MA204: Mathematical Statistics

Suggested Solutions to Assignment 4

4.1 Proof. Define a new random variable $W = S_1^2/n_1 + S_2^2/n_2$. Since

$$\begin{aligned} W &= \frac{\sigma_1^2}{n_1(n_1-1)} \cdot \frac{(n_1-1)S_1^2}{\sigma_1^2} + \frac{\sigma_2^2}{n_2(n_2-1)} \cdot \frac{(n_2-1)S_2^2}{\sigma_2^2} \\ &\hat{=} a_1\chi_1^2 + a_2\chi_2^2 \end{aligned}$$

is a linear combination of two independent chi-squared random variables, where $\chi_k^2 \sim \chi^2(f_k),\ f_k = n_k - 1,\ k = 1, 2$, we could approximate W/g by a chi-squared distribution with f degrees of freedom, i.e.,

$$\frac{W}{g} \dot{\sim} \chi^2(f) \quad \text{or} \quad a_1\chi_1^2 + a_2\chi_2^2 \dot{\sim} g \cdot \chi^2(f). \tag{4.1}$$

To determine the g and f, let the corresponding means and variances in both sides of (4.1) be equal, i.e.,

$$a_1 f_1 + a_2 f_2 = gf \quad \text{and} \quad a_1^2 \cdot 2f_1 + a_2^2 \cdot 2f_2 = g^2 \cdot 2f. \tag{4.2}$$

We obtain

$$g = \frac{a_1^2 f_1 + a_2^2 f_2}{a_1 f_1 + a_2 f_2}$$

and

$$f = \frac{(a_1 f_1 + a_2 f_2)^2}{a_1^2 f_1 + a_2^2 f_2} = \frac{\left(\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}\right)^2}{\left(\dfrac{\sigma_1^2}{n_1}\right)^2 \dfrac{1}{n_1 - 1} + \left(\dfrac{\sigma_2^2}{n_2}\right)^2 \dfrac{1}{n_2 - 1}}. \tag{4.3}$$

From the definition of T_{Welch}, we have

$$\begin{aligned}
T_{\text{Welch}} &= \frac{(\bar{X}_1 - \bar{X}_2 - \mu_1 + \mu_2)/\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}{\sqrt{W}/\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \\
&= \frac{N(0,1)}{\sqrt{W/(a_1 f_1 + a_2 f_2)}} \\
&\overset{(4.2)}{=} \frac{N(0,1)}{\sqrt{\frac{W}{g}/f}} \\
&\doteq \frac{N(0,1)}{\sqrt{\chi^2(f)/f}} \\
&\sim t(f).
\end{aligned}$$

Finally, since f is a function of both σ_1^2 and σ_2^2, we replace σ_k^2 in (4.3) by S_k^2 $(k = 1, 2)$ and obtain the estimate of f, denoted by

$$\nu = \frac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\left(\dfrac{S_1^2}{n_1}\right)^2 \dfrac{1}{n_1 - 1} + \left(\dfrac{S_2^2}{n_2}\right)^2 \dfrac{1}{n_2 - 1}}.$$

4.2 Solution. (a) Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{Poisson}(\lambda)$. Note that $E(X) = \text{Var}(X) = \lambda$, by the Central Limit Theorem (Theorem 2.9),

$$\frac{\sqrt{n}(\bar{X}_n - \lambda)}{\sqrt{\lambda}} \xrightarrow{\text{L}} Z \sim N(0,1).$$

Therefore, for large n, we have

$$1 - \alpha = \Pr(|Z| \leqslant z_{\alpha/2}) = \Pr\left\{\left|\frac{\sqrt{n}(\bar{X}_n - \lambda)}{\sqrt{\lambda}}\right| \leqslant z_{\alpha/2}\right\}.$$

Now $n = 100$, $\bar{X}_n = 6.25$, $\alpha = 0.05$, $z_{0.025} = 1.96$, an approximate and equal-tail 95% CI of λ is determined by

$$\left|\frac{10(6.25 - \lambda)}{\sqrt{\lambda}}\right| \leqslant 1.96$$

or

$$\lambda^2 - 12.5384\lambda + 39.0625 \leqslant 0.$$

There are two roots

$$\lambda_1 = \frac{12.5384 - \sqrt{12.5384^2 - 4 \times 39.0625}}{2} = 5.7789$$

and

$$\lambda_2 = \frac{12.5384 + \sqrt{12.5384^2 - 4 \times 39.0625}}{2} = 6.7595.$$

Finally, an approximate and equal-tail 95% CI of α is given by $[5.7789, 6.7595]$.

(b) The shortest Wilson CI for the parameter λ in a Poisson distribution is constructed as follows. Suppose that we have n random samples $X_1, \ldots, X_n$ from Poisson(λ), and want to construct a $(1-\alpha)100\%$ CI for λ. According to the Central Limit Theorem, we have

$$\frac{\bar{X} - \lambda}{\sqrt{\lambda/n}} \xrightarrow{\text{L}} N(0,1), \qquad \text{as } n \to \infty.$$

Let $\alpha_1 + \alpha_2 = \alpha$ so that $\alpha_2 = \alpha - \alpha_1$. Approximately, we obtain

$$\Pr\left(-z_{\alpha_1} \leqslant \frac{\bar{X} - \lambda}{\sqrt{\lambda/n}} \leqslant z_{\alpha-\alpha_1}\right) = 1 - \alpha.$$

If $-z_{\alpha_1} \leqslant \frac{\bar{X}-\lambda}{\sqrt{\lambda/n}} \leqslant 0$, then $\lambda \geq \bar{X}$ and

$$\bar{X} + \frac{z_{\alpha_1}^2}{2n} - z_{\alpha_1}\sqrt{\frac{\bar{X}}{n} + \frac{z_{\alpha_1}^2}{4n^2}} \leqslant \lambda \leqslant \bar{X} + \frac{z_{\alpha_1}^2}{2n} + z_{\alpha_1}\sqrt{\frac{\bar{X}}{n} + \frac{z_{\alpha_1}^2}{4n^2}}.$$

