Ann. Inst. Statist. Math.
35 (1983), Part A, 365-373

ADMISSIBILITY OF SOME PRELIMINARY TEST ESTIMATORS FOR THE MEAN OF NORMAL DISTRIBUTION

YASUSHI NAGATA

(Received Nov. 9, 1982; revised May 2, 1982)

Summary

Let a random variable X follow a p-variate normal distribution $N_p(\theta, I_p)$ with an unknown $p\times 1$ vector θ and $p\times p$ identity matrix I_p. The admissibility of a preliminary test estimator using AIC (Akaike's Information Criterion) procedure will be shown if $p=1$ and its inadmissibility will be shown if $p\geqq 3$ under the loss function based on Kullback-Leibler information measure. Furthermore the two sample case is also considered.

1. Introduction

A preliminary test estimation is described, as is well known, as follows: After a preliminary test of a certain null hypothesis, estimation is made under the alternative hypothesis if the null hypothesis is rejected, and made under the null hypothesis otherwise. For instance, suppose a random variable X follows a univariate normal distribution $N(\theta, 1)$. Let $H_0: \theta=0$ be the null hypothesis and $H_1: \theta\neq 0$ the alternative hypothesis. A familiar preliminary test estimator is given by

$$(1.1)\qquad d_c(X)=\begin{cases} 0 & \text{if } |X|\leqq c\,, \\ X & \text{if } |X|>c\,. \end{cases}$$

It looks appealing at a glance, but there are some inadequacies. First of all it is inadmissible under many standard loss functions for estimation including a quadratic loss function. For an admissible estimator must be a (proper) Bayes estimator or its limit and therefore it must be a smooth function in X under above loss functions. But the estimator (1.1) is not smooth and is neither Bayes estimator nor its limit, which implies its inadmissibility. Secondly when $|X|\leqq c$, we should

Key words: Preliminary test estimator; Admissibility.

consider not to be able to reject H_0 instead of accepting it and it seems unnatural to put zero as an estimate. So it may be appropriate to modify so that

$$d_c(X)=\begin{cases} d_1(X) & \text{if } |X|\leqq c\,, \\ X & \text{if } |X|>c\,. \end{cases} \tag{1.2}$$

Inada [4] adopts aX $(0\leqq a<1)$ as $d_1(X)$ in (1.2) and has obtained the optimal value of a by minimax regret criterion (He takes $c=\sqrt{2}$ by AIC procedure which will be discussed later.). However, even if we modify (1.1) in this way, the estimator (1.2) is still inadmissible if it is not smooth in X.

Preliminary test estimation should not be understood to be a procedure to be used merely to obtain a sharp estimate, but a procedure that includes both model selection and parameter estimation. Therefore we should discuss it under an appropriate loss function which fits such a situation. Meeden and Arnold [7] assume the following loss function and have shown the admissibility of the estimator (1.1):

$$L(d,\theta)=W(d,\theta)+\alpha^2 I(d)\,, \tag{1.3}$$

where $I(d)=0$, if $d=0$ and $I(d)=1$, if $d\neq 0$, and $W(d,\theta)$ is a standard loss function for estimation. The coefficient α^2 in (1.3) is a cost associated with the complexity of the estimator. This loss function may be thought of as the incorporation of the evaluation of an estimate and the simplicity of an estimator, and seems to fit well for some situations. We should note that the value of c in (1.1) is determined by α (>0) in (1.3). If, in particular, $W(d,\theta)=(d-\theta)^2$ (a quadratic loss function), (1.1) is admissible with $c=\alpha$.

In this paper we discuss the admissibility (or inadmissibility) of the estimator (1.1) under another loss function which incorporates model fitting and evaluation of an estimate. The loss function is based on Kullback-Leibler information measure and has been introduced by Inagaki [5] to show that AIC (Akaike's Information Criterion) statistic in normal linear regression is a generalized Bayes solution with respect to Lebesgue measure. Section 2 deals with the one-sample problem, where X follows a p-variate normal distribution $N_p(\theta, I_p)$ with an unknown $p\times 1$ vector θ and a $p\times p$ identity matrix I_p. The admissibility of the preliminary test estimator due to Hirano [2] determined with AIC procedure will be considered. In Theorem 2.1 it is shown to be admissible if $p=1$ and in Theorem 2.2 it is shown to be inadmissible if $p\geqq 3$, which is a kind of Stein problem (see e.g. James and Stein [6]). Section 3 deals with the two sample problem; when X_1 follows $N_p(\theta_1, \sigma_1^2 I_p)$ and X_2 follows $N_p(\theta_2, \sigma_2^2 I_p)$, σ_1^2 and σ_2^2 being known scalars. Only the inadmissibility if $p \geqq 3$ under Inagaki's [5] loss function is proved. Extention of (1.3) to the two-sample problem is treated similarly.

