Kumamoto J. Math.
Vol. 12, 81-89 March (1999)

Existence of estimators with bounded risks under an asymmetric loss function

Yoshikazu Takada

(Received February 1, 1999)

1. Introduction

Let $\{X_i; i\geq 1\}$ be independent and identically distributed (i.i.d.) random variables with the probability density function $\sigma^{-1}\left(\frac{x-\mu}{\sigma}\right)$ with respect to Lebesgue measure, where f is known and $\theta=(\mu, \sigma)$ $(\sigma>0)$ is unknown. We consider estimating μ under the loss function

$$L(\theta, d)=\rho(d-\mu),$$

where $\rho(u)$ is nonincreasing for $u<0$, nondecreasing for $u>0$ and $\rho(0)=0$. Let

$$M_{min}=\min(m_-, m_+), \quad M_{max}=\max(m_-, m_+),$$

where $m_-=\lim_{u\to-\infty}\rho(u)$ and $m_+=\lim_{u\to\infty}\rho(u)$, which may be infinity. Let $\delta(X)$ be an estimator based on $X=(X_1, \cdots, X_n)$. Then the purpose of this paper is to find if there exists an estimator $\delta(X)$ such that

$$E_\theta\{L(\theta, \delta(X))\}\leq W, \tag{1.1}$$

for all θ, where $W(>0)$ is an given constant.

Under a symmetric loss function Lehmann [3] showed that if the sample size n is predetermined, then no estimators meet (1.1) for $W<M_{min}(=M_{max})$. In Section 2 we shall show that the nonexistence is also true for an asymmetric loss function if $W<M_{min}$, but there does exist such an estimator if $M_{min}<W<M_{max}<\infty$. However, we can not show if the result also holds when $M_{max}=\infty$. In Section 3 employing two-stage procedures by Stein [6], we shall show that it is possible to construct an estimator meeting the requirement (1.1) for some distributions. Section 4 is devoted to the normal distribution and discuss the property of the proposed two-stage procedure.

2. Fixed sample size estimators

In this section we assume that the sample size n is predetermined. Then we get the following results.

Theorem 1. *If* $0<W<M_{min}$, *then there do not exist any estimators whose risk is bounded by* W *for all* θ.

Proof. Suppose that an estimator δ satisfies (1.1) for $0<W<M_{min}$. Then for any $d>0$

$$\begin{aligned} W &\geq E_\theta\{\rho(\delta(X)-\mu)\} \\ &\geq \rho(d)P_\theta(\delta(X)-\mu\geq d)+\rho(-d)P_\theta(\delta(X)-\mu\leq -d) \\ &\geq \min(\rho(d),\rho(-d))P_\theta(|\delta(X)-\mu|\geq d). \end{aligned}$$

Since $\lim_{d\to\infty}\min(\rho(d),\rho(-d))=M_{min}$, there exists d such that $0<\alpha=W/\min(\rho(d),\rho(-d))<1$. Hence

$$P_\theta(|\delta(X)-\mu|<d)>1-\alpha,$$

which contradicts the fact that there do not exist fixed-width confidence intervals of μ when the sample size is fixed (e.g. Lehmann [3] and Takada [7]). Hence the proof is completed.

Theorem 2. *If* $M_{min}<W<M_{max}<\infty$, *then there exists an estimator whose risk is bounded by* W *for all* θ.

Proof. Without loss of generality, let us assume $m_-<m_+$. Let $\hat{\mu}(X)$ and $\hat{\sigma}(X)(>0)$ be any statistics satisfying

$$\hat{\mu}(aX+b)=a\hat{\mu}(X)+b,$$
$$\hat{\sigma}(aX+b)=a\hat{\sigma}(X)$$

for all $a(>0)$ and b, where $aX+b=(aX_1+b,\cdots,aX_n+b)$. For example, $\hat{\mu}$ is the sample mean and $\hat{\sigma}$ the sample standard deviation. Let $\delta(X)=\hat{\mu}(X)+t\hat{\sigma}(X)$. Then

$$\begin{aligned}E_\theta\{L(\theta, \delta(X))\} &= E_\theta\{\rho(\hat{\mu}(X)+t\hat{\sigma}(X)-\mu)\}\\ &= E_{(0,1)}\{\rho(\hat{\mu}(\mu+\sigma X)+t\hat{\sigma}(\mu+\sigma X)-\mu)\}\\ &= E_{(0,1)}\{\rho(\sigma(\hat{\mu}(X)+t\hat{\sigma}(X)))\}\\ &\le m_+P_{(0,1)}(\hat{\mu}(X)+t\hat{\sigma}(X)\ge 0)\\ &\quad + m_-P_{(0,1)}(\hat{\mu}(X)+t\hat{\sigma}(X)<0)\\ &= m_- + (m_+ - m_-)P_{(0,1)}(\hat{\mu}(X)+t\hat{\sigma}(X)\ge 0).\end{aligned}$$

