[14강]

베이즈 추론의 기초

서울대학교 통계학과 이재용 교수

학습목표

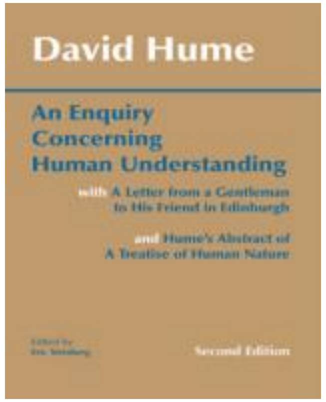
1. 베이즈 추론의 아버지: 베이즈
2. 베이즈 추론
3. 사후분포의 계산
4. 베이즈추정량과 신용집합

[14강]

베이즈 추론의 기초

1 베이즈 추론의 아버지: 베이즈

1.1 토마스 베이즈(Thomas Bayes)



- 토마스 베이즈는 1702년에 태어나 1761년 4월 17일에 사망하였다.
- 베이즈는 영국 장로교 목사이자 아마추어 수학자였다.
- 흄 (*David Hume*) 은 인과관계를 인간이 파악할 수 있다는 믿음을 의심했다. 이에 대한 반박으로 베이즈는 논문을 썼다.

2 베이즈 추론

2.1 통계적 추론의 문제

- 모수 (θ) 자연의 법칙을 나타내는 미지의 값
- 관측치 (x) 확률분포 $p(x|\theta)$ 를 따르는 확률변수
- 목표 x를 기초로 모수 θ 에 대한 추론을 하고자 한다.

2.1 통계적 추론의 문제

예 압정

- 문제

  λ의 확률을 알고자 한다.

2.1 통계적 추론의 문제

예 압정

- 압정 10 개를 던져서
  7 개의 ⋏ 와
  3 개의 ⊥ 를 얻었다.
- ⋏ 의 확률에 대해 어떤 결론을 내릴 수 있나?

2.1 통계적 추론의 문제

예 압정

- 모수 θ = ⋋의 확률
- 관측치 x =압정을 10 번 던졌을 때 ⋋가 나온 횟수
- 문제 x = 7를 관측했을 때 θ 에 관한 추론은?

2.2 베이즈 추론

베이즈 추론은 다음의 세 가지 요소로 이루어져 있다.

- 사전분포: θ의 분포로 자료를 보기 전의 분석자가 갖고 있는 θ에 관한 정보를 나타낸다. $\pi(\theta)$ 로 나타낸다.
- 확률모형: $x|\theta \sim p(x|\theta)$
- 사후분포: 자료가 주어졌을 때 θ의 확률분포로, 자료를 본 후의 분석자가 갖고 있는 θ에 관한 정보를 나타낸다. $\pi(\theta|x)$ 로 나타낸다.

[14강]

베이즈 추론의 기초

3 사후분포의 계산

3.1 베이즈 법칙: 사후분포의 계산방법

- 베이즈 법칙 혹은 베이즈 정리는 사후분포를 계산하는 수학적 방법이다.

$$\pi(\theta|x) \propto \pi(\theta)p(x|\theta)$$

혹은 사후분포 ∝ 사전분포 × 가능도

3.2 압정의 예: 사후분포의 계산

- 사전분포: $\theta \sim Uniform(0,1)$

  자료 x를 보기 전에 모든 θ 값이 동일한 가능성을 갖고 있다면

  즉, 모든$\theta_1 \neq \theta_2$에 대하여 $\pi(\theta_1) = \pi(\theta_2)$라면, $\theta \sim Uniform(0,1)$

