Stat 252 Winter 2007
Assignment #10

This assignment is due at the beginning of class on Wednesday, April 11, 2007. You must submit all problems that are marked with an asterix (*).

1. * Carefully solve parts **(a)**, **(c)**, **(d)**, **(e)**, **(f)**, **(g)**, and **(i)** from the Generalized Likelihood Ratio Test handout distributed in class on Wednesday, April 2, 2007.

2. A continuous random variable Y is said to have the Rayleigh(θ) distribution if the probability density function of Y is

$$f_Y(y|\theta) = \frac{y}{\theta^2} \exp\left\{-\frac{y^2}{2\theta^2}\right\}, \quad y > 0,$$

where $\theta > 0$ is a parameter. It turns out that

$$\mathbb{E}(Y) = \sqrt{\frac{\pi}{2}}\,\theta \quad \text{and} \quad \mathbb{E}(Y^2) = 2\theta^2.$$

(a) Determine the Fisher information $I(\theta)$ for the Rayleigh(θ) distribution.

Suppose that $Y_1, Y_2, \ldots, Y_n$ are independent and identically distributed Rayleigh(θ) random variables.

(b) Compute $\hat{\theta}_{\text{MOM}}$, the method of moments estimator of θ.

(c) Compute the variance of $\hat{\theta}_{\text{MOM}}$.

(d) Determine the maximum likelihood estimator of θ. You do *not* need to verify that your critical point is actually a maximum.

(e) Suppose that a random sample of size $n = 100$ is observed. From this random sample,

$$\sum_{i=1}^{100} y_i^2 = 80000$$

is calculated. Construct an approximate 95% confidence interval for θ.

3. Let X and Y be uncorrelated random variables. Suppose that X has mean 2θ and variance 4. Suppose that Y has mean θ and variance 2. The parameter θ is unknown.

(a) Compute the bias and mean squared error for each of the following estimators of θ:

$$\hat{\theta}_1 = \frac{1}{4}X + \frac{1}{2}Y \quad \text{and} \quad \hat{\theta}_2 = X - Y.$$

(b) Compute the efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_2$. Which estimator do you prefer? Why?

(c) Verify that the estimator

$$\hat{\theta}_c = \frac{c}{2}X + (1-c)Y$$

is unbiased. Find the value of c which minimizes $\text{Var}(\hat{\theta}_c)$.

(continued)

4. Suppose that $Y_1, Y_2, \ldots, Y_n$ are independent and identically distributed random variables, each having density function

$$f_Y(y|\theta) = \frac{\theta^{-252}}{251!}\, y^{251}\, e^{-y/\theta}, \quad y > 0,$$

where $\theta > 0$ is a parameter. It is known that if $Y \sim f_Y(y|\theta)$, then $\mathbb{E}(Y) = 252\,\theta$ and $\mathrm{Var}(Y) = 252\,\theta^2$. Let $\hat{\theta} = \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$.

(a) Find a function of $\hat{\theta}$ which is an unbiased estimator of θ. Call it $\hat{\theta}_A$

(b) Compute the Fisher information in a single observation from this density.

(c) Carefully explain why $\hat{\theta}_A$ must be the minimum variance unbiased estimator of θ.

5. A random variable Y has the Laplace distribution if its density function

$$f_Y(y|\lambda,\theta) = \frac{\lambda}{2}\, e^{-\lambda|y-\theta|}, \quad 0 < y < \infty,$$

with parameters $\lambda > 0$ and $-\infty < \theta < \infty$. The expectation and variance of Y are $\mathbb{E}(Y) = \theta$ and $\mathrm{Var}(Y) = 2\lambda^{-2}$. Find the method of moments estimators of both λ and θ.

6. Suppose that the random variables $Y_1, Y_2, \ldots, Y_{10}$ are independent and identically distributed Uniform$(0,\theta)$ random variables.

(a) Find $\hat{\theta}_{\text{MOM}}$, the method of moments estimator of θ. Show that $\hat{\theta}_{\text{MOM}}$ is an unbiased estimator of θ.

(b) From previous work, we know that $\hat{\theta}_{\text{MLE}}$, the maximum likelihood estimator of θ, is

$$\hat{\theta}_{\text{MLE}} = \max\{Y_1, \ldots, Y_{10}\}.$$

Find an unbiased estimator which is a function of the MLE. Call it $\hat{\theta}_B$.

(c) Find the efficiency of $\hat{\theta}_{\text{MOM}}$ relative to $\hat{\theta}_B$. Which estimator do you prefer? Why?

(d) Find the Fisher Information in a single observation from the Uniform$(0,\theta)$ distribution.

(e) Based on your previous answers, what can you say about the minimum variance unbiased estimator of θ? Why is the Cramer-Rao inequality of no use in this case?

