King Fahd University of Petroleum & Minerals

DEPARTMENT OF MATHEMATICAL SCIENCES

Technical Report Series

TR 384

Mar 2008

Improved Estimation and Large Sample Theory in Ranked Set Sampling

Hassen A. Muttlak, Sayed E. Ahmed and Ibrahim Rahimov

Improved Estimation and Large Sample Theory in Ranked Set Sampling

Hassen A. Muttlak[1], Sayed E. Ahmed[2] and Ibrahim Rahimov[1]

[1]Department of Mathematics and Statistics, King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia
[2]Department of Mathematics and Statistics, University of Windsor, Windsor, Ontario, Canada

Abstract

The purpose of this article is two-fold. First, we consider the ranked set sampling (RSS) estimation and test of hypothesis for the parameter of interest. Then, we suggest some improved estimation strategies for the mean parameter based on shrinkage and pretest principles. Generally speaking, the shrinkage and pretest methods use the non-sample information (NSI) regarding the parameter of interest. In practice, NSI is readily available in the form of a realistic conjecture based on the experimenter's knowledge and experience with the problem under consideration. It is advantageous to utilize NSI in the estimation process to construct improved estimation for the parameter of interest. In this contribution, the large sample properties of the suggested estimators will be assessed both analytically and numerically. More importantly, Monte Carlo simulation will be conducted to investigate the relative performance of the estimators for moderate and large samples. For illustration purposes, the proposed methodology is applied to a published data set.

Keywords: Asymptotic properties, local alternatives, ranked set sampling, relative precision, replications, and shrinkage and pretest estimation.

Mathematics Subject Classification 62F10, 62F12.

[1]Corresponding author, E-mail: hamstat@kfupm.edu.sa

1. Introduction

McIntyre (1952) was first to suggest using ranked set sampling (RSS) to estimate the population mean instead of the usual simple random sampling (SRS). Takahasi and Wakimoto (1968) supplied the necessary mathematical theory. Dell and Clutter (1972) studied the case in which the ranking may not be perfect i.e. there are errors in ranking the units with respect to the variable of interest. To over come the problem of errors in ranking, Muttlak (1997) proposed using the median ranked set sampling. For classified and extensively reviewed work in the area of RSS see Patil et al (1994), Kaur et al. (1995), Muttlak and Al-Saleh (2000) and recent book by Chen et al. (2004).

The ranked set sampling (RSS) procedure can be summarized as follows: Select n random sets, each of size n units from the population, and rank the units within each set with respect to a variable of interest. Then an actual measurement is taken from the unit with the smallest rank from the first set. From the second set, an actual measurement is taken from the unit with the second smallest rank, and the procedure is continued until the unit with the largest rank is chosen for actual measurement from the n-th set. In this way, we obtain a sample of n measured units, one from each set. The cycle may be repeated r times until nr units have been measured. These nr units form the RSS data.

Thompson (1968) advocated that the shrinkage of sample estimator towards a natural origin. In various context, this estimator is suggested by a host of researchers as evident by a number of publications in statistical and related journals. Recently, Schäfer and Strimmer, (2005) applied this strategy to large-scale covariance matrix estimation with implications to functional genomics. Ahmed and Krzanowski (2004), Bickel and Doksum (2001) and other pointed out that shrinkage estimator yields smaller mean squared error (MSE) when a priori information is correct or nearly correct, however at the expense of poorer performance in the rest of the parameter space induce by the prior information. In recent literature, a discussion about pretesting can be found in Giles and Giles (1993), Magnus (1999), Ohanti (1999), Reif and VIcek (2002), Khan and Ahmed (2003), among many others.

In the present investigation, we develop the inferential methods for the population mean μ, based on ranked set sampling. The estimation and testing procedures are presented in Section 2. The improved estimation procedures based on shrinkage and the preliminary test method are considered in Section 3 along with deriving analytic mean squared error of the estimators. Also, in Section 3, we compare our estimators with the natural estimator and show that our methods are asymptotically more efficient when the uncertain prior information regarding the parameter is correct or nearly correct. The results of simulation experiment are given in Section 4. In Section 5 we present the results of the analysis of a data set. We provide concluding remarks in Section 6.

2. Test of Hypotheses and Improve Estimation

Let $X_{(i:n)j}$ denote the $i-th$ order statistic from the $i-th$ set of size n in the $j-th$ cycle; $\mathrm{i}=1, 2, ..., \mathrm{n}$ and $\mathrm{j}=1, 2, ..., \mathrm{r}$. Then the unbiased estimator of the population mean μ, see Takahasi and Wakimoto (1968) using RSS data is given by

$$\hat{\mu}_{rss} = \frac{1}{r\,n}\sum_{j=1}^{r}\sum_{i=1}^{n} X_{(i:n)j}$$

with variance

$$\mathrm{var}\left(\hat{\mu}_{rss}\right) = \frac{1}{r\,n^2}\sum_{i=1}^{n}\sigma^2_{(i:n)} = \frac{\upsilon^2}{r},$$

where $\sigma^2_{(i:n)} = E\left[X_{(i:n)} - E\left(X_{(i:n)}\right)\right]^2$, i.e. the variance of the $i-th$ order statistics from set of size n.

2.1 The Asymptotic normality for the RSS estimator

In the RSS we may let the set size n or the number of cycles r to be large. Now we establish the asymptotic normality of $\hat{\mu}_{rss}$ if $r \to \infty$ using the following lemma.

Lemma 1: under assumed regularity conditions as $r \to \infty$, then

$$\sqrt{r}\,(\hat{\mu}_{rss} - \mu) \xrightarrow[D]{} N\,(0, \upsilon^2)$$

where $\upsilon^2 = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2_{(i:n)}$ and the notation $\xrightarrow[D]{}$ means convergence in distribution.

For the case r =1 and $n \to \infty$, let $B_n^2 = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2_{(i:n)}$ which the variance of $\hat{\mu}_{rss}$ if r =1, and we will denoted $\widehat{\mu}_{rss}$ by. Then

$$\frac{\widehat{\mu}_{rss}-\mu}{B_n}=\frac{(1/n)\sum_{i=1}^{n}X_{(i:n)}-(1/n)\sum_{i=1}^{n}E(X_{(i:n)})}{(1/n)\sqrt{\sum_{i=1}^{n}\sigma_{(i:n)}^2}}=\frac{\sum_{i=1}^{n}X_{(i:n)}-\sum_{i=1}^{n}E(X_{(i:n)})}{\sqrt{\sum_{i=1}^{n}\sigma_{(i:n)}^2}}.$$

Thus

$$\frac{n}{\sqrt{\sum_{i=1}^{n}\sigma_{(i:n)}^2}}(\widehat{\mu}_{rss}-\mu)=\frac{\widehat{\mu}_{rss}-\mu}{B_n}=\frac{\sum_{i=1}^{n}X_{(i:n)}-\sum_{i=1}^{n}E(X_{(i:n)})}{\sqrt{\sum_{i=1}^{n}\sigma_{(i:n)}^2}}.$$

Denote $\lambda(n)=\dfrac{n}{\sqrt{\sum_{i=1}^{n}\sigma_{(i:n)}^2}}$, the following theorem will be used to establish the asymptotic property of $\widehat{\mu}_{rss}$

Theorem 2.1 If r =1 and $n\to\infty$, then

$$\lambda(n)(\widehat{\mu}_{rss}-\mu)\xrightarrow[D]{}N(0,1)$$

If the following condition is holding

$$\frac{\sum_{i=1}^{n}E\left|X_{(i:n)}-E(X_{(i:n)})\right|^3}{\left(\sum_{i=1}^{n}\sigma_{(i:n)}^2\right)^{3/2}}\to 0$$

Property: Since $\frac{1}{n}\sum_{i=1}^{n}\sigma_{(i:n)}^2>\frac{1}{n+1}\sum_{i=1}^{n+1}\sigma_{(i:n)}^2$, we have $\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sigma_{(i:n)}^2=0$; Therefore $\frac{\lambda(n)}{\sqrt{n}}\to\infty$ as $n\to\infty$.

