Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT®

Journal of Statistical Planning and Inference 128 (2005) 95–107

journal of statistical planning and inference

www.elsevier.com/locate/jspi

Improved confidence estimators for confidence sets of location parameters

Hsiuying Wang

Academia Sinica, Institute of Statistical Science, Taipei, Taiwan 115, Taiwan

Received 21 October 2001; accepted 6 September 2003

Abstract

Consider a p-dimensional location family symmetrical about θ. Let $C_{t(X)}$ be a $1-\alpha$ confidence set $\{\theta\colon |t(X)-\theta|\leqslant c\}$ of θ, where $t(X)$ is some reasonable estimator of θ. Traditionally, the confidence coefficient $1-\alpha$, which is data independent, is used to be the report for the confidence of $C_{t(X)}$. In this paper, some improved confidence reports are provided for $p\geqslant 5$. These results are related to Robinson (Ann. Statistic 7 (1979) 756). The normal case discussed in Robinson (Ann. Statistic 7 (1979) 756) is a special case of the results of this paper. Moreover, some admissibility results when $p\leqslant 4$ are also present in this paper.

MSC: primary 62C15; secondary 62C10

Keywords: t-distribution; Confidence coefficient; Admissibility

1. Introduction

Traditionally, the confidence coefficient is used to be a confidence report for a confidence set. However, the confidence coefficient is a data-independent report. In the following, examples will be given to explain why the data-independent estimator is not feasible in some situations. For example, assume that the possible values of a random variable X are $\theta-1$ and $\theta+1$. The probability of X assigned to these two values is

$$p(X=\theta-1)=p(X=\theta+1)=\tfrac{1}{2}.$$

E-mail address: hywang@stat.sinica.edu.tw (H. Wang).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2003.09.029

Let x_1 and x_2 be two independent observations of X. A confidence interval of θ based on x_1 and x_2 is

$$C_x = \left\{ \theta : \left| \frac{x_1 + x_2}{2} - \theta \right| < \frac{1}{2} \right\}.$$

In fact, if we have two different observations of X, then we know that $I(\theta \in C_x) = 1$. If the two observations are the same value, then $I(\theta \in C_x) = 0$. It can be seen that the value of $I(\theta \in C_x)$ depends on the observations. Therefore, it is not feasible to use a data independent estimator to report the confidence of C_x. A data-dependent estimator is more feasible for estimating the confidence than a data independent one. This is an example that shows why we should think about the data-dependent confidence report.

Let X be a p-dimensional random variable with a p-dimensional unknown parameter θ. For a confidence set C_X of parameter θ, determining a confidence report for the confidence of C_X can be viewed as an estimation problem of estimation of the coverage function

$$I(\theta \in C_X) = \begin{cases} 1 & \text{if } \theta \in C_X \\ 0 & \text{if } \theta \notin C_X \end{cases} \tag{1}$$

of C_X.

Usually, the constant coverage probability estimator $1-\alpha$ is used as the estimator of $I(\theta \in C_X)$. Note that $1-\alpha$ is data-independent. However, according to the example given in the above paragraph, it might be better to consider a data-dependent estimator for (1), unless there are no existent data-dependent estimators better than $1-\alpha$. Therefore, we are interested to discover if there exist some estimators better than $1-\alpha$.

Robinson (1979) and Robert and Casella (1994) have some results relating to the confidence reports in the multivariate normal distribution. Robinson (1979) points out the relationship of confidence report and some conditional properties of statistical procedures and show that for estimating

$$I(\theta\colon |\bar{X} - \theta| \leqslant c), \tag{2}$$

where $X_{p\times 1} \sim N_{p\times 1}(\theta, I)$, and c satisfies that $P(|\bar{X} - \theta| \leqslant c) = 1 - \alpha$, $\bar{X}$ is the mean of the observations, the estimator $\gamma(X) = 1 - \alpha + \delta/(1 + |X|^2)$ is better than the usual constant coverage probability estimator if $p = 5$ under the squared error loss for some positive constants δ. Robert and Casella (1994) extend the results to $p \geqslant 5$ cases. Note that (2) is a rectangular confidence set, not an interval. From these two papers, we can predict that there might exist some better estimation than the usual constant coverage probability estimator for the location families. Besides, Wang (1999) also show that there exists a better data-dependent estimator than the usual constant coverage probability estimator for the confidence report in the regression model, where the distribution of error term is a normal distribution. Thus, in this literature, the normal cases are widely discussed, and the results are seldom applied to other distributions. In this paper, we extend the results to more general cases.

