J. Japan Statist. Soc.
Vol. 34 No. 1 2004 87–99

ASYMPTOTIC CONFIDENCE INTERVALS BASED ON M-PROCEDURES IN ONE- AND TWO-SAMPLE MODELS

Taka-aki Shiraishi*

Asymptotic confidence intervals of location parameters are proposed in one- and two-sample models. These are robust procedures based on scale-invariant M-statistics. The one-sample procedures have the same robustness as Huber's M-estimators. Furthermore although the symmetry of the underlying distribution is needed in the asymptotic theory of Huber's M-estimators, the proposed procedures do not demand the symmetry in the two-sample model. The asymptotic efficiency of the proposed confidence intervals is given by a numerical integration.

Key words and phrases: Asymptotics, confidence region, M-estimators, robustness.

1. Introduction

Let $X_1, \dots, X_n$ be a random sample from an absolutely continuous distribution function $F((x-\mu)/\sigma)$. We denote the density of $F(x)$ by $f(x)$. For convenience, we assume

$$\int_{-\infty}^{\infty} tf(t)dt = 0 \quad \text{and} \quad \int_{-\infty}^{\infty} t^2 f(t)dt = 1. \tag{1.1}$$

Huber (1964) proposed solution $\hat{\theta} = \theta$ of the equation:

$$\sum_{i=1}^{n} \psi(X_i - \theta) = 0 \tag{1.2}$$

as an estimator of μ and called it M-estimator, where $\psi(x)$ is monotone increasing and strictly negative (positive) for large negative (positive) values of x. Furthermore, he showed that the M-estimator given by taking $\psi(x) = \max\{\min\{x, c_0\}, -c_0\}$ for some positive constant c_0 has the minimax asymptotic variance among a class of estimators defined by the solution of (1.2) through the function $\psi(\cdot)$ over the class of distributions that the underlying distribution is in ϵ-contamination neighborhood of a normal distribution. Huber (1981) reviewed further progressive results of M-estimators. Shiraishi (2003) showed that (i) the M-estimator is a little less efficient than the sample mean for the case where the underlying distribution is normal, and that (ii) the M-estimator is more efficient than the sample mean for the case where the underlying distribution is not normal. Jurečková and Sen (1996) discussed robust confidence intervals based on M-statistics. However

Received December 26, 2003. Revised March 19, 2004. Accepted March 24, 2004.

*Department of Mathematical Sciences, Yokohama City University, Seto 22-2, Kanazawa-ku, Yokohama 236-0027, Japan.

the statistics are not scale-invariant. Since the scale parameter of the underlying distribution is unknown in the data analysis, we discuss robust confidence intervals which are scale-invariant. The optimum choice of parameter for confidence interval is discussed by using a Monte Carlo simulation.

Next let $X_1, \dots, X_{n_1}$ and $Y_1, \dots, Y_{n_2}$ be two samples from populations with absolutely continuous distribution functions $F((x-\mu_1)/\sigma)$ and $F((x-\mu_2)/\sigma)$ respectively. Shiraishi (1996) proposed scale-invariant M-estimators for difference of the two means $\delta = \mu_1 - \mu_2$. The asymptotic property of the proposed estimators was discussed. For the two-way layouts, Shiraishi (1993, 1998, 1999, 2001) discussed scale-invariant M-estimators of location parameters. In the two-sample model, we discuss asymptotic confidence intervals for δ, based on Shiraishi's M-estimators. The proposed statistics are scale-invariant. Furthermore although the symmetry of the underlying distribution is needed in the asymptotic theory of Huber's M-estimators, the proposed procedures do not demand the symmetry in the two-sample model.

Lastly the asymptotic efficiency of the proposed procedures relative to the classical normal procedures is expressed, and it is calculated by numerical value integration. It can be seen that the proposed procedures are more efficient than the classical normal procedures except for the case where the underlying distribution is normal. Especially, the proposed procedures are fairly efficient for the asymmetric distributions in the two-sample model.

The present paper generalizes the confidence intervals stated in textbook of Shiraishi (2003). Furthermore this gives the proof for the asymptotic results of the textbook.

2. One-sample confidence interval

For function $\Psi(x)$ defined on R and for constants Δ, ω and $\rho > 0$, let us put

(2.1)
$$W(\Delta,\omega) = \sum_{i=1}^{n} \left\{ \Psi\left(\frac{X_i - \mu - \Delta/\sqrt{n}}{\rho e^{\omega/\sqrt{n}}}\right) - \Psi\left(\frac{X_i - \mu}{\rho}\right) \right\} \Big/ \sqrt{n} + d(\Psi)\Delta/\sigma + e(\Psi)\omega,$$

and

(2.2)
$$W^*(\Delta,\omega) = \sum_{i=1}^{n} \left\{ \Psi\left(\frac{X_i - \mu - \Delta/\sqrt{n}}{\rho e^{\omega/\sqrt{n}}}\right) - \Psi\left(\frac{X_i - \mu}{\rho}\right) \right\} \Big/ \sqrt{n} + d(\Psi)\Delta/\sigma,$$

respectively, where

$$d(\Psi) = -\int_{-\infty}^{\infty} \Psi(\sigma x/\rho) f'(x)dx, \qquad (2.3)$$

and

$$e(\Psi) = -\int_{-\infty}^{\infty} \Psi(\sigma x/\rho) \left\{1 + \frac{xf'(x)}{f(x)}\right\} f(x)dx. \tag{2.4}$$

We impose the following conditions.

