Adv. & Appl. in Stat. **6**(3) (2006), 261-276
This paper is available online at http://www.pphmj.com

DOUBLE PERCENTILE RANKED SET SAMPLES FOR ESTIMATING THE POPULATION MEAN

ABDUL AZIZ JEMAIN and AMER IBRAHIM Al-OMARI

School of Mathematical Sciences
Faculty of Science and Technology
43600 UKM Bangi, Selangor, Malaysia
e-mail: azizj@pkrisc.cc.ukm.my
alomari_amer@yahoo.com

Abstract

Double percentile ranked set sampling procedure (DPRSS) is introduced as a modification of ranked set sampling (RSS) for estimating the mean of the population of interest. The efficiency of the DPRSS estimator of the mean relative to the simple random sampling (SRS), ranked set sampling (RSS), median ranked set sampling (MRSS), extreme ranked set sampling (ERSS) and percentile ranked set sampling (PRSS) methods is obtained. It turns out that DPRSS is an unbiased and more efficient than its counterpart SRS, RSS, MRSS, ERSS, and PRSS for the symmetric distributions. For non-symmetric distributions considered in this study, the DPRSS estimator has a smaller bias and it is more efficient than the SRS, ERSS and PRSS methods.

1. Introduction

McIntyre [4] was first introduced ranked set sampling procedure for estimating the mean of pasture yields. In situations where the experimental or sampling units in a study can be more easily ranked than quantified, McIntyre proposed that the mean of m sample units

2000 Mathematics Subject Classification: 62D05.

Keywords and phrases: ranked set sampling, median ranked set sampling, extreme ranked set sampling, percentile ranked set sampling, double percentile ranked set sampling.

Received March 11, 2006

based on a RSS as an estimator of the population mean. Takahasi and Wakimoto [8] provided the mathematical properties of RSS. They showed that the sample mean of the RSS is an unbiased estimator of the population mean with smaller variance than the sample mean of a simple random sample with the same size. Dell and Clutter [3] showed that the RSS estimator is an unbiased for the population mean regardless of error in ranking. Samawi et al. [7] investigated using extreme ranked set sampling for estimating a population mean, and showed that for symmetric distributions the ERSS estimator is an unbiased and has a smaller variance than the SRS estimator. Muttlak [5] suggested using median ranked set sampling (MRSS) to estimate the population mean. Al-Saleh and Al-Kadiri [1] introduced double ranked set sampling for estimating the population mean. Al-Saleh and Al-Omari [2] introduced multistage ranked set sampling, that increase the relative efficiency for estimating the population mean for fixed sample size. Muttlak [6] proposed percentile ranked set sampling procedure for estimating the population mean, he showed that by using PRSS we can improve the relative efficiency and reduce the errors in ranking.

In this article, modification for RSS, namely, double percentile ranked set samples. It is shown that DPRSS has smaller variance than its counterpart SRS, RSS, MRSS, ERSS and PRSS for estimating the population mean for symmetric distributions. For non-symmetric distributions considered in this study, the DPRSS estimator has smaller bias with variance smaller than that of the SRS estimator. An illustration of the DPRSS using real data is considered.

2. Sampling Methods

2.1. Ranked set sampling

The ranked set sampling procedure consists of choosing m random samples, each of size m, from a population of interest. The m units, in each of the m samples, are ranked by a judgment without any actual measurement. The smallest ranked unit from the first sample is measured, then the second smallest ranked unit from the second sample is measured and so on. This process continues until the largest ranked unit from the last sample is measured. Note that m^2 units are identified from a population of interest, but only m units are measured. These m independent but not identically distributed measured units constitute the ranked set sample.

2.2. Double ranked set sampling

DRSS can be described as follows:

(1) Identify m^3 elements from the target population and divide these elements randomly into m sets each of size m^2 elements.

(2) Use the usual RSS procedure on each set to obtain m ranked set samples of size m each.

(3) Apply the RSS procedure again on Step 2 to obtain a DRSS of size m.