Consider the confidence sets of location parameters of a p-multidimensional location family with marginal density functions $(f_{\theta_1}(x),\dots,f_{\theta_p}(x))$. Let $\theta=(\theta_1,\dots,\theta_p)$ and $X_i=(X_{i1},\dots,X_{ip})$ $i=1,\dots,n$ be a sample such that $X_{ij}-\theta_j$ are independent random variables with density function $f_0(x)$, $j=1,\dots,p$. Let $t(X)$ denote a reasonable estimator for θ and $C_{t(X)}$ denote the confidence set

$$C_{t(X)}=\{\theta\colon\ |t(X)-\theta|\leqslant c\},$$

where c satisfies $P(\theta\in C_{t(X)})=1-\alpha$. Consider the squared error loss function

$$L(\phi(X),\theta)=(\phi(X)-I(\theta\in C_{t(X)}))^2. \tag{3}$$

Some estimators better than $1-\alpha$ are provided for the location families when $p\geqslant 5$ under the loss function (3) in Section 2. However, these improved estimators are not completely specified in this paper due to the difficulty in determining two constants in the estimators. Nevertheless, Theorem 1 provides a guideline to choose the constants.

Moreover, the admissibility results for $p\leqslant 4$ are also included in Section 3. Section 4 gives some examples of the results of Section 2. And some simulation results are provided in Section 5.

2. Improved confidence estimators

In this section, some estimators will be shown to be better than the usual constant coverage probability estimator for estimating (1).

Theorem 1. *Let $X_i=(X_{i1},\dots,X_{ip})$, $i=1,\dots,n$, be a sample from a p dimensional location family with density functions $(f_{\theta_1}(x),\dots,f_{\theta_p}(x))$. Here $X_{ij}-\theta_j$, $j=1,\dots,p$, are identically distributed and follow the distribution with density function $f_0(x)$. Let $Y_i=(X_{1i},\dots,X_{ni})'$, $i=1,\dots,p$, and $t(X)=(e(Y_1),\dots,e(Y_p))$, where $X=(Y_1,\dots,Y_p)$ denotes an estimator of θ such that $e(Y_i)-\theta_i$ is symmetrical about 0 and the density function $h(s)$ of $s=t(X)-\theta$ satisfies*

$$h(s)=\mathrm{o}\left(\frac{1}{\prod_{i=1}^{p} s_i^4}\right)\quad \text{as } |s_i|\to\infty. \tag{4}$$

Then for estimating (1),

$$\gamma_1(X)=1-\alpha+\frac{a}{b+t'(X)t(X)}$$

has smaller risk than $1-\alpha$ for $p\geqslant 5$ under the loss function (3), where $a>0$ and $b>0$ are sufficiently small and sufficiently large constants, respectively.

Proof.

$$E(1-\alpha-I(|t(X)-\theta|\leqslant c))^2-E(\gamma_1(X)-I(|t(X)-\theta|\leqslant c))^2$$

$$= -2E\left[\frac{a}{b+t'(X)t(X)}(1-\alpha-I(|t(X)-\theta|\leqslant c))\right]$$
$$-E\left(\frac{a}{b+t'(X)t(X)}\right)^2 \qquad (5)$$

Now it will be shown that (5) > 0. Let $t(X)-\theta=Z=(Z_1,\ldots,Z_p)$. Then $(Z_1,\ldots,Z_p)$ are identically distributed since $t_i(X)$ are identically distributed. By using Taylor's expansion

$$\frac{1}{b+t'(X)t(X)}$$
$$=\frac{1}{b+(Z+\theta)'(Z+\theta)}$$
$$=(b+|\theta|^2)^{-1}-2\sum_{i=1}^{p}\frac{Z_i\theta_i}{(b+|\theta|^2)^2}$$
$$+4\sum_{\substack{i=1\\j=1}}^{p}Z_iZ_j\frac{\theta_i\theta_j}{(b+|\theta|^2)^3}-\sum_{i=1}^{p}\frac{Z_i^2}{(b+|\theta|^2)^2}+R(\eta,\theta),$$

where $R(\eta,\theta)$ is the remainder term and η is a point on the line segment joining z to the origin.

