Ann. Inst. Statist. Math.
Vol. 54, No. 3, 641–658 (2002)

MODIFIED MAXIMUM LIKELIHOOD ESTIMATORS BASED ON RANKED SET SAMPLES

GANG ZHENG[1] AND MOHAMMAD F. AL-SALEH[2]

[1] *Office of Biostatistics Research, National Heart, Lung, and Blood Institute, II Rockledge Centre, 6701 Rockledge Drive, Bethesda, MD 20892-7938, U.S.A.*, e-mail: zhengg@nhlbi.nih.gov
[2] *Department of Mathematics and Statistics, Sultan Qaboos University, P.O. Box 36 Al-Khodh, Postal Code 123, Sultanate of Oman*, e-mail: malsaleh@squ.edu.om

(Received September 13, 2000; revised March 26, 2001)

Abstract. The maximum likelihood estimator (MLE) using a ranked set sample (RSS) usually has no closed expression because the maximum likelihood equation involves both hazard and inverse hazard functions, and may no longer be efficient when the judgment ranking is imperfect. In this paper, we consider a modified MLE (MMLE) using RSS for general parameters, which has the same expression as the MLE using a simple random sample (SRS), except that the SRS in the MLE is replaced by the RSS. The results show that, for the location parameter, the MMLE is always more efficient than the MLE using SRS, and for the scale parameter, the MMLE is at least as efficient as the MLE using SRS, when the same sample size is used. Under the perfect judgment ranking, numerical examples also show that the MMLE has good efficiency relative to the MLE based on RSS. When the judgment error is present, we conduct simulations to show that the MMLE is more robust than the MLE using RSS.

Key words and phrases: Asymptotic relative efficiency, estimating equation, judgment error, modified maximum likelihood equation, order statistics, ranked set sampling, robustness.

1. Introduction

Suppose $X_1, \ldots, X_n$ is a simple random samples (SRS) of size n from the cumulative distribution function (cdf) $F(x;\theta)$ with the probability density function (pdf) $f(x;\theta)$. Denote its order statistics by $X_{1:n} \leq \cdots \leq X_{n:n}$. To estimate the mean and variance of a characteristic of a population, if sampling units are difficult or expensive to measure, but it is relatively easy to rank them by judgment without actual measurements, a ranked set sample (RSS) can be used. McIntyre (1952) first proposed the RSS for the estimation of the population mean. The RSS of McIntyre (1952) is described as follows. Suppose k^2 sampling units are identified and partitioned into k sets, each having k units. In the i-th set, sampling units are ranked by judgment in an ascending order, and the i-th judgment ranked unit, denoted by $X_{(i)}$, is selected for the actual measurement, $i = 1, \ldots, k$. Thus the RSS consists of $X_{(1)}, \ldots, X_{(n)}$, which are independent random variables. This entire cycle can be repeated for m times to obtain an m cycle RSS with total sample size $n = mk$, denoted by $X_{(i)j}$, $i = 1, \ldots, k$ and $j = 1, \ldots, m$. Under the perfect judgment ranking, $X_{(i)j}$ follows the same distribution as $X_{i:k}$, i.e., its pdf can be written as

$$f_{i:k}(x;\theta) = \frac{k!}{(i-1)!(k-i)!}[F(x;\theta)]^{i-1}f(x;\theta)[1-F(x;\theta)]^{k-i}. \tag{1.1}$$

The RSS has many applications in biology, agricultural and environmental studies (e.g., McIntyre (1952), Takahasi and Wakimoto (1968), Dell and Clutter (1972), Stokes (1980), Stokes and Sager (1988), Yu and Lam (1997), Barnett (1999), Al-Saleh and Al-Kadiri (2000), and Al-Saleh and Zheng (2002)). Kvam and Samaniego (1993, 1994) and El-Neweihi and Sinha (2000) also applied the RSS to the reliability theory. Recently, many parametric, non-parametric and Bayesian methods based on RSS are proposed. For a bibliography, see Patil *et al.* (1999).

When the underlying distribution is unknown, to estimate the population mean based on RSS of size $n = mk$, McIntyre (1952) proposed the following sample mean,

$$\hat{\mu}_{\text{RSS}} = \bar{X}_{\text{RSS}} = \frac{1}{mk}\sum_{j=1}^{m}\sum_{i=1}^{k} X_{(i)j}. \tag{1.2}$$

Takahasi and Wakimoto (1968) showed that this unbiased estimator using RSS has smaller variance than the sample mean based on SRS, $\bar{X}_{\text{SRS}}$, with the same size $n = mk$. To estimate the variance, a natural estimator based on RSS is

$$\hat{\sigma}^2_{\text{RSS}} = \frac{1}{mk-1}\sum_{j=1}^{m}\sum_{i=1}^{k}(X_{(i)j} - \bar{X}_{\text{RSS}})^2. \tag{1.3}$$

The properties of $\hat{\sigma}^2_{\text{RSS}}$ were discussed by Stokes (1980). When m is large enough, Stokes (1980) showed that the variance of $\hat{\sigma}^2_{\text{RSS}}$ is smaller than that of $\hat{\sigma}^2_{\text{SRS}}$, where

$$\hat{\sigma}^2_{\text{SRS}} = \frac{1}{mk-1}\sum_{i=1}^{mk}(X_i - \bar{X}_{\text{SRS}})^2.$$

Stokes (1980) also calculated the limiting relative precision (RP),

$$\text{RP} = \lim_{n\to\infty} \text{Var}(\hat{\sigma}^2_{\text{SRS}})/\text{MSE}(\hat{\sigma}^2_{\text{RSS}}),$$

for several distributions with set size $2 \le k \le 5$, where $\text{MSE}(\hat{\sigma}^2_{\text{RSS}})$ is the mean square error (MSE) of the estimate.

When the underlying distribution is known, other parametric estimators have been proposed recently. For example, the maximum likelihood estimator (MLE) and the best linear unbiased estimator (e.g., Stokes (1995), Sinha *et al.* (1996), and Chen (2000)). For the normal distribution $N(\mu, \sigma^2)$, $\bar{X}_{\text{SRS}}$ and $[(mk-1)/(mk)]\hat{\sigma}_{\text{SRS}}$ are MLE's for μ and σ^2, respectively, based on SRS of size $n = mk$. Generally, since both hazard function $f(x;\theta)/(1-F(x;\theta))$ and inverse hazard functions $f(x;\theta)/F(x;\theta)$ are involved in the likelihood equation of RSS, there is no closed expression for MLE for parameters of the underlying distribution. Besides, the judgment error is inevitable in practice. When the judgment ranking error is present in RSS, the MLE may no longer be efficient. Therefore, we consider a modified MLE (MMLE) for general parameters. We will show the MMLE is relatively easy to obtain and, for a small set size k, is efficient relative to the MLE based on the RSS when there is no judgment error, and that it is more robust than the MLE based on RSS when the judgment ranking is imperfect.

The rest of the paper is organized as follows. We define the MMLE in Section 2. In Section 3, the properties of the MMLE are studied under the perfect judgment ranking. The asymptotic relative efficiencies (ARE) of the MMLE to the MLE using SRS and MLE using RSS are given in Section 4. In Section 5, we consider the MMLE under the imperfect ranking. Simulation results to compare the MMLE and MLE using RSS are reported. Concluding remarks are given in Section 6.

