11.4 Entropy

Friday, September 28, 2012
11:57 AM

[bits (per symbol)]

$$H(X) = \sum_x p_X(x)\left(-\log_2 p_X(x)\right) = \mathbb{E}\left[-\log_2 p_X(X)\right]$$

convention: $0 \log 0 = 0$ $\left(\lim_{x\to 0} x \log x = 0\right)$

quantify/measure - how random a RV is
- the amount of uncertainty a RV has

computation: In MATLAB, define a row vector $\underline{p}$ of pmf of X

Can calculate entropy

HX = -p * (log2(p).')

Ex. Uniform discrete X on 1, ..., n

$$\text{pmf}: \quad p_X(x) = \begin{cases} 1/n, & x = 1, \ldots, n \\ 0, & \text{otherwise} \end{cases}$$

$$H(X) = \mathbb{E}\left[-\log_2 p_X(X)\right] = \mathbb{E}\left[-\log_2 \frac{1}{n}\right] = -\log_2 \frac{1}{n} = \log_2 n$$

another way

$$= -\sum_x p_X(x) \log_2 p_X(x) = -\sum_{x=1}^{n} \frac{1}{n} \log_2 \frac{1}{n} = -\cancel{n} \times \frac{1}{\cancel{n}} \log_2 \frac{1}{n} = \log_2 n$$

Fact :

$$0 \le H(X) \le \log_2 |S_X|$$

with equality iff X is degenerated (not random)

with equality iff $X \sim$ uniform

Multiple RVs :

$$H(X,Y) = -\mathbb{E}\left[\log_2 p_{X,Y}(X,Y)\right]$$

$$= -\sum\sum p_{X,Y}(x,y) \log_2 p_{X,Y}(x,y)$$

$\propto y$

$$H(X_1, \ldots, X_n) = -\mathbb{E}\left[\log_2 p_{X_1, X_2, \ldots, X_n}(X_1, X_2, \ldots, X_n)\right]$$

$$H(X_1^n) = -\mathbb{E}\left[\log_2 p_{X_1^n}(X_1^n)\right]$$

$$\left[I(X;Y) = H(X) + H(Y) - H(X,Y) \right]$$

$\uparrow$ mutual information

when $X_1, \ldots, X_n$ are indp.,

$$H(X_1^n) = -\mathbb{E}\left[\log_2 \prod_{k=1}^{n} p_{X_k}(X_k)\right] = -\mathbb{E}\left[\sum_{k=1}^{n} \log_2 p_{X_k}(X_k)\right]$$

$$= \sum_{k=1}^{n} \underbrace{-\mathbb{E}\left[\log_2 p_{X_k}(X_k)\right]}_{H(X_k)} = \sum_{k=1}^{n} H(X_k)$$

When $X_1, \ldots, X_n$ are i.i.d.,

$$H(X_1^n) = n H(X_1)$$

Fact: $H(X_1^n) \leq \sum_{k=1}^{n} H(X_k)$

$\uparrow$ with equality when $X_1, X_2, \ldots, X_n$ are indp.

$$H(X,Y) \leq H(X) + H(Y)$$

Source coding theorem:

Let $L(c, X) = \mathbb{E}[\ell(X)]$ when code c is used to encode R.V. X

$$L^*(X) = \min_{c} L(c, X) = L(c_{\text{Huffman}}, X)$$

$\uparrow$ over all UD codes

$$H(X) \leq L^*(X) < H(X) + 1$$

source extension

$$H(X_1^n) \leq L^*(X_1^n) < H(X_1^n) + 1$$

$$nH(X_1) \leq L^*(X_1^n) < nH(X_1) + 1$$

$$H(X_1) \leq \frac{L^*(X_1^n)}{n} < H(X_1) + \frac{1}{n}$$

As $n \to \infty$, $\frac{L^*(X_1^n)}{n} \to H(X_1)$