Statistical Inference Based on Partially Rank Ordered Set Samples

Omer Ozturk

The Ohio State University

Department of Statistics

June 20, 2011

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Ranked set sampling

- Select M units at random from a specified population.
- Rank these M units with some expert judgment without measuring them.
- Retain the smallest judged unit and return the others.
  Select the second M units and retain the second smallest unit judged.
- Continue to the process until M ordered units are measured.
- Note: These M ordered observations $X_{(1)i}, ..., X_{(M)i}$ are called a <u>cycle</u>.
- Note: Process repeated $i = 1, \cdots , n$ cycle to get Mn observations. These nM observations are called a <u>standard ranked set sample</u>.

Diagram

Let $M = 4$ and $n = 2$.

Cycle	Judgment Rank 1	2	3	4
1	$X_{(1)1}$	$X_{(2)1}$	$X_{(3)1}$	$X_{(4)1}$
	$X_{(1)1}$	$X_{(2)1}$	$X_{(3)1}$	$X_{(4)1}$
	$X_{(1)1}$	$X_{(2)1}$	$X_{(3)1}$	$X_{(4)1}$
	$X_{(1)1}$	$X_{(2)1}$	$X_{(3)1}$	$X_{(4)1}$
2	$X_{(1)2}$	$X_{(2)2}$	$X_{(3)2}$	$X_{(4)2}$
	$X_{(1)2}$	$X_{(2)2}$	$X_{(3)2}$	$X_{(4)2}$
	$X_{(1)2}$	$X_{(2)2}$	$X_{(3)2}$	$X_{(4)2}$
	$X_{(1)2}$	$X_{(2)2}$	$X_{(3)2}$	$X_{(4)2}$

$X_{(1)1}, \cdots, X_{(4)2}$ is called a ranked set sample.

Why ranked-set sampling?

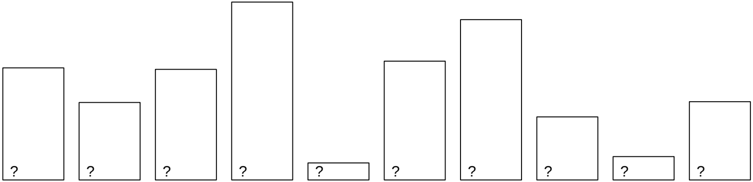
- Let X_i, $i = 1, \cdots, m$ be a SRS, and let $\bar{X}_{RSS}$ and $\bar{X}_{SRS}$ denote the sample averages based on RSS and SRS.
- It is easy to observe that

$$
\begin{aligned}
var(\bar{X}_{SRS}) &= \frac{1}{m^2} var(\sum_{i=1}^{m} X_i) = \frac{1}{m^2} var(\sum_{i=1}^{m} X_{(i)}) \\
&= \frac{1}{m^2} \left\{ \sum_{i=1}^{m} \sigma^2_{(i)} + \sum_{i \neq j} \sigma_{ij} \right\} = var(\bar{X}_{RSS}) + cov \\
var(\bar{X}_{SRS}) &\geq var(\bar{X}_{RSS})
\end{aligned}
$$

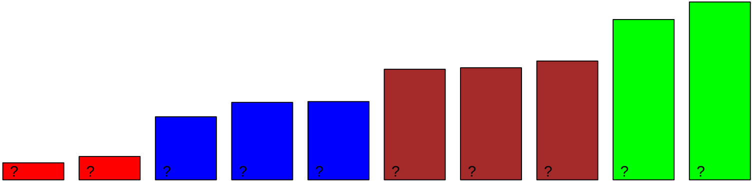
- Inequality becomes an equality when the ranking is completely random.
- This improved efficiency result holds for almost all statistical procedures based on RSS.

Problems?

- In standard ranked set sampling, the ranker must assign a rank to the unit selected for full measurement. Ties are not allowed.
- Even if the ranker has no confidence or little confidence to rank the units accurately, he/she is forced to produce a single rank for the selected unit.
- This leads to ranking error.
- Ranking error does not only reduce the efficiency, it may also provide invalid inference.
- We introduce an alternative sampling scheme that reduce the impact of ranking error when the rankers have little or no confidence to rank the units accurately.



