MA204: Mathematical Statistics

Suggested Solutions to Assignment 3

3.1 **Solution**. The parameter space $\mathbf{\Theta} = \{\boldsymbol{\theta} = (\theta_1, \theta_2)^{\top}\colon -\infty < \theta_1 \leqslant \theta_2 < +\infty\}$. The joint density of $\mathbf{x} = (X_1, \ldots, X_n)^{\top}$ is

$$f(\boldsymbol{x}; \boldsymbol{\theta}) = \frac{1}{(\theta_2 - \theta_1)^n}, \quad \theta_1 \leqslant x_i \leqslant \theta_2,$$

so that the log-likelihood function is

$$\ell(\boldsymbol{\theta}) = -n \log(\theta_2 - \theta_1), \quad \theta_1 \leqslant x_{(1)} \quad \text{and} \quad \theta_2 \geqslant x_{(n)}.$$

Since $\partial \ell(\boldsymbol{\theta})/\partial \theta_2 = -n/(\theta_2 - \theta_1) < 0$; i.e., $\ell(\boldsymbol{\theta})$ is a monotonic decreasing function of θ_2 when θ_1 is fixed, so that the MLE os θ_2 is $X_{(n)}$.

Since $\partial \ell(\boldsymbol{\theta})/\partial \theta_1 = n/(\theta_2 - \theta_1) > 0$; i.e., $\ell(\boldsymbol{\theta})$ is a monotonic increasing function of θ_1 when θ_2 is fixed, so that the MLE os θ_1 is $X_{(1)}$.

3.2 **Solution**. (a) We know that the MLE of μ_1 is $\hat{\mu}_1 = \bar{X}_1$. Similarly, the MLE of μ_2 is $\hat{\mu}_2 = \bar{X}_2$. Then, by using Theorem 3.2, we obtain

$$\hat{\theta} = \hat{\mu}_1 - \hat{\mu}_2 = \bar{X}_1 - \bar{X}_2.$$

(b) Note that the two samples are independent, we have

$$\begin{aligned}
\mathrm{Var}(\hat{\theta}) &= \mathrm{Var}(\bar{X}_1) + \mathrm{Var}(\bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \\
&= \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n - n_1}.
\end{aligned}$$

To minimize $\mathrm{Var}(\hat{\theta})$, we treat n_1 as a continuous variable, differentiate $\mathrm{Var}(\hat{\theta})$ with respect to n_1 and set it to zero:

$$\frac{\mathrm{dVar}(\hat{\theta})}{\mathrm{d}n_1} = -\frac{\sigma_1^2}{n_1^2} + \frac{\sigma_2^2}{(n-n_1)^2} = 0.$$

By solving this equation, we obtain

$$n_1 = \frac{n\sigma_1}{\sigma_1+\sigma_2} \quad \text{and} \quad n_2 = \frac{n\sigma_2}{\sigma_1+\sigma_2}.$$

3.3 Solution. The likelihood function of (α, β) is

$$\begin{aligned} L(\alpha,\beta) &\propto (\alpha\beta)^{n_1}[\alpha(1-\beta)]^{n_2}[(1-\alpha)\beta]^{n_3}[(1-\alpha)(1-\beta)]^{n_4} \\ &= \alpha^{n_1+n_2}(1-\alpha)^{n_3+n_4}\cdot\beta^{n_1+n_3}(1-\beta)^{n_2+n_4}. \end{aligned}$$

The log-likelihood function is given by

$$\begin{aligned} \ell(\alpha,\beta) &= (n_1+n_2)\log\alpha + (n_3+n_4)\log(1-\alpha) \\ &\quad + (n_1+n_3)\log\beta + (n_2+n_4)\log(1-\beta). \end{aligned}$$

By partially differentiating $\ell(\alpha, \beta)$ with respect to both α and β and setting them to be zeros, we have

$$\begin{aligned} \frac{\partial\ell(\alpha,\beta)}{\partial\alpha} &= \frac{n_1+n_2}{\alpha} - \frac{n_3+n_4}{1-\alpha} = 0, \\ \frac{\partial\ell(\alpha,\beta)}{\partial\beta} &= \frac{n_1+n_3}{\beta} - \frac{n_2+n_4}{1-\beta} = 0. \end{aligned}$$

Hence,

$$\hat{\alpha} = \frac{n_1+n_2}{n} \quad \text{and} \quad \hat{\beta} = \frac{n_1+n_3}{n}.$$

3.4 Solution. Let $\boldsymbol{\theta} = (\mu_1, \mu_2, \mu_3, \mu_4, \sigma^2)^\top$, where

$$\begin{aligned} \mu_1 &= a+b+c, \quad \mu_2 = a+b-c, \\ \mu_3 &= a-b+c, \quad \mu_4 = a-b-c. \end{aligned}$$

Hence

$$\begin{aligned}
\frac{\partial\mu_i}{\partial a} &= 1, \quad i=1,2,3,4, \\
\frac{\partial\mu_i}{\partial b} &= 1, \quad i=1,2, \qquad \frac{\partial\mu_i}{\partial b} = -1. \quad i=3,4, \\
\frac{\partial\mu_i}{\partial c} &= 1, \quad i=1,3, \qquad \frac{\partial\mu_i}{\partial c} = -1. \quad i=2,4,
\end{aligned}$$

Since $X_{i1},\ldots,X_{in} \overset{\text{iid}}{\sim} N(\mu_i,\sigma^2)$ for $i=1,\ldots,4$ and the four random samples are independent, the likelihood function is

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{4}\prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x_{ij}-\mu_i)^2}{2\sigma^2}\right\}.$$

