t, \dots\}$, se obtiene que la varianza queda dada por:

$$\mathbb{E}[(X_t-\mu_t)^2] = \mathbb{E}\left[\left(\sum_{j=0}^{\infty} a_j u_{t-j}\right)^2\right] = \mathbb{E}\left[\sum_{j=0}^{\infty}\left(a_j u_{t-j}\right)^2 + \sum_{j=0}^{\infty}\sum_{i=0, i\neq j}^{\infty}(a_i u_{t-i})(a_j u_{t-j})\right]$$

$$= \sum_{j=0}^{\infty}\mathbb{E}\left[a_j^2 u_{t-j}^2\right] + \sum_{j=0}^{\infty}\sum_{i=0, i\neq j}^{\infty}\mathbb{E}\left[a_i\, a_j\, u_{t-i}\, u_{t-j}\right] = \sum_{j=0}^{\infty} a_j^2\,\mathbb{E}\left[u_{t-j}^2\right] = \sigma^2\sum_{j=0}^{\infty} a_j^2$$

Por lo tanto, haciendo $\gamma(n) = \sigma^2 \sum_{j=0}^{\infty} a_{j+n} a_j$, resulta que $\gamma(0)$ es constante y finita. Así

$$\boxed{\mathbb{E}[(X_t-\mu_t)^2] = \gamma(0)}$$

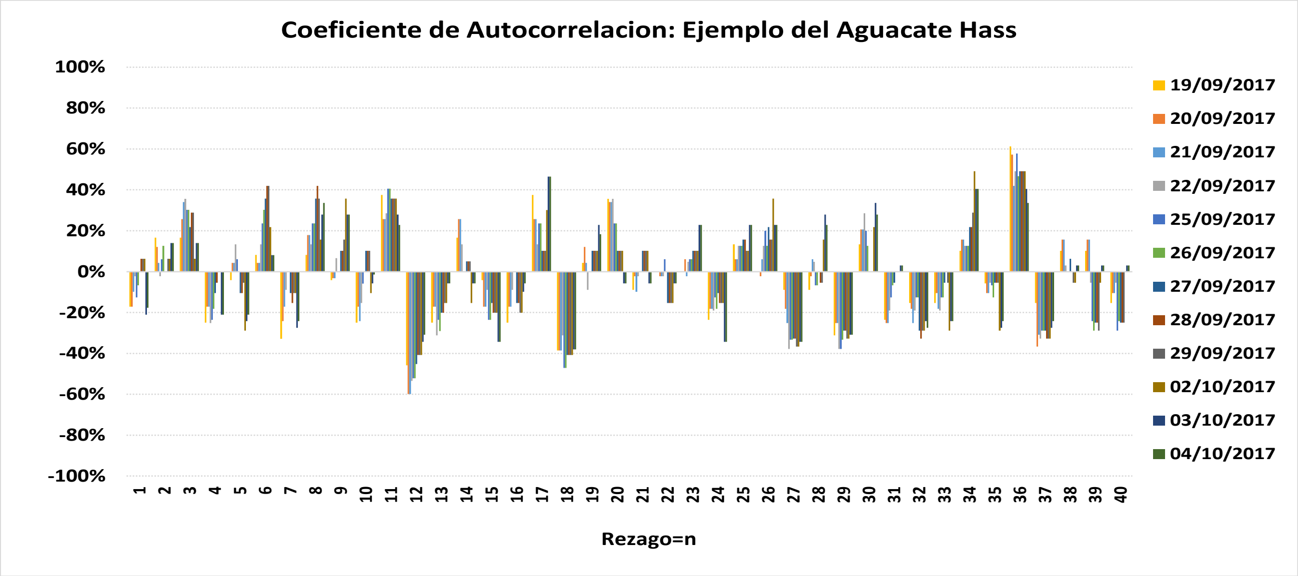
- **Conclusión:** Las series de tiempo con varianza estacionaria se simplifican enormemente mediante la descomposición de Wold asumiendo una representación mediante procesos estocásticos puros.

