1. CLT 복습

- $Y_1, \cdots, Y_n \sim iid\ f(y) = 2y,\ 0<y<1$

$S = Y_1 + \cdots + Y_{30}$, use CLT to approximate $P(S \le 22)$

$E(Y) = \frac{2}{3}$
$Var(Y) = \frac{1}{18}$

$E(S) = 30E(Y) = 30 \times \frac{2}{3} = 20$
$Var(S) = Var(Y_1 + \cdots + Y_{30}) \overset{iid}{=} 30\,Var(Y) = \frac{30}{18}$

$$S \overset{D}{\rightsquigarrow} N(20, \tfrac{30}{18})$$

$$\frac{S - E(S)}{\sqrt{Var(S)}} \overset{D}{\rightsquigarrow} N(0.1)$$

$$\frac{S-20}{\sqrt{30/18}} \overset{D}{\rightsquigarrow} N(0.1)$$

$$\therefore P(S \le 22) = P\left(\frac{S-20}{\sqrt{30/18}} \le \frac{22-20}{\sqrt{30/18}}\right)$$
$$= P(Z \le 1.5491)$$

2. efficiency

- efficiency of $\hat{\theta}_1$ related to $\hat{\theta}_2$ = $eff(\hat{\theta}_1, \hat{\theta}_2) = \frac{Var(\hat{\theta}_2)}{Var(\hat{\theta}_1)}$

외우기. 앞에 것이 분모가 된다!

ex) $Y_1, \cdots Y_n \sim iid\ N(\mu, \sigma^2)$

find efficiency of S^2 relative to $\hat{\sigma}^2$,

$$\hat{\sigma}^2 = \frac{1}{2}(Y_1 - Y_2)^2$$

$$eff(S^2, \hat{\sigma}^2) = \frac{Var(\hat{\sigma}^2)}{Var(S^2)}$$

i) $Var(S^2)$

$$Var(S^2) = Var\left(\frac{\sigma^2}{n-1} \times \underbrace{\frac{(n-1)S^2}{\sigma^2}}_{\text{우리가 아는 분포로 만들어주기}}\right) = \left(\frac{\sigma^2}{n-1}\right)^2 Var\left(\frac{(n-1)S^2}{\sigma^2}\right)$$

σ는 상수로 고정. S는 sample을 뽑을 때마다 변하는 값!

$\rightarrow Var(cX) = c^2 Var(X)$

$\chi^2_{(n-1)}$

카이제곱 평균 : 자유도
분산 : 자유도 X 2

$$= \left(\frac{\sigma^2}{n-1}\right)^2 \times 2(n-1)$$

$$= \frac{2}{n-1}\sigma^4$$

ii) $Var(\hat{\sigma}^2) = Var\left(\frac{1}{2}(Y_1 - Y_2)^2\right)$

그냥 표준화 not CLT

$(Y_1 - Y_2) \sim N(0, 2\sigma^2)$

$$\frac{(Y_1 - Y_2) - 0}{\sqrt{2\sigma^2}} \sim N(0,1)$$

$Y_1 \sim N(\mu_1, \sigma_1^2)$
$Y_2 \sim N(\mu_2, \sigma_2^2)$
$\downarrow$
$Y_1 + Y_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
$Y_1 - Y_2 \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$

$$\therefore Z = \frac{(Y_1 - Y_2)}{\sqrt{2\sigma^2}}$$

$$Z^2 = \frac{(Y_1 - Y_2)^2}{2\sigma^2}$$

$$(Y_1 - Y_2)^2 = (2\sigma^2)Z^2$$

$$Var(\hat{\sigma}^2) = Var\left(\frac{1}{2} \times 2\sigma^2 \cdot Z^2\right) = \sigma^4 \, Var(Z^2)$$

$\chi^2_{(1)}$

카이제곱 평균 : 자유도
분산 : 자유도 X 2

$$= \sigma^4 \times 2$$

$\chi^2_{(n)} = Z_1^2 + Z_2^2 + \cdots + Z_n^2$

$$\therefore eff(S^2, \hat{\sigma}^2) = \frac{Var(\hat{\sigma}^2)}{Var(S^2)} = \frac{2\sigma^4}{\frac{2}{(n-1)}\sigma^4} = n-1$$

☆ 매우 중요함

3. Consistency $(n \rightarrow \infty)$ 문제 : show @@ is constant!

암기하기

- Tool 1. WLLN (Weak Law of Large Numbers)
- Tool 2. Theorems on Limiting distribution
- Tool 3. $\hat{\theta}_n$ is U.E and $V(\hat{\theta}_n) \rightarrow 0$, $\hat{\theta}_n$ is consistent for θ

