

# Introducción a la Teoría de la Información

## Tasa de Entropía de un Proceso Estocástico.

Facultad de Ingeniería, UdelaR

# Agenda

1 Procesos Estocásticos

2 Tasa de Entropía

## Definición (Proceso estocástico)

Un *proceso estocástico* (discreto) es una secuencia indexada de variables aleatorias caracterizado por una distribución conjunta  $p(x_1 \dots x_n)$ ,  $n = 1, 2 \dots$

## Definición (Proceso estacionario)

Un proceso estocástico es *estacionario* si las distribuciones conjuntas no cambian en el tiempo, es decir

$$P(X_1 = x_1, \dots, X_n = x_n) = P(X_{1+r} = x_1, \dots, X_{n+r} = x_n),$$

para todo  $r, n$  y todo  $x_1 \dots x_n \in \mathcal{X}$ .

# Proceso de Markov

## Definición (Proceso de Markov)

Un *proceso o cadena de Markov* es un proceso estocástico en el cual para todo  $n > 0$  se cumple

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

En ese caso, podemos escribir  $p(x_1 \dots x_n) = p(x_1)p(x_2|x_1) \dots p(x_n|x_{n-1})$ .

## Definición (Estado)

Al valor que toma  $X_n$  se le llama *estado* de la cadena en el tiempo  $n$ .

## Definición (Invariancia en el tiempo)

Asumiremos que las probabilidades condicionales no dependen de  $n$ , es decir, para todo  $n > 0$  se cumple

$$P(X_{n+1} = b | X_n = a) = P(X_2 = b | X_1 = a).$$

En este caso decimos que el proceso de Markov es *invariante en el tiempo*.

# Proceso de Markov

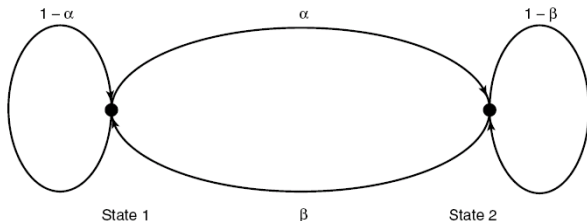
## Definición (Matriz de probabilidades de transición)

Definimos la *matriz de probabilidades de transición* como

$$\mathbf{P}_{ij} = P(X_{n+1} = j | X_n = i).$$

## Observación

Sea  $\mathbf{p}^{(n)}$  el vector de probabilidades definido como  $\mathbf{p}_i^{(n)} = P(X_n = i)$ . Entonces,  
 $\mathbf{p}^{(n+1)} = \mathbf{p}^{(n)} \times \mathbf{P} = \mathbf{p}^{(0)} \times \mathbf{P}^n$



# Proceso de Markov Estacionario

## Definición (Distribución estacionaria)

Una *distribución estacionaria* para los estados de una cadena de Markov es una distribución que se mantiene en el tiempo. Si  $\pi$  es una distribución estacionaria, entonces debe cumplirse

$$\pi = \pi \times \mathbf{P} .$$

## Observación

Si el estado inicial se elige con distribución estacionaria, el proceso de Markov resulta estacionario.

### Definición (Cadena de Markov irreducible)

Una cadena de Markov es *irreducible* si es posible llegar a cualquier estado desde cualquier estado en una cantidad finita de pasos. Es decir, para todo  $i, j$  existe  $k$  tal que se verifica

$$P(X_{n+k} = j | X_n = i) > 0.$$

### Definición (Cadena de Markov aperiódica)

Una cadena de Markov es *aperiódica* si para cada estado  $i$  existe una constante  $M_i$  tal que, para todo  $n > M_i$ , existe una trayectoria de probabilidad positiva que conduce de  $i$  a  $i$  en  $n$  pasos.