Descomposición de Wold (4/4)

Coeficiente de Autocorrelación:

- Asumiendo la descomposición de Wold para la sucesión $\{X_0, X_1, \dots, X_t, \dots\}$, se obtiene de manera similar la función de autocorrelación con rezago n:

$$\boxed{\mathrm{Cov}(X_t, X_{t-n}) = \gamma(n) \quad \text{y} \quad \rho_\tau = \frac{\gamma(n)}{\gamma(0)}}$$



RECESO

5 minutos ...

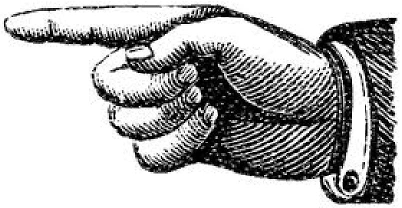
PLAN DE LA PRESENTACIÓN

PARTE 1

- **ANTECEDENTES**
- **SERIES DE TIEMPO**
- **DISTRIBUCIÓN DE SERIES DE TIEMPO**
- **CARACTERÍSTICAS DE LAS SERIES DE TIEMPO**

PARTE 2

- **DESCOMPOSICIÓN**
- **MODELOS DE SERIES DE TIEMPO**
- **EJEMPLO**



PARTE 3

- **MODELOS DE SERIES DE TIEMPO**
- **OTRAS CARACTERÍSTICAS**
- **PRONÓSTICOS**

Modelos Auto-Regresivos: AR(1)

- Consideremos una sucesión $\{X_0, X_1, \dots, X_t, \dots\}$ de datos que cumplen la siguiente relación iterativa:
$$X_t = \delta + \alpha X_{t-1} + u_t$$
con α y δ constantes y u_t un proceso aleatorio puro.

- **Definición:** A este tipo de serie de tiempo se le denomina ***Proceso Autoregresivo de primer grado***, denominado **AR(1).**

- **Comentario**: Un modelo AR(1) es una ecuación en diferencias estocástica, para la cual es de interés identificar la distribución del proceso X_t y otras de sus características, como por ejemplo si es estacionario.

Modelos Auto-Regresivos: AR(1)

- Tomando como valor inicial X_0 de la serie como dado y sustituyendo para la ecuación de el X_1 el término anterior X_0, obtenemos una relación que permite obtener una representación general para X_t:

$$\boldsymbol{X_1} = \boldsymbol{\delta} + \boldsymbol{\alpha X_0} + \boldsymbol{u_1}$$

$$\begin{aligned} X_2 &= \delta + \alpha \boldsymbol{X_1} + u_2 \\ &= \delta + \alpha(\boldsymbol{\delta} + \boldsymbol{\alpha X_0} + \boldsymbol{u_1}) + u_2 = (1+\alpha)\delta + \alpha^2 X_0 + \alpha\, u_1 + u_2 \end{aligned}$$

$$\vdots$$

$$X_t = \delta + \alpha X_{t-1} + u_t = \delta \sum_{s=0}^{t-1} \alpha^s + \alpha^t X_0 + \sum_{j=0}^{t-1} \alpha^j u_{t-j}$$