2. MMLE based on the RSS

Let $X_1, \ldots, X_n$ be a SRS of size n from the cdf $F(x;\theta)$ with pdf $f(x;\theta)$, where $x \in R^1$ and $\theta \in \Theta$, an open set in R^p. Assume the distribution $F(x;\theta)$ has the same support and that the MLE for θ, denoted by $\hat{\theta}_{\text{MLE,SRS}}$, satisfies the maximum likelihood equation

$$\sum_{i=1}^{n} \frac{f'(X_i;\theta)}{f(X_i;\theta)} = 0, \tag{2.1}$$

where the derivative is with respect to $\theta = [\theta_1 \cdots \theta_p]^T$, i.e.,

$$f'(x;\theta) = \left[\frac{\partial}{\partial\theta_1} f(x;\theta) \cdots \frac{\partial}{\partial\theta_p} f(x;\theta)\right]^T.$$

Generally, solving MLE from (2.1) needs iterative numerical calculations. Under suitable conditions (Lehmann (1983), p. 429), if θ is the true value, the MLE is strongly consistent, $P(\lim_{n\to\infty} \hat{\theta}_{\text{MLE,SRS}} = \theta) = 1$, and asymptotically normal,

$$n^{1/2}(\hat{\theta}_{\text{MLE,SRS}} - \theta) \xrightarrow{D} N_p(0, I^{-1}(\theta)), \tag{2.2}$$

where $I(\theta)$ is the Fisher information matrix about θ contained in the random variable X, defined as

$$I(\theta) = E_\theta\left\{\left[\frac{\partial}{\partial\theta} \log f(X;\theta)\right]\left[\frac{\partial}{\partial\theta} \log f(X;\theta)\right]^T\right\}. \tag{2.3}$$

We will use the following two basic identities of order statistics and RSS. If we assume $\int_{R^1} f'(x;\theta)dx = 0$, where the derivative is with respect to θ, then

$$\sum_{i=1}^{k} f_{i:k}(x;\theta) = kf(x;\theta) \quad \text{and} \tag{2.4}$$

$$\sum_{i=1}^{k} E_\theta \frac{f'(X_{(i)1};\theta)}{f(X_{(i)1};\theta)} = 0, \tag{2.5}$$

where (2.4) was given by Takahasi and Wakimoto (1968), and (2.5) can be obtained from (2.4) and the assumption.

The log-likelihood function based on the RSS of size $n = mk$, $X_{(i)j}$, $i = 1, \ldots, k$ and $j = 1, \ldots, m$, subject to a constant, is given by

$$\sum_{j=1}^{m}\sum_{i=1}^{k}\{\log f(X_{(i)j};\theta) + (i-1)\log F(X_{(i)j};\theta) + (k-i)\log \bar{F}(X_{(i)j};\theta)\},$$

where $\bar{F}(x;\theta) = 1 - F(x;\theta)$. Then, under the suitable conditions, the MLE for θ using the RSS satisfies the maximum likelihood equation

$$\sum_{j=1}^{m}\sum_{i=1}^{k}\left\{\frac{f'(X_{(i)j};\theta)}{f(X_{(i)j};\theta)} + (i-1)\frac{F'(X_{(i)j};\theta)}{F(X_{(i)j};\theta)} + (k-i)\frac{\bar{F}'(X_{(i)j};\theta)}{\bar{F}(X_{(i)j};\theta)}\right\} = 0, \tag{2.6}$$

where the derivative is with respect to $\theta = [\theta_1 \cdots \theta_p]^T$. Obviously, (2.6) is more difficult to solve for θ than (2.1), even for the normal and exponential distributions, since $\bar{F}'/\bar{F}$ involves the hazard function and F'/F involves the inverse hazard function. Once the MLE for θ from (2.6) is obtained, denoted by $\hat{\theta}_{\mathrm{MLE,RSS}}$, it is strongly consistent and asymptotically normal under the suitable conditions (Stokes (1995), Chen (2000), and Bai and Chen (2000)).

In this paper, we modify (2.6) to obtain a modified maximum likelihood equation by replacing the second and third terms on the left hand side of (2.6) by their expectations. The idea of replacing the hazard rate in the maximum likelihood equation by its expectation was proposed by Mehrotra and Nanda (1974), who estimated parame ters of normal and gamma distributions based on Type II censored data. Mehrotra and Nanda (1974) considered a single location or scale parameter of the normal and Gamma distributions. In the following, we will consider general parameters. The asymptotics of the results of Mehrotra and Nanda (1974) were considered by Bhattacharyya (1985).

Two things are worth noting about the modified maximum likelihood equation. First, one referee pointed out that the modified maximum likelihood equation is still an unbiased estimating equation (Godambe (1960, 1991)). To see this, denote (2.6) by $w(\boldsymbol{X};\theta) = 0$. Then we have $E_\theta(w(\boldsymbol{X};\theta)) = 0$. Thus, from (2.5), the modified maximum likelihood equation can be simplified as

$$\sum_{j=1}^{m}\sum_{i=1}^{k}\frac{f'(X_{(i)j};\theta)}{f(X_{(i)j};\theta)} = 0, \tag{2.7}$$

i.e., the sum of the expectations of the second and third terms on the left hand side of (2.6) is zero. Thus (2.5) implies that the modified maximum likelihood equation given by (2.7) is still an unbiased estimating equation and (2.7) implies that we actually use the partial likelihood for estimation. Second, in practice, the set size of RSS k is usually small ($2 \le k \le 5$), because it is difficult to rank a large number of sampling units without actual measurements. From Bhattacharyya (1985), using the modified maximum likelihood equation for the Type II censored data, the estimates for the location and scale parameters of the normal distribution are highly efficient relative to the MLE's based on the complete maximum likelihood equation. For example, the asymptotic relative efficiency of the estimate based on the modified maximum likelihood equation relative to the MLE based on the complete maximum likelihood equation is .80 for the normal location parameter when only 20% of the data is observed. In the RSS setting, for each

$i = 1, \ldots, k$ in (2.6), we have either Type II (if $i = 1$), or doubly (if $1 < i < k$), or left (if $i = k$) censored data. But we still expect the MMLE will retain good efficiency relative to the MLE using the RSS when k is small. On the other hand, for each $i = 1, \ldots, k$, the first term in (2.6) corresponds to the partial log-likelihood for the observed data while the second and third terms correspond to the probabilities for censored data, i.e., there are $i - 1$ and $k - i$ observations less and greater than the smallest and largest observed order statistics, respectively. Thus, intuitively, the first term in (2.6) contributes more information into the complete likelihood than the other two terms. However, as k increases, the contribution of the first term decreases, so the loss of efficiency of MMLE relative to MLE using RSS may increase as k increases.

Suppose the MMLE satisfies (2.7). From (2.1) and (2.7), we obtain:

THEOREM 2.1. *The MMLE based on RSS for general parameters has the same expression as the MLE based on SRS, but the SRS is replaced by the RSS. Moreover, if the MLE based on SRS is unique, then the MMLE is also unique.*

For example, for the normal distribution $N(\mu, \sigma^2)$, we obtain the MMLE for μ and σ^2 as

$$\hat{\mu}_{\mathrm{MMLE}} = \frac{1}{mk} \sum_{j=1}^{m} \sum_{i=1}^{k} X_{(i)j},$$

and

$$\hat{\sigma}^2_{\mathrm{MMLE}} = \frac{1}{mk} \sum_{j=1}^{m} \sum_{i=1}^{k} (X_{(i)j} - \hat{\mu}_{\mathrm{MMLE}})^2,$$

respectively. Note that, from (1.2) and (1.3), $\hat{\mu}_{\mathrm{MMLE}} = \hat{\mu}_{\mathrm{RSS}}$, and $\hat{\sigma}^2_{\mathrm{MMLE}}$ is asymptotically equivalent to $\hat{\sigma}^2_{\mathrm{RSS}}$ for the normal distribution. Thus it is very easy to obtain the MMLE using RSS, when the MLE using SRS has a closed expression, or, if the iterative procedures to solve the MLE using SRS have already been programmed, because we can use these programs to find the MMLE without doing new iterative calculations, which are required for the MLE using RSS from the original likelihood equation (2.6).

