



Análisis de series de tiempo

Técnicas básicas y nuevas tendencias

Dr. Raúl Salgado García

Centro de Investigación en Ciencias - UAEM
Instituto Potosino de Investigación Científica y Tecnológica A. C.

07 de julio de 2025

**Escuela Nacional de Supercómputo:
07-11 de julio de 2025.**

Contenido temático

Día 1 **Introducción al análisis de series temporales**

Día 2 **Procesos estocásticos**

Día 3 **Técnicas básicas de análisis**

Día 4 **Métodos basados en entropía**

Día 1

Introducción al análisis de series temporales

Contenido

1. Introducción: ¿Qué son las series de tiempo?
2. Métodos de análisis de series temporales
3. Fundamentos de probabilidad

Introducción

¿Qué son las series de tiempo?

Introducción

Una **serie de tiempo** es un conjunto de mediciones de alguna variable **cuantificable** de un fenómeno observado en momentos sucesivos y normalmente equidistantes. Se utiliza para analizar y comprender cómo varía una variable a lo largo del tiempo. Las series de tiempo pueden ser unidimensionales (una sola variable) o multidimensionales (varias variables), y son comunes en campos como la economía, la meteorología, la ingeniería y las finanzas.

PÁGINA PRINCIPAL > .INX · ÍNDICE

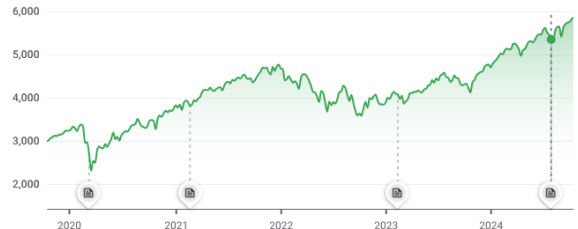
S&P 500

5,859.56 ↑ 96.22% +2,873.36 5 años

14 oct, 2:19:49 p.m. UTC-4 · INDEXSP · Renuncia de responsabilidad

1 d. 5 d. 1 m. 6 m. YTD 1 a. 5 a. MÁX.

× Eventos clave



Introducción

Características de una serie de tiempo:

- **Secuencialidad:** Los datos están organizados cronológicamente.
- **Dependencia temporal:** Las observaciones pueden depender entre sí a lo largo del tiempo.
- **Estacionalidad:** Muchas series de tiempo muestran patrones que se repiten en intervalos regulares, como estacionalidades mensuales o anuales.
- **Tendencia:** Pueden mostrar una dirección general a largo plazo (ascendente o descendente).

Introducción

Ejemplos:

1. Finanzas:

- Precios de acciones: El precio diario de una acción a lo largo del tiempo.
- Índices bursátiles: Por ejemplo, el valor del SP 500 o el *Dow Jones Industrial Average* a lo largo de los meses o años.

2. Economía:

- Producto Interno Bruto (PIB): El PIB trimestral o anual de un país a lo largo del tiempo.
- Tasa de desempleo: La tasa de desempleo mensual o anual.

3. Meteorología:

- Temperaturas diarias: Registro de temperaturas diarias en una ciudad a lo largo de los años.
- Precipitaciones mensuales: La cantidad de lluvia registrada cada mes durante varios años.

4. Ventas y Marketing.

- Ventas mensuales: Ventas mensuales de un producto o servicio.

Introducción

Ejemplos (continuación):

5. **Visitantes a un sitio web:**

- Número de visitantes diarios o mensuales a un sitio web.

6. **Salud:**

- Casos de enfermedades: Registro diario de casos de una enfermedad (como la gripe) a lo largo de una temporada.
- Tasa de mortalidad: Tasa de mortalidad anual en una población.

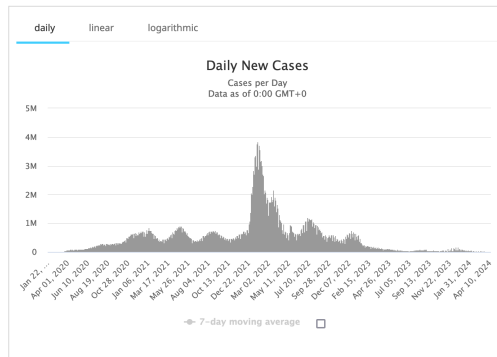
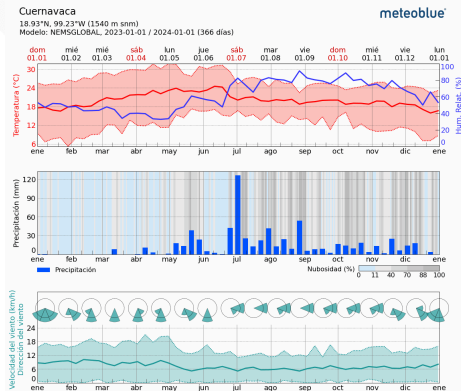
7. **Producción y Operaciones:**

- Producción diaria: Cantidad de unidades producidas en una fábrica cada día.
- Inventario mensual: Niveles de inventario de un producto a lo largo de los meses.

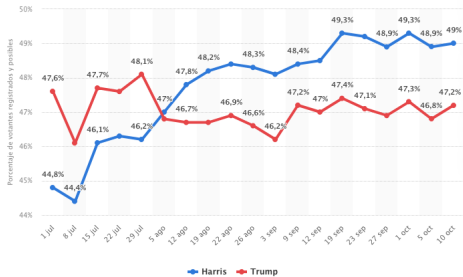
8. **Redes Sociales:**

- Interacciones diarias: Número de “me gusta”, comentarios o compartidos en una publicación a lo largo del tiempo.
- Crecimiento de seguidores: Número de seguidores en una cuenta de redes sociales cada mes.

