STAT3602 Statistical Inference

Example Class 3

LIU Chen

Department of Statistics and Actuarial Science, HKU

14/09/2020

Outline

Bayesian Approach: Prior to posterior

Bayesian approach

- X: observable random variate with probability function $f(x \mid \theta)$
- θ: unobservable random variate with a specified prior probability function $\pi(\theta)$:

$$\int_{\Theta} \pi(\theta) d\theta = 1 \quad (\text{ continuous } \theta), \qquad \text{or}$$

$$\sum_{\theta \in \Theta} \pi(\theta) = 1 \quad (\text{ discrete } \theta)$$

Bayesian Approach: Prior to posterior

Posterior Probability Function

The posterior probability function of θ given the observed data $\boldsymbol{x}$ is defined to be the conditional probability function of θ given $\boldsymbol{X} = \boldsymbol{x}$, that is

$$\begin{aligned}\pi(\theta \mid x) &= \frac{f(\boldsymbol{x} \mid \theta)\pi(\theta)}{\int_\Theta f\left(x \mid \theta'\right)\pi\left(\theta'\right) d\theta'} \\ &\propto f(\boldsymbol{x} \mid \theta)\pi(\theta)\end{aligned}$$

Bayesian Approach: **Expected Posterior Loss**

Let the prior $\pi(\theta)$ be given for $\theta \in \Theta$. Consider a decision problem with loss function $L(\theta, a)$ for $\theta \in \Theta$ and action $a \in \mathcal{A}$ (action space). Definition. The expected posterior loss given data $\boldsymbol{x}$, incurred by taking action a, is

$$\mathbb{E}[L(\theta, a) \mid \boldsymbol{x}] = \int_{\Theta} L(\theta, a)\pi(\theta \mid \boldsymbol{x})d\theta$$

Definition.

A Bayesian decision is to take an action $a \in \mathcal{A}$ which minimises the expected posterior loss $\mathbb{E}[L(\theta, a) \mid \boldsymbol{x}]$

Writing $f(\boldsymbol{x}) = \int_\Theta \pi(\theta') f(x \mid \theta') d\theta'$, we have

$$\begin{aligned}
\mathbb{E}[L(\theta, a) \mid \boldsymbol{x}] &= \int_\Theta L(\theta, a) \frac{f(\boldsymbol{x} \mid \theta)\pi(\theta)}{f(\boldsymbol{x})} d\theta \\
&= \frac{1}{f(\boldsymbol{x})} \int_\Theta L(\theta, a) f(\boldsymbol{x} \mid \theta)\pi(\theta) d\theta
\end{aligned}$$

Thus, minimising $\mathbb{E}[L(\theta, a) \mid \boldsymbol{x}]$ w.r.t. $a \in \mathcal{A}$ is equivalent to minimising $\int_\Theta L(\theta, a) f(\boldsymbol{x} \mid \theta)\pi(\theta) d\theta$ w.r.t. $a \in \mathcal{A}$

Outline

Exercise 1

We are faced with a shipment of N manufactured items. An unknown number D of these items are defective. A sample of n is drawn without replacement and inspected. The number X of defective items is recorded

1. What is the distribution of X?
2. Now suppose D has a binomial prior distribution with parameters N and p; that is

$$P(D = d) = \binom{N}{d} p^d (1-p)^{N-d}, d = 0, 1, \ldots, N \tag{1}$$

Show the posterior distribution of D given $X = x$ is that of $x + Y$ where Y has a binomial distribution with parameters $N - n$ and p.

Exercise 1: Solution I

The distribution of X given $D = d$ is the hypergeometric distribution, therefore,

$$f(x \mid d) = \frac{\begin{pmatrix} d \\ x \end{pmatrix} \begin{pmatrix} N-d \\ n-x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}, \quad 0 \leq x \leq n \wedge d, n-x \leq N-d, \tag{2}$$

and the prior pmf for D is

$$\pi(d) = \begin{pmatrix} N \\ d \end{pmatrix} p^d (1-p)^{N-d}, \quad 0 \leq d \leq N. \tag{3}$$

