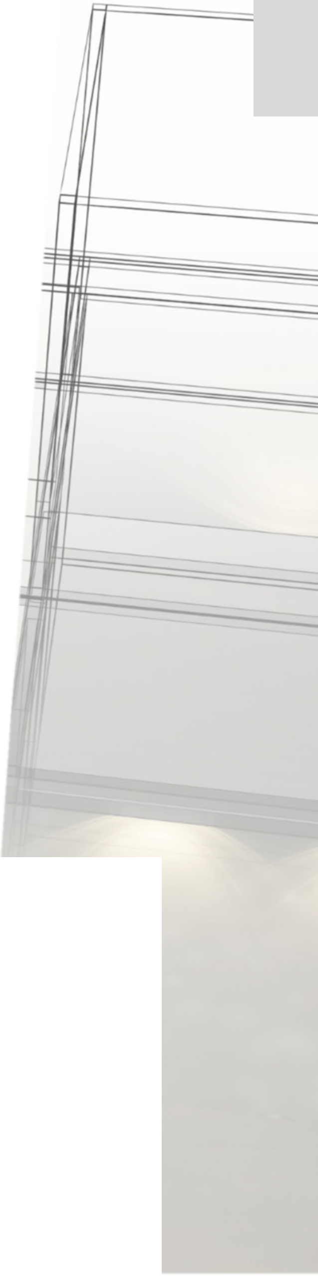
12강

표본분포 2

통계·데이터과학과
이긍희 교수

학습목표

1. 대수의 법칙을 이해할 수 있다.
2. 중심극한정리를 이해할 수 있다.
3. 표본분산의 분포를 이해할 수 있다.
4. 표본평균의 분포를 이해할 수 있다.



