1 Chapter 1

Exercise 1.1(a)

Let X be the sum of the two number obtained.
Let X_1 be the number obtained on Die 1.
Let X_2 be the number obtained on Die 2.
Thus, $X = X_1 + X_2$, and

$$E_x = \{X = x\} = \{X_1 + X_2 = x\} = \{\ X_1 = x_1, X_2 = x - x_1 \mid 1 \leq x_1, x - x_1 \leq 6\ \}$$

Now,

$$1 \leq x - x_1 \leq 6 \quad \Longleftrightarrow \quad -1 \geq x_1 - x \geq -6 \quad \Longleftrightarrow \quad x - 1 \geq x_1 \geq x - 6 \quad \Longleftrightarrow \quad x - 6 \leq x_1 \leq x - 1$$

Hence,

$$E_x = \{X = x\} = \{X_1 + X_2 = x\} = \{\ X_1 = x_1, X_2 = x - x_1 \mid \max\{1, x-6\} \leq x_1 \leq \min\{6, x-1\}\ \}$$

$$\begin{aligned} P(E_x) &= \sum_{x_1=\max\{1,x-6\}}^{\min\{6,x-1\}} P(X_1 = x_1, X_2 = x - x_1) = \sum_{x_1=\max\{1,x-6\}}^{\min\{6,x-1\}} \frac{1}{6^2} \\ &= \frac{1}{6^2}\left(\min\{6, x-1\} - \max\{1, x-6\} + 1\right) \end{aligned}$$

Next, note that

$$\min\{6, x-1\} = \begin{cases} x-1, & \text{if } x = 2, 3, \ldots, 6 \\ 6, & \text{if } x = 7, 8, \ldots, 12 \end{cases} \quad \text{and} \quad \max\{1, x-6\} = \begin{cases} 1, & \text{if } x = 2, 3, \ldots, 6 \\ x-6, & \text{if } x = 7, 8, \ldots, 12 \end{cases}$$

Hence,

$$\begin{aligned} P(E_x) &= \frac{1}{6^2}\left(\min\{6, x-1\} - \max\{1, x-6\} + 1\right) = \frac{1}{36}\begin{cases} (x-1) - 1 + 1, & \text{if } x = 2, 3, \ldots, 6 \\ 6 - (x-6) + 1, & \text{if } x = 7, 8, \ldots, 12 \end{cases} \\ &= \frac{1}{36}\begin{cases} x-1, & \text{if } x = 2, 3, \ldots, 6 \\ 13 - x, & \text{if } x = 7, 8, \ldots, 12 \end{cases} \end{aligned}$$

□

Exercise 1.18

Recapitulation of the rules of craps: *Let x be the number obtained on the first roll. If $x \in \{7, 11\}$, then the player wins. If $x \in \{2, 3, 12\}$, then the player loses. If $x \in \{4, 5, 6, 8, 9, 10\}$, then the player keeps rolling, until either 7 is rolled or x is rolled. If x is rolled first (before 7 is rolled), then the player wins. If 7 is rolled first (before x is rolled), then the player loses.*

Let W be the $\{0, 1\}$-valued random variable such that $W = 1$ if the player wins, and $W = 0$ if the player loses. We thus seek to compute $P(W = 1)$. Let X be (the random variable of) the sum of the two numbers obtained on the first roll. Note that $\text{Range}(X) = \{\, 2, 3, 4, \ldots, 12 \,\}$. Then,

$$\begin{aligned}
P(W = 1) &= \sum_{x=2}^{12} P(W = 1|X = x) \cdot P(X = x) \\
&= P(W = 1|X = 7)\, P(X = 7) + P(W = 1|X = 11)\, P(X = 11) + \sum_{x \in \{4,5,6,8,9,10\}} P(W = 1|X = x) \cdot P(X = x)
\end{aligned}$$

Now, note that $P(W = 1|X = 7) = P(W = 1|X = 11) = 1$, $P(X = 7) = \dfrac{6}{36} = \dfrac{1}{6}$, and $P(X = 11) = \dfrac{2}{36} = \dfrac{1}{18}$.

From Exercise 1.1(a), we have:

$$\begin{aligned}
P(X = x) &= \frac{1}{6^2}\left(\min\{6, x-1\} - \max\{1, x-6\} + 1\right) = \frac{1}{36}\begin{cases} (x-1) - 1 + 1, & \text{if } x = 2,\, 3,\, \ldots,\, 6 \\ 6 - (x-6) + 1, & \text{if } x = 7,\, 8,\, \ldots,\, 12 \end{cases} \\
&= \frac{1}{36}\begin{cases} x - 1, & \text{if } x = 2,\, 3,\, \ldots,\, 6 \\ 13 - x, & \text{if } x = 7,\, 8,\, \ldots,\, 12 \end{cases}
\end{aligned}$$

Next, let Y_n be the random variable of the sum of the two numbers obtained on the $(n + 1)$st roll. Then,

$$\begin{aligned}
P(W = 1|X = x) &= \sum_{n=1}^{\infty} \left[1 - P(Y_n = 7) - P(Y_n = x)\right]^{n-1} \cdot P(X = x) \\
&= P(X = x) \cdot \sum_{n=1}^{\infty} \left[1 - P(Y_n = 7) - P(Y_n = x)\right]^{n-1} \\
&= P(X = x) \cdot \frac{1}{1 - \left[1 - P(Y = 7) - P(Y = x)\right]} \\
&= \frac{P(X = x)}{P(Y = 7) + P(Y = x)} \\
&= \frac{P(X = x)}{\frac{1}{6} + P(Y = x)}
\end{aligned}$$

Hence,

$$
\begin{aligned}
P(W=1) &= \sum_{x=2}^{12} P(W=1|X=x)\cdot P(X=x) \\
&= P(W=1|X=7)\,P(X=7) + P(W=1|X=11)\,P(X=11) + \sum_{x\in\{4,5,6,8,9,10\}} P(W=1|X=x)\cdot P(X=x) \\
&= \frac{6}{36} + \frac{2}{36} + \sum_{x\in\{4,5,6,8,9,10\}} \frac{P(X=x)^2}{\frac{1}{6}+P(Y=x)} \\
&= \frac{6}{36} + \frac{2}{36} + \frac{(\frac{4-1}{36})^2}{\frac{1}{6}+\frac{4-1}{36}} + \frac{(\frac{5-1}{36})^2}{\frac{1}{6}+\frac{5-1}{36}} + \frac{(\frac{6-1}{36})^2}{\frac{1}{6}+\frac{6-1}{36}} + \frac{(\frac{13-8}{36})^2}{\frac{1}{6}+\frac{13-8}{36}} + \frac{(\frac{13-9}{36})^2}{\frac{1}{6}+\frac{13-9}{36}} + \frac{(\frac{13-10}{36})^2}{\frac{1}{6}+\frac{13-10}{36}} \\
&= \frac{6}{36} + \frac{2}{36} + \frac{(1/36)^2}{1/36}\left(\frac{3^2}{6+3} + \frac{4^2}{6+4} + \frac{5^2}{6+5} + \frac{5^2}{6+5} + \frac{4^2}{6+4} + \frac{3^2}{6+3}\right) \\
&= \frac{6}{36} + \frac{2}{36} + \frac{2}{36}\left(\frac{3^2}{6+3} + \frac{4^2}{6+4} + \frac{5^2}{6+5}\right) \;=\; \frac{1}{36}\left[6+2+2\left(\frac{9}{9}+\frac{16}{10}+\frac{25}{11}\right)\right] \\
&= \frac{1}{36}\left[8+2\left(\frac{536}{110}\right)\right] \;=\; \frac{1}{36}\left[\frac{1952}{110}\right] \;=\; \frac{1}{2^2\cdot 3^2}\left[\frac{2^5\cdot 61}{2\cdot 5\cdot 11}\right] \\
&= \frac{2^2\cdot 61}{3^2\cdot 5\cdot 11} \;\approx\; 0.4929293
\end{aligned}
$$

□

Exercise 1.19(a)

Let n be the number of workers in the sample. Let X_i, $i = 1, 2, \ldots, n$, be $\{0,\ 1\}$-valued random variables defined by:

$$X_i \;=\; \begin{cases} 1, & \text{if the } i\text{th subject is highly exposed,} \\ 0, & \text{if the } i\text{th subject is NOT highly exposed} \end{cases}$$

Define

$$S_n \;:=\; \sum_{i=1}^{n} X_i\,, \quad \text{and} \quad S_{n-1} \;:=\; \sum_{i=1}^{n-1} X_i\,.$$

First, note that

$$\theta_n \;=\; P(S_n \text{ is even})\,, \quad \text{and} \quad \theta_{n-1} \;=\; P(S_{n-1} \text{ is even})\,.$$

Note also that

$$\begin{aligned} \theta_n &= P(S_n \text{ is even}) \;=\; P(X_n = 1)P(S_{n-1} \text{ is odd}) \;+\; P(X_n = 0)P(S_{n-1} \text{ is even}) \\ &= \pi_h\,(1-\theta_{n-1}) + (1-\pi_h)\,\theta_{n-1} \;=\; \pi_h + (1-2\pi_h)\,\theta_{n-1} \end{aligned}$$

Thus, the desired difference equation is:

$$\theta_n \;=\; \pi_h + (1-2\pi_h)\,\theta_{n-1} \tag{1.1}$$

Exercise 1.19(b)

To solve the difference equation (1.1) obtained in Exercise 1.19(a), we assume that θ_n has the following form:

$$\theta_n = \alpha + \beta\,\gamma^n \tag{1.2}$$

where α, β, and γ are unknown constants to be determined. We first make the following:

Observation: $\beta \neq 0$ and $\gamma \notin \{0,\ 1\}$.
Indeed, if $\beta = 0$ or $\gamma \in \{0,\ 1\}$, then θ_n would be constant in n. In that case, define $\theta := \theta_n = \theta_{n-1} = \cdots$. By the difference equation (1.1), we would then have

$$\theta = \pi_h + (1-2\pi_h)\,\theta \quad\Longrightarrow\quad 0 = \pi_h\,(1-2\theta) \quad\Longrightarrow\quad \theta = \frac{1}{2} \quad (\text{since } \pi_h > 0)$$

However, this contradicts the initial condition that $\theta_0 = 1$. Thus, this proves the assertion that $\beta \neq 0$ and $\gamma \notin \{0,\ 1\}$. (Note that if the sample size is 0, then the number of highly exposed subjects must be 0; hence $\theta_0 = P(S_0 \text{ is even}) = 1$, since we have here adopted the convention that 0 is "even.")

Now, substituting (1.2) into (1.1) yields:

$$\begin{aligned} \alpha + \beta\gamma^n &= \theta_n \;=\; \pi_h + (1-2\pi_h)\,\theta_{n-1} \\ &= \pi_h + (1-2\pi_h)\left(\alpha + \beta\gamma^{n-1}\right) \\ &= \alpha + \pi_h(1-2\alpha) + \beta\gamma^{n-1}(1-2\pi_h) \end{aligned}$$

Collecting terms involving γ on the right-hand side yields:

$$\pi_h(2\alpha-1) \;=\; \beta\gamma^{n-1}\,(1-2\pi_h-\gamma)$$

Now, note that the left-hand side of the preceding equation is independent of γ, while the right-hand side is a scalar multiple of the $(n-1)$th power of γ; in other words, the right-hand side is a scalar multiple of a power of γ which is constant in n.

This happens if and only if either $\gamma \in \{0,\,1\}$, or if the coefficient $\beta(1-2\pi_h-\gamma)=0$. The preceding Observation (i.e. $\beta \neq 0$ and $\gamma \notin \{0,\,1\}$) thus implies:

$$\gamma = 1 - 2\pi_h$$

Since $\pi_h > 0$, we furthermore conclude that

$$\alpha = \frac{1}{2}$$

We therefore have:

$$\theta_n \;=\; \frac{1}{2} + \beta\,(1-2\pi_h)^n$$

The initial condition $\theta_0 = 1$ now implies:

$$1 = \theta_0 = \frac{1}{2} + \beta\,(1-2\pi_h)^0 = \frac{1}{2} + \beta \quad\Longrightarrow\quad \beta = \frac{1}{2}$$

We may now conclude:

$$\theta_n \;=\; \frac{1}{2} + \frac{1}{2}(1-2\pi_h)^n$$

Lastly, if $\pi_h = 0.05$, then

$$\theta_{50} \;=\; \frac{1}{2} + \frac{1}{2}(1-2\times 0.05)^{50} \;\approx\; 0.5025769$$

$\square$

Comment: For $0 < \pi_h < \frac{1}{2}$, the formula $\theta_n = \frac{1}{2} + \frac{1}{2}(1-2\pi_h)^n$ implies that $\theta_n > \frac{1}{2}$, for any $n = 1, 2, 3, \ldots$; in other words, there is a higher than $50:50$ chance that the number of highly exposed subjects in the sample is "even", whenever $0 < \pi_h < \frac{1}{2}$. This apparent asymmetry between odd and even is NOT surprising given the fact that 0 is regarded as "even" here, and that the probability that there are no highly exposed workers in the sample is high if π_h is "small" (e.g. $0 < \pi_h < \frac{1}{2}$).

Exercise 1.20(a)

$$p(D|S,x) = \frac{p(D,S,x)}{p(S,x)} = \frac{p(D,S,x)}{p(D,x)}\,\frac{p(D,x)}{p(S,x)} = p(S|D,x)\,\frac{p(D,x)/p(x)}{p(S,x)/p(x)} = p(S|D,x)\,\frac{p(D|x)}{p(S|x)}$$

Now, we are given that

$$p(S|D,x) = \pi_1\,, \quad \text{and} \quad p(D|x) = \frac{\exp\left(\beta_0+\beta^T x\right)}{1+\exp\left(\beta_0+\beta^T x\right)}$$

So, we now proceed to compute $p(S|x)$. To this end,

$$\begin{aligned} p(S|x) &= \frac{p(S,x)}{p(x)} = \frac{1}{p(x)}\left(p(S,D,x)+p(S,\overline{D},x)\right) = \frac{1}{p(x)}\left(\frac{p(S,D,x)}{p(D,x)}p(D,x)+\frac{p(S,\overline{D},x)}{p(\overline{D},x)}p(\overline{D},x)\right) \\ &= p(S|D,x)p(D|x)+p(S|\overline{D},x)p(\overline{D}|x) \end{aligned}$$

Hence,

$$\begin{aligned} p(D|S,x) &= p(S|D,x)\frac{p(D|x)}{p(S|x)} = \frac{p(S|D,x)\,p(D|x)}{p(S|D,x)\,p(D|x)+p(S|\overline{D},x)\,p(\overline{D}|x)} = \frac{\pi_1\cdot p(D|x)}{\pi_1\cdot p(D|x)+\pi_0\cdot p(\overline{D}|x)} \\ &= \frac{\pi_1\cdot\dfrac{\exp\left(\beta_0+\beta^T x\right)}{1+\exp\left(\beta_0+\beta^T x\right)}}{\pi_1\cdot\dfrac{\exp\left(\beta_0+\beta^T x\right)}{1+\exp\left(\beta_0+\beta^T x\right)}+\pi_0\cdot\dfrac{1}{1+\exp\left(\beta_0+\beta^T x\right)}} = \frac{\pi_1\cdot\exp\left(\beta_0+\beta^T x\right)}{\pi_1\cdot\exp\left(\beta_0+\beta^T x\right)+\pi_0} \\ &= \frac{\dfrac{\pi_1}{\pi_0}\cdot\exp\left(\beta_0+\beta^T x\right)}{1+\dfrac{\pi_1}{\pi_0}\cdot\exp\left(\beta_0+\beta^T x\right)} = \frac{\exp\left[\,\log(\pi_1/\pi_0)+\beta_0+\beta^T x\,\right]}{1+\exp\left[\,\log(\pi_1/\pi_0)+\beta_0+\beta^T x\,\right]}\,, \end{aligned}$$

as required.

