Inference on the Skew Normal Distribution Using Ranked Set Sampling

Mohammad Fraiwan Al-Saleh[1] and Mahmood Zuhier Aldarabseh[2]

[1] Dep. of Statistics, Yarmouk University, Jordan
[2] Dept. of Mathematics, Aljouf University- Saudi Arabia

Abstract: The estimation of the skewness parameter of the skew Normal Distribution (SND) using the information in a Ranked Set Sample (RSS) is considered. The estimators obtained are compared with respect to their counterparts using a Simple Random Sample (SRS). One main object is to study the unfavorable properties of some well-known estimators of the skewness parameter. Numerical comparisons among the estimators are conducted.

Keywords: Skew Normal Distribution; Simple Random Sampling; Ranked Set Sampling; Skewness Parameter.

1. INTRODUCTION

Simple Random Sampling (SRS) is the basic sampling technique and most important one. In this method, the researcher selects n units randomly from the population of size N in such a way so that all subsets of size n elements have the same probability of being selected.

The ranked set sampling is an important method of sampling that can reduce cost and time. This method was proposed by McIntyre (1952). It relies mainly on simple random sampling method in drawing the sample and judgment in ranking them. The procedure of RSS involves drawing m^2 sample units by simple random sampling method from the population, allocate the m^2 units into m sets randomly; now we have m sets each of m units, judgmentally rank the m units within each set from the lowest to the largest. Judgment ranking is based on the perception of the relative values for the variable of interest. Choose the lowest ranked unit from the first set, and the second lowest ranked unit from the second set, continuing until choosing the largest ranked unit from the last set and repeat the previous steps r times (cycles) if necessary to get a RSS of size $n = mr$ units. The elements of RSS are denoted by $X_{(i:m)}^{j}$, where $X_{(i:m)}^{j}$ is the ith lowest element from the ith set of the jth cycle, where, $j = 1,2,...,r$ and $i = 1,2,...,m$.

McIntyre (1952) used RSS method for estimating mean yields in pasture. Takahasi and Wakimoto (1968) estimated the population mean based on RSS and found a basis for the theory of RSS. Samawi et al. (1996) introduced the extreme ranked set sampling (ERSS) method. Al-Saleh and Al-Kadiri (2000) introduced the double ranked set sampling method (DRSS), this method increases the efficiency of RSS without increasing the set size m; which should be kept small. Al-Saleh and Muttlak (2000) considered Bayesian estimation using RSS method. Al-Saleh and Al-Omari (2002) introduced the multistage ranked set sampling (MSRSS) method as a generalization of DRSS method. The efficiency of estimators using RSS for all distributions is found to be between 1 and m^2. See also, Al-Saleh (2004), Al-Saleh and Samuh (2008). Al-Odat and Al-Saleh (2001) introduced the moving extreme ranked set sampling by varying the set size m in RSS method. See also, Al-Saleh and Ababneh (2015) and Al-Saleh and Naamneh (2016). More details about RSS are given in section(4).

E-mail address: m-saleh@yu.edu.jo, mah.darabseh@gmail.com

2. SKEW NORMAL DISTRIBUTION

The normal distribution (Gaussian distribution) has the following probability density function (pdf):

$$f(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \; -\infty < x < \infty, -\infty < \mu < \infty, \; \sigma^2 > 0,$$

μ is the population mean and σ² is the population variance. If $\mu = 0$ and $\sigma^2 = 1$, then the distribution is the standard normal distribution, denoted by $N(0,1)$ with pdf and cdf given by

$$\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}, \; \Phi(x) = \int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz, \; -\infty < x < \infty.$$

The univariate skew normal distribution $SN(\lambda)$ was proposed by Azzalini (1985). If $X \sim SN(\lambda)$, then X has the following pdf:

$$f(x;\lambda) = 2\phi(x)\Phi(\lambda x), -\infty < \lambda < \infty, -\infty < x < \infty,$$

λ is called the **_Skewness parameter_**. (See Brown (2001)). The graphs of $SN(\lambda)$ for some values of the Skewness parameter are given below (they are graphed using Scientific Work Place, SWP). The graphs indicate that as $|\lambda|$ gets large the graph becomes more skewed (to the right for positive λ and to the left for negative λ).

