Journal of Reliability and Statistical Studies; ISSN (Print): 0974-8024, (Online):2229-5666,
Vol. 5, Issue 1 (2012): 43-54

LINEAR ALLOCATION MODELS FOR SYMMETRIC DISTRIBUTION IN RANKED SET SAMPLING

Neeraj Tiwari and Girja Shankar Pandey

Department of Statistics, Kumaun University, S.S.J. Campus Almora,
Uttarakhand, 263601, India
E Mail:pandeygirja@yahoo.co.in
(Received November 22, 2011)

Abstract

Ranked Set Sampling (RSS) is one method to potentially increase precision and reduce cost by using simple judgment or qualitative information. For symmetric distributions, an optimal allocation model was suggested by Kaur et al. (1995) (for simplicity in notation we call it by KPT model). This allocation model measures either only mid or extreme rank orders. This results in an estimator, which is not sufficient and hence unreliable in most of the situations, although it is more precise then Neyman's allocation.

In this paper, we have proposed a Linear allocation model for two classes of symmetric distributions. These two classes of symmetric distribution are mound shaped and U-shaped, depending upon the plots of the variances of the order statistics against the rank order. The proposed allocation model is opposite to the Neyman allocation model and has an advantage over KPT model in the sense that measurements are made upon each rank orders.

Keywords: Ranked Set Sampling, Relative Precision, Neyman's allocation, KPT Model, Ordered Statistics.

1. Introduction

Researchers, mainly those engaged in field and laboratory work, want optimum precision against low cost. Kaur et al. (1994) first gave the optimum allocation model for unequal RSS when the underlying distribution is skewed. In a similar way, Kaur et al. (1995) also proposed an optimal allocation model for symmetric distributions. Their allocation outperforms both equal allocations and Neyman allocations in terms of the precision of the estimator of population mean. They also examined the effect of population kutosis upon the precision of the estimator for their model. However, in the case of symmetric distributions, the gains due to the Neyman's allocation are marginal.

For symmetric distributions, the performance of the Neyman allocation remains very close to that of equal allocation. Yanagawa and Chen (1980) suggested a minimum variance liner unbiased median-mean estimator of population mean for a family of symmetric distribution. Shirahata (1982) examined more general procedures that are unbiased for symmetric distributions. For symmetric distributions, an optimal allocation model was suggested by Kaur et al. (1995) (for simplicity in notation we call it by KPT model). This allocation model measures either only mid or extreme rank orders. This results in an estimator, which is not sufficient and hence unreliable in most of the situations, although it is more precise then Neyman's allocation. Kaur et al.

(1997) derived unequal allocation models for ranked set sampling with skew distributions. Yu et al. (1999) studied some unbiased estimate of σ^2 in the parametric case of a normal population. Tiwari and Kvam (2001) proposed unbiased estimator for σ^2 for location-scale families of symmetric distribution. MacEachern (2002) developed an unbiased estimator of the variance of a population based on RSS.

The general RSS scheme proposed by Wang et al. (2004) takes more than one units in a ranked set with select pre-specified ranks for the full measurement. Perron and Sinha (2004) showed that for more than one cycle, it is possible to construct a class of quadratic unbiased estimates of σ^2 in both balanced and unbalanced cases. Ahmed (2004) suggested some bootstrap techniques for estimation of variance under RSS. Sengupta and Mukhuti (2006) proposed some unbiased estimators of the variance of exponential distribution. Jemain and Al-omari (2006) suggested multistage median ranked set sampling (MMRSS) method for estimating the population mean. Frey (2007) demonstrated a new imperfect ranking model for ranked set sampling. Some variations of ranked set sampling studied by Jamain et al. (2008). Ozturk (2008) have proposed a inference in the presence of ranking error in ranked set sampling. Baklizi (2009) described empirical likelihood intervals for the population mean and quantiles based on balanced RSS. Liu et al. (2009) studied the problem of empirical likelihood for balanced ranked set sampled data. Estimation of population variance using ranked set sampling with auxiliary variable was studied by Hadhrami (2010).