Let $\alpha=(W-m_-)/(m_+-m_-)$. Then $0<\alpha<1$. So there exists t such that

$$P_{(0,1)}(\hat{\mu}(X)+t\hat{\sigma}(X)\ge 0)=\alpha.$$

Hence we get

$$\begin{aligned}E_\theta\{L(\theta, \delta(X))\} &\le m_- + (m_+-m_-)\alpha\\ &= W,\end{aligned}$$

which completes the proof.

In the proof we used the condition that $M_{max}<\infty$. We do not know if Theorem 2 holds without the condition. But the next example suggests that such an estimator may exist for some problems.

Example 1. Let $X_1, \cdots, X_n$ be i.i.d. according to the exponential distribution $(E(\mu, \sigma))$ with density $\sigma^{-1}\exp(-(x-\mu)/\sigma)$, $x>\mu$. We want to estimate μ under the loss function

$$\rho(u)=\begin{cases} 1, & \text{for } u>0\\ -u, & \text{for } u\le 0.\end{cases}$$

Hence $M_{min}=1$ and $M_{max}=\infty$. Let $\delta(X)=\min(X_1, \cdots, X_n)$. Then the risk function of δ is one since $\delta(X)>\mu$, so that for any $W(1<W<\infty)$

$$E_\theta\{L(\theta, \delta(X))\}\le W.$$

3. Two-stage estimation procedures

In this section we shall show that utilizing a two-stage procedure by Stein [6] may enable us to construct an estimator with bounded risk in some cases when there do not exist fixed sample size estimators.

For each n, let $X_n=(X_1, \cdots, X_n)$. We suppose that there exist two sequences of statistics $\{t_n(X_n);\ n\ge n_0\}$ and $\{s_n(X_n);\ n\ge n_0\}$ $(s_n(X_n)>0)$ which satisfy the following assumptions.

Assumption 1. For any $a(>0)$, b and $n \geq n_0$

$$t_n(aX_n+b)=at_n(X_n)+b,$$
$$s_n(aX_n+b)=as_n(X_n).$$

Assumption 2. There exists a positive constant β such that

$$P_{(0,1)}(n^{\beta}t_n(X_n) \leq x)=G(x)$$

is independent of n.

Assumption 3. There exists a positive integer $m(\geq n_0)$ such that $t_n(X_n)$ is independent of $s_m(X_m)$ for $n \geq m$.

The similar assumptions are used in Ghurye [2] who considered an application of Stein's two-stage procedure of testing for a location parameter of the location-scale family. See also Mukhopadhyay [5].

Let $S_m=s_m(X_m)$ and we assume that for each $z>0$

$$h_m(z)=\iint \rho\left(\frac{zx}{y}\right)G(dx)H_m(dy)<\infty, \tag{3.1}$$

where $H_m(y)=P_{(0,1)}(S_m \leq y)$. Then $h_m(z)$ is a nondecreasing function of $z>0$ and $h_m(0)=0$,

$$\lim_{z\to\infty} h_m(z)=m_+\int_{x\geq 0}G(dx)+m_-\int_{x<0}G(dx)$$
$$=\alpha m_+ +(1-\alpha)m_-$$

with $\alpha=\int_{x\geq 0}G(dx)$. Hence if $W<\alpha m_+ +(1-\alpha)m_-$, then there exists a positive constant z_m such that $h_m(z_m)=W$.

Now we shall give a two-stage estimator whose risk is bounded by W. Let $X_m=(X_1, \cdots, X_m)$ be the first sample and calculate S_m. Define the total sample size N by

$$N=\max\left\{m, \left[\left(\frac{S_m}{z_m}\right)^{1/\beta}\right]+1\right\}, \tag{3.2}$$

where $[u]$ denotes the largest integer less than u. If $N>m$, take the second sample $X_{m+1}, \cdots, X_N$. Then μ is estimated by $T_N=t_N(X_N)$.

