Statistics 252–Mathematical Statistics
Winter 2006 (200610)
Final Exam Solutions

Instructor: Michael Kozdron

1. (a) To find the method of moments estimator, we equate the first population moment with the first sample moment. Since $\mathbb{E}(Y_i) = 3\theta$, we conclude

$$\hat{\theta}_{\text{MOM}} = \frac{\overline{Y}}{3}.$$

1. (b) By definition, the likelihood function $L(\theta)$ is given by

$$L(\theta) = \prod_{i=1}^{n} f(y_i|\theta) = \prod_{i=1}^{n} \frac{y_i^2}{2\theta^3} \exp(-y_i/\theta) = 2^{-n}\theta^{-3n} \left(\prod_{i=1}^{n} y_i\right)^2 \exp\left(-\frac{1}{\theta}\sum_{i=1}^{n} y_i\right).$$

1. (c) If $L(\theta)$ is as in **(b)**, then the log-likelihood function is

$$\ell(\theta) = \log L(\theta) = -n\log 2 - 3n\log\theta + 2\sum_{i=1}^{n}\log y_i - \frac{1}{\theta}\sum_{i=1}^{n} y_i.$$

Taking derivatives gives

$$\ell'(\theta) = -\frac{3n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} y_i$$

and

$$\ell''(\theta) = \frac{3n}{\theta^2} - \frac{2}{\theta^3}\sum_{i=1}^{n} y_i.$$

Setting $\ell'(\theta) = 0$ gives

$$\theta = \frac{1}{3n}\sum_{i=1}^{n} y_i = \frac{\overline{y}}{3}.$$

Since

$$\ell''(\overline{y}/3) = -\frac{27n}{\overline{y}^2} < 0$$

we conclude that

$$\hat{\theta}_{\text{MLE}} = \frac{\overline{Y}}{3}.$$

1. (d) Since $\mathbb{E}(Y_i) = 3\theta$, we find

$$\mathbb{E}(\hat{\theta}_{\text{MLE}}) = \mathbb{E}(\hat{\theta}_{\text{MOM}}) = \mathbb{E}\left(\frac{\overline{Y}}{3}\right) = \frac{\mathbb{E}(Y_1)}{3} = \frac{3\theta}{3} = \theta.$$

In other words, both $\hat{\theta}_{\text{MLE}}$ and $\hat{\theta}_{\text{MOM}}$ are unbiased estimators of θ.

1. (e) Since $\hat{\theta}_{\text{MLE}} = \hat{\theta}_{\text{MOM}}$, we conclude that $\text{RE}(\hat{\theta}_{\text{MLE}}, \hat{\theta}_{\text{MOM}}) = 1$. In other words, we have no preference, since both estimators are the same!

1. (f) If

$$h(y_1, \ldots, y_n) = 2^{-n} \left(\prod_{i=1}^{n} y_i \right)^2 \quad \text{and} \quad g(u, \theta) = \theta^{-3n} \exp\left(-u/\theta\right)$$

then since $L(\theta) = h(y_1, \ldots, y_n) \cdot g(u, \theta)$ we conclude from the Factorization Theorem that

$$U = \sum_{i=1}^{n} Y_i$$

is sufficient for the estimation of θ.

1. (g) Recall that any 1–1 function of a sufficient statistic is also sufficient. Suppose that

$$T(U) = \frac{U}{3n}.$$

Since T is a 1–1 function, we conclude that

$$T(U) = \frac{1}{3n} \sum_{i=1}^{n} Y_i = \frac{\overline{Y}}{3} = \hat{\theta}_{\text{MLE}}$$

is also a sufficient statistic for the estimation of θ.

1. (h) Since

$$\log f(y|\theta) = -\log 2 - 3 \log \theta + 2 \log y - \frac{y}{\theta}$$

we compute

$$\frac{\partial}{\partial \theta} \log f(y|\theta) = -\frac{3}{\theta} + \frac{y}{\theta^2} \quad \text{and} \quad \frac{\partial^2}{\partial \theta^2} \log f(y|\theta) = \frac{3}{\theta^2} - \frac{2y}{\theta^3}.$$

Therefore,

$$I(\theta) = -\mathbb{E}\left(\frac{\partial^2}{\partial \theta^2} \log f(Y|\theta) \right) = \frac{2}{\theta^3} \mathbb{E}(Y) - \frac{3}{\theta^2} = \frac{2 \cdot 3\theta}{\theta^3} - \frac{3}{\theta^2} = \frac{3}{\theta^2}.$$