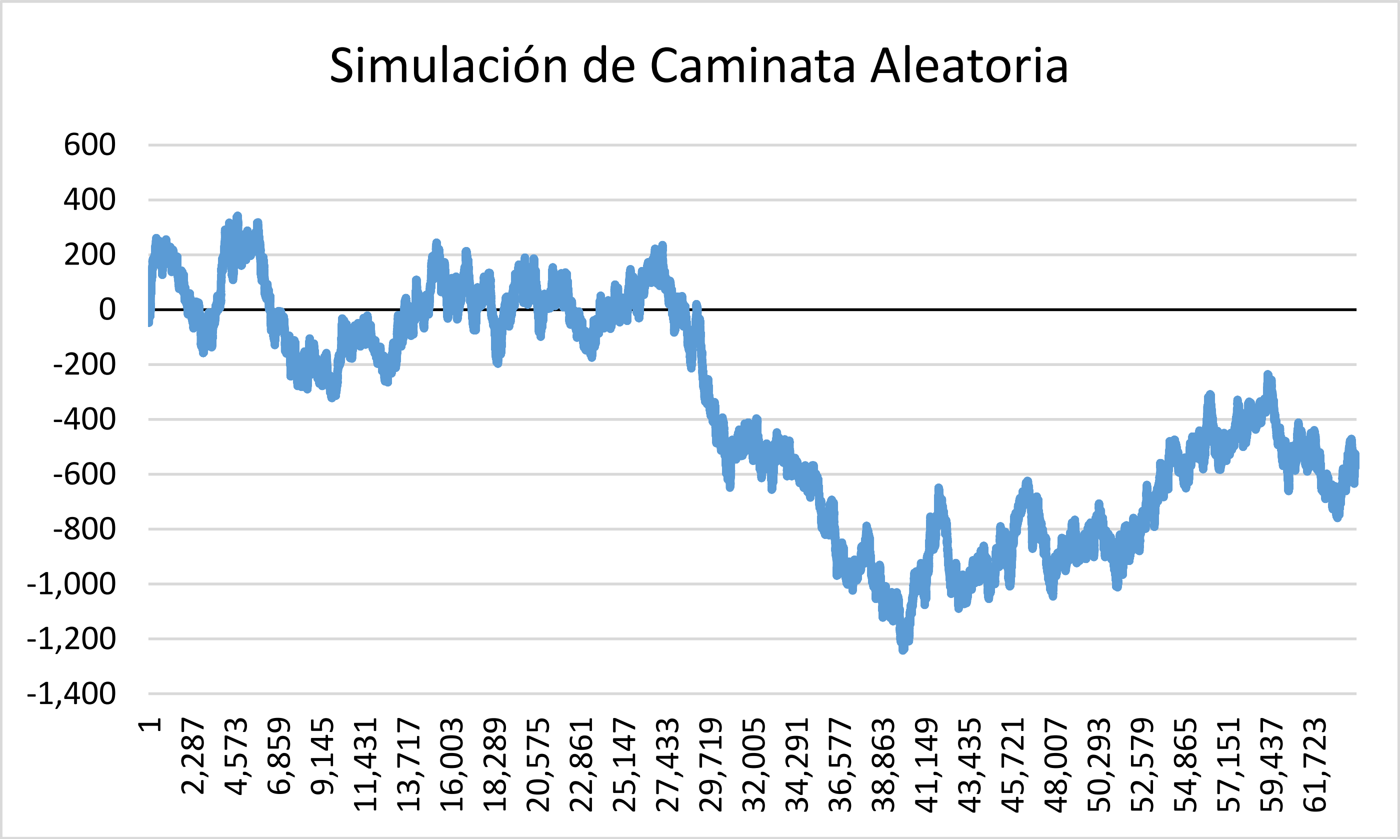
- Esta expresión puede simplificarse hasta obtener :

$$\boxed{X_t = \alpha^t X_0 + \delta \frac{1-\alpha^t}{1-\alpha} + \sum_{j=0}^{t-1} \alpha^j u_{t-j}}$$

- **Comentario:** Un proceso AR(1) no es estacionario, ya que depende de procesos estocásticos de otros periodos, obteniéndose distribuciones diferentes en el tiempo.

Modelos Auto-Regresivos: AR(1)

- **Ejemplo**: Tomando el siguiente modelo AR(1): $X_t = X_{t-1} + u_t$, con $u_t \sim N(0,1)$ y $X_t = 0$, se obtiene una caminata aleatoria.



Modelos Auto-Regresivos: AR(1)

- Tomando la expresión anterior y comenzando en el periodo $t_0 \leq 0$, para $\alpha < 1$ resulta que:

$$X_t = \alpha^{t-t_0} X_{t_0} + \delta \frac{1-\alpha^{t-t_0}}{1-\alpha} + \sum_{j=0}^{t-t_0-1} \alpha^j u_{t-j}$$

Tomando el límite cuando $t_0 \to -\infty$, tendríamos una **serie con una historia infinita**:

$$\boxed{X_t = \frac{\delta}{1-\alpha} + \sum_{j=0}^{\infty} \alpha^j u_{t-j}}$$

- **Comentarios:**

1. Este tipo de proceso AR(1) sí es estacionario, ya que la parte aleatoria tiene la misma distribución para cualquier tiempo t.
2. Coincide con la descomposición de Wold.
3. Por lo general, tomar toda la historia suele no ser fácil en la práctica.