In this paper, we have proposed a Linear allocation model for two classes of symmetric distributions. These two classes of symmetric distribution are mound shaped and U-shaped, depending upon the plots of the variances of the order statistics against the rank order. The details of these two classes of symmetric distributions are discussed in Section 3. The proposed Linear allocation model for both the classes of symmetric distributions overcomes the drawback of Neyman and KPT model. The proposed allocation model is opposite to the Neyman allocation model and has an advantage over KPT model in the sense that measurements are made upon each rank orders.

In Section 2, we discuss in brief the expressions of RP of estimates of population mean for KPT model with respect to simple random sampling for mound shaped and U-shaped symmetric distributions. In Section 4, we discuss some examples from the two classes of symmetric distributions to demonstrate the utility of the proposed procedure.

2. Comparison of KPT Model with Simple Random Sampling

When underlying distribution is symmetric rather than skewed, the resulting optimal allocation strategy is precisely the opposite of the Neyman strategy. Kaur et al. (1995) derived the expressions of asymptotic RP for both the classes of symmetric distributions. In mound shaped symmetric distributions, they ignored the rank orders with large variances and measured only the rank orders having the smallest variances. This becomes the optimal allocation for finding the optimal variance of the best linear unbiased estimator of population mean μ. Their optimal variance of the estimator for large n is

$$\sigma^2_{Mound}(KPT) = \frac{\sigma^2_{(1:k)}}{n} \quad . \tag{2.1}$$

After comparing the asymptotic variance in (2.1) with the variance of sample mean under SRS, the asymptotic RP is

$$RP_{Mound}(KPT) = \frac{\sigma^2}{\sigma^2_{(1:k)}} . \tag{2.2}$$

While in the case of U-shaped symmetric distributions Kaur et al. (1995) derived the asymptotic variance of the best linear unbiased estimator of population mean μ, and is

$$\sigma^2_U(KPT) = \begin{cases} \dfrac{\sigma^2_{\left(\frac{k}{2}:k\right)}}{n}, & if\ k\ is\ even \\ \dfrac{\sigma^2_{\left(\frac{k+1}{2}:k\right)}}{n}, & if\ k\ is\ odd \end{cases} \quad . \tag{2.3}$$

The asymptotic RP compared with SRS is

$$RP_U(KPT) = \begin{cases} \dfrac{\sigma^2}{\sigma^2_{\left(\frac{k}{2}:k\right)}}, & if\ k\ is\ even \\ \dfrac{\sigma^2}{\sigma^2_{\left(\frac{k+1}{2}:k\right)}}, & if\ k\ is\ odd \end{cases} \quad . \tag{2.4}$$

3. The Linear Allocation Models for Symmetric Distributions

For symmetric distributions, an optimal allocation model (denoted by KPT model) was suggested by Kaur et al. (1995). However, as discussed in Section 2, for finding the estimates of the population mean, the estimates suggested by Kaur et al. (1995) are not reliable. Making use of the fact that in symmetric distributions the optimal allocation strategy is precisely the opposite of the Neyman strategy, we propose a simple and systematic approach, which measures more heavily those rank orders having the smallest variances and the number of measurements on each rank order is also an integer. The proposed approach can be easily implemented upon the practical situations and it performs better than SRS, RSS with equal allocation, Neyman allocation, and quite close to KPT model.

On the basis of graph plots of the variances of the order statistics against the rank orders, the symmetric distributions can be classified into two classes, namely- a) mound shaped class, i.e. the distributions for which $\sigma^2_{(i:k)}$ is increasing in i for $1 \le i \le M$ and $\sigma^2_{(i:k)}$ is decreasing in i for $M \le i \le k$, where $M = \frac{k+1}{2}$, is the

unique middle rank order when k is odd, and b) U-shaped class, i.e. for which $\sigma^2_{(i:k)}$ is decreasing in i for $1 \le i \le M$ and $\sigma^2_{(i:k)}$ is increasing in i for $M \le i \le k$.
For symmetric distributions, we have

$$\mu = \begin{cases} \frac{1}{2}\left(\mu_{(i:k)} + \mu_{(k-i+1:k)}\right), & for\ 1 \le i < M \\ \mu_{(i:k)}, & for\ i = M\ \&\ k\ is\ odd \end{cases} \tag{3.1}$$

and

$$\sigma^2_{(i:k)} = \sigma^2_{(k-i+1:k)}\ ,\ for\ 1 \le i \le M \tag{3.2}$$

In this proposed allocation model the largest order statistic allocated minimum time and subsequent order statistics are allocated in linearly increasing pattern where the linear terms may be assumed as per the requirement. Thus in our proposed model, we have two cases:

(a) For mound shaped symmetric distribution:

$$m_i = m_{k-i+1} = \begin{cases} \left(\frac{k+2}{2} - i\right)a + b, & for\ 1 \le i < M\ \&\ k\ is\ even \\ \left(\frac{k+3}{2} - i\right)a + b, & for\ 1 \le i < M\ \&\ k\ is\ odd \\ a + b, & for\ i = M\ \&\ k\ is\ odd \end{cases} \tag{3.3}$$

Where a and b can take any integer values, positive or negative. The values of a and b depend on the set size k, satisfying the conditions $a \le k$ and $a + b = 1$.

In this case the number of units for measurement will be

$$n = \begin{cases} 2\left(\frac{ak}{2} + b\right) + \frac{a(k-2)}{2} + b + \frac{a(k-4)}{2} + b + \ldots + (2a+b) + (a+b), & if\ k\ is\ even \\ 2\left(\frac{a(k+1)}{2} + b + \frac{a(k-1)}{2} + b + \frac{a(k-3)}{2} + b + \ldots + (2a+b)\right) + (a+b), & if\ k\ is\ odd \end{cases}$$

$$= \begin{cases} 2\left((a+b)+(2a+b)+(3a+b)+\ldots+\left(\dfrac{a(k)}{2}+b\right) \right), if\ k\ is\ even \\ 2\left((a+b)+(2a+b)+\ldots+\left(\dfrac{a(k+1)}{2}+b\right)-(a+b) \right), if\ k\ is\ odd \end{cases}$$

$$= \begin{cases} 2\left(a+2a+3a+\ldots+\dfrac{ak}{2}+\left(\dfrac{k}{2}\right)b \right), \qquad if\ k\ is\ even \\ 2\left(a+2a+\ldots+a\left(\dfrac{k+1}{2}\right)+\left(\dfrac{k+1}{2}\right)b-(a+b) \right)+(a+b),\ if\ k\ is\ odd \end{cases}$$

or

$$n = \begin{cases} \dfrac{k(k+2)a}{4}+kb, \qquad if\ k\ is\ even \\ \left(\dfrac{(k+1)(k+3)-4}{4}\right)a+kb,\ \ if\ k\ is\ odd \end{cases} \tag{3.4}$$

and,

(b) For U-shaped symmetric distribution

$$m_i = m_{k-i+1} = \begin{cases} ai+b, \quad for\ 1 \le i < M\ \&\ for\ all\ k \\ Ma+b,\ for\ \ i=M\ \&\ k\ is\ odd \end{cases} \tag{3.5}$$

In this case the number of selected unit n for measurements is

$$n = \begin{cases} 2\left(\left(\dfrac{ak}{2}+b\right)+\dfrac{a(k-2)}{2}+b+\dfrac{a(k-4)}{2}b+\ldots+(2a+b)+(a+b)\right), \qquad if\ k\ is\ even \\ 2\left(\left(\dfrac{a(k+1)}{2}\right)+b+\left(\dfrac{a(k-1)}{2}\right)+b+\left(\dfrac{a(k-3)}{2}\right)+b+\ldots+(2a+b)\right)+Ma+b,\ if\ k\ is\ odd \end{cases}$$

$$= \begin{cases} 2\left(a+2a+\ldots+\dfrac{k}{2}a+\dfrac{k}{2}b \right), \qquad if\ k\ is\ even \\ 2\left(a+2a+\ldots+\dfrac{k-1}{2}a+\dfrac{k-1}{2}b \right)+Ma+b,\ if\ k\ is\ odd \end{cases}$$

$$= \begin{cases} 2\left(a\dfrac{k/2\left(k/2+1\right)}{2} + \dfrac{k}{2}b \right), & if\ k\, is\, even \\ 2\left(a\dfrac{k-1/2\left(k-1/2+1\right)}{2} + \dfrac{k-1}{2}b \right) + Ma + b, & if\ k\, is\, odd \end{cases}$$

or

$$n = \begin{cases} \dfrac{k(k+2)}{4}a + kb, & if\ k\, is\, even \\ a\left(\dfrac{k+1}{2}\right)^2 + kb, & if\ k\, is\, odd \end{cases} \tag{3.6}$$

In what follows we shall make use of the following Theorem 3.1 of Kaur et al. (1995) for finding the minimum variance of the proposed linear unbiased estimator of population mean.