3. Properties of MMLE based on RSS

In this section, we show that, under the usual regularity conditions, the MMLE, $\hat{\theta}_{\mathrm{MMLE}}$, is strongly consistent and asymptotically normal. Let

$$T_m(\theta) = \frac{1}{m} \sum_{j=1}^{m} \sum_{i=1}^{k} \frac{f'(X_{(i)j}; \theta)}{f(X_{(i)j}; \theta)} = \frac{1}{m} \sum_{j=1}^{m} g(Y_j; \theta),$$

where $g(Y_j; \theta) = \sum_{i=1}^{k} f'(X_{(i)j}; \theta)/f(X_{(i)j}; \theta)$, and $\theta = [\theta_1 \cdots \theta_p]^T$. First we have,

LEMMA 3.1. *Suppose the judgment ranking is perfect. Assume that each coordinate of $g(Y_j; \theta) = [g_1 \cdots g_p]^T$ is a continuous function, that $0 < \int_{R^1} g_l^2(x) f(x) dx < \infty$, $l = 1, \ldots, p$, and that $\int_{R^1} f'(x; \theta) dx = 0$. Let $n = mk$. Then, as $m \to \infty$ and k is fixed,*

$$n^{1/2} T_m(\theta) \xrightarrow{D} N_p(0, \Omega(\theta)),$$

where $\Omega(\theta) = k^2 I(\theta) - k\sum_{i=1}^k E_\theta[b_i(\theta)]E_\theta[b_i(\theta)]^T$, $I(\theta)$ *is the Fisher information matrix given by* (2.3), *and* $b_i(\theta) = f'(X_{(i)1};\theta)/f(X_{(i)1};\theta)$, $i = 1,\ldots,k$.

PROOF. The normality follows from that $g(Y_j;\theta)$, $j = 1,\ldots,m$, are independent identically distributed (iid) random vectors, that, from (2.5),

$$E_\theta[T_m(\theta)] = E_\theta[g(Y_1;\theta)] = 0,$$

and that, by the independence of RSS,

$$
\begin{aligned}
\text{(3.1)} \qquad \mathrm{Var}_\theta(T_m(\theta)) &= \frac{1}{m}E_\theta\left\{\left[\sum_{i=1}^k \frac{f'(X_{(i)1};\theta)}{f(X_{(i)1};\theta)}\right]\left[\sum_{i=1}^k \frac{f'(X_{(i)1};\theta)}{f(X_{(i)1};\theta)}\right]^T\right\} \\
&= \frac{1}{m}\sum_{i=1}^k E_\theta\left\{\left[\frac{f'(X_{(i)1};\theta)}{f(X_{(i)1};\theta)}\right]\left[\frac{f'(X_{(i)1};\theta)}{f(X_{(i)1};\theta)}\right]^T\right\} \\
&\quad + \frac{1}{m}\sum_{i\neq i'} E_\theta\left[\frac{f'(X_{(i)1};\theta)}{f(X_{(i)1};\theta)}\right] E_\theta\left[\frac{f'(X_{(i')1};\theta)}{f(X_{(i')1};\theta)}\right]^T \\
&= \frac{k}{m}E_\theta\left\{\left[\frac{f'(X;\theta)}{f(X;\theta)}\right]\left[\frac{f'(X;\theta)}{f(X;\theta)}\right]^T\right\} \\
&\quad - \frac{1}{m}\sum_{i=1}^k E_\theta\left[\frac{f'(X_{(i)1};\theta)}{f(X_{(i)1};\theta)}\right] E_\theta\left[\frac{f'(X_{(i)1};\theta)}{f(X_{(i)1};\theta)}\right]^T, \\
&= \frac{k}{m}I(\theta) - \frac{1}{m}\sum_{i=1}^k E_\theta[b_i(\theta)]E_\theta[b_i(\theta)]^T,
\end{aligned}
$$

where (3.1) is obtained by using (2.4) and (2.5). □

From the proof, we can also obtain another expression of $\Omega(\theta)$ as

$$\Omega(\theta) = k\sum_{i=1}^k \{E_\theta[b_i(\theta)b_i(\theta)^T] - E_\theta[b_i(\theta)]E_\theta[b_i(\theta)]^T\} = k\sum_{i=1}^k \mathrm{Var}_\theta(b_i(\theta)).$$

The following lemma of uniformly strong consistency will be used to show the strong consistency of MMLE.

LEMMA 3.2. *Let* θ_0 *be the true value of* $\theta \in \Theta$ *and* $B \subset \Theta$ *be an arbitrary compact set such that* $\theta_0 \in B$. *Let the judgment ranking be perfect and* $h(Y_j;\theta) = \sum_{i=1}^k \log f(X_{(i)j};\theta)$ *and* $L_m(\theta) = \sum_{j=1}^m h(Y_j;\theta)/m$. *We assume:*

(1) $h(y;\theta)$ *is continuous in* $\theta \in B$ *for every* y.

(2) $|h(y;\theta)| < M(y)$ *and* $\int_{R^1} M(x)f(x)dx < \infty$, *for* $\theta \in \Theta$ *and all* y.

Then, with probability 1,

$$\sup_{\theta\in B}\left|L_m(\theta) - E_\theta\sum_{i=1}^k \log f(X_{(i)1};\theta)\right| \to 0.$$

PROOF. Using (2.4), Lemma 3.2 is a special case of Theorem 2 of Bhattacharyya (1985). □

Before we state the properties of the MMLE, we need the following regularity conditions.

(C1)

$$\int_{R^1} \frac{\partial}{\partial \theta_l} f(x;\theta)dx = 0, \qquad l = 1, \ldots, p.$$

(C2) For $\theta \in \Theta$, the Fisher information $I(\theta)$ satisfies $0 < I(\theta) = (I_{ij}(\theta))_{p\times p} < \infty$.

(C3) There exists $M(x)$ such that $\int_{R^1} M(x)f(x;\theta)dx < \infty$ for $\theta \in \Theta$ and

$$\left| \frac{\partial^3 \log f(x;\theta)}{\partial\theta_a \partial\theta_b \partial\theta_c} \right| \leq M(x), \qquad a, b, c = 1, \ldots, p.$$

(C4) The distribution is identifiable, i.e.,

$$P_{\theta_0}(f(x;\theta) \neq f(x;\theta_0)) > 0.$$

THEOREM 3.1. *We assume the judgment ranking is perfect. Let* $n = mk$. *Under the conditions* (C1) *to* (C4), *the MMLE*, $\hat{\theta}_{\text{MMLE}}$, *solved from* (2.7) *has the following properties*:

(i) $\hat{\theta}_{\text{MMLE}}$ *is strongly consistent as* $m \to \infty$ *and* k *is fixed.*

(ii) *Let* θ_0 *be the true value of* θ. *As* $m \to \infty$ *and* k *is fixed,*

$$n^{1/2}(\hat{\theta}_{\text{MMLE}} - \theta_0) \xrightarrow{D} N_p(0, \Delta(\theta_0)), \tag{3.2}$$

where $\Delta(\theta_0) = I^{-1}(\theta_0) - \sum_{i=1}^{k} I^{-1}(\theta_0)E_{\theta_0}[b_i(\theta_0)]E_{\theta_0}[b_i(\theta_0)]^T I^{-1}(\theta_0)/k$, *and* $b_i(\theta)$ *is defined in Lemma* 3.1.

PROOF. Let B and $L_m(\theta)$ be defined as in Lemma 3.2. To show (i), we consider the log-likelihood ratio $L_m(\theta) - L_m(\theta_0)$ and obtain, from Lemma 3.2, as $m \to \infty$,

$$\begin{aligned} d(\theta) &= L_m(\theta) - L_m(\theta_0) \\ &\to E_{\theta_0} \sum_{i=1}^{k} \log f(X_{(i)1};\theta) - E_{\theta_0} \sum_{i=1}^{k} \log f(X_{(i)1};\theta_0), \end{aligned} \tag{3.3}$$

with probability 1 and uniformly in $\theta \in B$. We want to show that $d(\theta)$ has a local maximum at θ_0. Note that $d(\theta_0) = 0$. Thus we need to show that $d(\theta) < 0$ for $\theta \neq \theta_0$ and $\theta \in B$. From (2.4) and (3.3), we have

$$\begin{aligned} d(\theta) &\to k[E_{\theta_0} \log f(X;\theta) - E_{\theta_0} \log f(X;\theta_0)] \\ &= kE_{\theta_0}\left[\log \frac{f(X;\theta)}{f(X;\theta_0)}\right] \leq k \log \left[E_{\theta_0} \frac{f(X;\theta)}{f(X;\theta_0)}\right] = 0, \end{aligned}$$

by Jensen's inequality, where the equality holds if and only if $\theta = \theta_0$. By (C4), we obtain $d(\theta)$ has a local maximum at θ_0. Since B is arbitrary, the strong consistency of $\hat{\theta}_{\text{MMLE}}$ solved from (2.7) follows from the standard arguments.