(Modelo: $X_t = \delta + \alpha X_{t-1} + u_t$)

Modelos Auto-Regresivos: ACF para AR(1)

- La Función de Auto-correlación (**ACF**) o "Auto-Correlograma", para modelos autoregresivos AR(1), muestran diferentes comportamientos cualitativos en función del coeficiente α.

Realización

Autocorrelación de Wold

Autocorrelación Empírica

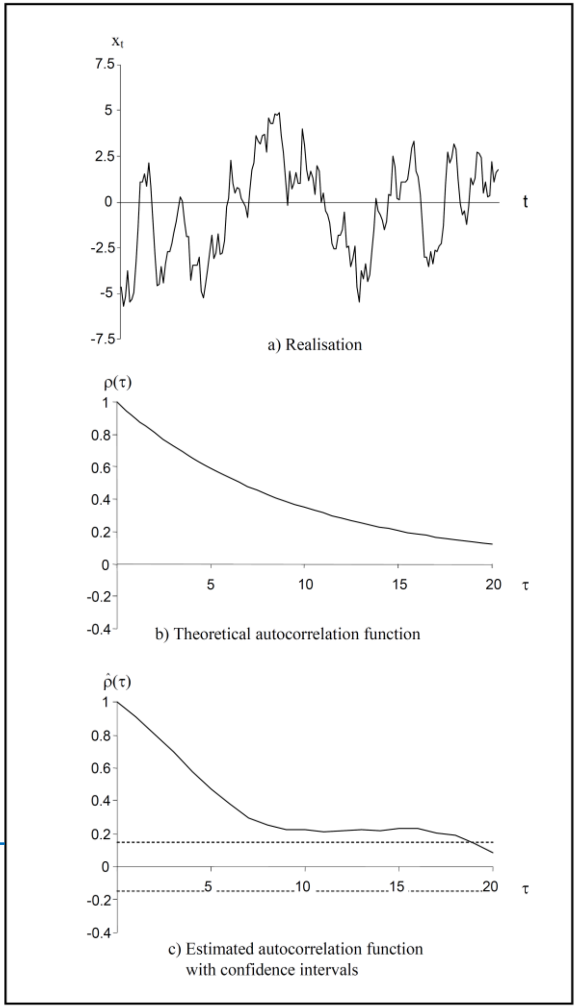


Figure 2.1: AR(1) process with α = 0.9

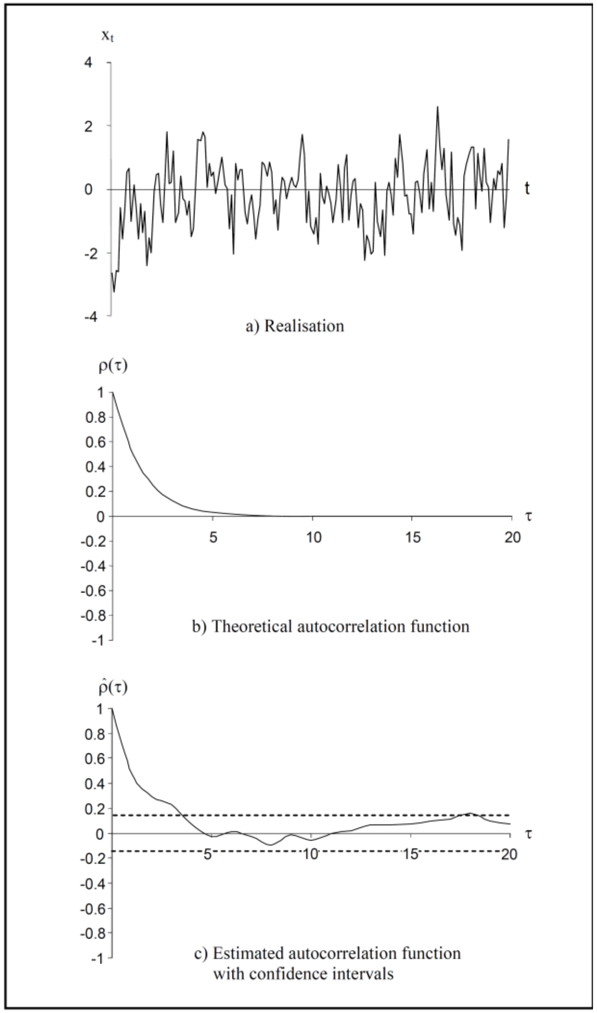


Figure 2.2: AR(1) process with α= 0.5

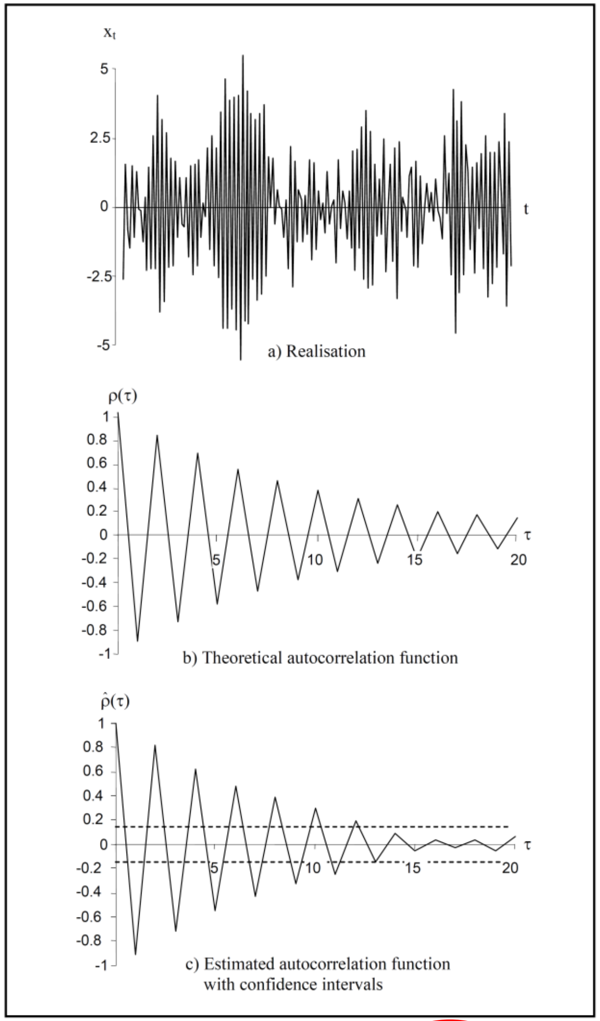


Figure 2.3: AR(1) process with α = -0.9

Fuente: [Kirchgässner Gebhard, Wolters Jürgen, Hassler Uwe]

Promedio Móvil
(MA: Moving-Average)

- Asumamos que u_t es un proceso estocástico puro i.i.d., con $\mathbb{E}[u_t] = 0$ y $\mathrm{Var}[u_t] = \sigma^2$.
