Homework I Bayesian Methods of Inference, Problem Sheet

Tianyu Zhang[1]

Abstract:

This homework sheet aims at the decision theory part, or, the Bayesian estimator part in both our lecture notes. The problems are labeled by the keywords in characterizing what concepts we are dealing with along with the symbol ⋆ which measures the hardness of the problem, the more ⋆, the harder the problem. For the solutions in finding exact result and the remarks, the end of solution is labeled by "||", for those problems asking us in proving or disproving some statements, the end is labeled by the usual □.

The problem sheet is divided into two sections. In the first section we have three problems treating the concept of loss and risk, where, of course, involving optimizing the decision. In the second section we focus on finding the posterior distribution and the Bayesian estimator. We do not cover a section in solving problems with respect to the interval estimation and hypothesis testing for Bayesian estimators since it involves too many calculations.

Section I: Loss and Risk

Problem 1: ⋆

(**Keywords**: <u>*Loss Function*</u>, <u>*Risk Function*</u>)

A random variable has the uniform density given by

$$f(x) = \begin{cases} \frac{1}{\theta}, & \text{for } 0 < x < \theta \\ 0, & \text{otherwise} \end{cases}$$

and we want to estimate the parameter θ on the basis of a single observation. If the decision function is to be of the form $d(x) = kx$, where $k \geq 1$ and the losses are proportional to the absolute allure of the errors, i.e.

$$L(kx, \theta) = c|kx - \theta|,$$

where c is a positive constant, find the value of k that minimizes the risk.

Problem 2: ⋆

(Keywords: <u>*Loss Function*</u>, <u>*Risk Function*</u>)

Use the minimal criterion to estimate the parameter θ of a binomial distribution on the basis of the random variable X, the observed number of successes in n trials, when the decision function is of the form

$$d(x) = \frac{x+a}{n+b},$$

where a and b are constants, and the loss function is given by

$$L\Big(\frac{x+a}{n+b}, \theta\Big) = c\Big(\frac{x+a}{n+b} - \theta\Big)^2,$$

where c is a positive constant.

---

[1] YMSC, BIMSA, bidenbaka@gmail.com

Problem 3: ⋆ ⋆

(**Keyword**: Bayes Risk, Bayes Rule)

With reference to **Problem 1**, suppose that the parameter of the uniform density is looked upon as a random variable with the probability density given by

$$h(\theta) = \begin{cases} \theta \cdot e^{-\theta}, & \text{if } \theta > 0 \\ 0, & \text{otherwise} \end{cases}.$$

If there is no restriction on the form of the decision function and the loss function is quadratic, i.e.

$$L\big(d(x), \theta\big) = c\big(d(x) - \theta\big)^2.$$

Find the decision function that minimizes the Bayes risk.

Section II: Prior and Posterior of Bayesian Estimators

The following three problems are treatment with respect to finding the posterior distribution of the Bayesian statistics. We introduce again the algorithm in finding the Bayesian posterior.

Algorithm 1: Finding Poseterior

Given $Y_1, Y_2, \cdots, Y_n$ random variables wih likelihood function $L(\theta \,|\, y_1, y_2, \cdots, y_n)$ and θ has density $g(\theta)$, then

Step I: Joint Density

$$f(y_1, y_2, \cdots, y_n, \theta) = L(\theta \,|\, y_1, y_2, \cdots, y_n) \cdot g(\theta).$$

Step II: Marginal Density

$$m(y_1, y_2, \cdots, y_n) = \int_{-\infty}^{\infty} L(y_1, y_2, \cdots, y_n \,|\, \theta) \cdot g(\theta) d\theta.$$

Step III: Posterior Density

$$g^*(\theta \,|\, y_1, y_2, \cdots, y_n) = \frac{L(\theta \,|\, y_1, y_2, \cdots, y_n) \cdot g(\theta)}{\int_{-\infty}^{\infty} L(\theta \,|\, y_1, y_2, \cdots, y_n) \cdot g(\theta) d\theta}.$$ ||

Problem 4 is solved by applying this algorithm directly.

Problem 4: ⋆⋆

(**Keywords**: Prior, Bernoulli, Beta, Posterior)

Let $Y_1, Y_2, \cdots, Y_n$ denote a random sample from a Bernoulli distribution where $\mathbb{P}(Y_i = 1) = p$ and $\mathbb{P}(Y_i = 0) = 1 - p$ and assume that the prior distribution for p is $\text{Beta}(\alpha, \beta)$. Find the posterior distribution for p.

In practice, having the prior density and the distribution of the data does not secure us the posterior distribution, we need further, as (2.6) in Notes 2 suggests, find the marginal distribution of the data, this could be cumbersome in some cases, especially when we are not dealing with the familiar distribution families.

In this problem sheet, we in saying a Bayesian estimator we use the following convention.

Definition: Bayesian Estimator

Let $Y_1, Y_2, \cdots, Y_n$ be a random sample with likelihood function $L(\theta \,|\, y_1, y_2, \cdots, y_n)$ and let θ have prior density $g(\theta)$. The posterior Bayes estimator for $t(\theta)$ is given by $\widehat{t(\theta)} = \mathbb{E}\left(t(\theta) \middle| Y_1, Y_2, \cdots, Y_n\right)$.

Problem 5: ★★

(**Keyword**: <u>*Prior*</u>, <u>*Gamma*</u>, <u>*Posterior*</u>, <u>*Bayesian Estimator*</u>)

In **Problem 4**, we found the posterior distribution of the Bernoulli parameter p based on a beta prior distribution with parameters α and β. In this problem we want to find the Bayes estimators for p and $p(1-p)$.

The last problem treats the problem in estimating the parameter with unknown mean but known variance. Before we proceed to the last problem in this section, let us recall the Bayesian conjugate.

Definition: Conjugate

A family of prior distributions π is conjugate to the model $f(x \mid \theta)$ if the posterior distribution belongs to the same family.

Also it is helpful ro recall the definition of conjugate family.

Definition: Conjugate Family

Let $\mathscr{F}$ denote the class of pdfs or pmfs $f(x \mid \theta)$ indexed by θ. A class Π of prior distributions is a conjugate family for $\mathscr{F}$ if the posterior distribution is in the class Π $\forall f \in \mathscr{F}$, all priors in Π, and all $x \in X$.

Problem 6: ★★★★

(**Keywords**: <u>*Prior*</u>, <u>*Posterior*</u>, <u>*Conjugate Family*</u>, <u>*Bayesian Estimator*</u>)

Let $Y_1, Y_2, \cdots, Y_n$ denote a random sample from a normal population with unknown mean μ and known variance σ^2. The conjugate prior distribution for μ is a normal distribution with known mean η and knwon variance δ^2. Find the posterior distribution and the Bayes estimator for μ.

In this section we have introduced the concept of Bayesian inference. This approach involves both the data distribution and the experimenter's belief. The calculation could be very cumbersome, as we mentioned in our lecture notes 2, the invovment of the convolution may ease the calculation. However, such a treatment contains advanced materials in both mathematics and statistics, interested readers may find the references offered in the second lecture notes, pp. 21 valable.