Handout 17 | Information Theory and Coding
Solutions to Midterm | December 16, 2005

Problem 1.

(a) Since $\hat{p}(x) = \max_{1\le k\le K} p_k(x)$, for any k, $p_k(x) \le \hat{p}(x)$, Also, since $p_k(x) \ge 0$,

$$\hat{p}(x) \le \sum_{k=1}^{K} p_k(x).$$

We thus have

$$1 = \sum_x p_k(x) \le \sum_x \hat{p}(x) \le \sum_k \sum_x p_k(x) = K.$$

(b) Since $l(x) \ge -\log_2 \hat{p}(x) + \log_2 A$,

$$\sum_x 2^{-l(x)} \le \sum_x \hat{p}(x)/A = 1,$$

We see that $l(x)$ satisfies the Kraft's inequality and we conclude that there exists a prefix free code with these lengths.

(c) We know that for any uniquely decodable code $\bar{L}_k = \sum_x p_k(x)l(x) \ge H_k$, so the left hand inequality follows. For the right hand, observe that

$$l(x) < -\log_2 \hat{p}(x) + \log_2 A + 1,$$

and since $\hat{p}(x) \ge p_k(x)$, we have

$$l(x) < -\log_2 p_k(x) + \log_2 A + 1,$$

and thus

$$\bar{L}_k < H_k + \log_2 A + 1.$$

The approach outlined here to universal coding is explored further in A. Orlitsky, N. P. Santhanam, J. Zhang "Universal compression of memoryless sources over unknown alphabets", *IEEE Transactions on Information Theory,* vol. 50, no. 7, pp. 1469–1481, July 2004. The quantity A is known as the "Shtarkov sum", and was introduced by Y. Shtarkov in "Universal sequential coding of single messages," *Problems of Information Transmission,* vol. 23 no. 3, pp. 3–17, 1987.

Problem 2.

(a) Note that if the gamber distributes a unit fortune among all outcomes by betting $p(k)$ on outcome k, he will exactly recover his unit investment at the end of the game: if the outcome is k, he will receive $1/p(k)$ times his bet on this outcome, namely 1. Thus, instead of holding $q(0)$ in reserve he may as well distribute this among all outcomes by betting $p(k)q(0)$ on outcome k. Formally, with $\hat{q}(k) = q(0)p(k) + q(k)$ we satisfy

$$\sum_{k=1}^{K} \hat{q}(k) = 1,$$

and the gamber will have $\hat{q}(k)/p(k) = q(0) + q(k)/p(k)$ amount of money if $X = k$, which is exactly the same as what he would have if he used strategy q.

(b) For a strategy q with $q(0) = 0$, the fortune of the gambler is multiplied by $q(X_i)/p(X_i)$ at round i. Thus $C_n/C_0 = \prod q(X_i)/p(X_i)$, and

$$R_n = \frac{1}{n} \sum_{i=1}^{n} \log \frac{q(X_i)}{p(X_i)}.$$

By the law of large numbers R_n converges to

$$r = E\left[\log \frac{q(X)}{p(X)}\right] = \sum_x p(x) \log \frac{q(x)}{p(x)} = -D(p\|q).$$

(c) If the gamber allocates $q(k|y)$ amount of his fortune on outcome k, his fortune will be multiplied by $q(X_i|Y_i)/p(X_i)$ at round i. Thus,

$$R_n = \frac{1}{n} \sum_{i=1}^{n} \log \frac{q(X_i|Y_i)}{p(X_i)}$$

will converge, by the law of large numbers, to

$$\begin{aligned} r = E\left[\log \frac{q(X|Y)}{p(X)}\right] &= \sum_{x,y} p(x,y) \log \frac{q(x|y)}{p(x)} = \sum_{x,y} p(x,y) \log \frac{q(x|y)p(y|x)}{p(x,y)} \\ &= \sum_{x,y} p(x,y) \log \frac{q(x|y)}{p(x|y)} \frac{p(y|x)}{p(y)} = -\sum_y p(y) D\big(p(\cdot|y) \parallel q(\cdot|y)\big) + I(X;Y) \end{aligned}$$

(d) Since $D\big(p(\cdot|y) \parallel q(\cdot|y)\big) \geq 0$ with equality only if $q(x|y) = p(x|y)$, we see that the best strategy is to use allocate a fraction $p(k|y)$ of the gamblers fortune on outcome k when the side information Y is y. In this case, the growth rate of the fortune is given by $I(X;Y)$.

This connection between gambling/investing and information theory was first noticed by J. L. Kelly in his paper "A new interpretation of information rate" *Bell System Technical Journal,* vol. 35, pp. 917–926, July 1956.

Problem 3.

(a) Since $I(X;Y,Z) = I(X;Y) + I(X;Z|Y)$, and since given Y, Z is determined, we have $I(X;Z|Y) = 0$, and we see that $I(X;Y,Z) = I(X;Y)$. On the other hand, $I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \geq I(X;Z)$. Thus we see that

$$I(X;Y) \geq I(X;Z)$$

with equality if and only if $I(X;Y|Z) = 0$, which is to say that X and Y are independent when conditioned on Z, that is, $p(x,y|z) = p(x|z)p(y|z)$.

(b) Since Z is a function of Y, $p(x|y,z) = p(x|y)$ (for $z = g(y)$), thus $p(xy|z) = p(y|z)p(x|yz) = p(y|z)p(x|y)$, and so to check if $p(x,y|z) = p(x|z)p(y|z)$, we only need to check if $p(x|z) = p(x|y)$ whenever $z = g(y)$.

With $Z = g^*(Y) = \log[p(Y|1)/p(Y|0)]$, we see that given $Z = z$, we know that that y is such that $p_{Y|X}(y|1)/p_{Y|X}(y|0) = 2^z$. Let the set of such y's be S_z. Notice that for such y's, the values of $p(x|y)$ are given by

$$p_{X|Y}(0|y) = \frac{p_X(0)p_{Y|X}(y|0)}{p_Y(y)} = \frac{p_X(0)p_{Y|X}(y|0)}{p_X(1)p_{Y|X}(y|1) + p_X(0)p_{Y|X}(y|0)}$$
$$= \frac{p_X(0)}{p_X(1)2^z + p_X(0)} =: f(0|z)$$

and

$$p_{X|Y}(1|y) = \frac{p_X(1)2^z}{p_X(1)2^z + p_X(0)} =: f(1|z)$$

are thus not depend of the specific y in S_z.

Observe now that

$$p(x|z) = \sum_{y \in S_z} p(x,y|z) = \sum_{y \in S_z} p(x|y,z)p(y|z)$$
$$= \sum_{y \in S_z} p(x|y)p(y|z) = f(x|z) \sum_{y \in S_z} p(y|z) = f(x|z).$$

We thus see that $p(x|z)$ and $p(x|y)$ are given by the same quantity, so we have verified that X and Y are independent conditional on $Z = g^*(Y)$.

(c) Suppose now that $Z = g(Y)$ is such that $I(X;Y) = I(X;Z)$. From the discussion above, we see that $p(x|y) = p(x|z)$, thus given z, we can compute $p_{X|Y}(1|y)$ and $p_{X|Y}(0|y)$, and further compute

$$Z^* = \log \frac{p_{Y|X}(y|1)}{p_{Y|X}(y|0)} = \log \frac{p_X(0)}{p_X(1)} + \log \frac{p_{X|Y}(1|y)}{p_{X|Y}(0|y)}.$$