Chapter 5

Huy Nguyen, Hoang Nguyen

February 28, 2020

Problem 1.

1.

$$S_n = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X_n})^2 = \frac{1}{n-1}(\sum_{i=1}^{n} X_i^2 - 2\bar{X_n}\sum_{i=1}^{n} X_i + \sum_{i=1}^{n}\bar{X_n}^2) = \frac{1}{n-1}(\frac{1}{n-1}(X_i^2 - 2n\bar{X_n}\bar{X_n} + n\bar{X_n}^2)$$

$$= \frac{\sum_{i=1}^{n} X_i^2}{n-1} - \frac{n}{n-1}\bar{X_n}^2$$

a)

We know

$$\begin{cases} \mathbb{E}[X_i^2] = \mathbb{V}(X_i) + \mathbb{E}[X_i]^2 = \sigma^2 + \mu^2 \\ \mathbb{E}[\bar{X_n}^2] = \mathbb{V}(\bar{X_n}) + \mathbb{E}[\bar{X_n}]^2 = \frac{\sigma^2}{n} + \mu^2 \end{cases} \tag{1}$$

$$\Rightarrow \mathbb{E}(S_n^2) = \frac{n\sigma^2 + n\mu^2 - n\left(\frac{\sigma^2}{n} + \mu^2\right)}{n-1} = \sigma^2$$

b)

$$S_n^2 = \frac{\sum_{i=1}^{n} X_i^2}{n-1} - \frac{n}{n-1}\bar{X_n}^2 = \frac{n}{n-1}\frac{\sum_{i=1}^{n} X_i^2}{n} - \frac{n}{n-1}\bar{X_n}^2$$

$$\begin{cases} \frac{n}{n-1} \to 1 \\ \frac{\sum_{i=1}^{n} X_i^2}{n} \xrightarrow{\text{P}} \sigma^2 + \mu^2 \text{ (Law of large number)} \\ \bar{X_n}^2 \xrightarrow{\text{P}} \mu^2 \text{ (Law of large number)} \end{cases} \tag{2}$$

$$\Rightarrow S_n^2 \xrightarrow{\text{P}} \sigma^2$$

We know $\mathbb{E}[X_i] = 1.p + 0.(1-p) = p$ and $\mathbb{E}[X_i^2] = 1.p + 0.(1-p) = p$. Hence, $\mathbb{V}[X_i] = \mathbb{E}[X_i^2] - (\mathbb{E}[X_i])^2 = p(1-p)$.

2. *Theorem*: If bias $\to 0$ ans se (standard error) $\to 0$ as n $\to \infty$ then $\hat{\theta}$ is consistent, that is, $\hat{\theta} \xrightarrow{\mathbf{P}} \theta$.

$\mathbb{V}[\bar{X}_i(1-\bar{X}_i)] = \mathbb{V}[\frac{\sum_{i=1}^{n} X_i}{n}(1 - \frac{\sum_{i=1}^{n} X_i}{n})] = \mathbb{V}[\frac{1}{n^2}\sum_{i=1}^{n} X_i\left(n - \sum_{i=1}^{n} X_i\right)]$

$$= \frac{1}{n^4}\mathbb{V}[\sum_{i=1}^{n} X_i\left(n - \sum_{i=1}^{n} X_i\right)]$$

$$= \frac{1}{n^4}\mathbb{V}\Big[\sum_{i=1}^{n} X_i \sum_{i=1}^{n}(1 - X_i)\Big]$$

$$= \frac{1}{n^4}\mathbb{V}\Big[\sum_{i=1}^{n} X_i(1-X_i) + \sum_{i\neq j} X_i(1-X_j)\Big]$$

$X_i \in 0,1 \Rightarrow \sum_{i=1}^{n} X_i(1-X_i) = 0$
Hence,

$$\mathbb{V}[\bar{X}_i(1-\bar{X}_i)] = \frac{1}{n^4}\mathbb{V}\Big[\sum_{i\neq j} X_i(1-X_j)\Big]$$

