STA5002: Mathematical Statistics

Assignment 3 Solution (Dec 4th – Dec13th)

Note: The solutions only serve as a reference. Some problems may have different methods to reach the same answer.

1. Randomly generate two independent samples from population $N(100, 4)$, the sample means of the two samples are denoted by $\bar{X}_1$ and $\bar{X}_2$, respectively. The sample sizes are 20 and 25, compute $P(|\bar{X}_1 - \bar{X}_2| > 0.1)$. (5 points)

   Solution: By the description of the problem, we know that $\bar{X}_1 \sim N(100, 4/20)$ and $\bar{X}_2 \sim N(100,\ 4/25)$. Since the two samples are independent, $\bar{X}_1$ and $\bar{X}_2$ are independent and

$$\bar{X}_1 - \bar{X}_2 \sim N\left(0, \frac{4}{20} + \frac{4}{25}\right) = N\left(0, \frac{9}{25}\right).$$

   Therefore:

$$P(|\bar{X}_1 - \bar{X}_2| > 0.1) = P\left(\frac{|\bar{X}_1 - \bar{X}_2|}{3/5} > \frac{0.1}{3/5}\right) = P\left(|Z| > \frac{1}{6}\right) \approx 2[1 - \Phi(0.17)] = 0.8650.$$

2. Suppose that $X_1, X_2, \ldots, X_{15}$ is a sample from population $X \sim N(0, \sigma^2)$, define

$$Y = \frac{X_1^2 + X_2^2 + \cdots + X_{10}^2}{2(X_{11}^2 + X_{12}^2 + \cdots + X_{15}^2)}.$$

   Compute $P(Y > 1)$. (5 points)

   Solution: It is obvious that X_i/σ are i.i.d. random variables which follow the standard normal distribution. So

$$\frac{X_1^2 + X_2^2 + \cdots + X_{10}^2}{\sigma^2} \sim \mathcal{X}^2(10), \frac{X_{11}^2 + X_{12}^2 + \cdots + X_{15}^2}{\sigma^2} \sim \mathcal{X}^2(5),$$

   moreover, the two terms are independent. Therefore:

$$Y = \frac{X_1^2 + X_2^2 + \cdots + X_{10}^2}{2(X_{11}^2 + X_{12}^2 + \cdots + X_{15}^2)} = \frac{\frac{1}{\sigma^2}(X_1^2 + X_2^2 + \cdots + X_{10}^2)/10}{\frac{1}{\sigma^2}(X_{11}^2 + X_{12}^2 + \cdots + X_{15}^2)/5} \sim F(10, 5).$$

   Using R, we have:

$$P(Y > 1) \approx 1 - 0.4651 = 0.5349.$$

3. Suppose that $X_1, X_2, \dots, X_n, X_{n+1}$ is a sample from population $X \sim N(\mu, \sigma^2)$. Let

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2.$$

Compute constant c such that $T_c = c(X_{n+1} - \bar{X}_n)/S_n$ follows a $t-$distribution and specify the degree of freedom of the $t-$distribution. (10 points)

Solution: By the description of the problem, we have

$$X_{n+1} \sim N(\mu, \sigma^2), \bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right), \frac{(n-1)S_n^2}{\sigma^2} \sim \mathcal{X}^2(n-1),$$

and $X_{n+1}, \bar{X}_n, S_n^2$ are independent. Therefore

$$X_{n+1} - \bar{X}_n \sim N\left(0, \sigma^2 + \frac{\sigma^2}{n}\right) = N\left(0, \frac{n+1}{n}\sigma^2\right)$$

$$\Longrightarrow T = \frac{(X_{n+1} - \bar{X}_n)/\sqrt{\frac{n+1}{n}\sigma^2}}{\sqrt{\frac{(n-1)S_n^2}{\sigma^2}/(n-1)}} = \sqrt{\frac{n}{n+1}}\frac{(X_{n+1} - \bar{X}_n)}{S_n} \sim t(n-1).$$

This indicates that when $c = \sqrt{n/(n+1)}$, $T_c = c(X_{n+1} - \bar{X}_n)/S_n$ follows a t-distribution and the degree of freedom is $n-1$.