Taking them together, we have

$$\bar{X} \leqslant \lambda \leqslant \bar{X} + \frac{z_{\alpha_1}^2}{2n} + z_{\alpha_1}\sqrt{\frac{\bar{X}}{n} + \frac{z_{\alpha_1}^2}{4n^2}}.$$

Similarly, if $0 \leqslant \frac{\bar{X}-\lambda}{\sqrt{\lambda/n}} \leqslant z_{\alpha-\alpha_1}$, we have

$$\bar{X} + \frac{z_{\alpha_1}^2}{2n} - z_{\alpha_1}\sqrt{\frac{\bar{X}}{n} + \frac{z_{\alpha_1}^2}{4n^2}} \leqslant \lambda \leqslant \bar{X}.$$

Thus, $-z_{\alpha_1} \leqslant \frac{\bar{X}-\lambda}{\sqrt{\lambda/n}} \leqslant z_{\alpha-\alpha_1}$ if and only if

$$\bar{X} + \frac{z_{\alpha-\alpha_1}^2}{2n} - z_{\alpha-\alpha_1}\sqrt{\frac{\bar{X}}{n} + \frac{z_{\alpha-\alpha_1}^2}{4n^2}} \leqslant \lambda \leqslant \bar{X} + \frac{z_{\alpha_1}^2}{2n} + z_{\alpha_1}\sqrt{\frac{\bar{X}}{n} + \frac{z_{\alpha_1}^2}{4n^2}}.$$

Therefore,

$$\left[\bar{X} + \frac{z_{\alpha-\alpha_1}^2}{2n} - z_{\alpha-\alpha_1}\sqrt{\frac{\bar{X}}{n} + \frac{z_{\alpha-\alpha_1}^2}{4n^2}}, \bar{X} + \frac{z_{\alpha_1}^2}{2n} + z_{\alpha_1}\sqrt{\frac{\bar{X}}{n} + \frac{z_{\alpha_1}^2}{4n^2}}\right]$$

is a $(1-\alpha)100\%$ CI for λ with length

$$l(\alpha_1) = \frac{z_{\alpha_1}^2 - z_{\alpha-\alpha_1}^2}{2n} + z_{\alpha-\alpha_1}\sqrt{\frac{\bar{X}}{n} + \frac{z_{\alpha-\alpha_1}^2}{4n^2}} + z_{\alpha_1}\sqrt{\frac{\bar{X}}{n} + \frac{z_{\alpha_1}^2}{4n^2}}.$$

The grid-point method can be used to find the shortest $l(\alpha_1)$ on $[0, \alpha]$. The corresponding R code is as follows.

```
function (x, alpha, error = 0.00001)
{
# Shortest.Wilson.CI.for.Poisson(x, alpha, error=0.00001)
# ---------------------------------------------------------
# Let X_1, ..., X_n ~iid Poisson(lambda),
# Aim: To find 100(1-alpha)% shortest Wilson CI for lambda,
# Input:
# x = a sequence of sample values,
# alpha = size, usually 0.05.
# error = increment of searching alpha_1, default is 0.00001
# Output:
# CI is a matrix.
```

```
# CI[1, ]: Lower & upper bounds, and the length of
#          the equal-tail CI (i.e. alpha1=alpha/2)
# CI[2, ]: Lower & upper bounds, and the shortest
#          length of the CI for lambda
# ----------------------------------------------------
n <- length(x)
xbar <- sum(x)/n
alpha1 <- seq(0, alpha, error)
z1 <- qnorm(alpha1)
z2 <- qnorm(1-alpha+alpha1)
LB <- xbar + z2^2/2/n - z2*sqrt(xbar/n+z2^2/(4*n*n))
UB <- xbar + z1^2/2/n - z1*sqrt(xbar/n+z1^2/(4*n*n))
length <- UB - LB
item <- order(length)[1]
length.alpha1 <- length(alpha1)
CI <- matrix(0, 3, 4)
CI[1, ] <- c(alpha1[length.alpha1/2+1], LB[length.alpha1/2+1],
UB[length.alpha1/2 + 1], length[length.alpha1/2 + 1])
CI[2, ] <- c(alpha1[item], LB[item], UB[item], length[item])
# -----------------------------------------------------------
Min <- 0
Max <- alpha
alpha_1 <- (Max + Min)/2
while(Max-Min > error){
    z1 <- qnorm(alpha_1)
    z2 <- qnorm(1-alpha+alpha_1)
    a1 <- (xbar/n+z1^2/4/n^2)^0.25
    a2 <- (xbar/n+z2^2/4/n^2)^0.25
    test <- exp(-z1*z1/2)/(a1-z1/(2*n*a1))^2
    test <- test - exp(-z2*z2/2)/(a2-z2/(2*n*a2))^2
```

```
    if(test<=0) Min <- alpha_1 else Max <- alpha_1
    alpha_1 <- (Max + Min)/2
}
z1 <- qnorm(alpha_1)
z2 <- qnorm(1-alpha+alpha_1)
L_B <- xbar + z2^2/2/n - z2*sqrt(xbar/n+z2^2/(4*n*n))
U_B <- xbar + z1^2/2/n - z1*sqrt(xbar/n+z1^2/(4*n*n))
CI[3, ] <- c(alpha_1, L_B, U_B, U_B - L_B)
dimnames(CI) <- list(c("Equal-tail CI: ",
"Shortest CI (Grid-Point): ", "Shortest CI (Bisection): "),
c("alpha1", "Lower.Bound","Upper.Bound", "UB.minus.LB" ))
return (CI)
}
```

