Ann. Inst. Statist. Math.
36 (1984), Part A, 207-215

A MINIMAX REGRET ESTIMATOR OF A NORMAL MEAN AFTER PRELIMINARY TEST

KÔICHI INADA

(Received Mar. 19, 1983; revised July 26, 1983)

Summary

This paper considers the problem of estimating a normal mean from the point of view of the estimation after preliminary test of significance. But our point of view is different from the usual one. The difference is interpretation about a null hypothesis. Let $\bar{X}$ denote the sample mean based on a sample of size n from a normal population with unknown mean μ and known variance σ^2. We consider the estimator that assumes the value $\omega\bar{X}$ when $|\bar{X}|<C\sigma/\sqrt{n}$ and the value $\bar{X}$ when $|\bar{X}|\geqq C\sigma/\sqrt{n}$ where ω is a real number such that $0\leqq\omega\leqq 1$ and C is some positive constant. We prove the existence of ω, satisfying the minimax regret criterion and make a numerical comparison among estimators by using the mean square error as a criterion of goodness of estimators.

1. Introduction

The estimation after preliminary test of significance has been studied by various authors. A bibliography in Bancroft and Han [2] is a good source of references. Let $X_1, X_2, \cdots, X_n$ be a random sample of size n from a normal population with unknown mean μ and known variance σ^2. We shall consider the problem of estimating a normal mean μ with known variance σ^2 from the point of view of the estimation after preliminary test of significance. Let α be the significance level for testing a null hypothesis H_0: $\mu=0$ against H_1: $\mu\neq 0$. A preliminary test estimator for μ can be given by

$$(1) \qquad \hat{\mu}=\begin{cases} 0 & \text{if } |\bar{X}|<C\sigma/\sqrt{n} \\ \bar{X} & \text{if } |\bar{X}|\geqq C\sigma/\sqrt{n} \end{cases}$$

where $\bar{X}=\frac{1}{n}\sum_{i=1}^{n} X_i$, C is a positive number and $\Pr(|\bar{X}|\geqq C\sigma/\sqrt{n}\,|H_0)=\alpha$.

The preliminary test estimator always depends on the significance level of the preliminary test. The methods to seek the optimal level of significance for the preliminary test have been investigated by Sawa and Hiromatsu [6], Toyoda and Wallace [7] and Ohtani and Toyoda [5]. Hirano [3] applied AIC (Akaike's Information Criterion [1]) to determine the optimal level of significance for the preliminary test. AIC is defined as $-2\log_e L(\hat{\mu})+2k$, where $L(\hat{\mu})$ is the maximum of the likelihood and k is the number of unknown parameters. By applying AIC to $\hat{\mu}$, $C=\sqrt{2}$ and $\alpha=0.1572\cdots$ were obtained (see Hirano [3]).

By the way, the hypothesis $H_0: \mu=0$ can be considered as an idealized model for the situation that μ is close enough to zero, so in case of $|\bar{X}|<C\sigma/\sqrt{n}$ it may be natural to use an estimator close enough to zero, but not always equal to zero. Thus we propose an estimator as follows:

$$\text{(2)} \qquad \mathrm{PT}(\omega, C)=\begin{cases} \omega\bar{X} & \text{if } |\bar{X}|<C\sigma/\sqrt{n} \\ \bar{X} & \text{if } |\bar{X}|\geqq C\sigma/\sqrt{n} \end{cases}$$

where the weight ω is a real number such that $0\leqq\omega\leqq1$. That is, for a small value of $|\bar{X}|$ we use $\omega\bar{X}$ as our estimate of μ; otherwise, we use the usual estimator $\bar{X}$. This estimator has a merit of never excessively changing in the neighborhood of $\pm C\sigma/\sqrt{n}$. From the point of view of the minimax regret criterion we shall prove the existence of optimal ω, ω^*, in Section 2.

In Section 3 the optimal value of ω, ω^*, is given and we make a numerical comparison among estimator and discuss this property.

If we use the mean square error as a criterion of goodness of estimators, our minimax regret estimator $\mathrm{PT}(\omega^*, C)$ will be appropriate when we know in advance that μ is probably small but we cannot be completely sure of it.