By X_i is IID:

$$\mathbb{V}[\bar{X}_i(1-\hat{\bar{X}}_i)] = \frac{1}{n^4}\sum_{i\neq j}\mathbb{V}[X_i]\mathbb{V}[1-X_i]$$

$$= \frac{1}{n^4}n(n-1)p(p-1)$$

The last equality tends to 0 as n $\to \infty$. So we have $\mathbb{V}[\bar{X}_i(1-\bar{X}_i)] \to 0$.
Hence, se$(\bar{X}_i(1-\bar{X}_i))$=$\sqrt{\mathbb{V}[\bar{X}_i(1-\bar{X}_i)]} \to 0$. (1)
By similar method, we can esaily get: $\mathbb{E}[\bar{X}_i(1-\bar{X}_i)] = \frac{1}{n^2}\sum_{i\neq j}\mathbb{E}[X_i(1-X_j)] = \frac{n(n-1)}{n^2}p(1-p) \to p(1-p)$ as n $\to \infty$.
Hence, bias$\Big(\bar{X}_i(1-\bar{X}_i)\Big) = \mathbb{E}[\bar{X}_i(1-\bar{X}_i)] - p(1-p) \to 0$ as n $\to \infty$ (2)
From (1), (2) and the theorem above, we get $\bar{X}_i(1-\bar{X}_i)$ is a consistent estimator of p(1-p).

3. We actually complete this exercise in previous solution.
bias$\Big(\bar{X}_i(1-\bar{X}_i)\Big) = \mathbb{E}[\bar{X}_i(1-\bar{X}_i)] - p(1-p) = \frac{n(n-1)}{n^2}p(1-p) - p(1-p)$.

4. To find an unbiased estimator, we have to find x such that $\frac{xn(n-1)}{n^2} = 1 \Rightarrow x = \frac{n}{n-1}$.
Hence, an unbiased estimator can be $\frac{n}{n-1}\bar{X}_i(1-\bar{X}_i)$

Problem 2.

$$\mathbb{E}[(\bar{X_n}-b)^2] = \mathbb{E}[(\bar{X_n}-\mathbb{E}[\bar{X_n}]+\mathbb{E}[\bar{X_n}]-b)^2] = \mathbb{V}(X_n) + 2(\mathbb{E}[\bar{X_n}]-b)\mathbb{E}(\bar{X_n}-\mathbb{E}[\bar{X_n}]) + (\mathbb{E}[X_n]-b)^2$$

$$\Rightarrow \mathbb{E}[(\bar{X_n}-b)^2] = \mathbb{V}(X_n) + (\mathbb{E}[X_n]-b)^2$$

Thus if $mathbbE[(\bar{X_n}-b)^2] \to 0$, then $\mathbb{V}(X_n) \to 0$ and $\mathbb{E}[X_n] \to 0$ (because they are non-negative). On the other hands, if $\mathbb{V}(X_n) \to 0$ and $\mathbb{E}[X_n] \to 0$, then $\mathbb{E}[(\bar{X_n}-b)^2] \to 0$

1. $(\mathbb{N}, (Poiss(\lambda))_{\lambda>0})$. This paramter is identified.
2. $(\mathbb{R}_+, (Exp(\lambda))_{10>\lambda>0})$. This parameter is identified.
3. $(\mathbb{R}_+, (Uni(0,\theta))_{\theta>0})$. This parameter is identified.
4. $(\mathbb{R}, (N(\mu,\sigma^2))_{(\mu,\sigma^2)\in\mathbb{R}\times\mathbb{R}_+})$. These parater are identified.
5.

$$\mathbb{P}(N(\mu,\sigma^2)>0) = \mathbb{P}\Big(N(0,1) > \frac{-\mu}{\sigma^2}\Big) = \phi(\frac{\mu}{\sigma^2})$$

Hence, the statistical model is: $(\{0,1\}, (Ber(\phi(\frac{\mu}{\sigma^2}))_{(\mu,\sigma^2)\in\mathbb{R}\times\mathbb{R}_+})$. This model depends on $\frac{\mu}{\sigma^2} \Rightarrow$ these parameters are not identified.

6. Same for 3.
7. Let X $\sim Exp(\lambda) \Rightarrow \mathbb{P}(X>20) = e^{-20\lambda}$. Hence, the statistical model is:

$$(\{0,1\}, (Ber(e^{-20\lambda}))_{\lambda>0})$$

This parameter is identified.

8. Let X $\sim Ber(p)$ such that:

$$\begin{cases} X_i = 1 \text{ if machine i has timelife less than 500 days} \\ X_i = 0 \text{ otherwise} \end{cases} \tag{3}$$

Hence:

$$p = \mathbb{P}(X_i = 1) = 1 - e^{-500\lambda}$$

The number of machines that have stopped working before 500 days is a binominal random variable with parameter $(67,\, 1 - e^{-500\lambda})$

The statistical model is $(\{1, 2, 3, .., 67\}, (Binominal(67, 1 - e^{-500\lambda}))_{\lambda>0})$. This parameter is identified.

