

TDAB01 Probability and Statistics

Maryna Prus
IDA, Linköping University

Lecture 10: Bayesian Inference

Overview

- **Prior and posterior**
- **Conjugate distribution families**
- **Bayesian estimator, confidence intervals and hypothesis testing**

Prior and posterior distributions

- ▶ X_1, \dots, X_n iid, $\theta \in \Theta$ unknown parameter
- ▶ Consider θ as random variable
- ▶ Discretization of Θ : $\theta \in \{\theta_1, \theta_2, \dots, \theta_K\}$

$$P(\theta = \theta_i | \mathbf{x}) = \frac{P(\mathbf{x} | \theta = \theta_i) \mathbf{P}(\theta = \theta_i)}{\sum_{j=1}^K P(\mathbf{x} | \theta = \theta_j) \mathbf{P}(\theta = \theta_j)}$$

- ▶ For $(\theta_{i+1} - \theta_i \rightarrow 0)$ we obtain

$$f(\theta | \mathbf{x}) = \frac{f(\mathbf{x} | \theta) f(\theta)}{\int f(\mathbf{x} | \theta) f(\theta) d\theta}$$

- ▶ **Bayes Theorem** for **continuous** parameter θ

$$\pi(\theta | \mathbf{x}) = \frac{f(\mathbf{x} | \theta) \pi(\theta)}{\int f(\mathbf{x} | \theta) \pi(\theta) d\theta}$$

- ▶ **Prior**: $\pi(\theta)$
- ▶ **Likelihood**: $f(\mathbf{x} | \theta)$
- ▶ **Posterior**: $\pi(\theta | \mathbf{x})$

Bernoulli model with beta prior

- ▶ Bernoulli model: $X_1, \dots, X_n | \theta \sim \text{Bernoulli}(\theta)$, $\theta \in (0, 1)$
- ▶ **Likelihood:** $\theta^s (1 - \theta)^f$, $f + s = n$ & **Prior:** $\theta \sim \text{Beta}(\alpha, \beta)$, i.e

$$\pi(\theta) = \frac{1}{B(\alpha, \beta)} \theta^{\alpha-1} (1 - \theta)^{\beta-1}$$

- ▶ **Posterior:**

$$\begin{aligned} \pi(\theta | \mathbf{x}) &= \frac{f(\mathbf{x} | \theta) \pi(\theta)}{\int f(\mathbf{x} | \theta) \pi(\theta) d\theta} = \frac{\theta^s (1 - \theta)^f \frac{1}{B(\alpha, \beta)} \theta^{\alpha-1} (1 - \theta)^{\beta-1}}{\int \theta^s (1 - \theta)^f \frac{1}{B(\alpha, \beta)} \theta^{\alpha-1} (1 - \theta)^{\beta-1} d\theta} \\ &= \frac{\theta^{\alpha+s-1} (1 - \theta)^{\beta+f-1}}{\int \theta^{\alpha+s-1} (1 - \theta)^{\beta+f-1} d\theta} = c \cdot \theta^{\alpha+s-1} (1 - \theta)^{\beta+f-1} \end{aligned}$$

where $c = 1 / \int \theta^{\alpha+s-1} (1 - \theta)^{\beta+f-1} d\theta$ is constant (does not depend on θ)

- ▶ Density of form $c \cdot \theta^{\alpha+s-1} (1 - \theta)^{\beta+f-1}$ is recognized as $\text{Beta}(\alpha + s, \beta + f)$:

$$\pi(\theta | \mathbf{x}) = \frac{1}{B(\alpha + s, \beta + f)} \theta^{\alpha+s-1} (1 - \theta)^{\beta+f-1}$$

Bayes Theorem on proportional form

- ▶ There is no need to calculate the denominator in Bayes Theorem, $\int f(\mathbf{x}|\theta)\pi(\theta)d\theta$
- ▶ We recognized the Beta distribution anyway
- ▶ Integral of any density function is 1 \Rightarrow constant c is unique
- ▶ Simple form of Bayes Theorem:

$$\pi(\theta|\mathbf{x}) \propto f(\mathbf{x}|\theta)\pi(\theta)$$

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior}$$

Normal model with normal prior

- ▶ **Model:** $X_1, \dots, X_n \stackrel{iid}{\sim} N(\theta, \sigma^2)$ with σ^2 **known**

- ▶ **Prior:**

$$\theta \sim N(\mu, \tau^2)$$

- ▶ **Posterior:**

$$\begin{aligned} f(\theta|x_1, \dots, x_n) &\propto f(x_1, \dots, x_n|\theta, \sigma^2)f(\theta) \\ &\propto N(\theta|\mu_x, \tau_x^2) \end{aligned}$$

where

$$\frac{1}{\tau_x^2} = \frac{n}{\sigma^2} + \frac{1}{\tau^2}$$

$$\mu_x = w\bar{x} + (1 - w)\mu$$

and

$$w = \frac{\frac{n}{\sigma^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}$$

Classical conjugate families

Model $f(\mathbf{x} \theta)$	Prior $\pi(\theta)$	Posterior $\pi(\theta \mathbf{x})$
Poisson(θ)	Gamma(α, λ)	Gamma($\alpha + n\bar{X}, \lambda + n$)
Binomial(k, θ)	Beta(α, β)	Beta($\alpha + n\bar{X}, \beta + n(k - \bar{X})$)
Normal(θ, σ)	Normal(μ, τ)	Normal $\left(\frac{n\bar{X}/\sigma^2 + \mu/\tau^2}{n/\sigma^2 + 1/\tau^2}, \frac{1}{\sqrt{n/\sigma^2 + 1/\tau^2}}\right)$

Bayesian estimator, confidence interval and hypothesis test

- ▶ Bayesian estimator of θ :

$$\hat{\theta}_B = E(\theta|X) = \int \theta \pi(\theta|X) d\theta$$

Posterior risk for $\hat{\theta}_B$:

$$\rho(\hat{\theta}_B) = \text{Var}(\theta|X)$$

- ▶ $[a, b]$ is a 95 % Bayesian confidence interval if

$$P(\theta \in [a, b]|X) = \int_a^b \pi(\theta|X) d\theta = 0.95$$

- ▶ Bayesian hypothesis test $H_0 : \theta \in \Theta_0$ vs $H_A : \theta \in \Theta_A$:

Reject H_0 if

$$\pi(\Theta_0|X) < \pi(\Theta_A|X)$$

Thank you for your attention!