9.19. In this problem, $Y_1, Y_2, ..., Y_n$ is an iid sample from a beta$(\theta, 1)$ population, where $\theta > 0$ is unknown. The population pdf is

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

In general, the WLLN says

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow{p} E(Y) = \mu,$$

as $n \to \infty$. The expected value of $Y \sim$ beta$(\theta, 1)$ is

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

Therefore,

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow{p} \frac{\theta}{\theta + 1},$$

as $n \to \infty$. That is, $\overline{Y}$ is a consistent estimator of $\theta/(\theta + 1)$.

9.24. In this problem, $Y_1, Y_2, ..., Y_n$ is an iid sample from a $\mathcal{N}(0, 1)$ population distribution. For part (a), we know

$$U_i = Y_i^2 \sim \chi^2(1).$$

Therefore, $U_1, U_2, ..., U_n$ are iid $\chi^2(1)$ and therefore

$$\sum_{i=1}^{n} Y_i^2 = \sum_{i=1}^{n} U_i \sim \chi^2(n).$$

Recall the "degrees of freedom add" because of independence.

(b) Note that

$$W_n = \frac{1}{n}\sum_{i=1}^{n} Y_i^2 = \frac{1}{n}\sum_{i=1}^{n} U_i$$

is the sample mean of $U_1, U_2, ..., U_n$, which are iid $\chi^2(1)$. From the WLLN, we know

$$W_n = \frac{1}{n}\sum_{i=1}^{n} U_i \xrightarrow{p} E(U) = 1,$$

as $n \to \infty$. Recall the mean of a χ^2 random variable is equal to its degrees of freedom.

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

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

In general, the WLLN says

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow{p} E(Y) = \mu,$$

as $n \to \infty$. However, note that

$$E(Y) = \int_{\mathbb{R}} y f_Y(y) dy = \int_2^{\infty} \frac{2}{y}\, dy = 2\left(\ln y\Big|_2^{\infty}\right) = +\infty.$$

Therefore, the WLLN does not apply in this case. Recall when we proved the WLLN, we assumed $\sigma^2 = V(Y) < \infty$ which requires finite second moments; i.e., $E(Y^2) < \infty$. In this population, the first moment $E(Y)$ isn't even finite.

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

- $X_1, X_2, ..., X_n$ is an iid sample from a Poisson(λ_1) population (Bacteria A counts)
- $Y_1, Y_2, ..., Y_n$ is an iid sample from a Poisson(λ_2) population (Bacteria B counts).

The goal is to estimate the parameter

$$\frac{\lambda_1}{\lambda_1 + \lambda_2} = \frac{\text{mean number of Bacteria A}}{\text{mean number of both bacteria combined}}.$$

Define the sample means

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{and} \quad \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i.$$

Because $\overline{X}$ is an unbiased estimator for $E(X) = \lambda_1$ and $\overline{Y}$ is an unbiased estimator for $E(Y) = \lambda_2$, an obvious estimator to suggest is

$$\frac{\overline{X}}{\overline{X} + \overline{Y}}.$$

This is *not* an unbiased estimator of $\lambda_1/(\lambda_1 + \lambda_2)$ because

$$E\left(\frac{\overline{X}}{\overline{X} + \overline{Y}}\right) \neq \frac{E(\overline{X})}{E(\overline{X} + \overline{Y})} = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

Note that expectations are linear (i.e., the expectation of a ratio is *not* the ratio of the expectations). However, our proposed estimator is a consistent estimator. From the WLLN, we know

$$\overline{X} \xrightarrow{p} \lambda_1 \quad \text{and} \quad \overline{Y} \xrightarrow{p} \lambda_2.$$

Therefore,

$$\overline{X} + \overline{Y} \xrightarrow{p} \lambda_1 + \lambda_2$$

and therefore

$$\frac{\overline{X}}{\overline{X} + \overline{Y}} \xrightarrow{p} \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

All probability limits above apply as $n \to \infty$. This example illustrates an important point−even biased estimators can be consistent! This is not a contradiction. Bias is a finite-sample concept (that is, for fixed sample size n). On the other hand, consistency is a large-sample concept; i.e., a concept that explores what is happening when the sample size $n \to \infty$.

