14 강

통계적추론

구간추정

한국방송통신대학교 통계·데이터과학과
이긍희 교수

한국방송통신대학교 대학원

학습내용

1. 구간추정의 기본 개념을 이해한다 .
2. 신뢰구간의 추정 방법을 이해한다.
3. 모평균에 대한 신뢰구간을 추정한다.
4. 모분산에 대한 신뢰구간을 추정한다.

01

구간추정의 개요

1 구간추정의 개요

구간추정의 정의

- 구간추정 : 모수를 포함할 것으로 기대되는 구간으로 모수를 추정

  【 예 】 여론조사, 경제예측

- $\boldsymbol{X} = (X_1, \cdots, X_n) \sim N(\theta, \sigma^2)$ 확률표본
  - θ의 점추정 : $T(\boldsymbol{X})$
  - θ의 구간추정 : $(L(\boldsymbol{X}),\ U(\boldsymbol{X}))$

1 구간추정의 개요

모평균의 구간추정

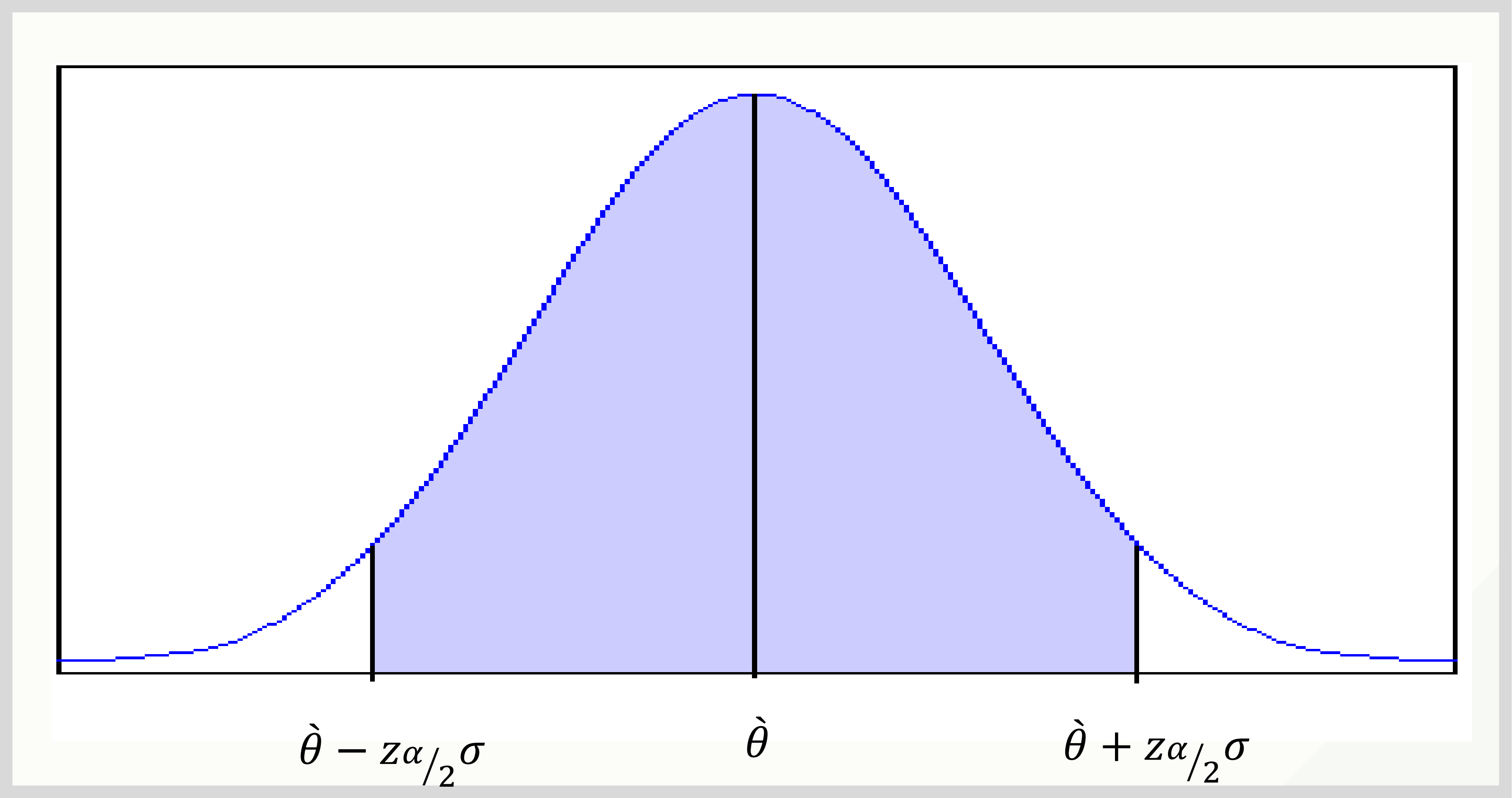
- $\hat{\theta} \sim N(\theta, \sigma^2)$, σ 기지
- $P(\theta - z_{\alpha/2}\sigma \le \hat{\theta} \le \theta + z_{\alpha/2}\sigma)$