4.3 Solution. (a) When $\sigma = \sigma_0$ is known, from (4.4) of Chapter 4, we know that

$$\left[\bar{X} - z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}},\ \bar{X} + z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}\right] = [-2.0925,\ 2.8425]$$

is a $100(1-\alpha)\%$ confidence interval for the mean μ, where $n = 4$, $\alpha = 0.1$, $z_{\alpha/2} = z_{0.05} = 1.645$, $\sigma_0 = 3$, and

$$\bar{X} = \frac{3.3 - 0.3 - 0.6 - 0.9}{4} = 0.375.$$

(b) When σ is unknown, from (4.6) of Chapter 4, we know that

$$\left[\bar{X} - t\left(\frac{\alpha}{2}, n-1\right)\frac{S}{\sqrt{n}},\ \bar{X} + t\left(\frac{\alpha}{2}, n-1\right)\frac{S}{\sqrt{n}}\right] = [-1.937, 2.687]$$

is a $100(1-\alpha)\%$ confidence interval for the mean μ, where $\bar{X} = 0.375$, $n = 4$, $t(\alpha/2, n-1) = t(0.05, 3) = 2.3534$, and

$$S = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}} = \sqrt{3.863} = 1.965.$$

4.4 Solution. Since σ^2 is unknown, from (4.6) of Chapter 4, we know that

$$\left[\bar{X} - t\left(\frac{\alpha}{2}, n-1\right)\frac{S}{\sqrt{n}},\ \bar{X} + t\left(\frac{\alpha}{2}, n-1\right)\frac{S}{\sqrt{n}}\right]$$
$$= \left[\bar{X} - t(0.05, n-1)\frac{S}{\sqrt{n}},\ \bar{X} + t(0.05, n-1)\frac{S}{\sqrt{n}}\right]$$

is a 90% CI for the mean μ. The length of the CI is

$$L = 2t(0.05, n-1)\frac{S}{\sqrt{n}}.$$

Then, we have

$$\begin{aligned} 0.95 &= \Pr(L \leqslant \sigma/5) \\ &= \Pr\left\{2t(0.05, n-1)\frac{S}{\sqrt{n}} \leqslant \frac{\sigma}{5}\right\} \\ &= \Pr\left\{4t^2(0.05, n-1)\frac{S^2}{n} \leqslant \frac{\sigma^2}{25}\right\} \\ &= \Pr\left\{\frac{(n-1)S^2}{\sigma^2} \leqslant \frac{n(n-1)}{100 \times t^2(0.05, n-1)}\right\} \\ &= \Pr\left\{\chi^2(n-1) \leqslant \frac{n(n-1)}{100 \times t^2(0.05, n-1)}\right\} \end{aligned}$$

or

$$\begin{aligned} 0.05 &= \Pr\left\{\chi^2(n-1) \geqslant \frac{n(n-1)}{100 \times t^2(0.05, n-1)}\right\} \\ &= \Pr\left\{\chi^2(n-1) \geqslant \chi^2(0.05, n-1)\right\}. \end{aligned}$$

Therefore, the sample size n should satisfy

$$\frac{n(n-1)}{100 \times t^2(0.05, n-1)} = \chi^2(0.05, n-1).$$

When $n = 309.228$, we obtain

$$\left|\frac{n(n-1)}{100 \times t^2(0.05, n-1)} - \chi^2(0.05, n-1)\right| \leqslant 0.00002.$$

Then, $n = 309$.

4.5 Solution. Because

$$\begin{pmatrix} A \\ B \end{pmatrix} \sim N_2\left(\begin{pmatrix} \mu_A \\ \mu_B \end{pmatrix}, \begin{pmatrix} \sigma_A^2 & \rho\sigma_A\sigma_B \\ \rho\sigma_A\sigma_B & \sigma_B^2 \end{pmatrix}\right),$$

we have

$$D \hat{=} A - B \sim N(\mu_A - \mu_B,\ \sigma^2),$$

where $\sigma^2 \hat{=} \sigma_A^2 + \sigma_B^2 - 2\rho\sigma_A\sigma_B$ is unknown. The objective is to find a 95% CI for $\mu_A - \mu_B$.

Now the random sample of D is: 6, 8, -2, 2, 7, 11, 1, 13. The sample mean $\bar{D} = 5.75$ and

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(D_i - \bar{D})^2} = \sqrt{26.2143} = 5.12.$$

Since σ^2 is unknown, from (4.6) of Chapter 4 (page 167), we know that

$$\begin{aligned}
& \left[\bar{D} - t\left(\frac{\alpha}{2}, n-1\right)\frac{S}{\sqrt{n}},\ \bar{D} + t\left(\frac{\alpha}{2}, n-1\right)\frac{S}{\sqrt{n}}\right] \\
= & \left[5.75 - t(0.025, 7)\frac{5.12}{\sqrt{8}},\ 5.75 + t(0.025, 7)\frac{5.12}{\sqrt{8}}\right] \\
= & \left[5.75 - 2.3646 \times \frac{5.12}{\sqrt{8}},\ 5.75 + 2.3646 \times \frac{5.12}{\sqrt{8}}\right] \\
= & [1.4696,\ 10.0304].
\end{aligned}$$

is a 95% CI for the difference $\mu_A - \mu_B$.

4.6 Solution. (a) When $f(x;\theta) = \theta x^{\theta-1} \cdot I_{(0,1)}(x)$, we have

$$F(x;\theta) = \int_0^x \theta t^{\theta-1}\,\mathrm{d}t = x^\theta, \qquad 0 < x < 1.$$

From (4.3) of Chapter 4, we have

$$-2\sum_{i=1}^{n}\log(X_i^\theta) = -2\theta\sum_{i=1}^{n}\log(X_i) \sim \chi^2(2n).$$

Thus, $-2\theta\sum_{i=1}^{n}\log(X_i)$ is a pivotal quantity. A $100(1-\alpha)\%$ equal-tail CI of θ can be constructed based on

$$\begin{aligned} & 1-\alpha \\ = \ & \Pr\left\{\chi^2(1-\alpha/2,2n) \leqslant -2\theta\sum_{i=1}^{n}\log(X_i) \leqslant \chi^2(\alpha/2,2n)\right\} \\ = \ & \Pr\left\{\frac{\chi^2(1-\alpha/2,2n)}{-2\sum_{i=1}^{n}\log(X_i)} \leqslant \theta \leqslant \frac{\chi^2(\alpha/2,2n)}{-2\sum_{i=1}^{n}\log(X_i)}\right\}. \end{aligned}$$

where $-2\sum_{i=1}^{n}\log(X_i) > 0$ since $0 < X_i < 1$ for $i = 1, \ldots, n$.