7. Each of the following require short written answers. Very clear and brief solutions are required for full credit.

(a) Describe how to interpret a 93% confidence interval.

(b) Why do you think it is desirable to find a *minimum variance unbiased estimator*?

(continued)

8. Suppose that $Y_1, \ldots, Y_n$ are independent and identically distributed $\mathcal{N}(\mu, \sigma^2)$ random variables. It is known that $\sigma^2 = 9$, but μ is unknown. You want to test $H_0 : \mu = 0$ against $H_A : \mu > 0$ by rejecting H_0 when $\overline{Y} > c$.

(a) Find c so that this test has significance level $\alpha = 0.05$. (Of course, c will depend on the sample size n.)

(b) Using the test determined in **(a)**, find the power of the test when $\mu = 1$ and $n = 36$.

(c) Again, using the test determined in **(a)**, show that the power when $\mu = 1$ will increase as the sample size increases. Regardless of whether or not you can show this result, does it make sense intuitively? Comment **very** briefly.

9. Suppose that $Y_1, \ldots, Y_n$ are independent and identically distributed $\mathcal{N}(\mu, \sigma^2)$ random variables. It is known that $\sigma^2 = 4$, but μ is unknown. You want to test $H_0 : \mu = 0$ against $H_A : \mu > 0$ by rejecting H_0 when $\overline{Y} > 3.92/\sqrt{n}$.

(a) Verify that this test has significance level $\alpha = 0.025$.

(b) Say you are planning an experiment for which the data will be analyzed by this hypothesis test. How large a sample should you collect if you would like the test to have power 0.9 when $\mu = 0.5$?

10. John carries out hypothesis tests using 0.01 as the significance level, while George uses 0.05 as his significance level. Ringo carries out an experiment to compare two hypotheses, and computes the p-value. But all Ringo tells John and George is that the p-value is smaller than 0.03. Can John make his decision to accept or reject the null? If so, what is the decision? What about George? Justify your answers **very** briefly.

11. In the context of Stat 252, define what is meant by a *significance level* α *hypothesis test.*

12. Let $Y_1, \ldots, Y_{100}$ be independent and identically distributed $\mathcal{N}(\mu, \sigma^2)$ random variables. It is known that $\sigma^2 = 25$, but μ is unknown. Consider testing $H_0 : \mu = 0$ against $H_A : \mu < 0$ by rejecting H_0 for small values of the sample mean $\overline{Y}$. The user of the test would like to treat the hypotheses somewhat symmetrically. In particular, she wants the probability of Type I error to be equal to the probability of Type II error when $\mu = -1/2$. Construct such a test, and give its significance level.

13.

(a) Suppose you have been asked to analyze a data set, and you are planning to use a particular estimator to estimate a parameter. Explain why it is important to determine (either exactly or approximately) the sampling distribution of this estimator.

(b) Suppose that you want to compare two simple hypotheses in light of a data set. You use a test with significance level $\alpha = 0.05$. The data yield a test statistic for which you accept the null hypothesis (or, if you prefer, you do not reject the null hypothesis). You would like to use this result to argue that the null hypothesis is likely to be true. What additional information about the test might you want first? Explain.

(continued)

14.

(a) Assume that the outcome of an experiment is a single random variable X. An 80% confidence interval for a parameter θ has the form $(X-1, X+2)$. From this, determine a rejection rule for testing $H_0 : \theta = 5$ against $H_A : \theta \neq 5$ at significance level 0.2.

(b) Let Y_1 and Y_2 be independent Uniform$(0, \theta)$ random variables. Consider testing $H_0 : \theta = 1$ against $H_A : \theta > 1$ by rejecting H_0 when $\max\{Y_1, Y_2\} > c$. Find c so that this test has significance level $\frac{19}{100}$. What is the power of this test (as a function of θ)?

15. An electrical circuit consists of three batteries connected in series to a lightbulb. We model the battery lifetimes X_1, X_2, X_3 as independent and identically distributed Exponential(λ) random variables (where $\lambda > 0$ is a parameter). Our experiment to measure the lifetime of the lightbulb Y is stopped when any one of the batteries fails. Hence, the only random variable we observe is $Y = \min\{X_1, X_2, X_3\}$.

(a) Determine the distribution of the random variable Y.

(b) Compute $\hat{\lambda}_{\text{MLE}}$, the maximum likelihood estimator of λ. *Hint:* Use the result of **(a)**.

(c) Determine the mean square error of $\hat{\lambda}_{\text{MLE}}$.

(d) Use the Cramer-Rao lower bound to prove that $\hat{\lambda}_{\text{MLE}}$ is the minimum variance unbiased estimator of λ.