To show (ii), let T_m be defined as in Lemma 3.1 and let $\hat{\theta}_{\text{MMLE}}$ satisfy (2.7). We consider the case of a single parameter. By the expansion of $T_m(\hat{\theta}_{\text{MMLE}})$ at θ_0,

$$T_m(\hat{\theta}_{\text{MMLE}}) - T_m(\theta_0) = T'_m(\theta_0)(\hat{\theta}_{\text{MMLE}} - \theta_0) + \frac{1}{2}T''_m(\epsilon)(\hat{\theta}_{\text{MMLE}} - \theta_0)^2,$$

where $T_m(\hat{\theta}_{\text{MMLE}}) = 0$ and ϵ is between $\hat{\theta}_{\text{MMLE}}$ and θ_0. Hence

$$n^{1/2}(\hat{\theta}_{\text{MMLE}} - \theta_0) = -\frac{n^{1/2}T_m(\theta_0)}{T'_m(\theta_0) + \frac{1}{2}T''_m(\epsilon)(\hat{\theta}_{\text{MMLE}} - \theta_0)}, \tag{3.4}$$

where $T''_m(\epsilon)$ is bounded from (C3) and by (i), $\hat{\theta}_{\text{MMLE}} - \theta_0 \to 0$, with probability 1, and from Lemma 3.1,

$$n^{1/2}T_m(\theta_0) \xrightarrow{D} N(0, \Omega(\theta_0)).$$

Finally, in (3.4), $T'_m(\theta_0) \to -kI(\theta_0)$ in probability by the law of large number. Therefore, we have

$$n^{1/2}(\hat{\theta}_{\text{MMLE}} - \theta_0) \xrightarrow{D} N\left(0, \frac{\Omega(\theta_0)}{k^2I^2(\theta_0)}\right),$$

i.e., $\Delta(\theta_0) = \Omega(\theta_0)/[k^2I^2(\theta_0)] = I^{-1}(\theta_0) - \sum_{i=1}^k E^2_{\theta_0}[b_i(\theta_0)]/[kI^2(\theta_0)]$.

For the multiparameter case, the proof of (ii) is basically the same as that of the single parameter case. However, the single equation derived from the expansion of $\hat{\theta}_{\text{MMLE}} - \theta_0$ will be replaced by a system of p equations which must be solved for $\hat{\theta}_{\text{MMLE},j} - \theta_{0,j}$, $j = 1, \ldots, p$. Thus the proof of the multiparameter case using RSS is an analogy of the proof of the multiparameter case from a single parameter using SRS (Lehmann (1983), p. 433). □

4. Asymptotic relative efficiency of MMLE

In this section, under the perfect judgment ranking, we compare the ARE's of MMLE relative to MLE using SRS and MLE using RSS with the same size. From Theorem 3.1, we have

$$I^{-1}(\theta) - \Delta(\theta) = \frac{1}{k}\sum_{i=1}^{k} I^{-1}(\theta)E[b_i(\theta)]E[b_i(\theta)]^T I^{-1}(\theta),$$

where the right side is a non-negative definite matrix. Thus $I^{-1}(\theta) \geq \Delta(\theta)$.

We are interested in the estimation of a function of θ, say, $\phi = \phi(\theta)$. From (2.2) and (3.2), the ARE of $\hat{\phi}_{\text{MMLE}} = \phi(\hat{\theta}_{\text{MMLE}})$ relative to $\hat{\phi}_{\text{MLE,SRS}} = \phi(\hat{\theta}_{\text{MMLE,SRS}})$ is given by

$$\begin{aligned} \text{ARE}(\hat{\phi}_{\text{MMLE}}, \hat{\phi}_{\text{MLE,SRS}}) &= \frac{\phi'(\theta)^T I^{-1}(\theta)\phi'(\theta)}{\phi'(\theta)^T\Delta(\theta)\phi'(\theta)} \\ &= 1 + \frac{1}{k}\sum_{i=1}^{k}\frac{[\phi'(\theta)^T I^{-1}(\theta)E(b_i)]^2}{\phi'(\theta)^T\Delta(\theta)\phi'(\theta)} \geq 1. \end{aligned} \tag{4.1}$$

This shows that the MMLE is at least as efficient as the MLE using SRS with the same size. The above equality holds if and only if $E(b_i) = 0$, for all $i = 1, \ldots, k$. We have the following result for the location and scale parameters.

THEOREM 4.1. *For the location parameter, $E(b_k) \neq 0$ for $k \geq 2$. For the scale parameter and $k = 2$, if the distribution F is symmetric about the location, then $E(b_i) = 0$, $i = 1, 2$.*

PROOF. For the location parameter, $E(b_k) = 0$ is equivalent to

$$0 = \int_{R^1} F^{k-1}(x) df(x) = f(x)F^{k-1}(x)|_{R^1} - (k-1)\int_{R^1} f^2(x)F^{k-2}(x)dx$$
$$= -(k-1)\int_{R^1} f^2(x)F^{k-2}(x)dx,$$

which shows that $E(b_k) \neq 0$ for $k \geq 2$. For the scale parameter and $k = 2$, we have

$$E(b_1 + b_2) = (2/\theta^2)\int_{R^1}(f(x) + xf'(x))dx = 0,$$

by integration by parts and

$$E(b_2 - b_1) = (2/\theta^2)\int_{R^1}(f(x) + xf'(x))(1 - 2F(x))dx = 0,$$

since the integrand is an odd function. Thus $E(b_1) = E(b_2) = 0$. □

Theorem 4.1 shows that, for the location parameter, the MMLE is always more efficient than the MLE using SRS, and that, for the scale parameter, the MMLE with set size 2 is as efficient as the MLE using SRS, with the same sample size.

Chen (2000) obtained the following asymptotic normality for the MLE of θ using RSS, $\hat{\theta}_{\text{MLE,RSS}}$,

$$n^{1/2}(\hat{\theta}_{\text{MLE,RSS}} - \theta) \xrightarrow{D} N(0, [I(\theta) + (k-1)\epsilon(\theta)]^{-1})$$

where

$$\epsilon(\theta) = E\left[\frac{\frac{\partial}{\partial\theta}F(X;\theta)\frac{\partial}{\partial\theta}F(X;\theta)^T}{F(X;\theta)\bar{F}(X;\theta)}\right]$$

is a non-negative definite matrix. To examine the loss of efficiency when the MMLE is used instead of the MLE based on RSS, we obtain the ARE of $\hat{\phi}_{\text{MMLE}}$ relative to $\hat{\phi}_{\text{MLE,RSS}}$ as

$$\text{ARE}(\hat{\phi}_{\text{MMLE}}, \hat{\phi}_{\text{MLE,RSS}}) = \frac{\phi'(\theta)^T[I(\theta) + (k-1)\epsilon(\theta)]^{-1}\phi'(\theta)}{\phi'(\theta)^T\Delta(\theta)\phi'(\theta)}. \tag{4.2}$$

Here, $kI(\theta) + k(k-1)\epsilon(\theta)$ is also the total Fisher information about θ contained in one cycle of RSS $X_{(i)1}$, $i = 1, \ldots, k$ (Chen (2000) and Zheng (2000)). For a single parameter θ_r, $r = 1, \ldots, p$, the above ARE's in (4.1) and (4.2) can be simplified as,

$$\text{ARE}(\hat{\theta}_{r,\text{MMLE}}, \hat{\theta}_{r,\text{MLE,SRS}}) = \frac{I_{rr}(\theta)}{I_{rr}(\theta) - \sum_{i=1}^k E^2(b_i)/k}, \tag{4.3}$$

$$\text{ARE}(\hat{\theta}_{r,\text{MMLE}}, \hat{\theta}_{r,\text{MLE,RSS}}) = \frac{[I_{rr}(\theta) + (k-1)\epsilon(\theta)]^{-1}}{I_{rr}^{-1}(\theta) - \sum_{i=1}^k E^2(b_i)/[kI_{rr}^2(\theta)]}, \tag{4.4}$$

respectively. We consider three examples.

