9.56. In this problem, $Y_1, Y_2, ..., Y_n$ are iid $\mathcal{N}(\mu_0, \sigma^2)$, where $\mu = \mu_0$ is known and $\sigma^2 > 0$ is unknown. We want find the MVUE for σ^2. We start by finding a sufficient statistic. The likelihood function is

$$\begin{aligned} L(\sigma^2|\mathbf{y}) &= f_Y(y_1|\sigma^2) \times f_Y(y_2|\sigma^2) \times \cdots \times f_Y(y_n|\sigma^2) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_1-\mu_0)^2} \times \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_2-\mu_0)^2} \times \cdots \times \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_n-\mu_0)^2} \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i-\mu_0)^2}. \end{aligned}$$

Note we can write

$$L(\sigma^2|\mathbf{y}) = \underbrace{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i-\mu_0)^2}}_{g(t,\sigma^2)} \times \underbrace{1}_{h(y_1,y_2,...,y_n)},$$

where $t = \sum_{i=1}^n (y_i - \mu_0)^2$. By the Factorization Theorem, it follows that

$$T = \sum_{i=1}^n (Y_i - \mu_0)^2$$

is a sufficient statistic for σ^2 (when $\mu = \mu_0$ is known). From the Rao-Blackwell Theorem, we know the MVUE of σ^2 must be a function of T. Therefore, let's calculate the expectation of T. We have

$$E(T) = E\left[\sum_{i=1}^n (Y_i - \mu_0)^2\right] = \sum_{i=1}^n E[(Y_i - \mu_0)^2] = \sum_{i=1}^n \sigma^2 = n\sigma^2.$$

Recall $\sigma^2 = V(Y) = E[(Y - \mu_0)^2]$; i.e., this is the definition of the variance of a random variable Y. Therefore,

$$E(T) = n\sigma^2 \implies E\left(\frac{T}{n}\right) = \sigma^2.$$

Therefore,

$$\frac{T}{n} = \frac{1}{n}\sum_{i=1}^n (Y_i - \mu_0)^2$$

is the MVUE of σ^2. It is a function of a sufficient statistic $T = \sum_{i=1}^n (Y_i - \mu_0)^2$ and it is unbiased.

9.59. In this problem, $Y_1, Y_2, ..., Y_n$ is an iid sample from a Poisson(λ) population distribution where $\lambda > 0$ is unknown. We want to find the MVUE of

$$E(C) = E(3Y^2) = 3E(Y^2) = 3\{V(Y) + [E(Y)]^2\} = 3(\lambda + \lambda^2).$$

We start by finding a sufficient statistic. The likelihood function is given by

$$\begin{aligned} L(\lambda|\mathbf{y}) &= p_Y(y_1|\lambda) \times p_Y(y_2|\lambda) \times \cdots \times p_Y(y_n|\lambda) \\ &= \frac{\lambda^{y_1} e^{-\lambda}}{y_1!} \times \frac{\lambda^{y_2} e^{-\lambda}}{y_2!} \times \cdots \times \frac{\lambda^{y_n} e^{-\lambda}}{y_n!} = \frac{\lambda^{\sum_{i=1}^n y_i} e^{-n\lambda}}{y_1! y_2! \cdots y_n!} = \frac{\lambda^{\sum_{i=1}^n y_i} e^{-n\lambda}}{\prod_{i=1}^n y_i!}. \end{aligned}$$

Note that we can write the likelihood function as

$$L(\lambda|\mathbf{y}) = \frac{\lambda^{\sum_{i=1}^n y_i} e^{-n\lambda}}{\prod_{i=1}^n y_i!} = \underbrace{\lambda^{\sum_{i=1}^n y_i} e^{-n\lambda}}_{g(t,\lambda)} \times \underbrace{\frac{1}{\prod_{i=1}^n y_i!}}_{h(y_1,y_2,...,y_n)},$$

where $t = \sum_{i=1}^n y_i$. By the Factorization Theorem, it follows that $T = \sum_{i=1}^n Y_i$ is a sufficient statistic for λ. From the Rao-Blackwell Theorem, we know the MVUE of $E(C) = 3(\lambda + \lambda^2)$ must be a function of T. We know

$$E(T) = E\left(\sum_{i=1}^n Y_i\right) = \sum_{i=1}^n E(Y_i) = \sum_{i=1}^n \lambda = n\lambda.$$

Therefore,

$$E(\overline{Y}) = E\left(\frac{T}{n}\right) = \lambda.$$

We also need to estimate λ^2 unbiasedly. Let's try $\overline{Y}^2$. Note that

$$E(\overline{Y}^2) = V(\overline{Y}) + [E(\overline{Y})]^2 = \frac{\lambda}{n} + \lambda^2.$$

Therefore, $\overline{Y}^2$ is a biased estimator of λ^2. However, note that

$$E(\overline{Y}^2) = \frac{\lambda}{n} + \lambda^2 \implies E\left(\overline{Y}^2 - \frac{\overline{Y}}{n}\right) = \frac{\lambda}{n} + \lambda^2 - \frac{\lambda}{n} = \lambda^2.$$

Therefore, $\overline{Y}^2 - \frac{\overline{Y}}{n}$ is an unbiased estimator of λ^2. Therefore,

$$E\left[3\left(\overline{Y} + \overline{Y}^2 - \frac{\overline{Y}}{n}\right)\right] = 3\left[E(\overline{Y}) + E\left(\overline{Y}^2 - \frac{\overline{Y}}{n}\right)\right] = 3(\lambda + \lambda^2).$$

This shows

$$3\left(\overline{Y} + \overline{Y}^2 - \frac{\overline{Y}}{n}\right)$$

is the MVUE of $E(C) = 3(\lambda + \lambda^2)$. It is a function of a sufficient statistic $T = \sum_{i=1}^n Y_i$ and it is unbiased.