Comment: The above derivations show that, in a case-control study, if one has knowledge (or good estimate) of the ratio π_1/π_0, one can obtain an estimate for $p(D|x)$, the disease risk associated to covariate value x, from the quantity $p(D|S,x)$, which can be estimated from case-control study data as follows:

$$p(D|S,x) \;\approx\; \frac{\#(\text{subjects in sample with disease and covariate value } x)}{\#(\text{subjects in sample with covariate value } x)}$$

However, in practice, the ratio π_1/π_0 is rarely, if ever, known. And, without knowledge or estimate of π_1/π_0, the disease risk $p(D|x)$ associated to covariate value x can NOT be estimated based on data from a case-control study.

Exercise 1.20(b)

First, note that

$$\frac{p(D|x^*)}{p(\overline{D}|x^*)} = \frac{\exp\left(\beta_0+\beta^T x^*\right)/(1+\exp\left(\beta_0+\beta^T x^*\right))}{1/(1+\exp\left(\beta_0+\beta^T x^*\right))} = \exp\left(\beta_0+\beta^T x^*\right)$$

Similarly,

$$\frac{p(D|x)}{p(\overline{D}|x)} = \exp\left(\beta_0+\beta^T x\right)$$

Hence,

$$\theta_r = \theta_r(x^*,x) = \frac{p(D|x^*)/p(\overline{D}|x^*)}{p(D|x)/p(\overline{D}|x)} = \frac{\exp\left(\beta_0+\beta^T x^*\right)}{\exp\left(\beta_0+\beta^T x\right)} = \exp\left[\,\beta^T(x^*-x)\,\right]\,,$$

as required. Next,

$$\theta_c \;=\; \theta_c(x^*,x) \;=\; \frac{p(D|S,x^*)/p(\overline{D}|S,x^*)}{p(D|S,x)/p(\overline{D}|S,x)} \;=\; \frac{\exp\left[\,\log(\pi_1/\pi_0)+\beta_0+\beta^T x^*\,\right]}{\exp\left[\,\log(\pi_1/\pi_0)+\beta_0+\beta^T x\,\right]} \;=\; \exp\left[\,\beta^T(x^*-x)\,\right]\,,$$

as required.

Comment: Exercise 1.20(a) showed that, without knowledge or estimate of the ratio π_1/π_0, case-control study data can NOT be used to estimate the disease $p(D|x)$ associated to covariate value x. On the other hand, case-control study data can be readily used to estimate the odds ratio

$$\theta_c \;=\; \theta_c(x^*,x) \;:=\; \frac{p(D|S,x^*)/p(\overline{D}|S,x^*)}{p(D|S,x)/p(\overline{D}|S,x)}$$

Exercise 1.20(b) shows that θ_c is equal to

$$\theta_r \;=\; \theta_r(x^*,x) \;:=\; \frac{p(D|x^*)/p(\overline{D}|x^*)}{p(D|x)/p(\overline{D}|x)}$$

Thus, Exercise 1.20(a) and Exercise 1.20(b) together show that, while case-control study data can NOT be used to estimate disease risk $p(D|x)$ associated to covariate value x, they can be used to estimate the disease odds ratio

$$\theta_r \;=\; \theta_r(x^*,x) \;:=\; \frac{p(D|x^*)/p(\overline{D}|x^*)}{p(D|x)/p(\overline{D}|x)}$$

associated to the covariate value x^* against x.

Exercise 1.21(a)

Let D be the random variable defined by:

$$D \;:=\; \begin{cases} 1, & \text{if a given individual has IBD,} \\ 0, & \text{otherwise.} \end{cases}$$

Let S_1 be the random variable defined by:

$$S_1 \;:=\; \begin{cases} 1, & \text{Strategy \#1 asserts that a given individual has IBD,} \\ 0, & \text{otherwise.} \end{cases}$$

Let S_2 be the random variable defined by:

$$S_2 \;:=\; \begin{cases} 1, & \text{Strategy \#2 asserts that a given individual has IBD,} \\ 0, & \text{otherwise.} \end{cases}$$

Note that

$$\begin{aligned} P(S_1 = D) &= P(S_1 = D, D = 1) + P(S_1 = D, D = 0) \;=\; P(S_1 = D|D = 1)P(D = 1) + P(S_1 = D|D = 0)P(D = 0) \\ &= P(S_1 = D|D = 1)\theta + P(S_1 = D|D = 0)(1-\theta) \end{aligned}$$

Next, note that

$$\begin{aligned} P(S_1 = D|D = 1) &= P(X \geq 2), \quad \text{where } X \sim \text{Binomial}(n = 3, p = \pi_1) \\ &= \binom{3}{2}\pi_1^2(1-\pi_1)^1 + \binom{3}{3}\pi_1^3(1-\pi_1)^0 \\ &= 3\pi_1^2(1-\pi_1) + \pi_1^3 \;=\; \pi_1^2\,(3 - 2\pi_1) \end{aligned}$$

Similarly,

$$P(S_1 = D|D = 0) \;=\; \pi_0^2\,(3 - 2\pi_0)$$

Therefore,

$$P(S_1 = D) \;=\; P(S_1 = D|D = 1)\theta + P(S_1 = D|D = 0)(1-\theta) \;=\; \theta\pi_1^2\,(3 - 2\pi_1) + (1-\theta)\pi_0^2\,(3 - 2\pi_0)$$

On the other hand, note that

$$P(S_2 = D|D = 1) \;=\; \pi_1 \quad \text{and} \quad P(S_2 = D|D = 0) \;=\; \pi_0$$

Hence,

$$\begin{aligned} P(S_2 = D) &= P(S_2 = D|D = 1)P(D = 1) + P(S_2 = D|D = 0)P(D = 0) \\ &= P(S_2 = D|D = 1)\theta + P(S_2 = D|D = 0)(1-\theta) \\ &= \theta\pi_1 + (1-\theta)\pi_0 \end{aligned}$$

Thus, a sufficent condition for $P(S_1 = D) \geq P(S_2 = D)$ is the following:

$$\pi_1^2\,(3 - 2\pi_1) \geq \pi_1 \quad \text{and} \quad \pi_0^2\,(3 - 2\pi_0) \geq \pi_0$$

Now,

$$\begin{aligned} \pi_1^2\,(3 - 2\pi_1) \geq \pi_1 \quad &\Longleftrightarrow \quad \pi_1\,(3 - 2\pi_1) \geq 1 \\ &\Longleftrightarrow \quad 2\pi_1^2 - 3\pi_1 + 1 \leq 0 \\ &\Longleftrightarrow \quad (2\pi_1 - 1)(\pi_1 - 1) \leq 0 \\ &\Longleftrightarrow \quad \frac{1}{2} \leq \pi_1 \leq 1 \end{aligned}$$

Similarly,

$$\pi_0^2\,(3-2\pi_0) \geq \pi_1 \quad \Longleftrightarrow \quad \frac{1}{2} \leq \pi_0 \leq 1$$

We may now conclude that a sufficient condition for $P(S_1 = D) \geq P(S_2 = 0)$ is

$$\frac{1}{2} \leq \pi_0\,,\, \pi_1 \leq 1$$

Comment: The above sufficient condition shows that as long as the probability of each doctor giving a correct diagnosis is at least $\frac{1}{2}$ (i.e. $\frac{1}{2} \leq \pi_0\,,\, \pi_1 \leq 1$), Strategy #1 will outperform Strategy #2, in the sense that the probability that Strategy #1 giving a correct diagnosis will exceed that of Strategy #2.

Exercise 1.21(b)

Let S_3 be the random variable defined by:

$$S_3 \;:=\; \begin{cases} 1, & \text{Strategy \#3 asserts that a given individual has IBD,} \\ 0, & \text{otherwise.} \end{cases}$$

Then,

$$\begin{aligned} P(S_3 = D | D = 1) &= P(Z \geq 3)\,, \quad \text{where } Z \sim \text{Binomial}(n=4, p=\pi_1) \\ &= \binom{4}{3}\pi_1^3(1-\pi_1)^1 + \binom{4}{4}\pi_1^4(1-\pi_1)^0 \\ &= 4\pi_1^3(1-\pi_1) + \pi_1^4 \\ &= \pi_1^3\,(4-3\pi_1) \end{aligned}$$

Similarly,

$$P(S_3 = D | D = 0) \;=\; \pi_0^3\,(4-3\pi_0)$$

Hence,

$$\begin{aligned} P(S_3 = D) &= P(S_3 = D, D = 1) + P(S_3 = D, D = 0) \\ &= P(S_3 = D | D = 1)P(D=1) + P(S_3 = D | D = 0)P(D=0) \\ &= \theta\pi_1^3\,(4-3\pi_1) + (1-\theta)\pi_0^3\,(4-3\pi_0) \end{aligned}$$

Now, observe that

$$\begin{aligned} P(S_1 = D) - P(S_3 = D) &= \left[\theta\pi_1^2\,(3-2\pi_1) + (1-\theta)\pi_0^2\,(3-2\pi_0)\right] - \left[\theta\pi_1^3\,(4-3\pi_1) + (1-\theta)\pi_0^3\,(4-3\pi_0)\right] \\ &= \theta\pi_1^2\left(3-2\pi_1-4\pi_1+3\pi_1^2\right) + (1-\theta)\pi_0^2\left(3-2\pi_0-4\pi_0+3\pi_0^2\right) \\ &= 3\theta\pi_1^2\left(\pi_1^2-2\pi_1+1\right) + 3(1-\theta)\pi_0^2\left(\pi_0^2-2\pi_0+1\right) \\ &= 3\theta\pi_1^2\,(\pi_1-1)^2 + 3(1-\theta)\pi_0^2\,(\pi_0-1)^2 \\ &\geq 0 \end{aligned}$$

Comment: This shows that Strategy #1 is always preferable over Strategy #3, regardless of the values of π_0 and π_1 (despite the latter involving more doctors). $\square$

Exercise 1.22

Let

- A be the event that an individual has Alzheimer's Disease.
- D be the event that an individual has diabetes.
- M be the event that an individual is male.

Note that

$$\begin{aligned}
\pi_1 \;:=\; P(A|D) \;=\; \frac{P(A,D)}{P(D)} &= \frac{P(A,D,M)+P(A,D,\overline{M})}{P(D)} \\
&= \frac{P(A,D,M)}{P(D,M)}\frac{P(D,M)}{P(D)} + \frac{P(A,D,\overline{M})}{P(D,\overline{M})}\frac{P(D,\overline{M})}{P(D)} \\
&= P(A|D,M)P(M|D) + P(A|D,\overline{M})P(\overline{M}|D) \\
&= \pi_{11}\cdot P(M|D) + \pi_{10}\cdot P(\overline{M}|D)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\pi_0 \;:=\; P(A|\overline{D}) \;=\; \frac{P(A,\overline{D})}{P(\overline{D})} &= \frac{P(A,\overline{D},M)+P(A,\overline{D},\overline{M})}{P(\overline{D})} \\
&= \frac{P(A,\overline{D},M)}{P(\overline{D},M)}\frac{P(\overline{D},M)}{P(\overline{D})} + \frac{P(A,\overline{D},\overline{M})}{P(\overline{D},\overline{M})}\frac{P(\overline{D},\overline{M})}{P(\overline{D})} \\
&= P(A|\overline{D},M)P(M|\overline{D}) + P(A|\overline{D},\overline{M})P(\overline{M}|\overline{D}) \\
&= \pi_{01}\cdot P(M|\overline{D}) + \pi_{00}\cdot P(\overline{M}|\overline{D})
\end{aligned}$$

We ASSUME

- $\pi_{00} \neq 0$, $\pi_{01} \neq 0$, and $\pi_0 \neq 0$.
- *homogeneity of risk ratio across gender groups*, i.e.

$$R_1 \;=\; R_0 \;=:\; R\,, \quad \text{where} \quad R_1 \;:=\; \frac{\pi_{11}}{\pi_{01}}\,, \quad R_0 \;:=\; \frac{\pi_{10}}{\pi_{00}}\,. \tag{1.3}$$

We seek to derive sufficient conditions for

$$R_c \;=\; R\,, \quad \text{where} \quad R_c \;:=\; \frac{\pi_1}{\pi_0}\,. \tag{1.4}$$

Now, it follows immediately from (1.3) and (1.4) that

$$\pi_{11} \;=\; R\cdot\pi_{01} \quad \text{and} \quad \pi_{10} \;=\; R\cdot\pi_{00}$$

Hence,

$$\pi_1 \;=\; R\cdot\left(\pi_{01}\cdot P(M|D) + \pi_{00}\cdot P(\overline{M}|D)\right)$$

which in turn implies:

$$\frac{\pi_1}{\pi_0} \;=\; R\cdot\left(\frac{\pi_{01}\,P(M|D)+\pi_{00}\,P(\overline{M}|D)}{\pi_{01}\,P(M|\overline{D})+\pi_{00}\,P(\overline{M}|\overline{D})}\right)$$

Thus, (1.4) will follow if the following holds:

$$\frac{\pi_{01}\,P(M|D)+\pi_{00}\,P(\overline{M}|D)}{\pi_{01}\,P(M|\overline{D})+\pi_{00}\,P(\overline{M}|\overline{D})} \;=\; 1 \tag{1.5}$$

Now, note:

$$\begin{aligned}
(1.5) \quad &\Longleftrightarrow \quad \pi_{01}\,P(M|D) + \pi_{00}\,P(\overline{M}|D) - \pi_{01}\,P(M|\overline{D}) - \pi_{00}\,P(\overline{M}|\overline{D}) = 0 \\
&\Longleftrightarrow \quad \pi_{01}\,\left[P(M|D) - P(M|\overline{D})\right] + \pi_{00}\,\left[P(\overline{M}|D) - P(\overline{M}|\overline{D})\right] = 0 \\
&\Longleftrightarrow \quad \pi_{01}\,\left[P(M|D) - P(M|\overline{D})\right] + \pi_{00}\,\left[1 - P(M|D) - 1 + P(M|\overline{D})\right] = 0 \\
&\Longleftrightarrow \quad [\,\pi_{01} - \pi_{00}\,]\cdot\left[\,P(M|D) - P(M|\overline{D})\,\right] = 0
\end{aligned}$$

Thus, two separate sufficient conditions for (1.4) are:

$$\pi_{01} \;=\; \pi_{00} \quad \text{and} \quad P(M|D) \;=\; P(M|\overline{D})$$

Furthermore,

$$\begin{aligned}
&\text{independence of } M \text{ and } D\text{, i.e. } P(M|D) = P(M) \\
\Longrightarrow \quad & \frac{P(M,D)}{P(D)} \;=\; P(M,D) + P(M,\overline{D}) \\
\Longrightarrow \quad & P(M,D) \;=\; P(M,D)P(D) + P(M,\overline{D})P(D) \\
\Longrightarrow \quad & P(M,D)\,[\,1 - P(D)\,] \;=\; P(M,\overline{D})P(D) \\
\Longrightarrow \quad & \frac{P(M,D)}{P(D)} \;=\; \frac{P(M,\overline{D})}{P(\overline{D})} \\
\Longrightarrow \quad & P(M|D) \;=\; P(M|\overline{D})
\end{aligned}$$

Therefore, we may now conclude that two separate sufficient conditions for (1.4) are:

- independence of M and D, i.e. $P(M|D) \;=\; P(M)$, and
- $\pi_{01} \;=\; \pi_{00}$, i.e. $P(A|\overline{D},M) = P(A|\overline{D},\overline{M})$.