$$\lambda = 0 \qquad\qquad\qquad \lambda = 3$$
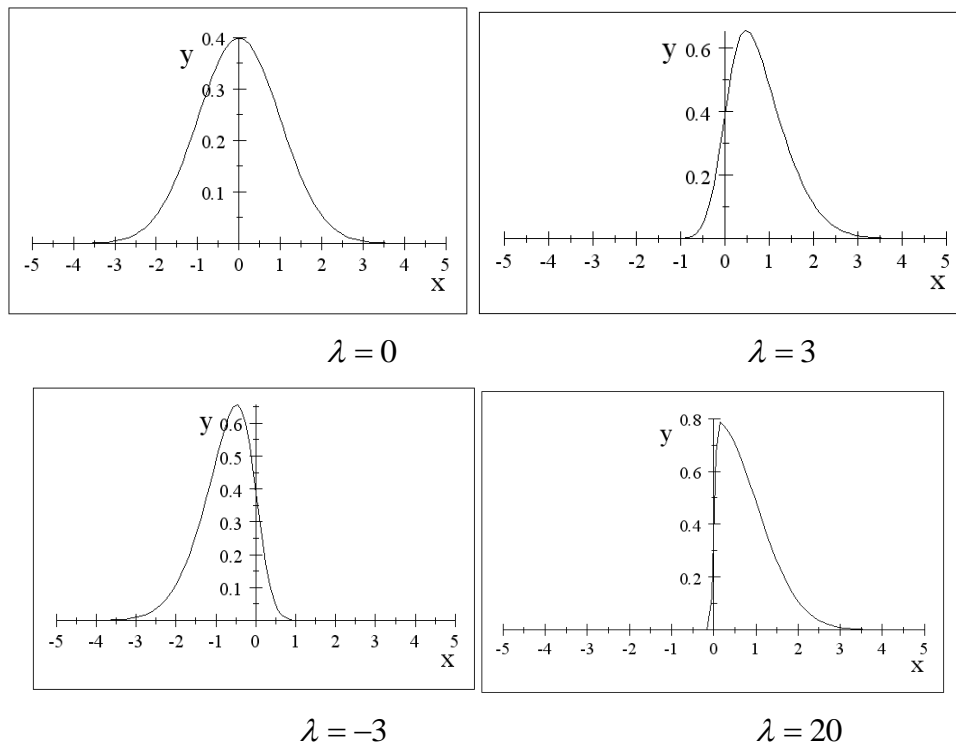

$$\lambda = -3 \qquad\qquad\qquad \lambda = 20$$

Figure 1.

If $Y = \mu + \sigma X$, then Y has the skew normal, $SN(\lambda, \mu, \sigma^2)$, with pdf given by

$$f(y;\lambda,\mu,\sigma) = \frac{2}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\Phi\left(\lambda\frac{y-\mu}{\sigma}\right).$$

As $\lambda \to \infty$, $\Phi\left(\lambda \dfrac{x-\mu}{\sigma}\right) \to 1$ when $x > \mu$, and $\Phi\left(\lambda \dfrac{x-\mu}{\sigma}\right) \to 0$ when $x < \mu$. Thus, the skew normal density

tends to the half normal density, denoted by $X \sim HN(\mu, \sigma^2)$, with pdf

$$f(x; \mu, \sigma^2) = \frac{2}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right), \ x > \mu \ and \ 0 \ o.w.$$

If U and V are iid $N(0,1)$ then (see Brown 2001):

$$X = \frac{\lambda}{\sqrt{1+\lambda^2}} |U| + \frac{1}{\sqrt{1+\lambda^2}} V \ \text{is } SN(\lambda).$$

The following are some properties of the skew normal distribution (see Brown (2001) :

1. The standard normal distribution is a special case of the skew normal distribution; it is $SN(0)$.

2. As $\lambda \to \infty$, $f(x; \lambda)$ approaches the half normal density.

3. If $X \sim SN(\lambda)$, then $-X \sim SN(-\lambda)$.

4. If $X \sim SN(\lambda)$, then $X^2 \sim \chi^2_{(1)}$. Thus, the chi-square random variable may be generated from the skew normal not only from the normal.

If $X \sim SN(\lambda)$, then the moment generating function of X is

$$M_X(t) = 2 e^{\frac{t^2}{2}} \Phi\left(\frac{\lambda t}{\sqrt{1+\lambda^2}}\right).$$

Using $M_X(t)$, we have

$$E(X) = \mu = M_X^{(1)}(0) = \sqrt{\frac{2}{\pi}} \left(\frac{\lambda}{\sqrt{1+\lambda^2}}\right); \ E(X^2) = M_X^{(2)}(0) = 2\Phi(0) = 1;$$

thus, the variance of X is

$$Var(X) = \sigma^2 = E(X^2) - (E(X))^2 = 1 - \frac{2\lambda^2}{\pi(1+\lambda^2)} = 1 - \mu^2.$$

Clearly, $|\mu| \leq 1$. Actually, it can be verified that $|\mu| < \sqrt{\dfrac{2}{\pi}}$.