Problem 3.

$$\begin{cases} \mathbb{E}[\bar{X_n}] = \mu \\ \mathbb{V}[\bar{X_n}] = \frac{\sigma^2}{n} \end{cases} \tag{4}$$

$$\Rightarrow \mathbb{E}[\bar{X_n}^2] = \frac{\sigma^2}{n} + \mu^2$$

$$\Rightarrow \mathbb{E}[(\bar{X_n} - \mu)^2] = \mathbb{E}[\bar{X_n}^2] - 2\mathbb{E}[\bar{X_n}]\mu + \mu^2 = \frac{\sigma^2}{n} + \mu^2 - 2\mu^2 + \mu^2 = \frac{\sigma^2}{n} \to 0$$

1. By central limit theorem (CLT), we have:

$$\frac{\sqrt{n}(\bar{X_i} - \mu)}{\sigma} \sim (N(0, 1))$$

Hence, $(a_n)_{n\in\mathbb{N}}$ can be $\frac{\sqrt{n}}{\sigma}$ and $(b_n)_{n\in\mathbb{N}}$ can be μ.

2. We have: $Z \backsim N(0, 1)$

Hence, $\mathbb{P}[|Z| \leqslant t] = \mathbb{P}[-t \leqslant Z \leqslant t] = \phi(t) - \phi(-t) = \phi(t) - (1 - \phi(t)) = 2\phi(t) - 1 = 2\mathbb{P}[Z \leqslant t] - 1$.

3. From part 1 we get:

$$\frac{\sqrt{n}(\bar{X_i} - \mu)}{\sigma} \sim (N(0, 1))$$

from part 2 we get:

$$\mathbb{P}[|Z| \leqslant t] = 2\mathbb{P}[Z \leqslant t] - 1$$

Substitution:

$$\mathbb{P}\Big[|\frac{\sqrt{n}(\bar{X_i} - \mu)}{\sigma}| \leqslant t\Big] = 2\mathbb{P}[Z \leqslant t] - 1$$

We have $2\mathbb{P}[Z \leqslant t] - 1 = 0.95 \Rightarrow t = \phi^{-1}(\frac{0.95+1}{2}) = 1.96$.

Hence,

$$\mathbb{P}\Big[|\frac{\sqrt{n}(\bar{X_i} - \mu)}{\sigma}| \leqslant 1.96\Big] = 0.95$$

$$\mathbb{P}\Big[-\frac{\sqrt{n}(\bar{X_i} - \mu)}{\sigma} \leqslant 1.96 \leqslant \frac{\sqrt{n}(\bar{X_i} - \mu)}{\sigma}\Big] = 0.95$$

Because X_i is Poisson random variable with parameter λ, so $\mu = \lambda$ and $\sigma = \sqrt{\lambda}$

We get:

$$\mathbb{P}\Big[-\frac{\sqrt{n}(\bar{X_i} - \lambda)}{\sqrt{\lambda}} \leqslant 1.96 \leqslant \frac{\sqrt{n}(\bar{X_i} - \lambda)}{\sqrt{\lambda}}\Big] = 0.95$$

$$\mathbb{P}\Big[\bar{X_i} - \frac{1.96\sqrt{\lambda}}{\sqrt{n}} \leqslant \lambda \leqslant \bar{X_i} + \frac{1.96\sqrt{\lambda}}{\sqrt{n}}\Big] = 0.95$$

We know: $\bar{X}_i \xrightarrow{P} \mathbb{E}[\bar{X}_i] = \lambda$
Hence,

$$\mathbb{P}\Big[\bar{X}_i - \frac{1.96\sqrt{\bar{X}_i}}{\sqrt{n}} \leqslant \lambda \leqslant \bar{X}_i + \frac{1.96\sqrt{\bar{X}_i}}{\sqrt{n}}\Big] \geqslant 0.95$$

$\Rightarrow$ L=$[\bar{X}_i - \frac{1.96\sqrt{\bar{X}_i}}{\sqrt{n}}, \bar{X}_i + \frac{1.96\sqrt{\bar{X}_i}}{\sqrt{n}}]$

