Assignment 7

12310520

南煜涵

T1 (1) Let X be the number of concurrent users

$$\bar{X} = \frac{1}{10} \times (17.2+22.1+18.5+17.2+18.6+14.8+21.7+15.8+16.3+22.8)$$
$$= 18.5k = 1.85 \times 10^4$$

$$S^2 = \frac{1}{9} \times [(17.2-18.5)^2+(22.1-18.5)^2+(18.5-18.5)^2+(17.2-18.5)^2+ (18.6-18.5)^2+(14.8-18.5)^2+(21.7-18.5)^2+(15.8-18.5)^2+(16.3-18.5)^2 + (22.8-18.5)^2] \approx 7.88k^2 = 7.88 \times 10^6$$

$$S = \sqrt{S^2} = 2.81 \times 10^3$$

(2) sort the sample: 14.8, 15.8, 16.3, 17.2, 17.2, 18.5, 18.6, 21.7, 22.1, 22.8

lower quartile: $0.25 \times 10 = 2.5$ $\quad \lfloor 2.5 \rfloor + 1 = 3$

$$Q_{0.25} = 16.3k = 1.63 \times 10^4$$

upper quartile: $0.75 \times 10 = 7.5$ $\quad \lfloor 7.5 \rfloor + 1 = 8$

$$Q_{0.75} = 21.7k = 2.17 \times 10^4$$

interquartile range: $Q_{0.75} - Q_{0.25} = 5.4 \times 10^3$

T2. (1) $f_{max}(x) = 3f(x)[F(x)]^2 = \frac{3x^2}{\theta^3}, \ 0<x<\theta$

$$E(\tfrac{4}{3}X_{(3)}) = \frac{4}{3}\frac{3}{\theta^3}\int_0^\theta x^3 dx = \frac{4}{\theta^3}\left[\frac{x^4}{4}\right]_0^\theta = \theta$$

$\hat{\theta}_1$ is unbiased estimator of θ

$$f_{min}(x) = 3f(x)[1-F(x)]^2 = \frac{3}{\theta}(1-\frac{x}{\theta})^2, \ 0<x<\theta$$

$$E(4X_{(1)}) = 4\frac{3}{\theta}\int_0^\theta x(1-\frac{x}{\theta})^2 d\theta = \frac{12}{\theta}\left[\frac{\theta^2}{2} - \frac{2\theta^2}{3} + \frac{\theta^2}{4}\right] = \theta$$

$\hat{\theta}_2$ is unbiased estimator of θ

(2) $Var(\hat{\theta}_1) = Var(\frac{4}{3}X_{(3)}) = (\frac{4}{3})^2 Var(X_{(3)}) = \frac{16}{9}\{E(X_{(3)}^2) - E(X_{(3)})^2\}$

$$E(X_{(3)}^2) = \int_0^\theta x^2 \frac{3}{\theta^3} x^2 dx = \frac{3}{\theta^3}\left[\frac{x^5}{5}\right]_0^\theta = \frac{3}{5}\theta^2$$

$\Rightarrow Var(\hat{\theta}_1) = \frac{16}{9}(\frac{3}{5}\theta^2 - (\frac{3}{4}\theta)^2) = \frac{1}{15}\theta^2$

$Var(\hat{\theta}_2) = 16\, Var(X_{(1)}) = 16\{E(X_{(1)}^2) - E(X_{(1)})^2\}$

$E(X_{(1)}^2) = \frac{3}{\theta}\int_0^\theta x^2(1-\frac{x}{\theta})^2 dx = \frac{3}{\theta}(\frac{x^3}{3} - \frac{x^4}{2\theta} + \frac{x^5}{5\theta^2})]_0^\theta = \frac{1}{10}\theta^2$

$\Rightarrow Var(\hat{\theta}_2) = 16(\frac{1}{10}\theta^2 - (\frac{1}{4}\theta)^2) = \frac{3}{5}\theta^2$

$\Rightarrow Var(\hat{\theta}_1) < Var(\hat{\theta}_2)$

$\hat{\theta}_1$ is more efficient.