Sartori (2005) noted some problems in estimating the skew parameter λ: If the observations are all positive then the maximum likelihood estimate of the skew parameter is infinite. Also, the same problem occurs for $SN(\mu, \sigma^2, \lambda)$. This problem is investigated in more details in the next section.

3. Estimation the Skewness Parameter Using Method of Moments (MME) Based on SRS, RSS

Let $X_1, X_2, ..., X_n$ be iid $SN(\lambda)$. To obtain the MME of λ we need to equate the first population moment to the first sample moment and solve for λ :

$$E(X) = \overline{X} \Leftrightarrow \sqrt{\frac{2}{\pi}} \frac{\lambda}{\sqrt{1+\lambda^2}} = \overline{X} \Leftrightarrow \lambda = \frac{\overline{X}}{\sqrt{\dfrac{2}{\pi}}} \sqrt{1+\lambda^2} \Rightarrow \lambda^2 = \frac{\overline{X}^2}{\dfrac{2}{\pi} - \overline{X}^2}.$$

Clearly, λ has the same sign of $\overline{X}$; thus the MME of λ is

$$\hat{\lambda}_{SRS(MME)} = \overline{X}\sqrt{\dfrac{1}{\dfrac{2}{\pi} - \overline{X}^2}} \;,$$

provided that $-\sqrt{\dfrac{2}{\pi}} < \overline{X} < \sqrt{\dfrac{2}{\pi}}$.

Thus, MME of λ exists only if $-\sqrt{\dfrac{2}{\pi}} < \overline{X} < \sqrt{\dfrac{2}{\pi}}$. If $\overline{X} = \pm\sqrt{\dfrac{2}{\pi}}$, then its value is $\pm\infty$, while if $\overline{X} > \sqrt{\dfrac{2}{\pi}}$ or

$\overline{X} < -\sqrt{\dfrac{2}{\pi}}$ then the value is imaginary (not a real number). Thus, this estimator may have a problem of assuming

values outside the domain of λ ($\pm\infty$ or imaginary values). Now, we need to know how often such improper thing occurs. The probability of such thing is

$$P^* = 1 - \Pr\left(\hat{\lambda}_{MME}\, finite\right) = 1 - p\left(-\sqrt{\dfrac{2}{\pi}} < \overline{X} < \sqrt{\dfrac{2}{\pi}}\right) = p\left(\left|\overline{X}\right| \geq \sqrt{\dfrac{2}{\pi}}\right).$$

Approximate values of this probability are obtained based on simulation. To find this value we need to get random values from the skew normal distribution. For a given value of λ, this is done by the following steps:

1) Simulate two independent random values, U and V, from N(0,1), then the random variable

$$\dfrac{\lambda}{\sqrt{1+\lambda^2}}\left|U\right| + \dfrac{1}{\sqrt{1+\lambda^2}}V \text{ is } SN(\lambda).$$

2) Step 1 is repeated n times to get a random sample of size n from $SN(\lambda)$. The average of this sample, say $\overline{X}_1$ is obtained.

3) Steps 1-2 can be repeated L times to obtain $\overline{X}_1, \overline{X}_2, ..., \overline{X}_L$.

4) For large L and based on the strong law of large numbers, P^* can be approximated by

$$P^* \approx \dfrac{1}{L}\sum_{i=1}^{L} I\left(\left|\overline{X}_i\right| \geq \sqrt{\dfrac{2}{\pi}}\right).$$

Table(1) gives simulated values of P^* for some values of n and λ. Where $I\left(\left|\overline{X}_i\right| \geq \sqrt{\dfrac{2}{\pi}}\right)$ is 1 if $\left(\left|\overline{X}_i\right| \geq \sqrt{\dfrac{2}{\pi}}\right)$

and zero otherwise.

Let $Y_1, Y_2, ..., Y_m$ be a ranked set sample from the $SN(\lambda)$ distribution then

$$\hat{\lambda}_{RSS(MME)} = \sqrt{\dfrac{\overline{Y}^2}{\dfrac{2}{\pi} - \overline{Y}^2}}\;.$$

Again this estimator may have a problem of assuming values outside the domain of λ ($\pm\infty$ or imaginary values). The

probability of such thing is again, $\Pr\left(\left|\overline{Y}\right| \geq \sqrt{\dfrac{2}{\pi}}\right)$.