$\square$

2 Chapter 2

Exercise 2.1(a)

Note that our "stopping criterion" here is that the sequence of selected individuals contains at least one individual with the rare blood disorder and at least one individual without the disorder. Thus, the following must hold: Let n be the length of a stopping sequence; the first $n-1$ individuals of the stopping sequence must all be of one type, while the n^{th} individual is of the other type.

Let 1 represent an individual with the rare blood disorder, while 0 an individual without the disorder. Then, since there are only four inddividuals with the disorder and three who do not, the following sequences are the only possible stopping sequences:

01, 001, 0001, 10, 110, 1110, 11110

The probabilities of the above admissible stopping sequences are tabulated below:

sequence s	$P(s)$	length of s	sequence s	$P(s)$	length of s
01	$\frac{3}{7}\cdot\frac{4}{6}$	2	10	$\frac{4}{7}\cdot\frac{3}{6}$	2
001	$\frac{3}{7}\cdot\frac{2}{6}\cdot\frac{4}{5}$	3	110	$\frac{4}{7}\cdot\frac{3}{6}\cdot\frac{3}{5}$	3
0001	$\frac{3}{7}\cdot\frac{2}{6}\cdot\frac{1}{5}\cdot\frac{4}{4}$	4	1110	$\frac{4}{7}\cdot\frac{3}{6}\cdot\frac{2}{5}\cdot\frac{3}{4}$	4
			11110	$\frac{4}{7}\cdot\frac{3}{6}\cdot\frac{2}{5}\cdot\frac{1}{4}\cdot\frac{3}{3}$	5

We thus see that, letting N denote (the random variable of) the stopping sequence:

$$\begin{aligned}
P(N=2) &= P(01)+P(10) &&= \frac{3\cdot 4+4\cdot 3}{7\cdot 6} &&= \frac{4}{7} \\
P(N=3) &= P(001)+P(110) &&= \frac{3\cdot 2\cdot 4+4\cdot 3\cdot 3}{7\cdot 6\cdot 5} &&= \frac{2}{7} \\
P(N=4) &= P(0001)+P(1110) &&= \frac{3\cdot 2\cdot 1\cdot 4+4\cdot 3\cdot 2\cdot 3}{7\cdot 6\cdot 5\cdot 4} &&= \frac{4}{35} \\
P(N=5) &= P(11110) &&= \frac{4\cdot 3\cdot 2\cdot 1\cdot 3}{7\cdot 6\cdot 5\cdot 4\cdot 3} &&= \frac{1}{35}
\end{aligned}$$

Hence,

$$E\,[\,N\,] \;=\; \sum_{n=2}^{5} n\cdot P(N=n) \;=\; 2\cdot\left(\frac{4}{7}\right)+3\cdot\left(\frac{2}{7}\right)+4\cdot\left(\frac{4}{35}\right)+5\cdot\left(\frac{1}{35}\right) \;=\; \frac{13}{5} \;=\; 2.6$$

Comment (the underlying probability space used in Exercise 2.1(a)):

Let

$$\Omega \;:=\; \{\,(\,x_i\,)_{i=1}^{\infty}\,|\,x_i\in\{0,1\}\,\} \;=\; \left\{\begin{array}{c}\text{all infinite sequences}\\ \text{of 0's and 1's}\end{array}\right\}$$

Note that each finite sequence $y = (y_1, y_2, \ldots, y_n)$ of zeros and ones can be regarded as a subset of Ω as follows:

$$y = (y_1, y_2, \ldots, y_n) \quad \longleftrightarrow \quad \{\, (\, x_i\,)_{i=1}^{\infty} \in \Omega \mid x_i = y_i,\ i = 1, 2, \ldots, n \,\}$$

Let Θ be the set of all finite sequences of zeros and ones. Then, by the preceding convention, we have that $\Theta \subset \text{PowerSet}\,(\Omega)$. Note the the underlying set of the probability space used in Exercise 2.1(a) is Ω, the probability measure on Ω is first defined subsets of Ω belonging to Θ, and then extend to the σ-algebra generated by Θ. Note also that, given any two members of Θ (as subsets of Ω), either they are disjoint or one is a subset of the other.

Exercise 2.1(b)

First, consider the concrete example that $M = 10$, $N = 100$, $k = 3$, and $X = 5$. Then,

$$\begin{aligned}
P(X = 5\,;\, 100, 10, 3) &= \left(\begin{array}{c} 5-1 \\ 3-1 \end{array}\right) \cdot \frac{10}{100} \cdot \frac{9}{99} \times \frac{90}{98} \cdot \frac{89}{97} \times \frac{8}{96} \;=\; \left(\begin{array}{c} 5-1 \\ 3-1 \end{array}\right) \cdot \frac{10}{100} \cdot \frac{10-1}{100-1} \cdot \frac{10-2}{100-2} \cdot \frac{90}{100-3} \cdot \frac{90-1}{100-4} \\
&= \left(\begin{array}{c} 5-1 \\ 3-1 \end{array}\right) \frac{10\,!}{(10-3)!} \cdot \frac{90\,!}{(90-(5-3))!} \cdot \frac{(100-5)!}{100\,!} \\
&= \left(\begin{array}{c} 5-1 \\ 3-1 \end{array}\right) \left(\begin{array}{c} 100-5 \\ 10-3 \end{array}\right) \Big/ \left(\begin{array}{c} 100 \\ 10 \end{array}\right)
\end{aligned}$$

In general, we therefore have:

$$P(X = x\,;\, N, M, k) \;=\; \frac{\left(\begin{array}{c} x-1 \\ k-1 \end{array}\right) \left(\begin{array}{c} N-x \\ M-k \end{array}\right)}{\left(\begin{array}{c} N \\ M \end{array}\right)}$$

$\square$

Exercise 2.23(a)

Since

$$U \;=\; \begin{cases} X, & \text{if } X \geq L, \\ g(L), & \text{if } X < L, \end{cases}$$

we have:

$$\begin{aligned}
E[\,U\,] &= \int_0^L g(L)\, f_X(x)\,\mathrm{d}x + \int_L^\infty x\, f_X(x)\,\mathrm{d}x \\
&= g(L)\cdot P(X \leq L) + P(X > L)\cdot \int_L^\infty x \cdot \frac{f_X(x)}{P(X > L)}\,\mathrm{d}x \\
&= g(L)\cdot(1-\pi) + \pi \cdot \int_L^\infty x \cdot f_{X\,|\,X>L}(x)\,\mathrm{d}x \\
&= g(L)\cdot(1-\pi) + \pi \cdot E[\,X\,|\,X > L\,]
\end{aligned}$$

Next, recall

$$V[\,U\,] \;:=\; E[\,(U - E[U])^2\,] \;=\; \cdots \;=\; E[\,U^2\,] - E[\,U\,]^2$$

So, we now compute $E[\,U^2\,]$:

$$\begin{aligned}
E[\,U^2\,] &= \int_0^L g(L)^2\, f_X(x)\,\mathrm{d}x \;+\; \int_L^\infty x^2 \cdot f_X(x)\,\mathrm{d}x \\
&= g(L)^2 \cdot P(X \leq L) \;+\; P(X > L)\cdot \int_L^\infty x^2 \cdot \frac{f_X(x)}{P(X > L)}\,\mathrm{d}x \\
&= g(L)^2 \cdot (1-\pi) \;+\; \pi \cdot \int_L^\infty x^2 \cdot f_{X\,|\,X>L}(x)\,\mathrm{d}x \\
&= g(L)^2 \cdot (1-\pi) \;+\; \pi \cdot E[\,X^2\,|\,X > L\,]
\end{aligned}$$

Hence,

$$\begin{aligned}
V[\,U\,] &= E[\,U^2\,] - E[\,U\,]^2 \\
&= \left\{ g(L)^2 \cdot (1-\pi) \;+\; \pi \cdot E[\,X^2\,|\,X > L\,] \right\} \;-\; \left\{ g(L)\cdot(1-\pi) + \pi \cdot E[\,X\,|\,X > L\,] \right\}^2 \\
&= g(L)^2 \cdot (1-\pi) \;+\; \pi \cdot E[\,X^2\,|\,X > L\,] \;-\; g(L)^2(1-\pi)^2 \;-\; 2\pi(1-\pi)g(L)E[\,X\,|\,X > L\,] \;-\; \pi^2 E[\,X\,|\,X > L\,]^2 \\
&= g(L)^2 \cdot (1-\pi)\cdot(1 - 1 + \pi) \;+\; \pi \cdot E[\,X^2\,|\,X > L\,] - \pi \cdot E[\,X\,|\,X > L\,]^2 \;+\; \pi \cdot E[\,X\,|\,X > L\,]^2 \\
&\quad - 2\pi(1-\pi)g(L)E[\,X\,|\,X > L\,] \;-\; \pi^2 E[\,X\,|\,X > L\,]^2 \\
&= \pi(1-\pi)g(L)^2 \;+\; \pi \cdot V[\,X\,|\,X > L\,] \;+\; \pi \cdot E[\,X\,|\,X > L\,]^2(1-\pi) \;-\; 2\pi(1-\pi)g(L)E[\,X\,|\,X > L\,] \\
&= \pi \cdot V[\,X\,|\,X > L\,] \;+\; \pi(1-\pi)\left\{ g(L)^2 \;-\; 2g(L)E[\,X\,|\,X > L\,] \;+\; E[\,X\,|\,X > L\,]^2 \right\} \\
&= \pi \cdot V[\,X\,|\,X > L\,] \;+\; \pi(1-\pi)\left(g(L) \;-\; E[\,X\,|\,X > L\,] \right)^2
\end{aligned}$$

Exercise 2.23(b)

Recall that

$$E[\,U\,] \;=\; g(L)(1-\pi) \;+\; \pi E[\,X\,|\,X > L\,]$$

Hence, setting $E[\,U\,] = E[\,X\,]$ yields:

$$E[\,X\,] \;=\; E[\,U\,] \;=\; g(L)(1-\pi) \;+\; \pi E[\,X\,|\,X > L\,]$$

which implies

$$\begin{aligned}
g(L) &= \frac{1}{1-\pi}\left(E[\,X\,] - \pi\,E[\,X \mid X > L\,]\right) \\
&= \frac{1}{P(X<L)}\left(\int_0^\infty x\,f_X(x)\,\mathrm{d}x \;-\; P(X>L)\int_L^\infty x\,\frac{f_X(x)}{P(X>L)}\,\mathrm{d}x\right) \\
&= \frac{1}{P(X<L)}\int_0^L x\,f_X(x)\,\mathrm{d}x \\
&= \int_0^L x\,\frac{f_X(x)}{P(X<L)}\,\mathrm{d}x \\
&= \int_0^L x\,f_{X\mid X<L}(x)\,\mathrm{d}x \\
&= E[\,X \mid X < L\,]
\end{aligned}$$

We conclude that the choice for $g(L)$ that will lead to $E[\,U\,] = E[\,X\,]$ is:

$$g(L) \;=\; E[\,X \mid X < L\,]$$

$$\begin{aligned}
g(L) &= \frac{\int_0^L x\,f_X(x)\,\mathrm{d}x}{P(X<L)} = \frac{\int_0^L x\,f_X(x)\,\mathrm{d}x}{\int_0^L f_X(x)\,\mathrm{d}x} = \frac{\int_0^L x\,e^{-x}\,\mathrm{d}x}{\int_0^L e^{-x}\,\mathrm{d}x} = \frac{\left[\,-e^{-x}(x+1)\,\right]_0^L}{\left[\,-e^{-x}\,\right]_0^L} \\
&= \frac{1-e^{-L}(L+1)}{1-e^{-L}}
\end{aligned}$$

For $L = 0.05$, we have

$$g(0.05) \;=\; \frac{1-e^{-0.05}(0.05+1)}{1-e^{-0.05}} \;\approx\; 0.02479168$$

Exercise 2.26(a)

Recall that:

$$h(t) \;:=\; \lim_{h\to 0}\frac{P(t\le T\le t+h\,|\,t\le T)}{h},\qquad F(t) \;=\; P(T<t),\qquad f(t) \;=\; F'(t),\qquad S(t)=1-F(t)$$

We want to prove:

$$h(t) \;=\; \frac{f(t)}{S(t)}.$$

Proof

$$\begin{aligned} P(t\le T\le t+h\,|\,t\le T) &= \frac{P(t\le T\le t+h\,\cap\, t\le T)}{P(t\le T)} \;=\; \frac{P(t\le T\le t+h)}{P(t\le T)} \;=\; \frac{P(T\le t+h)-P(T\le t)}{P(t\le T)} \\ &= \frac{F(t+h)-F(t)}{1-F(t)} \end{aligned}$$

Hence,

$$\frac{P(t\le T\le t+h\,|\,t\le T)}{h} \;=\; \frac{1}{1-F(t)}\cdot\frac{F(t+h)-F(t)}{h}$$

Hence,

$$h(t) \;=\; \lim_{h\to 0}\frac{P(t\le T\le t+h\,|\,t\le T)}{h} \;=\; \frac{1}{1-F(t)}\cdot\lim_{h\to 0}\frac{F(t+h)-F(t)}{h} \;=\; \frac{1}{1-F(t)}\cdot F'(t) \;=\; \frac{f(t)}{S(t)}$$

$\square$

Exercise 2.26(b)

Recall that $S(t)=1-F(t)$, hence $S'(t)=-F'(t)=-f(t)$. Consequently,

$$h(t) \;=\; \frac{f(t)}{S(t)} \;=\; -\frac{S'(t)}{S(t)} \;=\; -\,\frac{\mathrm{d}}{\mathrm{d}t}\Big(\log S(t)\Big)$$

Thus,

$$\int_0^t h(\tau)\,\mathrm{d}\tau \;=\; -\int_0^t \frac{\mathrm{d}}{\mathrm{d}\tau}\Big(\log S(\tau)\Big)\,\mathrm{d}\tau \;=\; -\,[\,\log S(\tau)\,]_0^t \;=\; -\log S(t)+\log S(0)$$

Now, $S(0)=1-F(0)=1-P(T<0)=1-0=1$; hence, $\log S(0)=\log(1)=0$. We thus have:

$$\log\left(S(t)\right) \;=\; -\int_0^t h(\tau)\,\mathrm{d}\tau \quad\Longrightarrow\quad S(t) \;=\; \exp\left\{-\int_0^t h(\tau)\,\mathrm{d}\tau\right\} \;=\; \exp\left\{-H(t)\right\}$$

$\square$

Exercise 2.26(c)

Assuming $\lim_{t\to\infty} t\cdot S(t) \;=\; \lim_{t\to\infty} t\,(1-F(t)) \;=\; 0$, we want to prove that

$$E\,[\,T\,] \;=\; \int_0^\infty S(t)\,\mathrm{d}t$$

PROOF

$$\begin{aligned}
E[T] &= \int_0^\infty t\cdot f(t)\,\mathrm{d}t \;=\; -\int_0^\infty t\cdot S'(t)\,\mathrm{d}t \\
&= -[\,t\cdot S(t)\,]_0^\infty + \int_0^\infty S(t)\,\mathrm{d}t\,, \quad \text{by integration by parts: } \int u\,\mathrm{d}v = uv - \int v\,\mathrm{d}u,\;\; u=t,\;\; \mathrm{d}v = S'(t)\,\mathrm{d}t \\
&= \int_0^\infty S(t)\,\mathrm{d}t
\end{aligned}$$

□

Exercise 2.26(d)

We want to prove:

$$E[\,H(X)\,] \;=\; F_T(\,c\,),$$

where $X = \min\{\,T,\,c\,\}$, $H(t) = \int_0^t h(\tau)\,\mathrm{d}\tau$.

