IEOR 3106: Introduction to Operations Research: Stochastic Models
SOLUTIONS to Part I of First Midterm Exam, October 4, 2007

Justify your answers; show your work.

There are three problems below, each worth 20 points. Each problem has parts. Along with these solutions, the grading scheme is explained.

1. A Two-Stage Coin-Flipping Experiment. (20 points)

In this problem we consider flips (tosses) of a fair coin, assuming that the two possible outcomes - heads and tails - are equally likely, and that successive flips are independent. In the first stage, you flip the fair coin twice; let \( N \) be the number of heads in these two coin flips. In the second stage, you flip the coin \( N \) times; i.e., given that \( N = n \), you flip the coin \( n \) more times. Let \( X \) be the number of heads in these \( N \) coin flips.

Determine the following quantities:

- [5] (a) \( P(X = 1) \),
- [5] (b) \( P(N = 1|X = 1) \),
- [3] (c) \( E[X] \),
- [2] (d) \( Var(X) \), the variance of \( X \),
- [2] (e) \( Cov(N, X) \), the covariance of \( N \) and \( X \),
- [3] (f) Are \( N \) and \( X \) independent random variables? Explain.

PURPOSE:

The purpose of this problem is to test knowledge of conditional probability, as reviewed in Chapter 1 of the textbook. Conditional probability is needed to do both parts (a) and (b). Conditional probability plays a key role in part (f) as well.

GRADING:

The greatest emphasis is given to parts (a), (b) and (f). Partial credit is given in these parts. Essentially no partial credit is given in parts (d) and (e). A possible error is misinterpreting the random variables. The grading tries to take that into account, by not penalizing too much if everything else is correct after a misinterpretation error.

ANSWERS:

It helps to draw figures. A good figure here is a probability tree. The tree for the two stages appears here in the figure on the top of the next page: (We do not show the results of individual coin flips.)

Then we have the following answers:

By looking at the tree, and adding over all the paths for which \( X = 1 \) (multiplying the two probabilities on each path), we get

- (a) \( P(X = 1) = (1/2)(1/2) + (1/4)(1/2) = (1/4) + (1/8) = 3/8. \)
- (b) \( P(N = 1|X = 1) = \frac{P(N=1,X=1)}{P(X=1)} = \frac{1/4}{3/8} = \frac{2}{3}. \)
(c) From the tree, we see that \( P(X = 0) = \frac{9}{16} \), \( P(X = 1) = \frac{6}{16} \) and \( P(X = 2) = \frac{1}{16} \). Hence, \( E[X] = 1\left(\frac{6}{16}\right) + 2\left(\frac{1}{16}\right) = \frac{8}{16} = \frac{1}{2} \).

(d) \( \text{Var}(X) = E[X^2] - (E[X])^2 \). Then, by above, \( P(X^2 = 0) = \frac{9}{16} \), \( P(X^2 = 1) = \frac{6}{16} \) and \( P(X^2 = 4) = \frac{1}{16} \), so that \( E[X^2] = 1\left(\frac{6}{16}\right) + 4\left(\frac{1}{16}\right) = \frac{10}{16} = \frac{5}{8} \). Hence, \( \text{Var}(X) = \left(\frac{5}{8}\right) - \left(\frac{1}{2}\right)^2 = \frac{3}{8} \).

(e) From the tree, we also can see the joint distribution of \( N \) and \( X \). We use \( \text{Cov}(N, X) = E[NX] - E[N]E[X] \); see page 53 of the textbook. Clearly, \( E[N] = 1 \). We have calculated that \( E[X] = \frac{1}{2} \). Hence, \( E[N]E[X] = \frac{1}{2} \). So it suffices to compute \( E[NX] \) from the joint distribution of \( (N, X) \). From the tree, we see that \( P(NX = 0) = \frac{9}{16} \), \( P(NX = 1) = \frac{4}{16} \), \( P(NX = 2) = \frac{2}{16} \) and \( P(NX = 4) = \frac{1}{16} \). Hence, \( E[NX] = 1\left(\frac{4}{16}\right) + 2\left(\frac{2}{16}\right) + 4\left(\frac{1}{16}\right) = \frac{12}{16} = \frac{3}{4} \). Finally, \( \text{Cov}(N, X) = (\frac{3}{4}) - (\frac{1}{2}) = \frac{1}{4} \).

