International Journal of Statistical Sciences *ISSN 1683–5603*
Vol. 9(Special Issue), 2009, pp 233-253

Exploratory Procedures after Searching the Underlying Distribution in Multi-Sample Models

T. Shiraishi
Department of Mathematical Sciences
Yokohama City University
Yokohama, Japan

[Received December 29, 2008; Accepted February 2, 2009]

Abstract

As statistical estimation procedures for location, least squares estimators, R-estimators, and M-estimators are introduced in a one-way analysis of variance model. The asymptotic distributional theory for the three estimators and simulated mean squared errors give the features of the respective estimators depending on the underlying distribution. Based on the features, we propose an estimation procedure selecting one of the three estimators after searching a distribution near to the underlying distribution. It is shown that the mean squared error of the new estimator is more stable than the three estimators. Next, as distribution-free test procedures, the permutation F-test, Kruskal-Wallis rank test, and the M-test are introduced. Asymptotic relative efficiency and simulated power of the respective tests are investigated. Based on their features, we propose a stable test procedure selecting one of the three tests after searching a distribution near to the underlying distribution. Surprisingly the new test is a little better than the permutation F-test when the underlying distribution is normal.

Keywords and Phrases: Robust Statistics, R-estimators, M-estimators, Asymptotic Property and Simulation Study.

AMS Classification: 62E20.

1 Introduction

In the present paper, we consider univariate k samples with n_i observations in the i-th population for $i = 1, \ldots, k$. The j-th observation X_{ij} in the i-th level is expressed as

$$X_{ij} = \mu_i + e_{ij} \quad (j = 1, \ldots, n_i, i = 1, \ldots, k) \tag{1.1}$$

where e_{ij} is a random variable with $E(e_{ij}) = 0$ for all i, j's. It is further assumed that e_{ij}'s are independent and identically distributed with continuous distribution function (d.f.) $F(x/\sigma)$ and $Var(e_{ij}) < \infty$. We denote the density of $F(x)$ by $f(x)$. For convenience, we assume

$$\int_{-\infty}^{\infty} x^2 f(x)dx = 1, \text{ that is, } Var(e_{ij}) = \sigma^2 > 0.$$

(1.1) is rewritten as usual by

$$X_{ij} = \nu + \tau_i + e_{ij},$$

where $\sum_{i=1}^{k} n_i \tau_i = 0$. Then ν and τ_i's are referred to as the grand mean and additive treatment constants, respectively. We put $n = \sum_{i=1}^{k} n_i$. The least squares estimator of τ_i is given by $\tilde{\tau}_i = \bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot}$, where $\bar{X}_{i\cdot} = \sum_{j=1}^{n_i} X_{ij}/n_i$ and $\bar{X}_{\cdot\cdot} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} X_{ij}/n$. When the underlying distribution is normal, that is, $F(x) = \Phi(x)$, $\tilde{\tau}_i$'s are the uniformly minimum variance unbiased estimator for τ_i's, where $\Phi(x)$ denotes the standard normal distribution function. Using a method similar to Hodges and Lehmann (1963), we propose as an estimator of $\eta_{ii'} = \boldsymbol{\tau}_i - \boldsymbol{\tau}_{i'}$

$$\hat{\eta}_{ii'} = \text{the sample median of } \{X_{ij} - X_{i'j'} : 1 \le j \le n_i,\ 1 \le j' \le n_{i'}\}.$$

Since the relation

$$\tau_i = (1/n) \sum_{i'=1}^{k} n_{i'} \eta_{ii'}, \tag{1.2}$$

holds, we may propose as an R-estimator of τ_i

$$\hat{\tau}_i = (1/n) \sum_{i'=1}^{k} n_{i'} \hat{\eta}_{ii'},$$

where we set $\hat{\eta}_{ii} = 0$ for convenience. Shiraishi (1990) proposed R-estimators which are asymptotically equivalent to $\hat{\tau}_i$'s. Setting $\hat{\boldsymbol{\tau}}_n = (\hat{\tau}_1, \cdots, \hat{\tau}_k)'$ and $\boldsymbol{\tau} = (\tau_1, \cdots, \tau_k)'$, from Shiraishi (1990), it follows

$$\sqrt{n}(\hat{\boldsymbol{\tau}}_n - \boldsymbol{\tau}) \xrightarrow{\text{L}} N_k(\mathbf{0}, \gamma^2 \Lambda) \tag{1.3}$$

where $\xrightarrow{\mathrm{L}}$ denotes convergence in law, $N_k(\boldsymbol{\theta}, \Sigma)$ stands for the k-dimensional normal variable with mean $\boldsymbol{\theta}$ and variance-covariance matrix Σ, $\gamma^2 = 12\sigma^2\{\int_{-\infty}^{\infty} f^2(x)dx\}^2$, and $\Lambda = (\delta_{ii'}/\lambda_i - 1)_{ii'=1,\ldots,k}$ with $\delta_{ii'}$ denoting Kronecker's delta.

For one-sample location model of $k = 1$, Huber (1964) proposed solution $\breve{\theta} = \theta$ of the equation:

$$\sum_{j=1}^{n} \psi(X_{1j} - \theta) = 0 \tag{1.4}$$

as an estimator of $E(X_{11})$ and called it M-estimator, where $\psi(x)$ is increasing and strictly negative (positive) for large negative (positive) values of x. For fixed ϵ such that $0 < \epsilon < 1$, choose positive constant c satisfying the equation:

$$\frac{2\varphi(c)}{c} - 2\Phi(-c) = \frac{\varepsilon}{1 - \varepsilon}, \tag{1.5}$$

where $\varphi(x)$ denotes the standard normal density function. Then Huber (1964) showed that the M-estimator given by taking

$$\psi(x) = \max\{\min\{x, c\}, -c\} \tag{1.6}$$

has the minimax asymptotic variance among a class of estimators defined by the solution of (1.4) through the function $\psi(\cdot)$ over the class of distributions that the underlying distribution is in ϵ-contamination neighborhood of a normal distribution: $U_\varepsilon = \{F(x/\sigma) = (1 - \varepsilon)\Phi(x) + \varepsilon H(x) : H(-x) = 1 - H(x)$ for any $x\}$. When $\varepsilon = 0.05, 0.03$ are given, the values of c satisfying (1.5) are respectively 1.398 and 1.579. Many values of (ε, c) were appeared in Table 1 of Shiraishi (2005). Furthermore Huber (1981) proposed a scale-invariant M-estimator.