2. A minimax regret estimate of a normal mean

Let $\bar{X}$ be the sample mean based on a sample of size n from a normal population with unknown mean μ and known variance σ^2. We consider the estimation of μ. By the reason stated in Section 1 we consider the estimator of the form:

$$\text{(3)} \qquad \mathrm{PT}(\omega, C)=\begin{cases} \omega\bar{X} & \text{if } |\bar{X}|<C\sigma/\sqrt{n} \\ \bar{X} & \text{if } |\bar{X}|\geqq C\sigma/\sqrt{n} \end{cases}$$

where the weight ω is a real number such that $0\leqq\omega\leqq1$. When $\omega=0$, this estimator $\mathrm{PT}(0, C)$ is the preliminary test estimator which is given by (1) and when $\omega=1$, PT$(1, C)$ is the usual estimator $\bar{X}$. In this section we prove the existence of ω', satisfying the minimax regret criterion. We denote the mean square error of PT(ω, C) as $M_c(\omega, \mu)$ and define the regret function by

$$\text{Reg}_c(\omega, \mu)=M_c(\omega, \mu)-\min_{0\le\omega\le1} M_c(\omega, \mu) . \tag{4}$$

Then the minimax regret criterion leads to the minimax regret weight ω^*, which attains the infimum of $\sup_{-\infty<\mu<\infty} \text{Reg}_c(\omega, \mu)$.

At first we evaluate the mean square error of PT(ω, C).

$$\begin{aligned} M_c(\omega, \mu)&=\int_{|\bar{x}|<C\sigma/\sqrt{n}} (\omega\bar{x}-\mu)^2\frac{\sqrt{n}}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2\right]d\bar{x} \\ &\quad+\int_{|\bar{x}|\ge C\sigma/\sqrt{n}} (\bar{x}-\mu)^2\frac{\sqrt{n}}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2\right]d\bar{x} \\ &=\frac{\sigma^2}{n}\left[\int_{-C-\theta}^{C-\theta} \{\omega(t+\theta)-\theta\}^2\phi(t)dt+\int_{-\infty}^{-C-\theta} t^2\phi(t)dt \right. \\ &\quad\left.+\int_{C-\theta}^{\infty} t^2\phi(t)dt\right] \\ &=\frac{\sigma^2}{n}\Psi_c(\omega, \theta) \end{aligned} \tag{5}$$

where $\theta=\sqrt{n}\,\mu/\sigma$, $\phi(x)=\frac{1}{\sqrt{2\pi}}\exp(-x^2/2)$, $\Phi(x)=\int_{-\infty}^{x}\phi(t)dt$ and

$$\begin{aligned} \Psi_c(\omega, \theta)=1&+(\omega^2-1)\{(-C-\theta)\phi(-C-\theta)-(C-\theta)\phi(C-\theta) \\ &+\Phi(C-\theta)-\Phi(-C-\theta)\}-2\omega(1-\omega)\theta \\ &\cdot\{\phi(-C-\theta)-\phi(C-\theta)\}+(1-\omega)^2\theta^2 \\ &\cdot\{\Phi(C-\theta)-\Phi(-C-\theta)\} . \end{aligned} \tag{6}$$

We observe that minimizing $M_c(\omega, \mu)$ is equivalent to minimizing $\Psi_c(\omega, \theta)$.

Let $\omega_0(\theta)$ and $\omega_1(\theta)$ be the function of θ defined by

$$\min_{-\infty<\omega<\infty} \Psi_c(\omega, \theta)=\Psi_c(\omega_0(\theta), \theta) , \tag{7}$$

$$\min_{0\le\omega\le1} \Psi_c(\omega, \theta)=\Psi_c(\omega_1(\theta), \theta) . \tag{8}$$

$\omega_0(\theta)$ and $\omega_1(\theta)$ maybe depend on C, but we abbreviate it.

THEOREM 1. *For any θ the function $\omega_0(\theta)$ is given by*

$$\omega_0(\theta)=\theta g(\theta)/h(\theta) \tag{9}$$

where

$$g(\theta)=\int_{-C-\theta}^{C-\theta}(t+\theta)\phi(t)dt \qquad \textit{and} \qquad h(\theta)=\int_{-C-\theta}^{C-\theta}(t+\theta)^2\phi(t)dt .$$

PROOF OF THEOREM 1. As the function $\Psi_c(\omega, \theta)$ is quadratic in ω with the positive coefficient to ω^2, the minimum is evidently attained at the value given in (9).