• Tool 1. WLLN

$Y_1, \cdots, Y_n \sim iid \quad E(Y_i) = \mu < \infty$

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow{p} \mu$$

$$\frac{1}{n}\sum_{i=1}^{n} \boxed{Y_i} \xrightarrow{p} E(Y) = \mu$$

<tip>
이 모양을 외워두기
$\frac{1}{n}\Sigma$ 이런 꼴이 나오면
tool 1을 사용할 준비하기

ex) $\frac{1}{n}\sum_{i=1}^{n} Y_i^2 \xrightarrow{p} E(Y^2) = \mu^2 + \sigma^2$

$\frac{1}{n}\sum_{i=1}^{n} \ln Y_i \xrightarrow{p} E(\ln y)$

• Tool 2. Theorems on Limiting Distribution

$W_n \xrightarrow{p} a, \quad V_n \xrightarrow{p} b \quad$ then

① $cW_n + dW_n \xrightarrow{p} ca + db$
② $W_n \times V_n \xrightarrow{p} ab$
③ $\frac{W_n}{V_n} \xrightarrow{p} \frac{a}{b}$
④ $h(W_n) \xrightarrow{p} h(a)$

ex) show $S_n^2 \xrightarrow{p} \sigma^2$

$$S_n^2 = \frac{1}{n-1}\Sigma(Y_i - \bar{Y}_n)^2 = \frac{1}{n-1}\left[\Sigma Y_i^2 - n(\bar{Y}_n)^2\right]$$

tool1 생각해보기

$$= \left(\frac{n}{n-1} \times \boxed{\frac{1}{n}\Sigma Y_i^2} - \frac{n}{n-1}\boxed{(\bar{Y}_n)^2}\right) \xrightarrow{p} \mu^2 + \sigma^2 - \mu^2 = \sigma^2$$

$\frac{n}{n-1} \to 1$

$\frac{1}{n}\Sigma Y_i^2 \xrightarrow{p} E(Y^2)$ tool 1

$\frac{n}{n-1} \to 1.$

$(\bar{Y}_n)^2 \xrightarrow{p} \mu^2$ tool 1 tool 2

$\bar{Y} \xrightarrow{p} \mu$ tool 1

$\bar{Y}^2 \xrightarrow{p} \mu^2$ tool 2

☆ WLLN 잘 활용해야한다

$S_n^2 \xrightarrow{p} \sigma^2$

↳ $S_n \xrightarrow{p} \sigma$ (by tool 2)

- Tool 3 $\underbrace{\hat\theta_n \text{ is U·E}}_{\text{조건 ①}}$ and $\underbrace{Var(\hat\theta_n) \to 0}_{\text{조건 ②}}$, $\hat\theta_n$ is constant for θ

  $\hat\theta_n \xrightarrow{p} \theta$

ex) $Y_1, \cdots, Y_n \sim iid\ Unif(0, \theta)$

① show $2\bar{Y}_n \xrightarrow{p} \theta$

tool 1.

$\bar{Y}_n \xrightarrow{p} E(Y) = \frac{\theta}{2}$ by tool1 (WLLN)

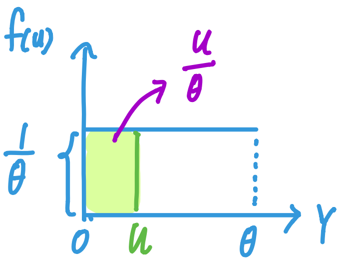
$2 \cdot \bar{Y}_n \xrightarrow{p} \theta$ by tool2

② show $\frac{n+1}{n} \max(Y_i) \xrightarrow{p} \theta$

직관적으로 tool 1을 쓰기 어려워보임.

i) $U = \max(Y_i)$

unbiased 보이려면 기댓값을 알아야하고 분포(pdf)를 뽑아내야한다.



기댓값 = $\frac{\theta}{2}$

$F_U(u) = Pr[U \le u] = P[\max(Y_i) \le u]$

$= P[Y_1 \le u, Y_2 \le u, \cdots, Y_n \le u]$

$= P[Y < u]^n$

$F_U(u) = \left(\frac{u}{\theta}\right)^n$

$f_U(u) = \frac{\partial}{\partial u} F_U(u) = n\left(\frac{u}{\theta}\right)^{n-1} \cdot \frac{1}{\theta}$

$= \frac{n}{\theta^n} u^{n-1}, \quad 0 < u < \theta$

$E(U) = \int_0^\theta u \cdot f(u)\, du = \int_0^\theta \frac{n}{\theta^n} u^n\, du = \frac{n}{n+1}\theta$

$E(aU) = aE(U)$이므로

$E\left[\frac{n+1}{n} U\right] = \frac{n+1}{n} E(U) = \theta$

결론: $\frac{n+1}{n} \max Y_i$가 θ에 대해 unbiased estimator이다.

ii) $Var\left(\frac{n+1}{n}U\right) \xrightarrow{n\to\infty} 0$

$E(U^2) - [E(U)]^2$

$\int_0^\theta u^2 f_U(u)\,du$ 구하기

$Var\left(\frac{n+1}{n}U\right) = \frac{1}{n(n+2)}\theta^2 \xrightarrow{n\to\infty} 0$

by tool 3

4. Convergence

· $\lim_{n\to\infty} F_n(x) = F(x)$, then $X_1, \cdots, X_n \xrightarrow{D} X$

☆ ① CLT
② mapping theorem
☆☆☆ ③ Slutsky's theorem

세트로 알아두기.