PROOF We first derive the cumulative probability function of the random variable $X := \min\{\,T,\,c\,\}$. First, note that range$(X) = (\,0,\,c\,]$.

$$\begin{aligned}
P(\,X\le x\,) &= P(\,\min\{T,\,c\}\le x\,) \;=\; P(\,T\le x \text{ or } c\le x\,) \\
&= P(\,T\le x\,) + P(\,c\le x\,) - P(\,T\le x \text{ and } c\le x\,) \\
&= \begin{cases} P(\,T\le x\,) + 1 - P(\,T\le x\,), & \text{fox } c\le x \\ P(\,T\le x\,) + 0 - 0, & \text{fox } c > x \end{cases} \\
&= \begin{cases} 1, & \text{for } c\le x \\ P(\,T\le x\,), & \text{for } c > x \end{cases} \\
&= \begin{cases} P(\,T\le x\,), & \text{for } x < c \\ 1, & \text{for } x = c \end{cases}
\end{aligned}$$

Hence,

$$\frac{\mathrm{d}}{\mathrm{d}t}P(X\le x) \;=\; \begin{cases} f_T(x), & \text{for } 0 < x < c \\ \text{undefined}, & \text{for } x = c \end{cases}$$

We thus see that the probability density μ_X of the random variable X is a probability measure on range$(X) = (\,0,\,c\,]$. Over the interval $x\in(\,0,\,c\,)$, μ_X is representable by a probability density function $f_X(x) = f_T(x)$ for $x\in(\,0,\,c\,)$. When restricted to the single point $x = c$, μ_X is the point mass measure with $\mu_X(\,c\,) = 1 - F_T(\,c\,)$. In other words, the probability measure μ_X of X is given by:

$$\begin{aligned}
\mu_X &= \begin{cases} f_T(x), & \text{for } 0 < x < c \\ P(X = c), & \text{for } x = c \end{cases} \\
&= \begin{cases} f_T(x), & \text{for } 0 < x < c \\ P(T \ge c), & \text{for } x = c \end{cases} \\
&= \begin{cases} f_T(x), & \text{for } 0 < x < c \\ 1 - F_T(\,c\,), & \text{for } x = c \end{cases}
\end{aligned}$$

Thus,

$$E[\,H(X)\,] \;=\; \int_{[\,0,\,c\,]} H(x)\,\mathrm{d}\mu_X \;=\; \int_0^c H(x)\cdot f_T(x)\,\mathrm{d}x \;+\; H(c)\,\mu_X(\,c\,)$$

Now, note that

$$H(c)\cdot\mu_X(\,c\,) \;=\; -\log S(c)\cdot(\,1 - F_T(\,c\,)\,)$$

Next, recall from Exercise 2.26(b) that $H(t) = -\log S(t)$, $f_T(x) = -S'(x)$.

$$\begin{aligned}
\int_0^c H(x)\cdot f_T(x)\,\mathrm{d}x &= \int_0^c (-\log S(x))\cdot(-S'(x))\;\mathrm{d}x \\
&= \int_0^c \frac{\mathrm{d}}{\mathrm{d}\,S(x)}\left(S(x)\log S(x) - S(x)\right)\cdot(S'(x))\;\mathrm{d}x \\
&= \int_0^c \frac{\mathrm{d}}{\mathrm{d}x}\left(S(x)\log S(x) - S(x)\right)\;\mathrm{d}x \\
&= [\,S(x)\log S(x) - S(x)\,]_0^c \\
&= S(c)\log S(c) - S(c) - S(0)\log S(0) + S(0) \\
&= S(c)\log S(c) - S(c) - 1\cdot\log(S(0)) + 1 \\
&= S(c)\log S(c) - S(c) + 1 \\
&= S(c)\log S(c) + F_T(\,c\,)
\end{aligned}$$

Thus, we now have

$$\begin{aligned}
E[\,H(X)\,] &= \int_{[\,0,\,c\,]} H(x)\,\mathrm{d}\mu_X \;=\; \int_0^c H(x)\cdot f_T(x)\,\mathrm{d}x \;+\; H(c)\cdot\mu_X(\,c\,) \\
&= S(c)\log S(c) + F_T(\,c\,) - \log S(c)\cdot(\,1 - F_T(\,c\,)\,) \\
&= S(c)\log S(c) + F_T(\,c\,) - \log S(c) + \log(S(c))\cdot F_T(\,c\,) \\
&= -\log S(c)\cdot(1 - S(c)) + F_T(\,c\,)\cdot(1 + \log S(c)) \\
&= -\log S(c)\cdot F_T(\,c\,) + F_T(\,c\,)\cdot(1 + \log S(c)) \\
&= F_T(\,c\,)\cdot(-\log S(c) + 1 + \log S(c)) \\
&= F_T(\,c\,)
\end{aligned}$$

□

Exercise 2.27(a)

Let $Y := \min\{\, X,\, N \,\}$ be the number of units actually sold. Note that $\text{range}(Y) = (\,0,\, N\,]$. Let Z be the profit. Then,

$$Z \;=\; GY - L(N-Y) \;=\; (G+L)Y - LN$$

Then,

$$E[\,Z\,] = (G+L)\,E[\,Y\,] - LN$$

We next compute $E[\,Y\,]$. To this end, we first need to compute the probability density measure of Y. Now,

$$\begin{aligned}
P(Y \leq y) &= P(\,\min\{\,X,\,N\,\} \leq y\,) \;=\; P(X \leq y \ \text{ or } \ N \leq y) \\
&= P(X \leq y) + P(N \leq y) - P(X \leq y \ \text{ and } \ N \leq y) \\
&= \begin{cases} P(X \leq y) + 1 - P(X \leq y), & \text{for } \ N \leq y \\ P(X \leq y) + 0 - 0, & \text{for } \ N > y \end{cases} \\
&= \begin{cases} P(X \leq y), & \text{for } \ 0 < y < N \\ \quad 1, & \text{for } \ y = N \end{cases}
\end{aligned}$$

$$\frac{\mathrm{d}}{\mathrm{d}y} P(Y \leq y) \;=\; \begin{cases} f_X(y), & \text{for } \ 0 < y < N \\ \text{undefined}, & \text{for } \ y = N \end{cases}$$

Thus, we now see that the probability measure μ_Y of Y is representable by the probability density function $f_X(y)$ for $y \in (\,0,\, N\,)$, and it restricts to the point mass measure at $y = N$ with $\mu_Y(N) = 1 - F_X(N)$. In other words,

$$\mu_Y \;=\; \begin{cases} f_X(y), & \text{for } \ 0 < y < N \\ 1 - F_X(N), & \text{for } \ y = N \end{cases}$$

Thus,

$$E[\,Y\,] \;=\; \int_{(0,\,N]} y \,\mathrm{d}\mu_Y \;=\; \int_0^N y\, f_X(y)\,\mathrm{d}y + N \cdot \mu_X(N) \;=\; \int_0^N y\, f_X(y)\,\mathrm{d}y + N \cdot (1 - F_X(N))$$

And,

$$\begin{aligned}
E[\,Z\,] &= E[\,(G+L)Y - LN\,] \;=\; (G+L)E[\,Y\,] - LN && (2.1) \\
&= (G+L)\left\{ \int_0^N y\, f_X(y)\,\mathrm{d}y + N \cdot (1 - F_X(N)) \right\} - LN && (2.2)
\end{aligned}$$

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}N} E[\,Z\,] &= (G+L)\,\{\,N\, f_X(N) + (1 - F_X(N)) - N\, f_X(N)\,\} - L \\
&= (G+L)\,(\,1 - F_X(N)\,) - L \\
&= G + L - (G+L)\cdot F_X(N) - L \\
&= G - (G+L)\cdot F_X(N)
\end{aligned}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}N^2} E[\,Z\,] \;=\; -(G+L)\cdot f_X(N) \;<\; 0$$

Noting that

$$\frac{\mathrm{d}}{\mathrm{d}N} E[\,Z\,] \;=\; 0 \quad\Longrightarrow\quad F_X(N) \;=\; \frac{G}{G+L}$$

we see that $N_0 := F_X^{-1}\left(\dfrac{G}{G+L}\right)$ is a local maximum of $E[\,Z\,]$ as a function of N (given by (2.2)).

Exercise 2.27(b)

Recall:

$$\int ax\exp\left(-bx^2\right)\,\mathrm{d}x \;=\; -\frac{a\exp\left(-bx^2\right)}{2b}$$

Hence, if

$$f_X(x) \;=\; \left(2\times 10^{-10}\right)\, x\, \exp\left\{-\left(10^{-10}\right)\, x^2\right\}, \quad x\in(\,0,\,\infty\,)$$

then

$$\begin{aligned}
F_X(x) &= \int_0^x f_X(\xi)\,\mathrm{d}\xi \;=\; \int_0^x \left(2\times 10^{-10}\right)\,\xi\,\exp\left\{-\left(10^{-10}\right)\,\xi^2\right\}\,\mathrm{d}\xi \\
&= \left[-\frac{2\times 10^{-10}\times\exp\left(-\left(10^{-10}\right)\xi^2\right)}{2\times 10^{-10}}\right]_0^x \;=\; -\left[\exp\left(-\left(10^{-10}\right)\xi^2\right)\right]_0^x \\
&= 1-\exp\left(-x^2/10^{10}\right)
\end{aligned}$$

Now, seek x_0 such that

$$\begin{aligned}
\frac{G}{G+L} &= F_X(x_0) \;=\; 1-\exp\left(-x_0^2/10^{10}\right) \\
\Longrightarrow\quad \exp\left(-x_0^2/10^{10}\right) &= 1-\frac{G}{G+L} \;=\; \frac{L}{G+L} \\
\Longrightarrow\quad -\frac{x_0^2}{10^{10}} &= \log\left(\frac{L}{G+L}\right) \\
\Longrightarrow\quad x_0^2 &= -10^{10}\cdot\log\left(\frac{L}{G+L}\right) \\
\Longrightarrow\quad x_0 &= 10^5\left(\log\left(\frac{G+L}{L}\right)\right)^{1/2}
\end{aligned}$$

With $G=4$, and $L=1$, we have:

$$x_0 \;=\; 10^5\left(\log\left(\frac{4+1}{1}\right)\right)^{1/2} \;=\; 10^5\left(\log 5\right)^{1/2} \;\approx\; 10^5\times 1.268636 \;\approx\; 126863.6$$

The number of units to produce that maximizes the expected value $E[\,Z\,]$ of profit Z is, roughly:

$$N = 126,864$$

$\square$

Exercise 2.28(a)

First, note that

$$E[\,P_2\,] \;=\; \alpha P_1\pi + \beta P_1(1-\pi) \;=\; P_1\,(\alpha\pi + \beta(1-\pi))$$

Similarly, note that

$$E[\,P_3\,|\,P_2] \;=\; P_2\,(\alpha\pi + \beta(1-\pi))$$

Hence,

$$E[\,P_3\,] \;=\; E[\,E[\,P_3\,|\,P_2\,]\,] \;=\; E[\,P_2\,(\alpha\pi + \beta(1-\pi))\,] \;=\; (\alpha\pi + \beta(1-\pi))\,E[\,P_2\,] \;=\; P_1\cdot(\alpha\pi + \beta(1-\pi))^2$$

We now see that an induction argument yields:

$$E[\,P_k\,] \;=\; P_1\cdot(\alpha\pi + \beta(1-\pi))^{k-1}\,, \quad k = 2,3,4,\ldots$$

Exercise 2.28(b)

We seek the smallest value of α (treating all other quantities fixed) such that:

$$P_1\cdot(\alpha\pi + \beta(1-\pi))^{k-1} \;=\; E[\,P_k\,] \;\geq\; P^*$$

which in turn implies:

$$\begin{aligned} & \alpha\pi + \beta(1-\pi) \;\geq\; \left(\frac{P^*}{P_1}\right)^{\frac{1}{k-1}} \\ \Longrightarrow \qquad & \alpha \;\geq\; \frac{1}{\pi}\left\{\left(\frac{P^*}{P_1}\right)^{1/(k-1)} - \beta(1-\pi)\right\} \end{aligned}$$

Thus, the smallest value α^* of α we are seeking is:

$$\alpha^* \;=\; \frac{1}{\pi}\left\{\left(\frac{P^*}{P_1}\right)^{1/(k-1)} - \beta(1-\pi)\right\}$$

And, it is clear that

$$\lim_{k\to\infty}\alpha^* \;=\; \frac{1}{\pi}\,(1-\beta(1-\pi))$$

For $\pi = 0.90$ and $\beta = 1.05$,

$$\lim_{k\to\infty}\alpha^* \;=\; \frac{1}{0.90}\,(1 - 1.05(1-0.90)) \;\approx\; 0.9944444$$

3 Chapter 3

Exercise 3.12

Recall that

$$\mathrm{Corr}\left(\overline{Y}_a, \overline{Y}_g\right) = \frac{\mathrm{Cov}\left(\overline{Y}_a, \overline{Y}_g\right)}{\sqrt{V\left[\,\overline{Y}_a\,\right] \cdot V\left[\,\overline{Y}_g\,\right]}}$$

We first compute $V\left[\,\overline{Y}_a\,\right]$, followed by $V\left[\,\overline{Y}_g\,\right]$, and then $\mathrm{Cov}\left(\overline{Y}_a, \overline{Y}_g\right)$. Now, if $X \sim N(\mu, \sigma^2)$ and $Y = \exp(X)$, then

$$Y \sim \mathrm{LogNormal}(\mu, \sigma^2)$$

and we then have

$$E[\,Y\,] = e^{\mu + \frac{\sigma^2}{2}} \quad \text{and} \quad V[\,Y\,] = E[\,Y\,]^2 \cdot \left(e^{\sigma^2} - 1\right) = \left(e^{\mu+\frac{\sigma^2}{2}}\right)^2 \cdot \left(e^{\sigma^2} - 1\right) = \left(e^{2\mu+\sigma^2}\right) \cdot \left(e^{\sigma^2} - 1\right)$$

Hence,

$$V\left[\,\overline{Y}_a\,\right] = V\left[\frac{1}{n}\sum_{i=1}^{n} Y_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} V\left[\,Y_i\,\right] = \frac{1}{n} V\left[\,Y\,\right] = \frac{1}{n}\left(e^{2\mu+\sigma^2}\right) \cdot \left(e^{\sigma^2} - 1\right)$$

Next, note that, since $\overline{Y}_g = \left(\prod_{i=1}^{n} Y_i\right)^{1/n}$, we have

$$\log\left(\overline{Y}_g\right) = \log\left[\left(\prod_{i=1}^{n} Y_i\right)^{1/n}\right] = \frac{1}{n}\sum_{i=1}^{n} \log\left(Y_i\right) = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}_n \sim N\left(\mu\,,\,\frac{\sigma^2}{n}\right)$$

Thus,

$$\overline{Y}_g = e^{\overline{X}_n} = \mathrm{LogNormal}\left(\mu\,,\,\frac{\sigma^2}{n}\right)$$