Métodos



Métodos



Información adicional

© Statista 2024

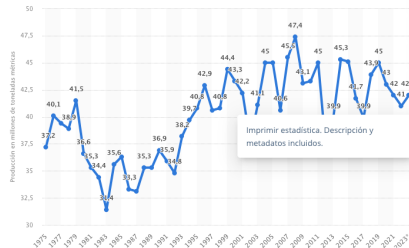
Ver fuente

Productos químicos y materias primas > Productos minerales no metálicos

Producción de sal en Estados Unidos de 1975 a 2023

(en millones de toneladas métricas)

Estadística ampliable: Seleccione la sección del diagrama que desee ampliar.



Información adicional

© Statista 2024

Ver fuente

Métodos

¿Por qué realizar análisis de series de tiempo?

Métodos

¿Por qué realizar análisis de series de tiempo?

Objetivos de los análisis de series temporales

- **Pronóstico.** Para predecir el comportamiento de alguna observación de interés. Como pronósticos de clima o predicción de la tendencia de ventas.
- **Clasificación.** Buscamos clasificar las señales observadas para decidir a qué categoría pertenece. Diagnóstico clínico basado en series fisiológicas (ECG, EEG, nivel de glucosa en la sangre, etc.)
- **Fundamentos.** Para realizar inferencia sobre los mecanismos fundamentales de una serie observada y comprender la naturaleza del fenómeno en cuestión. Verificar o rechazar hipótesis sobre algún sistema, e.g., la hipótesis débil de mercados eficientes; o para entender los mecanismos de propagación de la covid-19.

Métodos

Métodos elementales:

1. Descomposición de Series de Tiempo

Este enfoque descompone una serie de tiempo en componentes subyacentes que ayudan a comprender mejor su estructura. La descomposición puede ser aditiva (cuando los componentes se suman) o multiplicativa (cuando los componentes se multiplican).

- **Tendencia (Trend):** Componente que refleja el movimiento general hacia arriba o hacia abajo a lo largo del tiempo.
- **Estacionalidad (Seasonality):** Componente cíclico que ocurre a intervalos regulares de tiempo (por ejemplo, patrones diarios, mensuales, anuales).
- **Ruido (Residual/Irregular):** Componente que representa las fluctuaciones aleatorias no explicadas por la tendencia o la estacionalidad.

Métodos

Métodos elementales (continuación):

2. **Medias Móviles (Moving Average):** Las medias móviles se utilizan para suavizar los datos y destacar las tendencias subyacentes eliminando fluctuaciones de corto plazo.
 - **Media Móvil Simple (SMA):** Calcula la media de los valores en una ventana móvil de tiempo para suavizar los datos y reducir la volatilidad.
 - **Media Móvil Ponderada (WMA):** Este método es similar al SMA pero pondera con mayor peso los valores recientes dentro de la ventana para reflejar mejor las tendencias recientes.
 - **Media Móvil Exponencial (EMA):** Aplica un factor de suavización exponencial que prioriza los datos más recientes.

Métodos

Métodos elementales (continuación):

3. **Modelos Autorregresivos (AR)** (Hamilton, 1994; Lütkepohl, 2005): Los modelos autorregresivos (AR) utilizan observaciones pasadas de la serie de tiempo para predecir los valores futuros. Es útil cuando los valores futuros dependen de sus propios valores anteriores. Un *modelo autorregresivo de orden p* usa las p observaciones previas para predecir el valor actual.
4. **Modelos de Media Móvil (MA)** (Lütkepohl, 2005): Un modelo de media móvil (MA) utiliza los errores (residuales) pasados de la serie de tiempo para predecir los valores futuros. Este enfoque es útil para series que presentan patrones de error autocorrelacionados.

Métodos

Otros métodos:

1. **Modelos Autorregresivos Integrados de Media Móvil (ARIMA)** (Lütkepohl, 2005): ARIMA es uno de los modelos más poderosos para el análisis y predicción de series de tiempo. Combina los componentes de un modelo autorregresivo (AR) y un modelo de media móvil (MA), junto con la diferenciación para manejar la no estacionariedad.
2. **Modelos de Suavización Exponencial (Exponential Smoothing)** (Hyndman *et al.*, 2008): El método de Suavización Exponencial Simple, por ejemplo, asigna un peso decreciente exponencialmente a los datos pasados. Es útil para datos sin tendencia ni estacionalidad.
3. **Modelos SARIMA (Seasonal ARIMA)** (Lütkepohl, 2005): El modelo SARIMA es una extensión de ARIMA que incluye un componente estacional para manejar patrones que se repiten periódicamente.

Métodos

Otros métodos (continuación):

4. **Modelos ARCH/GARCH (Autoregressive Conditional Heteroskedasticity)** (Engle; Patton, 1995) Estos modelos se utilizan principalmente en el análisis financiero para modelar la volatilidad. Los modelos ARCH y GARCH capturan la heterocedasticidad condicional, es decir, cuando la varianza de los errores no es constante en el tiempo.
5. **Modelos de Espacio de Estados y Filtros de Kalman** (Hamilton, 1989): Estos modelos son útiles para sistemas dinámicos que evolucionan con el tiempo. Los modelos de espacio de estados representan una serie de tiempo como un sistema de ecuaciones de estado, y el Filtro de Kalman es una técnica recursiva para estimar los estados no observados. Son ampliamente usados en robótica, control de sistemas, y series de tiempo con componentes ocultos.

