Assignment #9 — Due Wednesday, November 4, 2009, by 5:00 P.M.

Turn in homework in lecture, discussion, or your TA's mailbox (just inside the main entrance of MSC, 1300 University Avenue) Please circle the discussion section you expect to attend to pick up this assignment.

311: Monday 1:20–2:10 **312:** Monday 12:05–12:55 **313:** Tuesday 8:25–9:15

1. Suppose that $X \sim \text{Geometric}(0.4)$ on $0, 1, \ldots$ and $Y \sim \text{Geometric}(0.7)$ on $0, 1, \ldots$. Let $Z = X - Y$ and let $W = 3X + Y$.

   (a) Find $\mathsf{E}(Z)$ and $\mathsf{E}(W)$.

   (b) Note that the problem does not specify that X and Y are independent. Does this matter? Briefly explain.

   (c) Assume that X and Y are independent. Find $\mathsf{E}(XY)$. Does the independence assumption matter here? Briefly explain.

2. Suppose that X is a discrete random variable on $k = 0, 1, 2, \ldots$ with $\mathsf{P}(X = k) = p_k$. Prove that $\sum_{k=0}^{\infty} \mathsf{P}(X > k) = \mathsf{E}(X)$. (*Hint: Write* $\mathsf{P}(X > k)$ *as a sum and then (carefully) change the order of summation.*)

3. Use the result from the previous problem to verify $\mathsf{E}(X) = (1 - \theta)/\theta$ if $X \sim \text{Geometric}(\theta)$.

4. 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,

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

   (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 $\mathsf{E}(X + 1)$. We can begin with

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

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

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

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

   and that

$$\mathsf{E}((X + 2)(X + 1)) = \sum_{k=0}^{\infty} (k + 2)(k + 1)(1 - \theta)^k \theta \, .$$

   Show that this expectation is $2/\theta^2$.

6. Ten slips of paper with the digits from 1 to 10 are mixed and then placed in a random order. Assume that all permutations are equally likely. Let X be the number of times the sequence needs to be read from left to right to observe all of the numbers in order. For example, the sequence $1, 2, 3, 4, 5, 6, 7, 8, 9, 10$ only needs to be read once, while the sequence $2, 10, 9, 5, 8, 6, 3, 4, 1, 7$ needs to be read six times. $(1 - - - 2, 3, 4 - - - 5, 6, 7 - - - 8 - - - 9 - - - 10)$. Find $\mathsf{E}(X)$. (*Hint: Write* X *as a sum of simpler random variables.*)

7. Shuffle a standard deck of cards. Find the following.
   (a) The expected number of jacks in the first five cards.
   (b) The expected number of jacks before the first queen.
   (c) The expected number of jacks before the second queen.
   (d) The expected number of jacks before the $Q\spadesuit$.

8. **(Computer Problem)**. The program R has built in functions for computing probabilities, densities, and quantiles (the inverse cdf function) for many distributions. This problem asks you to solve several normal probability problems both using R and the normal table. For each problem, write down the R command and the numerical answer to four decimal places.

   Let $Y \sim \mathrm{N}(250, 25^2)$.

   (a) Use the normal table and `pnorm` to find $\mathsf{P}(Y > 240)$;
   (b) Use the normal table and `pnorm` to find $\mathsf{P}(Y > 212)$;
   (c) Use the normal table and `pnorm` to find $\mathsf{P}(210 < Y < 287)$;
   (d) Use the normal table and `qnorm` to find the constant y such that $\mathsf{P}(Y > y) = 0.20$;
   (e) Use the normal table and `pnorm` to find $\mathsf{P}(|Y - 250| > 30)$.
   (f) Use the normal table and `qnorm` to find the constant y such that the constant y (accurate to two decimal places) such that $\mathsf{P}(|Y - 250| < y) = 0.75$.

9. **(Computer Problem)**. Let $X \sim \text{Binomial}(1200, 0.38)$. For each problem, write an expression using the binomial probability function for the answer, but evaluate it numerically using R.
   (a) $\mathsf{P}(X > 430)$.
   (b) $\mathsf{P}(430 \le X \le 470)$.
   (c) $\mathsf{P}(X < \mathsf{E}(X))$.
   (d) $\mathsf{P}(X = \mathsf{E}(X))$.
   (e) Find the an integer a such that $\mathsf{P}(X \ge a) \ge 0.05$ and $\mathsf{P}(X \le a) \ge 0.95$ (use `qbinom()`).

Work to do, but not turn in.

- See the handouts and install R on your computer.
- Read Chapter 3.