LEMMA 1. *Let* $f(\theta)=(d/d\theta)g(\theta)$ *for any* θ, *then we have*

$$f(\theta)=f(-\theta)\,, \qquad f(0)>0 \tag{10}$$

and there exists a positive number θ_1 *such that*

$$\begin{aligned} &f(\theta)>0 \quad &&\text{if } |\theta|<\theta_1\,,\\ &f(\theta)=0 \quad &&\text{if } |\theta|=\theta_1\,,\\ &f(\theta)<0 \quad &&\text{if } |\theta|>\theta_1\,. \end{aligned} \tag{11}$$

PROOF OF LEMMA 1. (10) is obvious. So we consider only the case of $\theta \geqq 0$.

$$\begin{aligned} \frac{d}{d\theta}f(\theta)&=(-C^2-1+C\theta)\phi(C-\theta)+(C^2+1+C\theta)\phi(-C-\theta)\\ &=\phi(C+\theta)\{g_1(\theta)-g_2(\theta)\} \end{aligned}$$

where $g_1(\theta)=C\theta+1+C^2$ and $g_2(\theta)=(C^2+1-C\theta)e^{2C\theta}$. As $g_1(0)=g_2(0)=C^2+1$ and $(d/d\theta)g_2(\theta)>(d/d\theta)g_1(\theta)$ at $\theta=0$, $g_2(\theta)$ increases up the line $g_1(\theta)$ and decreases after $\theta=C+1/2C$. $g_2(\theta)$ stays negative after $\theta=C+1/C$. Therefore except for $\theta=0$ the straight line of $g_1(\theta)$ and the curve of $g_2(\theta)$ cross at a single point between $C+1/2C$ and $C+1/C$. Thus there is a point $\theta_0\,(>0)$ such that

$$\begin{aligned} &\frac{d}{d\theta}f(\theta)<0 \quad &&\text{if } 0<\theta<\theta_0\,,\\ &\frac{d}{d\theta}f(\theta)=0 \quad &&\text{if } \theta=\theta_0\,,\\ &\frac{d}{d\theta}f(\theta)>0 \quad &&\text{if } \theta>\theta_0\,. \end{aligned} \tag{12}$$

The fact that $f(0)>0$ and $\lim_{\theta\to\infty} f(\theta)=0$ and (12) guarantee the existence of a positive number θ_1 satisfying (11). This completes the proof of Lemma 1.

THEOREM 2. *The function* $\omega_0(\theta)$ *which is given in* (9) *is a non-negative even function satisfying* $\omega_0(\theta)=0$ *if and only if* $\theta=0$ *and there exists a positive number* θ_1 *such that*

$$\begin{aligned} &\omega_0(\theta)<1 \qquad \textit{if } |\theta|<\theta_1\,, \\ (13) \qquad &\omega_0(\theta)=1 \qquad \textit{if } |\theta|=\theta_1\,, \\ &\omega_0(\theta)>1 \qquad \textit{if } |\theta|>\theta_1\,. \end{aligned}$$

PROOF OF THEOREM 2. As $g(0)=0$ and $\lim_{\theta\to\infty} g(\theta)=0$, Lemma 1 implies that $g(\theta)$ is positive in $\theta>0$, and hence $\omega_0(\theta)$ is positive except for $\theta=0$ because of $\theta g(\theta)=-\theta g(-\theta)$ and $h(\theta)=h(-\theta)$. Now, $\omega_0(\theta)<1$, $=1$ and >1 is equivalent to $h(\theta)-\theta g(\theta)>0$, $=0$ and <0, respectively, because of $h(\theta)>0$. On the other hand it turns out

$$f(\theta)=h(\theta)-\theta g(\theta) \tag{14}$$

where $f(\theta)$ is given in Lemma 1, and because of Lemma 1, the existence of a positive number θ_1 is concluded.

COROLLARY. *$\omega_1(\theta)$ defined by* (8) *is given by*

$$\omega_1(\theta)=\begin{cases} \omega_0(\theta) & \textit{if } |\theta|<\theta_1 \\ 1 & \textit{if } |\theta|\geqq\theta_1 \end{cases} \tag{15}$$

where $\omega_0(\theta)$ is defined by (7) *and θ_1 is given in Theorem* 2.

