
Entropía

LAB-PEP

Historia

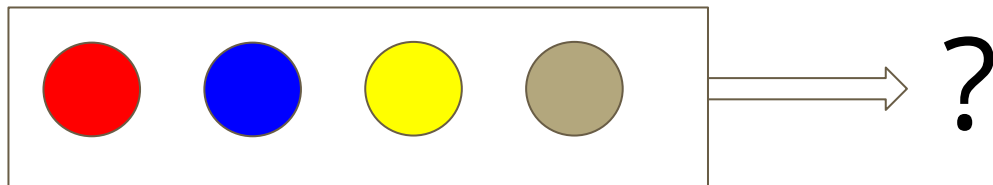
Concepto de la disipación de energía útil (Lazare Carnot - 1803)

Segunda ley de la termodinámica (Rudolf Clausius - 1854): Todo sistema aislado evoluciona a un estado de mayor *desorden*

Mecánica estadística (Ludwig Boltzmann - 1872): Todo sistema aislado evoluciona a un estado de más probable

Teoría de la información (Claude Shannon - 1948): Todo sistema aislado evoluciona a un estado de mayor información perdida (Information loss increases)

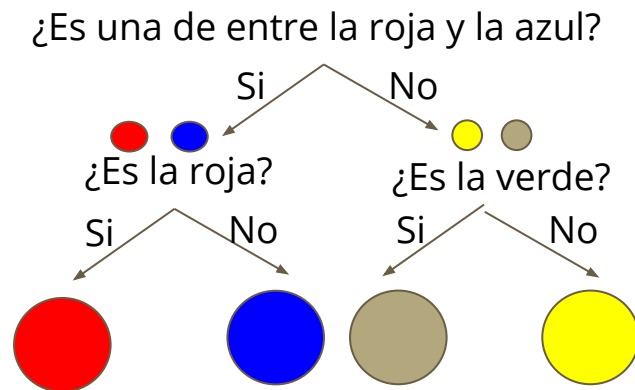
Máxima entropía: Todas igual probabilidad



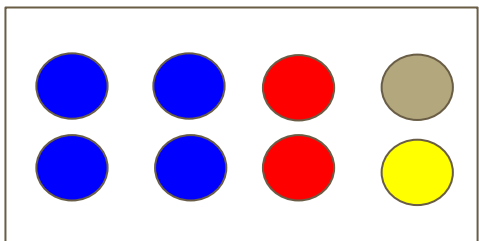
- 4 bolillas metidas en una bolsa.
- Se saca una de ellas al azar
- Todas las bolillas tienen la misma probabilidad: $\frac{1}{4}$

¿Cuál es el mínimo de preguntas (si/no) que debo hacer para averiguar cuál es la que se sacó?

$$S = 2 \times \frac{1}{4} + 2 \times \frac{1}{4} + 2 \times \frac{1}{4} + 2 \times \frac{1}{4} = 2$$

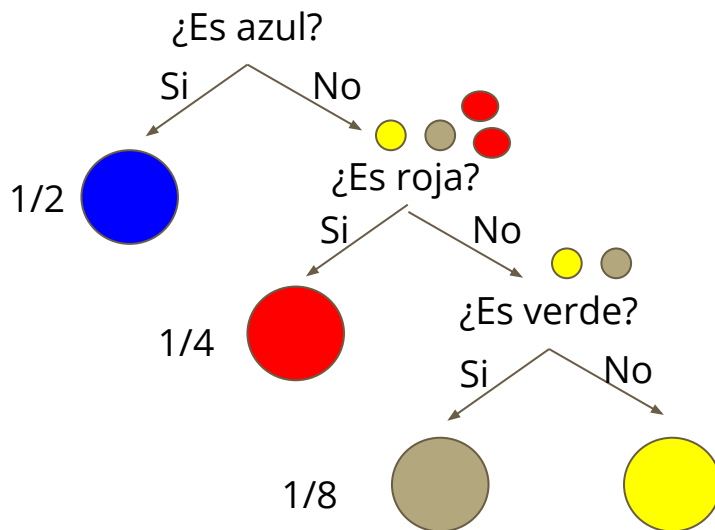


Disminuimos entropía



$$S = 1 \times 1/2 + 2 \times 1/4 + 3 \times 1/8 + 3 \times 1/8 = 1.75$$

$$S = \frac{1}{2} \log_2(2) + \frac{1}{4} \log_2(4) + \frac{1}{8} \log_2(8) + \frac{1}{8} \log_2(8)$$



Si saque azul -> una pregunta
Si saque rojo -> 2 preguntas
Si saque verde -> 3 preguntas
Si saque amarillo -> 3 preguntas

Teoría de la información

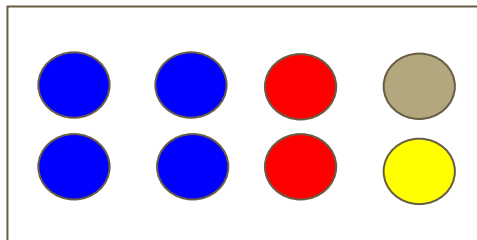
Cuanto vale la entropía si todas las bolillas son azules? Cuantas preguntas tengo que hacer para saber el color?

$$H(X) = \mathbb{E}_X[I(x)] = - \sum_{x \in \mathbb{X}} p(x) \log p(x)$$

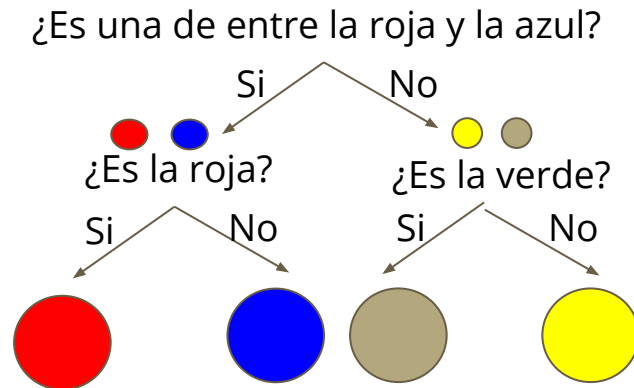
log 1/p es la cantidad de información y si el logaritmo está en base 2 las unidades son bits

Entropía cruzada

Cantidad de preguntas promedio utilizando una estrategia óptima de otra distribución



$$S = 2 \times 1/8 + 2 \times 1/8 + 2 \times 1/4 + 2 \times 1/2 = 2$$



Definición formal

Dadas dos distribuciones de probabilidad, la entropía cruzada mide el grado de similitud entre las distribuciones

Se basa en KL-divergence $D_{KL}(p||q)$ (Kullback-Leibler)

$$H(p, q) = - \sum_{i=1}^N p_i \log(q_i) = H(p) + D_{KL}(p||q)$$

$$D_{KL}(p||q) = \sum_i p_i \log(p_i/q_i)$$

La divergencia KL es cero cuando las dos distribuciones son iguales.