· Slutsky's theorem

$U_n \xrightarrow{D} U$ and $W_n \xrightarrow{P} 1$ then

$\frac{U_n}{W_n} \xrightarrow{D} U$

WLLN 잘 활용하기

☆ t분포 $\xrightarrow{D}$ z분포

$T_{(n)} \overset{\text{정의}}{=} \frac{Z}{\sqrt{\frac{u}{n}}}$, when $Z \sim N(0,1)$, $U \sim \chi^2(n)$

$\sim Z$

$\xrightarrow{P} 1$???

$U = Z_1^2 + Z_2^2 + \cdots + Z_n^2$ (χ^2 정의)

카이제곱 평균 : 자유도
분산 : 자유도 X 2

$\frac{u}{n} = \frac{1}{n}\underline{\sum_{i=1}^{n} z_i^2} \xrightarrow{p} E(z^2) = 1$ by tool1 (WLLN)

카이제곱 평균

$\sqrt{\frac{u}{n}} \xrightarrow{p} 1$ by tool2

$\therefore T_{(n)} = \frac{z}{\sqrt{\frac{u}{n}}} \xrightarrow{D} N(0,1)$ by Slutsky's thm.

then $P(C_1 < \frac{\bar{Y}_n - \mu}{S_n/\sqrt{n}} < C_2) = 95\%$

$\sim t_{(n)}$ 우리가 아는 분포

C_1: $-t_{0.975,n}$ C_2: $t_{0.975,n}$

$P(C_1 < \frac{\bar{Y}_n - \mu}{S_n/\sqrt{n}} < C_2) \doteq 95\%$

$\sim Z$

C_1: -1.96 C_2: 1.96 $n \to \infty$ 근사한다!

☆ $Y_1 \cdots Y_n \sim iid\ Ber(p)$, $(1-\alpha)$ CI for p?

p: target parameter

확률

$Y_i = \begin{bmatrix} 1 & p \\ 0 & 1-p \end{bmatrix}$ $E(Y) = p$

$Var(Y) = p(1-p)$

$\hat{p} = \bar{Y}$

by CLT $\boxed{\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}} \xrightarrow{D} N(0,1)$

Q. pivotal quantity for p?

A. Yes!

① p : only parameter

② 분포에 p 없음.

$E(\hat{p})$, $Var(\bar{Y}) = \frac{Var(Y)}{n}$

$P(C_1 < \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} < C_2) \doteq (1-\alpha)$

C_1: $-Z_{\alpha/2}$ C_2: $Z_{\alpha/2}$

$$P\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} < p < \hat{p} + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right) \fallingdotseq (1-\alpha)$$

$$\frac{\dfrac{\hat{p}-p}{\sqrt{\dfrac{p(1-p)}{n}}}}{\boxed{}} \xrightarrow{P} N(0,1) \quad \text{by slutsky}$$

$$\frac{\hat{p}-p}{\sqrt{\dfrac{p(1-p)}{n}}} \xrightarrow{D} z \quad \text{by CLT}$$

$$\boxed{} \xrightarrow{P} 1$$ 결과론적으로 풀기

목표 : $\dfrac{\hat{p}-p}{\sqrt{\dfrac{p(1-p)}{n}}} \longrightarrow \dfrac{\hat{p}-p}{\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}}$

$$\frac{\hat{p}-p}{\frac{1}{\sqrt{n}}\sqrt{p(1-p)}\cdot\dfrac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{p(1-p)}}} = \frac{\hat{p}-p}{\sqrt{\dfrac{p(1-p)}{n}}}\cdot\frac{\sqrt{p(1-p)}}{\sqrt{\hat{p}(1-\hat{p})}} = \frac{\dfrac{\hat{p}-p}{\sqrt{\dfrac{p(1-p)}{n}}}}{\sqrt{\dfrac{\hat{p}(1-\hat{p})}{p(1-p)}}}$$

$$\frac{\hat{p}-p}{\sqrt{\dfrac{p(1-p)}{n}}} \xrightarrow{D} z$$

$$\sqrt{\frac{\hat{p}(1-\hat{p})}{p(1-p)}} \xrightarrow{P} 1$$ 보이기!