Theorem 3.1: Let $X_1, X_2, ..., X_n$ are n independent random variables with a common mean μ and with variances $\sigma_1^2, \sigma_2^2, ..., \sigma_n^2$. The linear combination $C_1X_1 + C_2X_2 + ... + C_nX_n$, with $C_1 + C_2 + ... + C_n = 1$, that has the smallest variance and is obtained by taking C_i inversely proportional to σ_i^2. The resulting minimum variance is

$$\frac{1}{\dfrac{1}{\sigma_1^{\ 2}} + \dfrac{1}{\sigma_2^{\ 2}} + ... + \dfrac{1}{\sigma_n^{\ 2}}} \tag{3.7}$$

We shall now discuss in detail the properties of the estimates with derivations for both the classes of symmetric distributions.

3.1 Mound Shaped Symmetric Distribution

Denoting the measured units by $Y_{(i:k)j}$, i=1, 2, …,k, j=1, 2, …, m_i (m_i is given in (3.3)), an unbiased estimator of the population mean μ based on i^{th} and $(k-i+1)^{th}$ order statistics is given by

$$\hat{\mu}_{(i:k)} = \begin{cases} \frac{1}{2}\left(\frac{T_i + T_{(k-i+1)}}{\left(\frac{K+2}{2} - i\right)a + b}\right), & for \quad 1 \le i < M \quad \& \quad k \text{ is even} \\ \frac{1}{2}\left(\frac{T_i + T_{(k-i+1)}}{\left(\frac{K+3}{2} - i\right)a + b}\right), & for \quad 1 \le i < M \quad \& \quad k \text{ is odd} \\ \frac{T_M}{a+b}, & for \quad i = M \quad \& \quad k \text{ is odd} \end{cases} \qquad (3.8)$$

where $T_i = \sum_{j=1}^{m_i} Y_{(i:k)j}$.

or

$$RP_{Mound}(Linear) = \begin{cases} \dfrac{4\sigma^2 \sum_{i=1}^{k/2}\left(\frac{(k+2-2i)a+2b}{\sigma^2_{(i:k)}}\right)}{k(k+2)a+4kb}, & if\ k\ is\ even \\ \dfrac{4\sigma^2\left(\sum_{i=1}^{(k-1)/2}\left(\frac{(k+3-2i)a+2b}{\sigma^2_{(i:k)}}\right) + \frac{a+b}{\sigma^2_{(M:k)}}\right)}{((k+1)(k+3)-4)a+4kb}, & if\ k\ is\ odd \end{cases} \qquad (3.9)$$

3.2 U-Shaped Symmetric Distribution

For U-shaped symmetric distribution with allocation model (3.5), an unbiased estimator of the population mean μ based on i^{th} and $(k-i+1)^{th}$ order statistics is given by

$$\breve{\mu}_{(i:k)} = \begin{cases} \frac{1}{2}\left(\frac{T_i + T_{k-i+1}}{ai+b}\right), & for\ 1 \le i < M\ \&\ for\ all\ k \\ \frac{T_M}{Ma+b}, & for\ i = M\ \&\ k\ is\ odd \end{cases} \qquad (3.10)$$

or

$$RP_U(Linear)=\begin{cases}\dfrac{8\sigma^2}{k(k+2)a+4kb}\displaystyle\sum_{i=1}^{k/2}\frac{ai+b}{\sigma^2_{(i:k)}}, & if\ k\ is\ even\\ \dfrac{8\sigma^2}{(k+1)^2a+4kb}\displaystyle\sum_{i=1}^{(k-1)/2}\frac{(ai+b)}{\sigma^2_{(i:k)}}+\frac{Ma+b}{2\sigma^2_{(M:k)}}, & if\ k\ is\ odd\end{cases} \quad (3.11)$$

