Bayesian Models

MIT 18.650

Dr. Kempthorne

Spring 2016

Outline

Bayesian Statistical Models

Statistical Model (as before)

- A random variable X

  $\mathcal{X}$: Sample Space $= \{$outcomes $x\}$

  $\mathcal{F}_X$: sigma-field of measurable events

  $P(\cdot)$ probability distribution defined on $(\mathcal{X}, \mathcal{F}_X)$
- Statistical Model

  $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$

  Parameter θ identifies/specifies distribution in $\mathcal{P}$.

Bayesian Principle

- Assume that the true value of the parameter θ is the realization of a random variable:

  $\theta \sim \pi(\cdot)$, where $\pi(\cdot)$ is a distribution on (Θ, σ_Θ).
- The distribution $(\Theta, \sigma_\Theta, \pi)$ is the **Prior Distribution** for θ.
- The specification of $\pi(\cdot)$ may be

  purely subjective (personalistic)

  based on actual data (empirical Bayes)

Bayesian Statistical Models

Bayesian Framework

- Prior distribution for θ with density/pmf function
  $$\pi(\theta), \quad \theta \in \Theta$$
- Conditional distributions for X given θ,P_θ, with density/pmf function
  $$p(x \mid \theta), \quad x \in \mathcal{X}$$
- Joint distribution for (θ, X) with joint density/pmf function
  $$f(\theta, x) = \pi(\theta)p(x \mid \theta)$$
- **Posterior distribution** for θ given $X = x$ with density/pmf function
  $$\pi(\theta \mid x) = \frac{\pi(\theta)p(x|\theta)}{\sum_t \pi(t)p(x|t)} \quad \text{(discrete prior)}$$
  $$\pi(\theta \mid x) = \frac{\pi(\theta)p(x|\theta)}{\int_\Theta \pi(t)p(x|t)dt} \quad \text{(continuous prior)}$$

Outline

1. Bayesian Models
   - Bayesian Framework
   - Examples

Bayesian Model for Sampling Inspection

Example 1.1.1 Sampling Inspection

- Shipment of manufactured items inspected for defects
- $N =$ Total number of items
- $N\theta =$ Number of defective items
- Sample $n < N$ items without replacement and inspect for defects
- $X =$ Number of defective items in the sample

Bayesian Model for Sampling Inspection

Probability Model for X (Number of defectives in sample)

- Sample Space: $\mathcal{X} = \{x\} = \{0, 1, \dots, n\}$.
- Parameter θ: proportion of defective items in shipment
  $$\Theta = \{\theta\} = \{0, \frac{1}{N}, \frac{2}{N}, \dots, \frac{N}{N}\}.$$
- Probability distribution of X
  $$P(X = k) = \frac{\binom{N\theta}{k}\binom{N - N\theta}{n - k}}{\binom{N}{n}}$$
- Range of X depends on θ, n, and N
  $$k \leq n \text{ and } k \leq N\theta$$
  $$(n - k) \leq n \text{ and } (n - k) \leq N(1 - \theta)$$
  $$\implies max(0, n - N(1 - \theta)) \leq k \leq min(n, N\theta).$$
- $X \sim Hypergeometric(N\theta, N, n)$.

Bayesian Model for Sampling Inspection: Prior Distribution

Case 1: Empirical Specification of π

- Data on past shipments provides a frequency distribution for proportion of defectives:
$$P(\theta = \frac{i}{N}) = \pi_i, \ i = 0, 1, 2, \ldots, N$$
- Before inspecting the current shipment, assume that the proportion of defectives in the current shipment is a realization from this distribution.

Case 2: Parametric model

- Conclude from past experience that each item in a shipment is defective with probability 0.20, independently of each other.
- For a shipment of size $N = 100$ the prior distribution is
$$\pi_i = \binom{100}{i} (0.2)^i (0.8)^{100-i},$$
$i = 0, 1, \ldots, 100$

Bayesian Model for Sampling Inspection: Posterior

- The **Joint Distribution** of (θ, X) has probability mass function:

$$P(\theta = \tfrac{i}{N}, X = x) = \pi(\theta = \tfrac{i}{N}) p(X = x \mid \theta = \tfrac{i}{N})$$

$$= \pi_i \cdot \frac{\binom{i}{x}\binom{N-i}{n-x}}{\binom{N}{n}}$$

- The **Posterior distribution** is the conditional distribution with $\pi(\theta \mid X = x) = \pi(\theta)p(x \mid \theta) / \sum_{t \in \Theta} \pi(t)p(x \mid t)$

