Homework Set #3

Due 16 October 2008, before 14:00, INR 031

Problem 1 (Weak Law of Large Numbers)

In this exercise we will prove the weak law of large numbers using the Markov's inequality. Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables, with $\mathbb{E}[X_i] = \mu < \infty$ and $\mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2 = \sigma^2 < \infty$. Define the new sequence of random variables

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n}, \qquad n = 1, 2, \ldots$$

(a) Compute the mean and variance of random variable S_n.

(b) Find the probability that S_n lies outside of the interval $(\mu - \varepsilon, \mu + \varepsilon)$, that is,

$$\Pr[(S_n - \mu)^2 > \varepsilon^2].$$

How does this probability behaves as n increases?
Hint: Define a new random variable $Y_n = (S_n - \mu)^2$, and use the Markov's inequality, *i.e.,* for any non-negative random variable $X \geq 0$ and a positive number $a > 0$, we have

$$\Pr[X > a] \leq \frac{\mathbb{E}[X]}{a}.$$

Problem 2 (Coin Weighing)

Let we have n coins, out of which one is "fake", *i.e.,* has a different weight than the others, and the goal is to find the fake coin. We have two kinds of devices:

- *balance*: You can put two sets of coins on its arms, and the device compares their weight and determines whether which side is heavier, or they have equal weight.
- *detector*: You can put some number of coins on it, and the device tells you whether all the coins have the correct weight, or there is a fake coin among them.

(a) Assume we know that the fake coin is lighter than the real ones. How many different combinations for the coins are possible? (a combination of coins is a sequence of length n, whose i-th elements can be either R or F to denote that the i-th coin is a real or fake coin.)

(b) By each use of a balance, how much information do we get (by how much we can decrease the size of the set of all possible combinations)? What is the minimum number of times we have to use a balance to find the light coin?

(c) What is the minimum number of times we should use a detector to find the fake coin?

(d) Now, assume that we do not know whether the fake coin is lighter or heavier than the real ones. What is the number of different combinations? Note that here each element of the combination sequence can be R, L, or H, for real, lighter, or heavier coins, respectively.

(e) Repeat part (b) for assumption in (d).

(f) Repeat part (c) for assumption in (d).

Problem 3 (The Guessing Game)

You are playing a guessing game: your friend thinks of an object from the set $\{A, B, \ldots, F\}$ and your job is to guess this object by asking questions that have "Yes" or "No" answers. You know that your friend picks the object according to distribution P with probability $\frac{1}{2}$, otherwise he picks it according to distribution Q. Your optimal strategy is to find the object with the minimum average number of questions asked (minimum average cost). Let $P = (p_A, p_B, \ldots, p_F) = (\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}, 0, 0)$ and $Q = (q_A, q_B, \ldots, q_F) = (\frac{1}{2}, \frac{1}{4}, 0, 0, \frac{1}{8}, \frac{1}{8})$.

(a) What is the optimal guessing strategy and calculate the average number of questions needed to find the object.

(b) Suppose a genie tells you that your friend is choosing the object according to distribution P. Come up with an optimal guessing strategy and calculate the average number of questions.

(c) Suppose that the genie always tells you which distribution your friend is using. How much information is the genie giving you?
Hint: compare the average number of questions with or without the genie.)

(d) If the genie is not free, then using it every time is equivalent to asking the question "Are you using distribution P?", and increases your cost by 1 question. Suppose the genie is optional. Can you still use it (in a smart way) and achieve the same average number of questions as in Part (a)?

Problem 4 (Random Box Size)

An n-dimensional rectangular box with sides $X_1, X_2, \cdots, X_n$ is to be constructed . The volume is $V_n = \prod_{i=1}^{n} X_i$. The edge length l of a n-cube with the same volume as the random box is $l = V_n^{1/n}$. Let $X_1, X_2, \cdots$ be i.i.d. uniform random variables over the unit interval $[0, 1]$. Find $\lim_{n\to\infty} V_n^{1/n}$ and compare to $(\mathbb{E}V_n)^{\frac{1}{n}}$. Clearly, the expected edge length does not capture the idea of the volume of the box. The geometric mean, rather than the arithmetic mean, characterizes the behavior of the products.