Remark: In practical setting n is small, so it does not make sense to consider the asymptotic properties of $\widehat{\mu}_{rss}$ with one replication and large n. However, it makes sense to take large r. In addition, it is not easy to satisfy the condition of the above theorem; we need to know the underling distribution to check this condition.

2.2 Asymptotic interval estimation

First, consider

$$Z_1 = \frac{\sqrt{r}\,(\hat{\mu}_{rss} - \mu)}{\upsilon} \xrightarrow[D]{} N(0,1)$$

However υ depends on the variance of the $i-th$ order statistic from a set of size *n*, and this will depend on the underling distribution. This result does not immediately yield confidence intervals for μ, since the parameter υ is generally unknown. Another pivotal quantity can be defined by replacing υ by its empirical estimate $\hat{\upsilon}$ in the denominator of Z_1 and is given by

$$Z_2 = \frac{\sqrt{r}\,(\hat{\mu}_{rss} - \mu)}{\hat{\upsilon}} \xrightarrow[D]{} N(0,1)$$

where $\hat{\upsilon} = \sqrt{\frac{1}{n}\hat{\sigma}^2_{rss} - \frac{1}{n^2}\sum_{i=1}^{n}(\widetilde{X}_{(i)} - \hat{\mu}_{rss})^2}$, with $\hat{\sigma}^2_{rss} = \frac{1}{nr-1}\sum_{j=1}^{r}\sum_{i=1}^{n}(X_{(i:n)j} - \hat{\mu}_{rss})^2$ and

$\widetilde{X}_{(i)} = \frac{1}{r}\sum_{j=1}^{r} X_{(i:n)j}$ is an estimator for the population mean of the *i*-th order statistic.

Further, $\hat{\upsilon}$ is a consistent estimator of υ thus $Z_2 \xrightarrow[D]{} N(0,1)$. Therefore, for large r, it is appropriate to approximate the distribution of Z_2 by the standard normal distribution. An asymptotic $(1-\alpha)100\%$ confidence interval for μ is

$$\Pr\left\{\hat{\mu}_{rss} - z_{\alpha/2}\left(\hat{\upsilon}^2/2\right)^{1/2} \le \mu \le \hat{\mu}_{rss} + z_{\alpha/2}\left(\hat{\upsilon}^2/2\right)^{1/2}\right\} = 1-\alpha \text{ as } r \to \infty$$

where $z_{\alpha/2}$ is the $(1-\alpha)100$ percentile point of the standard normal distribution.

2.3 Asymptotic tests and power

In this section, use the result of Lemma 1, we propose the following test statistic for

$$H_o : \mu = \mu_o \text{ vs. } H_a : \mu \ne \mu_o \text{ or } \mu < \mu_o \text{ or } \mu > \mu_o$$

$$D_r = \frac{\left\{\sqrt{r}\,(\hat{\mu}_{rss} - \mu_o)\right\}^2}{\hat{\upsilon}^2},$$

where $\hat{\upsilon}^2 = \frac{1}{n}\hat{\sigma}^2_{rss} - \frac{1}{n^2}\sum_{i=1}^{n}(\widetilde{X}_{(i)} - \hat{\mu}_{rss})^2$. For large r and under the null hypothesis, D_r follows a χ^2-distribution with one degree of freedom, which provides the asymptotic critical values. The simulation results in Section 4 indicate that the test statistics, the asymptotic critical values are useful for r ≥15.

It is important to note that for a fixed alternative that is different from the null hypothesis the power of all three tests statistics will converge to one as $r \to \infty$. Hence, to explore the asymptotic power properties of D_r, we confine ourselves to a sequence of local alternative $\{K_r\}$. In the present work such a sequence is specified by

$$K_r : \mu = \mu_o + \frac{\delta}{\sqrt{r}} \tag{1}$$

where δ is a fixed real number. Obviously, μ approaches μ_o at a rate of $r^{-1/2}$. Stochastic convergence of $\hat{\mu}_{rss}$ to the parameter μ ensures that $\hat{\mu}_{rss} \xrightarrow[p]{} \mu$ under local alternatives as well, where the notation $\xrightarrow[p]{}$ means convergence in probability.

The following theorem, which we present without proof, characterizes the asymptotic powers of the three test statistics under local alternatives.

Theorem 2.2 Under local alternatives in (1) the following results hold:

1. $\sqrt{r}\,(\hat{\mu}_{rss} - \mu) \xrightarrow[D]{} N(\delta, \upsilon^2)$,
2. D_r has asymptotically a non-central χ^2-distribution with 1 degree of freedom and non-centrality parameter $\Delta = \frac{\delta^2}{\upsilon^2}$.

Hence, the power calculations of the proposed test statistic can be accomplished by using non-central χ^2-distribution. A simulation study is carried out in Section 4 to observe the applicability of the asymptotic results for moderate samples.

3. Improved Estimation Strategies

We are mainly interested here to improve the estimation of μ when it is suspected a priori that $\mu = \mu_o$, which could be any sort of a priori information about μ. In many instances the practitioners have some conjectures about the value of parameter μ based on the past experience or familiarity with the experiment under investigation. It is reasonable then to move the $\hat{\mu}_{rss}$ of μ close to μ_o. However, the information regarding the parameter μ very well regarded as uncertain prior information (UPI). We try resolve this uncertainty by considering the linear combinations estimator may be defined as

$$\hat{\mu}_{rss}^{S} = \pi\mu_o + (1-\pi)\,\hat{\mu}_{rss}\,, \tag{2}$$

in which we would choose the coefficient π to minimize the mean squared error (MSE). Further, π may be also defined as the degree of trust in the prior information μ_o. The value of $\pi \in [0, 1]$ may be assigned by the experimenter according to trust in the prior value μ_o. If $\pi = 0$, we would rely on the sample data entirely. Estimators constructed as linear (or, more precisely, convex) combinations of other estimators or guessed values as in (2), are called composite estimators. The composite estimator $\hat{\mu}_{rss}^{S}$ can be interpreted as shrinkage estimator (SE), as it moves the sample estimator $\hat{\mu}_{rss}$ toward μ_o.

We will demonstrate that $\hat{\mu}_{rss}^{S}$ will have a smaller MSE than $\hat{\mu}_{rss}$ near the restriction, that is μ_o. However, $\hat{\mu}_{rss}^{S}$ becomes considerably biased and inefficient when the restriction may not be judiciously justified. Thus, the performance of this shrinkage procedure depends upon the correctness of the uncertain prior information (UPI). As such, when the prior information is rather suspicious, it may be reasonable to construct a shrinkage pretest estimator (SPE) denoted by $\hat{\mu}_{rss}^{SP}$ which incorporates a preliminary test on μ_o. Thus, the estimator $\hat{\mu}_{rss}$ and $\hat{\mu}_{rss}^{SP}$ is selected depending upon the outcome of the preliminary test. If the prior information is tenable, one may use $\hat{\mu}_{rss}^{SP}$, while $\hat{\mu}_{rss}$ may be chosen otherwise.

Thus, we consider the shrinkage pretest estimator (SPE) which is defined as

$$\hat{\mu}_{rss}^{SP} = \hat{\mu}_{rss}\, I\,(D_r \geq c_\alpha) + [(1-\pi)\,\hat{\mu}_{rss} + \pi\,\mu_o\,]\, I\,(D_r < c_\alpha) \tag{3}$$

where D_r is the test statistic for the null hypothesis $H_o : \mu = \mu_o$, which is defined in the previous section, and I (A) is the indicator function of a set A. The critical value c_α converges to $\chi^2_{1,\alpha}$ as $r \to \infty$. Thus, the critical value c_α of D_r may be approximated by $\chi^2_{1,\alpha}$, the upper $100\alpha\%$ critical value of the χ^2-distribution with 1 degree of freedom. Further, $\hat{\mu}^{SP}_{rss}$ can be written in the following computationally more attractive form as follows:

$$\hat{\mu}^{SP}_{rss} = \hat{\mu}_{rss} - \pi(\hat{\mu}_{rss} - \mu_o)\, I\,(D_r < c_\alpha). \tag{4}$$

Thus, the classical pretest estimator (PE) is readily obtained, by substituting $\pi = 1$ the in above relation,

$$\hat{\mu}^{P}_{rss} = \hat{\mu}_{rss} - (\hat{\mu}_{rss} - \mu_o)\, I\,(D_r < c_\alpha). \tag{5}$$

The PE is due to Bancroft (1944). The proposed SPE may be viewed as an improved PE which represents both $\hat{\mu}_{rss}$ and $\hat{\mu}^{P}_{rss}$ for $\pi = 0$ and $\pi = 1$ respectively.

