

Procesos Estocásticos y Series de Tiempo

Introducción Conceptual

Daniel Camarena



17 de noviembre de 2021

Contenido

- 1 Croquis de Estadística
- 2 La Ruina del Apostador
 - El problema de la ruina del apostador
 - Un primer modelo
- 3 ¿Qué es un Proceso Estocástico?
 - Proceso estocástico vs. proceso determinista
 - Modelos discretos vs Modelos continuos
- 4 Cadenas de Markov
 - Un segundo modelo para la ruina del apostador
- 5 ¿Qué es una serie de tiempo?
 - Idea de serie de tiempo
 - Modelo de serie de tiempo
 - Análisis de series de tiempo

Croquis de Estadística

CROQUIS DE ESTADÍSTICA

Enfoque
Frecuentista
VS
Enfoque
Bayesiano

→ Estadística
Univariada
Inferencia

Y

→ Estadística
Bivariada
Dependencia

$Y \sim X$

No Restringida
Ordenamiento
Agrupamiento

+

Restringida
Causalidad

\approx

Estadística
Multivariada

← Teoría de la
Probabilidad

← Estadística
Descriptiva

Aprendizaje
Estadístico

\sim

No Supervisado
Estructura

+

Supervisado
Predicción

+ ...

ΔY_t

$\Delta Y_t \sim \Delta X_t$

Estadística
Temporal

$\Delta Y_{i,j}$

$\Delta Y_{i,j} \sim \Delta X_{i,j}$

Estadística
Espacial

⋮

Dimensiones

Estadística
Paramétrica

→ Estadística
Semiparamétrica

← Estadística
No Paramétrica

Complejidad

La Ruina del Apostador

La ruina del apostador

The “Gambler’s Ruin”

Gambler A



Gambler B



50%



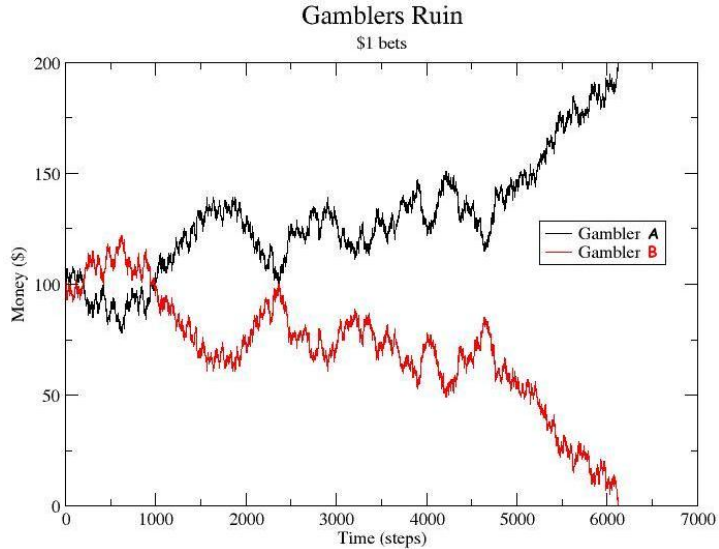
Chance
of winning



50%



La ruina del apostador



Modelando la ruina del apostador

Problema:

¿Cuál es la probabilidad de que el jugador A gane el juego?

Supongamos que se distribuye k monedas al jugador A y $N - k$ monedas al jugador B, entonces se requiere hallar

$$p_k = P(\text{A gane el juego})$$

donde hay una dependencia de k .

Modelando la ruina del apostador

Sea C el evento que el jugador A gane la primera moneda entonces por regla de probabilidad condicional

$$p_k = P(A \text{ gane el juego} | C) \times P(C) + P(A \text{ gane el juego} | C^c) \times P(C^c)$$

Como $P(C) = 0.5$ se tiene

$$p_k = p_{k+1} \times 0.5 + p_{k-1} \times 0.5$$

Considerando condiciones de frontera se llega a la siguiente ecuación

$$\begin{cases} p_0 = P(A \text{ gane el juego empezando con 0 monedas}) = 0 & k = 0 \\ p_k = 0.5p_{k+1} + 0.5p_{k-1} & 0 < k < N \\ p_N = P(A \text{ gane el juego empezando con } N \text{ monedas}) = 1 & k = N \end{cases}$$

Modelando la ruina del apostador

Otro problema:

¿Cuánto durará el juego en promedio hasta que el jugador A gane (o pierda)?