The values in brackets in Table(1) are simulated values of P^{**} for some values of n and λ, where $m = 3$, $n = mr = 3r$, $r = 1,2,4,6$. The values of the table indicate how often $\hat{\lambda}_{SRS(MME)}$ $\hat{\lambda}_{RSS(MME)}$ are improper. We note that P^* is approximately the same for positive and negative λ; for fixed n, P^* tends to increase in $\left|\lambda\right|$. For

fixed λ, P^* tends to decrease in n if $|\lambda|$ is small and increase in n if $|\lambda|$ is large. Clearly, the MME has a major problem for samples of small size. The same thing can be said about P^{**}; MME has a major problem for samples of small size or large $|\lambda|$.

Table 1. Some simulated values of P^* and P^{**}

n \ λ	3	6	12	24
0	0.1685 (0.0560)	0.0485 (0.0070)	0.0044 0.0000	0.0002 0.0000
1	0.3111 (0.2830)	0.2430 (0.1700)	0.1615 (0.1030)	0.0846 (0.0230)
-1	0.3108 (0.2540)	0.2424 (0.1640)	0.1633 (0.0670)	0.0892 (0.0260)
3	0.4266 (0.4350)	0.4220 (0.3720)	0.4048 (0.3860)	0.3667 (0.3430)
-3	0.4285 (0.4156)	0.4188 (0.3520)	0.4034 (0.3466)	0.3770 (0.3260)
5	0.4417 (0.4476)	0.4499 (0.4640)	0.4555 (0.4590)	0.4391 (0.4110)
-5	0.4432 (0.4486)	0.4509 (0.4360)	0.4527 (0.4860)	0.4363 (0.4690)
10	0.4556 (0.4640)	0.4732 (0.4470)	0.4687 (0.4400)	0.4731 (0.5360)
-10	0.4518 (0.4820)	0.4666 (0.4590)	0.4681 (0.4840)	0.4668 (0.5080)
20	0.4656 (0.4510)	0.4711 (0.4810)	0.4730 (0.5010)	0.4828 (0.4890)
-20	0.4565 (0.4440)	0.4753 (0.4860)	0.4832 (0.5190)	0.4843 (0.4960)
100	0.4652 (0.4860)	0.4689 (0.4680)	0.4836 (0.4560)	0.4812 (0.5120)
-100	0.4625 (0.4730)	0.4663 (0.4900)	0.4799 (0.4770)	0.4887 (0.4600)

Again, if $X_1, X_2, ..., X_n$ is a random sample from $SN(\lambda)$. The log likelihood function for the parameter λ based on the observed sample is given by

$$L(\lambda) = n \log 2 + \sum_{i=1}^{n} \log \phi(x_i) + \sum_{i=1}^{n} \log \Phi(\lambda x_i).$$

Now,

$$\frac{dL(\lambda)}{d\lambda} = \sum_{i=1}^{n} \frac{\phi(\lambda x_i)}{\Phi(\lambda x_i)} x_i \ .$$

As noted by Martinez et al. (2008), if $x_1, x_2, ..., x_n$ are all positive then $\sum_{i=1}^{n} \frac{\phi(\lambda x_i)}{\Phi(\lambda x_i)} x_i$ is positive and increasing

in λ, hence there is no MLE for λ. A similar problem occurs if all $x_1, x_2, ..., x_n$ are negative. Otherwise, the MLE

of λ can be obtained numerically, by setting $\dfrac{dL(\lambda)}{d\lambda}$ to zero and solve for λ.

To investigate how often this problem (MLE doesn't exist) occurs, we need to evaluate the quantity

$\prod_{i=1}^{n} P(X_i < 0) + \prod_{i=1}^{n} P(X_i > 0)$:

$$\prod_{i=1}^{n} P(X_i < 0) + \prod_{i=1}^{n} P(X_i > 0) = \left[\int_{-\infty}^{0} 2\phi(x)\Phi(\lambda x)dx\right]^n + \left[\int_{0}^{\infty} 2\phi(x)\Phi(\lambda x)dx\right]^n$$

This quantity is a function of λ, say $g(\lambda)$;

If $g^{+}(\lambda) = P(X > 0) = \int_{0}^{\infty} 2\phi(x)\Phi(\lambda x)dx$, then,

$$\frac{d}{d\lambda} g^{+}(\lambda) = \frac{d}{d\lambda} 2\int_{0}^{\infty} x\phi(x)\phi(\lambda x)dx = \frac{d}{d\lambda} 2\int_{0}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\lambda^2 x^2} dx = \frac{1}{\pi(1+\lambda^2)} \ .$$

