TEMARIO

1. Estadística e implementación del lenguaje R en Jupyter Notebook.
2. Introducción a las series de tiempo.
3. Estacionariedad.
4. Introducción a procesos estocásticos y procesos gaussianos.
5. Promedios móviles y autorregresivos. Modelos integrados: ARIMA.
6. Ecuaciones de Yule-Walker.
7. Pronósticos de series de tiempo (Forecasting).

¿Qué es una serie de tiempo?

Una serie temporal (o simplemente una serie de tiempo) es una secuencia de observaciones (datos) ordenadas y equidistantes cronológicamente sobre una característica (serie univariante o escalar) o sobre varias características (serie multivariante o vectorial) de una unidad observable en diferentes momentos.

Los datos se pueden comportar de diferentes maneras a través del tiempo: puede que se presente una **tendencia, estacionalidad o simplemente no presenten una forma definida**.

Debido a que los datos observvbles en una serie de tiempo se concentran en períodos de tiempo adyacentes, existe una potencial correlación entre las observaciones.

¿Qué es una serie de tiempo?

Las series de tiempo tienen aplicaciones en múltiples áreas del conocimiento, por ejemplo en economía, ciencias sociales, finanzas, epidemiología, y en la física.

Field	Example topics	Example dataset
Economics	Gross Domestic Product (GDP), Consumer Price Index (CPI), S&P 500 Index, and unemployment rates	U.S. GDP from the Federal Reserve Economic Data
Social sciences	Birth rates, population, migration data, political indicators	Population without citizenship from Eurostat
Epidemiology	Disease rates, mortality rates, mosquito populations	U.S. Cancer Incidence rates from the Center for Disease Control
Medicine	Blood pressure tracking, weight tracking, cholesterol measurements, heart rate monitoring	MRI scanning and behavioral test dataset
Physical sciences	Global temperatures, monthly sunspot observations, pollution levels.	Global air pollution from the Our World in Data

¿Qué es una serie de tiempo?

Ejemplo

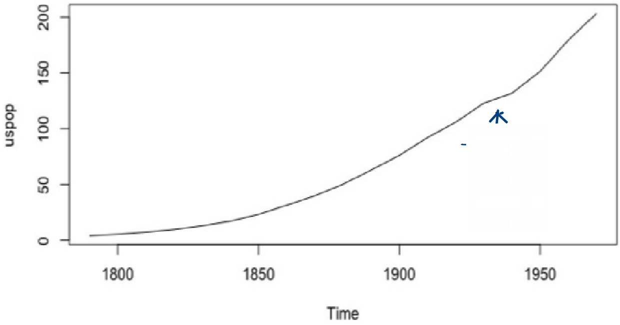


Figura: Tamaño de la población (en millones de habitantes) en USA entre 1790-1990

¿Qué es una serie de tiempo?

Ejemplo

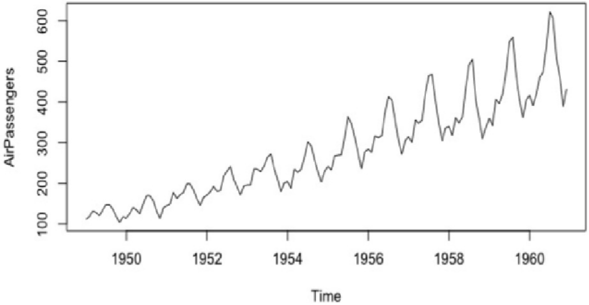


Figura: Número mensual de pasajeros (en millones) entre 1949-1960 en líneas aéreas.

¿Qué es una serie de tiempo?