2.3. Extreme ranked set sampling

The ERSS method can be summarized as follows: Randomly select m samples each of size m units from the target population and rank the units within each set with respect to a variable of interest by visual inspection. If the sample size m is even, then select the smallest ranked unit from the first $m/2$ sets and the largest ranked unit from the other $m/2$ sets for measurement. If the sample size is odd, then select the smallest unit from the first $(m-1)/2$ samples and the largest ranked unit from the second $(m-1)/2$ samples and the median from the remaining set. The cycle may be repeated n times, to obtain an ERSS nm units.

2.4. Median ranked set sampling

In the MRSS method, randomly select m samples each of size m units from the population then ranked the units within each sample with respect to a variable of interest by visual inspection. If the sample size is odd, then select the $((m+1)/2)$th smallest rank (the median of the sample) for measurement. For even sample size, select the $(m/2)$th smallest rank from the first $m/2$ samples and the $((m+2)/2)$th smallest rank from the second $m/2$ samples. The cycle may be repeated n times, to obtain a MRSS of size nm.

2.5. Percentile ranked set sampling

The PRSS method consists of drawing m random samples each of size m from the population and rank the units within each sample with respect to a variable of interest. If the sample size is even, then select the $(p(m+1))$th smallest rank from the first $m/2$ samples and the $(q(m+1))$th smallest rank from the other $m/2$ samples for measurement, note that $0 \leq p \leq 1$ and $p + q = 1$. If the sample size is odd, then select the $(p(m+1))$th smallest rank from the first $(m-1)/2$ samples and the $(q(m+1))$th smallest rank from the second $(m-1)/2$ samples, and from one sample the median for that sample for actual measurement. The cycle may be repeated n times, to obtain a PRSS of size nm.

3. Double Percentile Ranked Set Sampling

The DPRSS procedure can be described as follows:

Step 1. Select m^3 units from the population and allocate them into m^2 samples each of size m.

Step 2. If the sample size is even, then select the $(p(m+1))$th smallest rank from the first $m^2/2$ samples and the $(q(m+1))$th smallest rank from the second $m^2/2$ samples for measurement. If the sample size is odd, then select the $(p(m+1))$th smallest rank from the first $m(m-1)/2$ samples and the median from the next m samples and the $(q(m+1))$th smallest rank from the other $m(m-1)/2$ samples. This step yields m samples each of size m.

Step 3. Apply the PRSS procedure on the m samples obtained in Step 2 to get a DPRSS sample of size m.

The whole cycle may be repeated n times to obtain a sample of size mn units from DPRSS. Note that the nearest integer of the $(p(m+1))$th and $(q(m+1))$th is considered, where $q = 1 - p$ and $0 \leq p \leq 1$.

4. Estimation of the Population Mean

Let $X_1, X_2, \ldots, X_m$ be a random sample with probability density function $f(x)$ with finite mean and variance σ^2. Let $X_{11}, X_{12}, \ldots, X_{1m}$; $X_{21}, X_{22}, \ldots, X_{2m}; \ldots; X_{m1}, X_{m2}, \ldots, X_{mm}$; be independent random variables all with the same cumulative distribution function $F(x)$. Let $X_{i(1)}, X_{i(2)}, \ldots, X_{i(m)}$ $(i = 1, 2, \ldots, m)$ be the ordered statistics of the ith sample, $X_{i1}, X_{i2}, \ldots, X_{im}$, $(i = 1, 2, \ldots, m)$. Let $Y_1, Y_2, \ldots, Y_m$ be RSS. Then $Y_i \overset{d}{=} X_{(i)}$. The *estimator of the population mean* μ using RSS is defined as $\hat{\mu}_{RSS} = \frac{1}{m}\sum_{i=1}^{m} Y_i$ and variance given by $\mathrm{Var}(\hat{\mu}_{RSS}) = \frac{\sigma^2}{m} - \frac{1}{m^2}\sum_{i=1}^{m}(\mu_{(i)} - \mu)^2$.

The *estimator of the population mean* μ using SRS is defined as $\hat{\mu}_{SRS} = \frac{1}{m}\sum_{i=1}^{m} X_i$, with variance σ^2/m.