$\hat{p} = \bar{Y} \xrightarrow{P} E(Y) = p$ by tool 1 (WLLN)

① $\hat{p} \xrightarrow{P} p$

② $1-\hat{p} \xrightarrow{P} 1-p$ by tool 2

①×② $\hat{p}(1-\hat{p}) \xrightarrow{P} p(1-p)$ by tool 2

곱하기 $\dfrac{\hat{p}(1-\hat{p})}{\boxed{p(1-p)}} \xrightarrow{P} 1$ by tool 2

$\sqrt{\dfrac{\hat{p}(1-\hat{p})}{p(1-p)}} \xrightarrow{P} \sqrt{1} = 1$ by tool 2

$$\therefore \frac{\dfrac{\hat p - p}{\sqrt{\dfrac{p(1-p)}{n}}}}{\dfrac{\hat p(1-\hat p)}{p(1-p)}} = \boxed{\frac{\hat p - p}{\sqrt{\dfrac{\hat p(1-\hat p)}{n}}}} \xrightarrow{D} N(0.1)$$

by slutsky

$$\frac{\hat p - p}{\sqrt{\frac{p(1-p)}{n}}} \xrightarrow{D} N(0.1) \qquad \frac{\hat p(1-\hat p)}{p(1-p)} \xrightarrow{P} 1.$$

↳ 이 값을 새로운 pivotal quantity로 씀.

$$P\left(c_1 < \frac{\hat p - p}{\sqrt{\frac{\hat p(1-\hat p)}{n}}} < c_2\right) = 0.95 \qquad \frac{\hat p - p}{\sqrt{\frac{\hat p(1-\hat p)}{n}}} \xrightarrow{D} N(0.1)$$

$$P\left(-z_{\alpha/2} < \frac{\hat p - p}{\sqrt{\frac{\hat p(1-\hat p)}{n}}} < z_{\alpha/2}\right) \doteqdot 0.95$$

$$P\left(\hat p - z_{\alpha/2}\sqrt{\frac{\hat p(1-\hat p)}{n}} < p < \hat p + z_{\alpha/2}\sqrt{\frac{\hat p(1-\hat p)}{n}}\right) \doteqdot 0.95$$

ex) $X_i \sim$ iid Gamma$(3,\theta)$, $E(X)=3\theta$, $Var(X)=3\theta^2$

1. Show that $\sqrt{3n}(\bar x - 3\theta)/\bar x$ converge in distribution $N(0,1)$

↓

① CLT
② mapping
③ slutsky

] 3개 떠올리기

<tip>

slutsky $\dfrac{\square}{\square}$ $\quad \square \xrightarrow{D} (\quad)$ by CLT $\quad \square \xrightarrow{P} 1$ by WLLN

이런 경우가 많다.

$$E(\bar x) = 3\theta, \quad Var(\bar x) = \frac{Var(X)}{n} = \frac{3\theta^2}{n}$$

$$\frac{\dfrac{\bar x - 3\theta}{\frac{\sqrt3\theta}{\sqrt n}}}{(*)} \qquad \frac{\bar x - 3\theta}{\frac{\sqrt3\theta}{\sqrt n}} \xrightarrow{D} N(0.1) \text{ by CLT}$$

비교하기

$$\frac{\dfrac{(\bar x - 3\theta)}{\frac{1}{\sqrt{3n}}}}{\bar x}$$

2. Find approximate 95% CI of θ using result in 1. for a large n

$\frac{3\theta}{\sqrt{3n}}$ 곱하면 ?

$$\frac{\left(\frac{\bar{X}-3\theta}{\frac{3\theta}{\sqrt{3n}}}\right)}{\left(\frac{\bar{X}}{3\theta}\right)}$$

$\frac{\bar{X}-3\theta}{\frac{3\theta}{\sqrt{3n}}} \xrightarrow{D} N(0,1)$

$\frac{\bar{X}}{3\theta} \xrightarrow{P} 1.$

$\bar{X} \xrightarrow{P} (EX)=3\theta$ by tool1

$\frac{\bar{X}}{3\theta} \xrightarrow{P} 1$ by tool 2.

$$\underbrace{\sqrt{3n}(\bar{X}-3\theta)/\bar{X}}_{\text{pivotal quantity}} \xrightarrow[\text{근사하므로}]{D} N(0,1)$$

pivotal quantity

① θ : only parameter

② $N(0,1)$

$$P\left[\underset{-1.96}{c_1} < \underbrace{\frac{\sqrt{3n}(\bar{X}-3\theta)}{\bar{X}}}_{Z} < \underset{1.96}{c_2}\right] \fallingdotseq 95\%$$

P [정리 $< \theta <$ 정리] $\fallingdotseq$ 95%