Probability and Statistical Theory

MS Comprehensive Examination

April 14, 2007

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.

Comprehensive Examination April 14, 2007

1. (50 points) Suppose X_i's $(1 \leq i \leq n)$ are iid with the density function

$$f(x|a,b) = \frac{ba^b}{x^{b+1}}, x \geq a > 0, b > 0.$$

Suppose a is **KNOWN.** Answer the following questions.

(1) Find $\mathrm{E}(X)$ $(b > 1)$.

(2) Show that

$$2bG \sim \chi^2(2n), \tag{0.1}$$

where $G = \log\left(\frac{\prod_{i=1}^n X_i}{a^n}\right)$.

(3) Find MLE $\hat{b}$.

(4) Find a complete and sufficient and statistic for b.

(5) Find the UMVUE $\tilde{b}$.

(6) Calculate the variance of $\tilde{b}$.

(7) Calculate the Cramér-Rao Lower Bound. Does UMVUE reach it?

(8) Use the K-R Theorem to find a UMP level α test $\phi_1(\vec{x})$ for

$$H_0 : b = b_0 \text{ vs } H_1 : b > b_0.$$

(9) By the test in (8), find a UMA $1 - \alpha$ lower confidence bound for b.

(10) Consider $H_0 : b = b_0$ vs. $H_1 : b \neq b_0$. Show that an LRT $\phi_2(\vec{x})$ with the rejection region

$$R = \left\{\vec{x} : 2b_0G \leq \chi_{1-\frac{\alpha}{2}}(2n) \text{ or } 2b_0G \geq \chi_{\frac{\alpha}{2}}(2n)\right\}$$

is a level $\alpha \in (0,1)$ test, where G is defined in (0.1).

2. A bag has 10 balls, R of which are red; the remainder are white. A random sample, *with replacement,* of size n = 4 is taken from the bag. Denote the number of red balls in this sample by X.

Part A. (20 points)

For this part, the number of red balls in the bag, R, is random, with probability distribution given by $f(r) = P(R = r) = (r+1)/15$, for $r \in \{0,1,2,3,4\}$.

a. Identify the conditional distribution of X given R. Also state the conditional mean and variance of X given R.
b. Find the expectation and variance of R. Show your work. Answers: 8/3 and 14/9.
c. Use parts a and b and the appropriate conditional expectation and variance formulas to find E(X) and var(X). Answers: 16/15 and 218/225.
d. The first six columns of the table at the bottom of the page give the appropriate conditional distributions P(X=x | R=r). The last three columns are as labeled, with x^2 standing for "x squared". The row labeled "Sum" is the sum for each column. The last two columns are products of the two terms. For each of the 5 values in **bold-faced type and underlined,** give the appropriate formula and show the calculations that lead to that number.
e. Use the table to write down 1) the distribution of X, 2) E(X) and 3) var(X), and confirm that you arrive at the same answers you found in part c.

Part B. (20 points)

Here we regard R as an unknown fixed parameter. Using one observation on X, we will test $H_0{:}R=4$ versus $H_A{:}R \in \{0,1,2,3\}$ using a level of significance α no larger than .20.

f. State the domain of the likelihood function L. Then state clearly how we can find L from the table below.
g. Use the table below to find the maximum likelihood estimate, $\hat{R}$, of R as a function of x. State precisely how you do this.
h. For the test given above, use the table below to find the likelihood ratio test statistic λ(x). State the definition of λ and clearly show your work.
i. Using the table below, find the critical region associated with the likelihood ratio test that satisfies the restriction on α given above. State clearly what is the test, and state what is your actual level of significance α. State the test in terms of λ and the equivalent test in terms of X.
j. Using the table below, find and graph the power function for this test.

Table for Parts d - j:

	P(X=x | R=r)							
x	r=0	r=1	r=2	r=3	r=4	P(X=x)	x P(X=x)	x^2 P(X=x)
0	1.0	**<u>0.6561</u>**	0.4096	0.2401	0.1296	0.3433	0.0000	0.0000
1	0.0	0.2916	0.4096	0.4116	0.3456	**<u>0.3458</u>**	0.3458	0.3458
2	**<u>0.0</u>**	0.0486	**<u>0.1536</u>**	0.2646	0.3456	0.2230	0.4459	0.8918
3	0.0	0.0036	0.0256	0.0756	0.1536	0.0770	0.2309	**<u>0.6926</u>**
4	0.0	0.0001	0.0016	0.0081	0.0256	0.0110	0.0441	0.1764
Sum	1.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0667	2.1070

3. **(10 points)** Let Range(X,Y) be the unit square $[0,1]\times[0,1]$ and let f(x,y) = $c(x+\theta y)$ for $\theta \geq 0$.

a. Find c as a function of θ.
b. Find the marginal distributions of X and Y. Check your answers by confirming that the functions are actually densities.
c. For what values of θ in the parameter space, if any, are X and Y independent? Show why or why not.
d. Give at least one example of how you could define a method of moments estimator for θ based on one observation of the vector (X,Y). Show your work, i.e., why this makes sense as a method of moments estimator.