[15강]

사전분포와 베이즈 검정

서울대학교 통계학과 이재용 교수

학습목표

1. 사전분포
2. 정규모형
3. 베이즈 가설 검정

[15강]

사전분포와 베이즈 검정

1 사전분포

1.1 켤레사전분포

- 어떤 통계모형의 사후분포가 사전분포의 집합에 포함될 때, 사전분포의 집합을 그 통계모형의 켤레사전분포의 집합이라 하고, 그 원소를 켤레사전분포라 한다.

1.1 켤레사전분포

- $X|\theta \sim B(n,\theta)$이고, $\theta \sim Beta(\alpha,\beta)$이면,
  $\theta|X = x \sim Beta(\alpha + x, \beta + n - x)$이다.
  사전분포와 사후분포 모두 베타분포이므로,
  베타분포는 이항모형의 켤레사전분포이다.
- $X_1, X_2, \ldots, X_n \,|\, \theta \overset{i.i.d.}{\sim} N(\theta, \sigma^2), \sigma^2 > 0$는 알려진 값이라 하자.
  $\theta \sim N(\mu, \tau^2)$이면

$$\theta|X_1,\ldots,X_n \sim N\left(\frac{\frac{1}{\tau^2}\mu + \frac{n}{\sigma^2}\bar{x}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}}, \left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{-1}\right)$$ 이다.

따라서 정규분포는 정규모형의 켤레사전분포이다.

1.2 이항모형에서 베타사전분포 모수의 결정

- $X|\theta \sim B(n, \theta)$이고, $\theta \sim Beta(\alpha, \beta)$이면,
- $\theta|X = x \sim Beta(\alpha + x, \beta + n - x)$이다.

이 때, 사후평균은 다음과 같다.

$$\hat{\theta}^B = \frac{\alpha + x}{\alpha + \beta + n} = \frac{\alpha + \beta}{\alpha + \beta + n}\frac{\alpha}{\alpha + \beta} + \frac{n}{\alpha + \beta + n}\frac{x}{n}$$

- 베이즈 추정량은 사전분포의 평균과 자료로 이루어진 최대가능도추정량의 가중평균이고, 가중치는 각각 $\alpha + \beta$ 와 n이다.
- $\alpha + \beta$ 를 사전자료의 크기라 하고 사전분포가 가지고 있는 정보의 양으로 해석한다.

1.2 이항모형에서 베타사전분포 모수의 결정

- 이항 – 베타모형에서 자료를 보기 전에 θ의 그럴듯한 값이 θ_0이고, 이 정보는 관측치 k개가 갖고 있는 정보량 만큼 확신을 갖고 있다면

$$\frac{\alpha}{\alpha+\beta}=\theta_0,\ \alpha+\beta=k$$ 라 놓고,

이를 풀면 $\alpha=\theta_0 k,\ \beta=(1-\theta_0)k$ 를 얻는다.

1.2 이항모형에서 베타사전분포 모수의 결정

예 이항모형

- $X|\theta \sim B(n,\theta)$이고, $\theta \sim Beta(\alpha,\beta)$인 이항 – 베타모형을 고려하기로 하였다. 자료를 보기 전에 θ는 1 / 4근처일 것이라 예상되고 이는 약 0.1개의 관측치 만큼 확신이 든다고 할 때 사전분포의 모수 α,β를 구하라.

1.3 정규모형에서 정규사전분포 모수의 결정

- $X_1, X_2, \ldots, X_n | \theta \overset{i.i.d.}{\sim} N(\theta, \sigma^2), \sigma^2 > 0$는 알려진 값이라 하자. $\theta \sim N(\mu, \tau^2)$이라 하자. 이 때, 사후평균은 다음과 같다.

$$\hat{\theta}^B = \frac{\frac{1}{\tau^2}\mu + \frac{n}{\sigma^2}\bar{X}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}}$$

- 베이즈 추정량은 사전분포의 평균과 최대가능도추정량의 가중평균이고, 가중치는 각각 $1/\tau^2$ 와 n/σ^2 이다.
- 분산의 역수는 정밀도라 불리고 정규분포가 가지고 있는 정보의 크기로 해석한다.