01

12강 표본분포 2

대수의 법칙

확률변수의 수렴

- 확률적 수렴 : $X_n \xrightarrow{p} X$

  임의의 $\varepsilon > 0,\ \lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0$

- 거의 확실 한 수렴 : $X_n \xrightarrow{a.s.} X$

$$P(\lim_{n\to\infty} X_n = X) = 1$$

확률변수의 수렴

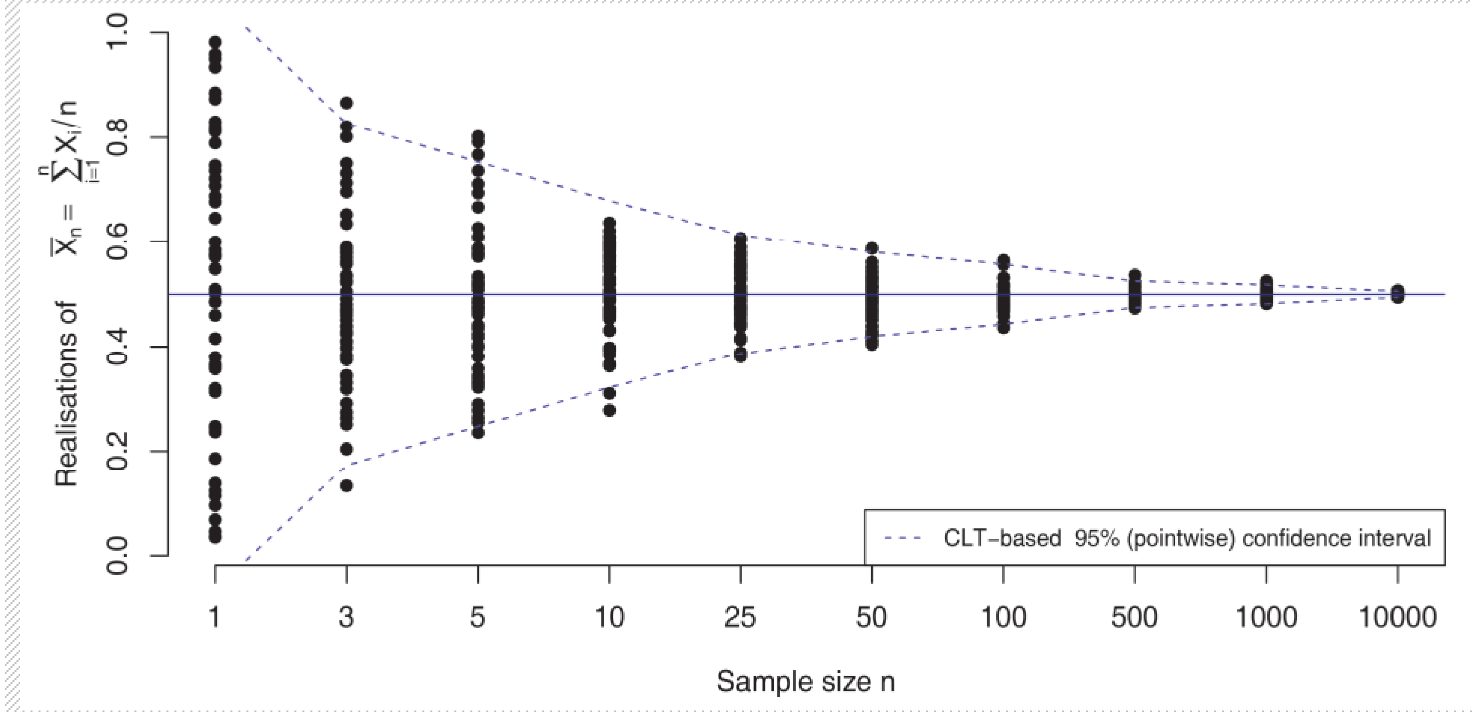
- 분포적 수렴 : $X_n \xrightarrow{d} X$

$$\lim_{n\to\infty} P(X_n < x) = P(X < x)$$

확률변수의 수렴

- $X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{p} X \Rightarrow X_n \xrightarrow{d} X$

대수의 법칙



Realisations of $\overline{X}_n = \sum_{i=1}^{n} X_i / n$
0.0
0.2
0.4
0.6
0.8
1.0
CLT–based 95% (pointwise) confidence interval
1
3
5
10
25
50
100
500
1000
10000
Sample size n

대수의 법칙

- 약대수의 법칙(weak law of large number) : 확률적 수렴
- 강대수의 법칙(strong law of large number) : 거의 확실한 수렴

약대수의 법칙

- $X_1, X_2, \cdots, X_n \sim (\mu, \sigma^2)$ 의 확률표본.
  표본 수 커지면서 $\bar{X}_n$는 μ 에 확률적으로 수렴

  임의의 $\varepsilon > 0$, $\lim_{n\to\infty} P(|\bar{X}_n - \mu| > \varepsilon) = 0$

약대수의 법칙

◆ 임의의 $\varepsilon > 0, \lim_{n \to \infty} P(|\bar{X}_n - \mu| > \varepsilon) = 0$

강대수의 법칙

- $X_1, X_2, \cdots, X_n \sim (\mu, \sigma^2)$ 의 확률표본.

  표본 수 커지면서 $\bar{X}$는 μ 에 거의 확실하게 수렴

$$P(\lim_{n\to\infty} \bar{X}_n = \mu) = 1$$

02

12강 표본분포 2

중심극한정리

표본평균의 분포

- $X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2)$ 확률표본

$$\Rightarrow \sqrt{n}(\bar{X}_n - \mu) \sim N(0, \sigma^2)$$

중심극한정리

- 평균 μ, 분산 σ^2 인 무한 모집단의 n 개 확률표본.
  n 이 무한히 커지면서 표본평균은 모집단의
  분포 관계없이 근사적으로 정규분포를 따름

