UNIVERSITY OF MASSACHUSETTS
Department of Mathematics and Statistics
Basic Exam - Probability
Friday, September 2, 2011

Work all problems. 60 points are needed to pass at the Masters Level and 75 to pass at the Ph.D. level.

1. (20 PTS) Suppose that Y is uniformly distributed on the interval (0,1).
   (a) Find the moment generating function for Y.
   (b) If a is a positive constant, derive the moment generating function for $W = aY$. What is the distribution of W? Why?
   (c) If a is a positive constant and b is a fixed constant, derive the moment generating function of $V = aY + b$. What is the distribution of V? Why?

2. (20 PTS) Let X_1 and X_2 be independent variables each having an exponential distribution with mean 1. Let $Y_1 = X_1/X_2$ and $Y_2 = X_2$.
   (a) Without any calculation, give the marginal distribution of Y_2.
   (b) Find the joint probability density function of Y_1 and Y_2.
   (c) Find the conditional distribution of Y_1 given $Y_2 = y_2$.
   (d) Find the marginal probability density function of Y_1.

3. (20 PTS) A blood test is 99 percent effective in detecting a certain disease when the disease is present. However, the test also yields a false-positive result for 2 percent of healthy patients tested, who do not have the disease. Suppose 0.5 percent of the population has the disease.
   (a) Find the conditional probability that a randomly tested individual actually has the disease given that his or her test result is positive.
   (b) Suppose instead that an individual is tested only if he or she has symptoms. Among those with symptoms, 50% are known to have the disease. Find the conditional probability that a person with symptoms actually has the disease given that his or her test result is positive.

4. (25 PTS) (X_1, X_2) is a bivariate random variable and define $\theta = P(X_1 > X_2)$. Define the function $g(X_1, X_2)$ as follows:

$$g(X_1, X_2) = \begin{cases} 1 & \text{if } X_1 > X_2 \\ 0 & \text{otherwise} \end{cases}$$

(a) What is the expected value of $g(X_1, X_2)$, $E(g(X_1, X_2))$?

(b) Let the pairs (X_{1i}, X_{2i}) be independent and identically distributed samples with the same distribution as (X_1, X_2) where $i = 1, \ldots, n$. Define $Q = \sum_{i=1}^{n} g(X_{1i}, X_{2i})$. Find the distribution of Q.

(c) Show Q/n converges in probability to θ as n goes to infinity.

(d) Obtain the asymptotic distribution of $n^{-1/2}(Q - n\theta)$.

5. (15 PTS) Consider a population consisting of 3 units with known sizes, $s_1 = 1, s_2 = 2, s_3 = 3$, from which we will select $n = 2$ units using a procedure known as *probability proportional to size without replacement* (PPSWOR) sampling. The PPSWOR algorithm proceeds as follows:

i. Select the first unit with probability proportional to size.

ii. Select the next unit with probability proportional to size from among the remaining un-sampled units.

iii. Repeat step (ii) until a sample of size n is obtained.

(a) What is the probability that the first unit selected is the unit of size 2 ($s_2 = 2$)?

(b) What is the probability that the n=2 units selected are the units of sizes 1 and 2 (in any order)?

(c) Give numerical expressions for the probabilities of sampling each possible pair of units (you need not simplify these expressions completely).

(d) Give numerical expressions for the probabilities of sampling each unit in the population, that is, for π_1, π_2, and π_3, where π_i is the probability that unit i is selected (again you need not simplify these expressions completely).

(e) This sampling scheme is sometimes used to approximate a *probability proportional to size* (PPS) sample. In a PPS sample, each unit is sampled with probability proportional to its size. That is:

$$\frac{\pi_i}{\pi_j} = \frac{s_i}{s_j}.$$

Simplify the expressions in the previous part (d) as necessary to show that the PPSWOR sampling algorithm does not result in a PPS sample.