12강

통계적추론

가설검정 1

한국방송통신대학교 통계·데이터과학과
이긍희 교수

학습내용

1. 가설검정의 개념을 이해한다.
2. 최강력 검정의 개념을 이해한다.
3. 네이만-피어슨 보조정리를 이해한다.
4. 균일최강력 검정의 개념을 이해한다.

01

가설검정의 개념

1 통계적 추론

통계적 추론 : 추정과 검정

- 가설검정 : 주어진 확률표본으로부터 모집단의 가설에 대한 채택여부 결정
- 칼 피어슨의 카이제곱 검정, 고셋의 t-검정, 피셔의 검정과 네이만-피어슨의 검정

2 가설

- 가설 : 현상이나 모집단에 대한 진술
  - 대립가설(H_1) : 입증하고자 하는 내용의 가설
  - 귀무가설(H_0) : 대립가설에 반대되는 가설

2 가설

예 두 모집단의 평균이 차이가 있음을 알고자 한다.

$$H_0 : \mu_1 = \mu_2 \quad vs \quad H_1 : \mu_1 \neq \mu_2$$

2 가설

단순가설과 복합가설

- $\Omega = \Omega_0 \cup \Omega_1 , \Omega_0 \cap \Omega_1 = \phi \rightarrow H_0 : \theta \in \Omega_0 \ , \ H_1 : \theta \in \Omega_1$
- 단순가설 : $\Omega_0 = \{\theta_0\}$, $\Omega_1 = \{\theta_1\}$ 인 경우
- 복합가설 : Ω_i가 두 개 이상의 값을 포함하는 경우

2 가설

예

- 단순가설 $X \sim Poisson(\lambda)$

$$\mathrm{H}_0 : \lambda = 2 \;\; vs \; H_1 : \lambda = 8$$

- 복합가설 $\mathrm{H}_0 : \theta \leq 0.5 \;\; vs \; H_1 : \theta > 0.5$

3 검정의 기초

검정

- $H_0 : \theta \in \Omega_0 \;\; vs \;\; H_1 : \theta \in \Omega_1$
- $X = \{X_1, \dots, X_n\} : \theta$ 포함 확률분포의 확률표본
- 검정 : 확률표본 이용, H_0 를 기각하거나 채택

$$\delta(X) = \begin{cases} 0 : H_0 \text{ 채택} \\ 1 : H_0 \text{ 기각} \end{cases}$$

3 검정의 기초

기각역과 채택역

- 표본공간 $S = R \cup A$

  - 기각역 R : H_0 기각하는 S 부분집합

$$R = \{X \in S \,|H_0 \text{기각}, \delta(X) = 1\}$$

  - 채택역 A : H_0 채택하는 S 부분집합

$$A = \{X \in S \,|H_0 \text{채택}, \delta(X) = 0\}$$

3 검정의 기초

예

$$X_1, X_2 \sim Poisson(\theta)$$

$$H_0: \theta \leq 1 \ \ vs \ \ H_1: \theta > 1$$

$S = \{(x_1, x_2) \mid x_1, x_2$는 0 이상인 정수$\}$

$$R = \{(x_1, x_2) \in S \mid x_1 + x_2 > 2\}$$

$$A = \{(x_1, x_2) \in S \mid x_1 + x_2 \leq 2\}$$

4 검정의 오류

제1종의 오류와 제2종의 오류

		검정의 결과	
		H_0기각하지 않음	H_0기각
실제	H_0참	올바른 판단	제1종 오류
	H_1참	제2종 오류	올바른 판단

- 제1종 오류(type I error) : 참 귀무가설 기각

$$\alpha = P(\boldsymbol{X} \in R | H_0) = E(\delta(X) | H_0)$$

- 제2종 오류 (type II error) : 거짓 귀무가설 채택

$$\beta = P(\boldsymbol{X} \notin R | H_1) = 1 - P(X \in R | H_1)$$
$$= 1 - E(\delta(X) | H_1)$$