Thus, the log-likelihood function is

$$\ell(\boldsymbol{\theta}) = -2n\log(2\pi) - 2n\log(\sigma^2) - \frac{\sum_{i=1}^{4}\sum_{j=1}^{n}(x_{ij}-\mu_i)^2}{2\sigma^2}.$$

By partially differentiating $\ell(\boldsymbol{\theta})$ with respect to a, b, c, σ^2 and setting them to be zeros, we have

$$\begin{aligned}
\frac{\partial\ell(\boldsymbol{\theta})}{\partial a} &= -\frac{1}{2\sigma^2}\sum_{i=1}^{4}\sum_{j=1}^{n}(-2)(x_{ij}-\mu_i) = 0, && (3.1) \\
\frac{\partial\ell(\boldsymbol{\theta})}{\partial b} &= -\frac{1}{2\sigma^2}\left[\sum_{i=1}^{2}\sum_{j=1}^{n}(-2)(x_{ij}-\mu_i)\right. \\
&\quad \left. + \sum_{i=3}^{4}\sum_{j=1}^{n}(-2)(-1)(x_{ij}-\mu_i)\right] = 0, && (3.2) \\
\frac{\partial\ell(\boldsymbol{\theta})}{\partial c} &= -\frac{1}{2\sigma^2}\left[\sum_{i=1,3}\sum_{j=1}^{n}(-2)(x_{ij}-\mu_i)\right. \\
&\quad \left. + \sum_{i=2,4}\sum_{j=1}^{n}(-2)(-1)(x_{ij}-\mu_i)\right] = 0, && (3.3) \\
\frac{\partial\ell(\boldsymbol{\theta})}{\partial \sigma^2} &= -\frac{2n}{\sigma^2} + \frac{\sum_{i=1}^{4}\sum_{j=1}^{n}(x_{ij}-\mu_i)^2}{2\sigma^4}. && (3.4)
\end{aligned}$$

From (3.1), we have

$$\begin{aligned}
0 &= \sum_{i=1}^{4}\sum_{j=1}^{n}(x_{ij}-\mu_i) \\
&= \sum_{j=1}^{n}[x_{1j}-(a+b+c)] + \sum_{j=1}^{n}[x_{2j}-(a+b-c)] \\
&\quad + \sum_{j=1}^{n}[x_{3j}-(a-b+c)] + \sum_{j=1}^{n}[x_{4j}-(a-b-c)] \\
&= n\bar{x}_1 - n(a+b+c) + n\bar{x}_2 - n(a+b-c) \\
&\quad + n\bar{x}_3 - n(a-b+c) + n\bar{x}_4 - n(a-b-c) \\
&= n(\bar{x}_1+\bar{x}_2+\bar{x}_3+\bar{x}_4) - 4na,
\end{aligned}$$

i.e.,

$$\hat{a} = \frac{\bar{X}_1+\bar{X}_2+\bar{X}_3+\bar{X}_4}{4}. \tag{3.5}$$

From (3.2), we have

$$\begin{aligned}
0 &= \sum_{i=1}^{2}\sum_{j=1}^{n}(x_{ij}-\mu_i) - \sum_{i=3}^{4}\sum_{j=1}^{n}(x_{ij}-\mu_i) \\
&= \sum_{j=1}^{n}[x_{1j}-(a+b+c)] + \sum_{j=1}^{n}[x_{2j}-(a+b-c)] \\
&\quad - \sum_{j=1}^{n}[x_{3j}-(a-b+c)] - \sum_{j=1}^{n}[x_{4j}-(a-b-c)] \\
&= n\bar{x}_1 - n(a+b+c) + n\bar{x}_2 - n(a+b-c) \\
&\quad - n\bar{x}_3 + n(a-b+c) - n\bar{x}_4 + n(a-b-c) \\
&= n(\bar{x}_1+\bar{x}_2-\bar{x}_3-\bar{x}_4) - 4nb,
\end{aligned}$$

i.e.,

$$\hat{b} = \frac{\bar{X}_1+\bar{X}_2-\bar{X}_3-\bar{X}_4}{4}. \tag{3.6}$$

From (3.3), we have

$$
\begin{aligned}
0 &= \sum_{i=1,3}\sum_{j=1}^{n}(x_{ij}-\mu_i) - \sum_{i=2,4}\sum_{j=1}^{n}(x_{ij}-\mu_i) \\
&= \sum_{j=1}^{n}[x_{1j}-(a+b+c)] - \sum_{j=1}^{n}[x_{2j}-(a+b-c)] \\
&\quad + \sum_{j=1}^{n}[x_{3j}-(a-b+c)] - \sum_{j=1}^{n}[x_{4j}-(a-b-c)] \\
&= n\bar{x}_1 - n(a+b+c) - n\bar{x}_2 + n(a+b-c) \\
&\quad + n\bar{x}_3 - n(a-b+c) - n\bar{x}_4 + n(a-b-c) \\
&= n(\bar{x}_1 - \bar{x}_2 + \bar{x}_3 - \bar{x}_4) - 4nc,
\end{aligned}
$$

i.e.,

$$
\hat{c} = \frac{\bar{X}_1 - \bar{X}_2 + \bar{X}_3 - \bar{X}_4}{4}. \tag{3.7}
$$

From (3.4), we have

$$
\hat{\sigma}^2 = \frac{\sum_{i=1}^{4}\sum_{j=1}^{n}(X_{ij}-\bar{X}_i)^2}{4n}. \tag{3.8}
$$

3.5 Solution. The density of X is

$$
f(x;\mu,\sigma) = \frac{1}{2\sqrt{3}\sigma}\cdot I_{[\mu-\sqrt{3}\sigma,\,\mu+\sqrt{3}\sigma]}(x). \tag{3.9}
$$

Using the formulae in Appendix A.2.1, we have

$$
E(X) = \mu \quad \text{and} \quad \mathrm{Var}(X) = \sigma^2.
$$

Let $X_{(1)} = \min(X_1,\ldots,X_n)$ and $X_{(n)} = \max(X_1,\ldots,X_n)$. Furthermore, let $x_{(1)} = \min(x_1,\ldots,x_n)$ and $x_{(n)} = \max(x_1,\ldots,x_n)$ denote the realizations of $X_{(1)}$ and $X_{(n)}$, respectively.