Métodos

Otros métodos (continuación):

6. **Modelos de Aprendizaje Automático (Machine Learning)** (Taylor; Letham, 2018; Chung *et al.*, 2014): Los métodos basados en machine learning se están utilizando cada vez más para el análisis y pronóstico de series de tiempo. Algunos algoritmos comunes incluyen Análisis de Regresión Lineal, Redes Neuronales Recurrentes (RNN) y LSTM así como Modelos de Árboles de Decisión
7. **Análisis de Fourier y Ondeletas (Wavelet)** (Bracewell, 2000; Mallat, 1999): Estos métodos transforman una serie de tiempo del dominio temporal al dominio de la frecuencia, lo que permite analizar ciclos y patrones repetitivos en la serie de tiempo. Las Transformadas de Fourier y las Transformadas de Ondeletas son útiles para estudiar componentes periódicos y estacionales.

Métodos

Elección del Método

La elección del método de análisis de series de tiempo depende de varios factores, como el tipo de patrón en los datos (tendencia, estacionalidad), la necesidad de predicción o simplemente el análisis de comportamiento, y el grado de complejidad que estés dispuesto a asumir.

Elección del método

- **Los métodos más simples de análisis** de series de tiempo son los métodos de suavización y descomposición.
- **Para pronósticos más complejos**, los modelos ARIMA y SARIMA son ideales, mientras que para series muy volátiles o financieras, los modelos ARCH/GARCH son recomendables.
- **Los métodos de machine learning** son más apropiados cuando se cuenta con grandes volúmenes de datos y se buscan patrones más complejos.

Métodos

Métodos avanzados

- Modelos de Estados Espacio y Filtro de Kalman (Hamilton, 1989)
- Modelos VAR (Vectores Autorregresivos) (Durbin; Koopman, 2012)
- Modelos de Cambio de Régimen (*Markov Switching Models*) (Rasmussen; Williams, 2006).
- Modelos de Procesos Gaussianos (Hochreiter; Schmidhuber, 1997).
- Modelos de Redes Neuronales Convolucionales (CNN) aplicadas a Series de Tiempo (Taylor; Letham, 2018).

Probabilidad

Fundamentos de probabilidad

Probabilidad

Experimento aleatorio

Definición: (Experimento aleatorio)

Decimos que un experimento es **aleatorio** si al ejecutarse repetidas veces el resultado no es siempre idéntico. Cada ejecución del experimento se denomina **realización del experimento aleatorio**.

- El conjunto de todos los posibles resultados de un experimento aleatorio se le denomina **espacio muestral** y se denota por Ω .
- Una colección A de posibles resultados se le denomina **evento** y se suele indicar como $A \subset \Omega$.
- El resultado ω de una realización de un experimento aleatorio se le llama **suceso** y se denota como $\omega \in \Omega$.

Probabilidad

Álgebra de eventos

- Los eventos son representaciones de los “objetos” a los que podemos asignar probabilidad.
- Los eventos se pueden componer entre sí para genera nuevos eventos a partir de operaciones de unión, intersección y complementación.
- Las reglas para generar nuevos eventos nos permiten definir una colección de eventos, denominada σ -**álgebra**, a los que podemos asignar probabilidad.

Probabilidad

Álgebra de eventos

Definición: (σ -álgebra)

Una colección \mathcal{F} de subconjuntos de Ω es una σ -álgebra si cumple las siguientes condiciones:

- i. $\Omega \in \mathcal{F}$
- ii. Si $A \in \mathcal{F}$, entonces $A^c \in \mathcal{F}$.
- iii. Si $A_1, A_2, \dots \in \mathcal{F}$, entonces

$$\bigcup_{j \in \mathbb{N}} A_j \in \mathcal{F}.$$

A la pareja (Ω, \mathcal{F}) se le llama **espacio medible** y a los elementos de \mathcal{F} se les llama *eventos* o *conjuntos medibles*.

Probabilidad

Medida de probabilidad

Definición: (Medida de probabilidad)

Sea (Ω, \mathcal{F}) un espacio medible. Una medida de probabilidad es una función $P : \mathcal{F} \rightarrow [0, 1]$ que satisface

- i. $P(\Omega) = 1$.
- ii. $P(A) \geq 0$, para cualquier $A \in \mathcal{F}$
- iii. Si $A_1, A_2, \dots \in \mathcal{F}$, son disjuntos a pares, esto es, si $A_n \cap A_m = \emptyset$ para $n \neq m$, entonces:

$$P\left(\bigcup_{j \in \mathbb{N}} A_j\right) = \sum_{j \in \mathbb{N}} P(A_j).$$

Al espacio muestral Ω dotado de una σ -álgebra \mathcal{F} y una medida de probabilidad P , (Ω, \mathcal{F}, P) , se le llama **espacio de probabilidad**.

Probabilidad

Medida de probabilidad

Proposición:

Sea (Ω, \mathcal{F}, P) un espacio de probabilidad. Entonces:

1. $P(\emptyset) = 0$.
2. Si $A_1, A_2, \dots, \in \mathcal{F}$, son disjuntos a pares, entonces

$$P\left(\bigcup_{j=1}^n A_j\right) = \sum_{j=1}^n P(A_j).$$

3. $P(A^c) = 1 - P(A)$.
4. Si $A \subset B$, entonces $P(B - A) = P(B) - P(A)$.
5. Si $A \subset B$, entonces $P(A) \leq P(B)$.
6. $0 \leq P(A) \leq 1$.
7. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
8. $P(A \cup B) \geq P(A) + P(B)$.

Variables aleatorias

Definición de variable aleatoria

Definición: (Variable aleatoria)

Sea (Ω, \mathcal{F}, P) un espacio de probabilidad. Una **variable aleatoria real** (o simplemente, *variable aleatoria*) se define como una función $X : \Omega \rightarrow \mathbb{R}$ tal que para cualquier conjunto B en la σ -álgebra de Borel, se cumple que el conjunto $X^{-1}(B)$ es un elemento de \mathcal{F} .