Theorem 3. *If Assumptions 1 to 3 are satisfied and* $W<\alpha m_+ +(1-\alpha)m_-$, *then the estimator* T_N *meets the requirement.*

Proof. It follows from Assumption 3 that

$$\begin{aligned} E_\theta\{L(\theta, T_N)\} &= E_\theta\{\rho(T_N-\mu)\} \\ &= \sum_{n=m}^{\infty} E_\theta\{I_{(N=n)}\rho(t_n(X_n)-\mu)\} \\ &= \sum_{n=m}^{\infty} P_\theta(N=n)E_\theta\{\rho(t_n(X_n)-\mu)\}, \end{aligned}$$

where $I_{(N=n)}$ denotes the indicator function of the set $\{N=n\}$. Then from Assumptions 1 and 2 and the definition of N we get

$$\begin{aligned} E_\theta\{L(\theta, t_N(X_N))\} &= \sum_{n=m}^{\infty} P_\theta(N=n)E_{(0,1)}\{\rho(\sigma t_n(X_n))\} \\ &= \sum_{n=m}^{\infty} P_\theta(N=n)\int\rho\Big(\frac{\sigma x}{n^\beta}\Big)G(dx) \\ &= E_\theta\Big\{\int\rho\Big(\frac{\sigma x}{N^\beta}\Big)G(dx)\Big\} \\ &\le E_\theta\Big\{\int\rho\Big(\frac{z_m\sigma x}{S_m}\Big)G(dx)\Big\} \\ &= h_m(z_m)=W. \end{aligned}$$

Hence the proof is completed.

Example 2. Let $\{X_i, i\ge 1\}$ be i.i.d. according to the exponential distribution $E(\mu, \sigma)$, and let $t_n(X_n)=\min(X_1, \cdots, X_n)$ and $s_n(X_n)=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-t_n(X_n))$. Then $t_n(X_n)$ and $s_n(X_n)$ are independently distributed as $E(\mu, \sigma/n)$ and $\frac{1}{2(n-1)}\sigma\chi^2_{2n-2}$, respectively, where χ^2_{2n-2} denotes the chi squared distribution with $2n-2$ degrees of freedom. It is easy to see that Assumptions 1 to 3 are satisfied with $\beta=1$, and that (3.1) is given by

$$h_m(z)=\int_0^\infty \rho\Big(\frac{zu}{m-1}\Big)f_{2,2(m-1)}(u)du,$$

where $f_{2,2(m-1)}(u)$ is the density of F distribution with $(2, 2(m-1))$ degrees of freedom. So we can get a two-stage estimator $\min(X_1, \cdots, X_N)$ whose risk is bounded by $W(<m_+)$. The total sample size N by (3.2) is

$$N=\max\Big\{m, \Big[\frac{S_m}{z_m}\Big]+1\Big\},$$

wher $h_m(z_m)=W$ and $S_m=s_m(X_m)$.

4. Normal distribution

Let $\{X_i, i\geq 1\}$ be i.i.d. according to the normal distribution with mean μ and variance σ^2, and let $t_n(X_n)=\bar{X}_n \left(=\frac{1}{n}\sum_{i=1}^{n} X_i\right)$, $s_n(X_n)=\sqrt{U_n}$ with $U_n=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2$. Then it is easy to see that Assumptions 1 to 3 are satisfied with $\beta=1/2$, and that (3.1) is given by

$$h_m(z)=\int_{-\infty}^{\infty}\rho(zu)g_{m-1}(u)du,$$

where $g_{m-1}(u)$ is the density of t distribution with $m-1$ degrees of freedom. So we can get a two-stage estimator $\bar{X}_N$ whose risk is bounded by $W(<(m_-+m_+)/2)$. The total sample size N by (3.2) is

$$N=\max\left\{m, \left[\frac{U_m}{z^2_m}\right]+1\right\}, \tag{4.1}$$

where $h_m(z_m)=W$.

Now we consider properties of the two-stage procedure (4.1). Let

$$h(z)=\int_{-\infty}^{\infty}\rho(zu)\phi(u)du,$$

where ϕ is the density function of the standard normal distribution. If σ were known, then

$$E_\theta\rho(\bar{X}_n-\mu)\leq W$$

if and only if $n\geq n^*=(\sigma/z^*)^2$, where $h(z^*)=W$. So n^* would be the optimal sample size if σ were known.

Lemma 1. $z^*\geq z_m$.

Proof. Suppose that X and Y are random variables with the standard normal distribution and t distribution with $\nu(=m-1)$ degrees of freedom, respectively. The it follows from Theorem 4 of Ghosh [1] that

$$P(|X|>c)\leq P(|Y|>c)$$

for any $c>0$. That is, $|Y|$ is stochastically larger than $|X|$. Since $h(z)=E\rho(zX)$ and $h_m(z)=E\rho(zY)$, it can be shown that $h_m(z)\geq h(z)$ (e.g. Lehmann [4], p. 116), from which the theorem is proved.