(a) The likelihood function is given by

$$\begin{aligned} L(\mu,\sigma) &= \left(\frac{1}{2\sqrt{3}\sigma}\right)^n \prod_{i=1}^n I_{[\mu-\sqrt{3}\sigma,\, \mu+\sqrt{3}\sigma]}(x_i) \\ &= \left(\frac{1}{2\sqrt{3}\sigma}\right)^n \cdot I_{[\mu-\sqrt{3}\sigma,\, x_{(n)}]}(x_{(1)}) \cdot I_{[x_{(1)},\, \mu+\sqrt{3}\sigma]}(x_{(n)}) \\ &= \left(\frac{1}{2\sqrt{3}\sigma}\right)^n \cdot I_{[(\mu-x_{(1)})/\sqrt{3},\, \infty]}(\sigma) \cdot I_{[(x_{(n)}-\mu)/\sqrt{3},\, \infty]}(\sigma). \end{aligned}$$

Note that $L(\mu,\sigma)$ is $(2\sqrt{3}\sigma)^{-n}$ (a decreasing function of σ) if $\sigma \geqslant \max\{(\mu-x_{(1)})/\sqrt{3},\ (x_{(n)}-\mu)/\sqrt{3}\}$ and 0 elsewhere. Thus, when σ is smallest, which is the intersection of the lines $\mu-\sqrt{3}\sigma = x_{(1)}$ and $\mu+\sqrt{3}\sigma = x_{(n)}$. Hence, the mles of μ and σ are

$$\hat{\mu} = \frac{x_{(1)}+x_{(n)}}{2} \quad \text{and} \quad \hat{\sigma} = \frac{x_{(n)}-x_{(1)}}{2\sqrt{3}}.$$

Thus, the MLEs of μ and σ are

$$\hat{\mu}^{\mathrm{MLE}} = \frac{X_{(1)}+X_{(n)}}{2} \quad \text{and} \quad \hat{\sigma}^{\mathrm{MLE}} = \frac{X_{(n)}-X_{(1)}}{2\sqrt{3}}. \tag{3.10}$$

(b) The moment estimators of μ and σ must satisfy

$$\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i = E(X) = \mu, \quad \text{and}$$

$$\frac{1}{n}\sum_{i=1}^n X_i^2 = E(X^2) = \mathrm{Var}(X) + [E(X)]^2 = \sigma^2 + \mu^2.$$

Thus,

$$\hat{\mu}^{\mathrm{M}} = \bar{X} \quad \text{and} \quad \hat{\sigma}^{\mathrm{M}} = \sqrt{\frac{1}{n}\sum_{i=1}^n (X_i-\bar{X})^2} \tag{3.11}$$

are the corresponding moment estimators of μ and σ.

3.6 Solution. (a) The likelihood function

$$
\begin{aligned}
L(\theta) &= \prod_{i=1}^{n} \mathrm{e}^{-(x_i-\theta)} \cdot I_{[\theta,\infty)}(x_i) \\
&= \mathrm{e}^{-\sum_{i=1}^{n} x_i + n\theta} \prod_{i=1}^{n} I_{[\theta,\infty)}(x_i) \\
&= \mathrm{e}^{-n\bar{x}+n\theta} \cdot I_{[\theta,\infty)}(x_{(1)}) \\
&= \mathrm{e}^{-n\bar{x}+n\theta} \cdot I_{(-\infty,x_{(1)}]}(\theta)
\end{aligned}
$$

is an increasing function of θ. When $\theta = x_{(1)}$, $L(\theta)$ reaches its maximum. Thus, the MLE of θ is $X_{(1)}$.

(b) Let $y = x - \theta$, we obtain

$$
E(X) = \int_{\theta}^{\infty} x\mathrm{e}^{-(x-\theta)}\,\mathrm{d}x = \int_{0}^{\infty} (y+\theta)\mathrm{e}^{-y}\,\mathrm{d}y = 1+\theta.
$$

The moment estimator of θ must satisfy

$$
\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = E(X) = 1+\theta.
$$

We have $\hat{\theta}^{\mathrm{M}} = \bar{X} - 1$.

