Available online at www.elixirpublishers.com (Elixir International Journal)

Statistics

Elixir Statistics 123 (2018) 51951-51955

ISSN: 2229-712X

Loss and Risk in Bayesian inference and its Application

Abu Elgasim Abbas A bow Mohammed

[1]Quant method unit, Faculty of Business and Economics, Qassim University, Buriadah-Kingdom of Saudi Arabic.

[2]Department of statistics and econometrics, Faculty of Economics and Political Science, Omdurman Islamic University, Khartoum, Sudan.

ARTICLE INFO

Article history:

Received: 10 September 2018;
Received in revised form:
3 October 2018;
Accepted: 12 October 2018;

Keywords

Estimate,
Loss,
Risk,
Prior distribution,
Posterior distribution.

ABSTRACT

This paper was taken the loss and risk concepts, the objective of this paper is to estimate the loss and the risk in the Bayes method, as illustrated in the related case Where the prior distribution of the random variable is discrete and continuous, it applied on crop sales data in the continuous case, while it applied by two examples in the discrete case. Then the estimator of the loss function is the mean posterior distribution of the random variable, so it's the same with median posterior distribution, hence is also the Bayes estimator with respect to a loss function equal to absolute deviation. The Bayes risk is a real value is the average of the loss function, so the smallest estimator who is the best estimator of the Bayes risk.

1. Introduction

In our consideration of point estimation we have assumed that our sample $X_1, \dots, X_n$came from some density $f(x,\theta)$was assumed known to us. Moreover, we assumed the parameter θ was fixed value, but it'sunknown to take any possible value in sample space. The statistical inference methods based on these assumptions are called Classical Approaches Methods. In some practical application we may based on previous experience, have additional information about the parameter θ. We may have evidence that the parameter θ as a random variable with a density distribution now ask how this additional information is used to estimate it. To deal with this issue we will assume that θ is the value of a random variable denoted by Θ with a density function $g_\theta(\theta)$) and cumulative function $g_\theta(\theta)$. These function do not contain unknown parameter; we also assume that the set of possible value for the random variable Θ is the sample space Ω. The statistical inference methods based on these assumption is called the Bayes Approach. Thus, the Bayes method is different from the classical method, where the classical looks at parameter θ as a fixed value is unknown while Bayes method is seen as a variable quantity according a previous distribution called a prior distribution. This distribution depends on previous experience about the parameter θ before the sample is withdrawn. The prior distribution is up dated by the information available in sample about the parameter θ derive a posterior distribution, the process of modernization of prior distribution is done by the base of Bayes, this type which depend on this method called Bayesian statistic.

Before dealing with concepts of loss and risk in Bayes method we defined prior and posterior distribution in followed section.

2. Prior and Posterior Distribution

Here to for we have used the notation $f(X;\theta)$to indicate the density of a random variable x for each θ in Ω. Whenever we want to indicate that the parameter θ is the value of a random variable Θ, we shall Write the density of X as $f(x|\theta)$ instead of $f(x;\theta)$ we should note that $f(x|\theta)$is a conditional density; it is the density of given Θ = θ. Amore complete notation for $f(x|\theta)$ would be$f(X|\theta) = \theta(x|\theta)$. Let $X_1, \dots, X_n$ be a random sample of a size n from density$f(X|\theta)$; where θ is a value of a random variable Θ. Assume that the density of Θ, $g_\theta(\theta)$, is known and contains no un known parameters, and suppose that we want to estimate $\tau(\theta)$. How do we incorporate the additional information of known $g_\theta(\theta)$ in to our estimation procedures? In the past, we want thought of the likelihood function as a single expression that contained all our information; the likelihood function included observed sample $x_1, \dots, x_n$ as well as the form of density $f(x;\theta)$ we sampled from in its expression. Now we need an expression that contains all information that the likelihood function contains plus the added information of the known density $g_\theta(\theta).g_\theta(\theta)$ is called prior distribution of Θ. It is minimizes what we know about θafter take a random sample we seek about the posterior distribution of$\theta X_1 = x_1, \dots, X_n = x_n$.

