Homework Set #3

Due 15 October 2009, 6 pm, in INR036

Problem 1 (Convergence in probability)

Let $X_1, X_2, \cdots, X_n, \cdots$ be independent, identically distributed random variables drawn from $\mathcal{X} = \{0, 1, 2, 3, 4, 5\}$ according to the probability distribution $\{8/23, 6/23, 4/23, 2/23, 2/23, 1/23\}$ which is ordered according to $\mathcal{X}$ above.
Define $Y_n = \frac{1}{n} \log p(X_1, X_2, \cdots, X_n)$, and $Z_n = \frac{1}{n} \sum_{i=1}^{n} X_i^2$.

(a) Does Y_n converge in probability? If so, calculate the value Y it converges to.

(b) Does Z_n converge in probability? If so, calculate the value Z it converges to.

(c) Compare Z and $(\mathbb{E}[X])^2$ and explain why that relationship holds for an arbitrary choice of $\mathcal{X}$ and $p(x)$.

Problem 2 (Initial conditions of a Markov chain)

Suppose $\{X_i\}$ is a Markov chain, *i.e.,* $X_0 \leftrightarrow X_1 \leftrightarrow \cdots \leftrightarrow X_n$. In other words

$$p(x_0, \ldots, x_n) = p(x_0)p(x_1|x_0) \ldots p(x_n|x_{n-1}).$$

Show that $H(X_0|X_n) \geq H(X_0|X_{n-1})$. In other words, the initial conditions of the Markov chain becomes more and more difficult to recover as time (or process $\{X_i\}$) unfolds.

Problem 3 (Prediction of future block from past block)

For a stationary stochastic process, show that

$$\lim_{n\to\infty} \frac{1}{2n} I(X_1, X_2 \ldots, X_n; X_{n+1}, \ldots, X_{2n}) = 0.$$

This can be interpreted that asymptotically, the dependence between adjacent n-length blocks of a stationary process grows sub-linearly in n.

Problem 4 (Alternative view of AEP)

Let $X_1, X_2, \ldots$ be independent, identically distributed random variables drawn according to the probability mass function $p(x), x \in \{1, \ldots, m\}$. Thus we have $p(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} p(x_i)$. We know that from the law of large numbers that $-\frac{1}{n} \log p(x_1, \ldots, x_n) \to H(X)$ in probability. Let $q(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} q(x_i)$, where q is another probability mass function on $\{1, \ldots, m\}$.

(a) Evaluate $\lim_{n\to\infty} -\frac{1}{n} \log q(x_1, x_2, \ldots, x_n)$, where $X_1, X_2, \ldots$ are i.i.d. according to $p(x)$.

(b) Now evaluate the limit of the log-likelihood ratio $\lim_{n\to\infty} -\frac{1}{n} \log \frac{q(x_1,x_2,\ldots,x_n)}{p(x_1,x_2,\ldots,x_n)}$, when $X_1, X_2, \ldots$ are i.i.d. according to $p(x)$. Thus we have another interpretation of a familiar information-theoretic quantity in terms of the odds of favoring q when the true distribution is p.