(c) The joint pdf of $X_1, \ldots, X_n$ and θ is

$$
\begin{aligned}
f(x_1,\ldots,x_n,\theta) &= L(\theta) \times \pi(\theta) \\
&= \mathrm{e}^{-n\bar{x}+n\theta} \cdot I_{(-\infty,\,x_{(1)}]}(\theta) \times \mathrm{e}^{-\theta} I_{(0,\infty)}(\theta) \\
&= \mathrm{e}^{-n\bar{x}+(n-1)\theta} \cdot I_{(0,\,x_{(1)}]}(\theta).
\end{aligned}
$$

Thus, the posterior density is

$$
p(\theta|x_1,\ldots,x_n) \propto f(x_1,\ldots,x_n,\theta) \propto \mathrm{e}^{(n-1)\theta}, \quad 0 < \theta \leqslant x_{(1)}.
$$

That is, $p(\theta|x_1,\ldots,x_n) = c^{-1}\mathrm{e}^{(n-1)\theta}, \quad 0 < \theta \leqslant x_{(1)}$, where

$$
c = \int_{0}^{x_{(1)}} \mathrm{e}^{(n-1)\theta}\,\mathrm{d}\theta
$$

$$
\begin{aligned}
&= \frac{1}{n-1} e^{(n-1)\theta}\Big|_0^{x_{(1)}} \\
&= \frac{1}{n-1}[e^{(n-1)x_{(1)}} - 1]. \qquad (3.12)
\end{aligned}
$$

Therefore, the Bayesian estimator of θ is given by

$$
\begin{aligned}
& E(\theta|x_1, \ldots, x_n) \\
= \;& c^{-1} \int_0^{x_{(1)}} \theta e^{(n-1)\theta} \, d\theta \\
= \;& c^{-1} \int_0^{x_{(1)}} \theta \, d\left[\frac{1}{n-1} e^{(n-1)\theta}\right] \\
= \;& c^{-1} \left[\frac{\theta}{n-1} e^{(n-1)\theta}\Big|_0^{x_{(1)}} - \int_0^{x_{(1)}} \frac{e^{(n-1)\theta}}{n-1} \, d\theta\right] \\
= \;& c^{-1} \left[\frac{x_{(1)} e^{(n-1)x_{(1)}}}{n-1} - \frac{c}{n-1}\right] \\
= \;& \frac{c^{-1} x_{(1)} e^{(n-1)x_{(1)}} - 1}{n-1},
\end{aligned}
$$

where c is defined by (3.12).

3.7 Solution. (a) Note that

$$
\begin{aligned}
E[t_1(X)] &= E(X) = 0 \cdot (1-\theta) + 1 \cdot \theta = \theta, \quad \text{and} \\
E[t_2(X)] &= E(1/2) = 1/2.
\end{aligned}
$$

Thus, $t_1(X)$ is unbiased estimator of θ and $t_2(X)$ is biased estimator of θ.

(b) Note that

$$
\begin{aligned}
\text{MSE}[t_1(X)] &= E(X-\theta)^2 = \text{Var}(X) = \theta(1-\theta), \quad \text{and} \\
\text{MSE}[t_2(X)] &= E(1/2-\theta)^2 = (1/2-\theta)^2.
\end{aligned}
$$

When $\frac{2-\sqrt{2}}{4} \leqslant \theta \leqslant \frac{2+\sqrt{2}}{4}$, we have

$$
\text{MSE}[t_1(X)] \geqslant \text{MSE}[t_2(X)].
$$

When $0 < \theta < \frac{2-\sqrt{2}}{4}$ or $\frac{2+\sqrt{2}}{4} < \theta < 1$, we have

$$\mathrm{MSE}[t_1(X)] < \mathrm{MSE}[t_2(X)].$$

3.8 **Solution**. (a) Let $Y_i = 1$ if the i-th respondent puts a tick in the triangle and $Y_i = 0$ if the i-th respondent puts a tick in the circle. Let y_i denote Y_i's realization for $i = 1, \ldots, n$. Then, we have

$$\begin{aligned} & \Pr\{Y_i = 1\} \\ = \; & \Pr\{\text{The } i\text{-th respondent puts a tick in the triangle}\} \\ = \; & \pi + (1-\pi)p \hat{=} \theta. \end{aligned}$$

Therefore, $\pi = (\theta - p)/(1-p)$. Since $Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} \text{Bernoulli}(\theta)$, then $Y \hat{=} \sum_{i=1}^n Y_i \sim \text{Binomial}(n, \theta)$. Thus, the MLE of θ is given by $\hat{\theta} = \frac{1}{n}Y$. By Theorem 3.1, the MLE of π is

$$\hat{\pi} = \begin{cases} \dfrac{\hat{\theta} - p}{1-p} = \dfrac{\frac{1}{n}Y - p}{1-p}, & \text{if } Y > np, \\ 0, & \text{if } Y \leqslant np. \end{cases}$$

(b) Since $\hat{\pi} = (Y/n - p)/(1-p) \cdot I_{(Y>np)}$, we have

$$E(\hat{\pi}) = \sum_{y>np} \frac{\frac{1}{n}y - p}{1-p} \cdot \binom{n}{y} \theta^y (1-\theta)^{n-y}.$$

3.9 **Solution**. Let $Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} \text{ZTB}(m, \pi)$ and the observed data be $Y_{\text{obs}} = \{y_1, \ldots, y_n\}$. Then, the observed-data likelihood function is given by

$$\begin{aligned} L(\pi | Y_{\text{obs}}) &= \prod_{i=1}^n \frac{\binom{m}{y_i} \pi^{y_i} (1-\pi)^{m-y_i}}{1 - (1-\pi)^m} \\ &\propto \pi^{n\bar{y}} (1-\pi)^{n(m-\bar{y})} \cdot [1 - (1-\pi)^m]^{-n}, \end{aligned}$$

where $\bar{y} = (1/n)\sum_{i=1}^{n} y_i$ is a sufficient statistic of π, and the log-likelihood function is

$$\begin{aligned}\ell(\pi|Y_{\text{obs}}) &= n\Big\{\bar{y}\log(\pi) + (m-\bar{y})\log(1-\pi) \\ &\quad - \log[1-(1-\pi)^m]\Big\}. \qquad (3.13)\end{aligned}$$