1.3 정규모형에서 정규사전분포 모수의 결정

- 정규 – 정규모형에서 자료를 보기 전에
  θ의 그럴듯한 값이θ_0이고,
  이 정보는 관측치 k개가 갖고 있는 정보량 만큼
  확신을 갖고 있다면 θ 의 사전분포는

  $\theta \sim N\,(\theta_0, \sigma^2/\mathrm{k})$ 가 된다.

1.3 정규모형에서 정규사전분포 모수의 결정

예 정규모형

- $X_1, X_2, \ldots, X_n | \theta \overset{i.i.d.}{\sim} N(\theta, 4)$를 따르는 확률표본이고 θ의 사전분포를 구하라.
  θ의 사전분포는 정규분포를 쓰기로 하였다.
  자료를 보기 전에 θ는 10 정도로 예상되고 이는 관측치 0.5개 만큼의 확신이 든다.
  θ의 사전분포를 구하라.

1.4 무정보사전분포

- 주관적 사전분포를 이용하는 것이 이상적이지만 실제 분석에 사용할 때는 어려움이 따른다. 자신의 의견을 정확하게 확률분포로 나타내는 것은 많은 시간과 노력이 필요하다.

- 분석결과를 공유할 때는 모든 사람들이 수용할 수 있는 사전분포가 유용하다.
- 주관적 정보를 전혀 포함하지 않는 사전분포를 무정보사전분포라한다.
- $\pi(\theta) = 1, \forall\theta$를 균일사전분포라 한다.
- 많이 쓰는 무정보사전분포에는 균등사전분포와 제프리스의 사전분포가 있다.

2 정규모형

2.1 정규모형과 사전분포

- $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} N(\theta, \sigma^2)$을 따를 때, θ와 σ^2 모두 추론하고자 한다.
- θ 와 σ^2 의 사전분포로 다음의 분포를 고려한다.

$$\theta|\sigma^2 \sim N(\mu_0, \frac{\sigma^2}{K_0})$$

$$\sigma^2 \sim Inv - Gamma\ (\frac{v_0}{2}, \frac{v_0}{2}\sigma_0^2)$$

μ_0는 자료를 관측하기 전 θ 의 예측값, K_0는 θ 의 정보에 대한 사전정보의 크기, σ_0^2은 자료를 관측하기 전 σ^2 의 예상값이다.

- $X \sim Gamma\,(\alpha, \beta)$일 때, $1/X$의 분포를 역감마분포라 하고, $Inv - Gamma(\alpha, \beta)$라 표시한다.

2.2 사후분포

$$\theta \mid \sigma^2 \sim N(\mu_n, \frac{\sigma^2}{K_n})$$

$$\sigma^2 \sim Inv-Gamma\ (\frac{v\ n}{2}, \frac{v\ n}{2}\sigma_n^2)$$

사후분포의 모수들은 다음과 같이 정의된다.

$$\mu_n = \frac{k_0\mu_0 + n\bar{x}}{k_0 + n}$$

$$k_n = k_0 + n$$

$$v_n = v_0 + n$$

$$\sigma_n^2 = \frac{1}{v_n}\left[v_0\sigma_0^2 + (n-1)s^2 + \frac{k_0 n}{k_n}(\bar{x} - \mu_0)^2\right]$$

여기서 $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ 이다.

2.3 θ의 추정

- θ의 주변사후분포는

$$\theta\ |X_1,\ldots,X_n \sim t_{v_n}(\mu_n, \frac{\sigma^2}{K_n})$$ 이다.

여기서 $t_{v_n}(\mu, \sigma^2) \overset{d}{=} \mu + \sigma t_{v_n}$을 의미한다.

- $\theta^B = \mu_n$
- θ의 $100(1-\alpha)\%$ 신용구간은

$$\mu_n + t_{a/2}(v_n)\frac{\sigma_n}{\sqrt{K_n}}$$ 이다.