Ejemplo

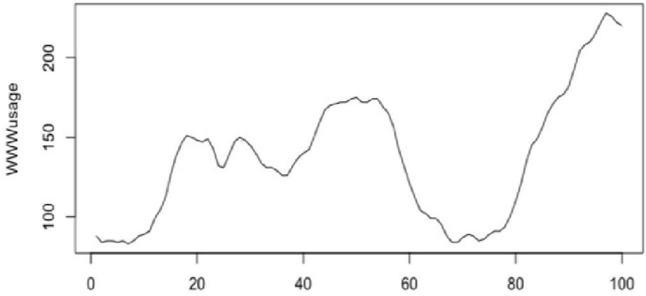


Figura: Número de usuarios conectados a internet a través de un servidor todos los minutos.

¿Qué es una serie de tiempo?

Ejemplo

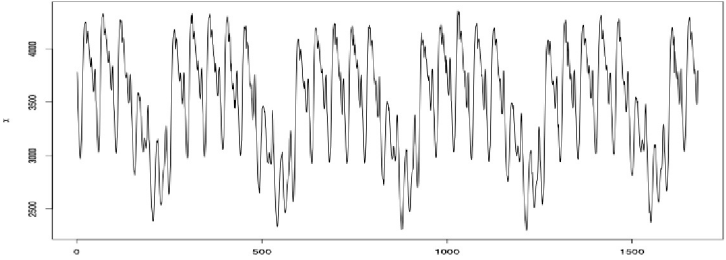


FIGURA: Consumo de electricidad en un país

¿Qué es una serie de tiempo?

Ejemplo

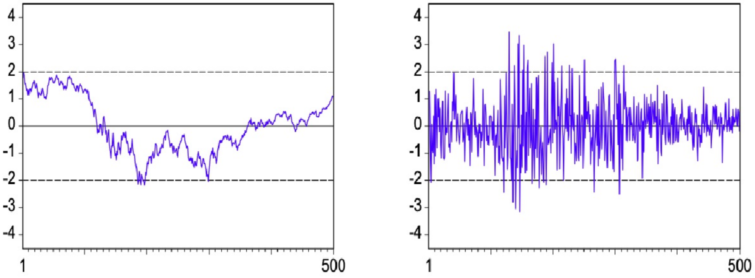


Figura: Izq: IBEX (log) al cierre en la Bolsa de Madrid 2 enero 2002-30 dic de 2003. Der. Tasa logarítmica de variación diaria del IBEX 35 3 de enero de 2002 – 30 de diciembre de 2003.

Ejemplo

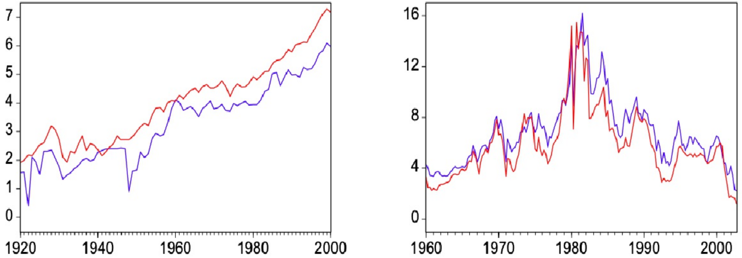


Figura: Izq: Índices bursátiles anuales en Ale y USA 1920-2000. Der. Tipos de interés de la deuda pública en USA 1960:I-2002:IV.

Componentes en una TS:

- La **tendencia** es el movimiento de los datos hacia arriba o hacia abajo a lo largo del tiempo. También, ocurre que los datos se mantienen estables, esto significa que las ventas no aumentan ni disminuyen conforme pasa el tiempo.
- La **estacionalidad** se identifica como el patrón que muestran los datos en intervalos regulares, por encima o por debajo de la estación promedio. Una estación con un factor estacional igual a uno de interpreta como una estación promedio; una estación con factor estacional mayor que uno se interpreta como una estación por encima del promedio y, una estación con un factor estacional menor que uno se interpreta como una estación por debajo del promedio.
- La **ciclicidad** son los patrones que se identifican en ciertos intervalos de tiempo, se asocia la ciclicidad al ciclo económico.
- Las **variaciones aleatorias (conocido como error)** son irregularidades que se suponen explica el azar. No muestran un patrón y presentan una distribución normal con media igual a cero.