3.1 Asymptotic bias and mean squared error

We will assess the performance of all these listed estimators using the mean squared error (MSE) criterion. The MSE of an estimator $\tilde{\mu}$ aimed at the target μ is defined as

$$MSE\,(\tilde{\mu}; \mu) = E\left\{(\tilde{\mu} - \mu)^2\right\},$$

where the notation E is with reference to hypothetical replications of the sampling process. The bias of an estimator $\tilde{\mu}$ of μ is denoted by $B\,(\tilde{\mu}; \mu)$, so

$$MSE\,(\tilde{\mu}; \mu) = \mathrm{var}(\tilde{\mu}) + B\,(\tilde{\mu}; \mu)^2.$$

We regard MSE as the measure of efficiency. However, MSE is usually not known and its value may depend on one or several parameters, sometime on the target itself. An estimator may be efficient for some values of the parameters but not for others. All this makes the search for the most efficient estimator challenging. To meet some of these challenges, we will express the MSE expressions for the listed estimators as a function of non-centrality parameter Δ for a smooth reading and comparison.

Further, noting that our results are based on the asymptotic normality of $\hat{\mu}_{rss}$, so we work with asymptotic bias (AB) and asymptotic MSE (AMSE). The asymptotic bias of an estimator $\tilde{\mu}$ of μ is defined as

$$AB(\tilde{\mu};\mu) = \lim_{r\to\infty} E\left\{\sqrt{r}\,(\tilde{\mu}-\mu)\right\}. \tag{6}$$

Thus, under local alternatives $AB(\hat{\mu}_{rss}^{S}) = -\pi\delta$. The expression of $AB(\hat{\mu}_{rss}^{SP})$ will be obtained with the aid of the following lemma.

Lemma 2: If the random variable Z is normally distributed with mean μ and variance 1, then

$$E\left\{Z\ I(0<Z^2<x)\right\} = \mu\,P(\chi^2_{3,\mu^2/2} < x),$$

where $\chi^2_{3,\mu^2/2}$ is random variable with the non-central chi-square distribution with 3 degrees of freedom and non-centrality parameter $\mu^2/2$. For proof of the lemma, readers are referred to Judge and Bock (1978).

Using Lemma 2, the following relation is established.

$$AB(\hat{\mu}_{rss}^{SP};\mu) = -\pi\delta G_3(\chi^2_{1,\alpha;\Delta}),$$

where $G_q(.;\Delta)$ is the cumulative distribution of a non-central χ^2-distribution with q degrees of freedom and non-centrality parameter Δ. Since $\lim_{\delta\to\infty} \delta\, G_3(\chi^2_{1,\alpha};\Delta) = 0$, it can be concluded that $\hat{\mu}_{rss}^{SP}$ is asymptotically unbiased, with respect to δ. On the other hand, $AB(\hat{\mu}_{rss}^{S};\mu)$ is an unbounded function of δ. Further,

$$AB(\hat{\mu}_{rss}^{P};\mu) = -\delta G_3(\chi^2_{1,\alpha};\Delta).$$

The $AB(\hat{\mu}_{rss}^{SP};\mu)$ and $AB(\hat{\mu}_{rss}^{P};\mu)$ are 0 at $\Delta = 0$. The bias functions of both pretest estimators increases to maximum as δ increases, then decreases towards 0 as Δ further increases. Ahmed et al. (2000) among others has pointed out that the estimators based on the pretest principle possess substantially smaller asymptotically mean square error (AMSE) than $\hat{\mu}_{rss}$ in a shrinkage neighborhood of UPI in the parameter space. For this reason, a sequence $\{K_r\}$ of local alternatives is considered for asymptotic analysis. Now, under the local alternative in (1) we present the expressions for the AMSE for the estimators under consideration. The AMSE of the first three estimators are

$$AMSE\,(\hat{\mu}_{rss}\,;\mu) = \upsilon^2\,;$$

$$AMSE\,(\hat{\mu}_{rss}^{S}\,;\mu) = \upsilon^2 - \upsilon^2\pi(2-\pi) + \upsilon^2\pi^2\Delta\,;$$

$$AMSE\,(\hat{\mu}_{rss}^{SP}\,;\mu) = \upsilon^2 - \upsilon^2\pi(2-\pi)G_3(\chi^2_{1,\alpha}\,;\Delta) + \upsilon^2\pi\,\Delta\{2G_3(\chi^2_{1,\alpha};\Delta) - (2-\pi)G_5(\chi^2_{1,\alpha};\Delta)\}.$$

The expression of $AMSE\,(\hat{\mu}_{rss}^{SP}\,;\mu)$ is readily obtained with the use of the following lemma.

Lemma 3: If the random variable Z is normally distributed with mean μ and variance 1, then

$$E\left\{Z^2\, I\,(0 < Z^2 < x)\right\} = P\,(\chi^2_{3,\,\mu^2/2} < x) + \mu^2\, P\,(\chi^2_{5,\,\mu^2/2} < x)\,.$$

The proof of this lemma can be found in Judge and Bock (1978).

The $AMSE\,(\hat{\mu}_{rss}^{S}\,;\mu)$ is a straight line in as a function of Δ which intersects the $AMSE\,(\hat{\mu}_{rss}\,;\mu)$ at $\Delta = (2\,\text{-}\pi)\,/\,\pi$. Under the null hypothesis the AMSE of $\hat{\mu}_{rss}^{S}$ is less than the AMSE of $\hat{\mu}_{rss}$. Specifically, $AMSE\,(\hat{\mu}_{rss}^{S}\,;\mu) \le AMSE\,(\hat{\mu}_{rss}\,;\mu)$ whenever $\Delta \in [0,(2-\pi)/\pi]$. On the other hand, $AMSE\,(\hat{\mu}_{rss}^{SP}\,;\mu) \ge AMSE\,(\hat{\mu}_{rss}\,;\mu)$ if

$$\Delta \ge\ (2-\pi)G_3(\chi^2_{1,\alpha}\,;\Delta) + \{2G_3(\chi^2_{1,\alpha};\Delta) - (2-\pi)G_5(\chi^2_{1,\alpha};\Delta)\}^{-1} \tag{7}$$

Alternatively, $AMSE\,(\hat{\mu}_{rss}^{SP}\,;\mu)$ performs better than $\hat{\mu}_{rss}$ if

$$\Delta <\ (2-\pi)G_3(\chi^2_{1,\alpha}\,;\Delta) + \{2G_3(\chi^2_{1,\alpha};\Delta) - (2-\pi)G_5(\chi^2_{1,\alpha};\Delta)\}^{-1}.$$

We noticed that the AMSE of pretest estimators are function of α, the level of the statistical significance. As α approaches, one $AMSE\,(\hat{\mu}_{rss}^{SP}\,;\mu)$ tends to $AMSE\,(\hat{\mu}_{rss}\,;\mu)$. In terms of Δ, when Δ increases and tends to infinity, the $AMSE\,(\hat{\mu}_{rss}^{SP}\,;\mu)$ approaches the $AMSE\,(\hat{\mu}_{rss}\,;\mu)$. Broadly speaking, for larger values of Δ, the value of the $AMSE\,(\hat{\mu}_{rss}^{SP}\,;\mu)$ increases, reaches its maximum after crossing the

$AMSE\,(\hat{\mu}_{rss}\,;\mu)$ and then monotonically decreases and approaches the $AMSE\,(\hat{\mu}_{rss}\,;\mu)$.

