CS 286a | **HW 1: Probability Refresher** | **Guru: Adam Wierman**

Assigned: 9/29/08 | Due: 10/6/08 10:30am

I encourage you to discuss these problems with others, but you need to write up the actual solutions alone. At the top of your homework sheet, please list all the people with whom you discussed. Crediting help from other classmates will not take away any credit from you. Also, please limit your use of the web as it is difficult to develop good problems, and so enough search will probably lead you to find the solutions online.

1 Alternative definition of expectation [25 points]

In this problem, we will investigate alternative ways to calculate the expectation and higher moments of a random variable. This dual view of expectation will be important later in the course.

(a) Consider a non-negative, discrete, integer valued random variable, X. Traditionally, the expectation is defined as

$$E[X] = \sum_{n=0}^{\infty} nPr(X = n).$$

Prove that we can also calculate $E[X]$ using

$$E[X] = \sum_{n=0}^{\infty} Pr(X > n).$$

(b) For a continuous, non-negative random variable having p.d.f. $f(y)$ the expectation is typically defined as

$$E[Y] = \int_0^{\infty} yf(y)dy.$$

Prove that we can also calculate $E[Y]$ using

$$E[Y] = \int_0^{\infty} \bar{F}(y)dy,$$

where $\bar{F}(y) = Pr(Y \geq y)$.

(c) The i-th moment of Y is typically defined as

$$E[Y^i] = \int_0^{\infty} y^i f(y)dy.$$

Extend the analysis in the previous two parts to provide an equation for the i-th moment of Y in terms of $\bar{F}(y)$.

(d) Do (b) and (c) hold for general random variables (ones that are not non-negative), why or why not?

2 Uniform sampling of a stream [10 points]

Suppose you are observing a stream of data and you want to maintain a sample of one item such that, at all times, the sample is uniformly distributed over all items that have passed so far. You need to do this without knowing the total number of items you'll see and without storing all the items you see. Can this be done?

Amazingly, it can, and with a very simple algorithm: When observing the k-th item, store it as the current sample with probability $1/k$. Prove that this algorithm is correct.

3 Poisson approximation [20 points]

The Poisson distribution will (hopefully) become one of your favorite distributions. We'll use it a lot, so this problem will provide a refresher of some of it's important properties. Recall that a Poisson distribution with parameter λ is a discrete distribution with p.m.f.

$$Pr(X = n) = \frac{e^{-\lambda}\lambda^n}{n!}$$

(a) Calculate the mean of a $Poisson(\lambda)$ random variable X.

(b) Determine the distribution of $X_1 + X_2$ where $X_1 \sim Poisson(\lambda_1)$ and $X_2 \sim Poisson(\lambda_2)$.

(c) Prove that a $Binomial(n, p)$ is well-approximated by a $Poisson(\lambda)$ when $\lambda = np$ is fixed while $n \to \infty$ and $p \to 0$.

This approximation is a fundamental technique for studying balls-in-bins problems. Consider throwing m balls into n bins, where each ball is equally likely to fall into any bin. Using this result, we can approximate the the distribution of the number of balls in a given by $Poisson(m/n)$. It turns out that you can often also approximate the joint distribution of the number of balls in all the bins by independent Poisson random variables, though proving this requires some work! There is a nice book on this topic titled "Poisson Approximation" by Barbour, Holst, and Janson.

4 The weak law of large numbers [15 points]

Let $X_1, X_2, \ldots$ be i.i.d. with finite mean $E[X]$ and finite variance σ_X^2. Let $S_n = \sum_{i=1}^n X_i$. The goal of this question is to prove the weak law of large numbers, i.e., to prove that

$$\lim_{n\to\infty} Pr\left(\left|\frac{S_n}{n} - E[X]\right| \geq \epsilon\right) = 0.$$

(a) Prove Markov's inequality, which says that if X is non-negative with finite mean then for all t,

$$Pr(X > t) \leq \frac{E[X]}{t}.$$

Provide an example illustrating that Markov's Inequality is tight.

(b) Prove Chebyshev's inequality, which says that if X has finite mean and variance then

$$Pr(|X - E[X]| \geq t) \leq \frac{\sigma_X^2}{t^2}.$$

(c) Use Chebyshev's inequality to prove the weak law of large numbers.

5 Slowdown vs. Response Time [10 points]

Consider a single server system that serves arriving jobs according to FCFS. The average arrival rate is 1/2 job/sec and the job sizes (service requirements) are chosen $i.i.d.$ such that with probability 3/4 the size is 1 and with probability 1/4 the size is 2. You have measured the mean response time of the system and found it to be $29/12$.

(a) Based on this information, can you compute the mean slowdown? The slowdown of a job is defined as a job's response time divided by its size.

(b) If the service order were Shortest Job First (SJF) would you be able to compute the mean slowdown?

6 Simple, Greedy, and Optimal [20 points]

(a) Consider a single server to which jobs arrive according to some arbitrary arrival process. Assume that we know the size (service requirement) of each job. Prove that Shortest Remaining Processing Time (SRPT) minimizes the mean response time. SRPT operates by greedily giving the full server to the job in the system with the smallest remaining service time.

Important: There is a fundamental queueing result known as "Little's law," which says that

$$E[N] = \lambda E[T],$$

where N is the number of jobs in the system, T is the response time, and λ is the arrival rate. We will prove this later in the course, but you should use it for this problem. In other words, it is enough to show that SRPT minimizes the mean number of jobs in the system.

(b) The slowdown of a job is defined as a job's response time divided by its size. Mean slowdown is thought by many to be a more important performance metric than mean response time. Why? It seems intuitive that SRPT should minimize mean slowdown as well as mean response time. Prove or disprove this claim.