4 검정의 오류

검정력

- 대립가설이 참일 때 귀무가설을 기각할 확률

$$P(\boldsymbol{X} \in R | H_1) = 1 - \beta$$

4 검정의 오류

검정력

- 제1종의 오류 α와 제2종 오류 β, R={2,3}

관측결과	0	1	2	3
$H_0: p = \frac{1}{3}$	8/27	12/27	6/27	1/27
$H_1: p = \frac{2}{3}$	1/27	6/27	12/27	8/27

5 검정의 선택

검정의 선택

- 제1종의 오류와 제2종의 오류를 모두 줄이는 검정
  - 제1종의 오류와 제2종의 오류 간 상충 관계

→ 유의수준α에서 제2종 오류를 최소로 하는 검정
  - 유의수준 : 유의확률의 허용 상한

5 검정의 선택

검정의 수준

- 수준 α 검정(level α test) : $E(\delta(X)|H_0) \leq \alpha$
  - 제1종의 오류를 범할 확률이 α이하인 검정 $\delta(X)$

5 검정의 선택

예 $X \sim N(\mu, 1)$, $H_0 : \mu = 1, H_1 : \mu = 4$,

$R = \{x\,; x \geq 3\}$ 인 검정의 제1종 오류를 범할 확률,

제2종 오류를 범할 확률, 검정력을 구하시오.

5 검정의 선택

예

$X \sim N(\mu, 1)$, $H_0 : \mu = 1, H_1 : \mu = 4$,

$R = \{x\,;\, x \geq 3\}$ 인 검정의 제1종 오류를 범할 확률, 제2종 오류를 범할 확률, 검정력을 구하시오.

02

최강력 검정

1 최강력 검정의 정의

최강력 검정

- α에서의 최강력검정 $\delta^*(X)$
  - 단순가설 : $H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta = \theta_1$
  - $E(\delta(X)|\theta_0) \leq \alpha$의 모든 검정 $\delta(X)$ 중 다음을 만족하는 검정 $\delta^*(X)$

$$E(\delta^*(X)|\theta_0) \leq \alpha$$

$$E(\delta^*(X)|\theta_1) \geq E(\delta(X)|\theta_1)$$

2 네이만-피어슨 보조정리

네이만-피어슨 보조정리

- α에서의 최강력 검정 $\delta^*(X)$
  - **단순가설 :** $H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta = \theta_1$
  - $\boldsymbol{X} = (X_1, \dots, X_n) \sim f(x|\theta)$ 확률표본
  - 검정 δ는 유의수준 α_0의 최강력 검정

$$\frac{f(\boldsymbol{x}|\theta_1)}{f(\boldsymbol{x}|\theta_0)} > k, \quad (k > 0) \Rightarrow \delta(\boldsymbol{X}) = 1$$

$$\mathrm{E}(\delta(\boldsymbol{X})|\theta_0) = \alpha$$

2 네이만-피어슨 보조정리

예 $X_1, \cdots, X_n \sim N(\theta, 1)$ 확률표본. 최강력 검정은?

$$H_0 : \theta = 0 \quad vs \quad H_1 : \theta = 1$$

2 네이만-피어슨 보조정리

예 $X_1, \cdots, X_n \sim N(\theta, 1)$ 확률표본. 최강력 검정은?

$$H_0 : \theta = 0 \quad vs \quad H_1 : \theta = 1$$

2 네이만-피어슨 보조정리

예 $X_1, \cdots, X_n \sim N(\theta, 1)$ 확률표본. 최강력 검정은?

$$H_0 : \theta = 0 \quad vs \quad H_1 : \theta = 1$$

2 네이만-피어슨 보조정리

랜덤화 검정

- **랜덤화 검정 :** $H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta = \theta_1$
  - $X_1, \ldots, X_n \sim f(x|\theta)$의 확률표본
  - 유의수준 α의 검정 : $\delta(\boldsymbol{X}) = \begin{cases} 1, & \dfrac{f(\boldsymbol{x}|\theta_1)}{f(\boldsymbol{x}|\theta_0)} > k \\ \gamma, & \dfrac{f(\boldsymbol{x}|\theta_1)}{f(\boldsymbol{x}|\theta_0)} = k \\ 0, & \dfrac{f(\boldsymbol{x}|\theta_1)}{f(\boldsymbol{x}|\theta_0)} < k \end{cases}$
  - $E(\delta(X)|\theta_0) = \alpha$