For the kth cycle, $(k = 1, 2, \ldots, n)$, if the sample size is even, let $Y^*_{i(p(m+1))k}$ be the $(p(m+1))$th smallest rank of the ith sample $(i = 1, 2, \ldots, l;\ l = m/2)$, and let $Y^*_{i(q(m+1))k}$ be the $(q(m+1))$th smallest rank of the ith sample $(i = l+1, \ldots, m)$. The only quantified sample $Y^*_{1(p(m+1))k}, Y^*_{2(p(m+1))k}, \ldots, Y^*_{\frac{m}{2}(p(m+1))k}, Y^*_{\frac{m}{2}+1(q(m+1))k}, \ldots, Y^*_{m(q(m+1))k}$, will denote the DPRSSE.

For odd sample size, let $Y^*_{i(p(m+1))k}$ be the $(p(m+1))$th smallest rank of the ith sample, $(i = 1, 2, \ldots, j;\ j = (m-1)/2)$, and let $Y^*_{i((m+1)/2)k}$ be the median of the ith sample of the rank $i = (m+1)/2$ and let $Y^*_{i(q(m+1))k}$ be the $(q(m+1))$th smallest rank of the ith sample $(i = j+2, \ldots, m)$. Then the only quantified sample $Y^*_{1(p(m+1))k}, Y^*_{2(p(m+1))k}, \ldots, Y^*_{\frac{m-1}{2}(p(m+1))k}$, $Y^*_{\frac{m-1}{2}+1\left(\frac{m+1}{2}\right)k}, Y^*_{\frac{m-1}{2}+2(q(m+1))k}, \ldots, Y^*_{m(q(m+1))k}$ will denote the DPRSSO.

The *estimator of the population mean* using DPRSS can be defined as

$$\hat{\mu}_{DPRSS} = \begin{cases} \hat{\mu}_{DPRSSE} = \dfrac{1}{mn}\displaystyle\sum_{k=1}^{n}\left(\sum_{i=1}^{l} Y^{*}_{i(p(m+1))k} + \sum_{i=l+1}^{m} Y^{*}_{i(q(m+1))k}\right), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad l = m/2 \\ \hat{\mu}_{DPRSSO} = \dfrac{1}{mn}\displaystyle\sum_{k=1}^{n}\left(\sum_{i=1}^{j} Y^{*}_{i(p(m+1))k} + Y^{*}_{(j+1)\left(\frac{m+1}{2}\right)k} + \sum_{i=j+2}^{m} Y^{*}_{i(q(m+1))k}\right), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad j = (m-1)/2. \end{cases}$$

Assume that Y_i^{*} has the mean μ_i^{*} and the variance, $\sigma^{*2}_{(i)}$ $(i = 1, 2, \ldots, m)$, the *variance of* $\hat{\mu}_{DPRSS}$ in the case of an even and odd sample size can be defined respectively as

$$\sigma^2_{DPRSSE} = \frac{1}{nm^2}\sum_{k=1}^{n}\left(\sum_{i=1}^{l}\sigma^{*2}_{i(p(m+1))k} + \sum_{i=l+1}^{m}\sigma^{*2}_{i(q(m+1))k}\right), \; l = m/2.$$

$$\sigma^2_{DPRSSO} = \frac{1}{nm^2}\sum_{k=1}^{n}\left(\sum_{i=1}^{j}\sigma^{*2}_{i(p(m+1))k} + \sigma^{*2}_{(j+1)\left(\frac{m+1}{2}\right)k} + \sum_{i=j+2}^{m}\sigma^{*2}_{i(q(m+1))k}\right),$$

$$j = (m-1)/2.$$

Al-Saleh and Al-Kadiri [1] showed that $\mu = \sum_{i=1}^{m}\mu^{*}_{(i)}$ and

$$\sigma^2 = \frac{1}{m}\left[\sum_{i=1}^{m}\sigma^{*2}_{(i)} + \sum_{i=1}^{m}(\mu^{*}_{(i)} - \mu)^2\right],$$

where μ and σ^2 are the mean and the variance of the population, respectively.