9.63. In this problem, $Y_1, Y_2, ..., Y_n$ is an iid sample from a population with pdf

$$f_Y(y) = \begin{cases} \frac{3y^2}{\theta^3}, & 0 \le y \le \theta \\ 0, & \text{otherwise,} \end{cases}$$

where the population parameter $\theta > 0$ is unknown. We want to find the MVUE of θ. We start by finding a sufficient statistic. The likelihood function is

$$\begin{aligned} L(\theta|\mathbf{y}) &= f_Y(y_1|\theta) \times f_Y(y_2|\theta) \times \cdots \times f_Y(y_n|\theta) \\ &= \frac{3y_1^2}{\theta^3} I(0 \le y_1 \le \theta) \times \frac{3y_2^2}{\theta^3} I(0 \le y_2 \le \theta) \times \cdots \times \frac{3y_n^2}{\theta^3} I(0 \le y_n \le \theta) \\ &= \left(\frac{3}{\theta^3}\right)^n \left(\prod_{i=1}^n y_i^2\right) \prod_{i=1}^n I(0 \le y_i \le \theta) \\ &= \left(\frac{3}{\theta^3}\right)^n \left(\prod_{i=1}^n y_i^2\right) I(0 \le y_{(n)} \le \theta). \end{aligned}$$

Note we can write

$$L(\theta|\mathbf{y}) = \underbrace{\left(\frac{3}{\theta^3}\right)^n I(0 \leq y_{(n)} \leq \theta)}_{g(t,\theta)} \times \underbrace{\prod_{i=1}^n y_i^2}_{h(y_1,y_2,\ldots,y_n)},$$

where $t = y_{(n)}$. By the Factorization Theorem, it follows that $T = Y_{(n)}$ is a sufficient statistic for θ. From the Rao-Blackwell Theorem, we know the MVUE of θ must be a function of T. Therefore, let's calculate the expectation of T.

In part (a), the authors ask you to derive the pdf of $T = Y_{(n)}$; you need to know this so that you can calculate $E(T) = E(Y_{(n)})$. Recall that in general,

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

where $F_Y(y)$ is the population cdf. We calculate

$$F_Y(y) = \begin{cases} 0, & y < 0 \\ \left(\frac{y}{\theta}\right)^3, & 0 \leq y \leq \theta \\ 1, & y > \theta. \end{cases}$$

Therefore, for $0 \leq y \leq \theta$, the pdf of the maximum order statistic $Y_{(n)}$ is

$$f_{Y_{(n)}}(y) = n\left(\frac{3y^2}{\theta^3}\right)\left[\left(\frac{y}{\theta}\right)^3\right]^{n-1} = \frac{3ny^{3n-1}}{\theta^{3n}}.$$

Summarizing,

$$f_{Y_{(n)}}(y) = \begin{cases} \frac{3ny^{3n-1}}{\theta^{3n}}, & 0 \leq y \leq \theta \\ 0, & \text{otherwise.} \end{cases}$$

(b) To find the MVUE, let's first find $E(Y_{(n)})$. We have

$$\begin{aligned} E(Y_{(n)}) = \int_{\mathbb{R}} yf_{Y_{(n)}}(y)dy = \int_0^\theta \frac{3ny^{3n}}{\theta^{3n}}dy &= \frac{3n}{\theta^{3n}}\left(\frac{1}{3n+1}\right)y^{3n+1}\Big|_0^\theta \\ &= \frac{3n}{3n+1}\frac{\theta^{3n+1}}{\theta^{3n}} = \left(\frac{3n}{3n+1}\right)\theta. \end{aligned}$$

Therefore, $Y_{(n)}$ is a biased estimator of θ; however, note that

$$E(Y_{(n)}) = \left(\frac{3n}{3n+1}\right)\theta \implies E\left[\left(\frac{3n+1}{3n}\right)Y_{(n)}\right] = \theta.$$

Therefore,

$$\left(\frac{3n+1}{3n}\right)Y_{(n)}$$

is the MVUE of θ. It is a function of a sufficient statistic $T = Y_{(n)}$ and it is unbiased.

