STAT3602 Statistical Inference

Example Class 5

LIU Chen

Department of Statistics and Actuarial Science, HKU

07/10/2020

Outline

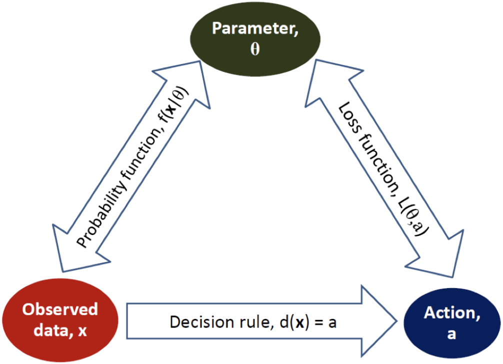


Figure: Frequentist Approach: Formulation

Frequentist Approach: Formulation

- Formulation and Risk Function
- Criteria for a good decision rule
  - Admissibility
  - Minimaxity
  - Unbiased rule
- Randomised decision rules

Bayesian Decision

- Prior and Posterior Probability Function.
- Expected Posterior Loss
- Bayesian Decision

Bayesian Statistical Inference

- Point Estimation
- Hypothesis Testing about θ
- Interval Estimation of θ:
  - Given fixed interval length, maximise posterior coverage probability.
  - Given fixed posterior coverage probability, minimise interval length.
  - Given fixed posterior coverage probability, require "equal-tailed" inteval.
- Predictive Distribution

Outline

Papet, an old greedy man living in village A, wants to purchase a plot of land in his village. The land is currently owned by Jean, a hunchback recently arriving from Village B, who inherits the land from his deceased mother Florette, Papet's ex-lover.

The price of the land depends on the availability of water. Papet knows that there is a water source 1 mile from the land, and is considering to block it.

If Papet does not block the source (action a_0), then Jean will have access to water from this source and will sell his land at a price of 1 unit. If Papet blocks the source (action a_1), then Jean will need to search for an alternative water source, at an unknown distance of θ miles away, and will sell his land at a price of $e^{1-\theta}$ units. The following questions concern

Papet's decision problem of whether to block the water source.

1. Papet's loss function $L(\theta, a)$ can be defined as the price of the land. Write down the expressions for $L(\theta, a_0)$ and $L(\theta, a_1)$, for any $\theta > 0$
2. Papet tries to make his decision based on an observation X, which is the growth rate of a plant found near the land plot. It is known that X is distributed under the density function

$$f(x \mid \theta) = \theta e^{-\theta x}, \quad x > 0.$$

Show that the risk function of Papet's decision rule d is given by

$$R(\theta, d) = \mathbb{P}_\theta(d(X) = a_1)\left(e^{1-\theta} - 1\right) + 1, \quad \theta > 0$$

3. Find a minimax decision rule for Papet.

4. Suppose that Papet has a prior belief about θ, represented by the prior density function

$$\pi(\theta) = ce^{-c\theta}, \quad \theta > 0,$$

for some fixed constant $c > 0$. Show that Papet's Bayes rule is to block the water source if and only if

$$c + X < \frac{1}{\sqrt{e} - 1}$$

[You may find the following identity useful: $\int_0^\infty \theta e^{-k\theta} d\theta = k^{-2}$ **for any** $k > 0$**.]**

Exercise 1: Solution I

1. Write down the expressions for $L(\theta, a_0)$ and $L(\theta, a_1)$, for any $\theta > 0$.

$$L(\theta, a_0) = 1, \quad L(\theta, a_1) = e^{1-\theta}$$

2. The risk function of Papet's decision rule d is

$$\begin{aligned} R(\theta, d) &= \mathbb{E}_\theta[L(\theta, d(X))] \\ &= L(\theta, a_0)\,\mathbb{P}_\theta(d(X) = a_0) + L(\theta, a_1)\,\mathbb{P}_\theta(d(X) = a_1) \\ &= 1 - \mathbb{P}_\theta(d(X) = a_1) + e^{1-\theta}\mathbb{P}_\theta(d(X) = a_1) \\ &= \mathbb{P}_\theta(d(X) = a_1)\left(e^{1-\theta} - 1\right) + 1 \end{aligned}$$

Exercise 1: Solution II

3. Find a minimax decision rule for Papet.
   Denote by d_0 the rule of always choosing action a_0. For any rule d,

$$
\begin{aligned}
\sup_{\theta>0} R(\theta, d) &\geq R(\theta=0, d) \\
&= \mathbb{P}_\theta\left(d(X)=a_1\right)\left(e^{1-0}-1\right)+1 \\
&\geq 1 \\
&= \sup_{\theta>0} R\left(\theta, d_0\right)
\end{aligned}
$$

so that d_0 is minimax.