By substituting Taylor's expansion of $1/(b+t'(X)t(X))$ into (5) and using the fact that $(Z_1,\ldots,Z_p)$ are identically distributed and $E(Z_1(1-\alpha-I(|Z|\leqslant c)))=0$ (Z_i is symmetrical about 0 and $t(X)$ satisfies (4)), (5) is equal to

$$2aE[Z_1^2(1-\alpha-I(|Z|\leqslant c))]\frac{(p-4)|\theta|^2+pb}{(b+|\theta|^2)^3}-a^2\frac{1}{(b+|\theta|^2)^2}+e(\theta), \qquad (6)$$

where

$$e(\theta)=aE[R(Z,\eta,\theta)(1-\alpha-I(|Z|\leqslant c))]-4a^2E\left(\sum_{i=1}^{p}\frac{Z_i\theta_i}{\left(b+|\theta|^2\right)^2}\right)^2.$$

$e(\theta)$ is shown in the Appendix to be $o(1/[(b+|\theta|^2)^2]$ when b is large enough. Note that the term $E[Z_1^2(1-\alpha-I(|Z|\leqslant c))]$ in (6) is positive since Z_1^2 and $(1-\alpha-I(|Z|\leqslant c))$ are increasing in Z_1^2 and the density function of $t(X)$ satisfies (4). Therefore, for $p>4$, (6) is positive if a is chosen small enough and b is chosen large enough. □

Remark 1. In Theorem 1, when the location family is a normal distribution, the improved confidence estimator $\gamma(X)=1-\alpha+a/(b+X'X)$ is derived from empirical Bayes aspect (see Robert and Casella (1994) and Wang (1999)). For general location distributions, however, it is hard to find a general conjugate prior. Therefore, a similar argument from empirical Bayes perspective for normal distributions cannot apply to location families. Thus, we try to construct $\gamma_1(X)$ based on the form of $\gamma(X)$ since normal distribution is a location family. And fortunately, it is an accurate presumption.

When the random variables $X_{ij} - \theta_j$, $j = 1, \ldots, p$, in Theorem 1 are not independent and each pair has the same correlation, Theorem 2 gives a necessary condition for p such that $\gamma_1(X)$ is better than $1 - \alpha$ under the loss function (3).

Theorem 2. *Assumed that the random variables $X_{ij} - \theta_j$, $j = 1, \ldots, p$, in Theorem 1 are not independent, and have a covariance matrix Σ, where the diagonal terms in Σ are 1 and the other terms in Σ are the same. Then when $p > 4 + 2E[Z_1 Z_2(1 - \alpha - I(|Z| \leqslant c))]/E[Z_1^2(1 - \alpha - I(|Z| \leqslant c))]$, the results of Theorem 1 hold.*

Proof. Since $X_{ij} - \theta_j$, $j = 1, \ldots, p$, are not independent and $E(X_{ij} - \theta_j)(X'_{ij'} - \theta'_{j'})$ are a constant for any pair (j, j'), $j \neq j'$, by an argument similar to that in Theorem 1 and the fact that $\sum_{i \neq j} \theta_i \theta_j \leqslant |\theta|^2/2$, (5) is greater than

$$\begin{aligned}
&2aE[Z_1^2(1 - \alpha - I(|Z| \leqslant c))] \\
&\times[(p-4-2E(Z_1 Z_2(1-\alpha-I(|Z|\leqslant c)))/E[Z_1^2(1-\alpha-I(|Z|\leqslant c))])|\theta|^2 + pb]/ \\
&(b + |\theta|^2)^3 - a^2/(b + |\theta|^2)^2 + e(\theta). \qquad (7)
\end{aligned}$$

Hence when $p > 4 + 2E[Z_1 Z_2(1 - \alpha - I(|Z| \leqslant c))]/E[Z_1^2(1 - \alpha - I(|Z| \leqslant c))]$, by a similar argument as Theorem 1, (7) is greater than zero if a is chosen small enough and b is chosen large enough. □

In Theorems 1 and 2, constants a and b are not specified. In the following, Lemmas 1 and 2 give a rough bound of a when b is given. And b can be chosen as 1 according to simulation results and literature.

Lemma 1. *When $\theta = 0$ and X_{ij}, $j = 1, \ldots, p$, are independent, $\gamma_1(X)$ has a small risk than $1 - \alpha$ if and only if*

$$a \leqslant 2E_0[[I(|t(X)| \leqslant c) - (1 - \alpha)]]/E_0[1/(b + t'(X)t(X))^2].$$

Proof. When $\theta = 0$, (5) can be rewritten as

$$\begin{aligned}
&a(-2E_0[(1 - \alpha - I(|t(X)| \leqslant c))/(b + t'(X)t(X))] \\
&\quad - aE_0[1/(b + t'(X)t(X))^2]). \qquad (8)
\end{aligned}$$

(8) is greater than zero if and only if

$$0 < a \leqslant 2E_0[(I(|t(X)| \leqslant c) - (1 - \alpha))/(b + t'(X)t(X))]/E_0[1/(b + t'(X)t(X))^2].$$

