Exam 3

14.30, Fall 2000

This exam is closed-notes and closed-book, although you may bring in one page of notes. You may use a calculator. Please consult the sixth page of this exam for notation, PMFs/PDFs, means, and variances of special distributions.

READ THROUGH THE EXAM FIRST, in order to ask clarifying questions and to allocate your time appropriately. In order to receive partial credit in the case of computational errors, please show all work. You have approximately 3 hours in which to complete the exam. Good luck!

1. (20 pts.) True or False with short explanation. Only write a sentence or two for each. I will not read long explanations.

a) Knowing the distribution of a test statistic under the null allows one to calculate α and β (the probabilities of type I and type II error).

b) A 90% confidence interval for an unknown parameter contains that parameter with 90% probability.

c) A minimum mean squared error estimator is defined as the most efficient estimator among the class of unbiased estimators.

d) Maximum likelihood estimators are asymptotically normally distributed.

2. (20 pts.) While making toffee I have a constant probability (p) of ruining a batch. Assume that batches are independent.

a) My kids saw me attempt four batches and ruin two. How could they estimate p and what would their estimate be? Would their answer change depending on the estimation framework they used? Explain.

b) My husband, quite familiar with my toffee-making problems, knows that two-thirds of the batches end up in the trashcan. He did not observe the number of batches I attempted (n), but if he sees two batches in the trashcan, what is his estimate of n? Would the answer vary depending on the estimation framework he used? Explain.

3. (20 pts.) You have one observation z that takes on one of four values according to one of the three distributions shown below:

	z_1	z_2	z_3	z_4
$f_1(z)$	0.4	0.1	0.2	0.3
$f_2(z)$	0.2	0.2	0.3	0.3
$f_3(z)$	0.2	0.0	0.0	0.8

a) Determine the possible critical regions and corresponding values for α for a generalized likelihood ratio test of $H_0: \ f(z) = f_1(z)$ against $H_A: \ f(z) = f_2(z)$ or $f_3(z)$.

b) Keeping in mind that knowing the distribution of your test statistic under the null is an important practical consideration, could you perform a similar test to the one above if you had 500 observations? Explain.

4. (20 pts.) A graduate student in political science is performing a study on voter apathy. He polls a randomly selected group of registered voters about how far they would be willing to travel to vote, denoting their responses $X_i,\ i = 1, \ldots 10$. He computes

$$\bar{X} = \frac{1}{10} \sum_{i=1}^{10} X_i = 4.3$$

$$s^2 = \frac{1}{9} \sum_{i=1}^{10} (X_i - \bar{X})^2 = 2.7.$$

He then infers that for 95% of the registered voters the max they are willing to travel is within $[2.7, 5.9]$

a) What assumptions underly this inference?

b) Taking the assumptions you outlined above as given, construct a 95% confidence interval for the mean of the X_i's. How many voters would he have to poll to limit the width of the confidence interval to a quarter of a mile?

c) Suppose that the graduate student is given the additional information that the X_i's have a uniform distribution on $[0, \theta]$. Unfortunately, he has already thrown away his original data and kept only the sample mean and sample variance he computed. Can he still construct a 95% confidence interval for θ? Explain.

5. (20 pts.) Find the power function ($r(\theta) = 1 - \beta(\theta)$) of the test which rejects $p = 1/2$ when fewer than two or more than three successes occur in five independent Bernoulli(p) trials. Compare this with the power function of the test which rejects $p = 1/2$ when more than three successes occur and rejects randomly 60% of the time when three successes occur. Is α the same for the two tests? Is there any region of the parameter space where the second test gives higher power?

Have a great break!