In order to introduce robust estimators for τ_i's, we put, for $i \neq i'$, $\tilde{\nu}_{ii'} = (n_i\bar{X}_{i\cdot} + n_{i'}\bar{X}_{i'\cdot})/N_{ii'}$, $N_{ii'} = n_i + n_{i'}$ and $\breve{\sigma}_n = \sqrt{\pi}\sum_{i=1}^{k}\sum_{j=1}^{n_i}|X_{ij} - \bar{X}_{i\cdot}|/(\sqrt{2}n)$. $\breve{\sigma}_n$ is a consistent estimator of $\rho = (\sqrt{\pi}\sigma/\sqrt{2})\int_{-\infty}^{\infty}|x|dF(x)$. Moreover, we put

$$T_{ii'}(\theta) = \frac{1}{n_i}\sum_{j=1}^{n_i}\psi\left(\frac{X_{ij} - \tilde{\nu}_{ii'} - (n_{i'}/N_{ii'})\cdot\theta}{\breve{\sigma}_n}\right) - \frac{1}{n_{i'}}\sum_{j=1}^{n_{i'}}\psi\left(\frac{X_{i'j} - \tilde{\nu}_{ii'} + (n_i/N_{ii'})\cdot\theta}{\breve{\sigma}_n}\right).$$

Then Shiraishi (2007) denoted solution of the equation: $T_{ii'}(\theta) = 0$ by $\breve{\eta}_{ii'}$. He proposed $\breve{\eta}_{ii'}$ as a robust estimator for $\boldsymbol{\tau}_i - \boldsymbol{\tau}_{i'}$. Hence, from (1.2), as a robust estimator for τ_i, he might propose

$$\breve{\tau}_i = \frac{1}{n}\sum_{i'=1}^{k} n_{i'}\breve{\eta}_{ii'}, \quad (i = 1, \cdots, k)$$

where $\breve{\eta}_{ii} = 0$. Setting $\breve{\boldsymbol{\tau}}_n = (\breve{\tau}_1, \cdots, \breve{\tau}_k)'$, Shiraishi (2007) showed

$$\sqrt{n}(\breve{\boldsymbol{\tau}}_n - \boldsymbol{\tau}) \xrightarrow{\text{L}} N_k(\mathbf{0}, \sigma^2(\psi)\Lambda) \tag{1.7}$$

where $\sigma^2(\psi) = \sigma^2\{\int_{-\infty}^{\infty} \psi(\sigma x/\rho)f'(x)dx\}^2 / \int_{-\infty}^{\infty}\{\psi(\sigma x/\rho) - \bar{\psi}\}^2 f(x)dx$ and $\bar{\psi} = \int_{-\infty}^{\infty} \psi(\sigma x/\rho)f(x)dx$.

We give the values of the asymptotic relative efficiency among the three estimators $\tilde{\boldsymbol{\tau}}_n = (\tilde{\tau}_1, \cdots, \tilde{\tau}_k)'$, $\hat{\boldsymbol{\tau}}_n$ and $\breve{\boldsymbol{\tau}}_n$ for many underlying distributions. Furthermore, we give values of simulated mean squared errors by using a Monte Carlo simulation. Then we may get the features of the three type estimators on the underlying distributions. Based on the features, we propose the estimation procedure selecting one of the three type estimators after searching a distribution near to the underlying distribution. It is shown that the mean squared error of the new estimator is more stable than the estimators $\tilde{\boldsymbol{\tau}}_n$, $\hat{\boldsymbol{\tau}}_n$ and $\breve{\boldsymbol{\tau}}_n$.

Next, we consider distribution-free test procedures for the null hypothesis of homogenuity

$$H_0;\ \tau_1 = \ldots = \tau_k = 0.$$

The multi-sample F-test is the optimum when the underlying distribution is normal. The permutation F-test statistic is distribution-free under H_0. When the underlying distribution is logistic, Hájek *et al.* (1999) reviewed that Kruskal-Wallis rank test is the asymptotically optimum test. Shiraishi (1996) proposed an M-test based on the statistic similar to M-estimator. It is shown that the asymptotic relative efficiency (ARE) among the three tests of the F-test, the Kruskal-Wallis test and the M-test agrees with the ARE among the three estimators $\tilde{\boldsymbol{\tau}}_n$, $\hat{\boldsymbol{\tau}}_n$ and $\breve{\boldsymbol{\tau}}_n$. The simulated power of the respective tests is investigated. Based on their features, we propose the test procedure selecting one of the three tests after searching a distribution near to the underlying distribution. It is shown that the power of the proposed test is more stable than the three tests of the permutation F-test, the Kruskal-Wallis test and the permutation M-test. Surprisingly the new test is a little better than the permutation F-test when the underlying distribution is normal.

2 Searching the underlying distribution

Since the power of goodness of fit tests is low, we consider to search a distribution near to the underlying distribution by using an empirical distribution function. For integers m and

$j = m - \sum_{i'=0}^{i-1} n_{i'}$ such that $\sum_{i'=0}^{i-1} n_{i'} + 1 \leq m \leq \sum_{i'=0}^{i} n_{i'}$, we define $Z_1, \cdots, Z_n$ by

$$Z_m = (X_{ij} - \bar{X}_{i\cdot})/\sqrt{1 - 1/n_i},$$

where $n_0 = 0$. Then $E(Z_i) = 0$ and $Var(Z_i) = \sigma^2$ hold. Let us define $\hat{G}_n(x)$ by the empirical distribution function of $\{Z_1, \cdots, Z_n\}$. That is, $n \cdot \hat{G}_n(x)$ equals the number of observations $Z_1, \cdots, Z_n$ that are less than or equal to x. $\hat{G}_n(x)$ is an unbiased and consistent estimator of $F((x - \boldsymbol{\tau})/\sigma)$. Hence for specified distribution function F_0, we put

$$D_{F_0} = \sup_{-\infty < x < \infty} \left| \hat{G}_n(x) - F_0\left(\frac{x}{\breve{\sigma}_n}\right) \right| \tag{2.1}$$

where

$$\breve{\sigma}_n = \sum_{i=1}^{n} |Z_i| / \left(n \int_{-\infty}^{\infty} |x| f_0(x) dx \right).$$

We may rewrite

$$D_{F_0} = \max_{1 \leq i \leq n} \left[\max \left\{ \left| \frac{i}{n} - F_0\left(\frac{Z_{(i)}}{\breve{\sigma}_n}\right) \right|, \left| F_0\left(\frac{Z_{(i)}}{\breve{\sigma}_n}\right) - \frac{i-1}{n} \right| \right\} \right].$$