Exercise 1: Solution III

4. Suppose that Papet has a prior belief about θ, represented by the prior density function $\pi(\theta) = ce^{-c\theta}, \quad \theta > 0$ for some fixed constant $c > 0$. Show that Papet's Bayes rule is to block the water source if and only if $c + X < \frac{1}{\sqrt{e}-1}$

Expected posterior loss:

$$\mathbb{E}[L(\theta, a) \mid X] \propto \int_0^\infty L(\theta, a)\left(\theta e^{-\theta X}\right)\left(ce^{-c\theta}\right) d\theta$$
$$= \begin{cases} c\int_0^\infty \theta e^{-\theta(c+X)} d\theta = c(c+X)^{-2}, & a = a_0 \\ ce\int_0^\infty \theta e^{-\theta(1+c+X)} d\theta = ce(1+c+X)^{-2}, & a = a_1 \end{cases}$$

Bayes rule is to choose a_1 iff

$$c(c+X)^{-2} > ce(1+c+X)^{-2} \Leftrightarrow c + X < \frac{1}{\sqrt{e}-1}$$

Outline

Exercise 2

Each customer buys one of the K products in a shop, buying product i with probability θ_i. Suppose there are n customers. X_i of them buy product i. X $= (X_1, \ldots, X_K)$ follows multinomial distribution with probability mass function

$$f(x \mid \theta) = \frac{n!}{x_1! \ldots x_K!} \theta_1^{x_1} \ldots \theta_K^{x_K}.$$

(Mean for X_i is $n\theta_i$).
Our prior distribution for $\theta = (\theta_1, \ldots, \theta_K)$ is $\theta \sim \text{Dirichlet}\left(\alpha_1, \ldots, \alpha_K\right)$

$$\pi(\theta) = \frac{\Gamma\left(\alpha_1 + \cdots + \alpha_K\right)}{\Gamma\left(\alpha_1\right) \ldots \Gamma\left(\alpha_K\right)} \theta_1^{\alpha_1 - 1} \ldots \theta_K^{\alpha_K - 1}.$$

1. Find the posterior distribution of $\theta \mid x$
2. Suppose there are m customers independent of the first n customers. Y_i of the m customers buy product i. Find the mean of the posterior predictive distribution of $Y_i \mid \mathrm{x}, i = 1, \ldots, K$

Exercise 2: Solution I

1. Please refer to the page 28 – 29 in the note of Chapter 2

$$\begin{aligned}\pi(\theta \mid \mathrm{x}) &\propto \pi(\theta) f(\mathrm{x} \mid \theta) \\ &\propto \theta_1^{\alpha_1-1} \ldots \theta_K^{\alpha_K-1} \theta_1^{x_1} \ldots \theta_K^{x_K} \\ &= \theta_1^{x_1+\alpha_1-1} \ldots \theta_K^{x_K+\alpha_K-1}\end{aligned}$$

which follows $\operatorname{Dir}(x_1+\alpha_1, \ldots, x_K+\alpha_K)$

2. $$\begin{aligned}E(Y_i \mid \mathrm{x}) &= E(E(Y_i \mid \theta) \mid \mathrm{x}) \\ &= E(m\theta_i \mid \mathrm{x}) = \frac{m(x_i+\alpha_i)}{\sum_{j=1}^K (x_j+\alpha_j)}\end{aligned}$$

[It can be derived that the mean of θ_i is $\frac{\alpha_i}{\sum_{j=1}^K \alpha_j}$ for $\theta \sim \text{Dirichlet}(\alpha_1, \ldots, \alpha_K)$]

Outline

Let θ be an unknown parameter in $[0, 1]$, representing the pollution index of a city. In a sample of air collected from the city, the content of a particular pollutant is observed to be $100X\%$, where $X \in [0, 1]$ is a random variable with the density function

$$f(x \mid \theta) = 1 + 2\theta \left(x - \frac{1}{2} \right), \quad x \in [0, 1]$$

Before the measurement of X, non-informative prior distribution is given to θ, so that θ is uniformly distribution over [0,1]

1. Plot the functions $f(x \mid 0)$ and $f(x \mid 1)$ against $x \in [0, 1]$ on the same graph. Comment briefly on the likely magnitude of the observation X if the pollution index θ is high in the city.