The AMSE of $AMSE\,(\hat{\mu}_{rss}^{P}\,;\mu)$ is defined as follows

$$AMSE\,(\hat{\mu}_{rss}^{P}\,;\mu) = \upsilon^2 + \upsilon^2\Delta\{2G_3(\chi^2_{1,\alpha};\Delta) - G_5(\chi^2_{1,\alpha};\Delta)\} - \upsilon^2 G_3(\chi^2_{1,\alpha};\Delta)\,,$$

and

$AMSE\,(\hat{\mu}_{rss}^{P}\,;\mu) \geq AMSE\,(\hat{\mu}_{rss}\,;\mu)$ if

$$\Delta \geq \; G_3(\chi^2_{1,\alpha};\Delta) + \{2G_3(\chi^2_{1,\alpha};\Delta) - G_5(\chi^2_{1,\alpha};\Delta)\}^{-1}\,. \tag{8}$$

By comparing the right hand side of equation (7) to the right hand side of (8), we noticed that the range of the parameter space in (8) is smaller than that in (7).

The risk difference

$$\begin{aligned} AMSE\,(\hat{\mu}_{rss}^{P}\,;\mu) - AMSE\,(\hat{\mu}_{rss}^{SP}\,;\mu) &= \Delta\upsilon^2\{2(1-\pi)G_3(\chi^2_{1,\alpha};\Delta) - (1-\pi)^2 G_5(\chi^2_{1,\alpha};\Delta\} \\ &\quad - \upsilon^2(1-\pi)^2 G_3(\chi^2_{1,\alpha};\Delta) \end{aligned}$$

suggests that $AMSE\,(\hat{\mu}_{rss}^{P}\,;\mu) \leq AMSE\,(\hat{\mu}_{rss}^{SP}\,;\mu)$ as according

$$\Delta \quad \leq (1-\pi)G_3(\chi^2_{1,\alpha};\Delta)\{2G_3(\chi^2_{1,\alpha};\Delta) - (1-\pi)G_5(\chi^2_{1,\alpha};\Delta)\}^{-1}$$

Thus $\hat{\mu}_{rss}^{SP}$ outshines $\hat{\mu}_{rss}^{P}$

$$\Delta \quad > (1-\pi)G_3(\chi^2_{1,\alpha};\Delta)\{2G_3(\chi^2_{1,\alpha};\Delta) - (1-\pi)G_5(\chi^2_{1,\alpha};\Delta)\}^{-1}$$

Based on above analytical findings, it can be safely concluded that none of the four estimators performs better than the other three. However, at $\Delta = 0$, the shrinkage estimator will be best choice. In addition, both pretest estimators have smaller AMSE than that of $\hat{\mu}_{rss}$ when the null hypothesis is tenable.

4. A Simulation Study

We now conduct a Monte Carlo simulation study to provide empirical outcomes to the theory developed in the early sections of this paper. The objective of this simulation study is bi-fold:

1) to examine the null distribution of the test statistics D_r given in Section 2, and study the power of the test statistics,

2) to examine the behavior of the relative precisions of $\hat{\mu}_{rss}^{S}$, $\hat{\mu}_{rss}^{SP}$ and $\hat{\mu}_{rss}^{P}$ with respect

to $\hat{\mu}_{rss}$ and $\hat{\mu}_{srs}$, were $\hat{\mu}_{srs}$ is the sample mean of simple random sample (SRS) of size nr.

The samples are drawn from two distributions namely, normal, and logistic. To study the null distributions of D_r , we estimated their sample critical values by simulating M = 10,000 samples from underling distributions. The values of the test statistic D_r are computed and arranged in ascending order. The cut off point at position (1-α)10,000 estimates the (1-α)$^{\text{th}}$100 percentile for α = 0.01, 0.05 and 0.10. The process is repeated for different choices of n=3, 5 and for r = 10, 15, 20, 25. Figures 1 and 2 present the differences between the critical values and there different asymptotic values: $\chi^2_{0.01,}(1) = 6.6349$, $\chi^2_{0.05,}(1) = 3.8415$ and $\chi^2_{0.1,}(1) = 2.7055$ respectively, using the logistic distribution for $\pi = 0.5$ and n=3 and 5. Therefore, our simulation results support the theoretical development of Section 2 and demonstrate that under the null hypothesis $\mathrm{H_o}$, the test statistic D_r is approximately distributed as $\chi^2_{(1)}$ for moderate samples. Note that using the same values of π and n, the normal distribution exhibit similar results that we get using the logistic distribution.

The powers of each test were estimated by proportions of trials resulting from rejecting the null hypothesis. Table 1 presents the powers of the tests at the selected values of the parameters with $\alpha = 0.01$, 0.05 and 0.1 for the underling normal and logistic distributions, for n=3 and r=25. Very close power results were obtained when the values of the test statistics D_r were compared to their asymptotic critical value at α = 0.05. Similar results are obtained for α = 0.01 and 0.10.

Finally, another Monte Carlo simulation study was performed in an effort to investigate the relative performance of the $\hat{\mu}^S_{rss}$, $\hat{\mu}^{SP}_{rss}$ and $\hat{\mu}^P_{rss}$ to $\hat{\mu}_{rss}$ and $\hat{\mu}_{srs}$ .We considered $\mathrm{H_o}: \mu = \mu_\mathrm{o}$ against $\mathrm{H_a}: \mu = \mu_\mathrm{o} + \omega$ here ω is a shift real number in the neighborhood domain of μ . Using M = 10,000 realizations of for r = 25 and n = 3, 5. Figures 3- 26 present the simulated relative precisions (SRP) of the estimators for various values of $\Delta = \omega\sqrt{r} / \upsilon^2$, where

$$SRP(\hat{\mu}_{srs}, \tilde{\mu}) = \frac{SMSE(\hat{\mu}_{srs})}{SMSE(\tilde{\mu})}$$

is the simulated precision of $\tilde{\mu}$ relative $\hat{\mu}_{srs}$. Further, $SMSE(\tilde{\mu})$ is the empirical mean squared errors of $\tilde{\mu}$. Finally, $SMSE(\hat{\mu}_{rss})$ and $SMSE(\hat{\mu}_{srs})$ are equal to standard error of $\hat{\mu}_{rss}$ and $\hat{\mu}_{srs}$ respectively.

Considering the results showing in Figures 3-26 the following remarks can be made

1. As we have seen on the above analytical finding, none of the four estimators dominated the other. However, if Δ between 0 and 3 and $\alpha \geq 0.05$ the $\hat{\mu}_{rss}^{S}$ (SE) dominated all other estimators. The gain in the SRP is substantial, e.g. if n=3, r =25, $\Delta = 0.8$ $\pi = 0.5$, $\alpha = 0.05$ and the underling distribution is the normal, the gain in the SRP of $\hat{\mu}_{rss}^{S}$, $\hat{\mu}_{rss}^{SP}$, $\hat{\mu}_{rss}^{P}$ and $\hat{\mu}_{rss}$ with respect to $\hat{\mu}_{srs}$ are 4.2, 2.3, 1.7 and 2 respectively. The major disadvantage of $\hat{\mu}_{rss}^{S}$ its SRP become the lowest (close to zero) if Δ getting larger.
2. The SRP of $\hat{\mu}_{rss}^{SP}$ started lower than $\hat{\mu}_{rss}^{S}$ and $\hat{\mu}_{rss}^{P}$ estimators if Δ between 0 and 0.3, but larger than $\hat{\mu}_{rss}$ and $\hat{\mu}_{srs}$ estimators. However, if Δ get larger the SRP of $\hat{\mu}_{rss}^{SP}$ will be more than $\hat{\mu}_{rss}^{S}$ and $\hat{\mu}_{rss}^{P}$ estimators e.g. if n=3, r =25, $\Delta = 0.1$ $\pi = 0.5$, $\alpha = 0.05$ and using the normal distribution, the gain in the SRP of $\hat{\mu}_{rss}^{S}$, $\hat{\mu}_{rss}^{SP}$, $\hat{\mu}_{rss}^{P}$ and $\hat{\mu}_{rss}$ with respect to $\hat{\mu}_{srs}$ are 6.9, 3.6, 4.6 and 2 respectively. Compare to 0.74, 1.3, 0.95 and 2 if $\Delta = 8.5$ and every thing else remain the same as above.
3. The $\hat{\mu}_{rss}^{P}$ enjoy the highest gain in the SRP if $\alpha = 0.01$ and Δ close to zero, e.g. if n=3, r =25, $\Delta = 0$ $\pi = 0.5$, $\alpha = 0.01$ and the underling distribution is the normal, the gain in the SRP of $\hat{\mu}_{rss}^{S}$, $\hat{\mu}_{rss}^{SP}$, $\hat{\mu}_{rss}^{P}$ and $\hat{\mu}_{rss}$ with respect to $\hat{\mu}_{srs}$ are 7.6, 5.9, 19.5 and 2 respectively. However, if Δ get larger the SRP of $\hat{\mu}_{rss}^{P}$ will go below $\hat{\mu}_{rss}$, but not below $\hat{\mu}_{srs}$ and start to get closer to $\hat{\mu}_{rss}$ if Δ get larger.
4. Increasing the set size from 3 to 5 will increase the SRP for the all estimators in this study.
5. If the underling distribution is the logistic, we get similar results that we have got for the normal distribution, but the SRPs are less for all estimators considered in this study.

