

8.1. In this problem, we prove

$$\text{MSE}(\hat{\theta}) = E[(\hat{\theta} - \theta)^2] = V(\hat{\theta}) + [B(\hat{\theta})]^2,$$

where

$$B(\hat{\theta}) = E(\hat{\theta}) - \theta$$

is the bias of $\hat{\theta}$ as a point estimator of θ . Write the expectation above as

$$E[(\hat{\theta} - \theta)^2] = E\{\underbrace{[\hat{\theta} - E(\hat{\theta}) + E(\hat{\theta}) - \theta]^2}_{\text{subtract and add}}\}.$$

Now write

$$[\hat{\theta} - E(\hat{\theta}) + E(\hat{\theta}) - \theta]^2 = (a + b)^2,$$

where $a = \hat{\theta} - E(\hat{\theta})$ and $b = E(\hat{\theta}) - \theta$. Note that

$$(a + b)^2 = a^2 + 2ab + b^2 = \underbrace{[\hat{\theta} - E(\hat{\theta})]^2}_{(*)} + 2\underbrace{[\hat{\theta} - E(\hat{\theta})][E(\hat{\theta}) - \theta]}_{(**)} + \underbrace{[E(\hat{\theta}) - \theta]^2}_{(***)}.$$

Now, take expectations of each piece and add up the results. For $(*)$, we have

$$E\{[\hat{\theta} - E(\hat{\theta})]^2\} = V(\hat{\theta}).$$

This follows from the definition of the variance of a random variable (here, $\hat{\theta}$). For $(**)$, note that $E(\hat{\theta})$ and θ are constants; therefore, we have

$$\begin{aligned} E\{2[\hat{\theta} - E(\hat{\theta})][E(\hat{\theta}) - \theta]\} &= 2\{E[\hat{\theta}E(\hat{\theta})] - E(\hat{\theta})\theta - E[E(\hat{\theta})E(\hat{\theta})] + E[E(\hat{\theta})\theta]\} \\ &= 2\{E(\hat{\theta})E(\hat{\theta}) - \theta E(\hat{\theta}) - E(\hat{\theta})E(\hat{\theta}) + E(\hat{\theta})\theta\} = 2(0) = 0. \end{aligned}$$

Therefore, the cross product term $2[\hat{\theta} - E(\hat{\theta})][E(\hat{\theta}) - \theta]$ has expectation zero. For $(***)$, we have

$$E\{[E(\hat{\theta}) - \theta]^2\} = [E(\hat{\theta}) - \theta]^2 = [B(\hat{\theta})]^2,$$

because $E(\hat{\theta})$ and θ are both constants; hence $[E(\hat{\theta}) - \theta]^2$ is constant. This shows algebraically that

$$\text{MSE}(\hat{\theta}) = E[(\hat{\theta} - \theta)^2] = V(\hat{\theta}) + [B(\hat{\theta})]^2.$$

Of course, if $\hat{\theta}$ is an unbiased estimator of θ , then $B(\hat{\theta}) = E(\hat{\theta}) - \theta = 0$ and $\text{MSE}(\hat{\theta}) = V(\hat{\theta})$.

8.6. We are given

$$\begin{aligned} E(\hat{\theta}_1) &= E(\hat{\theta}_2) = \theta \\ V(\hat{\theta}_1) &= \sigma_1^2 \\ V(\hat{\theta}_2) &= \sigma_2^2. \end{aligned}$$

Note that both $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased estimators of θ . Suppose $a \in \mathbb{R}$ is a constant. For part (a), we have

$$E(\hat{\theta}_3) = E[a\hat{\theta}_1 + (1 - a)\hat{\theta}_2] = aE(\hat{\theta}_1) + (1 - a)E(\hat{\theta}_2) = a\theta + (1 - a)\theta = \theta.$$

Therefore, the linear (convex) combination $\hat{\theta}_3 = a\hat{\theta}_1 + (1 - a)\hat{\theta}_2$ is also an unbiased estimator of θ .