4. Suppose that $X_1, X_2, \dots, X_n$ is a sample from population $X \sim U[\theta_1, \theta_2](\theta_2 > \theta_1)$. Try to obtain the sufficient statistic of (θ_1, θ_2). (10 points)

Solution: The joint PDF of $\boldsymbol{X} = (X_1, X_2, \dots, X_n)$ is

$$\begin{aligned} f(x_1, x_2, \cdots, x_n; \theta_1, \theta_2) &= f(\boldsymbol{x}; \theta_1, \theta_2) \\ &= \begin{cases} \left(\frac{1}{\theta_2 - \theta_1}\right)^n, \text{if } \theta_1 \le x_1, x_2, \dots, x_n \le \theta_2 \\ 0, \qquad \text{otherwise} \end{cases} \\ &= \begin{cases} \left(\frac{1}{\theta_2 - \theta_1}\right)^n, \text{if } \theta_1 \le x_{(1)} \le x_{(n)} \le \theta_2 \\ 0, \qquad \text{otherwise} \end{cases}. \end{aligned}$$

Let $T_1 = T_1(\boldsymbol{X}) = X_{(1)}$, $T_2 = T_2(\boldsymbol{X}) = X_{(n)}$, $h(\boldsymbol{x}) = 1$, and ($I(\cdot)$ is the indicator function)

$$g(T_1, T_2, \theta_1, \theta_2) = \left(\frac{1}{\theta_2 - \theta_1}\right)^n I(\theta_1 \le T_1 \le T_2 \le \theta_2),$$

then $f(\boldsymbol{x}; \theta_1, \theta_2) = g(T_1, T_2, \theta_1, \theta_2)h(\boldsymbol{x})$. By the factorization theorem, $\boldsymbol{T} = (T_1, T_2) =$

$\left(X_{(1)}, X_{(n)}\right)$ is a sufficient statistic of (θ_1, θ_2).

5. For each of the following PDFs, assume $X_1, X_2, \dots, X_n$ is a sample from each PDF, compute the moment estimators of the unknown parameters.

   (1) $f(x;\theta) = (\theta+1)x^{\theta},\ 0 < x < 1, \theta > 0$. (5 points)

   (2) $f(x;\theta,\mu) = \exp\{-(x-\mu)/\theta\}/\theta,\ x > \mu,\ \theta > 0$. (5 points)

   Solution:

   (1) Compute the population 1st moment of $X \sim f(x;\theta)$:

$$E(X) = \int_0^1 x(\theta+1)x^{\theta}dx = (\theta+1)\int_0^1 x^{\theta+1}dx = \frac{\theta+1}{\theta+2}.$$

$$\Longrightarrow \theta = \frac{1-2E(X)}{E(X)-1} \Longrightarrow \text{moment estimator of } \theta \text{ is } \hat{\theta} = \frac{1-2\bar{X}}{\bar{X}-1}.$$

   (2) Compute the population 1st and 2nd moments of $X \sim f(x;\theta,\mu)$:

$$E(X) = \int_{\mu}^{\infty} \frac{x}{\theta}\exp\left\{-\frac{x-\mu}{\theta}\right\}dx = \frac{1}{\theta}\left[\int_0^{\infty} te^{-\frac{t}{\theta}}dt + \mu\int_0^{\infty} e^{-\frac{t}{\theta}}dt\right] = \theta+\mu.$$

$$E(X^2) = \int_{\mu}^{\infty} \frac{x^2}{\theta}\exp\left\{-\frac{x-\mu}{\theta}\right\}dx = \frac{1}{\theta}\int_0^{\infty}(t+\mu)^2 e^{-\frac{t}{\theta}}dt$$

$$= \frac{1}{\theta}\left[\int_0^{\infty} t^2 e^{-\frac{t}{\theta}}dt + 2\mu\int_0^{\infty} te^{-\frac{t}{\theta}}dt + \mu^2\int_0^{\infty} e^{-\frac{t}{\theta}}dt\right]$$

$$= 2\theta^2 + 2\mu\theta + \mu^2.$$

$$\Longrightarrow \mathrm{Var}(X) = E(X^2) - [E(X)]^2 = (2\theta^2 + 2\mu\theta + \mu^2) - (\theta+\mu)^2 = \theta^2.$$

$$\Longrightarrow \theta = \sqrt{\mathrm{Var}(X)}, \mu = E(X) - \sqrt{\mathrm{Var}(X)}.$$