Definition 1

The density $g_\theta(\theta)$ is called the prior distribution of Θ. the conditional density Θ given $X_1 = x_1, \dots, X_n = x_n, denoted f_{\theta|X_1=x_1,\dots,X_n=x_n}(\theta|X_1, \dots, X_n$ is called the posterior distribution

Remark 1

$$f_{\theta|X_1=x_1,\dots,X_n=x_n}(\theta|x_1, \dots, x_n) = \frac{f_{x_1,\dots,x_n|\theta=\theta(x_1,\dots x_n|\theta}g_\theta(\theta)}{f_{x_1,\dots,x_n}(x_1, \dots, x_n)}$$

$$\frac{[\prod_{i=1}^{n} f(X_i|\theta)]g_\Theta(\theta)}{\int[\prod_{i=1}^{n} f(X_i|\theta)]g_\Theta(\theta)d\theta}$$

For rand sampling, $.\left[Recall\ that \int y|x = x\,(y|X) = f_{X,Y}(x,y)|f_X(x)\right] y(x|yf_Y(y)|f_X(x)$.

Tele:
E-mail address: gasimtas@gmail.com .

The posterior distribution replaces the likelihood function as an expression that incorporates all information. If we want to estimate θ and parallel the development of maximum of likelihood estimator of θ, we could take as our estimator of θ that θ which maximizes the posterior distribution that is estimated θ with the mode of posterior distribution, however unlike the likelihood function (as function of θ) the posterior distribution

Definition 2

Let $X_1, \ldots, X_n$ be a random sample from density $f(X|\theta)$, where θ is a value of a random variable Θ with known density $g_\theta(\theta)$. The posterior Bayes estimator of τ (θ) with respect to the prior $g_\theta(\theta)$ is defined to be $E[\tau(\Theta|X_1, \ldots, X_n]$

Remark 2

$$E[\tau(\Theta|X_1 = x_1, \ldots, X_n = x_n] = \int \tau(\theta) f\theta| X_1 = x_1, \ldots, X_n = x_n(\theta|x_1, \ldots x_n) d\theta$$

$$\frac{\int \tau(\theta)[\prod_{i=1}^{n} f(x_i|\theta)] g_\theta(\theta) d\theta}{\int [\prod_{i=1}^{n} f(x_i|\theta)] g_\theta(\theta) d\theta} \quad (1)$$

3. Loss function in Bayesian approach

Loss and risk these two concepts used to assess goodness of estimators, inside that we discuss how the additional information of knowledge of prior distributionof Θ can be used in conjunction loss and risk to define or select on optimum estimator. Paper commence with a review of the problem want to solve. Let $X_1, \ldots, X_n$ be a random sample from a density$f(\mathbf{x}|\theta)$, θ belonging to Ω , where the function $f(x|\theta)$is assumed known except for θ, we assumed that the unknown θ is a value of some random variable Θ and the distribution of Θ is known and contains no unknown parameters . On the basis of the random sample $X_1, \ldots, X_n$was taken to estimate τ (θ), some function of θ. In addition, assumed that a loss function $l(t; \theta)$ has been specified, where $l(t; \theta)$ represents the loss incurred if estimate τ (θ) to be t then θ is the parameter of density from which we sampled. For any estimator$T = t(X_1, \ldots, X_n)$ the $E_\theta[l(T; \theta)]$ represented the average loss of that estimator, and define this average loss to be the risk denoted by $R_t(\theta)$, of the estimator $t(X_1, \ldots, X_n)$further noted that two estimators, say $T_1 = t_1(X_1, \ldots, X_n$ and$T_2 = t_2(X_1, \ldots, X_n)$ could be compared by looking at their respective risks $R_1(\theta)$and $R_2(\theta)$, preference being given to that estimator with smaller risk. In general, the function as functions of θ of two estimators may cross one risk function being smaller for some θ and other smaller for other θ. Then, since θ is unknown, it is difficult to make a choice between the two estimators. The difficulty is caused by the dependence of the risk function θ now, since have assumed that θ is a value of some random variable Θ, the distribution of which is also assumed known, have a natural way of removing the dependence of the risk function on θ, namely, by averaging out the θ, using the density of Θ as our weight function. Mood, Gray bill, Boes [1].