Recuento de conceptos de probabilidad

X variable aleatoria $\rightsquigarrow$ conocer $\underbrace{F_X(x) = \mathbb{P}(X \leq x)}_{\text{Función de distribución}}$ $x \in \mathbb{R}$

distribución o ley

- Discretas
  $\mathbb{P}(X=x)>0$
- Continuas $F_X(x) = \int_{-\infty}^{x} \underbrace{f_X(y)}_{\text{función de densidad}} dy$
  $\mathbb{P}(X=x)=0$ ¿?

$\begin{cases} \text{Bernoulli} \\ \text{Binomial (Neg)} \\ \text{Geométrica} \\ \text{Poisson} \end{cases}$

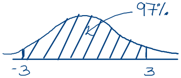
Exponencial
Gamma
Pareto
$\boxed{\text{Uniforme}}$ genera v.a.
$\boxed{\text{Normal o gaussiana}}$

- $\mathcal{N}(\mu, \sigma^2)$
- $\mathcal{N}(0,1)$

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right\}$$

$Z \sim \mathcal{N}(0,1)$

$$\mathbb{P}(Z \leq z) = \Phi(z)$$



Nota. Dada X v.a

$$\mathbb{E}[X^n] = \begin{cases} \sum_{x \in R_X} x^n \mathbb{P}(X=x) & X \text{ discreta} \\ \int_{\mathbb{R}} x^n f_X(x)\,dx & X \text{ continua} \end{cases} \qquad n \geq 1$$

$$Var[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

$\{X_t\}_{\underbrace{t \in I}_{\text{tiempo}}}$ sucesión de v.a

Si $X_1, X_2, X_3, \ldots, X_n, \ldots$ v.a independientes e idénticamente distribuidas $(\mathbb{E}[X^n] < \infty)$

$$S_n = X_1 + \ldots + X_n$$

$$\frac{S_n}{n} \xrightarrow{\mathbb{P}} \mathbb{E}[X_i] \quad \text{LGN}$$

$$\boxed{\mathbb{P}\left(\frac{S_n - \mathbb{E}[S_n]}{\sqrt{Var[S_n]}} \leq x\right) \xrightarrow[n\to\infty]{} \Phi(x)}$$

Teorema central del límite.

"Si no se conoce el comportamiento de una sucesión de datos (a gran escala) los podemos considerar como normales"

Definición.- (Serie de tiempo) Una serie de tiempo es una colección $\{X_t\}_{t\in I}$ de v.a.s con alguna característica común.

Recordemos que existe una potencial correlación entre las vas

Definición.- Sea $\{X_t\}$ una st con $\mathbb{E}[X_t^2]<\infty$

- La función de medias de $\{X_t\}$

$$\mu_X(t)=\mathbb{E}[X_t]$$

- La función de covarianza de $\{X_t\}$

$$\gamma_X(r,s)=Cov(X_r,X_s)=\mathbb{E}\left[(X_r-\mu_X(r))(X_s-\mu_X(s))\right]$$

Definición.- Decimos que $\{X_t\}$ (st) es estacionaria si

1.- $\mu_X(t)$ no depende de t

2.- $\gamma_X(t+h,t)$ no depende de t para cada $h>0$.

Ruido blanco (inovaciones)

Si $\{X_t\}$ es una sucesión de v.a.s no correlacionadas con $\mathbb{E}[X_t]=0$ y $Var[X_t]=\sigma^2$, entonces a $\{X_t\}$ le llamaremos ruido blanco

Una caminata aleatoria

$\{S_t : t=0,1,2,\ldots\}$ $S_0=0$

$$S_t = X_1+X_2+\ldots+X_t \quad t=1,2,\ldots$$

donde $\{X_t\}$ es un ruido blanco.

