A MINIMAX REGRET ESTIMATOR OF THE NORMAL MEAN WITH UNKNOWN VARIANCE AFTER PRELIMINARY TEST

Rep. Fac. Sci., Kagoshima Univ., (Math., Phys. & Chem.),
No. 24, p. 1–7, 1991.

A MINIMAX REGRET ESTIMATOR OF THE NORMAL MEAN WITH UNKNOWN VARIANCE AFTER PRELIMINARY TEST

Koichi INADA*

(Received September 10, 1991)

Abstract

For the problem of estimating the normal mean μ based on a random sample $X_1,\dots,X_n$ with known variance σ^2 when the mean μ is expected to lie near a prior value μ_0, a class of estimators $\hat{\mu}(k) = k(U)\bar{X} + (1 - k(U))\mu_0$ is considered, where $\bar{X}$ is the sample mean, $U = \sqrt{n}(\bar{X} - \mu_0)/\sigma$ and k is a weight function. Various choices of k have been considered to search the estimator optimal in some sence. Inada [2] combined the idea of preliminary test estimator's and shrinkage estimator's and proposed $\hat{\mu}(k)$ with $k(U) = w^*I(|U| < C) + I(|U| \geq C)$, where for fixed C, $w^* \in [0, 1]$ is chosen by a minimax regret criterion. In this paper we investigate Inada's minimax regret estimator when the variance σ^2 is unknown. The existence of a minimax regret weight is proved and numerical considerations are given.

1. Introduction

Let $X_1,\dots,X_n$ be a random sample of size n from a normal population with unknown mean μ and known variance σ^2. For the problem of estimating the normal mean μ when μ is expected to lie near a prior value μ_0, a class of estimators

$$\hat{\mu}(k) = k(U)\bar{X} + (1 - k(U))\mu_0 \tag{1}$$

is considered, when $\bar{X}$ is the sample mean, $U = \sqrt{n}(\bar{X} - \mu_0)/\sigma$ and k is a weight function. For certain choice of k, $\hat{\mu}(k)$ coincides with previously studied preliminary test and shrinkage estimators. Hirano [3] studied a special type of preliminary estimator, by taking k in (1) as $k_1(U) = I(|U| \geq z_{\frac{\alpha}{2}})$, where $z_{\frac{\alpha}{2}}$ is the $100 \times \frac{\alpha}{2}$ percentage point of N(0, 1). He applied Akaike's [1] information criterion to determine the optimal level of significance for the preliminary test. Thompson [5] proposed a shrinkage estimator of the form (1) with $k_2(U) = U^2/(1 + U^2)$. Mehta and Srinivasan [4] considered $k_3(U) = 1 - ae^{-bU^2}$, where a and b are adjustable constants. Inada [2] combined the idea of preliminary test estimator's and shrinkage estimator's and proposed (1) with $k_4(U) = w^*I(|U| < C) + I(|U| \geq C)$, where for fixed C, $w^* \in [0, 1]$ is chosen by a minimax regret criterion. Looking at these estimators, one observes that they

* Department of Mathematics, Faculty of Science, Kagoshima University, Kagoshima, 890, Japan.

search the estimator optimal in some sence.

In this paper we investigate Inada's minimax regret estimator when the variance σ^2 is nuknown. In this case we consider the estimator

$$\hat{\hat{\mu}}(k) = k(T)\bar{X} + (1 - k(T))\mu_0 \tag{2}$$

where $k(T) = wI(|T| < C) + I(|T| \geq C)$, $T = \sqrt{n}(\bar{X} - \mu_0)/\sqrt{\Sigma(X_i - \bar{X})^2/(n-1)}$ and w is a real number such that $0 \leq w \leq 1$. From the point of view of the minimax regret criterion we shall prove the existence of a minimax regret weight w^* in section 2 and numerical values of minimax regret weight are given in section 3.

2. A minimax regret estimate of a normal mean with unkown variance

Let $X_1, \ldots, X_n$ be a random sample of size n from a normal population with unknown mean μ and variance σ^2. We consider the estimator $\hat{\hat{\mu}}(k)$ given in (2) and prove the existence of w, satisfying a minimax regret criterion. Since a normal distribution is invariant under translations of the mean, we may assume without loss of generality, that $\mu_0 = 0$. We denote the mean squared error of $\hat{\hat{\mu}}(k)$ as $M(w, \mu, \sigma, n)$. Taking $M(w, \mu, \sigma, n)$ as a risk function, a regret function can be written as

$$Reg(w, \mu, \sigma, n) = M(w, \mu, \sigma, n) - \min_{0 \leq w \leq 1} M(w, \mu, \sigma, n). \tag{3}$$