(c.1); $f(x)$ is symmetrical about 0, i.e.,

$$f(-x) = f(x). \tag{2.5}$$

(c.2); $f(x)$ have finite Fisher's informations, i.e.,

$$0 < \int_{-\infty}^{\infty} \{-f'(x)/f(x)\}^2 f(x)dx < \infty$$

and

$$0 < \int_{-\infty}^{\infty} \{-1 - xf'(x)/f(x)\}^2 f(x)dx < \infty.$$

(c.3); $\Psi(x) = \Psi_1(x) + \Psi_2(x)$, $\Psi_1(x)$ is nondecreasing and $\Psi_2(x)$ is nonincreasing. There exists a constant c such that $\Psi(x) = \Psi(-c)$ for $x \le -c$; $= \Psi(c)$ for $x \ge c$.

If (c.2) is satisfied, from Shiraishi (1989), the densities $\{\prod_{k=1}^{n}[1/(\sigma e^{\omega/\sqrt{n}}) f((x_k - \Delta/\sqrt{n})/(\sigma e^{\omega/\sqrt{n}}))]\}$ are contiguous to the densities $\{\prod_{k=1}^{n}[(1/\sigma) f(x_k/\sigma)]\}$ as n tends to infinity.

Proceeding as in the proof of Lemma 3.1 of Shiraishi (1996), we get Theorem 2.1.

THEOREM 2.1. *Let* $(X_1, \dots, X_n)$ *have joint distribution function* $\prod_{k=1}^{n} F(x_k/\sigma)$. *Then under the conditions* (c.1)–(c.3), *we have, for any positive* C_1, C_2, *and* ϵ,

$$\lim_{n\to\infty} P\left\{\sup_{|\Delta|<C_1, |\omega|<C_2} |W(\Delta,\omega)| > \epsilon\right\} = 0.$$

Furthermore, we get Corollary 2.2.

COROLLARY 2.2. *Suppose that* $\Psi(x)$ *is skew symmetrical, i.e.,* $\Psi(-x) = -\Psi(x)$. *Then under the assumptions of Theorem* 2.1, *we have, for any positive* C_1, C_2, *and* ϵ,

$$\lim_{n\to\infty} P\left\{\sup_{|\Delta|<C_1, |\omega|<C_2} |W^*(\Delta,\omega)| > \epsilon\right\} = 0.$$

PROOF. The condition (c.1) and the skew symmetry of $\Psi(x)$ give $e(\Psi) = 0$. Combining this fact with Theorem 2.1, we get the conclusion. □

Also from a direct application of Corollary 2.2, we have Corollary 2.3.

COROLLARY 2.3. *Suppose that $\Psi(x)$ is skew symmetrical. Then under the assumptions of Theorem 2.1, we get, for any positive C_1, C_2, and ϵ,*

$$\lim_{n\to\infty} P\left\{\sup_{|\Delta_1|<C_1,|\omega|<C_2} |W^*(\Delta_1+\Delta_2,\omega)| > \epsilon\right\} = 0.$$

Let us put

$$T_M(\mu) = \sum_{i=1}^{n} \psi\left(\frac{X_i-\mu}{\hat{\sigma}_n}\right),$$

where $\hat{\sigma}_n$ is a consistent estimator of some constant $\rho > 0$ satisfying the condition (c.4).

(c.4); $\sqrt{n}(\hat{\sigma}_n - \rho) = O_p(1)$.

We refer to the solution $\hat{\mu}_n$ of $T_M(\mu) = 0$ as M-estimator. We impose the condition (c.5) on $\psi(\cdot)$.

(c.5); $\psi(x)$ is nondecreasing and skew symmetrical. There exists a constant c such that $\psi(x) = \psi(-c)$ for $x \le -c$; $= \psi(c)$ for $x \ge c$. $d(\psi) = -\int_{-\infty}^{\infty} \psi(\sigma x/\rho) f'(x)dx > 0$.

