Probability and Statistical Theory

MS Comprehensive Examination

April 11, 2015

Instructions:

Please answer all three questions.

Record your answers in your blue books.

Show all of your computations.
Prove all of your assertions or quote the appropriate theorems.
Books, notes, and calculators *may be used.*

You have three hours.

1. (25 points) $X_1, \cdots, X_n \overset{iid}{\sim} P(X_i = j) = p_j$ $(j = 1, \cdots, m)$ are a total of n independent trials, each of the trials results in exactly one of m possible outcomes, and each outcome has a success probability $0 \le p_j \le 1$ $(1 \le j \le m, \sum_{j=1}^{m} p_j = 1)$. Define $Y_j = \sum_{i=1}^{n} I(X_i = j)$, where $I(X_i = j) = 1$ if $X_i = j$ and 0 otherwise. Y_j is the number of X's equal to j and $\sum_{j=1}^{m} Y_j = n$.

   a. Give the names of the joint distribution of $(Y_1, \cdots, Y_m)$ and marginal distribution of Y_j.
   b. Show that $\mathrm{E}(Y_j) = np_j$ and $\mathrm{V}(Y_j) = np_j(1 - p_j)$.
   c. Show that $\mathrm{Cov}(Y_j, Y_k) = -np_jp_k$ $(j \ne k)$.
   d. Find the likelihood ratio test (LRT) statistic for $H_0 : p_1 = \cdots = p_m$.
   e. Describe how you use the LRT statistic to test $H_0 : p_1 = \cdots = p_m$, that is how you calculate the p-value or the rejection region.

2. (25 points) $X_1, \cdots, X_n \overset{iid}{\sim} f(x|\theta) = \frac{4x^3}{\theta} \exp\left(-\frac{x^4}{\theta}\right)$ for $x \ge 0$ and $\theta > 0$.

   a. Find a complete and sufficient statistic for θ.
   b. Find the UMVUE $\tilde{\theta}$.
   c. Calculate the variance of $\tilde{\theta}$.
   d. Calculate the Cramér-Rao Lower Bound. Does the UMVUE reach it?
   e. Use the Karlin-Rubin Theorem to find a UMP level α test $\phi(T)$ for

$$H_0 : \theta = \theta_0 \text{ vs } H_1 : \theta > \theta_0.$$

   Write the rejection region R of this test using the test statistic T and the quantile of a well-known distribution.

3. (50 points) Throughout this problem, we have independent observations $X_1,\ldots,X_n$ from a density function $f(x) = {}^{c}/_{x^4}$ for $0 < \theta \le x < \infty$. Answers in parentheses are given for some parts so that if you get stuck on one part, you can still do the next. So if you can't do one part, look at the answers and look ahead to later parts to see what you can do.

   a. Find c.

   b. Find the CDF F(x) for all x. Check your answer by checking the lower and upper limits on Range(X) and the density f(x).

   c. Find the mean μ and variance σ^2 for this distribution. (answers: $\mu = 3\theta/2$ and $\sigma^2 = 3\theta^2/4$)

   d. For the random sample, find, simplify, and sketch the likelihood function $L(\theta; x_1,\ldots,x_n)$.

   e. Identify a sufficient statistic for θ. State why. (answer: the minimum, $X_{(1)}$)

   f. Find the maximum likelihood estimator (MLE) for θ.

   g. Find the CDF, density, mean, bias and variance of this MLE. Show that it is asymptotically unbiased. (partial answer: $F_{(1)}(y) = 1 - (\theta / y)^{3n}$ and $\mathrm{Var}(X_{(1)}) = {}^{3n\theta^2}/_{(3n-2)(3n-1)^2}$)

   h. Find the method of moments estimator (MME) for θ.

   i. Find the mean and variance of the MME. Show that it is unbiased.

   j. Find the asymptotic relative efficiency of these estimators, that is, the limit as $n \to \infty$ of the ratio of Var(MLE) / Var(MME).

   k. We wish to test H_0: $\theta = 1$ vs. H_A: $\theta > 1$ at level of significance $\alpha > 0$. Find the likelihood ratio test (LRT) statistic λ as a function of the data.

   l. The LRT statistic can be written in terms of the MLE for θ whose distribution we know from part g. Find the test in terms of this statistic.

   m. Solve for the critical value for the test in part l in terms of α, so that if we have n and α we can find the actual number. Do this for level of significance $\alpha = 0.10$ and $n = 4$. (answer for the formula: $c = \alpha^{-1/3n}$)

   n. Find and sketch the power function for this test. Be careful, there are cases.