

Lecture 1 Basic Concepts of Probability and Inference

Fei Tan

Department of Economics
Chaifetz School of Business
Saint Louis University

E6930 Applied Bayesian Statistics

December 18, 2022

What Is the Course About?

- ▶ Introduce Bayesian inferential methods & develop hands-on skills for Python data science
- ▶ **Why Bayesian paradigm?** Handle sophisticated models & uncertainty in decision making
- ▶ Main references
 - ▶ Required: Greenberg (2008), *“Introduction to Bayesian Econometrics”*
 - ▶ Optional: Geweke (2005), *“Contemporary Bayesian Econometrics and Statistics”*
 - ▶ Optional: Herbst & Schorfheide (2015), *“Bayesian Estimation of DSGE Models”*
- ▶ Homework production
 - ▶ L^AT_EX typesetting: www.overleaf.com
 - ▶ Python programming: www.anaconda.com

The Road Ahead...

- ▶ Frequentist v.s. Bayesian views of probability
- ▶ Prior, likelihood, and posterior
- ▶ Coin-tossing example

Frequentist v.s. Bayesian

Probability axioms

1. $0 \leq \mathbb{P}(A) \leq 1$ for any event A
2. $\mathbb{P}(A) = 1$ if event A represents logical truth
3. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ for disjoint events A and B
4. $\mathbb{P}(A|B) = \mathbb{P}(A \cap B)/\mathbb{P}(B)$ (conditional probability)

- ▶ Satisfied by any assignment of probabilities
 - ▶ Frequentists assign probabilities to events describing outcome of *repeated* experiment
 - ▶ Bayesians assign 'subjective' probabilities to uncertain events [de Finetti's (1990) coherency principle]
- ▶ How likely it rains tomorrow?

Prior, Likelihood, and Posterior

Bayes Theorem

$$\pi(\theta|y) = \frac{f(y|\theta)\pi(\theta)}{m(y)} \propto f(y|\theta)\pi(\theta)$$

- ▶ Learning of random vector $Y = [Y_1, \dots, Y_n]'$
 - ▶ call their realizations $y = [y_1, \dots, y_n]'$ *data*
 - ▶ parametric distribution \mathbb{P}_θ for Y
 - ▶ learning of unknown parameter θ
- ▶ Bayesian approach treats θ as being random
 - ▶ start with *prior* density $\pi(\theta)$
 - ▶ update by *likelihood* function $f(y|\theta)$
 - ▶ *posterior* density $\pi(\theta|y)$ proportional to prior \times likelihood
 - ▶ *marginal likelihood* $m(y) = \int f(y|\theta)\pi(\theta)d\theta$

Coin-Tossing Example

- ▶ Likelihood function

- ▶ single toss: $y_i = 1$ if head or 0 if tail

$$y_i \sim \text{Bernoulli}(\theta) \quad \Rightarrow \quad f(y_i|\theta) = \theta^{y_i}(1 - \theta)^{1-y_i}$$

- ▶ n independent tosses

$$f(y_1, \dots, y_n|\theta) = \theta^{\sum y_i} (1 - \theta)^{n - \sum y_i}$$

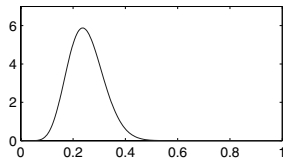
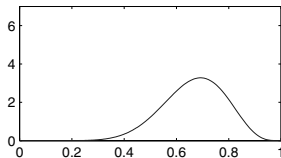
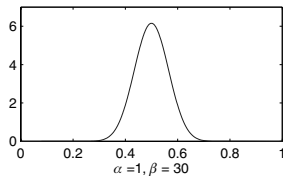
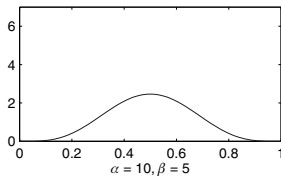
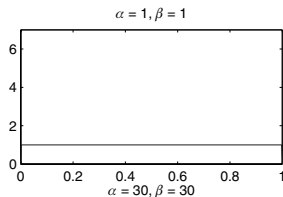
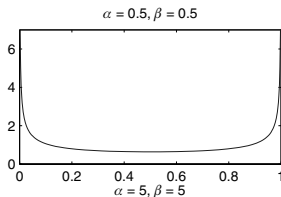
- ▶ (Conjugate) prior density: $\theta \sim \text{Beta}(\alpha, \beta)$

$$\pi(\theta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1 - \theta)^{\beta-1}, \quad 0 \leq \theta \leq 1, \quad \alpha, \beta > 0$$

- ▶ Posterior density: $\theta \sim \text{Beta}(\alpha + \sum y_i, \beta + n - \sum y_i)$

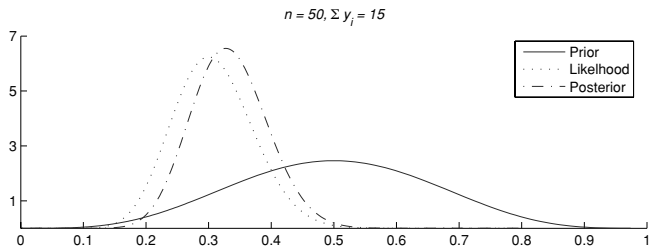
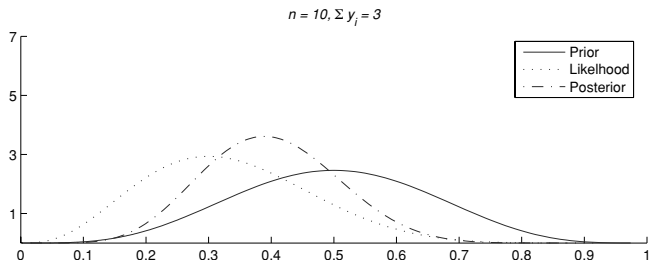
$$\pi(\theta|y) \propto \theta^{\alpha + \sum y_i - 1} (1 - \theta)^{\beta + n - \sum y_i - 1}$$

Hyperparameters



► Shape of Beta: $\mathbb{E}(\theta) = \frac{\alpha}{\alpha+\beta}$, $\mathbb{V}(\theta) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

Sample Size



► $\mathbb{E}(\theta|y) = \frac{\alpha+\beta}{\alpha+\beta+n}\mathbb{E}(\theta) + \frac{n}{\alpha+\beta+n}\bar{y} \rightarrow_{n \rightarrow \infty} \bar{y} \text{ (MLE)}$

Readings

- ▶ de Finetti (1990), "*Theory of Probability*"
- ▶ Hacking (2001), "*An Introduction to Probability and Inductive Logic*"
- ▶ Howson & Urbach (1993), "*Scientific Reasoning: The Bayesian Approach*"
- ▶ Zellner (1971), "*An Introduction to Bayesian Inference in Econometrics*"