D_{F_0} is a distance between the underlying distribution F and specified F_0. Symmetric distributions chosen here as F_0 are normal $N(0,1)$, contaminated normal $CN(\varepsilon) = (1 - \varepsilon)N(0, 1/(1+8\varepsilon)) + \varepsilon N(0, 9/(1+8\varepsilon))$, logistic $LG(0, \sqrt{3}/\pi)$ with density function $\exp(-\pi x/\sqrt{3})/\{1 + \exp(-\pi x/\sqrt{3})\}^2$, and double exponential $DE(0, 1/\sqrt{2})$ with density function $(1/\sqrt{2}) \exp(-\sqrt{2}|x|)$. As asymmetric distributions, we provide exponential, Weible, lognormal and asymmetric contaminated normal distributions. The normalized exponential EX density function with mean 0 and variance 1 is given by $f_e(x) = e^{-(x+1)} I_{[-1,\infty)}(x)$. The Weible density function is expressed as $g_w(x|a,b) = (a/b)(x/b)^{a-1} \exp\{-(x/b)^a\}$ $(0 < x < \infty)$ with mean $\boldsymbol{\tau} = b\Gamma(\frac{1}{a} + 1)$ and variance $\sigma^2 = b^2\{\Gamma(\frac{2}{a} + 1) - \Gamma^2(\frac{1}{a} + 1)\}$. The normalized Weible $W(a)$ density function is given by $f_w(x|a) = \sigma g_w(\sigma x + \boldsymbol{\tau}|a, b)$, which does not depend on parameter b. The lognormal density function is expressed as $g_\ell(x|a,b) = \frac{1}{\sqrt{2\pi}bx} \exp\{-\frac{(\log x - a)^2}{2b^2}\}$ $(0 < x < \infty)$ with mean $\boldsymbol{\tau} = \exp(a + \frac{b^2}{2})$ and variance $\sigma^2 = \exp(2a + 2b^2) - \exp(2a + b^2)$. The normalized lognormal $LN(b)$ density function is given by $f_\ell(x|b) = \sigma g_\ell(\sigma x + \boldsymbol{\tau}|a, b)$ which does not depend on parameter a. The asymmetric contaminated normal $ACN = 0.98N(-0.1, 1/1.47) + 0.02N(4.9, 0.01/1.47)$ has outlier with probability 0.02. The mean and variance of ACN are respectively 0 and 1. We may seek F_0 minimizing D_{F_0} in the distributions. To compute $\breve{\sigma}_n$, the values of $\int_{-\infty}^{\infty} |x| f_0(x) dx$ are appeared in Table 1.

Table 1: The values of $\int_{-\infty}^{\infty}|x|f_0(x)dx$, where $\mu_0 = \exp(\frac{b^2}{2})$.

$f_0(x)$	$N(0,1)$	$CN(\varepsilon)$	$LG(0,\sqrt{3}/\pi)$	$DE(0,1/\sqrt{2})$
$\int_{-\infty}^{\infty}\|x\|f_0(x)dx$	$\sqrt{\frac{2}{\pi}}$	$\left\{\frac{(1+2\varepsilon)}{\sqrt{(1+8\varepsilon)}}\right\}\sqrt{\frac{2}{\pi}}$	$2\log 2\times\frac{\sqrt{3}}{\pi}$	$\frac{1}{\sqrt{2}}$

$f_0(x)$	EX	$W(a)$	$LN(b)$	ACN
$\int_{-\infty}^{\infty}\|x\|f_0(x)dx$	$\frac{2}{e}$	$\frac{2\mu}{\sigma}-\frac{2}{\sigma}\int_0^{\mu}\exp\{-(x/b)^a\}dx$	$\frac{2}{\sigma}\int_0^{\mu_0}\Phi\left(\frac{\log x}{b}\right)dx$	0.72118

Moreover, as a characterization of the distribution, we may use the skewness: $\ell_1 = \int_{-\infty}^{\infty}x^3dF(x)$ and the kurtosis: $\ell_2 = \int_{-\infty}^{\infty}x^4dF(x)-3$. The values for the skewness and kurtosis of the respective distributions: $N(0,1)$, $CN(0.05)$, $LG(0,\sqrt{3}/\pi)$, $DE(0,1/\sqrt{2})$, EX, and ACN are appeared in Table 2.

Table 2: The values for the skewness and kurtosis of $N(0,1)$, $CN(0.05)$, $LG(0,\sqrt{3}/\pi)$, $DE(0,1/\sqrt{2})$, EX, and ACN.

$F(x)$	skewness	kurtosis	$F(x)$	skewness	kurtosis
$N(0,1)$	0	0	$DE(0,1/\sqrt{2})$	0	3
$CN(0.05)$	0	4.65	EX	2	6
$LG(0,\sqrt{3}/\pi)$	0	1.2	ACN	-0.16	-1.61

The skewness and kurtosis of $W(a)$ become

$$\ell_1 = \frac{\Gamma(\frac{3}{a}+1)-3\Gamma(\frac{1}{a}+1)\Gamma(\frac{2}{a}+1)+2\Gamma^3(\frac{1}{a}+1)}{\{\Gamma(\frac{2}{a}+1)-\Gamma^2(\frac{1}{a}+1)\}^{3/2}}$$

and

$$\ell_2 = \frac{\Gamma(\frac{4}{a}+1)-4\Gamma(\frac{1}{a}+1)\Gamma(\frac{3}{a}+1)+6\Gamma^2(\frac{1}{a}+1)\Gamma(\frac{2}{a}+1)-3\Gamma^4(\frac{1}{a}+1)}{\{\Gamma(\frac{2}{a}+1)-\Gamma^2(\frac{1}{a}+1)\}^2}-3.$$

The values for some a's are appeared in Table 3.

Table 3: The values for the skewness and kurtosis of $W(a)$ for $a = 1(1)6$

a	skewness	kurtosis	a	skewness	kurtosis
1	2	6	4	-0.09	-0.25
2	0.63	0.25	5	-0.25	-0.12
3	0.17	-0.27	6	-0.37	-0.36

Le us put $a_0 = \exp(b^2)$. Then the skewness and kurtosis of $LN(b)$ become $\ell_1 = (a_0 - 1)^{1/2}(a_0 + 2)$ and $\ell_2 = (a_0 - 1)(a_0^3 + 3a_0^2 + 6a_0 + 6)$. The values for some b^2's are appeared in Table 4.

Table 4: The values for the skewness and kurtosis of $LN(b)$ for $b^2 = 0.1(0.1)1.0$

b^2	skewness	kurtosis	b^2	skewness	kurtosis
0.1	1.00	1.86	0.6	3.47	27.08
0.2	1.52	4.35	0.7	4.04	38.94
0.3	1.98	7.71	0.8	4.68	55.44
0.4	2.45	12.27	0.9	5.39	78.51
0.5	2.94	18.51	1.0	6.18	110.94

The estimators for ℓ_1 and ℓ_2 are respectively given by

$$\hat{\ell}_1 = \frac{n\sqrt{n-1}}{n-2} \cdot \frac{\sum_{i=1}^n (Z_i - \bar{Z})^3}{\{\sum_{i=1}^n (Z_i - \bar{Z})^2\}^{3/2}} \tag{2.2}$$

and

$$\hat{\ell}_2 = \frac{n(n+1)(n-1)}{(n-2)(n-3)} \cdot \frac{\sum_{i=1}^n (Z_i - \bar{Z})^4}{\{\sum_{i=1}^n (Z_i - \bar{Z})^2\}^2} - \frac{3(n-1)^2}{(n-2)(n-3)}. \tag{2.3}$$

$\hat{\ell}_1$ and $\hat{\ell}_2$ are used in the SAS system and Microsoft Excel.