4. We can easily see min$(X_i)\leqslant \bar{X}_i \leqslant$ max(X_i). Hence, a new interval can be:

$$L = [min(X_i) - \frac{1.96\sqrt{\bar{X}_i}}{\sqrt{n}}, max(X_i) + \frac{1.96\sqrt{\bar{X}_i}}{\sqrt{n}}]$$

Problem 4. We have:

$$\mathbb{E}[X] = \frac{1}{n}\big(1 - \frac{1}{n^2}\big) + n\frac{1}{n^2} = \frac{1}{n} - \frac{1}{n^3} + \frac{1}{n^2}$$

$$\Rightarrow \mathbb{P}(X_n > t) \leqslant \frac{\mathbb{E}[X]}{t} = \frac{\frac{1}{n} - \frac{1}{n^3} + \frac{1}{n^2}}{t} \to 0$$

Hence, $X_n \xrightarrow{\text{P}} 0$

$$\mathbb{E}[X_n^2] = \frac{1}{n^2}\big(1 - \frac{1}{n^2}\big) + 1 \to 1$$

Hence, X_n does not converge in quadratic mean.
We have X_i is IID. Hence, $\mathbb{P}(M_n \leqslant t) = \prod_{n=1}^{n} \mathbb{P}(X_i \leqslant t)$.
By uniform distribution, the CDF of M_n:

$$\mathbb{P}(M_n \leqslant t) = F(t) = \Big(\frac{t}{\theta}\Big)^n$$

Hence, the PDF of M_n is:

$$f(t) = \frac{dF}{dt} = n\theta^{-n}t^{n-1}$$

We can easily get:

$$\mathbb{E}[M_n] = \int_0^\theta tn\theta^{-n}t^{n-1}dt = \frac{n}{n+1}\theta \to \theta \textbf{ as n} \to \infty$$

By Markov's Inequality:

$$\mathbb{P}\Big[|M_n - \theta| > \epsilon\Big] \leqslant \mathbb{P}[M_n - \theta > \epsilon] \leqslant \frac{\mathbb{E}[M_n - \theta]}{\epsilon} = \frac{\mathbb{E}[M_n] - \theta}{\epsilon} \to 0$$

Hence, M_n converages in probility to θ.

2. From part 1 we get: M_n: $\mathbb{P}[M_n \leqslant t] = \Big(\frac{t}{\theta}\Big)^n$. Hence, CDF of $n(1 - \frac{M_n}{\theta})$ is:

$$P\Big[n(1 - \frac{M_n}{\theta}) \leqslant t\Big] = \mathbb{P}\Big[M_n \geqslant \frac{(n-t)\theta}{n}\Big] = 1 - \Big(\frac{n-t}{n}\Big)^n \to 1 - e^{-t} \textbf{ as n} \to \infty$$

Hence, $n(1 - \frac{M_n}{\theta})$ converages in distribution to an exponential random variable with parameter 1.

3. Let A is an exponential random variable with parameter 1. Because $n(1 - \frac{M_n}{\theta})$ converages in distribution to X, we have:

$$\mathbb{P}\Big[n(1 - \frac{M_n}{\theta}) \leqslant t\Big] \to \mathbb{P}[X \leqslant t] = 1 - e^{-t}$$

$1 - e^{-t} = 0.95 \Rightarrow t = 3$. We have:

$$\mathbb{P}\Big[n(1 - \frac{M_n}{\theta}) \leqslant 3\Big] \to 0.95$$

which is:

$$\mathbb{P}\Big[\theta \leqslant \frac{nM_n}{n-3}\Big] \to 0.95$$

On the other hand, we always have $\theta \geqslant M_n$ (uniform distribution). Hence, we get:

$$\mathbb{P}\Big[M_n \leqslant \theta \leqslant \frac{nM_n}{n-3}\Big] \to 0.95 \textbf{ as n} \to \infty$$

We conclude $L = \Big[M_n, \frac{nM_n}{n-3}\Big] = \Big[M_n, M_n + \frac{3M_n}{n-3}\Big]$.

4. $bias(M_n) = \mathbb{M}_{\ltimes} - \theta = \frac{n}{n+1}\theta - \theta \neq 0$. Hence, M_n is biased.