From (3.13), the first and second derivatives of the log-likelihood function are given by

$$\frac{\mathrm{d}\ell(\pi|Y_{\text{obs}})}{\mathrm{d}\pi} = n\left[\frac{\bar{y}}{\pi} - \frac{m-\bar{y}}{1-\pi} - \frac{m(1-\pi)^{m-1}}{1-(1-\pi)^m}\right] \quad \text{and}$$

$$\frac{\mathrm{d}^2\ell(\pi|Y_{\text{obs}})}{\mathrm{d}\pi^2} = n\left[-\frac{\bar{y}}{\pi^2} - \frac{m-\pi}{(1-\pi)^2} + \frac{m(1-\pi)^{m-2}\cdot A}{[1-(1-\pi)^m]^2}\right],$$

respectively, where

$$A = (m-1)[1-(1-\pi)^m] + m(1-\pi)^m.$$

Let $Y \sim \text{ZTB}(m, \pi)$, then $E(Y) = m\pi/[1-(1-\pi)^m] = E(\bar{Y})$. Thus, the Fisher information is

$$\begin{aligned} & J(\pi) \\ = \; & E\left[-\frac{\mathrm{d}^2\ell(\pi|Y_{\text{obs}})}{\mathrm{d}\pi^2}\right] \\ = \; & \frac{nm}{1-(1-\pi)^m}\left\{\frac{1}{\pi} + \frac{1-(1-\pi)^{m-1}}{1-\pi} - \frac{(1-\pi)^{m-2}\cdot A}{1-(1-\pi)^m}\right\}.\end{aligned}$$

Let $\pi^{(0)}$ be initial value of the MLE $\hat{\pi}$. If $\pi^{(t)}$ denotes the t-th approximation of $\hat{\pi}$, then, its $(t+1)$-th approximation can be obtained by the following Fisher scoring algorithm:

$$\pi^{(t+1)} = \pi^{(t)} + J^{-1}(\pi^{(t)})\frac{\mathrm{d}\ell(\pi^{(t)}|Y_{\text{obs}})}{\mathrm{d}\pi}$$

3.10 Solution. (a) The likelihood function is

$$L(\theta) = \left(\frac{1}{2\pi\theta}\right)^{n/2} \exp\left\{-\frac{\sum_{i=1}^{n}(x_i - \mu_0)^2}{2\theta}\right\}$$

so that

$$\ell(\theta) = \log L(\theta) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\theta) - \frac{\sum_{i=1}^{n}(x_i - \mu_0)^2}{2\theta}.$$

Therefore, the solution to

$$0 = \frac{\mathrm{d}\ell(\theta)}{\mathrm{d}\theta} = -\frac{n}{2\theta} + \frac{\sum_{i=1}^{n}(x_i - \mu_0)^2}{2\theta^2}$$

yields the MLE of θ, given by

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2.$$

(b) Note that the sample size is n, we then denote $\hat{\theta}$ by $\hat{\theta}_n$. From Example 3.19, we have $I(\theta) = 1/(2\theta^2)$. From (3.34) of Chapter 3 (page 151), we obtain

$$[nI(\theta)]^{1/2}(\hat{\theta}_n - \theta) = \sqrt{\frac{n}{2\theta^2}}\,(\hat{\theta}_n - \theta) \xrightarrow{\mathrm{L}} N(0, 1).$$

Hence

$$\sqrt{n}(\hat{\theta}_n - \theta) = \sqrt{2\theta^2} \cdot \sqrt{\frac{n}{2\theta^2}}\,(\hat{\theta}_n - \theta) \xrightarrow{\mathrm{L}} N(0, 2\theta^2).$$

3.11 Solution. (a) The joint density of $X_1, \ldots, X_n$ is

$$f(\boldsymbol{x}; \theta) = \left\{\prod_{i=1}^{n} g(x_i)\right\} \times h^{-n}(\theta) \prod_{i=1}^{n} I_{[a(\theta),\, b(\theta)]}(x_i). \qquad (3.14)$$

Note that

$$\begin{aligned}\prod_{i=1}^{n} I_{[a(\theta),\, b(\theta)]}(x_i) = 1 \quad &\Longleftrightarrow \quad a(\theta) \leqslant x_{(1)}, x_{(n)} \leqslant b(\theta) \\ &\Longleftrightarrow \quad \theta \leqslant \min\{a^{-1}(x_{(1)}),\ b^{-1}(x_{(n)})\}.\end{aligned}$$

Define $\tilde{\theta} = \min\{a^{-1}(x_{(1)}),\ b^{-1}(x_{(n)})\}$, we have

$$f(\boldsymbol{x};\theta) = \left\{ h^{-n}(\theta) \prod_{i=1}^{n} I_{[\theta,\,\infty)}(\tilde{\theta}) \right\} \times \prod_{i=1}^{n} g(x_i).$$

Thus $\hat{\theta} \hat{=} \min\{a^{-1}(X_{(1)}),\ b^{-1}(X_{(n)})\}$ is a sufficient statistic for θ.

(b) The joint density is still given by (3.14). Note that

$$\begin{aligned} \prod_{i=1}^{n} I_{[a(\theta),\, b(\theta)]}(x_i) = 1 \quad &\Longleftrightarrow \quad a(\theta) \leqslant x_{(1)}, x_{(n)} \leqslant b(\theta) \\ &\Longleftrightarrow \quad \theta \geqslant \max\{a^{-1}(x_{(1)}),\ b^{-1}(x_{(n)})\}. \end{aligned}$$

Define $\tilde{\theta} = \max\{a^{-1}(x_{(1)}),\ b^{-1}(x_{(n)})\}$, we have

$$f(\boldsymbol{x};\theta) = \left\{ h^{-n}(\theta) \prod_{i=1}^{n} I_{(-\infty,\,\theta]}(\tilde{\theta}) \right\} \times \prod_{i=1}^{n} g(x_i).$$

Thus $\hat{\theta} \hat{=} \max\{a^{-1}(X_{(1)}),\ b^{-1}(X_{(n)})\}$ is a sufficient statistic for θ.