5. An Example

To illustrate the computation of the new estimators introduced in the previous sections, we use the data collected by Muttlak and Al-Sabah (2003) from the Pepsi Cola production company in Al-Khobar, Saudi Arabia. The company has many production lines; the lines of filling bottles with soft drink were chosen as the population of interest, the variable of interest is the quantity of soft drink contained in the bottle.

They collected data sets using RSS method with perfect ranking for sets of size n=3. The RSS procedure for a set size of 3 bottles may be summarized as follows: Select 9 bottles at random from the production line and then divide them randomly into three sets of size three bottles each, rank the bottles in each set with respect to the level of soft drink by visual inspection. From the first ranked set the bottle with shortest empty part among the three bottles is selected and measured the empty part using an instrument designed for this purpose. From the second set of three bottles, the bottle with the second shortest empty part is selected for measurement. Finally from the last set the bottle with the biggest empty part is chosen for measurement. We used the first 30 replications of this process, i.e. r =30 to test the hypothesis of interest in this study.

We are interested in testing the null hypothesis against the alternative hypothesis that

$$H_o : \mu = 6.0 \text{ vs. } \mathrm{H_a} : \mu \neq 6.0$$

We calculate from the data $\hat{\mu}_{rss}$ = 5.9967 and $\hat{\mu}_{rss}^{S}$ = 5.999 with π=0.5, and $D_r = 0.0044176$. The critical value is $\chi^2_{0.05,1} = 3.841$, so the null hypothesis is not rejected at 0.05 level of significance. The other estimators of μ are $\hat{\mu}_{rss}^{P}$ and $\hat{\mu}_{rss}^{SP}$ with values equal to 5.999 and 6, respectively. Based on our earlier discussion we suggest using $\hat{\mu}_{rss}^{SP}$.

6. Conclusions

The goal of the paper was to succinctly develop the basic formulas, relationships and issues involved in estimation and testing of the population mean using replicated ranked set samples. This task is accomplished by establishing the asymptotic normal theory of the natural estimator. Further, a class of point estimators is introduced when there is uncertainty concerning the appropriate statistical model-estimator to use in representing data sampling process; we considered a basis for optimally combining estimation problems. In this regard, we constructed well-defined data-based shrinkage and pretest estimators. Asymptotic and moderate sample mean squared error results are established. Our analytical results are well supported by simulation results.

In this communication, we have developed our formulations and evaluations in a one-sample-problem context. The formulations of the present work can be extended to multiple estimation problems and some improved estimation can be adopted. Research on the statistical implications of these and other combining possibilities for a range of statistical models and distance measures is ongoing and continue to play in the arena of statistical inference.

Acknowledgements

The Authors are in dept for facilities provided by King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia. The work of Professor Ahmed is supported by a grant from the Natural Sciences and Engineering Council of Canada and a part of this investigation has been done while he was visiting the King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia.

References

Ahmed, S. E. (1999). Simultaneous estimation of survivor functions in exponential lifetime models. Journal of Statistical Computation and Simulation, 45, 689-696.

Ahmed, S. E. and Krzanowski, W. J. (2004). Biased estimation in a simple multivariate regression model, Computational Statistics and Data Analysis, 45, No. 4, 689-696.

Ahmed, S. E., Ahmed, Gupta, A. K. Khan, S. M. and Nicol, C. (2000). Simultaneous estimation of several interclass correlation coefficients. Ann. Inst. Statist. Math., 53, 354-369.

Bancroft, T. A. (1944). On biases in estimation due to use of preliminary test of significance. Ann. Math. Statist. 15, 190-204.

Bickel, P. J. and Doksum, K. A. (2001). Mathematical Statistics, Vol. 1, 2end edition. Prentice Hall, New Jersey.

Chen, Z., Bai, Z. and Sinha, B. K. (2004). Ranked Set Sampling. Theory and Applications, Springer-Verlag, New York.

Dell, T. R. and Clutter, J. L. (1972). Ranked set sampling theory with order statistics background. Biometrics, 28, 545-553.

Giles, J. A. and Giles, D. E. A. (1993). Pre-test estimation and testing in econometrics: recent developments. Journal of Economic Surveys, 7, 145-197.

Harter, H. L. and Balakrishnan, N. (1996). CRC Handbook of Tables for the use of Order Statistics in Estimation, CRC Press, Boca Raton.

Khan, B. U. and Ahmed, S. E. (2003). Improved estimation of coefficient vector in a regression model. Communication in Statistics: Simulation and Computation, 32, 747-769.

Kaur, A., Patil, G. P., Sinha, B. K. and Taillie, C. (1995). Ranked set sampling: an annotated bibliography. *Environmental and Ecological Statistics*, 2, 25-54.

Judge, G. G., and Bock, M. E. (1978). The statistical implication of pre-test and Stein-rule estimators in econometrics. Amsterdam: North-Holland.

Magnus, J. R. (1999). The traditional pretest estimator. Theory of Probability and its Applications, 44, 293-308.

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

Muttlak, H. A. (1997). Median ranked set sampling. Journal of Applied Statistical Sciences 6, 245-255.

Muttlak, H. A. and Al-Saleh (2000). Recent development in ranked set sampling, *Pakistan Journal of Statistics,* 16, 269-290.

Ohanti, K.(1999). MSE performance of a heterogeneous pre-test estimator. Statistics & Probability Letters, 41, 65-71.

Patil, G. P., Sinha, A. K. and Taillie, C. (1994). Ranked set sampling. *A Handbook of Statistics*, vol. 12, pp167-200.

Reif, J. and VIcek, K. (2002). Optimal pre-test estimators in regression. Journal of Econometrics, 110, 91-102.

Schäfer, J. and Strimmer, K. (2005) A Shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics, statistical applications in genetics and molecular biology 4, Article 32.

Takahasi, K. and Wakimoto, K. (1968). On the unbiased estimates of the population mean based on the sample stratified by means of ordering. Annals of Institute of Statistical Mathematics 20, 1-31.

Thompson, J. R. (1968). Some shrinkage techniques for estimating the mean. Journal of American Statistical Association 63, 113-122.

Table 1 The power of the test statistic D_r for different values of the noncentral parameter with n=3 and r=15 for α =0.01, 0.5 and 0.1 for the normal and logistic distributions

	Normal			Logistic		
W	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
0	0.0123	0.0514	0.1070	0.0123	0.0506	0.1066
0.025	0.0160	0.0627	0.1214	0.014	0.0564	0.1135
0.050	0.0273	0.0968	0.1641	0.0144	0.0609	0.1278
0.075	0.0471	0.1466	0.2274	0.0205	0.0821	0.1452
0.1	0.0833	0.2224	0.3218	0.0281	0.1013	0.1781
0.2	0.3489	0.5988	0.7151	0.0908	0.2478	0.3635
0.3	0.7084	0.8888	0.9440	0.2286	0.4519	0.5832
0.5	0.9925	0.9994	0.9997	0.6153	0.8235	0.9019
0.7	1.0000	1.0000	1.0000	0.9025	0.9753	0.9887
0.8	1.0000	1.0000	1.0000	0.9600	0.9912	0.9982
0.9	1.0000	1.0000	1.0000	0.9874	0.9989	0.9997
1.0	1.0000	1.0000	1.0000	0.9971	0.9998	1.0000