En lo sucesivo abreviaremos “variable aleatoria” por **v.a.**

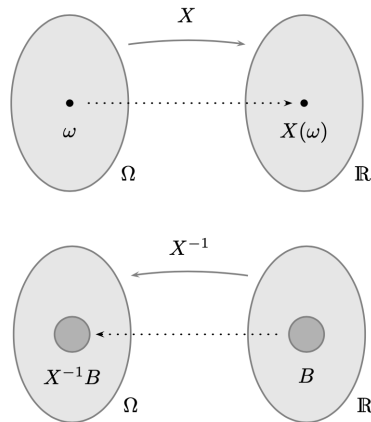
Variables aleatorias

Definición de variable aleatoria

Definición: (Variable aleatoria)

Sea (Ω, \mathcal{F}, P) un espacio de probabilidad. Una **variable aleatoria real** (o simplemente, *variable aleatoria*) se define como una función $X : \Omega \rightarrow \mathbb{R}$ tal que para cualquier conjunto B en la σ -álgebra de Borel, se cumple que el conjunto $X^{-1}(B)$ es un elemento de \mathcal{F} .

En lo sucesivo abreviaremos “variable aleatoria” por **v.a.**



Variables aleatorias

Función de distribución

Definición: (Función de distribución)

La función de distribución de una variable aleatoria X es la función $F(x) : \mathbb{R} \rightarrow [0, 1]$, definida como:

$$F(x) := P(X \leq x).$$

Proposición:

Sea $F(x)$ la función de distribución de una variable aleatoria. Entonces

1. $\lim_{x \rightarrow +\infty} F(x) = 1.$
2. $\lim_{x \rightarrow -\infty} F(x) = 0.$
3. Si $x_1 \leq x_2$, entonces $F(x_1) \leq F(x_2).$
4. $F(x)$ es continua por la derecha, es decir, $F(x^+) = F(x).$

Variables aleatorias

Variable aleatoria discreta

Definición: (Variable aleatoria discreta)

Una variable aleatoria X se le llama discreta si su correspondiente función de distribución $F(x)$ es una función constante por tramos. Sea $D = \{x_1, x_2, \dots\}$ el conjunto de todos los puntos de discontinuidad de $F(x)$. En cada uno de estos puntos el tamaño de la discontinuidad es $P(X = x_i) = F(x_i) - F(x_i-) > 0$. La función $f(x)$ que indica estos incrementos recibe el nombre de **función de probabilidad de X** , y se define como:

$$f(x) = \begin{cases} P(X = x) & \text{si } x \in D \\ 0 & \text{si } x \notin D \end{cases}$$

La función de distribución se reconstruye de la forma siguiente

$$F(x) = \sum_{u \leq x} f(u).$$

Variables aleatorias

Ejemplos de variables aleatorias discretas

- 1 **Distribución de Bernoulli:** X tiene distribución de Bernoulli con probabilidad de éxito p , $X \sim \text{Ber}(p)$, si,

$$f(x) = \begin{cases} 1-p & \text{si } x=0 \\ p & \text{si } x=1 \\ 0 & \text{si c.o.c.} \end{cases}$$

- 2 **Distribución binomial:** X tiene distribución binomial con parámetros $n \in \mathbb{N}$ y $p \in [0, 1]$, $X \sim \text{Binom}(n, p)$,

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x},$$

para $x = 0, 1, 2, \dots, n$.

- 3 **Distribución geométrica:** X tiene distribución geométrica con parámetro $p \in [0, 1]$, $X \sim \text{Geom}(p)$,

$$f(x) = p(1-p)^x,$$

para $x = 0, 1, 2, \dots$.

- 4 **Distribución Poisson:** X tiene distribución de Poisson con parámetro $\lambda > 0$, $X \sim \text{Poisson}(\lambda)$,

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

para $x = 0, 1, 2, \dots$.

Variables aleatorias

Variable aleatoria absolutamente continua

Definición: (Variable aleatoria absolutamente continua)

La variable aleatoria continua X con función de distribución $F(x)$ se llama absolutamente continua, si existe una función no negativa e integrable $f : \mathbb{R} \rightarrow \mathbb{R}$ tal que para cualquier valor de $x \in \mathbb{R}$ se cumple

$$F(x) = \int_{-\infty}^x f(u) du.$$

En tal caso a la función $f(x)$ se le llama función de densidad de X .

Variables aleatorias

Ejemplos de variables aleatorias continuas

- 1 **Distribución normal:** X tiene distribución normal con media μ y varianza σ^2 ,
 $X \sim \mathcal{N}(\mu, \sigma^2)$, si,

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}.$$

- 2 **Distribución exponencial:** X tiene distribución exponencial parámetro λ , $X \sim \exp(\lambda)$, si,

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

Variables aleatorias

Ejemplos de variables aleatorias continuas

- 1 **Distribución normal:** X tiene distribución normal con media μ y varianza σ^2 ,
 $X \sim \mathcal{N}(\mu, \sigma^2)$, si,

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}.$$

- 2 **Distribución exponencial:** X tiene distribución exponencial parámetro λ , $X \sim \text{exp}(\lambda)$, si,

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

- 3 **Distribución uniforme:** X tiene distribución uniforme en el intervalo $[a, b]$, $X \sim \text{unif}(a, b)$, si,

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{si } x \notin [a, b] \end{cases}$$

- 4 **Distribución gamma:** X tiene distribución gamma con parámetros k y λ ,
 $X \sim \text{gamma}(k, \lambda)$, si,

$$f(x) = \begin{cases} \frac{(\lambda x)^{k-1}}{\Gamma(k)} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

Variables aleatorias

Esperanza

Definición: Esperanza

Sea X una variable aleatoria con función de distribución $F(x)$. La esperanza de X , denotada por $E(X)$, se define como:

$$E[X] = \int_{-\infty}^{\infty} x dF(x),$$

cuando esta integral sea absolutamente convergente, es decir, cuando $\int_{-\infty}^{\infty} |x| dF(x) < \infty$, y en tal caso se dice que X es **integrable**, o que X **tiene esperanza finita**.

