

# T2.4 Entropía y entropía cruzada

## Índice

### 1 Entropía para variables aleatorias discretas

#### 1.1 Definición

#### 1.2 Distribución de máxima entropía

#### 1.3 Distribución de mínima entropía

#### 1.4 Función entropía binaria

### 2 Entropía cruzada

# 1 Entropía para variables aleatorias discretas

## 1.1 Definición

**Entropía de una variable aleatoria discreta  $X$ :** con distribución  $p$  sobre  $K$  estados

$$\mathbb{H}(X) = - \sum_k p(X = k) \log_2 p(X = k) = -\mathbb{E}_X[\log p(X)]$$

**Unidades de medida:** **bits** si usamos logaritmos binarios; **nats** si usamos logaritmos naturales

**Ejemplo:**  $X \in \{1, \dots, 5\}$  con  $p = [0.25, 0.25, 0.2, 0.15, 0.15]$

$$\begin{aligned}\mathbb{H}(X) &= -2 \frac{1}{4} \log_2 \frac{1}{4} - \frac{1}{5} \log_2 \frac{1}{5} - 2 \frac{3}{20} \log_2 \frac{3}{20} \\ &= -\frac{1}{2}(-2) + \frac{1}{5} \log_2 5 - \frac{3}{10} (\log_2 3 - \log_2 20) \\ &= 1 + 0.2 \cdot 2.3219 - 0.3 (1.5850 - 4.3219) \\ &= 2.29 \text{ bits}\end{aligned}$$

## 1.2 Distribución de máxima entropía

**Distribución de máxima entropía:** es la uniforme

$$p(X = k) = \frac{1}{K} \quad \text{para } X \text{ } K\text{-aria}$$

$$\mathbb{H}(X) = - \sum_{k=1}^K \frac{1}{K} \log \frac{1}{K} = - \log \frac{1}{K} = \log K$$

## 1.3 Distribución de mínima entropía

**Distribución de mínima entropía:** cualquier función delta que asigne toda la masa en un estado  $k^*$

$$p(X = k) = \delta(X = x^*)$$

$$\mathbb{H}(X) = -1 \log 1 - \sum_{k \neq k^*} 0 \log 0 = 0$$

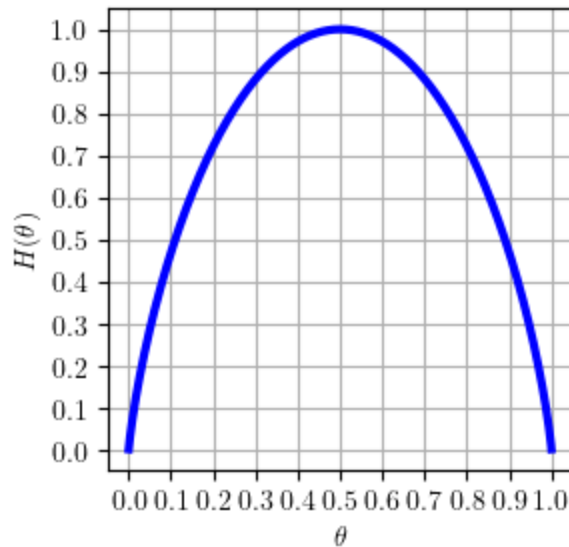
## 1.4 Función entropía binaria

**Función entropía binaria:** sea la variable binaria  $X \in \{0, 1\}$  tal que  $p(X = 1) = \theta$  y  $p(X = 0) = 1 - \theta$

$$\begin{aligned}\mathbb{H}(\theta) &= \mathbb{H}(X) \\ &= -[p(X = 1) \log_2 p(X = 1) + p(X = 0) \log_2 p(X = 0)] \\ &= -[\theta \log_2 \theta + (1 - \theta) \log_2 (1 - \theta)]\end{aligned}$$

$\mathbb{H}(\theta)$  alcanza su máximo de 1 bit con  $\theta = 0.5$ :

```
In [4]: import numpy as np; import matplotlib.pyplot as plt
plt.rcParams['text.usetex'] = True; x = np.linspace(1e-4, 1-1e-4, 1000)
H = lambda x: -(x*np.log2(x) + (1-x) * np.log2(1-x))
plt.figure(figsize=(3, 3)); plt.plot(x, H(x), '-b', lw=3); plt.grid();
ticks = np.linspace(0, 1, 11); plt.xticks(ticks); plt.yticks(ticks);
plt.xlabel(r'$\theta$'); plt.ylabel(r'$H(\theta)$');
```



## 2 Entropía cruzada

**Entropía cruzada:** entre una distribución  $p$  y otra  $q$  es

$$\mathbb{H}(p, q) = - \sum_k p_k \log q_k$$

**Interpretación:** número esperado de bits necesario para comprimir muestras de  $p$  mediante un cuantificador basado en  $q$

**Mínimo:** con  $q = p$ , cumpliéndose  $\mathbb{H}(p, p) = \mathbb{H}(p)$