A Note on Bayes Risks of Estimators

Se-youhl Lee

추정량의 베이즈 위험에 관한 소고

李 世 烈

Summary

The Bayes estimation theory is based on the prior space, the sample space, the loss function and the posterior distribution, etc.

In this paper we observe some loss functions which are concerned with mean, median and mode of the posterior distribution of the parameter, and we derive, using the predictive distribution and the posterior distribution, the retation of Bayes risks.

Introduction

Let $(\Omega, \mathcal{H}, Q)$ be a prior space of the parameter θ and $(\mathscr{X}\ \mathcal{U},\ P_\theta)$, $\theta\epsilon\Omega$ be a sample space of the random variable or the random vector x. Let the prior distribution Q be absolutely continuous with respect to the σ-finite measure ν defined on $\mathcal{H}$, and let the σ-finite measure μ defined on $\mathcal{U}$ dominate the family of sample distributions $\{P_\theta, \theta\epsilon\Omega\}$.

By Radon-Nikodym theorem and Fubini theorem, we can define densities relative to ν or μ and the predictive distribution P^* as follows;

(1.1) $q(\theta) = \frac{dQ}{d\nu}(\theta)$,

(1.2) $f(x;\theta)\frac{dP_\theta}{d\mu}(x)$,

(1.3) $m(x) = \int_\Omega f(x;\theta)\, q(\theta)\, d\nu(\theta)$,

(1.4) $P^*(A) = \int_A m(x) d\mu(x)$ for all $A\epsilon\mathcal{U}$.

And we shall define the posterior density $\pi(\cdot;x)$ relative to ν and the posterior distribution Π_x as follows;

(1.5) $\pi(\theta;x)\, m(x) = f(x;\theta)\, q(\theta)$,

(1.6) $\Pi_x(D) = \int_D \pi(\theta;x)\, d\nu(\theta)$ for all $D\epsilon\mathcal{H}$.

In the above statements, $m(x)>o$ and $q(\theta)>O$ for all $x\epsilon\mathscr{X}$ and $\theta\epsilon\Omega$ are regarded as to be assumed. Then for expectations

(1.7) $E^Q E^x_\theta\ \Psi(x,\theta) = \int_\Omega \int_{\mathscr{X}} \Psi(x,\theta)\, f(x;\theta)\, q(\theta)\, d\mu(x)\, d\nu(\theta)$, (1.8) $E^{P^*}E_x^{\theta}\,\Psi(x,\theta)=\int_{\mathscr{X}}\int_{\Omega}\Psi(x,\theta)\,\pi(\theta;x)\,m(x)\,d\nu(\theta)\,d\mu(x),$

the following equations are established by Fubini theorem;

(1.9) $E^{Q}E_{\theta}^{x}\,\Psi(x,\theta)=E^{P^*}E_x^{\theta}\,\Psi(x,\theta),$

(1.10) $E^{Q}E_{\theta}^{x}\,g(x)=E^{P^*}E_x^{\theta}\,g(x)=E^{P^*}\,g(x),$

(1.11) $E^{P^*}E_x^{\theta}\,h(\theta)=E^{Q}E_{\theta}^{x}\,h(\theta)=E^{Q}\,h(\theta).$

We shall represent the estimator of $\varphi(\theta)$ to T(x), the loss function to $W(T(x);\theta)$, it's weight to $\alpha(\theta)$ and the indicator function of the .set A to I_A. We shall assume, as a matter of convenience, that T(x) is a μ-measurable function defined on $\mathscr{X}$, $\varphi(\theta)$ and $\alpha(\theta)$ are ν-measurable functions defined on Ω, and $W(T(x);\theta)$ is a $(\mu x \nu)$-measurable function defined on $\mathscr{X} X\,\Omega$.

The Bayes riske and the Bayes estimator are defined as follows:

Definition 1.

(1.12) $\gamma_W(T)=E^{Q}E_{\theta}^{x}\,W(T(x);\theta)$

is called the Bayes riske of T(with respect to Q). If an estimator T^Q which minimizes $\gamma_W(T)$ exists, then T^Q is called the Bayes estimator and the quantity $\gamma_W(T^Q)$ is called the Bayes risk of Q.

Theorems

The squared-erro loss, the linear loss and the "0-1" loss are concerned with mean, median and mode of the posterior distribution of the parameter θ, respectively.

We deal with these loss functions. And we shall describe the relation of Bayes risks, using the predictive distribution P* and the posterior distribution Π_x.