□

Lemma 2. *When* $\|\theta\|$ *goes to infinity and* X_{ij}, $j=1,\ldots,p$, *are independent*, $\gamma_1(X)$ *has a risk smaller than* $1-\alpha$ *asymptotically if*

$$a \leqslant 2(p-4)E_0[t_1^2(X)(1-\alpha-I(|t(X)|\leqslant c))].$$

Proof. If $\|\theta\|$ goes to infinity, then

$$\begin{aligned}(6) &\approx 2aE_0[t_1^2(X)(1-\alpha-I(|t(X)|\leqslant c))](p-4)/|\theta|^4 - a^2/|\theta|^4\\ &= a\{2(p-4)E_0[t_1^2(X)(1-\alpha-I(|t(X)|\leqslant c))]-a\}/|\theta|^4.\end{aligned}$$

Thus (4) is greater than zero asymptotically if

$$a \leqslant 2(p-4)E_0[t_1^2(X)(1-\alpha-I(|t(X)|\leqslant c))]. \qquad \square$$

When $X_{ij}-\theta_j$, $j=1,\ldots,p$, are independent, combining Lemmas 1 and 2,

$$\begin{aligned}\ell = \min\{&2E_0[(I(|t(X)|\leqslant c)-(1-\alpha))/(b+t'(X)t(X))]/E_0[1/(b+t'(X)t(X))^2],\\ &2(p-4)E_0[t_1^2(X)(1-\alpha-I(|t(X)|\leqslant c))]\}\end{aligned}$$

is a rough bound of a. Although it is not a exact bound of a, it can provide information in determining a. When $X_{ij}-\theta_j$ are pairwise dependent as in Theorem 2, a rough bound of a is

$$\begin{aligned}\ell = \min\{&2E_0[(I(|t(X)|\leqslant c)-(1-\alpha))/(b+t'(X)t(X))]/E_0[1/(b+t'(X)t(X))^2],\\ &2[p-4-2E_0(t_1(X)t_2(X)(1-\alpha-I(|t(X)|\leqslant c)))]\\ &\times E_0[t_1^2(X)(1-\alpha-I(|t(X)|\leqslant c))]\}.\end{aligned}$$

3. Admissibility results when $p \leqslant 4$

In Section 2, the scenarios for $p \geqslant 5$ are discussed. In this section, the admissibility results of confidence coefficient for $p \leqslant 4$ are concluded in Theorem 3. Moreover, here we assume that the confidence coefficient is a Bayes estimator of (1) with respect to the noninformative prior $\pi(\theta)=1$. For t distribution, if $t(X)$ is chosen to be the mean vector of observations, then $1-\alpha$ is a Bayes estimator of (1) with respect to prior $\pi(\theta)=1$. The proof will be by Blyth's method (1951), which is a sufficient condition for admissibility. Another version of this method can be found in Berger (1985) and Brown (1971).

Theorem (Berger, 1985). *Consider a decision problem in which* Θ *is a nondegenerate convex subset of Euclidean space* (*i.e.*, Θ *has positive Lebesgue measure*), *and in which the decision rules with continuous risk functions form a complete class. Then an estimator* δ_0 (*with a continuous risk function*) *is admissible if there exists a* *sequence* $\{\pi_n\}$ *of (generalized) priors such that*

(a) *the Bayes risks* $r(\pi_n, \delta_0)$ *and* $r(\pi_n, \delta^n)$ *are finite for all* n, *where* δ^n *is the Bayes rule with respect to* π_n;

(b) *for any nondegenerate convex set* $C \subset \Theta$, *there exists a* $K > 0$ *and an integer* N *such that, for* $n \geqslant N$,

$$\int_C \mathrm{d}F^{\pi_n}(\theta) \geqslant K;$$

(c) $\lim_{n\to\infty} [r(\pi_n, \delta_0) - r(\pi_n, \delta^n)] = 0$.

First, consider a sequence of priors $\pi_n(\theta) = \eta_n(|\theta|^2)$, with

$$\eta_n(v) = \begin{cases} 1 & \text{if } 0 \leqslant v \leqslant 1 \\ \left(1 - \dfrac{\ln v}{\ln n}\right)^2 & \text{if } 1 \leqslant v \leqslant n/2 \\ \dfrac{a_n}{4v^2/n^2 - b} & \text{if } n/2 \leqslant v < \infty, \end{cases}$$

$a_n = \ln^3 2/\ln^2 n$, $b = 1 - \ln 2$. Define

$$\Delta_n = \int\int [(1 - \alpha - I(\theta \in C_{t(x)}))^2 - (\delta_{(t)}^{\pi_n} - I(\theta \in C_{t(x)}))^2] h(t - \theta)\pi_n(\theta)\,\mathrm{d}t\,\mathrm{d}\theta,$$

where $\delta_{(t)}^{\pi_n}$ is the Bayes estimator of (1) with respect to prior $\pi_n(\theta)$. Note that $\pi_n(\theta) \to 1$ as $n \to \infty$.