Bayesian Model for Bernoulli Trials

Example 1.2.1 Bernoulli Trials

- $X_1, X_2, \ldots, X_n$ are i.i.d. $Bernoulli(\theta)$ r.v.s
- $\mathcal{X} = \{Success(1), Failure(0)\}$

$$P(X_i = 1 \mid \theta) = \theta$$
$$P(X_i = 0 \mid \theta) = 1 - \theta$$

- Parameter Space: $\Theta = \{\theta : 0 \le \theta \le 1\}$
- Prior Distribution for θ: density $\pi(\theta)$
- Posterior Distribution for θ :

$$\pi(\theta \mid x_1, \ldots, x_n) = \frac{\pi(\theta)\theta^k(1-\theta)^{n-k}}{\int_0^1 \pi(t)t^k(1-t)^{n-k}dt}$$
$$0 < \theta < 1,$$
$$x_i = 0 \text{ or } 1,\ i = 1, \ldots, n$$
$$k = \sum_{i=1}^n x_i.$$

Bayesian Model for Bernoulli Trials

Note

- Posterior distribution depends on $X = (X_1, \ldots, X_n)$ through
  $$T(X) = \quad {}_{1}^{n} X_i.$$
- Given θ, $T(X) \sim Binomial(n, \theta)$.
- Consider the posterior distribution if we only observe $T(X)$.
  By Exercise 1.2.9 the same distribution obtains.

Conjugate Prior Distribution (Prior and Posterior in same family)

- A priori, assume $\theta \sim Beta(r, s)$ distribution, with density
  $$\pi(\theta) = \frac{\theta^{r-1}(1-\theta)^{s-1}}{\beta(r,s)}, \quad 0 < \theta < 1$$
  $$\text{where } \beta(r,s) = \int_0^1 \theta^{r-1}(1-\theta)^{s-1} d\theta = \Gamma(r)\Gamma(s)/\Gamma(r+s)$$
- A priori, $E[\theta] = r/(r+s)$ and $Var(\theta) = rs/[(r+s)^2(r+s+1)]$
- A posteriori, $\pi(\theta \mid T(X) = k) \sim Beta(r+k, s+(n-k))$

Alternate Sampling Models for Bernoulli Trials

Bernoulli Trials: $X_1, X_2, \ldots$ i.i.d. *Bernoulli*(θ) r.v.s
Suppose $(X_1, X_2, X_3, X_4, X_5) = (0, 1, 0, 1, 0)$.

Possible sample models for the data:

- Sample $n = 5$ trials (regardless of the outcomes).
  $$Y = X_1 + X_2 + \cdots + X_5 \text{ is } \textit{Binomial}(n = 5, p = \theta).$$
- Sample trials until three failures are realized.
  $$S = \text{Number of successes before } r(= 2) \text{ failures}$$
  $$S \sim \text{Negative Binomial Distribution}$$
  $$p(S = s \mid \theta) = \binom{r + s - 1}{s} (1 - \theta)^r \theta^s,$$
  $$s = 0, 1, 2, \ldots$$
- ... (sampling protocols applying an operational stopping rule)

Significant Property:

Bayes posterior distributions are all the same!

Problems

Problem 1.2.1 Merging opinions. Two-model case of Bernoulli Trials. Convergence of posterior distributions.

Problem 1.2.2 Half-triangular distributions. Mean of posterior distributions for alternate prior distributions. Non-informative prior distributions.

Problem 1.2.3 Geometric distribution (number of Bernoulli trials until first success). Solving for the posterior distribution under alternative prior distributions, including conjugate prior.

Problem 1.2.6 Conjugate priors for Poisson distribution (Gamma distributions).

Problems (continued)

Problem 1.2.9 Bayesian model using summary statistic from Bernoulli trials.

Problem 1.2.12 Bayesian model of a Gaussian distribution with known mean and unknown variance. Inverse chi-squared distributions.

Problem 1.2.13 Computation of posterior distribution using an improper prior.

Problem 1.2.15 Conjugate prior for multinomial distributions (Dirichlet distributions).