

# 理解交叉熵

也是：“不确定”的程度

熵：一个系统从不确定  $\rightarrow$  确定的难度

熵越大，越“混乱”，或就难度越大

	C9	Faze	NAVI	FPX
事物的 A组	$\frac{1}{2}$	$\frac{1}{2}$	B组 $\frac{99}{100}$	$\frac{1}{100}$
单个信息量：	$-\log_2 \frac{1}{2} = 1$	$-\log_2 \frac{1}{2} = 1$	$-\log_2 \frac{99}{100} \approx 0.0145$	$-\log_2 \frac{1}{100} \approx 6.6439$

对系统“最大”的信息量：  
 $1 \times \frac{1}{2}$        $1 \times \frac{1}{2}$        $0.0145 \times \frac{99}{100}$        $6.6439 \times \frac{1}{100}$

+

+

所以熵：

相对熵 } 以 P 系统为基准，到 P 系统的“难度程度”  
 KL 散度 } 以人为基准 / 某系统

$X - Y > 0$  一定

$X \downarrow$  越好

交叉熵

极大似然估计法

形式一致，含义不同

True

Predict

$$H(P, Q) = \sum_{i=1}^n P_i (-\log_2 Q_i)$$

$$= -\sum_{i=1}^n (x_i \cdot \log_2 y_i + (1-x_i) \cdot \log_2 (1-y_i))$$