Lemma 3.

$$\int_{C_{t(x)}} \pi_n(\theta)h(t - \theta)\,\mathrm{d}\theta = \pi_n(t)(1 - \alpha) + \mathrm{o}(1). \tag{9}$$

for n *sufficiently large.*

Proof. This proof is referred to Lemma 1 in Wang (1998), where h denotes the density function of a p-dimensional normal random variable. For a density function h satisfying (4), Eq. (9) can be derived by an argument similar that of Lemma 1 of Wang (1998). □

Lemma 4. *There exists a sequence* $B_n \to \infty$ *such that as* $n \to \infty$

$$\Delta_n \leqslant \int_{|t|>B_n} (1 - \alpha - \delta_*^{\pi_n}(t))^2 \int_{|\theta - t| \leqslant |t|/2} \pi_n(\theta)h(t - \theta)\,\mathrm{d}\theta\,\mathrm{d}t + \mathrm{o}(1),$$

where

$$\delta_*^{\pi_n}(t) = \frac{\int_{|\theta - t| < c} \pi_n(\theta)h(t - \theta)\,\mathrm{d}\theta}{\int_{|\theta - t| \leqslant |t|/2} \pi_n(\theta)h(t - \theta)\,\mathrm{d}\theta}.$$

The proof of Lemma 4 can be followed as Brown and Hwang (1990) except that inequality (2.4) in Brown and Hwang should be changed to

$$\pi_n(\theta)h(t-\theta) \leqslant kh((|\theta|-B)^+).$$

Lemma 5.

$$(\delta_*^{\pi_n}(t)-(1-\alpha))^2 \leqslant H\left[\left[\frac{\eta_n'(R^2)}{\eta_n(R^2)}\right]^2+\left[\frac{R^2\bar{\bar{\eta}}_n(R^2)}{\eta_n(R^2)}\right]^2+\mathrm{O}\left(h\left(\frac{t}{3}\right)\right)\right],$$

where $R=|t|$, $\bar{\bar{\eta}}_n(u^2)=\sup\{|\eta_n''(v^2)|: v \geqslant u/2\}$, *and* H *is some positive constant.*

Since $h(s)$ satisfies (4), the proof of Lemma 5 can be demonstrated by similar arguments as Lemma 4 in Brown and Hwang (1990).

With Lemmas 3–5, we have the following theorem.

Theorem 3. *The assumptions on* X_i, $i=1,\dots,n$, *are the same as in Theorem* 1, *except that the dimension of* X_i *is* $p \leqslant 4$. *And also assume that the confidence coefficient* $1-\alpha$ *is a Bayes estimator of* (1) *with respect to a prior* 1. *Then* $1-\alpha$ *is an admissible estimator of* (1) *under the loss function* (3).

This proof will be by Blyth's method to show that $\Delta_n \to 0$ as $n \to \infty$. Combining Lemmas 3–5 and then making the change of variables $r=|t|^2$, we need to show that

$$\int_{B_n}^{\infty}\left[\left[\frac{\eta_n'(r)}{\eta_n(r)}\right]^2+\left[\frac{r\bar{\bar{\eta}}_n(r)}{\eta_n(r)}\right]^2\right]\eta_n(r)r^{p/2-1}\,\mathrm{d}r \to 0 \quad \text{as } n\to\infty \tag{10}$$

By straightforward calculation (see the proof of Theorem 6 in Brown and Hwang (1990) and Theorem 1 in Wang (2001)), (10) $\to 0$ as $n \to \infty$ and the proof is completed.

4. Examples

In Theorem 1, it is not difficult to find estimators to satisfy condition (4). The following are some examples.

Example 1. Cauchy distribution: Let $X_i=(X_{i1},\dots,X_{ip})$, $i=1,\dots,n$, be a sample of p dimensional Cauchy distribution with density function $(f_{\theta_1}(x),\dots,f_{\theta_p}(x))$, where $f_\theta(x)=1/[\pi(1+(x-\theta)^2)]$ and $t(X)$ is the median vector of the sample. Suppose that n is odd and $n=2m+1$. Then the density function of $t(X)-\theta$ is

$$h(s)=\prod_{i=1}^{p}\frac{n!}{m!1!m!}[F(s_i)]^m[1-F(s_i)]^m\frac{1}{\pi(1+s_i^2)},$$

see David (1981). Note that when s_i goes to infinity,

$$F(s_i) \to 1$$

and

$$1 - F(s_i) = \mathrm{o}\left(\frac{1}{s_i}\right).$$

The last equality is due to

$$\int_s^\infty \frac{1}{\pi(1+x^2)}\,\mathrm{d}x \leqslant \int_s^\infty \frac{1}{\pi x^2}\,\mathrm{d}x = \frac{1}{\pi s}.$$