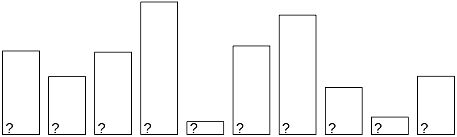
Unranked experimental units

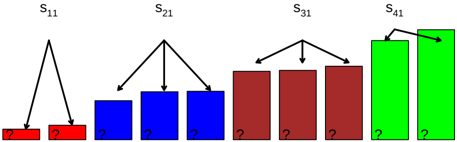


Judgment subsetting

- Select M units at random from a specified population.
- Rank these M units with some expert judgment without measurement by allowing ties among the units whenever their ranks can not be assigned with high confidence.
- The smallest ranked unit(s) is replaced in set $s_{1,j}$, second smallest unit(s) is replaced in set $s_{2,j}$ and so on until the d_j-th smallest unit(s) is replaced in set $s_{d_j,j}$.
- The subsets, $s_{1,j}, \cdots, s_{d_j,j}$, are partially ordered since any unit in $s_{h,j}$ has a smaller rank than any other unit in $s_{h',j}$ as long as $h < h'$.
- Let $D_j = \{s_{1,j}, \cdots, s_{d_j,j}\}$ be a set of the judgment subsets.
- This process is called judgment subsetting.

Judgment Subsetting :Example





- $D_1 = \{s_{1,1}, s_{2,1}, s_{3,1}, s_{4,1}\}$
- $s_{h,1} \leq s_{h',1},\ h < h'$

- Select a sample size N and set size M.
- Select N sets each of size M from the population of interest.
- Apply judgment subsetting process to each one of the N sets to create design parameters $\{D_j, d_j\}$, $j = 1, \cdots, N$
- Group these N sets into K different groups G_i, $i = 1, \cdots, K$,

$$G_i = \{D_{1,i}, \cdots, D_{n_i,i}\}, \quad \sum_{i=1}^{K} n_i = N,$$

  where $D_{h,i} = D_j$ for some j.
- In each group G_i, select a unit at random for full measurement from subset $s_{h,h,i}$ in set $D_{h,i}$, $h = 1, \cdots, n_i$, where $s_{h,r,i}$ is the h-th subset in set $D_{r,i}$ and group G_i
- Th fully measured units, $X_{[s_{h,h,i}]}, h = 1, \cdots, n_i, i = 1, \cdots, K$, are called partially rank ordered set (PROS) sample.

Definition

Sampling design will be called balanced if the sets G_i, $i = 1, \cdots, K$ satisfy the following condition

$$\cup_{i=1}^{K} (\cup_{h=1}^{n_i} s_{h,h,i}) = L\mathcal{M},$$

where $\mathcal{M}$ is the set of integers $1, \cdots, M$ and L is an integer number.

- This definition indicates that union of subsets $s_{h,h,i}$, $h = 1, \cdots, n_i$, $i = 1, \cdots, K$ can provide L sets of integer $1, \cdots, M$.
- Each one of the L sets can be considered as a cycle that contains all order statistics in a set of size M.
- Design G: the collection of G_i, $i = 1, \cdots, K$, is called design.

Design G: $M = 10$ and $K = 2$.

Group(i)	$S_{r,i}$	Judgment subsets	m_{hi}	$X_{[s_{h,h,i}]}$
1	$s_{1,1,1}, s_{2,1,1}$	$\{1,2,3,4,5\}, \{6,7,8,9,10\}$	5	$X_{[s_{1,1,1}]}$
	$s_{1,2,1}, s_{2,2,1}, s_{3,2,1}$	$\{1,2,3\}, \{4,5,6,7\}, \{8,9,10\}$	4	$X_{[s_{2,2,1}]}$
	$s_{1,3,1}, s_{2,3,1}, s_{3,3,1}$	$\{1,2,3\}, \{4,5\}, \{6,7,8,9,10\}$	5	$X_{[s_{3,3,1}]}$
2	$s_{1,1,2}, s_{2,1,2}, s_{3,1,2}$	$\{1,2,3\}, \{4,5,6,7\}, \{8,9,10\}$	3	$X_{[s_{1,1,2}]}$
	$s_{1,2,2}, s_{2,2,2}$	$\{1,2,3,4,5,6,7\}, \{8,9,10\}$	3	$X_{[s_{2,2,2}]}$