Hence,

$$E\left[\,\overline{Y}_g\,\right] = \exp\left\{\mu + \frac{\sigma^2}{2n}\right\} \quad \text{and} \quad V\left[\,\overline{Y}_g\,\right] = \left[e^{\mu+\frac{\sigma^2}{2n}}\right]^2 \cdot \left(e^{\frac{\sigma^2}{n}} - 1\right) = \left[e^{2\mu+\frac{\sigma^2}{n}}\right] \cdot \left(e^{\frac{\sigma^2}{n}} - 1\right)$$

Next,

$$\mathrm{Cov}\left(\overline{Y_a}\,,\,\overline{Y}_g\right) = \mathrm{Cov}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\,,\,\overline{Y}_g\right) = \frac{1}{n}\sum_{i=1}^{n}\mathrm{Cov}\left(Y_i\,,\,\overline{Y}_g\right) = \frac{1}{n}\sum_{i=1}^{n}\left\{E\left[\,Y_i \cdot \overline{Y}_g\,\right] - E[\,Y_i\,] \cdot E\left[\overline{Y}_g\,\right]\right\}$$

Now,

$$\begin{aligned} E\left[\,Y_i \cdot \overline{Y}_g\,\right] &= E\left[Y_i \cdot \left(\prod_{k=1}^{n} Y_k\right)^{1/n}\right] = E\left[Y_i \cdot \prod_{k=1}^{n} Y_k^{1/n}\right] = E\left[Y_i^{1+\frac{1}{n}} \cdot \prod_{\substack{k=1\\k\neq i}}^{n} Y_k^{1/n}\right] \\ &= E\left[\,Y_i^{1+\frac{1}{n}}\,\right] \cdot E\left[\prod_{\substack{k=1\\k\neq i}}^{n} Y_k^{1/n}\right] \quad \text{(by independence of } Y_1, \ldots, Y_n\text{)} \end{aligned}$$

Now, note that

$$\log\left(Y^{1+\frac{1}{n}}\right) = \left(1 + \frac{1}{n}\right)\log(Y) = \left(1 + \frac{1}{n}\right) X \sim N\left(\left(1 + \frac{1}{n}\right)\mu\,,\,\left(1 + \frac{1}{n}\right)^2 \sigma^2\right)$$

Hence,

$$Y^{1+\frac{1}{n}} \sim \text{LogNormal}\left(\left(1+\frac{1}{n}\right)\mu\,,\ \left(1+\frac{1}{n}\right)^2\sigma^2 \right)$$

and

$$E\left[\,Y^{1+\frac{1}{n}}\,\right] \;=\; \exp\left\{ \left(1+\frac{1}{n}\right)\mu \;+\; \left(1+\frac{1}{n}\right)^2\frac{\sigma^2}{2} \right\}$$

On the other hand,

$$\log\left[\prod_{\substack{k=1\\k\neq i}}^{n} Y_k^{1/n}\right] \;=\; \log\left[\left(\prod_{\substack{k=1\\k\neq i}}^{n} Y_k\right)^{1/n}\right] \;=\; \frac{1}{n}\sum_{\substack{k=1\\k\neq i}}^{n}\log Y_k \;=\; \frac{1}{n}\sum_{\substack{k=1\\k\neq i}}^{n} X_k \;\sim\; N\left(\left(\frac{n-1}{n}\right)\mu\,,\ \left(\frac{n-1}{n^2}\right)\sigma^2 \right)$$

Thus,

$$\prod_{\substack{k=1\\k\neq i}}^{n} Y_k^{1/n} \;\sim\; \text{LogNormal}\left(\left(\frac{n-1}{n}\right)\mu\,,\ \left(\frac{n-1}{n^2}\right)\sigma^2 \right)$$

Hence,

$$E\left[\prod_{\substack{k=1\\k\neq i}}^{n} Y_k^{1/n}\right] \;=\; \exp\left\{ \left(\frac{n-1}{n}\right)\mu \;+\; \left(\frac{n-1}{n^2}\right)\frac{\sigma^2}{2} \right\}$$

Hence,

$$\begin{aligned}
\text{Cov}\left(Y_i\,,\,\overline{Y}_g\right) \;&=\; E\left[\,Y_i\cdot\overline{Y}_g\,\right] \;-\; E[\,Y_i\,]\cdot E\left[\overline{Y}_g\,\right] \;=\; E\left[\,Y_i^{1+\frac{1}{n}}\,\right]\cdot E\left[\prod_{\substack{k=1\\k\neq i}}^{n} Y_k^{1/n}\right] \;-\; E[\,Y_i\,]\cdot E\left[\overline{Y}_g\,\right] \\
&=\; \exp\left\{ \left(1+\frac{1}{n}\right)\mu \;+\; \left(1+\frac{1}{n}\right)^2\frac{\sigma^2}{2} \right\}\cdot\exp\left\{ \left(\frac{n-1}{n}\right)\mu + \left(\frac{n-1}{n^2}\right)\frac{\sigma^2}{2} \right\} \;-\; \exp\left\{\mu+\frac{\sigma^2}{2}\right\}\exp\left\{\mu+\frac{\sigma^2}{2n}\right\} \\
&=\; \exp\left\{ \left(1+\frac{1}{n}\right)\mu \;+\; \left(1+\frac{2}{n}+\frac{1}{n^2}\right)\frac{\sigma^2}{2} \;+\; \left(1-\frac{1}{n}\right)\mu + \left(\frac{1}{n}-\frac{1}{n^2}\right)\frac{\sigma^2}{2} \right\} \;-\; \exp\left\{2\mu+\frac{\sigma^2}{2}\left(1+\frac{1}{n}\right)\right\} \\
&=\; \exp\left\{2\mu \;+\; \frac{\sigma^2}{2}\left(1+\frac{3}{n}\right)\right\} \;-\; \exp\left\{2\mu+\frac{\sigma^2}{2}\left(1+\frac{1}{n}\right)\right\} \\
&=\; \exp\left\{2\mu \;+\; \frac{\sigma^2}{2}\left(1+\frac{1}{n}\right)\right\}\cdot\left\{\exp\left(\frac{\sigma^2}{n}\right)-1\right\}
\end{aligned}$$

So,

$$\begin{aligned}
\text{Cov}\left(\overline{Y_a}\,,\,\overline{Y}_g\right) \;&=\; \text{Cov}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\,,\,\overline{Y}_g\right) \;=\; \frac{1}{n}\sum_{i=1}^{n}\text{Cov}\left(Y_i\,,\,\overline{Y}_g\right) \;=\; \frac{n}{n}\exp\left\{2\mu \;+\; \frac{\sigma^2}{2}\left(1+\frac{1}{n}\right)\right\}\cdot\left\{\exp\left(\frac{\sigma^2}{n}\right)-1\right\} \\
&=\; \exp\left\{2\mu \;+\; \frac{\sigma^2}{2}\left(1+\frac{1}{n}\right)\right\}\cdot\left\{\exp\left(\frac{\sigma^2}{n}\right)-1\right\}
\end{aligned}$$

Lastly,

$$
\begin{aligned}
\mathrm{Corr}\left(\,\overline{Y_a}\,,\,\overline{Y}_g\,\right) &= \frac{\mathrm{Cov}\left(\,\overline{Y_a}\,,\,\overline{Y}_g\,\right)}{\sqrt{V\left[\,\overline{Y}_a\,\right]\cdot V\left[\,\overline{Y}_g\,\right]}} \\
&= \frac{\exp\left\{2\mu + \dfrac{\sigma^2}{2}\left(1+\frac{1}{n}\right)\right\}\cdot\left\{\exp\left(\dfrac{\sigma^2}{n}\right)-1\right\}}{\sqrt{\dfrac{1}{n}\left(e^{2\mu+\sigma^2}\right)\cdot\left(e^{\sigma^2}-1\right)\cdot\left(e^{2\mu+\frac{\sigma^2}{n}}\right)\cdot\left(e^{\frac{\sigma^2}{n}}-1\right)}} \\
&= \frac{\sqrt{n}\cdot\exp\left\{2\mu + \dfrac{\sigma^2}{2}\left(1+\frac{1}{n}\right)\right\}\cdot\left\{\exp\left(\dfrac{\sigma^2}{n}\right)-1\right\}}{\left(e^{\mu+\frac{\sigma^2}{2}}\right)\cdot\sqrt{e^{\sigma^2}-1}\cdot\left(e^{\mu+\frac{\sigma^2}{2n}}\right)\cdot\sqrt{e^{\frac{\sigma^2}{n}}-1}} \\
&= \frac{\sqrt{n}\cdot\left\{\exp\left(\dfrac{\sigma^2}{n}\right)-1\right\}}{\sqrt{e^{\sigma^2}-1}\cdot\sqrt{e^{\frac{\sigma^2}{n}}-1}} \\
&= \sqrt{\frac{n\cdot\left(e^{\frac{\sigma^2}{n}}-1\right)}{e^{\sigma^2}-1}}
\end{aligned}
$$

Exercise 3.14

Solution 1 (without using moment generating functions)

First, note that for $k = 0, 1, 2, \ldots$, we have:

$$
\begin{aligned}
P(\,Y_p = k\,) &= \sum_{y=0}^{\infty} P(\,Y_p = k \,|\, Y = y\,) \cdot P(Y = y) \;=\; \sum_{y=k}^{\infty} P(\,Y_p = k \,|\, Y = y\,) \cdot P(Y = y) \\
&= \sum_{y=k}^{\infty} \binom{y}{k} \pi^k (1-\pi)^{y-k} \cdot e^{-\lambda L} \frac{(\lambda L)^y}{y!} \\
&= e^{-\lambda L} \left(\frac{\pi}{1-\pi}\right)^k \cdot \sum_{y=k}^{\infty} \frac{y!}{k!\,(y-k)!} \cdot \frac{((1-\pi)\lambda L)^y}{y!} \\
&= e^{-\lambda L} \left(\frac{\pi}{1-\pi}\right)^k \frac{1}{k!} \cdot \sum_{y=k}^{\infty} \frac{((1-\pi)\lambda L)^y}{(y-k)!} \\
&= e^{-\lambda L} \left(\frac{\pi}{1-\pi}\right)^k \frac{1}{k!} \cdot \sum_{m=0}^{\infty} \frac{((1-\pi)\lambda L)^{m+k}}{m!} \\
&= e^{-\lambda L} \left(\frac{\pi}{1-\pi}\right)^k \frac{1}{k!} \cdot ((1-\pi)\lambda L)^k \cdot \sum_{m=0}^{\infty} \frac{((1-\pi)\lambda L)^m}{m!} \\
&= e^{-\lambda L} \left(\frac{\pi}{1-\pi}\right)^k \frac{1}{k!} \cdot ((1-\pi)\lambda L)^k \cdot e^{(1-\pi)\lambda L} \\
&= e^{-\lambda L + \lambda L - \pi\lambda L} \frac{(\pi\lambda L)^k}{k!} \\
&= e^{-\pi\lambda L} \frac{(\pi\lambda L)^k}{k!}
\end{aligned}
$$

Thus, we see that $Y_p \sim \text{Poisson}(\pi\lambda L)$.

4 Chapter 4

Exercise 4.8(a)

The likelihood function is:

$$L(\theta \,|\, y_1,\ldots,y_n) \;=\; \prod_{i=1}^{n} \frac{1}{\alpha}\exp\Big(-\,\frac{y_i}{\alpha}\Big) \;=\; \prod_{i=1}^{n} \theta^{-1/2}\exp\Big(-\,\theta^{-1/2}\cdot y_i\Big) \;=\; \theta^{-n/2}\exp\left(-\,\theta^{-1/2}\sum_{i=1}^{n} y_i\right)$$

Hence, the log-likelihood function is:

$$l(\theta \,|\, y_1,\ldots,y_n) \;=\; \log L(\theta \,|\, y_1,\ldots,y_n) \;=\; -\frac{n}{2}\log\theta - \theta^{-1/2}\sum_{i=1}^{n} y_i$$

Thus,

$$l'(\theta \,|\, y_1,\ldots,y_n) \;=\; -\frac{n}{2}\theta^{-1} + \frac{1}{2}\theta^{-3/2}\sum_{i=1}^{n} y_i \;=\; -\frac{\theta^{-1}}{2}\left(n - \theta^{-1/2}\sum_{i=1}^{n} y_i\right)$$

Thus,

$$l'(\theta \,|\, y_1,\ldots,y_n) = 0 \quad\Longrightarrow\quad \theta = \left(\frac{1}{n}\sum_{i=1}^{n} y_i\right)^2$$

We may now conclude that:

$$\widehat{\theta}_{\mathrm{MLE}} \;=\; \left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)^2$$

Next, we compute an asymptotic estimate for $\mathrm{Var}\Big(\widehat{\theta}_{\mathrm{MLE}}\Big)$. To this end, we recall from general MLE theory that $\mathrm{Var}\Big(\widehat{\theta}_{\mathrm{MLE}}\Big)$ asymptotically attains the Cramér-Rao Lower Bound (CRLB) for any unbiased estimator of θ. Now,

$$\mathrm{CRLB}(\theta) \;=\; \frac{1}{\mathrm{E}\left[l'(\theta \,|\, y_1,\ldots,y_n)^2\right]} \;=\; \frac{-1}{\mathrm{E}\left[l''(\theta \,|\, y_1,\ldots,y_n)\right]}$$

Now,

$$l''(\theta \,|\, y_1,\ldots,y_n) \;=\; \frac{n}{2}\theta^{-2} - \frac{3}{4}\theta^{-5/2}\sum_{i=1}^{n} y_i$$

Hence,

$$E\left[l''(\theta \,|\, y_1,\ldots,y_n)\right] \;=\; E\left[\frac{n}{2}\theta^{-2} - \frac{3}{4}\theta^{-5/2}\sum_{i=1}^{n} y_i\right] \;=\; \frac{n}{2}\theta^{-2} - \frac{3}{4}\theta^{-5/2}\sum_{i=1}^{n} E\left[y_i\right] \;=\; \frac{n}{2}\theta^{-2} - \frac{3}{4}\theta^{-5/2}\cdot n\,\theta^{1/2} \;=\; -\,\frac{n}{4\theta^2}$$

Thus,

$$\mathrm{CRLB}(\theta) \;=\; \frac{-1}{\mathrm{E}\left[l''(\theta \,|\, y_1,\ldots,y_n)\right]} \;=\; \frac{4\theta^2}{n}$$

Since $\widehat{\theta}_{\mathrm{MLE}}$ is asymptotically Gaussian, with an asymptotic 95% confidence interval for $\widehat{\theta}_{\mathrm{MLE}}$ is thus:

$$\widehat{\theta}_{\mathrm{MLE}} \pm 1.96\sqrt{\mathrm{Var}\Big[\widehat{\theta}_{\mathrm{MLE}}\Big]} \;=\; \widehat{\theta}_{\mathrm{MLE}} \pm 1.96\cdot\sqrt{\frac{4\,\widehat{\theta}^2_{\mathrm{MLE}}}{n}} \;=\; \left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)^2 \pm \frac{3.92}{\sqrt{n}}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)^2 \;=\; \left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)^2\left[\,1 \pm \frac{3.92}{\sqrt{n}}\,\right]$$

For $n = 50$ and $S = \sum_{i=1}^{n} Y_i = 40$, we thus have:

$$\widehat{\theta}_{\mathrm{MLE}} \pm 1.96\sqrt{\mathrm{Var}\Big[\widehat{\theta}_{\mathrm{MLE}}\Big]} \;=\; \left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)^2\left[\,1 \pm \frac{3.92}{\sqrt{n}}\,\right] \;=\; \left(\frac{40}{50}\right)^2\left(1 \pm \frac{3.92}{\sqrt{50}}\right) \;\approx\; (0.285,\ 0.995)$$

Exercise 4.8(b)

The probability density function of the sample observations, given θ, is:

$$f_{\mathbf{Y}}(y_1,\ldots,y_n\,|\,\theta) \;=\; \prod_{i=1}^{n}\frac{1}{\alpha}\exp\left(-\,\frac{y_i}{\alpha}\right) \;=\; \prod_{i=1}^{n}\theta^{-1/2}\exp\left(-\,\theta^{-1/2}\cdot y_i\right) \;=\; \theta^{-n/2}\exp\left(-\,\theta^{-1/2}\sum_{i=1}^{n}y_i\right)$$

As θ varies, this gives an exponential family of probability distributions. Indeed,

$$f_{\mathbf{Y}}(y_1,\ldots,y_n\,|\,\theta) \;=\; \theta^{-n/2}\exp\left\{(-\,\theta^{-1/2})\cdot\sum_{i=1}^{n}y_i\right\} \;=\; (-\,\tau)^n\exp\left(\tau\,S(\mathbf{y})\right)\,,$$

where $\tau := -\,\theta^{-1/2}$ and $S(\mathbf{y}) := \sum_{i=1}^{n}y_i$. By the general theory of exponential families, we now see that $S(\mathbf{Y}) := \sum_{i=1}^{n}Y_i$ is a complete and sufficient statistic of $\tau = -\,\theta^{-1/2}$, and hence of θ. Consequently, by the Rao-Blackwell Theorem, the unique MVUE (minimum variance unbiased estimator) of θ can be given as $E\left[\,U\,|\,S\,\right]$, for any unbiased estimator U of the θ.