When $s \to -\infty$, by a similar argument, we have $F(s) = \mathrm{o}(1/s)$ and $1 - F(s) \to 1$. It leads to

$$h(s) = \prod_{i=1}^{p} \mathrm{o}\left(\frac{1}{s_i}\right)^m \frac{1}{1+s_i^2} = \mathrm{o}\left(\frac{1}{s_i^m + s_i^{2+m}}\right).$$

Then the necessary condition of

$$h(s) = \mathrm{o}\left(\prod_{i=1}^{p} \frac{1}{s_i^4}\right)$$

is $m \geqslant 3$ $(n \geqslant 7)$.

Example 2. t distributions: In Example 1, if the Cauchy distribution is changed to t distribution, $t(X)$ could be chosen to be the median if the number of observations is greater than 7, since the tail of the Cauchy distribution, t distribution with degree of freedom 1, is heavier than t distributions with degrees of freedom greater than 2. For the t distribution with degrees of freedom greater than 3, $t(X)$ could be chosen to be the mean for any number of observations because the distribution of mean is the same as that of one observation distribution and the tail of t distribution satisfies (4) with degrees of freedom greater than 3. Therefore, the normal case shown in Robinson (1979) is a special case of Theorem 1.

Example 3. Uniform distribution: Let $X_i = (X_{i1}, \ldots, X_{ip})$, $i = 1, \ldots, n$, be a sample of p-dimensional uniform distribution with density function $(f_{\theta_1}(x), \ldots, f_{\theta_p}(x))$, where $f_\theta(x) = I_{[\theta-1,\theta+1]}(x)$. Since $f_\theta(x)$ is zero outside the interval $[\theta-1, \theta+1]$, for any estimator $t(X)$ of θ, the density function of $t(X)$ satisfies condition (4).

Example 4. Exponential distribution: Let $X_i = (X_{i1}, \ldots, X_{ip}), i = 1, \ldots, n$, be a sample of p-dimensional exponential distribution with density function $(f_{\theta_1}(x), \ldots, f_{\theta_p}(x))$, where $f_\theta(x) = \mathrm{e}^{-(x-\theta)}$, $x > \theta$. Then the UMVUE estimator of θ_j is $t_j(X) = X_{(1)j} - 1/n$, where $X_{(1)j}$ denotes the minimal value of $(X_{1j}, X_{2j}, \ldots, X_{nj})$. The density function of

$t_j(X)$ is

$$\frac{n!}{(n-1)!} f\left(s+\frac{1}{n}\right)\left(1-F\left(s+\frac{1}{n}\right)\right)^{n-1} \quad \text{if } \theta-\frac{1}{n}<s$$

$$= n\mathrm{e}^{-(s+\frac{1}{n}-\theta)} \int_{s+\frac{1}{n}}^{\infty} \mathrm{e}^{-(t-\theta)}\,\mathrm{d}t$$

$$= -n\mathrm{e}^{-(s+\frac{1}{n}-\theta)} \mathrm{e}^{-(t-\theta)}|_{s+\frac{1}{n}}^{\infty}$$

$$= n\mathrm{e}^{-(s+\frac{1}{n}-\theta)} \mathrm{e}^{-(s+\frac{1}{n}-\theta)} = n\mathrm{e}^{-2(s+\frac{1}{n}-\theta)}$$

$$= \mathrm{o}\left(\frac{1}{s^4}\right) \quad \text{as } |s| \to \infty.$$

Therefore, for Examples 3 and 4, $\gamma_1(X)$ can be used as an indicator of θ belonging to $C_{t(X)}$ for all sample size n if $p \geqslant 5$.

5. Simulation results

In Section 2, the form of improved estimators $\gamma_1(X)$ has been provided. However, the coefficients of $\gamma_1(X)$ are not specified in the proof of Theorem 1. Although the values of the coefficients are hard to provide by theoretical deduction, they can be obtained from statistical simulations. Figs. 1–4 show the values of the ratios of the mean squared error $R(\gamma_1(X), I(\theta \in C_{t(X)}))$ of improved estimators and the mean squared error $R(1-\alpha, I(\theta \in C_{t(X)}))$ of confidence coefficients through the norm of θ for the cases that $X_{ij}-\theta$, $j=1,\ldots,p$, are independent and identical t distributions with degrees of freedom k. The improvement is seen to be substantial when the norm of θ is not large. It can also be seen that the improvement of the new estimators for estimating

