STAT 576 Bayesian Analysis

Lecture 0: Overview

Chencheng Cai

Washington State University

Course Information

Lecture Tu/Th 12:05 — 1:20 PM @SLOAN 7

Office Hours Tu/Wed 1:30 — 3:30 PM @Neill 405 or by appointment

Contact
- email: chencheng.cai@wsu.edu
- zoom: https://wsu.zoom.us/my/chenchengcai
- phone: (509)-335-3141

Site
- Canvas
- http://math.wsu.edu/faculty/ccai/stat576.html

Course Requirements

- Prerequisites: STAT 536, STAT 556, and R/Python programming.
- Textbook:
  Bayesian Data Analysis, 3rd Edition. Gelman, Carlin, Stern, Dunson, Vehtari and Rubin. 2013.
  Free online access from the book website
  (http://www.stat.columbia.edu/ gelman/book/)
- Recommended reading:
  The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation, 2nd Edition. Robert. 2007.

  Monte Carlo Methods for Scientific Computing. Liu. Springer. 2008.

Assessment

- Homework: 40%
  Around six homework in total.
- Mid-term Exam: 30%
  One closed-book exam.
- Project: 30%
  One data analytic project.

Tenative Schedule

- (1 week): Introduction and review.
- (5 weeks): Foundations of Bayesian inference
- (4 weeks): Bayesian computation.
- (2 weeks): Bayesian regression models.
- (2 weeks): State-space models and sequential Monte Carlo.

Bayesian Analysis — Example

- Image you are given a coin and you flip it 10 times.
- What you read from the outcomes:

$$\mathcal{S} = \text{H, T, H, H, H, H, H, H, H, H}$$

- Do you think the coin is fair?
  - No: **Hypothesis testing** gives a rejection of the null that the coin is fair.
  - Yes: if there is no sign that the coin is defective, 9H/1T case just happens **by chance**.
- Will you change your mind if the following scenarios are given:
  - It is a **standard quarter coin** manufactured by U.S. Mint.
  - It is a coin you picked up from a **casino**.
  - It is a coin that your **magician** friend gave you.

Frequentist Inference

- Flipping a coin is a **Bernoulli** event with success (head) probability $\theta \in [0, 1]$.
- θ is **unknown** and **fixed**.
- The conditional probability of observing the H/T sequence:

$$\mathbb{P}[\mathcal{S} \mid \theta] = 10 \times \theta^9(1-\theta)$$

- The above can be viewed as a function of θ given $\mathcal{S}$. The function is called the **likelihood** function:

$$L(\theta; \mathcal{S}) := \mathbb{P}[\mathcal{S} \mid \theta] = 10 \times \theta^9(1-\theta)$$

- Maximize the likelihood function to get **Maximum Likelihood Estimator**:

$$\hat{\theta} = \arg\max_{\theta \in [0,1]} \; L(\theta; \mathcal{S}) = 0.9$$

- We can also construct a **confidence interval** (l, r) such that:

$$\mathbb{P}[\theta \in (l, r)] \geq 1 - \alpha$$

Bayesian Inference

- Flipping a coin is a **Bernoulli** event with success (head) probability $\theta \in [0, 1]$.
- θ is **unknown** and **random** by following a **prior** distribution $\theta \sim \pi(\cdot)$
- The probability of observing $\mathcal{S}$ is the **sampling** distribution:

$$\mathbb{P}[\mathcal{S} \mid \theta] = 10 \times \theta^9(1 - \theta)$$

- Use **Bayes' rule** to get the **posterior** distribution for θ:

$$\pi(\theta \mid \mathcal{S}) = \frac{\mathbb{P}[\mathcal{S} \mid \theta]\pi(\theta)}{\int_\theta \mathbb{P}[\mathcal{S} \mid \theta]\pi(\theta)d\theta}$$

- We may use the **maximum a posteriori (MAP)** estimator:

$$\hat{\theta} = \arg\max_\theta \ \pi(\theta \mid \mathcal{S})$$

- A **credible interval** can be constructed as $[l, u]$ such that

$$\int_l^u \pi(\theta \mid \mathcal{S}) \ d\theta \geq 1 - \alpha$$

Frequentist v.s. Bayesian

	Frequentist	Bayesian
Parameter	fixed	random
inference based on	likelihood	posterior
point estimator	MLE	MAP
interval estimation	confidence interval	credible interval

- The likelihood function is **NOT** a distribution of θ, while the posterior **is**.
- Interpretation of confidence interval: if you repeat the experiment (**same** θ) many times, in at least $1-\alpha$ of the experiments, the confidence interval **covers** the true parameter θ.
- Interpretation of credible interval: if you repeat the experiment (**random** θ) many times, in at least $1-\alpha$ of the experiments, the true parameter θ is **in** the credible interval. **The statement can be made when conditioned on the observations.**

Frequentist v.s. Bayesian

Bayesian

Thomas Bayes (1702–1761)

Pierre-Simon Laplace (1749–1827)

Frequentist

Ronald Fisher (1890–1962)

Jerzy Neyman (1894–1981)

Egon Pearson (1895–1980)

Frequentist v.s. Bayesian

- Bayesian inference was developed much earlier than the frequentist.
- But it was neglected for centuries because of the difficulties in estimation.
- Modern statistics are mostly built on frequentist inference.
  - Likelihood, confidence interval, MLE, hypothesis testing, etc..
- Bayesian inference regains popularity around 1980s because of the developments and advances in Markov Chain Monte Carlo (MCMC).
  - Bayesian inference through Monte Carlo sampling instead of theoretical calculation.
- The third paradigm: **fiducial inference**
  - Fisher developed fiducial inference as a compromise of frequentist and Bayesian.
  - Fisher's try was not successful.
  - David Cox (1924–2022) developed the confidence distribution (CD).
  - Inference based on the confidence distributions is a new area of research.

Bayesian, **F**requentist, **F**iducial

Why Bayesian Statistics?

- Bayesians argue that it is the only correct form of inference.
- It allows a combination of prior knowledge with observations.
- Can solve problems with limited sample size (small sample problem, high-dimensional inference, etc..)
- Consistent with frequentist statistics under certain settings.
- Bayesian inference is decision-theoretical optimal.

Topics in Bayesian Statistics

- Prior elicitation.
  - Subjective Bayes
  - Objective Bayes
- Estimation from the posterior and prediction.
- Decision-theoretical properties.
- Large-sample properties.
- Bayesian hypothesis testing.
- Hierarchical models, sequential models...
- Bayesian computation.
  - Direct sampling from posterior.
  - Expectation-Maximization algorithm.
  - Markov Chain Monte Carlo.
  - Approximate Bayesian Computation.