Homework 11

Mengxiang Jiang

Stat 610 Distribution Theory

November 24, 2025

Problem 1. Suppose $X_i \overset{iid}{\sim}$ exponential$(\beta), i = 1, 2, \ldots$.

(a) Identify (with all parameters) the asymptotic bivariate normal distribution of the first two sample moments, $\hat{\mu}_1'$ and $\hat{\mu}_2'$. Recall Corollary 5.22 in the notes.

(b) Use the bivariate delta method to get a limit distribution for $\sqrt{n}\left(\hat{\mu}_1' - \beta, \sqrt{\hat{\mu}_2'/2} - \beta\right)$.

(c) Of the two potential estimators for β, $\hat{\mu}_1'$ and $\sqrt{\hat{\mu}_2'/2}$, which has the smaller asymptotic variance? Note: it is not possible to get the exact variance of $\sqrt{\hat{\mu}_2'/2}$. [Such a question is relevant for statisticians wanting to find the most efficient estimator of β.]

(a) We have $\mu_1' = \beta$ and $\mu_2' = 2\beta^2$. The variance and covariance terms are

$$\sigma_{11} = \mathsf{Var}(X) = \beta^2,$$

$$\sigma_{22} = \mathsf{Var}(X^2) = \mathsf{E}(X^4) - (\mathsf{E}(X^2))^2 = 4!\beta^4 - (2!\beta^2)^2 = 20\beta^4,$$

$$\sigma_{12} = \sigma_{21} = \mathsf{Cov}(X, X^2) = \mathsf{E}(X^3) - \mathsf{E}(X)\mathsf{E}(X^2) = 3!\beta^3 - 2!\beta^3 = 4\beta^3.$$

Therefore, by Corollary 5.22, we have

$$\sqrt{n}\begin{pmatrix} \hat{\mu}_1' - \beta \\ \hat{\mu}_2' - 2\beta^2 \end{pmatrix} \xrightarrow{D} \text{normal}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \beta^2 & 4\beta^3 \\ 4\beta^3 & 20\beta^4 \end{pmatrix}\right).$$

(b) Let

$$g(x, y) = \begin{pmatrix} x \\ \sqrt{y/2} \end{pmatrix}.$$

Then we have

$$g'(x, y) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2\sqrt{2}}y^{-1/2} \end{pmatrix}$$

$$g'(\beta, 2\beta^2) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2\sqrt{2}}(2\beta^2)^{-1/2} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{4\beta} \end{pmatrix}.$$

1

By the bivariate delta method, we have the asymptotic covariance matrix

$$\Sigma^{**} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{4\beta} \end{pmatrix} \begin{pmatrix} \beta^2 & 4\beta^3 \\ 4\beta^3 & 20\beta^4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{4\beta} \end{pmatrix}$$

$$= \begin{pmatrix} \beta^2 & \beta^2 \\ \beta^2 & \frac{5\beta^2}{4} \end{pmatrix}.$$

Therefore,

$$\sqrt{n}\left(\begin{matrix} \hat{\mu}_1' - \beta \\ \sqrt{\hat{\mu}_2'/2} - \beta \end{matrix} \right) \xrightarrow{D} \text{normal} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \beta^2 & \beta^2 \\ \beta^2 & \frac{5\beta^2}{4} \end{pmatrix} \right).$$

(c) The asymptotic variance of $\hat{\mu}_1'$ is β^2, and the asymptotic variance of $\sqrt{\hat{\mu}_2'/2}$ is $\frac{5\beta^2}{4}$. Since $\beta^2 < \frac{5\beta^2}{4}$, $\hat{\mu}_1'$ has the smaller asymptotic variance.

Problem 2. Suppose $T_i \overset{iid}{\sim} \text{Poisson}(\lambda), i = 1, 2, \ldots$. Note that $\mu_3' = \mathsf{E}(T_i^3) = \lambda^3 + 3\lambda^2 + \lambda$ and $\mu_4' = \mathsf{E}(T_i^4) = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda$.

(a) Identify the asymptotic joint distribution of the sample mean $\overline{T}_n$ and the sample variance $\hat{\sigma}_n^2$. Recall Theorem 5.23 in the notes.

(b) Both $\overline{T}_n$ and $\hat{\sigma}_n^2$ are unbiased for λ, that is, their expectations equal λ exactly. Show that $a\overline{T}_n + (1 - a)\hat{\sigma}_n^2$ is also unbiased for λ, for any $a \in [0, 1]$.

(c) What is $\mathsf{Var}(a\overline{T}_n + (1 - a)\hat{\sigma}_n^2)$, and what value of a minimizes this variance? [Again, the point here is to identify what is best for estimation purposes.]

