23 – 2 DSL 정규 세션 기초과제 1

기수: 10 기 이름: 윤희찬

문제 1

---

1 - 1) 중심극한정리의 정의와 그 의미를 서술하시오

중심극한정리의 정의 : 독립적이고 동일하게 분포된 확률 변수들의 합 또는 평균이 표본 크기가 커질수록 정규분포에 근사한다

중심극한정리의 의미 : 중심극한정리는 모집단 변수의 분포형태에 좌우되지 않아, 표본을 통해 가설검정과 추정을 하는 데 있어 매우 유용합니다. 예를 들어, 사회과학의 조사에 있어 모집단의 분포를 몰라도, 표본평균의 분포가 근사적으로 정규분포를 따른다는 가정을 통해 이론적인 표본평균 분포값에 따라 가설을 검정할 수 있습니다.

1 – 2) 중심극한정리가 통계적 추론 중 "구간추정"에서 어떻게 유용한지 서술하시오

표본 크기가 모집단 크기보다 훨씬 작더라도, 표본 크기가 충분히 크다면 CLT 를 이용하여 정규분포를 기반으로 구간을 추정할 수 있습니다. 또한 이를 통해 신뢰구간을 계산하여 모수의 값을 추정하고 추정의 신뢰성을 통계적으로 평가할 수 있습니다.

1 – 3) 중심극한정리를 이용하여 모평균에 대한 근사신뢰구간을 만들 때, 표준오차 부분의 수식적으로 증명하시오.

$p(|\bar{X}-\mu| > c) \to 0$ as $n \to \infty$ for any $c > 0$

$$P(|\bar{X}-\mu| > c) \le \frac{Var(\bar{X})}{c^2} = \frac{\frac{1}{n^2}\cdot n\sigma^2}{c^2} = \frac{\sigma^2}{nc^2}$$

$$Var(\bar{X}) = Var\left(\frac{1}{n}\sum X_i\right)$$
$$= \frac{1}{n^2} Var\left(\sum X_i\right)$$
$$= \frac{1}{n^2} \times n \times \sigma^2 \qquad \therefore S^2 \xrightarrow{p} \sigma^2$$

Slutsky's theorem

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} = \frac{\sigma}{S} \cdot \frac{(\bar{X}-\mu)}{\sigma/\sqrt{n}}$$

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} \xrightarrow{d} Z \quad (N(0,1^2))$$

따라서 CLT에의해 $n \to \infty$ 이므로 이 때

표준오차를 표본분산으로 대체할 수 있다.

문제 2

2 – 1) 스튜던트 정리 3 번 증명

③. $(n-1)\frac{S^2}{\sigma^2} \sim \chi^2(n-1)$

증명 $V = \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2$ $\quad \frac{X_i-\mu}{\sigma} \sim N(0,1)$

$V \sim \chi^2(n)$

$$V = \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2$$

$$= \sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 + \left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2$$

$$\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 = \frac{n-1}{\sigma^2}\sum_{i=1}^{n}\frac{(X_i-\bar{X})^2}{n-1}$$

$$V = (n-1)\frac{S^2}{\sigma^2} + \left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2$$

$$V \sim \chi^2(n) \quad \left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right) \sim N(0,1)$$

$$\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2 \sim \chi^2(1)$$

$$(1-2t)^{-n/2} = E\{e^{((n-1)\frac{S^2}{\sigma^2}t)}\}(1-2t)^{-1/2}$$

따라서 $(n-1)\frac{S^2}{\sigma^2}$의 적률생성함수가
$\chi^2(n-1)$의 적률생성함수이다 unique한 확률밀도 pdf가
결정된다. 그러므로 ③은 증명된다.

2 – 2) 스튜던트 정리 4 번 증명

④: $T = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$

$$T = \frac{\bar{X}-\mu}{S/\sqrt{n}} = \frac{(\bar{X}-\mu)/(\sigma/\sqrt{n})}{\sqrt{(n-1)S^2/(\sigma^2(n-1))}}$$

따라서 ①과 ③에 의해 $T \sim t(n-1)$

문제 3

3 – 1)

- a + b

μ_1 : DSL 학회원들의 평균키

μ_2 : DSL 학회원이 아닌 사람들의 평균키

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 > \mu_2$

b) 등분산 검정

검정통계량 $\dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} = -1.411 \sim Z$

가설검정 : $Z > Z_{0.05}$ reject H_0

$-1.411 < 1.645$ not reject H_0

통계적결론 : DSL 학회원들의 평균키가 DSL 학회원이 아닌 사람들의 평균키보다 크다고 할 수 없다.

3 – 2)

- a

μ_A : 해제 이전 평균 이용객수

μ_B : 해제 이후 평균 이용객수

$H_0: \mu_A - \mu_B = 0$

$H_1: \mu_A - \mu_B < 0$