[8강]

점추정

정보통계학과 이긍희 교수

학습목표

1. 통계적 추정의 개념을 이해한다.
2. 적률추정법을 이해한다.
3. 최대가능도 추정법을 이해한다.

[8강]

점추정

1 통계적 추정의 개념

1.1 추정

- 모집단에서 확률표본을 추출 → 모집단 추정
  - $X \sim f(x|\theta)$
  - 확률표본 $X_1, X_2, \cdots, X_n$을 모집단에서 추출

1.1 추정

- 추정 : 통계량의 분포를 이용하여 모수를 추정
  - 점추정 : 모수 추정값
  - 구간추정 : 모수 추정값 + 정확도

1.2 추정량

- 추정량 : 모수 추정을 위한 통계량
  - 적률추정량, 최대가능도추정량, 베이즈추정량 등

2 칼 피어슨과 피셔의 점추정

2.1 칼 피어슨(K. Pearson)의 추정법

- 관측값이 미지의 모수를 확률분포를 가짐
- 모집단의 확률분포를 데이터 기반의 4 개의 적률(평균, 분산, 왜도, 첨도)로 추정
- 적률추정량 제시

2.2 피셔(R.A. Fisher)의 추정법

- 확률분포는 수학적 함수, 수집된 데이터로 확률분포 추정
- 모집단에 임의로 추출된 표본의 통계량으로 모집단 추정
- 최대가능도 추정량 제시

3 적률추정량

3.1 적률추정법

- 적률이 존재하는 경우 적률추정량을 쉽게 구할 수 있음
  - 적률추정량보다 더 좋은 추정량도 있는 경우도 많음

3.1 적률추정법

- $X_1, \cdots, X_n \sim f(x|\theta_1, \cdots, \theta_k)$ 확률표본
  - r차 모적률 : $\mu_r = E(X^r)$
  - r차 표본적률 : $m_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r$

3.2 약대수법칙

- $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \mu_1 = E(X)$
- $m_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r \xrightarrow{p} \mu_r = E(X^r)$

3.3 적률추정량 구하는 방법

① $\mu_1, \cdots, \mu_k$는 $\theta_1, \cdots, \theta_k$의 함수

$$\mu_1 = g_1(\theta_1, \cdots, \theta_k),$$
$$\vdots,$$
$$\mu_k = g_k(\theta_1, \cdots, \theta_k)$$

3.3 적률추정량 구하는 방법

② 모수의 해를 구함

$$\theta_1 = h_1(\mu_1, \cdots, \mu_k),$$

$$\vdots,$$

$$\theta_k = h_k(\mu_1, \cdots, \mu_k)$$

3.3 적률추정량 구하는 방법

③ 적률추정량 : 표본적률을 대입하여 $\theta_1, \cdots, \theta_k$를 추정

$$\hat{\theta}_1 = h_1(m_1, \cdots, m_k),$$

$$\vdots,$$

$$\hat{\theta}_k = h_k(m_1, \cdots, m_k)$$

3.4 적률추정량의 특징

- 계산이 간단, 일치추정량
- 불편추정량이 아닌 경우도 있고 유일하지 않는 경우도 있음

3.4 적률추정량의 특징

예 5.1

$X_1, \cdots, X_n \sim (\mu, \sigma^2)$ 확률표본, μ, σ^2의 적률추정량은?

3.4 적률추정량의 특징

예 5.1

$X_1, \cdots, X_n \sim (\mu, \sigma^2)$ 확률표본, μ, σ^2의 적률추정량은?

3.4 적률추정량의 특징

예 5.2

$X_1, \cdots, X_n \sim B(k, p)$ 확률표본, p, k의 적률추정량은?

$$f(x;\theta) = \binom{k}{x} p^x (1-p)^{k-x},$$

$$0 < p < 1, \qquad x = 1,2,\cdots,k$$

3.4 적률추정량의 특징

예 5.2

$X_1, \cdots, X_n \sim B(k, p)$ 확률표본, p, k의 적률추정량은?

3.4 적률추정량의 특징

예 5.4

$X_1, \cdots, X_n \sim U(0, \theta)$ 확률표본, θ의 적률추정량은?