3 Asymptotic efficiency for estimators

We investigate the asymptotic relative efficiency (ARE) of R-estimator $\hat{\boldsymbol{\tau}}_n$ and M-estimator $\breve{\boldsymbol{\tau}}_n$ with respect to least squares estimator $\tilde{\boldsymbol{\tau}}_n$. For two sequences of estimators $\{\boldsymbol{T}_{1n}\}$ and $\{\boldsymbol{T}_{2n}\}$ such that

$$\sqrt{n}(\boldsymbol{T}_{1n} - \boldsymbol{\tau}) \xrightarrow{\text{L}} N_k(\mathbf{0}, \sigma_1^2 \Lambda) \text{ and } \sqrt{n}(\boldsymbol{T}_{2n} - \boldsymbol{\tau}) \xrightarrow{\text{L}} N_k(\mathbf{0}, \sigma_2^2 \Lambda),$$

where $\xrightarrow{\text{L}}$ denotes convergence in law, we define the ARE of $\boldsymbol{T}_{1n}$ with respect to $\boldsymbol{T}_{2n}$ by $ARE(\boldsymbol{T}_{1n}, \boldsymbol{T}_{2n}) = \sigma_2^2/\sigma_1^2$. Saleh and Shiraishi (1993) defined the asymptotic distributional quadratic risk (ADQR) of $\boldsymbol{T}_{in}$ by

$$AR(T_{in}) = \lim_{b\to\infty}\lim_{n\to\infty} E[\min\{n(\boldsymbol{T}_{in} - \boldsymbol{\tau})'(\boldsymbol{T}_{in} - \boldsymbol{\tau}), b\}].$$

$ARE(\boldsymbol{T}_{1n}, \boldsymbol{T}_{2n})$ is given by the ratio of $AR(\boldsymbol{T}_{2n})$ to $AR(\boldsymbol{T}_{1n})$, that is, $ARE(\boldsymbol{T}_{1n}, \boldsymbol{T}_{2n}) = AR(\boldsymbol{T}_{2n})/AR(\boldsymbol{T}_{1n})$. Hence if $ARE(\boldsymbol{T}_{1n}, \boldsymbol{T}_{2n}) > (<) 1$ holds, $\boldsymbol{T}_{1n}$ is better (worse) than $\boldsymbol{T}_{2n}$. From the direct application of central limit theorem, we get

$$\sqrt{n}(\tilde{\boldsymbol{\tau}}_n - \boldsymbol{\tau}) \xrightarrow{\text{L}} N(0, \sigma^2 \Lambda). \tag{3.1}$$

Hence, from (1.3), (1.7) and (3.1), the asymptotic relative efficiency (ARE) of $\hat{\boldsymbol{\tau}}_n$ and $\breve{\boldsymbol{\tau}}_n$ with respect to $\tilde{\boldsymbol{\tau}}_n$ yield

$$ARE(\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n) = 12\left\{\int_{-\infty}^{\infty} f^2(x)dx\right\}^2,$$

and

$$ARE(\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n) = \left[\int_{-\infty}^{\infty}\{\psi(\sigma x/\rho) - \bar{\psi}\}f'(x)dx\right]^2 \Big/ \int_{-\infty}^{\infty}\{\psi(\sigma x/\rho) - \bar{\psi}\}^2 f(x)dx. \tag{3.2}$$

Let us put $\xi = (\sqrt{\pi}/\sqrt{2})\int_{-\infty}^{\infty}|x|dF(x)$. Then (3.2) becomes

$$ARE(\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n) = \frac{\left\{\int_{-c\xi}^{c\xi}(x/\xi - \bar{\psi})f'(x)dx + \int_{c\xi}^{\infty}(c - \bar{\psi})f'(x)dx - \int_{-\infty}^{-c\xi}(c + \bar{\psi})f'(x)dx\right\}^2}{\int_{-c\xi}^{c\xi}(x/\xi - \bar{\psi})^2 f(x)dx + \int_{c\xi}^{\infty}(c - \bar{\psi})^2 f(x)dx + \int_{-\infty}^{-c\xi}(c + \bar{\psi})^2 f(x)dx},$$

where $\bar{\psi} = \int_{-c\xi}^{c\xi}(x/\xi)f(x)dx + \int_{c\xi}^{\infty} cf(x)dx - \int_{-\infty}^{-c\xi} cf(x)dx$. Since the M-function ψ of (1.6) depends on constant c, $ARE(\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$ is a function of c. From Table 1, we choose the value of 1.579 as c. We denote the M-estimator by $\breve{\boldsymbol{\tau}}_n$. The values of $ARE(\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$, $ARE(\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$ are appeared in Table 5. 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, and exponential. The values of $ARE(\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$ in Table 5 are stated in Lehmann (1975) and others.

Table 5: The ARE of the R-estimator and two M-estimators relative to the least squares estimators.

$F(x)$	$ARE(\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$	$ARE(\check{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$
normal	0.955	0.970
contaminated normal	1.196	1.212
logistic	1.097	1.081
double exponential	1.5	1.262
exponential	3	1.288

When the underlying distribution is ϵ-contaminated normal $CN(\varepsilon)$ distribution , the values of $ARE(\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$ and $ARE(\check{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$ are appeared in Table 6. When the underlying distribution is t distribution with k degrees of freedom, the values of ARE's are respectively appeared in Table 7. Then the density of the normalized t distribution is given by

$$f(x) = \frac{\Gamma(\frac{k+1}{2})}{\sqrt{\pi(k-2)}\Gamma(\frac{k}{2})} \left(1 + \frac{x^2}{k-2}\right)^{-\frac{k+1}{2}} .$$

Table 6: The ARE of the R-estimator and two M-estimators relative to the LSE when the underlying distribution is ϵ-contaminated normal $CN(\varepsilon)$ distribution.

ϵ	0.01	0.02	0.03	0.04	0.05	0.06	0.08	0.10
$ARE(\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$	1.009	1.060	1.108	1.153	1.196	1.236	1.309	1.373
$ARE(\check{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$	1.025	1.077	1.125	1.170	1.212	1.250	1.316	1.370
ϵ	0.12	0.14	0.16	0.18	0.20	0.22	0.25	0.30
$ARE(\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$	1.429	1.476	1.516	1.548	1.575	1.595	1.616	1.627
$ARE(\check{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$	1.412	1.443	1.465	1.478	1.484	1.484	1.473	1.437

Table 7: The ARE of the R-estimator and two M-estimators relative to the least squares estimators when the underlying distribution is t-distribution with k degrees of freedom.

k	3	4	5	6	7	8	9	10	11
$ARE(\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$	1.900	1.401	1.241	1.164	1.119	1.089	1.069	1.054	1.042
$ARE(\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$	1.737	1.336	1.207	1.144	1.107	1.083	1.066	1.053	1.043
k	12	13	14	15	16	17	18	19	20
$ARE(\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$	1.033	1.025	1.019	1.014	1.009	1.006	1.002	0.999	0.997
$ARE(\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$	1.036	1.030	1.024	1.020	1.016	1.013	1.010	1.008	1.006

Since

$$ARE(\breve{\boldsymbol{\tau}}_n, \hat{\boldsymbol{\tau}}_n) = ARE(\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)/ARE(\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$$

holds, we may get the values of the ARE of the M-estimator relative to the R-estimator from Tables 5-7.