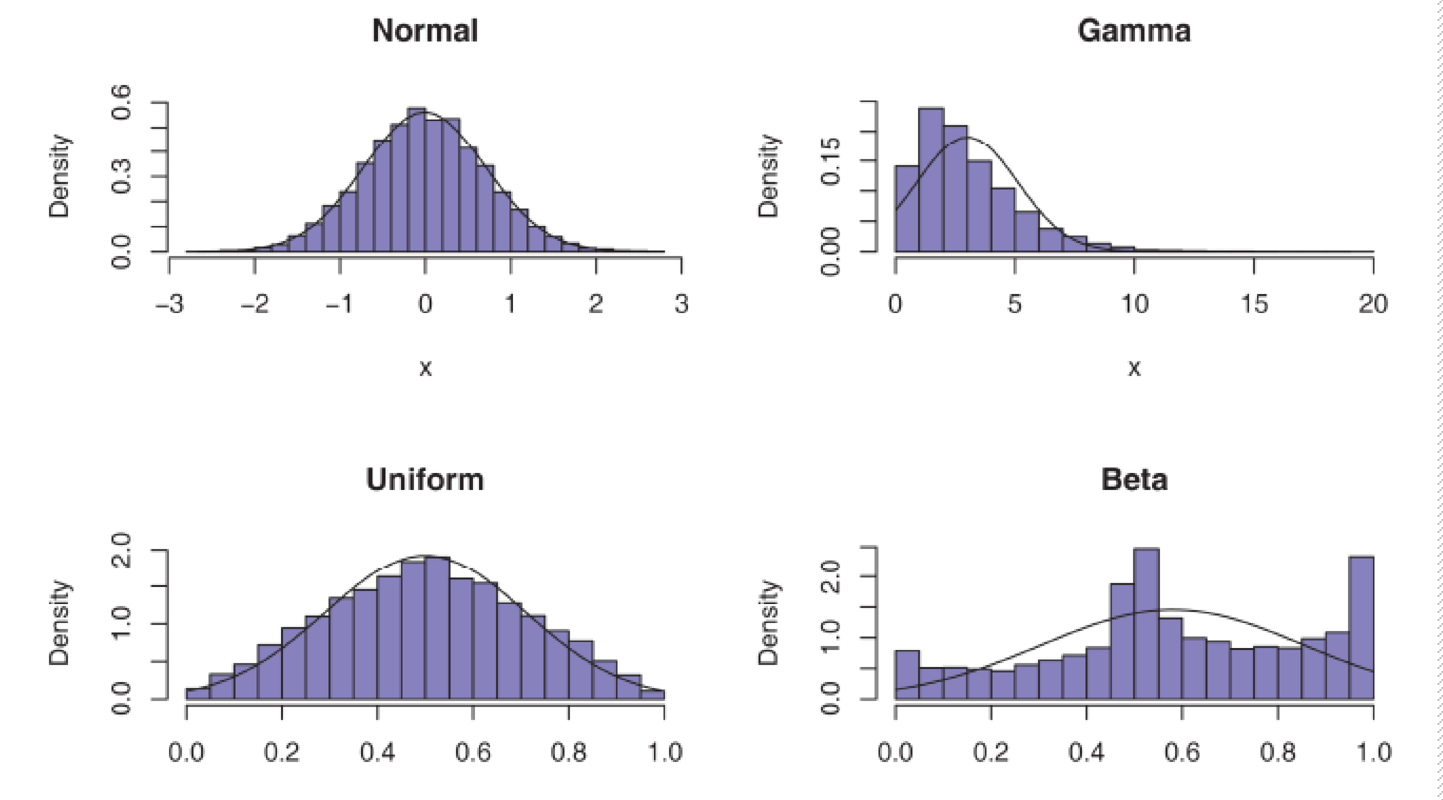
중심극한 정리

- $X_1, X_2, \cdots, X_n \sim (\mu, \sigma^2)$ 확률표본

$$\Rightarrow \sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} N(0, \sigma^2)$$