(b) The variance of $\hat{\theta}_3 = a\hat{\theta}_1 + (1-a)\hat{\theta}_2$ is

$$V(\hat{\theta}_3) = V[a\hat{\theta}_1 + (1-a)\hat{\theta}_2] = a^2V(\hat{\theta}_1) + (1-a)^2V(\hat{\theta}_2) + 2a(1-a)\text{Cov}(\hat{\theta}_1, \hat{\theta}_2).$$

However, because $\hat{\theta}_1$ and $\hat{\theta}_2$ are independent (by assumption), we have $\text{Cov}(\hat{\theta}_1, \hat{\theta}_2) = 0$ and therefore

$$V(\hat{\theta}_3) = a^2V(\hat{\theta}_1) + (1-a)^2V(\hat{\theta}_2) = \underbrace{a^2\sigma_1^2 + (1-a)^2\sigma_2^2}_{=g(a), \text{ say}}.$$

Let $g(a) = a^2\sigma_1^2 + (1-a)^2\sigma_2^2$; i.e., view $V(\hat{\theta}_3)$ as a function of a . We want to find the value of a that minimizes $g(a)$. Note that

$$\begin{aligned} \frac{d}{da}g(a) &= 2a\sigma_1^2 + 2(1-a)(-1)\sigma_2^2 \stackrel{\text{set}}{=} 0 \implies 2a\sigma_1^2 - 2\sigma_2^2 + 2a\sigma_2^2 = 0 \\ &\implies 2a\sigma_1^2 + 2a\sigma_2^2 = 2\sigma_2^2 \\ &\implies a\sigma_1^2 + a\sigma_2^2 = \sigma_2^2 \implies a = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}. \end{aligned}$$

This is the solution to

$$\frac{d}{da}g(a) = 0;$$

i.e.,

$$a = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

is a first-order critical point of $g(a)$. Use the second derivative test to determine if this critical point is a minimizer; we have

$$\frac{d^2}{da^2}g(a) = \frac{d}{da}[2a\sigma_1^2 + 2(1-a)(-1)\sigma_2^2] = 2\sigma_1^2 + 2(-1)(-1)\sigma_2^2 = 2(\sigma_1^2 + \sigma_2^2) > 0,$$

because variances are positive. Because $(d^2/da^2)g(a) > 0$ for all $a \in \mathbb{R}$, the function $g(a)$ is concave up and therefore the first-order critical point above is a minimizer of $g(a)$. Therefore, $a = \sigma_2^2/(\sigma_1^2 + \sigma_2^2)$ minimizes $V(\hat{\theta}_3)$.

8.9. In this problem, Y_1, Y_2, \dots, Y_n is an iid sample from an exponential population distribution with mean $\beta = \theta + 1$. Therefore, we know

$$E(\bar{Y}) = \mu = \beta = \theta + 1.$$

Therefore,

$$E(\bar{Y} - 1) = \theta + 1 - 1 = \theta.$$

This shows $\hat{\theta} = \bar{Y} - 1$ is an unbiased estimator of θ .

8.12. In this problem, Y_1, Y_2, \dots, Y_n is an iid sample from a uniform population distribution from θ to $\theta + 1$; i.e., the population distribution is $\mathcal{U}(\theta, \theta + 1)$. For part (a), note the population mean is

$$\mu = E(Y) = \frac{\theta + (\theta + 1)}{2} = \theta + \frac{1}{2}.$$

Therefore, we know

$$E(\bar{Y}) = \mu = \theta + \frac{1}{2}.$$

This shows \bar{Y} is a biased estimator of θ . The bias of \bar{Y} as a point estimator of θ is

$$B(\bar{Y}) = E(\bar{Y}) - \theta = \theta + \frac{1}{2} - \theta = \frac{1}{2}.$$

(b) Note that

$$E(\bar{Y}) = \theta + \frac{1}{2} \implies E\left(\bar{Y} - \frac{1}{2}\right) = \theta + \frac{1}{2} - \frac{1}{2} = \theta.$$

This shows $\hat{\theta} = \bar{Y} - \frac{1}{2}$ is an unbiased estimator of θ .

(c) The MSE of \bar{Y} is

$$\text{MSE}(\bar{Y}) = V(\bar{Y}) + [B(\bar{Y})]^2.$$

We have already calculated $B(\bar{Y}) = \frac{1}{2}$. Recall

$$V(\bar{Y}) = \frac{\sigma^2}{n},$$

where $\sigma^2 = V(Y)$ is the population variance. Using what we know about the uniform distribution, we have

$$\sigma^2 = V(Y) = \frac{[(\theta + 1) - \theta]^2}{12} = \frac{1}{12}.$$

Therefore,

$$\text{MSE}(\bar{Y}) = V(\bar{Y}) + [B(\bar{Y})]^2 = \frac{(1/12)}{n} + \left(\frac{1}{2}\right)^2 = \frac{1}{12n} + \frac{1}{4}.$$