T3. for Chebyshev's inequality, $P(|x-\mu| \geq k\sigma) \leq \frac{1}{k^2}$

$\mu = 1.3\times10^9$, $\sigma = 0.7\times10^9$

$9.4\times10^9 = \mu + 3\sigma$, $5.2\times10^9 = \mu - 3\sigma \Rightarrow k=3$

$P(|x-\mu| \leq 3\sigma) = 1 - P(|x-\mu| \geq 3\sigma) \geq 1 - \frac{1}{3^2} = 88.9\%$

$\Rightarrow$ lower bound is 88.9%

T4. $\hat{\mu} = \frac{2}{n(n+1)}\sum_{k=1}^n kX_k = \frac{2}{n(n+1)}\sum_{k=1}^n k\mu$

$= \frac{2}{n(n+1)}\frac{(1+n)n}{2}\mu = \mu \Rightarrow$ unbiasedness

$Var(\hat{\mu}) = Var(\frac{2}{n(n+1)}\sum_{k=1}^n kX_k) = [\frac{2}{n(n+1)}]^2 Var(\sum_{k=1}^n kX_k)$

for X_i i.i.d $\Rightarrow Var(\hat{\mu}) = \frac{4}{n^2(n+1)^2}\sum_{k=1}^n Var(kX_k)$

$= \frac{4}{n^2(n+1)^2}\sum_{k=1}^n Var(X_k)\cdot k^2 = \frac{4}{n^2(n+1)^2}\sum_{k=1}^n k^2\sigma^2$

$= \frac{4}{n^2(n+1)^2}\frac{n(n+1)(2n+1)}{6}\sigma^2$

$= \frac{2\sigma^2(2n+1)}{3n(n+1)}$ when $n\to\infty$, $Var(\hat{\mu})\to 0$

$\Rightarrow$ vanishing variance

$\Rightarrow$ consistency

T5. (1) $\mu_1 = \bar{X} = 3\theta + 7(1-\theta) = 7-4\theta$

$\Rightarrow \hat{\theta}_1 = \dfrac{7-\mu_1}{4}$

$\mu_1 = \frac{1}{8} \times (3+3+3+3+3+7+7+7) = \frac{9}{2}$

$\hat{\theta}_1 = \dfrac{7-\frac{9}{2}}{4} = \dfrac{5}{8}$

(2) $E(\hat{\theta}_1) = E(\dfrac{7-\mu_1}{4}) = \dfrac{7-E(\bar{X})}{4} = \dfrac{7-(7-4\theta)}{4} = \theta$

$\Rightarrow \hat{\theta}_1$ is unbiased

$Var(\hat{\theta}_1) = \frac{1}{16} Var(\mu_1) = \frac{1}{16} Var(\bar{X})$

$= \dfrac{1}{16\times 8} Var(X) = \dfrac{1}{128}\left(E(X^2) - E^2(X)\right) = \dfrac{1}{128} \times \left(9\theta + 49(1-\theta) - (7-4\theta)^2\right)$

$= \dfrac{1}{128}(16\theta - 16\theta^2) = \dfrac{\theta-\theta^2}{8}$

$\hat{\theta}_1$ is unbiased estimator

(3) $L(\theta; x) = \theta^5 (1-\theta)^3$

$l(\theta; x) = 5\log\theta + 3\log(1-\theta)$

$\dfrac{dl(\theta;x)}{d\theta} = \dfrac{5}{\theta} - \dfrac{3}{1-\theta} = 0$

$\Rightarrow \hat{\theta} = \frac{5}{8}$

the maximum likehood estimate of θ is $\frac{5}{8}$

T6. (1) $\mu_1 = \bar{x} = \int_0^\infty \frac{1}{\theta} e^{-\frac{x}{\theta}} \cdot x\,dx = \int_0^\infty (-x)\,d(e^{-\frac{x}{\theta}})$