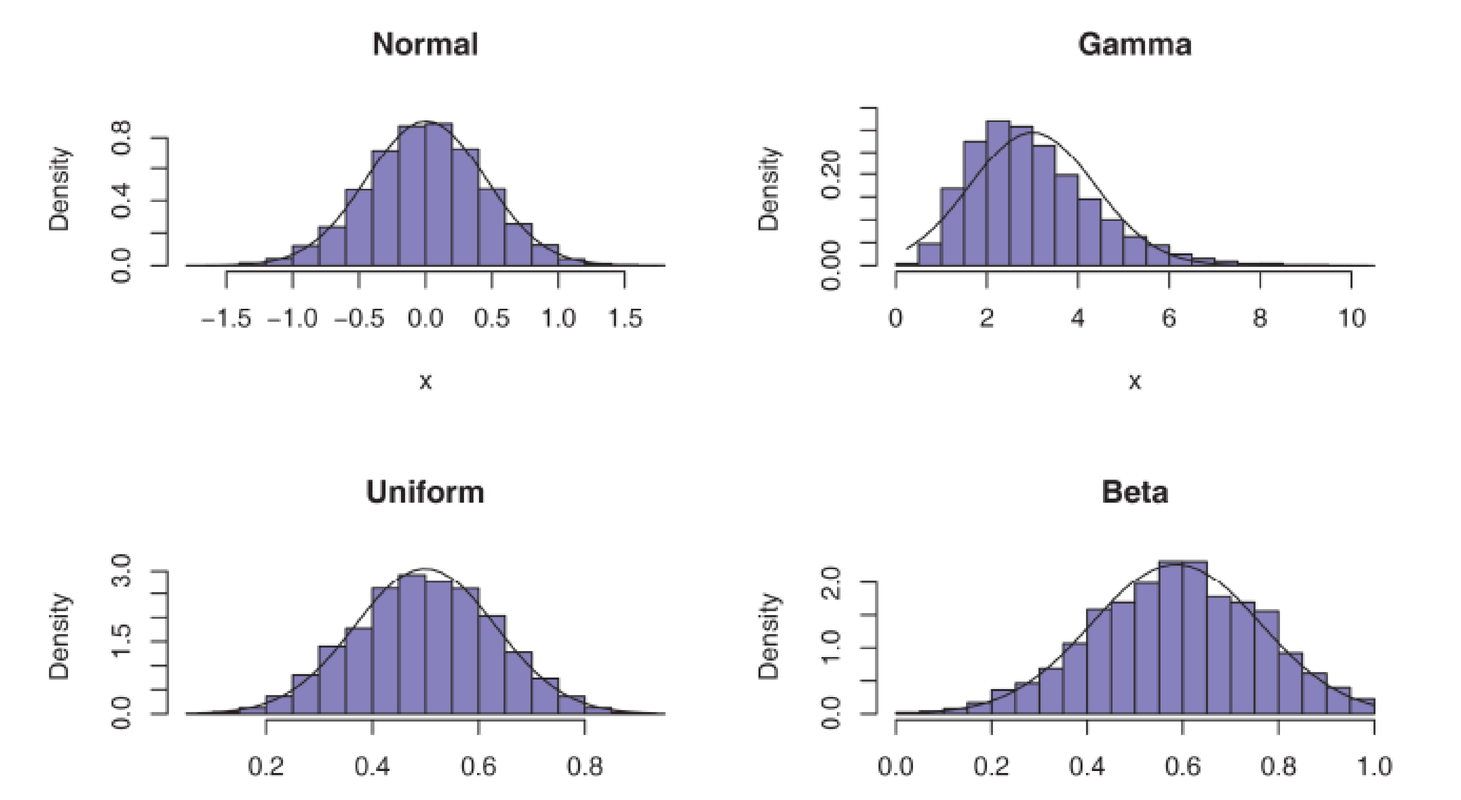
중심극한정리

- $n = 2$


Normal
Density
0.0
0.3
0.6
−3
−2
−1
0
1
2
3
x
Gamma
Density
0.00
0.15
0
5
10
15
20
x
Uniform
Density
0.0
1.0
2.0
0.0
0.2
0.4
0.6
0.8
1.0
Beta
Density
0.0
1.0
2.0
0.0
0.2
0.4
0.6
0.8
1.0