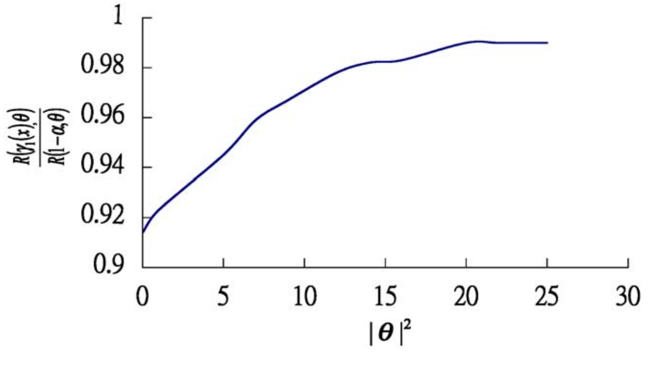

Fig. 1. The plot of $R(\gamma_1(x),\theta)/R(1-\alpha,\theta)$ through $|\theta|^2$ with $k=1$, $c=3.74$, $1-\alpha=0.8$, $n=5$, $p=10$, $a=0.6$, $b=1$, $\ell=1.48753$ and $t(x)$ is the median.

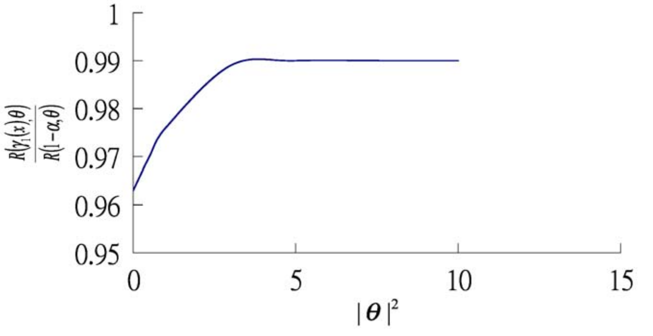

Fig. 2. The plot of $R(\gamma_1(x),\theta)/R(1-\alpha,\theta)$ through $|\theta|^2$ with $k=4$, $c=2.16$, $1-\alpha=0.9$, $n=6$, $p=8$, $a=0.2$, $b=1$, $\ell=0.31214$ and $t(x)$ is the mean.

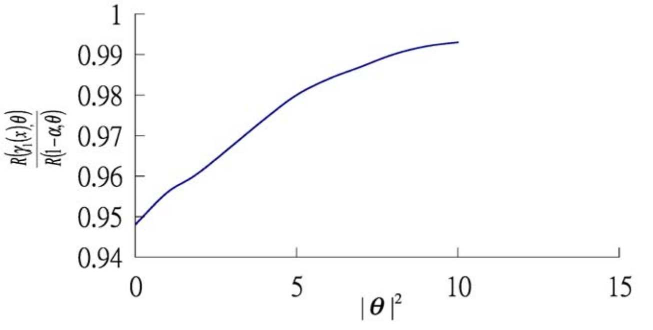

Fig. 3. The plot of $R(\gamma_1(x),\theta)/R(1-\alpha,\theta)$ through $|\theta|^2$ with $k=1$, $c=2.81$, $1-\alpha=0.9$, $n=6$, $p=5$, $a=0.2$, $b=1$, $\ell=0.28233$ and $t(x)$ is the median.

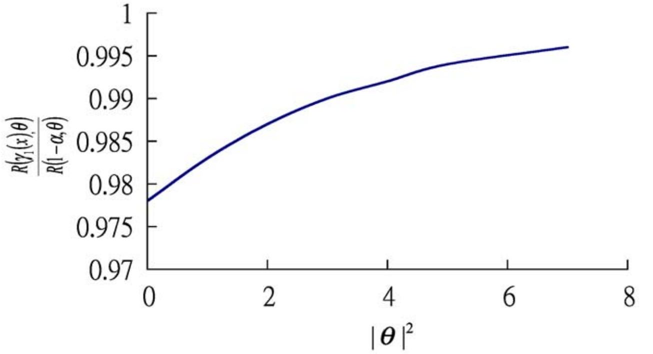

Fig. 4. The plot of $R(\gamma_1(x),\theta)/R(1-\alpha,\theta)$ through $|\theta|^2$ with $k=5$, $c=2.55$, $1-\alpha=0.95$, $n=4$, $p=7$, $a=0.1$, $b=1$, $\ell=0.17163$ and $t(x)$ is the mean.

the coverage function of the confidence set based on the median is more significant than that based on the mean.

All simulations were based on 10,000 replications.

Acknowledgements

The author wishes to thank the Editors and the Referees for their many valuable comments, which have substantially improved this paper.