(b) Let $\alpha_1 + \alpha_2 = \alpha$ so that $\alpha_2 = \alpha - \alpha_1$. The $100(1-\alpha)\%$ shortest CI of θ can be constructed based on

$$\begin{aligned} 1-\alpha &= \Pr\left\{\chi^2(1-\alpha_2,2n) \leqslant -2\theta\sum_{i=1}^{n}\log(X_i) \leqslant \chi^2(\alpha_1,2n)\right\} \\ &= \Pr\left\{\frac{\chi^2(1-\alpha_2,2n)}{-2\sum_{i=1}^{n}\log(X_i)} \leqslant \theta \leqslant \frac{\chi^2(\alpha_1,2n)}{-2\sum_{i=1}^{n}\log(X_i)}\right\}. \end{aligned}$$

The width of this CI is

$$\begin{aligned} l(\alpha_1) &= \frac{\chi^2(\alpha_1,2n)}{-2\sum_{i=1}^{n}\log(X_i)} - \frac{\chi^2(1-\alpha_2,2n)}{-2\sum_{i=1}^{n}\log(X_i)} \\ &= \frac{\chi^2(\alpha_1,2n) - \chi^2(1-\alpha+\alpha_1,2n)}{-2\sum_{i=1}^{n}\log(X_i)} \end{aligned}$$

Thus, we can find α_1^* numerically such that

$$l(\alpha_1^*) = \min_{\alpha_1\in[0,\alpha]} l(\alpha_1) \quad \text{or} \quad \alpha_1^* = \arg\min_{\alpha_1\in[0,\alpha]} l(\alpha_1).$$

Therefore, The $100(1-\alpha)\%$ shortest CI of θ is

$$\left[\frac{\chi^2(1-\alpha+\alpha_1^*,2n)}{-2\sum_{i=1}^{n}\log(X_i)}, \frac{\chi^2(\alpha_1^*,2n)}{-2\sum_{i=1}^{n}\log(X_i)}\right].$$

4.7 Solution. (a) We know from Example 4.1 that $2\theta n\bar{X}$ is a pivotal quantity, and

$$\begin{aligned}[L_p, U_p] &= \left[\frac{\chi^2(1-\alpha/2, 2n)}{2n\bar{X}}, \frac{\chi^2(\alpha/2, 2n)}{2n\bar{X}}\right] \\ &= \left[\frac{9.591}{20 \times 55.087}, \frac{34.170}{20 \times 55.087}\right] = [0.00871, 0.03101]\end{aligned}$$

is an exact 95% equal-tail CI for θ.

(b) An exact 95% equal-tail CI for $1/\theta$ is

$$\left[\frac{2n\bar{X}}{\chi^2(\alpha/2, 2n)}, \frac{2n\bar{X}}{\chi^2(1-\alpha/2, 2n)}\right] = [32.24766, 114.8106].$$

This interval is obviously quite wide, reflecting substantial variability in breakdown times and a small sample size.

4.8 Solution. From (1.42) of Lecture Notes Chapter 1, we know that the mgf of $U \sim \text{Gamma}(\alpha, \beta)$ is

$$M_U(t) = \left(\frac{\beta}{\beta - t}\right)^{\alpha}, \quad t < \beta.$$

Since $\chi^2(1) = \text{Gamma}(1/2, 1/2)$, then the mgf of $V \sim \chi^2(1)$ is $M_V(t) = [1/(1-2t)]^{1/2}$. Thus, we only need to prove that the mgf of $Y \hat{=} \lambda(X-\mu)^2/(\mu^2 X)$ is

$$M_Y(t) = \sqrt{\frac{1}{1-2t}}, \quad t < 0.5.$$

(a) Let $\mu = \sqrt{a}$ and $\lambda = 2ab$, where $a > 0$ and $b > 0$, then we have

$$Y = \frac{2b}{X}(X^2 - 2\mu X + \mu^2) = -4b\sqrt{a} + 2b(X + a/X). \tag{4.4}$$

Using (3.15) and (3.16) in Suggested Solutions to Assignment 3, we obtain

$$\begin{aligned}
M_Y(t) &= E(\mathrm{e}^{tY}) \overset{(4.4)}{=} \mathrm{e}^{-4bt\sqrt{a}} E[\mathrm{e}^{2bt(X+a/X)}] \\
&\overset{(3.15)}{=} \mathrm{e}^{-4bt\sqrt{a}} \sqrt{\frac{ab}{\pi}} \mathrm{e}^{2b\sqrt{a}} \int_0^\infty \mathrm{e}^{2bt(x+a/x)} x^{-3/2} \mathrm{e}^{-b(x+a/x)}\, \mathrm{d}x \\
&= \mathrm{e}^{-4bt\sqrt{a}} \sqrt{\frac{ab}{\pi}} \mathrm{e}^{2b\sqrt{a}} \int_0^\infty x^{-3/2} \mathrm{e}^{-(b-2bt)(x+a/x)}\, \mathrm{d}x \\
&\overset{(3.16)}{=} \mathrm{e}^{-4bt\sqrt{a}} \sqrt{\frac{ab}{\pi}} \mathrm{e}^{2b\sqrt{a}} \cdot \sqrt{\frac{\pi}{a(b-2bt)}}\, \mathrm{e}^{-2(b-2bt)\sqrt{a}} \\
&= \sqrt{\frac{1}{1-2t}}, \quad t < 0.5.
\end{aligned}$$

(b) The formula (3.15) in Suggested Solutions to Assignment 3 says that the pdf of $X \sim \mathrm{IG}(\mu, \mu^3/\lambda)$ can be rewritten as

$$\sqrt{\frac{ab}{\pi}} \mathrm{e}^{2b\sqrt{a}} \cdot x^{-3/2} \mathrm{e}^{-b(x+a/x)}, \quad x > 0. \tag{4.5}$$

The joint density of $X_1, \dots, X_n \overset{\text{iid}}{\sim} \mathrm{IG}(\mu, \mu^3/\lambda)$ can be factorized into

$$\begin{aligned}
& f(x_1, \dots, x_n; a, b) \\
= & \prod_{i=1}^n \left\{ \sqrt{\frac{ab}{\pi}} \mathrm{e}^{2b\sqrt{a}} \cdot x_i^{-3/2} \mathrm{e}^{-b(x_i+a/x_i)} \right\} \\
= & \left(\frac{ab}{\pi}\right)^{n/2} \mathrm{e}^{2nb\sqrt{a}} \mathrm{e}^{-b(t_1+at_2)} \times (\textstyle\prod_{i=1}^n x_i)^{-3/2},
\end{aligned}$$

where $t_1 = \sum_{i=1}^n x_i$ and $t_2 = \sum_{i=1}^n x_i^{-1}$, so that (T_1, T_2) are jointly sufficient statistics of (a, b) or (μ, λ), where $T_1 = \sum_{i=1}^n X_i$ and $T_2 = \sum_{i=1}^n X_i^{-1}$.