Another fact about the estimator

$$\frac{\overline{X}}{\overline{X}+\overline{Y}}$$

is that it is the MLE of $\lambda_1/(\lambda_1+\lambda_2)$. Recall we showed in the notes (Example 9.17) that the sample mean is the MLE of a Poisson population mean. Therefore, $\overline{X}$ is the MLE of λ_1 and $\overline{Y}$ is the MLE of λ_2. By the **invariance property** of MLEs,

$$\frac{\overline{X}}{\overline{X}+\overline{Y}} \quad \text{is the MLE of} \quad \frac{\lambda_1}{\lambda_1+\lambda_2}.$$

9.81. In this problem, $Y_1, Y_2, ..., Y_n$ is an iid sample from an exponential(θ) population distribution, where $\theta > 0$ is unknown. The population pdf is

$$f_Y(y|\theta) = \begin{cases} \frac{1}{\theta}e^{-y/\theta}, & y > 0 \\ 0, & \text{otherwise.} \end{cases}$$

The likelihood function is given by

$$L(\theta|\mathbf{y}) = \frac{1}{\theta}e^{-y_1/\theta} \times \frac{1}{\theta}e^{-y_2/\theta} \times \cdots \times \frac{1}{\theta}e^{-y_n/\theta} \;=\; \left(\frac{1}{\theta}\right)^n e^{-\sum_{i=1}^n y_i/\theta}.$$

The log-likelihood function is given by

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

The derivative of the log-likelihood function is given by

$$\begin{aligned} \frac{\partial}{\partial\theta}\ln L(\theta|\mathbf{y}) &= -\frac{n}{\theta} + \frac{\sum_{i=1}^n y_i}{\theta^2} \stackrel{\text{set}}{=} 0 \\ &\Longrightarrow\; -n\theta + \sum_{i=1}^n y_i = 0 \;\Longrightarrow\; \sum_{i=1}^n y_i = n\theta \;\Longrightarrow\; \widehat{\theta} = \frac{1}{n}\sum_{i=1}^n y_i = \overline{y}. \end{aligned}$$

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

$$\frac{\partial^2}{\partial\theta^2}\ln L(\theta|\mathbf{y}) = \frac{n}{\theta^2} - \frac{2\sum_{i=1}^n y_i}{\theta^3}.$$

Note that

$$\frac{\partial^2}{\partial\theta^2}\ln L(\theta|\mathbf{y})\Big|_{\theta=\overline{y}} = \frac{n}{\overline{y}^2} - \frac{2\sum_{i=1}^n y_i}{\overline{y}^3} = \frac{n\overline{y}}{\overline{y}^3} - \frac{2n\overline{y}}{\overline{y}^3} = -\frac{n\overline{y}}{\overline{y}^3} = -\frac{n}{\overline{y}^2} < 0.$$

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

$$\widehat{\theta} = \frac{1}{n}\sum_{i=1}^n Y_i = \overline{Y},$$

the sample mean. We have shown $\overline{Y}$ is the MLE of θ. Therefore, $\overline{Y}^2$ is the MLE of θ^2, by the **invariance property** of MLEs.

9.82. In this problem, $Y_1, Y_2, ..., Y_n$ is an iid sample from a Weibull(r, θ) population, where r is known and $\theta > 0$ is unknown.

Note: The Rayleigh population distribution is a special case of the Weibull when $r = 2$; in the notes (Example 9.18, pp 144), we determined the MLE of θ is

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

In Problem 9.81 (above), the exponential distribution is a special case of the Weibull when $r = 1$; we determined the MLE of θ is

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

For the general Weibull(r, θ) population distribution, the MLE of θ (when r is known) is

$$\widehat{\theta} = \frac{1}{n}\sum_{i=1}^{n} Y_i^r,$$

the rth sample moment. We now show this. The likelihood function is given by

$$L(\theta|\mathbf{y}) = \frac{ry_1^{r-1}}{\theta}e^{-y_1^r/\theta} \times \frac{ry_2^{r-1}}{\theta}e^{-y_2^r/\theta} \times \cdots \times \frac{ry_n^{r-1}}{\theta}e^{-y_n^r/\theta} = \left(\frac{r}{\theta}\right)^n \left(\prod_{i=1}^{n} y_i^{r-1}\right) e^{-\sum_{i=1}^{n} y_i^r/\theta}.$$

Note that we can write

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

where $t = \sum_{i=1}^{n} y_i^r$. By the Factorization Theorem, it follows that $T = \sum_{i=1}^{n} Y_i^r$ is a sufficient statistic; this is part (a).