8.13. In this problem, we have $Y \sim b(n, p)$. We know

$$E\left(\frac{Y}{n}\right) = \frac{E(Y)}{n} = \frac{np}{n} = p;$$

i.e., the usual sample proportion Y/n is an unbiased estimator of p . In part (a), consider the estimator

$$n\left(\frac{Y}{n}\right)\left(1 - \frac{Y}{n}\right) = Y\left(1 - \frac{Y}{n}\right) = Y - \frac{Y^2}{n}.$$

We have

$$E\left[n\left(\frac{Y}{n}\right)\left(1 - \frac{Y}{n}\right)\right] = E\left(Y - \frac{Y^2}{n}\right) = E(Y) - \frac{E(Y^2)}{n}.$$

We know $E(Y) = np$ and

$$E(Y^2) = V(Y) + [E(Y)]^2 = np(1 - p) + (np)^2.$$

Therefore,

$$\begin{aligned} E\left[n\left(\frac{Y}{n}\right)\left(1 - \frac{Y}{n}\right)\right] &= E(Y) - \frac{E(Y^2)}{n} = np - \frac{np(1 - p) + (np)^2}{n} \\ &= np - p(1 - p) - np^2 \\ &= np - np^2 - p(1 - p) \\ &= np(1 - p) - p(1 - p) = V(Y) - p(1 - p). \end{aligned}$$

This shows that

$$n \left(\frac{Y}{n} \right) \left(1 - \frac{Y}{n} \right)$$

is a biased estimator of $V(Y) = np(1-p)$.

(b) In part (a), we showed

$$E \left[n \left(\frac{Y}{n} \right) \left(1 - \frac{Y}{n} \right) \right] = V(Y) - p(1-p).$$

Pre-multiply the estimator in part (a) by n and consider

$$n^2 \left(\frac{Y}{n} \right) \left(1 - \frac{Y}{n} \right).$$

We have

$$\begin{aligned} E \left[n^2 \left(\frac{Y}{n} \right) \left(1 - \frac{Y}{n} \right) \right] &= n E \left[n \left(\frac{Y}{n} \right) \left(1 - \frac{Y}{n} \right) \right] = n [V(Y) - p(1-p)] \\ &= nV(Y) - np(1-p) \\ &= nV(Y) - V(Y) = (n-1)V(Y). \end{aligned}$$

Therefore,

$$E \left[\frac{n^2 \left(\frac{Y}{n} \right) \left(1 - \frac{Y}{n} \right)}{n-1} \right] = \left(\frac{1}{n-1} \right) E \left[n^2 \left(\frac{Y}{n} \right) \left(1 - \frac{Y}{n} \right) \right] = \left(\frac{1}{n-1} \right) (n-1)V(Y) = V(Y).$$

This shows

$$\frac{n^2 \left(\frac{Y}{n} \right) \left(1 - \frac{Y}{n} \right)}{n-1} = \left(\frac{n^2}{n-1} \right) \left(\frac{Y}{n} \right) \left(1 - \frac{Y}{n} \right)$$

is an unbiased estimator of the population variance $V(Y) = np(1-p)$.

8.15. In this problem, Y_1, Y_2, \dots, Y_n is an iid sample from a Pareto population distribution with pdf

$$f_Y(y) = \begin{cases} \frac{3\beta^3}{y^4}, & y \geq \beta \\ 0, & \text{otherwise,} \end{cases}$$

where $\beta > 0$ is unknown. We consider the point estimator $\hat{\beta} = Y_{(1)}$, the minimum order statistic. If we are going to derive the bias and MSE of $\hat{\beta}$, then we need to know its sampling distribution. Recall the pdf of $Y_{(1)}$, in general, is

$$f_{Y_{(1)}}(y) = n f_Y(y) [1 - F_Y(y)]^{n-1},$$

where $F_Y(y)$ is the cdf of Y . Therefore, we need to derive the cdf of Y first. For $y < \beta$, we have

$$F_Y(y) = \int_{-\infty}^y f_Y(t) dt = \int_{-\infty}^y 0 dt = 0.$$

For $y \geq \beta$, we have

$$\begin{aligned} F_Y(y) &= \int_{-\infty}^y f_Y(t) dt = \int_{-\infty}^{\beta} 0 dt + \int_{\beta}^y \frac{3\beta^3}{t^4} dt \\ &= 0 + 3\beta^3 \left(-\frac{1}{3} \right) \frac{1}{t^3} \Big|_{\beta}^y = \frac{\beta^3}{t^3} \Big|_{\beta}^y = 1 - \left(\frac{\beta}{y} \right)^3. \end{aligned}$$

