Statistics G8243
Thursday, February 29, 2009
Handout #9

HW Set # 3

Homework due by Tuesday March 24, at 2:40pm, at the *beginning* of class.

1. Read the notes from the last two lectures.

2. *Strong converse in data compression.* Here you will prove the counterpart to the "error exponents" theorem that was stated in class. You'll show that ANY sequence of codebooks with rate below the entropy, has probability of error that goes to 1 exponentially fast.

   Let Q be a source distribution on A, and suppose that $\{B_n \subseteq A^n \ ; \ n \geq 1\}$ is a sequence of codebooks with asymptotic rate

   $$R_n \triangleq \frac{\log |B_n|}{n} \to R < H(Q) \quad \text{bits/symbol.}$$

   Show that the probability of error $Q^n(B_n^c)$ goes to 1 exponentially fast, namely, that

   $$\limsup_{n\to\infty} \frac{1}{n} \log Q^n(B_n) \leq -D_*(Q,R),$$

   where

   $$D_*(Q,R) \ \triangleq \min_{P:H(P)\leq R} D(P\|Q) > 0.$$

   Hints. a. Read the proof of the "error exponents" theorem in the notes.

   b. Show that for any n-type P, the proportion of sequences in $T(P)$ that belong to B_n is at most $(n+1)^{-m}2^{-n|H(P)-R_n|^+}$ [where $|x|^+ = x$ if $x \geq 0$ and $= 0$ otherwise].

   c. Conclude that

   $$\limsup_{n\to\infty} \frac{1}{n} \log Q^n(B_n) \leq -\min_{\text{all } P} [D(P\|Q) + |H(P) - R|^+].$$

   d. Argue (informally) that the last minimum is achieved when $H(P) \leq R$, so that it is actually equal to $D_*(Q,R)$.

   [If you want to be 100% rigorous, note that it suffices to consider the case $R = 0$. Then, however, we have the identity $\min_P [D(P\|Q) + H(P)] = -\log[\max_x Q(x)] = \min_{P:H(P)=0} D(P\|Q)$.]

3. *Counting.* Let $A = \{1, 2, \ldots, m\}$. Show that the number of sequences $x_1^n \in A^n$ 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, where

   $$H^* = \max_{P:\sum_{i=1}^m P(i)g(i)\geq\alpha} H(P).$$

4. *Large deviations.* Let $X_1, X_2, \ldots$ be IID random variables drawn according to the geometric distribution, $\Pr\{X = k\} = p^{k-1}(1-p), k = 1, 2, \ldots$. Find good estimates for:

   (a) $\Pr\{\frac{1}{n}\sum_{i=1}^n X_i \geq \alpha\}$.

   (b) $\Pr\{X_1 = k | \frac{1}{n}\sum_{i=1}^n X_i \geq \alpha\}$.

   (c) Evaluate (a) and (b) for $p = \frac{1}{2}, \alpha = 4$.

5. *The running problem.* Let $X_1, X_2, \ldots, X_n$ be IID $\sim Q_1(x)$, and $Y_1, Y_2, \ldots, Y_n$ be IID $\sim Q_2(x)$. Let X_1^n and Y_1^n be independent. Find an expression for $\Pr\{\sum_{i=1}^n X_i - \sum_{i=1}^n Y_i \geq nt\}$, good to first order in the exponent. The answer can be left in parametric form.

6. *Bent coins.* Let $\{X_i\}$ be IID $\sim Q$ where Q is the Binomial(m, q) distribution. Show that, as $n \to \infty$,

$$\Pr\left\{X_1 = k \;\middle|\; \frac{1}{n}\sum_{i=1}^n X_i \geq \alpha\right\} \to P^*(k),$$

   where P^* is Binomial(m, λ) for some $\lambda \in [0, 1]$. Identify λ.

7. *Large deviations below the mean.* Let $\{X_n\}$ be IID random variables with distribution Q on a finite alphabet A. State, without proof:

   (a) The corresponding result to Cramèr's theorem for the probabilities:

$$\Pr\left\{\frac{1}{n}\sum_{i=1}^n X_i \leq c\right\}, \qquad \text{when } \min\{A\} < c < E_Q[X].$$

   (b) The definition of the relevant exponent $\Lambda^*(c)$ here.

   (c) The properties of $\Lambda(\lambda)$ and $\Lambda^*(c)$ in this case, corresponding to those proved in the Lemma in the proof of Cramèr's theorem.

8. *Simulating rare events.* Suppose you are a public health official, and you want to get some statistical information about what would happen in the event of a major biological disaster, such as a major epidemic, or in case of a biological weapons attack. How do you simulate events in a very unlikely scenario?

   Here's a mickey-mouse version of this problem: Suppose you roll a fair die 10,000 times, and instead of getting a 6 approximately 1/6 of the time, you get a 6 25% of time or more! Clearly this is a very rare event. Now suppose someone asks you, "If I rolled a die 10,000 times and got a 6 at least 25% of the time, what should I expect the overall average of my 10,000 die rolls to be, approximately?" If you tried to simulate this situation, you'd be dead in the water – it would almost never happen that you get ≥25% 6s in 10,000 simulated rolls. Instead, use the Conditional Limit Theorem to:

   (a) Answer this question analytically.

   (b) Answer this question using simulation.

9. Read *again* the notes from the last two lectures before coming to class!