2. Admissibility in the one-sample problem

We shall first describe the loss function introduced by Inagaki [5] (eq. (5.6)). Let X be a random variable with p.d.f. (probability density function) $f(x, \theta) \in \mathcal{F} = \{f(x, \theta)\,;\ \theta \in \Theta\}$, where Θ is a parameter space. Suppose $\mathcal{F}_\gamma = \{f_\gamma(x, \zeta);\ \zeta \in \Theta_\gamma\}$ is a model for $\mathcal{F}$ and Θ_γ is a parameter space indexed by γ. And suppose $\zeta_\gamma(\theta)$ is defined by following equation,

$$(2.1) \qquad \int \log \{f(x, \theta)/f_\gamma(x, \zeta_\gamma(\theta))\} f(x, \theta) dx := \min_{\zeta \in \Theta_\gamma} \int \log \{f(x, \theta)/f_\gamma(x, \zeta)\} f(x, \theta) dx\,.$$

The loss function has the following form,

$$(2.2) \qquad L((k, d), \theta, x) = \log \{f(x, \theta)/f_k(x, \zeta_k(\theta))\} + \int \log \{f_k(y, \zeta_k(\theta))/f_k(y, \zeta_k(d))\} f_k(y, \zeta_k(\theta)) dy\,.$$

Let $k(X)$ be an estimator for index γ, $d(X)$ be an estimator for θ and $\zeta_\gamma(d(X))$ be an estimator for $\zeta_\gamma(\theta)$. The loss function (2.2) is based on Kullback-Leibler information measure, and the first term of the right-hand-side of (2.2) is a loss for a model fitting and the second term is a loss for an evaluation for an estimate.

Now, let X be a random sample from $N_p(\theta, I_p)$ with the p.d.f. $f(x, \theta)$, where θ is an unknown $p \times 1$ vector and I_p is a $p \times p$ identity matrix. We consider two models, $\mathcal{F}_0 = \{f(x, 0)\}$ and $\mathcal{F}_1 = \{f(x, \theta);\ \theta \in \Theta\}$. In this problem Hirano [2] proposed the following as a preliminary test estimator by using AIC procedure,

$$(2.3) \qquad d_0(X) = \begin{cases} 0 & \text{if } X'X \leqq 2p\,, \\ X & \text{if } X'X > 2p\,. \end{cases}$$

In (2.3) the upper formula is considered to choose the model with $k=0$ and lower formula the model with $k=1$. The loss function (2.2) becomes,

$$(2.4) \qquad L((k, d), \theta, x) = \begin{cases} x'x/2 - (x-\theta)'(x-\theta)/2 & \text{if } k=0\,, \\ (d-\theta)'(d-\theta)/2 & \text{if } k=1\,. \end{cases}$$

The upper formula in (2.4) denotes the loss accompanied with model fitting and implies $d=0$ in (2.2), whereas the lower formula has been derived from the second term of (2.2). We shall discuss the admissibility of the estimator (2.3) under the loss function (2.4).

LEMMA 2.1. *The estmator* (2.3) *is the limit of a sequence of* (*proper*) *Bayes solutions under the loss function of* (2.4).