Thus,
$$g^{+}(\lambda) = \int_{0}^{\infty} \frac{1}{\pi(1+\lambda^2)} d\lambda = \frac{1}{\pi} \tan^{-1}(\lambda) + c \ ,$$

$$g^{+}(0) = 2\int_{0}^{\infty} \phi(x)\Phi(0)dx = 2\int_{0}^{\infty} \phi(x)\frac{1}{2}dx = \frac{1}{2} \ .$$

Now, $\dfrac{1}{2} = \dfrac{1}{\pi}\tan^{-1}(0) + c \Rightarrow c = \dfrac{1}{2}$. Hence,

$$g^{+}(\lambda) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(\lambda).$$

Similarly,

$$g^{-}(\lambda) = P(x < 0) = 1 - g^{+}(\lambda) = \frac{1}{2} - \frac{1}{\pi}\tan^{-1}(\lambda)$$

Thus,

$$g(\lambda) = \left[\frac{1}{2} + \frac{1}{\pi}\tan^{-1}(\lambda)\right]^n + \left[\frac{1}{2} - \frac{1}{\pi}\tan^{-1}(\lambda)\right]^n.$$

It can be shown that for $0 < p < 1$, $p^n + (1-p)^n \geq \left(\dfrac{1}{2}\right)^{n-1}$.

The above results are the content of the following theorem:

Theorem (1): Assume that $X_1, X_2, ..., X_n$ is random sample from $SN(\lambda)$. Then:

(1) The probability that the MLE of λ is infinite is given by

$$g(\lambda) = \left[\frac{1}{2} + \frac{1}{\pi}\tan^{-1}(\lambda)\right]^{n} + \left[\frac{1}{2} - \frac{1}{\pi}\tan^{-1}(\lambda)\right]^{n}.$$

(2) $g(\lambda) \geq \left(\frac{1}{2}\right)^{n-1}$.

Table (2) gives some numerical value of $g(\lambda)$. The values of the table indicate how often $\hat{\lambda}_{SRS(MLE)}$ is infinite. We note that the probability of occurrence of this problem, $g(\lambda)$, is the same for positive and negative λ because $\tan^{-1}(\lambda) = -\tan^{-1}(-\lambda)$. At fixed n $g(\lambda)$ is increasing in $|\lambda|$. For fixed λ, $g(\lambda)$ is decreasing in n. We can conclude that if we feel that $|\lambda|$ is large, then we need a large sample to be able to estimate λ using MLE.

Table (2). Probability of infinite values of MLE of λ

n / λ	3	5	8	12	15	20	50	100	10000
0	0.2500	0.0625	0.0078	0.0005	0.0001	0.0000	0.0000	0.0000	0.0000
±1	0.4375	0.2383	0.1001	0.0317	0.0134	0.0032	0.0000	0.0000	0.0000
±3	0.7242	0.5826	0.4213	0.2735	0.1978	0.1152	0.0045	0.0000	0.0000
±5	0.8234	0.7229	0.5950	0.4590	0.3778	0.2731	0.0390	0.0015	0.0000
±7	0.8706	0.7937	0.6909	0.5743	0.4999	0.3968	0.0992	0.0098	0.0000
±10	0.9078	0.8511	0.7727	0.6792	0.6166	0.5248	0.1995	0.0398	0.0000
±15	0.9378	0.8985	0.8425	0.7734	0.7252	0.6516	0.3427	0.1175	0.0000
±25	0.9623	0.9380	0.9026	0.8575	0.8252	0.7740	0.5271	0.2778	0.0000
±40	0.9763	0.9609	0.9381	0.9086	0.8871	0.8524	0.6707	0.4499	0.0000
±100	0.9905	0.9842	0.9748	0.9625	0.9533	0.9382	0.8526	0.7270	0.0000
±10000	0.9999	0.9998	0.9997	0.9996	0.9995	0.9994	0.9984	0.9968	0.7274

Based on the previous results obtained, the Skewness of the skew normal distribution depends on the Skewness parameter, and there are problems in estimating the Skewness parameter using MM or ML estimation. The problems can be roughly resolved by increasing the sample size whenever the Skewness parameter is believed to be large.

4. Estimation of the Mean of the Skew Normal Based on SRS, RSS

Ranked set sampling is an important method of sampling that can reduce cost and time. This method was proposed by **McIntyre (1952).** It relies mainly on simple random sampling method in drawing the sample and judgment in ranking them. The following steps describe the RSS technique:

Step 1: Draw randomly m sets, every set consists of m units using simple random sampling method from the population.

Step 2: Judgmentally rank the m units within each set from lowest to largest with respect to a variable of interest.

Step 3: Choose the lowest ranked unit from the first set, and the second lowest ranked unit from the second set, continuing until choosing the largest ranked unit from the last set.