PROOF OF COROLLARY. The function $\Psi_c(\omega, \theta)$ is quadratic in ω with the positive coefficient to ω^2 and by Theorem 2 we have $0\leqq\omega_0(\theta)<1$ in case of $|\theta|<\theta_1$, so in this case $\omega_1(\theta)$ is equal to $\omega_0(\theta)$. And from (13) we have $\omega_0(\theta)\geqq 1$ in case of $|\theta|\geqq\theta_1$, therefore $\min_{0\leqq\omega\leqq 1} \Psi_c(\omega, \theta)$ is attained at $\omega=1$, that is, $\omega_1(\theta)=1$.

THEOREM 3. *The minimax regret weight exists, whose value is independent of n and σ.*

PROOF OF THEOREM 3. By the corollary of Theorem 2, the regret function (4) can be written as $\mathrm{Reg}_c(\omega, \mu)=(\sigma^2/n)R_c(\omega, \theta)$ where $\theta=\sqrt{n}\,\mu/\sigma$ and $R_c(\omega, \theta)=\Psi_c(\omega, \theta)-\Psi_c(\omega_1(\theta), \theta)$. As $\Psi_c(\omega, \theta)$ and $\omega_1(\theta)$ are continuous, $R_c(\omega, \theta)$ is continuous both in ω and θ. Examination of (6) indicates $\lim_{\theta\to\pm\infty} R_c(\omega, \theta)=0$. Therefore the supremum of $R_c(\omega, \theta)$ on $\theta\in(-\infty, +\infty)$ is attained at a finite value of θ, or equivalently there is a function $\theta(\omega)$ with $\sup_{-\infty<\theta<\infty} R_c(\omega, \theta)=R_c(\omega, \theta(\omega))$ and $\theta(\omega)\in(-\infty, +\infty)$. And $R_c(\omega, \theta(\omega))$ is a lower semicontinuous function of ω defined on [0, 1]. As [0, 1] is compact, the infimum of $R_c(\omega, \theta(\omega))$ is attained at a point in [0, 1], which does not depend on n and σ. This completes the proof of Theorem 3.

3. Numerical comparisons and discussions

At first we calculate the value of the minimax regret weight in case of $C=\sqrt{2}$. After rather extensive numerical examinations, we arrived at the value of the minimax regret weight, which is $\omega^*=0.6396$. We also found that $\omega_1(\theta)=1$ when $|\theta|\geqq 1.22356$. In Fig. 1 the curves of $\Psi_{\sqrt{2}}(0, \theta)$, $\Psi_{\sqrt{2}}(1, \theta)$, $\Psi_{\sqrt{2}}(\omega_0(\theta), \theta)$, $\Psi_{\sqrt{2}}(\omega_1(\theta), \theta)$ and $\Psi_{\sqrt{2}}(\omega^*, \theta)$ are plotted as functions of $|\theta|$.

For an estimator $\tilde{\theta}$, let θ^* be defined as the largest value of $|\theta|$ such that for all $|\theta|<\theta^*$ the mean square error of $\tilde{\theta}$ is smaller than one. θ^* will be referred to as the effective interval when compared with the well-known estimator $\mathrm{PT}(1, C)=\bar{X}$.

