Math 336.01 **Exam I** **February 27, 2006**

Directions: Work all problems on separate paper. Be neat and show your work.

1. (24 points) Suppose busses in a city are uniquely labeled with ID numbers from the consecutive positive integers 1, 2, ..., N. Jack has read in the newspaper that there are 20 busses in the fleet, but he suspects there are actually more. He is going to use a significance test to decide which is the case, the newspaper report (H_0) or his own hunch (H_1) and he will base the test on observing the next two busses he rides into town. Assume that the two ID numbers he sees, X_1 and X_2, are iid from the discrete, uniform distribution on the set $1, 2, ..., N$. He decides to use the following decision rule: Reject H_0 in favor of H_1 iff the maximum of X_1 and X_2 is 19 or greater.
   (a) Find the probability of a type I error for Jack's test.
   (b) Find the power of the test if there are 25 busses in the fleet.
   (c) Find the probability of a type II error if there are 25 busses in the fleet.
   (d) Without doing any calculations, would the power when there are 30 busses in the fleet be bigger or smaller than the answer to (b)? Explain your answer.
   (e) Is $X_{max} = X_2'$ sufficient for N? Prove your answer.

2. (24 points) Suppose $X_1, X_2, ..., X_n$ are iid from the discrete, geometric pdf

$$f_X(x) = \begin{cases} (1-p)^{x-1}p, & x = 1, 2, ..., \\ 0, & \text{elsewhere.} \end{cases}$$

   Recall that $E(X) = 1/p$.
   (a) Find the maximum likelihood estimator for p.
   (b) Find the method of moments estimator for p and compare it to the MLE.
   (c) Suppose it can be assumed that the number of cars observed in a random passenger car at a particular busy intersection in Los Angeles can be modeled by the geometric pdf given in the problem statement. After observing a sample of 1011 cars (which we'll assume are iid) we obtain the data set below. Estimate p using the MLE. (You may assume the 6+ category are all 6's.)

Number of occupants	Frequency
1	678
2	227
3	56
4	28
5	8
6	14
	1011

3. (16 points) Suppose an auditor is auditing the accounts in a very large tax return preparation company that has on file millions of accounts. One parameter of interest is the proportion, p, of the accounts that have been audited by the IRS in the past year.

   (a) Suppose a random sample of 100 accounts reveals 12 audits. Find a 98% confidence interval for p.

   (b) Suppose the auditor would like to estimate p to within $\pm.02$ with 98% confidence. How large a sample size at a minimum would guarantee such a level of precision? Assume here that the auditor is essentially sure that $p < .20$.

4. (18 points) Let Y_1, Y_2, Y_3, Y_4 be iid $N(0, 1)$ random variables. Let $\bar{Y} = \sum_{i=1}^{4} Y_i/4$ and $S^2 = \sum_{i=1}^{4}(Y_i - \bar{Y})^2/3$.

   Answer these questions, explaining your answers carefully and stating appropriate theorems.

   (a) What is $E(S^2)$?

   (b) Find $P(2\bar{Y} > 1.25S)$.

   (c) Find $Var(\bar{Y}^2 + S^2)$. Make particularly sure you explain this one clearly at every step.

5. (12 points) Suppose I weigh an object 3 times on a spring scale and get 10.2, 10.3, and 10.7 ounces. Suppose the distribution of weighings is like an iid sample from a $N(\mu, \sigma^2)$ distribution where the true weight of the object is μ.

   (a) Suppose σ is known to be 0.2 ounces. Find a 95% confidence interval for μ.

   (b) Find a 95% confidence interval for μ if σ is unknown.