2 네이만-피어슨 보조정리

예 $X_1, \cdots, X_n \sim Ber(\theta)$ 확률표본. 다음 가설에 대한 유의수준 0.05 에서의 최강력 검정은?

$$H_0 : \theta = 0.5 \quad vs \quad H_1 : \theta = 0.6$$

2 네이만-피어슨 보조정리

예 $X_1, \cdots, X_n \sim Ber(\theta)$ 확률표본. 다음 가설에 대한 유의수준 0.05 에서의 최강력 검정은?

$$H_0 : \theta = 0.5 \quad vs \quad H_1 : \theta = 0.6$$

2 네이만-피어슨 보조정리

예 $X_1, \cdots, X_n \sim Ber(\theta)$ 확률표본. 다음 가설에 대한 유의수준 0.05 에서의 최강력 검정은?

$$H_0 : \theta = 0.5 \quad vs \quad H_1 : \theta = 0.6$$

03

균일최강력 검정

1 균일최강력 검정의 정의

균일최강력 검정의 정의

- 복합가설 $H_0 : \theta \in \Omega_0 \quad vs \quad H_1 : \theta \in \Omega_1$
- $\alpha(\delta) \le \alpha$ 만족하는 모든 검정 $\delta(x)$에 대해 다음 성립

  - $\alpha(\delta) = \max_{\theta \in \Omega_0} E(\delta(X) \mid \theta)$

  - $\alpha(\delta^*) \le \alpha,\ E(\delta^*(X)|\theta) \ge E(\delta(X)|\theta),\ \theta \in \Omega_1$

  → 검정 $\delta^*(x)$를 유의수준α의 균일최강력 검정

1 균일최강력 검정의 정의

예 $X_1, \cdots, X_n \sim Ber(\theta)$ 확률표본. 다음 가설에 대한 균일최강력 검정은?

$$H_0 : \theta = 0.5 \quad vs \quad H_1 : \theta > 0.5$$

1 균일최강력 검정의 정의

예 $X_1, \cdots, X_n \sim Ber(\theta)$ 확률표본. 다음 가설에 대한 균일최강력 검정은?

$$H_0 : \ \theta = 0.5 \quad vs \quad H_1 : \theta > 0.5$$

2 단조가능도비와 균일최강력 검정

단조가능도비

- $\theta_0 < \theta_1$에 대하여 $\dfrac{f(\boldsymbol{x}|\theta_1)}{f(\boldsymbol{x}|\theta_0)}$: $T(\boldsymbol{x})$ 증가함수

  $\Rightarrow f(x|\theta)$: T에 대해 단조가능도비를 가짐

2 단조가능도비와 균일최강력 검정

단조가능도비

- $H_0 : \theta = \theta_0 \;\; vs \;\; H_1 : \theta > \theta_0$
- $f(\boldsymbol{x}|\theta)$가 $T(\boldsymbol{x})$에서 단조가능도비
- δ는 유의수준α의 균일최강력 검정

$$\delta(\boldsymbol{X}) = \begin{cases} 1, & T(\boldsymbol{X}) > k \\ \gamma, & T(\boldsymbol{X}) = k \\ 0, & T(\boldsymbol{X}) < k \end{cases}, \qquad E(\delta(\boldsymbol{X})|\theta_0) = \alpha$$

2 단조가능도비와 균일최강력 검정

예 $X_1, \cdots, X_n \sim N(\theta, 1)$ 확률표본. 다음 가설에 대한 유의수준 α에서의 균일최강력 검정은?

$$H_0 : \theta \le \theta_0 \quad vs \quad H_1 : \theta > \theta_0$$

2 단조가능도비와 균일최강력 검정

예 $X_1, \cdots, X_n \sim N(\theta, 1)$ 확률표본. 다음 가설에 대한 유의수준 α에서의 균일최강력 검정은?

