

量化一个信号包含了多少信息

基本想法：小概率事件发生能提供更懂信息

我们想要通过这种基本想法来量化信息。特别地，

- 非常可能发生的事件信息量要比较少，并且极端情况下，确保能够发生的事件应该没有信息量。
- 较不可能发生的事件具有更高的信息量。
- 独立事件应具有增量的信息。例如，投掷的硬币两次正面朝上传递的信息量，应该是投掷一次硬币正面朝上的信息量的两倍。

自信息 $I(x)$ 的单位是 nats

一奈特是以 $\frac{1}{e}$ 的概率观测到一个事件时获得的信息量。

KL散度

衡量两个概率分布的差异

同一个随机变量 x 有两个单独的概率分布 $P(x)$ 和 $Q(x)$

$$D_{\text{KL}}(P||Q) = \mathbb{E}_{x \sim P} \left[\log \frac{P(x)}{Q(x)} \right] = \mathbb{E}_{x \sim P} [\log P(x) - \log Q(x)].$$