HOMEWORK N° 13

Chapter 7

A-12

Suppose that the life X of a light bulb manufactured by a certain factory complies with normal distribution:$N(\mu, \sigma^2)$.

1. Find the 95% confidence interval for μ.
2. Find the 95% one-sided confidence level lower bound for μ.

Solution:

1. According to Equation (7.4.3), the confidence interval of $\hat{\mu}$ is:

$$(\bar{x} \pm \frac{s}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1))$$

In this example:

$$\bar{X} \approx 3400.933, S = 412.795, n = 15$$

Substituting those figures back into the aforementioned equation would we obtain the result:

$$(3172.34, 3629.53) \quad \checkmark$$

2. Likewise:

$$\hat{\theta}_{\mathrm{L}} = \bar{x} - \frac{s}{\sqrt{n}} t_{\alpha}(n-1) = 3213.21 \quad \checkmark$$

A-16

Suppose that the life expectancy of a vacuum tube set complies with the normal distribution $X \sim N(\mu, \sigma^2)$ of which the parameters μ, σ are unknown. Randomly extracting 16 samples and one had obtained $s = 300$. Find these under 95% confidence level:

1. Confidence interval.
2. One-sided confidence level upper bound.

Solution:

1. According to equation (7.4.5), the confidence interval of $\hat{\sigma}$ is:

$$(s\sqrt{\frac{(n-1)}{\chi^2_{\frac{\alpha}{2}}(n-1)}}, s\sqrt{\frac{(n-1)}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}})$$

In this example:

$$n = 16, s = 300$$

Therefore:

$$(\hat{\theta}_{\mathrm{L}}, \hat{\theta}_{\mathrm{U}}) = (221.61, 464.31)$$ ✓

2. Likewise:

$$\hat{\theta}_{\mathrm{U}} = s\sqrt{\frac{(n-1)}{\chi^2_{\alpha}(n-1)}} = 431.19$$

✓

A-17

To understand the spending habits of students in two universities in a certain city, a random survey of 100 students was conducted at each university. The results showed:

University A: Monthly average spending = 803 yuan, standard deviation = 75 yuan.
University B: Monthly average spending = 938 yuan, standard deviation = 102 yuan.
Assume:
The monthly spending of students at University A follows $X \sim N(\mu_1, \sigma^2)$.
The monthly spending of students at University B follows $Y \sim N(\mu_2, \sigma^2)$.
μ_1, μ_2, σ^2 are unknown.
The two samples are independent.
Find:

1. The 95% confidence interval for the difference in average monthly spending $\mu_1 - \mu_2$.
2. The one-sided 95% confidence lower bound for $\mu_1 - \mu_2$.

Solution:

1. According to equation (7.4.5), the confidence interval of $\mu_1 \hat{-} \mu_2$ is:

$$(\bar{x} - \bar{y} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}})$$

   In this example:

$$n_1 = n_2 = 100; \bar{x} = 803, \bar{y} = 938; s_1 = 75, s_2 = 102$$

   Therefore:

$$(\hat{\theta}_{\mathrm{L}}, \hat{\theta}_{\mathrm{U}}) = (-159.81, -110.19) \quad \checkmark$$

2. Likewise:

$$\hat{\theta}_{\mathrm{L}} = \bar{x} - \bar{y} - z_{\alpha} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = -155.82 \quad \checkmark$$

B-5

Let the density function of a population X be

$$f(x;\theta) = \begin{cases} \frac{2x}{\theta^2}, & 0 < x < \theta, \\ 0, & \text{otherwise}, \end{cases}$$

where $\theta > 0$ is unknown, and $X_1, X_2, \cdots, X_n$ $(n \geq 4)$ is a simple random sample drawn from the population X.

1. Find the method of moments estimator $\hat{\theta}_1$ and the maximum likelihood estimator $\hat{\theta}_2$ for θ;
2. Under the mean squared error (MSE) criterion, determine which estimator is more efficient;
3. Determine whether the two estimators are consistent estimators of θ.

Solution:

1. The method of moments estimator is:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

   And:

$$\mathbb{E}[X] = \int_0^{\theta} x \cdot \frac{2x}{\theta^2} dx = \frac{2}{3}\theta$$

   Therefore:

$$\hat{\theta}_1 = \frac{3}{2n} \sum_{i=1}^{n} X_i$$

For the maximum likelihood estimator, its maximum log-likelihood function is:

$$\ell(\theta, x_i) = n \ln 2 - 2n \ln \theta + \sum_{i=1}^{n} \ln x_i$$

The argument of the maxima is the maximum likelihood estimator:

$$\hat{\theta}_2 = \arg\max \ell(\theta, x_i) = \max\{X_1, \ldots, X_n\}$$

2. We already have:

$$\mathbb{E}[\bar{X}] = \frac{2\theta}{3}, \quad \mathrm{Var}(\bar{X}) = \frac{\mathrm{Var}(X)}{n}$$