(c) If $\lambda = \lambda_0$ is known, the joint density of $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{IG}(\mu, \mu^3/\lambda_0)$ can be factorized into

$$\begin{aligned}
& f(x_1, \ldots, x_n; \mu) \\
= & \prod_{i=1}^{n} \left\{ \sqrt{\frac{\lambda_0}{2\pi}} x_i^{-3/2} \exp\left[-\frac{\lambda}{2\mu^2 x_i} (x_i - \mu)^2 \right] \right\} \\
\propto & \ \mathrm{e}^{-\lambda_0 t_1/(2\mu^2) + n\lambda_0/\mu} \times (\textstyle\prod_{i=1}^{n} x_i)^{-3/2} \mathrm{e}^{-\lambda_0 t_2/2}.
\end{aligned}$$

Therefore, T_1 is sufficient for μ.

(d) From Q3.18(c), we obtain

$$T_1 = \sum_{i=1}^{n} X_i \sim \text{IG}(\mu^*, \mu^{*3}/\lambda^*) = \text{IG}(n\mu, n\mu^3/\lambda_0),$$

where $\mu^* = n\mu$ and $\lambda^* = n^2\lambda_0$. From Q4.8(a), we have

$$P \hat{=} \frac{\lambda^*(T_1 - \mu^*)^2}{\mu^{*2} T_1} = \frac{\lambda_0 (T_1 - n\mu)^2}{\mu^2 T_1} \sim \chi^2(1),$$

i.e., P is a pivotal quantity. Let $\alpha = 0.05$,

$$a_1 \hat{=} \chi^2(1 - \alpha/2, 1) - \frac{n^2 \lambda_0}{T_1} \quad \text{and} \quad b_1 \hat{=} \chi^2(\alpha/2, 1) - \frac{n^2\lambda_0}{T_1},$$

then the equal-tail 95% CI of μ is given by

$$\begin{aligned}
& 1 - \alpha \\
= & \Pr\left\{ \chi^2(1 - \alpha/2, 1) \leqslant \frac{\lambda_0 (T_1 - n\mu)^2}{\mu^2 T_1} \leqslant \chi^2(\alpha/2, 1) \right\} \\
= & \Pr\left\{ \chi^2(1 - \alpha/2, 1) \leqslant \frac{\lambda_0 (T_1 - 2n\mu)}{\mu^2} + \frac{n^2 \lambda_0}{T_1} \leqslant \chi^2(\alpha/2, 1) \right\} \\
= & \Pr\left\{ a_1 \leqslant \frac{\lambda_0 (T_1 - 2n\mu)}{\mu^2} \leqslant b_1 \right\} \\
= & \Pr(a_1 \mu^2 \leqslant \lambda_0 T_1 - 2n\lambda_0 \mu \leqslant b_1 \mu^2) \\
= & \Pr(L \leqslant \mu \leqslant U).
\end{aligned}$$

4.9 Solution. (a) Let $X \sim \mathrm{CS}(m, p, \lambda)$, then the pmf of X is

$$\begin{aligned}
\Pr(X = x) &= \Pr(X_1 + X_2 = x) \\
&= \sum_{k \geqslant 0} \Pr(X_1 = k) \cdot \Pr(X_1 + X_2 = x | X_1 = k) \\
&= \sum_{k \geqslant 0} \Pr(X_1 = k) \cdot \Pr(X_2 = x - k | X_1 = k) \\
&= \sum_{k=0}^{\min(m,x)} \Pr(X_1 = k) \cdot \Pr(X_2 = x - k) \\
&= \sum_{k=0}^{\min(m,x)} \binom{m}{k} p^k (1-p)^{m-k} \cdot \frac{\lambda^{x-k} \mathrm{e}^{-\lambda}}{(x-k)!},
\end{aligned}$$

for $x = 0, 1, \ldots, \infty$, where m is a known positive integer, $p \in (0, 1)$ and $\lambda > 0$.

(b) Since $X = X_1 + X_2$, we obtain

$$E(X) = E(X_1) + E(X_2) = mp + \lambda \hat{=} \mu$$

and

$$\mathrm{Var}(X) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) = mp(1-p) + \lambda = \mu - mp^2.$$

(c) Define $\bar{X}_n = (1/n) \sum_{i=1}^n X_i$, we have

$$E(\bar{X}_n) = E(X_1) = \mu \quad \text{and} \quad \mathrm{Var}(\bar{X}_n) = \frac{\mu - mp^2}{n}.$$

From the Central Limit Theorem, we obtain

$$\frac{\bar{X}_n - E(\bar{X}_n)}{\sqrt{\mathrm{Var}(\bar{X}_n)}} = \frac{\bar{X}_n - \mu}{\sqrt{(\mu - mp^2)/n}} \xrightarrow{\mathrm{D}} Z \sim N(0, 1), \quad \text{as } n \to \infty.$$