- 가능도: $x|\theta \sim Binomial(n,\theta)$

- 사후분포: $\pi(\theta|x) \propto \pi(\theta) \times p(x|\theta)$

$$= 1 \times \binom{n}{x}\theta^{x},(1-\theta)^{n-x}$$

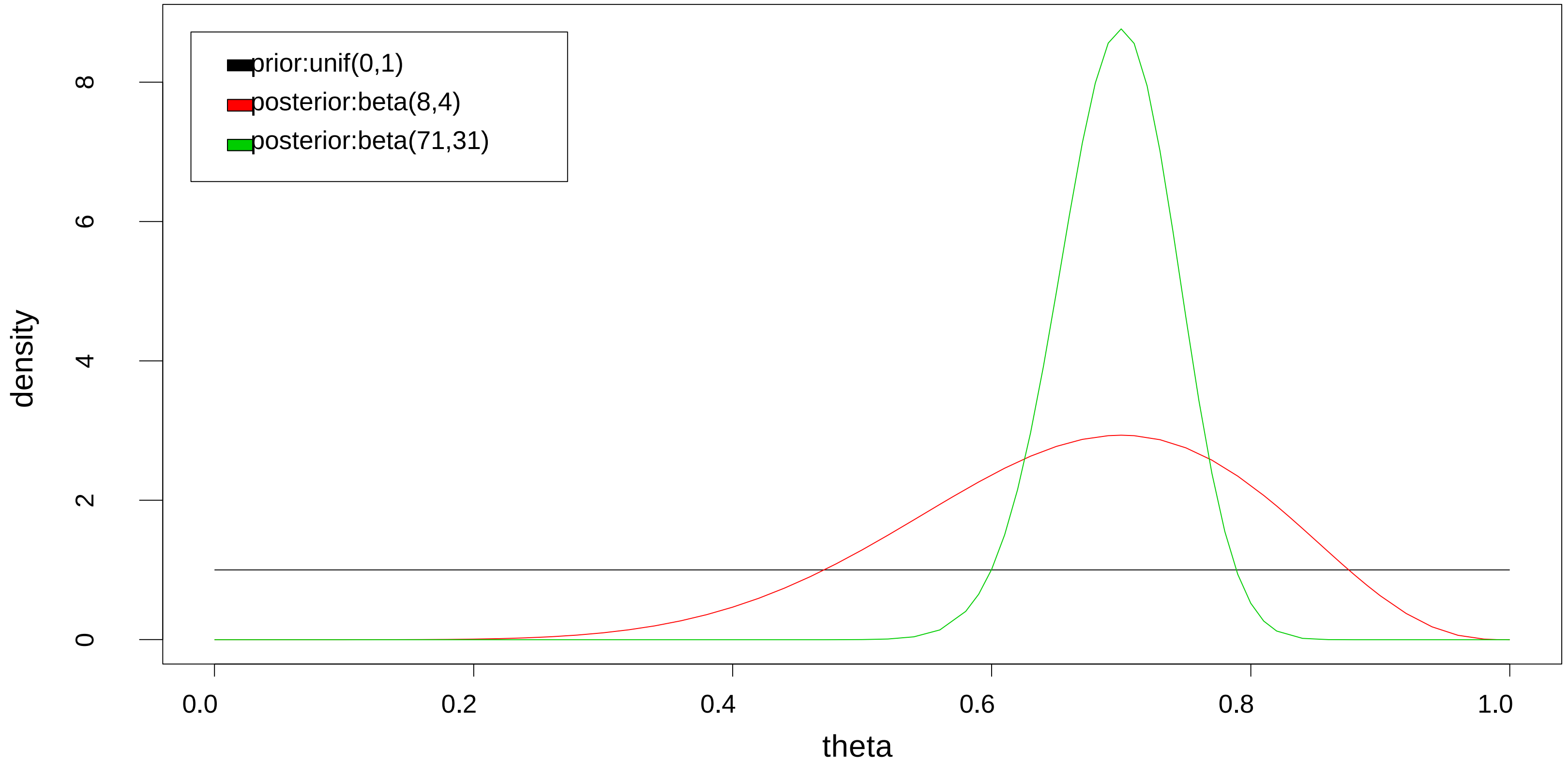
$$\propto \theta^{x}(1-\theta)^{n-x}, 0 < \theta < 1$$

$$\theta|x \sim Beta(x+1, n-x+1)$$

- 압정의 예에서는, $\theta|x = 7 \sim Beta(8,4)$

3.3 압정의 예: 사후분포의 계산

- 사후확률밀도함수



3.3 압정의 예: 사후분포의 계산

예 이항모형

확률변수 X 가 이항분포 $B(n,\theta)(n = 1,2,\ldots,0 < \theta < 1)$ 를 따르고 θ 의 사전분포가 $Beta(\alpha,\beta)(\alpha,\beta > 0)$ 를 따른다고 하자. $X = x(x = 0,1,2,\ldots,n)$ 를 관측했을 때, θ 의 사후분포를 구하라.