1. (i) Recall that the Cramer-Rao lower bound tells us that if $\hat{\theta}$ is an unbiased estimator of θ, then

$$\text{Var}(\hat{\theta}) \geq \frac{1}{nI(\theta)}.$$

Since $\hat{\theta}_{\text{MLE}}$ is unbiased (as shown in **(d)**) with

$$\text{Var}\left(\hat{\theta}_{\text{MLE}}\right) = \text{Var}\left(\frac{\overline{Y}}{3}\right) = \frac{1}{9n}\text{Var}(Y_1) = \frac{3\theta^2}{9n} = \frac{\theta^2}{3n},$$

and since (using **(h)**)

$$\frac{1}{nI(\theta)} = \frac{1}{3n/\theta^2} = \frac{\theta^2}{3n}$$

we see that the lower bound of the Cramer-Rao inequality is attained, and so we deduce that $\hat{\theta}_{\text{MLE}}$ is the minimum variance unbiased estimator of θ.

1. (j) Recall that a $(1-\alpha)$ confidence interval based on the Fisher information and the maximum likelihood estimator is given by

$$\hat{\theta}_{\text{MLE}} \pm z_{\alpha/2} \cdot \frac{1}{\sqrt{nI(\hat{\theta}_{\text{MLE}})}}.$$

Using the results of **(c)** and **(h)** we conclude that the required confidence interval is

$$\frac{\overline{Y}}{3} \pm \frac{1.65 \cdot \overline{Y}}{3\sqrt{3n}}$$

since $z_{0.05} = 1.65$ from Table 2.

1. (k) The generalized likelihood ratio test statistic is given by

$$\Lambda(y) = \frac{L(\theta_0)}{L(\hat{\theta}_{\text{MLE}})} = \frac{2^{-n}\theta_0^{-3n}\left(\prod_{i=1}^n y_i\right)^2 \exp\left(-\frac{1}{\theta_0}\sum_{i=1}^n y_i\right)}{2^{-n}\hat{\theta}_{\text{MLE}}^{-3n}\left(\prod_{i=1}^n y_i\right)^2 \exp\left(-\frac{1}{\hat{\theta}_{\text{MLE}}}\sum_{i=1}^n y_i\right)} = \left(\frac{\overline{Y}}{3\theta_0}\right)^{3n} \exp\left(3n - \frac{n\overline{Y}}{\theta_0}\right).$$

1. (l) Using the data given, we find that a 90% confidence interval for θ is

$$\left(\frac{4}{3} - \frac{1.65 \cdot 4}{27}, \frac{4}{3} + \frac{1.65 \cdot 4}{27}\right) \quad \text{or, approximately,} \quad (1.09, 1.58).$$

Since $\theta_0 = 1$ does not lie in the confidence interval, we can reject the hypothesis $H_0 : \theta = 1$ in favour of $H_A : \theta \neq 1$ at the significance level $\alpha = 0.10$.

1. (m) Using the data given, we compute

$$-2\log\Lambda = -2\left[3n\log\overline{y} - 3n\log(3\theta_0) + 3n - \frac{n\overline{y}}{\theta_0}\right] = -2\left[81\log 4 - 81\log 3 + 81 - 108\right] \approx 7.3955.$$

Since $-2\log\Lambda \sim \chi^2_1$ (approximately), we find from Table 6 that $\chi^2_{1,0.1} = 2.70554$ and so we reject H_0 in favour of H_A at significance level $\alpha = 0.1$.

2. (a) Since $X_i \sim \mathcal{N}(200, 10^2)$ we know that $\overline{X} \sim \mathcal{N}(200, 10^2/4)$. Therefore, the probability that the total resistance of these four resistors does not exceed 840 ohms is given by

$$P(X_1 + X_2 + X_3 + X_4 \leq 840) = P(\overline{X} \leq 210) = P\left(\frac{\overline{X} - 200}{10/\sqrt{4}} \leq \frac{210 - 200}{10/\sqrt{4}}\right) = P(Z \leq 2)$$

where $Z \sim \mathcal{N}(0,1)$. From Table 2 we find $P(Z \leq 2) = 1 - 0.0228 = 0.9772$ and so the required probability is

$$P(X_1 + X_2 + X_3 + X_4 \leq 840) = 0.9772.$$

2. (b) The probability that the circuit does not fail within the first 15 hours of operation is given by

$$P(\min\{Y_1, Y_2, Y_3, Y_4\}) > 15) = P(Y_1 > 15, Y_2 > 15, Y_3 > 15, Y_4 > 15) = [P(Y_1 > 15)]^4$$

since the Y_i are assumed to be independent and identically distributed random variables. We find

$$P(Y_1 > 15) = \int_{15}^{\infty} \frac{1}{12} e^{-y/12}\, dy = -e^{-y/12}\Big|_{15}^{\infty} = e^{-15/12}.$$