- This sample is balanced since the union of $s_{h,h,i}$, $h = 1, n_i$, $i = 1, 2$ (integers having the red colors) creates two cycles.
- $X_{[s_{h,h,i}]}$ is selected at random from the units in partially ordered set $s_{h,h,i}$.
- m_{hi} is the number of unranked units in judgment subset $s_{h,h,i}$.

Design G^*: If the within group sets have the same partition, the collection of G_i is called design G^*.

$N = 5,\ K = 2,\ M = 6$

Goup(i)	$S_{r,i}$	Judgment Subsets	m_{hi}	Obs
1	$\{s_{1,1,1}, s_{1,2,1}, s_{1,3,1}\}$	$\{1,2\}, \{3,4\}, \{5,6\}$	2	$X_{s_{1,1,1}}$
	$\{s_{2,1,1}, s_{2,2,1}, s_{2,3,1}\}$	$\{1,2\}, \{3,4\}, \{5,6\}$	2	$X_{s_{2,2,1}}$
	$\{s_{3,1,1}, s_{3,2,1}, s_{3,3,1}\}$	$\{1,2\}, \{3,4\}, \{5,6\}$	2	$X_{s_{3,3,1}}$
2	$\{s_{1,1,2}, s_{1,2,2}\}$	$\{1,2,3\}, \{4,5,6\}$	3	$X_{s_{1,1,2}}$
	$\{s_{2,1,2}, s_{2,2,2}\}$	$\{1,2,3\}, \{4,5,6\}$	3	$X_{s_{2,2,2}}$

Design G^{**}: If all sets $S_{r,i}$ have the same partition, the collection of G_i is called design G^{**}.

$$N = 4,\ K = 2,\ M = 6$$

Goup(i)	$S_{r,i}$	Judgment Subsets	m_{hi}	Obs
1	$\{s_{1,1,1}, s_{1,2,1}\}$	$\{1,2,3\},\{4,5,6\}$	3	$X_{s_{1,1,1}}$
	$\{s_{2,1,1}, s_{2,2,1}\}$	$\{1,2,3\},\{4,5,6\}$	3	$X_{s_{2,2,1}}$
2	$\{s_{1,1,2}, s_{1,2,2}\}$	$\{1,2,3\},\{4,5,6\}$	3	$X_{s_{1,1,2}}$
	$\{s_{2,1,2}, s_{2,2,2}\}$	$\{1,2,3\},\{4,5,6\}$	3	$X_{s_{2,2,2}}$

Some Observations

- Subsets $s_{h,h,i}$ are random subsets, and the number of subsets in each set is a random integer.
- Balanced design leads to robust inference against imperfect ranking.
- Construction of G may pose some logistical challenges in practice, but it produces the least amount of ranking error.
- Design G^* and G^{**} are easy to construct, but they may have slightly elevated ranking error.

Lemma

Let $X_{[s_{h,h,i}]}$ be an observation from a PROS sample from a population having a finite second moment. The probability density function (pdf), mean and variance of $X_{[s_{r,r,i}]}$ are given by

$$
\begin{aligned}
g_{[h;m_{h,i}]}(y) &= \frac{1}{m_{h,i}} \sum_{r \in s_{h,h,i}} f_{[r:M]}(y), \\
EX_{[h;m_{h,i}]} &= \gamma_{[h;m_{h,i}]} = \frac{1}{m_{h,i}} \sum_{r \in s_{h,h,i}} \mu_{[r:M]}, \\
var(X_{[r;m_{h,i}]}) &= \tau^2_{[h;m_{h,i}]} = \frac{1}{m_{h,i}} \sum_{r \in s_{h,h,i}} (\sigma^2_{[r:M]} + \mu^2_{[r:M]}) \\
&\quad - \left(\frac{1}{m_{h,i}} \sum_{r \in s_{h,h,i}} \mu_{[r:M]} \right)^2 .
\end{aligned}
$$