- $\mathbb{E}[S_t]=\mathbb{E}[X_1+\ldots+X_t]=\mathbb{E}[X_1]+\ldots+\mathbb{E}[X_t]=0$
- $\gamma_S(t+h,t)=Cov(S_{t+h},S_t)$ $\quad S_{t+h}=\underbrace{X_1+\ldots+X_t}_{S_t}+X_{t+1}+\ldots+X_{t+h}$

$$= Cov(S_t+X_{t+1}+\ldots+X_{t+h},S_t)$$

Cov bilineal

$$= Cov(S_t,S_t)+\underbrace{Cov(X_{t+1},S_t)+\ldots+Cov(X_{t+h},S_t)}_{0}$$

$$= Cov(S_t,S_t)$$

$$= Var[S_t]=t\sigma^2$$

$\gamma_S(t+h,t)=t\sigma^2$ Por lo tanto no es estacionaria.

Nota. $S_t = X_1+\ldots+X_{t-1}+X_t$

$S_{t-1}=X_1+\ldots+X_{t-1}$

$S_t-S_{t-1}=X_t \quad t\geq 1$

$\Delta S_t = X_t$

Definición. Sea $\{X_t\}$ una st estacionaria

La función de autocovarianza (acf) de $\{X_t\}$ en el lag h es

$$\gamma_X(h)=Cov(X_{t+h},X_t) \qquad \gamma_X(0)=Cov(X_t,X_t)=Var(X_t)$$

La función de autocorrelación (acf) de $\{X_t\}$ en el lag h es

$$\boxed{\rho_X(h)=\frac{\gamma_X(h)}{\gamma_X(0)}}$$

Va a ser muy importante a lo largo del curso.

Nota. Existe una forma de volver estacionaria a cualquier serie de tiempo (remover tendencia).

Ejemplo Proceso MA(1) Moving-Averge of First-order

$X_t = Z_t+\theta Z_{t-1}$ $\quad$ Z_t es ruido blanco. $\theta\in\mathbb{R}$

precio al día de hoy

- $\mathbb{E}[X_t]=\mathbb{E}[Z_t]+\theta\,\mathbb{E}[Z_{t-1}]=0$
- $\mathbb{E}[X_t^2]=Var[X_t]=Var[Z_t]+\theta^2 Var[Z_{t-1}]$

$$=\sigma^2+\theta^2\sigma^2=\sigma^2(1+\theta^2)<\infty$$

- $Cov(X_{t+h},X_t)=\mathbb{E}[X_{t+h}X_t]$

$\gamma_X(t+h,t)$

$$=\mathbb{E}\left[(Z_{t+h}+\theta Z_{t+h-1})(Z_t+\theta Z_{t-1})\right]$$

$$=\mathbb{E}[Z_{t+h}Z_t]+\theta\,\mathbb{E}[Z_{t+h}Z_{t-1}]+$$

$$\theta\,\mathbb{E}[Z_{t+h-1}Z_t]+\theta^2\,\mathbb{E}[Z_{t+h-1}Z_{t-1}]$$

$$=\begin{cases}\sigma^2(1+\theta^2) & h=0\\ \sigma^2\theta & h=1,\ h=-1\\ 0 & \text{en otro caso}\end{cases}$$

MA(1) es estacionario.

$$\boxed{\rho_X(h)=\begin{cases}1 & h=0\\ \theta/1+\theta^2 & h=1, h=-1\\ 0 & \text{en otro caso}\end{cases}}$$

¿Cómo remover la tendencia? (Modelo de caminata aleatoria)

$X_t = X_{t-1}+Z_t \longrightarrow \boxed{Z_t = X_t - X_{t-1}}$

$\downarrow$

$\nabla X_t = X_t - X_{t-1}$

$\boxed{\text{Obtenemos un proceso puramente aleatorio}} \longleftarrow \boxed{Z_t=\nabla X_t \quad \text{operador diferencial}}$

Se elimina la tendencia.