Statistics Masters Comprehensive Exam

November 3, 2001

Student Name: ____________________

1. Answer 8 out of 12 problems. Mark the problems you selected in the following table.

1	2	3	4	5	6	7	8	9	10	11	12

2. Write your answer right after each problem selected, attach more pages if necessary.

3. Assemble your work in right order and in the original problem order.

1. Let $f(x, y) = 3/2, x^2 \le y \le 1, 0 \le x \le 1$ be the joint p.d.f. of X and Y. Find

   (a) $P(1/2 \le X)$.

   (b) $P(X \le Y)$.

   (c) Marginal p.d.f. of X and Y.

2. Let X_1 and X_2 be independently and identically distributed from a geometric distribution with probability mass function given by

$$P(X = x) = p(1 - p)^{x-1}, \text{ where } x = 1, 2, 3, \ldots.$$

(a) Find the UMVUE for $1/p$.

(b) Find the UMVUE for p.

3. Let $\{X_1, \ldots, X_n\}$ be independently and identically distributed with density $f(x; \mu, \sigma^2)$, where μ is the mean value and σ^2 the variance. Put $Y = \sum_{i=1}^n X_i$ and $S^2 = \sum_{i=1}^n X_i^2$.

   (a) Define the central limit theorem for Y.

   (b) If $f(x; \mu, \sigma^2)$ is normal with $\mu = 0$, show that the variable $Z = \{S^2 - n\sigma^2\}/\{\sigma^2(2n)^{-1/2}\}$ converges in distribution to $N(0, 1)$ as $n \to \infty$, where $N(0, 1)$ denotes the normal distribution with mean 0 and variance 1.

4. Let X_1, X_2, X_3 be random variables.

   (a) If $X_1 \sim Poisson(\lambda),\ X_2 \sim Poisson(\mu)$, and X_1, X_2 are independent find $P(X_1 = k|X_1 + X_2 = 2k)$

   (b) If $(X_1, X_2, X_3) \sim \text{Trinomial}(n, p_1, p_2, p_3)$, that is,

$$P(X_1 = x_1, X_2 = x_2, X_3 = x_3) = \frac{n!}{x_1!x_2!x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3},$$

   where $p_3 = 1 - p_1 - p_2,\ X_3 = n - X_1 - X_2$, find $P(X_1 = k|X_1 + X_2 = m)$.

5. Let $\{X_1, \ldots, X_n\}$ be a random sample from the population with density $f(x, \theta) = \theta^{-1}\, e^{-x/\theta}, x > 0, \theta > 0$.

   (a) Derive the UMP (Uniformly Most Powerful) size α test for testing $H_0 : \theta = 1$ versus $H_1 : \theta > 1$.

   (b) What is the power function of your test ?

6. Let a discrete random variable X with probability mass function $f(x;\theta)$, where $\theta \in \{1,2,3\}$ and

x	$f(x;1)$	$f(x;2)$	$f(x;3)$
1	0.4	0.25	0.1
2	0.3	0.25	0.2
3	0.2	0.25	0.3
4	0.1	0.25	0.4

Find the MLE of θ, if we observe the following sample:

(a) $1, 2$

(b) $2, 4$

(c) $1, 2, 2, 4$.

7. Suppose that diseased trees are distributed randomly and uniformly throughout a large forest with an average of λ per acre. The numbers of diseased trees observed in ten four-acre plots were $1, 1, 3, 2, 0, 2, 2, 0, 1, 1$.

   (a) Find the maximum likelihood estimate of λ.

   (b) Test $H_0 : \lambda = 0.2$ versus $H_1 : \lambda > 0.2$.

8. Let X be a central F-distribution with degrees of freedoms $\{f_1, f_2\}$. That is, the density of X is:

$$f_X(x) = \frac{1}{B(f_1/2, f_2/2)}(f_1/f_2)^{f_1/2}x^{f_1/2-1}\left(1+\frac{f_1}{f_2}x\right)^{-(f_1+f_2)/2}, x > 0,$$

where the f_i's are positive integers.

(a) Show that $Z = \frac{f_1 X}{f_2 + f_1 X}$ is distributed as a Beta- variable with parameters $\{f_1/2, f_2/2\}$. That is, the density of Z is:

$$g_Z(z) = \frac{1}{B(f_1/2, f_2/2)} z^{f_1/2-1}(1-z)^{f_2/2-1}, 0 < z < 1.$$

(b) Given that X is distributed as a central F-distribution with degrees of freedoms $\{f_1 = 5, f_2 = 10\}$, show that the probability $P(X \le 4)$ is given by:

$$P(X \le 4) = \frac{1}{B(2.5, 5)} \int_0^{\frac{2}{3}} x^{1.5}(1-x)^4 dx.$$

9. Let $X_1, X_2, \cdots, X_n$ be a random sample taken from the distribution with the p.d.f.

$$f(x;\theta) = 1/\theta\ x^{1/\theta - 1}, \quad 0 < x < 1, 0 < \theta < \infty.$$

(a) Find the moment estimator of θ.

(b) Find the maximum likelihood estimator of θ.

(c) Find Rao-Cramér lower bound of any unbiased estimator $\hat{\theta}$ for θ.

10. In a certain class, $(X_1, Y_1), \ldots, (X_7, Y_7)$ were measured where X_i was the score on homework of student i and Y_i was student i's subsequent score on an examination. The data are

i	1	2	3	4	5	6	7
x_i	0	10	49	52	59	64	64
y_i	12	20	60	43	37	58	63

In fact, $\sum_{i=1}^{7} x_i = 298$, $\sum_{i=1}^{7} y_i = 293$, $\sum_{i=1}^{7} x_i^2 = 16878$, and $\sum_{i=1}^{7} x_i y_i = 15303$.

(a) Find the estimated regression line of Y on X: $y = \hat{\alpha} + \hat{\beta}x$.

(b) Plot the seven data points and the estimated regression line on the same graph.

(c) If a similar student in a similar situation in the future obtained a homework score of 30, what would be the best linear estimate of his subsequent examination score?

11. Let $\{X_1, \ldots, X_n\}$ be a random sample from the population with density $f(x;\theta) = \theta^x(1-\theta)^{1-x}, x = 0, 1; 0 < \theta < 1$.

(a) Put $\bar{X} = \frac{Y}{n}$, where $Y = \sum_{i=1}^{n} X_i$. Show that $\bar{X}$ is the UMUVE (Uniformly Minimum Variance Unbiased Estimator) of θ.

(b) Show that $\frac{Y(Y-1)}{n(n-1)}$ is the UMUVE of θ^2.

12. For a Poisson population, it is desired to test $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda = \lambda_1$ $(\lambda_0 < \lambda_1)$.

(a) How large a sample size is needed to obtain an α-level most powerful test with power $1 - \beta$? (Assume the sample size is large enough to apply the Central Limit Theorem.)

(b) Compute the sample size in part (a) if $\lambda_0 = 2, \lambda_1 = 1, \alpha = 0.05$ and $1 - \beta = 0.90$.