10 강

통계적추론

점추정 2

한국방송통신대학교 통계·데이터과학과
이긍희 교수

학습내용

1. 불편추정량을 이해한다.
2. 일치추정량을 이해한다.
3. 추정량의 효율성을 이해한다.
4. 평균제곱오차를 이해한다.
5. 크래머-라오 하한을 이해한다.

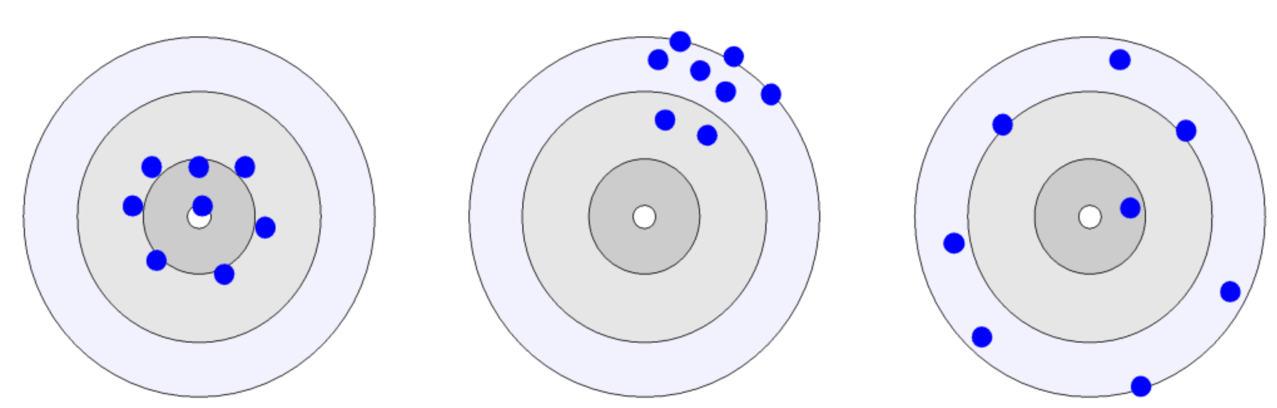
01

불편성

1 통계적 추정

통계적 추정

- 모집단 : 확률변수 $X \sim f(x|\theta)$
- 표본 : $X_1, X_2, \cdots X_n$ 독립 추출$\sim f(x|\theta)$
- 추정 : 통계량으로 모수 θ 추정 → 추정량
- 좋은 추정량 : 불편성, 효율성, 일치성



(a) (b) (c)

2 불편추정량

불편추정량의 정의

- 불편추정량 : 추정량의 모든 가능한 값을 평균하면 모수와 같아지는 추정량

$$E(T_n(\boldsymbol{X})) = \theta$$

- 편의 : $bias = E(T_n(\boldsymbol{X})) - \theta$

2 불편추정량

예 $X_1, X_2, \cdots, X_n \sim Poisson(\lambda)$ 의 확률표본일 때 $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ 의 편의는?

2 불편추정량

예 $X_1, X_2, \cdots, X_n \sim Poisson(\lambda)$ 의 확률표본일 때
$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ 의 편의는?

2 불편추정량

예 $X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2)$ 의 확률표본일 때
$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$, $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$
의 편의는?

2 불편추정량

예

$X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2)$의 확률표본일 때

$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$, $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$

의 편의는?

02

일치성

1 일치추정량

일치추정량의 정의

- 일치성 : 표본크기가 증가할수록 추정량이 확률적으로 모숫값으로 집중되어 가는 성질
- 일치추정량 : $T_n(\boldsymbol{X}) \xrightarrow{p} \theta$
  - 임의의 $\varepsilon > 0$, $\lim_{n\to\infty} P(|T_n(\boldsymbol{X}) - \theta| < \varepsilon) = 1$

1 일치추정량

예 $X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2)$ 의 확률표본일 때
$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ 가 μ 의 일치추정량임을 보이시오.

1 일치추정량

예 $X_1, X_2, \cdots, X_n \sim (\mu, \sigma^2)$ 의 확률표본일 때 $S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$ 가 σ의 일치추정량임을 보이시오.

1 일치추정량

예 $X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2)$ 의 확률표본일 때
$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$ 가 σ의 일치추정량임을
보이시오.

2 적률추정량

적률추정량과 일치추정량

- $X_1, X_2, \cdots X_n$: 모집단의 확률표본
- $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} E(X_1) = \mu$
- $m_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r \xrightarrow{p} E(X_1^r)$

03

효율성과 평균제곱오차

1 효율성

효율성의 정의

- 추정량의 변동성이 작다면 추정량 값의 신뢰도는 높아짐
- 불편추정량 $\boldsymbol{T}$의 효율성 : $eff(T) = \dfrac{1}{Var(T)}$

1 효율성

예 $X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2)$의 확률표본일 때 다음 두 추정량의 효율성을 비교하시오.

