UNIVERSITY OF MASSACHUSETTS
Department of Mathematics and Statistics
Basic Exam - Statistics
Wednesday, August 29, 2012

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

1. (15 points) Jane is trapped in a mine containing 3 doors. The first door leads to a tunnel that will take her to safety after 3 hours of travel. The second door leads to a tunnel that will return her to her starting point in the mine after 5 hours of travel. The third door leads to a tunnel that will return her to her starting point in the mine after 7 hours. If we assume that Jane is at all times equally likely to choose any one of the doors, what is the expected length of time until she reaches safety?

2. Suppose we are in a situation where the value of a random variable X is observed and then, based on the observed value, an attempt is made to predict the value of another random variable Y. Let $g(X)$ denote the predictor (i.e., if X is observed to be equal to x, then $g(x)$ is our prediction for the value of Y). We would like to choose g so that $g(X)$ tends to be close to Y. Suppose we decide to choose g to minimize $E((Y - g(X))^2)$. (That defines "best" below.) Assume that the means and variances of X and Y, denoted by $\mu_X = E(X)$, $\mu_Y = E(Y)$, $\sigma_X^2 = Var(X)$ and $\sigma_Y^2 = Var(Y)$, and the correlation of X and Y, denoted by $\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{\sigma_X^2 \sigma_Y^2}}$, are known.

   (a) (10 points) What is the best predictor of Y? (show your calculation in detail)

   (b) (5 points) Consider linear predictors of Y, i.e., $g(X) = a + bX$. In that case, what is the best linear predictor of Y with respect to X? In other words, choose a and b in $g(X) = a + bX$ as functions of the means and variances of X and Y and the correlation of X and Y.

3. The number of defects per yard in a certain fabric, Y, is known to have a Poisson distribution with parameter λ. The probability mass function is:

$$Pr(Y = y|\lambda) \quad = \quad \frac{\exp(-\lambda)\lambda^y}{y!}, \lambda > 0, y = 0, 1, 2, \ldots$$

   Suppose that λ is also an exponential random variable with mean 1 (pdf: $f(\lambda) = \exp(-\lambda), \lambda \geq 0$ and 0 otherwise.)

   (a) (5 points) Write down and solve an integral expression for the unconditional probability mass function for Y.

   (b) (5 points) Without using your answer from part a above, what is the unconditional expectation of Y?

(c) (5 points) Without using your answer from part a above, what is the unconditional variance of Y?

4. Suppose Y is a random variable with pdf $g(y)$, X is a random variable with pdf $f(x)$, and U has a $U(0,1)$ distribution. Further, let $M > 1$ be a constant where $f(x) < Mg(x)$ for all x. Consider the following algorithm.

   (i) Sample y from $g(y)$ and u from a $U(0,1)$.

   (ii) If $u < f(y)/(Mg(y))$ then "accept" y and stop the algorithm. If not, go to (i).

   (a) (5 points) Show that $Pr(U < f(Y)/(Mg(Y))) = E(f(Y)/(Mg(Y)))$ where the expectation is with respect to Y.

   (b) (5 points) Show that the probability that the algorithm accepts y on the first try is $1/M$.

   (c) (5 points) What is the expected number of tries the algorithm will make until it accepts for the first time?

   (d) (5 points) Derive the density of the accepted y.

5. Central Limit Theorem and related topics

   (a) (10 points) State carefully a Central Limit Theorem for a sequence of i.i.d. random variables.

   (b) (10 points) Suppose $X_i, i = 1, \ldots, 100$ are i.i.d. $Poisson(0.0001)$. What is the standard error of the sample mean?

   (c) (5 points) Let $Y = \sum_{i=1}^{100} X_i$. Use the Central Limit Theorem to write an expression to approximate $Pr(Y \geq 1)$. (You do not need a number.)

   (d) (5 points) Is the answer to part c close to zero or close to one? Why?

   (e) (5 points) The $Poisson(\lambda)$ moment generating function is $\exp(\lambda(\exp(t) - 1))$. Use that result to derive an answer part c in another way. Does that support or disagree with your answer to part d?