Stat 252.01 Winter 2007
Assignment #10 Solutions

1. (a) To find the method of moments estimator of θ we must solve the equation $\mathbb{E}(Y) = \overline{Y}$ for θ. Since $\mathbb{E}(\overline{Y}) = \theta$, we conclude

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

1. (c) The likelihood function is

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

so that the log-likelihood function is

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

Since $\ell'(\theta) = 0$ implies

$$0 = \frac{2n}{\theta} - \sum_{i=1}^{n} \frac{1}{y_i}$$

so that

$$\theta = \frac{2n}{\sum_{i=1}^{n} \frac{1}{y_i}},$$

and since

$$\ell''(\theta) = -\frac{2n}{\theta^2} < 0,$$

we conclude from the second derivative test that

$$\hat{\theta}_{\text{MLE}} = \frac{2n}{\displaystyle\sum_{i=1}^{n} \frac{1}{Y_i}}.$$

1. (d) Since

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

we see that if we let $u = \sum_{i=1}^{n} \frac{1}{y_i}$, $g(u, \theta) = \theta^{2n} e^{-\theta u}$, and $h(y_1, \ldots, y_n) = \left(\prod_{i=1}^{n} y_i\right)^{-3}$, then

$$L(\theta) = g(u, \theta) \cdot h(y_1, \ldots, y_n)$$

so from the Factorization Theorem we conclude that

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

is sufficient for the estimation of θ.

1. (e) If $T(U) = \frac{2n}{U}$, then since T is a one-to-one function and since any one-to-one function of a sufficient statistic is also sufficient, we conclude that

$$T\left(\sum_{i=1}^{n} \frac{1}{Y_i}\right) = \frac{2n}{\displaystyle\sum_{i=1}^{n} \frac{1}{Y_i}} = \hat{\theta}_{\text{MLE}}$$

is also sufficient.

1. (f) Since

$$\log f_Y(y|\theta) = 2\log\theta - 3\log y - \frac{\theta}{y},$$

we find

$$\frac{\partial}{\partial\theta}\log f_Y(y|\theta) = \frac{2}{\theta} - \frac{1}{y} \quad \text{and} \quad \frac{\partial^2}{\partial\theta^2}\log f_Y(y|\theta) = -\frac{2}{\theta^2}$$

Thus,

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

1. (g) An approximate 90% confidence interval for θ is given by

$$\left[\hat{\theta}_{\text{MLE}} - 1.645\,\frac{1}{\sqrt{n\,I(\hat{\theta}_{\text{MLE}})}}\,,\,\hat{\theta}_{\text{MLE}} + 1.645\,\frac{1}{\sqrt{n\,I(\hat{\theta}_{\text{MLE}})}}\right]$$

or

$$\left[\frac{2n}{\sum_{i=1}^{n}\frac{1}{Y_i}} - 1.645\frac{\sqrt{2n}}{\sum_{i=1}^{n}\frac{1}{Y_i}}\,,\,\frac{2n}{\sum_{i=1}^{n}\frac{1}{Y_i}} + 1.645\frac{\sqrt{2n}}{\sum_{i=1}^{n}\frac{1}{Y_i}}\right].$$

1. (i) If $n = 8$ observations produce $\sum_{i=1}^{8}\frac{1}{y_i} = 10$, then based on this data, an approximate 90% confidence interval for θ is

$$[1.6 - 0.66, 1.6 + 0.66] \quad \text{or} \quad [0.94, 2.26].$$

Since $\theta_0 = 1$ falls in this interval, we conclude from the confidence interval–hypothesis test duality that we do not reject $H_0 : \theta = 1$ in favour of $H_A : \theta \neq 1$ at significance level $\alpha = 0.10$.