(b) The log-likelihood function is given by

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

The derivative of the log-likelihood function is given by

$$\begin{aligned}\frac{\partial}{\partial\theta}\ln L(\theta|\mathbf{y}) &= -\frac{n}{\theta}+\frac{\sum_{i=1}^n y_i^r}{\theta^2} \stackrel{\text{set}}{=} 0 \\ &\Longrightarrow -n\theta+\sum_{i=1}^n y_i^r = 0 \Longrightarrow \sum_{i=1}^n y_i^r = n\theta \Longrightarrow \widehat{\theta}=\frac{1}{n}\sum_{i=1}^n y_i^r.\end{aligned}$$

Showing this first-order critical point $\widehat{\theta}$ maximizes $\ln L(\theta|\mathbf{y})$ is done in the same way as in the exponential and Rayleigh cases. The second derivative of the log-likelihood function is given by

$$\frac{\partial^2}{\partial\theta^2}\ln L(\theta|\mathbf{y}) = \frac{n}{\theta^2}-\frac{2\sum_{i=1}^n y_i^r}{\theta^3}.$$

Note that

$$\frac{\partial^2}{\partial\theta^2}\ln L(\theta|\mathbf{y})\Big|_{\theta=\widehat{\theta}} = \frac{n}{\widehat{\theta}^2}-\frac{2\sum_{i=1}^n y_i^r}{\widehat{\theta}^3} = \frac{n\widehat{\theta}}{\widehat{\theta}^3}-\frac{2n\widehat{\theta}}{\widehat{\theta}^3} = -\frac{n\widehat{\theta}}{\widehat{\theta}^3} = -\frac{n}{\widehat{\theta}^2} < 0.$$

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

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

(c) The MLE $\widehat{\theta}$ above is the MVUE of θ (when r is known). First note that $\widehat{\theta}$ is a function of the sufficient statistic $T=\sum_{i=1}^n Y_i^r$. Now, let's show that it is an unbiased estimator of θ. We have

$$E(\widehat{\theta}) = E\left(\frac{1}{n}\sum_{i=1}^n Y_i^r\right) = \frac{1}{n}\sum_{i=1}^n E(Y_i^r).$$

The rth moment of $Y\sim \text{Weibull}(r,\theta)$ is

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

In the last integral, let

$$u=y^r \quad\Longrightarrow\quad du = ry^{r-1}dy.$$

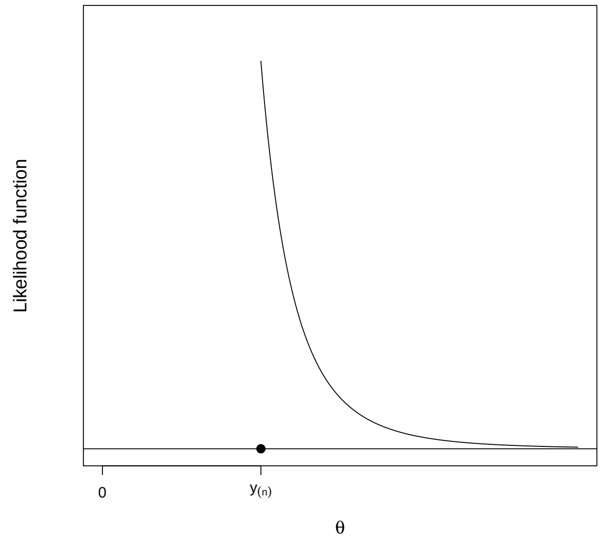
The limits on the integral do not change under this transformation. Therefore,

$$E(Y^r)=\int_0^\infty y^r\times\frac{ry^{r-1}}{\theta}e^{-y^r/\theta}dy \;=\; \int_0^\infty u\;\frac{ry^{r-1}}{\theta}e^{-u/\theta}\;\frac{du}{ry^{r-1}} = \int_0^\infty u\times\frac{1}{\theta}e^{-u/\theta}du = E(U),$$

where $U\sim\text{exponential}(\theta)$. Therefore $E(Y^r)=E(U)=\theta$. Therefore,

$$E(\widehat{\theta}) = \frac{1}{n}\sum_{i=1}^n\theta = \frac{1}{n}(n\theta)=\theta.$$