$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$, $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$

1 효율성

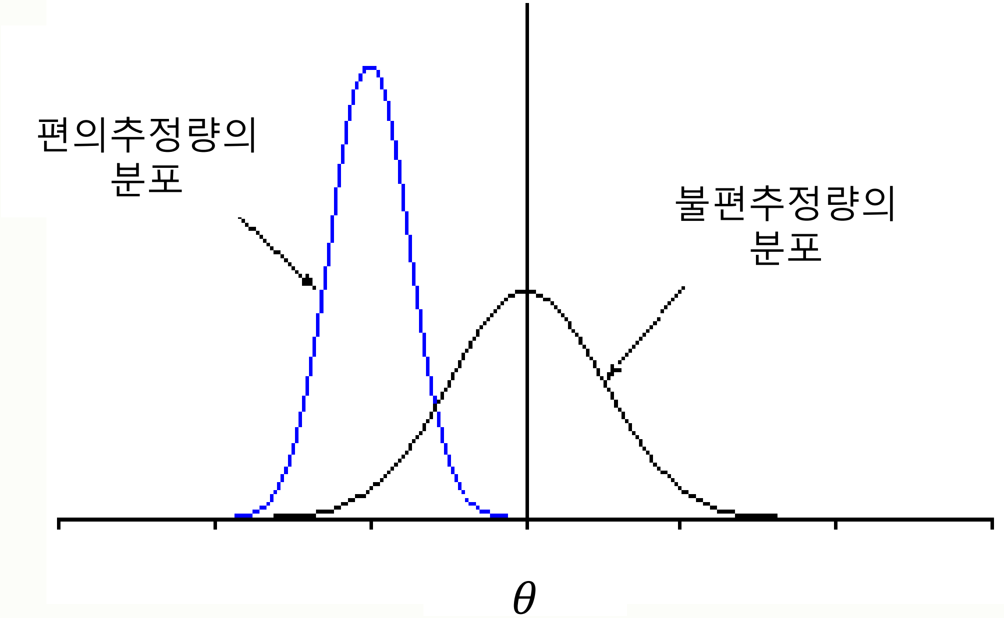
예 $X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2)$의 확률표본일 때 다음 두 추정량의 효율성을 비교하시오.

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2, \ \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

2 평균제곱오차

불편성과 효율성

- 불편성과 분산을 모두 고려한 추정량을 찾을 필요



2 평균제곱오차

평균제곱오차(mean square error)

- 편의와 효율성을 동시에 고려한 기준

$$MSE(T) = E[(T-\theta)^2]$$

2 평균제곱오차

- 평균제곱오차(mean square error)
  - $MSE(T) = Var(T) + bias(T)^2$

2 평균제곱오차

예 $X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2)$의 확률표본일 때 다음 두 추정량의 평균제곱오차를 비교하시오.

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2,\ \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

2 평균제곱오차

예 $X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2)$의 확률표본일 때 다음 두 추정량의 평균제곱오차를 비교하시오.

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2,\ \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

3 모수공간과 추정량

평균제곱오차 기반 추정량

- 모수공간에서 평균제곱오차를 최대로 하는 추정량 :

$$Max_{\theta\in\Omega}MSE(\theta,T)$$

- 모수공간에서 평균손실을 최소로 하는 추정량 :

$$\int_{\theta\in\Omega}MSE(\theta,T)\pi(\theta)d\theta$$

3 모수공간과 추정량

예

$X_1, X_2, \cdots, X_n \sim N(\mu, 1)$ 확률표본, $\pi(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\mu^2}$,

$-\infty < \mu < \infty$일 때 $\bar{X}$과 $\frac{n-1}{n}\bar{X}$의 평균손실을 비교.

3 모수공간과 추정량

예

$X_1, X_2, \cdots, X_n \sim N(\mu, 1)$ 확률표본, $\pi(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\mu^2}$,

$-\infty < \mu < \infty$일 때 $\bar{X}$과 $\frac{n-1}{n}\bar{X}$의 평균손실을 비교.

3 모수공간과 추정량

예

$X_1, X_2, \cdots, X_n \sim N(\mu, 1)$ 확률표본, $\pi(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\mu^2}$,

$-\infty < \mu < \infty$일 때 $\bar{X}$과 $\frac{n-1}{n}\bar{X}$의 평균손실을 비교.