3.3 압정의 예: 사후분포의 계산

예 정규모형

$X_1, X_2, \ldots, X_n$이 $N(\theta, \sigma^2)(\theta \in \mathbb{R},\ \sigma^2 > 0)$을 따르는 확률표본이라 하자. σ^2 은 알려져 있다고 하자.

θ의 사전분포가 $N(\mu, \tau^2)(\mu \in \mathbb{R},\ \tau^2 > 0)$이라 하자.

$X_1, X_2, \ldots, X_n$가 주어졌을 때, θ의 사후분포를 구하라.

4 베이즈추정량과 신용집합

4.1 베이즈 추정량

- 베이즈 추정량은 사후분포의 한 점 요약이다.
- 보통 사후분포의 평균(*Posterior mean*),
  최대사후분포추정량(*MAP, Maximum a Posteriori*),
  사후분포의 중앙값(*Posterior median*) 등을 쓴다.
- 압정의 예
  사후분포는 $\theta|x \sim Beta(x+1,\ n-x+1)$ 이고, 사후분포의 평균은

$$\mathbb{E}(\theta|x) = \frac{x+1}{n+2} = \frac{7+1}{10+2} = 0.667$$ 이다.

4.1 베이즈 추정량

예 이항모형

- 확률변수 X가 이항분포$B(n,\theta)(n = 1,2,\ldots,0 < \theta < 1)$를 따르고 θ의 사전분포가$Beta(\alpha,\beta)(\alpha,\beta > 0)$를 따른다고 하자. $X = x(x = 0,1,2,\ldots,n)$를 관측했을 때, θ의 사후분포의 평균값, 최빈값을 구하라.
- 위의 문제에서 θ의 사전분포가$Beta(1,1)$이고, $X = 3$을 관측하였을때, 사후분포의 중앙값을 구하라.

4.1 베이즈 추정량

예 정규모형

$X_1, X_2, \ldots, X_n$이 $N(\theta, \sigma^2)(\theta \in \mathbb{R}, \sigma^2 > 0)$을

따르는 확률표본이라 하자. σ^2은 알려져 있다고 하자.

θ의 사전분포가 $N(\mu, \tau^2)$ $(\mu \in \mathbb{R}, \tau^2 > 0)$이라 하자.

$X_1, X_2, \ldots, X_n$가 주어졌을 때,

θ의 사후분포의 평균값, 중앙값, 최빈값을 구하라.

4.2 신용구간 (Credible Interval)

- 베이즈 구간 추정량을 신용구간이라 부른다.
- 100 $(1-\alpha)$% 신용구간은 다음을 만족하는 (L, U)이다.