This proves $\widehat{\theta}$ is the MVUE of θ. It is a function of a sufficient statistic T and it is unbiased.



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

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

where the population parameter $\theta > 0$ is unknown. In Problem 9.63 (HW9), we calculated the likelihood function:

$$L(\theta|\mathbf{y}) = \left(\frac{3}{\theta^3}\right)^n \left(\prod_{i=1}^n y_i^2\right) I(0 \le y_{(n)} \le \theta).$$

We see the support depends on θ in the population pdf $f_Y(y)$ so $L(\theta|\mathbf{y})$ is not a differentiable function of θ. A graph of the likelihood function $L(\theta|\mathbf{y})$ is shown above. Note that

- For $\theta \ge y_{(n)}$, $L(\theta|\mathbf{y}) = \left(\frac{3}{\theta^3}\right)^n \prod_{i=1}^n y_i^2$, which is a decreasing function of θ (see above).
- For $\theta < y_{(n)}$, $L(\theta|\mathbf{y}) = 0$.

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

(b) In Problem 9.63 (HW9), we derived the pdf of $Y_{(n)}$ to be

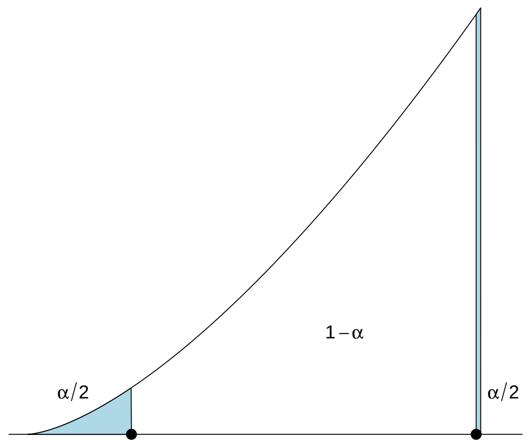
$$f_{Y_{(n)}}(y) = \begin{cases} \dfrac{3ny^{3n-1}}{\theta^{3n}}, & 0 \le y \le \theta \\ 0, & \text{otherwise.} \end{cases}$$

Consider the function

$$Q = \frac{Y_n}{\theta}.$$

We will now show the distribution of Q is free of θ; i.e., Q is a pivotal quantity. The support of Q is

$$R_Q = \{q : 0 \le q \le 1\}.$$



Therefore, the pdf of Q is nonzero over $[0,1]$. For $0 \leq q \leq 1$, the cdf of Q is

$$\begin{aligned} F_Q(q) = P(Q \leq q) = P\left(\frac{Y_{(n)}}{\theta} \leq q\right) &= P(Y_{(n)} \leq q\theta) \\ &= \int_0^{q\theta} f_{Y_{(n)}}(y)dy \\ &= \int_0^{q\theta} \frac{3ny^{3n-1}}{\theta^{3n}}dy \\ &= \frac{1}{\theta^{3n}}\left(y^{3n}\Big|_0^{q\theta}\right) = \frac{(q\theta)^{3n}}{\theta^{3n}} = q^{3n}. \end{aligned}$$

Summarizing,

$$F_Q(q) = \begin{cases} 0 & q < 0 \\ q^{3n}, & 0 \leq q \leq 1 \\ 1, & q > 1. \end{cases}$$

Therefore, Q is a pivotal quantity because its distribution does not depend on θ. Taking derivatives, the pdf of Q is

$$f_Q(q) = \begin{cases} 3nq^{3n-1}, & 0 \leq q \leq 1 \\ 0, & \text{otherwise.} \end{cases}$$

We recognize $f_Q(q)$ as a beta pdf with $\alpha = 3n$ and $\beta = 1$.

