

Assignment 1

1. (a) $P(X=1; a, b) = \int_0^1 P(X=1|\theta) f(\theta; a, b) d\theta$

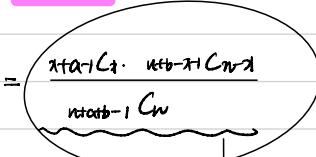
$$= \int_0^1 \binom{n}{1} \theta^1 (1-\theta)^{n-1} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1-\theta)^{b-1} \mathbb{1}_{0 \leq \theta \leq 1} d\theta$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \binom{n}{1} \int_0^1 \theta^{1+a-1} (1-\theta)^{n+b-1-1} \mathbb{1}_{0 \leq \theta \leq 1} d\theta$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \binom{n}{1} \underbrace{\frac{\Gamma(1+a)\Gamma(n+b-1)}{\Gamma(n+a)\Gamma(n+b)}}_{=1} \int_0^1 \theta^{1+a-1} (1-\theta)^{n+b-1-1} d\theta$$

$$= \binom{n}{1} \cdot \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(1+a)\Gamma(n+b-1)}{\Gamma(n+a)\Gamma(n+b)}$$

$$= \frac{n!}{1!(n-1)!} \cdot \frac{(a+b-1)!}{(a-1)!(b-1)!} \cdot \frac{(1+a-1)!(n+b-1-1)!}{(n+a-1)!}$$

$$= \frac{\binom{1+a-1}{1} \binom{n+b-1-1}{n-1}}{\binom{n+a+b-1}{n}} = \frac{\frac{n+a-1}{1} C_1 \cdot \frac{n+b-1}{n} C_{n-1}}{n+a+b-1 C_n}$$


(b) calculate Bayes Factor $\frac{P(X=1|M_1)}{P(X=1|M_2)}$

$$\rightarrow P(X=1|M_1) = \int_0^1 \binom{n}{1} \theta^1 (1-\theta)^{n-1} \mathbb{1}_{0 \leq \theta \leq 1} d\theta$$

$$= \binom{n}{1} \left(\frac{1}{2}\right)^n$$

$$\rightarrow P(X=1|M_2) = \frac{1}{N+1} \quad \text{← when } a=b=1$$

$$BF = \frac{\binom{10}{9} \left(\frac{1}{2}\right)^{10}}{\frac{1}{11}} = \frac{110}{1024} = \frac{55}{512}$$

2. (a) $E = \{1, 2, 3, 4\}$, $A = \{\{1\}, \{1, 2\}\}$ Find $\sigma(A)$.

$$\sigma(A) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}, \{3, 4\}, \{1, 2, 3, 4\}, \{1, 3, 4\}, \{1, 2, 3, 4\}\}$$

(b) 1. $ACC \Rightarrow \sigma(A) \subset \sigma(C)$

Let B be another collection of subsets of set E which belongs to C but not to A .

Then I can express C as below

$$\Rightarrow C = \{A, B\}$$

Then, $\sigma(C)$ is $\{\sigma(A), \sigma(B)\}$.

$$\sigma(A) \subset \{\sigma(A), \sigma(B)\}$$

$$= \sigma(C)$$

2. $A \subset \sigma(C)$, $C \subset \sigma(A) \Rightarrow \sigma(A) = \sigma(C)$



$\circ A \subset \sigma(C) \Rightarrow \sigma(A) \subset \sigma(\sigma(C)) = \sigma(C)$

$$\Rightarrow \sigma(A) \subset \sigma(C)$$

$\circ C \subset \sigma(A) \Rightarrow \sigma(C) \subset \sigma(\sigma(A)) = \sigma(A)$
 $\Rightarrow \sigma(C) \subset \sigma(A)$

$$\sigma(A) \subset \sigma(C), \sigma(C) \subset \sigma(A)$$

$$\Rightarrow \therefore \sigma(A) = \sigma(C)$$

3. $A \subset C \subset \sigma(A)$

$\boxed{A \subset C \Rightarrow \sigma(A) \subset \sigma(C) \text{ according to } \#1.}$

$$\boxed{C \subset \sigma(A) \Rightarrow \sigma(C) \subset \sigma(A)}$$

$$\Rightarrow \sigma(A) \subset \sigma(C) \subset \sigma(A)$$

$$\Rightarrow \therefore \sigma(A) = \sigma(C)$$

$$3. M_X(t) = \mathbb{E}[e^{tX}]$$

$$\begin{aligned}
 (a) \quad \mathbb{E}[e^{tX}] &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
 &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2} dx \\
 &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2 + \frac{1}{2}t^2} dx \\
 &= e^{\frac{1}{2}t^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x+t)^2} dx \quad x-t \Rightarrow z \\
 &= e^{\frac{1}{2}t^2} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz}_{=1} = \underbrace{e^{\frac{1}{2}t^2}}_{=1}
 \end{aligned}$$