$$= P\left(-z_{\alpha/2} \le \frac{\hat{\theta}-\theta}{\sigma} \le z_{\alpha/2}\right)$$

$$= P(\hat{\theta} - z_{\alpha/2}\sigma \le \theta \le \hat{\theta} + z_{\alpha/2}\sigma)$$

1 구간추정의 개요

- 모평균의 구간추정



1 구간추정의 개요

구간추정의 구분

- 신뢰구간(Confidence Interval)
- 신용구간(Credible Interval)

2 신뢰구간

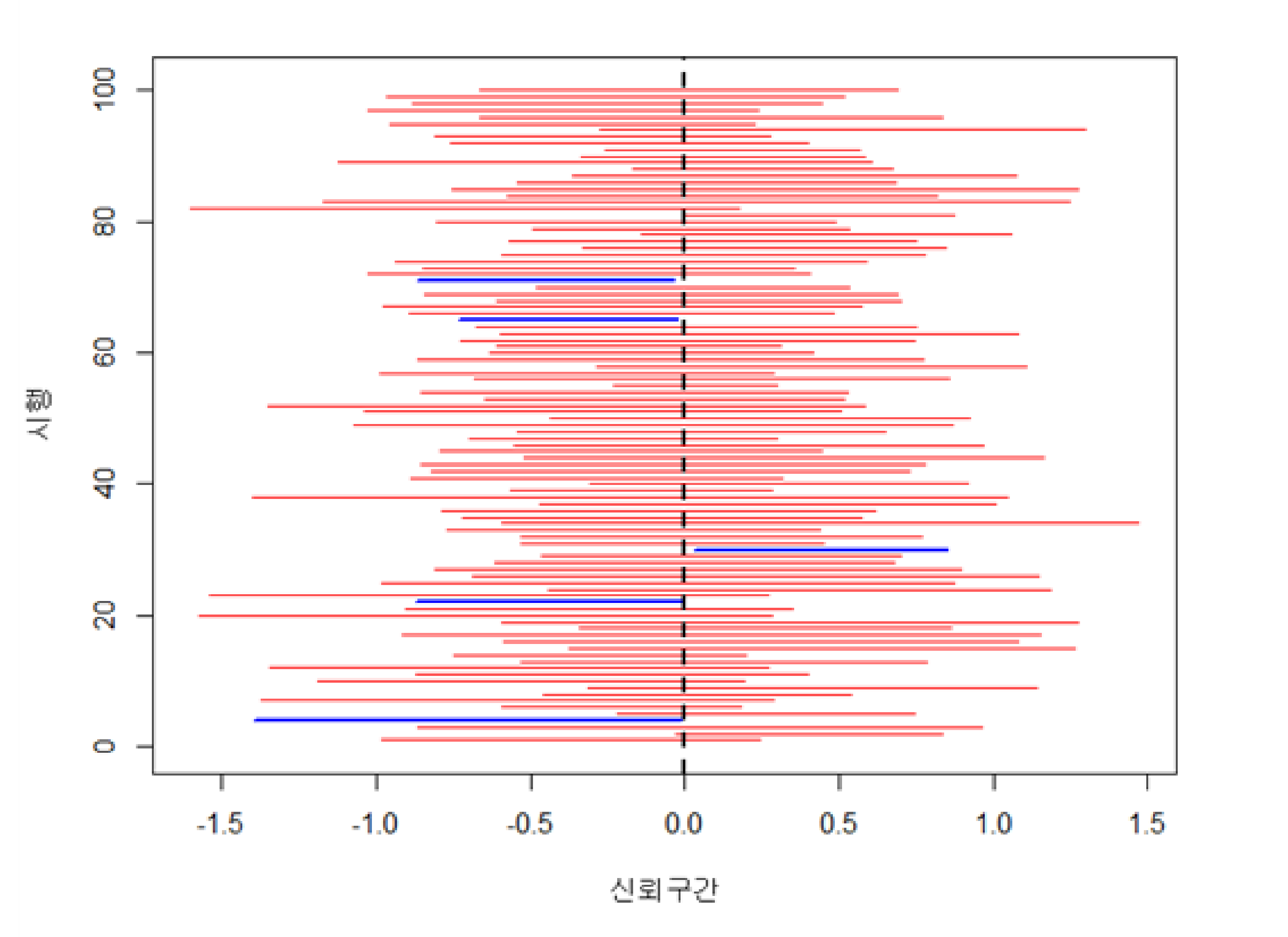
신뢰구간의 의미

- $100(1-\alpha)\%$ 신뢰구간
  - $100(1-\alpha)\%$: 신뢰구간을 구하는 과정을 여러 번 반복할 때 그 중에서 모수를 포함하는 신뢰구간의 비율의 극한
- θ의 $100(1-\alpha)\%$ 신뢰구간 : $[c(\hat{\theta}), d(\hat{\theta})]$

$$P[c(\hat{\theta}) \le \theta \le d(\hat{\theta})] = 1-\alpha$$

2 신뢰구간

95% 신뢰구간의 의미



02

신뢰구간의 추정 방법

1 신뢰구간의 추정 방법의 구분

신뢰구간의 추정 방법

- 가설검정의 채택역을 이용하는 방법
- 피봇(Pivot)을 이용하는 방법
- 누적분포함수를 이용하는 방법

2 검정 이용법

모평균의 가설검정

- $X_1, \cdots, X_n \sim N(\mu, \sigma^2)$ 확률표본, σ^2기지
- $H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu \neq \mu_0$
- 기각역 : $\mathrm{R}(\mu_0) = \{\bar{X} : | \bar{X} - \mu_0 | > z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\}$