Let us define the solution θ of the following equation by $\hat{\theta}_n$.

$$\sum_{i=1}^{n} \psi\left(\frac{X_i-\mu}{\rho}\right) \Big/ \sqrt{n} = \sqrt{n}d(\psi)(\theta-\mu)/\sigma.$$

By using Corollary 2.2 given by $\Psi(x) = \psi(x)$, along the lines on the proofs of Lemma 4.1–4.5 of Jurečková (1971), we can show

$$\sqrt{n}|\hat{\mu}_n - \hat{\theta}_n| \approx 0,$$

where $A_n \approx B_n$ denotes $A_n - B_n \xrightarrow{\text{P}} 0$ and $\xrightarrow{\text{P}}$ denotes convergence in probability. Therefore we get

(2.6)

$$\sqrt{n}(\hat{\mu}_n - \mu) \approx (\sigma/d(\psi)) \sum_{i=1}^{n} \psi\left(\frac{X_i-\mu}{\rho}\right) \Big/ \sqrt{n} \xrightarrow{\text{L}} N\left(0, c(\psi,f)\sigma^2/d^2(\psi)\right),$$

where

$$c(\psi, f) = \int_{-\infty}^{\infty} \psi^2(\sigma x/\rho) f(x)dx. \tag{2.7}$$

Let us put

$$\hat{\eta}_n = \{T_M(\hat{\mu}_n - \Delta/\sqrt{n}) - T_M(\hat{\mu}_n + \Delta/\sqrt{n})\}/(2\sqrt{n}\Delta).$$

Then by applying $\Psi(\cdot) = \psi(\cdot)$ and $\Delta_2 = \Delta$ in Corollary 2.3, we get

$$\begin{aligned}
(2.8)\qquad & P\{|W^*(\sqrt{n}(\hat{\mu}_n - \mu) + \Delta, \sqrt{n}(\hat{\sigma}_n - \rho))| > \epsilon\} \\
&\le P\{|W^*(\sqrt{n}(\hat{\mu}_n - \mu) + \Delta, \sqrt{n}(\hat{\sigma}_n - \rho))| > \epsilon, \\
&\qquad\qquad \sqrt{n}|\hat{\mu}_n - \mu| < C_1,\ \sqrt{n}|\hat{\sigma}_n - \rho| < C_2\} \\
&\quad + P\{\sqrt{n}|\hat{\mu}_n - \mu| \ge C_1\} + P\{\sqrt{n}|\hat{\sigma}_n - \rho| \ge C_2\} \\
&\le P\left\{\sup_{|\Delta_1|<C_1, |\omega|<C_2} |W^*(\Delta_1 + \Delta, \omega)| > \epsilon\right\} \\
&\quad + P\{\sqrt{n}|\hat{\mu}_n - \mu| \ge C_1\} + P\{\sqrt{n}|\hat{\sigma}_n - \rho| \ge C_2\}.
\end{aligned}$$

By choosing C_1, C_2 and n sufficiently large for any positive ϵ and ϵ_0, Corollary 2.2, the condition (c.4), and (2.6) give

$$(\text{the right hand side of } (2.8)) \le \epsilon_0.$$

Hence, we have

$$\sum_{i=1}^{n} \left\{\psi\left(\frac{X_i - \hat{\mu}_n + \Delta/\sqrt{n}}{\hat{\sigma}_n}\right) - \psi\left(\frac{X_i - \mu}{\rho}\right)\right\} \Big/ \sqrt{n} + d(\psi)\{\sqrt{n}(\hat{\mu}_n - \mu) - \Delta\}/\sigma \xrightarrow{\text{P}} 0$$

and

$$\sum_{i=1}^{n} \left\{\psi\left(\frac{X_i - \hat{\mu}_n - \Delta/\sqrt{n}}{\hat{\sigma}_n}\right) - \psi\left(\frac{X_i - \mu}{\rho}\right)\right\} \Big/ \sqrt{n} + d(\psi)\{\sqrt{n}(\hat{\mu}_n - \mu) + \Delta\}/\sigma \xrightarrow{\text{P}} 0,$$

which imply

$$(2.9)\qquad \hat{\eta}_n \xrightarrow{\text{P}} d(\psi)/\sigma.$$

Moreover, let us put

$$\hat{c}_n(\psi, f) = \frac{1}{n}\sum_{i=1}^{n} \psi^2\left(\frac{X_i - \hat{\mu}_n}{\hat{\sigma}_n}\right).$$

Then by applying $\Psi(x) = \{\psi(x)\}^2$ to Theorem 2.1, we get

$$\begin{aligned}
&\sum_{i=1}^{n} \psi^2\left(\frac{X_i - \hat{\mu}_n}{\hat{\sigma}_n}\right) \Big/ \sqrt{n} \\
&\approx \sum_{i=1}^{n} \psi^2\left(\frac{X_i - \mu}{\rho}\right) \Big/ \sqrt{n} - \sqrt{n}d(\psi^2)(\hat{\mu}_n - \mu)/\sigma - \sqrt{n}e(\psi^2)(\log\hat{\sigma}_n - \log\rho).
\end{aligned}$$

Hence $\hat{\mu}_n \xrightarrow{\text{P}} \mu$ and $\hat{\sigma}_n \xrightarrow{\text{P}} \rho$ imply

$$\hat{c}_n(\psi, f) \approx \sum_{i=1}^{n} \psi^2\left(\frac{X_i - \mu}{\rho}\right) \Big/ n \xrightarrow{\text{P}} c(\psi, f). \tag{2.10}$$

From (2.6), (2.9) and (2.10), we get

$$\frac{\sqrt{n}\hat{\eta}_n}{\sqrt{\hat{c}_n(\psi, f)}}(\hat{\mu}_n - \mu) \xrightarrow{\text{L}} N(0, 1),$$

where $\xrightarrow{\text{L}}$ denotes convergence in law.