(c) We only consider Case (a). The log-likelihood is

$$\ell(\theta) = -n \log h(\theta) + \sum_{i=1}^{n} \log g(x_i), \quad \theta \leqslant \tilde{\theta}.$$

Let $\theta_2 \geqslant \theta_1$. Since

$$\begin{aligned} h(\theta_2) - h(\theta_1) &= \int_{a(\theta_2)}^{b(\theta_2)} g(x)\,\mathrm{d}x - \int_{a(\theta_1)}^{b(\theta_1)} g(x)\,\mathrm{d}x \\ &= -\int_{a(\theta_1)}^{a(\theta_2)} g(x)\,\mathrm{d}x - \int_{b(\theta_2)}^{b(\theta_1)} g(x)\,\mathrm{d}x \\ &\leqslant 0, \\ &\Rightarrow \ell(\theta_2) \geqslant \ell(\theta_1), \end{aligned}$$

$\ell(\theta)$ is an increasing function of θ. Thus $\tilde{\theta}$ is the mle of θ and $\hat{\theta}$ is the MLE of θ.

3.12 Solution. (a) Since $Y \sim \text{Bernoulli}(\theta)$, we have $E(Y) = \theta$ and $E(Y^2) = \theta$. On the other hand, from $U \sim \text{Poisson}(\lambda)$, we obtain

$$E(U) = \lambda \quad \text{and} \quad E(U^2) = \text{Var}(U) + (EU)^2 = \lambda + \lambda^2.$$

Let

$$\begin{aligned}
\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i &= E(X) = E(Y) + E(U) = \theta + \lambda, \\
\Delta \hat{=} \frac{1}{n}\sum_{i=1}^{n} X_i^2 &= E(X^2) = E(Y^2) + E(U^2) + 2E(YU) \\
&= \theta + \lambda + \lambda^2 + 2\theta\lambda \\
&= (\theta + \lambda) + \lambda[\theta + (\theta + \lambda)],
\end{aligned}$$

we obtain the moment estimators as

$$\hat{\lambda}^{\mathrm{M}} = \frac{\Delta - \bar{X}}{\hat{\theta}^{\mathrm{M}} + \bar{X}} \quad \text{and} \quad \hat{\theta}^{\mathrm{M}} = \sqrt{\bar{X}(1 + \bar{X}) - \Delta}.$$

(b) We first find the distribution of $X = Y + U$. We consider two cases. If $x = 0$, then

$$\Pr(X = x) = \Pr(Y + U = 0) = \Pr(Y = 0, U = 0) = (1 - \theta)\mathrm{e}^{-\lambda}.$$

If $x \geqslant 1$, then

$$\begin{aligned}
\Pr(X = x) &= \Pr(Y + U = x) \\
&= \sum_{y=0}^{1} \Pr(Y = y, U = x - y) \\
&= \sum_{y=0}^{1} \theta^y (1 - \theta)^{1-y} \cdot \frac{\lambda^{x-y}}{(x-y)!}\mathrm{e}^{-\lambda} \\
&= (1 - \theta)\frac{\lambda^x}{x!}\mathrm{e}^{-\lambda} + \theta\frac{\lambda^{x-1}}{(x-1)!}\mathrm{e}^{-\lambda}.
\end{aligned}$$

Without loss of generality, we assume $X_i = 0$ for $i = 1, \ldots, m$ and $X_i \geqslant 1$ for $i = m+1, \ldots, n$. Thus, the likelihood function is

$$\begin{aligned}
& L(\theta, \lambda | x_1, \ldots, x_n) \\
= & \prod_{i=1}^{m} \Pr(X_i = 0) \times \prod_{i=m+1}^{n} \Pr(X_i = x_i) \\
= & (1-\theta)^m \mathrm{e}^{-m\lambda} \prod_{i=m+1}^{n} \left[(1-\theta) \frac{\lambda^{x_i}}{x_i!} \mathrm{e}^{-\lambda} + \theta \frac{\lambda^{x_i - 1}}{(x_i - 1)!} \mathrm{e}^{-\lambda} \right].
\end{aligned}$$

Define $\ell(\theta, \lambda) = \log L(\theta, \lambda | x_1, \ldots, x_n)$. Let

$$\frac{\partial \ell(\theta, \lambda)}{\partial \theta} = 0 \quad \text{and} \quad \frac{\partial \ell(\theta, \lambda)}{\partial \lambda} = 0,$$

we could obtain the mles of θ and λ. However, for the current situation, explicit solutions are not available. We need to use an iterative method such as Newton–Raphson algorithm.