The following features 1-3 are drawn from Table 5.

1. When the underlying distribution is normal, the least squares estimator is a little better than the other estimators. The relation of ADQR's is given by

$$AR(\tilde{\boldsymbol{\tau}}_n) < AR(\breve{\boldsymbol{\tau}}_n) < AR(\hat{\boldsymbol{\tau}}_n).$$

2. When the underlying distribution is 0.05-contaminated normal, the M-estimator $\breve{\boldsymbol{\tau}}_n$ is better than the other estimators. The least squares estimator is extremely bad. The relation of ADQR's is given by

$$AR(\breve{\boldsymbol{\tau}}_n) < AR(\hat{\boldsymbol{\tau}}_n) < AR(\tilde{\boldsymbol{\tau}}_n).$$

3. When the underlying distribution is logistic, double exponential or exponential, the R-estimator $\hat{\boldsymbol{\tau}}_n$ is better than the other estimators and the least squares estimator is the worst. The relation of ADQR's is given by

$$AR(\hat{\boldsymbol{\tau}}_n) < AR(\breve{\boldsymbol{\tau}}_n) < AR(\tilde{\boldsymbol{\tau}}_n).$$

From Table 6, when the underlying distribution is ϵ-contaminated normal, we conclude that (i) $ARE(\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$ and $ARE(\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$ increase in ϵ, (ii) all the ARE's are larger than 1, and (iii) the relation of ADQR's is given by

$$AR(\breve{\boldsymbol{\tau}}_n) < AR(\hat{\boldsymbol{\tau}}_n) < AR(\tilde{\boldsymbol{\tau}}_n).$$

From Table 7, when the underlying distribution is t-distribution with k degrees of freedom, we conclude (i) $ARE(\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$ and $ARE(\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n)$ decrease in k,

$$\text{(ii)} \quad AR(\hat{\boldsymbol{\tau}}_n) > \text{ or } < AR(\tilde{\boldsymbol{\tau}}_n) \quad \text{according as} \quad k > \text{ or } \leq 18,$$

(iii) $AR(\breve{\boldsymbol{\tau}}_n) < AR(\tilde{\boldsymbol{\tau}}_n)$, (iv) $AR(\hat{\boldsymbol{\tau}}_n) < AR(\breve{\boldsymbol{\tau}}_n) < AR(\tilde{\boldsymbol{\tau}}_n)$ for $k \leq 10$, and (iv) $AR(\breve{\boldsymbol{\tau}}_n) < AR(\hat{\boldsymbol{\tau}}_n) < AR(\tilde{\boldsymbol{\tau}}_n)$ for $11 \leq k \leq 18$.

4 New estimator

Using the distance D_{F_0}, the sample skewness $\hat{\ell}_1$ and the sample kurtosis $\hat{\ell}_2$ given by (2.1)-(2.3) respectively, we may propose a new estimator selecting one of the three estimators; $\tilde{\boldsymbol{\tau}}_n$, $\hat{\boldsymbol{\tau}}_n$, $\breve{\boldsymbol{\tau}}_n$ after searching a distribution near to the underlying distribution. First, we compute D_{F_0} for $F_0 = N(0,1),\ CN(0.05),\ LG(0, \sqrt{3}/\pi),\ DE(0, 1/\sqrt{2})$ and the respective values are denoted by d_{NM}, d_{CN}, d_{LG}, and d_{DE}. Next we put

$$d_0 = \text{the minimum of } \{d_{NM},\ d_{CN},\ d_{LG},\ d_{DE}\}.$$

Furthermore we compute $\hat{\ell}_1$ for the sake of detecting the underlying asymmetric distribution. Lastly we compute $\hat{\ell}_2$. From the features of Section 3, we recommend using $\tilde{\boldsymbol{\tau}}_n$ when the underlying distribution is close to the normal distribution. We recommend using $\hat{\boldsymbol{\tau}}_n$ when the underlying distribution is remote from the normal distribution. Otherwise we may use $\breve{\boldsymbol{\tau}}_n$. Hence as a new estimator, we propose

$$\widehat{\boldsymbol{\tau}_n^*} = \begin{cases} \tilde{\boldsymbol{\tau}}_n & \text{if} \quad d_0 = d_{NM}, \\ \hat{\boldsymbol{\tau}}_n & \text{if} \quad d_0 = d_{DE}, \quad \text{or} \quad \hat{\ell}_1 \geq 1.0, \quad \text{or} \quad \hat{\ell}_2 \geq 0.5, \\ \breve{\boldsymbol{\tau}}_n & \text{otherwise}. \end{cases} \tag{4.1}$$

Let $MS(\boldsymbol{T}_n)$ be the mean squared error of the estimator $\boldsymbol{T}_n$, that is, $MS(\boldsymbol{T}_n) = E\{(\boldsymbol{T}_n - \boldsymbol{\tau})'(\boldsymbol{T}_n - \boldsymbol{\tau})\}$. Then for two sequences of estimators $\{\boldsymbol{T}_{1n}\}$ and $\{\boldsymbol{T}_{2n}\}$, we define the relative risk efficiency (RRE) of $\boldsymbol{T}_{1n}$ with respect to $\boldsymbol{T}_{2n}$ by $RRE(\boldsymbol{T}_{1n}, \boldsymbol{T}_{2n}) = MS(\boldsymbol{T}_{2n})/MS(\boldsymbol{T}_{1n})$. Under certain regularity conditions,

$$\lim_{n\to\infty} RRE(\boldsymbol{T}_{1n}, \boldsymbol{T}_{2n}) = ARE(\boldsymbol{T}_{1n}, \boldsymbol{T}_{2n})$$

holds. We simulate the RRE's among the least squares estimator, the R-estimator, the M-estimator and the new estimator. We limited attention to $k = 3, 5, n_1 = \cdots = n_k = 10, 20, 30$ and $F(x) = N(0, 1)$, $CN(0.05)$, $LG(0, \sqrt{3}/\pi)$, $DE(0, 1/\sqrt{2})$, $LN(1)$, $W(2)$, EX. We denote M-estimator satisfying the equation (1.5) with $c = 1.579$ by $\breve{\boldsymbol{\tau}}_n$. The RRE's among the estimators $\tilde{\boldsymbol{\tau}}_n$, $\hat{\boldsymbol{\tau}}_n$, $\breve{\boldsymbol{\tau}}_n$, and $\widehat{\boldsymbol{\tau}_n^*}$ are appeared in Tables 8 and 9 for $k = 3, 5$, respectively. The values of the RRE are estimated by Monte-Carlo simulation of $5,000$ samples.

The following features I-V are drawn from Tables 8 and 9.