04

크래머-라오 하한

1 피셔의 정보량

피셔의 정보량

- $I(\theta) = Var\left[\frac{\partial}{\partial\theta} logf(X_1;\theta)\right] = E\left[\frac{\partial}{\partial\theta} logf(X_1;\theta)\right]^2$

1 피셔의 정보량

피셔의 정보량

- 로그가능도함수가 2차 미분가능할 때

$$I(\theta) = E\left[\frac{\partial}{\partial\theta} logf(X_1;\theta)\right]^2 = E\left[-\frac{\partial^2}{\partial^2\theta} logf(X_1;\theta)\right]$$

1 피셔의 정보량

피셔의 정보량

- 로그가능도함수가 2차 미분가능할 때

$$I(\theta) = E\left[\frac{\partial}{\partial\theta} logf(X_1;\theta)\right]^2 = E\left[-\frac{\partial^2}{\partial^2\theta} logf(X_1;\theta)\right]$$

1 피셔의 정보량

피셔의 정보량

- $X_1, \ldots, X_n \sim f(x|\theta)$ 의 확률표본
- n개의 표본에 대한 피셔의 정보량 : $nI(\theta)$

2 크래머-라오 하한

크래머-라오 하한

- $\hat{\theta}$는 θ의 불편추정량일 때 하한

$$Var(\hat{\theta}) \geq \frac{1}{nVar\left(\frac{\partial}{\partial\theta}logf(X_1;\theta)\right)} = \frac{1}{nI(\theta)}$$

- 어떤 불편추정량의 분산이 크래머-라오 하한과 일치
  → 균일 최소분산 불편추정량

2 크래머-라오 하한

예

$X_1, X_2, \cdots, X_n \sim N(\mu, 1)$ 의 확률표본.

$\hat{\mu} = \bar{X}$의 크래머-라오 하한을 구하시오.

2 크래머-라오 하한

예

$X_1, X_2, \cdots, X_n \sim N(\mu, 1)$ 의 확률표본.

$\hat{\mu} = \bar{X}$의 크래머-라오 하한을 구하시오.

2 크래머-라오 하한

예 $X_1, X_2, \cdots, X_n \sim Ber(p)$ 의 확률표본. $\hat{p} = \bar{X}$ 가 불편추정량임을 보이고, $\mathrm{Var}(\hat{p})$ 와 크래머-라오 하한과 비교하시오.

2 크래머-라오 하한

예

$X_1, X_2, \cdots, X_n \sim Ber(p)$ 의 확률표본. $\hat{p} = \bar{X}$ 가 불편추정량임을 보이고, $\mathrm{Var}(\hat{p})$ 와 크래머-라오 하한과 비교하시오.

2 크래머-라오 하한

예

$X_1, X_2, \cdots, X_n \sim Ber(p)$ 의 확률표본. $\hat{p} = \bar{X}$ 가 불편추정량임을 보이고, Var$(\hat{p})$ 와 크래머-라오 하한과 비교하시오.

3 최대가능도추정량

최대가능도추정량의 근사분포

- $X_1, \dots, X_n \sim f(x|\theta)$ 의 확률표본
- $\hat{\theta}$: θ의 최대가능도추정량
- $\sqrt{n}(\hat{\theta}-\theta) \xrightarrow{d} N(0, \frac{1}{I(\theta)})$

3 최대가능도추정량

예

$X_1, X_2, \cdots, X_n \sim Ber(p)$ 의 확률표본. 최대가능도추정량 $\hat{p}$ 의 점근적 분포를 구하시오.

정리하기

- 불편성은 추정량 T의 기댓값이 모수 θ가 되어 치우침이 없는 성질을 의미한다.

$$E(T_n(\boldsymbol{X})) = \theta, \theta \in \Omega$$

- 일치성은 추정량이 모수에 확률적으로 수렴하는 성질을 가지는 것을 의미한다.

$$\lim_{n\to\infty} P(|T_n(\boldsymbol{X}) - \theta| < \varepsilon) = 1$$

정리하기

- 추정량 T의 효율성은 추정량 분산의 역수로 정의된다.

$$eff(T) = 1/Var(T)$$

- T가 θ의 추정통계량일 때 T의 평균제곱오차(MSE)는 편의와 분산을 모두 고려한 추정량의 비교기준이다.

$$MSE(T) = E[(T-\theta)^2]$$

정리하기

- 크래머-라오 하한은 불편추정량이 취할 수 있는 분산의 하한이다.

$$Var(\hat{\theta}) \geq \frac{1}{nVar\left(\frac{\partial}{\partial\theta} logf(X_1;\theta)\right)} = \frac{1}{nI(\theta)}$$

11 강

다음시간 안내

통계적 추정의 원리