Lemmas

1. $\hat{\mu}_{DPRSS}$ *is unbiased estimator of the population mean if the underling distribution is symmetric about the population mean* μ.

2. $\mathrm{Var}(\hat{\mu}_{DPRSS})$ *is less than each of* $\mathrm{Var}(\hat{\mu}_{SRS})$ *and* $\mathrm{Var}(\hat{\mu}_{PRSS})$.

3. *The mean square error of* $\hat{\mu}_{DPRSS}$ *is less than the variance of SRS estimator for asymmetric distributions about* μ, *i.e.*, $\text{MSE}(\hat{\mu}_{DPRSS}) < \text{Var}(\hat{\mu}_{SRS})$.

The above properties can be proved based on Takahasi and Wakimoto [8], Muttlak [6] and Al-Saleh and Al-Kadiri [1].

To compare the DPRSS estimators for the population mean with SRS, RSS, MRSS, ERSS and PRSS procedures, three symmetric distributions, namely uniform, normal and logistic and four non-symmetric distributions, namely lognormal, exponential, gamma and weibull are considered for comparison. If the parent distribution is symmetric about μ, then the *relative efficiency of the* RSS, MRSS, ERSS, PRSS *and* DPRSS with respect to SRS can be defined as

$$eff(\hat{\mu}_{SRS}, \hat{\mu}_{RSS}) = \frac{\text{Var}(\hat{\mu}_{SRS})}{\text{Var}(\hat{\mu}_{RSS})},$$

and if the distribution is not symmetric the *relative efficiency* is defined as follows:

$$eff(\hat{\mu}_{SRS}, \hat{\mu}_{RSS}) = \frac{\text{Var}(\hat{\mu}_{SRS})}{\text{MSE}(\hat{\mu}_{RSS})}.$$

Assume the cycle is repeated once, in Tables 1 and 2, we compute the relative efficiency of the estimators of RSS, MRSS, ERSS and DPRSS using $P = 20\%$, 30% and 40% with sample sizes $m = 7$ and 8, respectively. In Tables 3 and 4, we use $P = 25\%$, 35% and 45% with sample sizes $m = 10$ and 11, respectively. Finally, in Table 5, we compute the relative efficiency of RSS, MRSS, ERSS, PRSS and DPRSS using $P = 20\%$, 30% and 40%, with sample size, $m = 12$. We compared the average of the 70,000 sample estimates, namely,

$$\bar{\hat{\mu}}_{RSS} = \frac{1}{70{,}000}\sum_{i=1}^{70{,}000}\hat{\mu}_{RSS,\,i} \text{ and } \bar{\hat{\mu}}_{SRS} = \frac{1}{70{,}000}\sum_{i=1}^{70{,}000}\hat{\mu}_{SRS,\,i}.$$

Table 1. The relative efficiency of RSS, MRSS, ERSS and DPRSS with respect to SRS for estimating the population mean with sample size $m = 7$

Distribution		RSS	MRSS	ERSS	DPRSS		
					P20%	P30%	P40%
Uniform (0, 1)	*eff*	4.000	2.986	5.711	23.263	23.263	14.159
	Bias						
Normal (0, 1)	*eff*	3.654	4.741	2.762	14.710	14.710	20.812
	Bias						
Logistic (–1, 1)	*eff*	3.258	5.787	1.997	13.590	13.590	24.468
	Bias						
Lognormal (0, 1)	*eff*	1.819	2.021	0.465	12.182	12.182	2.516
	Bias		0.538	0.696	0.027	0.027	0.505
Exponential (1)	*eff*	2.700	1.791	0.812	8.515	8.515	2.645
	Bias		0.241	0.280	0.060	0.060	0.220
Exponential (2)	*eff*	2.711	1.813	0.810	8.542	8.542	2.663
	Bias		0.120	0.141	0.030	0.030	0.110
Gamma (1, 2)	*eff*	2.718	1.808	0.812	8.639	8.639	2.667
	Bias		0.480	0.563	0.118	0.118	0.439
Gamma (1, 3)	*eff*	2.693	1.790	0.817	8.541	8.541	2.649
	Bias		0.722	0.838	0.178	0.178	0.659
Weibull (1, 4)	*eff*	2.739	1.800	0.812	8.484	8.484	2.659
	Bias		0.963	1.126	0.238	0.238	0.879
Weibull (1, 3)	*eff*	2.717	1.791	0.814	8.440	8.440	2.639
	Bias		0.721	0.838	0.179	0.179	0.659