- 채택역 : $\mathrm{A}(\mu_0) = \{\bar{X} : |\bar{X} - \mu_0| \leq z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\}$

2 검정 이용법

모평균의 가설검정과 신뢰구간 추정

- 유의수준 α에서 $H_0 : \mu = \mu_0$를 기각 못하는 μ_0의 값의 범위
  - σ기지 : $\left[\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$
  - σ 미지 : $\left[\bar{X} - t_{\alpha/2}(n-1)\frac{\sigma}{\sqrt{n}},\ \bar{X} + t_{\alpha/2}(n-1)\frac{\sigma}{\sqrt{n}}\right]$

→ 모평균에 대한 $100(1-\alpha)\%$신뢰구간

2 검정 이용법

θ의 신뢰구간

- $A(\theta_0)$: 유의수준 α에서 $H_0 : \theta = \theta_0$에 대한 채택역
- $C(\boldsymbol{X})$: θ의 $100(1-\alpha)\%$ 신뢰구간

$$C(\boldsymbol{X}) = \{\theta : \boldsymbol{x} \in A(\theta)\}$$

3 피봇 이용법

피봇(Pivotal Quantities)

- $Q(\boldsymbol{X},\theta)$: Q의 분포가 모든 모수와 독립
- $\boldsymbol{X} \sim F(\boldsymbol{x}|\theta) \rightarrow Q(\boldsymbol{X},\theta)$의 θ의 모든 값에 대해 같은 분포
  - $P_\theta(Q(\boldsymbol{X},\theta) \in A)$는 θ에 의존하지 않음

3 피봇 이용법

예 $X_1, \cdots, X_n \sim N(\mu, \sigma^2)$ 확률표본, σ^2 미지. 피봇을 이용하여 μ의 $100(1-\alpha)\%$ 신뢰구간을 구하시오.