Hence we can obtain Theorem 2.4.

Theorem 2.4. *Suppose that the conditions* (c.1), (c.2), (c.4) *and* (c.5) *are satisfied. Then*

$$\left(\hat{\mu}_n - \frac{\sqrt{\hat{c}_n(\psi, f)}z_{(\alpha/2)}}{\sqrt{n}\hat{\eta}_n},\ \hat{\mu}_n + \frac{\sqrt{\hat{c}_n(\psi, f)}z_{(\alpha/2)}}{\sqrt{n}\hat{\eta}_n}\right)$$

is an asymptotically distribution-free $100(1 - \alpha)$ *percent confidence interval for* μ, *where* $z_{(\alpha/2)}$ *is the upper* $100 \times (\alpha/2)$ *percentile of the standard normal distribution.*

Theorem 2.4 implies that the asymptotic confidence interval does not depend on Δ. However we must decide the value of Δ. Hence a simulation study for the goodness of $\hat{\eta}_n$ estimating $d(\psi)/\sigma$ is done, based on

$$\psi(x) = \max\{\min\{x, 1.399\}, -1.399\} \tag{2.11}$$

and

$$\hat{\sigma}_n = \frac{1}{\Phi^{-1}(0.75)} \cdot \text{med}\{|X_1 - \text{med}(X)|, \dots, |X_n - \text{med}(X)|\}, \tag{2.12}$$

where $\text{med}(X)$ denotes the sample median among $\{X_1, \dots, X_n\}$, and $\Phi(x)$ denotes the standard normal distribution function. From Table 5.2 of Shiraishi (2003), this score function $\psi(x)$ is approximately the optimum choice which gives the minimax asymptotic variance on 0.05-contaminated normal neighborhood $\{f(x) = 0.95\phi(x) + 0.05h(x) : \phi(x)$ is a standard normal density and $h(x)$ is any symmmetric density$\}$. Welsh (1986) showed that $\hat{\sigma}_n$ satisfies (c.4) for $\rho = \sigma F^{-1}(0.75)/\Phi^{-1}(0.75)$. Hence $\hat{\sigma}_n$ is a consistent estimator of ρ. Further discussion for $\hat{\sigma}_n$ is seen in Ando and Kimura (2003) and Rousseeuw and Croux (1993). The underlying distributions $F(x)$ chosen here are normal; $N(0, 1)$, logistic distribution, contaminated normal; $0.95N(0, 5/7) + 0.05N(0, 45/7)$, and double exponential. $\hat{\eta}_n$ depends on Δ. From (2.9), $\hat{\eta}_n$ is a consistent estimator for $\eta = d(\psi)/\sigma$. We simulate the mean squared error of $\hat{\eta}_n$ (MSE) given by $E\{(\hat{\eta}_n - \eta)^2\}$ in Table 1 for $n = 20, 30, 50$ and $\Delta = 3.5, 7.0$ (0.5). The values of the MSE are estimated by Monte-Carlo simulation from $2,000$ samples. From Table 1, we may decide $\Delta = 5.5$ as the best choice.

Table 1. The simulated mean squared error of $\hat{\eta}_n$.

(i) $F(x)$ = normal

$n = 20$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0225	0.0209	0.0179	0.0157	0.0167	0.0184	0.0200	0.0244
$n = 30$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0195	0.0154	0.0148	0.0131	0.0126	0.0128	0.0137	0.0145
$n = 50$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0103	0.0102	0.0098	0.0092	0.0089	0.0089	0.0082	0.0098

(ii) $F(x)$ = logistic

$n = 20$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0307	0.0233	0.0209	0.0218	0.0217	0.0261	0.0315	0.0377
$n = 30$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0238	0.0198	0.0183	0.0172	0.0163	0.0170	0.0180	0.0221
$n = 50$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0154	0.0147	0.0133	0.0134	0.0118	0.0110	0.0114	0.0123

(iii) $F(x)$ = contaminated normal

$n = 20$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0265	0.0252	0.0219	0.0217	0.0264	0.0325	0.0371	0.0467
$n = 30$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0239	0.0195	0.0184	0.0172	0.0177	0.0192	0.0224	0.0255
$n = 50$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0140	0.0134	0.0126	0.0122	0.0119	0.0123	0.0122	0.0149

(iv) $F(x)$ = double exponential

$n = 20$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0408	0.0345	0.0343	0.0399	0.0455	0.0593	0.0735	0.0897
$n = 30$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0388	0.0320	0.0284	0.0277	0.0281	0.0321	0.0384	0.0490
$n = 50$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0294	0.0268	0.0235	0.0226	0.0201	0.0188	0.00206	0.0225

3. Two-sample confidence interval

We suppose (1.1). Then we get $E(X_i) = \mu_1$, $E(Y_j) = \mu_2$, and $V(X_i) = V(Y_j) = \sigma^2$. We do not impose the symmetry on $f(x)$.