Supongamos que se distribuye k monedas al jugador A y $N - k$ monedas al jugador B, entonces se requiere hallar

$$m_k = E(\text{\#lanzamientos hasta que A gane o pierda})$$

donde hay una dependencia de k .

Modelando la ruina del apostador

Como en el problema de calcular la probabilidad de que el jugador A gane: Sea C el evento que el jugador A gane la primera moneda entonces por regla de esperanza condicional

$$m_k = (1 + m_{k+1}) \times P(C) + (1 + m_{k-1}) \times P(C^c)$$

Considerando condiciones de frontera se llega a la siguiente ecuación

$$\begin{cases} m_0 = 0 & k = 0 \\ m_k = 1 + 0.5m_{k+1} + 0.5m_{k-1} & 0 < k < N \\ m_N = 0 & k = N \end{cases}$$

Modelando la ruina del apostador

La probabilidad de que el jugador A gane:

$$\begin{cases} p_0 = 0 & k = 0 \\ p_k = 0.5p_{k+1} + 0.5p_{k-1} & 0 < k < N \\ p_N = 1 & k = N \end{cases} \longrightarrow p_k = \frac{k}{N}$$

La duración promedio del juego:

$$\begin{cases} m_0 = 0 & k = 0 \\ m_k = 1 + 0.5m_{k+1} + 0.5m_{k-1} & 0 < k < N \\ m_N = 0 & k = N \end{cases} \longrightarrow m_k = k(N - k)$$

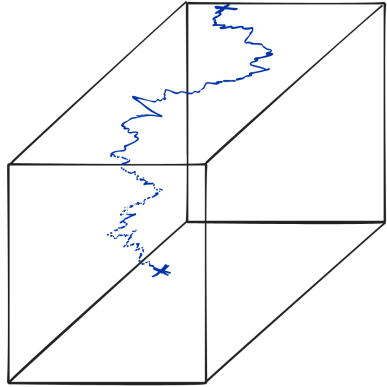
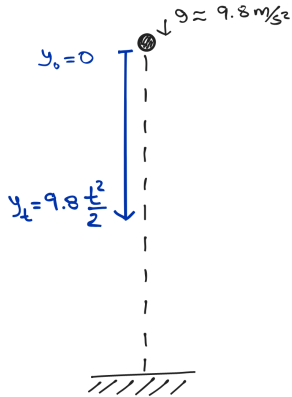
¿Qué es un Proceso Estocástico?

¿Qué es un Proceso Estocástico?