Exercise 3 II

2. Show that the posterior distribution of θ given $X = x$ has the density function

$$\pi(\theta \mid x) = \left(x + \frac{1}{2}\right)^{-1} (1 + (2x - 1)\theta), \quad \theta \in [0, 1]$$

   Describe briefly how the observation x affected our posterior belief about θ in the two cases where $x < 1/2$ and $x > 1/2$

3. Based on the observation $X = x$, find the Bayes interval for θ which has length $1/3$ and the maximum posterior coverage probability if
   1. $x < 1/2$
   2. $x > 1/2$

Exercise 3 III

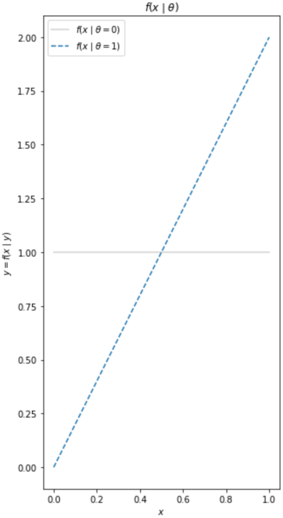
4. If a second sample of air is to be collected from the city independently of the first sample, and its pollutant content observed to be $100Y\%$, show that the posterior predictive distribution of Y given $X = x$ has the density function

$$g^*(y \mid x) = \frac{1}{3}\left(x + \frac{1}{2}\right)^{-1}(x + 1 + (4x + 1)y), \quad y \in [0, 1].$$

5. If the first sample is not polluted, that is $X = 0$, predict the pollutant content Y of the second sample using the mean of the posterior predictive distribution g^* given in (d) above.

Exercise 3: Solution I

1. The density functions are plotted as



Exercise 3: Solution II

2. The posterior density is

$$\pi(\theta \mid x) \propto f(x \mid \theta)\pi(\theta) = 1 + 2\theta \left(x - \frac{1}{2}\right)$$

with normalizing constant

$$\int_0^1 f(x \mid \theta)\pi(\theta) d\theta = x + \frac{1}{2}$$

That is

$$\pi(\theta \mid x) = \left(x + \frac{1}{2}\right)^{-1} \left(1 + 2\theta \left(x - \frac{1}{2}\right)\right)$$

If $x < 1/2, \pi(\theta \mid x)$ is a decreasing function in θ, so we have a stronger belief in small θ. If $x > 1/2$, $\pi(\theta \mid x)$ is an increasing function in θ and we have a stronger belief in large θ.

3. For $x < 1/2, \pi(\theta \mid x)$ is a downward sloping straight line from 0 to 1, so the Bayes interval with length $1/3$ is $[0, 1/3]$; For $x > 1/2 \pi(\theta \mid x)$ is an upward sloping straight line from 0 to 1, so the Bayes interval with length $1/3$ is $[2/3, 1]$.
4. The predictive density is

$$
\begin{aligned}
g^*(y \mid x) &= \int_0^1 f(y \mid \theta)\pi(\theta \mid x) d\theta \\
&= \int_0^1 [1 + (2y-1)\theta](x+1/2)^{-1}[1 + (2x-1)\theta] d\theta \\
&= (x+1/2)^{-1}\left\{1 + 2(x+y-1)\frac{1}{2} + \frac{1}{3}(2y-1)(2x-1)\right\} \\
&= \frac{4xy + x + y + 1}{3(x+1/2)}
\end{aligned}
$$

Exercise 3: Solution IV

5. Given the observed $X = 0$, the mean of Y is

$$\int_0^1 yg^*(y \mid x)dy = \frac{1}{3}(x + 1/2)^{-1}\left[\frac{1}{2}(x+1) + \frac{1}{3}(4x+1)\right], \quad x = 0$$

Final answer is $\frac{5}{9}$.