So, the posterior pmf of D is given by

Exercise 1: Solution II

$$
\begin{aligned}
\pi(d \mid x) &\propto \binom{d}{x}\binom{N-d}{n-x}\binom{N}{d} p^d(1-p)^{N-d}, && 0 \le x \le d, d \le N-(n-x) \\
&\propto \frac{d!}{(d-x)!}\frac{(N-d)!}{(N-d-n+x)!}\frac{1}{d!(N-d)!} p^d(1-p)^{N-d}, && 0 \le x \le d \le N-n+x \\
&= \frac{1}{(d-x)!(N-d-n+x)!} p^d(1-p)^{N-d}, && 0 \le x \le d \le N-n+x \\
&\propto \frac{(N-n)!}{(d-x)!(N-n-d+x)!} p^{d-x}(1-p)^{N-n-d+x}, && 0 \le x \le d \le N-n+x
\end{aligned}
$$

Therefore, given $X = x$, $Y = D - x$ has a binomial distribution with parameters $N - n$ and p.

$$
\pi(d \mid x) = \binom{N-n}{d-x} p^{d-x}(1-p)^{N-n-(d-x)}, \quad d = x, x+1, \dots, N-n+x. \tag{4}
$$

Outline

Exercise 2

Suppose $X_1, ..., X_n$ have a Poisson distribution with mean $\theta(> 0)$, and the prior distribution of is $\Gamma(\alpha, \beta)$, where $\alpha, \beta > 0$, i.e.

$$\pi(\theta) = \frac{1}{\Gamma(\alpha)} \beta^\alpha \theta^{\alpha-1} e^{-\beta\theta}, \beta, \alpha > 0 \tag{5}$$

1. Determine the posterior distribution of θ given the random sample $x_1, ..., x_n$.
2. If we regard the problem of estimating θ based on a size-n sample of X as a statistical decision problem and we adopt the square loss function, then what is the Bayes rule for estimating θ?

Exercise 2: Solution I

The pmf of $X_1, ..., X_n$ given θ is

$$f\left(x_1, \ldots, x_n \mid \theta\right)=\prod_{i=1}^n \frac{\theta^{x_i} e^{-\theta}}{x_i!}=\frac{\theta^{\sum x_i} e^{-n\theta}}{\prod x_i!}, x_i=0,1, \ldots \tag{6}$$

and the prior density function

$$\pi(\theta)=\frac{1}{\Gamma(\alpha)} \beta^\alpha \theta^{\alpha-1} e^{-\beta\theta}, \quad \alpha, \beta>0 \tag{7}$$

So the posterior density of is

$$\begin{aligned}
\pi(\theta \mid \vec{x}) & \propto \theta^{\sum x_i} e^{-n\theta} \times \theta^{\alpha-1} e^{-\beta\theta} && (8)\\
& \propto \theta^{\sum x_i+\alpha-1} e^{-(n+\beta)\theta}, \quad \vec{x}=x_1, x_2, \ldots, x_n && (9)
\end{aligned}$$

Therefore, it is easy to determine $\theta \mid \vec{x} \sim \operatorname{Gamma}\left(\sum_{i=1}^n x_i+\alpha, n+\beta\right)$.

Exercise 2: Solution II

If the square loss function is used, then the Bayes estimator is simply the posterior mean:

$$E[\theta \mid \vec{x}] = \frac{n\bar{x} + \alpha}{n + \beta}$$

Outline

Exercise 3

Suppose X has a Bernoulli distribution with parameter θ where θ has prior uniform distribution on $[0, 1]$. A random sample of size n is taken.

1. What is the posterior distribution of θ?
2. Find the Bayes estimate of θ when the loss function is $(\theta - \hat{\theta})^2$

Exercise 3: Solution I

1. The pmf of X given θ is

$$f(x \mid \theta) = \theta^x (1-\theta)^{1-x}, \quad x = 0, 1 \tag{10}$$

and the prior density function

$$\pi(\theta) = 1\{0 < \theta < 1\} \tag{11}$$

So the posterior density of θ given $x_1, ..., x_n$ is

$$\pi(\theta \mid \vec{x}) \propto \theta^{\sum x_i} (1-\theta)^{n - \sum x_i}, \quad \theta \in (0, 1) \tag{12}$$

Therefore, it is easy to determine

$$\theta \mid \vec{x} \sim Beta\left(\sum_{i=1}^{n} x_i + 1, n - \sum_{i=1}^{n} x_i + 1\right) \tag{13}$$

2. When the loss function is $\left(\theta - \hat{\theta}\right)^2$, the Bayes estimate of θ is the posterior mean. Therefore, the Bayes estimate is $\frac{n\bar{x}+1}{n+2}$

Outline

Exercise 4

Let $X_1, ..., X_n$ iid samples from $N(\theta, 1)$, suppose the prior distribution of θ is a standard normal distribution. Find the equal-tailed interval for θ of fixed posterior coverage probability 0.95.