For computing the regret function, we first evaluate the mean squared error of $\hat{\hat{\mu}}(k)$.

$$M(w, \mu, \sigma, n) = \iint_{\frac{\sqrt{n}|\bar{x}|}{\sqrt{\Sigma(x_i - \bar{x})^2/(n-1)}} < C} (w\bar{x} - \mu)^2 f(\bar{x}) g(s^2)\, d\bar{x}\, ds^2 \tag{4}$$

$$+ \iint_{\frac{\sqrt{n}|\bar{x}|}{\sqrt{\Sigma(x_i - \bar{x})^2/(n-1)}} \geq C} (\bar{x} - \mu)^2 f(\bar{x}) g(s^2)\, d\bar{x}\, ds^2$$

where $f(\bar{x}) = \dfrac{\sqrt{n}}{\sqrt{2\pi}\sigma} exp\left[-\dfrac{n(\bar{x} - \mu)^2}{2\sigma^2} w\right]$, $g(s^2) = \dfrac{1}{2\Gamma(\frac{n-1}{2})}\left[\dfrac{(n-1)s^2}{2\sigma^2}\right]^{\frac{n-1}{2}-1}$ $exp\left[-\dfrac{(n-1)s^2}{2\sigma^2}\right]\left(\dfrac{n-1}{\sigma^2}\right)$ and $s^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$.

The transformation defined by $u = \sqrt{n}(\bar{x} - \mu)/\sigma$ and $v = \dfrac{(n-1)s^2}{2\sigma^2}$, will yield us

$$M(w, \mu, \sigma, n) = \frac{\sigma^2}{n} \Psi(w, \delta, n) \tag{5}$$

where

$$\Psi(w, \delta, n) = \iint_{(n-1)C^2 v > 2(u+\delta)^2} \{w(u + \delta) - \delta\}^2 \phi(u)k(v)\, du dv \tag{6}$$

$$+ \iint_{(n-1)C^2 v \leq 2(u+\delta)^2} u^2 \phi(u)k(v)\, du dv,$$

$\phi(u) = exp[-v^2/2]/\sqrt{2\pi}$, $k(v) = v^{\frac{n-1}{2}-1} exp[-v]/\Gamma(\frac{n-1}{2})$ and $\delta = \sqrt{n}\,\mu/\sigma$.

The regret function $Reg(w, \mu, \sigma, n)$ can be rewritten as

$$Reg(w, \mu, \sigma, n) = \frac{\sigma^2}{n}\{\Psi(w, \delta, n) - \min_{0 \leq w \leq 1} \Psi(w, \delta, n)\}. \tag{7}$$

$$= \frac{\sigma^2}{n} R(w, \delta, n)$$

This is nothing but the actual mean squared error minus its minimum value, and thus the regret function is the shortcoming in the mean squared error caused by the absence of knowledge on μ/σ. Then the minimax regret criterion leads to the minimax regret weight w^*, which attains

$$\inf_{0 \leq w \leq 1} \sup_{-\infty < \delta < \infty} R(w, \delta, n).$$

Let $w_0(\delta)$ and $w_1(\delta)$ be the function of δ defined by

$$\min_{-\infty < w < \infty} \Psi(w, \delta, n) = \Psi(w_0(\delta), \delta, n), \tag{8}$$

$$\min_{0 \leq w \leq 1} \Psi(w, \delta, n) = \Psi(w_1(\delta), \delta, n). \tag{9}$$

We need Lemma 1 in order to prove Theorem 1.

Lemma 1. *Let* $m(\delta) = x(e^{x\delta} + e^{-x\delta}) - \delta(e^{x\delta} - e^{-x\delta})$, *then we have for a positive number* x

$$m(\delta) = m(-\delta), \; m(0) > 0 \tag{10}$$

and there exists a positive number δ_1 *such that*

$$\begin{aligned} m(\delta) &> 0 \quad \text{if } |\delta| < \delta_1, \\ m(\delta) &= 0 \quad \text{if } |\delta| = \delta_1, \\ m(\delta) &< 0 \quad \text{if } |\delta| > \delta_1. \end{aligned} \tag{11}$$