Table 2. The relative efficiency of RSS, MRSS, ERSS and DPRSS procedures for estimating the population mean with sample size $m = 8$

Distribution		RSS	MRSS	ERSS	DPRSS		
					P20%	P30%	P40%
Uniform (0, 1)	*eff*	4.500	3.357	8.403	41.570	20.873	16.003
	Bias						
Normal (0, 1)	*eff*	4.048	5.343	2.675	15.404	23.893	28.440
	Bias						
Logistic (–1, 1)	*eff*	3.604	6.529	1.841	13.136	27.163	35.844
	Bias						
Lognormal (0, 1)	*eff*	1.963	1.804	0.280	3.382	3.684	1.563
	Bias		0.539	1.033	0.332	0.383	0.607
Exponential (1)	*eff*	2.935	1.665	0.455	2.165	4.701	1.517
	Bias		0.240	0.422	0.211	0.147	0.281
Exponential (2)	*eff*	2.950	1.675	0.456	2.160	4.674	1.534
	Bias		0.120	0.211	0.106	0.074	0.140
Gamma (1, 2)	*eff*	2.945	1.656	0.456	2.157	4.673	1.523
	Bias		0.483	0.841	0.423	0.295	0.561
Gamma (1, 3)	*eff*	2.947	1.659	0.387	2.169	4.661	1.517
	Bias		0.721	1.426	0.632	0.443	0.842
Weibull (1, 4)	*eff*	2.949	1.687	0.462	2.209	4.725	1.540
	Bias		0.960	1.684	0.840	0.590	1.122
Weibull (1, 3)	*eff*	2.973	1.684	0.391	2.213	4.714	1.608
	Bias		0.721	1.426	0.630	0.444	0.823

Table 3. The relative efficiency of RSS, MRSS, ERSS and DPRSS procedures for estimating the population mean with sample size $m = 10$

Distribution		RSS	MRSS	ERSS	DPRSS		
					P25%	P35%	P45%
Uniform (0, 1)	*eff*	5.500	4.023	12.282	39.219	27.829	23.865
	Bias						
Normal (0, 1)	*eff*	4.859	6.606	2.938	30.809	39.541	43.454
	Bias						
Logistic (–1, 1)	*eff*	4.258	8.247	1.899	31.238	45.997	54.508
	Bias						
Lognormal (0, 1)	*eff*	2.091	1.397	0.186	12.281	1.896	1.214
	Bias		0.562	1.281	0.164	0.495	0.622
Exponential (1)	*eff*	3.453	1.335	0.224	22.465	2.137	1.173
	Bias		0.254	0.516	0.027	0.212	0.290
Exponential (2)	*eff*	3.486	1.344	0.269	22.120	2.141	1.175
	Bias		0.127	0.257	0.014	0.106	0.145
Gamma (1, 2)	*eff*	3.452	1.329	0.293	22.277	2.119	1.164
	Bias		0.508	1.029	0.055	0.425	0.581
Gamma (1, 3)	*eff*	3.506	1.343	0.297	22.303	2.118	1.163
	Bias		0.762	1.540	0.082	0.637	0.871
Weibull (1, 4)	*eff*	3.390	1.308	0.290	2.624	2.121	1.167
	Bias		1.019	2.057	0.108	0.850	1.161
Weibull (1, 3)	*eff*	3.442	1.332	0.293	22.022	2.103	1.159
	Bias		0.762	1.547	0.081	0.638	0.871

Table 4. The relative efficiency of RSS, MRSS, ERSS and DPRSS procedures for estimating the population mean with sample size $m = 11$