3 피봇 이용법

예 $X_1, \cdots, X_n \sim N(\mu, \sigma^2)$ 확률표본, σ^2 미지. 피봇을 이용하여 μ의 $100(1-\alpha)\%$ 신뢰구간을 구하시오.

03

모평균에 대한 신뢰구간

1 모평균의 신뢰구간

표본평균의 분포 : σ^2 기지

- $X_1, \cdots, X_n \sim N(\mu, \sigma^2)$ 확률표본, σ^2기지
- $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$
- $P\left[-z_{\alpha/2} \le \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le z_{\alpha/2}\right] = 1 - \alpha$

1 모평균의 신뢰구간

모평균의 신뢰구간 : σ^2 기지

- $X_1, \cdots, X_n \sim N(\mu, \sigma^2)$ 확률표본, σ^2기지
- μ에 대한 $100(1-\alpha)\%$ 신뢰구간 :

$$\left[\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

1 모평균의 신뢰구간

표본평균의 분포 : σ^2미지

- $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ 확률표본, σ^2미지
- $T = \dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$
- $P\left[-t_{\alpha/2}(n-1) \le \dfrac{\bar{X}-\mu}{S/\sqrt{n}} \le t_{\alpha/2}(n-1)\right] = 1-\alpha$

1 모평균의 신뢰구간

모평균의 신뢰구간 : σ^2미지

- $X_1, \cdots, X_n \sim N(\mu, \sigma^2)$ 확률표본, σ^2미지
- μ에 대한 $100(1-\alpha)\%$ 신뢰구간 :

$$\left[\bar{X} - t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}}, \bar{X} + t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right]$$

1 모평균의 신뢰구간

일반 모집단 모평균의 신뢰구간 : n 이 클 때

- $X_1, \cdots, X_n \sim (\mu, \sigma^2)$ 확률표본
- μ에 대한 $100(1-\alpha)\%$ 신뢰구간 : σ^2미지

$$P\{\mu \in [\bar{X} - z_{\alpha/2} S/\sqrt{n}, \bar{X} + z_{\alpha/2} S/\sqrt{n}]\} = 1 - \alpha$$

2 모비율의 신뢰구간

모비율의 신뢰구간

- $X_1, \cdots, X_n \sim Ber(p)$ 확률표본, $\hat{p} = \sum_{i=1}^{n} X_i / n$

$$P\left\{p \in \left[\hat{p} - z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}, \hat{p} + z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}\right]\right\} = 1 - \alpha$$

2 모비율의 신뢰구간

모비율의 신뢰구간

- $X_1, \cdots, X_n \sim Ber(p)$ 확률표본, $\hat{p} = \sum_{i=1}^{n} X_i / n$

$$P\left\{p \in \left[\hat{p} - z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}, \hat{p} + z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}\right]\right\} = 1 - \alpha$$

3 모평균 차의 신뢰구간

모평균 차의 신뢰구간 : 독립표본

- $X_1, \cdots, X_m \sim N(\mu_1, \sigma^2)$, $Y_1, \cdots, Y_n \sim N(\mu_2, \sigma^2)$ 확률표본, 독립

$$P\left\{\mu_1 - \mu_2 \in \left(\bar{X} - \bar{Y} \pm t_{\alpha/2}(m+n-2) S_p \sqrt{\frac{1}{m} + \frac{1}{n}}\right)\right\} = 1 - \alpha$$

단, $S_p^2 = \left\{\sum_{i=1}^{n}(X_i - \bar{X})^2 + \sum_{j=1}^{m}\left(Y_j - \bar{Y}\right)^2\right\}/(m+n-2)$

3 모평균 차의 신뢰구간

모평균 차의 신뢰구간 : 쌍체표본

- $(X_1, Y_1), \cdots, (X_n, Y_n)$ 확률표본, $D_i = X_i - Y_i \sim N(\mu_1 - \mu_2, \sigma_D^2)$ 독립

$$P\left\{\mu_1 - \mu_2 \in \left(\bar{D} \pm t_{\alpha/2,(n-1)}\, S_D/\sqrt{n}\right)\right\} = 1 - \alpha$$