   Therefore, the moment estimators of θ and μ are:

$$\hat{\theta} = \tilde{S}, \hat{\mu} = \bar{X} - \tilde{S}.$$

6. Suppose that the number of words in a sentence from a book X approximately follows a log-normal distribution, i.e., $Y = \ln X \sim N(\mu, \sigma^2)$. 20 sentences are randomly picked from the book and the number of words in them are

   50 13 13 61 14 5 26 5 8 57

28 4 27 12 31 30 24 20 65 22

Compute the maximum likelihood estimate of $\theta = E(X) = e^{\mu+\sigma^2/2}$, the expected number of words of a sentence from the book. (10 points)

Solution: The maximum likelihood estimators of μ and σ^2 of $N(\mu, \sigma^2)$ are the sample mean and adjusted sample variance. So, the estimates are

$$\hat{\mu} = \frac{1}{20}\sum_{i=1}^{20} y_i = \frac{1}{20}\sum_{i=1}^{20} \ln x_i \approx 2.9582.$$

$$\hat{\sigma}^2 = \frac{1}{20}\sum_{i=1}^{20}(y_i - 2.9582)^2 = \frac{1}{20}\sum_{i=1}^{20}(\ln x_i - 2.9582)^2 \approx 0.6622.$$

Due to the invariance property of MLE, the MLE of $\theta = E(X) = e^{\mu+\sigma^2/2}$ is

$$\hat{\theta} = e^{2.9582+0.6622/2} \approx 26.8241.$$

7. Assume that $X_1, X_2, \ldots, X_n$ is a sample from population X with PDF $f(x;\theta) = \theta x^{\theta-1}$, $0 < x < 1,\ \theta > 0$.

(1) Compute the maximum likelihood estimator of $g(\theta) = 1/\theta$. (5 points)

(2) Compute the C-R lower bound of any unbiased estimator of $g(\theta)$ and show that the estimator in (1) is an efficient estimator of $g(\theta)$. (5 points)

Solution:

(1) The likelihood function is

$$L(\theta; \boldsymbol{x}) = (\theta)^n (x_1 x_2 \cdots x_n)^{\theta-1}.$$

So, the log-likelihood function is

$$\ell(\theta; \boldsymbol{x}) = n \ln\theta + (\theta - 1)(\ln x_1 + \ln x_2 + \cdots + \ln x_n).$$

Take the first derivative of $\ell(\theta)$, set it to zero and solve the equation:

$$\frac{\partial \ell}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^{n} \ln x_i = 0 \Longrightarrow \hat{\theta} = -\frac{n}{\sum_{i=1}^{n} \ln x_i}.$$

Consider the second derivative of ℓ evaluated at $\hat{\theta}$:

$$\left.\frac{\partial^2 \ell}{\partial \theta^2}\right|_{\hat{\theta}} = \left.\left(-\frac{n}{\theta^2}\right)\right|_{\hat{\theta}} = -\frac{n}{\hat{\theta}^2} < 0.$$

So, $\hat{\theta}$ is the maximum likelihood estimator of θ. By the invariance property of MLE, the MLE of $g(\theta) = 1/\theta$ is

$$\hat{g} = -\frac{1}{n}\sum_{i=1}^{n} \ln X_i.$$

(2) First compute the fisher information of θ, since $\ln f(x;\theta) = \ln\theta + (\theta - 1)\ln x$, so:

$$\frac{\partial \ln f(x;\theta)}{\partial\theta} = \frac{1}{\theta} + \ln x\,, \frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2} = -\frac{1}{\theta^2}.$$

$$\Longrightarrow I(\theta) = -E\left(\frac{\partial^2 \ln f(X;\theta)}{\partial\theta^2}\right) = \frac{1}{\theta^2}.$$

Since $g(\theta) = 1/\theta$, so $g'(\theta) = -1/\theta^2$, consequently, the C-R lower bound of any unbiased estimator of $g(\theta)$ is

$$\frac{[g'(\theta)]^2}{nI(\theta)} = \frac{1}{n\theta^2}.$$

Then compute the expectation and variance of $\hat{g}$ in (1). Let $Y = -\ln X$, then

$$P(Y < y) = P(-\ln X < y) = P(X > e^{-y}) = \int_{e^{-y}}^{1} \theta x^{\theta-1} dx = 1 - e^{-\theta y}.$$

So, $Y \sim Exp(\theta)$, $\Longrightarrow E(Y) = 1/\theta$ and $\mathrm{Var}(Y) = 1/\theta^2$, consequently:

$$E(\hat{g}) = \frac{1}{n}\sum_{i=1}^{n} E(-\ln X_i) = E(Y) = \frac{1}{\theta}, \mathrm{Var}(\hat{g}) = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(\ln X_i) = \frac{1}{n}\mathrm{Var}(Y) = \frac{1}{n\theta^2}.$$

This indicate that $\hat{g}$ is an unbiased estimator of $g(\theta) = 1/\theta$ and it attains the C-R lower bound. So $\hat{g}$ is an efficient estimator of $g(\theta)$.