Example 1. Normal distribution $N(\mu, \sigma^2)$. For the location parameter, $I_{11}(\theta) = 1/\sigma^2$, and $\epsilon(\theta) = 0.4805/\sigma^2$ (Chen (2000)). Hence, from (4.3) and (4.4),

$$\mathrm{ARE}(\hat{\mu}_{\mathrm{MMLE}}, \hat{\mu}_{\mathrm{MLE,SRS}}) = \frac{1}{1 - \sum_{i=1}^{k} E^2(Y_i)/k},$$

$$\mathrm{ARE}(\hat{\mu}_{\mathrm{MMLE}}, \hat{\mu}_{\mathrm{MLE,RSS}}) = \frac{[1 + 0.4805(k-1)]^{-1}}{1 - \sum_{i=1}^{k} E^2(Y_i)/k},$$

where Y_i is the i-th order statistic of a SRS from $N(0,1)$. For the scale parameter, $I_{22}(\theta) = 2/\sigma^2$, and $\epsilon(\theta) = 0.2701/\sigma^2$. Thus

$$\mathrm{ARE}(\hat{\sigma}^2_{\mathrm{MMLE}}, \hat{\sigma}^2_{\mathrm{MLE,SRS}}) = \frac{2}{2 - \sum_{i=1}^{k}[1 - \mathrm{Var}(Y_i) - E^2(Y_i)]^2/k},$$

$$\mathrm{ARE}(\hat{\sigma}^2_{\mathrm{MMLE}}, \hat{\sigma}^2_{\mathrm{MLE,RSS}}) = \frac{[2 + 0.2701(k-1)]^{-1}}{0.5 - \sum_{i=1}^{k}[1 - \mathrm{Var}(Y_i) - E^2(Y_i)]^2/(4k)},$$

where Y_i is the same as before. Note that the $\mathrm{ARE}(\hat{\sigma}^2_{\mathrm{MMLE}}, \hat{\sigma}^2_{\mathrm{MLE,SRS}})$ for the normal distribution is the same as the limiting RP, reported in Stokes (1980).

Example 2. Logistic distribution with pdf $f(x; \mu, \sigma) = \frac{1}{\sigma}\exp((x-\mu)/\sigma)/\{1 + \exp[(x-\mu)/\sigma]\}^2$. Thus $I_{11}(\theta) = 1/(3\sigma^2)$ and $\epsilon(\theta) = 1/(6\sigma^2)$. From (4.3) and (4.4), it can be shown that

$$\mathrm{ARE}(\hat{\mu}_{\mathrm{MMLE}}, \hat{\mu}_{\mathrm{MLE,RSS}}) = \frac{2}{k+1}\mathrm{ARE}(\hat{\mu}_{\mathrm{MMLE}}, \hat{\mu}_{\mathrm{MLE,SRS}}),$$

where

$$\mathrm{ARE}(\hat{\mu}_{\mathrm{MMLE}}, \hat{\mu}_{\mathrm{MLE,SRS}}) = \left[1 - \frac{3}{k}\sum_{i=1}^{k} E^2(b_i)\right]^{-1}$$

and $b_i = (\exp(Y_i) - 1)/(\exp(Y_i) + 1)$ where Y_i is the i-th order statistic from $f(x; 0, 1)$. It can be shown that b_i is the i-th order statistic of a SRS from the uniform distribution $U(-1, 1)$. Thus $E(b_i) = 2i/(k+1) - 1$, and

$$\sum_{i=1}^{k} E^2(b_i) = \frac{4}{(k+1)^2}\sum_{i=1}^{k} i^2 - \frac{4}{k+1}\sum_{i=1}^{k} i + k = \frac{k(k-1)}{3(k+1)}.$$

Hence $\mathrm{ARE}(\hat{\mu}_{\mathrm{MMLE}}, \hat{\mu}_{\mathrm{MLE,SRS}}) = (k+1)/2$, which implies that there is no loss of efficiency for the MMLE of μ for the logistic distribution relative to the MLE using RSS.

Example 3. Weibull distribution $W(\theta, \alpha)$ with cdf $F(x; \theta) = 1 - \exp(-(x/\theta)^\alpha)$, where α is known. Thus $I(\theta) = (\alpha/\theta)^2$ and $\epsilon(\theta) = .4041(\alpha/\theta)^2$. From (2.7), the MMLE for θ has the closed form,

$$\hat{\theta}_{\mathrm{MMLE}} = \left(\frac{1}{mk}\sum_{j=1}^{m}\sum_{i=1}^{k} X^{\alpha}_{(i)j}\right)^{1/\alpha}.$$

Table 1. Asymptotic efficiencies of the MMLE relative to the MLE using SRS and MLE using RSS.

Parameter θ	Distribution	Set size k	ARE($\hat{\theta}_{\mathrm{MMLE}}, \hat{\theta}_{\mathrm{MLE,SRS}}$)	ARE($\hat{\theta}_{\mathrm{MMLE}}, \hat{\theta}_{\mathrm{MLE,RSS}}$)
Location	Extreme value	2	1.3333	.9496
		3	1.6364	.9050
		4	1.9200	.8679
		5	2.1898	.8369
	Logistic	2	1.5000	1.0000
		3	2.0000	1.0000
		4	2.5000	1.0000
		5	3.0000	1.0000
	Normal	2	1.4669	.9906
		3	1.9137	.9759
		4	2.3496	.9623
		5	2.7702	.9480
Scale	Extreme value	2	1.0101	.8875
		3	1.0898	.8539
		4	1.1812	.8539
		5	1.2743	.8208
	Exponential	2	1.3333	.9496
	(Weibull)	3	1.6364	.9050
		4	1.9200	.8679
		5	2.1898	.8369
	Logistic	2	1.0000	.8693
		3	1.0958	.8425
		4	1.2119	.8352
		5	1.3340	.8330
	Normal	2	1.0000	.8810
		3	1.0822	.8521
		4	1.1792	.8392
		5	1.2791	.8305

To evaluate ARE's, we obtain

$$\mathrm{ARE}(\hat{\theta}_{\mathrm{MMLE}}, \hat{\theta}_{\mathrm{MLE,SRS}}) = \frac{1}{1-\sum_{i=1}^{k}[1-E(Y_i^{\alpha})]^2/k},$$
$$\mathrm{ARE}(\hat{\theta}_{\mathrm{MMLE}}, \hat{\theta}_{\mathrm{MLE,RSS}}) = \frac{[1+.4041(k-1)]^{-1}}{1-\sum_{i=1}^{k}[1-E(Y_i^{\alpha})]^2/k},$$

where Y_i is the i-th order statistic of a SRS from $W(1,\alpha)$. Thus the denominators in the above ARE's are independent of α, since Y_i^{α} is the i-th order statistic from $W(1,1)$, which is the exponential distribution. Hence the ARE's for θ do not change with the shape parameter α.