Next, we claim that $\dfrac{n}{n+1}\cdot\widehat{\theta}_{\mathrm{MLE}}$ is an unbiased estimator of θ. First, observe that:

$$\begin{aligned}
E\left[\widehat{\theta}_{\mathrm{MLE}}\right] &= E\left[\left(\frac{1}{n}\sum_{i=1}^{n}Y_i\right)^2\right] \;=\; E\left[\frac{1}{n^2}\left(\sum_{i=1}^{n}Y_i^2+\sum_{i\neq j}Y_iY_j\right)\right] \\
&= \frac{1}{n^2}\left\{\,n\,E[\,Y^2\,]+(n^2-n)E[\,Y\,]\,E[\,Y\,]\,\right\} \;=\; \frac{1}{n^2}\left\{\,n\left(\mathrm{Var}(\,Y\,)+E(Y)^2\right)+n(n-1)E[\,Y\,]^2\,\right\} \\
&= \frac{1}{n^2}\left\{\,n\left(\alpha^2+\alpha^2\right)+n(n-1)\alpha^2\,\right\} \;=\; \frac{n\,\alpha^2}{n^2}\left\{\,2+n-1\,\right\} \;=\; \left(\frac{n+1}{n}\right)\alpha^2 \;=\; \left(\frac{n+1}{n}\right)\theta
\end{aligned}$$

which implies

$$E\left[\left(\frac{n}{n+1}\right)\widehat{\theta}_{\mathrm{MLE}}\right] \;=\; \theta$$

Hence, $\left(\dfrac{n}{n+1}\right)\widehat{\theta}_{\mathrm{MLE}}$ is indeed an unbiased estimator of θ, and

$$\widehat{\theta}^{\,*} \;:=\; E\left[\left(\frac{n}{n+1}\right)\widehat{\theta}_{\mathrm{MLE}}\;\middle|\;S\right] \;=\; E\left[\left(\frac{n}{n+1}\right)\left(\frac{1}{n}\sum_{i=1}^{n}Y_i\right)^2\;\middle|\;S\right] \;=\; \frac{1}{n(n+1)}E\left[\,S^2\;\middle|\;S\,\right] \;=\; \frac{S^2}{n(n+1)}$$

is the unique MVUE of θ. Next, we compute $\mathrm{Var}\left(\widehat{\theta}^{\,*}\right)$:

$$\mathrm{Var}\left(\widehat{\theta}^{\,*}\right) \;=\; E\left[\left(\widehat{\theta}^{\,*}\right)^2\right]-E\left[\widehat{\theta}^{\,*}\right]^2 \;=\; E\left[\left(\frac{S^2}{n(n+1)}\right)^2\right]-E\left[\widehat{\theta}^{\,*}\right]^2 \;=\; \frac{1}{n^2(n+1)^2}\,E\left[\,S^4\,\right]-\theta^2$$

Thus, we need to compute $E\left[\,S^4\,\right]$. To this end, we claim that $S\sim\mathrm{Gamma}(\alpha=\theta^{1/2},\,\beta=n)$, and we will the moment generating function $M_S(t)$ of S to compute its fourth moment $E\left[\,S^4\,\right]$.

$$\begin{aligned}
M_S(t) \;&:=\; E\left[\,e^{tS}\,\right] \;=\; \int e^{ts}f_{\mathbf{Y}}(\mathbf{y})\,\mathrm{d}y_1\mathrm{d}y_2\cdots\mathrm{d}y_n \;=\; \int\exp\left\{t\cdot\sum_{i=1}^{n}y_i\right\}\cdot\prod_{i=1}^{n}\theta^{-1/2}\exp(-\,\theta^{-1/2}y_i)\,\mathrm{d}y_1\mathrm{d}y_2\cdots\mathrm{d}y_n \\
&= \int\prod_{i=1}^{n}\exp(ty_i)\cdot\theta^{-1/2}\exp(-\,\theta^{-1/2}y_i)\,\mathrm{d}y_1\mathrm{d}y_2\cdots\mathrm{d}y_n \;=\; \prod_{i=1}^{n}\left(\int\exp(ty_i)\cdot\theta^{-1/2}\exp(-\,\theta^{-1/2}y_i)\,\mathrm{d}y_i\right) \\
&= \prod_{i=1}^{n}E\left[\,e^{ty_i}\,\right] \;=\; \prod_{i=1}^{n}M_{Y_i}(t) \;=\; \prod_{i=1}^{n}\left(1-\theta^{1/2}\,t\right)^{-1} \;=\; \left(1-\theta^{1/2}\,t\right)^{-n} \;=\; M_X(t),
\end{aligned}$$

where $X \sim \text{Gamma}(\alpha = \theta^{1/2},\, \beta = n)$. This shows that $S := \sum_{i=1}^{n} Y_i \sim \text{Gamma}(\alpha = \theta^{1/2},\, \beta = n)$. Hence,

$$\begin{aligned}
M_S(t) &= \left(1 - \theta^{1/2}\,t\right)^{-n} \\
M'_S(t) &= n\,\theta^{1/2}\left(1 - \theta^{1/2}\,t\right)^{-(n+1)} \\
M''_S(t) &= n(n+1)\,\theta\left(1 - \theta^{1/2}\,t\right)^{-(n+2)} \\
M^{(3)}_S(t) &= n(n+1)(n+2)\,\theta^{3/2}\left(1 - \theta^{1/2}\,t\right)^{-(n+3)} \\
M^{(4)}_S(t) &= n(n+1)(n+2)(n+3)\,\theta^{2}\left(1 - \theta^{1/2}\,t\right)^{-(n+4)}
\end{aligned}$$

And, we now have

$$E\left[\,S^4\,\right] = M^{(4)}_S(0) = n(n+1)(n+2)(n+3)\,\theta^2\left(1 - \theta^{1/2}\cdot 0\right)^{-(n+4)} = n(n+1)(n+2)(n+3)\,\theta^2$$

Hence, an explicit expression for $\text{Var}\left(\widehat{\theta}^{\,*}\right)$ is:

$$\text{Var}\left(\widehat{\theta}^{\,*}\right) = \frac{1}{n^2(n+1)^2}\,E\left[\,S^4\,\right] - \theta^2 = \frac{n(n+1)(n+2)(n+3)\,\theta^2}{n^2(n+1)^2} - \theta^2 = \theta^2\left(\frac{(n+2)(n+3)}{n(n+1)} - 1\right) = \theta^2\cdot\frac{2(2n+3)}{n(n+1)}$$

Exercise 4.8(c)

No, $\widehat{\theta}^{\,*}$ does NOT achieve the Cramér-Rao lower bound for the variance of unbiased estimators of θ. The following calculation shows indeed that $\text{Var}\left(\widehat{\theta}^{\,*}\right) > \text{CRLB}(\theta)$:

$$\text{Var}\left(\widehat{\theta}^{\,*}\right) = \theta^2\cdot\frac{2(2n+3)}{n(n+1)} = \frac{4\,\theta^2}{n}\cdot\frac{2n+3}{2n+2} > \frac{4\,\theta^2}{n} = \text{CRLB}(\theta)$$

Exercise 4.8(d)

Recall that for any random variable X and any constant θ, we have:

$$\text{MSE}(X,\theta) = E\left[\,(X-\theta)^2\,\right] = \cdots = E\left[\,(X-\mu_X+\mu_X-\theta)^2\,\right] = E\left[\,(X-\mu_X)^2\,\right] + (\mu_X-\theta)^2 = \text{Var}(X) + (\mu_X-\theta)^2$$

Since $\widehat{\theta}^{\,*}$ is an unbiased estimator of θ, we have

$$\text{MSE}\left(\widehat{\theta}^{\,*},\theta\right) = \text{Var}\left(\widehat{\theta}^{\,*}\right) + \left(E\left[\,\widehat{\theta}^{\,*}\,\right]-\theta\right)^2 = \text{Var}\left(\widehat{\theta}^{\,*}\right) + (\,0\,)^2 = \theta^2\cdot\frac{2(2n+3)}{n(n+1)} = \frac{\theta^2}{n}\left(\frac{4n+6}{n+1}\right) = \frac{\theta^2}{n}\left(4 + \frac{2}{n+1}\right)$$

On the other hand,

$$\begin{aligned}
\text{MSE}\left(\widehat{\theta}_{\text{MLE}},\theta\right) &= \text{Var}\left(\widehat{\theta}_{\text{MLE}}\right) + \left(E\left[\,\widehat{\theta}_{\text{MLE}}\,\right]-\theta\right)^2 = \text{Var}\left[\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)^2\right] + \left(\left(\frac{n+1}{n}\right)\theta - \theta\right)^2 \\
&= \text{Var}\left[\frac{1}{n^2}S^2\right] + \frac{\theta^2}{n^2} = \text{Var}\left[\frac{n(n+1)}{n^2}\,\frac{S^2}{n(n+1)}\right] + \frac{\theta^2}{n^2} = \frac{n^2(n+1)^2}{n^4}\text{Var}\left[\,\widehat{\theta}^{\,*}\,\right] + \frac{\theta^2}{n^2} \\
&= \frac{(n+1)^2}{n^2}\,\frac{\theta^2}{n}\left(\frac{4n+6}{n+1}\right) + \frac{\theta^2}{n^2} = \frac{\theta^2}{n}\left\{\frac{(n+1)(4n+6)}{n^2} + \frac{1}{n}\right\} \\
&= \frac{\theta^2}{n}\left\{\frac{4n^2+10n+6+n}{n^2}\right\} = \frac{\theta^2}{n}\left\{4 + \frac{11}{n} + \frac{6}{n^2}\right\}
\end{aligned}$$

To compare $\mathrm{MSE}\left(\widehat{\theta}_{\mathrm{MLE}},\theta\right)$ and $\mathrm{MSE}\left(\widehat{\theta}^{\,*},\theta\right)$, note that

$$\mathrm{MSE}\left(\widehat{\theta}_{\mathrm{MLE}},\theta\right) \;=\; \cdots \;=\; \frac{n^2(n+1)^2}{n^4}\mathrm{Var}\left[\,\widehat{\theta}^{\,*}\,\right] + \frac{\theta^2}{n^2} \;=\; \frac{(n+1)^2}{n^2}\,\mathrm{MSE}\left(\widehat{\theta}^{\,*},\theta\right) + \frac{\theta^2}{n^2} \;>\; \mathrm{MSE}\left(\widehat{\theta}^{\,*},\theta\right)$$

So, for finite n, the estimator $\widehat{\theta}^{\,*}$ is preferred since it has smaller mean squared error. However, we also easily see that

$$\lim_{n\to\infty}\mathrm{MSE}\left(\widehat{\theta}_{\mathrm{MLE}},\theta\right) \;=\; \lim_{n\to\infty}\left(\frac{(n+1)^2}{n^2}\,\mathrm{MSE}\left(\widehat{\theta}^{\,*},\theta\right) + \frac{\theta^2}{n^2}\right) \;=\; \mathrm{MSE}\left(\widehat{\theta}^{\,*},\theta\right)$$

Thus, asymptotically, there is no difference in the mean squared error between $\widehat{\theta}_{\mathrm{MLE}}$ and $\widehat{\theta}^{\,*}$.

Exercise 4.12(a)

The likelihood function is:

$$\begin{aligned}
L(\lambda_0,\lambda_1) &= \prod_{i=0}^{1}\prod_{j=1}^{n_i} P(Y_{ij}=y_{ij}) \;=\; \prod_{i=0}^{1}\prod_{j=1}^{n_i} \exp(-\,L_{ij}\lambda_i)\,\frac{\lambda_i^{y_{ij}}}{y_{ij}!} \\
&= \frac{1}{\prod_{i=0}^{1}\prod_{j=1}^{n_i}\, y_{ij}!}\,\exp\left(-\sum_{i=0}^{1}\sum_{j=1}^{n_i} L_{ij}\lambda_i\right)\prod_{i=0}^{1}\lambda_i^{\sum_{j=1}^{n_i} y_{ij}}
\end{aligned}$$

The log likelihood function is thus:

$$l(\lambda_0,\lambda_1) \;=\; \log\, L(\lambda_0,\lambda_1) \;=\; -\,\log K - \sum_{i=0}^{1}\left(\lambda_i\sum_{j=1}^{n_i} L_{ij}\right) + \sum_{i=0}^{1}\left(\sum_{j=1}^{n_i} y_{ij}\right)\log(\lambda_i)$$

The partial derivatives of $l(\lambda_0,\lambda_1)$ with respect to λ_0 and λ_1 are:

$$\frac{\partial l}{\partial\lambda_0} \;=\; -\sum_{j=1}^{n_0} L_{0j} + \frac{1}{\lambda_0}\sum_{j=1}^{n_0} y_{0j}\,, \qquad \frac{\partial l}{\partial\lambda_1} \;=\; -\sum_{j=1}^{n_1} L_{1j} + \frac{1}{\lambda_1}\sum_{j=1}^{n_1} y_{1j}$$

Setting these two partial derivatives to zero yields:

$$\frac{\partial l}{\partial\lambda_0} = 0 \;\Longrightarrow\; \lambda_0 = \frac{\sum_{j=1}^{n_0} y_{0j}}{\sum_{j=1}^{n_0} L_{0j}}\,, \qquad \frac{\partial l}{\partial\lambda_1} = 0 \;\Longrightarrow\; \lambda_1 = \frac{\sum_{j=1}^{n_1} y_{1j}}{\sum_{j=1}^{n_1} L_{1j}}$$

Hence,

$$\widehat{\lambda}_0 = \frac{\sum_{j=1}^{n_0} Y_{0j}}{\sum_{j=1}^{n_0} L_{0j}}\,, \qquad \widehat{\lambda}_1 = \frac{\sum_{j=1}^{n_1} Y_{1j}}{\sum_{j=1}^{n_1} L_{1j}}$$