9.64. In this problem, $Y_1, Y_2, ..., Y_n$ is an iid sample from a $\mathcal{N}(\mu, 1)$ population distribution where $-\infty < \mu < \infty$ is unknown and the population variance is 1. In part (a), we want to find the MVUE for μ^2. We start by finding a sufficient statistic. The likelihood function is

$$\begin{aligned} L(\mu|\mathbf{y}) &= f_Y(y_1|\mu) \times f_Y(y_2|\mu) \times \cdots \times f_Y(y_n|\mu) \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_1-\mu)^2} \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_2-\mu)^2} \times \cdots \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_n-\mu)^2} \\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^n (y_i-\mu)^2}. \end{aligned}$$

Now, write

$$\sum_{i=1}^n (y_i - \mu)^2 = \sum_{i=1}^n (y_i^2 - 2\mu y_i + \mu^2) = \sum_{i=1}^n y_i^2 - 2\mu \sum_{i=1}^n y_i + n\mu^2.$$

Therefore,

$$\begin{aligned} L(\mu|\mathbf{y}) &= \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\left(\sum_{i=1}^n y_i^2 - 2\mu\sum_{i=1}^n y_i + n\mu^2\right)} \\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^n y_i^2}\ e^{\mu\sum_{i=1}^n y_i}\ e^{-\frac{n\mu^2}{2}} \\ &= \underbrace{e^{\mu\sum_{i=1}^n y_i} e^{-\frac{n\mu^2}{2}}}_{g(t,\mu)} \times \underbrace{\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^n y_i^2}}_{h(y_1,y_2,...,y_n)}, \end{aligned}$$

where $t = \sum_{i=1}^n y_i$. By the Factorization Theorem, it follows that $T = \sum_{i=1}^n Y_i$ is a sufficient statistic for μ. From the Rao-Blackwell Theorem, we know the MVUE of μ^2 must be a function of T. We know

$$E(\overline{Y}) = E\left(\frac{T}{n}\right) = \mu,$$

so (to estimate μ^2) let's try working with $\overline{Y}^2$. Note that

$$E(\overline{Y}^2) = V(\overline{Y}) + [E(\overline{Y})]^2 = \frac{1}{n} + \mu^2.$$

Therefore, $\overline{Y}^2$ is a biased estimator of μ^2. However, note that

$$E(\overline{Y}^2) = \frac{1}{n} + \mu^2 \implies E\left(\overline{Y}^2 - \frac{1}{n}\right) = \mu^2.$$

Therefore,

$$\overline{Y}^2 - \frac{1}{n}$$

is the MVUE of μ^2. It is a function of a sufficient statistic $T = \sum_{i=1}^n Y_i$ and it is unbiased.

(b) In this part, we want to calculate

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

The only way I know how to find $V(\overline{Y}^2)$ is to write (via the variance computing formula)

$$V(\overline{Y}^2) = E(\overline{Y}^4) - [E(\overline{Y}^2)]^2.$$

Now, $E(\overline{Y}^4)$ is the fourth moment of

$$\overline{Y} \sim \mathcal{N}\left(\mu, \frac{1}{n}\right).$$

The mgf of $\overline{Y}$ is

$$m_{\overline{Y}}(t) = \exp\left[\mu t + \frac{\left(\frac{1}{n}\right)t^2}{2}\right] = \exp\left(\mu t + \frac{t^2}{2n}\right).$$

Therefore, we can get $E(\overline{Y}^4)$ by calculating

$$E(\overline{Y}^4) = \frac{d^4}{dt^4} m_{\overline{Y}}(t)\Big|_{t=0}.$$

Here are the derivatives of $m_{\overline{Y}}(t)$:

$$\begin{aligned}
\frac{d}{dt} m_{\overline{Y}}(t) &= \left(\mu + \frac{t}{n}\right)\exp\left(\mu t + \frac{t^2}{2n}\right) \\
\frac{d^2}{dt^2} m_{\overline{Y}}(t) &= \frac{1}{n}\exp\left(\mu t + \frac{t^2}{2n}\right) + \left(\mu + \frac{t}{n}\right)^2 \exp\left(\mu t + \frac{t^2}{2n}\right) \\
&= \left[\frac{1}{n} + \left(\mu + \frac{t}{n}\right)^2\right]\exp\left(\mu t + \frac{t^2}{2n}\right) \\
\frac{d^3}{dt^3} m_{\overline{Y}}(t) &= 2\left(\mu + \frac{t}{n}\right)\frac{1}{n}\exp\left(\mu t + \frac{t^2}{2n}\right) + \left[\frac{1}{n} + \left(\mu + \frac{t}{n}\right)^2\right]\left(\mu + \frac{t}{n}\right)\exp\left(\mu t + \frac{t^2}{2n}\right) \\
&= \left(\mu + \frac{t}{n}\right)^3 \exp\left(\mu t + \frac{t^2}{2n}\right) + 3\left(\mu + \frac{t}{n}\right)\frac{1}{n}\exp\left(\mu t + \frac{t^2}{2n}\right) \\
&= \left[\left(\mu + \frac{t}{n}\right)^3 + 3\left(\mu + \frac{t}{n}\right)\frac{1}{n}\right]\exp\left(\mu t + \frac{t^2}{2n}\right)
\end{aligned}$$

and

$$\begin{aligned}
\frac{d^4}{dt^4} m_{\overline{Y}}(t) = &\left[3\left(\mu + \frac{t}{n}\right)^2 \frac{1}{n} + \frac{3}{n^2}\right]\exp\left(\mu t + \frac{t^2}{2n}\right) \\
&+ \left[\left(\mu + \frac{t}{n}\right)^3 + 3\left(\mu + \frac{t}{n}\right)\frac{1}{n}\right]\left(\mu + \frac{t}{n}\right)\exp\left(\mu t + \frac{t^2}{2n}\right).
\end{aligned}$$