Step 4: Repeat, if necessary, the previous steps r times (cycles) to obtain a RSS of size $n = mr$ units.

For $r = 1$, the ranked set sample is $X_{(1:m)}, X_{(2:m)}, ..., X_{(m:m)}$. The ranked set sampling mean estimator is

$$\hat{\mu}_{RSS} = \frac{1}{m} \sum_{i=1}^{m} X_{(i:m)} \cdot$$

In the case of r cycles the above estimator will be

$$\hat{\mu}_{RSS} = \frac{1}{mr} \sum_{j=1}^{r} \sum_{i=1}^{m} X_{(i:m)}^{j} \cdot$$

Let the underlying pdf be $f(x)$, with mean μ and variance σ^2, and the $(i:m)^{th}$ order statistic have density function $f_{(i:m)}(x)$ with mean $\mu_{(i:m)}$ and variance $\sigma_{(i:m)}^2$, based on **Takahasi and Wakimoto (1968)** we have

$$f(x) = \frac{1}{m} \sum_{i=1}^{m} f_{(i:m)}(x) \cdot$$

Thus,

$$\mu = \frac{1}{m} \sum_{i=1}^{m} \mu_{(i:m)} \text{ and } \quad \sigma^2 = \frac{1}{m} \sum_{i=1}^{m} \sigma_{(i:m)}^2 + \frac{1}{m} \sum_{i=1}^{m} \left(\mu_{(i:m)} - \mu \right)^2 \cdot$$

$$\frac{\sigma^2}{m} = \frac{1}{m^2} \sum_{i=1}^{m} \sigma_{(i:m)}^2 + \frac{1}{m^2} \sum_{i=1}^{m} \left(\mu_{(i:m)} - \mu \right)^2,$$

.

When $r = 1$, and if $\hat{\mu}_{SRS}$ is the mean of a SRS of size m, then

$$VAR(\hat{\mu}_{SRS}) = \frac{\sigma^2}{m}$$

and

$$Var(\hat{\mu}_{RSS}) = \frac{1}{m^2} \sum_{i=1}^{m} \sigma_{(i::m)}^2 = \frac{\sigma^2}{m} - \frac{1}{m^2} \sum_{i=1}^{m} \left(\mu_{(i:m)} - \mu \right)^2 \cdot$$

The efficiency of $\hat{\mu}_{RSS}$ w.r.t. $\hat{\mu}_{SRS}$ is

$$Eff(\hat{\mu}_{RSS}, \hat{\mu}_{SRS}) = \frac{Var(\hat{\mu}_{SRS})}{Var(\hat{\mu}_{RSS})} = \frac{1}{1 - \frac{1}{m\sigma^2} \sum_{i=1}^{m} \left(\mu_{(i:m)} - \mu \right)^2} = \frac{1}{1 - \frac{1}{m} \sum_{i=1}^{m} \left(\frac{\mu_{(i:m)} - \mu}{\sigma} \right)^2} \cdot$$

Let $X \sim SN(\lambda)$, then

$$f(x, \lambda) = 2\phi(x)\Phi(\lambda x), \ F(x, \lambda) = 2 \int_{-\infty}^{x} \phi(y)\Phi(\lambda y) dy$$

and

$$f_{(i:m)}(x) = m \binom{m-1}{i-1} (F(x,\lambda))^{i-1} (1 - F(x,\lambda))^{m-i} f(x,\lambda)$$

$$Eff(\hat{\mu}_{RSS}, \hat{\mu}_{SRS}) = \frac{1}{1 - \dfrac{\sum_{i=1}^{m} \left(\mu_{(i:m)} - \sqrt{\dfrac{2}{\pi}} \dfrac{\lambda}{\sqrt{1+\lambda^2}} \right)^2}{m \times \left(1 - \dfrac{2\lambda^2}{\pi(1+\lambda^2)} \right)}} \, .$$

Table(3) gives some values for the efficiency of $\hat{\mu}_{RSS}$ with respect to $\hat{\mu}_{SRS}$ for some values of m and λ.

Table (3): Efficiency of $\hat{\mu}_{RSS}$ w.r.t. $\hat{\mu}_{SRS}$ for $SN(\lambda)$ distribution

m \ λ	3	4	5
0	1.9388	2.3197	2.7549
1	1.8679	2.2748	2.6802
-1	1.9152	2.3755	2.7197
3	1.8122	2.3572	2.6349
-3	1.8430	2.2058	2.5605
5	1.8977	2.2854	2.6439
-5	1.8833	2.2489	2.5793
10	1.9430	2.3724	2.6229
-10	1.8560	2.2459	2.2783
20	1.7996	2.2351	2.4738
-20	1.8391	2.2287	2.5994
100	1.8390	2.1880	2.6926
-100	1.7954	2.2720	2.7215

The values of table indicate that the efficiency of $\hat{\mu}_{RSS}$ w.r.t. $\hat{\mu}_{SRS}$ is increasing in m for fixed λ, and almost the same or fixed m and varied λ. Overall, $\hat{\mu}_{RSS}$ is more efficient than $\hat{\mu}_{SRS}$

5. Half Normal Distribution

In this section, we study the Half Normal Distribution (HND). We derive and study some properties of the MME estimators of the parameter μ for HND based on SRS, RSS.