Variables aleatorias

Esperanza

Esperanza de v.a. discretas y continuas

Si X es una **v.a. discreta** cuya **función de probabilidad** $f(x)$ es positiva en $\{x_1, x_2, \dots\}$, entonces la esperanza X , $E[X]$ está dada por:

$$E[X] = \sum_{n=1}^{\infty} x_n f(x_n).$$

Si X es una **v.a. continua** con **función de densidad** $f(x)$, entonces la esperanza X se puede escribir como:

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx.$$

Variables aleatorias

Esperanza de algunas variables aleatorias

Esperanza de algunas v.a.'s discretas:

- Si $X \sim \mathbf{Ber}(p)$, entonces $E[X] = p$.
- Si $X \sim \mathbf{Binom}(n, p)$, entonces $E[X] = np$.
- Si $X \sim \mathbf{Geom}(p)$, entonces

$$E[X] = \frac{1-p}{p}.$$

- Si $X \sim \mathbf{Poisson}(\lambda)$, entonces $E[X] = \lambda$.

Esperanza de algunas v.a.'s continuas

- Si $X \sim \mathcal{N}(\mu, \sigma)$, entonces $E[X] = \mu$.
- Si $X \sim \mathbf{exp}(\lambda)$, entonces $E[X] = 1/\lambda$.
- Si $X \sim \mathbf{unif}(a, b)$, entonces

$$E[X] = \frac{a+b}{2}$$

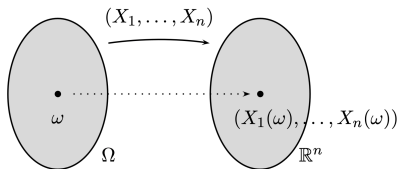
- Si $X \sim \mathbf{gamma}(k, \lambda)$, entonces $E[X] = k/\lambda$.

Vectores aleatorios

Vector aleatorio

Definición: (Vector aleatorio)

Sea (Ω, \mathcal{F}, P) un espacio de probabilidad. Un **vector aleatorio** es una función $X : \Omega \rightarrow \mathbb{R}^n$ tal que para cualquier conjunto B en la σ -álgebra de Borel $\mathcal{B}(\mathbb{R}^n)$, se cumple que $X^{-1}(B)$ es un elemento de \mathcal{F} .

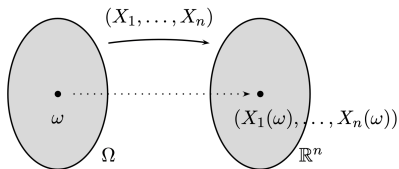


Vectores aleatorios

Vector aleatorio

Definición: (Vector aleatorio)

Sea (Ω, \mathcal{F}, P) un espacio de probabilidad. Un **vector aleatorio** es una función $X : \Omega \rightarrow \mathbb{R}^n$ tal que para cualquier conjunto B en la σ -álgebra de Borel $\mathcal{B}(\mathbb{R}^n)$, se cumple que $X^{-1}(B)$ es un elemento de \mathcal{F} .



Definición: (Función de distribución conjunta)

Sea (X, Y) un vector aleatorio bivariado ($n = 2$) definido sobre un espacio de probabilidad (Ω, \mathcal{F}, P) . La función de distribución $F(x, y)$ del vector aleatorio (X, Y) , es una función $F(x, y) : \mathbb{R}^2 \rightarrow [0, 1]$ definida como:

$$F(x, y) = P(X \leq x; Y \leq y).$$

para todo $(x, y) \in \mathbb{R}^2$

Vectores aleatorios

Vector aleatorio

Definición: (Función de probabilidad conjunta)

La función de probabilidad de un **vector aleatorio discreto** (X, Y) se define como la función $f(x, y) : \mathbb{R}^2 \rightarrow [0, \infty)$ dada por:

$$f(x, y) = P(X = x, Y = y)$$

Vectores aleatorios

Vector aleatorio

Definición: (Función de probabilidad conjunta)

La función de probabilidad de un **vector aleatorio discreto** (X, Y) se define como la función $f(x, y) : \mathbb{R}^2 \rightarrow [0, \infty)$ dada por:

$$f(x, y) = P(X = x, Y = y)$$

Definición: (Función de densidad conjunta)

Sea (X, Y) un vector continuo con función de distribución $F(x, y)$. Se dice que (X, Y) es **vector aleatorio absolutamente continuo** si existe una función no negativa e integrable $f(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$ tal que para todo (x, y) en \mathbb{R}^2 se cumple la igualdad

$$F(x, y) : \int_{-\infty}^x \int_{-\infty}^y f(u, v) dv du.$$

Vectores aleatorios

Marginales

Definición: (Marginales para el caso discreto)

Sea (X, Y) un vector aleatorio. Si (X, Y) es un v.a. discreto con función de probabilidad conjunta $f_{X,Y}(x, y)$. Definimos la **función de probabilidad marginal de X** denotada por $f_X(x)$ se define como

$$f_X(x) = \sum_y f_{X,Y}(x, y)$$

mientras que la **función de probabilidad marginal de Y** denotada por $f_Y(y)$ se define como

$$f_Y(y) = \sum_x f_{X,Y}(x, y)$$

Vectores aleatorios

Marginales

Definición: (Marginales para el caso continuo)

Sea (X, Y) un vector aleatorio abs. continuo con densidad de probabilidad conjunta $f_{X,Y}(x, y)$. Definimos la **función densidad marginal de X** denotada por $f_X(x)$ se define como

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$$

mientras que la **función densidad marginal de Y** denotada por $f_Y(y)$ se define como

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx$$

Vectores aleatorios

Independencia y esperanza

- Dos v.a.'s son independientes si y solo si su función de probabilidad conjunta (o densidad conjunta) se puede factorizar como el producto de sus marginales:

$$f_{X,Y}(x, y) = f_X(x)f_Y(y)$$

- La **esperanza** de $\phi(X, Y)$ está dada por

$$E[\phi(X, Y)] = \sum_x \sum_y \phi(x, y) f_{X,Y}(x, y)$$

donde (X, Y) es un v.a. discreto con función de probabilidad $f_{X,Y}(x, y)$.