Therefore,

$$E(\overline{Y}^4) = \frac{d^4}{dt^4} m_{\overline{Y}}(t)\Big|_{t=0} = \left(\frac{3\mu^2}{n} + \frac{3}{n^2}\right) + \left(\mu^3 + \frac{3\mu}{n}\right)\mu \;=\; \mu^4 + \frac{6\mu^2}{n} + \frac{3}{n^2}.$$

Also,

$$E(\overline{Y}^2) = \frac{d^2}{dt^2} m_{\overline{Y}}(t)\Big|_{t=0} = \frac{1}{n} + \mu^2.$$

Therefore, the variance of the MVUE of μ^2 is

$$\begin{aligned} V\left(\overline{Y}^2 - \frac{1}{n}\right) = V(\overline{Y}^2) &= E(\overline{Y}^4) - [E(\overline{Y}^2)]^2 \\ &= \mu^4 + \frac{6\mu^2}{n} + \frac{3}{n^2} - \left(\frac{1}{n} + \mu^2\right)^2 \\ &= \mu^4 + \frac{6\mu^2}{n} + \frac{3}{n^2} - \left(\frac{1}{n^2} + \frac{2\mu^2}{n} + \mu^4\right) \\ &= \mu^4 + \frac{6\mu^2}{n} + \frac{3}{n^2} - \frac{1}{n^2} - \frac{2\mu^2}{n} - \mu^4 \ = \ \frac{4\mu^2}{n} + \frac{2}{n^2}. \end{aligned}$$

9.72. In this problem, $Y_1, Y_2, ..., Y_n$ are iid from a $\mathcal{N}(\mu, \sigma^2)$ population distribution, where both parameters are unknown; i.e., there are $d = 2$ parameters to estimate. Therefore, to find the MOM estimators of μ and σ^2, we need two equations. The first two population moments are

$$\begin{aligned} E(Y) &= \mu \\ E(Y^2) = V(Y) + [E(Y)]^2 &= \sigma^2 + \mu^2. \end{aligned}$$

The first two sample moments are

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} Y_i &= \overline{Y} \\ \frac{1}{n}\sum_{i=1}^{n} Y_i^2 &= m_2'. \end{aligned}$$

Therefore, the MOM estimators of μ and σ^2 are found by solving

$$\begin{aligned} \mu &\stackrel{\text{set}}{=} \overline{Y} \\ \sigma^2 + \mu^2 &\stackrel{\text{set}}{=} m_2'. \end{aligned}$$

The solution to the first equation is obvious; i.e.,

$$\widehat{\mu} = \overline{Y}.$$

Substituting $\widehat{\mu} = \overline{Y}$ into the second equation, we get

$$\begin{aligned} \widehat{\sigma}^2 = m_2' - \overline{Y}^2 = \frac{1}{n}\sum_{i=1}^{n} Y_i^2 - \overline{Y}^2 &= \frac{1}{n}\sum_{i=1}^{n} Y_i^2 - \frac{n\overline{Y}^2}{n} \\ &= \frac{1}{n}\left(\sum_{i=1}^{n} Y_i^2 - n\overline{Y}^2\right) = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2. \end{aligned}$$

Therefore, the MOM estimator of $\boldsymbol{\theta} = (\mu, \sigma^2)$ is

$$\widehat{\boldsymbol{\theta}} = \begin{pmatrix} \overline{Y} \\ \frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2 \end{pmatrix}.$$

Note that the MOM estimator of $\boldsymbol{\theta}$ and the MLE of $\boldsymbol{\theta}$ are the same under $\mathcal{N}(\mu, \sigma^2)$ assumption; see Example 9.20 (notes).

9.74. In this problem, $Y_1, Y_2, ..., Y_n$ is an iid sample from a population with pdf

$$f_Y(y|\theta) = f_Y(y) = \begin{cases} \dfrac{2}{\theta^2}(\theta - y), & 0 \le y \le \theta \\ 0, & \text{otherwise}, \end{cases}$$

where $\theta > 0$ is unknown. In part (a), we want to find the MOM estimator of θ. There is only 1 parameter in this population pdf, so to find the MOM estimator we only need one equation. The first population moment is

$$\begin{aligned} E(Y) = \int_{\mathbb{R}} y f_Y(y) dy = \int_0^\theta \frac{2y}{\theta^2}(\theta - y) dy &= \int_0^\theta \left(\frac{2y}{\theta} - \frac{2y^2}{\theta^2} \right) dy \\ &= \left(\frac{y^2}{\theta} - \frac{2y^3}{3\theta^2} \right) \Bigg|_0^\theta = \theta - \frac{2\theta}{3} = \frac{\theta}{3}. \end{aligned}$$