- El **promedio móvil** está dado por la siguiente relación: $X_t = a_1 X_{t-1} + \cdots + a_n X_{t-n} + u_t$
- Para el caso en que la ventana de tiempo sea de $n = 2$ periodos, se obtiene que:

$$\boxed{X_t = a_1 X_{t-1} + a_2 X_{t-2} + u_t}$$

cumpliéndose lo siguiente:

$$\begin{aligned}
&X_0 \quad \text{dado} \\
&X_1 = a_1 X_0 + u_1 \\
&X_2 = a_1 X_1 + a_2 X_0 + u_2 \\
&\quad\; = a_1(a_1 X_0 + u_1) + a_2 X_0 + u_2 \\
&\quad\; = ({a_1}^2 X_0 + a_1 u_1) + a_2 X_0 + u_2 \\
&\quad\; = ({a_1}^2 + a_2) X_0 + a_1 u_1 + u_2 \\
&\quad\; = \mu - (-a_1) u_1 + u_2 = \mu + u_2 - \beta u_1
\end{aligned}$$

- De esta manera, el promedio móvil con $n = 2$ periodos de rezago se denota como **MA(2)**, cumpliendo que:

$$\boxed{X_t = \mu + u_t - \beta u_{t-1}}$$

$$\begin{cases} \mathbb{E}[X_t] = \mu \\ \mathrm{Var}[X_t] = (1 + \beta^2)\sigma^2 \\ \rho_\tau = -\beta\sigma^2 \end{cases}$$