- La **esperanza** v $\phi(X, Y)$ está dada por

$$E[\phi(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(x, y) f_{X,Y}(x, y) dx dy.$$

donde (X, Y) es un v.a. continuo con función densidad de probabilidad $f_{X,Y}(x, y)$.

Vectores aleatorios

Independencia y esperanza

- Para dos v.a.'s X y Y , con esperanza finita, tenemos que:

$$E[X + Y] = E[X] + E[Y].$$

- Para dos v.a.'s X y Y independientes tenemos que:

$$E[g(X)h(Y)] = E[g(X)]E[h(Y)].$$

- En particular, si X y Y son independientes tenemos que

$$E[XY] = E[X]E[Y].$$

Vectores aleatorios

Covarianza

Definición (Covarianza)

La covarianza de dos v.a.'s X y Y , denotada por $\text{Cov}(X, Y)$, está definida como:

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))]$$

Para dos v.a.'s. X y Y tenemos que:

$$\text{Cov}(X, Y) = E(XY) - E(X)E[Y].$$

Además, si X y Y son independientes, entonces

$$\text{Cov}(X, Y) = 0,$$

(pero $\text{Cov}(X, Y) = 0$ no implica independencia de X y Y).

Vectores aleatorios

Varianza

Definición: (Varianza)

La varianza de una v.a. X se define como:

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

Nótese que $\text{Var}(X) = \text{Cov}(X, X)$ y por tanto $\text{Var}(X) = E(X^2) - (E(X))^2$.

Proposición:

Si X y Y son independientes, entonces,

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

Vectores aleatorios

Coeficiente de correlación

Definición: (Coeficiente de correlación)

El coeficiente de correlación de las variables aleatorias X y Y , denotado por $\rho(X, Y)$, es definido como:

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$$

Proposición:

1. Si X y Y son independientes, entonces $\rho(X, Y) = 0$.
2. $-1 \leq \rho(X, Y) \leq 1$.
3. $|\rho(X, Y)| = 1$ i y solo si existen constantes reales a y b tales que, con probabilidad uno, $Y = aX + b$ con $a > 0$ si $\rho(X, Y) = 1$ y $a < 0$ si $\rho(X, Y) = -1$.

Variables aleatorias

Teorema de límite central

Teorema de límite central

Sea $\{X_n : n \in \mathbb{N}\}$ una colección de v.a.'s independientes e idénticamente distribuidas tales que $E(X_n) = \mu$ y $\text{Var}(X_n) = \sigma^2 < \infty$. Entonces, para cualquier $x \in \mathbb{R}$, se tiene que:

$$\lim_{n \rightarrow \infty} P\left(\frac{(X_1 + X_2 + \cdots + X_n) - n\mu}{\sqrt{n}\sigma} \leq x\right) = P(Z \leq x)$$

en donde Z es una v.a. con distribución normal estándar.

O en otras palabras, se puede considerar que la v.a.

$$Z_n := \frac{\frac{1}{n}(X_1 + X_2 + \cdots + X_n) - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} Z \sim \mathcal{N}(0, 1).$$

tiende en distribución a una v.a.' con distribución normal estándar.

Estadística

Estadística básica

Variables aleatorias

Problema central de la estadística

La **probabilidad** ofrece una amplia variedad de modelos matemáticos para describir un **fenómeno observado**.

La **estadística** se encarga de escoger el modelo probabilístico para describir el **fenómeno observado** de manera *óptima*.

El conjunto de técnicas para este propósito se denomina *inferencia estadística*.

Variables aleatorias

Problema central de la estadística

La estadística matemática hace uso de la hipótesis de que una serie de observaciones *cuantitativas* (y a menudo también *cualitativas*) de algún fenómeno son consideradas **realizaciones independientes de una cierta variable aleatoria**.

Al iniciar el estudio estadístico de una serie de observaciones, normalmente las propiedades de dicha variable aleatoria son **desconocidas**.

(A la postre, las “observaciones cuantitativas” las relacionaremos con el concepto de serie de tiempo)

Variables aleatorias

Niveles de inferencia

Cuando *sabemos* qué tipo de distribución describe la variable aleatoria, tenemos dos opciones:

- Proponer una familia paramétrica de la posible distribución para la variable aleatoria. En esta situación requerimos conocer el parámetro óptimo que describe las observaciones. El conjunto de técnicas para este propósito se denomina **teoría de la estimación**.
- Suponer que un parámetro específico, de la familia paramétrica, es el valor “correcto” o “adecuado” que describe una serie de observaciones. Tenemos que probar que eso es así o refutar la hipótesis. El conjunto de técnicas para este propósito se denomina **prueba de hipótesis o contraste de hipótesis**.

Variables aleatorias

Niveles de inferencia

Cuando *no sabemos* qué tipo de distribución describe la variable aleatoria, tenemos otras opciones:

- Suponer que una cierta distribución describe las observaciones. En este caso tenemos que **contrastar hipótesis a nivel de distribuciones** (no solamente sobre parámetros).
- Suponer que dos observaciones siguen o no siguen una misma distribución o alguna propiedad específica de la distribución.

