문제 1	Central Limit Theorem

중심극한정리는 확률변수의 합 형태(sum of random variables)의 극한분포를 손쉽게 구할 수 있도록 해주기에 통계학에서 가장 자주 사용하는 정리입니다. 이 문제에서는 중심극한정리의 정의와 그 활용에 대해 짚어보겠습니다.

(1-1) ✎ 중심극한정리(Central Limit Theorem)의 정의와 그 의미를 서술하시오.

- ☝ 통계학입문(3판) 7장 참고
- ☝ Hogg(8판) 4장 2절, 5장 3절 참고

Def: 평균이 μ이고 분산이 σ^2인 모집단에서 랜덤추출한 서로 독립적인 확률변수 $X_1, X_2 \cdots X_n$이 존재할 때, $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$은 n이 ∞에 가까워짐에 따라 $N(0,1)$인 표준정규분포를 따른다.

$$\left(\begin{array}{l} \text{As } n \to \infty \\ \bar{X}_n \sim N(\mu, \frac{\sigma^2}{\sqrt{n}} \end{array} \right)$$

Meaning: 표본의 크기가 충분히 크다면, 모집단이 어떤 분포를 갖고있던지 $N(\mu, \sigma^2)$를 따른다.

→ 따라서 수집한 표본의 통계량을 통하여 모수 ($\mu, \sigma^2 \cdots$)을 확률적으로 추정할수 있다는 점에 의미가 있는 정리이다.

(1-2) ✎ 중심극한정리가 통계적 추론 중 “구간추정”에서 어떻게 유용한지 서술하시오.

☝ Hogg(8판) 4장 2절 참고

(1-3) ✎ 중심극한정리를 이용하여 모평균에 대한 근사신뢰구간을 만들 때, 표준오차($\sqrt{Var(\overline{X})}$) 부분의 모분산을 표본분산으로 대체할 수 있는 이유를 수식적으로 증명하시오.

☝ 표본분산 s^2는 모분산 σ^2로 확률수렴한다는 사실을 이용할 수 있습니다.
☝ Slutsky's theorem을 이용할 수 있습니다.
☝ Hogg(8판) 5장 1~3절 참고

1-2) 모집단이 정규분포를 따르지 않아도 신뢰구간 추정이 가능하다.

CLT에 의해 표본의 크기가 충분히 크다면 표본평균의 분포는 정규분포를 따르므로, $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$을 이용해

근사신뢰구간을 구할수있다.

예) 신뢰도 95%로 μ(모평균)을 추정하면 $\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$로 구간추정이 가능하다.

(모분산을 모르는 경우 S)

1-3) $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 = \frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{X}_n^2\right)$

$$= \frac{n}{n-1}\left(E(X_i^2) - \mu^2\right)$$

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} \frac{n}{n-1}\left(E(X_i^2) - \mu^2\right) = E(X_i^2) - \mu^2 = \sigma^2$$

As $n \to \infty$,

∴ 표본분산 S^2는 모분산 σ^2에 수렴한다.

따라서 CLT가 적용되면 표본분산 S^2이 σ^2에 확률수렴 하므로 대체하여 사용할 수 있다.

문제 2 Student's Theorem

스튜던트 정리는 통계적 추정에서 필요한 정리 중 하나로, 표본평균과 표본분산이 어떤 분포를 갖는지 알려줍니다. 이 문제에서는 스튜던트 정리의 내용을 어떻게 수식적으로 유도할 수 있는지 짚어보겠습니다.

스튜던트 정리는 다음과 같이 총 4개의 내용으로 구성되어 있습니다.

① $\overline{X} \approx N(\mu, \frac{\sigma^2}{n})$ → 〈문제 1〉에서 증명함

② 표본평균 $\overline{X}$ 와 표본분산 s^2은 서로 독립이다.

③ ???

④ ???

(2-1) ✎ ③의 내용을 쓰고 증명하시오.

☝ Hogg(8판) 3장 6절 참고

☝ 무작위표본 $X_1, \cdots, X_n$이 독립적으로 동일하게(independently and identically distributed) 평균이 μ이고 분산이 σ^2인 정규분포를 따를 때, 자유도가 n인 카이제곱분포를 따르는 새로운 확률변수 V를 아래와 같이 두어 증명에 활용할 수 있습니다.

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

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

Proof: $Z^2 = \left(\frac{X-\mu}{\sigma}\right)^2$. So, $V = \sum_{i=1}^{n} \left(\frac{X_i-\mu}{\sigma}\right)^2 = \sum_{i=1}^{n} Z_i^2$

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

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

by ①, $\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right) \sim N(0,1)$ (표준화)이므로 $\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right)^2 \sim \chi^2(1)$

by ②, $\frac{S^2(n-1)}{\sigma^2}$과 $\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right)^2$ 독립이며 $\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right)^2$가 자유도가 1인 카이제곱 ~~분포를~~

따라서 V는 자유도가 n인 카이제곱 분포를 따르므로 $\frac{(n-1)S^2}{\sigma^2}$은 자유도가 (n-1)인 카이제곱분포를 따른다.

(2-2) ✎ ④의 내용을 쓰고, (2-1)을 이용하여 증명하시오.

> ☝ Hogg(8판) 3장 6절 참고
> ☝ t분포의 정의에 따르면, 표준정규분포를 따르는 확률변수와 카이제곱분포를 따르는 확률변수를 이용하여 t분포를 유도할 수 있습니다.

---

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

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

by ① $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$, by 2-1 $\dfrac{S^2(n-1)}{\sigma^2} \sim \chi^2(n-1)$

T분포의 정의: W, V 독립, $W \sim N(0,1)$, $V \sim \chi^2(r)$

$T = \dfrac{W}{\sqrt{V/r}} \sim t(r)$

따라서 $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} = W$, $\dfrac{(n-1)S^2}{\sigma^2} = V$라 하면, r은 $n-1$이 되며

$T = \dfrac{\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2} \cdot \dfrac{1}{n-1}}} = \dfrac{W}{\sqrt{V/r}} \sim t(r) \ (= t(n-1))$