PROOF. Take $N_p(0, \tau^2 I_p)$ as a prior distribution of θ. Then the posterior distribution of θ given X is $N_p(a(\tau)X, a(\tau)I_p)$, where $a(\tau)=\tau^2/(1+\tau^2)$. Let $g(\theta|X)$ be the conditional p.d.f. Then the posterior risk $\rho((k, d), \tau, x)$ can be written as

$$\rho((0, d), \tau, x)=\int L((0, d(x)), \theta, x)g(\theta|x)d\theta =(1/2)((2a(\tau)-a(\tau)^2)x'x-pa(\tau)) , \tag{2.5}$$

and

$$\rho((1, d), \tau, x)=(1/2)\int (d(x)-\theta)'(d(x)-\theta)g(\theta|x)d\theta , \tag{2.6}$$

which is minimized by putting $d(x)=a(\tau)x$ and we obtain

$$\min_d \rho((1, d), \tau, x)=pa(\tau)/2 . \tag{2.7}$$

Therefore comparing (2.5) with (2.7) we obtain a following Bayes solution,

$$d_\tau(X)=\begin{cases} 0 & \text{if } X'X\leqq 2p/(2-a(\tau)) , \\ a(\tau)X & \text{if } X'X>2p/(2-a(\tau)) . \end{cases} \tag{2.8}$$

Since $a(\tau)\to 1$ as $\tau\to\infty$, (2.8) converges to (2.3).

Now since the estimator (2.3) is a limit of Bayes solutions under our framework of the loss function, it is interesting to consider its admissibility. We shall show below its admissibility when $p=1$ and its inadmissibility when $p\geqq 3$.

THEOREM 2.1. *The estimator* (2.3) *is admissible when* $p=1$.

PROOF. We shall use the method of Blyth [1]. By straightforward calculation the risk functions of d_0 and d_τ become respectively,

$$R(\theta, d_0)=\int L((k(x), d_0(x)), \theta, x)f(x, \theta)dx =1/2+(1/2)\int_{-\sqrt{2}}^{\sqrt{2}} (4x\theta-x^2-2\theta^2)f(x, \theta)dx \tag{2.9}$$

and

$$(2.10)\quad R(\theta, d_\tau)=\int L((k(x), d_\tau(x)), \theta, x)f(x, \theta)dx$$
$$=(1/2)\{a(\tau)^2+\theta^2(a(\tau)-1)^2\}$$
$$+(1/2)\int_{-\sqrt{2/(2-a(\tau))}}^{\sqrt{2/(2-a(\tau))}}\{2x\theta(1+a(\tau))-a(\tau)^2x^2-2\theta^2\}f(x, \theta)dx\,.$$

Next we calculate the Bayes risk functions of d_0 and d_τ with respect to $N(0, \tau^2)$, obtaining

$$(2.11)\quad r(\tau^2, d_0)=1/2+(1/2)\int_{-\sqrt{2}}^{\sqrt{2}}\{x^2(4a(\tau)-1-2a(\tau)^2)-2a(\tau)\}f_\tau(x)dx\,,$$

where $f_\tau(x)$ is the p.d.f. of marginal distribution of X, $N(0, 1+\tau^2)$, and

$$(2.12)\quad r(\tau^2, d_\tau)=a(\tau)/2+(1/2)\int_{-\sqrt{2/(2-a(\tau))}}^{\sqrt{2/(2-a(\tau))}}\{x^2(2a(\tau)-a(\tau)^2)-2a(\tau)\}f_\tau(x)dx\,.$$

Suppose that d_0 is inadmissible. Then there exists another estimator d^* such that for all $\theta\in\Theta$

$$(2.13)\quad R(\theta, d^*)\leqq R(\theta, d_0)$$

and for some $\theta_0\in\Theta$

$$(2.14)\quad R(\theta_0, d^*)<R(\theta_0, d_0)\,.$$

Since the loss function (2.4) is continuous in θ with fixed d, there exists ε (>0) and δ (>0) such that for all $\theta\in(\theta_0-\delta,\ \theta_0+\delta)$,

$$(2.15)\quad R(\theta, d^*)<R(\theta, d_0)-\varepsilon\,.$$

Therefore from (2.13) and (2.15) we obtain

$$(2.16)\quad r(\tau^2, d_0)-r(\tau^2, d^*)>\varepsilon\int_{\theta_0-\delta}^{\theta_0+\delta}(1/\sqrt{2\pi}\,\tau)\exp(-\theta^2/2\tau^2)d\theta=\varepsilon K/\tau\,,$$

where K is a positive constant number. From (2.11) and (2.12) it follows that

$$(2.17)\quad r(\tau^2, d_0)-r(\tau^2, d_\tau)$$
$$=(1/2)\int_{-\sqrt{2}}^{\sqrt{2}}\{x^2(4a(\tau)-1-2a(\tau)^2)-2a(\tau)\}f_\tau(x)dx$$
$$-(1/2)\int_{-\sqrt{2/(2-a(\tau))}}^{\sqrt{2/(2-a(\tau))}}\{x^2(2a(\tau)-a(\tau)^2)-2a(\tau)\}f_\tau(x)dx+1/2(1+\tau^2)\,.$$

