Homework 8

Solutions!

Note: For problems where the $1 - \alpha$ confidence level is specified (i.e. 95% confidence level), you are to find the appropriate percentiles. You can look them up in the correct table in appendix 3 of the book. You may laso be able to find these tables online. In some cases, there may not be a table, because you are able to determine the percentile by evaluating an integral in the same way that the book does in example 8.4 and example 8.5.

1. Problem 8.39 from the book (p 409)
   Solution:

   Using Table 6 with 4 degrees of freedom, $P\left(.71072 \le 2Y/\beta \le 9.48773\right) = .90$. So,
   $$P\left(\tfrac{2Y}{9.48773} \le \beta \le \tfrac{2Y}{.71072}\right) = .90$$
   and $\left(\tfrac{2Y}{9.48773}, \tfrac{2Y}{.71072}\right)$ forms a 90% CI for β.

2. Problem 8.40 from the book (p 409)
   Solution:
   Use the fact that $Z = \frac{Y-\mu}{\sigma}$ has a standard normal distribution. With $\sigma = 1$:
   a. The 95% CI is $(Y - 1.96, Y + 1.96)$ since
   $$P(-1.96 \le Y - \mu \le 1.96) = P(Y - 1.96 \le \mu \le Y + 1.96) = .95.$$
   b. The value $Y + 1.645$ is the 95% upper limit for μ since
   $$P(Y - \mu \le 1.645) = P(\mu \le Y + 1.645) = .95.$$
   c. Similarly, $Y - 1.645$ is the 95% lower limit for μ.

3. Problem 8.41 from the book (p 409)
   Solution:
   Using Table 6 with 1 degree of freedom:
   a. $.95 = P(.0009821 \le Y^2/\sigma^2 \le 5.02389) = P(Y^2/5.02389 \le \sigma^2 \le Y^2/.0009821)$.
   b. $.95 = P(.0039321 \le Y^2/\sigma^2) = P(\sigma^2 \le Y^2/.0039321)$.
   c. $.95 = P(Y^2/\sigma^2 \le 3.84146) = P(Y^2/3.84146 \le \sigma^2)$.

4. Problem 8.42 from the book (p 410)
   Solution:

Using the results from Ex. 8.41, the square–roots of the boundaries can be taken to obtain interval estimates σ:

a. $Y/2.24 \le \sigma \le Y/.0313$.
b. $\sigma \le Y/.0627$.
c. $\sigma \ge Y/1.96$.

5. Problem 8.43 from the book (p 410)
Solution:

a. The distribution function for $Y_{(n)}$ is $G_n(y) = \left(\frac{y}{\theta}\right)^n$, $0 \le y \le \theta$, so the distribution function for U is given by

$$F_U(u) = P(U \le u) = P(Y_{(n)} \le \theta u) = G_n(\theta u) = u,\ 0 \le y \le 1.$$

b. (Similar to Example 8.5) We require the value a such that $P\left(\frac{Y_{(n)}}{\theta} \le a\right) = F_U(a) = .95$.

Therefore, $a^n = .95$ so that $a = (.95)^{1/n}$ and the lower confidence bound is $[Y_{(n)}](.95)^{-1/n}$.

6. Problem 8.44 from the book (p 410)
Solution:

a. $F_Y(y) = P(Y \le y) = \int_0^y \frac{2(\theta - t)}{\theta^2} dt = \frac{2y}{\theta} - \frac{y^2}{\theta^2}$, $0 < y < \theta$.

b. The distribution of $U = Y/\theta$ is given by
$F_U(u) = P(U \le u) = P(Y \le \theta u) = F_Y(\theta u) = 2u - u^2 = 2u(1-u)$, $0 < u < 1$. Since this distribution does not depend on θ, $U = Y/\theta$ is a pivotal quantity.

c. Set $P(U \le a) = F_Y(a) = 2a(1 - a) = .9$ so that the quadratic expression is solved at $a = 1 - \sqrt{.10} = .6838$ and then the 90% lower bound for θ is $Y/.6838$.

7. Problem 8.46 from the book (p 410)
Solution:
Let $U = 2Y/\theta$ and let $m_Y(t)$ denote the mgf for the exponential distribution with mean θ. Then:

a. $m_U(t) = E(e^{tU}) = E(e^{t2Y/\theta}) = m_Y(2t/\theta) = (1-2t)^{-1}$. This is the mgf for the chi–square distribution with one degree of freedom. Thus, U has this distribution, and since the distribution does not depend on θ, U is a pivotal quantity.

b. Using Table 6 with 2 degrees of freedom, we have
$$P(.102587 \le 2Y/\theta \le 5.99147) = .90.$$
So, $\left(\frac{2Y}{5.99147}, \frac{2Y}{.102587}\right)$ represents a 90% CI for θ.

c. They are equivalent.

8. Problem 8.47 from the book (p 410)

Solution:

Note that for all i, the mgf for Y_i is $m_Y(t) = (1-\theta t)^{-1}$, $t < 1/\theta$.

a. Let $U = 2\sum_{i=1}^{n} Y_i / \theta$. The mgf for U is

$$m_U(t) = E(e^{tU}) = [m_Y(2t/\theta)]^n = (1-2t)^{-n}, \; t < 1/2 .$$

This is the mgf for the chi–square distribution with $2n$ degrees of freedom. Thus, U has this distribution, and since the distribution does not depend on θ, U is a pivotal quantity.

b. Similar to part b in Ex. 8.46, let $\chi^2_{.975}$, $\chi^2_{.025}$ be percentage points from the chi–square distribution with $2n$ degrees of freedom such that

$$P\left(\chi^2_{.975} \le 2\sum_{i=1}^{n} Y_i / \theta \le \chi^2_{.025}\right) = .95 .$$

So, $\left(\dfrac{2\sum_{i=1}^{n} Y_i}{\chi^2_{.975}}, \dfrac{2\sum_{i=1}^{n} Y_i}{\chi^2_{.025}} \right)$ represents a 95% CI for θ.

c. The CI is $\left(\dfrac{2(7)(4.77)}{26.1190}, \dfrac{2(7)(4.77)}{5.62872} \right)$ or (2.557, 11.864).