Problem 5. $X_i \sim \text{Ber(p)} \iff \frac{1}{n}\sum_{i=1}^{n} X_i^2 = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X_n} \Rightarrow$

$$\begin{cases} \bar{X_n} \xrightarrow{\text{P}} p \text{ Law of Large Number} \\ \bar{X_n} \xrightarrow{\text{qm}} p \text{ From proof of problem 3} \end{cases} \tag{5}$$

Problem 6. By Central Limit Theorem:

$$\frac{\sqrt{100}(\bar{X_n} - 68)}{2.6} \sim \mathbb{N}(0,1) \Longrightarrow \mathbb{P}(\bar{X_n} > 68) = \mathbb{P}(\mathbb{N}(0,1) > 0) = \frac{1}{2}$$

Problem 7.

$$X_n \sim Poisson(\lambda_n) \Rightarrow \mathbb{E}[X_n] = \mathbb{V}[X_n] = \lambda_n = \frac{1}{n}$$

a.)

$$\mathbb{P}(X_n > t) \leqslant \frac{\mathbb{E}[X_n]}{t} = \frac{1}{nt} \to 0$$

$$\Rightarrow X_n \xrightarrow{\text{P}} 0$$

b.)

$$\mathbb{P}(Y_n > t) = \mathbb{P}(X_n > \frac{t}{n}) = \sum_{i=\frac{t}{n}}^{\infty} \frac{\left(\frac{1}{n}\right)^i \exp(\frac{-1}{n})}{i!} \xrightarrow{n\to\infty} \exp(\frac{-1}{n}) \sum_{i=0}^{\infty} \frac{\left(\frac{1}{n}\right)^i}{i!} =$$

Problem 8. By Central Limit Theorem:

$$Z_n = \frac{\sqrt{100}(\bar{X_n} - 1)}{1} \sim \mathbb{N}(0,1)$$

$$\mathbb{P}(Y = \sum_{i=1}^{n} X_i < 90) = \mathbb{P}(\bar{X_n} < 0.9) = P(Z_n < -1) = \phi(-1.01)$$

Problem 9.

$$\mathbb{P}(|X_n - X| > t) = \mathbb{P}(X_n \neq X) = \mathbb{P}(X_n = \exp(n)) = \frac{1}{n} \longrightarrow 0$$

$$\longrightarrow X_n \xrightarrow{(P)} X \longrightarrow X_n \rightsquigarrow X$$

$$\mathbb{E}[(X - X_n)^2] = (\exp(n) - 1)^2 \frac{1}{2}\frac{1}{n} + (\exp(n) = 1)^2 \frac{1}{2}\frac{1}{n} = \frac{\exp(2n) + 1}{2n} \to \infty$$

Problem 11.

$$X_n \sim \mathbb{N}(0, \frac{1}{n}) \Rightarrow \sqrt{n}X_n \sim \mathbb{N}(0,1)$$

$$\mathbb{P}(|X_n - X| > t) = \mathbb{P}(|\sqrt{n}X_n - \sqrt{n}X| > \sqrt{n}t) \leqslant \frac{\mathbb{E}[|\mathbb{N}(0,1) - \sqrt{n}X|]}{\sqrt{n}t} = \frac{\mathbb{E}[\mathbb{N}(-\sqrt{n}X, 1)]}{\sqrt{n}t} = \frac{-\sqrt{n}X}{\sqrt{n}t} = \frac{-X}{t} < 0 \text{ (disprove)}$$

Problem 12. if $X_n \rightsquigarrow X$, then :

$$F_n(k) \to F(k) \Leftrightarrow$$

$$\begin{cases} \mathbb{P}(X_n = k) - \mathbb{P}(X_n = k-1) \to \mathbb{P}(X = k) - \mathbb{P}(X = k-1) \\ \mathbb{P}(X_n = k-2) - \mathbb{P}(X_n = k-2) \to \mathbb{P}(X = k-1) - \mathbb{P}(X = k-2) \\ ... \\ ... \\ ... \\ \mathbb{P}(X_n = 0) = \mathbb{P}(X = 0) = 0 \end{cases} \tag{6}$$

Add those equations above we get:

$$\Rightarrow \mathbb{P}(X_n = k) \to \mathbb{P}(X = k)$$

if $\mathbb{P}(X_n = k) \to \mathbb{P}(X = k)$, then:

$$\begin{cases} \mathbb{P}(X_n = k) \to \mathbb{P}(X = k) \\ \mathbb{P}(X_n = k-1) \to \mathbb{P}(X = k-1) \\ ... \\ ... \\ ... \end{cases} \tag{7}$$

Add those equations above we get:

$$F_n(k) \to F(k)$$