Using (2.16), (2.17) and Lebesgue's dominating convergence theorem, we have

$$(2.18)\quad \{r(\tau^2, d_0)-r(\tau^2, d^*)\}/\{r(\tau^2, d_0)-r(\tau^2, d_\tau)\}\to\infty \qquad \text{as } \tau\to\infty\,.$$

So the left-hand-side of (2.18) is larger than one for a large τ, which implies

$$r(\tau^2, d^*) < r(\tau^2, d_\tau) . \tag{2.19}$$

This contradicts the fact that d_τ is a Bayes solution and we complete the proof.

THEOREM 2.2. *The estimator* (2.3) *is inadmissible when* $p \geqq 3$.

PROOF. We take the following estimator as a candidate to improve d_0:

$$d_1(X) = \begin{cases} 0 & \text{if } X'X \leqq 2p , \\ (1 - c/X'X)X & \text{if } X'X > 2p , \end{cases} \tag{2.20}$$

where c is a constant. Let $S_1 = \{x;\ x'x \leqq 2p\}$ and $S_2 = \{x;\ x'x > 2p\}$. We shall show that there exists a constant c such that $R(\theta, d_0) \geqq R(\theta, d_1)$ for all θ and $R(\theta_0, d_0) > R(\theta_0, d_1)$ for at least one θ_0. It follows that

$$\begin{aligned} (2.21)\quad & R(\theta, d_0) - R(\theta, d_1) \\ &= (1/2) \int_{S_2} (x-\theta)'(x-\theta) f(x, \theta) dx \\ &\quad - (1/2) \int_{S_2} (x - cx/x'x - \theta)'(x - cx/x'x - \theta) f(x, \theta) dx \\ &= \int_{S_2} (x-\theta)'(cx/x'x) f(x, \theta) dx - (1/2) \int_{S_2} (c^2/x'x) f(x, \theta) dx . \end{aligned}$$

By integration by parts we obtain

$$\int (x_i - \theta_i)(cx_i/x'x) f(x, \theta) dx = \int (\partial/\partial x_i)(cx_i/x'x) f(x, \theta) dx , \tag{2.22}$$

where x_i and θ_i are respectively i-th coordinates of x and θ. Since S_1 is a region of x such that $x'x \leqq 2p$, putting $A_i(x) := \left(2p - \sum_{j \neq i} x_j^2\right)^{1/2}$, we have

$$-A_i(x) \leqq x_i \leqq A_i(x) \qquad \text{for any } x \in S_1 . \tag{2.23}$$

Similarly by integration by parts we obtain

$$\begin{aligned} (2.24)\quad & \int_{S_1} (x_i - \theta_i)(cx_i/x'x) f(x, \theta) dx \\ &= \int_{S_1'} [-(cx_i/x'x)(1/\sqrt{2\pi}) \exp\{-(x_i - \theta_i)^2/2\}]_{-A_i(x)}^{A_i(x)} f(x_i', \theta_i') dx_i' \\ &\quad + \int_{S_1} (\partial/\partial x_i)(cx_i/x'x) f(x, \theta) dx \\ &\leqq \int_{S_1} (\partial/\partial x_i)(cx_i/x'x) f(x, \theta) dx , \end{aligned}$$

where x_i' and θ_i' denote the vectors obtained by deleting the i-th components in x and θ respectively, and S_1' is the region after integrating out x_i. Substituting (2.22) and (2.24) into (2.21), we obtain

$$R(\theta, d_0)-R(\theta, d_1)\geqq\sum_{i=1}^{p}\int_{S_2}(\partial/\partial x_i)(cx_i/x'x)f(x, \theta)dx \tag{2.25}$$
$$-(1/2)\int_{S_2}(c^2/x'x)f(x, \theta)dx$$
$$=\int_{S_2}\{(pc-2c-c^2/2)/x'x\}f(x, \theta)dx\ .$$

Therefore if we choose $0<c<2(p-2)$, the right-hand-side of (2.25) is positive, which implies that d_1 dominates d_0.

