Assignment #9 (corrected and revised) — Due Wednesday, November 5, 2008, by 5:00 P.M.

Turn in homework in lecture, discussion, or your TA’s mailbox. Indicate the discussion section in which you expect to attend to pick up this assignment on the assignment.

311: Monday 1:20–2:10  
312: Monday 12:05–12:55

Some of the following problems require the program R. The motivation for using R is described briefly in your course syllabus. There are handouts which explain how to install R and how to use R for these problems on the course web page.

1. Use the fact from the previous homework assignment that if X is a discrete random variable on the nonnegative integers, then \( E(X) = \sum_{k=0}^{\infty} P(X > k) \), to find the mean of \( X \sim \text{Geometric}(\theta) \).

2. Another technique for finding expected values is to use calculus and to exchange the order of summation and differentiation. (The derivative of a sum is the sum of the derivatives.) Specifically,

\[
\sum_{k=0}^{\infty} \frac{d}{d\theta} (g_k(\theta)) = \frac{d}{d\theta} \left( \sum_{k=0}^{\infty} g_k(\theta) \right)
\]

(which, to be true for infinite sums, requires some conditions which are met in our example). Suppose that \( X \sim \text{Geometric}(\theta) \) and we desire to compute \( E(X + 1) \). We can begin with

\[
E(X + 1) = \sum_{k=0}^{\infty} (k+1)(1-\theta)^k \theta = \theta \sum_{k=0}^{\infty} \frac{d}{d\theta} \left( -(1-\theta)^{k+1} \right)
\]

Exchange the order of summation and differentiation, simplify the sum, take the derivative, and simplify. Note that \( E(X) = E(X + 1) - 1 \) to find, yet again, the mean of the geometric distribution.

3. The technique of the previous problem can also be used for second derivatives. If \( X \sim \text{Geometric}(\theta) \), notice that

\[
\frac{d^2}{d\theta^2} \left( (1-\theta)^{k+2} \right) = (k+2)(k+1)(1-\theta)^k
\]

and that

\[
E((X + 2)(X + 1)) = \sum_{k=0}^{\infty} (k+2)(k+1)(1-\theta)^k \theta.
\]

Show that this expectation is \( 2/\theta^2 \).

4. Use the result of the previous problem to find \( E(X^2) \) and \( \text{Var}(X) \) if \( X \sim \text{Geometric}(\theta) \).

5. Suppose that \( X \) is an absolutely continuous random variable with density \( f \) and that \( X > 0 \) (in other words, \( f(x) = 0 \) for \( x < 0 \)). Show that

\[
E(X) = \int_0^\infty P(X > x) \, dx.
\]

(Hint: \( P(X > x) = \int_x^\infty f(t) \, dt \).)

6. Let \( X \) and \( Y \) have joint density

\[
f_{X,Y}(x, y) = \begin{cases} 
4x^2y + 2y^5 & \text{if } 0 < x < 1, \ 0 < y < 1 \\
0 & \text{otherwise}
\end{cases}
\]

Find the following.
(a) $E(X)$
(b) $E(Y)$
(c) $E(2X - 4Y)$
(d) $E(X^2)$
(e) $E(Y^2)$
(f) $E(XY)$

7. **(Computer Problem)**. The program R has built in functions for generating random variables and for computing sample means, variances, and standard deviations. This problem asks you to simulate samples of random variables to compare sample statistics with theoretically determined values.

(a) The command `x = rnorm(100000,100,5)` generates a random sample of 100,000 random variables from a normal distribution with mean $\mu = 100$ and $\sigma = 5$ and saves the sample in an object `x`. The command `mean(x)` calculates the sample mean of the values saved in object `x`. If you do not want to save the sample, you can do this in one step with `mean(rnorm(100000,100,5))`. Find and report the means from five separate samples. How close are these means to 100? *(Hint: In R you can use the arrow keys to recall and edit previous commands.)*

(b) The command `x = rgeom(100000,1/3)` generates 100,000 random geometric random variables with $\theta = 1/3$. Find the mean, variance, and standard deviation of this sample using the commands `mean`, `var`, and `sd`, respectively. Repeat for a total of five samples. Compare the simulated values to the theoretical values.

**Work to do, but not turn in.**

- Read Chapter 3 sections 3.4–3.6.