Entropía

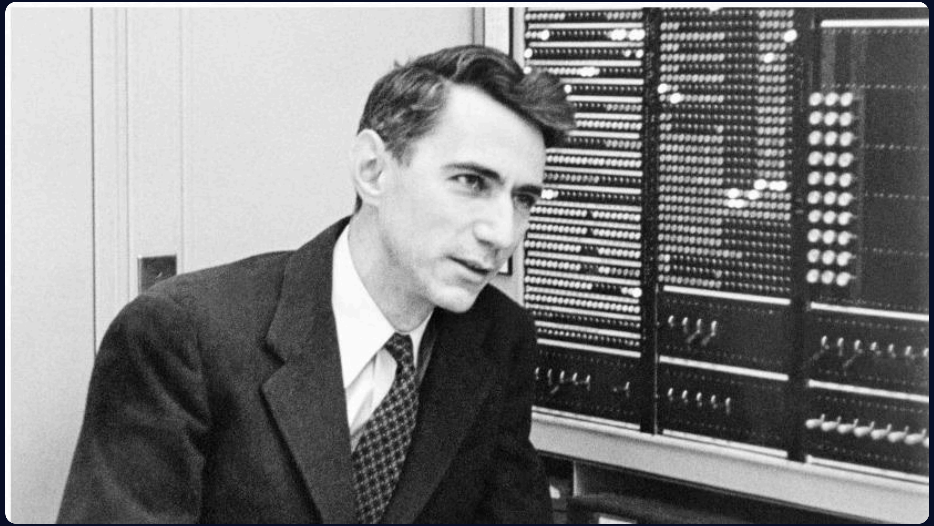
MÓDULO 4

Técnicas computacionales avanzadas para modelar fenómenos sociales

Concentración en Economía Aplicada y Ciencia de Datos
ITESM-SF | CDMX, Feb-Jun 2026

HISTORIA

Entropía de Shannon



Claude Shannon (1948)

En julio de 1948, Claude Shannon publicó "A Mathematical Theory of Communication" en el Bell System Technical Journal, estableciendo las bases de la teoría de la información moderna.

El problema fundamental de la comunicación es reproducir en un punto un mensaje seleccionado en otro punto, ya sea exacta o aproximadamente.

El Problema Fundamental

> "El aspecto significativo es que el mensaje real es uno **seleccionado de un conjunto de mensajes posibles**. El sistema debe diseñarse para operar con cada selección posible, no solo con la que realmente se elegirá."

Los aspectos semánticos de la comunicación son irrelevantes para el problema de ingeniería. Lo importante es cuantificar la información usando una medida logarítmica, como señaló Hartley.

Intuición: El Cumpleaños

Consideremos tres declaraciones sobre el cumpleaños de mi hermano:

Declaración 1

El cumpleaños es en un día particular del año

Probabilidad: $P_1 = 1$

Declaración 2

El cumpleaños es en la segunda mitad del año

Probabilidad: $P_2 = 1/2$

Declaración 3

El cumpleaños es el día 25 de algún mes

Probabilidad: $P_3 = 12/365$

¿Cómo cuantificamos la utilidad de esta información?

Si las variables aleatorias X y Y son independientes, la probabilidad conjunta de los eventos 2 y 3 es: $P_{2,3} = 1/2 \times 12/365 = 6/365$