(a) We have $\mu_1' = \lambda$, $\mu_2' = \lambda^2 + \lambda$, $\mu_3' = \lambda^3 + 3\lambda^2 + \lambda$, and $\mu_4' = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda$. The variance and covariance terms are

$$\sigma_{11} = \mathsf{Var}(T_i) = \lambda,$$

$$\begin{aligned} \sigma_{22} = \mathsf{Var}(T_i^2) &= \mathsf{E}(T_i^4) - (\mathsf{E}(T_i^2))^2 \\ &= (\lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda) - (\lambda^2 + \lambda)^2 \\ &= 4\lambda^3 + 6\lambda^2 + \lambda, \end{aligned}$$

$$\begin{aligned} \sigma_{12} = \sigma_{21} = \mathsf{Cov}(T_i, T_i^2) &= \mathsf{E}(T_i^3) - \mathsf{E}(T_i)\mathsf{E}(T_i^2) \\ &= (\lambda^3 + 3\lambda^2 + \lambda) - \lambda(\lambda^2 + \lambda) \\ &= 2\lambda^2 + \lambda. \end{aligned}$$

Therefore, by Theorem 5.23, we have

$$\sqrt{n}\left(\begin{matrix} \overline{T}_n - \lambda \\ \hat{\sigma}_n^2 - \lambda \end{matrix} \right) \xrightarrow{D} \text{normal} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma^* \right),$$

2

where

$$\Sigma^* = \begin{pmatrix} \sigma_{11} & -2\mu\sigma_{11} + \sigma_{12} \\ -2\mu\sigma_{11} + \sigma_{12} & 4\mu^2\sigma_{11} - 4\mu\sigma_{12} + \sigma_{22} \end{pmatrix}$$
$$= \begin{pmatrix} \lambda & -2\lambda^2 + (2\lambda^2 + \lambda) \\ -2\lambda^2 + (2\lambda^2 + \lambda) & 4\lambda^3 - 4\lambda(2\lambda^2 + \lambda) + (4\lambda^3 + 6\lambda^2 + \lambda) \end{pmatrix}$$
$$= \begin{pmatrix} \lambda & \lambda \\ \lambda & 2\lambda^2 + \lambda \end{pmatrix}.$$

(b) Since both $\overline{T}_n$ and $\hat{\sigma}_n^2$ are unbiased for λ, we have

$$\mathsf{E}(a\overline{T}_n + (1-a)\hat{\sigma}_n^2) = a\mathsf{E}(\overline{T}_n) + (1-a)\mathsf{E}(\hat{\sigma}_n^2)$$
$$= a\lambda + (1-a)\lambda$$
$$= \lambda.$$

(c) We have

$$\mathsf{Var}(a\overline{T}_n + (1-a)\hat{\sigma}_n^2) = a^2\mathsf{Var}(\overline{T}_n) + (1-a)^2\mathsf{Var}(\hat{\sigma}_n^2)$$
$$+ 2a(1-a)\mathsf{Cov}(\overline{T}_n, \hat{\sigma}_n^2)$$
$$= a^2\frac{\lambda}{n} + (1-a)^2\frac{2\lambda^2 + \lambda}{n} + 2a(1-a)\frac{\lambda}{n}$$
$$= \frac{\lambda}{n}\left(a^2 + (1-a)^2(2\lambda+1) + 2a(1-a)\right)$$
$$= \frac{\lambda}{n}\left(a^2 + (1-a)^2(2\lambda+1) + 2a - 2a^2\right)$$
$$= \frac{\lambda}{n}\left((1-a)^2(2\lambda+1) + 2a - a^2\right)$$
$$= \frac{\lambda}{n}\left((1-a)^2(2\lambda+1) - (1 - 2a + a^2) + 1\right)$$
$$= \frac{\lambda}{n}\left((1-a)^2(2\lambda+1) - (1-a)^2 + 1\right)$$
$$= \frac{\lambda}{n}\left((1-a)^2(2\lambda+1-1) + 1\right)$$
$$= \frac{\lambda}{n}\left(2\lambda(1-a)^2 + 1\right).$$

To minimize this variance, we need to minimize $(1-a)^2$, which occurs at $a = 1$.