Summarizing, the population cdf is

$$F_Y(y) = \begin{cases} 0, & y < \beta \\ 1 - \left(\frac{\beta}{y} \right)^3, & y \geq \beta. \end{cases}$$

Therefore, the pdf of $Y_{(1)}$, for $y \geq \beta$, is

$$f_{Y_{(1)}}(y) = n f_Y(y) [1 - F_Y(y)]^{n-1} = n \times \frac{3\beta^3}{y^4} \times \left\{ 1 - \left[1 - \left(\frac{\beta}{y} \right)^3 \right] \right\}^{n-1} = \frac{3n\beta^3}{y^4} \times \left(\frac{\beta}{y} \right)^{3n-3} = \frac{3n\beta^{3n}}{y^{3n+1}}.$$

Summarizing,

$$f_{Y_{(1)}}(y) = \begin{cases} \frac{3n\beta^{3n}}{y^{3n+1}}, & y \geq \beta \\ 0, & \text{otherwise.} \end{cases}$$

(a) Let's calculate the expected value of $\hat{\beta} = Y_{(1)}$. We have

$$\begin{aligned} E(\hat{\beta}) &= E(Y_{(1)}) = \int_{\mathbb{R}} y f_{Y_{(1)}}(y) dy = \int_{\beta}^{\infty} y \times \frac{3n\beta^{3n}}{y^{3n+1}} dy \\ &= 3n\beta^{3n} \int_{\beta}^{\infty} \frac{1}{y^{3n}} dy \\ &= 3n\beta^{3n} \left(-\frac{1}{3n-1} \right) \frac{1}{y^{3n-1}} \Big|_{\beta}^{\infty} \\ &= \left(\frac{3n\beta^{3n}}{3n-1} \right) \frac{1}{y^{3n-1}} \Big|_{\infty}^{\beta} = \left(\frac{3n\beta^{3n}}{3n-1} \right) \left(\frac{1}{\beta^{3n-1}} - 0 \right) = \left(\frac{3n}{3n-1} \right) \beta. \end{aligned}$$

Therefore, $\hat{\beta} = Y_{(1)}$ is a (positively) biased estimator of β . The bias of $\hat{\beta} = Y_{(1)}$ is

$$B(\hat{\beta}) = B(Y_{(1)}) = E(Y_{(1)}) - \beta = \left(\frac{3n}{3n-1} \right) \beta - \beta = \left(\frac{1}{3n-1} \right) \beta.$$

(b) We want to calculate

$$\text{MSE}(\hat{\beta}) = \text{MSE}(Y_{(1)}) = V(Y_{(1)}) + [B(Y_{(1)})]^2.$$

Therefore, we need to calculate

$$V(Y_{(1)}) = E(Y_{(1)}^2) - [E(Y_{(1)})]^2$$

first. The second moment of $Y_{(1)}$ is

$$\begin{aligned}
 E(Y_{(1)}^2) &= \int_{\mathbb{R}} y^2 f_{Y_{(1)}}(y) dy = \int_{\beta}^{\infty} y^2 \times \frac{3n\beta^{3n}}{y^{3n+1}} dy \\
 &= 3n\beta^{3n} \int_{\beta}^{\infty} \frac{1}{y^{3n-1}} dy \\
 &= 3n\beta^{3n} \left(-\frac{1}{3n-2} \right) \frac{1}{y^{3n-2}} \Big|_{\beta}^{\infty} \\
 &= \left(\frac{3n\beta^{3n}}{3n-2} \right) \frac{1}{y^{3n-2}} \Big|_{\infty}^{\beta} = \left(\frac{3n\beta^{3n}}{3n-2} \right) \left(\frac{1}{\beta^{3n-2}} - 0 \right) = \left(\frac{3n}{3n-2} \right) \beta^2.
 \end{aligned}$$

Therefore,

$$\begin{aligned}
 V(Y_{(1)}) &= E(Y_{(1)}^2) - [E(Y_{(1)})]^2 = \left(\frac{3n}{3n-2} \right) \beta^2 - \left[\left(\frac{3n}{3n-1} \right) \beta \right]^2 \\
 &= \left[\left(\frac{3n}{3n-2} \right) - \left(\frac{3n}{3n-1} \right)^2 \right] \beta^2 = \left[\frac{3n}{(3n-1)^2(3n-2)} \right] \beta^2.
 \end{aligned}$$