I. When the underlying distribution is normal, although $\tilde{\boldsymbol{\tau}}_n$ is the best estimator, $\widehat{\boldsymbol{\tau}_n^*}$ is quite a little worse than $\tilde{\boldsymbol{\tau}}_n$. $\hat{\boldsymbol{\tau}}_n$ is the worst. The relation of MSE's is given by

$$MS(\tilde{\boldsymbol{\tau}}_n) < MS(\widehat{\boldsymbol{\tau}_n^*}) < MS(\breve{\boldsymbol{\tau}}_n) < MS(\hat{\boldsymbol{\tau}}_n). \tag{4.2}$$

II. When the underlying distribution is 0.05-contaminated normal, the M-estimators $\breve{\boldsymbol{\tau}}_n$ are better than the other estimators. The new estimator $\widehat{\boldsymbol{\tau}_n^*}$ is quite a little worse than the M-estimators and $\tilde{\boldsymbol{\tau}}_n$ is the worst. The relation of MSE's is given by

$$MS(\breve{\boldsymbol{\tau}}_n) < MS(\widehat{\boldsymbol{\tau}_n^*}) \approx MS(\hat{\boldsymbol{\tau}}_n) < MS(\tilde{\boldsymbol{\tau}}_n).$$

III. When the underlying distribution is logistic, the $\hat{\boldsymbol{\tau}}_n$ is better than the other estimators. The new estimator $\widehat{\boldsymbol{\tau}_n^*}$ is quite a little worse than $\hat{\boldsymbol{\tau}}_n$. $\tilde{\boldsymbol{\tau}}_n$ is the worst. The relation of MSE's is given by

$$MS(\hat{\boldsymbol{\tau}}_n) < MS(\breve{\boldsymbol{\tau}}_n) < MS(\widehat{\boldsymbol{\tau}_n^*}) < MS(\tilde{\boldsymbol{\tau}}_n).$$

IV. When the underlying distribution is double exponential, the $\hat{\boldsymbol{\tau}}_n$ is better than the other estimators. The $\widehat{\boldsymbol{\tau}_n^*}$ has maxmum 5 percent loss in comparison with $\hat{\boldsymbol{\tau}}_n$. $\tilde{\boldsymbol{\tau}}_n$ is extremely bad. The relation of MSE's is given by (4.2).

V. When the underlying distribution is asymmetric, the $\widehat{\boldsymbol{\tau}_n^*}$ is as good as $\hat{\boldsymbol{\tau}}_n$. $\tilde{\boldsymbol{\tau}}_n$ and $\breve{\boldsymbol{\tau}}_n$ are extremely worse than the former two estimators.

The double exponential distribution and asymmetric distributions $LN(1)$, $W(2)$, EX are very far from the normal distribution. From the above I-V, we conclude that the MSE's of the new estimators are stable in comparison with the other estimators.

Table 8: The relative risk efficiency (RRE) among the estimators $\tilde{\boldsymbol{\tau}}_n,\ \hat{\boldsymbol{\tau}}_n,\ \breve{\boldsymbol{\tau}}_n$, and $\widehat{\boldsymbol{\tau}_n^*}$ for $k = 3$.

(i) $F(x)$ is normal

n_1	RRE($\widehat{\boldsymbol{\tau}_n^*}, \tilde{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \hat{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \breve{\boldsymbol{\tau}}_n$)	RRE($\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)	RRE($\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)
10	0.970	1.035	0.991	0.938	0.979
20	0.980	1.033	1.004	0.949	0.976
30	0.980	1.043	1.014	0.939	0.967

(ii) $F(x)$ is contaminated normal

n_1	RRE($\widehat{\boldsymbol{\tau}_n^*}, \tilde{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \hat{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \breve{\boldsymbol{\tau}}_n$)	RRE($\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)	RRE($\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)
10	1.163	1.009	0.990	1.152	1.175
20	1.213	1.002	0.993	1.210	1.221
30	1.178	0.997	0.983	1.181	1.199

(iii) $F(x)$ is logistic

n_1	RRE($\widehat{\boldsymbol{\tau}_n^*}, \tilde{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \hat{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \breve{\boldsymbol{\tau}}_n$)	RRE($\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)	RRE($\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)
10	1.047	0.997	0.992	1.050	1.056
20	1.065	0.980	0.986	1.087	1.080
30	1.069	0.978	0.982	1.093	1.088

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

n_1	RRE($\widehat{\boldsymbol{\tau}_n^*}, \tilde{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \hat{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \breve{\boldsymbol{\tau}}_n$)	RRE($\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)	RRE($\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)
10	1.269	0.955	1.053	1.328	1.205
20	1.394	0.963	1.112	1.448	1.253
30	1.385	0.972	1.114	1.425	1.243

(v) $F(x)$ is lognormal

n_1	RRE($\widehat{\boldsymbol{\tau}_n^*}, \tilde{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \hat{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \breve{\boldsymbol{\tau}}_n$)	RRE($\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)	RRE($\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)
10	4.529	0.993	1.879	4.559	2.411
20	5.772	0.999	2.179	5.775	2.648
30	5.899	1.000	2.291	5.899	2.575

(vi) $F(x)$ is Weible

n_1	RRE($\widehat{\boldsymbol{\tau}_n^*}, \tilde{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \hat{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \breve{\boldsymbol{\tau}}_n$)	RRE($\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)	RRE($\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)
10	15.345	0.997	4.501	15.393	3.409
20	28.584	1.000	7.453	28.584	3.835
30	41.529	1.000	10.769	41.529	3.856

(vii) $F(x)$ is exponential

n_1	RRE($\widehat{\boldsymbol{\tau}_n^*}, \tilde{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \hat{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \breve{\boldsymbol{\tau}}_n$)	RRE($\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)	RRE($\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)
10	1.767	0.998	1.305	1.770	1.354
20	2.114	0.999	1.490	2.116	1.419
30	2.306	0.998	1.626	2.311	1.418

Table 9: The relative risk efficiency (RRE) among the estimators $\tilde{\boldsymbol{\tau}}_n,\ \hat{\boldsymbol{\tau}}_n,\ \breve{\boldsymbol{\tau}}_n$, and $\widehat{\boldsymbol{\tau}_n^*}$ for $k = 5$.

(i) $F(x)$ is normal

n_1	RRE($\widehat{\boldsymbol{\tau}_n^*}, \tilde{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \hat{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \breve{\boldsymbol{\tau}}_n$)	RRE($\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)	RRE($\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)
10	0.979	1.039	1.006	0.943	0.973
20	0.989	1.041	1.016	0.950	0.974
30	0.990	1.046	1.023	0.947	0.968

(ii) $F(x)$ is contaminated normal

n_1	RRE($\widehat{\boldsymbol{\tau}_n^*}, \tilde{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \hat{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \breve{\boldsymbol{\tau}}_n$)	RRE($\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)	RRE($\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)
10	1.170	1.002	0.992	1.168	1.179
20	1.190	1.000	0.989	1.190	1.204
30	1.179	0.992	0.981	1.188	1.202

(iii) $F(x)$ is logistic

n_1	RRE($\widehat{\boldsymbol{\tau}_n^*}, \tilde{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \hat{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \breve{\boldsymbol{\tau}}_n$)	RRE($\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)	RRE($\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)
10	1.049	0.984	0.986	1.066	1.063
20	1.055	0.980	0.993	1.077	1.063
30	1.059	0.974	0.984	1.086	1.076