4. Bayes risk

Definition 3

Let $X_1, \ldots, X_n$ be a random sample from a density $f(X|\theta)$, where θ is a value of a random variable Θ with cumulative distribution function $G(\theta) = G_\Theta(\theta)$ and corresponding density $g(\theta) = g_\Theta(\theta)$ in estimating τ (θ), let $l(t; \theta)$be the loss function. The risk of estimate $T = t(X_1, \ldots, X_n$ is denoted by $R_t(\theta)$. The Bayes risk of estimator $T = t(X_1, \ldots, X_n)$ with respect to loss function $l(t; \theta)$ and prior cumulative distribution G(θ)denoted by $r(t) = r_{L,G}(t)$ is defined to be

$$r(t) = \tau_{t,G}(t) = E_\Theta[R_T(\Theta)] = \int_\Omega R_T(\theta) g_\Theta(\theta) d\theta \quad (2)$$

$$r(t) = r_{L,G}(t) = E_\Theta[R_T(\Theta)] = \sum_{\theta \in \Omega} R_T(\theta) g_\Theta(\theta)$$

If $g_\Theta(\theta)$is discrete. Myers The Bayes risk of an estimator is an average risk , the average being over the parameter space Ω with respect to prior density $g(\theta)$ for given loss function $l(t; \theta)$ and prior density $g(\theta)$. The Bayes risk of an estimator is a real number; so now two competing estimator can be readily compared by comparing their respective Bayes risk, still preferring that estimator with smaller Bayes risk. In fact, we can now define the "best" estimator of τ (θ) to be that estimator with smallest Bayes risk

5. Bayes estimator

Definition 4

The Bayes estimator of τ (θ), denoted by $\mathbf{T}^*_{L,G} = \mathbf{t}^*_{L,G}(\mathbf{X}_1, \ldots, \mathbf{X}_n)$ with respect to the loss function $\mathbf{l}(\mathbf{t}; \boldsymbol{\theta})$ and prior cumulative distribution G(θ) , is denoted to be that estimator with smallest Bayes risk . Or the Bayes estimator of τ (θ) is that estimator $\mathbf{t}^*_{L,G}$ satisfying

$$\mathbf{r}_{L,G}(\mathbf{t}^*) = \mathbf{r}_{L,G}(\mathbf{t}^*_{t,G}) \le \mathbf{r}_{L,G}(\mathbf{t})$$

For every other estimator $\mathbf{T} = \mathbf{t}(\mathbf{X}_1, \ldots, \mathbf{X}_n)$ of τ(θ) the posterior Bayes estimator of τ(θ) , define in Definition (2) was defined without regard to a loss function , whereas the definition given above requires specification of a loss function . The definition leaves the problem of actually finding the Bayes estimator which may not be easy an arbitrary loss function,unsolved. However, for square-error loss, finding the Bayes estimator is relatively easy we seek that estimator, say $\mathbf{t}^*(\mathbf{X}_1, \ldots, \mathbf{X}_n)$which minimizes the expression

$$\int_\Omega \mathbf{R}_t(\theta) \mathbf{g}(\theta) \mathbf{d}\theta$$

$$= \int_\Omega \mathbf{E}_\theta[[\mathbf{t}(\mathbf{X}_1, \ldots, \mathbf{X}_n) - \tau(\theta)]^2] \mathbf{g}(\theta) \mathbf{d}\theta$$

$$= \int_\Omega \left\{ \int_x [\mathbf{t}(\mathbf{X}_1, \ldots, \mathbf{X}_n) - \tau(\theta)]^2 \mathbf{f}\mathbf{x}_1 \ldots \mathbf{f}\mathbf{x}_n(\mathbf{X}_1, \ldots, \mathbf{X}_n \middle| \theta) \prod_{i=1}^{n} \mathbf{d}\mathbf{X}_i \right\} \mathbf{g}(\theta) \mathbf{d}(\theta)$$

$$\int_x \left\{ \int_\Omega [\tau(\theta) - \mathbf{t}(\mathbf{X}_1, \ldots, \mathbf{X}_n)]^2 \mathbf{f}\mathbf{X}_1, \ldots, \mathbf{f}\,\mathbf{X}_n(\mathbf{X}_1, \ldots, \mathbf{X}_n|\theta \right\} \frac{\mathbf{g}(\theta)\mathbf{d}(\theta)}{\mathbf{f}\mathbf{X}_1 \ldots, \mathbf{f}\mathbf{X}_n(\mathbf{X}_1, \ldots, \mathbf{X}_n)} \mathbf{f}\mathbf{X}_1, \ldots,$$

$$\int_\Omega \left\{ \int_\Omega [\tau(\theta) - \mathbf{t}(\mathbf{X}_1, \ldots \mathbf{X}_n)]^2 \mathbf{f}(\theta|\mathbf{X}_1 = \mathbf{x}_1, \ldots, \mathbf{X}_n = \mathbf{x}_n(\theta|\mathbf{X}_1, \ldots, \mathbf{X}_n) \mathbf{d}\theta \right\} \mathbf{f}\mathbf{X}_1, \ldots, \mathbf{X}_n\mathbf{x}_n(\mathbf{X}_1, \ldots, \mathbf{X}_n) \prod_{i=1}^{n} \mathbf{d}\mathbf{X}_i$$