Appendix

Proof of the following equality

$$e(\theta) = \mathrm{o}\left(\frac{1}{b+|\theta|^2}\right)^2.$$

By definition, the term $E[R(\eta,\theta)(1-\alpha-I(|z|\leqslant c))]$ in $e(\theta)$ can be written as

$$\int_{|\Delta|>\frac{1}{2}(b+|\theta|^2)} R(\eta,\theta)(1-\alpha-I(|z|\leqslant c))f(z)\,\mathrm{d}z + \int_{|\Delta|<\frac{1}{2}(b+|\theta|^2)} R(\eta,\theta)(1-\alpha-I(|z|\leqslant c))f(z)\,\mathrm{d}z = K_1 + K_2,$$

where

$$\Delta = (\eta+\theta)'(\eta+\theta) - |\theta|^2.$$

For $|\Delta| < \frac{1}{2}(b+|\theta|^2)$,

$$\frac{1}{b+|\eta+\theta|^2} = \frac{1}{b+\Delta+|\theta|^2} < \frac{2}{b+|\theta|^2}.$$

Therefore $K_2 = \mathrm{o}(\frac{1}{b+|\theta|^2})^2$.

For $|\Delta| > \frac{1}{2}(b+|\theta|^2)$, we have

$$|\eta|^2 + 2|\eta||\theta| > \frac{1}{2}(b+|\theta|^2),$$

which implies

$$|\eta| > [(b+|\theta|^2)/2 + |\theta|^2]^{1/2} - |\theta| \geqslant \left(\frac{b+|\theta|^2}{24}\right)^{1/2}.$$

Since η is on the line segment joining z to the origin, we have

$$|z| > \left(\frac{b+|\theta|^2}{24}\right)^{1/2} \tag{A.1}$$

for $|\Delta| > \frac{1}{2}(b+|\theta|^2)$.

Now using (9) and the fact that the density of z_i is $\mathrm{o}(1/z_i^4)$ if $|z_i| \to \infty$, for b large enough, we have

$$K_1 = \int_{|\Delta| > \frac{1}{2}(b+|\theta|^2)} R(z,\eta,\theta)(1-\alpha-I(|z| \leqslant c))\mathrm{o}\left(\frac{1}{\prod_{i=1}^{p}(z_i^4)}\right) \mathrm{d}z$$

$$\leqslant \frac{1}{(b+|\theta|^2)^3} \int_{|\Delta| > \frac{1}{2}(b+|\theta|^2)} \left(\sum_{i,j} |z_i|^2|z_j| + \sum_i |z_i|^3\right)$$

$$\times(1-\alpha-I(|z| \leqslant c))\mathrm{o}\left(\frac{1}{\prod_{i=1}^{p} z_i^4}\right) \mathrm{d}z$$

$$= \mathrm{o}\left(\frac{1}{b+|\theta|^2}\right)^2.$$

The last equality holds because the above integration is bounded.

For the other term

$$E\left(\sum_{i=1}^{p} \frac{z_i\theta_i}{(b+|\theta|^2)^2}\right)^2$$

in $e(\theta)$, it is obviously equal to $\mathrm{o}(1/(b+|\theta|^2))^2$. Therefore, combining the above results, we have

$$e(\theta) = \mathrm{o}\left(\frac{1}{b+|\theta|^2}\right)^2.$$

References

Berger, J.O., 1985. Statistical Decision Theory and Bayesian Analysis. 2nd Edition, Springer Series in Statistics. Springer, New York.

Blyth, C.R., 1951. On minimax statistical decision procedures and their admissibility. Ann. Math. Statist. 22, 22–42.

Brown, L.D., 1971. Admissible estimators, recurrent diffusions, and insoluble boundary-value problems. Ann. Math. Statist. 42, 855–903.

Brown, L.D., Hwang, J.T.G., 1990. Admissibility of confidence estimators. In: Chao, M.T., Cheng, P.E. (Eds.), Proceedings of the 1990 Taipei Symposium in Statistics, 1–10.

David, H.A., 1981. Order Statistics. Wiley, New York.

Robert, C., Casella, G., 1994. Improved confidence estimators for the usual multivariate normal confidence set. In: Gupta, S.S., Berger, J.O. (Eds.), Statistical Decision Theory and Related Topics V, Springer, New York, pp. 351–368.

Robinson, G.K., 1979. Conditional properties of statistical procedures for location and scale parameters. Ann. Statist. 7, 756–771.

Wang, H., 1998. Admissibility of the constant coverage probability estimator for estimating the coverage function of certain confidence interval. Statist. Probab. Lett. 36, 365–372.

Wang, H., 1999. Brown's paradox in the estimated confidence approach. Ann. Statist. 27, 610–626.

Wang, H., 2001. Admissibility of confidence estimators in the regression model. J. Multivariate Anal. 76, 267–276.