Effect of Nonnormality on Bayes Decision Function for Testing Normal Mean

Ashok K. Bansal*

ABSTRACT

A zone of sensitivity is developed to investigate the effect of nonnormality on the Bayes decision function for testing mean of a normal population when either parent or prior belongs to Edgeworthian family of moderately nonnormal probability density functoins.

1. Introduction

In classical normal theory Bayes decision problems, a random sample is assumed to be drawn from a normal population and then the conjugate prior is chosen for mathematical simplicity. Quite often, in the absence of the precise form of the overall distribution, the researcher finds only a finite number of important assignable sources of variation. In such a situation, the central limit theorem does not allow him to assume exact normality. However, it does provide a basis for thinking of his distribution belonging to a family of moderately nonnormal distributions. Following Box and Tiao (1962), earlier workers considered a class of symmetric exponential power distributions to study inference robustness of normal theory Bayesian procedures. As the inference concerning unknown mean is affected by skewness as well, Bansal (1978a,b) found Edgeworth series approach useful in investigating the effect of nonnormality in

* Department of Mathematical Statistics, Faculty of Mathematics, University of Delhi, Delhi 110007, INDIA. I am grateful to the U.G.C. of India for partially financing this investigation and the referee for suggestions.

samples (or assumed prior) on the Bayes estimator for the population mean.

It is well known that two sided test of simple hypothesis exist in Bayesian inference. If the prior information on the unknown mean θ is substantial, according to Jeffreys (1961), the particular value θ_0 has a different order of importance from the other values of θ. Thus our prior information about θ should be in two parts; a prior discrete probability $p(>0)$ for $\theta=\theta_0$, together with some conditional prior distribution over $\theta \neq \theta_0$. Following this approach, in this paper, we reconsider the Bayes decision function (BDF) analogue of the significance test for the mean of a normal population with known variance. The effect of nonnormality on such a BDF is illustrated in terms of a zone of sensitivity (ZS) around the true probability p for $\theta=\theta_0$.

2. BDF with Edgeworth Conditional Prior

Suppose that a random sample $\boldsymbol{x}=(x_1, x_2, \cdots, x_n)$ of size n is drawn from a normal population for which the value of mean θ is unknown but the precision is known to be r.

Consider the problem of testing $H_0 : \theta=\theta_0$ against the two sided alternative $H_1 : \theta \neq \theta_0$ as a two-decision problem in which decision d_i $(i=0, 1)$ amounts to acceptance of H_i. Further, let the loss function be quadratic and specified by

$$L(d_0,\theta)=a(\theta-\theta_0)^2 \geq 0 \text{ for } \theta \varepsilon (-\infty, \infty), \qquad (1)$$

$$L(d_1,\theta)=\begin{cases} b \text{ for } \theta=\theta_0 \\ 0 \text{ otherwise.} \end{cases}$$

with constants $a>0$ and $b>0$.

Following Jeffreys, we take the prior distribution of the unknown mean Θ as mixed type, comprising discrete probability mass p at $\theta=\theta_0$ and a continuous distribution of total probability $(1-p)$ with density $\xi(\theta)$ such that $\int_{-\infty}^{\infty} \xi(\theta) d\theta = 1-p$. In the absence of the true conditional prior, assume that $\xi(\theta)$ for $\theta \neq \theta_0$ belongs to a family of Edgeworth series distributions (ESD) represented by the first four terms

$$\xi(\theta)=[1+\frac{1}{6}\lambda_3 H_3\{\sqrt{\tau}(\theta-\mu)\}+\frac{1}{24}\lambda_4 H_4\{\sqrt{\tau}(\theta-\mu)\}+\frac{1}{72}\lambda_3{}^2 H_6\{\sqrt{\tau}(\theta-\mu)\}](\tau/2\pi)^{\frac{1}{2}}\exp[-\frac{1}{2}\tau(\theta-\mu)^2], \quad (2)$$

Here $H_k(\cdot)$ is a Hermite polynomial of degree k; $\mu,\tau,\lambda_3{}^2$ and λ_4 are know mean, precision, skewness and kurtosis of the density.