Relación entre Probabilidad e Información

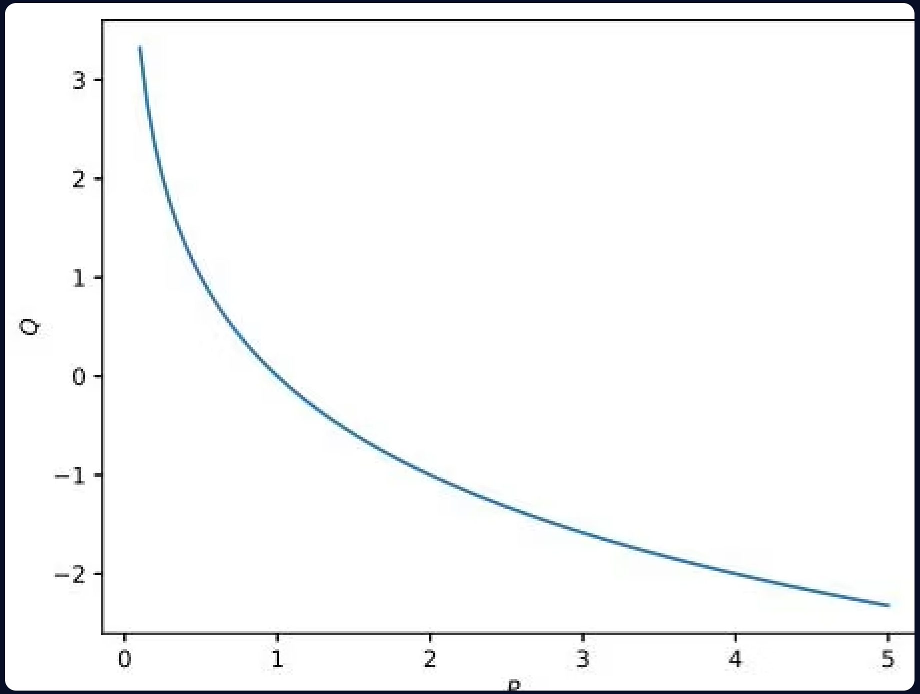
Definición Matemática

La **información** Q contenida en una declaración está definida por:

$$Q = -\log_2 P$$

donde P es la probabilidad de que esa declaración sea cierta.

- Si P aumenta, Q disminuye
- Si P disminuye, Q aumenta
- Se usa $\log_2$ porque la información se mide en **bits**



El gráfico muestra la función convexa decreciente $Q = -\log_2 P$, ilustrando cómo la información aumenta cuando la probabilidad disminuye.

FÓRMULA CLAVE

La Entropía de Shannon

Para un conjunto de declaraciones con probabilidad P_i y correspondiente información $Q_i = -\log P_i$, el promedio ponderado de la información contenida está dado por:

$$H = \sum_{i=1}^{n} Q_i P_i = -\sum_{i=1}^{n} P_i \log P_i$$

Condiciones

- $\sum_i P_i = 1$
- $P_i \geq 0$ con $1 \leq i \leq n$

Significado

H es la **entropía de Shannon**, que cuantifica la cantidad promedio de información obtenida al medir una cantidad particular.

Ejemplo: Dados Normal vs. Cargado

Dado Normal

Probabilidades iguales:
1/6 para cada salida (1, 2, 3, 4, 5, 6)

Información por salida:
$Q = -\log_2(1/6) = \log_2 6$

Entropía de Shannon:

$$H = \sum_{i=1}^{6} \frac{1}{6} \log_2 6 = 2.58$$

Dado Cargado

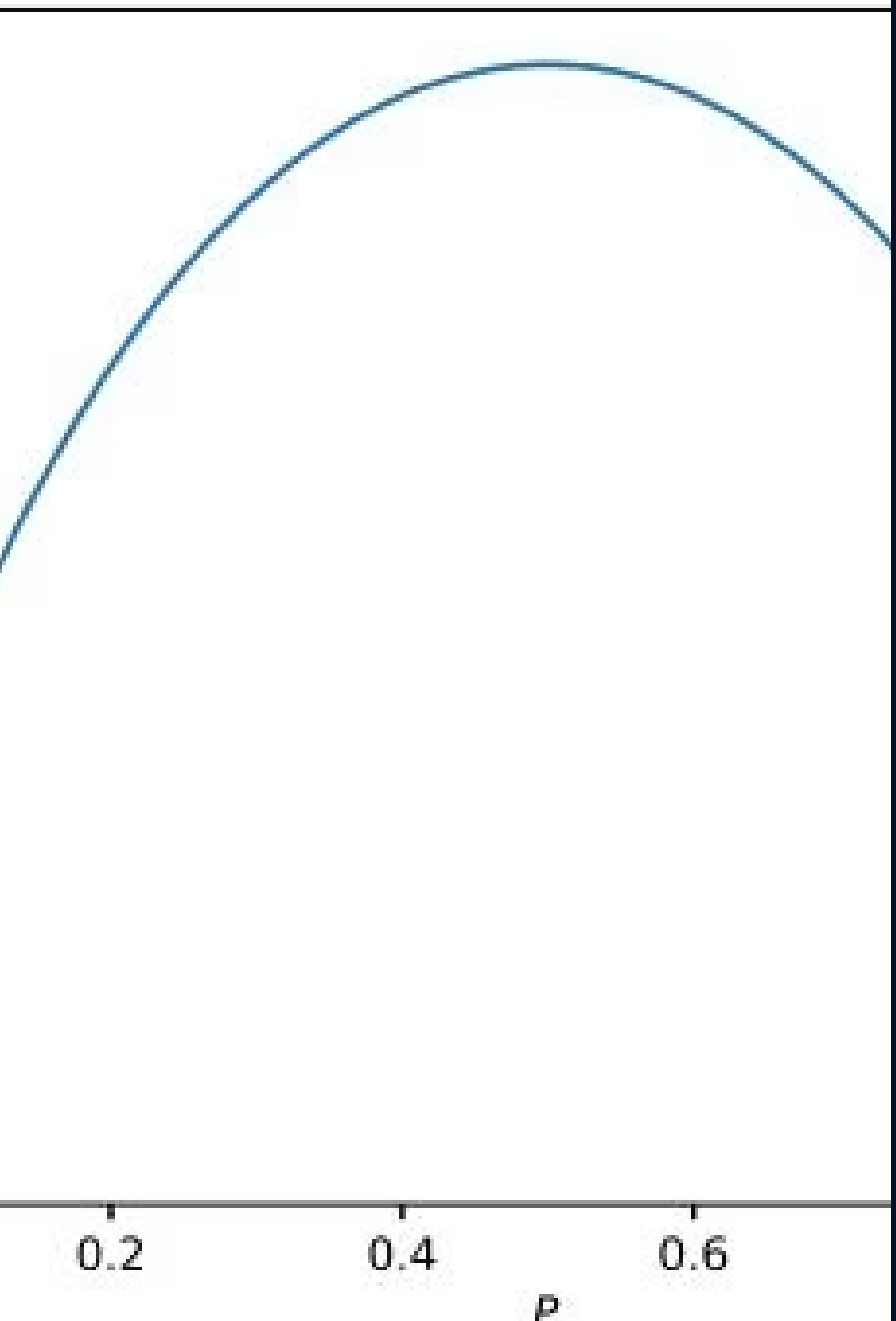
Probabilidades desiguales:
1/10, 1/10, 1/10, 1/10, 1/10, 1/2

Información variable:
$\log_2 10$ para cinco salidas
$\log_2 2$ para una salida

Entropía de Shannon:

$$H = 5 \times \frac{1}{10} \log_2 10 + \frac{1}{2} \log_2 2 = 1.16$$

Conclusión: La entropía disminuye cuando las probabilidades no son iguales en el dado. Mayor certeza = menor entropía.



Ejemplo: La Moneda

Consideremos la tirada de una moneda con una cara de probabilidad P y la otra de probabilidad 1-P. La entropía del sistema es:

$$H(P) = -P\log_2 P - (1-P)\log_2(1-P)$$

1

Máximo en P = 1/2

Mucha incertidumbre sobre la salida, mayor información ganada

2

Mínimo en P = 0 o P = 1

Poca incertidumbre sobre la salida, menor información ganada

DEFINICIÓN FORMAL

Entropía de Variable Aleatoria

Supongamos que X es una variable aleatoria discreta que toma valores en un conjunto finito $\mathcal{X}$. La entropía de la variable aleatoria X está definida por:

$$H(X) = -\sum_{x \in \mathcal{X}} \Pr[x] \log_2 \Pr[x]$$

Ejemplo 1

$\mathcal{X}_1 = \{\text{águila, sol}\}$

Ejemplo 2

$\mathcal{X}_2 = \{1, 2, 3, 4, 5, 6\}$

Entropía Condicional

Supongamos que X y Y son dos variables aleatorias. Para cualquier realización y de Y, la probabilidad condicional X|y tiene entropía:

$$H(X|y) = -\sum_{x} \Pr[x|y] \log_2 \Pr[x|y]$$

La **entropía condicional** H(X|Y) es el promedio ponderado (respecto a las probabilidades Pr[y]) de la entropía H(X|y) sobre todos los valores posibles y:

$$H(X|Y) = -\sum_{y} \sum_{x} \Pr[y] \Pr[x|y] \log_2 \Pr[x|y]$$

> La entropía condicional mide el promedio de la información acumulada acerca de X que no es revelada por Y.

Referencia: Stinson, D. R. (2005). Cryptography: theory and practice. Chapman and Hall/CRC.