Let us put

$$W_1(\Delta_1 + \Delta_2, \omega) = \frac{\sqrt{n}}{n_1} \sum_{i=1}^{n_1} \left\{ \Psi\left(\frac{X_i - \mu_1 - (\Delta_1 + \Delta_2)/\sqrt{n}}{\rho e^{\omega/\sqrt{n}}} \right) - \Psi\left(\frac{X_i - \mu_1}{\rho} \right) \right\} + d(\Psi)(\Delta_1 + \Delta_2)/\sigma + e(\Psi)\omega,$$

$$
\begin{aligned}
&W_2\left(\Delta_1 - \frac{n_1}{n_2}\Delta_2, \omega\right) \\
&= \frac{\sqrt{n}}{n_2}\sum_{j=1}^{n_2}\left\{\Psi\left(\frac{Y_j - \mu_2 - (\Delta_1 - \frac{n_1}{n_2}\Delta_2)/\sqrt{n}}{\rho e^{\omega/\sqrt{n}}}\right) - \Psi\left(\frac{Y_j - \mu_2}{\rho}\right)\right\} \\
&\quad + d(\Psi)\left(\Delta_1 - \frac{n_1}{n_2}\Delta_2\right)\Big/\sigma + e(\Psi)\omega,
\end{aligned}
$$

and

$$
W(\Delta_1, \Delta_2, \omega) = W_1(\Delta_1 + \Delta_2, \omega) - W_2\left(\Delta_1 - \frac{n_1}{n_2}\Delta_2, \omega\right),
$$

where $n = n_1 + n_2$, $d(\Psi)$ and $e(\Psi)$ are defined by (2.3) and (2.4) respectively.

We add the condition

(c.6); $0 < \lim_{n\to\infty} n_1/n = \lambda < 1$.

Then from the discussion similar to the proof of Theorem 2.1, we can derive Lemma 3.1.

LEMMA 3.1. *Suppose that* (c.2), (c.3) *and* (c.6) *are satisfied. Then we get, for positive* $C_1,\ C_2,\ C_3,$ *and* ϵ,

$$
\lim_{n\to\infty} P\left\{\sup_{|\Delta_1|<C_1, |\Delta_2|<C_2, |\omega|<C_3} |W_1(\Delta_1 + \Delta_2, \omega)| > \epsilon\right\} = 0,
$$

and

$$
\lim_{n\to\infty} P\left\{\sup_{|\Delta_1|<C_1, |\Delta_2|<C_2, |\omega|<C_3} \left|W_2\left(\Delta_1 - \frac{n_1}{n_2}\Delta_2, \omega\right)\right| > \epsilon\right\} = 0.
$$

Hence from a direct application of Lemma 3.1, we have Theorem 3.2.

THEOREM 3.2. *Suppose that the assumptions of Lemma* 3.1 *are satisfied. Then we get, for positive* $C_1,\ C_2,\ C_3,$ *and* ϵ,

$$
\lim_{n\to\infty} P\left\{\sup_{|\Delta_1|<C_1, |\Delta_2|<C_2, |\omega|<C_3} |W(\Delta_1, \Delta_2, \omega)| > \epsilon\right\} = 0.
$$

Let us put

$$
T_M^*(\theta) = \frac{1}{n_1}\sum_{i=1}^{n_1}\psi\left(\frac{X_i - \tilde{\mu} - \theta}{\hat{\sigma}_n}\right) - \frac{1}{n_2}\sum_{j=1}^{n_2}\psi\left(\frac{Y_j - \tilde{\mu} + \left(\frac{n_1}{n_2}\right)\cdot\theta}{\hat{\sigma}_n}\right),
$$

where $\tilde{\mu} = (n_1\bar{X} + n_2\bar{Y})/n$ and $\bar{X} = \sum_{i=1}^{n_1} X_i/n_1$ and $\bar{Y} = \sum_{j=1}^{n_2} Y_j/n_2$.

We can derive the solution $\breve{\theta}_n$ of $T_M^*(\theta) = 0$. Shiraishi (2003) proposed $\breve{\delta}_n = (1 + n_1/n_2) \cdot \breve{\theta}_n$ as a robust estimator of $\delta = \mu_1 - \mu_2$. $\hat{\sigma}_n$ is a consistent estimator of some constant ρ satisfying (c.4). By applying $\Psi(x) = \psi(x)$ to Theorem 3.2, we get

$$\begin{aligned} 0 &= \sqrt{n} T_M(\breve{\theta}_n) \\ &\approx \frac{\sqrt{n}}{n_1} \sum_{i=1}^{n_1} \left\{ \psi\left(\frac{X_i - \mu_1}{\rho} \right) - \bar{\psi} \right\} - \frac{\sqrt{n}}{n_2} \sum_{j=1}^{n_2} \left\{ \psi\left(\frac{Y_j - \mu_2}{\rho} \right) - \bar{\psi} \right\} \\ &\quad - \sqrt{n} d(\psi)(\breve{\delta}_n - \delta)/\sigma, \end{aligned}$$

where $\bar{\psi} = \int_{-\infty}^{\infty} \psi(\sigma x/\rho) dF(x)$.

