MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering and Computer Science
6.262 – Discrete Stochastic Processes

Problem Set #1

Issued: February 4, 2009
Due: February 11, 2009

Reading:
For this week: Pages 1-30 of the handout for Chapter 1.

For next week: Finish Chapter 1.

1. a) Exercise 1.12 on p. 48 of handout. Parts b) – e) of this problem (below) all refer to the random variables in Ex. 1.12.

b) Find the probability that X_n is a ***record to date***, i.e., the probability that X_n is greater than or equal to all preceeding random variables. To state it more formally, find:

$$P\left(X_n > X_1, X_n > X_2, \cdots X_n > X_{n-1}\right)$$

for each $n \geq 1$. Hint: Again, no computation is necessary – use symmetry.

c) Is the expected number of records to date in the infinite sequence $X_1, X_2, X_3, \cdots$ finite or infinite?

d) Find the conditional probability that X_{n+1} is a record to date, given that X_n is a record to date. Are these events independent?

e) Is the expected number of adjacent pairs of records to date in the infinite sequence $X_1, X_2, X_3, \cdots$ finite or infinite?

2.a) Exercise 1.16 on page 49 of handout. Assume X and Y are independent.

b) Use the central limit theorem to find

$$\lim_{n\to\infty} P\left(X_n \leq n + \sqrt{n}\right)$$

where X_n is a Poisson random variable with expectation equal to n.

3. Exercise 1.23 (It is worthwhile to think very carefully here.)

4. Problem A: **The Double-or-Quarter Game (Part I)**

Consider the ***"Double or Quarter"*** gambling game, which is based on a sequence of independent tosses of a fair coin. Immediately before the k-th toss you bet b dollars. If the toss comes up heads, you receive back 2b dollars, but if the toss comes up tails you receive back (b/4) dollars. In other words, you receive back $b \cdot \mathbf{X}_k$ dollars after the k-th toss, where the random variables $\mathbf{X}_k$ are independent with

$$P(X_k = 1/4) = P(X_k = 2) = 1/2 .$$

We initially consider two strategies for playing the game. In the ***non-compounding strategy***, you bet $1.00 on each toss. Regardless of the outcome of that toss, you then bet $1.00 on the next toss, until you quit playing.

a) It costs you n dollars to play n games using the non-compounding strategy. Represent your net change in wealth (winnings minus cost) after $n \geq 1$ games, $\Delta \mathbf{W}n$, as a function of $\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n$, and find the mean and standard deviation of $\Delta \mathbf{W}n$.

The second strategy we consider is the ***compounding*** strategy, in which you bet $1.00 on the first toss, then bet the entire return from the first toss (i.e., 25¢ or $2.00) on the second toss, and in general bet the entire return from the (n-1)-st toss on the n-th toss.

b) It costs you 1 dollar to play n games using the compounding strategy. Represent your net change in wealth (winnings minus cost) after $n \geq 1$ games, $\Delta \mathbf{W}n$, as a function of $\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n$, and find the mean and standard deviation of $\Delta \mathbf{W}n$.

c) Which strategy gives you the larger expected increase in wealth E[$\Delta \mathbf{W}n$] for n>1?

d) Show that you win or break even (i.e., $\Delta \mathbf{W}n \geq 0$) following n tosses using the *non-compounding* strategy, if

$$\# \text{ heads} \geq 3n/7,$$

while you win or break even (i.e., $\Delta \mathbf{W}n \geq 0$) following n tosses using the *compounding* strategy if

$$\# \text{ heads} \geq 2n/3.$$

e) Use the central limit theorem to estimate the probability you win or break even after 100 tosses using the *non-compounding* strategy. Also use it to estimate the probability you win or break even after 100 tosses using the *compounding* strategy.

f) Briefly and qualitatively describe the relative advantages and disadvantages of the compounding and non-compounding strategies from your conclusions in parts c) and e).

g) This problem concerns the probability p that you win or break even after 100 tosses using the compounding strategy. Find a rigorous upper bound on p using the Chebyshev inequality. Find the best rigorous upper bound on p that you can by using the exponential (aka Chernoff) bound. Briefly compare these values obtained from these bounds with each other and with the estimate you found in part e) using the central limit theorem.
Briefly explain the advantage of using that approximation and/or these bounds in place of a direct calculation of the exact value of p.

5) Problem B. **Stable Distributions**

Some classes of probability distributions for random variables have the special property that for any two independent random variables **X** and **Y** with probability distributions in that class, the sum

$$\mathbf{Z} = \mathbf{X} + \mathbf{Y}$$

also has a probability distribution in that class. These classes of probability distributions are called ***stable distributions.*** We need to be careful about the set of probability distributions we consider to be a class*,* since, for example, the class of all Gaussian distributions is stable, the class of all Gaussian distributions with mean zero is stable, but the class of all Gaussian distributions with variance 1 is *not* stable.

The definition extends without change to the sum of any finite number of independent random variables.

a) We can also characterize a stable class of distributions by the properties of the moment generating functions of the distributions in that class. What property must the class of MGF's have for the original class of distributions to be stable?

Page 45 of the Chapter 1 handout gives 8 classes of probability distributions, along with their MGF's. A ninth one of interest is the class of Cauchy distributions, in various widths, which have probability densities

$$f_X(x) = \frac{a}{\pi(x^2 + a^2)},\ a > 0,\ -\infty < x < \infty$$

and corresponding moment generating functions (characteristic functions)

$$g_X(i\omega) = e^{-a|\omega|}$$

defined only for $r = i\omega$ pure imaginary. (The means of the Cauchy distributions are undefined and the variances are infinite for every a > 0.)

b) Use the moment generating functions to decide whether each of these 9 classes of distributions is stable or not.

(To be completely explicit, the exponential class and Poisson class on p. 45 are each defined as having all values of $\lambda > 0$. The Erlang class is defined as having all values of $\lambda > 0,\ n \geq 1$. The Gaussian class has all values of $a \in R,\ \sigma > 0$. The uniform class has all values of a>0. The binary class has all values of q $\text{with } 0 \leq q \leq 1.$ The binomial class has all values of q $\text{with } 0 \leq q \leq 1,$ and all values of $n \geq 1$. The geometric class has all values of q with $0 < q \leq 1.$)

c) Is the class of binomial distributions with p = 0.6 (and $n \geq 1$ arbitrary) a stable class of distributions?

d) One apparently plausible reading of the central limit theorem suggests that the class of all Gaussian distributions (or all Gaussian distributions with zero mean) is the only nontrivial stable class, since an appropriately weighted sum of large numbers of iid random variables must have an approximately Gaussian distribution function. Your answer to part b) shows that this reading is mistaken. Briefly explain how each of your positive answers to part b) can be consistent with the central limit theorem.