DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF MASSACHUSETTS
BASIC EXAM - STATISTICS
August 29, 2007

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

1. (40 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 n ($n > 100$) units from a lot large enough to treat $X_1, \ldots, X_n$ as i.i.d. Bernoulli (θ), where $X_i = 1$ if unit i in the sample is defective and is 0 otherwise. Hence, θ is the probability of getting a defective unit or equivalently, the proportion of defective units in the population. Let $Y = X_1 + \cdots + X_n$, and

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

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

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

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

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

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

   (f) Consider a Bayes approach for estimating θ. Suppose that the prior density for θ is $\pi(\theta) = 12\theta(1-\theta)^2, \quad 0 < \theta < 1$.

      i. Find the posterior density of θ given $Y = y$. (Make sure what you get is a pdf.)

      ii. Find the Bayes estimator of θ given $Y = y$.

2. (20 points) Let $X_i, i = 1, \cdots, n$ be independently distributed as $N(\alpha + \beta t_i, \sigma^2)$ where α, β, and σ^2 are unknown, the t's are known constants that are not all equal, and $\sum_{i=1}^n t_i = 0$.

   (a) Show that the joint distribution of X's belongs to the exponential family. Find the complete and sufficient statistic for $(\alpha, \beta, \sigma^2)$.

   (b) Show that $\hat{\alpha} = \bar{X}$ and $\hat{\beta} = \sum_{i=1}^n t_i X_i / \sum_{i=1}^n t_i^2$ are UMVUEs of α and β.

   (c) Show that $\hat{\beta}$ is aysmptotically efficient by showing that it is the MLE of β.

3. (20 points) Let $X_i, i = 1, \cdots, n$ be a random sample from an $N(0, \sigma^2)$ distribution.

   (a) Show that the uniformly most powerful (UMP) test for

   $$H_0 : \sigma = \sigma_0 \quad H_1 : \sigma > \sigma_0.$$

   is based on the statstic $T = \sum_{i=1}^{n} X_i^2$. (You need to explain why the test is UMP.)

   (b) Specify the distribution of T, with justifications.

   (c) Determine the rejection region of the UMP test at level α.

   (d) Calculate and plot the power function of the UMP test.

4. (20 points) A random sample $X_i, i = 1, \cdots, n$ is drawn from a Pareto population with pdf

   $$f(x|\theta, \nu) = \theta \nu^{\theta} / x^{\theta+1}, \quad x > \nu, \theta > 0, \nu > 0.$$

   (a) Find the MLEs for θ and ν.

   (b) Show that the LRT of

   $$H_0 : \theta = 1, \nu \text{ unknown} \quad H_1 : \theta \neq 1, \nu \text{ unknown},$$

   has critical region of the form $\{\mathbf{x} : T(\mathbf{x}) \leq c_1 \text{ or } T(\mathbf{x}) \geq c_2\}$, where $0 < c_1 < c_2$ and

   $$T = \sum_{i=1}^{n} \log X_i - n \log(\min_i X_i).$$

Appendix: Possible Useful Facts

1. Beta(α, β) variable pdf:

   $$f(x|\alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}; \quad 0 \leq x \leq 1; \quad \alpha > 0, \quad \beta > 0$$

   $$E(X) = \frac{\alpha}{\alpha+\beta}, \quad Var(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}, \quad \Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$$

2. Chi squared vriable pdf:

   $$f(x|k) = \frac{1}{\Gamma(k/2)2^{k/2}} x^{k/2-1} e^{-x/2}; \quad x \geq 0; \quad k = 1, 2, \cdots$$

   $$E(X) = k, \quad Var(X) = 2k$$