STA219 Assignment 7

12312706 Zhou Liangyu

1. (1) $\because X \sim \text{Bernoulli}(p)$

$\therefore \mathrm{E}(X_i) = p,\ \mathrm{Var}(X_i) = p(1-p).$

$\therefore \mathrm{E}(\overline{X}) = \mathrm{E}(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}(X_i) = p,\ \mathrm{Var}(\overline{X}) = \mathrm{Var}(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(X_i) = \frac{p(1-p)}{n}.$

$\therefore \mathrm{E}(\overline{X}^2) = \mathrm{Var}(\overline{X}) + [\mathrm{E}(\overline{X})]^2 = \frac{p(1-p)}{n} + p^2.$

(2) $\because \mathrm{E}(\overline{X}^2) = \frac{p(1-p)}{n} + p^2$

$\therefore p^2 = \frac{n\mathrm{E}(\overline{X}^2) - p}{n-1} = \frac{n\mathrm{E}(\overline{X}^2) - \mathrm{E}(\overline{X})}{n-1}.$

$\therefore \hat{p}^2 = \frac{n\overline{X}^2 - \overline{X}}{n-1}$ is an unbiased estimator of p^2.

2. $\mu = \mathrm{E}(X) = 7.3 \times 10^9,\ \sigma = 0.7 \times 10^9.$

$\because 7.3 \times 10^9 - 5.2 \times 10^9 = 9.4 \times 10^9 - 7.3 \times 10^9 = 2.1 \times 10^9$

$\therefore k = \frac{2.1 \times 10^9}{0.7 \times 10^9} = 3.$

$\because P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$

$\therefore P(|X-\mu| < k\sigma) > 1 - \frac{1}{k^2} = 1 - \frac{1}{9} = \frac{8}{9}.$

$\therefore$ The lower bound for the probability that the white blood cell count per liter of blood is between 5.2×10^9 and 9.4×10^9 is $\frac{8}{9}$.

3. $\because \mathrm{E}(\hat{\mu}) = \mathrm{E}(\frac{2}{n(n+1)}\sum_{k=1}^{n} kX_k) = \frac{2}{n(n+1)}\sum_{k=1}^{n} k\mathrm{E}(X_k) = \frac{2}{n(n+1)} \cdot \frac{(n+1)n}{2} \cdot \mu = \mu.$

$\therefore \hat{\mu}$ is an unbiased estimator of μ.

Let $\sigma^2 = \mathrm{Var}(X)$,

then $\mathrm{Var}(\hat{\mu}) = \mathrm{Var}(\frac{2}{n(n+1)}\sum_{k=1}^{n} kX_k) = \frac{4}{n^2(n+1)^2}\sum_{k=1}^{n} k^2\mathrm{Var}(X_k) = \frac{4}{n^2(n+1)^2} \cdot \frac{n(n+1)(2n+1)}{6} \cdot \sigma^2 = \frac{2\sigma^2(2n+1)}{3n(n+1)}.$

$\because \mathrm{Var}(\hat{\mu}) = \frac{2\sigma^2(2n+1)}{3n(n+1)} \to \frac{4\sigma^2 n}{3n^2} \to 0$ as $n \to \infty$

$\therefore \hat{\mu}$ is an consistent estimator of μ.

4. (1) $\because \mu_1 = \mathrm{E}(X) = 3\theta + 7(1-\theta) = 7 - 4\theta \Rightarrow \hat{\theta} = \frac{7 - M_1}{4} = \frac{7 - \overline{X}}{4},\ \overline{X} = \frac{5 \times 3 + 3 \times 7}{8} = \frac{9}{2}.$

$\therefore$ The moment estimate of θ is $\hat{\theta}_1 = \frac{7 - \overline{X}}{4} = \frac{5}{8}.$

(2) $\because \mathrm{E}(X^2) = 3^2\theta + 7^2(1-\theta) = 49 - 40\theta$

$\therefore \mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = (49 - 40\theta) - (7 - 4\theta)^2 = 16\theta - 16\theta^2,\ \mathrm{Var}(\overline{X}) = \frac{\mathrm{Var}(X)}{8} = 2\theta - 2\theta^2.$

$\therefore \mathrm{E}(\hat{\theta}_1) = \mathrm{E}(\frac{7-\overline{X}}{4}) = \frac{7 - \mathrm{E}(\overline{X})}{4} = \frac{7 - (7 - 4\theta)}{4} = \theta,\ \mathrm{Var}(\hat{\theta}_1) = \mathrm{Var}(\frac{7-\overline{X}}{4}) = \frac{\mathrm{Var}(\overline{X})}{16} = \frac{\theta - \theta^2}{8}.$

$\therefore \hat{\theta}_1$ is an unbiased estimator.

