DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF MASSACHUSETTS
BASIC EXAM - STATISTICS
August 31, 2009

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

1. (25 pts) Let $X_1, \ldots, X_n$ be a random sample from a uniform distribution on $(0, \theta)$ and let $Y = \max\{X_1, \ldots, X_n\}$. We are interested in constructing an interval estimator for θ. You may use, without proof, the fact that the probability density function (pdf) of Y is $f_Y(y) = \frac{ny^{n-1}}{\theta^n}$ where $0 < y < \theta$.

   (a) Find the pdf of $T = \frac{Y}{\theta}$.

   (b) Is T in (a) a pivotal quantity? (In other words, does the distribution of $T = Y/\theta$ depend on θ?)

   (c) Use T to find a 95 lower one-sided confidence interval for θ. (Here, a lower one-sided confidence interval means an interval placing a lower bound on the value of θ.)

   (d) Suppose that in a sample of size 20 we obtain the interval $(947.5402, \infty)$ in part (c). Explain the meaning of the statement "$(947.5402, \infty)$ is a 95% lower one-sided confidence interval for θ".

2. (25 pts) The EPA conducts occasional reviews of its standards for airborne asbestos. During a review, the EPA examines data from 20 studies. Different studies keep track of different groups of people, and different groups have different exposures to asbestos. Let n_i be the number of people in the ith study, let x_i be their asbestos exposure, and let y_i be the number who develop lung cancer in the ith study. The EPA's model is $y_i \sim Poisson(\lambda_i)$ where $\lambda_i = n_i x_i \lambda$ and where λ is the typical rate at which asbestos causes cancer. The n_is and x_is are known constants, and the y_is are random variables. The EPA wants a posterior distribution for λ. (The Poisson probability mass function is $f(y_i|\lambda_i) = \lambda_i^{y_i} \exp(-\lambda_i)/y_i!, \lambda_i > 0, y_i = 0, 1, 2, \ldots$)

   (a) Write down the likelihood function for λ.

   (b) Find a one-dimensional sufficient statistic for lambda.

   (c) What one parameter distribution would be convenient prior distribution for λ?

   (d) If the EPA used the prior $\lambda \sim gamma(a, b)$, what would be the EPA's posterior? (You may use that the the gamma probability density function is $f(z|a, b) = \frac{1}{b^a \Gamma(a)} z^{a-1} \exp(-z/b), a > 0, b > 0, z > 0.$)

3. (25 pts) A measuring instrument is run n times on a known standard which has a known value μ_0. The resulting observations $X_1, \ldots, X_n$ are assumed to be independent and normally distributed with mean μ_0 and an unknown variance σ^2.

   (a) Find a maximum likelihood estimator for σ^2 and show that it is a maximum likelihood estimator.

   (b) Is the estimator that you found in (a) unbiased for σ^2? Why or why not?

   (c) What is the standard error of the estimator you found in (a)?

   (d) Is the estimator that you found in (a) consistent for σ^2? Why or why not?

4. (25 pts) Suppose $X_1, \ldots, X_n$ is a random sample from the distribution with density

$$f(x|p) = \binom{n}{x} p^x (1-p)^{n-x}, 0 \leq p \leq 1, 0 \leq x \leq n, x \text{ an integer.}$$

   (a) Find a maximum likelihood estimator for p and show that it is a maximum likelihood estimator.

   (b) What is the variance of the estimator you found in (a)?

   (c) Assuming $0 < p < 1$, what is the approximate distribution of the estimator you found in (a) as n gets large?

   (d) Use your result from part (c) to find an approximate 95% confidence interval for p.