4 최대가능도추정량

4.1 최대가능도법

- 1821년 독일 수학자 가우스에 의해 제안
- 1922년 피셔에 의해 재발견되어 발전

4.1 최대가능도법

예 5.6

상자 안에 빨간색과 파란색 공 3개, 빨간 공과 파란색 공은 적어도 1개 이상.
공을 1개씩 3번 복원추출해서 나타난 파란색 공의 개수로 상자 안 공의 구성을 추정하라.

4.1 최대가능도법

예 5.6

파란색 공의 개수로 상자 안 공의 구성을 추정하라.

4.2 가능도함수

- $X_1, \cdots, X_n$: $f(x|\theta)$의 확률표본일 때 가능도함수

$$L(\theta) = L(\theta|x_1, \cdots, x_n) = \prod_{i=1}^{n} f(x_i|\theta)$$

- 가능도 함수에 확률표본에서 얻을 수 있는 모수의 모든 정보를 가지고 있음

4.3 최대가능도추정량의 정의

- θ에 대해 가능도함수 $L(\theta|\boldsymbol{x})$ 최대로 하는 통계량

$$\hat{\theta}(X_1, \cdots, X_n) = argmax_\theta L(\theta | x_1, \cdots, x_n)$$

- $\hat{\theta}(\boldsymbol{X})$: $\boldsymbol{X} = (X_1, \cdots, X_n)'$에 대한 모수 θ의

  최대가능도추정량

4.4 최대가능도추정법

- 가능도방정식을 풀어서 최대가능도추정량을 구함

$$\frac{d}{d\theta} logL(\theta) = 0$$

4.4 최대가능도추정법

- $f(x|\theta_1, \cdots, \theta_k)$ 확률표본의 가능도함수

$$L(\boldsymbol{\theta}|x) = \prod_{i=1}^{n} f(x_i|\theta_1, \cdots, \theta_k)$$

- $\theta_1, \cdots, \theta_k$에 대한 최대가능도 추정량

$$\frac{d}{d\theta_1} logL(\theta_1, \cdots, \theta_k) = 0$$

$$\vdots,$$

$$\frac{d}{d\theta_k} logL(\theta_1, \cdots, \theta_k) = 0$$

4.4 최대가능도추정법

예 5.7

$X_1, \cdots, X_n \sim B(1, p)$ 확률표본, p의 최대가능도추정량은?