Determinismo

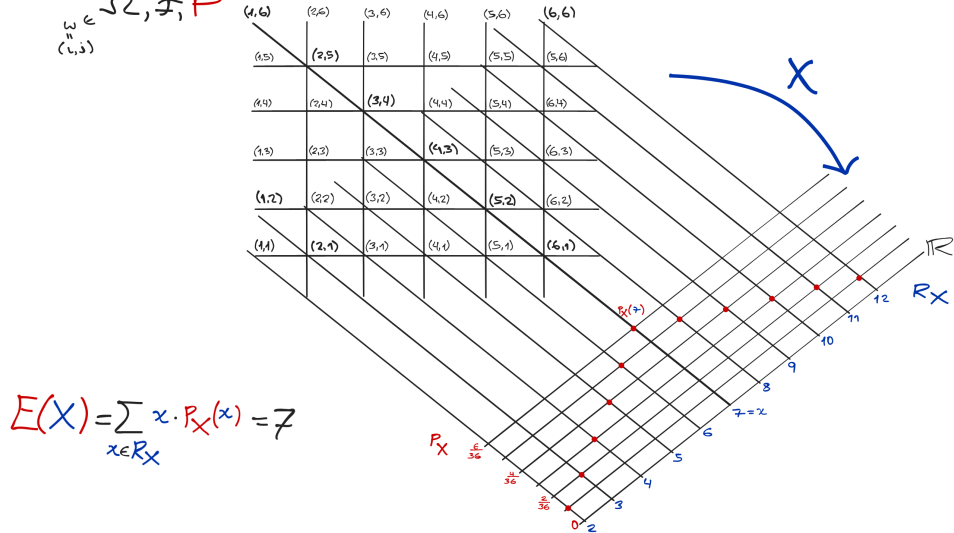
vs

Aleatoriedad



Modelando la aleatoriedad

$\omega \in \Omega, \mathcal{F}, \mathbb{P}$
 (i, j)



$$E(X) = \sum_{x \in R_X} x \cdot P_X(x) = 7$$

Definición de proceso estocástico

Proceso estocástico: $X = \{X_t : t \in T\}$

- Espacio de probabilidad:

$$(\Omega, \mathcal{F}, \mathbb{P})$$

- Horizonte temporal (infinito):

$$T = \{0, 1, 2, \dots\} = \mathbb{N} \quad \text{o} \quad T = [0, +\infty[$$

- Espacio de estados:

$$E = (\mathbb{Z}, \mathcal{P}(\mathbb{Z})) \quad \text{o} \quad E = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

- Variable de estado al tiempo $t \in T$:

$$(\Omega, \mathbb{P}) \xrightarrow{X_t} (E, P_{X_t})$$

Modelos discretos y modelos continuos

Proceso estocástico: $\left\{ \Omega \xrightarrow{X_t} E \right\}_{t \in T}$

- Proceso estocástico en tiempo discreto

$$T = \{0, 1, 2, \dots\} = \mathbb{N} \quad \text{y} \quad \begin{cases} E \subset \mathbb{Z} & \text{espacio de estados discreto} \\ E = \mathbb{R} & \text{espacio de estados continuo} \end{cases}$$

En este caso, las trayectorias o realizaciones del proceso son sucesiones

$$(X_t(\omega))_{t \in \mathbb{N}} \in E^{\mathbb{N}}$$

- Proceso estocástico tiempo continuo

$$T = [0, +\infty[\quad \text{y} \quad \begin{cases} E \subset \mathbb{Z} & \text{espacio de estados discreto} \\ E = \mathbb{R} & \text{espacio de estados continuo} \end{cases}$$

En este caso, las trayectorias o realizaciones del proceso son funciones

$$(t \mapsto X_t(\omega))_{t \in [0, \infty[} \in E^{[0, \infty[}$$

Ejemplos de modelos discretos y modelos continuos

- P.E. en tiempo discreto: La ruina del apostador

$$X_t = \text{reserva al tiempo } t = k + \sum_{i=1}^t \xi_i, \quad \forall t = 0, 1, 2, \dots$$

- P.E. en tiempo discreto: Caminata aleatoria en \mathbb{R}

$$X_t = \text{posición al tiempo } t = \sum_{i=1}^t \xi_i, \quad \forall t = 0, 1, 2, \dots$$

- P.E. en tiempo continuo: Proceso Poisson

$$N_t = \# \text{ eventos al tiempo } t \sim \text{Poi}(\lambda t), \quad \forall t \geq 0$$

- P.E. en tiempo continuo: Movimiento Browniano

$$W_t = \text{posición al tiempo } t \sim \mathcal{N}(0, t), \quad \forall t \geq 0$$