## Teorema

Una cadena de Markov irreducible y aperiódica tiene una única distribución estacionaria que aparece como distribución límite cualquiera sea la distribución del estado inicial. Es decir

$$\lim_{n \rightarrow \infty} \mathbf{p}^{(0)} \times \mathbf{P}^n = \pi, \quad \text{para cualquier } \mathbf{p}^{(0)},$$

donde  $\pi$  es la única distribución estacionaria del proceso y puede encontrarse resolviendo el sistema de ecuaciones dado por

$$\begin{aligned} \pi \times (\mathbf{I} - \mathbf{P}) &= 0, \\ \sum \pi_i &= 1. \end{aligned}$$



# Tasa de Entropía

## Definición (Tasa de entropía: $H(\mathcal{X})$ )

La *tasa de entropía* de un proceso  $\{X_i\}$  es el límite

$$H(\mathcal{X}) = \lim_{n \rightarrow \infty} \frac{1}{n} H(X_1, \dots, X_n).$$

## Ejemplo

Si  $\{X_i\}$  son i.i.d,

$$H(\mathcal{X}) = \lim_{n \rightarrow \infty} \frac{1}{n} H(X_1, \dots, X_n) = \lim_{n \rightarrow \infty} \frac{1}{n} n H(X_1) = H(X_1).$$

## Tasa de Entropía (alternativa)

### Definición (Tasa de entropía: $H'(\mathcal{X})$ )

Una definición alternativa de *tasa de entropía*, que intenta capturar un concepto similar es

$$H'(\mathcal{X}) = \lim_{n \rightarrow \infty} H(X_n | X_{n-1}, \dots, X_1).$$

En general  $H'(\mathcal{X})$  y  $H(\mathcal{X})$  no tienen por qué coincidir.

### Teorema

Si  $\{X_i\}$  es un proceso estocástico **estacionario**,  $H'(\mathcal{X})$  y  $H(\mathcal{X})$  existen y son iguales.

## Teorema

Si  $\{X_i\}$  es un proceso estocástico estacionario,  $H(X_n|X_{n-1}, \dots, X_1) \searrow H'(\mathcal{X})$ .

## Demostración.

$$\begin{aligned} H(X_{n+1}|X_n, \dots, X_2, X_1) &\leq H(X_{n+1}|X_n, \dots, X_2) \\ &= H(X_n|X_{n-1}, \dots, X_1) \end{aligned}$$



## Tasa de Entropía de Procesos Estacionarios

### Teorema

Si  $\{X_i\}$  es un proceso estocástico estacionario,  $H'(\mathcal{X})$  y  $H(\mathcal{X})$  existen y son iguales.

### Demostración.

$$\begin{aligned}\lim_{n \rightarrow \infty} \frac{1}{n} H(X_n, \dots, X_1) &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n H(X_i | X_{i-1}, \dots, X_1) \\ &\rightarrow H'(\mathcal{X}) \quad (\text{Cesáro}).\end{aligned}$$



### Teorema (Promedio a la Cesáro)

Si  $a_n \rightarrow a$ , entonces  $\bar{a}_n = \frac{1}{n} \sum_{i=1}^n a_n \rightarrow a$ .

## Tasa de Entropía de Procesos de Markov

### Teorema

Si  $\{X_i\}$  es un proceso de Markov estacionario con distribución estacionaria  $\pi$ ,

$$H(\mathcal{X}) = H'(\mathcal{X}) = - \sum_{ij} \pi_i \mathbf{P}_{ij} \log \mathbf{P}_{ij}.$$

### Demostración.

$$\begin{aligned} \lim_{n \rightarrow \infty} H(X_n | X_{n-1}, \dots, X_1) &= \lim_{n \rightarrow \infty} H(X_n | X_{n-1}) \\ &= H(X_2 | X_1) \\ &= \sum_i \pi_i H(X_2 | X_1 = i) \\ &= \sum_i \pi_i \sum_j \mathbf{P}_{ij} \log \frac{1}{\mathbf{P}_{ij}}. \end{aligned}$$