3.13 Solution. (a) Example 3.24 showed that Y_1 is a sufficient statistic for θ. The cdf of X is

$$F(y; \theta) = \int_{\theta}^{y} f(x; \theta) \, \mathrm{d}x = 1 - \mathrm{e}^{-(y-\theta)}, \quad y \geqslant \theta.$$

Then, the density of Y_1 is

$$g_1(y) = n[1 - F(y; \theta)]^{n-1} f(y; \theta) = n\mathrm{e}^{-n(y-\theta)}, \quad y \geqslant \theta.$$

We can prove that Y_1 is also complete. According to Definition 3.9 on page 146, if

$$E[h(Y_1)] = 0 \quad \text{for all } \theta \in (-\infty, \infty),$$

then

$$E[h(Y_1)] = \int_{\theta}^{\infty} h(y) \cdot n\mathrm{e}^{-n(y-\theta)} \, \mathrm{d}y = 0.$$

This implies that

$$\int_\theta^\infty h(y)\mathrm{e}^{-ny}\,\mathrm{d}y = 0 \quad \text{for all} \;\; \theta \in (-\infty, \infty).$$

Differentiating both sides of the above identity with respect to θ yields

$$h(\theta)\mathrm{e}^{-n\theta} = 0,$$

i.e., $h(Y_1) = 0$ with probability 1. Therefore, Y_1 is complete.

(b) Now

$$E(Y_1) = \int_\theta^\infty y \cdot n\mathrm{e}^{-n(y-\theta)}\,\mathrm{d}y = \theta + 1/n.$$

Then $g(Y_1) = Y_1 - 1/n$ is an unbiased estimator of θ. Using Theorem 3.7, we know that $Y_1 - 1/n$ is the unique UMVUE of θ.

3.14 Solution. (a) The joint pmf of $X_1, \ldots, X_n$ is

$$f(x_1, \ldots, x_n; \theta) = \prod_{i=1}^n \frac{1}{\theta} \cdot I_{(1 \leqslant x_i \leqslant \theta)} = \left\{ \frac{1}{\theta^n} I_{(x_{(n)} \leqslant \theta)} \right\} \cdot I_{(x_{(1)} \geqslant 1)}.$$

By Theorem 3.5 (Factorization Theorem), $Y \mathrel{\hat{=}} X_{(n)}$ is sufficient for θ. The cdf of X is

$$F(m) = \Pr(X \leqslant m) = \sum_{x=1}^m \frac{1}{\theta} = \frac{m}{\theta}, \quad m = 1, 2, \ldots, \theta.$$

and the cdf of Y is

$$G_n(y) = \Pr(Y \leqslant y) = \Pr(X_{(n)} \leqslant y) = [F(y)]^n = [y/\theta]^n.$$

To prove that Y is also complete, we need to derive the pmf of Y:

$$\begin{aligned} g_n(y) &= \Pr(Y = y) \\ &= \Pr(Y \leqslant y) - \Pr(Y \leqslant y - 1) \\ &= G_n(y) - G_n(y-1) \\ &= \left(\frac{y}{\theta}\right)^n - \left(\frac{y-1}{\theta}\right)^n, \quad y = 1, 2, \ldots, \theta. \end{aligned}$$

If a function $h(y)$ satisfies

$$E[h(Y)] = 0 \quad \text{for all } \theta = 1, 2, \ldots$$

then

$$E[h(Y)] = \sum_{y=1}^{\theta} h(y) \, \frac{y^n - (y-1)^n}{\theta^n} = 0 \quad \text{for all } \theta = 1, 2, \ldots$$

If $\theta = 1$, then, we have $y = 1$ and

$$h(1) \cdot \frac{1^n - 0^n}{1^n} = 0,$$

i.e., $h(1) = 0$. If $\theta = 2$, then we have

$$h(1) \cdot \frac{1^n - 0^n}{1^n} + h(2) \cdot \frac{2^n - 1^n}{2^n} = 0,$$

i.e., $h(2) = 0$. By induction $h(y) = 0$ for $y = 1, 2, \ldots, \theta$. Therefore, Y is also complete.

(b) Let

$$g(Y) = \frac{Y^{n+1} - (Y-1)^{n+1}}{Y^n - (Y-1)^n},$$

then

$$\begin{aligned} E[g(Y)] &= \sum_{y=1}^{\theta} g(y) \cdot g_n(y) \\ &= \sum_{y=1}^{\theta} \frac{y^{n+1} - (y-1)^{n+1}}{y^n - (y-1)^n} \cdot \frac{y^n - (y-1)^n}{\theta^n} \\ &= \theta^{-n} \sum_{y=1}^{\theta} [y^{n+1} - (y-1)^{n+1}] \\ &= \theta. \end{aligned}$$

Then $g(Y)$ is an unbiased estimator of θ. By Theorem 3.7, we know that $g(Y)$ is the unique UMVUE of θ.

3.15 Solution. (a) The pmf of $X \sim$ Bernoulli(θ) is

$$f(x;\theta) = \theta^x (1-\theta)^{1-x}, \quad x = 0, 1,$$

so that $\log f(x;\theta) = x \log\theta + (1-x)\log(1-\theta)$ and $E(X) = \theta$. From Example 3.17, $I_n(\theta) = nI(\theta) = n/[\theta(1-\theta)]$. From Theorem 3.3 on page 129, we know the CR lower bound is given by

$$\frac{[\tau'(\theta)]^2}{I_n(\theta)} = \frac{\theta(1-\theta)(1-2\theta)^2}{n}.$$

(b) From Example 3.28, we know that $T = \sum_{i=1}^n X_i$ is a complete sufficient statistic for θ. Let $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$ denote the sample variance. Note that X_i only takes value 0 or 1, then

$$S^2 = \frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{n-1} = \frac{\sum_{i=1}^n X_i - n\bar{X}^2}{n-1} = \frac{T - T^2/n}{n-1} \hat{=} g(T)$$

is a function of T. Since

$$E(S^2) = \frac{\sum_{i=1}^n E(X_i) - nE(\bar{X}^2)}{n-1} = \frac{n\theta - n[\mathrm{Var}(\bar{X}) + (E\bar{X})^2]}{n-1}$$

$$= \frac{n\theta - n[\theta(1-\theta)/n + \theta^2]}{n-1} = \theta(1-\theta),$$

i.e., $S^2 = g(T)$ is an unbiased estimator of $\tau = \theta(1-\theta)$, According to Lehmann–Scheffe Theorem, S^2 is the unique UMVUE of $\tau(\theta)$.