Now we extend the loss function (1.3) by Meeden and Arnold [7] to a multivariate case, which becomes

$$L(d, \theta)=(d-\theta)'(d-\theta)+\alpha^2 I(d)\ , \tag{2.26}$$

where $I(d)=0$, if $d=0$ and $I(d)=1$, if $d\neq 0$. As the first term of (1.3) we take a quadratic loss function for simplicity. Then the following result holds similarly as Theorem 2.2.

COROLLARY 2.1. *The estimator*

$$d(X)=\begin{cases} 0 & \text{if } X'X\leqq\alpha^2\ , \\ X & \text{if } X'X>\alpha^2 \end{cases} \tag{2.27}$$

is the limit of some sequence of Bayes solutions but is inadmissible when $p\geqq 3$ *under the loss function* (2.26).

PROOF. The estimator (2.27) is dominated by the estimator with α^2 instead of $2p$ in (2.20).

Remark. i) In the case of $p=2$ the admissibility of the estimator (2.3) remains open.

ii) We should note that Theorem 2.2 and Corollary 2.1 correspond to the result of Sclove, Morris and Radhakrishnan [8] although our loss function is different from theirs (Recall that the estimator (2.3) is the limit of a sequence of (proper) Bayes solutions under our loss function.). Here we have given another proof using integration by parts.

3. Admissibility in the two-sample problem

In this section we deal with the two-sample case. Let X_1 and X_2 be random variables following independently $N_p(\theta_1, \sigma_1^2 I_p)$ and $N_p(\theta_2, \sigma_2^2 I_p)$, where σ_1^2 and σ_2^2 are known scalars. Let $X=(X_1', X_2')'$ and $\theta=(\theta_1', \theta_2')'$.

Let $f(x, \theta)$ be the p.d.f. of X. Our models considered here are $\mathcal{F}_0=\{f(x, \theta);\ \theta_1=\theta_2\}$ and $\mathcal{F}_1=\{f(x, \theta)\}$. The former is the model with equal population means and the latter is the full model. For simplicity, we write in the sequel $X=(X_1, X_2)$ etc. instead of $X=(X_1', X_2')'$. In this problem Hirano [3] proposed the following preliminary test estimator using AIC procedure:

$$(3.1)\qquad d_0(X)=\begin{cases} ((\sigma_2^2X_1+\sigma_1^2X_2)/(\sigma_1^2+\sigma_2^2),\ (\sigma_2^2X_1+\sigma_1^2X_2)/(\sigma_1^2+\sigma_2^2)) \\ \qquad\qquad \text{if } (X_1-X_2)'(X_1-X_2)/(\sigma_1^2+\sigma_2^2)\leqq 2p\,, \\ (X_1, X_2) \qquad \text{if } (X_1-X_2)'(X_1-X_2)/(\sigma_1^2+\sigma_2^2)>2p\,. \end{cases}$$

Now we consider the loss function by Inagaki [5] again. By some calculation (2.2) becomes

$$(3.2)\qquad L((k, d), \theta, x)=\begin{cases} (1/2\sigma_1^2)[-(x_1-\theta_1)'(x_1-\theta_1)+(x_1-\theta^*)'(x_1-\theta^*) \\ \qquad +(d_0-\theta^*)'(d_0-\theta^*)] \\ \quad +(1/2\sigma_2^2)[-(x_2-\theta_2)'(x_2-\theta_2)+(x_2-\theta^*)'(x_2-\theta^*) \\ \qquad +(d_0-\theta^*)'(d_0-\theta^*)] \\ \qquad\qquad \text{if } k=0\,, \\ (1/2\sigma_1^2)(d_1-\theta_1)'(d_1-\theta_1)+(1/2\sigma_2^2)(d_2-\theta_2)'(d_2-\theta_2) \\ \qquad\qquad \text{if } k=1\,, \end{cases}$$

where $d=(d_1, d_2)$ or (d_0, d_0), d_1, d_2 and d_0 are $p\times 1$ vectors and $\theta^*=(\sigma_2^2\theta_1+\sigma_1^2\theta_2)/(\sigma_1^2+\sigma_2^2)$.