9. Problem 8.58 from the book (p 417)
   Solution:

   The parameter of interest is μ = mean number of days required for treatment. The 95% CI is approximately $\bar{y} \pm z_{.025}\left(s/\sqrt{n}\right)$, or $5.4 \pm 1.96(3.1/\sqrt{500})$ or (5.13, 5.67).

10. Problem 8.59 from the book (p 418)
    Solution:

    a. With $z_{.05} = 1.645$, the 90% interval is $.78 \pm 1.645\sqrt{\frac{.78(.22)}{1030}}$ or $.78 \pm .021$.

    b. The lower endpoint of the interval is $.78 - .021 = .759$, so there is evidence that the true proportion is greater than 75%.

11. Let $X_1, X_2, ..., X_n$ be i.i.d.$Exp(\delta)$
    a) Construct a pivot for a CI for δ, based on $X_1, ..., X_n$ and identify the distribution of your pivot.

Solution 1:

$$\begin{aligned} X_i &\sim Exp(\delta) \\ \Rightarrow \frac{X_i}{\delta} &\sim Exp(1) \\ \Rightarrow \frac{\sum_{i=1}^n X_i}{\delta} = \sum_{i=1}^n \frac{X_i}{\delta} &\sim \Gamma(n,1) \end{aligned}$$

Solution 2:

$$\begin{aligned} \frac{\sum_{i=1}^n X_i}{\delta} = \sum_{i=1}^n \frac{X_i}{\delta} &\sim \Gamma(n,1) \\ \Rightarrow \frac{\bar{X}}{\delta} &\sim \Gamma(n, \frac{1}{n}) \end{aligned}$$

b) Construct a one-sided lower bound $1-\alpha$ CI for δ that is based on the MLE or δ.
solution:
Let g_α be the $1-\alpha$ percentile of $\Gamma(n,1)$, and let $\overline{X} = \sum_{i=1}^n X_i/n$, which is the MLE for δ.

$$\begin{aligned} X_1, \ldots, X_n &\quad iid \quad Exp(\delta) \\ \Rightarrow \sum_{i=1}^n X_i &\sim \Gamma(n,\delta) \\ \Rightarrow \overline{X} &\sim \Gamma(n, \delta/n) \\ \Rightarrow n\overline{X}/\delta &\sim \Gamma(n,1) \\ \Rightarrow P(\frac{n}{\delta}\overline{X} < g_\alpha) &= 1-\alpha \\ \Rightarrow P(\frac{n}{g_\alpha}\overline{X} < \delta) &= 1-\alpha \end{aligned}$$

So, our $1-\alpha$ lower bond CI for δ is $(\frac{n}{g_\alpha}\overline{X}, \infty)$

c) Assuming that n is sufficiently large,Construct an approximate one-sided upper bound $1-\alpha$ CI for δ using $X_1, ..., X_n$ and an appropriate estimator for $V[X_1]$.
Solution:
Let Z_α is the $1-\alpha$ percentile of the $N(0,1)$ distribution. We know that the MOM estimator and the MLE are both $\hat{\delta} = \overline{X} = \sum_{i=1}^n X_i/n$, so by the invariance property of the MLE, we know that the MLE for $V[X_1] = \delta^2$ (which is an appropriate estimator to use when constructing an approximate CI) is $\overline{X}^2$. For this solution we will simply use the MLE estimator, but the sample variance $S^2 = \sum_{i=1}^n (X_i - \overline{X})^2/(n-1)$ is another appropriate estimator for $V[X_1]$ when constructing an approximate CI.

$$\sqrt{n}\frac{\overline{X}-\delta}{\delta} \sim N(0,1) \text{ (approximately, by CLT)}$$

$$
\begin{aligned}
\overline{X} \text{ is the MLE for } \delta & \\
\Rightarrow \sqrt{n}\frac{\overline{X}-\delta}{\overline{X}} &\sim N(0,1) \text{ (approximately)} \\
\Rightarrow P(-Z_\alpha < \frac{\overline{X}-\delta}{\sqrt{\overline{X}^2/n}}) &\approx 1-\alpha \\
\Rightarrow P(-Z_\alpha \overline{X}/\sqrt{n} < \overline{X}-\delta) &\approx 1-\alpha \\
\Rightarrow P(\delta < \overline{X} + Z_\alpha \overline{X}/\sqrt{n}) &\approx 1-\alpha
\end{aligned}
$$

So, our approximate upper bound $1-\alpha$ CI is $(-\infty, \overline{X} + Z_\alpha \overline{X}/\sqrt{n})$ (we can also turn this into $(0, \sqrt{n}\frac{\bar{X}}{-Z_\alpha + \sqrt{n}})$ since $\delta > 0$, by definition).

Challenge Question:

Let $X_1, ...X_{n_1}$ be i.i.d. $N(\mu, \sigma_x^2)$ and let $Y_1, ..., Y_{n_2}$ be i.i.d. $N(\mu, \sigma_y^2)$ where μ, σ_x, & σ_y are unknown, and assume that all of the $X's$ and all of the $Y's$ are independent from each other. Construct a $1-\alpha$ two-sided confidence interval for $\frac{\sigma_x}{\sigma_y}$ that is a function of all of the $X's$ and all of the $Y's$.