Consequently,

$$\widehat{\log\psi} \;=\; \log\left(\widehat{\lambda}_1\right) - \log\left(\widehat{\lambda}_0\right) \;=\; \log\left(\frac{\sum_{j=1}^{n_0} Y_{1j}}{\sum_{j=1}^{n_1} L_{1j}}\right) - \log\left(\frac{\sum_{j=1}^{n_1} Y_{0j}}{\sum_{j=1}^{n_0} L_{0j}}\right)$$

Next,

$$\begin{aligned}
\mathrm{Var}\left[\widehat{\log\psi}\right] &= \mathrm{Var}\left[\log\left(\widehat{\lambda}_1\right) - \log\left(\widehat{\lambda}_0\right)\right] \;=\; \mathrm{Var}\left[\log\left(\widehat{\lambda}_1\right)\right] + \mathrm{Var}\left[\log\left(\widehat{\lambda}_0\right)\right] \;=\; \left(\frac{1}{\widehat{\lambda}_1}\right)^2 \mathrm{Var}\left[\widehat{\lambda}_1\right] + \left(\frac{1}{\widehat{\lambda}_0}\right)^2 \mathrm{Var}\left[\widehat{\lambda}_0\right] \\
&= \frac{\left(\sum_{j=1}^{n_1} L_{1j}\right)^2}{\left(\sum_{j=1}^{n_1} Y_{1j}\right)^2}\,\frac{\sum_{j=1}^{n_1} Y_{1j}}{\left(\sum_{j=1}^{n_1} L_{1j}\right)^2} + \frac{\left(\sum_{j=1}^{n_0} L_{0j}\right)^2}{\left(\sum_{j=1}^{n_0} Y_{0j}\right)^2}\,\frac{\sum_{j=1}^{n_0} Y_{0j}}{\left(\sum_{j=1}^{n_0} L_{0j}\right)^2} \\
&= \frac{1}{\sum_{j=1}^{n_1} Y_{1j}} + \frac{1}{\sum_{j=1}^{n_0} Y_{0j}}
\end{aligned}$$

Exercise 4.12(b)

$$\begin{aligned}
\widehat{\log\psi} &= \log\left(\frac{\sum_{j=1}^{n_0} Y_{1j}}{\sum_{j=1}^{n_1} L_{1j}}\right) - \log\left(\frac{\sum_{j=1}^{n_1} Y_{0j}}{\sum_{j=1}^{n_0} L_{0j}}\right) \\
&= \log\left(\frac{40}{350}\right) - \log\left(\frac{35}{400}\right) \;=\; \log\left(\frac{40}{350}\,\frac{400}{35}\right) \;=\; \log\left(\frac{8}{7}\right)^2 \;=\; 2\,\log\left(\frac{8}{7}\right) \;\approx\; 0.2670
\end{aligned}$$

$$\mathrm{Var}\left[\widehat{\log\psi}\right] \;=\; \frac{1}{\sum_{j=1}^{n_1} Y_{1j}} + \frac{1}{\sum_{j=1}^{n_0} Y_{0j}} \;=\; \frac{1}{40} + \frac{1}{35} \;\approx\; 0.05357$$

Approximate 95% confidence interval for $\log\psi$:

$$2\cdot\log\left(\frac{8}{7}\right) \;\pm\; 1.96 \times \sqrt{\frac{1}{40}+\frac{1}{35}} \;\approx\; (-\,0.18659,\, 0.72071)$$

Approximate 95% confidence interval for ψ:

$$(e^{-\,0.18659},\, e^{0.72071}) \;\approx\; (0.8298,\, 2.05589)$$

Exercise 4.13(a)

The likelihood function is:

$$L(\theta) \;=\; \prod_{i=1}^{n} f_T(t_i;\theta) \;=\; \prod_{i=1}^{n} \theta\, e^{-\theta\, t_i} \;=\; \theta^n\, e^{-\theta \sum_{i=1}^{n} t_i}$$

The log likelihood function is:

$$l(\theta) \;=\; \log L(\theta) \;=\; n\,\log(\theta) - \theta \sum_{i=1}^{n} t_i$$

$$\frac{\partial l}{\partial \theta} \;=\; \frac{n}{\theta} \;-\; \sum_{i=1}^{n} t_i$$

Thus,

$$\frac{\partial l}{\partial \theta} \;=\; 0 \quad\Longrightarrow\quad \theta \;=\; \frac{n}{\sum_{i=1}^{n} t_i}$$

Hence,

$$\widehat{\theta}_{\mathrm{MLE}} \;=\; \frac{n}{\sum_{i=1}^{n} T_i}$$

Next, note that:

$$\frac{\partial^2 l}{\partial \theta^2} \;=\; -\,\frac{n}{\theta^2}$$

Hence,

$$n\,\mathrm{Var}\Big[\widehat{\theta}_{\mathrm{MLE}}\Big] \;\approx\; n\left(-E\left[\frac{\partial^2 l}{\partial\theta^2}\right]\right)^{-1} \;=\; n\left(-E\Big[-\frac{n}{\theta^2}\Big]\right)^{-1} \quad\Longrightarrow\quad \mathrm{Var}\Big[\widehat{\theta}_{\mathrm{MLE}}\Big] \;=\; \frac{\theta^2}{n}$$

Exercise 4.13(b)

First, note that

$$P(Y=1) \;=\; P(T>t^*) \;=\; \int_{t^*}^{\infty} f_T(t;\theta)\,\mathrm{d}t \;=\; \int_{t^*}^{\infty} \theta\, e^{-\theta t}\,\mathrm{d}t \;=\; \theta\left[-\frac{e^{-\theta t}}{\theta}\right]_{t^*}^{\infty} \;=\; e^{-\theta t^*}$$

Hence,

$$P(Y=0) \;=\; 1 - P(Y=1) \;=\; 1 - e^{-\theta t^*}$$

The likelihood function is thus:

$$L^*(\theta) \;=\; \prod_{i=1}^{n} P(Y_i=1)^{Y_i}\,P(Y_i=0)^{1-Y_i} \;=\; P(Y=1)^{\sum_{i=1}^{n} Y_i}\cdot P(Y=0)^{n-\sum_{i=1}^{n} Y_i} \;=\; \left(e^{-\theta t^*}\right)^{S}\cdot\left(1-e^{-\theta t^*}\right)^{n-S}\,,$$

where $S := \sum_{i=1}^{n} Y_i$. Hence, the log likelihood function is:

$$l^*(\theta) \;=\; \log L^*(\theta) \;=\; -\,\theta t^* S \,+\, (n-S)\log\left(1-e^{-\theta t^*}\right)$$

Hence,

$$\frac{\partial l^*}{\partial\theta} \;=\; -t^*S + \frac{n-S}{1-e^{-\theta t^*}}\left(-\,e^{-\theta t^*}\cdot(-t^*)\right) \;=\; -t^*S + \frac{t^* e^{-\theta t^*}(n-S)}{1-e^{-\theta t^*}}$$

And,

$$\frac{\partial l^*}{\partial\theta} \;=\; 0 \quad\Longrightarrow\quad \cdots \quad\Longrightarrow\quad \widehat{\theta}^* \;=\; \frac{1}{t^*}\log\left(\frac{n}{S}\right) \;=\; \frac{1}{t^*}\log\left(\frac{n}{\sum_{i=1}^{n} Y_i}\right)$$

Exercise 4.13(c)

$$\frac{\partial l^*}{\partial \theta} \;=\; -t^* S + \frac{t^* e^{-\theta t^*}(n-S)}{1-e^{-\theta t^*}} \;=\; -t^* S + (n-S)\cdot t^* \cdot e^{-\theta t^*} \cdot \left(1 - e^{-\theta t^*}\right)^{-1}$$

And,

$$\frac{\partial^2 l^*}{\partial \theta^2} \;=\; \cdots \;=\; = \;-\,(n-S)\,\frac{e^{-\theta t^*}(t^*)^2}{\left(1-e^{-\theta t^*}\right)^2}$$

Now, note that: $S \;\sim\; \text{Binomial}\big(\,n,\, e^{-\theta t^*}\,\big)$. Hence,

$$E\left[\,\frac{\partial^2 l^*}{\partial \theta^2}\,\right] \;=\; -\,(n - E[\,S\,])\,\frac{e^{-\theta t^*}(t^*)^2}{\left(1-e^{-\theta t^*}\right)^2} \;=\; -\,(n - n\,e^{-\theta t^*})\,\frac{e^{-\theta t^*}(t^*)^2}{\left(1-e^{-\theta t^*}\right)^2} \;=\; \frac{n\,e^{-\theta t^*}(t^*)^2}{(e^{-\theta t^*}-1)} \;=\; \frac{n(t^*)^2}{(1-e^{+\theta t^*})}$$

Hence,

$$\text{Var}\left[\,\widehat{\theta}^*\,\right] \;=\; -\left(E\left[\,\frac{\partial^2 l^*}{\partial \theta^2}\,\right]\right)^{-1} \;=\; \frac{\left(e^{\theta t^*}-1\right)}{n(t^*)^2}$$

Thus,

$$\frac{\text{Var}\left[\,\widehat{\theta}_{\text{MLE}}\,\right]}{\text{Var}\left[\,\widehat{\theta}^*\,\right]} \;=\; \frac{\theta^2}{n}\cdot\frac{n(t^*)^2}{(e^{\theta t^*}-1)} \;=\; \frac{(\theta t^*)^2}{(e^{\theta t^*}-1)}$$

Exercise 4.14(a)

First, note that $\theta = P(Y=0) = e^{-\lambda}$; hence, $\lambda = -\log(\theta)$. The likelihood function is thus:

$$L(\theta) \;=\; \prod_{i=1}^{n} P_Y(y_i;\lambda) \;=\; \prod_{i=1}^{n} e^{-\lambda}\frac{\lambda^{y_i}}{y_i!} \;=\; \left(e^{-\lambda}\right)^n \frac{\lambda^{\sum_{i=1}^n y_i}}{\prod_{i=1}^n y_i!} \;=\; \theta^n \frac{(-\log\theta)^{\sum_{i=1}^n y_i}}{\prod_{i=1}^n y_i!}$$

The log likelihood is thus:

$$l(\theta) \;=\; \log L(\theta) \;=\; -\log K + n\log\theta + \left(\sum_{i=1}^n y_i\right)\log(-\log\theta) \;=\; -\log K + n\log\theta + S\log(-\log\theta)\ ,$$

where $S := \sum_{i=1}^n Y_i$. Note that $S \sim \text{Poisson}(n\lambda)$.

$$\frac{\partial l}{\partial\theta} \;=\; \frac{n}{\theta} + S\frac{1}{-\log\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}(-\log\theta) \;=\; \frac{n}{\theta} + S\cdot\frac{1}{(-\log\theta)}\left(-\frac{1}{\theta}\right) \;=\; \frac{n}{\theta} + \frac{S}{\theta\log\theta}$$

$$\left(\frac{\partial l}{\partial\theta}\right)^2 \;=\; \left(\frac{n}{\theta} + \frac{S}{\theta\log\theta}\right)^2 \;=\; \frac{n^2}{\theta^2} - 2\cdot\left(\frac{n}{\theta}\right)\cdot\left(\frac{S}{\theta\log\theta}\right) + \frac{S^2}{(\theta\log\theta)^2} \;=\; \frac{n^2}{\theta^2} - \frac{2nS}{\theta^2\log\theta} + \frac{S^2}{(\theta\log\theta)^2}$$

$$\begin{aligned}
E\left[\left(\frac{\partial l}{\partial\theta}\right)^2\right] \;&=\; \frac{n^2}{\theta^2} + \frac{2n\,E[\,S\,]}{\theta^2\log\theta} + \frac{E\left[\,S^2\,\right]}{(\theta\log\theta)^2} \;=\; \frac{n^2}{\theta^2} + \frac{2n\,E[\,S\,]}{\theta^2\log\theta} + \frac{\mathrm{Var}[\,S\,] + E[\,S\,]^2}{(\theta\log\theta)^2} \\
&=\; \frac{n^2}{\theta^2} + \frac{2n\,(n\lambda)}{\theta^2\log\theta} + \frac{n\lambda + n^2\lambda^2}{(\theta\log\theta)^2} \;=\; \frac{n^2}{\theta^2} - \frac{2n^2\log\theta}{\theta^2\log\theta} - \frac{n\log\theta\,(1-n\log\theta)}{\theta^2\,(\log\theta)^2} \\
&=\; \frac{n}{\theta^2}\left(-n - \frac{1-n\log\theta}{\log\theta}\right) \;=\; \frac{n}{\theta^2}\left(\frac{-n\log\theta - 1 + n\log\theta}{\log\theta}\right) \\
&=\; -\frac{n}{\theta^2\log\theta} \;=\; \frac{n}{\theta^2\,(-\log\theta)}
\end{aligned}$$

Thus, the Cramér-Rao Lower Bound for the variance of any unbiased estimator of θ is:

$$\mathrm{CRLB}(\theta) \;=\; \left(E\left[\left(\frac{\partial l}{\partial\theta}\right)^2\right]\right)^{-1} \;=\; \frac{\theta^2\,(-\log\theta)}{n} \;=\; \frac{\lambda}{ne^{2\lambda}} \;=\; \frac{1}{e^{2\lambda}}\cdot\frac{\lambda}{n}$$

Next, we claim that A^S is an unbiased estimator of θ for a suitably chosen constant A. To this end, note that, since $S = \sum_{i=1}^n Y_i \sim \text{Poisson}(n\lambda)$, the moment generating function of S is:

$$M_S(t) \;=\; E\left[\,e^{tS}\,\right] \;=\; e^{n\lambda\left(e^t - 1\right)}$$

Hence,

$$E\left[\,A^S\,\right] \;=\; E\left[e^{(\log A)S}\right] \;=\; M_S(\log A) \;=\; e^{n\lambda\left(e^{\log A} - 1\right)} \;=\; \left(\theta^{-n}\right)^{A-1}$$

Now, we seek constant A such that $E\left[\,A^S\,\right] = \theta$; so we set

$$\theta \;=\; E\left[\,A^S\,\right] \;=\; \left(\theta^{-n}\right)^{A-1}\ ,$$

which implies

$$\log\theta \;=\; -n\,(A-1)\log\theta \quad\Longrightarrow\quad A \;=\; 1 - \frac{1}{n} \;=\; \frac{n-1}{n}$$

Thus, we may now conclude:

$$\widehat{\theta} \;:=\; \left(\frac{n-1}{n}\right)^{S}$$

is an unbiased estimator of θ. Now,

$$\begin{aligned}
\mathrm{Var}\Big[\,\widehat{\theta}\,\Big] &= E\Big[\,\widehat{\theta}^{\,2}\Big] - E\Big[\,\widehat{\theta}\,\Big]^2 \;=\; E\left[\left(\frac{n-1}{n}\right)^{2S}\right] - E\left[\left(\frac{n-1}{n}\right)^{S}\right]^2 \\
&= E\left[\exp\left\{S\cdot 2\log\left(\frac{n-1}{n}\right)\right\}\right] - E\left[\exp\left\{S\cdot\log\left(\frac{n-1}{n}\right)\right\}\right]^2 \\
&= M_S\left(2\log\left(\frac{n-1}{n}\right)\right) - M_S\left(\log\left(\frac{n-1}{n}\right)\right)^2 \\
&= \exp\left\{n\lambda\left(\left(1-\frac{1}{n}\right)^2 - 1\right)\right\} - \exp\left\{2n\lambda\left(\left(1-\frac{1}{n}\right) - 1\right)\right\} \\
&= \exp\left\{-2\lambda + \frac{\lambda}{n}\right\} - \exp\{-2\lambda\} \;=\; \exp\{-2\lambda\}\left(\exp\frac{\lambda}{n} - 1\right) \;=\; \frac{1}{e^{2\lambda}}\cdot\left(e^{\lambda/n} - 1\right) \\
&= \frac{1}{e^{2\lambda}}\cdot\left(1 + \frac{\lambda}{n} + \frac{(\lambda/n)^2}{2!} + \sum_{k=3}^{\infty}\frac{(\lambda/n)^k}{k!} - 1\right) \;=\; \frac{1}{e^{2\lambda}}\cdot\left(\frac{\lambda}{n} + \frac{(\lambda/n)^2}{2!} + \sum_{k=3}^{\infty}\frac{(\lambda/n)^k}{k!}\right) \\
&> \frac{1}{e^{2\lambda}}\cdot\frac{\lambda}{n} \;=\; \mathrm{CRBL}(\theta)
\end{aligned}$$

Hence, we see that the variance of the unbiased estimator $\widehat{\theta}$ of θ strictly exceeds the Cramér-Rao Lower Bound, for each n. Since $S = \sum_{i=1}^{n} Y_i$ is a complete sufficient statistic for θ (look up complete statistic for exponential families), we have

$$E\Big[\,\widehat{\theta}\;\Big|\;S\,\Big] \;=\; E\left[\left(\frac{n-1}{n}\right)^{S}\;\middle|\;S\right] \;=\; \left(\frac{n-1}{n}\right)^{S} \;=\; \widehat{\theta}$$

is the unique minimum variance unbiased estimator (MVUE) of θ. Since $\widehat{\theta}$ does NOT attain the Cramér-Rao Lower Bound for the variance of unbiased estimators of θ, we may now conclude that no unbiased estimators of θ attain that bound.