LEMMA 3.1. *The estimator* (3.1) *is the limit of some sequence of Bayes solutions under the loss function* (3.2).

The lemma can be proved as Lemma 2.1, so we omit a proof. Since the estimator (3.1) is a limit of Bayes solutions by Lemma 3.1, we are interested in its admissibility.

THEOREM 3.1. *The estimator* (3.1) *is inadmissible when* $p\geqq 3$ *under the loss function* (3.2).

PROOF. The estimator (3.1) is dominated by

$$(3.3)\qquad d_1(X)=\begin{cases} ((\sigma_2^2X_1+\sigma_1^2X_2)/(\sigma_1^2+\sigma_2^2),\ (\sigma_2^2X_1+\sigma_1^2X_2)/(\sigma_1^2+\sigma_2^2)) \\ \qquad\qquad \text{if } (X_1-X_2)'(X_1-X_2)/(\sigma_1^2+\sigma_2^2)\leqq 2p\,, \\ (X_1-c(X_1-X_2)/(X_1-X_2)'(X_1-X_2), \\ \qquad X_2-c(X_2-X_1)/(X_1-X_2)'(X_1-X_2)) \\ \qquad\qquad \text{if } (X_1-X_2)'(X_1-X_2)/(\sigma_1^2+\sigma_2^2)>2p\,, \end{cases}$$

where c satisfies $0<c<4\sigma_1^2\sigma_2^2(p-2)/(\sigma_1^2+\sigma_2^2)$. The proof that d_0 is dominated by d_1 is similar to that of Theorem 2.2.

Furthermore we consider a multivariate extension of the loss function (1.3) by Meeden and Arnold [7], i.e.

$$(3.4)\qquad L(d,\theta)=(1/\sigma_1^2)(d_1-\theta_1)'(d_1-\theta_1)+(1/\sigma_2^2)(d_2-\theta_2)'(d_2-\theta_2)+\alpha^2 I(d_1,d_2)\,,$$

where $I(d_1, d_2)=0$, if $d_1=d_2$ and $=1$, if $d_1\neq d_2$. In this case the next corollary holds.

COROLLARY 3.1. *The estimator*

$$(3.5)\qquad d(X)=\begin{cases}((\sigma_2^2X_1+\sigma_1^2X_2)/(\sigma_1^2+\sigma_2^2),\ (\sigma_2^2X_1+\sigma_1^2X_2)/(\sigma_1^2+\sigma_2^2)) & \\ \qquad\qquad \text{if } (X_1-X_2)'(X_1-X_2)/(\sigma_1^2+\sigma_2^2)\leqq\alpha^2\,, & \\ (X_1, X_2) \qquad \text{if } (X_1-X_2)'(X_1-X_2)/(\sigma_1^2+\sigma_2^2)>\alpha^2 & \end{cases}$$

is the limit of some sequence of Bayes solutions but is inadmissible when $p\geqq 3$ *under the loss function* (3.4).

PROOF. The estimator (3.5) is dominated by the estimator with α^2 instead of $2p$ in (3.5).

Remark. In the cases $p=1$ and 2 the admissibility of the estimator (3.1) remains open.

OSAKA UNIVERSITY

REFERENCES

[1] Blyth, C. R. (1951). On minimax statistical decision procedures and their admissibility, *Ann. Math. Statist.* **22**, 22-42.

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

[3] Hirano, K. (1978). On level of significance of the preliminary test in pooling means, *Ann. Inst. Statist. Math.*, **30**, 1-8.

[4] Inada, K. (1981). A minimax regret estimator of a normal mean after preliminary test, *Proc. 50th Annual Meeting of Japan Statist. Soc.*, Kansai Univ. (in Japanese).

[5] Inagaki, N. (1977). Two errors in statistical model fitting, *Ann. Inst. Statist. Math.*, **29**, 131-152.

[6] James, W. and Stein, C. (1961). Estimation with quadratic loss, *Proc. 4th Berkeley Symp. Math. Statist. Prob.*, **1**, 361-379.

[7] 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.

[8] Sclove, S. L., Morris, C. and Radhakrishnan, R. (1972). Non-optimality of preliminary-test estimators for the mean of a multivariate normal distribution, *Ann. Math. Statist.*, **43**, 1481-1490.