EE376B
Information Theory
Prof. T. Cover

Handout #13
Tuesday, May 5, 2009
Due Thursday, May 14, 2009

Homework Set #5

1. **Counting**
Let $\mathcal{X} = \{1, 2, \ldots, m\}$. Show that the number of sequences $x^n \in \mathcal{X}^n$ satisfying $\frac{1}{n}\sum_{i=1}^n g(x_i) \geq \alpha$ is approximately equal to 2^{nH^*}, to first order in the exponent, for n sufficiently large, where
$$H^* = \max_{P:\sum P(i)g(i)\geq\alpha} H(P).$$

2. **Counting states**
Suppose an atom is equally likely to be in each of 6 states, $X \in \{s_1, s_2, \ldots, s_6\}$. One observes n atoms $X_1, X_2, \ldots, X_n$ independently drawn according to this uniform distribution. It is observed that the frequency of occurrence of state s_1 is twice the frequency of occurrence of state s_2.
   (a) To first order in the exponent, what is the probability of observing this event?
   (b) Assume n large, find the conditional distribution of the state of the first atom X_1, given this observation.

3. **Sanov**
Let X_i be i.i.d. $\sim N(0, \sigma^2)$.
   (a) Find the exponent in the behavior of $\Pr\left\{\frac{1}{n}\sum_{i=1}^n X_i^2 \geq \alpha^2\right\}$. This can be done from first principles (since the normal distribution is nice) or by using Sanov's theorem.
   (b) What does the data look like if $\frac{1}{n}\sum_{i=1}^n X_i^2 \geq \alpha^2$? That is, what is the P^* that minimizes $D(P||Q)$ subject to the appropriate constraint?

4. **Large deviations**
   Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables drawn according to the geometric distribution
   $$\Pr\{X = k\} = p^{k-1}(1-p), k = 1, 2, \ldots.$$
   Find good estimates (to first order in the exponent) of
   (a) $\Pr\left\{\frac{1}{n}\sum_{i=1}^{n} X_i \geq \alpha\right\}$
   (b) $\Pr\left\{X_1 = k \,\middle|\, \frac{1}{n}\sum_{i=1}^{n} X_i \geq \alpha\right\}$
   (c) Evaluate (a) and (b) for $p = \frac{1}{2}$, $\alpha = 4$.

5. **A relation between $D(P \parallel Q)$ and Chi-square**
   Show that the χ^2 statistic
   $$\chi^2 = \Sigma_x \frac{(P(x) - Q(x))^2}{Q(x)}$$
   is (twice) the first term in the Taylor series expansion of $D(P \parallel Q)$ about Q. Thus $D(P \parallel Q) = \frac{1}{2}\chi^2 + \ldots$
   Hint: Write $\frac{P}{Q} = 1 + \frac{P-Q}{Q}$ and expand the log.

6. **Stein's lemma.**
   In connection with the two hypothesis test
   $$H_1 : f = f_1 \quad \text{vs.} \quad H_2 : f = f_2,$$
   find the relative entropy $D(f_1 \parallel f_2) = \int f_1 \ln \frac{f_1}{f_2}$ if
   (a) $f_i(x) = N(0, \sigma_i^2), i = 1, 2$
   (b) $f_i(x) = \lambda_i e^{-\lambda_i x}, x \geq 0, i = 1, 2$
   (c) $f_1(x)$ is the uniform density over the interval [0,1] and $f_2(x)$ is the uniform density over $[a, a+1]$. Assume $0 < a < 1$.
   (d) f_1 corresponds to a fair coin and f_2 corresponds to a two-headed coin.

7. **Error exponent for universal codes.**
   A universal source code of rate R achieves a probability of error $P_e^{(n)} \doteq e^{-nD(P^* \| Q)}$, where Q is the true distribution and P^* achieves $\min D(P \parallel Q)$ over all P such that $H(P) \geq R$. Find P^* in terms of Q and R.