From (4.1) we get

$$\frac{U_m}{z_m^2} \leq N \leq \frac{U_m}{z_m^2} + m. \tag{4.2}$$

Hence

$$\left(\frac{z^*}{z_m}\right)^2 \frac{U_m}{\sigma^2} \leq \frac{N}{n^*} \leq \left(\frac{z^*}{z_m}\right)^2 \frac{U_m}{\sigma^2} + \frac{mz^{*2}}{\sigma^2}. \tag{4.3}$$

From the left inequality of (4.3) and Lemma 1 we get the following result.

Theorem 4. *For any fixed W and m*

$$\frac{E_\theta(N)}{n^*} \geq \left(\frac{z^*}{z_m}\right)^2 \geq 1.$$

Next we consider the asymptotic properties of the two-stage procedure as $W \to 0$. We call the two-stage procedure asymptotically efficient if

$$\lim_{W \to 0} \frac{E_\theta(N)}{n^*} = 1.$$

It follows from (4.3) that

$$\left(\frac{z^*}{z_m}\right)^2 \leq \frac{E_\theta(N)}{n^*} \leq \left(\frac{z^*}{z_m}\right)^2 + \frac{mz^{*2}}{\sigma^2}.$$

So in order for the two-stage procedure to be asymptotically efficient it is necessary that m must be chosen such that

$$\lim_{W \to 0} \frac{z^*}{z_m} = 1 \tag{4.4}$$

and

$$\lim_{W \to 0} mz^{*2} = 0. \tag{4.5}$$

Then the next result is obtained.

Theorem 5. *If the initial sample size m is chosen such that (4.4) and (4.5) are satisfied, then the two-stage procedure is asymptotically efficient.*

Example 3. Suppose that

$$\rho(u)=\begin{cases} au^2, & \text{for } u>0 \\ -bu, & \text{for } u\leq 0, \end{cases}$$

where $a>0$ and $b>0$. Then straightforward calculations show that

$$h(z)=cz+dz^2$$

$$h_m(z)=c_m z+d_m z^2,$$

where $c=\frac{a}{2}$, $d=\frac{b}{\sqrt{2\pi}}$, $c_m=c\frac{\nu}{\nu-2}$ and $d_m=d\frac{\Gamma\left(\frac{\nu-1}{2}\right)\sqrt{\nu}}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{2}}$ with $\nu=m-1$.

Hence

$$z^*=\frac{-d+\sqrt{d^2+4cW}}{2c}, \quad z_m=\frac{-d_m+\sqrt{d_m^2+4c_m W}}{2c_m}.$$

So

$$\frac{z^*}{z_m}=\frac{\sqrt{d_m^2+4c_m W}+d_m}{\sqrt{d^2+4cW}+d}$$

and

$$z^*=\frac{2W}{\sqrt{d^2+4cW}+d}.$$

Note that $\lim_{m\to\infty}c_m=c$ and $\lim_{m\to\infty}d_m=d$. If m is chosen such that $m=o(W^{-2})$ as $W\to 0$, then (4.4) and (4.5) are satisfied and the two-stage procedure becomes asymptotically efficient.

References

[1] Ghosh, B. K., Some monotonicity theorem for χ^2, F and t distribution with applications, J. Roy. Statist. Soc., B, **35** (1973), 48-492.

[2] Ghurye, S. G., Note on sufficient statistics and two-stage procedures, Ann. Math. Statist. **29** (1958), 155-166.

[3] Lehmann, E. L., Notes on the theory of estimation, University of Californian Press, Berkeley, 1951.

[4] Lehmann, E. L., Testing Statistical Hypotheses, 2nd ed., Wiley, New York, 1986.

[5] Mukhopadhyay, N., Stein's two-stage procedure and exact consistency, Scand. Actuarial J. (1982), 110-122.

[6] Stein, C., A two-sample test for a linear hypothesis whose power is independent of the variance, Ann. Math. Statist., **16** (1945), 245-258.

[7] Takada, Y., The nonexistence of procedures with bounded performance characteristics in certain parametric inference problems, Ann. Inst. Statist. Math., **50** (1998), 325-335.

Yoshikazu Takada
Department of Mathematics
Faculty of Science
Kumamoto University
Kumamoto 860-855, Japan
E-mail: takada@sci. kumamoto-u. ac. jp