Hence we have

$$\begin{aligned} \sqrt{n}(\breve{\delta}_n - \delta) &\approx (\sigma/d(\psi)) \\ &\quad \times \left[\frac{\sqrt{n}}{n_1} \sum_{i=1}^{n_1} \left\{ \psi\left(\frac{X_i - \mu_1}{\rho} \right) - \bar{\psi} \right\} - \frac{\sqrt{n}}{n_2} \sum_{j=1}^{n_2} \left\{ \psi\left(\frac{Y_j - \mu_2}{\rho} \right) - \bar{\psi} \right\} \right] \\ &\xrightarrow{\text{L}} N(0, c^*(\psi, f)\sigma^2/\{\lambda(1-\lambda)d^2(\psi)\}), \end{aligned} \tag{3.1}$$

where $c^*(\psi, f) = \int_{-\infty}^{\infty} \{\psi(\sigma x/\rho) - \bar{\psi}\}^2 f(x) dx$. Let us put

$$\breve{\eta}_n = \sqrt{n}\{T_M(\breve{\theta}_n - \Delta/\sqrt{n}) - T_M(\breve{\theta}_n + \Delta/\sqrt{n})\} \Big/ \left\{ 2\left(1 + \frac{n_1}{n_2}\right)\Delta \right\}.$$

Then by applying $\Psi(x) = \psi(x)$ to Theorem 3.2, as in the proof of (2.9), we get

$$\breve{\eta}_n \xrightarrow{\text{P}} d(\psi)/\sigma. \tag{3.2}$$

Let us put

$$\begin{aligned} &\breve{c}_n(\psi, f) \\ &= \frac{1}{n}\left[\sum_{i=1}^{n_1} \left\{ \psi\left(\frac{X_i - \bar{X}}{\hat{\sigma}_n} \right) - \bar{\psi}(X,Y) \right\}^2 + \sum_{j=1}^{n_2} \left\{ \psi\left(\frac{Y_j - \bar{Y}}{\hat{\sigma}_n} \right) - \bar{\psi}(X,Y) \right\}^2 \right], \end{aligned}$$

where

$$\bar{\psi}(X,Y) = \frac{1}{n}\left\{ \sum_{i=1}^{n_1} \psi\left(\frac{X_i - \bar{X}}{\hat{\sigma}_n} \right) + \sum_{j=1}^{n_2} \psi\left(\frac{Y_j - \bar{Y}}{\hat{\sigma}_n} \right) \right\}.$$

Then by the discussion similar to the proof of (2.10), we find

$$\breve{c}_n(\psi, f) \xrightarrow{\text{P}} c^*(\psi, f). \tag{3.3}$$

From (3.1)–(3.3), we get

$$\frac{\sqrt{n_1 n_2}\breve{\eta}_n}{\sqrt{n\breve{c}_n(\psi, f)}}(\breve{\delta}_n - \delta) \xrightarrow{\text{L}} N(0,1).$$

Hence we can obtain Theorem 3.3.

THEOREM 3.3. *Suppose that the conditions* (c.2), (c.4), (c.5) *and* (c.6) *are satisfied. Then*

$$\left(\breve{\delta}_n - \frac{\sqrt{n\breve{c}_n(\psi, f)}z_{(\alpha/2)}}{\sqrt{n_1 n_2}\breve{\eta}_n},\ \breve{\delta}_n + \frac{\sqrt{n\breve{c}_n(\psi, f)}z_{(\alpha/2)}}{\sqrt{n_1 n_2}\breve{\eta}_n}\right)$$

is an asymptotically distribution-free $100(1-\alpha)$ *percent confidence interval for* δ.

Theorem 3.3 implies that the asymptotic confidence interval does not depend on Δ. However we must decide the value of Δ. Hence a simulation study for the goodness of $\hat{\eta}_n$ estimating $d(\psi)/\sigma$ is done, based on $\psi(x)$ given by (2.11) and

$$\hat{\sigma}_n = \frac{\sqrt{\pi}}{\sqrt{2}\cdot n}\sum_{i=1}^{n}|Z_i|, \tag{3.4}$$

where we define $Z_1, \dots, Z_n$ by

$$Z_i = \begin{cases} X_i - \bar{X} & (i = 1, \dots, n_1) \\ Y_{i-n_1} - \bar{Y} & (i = n_1 + 1, \dots, n). \end{cases} \tag{3.5}$$

Table 2. The simulated mean squared error of $\breve{\eta}_n$.