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

n_1	RRE($\widehat{\boldsymbol{\tau}_n^*}, \tilde{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \hat{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \breve{\boldsymbol{\tau}}_n$)	RRE($\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)	RRE($\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)
10	1.283	0.947	1.059	1.354	1.211
20	1.355	0.972	1.106	1.394	1.225
30	1.402	0.980	1.140	1.430	1.230

(v) $F(x)$ is lognormal

n_1	RRE($\widehat{\boldsymbol{\tau}_n^*}, \tilde{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \hat{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \breve{\boldsymbol{\tau}}_n$)	RRE($\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)	RRE($\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)
10	5.017	0.999	1.962	5.020	2.558
20	5.589	1.000	2.173	5.589	2.573
30	5.952	1.000	2.278	5.952	2.613

(vi) $F(x)$ is Weible

n_1	RRE($\widehat{\boldsymbol{\tau}_n^*}, \tilde{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \hat{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \breve{\boldsymbol{\tau}}_n$)	RRE($\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)	RRE($\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)
10	14.421	1.000	4.305	14.421	3.350
20	30.262	1.000	8.085	30.262	3.743
30	41.788	1.000	10.530	41.788	3.969

(vii) $F(x)$ is exponential

n_1	RRE($\widehat{\boldsymbol{\tau}_n^*}, \tilde{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \hat{\boldsymbol{\tau}}_n$)	RRE($\widehat{\boldsymbol{\tau}_n^*}, \breve{\boldsymbol{\tau}}_n$)	RRE($\hat{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)	RRE($\breve{\boldsymbol{\tau}}_n, \tilde{\boldsymbol{\tau}}_n$)
10	1.834	0.999	1.335	1.836	1.375
20	2.132	1.000	1.515	2.132	1.407
30	2.281	1.000	1.598	2.281	1.427

5 Test procedures

We shall introduce distribution-free tests for the null hypothesis H_0. Let $X_{(1)}, \cdots, X_{(N)}$ be the order statistics of the independent observations $X_{11}, \cdots, X_{kn_k}$. Then we set $\boldsymbol{X}_{(\cdot)} = (X_{(1)}, \cdots, X_{(N)})$. The permutation test statistic is given by

$$T(\boldsymbol{X}) = \sum_{i=1}^{k} n_i (\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2,$$

where $\bar{X}_{\cdot\cdot}$ stands for the overall sample mean. Let us put

$$\mathcal{V}_n \equiv \{\boldsymbol{v} : \boldsymbol{v} \text{ is a permutation of } \boldsymbol{x}_{(\cdot)} = (x_{(1)}, \cdots, x_{(n)})\} \tag{5.1}$$

Then the conditional distribution of $T(\boldsymbol{X})$ given $\boldsymbol{X}_{(\cdot)} = \boldsymbol{x}_{(\cdot)}$ under H_0 is expressed as

$$P_0(T(\boldsymbol{X}) \le t | \, \boldsymbol{X}_{(\cdot)} = \boldsymbol{x}_{(\cdot)}) = \frac{1}{n!} \#\{\boldsymbol{v} : \sum_{i=1}^{k} n_i (\bar{v}_{i\cdot} - \bar{v}_{\cdot\cdot})^2 \le t, \boldsymbol{v} \in \mathcal{V}_n\}, \tag{5.2}$$

where $\#A$ stands for the number of elements of the set A. The test procedure based on $T(\boldsymbol{X})$ under the conditional probability measure $P_0(\cdot| \, \boldsymbol{X}_{(\cdot)} = \boldsymbol{x}_{(\cdot)})$ is distribution-free. The test based on $T(\boldsymbol{X})$ is equivalent to the conditional F-test.

Next let R_{ij} be the rank of X_{ij} among $\{X_{ij} : \ j = 1, \ldots, n_i, i = 1, \ldots, k\}$. The Kruskal-Wallis rank test statistic is given by

$$T(\boldsymbol{R}) = \sum_{i=1}^{k} n_i \left(\bar{R}_{i\cdot} - \frac{n+1}{2} \right)^2,$$

where $\bar{R}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} R_{ij}$. The rank test based on $T(\boldsymbol{R})$ is distribution-free.

The M-test statistic stated in Shiraishi (1996) is given by

$$T(\boldsymbol{\psi}) = \sum_{i=1}^{k} n_i \{\bar{\psi}_{i\cdot}(\boldsymbol{X}) - \bar{\psi}_{\cdot\cdot}(\boldsymbol{X})\}^2,$$

where

$$\bar{\psi}_{i\cdot}(\boldsymbol{X}) = \frac{1}{n_i} \sum_{j=1}^{n_i} \psi \left(\frac{X_{ij} - \tilde{\nu}}{\breve{\sigma}_n} \right), \text{ and } \bar{\psi}_{\cdot\cdot}(\boldsymbol{X}) = \frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{n_i} \psi \left(\frac{X_{ij} - \tilde{\nu}}{\breve{\sigma}_n} \right).$$

The conditional distribution of $T(\boldsymbol{\psi})$ given $\boldsymbol{X}_{(\cdot)} = \boldsymbol{x}_{(\cdot)}$ under H_0 is expressed on a parallel to the equation (5.2) as

$$P_0(T(\boldsymbol{\psi}) \le t | \, \boldsymbol{X}_{(\cdot)} = \boldsymbol{x}_{(\cdot)}) = \frac{1}{n!} \#\{\boldsymbol{v} : \sum_{i=1}^{k} n_i \{\bar{\psi}_{i\cdot}(\boldsymbol{v}) - \bar{\psi}_{\cdot\cdot}(\boldsymbol{v})\}^2 \le t, \boldsymbol{v} \in \mathcal{V}_n\}.$$

Hence the M-test based on $T(\boldsymbol{\psi})$ is also distribution-free.

The asymptotic relative efficiency (ARE) among the F-test, the Kruskal-Wallis rank test and the M-test agrees with the ARE among the least squares estimators, the R-estimators and the M-estimators appeared in Tables 5-7. Therefor we propose the exploratory test procedure based on

$$\begin{cases} T(\boldsymbol{X}) & \text{if } \ d_0 = d_{NM}, \\ T(\boldsymbol{R}) & \text{if } \ d_0 = d_{DE}, \quad \text{or} \quad \hat{\ell}_1 \geq 1.0, \quad \text{or} \quad \hat{\ell}_2 \geq 0.5, \\ T(\boldsymbol{\psi}) & \text{otherwise}, \end{cases} \tag{5.3}$$

and we refer to this test as T^*-test. We simulate the power for the tests based on $T(\boldsymbol{X})$, $T(\boldsymbol{R})$, $T(\boldsymbol{\psi})$, and T^*-test under the alternative $A_n : \ \tau_i = 5i/\sqrt{kn} \ \ (i = 1, \ldots, k)$. For k, n_i's and $F(x)$, we limited the same attention as in Section 4. We set $c = 1.579$ in $T(\boldsymbol{\psi})$. The values of the power for the tests based on $T(\boldsymbol{X}),\ T(\boldsymbol{R}),\ T(\boldsymbol{\psi})$, and T^*-test are appeared in Tables 10 and 11 for $k = 3, 5$, respectively. The values of the power are estimated by Monte-Carlo simulation of $5,000$ samples.