$$H_0 : \theta \leq \theta_0 \quad vs \quad H_1 : \theta > \theta_0$$

2 단조가능도비와 균일최강력 검정

예 $X_1, \cdots, X_n \sim N(\theta, 1)$ 확률표본. 다음 가설에 대한 유의수준 α에서의 균일최강력 검정은?

$$H_0 : \theta \leq \theta_0 \quad vs \quad H_1 : \theta > \theta_0$$

2 단조가능도비와 균일최강력 검정

지수족과 단조가능도비

- 지수족 : 확률밀도(질량)함수가 다음과 같은 꼴

$$f(x|\theta) = \exp\big(Q(\theta)T(x) + S(x) + C(\theta)\big)$$

- $\boldsymbol{X} = (X_1, \cdots, X_n)$의 확률분포가 지수족

  $Q(\theta)$: θ 증가함수, $f(\boldsymbol{x}|\theta)$: $T(\boldsymbol{X})$에서 단조가능도비

2 단조가능도비와 균일최강력 검정

지수족과 단조가능도비

- $\boldsymbol{X}$의 확률분포 : 지수족, $Q(\theta)$가 θ의 증가함수
- $H_0 : \theta \leq \theta_0 \;\; vs \;\; H_1 : \theta > \theta_1$
- 검정 δ는 유의수준 α의 균일최강력 검정

$$\delta(\boldsymbol{X}) = \begin{cases} 1, & T(\boldsymbol{X}) > k \\ \gamma, & T(\boldsymbol{X}) = k, \\ 0, & T(\boldsymbol{X}) < k \end{cases} \qquad E(\delta(\boldsymbol{X})|\theta_0) = \alpha$$

2 단조가능도비와 균일최강력 검정

예 $X_1, \cdots, X_n \sim Ber(\theta)$ 확률표본. 다음 가설에 대한 유의수준 α에서의 균일최강력 검정은?

$$H_0 : \theta \le \theta_0 \quad vs \quad H_1 : \theta > \theta_0$$

2 단조가능도비와 균일최강력 검정

예 $X_1, \cdots, X_n \sim Ber(\theta)$ 확률표본. 다음 가설에 대한 유의수준 α에서의 균일최강력 검정은?

$$H_0 : \theta \leq \theta_0 \quad vs \quad H_1 : \theta > \theta_0$$

정리하기

- 제1종 오류와 제2종 오류로 구분되는데 제1종 오류는 참인 귀무가설을 기각하는 오류이며, 제2종 오류는 거짓인 귀무가설을 채택하는 오류이다.
- 유의확률은 참 H_0을 잘못 기각할 확률이며, 검정력은 거짓 H_0을 기각할 확률이다.

정리하기

- 단순가설 $H_0 : \theta = \theta_0 \;\; vs \;\; H_1 : \theta = \theta_1$에 대하여 $E(\delta(X)|\theta_0) \leq \alpha$을 만족하는 모든 검정 중 가장 검정력이 큰 검정을 유의수준 α에서 최강력 검정이라 한다.
- 네이만-피어슨 보조정리 : δ는 $H_0 : \theta = \theta_0 \;\; vs \;\; H_1 : \theta = \theta_1$에 대한 유의수준 α의 최강력 검정이다.

$$\frac{f(x|\theta_1)}{f(x|\theta_0)} > k, (k > 0)\text{이면 } \delta(X) = 1$$

$$E(\delta(X)|\theta_0) = \alpha$$

정리하기

- 복합가설 $H_0 : \theta \in \Omega_0$ vs $H_1 : \theta \in \Omega_1$에 대한 $\delta^*(X)$를 유의수준 α에서 균일최강력 검정이라 한다

  - $\alpha(\delta) = \max_{\theta \in \Omega_0} E(\delta(X)\,|\theta)$
  - $\alpha(\delta) \le \alpha$ 만족 모든 검정 $\delta(x), \alpha(\delta^*) \le \alpha$
  - $E(\delta^*(X)|\theta) \ge E(\delta(X)|\theta),\ \theta \in \Omega_1$

13 강

가설검정 2

다음시간 안내