(i) $F(x) =$ normal

$n_1 = n_2 = 20$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0104	0.0089	0.0088	0.0082	0.0083	0.0081	0.0086	0.0091
$n_1 = n_2 = 30$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0068	0.0063	0.0061	0.0057	0.0058	0.0055	0.0061	0.0061

(ii) $F(x) =$ logistic

$n_1 = n_2 = 20$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0137	0.0127	0.0119	0.0113	0.0106	0.0103	0.0106	0.0111
$n_1 = n_2 = 30$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0099	0.0087	0.0086	0.0079	0.0078	0.0076	0.0075	0.0079

(iii) $F(x) =$ contaminated normal

$n_1 = n_2 = 20$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0154	0.0132	0.0130	0.0123	0.0125	0.0134	0.0140	0.0150
$n_1 = n_2 = 30$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0102	0.0094	0.0090	0.0086	0.0091	0.0092	0.0097	0.0095

(iv) $F(x) =$ double exponential

$n_1 = n_2 = 20$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0222	0.0205	0.0189	0.0177	0.0162	0.0151	0.0150	0.0154
$n_1 = n_2 = 30$	Δ	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
	MSE	0.0162	0.0143	0.0141	0.0129	0.0124	0.0120	0.0114	0.0115

Hence $\hat{\sigma}_n$ is a consistent estimator of $\rho = (\sqrt{\pi}\sigma/\sqrt{2})\int_{-\infty}^{\infty}|x|dF(x)$. $\hat{\sigma}_n$ defined by (3.4) satisfies (c.4) from the relation

$$\sqrt{n}(\hat{\sigma}_n - \rho) \approx \left\{\sum_{i=1}^{n_1}\left(\frac{\sqrt{\pi}}{\sqrt{2}}|X_i - \mu_1| - \rho\right) + \sum_{j=1}^{n_2}\left(\frac{\sqrt{\pi}}{\sqrt{2}}|Y_j - \mu_2| - \rho\right)\right\} \Big/ \sqrt{n}$$
$$+ \sqrt{n}\int_{-\infty}^{\infty}|x|f'(x)dx\left\{(\bar{X} - \mu_1) + (\bar{Y} - \mu_2)\right\}/\sigma.$$

The underlying distributions $F(x)$ chosen here are normal; $N(0,1)$, logistic distribution, contaminated normal; $0.95N(0,5/7) + 0.05N(0,45/7)$, and double exponential. $\hat{\eta}_n$ depends on Δ. From (3.2), $\breve{\eta}_n$ is a consistent estimator for $\eta = d(\psi)/\sigma$. We simulate the mean squared error of $\breve{\eta}_n$ (MSE) given by $E\{(\breve{\eta}_n - \eta)^2\}$ in Table 2 for $n = 20, 30$ and $\Delta = 3.5, 7.0\ (0.5)$. The values of the MSE are estimated by Monte-Carlo simulation from $2,000$ samples. From Table 2, we may decide $\Delta = 5.0$ as the best choice.

4. Asymptotic efficiency

For two sequences of $100(1-\alpha)$ percent confidence intervals $CI_{1n} = (L_{1n}, U_{1n})$ and $CI_{2n} = (L_{2n}, U_{2n})$, we assume

$$(U_{2n} - L_{2n})^2/(U_{1n} - L_{1n})^2 \xrightarrow{\text{P}} \gamma,$$

where γ is a nonnegative constant. Then we define the asymptotic relative efficiency (ARE) of CI_{1n} relative to CI_{2n} by

$$ARE(CI_{1n}, CI_{2n}) = \gamma.$$

(i) One-sample case: CR_1 denotes the robust confidence interval of Theorem 2.4 based on $\psi(x)$ and $\hat{\sigma}_n$ defined by (2.11) and (2.12) respectively. The normal theory confidence interval is given by

$$CR_2 = \left(\bar{X}_n - \frac{\sigma z_{(\alpha/2)}}{\sqrt{n}},\ \bar{X}_n + \frac{\sigma z_{(\alpha/2)}}{\sqrt{n}}\right),$$

where $\bar{X}_n$ denotes the sample mean. Then by using (2.9) and (2.10), $ARE(CR_1, CR_2)$ is equal to

$$d^2(\psi)/c(\psi, f) = \left\{\int_{-\infty}^{\infty}\psi(\sigma x/\rho)f'(x)dx\right\}^2 \Big/ \int_{-\infty}^{\infty}\psi^2(\sigma x/\rho)f(x)dx. \tag{4.1}$$

Let us put $\xi = \Phi^{-1}(0.75)/F^{-1}(0.75)$. Then using integration by parts, (4.1) becomes

$$ARE(CR_1, CR_2) = \frac{2\xi^2\{F(c/\xi) - 0.5\}^2}{c^2 - 2\xi^2\displaystyle\int_0^{c/\xi} xF(x)dx}.$$

The values of $ARE(CR_1, CR_2)$ are given in Table 3. The underlying distributions chosen here are normal; $N(0,1)$, logistic distribution, contaminated normal; $0.95N(0,5/7) + 0.05N(0,45/7)$, and double exponential. From Table 3, we can see that the proposed confidence interval is more efficient than the normal theory confidence interval except for the case where the underlying distribution is normal. Under the normal distribution, the proposed confidence interval is nearly efficient to the normal theory confidence interval.

Table 3. The asymptotic relative efficiency of the proposed confidence interval relative to the normal theory confidence interval in one-sample model.