Let

$$\bar{X}_{PROS} = \frac{1}{LM}\sum_{i=1}^{K}\sum_{h=1}^{n_i} m_{h,i} X_{[s_{h,h,i}]}$$

Lemma

Let $X_{[s_{r,r,i}]}$, $r = 1, \cdots, n_i, i = 1, \cdots, K$ be a balanced random sample from design G. The estimator $\bar{X}_{PROS}$ is unbiased for μ and its conditional variance given design G is equal to

$$\eta^2 = \frac{1}{L^2M^2}\sum_{j=1}^{K}\sum_{r=1}^{n_j} m_{r,j} \sum_{h\in s_{r,r,j}} \{\sigma^2_{h:M} + (\mu^2_{[h:M]} - \mu^2)\}$$
$$-\frac{1}{L^2M^2}\sum_{j=1}^{K}\sum_{r=1}^{n_j} m^2_{r,j}(\gamma^2_{[r;m_{r,j}]} - \mu^2).$$

- For design G^*, the variance of $\bar{X}_{PROS}$ simplifies to

$$\eta^2 = \frac{\sigma^2 \sum_{j=1}^{K} m_j}{ML^2} - \frac{1}{M^2 L^2} \sum_{j=1}^{K} m_j^2 \sum_{i=1}^{n_j} (\gamma_{[i;m_j]} - \bar{\gamma}_{n_j})^2,$$

where $\bar{\gamma}_{n_j} = \sum_{i=1}^{n_j} \gamma_{[i;m_j]} / n_j$.

- For design G^{**}, the variance of $\bar{X}_{PROS}$ simplifies to

$$\eta^2 = \frac{\sigma^2}{Kn} - \frac{1}{n^2 K} \sum_{i=1}^{n} (\gamma_{[i;m]} - \bar{\gamma}_n)^2.$$

Relative efficiency with respect to RSS

- Assume that ranking is free.
- A fair comparison can be made by matching the number of fully measured units in both designs.
- Let $\bar{X}_{RSS} = \frac{1}{H}\sum_{h=1}^{H} X_{[1:H]}$ be a ranked set sample with just one cycle.
- We consider $\bar{X}_{PROS} = \frac{1}{M}\sum_{r=1}^{H} m_{r,1} X_{[s_{r,r,1}]}$ with just one group.
- Let $D = \{m_{1,1}, \cdots, m_{H,1}\}$ be the set of integers that contains number of units in subsets $s_{r,r,1}$, $r = 1, \cdots, H$.
- Let $RE = \frac{var(\bar{X}_{RSS})}{var(\bar{X}_{PROS})}$.

Relative efficiencies for several designs

M	D	N RE	L RE	U RE	E RE
6	$\{3,3\}$	1.350	1.295	1.485	1.210
	$\{4,2\}$	1.087	1.081	1.089	1.645
	$\{5,1\}$	0.656	0.686	0.598	1.182
	$\{2,2,2\}$	1.331	1.278	1.441	1.212
	$\{1,4,1\}$	0.724	0.817	0.563	0.931
8	$\{4,4\}$	1.438	1.363	1.636	1.255
	$\{5,3\}$	1.260	1.225	1.333	1.687
	$\{6,2\}$	0.902	0.920	0.851	1.708
	$\{3,2,3\}$	1.316	1.222	1.600	1.125
	$\{2,4,2\}$	1.363	1.415	1.176	1.441
	$\{1,6,1\}$	0.559	0.64	0.426	0.755
	$\{2,2,2,2\}$	1.402	1.333	1.543	1.261

Summary:

- RE increases with M.
- For symmetric distributions, design G^* or design G^{**} provides high efficiency with respect to a comparable RSS design.
- If we high confidence for correct ranking, we select large n (number of judgment subsets).
- For skewed distributions, exponential, gamma, log-normal, G^* and G^{**} are not optimal.
- Optimal design for skewed distribution may depend on the shape and the location(s) of the mode(s) of the underlying distribution.