We compute $\mathrm{Var}(X)$:

$$\mathbb{E}[X^2] = \int_0^\theta x^2 \cdot \frac{2x}{\theta^2}\, dx = \frac{2}{\theta^2} \int_0^\theta x^3\, dx = \frac{2}{\theta^2} \cdot \frac{\theta^4}{4} = \frac{\theta^2}{2}$$

Then:

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{\theta^2}{2} - \left(\frac{2\theta}{3}\right)^2 = \frac{\theta^2}{18}$$

$$\mathrm{Var}(\hat{\theta}_1) = \left(\frac{3}{2}\right)^2 \cdot \frac{\mathrm{Var}(X)}{n} = \frac{9}{4} \cdot \frac{\theta^2}{18n} = \frac{\theta^2}{8n}$$

Thus, since $\hat{\theta}_1$ is unbiased:

$$\mathrm{MSE}(\hat{\theta}_1) = \mathrm{Var}(\hat{\theta}_1) = \frac{\theta^2}{8n}$$

For $\hat{\theta}_2$, this estimator is biased. Its CDF is:

$$F_{\hat{\theta}_2}(x) = P(\max X_i \le x) = (F_X(x))^n = \left(\frac{x^2}{\theta^2}\right)^n$$

Then the expectation is:

$$\mathbb{E}[\hat{\theta}_2] = \int_0^\theta x \cdot f_{\hat{\theta}_2}(x)\, dx = \frac{2n}{\theta^{2n}} \int_0^\theta x^{2n}\, dx = \frac{2n\theta}{2n+1}$$

Bias: $\mathrm{Bias} = \frac{2n\theta}{2n+1} - \theta = -\frac{\theta}{2n+1}$

Variance (standard result):

$$\mathrm{Var}(\hat{\theta}_2) = \theta^2 (\frac{2n}{2n+2} - \frac{4n^2}{(2n+1)^2})$$

Hence, MLE is more efficient than MoM:

$$\mathrm{MSE}(\hat{\theta}_2) = \mathrm{Var} + \mathrm{Bias}^2 = \theta^2 (\frac{2n}{2n+2} - \frac{4n^2 - 1}{(2n+1)^2}) < \mathrm{MSE}(\hat{\theta}_1)$$

3.

$$\hat{\theta}_1 = \frac{3}{2}\bar{X} \xrightarrow{P} \frac{3}{2} \cdot \mathbb{E}[X] = \theta \Rightarrow \text{consistent}$$

$$\hat{\theta}_2 = \max X_i \xrightarrow{P} \theta \quad \text{(standard result for max of bounded distribution)} \Rightarrow \text{consistent}$$

B-6

Let the density function of the population X be

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}e^{-\frac{x}{\theta}}, & x > 0, \\ 0, & \text{otherwise}, \end{cases}$$

where $\theta > 0$ is unknown, and $X_1, X_2, \cdots, X_n$ is a simple random sample from the population X.

1. Prove that the sample mean is both the moment estimator and the maximum likelihood estimator of θ;
2. Among the estimators of the form $c\sum_{i=1}^{n} X_i$, find the value of c that is optimal under the mean squared error criterion;
3. Determine whether the estimator obtained in (2) is a consistent estimator of θ.

Solution:

1. The first population moment (expected value) of X is:

$$E[X] = \int_0^\infty x \cdot \frac{1}{\theta}e^{-\frac{x}{\theta}}\, dx = \theta.$$

The sample moment is the sample mean:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Setting the population moment equal to the sample moment:

$$E[X] = \overline{X} \implies \theta = \overline{X}.$$

Thus, the moment estimator of θ is $\hat{\theta}_{\text{MM}} = \overline{X}$.

The likelihood function is:

$$L(\theta) = \prod_{i=1}^{n} f(X_i;\theta) = \theta^{-n}e^{-\frac{1}{\theta}\sum_{i=1}^{n} X_i}.$$

The log-likelihood function is:

$$\ln L(\theta) = -n\ln\theta - \frac{1}{\theta}\sum_{i=1}^{n} X_i.$$

Taking the derivative w.r.t. θ and setting it to zero:

$$\frac{d}{d\theta}\ln L(\theta) = -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} X_i = 0.$$

✓

Solving for θ:

$$\hat{\theta}_{\text{MLE}} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}.$$

Thus, the MLE of θ is also $\overline{X}$.