Figure 1. The difference between the critical values and there different asymptotic values for the Logistic distribution with n=3 and $\pi = 0.5$ for different values of r and α

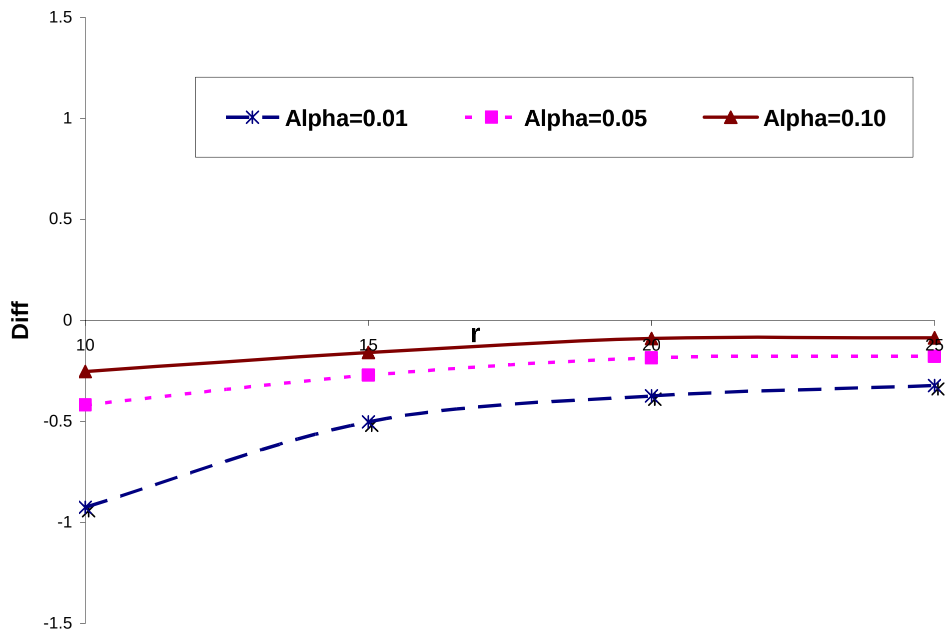
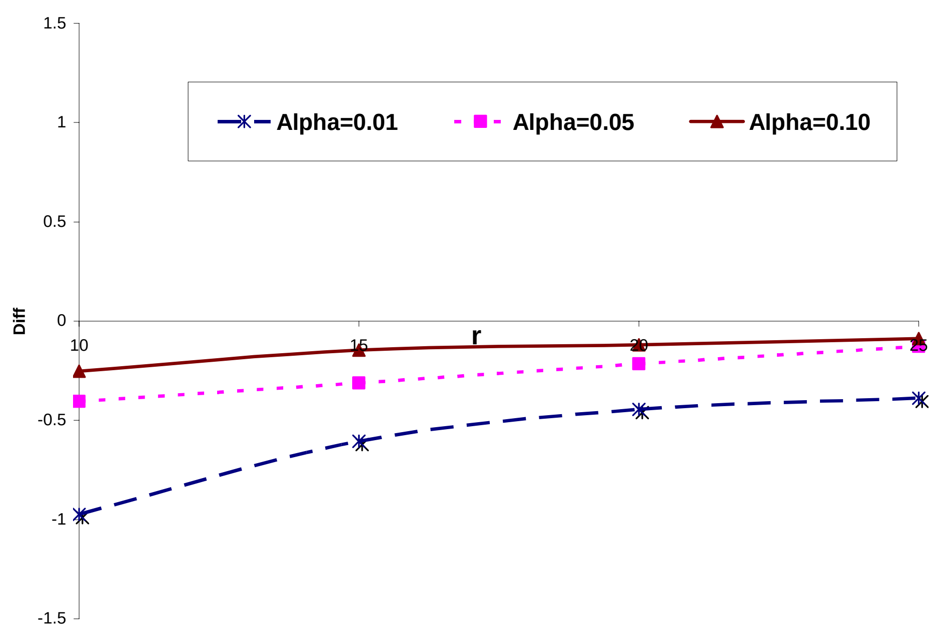


Figure 2. The difference between the critical values and there different asymptotic values for the Logistic distribution with n=5 and $\pi = 0.5$ for different values of r and α



Figures 3- 8. The simulated relative precisions (SRP) of $\hat{\mu}^{S}_{rss}$, $\hat{\mu}^{SP}_{rss}$ and $\hat{\mu}^{P}_{rss}$ with respect to $\hat{\mu}_{srs}$ denoted by SRPS, SRPSP and SRPP respectively using $\pi = 0.5$, n =3, r =25, and the relative precision of $\hat{\mu}_{rss}$ (RPR) using the normal distribution.