The sample data $\boldsymbol{x}$ refines the probability that $\theta=\theta_0$ and the density of Θ over $\theta\neq\theta_0$, to produce the mixed type posterior distribution of the unknown mean Θ. The Bayes risk in accepting H_0 (decision d_0) is

$$\rho(d_0)=E[L(d_0,\ \Theta)\,|\,\boldsymbol{x}=\boldsymbol{x}]=a(1-p)\int_{-\infty}^{\infty}(\theta-\theta_0)^2\xi_{\boldsymbol{x}}(\theta)\,d\theta$$
$$=a(1-p)[\rho_{\boldsymbol{x}}{}^*(\xi)+\{\delta^*(\boldsymbol{x})-\theta_0\}^2], \quad (3)$$

where $\xi_x(\theta)$, $\delta^*(\boldsymbol{x})$, $\rho_{\boldsymbol{x}}{}^*(\xi)$ are given in Bansal (1978a) as posterior density, Bayes estimator for Θ and the associated Bayes risk, respectively. The Bayes risk in accepting H_1 is

$$\rho(d_1)=bp[\Theta=\theta_0\,|\,\boldsymbol{x}=\boldsymbol{x}]$$
$$=bpf_n(\boldsymbol{x}\,|\,\theta_0)/[pf_n(\boldsymbol{x}\,|\,\theta_0)+(1-p)\int_{-\infty}^{\infty}f_n(\boldsymbol{x}\,|\,\theta)\xi(\theta)\,d\theta]$$
$$=bp/[p+(1-p)c], \quad (4)$$

where $f_n(\boldsymbol{x}\,|\,\theta)$ is the likelihood function and

$$c=G\sqrt{\tau\tau'}\exp[-\frac{1}{2}nr\{\tau\tau'(\bar{x}-\mu)^2-(\bar{x}-\theta_0)^2\}],$$
$$\tau'=(\tau+nr)^{-1}. \quad (5)$$

We observe that the risk curve for d_0 in the $p\rho$-plane is a straight line with slope

$$A=-a[\rho_{\boldsymbol{x}}{}^*(\xi)+\{\delta^*(\boldsymbol{x})-\theta_0\}^2], \quad (6)$$

and passes through the point $(1, 0)$. For decision d_1, the risk curve is a segment of a rectangular hyperbola

$$\left(p+\frac{c}{1-c}\right)\left(\rho-\frac{b}{1-c}\right)=-cb(1-c)^{-2}, \quad (7)$$

lying in the first quadrant of $p\rho$-plane with end points $(0, 0)$ and $(1, b)$. The two curves intersect at $(p^*,\ \rho^*)$ where p^* is the positive root (≤ 1) of the quadratic equation

$$A(c-1)p^2+(A+b-2Ac)p+Ac=0. \quad (8)$$

In fact

$$p^*=[(2Ac-b-A)+\{(b+A)^2-4Abc\}^{\frac{1}{2}}]/[2A(c-1)]. \quad (9)$$

Now the BDF for testing a simple two sided hypothesis

$$D(\boldsymbol{x})=\begin{cases} d_0 \text{ if } \rho(d_0)<\rho(d_1), \\ d_1 \text{ otherwise,} \end{cases} \quad (9)$$

may be reinterpreted in terms of the subjective a priori probability p in favour of H_0 as

$$D(\boldsymbol{x})=\begin{cases} d_0 \text{ if } p>p^* \\ d_1 \text{ otherwise.} \end{cases} \quad (10)$$

Zone of Sensitivity (ZS)