Finally,

$$\begin{aligned}
 \text{MSE}(\hat{\beta}) &= \text{MSE}(Y_{(1)}) = V(Y_{(1)}) + [B(Y_{(1)})]^2 \\
 &= \left[\frac{3n}{(3n-1)^2(3n-2)} \right] \beta^2 + \left[\left(\frac{1}{3n-1} \right) \beta \right]^2 = \left[\frac{2}{(3n-1)(3n-2)} \right] \beta^2.
 \end{aligned}$$

Note: In the calculations for $V(Y_{(1)})$ and $\text{MSE}(\hat{\beta})$ above, I did not show the algebra that led to the final simplified answers.

8.16. In this problem, Y_1, Y_2, \dots, Y_n is an iid sample from a $\mathcal{N}(\mu, \sigma^2)$ population distribution where both μ and σ^2 are unknown. We know the sample variance S^2 is an unbiased estimator of σ^2 ; i.e.,

$$E(S^2) = \sigma^2.$$

However, the sample standard deviation S is not an unbiased estimator of σ ; this makes sense because $S = \sqrt{S^2}$ and the square-root function is not linear. Therefore, bias is introduced when we take square roots. In part (a), we want to calculate $E(S)$. From Result 6 in Chapter 7, recall that when Y_1, Y_2, \dots, Y_n is an iid sample from a $\mathcal{N}(\mu, \sigma^2)$, we know

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \stackrel{d}{=} \text{gamma}\left(\frac{n-1}{2}, 2\right).$$

Therefore,

$$S^2 = \frac{\sigma^2}{n-1} \left[\frac{(n-1)S^2}{\sigma^2} \right] \sim \text{gamma}\left(\frac{n-1}{2}, \frac{2\sigma^2}{n-1}\right).$$

Therefore, calculating $E(S)$ results from calculating

$$E(X^{\frac{1}{2}}), \text{ where } X \sim \text{gamma}\left(\frac{n-1}{2}, \frac{2\sigma^2}{n-1}\right).$$

We have

$$\begin{aligned}
 E(X^{\frac{1}{2}}) &= \int_{\mathbb{R}} x^{\frac{1}{2}} f_X(x) dx = \int_0^{\infty} x^{\frac{1}{2}} \underbrace{\frac{1}{\Gamma(\frac{n-1}{2}) \left(\frac{2\sigma^2}{n-1}\right)^{\frac{n-1}{2}}} x^{\frac{n-1}{2}-1} e^{-x/\left(\frac{2\sigma^2}{n-1}\right)}}_{\text{gamma}\left(\frac{n-1}{2}, \frac{2\sigma^2}{n-1}\right) \text{ pdf}} dx \\
 &= \frac{1}{\Gamma(\frac{n-1}{2}) \left(\frac{2\sigma^2}{n-1}\right)^{\frac{n-1}{2}}} \int_0^{\infty} \underbrace{x^{\frac{n-1}{2}-1} e^{-x/\left(\frac{2\sigma^2}{n-1}\right)}}_{\text{gamma}\left(\frac{n}{2}, \frac{2\sigma^2}{n-1}\right) \text{ kernel}} dx \\
 &= \frac{1}{\Gamma(\frac{n-1}{2}) \left(\frac{2\sigma^2}{n-1}\right)^{\frac{n-1}{2}}} \Gamma\left(\frac{n}{2}\right) \left(\frac{2\sigma^2}{n-1}\right)^{\frac{n}{2}} \\
 &= \frac{\Gamma\left(\frac{n}{2}\right) \left(\frac{2\sigma^2}{n-1}\right)^{\frac{1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)} = \left[\frac{\sqrt{2}\Gamma\left(\frac{n}{2}\right)}{\sqrt{n-1}\Gamma\left(\frac{n-1}{2}\right)} \right] \sigma.
 \end{aligned}$$

Therefore, the sample standard deviation S is a biased estimator of σ because

$$E(S) = \left[\frac{\sqrt{2}\Gamma\left(\frac{n}{2}\right)}{\sqrt{n-1}\Gamma\left(\frac{n-1}{2}\right)} \right] \sigma.$$

(b) Adjusting S to “make it unbiased” is easy. Note that

$$E(S) = \left[\frac{\sqrt{2}\Gamma\left(\frac{n}{2}\right)}{\sqrt{n-1}\Gamma\left(\frac{n-1}{2}\right)} \right] \sigma \implies E\left[\frac{\sqrt{n-1}\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{2}\Gamma\left(\frac{n}{2}\right)} S \right] = \sigma;$$

i.e., we just multiplied by the reciprocal. This shows

$$\hat{\sigma} = \frac{\sqrt{n-1}\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{2}\Gamma\left(\frac{n}{2}\right)} S$$

is an unbiased estimator of σ .