, and since the integral is nonnegative, the double integral can be minimized if the expression within the braces is minimized for each $\mathbf{X}_1, \ldots, \mathbf{X}_n$. But the expression with in the braces is conditional expectation of $[\tau(\Theta) - \mathbf{t}(\mathbf{X}_1, \ldots, \mathbf{X}_n]^2$ with respect to the posterior distribution of Θ given $(\mathbf{X}_i = \mathbf{X}_1, \ldots, \mathbf{X}_n)$, which is minimized as a function of $\mathbf{t}(\mathbf{X}_1, \ldots, \mathbf{X}_n)$ for $\mathbf{t}^*(\mathbf{X}_1, \ldots, \mathbf{X}_n)$ equal to the conditional expectation of τ(Θ)with respect to the posterior distribution of Θ given $\mathbf{X}_1 = \mathbf{x}_1, \ldots, \mathbf{X}_n = \mathbf{x}_n$

$\{\mathbf{Recall\ that\ E[(Z - a)^2]\ is\ minimized\ as\ afunction\ of\ a\ for\ a^* = E(Z)}\}$

Hence Bayes estimator of τ (θ) with respect to the square-error loss function is given. Linden, Dose, Toussaint [3]

$$\mathbf{E[\tau\Theta|X_1 = x_1, \ldots, X_n = x_n] = \frac{\int \tau(\theta)[\prod_{i=1}^{n} f(x_i|\theta)]g(\theta)d\theta}{\int[\prod_{i=1}^{n} f(x_i|\theta)]g(\theta)d\theta}} \quad (4)$$

This is identical to the estimator given in Eq (2). For a general loss function, we seek that estimator which minimizes $\int_{\Omega} \mathbf{R_t(\theta)g(\theta)d\theta}$

Again,

$$\int_{\Omega} \mathbf{R_t(\theta)g(\theta)d\theta}$$

$$= \int \left[\int_x \mathbf{l(t(x_1, \ldots, x_n); \theta) f x_1, \ldots, f x_n(x_1, \ldots, x_n} \middle| \theta) \prod_{i=1}^{n} \mathbf{dx_i}\right] \mathbf{g(\theta)d\theta}$$

$$\int_x \left[\int_{\Omega} \mathbf{l(t(X_1, \ldots, X_n); \theta) f\theta|X_1 = x_1, \ldots, X_n} \right.$$

$$\left. = \mathbf{x_n(\theta|x_1, \ldots, x_n)d\theta}\right] \mathbf{fX_1, \ldots, X_n(x_1, \ldots, x_n)} \prod_{i=1}^{n} \mathbf{dx_i}$$

And minimizing the double integral is equivalent to minimizing the expression within the brackets, which is sometimes called the posterior risk. So, in general, the Bayes estimator of τ(θ) with respect to the loss function $\mathbf{l(t;\theta)}$ and prior density $\mathbf{g(\theta)}$ is that estimator which minimizes the posterior risk, which is the expected loss with respect to the posterior distribution of Θ given observation , $\mathbf{x_1, \ldots, x_n}$ we have the following theorem and corollaries

Theorem 5.1.

Let $X_1, \ldots, X_n$ be a random sample from the density $f(X|\theta)$ and let be $g(\theta)$be the density of Θ. Further let $l(t;\theta)$ be the loss function for estimating τ(θ).The Bayes estimator of τ(θ) is that estimator $t^*(X_1, \ldots, X_n)$ which minimizes $\int_{\Omega} l(t(x_1, \ldots, x_n); \theta) f(\theta|X_1 = x_1, \ldots, X_n = x_n)(\theta|x_1, \ldots, x_n)d\theta$

