Comprehensive Examination on the Theory of Mathematical Statistics
Department of Statistics, University of Florida

January 7, 2002

Instructions:

1. You have three hours to answer questions in this examination.
2. There are 6 problems of which you must answer 5.
3. While the questions are equally weighted, some problems are more difficult than others.
4. Write only on one side of the paper, and start each question on a new page.

You may use the following facts/formulas without proof:

Iterated Expectation Formula: $\mathrm{E}(X) = \mathrm{E}\left[\mathrm{E}(X|Y)\right]$.

Iterated Variance Formula: $\mathrm{Var}(X) = \mathrm{E}\left[\mathrm{Var}(X|Y)\right] + \mathrm{Var}\left[\mathrm{E}(X|Y)\right]$.

Delta Method: Let Y_n be a sequence of random variables such that $\sqrt{n}\,(Y_n - \theta) \xrightarrow{d} \mathrm{N}(0, \sigma^2)$. For a given function g and a specific value of θ, suppose that $g'(\theta)$ exists and is not 0. Then

$$\sqrt{n}\left[g(Y_n) - g(\theta)\right] \xrightarrow{d} \mathrm{N}\Big(0, \sigma^2 \left[g'(\theta)\right]^2\Big) \ .$$

1. Let X and Y be jointly continuous random variables with joint density function given by

$$f(x,y) = \begin{cases} y\, e^{-y(x+1)} & x > 0, y > 0 \\ 0 & \text{otherwise} \end{cases}$$

(a) Are X and Y independent? Why?

(b) Find the marginal density functions of X and Y.

(c) Find $P(X > 1 | Y > \pi)$.

(d) Find $P(X > 1 | Y = \pi)$.

(e) Find the expectations of X and Y?

2. In this question, we will derive the mean and variance of a hypergeometric random variable. Imagine an urn with M white balls and $N - M$ black balls. Suppose K balls are draw from the urn at random *without replacement*. Let Y denote the number of white balls in the sample. Clearly, $Y \sim \text{HG}(N, M, K)$. For $i = 1, 2, \ldots, K$, define

$$X_i = \begin{cases} 1 & \text{if } i\text{th ball drawn is white} \\ 0 & \text{if } i\text{th ball drawn is black .} \end{cases}$$

Throughout this problem, you may use the following two facts:

- The random variables $X_1, X_2, \ldots, X_n$ are identically distributed.
- The joint distribution of (X_i, X_j) is the same for all $i \neq j$.

(a) Show that Y can be written as a simple function of the X_i's.

(b) Find the probability mass function (PMF) of X_1 and use it to calculate $\text{E}(Y)$.

(c) Prove that X_1 and X_2 are *not* independent.

(d) Find the joint PMF of (X_1, X_2) and use it to calculate $\text{Cov}(X_1, X_2)$.

(e) Let $W_1, \ldots, W_n$ be a set of random variables each with a finite second moment. Show that

$$\text{Var}\left(\sum_{i=1}^{n} W_i\right) = \sum_{i=1}^{n} \text{Var}(W_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(W_i, W_j) \ .$$

(f) Find $\text{Var}(Y)$.

3. Suppose $X_1, \ldots, X_n$ are iid Bernoulli(p) and that $n \geq 4$.

(a) Find the MLE of p and call it $\hat{p}$.

(b) Show that the variance of $\hat{p}$ attains the Cramér-Rao Lower Bound.

(c) Show that $\prod_{i=1}^{4} X_i$ is an unbiased estimator of p^4.

(d) Find the UMVUE of p^4.

4. (a) Suppose that $X_1, \ldots, X_n$ are independent random variables with $X_i \sim \text{Gamma}(\alpha_i, \beta)$; that is, X_i has probability density function given by

$$f_{X_i}(x) = \begin{cases} \frac{1}{\Gamma(\alpha_i)\,\beta^{\alpha_i}}\, x^{\alpha_i - 1} \exp\{-x/\beta\} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha_i, \beta > 0$. Derive the distribution of $\sum_{i=1}^n X_i$.

(b) Define $U_i = X_i/(X_1 + X_2 + \cdots + X_n)$ for $i = 1, \ldots, n$. Show that

$$U_i \sim \text{Beta}\left(\alpha_i, \sum_{j \neq i} \alpha_j\right) .$$

(Hint: Think of U_i as $X_i/(X_i + W)$ where $W = \sum_{j \neq i} X_j$ is independent of X_i.)