The following lemma is well-known [1].

Lemma 2. For the weighted squared-erro loss function

$$W(T(x);\theta)=\alpha(\theta)\,(T(x)-\varphi(\theta))^2,\ \alpha(\theta)>0,$$

the Bayes estimator of $\varphi(\theta)$ is given by

(2.1) $$T^Q(x)=\frac{E^{\Pi x}\varphi(\theta)\alpha(\theta)}{E^{\Pi x}\alpha(\theta)}=\frac{\int_{\Omega}\varphi(\theta)\alpha(\theta)f(x;\theta)q(\theta)d\nu(\theta)}{\int_{\Omega}\alpha(\theta)f(x;\theta)q(\theta)d\nu(\theta).}$$

Theorem 3. If

$$W(T(x);\theta)=\alpha(\theta)(T(x)-\varphi(\theta))^2,\ \alpha(\theta)>0,$$
$$\gamma_W(T)<\infty,$$

then

(3.1) $$\gamma_W(T)=\gamma_W(T^Q)+E^{P^*}(E_x^{\theta}\,\alpha(\theta))(T(x)-T^Q(x))^2.$$

Proof :

(3.2) $$\begin{aligned}\gamma_W(T) &= E^{Q}E_{\theta}^{x}\,\alpha(\theta)\,(T(x)-\varphi(\theta))^2\\ &= E^{P^*}E_x^{\theta}\,\alpha(\theta)(T(x)-\varphi(\theta))^2\\ &= E^{P^*}E_x^{\theta}\,\alpha(\theta)(T^Q(x)-\varphi(\theta))^2\\ &+ E^{P^*}E_x^{\theta}\,\alpha(\theta)\,(T(x)-T^Q(x))^2\\ &+ 2E^{P^*}E_x^{\theta}\,\alpha(\theta)\,(T(x)-T^Q(x))(T^Q(x)-\varphi(x))\\ &= E^{Q}E_{\theta}^{x}\,\alpha(\theta)\,(T^Q(x)-\varphi(x))^2\\ &+ E^{P^*}(E_x^{\theta}\,\alpha(\theta))(T(x)-T^Q(x))^2\\ &+ 2E^{P^*}(T^Q(x)E_x^{\theta}\,\alpha(\theta))\,(T(x)-T^Q(x))\\ &- 2E^{P^*}\{E_x^{\theta}\,\varphi(\theta)\alpha(\theta)\alpha(\theta)\}\,(T(x)-T^Q(x)).\end{aligned}$$

Since, on the last of (3.2),

$$E^Q E_\theta^x \alpha(\theta)(T^Q(x)-\varphi(x)^2 = \gamma_W(T^Q)$$

by definition 1 and

$$T^Q(x) E_x^\theta \alpha(\theta) = E_\theta^x \varphi(\theta)\alpha(\theta)$$

by lemma 2, we obtain

$$\gamma_W(T)=\gamma_W(T^Q)+E^{P^*}(E_x^\theta \alpha(\theta))(T(x)-T^Q(x))^2.$$

Corrolary 4. If

$$W(T(x)\theta)=(T(x)-\theta)^2,$$
$$\gamma_W(T)<\infty,$$

then the Bayes estimator of θ is given by the mean of the posterior distribution of θ, i.e.,

(4.1) $T^Q(x)=E^{\Pi x}(\theta)$,

and

(4.2) $\gamma_W(T)=\gamma_W(T^Q)+E^{P^*}(T(x)-T^Q(x))^2$.

The following lemmas are well-known [1].

Lemma 5.1. For the weighted linear loss function.

$$W(T(x)\theta)=\alpha(\theta)|T(x)-\varphi(\theta)|, \alpha(\theta)>0,$$

if

$$E^{\Pi x}\alpha(\theta)<\infty,$$
$$E^{\Pi x}\alpha(\theta) I_{(-\infty, M(x))}(\varphi(\theta)) \leq \frac{1}{2}E^{\Pi x}\alpha(\theta),$$
$$E^{\Pi x}\alpha(\theta) I_{(-\infty, M(x)]}(\varphi(\theta)) \geq \frac{1}{2}E^{\Pi x}\alpha(\theta),$$

then the Bayes estimator of $\varphi(\theta)$ is given by

(5.1.1) $T^Q(x)=M(x)$.