(f) Are \( N \) and \( X \) independent random variables? Explain.

No, the random variables \( N \) and \( X \) are dependent. That is implied by having the covariance not be 0. See pages 52-53 in the textbook.

2. Repeated Plays of a Game. (20 points)

Consider a game with three possible outcomes: You win 8 dollars with probability \( \frac{11}{48} \); you win (and lose) nothing with probability \( \frac{3}{48} \); and you lose 4 dollars with probability \( \frac{34}{48} \). Suppose that you play this same game 100 times, with each successive play being independent of the previous ones.

[5] (a) What is the expected value of the total amount you win (or lose) in the end?
(b) What is the (approximate) probability that you come out ahead (have positive winnings) in the end?

PURPOSE:

The purpose of this problem is to test knowledge of the normal approximation for a sum of IID random variables, as justified by the central limit theorem. This is covered in Section 2.7 and was heavily emphasized in the lectures. Indeed, in class, I even promised that a CLT problem would appear on the exam. We apply that in part (b).

GRADING:

The greatest emphasis is given to part (b). Part (a) is there only to serve as an easy warmup. Roughly, the 15 points in part (b) are allocated as follows: (i) 7 points for recognizing that we should be doing a normal approximation, (ii) 3 points for explicitly mentioning that the justification follows from the central limit theorem. (At the top of the exam it said “justify your answers.”) Finally, (iii) 5 points are given for getting the correct answer.

ANSWERS: We have a sum of IID random variables. This is mostly about the central limit theorem (CLT), as discussed on pages 79-83 of the textbook. That needs to be said, as part of justifying your work. Let $X$ be the amount won on one play of the game. Let $X_k$ be the amount won on play $k$, $1 \leq k \leq 100$, with each $X_k$ distributed as $X$. Let the total winnings be

$$S_{100} = X_1 + \cdots + X_{100}.$$  

(a) First observe that $E[X] = 8(11/48) + 0(3/48) - 4(34/48) = (88 - 136)/48 = -48/48 = -1$. We use the fact that the expectation of a sum is the sum of the expectations. Hence, the expected total amount won is $E[S_{100}] = -100$ dollars. That is, we expect to lose 100 dollars.

(b) Now we apply the central limit theorem. To find the variance, it is perhaps easiest to compute the second moment and then subtract the square of the mean. We first find that $E[X^2] = 26$. Then we find that the $Var(X) = 25$, so that $Var(S_{100}) = 2500$, and the associated standard deviation is 50. (We have worked to make the arithmetic clean!)

$$P(S_{100} > 0) = P \left( \frac{S_{100} - E[S_{100}]}{SD(S_{100})} > \frac{0 - E[S_{100}]}{SD(S_{100})} \right)$$

$$\approx P \left( N(0, 1) > \frac{0 - E[S_{100}]}{SD(S_{100})} \right)$$

$$= P(N(0, 1) > 2) \approx 1.000 - 0.9772 = 0.0228 \approx 0.023,$$  

from the table on page 81.

3. Random Meeting. (20 points)

Two students - Albert Lee and Blake Rego - have planned to meet at the Mudd Building one day between 9am and 10 am. Suppose that they arrive independently in this one-hour
interval, each at a random time that is uniformly distributed over the interval. We will want to say something about the time the first arrival has to wait for the second arrival. Toward that end, let $A$ be the random time that Albert arrives and let $B$ be the random time that Blake arrives. For simplicity assume that these are uniform in the interval $[0, 1]$. Then the waiting time is $W \equiv |A - B|$.