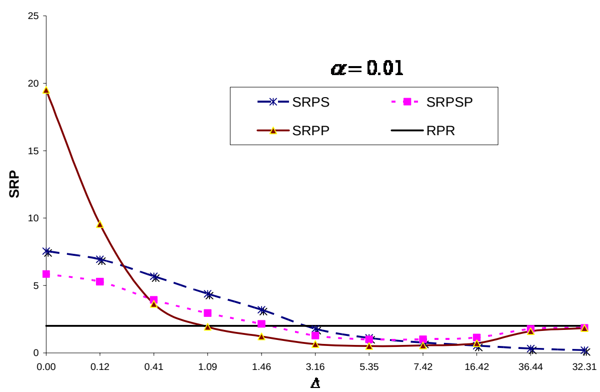


Figure 3

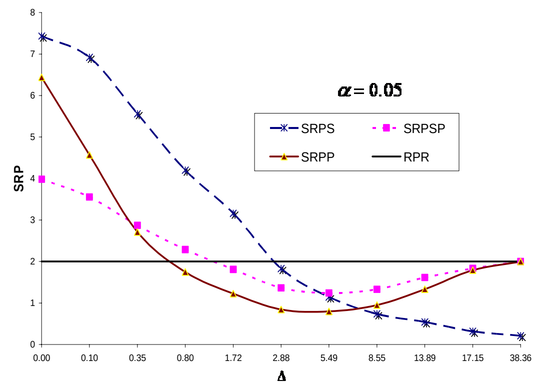


Figure 4

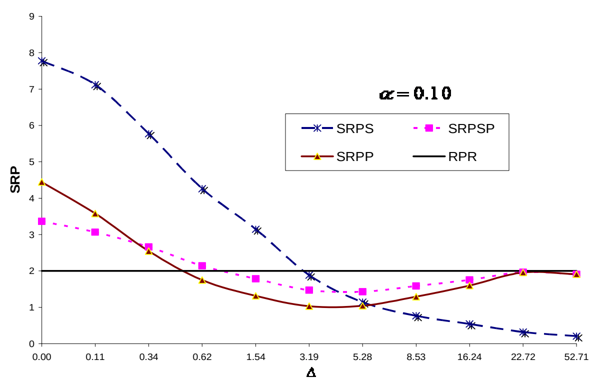


Figure 5

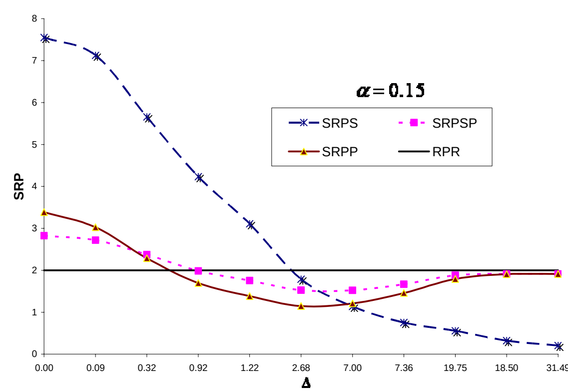


Figure 6

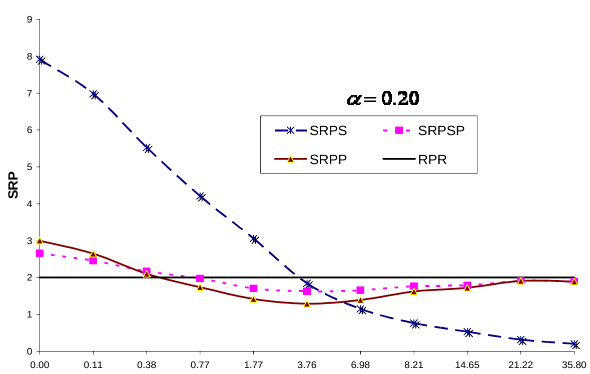


Figure 7

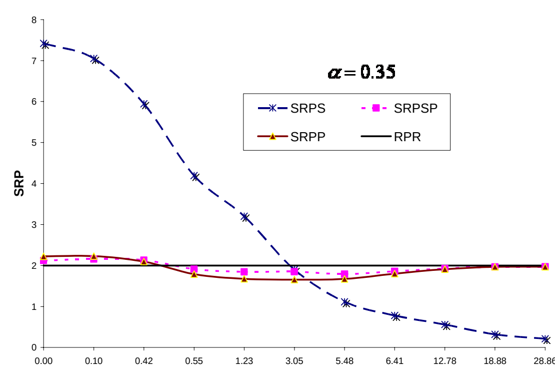


Figure 8

Figures 9- 14 The simulated relative precisions (SRP) of $\hat{\mu}^{S}_{rss}$, $\hat{\mu}^{SP}_{rss}$ and $\hat{\mu}^{P}_{rss}$ with respect to $\hat{\mu}_{srs}$ denoted by SRPS, SRPSP and SRPP respectively using $\pi = 0.5$, n =5, r =25, and the relative precision of $\hat{\mu}_{rss}$ (RPR) using the normal distribution.

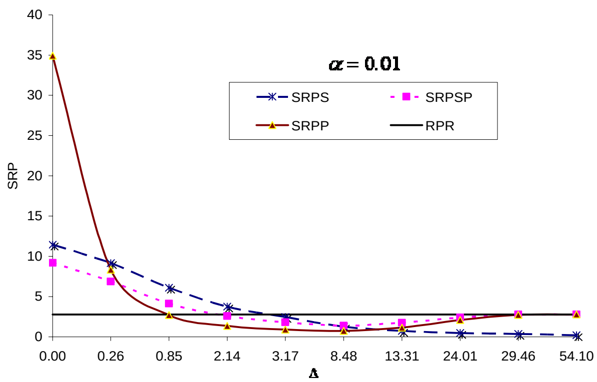


Figure 9

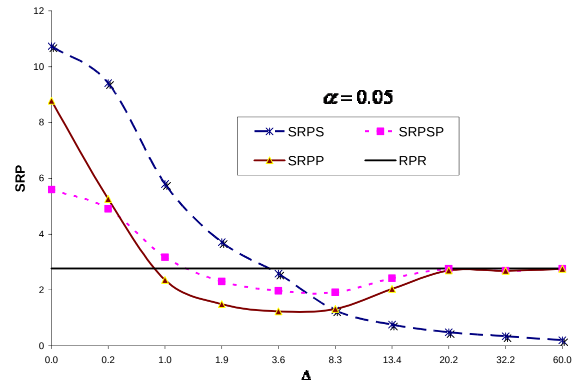


Figure 10

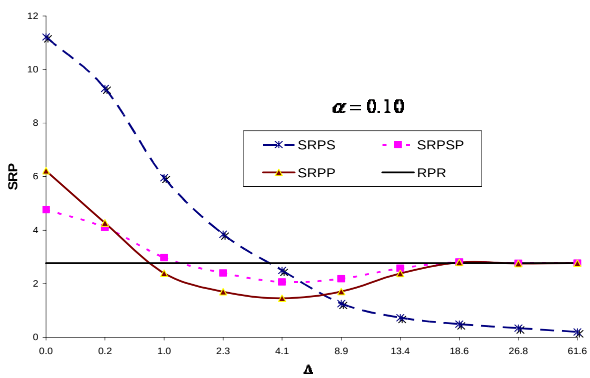


Figure 11

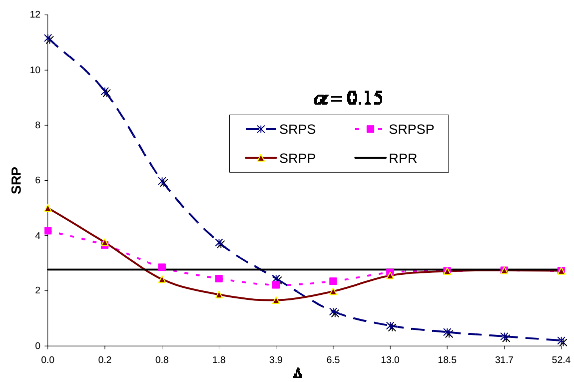


Figure 12

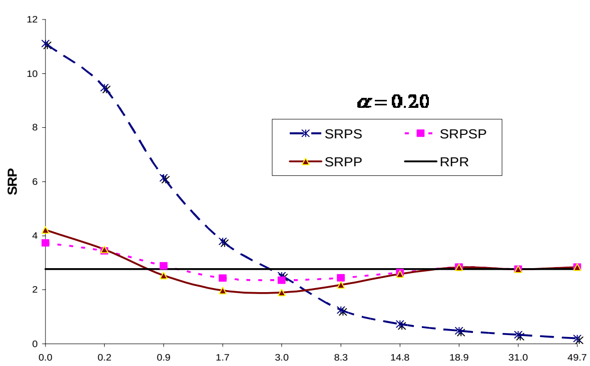


Figure 13

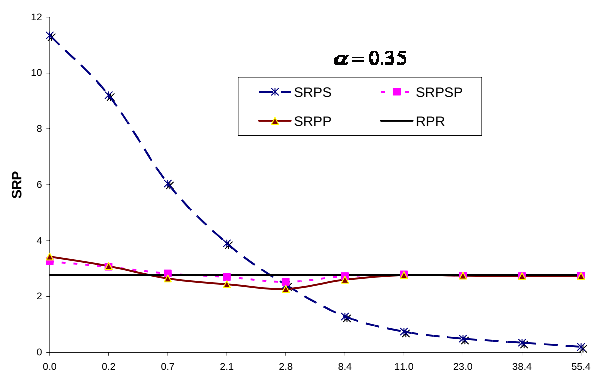


Figure 14

Figures 15- 20. The simulated relative precisions (SRP) of $\hat{\mu}_{rss}^{S}$, $\hat{\mu}_{rss}^{SP}$ and $\hat{\mu}_{rss}^{P}$ with respect to $\hat{\mu}_{srs}$ denoted by SRPS, SRPSP and SRPP respectively using $\pi = 0.5$, n =3, r =25, and the relative precision of $\hat{\mu}_{rss}$ (RPR) using the logistic distribution.