2. (a) Since

$$\log f_Y(y|\theta) = \log y - 2\log\theta - \frac{y^2}{2\theta^2},$$

we find

$$\frac{\partial^2}{\partial\theta^2}\log f_Y(y|\theta) = \frac{2}{\theta^2} - \frac{3y^2}{\theta^4}.$$

Thus,

$$I(\theta) = -\mathbb{E}\left(\frac{\partial^2}{\partial\theta^2}\log f(Y|\theta)\right) = \frac{3\mathbb{E}(Y^2)}{\theta^4} - \frac{2}{\theta^2} = \frac{4}{\theta^2}.$$

2. (b) To find $\hat{\theta}_{\text{MOM}}$ we solve the equation $\mathbb{E}(Y) = \overline{Y}$ for θ. This implies

$$\hat{\theta}_{\text{MOM}} = \sqrt{\frac{2}{\pi}}\,\overline{Y}.$$

2. (c)

$$\mathrm{Var}(\hat{\theta}_{\mathrm{MOM}}) = \frac{2}{\pi}\ \mathrm{Var}(\overline{Y}) = \frac{2}{n\pi}\ \mathrm{Var}(Y_1) = \frac{2}{n\pi}\ \left(\mathbb{E}(Y_1^2) - [\mathbb{E}(Y)]^2\right) = \frac{2}{n\pi}\ \left(2 - \frac{\pi}{2}\right)\theta^2$$

$$= \left(\frac{4-\pi}{n\pi}\right)\ \theta^2$$

2. (d) The likelihood function is

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

so that the log-likelihood function is

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

Hence $\ell'(\theta) = 0$ implies

$$0 = -\frac{2n}{\theta} + \frac{1}{\theta^3}\sum_{i=1}^{n} y_i^2$$

so that

$$\hat{\theta}_{\mathrm{MLE}} = \sqrt{\frac{1}{2n}\sum_{i=1}^{n} Y_i^2}.$$

2. (e) An approximate 95% confidence interval for θ is given by

$$\left[\hat{\theta}_{\mathrm{MLE}} - 1.96\ \frac{1}{\sqrt{n\, I(\hat{\theta}_{\mathrm{MLE}})}}\ ,\ \hat{\theta}_{\mathrm{MLE}} + 1.96\ \frac{1}{\sqrt{n\, I(\hat{\theta}_{\mathrm{MLE}})}}\right].$$

Since $n = 100$ and $\sum_{i=1}^{100} y_i^2 = 80000$, we conclude that

$$\hat{\theta}_{\mathrm{MLE}} = \sqrt{\frac{80000}{200}} = \sqrt{400} = 20$$

and

$$I(\hat{\theta}_{\mathrm{MLE}}) = \frac{4}{\hat{\theta}_{\mathrm{MLE}}^2} = \frac{4}{400} = \frac{1}{100}.$$

Hence, an approximate 95% confidence interval for θ is

$$[20 - 1.96, 20 + 1.96] \quad \text{or} \quad [18.04, 21.96].$$

3. (a) Since

$$\mathbb{E}(\hat{\theta}_1) = \frac{1}{4}\mathbb{E}(X) + \frac{1}{2}\mathbb{E}(Y) = \frac{2\theta}{4} + \frac{\theta}{2} = \theta$$

and

$$\mathbb{E}(\hat{\theta}_2) = \mathbb{E}(X) - \mathbb{E}(Y) = 2\theta - \theta = \theta$$

we conclude that $B(\hat{\theta}_1) = B(\hat{\theta}_2) = 0$. Thus,

$$\mathrm{MSE}(\hat{\theta}_1) = \mathrm{Var}(\hat{\theta}_1) = \frac{1}{16}\ \mathrm{Var}(X) + \frac{1}{4}\ \mathrm{Var}(Y) = \frac{3}{4}$$

and

$$\mathrm{MSE}(\hat\theta_2) = \mathrm{Var}(\hat\theta_2) = \mathrm{Var}(X) + \mathrm{Var}(Y) = 6.$$

3. (b) We find

$$\mathrm{eff}(\hat\theta_1, \hat\theta_2) = \frac{\mathrm{Var}(\hat\theta_2)}{\mathrm{Var}(\hat\theta_1)} = 8.$$

Since both $\hat\theta_1$ and $\hat\theta_2$ are unbiased, the one with the smaller variance is preferrable, namely $\hat\theta_1$.

3. (c) Since

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

we see $\hat\theta_c$ is unbiased. Since

$$\mathrm{Var}(\hat\theta_c) = \frac{c^2}{4}\,\mathrm{Var}(X) + (1-c)^2\,\mathrm{Var}(Y) = c^2 + 2(1-c)^2 = 3c^2 - 4c + 2$$

the value that minimizes $\mathrm{Var}(\hat\theta_c)$ is the same value that minimizes the polynomial $g(c) = 3c^2 - 4c + 2$. Since $g'(c) = 0$ implies $c = 2/3$, and since $g''(2/3) > 0$, the minimal value of c is $2/3$.

4. (a) We find that

$$\mathbb{E}(\overline{Y}) = \mathbb{E}(Y_1) = 252\,\theta.$$

Thus, if

$$\hat\theta_A = \frac{\overline{Y}}{252} = \frac{1}{252n}\sum_{i=1}^{n} Y_i$$

then $\hat\theta_A$ is an unbiased estimator of θ.

4. (b) Since

$$\log f_Y(y|\theta) = -252\log\theta - \log(251!) + 251\log y - \frac{y}{\theta},$$

we find

$$\frac{\partial^2}{\partial\theta^2}\log f_Y(y|\theta) = \frac{252}{\theta^2} - \frac{2y}{\theta^3}.$$

Thus,

$$I(\theta) = -\mathbb{E}\left(\frac{\partial^2}{\partial\theta^2}\log f_Y(Y|\theta)\right) = -\frac{252}{\theta^2} + \frac{2\mathbb{E}(Y)}{\theta^3} = \frac{252}{\theta^2}.$$

4. (c) The Cramer-Rao inequality tells us that any unbiased estimator $\hat\theta$ of θ must satisfy

$$\mathrm{Var}(\hat\theta) \geq \frac{1}{nI(\theta)} = \frac{\theta^2}{252n}.$$

Since

$$\mathrm{Var}(\hat\theta_A) = \frac{1}{252^2 n}\,\mathrm{Var}\,Y_1 = \frac{1}{252^2}\cdot(252\,\theta^2) = \frac{\theta^2}{252n},$$

we have found an unbiased estimator whose variance attains the lower bound of the Cramer-Rao inequality. Hence, $\hat\theta_A$ must be the MVUE of θ.

5. To find the method of moments estimators for λ and θ, we must solve the system of equations

$$\mathbb{E}(Y) = \overline{Y} \quad \text{and} \quad \mathbb{E}(Y^2) = \frac{1}{n}\sum_{i=1}^{n} Y_i^2$$

Thus, some trivial algebra gives

$$\hat{\theta}_{\text{MOM}} = \overline{Y} \quad \text{and} \quad \hat{\lambda}_{\text{MOM}} = \sqrt{\frac{2n}{\sum_{i=1}^{n} Y_i^2}}.$$

6. (a) If $Y \sim \text{Uniform}(0,\theta)$, then $\mathbb{E}(Y) = \theta/2$ and $\text{Var}(Y) = \theta^2/12$. Thus,

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

Since

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

we conclude that $\hat{\theta}_{\text{MOM}}$ is an unbiased estimator of θ.

6. (b) In order to find $\mathbb{E}(\hat{\theta}_{\text{MLE}})$ we must find the density function of $\hat{\theta}_{\text{MLE}}$. Now,

$$P(\hat{\theta}_{\text{MLE}} \le t) = \left[\int_0^t \theta^{-1}\, dy\right]^{10} = \frac{t^{10}}{\theta^{10}}, \quad 0 \le t \le \theta,$$

so that $f(t) = 10\,\theta^{-10}\,t^9$, $0 \le t \le \theta$. Thus,

$$\mathbb{E}(\hat{\theta}_{\text{MLE}}) = \int_0^{\theta} 10\,\theta^{-10}\,t^{10}\, dt = \frac{10}{11}\,\theta.$$

An unbiased estimator of θ which is a function of the MLE is given by

$$\hat{\theta}_B = \frac{11}{10}\max\{Y_1,\ldots,Y_{10}\}.$$

Also, note that

$$\mathbb{E}(\hat{\theta}^2_{\text{MLE}}) = \int_0^{\theta} 10\,\theta^{-10}\,t^{11}\, dt = \frac{10}{12}\,\theta^2.$$

6. (c) From **(a)**, we conclude

$$\text{Var}(\hat{\theta}_{\text{MOM}}) = 4\,\text{Var}(\overline{Y}) = \frac{4}{10}\,\text{Var}(Y_1) = \frac{4\theta^2}{10\cdot 12} = \frac{\theta^2}{30}.$$

From **(b)**, we conclude

$$\text{Var}(\hat{\theta}_B) = \frac{121}{100}\,\text{Var}(\max\{Y_1,\ldots,Y_{10}\}) = \frac{121}{100}\left(\frac{10}{12} - \frac{100}{121}\right)\theta^2 = \frac{\theta^2}{120}.$$

Thus,

$$\text{eff}(\hat{\theta}_{\text{MOM}}, \hat{\theta}_B) = \frac{\text{Var}(\hat{\theta}_B)}{\text{Var}(\hat{\theta}_{\text{MOM}})} = \frac{1}{4}.$$

Since both $\hat{\theta}_{\text{MOM}}$ and $\hat{\theta}_B$ are unbiased, the one with the smaller variance is preferrable, namely $\hat{\theta}_B$.

6. (d) Since

$$\log f_Y(y|\theta) = -\log \theta,$$

we find

$$\frac{\partial^2}{\partial \theta^2} \log f_Y(y|\theta) = \frac{1}{\theta^2}$$

Thus,

$$I(\theta) = -\mathbb{E}\left(\frac{\partial^2}{\partial \theta^2} \log f_Y(Y|\theta)\right) = -\frac{1}{\theta^2}.$$

6. (e) The Cramer-Rao inequality tells us that that if $\hat{\theta}$ is any unbiased estimator of θ based on $Y_1, \ldots, Y_{10}$, then

$$\text{Var}(\hat{\theta}) \geq \frac{1}{10\, I(\theta)} = \frac{-\theta^2}{10}.$$

Of course, for any random variable X, $\text{Var}(X) \geq 0$. Thus, having a negative lower bound in the C-R inequality is useless. It give us no new information. *The problem in this question arises from the fact that the density function of a uniform random variable is discontinuous. Therefore, technically, the computation of the Fisher inequality is invalid. My reason for asking you this question was to draw your attention to this important fact.*

7. (a) Suppose that $\hat{\theta}$ is an estimator of θ. The random interval $[L(\hat{\theta}), U(\hat{\theta})]$ is a 93% confidence interval for θ if

$$P(L(\hat{\theta}) \leq \theta \leq U(\hat{\theta})) = 0.93.$$

Hence, we interpret a 93% confidence interval to mean that before the data have been observed, there is a 93% chance that the parameter will lie in the random interval. However, once the data have been observed, no such probability statement is true. Either the given interval does or does not contain θ. Alternatively, if many, many intervals are observed, each constructed using the same formula, then the long-run average that will contain θ is 0.93.

7. (b) It is desirable to find unbiased estimators because by having an unbiased estimator we know $\mathbb{E}(\hat{\theta}) = \theta$. Thus, the most likely "value" of $\hat{\theta}$ is θ. If we have the unbiased estimator with the smallest variance, then the distribution of $\hat{\theta}$ is clustered as tightly as possible about its mean, namely θ. Thus, the MVUE is the "most likely" of all unbiased estimators to be "closest" to θ.

8. (a) By definition, the significance level α is the probability of a Type I error; that is, the probability under H_0 that H_0 is rejected. Hence, since $\overline{Y} \sim \mathcal{N}(\mu, 9/n)$,

$$0.05 = P_{H_0}(\text{reject } H_0) = P_{\mu=0}(\overline{Y} > c) = P\left(\frac{\overline{Y} - 0}{3/\sqrt{n}} > \frac{c - 0}{3/\sqrt{n}}\right) = P(Z > c\sqrt{n}\,/\,3),$$

where $Z \sim \mathcal{N}(0,1)$. From Table 4, we find that $P(Z > 1.65) = 0.05$. Therefore, we must have

$$\frac{c\sqrt{n}}{3} = 1.96 \quad \text{or} \quad c = \frac{4.95}{\sqrt{n}}.$$

8. (b) By definition, the power of a hypothesis test is the probability under H_A that H_0 is rejected. Hence, when $\mu = 1$, $n = 36$, we find $c = 0.825$, so that

$$\text{power} = P_{\mu=1}(\overline{Y} > 0.825) = P\left(\frac{\overline{Y}-1}{3/\sqrt{36}} > \frac{0.825-1}{3/\sqrt{36}}\right) = P(Z > -0.36) = 1 - 0.3594 = 0.6406$$

where $Z \sim \mathcal{N}(0,1)$. (The last step follows from Table 4.)

8. (c) As in **(a)** and **(b)**,

$$\text{power} = P_{\mu=1}\left(\overline{Y} > \frac{4.95}{\sqrt{n}}\right) = P\left(Z > \frac{4.95/\sqrt{n}-1}{3/\sqrt{n}}\right) = P(Z > 1.65 - \sqrt{n}/3).$$

Hence, as n increases ($\to \infty$), $1.65 - \sqrt{n}/3$ decreases monotonically ($\to -\infty$), so that the power increases monotonically ($\to 1$). In particular, if $m > n$, then

$$P(Z > 1.65 - \sqrt{n}/3\,) < P(Z > 1.65 - \sqrt{m}/3\,).$$

This indeed makes sense intuitively. As the sample size increases, it becomes easier to detect that $\mu = 1$ is false.

9. (a) By definition, the significance level α is the probability of a Type I error; that is, the probability under H_0 that H_0 is rejected. Hence, since $\overline{Y} \sim \mathcal{N}(\mu, 4/n)$,

$$\begin{aligned}\alpha = P_{H_0}(\text{reject } H_0) = P_{\mu=0}(\overline{Y} > 3.92/\sqrt{n}) &= P\left(\frac{\overline{Y}-0}{2/\sqrt{n}} > \frac{3.92/\sqrt{n}-0}{2/\sqrt{n}}\right)\\ &= P(Z > 1.96) = 0.025,\end{aligned}$$

where $Z \sim \mathcal{N}(0,1)$. (The last step follows from Table 4.) Hence, we see that the hypothesis test does, in fact, have significance level $\alpha = 0.025$.

9. (b) By definition, the power of a hypothesis test is the probability under H_A that H_0 is rejected. Hence, when $\mu = 0.5$, we find

$$\begin{aligned}\text{power} = P_{H_A}(\text{reject } H_0) = P_{\mu=0.5}(\overline{Y} > 3.92/\sqrt{n}) &= P\left(\frac{\overline{Y}-0.5}{2/\sqrt{n}} > \frac{3.92/\sqrt{n}-0.5}{2/\sqrt{n}}\right)\\ &= P(Z > 1.96 - 0.25\,\sqrt{n}\,)\end{aligned}$$

where $Z \sim \mathcal{N}(0,1)$. If we desire the test to have power 0.9, then using Table 4, we find $P(Z > -1.28) = 0.90$. Thus, we require that n satisfy

$$1.96 - 0.25\,\sqrt{n} = -1.28 \quad \text{or} \quad n \approx 168.$$

(In fact, we can take $n \geq 168$ to guarantee that the test will have power (at least) 0.9 when $\mu = 0.5$.

10. Draw a picture! From the scenario presented, we know that John rejects H_0 if and only if $p \leq 0.01$, and that George rejects H_0 if and only if $p \leq 0.05$. Since Ringo's p-value is smaller than 0.03, we can conclude immediately that George will reject the null hypothesis. However, John cannot make a decision. We are only told that Ringo's p-value is smaller than 0.03. We do not know, therefore, how it compares to John's desired significance level of $\alpha = 0.01$. (It could be the case that $0.01 < p < 0.03$ or it could be the case that $p < 0.01 < 0.03$. These yield different conclusions for John.)

11. Consider a hypothesis test of $H_0 : \theta = \theta_0$ against H_A where H_A could be any one of $H_A : \theta \neq \theta_0$, $H_A : \theta > \theta_0$, or $H_A : \theta < \theta_0$. The significance level α is simply the probability of a Type I error. A Type I error occurs if H_0 is rejected when, in fact, H_0 is true. Thus,

$$\alpha = P(\text{Type I error}) = P_{H_0}(\text{reject } H_0).$$

12. In this problem, we find that $\alpha = P_{\mu=0}(\overline{Y} < c)$ and $\beta = P_{\mu=-1/2}(\overline{Y} > c)$. Since $\overline{Y} \sim \mathcal{N}(\mu, \sigma^2/n) = \mathcal{N}(\mu, 0.25)$, we conclude that

$$\alpha = P_{\mu=0}(\overline{Y} < c) = P\left(\frac{\overline{Y}-0}{\sqrt{0.25}} < \frac{c-0}{\sqrt{0.25}}\right) = P(Z < 2c)$$

and

$$\beta = P_{\mu=-1/2}(\overline{Y} > c) = P\left(\frac{\overline{Y}+1/2}{\sqrt{0.25}} > \frac{c+1/2}{\sqrt{0.25}}\right) = P(Z > 2c+1)$$

where $Z \sim \mathcal{N}(0,1)$. In order for $\alpha = \beta$, we require that $P(Z < 2c) = P(Z > 2c+1)$. Since the standard normal distribution is symmetric about 0, we see that we must have $-2c = 2c+1$ or $c = -1/4$. (DRAW A PICTURE TO SEE WHERE THE MINUS SIGN COMES FROM!) Consulting Table 4, we find that with $c = -1/4$, the significance level of this test is

$$\alpha = P(Z < -1/2) = 0.3085.$$

13. (a) The samplig distribution of this estimator is vital in order to construct confidence intervals (either exactly by the pivotal method or approximately using the MLE and Fisher information) and to conduct hypothesis tests (either exactly or using the likelihood ratio test approximation). The sampling distribution is also needed so that the accuracy (bias, mean-squared error, etc.) of the estimator can be evaluated.

13. (b) You might want additional pieces of information such as the p-value of the test, the power (which can be computed exactly since both hypotheses are simple), how the data was collected, the sampling distribution of the test statistic, how the test was conducted (likelihood ratio test, CI-HT duality, Z-test, T-test, χ^2-test, etc.), and whether or not any approximations were made.

14. (a) By the confidence interval–hypothesis test duality, we do not reject $H_0 : \theta = 5$ if and only if $5 \in (X-1, X+2)$. In other words, we reject H_0 if $5 \leq X-1$ or if $5 \geq X+2$. Hence, the required rejection region is

$$RR = \{X \leq 3 \text{ or } X \geq 6\}.$$

14. (b) By definition, the significance level α is the probability of a Type I error; that is, the probability under H_0 that H_0 is rejected. Hence, we must find c so that

$$\frac{19}{100} = P_{H_0}(\text{reject } H_0) = P_{\theta=1}(\max\{Y_1, Y_2\} > c).$$

Since Y_1, Y_2 are independent Uniform$(0, \theta)$ random variables, we find

$$P_{\theta=1}(\max\{Y_1, Y_2\} \le c) = [P_{\theta=1}(Y_1 \le c)]^2 = \left[\int_0^c 1\,dy\right]^2 = c^2$$

and so $P_{\theta=1}(\max\{Y_1, Y_2\} > c) = 1 - c^2$. Setting

$$\frac{19}{100} = 1 - c^2 \quad \text{implies} \quad c = \frac{9}{10}.$$

By definition, the power of a hypothesis test is the probability under H_A that H_0 is rejected. Hence, we find

$$\begin{aligned}
\text{power} = P_{H_A}(\text{reject } H_0) = P_{\theta>1}\left(\max\{Y_1, Y_2\} > \frac{9}{10}\right) &= 1 - P_{\theta>1}\left(\max\{Y_1, Y_2\} \le \frac{9}{10}\right) \\
&= 1 - \left[P_{\theta>1}\left(Y_1 \le \frac{9}{10}\right)\right]^2 \\
&= 1 - \left[\int_0^{9/10} \frac{1}{\theta}\,dy\right]^2 \\
&= 1 - \frac{81}{100\,\theta^2}.
\end{aligned}$$

15. (a) If X_1, X_2, X_3 are i.i.d. Exponential(λ) random variables and $Y = \min\{X_1, X_2, X_3\}$, then for $y > 0$,

$$P(Y > y) = [P(X_1 > y)]^3 = [1 - P(X_1 \le y)]^3 = [1 - (1 - e^{-y/\lambda})]^3 = e^{-3y/\lambda}.$$

That is,

$$F_Y(y) = 1 - e^{-3y/\lambda} \quad \text{and} \quad f_Y(y) = \frac{3}{\lambda}e^{-3y/\lambda},\ y > 0,$$

implying that $Y \sim$ Exponential$(\lambda/3)$.

15. (b) The likelihood function is

$$L(\lambda) = f_Y(y|\lambda) = \frac{3}{\lambda}e^{-3y/\lambda}$$

(since there is $n = 1$ random variable, namely Y). In order to maximize the likelihood function, we attempt to maximize the log-likelihood function

$$\ell(\lambda) = \log 3 - \log\lambda - \frac{3y}{\lambda}.$$

Since

$$\ell'(\lambda) = -\frac{1}{\lambda} + \frac{3y}{\lambda^2}$$

so that $\ell'(\lambda) = 0$ implies $\lambda = 3y$, and since

$$\ell''(\lambda) = \frac{1}{\lambda^2} - \frac{6y}{\lambda^3}$$

so that

$$\ell''(3y) = -\frac{1}{9y^2} < 0,$$

we conclude from the second derivative test that

$$\hat{\lambda}_{\text{MLE}} = 3Y.$$

15. (c) We begin by noting that $\text{MSE}(\hat{\lambda}_{\text{MLE}}) = \text{Var}(\hat{\lambda}_{\text{MLE}}) + [B(\hat{\lambda}_{\text{MLE}})]^2$. Since $Y \sim \text{Exponential}(\lambda/3)$, we find

$$\mathbb{E}(Y) = \frac{\lambda}{3} \quad \text{and} \quad \text{Var}(Y) = \frac{\lambda^2}{9}.$$

This implies that $\text{Var}(\hat{\lambda}_{\text{MLE}}) = \text{Var}(3Y) = 9\,\text{Var}(Y) = \lambda^2$ and $\mathbb{E}(\hat{\lambda}_{\text{MLE}}) = \mathbb{E}(3Y) = 3\mathbb{E}(Y) = \lambda$ so that $B(\hat{\lambda}_{\text{MLE}}) = 0$. Hence,

$$\text{MSE}(\hat{\lambda}_{\text{MLE}}) = \text{Var}(\hat{\lambda}_{\text{MLE}}) + [B(\hat{\lambda}_{\text{MLE}})]^2 = \lambda^2 + 0 = \lambda^2.$$

15. (d) We find

$$\log f_Y(y|\lambda) = \log 3 - \log \lambda - \frac{3y}{\lambda}$$

and so

$$\frac{\partial^2}{\partial \lambda^2} \log f_Y(y|\lambda) = \frac{1}{\lambda^2} - \frac{6y}{\lambda^3}.$$

Thus,

$$I(\lambda) = -\mathbb{E}\left(\frac{\partial^2}{\partial \lambda^2} \log f_Y(Y|\lambda)\right) = -\mathbb{E}\left(\frac{1}{\lambda^2} - \frac{6Y}{\lambda^3}\right) = \frac{6\mathbb{E}(Y)}{\lambda^3} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

The Cramer-Rao inequality tells us that any unbiased estimator $\hat{\lambda}$ of λ must satisfy

$$\text{Var}(\hat{\lambda}) \geq \frac{1}{I(\lambda)} = \lambda^2$$

(since $n = 1$ in this problem). Since

$$\text{Var}(\hat{\lambda}_{\text{MLE}}) = \lambda^2 = \frac{1}{I(\lambda)}$$

we have found an unbiased estimator whose variance attains the lower bound of the Cramer-Rao inequality. Hence, $\hat{\lambda}_{\text{MLE}}$ must be the MVUE of λ.