중심극한정리

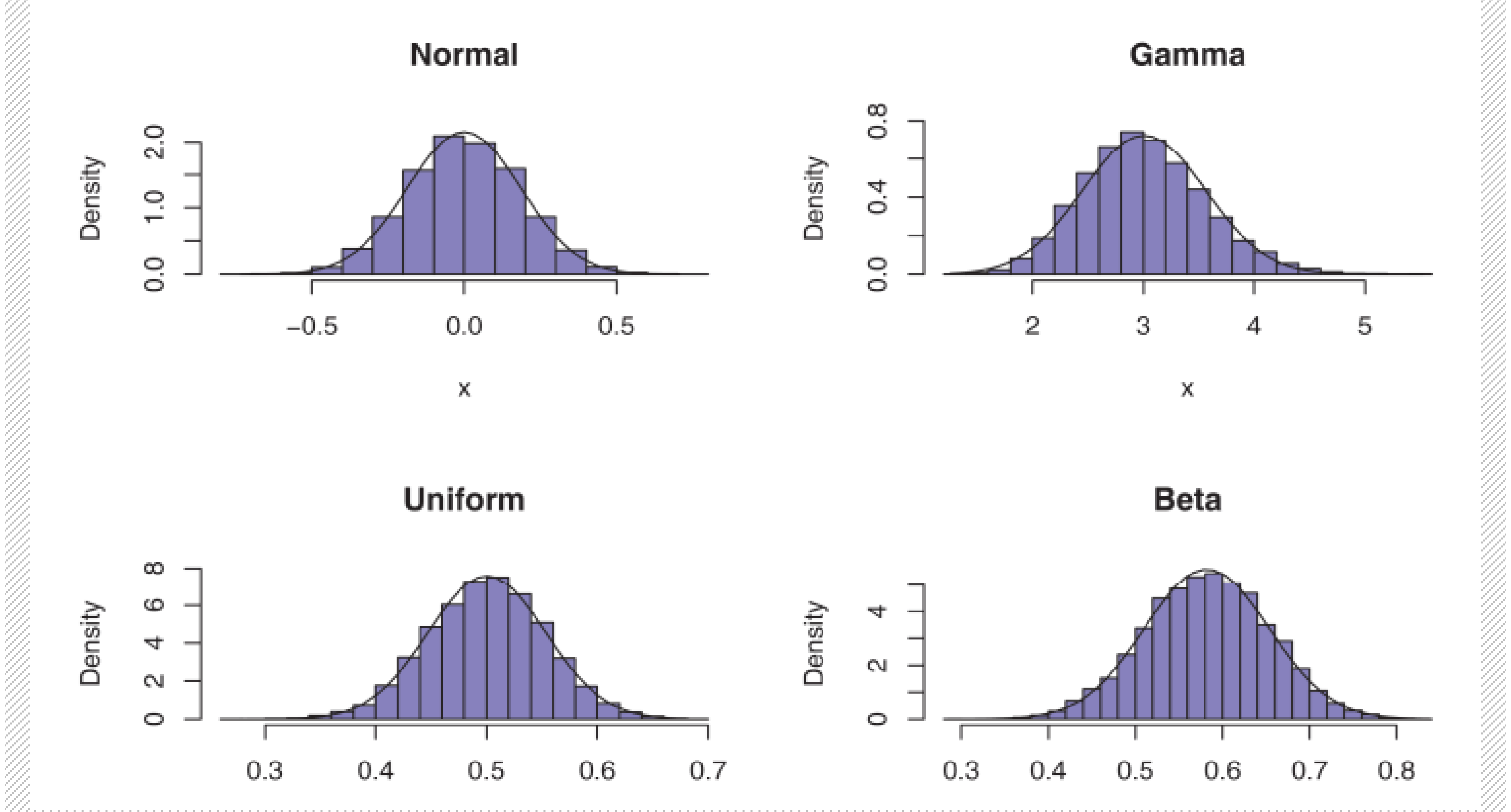
- $n = 5$



Normal
Density
0.0 0.4 0.8
−1.5 −1.0 −0.5 0.0 0.5 1.0 1.5
x
Gamma
Density
0.00 0.20
0 2 4 6 8 10
x
Uniform
Density
0.0 1.5 3.0
0.2 0.4 0.6 0.8
Beta
Density
0.0 1.0 2.0
0.0 0.2 0.4 0.6 0.8 1.0

중심극한정리

- $n = 30$


Normal
Density
0.0
1.0
2.0
−0.5
0.0
0.5
x
Gamma
Density
0.0
0.4
0.8
2
3
4
5
x
Uniform
Density
0
2
4
6
8
0.3
0.4
0.5
0.6
0.7
Beta
Density
0
2
4
0.3
0.4
0.5
0.6
0.7
0.8

중심극한정리의 예

예

성인 남성 콜레스톨 수치는 평균 200, 표준 편차 16. 성인 남성 64명을 임의로 뽑아서 구한 콜레스톨 수치의 평균이 196에서 204 사이일 확률은?