Distribution		RSS	MRSS	ERSS	DPRSS		
					P25%	P35%	P45%
Uniform (0, 1)	*eff*	6.000	4.380	12.063	48.545	35.343	29.709
	Bias						
Normal (0, 1)	*eff*	5.263	7.290	3.234	35.437	45.018	51.566
	Bias						
Logistic (–1, 1)	*eff*	4.632	9.163	2.078	34.506	51.847	63.744
	Bias						
Lognormal (0, 1)	*eff*	2.085	1.195	0.186	18.777	2.099	1.196
	Bias		0.580	1.219	0.110	0.446	0.595
Exponential (1)	*eff*	3.631	1.162	0.306	28.452	2.589	1.204
	Bias		0.263	0.481	0.001	0.183	0.273
Exponential (2)	*eff*	3.661	1.157	0.308	28.706	2.568	1.198
	Bias		0.132	0.240	0.000	0.092	0.137
Gamma (1, 2)	*eff*	3.731	1.322	0.311	28.892	2.566	1.195
	Bias		0.526	0.960	0.002	0.367	0.547
Gamma (1, 3)	*eff*	3.681	1.161	0.307	28.439	2.554	1.193
	Bias		0.791	1.443	0.002	0.551	0.820
Weibull (1, 4)	*eff*	3.686	1.167	0.310	28.644	2.552	1.255
	Bias		1.053	1.918	0.002	0.733	1.063
Weibull (1, 3)	*eff*	3.715	1.167	0.308	28.621	2.568	1.195
	Bias		0.791	1.446	0.002	0.550	0.820

Table 5. The relative efficiency for estimating the population mean using RSS, MRSS, ERSS, PRSS and DPRSS with sample size $m = 12$

Distribution		RSS	MRSS	ERSS	PRSS			DPRSS		
					P20%	P30%	P40%	P20%	P30%	P40%
Uniform	*eff*	6.500	4.710	16.895	6.800	5.581	4.993	67.621	45.603	36.014
(0, 1)	*Bias*									
Normal	*eff*	5.621	7.903	3.130	6.423	7.184	7.759	36.774	49.740	47.580
(0, 1)	*Bias*									
Logistic	*eff*	4.897	9.748	1.926	6.684	8.350	8.248	34.205	54.543	54.223
(–1, 1)	*Bias*									
Lognormal	*eff*	2.200	1.118	0.126	4.389	1.958	1.316	28.530	3.000	1.868
(0, 1)	*Bias*		0.579	1.506	0.209	0.418	0.529	0.048	0.355	0.453
Exponential	*eff*	3.924	1.092	0.202	4.717	2.092	1.320	8.396	4.504	10.673
(1)	*Bias*		0.263	0.592	0.061	0.173	0.235	0.083	0.130	0.187
Exponential	*eff*	3.885	1.073	0.199	4.646	2.090	1.316	8.227	4.436	2.867
(2)	*Bias*		0.132	0.296	0.031	0.086	0.117	0.042	0.065	0.093
Gamma	*eff*	3.930	1.084	0.200	4.697	2.109	1.313	8.435	4.472	2.286
(1, 2)	*Bias*		0.528	1.187	0.123	0.344	0.471	0.165	0.261	0.374
Gamma	*eff*	3.902	1.087	0.201	4.703	1.679	1.314	8.358	4.479	2.283
(1, 3)	*Bias*		0.790	1.777	0.184	0.516	0.706	0.249	0.391	0.561
Weibull	*eff*	3.810	1.074	0.198	4.631	2.074	2.083	8.290	4.434	2.258
(1, 4)	*Bias*		1.054	2.371	0.245	0.691	0.689	0.332	0.521	0.748
Weibull	*eff*	3.810	1.070	0.196	4.619	2.060	2.050	8.250	4.376	2.241
(1, 3)	*Bias*		0.789	1.782	0.185	0.517	0.519	0.248	0.392	0.561

5. Concluding Remarks

From simulation results, we conclude the following:

1. The estimator of the population mean obtained by using DPRSS procedure is more efficient than that obtained using the usual SRS.

