Entropy quick reference

Paul Dumais, October 2002

H1 DEFINITION (*entropy*) Let X be a random variable taking the values $x_1, \ldots, x_m$ with probabilities $p_1, \ldots, p_m$ respectively, with $\forall i : 0 < p_i \leq 1$. Then

$$H(X) \stackrel{\text{def}}{=} H(p_1, \ldots, p_m) \stackrel{\text{def}}{=} \sum_{i=1}^{m} p_i \lg(\frac{1}{p_i}).$$

H2 $0 \leq H(X) \leq \lg(m)$. $H(X) = 0$ iff $m = 1$, $p_1 = 1$. $H(X) = \lg(m)$ iff $p_1 = \ldots = p_m = \frac{1}{m}$.

H3 If $a = p_1 + \ldots + p_k$ and $1 - a = p_{k+1} + \ldots + p_m$ then

$$H(X) = H(a, 1-a) + aH(\frac{p_1}{a}, \ldots, \frac{p_k}{a}) + (1-a)H(\frac{p_{k+1}}{1-a}, \ldots, \frac{p_m}{1-a}).$$

H4 (*Concavity*) Let $(p_1, \ldots, p_m)$ and $(q_1, \ldots, q_m)$ be two distributions, and $0 \leq a \leq 1$. Then

$$H(a(p_1, \ldots, p_m) + (1-a)(q_1, \ldots, q_m)) \geq aH(p_1, \ldots, p_m) + (1-a)H(q_1, \ldots, q_m).$$

H5 DEFINITIONS (*entropy of a joint distribution, conditional entropy, mutual information*) Let X and Y be a pair of random variables taking the value (x_i, y_j) for $1 \leq i \leq m$, $1 \leq j \leq n$, with $\forall i, j : 0 < \text{Prob}[X = x_i, Y = y_j] \leq 1$. Then

$$\begin{aligned}
H(X,Y) &\stackrel{\text{def}}{=} \sum_{i=1}^{m}\sum_{j=1}^{n} \text{Prob}[X = x_i, Y = y_j] \lg(\frac{1}{\text{Prob}[X = x_i, Y = y_j]}), \\
H(X|Y = y) &\stackrel{\text{def}}{=} \sum_{i=1}^{m} \text{Prob}[X = x_i | Y = y] \lg(\frac{1}{\text{Prob}[X = x_i | Y = y]}), \\
H(X|Y) &\stackrel{\text{def}}{=} \sum_{j=1}^{n} \text{Prob}[Y = y_j] H(X|Y = y_j), \\
I(X:Y) &\stackrel{\text{def}}{=} H(X) + H(Y) - H(X,Y)
\end{aligned}$$

H6 $H(X,Y) \leq H(X) + H(Y)$. $H(X,Y) = H(X) + H(Y)$ iff X and Y are independant.

H7 $H(X|Y) = H(X,Y) - H(Y)$.

H8 $H(X|Y) \leq H(X)$.