As a function of $l(X_1, \ldots, X_n)$

Proof: The Bayes estimator is that minimizes bellow (for continuous case)

$$E_{\theta}[R_T(\theta)] = \int_{\Omega} R_T(\theta) g_{\theta}(\theta) d\theta = \int_{\Omega} E[L(T(X;\theta)] g_{\theta}(\theta) d\theta$$

$$= \int_{\Omega} \left[\int \ldots_X \int L(t(X;\theta) f_{X|\theta}(X|\theta) dx_1, \ldots, dx_n\right] g_{\theta} d\theta$$

$$= \int_X \ldots \int \left[\int_{\Omega} L(t(X);\theta) f_{X|\theta}(X|\theta) g_{\theta}(\theta) d\theta\right] dx_1, \ldots, dx_n$$

$$= \int_X \ldots \int \left[\int_{\Omega} L(t);\theta) \frac{f_{X|\theta}(X|\theta) g_{\theta}(\theta)}{f_X(X)} d\theta\right] f_X(x) dx_1, \ldots . dx_n$$

$$= \int_X \ldots \int \left[\int_{\Omega} L(t(X);\theta) f_{\theta|X}(\theta|X) d\theta\right] f_X(x) dx_1, \ldots, dx_n$$

$$= \int_X \ldots \int \{E_{\theta|X}[L(t(X);\theta|X = x]\} f_X(x) dx_1, \ldots, dx_n$$

Minimizing integral above equivalent estimator in the blacked it is $E_{\theta|x}[L(t(x;\theta)|X = x]$. Shiaha [4].

Remark: Theestimator$E_{\theta|X}[L(t(x);\theta|X = x] = \int_{\Omega} L(t(x);\theta f_{\theta|x}(\theta)d\theta$ is called by (posterior risk) and it represents the expected loss with respect to posterior distribution$f_{x|\theta}(x|\theta)$.

Corollary5.1

Under the assumption of theorem 5.1, the Bayes estimator of τ (θ) is given by

$$E[\tau(\theta|X_1 = x_1, \ldots, X_n = x_n)] \frac{\int \tau(\theta)[\prod_{i=1}^{n} f(X_i|\theta)] g(\theta) d\theta}{\int[\prod_{i=1}^{n} f(X_i|\theta)] g(\theta) d\theta}$$

For a square-error loss function

Corollary 5.2

Under the assumption of theorem 5.1 the Bayes estimator of θ is given by the median of the posterior distribution of Θ for a loss function equal to absolutedeviation the proofs of the theorem 5.1 and the first corollary preceded the statement of the theorem. The second corollary follows from observation that

$$\int_{\Omega} \mathbf{|\theta - t(X_1, \ldots, X_n)| \, f\theta|X_1 = x_1, \ldots, X_n = x_n(\theta|X_1, \ldots X_n)d\theta}$$

Is minimized as a function of$t(x_1, \ldots, x_n)$ for $t^*(x_1, \ldots, x_n)$ equal to the median of the posterior distribution ofΘ.$\{Recall\ that\ E[Z - a^2]\ is\ minimized\ as\ afunction\ of\ a\ for\ a^* = median\ of\ Z.\}$

6. Minimax estimator

Definition 5.

Minimax estimator as an estimator whose maximum risk less than or equal to the maximum risk of any other estimator. The theorem is some time useful in finding a minimax estimator. The following theorem is useful in finding a minimax estimator

Theorem 6.1

If$\mathbf{T^* = t^*(X_1, \ldots, X_n)}$ is a Bayes estimator having constant risk, that is $R_{t^*}(\theta) = constat$, then T^* is a minimax estimator

Proof:

Let $g^*(\theta)$ be the prior density corresponding to the Bayes estimator

$t^*(X_1, \ldots, X_n) \underset{\theta \in \Omega}{Sup} R_{t^*}(\theta) = costant = R_{t^*}(\theta)$

$\int_{\Omega} R_{t^*}(\theta) g^*(\theta) d\theta \le \int_{\Omega} R_t(\theta) g^*(\theta) d\theta \le \underset{\theta \in \Omega}{Sup} R_t(\theta)$ for any other estimator$t(X_1, \ldots, X_n)$.Shelemyahu, Zack [5]

Another criterion that sometimes is used to select an estimator from class of admissible estimators is the minimax criterion.