2. A gain in efficiency is attained using DPRSS for estimating the population mean. As an example for normal distribution with mean 0 and variance 1, with $m = 12$ using $P = 20\%$, 30% and 40% the relative efficiency of DPRSS estimator is 36.774, 49.740 and 47.580, respectively,

and by using PRSS the relative efficiency is 6.423, 7.184 and 7.759, respectively, and the relative efficiency using RSS, MRSS and ERSS comes out to be 5.621, 7.903 and 3.130, respectively.

3. If the underlying distribution is non-symmetric, then a gain in efficiency is attained using DPRSS. For example, if the underlying distribution is weibull with parameters 1 and 4, then the relative efficiency for estimating the population mean using RSS, MRSS and ERSS comes out to be 3.686, 1.167 and 0.310, respectively, and by using DPRSS with $P = 25\%,\ 35\%$ and 45% comes out to be 28.644, 2.552 and 1.255 with values of bias 0.002, 0.733 and 1.063, respectively.

4. The gain in efficiency for mean estimation using DPRSS is greater for symmetric distributions than for the asymmetric distributions.

6. Double Percentile Ranked Set Sampling with Errors in Ranking

Dell and Clutter [3] showed that the sample mean using RSS is unbiased estimator of the population mean regardless of whatever ranking is perfect or not, and has a smaller variance than its counterpart SRS with the same sample size.

Muttlak [6] showed that PRSS with errors in ranking is unbiased estimator of the population mean when the underlying distribution is assumed symmetric about its mean.

Let $Y^*_{i[p(m+1)]k}$ and $Y^*_{i[q(m+1)]k}$, $(k = 1, 2, \ldots, n)$ be the $(p(m+1))$th and $(q(m+1))$th judgment order statistics respectively of the ith sample $(i = 1, 2, \ldots, m)$ with errors in ranking. The *estimator of the population mean* with error in ranking using DPRSS can be defined as

$$\hat{\mu}_{DPRSS_e} = \begin{cases} \hat{\mu}_{DPRSSE_e} = \dfrac{1}{mn}\displaystyle\sum_{k=1}^{n}\left(\sum_{i=1}^{l} Y^*_{i[q_1(m+1)]k} + \sum_{i=l+1}^{m} Y^*_{i[q_3(m+1)]k}\right), \\ \hfill l = m/2 \\ \hat{\mu}_{DPRSSO_e} = \dfrac{1}{mn}\displaystyle\sum_{k=1}^{n}\left(\sum_{i=1}^{j} Y^*_{i[q_1(m+1)]k} + Y^*_{(j+1)\left[\left(\frac{m+1}{2}\right)\right]k} + \sum_{i=j+2}^{m} Y^*_{i[q_3(m+1)]k}\right), \\ \hfill j = (m-1)/2. \end{cases}$$

The estimator of the population mean $\hat{\mu}_{DPRSS_e}$ with errors in ranking has the following properties:

1. $\hat{\mu}_{DPRSS_e}$ is unbiased estimator of the population mean if the population is symmetric about its mean.

2. $\mathrm{Var}(\hat{\mu}_{DPRSS_e})$ is less than $\mathrm{Var}(\hat{\mu}_{SRS})$.

3. For asymmetric distribution about its mean, $\mathrm{MSE}(\hat{\mu}_{DPRSS_e}) < \mathrm{Var}(\hat{\mu}_{SRS})$.

The above properties can be proved using Takahasi and Wakimoto [8], Dell and Clutter [3], Muttlak [6] and Al-Saleh and Al-Kadiri [1].