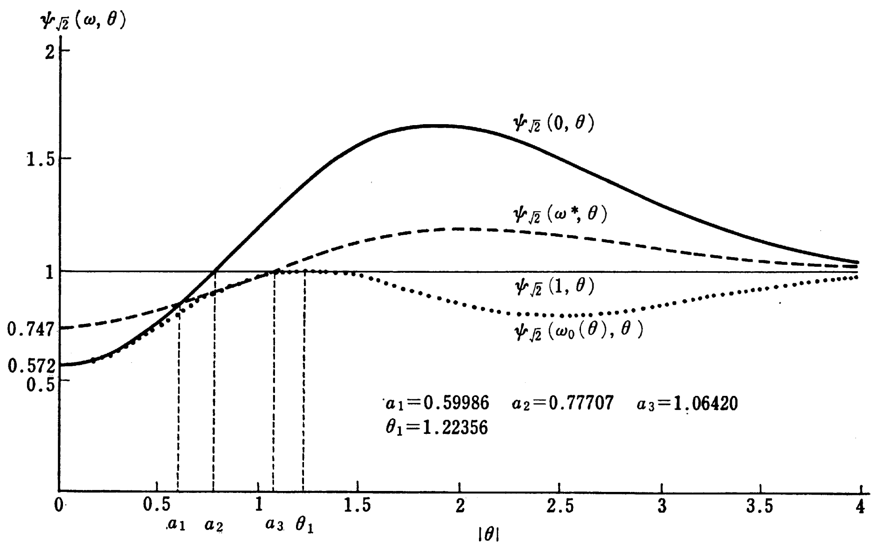

Fig. 1. Curves of $\Psi_{\sqrt{2}}(\omega, \theta)$

The mean square error of $\mathrm{PT}(\omega, \sqrt{2})$ is equal to $(\sigma^2/n)\Psi_{\sqrt{2}}(\omega, \theta)$. The curves of $\Psi_{\sqrt{2}}(0, \theta)$ and $\Psi_{\sqrt{2}}(1, \theta)$ correspond to $\mathrm{PT}(0, \sqrt{2})$ and $\mathrm{PT}(1, \sqrt{2})$. $\Psi_{\sqrt{2}}(\omega_0(\theta), \theta)=\min_{-\infty<\omega<\infty} \Psi_{\sqrt{2}}(\omega, \theta)$.

$$\Psi_{\sqrt{2}}(\omega_1(\theta), \theta)=\min_{0\leqq\omega\leqq 1} \Psi_{\sqrt{2}}(\omega, \theta)=\begin{cases} \Psi_{\sqrt{2}}(\omega_0(\theta), \theta) & \text{if } |\theta|\leqq 1.22356. \\ 1 & \text{if } |\theta|>1.22356. \end{cases}$$

The curve $\Psi_{\sqrt{2}}(\omega^*, \theta)$ corresponds to $\mathrm{PT}(\omega^*, \sqrt{2})$, the minimax regret estimator, which minimizes the maximum of $\Psi_{\sqrt{2}}(\omega, \theta)-\Psi_{\sqrt{2}}(\omega_1(\theta), \theta)$.

The curves of Fig. 1 indicate: (I) the maximum possible value of the mean square error is smaller for $\mathrm{PT}(\omega^*, \sqrt{2})$ than $\mathrm{PT}(0, \sqrt{2})$; (II) the effective interval is greater for $\mathrm{PT}(\omega^*, \sqrt{2})$ than $\mathrm{PT}(0, \sqrt{2})$; (III) for larger values of $|\theta|$, $\mathrm{PT}(\omega^*, \sqrt{2})$ is better than $\mathrm{PT}(0, \sqrt{2})$

and (IV) for smaller values of $|\theta|$, PT$(0, \sqrt{2})$ is better than PT$(\omega^*, \sqrt{2})$. PT$(\omega_0(\theta), \sqrt{2})$ is always better than PT$(0, \sqrt{2})$, PT$(1, \sqrt{2})$ and PT$(\omega^*, \sqrt{2})$, but it depends on an unknown parameter and hence it is not an estimator. Furthermore the region of $|\theta|$ that $\Psi_{\sqrt{2}}(\omega^*, \theta) \leqq \Psi_{\sqrt{2}}(0, \theta)$ is $|\theta| \geqq 0.59986$.