$= (-x \cdot e^{-\frac{x}{\theta}})\Big]_0^\infty - \int_0^\infty e^{-\frac{x}{\theta}}\, d(-x)$

$= (-\theta e^{-\frac{x}{\theta}})\Big]_0^\infty = \theta$

$\Rightarrow \mu_1$ is an unbiased estimator of θ

$\hat{\theta}_1 = \bar{X}$, and based on the sample

$\hat{\theta}_1 = \dfrac{150}{10} = 15$

(2) $\sigma(\hat{\theta}_1) = \sqrt{\frac{Var(X)}{n}} = \frac{1}{\sqrt{n}}\sqrt{Var(X)}$

SE

$Var(X) = E(X^2) - E^2(X)$

$E(X^2) = \int_0^{\infty} \frac{1}{\theta} e^{-\frac{x}{\theta}} x^2 dx = \int_0^{\infty} d(e^{-\frac{x}{\theta}})(-x^2) = e^{-\frac{x}{\theta}}(-x^2)\Big]_0^{\infty} + \int_0^{\infty} e^{-\frac{x}{\theta}} 2x dx$

$= 2\theta \int_0^{\infty} \frac{1}{\theta} e^{-\frac{x}{\theta}} x dx = 2\theta \bar{X} = 2\theta^2$

$\Rightarrow Var(X) = 2\theta^2 - \theta^2 = \theta^2$

$\sigma(\hat{\theta}_1) = \frac{\theta}{\sqrt{n}} = \frac{\theta}{\sqrt{10}}$

(3) $L(\theta; x) = \prod_{i=1}^{n} \frac{1}{\theta} e^{-\frac{x_i}{\theta}} = \left(\frac{1}{\theta}\right)^n e^{-\frac{\sum_{i=1}^{n} x_i}{\theta}}$

$\Rightarrow l(\theta; x) = -n\log\theta - \frac{\sum_{i=1}^{n} x_i}{\theta}$

$\frac{d\,l(\theta; x)}{d\theta} = -\frac{n}{\theta} + \frac{\sum_{i=1}^{n} x_i}{\theta^2} = 0 \Rightarrow \hat{\theta}_2 = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{X}$

based on the sample, $\hat{\theta}_2 = 15$

T7. (1) $Z_{0.025} = 1.96$, $\bar{x} = 42$ $\sigma = 5$

the 95% CI of μ is $\bar{x} \pm Z_{0.025}\frac{\sigma}{\sqrt{n}} \approx 42 \pm 1.96 \times \frac{5}{\sqrt{64}}$

$= (40.775, 43.225)$ min

(2) $\mu = 40$, $X \sim N(40, 25)$

$P(40.775 < X < 43.225) = P\left(\frac{40.775-40}{5} < \frac{X-40}{5} < \frac{43.225-40}{5}\right)$

$= P(0.155 < Z < 0.645) = \Phi(0.645) - \Phi(0.155)$

$= 0.7422 - 0.5636 = 0.1786$

T8. Let X be the height of woman aged 18 to 25 in region A,

Y be the height of woman aged 18 to 25 in region B.

$\bar{X} - \bar{Y} \overset{approx.}{\sim} N\left(\mu_X - \mu_Y, \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}\right)$

$\mu_X - \mu_Y = 0.02$ $\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m} = \frac{0.2^2}{40} + \frac{0.4^2}{50} = 0.0042$

$Z_{0.05} = 1.645$ the 90% CI of $\bar{X}-\bar{Y}$ is:

$$\left(\bar{X}-\bar{Y} - Z_{0.05}\sqrt{\frac{S_X^2}{n}+\frac{S_Y^2}{m}},\ \bar{X}-\bar{Y} + Z_{0.05}\sqrt{\frac{S_X^2}{n}+\frac{S_Y^2}{m}}\right) = 0.02 \pm 1.645 \times 0.0042$$

$$= (0.013,\ 0.027)\ m$$