The normal theory critical point $p_0{}^*$ is obtained by making $\lambda_3=\lambda_4=0$ in (9). We get

$$p_0{}^*=[2A_0c_0-b-A_0+\{(b+A_0)^2-4A_0bc_0\}^{\frac{1}{2}}]/[2A_0(c_0-1)] \quad (11)$$

with $A_0=-a[\tau'+(\mu'-\theta_0)^2]$, $\mu'=(\tau\mu+nr\bar{x})\tau'$,

$$c_0=\sqrt{\tau\tau'}\exp[-12nr\{\tau\tau'(\bar{x}-\mu)^2-(\bar{x}-\theta_0)^2\}] \quad (12)$$

Write

$$\min(p^*, p_0{}^*)=p_1 \text{ and } \max\ (p^*, p_0{}^*)=p_2.$$

Now any $p\in(p_1, p_2)$ will lead to an incorrect decision and it will be due to the wrong assumption that the true conditional distribution of Θ ($\neq\theta_0$) is normal.

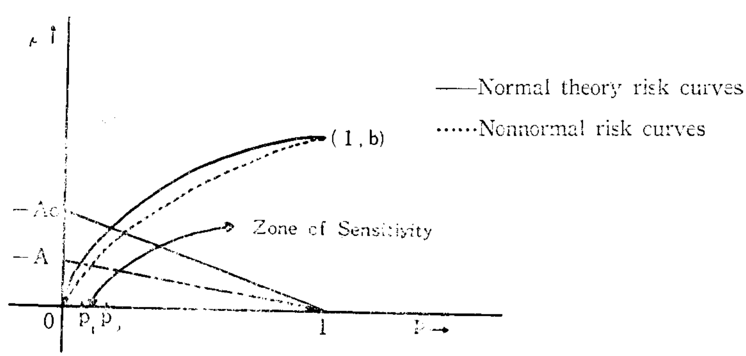

Fig. The Zone of Sensitivity

We shall call the interval (p_1, p_2) as zone of sensitivity (to nonnormality), as illustrated in the figure above, for the BDF to accept/reject the null hypothesis against the two sided alternative. Further the BDF will be robust in the complementary region

$$R_x(\lambda_3, \lambda_4) = \{p : p \in (0, p_1) \cup (p_2, 1)\} \quad (13)$$

This concept of ZS is relevant for investigating the effect of nonnormality on BDF in Jeffreys' framework. Any change in the extent of ZS reflects nonnormality in the same way as in criterion robustness of t- and F- tests where we are interested in the effect of nonnormality on significance level.

Illustration

Consider an observed random sample of size 5 from N(0, 1) given as

$$x = (1.9619,\ -1.6159,\ 2.1711,\ -1.4922,\ 2.2688).$$

For $a=b=r=\tau=1$, $\mu=\theta_0=0$, $\lambda_3=0.4$ and $\lambda_4=1.0$ we find $\rho(d_0)=0.4269(1-p)$ and $\rho(d_1)=p/[p+0.5543(1-p)]$, and these two curves intersect at $p^*=0.1910$. The normal theory risk curves are found to be $\rho(d_0)=0.4670(1-p)$ and $\rho(d_1)=p/[p+0.5855(1-p)]$ and they intersect for $p_0^*=0.2426$. Under normality assumption, the BDF will choose decision d_1 (reject H_0) for any choice of $p<0.2426$, whereas, ESD prior leads to the choice of d_0 (accept H_0) for all $p>0.1910$. The interval (0.1910, 0.2426) is the ZS for the BDF and the region of robustness is the set $R_x(0.4, 1.0)=\{p : p \in (0, 0.1910) \cup (0.2426, 1.0\}$.

Earlier studies of the author (Bansal, 1978a) concerning asymptotic behaviour of the Bayes estimate and the risk for increasing sample size and vagueness of the prior suggest that ZS will shrink in size, i.e. $p^* \to p_0^*$ as $n \to \infty$ and $\tau^{-1} \to \infty$.