무정보사전분포를 쓰면, 신용구간은 빈도론자들의 신뢰구간과 동일하게 된다.

3 베이즈 가설 검정

3.1 베이즈 가설검정

- 모형

$$x \mid \theta \sim f(x \mid \theta), \theta \in \Theta$$

- 가설

$$H_0 : \theta = \theta_0 \ vs \ H_1 : \theta = \theta_1$$

- 혹은

$$H_0 : x \sim f(x \mid \theta_0) \ vs \ H_1 : x \sim f(x \mid \theta_1)$$

3.2 가설검정의 사전분포

- $\pi_0 = \pi(H_0)\pi_1 = \pi(H_1)$
  여기서 $\pi_0 + \pi_1 = 1, 0 < \pi_0, \pi_1 < 1$이다
- 자료를 보기 전에 H_0와 H_1가 동일한 가능성을 갖는다고 생각되면 다음과 같이 놓는다.

$$\pi_0 = \pi_1 = \frac{1}{2}$$

3.3 사후확률

$$\pi(H_0|x) = \frac{\pi_0 f\,(x|H_0)}{\pi_0 f\,(x|H_0) + \pi_1 f\,(x|H_1)}$$

여기서 $f\,(x|H_0) = f\,(x|\theta_0)$이고 $f\,(x|H_1) = f\,(x|\theta_1)$이다.

3.4 베이즈 인수(Bayes Factor)

- 사후확률의 비(Posterior odds)

$$\frac{\pi(H_1|x)}{\pi(H_0|x)} = \frac{\pi_1 f\ (x|H_1)}{\pi_0 f\ (x|H_0)}$$

혹은,

$$\text{사후확률의 비} = \frac{f(x\ |H_1)}{f(x\ |H_0)} \times \text{사전확률의 비}$$

- 베이즈 인수(베이즈 인자, 베이즈 팩터)

$$B_{10} = \frac{\pi_1 f(x\ |H_1)}{\pi_0 f(x\ |H_0)},\ B_{01} = \frac{1}{B_{10}}$$

- 사후확률의 비 = 베이즈 인수 × 사전확률의 비

- 사전확률의 비가 1일때 즉, $\pi_0 = \pi_1 = 1/2$일때, 베이즈 인수는 사후확률의 비와 같다.

3.4 베이즈 인수(Bayes Factor)

압정의 예

- 모형 $x \mid \theta \sim Bin(n = 10, \theta)$
- $x = 7$을 관측했다고 하자.
- 다음의 가설 검정을 고려해보자.

  $H_0 : \theta = 1/2 \ vs \ H_1 : \theta = 2/3$
- 사전확률

  $\pi_0 = \pi_1 = 1/2$

3.4 베이즈 인수(Bayes Factor)

- 베이즈 인수

$$B_{10} = \frac{f\ (x\ =\ 7|\theta\ =\ 2/3)}{f\ (x\ =\ 7|\theta\ = 1/2)}$$
$$= \frac{\binom{10}{7}(2/3)^7(1/3)^3}{\binom{10}{7}(1/2)^7(1/2)^3}$$
$$=\ 2.2197$$

- 가설들의 사후확률

$$\pi(H_1|x) = \frac{\pi_1 B_{10}}{\pi_1 B_{10} + \pi_0} =\ 0.6894$$
$$\pi(H_0|x)\ =\ 1\ -\ \pi(H_1|x)\ =\ 0.3106$$

3.5 베이즈 팩터의 값에 대한 제프리스(Jeffreys)의 기준

$\log_{10} B_{10}$	B_{10}	H_1에 대한 증거의 정도
$0 - \frac{1}{2}$	$1 - 3.2$	언급할 만한 가치가 없는 증거 (*not worth a bare mention*)
$\frac{1}{2} - 1$	$3.2 - 10$	상당한 증거(*substantial*)
$1 - 2$	$10 - 100$	강한 증거(*strong*)
> 2	> 100	결정적 증거(*decisive*)

한 학기 동안

수고하셨습니다.

감사합니다.