$$(b) \quad Y_n := c(X_1 + X_2 + \dots + X_n)$$

$$\begin{aligned}
 \mathbb{E}[e^{tY}] &= \mathbb{E}(e^{ct(X_1 + \dots + X_n)}) \\
 &= \mathbb{E}(e^{cX_1} \cdot e^{cX_2} \cdots e^{cX_n}) \quad X_i's \text{ all i.i.d.}
 \end{aligned}$$

$$\begin{aligned}
 &= \underbrace{\mathbb{E}(e^{cX_1})}_{M_{X_1}(ct)} \underbrace{\mathbb{E}(e^{cX_2})}_{M_{X_2}(ct)} \cdots \underbrace{\mathbb{E}(e^{cX_n})}_{M_{X_n}(ct)}
 \end{aligned}$$

$$= \left\{ M_X(ct) \right\}^n \text{ because } X_1, \dots, X_n \text{ are i.i.d.}$$

(c) Additionally need to assume that all X_i 's are $\perp \!\! \perp$ as stated in (b). *

$$X \sim (\mu, \sigma^2)$$

$$Y_i := \frac{X_i - \mu}{\sigma}, \quad Z_n := \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i$$

$$\begin{aligned}
 M_{Z_n}(t) &= \mathbb{E}(e^{tZ_n}) = \mathbb{E}\left(e^{\frac{t}{\sqrt{n}} \sum_{i=1}^n Y_i}\right) \\
 &= \left\{ \mathbb{E}\left(e^{\frac{t}{\sqrt{n}} Y_i}\right) \right\}^n \quad \text{Let it as } O(\frac{t}{\sqrt{n}}) \\
 &= \left\{ \mathbb{E}\left(1 + \frac{tY_i}{\sqrt{n}} + \frac{1}{2!} \frac{t^2 Y_i^2}{n} + \frac{1}{3!} \frac{(tY_i)^3}{\sqrt{n}} + \dots\right) \right\}^n
 \end{aligned}$$

$$\mathbb{E}Y_i = \mathbb{E}\left(\frac{X_i - \mu}{\sigma}\right) = 0$$

$$\mathbb{E}Y_i^2 = \mathbb{E}\left(\frac{X_i - \mu}{\sigma}\right)^2 = \mathbb{E}\left(\frac{X_i^2 - 2\mu X_i + \mu^2}{\sigma^2}\right)$$

$$\begin{aligned}
 &= \mathbb{E} \left(\frac{\sigma^2 + M^2 - 2\mu^2 + \mu^2}{n} \right) \\
 &= 1
 \end{aligned}$$

$$= \left[1 + \frac{t^2}{2n} + O\left(\frac{t}{\sqrt{n}}\right) \right]^n$$

In statistics, O notation is used to represent really tiny terms, which are so small that they will not influence on the result of computation.

Using the fact that

$$\lim_{n \rightarrow \infty} \left(1 + \frac{t}{n}\right)^n = e.$$

$$\lim_{n \rightarrow \infty} \left(1 + \frac{k}{n}\right)^n = e^k.$$

$$M_{Z_n}(t) \rightarrow e^{\frac{1}{2}t^2} \text{ as } n \rightarrow \infty$$

This is same as the Mgf of standard normal dist'n.

$$\therefore Z_n \xrightarrow{d} N(0,1)$$

4. (a) $\lim_{n \rightarrow \infty} (\underbrace{|\hat{\sigma}_n^2 - \sigma^2|}_{\geq \varepsilon}) = 0$

$$\begin{aligned}\hat{\sigma}_n^2 &= \frac{1}{n} \sum (X_i - \bar{X}_n)^2 \\ &= \frac{1}{n} \sum \{(X_i - \mu) - (\bar{X} - \mu)\}^2 \\ &= \frac{1}{n} \sum \{ (X_i - \mu)^2 - 2(X_i - \mu)(\bar{X} - \mu) + (\bar{X} - \mu)^2 \} \\ &= \frac{1}{n} \sum (X_i - \mu)^2 - 2(\bar{X} - \mu) \frac{1}{n} \sum (X_i - \mu) + (\bar{X} - \mu)^2 \\ &= \sigma^2 - (\bar{X} - \mu)^2\end{aligned}$$

$$\begin{aligned}|\hat{\sigma}_n^2 - \sigma^2| &= |-(\bar{X} - \mu)^2| \\ &= (\bar{X}_n - \mu)^2\end{aligned}$$