Exercise 4.14(b)

Exercise 4.14(c)

5 Chapter 5

Exercise 5.4(a)

The likelihood function is:

$$L(\theta_1, \theta_2) \; = \; \prod_{i=1}^{n} \theta_1^{x_i} \, (1-\theta_1)^{1-x_i} \cdot \prod_{j=1}^{n} \theta_2^{y_j} \, (1-\theta_2)^{1-y_j} \; = \; \theta_1^{S_X} \cdot (1-\theta_1)^{n-S_X} \cdot \theta_2^{S_Y} \cdot (1-\theta_2)^{n-S_Y}$$

where $S_X := \sum_{i=1}^{n} X_i$ and $S_Y := \sum_{j=1}^{n} Y_j$. Hence

$$\frac{L(0.5, 0.5)}{L(0.6, 0.6)} \; = \; \frac{0.5^{S_X} \cdot (1-0.5)^{n-S_X} \cdot 0.5^{S_Y} \cdot (1-0.5)^{n-S_Y}}{0.6^{S_X} \cdot (1-0.6)^{n-S_X} \cdot 0.6^{S_Y} \cdot (1-0.6)^{n-S_Y}} \; = \; \left(\frac{5}{4}\right)^{2n} \cdot \left(\frac{2}{3}\right)^{S_X+S_Y}$$

By the Neyman-Pearson Lemma, the most powerful test has rejection region of the form:

$$\frac{L(0.5, 0.5)}{L(0.6, 0.6)} < k \quad \Longrightarrow \quad \left(\frac{5}{4}\right)^{2n} \cdot \left(\frac{2}{3}\right)^{S_X+S_Y} < k \quad \Longrightarrow \quad S_X + S_Y > -\frac{1}{\log(3/2)} \left(\log k - 2n \log(5/4)\right)$$

Thus, we see that the most powerful test has rejection region of the form:

$$S_X + S_Y > k'$$

for some suitable k'.

Now, assume $n = 30$ and $\alpha = 0.05$. Under the null hypothesis $H_0 : \theta_1 = \theta_2 = 0.5$, note that $S_X + S_Y \sim \text{Binomial}(N = 60, p = 0.5)$. We see the threshold τ such that $P(S_X + S_Y \geq \tau \,|\, H_0 : \theta_1 = \theta_2 = 0.5) = \alpha = 0.05$. In other words, τ is the smallest integer such that $P(W \geq \tau) \leq 0.05$, where $W \sim \text{Binomial}(N = 60, p = 0.5)$. Now, note that $P(W \geq 36) \approx 0.07750095$ and $P(W \geq 37) \approx 0.04623049$. We conclude that the threshold $\tau = 37$.

The power of the test is thus:

$$P(\, S_X + S_Y \geq 37 \,|\, H_1 : \theta_1 = \theta_2 = 0.6 \,) \; = \; P(V \geq 37) \; \approx \; 0.4511064,$$

where $V \sim \text{Binomial}(N = 60, p = 0.6)$.

Exercise 5.4(b)

The likelihood function is:

$$L(\theta_1, \theta_2) \; = \; \prod_{i=1}^{n} \theta_1^{x_i} \, (1-\theta_1)^{1-x_i} \cdot \prod_{j=1}^{n} \theta_2^{y_j} \, (1-\theta_2)^{1-y_j} \; = \; \theta_1^{S_X} \cdot (1-\theta_1)^{n-S_X} \cdot \theta_2^{S_Y} \cdot (1-\theta_2)^{n-S_Y}$$

where $S_X := \sum_{i=1}^{n} X_i$ and $S_Y := \sum_{j=1}^{n} Y_j$. The log likelihood is thus:

$$l(\theta_1, \theta_2) \; = \; \log L(\theta_1, \theta_2) \; = \; S_X \log \theta_1 + (n - S_X) \log(1-\theta_1) + S_Y \log \theta_2 + (n - S_Y) \log(1-\theta_2)$$

Thus,

$$\frac{\partial l}{\partial \theta_1} \; = \; \frac{S_X}{\theta_1} - \frac{n - S_X}{1-\theta_1}\,, \quad \text{and} \quad \frac{\partial l}{\partial \theta_2} \; = \; \frac{S_Y}{\theta_2} - \frac{n - S_Y}{1-\theta_2}$$

Hence,

$$\frac{\partial l}{\partial \theta_1} \; = \; 0 \;\; \Longrightarrow \;\; \widehat{\theta}_1 \; = \; \frac{1}{n} S_X \; = \; \frac{1}{n} \sum_{i=1}^{n} X_i \quad \text{and} \quad \frac{\partial l}{\partial \theta_2} \; = \; 0 \;\; \Longrightarrow \;\; \widehat{\theta}_2 \; = \; \frac{1}{n} S_Y \; = \; \frac{1}{n} \sum_{j=1}^{n} Y_j$$

Now, define

$$\widehat{\tau} \;:=\; \widehat{\theta}_1 - \widehat{\theta}_2 \;=\; \frac{1}{n}\left(\sum_{i=1}^{n} X_i + \sum_{j=1}^{n} Y_j\right)$$

Then, $E[\,\widehat{\tau}\,] = \cdots = \theta_1 - \theta_2$, and

$$\begin{aligned}
\mathrm{Var}[\,\widehat{\tau}\,] &= \mathrm{Var}\left[\widehat{\theta}_1 - \widehat{\theta}_2\right] \;=\; \mathrm{Var}\left[\widehat{\theta}_1\right] + \mathrm{Var}\left[\widehat{\theta}_2\right] \;=\; \mathrm{Var}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] + \mathrm{Var}\left[\frac{1}{n}\sum_{j=1}^{n} Y_j\right] \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}[\,X_i\,] + \frac{1}{n^2}\sum_{j=1}^{n}\mathrm{Var}[\,Y_j\,] \;=\; \frac{1}{n^2}\sum_{i=1}^{n}\theta_1(1-\theta_1) + \frac{1}{n^2}\sum_{j=1}^{n}\theta_2(1-\theta_2) \\
&= \frac{1}{n}\left[\theta_1(1-\theta_1) + \theta_2(1-\theta_2)\right]\,,
\end{aligned}$$

since, for each $i = 1,\ldots,n$, we have $X_i \sim \mathrm{Bernoulli}(p = \theta_1)$, hence $\mathrm{Var}[\,X_i\,] = \theta_1(1-\theta_1)$, and similarly, for each $j = 1,\ldots,n$, we have $Y_i \sim \mathrm{Bernoulli}(p = \theta_2)$, hence $\mathrm{Var}[\,Y_j\,] = \theta_2(2-\theta_2)$.

Since $\tau := \widehat{\theta}_1 - \widehat{\theta}_2$ is the maximum likelihood estimator of $\theta_1 - \theta_2$, we know, by the general theory of maximum likelihood estimators, that

$$\frac{\widehat{\tau} - E[\,\widehat{\tau}\,]}{\sqrt{\mathrm{Var}[\,\widehat{\tau}\,]}} \;\overset{\text{asymp.}}{\sim}\; N(0,1)$$

Under that null hypothesis $H_0 : \theta_1 = \theta_2 = \theta_0$, we have:

$$E[\,\widehat{\tau}\,] \;=\; 0 \quad\text{and}\quad \mathrm{Var}[\,\tau\,] \;=\; \frac{2}{n}\theta_0(1-\theta_0)$$

So, a one-sided $\alpha = 0.05$ test has rejection region:

$$\frac{\widehat{\tau}}{\sqrt{(2/n)\theta_0(1-\theta_0)}} \;>\; z_{1-\alpha} \;=\; z_{0.95} \;\approx\; 1.644854 \quad\Longleftrightarrow\quad \widehat{\tau} \;>\; z_{0.95}\sqrt{\frac{2}{n}\,\theta_0(1-\theta_0)}$$

Next, the power of the test is:

$$\begin{aligned}
\text{power} &= P\left(\widehat{\tau} \geq z_{1-\alpha}\sqrt{\frac{2}{n}\,\theta_0(1-\theta_0)}\;\middle|\; H_1 : \theta_1 - \theta_2 \geq 0.2\right) \\
&\geq P\left(\widehat{\tau} \geq z_{1-\alpha}\sqrt{\frac{2}{n}\,\theta_0(1-\theta_0)}\;\middle|\; H_1 : \theta_1 - \theta_2 = 0.2\right) \\
&= P\left(\frac{\widehat{\tau} \;-\; 0.2}{\sqrt{\frac{1}{n}\left[\,\theta_1(1-\theta_1)+\theta_2(1-\theta_2)\,\right]}} \geq \frac{z_{1-\alpha}\sqrt{\frac{2}{n}\,\theta_0(1-\theta_0)} \;-\; 0.2}{\sqrt{\frac{1}{n}\left[\,\theta_1(1-\theta_1)+\theta_2(1-\theta_2)\,\right]}}\;\middle|\; H_1 : \theta_1 - \theta_2 = 0.2\right) \\
&\approx P\left(Z \geq \frac{z_{1-\alpha}\sqrt{\frac{2}{n}\,\theta_0(1-\theta_0)} \;-\; 0.2}{\sqrt{\frac{1}{n}\left[\,\theta_1(1-\theta_1)+\theta_2(1-\theta_2)\,\right]}}\;\middle|\; Z \sim N(0,1)\right) \\
&\geq P\left(Z \geq \frac{z_{1-\alpha}\sqrt{\frac{2}{n}\,\theta_0(1-\theta_0)} \;-\; 0.2}{\frac{1}{\sqrt{n}}\cdot\max\limits_{0<\theta_2<1}\left\{\sqrt{\left[\,(\theta_2+0.2)(1-\theta_2-0.2)+\theta_2(1-\theta_2)\,\right]}\right\}}\;\middle|\; Z \sim N(0,1)\right)
\end{aligned}$$

Now, let

$$f(\theta_2) \;:=\; (\theta_2+0.2)(1-\theta_2-0.2)+\theta_2(1-\theta_2) \;=\; -2\theta_2^2+1.6\theta_2+0.16$$

Then,

$$f'(\theta_2) \;=\; -4\theta_2+1.6$$

Hence,

$$f'(\theta_2) \;=\; 0 \quad\Longrightarrow\quad \theta_2 \;=\; \frac{1.6}{4} \;=\; \frac{2}{5}$$

And,

$$f\left(\frac{2}{5}\right) \;=\; -2\times\left(\frac{2}{5}\right)^2+1.6\times\left(\frac{2}{5}\right)+0.16 \;=\; -\frac{8}{25}+\frac{16}{10}\times\frac{2}{5}+\frac{16}{100} \;=\; -\frac{8}{25}+\frac{16}{25}+\frac{4}{25} \;=\; \frac{12}{25}$$

Therefore,

$$\begin{aligned}
\text{power} \;&=\; P\left(\widehat{\tau} \,\geq\, z_{1-\alpha}\sqrt{\frac{2}{n}\,\theta_0(1-\theta_0)}\;\middle|\; H_1:\theta_1-\theta_2\geq 0.2\right) \\
&\geq\; P\left(Z \,\geq\, \frac{z_{1-\alpha}\sqrt{\frac{2}{n}\,\theta_0(1-\theta_0)}\;-\;0.2}{\frac{1}{\sqrt{n}}\cdot\max\limits_{0<\theta_2<1}\left\{\sqrt{[\,(\theta_2+0.2)(1-\theta_2-0.2)+\theta_2(1-\theta_2)\,]}\right\}}\;\middle|\; Z\sim N(0,1)\right) \\
&=\; P\left(Z \,\geq\, \frac{z_{1-\alpha}\sqrt{\frac{2}{n}\,\theta_0(1-\theta_0)}\;-\;0.2}{\frac{1}{\sqrt{n}}\cdot\sqrt{\frac{12}{25}}}\;\middle|\; Z\sim N(0,1)\right) \\
&=\; P\left(Z \,\geq\, \sqrt{\frac{25}{12}}\left(z_{1-\alpha}\sqrt{2\theta_0(1-\theta_0)}\;-\;0.2\sqrt{n}\right)\;\middle|\; Z\sim N(0,1)\right) \\
&=\; P\left(Z \,\geq\, \frac{5}{2\sqrt{3}}\left(z_{1-\alpha}\sqrt{2\times 0.1(1-0.1)}\;-\;0.2\sqrt{n}\right)\;\middle|\; Z\sim N(0,1)\right) \\
&=\; P\left(Z \,\geq\, \frac{5}{2\sqrt{3}}\left(z_{1-\alpha}\sqrt{2\times\frac{9}{100}}\;-\;0.2\sqrt{n}\right)\;\middle|\; Z\sim N(0,1)\right) \\
&=\; P\left(Z \,\geq\, \frac{5}{2\sqrt{3}}\left(\frac{3\sqrt{2}}{10}z_{1-\alpha}\;-\;0.2\sqrt{n}\right)\;\middle|\; Z\sim N(0,1)\right) \\
&=\; P\left(Z \,\geq\, \frac{5}{2\sqrt{3}}\left(\frac{3}{5\sqrt{2}}z_{1-\alpha}\;-\;\frac{\sqrt{n}}{5}\right)\;\middle|\; Z\sim N(0,1)\right)
\end{aligned}$$

Since we want to impose power $\geq \beta = 0.9$, we seek n such that

$$\begin{aligned}
\frac{1}{2\sqrt{3}}\left(\frac{3}{\sqrt{2}}\,z_{1-\alpha}-\sqrt{n}\right) \;&=\; z_{1-\beta} \;=\; z_{1-0.9} \;=\; z_{0.1} \\
\Longrightarrow\quad n \;=\; \left(\frac{3}{\sqrt{2}}\cdot z_{0.95}-2\sqrt{3}\cdot z_{0.1}\right)^2 \;&=\; \left(\frac{3}{\sqrt{2}}\cdot 1.644854-2\sqrt{3}\cdot(-1.281552)\right)^2 \;=\; 62.86407
\end{aligned}$$

Thus, the required sample size is at least:

$$n \;\geq\; 63$$

References