2. We seek an estimator of the form $\hat{\theta} = c\sum_{i=1}^{n} X_i$.

   Since $X_i \sim \text{Exp}(\theta^{-1})$, we have:

$$E\left[\sum_{i=1}^{n} X_i\right] = n\theta, \quad \text{Var}\left(\sum_{i=1}^{n} X_i\right) = n\theta^2.$$

   The bias of $\hat{\theta}$ is:

$$\text{Bias}(\hat{\theta}) = E[\hat{\theta}] - \theta = cn\theta - \theta = \theta(cn-1).$$

   The variance is:

$$\text{Var}(\hat{\theta}) = c^2 \cdot n\theta^2.$$

   The MSE is:

$$\text{MSE}(\hat{\theta}) = \text{Var}(\hat{\theta}) + \text{Bias}^2(\hat{\theta}) = c^2 n\theta^2 + \theta^2(cn-1)^2.$$

   To minimize $\text{MSE}(\hat{\theta})$, take the derivative w.r.t. c and set it to zero:

$$\frac{d}{dc}\text{MSE}(\hat{\theta}) = 2cn\theta^2 + 2\theta^2(cn-1)n = 0.$$

   Simplify:

$$2cn + 2n(cn-1) = 0 \implies cn + cn^2 - n = 0 \implies c(n+n^2) = n.$$

   Thus:

$$c = \frac{n}{n+n^2} = \frac{1}{1+n}.$$

   ✓

   The optimal estimator under the MSE criterion is:

$$\hat{\theta}_{\text{opt}} = \frac{1}{1+n}\sum_{i=1}^{n} X_i.$$

3. An estimator is consistent if it converges in probability to the true parameter as $n \to \infty$.

   The estimator is:

$$\hat{\theta}_{\text{opt}} = \frac{1}{1+n}\sum_{i=1}^{n} X_i = \frac{n}{1+n}\overline{X}.$$

   As $n \to \infty$:

$$\frac{n}{1+n} \to 1, \quad \overline{X} \xrightarrow{P} \theta \quad \text{(by the Law of Large Numbers)}.$$

   Thus:

$$\hat{\theta}_{\text{opt}} \xrightarrow{P} \theta.$$

   $\hat{\theta}_{\text{opt}}$ is a consistent estimator of θ.

✓

B-7

Let the density function of the population X be

$$f(x;\theta,\lambda)=\begin{cases}\frac{\lambda x^{\lambda-1}}{\theta^{\lambda}}, & 0<x<\theta,\\ 0, & \text{otherwise,}\end{cases}$$

where $\theta>0,\lambda>0$, and $X_1,X_2,\cdots,X_n$ is a simple random sample from the population X.

1. If $\lambda=3$ and θ is unknown, find the method of moment estimator of θ and determine whether it is biased or not.
2. Similarly, $\theta=3$, λ unknown, find the maximum likelihood estimator, determining whether it is biased.

Solution:

1. Given the density function:

$$f(x;\theta,3)=\begin{cases}\frac{3x^2}{\theta^3}, & 0<x<\theta,\\ 0, & \text{otherwise,}\end{cases}$$

   the expected value is:

$$E[X]=\int_0^{\theta} x\cdot\frac{3x^2}{\theta^3}\,dx=\frac{3}{\theta^3}\int_0^{\theta}x^3\,dx=\frac{3}{\theta^3}\cdot\frac{\theta^4}{4}=\frac{3\theta}{4}.$$

   The sample moment is the sample mean $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$. Setting $E[X]=\overline{X}$:

$$\frac{3\theta}{4}=\overline{X}\implies\hat{\theta}_{\text{MM}}=\frac{4\overline{X}}{3}.$$

   Compute the expected value of $\hat{\theta}_{\text{MM}}$:

$$E[\hat{\theta}_{\text{MM}}]=E\left[\frac{4\overline{X}}{3}\right]=\frac{4}{3}E[\overline{X}]=\frac{4}{3}\cdot\frac{3\theta}{4}=\theta.$$

   Since $E[\hat{\theta}_{\text{MM}}]=\theta$, the estimator is unbiased.

2. Given $\theta=3$, the density becomes:

$$f(x;3,\lambda)=\begin{cases}\frac{\lambda x^{\lambda-1}}{3^{\lambda}}, & 0<x<3,\\ 0, & \text{otherwise.}\end{cases}$$

   The likelihood function is:

$$L(\lambda)=\prod_{i=1}^{n}f(X_i;3,\lambda)=\lambda^n\left(\prod_{i=1}^{n}X_i\right)^{\lambda-1}\cdot 3^{-n\lambda}.$$

   The log-likelihood is:

$$\ln L(\lambda)=n\ln\lambda+(\lambda-1)\sum_{i=1}^{n}\ln X_i-n\lambda\ln 3.$$

Take the derivative w.r.t. λ and set it to zero:

$$\frac{d}{d\lambda}\ln L(\lambda) = \frac{n}{\lambda} + \sum_{i=1}^{n}\ln X_i - n\ln 3 = 0.$$