Let $Y \sim SN(\lambda)$, then $X = \mu + \sigma Y \sim SN(\lambda, \mu, \sigma^2)$ with pdf

$$f(x, \lambda, \mu, \sigma^2) = \frac{2}{\sigma} \phi\left(\frac{x-\mu}{\sigma} \right) \Phi\left(\lambda \frac{x-\mu}{\sigma} \right),$$

as $\lambda \to \infty$, $\Phi\left(\lambda \dfrac{x-\mu}{\sigma}\right) \to 1$, when $x > \mu$ and $\Phi\left(\lambda \dfrac{x-\mu}{\sigma}\right) \to 0$, when $x < \mu$.

Thus, the above density tends to the half normal density, denoted by $X \sim HN(\mu, \sigma^2)$, where

$$f(x, \mu, \sigma^2) = \frac{2}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right), \qquad\qquad x > \mu.$$

The following facts can be easily verified:

The moment generating function for HND is given by

$$M_X(t) = 2\Phi(\sigma t)e^{\frac{1}{2}\sigma^2 t^2 + \mu t},$$

thus,

$$E(X) = \mu + \sigma\sqrt{\frac{2}{\pi}}, \quad Var(X) = \left(1 - \frac{2}{\pi}\right)\sigma^2.$$

The distribution function for HND is

$$F(x; \mu, \sigma^2) = \int_{\mu}^{x} \frac{2}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)dy = 2\Phi\left(\frac{x-\mu}{\sigma}\right) - 1.$$

Let $X_1, X_2, ..., X_n$ be a random sample from $HN(\mu, 1)$. Equating the first population moment to the first sample moment to get the estimator:

$$\hat{\mu}_{SRS(MME)} = \overline{X} - \sqrt{\frac{2}{\pi}}.$$

Now,

$$E\left(\hat{\mu}_{SRS(MME)}\right) = E\left(\overline{X} - \sqrt{\frac{2}{\pi}}\right) = E(\overline{X}) - \sqrt{\frac{2}{\pi}} = \left(\mu + \sqrt{\frac{2}{\pi}}\right) - \sqrt{\frac{2}{\pi}} = \mu.$$

Thus, $\hat{\mu}_{SRS(MME)}$ is an unbiased estimator of μ with variance:

$$Var\left(\hat{\mu}_{SRS(MME)}\right) = V\left(\overline{X} - \sqrt{\frac{2}{\pi}}\right) = V(\overline{X}) = \frac{Var(X)}{n} = \frac{\sigma^2}{n}\left(1 - \frac{2}{\pi}\right).$$

Similarly, if $Y_1, Y_2, ..., Y_m$ be a ranked set sample from $HN(\mu, 1)$ then

$$\hat{\mu}_{RSS(MME)} = \overline{Y} - \sqrt{\frac{2}{\pi}}.$$

$\hat{\mu}_{RSS(MME)}$ is an unbiased estimator of μ :

$$E\left(\hat{\mu}_{RSS(MME)}\right) = E\left(\overline{Y} - \sqrt{\frac{2}{\pi}}\right) = \frac{1}{m}\sum_{i=1}^{m} E(Y_{(i:m)}) - \sqrt{\frac{2}{\pi}} = \frac{1}{m}\sum_{i=1}^{m}\mu_{(i:m)} - \sqrt{\frac{2}{\pi}}\ \mu + \sqrt{\frac{2}{\pi}} - \sqrt{\frac{2}{\pi}} = \mu.$$

$$\textit{Eff}\ (\hat{\mu}_{RSS(MME)}; \hat{\mu}_{SRS(MME)}) = \frac{Var\left(\hat{\mu}_{SRS(MME)}\right)}{Var\left(\hat{\mu}_{RSS(MME)}\right)},$$

where ,

$$Var\left(\hat{\mu}_{SRS(MME)}\right) = Var\left(\overline{Y} - \sqrt{\frac{2}{\pi}}\right) = Var(\overline{Y}) = \frac{\sigma^2}{m} = \frac{1}{m}\left(1 - \frac{2}{\pi}\right),$$

$$Var\left(\hat{\mu}_{RSS(MME)}\right) = \frac{\sigma^2}{m} - \frac{1}{m^2}\sum_{i=1}^{m}\left(\mu_{(i:m)} - \mu\right)^2 ; \ \sigma^2 = Var(Y) = \left(1 - \frac{2}{\pi}\right), \mu = E(Y) = \sqrt{\frac{2}{\pi}} .$$