단, $S_D^2 = \sum_{i=1}^{n} (D_i - \bar{D})^2/(n-1)$

4 모비율 차의 신뢰구간

모비율 차의 신뢰구간

- $X_1, \cdots, X_m \sim Ber(p_1),\ Y_1, \cdots, Y_n \sim Ber(p_2)$ 확률표본, 독립
- $\hat{p}_1 = \sum_{i=1}^{n} X_i/m\,, \hat{p}_2 = \sum_{i=1}^{m} Y_i/n$
- $P\left\{p_1 - p_2 \in \left(\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2}\sqrt{\hat{p}_1(1-\hat{p}_1)/m + \hat{p}_2(1-\hat{p}_2)/n}\right)\right\} = 1-\alpha$

4 모비율 차의 신뢰구간

모비율의 신뢰구간

- $P\left\{p_1 - p_2 \in \left(\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2}\sqrt{\hat{p}_1(1-\hat{p}_1)/m + \hat{p}_2(1-\hat{p}_2)/n}\right)\right\} = 1 - \alpha$

04

모분산에 대한 신뢰구간

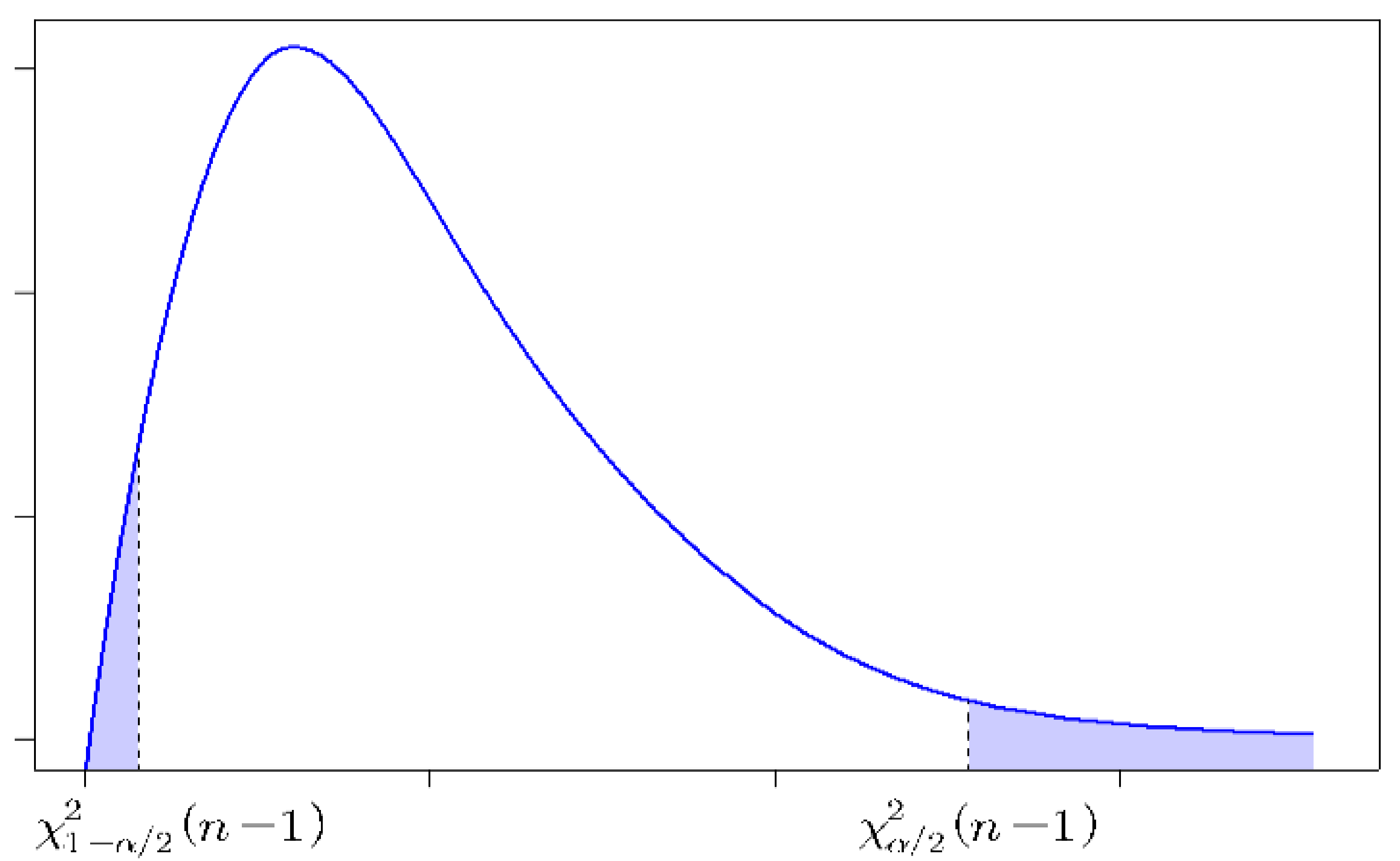
1 모분산의 신뢰구간

모분산의 신뢰구간

- $X_1, \cdots, X_n \sim N(\mu, \sigma^2)$ 확률표본
- $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$
- $P\left\{\sigma^2 \in \left(\frac{n-1}{\chi^2_{\alpha/2}(n-1)} S^2, \frac{n-1}{\chi^2_{1-\alpha/2}(n-1)} S^2\right)\right\} = 1 - \alpha$

1 모분산의 신뢰구간

모분산의 신뢰구간



2 모분산 비의 신뢰구간

모분산 비의 신뢰구간

- $X_1, \cdots, X_m \sim N(\mu_1, \sigma_1^2), Y_1, \cdots, Y_n \sim N(\mu_2, \sigma_2^2)$, 확률표본, 독립
- 두 표본분산의 비 : F분포

$$\frac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2} \sim F(n-1, m-1)$$

2 모분산 비의 신뢰구간

모분산 비의 신뢰구간