Proof of Lemma 1. Proof of (10) is clear. So we suppose without loss of generality that $\delta > 0$. $m(\delta)$ can be written as

$$m(\delta) = e^{-x\delta}\{g_1(\delta) - g_2(\delta)\}$$

where $g_1(\delta) = \delta + x$ and $g_2(\delta) = (\delta - x)e^{2x\delta}$. As $\frac{d}{d\delta}g_2(\delta) = 2xe^{2x\delta}\left(\delta - \frac{2x^2-1}{2x}\right)$, when $0 < x < \frac{\sqrt{2}}{2}$, $g_2(\delta)$ is strictly increasing in $\delta > 0$ and when $x \geq \frac{\sqrt{2}}{2}$, $g_2(\delta)$ is strictly decreasing in $0 < \delta < \frac{2x^2-1}{2x} < x$ and strictly increasing in $\delta > \frac{2x^2-1}{2x}$. And as $m(0) = m(x) = 2x > 0$, the straight line of $g_1(\delta)$ and the curve of $g_2(\delta)$ cross at a single point δ_1. This completes the proof of lemma 1.

Theorem 1. *The function $w_0(\delta)$ is given by*

$$w_0(\delta) = \frac{\delta \iint_{(n-1)C^2v > 2(u+\delta)^2}(u+\delta)\phi(u)k(v)\,du\,dv}{\iint_{(n-1)C^2v > 2(u+\delta)^2}(u+\delta)^2\phi(u)k(v)\,du\,dv} \tag{12}$$

which is a non-negative even function satisfying $w_0(\delta) = 0$ if and only if $\delta = 0$ and there exists a positive number δ_1 such that

$$\begin{aligned} w_0(\delta) &< 1 \quad \text{if } |\delta| < \delta_1, \\ w_0(\delta) &= 1 \quad \text{if } |\delta| = \delta_1, \\ w_0(\delta) &> 1 \quad \text{if } |\delta| > \delta_1. \end{aligned} \tag{13}$$

Proof of Theorem 1. As the function $\Psi(w, \delta, n)$ is quadratic in w with the positive coefficient to w^2, the minimum is evidently attained at the value given in (12). Let $w_0(\delta) = \delta g(\delta)/h(\delta)$. After the transformation, $u + \delta = x$, we have

$$g(\delta) = \iint_{(n-1)C^2v > x^2} x\phi(x-\delta)k(v)\,dx\,dv \tag{14}$$

$$= e^{-\frac{\delta^2}{2}}\int_0^\infty \int_{\frac{x}{C\sqrt{n-1}}}^\infty xe^{-\frac{x^2}{2}}(e^{x\delta} - e^{-x\delta})k(v)\,dv\,dx.$$

Simiraly we have

$$h(\delta) = e^{-\frac{\delta^2}{2}}\int_0^\infty \int_{\frac{x}{C\sqrt{n-1}}}^\infty x^2e^{-\frac{x^2}{2}}(e^{x\delta} + e^{-x\delta})k(v)\,dv\,dx. \tag{15}$$

Therefore it is easy to check $g(\delta) = -g(-\delta)$ and $h(\delta) = h(-\delta)$. Since $w_0(\delta) = w_0(-\delta)$, we suppose without loss of generality that $\delta > 0$. (14) imply that $g(\delta)$ is positive, and hence $w_0(\delta)$ is positive except for $\delta = 0$. Now $w_0(\delta) < 1$, $= 1$ and > 1 is equivalent to $h(\delta) - \delta g(\delta) > 0$, $= 0$ and < 0, respectively. On the other hand it turns out

$$h(\delta) - \delta g(\delta) = e^{-\frac{\delta^2}{2}} \int_0^\infty \int_{\frac{x}{C\sqrt{n-1}}}^\infty x e^{-\frac{x^2}{2}} m(\delta) k(v)\, dv\, dx$$

and because of Lemma 1, the existence of a positive number δ_1 is concluded.

We thus have

$$R(w, \delta, n) = \Psi(w, \delta, n) - \Psi(w_1(\delta), \delta, n) \tag{16}$$

where $\delta = \sqrt{n}\mu/\sigma$ and

$$w_1(\delta) = \begin{cases} w_0(\delta) & \text{if } |\delta| \leq \delta_1 \\ 1 & \text{if } |\delta| > \delta_1 \end{cases}$$

and $w_0(\delta)$ and δ_1 are given in Theorem 1.

Theorem 2. *The minimax regret weight exists.*

Proof of Theorem 2. As $\Psi(w, \delta, n)$ and $w_1(\delta)$ are continuous, $R(w, \delta, n) = \Psi(w, \delta, n) - \Psi(w_1(\delta), \delta, n)$ is continuous both in w and δ. Further examination of (6) indicates

$$\lim_{\delta \to \pm\infty} R(w, \delta, n) = 0.$$

Therefore the supremum of $R(w, \delta, n)$ in $\delta \in (-\infty, \infty)$ is attained at a finite value of δ, or equivalently there is a function $\delta(w)$ with

$$\sup_{-\infty < \delta < \infty} R(w, \delta, n) = R(w, \delta(w), n)$$

and $\delta(w) \in (-\infty, \infty)$. And $R(w, \delta(w), n)$ is a lower semicontinous function of w defined on [0, 1]. As [0, 1] is compact, the infimum of $R(w, \delta(w), n)$ is attained at a point in [0, 1].