7. Evaluation of DPRSS for Estimating the Mean Weight of 342 Students

In this study, balanced RSS is considered, to illustrate the performance of DPRSS estimator for estimating the population mean of a real data set, we take the weights of 342 students in Wishah School, in UAE. We obtained the mean and the variance of the sample mean using SRS, RSS, MRSS, ERSS, PRSS and DPRSS methods with set sizes $m = 7$, 8, 10, 11, and 12. The samplings were carried out without replacement. Let ν_i, $i = 1, 2, \ldots, 342$ be the weight of the ith student in the population. Then the mean μ and the variance σ^2 of the population:

$$\mu = \frac{1}{342}\sum_{i=1}^{342}\nu_i = 50.047\text{kg} \text{ and } \sigma^2 = \frac{1}{342}\sum_{i=1}^{342}(\nu_i - \mu)^2 = 258.93\text{kg}^2.$$

Table 6. Empirical mean, variance and relative precision of RSS, MRSS, ERSS, PRSS and DPRSS with respect to SRS in the case of perfect ranking

Sampling methods		$m = 7$	$m = 8$	$m = 12$	$m = 10$	$m = 11$
SRS	Mean	50.093	50.011	50.046	50.086	49.999
	Variance	36.387	31.152	20.762	25.474	22.321

RSS		Mean	50.065	50.069	50.066		50.068	50.061
		Variance	12.051	9.472	4.827		6.531	5.616
		RP	3.019	3.289	4.301		3.900	3.975
MRSS		Mean	47.949	47.916	47.885		47.901	47.907
		Variance	5.617	4.389	2.039		2.786	2.363
		RP	3.632	3.488	7.135		3.448	3.216
ERSS		Mean	52.949	54.430	56.329		55.3667	55.227
		Variance	20.378	19.580	12.234		14.7304	12.555
		RP	1.263	0.803	0.406		0.592	0.567
PRSS	P20%	Mean	49.091	49.324	48.960	P25%	48.624	48.624
		Variance	9.426	8.227	3.518		3.524	3.524
		RP	3.519	3.481	4.418		4.591	4.022
	P30%	Mean	49.128	48.235	48.162	P35%	48.067	48.067
		Variance	9.505	5.254	2.342		2.580	2.580
		RP	3.516	3.650	3.522		3.919	3.434
	P40%	Mean	48.143	47.954	47.926	P45%	47.898	47.898
		Variance	6.286	4.522	2.045		2.369	2.369
		RP	3.672	3.499	3.173		3.646	3.194
DPRSS	P20%	Mean	50.139	51.738	49.778	P25%	49.003	49.003
		Variance	4.030	4.0136	0.907		0.865	0.865
		RP	9.010	4.532	21.190		13.034	11.421
	P30%	Mean	50.139	47.907	47.888	P35%	47.842	47.842
		Variance	4.030	0.988	0.350		0.351	0.351
		RP	9.010	5.595	4.143		4.888	4.283
	P40%	Mean	47.961	47.903	47.808	P45%	47.867	47.867
		Variance	0.915	0.890	0.215		0.338	0.338
		RP	6.913	5.678	3.970		5.007	4.387

It is noted that the DPRSS mean estimate is close to the population mean $\mu = 50.047\text{kg}$ computed from the entire population of the 342 students. It is obvious that the DPRSS procedure is more efficient than the SRS procedure for different cases that considered. In general, the recommendation is to use DPRSS for estimating the population mean of symmetric distributions with large sample sizes. In addition, using DPRSS will reduce the errors in ranking comparing to usual RSS. For asymmetric distributions, regardless of the smaller bias one can use DPRSS for estimating the population mean.

References

[1] M. F. Al-Saleh and M. Al-Kadiri, Double ranked set sampling, Statist. Probab. Lett. 48(2) (2000), 205-212.

[2] M. F. Al-Saleh and A. I. Al-Omari, Multistage ranked set sampling, J. Statist. Plann. Inference 102(2) (2002), 273-286.

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

[4] G. A. McIntyre, A method for unbiased selective sampling using ranked sets, Australian J. Agricultural Research 3 (1952), 385-390.

[5] H. A. Muttlak, Median ranked set sampling, J. Appl. Statist. Sci. 6(4) (1997), 245-255.

[6] H. A. Muttlak, Modified ranked set sampling methods, Pakistan J. Statist. 19(3) (2003), 315-323.

[7] H. Samawi, W. Abu-Dayyeh and S. Ahmed, Extreme ranked set sampling, Biom. J. 30 (1996), 577-586.

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

■