(continued)

Therefore, the required probability is

$$P(\min\{Y_1, Y_2, Y_3, Y_4\}) > 15) = \left[e^{-15/12}\right]^4 = e^{-5}.$$

3. (a) Recall that the significance level of an hypothesis test is the probability of a Type I error. That is,

$$\alpha = P(\text{Type I error}) = P_{H_0}(\text{reject } H_0) = P(\overline{Y} > 1.55 \,|\, \mu = 0).$$

Since $Y_i \sim \mathcal{N}(\mu, 4)$ we know that $\overline{Y} \sim \mathcal{N}(\mu, 4/9)$. Therefore,

$$P(\overline{Y} > 1.55 \,|\, \mu = 0) = P\left(\frac{\overline{Y} - 0}{2/3} > \frac{1.55 - 0}{2/3}\right) = P(Z > 2.325) \approx 0.01$$

where $Z \sim \mathcal{N}(0, 1)$ and the last step followed from Table 2. In other words, this test has significance level $\alpha = 0.01$.

3. (b) Recall that the power of an hypothesis test is $P_{H_A}(\text{reject } H_0)$. If the alternative is $H_A : \mu = 1$, then this test has power

$$P_{H_A}(\text{reject } H_0) = P(\overline{Y} > 1.55 \,|\, \mu = 1) = P\left(\frac{\overline{Y} - 1}{2/3} > \frac{1.55 - 1}{2/3}\right) = P(Z > 0.825) \approx 0.2033$$

where $Z \sim \mathcal{N}(0, 1)$ and the last step followed from Table 2. In other words, this test has power (approximately) 0.2033 when $\mu = 1$.

4. (a) To find the method of moments estimator, we equate the first population moment with the first sample moment. If $Y \sim \text{Unif}(0, \theta)$, we find

$$\mathbb{E}(Y) = \int_0^\theta \frac{y}{\theta}\, dy = \frac{\theta}{2}$$

and so we conclude

$$\hat{\theta}_{\text{MOM}} = 2\overline{Y}.$$

4. (b) Recall that the mean-squared error of an estimator $\hat{\theta}$ is given by

$$\text{MSE}(\hat{\theta}) = \text{Var}(\hat{\theta}) + \text{bias}(\hat{\theta})^2.$$

We compute

$$\text{bias}(\hat{\theta}_{\text{MOM}}) = \mathbb{E}(\hat{\theta}_{\text{MOM}}) - \theta = 2\mathbb{E}(\overline{Y}) - \theta = 2\mathbb{E}(Y_1) - \theta = 2 \cdot \frac{\theta}{2} - \theta = 0$$

and

$$\text{Var}(\hat{\theta}_{\text{MOM}}) = \text{Var}(2\overline{Y}) = \frac{4}{n}\,\text{Var}(Y_1).$$

Since $Y_1 \sim \text{Unif}(0, \theta)$, we find

$$\text{Var}(Y_1) = \mathbb{E}(Y_1^2) - \mathbb{E}(Y_1)^2 = \int_0^\theta \frac{y^2}{\theta}\, dy - \frac{\theta^2}{4} = \frac{\theta^2}{3} - \frac{\theta^2}{4} = \frac{\theta^2}{12}.$$

Combining the above, we conclude

$$\text{MSE}(\hat{\theta}_{\text{MOM}}) = \text{Var}(\hat{\theta}_{\text{MOM}}) + \text{bias}(\hat{\theta}_{\text{MOM}})^2 = \frac{4}{n} \cdot \frac{\theta^2}{12} + 0 = \frac{\theta^2}{3n}.$$

4. **(c)** We begin by finding the density function of $\hat{\theta}_{\text{MLE}}$. Since $Y_1, Y_2, \dots, Y_n$ are independent and identically distributed we conclude that for $0 \le t \le \theta$,

$$P(\hat{\theta}_{\text{MLE}} \le t) = P(\max\{Y_1, \dots, Y_n\} \le t) = [P(Y_1 \le t)]^n = \left[\int_0^t \frac{1}{\theta}\, dy\right]^n = \frac{t^n}{\theta^n}.$$

Therefore, the density of $\hat{\theta}_{\text{MLE}}$ is

$$f(t) = \frac{nt^{n-1}}{\theta^n}, \quad 0 \le t \le \theta.$$

We now compute

$$\mathbb{E}(\hat{\theta}_{\text{MLE}}) = \int_0^\theta \frac{nt^n}{\theta^n}\, dt = \frac{n}{n+1}\theta \quad \text{and} \quad \mathbb{E}(\hat{\theta}^2_{\text{MLE}}) = \int_0^\theta \frac{nt^{n+1}}{\theta^n}\, dt = \frac{n}{n+2}\theta^2.$$