03

12강 표본분포 2

표본 분산의 분포

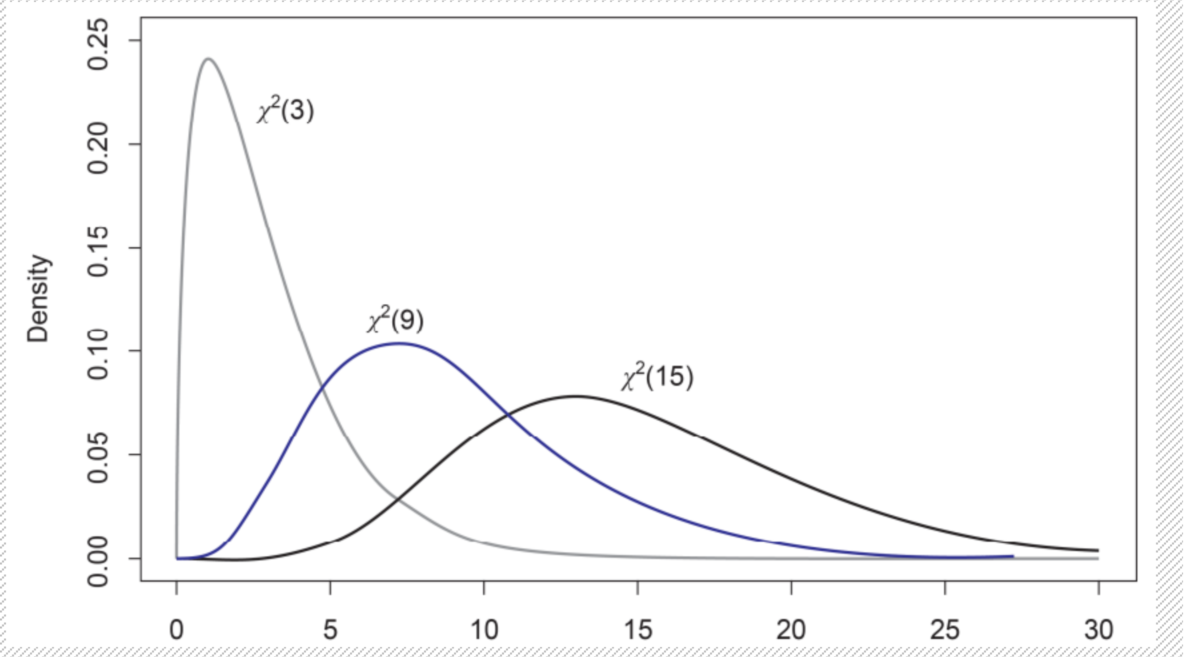
카이제곱 분포

- $X \sim \chi^2(n)$

$$f(x) = \frac{1}{\Gamma\left(\frac{n}{2}\right)} \left(\frac{1}{2}\right)^{\frac{n}{2}} x^{\frac{n}{2}-1} e^{-\frac{1}{2}x}, \quad x > 0$$

카이제곱 분포

- $X \sim \chi^2(n),\ E(X) = n, Var(X) = 2n$



카이제곱 분포의 가법성

- $X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2)$ 확률표본

  $\Rightarrow (\frac{X_i - \mu}{\sigma})^2 \sim \chi^2(1),\quad \sum_{i=1}^{n} (\frac{X_i - \mu}{\sigma})^2 \sim \chi^2(n)$

표본분산의 분포

- $X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2)$ 확률표본
- 표본분산 $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ 의 분포 :