Definition. 6

Minimax estimator an estimator T_i is a minimax estimator if $\underset{\theta}{max} R_{T_1}(\theta) \le \underset{\theta}{max} R_T$ for every estimator T

In the other words, $T_1$is an estimator that minimizes the maximum risk, or

$$\underset{\theta}{max} R_{T_1}(\theta) = \underset{T}{min} \underset{\theta}{max} R_T(\theta) \quad (8)$$

Of course, this assumes that the risk function attains a maximum value for some θ and that such maximum values attain a minimum for some T. In a more general treatment of the topic, the maximum and minimum could be replaced with the more general concepts of least upper bound and greatest lower bound, respectively. Bain [6]

7. Application

In this section firstly we applied with data on a value of crop sales by million Sudanese pound from marketing center, this data as a random sample of size 30 as follows

Table 1. sales of crop by million Sudanese bound.

5	4.5	5
4.5	4.5	4
5	2	5
5	4.5	4
4.5	4.5	3
5	4	4
3	5	4.5
4.5	4	4
5	5	4.5
5	6	7

Let this random sample $\mathbf{5}$, $\mathbf{4.5}$, ... , $\mathbf{7}$from normal density with mean θ and variance 1, and then we use this to estimate θ with a squared-error loss function. Assume that Θ has a normal density function with mean μ_0 andvariance$\sigma^2 = 1$, now write $\mu_0 = x_0$ when convenient according E q (4) the Bayes estimator is given as the mean of the posterior distribution of Θ.

$$E[\tau(\Theta)|X = x] = E[(\Theta)|X_1 = 5, \dots, X_{30} = 7]$$

$$= \int_{\Omega} \tau(\theta) f_{\Theta|X}(\theta|X) d\theta$$

$\Theta \sim Normal(\mu_0 , 1)$

$g_{\Theta}(\theta) = \frac{1}{\sqrt{2\pi}} exp\{-\frac{1}{2}(\theta - \mu_0)^2\} I_{(-\infty,\infty)}(\theta) \quad , \theta \in \Omega = \{\theta: -\infty < \theta < \infty\}$

$$f(X;\theta) = f(X|\theta) = \frac{1}{2} exp\left\{-\frac{1}{2}(x-\theta)^2\right\} I_{(-\infty,\infty)}(x)$$

$$f_{\Theta|X}(\theta|X) = f_{\Theta|x_1,\dots,x_{30}}(\theta|X_1, \dots, X_{30}) = \frac{f_{X|\Theta}(X|\theta) g_{\Theta}(\theta)}{f_X(X)}$$

$$= \frac{f_{X_1,\dots,X_{30}|\Theta}(x_1,\dots,x_{30}|\theta) g_{\Theta}(\theta)}{f_{X_1,\dots,X_{30}}(x_1,\dots,x_{30})}$$

$$= \frac{\left[\prod_{i=1}^{30} f(x_i|\theta)\right] g_{\Theta}(\theta)}{\int_{\Omega}\left[\prod_{i=1}^{30} f(x_i|\theta)\right] g_{\Theta}(\theta) d\theta}$$

$$= \frac{(1/\sqrt{2\pi})^n exp[-1/2\sum_1^n (x_i-\theta)^2](1/\sqrt{2\pi}) exp[-1/2(\theta-\mu_0)^2]}{\int_{-\infty}^{\infty}(1/\sqrt{2\pi})^n exp[-1/2\sum_1^n (x_i-\theta)^2](1/\sqrt{2\pi}) exp[-1/2(\theta-\mu_0)^2] d\theta}$$

$$= \frac{exp[-1/2\sum_{i=0}^n (x_i-\theta)^2]}{\int_{-\infty}^{\infty} exp[-1/2\sum_{i=0}^n (x_i-\theta)^2] d\theta}$$

$$= \frac{exp\{[-1/2(n+1)\theta^2 - 2\theta\sum_{i=1}^n x_i + \sum_{i=0}^n x_i^2]\}}{\int_{-\infty}^{\infty} exp\{-1/2[(n+1)\theta^2 - 2\theta\sum_{i=0}^n x_i + \sum_{i=0}^n x_i^2]\} d\theta}$$

$$= \frac{exp\left[(-(n+1)/2\left\{\theta^2 - 2\theta\sum_{i=0}^n x_i/(n+1) + (\sum_{i=0}^n x_i/(n+1)^2)\right\}\right]}{\int_{-\infty}^{\infty} \exp(-[n+1/2]\left\{\theta^2 - 2\theta\sum_{i=0}^n x_i/(n+1) + [\sum_{i=0}^n x_i/n+1]^2\right\} d\theta}$$

$$= \frac{[1/\sqrt{2\pi(n+1)}] exp\left\{-[(n+1)/2]\left[\theta - \sum_{i=0}^n x_i/(n+1)^2\right]\right\}}{\int_{-\infty}^{\infty}[1/\sqrt{2\pi(n+1)}] exp\left\{-[(n+1)/2]\left[\theta - \sum_{i=0}^n x_i/n+1^2\right]\right\} d\theta}$$

$$= \frac{1}{\sqrt{2\pi(n+1)}} exp\left\{\frac{-(n+1)}{2}\left[\theta - \sum_{i=0}^n x_i/(n+1)^2\right]\right\};$$