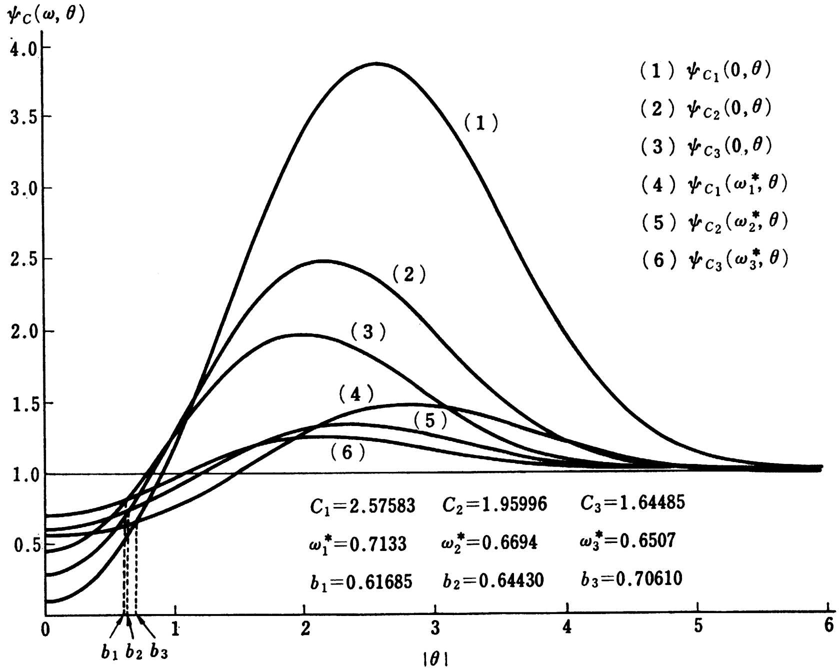

Fig. 2. Curves of $\Psi_c(0, \theta)$ and $\Psi_c(\omega^*, \theta)$ for $C=C_1, C_2, C_3$.

Let $C_1=2.57583$, $C_2=1.95996$ and $C_3=1.64485$. $C=C_1, C_2, C_3$ correspond to $\alpha=0.01, 0.05, 0.1$ respectively, in the preliminary test estimator given by (1). For $C=C_1, C_2, C_3$, we have the values of the minimax regret weight, which are $\omega^*=0.7133, 0.6694, 0.6507$ respectively. In Fig. 2 the curves of $\Psi_c(0, \theta)$ and $\Psi_c(\omega^*, \theta)$ for $C=C_1, C_2, C_3$ are plotted as functions of $|\theta|$, and in these casses similar results to the case of $C=\sqrt{2}$ are obtained. Thus it seems that the minimax regret estimator PT(ω^*, C) is an adequate choice.

Next we shall examine another type of the estimator of the form:

$$(16) \qquad d(\omega, C)=\begin{cases} 0 & \text{if } |\bar{X}|<C\sigma/\sqrt{n} \\ \omega\bar{X} & \text{if } |\bar{X}|\geqq C\sigma/\sqrt{n} \end{cases}$$

where the weight ω is a real number such that $0<\omega\leqq 1$. The estimator $d(\omega, C)$ is treated by Meeden and Arnold [4]. They exhibit the loss function for which this estimator is admissible. The loss function is given by

$$L(\theta, d) = W(|d-\theta|) + b\gamma(d)$$

where $\gamma(d)=0$ if $d=0$ and $\gamma(d)=1$ if $d\neq 0$, b is some positive real number that measures the cost of the complexity of using an estimate different from zero, and W is a nonconstant, nondecreasing function that measures the loss due to the inaccuracy of the estimate. If we denote the mean square error of $d(\omega, C)$ as $\mathrm{MD}_c(\omega, \theta)$, we have $\mathrm{MD}_c(\omega, \theta) = (\sigma^2/n)\kappa_c(\omega, \theta)$ where $\theta = \sqrt{n}\,\mu/\sigma$ and $\kappa_c(\omega, \theta) = \int_{-C-\theta}^{C-\theta} \theta^2\phi(t)dt + \int_{-\infty}^{-C-\theta} \{\omega(t+\theta) - \theta\}^2\phi(t)dt + \int_{C-\theta}^{\infty} \{\omega(t+\theta)-\theta\}^2\phi(t)dt$.