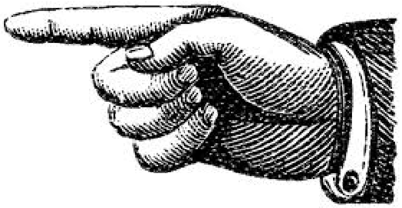
PLAN DE LA PRESENTACIÓN

PARTE 1

- ANTECEDENTES
- SERIES DE TIEMPO
- DISTRIBUCIÓN DE SERIES DE TIEMPO
- CARACTERÍSTICAS DE LAS SERIES DE TIEMPO

PARTE 2

- DESCOMPOSICIÓN
- MODELOS DE SERIES DE TIEMPO
- EJEMPLO

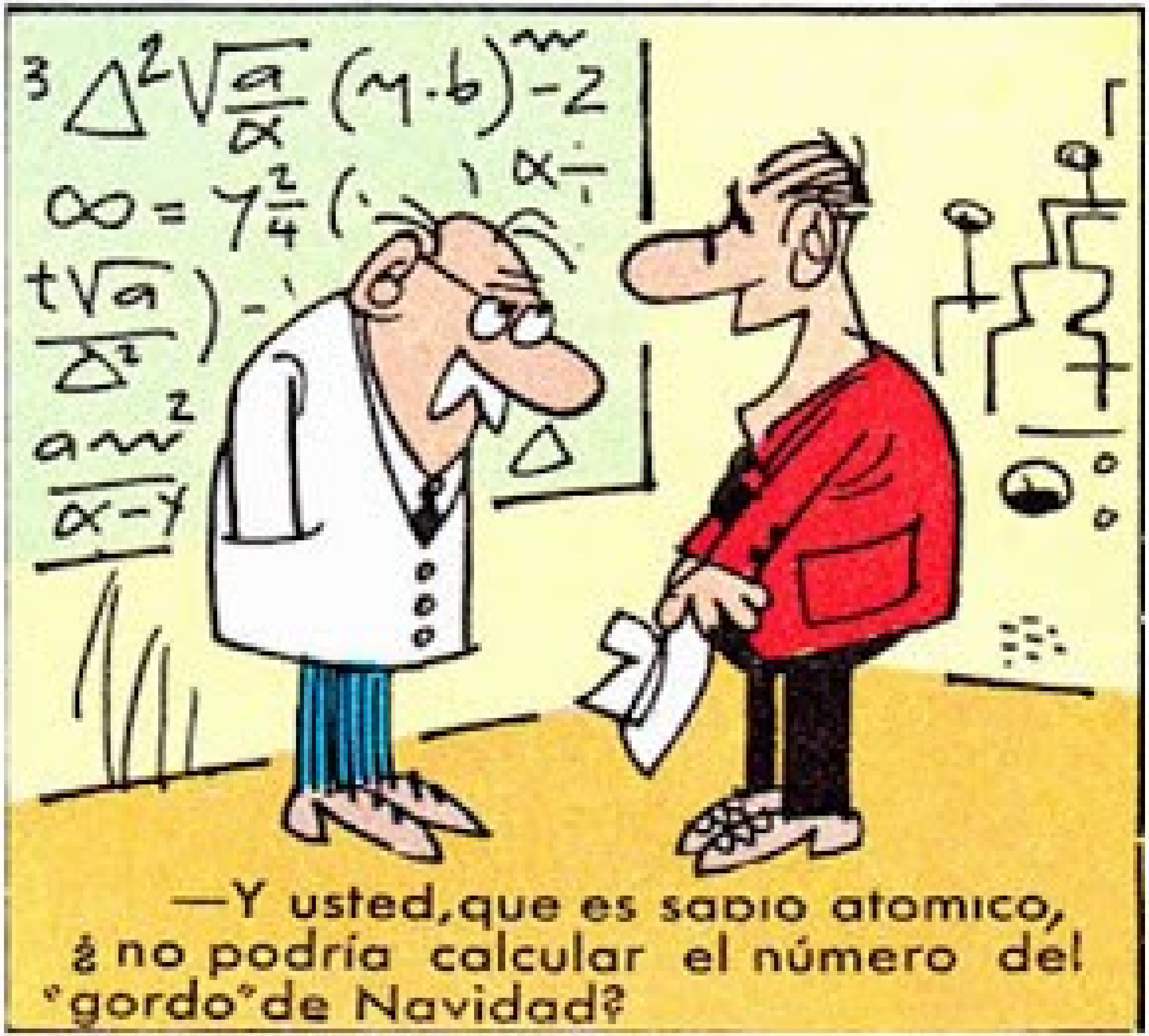


PARTE 3

- MODELOS DE SERIES DE TIEMPO
- OTRAS CARACTERÍSTICAS
- PRONÓSTICOS

EJEMPLO

Ejemplo de capacidad de predicción con series de tiempo…



Fuente: http://estadisticaavanzadaus.blogspot.mx/2015/10/comentario-noticia-que-probabilidad-hoy.html

Series de tiempo y predicción

Sornette D. y Andersen J.V.; International Journal of Modern Physics, junio/2000.

"El signo del incremento de series de tiempo no-correlacionadas puede predecirse con una probabilidad de éxito universal del 75%"

https://arxiv.org/pdf/cond-mat/0001324.pdf

Predicción: Resultado de Sornette y Andersen (1/2)

SÍNTESIS DE ALGUNOS DE LOS RESULTADOS:

- **Supuestos:** Datos X_i sin auto-correlación. Comenzaremos asumiendo X_i **i.i.d.**, $X_i \sim \mathrm{U}[0,1]$.
- **Pregunta:** Dados los valores anteriores hasta X_i, ¿cuál es el mejor predictor de X_{i+1}?
- **Incrementos:** Si consideramos la diferencia y que el valor esperado de la distribución uniforme es $1/2$:

$$\mathbb{E}[X_{i+1} - X_i] = \mathbb{E}[X_{i+1}] - \mathbb{E}[X_i] = \frac{1}{2} - \frac{1}{2} = 0$$

$$\mathbb{E}[X_{i+1} - X_i | X_i] = \mathbb{E}[X_{i+1} | X_i] - \mathbb{E}[X_i | X_i] = \frac{1}{2} - X_i$$

- **Conclusión 1**: Parece haber predictibilidad al considerar la información del periodo anterior.