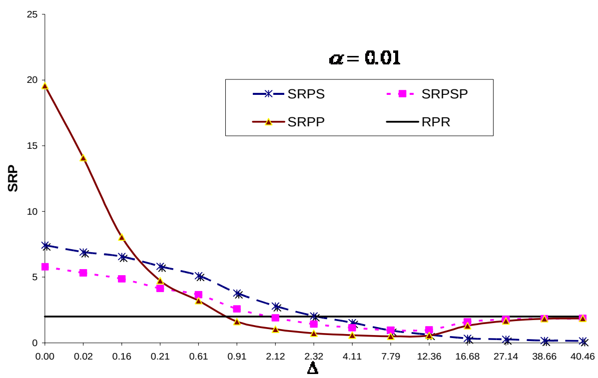


Figure 15

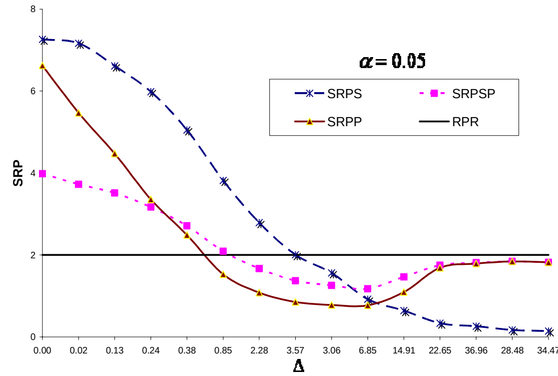


Figure 16

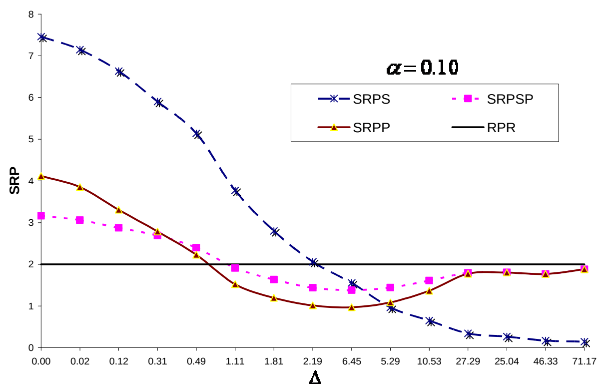


Figure 17

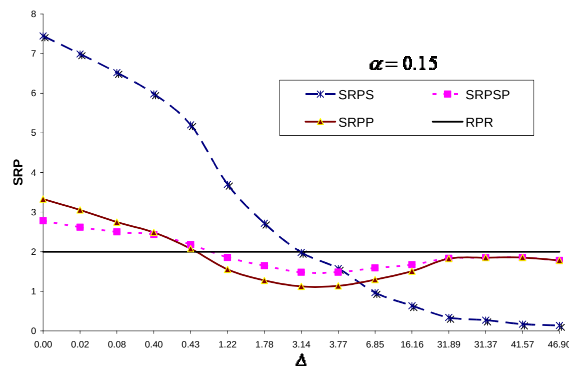


Figure 18

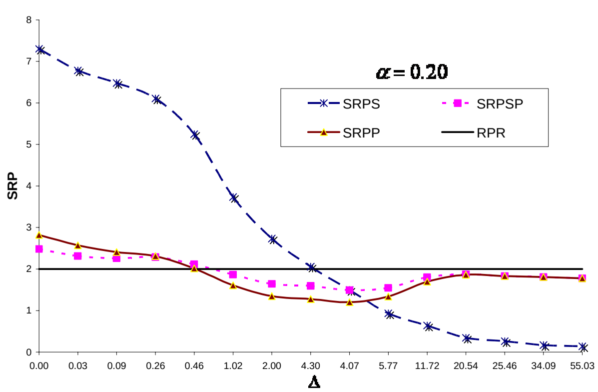


Figure 19

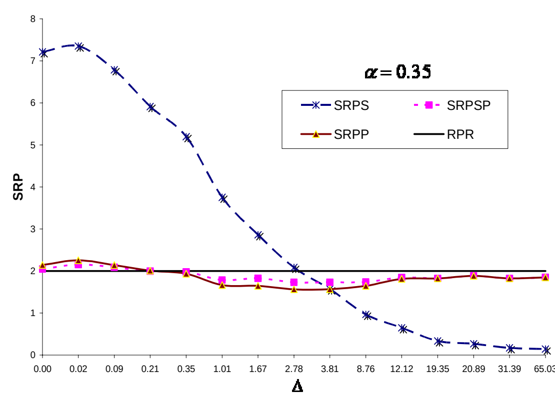


Figure 20

Figures 21- 26 The simulated relative precisions (SRP) of $\hat{\mu}_{rss}^{S}$, $\hat{\mu}_{rss}^{SP}$ and $\hat{\mu}_{rss}^{P}$ with respect to $\hat{\mu}_{srs}$ denoted by SRPS, SRPSP and SRPP respectively using $\pi = 0.5$, n =5, r =25, and the relative precision of $\hat{\mu}_{rss}$ (RPR) using the logistic distribution.

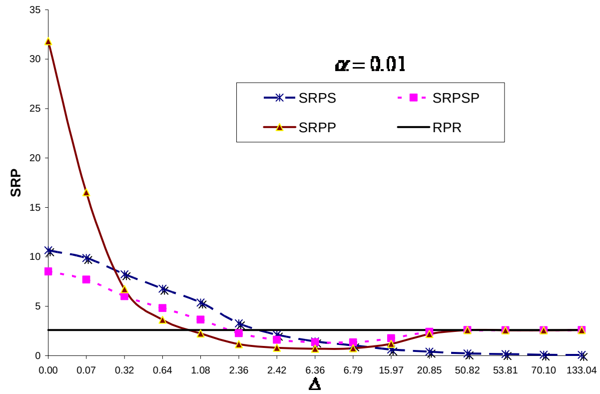


Figure 21

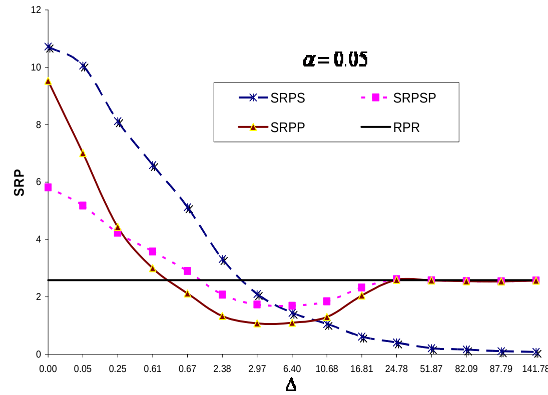


Figure 22

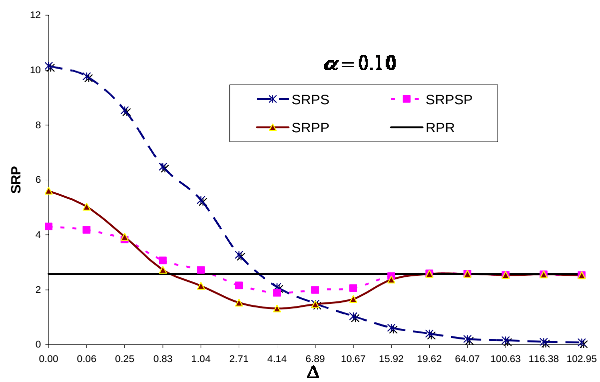


Figure 23

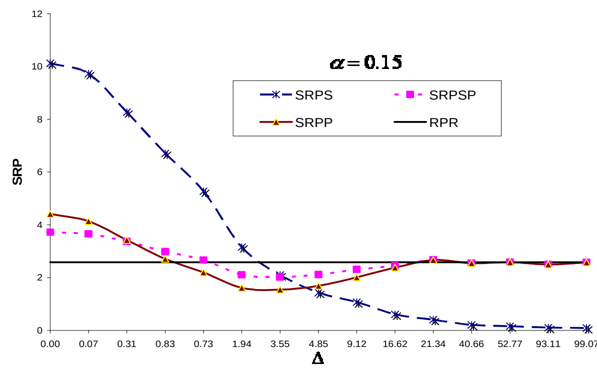


Figure 24

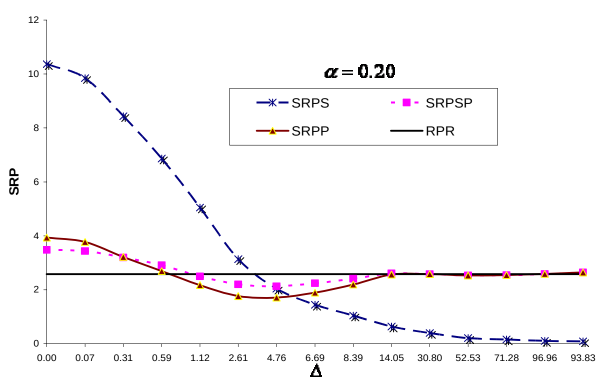


Figure 25

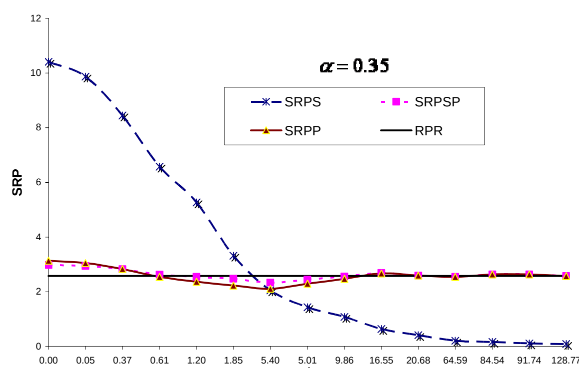


Figure 26