4.4 최대가능도추정법

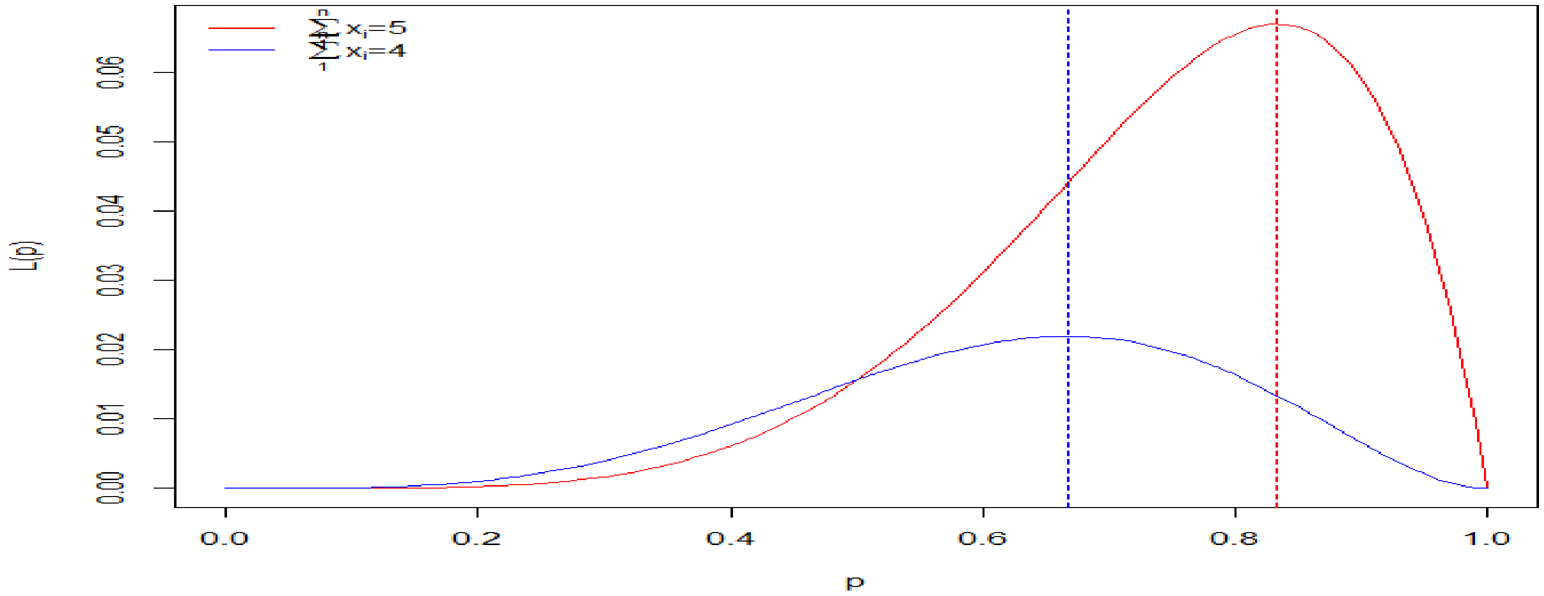
예 5.7

$X_1, \cdots, X_n \sim B(1, p)$ 확률표본, p의 최대가능도추정량은?

4.4 최대가능도추정법

예 5.8

$X_1, \cdots, X_6 \sim B(1, p)$ 확률표본, $\sum_{i=1}^{6} x_i = 4, 5$ 일 때 p의 최대가능도추정량 값은?



4.4 최대가능도추정법

예 5.9

$X_1, \cdots, X_n \sim Exp(\lambda)$ 확률표본, λ의 최대가능도추정량은?

4.4 최대가능도추정법

예 5.9

$X_1, \cdots, X_n \sim Exp(\lambda)$ 확률표본, λ의 최대가능도추정량은?

4.4 최대가능도추정법

예 5.11

$X_1, \cdots, X_n \sim N(\mu, \sigma^2)$ 확률표본, μ와 σ^2의 최대가능도추정량은?

4.4 최대가능도추정법

예 5.11

$X_1, \cdots, X_n \sim N(\mu, \sigma^2)$ 확률표본, μ와 σ^2의 최대가능도추정량은?

4.4 최대가능도추정법

예 5.11

$X_1, \cdots, X_n \sim N(\mu, \sigma^2)$ 확률표본, μ와 σ^2의 최대가능도추정량은?

4.4 최대가능도추정법

예 5.13

$X_1, \cdots, X_n \sim U(\alpha, \beta)$ 확률표본, α와 β의 최대가능도 추정량은?

4.4 최대가능도추정법

예 5.13

$X_1, \cdots, X_n \sim U(\alpha, \beta)$ 확률표본, α와 β의 최대가능도 추정량은?

4.5 최대가능도추정량의 특성

- 최대가능도추정량의 불변성
  - $\eta = g(\theta)$ 최대가능도추정량 : $\hat{\eta}^{MLE} = g\left(\hat{\theta}^{MLE}\right)$
- 최대가능도추정량은 근사적으로 불편성, 효율성, 일치성을 가지며 정규분포를 따름

4.5 최대가능도추정량의 특성

예 5.16

$X_1, \cdots, X_n \sim Exp(\lambda)$ 확률표본, $\theta = \frac{1}{\lambda}$의 최대가능도 추정량은?

5 정리하기

정리하기

1. 통계적 추정은 표본에 근거하여 모집단 모수를 추측하는 과정이다.
2. 적률추정법은 표본적률을 이용하여 모적률을 추정하는 방법이고, 적률추정량은 모수를 표본적률의 함수로 추정한 추정량이다.

정리하기

3. 최대가능도법은 가능도함수를 최대로 하는 모수값으로 모수를 추정하는 방법이며, 최대가능도추정량은 최대가능도법으로 구한 추정량이다.

다음 시간 안내 ▼

9강. 점추정량의 비교(1)

수고하셨습니다.