Then, what I need to show is that $\lim_{n \rightarrow \infty} (\underbrace{(\bar{X}_n - \mu)^2}_{\geq \varepsilon}) = 0$

By Chebychev's inequality,

$$\lim_{n \rightarrow \infty} (\underbrace{(\bar{X}_n - \mu)^2}_{\geq \varepsilon}) \leq \frac{\sigma^2}{n\varepsilon}$$

$\Rightarrow 0 \text{ as } n \rightarrow \infty$

$\therefore \hat{\sigma}_n^2 \xrightarrow{P} \sigma^2 \text{ as } n \rightarrow \infty$

(b) $E(\hat{\sigma}_n^2)$

$$\begin{aligned}&= E\{\sigma^2 - (\bar{X} - \mu)^2\} = \sigma^2 - E(\bar{X} - \mu)^2 \\ &= \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n} \sigma^2\end{aligned}$$

$$\text{Bias} = E(\hat{\sigma}_n^2) - \sigma^2 = -\frac{\sigma^2}{n} \neq 0 \Rightarrow \text{If it is biased,}$$

(c) Letting $\hat{\sigma}_n^2 = \frac{1}{n-1} \sum (X_i - \bar{X}_n)^2$ is the solution.

$$\begin{aligned}E(\hat{\sigma}_n^2) &= E\left(\frac{n}{n-1} \cdot \frac{1}{n} \sum (X_i - \bar{X}_n)^2\right) \\ &= \frac{n}{n-1} \cdot E\left(\frac{1}{n} \sum (X_i - \bar{X}_n)^2\right) \\ &= \frac{n}{n-1} \cdot \frac{n-1}{n} \sigma^2 = \sigma^2\end{aligned}$$

$$\text{Bias} = \sigma^2 - \sigma^2 = 0 \Rightarrow \text{It is unbiased now.}$$

5. What I need to show is $X_n \xrightarrow{d} c$

For all $\varepsilon > 0$,

$$\bullet 1) F_{X_n}(c) = P(X_n \leq c, X \leq c+\varepsilon) + P(X_n \leq c, X > c+\varepsilon)$$

$$\leq P(X \leq c+\varepsilon) + P(X_n < X - \varepsilon)$$

$$= P(X \leq c+\varepsilon) + P(X - X_n > \varepsilon)$$

$$\leq P(X \leq c+\varepsilon) + P(|X_n - X| > \varepsilon)$$

idea:

$$= F_X(c+\varepsilon) + P(|X_n - X| > \varepsilon) \quad \text{with } X - \varepsilon < X_n < X + \varepsilon$$

$$\bullet 2) P(X \leq c-\varepsilon) = P(X \leq c-\varepsilon, |X_n - X| < \varepsilon) + P(X \leq c-\varepsilon, |X_n - X| \geq \varepsilon)$$

$$= P(X - \varepsilon < X_n < X + \varepsilon \leq c) + P(X \leq c-\varepsilon, |X_n - X| \geq \varepsilon)$$

$$\leq P(X_n < c) + P(|X_n - X| \geq \varepsilon)$$

$$= F_{X_n}(c) + P(|X_n - X| \geq \varepsilon)$$

$$\Rightarrow 1) : F_{X_n}(c) \leq F_X(c+\varepsilon) + P(|X_n - X| > \varepsilon)$$

$$\Rightarrow 2) : P(X \leq c-\varepsilon) - P(|X_n - X| \geq \varepsilon) \leq F_{X_n}(c)$$

$$1) \cup 2) \Rightarrow F_X(c-\varepsilon) - P(|X_n - X| \geq \varepsilon) \leq F_{X_n}(c) \leq F_X(c+\varepsilon) + P(|X_n - X| > \varepsilon)$$

as : $\lim_{n \rightarrow \infty}$

$\lim_{n \rightarrow \infty}$

$\lim_{n \rightarrow \infty}$

$$\Rightarrow \lim_{n \rightarrow \infty} \left\{ F_X(c-\varepsilon) - P(|X_n - X| \geq \varepsilon) \right\} \leq \lim_{n \rightarrow \infty} F_{X_n}(c) \leq \lim_{n \rightarrow \infty} \left\{ F_X(c+\varepsilon) + P(|X_n - X| > \varepsilon) \right\}$$

$\xrightarrow{\text{P}} \text{by definition}$

$\xrightarrow{\text{P}} \text{by definition}$

By Sandwich Theorem $\Rightarrow \lim_{n \rightarrow \infty} F_{X_n}(c) = F_X(c)$

$$\therefore \underline{X_n \xrightarrow{d} c}$$