Exercise 4: Solution I

In general, when $X_1, ..., X_n$ iid samples from $N(\theta, \sigma^2)$, the pdf of $X_1, ..., X_n$ given θ is

$$\begin{aligned} f(\vec{x} \mid \theta) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x_i-\theta)^2}{2\sigma^2}\right] \\ &= \frac{1}{\left(\sqrt{2\pi\sigma^2}\right)^n} \exp\left[-\frac{\sum (x_i-\theta)^2}{2\sigma^2}\right] \end{aligned}$$

and the prior distribution of θ is a normal distribution with mean μ and variance γ^2, its density function is

$$\pi(\theta) = \frac{1}{\sqrt{2\pi\gamma^2}} \exp\left[-\frac{(\theta-\mu)^2}{2\gamma^2}\right] \tag{14}$$

Exercise 4: Solution II

Then the posterior density of θ is

$$\begin{aligned}
\pi(\theta \mid \vec{x}) &\propto \exp\left\{-\frac{\sum(x_i-\theta)^2}{2\sigma^2}\right\} \times \exp\left\{-\frac{(\theta-\mu)^2}{2\gamma^2}\right\} \\
&= \exp\left\{-\left[\frac{\sum(x_i-\theta)^2}{2\sigma^2} + \frac{(\theta-\mu)^2}{2\gamma^2}\right]\right\} \\
&= \exp\left\{-\left[\frac{\sum x_i^2 - 2\theta\sum x_i + n\theta^2}{2\sigma^2} + \frac{\theta^2 - 2\mu\theta + \mu^2}{2\gamma^2}\right]\right\} \\
&\propto \exp\left\{-\frac{-2\theta\sum x_i\gamma^2 + n\theta^2\gamma^2 + \theta^2\sigma^2 - 2\mu\theta\sigma^2}{2\sigma^2\gamma^2}\right\}
\end{aligned}$$

$$= \exp\left\{-\frac{\left(n\gamma^2+\sigma^2\right)\theta^2 - 2\left(\sum_i^2 x_i\gamma^2+\mu\sigma^2\right)\theta}{2\sigma^2\gamma^2}\right\}$$

$$= \exp\left\{-\frac{\theta^2 - 2\frac{\sum x_i\gamma^2+\mu\sigma^2}{n\gamma^2+\sigma^2}\theta}{2\frac{\sigma^2\gamma^2}{n\gamma^2+\sigma^2}}\right\}$$

$$\propto \exp\left\{-\frac{(\theta-\alpha)^2}{2\beta^2}\right\}$$

Where, $\alpha = \frac{\sum x_i\gamma^2+\mu\sigma^2}{n\gamma^2+\sigma^2}$ and $\beta^2 = \frac{\sigma^2\gamma^2}{n\gamma^2+\sigma^2}$.
Hence, it is easy to determine $\theta \mid \vec{x} \sim N\left(\alpha, \beta^2\right)$.
In our case, $\sigma^2 = 1$, $\mu = 0$ and $\gamma^2 = 1$. Therefore, the posterior distribution

$$\theta \mid \vec{x} \sim N\left(\frac{\sum x_i}{n+1}, \frac{1}{n+1}\right)$$

Exercise 4: Solution IV

The equal-tailed interval for θ of fixed posterior coverage probability 0.95 is

$$\left[\frac{\sum_{i=1}^{n} x_i}{n+1} - \frac{1.96}{\sqrt{n+1}}, \frac{\sum_{i=1}^{n} x_i}{n+1} + \frac{1.96}{\sqrt{n+1}}\right] \tag{15}$$