- **Autocorrelación:** Buscando intuición sobre la autocorrelación de los incrementos:

$$\mathbb{E}[(X_{i+1} - X_i)(X_i - X_{i-1})]$$

$$= \mathbb{E}[X_{i+1}X_i] - \mathbb{E}[X_{i+1}X_{i-1}] - \mathbb{E}\left[{X_i}^2\right] + \mathbb{E}[X_iX_{i-1}]$$

$$= \frac{1}{4} - \frac{1}{4} - \frac{1}{2} + \frac{1}{4} = -\frac{1}{4} = -0.25$$

- **Conclusión 2**: Autocorrelación negativa habla de que habrá frecuentemente cambio de signo.
- **Estrategia:** Considerando lo anterior, la estrategia sugerida para X_{i+1} es el estimador $\hat{X}_{i+1} = X_i - \frac{1}{2}$

EJEMPLO

Predicción: Resultado de Sornette y Andersen (2/2)

- **Probabilidad de predecir la dirección:** Se establece la siguiente doble relación que cumple γ, representando el valor esperado de que coincidan los signos de la serie con el estimador propuesto.

$$\begin{cases} \gamma = \mathbb{E}\left[\operatorname{sign}(X_{i+1} - X_i)\ \operatorname{sign}\left(\frac{1}{2} - X_i\right)\right] = \frac{1}{2} \\ \gamma = \mathbb{E}\left[\operatorname{sign}(X_{i+1} - X_i)\ \operatorname{sign}\left(\frac{1}{2} - X_i\right)\right] = (+1)p_+ + (-1)p_- \end{cases}$$

- Se obtiene por un lado que $\gamma = 1/2$ tras realizar los cálculos y por otro que γ es igual a la probabilidad de coincidencia de signos (p_+) más aquella de no coincidencia de signos (p_-).

- La ecuación resultante sería:

$$p_+ - (1 - p_+) = \frac{1}{2}$$

- Despejando p_+, se concluye que la probabilidad de predecir los signos sería $p_+ = 75\%$.

- A la derecha un ejemplo de realización del proceso $(X_i)_{i=0,\dots,1500}$.