The first sample moment is

$$\frac{1}{n} \sum_{i=1}^n Y_i = \overline{Y}.$$

Therefore, the MOM estimator of θ is found by solving

$$\frac{\theta}{3} \stackrel{\text{set}}{=} \overline{Y} \implies \widehat{\theta} = 3\overline{Y}.$$

(b) Note that the support of Y depends on θ; i.e., $0 \le y \le \theta$, so the sufficient statistic will be an order statistic or a function of the order statistics $Y_{(1)}, Y_{(2)}, ..., Y_{(n)}$. The MOM estimator $3\overline{Y}$ is not sufficient because it is not a 1:1 function of the order statistics; i.e., if you know $Y_{(1)}, Y_{(2)}, ..., Y_{(n)}$, you can calculate $3\overline{Y}$, but not the other way–if you know $3\overline{Y}$, you can not determine the order statistics.

Interesting: In this population-level model, the order statistics $Y_{(1)}, Y_{(2)}, ..., Y_{(n)}$ are sufficient and you cannot reduce any further. If you do, then you will start to lose information about θ.

9.75. In this problem, $Y_1, Y_2, ..., Y_n$ is an iid sample from a beta population distribution where $\alpha = \theta$ and $\beta = \theta$; i.e., the population parameters α and β are both equal. In other words, the population distribution is $Y \sim \text{beta}(\theta, \theta)$. We want to find the MOM estimator of θ. There is only 1 unknown parameter in this model, so we only need 1 equation.

However, we quickly encounter a problem. Recall that the first population moment is

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

Therefore, if we were to set the first population moment equal to the first sample moment $\overline{Y}$, we would get

$$\frac{1}{2} \stackrel{\text{set}}{=} \overline{Y},$$

which is not helpful (because this equation does not involve θ).

Q: What do we do in this situation?
A: Move to second moments.

The second moment of Y is

$$\begin{aligned} E(Y^2) = V(Y) + [E(Y)]^2 &= \frac{\theta(\theta)}{(\theta+\theta+1)(\theta+\theta)^2} + \left(\frac{1}{2}\right)^2 \\ &= \frac{\theta^2}{(2\theta+1)4\theta^2} + \frac{1}{4} \\ &= \frac{1}{4(2\theta+1)} + \frac{1}{4}. \end{aligned}$$

The second sample moment is

$$\frac{1}{n}\sum_{i=1}^{n} Y_i^2 = m_2'.$$

Therefore, to find the MOM estimator of θ, we set

$$\frac{1}{4(2\theta+1)} + \frac{1}{4} \overset{\text{set}}{=} m_2',$$

and solve for θ. We have

$$\begin{aligned} \frac{1}{4(2\theta+1)} = m_2' - \frac{1}{4} &\implies 4(2\theta+1) = \frac{1}{m_2' - \frac{1}{4}} \\ &\implies 2\theta + 1 = \frac{1}{4m_2' - 1} \\ &\implies 2\theta = \frac{1}{4m_2' - 1} - 1 = \frac{2 - 4m_2'}{4m_2' - 1} = \frac{2(1 - 2m_2')}{4m_2' - 1} \implies \widehat{\theta} = \frac{1 - 2m_2'}{4m_2' - 1}. \end{aligned}$$

The MOM estimator of θ is

$$\widehat{\theta} = \frac{1 - 2m_2'}{4m_2' - 1} = \frac{1 - \frac{2}{n}\sum_{i=1}^{n} Y_i^2}{\frac{4}{n}\sum_{i=1}^{n} Y_i^2 - 1}.$$

9.77. In this problem, $Y_1, Y_2, ..., Y_n$ is an iid sample from a $\mathcal{U}(0, 3\theta)$ population distribution, where $\theta > 0$ is unknown. We want to find the MOM estimator of θ. There is only 1 parameter in this population pdf, so to find the MOM estimator we only need one equation. The first population moment is

$$E(Y) = \frac{3\theta}{2};$$

i.e., the midpoint of 0 and 3θ. The first sample moment is

$$\frac{1}{n}\sum_{i=1}^{n} Y_i = \overline{Y}.$$

Therefore, the MOM estimator of θ is found by solving

$$\frac{3\theta}{2} \overset{\text{set}}{=} \overline{Y} \implies \widehat{\theta} = \frac{2\overline{Y}}{3}.$$

9.88. In this problem, $Y_1, Y_2, ..., Y_n$ is an iid sample from a population with pdf

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

where $\theta > -1$ is unknown. Note that this is of pdf of $Y \sim \text{beta}(\theta+1, 1)$. The likelihood function is given by

$$\begin{aligned} L(\theta|\mathbf{y}) &= f_Y(y_1|\theta) \times f_Y(y_2|\theta) \times \cdots \times f_Y(y_n|\theta) \\ &= (\theta+1)y_1^\theta \times (\theta+1)y_2^\theta \times \cdots \times (\theta+1)y_n^\theta \;=\; (\theta+1)^n \left(\prod_{i=1}^n y_i\right)^\theta. \end{aligned}$$