(3) $L(\theta; x) = \theta^5(1-\theta)^3 \to l(\theta; x) = \log L(\theta; x) = 5\log\theta + 3\log(1-\theta).$

$$0=\frac{dl(\theta;x)}{d\theta}=\frac{5}{\theta}-\frac{3}{1-\theta}=\frac{5-8\theta}{\theta(1-\theta)}\to\hat{\theta}=\frac{5}{8}.$$

$\therefore$ The maximum likelihood estimate of θ is $\hat{\theta}=\frac{5}{8}$.

5. (1) $f(x;\theta)=\frac{1}{\theta}e^{-\frac{x}{\theta}}\to X\sim \mathrm{Exp}(\theta),\ \mathrm{E}(X)=\theta,\ \mathrm{Var}(X)=\theta^2.$

$\because \mu_1=\mathrm{E}(X)=\theta.$

$\therefore \hat{\theta}_1=\overline{X}=\frac{150}{10}=15.$

(2) $\because \mathrm{SD}(X)=\sqrt{\mathrm{Var}(X)}=\theta.$

$\therefore$ The standard error of $\hat{\theta}_1$ is $\sigma(\hat{\theta}_1)=\mathrm{SD}(\hat{\theta}_1)=\mathrm{SD}(\overline{X})=\frac{\mathrm{SD}(X)}{\sqrt{n}}=\frac{\sqrt{10}}{10}\hat{\theta}_1\approx 4.74.$

(3) $L(\theta;x)=\prod_{i=1}^{n}\frac{1}{\theta}e^{-\frac{x_i}{\theta}}=\frac{1}{\theta^n}e^{-\frac{\sum_{i=1}^{n}x_i}{\theta}}\to l(\theta;x)=\log L(\theta;x)=-n\log\theta-\frac{\sum_{i=1}^{n}x_i}{\theta}.$

$$0=\frac{dl(\theta;x)}{d\theta}=-\frac{n}{\theta}+\frac{\sum_{i=1}^{n}x_i}{\theta^2}\to\hat{\theta}=\frac{1}{n}\sum_{i=1}^{n}X_i=\overline{X}=15.$$

$\therefore$ The maximum likelihood estimate of θ is $\hat{\theta}=15$.

6. (1) Let X denote the population of the installation time(in minute). $\mu=42$, $\sigma=5$, $\alpha=0.05$.

$\because n=64>30$

$\therefore$ By the CLT, $\overline{X}\overset{\text{approx.}}{\sim}N(\mu,\frac{\sigma^2}{n})=N(42,\frac{25}{64}).$

$\because$ From the standard normal distribution table, $z_{\frac{\alpha}{2}}=z_{0.025}=1.96$

$\therefore$ A large sample 95% CI of $\overline{X}$ is $(\overline{X}-z_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}},\overline{X}+z_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}})=(42-1.96\times\frac{5}{8},42+1.96\times\frac{5}{8})=(40.775,43.225).$

(2) $X\sim N(40,25).$

$$\begin{aligned}\therefore P(40.775<X<43.225)&=P(\frac{40.775-40}{5}<\frac{X-\mu}{\sigma}<\frac{43.225-40}{5})\\&=P(0.155<Z<0.645)\\&=\frac{\Phi(0.64)+\Phi(0.65)}{2}-\frac{\Phi(0.15)+\Phi(0.16)}{2}\\&=\frac{0.7389+0.7422}{2}-\frac{0.5596+0.5636}{2}\\&=0.17895.\end{aligned}$$

$\therefore$ The probability that the installation time will be within the interval computed in (1) is 17.895%.

7. Let X and Y denote the height of a woman in region A and B. $\overline{X}=1.64,\ \overline{Y}=1.62,\ S_X=0.2,\ S_Y=0.4.$

$\because n=40>30,\ m=50>30,\ \alpha=0.1$

$\therefore$ By the CLT, $\overline{X}-\overline{Y}\overset{\text{approx.}}{\sim}N(\mu_X-\mu_Y,\frac{S_X^2}{n}+\frac{S_Y^2}{m})=N(\mu_X-\mu_Y,0.0042),\ z_{\frac{\alpha}{2}}=z_{0.05}=1.645.$

$\therefore$ A large amount 90% CI of $\mu_X-\mu_Y$ is

$$(\overline{X}-\overline{Y}-z_{\frac{\alpha}{2}}\sqrt{\frac{S_X^2}{n}+\frac{S_Y^2}{m}},\overline{X}-\overline{Y}+z_{\frac{\alpha}{2}}\sqrt{\frac{S_X^2}{n}+\frac{S_Y^2}{m}})=(1.64-1.62)\pm1.645\times\sqrt{\frac{0.2^2}{40}+\frac{0.4^2}{50}})\approx(-0.0866,0.1266).$$