An approximate $100(1-\alpha)\%$ CI for the mean μ is given by

$$\begin{aligned}
1-\alpha &= \Pr\left\{\left|\frac{\bar{X}_n-\mu}{\sqrt{(\mu-mp^2)/n}}\right| \leqslant z_{\alpha/2}\right\} \\
&= \Pr\left\{\frac{n(\bar{X}_n-\mu)^2}{\mu-mp^2} \leqslant z^2_{\alpha/2}\right\} \\
&= \Pr\left\{\mu^2-(2\bar{X}_n+z_*)\mu+\bar{X}_n^2+mp^2z_* \leqslant 0\right\} \\
&= \Pr(L \leqslant \mu \leqslant U),
\end{aligned}$$

where $z_* \mathrel{\hat{=}} z^2_{\alpha/2}/n$,

$$\begin{aligned}
L &= \frac{2\bar{X}_n+z_*-\sqrt{z_*^2+4z_*(\bar{X}_n-mp^2)}}{2} \quad \text{and} \\
U &= \frac{2\bar{X}_n+z_*+\sqrt{z_*^2+4z_*(\bar{X}_n-mp^2)}}{2}.
\end{aligned}$$

4.10 Solution. (a) The cdf of X_1 with density

$$f(x;\mu)=\frac{1}{\sigma_0}\mathrm{e}^{-\frac{x-\mu}{\sigma_0}}\exp(-\mathrm{e}^{-\frac{x-\mu}{\sigma_0}}), \quad x\in\mathbb{R}, \quad \mu\in\mathbb{R}, \quad \sigma_0>0,$$

is given by

$$F(x;\mu)=\exp\left(-\mathrm{e}^{-\frac{x-\mu}{\sigma_0}}\right).$$

(b) From (4.3), we have

$$-2\sum_{i=1}^n \log F(X_i;\mu)=2\mathrm{e}^{\mu/\sigma_0}\sum_{i=1}^n \mathrm{e}^{-X_i/\sigma_0} \sim \chi^2(2n),$$

so that the $100(1-\alpha)\%$ equal-tail CI of μ is given by

$$\begin{aligned}
& 1-\alpha \\
= \;& \Pr\left\{\chi^2(1-\alpha/2;2n) \leqslant 2\mathrm{e}^{\mu/\sigma_0}\sum_{i=1}^n \mathrm{e}^{-x_i/\sigma_0} \leqslant \chi^2(\alpha/2;2n)\right\} \\
= \;& \Pr\left\{\frac{\chi^2(1-\alpha/2;2n)}{2\sum_{i=1}^n \mathrm{e}^{-x_i/\sigma_0}} \leqslant \mathrm{e}^{\mu/\sigma_0} \leqslant \frac{\chi^2(\alpha/2;2n)}{2\sum_{i=1}^n \mathrm{e}^{-x_i/\sigma_0}}\right\} \\
= \;& \Pr(L \leqslant \mu \leqslant U),
\end{aligned}$$

where

$$\begin{aligned}
L &= \sigma_0 \log\left[\frac{\chi^2(1-\alpha/2;2n)}{2\sum_{i=1}^n \mathrm{e}^{-x_i/\sigma_0}}\right] \quad \text{and} \\
U &= \sigma_0 \log\left[\frac{\chi^2(\alpha/2;2n)}{2\sum_{i=1}^n \mathrm{e}^{-x_i/\sigma_0}}\right].
\end{aligned}$$

4.11 Solution. The result in Q3.20(a) of Assignment 3 states that

$$Y_i \mathrel{\hat{=}} X_i^2 \overset{\text{iid}}{\sim} \text{Exponential}(\theta), \quad \text{where} \quad \theta = 1/(2\sigma^2).$$

From Example 4.1 of the textbook "Mathematical Statistics", we have

$$2\theta n\bar{Y} \sim \chi^2(2n), \quad \text{where} \quad \bar{Y} = \frac{\sum_{i=1}^n Y_i}{n} = \frac{\sum_{i=1}^n X_i^2}{n}.$$

Thus, by using the equal-tail method, we have

$$\begin{aligned}
1-\alpha &= \Pr\left\{\chi^2(1-\alpha/2,2n) \leqslant 2\theta n\bar{Y} \leqslant \chi^2(\alpha/2,2n)\right\} \\
&= \Pr\left\{\chi^2(1-\alpha/2,2n) \leqslant \frac{n\bar{Y}}{\sigma^2} \leqslant \chi^2(\alpha/2,2n)\right\} \\
&= \Pr\left\{\frac{n\bar{Y}}{\chi^2(\alpha/2,2n)} \leqslant \sigma^2 \leqslant \frac{n\bar{Y}}{\chi^2(1-\alpha/2,2n)}\right\},
\end{aligned}$$

that is,

$$[L_p, U_p] = \left[\frac{n\bar{Y}}{\chi^2(\alpha/2, 2n)}, \frac{n\bar{Y}}{\chi^2(1-\alpha/2, 2n)}\right]$$

is a $100(1-\alpha)\%$ CI for σ^2.

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

$$f(x_1, \ldots, x_n; \theta) = \prod_{i=1}^{n} \frac{1}{\theta} I(0 < x_i < \theta) = \theta^{-n} I(\theta > x_{(n)}) \times 1.$$

Factorization Theorem indicates that $X_{(n)}$ is a sufficient statistic of θ.

(b) From Example 2.16 of the textbook "Mathematical Statistic-s" on page 89, we know that the cdf of $X_{(n)}$ is

$$F_{X_{(n)}}(x) = 0 \cdot I(x \leqslant 0) + \frac{x^n}{\theta^n} \cdot I(0 < x < \theta) + 1 \cdot I(x \geqslant \theta).$$

Thus, the cdf of $U = X_{(n)}/\theta$ is

$$\begin{aligned} F_U(u) &= \Pr(U \leqslant u) = \Pr(X_{(n)} \leqslant u\theta) = F_{X_{(n)}}(u\theta) \\ &= 0 \cdot I(u \leqslant 0) + u^n \cdot I(0 < u < 1) + 1 \cdot I(u \geqslant 1), \end{aligned}$$

and the pdf of U is

$$f_U(u) = nu^{n-1}, \quad 0 < u < 1,$$

indicating that $U \sim \text{Beta}(n, 1)$. Since U is a function of both the sufficient statistic $X_{(n)}$ and the parameter θ, and the distribution of U does not depend on θ, U is a pivotal quantity.