3.16 Solution. (a) The joint pmf of (Y_1, Y_2) is given by

$$\begin{aligned}
& \Pr(Y_1 = y_1, Y_2 = y_2) \\
= \ & \Pr(X_0 + X_1 = y_1, X_0 + X_2 = y_2) \\
= \ & \sum_{k=0}^{\min(y_1,y_2)} \Pr(X_0 = k) \cdot \Pr(X_0 + X_1 = y_1, X_0 + X_2 = y_2 | X_0 = k)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{\min(y_1,y_2)} \Pr(X_0 = k)\cdot \Pr(X_1 = y_1 - k, X_2 = y_2 - k | X_0 = k) \\
&= \sum_{k=0}^{\min(y_1,y_2)} \Pr(X_0 = k)\cdot \Pr(X_1 = y_1 - k)\cdot \Pr(X_2 = y_2 - k) \\
&= \sum_{k=0}^{\min(y_1,y_2)} \frac{\lambda_0^k \mathrm{e}^{-\lambda_0}}{k!}\cdot \frac{\lambda_1^{y_1-k}\mathrm{e}^{-\lambda_1}}{(y_1-k)!}\cdot \frac{\lambda_2^{y_2-k}\mathrm{e}^{-\lambda_2}}{(y_2-k)!} \\
&= \mathrm{e}^{-\lambda_0-\lambda_1-\lambda_2} \sum_{k=0}^{\min(y_1,y_2)} \frac{\lambda_0^k \lambda_1^{y_1-k}\lambda_2^{y_2-k}}{k!(y_1-k)!(y_2-k)!}.
\end{aligned}
$$

(b) Since $\min(\boldsymbol{y}_j) = \min(y_{1j}, y_{2j}) = 0$ for all j, the likelihood function of $(\lambda_0, \lambda_1, \lambda_2)$ is given by

$$
\begin{aligned}
L(\lambda_0, \lambda_1, \lambda_2) &= \prod_{j=1}^{n} \mathrm{e}^{-\lambda_0-\lambda_1-\lambda_2} \sum_{k=0}^{\min(y_{1j},y_{2j})} \frac{\lambda_0^k \lambda_1^{y_{1j}-k}\lambda_2^{y_{2j}-k}}{k!(y_{1j}-k)!(y_{2j}-k)!} \\
&= \prod_{j=1}^{n} \mathrm{e}^{-\lambda_0-\lambda_1-\lambda_2} \frac{\lambda_1^{y_{1j}}\lambda_2^{y_{2j}}}{y_{1j}!y_{2j}!} \\
&\propto \lambda_1^{n\bar{y}_1}\lambda_2^{n\bar{y}_2}\mathrm{e}^{-n(\lambda_0+\lambda_1+\lambda_2)},
\end{aligned}
$$

where $\bar{y}_i = (1/n)\sum_{j=1}^n y_{ij}$ for $i = 1, 2$, so that the log-likelihood function is

$$\ell(\lambda_0, \lambda_1, \lambda_2) = n[\bar{y}_1 \log \lambda_1 + \bar{y}_2 \log \lambda_2 - \lambda_0 - \lambda_1 - \lambda_2].$$

Let $\partial\ell(\lambda_0, \lambda_1, \lambda_2)/\partial\lambda_i = 0$, then, the MLE of λ_i is

$$\hat{\lambda}_i = \bar{Y}_i = \frac{\sum_{j=1}^n Y_{ij}}{n}, \qquad i = 1, 2.$$

Given λ_1 and λ_2, since $\ell(\lambda_0, \lambda_1, \lambda_2)$ is a monotone decreasing function of λ_0, so the MLE of λ_0 is $\hat{\lambda}_0 = 0$.

3.17 Solution. (a) In the last row of Table 1.2 of Chapter 1, we set $r = 1$, then the negative binomial distribution reduces to the geometric distribution. Let X denote the population random variable of the geometric distribution, then $E(X) = 1/p$. Let the sample mean be equal to the population mean; i.e., $\bar{X} = E(X) = 1/p$, we obtain the moment estimator

$$\hat{p}^{\mathrm{M}} = \frac{1}{\bar{X}}.$$

(b) The joint density of $X_1, \ldots, X_n$ is

$$f(\boldsymbol{x}; p) = p^n (1-p)^{n\bar{x}-n}, \quad x_i = 1, 2, \ldots,$$

so that the log-likelihood function of p is

$$\ell(p) = n \log(p) + n(\bar{x} - 1) \log(1 - p).$$

Therefore, the MLE of p is $\hat{p} = 1/\bar{X}$.

(c) Since $p \sim U[0, 1]$, the posterior density of p is

$$p(p|\boldsymbol{x}) \propto p^n (1-p)^{n\bar{x}-n},$$

so that $p|\boldsymbol{x} \sim \text{Beta}(n + 1, n\bar{x} - n + 1)$. Therefore,

$$E(p|\boldsymbol{x}) = \frac{n+1}{n\bar{x}+2}$$

is the Bayesian estimate of p, and $(n+1)/(n\bar{X}+2)$ is the Bayesian estimator of p.