3. Numerical considerations

To obtain the value of the minimax regret weight, we must calculate $\Psi(w, \delta, n)$. The value of this quantity can be evaluated by the expansion (17), (18) and (19).

$$\iint_{(n-1)C^2 v > 2(u+\delta)^2} u^2 \phi(u) k(v)\, du\, dv = \sum_{i=0}^{\frac{n-3}{2}} \frac{\sqrt{d}}{i!\sqrt{2+d}} e^{-\frac{\delta^2}{2+d}} (d\mu_{2i+2} - 2\sqrt{d}\,\delta\mu_{2i+1} + \delta^2 \mu_{2i}), \tag{17}$$

$$\iint_{(n-1)C^2 v > 2(u+\delta)^2} u \phi(u) k(v)\, du\, dv = \sum_{i=0}^{\frac{n-3}{2}} \frac{\sqrt{d}}{i!\sqrt{2+d}} e^{-\frac{\delta^2}{2+d}} (\sqrt{d}\,\mu_{2i+1} - \delta\mu_{2i}) \tag{18}$$

and

$$\iint_{(n-1)C^2 v > 2(u+\delta)^2} \phi(u)k(v)\,du\,dv = \sum_{i=0}^{\frac{n-3}{2}} \frac{\sqrt{d}}{i!\sqrt{2+d}} e^{-\frac{\delta^2}{2+d}} \mu_{2i} \tag{19}$$

where

$$\mu_k = \int_{-\infty}^{\infty} t^k \frac{1}{\sqrt{2\pi}\tilde{\sigma}} e^{-\frac{(t-\tilde{\mu})^2}{2\tilde{\sigma}^2}} dt,\ \tilde{\mu} = \frac{\sqrt{d}}{2+d},\ \tilde{\sigma} = \frac{1}{\sqrt{2+d}},\ d = \frac{(n-1)C^2}{2}.$$

These three expansions are based on integrating first in u and then applying the expansion (20)

$$\int_a^{\infty} \frac{1}{\Gamma(p)} v^{p-1} e^{-v} dv = \sum_{i=0}^{p-1} \frac{1}{i!} a^i e^{-a} \tag{20}$$

where p is a positive integer. Finally from (17), (18) and (19) we arrive at a simple expression of $\Psi(w, \delta, n)$.

$$\Psi(w, \delta, n) = 1 + \sum_{i=0}^{\frac{n-3}{2}} \frac{\sqrt{d}}{i!\sqrt{2+d}} e^{-\frac{\delta^2}{2+d}} \{d(w^2-1)\mu_{2i+2} + 2\sqrt{d}\,\delta(1-w)\mu_{2i+1}\} \tag{21}$$

After rather extensive numerical examinations, values of the minimax regret weight w^* were obtained for $C = \sqrt{(n-1)(e^{\frac{2}{n}}-1)}$, which was given by Hirano by applying Akaike's information criterion, and are given in Table 1.

Table 1. Values of the minimax regret weight.

n	5	7	9	11	13	15
w^*	0.6461	0.6444	0.6434	0.6428	0.6423	0.6420

n	17	19	21	23	25
w^*	0.6417	0.6415	0.6413	0.6412	0.6411

References

[1] Akaike, H. (1973). Information theory and an extention of the maximum likelihood principle, *2nd International Symposium on Information Theory* (eds. B.N.Petrov and F.Csaki), Akademiai Kiado, Budapest, 267–281.

[2] Inada, K. (1984). A minimax regret estimator of a normal mean after preliminary test. *Ann. Inst. Staist. Math.*, **36**, A, 207–215.

[3] Hirano, K. (1977). Estimation procedure based on preliminary test, shrinkage technique and information criterion. *Ann. Inst. Statist. Math.*, **29**, A, 21–34.

[4] Mehta, J. S. and Srinivasan, R. (1971). Estimation of the mean by shrinkage to a point. *J. Amer. Statist. Assoc.*, **66**, 86–90.

[5] Thompson, J. R. (1968). Some shrinkage techniques for estimating the mean. *J. Amer. Statist. Assoc.*, **63**, 113–122.