In Table 1, we report the results of the ARE's for the location parameter of the extreme value, logistic and normal distributions, and the scale parameter of the extreme value, exponential (Weibull), logistic, and normal distributions for the set size $2 \leq k \leq 5$. From Table 1, the MMLE has higher ARE relative to the MLE using SRS or RSS for the location parameter than the scale parameter. ARE($\hat{\theta}_{\mathrm{MMLE}}, \hat{\theta}_{\mathrm{MLE,SRS}}$) increases and ARE($\hat{\theta}_{\mathrm{MMLE}}, \hat{\theta}_{\mathrm{MLE,RSS}}$) decreases with the set size k. From Section 2, the decreasing of ARE($\hat{\theta}_{\mathrm{MMLE}}, \hat{\theta}_{\mathrm{MLE,RSS}}$) as k increases may be due to the less contribution of the partial log-likelihood (the first term in (2.6)) as k increases. In Table 1, over 10% efficiency is lost using the MMLE for the scale parameter relative to the MLE. Overall ARE's are over 82% for these distributions with $2 \leq k \leq 5$. From the numerical results, the ARE for the location parameter of extreme value distribution is the same as that for the scale parameter of exponential distribution. In fact, they are exactly same for any k, because they have the same $\epsilon(\theta)$ and $I(\theta)$ and because $E^2(b_i)$ of the extreme value distribution is the same as $E^2(b_i)$ of the exponential distribution.

In the following we consider general parameters. We are interested in the MMLE for the mean and variance of Gamma and Weibull distributions where the shape parameter is an unknown parameter. For the Gamma family with pdf $f(x;\theta,\alpha) = (x/\theta)^{\alpha-1} \exp(-x/\theta)/[\Gamma(\alpha)\theta]$, from Chen (2000), the mean and variance are $\mu = \alpha\theta$ and $\sigma^2 = \alpha\theta^2$, respectively, and its Fisher information matrix about (α,θ) is given by

$$\begin{bmatrix} \{\Gamma(\alpha)\Gamma''(\alpha) - [\Gamma'(\alpha)]^2\}[\Gamma(\alpha)]^{-2} & \theta^{-1} \\ \theta^{-1} & \alpha\theta^{-2} \end{bmatrix}.$$

For the Weibull distribution with cdf $F(x,\theta,\alpha) = 1-\exp(-(x/\theta)^{\alpha})/\theta$, from Chen (2000), the mean and variance are, respectively, $\mu = \theta\Gamma(1+1/\alpha)$ and $\sigma^2 = \theta^2[\Gamma(1+2/\alpha)-\Gamma^2(1+1/\alpha)]$, and the Fisher information of (α,θ) is given by

$$\begin{bmatrix} \alpha^{-2}+\tau_2 & \theta^{-1}(1-\alpha\tau_1-\tau_0) \\ \theta^{-1}(1-\alpha\tau_1-\tau_0) & \alpha\theta^{-2}[(\alpha+1)\tau_0-1] \end{bmatrix},$$

where $\tau_i = E(X^{\alpha}(\log X)^i)$, $i = 0,1,2$, and $X \sim F(x;\alpha,1)$. Chen (2000) also calculated the information gain matrix $\epsilon(\theta,\alpha)$ for these two distributions. Using (4.1) and (4.2), in Tables 2 and 3, we present the results of the ARE for the mean and variance of Gamma and Weibull distributions for various values of shape parameter.

Table 2. Relative efficiency for the Gamma distribution ($k = 2, \theta = 1$).

α	ARE($\hat{\mu}_{\mathrm{MMLE}}$, $\hat{\mu}_{\mathrm{MLE,SRS}}$)	ARE($\hat{\mu}_{\mathrm{MMLE}}$, $\hat{\mu}_{\mathrm{MLE,RSS}}$)	ARE($\hat{\sigma}^2_{\mathrm{MMLE}}$, $\hat{\sigma}^2_{\mathrm{MLE,SRS}}$)	ARE($\hat{\sigma}^2_{\mathrm{MMLE}}$, $\hat{\sigma}^2_{\mathrm{MLE,RSS}}$)
1.5	1.3702	.9654	1.0867	.8817
2	1.3913	.9727	1.0766	.8984
2.5	1.4049	.9750	1.0682	.8846
3	1.4144	.9782	1.0612	.8951
3.5	1.4213	.9799	1.0554	.8939
4	1.4267	.9812	1.0506	.8928
4.5	1.4309	.9822	1.0465	.8926
5	1.4343	.9831	1.0431	.8923
10	1.4502	.9874	1.0245	.8881

Table 3. Relative efficiency for the Weibull distribution ($k = 2, \theta = 1$).

α	ARE($\hat{\mu}_{\text{MMLE}}$, $\hat{\mu}_{\text{MLE,SRS}}$)	ARE($\hat{\mu}_{\text{MMLE}}$, $\hat{\mu}_{\text{MLE,RSS}}$)	ARE($\hat{\sigma}^2_{\text{MMLE}}$, $\hat{\sigma}^2_{\text{MLE,SRS}}$)	ARE($\hat{\sigma}^2_{\text{MMLE}}$, $\hat{\sigma}^2_{\text{MLE,RSS}}$)
1.5	1.3943	.9727	1.0660	.9135
2	1.4025	.9731	1.0598	.9208
2.5	1.3985	.9719	1.0499	.9269
3	1.3923	.9706	1.0408	.9326
3.5	1.3862	.9693	1.0334	.9379
4	1.3808	.9682	1.0275	.9425
4.5	1.3762	.9672	1.0230	.9468
5	1.3722	.9663	1.0195	.9505
10	1.3528	.9607	1.0061	.9720

5. Under imperfect judgment ranking

In the previous sections, we consider the MMLE and its properties under the perfect judgment ranking. In this section, we consider the imperfect judgment ranking. In practice, without the actual measurements, the judgment error is inevitable, especially, for the large set size k. We use the model of the imperfect ranking considered by Bohn and Wolfe (1994), Hettmansperger (1995), and Chen (2000). Suppose the probability that the unit in SRS of size k with actual rank r is judgment ranked as s is p_{rs}. Thus $\sum_{s=1}^{k} p_{rs} = 1$, for $r = 1, \dots, k$. Denote the RSS by $X_{[i]j}$, $i = 1, \dots, k$ and $j = 1, \dots, m$, where $[i]$ is used to indicate the judgment error in ranking r. Assume $p_{rs} = p_{sr}$. Then the pdf of $X_{[i]1}$, denoted by $f_{[i]:k}(x;\theta)$, is given by

$$f_{[i]:k}(x;\theta) = \sum_{t=1}^{k} p_{it} f_{t:k}(x;\theta)$$

where $f_{t:k}(x;\theta)$ is given by (1.1). Note that $f_{[i]:k}(x;\theta) = f_{i:k}(x;\theta)$ if $p_{ii} = 1$ and that $f_{[i]:k}(x;\theta) = f(x;\theta)$ if $p_{ij} = 1/k$, for $j = 1, \dots, k$. It can be shown that the analogies of (2.4) and (2.5) still hold under the imperfect judgment ranking, i.e.,

$$\sum_{i=1}^{k} f_{[i]:k}(x;\theta) = kf(x;\theta) \quad \text{and} \quad \sum_{i=1}^{k} E\frac{f'(X_{[i]i};\theta)}{f(X_{[i]1};\theta)} = 0. \tag{5.1}$$

However, the sum of the expectations of the second and third terms in (2.6) is no longer zero if the ranking is imperfect. In practice, we do not know whether the judgment ranking is perfect or not. Therefore, we will still solve the MMLE for θ, denote by $\hat{\theta}^*_{\text{MMLE}}$, from

$$\sum_{j=1}^{m}\sum_{i=1}^{k} \frac{f'(X_{[i]j};\theta)}{f(X_{[i]j};\theta)} = 0. \tag{5.2}$$

When the ranking is imperfect, Bai and Chen (2000) show that the MLE based on RSS for θ is still asymptotically normal, and that it is at least as efficient as the MLE based on SRS. Here, the properties of $\hat{\theta}^*_{\text{MMLE}}$ are summarized in:

THEOREM 5.1. *Under the above judgment ranking model and the conditions* (C1) *to* (C4), *for* $n = mk$, *the MMLE*, $\hat{\theta}^*_{\text{MMLE}}$, *solved from* (5.2) *has the following properties*:

(i) $\hat{\theta}^*_{\text{MMLE}}$ *is strongly consistent as* $m \to \infty$ *and* k *is fixed.*

(ii) *Let* θ_0 *be the true value of* θ. *As* $m \to \infty$ *and* k *is fixed,*

$$n^{1/2}(\hat{\theta}^*_{\text{MMLE}} - \theta_0) \xrightarrow{D} N(0, \Delta^*(\theta_0)),$$

where $\Delta^*(\theta_0) = I^{-1}(\theta_0) - \sum_{i=1}^{k} I^{-1}(\theta_0)E_{\theta_0}[b_i(\theta_0)]E_{\theta_0}[b_i(\theta_0)]^T I^{-1}(\theta_0)/k$, *and where*

$$b_i(\theta_0) = f'(X_{[i]1}; \theta_0)/f(X_{[i]1}; \theta_0).$$

Theorem 5.1 can be proved similarly as Theorem 3.1 using (5.1). Theorem 5.1 shows that, under the imperfect ranking, the MMLE is still at least as efficient as the MLE using SRS with the same size. When $p_{ii} = 1$, i.e., no judgment error, Theorem 5.1 reduces to Theorem 3.1. When $p_{ij} = 1/k$, i.e., the RSS is equivalent to the SRS, then $E_{\theta_0}[b_i(\theta_0)] = 0$ and Theorem 5.1 reduces to (2.2).