El conjunto de técnicas para estos propósitos forman parte del *corpus* de la teoría de **inferencia no paramétrica**.

Variables aleatorias

Estimación

Hay dos grandes formas de realizar estimación de parámetros a partir de observaciones:

- **Estimación puntual:** se calcula un valor específico, a partir de las observaciones, para el parámetro óptimo de la familia paramétrica de la distribución propuesta.
- **Estimación por intervalo:** se calcula todo un intervalo en el cual es posible hallar el valor “correcto” del parámetro buscado, con cierta probabilidad.

Variables aleatorias

Estimación

Una **estimación puntual** se lleva a cabo a través de un **estimador** (algunas veces también llamado *estadística*). Un estimador es una v.a. definida a partir de la *serie de observaciones* que se denomina *muestra*.

Si $\{X_1, X_2, \dots, X_n\}$ es una muestra, que suponemos es descrita por una distribución parametrizada $f(x; \theta)$, entonces un estimador para el parámetro (o conjunto de parámetros) θ es una v.a. función de la muestra:

$$\Theta = \Theta(X_1, X_2, \dots, X_n).$$

Variables aleatorias

Estimación

Propiedades *deseables* de un estimador:

- Insesgado:

$$E[\Theta(X_1, X_2, \dots, X_n)] = \theta.$$

- Consistente:

$$\lim_{n \rightarrow \infty} P\left(\left|\Theta(X_1, X_2, \dots, X_n) - \theta\right| < \epsilon\right) = 0.$$

- Eficiente:

$$\text{Var}(\Theta(X_1, X_2, \dots, X_n)) \leq \text{Var}(\Theta'(X_1, X_2, \dots, X_n)).$$

Procesos estocásticos

Procesos estocásticos

Procesos estocásticos

Generalidades

Un **proceso estocástico** es un concepto matemático que describe la evolución de un sistema que cambia con el tiempo de manera aleatoria o incierta. A diferencia de los procesos deterministas, donde el futuro del sistema está completamente determinado por las condiciones iniciales, en un proceso estocástico el comportamiento futuro no puede predecirse con certeza, aunque se pueden estudiar sus propiedades probabilísticas.

La principal característica es la **dependencia temporal y aleatoriedad**: Un proceso estocástico es una colección de variables aleatorias indexadas por el tiempo (o cualquier otra variable). A cada instante de tiempo, el proceso puede tomar distintos valores, que dependen de ciertas probabilidades, lo que introduce aleatoriedad en su comportamiento.

Procesos estocásticos

Clasificación

■ En función del tiempo:

Discretos en el tiempo: El proceso se observa en momentos discretos (es decir, en instantes separados). Un ejemplo es el valor de una acción en la bolsa de valores al final de cada día.

Continuos en el tiempo: El proceso se observa en todos los instantes de tiempo, como la temperatura medida en un punto en el tiempo que varía de manera continua.

■ En función del espacio de estados:

Discretos en el espacio: El proceso solo puede tomar un número finito o contable de valores posibles. Un ejemplo es el número de llamadas que recibe una central telefónica en una hora.

Continuos en el espacio: El proceso puede tomar un conjunto continuo de valores. Un ejemplo sería el precio de una acción, que puede variar en cualquier valor dentro de un rango continuo.

Procesos estocásticos

Algunos procesos estocásticos de uso común

- **Proceso de Poisson:** Es un proceso de tiempo continuo donde los eventos ocurren de manera aleatoria pero con una tasa promedio constante. Es muy útil para modelar fenómenos como la llegada de llamadas telefónicas a una central, o los intervalos entre accidentes de tráfico.
- **Proceso Markoviano (Cadena de Markov):** En un proceso de Markov, el futuro del proceso depende solo del estado presente y no del camino seguido para llegar a ese estado (propiedad de "sin memoria"). Las cadenas de Markov son muy usadas en la modelización de sistemas donde solo el estado actual es relevante para predecir el futuro.
- **Proceso de Wiener (Movimiento browniano):** Es un ejemplo clásico de proceso estocástico continuo en el tiempo y en el espacio, que modela el movimiento aleatorio de partículas suspendidas en un fluido. Se usa también para modelar la evolución de los precios de activos financieros.
- **Ruido blanco:** Es un proceso estocástico en el cual las observaciones en diferentes instantes de tiempo son independientes y tienen una distribución de probabilidad constante. Es la base de muchos modelos en teoría de señales y econometría.

Procesos estocásticos

Aplicaciones de los procesos estocásticos

- **Economía y finanzas:** Los procesos estocásticos se usan para modelar la evolución de precios de activos financieros, tasas de interés y otros factores económicos. Por ejemplo, el *modelo de Black-Scholes* para la valoración de opciones se basa en la idea de que los precios de los activos siguen un movimiento browniano.
- **Ciencias naturales:** En biología, los procesos estocásticos modelan fenómenos como la evolución de poblaciones, las mutaciones genéticas o la difusión de moléculas en biología celular.
- **Física:** El movimiento browniano es un ejemplo importante en la física estadística, donde las trayectorias de partículas en fluidos se modelan como procesos estocásticos.
- **Ingeniería:** En telecomunicaciones y procesamiento de señales, los procesos estocásticos se usan para modelar ruido, señales y otros fenómenos aleatorios.

Cadenas de Markov

Cadenas de Markov

Cadenas de Markov

Definición conceptual

Una **cadena de Markov** es un modelo matemático que describe un sistema que cambia de un estado a otro dentro de un conjunto de posibles estados. El principio fundamental de una cadena de Markov es que la **probabilidad de pasar a un estado futuro depende solo del estado actual** en el que se encuentra el sistema, y no de los estados anteriores. A esto se le llama la propiedad de Markov o memoria a corto plazo.