(c) Let $X_1, \ldots, X_n$ be iid $\text{Gamma}(\alpha, \beta)$. Suppose that, conditional $X_1, \ldots, X_n$, $Y_1, \ldots, Y_n$ are independent and such that $Y_i | X_i \sim \text{Gamma}(\alpha, \beta\, X_i)$. Show that

$$\text{E}\left(\frac{\bar{Y}}{\bar{X}}\right) = \alpha\beta \quad \text{and} \quad \text{Var}\left(\frac{\bar{Y}}{\bar{X}}\right) = \alpha\beta^2\, \text{E}\left[\frac{\sum X_i^2}{(\sum X_i)^2}\right] .$$

(Hint: Use the iterated expectation and variance formulas.)

(d) Now use (b) to evaluate

$$\text{E}\left[\frac{\sum X_i^2}{(\sum X_i)^2}\right] ,$$

leading to a formula for $\text{Var}(\bar{Y}/\bar{X})$.

5. Let $X_1, \ldots, X_n$ be iid $\text{Uniform}\big(\theta - \frac{1}{2}, \theta + \frac{1}{2}\big)$, where θ is an unknown parameter.

(a) Write down the likelihood function for θ and use it to show that $(X_{(1)}, X_{(n)})$ is sufficient for θ. (Of course, $X_{(1)}$ and $X_{(n)}$ are the smallest and largest order statistics, respectively)

(b) Show that the likelihood function can be written as

$$L(\theta; x_1, \ldots, x_n) = \begin{cases} 1 & \text{if } \theta \in \big(x_{(n)} - \frac{1}{2}, x_{(1)} + \frac{1}{2}\big) \\ 0 & \text{otherwise} . \end{cases}$$

(c) Show that $\hat{\theta} = (X_{(1)} + X_{(n)})/2$ is an unbiased maximum likelihood estimator of θ.

(d) Derive the density of $\hat{\theta}$ when $n = 2$.

(e) Find formulas for $P(\hat{\theta} > \theta + t)$ and $P(\hat{\theta} < \theta - t)$ for $0 < t < 1/2$ and use them to derive an exact 90% confidence interval for θ. (Hint: You don't need any integrals here.)

(f) Suppose that $X_1 = 5.7$ and $X_2 = 4.9$. What are the minimum and maximum possible values of θ?

(g) Calculate an exact 90% two-sided confidence interval for θ based on the sample data in part (e). Comment on your interval in light of your answer to part (e) and comment on your answer.

6. Suppose $Z|\theta \sim$ Geometric(θ); that is,

$$P(Z = z|\theta) = \theta(1-\theta)^z$$

for $z \in \mathbb{Z}_+ = \{0, 1, 2, \dots\}$ and $\theta \in (0, 1)$.

(a) Show that $\mathrm{E}(Z|\theta) = \frac{1-\theta}{\theta}$ and that $\mathrm{Var}(Z|\theta) = \frac{1-\theta}{\theta^2}$.

(b) Suppose that $\theta \sim$ Beta$(\alpha, 1)$ with $\alpha > 2$. Use the iterated expectation and variance formulas to find $\mathrm{E}Z$ and $\mathrm{Var}(Z)$. Also, show that marginally, for $z \in \mathbb{Z}_+$

$$P_\alpha(Z = z) = \frac{\alpha^2\,\Gamma(\alpha)\,z!}{\Gamma(z+\alpha+2)}\ . \tag{1}$$

(c) Let $Z_1, Z_2, \dots, Z_n$ be an iid sequence from the mass function in (1) with $\alpha > 2$. Let $\tilde{\alpha}_n$ denote the method of moments estimator of α. Find $\tilde{\alpha}_n$ and show that

$$\sqrt{n}\,(\tilde{\alpha}_n - \alpha) \xrightarrow{d} \mathrm{N}\left(0, \sigma^2(\alpha)\right)$$

where

$$\sigma^2(\alpha) = \frac{\alpha^2\,(\alpha-1)^2}{\alpha-2}\ .$$

(Hint: Use the delta method.)

(d) Suppose that $n = 1$. Construct a UMP size 0.75 test of $H_0 : \alpha \le 3$ versus $H_A : \alpha > 3$. You may use the fact that for fixed $0 < a < b$, the function

$$g(t) = \frac{\Gamma(t+a)}{\Gamma(t+b)}$$

is decreasing in t.