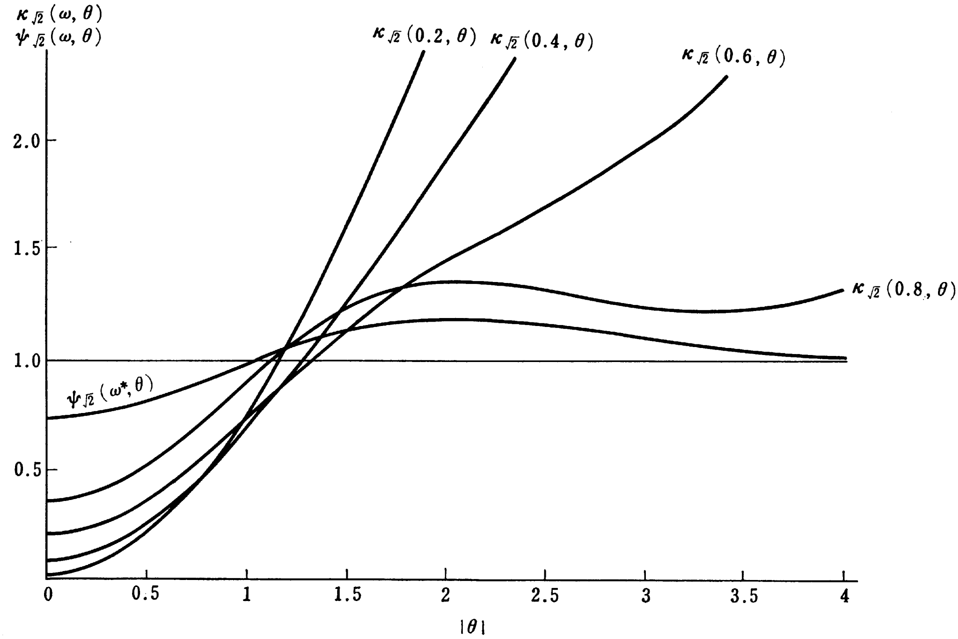

Fig. 3. Curves of $\kappa_{\sqrt{2}}(\omega, \theta)$ and $\Psi_{\sqrt{2}}(\omega^*, \theta)$

The mean square error of $d(\omega, \sqrt{2})$ is equal to $(\sigma^2/n)\kappa_{\sqrt{2}}(\omega, \theta)$. The curves of $\kappa_{\sqrt{2}}(\omega, \theta)$ and $\Psi_{\sqrt{2}}(\omega^*, \theta)$ correspond to $d(\omega, \sqrt{2})$ and PT $(\omega^*, \sqrt{2})$.

As the mean square error is used as a criterion of goodness of estimators in this paper, in Fig. 3 we show the curves of $\Psi_{\sqrt{2}}(\omega^*, \theta)$ and $\kappa_{\sqrt{2}}(\omega, \theta)$ as functions of $|\theta|$ for $\omega = 0.2, 0.4, 0.6, 0.8$. The curves of Fig. 3 indicate: (I) the maximum possible value of the mean square error is smaller for PT $(\omega^*, \sqrt{2})$ than $d(\omega, \sqrt{2})$, in fact for any C values of the mean square error of $d(\omega, C)$ increase infinitely as $|\theta|$ increase; (II) for larger values of $|\theta|$, PT $(\omega^*, \sqrt{2})$ is better than $d(\omega, \sqrt{2})$ and (III) for smaller values of $|\theta|$, $d(\omega, \sqrt{2})$ is better than PT $(\omega^*, \sqrt{2})$. For any other C similar results are obtained.

Therefore when the experimenter has the information that $|\theta|$ is close to 0, he should use $\mathrm{PT}(0, C)$; otherwise the minimax regret estimator $\mathrm{PT}(\omega^*, C)$ will be appropriate since it controls the mean square error for small as well as moderate values of $|\theta|$.

Acknowledgements

The author is deeply indebted to Professor A. Kudô of Kyushu University for his helpful advices and critical readings of the original manuscript. The author is also grateful to Professor H. Yamato of Kagoshima University and Dr. T. Yanagawa of Kyushu University for their advices and encouragements and to the referee for helpful comments.

KAGOSHIMA UNIVERSITY

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] Bancroft, T. A. and Han, Chien-Pai (1977). Inference based on conditional specification: a note and a bibliography, *International Statistical Review*, **45**, 117-127.

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

[4] Meeden, G. and Arnold, B. C. (1979). The admissibility of a preliminary test estimator when the loss incorporates a complexity cost, *J. Amer. Statist. Ass.*, **74**, 872-874.

[5] Ohtani, K. and Toyoda, T. (1978). Minimax regret critical values for a preliminary test in pooling variances, *J. Japan Statist. Soc.*, 8, 15-20.

[6] Sawa, T. and Hiromatsu, T. (1973). Minimax regret significance points for a preliminary test in regression analysis, *Econometrica*, **41**, 1093-1101.

[7] Toyoda, T. and Wallace, T. D. (1975). Estimation of variance after a preliminary test of homogeneity and optimal levels of significance for the pre-test, *J. Econometrics*, **3**, 395-404.