Given (5.1), the modified maximum likelihood equation (5.2) with the judgment error is still an unbiased estimating equation. However, the original maximum likelihood equation (2.6) may not be unbiased when the ranking is imperfect. In this sense, the MMLE should be more robust than the MLE when the ranking is imperfect. Thus we conduct simulations to compare MLE using RSS and MMLE when the ranking is imperfect. We use the simulation method considered by Dell and Clutter (1972) and David and Levine (1972). In the simulation, we choose $k = 2, 4$ and $m = 10$. So the total sample size is $n = 20, 40$. We consider two underlying distributions, the normal distribution $N(\theta, 1)$ with the location parameter θ and the exponential distribution $E(\theta)$ with the scale parameter θ. We choose two random error variables, $N(0, \sigma^2)$ and the Laplace distribution $L(0, \sigma^2)$ with zero mean and variance σ^2, where $\sigma^2 = 0.25, 0.50, 1.00, 3.00$ for each distribution. Given k, m, $F(x;\theta)$ and the random error distribution, in the first stage, we generate k sets of simple random samples, $X_{s,t}$, of size k each from $F(x;\theta)$, $s, t = 1, \ldots, k$, and also k sets of random error variables $e_{s,t}$, of size k each, where $X_{s,t}$ and $e_{s,t}$ are independent. Define $Y_{s,t} = X_{s,t} + e_{s,t}$. Then we can obtain two one cycle perfectly ranked RSS's $X_{(i)1}$ and $Y_{(i)1}$, $i = 1, \ldots, k$. We repeat this process for m times and obtain m cycles of RSS's: $X_{(i)j}$ and $Y_{(i)j}$. For the pairs $(Y_{(i)j}, X_{(i)j})$, $i = 1, \ldots, k$ and $j = 1, \ldots, m$, define the concomitants of $Y_{(i)j}$ as $X_{[i]j}$ for $i = 1, \ldots, k$ and $j = 1, \ldots, m$, i.e., for each j, if $Y_{(i)j} = Y_{i^*,j}$ then $X_{[i]j} = X_{i^*,j}$. Thus $X_{[i]j}$, $i = 1, \ldots, k$ and $j = 1, \ldots, m$, are the RSS with judgment error unless $e_{s,t} = 0$ for all $s, t = 1, \ldots, k$. For each replication, we find estimators $\hat{\theta}_{\text{MLE,RSS}}$ and $\hat{\theta}_{\text{MMLE}}$ based on $\{X_{(i)j}, i = 1, \ldots, k, j = 1, \ldots, m\}$ and $\hat{\theta}^*_{\text{MLE,RSS}}$ and $\hat{\theta}^*_{\text{MMLE}}$ based on $\{X_{[i]j}, i = 1, \ldots, k, j = 1, \ldots, m\}$. We repeat this process for 30,000 times. The estimate for the θ is the average of the estimates from these replications and the variance (Var) of the estimate is the sample variance. We compare the MSE of the estimates for θ by MLE and MMLE. The simulation results are reported in Tables 4 to 7. In Tables 4 and 5, we present the MSE and the estimates with their variances for the location parameter of the normal distribution, respectively. The results for the scale parameter of the exponential distribution are presented in Tables 6 and 7.

We choose the normal and Laplace as the random error distributions because the Laplace has heavier tail than the normal distribution. From Tables 4 and 6, not surprisingly, the MLE has smaller MSE than the MMLE when there is no judgment error.

Table 4. Compare the mean square error (MSE) of MLE and MMLE based on RSS for the location parameter of the normal distribution $N(\theta, 1)$ where $\theta = 0$.

		Random error from			
		$N(0,\sigma^2)$		$L(0,\sigma^2)$	
		MSE			
k	σ^2	MLE*	MMLE*	MLE*	MMLE*
2	.25	.03732 [.03384]	.03733 [.03406]	.03898 [.03381]	.03890 [.03405]
	.50	.03944 [.03365]	.03931 [.03388]	.04117 [.03340]	.04098 [.03371]
	1.0	.04202 [.03368]	.04178 [.03393]	.04382 [.03364]	.04347 [.03390]
	3.0	.04598 [.03338]	.04555 [.03361]	.04735 [.03369]	.04704 [.03402]
4	.25	.01358 [.01036]	.01372 [.01079]	.01519 [.01032]	.01520 [.01077]
	.50	.01550 [.01023]	.01547 [.01061]	.01763 [.01020]	.01739 [.01061]
	1.0	.01782 [.01019]	.01769 [.01063]	.01992 [.01024]	.01957 [.01067]
	3.0	.02168 [.01035]	.02134 [.01077]	.02280 [.01024]	.02239 [.01062]

* Here the first number is the MSE for the imperfect ranking and the second number in bracket is the corresponding MSE for the perfect ranking.

Table 5. Compare the estimates and their variances of MLE and MMLE based on RSS for the location parameter of the normal distribution $N(\theta, 1)$ where $\theta = 0$.

		Random error from							
		$N(0,\sigma^2)$				$L(0,\sigma^2)$			
		$\hat{\theta}$		Var of $\hat{\theta}$		$\hat{\theta}$		Var of $\hat{\theta}$	
k	σ^2	MLE	MMLE	MLE	MMLE	MLE	MMLE	MLE	MMLE
2	.25	−.00031	−.00045	.03732	.03733	.00253	.00265	.03897	.03889
	.50	.00165	.00163	.03943	.03931	.00061	.00067	.04117	.04098
	1.0	−.00036	−.00032	.04202	.04178	.00111	.00115	.04382	.04346
	3.0	−.00142	−.00134	.04598	.04554	−.00013	−.00030	.04735	.04704
4	.25	−.00027	−.00014	.01357	.01372	.00051	.00049	.01519	.01520
	.50	−.00028	−.00031	.01550	.01547	.00004	.00007	.01763	.01739
	1.0	.00061	.00057	.01782	.01769	.00096	.00094	.01992	.01957
	3.0	−.00062	−.00069	.02168	.02134	.00016	.00025	.02280	.02239

However, for some moderate judgment error ($\sigma^2 \geq 0.5$), the MMLE has smaller MSE than the MLE. For the normal case, the larger MSE for the MMLE is due to the larger variance of the MLE under the imperfect judgment ranking (Table 5), since the MLE and MMLE for the location parameter of the normal are very close to each other and seem to be unbiased. But for the scale parameter of the exponential distribution, in Table 7, the MLE is biased under the imperfect judgment ranking while the MMLE is still unbiased. In addition, the MMLE has a smaller variance than the MLE for a moderate judgment error. From Tables 4 to 7, we can also see that the heavier tailed distribution has more effect on the efficiencies of the MLE and MMLE. From Table 6, for $\sigma^2 = 3.00$, we also see that the MSE for the MLE is greater than .05 when $k = 2$ ($n = 20$) and greater than .025 when $k = 4$ ($n = 40$). This implies that when the judgment error is

Table 6. Compare the mean square error (MSE) of MLE and MMLE based on RSS for the scale parameter of the exponential distribution $E(\theta)$ where $\theta = 1$.