Cadenas de Markov

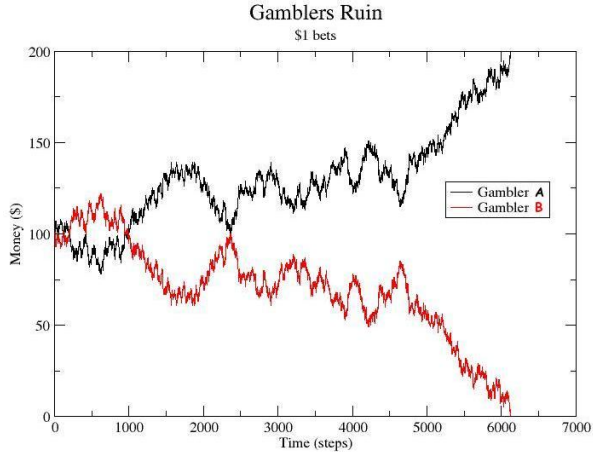
Procesos Estocásticos: Cadenas de Markov

- La ruina del apostador
- Caminatas Aleatorias
- Cadenas de Markov
- Procesos de Ramificación

Bibliografía recomendada.

- [**Rodriguez**] Pablo Rodriguez (2017). Modelos probabilisticos discretos y aplicaciones.
- [**Schinazi**] Rinaldo Schinazi (2014). Classical and Spatial Stochastic Processes.

La propiedad de pérdida de memoria



- Capital inicial del apostador: $X_0 = k$
- Después de un tiempo $t > 0$: $X_t = x_t$

La ruina del apostador como proceso estocástico

Proceso estocástico en tiempo discreto

$$\left\{ \Omega \xrightarrow{X_t} E \right\}_{t \in T}$$

donde

- Horizonte temporal: $T = \{0, 1, 2, \dots\} = \mathbb{N}$

- Apuesta:

$$\{\xi_i\}_{i \geq 1} \stackrel{iid}{\sim} 2\text{Ber}(p) - 1, \quad 0 < p < 1.$$

- Variable de estado:

$$X_t = \begin{cases} k & t = 0 \\ k + \sum_{i=1}^t \xi_i & t = 1, 2, \dots \end{cases}$$

- Espacio de estados: $E = \{\dots, -2, -1, 0, 1, 2, \dots\} = \mathbb{Z}$ (restringir a $E = \{0, 1, 2, \dots, N\}$)

La propiedad de Markov

Para nuestro proceso estocástico en *tiempo discreto* $\left\{ \Omega \xrightarrow{X_t} E \right\}_{t \in \mathbb{N}}$

$$X_{t+1} = k + \sum_{i=1}^{t+1} \xi_i = X_t + \xi_{t+1}$$

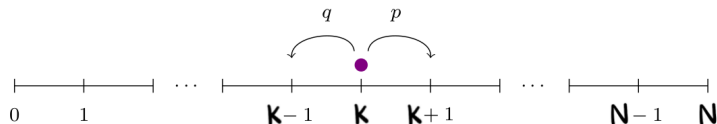
Propiedad de Markov (con eventos):

$$P_{x_0}(X_{t+1} = x | X_t = x_t, \dots, X_0 = x_0) = P_{x_0}(X_{t+1} = x | X_t = x_t)$$

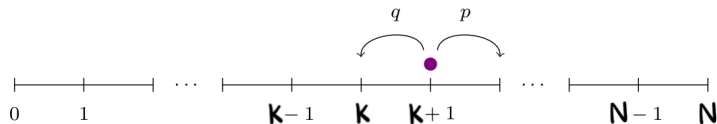
Propiedad de Markov (con procesos):

$\{Y_t = X_{t+\tau} : t \geq 0\} | \{X_\tau = x_\tau\}$ es independiente de $\{X_\tau = x_\tau, \dots, X_0 = x_0\}$
 $\{Y_t = X_{t+\tau} : t \geq 0\} | \{X_\tau = x_0\}$ tiene la misma dinámica que $\{X_s : s \geq 0\}$