4. Examples

In this Section we compare the numerical values of RP_{eql}, RP_{Ney}, $RP_{Mound}(KPT)/RP_U(KPT)$ and $RP_{Mound}(Linear)/RP_U(Linear)$ for some symmetric distributions to demonstrate the utility of the proposed procedure. Under mound shaped distributions, we have considered the uniform distribution and under U-shaped distribution, normal and standard special distributions are considered. The three distributions uniform, normal and standard special, considered have zero mean and unit variance. For the values of variances of order statistics $\sigma^2_{(i:k)}$ we refer to Hastings et al. (1947) and Sarhan and Greenberg (1962). These values for set size k =2, 3, …,10, in the case of uniform, normal and standard special distributions.

The relative precisions RP_{eql}, RP_{Ney}, $RP_{Mound}(KPT)/RP_U(KPT)$ and $RP_{Mound}(Linear)/RP_U(Linear)$ are computed for the uniform, normal and standard special distributions for set size k =2, 3, …,10 in Tables 4.1, 4.2 and 4.3 respectively. From these tables it is seen that for k=2 all methods of allocation for symmetric distribution are equivalent, that is,

$$RP_{eql} = RP_{Ney} = RP_{Mound}(KPT)/RP_U(KPT) = RP_{Mound}(Linear)/RP_U(Linear)$$

The proposed model is better than equal and Neyman allocation models for set size k greater than 2. Moreover, the proposed allocation model is quite close to the KPT model. To clear the variations between the different RP’s, we have plotted the bar diagrams for all the distributions for set size k=2, 3, …,10. These graphs are shown in the Figures 4.1 to 4.3.

From these results, it is clear that the Linear allocation model may be considered a good allocation model for selecting the sample when the underlying population is symmetric.

K	RP_{eql}	RP_{Ney}	$RP_{Mound}(KPT)$	$RP_{Mound}(Linear)$
2	1.499987	1.499987	1.499987	1.499987
3	1.999995	2.009614	2.222225	2.160491
4	2.499985	2.525496	3.124949	2.951308
5	3.000024	3.045789	4.200001	3.60662
6	3.500028	3.569228	5.444481	4.58896
7	3.999990	4.095032	6.857190	5.229903
8	4.499978	4.622724	8.437264	6.34659
9	5.000020	5.151980	10.185170	6.986506
10	5.500000	5.682400	12.099980	8.20913

Table 4.1: Relative precisions RP_{eql}, RP_{Ney}, $RP_{Mound}(KPT)$ and $RP_{Mound}(Linear)$ **for Uniform (0, 1) for k = 2(1)10.**

K	RP_{eql}	RP_{Ney}	$RP_U(KPT)$	$RP_U(Linear)$
2	1.466942	1.466942	1.466942	1.466942
3	1.913747	1.918730	2.228804	2.081675
4	2.346948	2.361036	2.774269	2.650807
5	2.770176	2.796752	3.486341	3.253601
6	3.185669	3.227568	4.061532	3.82106
7	3.594922	3.654567	4.751793	4.407152
8	3.998987	4.078489	5.342274	4.971227
9	4.398550	4.499770	6.021760	5.550253
10	4.794490	4.919100	6.623100	6.113327

Table 4.2: Relative precisions RP_{eql}, RP_{Ney}, $RP_U(KPT)$ and $RP_U(Linear)$ **for Normal (0, 1) for k = 2(1)10.**

K	RP_{eql}	RP_{Ney}	$RP_U(KPT)$	$RP_U(Linear)$
2	1.400843	1.400843	1.400843	1.400843
3	1.752002	1.793318	2.926355	2.43732
4	2.072350	2.176563	3.616811	3.25601
5	2.371180	2.552366	4.959630	4.186925
6	2.653780	2.922187	5.765164	5.000188
7	2.923583	3.287129	7.018352	5.881072
8	3.182948	3.648068	7.883330	6.691833
9	3.433420	4.005460	9.087210	7.554441
10	3.684160	4.376560	10.279440	8.727263

Table 4.3 Relative precisions RP_{eql}, RP_{Ney}, $RP_U(KPT)$ and $RP_U(Linear)$ for Standard Special for k =2(1)10.