		Random error from			
		$N(0,\sigma^2)$		$L(0,\sigma^2)$	
		MSE			
k	σ^2	MLE*	MMLE*	MLE*	MMLE*
2	0.25	.04024 [.03643]	.04118 [.03814]	.04163 [.03572]	.04166 [.03747]
	0.50	.04219 [.03572]	.04240 [.03757]	.04484 [.03575]	.04344 [.03746]
	1.00	.04572 [.03587]	.04389 [.03763]	.04964 [.03591]	.04572 [.03783]
	3.00	.05376 [.03542]	.04670 [.03715]	.05715 [.03545]	.04704 [.03728]
4	0.25	.01594 [.01140]	.01570 [.01311]	.01850 [.01129]	.01664 [.01289]
	0.50	.01939 [.01137]	.01727 [.01312]	.02292 [.01139]	.01855 [.01317]
	1.00	.02492 [.01146]	.01902 [.01309]	.02996 [.01132]	.02060 [.01293]
	3.00	.03740 [.01112]	.02165 [.01284]	.04312 [.01133]	.02275 [.01309]

* Here the first number is the MSE for the imperfect ranking and the second number in bracket is the corresponding MSE for the perfect ranking.

Table 7. Compare the estimates and their variances of MLE and MMLE based on RSS for the scale parameter of the exponential distribution $E(\theta)$ where $\theta = 1$.

		Random error from							
		$N(0,\sigma^2)$				$L(0,\sigma^2)$			
		$\hat{\theta}$		Var of $\hat{\theta}$		$\hat{\theta}$		Var of $\hat{\theta}$	
k	σ^2	MLE	MMLE	MLE	MMLE	MLE	MMLE	MLE	MMLE
2	0.25	1.01695	1.00068	.03995	.04118	1.02403	1.00127	.04105	.04166
	0.50	1.02265	.99818	.04168	.04240	1.03112	.99788	.04387	.04344
	1.00	1.03704	1.00124	.04434	.04389	1.04577	1.00083	.04755	.04573
	3.00	1.05919	1.00081	.05025	.04670	1.06730	1.00068	.05262	.04750
4	0.25	1.02793	.99964	.01516	.01570	1.04052	.99968	.01686	.01664
	0.50	1.04500	1.00056	.01737	.01727	1.05924	.99894	.01941	.01854
	1.00	1.06724	1.00087	.02040	.01902	1.08242	.99848	.02317	.02060
	3.00	1.10889	1.00057	.02554	.02165	1.12400	1.00055	.02774	.02275

present, the MLE based on RSS may be less efficient than that based on SRS, since .05 and .025 are the MSE's for the MLE using SRS of size $n = 20$ and $n = 40$, respectively. However, the MMLE based on RSS ha s smaller MSE than the MLE using SRS of same size, regardless of judgment errors.

6. Concluding remarks

In this paper, we consider the MMLE based on RSS for general parameters, and examine the large sample properties. Compared to the MLE using RSS, the MMLE has high efficiency (ARE is 1 for the location parameter of the logistic distribution) for the location parameter with small set size when the judgment ranking is perfect and is relatively easy to obtain, especially when the MLE using SRS has a closed expression or is already programmed. When the judgment ranking is imperfect, the simulations show that the MMLE is more robust than the MLE. In addition, the MMLE based on RSS is also more efficient than (at least as efficient as) the MLE for the location (scale) parameter based on SRS with the same sample size. Based on Theorem 3.1, we can also construct modified $100(1-\alpha)\%$ confidence intervals for the location and scale parameters (cf. Chen (2000)). When the judgment ranking is imperfect, this modified confidence interval is shorter than that of Chen (2000) based on MLE.

Acknowledgements

The authors are grateful to two referees for their extremely thorough and insightful suggestions that greatly improved the article. We also thank Professors Zehua Chen and Lucio Barabesi for helpful discussions and comments.

References

Al-Saleh, M. F. and Al-Kadiri, M. (2000). Double ranked set sampling, *Statist. Probab. Lett.*, **48**, 205–212.

Al-Saleh, M. F. and Zheng, G. (2002). Estimation of bivariate characteristics using ranked set sampling, *Australian & New Zealand Journal of Statistics*, **44**, 221–232.

Bai, Z. and Chen, Z. (2000). On the theory of ranked-set sampling and its ramifications, *J. Statist. Plann. Inference* (to appear).

Barnett, V. (1999). Ranked set sample design for environmental investigations, *Environmental and Ecological Statistics*, **6**, 59–74.

Bhattacharyya, G. K. (1985). The asymptotics of maximum likelihood and related estimators based on type II censored data, *J. Amer. Statist. Assoc.*, **80**, 398–404.

Bohn, L. L. and Wolfe, D. A. (1994). The effect of imperfect judgment rankings on properties of procedures based on the ranked set samples analog of the Mann-Whitey-Wilcoxon statistic, *J. Amer. Statist. Assoc.*, **89**, 168–176.

Chen, Z. (2000). The efficiency of ranked-set sampling relative to simple random sampling under multi-parameter families, *Statistica Sinica*, **10**, 247–263.

David, H. A. and Levine, D. N. (1972). Ranked set sampling in the presence of Judgment error, *Biometrics*, **28**, 553–555.

Dell, T. R. and Clutter, J. L. (1972). Ranked-set sampling theory with order statistics background, *Biometrics*, **28**, 545–555.

El-Neweihi, E. and Sinha, B. K. (2000). Reliability estimation based on ranked set sampling, *Comm. Statist. Theory Method*, **29**, 1583–1595.

Godambe, V. P. (1960). An optimum property of regular maximum likelihood estimation, *Ann. Math. Statist.*, **31**, 1208–1212.

Godambe, V. P. (1991). *Estimating Functions* (ed. V. P. Godambe), Clarendon Press, Oxford.

Hettmansperger, T. P. (1995). The ranked-set sampling sign test, *J. Nonparametr. Statist.*, **4**, 263–270.

Kvam, P. H. and Samaniego, F. J. (1993). On maximum likelihood estimation based on ranked set sampling, with applications to reliability, *Advances in Reliability* (ed. A. Basu), 215–229, North Holland, Amsterdam.

Kvam, P. H. and Samaniego, F. J. (1994). Nonparametric maximum likelihood estimation based on ranked set samples, *J. Amer. Statist. Assoc.*, **89**, 526–537.

Lehmann, E. L. (1983). *Theory of Point Estimation*, Wiley, New York.

McIntyre, G. A. (1952). A method for unbiased selective sampling using ranked sets, *Australian Journal of Agricultural Research*, **3**, 385–390.

Mehrotra, K. G. and Nanda, P. (1974). Unbiased estimation of parameters by order statistics in the case of censored samples, *Biometrika*, **61**, 601–606.

Patil, G. P., Sinha, A. K. and Taillie, C. (1999). Ranked set sampling: A bibliography, *Environmental and Ecological Statistics*, **6**, 91–98.

Sinha, B. K., Sinha, R. K. and Purkayastha, S. (1996). On some aspects of ranked set sampling for estimation of normal and exponential parameters, *Statist. Decisions*, **14**, 223–240.

Stokes, S. L. (1980). Estimation of variance using judgment ordered ranked set samples, *Biometrics*, **26**, 35–42.

Stokes, S. L. (1995). Parametric ranked set sampling, *Ann. Inst. Statist. Math.*, **47**, 465–482.

Stokes, S. L. and Sager, T. W. (1988). Characterizations of a ranked-set sample with application to estimating distribution functions, *J. Amer. Statist. Assoc.*, **83**, 374–381.

Takahasi, K. and Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering, *Ann. Inst. Statist. Math.*, **20**, 1–31.

Yu, P. L. H. and Lam, K. (1997). Regression estimator in ranked set sampling, *Biometrics*, **53**, 1070–1080.

Zheng, G. (2000). Some remarks on Fisher information in a ranked set sample (submitted).