From Tables 10 and 11, we get the following conclusion (1) similar to the conclusions of the estimators stated in Section 4. (1) The power of the new test is stable in comparison with the other tests. Furthermore we get the following conclusion (2). (2) Surprisingly the new test is a little better than the permutation F-test when the underlying distribution is normal.

Table 10: The power of the permutation F test, the Kruskal-Wallis test, the permutation M test and the new test for $k = 3$.

(i) $F(x)$ is normal

n_1	New test (T^*)	F-test ($T(\boldsymbol{X})$)	Rank test($T(\boldsymbol{R})$)	M-test ($T(\boldsymbol{\psi})$)
10	0.505	0.502	0.475	0.494
20	0.526	0.520	0.497	0.510
30	0.534	0.531	0.514	0.524

(ii) $F(x)$ is contaminated normal

n_1	New test (T^*)	F-test ($T(\boldsymbol{X})$)	Rank test($T(\boldsymbol{R})$)	M-test ($T(\boldsymbol{\psi})$)
10	0.601	0.568	0.582	0.590
20	0.614	0.562	0.600	0.609
30	0.628	0.570	0.619	0.630

(iii) $F(x)$ is logistic

n_1	New test (T^*)	F-test ($T(\boldsymbol{X})$)	Rank test($T(\boldsymbol{R})$)	M-test ($T(\boldsymbol{\psi})$)
10	0.544	0.516	0.525	0.531
20	0.565	0.528	0.557	0.556
30	0.581	0.544	0.576	0.570

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

n_1	New test (T^*)	F-test ($T(\boldsymbol{X})$)	Rank test($T(\boldsymbol{R})$)	M-test ($T(\boldsymbol{\psi})$)
10	0.608	0.537	0.604	0.580
20	0.642	0.542	0.647	0.597
30	0.693	0.549	0.696	0.638

(v) $F(x)$ is lognormal

n_1	New test (T^*)	F-test ($T(\boldsymbol{X})$)	Rank test($T(\boldsymbol{R})$)	M-test ($T(\boldsymbol{\psi})$)
10	0.934	0.719	0.934	0.851
20	0.981	0.682	0.981	0.902
30	0.994	0.668	0.994	0.923

(vi) $F(x)$ is Weible

n_1	New test (T^*)	F-test ($T(\boldsymbol{X})$)	Rank test($T(\boldsymbol{R})$)	M-test ($T(\boldsymbol{\psi})$)
10	0.237	0.074	0.237	0.087
20	0.307	0.059	0.307	0.073
30	0.351	0.057	0.351	0.077

(vii) $F(x)$ is exponential

n_1	New test (T^*)	F-test ($T(\boldsymbol{X})$)	Rank test($T(\boldsymbol{R})$)	M-test ($T(\boldsymbol{\psi})$)
10	0.726	0.565	0.726	0.638
20	0.814	0.564	0.814	0.685
30	0.845	0.569	0.845	0.701

Table 11: The power of the permutation F test, the Kruskal-Wallis test, the permutation M test and the new test for $k = 5$.

(i) $F(x)$ is normal

n_1	New test (T^*)	F-test ($T(\boldsymbol{X})$)	Rank test($T(\boldsymbol{R})$)	M-test ($T(\boldsymbol{\psi})$)
10	0.674	0.670	0.645	0.659
20	0.696	0.694	0.671	0.681
30	0.706	0.705	0.683	0.693

(ii) $F(x)$ is contaminated normal

n_1	New test (T^*)	F-test ($T(\boldsymbol{X})$)	Rank test($T(\boldsymbol{R})$)	M-test ($T(\boldsymbol{\psi})$)
10	0.754	0.693	0.738	0.751
20	0.774	0.701	0.770	0.775
30	0.783	0.715	0.780	0.790

(iii) $F(x)$ is logistic

n_1	New test (T^*)	F-test ($T(\boldsymbol{X})$)	Rank test($T(\boldsymbol{R})$)	M-test ($T(\boldsymbol{\psi})$)
10	0.716	0.685	0.705	0.705
20	0.729	0.692	0.729	0.724
30	0.739	0.703	0.741	0.734

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

n_1	New test (T^*)	F-test ($T(\boldsymbol{X})$)	Rank test($T(\boldsymbol{R})$)	M-test ($T(\boldsymbol{\psi})$)
10	0.783	0.689	0.787	0.755
20	0.818	0.705	0.821	0.783
30	0.842	0.698	0.846	0.794

(v) $F(x)$ is lognormal

n_1	New test (T^*)	F-test ($T(\boldsymbol{X})$)	Rank test($T(\boldsymbol{R})$)	M-test ($T(\boldsymbol{\psi})$)
10	0.990	0.790	0.990	0.948
20	0.999	0.787	0.999	0.977
30	1.000	0.780	1.000	0.983

(vi) $F(x)$ is Weible

n_1	New test (T^*)	F-test ($T(\boldsymbol{X})$)	Rank test($T(\boldsymbol{R})$)	M-test ($T(\boldsymbol{\psi})$)
10	0.292	0.059	0.292	0.079
20	0.396	0.066	0.396	0.082
30	0.466	0.064	0.466	0.077

(vii) $F(x)$ is exponential

n_1	New test (T^*)	F-test ($T(\boldsymbol{X})$)	Rank test($T(\boldsymbol{R})$)	M-test ($T(\boldsymbol{\psi})$)
10	0.884	0.706	0.883	0.811
20	0.948	0.713	0.948	0.853
30	0.961	0.714	0.961	0.864

References

[1] Hodges, J. L. Jr. and Lehmann, E. L. (1963). Estimates of location based on rank tests. *Ann. Math. Statist.* **34**. 598-611.

[2] Hájek, J., Šidák, Z. and Sen, P. K. (1999). *Theory of Rank Tests, 2nd Edition.* Academic Press.

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

[4] Huber, P. J. (1981). *Robust Statistics.* New York: John Wiley.

[5] Lehmann, E. L. (1975). *Nonparametrics: Statistical Methods Based on Ranks.* San Francisco: Holden-Day.

[6] Saleh, A.K.Md.E. and T. Shiraishi (1993). On robust estimation for the parameters of multiple-design multivariate linear models under general restriction. *J. Nonparametric Statist.* **2**, 295-305.

[7] Shiraishi, T. (1990). R-estimators and confidence regions in one-way MANOVA. *J. Statist. Plan. Infer.* **24**, 203-214.

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

[9] Shiraishi, T. (2005). One-sample exploratory procedures after searching the underlying distribution. *J. Jpn. Soc. Comp. Statist.* **18**, 47-60.

[10] Shiraishi, T. (2007). Multiple comparisons based on studenized M-statistics in the one-way layout. *J. Statistical Studies* **26**, 105-118.