The log-likelihood function is given by

$$\begin{aligned} \ln L(\theta|\mathbf{y}) &= \ln\left[(\theta+1)^n \left(\prod_{i=1}^n y_i\right)^\theta\right] \\ &= \ln[(\theta+1)^n] + \ln\left[\left(\prod_{i=1}^n y_i\right)^\theta\right] \\ &= n\ln(\theta+1) + \theta\ln\left(\prod_{i=1}^n y_i\right) \;=\; n\ln(\theta+1) + \theta\sum_{i=1}^n \ln y_i. \end{aligned}$$

The derivative of the log-likelihood function is given by

$$\begin{aligned} \frac{\partial}{\partial\theta}\ln L(\theta|\mathbf{y}) &= \frac{n}{\theta+1} + \sum_{i=1}^n \ln y_i \overset{\text{set}}{=} 0 \\ &\Longrightarrow \sum_{i=1}^n \ln y_i = -\frac{n}{\theta+1} \Longrightarrow \theta+1 = -\frac{n}{\sum_{i=1}^n \ln y_i} \Longrightarrow \widehat{\theta} = -\frac{n}{\sum_{i=1}^n \ln y_i} - 1. \end{aligned}$$

We now show this first-order critical point $\widehat{\theta}$ maximizes $\ln L(\theta|\mathbf{y})$. The second derivative of the log-likelihood function is given by

$$\frac{\partial^2}{\partial\theta^2}\ln L(\theta|\mathbf{y}) = -\frac{n}{(\theta+1)^2}.$$

Note that

$$\frac{\partial^2}{\partial\theta^2}\ln L(\theta|\mathbf{y})\Big|_{\theta=\widehat{\theta}} = -\frac{n}{(\widehat{\theta}+1)^2} < 0.$$

Therefore, $\widehat{\theta}$ maximizes $\ln L(\theta|\mathbf{y})$. The MLE of θ is

$$\widehat{\theta} = -\frac{n}{\sum_{i=1}^n \ln Y_i} - 1.$$

What is the MOM estimator of θ in this example? The first population moment is

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

recall the mean for a beta random variable. If you did not recognize $Y \sim \text{beta}(\theta+1, 1)$, then just calculate $E(Y)$ directly by using the pdf; i.e.,

$$
\begin{aligned}
E(Y) = \int_{\mathbb{R}} y f_Y(y) dy &= \int_0^1 y \times (\theta+1) y^{\theta} dy \\
&= \int_0^1 (\theta+1) y^{\theta+1} dy = (\theta+1) \left(\frac{1}{\theta+2}\, y^{\theta+2} \right) \Big|_0^1 = \frac{\theta+1}{\theta+2}.
\end{aligned}
$$

The first sample moment is

$$\frac{1}{n} \sum_{i=1}^{n} Y_i = \overline{Y}.$$

Therefore, the MOM estimator of θ is found by solving

$$
\begin{aligned}
\frac{\theta+1}{\theta+2} \overset{\text{set}}{=} \overline{Y} &\implies \theta + 1 = \overline{Y}(\theta+2) \\
&\implies \theta + 1 = \theta\overline{Y} + 2\overline{Y} \\
&\implies \theta - \theta\overline{Y} = 2\overline{Y} - 1 \implies \theta(1-\overline{Y}) = 2\overline{Y} - 1 \implies \widehat{\theta} = \frac{2\overline{Y}-1}{1-\overline{Y}}.
\end{aligned}
$$

How do these estimators compare? One observation is that the MLE depends on the sufficient statistic, and the MOM does not (a common occurrence). Note that we can write the likelihood function

$$L(\theta|\mathbf{y}) = (\theta+1)^n \left(\prod_{i=1}^{n} y_i \right)^{\theta} = \underbrace{(\theta+1)^n \left(\prod_{i=1}^{n} y_i \right)^{\theta}}_{g(t,\theta)} \times \underbrace{1}_{h(y_1, y_2, \ldots, y_n)},$$

where $t = \prod_{i=1}^{n} y_i$. By the Factorization Theorem, it follows that $T = \prod_{i=1}^{n} Y_i$ is a sufficient statistic for θ. Note that the MLE

$$\widehat{\theta} = -\frac{n}{\sum_{i=1}^{n} \ln Y_i} - 1 = -\frac{n}{\ln \prod_{i=1}^{n} Y_i} - 1 = -\frac{n}{\ln T} - 1,$$

which is a function of T. The MOM estimator is not a function of T; it is a function of the sample mean $\overline{Y}$ (which is not sufficient).