8. Assume that $X_1, X_2, \dots, X_n$ is a sample from population X with PDF $f(x|\theta) = \theta x^{\theta-1}$, $0 < x < 1$, $\theta > 0$. Let the prior distribution of θ be the Gamma distribution, i.e., $\theta \sim Gamma(\alpha, \beta)$, compute the posterior expectation as the Bayes' estimator of θ. (Hint: $\int_0^\infty x^{\alpha-1} e^{-x} dx = \Gamma(\alpha)$, the expectation of $Y \sim Gamma(\alpha, \beta)$ is $E(Y) = \alpha/\beta$) (10 points)

Solution: The joint distribution of $X_1, X_2, \dots, X_n$ and θ is

$$f(x_1, x_2, \ldots, x_n, \theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} \cdot \prod_{i=1}^{n} \theta x_i^{\theta-1}$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{n+\alpha-1} \exp\left\{-\theta\left(\beta - \sum_{i=1}^{n} \ln x_i\right)\right\} \prod_{i=1}^{n} \frac{1}{x_i}.$$

Then (for the second "=", set $\tilde{\theta} = (\beta - \sum_{i=1}^{n} \ln x_i)\theta$)

$$\int_0^\infty f(x_1, x_2, \ldots, x_n, \theta) d\theta = \frac{\beta^\alpha}{\Gamma(\alpha)} \prod_{i=1}^{n} \frac{1}{x_i} \int_0^{+\infty} \theta^{n+\alpha-1} \exp\left\{-\theta\left(\beta - \sum_{i=1}^{n} \ln x_i\right)\right\} d\theta$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \prod_{i=1}^{n} \frac{1}{x_i} \left(\beta - \sum_{i=1}^{n} \ln x_i\right)^{-(n+\alpha)} \int_0^{+\infty} \tilde{\theta}^{n+\alpha-1} \exp\{-\tilde{\theta}\} d\tilde{\theta}$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \prod_{i=1}^{n} \frac{1}{x_i} \left(\beta - \sum_{i=1}^{n} \ln x_i\right)^{-(n+\alpha)} \Gamma(n+\alpha)$$

Therefore, the posterior distribution of θ is

$$\pi(\theta | x_1, x_2, \ldots, x_n) = \frac{f(x_1, x_2, \ldots, x_n, \theta)}{\int_0^\infty f(x_1, x_2, \ldots, x_n, \theta) d\theta}$$

$$= \frac{(\beta - \sum_{i=1}^{n} \ln x_i)^{n+\alpha}}{\Gamma(n+\alpha)} \theta^{n+\alpha-1} \exp\left\{-\theta\left(\beta - \sum_{i=1}^{n} \ln x_i\right)\right\}.$$

It is not difficult to see that it is the Gamma distribution $Gamma(n + \alpha, \beta - \sum_{i=1}^{n} \ln x_i)$. Then the posterior expectation as the Bayes' estimator of θ is

$$\hat{\theta}_B = \frac{n+\alpha}{\beta - \sum_{i=1}^{n} \ln x_i}.$$

9. It is assumed that the compressive strength (抗压强度) of a type of material is $X \sim N(\mu, \sigma^2)$. Now randomly pick 10 test-piece and perform the compression test (抗压试验), the compressive strengths are: 479, 490, 454, 468, 507, 443, 432, 415, 396, 466.

(1) If it is known that $\sigma = 30$, compute the 95% confidence interval of μ. (5 points)

(2) Compute the 95% confidence interval of μ assuming σ^2 is unknown. (5 points)

(3) Compute the 95% confidence interval of σ. (5 points)

Solution:

(1) The sample mean is computed to be $\bar{x} = 455$, the $100(1-\alpha)\%$ confidence interval of μ when σ is known is:

$$\bar{X} \pm u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}.$$

Plugging in $\bar{x} = 455$, $\sigma = 30$, $n = 10$ and $u_{0.975} = 1.96$, we obtain the 95% confidence interval of μ to be $[436.4058, 473.5942]$.