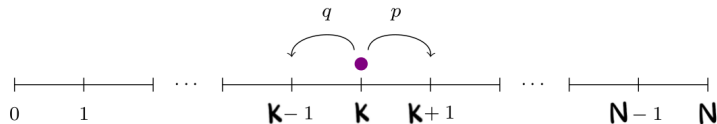
Cadenas de Markov sobre grafos



En $\mathbf{t} = 0$ la partícula está en el vértice \mathbf{k}



En $\mathbf{t} = 1$ la partícula salta para el vértice $\mathbf{k} + 1$

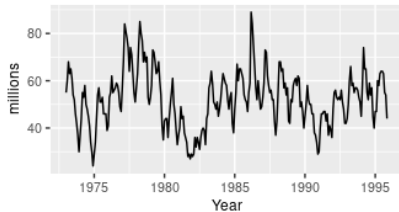


En $\mathbf{t} = 2$ la partícula salta para el vértice $\mathbf{k} - 1$

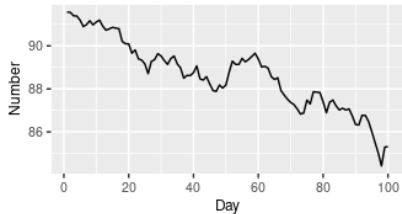
¿Qué es una serie de tiempo?

¿Series de tiempo?

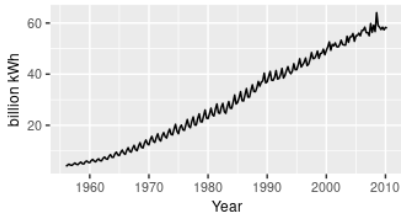
Sales of new one-family houses, USA



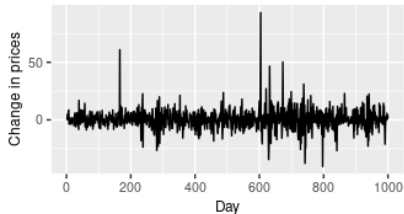
US treasury bill contracts



Australian quarterly electricity production



Google daily changes in closing stock price



¿Series de tiempo?

Serie de tiempo: *observaciones equiespaciadas en el tiempo de una magnitud*

$$y_{t_1}, y_{t_2}, \dots, y_{t_N}$$

Componentes de una serie de tiempo

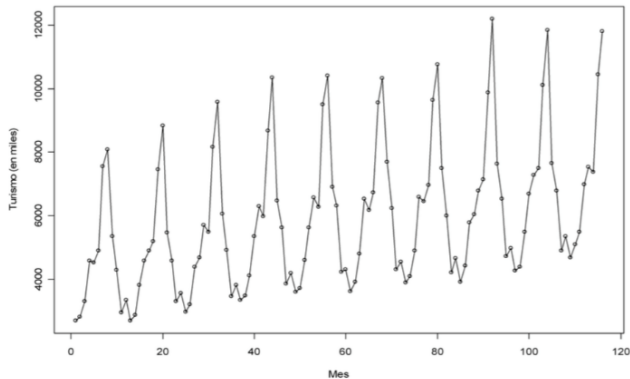
- Tendencia: movimiento a largo plazo
- Ciclo: oscilaciones de mediano plazo alrededor de la tendencia (mayor a 1 año)
- Estacionalidad: oscilaciones periódicas de corto plazo (menor a 1 año)
- Aleatoria: no siguen un patrón específico (ruido con correlación)

Lugar: España

Frecuencia de observ.: Mensual

Cantidad de observaciones: 116 (enero 1995-agosto 2004)

La serie presenta tendencia y componente estacional

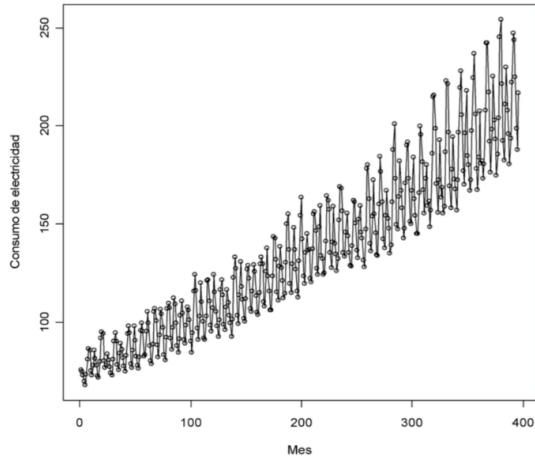


Lugar: EE.UU.