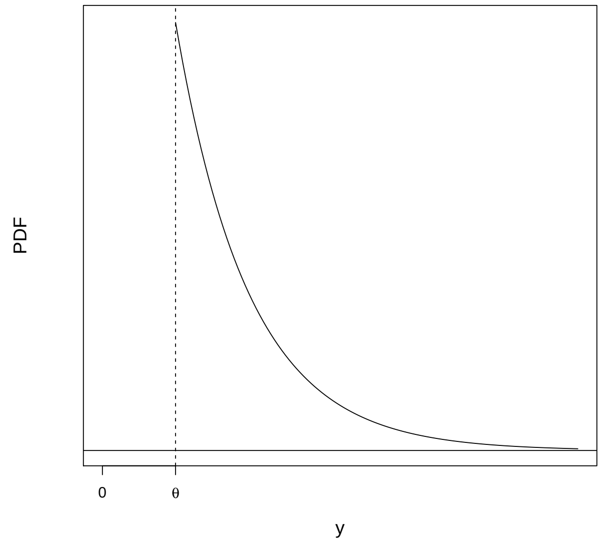
9.104. In this problem, $Y_1, Y_2, \ldots, Y_n$ is an iid sample from

$$f_Y(y) = \begin{cases} e^{-(y-\theta)}, & y \geq \theta \\ 0, & \text{otherwise}, \end{cases}$$

where the population parameter $\theta > 0$ is unknown. Note that this is a shifted exponential distribution; specifically, an exponential(1) pdf shifted to the right by θ units (since $\theta > 0$). This pdf is shown at the top of the next page.

(a) We want to find the MOM estimator. The first population moment is

$$E(Y) = \int_{\mathbb{R}} y f_Y(y) dy = \int_{\theta}^{\infty} y e^{-(y-\theta)} dy.$$



In this integral, let

$$u = y - \theta \quad \Longrightarrow \quad du = dy.$$

The limits change with this transformation; note that $y : \theta \to \infty \Longrightarrow u : 0 \to \infty$. Therefore,

$$E(Y) = \int_\theta^\infty y e^{-(y-\theta)} dy = \int_0^\infty (u+\theta) e^{-u} du = E(U+\theta),$$

where $U \sim$ exponential(1). Note that e^{-u} is the exponential(1) pdf and we are integrating over $(0, \infty)$. Therefore, $E(Y) = E(U+\theta) = E(U) + \theta = 1 + \theta$. The first sample moment is

$$\frac{1}{n}\sum_{i=1}^n Y_i = \overline{Y}.$$

Therefore, the MOM estimator of θ is found by solving

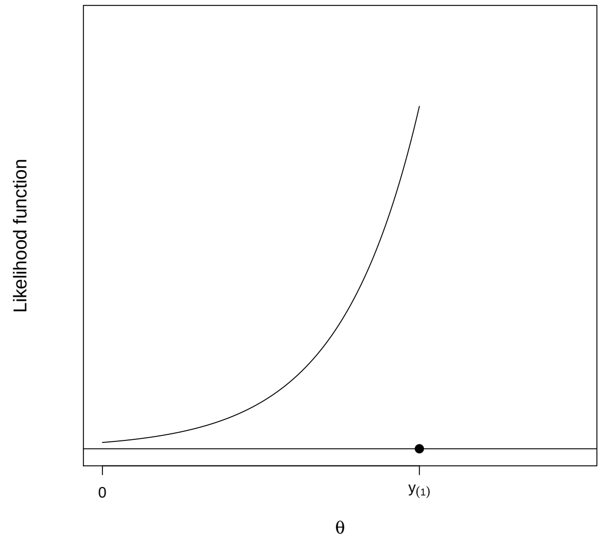
$$1 + \theta \stackrel{\text{set}}{=} \overline{Y} \quad \Longrightarrow \quad \widehat{\theta} = \overline{Y} - 1.$$

(b) The likelihood function is given by

$$\begin{aligned} L(\theta|\mathbf{y}) = \prod_{i=1}^n f_Y(y_i|\theta) &= f_Y(y_1|\theta) \times f_Y(y_2|\theta) \times \cdots \times f_Y(y_n|\theta) \\ &= e^{-(y_1-\theta)} I(y_1 \geq \theta) \times e^{-(y_2-\theta)} I(y_2 \geq \theta) \times \cdots \times e^{-(y_n-\theta)} I(y_n \geq \theta) \\ &= e^{-\sum_{i=1}^n (y_i-\theta)} \prod_{i=1}^n I(y_i \geq \theta). \end{aligned}$$

Note that

$$\prod_{i=1}^n I(y_i \geq \theta) = 1 \quad \Longleftrightarrow \quad I(y_{(1)} \geq \theta) = 1.$$



Therefore, we can write the likelihood function as

$$L(\theta|\mathbf{y}) = e^{-\sum_{i=1}^{n}(y_i-\theta)} I(y_{(1)} \geq \theta).$$

The likelihood function $L(\theta|\mathbf{y})$ is shown at the top of this page. Note that $L(\theta|\mathbf{y})$ is not differentiable for all θ; therefore, we cannot use a calculus argument. However, note that

- For $\theta \leq y_{(1)}$, $L(\theta|\mathbf{y}) = e^{-\sum_{i=1}^{n}(y_i-\theta)} = e^{n\theta - \sum_{i=1}^{n} y_i}$, which is an increasing function of θ (see above).
- For $\theta > y_{(1)}$, $L(\theta|\mathbf{y}) = 0$.

Clearly, the MLE of θ is $\widehat{\theta} = Y_{(1)}$.