Problem 3. Suppose $X_1, X_2, \ldots, X_n$ is a random sample from the Laplace$(0, \beta)$ distribution. Argue that

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_i, \frac{1}{n}\sum_{i=1}^n |X_i| - \beta\right) \xrightarrow{D} \text{normal}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma\right),$$

and find the matrix Σ. What is the asymptotic correlation between $\frac{1}{n}\sum_{i=1}^n X_i$ and $\frac{1}{n}\sum_{i=1}^n |X_i|$? [A more realistic scenario would be a sample from the Laplace(μ, β) distribution, with both

3

parameters needing to be estimated. It turns out that $\sqrt{n}\left(X_n - \mu, \frac{1}{n}\sum_{i=1}^{n}|X_i - X_n| - \beta\right)$ has the same limit distribution as the above.]

We have $\mu'_1 = \mathsf{E}(X_i) = 0$ and $\mu'_2 = \mathsf{E}(X_i^2) = 2\beta^2$. The variance and covariance terms are

$$\sigma_{11} = \mathsf{Var}(X_i) = 2\beta^2,$$

$$\sigma_{22} = \mathsf{Var}(|X_i|) = \mathsf{E}(X_i^2) - (\mathsf{E}(|X_i|))^2 = 2\beta^2 - \beta^2 = \beta^2,$$

$$\sigma_{12} = \sigma_{21} = \mathsf{Cov}(X_i, |X_i|) = \mathsf{E}(X_i|X_i|) - \mathsf{E}(X_i)\mathsf{E}(|X_i|) = 0 - 0 = 0.$$

Therefore, by Theorem 5.23, we have

$$\sqrt{n}\begin{pmatrix} \frac{1}{n}\sum_{i=1}^{n}X_i - 0 \\ \frac{1}{n}\sum_{i=1}^{n}|X_i| - \beta \end{pmatrix} \xrightarrow{D} \text{normal}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2\beta^2 & 0 \\ 0 & \beta^2 \end{pmatrix}\right).$$

The asymptotic correlation between $\frac{1}{n}\sum_{i=1}^{n}X_i$ and $\frac{1}{n}\sum_{i=1}^{n}|X_i|$ is

$$\mathsf{Corr}\left(\frac{1}{n}\sum_{i=1}^{n}X_i, \frac{1}{n}\sum_{i=1}^{n}|X_i|\right) = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}} = \frac{0}{\sqrt{2\beta^2 \cdot \beta^2}} = 0.$$

Problem 4. Let $\hat{F}(x)$ be the empirical cdf for a random sample $(X_1, \ldots, X_n)$ from distribution $F(x)$. Fix two values $x_1 < x_2$ in the support of F.

(a) Let Y_1 be the number of sample values that are less than or equal to x_1, Y_2 be the number that are greater than x_1 and less than or equal to x_2, and Y_3 be the number that are greater than x_2. Explain why (Y_1, Y_2, Y_3) has trinomial$(n, F(x_1), F(x_2) - F(x_1), 1 - F(x_2))$ joint distribution. (Recall Example 4.1 in the notes.)

(b) Observe that $Y_1/n = \hat{F}(x_1)$, $Y_2/n = \hat{F}(x_2) - \hat{F}(x_1)$, and $Y_3/n = 1 - \hat{F}(x_2)$. Use the conclusion of Example 4.1 (Slide 257) to deduce $\mathsf{Cov}(\hat{F}(x_1), \hat{F}(x_2) - \hat{F}(x_1))$ and thus $\mathsf{Cov}(\hat{F}(x_1), \hat{F}(x_2))$.

(a) Each X_i falls into one of the three categories:

- $X_i \leq x_1$,
- $x_1 < X_i \leq x_2$,
- $X_i > x_2$.

The probabilities of these three categories are

- $P(X_i \leq x_1) = F(x_1)$,
- $P(x_1 < X_i \leq x_2) = F(x_2) - F(x_1)$,
- $P(X_i > x_2) = 1 - F(x_2)$.

Since we have a random sample of size n, the joint distribution of (Y_1, Y_2, Y_3) is trinomial$(n, F(x_1), F(x_2) - F(x_1), 1 - F(x_2))$.

4

(b) By the conclusion of Example 4.1, we have

$$\text{Cov}(\hat{F}(x_1), \hat{F}(x_2) - \hat{F}(x_1)) = \text{Cov}\left(\frac{Y_1}{n}, \frac{Y_2}{n}\right)$$
$$= \frac{1}{n^2}\text{Cov}(Y_1, Y_2)$$
$$= \frac{1}{n^2}(-np_1p_2)$$
$$= -\frac{1}{n}F(x_1)(F(x_2) - F(x_1)).$$