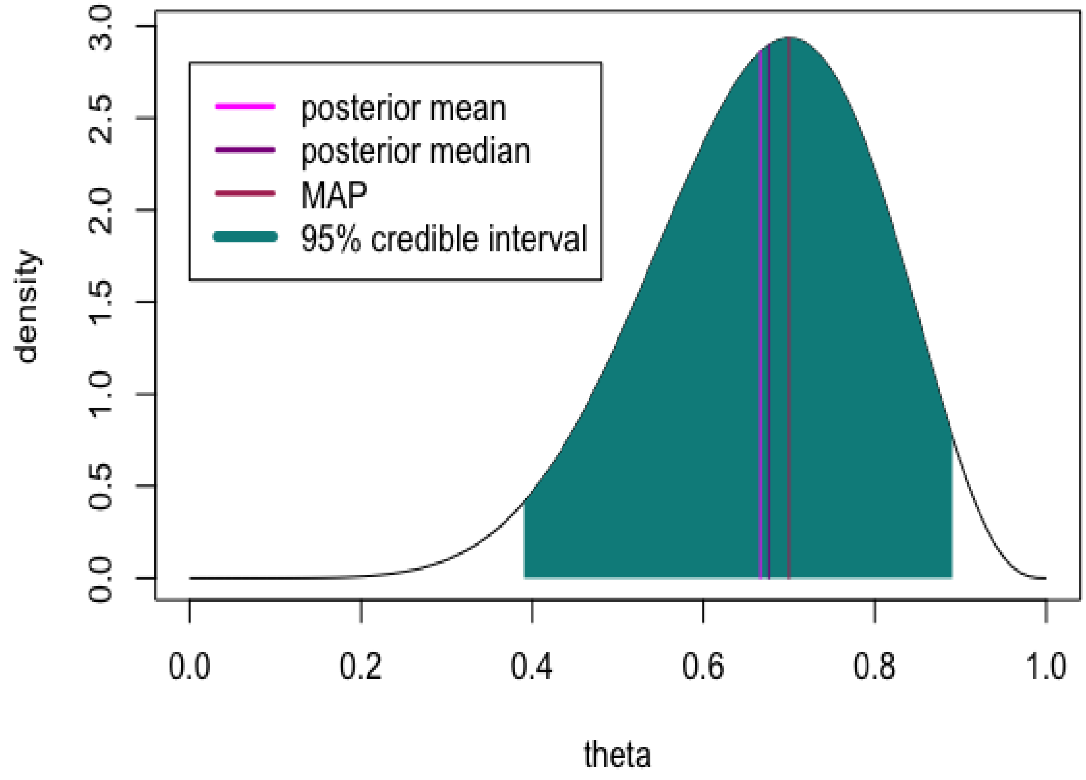
$$P[L < \theta < U|x] = 1 - \alpha$$

- 베이즈 추론에서 구간추정의 해석은 일반인들이 자연스럽게 생각하는 해석과 동일하다.

  즉, θ 가 $[L, U]$ 에 포함될 확률이 $1 - \alpha$ 이다.

4.2 신용구간 (Credible Interval)

압정의 예



- θ의 95% 신용구간은

$$(0.390, 0.891)$$

이다.

4.2 신용구간 (Credible Interval)

예 이항모형

확률변수 X가

이항분포 $B(n, \theta)(n = 1, 2, \ldots, 0 < \theta < 1)$를

따르고 θ의 사전분포가 $Beta(1, 1)$이고,

$X = 3$을 관측하였을 때,

사후분포의 95% 신용구간을 구하라.

4.2 신용구간 (Credible Interval)

예 정규모형

$X_1, X_2, \ldots, X_{10}$이

$N(\theta, 1)(\theta \in \mathbb{R})$을 따르는 확률표본이고,

θ의 사전분포가 $N(0, 10)$이라 하자.

관측된 표본평균이 $\bar{x} = 10.2$일 때,

θ의 95% 최고사후밀도구간을 구하라.

4.3 구간추정의 해석

- 빈도론자의 95% 신뢰구간 해석

  동일한 실험을 무한히 반복해서 동일한 방법으로 신뢰구간들을 계산하면 이 중 95%의 신뢰구간이 모수 θ를 포함한다.
  이 해석은 당장 우리 손 안에 있는 신뢰구간에 대해 아무런 말도 하지 못한다.

- 베이즈주의자의 95% 신용구간에 대한 해석

  주어진 자료에 대해, 계산된 신용구간이 모수를 포함할 확률이 95%이다.

4.4 확률의 의미

- 빈도론적 해석

  빈도론적 해석은 무한히 반복 가능한 실험에 대해서만 적용 가능하다. 사건 A의 확률은

$$\frac{\text{사건 } A\text{의 빈도}}{\text{실험의 반복수}}$$

- 주관적 확률의 해석

  개인적 믿음의 정도

- 베이즈주의자들은 주관적 확률의 해석을 믿고 모든 불확실성은 확률로 표현 가능하다고 믿는다.

4.5 베이즈주의자들이 믿는 것들

- 미지의 모수에 대한 추론은 자료(와 사전분포)에 대해서만 의존해야하고 관측할 수도 있었던 값들에 의존하면 안된다.

  놀랍게도 빈도론적 추론에서는 관측될 수도 있으리라 믿어지는 값들이지만 정작 관측되지는 않은 값들에 의해 추론이 의존한다.

- 모든 불확실성은 확률로 표현 가능하다.

  따라서 모수에 대한 불확실성은 자료가 관측되었건 안되었건 확률분포로 표현 가능하다.

다음 시간 안내▼

15강. 사전분포와 베이즈 검정

수고하셨습니다.