Solve for λ:

$$\hat{\lambda}_{\text{MLE}} = \frac{n}{\sum_{i=1}^{n}(\ln 3 - \ln X_i)}.$$

✓

To verify consistency, we need to show that $\hat{\lambda}_{\text{MLE}}$ converges in probability to λ as $n \to \infty$. The expected value $E[\ln X]$ is:

$$E[\ln X] = \int_0^3 \ln x \cdot \frac{\lambda x^{\lambda-1}}{3^\lambda}\,dx.$$

This integral evaluates to:

$$E[\ln X] = \ln 3 - \frac{1}{\lambda}.$$

By the LLN:

$$\frac{1}{n}\sum_{i=1}^{n}\ln X_i \xrightarrow{P} \ln 3 - \frac{1}{\lambda}.$$

Thus:

$$\hat{\lambda}_{\text{MLE}} = \frac{1}{\ln 3 - \frac{1}{n}\sum_{i=1}^{n}\ln X_i} \xrightarrow{P} \frac{1}{\ln 3 - (\ln 3 - \frac{1}{\lambda})} = \lambda.$$

The MLE $\hat{\lambda}_{\text{MLE}}$ is a consistent estimator of λ.

✓

B-8

Let the density function of the population X be

$$f(x;\theta) = \begin{cases} e^{-(x-\theta)}, & x \geq \theta, \\ 0, & \text{otherwise,} \end{cases}$$

where $\theta > 0$ is unknown, and $X_1, X_2, \cdots, X_n$ is a simple random sample from the population X.

1. Find the maximum likelihood estimator (MLE) $\hat{\theta}$ of θ.

2. Determine the density function of $\hat{\theta} - \theta$.

3. Determine whether $\hat{\theta} - \theta$ can be used as a pivotal quantity for interval estimation of θ.

4. Find the lower confidence bound for θ at confidence level $1 - \alpha$.

Solution:

1. The density is nonzero only when $x \geq \theta$. Thus, the likelihood function is:

$$L(\theta) = \prod_{i=1}^{n} f(X_i;\theta) = \prod_{i=1}^{n} e^{-(X_i-\theta)} = e^{-\sum_{i=1}^{n} X_i + n\theta}, \quad \text{for } X_i \geq \theta \text{ for all } i.$$

   The condition $X_i \geq \theta$ for all i implies that $\theta \leq \min(X_1, X_2, \ldots, X_n)$.

   The likelihood $L(\theta)$ increases with θ (since $n\theta$ grows as θ increases). However, θ is constrained by $\theta \leq \min(X_1, X_2, \ldots, X_n)$.

   Thus, the MLE is:

$$\hat{\theta} = \min\{X_1, X_2, \ldots, X_n\}.$$

2. Each X_i has CDF:

$$F(x;\theta) = \begin{cases} 1 - e^{-(x-\theta)}, & x \geq \theta, \\ 0, & x < \theta. \end{cases}$$

   The CDF of $\hat{\theta}$ is:

$$P(\hat{\theta} \leq t) = 1 - P(\hat{\theta} > t) = 1 - P(X_1 > t, \ldots, X_n > t) = 1 - [1 - F(t)]^n = 1 - e^{-n(t-\theta)}, \quad t \geq \theta.$$

   Thus, the PDF of $\hat{\theta}$ is:

$$f_{\hat{\theta}}(t) = \frac{d}{dt} P(\hat{\theta} \leq t) = ne^{-n(t-\theta)}, \quad t \geq \theta.$$

   Let $Y = \hat{\theta} - \theta$. Then:

$$P(Y \leq y) = P(\hat{\theta} \leq \theta + y) = 1 - e^{-ny}, \quad y \geq 0.$$

   Thus, the PDF of Y is:

$$f_Y(y) = ne^{-ny}, \quad y \geq 0.$$

   This is an exponential distribution with rate n.

3. A **pivotal quantity** must:

   1. Depend on the data and θ,
   2. Have a distribution that does not depend on θ.

   From (2), $\hat{\theta} - \theta$ follows an exponential distribution with rate n, which does not depend on θ. Therefore, $\hat{\theta} - \theta$ is a valid pivotal quantity for θ.

4. Using the pivotal quantity $Y = \hat{\theta} - \theta \sim \mathrm{Exp}(n)$, we have:

$$P(Y \leq c) = 1 - e^{-nc} = 1 - \alpha.$$

   Solving for c:

$$e^{-nc} = \alpha \implies c = -\frac{\ln \alpha}{n}.$$

   Thus:

$$P\left(\hat{\theta} - \theta \leq -\frac{\ln \alpha}{n}\right) = \alpha \implies P\left(\theta \geq \hat{\theta} + \frac{\ln \alpha}{n}\right) = 1 - \alpha.$$

$$\hat{\theta}_L = \hat{\theta} + \frac{\ln \alpha}{n}$$