Therefore,

$$\text{Cov}(\hat{F}(x_1), \hat{F}(x_2)) = \text{Cov}(\hat{F}(x_1), \hat{F}(x_2) - \hat{F}(x_1) + \hat{F}(x_1))$$
$$= \text{Cov}(\hat{F}(x_1), \hat{F}(x_2) - \hat{F}(x_1)) + \text{Cov}(\hat{F}(x_1), \hat{F}(x_1))$$
$$= -\frac{1}{n}F(x_1)(F(x_2) - F(x_1)) + \text{Var}(\hat{F}(x_1))$$
$$= -\frac{1}{n}F(x_1)(F(x_2) - F(x_1)) + \frac{1}{n}F(x_1)(1 - F(x_1))$$
$$= \frac{1}{n}F(x_1)(1 - F(x_2)).$$

Problem 5. Suppose $T_1, \ldots, T_{40}$ is a random sample from the exponential(1) distribution (which has 75-th percentile $t_{0.75} = \log(4) = 1.3863$).

(a) Based on Theorem 5.27 in the notes, find the exact probability that the sample 75-percentile, $\hat{t}_{0.75}$, exceeds 1.75.

(b) Find the approximate normal probability of the same event.

(c) Compare the two probabilities. Is the normal approximation reasonable in this case?

(a) By Theorem 5.27, we have

$$\text{P}(\hat{x}_p \le x) = 1 - B(\lceil np - 1\rceil; n, F(x))$$

Therefore,

$$\text{P}(\hat{t}_{0.75} > 1.75) = 1 - \text{P}(\hat{t}_{0.75} \le 1.75)$$
$$= B(29; 40, 1 - e^{-1.75})$$
$$= \sum_{k=0}^{29}\binom{40}{k}(1 - e^{-1.75})^k(e^{-1.75})^{40-k}$$
$$\approx 0.0748.$$

5

(b) Using the normal approximation, we have

$$P(\hat{t}_{0.75} > 1.75) = 1 - P(\hat{t}_{0.75} \leq 1.75)$$

$$\approx 1 - \left(1 - \Phi \left(\frac{\sqrt{n}(p - F(x))}{\sqrt{F(x)(1 - F(x))}} \right) \right)$$

$$= \Phi \left(\frac{\sqrt{40}(0.75 - (1 - e^{-1.75}))}{\sqrt{(1 - e^{-1.75})e^{-1.75}}} \right)$$

$$= \Phi(-1.272)$$

$$= 0.102.$$

(c) The two probabilities are 0.0748 (exact) and 0.102 (approximate), which are fairly close. Therefore, the normal approximation is reasonable in this case.

Problem 6. The Frechét(α, β) cdf is $F(t) = e^{-1/(\beta t^\alpha)}$ for $t > 0$, with $\alpha > 0$ and $\beta > 0$. (See also Example 5.16 in the notes.)

(a) What is the pdf?

(b) Prove that this distribution satisfies $\lim_{x \to \infty} x^\alpha (1 - F(x)) = c > 0$ for some c (and identify c).

(c) Suppose $T \sim$ Frechét(α, β). Show that $W = \frac{1}{T}$ has a Weibull distribution and identify its parameters.

(a) The pdf is

$$f(t) = \frac{d}{dt} F(t)$$

$$= \frac{d}{dt} e^{-1/(\beta t^\alpha)}$$

$$= e^{-1/(\beta t^\alpha)} \cdot \frac{\alpha}{\beta t^{\alpha+1}}$$

$$= \frac{\alpha}{\beta t^{\alpha+1}} e^{-1/(\beta t^\alpha)}.$$

(b) We have

$$\lim_{x \to \infty} x^\alpha (1 - F(x)) = \lim_{x \to \infty} x^\alpha \left(1 - e^{-1/(\beta x^\alpha)} \right)$$

$$= \infty - \infty \quad \text{(indeterminate form)}$$

$$= \lim_{x \to \infty} \frac{1 - e^{-1/(\beta x^\alpha)}}{1/x^\alpha}$$

$$= \lim_{x \to \infty} x^\alpha \cdot \frac{1}{\beta x^\alpha} \quad \text{(by L'Hospital's Rule)}$$

$$= \frac{1}{\beta}.$$

Therefore, $c = \frac{1}{\beta} > 0$.

(c) We have

$$
\begin{aligned}
F_W(w) &= P(W \le w) \\
&= P\left(\frac{1}{T} \le w\right) \\
&= P\left(T \ge \frac{1}{w}\right) \\
&= 1 - P\left(T < \frac{1}{w}\right) \\
&= 1 - F_T\left(\frac{1}{w}\right) \\
&= 1 - e^{-1/(\beta(1/w)^\alpha)} \\
&= 1 - e^{-(w^\alpha/\beta)}.
\end{aligned}
$$

Therefore, W has a Weibull distribution with parameters α and β.