- $\frac{\sigma_2^2}{\sigma_1^2}$에 대한 $100(1-\alpha)\%$ 신뢰구간 :

$$\left[\frac{S_2^2/S_1^2}{F_{\alpha/2}(n-1,m-1)}, S_2^2/S_1^2 \cdot F_{\alpha/2}(m-1,n-1)\right]$$

정리하기

- 구간추정은 모수를 포함할 것으로 기대되는 구간을 제시하여 모수를 추정하는 방법이다.
- $X_1, \dots, X_n$이 $N(\mu, \sigma^2)$의 확률표본, σ^2를 모를 때, μ에 대한 $100(1-\alpha)\%$ 신뢰구간은 다음과 같다.

$$\left[\bar{X} - t_{\alpha/2}(n-1)S/\sqrt{n}, \qquad \bar{X} + t_{\alpha/2}(n-1)S/\sqrt{n}\right]$$

정리하기

- $X_1, \ldots, X_m \sim N(\mu_1, \sigma^2)$의 확률표본, $Y_1, \ldots, Y_n \sim N(\mu_2, \sigma^2)$의 확률표본, 두 표본 서로 독립. σ^2 모를 때 $\mu_1 - \mu_2$에 대한 $100(1-\alpha)\%$ 신뢰구간은 다음과 같다.

$$\left(\bar{X} - \bar{Y} \pm t_{\alpha/2}(m+n-2) S_p \sqrt{\frac{1}{m} + \frac{1}{n}}\right)$$

정리하기

- 표본 크기가 클 때 모비율 p의 $100(1-\alpha)\%$ 신뢰구간은 다음과 같다.

$$\left[\hat{p} - z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}\,, \hat{p} + z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}\right]$$

- $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$의 확률표본, σ^2에 대한 $100(1-\alpha)\%$ 신뢰구간은 다음과 같다.

$$\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \quad \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right]$$

통계적 추론 강의 정리

강의 내용

▪ **통계적 추론의 개념**	▪ **점추정**
▪ 확률	▪ 추정의 원리
▪ 모집단의 분포	▪ 가설검정
▪ 표본분포	▪ 구간추정

수고하셨습니다.