Massachusetts Institute of Technology
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Spring 2005)

Problem Set 4G: Solutions
Due: February 23, 2005

G1. (a) Using the inequality $\ln a \le a - 1$, we have

$$-\sum_{i=1}^{n} p_i \ln p_i + \sum_{i=1}^{n} p_i \ln q_i = \sum_{i=1}^{n} p_i \ln\left(\frac{q_i}{p_i}\right) \le \sum_{i=1}^{n} p_i\left(\frac{q_i}{p_i} - 1\right) = 0,$$

with equality if and only if $p_i = q_i$ for all i. Since $\ln p = \log p \ln 2$, we obtain the desired relation $H(X) \le -\sum_{i=1}^{n} p_i \log q_i$. The inequality $H(X) \le \log n$ is obtained by setting $q_i = 1/n$ for all i.

(b) The numbers $p_X(x)p_Y(y)$ satisfy $\sum_x \sum_y p_X(x)p_Y(y) = 1$, so by part (a), we have

$$\sum_x \sum_y p_{X,Y}(x,y) \log\left(p_{X,Y}(x,y)\right) \ge \sum_x \sum_y p_{X,Y}(x,y) \log\left(p_X(x)p_Y(y)\right),$$

with equality if and only if

$$p_{X,Y}(x,y) = p_X(x)p_Y(y), \qquad \text{for all } x \text{ and } y,$$

which is equivalent to X and Y being independent.

(c) We have

$$I(X;Y) = \sum_x \sum_y p_{X,Y}(x,y) \log p_{X,Y}(x,y) - \sum_x \sum_y p_{X,Y}(x,y) \log\left(p_X(x)p_Y(y)\right),$$

and

$$\sum_x \sum_y p_{X,Y}(x,y) \log p_{X,Y}(x,y) = -H(X,Y)\ ,$$

$$\begin{aligned}
-\sum_x \sum_y p_{X,Y}(x,y) \log\left(p_X(x)p_Y(y)\right) &= -\sum_x \sum_y p_{X,Y}(x,y) \log p_X(x) \\
&\quad -\sum_x \sum_y p_{X,Y}(x,y) \log p_Y(y) \\
&= -\sum_x p_X(x) \log p_X(x) - \sum_y p_Y(y) \log p_Y(y) \\
&= H(X) + H(Y)\ .
\end{aligned}$$

Combining the above three relations, we obtain $I(X;Y) = H(X) + H(Y) - H(X,Y)$.

(d) From the calculation in part (c), we have

$$\begin{aligned}
I(X;Y) &= \sum_x \sum_y p_{X,Y}(x,y) \log p_{X,Y}(x,y) - \sum_x p_X(x) \log p_X(x) - \sum_x \sum_y p_{X,Y}(x,y) \log p_Y(y) \\
&= H(X) + \sum_x \sum_y p_{X,Y}(x,y) \log\left(\frac{p_{X,Y}(x,y)}{p_Y(y)}\right) \\
&= H(X) + \sum_x \sum_y p_Y(y) p_{X|Y}(x \mid y) \log p_{X|Y}(x \mid y) \\
&= H(X) - H(X \mid Y)\ .
\end{aligned}$$