Determine the following quantities:

4] (a) $P(1/3 \leq A \leq 2/3, 1/4 \leq B \leq 3/4)$,

4] (b) $P(A - B > t)$ for $t > 0$,

3] (c) $P(W > t)$ for $t \geq 0$,

3] (d) $f_W(t)$, the probability density function of $W$,

3] (e) $E[W]$,

3] (f) $Var(W)$, variance of $W$.

PURPOSE:

The purpose of this problem is to test knowledge of basic probability manipulations, in particular, the way to compute the distribution of the sum (or difference) of two independent random variables. That involves a convolution of the two distributions, as described on page 58 of the textbook. This was also to illustrate again that simple pictures help a lot.

GRADING:

The primary goal was to have you do part (c). Part (a) is a warmup. And part (b) is a key first step to do (c). This problem proved to be especially difficult. Accordingly, the grading gave credit for a good start. In particular, 4 points were given for (a). Then 4 points were given if the work headed toward computing the convolution. The detailed scoring is above, but we also awarded 2 points for a good effort overall beyond the initial 4 points for trying to do a convolution.

ANSWERS: Again, a good picture makes everything clear. Here we observe that the pair $(A, B)$ of arrival times is uniformly distributed in the unit square. Let $(a, b)$ be possible values of $(A, B)$. The set of points for which $a - b > t$ is the same as the set of points for which $a > b + t$. That set is the shaded triangle in the figure below.

(a) This is a simple warm-up question, included to help you see that the joint distribution is uniform on the square.

$$P(1/3 \leq A \leq 2/3, 1/4 \leq B \leq 3/4) = (1/3) \times (1/2) = 1/6.$$ 

(b) $P(A - B > t) = (1 - t)^2/2$ for $t > 0$, as can be seen from the picture: we just need to find the area of the triangle. We could proceed in a more systematic, analytic way, doing the two-dimensional integral. The general procedure is discussed on page 58 of the book. This
involves convolution. We can write
\[ P(A - B > t) = P(A > B + t) = \int_{0}^{1-t} f_B(y) \left( \int_{y+t}^{1} f_A(x) \, dx \right) \, dy \]
\[ = \int_{0}^{1-t} \left( \int_{y+t}^{1} dx \right) \, dy \]
\[ = \int_{0}^{1-t} (1 - t - y) \, dy \]
\[ = \int_{0}^{1-t} y \, dy \]
\[ = \frac{(1 - t)^2}{2}. \]

You really should be able to do this. This is finding the distribution of the sum of two random variables, where here one of the random variables is $A$, while the other is $-B$. This is a basic topic in an introductory probability course.

(c) It should be intuitively obvious that we simply multiply the answer in part (b) by 2. More formally, we can derive the value for $P(A - B < -t) = P(B - A > t)$ for $t > 0$, by multiplying by $-1$. But then, by simply changing the labels $A$ and $B$, we see that this must agree with what we did in part (b). Here too we have $P(A - B < -t) = (1 - t)^2/2$. This is just the other triangle at the top left, congruent to the one displayed on the lower right. Hence, we have
\[ P(W > t) = (1 - t)^2 \quad \text{for} \quad t \geq 0. \]

(d) $f_W(t) = 2(1 - t)$, by differentiating in part (c). (You need to multiply by minus 1 too.) Check that it integrates to 1.
(e) $E[W] = \int_0^1 x f_W(x) \, dx = \int_0^1 2x(1 - x) \, dx = 1/3$. Again, check that your answer makes sense. It must be a number between 0 and 1.

(f) $Var(W) = E[W^2] - (E[W])^2 = (1/6) - (1/3)^2 = 1/18$, because

$$E[W^2] = \int_0^1 2x^2(1 - x) \, dx = 1/6.$$ 

Honor Code: Students are expected to behave honorably, following the accepted code of academic honesty. After completing your exam, please affirm that you have done so by writing "I have neither given nor received improper help on this examination," on your examination booklet and sign your name. You may keep the exam itself. Solutions will eventually be posted on line.