Practica 2(2021)

Problema 1:

Enunciado: Demuestre que la divergencia de Kullback-Leibler puede no ser simétrica y no cumplir con la desigualdad triangular. ¿ En qué casos la divergencia de Kullback-Leibler se va a infinito?

Resolucion: Contraejemplo: Sistema de monedas cargadas:

$$\begin{aligned} x_1 &\in A_{X_1} = \{H, T\}, \text{Prob de } H = p_H = 1/10 \\ x_2 &\in A_{X_2} = \{H, T\}, \text{Prob de } H = q_H = 2/10 \\ x_3 &\in A_{X_3} = \{H, T\}, \text{Prob de } H = r_H = 3/10 \end{aligned}$$

La divergencia para dos monedas cargadas con distinto peso p y q de H:

$$D_{KL}(p \mid\mid q) = p \log \frac{p}{q} + (1-p) \log \frac{(1-p)}{(1-q)}$$

Para la desigualdad triangular:

$$D_{KL}(p \mid\mid q) + D_{KL}(q \mid\mid r) \approx 0,0271$$

$$D_{KL}(p \mid\mid r) \approx 0.0370$$

NO satisface la desigualdad triangular

$$D_{KL}(p \mid\mid q) + D_{KL}(q \mid\mid r) \approx 0,0271 \leq D_{KL}(p \mid\mid r)$$

Para simetría:

$$D_{KL}(p \mid\mid q) \approx 0.037, \quad D_{KL}(q \mid\mid p) \approx 0.019$$

son distintas, D_{KL} es simétrica.

Problema 2:

Enunciado: Utilizando la desigualdad de Jensen, demuestre que una variable aleatoria es determinista si y sólo si su varianza se anula.

Resolucion: Un variable aleatoria es determinista $\Leftrightarrow$ Su varianza es nula

Demo:

como $f(x) = x^2$ es estrictamente convexa:

$$X \text{ es det } \Leftrightarrow f(\langle X \rangle) = \langle f(X) \rangle \Leftrightarrow Var = \langle X^2 \rangle - \langle X \rangle^2 = 0$$

Problema 3

Enunciado: Demuestre que el promedio aritmético de un conjunto de nümeros reales positivos es mayor o igual que su promedio geométrico. **Resolucion:** Demuestre que el promedio aritmético de un conjunto de números reales positivos es mayor o igual que su promedio geométrico.

Demo:

$$\begin{aligned} \langle \log x_i \rangle &\leq \log \langle x_i \rangle \\ \frac{1}{N} \sum \log x_i &\leq \log \langle x_i \rangle \\ \frac{1}{N} \log \prod x_i &\leq \log \langle x_i \rangle \\ \log \sqrt[n]{\prod x_i} &\leq \log \langle x_i \rangle \\ \sqrt[n]{\prod x_i} &\leq \langle x_i \rangle \\ \sqrt[n]{\prod x_i} &\leq \frac{1}{N} \sum x_i \end{aligned}$$

que es lo que se buscaba

Problema 4

Enunciado: Demuestre que la entropia es una función cóncava de p. Es decir, si se consideran dos distribuciones p_1 y p_2, entonces se cumple que para todo $\lambda \in [0, 1]$

$$H\left[\lambda p_1 + \{1-\lambda) p_2\right] \geq \lambda H\left(p_1\right) + (1-\lambda) H\left(p_2\right).$$

Resolucion: Theorem: The cross-entropy is convex in the probability distribution q, i.e.

$$\mathrm{H}\left[p, \lambda q_1 + (1-\lambda) q_2\right] \leq \lambda \mathrm{H}\left[p, q_1\right] + (1-\lambda) \mathrm{H}\left[p, q_2\right]$$

where p is a fixed and q_1 and q_2 are any two probability distributions and $0 \leq \lambda \leq 1$. Proof: The relationship between Kullback-Leibler divergence, entropy and cross-entropy is:

$$\mathrm{KL}[P\|Q]=\mathrm{H}(P,Q)-\mathrm{H}(P)$$

Note that the KL divergence is convex in the pair of probability distributions (p,q) :

$$\mathrm{KL}\left[\lambda p_1+(1-\lambda)p_2\|\lambda q_1+(1-\lambda)q_2\right]\leq\lambda\mathrm{KL}\left[p_1\|q_1\right]+(1-\lambda)\mathrm{KL}\left[p_2\|q_2\right]$$

A special case of this is given by

$$\begin{aligned}\mathrm{KL}\left[\lambda p+(1-\lambda)p\|\lambda q_1+(1-\lambda)q_2\right]&\leq\lambda\mathrm{KL}\left[p\|q_1\right]+(1-\lambda)\mathrm{KL}\left[p\|q_2\right]\\ \mathrm{KL}\left[p\|\lambda q_1+(1-\lambda)q_2\right]&\leq\lambda\mathrm{KL}\left[p\|q_1\right]+(1-\lambda)\mathrm{KL}\left[p\|q_2\right]\end{aligned}$$

and applying equation (2), we have

$$\begin{aligned}\mathrm{H}\left[p,\lambda q_1+(1-\lambda)q_2\right]-\mathrm{H}[p]&\leq\lambda\left(\mathrm{H}\left[p,q_1\right]-\mathrm{H}[p]\right)+(1-\lambda)\left(\mathrm{H}\left[p,q_2\right]-\mathrm{H}[p]\right)\\ \mathrm{H}\left[p,\lambda q_1+(1-\lambda)q_2\right]-\mathrm{H}[p]&\leq\lambda\mathrm{H}\left[p,q_1\right]+(1-\lambda)\mathrm{H}\left[p,q_2\right]-\mathrm{H}[p]\\ \mathrm{H}\left[p,\lambda q_1+(1-\lambda)q_2\right]&\leq\lambda\mathrm{H}\left[p,q_1\right]+(1-\lambda)\mathrm{H}\left[p,q_2\right]\end{aligned}$$

which is equivalent to (1).

Problema 5

Enunciado: Cota de la independencia: Demuestre que

$$H\left(X_1,X_2,\ldots,X_N\right)\leq\sum_i H\left(X_i\right)$$

Resolucion: Usando

$$H\left(X_1,X_2,\ldots,X_n\right)=\sum_{i=1}^n H\left(X_i\mid X_{i-1},\ldots,X_1\right)$$

Sabemos que:

$$H(X\mid Y)\leq H(X)$$

Cuando X e Y son independientes: generalizando: