Chapter 9: Properties of Point Estimators and Methods of Estimation

9.1 Refer to Ex. 8.8 where the variances of the four estimators were calculated. Thus,
eff($\hat\theta_1, \hat\theta_5$) = 1/3 eff($\hat\theta_2, \hat\theta_5$) = 2/3 eff($\hat\theta_3, \hat\theta_5$) = 3/5.

9.2 **a.** The three estimators a unbiased since:

$$E(\hat\mu_1) = \tfrac{1}{2}\big(E(Y_1) + E(Y_2)\big) = \tfrac{1}{2}(\mu + \mu) = \mu$$

$$E(\hat\mu_2) = \mu/4 + \frac{(n-2)\mu}{2(n-2)} + \mu/4 = \mu$$

$$E(\hat\mu_3) = E(\overline{Y}) = \mu .$$

b. The variances of the three estimators are

$$V(\hat\mu_1) = \tfrac{1}{4}(\sigma^2 + \sigma^2) = \tfrac{1}{2}\sigma^2$$

$$V(\hat\mu_2) = \sigma^2/16 + \frac{(n-2)\sigma^2}{4(n-2)^2} + \sigma^2/16 = \sigma^2/8 + \frac{\sigma^2}{4(n-2)}$$

$$V(\hat\mu_3) = \sigma^2/n .$$

Thus, eff($\hat\mu_3, \hat\mu_2$) = $\dfrac{n^2}{8(n-2)}$, eff($\hat\mu_3$ $\hat\mu_1$) = n/2.

9.3 **a.** $E(\hat\theta_1) = E(\overline{Y}) - 1/2 = \theta + 1/2 - 1/2 = \theta$. From Section 6.7, we can find the density function of $\hat\theta_2 = Y_{(n)}$: $g_n(y) = n(y-\theta)^{n-1}$, $\theta \le y \le \theta + 1$. From this, it is easily shown that $E(\hat\theta_2) = E(Y_{(n)}) - n/(n+1) = \theta$.

b. $V(\hat\theta_1) = V(\overline{Y}) = \sigma^2/n = 1/(12n)$. With the density in part **a**, $V(\hat\theta_2) = V(Y_{(n)}) = \frac{n}{(n+2)(n+1)^2}$.
Thus, eff($\hat\theta_1, \hat\theta_2$) = $\frac{12n^2}{(n+2)(n+1)^2}$.

9.4 See Exercises 8.18 and 6.74. Following those, we have that $V(\hat\theta_1) = (n+1)^2 V(Y_{(n)}) = \frac{n}{n+2}\theta^2$. Similarly, $V(\hat\theta_2) = \left(\frac{n+1}{n}\right)^2 V(Y_{(n)}) = \frac{1}{n(n+2)}\theta^2$. Thus, the ratio of these variances is as given.

9.5 From Ex. 7.20, we know S^2 is unbiased and $V(S^2) = V(\hat\sigma_1^2) = \frac{2\sigma^4}{n-1}$. For $\hat\sigma_2^2$, note that $Y_1 - Y_2$ is normal with mean 0 and variance σ^2. So, $\frac{(Y_1-Y_2)^2}{2\sigma^2}$ is chi–square with one degree of freedom and $E(\hat\sigma_2^2) = \sigma^2$, $V(\hat\sigma_2^2) = 2\sigma^4$. Thus, we have that eff($\hat\sigma_1^2, \hat\sigma_2^2$) = $n - 1$.

9.6 Both estimators are unbiased and $V(\hat\lambda_1) = \lambda/2$ and $V(\hat\lambda_2) = \lambda/n$. The efficiency is $2/n$.

9.7 The estimator $\hat\theta_1$ is unbiased so MSE($\hat\theta_1$) = $V(\hat\theta_1) = \theta^2$. Also, $\hat\theta_2 = \overline{Y}$ is unbiased for θ (θ is the mean) and $V(\hat\theta_2) = \sigma^2/n = \theta^2/n$. Thus, we have that eff($\hat\theta_1, \hat\theta_2$) = $1/n$.

9.8 **a.** It is not difficult to show that $\frac{\partial^2 \ln f(y)}{\partial \mu^2} = -\frac{1}{\sigma^2}$, so $I(\mu) = \sigma^2/n$, Since $V(\overline{Y}) = \sigma^2/n$, $\overline{Y}$ is an efficient estimator of μ.

b. Similarly, $\frac{\partial^2 \ln p(y)}{\partial \lambda^2} = -\frac{y}{\lambda^2}$ and $E(-Y/\lambda^2) = 1/\lambda$. Thus, $I(\lambda) = \lambda/n$. By Ex. 9.6, $\overline{Y}$ is an efficient estimator of λ.

9.9 **a.** $X_6 = 1$.
b.-e. Answers vary.

9.10 **a.-b.** Answers vary.

9.11 **a.-b.** Answers vary.
c. The simulations are different but get close at $n = 50$.

9.12 **a.-b.** Answers vary.

9.13 **a.** Sequences are different but settle down at large n.
b. Sequences are different but settle down at large n.

9.14 **a.** the mean, 0.
b.-c. the variability of the estimator decreases with n.

9.15 Referring to Ex. 9.3, since both estimators are unbiased and the variances go to 0 with as n goes to infinity the estimators are consistent.

9.16 From Ex. 9.5, $V(\hat{\sigma}_2^2) = 2\sigma^4$ which is constant for all n. Thus, $\hat{\sigma}_2^2$ is not a consistent estimator.

9.17 In Example 9.2, it was shown that both $\overline{X}$ and $\overline{Y}$ are consistent estimators of μ_1 and μ_2, respectively. Using Theorem 9.2, $\overline{X} - \overline{Y}$ is a consistent estimator of $\mu_1 - \mu_2$.

9.18 Note that this estimator is the pooled sample variance estimator S_p^2 with $n_1 = n_2 = n$. In Ex. 8.133 it was shown that S_p^2 is an unbiased estimator. Also, it was shown that the variance of S_p^2 is $\frac{2\sigma^4}{n_1 + n_2 - 2} = \frac{\sigma^4}{n-1}$. Since this quantity goes to 0 with n, the estimator is consistent.

9.19 Given $f(y)$, we have that $E(Y) = \frac{\theta}{\theta+1}$ and $V(Y) = \frac{\theta}{(\theta+2)(\theta+1)^2}$ (Y has a beta distribution with parameters $\alpha = \theta$ and $\beta = 1$. Thus, $E(\overline{Y}) = \frac{\theta}{\theta+1}$ and $V(\overline{Y}) = \frac{\theta}{n(\theta+2)(\theta+1)^2}$. Thus, the conditions are satisfied for $\overline{Y}$ to be a consistent estimator.

9.20 Since $E(Y) = np$ and $V(Y) = npq$, we have that $E(Y/n) = p$ and $V(Y/n) = pq/n$. Thus, Y/n is consistent since it is unbiased and its variance goes to 0 with n.

9.21 Note that this is a generalization of Ex. 9.5. The estimator $\hat{\sigma}^2$ can be written as

$$\hat{\sigma}^2 = \frac{1}{k}\left[\frac{(Y_2 - Y_1)^2}{2} + \frac{(Y_4 - Y_3)^2}{2} + \frac{(Y_6 - Y_5)^2}{2} + \ldots + \frac{(Y_n - Y_{n-1})^2}{2}\right].$$

There are k independent terms in the sum, each with mean σ^2 and variance $2\sigma^4$.

a. From the above, $E(\hat{\sigma}^2) = (k\sigma^2)/k = \sigma^2$. So $\hat{\sigma}^2$ is an unbiased estimator.

b. Similarly, $V(\hat{\sigma}^2) = k(2\sigma^4)/k^2 = 2\sigma^4/k$. Since $k = n/2$, $V(\hat{\sigma}^2)$ goes to 0 with n and $\hat{\sigma}^2$ is a consistent estimator.

9.22 Following Ex. 9.21, we have that the estimator $\hat{\lambda}$ can be written as

$$\hat{\lambda} = \frac{1}{k}\left[\frac{(Y_2 - Y_1)^2}{2} + \frac{(Y_4 - Y_3)^2}{2} + \frac{(Y_6 - Y_5)^2}{2} + \ldots + \frac{(Y_n - Y_{n-1})^2}{2}\right].$$

For Y_i, Y_{i-1}, we have that:

$$\frac{E[(Y_i - Y_{i-1})^2]}{2} = \frac{E(Y_i^2) - 2E(Y_i)E(Y_{i-1}) + E(Y_{i-1}^2)}{2} = \frac{(\lambda + \lambda^2) - 2\lambda^2 + (\lambda + \lambda^2)}{2} = \lambda$$

$$\frac{V[(Y_i - Y_{i-1})^2]}{4} < \frac{V(Y_i^2) + V(Y_{i-1}^2)}{4} = \frac{2\lambda + 12\lambda^2 + 8\lambda^3}{4} = \gamma$$, since Y_i and Y_{i-1} are independent and non–negative (the calculation can be performed using the Poisson mgf).

a. From the above, $E(\hat{\lambda}) = (k\lambda)/k = \lambda$. So $\hat{\lambda}$ is an unbiased estimator of λ.

b. Similarly, $V(\hat{\lambda}) < k\gamma/k^2$, where $\gamma < \infty$ is defined above. Since $k = n/2$, $V(\hat{\lambda})$ goes to 0 with n and $\hat{\lambda}$ is a consistent estimator.

9.23 **a.** Note that for $i = 1, 2, \ldots, k$,

$$E(Y_{2i} - Y_{2i-1}) = 0 \qquad V(Y_{2i} - Y_{2i-1}) = 2\sigma^2 = E[(Y_{2i} - Y_{2i-1})^2.$$

Thus, it follows from methods used in Ex. 9.23 that $\hat{\sigma}^2$ is an unbiased estimator.

b. $V(\hat{\sigma}^2) = \frac{1}{4k^2}\sum_{i=1}^{k} V[(Y_{2i} - Y_{2i-1})^2] = \frac{1}{4k}V[(Y_2 - Y_1)^2]$, since the Y's are independent and identically distributed. Now, it is clear that $V[(Y_2 - Y_1)^2] \le E[(Y_2 - Y_1)^4]$, and when this quantity is expanded, only moments of order 4 or less are involved. Since these were assumed to be finite, $E[(Y_2 - Y_1)^4] < \infty$ and so $V(\hat{\sigma}^2) = \frac{1}{4k}V[(Y_2 - Y_1)^2] \to 0$ as $n \to \infty$.

c. This was discussed in part b.

9.24 **a.** From Chapter 6, $\sum_{i=1}^{n} Y_i^2$ is chi–square with with n degrees of freedom.

b. Note that $E(W_n) = 1$ and $V(W_n) = 1/n$. Thus, as $n \to \infty$, $W_n \to E(W_n) = 1$ in probability.

9.25 **a.** Since $E(Y_1) = \mu$, Y_1 is unbiased.
b. $P(|Y_1 - \mu| \le 1) = P(-1 \le Z \le 1) = .6826$.
c. The estimator is not consistent since the probability found in part b does not converge to unity (here, $n = 1$).

9.26 **a.** We have that $P(\theta - \varepsilon \le Y_{(n)} \le \theta + \varepsilon) = F_{(n)}(\theta + \varepsilon) - F_{(n)}(\theta - \varepsilon)$.

- If $\varepsilon > \theta$, $F_{(n)}(\theta + \varepsilon) = 1$ and $F_{(n)}(\theta - \varepsilon) = 0$. Thus, $P(\theta - \varepsilon \le Y_{(n)} \le \theta + \varepsilon) = 1$.
- If $\varepsilon < \theta$, $F_{(n)}(\theta + \varepsilon) = 1$, $F_{(n)}(\theta - \varepsilon) = \left(\frac{\theta-\varepsilon}{\theta}\right)^n$. So, $P(\theta - \varepsilon \le Y_{(n)} \le \theta + \varepsilon) = 1 - \left(\frac{\theta-\varepsilon}{\theta}\right)^n$.

b. The result follows from $\lim_{n\to\infty} P(\theta - \varepsilon \le Y_{(n)} \le \theta + \varepsilon) = \lim_{n\to\infty}\left[1 - \left(\frac{\theta-\varepsilon}{\theta}\right)^n\right] = 1$.

9.27 $P(|Y_{(1)} - \theta| \le \varepsilon) = P(\theta - \varepsilon \le Y_{(1)} \le \theta + \varepsilon) = F_{(1)}(\theta + \varepsilon) - F_{(1)}(\theta - \varepsilon) = 1 - \left(1 - \frac{\theta-\varepsilon}{\theta}\right)^n = \left(\frac{\varepsilon}{\theta}\right)^n$.
But, $\lim_{n\to\infty}\left(\frac{\varepsilon}{\theta}\right)^n = 0$ for $\varepsilon < \theta$. So, $Y_{(1)}$ is not consistent.

9.28 $P(|Y_{(1)} - \beta| \le \varepsilon) = P(\beta - \varepsilon \le Y_{(1)} \le \beta + \varepsilon) = F_{(1)}(\beta + \varepsilon) - F_{(1)}(\beta - \varepsilon) = 1 - \left(\frac{\beta}{\beta+\varepsilon}\right)^{\alpha n}$. Since $\lim_{n\to\infty}\left(\frac{\beta}{\beta+\varepsilon}\right)^{\alpha n} = 0$ for $\varepsilon > 0$, $Y_{(1)}$ is consistent.

9.29 $P(|Y_{(1)} - \theta| \le \varepsilon) = P(\theta - \varepsilon \le Y_{(1)} \le \theta + \varepsilon) = F_{(1)}(\theta + \varepsilon) - F_{(1)}(\theta - \varepsilon) = 1 - \left(\frac{\theta-\varepsilon}{\theta}\right)^{\alpha n}$. Since $\lim_{n\to\infty}\left(\frac{\theta-\varepsilon}{\theta}\right)^{\alpha n} = 0$ for $\varepsilon > 0$, $Y_{(1)}$ is consistent.

9.30 Note that Y is beta with $\mu = 3/4$ and $\sigma^2 = 3/5$. Thus, $E(\overline{Y}) = 3/4$ and $V(\overline{Y}) = 3/(5n)$. Thus, $V(\overline{Y}) \to 0$ and $\overline{Y}$ converges in probability to 3/4.

9.31 Since $\overline{Y}$ is a mean of independent and identically distributed random variables with finite variance, $\overline{Y}$ is consistent and $\overline{Y}$ converges in probability to $E(\overline{Y}) = E(Y) = \alpha\beta$.

9.32 Notice that $E(Y^2) = \int_2^\infty y^2 \frac{2}{y^2} dy = \int_2^\infty 2dy = \infty$, thus $V(Y) = \infty$ and so the law of large numbers does not apply.

9.33 By the law of large numbers, $\overline{X}$ and $\overline{Y}$ are consistent estimators of λ_1 and λ_2. By Theorem 9.2, $\frac{\overline{X}}{\overline{X}+\overline{Y}}$ converges in probability to $\frac{\lambda_1}{\lambda_1+\lambda_2}$. This implies that observed values of the estimator should be close to the limiting value for large sample sizes, although the variance of this estimator should also be taken into consideration.

9.34 Following Ex. 6.34, Y^2 has an exponential distribution with parameter θ. Thus, $E(Y^2) = \theta$ and $V(Y^2) = \theta^2$. Therefore, $E(W_n) = \theta$ and $V(W_n) = \theta^2/n$. Clearly, W_n is a consistent estimator of θ.

9.35 **a.** $E(\overline{Y}_n) = \frac{1}{n}(\mu + \mu + \cdots + \mu) = \mu$, so $\overline{Y}_n$ is unbiased for μ.

b. $V(\overline{Y}_n) = \frac{1}{n^2}(\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2$.

c. In order for $\overline{Y}_n$ to be consistent, it is required that $V(\overline{Y}_n) \to 0$ as $n \to \infty$. Thus, it must be true that all variances must be finite, or simply $\max_i\{\sigma_i^2\} < \infty$.

9.36 Let $X_1, X_2, \ldots, X_n$ be a sequence of Bernoulli trials with success probability p. Thus, it is seen that $Y = \sum_{i=1}^{n} X_i$. Thus, by the Central Limit Theorem, $U_n = \dfrac{\hat{p}_n - p}{\sqrt{\frac{pq}{n}}}$ has a limiting standard normal distribution. By Ex. 9.20, it was shown that $\hat{p}_n$ is consistent for p, so it makes sense that $\hat{q}_n$ is consistent for q, and so by Theorem 9.2 $\hat{p}_n\hat{q}_n$ is consistent for pq. Define $W_n = \sqrt{\dfrac{\hat{p}_n\hat{q}_n}{pq}}$ so that W_n converges in probability to 1. By Theorem 9.3, the quantity $\dfrac{U_n}{W_n} = \dfrac{\hat{p}_n - p}{\sqrt{\frac{\hat{p}_n\hat{q}_n}{n}}}$ converges to a standard normal variable.

9.37 The likelihood function is $L(p) = p^{\Sigma x_i}(1-p)^{n-\Sigma x_i}$. By Theorem 9.4, $\sum_{i=1}^{n} X_i$ is sufficient for p with $g(\Sigma x_i, p) = p^{\Sigma x_i}(1-p)^{n-\Sigma x_i}$ and $h(\mathbf{y}) = 1$.

9.38 For this exercise, the likelihood function is given by

$$L = \frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left[-\frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2\sigma^2}\right] = (2\pi)^{-n/2}\sigma^{-n}\exp\left[\frac{-1}{2\sigma^2}\left(\sum_{i=1}^{n} y_i^2 - 2\mu n\overline{y} + n\mu^2\right)\right].$$

a. When σ^2 is known, $\overline{Y}$ is sufficient for μ by Theorem 9.4 with

$$g(\overline{y}, \mu) = \exp\left(\frac{2\mu n\overline{y} - n\mu^2}{2\sigma^2}\right) \text{ and } h(\mathbf{y}) = (2\pi)^{-n/2}\sigma^{-n}\exp\left(\frac{-1}{2\sigma^2}\sum_{i=1}^{n} y_i^2\right).$$

b. When μ is known, use Theorem 9.4 with

$$g\left(\sum_{i=1}^{n}(y_i-\mu)^2, \sigma^2\right) = (\sigma^2)^{-n/2}\exp\left[-\frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2\sigma^2}\right] \text{ and } h(\mathbf{y}) = (2\pi)^{-n/2}.$$

c. When both μ and σ^2 are unknown, the likelihood can be written in terms of the two statistics $U_1 = \sum_{i=1}^{n} Y_i$ and $U_2 = \sum_{i=1}^{n} Y_i^2$ with $h(\mathbf{y}) = (2\pi)^{-n/2}$. The statistics $\overline{Y}$ and S^2 are also jointly sufficient since they can be written in terms of U_1 and U_2.

9.39 Note that by independence, $U = \sum_{i=1}^{n} Y_i$ has a Poisson distribution with parameter $n\lambda$. Thus, the conditional distribution is expressed as

$$P(Y_1 = y_1, \ldots, Y_n = y_n \mid U = u) = \frac{P(Y_1 = y_1, \ldots, Y_n = y_n)}{P(U = u)} = \frac{\prod_{i=1}^{n} \frac{\lambda^{y_i} e^{-\lambda}}{y_i!}}{\frac{(n\lambda)^u e^{-n\lambda}}{u!}} = \frac{\frac{\lambda^{\Sigma y_i} e^{-n\lambda}}{\Pi y_i!}}{\frac{(n\lambda)^u e^{-n\lambda}}{u!}}.$$

We have that $\Sigma y_i = u$, so the above simplifies to

$$P(Y_1 = y_1, \ldots, Y_n = y_n \mid U = u) = \begin{cases} \dfrac{u!}{n^u \Pi y_i!} & \text{if } \Sigma y_i = u \\ 0 & \text{otherwise} \end{cases}.$$

Since the conditional distribution is free of λ, the statistic $U = \sum_{i=1}^{n} Y_i$ is sufficient for λ.

9.40 The likelihood is $L(\theta) = 2^n \theta^{-n} \prod_{i=1}^{n} y_i \exp\left(-\sum_{i=1}^{n} y_i^2 / \theta\right)$. By Theorem 9.4, $U = \sum_{i=1}^{n} Y_i^2$ is sufficient for θ with $g(u, \theta) = \theta^{-n} \exp(-u/\theta)$ and $h(\boldsymbol{y}) = 2^n \prod_{i=1}^{n} y_i$.

9.41 The likelihood is $L(\alpha) = \alpha^{-n} m^n \left(\prod_{i=1}^{n} y_i\right)^{m-1} \exp\left(-\sum_{i=1}^{n} y_i^m / \alpha\right)$. By Theorem 9.4, $U = \sum_{i=1}^{n} Y_i^m$ is sufficient for α with $g(u, \alpha) = \alpha^{-n} \exp(-u/\alpha)$ and $h(\boldsymbol{y}) = m^n \left(\prod_{i=1}^{n} y_i\right)^{m-1}$.

9.42 The likelihood function is $L(p) = p^n (1-p)^{\Sigma y_i - n} = p^n (1-p)^{n\bar{y} - n}$. By Theorem 9.4, $\bar{Y}$ is sufficient for p with $g(\bar{y}, p) = p^n (1-p)^{n\bar{y}-n}$ and $h(\boldsymbol{y}) = 1$.

9.43 With θ known, the likelihood is $L(\alpha) = \alpha^n \theta^{-n\alpha} \left(\prod_{i=1}^{n} y_i\right)^{\alpha - 1}$. By Theorem 9.4, $U = \prod_{i=1}^{n} Y_i$ is sufficient for α with $g(u, \alpha) = \alpha^n \theta^{-n\alpha} \left(\prod_{i=1}^{n} y_i\right)^{\alpha-1}$ and $h(\boldsymbol{y}) = 1$.

9.44 With β known, the likelihood is $L(\alpha) = \alpha^n \beta^{n\alpha} \left(\prod_{i=1}^{n} y_i\right)^{-(\alpha+1)}$. By Theorem 9.4, $U = \prod_{i=1}^{n} Y_i$ is sufficient for α with $g(u, \alpha) = \alpha^n \beta^{n\alpha} (u)^{-(\alpha+1)}$ and $h(\boldsymbol{y}) = 1$.

9.45 The likelihood function is

$$L(\theta) = \prod_{i=1}^{n} f(y_i \mid \theta) = [a(\theta)]^n \left[\prod_{i=1}^{n} b(y_i)\right] \exp\left[-c(\theta) \sum_{i=1}^{n} d(y_i)\right].$$

Thus, $U = \sum_{i=1}^{n} d(Y_i)$ is sufficient for θ because by Theorem 9.4 $L(\theta)$ can be factored into, where $u = \sum_{i=1}^{n} d(y_i)$, $g(u, \theta) = [a(\theta)]^n \exp[-c(\theta) u]$ and $h(\boldsymbol{y}) = \prod_{i=1}^{n} b(y_i)$.

9.46 The exponential distribution is in exponential form since $a(\beta) = c(\beta) = 1/\beta$, $b(y) = 1$, and $d(y) = y$. Thus, by Ex. 9.45, $\sum_{i=1}^{n} Y_i$ is sufficient for β, and then so is $\bar{Y}$.

9.47 We can write the density function as $f(y \mid \alpha) = \alpha\theta^{\alpha} \exp[-(\alpha - 1)\ln y]$. Thus, the density has exponential form and the sufficient statistic is $\sum_{i=1}^{n} \ln\left(\overline{Y}_i\right)$. Since this is equivalently expressed as $\ln\left(\prod_{i=1}^{n} Y_i\right)$, we have no contradiction with Ex. 9.43.

9.48 We can write the density function as $f(y \mid \alpha) = \alpha\beta^{\alpha} \exp[-(\alpha + 1)\ln y]$. Thus, the density has exponential form and the sufficient statistic is $\sum_{i=1}^{n} \ln Y_i$. Since this is equivalently expressed as $\ln \prod_{i=1}^{n} Y_i$, we have no contradiction with Ex. 9.44.

9.49 The density for the uniform distribution on (0, θ) is $f(y \mid \theta) = \frac{1}{\theta}$, $0 \le y \le \theta$. For this problem and several of the following problems, we will use an <u>indicator function</u> to specify the support of y. This is given by, in general, for $a < b$,

$$I_{a,b}(y) = \begin{cases} 1 & \text{if } a \le y \le b \\ 0 & \text{otherwise} \end{cases}.$$

Thus, the previously mentioned uniform distribution can be expressed as

$$f(y \mid \theta) = \frac{1}{\theta} I_{0,\theta}(y).$$

The likelihood function is given by $L(\theta) = \frac{1}{\theta^n} \prod_{i=1}^{n} I_{0,\theta}(y_i) = \frac{1}{\theta^n} I_{0,\theta}(y_{(n)})$, since $\prod_{i=1}^{n} I_{0,\theta}(y_i) = I_{0,\theta}(y_{(n)})$. Therefore, Theorem 9.4 is satisfied with $h(\mathbf{y}) = 1$ and $g(y_{(n)}, \theta) = \frac{1}{\theta^n} I_{0,\theta}(y_{(n)})$.

(This problem could also be solved using the conditional distribution definition of sufficiency.)

9.50 As in Ex. 9.49, we will define the uniform distribution on the interval (θ_1, θ_2) as

$$f(y \mid \theta_1, \theta_2) = \frac{1}{(\theta_2 - \theta_1)} I_{\theta_1,\theta_2}(y).$$

The likelihood function, using the same logic as in Ex. 9.49, is

$$L(\theta_1, \theta_2) = \frac{1}{(\theta_2 - \theta_1)^n} \prod_{i=1}^{n} I_{\theta_1,\theta_2}(y_i) = \frac{1}{(\theta_2 - \theta_1)^n} I_{\theta_1,\theta_2}(y_{(1)}) I_{\theta_1,\theta_2}(y_{(n)}).$$

So, Theorem 9.4 is satisfied with $g(y_{(1)}, y_{(n)}, \theta_1, \theta_2) = \frac{1}{(\theta_2 - \theta_1)^n} I_{\theta_1,\theta_2}(y_{(1)}) I_{\theta_1,\theta_2}(y_{(n)})$ and $h(\mathbf{y}) = 1$.

9.51 Again, using the indicator notation, the density is

$$f(y \mid \theta) = \exp[-(y - \theta)] I_{a,\infty}(y)$$

(it should be obvious that $y < \infty$ for the indicator function). The likelihood function is

$$L(\theta) = \exp\left(-\sum_{i=1}^{n} y + n\theta\right)\prod_{i=1}^{n} I_{a,\infty}(y_i) = \exp\left(-\sum_{i=1}^{n} y + n\theta\right)I_{a,\infty}(y_{(1)}).$$

Theorem 9.4 is satisfied with $g(y_{(1)},\theta) = \exp(n\theta)I_{a,\infty}(y_{(1)})$ and $h(\mathbf{y}) = \exp\left(-\sum_{i=1}^{n} y\right)$.

9.52 Again, using the indicator notation, the density is

$$f(y\,|\,\theta) = \frac{3y^2}{\theta^3} I_{0,\theta}(y).$$

The likelihood function is $L(\theta) = \dfrac{3^n \prod_{i=1}^{n} y_i^2}{\theta^{3n}} \prod_{i=1}^{n} I_{0,\theta}(y_i) = \dfrac{3^n \prod_{i=1}^{n} y_i^2}{\theta^{3n}} I_{0,\theta}(y_{(n)})$. Then,

Theorem 9.4 is satisfied with $g(y_{(n)},\theta) = \theta^{-3n} I_{0,\theta}(y_{(n)})$ and $h(\mathbf{y}) = 3^n \prod_{i=1}^{n} y_i^2$.

9.53 Again, using the indicator notation, the density is

$$f(y\,|\,\theta) = \frac{2\theta^2}{y^3} I_{\theta,\infty}(y).$$

The likelihood function is $L(\theta) = 2^n \theta^{2n}\left(\prod_{i=1}^{n} y_i^{-3}\right)\prod_{i=1}^{n} I_{\theta,\infty}(y_i) = 2^n \theta^{2n}\left(\prod_{i=1}^{n} y_i^{-3}\right) I_{\theta,\infty}(y_{(1)})$

Theorem 9.4 is satisfied with $g(y_{(1)},\theta) = \theta^{2n} I_{\theta,\infty}(y_{(1)})$ and $h(\mathbf{y}) = 2^n\left(\prod_{i=1}^{n} y_i^{-3}\right)$.

9.54 Again, using the indicator notation, the density is

$$f(y\,|\,\alpha,\theta) = \alpha\theta^{-\alpha} y^{\alpha-1} I_{0,\theta}(y).$$

The likelihood function is

$$L(\alpha,\theta) = \alpha^n \theta^{-n\alpha}\left(\prod_{i=1}^{n} y_i\right)^{\alpha-1}\prod_{i=1}^{n} I_{0,\theta}(y_i) = \alpha^n \theta^{-n\alpha}\left(\prod_{i=1}^{n} y_i\right)^{\alpha-1} I_{0,\theta}(y_{(n)}).$$

Theorem 9.4 is satisfied with $g(\prod_{i=1}^{n} y_i, y_{(n)}, \alpha, \theta) = \alpha^n \theta^{-n\alpha}\left(\prod_{i=1}^{n} y_i\right)^{\alpha-1} I_{0,\theta}(y_{(n)})$, $h(\mathbf{y}) = 1$ so that $\left(\prod_{i=1}^{n} Y_i, Y_{(n)}\right)$ is jointly sufficient for α and θ.

9.55 Lastly, using the indicator notation, the density is

$$f(y\,|\,\alpha,\beta) = \alpha\beta^{\alpha} y^{-(\alpha+1)} I_{\beta,\infty}(y).$$

The likelihood function is

$$L(\alpha,\beta) = \alpha^n \beta^{n\alpha}\left(\prod_{i=1}^{n} y_i^{-(\alpha+1)}\right)\prod_{i=1}^{n} I_{\beta,\infty}(y_i) = \alpha^n \beta^{n\alpha}\left(\prod_{i=1}^{n} y_i^{-(\alpha+1)}\right) I_{\beta,\infty}(y_{(1)}).$$

Theorem 9.4 is satisfied with $g(\prod_{i=1}^{n} y_i, y_{(1)}, \alpha, \beta) = \alpha^n \beta^{n\alpha}\left(\prod_{i=1}^{n} y_i^{-(\alpha+1)}\right) I_{\beta,\infty}(y_{(1)})$, and $h(\mathbf{y}) = 1$ so that $\left(\prod_{i=1}^{n} Y_i, Y_{(1)}\right)$ is jointly sufficient for α and β.

9.56 In Ex. 9.38 (b), it was shown that $\sum_{i=1}^{n}(y_i - \mu)^2$ is sufficient for σ^2. Since the quantity $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \mu)^2$ is unbiased and a function of the sufficient statistic, it is the MVUE of σ^2.

9.57 Note that the estimator can be written as

$$\hat{\sigma}^2 = \frac{S_X^2 + S_Y^2}{2},$$

where $S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$, $S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$. Since both of these estimators are the MVUE (see Example 9.8) for σ^2 and E($\hat{\sigma}^2$) = σ^2, $\hat{\sigma}^2$ is the MVUE for σ^2.

9.58 From Ex. 9.34 and 9.40, $\sum_{i=1}^{n} Y_i^2$ is sufficient for θ and $E(Y^2) = \theta$. Thus, the MVUE is $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} Y_i^2$.

9.59 Note that $E(C) = E(3Y^2) = 3E(Y^2) = 3[V(Y) + (E(Y))^2] = 3(\lambda + \lambda^2)$. Now, from Ex. 9.39, it was determined that $\sum_{i=1}^{n} Y_i$ is sufficient for λ, so if an estimator can be found that is unbiased for $3(\lambda + \lambda^2)$ and a function of the sufficient statistic, it is the MVUE. Note that $\sum_{i=1}^{n} Y_i$ is Poisson with parameter $n\lambda$, so

$$E(\overline{Y}^2) = V(\overline{Y}) + [E(\overline{Y})]^2 = \tfrac{\lambda}{n} + \lambda^2, \text{ and}$$
$$E(\overline{Y}/n) = \lambda / n.$$

Thus $\lambda^2 = E(\overline{Y}^2) - E(\overline{Y}/n)$ so that the MVUE for $3(\lambda + \lambda^2)$ is

$$3\left[\overline{Y}^2 - \overline{Y}/n + \overline{Y}\right] = 3\left[\overline{Y}^2 + \overline{Y}\left(1 - \tfrac{1}{n}\right)\right].$$

9.60 **a.** The density can be expressed as $f(y \mid \theta) = \theta \exp[(\theta - 1)\ln y]$. Thus, the density has exponential form and $-\sum_{i=1}^{n} \ln y_i$ is sufficient for θ.

b. Let $W = -\ln Y$. The distribution function for W is

$$F_W(w) = P(W \le w) = P(-\ln Y \le w) = 1 - P(Y \le e^{-w}) = 1 - \int_0^{e^{-w}} \theta y^{\theta - 1} dy = 1 - e^{-\theta w},\ w > 0.$$

This is the exponential distribution function with mean 1/θ.

c. For the transformation $U = 2\theta W$, the distribution function for U is

$$F_U(u) = P(U \le u) = P(2\theta W \le u) = P(W \le \tfrac{u}{2\theta}) = F_W(\tfrac{u}{2\theta}) = 1 - e^{-u/2},\ u > 0.$$

Note that this is the exponential distribution with mean 2, but this is equivalent to the chi–square distribution with 2 degrees of freedom. Therefore, by property of independent chi–square variables, $2\theta\sum_{i=1}^{n} W_i$ is chi–square with $2n$ degrees of freedom.

d. From Ex. 4.112, the expression for the expected value of the reciprocal of a chi–square variable is given. Thus, it follows that $E\left[\left(2\theta\sum_{i=1}^{n} W_i\right)^{-1}\right] = \dfrac{1}{2n-2} = \dfrac{1}{2(n-1)}$.

e. From part d, $\dfrac{n-1}{\sum_{i=1}^{n} W_i} = \dfrac{n-1}{-\sum_{i=1}^{n} \ln Y_i}$ is unbiased and thus the MVUE for θ.

9.61 It has been shown that $Y_{(n)}$ is sufficient for θ and $E(Y_{(n)}) = \left(\frac{n}{n+1}\right)\theta$. Thus, $\left(\frac{n+1}{n}\right)Y_{(n)}$ is the MVUE for θ.

9.62 Calculate $E(Y_{(1)}) = \int_{\theta}^{\infty} nye^{-n(y-\theta)}dy = \int_{0}^{\infty} n(u+\theta)e^{-nu}du = \theta + \frac{1}{n}$. Thus, $Y_{(1)} - \frac{1}{n}$ is the MVUE for θ.

9.63 **a.** The distribution function for Y is $F(y) = y^3/\theta^3$, $0 \le y \le \theta$. So, the density function for $Y_{(n)}$ is $f_{(n)}(y) = n[F(y)]^{n-1} f(y) = 3ny^{3n-1}/\theta^{3n}$, $0 \le y \le \theta$.

b. From part **a**, it can be shown that $E(Y_{(n)}) = \frac{3n}{3n+1}\theta$. Since $Y_{(n)}$ is sufficient for θ, $\frac{3n+1}{3n} Y_{(n)}$ is the MVUE for θ.

9.64 **a.** From Ex. 9.38, $\overline{Y}$ is sufficient for μ. Also, since σ = 1, $\overline{Y}$ has a normal distribution with mean μ and variance 1/*n*. Thus, $E(\overline{Y}^2) = V(\overline{Y}) + [E(\overline{Y})]^2 = 1/n + \mu^2$. Therefore, the MVUE for μ^2 is $\overline{Y}^2 - 1/n$.

b. $V(\overline{Y}^2 - 1/n) = V(\overline{Y}^2) = E(\overline{Y}^4) - [E(\overline{Y}^2)]^2 = E(\overline{Y}^4) - [1/n + \mu^2]^2$. It can be shown that $E(\overline{Y}^4) = \frac{3}{n^2} + \frac{6\mu^2}{n} + \mu^4$ (the mgf for $\overline{Y}$ can be used) so that

$$V(\overline{Y}^2 - 1/n) = \tfrac{3}{n^2} + \tfrac{6\mu^2}{n} + \mu^4 -- [1/n + \mu^2]^2 = (2 + 4n\mu^2)/n^2.$$

9.65 **a.** $E(T) = P(T=1) = P(Y_1 = 1, Y_2 = 0) = P(Y_1 = 1)P(Y_2 = 0) = p(1-p)$.

b. $P(T = 1 \mid W = w) = \dfrac{P(Y_1 = 1, Y_2 = 0, W = w)}{P(W = w)} = \dfrac{P(Y_1 = 1, Y_2 = 0, \sum_{i=3}^{n} Y_i = w-1)}{P(W = w)}$

$$= \frac{P(Y_1 = 1)P(Y_2 = 0)P(\sum_{i=3}^{n} Y_i = w-1)}{P(W = w)} = \frac{p(1-p)\binom{n-2}{w-1}p^{w-1}(1-p)^{n-(w-1)}}{\binom{n}{w}p^{w}(1-p)^{n-w}}$$

$$= \frac{w(n-w)}{n(n-1)}.$$

c. $E(T \mid W) = P(T = 1 \mid W) = \dfrac{W}{n}\left(\dfrac{n-W}{n-1}\right) = \left(\dfrac{n}{n-1}\right)\dfrac{W}{n}\left(1 - \dfrac{W}{n}\right)$. Since T is unbiased by part (a) above and W is sufficient for p and so also for $p(1-p)$, $n\overline{Y}(1-\overline{Y})/(n-1)$ is the MVUE for $p(1-p)$.

9.66 **a.** **i.** The ratio of the likelihoods is given by

$$\frac{L(\boldsymbol{x} \mid p)}{L(\boldsymbol{y} \mid p)} = \frac{p^{\Sigma x_i}(1-p)^{n-\Sigma x_i}}{p^{\Sigma y_i}(1-p)^{n-\Sigma y_i}} = \frac{p^{\Sigma x_i}(1-p)^{-\Sigma x_i}}{p^{\Sigma y_i}(1-p)^{-\Sigma y_i}} = \left(\frac{p}{1-p}\right)^{\Sigma x_i - \Sigma y_i}$$

ii. If $\Sigma x_i = \Sigma y_i$, the ratio is 1 and free of p. Otherwise, it will not be free of p.

iii. From the above, it must be that $g(Y_1, \ldots, Y_n) = \sum_{i=1}^n Y_i$ is the minimal sufficient statistic for p. This is the same as in Example 9.6.

b. **i.** The ratio of the likelihoods is given by

$$\frac{L(\boldsymbol{x} \mid \theta)}{L(\boldsymbol{y} \mid \theta)} = \frac{2^n(\prod_{i=1}^n x_i)\theta^{-n}\exp(-\sum_{i=1}^n x_i^2/\theta)}{2^n(\prod_{i=1}^n y_i)\theta^{-n}\exp(-\sum_{i=1}^n y_i^2/\theta)} = \frac{\prod_{i=1}^n x_i}{\prod_{i=1}^n y_i}\exp\left[-\frac{1}{\theta}\left(\sum_{i=1}^n x_i^2 - \sum_{i=1}^n y_i^2\right)\right]$$

ii. The above likelihood ratio will only be free of θ if $\sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i^2$, so that $\sum_{i=1}^n Y_i^2$ is a minimal sufficient statistic for θ.

9.67 The likelihood is given by

$$L(\boldsymbol{y} \mid \mu, \sigma^2) = \frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left[-\frac{\sum_{i=1}^n (y_i-\mu)^2}{2\sigma^2}\right].$$

The ratio of the likelihoods is

$$\frac{L(\boldsymbol{x} \mid \mu, \sigma^2)}{L(\boldsymbol{y} \mid \mu, \sigma^2)} = \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^n (x_i-\mu)^2 - \sum_{i=1}^n (y_i-\mu)^2\right]\right\} =$$

$$\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^n x_i^2 - \sum_{i=1}^n y_i^2 - 2\mu\left(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i\right)\right]\right\}.$$

This ratio is free of (μ, σ^2) only if both $\sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i^2$ and $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$, so $\sum_{i=1}^n Y_i$ and $\sum_{i=1}^n Y_i^2$ form jointly minimal sufficient statistics for μ and σ^2.

9.68 For unbiased estimators $g_1(U)$ and $g_2(U)$, whose values only depend on the data through the sufficient statistic U, we have that $E[g_1(U) - g_2(U)] = 0$. Since the density for U is complete, $g_1(U) - g_2(U) \equiv 0$ by definition so that $g_1(U) = g_2(U)$. Therefore, there is only one unbiased estimator for θ based on U, and it must also be the MVUE.

9.69 It is easy to show that $\mu = \frac{\theta+1}{\theta+2}$ so that $\theta = \frac{2\mu-1}{1-\mu}$. Thus, the MOM estimator is $\hat{\theta} = \frac{2\overline{Y}-1}{1-\overline{Y}}$. Since $\overline{Y}$ is a consistent estimator of μ, by the Law of Large Numbers $\hat{\theta}$ converges in probability to θ. However, this estimator is not a function of the sufficient statistic so it can't be the MVUE.

9.70 Since μ = λ, the MOM estimator of λ is $\hat{\lambda} = m_1' = \overline{Y}$.

9.71 Since $E(Y) = \mu_1' = 0$ and $E(Y^2) = \mu_2' = V(Y) = \sigma^2$, we have that $\hat{\sigma}^2 = m_2' = \frac{1}{n}\sum_{i=1}^{n} Y_i^2$.

9.72 Here, we have that $\mu_1' = \mu$ and $\mu_2' = \sigma^2 + \mu^2$. Thus, $\hat{\mu} = m_1' = \overline{Y}$ and $\hat{\sigma}^2 = m_2' - \overline{Y}^2 = \frac{1}{n}\sum_{i=1}^{n} Y_i^2 - \overline{Y}^2 = \frac{1}{n}\sum_{i=1}^{2} (Y_i - \overline{Y})^2$.

9.73 Note that our sole observation Y is hypergeometric such that $E(Y) = n\theta/N$. Thus, the MOM estimator of θ is $\hat{\theta} = NY/n$.

9.74 **a.** First, calculate $\mu_1' = E(Y) = \int_0^\theta 2y(\theta - y)/\theta^2 dy = \theta/3$. Thus, the MOM estimator of θ is $\hat{\theta} = 3\overline{Y}$.

b. The likelihood is $L(\theta) = 2^n \theta^{-2n} \prod_{i=1}^{n} (\theta - y_i)$. Clearly, the likelihood can't be factored into a function that only depends on $\overline{Y}$, so the MOM is not a sufficient statistic for θ.

9.75 The density given is a beta density with $\alpha = \beta = \theta$. Thus, $\mu_1' = E(Y) = .5$. Since this doesn't depend on θ, we turn to $\mu_2' = E(Y^2) = \frac{\theta+1}{2(2\theta+1)}$ (see Ex. 4.200). Hence, with $m_2' = \frac{1}{n}\sum_{i=1}^{n} Y_i^2$, the MOM estimator of θ is $\hat{\theta} = \frac{1-2m_2'}{4m_2'-1}$.

9.76 Note that $\mu_1' = E(Y) = 1/p$. Thus, the MOM estimator of p is $\hat{p} = 1/Y$.

9.77 Here, $\mu_1' = E(Y) = \frac{3}{2}\theta$. So, the MOM estimator of θ is $\hat{\theta} = \frac{2}{3}\overline{Y}$.

9.78 For Y following the given power family distribution,

$$E(Y) = \int_0^3 \alpha y^\alpha 3^{-\alpha} dy = \alpha 3^{-\alpha} \frac{y^{\alpha+1}}{\alpha+1}\Big|_0^3 = \frac{3\alpha}{\alpha+1}.$$

Thus, the MOM estimator of θ is $\hat{\theta} = \frac{\overline{Y}}{3-\overline{Y}}$.

9.79 For Y following the given Pareto distribution,

$$E(Y) = \int_\beta^\infty \alpha\beta^\alpha y^{-\alpha} dy = \alpha\beta^\alpha \frac{y^{-\alpha+1}}{-\alpha+1}\Big|_\beta^\infty = \alpha\beta/(\alpha - 1).$$

The mean is not defined if $\alpha < 1$. Thus, a generalized MOM estimator for α cannot be expressed.

9.80 **a.** The MLE is easily found to be $\hat{\lambda} = \overline{Y}$.
b. $E(\hat{\lambda}) = \lambda$, $V(\hat{\lambda}) = \lambda/n$.

c. Since $\hat{\lambda}$ is unbiased and has a variance that goes to 0 with increasing n, it is consistent.
d. By the invariance property, the MLE for $P(Y = 0)$ is $\exp(-\hat{\lambda})$.

9.81 The MLE is $\hat{\theta} = \overline{Y}$. By the invariance property of MLEs, the MLE of θ^2 is $\overline{Y}^2$.

9.82 The likelihood function is $L(\theta) = \theta^{-n} r^n \left(\prod_{i=1}^n y_i\right)^{r-1} \exp\left(-\sum_{i=1}^n y_i^r / \theta\right)$.

a. By Theorem 9.4, a sufficient statistic for θ is $\sum_{i=1}^n Y_i^r$.

b. The log–likelihood is

$$\ln L(\theta) = -n \ln \theta + n \ln r + (r-1)\ln\left(\prod_{i=1}^n y_i\right) - \sum_{i=1}^n y_i^r / \theta .$$

By taking a derivative w.r.t. θ and equating to 0, we find $\hat{\theta} = \frac{1}{n}\sum_{i=1}^n Y_i^r$.

c. Note that $\hat{\theta}$ is a function of the sufficient statistic. Since it is easily shown that $E(Y^r) = \theta$, $\hat{\theta}$ is then unbiased and the MVUE for θ.

9.83 **a.** The likelihood function is $L(\theta) = (2\theta + 1)^{-n}$. Let $\gamma = \gamma(\theta) = 2\theta + 1$. Then, the likelihood can be expressed as $L(\gamma) = \gamma^{-n}$. The likelihood is maximized for small values of γ. The smallest value that can safely maximize the likelihood (see Example 9.16) without violating the support is $\hat{\gamma} = Y_{(n)}$. Thus, by the invariance property of MLEs,

$$\hat{\theta} = \tfrac{1}{2}\left(Y_{(n)} - 1\right).$$

b. Since $V(Y) = \frac{(2\theta+1)^2}{12}$. By the invariance principle, the MLE is $\left(Y_{(n)}\right)^2 / 12$.

9.84 This exercise is a special case of Ex. 9.85, so we will refer to those results.

a. The MLE is $\hat{\theta} = \overline{Y}/2$, so the maximum likelihood estimate is $\overline{y}/2 = 63$.

b. $E(\hat{\theta}) = \theta$, $V(\hat{\theta}) = V(\overline{Y}/2) = \theta^2/6$.

c. The bound on the error of estimation is $2\sqrt{V(\hat{\theta})} = 2\sqrt{(130)^2/6} = 106.14$.

d. Note that $V(Y) = 2\theta^2 = 2(130)^2$. Thus, the MLE for $V(Y) = 2(\hat{\theta})^2$.

9.85 **a.** For α > 0 known the likelihood function is

$$L(\theta) = \frac{1}{[\Gamma(\alpha)]^n \theta^{n\alpha}} \left(\prod_{i=1}^n y_i\right)^{\alpha-1} \exp\left(-\sum_{i=1}^n y_i / \theta\right).$$

The log–likelihood is then

$$\ln L(\theta) = -n \ln[\Gamma(\alpha)] - n\alpha \ln \theta + (\alpha - 1)\ln\left(\prod_{i=1}^n y_i\right) - \sum_{i=1}^n y_i / \theta$$

so that

$$\tfrac{d}{d\theta} \ln L(\theta) = -n\alpha / \theta + \sum_{i=1}^n y_i / \theta^2 .$$

Equating this to 0 and solving for θ, we find the MLE of θ to be

$$\hat{\theta} = \tfrac{1}{n\alpha}\sum_{i=1}^n Y_i = \tfrac{1}{\alpha}\overline{Y} .$$

b. Since $E(Y) = \alpha\theta$ and $V(Y) = \alpha\theta^2$, $E(\hat{\theta}) = \theta$, $V(\hat{\theta}) = \theta^2/(n\alpha)$.

c. Since $\overline{Y}$ is a consistent estimator of $\mu = \alpha\theta$, it is clear that $\hat{\theta}$ must be consistent for θ.
d. From the likelihood function, it is seen from Theorem 9.4 that $U = \sum_{i=1}^{n} Y_i$ is a sufficient statistic for θ. Since the gamma distribution is in the exponential family of distributions, U is also the minimal sufficient statistic.

e. Note that U has a gamma distribution with shape parameter $n\alpha$ and scale parameter θ. The distribution of $2U/\theta$ is chi–square with $2n\alpha$ degrees of freedom. With $n = 5$, $\alpha = 2$, $2U/\theta$ is chi–square with 20 degrees of freedom. So, with $\chi^2_{.95} = 10.8508$, $\chi^2_{.05} = 31.4104$, a 90% CI for θ is $\left(\frac{2\sum_{i=1}^{n} Y_i}{31.4104}, \frac{2\sum_{i=1}^{n} Y_i}{10.8508} \right)$.

9.86 First, similar to Example 9.15, the MLEs of μ_1 and μ_2 are $\hat{\mu}_1 = \overline{X}$ and $\hat{\mu}_2 = \overline{Y}$. To estimate σ^2, the likelihood is

$$L(\sigma^2) = \frac{1}{(2\pi)^{(m+n)/2}\sigma^{m+n}} \exp\left\{ -\tfrac{1}{2}\left[\sum_{i=1}^{m} \left(\frac{x_i - \mu_1}{\sigma} \right)^2 - \sum_{i=1}^{n} \left(\frac{y_i - \mu_2}{\sigma} \right)^2 \right] \right\}.$$

The log–likelihood is

$$\ln L(\sigma^2) = K - (m+n)\ln\sigma - \tfrac{1}{2\sigma^2}\left[\sum_{i=1}^{m} (x_i - \mu_1)^2 - \sum_{i=1}^{n} (y_i - \mu_2)^2 \right]$$

By differentiating and setting this quantity equal to 0, we obtain

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{m} (x_i - \mu_1)^2 - \sum_{i=1}^{n} (y_i - \mu_2)^2}{m+n}.$$

As in Example 9.15, the MLEs of μ_1 and μ_2 can be used in the above to arrive at the MLE for σ^2:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{m} (X_i - \overline{X})^2 - \sum_{i=1}^{n} (Y_i - \overline{Y})^2}{m+n}.$$

9.87 Let Y_1 = # of candidates favoring candidate A, Y_2 = # of candidate favoring candidate B, and Y_3 = # of candidates favoring candidate C. Then, (Y_1, Y_2, Y_3) is trinomial with parameters (p_1, p_2, p_3) and sample size n. Thus, the likelihood $L(p_1, p_2)$ is simply the probability mass function for the trinomial (recall that ($p_3 = 1 - p_1 - p_2$):

$$L(p_1, p_2) = \tfrac{n!}{n_1!n_2!n_3!} p_1^{y_1} p_2^{y_2} (1 - p_1 - p_2)^{y_3}$$

This can easily be jointly maximized with respect to p_1 and p_2 to obtain the MLEs $\hat{p}_1 = Y_1/n$, $\hat{p}_2 = Y_2/n$, and so $\hat{p}_3 = Y_3/n$.

For the given data, we have $\hat{p}_1 = .30$, $\hat{p}_2 = .38$, and $\hat{p}_3 = .32$. Thus, the point estimate of $p_1 - p_2$ is $.30 - .38 = -.08$. From Theorem 5.13, we have that $V(Y_i) = np_iq_i$ and $\mathrm{Cov}(Y_i, Y_j) = -np_ip_j$. A two–standard–deviation error bound can be found by

$$2\sqrt{V(\hat{p}_1 - \hat{p}_2)} = 2\sqrt{V(\hat{p}_1) + V(\hat{p}_2) - 2Cov(\hat{p}_1, \hat{p}_2)} = 2\sqrt{p_1q_1/n + p_2q_2/n + 2p_1p_2/n}.$$

This can be estimated by using the MLEs found above. By plugging in the estimates, error bound of .1641 is obtained.

9.88 The likelihood function is $L(\theta) = (\theta+1)^n \left(\prod_{i=1}^n y_i\right)^\theta$. The MLE is $\hat{\theta} = -n / \sum_{i=1}^n \ln Y_i$. This is a different estimator that the MOM estimator from Ex. 9.69, however note that the MLE is a function of the sufficient statistic.

9.89 Note that the likelihood is simply the mass function for Y: $L(p) = \binom{2}{y} p^y (1-p)^{2-y}$. By the ML criteria, we choose the value of p that maximizes the likelihood. If $Y = 0$, $L(p)$ is maximized at $p = .25$. If $Y = 2$, $L(p)$ is maximized at $p = .75$. But, if $Y = 1$, $L(p)$ has the same value at both $p = .25$ and $p = .75$; that is, $L(.25) = L(.75)$ for $y = 1$. Thus, for this instance the MLE is not unique.

9.90 Under the hypothesis that $p_W = p_M = p$, then Y = # of people in the sample who favor the issue is binomial with success probability p and $n = 200$. Thus, by Example 9.14, the MLE for p is $\hat{p} = Y/n$ and the sample estimate is 55/200.

9.91 Refer to Ex. 9.83 and Example 9.16. Let $\gamma = 2\theta$. Then, the MLE for γ is $\hat{\gamma} = Y_{(n)}$ and by the invariance principle the MLE for θ is $\hat{\theta} = Y_{(n)}/2$.

9.92 **a.** Following the hint, the MLE of θ is $\hat{\theta} = Y_{(n)}$.

b. From Ex. 9.63, $f_{(n)}(y) = 3ny^{3n-1}/\theta^{3n}$, $0 \le y \le \theta$. The distribution of $T = Y_{(n)}/\theta$ is

$$f_T(t) = 3nt^{3n-1},\ 0 \le t \le 1.$$

Since this distribution doesn't depend on θ, T is a pivotal quantity.

c. (Similar to Ex. 8.132) Constants a and b can be found to satisfy $P(a < T < b) = 1 - \alpha$ such that $P(T < a) = P(T > b) = \alpha/2$. Using the density function from part b, these are given by $a = (\alpha/2)^{1/(3n)}$ and $b = (1-\alpha/2)^{1/(3n)}$. So, we have

$$1 - \alpha = P(a < Y_{(n)}/\theta < b) = P\left(Y_{(n)}/b < \theta < Y_{(n)}/a\right).$$

Thus, $\left(\dfrac{Y_{(n)}}{(1-\alpha/2)^{1/(3n)}}, \dfrac{Y_{(n)}}{(\alpha/2)^{1/(3n)}}\right)$ is a $(1-\alpha)100\%$ CI for θ.

9.93 **a.** Following the hint, the MLE for θ is $\hat{\theta} = Y_{(1)}$.

b. Since $F(y \mid \theta) = 1 - 2\theta^2 y^{-2}$, the density function for $Y_{(1)}$ is easily found to be

$$g_{(1)}(y) = 2n\theta^{2n} y^{-(2n+1)},\ y > \theta.$$

If we consider the distribution of $T = \theta/Y_{(1)}$, the density function of T can be found to be

$$f_T(t) = 2nt^{2n-1},\ 0 < t < 1.$$

c. (Similar to Ex. 9.92) Constants a and b can be found to satisfy $P(a < T < b) = 1 - \alpha$ such that $P(T < a) = P(T > b) = \alpha/2$. Using the density function from part b, these are given by $a = (\alpha/2)^{1/(2n)}$ and $b = (1-\alpha/2)^{1/(2n)}$. So, we have

$$1 - \alpha = P(a < \theta/Y_{(1)} < b) = P\left(aY_{(1)} < \theta < bY_{(1)}\right).$$

Thus, $\left[(\alpha/2)^{1/(2n)}Y_{(1)},\ (1-\alpha/2)^{1/(2n)}Y_{(1)}\right]$ is a $(1 - \alpha)100\%$ CI for θ.

9.94 Let $\beta = t(\theta)$ so that $\theta = t^{-1}(\beta)$. If the likelihood is maximized at $\hat{\theta}$, then $L(\hat{\theta}) \geq L(\theta)$ for all θ. Define $\hat{\beta} = t(\hat{\theta})$ and denote the likelihood as a function of β as $L_1(\beta) = L(t^{-1}(\beta))$. Then, for any β,

$$L_1(\beta) = L(t^{-1}(\beta)) = L(\theta) \leq L(\hat{\theta}) = L(t^{-1}(\hat{\beta})) = L_1(\hat{\beta}).$$

So, the MLE of β is $\hat{\beta}$ and so the MLE of $t(\theta)$ is $t(\hat{\theta})$.

9.95 The quantity to be estimated is $R = p/(1 - p)$. Since $\hat{p} = Y/n$ is the MLE of p, by the invariance principle the MLE for R is $\hat{R} = \hat{p}/(1 - \hat{p})$.

9.96 From Ex. 9.15, the MLE for σ^2 was found to be $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$. By the invariance property, the MLE for σ is $\hat{\sigma} = \sqrt{\hat{\sigma}^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2}$.

9.97 **a.** Since $\mu_1' = 1/p$, the MOM estimator for p is $\hat{p} = 1/m_1' = 1/\overline{Y}$.

b. The likelihood function is $L(p) = p^n(1-p)^{\Sigma y_i - n}$ and the log–likelihood is

$$\ln L(p) = n\ln p + \left(\sum_{i=1}^{n} y_i - n\right)\ln(1-p).$$

Differentiating, we have

$$\tfrac{d}{dp}\ln L(p) = \tfrac{n}{p} - \tfrac{1}{1-p}\left(\sum_{i=1}^{n} y_i - n\right).$$

Equating this to 0 and solving for p, we obtain the MLE $\hat{p} = 1/\overline{Y}$, which is the same as the MOM estimator found in part a.

9.98 Since $\ln p(y \mid p) = \ln p + (y-1)\ln(1-p)$,

$$\tfrac{d}{dp}\ln p(y \mid p) = 1/p - (y-1)/(1-p)$$

$$\tfrac{d^2}{dp^2}\ln p(y \mid p) = -1/p^2 - (y-1)/(1-p)^2.$$

Then,

$$-E\left[\tfrac{d^2}{dp^2}\ln p(Y \mid p)\right] = -E\left[-1/p^2 - (Y-1)/(1-p)^2\right] = \frac{1}{p^2(1-p)}.$$

Therefore, the approximate (limiting) variance of the MLE (as given in Ex. 9.97) is given by

$$V(\hat{p}) \approx \frac{p^2(1-p)}{n}.$$

9.99 From Ex. 9.18, the MLE for $t(p) = p$ is $\hat{p} = Y/n$ and with $-E\left[\frac{d^2}{dp^2}\ln p(Y \mid p)\right] = \frac{1}{p(1-p)}$, a 100(1 – α)% CI for p is $\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$. This is the same CI for p derived in Section 8.6.

9.100 In Ex. 9.81, it was shown that $\overline{Y}^2$ is the MLE of $t(\theta) = \theta^2$. It is easily found that for the exponential distribution with mean θ,

$$-E\left[\frac{d^2}{dp^2}\ln f(Y \mid \theta)\right] = \frac{1}{\theta^2}.$$

Thus, since $\frac{d}{d\theta}t(\theta) = 2\theta$, we have an approximate (large sample) 100(1 – α)% CI for θ as

$$\overline{Y}^2 \pm z_{\alpha/2}\sqrt{\left.\left(\frac{(2\theta)^2}{n\frac{1}{\theta^2}}\right)\right|_{\theta=\hat{\theta}}} = \overline{Y}^2 \pm z_{\alpha/2}\left(\frac{2\overline{Y}^2}{\sqrt{n}}\right).$$

9.101 From Ex. 9.80, the MLE for $t(\lambda) = \exp(-\lambda)$ is $t(\hat{\lambda}) = \exp(-\hat{\lambda}) = \exp(-\overline{Y})$. It is easily found that for the Poisson distribution with mean λ,

$$-E\left[\frac{d^2}{dp^2}\ln p(Y \mid \lambda)\right] = \frac{1}{\lambda}.$$

Thus, since $\frac{d}{d\lambda}t(\lambda) = -\exp(-\lambda)$, we have an approximate 100(1 – α)% CI for λ as

$$\exp(-\overline{Y}) \pm z_{\alpha/2}\sqrt{\left.\frac{\exp(-2\lambda)}{n\frac{1}{\lambda}}\right|_{\lambda=\overline{Y}}} = \exp(-\overline{Y}) \pm z_{\alpha/2}\sqrt{\frac{\overline{Y}\exp(-2\overline{Y})}{n}}.$$

9.102 With $n = 30$ and $\overline{y} = 4.4$, the maximum likelihood estimate of p is 1/(4.4) = .2273 and an approximate 90% CI for p is

$$\hat{p} \pm z_{.025}\sqrt{\frac{\hat{p}^2(1-\hat{p})}{n}} = .2273 \pm 1.96\sqrt{\frac{(.2273)^2(.7727)}{30}} = .2273 \pm .0715 \text{ or } (.1558, .2988).$$

9.103 The Rayleigh distribution is a special case of the (Weibull) distribution from Ex. 9.82. Also see Example 9.7

a. From Ex. 9.82 with $r = 2$, $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} Y_i^2$.

b. It is easily found that for the Rayleigh distribution with parameter θ,

$$\frac{d^2}{dp^2}\ln f(Y \mid \theta) = \frac{1}{\theta^2} - \frac{2Y^2}{\theta^3}.$$

Since $E(Y^2) = \theta$, $-E\left[\frac{d^2}{dp^2}\ln f(Y \mid \theta)\right] = \frac{1}{\theta^2}$ and so $V(\hat{\theta}) \approx \theta^2/n$.

9.104 **a.** MOM: $\mu_1' = E(Y) = \theta + 1$, so $\hat{\theta}_1 = m_1' - 1 = \overline{Y} - 1$.

b. MLE: $\hat{\theta}_2 = Y_{(1)}$, the first order statistic.

c. The estimator $\hat{\theta}_1$ is unbiased since $E(\hat{\theta}_1) = E(\overline{Y}) - 1 = \theta + 1 - 1 = \theta$. The distribution of $Y_{(1)}$ is $g_{(1)}(y) = ne^{-n(y-\theta)}$, $y > \theta$. So, $E(Y_{(1)}) = E(\hat{\theta}_2) = \frac{1}{n} + \theta$. Thus, $\hat{\theta}_2$ is not unbiased but $\hat{\theta}_2^* = Y_{(1)} - \frac{1}{n}$ is unbiased for θ.

The efficiency of $\hat{\theta}_1 = \overline{Y} - 1$ relative to $\hat{\theta}_2^* = Y_{(1)} - \frac{1}{n}$ is given by

$$\text{eff}(\hat{\theta}_1, \hat{\theta}_2^*) = \frac{V(\hat{\theta}_2^*)}{V(\hat{\theta}_1)} = \frac{V(Y_{(1)} - \frac{1}{n})}{V(\overline{Y} - 1)} = \frac{V(Y_{(1)})}{V(\overline{Y})} = \frac{\frac{1}{n^2}}{\frac{1}{n}} = \frac{1}{n}.$$

9.105 From Ex. 9.38, we must solve

$$\tfrac{d^2 \ln L}{d\sigma^2} = \tfrac{-n}{2\sigma^2} + \tfrac{\Sigma(y_i - \mu)^2}{2\sigma^4} = 0, \quad \text{so} \quad \hat{\sigma}^2 = \tfrac{\Sigma(y_i - \mu)^2}{n}.$$

9.106 Following the method used in Ex. 9.65, construct the random variable T such that

$T = 1$ if $Y_1 = 0$ and $T = 0$ otherwise

Then, $E(T) = P(T = 1) = P(Y_1 = 0) = \exp(-\lambda)$. So, T is unbiased for $\exp(-\lambda)$. Now, we know that $W = \sum_{i=1}^n Y_i$ is sufficient for λ, and so it is also sufficient for $\exp(-\lambda)$. Recalling that W has a Poisson distribution with mean $n\lambda$,

$$E(T \mid W = w) = P(T = 1 \mid W = w) = P(Y_1 = 0 \mid W = w) = \frac{P(Y_1 = 0, W = w)}{P(W = w)}$$

$$= \frac{P(Y_1 = 0)P\left(\sum_{i=2}^n Y_i = w\right)}{P(W = w)} = \frac{e^{-\lambda}\left(e^{-(n-1)\lambda}\frac{[(n-1)\lambda]^w}{w!}\right)}{e^{-n\lambda}\frac{(n\lambda)^w}{w!}} = \left(1 - \tfrac{1}{n}\right)^w.$$

Thus, the MVUE is $\left(1 - \frac{1}{n}\right)^{\Sigma Y_i}$. Note that in the above we used the result that $\sum_{i=2}^n Y_i$ is Poisson with mean $(n-1)\lambda$.

9.107 The MLE of θ is $\hat{\theta} = \overline{Y}$. By the invariance principle for MLEs, the MLE of $\overline{F}(t)$ is $\hat{\overline{F}}(t) = \exp(-t/\overline{Y})$.

9.108 **a.** $E(V) = P(Y_1 > t) = 1 - F(t) = \exp(-t/\theta)$. Thus, V is unbiased for $\exp(-t/\theta)$.

b. Recall that U has a gamma distribution with shape parameter n and scale parameter θ. Also, $U - Y_1 = \sum_{i=2}^n Y_i$ is gamma with shape parameter $n - 1$ and scale parameter θ, and since Y_1 and $U - Y_1$ are independent,

$$f(y_1, u - y_1) = \left(\tfrac{1}{\theta}e^{-y_1/\theta}\right)\tfrac{1}{\Gamma(n-1)\theta^{n-1}}(u - y_1)^{n-2}e^{-(u-y_1)/\theta}, \ 0 \le y_1 \le u < \infty.$$

Next, apply the transformation $z = u - y_1$ such that $u = z + y_1$ to get the joint distribution

$$f(y_1, u) = \tfrac{1}{\Gamma(n-1)\theta^n}(u - y_1)^{n-2}e^{-u/\theta}, \ 0 \le y_1 \le u < \infty.$$

Now, we have

$$f(y_1 \mid u) = \frac{f(y_1, u)}{f(u)} = \left(\frac{n-1}{u^{n-1}}\right)(u - y_1)^{n-2},\ 0 \le y_1 \le u < \infty.$$

c. $E(V \mid U) = P(Y_1 > t \mid U = u) = \int_t^u \left(\frac{n-1}{u^{n-1}}\right)(u-y_1)^{n-2} dy_1 = \int_t^u \left(\frac{n-1}{u}\right)\left(1 - \frac{y_1}{u}\right)^{n-2} dy_1$

$$= -\left(1 - \frac{y_1}{u}\right)^{n-1}\Bigg|_t^u = \left(1 - \frac{t}{u}\right)^{n-1}.$$

So, the MVUE is $\left(1 - \frac{t}{U}\right)^{n-1}$.

9.109 Let $Y_1, Y_2, \ldots, Y_n$ represent the (independent) values drawn on each of the n draws. Then, the probability mass function for each Y_i is

$$P(Y_i = k) = \tfrac{1}{N},\ k = 1, 2, \ldots, N.$$

a. Since $\mu_1' = E(Y) = \sum_{k=1}^{N} kP(Y = k) = \sum_{k=1}^{N} k\,\frac{1}{N} = \frac{N(N+1)}{2N} = \frac{N+1}{2}$, the MOM estimator of N is $\frac{\hat{N}_1 + 1}{2} = \overline{Y}$ or $\hat{N}_1 = 2\overline{Y} - 1$.

b. First, $E(\hat{N}_1) = 2E(\overline{Y}) - 1 = 2\left(\frac{N+1}{2}\right) - 1 = N$, so $\hat{N}_1$ is unbiased. Now, since $E(Y^2) = \sum_{k=1}^{N} k^2\,\frac{1}{N} = \frac{N(N+1)(2N+1)}{6N} = \frac{(N+1)(2N+1)}{6}$, we have that $V(Y) = \frac{(N+1)(N-1)}{12}$. Thus, $V(\hat{N}_1) = 4V(\overline{Y}) = 4\left(\frac{(N+1)(N-1)}{12n}\right) = \frac{N^2-1}{3n}$.

9.110 **a.** Following Ex. 9.109, the likelihood is

$$L(N) = \tfrac{1}{N^n}\prod_{i=1}^{n} I(y_i \in \{1, 2, \ldots, N\}) = \tfrac{1}{N^n} I(y_{(n)} \le N).$$

In order to maximize L, N should be chosen as small as possible subject to the constraint that $y_{(n)} \le N$. Thus $\hat{N}_2 = Y_{(n)}$.

b. Since $P(\hat{N}_2 \le k) = P(Y_{(n)} \le k) = P(Y_1 \le k)\cdots P(Y_n \le k) = \left(\frac{k}{N}\right)^n$, so $P(\hat{N}_2 \le k-1) = \left(\frac{k-1}{N}\right)^n$ and $P(\hat{N}_2 = k) = \left(\frac{k}{N}\right)^n - \left(\frac{k-1}{N}\right)^n = N^{-n}[k^n - (k-1)^n]$. So,

$$E(\hat{N}_2) = N^{-n}\sum_{k=1}^{N} k[k^n - (k-1)^n] = N^{-n}\sum_{k=1}^{N}[k^{n+1} - (k-1)^{n+1} - (k-1)^n]$$
$$= N^{-n}\left[N^{n+1} - \sum_{k=1}^{N}(k-1)^n\right].$$

Consider $\sum_{k=1}^{N}(k-1)^n = 0^n + 1^n + 2^n + \ldots + (N-1)^n$. For large N, this is approximately the area beneath the curve $f(x) = x^n$ from $x = 0$ to $x = N$, or $\sum_{k=1}^{N}(k-1)^n \approx \int_0^N x^n dx = \frac{N^{n+1}}{n+1}$.

Thus, $E(\hat{N}_2) \approx N^{-n}[N^{n+1} - \frac{N^{n+1}}{n+1}] = \frac{n}{n+1}N$ and $\hat{N}_3 = \frac{n+1}{n}\hat{N}_2 = \frac{n+1}{n}Y_{(n)}$ is approximately unbiased for N.

c. $V(\hat{N}_2)$ is given, so $V(\hat{N}_3) = \left(\frac{n+1}{n}\right)^2 V(\hat{N}_2) = \frac{N^2}{n(n+2)}$.

d. Note that, for $n > 1$,

$$\frac{V(\hat{N}_1)}{V(\hat{N}_3)} = \frac{n(n+2)}{3n}\frac{(N^2-1)}{N^2} = \frac{n+2}{3}\left(1-\frac{1}{N^2}\right) > 1,$$

since for large N, $\left(1-\frac{1}{N^2}\right) \approx 1$

9.111 The (approximately) unbiased estimate of N is $\hat{N}_3 = \frac{n+1}{n}Y_{(n)} = \frac{6}{5}(210) = 252$ and an approximate error bound is given by

$$2\sqrt{V(\hat{N}_3)} \approx 2\sqrt{\frac{N^2}{n(n+2)}} \approx 2\sqrt{\frac{(252)^2}{5(7)}} = 85.192.$$

9.112 **a.** (Refer to Section 9.3.) By the Central Limit Theorem, $\dfrac{\overline{Y}-\lambda}{\sqrt{\lambda/n}}$ converges to a standard normal variable. Also, $\overline{Y}/\lambda$ converges in probability to 1 by the Law of Large Numbers, as does $\sqrt{\overline{Y}/\lambda}$. So, the quantity

$$W_n = \frac{\dfrac{\overline{Y}-\lambda}{\sqrt{\lambda/n}}}{\sqrt{\overline{Y}/\lambda}} = \frac{\overline{Y}-\lambda}{\sqrt{\overline{Y}/n}}$$

converges to a standard normal distribution.

b. By part **a**, an approximate $(1-\alpha)100\%$ CI for λ is $\overline{Y} \pm z_{\alpha/2}\sqrt{\overline{Y}/n}$.