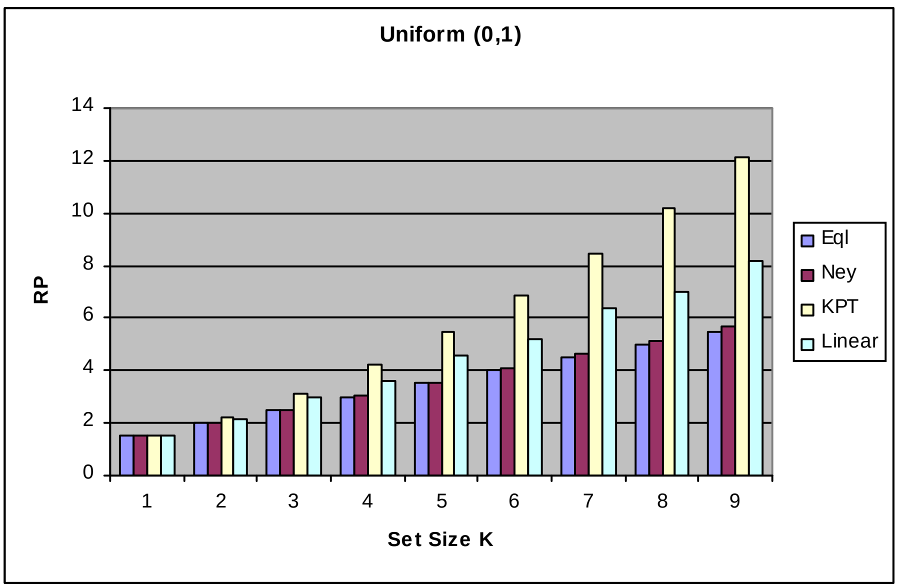

Figure 4.1 Bar diagrams for the RP's RP_{eql}, RP_{Ney}, $RP_{Mound}(KPT)$ and $RP_{Mound}(Linear)$ of Uniform (0, 1) for k = 2(1)10.

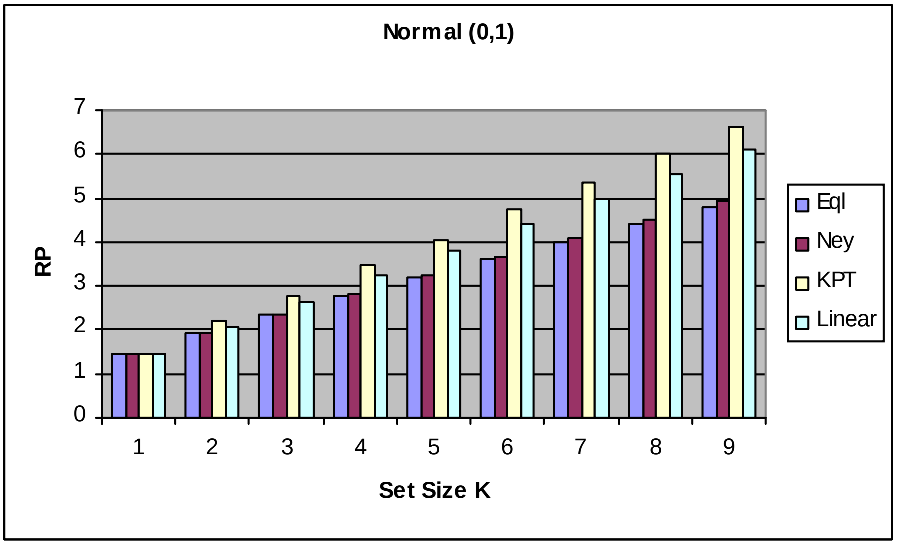

Figure 4.2 Bar diagrams for the RP's RP_{eql}, RP_{Ney}, $RP_U(KPT)$ and $RP_U(Linear)$ **of Normal (0, 1) for k = 2(1)10.**

