STAT 870

Problems: Assignment 1

1. A Binomial(n, p) random variable X can be written as $X = X_1 + \cdots + X_n$ where X_i is 1 if the ith trial is a success and 0 otherwise. Use this fact to find the mean, variance and moment generating function of X. You may use without proof standard properties of independent random variables (such as formulas for expected products and variances).

2. Suppose $X_1, X_2, \ldots$ are iid Uniform[0,1] random variables.

   (a) Fix an integer n and let M be the number of $\{X_j, 2 \le j \le n+1\}$ which are smaller than X_1. Find the mean and variance of M.

   (b) Now suppose that n is actually a random variable N (independent of all the X_i) with a Poisson distribution with mean λ. Find the mean and variance of M in this case. You may get moments of Poisson random variables from some source (providing you cite the source) rather than deriving them.

   (c) For the situation in the previous part find the distribution of M by computing $P(M = m)$. Your answer may involve integrals; you do not need to do these integrals unless they turn out to be easy.

3. Consider a population of 30 million people of whom 30 thousand have a certain condition. A test is available with the following properties. Assuming that a person has the condition the probability that the test detects the condition is 0.9. Assuming that a person does not have the condition the test detects (incorrectly) the condition with probability 0.001. A person is picked at random from the 30 million people and the test is administered.

   (a) What is the chance that the test detects the condition for this randomly selected person?

   (b) Assuming that the condition is detected by the test for this randomly selected person what is the chance that the person has the condition?

   (c) A mandatory testing program is contemplated. If all 30 million are tested about how many positive results should be expected? Of these about how many will not have the condition?

   (d) Write a short paragraph on the implications of these calculations for public policy. Think about screening tests for cancer, whole body MRI scans, corporate drug testing policies or whatever other application area interests you.

4. Suppose X and Y are independent Geometric(p) random variables. In other words

$$\mathrm{P}(X = j \text{ and } Y = k) = \mathrm{P}(X = j)\mathrm{P}(Y = k) = (1-p)^2 p^{j+k}.$$

   (a) Let $U = \min(X, Y)$, $V = \max(X, Y)$ and $W = V - U$. Express the event $U = j$ and $W = k$ in terms of X and Y.

(b) Compute $\mathrm{P}(U = j)$ and $\mathrm{P}(W = k)$ and prove that the event $U = j$ and the event $W = k$ are independent.

(c) A computer is waiting for two flags to be set. Both flags start out not set. For each flag the conditional probability that the flag is set next cycle given it is not yet set is 0.5. The two flags operate independently. How many cycles should you expect to wait before both flags are set including the cycle on which the last flag becomes set?

5. Show that if $Y > 0$ has cdf G then $\mathrm{E}(Y) = \int_0^\infty \{1 - G(y)\}\, dy$. Please try to do general Y but if not you may assume that Y is integer valued or has a density g.

6. Suppose A is a subset of Ω. What events are in the smallest σ-field containing A? [HINT: there aren't very many.]

7. Suppose that A and B are independent events. Let $\mathcal{F}_A$ be the smallest σ-field containing the event A and similarly defined $\mathcal{F}_B$. Show that $\mathcal{F}_A$ and $\mathcal{F}_B$ are independent. In particular show that A and B are independent if and only if A^c and B are independent.

8. Let X be uniformly distributed on the unit interval. Then X splits the interval into two intervals, namely, $[0, X)$ and $[X, 1]$. Let Y be independent of X and uniform on $[0, 1]$. Think of Y as a time of sampling (between 0 and 1 so that X splits an interval of 1 time unit into two pieces and Y picks a piece) and let Z be the length of the selected one of the two intervals determined by X. Determine the distribution of Z and its mean. What is the mean length of the two intervals $[0, X]$ and $[X, 1]$?

9. Suppose X and Y have joint density $f(x, y)$. Prove from the definition of density *given in class* that the density of X is $g(x) = \int f(x, y)\, dy$.

10. Suppose X is Poisson(θ). After observing X a coin landing Heads with probability p is tossed X times. Let Y be the number of Heads and Z be the number of Tails. Find the joint and marginal distributions of Y and Z.

11. **Warning: This is probably hard. Don't waste too much time on it.** Suppose X and Y are independent $\Gamma(p, 1)$ and $\Gamma(p + 1/2, 1)$ random variables. Show that $Z = 2(XY)^{1/2}$ is a $\Gamma(2p, 1)$ random variable.

12. Suppose n people throw their names in a hat at Christmas. Each person then picks a name to give a present to. What is the expected number of people who draw their own name? What is the variance?

13. Individuals are picked at random, one at a time, from a large population until a person has been picked with a birthday on each day of the year. Compute the mean and variance of the number of picks required. In this problem I want to see you set the problem up and make enough assumptions to get a rough answer.

14. Consider the Monte Hall problem discussed in class. Suppose $P(H = i) = p_i$ for $i = 1, 2, 3$. That is, consider the problem if you model Monte Hall as having probabilities

which might be different from 1/3 of hiding the prize behind specific doors. After Monte hides the prize you get to Choose a door; that is, you control $P(C = j)$. What probabilities should you use and should you switch, or not, when Monte offers you the choice?

Due Date: **May 25**