

Two States, A or B.

A: 50%, B: 50%

告诉你的是那个. 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:  $\log_2 \frac{1}{0.5} = 1 \text{ bit}$

Shannon: Reduce the  
uncertainty by a factor  
of 2.

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:  $0.25 \times 2 + 0.75 \times 0.41$

Entropy  $\uparrow$ :  $-\sum_i p_i \log p_i$

# Cross Entropy

A	B	C	D
0.25	0.25	0.25	0.25
00	01	10	11

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

发送信息用的资源:

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

A	B	C	D
0.5	0.25	0.125	0.125

$$\begin{aligned} \text{Entropy: } & -(0.5 \cdot \log 0.5 + 0.25 \cdot \log 0.25 \\ & + 0.125 \log 0.125 + 0.125 \log 0.125) \\ & = 1.75 \end{aligned}$$

若仍编码

A	B	C	D
00	01	10	11

则用了 2 bits, 传递了 1.75 bits 信息.

A	B	C	D
0	10	110	111

$$0.5 \times 1 + 0.25 \times 2 + 0.125 \times 3$$

$$+ 0.125 \times 3$$
$$= 1.75$$

问:

A	B	C	D	?
0	1	10	11	?

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
$\frac{1}{2^3}$	$\frac{1}{2^1}$	$\frac{1}{2^2}$	$\frac{1}{2^3}$

$$H(p, q) = 0.5 \log 0.125 + 0.25 \log 0.5 + 0.125 \log 0.25 + 0.125 \log 0.125$$

$\approx 2.375$   
 ~~$\approx 2.18$~~