(c) Define

$$\begin{aligned} b_{3n,1,1-\alpha/2} &= \textbf{lower } \alpha/2 \text{ quantile of beta}(3n,1) \\ b_{3n,1,\alpha/2} &= \textbf{upper } \alpha/2 \text{ quantile of beta}(3n,1); \end{aligned}$$

see the figure at the top of the previous page where I have graphed the beta$(3n, 1)$ pdf. Because $Q \sim \text{beta}(3n, 1)$, we can write

$$\begin{aligned}
1-\alpha &= P\left(b_{3n,1,1-\alpha/2} < Q < b_{3n,1,\alpha/2}\right) \\
&= \left(b_{3n,1,1-\alpha/2} < \frac{Y_{(n)}}{\theta} < b_{3n,1,\alpha/2}\right) \\
&= P\left(\frac{1}{b_{3n,1,1-\alpha/2}} > \frac{\theta}{Y_{(n)}} > \frac{1}{b_{3n,1,\alpha/2}}\right) \\
&= P\left(\frac{Y_{(n)}}{b_{3n,1,1-\alpha/2}} > \theta > \frac{Y_{(n)}}{b_{3n,1,\alpha/2}}\right) = P\Bigg(\underbrace{\frac{Y_{(n)}}{b_{3n,1,\alpha/2}}}_{\theta_L} < \theta < \underbrace{\frac{Y_{(n)}}{b_{3n,1,1-\alpha/2}}}_{\theta_U}\Bigg).
\end{aligned}$$

Therefore,

$$\left(\frac{Y_{(n)}}{b_{3n,1,\alpha/2}}, \ \frac{Y_{(n)}}{b_{3n,1,1-\alpha/2}}\right)$$

is a $100(1-\alpha)\%$ confidence interval for θ.

9.97. In this problem, $Y_1, Y_2, ..., Y_n$ is an iid sample from a geometric(p) population distribution, where the success probability p is unknown $(0 < p < 1)$. In part (a), we want to find the MOM estimator of p. 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

$$E(Y) = \frac{1}{p}$$

The first sample moment is

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

Therefore, the MOM estimator of p is found by solving

$$\frac{1}{p} \stackrel{\text{set}}{=} \overline{Y} \implies \widehat{p} = \frac{1}{\overline{Y}}.$$

(b) We now find the MLE. The likelihood function is

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

The log-likelihood function is

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

The derivative of the log-likelihood function is

$$\begin{aligned}
\frac{\partial}{\partial p}\ln L(p|\mathbf{y}) &= -\frac{(\sum_{i=1}^n y_i - n)}{1-p} + \frac{n}{p} \stackrel{\text{set}}{=} 0 \\
&\implies -p\left(\sum_{i=1}^n y_i - n\right) + n(1-p) = 0 \\
&\implies -p\sum_{i=1}^n y_i + np + n - np = 0 \\
&\implies -p\sum_{i=1}^n y_i + n = 0 \implies n = p\sum_{i=1}^n y_i \implies \widehat{p} = \frac{n}{\sum_{i=1}^n y_i} = \frac{1}{\overline{y}}.
\end{aligned}$$

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

$$\frac{\partial^2}{\partial p^2}\ln L(p|\mathbf{y}) = -\frac{(\sum_{i=1}^n y_i - n)}{(1-p)^2} - \frac{n}{p^2}.$$

Note that

$$\begin{aligned}
\frac{\partial^2}{\partial p^2}\ln L(p|\mathbf{y})\Big|_{p=\frac{1}{\overline{y}}} &= -\frac{(\sum_{i=1}^n y_i - n)}{\left(1-\frac{1}{\overline{y}}\right)^2} - \frac{n}{\left(\frac{1}{\overline{y}}\right)^2} \\
&= -\frac{(n\overline{y}-n)}{\left(\frac{\overline{y}-1}{\overline{y}^2}\right)^2} - n\overline{y}^2 = -\frac{n(\overline{y}-1)}{\left(\frac{\overline{y}-1}{\overline{y}^2}\right)^2} - n\overline{y}^2 = -\frac{n\overline{y}^2}{\overline{y}-1} - n\overline{y}^2 < 0,
\end{aligned}$$

because $\overline{y} > 1$. Therefore, $\widehat{p} = 1/\overline{y}$ maximizes $\ln L(p|\mathbf{y})$. The MLE of p is

$$\widehat{p} = \frac{1}{\overline{Y}}.$$

Therefore, the MOM estimator and MLE of p are the same in this example.