(c) To find the upper $(\alpha/2)$-th quantile $h_{\alpha/2}$ of the distribution of U, we use

$$\Pr(U > h_{\alpha/2}) = \frac{\alpha}{2} \quad \text{or} \quad 1 - F_U(h_{\alpha/2}) = \frac{\alpha}{2},$$

to obtain

$$h_{\alpha/2} = (1-\alpha/2)^{1/n}. \tag{4.6}$$

Thus, $h_{1-\alpha/2} = (\alpha/2)^{1/n}$.

(d) The $100(1-\alpha)\%$ equal-tail CI for θ is

$$\begin{aligned} 1-\alpha &= \Pr(h_{1-\alpha/2} \leqslant U \leqslant h_{\alpha/2}) \qquad (4.7)\\ &= \Pr\left(h_{1-\alpha/2} \leqslant \frac{X_{(n)}}{\theta} \leqslant h_{\alpha/2}\right) \\ &= \Pr\left(\frac{X_{(n)}}{h_{\alpha/2}} \leqslant \theta \leqslant \frac{X_{(n)}}{h_{1-\alpha/2}}\right). \qquad (4.8) \end{aligned}$$

(e) From the given data, we have $n = 5$, $\max(x_1, \ldots, x_5) = 4.2$, $1-\alpha = 0.95$, $\alpha/2 = 0.025$, and the 95% equal-tail CI for θ is $[4.22, 8.78]$.

(f) Let h_{α_2} be the upper α_2-th quantile of the distribution of $U = X_{(n)}/\theta$, from (4.6), we obtain $h_{\alpha_2} = (1-\alpha_2)^{1/n}$.

Define $\alpha_1 = \alpha - \alpha_2$. Similarly, we have $h_{1-\alpha_1} = (\alpha_1)^{1/n} = (\alpha-\alpha_2)^{1/n}$. Similar to (4.7)–(4.8), the $100(1-\alpha)\%$ CI for θ is

$$\begin{aligned} 1-\alpha &= \Pr(h_{1-\alpha_1} \leqslant U \leqslant h_{\alpha_2}) \\ &= \Pr\left(h_{1-\alpha_1} \leqslant \frac{X_{(n)}}{\theta} \leqslant h_{\alpha_2}\right) \\ &= \Pr\left(\frac{X_{(n)}}{h_{\alpha_2}} \leqslant \theta \leqslant \frac{X_{(n)}}{h_{1-\alpha_1}}\right) \\ &= \Pr\left\{\frac{X_{(n)}}{(1-\alpha_2)^{1/n}} \leqslant \theta \leqslant \frac{X_{(n)}}{(\alpha-\alpha_2)^{1/n}}\right\}. \qquad (4.9) \end{aligned}$$

The width of this CI is

$$w(\alpha_2) = X_{(n)} \left[\frac{1}{(\alpha - \alpha_2)^{1/n}} - \frac{1}{(1 - \alpha_2)^{1/n}} \right],$$

which is an increasing function of α_2, since

$$\frac{\mathrm{d}w(\alpha_2)}{\mathrm{d}\alpha_2} = \frac{X_{(n)}}{n} \left[\frac{1}{(\alpha - \alpha_2)^{1+1/n}} - \frac{1}{(1 - \alpha_2)^{1+1/n}} \right] > 0. \quad (4.10)$$

Note that $0 \leqslant \alpha_2 \leqslant \alpha$, then $w(\alpha_2)$ will arrive it minimum at $\alpha_2 = 0$. From (4.9), thus, the shortest $100(1-\alpha)\%$ CI of θ is $[X_{(n)}, X_{(n)}\alpha^{-1/n}]$.

4.13 Solution. (a) Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} f(x; \lambda)$, where

$$f(x; \lambda) = \frac{\lambda}{\mathrm{e}^\lambda - 1} \mathrm{e}^{\lambda x}, \quad 0 \leqslant x \leqslant 1,\ \lambda > 0.$$

The likelihood function of λ is

$$L(\lambda) = \prod_{i=1}^{n} \frac{\lambda}{\mathrm{e}^\lambda - 1} \mathrm{e}^{\lambda x_i} = \frac{\lambda^n}{(\mathrm{e}^\lambda - 1)^n} \mathrm{e}^{\lambda n \bar{x}},$$

where $\bar{x} = (1/n)\sum_{i=1}^{n} x_i$. The log-likelihood function of λ is

$$\ell(\lambda) = n \log \lambda - n \log(\mathrm{e}^\lambda - 1) + \lambda n \bar{x},$$

so that

$$\ell'(\lambda) = \frac{n}{\lambda} - \frac{n\mathrm{e}^\lambda}{\mathrm{e}^\lambda - 1} + n\bar{x} \quad \text{and} \quad \ell''(\lambda) = -\frac{n}{\lambda^2} + \frac{n\mathrm{e}^\lambda}{(\mathrm{e}^\lambda - 1)^2}.$$

Newton's method to calculate the MLE of λ is to update

$$\lambda^{(t+1)} = \lambda^{(t)} - \frac{\ell'(\lambda^{(t)})}{\ell''(\lambda^{(t)})}, \quad t = 0, 1, 2, \ldots.$$

(b) The cdf of X is

$$F(x; \lambda) = \int_0^x \frac{\lambda \mathrm{e}^{\lambda t}}{\mathrm{e}^\lambda - 1}\, \mathrm{d}t = \frac{1}{\mathrm{e}^\lambda - 1} \mathrm{e}^{\lambda t} \Big|_0^x = \frac{\mathrm{e}^{\lambda x} - 1}{\mathrm{e}^\lambda - 1} \cdot I(0 \leqslant x \leqslant 1).$$

(c) From (4.3) of the Textbook, we have

$$-2\sum_{i=1}^{n}\log F(X_i;\lambda) = -2\sum_{i=1}^{n}\log\left(\frac{\mathrm{e}^{\lambda X_i}-1}{\mathrm{e}^{\lambda}-1}\right) \sim \chi^2(2n). \quad (4.11)$$