(c) The function $\phi_{\alpha} = \mu - z_{\alpha}\sigma$ is the α th quantile of $Y \sim \mathcal{N}(\mu, \sigma^2)$. Note that

$$\alpha = P(Y \leq \phi_{\alpha}) = P\left(Z \leq \frac{\phi_{\alpha} - \mu}{\sigma}\right) \implies \frac{\phi_{\alpha} - \mu}{\sigma} = z_{\alpha} \implies \phi_{\alpha} = \mu - z_{\alpha}\sigma.$$

We already know \bar{Y} is an unbiased estimator of μ . We have an unbiased estimator of σ from above. Because z_{α} is a constant, we have

$$E(\bar{Y} - z_{\alpha}\hat{\sigma}) = E(\bar{Y}) - z_{\alpha}E(\hat{\sigma}) = \mu - z_{\alpha}\sigma.$$

This shows

$$\bar{Y} - z_{\alpha}\hat{\sigma} = \bar{Y} - z_{\alpha} \left[\frac{\sqrt{n-1}\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{2}\Gamma\left(\frac{n}{2}\right)} S \right]$$

is an unbiased estimator of the α th quantile of $Y \sim \mathcal{N}(\mu, \sigma^2)$.

8.18. In this problem, Y_1, Y_2, \dots, Y_n is an iid sample from a $\mathcal{U}(0, \theta)$ population distribution where $\theta > 0$ is unknown. We want to find $E(Y_{(1)})$, where $Y_{(1)}$ is the minimum order statistic. Recall the pdf of $Y_{(1)}$, in general, is

$$f_{Y_{(1)}}(y) = n f_Y(y) [1 - F_Y(y)]^{n-1},$$

where $F_Y(y)$ is the cdf of Y . The population pdf of $Y \sim \mathcal{U}(0, \theta)$ is

$$f_Y(y) = \begin{cases} \frac{1}{\theta}, & 0 < y < \theta \\ 0, & \text{otherwise} \end{cases}$$

and the population cdf is

$$F_Y(y) = \begin{cases} 0, & y \leq 0 \\ \frac{y}{\theta}, & 0 < y < \theta \\ 1, & y \geq \theta. \end{cases}$$

Therefore, for $0 < y < \theta$, the pdf of $Y_{(1)}$ is

$$f_{Y_{(1)}}(y) = n f_Y(y) [1 - F_Y(y)]^{n-1} = n \left(\frac{1}{\theta} \right) \left(1 - \frac{y}{\theta} \right)^{n-1} = \frac{n}{\theta} \left(1 - \frac{y}{\theta} \right)^{n-1}.$$

Summarizing,

$$f_{Y_{(1)}}(y) = \begin{cases} \frac{n}{\theta} \left(1 - \frac{y}{\theta} \right)^{n-1}, & 0 < y < \theta \\ 0, & \text{otherwise.} \end{cases}$$

Therefore,

$$E(Y_{(1)}) = \int_{\mathbb{R}} y f_{Y_{(1)}}(y) dy = \int_0^\theta y \times \frac{n}{\theta} \left(1 - \frac{y}{\theta} \right)^{n-1} dy.$$

In the last integral, let

$$u = \frac{y}{\theta} \implies du = \frac{dy}{\theta}.$$

Note also that with this u -substitution, the limits of integration change. As $y : 0 \rightarrow \theta$, the transformed variable $u : 0 \rightarrow 1$. Therefore, we have

$$\begin{aligned} E(Y_{(1)}) &= \int_0^\theta y \times \frac{n}{\theta} \left(1 - \frac{y}{\theta} \right)^{n-1} dy = \int_0^1 \theta u \times \frac{n}{\theta} (1 - u)^{n-1} \theta du \\ &= n\theta \int_0^1 u(1 - u)^{n-1} du = n\theta \left[\frac{\Gamma(2)\Gamma(n)}{\Gamma(n+2)} \right]. \end{aligned}$$