This gives

$$\text{bias}(\hat{\theta}_{\text{MLE}}) = \mathbb{E}(\hat{\theta}_{\text{MLE}}) - \theta = \frac{n}{n+1}\theta - \theta = -\frac{\theta}{n+1}$$

and

$$\text{Var}(\hat{\theta}_{\text{MLE}}) = \mathbb{E}(\hat{\theta}^2_{\text{MLE}}) - \mathbb{E}(\hat{\theta}_{\text{MLE}})^2 = \frac{n}{n+2}\theta^2 - \frac{n^2}{(n+1)^2}\theta^2,$$

and so

$$\text{MSE}(\hat{\theta}_{\text{MLE}}) = \left(\frac{n}{n+2} - \frac{n^2}{(n+1)^2} + \frac{1}{(n+1)^2}\right)\theta^2 = \left(\frac{n}{n+2} - \frac{n-1}{n+1}\right)\theta^2 = \frac{2\theta^2}{(n+2)(n+1)}.$$

5. The goal of this problem is to test whether or not you understand the pragmatic implications of Type I and Type II errors (and can translate them into everyday language). As always, a Type I error occurs when the null hypothesis is falsely rejected, and a Type II error occurs when the null hypothesis is falsely accepted. For hypothesis test A, the consequence of a Type I error is that Michael is found not guilty and does not go to prison when, in fact, he is guilty. The consequence of a Type II error is that Michael is found guilty and is sent to prison when, in fact, he is innocent. For hypothesis test B, these errors are switched. That is, the consequence of a Type I error is that Michael is found guilty and sent to prison when, in fact, he is innocent. The consequence of a Type II error is that Michael is found not guilty and does not go to prison when, in fact, he is guilty. In order to receive full credit for this problem, one needs to correctly identify the practical significance of the two errors (i.e., Michael is innocent but is sent to jail, or Michael is guilty but is set free), and one needs to provide an opinion as to which error is believed to be more serious.

6. The correct answers, in order, are: True, False, False, False, False, False, True, False.

7. Suppose that the 90% confidence interval for θ based on Y is given by $(Y - U, Y + L)$, where U and L are to be determined. Using the confidence interval–hypothesis test duality, we know that the rejection region of the significance level $\alpha = 0.10$ test $H_0 : \theta = 3$ vs. $H_A : \theta \ne 3$ is

$$RR = \{3 \notin (Y - U, Y + L)\} = \{3 < Y - U \;\text{ or }\; 3 > Y + L\} = \{Y > 3 + U \;\text{ or }\; Y < 3 - L\}.$$

Since the rejection region given is $\{Y > 7 \;\text{ or }\; Y < 2\}$ we conclude that $3 + U = 7$ and $3 - L = 2$. Therefore, $L = 1$ and $U = 4$, and so the required confidence interval is

$$(Y - 4, Y + 1).$$

8. Suppose that $U = Y/\theta$. The distribution function of U is

$$P(U \leq u) = P(Y \leq u\theta) = \int_0^{u\theta} f(y|\theta)\, dy = \int_0^{u\theta} \frac{2y}{\theta} \exp(-y^2/\theta^2)\, dy.$$

Make the substitution $v = y^2/\theta^2$ so that $dv = 2y/\theta^2 dy$ and therefore,

$$P(U \leq u) = \int_0^{u^2} e^{-v}\, dv = 1 - e^{-u^2}$$

so that the density function of U is

$$f_U(u) = 2ue^{-u^2}, \quad u > 0.$$

Suppose that a and b are chosen so that

$$\alpha_1 = \int_0^a f_U(u)\, du \quad \text{and} \quad \alpha_2 = \int_b^\infty f_U(u)\, du.$$

With this choice of a and b, we know that

$$1 - (\alpha_1 - \alpha_2) = P(a \leq U \leq b) = P\left(\frac{Y}{b} \leq \theta \leq \frac{Y}{a}\right)$$

which will give us the required confidence interval. Since

$$\int_0^a f_U(u)\, du = \int_0^a 2ue^{-u^2}\, du = 1 - e^{-a^2}$$

we find $a = \sqrt{-\log(1 - \alpha_1)}$ and since

$$\int_b^\infty f_U(u)\, du = \int_b^\infty 2ue^{-u^2}\, du = e^{-b^2}$$

we find $b = \sqrt{-\log(\alpha_2)}$. In conclusion,

$$\left(\frac{Y}{\sqrt{-\log(\alpha_2)}}, \frac{Y}{\sqrt{-\log(1 - \alpha_1)}}\right)$$

is a confidence interval for θ with coverage probability $1 - (\alpha_1 - \alpha_2)$.