Lemma 5.2. For the loss function

$$W(T(x)\theta)=\alpha(\theta)\left[\frac{k_0+k_1}{2}|T(x)-\varphi(\theta)|-\frac{k_0-k_1}{2}(T(x)-\varphi(\theta))\right], \alpha(\theta)>0, k_0>0, k_1>0,$$

if

$$E^{\Pi x}\alpha(\theta)<\infty,$$
$$E^{\Pi x}\alpha(\theta) I_{(-\infty, M(x))}(\varphi(\theta)) \leq \frac{k_0}{k_0+k_1}E^{\Pi x}\alpha(\theta),$$
$$E^{\Pi x}\alpha(\theta) I_{(-\infty, M(x)]}(\varphi(\theta)) \geq \frac{k_0}{k_0+k_1}E^{\Pi x}\alpha(\theta),$$

then

(5.2.1) $T^Q(x)=M(x)$.

Theorem 6. For the loss function

$$W(T(x)\theta)=\alpha(\theta)|T(x)-\varphi(\theta)|, \alpha(\theta)>0.$$

if

$$E^{\Pi x}\alpha(\theta) <\infty,$$
$$\gamma_W(T)<\infty,$$

then

(6.1) $\gamma_W(T)\leq\gamma_W(T^Q)+E^{P^*}(E_x^\theta \alpha(\theta))|T(x)-T^Q(x)|$.

Proof :

$$\begin{aligned}\gamma_W(T)&=E^Q E_\theta^x \alpha(\theta)|T(x)-\varphi(\theta)|\\ &=E^{P^*}E_x^\theta \alpha(\theta)|T(x)-\varphi(\theta)|\\ &\leq E^{P^*}E_x^\theta \alpha(\theta)|T^Q(x)-\varphi(\theta)|\\ &+E^{P^*}E_x^\theta \alpha(\theta)|T(x)-T^Q(x)|\\ &=E^Q E_\theta^x \alpha(\theta)|T^Q(x)-T^\varphi(x)|\\ &+E^{P^*}(E_x^\theta \alpha(\theta))|T(x)-T^Q(x)|\\ &=\gamma_W(T^Q)+E^{P^*}(E_x^\theta \alpha(\theta))|T(x)-T^Q(x)|.\end{aligned}$$

Corrolary 7. If

$W(T(x);\theta)=|T(x)-\theta|,$

$\gamma_W(T)<\infty,$

then the Bayes estimator of θ, $T^Q(x)$ is given by the median of the posterior distribution of θ, and

$$(7.1)\ \gamma_W(T)\leqq\gamma_W(T^Q)+E^{P^*}|T(x)-T^Q(x)|.$$

Corrolary 8. If

$$W(T(x);\theta)=\frac{k_0+k_1}{2}|T(x)-\theta|-\frac{k_0-k_1}{2}(T(x)-\theta),$$

$k_0>0,\ k_1>0,$

$\gamma_W(T)<\infty,$

then the Bayes estimator of θ, $T^Q(x)$ is given by the $\frac{k_0}{k_0+k_1}$ – fractile of the posterior distribution of θ, and

$$(8.1)\ \gamma_W(T)=\gamma_W(T^Q)+E^{P^*}\left[\frac{k_0+k_1}{2}|T(x)-T^Q(x)|-\frac{k_0-k_1}{2}(T(x)-T^Q(x))\right].$$

The following lemma can be obtained.

Lemma 9. For the weighted "0-1" loss function

$$W(T(X);\theta)=\begin{cases}\alpha(\theta) & \text{if } |T(x)-\varphi(\theta)|>\varepsilon\\ 0 & \text{if } |T(x)-\varphi(\theta)|<\varepsilon\end{cases},\ \alpha(\theta)>0$$

for arbitrary small $\varepsilon>0$, if

$\nu(\{\theta;|T(x)-\varphi(\theta)|=\omega\})=0,$

$A_{T.\varepsilon}=\{\theta;|T(x)-\varphi(\theta)|<\omega\}\cap\Omega,$

$E^{\Pi x}\alpha(\theta)<\infty,$

$E^{\Pi x}\alpha(\theta)I_{AM.\varepsilon}(\theta)\geqq E^{\Pi x}\alpha(\theta)I_{AT.\varepsilon}(\theta)$

