

Two States, A or B.

A: 50%, B: 50%

$$0.5 \rightarrow 1, \quad \frac{1}{0.5}$$

$$0.1 \rightarrow 1 \\ \leftarrow \frac{1}{0.1}$$

告诉你是那个. 0 代表 A, 1 代表 B. 也可以是 Rainy 代表 A, 但有效信息只有 1 bit.

不确定性减少.  $0.5 \rightarrow 1$

2 种状态减少到 1 种.

度量:  $\log_2 \frac{1}{0.5} = 1 \text{ bit}$

$$\log_2 \frac{1}{0.5} = 1 \text{ bit. 字节,}$$

Shannon: Reduce the uncertainty by a factor of 2.

基准.

$$\log_{10}$$

$$\log$$

A: 75%, B: 25%

告诉你 B 发生. 信息量  $\log_2 \frac{1}{0.25} = 2$

... A ...  $\log_2 \frac{1}{0.75} \approx 0.41$

平均信息含量:  $0.25 \times 2 + 0.75 \times 0.41$

$$\text{Entropy} \uparrow: \left[ - \sum_i p_i \log p_i \right]$$

$$\log_2 \frac{1}{p} = -\log_2 p$$

晴 A 多云 B 下雨 C 下雪 D

Entropy:  $(-0.25 \log 0.25) \times 4 = 2$  bits

$$(2 \times 0.25) \times 4 = 2 \text{ bits.}$$

Entropy :  $-(0.5 \cdot \log 0.5 + 0.25 \cdot \log 0.25 + 0.125 \log 0.125 + 0.125 \log 0.125)$   
 $= 1.75$



	A	B	C	D
A	00	01	10	11

例 用 3 2 bits, 传递 3 1.75 bits 信息。

$$\begin{aligned} & \underline{0.5 \times 1} + \underline{0.25 \times 2} + \underline{0.125 \times 3} \\ & + \underline{0.125 \times 3} \\ & = \underline{1.75} \end{aligned}$$

A B C D ?

110 0 10 111

# Cross Entropy

A	B	C	D
110	0	10	111
0.5	0.25	0.125	0.125

predicted dist.

A	B	C	D
$\frac{1}{2^3}$	$\frac{1}{2^1}$	$\frac{1}{2^2}$	$\frac{1}{2^3}$

$$H(p, q) = \underbrace{0.5}_{\text{orange}} \log \underbrace{0.125}_{\text{purple}} + \underbrace{0.25}_{\text{purple}} \log \underbrace{0.5}_{\text{green}} + \underbrace{0.125}_{\text{orange}} \log \underbrace{0.25}_{\text{green}} + \underbrace{0.125}_{\text{orange}} \log \underbrace{0.125}_{\text{purple}} \boxed{2.375}$$

$\approx 2.18$