Frecuencia de observ.: Mensual

Cantidad de observaciones: 396 (enero 1972-diciembre 2004)

La serie presenta tendencia, componente estacional y heterocedasticidad



Autocorrelación

Si consideramos en dos periodos de tiempo

$$\begin{cases} y_{t_1}, y_{t_2}, \dots, y_{t_{N-1}} & \text{observaciones de } Y_{t-1} \\ y_{t_2}, y_{t_3}, \dots, y_{t_N} & \text{observaciones de } Y_t \end{cases}$$

la correlación en serie o autocorrelación es

$$\rho(t-1, t) = \frac{\text{Cov}(Y_{t-1}, Y_t)}{\text{sd}(Y_{t-1})\text{sd}(Y_t)}$$

Más generalmente se puede definir la *autocorrealación entre s y t*

$$\rho(s, t) = \frac{\text{Cov}(Y_s, Y_t)}{\text{sd}(Y_s)\text{sd}(Y_t)}$$

Modelo de serie de tiempo

Proceso estocástico: $\dots, \mathbf{Y}_{-1}, \mathbf{Y}_0, \underbrace{\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_N}_{\downarrow}, \mathbf{Y}_{N+1}, \dots$

Serie temporal:

$\boxed{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N}$

Modelos de serie de tiempo

(Modelo de) Serie de tiempo: $\left\{ \Omega \xrightarrow{Y_t} E \right\}_{t \in \mathbb{N}}$

- Modelo aditivo:

$$Y_t = M_t + C_t + S_t + \epsilon_t$$

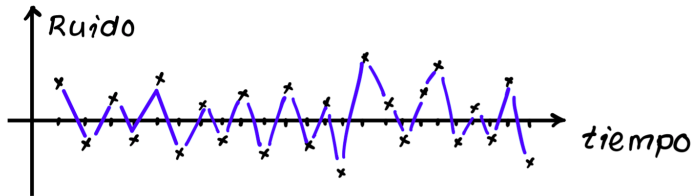
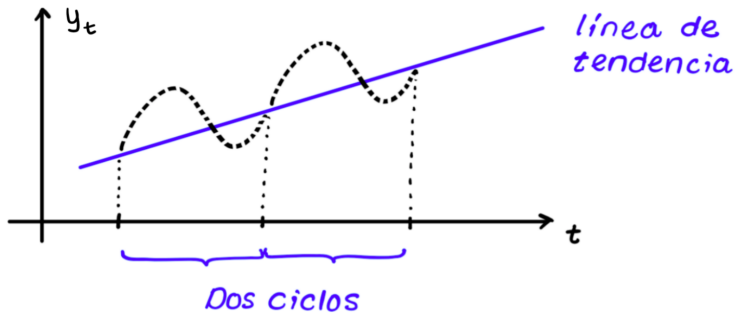
- Modelo multiplicativo:

$$Y_t = M_t \times C_t \times S_t \times \epsilon_t$$

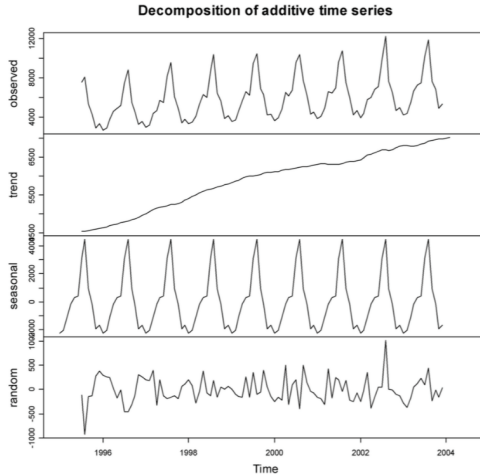
Componentes de una serie de tiempo

- Tendencia: M_t
- Ciclo: C_t
- Estacionalidad: S_t
- Aleatoria: ϵ_t

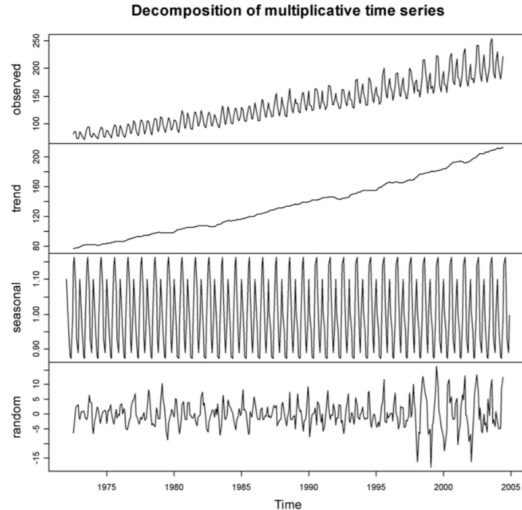
Componentes de serie de tiempo



Lugar: España , Frecuencia de observ.: Mensual, Cantidad de observaciones: 116 (enero 1995-agosto 2004). La serie presenta tendencia y componente estacional



Lugar: EE.UU., Frecuencia de observ.: Mensual, Cantidad de observaciones: 396 (enero 1972-diciembre 2004). La serie presenta tendencia, componente estacional y heterocedasticidad

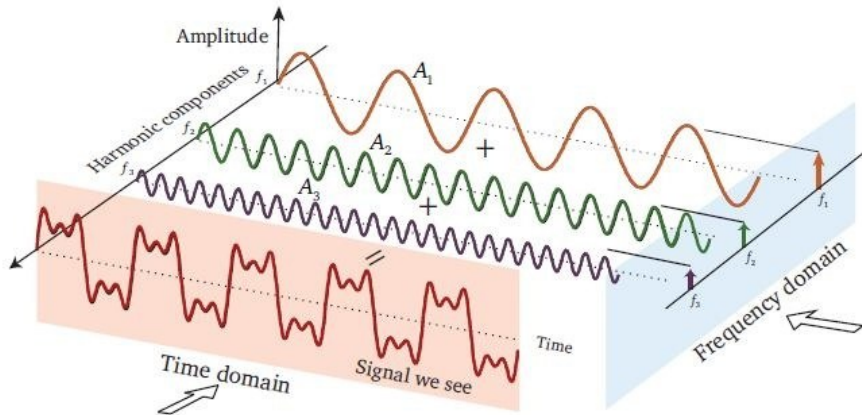


Análisis de series de tiempo

Un gráfico en el dominio del tiempo muestra cómo cambia una señal (serie de tiempo) con el tiempo, mientras que un gráfico en el dominio de la frecuencia muestra qué parte de la señal (serie de tiempo) se encuentra dentro de cada banda de frecuencia dada en un rango de frecuencias.

- **Dominio tiempo (τ):** predicción y pronóstico.
Se construye un modelo que permita extrapolar fuera del periodo cubierto por los datos.
- **Dominio frecuencia (ω):** diagnóstico y detección.
Se descompone la serie en subseries más tratables asociadas a una frecuencia específica.

Dominio tiempo vs. Dominio Frecuencia



Referencias sobre Series de Tiempo

- Modelo de series de tiempo: Procesos estacionarios.
- Dominio Tiempo: Modelos ARIMA
- Dominio Frecuencia: Análisis Fourier y Wavelet
- Procesos no estacionarios o no lineales*

Bibliografía recomendada.

- [**Hamilton**] Hamilton (1994). Time Series Analysis.
- [**Shumway**] Shumway & Stoffer (2017). Time Series Analysis and Its Applications.

¡Muchas Gracias!