3. Sample from Nonnormal Universe

Let us consider a random sample of size n from the moderately nonnormal population represented by the ESD given in (2). As before, assume that its mean θ is unknown but the precision r, λ_3 and λ_4 are known. After neglecting terms containing products and higher powers of λ_3 and λ_4 other than λ_3^2, the likelihood function for a sample of size n, given $\theta=\theta_0$ is

$$f_n(\boldsymbol{x}|\theta_0)=(r/2\pi)^{\frac{n}{2}}\exp[-\frac{1}{2}r\sum_{i=1}^{n}(x_i-\theta_0)^2]\times[1+\frac{1}{6}\lambda_3(s_3-s_1)$$

$$+\frac{1}{24}\lambda_4(s_4-6s_2+3n)+\frac{1}{72}\lambda_3{}^2(s_3{}^2-9s_4-6s_3s_1$$

$$+36s_2+9s_1{}^2-15n)]. \tag{14}$$

with $s_k=\sum_{i=1}^{n}[\sqrt{r}(x_i-\theta_0)]^k,\ k=1, 2, 3, 4.$

As in the previous section, consider mixed type prior distribution of the unknown mean such that $P[\Theta=\theta_0]=p>0$ but the conditional density of $\Theta(\neq\theta_0)$ as normal with known mean μ and precision τ. The risk in taking decision d_0 will be given by (3) where $n, r, \delta^*(x)$, $\rho_x{}^*(\xi)$ and G are T, $r, \hat{x}_s(\mathrm{T})$, $\rho_s{}^*$ and G_s respectively of Bansal (1978b). The risk in taking decision d_1 will be given by (4) with a new

$$c=\sqrt{\tau\tau' G}\ \exp[-\frac{1}{2}nr\{\tau\tau'(\bar{x}-\mu)^2-(\bar{x}-\theta_0)^2\}]\times[1+\frac{1}{6}\lambda_3(s_3-s_1)$$

$$+\frac{1}{24}\lambda_4(s_4-6s_2+3n)+\frac{1}{72}\lambda_3{}^2(s_3{}^2-9s_4-6s_3s_1+36s_2+9s_1{}^2-15n)].$$

Now the BDF and the ZS are defined in the same manner as those of the previous section.

Illustration

Let us reconsider the earlier example (discussed in section 2). Suppose that that the conditional prior $\xi(\theta)$ is standard normal but the sample is suspected to have come from ESD parent with nonnormality parameter $(\lambda_3,\ \lambda_4)=(0.3,\ 0.5)$. For the same constants $a, b, \tau, r, \mu, \theta_0$ and thhe observed sample, we find that the two risk curves intersect for $p^*=0.1020$. Thus $D(\boldsymbol{x})=d_0$ only if $p>0.1020$, otherwise decision d_1. As before, in this case, ZS becomes $(0.1020,\ 0.2426)$ and

$$R_x(0.3,\ 0.5)=\{p:\ p\in(0, 0.1020)\cup(0.2426,\ 1.0)\}.$$

It is interesting to observe that the extent of ZS is larger than that with nonnormal conditional prior even though the corresponding values of λ_3 and λ_4 are smaller. Further, on the basis of earlier study by the author (Bansal, 1978b), we may conclude that the BDF with nonnormal samples will depend heavily on the actually observed sample. Thus, in this case, the effect of nonnormality may not be predicted until the sample is obtained.

In the light of the numerical illustrations and discussions, we find that the zone of sensitivity shall serve to forewarn the decision maker against any indiscriminate choice of the prior probability in favour of the simple null hypothesis.

REFERENCES

[1] Bansal, A.K., "Robustness of Bayes Estimator for the Mean of Normal Population with Nonnormal Prior," *Communications in Statistics* 7A(5), (1978a) 453-60.

[2] ________, "Robustness of Bayes Forecast to Nonnormality," *The Journal of the Korean Statistical Society*, VII (1), (1978b), 11-16.

[3] Box, G.E.P. and Tiao, G.C. "A Further Look at Robustness via Bayes Theorem," *Biometrika* 49, (1962), 419-32.

[4] Jeffreys, H., *Theory of Probability*, Oxford University Press, (1961).