Note that the last integrand $u(1 - u)^{n-1}$ is the beta(2, n) kernel and we are integrating over $(0, 1)$. Therefore,

$$E(Y_{(1)}) = n\theta \left[\frac{\Gamma(2)\Gamma(n)}{\Gamma(n+2)} \right] = n\theta \left[\frac{\Gamma(n)}{(n+1)n\Gamma(n)} \right] = \left(\frac{1}{n+1} \right) \theta.$$

Finally, note that

$$E(Y_{(1)}) = \left(\frac{1}{n+1} \right) \theta \implies E[(n+1)Y_{(1)}] = \theta.$$

This shows $\hat{\theta} = (n+1)Y_{(1)}$ is an unbiased estimator of θ .

8.20. In this problem, Y_1, Y_2, Y_3, Y_4 is an iid sample of size $n = 4$ from an exponential(θ) population distribution where $\theta > 0$ is unknown. In part (a), we have

$$E(X) = E(\sqrt{Y_1 Y_2}) = E(\sqrt{Y_1} \sqrt{Y_2}) = E(\sqrt{Y_1}) E(\sqrt{Y_2}).$$

The last step follows because $Y_1 \perp\!\!\!\perp Y_2$, so $\sqrt{Y_1} \perp\!\!\!\perp \sqrt{Y_2}$; i.e., functions of independent random variables are also independent. Note that $E(\sqrt{Y_1}) = E(\sqrt{Y_2})$ because Y_1 and Y_2 are identically distributed. Therefore, let's calculate $E(\sqrt{Y})$, where $Y \sim \text{exponential}(\theta)$. We have

$$E(\sqrt{Y}) = \int_{\mathbb{R}} \sqrt{y} f_Y(y) dy = \int_0^{\infty} y^{\frac{1}{2}} \times \frac{1}{\theta} e^{-y/\theta} dy = \frac{1}{\theta} \int_0^{\infty} y^{\frac{1}{2}} e^{-y/\theta} dy.$$

We recognize $y^{\frac{1}{2}} e^{-y/\theta} = y^{\frac{3}{2}-1} e^{-y/\theta}$ as the kernel of a gamma pdf with shape parameter $\alpha = 3/2$ and scale parameter $\beta = \theta$. Therefore,

$$E(\sqrt{Y}) = \frac{1}{\theta} \Gamma\left(\frac{3}{2}\right) \theta^{\frac{3}{2}} = \frac{1}{2} \Gamma\left(\frac{1}{2}\right) \theta^{\frac{1}{2}} = \frac{\sqrt{\pi\theta}}{2},$$

because $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, a result we proved in STAT 511. Therefore,

$$E(X) = E(\sqrt{Y_1}) E(\sqrt{Y_2}) = \frac{\sqrt{\pi\theta}}{2} \frac{\sqrt{\pi\theta}}{2} = \left(\frac{\pi}{4}\right) \theta.$$

Finally, note that

$$E(X) = \left(\frac{\pi}{4}\right) \theta \implies E\left(\frac{X}{\pi/4}\right) = \theta.$$

Therefore,

$$\frac{X}{\pi/4} = \frac{4X}{\pi} = \frac{4\sqrt{Y_1 Y_2}}{\pi}$$

is an unbiased estimator of θ .

(b) The argument for this part is identical. We have

$$\begin{aligned} E(W) &= E(\sqrt{Y_1 Y_2 Y_3 Y_4}) = E(\sqrt{Y_1} \sqrt{Y_2} \sqrt{Y_3} \sqrt{Y_4}) \\ &= E(\sqrt{Y_1}) E(\sqrt{Y_2}) E(\sqrt{Y_3}) E(\sqrt{Y_4}) = \left(\frac{\sqrt{\pi\theta}}{2}\right)^4 = \left(\frac{\pi^2}{16}\right) \theta^2. \end{aligned}$$

Therefore,

$$E(W) = \left(\frac{\pi^2}{16}\right) \theta^2 \implies E\left(\frac{W}{\pi^2/16}\right) = \theta^2.$$

Therefore,

$$\frac{W}{\pi^2/16} = \frac{16W}{\pi^2} = \frac{16\sqrt{Y_1 Y_2 Y_3 Y_4}}{\pi^2}$$

is an unbiased estimator of θ^2 .