for all T, then the Bayes estimator of $\varphi(\theta)$ is given by

$$(9.1)\ T^Q(x)=M(x).$$

Proof: Let

$B_{T.\varepsilon}=\{\theta;|T(x)-\varphi(\theta)|>\varepsilon\}\cap\Omega.$

Then from the hypotheses,

$$\begin{aligned}(9.2)\ E^{\Pi x}W(M(x);\theta)&=\int_\Omega W(M(x);\theta)\pi(\theta;x)d\nu(\theta)\\ &=\int_{BM.\varepsilon}\alpha(\theta)\pi(\theta;x)d\nu(\theta)\\ &=\int_{\Omega\text{-}AE.\varepsilon}\alpha(\theta)\pi(\theta;x)d\nu(\theta)\\ &=E^{\Pi x}\alpha(\theta)-E^{\Pi x}\alpha(p)I_{AM.\varepsilon}(\theta),\end{aligned}$$

similarly

$$(9.3)\ E^{\Pi x}W(T(x);\theta)=E^{\Pi x}\alpha(\theta)-E^{\Pi x}\alpha(\theta)I_{AT.\varepsilon}(\theta),$$

and therefore

$$(9.4)\ E^{\Pi x}W(M(x);\theta)\leqq E^{\Pi x}W(T(x);\theta).$$

Theorem 10. Under the hypotheses of lemma 9, if

$\gamma_W(T)<\infty,$

then

$$(10.1)\ \gamma_W(T)=\gamma_W(T^Q)+E^{P^*}E_x^\theta\alpha(\theta)[I_{AT.\varepsilon}(\theta)-I_{AT\cdot\varepsilon}(\theta)]$$

Proof. Subtracting (9.2) from (9.3) and appling

$T^Q(x)=M(x),$

We obtain

$$\begin{aligned}&E^{\Pi x}W(T(x);\theta)-E^{\Pi x}(T^Q(x);\theta)\\ &\quad=E^{\Pi x}\alpha(\theta)[I_{ATQ.\varepsilon}(\theta)-I_{AT.\varepsilon}(\theta)],\end{aligned}$$

$$\begin{aligned}&E^{P^*}E_x^\theta W(T(x);\theta)-E^{P^*}E_x^\theta W(T^Q(x);\theta)\\ &\quad=E^{P^*}E_x^\theta\alpha(\theta)[I_{ATQ.\varepsilon}(\theta)-I_{AT\cdot\varepsilon}(\theta)]\end{aligned}$$

$$E^{Q}E_{\theta}^{x}\,W(T(x);\theta)-E^{Q}E_{\theta}^{x}\,W(T^{Q}(x);\theta)$$
$$=E^{P^*}E_{x}^{\theta}\,\alpha(\theta)[I_{A_{T^Q}\cdot\varepsilon}(\theta)-I_{A_{T}\cdot\varepsilon}(\theta)].$$

Therefore

$$\gamma_W(T)-\gamma_W(T^Q)=E^{P^*}E_x^{\theta}\,\alpha(\theta)[I_{A_{T^Q}\cdot\varepsilon}(\theta)-I_{A_T\cdot\varepsilon}(\theta)].$$

Corrolary 11. If

$$W(T(x);\theta)=\begin{cases}1 \text{ if } |T(x)-\theta|>\varepsilon\\ o \text{ if } |T(x)-\theta|<\varepsilon\end{cases}$$

for arbitrary small $\varepsilon>0$, ν is the usual measure on $\Omega=R$ and $\pi(\theta;x)$ is continuous, then the Bayes estimator of θ, $T^Q(x)$ is given by the mode of the posterior distribution of θ, and

(11.1)

$$\gamma_W(T)\approx\gamma_W(T^Q)+2\varepsilon E^{P^*}[\pi(T^Q(x);x)-\pi(T(x);x)].$$

Literatures Cited

[1] Berger, J., 1980, *Statistical Decision Theory,* Springer-Verlag.

[2] George E.P Box & George C. Tiao, 1973. *Bayesian Inference in Statistical Analysis,* Addison-Wesley.

[3] I.A.Ibragimov & R.Z. Has' minsPeii, 1981, *Statistical Estimation,* Springer-Verlay.

[4] S.H. Shipiro, 1973, "On Computing Bayes Risks" *Journal of the American Statistical Association* 68, June, No. 342.

국 문 초 록

추정량의 베이즈 위험에 관한 소고

베이즈 추정 이론은 실험 공간, 표본 공간, 손실 함수와 사후 분포 등에 바탕을 두고 있다.

본 논문에서는 모수의 사후 분포 평균, 사후 분포 중위수 및 사후 분포 최빈수에 관계되는 몇 가지 손실 함수를 생각하고, 베이즈 위험들 간의 관계를 예측 분포와 사후 분포를 이용해서 규명하였다.