베이즈 통계 심화스터디 week 2-2

prior & posterior distribution

배시예

Posterior, Likelihood, Prior

- posterior, likelihood, prior

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{\sum_{i=1}^{m} P(B|A_i)P(A_i)}$$

Prior

- 사전확률
- 주관적
- 예시: 동전 뒤집기(남동생에 대한 사전 정보)
- 0, 1로 두는 것은 좋지 않음

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{\sum_{i=1}^{m} P(B|A_i)P(A_i)}$$

- Hyperparameters에도 확장 가능

Posterior

- 사후확률
- 예시: 동전 뒤집기(additional observations)
- Information in prior + Information in data
- 데이터가 매우 많다면, prior의 영향력 少

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{\sum_{i=1}^{m} P(B|A_i)P(A_i)}$$

Sequential Analysis

- update, update, … 거듭한 것과 한번에 update한 것은 같은 결과

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{\sum_{i=1}^{m} P(B|A_i)P(A_i)}$$

θ

$\theta | y_1, \cdots, y_n$

$\theta | y_1, \cdots, y_{n+m}$

Predictive distribution

- Prior predictive distribution(사전 예측 분포)

$$f(\tilde{x}) = \int f(\tilde{x}, \theta) d\theta = \int f(\tilde{x}|\theta) \boxed{f(\theta)} d\theta$$

- Posterior predictive(사후 예측 분포)

$$f(\tilde{x}|x) = \int f(\tilde{x}, \theta|x) d\theta = \int f(\tilde{x}|\theta, x) \boxed{f(\theta|x)} d\theta$$

$If\ x\ and\ \tilde{x}\ are\ independent,$

$$f(\tilde{x}|x) = \int f(\tilde{x}|\theta, x) f(\theta|x) d\theta = \int f(\tilde{x}|\theta) f(\theta|x) d\theta$$

Credible Interval

- Confidence Interval(신뢰 구간) : Frequentist

-모수는 고정된 상수, 신뢰구간은 랜덤

-probability statement 불가

- Credible Interval(신용 구간) : Bayesian

-모수는 분포를 가진 변수, 신용구간은 사후분포로 구해짐

-actual probability of containing theta

Conjugate

- 사전분포와 가능도(likelihood)가 특정한 짝을 이루고 있다면, 이로부터 추출되는 사후분포는 사전분포와 동일한 형태를 가짐
- 매우 편리

$$P(A_1|B) \propto P(B|A_1)P(A_1)$$

Ex1) Bernoulli & Binomial

- 동전 뒤집기-H가 나온 횟수

$$f(\underset{\sim}{y}|\theta) = \theta^{\Sigma y_i} (1-\theta)^{n-\Sigma y_i}$$

- Beta prior, Beta posterior (conjugate)

$$f(x;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, x^{\alpha-1}(1-x)^{\beta-1}$$

Ex1) Bernoulli & Binomial

- Prior, Data의 contribution
- Effective sample size

$$f(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1} I_{\{0 \le \theta \le 1\}} \qquad (\theta \sim Beta(\alpha, \beta))$$

$$f(\theta|\underset{\sim}{y}) \propto f(\underset{\sim}{y}|\theta) f(\theta) = \theta^{\Sigma y_i} (1-\theta)^{n-\Sigma y_i} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1} I_{\{0 \le \theta \le 1\}}$$

$$= \theta^{\alpha+\Sigma y_i - 1} (1-\theta)^{\beta+n-\Sigma y_i - 1} I_{\{0 \le \theta \le 1\}}$$

$$\therefore \theta|\underset{\sim}{y} \sim Beta(\alpha+\Sigma y_i, \beta+n-\Sigma y_i)$$

Ex1) Bernoulli & Binomial

- Prior, Data의 contribution
- Effective sample size

mean of beta : $\frac{\alpha}{\alpha+\beta}$

posterior mean : $\frac{\alpha+\Sigma y_i}{\alpha+\Sigma y_i+\beta+n-\Sigma y_i}$

$$= \frac{\alpha+\beta}{\alpha+\beta+n} \cdot \frac{\alpha}{\alpha+\beta} + \frac{n}{\alpha+\beta+n} \cdot \frac{\Sigma y_i}{n}$$

Ex2) Poisson

- 초코칩 쿠키-쿠키 하나당 들어있는 초코칩 수

$$f(\underset{\sim}{y}|\lambda) = \frac{\lambda^{\Sigma y_i} e^{-n\lambda}}{\prod_{i=1}^{n} y_i!} \quad \text{for } \lambda > 0$$

- Gamma prior, Gamma posterior (conjugate)

$$f(x;k,\theta) = x^{k-1}\frac{e^{-x/\theta}}{\theta^k\,\Gamma(k)} \text{ for } x > 0$$

Ex2) Poisson

- Prior, Data의 contribution
- Effective sample size

$$f(\lambda) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} \qquad (\lambda \sim \Gamma(\alpha, \beta))$$

$$f(\lambda \mid \underset{\sim}{y}) \propto f(\underset{\sim}{y} \mid \lambda) f(\lambda) \propto \lambda^{\Sigma y_i} e^{-n\lambda} \lambda^{\alpha-1} e^{-\beta\lambda}$$

$$= \lambda^{\alpha + \Sigma y_i - 1} e^{-(\beta+n)\lambda}$$

$$\therefore \lambda \mid \underset{\sim}{y} \sim \Gamma(\alpha + \Sigma y_i, \beta + n)$$

Ex2) Poisson

- Prior, Data의 contribution
- Effective sample size

mean of gamma : $\frac{\alpha}{\beta}$

posterior mean : $\frac{\alpha + \Sigma y_i}{\beta + n} = \frac{\beta}{\beta + n} \cdot \frac{\alpha}{\beta} + \frac{n}{\beta + n} \cdot \frac{\Sigma y_i}{n}$

Ex2) Poisson

- Hyperparameters(alpha, beta)의 결정

1) prior mean $\frac{\alpha}{\beta}$

   a) prior standard deviation $\frac{\sqrt{\alpha}}{\beta}$

   b) effective sample size β

Ex2) Poisson

- Hyperparameters(alpha, beta)의 결정

2) vague prior

$\varepsilon > 0$ $\quad \Gamma(\varepsilon, \varepsilon) \Rightarrow \begin{cases} \text{mean} : \dfrac{\varepsilon}{\varepsilon} = 1 \\ \text{variance} : \dfrac{1}{\varepsilon} \to \infty \quad 大 \end{cases}$

posterior mean : $\dfrac{\varepsilon + \Sigma y_i}{\varepsilon + n} \longrightarrow \dfrac{\Sigma y_i}{n} \doteq$ data mean

as $\varepsilon \to 0$