8.133. In this problem, we have two independent random samples:

- $Y_{11}, Y_{12}, \dots, Y_{1n_1}$ is an iid sample from a $\mathcal{N}(\mu_1, \sigma^2)$ population distribution

- $Y_{21}, Y_{22}, \dots, Y_{2n_2}$ is an iid sample from a $\mathcal{N}(\mu_2, \sigma^2)$ population distribution,

where all population parameters are unknown. **Important:** Note that the population variance σ^2 is assumed to be the same in each population.

Define the sample means

$$\bar{Y}_{1+} = \frac{1}{n_1} \sum_{j=1}^{n_1} Y_{1j} \quad \text{and} \quad \bar{Y}_{2+} = \frac{1}{n_2} \sum_{j=1}^{n_2} Y_{2j}$$

and the sample variances

$$S_1^2 = \frac{1}{n_1 - 1} \sum_{j=1}^{n_1} (Y_{1j} - \bar{Y}_{1+})^2 \quad \text{and} \quad S_2^2 = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} (Y_{2j} - \bar{Y}_{2+})^2.$$

The pooled sample variance estimator is

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

In part (a), we want to show that S_p^2 is an unbiased estimator of the common population variance σ^2 ; i.e., we want to show

$$E(S_p^2) = \sigma^2.$$

Showing this is easy. Recall that the sample variance is always an unbiased estimator of the population variance (in any population distribution, provided that the population variance is finite, of course). Therefore, we have

$$\begin{aligned} E(S_1^2) &= \sigma^2 \\ E(S_2^2) &= \sigma^2. \end{aligned}$$

Therefore,

$$\begin{aligned} E(S_p^2) &= E \left[\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \right] = \frac{1}{n_1 + n_2 - 2} E[(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2] \\ &= \frac{1}{n_1 + n_2 - 2} [(n_1 - 1)E(S_1^2) + (n_2 - 1)E(S_2^2)] \\ &= \frac{1}{n_1 + n_2 - 2} [(n_1 - 1)\sigma^2 + (n_2 - 1)\sigma^2] \\ &= \left(\frac{n_1 + n_2 - 2}{n_1 + n_2 - 2} \right) \sigma^2 = \sigma^2. \end{aligned}$$

This shows S_p^2 is an unbiased estimator of σ^2 .

(b) To calculate $V(S_p^2)$, recall that

$$\frac{(n_1 - 1)S_1^2}{\sigma^2} \sim \chi^2(n_1 - 1) \quad \text{and} \quad \frac{(n_2 - 1)S_2^2}{\sigma^2} \sim \chi^2(n_2 - 1).$$

Therefore,

$$V \left(\frac{(n_1 - 1)S_1^2}{\sigma^2} \right) = 2(n_1 - 1) \implies \frac{(n_1 - 1)^2}{\sigma^4} V(S_1^2) = 2(n_1 - 1) \implies V(S_1^2) = \frac{2\sigma^4}{n_1 - 1},$$

a result we saw in Chapter 7 (Result 5). Similarly,

$$V(S_2^2) = \frac{2\sigma^4}{n_2 - 1}.$$

Therefore,

$$\begin{aligned} V(S_p^2) &= V\left[\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}\right] = \left(\frac{1}{n_1 + n_2 - 2}\right)^2 V[(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2] \\ &= \left(\frac{1}{n_1 + n_2 - 2}\right)^2 [(n_1 - 1)^2 V(S_1^2) + (n_2 - 1)^2 V(S_2^2)]. \end{aligned}$$

The last equality is true because $\text{Cov}(S_1^2, S_2^2) = 0$. This is true because the samples are independent (i.e., S_1^2 is a statistic from sample 1, and S_2^2 is a statistic from sample 2). Therefore,

$$\begin{aligned} V(S_p^2) &= \left(\frac{1}{n_1 + n_2 - 2}\right)^2 [(n_1 - 1)^2 V(S_1^2) + (n_2 - 1)^2 V(S_2^2)] \\ &= \left(\frac{1}{n_1 + n_2 - 2}\right)^2 \left[(n_1 - 1)^2 \left(\frac{2\sigma^4}{n_1 - 1}\right) + (n_2 - 1)^2 \left(\frac{2\sigma^4}{n_2 - 1}\right)\right] \\ &= 2\sigma^4 \left(\frac{1}{n_1 + n_2 - 2}\right)^2 [(n_1 - 1) + (n_2 - 1)] \\ &= 2\sigma^4 \left(\frac{1}{n_1 + n_2 - 2}\right)^2 (n_1 + n_2 - 2) = \frac{2\sigma^4}{n_1 + n_2 - 2}. \end{aligned}$$