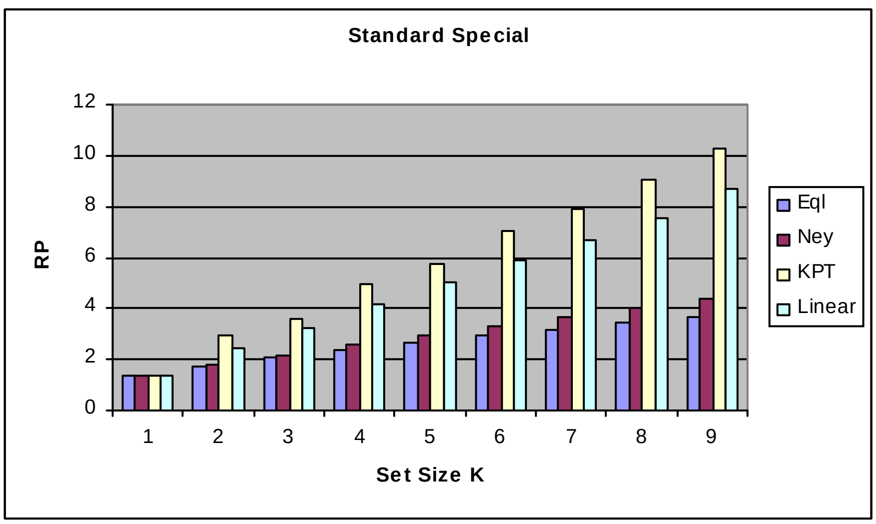

Figure 4.3 Bar diagrams for the RP's RP_{eql}, RP_{Ney}, $RP_U(KPT)$ and $RP_U(Linear)$ **of Standard Special for k = 2(1)10.**

References

1. Baklizi, A. (2009). Empirical likelihood intervals for the population mean and quantiles based on balanced ranked set sample, Stat Methods Appl., 18, p. 483-505.
2. Frey J.(2007). New imperfect ranking models for ranked set sampling, J. Stat Plan Inf, 137, p. 1433-1445.
3. Hastings, C. Jr., Mosteller, F., Tukey, J.W. and Winsor, C.P. (1947). Low moments for small samples: A comparative study of order statistics, The Annals of Mathematical Statistics 18(3), p. 413-426.
4. Jemain, A. A., Al-Omri, A. and Ibrahim, K. (2007). Multisatage median ranked set sampling for estimating the population mean, Journal of Mathematics and Statistics, 3(2), p. 58-64.
5. Jemain, A., A.A. Amer, A., Ibrahim, K. (2008). Some variations of ranked set sampling, Electronic Journal of Applied Statistics Analysis, 1, p. 1-18.
6. Kaur, A., Patil, G. P. and Taillie, C. (1994). Unequal allocation models for ranked set sampling with skew distributions. Technical Report 94-0930, Center for Statistical Ecology and Environmental Statistics, Pennsylvania State University, University Park.
7. Kaur, A., Patil, G. P. and Taillie, C. (1995). Ranked set sampling, an annotated bibliography. Environmental and Ecological Statistics, 2, p. 25-54.
8. Kaur, A., Patil, G. P. and Taillie, C. (1995). Optimal allocation models for symmetric distributions in ranked set sampling. Technical Report TR 95-0201, Center for Statistical Ecology and Environmental Statistics, Department of Statistics, Penn State University. University Park PA.
9. Kaur, A., Patil, G. P. and Taillie, C. (1997). Unequal allocation models for ranked set sampling with skew distributions, Biometrics, 53, p. 123-130
10. Ozturk, O. (2008). Inference in the presence of ranking error in ranked set sampling, Canadian Journal of Statistics, 36(4), 577.594.
11. .Sengupta, S., and Mukhuti S. (2006). Unbiased variance estimation in a simple exponential population using ranked set samples, Journal of Statistical Planning and Inference, 136, p. 1526-1553.
12. Shirahata, S. (1982). An extension of the ranked set sampling theory, Journal of Statistical Planning and Inference, 6, p. 65-72.
13. Wang, Y.G., Chen, Z., and Liu, J. (2004). General ranked set sampling with cost considerations, Biometrics, 60, p. 556-561.
14. Yanagawa, T. and Chen, S. (1980). The MG-procedure in Ranked Set Sampling, Journal of Statistical Planning and Inference, 4, p. 33-44.
15. Yanagawa, T. and Shirahata, S. (1976). Ranked Set Sampling Theory with Selective Probability Matrix, Australian Journal of Agricultural Research, 18, p. 45-52.