E0 232 Probability and Statistics *Due:* Oct 9, 2012

Assignment 4

1. In each of the following cases, find the moment generating function of the random variable X.

   (a) Let X be a random variable with CDF

   $$F_X(x) = \begin{cases} 0 & \text{if } x < -3 \\ 0.3 & \text{if } -3 \leq x < 0 \\ 0.5 & \text{if } 0 \leq x < 5 \\ 1 & \text{otherwise.} \end{cases}$$

   (b) Let X be a random variable with CDF

   $$F_X(x) = \begin{cases} 0 & \text{if } x < 0 \\ x^2 & \text{if } 0 \leq x < 1 \\ 1 & \text{otherwise.} \end{cases}$$

   (c) You have a fair coin that comes up heads with probability $\frac{1}{2}$. Your friend has a crooked coin that comes up heads with probability $\frac{3}{4}$. You toss your coin 100 times, while your friend tosses his coin 75 times; all tosses are independent of each other. Let X be the total number of heads obtained.

2. A shipping company called Probabilistic Logistics Inc. packs items sent by clients into boxes, and boxes into crates. The weight of an item (in kg), W, is a continuous random variable distributed uniformly between 1 and 10, independent of everything else. The number of items packed into a box, N, is a $\text{Bin}(20, \frac{1}{2})$ random variable, and is independent of the weights of the items packed or the number of items in other boxes. The number of boxes packed into a crate, K, is 12, 16, 20, or 24, each with equal probability, independent of the contents of the boxes or the number of boxes in other crates.

   (a) Find the MGF, expected value, and variance of the total weight of items in a box.

   (b) Find the MGF, expected value, and variance of the number of items in a crate.

   (c) Find the MGF, expected value, and variance of the total weight of items in a crate.

3. There is a pizzeria outside IISc that is especially popular with students. On a given day, the number of customers at the pizzeria is a Poisson random variable with mean 200. There are 16 different types of pizzas on the menu. Each person who arrives selects one of these 16 pizza types uniformly at random, independently of other customers.

   (a) Find the PMF of the number of customers who order a pizza of type 1.

   (b) Find the expected number of customers who order a pizza of type 1.

   (c) Find the expected number of different types of pizzas that the chef has to make.

4. Suppose you roll a fair 6-sided die repeatedly, each roll being independent of the others. Let X_n denote the outcome of the n-th roll. For each of the following sequences of random variables, investigate whether the sequence converges in probability to some number. If so, give a proof of convergence; if not, explain why not.

   (a) $Y_n = \max(X_1, \ldots, X_n)$

   (b) $Z_n = \min(X_1, \ldots, X_n)$

   (c) $U_n = \frac{X_1 + \ldots + X_n}{n}$

   (d) $V_n = X_{n+1} - X_n$

5. Consider a random variable $X \sim \text{Bin}(n, p)$. For each of the following values of n and p, compute (i) the exact probability $P(|X - np| < 10)$ using the PMF of X; (ii) a lower bound on the probability in (i) using Hoeffding's inequality; (iii) an approximation to the probability in (i) using the Central Limit Theorem; and (iv) for parts (c)-(e), an approximation to the probability in (i) using the PMF of an appropriate Poisson random variable. For (i) and (iv), you may need to write a small computer program.

   (a) $n = 1000$, $p = 0.5$.

   (b) $n = 1000$, $p = 0.1$.

   (c) $n = 1000$, $p = 0.01$.

   (d) $n = 1000$, $p = 0.001$.

   (e) $n = 1000000$, $p = 0.001$.

6. You would like to estimate the average height of adults in a population. You are told that the height of every adult in the population lies between 1 and 2 meters. You decide to randomly sample heights of n adults; each time, you select an adult uniformly at random from the population, independently of other selections, and record his/her height (in meters). Let H_i be the height recorded for the i-th adult selected. You report $\bar{H}_n = \frac{1}{n}\sum_{i=1}^{n} H_i$ as the estimated height. The true (unknown) average height (in meters) is h.

   (a) Based on what you have learned in class, find the smallest value of n that will guarantee that with probability at least 0.99, your estimate is within 5 cm of h, i.e. that will guarantee that

$$P\big(|\bar{H}_n - h| < 0.05\big) \geq 0.99\,.$$

   (b) Suppose you are also told that the variance of heights in the actual population is 0.2 meters. For $n = 1000$, find a good approximation to the probability

$$P\big(|\bar{H}_n - h| < 0.05\big)\,.$$

7. Show that the Chebyshev inequality is tight, i.e. for any values of $\mu \in \mathbb{R}$, $\sigma > 0$, and $\epsilon \geq \sigma$, construct a random variable X with mean μ and variance σ^2 such that

$$P\big((X - \mu)^2 \geq \epsilon^2\big) = \frac{\sigma^2}{\epsilon^2}\,.$$

   Hint: Consider constructing a random variable that takes only three values.

8. Let X_n denote the number of independent tosses of a coin of bias p that are needed to see n consecutive heads (i.e. the number of tosses until n consecutive heads are seen for the first time). Find $E[X_n]$.
   Hint: Consider $E[X_n|X_{n-1}]$.