Define $\boldsymbol{X} = (X_1,\ldots,X_n)^\top$,

$$\begin{aligned} u_i(\lambda|X_i) &\triangleq \frac{\mathrm{e}^{\lambda X_i}-1}{\mathrm{e}^{\lambda}-1} \in (0,1), \quad i=1,\ldots,n \quad \text{and} \\ u(\lambda|\boldsymbol{X}) &\triangleq \sum_{i=1}^{n}\log[u_i(\lambda|X_i)]. \end{aligned} \quad (4.12)$$

Given all $\{X_i\}_{i=1}^n$, in the follows, we can show that $u(\lambda|\boldsymbol{X})$ is a monotonic decreasing function of λ or

$$u'(\lambda|\boldsymbol{X}) \leqslant 0. \quad (4.13)$$

From (4.11), we obtain

$$-2u(\lambda|\boldsymbol{X}) \overset{(4.12)}{=} -2\sum_{i=1}^{n}\log[u_i(\lambda|X_i)] \sim \chi^2(2n), \quad (4.14)$$

which can be used to construct $100(1-\alpha)\%$ CIs of λ as

$$\begin{aligned} & 1-\alpha \\ = \; & \Pr\left\{\chi^2(1-\alpha_1,2n) \leqslant -2u(\lambda|\boldsymbol{X}) \leqslant \chi^2(\alpha-\alpha_1,2n)\right\} \\ = \; & \Pr\left\{-\frac{1}{2}\chi^2(1-\alpha_1,2n) \geqslant u(\lambda|\boldsymbol{X}) \geqslant -\frac{1}{2}\chi^2(\alpha-\alpha_1,2n)\right\}, \end{aligned} \quad (4.15)$$

where $\alpha_1 \in (0,\alpha)$, $\chi^2(\alpha,2n)$ denotes the upper α-th quantile of the $\chi^2(2n)$ distribution satisfying

$$\Pr\{\chi^2(2n) > \chi^2(\alpha,2n)\} = \alpha.$$

From (4.15), it is clear that the explicit solutions of the CIs of λ cannot be calculated directly. Given α, denoting $100(1-\alpha)\%$ CIs of λ by $[\lambda_L(\alpha_1),\ \lambda_U(\alpha_1)]$, we have

$$\begin{aligned} 1-\alpha &= \Pr\{\lambda_L(\alpha_1) \leqslant \lambda \leqslant \lambda_U(\alpha_1)\} \\ &\overset{(4.13)}{=} \Pr\Big\{u(\lambda_L(\alpha_1)|\boldsymbol{X}) \geqslant u(\lambda|\boldsymbol{X}) \geqslant u(\lambda_U(\alpha_1)|\boldsymbol{X})\Big\}. \end{aligned} \tag{4.16}$$

By comparing (4.16) with (4.15), we obtain

$$\begin{aligned} u(\lambda_L(\alpha_1)|\boldsymbol{X}) + \frac{1}{2}\chi^2(1-\alpha_1, 2n) = 0 \quad \text{and} \\ u(\lambda_U(\alpha_1)|\boldsymbol{X}) + \frac{1}{2}\chi^2(\alpha-\alpha_1, 2n) = 0. \end{aligned} \tag{4.17}$$

Hence, given $\{\alpha, \alpha_1\}$, solving the lower–bound $\lambda_L(\alpha_1)$ and the upper–bound $\lambda_U(\alpha_1)$ can be reduced to finding the roots of the non-linear equation

$$u(\lambda|\boldsymbol{X}) + c_1 = 0, \tag{4.18}$$

corresponding to $c_1 = 0.5\chi^2(1-\alpha_1, 2n)$ and $c_1 = 0.5\chi^2(\alpha-\alpha_1, 2n)$. In particular, let $\alpha_1 = \alpha/2$.

Proof of (4.13). If $X_i = 0$, then $u_i(\lambda|X_i) = 0$ so that $\log[u_i(\lambda|X_i)]$ and $u(\lambda|\boldsymbol{X})$ have no definition. If $\sum_{i=1}^n X_i = n$ (i.e., all $X_i = 1$), then $u_i(\lambda|X_i) = 1$ so that $\log[u_i(\lambda|X_i)] = 0$ and $u(\lambda|\boldsymbol{X}) = 0$.

First we assume that each X_i belongs to $(0,1)$. We only need to show that $u_i'(\lambda|X_i) \triangleq \mathrm{d}u_i(\lambda|X_i)/\,\mathrm{d}\lambda < 0$ for all $i = 1, \ldots, n$. In fact,

$$\begin{aligned} u_i'(\lambda|X_i) < 0 \quad &\Leftrightarrow \quad X_i \mathrm{e}^{\lambda X_i}(\mathrm{e}^{\lambda}-1) - \mathrm{e}^{\lambda}(\mathrm{e}^{\lambda X_i}-1) < 0 \\ &\Leftrightarrow \quad (1-X_i)(\mathrm{e}^{\lambda}-1) > \mathrm{e}^{\lambda(1-X_i)} - 1 \\ &\Leftrightarrow \quad z\left(\mathrm{e}^{\lambda}-1\right) > \mathrm{e}^{\lambda z} - 1, \end{aligned}$$

where $z = 1 - X_i \in (0, 1)$. By applying the Taylor expansion of $\exp(\cdot)$ around 0, we have

$$\begin{aligned} z(\mathrm{e}^{\lambda} - 1) &= \lambda z + \frac{1}{2}\lambda^2 z + \frac{1}{6}\lambda^3 z + \cdots \quad \text{and} \\ \mathrm{e}^{\lambda z} - 1 &= \lambda z + \frac{1}{2}\lambda^2 z^2 + \frac{1}{6}\lambda^3 z^3 + \cdots . \end{aligned}$$

Note that $z > z^r$ for any positive integer r, we obtain $z(\mathrm{e}^{\lambda} - 1) > \mathrm{e}^{\lambda z} - 1$.

Next, if $X_i = 1$, then $u_i'(\lambda|X_i) = 0$; however, as long as $\sum_{i=1}^{n} X_i \neq n$, then $u'(\lambda) < 0$ is still true. Thus $u(\lambda|\boldsymbol{X})$ is a monotonic decreasing function of λ. $\square$