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

표본분산의 분포

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

04

12강 표본분포 2

표본평균의 분포

확률변수 합의 분포

- $X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2)$ 확률표본

  $\Rightarrow X_1 + X_2 + \cdots + X_n \sim N(n\mu, n\sigma^2)$

표본평균의 분포

- $X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2)$ 확률표본

$$\Rightarrow \bar{X} \sim N(\mu, \sigma^2/n)$$

$$\Rightarrow Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

표본평균의 분포

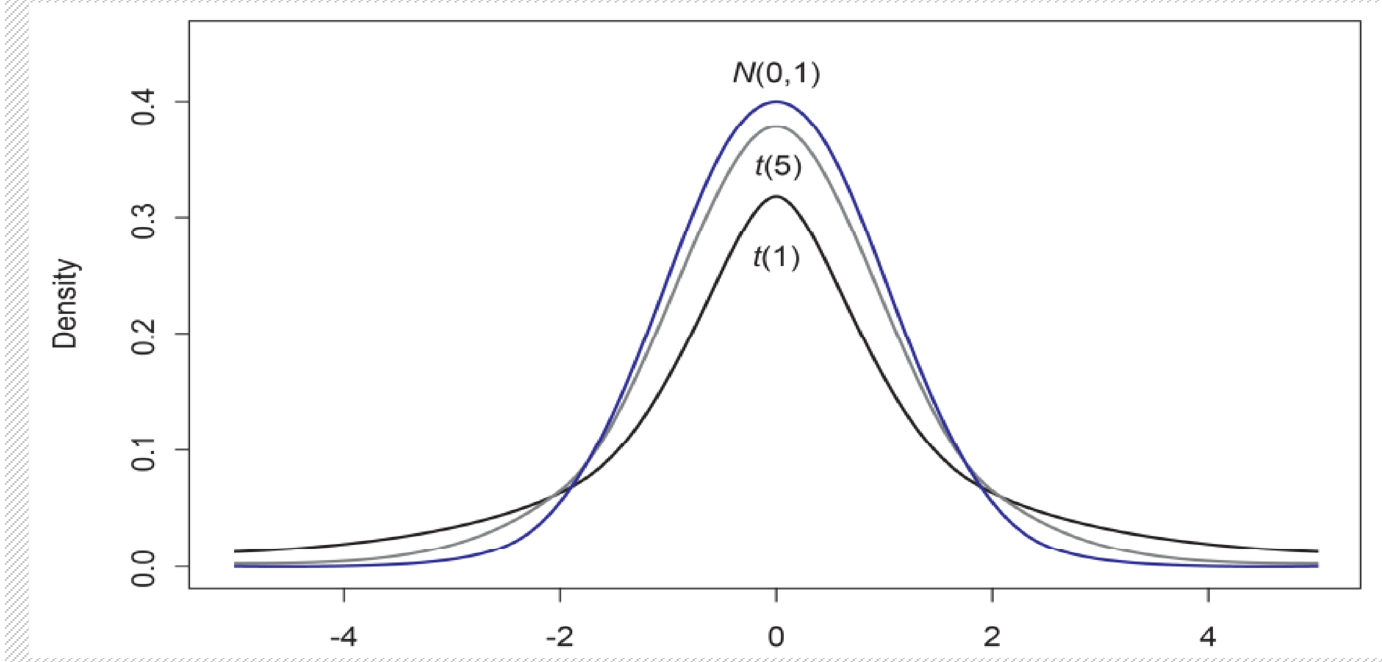
◆ $X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2)$ 확률표본

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

$$f(x) = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n-1}{2}\right)}\left(1 + \frac{x^2}{n-1}\right)^{-\frac{n}{2}} \frac{1}{\sqrt{n-1}}, \quad -\infty < x < \infty$$

t분포

◆ t분포의 모양



학습정리

- 모평균이 존재하는 분포에서 추출한 확률표본의 표본평균은 표본크기가 커지면서 모평균에 수렴한다.

$$\bar{X}_n \xrightarrow{p} \mu$$

- 표본 크기가 충분히 크면 모집단의 분포와 관계없이 표본평균은 근사적으로 정규분포를 따른다.

$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} N(0, \sigma^2)$$

정리하기

학습정리

- 모집단이 정규분포를 따를 때 표본분산은 카이제곱분포를 따른다.

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

- 모집단이 정규분포를 따르는데 모분산을 모른다면 표준화된 표본평균의 분포는 t분포를 따른다.

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

정리하기

다음시간에는

수고하셨습니다.

12강 표본분포 2

13강 확률과정 1