Therefore the relative efficiency $Eff\left(\hat{\mu}_{RSS(MME)}; \hat{\mu}_{SRS(MME)}\right)$ is

$$Eff = \frac{\dfrac{\sigma^2}{m}}{\dfrac{\sigma^2}{m} - \dfrac{1}{m^2}\sum_{i=1}^{m}\left(\mu_{(i:m)} - \mu\right)^2} = \frac{1}{1 - \dfrac{\pi}{m(\pi - 2)}\sum_{i=1}^{m}\left(\mu_{(i:m)} - \sqrt{\dfrac{2}{\pi}}\right)^2}$$

The following table gives some values for the efficiency of $\hat{\mu}_{RSS(MME)}$ with respect to $\hat{\mu}_{SRS(MME)}$ for some values of m.

Table (4). Efficiency of $\hat{\mu}_{RSS(MME)}$ with respect to $\hat{\mu}_{SRS(MME)}$ for $HN(\mu,1)$

m	3	4	5
Eff	1.8408	2.2390	2.6280

The values of this table indicate that $\hat{\mu}_{RSS(MME)}$ is more efficient than $\hat{\mu}_{SRS(MME)}$ and the efficiency is increasing in m.

6. Concluding Remarks

Different estimators of the skew parameter or a function of it were discussed and investigated based on the two sampling techniques SRS and RSS. It turned out that the estimators using RSS are more efficient than those using SRS. The half normal distribution is one extreme example of the skew normal; it is the limit of the skew normal when the skew parameter approaches $\pm\infty$. Inference about this distribution is studied in some more details.

One main problem in estimating the skew parameter using usual estimation techniques (MME and MLE) is that the possibility of obtaining values of the estimator that are outside the accepted values of the parameter (infinite or imaginary values). The probability of such improper values is investigated.

REFERENCES

[1] Al-Odat, M. and Al-Saleh, M.F. (2001). A variation of ranked set sampling. *Journal of Applied Statistical Science* **10**, 137–146.

[2] Al-Saleh, M. F. (2004). Steady state ranked set sampling and parametric estimation. *Journal of Statistical Planning and Inference* **123**, 83-95.

[3] Al-Saleh, M. F. and Ababneh, A. (2015). Test for Accuracy in Ranking in Moving Extreme Ranked Set Sampling. International Journal of Computational and Theoretical Statistics 2, 67-77.

[4] Al-Saleh, M.F. and Al-Kadri, M. (2000). Double Ranked Set Sampling. *Statistics and Probability Letters* **48**, 205-212.

[5] Al-Saleh, M.F. and Al-Omari, A.I (2002). Multistage ranked set sampling. *Journal of Statistical Planning and Inference* **102**, 273-286.

[6] Al-Saleh, M. F. and Naamneh, A. (2016). Properties of the Elements of SRS, RSS, MERSS. Journal of Applied Statistical Science(JASS) 22, 75-85.

[7] Al-Saleh, M.F. and Muttlak, H.A. (2000). Bayesian estimation using ranked set sampling. *Biometrical Journal.* **42**, 489-500.

[8] Al-Saleh, M.F. and Samuh, M. (2008). On multistage ranked set sampling for distribution and median estimation. *Computational Statistics and Data Analysis* **52**, 2066-2078.

[9] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics* **12**, 171-178.

[10] Brown, N.D. (2001). Reliability Studies of the Skew Normal distribution. Master Thesis, University of Maine.

[11] Martinez, E. H., Varela, H., Gomez, H.W. and Bolfarine, H. (2008). A note on the likelihood and moments of the skew-normal distribution. *SORT* **32**, 57-66.

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

[13]   Samawi, H.M. Ahmed, M. S. and Abu-Dayyeh, W. (1996). Estimating the population mean using extreme ranked set sample. *Biometrical Journal* **5**, 577-586.

[14]   Takahasi, K. and Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering. *Annals of the Institute of Statistical Mathematics* **20**, 421-428.

[15]   Sartori, N. (2005). Bias prevention of maximum likelihood estimates for scalar skew normal and skew t distributions. *Journal of Statistical Planning and Inference* **136**, 4259-4275.