9.107. In this problem, $Y_1, Y_2, ..., Y_n$ is an iid sample from an exponential(θ) population distribution, where $\theta > 0$ is unknown. In Problem 9.81, we already showed $\widehat{\theta} = \overline{Y}$ is the MLE of θ. Therefore, the MLE of the reliability at time t, that is, $\overline{F}(t) = e^{-t/\theta}$, is $e^{-t/\overline{Y}}$, by invariance.

9.112. In this problem, $Y_1, Y_2, ..., Y_n$ is an iid sample from a Poisson(λ) population distribution, where $\lambda > 0$ is unknown. In part (a), we want to show

$$W_n = \frac{\overline{Y} - \lambda}{\sqrt{\overline{Y}/n}} \xrightarrow{d} \mathcal{N}(0,1),$$

as $n \to \infty$. In this population, we know $\mu = E(Y) = \lambda$ and $\sigma^2 = V(Y) = \lambda$. Therefore, from the CLT, we know

$$U_n = \frac{\overline{Y} - \lambda}{\sqrt{\lambda/n}} \xrightarrow{d} \mathcal{N}(0,1),$$

as $n \to \infty$. From the WLLN, we know

$$\overline{Y} \xrightarrow{p} \lambda \implies \frac{\overline{Y}}{\lambda} \xrightarrow{p} 1 \implies V_n = \sqrt{\frac{\overline{Y}}{\lambda}} \xrightarrow{p} 1;$$

the last two implications follow from continuity. Now, simply note that

$$W_n = \frac{\overline{Y} - \lambda}{\sqrt{\overline{Y}/n}} = \frac{\dfrac{\overline{Y} - \lambda}{\sqrt{\lambda/n}}}{\sqrt{\dfrac{\overline{Y}}{\lambda}}} = \frac{U_n}{V_n} \xrightarrow{d} \mathcal{N}(0,1),$$

by Slutsky's Theorem.

(b) Because

$$W_n = \frac{\overline{Y} - \lambda}{\sqrt{\overline{Y}/n}} \xrightarrow{d} \mathcal{N}(0,1),$$

as $n \to \infty$, we can write

$$\begin{aligned}
1 - \alpha \approx P\left(-z_{\alpha/2} < \frac{\overline{Y} - \lambda}{\sqrt{\overline{Y}/n}} < z_{\alpha/2}\right) &= P\left(-z_{\alpha/2}\sqrt{\frac{\overline{Y}}{n}} < \overline{Y} - \lambda < z_{\alpha/2}\sqrt{\frac{\overline{Y}}{n}}\right) \\
&= P\left(z_{\alpha/2}\sqrt{\frac{\overline{Y}}{n}} > \lambda - \overline{Y} > -z_{\alpha/2}\sqrt{\frac{\overline{Y}}{n}}\right) \\
&= P\left(\overline{Y} + z_{\alpha/2}\sqrt{\frac{\overline{Y}}{n}} > \lambda > \overline{Y} - z_{\alpha/2}\sqrt{\frac{\overline{Y}}{n}}\right) \\
&= P\left(\overline{Y} - z_{\alpha/2}\sqrt{\frac{\overline{Y}}{n}} < \lambda < \overline{Y} + z_{\alpha/2}\sqrt{\frac{\overline{Y}}{n}}\right).
\end{aligned}$$

This argument proves

$$\left(\overline{Y} - z_{\alpha/2}\sqrt{\frac{\overline{Y}}{n}},\ \overline{Y} + z_{\alpha/2}\sqrt{\frac{\overline{Y}}{n}}\right)$$

is a large-sample $100(1-\alpha)\%$ confidence interval for the population mean λ.