(2) By computation, the sample mean and sample variance are $\bar{x} = 455$, $s^2 = 1172.222$, so the sample standard deviation is $s \approx 34.24$. When σ is unknown, the $100(1-\alpha)\%$ confidence interval of μ is

$$\bar{X} \pm t_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}}.$$

Plugging in $\bar{x} = 455$, $s = 34.24$, $n = 10$ and $t_{0.975}(9) = 2.262$, we obtain the 95% confidence interval of μ to be $[430.5095, 479.4905]$.

(3) The $100(1-\alpha)\%$ confidence interval of σ^2 is

$$\left[\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}\right].$$

From the chi-square distribution table, we have $\chi^2_{0.025}(9) = 2.700$, $\chi^2_{0.975}(9) = 19.02$. Plugging in $(n-1)s^2 = 10550$, the 95% confidence interval for σ^2 is

$$\left[\frac{10550}{19.02}, \frac{10550}{2.700}\right] \approx [554.679, 3907.407].$$

The 95% confidence interval for σ is then $[\sqrt{554.679}, \sqrt{3907.407}] \approx [23.552, 62.509]$.

10. Assume that population $X \sim N(\mu_1, \sigma_1^2)$ and population $Y \sim N(\mu_2, \sigma_2^2)$. Two independent samples with sample sizes $n_1 = 10$, $n_2 = 13$ are obtained from the two populations, the sample means and variances are computed as $\bar{x} = 82$, $s_1^2 = 56.5$, $\bar{y} = 76$, $s_2^2 = 52.4$.

(1) If it is known that $\sigma_1^2 = 64$, $\sigma_2^2 = 49$, compute the 95% confidence interval of $\mu_1 - \mu_2$. (5 points)

(2) If it is known that $\sigma_1^2 = \sigma_2^2$, compute the 95% confidence interval of $\mu_1 - \mu_2$. (5 points)

(3) Compute the 95% confidence interval of σ_1^2/σ_2^2. (5 points)

Solution:

(1) When σ_1^2 and σ_2^2 are known, the $100(1-\alpha)\%$ confidence interval is

$$(\bar{X} - \bar{Y}) \pm u_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

Plugging in that $\bar{x}$, $\bar{y}$, $u_{0.975}=1.96$, $\sigma_1^2=64$, $\sigma_2^2=49$, the 95% confidence interval of $\mu_1-\mu_2$ is

$$(82-76)\pm 1.96\times\sqrt{\frac{64}{10}+\frac{49}{13}}=[-0.2503, 12.2503].$$

(2) If $\sigma_1^2=\sigma_2^2$, the $100(1-\alpha)\%$ confidence interval of $\mu_1-\mu_2$ is

$$(\bar{X}-\bar{Y})\pm t_{1-\alpha/2}(n_1+n_2-2)S_\omega\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},$$

where

$$S_\omega^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}.$$

Plugging in $n_1=10$, $n_2=13$, $s_1^2=56.5$, $s_2^2=52.4$, we have $s_\omega^2=54.1571$. Moreover, from the t-distribution table, $t_{0.975}(21)=2.080$, so the 95% confidence interval for $\mu_1-\mu_2$ is

$$(82-76)\pm 2.080\times\sqrt{54.1571}\times\sqrt{\frac{1}{10}+\frac{1}{13}}=[-0.4385, 12.4385].$$

(3) The $100(1-\alpha)\%$ confidence interval for σ_1^2/σ_2^2 is

$$\left[\frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{1-\alpha/2}(n_1-1,n_2-1)},\frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{\alpha/2}(n_1-1,n_2-1)}\right].$$

From the F-square distribution table, we have $F_{0.975}(9,12)=3.44$, $F_{0.975}(12,9)=3.87$. By the triple-reverse formula,

$$F_{0.025}(9,12)=\frac{1}{F_{0.975}(12,9)}=\frac{1}{3.87}.$$

Plugging in $s_1^2=56.5$, $s_2^2=52.4$, $F_{0.025}(9,12)=1/3.87$, $F_{0.975}(9,12)=3.44$, the 95% confidence interval of σ_1^2/σ_2^2 is

$$\left[\frac{56.5}{52.4}\cdot\frac{1}{3.44},\frac{56.5}{52.4}\cdot 3.87\right]=[0.3134, 4.1728].$$