DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF MASSACHUSETTS
BASIC EXAM - STATISTICS
August 29, 2003

Work all problems. Sixty points are needed to pass at the Master's level and seventy-five at the Ph.D. level.

1. (30 points) The assessment of the proportion of defective units in a lot of units is an important problem. Suppose you take a random sample of 100 units from a lot large enough to treat $X_1, \ldots, X_{100}$ as i.i.d. Bernoulli (p), where $X_i = 1$ if unit i in the sample is defective and is 0 otherwise. Hence, p is the probability of getting a defective unit or equivalently, the proportion of defective units in the population. Let

$$\hat{p} = \text{proportion of defective items in the sample}$$

   (a) Define an unbiased estimator of p in general, and then justify that $\hat{p}$ is unbiased for p.

   (b) Define a consistent estimator of p in general, and then justify that $\hat{p}$ is consistent for p.

   (c) What is the approximate distribution of $\hat{p}$? What theorem did you use to get the result?

   (d) Find an approximate 95% confidence interval for p. Justify your answer.

   (e) Find an approximate 90% confidence interval for the odd-ratio $\theta = p/(1-p)$ assuming $p \neq 1$. Explain your answer.

2. (30 points) Let $X_1, \ldots, X_m$, i.i.d. $\sim N(\mu_1, \sigma^2)$, and $Y_1, \ldots, Y_n$, i.i.d. $\sim N(\mu_2, \sigma^2)$, be two independent samples corresponding to a treatment and a control group respectively. Let $\overline{X}$ and $\overline{Y}$ denote the two sample means and S_p^2 the pooled estimate of the variance.

(a) It is known that S_p^2 multiplied by some constant, say, K, will have a chi-squared distribution with M degrees of freedom. What are K and M, respectively? You can just state the result. Use this result to derive an exact $1-\alpha$ confidence interval for σ^2.

(b) State what the distribution of $\overline{X} - \overline{Y}$ is.

(c) Accepting that $\overline{X} - \overline{Y}$ is independent of S_p^2, justify what the distribution of

$$\frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{[S_p^2(m^{-1} + n^{-1})]^{1/2}}$$

is. You do not need to actually get a density function but show that it is a named distribution that arises from the results of the previous two parts.

(d) Use the result of the previous part to derive a confidence interval for $\mu_1 - \mu_2$, with confidence level $1-\alpha$.

3. (40 points) Let $X_1, \cdots, X_n$ denote random sample of lifetimes from an exponential distribution with mean θ.

(a) Find the MLE of θ. Be complete, clearly writing down what the likelihood function is and verifying that your solution maximizes the likelihood.

(b) Find a complete sufficient statistic and then use this to find the UMVUE (uniform minimum variance unbiased estimator) of θ. State clearly what general results you are applying.

(c) Derive the likelihood ratio test for testing $H_0: \quad \theta = \theta_0$ against $H_1: \quad \theta \neq \theta_0$ at the .05 significance level. Give the test in terms of a known distribution. (Hint: $2X_i/\theta$ is distributed Chi-square with $2n$ degrees of freedom.)

(d) Show how to calculate the power function for the test in the previous part; that is, how to obtain $\pi(\theta)$ = probability of rejecting H_0 when θ is the true value.