Para definir una cadena de Markov requerimos:

- **Un conjunto de estados:** Representan las diferentes situaciones o posiciones posibles en las que el sistema puede estar. Se denotan como
- **Definir las transiciones entre estados:** Representan el paso de un estado a otro. Estas transiciones ocurren con ciertas probabilidades que se colectan en una **matriz de transición**.
- **Una condición inicial:** Indica cómo “echar a andar la cadena”, es decir, cómo comenzar la evolución del sistema.

Cadenas de Markov

Definición matemática

Definición: (Cadena de Markov)

Una **cadena de Markov** es un proceso estocástico a tiempo discreto $\{X_n : n = 0, 1, 2, \dots\}$ con espacio de estados discreto S tal que para cualquier entero $n \geq 0$ y para cualesquiera $x_0, x_1, \dots, x_{n+1} \in S$ satisface

$$P(X_{n+1} = x_{n+1} | X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = P(X_{n+1} = x_{n+1} | X_n = x_n)$$

a esta propiedad se le conoce como propiedad de Markov.

La matriz de transición (también llamada matriz de probabilidad) es la probabilidad condicional del proceso:

$$M(a, b) := P(X_{n+1} = b | X_n = a)$$

Cadenas de Markov

Definición matemática

Para definir una cadena de Markov necesitamos tres ingredientes principalmente (“datos del problema”):

- Espacio de estados: \mathcal{S}
- Probabilidades de transición: la matriz $M(a, b)$.
- Una condición inicial: la variable aleatoria X_0 (la primera en la colección $\{X_n : n = 0, 1, 2, \dots\}$)

Cadenas de Markov

Definición matemática

Para definir una cadena de Markov necesitamos tres ingredientes principalmente (“datos del problema”):

- Espacio de estados: \mathcal{S}
- Probabilidades de transición: la matriz $M(a, b)$.
- Una condición inicial: la variable aleatoria X_0 (la primera en la colección $\{X_n : n = 0, 1, 2, \dots\}$)

Ejemplo:

Consideremos un espacio de dos estados
 $S = \{0, 1\}$.

Cadenas de Markov

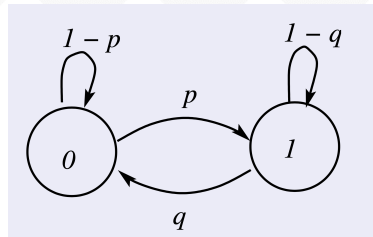
Definición matemática

Para definir una cadena de Markov necesitamos tres ingredientes principalmente (“datos del problema”):

- Espacio de estados: \mathcal{S}
- Probabilidades de transición: la matriz $M(a, b)$.
- Una condición inicial: la variable aleatoria X_0 (la primera en la colección $\{X_n : n = 0, 1, 2, \dots\}$)

Ejemplo:

Consideremos un espacio de dos estados $\mathcal{S} = \{0, 1\}$.



Asignamos a cada transición una probabilidad:

$$P = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}$$

La matriz P representa la **matriz de transición**.

Cadenas de Markov

Cantidades de interés

Desde un punto de vista práctico nos interesa conocer la evolución de las distribuciones, algo que se puede hacer a partir de la matriz de transición. Es decir, si el proceso se representa por $\{X_n : n = 0, 1, 2, \dots\}$, buscamos

$$P(X_n = s), \quad \text{para } s \in S.$$

Si la condición inicial se representa como un vector (fila):

$$\pi_0 = (\pi_0(1), \pi_0(2), \dots, \pi_0(s)).$$

(con $a \in [0, 1]$) entonces la evolución de la distribución está dada por

$$\pi_{n+1} = \pi_n P$$

donde usamos la notación:

$$\pi_n(s) := P(X_n = s)$$

Referencias

BRACEWELL, R. N. **The Fourier Transform and Its Applications**. 3rd edition. ed. [S.l.]: McGraw-Hill, 2000.

CHUNG, J.; GULCEHRE, C.; CHO, K.; BENGIO, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. **arXiv preprint arXiv:1412.3555**, 2014.

DURBIN, J.; KOOPMAN, S. J. **Time Series Analysis by State Space Methods**. 2nd edition. ed. [S.l.]: Oxford University Press, 2012.

ENGLE, R. F.; PATTON, A. J. **ARCH Models in Finance**. Oxford, UK: Oxford University Press, 1995.

HAMILTON, J. D. A new approach to the economic analysis of nonstationary time series and the business cycle. **Econometrica: Journal of the Econometric Society**, Wiley Online Library, p. 357–384, 1989.

_____. **Time Series Analysis**. Princeton, NJ: Princeton University Press, 1994.

HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. **Neural computation**, MIT Press, v. 9, n. 8, p. 1735–1780, 1997.

HYNDMAN, R. J.; KOEHLER, A. B.; ORD, J. K.; SNYDER, R. D. **Forecasting with Exponential Smoothing: The State Space Approach**. Berlin, Germany: Springer, 2008.

LÜTKEPOHL, H. **New Introduction to Multiple Time Series Analysis**. [S.l.]: Springer Science & Business Media, 2005.

MALLAT, S. **A Wavelet Tour of Signal Processing**. 2nd edition. ed. [S.l.]: Academic Press, 1999.

RASMUSSEN, C. E.; WILLIAMS, C. K. I. **Gaussian Processes for Machine Learning**. Cambridge, MA: MIT Press, 2006.

TAYLOR, S. J.; LETHAM, B. Forecasting at scale. **The American Statistician**, Taylor & Francis, v. 72, n. 1, p. 37–45, 2018.