The denominator is unity since it is the integral of a density. We have shown that posterior distribution Θ is normal with mean $\sum_0^n X_i/(n+1)$ and variance $1/(n+1)$; hence the Bayes estimator of θ with respect to square-error loss is $\frac{X_0+\sum_1^n X_i}{n+1} = \frac{\mu_0+\sum_1^n X_i}{n+1} = \frac{4.417+132.5}{30+1} = 4.417$ as the following tables

Table 2.one-sample statistic.

	N	Mean	Std Deviation	Std error Mean
sell	30	4.417	1	.18

Table 3. one sample test.

	t	Df	Sig 2tailed	MD	95%confidence interval of difference	
					Lower	Upper
sell	24.2	29	.000	4.42	4.043	4.790

Since the posterior distribution of Θ is normal, it's mean and median are the same, hence $\frac{\mu_0+\sum_1^n X_i}{n+1} = 4.417$ is also the Bayes estimator with respect to a loss function equal to the absolute deviation.

Secondly this application is based on the case where the prior distribution of the random variable is discrete bellow.

Let$X_1, \dots, X_n$ denote a random sample from Bernoulli density

$$f(X|\theta) = \theta^x(1-\theta)^{1-x} for x = 0, 1.$$

Assume that the prior distributed over the interval (0, 1). Consider estimating θ and τ (θ) =θ (1-θ)

Now$f\theta|X_1 = x_1, \dots, X_n = x_n(\theta|x_1, \dots, x_n)$= $\frac{\theta^{\sum x_i}(1-\theta)^{n-\sum x_i} I_{(0,1)}(\theta)}{\int_0^1 \theta^{\sum x_i}(1-\theta)^{n-\sum x_i} d\theta}$;

So the posterior Bayes estimator of θ with respect to the uniform prior distribution is given by

$$E[\Theta|X_1 = x_1, \dots, X_n = x_n] = \int \theta \int \Theta|X_1 = x_1, \dots, X_n = x_n(\theta|x_1, \dots, x_n) d\theta$$

$$= \frac{\int_0^1 \theta\theta^{\sum x_i}(1-\theta)^{n-\sum x_i} d\theta}{\int_0^1 \theta^{\sum x_i}(1-\theta)^{n-\sum x_i} d\theta} = \frac{B(\sum X_i+2, n-\sum X_i+1)}{B(\sum X_i+1, n-\sum X_i+1)}$$

$$= \frac{\Gamma(\sum X_i+2)\,\Gamma(n-\sum X_i+1)}{\Gamma(n+3)} \frac{\Gamma(n+2)}{\Gamma(\sum X_i+1)\,\Gamma(n-\sum X_i+1)}$$

$$= \frac{\sum X_i+1}{n+2}$$

Hence the posterior Bayes estimator of θ with respect to the uniform prior distribution is given by $(\sum X_i + 1)/(n+2)$. To obtain the posterior Bayes estimator of, say τ(θ)=θ(1-θ),

We calculate E $(\tau\Theta|X_i = x_i, \dots, X_i = x_i)$= $\int \theta(1-\theta) f\Theta|X_i = x_i, \dots, X_n = x_n(\theta|x_1, \dots, x_n) d\theta$

$$= \frac{\int \theta(1-\theta)^{\sum X_i}\theta(1-\theta)^{n-\sum x_i} d\theta}{\int_0^1 \theta^{\sum x_i}(1-\theta)^{n-\sum x_i} d\theta}$$

$$= \frac{\Gamma(X_i+2)\,\Gamma(2-\sum X_i+2)}{\Gamma(n+4)} \frac{\Gamma n+2}{\Gamma\sum X_i+1\,\Gamma 2-\sum X_i+1}$$

$$= \frac{(x_i+1)(n-\sum X_i+1)}{(n+3)(n+2)}$$

So the posterior Bayes estimator of θ (1-θ) with respect to a uniform prior distribution is $=\frac{(\sum X_i+1)(n-\sum X_i+1)}{(n+3)(n+2)}$.

