

LGE - SNU DS 고급 과정

확률통계 및 공통 알고리즘

과제3 - 표본분포

2022. 01

※필요하다면 다음의 값을 이용하시오

: $z_{0.977} = 2$, $t_{0.90}(24) = 1.318$, $\chi^2_{0.90}(24) = 33.20$

문제 1. 다음과 같은 분포를 가지는 무한모집단에서의 크기 2의 랜덤포본을 X_1, X_2 라 하자.

x	0	1	2
$P(X = x)$	0.2	0.4	0.4

(a) X 의 기댓값 $E(X)$ 와 분산 $Var(X)$ 를 구하여라.

(b) 표본평균 $\bar{X} = \frac{X_1 + X_2}{2}$ 의 확률분포를 구하여라.

(c) $E(\bar{X})$ 와 $Var(\bar{X})$ 를 계산하고 $E(X)$, $Var(X)$ 와의 관계에 대해 설명하여라.

[풀이].

(a) $E(X) = 0 \times 0.2 + 1 \times 0.4 + 2 \times 0.4 = 1.2$

$$E(X^2) = 0^2 \times 0.2 + 1^2 \times 0.4 + 2^2 \times 0.4 = 2.0$$

$$Var(X) = E(X^2) - \{E(X)\}^2 = 2.0 - 1.2^2 = 0.56$$

(b)

x	0	0.5	1	1.5	2
$P(\bar{X} = x)$	0.04	0.16	0.32	0.32	0.16

(c) $E(\bar{X}) = 1.2 = E(X)$

$$E(\{\bar{X}\}^2) = 1.72$$

$$Var(\bar{X}) = E(\{\bar{X}\}^2) - \{E(\bar{X})\}^2 = 1.72 - 1.2^2 = 0.28 = \frac{Var(X)}{2}$$

문제 2. 유한모집단 $\{1, 3, 3, 4\}$ 에서 크기 2인 표본을 단순랜덤비복원추출하였을 때, 표본평균의 확률분포를 구하여라.

[풀이].

x	2	2.5	3	3.5
$P(\bar{X} = x)$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$

문제 3. 하루 10,000개의 전구를 생산하는 공장이 있다. 이 공장에서 생산되는 전구의 불량률이 0.01이라 한다. 어느 날 생산한 전구 중 불량품이 120개 이상이 될 확률을 중심극한정리를 이용하여 근사적으로 구하여라. 이 때, $\sqrt{99}$ 의 값을 10 으로 근사하여도 좋다.

[풀이]. 생산한 전구 중 불량품이 개수를 X 라 하면 $X = X_1 + \cdots + X_{10000}$, $X_i \stackrel{i.i.d}{\sim} Ber(0.01)$ 로 나타낼 수 있다. $Ber(0.01)$ 의 평균과 분산은 각각 0.01, 0.01×0.99 이고, n 이 충분히 크므로 중심극한정리에 의해서 \bar{X} 의 분포는 정규분포 $N(0.01, \frac{0.01 \times 0.99}{10000})$ 에 근사한다. $X = 10000\bar{X}$ 이므로, 따라서 X 의 분포는 $N(100, 99)$ 에 근사한다.

$$\begin{aligned}
 P(X \geq 120) &= P\left(\frac{X - 100}{\sqrt{99}} \geq \frac{120 - 100}{\sqrt{99}}\right) \\
 &= P(Z \geq 2) \quad \text{where } Z \sim N(0, 1^2) \\
 &= 0.023
 \end{aligned}$$

문제 4. 평균이 2인 정규분포를 따르는 것으로 알려진 모집단에서 크기가 25인 랜덤표본들을 추출할 수 있다. 이렇게 뽑은 랜덤표본의 표본분산을 S^2 라 정의하자.

- (a) $P(\bar{X} \leq 2 + y \cdot S) = 0.90$ 가 성립하는 y 를 구하여라.
- (b) 모분산이 16이라 할 때, 표본분산 S^2 에 대하여 $P(S^2 \leq y) = 0.90$ 가 성립하는 y 를 구하여라.

[풀이].

$$(a) \quad T = \frac{\bar{X} - 2}{S/\sqrt{25}} = 5 \left(\frac{\bar{X} - 2}{S} \right) \sim t(24) \text{ 이므로}$$

$$\begin{aligned}
 P(\bar{X} \leq 2 + y \cdot S) &= 0.90 \Rightarrow P(T \leq 5y) = 0.90 \\
 &\Rightarrow P(T > 5y) = 0.10 \\
 &\Rightarrow 5y = t_{0.90}(24)
 \end{aligned}$$

$$y = \frac{1}{5} t_{0.90}(24) = \frac{1}{5} \times 1.318 = 0.2636$$

$$(b) \quad V = \frac{(25-1)S^2}{4^2} = \frac{3}{2}S^2 \sim \chi^2(24) \quad \text{이므로}$$

$$P\left(S^2 \leq y\right) = 0.90 \Rightarrow P\left(V \leq \frac{3}{2}y\right) = 0.90$$

$$\Rightarrow P\left(V > \frac{3}{2}y\right) = 0.10$$

$$\Rightarrow \frac{3}{2}y = \chi^2_{0.90}(24)$$

$$y = \frac{2}{3}\chi^2_{0.90}(24) = \frac{2}{3} \times 33.20 = 22.133$$