$F(x)$	$ARE(CR_1, CR_2)$
normal	0.955
logistic	1.090
contaminated normal	1.205
double exponential	1.381

(ii) Two-sample case: CR_1^* denotes the robust confidence interval of Theorem 3.3 based on $\psi(x)$ and $\hat{\sigma}_n$ defined by (2.11) and (3.4) respectively. The normal theory confidence interval is given by

$$CR_2^* = \left(\bar{X} - \bar{Y} - \frac{\sqrt{n}\sigma z_{(\alpha/2)}}{\sqrt{n_1 n_2}},\ \bar{X} - \bar{Y} + \frac{\sqrt{n}\sigma z_{(\alpha/2)}}{\sqrt{n_1 n_2}} \right).$$

Then by using (3.2) and (3.3), $ARE(CR_1^*, CR_2^*)$ is equal to

$$\begin{aligned} (4.2) \qquad & d^2(\psi)/c^*(\psi, f) \\ &= \left\{ \int_{-\infty}^{\infty} \psi(\sigma x/\rho) f'(x) dx \right\}^2 \Big/ \int_{-\infty}^{\infty} \left\{ \psi(\sigma x/\rho) - \bar{\psi} \right\}^2 f(x) dx. \end{aligned}$$

Let us put $\xi^* = (\sqrt{\pi}/\sqrt{2}) \int_{-\infty}^{\infty} |x| dF(x)$. Then using integration by parts, (4.2) becomes

$$\begin{aligned} & ARE(CR_1^*, CR_2^*) \\ &= \frac{\xi^* \{F(c/\xi^*) - F(-c/\xi^*)\}^2}{2c \int_{-c/\xi^*}^{c/\xi^*} F(x)dx - 2\xi^* \int_{-c/\xi^*}^{c/\xi^*} xF(x)dx - \xi^* \left\{ \int_{-c/\xi^*}^{c/\xi^*} F(x)dx \right\}^2}. \end{aligned}$$

The values of $ARE(CR_1^*, CR_2^*)$ are given in Table 4. The underlying distributions chosen here are normal; $N(0,1)$, logistic distribution, contaminated normal; $0.95N(0,5/7)+0.05N(0,45/7)$, double exponential, exponential, and asymmetric contaminated normal; $0.95N(-0.1, 0.76064^2) + 0.05N(1.9, 9 \times 0.76064^2)$. From Table 4, we can see that the proposed confidence interval is more efficient than the normal theory confidence interval except for the case where the underlying distribution is normal. Especially its efficiency is remarkably larger than 1 for the asymmetric underlying distributions. Under the normal distribution, the proposed confidence interval is nearly efficient to the normal theory confidence interval.

Table 4. The asymptotic relative efficiency of the proposed confidence interval relative to the normal theory confidence interval in two-sample model.

$F(x)$	$ARE(CR_1^*, CR_2^*)$
normal	0.955
logistic	1.088
contaminated normal	1.208
double exponential	1.307
asymmetric contaminated normal	1.417
exponential	1.536

Acknowledgements

The author would like to thank the referees for their valuable comments.

References

Ando, M. and Kimura, M. (2003). A characterization of the neighborhoods defined by certain special capacities and their applications to bias-robustness of estimates, *J. Statist. Plann. Infer.*, **116**, 61–90.

Huber, P. J. (1964). Robust estimation of a location parameter, *Ann. Math. Statist.*, **35**, 73–101.

Huber, P. J. (1981). *Robust Statistics*, John Wiley, New York.

Jurečková, J. (1971). Nonparametric estimate of regression coefficients, *Ann. Math. Statist.*, **42**, 1328–1338.

Jurečková, J. and Sen, P. K. (1996). *Robust Statistical Procedures: Asymptotics and Interrelations*, John Wiley, New York.

Rousseeuw, P. J. and Croux, C. (1993). Alternatives to the median absolute deviation, *J. Amer. Statist. Assoc.*, **88**, 1273–1283.

Shiraishi, T. (1989). Multivariate multi-sample rank tests for location-scale alternatives, *Commun. Statist.-Theory Meth.*, **18**, 85–105.

Shiraishi, T. (1993). Statistical inference based on M-statistics in two-way manova, *J. Japan Statist. Soc.*, **23**, 33–48.

Shiraishi, T. (1996). On scale-invariant M-statistics in multivariate k samples, *J. Japan Statist. Soc.*, **26**, 241–253.

Shiraishi, T. (1998). Studentized robust statistics in multivariate randomized block design, *J. Nonparametric Statist.*, **10**, 95–110.

Shiraishi, T. (1999). Studentized robust statistics for main effects in a two-factor MANOVA, *Commun. Statist.-Theory Meth.*, **28**, 809–823.

Shiraishi, T. (2001). Robust estimates of location parameters in two-way layouts with interaction, *J. Jpn. Soc. Comp. Statist.*, **14**, 49–57.

Shiraishi, T. (2003). *Toukeikagaku*, Nippn-Hyouron-Sha Co., Ltd., (in Japanese).

Welsh, A. H. (1986). Bahadur representations for robust scale estimators based on regression residuals, *Ann. Statist.*, **14**, 1246–1251.