Thirdly this application shows the case in which the loss is constant as fellow:

Considers the minimax estimator of θ in sampling from Bernoulli distribution using square-error loss function. We seek a Bayes estimator with constant risk. The family of beta distribution is a family of a possible prior distribution. We want that for one of beta prior distribution the corresponding Bayes estimator will have constant risk. A Bayes estimator is given by

$$\frac{\int_0^1 \theta\theta^{\sum X_i}(1-\theta)^{n-\sum X_i}\left[1/B(a,b)\theta^{a-1}(1-\theta)^{b-1}\right] d\theta}{\int_0^1 \theta^{\sum X_i}(1-\theta)^{n-\sum X_i}[1/B(a,b)]\theta^{a-1}(1-\theta)^{b-1} d\theta}$$

$$= \frac{\int_0^1 \theta^{\sum X_i+a-1}(1-\theta)^{n-\sum X_i+b-1} d\theta}{\int_0^1 \theta^{\sum X_i+a-1}(1-\theta)^{n-\sum X_i+b-1} d\theta}$$

$$= \frac{B(\sum X_i+a+1, n-\sum X_i+b)}{B(\sum X_i+a, n-\sum X_i+b)} = \frac{\sum X_i+a}{n+a+b}$$

So the Bayes estimator with respect to beta prior distribution having parameters a and b is given by $\frac{\sum X_i+a}{n+a+b}$.

We now evaluate the risk of $\sum X_i + a/n + a + b$ with hope that we will be able to select a and b so that the risk will be constant. Write

$t^*_{A,B}(x_1, \dots, x_n) - A\sum X_i + B - \sum X_i + a/n + a + b$;

Then

$$Rt_{A,B}(\theta) = E\left[(A\sum X_i + B - \theta)^2\right]$$

$$= E\left[\left[A(\sum X_i - n\theta) + B - \theta + nA\theta\right]^2\right]$$

$$= A^2 E\left[(\sum X_i - n\theta)^2\right] + (B - \theta + nA\theta)$$

$= nA^2\theta(1-\theta) + (B - \theta + nA\theta)^2 = \theta^2[(nA-1)^2 - nA^2] + \theta[nA^2 + 2(nA-1)B] + B^2$;

Which is constant if

$$(nA-1)^2 - nA^2 = 0 \text{ and } nA^2 + 2(nA-1)B = 0$$

Now $(nA-1)^2 - nA^2 = 0 \ \ if\ A = 1/\sqrt{n}\,(\sqrt{n} \pm 1)$;

And $nA^2 + 2(nA-1)B = 0 \ \ if\ B = -nA^2/2(nA-1)$, which is

$1/2(\sqrt{n}+1) \ \ for \ \ A = 1/\sqrt{n}(\sqrt{n}+1)$.

On solving for a and b, we obtain $a = b = \sqrt{n}/2$; so $(\sum X_i + \sqrt{n}/2)/(n+\sqrt{n})$

Is a Bayes estimator with constant risk and hence, minimax.

8. Conclusion

The paper conclusive that the loss function is the mean of posterior distribution of the random variable and it's also the median of it. Further wise the Bayes estimator with respect to a loss function equal to absolute deviation. The estimator Bayes risk is the average of the loss function as the smallest estimator is the best estimator.

References

[1]Mood, A. M,Graybill, F. A, and Boes, D.C,(1974) "introduction to theory of statistics", 3rd edition, New York, Mc Craw-Hill, p 343.

[2]Myers, Walpole, Ye,(2014), "Probability and Statistics for Engineers and Scientists", Ninth edition, Person Education Limited, United State of America, p 116.

[3]Linden, W.V, Volker Dose, Toussaint, U.V, (2014), "Bayesian Probability Theory", Clays, St Ives plc, United Kingdom, p 228.

[4]Shiaha, Abdullah Abd Elkreem, (2007),"bases of Estimate Theorem" King Saud University, Kingdom of Saudi Arabic, p 307.

[5]Shelemyahu, Zacks, (1981), "Parametric Statistical Inference Basic Theory and Modern Approaches", first edition, pergamon press, New York, U.S.A, p325

[6]Bain, Lee.J, (1992), "Introduction to Probability and Mathematical Statistics", second edition, Duxbury press, California, U.S.A, p 321