(c) In this part, we want to compare

$$\begin{aligned} \widehat{\theta}_1 &= \overline{Y} - 1 \quad \text{(MOM)} \\ \widehat{\theta}_2 &= Y_{(1)} \quad\quad \text{(MLE)}. \end{aligned}$$

The MOM estimator is unbiased so no "adjustment" is necessary. Note that

$$E(\widehat{\theta}_1) = E(\overline{Y} - 1) = E(\overline{Y}) - 1 = (1+\theta) - 1 = \theta.$$

The MLE is biased. Let's find the pdf of $\widehat{\theta}_2 = Y_{(1)}$ so we can calculate its expectation. Recall that in general,

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

The population cdf is

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

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

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

Summarizing,

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

The mean of $Y_{(1)}$ is

$$E(Y_{(1)}) = \int_{\mathbb{R}} y f_{Y_{(1)}}(y)dy = \int_{\theta}^{\infty} y \times ne^{-n(y-\theta)}dy.$$

In the last integral, let

$$u = y - \theta \implies du = dy$$

so that

$$E(Y_{(1)}) = \int_0^{\infty} (u+\theta)\ ne^{-nu}du = E(U+\theta),$$

where $U \sim \text{exponential}(1/n)$; note that ne^{-nu} is the exponential$(1/n)$ pdf and the last integral is over $(0,\infty)$. Therefore,

$$E(\widehat{\theta}_2) = E(Y_{(1)}) = E(U+\theta) = E(U) + \theta = \frac{1}{n} + \theta \implies E\left(Y_{(1)} - \frac{1}{n}\right) = \theta.$$

Therefore, the "adjusted" version of the MLE; i.e.,

$$Y_{(1)} - \frac{1}{n},$$

is an unbiased estimator of θ. We now want to calculate the efficiency of

$$\overline{Y} - 1 \ \ \text{(MOM)} \ \ \text{to} \ \ Y_{(1)} - \frac{1}{n} \ \ \text{(adjusted MLE)}.$$

That is, we want to calculate

$$\text{eff}\left(\overline{Y} - 1 \ \ \text{to} \ \ Y_{(1)} - \frac{1}{n}\right) = \frac{V(Y_{(1)} - \frac{1}{n})}{V(\overline{Y} - 1)}.$$

This calculation makes sense because both the MOM and the adjusted MLE are unbiased. First, we have

$$V(\overline{Y} - 1) = V(\overline{Y}) = \frac{\sigma^2}{n},$$

where $\sigma^2 = V(Y)$, the population variance. The population variance of Y is 1. Note that $f_Y(y)$ is the exponential(1) pdf shifted to the right by θ units. The right shift will not affect the variance, so $V(Y)$ is the same as it would be if Y were exponential(1). Therefore,

$$V(\overline{Y} - 1) = \frac{1}{n}.$$

Now, for the numerator. We have

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

Let's get the second moment of $Y_{(1)}$. We have

$$E(Y_{(1)}^2) = \int_{\mathbb{R}} y^2 f_{Y_{(1)}}(y)dy = \int_\theta^\infty y^2 \times ne^{-n(y-\theta)}dy.$$

In the last integral, let

$$u = y - \theta \implies du = dy$$

so that

$$E(Y_{(1)}^2) = \int_0^\infty (u+\theta)^2 \; ne^{-nu}du = E[(U+\theta)^2],$$

where $U \sim$ exponential$(1/n)$; note that ne^{-nu} is the exponential$(1/n)$ pdf and the last integral is over $(0,\infty)$. Therefore,

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

The second moment of $U \sim$ exponential$(1/n)$ is

$$E(U^2) = V(U) + [E(U)]^2 = \frac{1}{n^2} + \frac{1}{n^2} = \frac{2}{n^2}.$$

Therefore,

$$E(Y_{(1)}^2) = E(U^2) + 2\theta\left(\frac{1}{n}\right) + \theta^2 = \frac{2}{n^2} + \frac{2\theta}{n} + \theta^2.$$

Therefore,

$$\begin{aligned} V\left((Y_{(1)} - \frac{1}{n}\right) &= E(Y_{(1)}^2) - \left(\frac{1}{n}+\theta\right)^2 \\ &= \frac{2}{n^2} + \frac{2\theta}{n} + \theta^2 - \left(\frac{1}{n}+\theta\right)^2 = \frac{2}{n^2} + \frac{2\theta}{n} + \theta^2 - \left(\frac{1}{n^2} + \frac{2\theta}{n} + \theta^2\right) = \frac{1}{n^2}. \end{aligned}$$

Finally,

$$\text{eff}\left(\overline{Y} - 1 \;\text{ to }\; Y_{(1)} - \frac{1}{n}\right) = \frac{V(Y_{(1)} - \frac{1}{n})}{V(\overline{Y}-1)} = \frac{1/n^2}{1/n} = \frac{1}{n} < 1.$$

This means the adjusted MLE $Y_{(1)} - \frac{1}{n}$ is only $(1/n)$th as variable the MOM estimator $\overline{Y} - 1$; i.e., the adjusted MLE is n times *more* efficient!