Unconditional Variance of $\bar{X}_{PROS}$

- We assume that the number of judgment subsets n is an integer random variable.
- Note that n can take certain values, i.e M/n must be an integer.
- For example if $M = 12$, then $n = 1, 2, 3, 4, 6, 12$.
- Assume that n has a truncated binomial (p,M) random variable, with success probability p and sample size M, which puts all its probability masses on possible values of n
- Under this model, unconditional variance of $\bar{X}_{PROS}$ in design G^* is given by

$$\eta^2 = \frac{1}{K}\sum_n b(n,p,M)\left\{\frac{\sigma^2}{n} - \frac{\sum_{h=1}^{n}(\gamma_{[h;M/n]} - \mu)^2}{n^2}\right\}$$

Relative efficiency of $\bar{X}_{PROS}$, $M = 8$

p	$E(n)$	$var(\bar{X}_{SRS})/\eta^2$	$var(\bar{X}_{RSS})/\eta^2$
0.100	1.303	1.095	1.037
0.200	1.639	1.209	1.078
0.300	2.119	1.365	1.126
0.400	2.713	1.609	1.184
0.500	3.290	1.983	1.253
0.600	3.869	2.476	1.318
0.670	4.544	2.850	1.338
0.700	5.009	3.017	1.331
0.800	7.111	3.644	1.168
0.900	7.958	3.980	1.010

Lemma

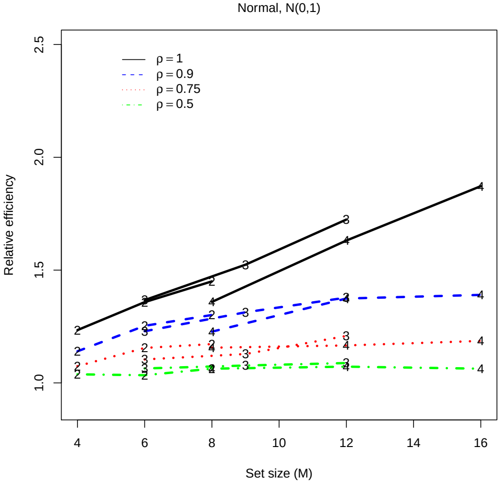
Under perfect ranking, assume $m_{h,i} = M/2$, $h = 1, \cdots, 2$, $i = 1, \cdots, K$. *For this design,* $D^{**} = \{M/2, M/2\}$, $RE_{D^{**}}(\bar{X}_{RSS}, \bar{X}_{PROS}) \geq 1$ *for any even* M.

- This result holds for perfect ranking.
- Under imperfect ranking, obtaining similar efficiency result depends on judgment ranking model.

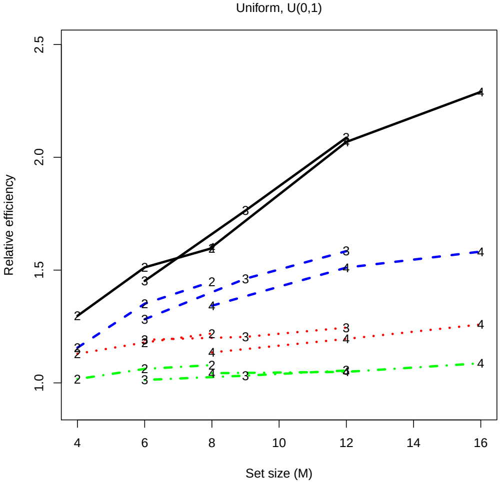
Dell Clutter model (1972)

- Ranking is performed based on the perceived values of the experimental units.
- Let $Y = (Y_1, \cdots, Y_M)$ be an M-dimensional random vector from underlying distribution F
- Let $\epsilon = (\epsilon_1 \cdots, \epsilon_M)$ be another random vector from a normal distribution with mean zero and variance τ^2.
- Let $X = Y + \epsilon$.
- The vector X is sorted and the unit $Y_{[i]}$ that corresponds $X_{(i)}$ is selected as the i-th judgment order statistic.
- Quality of ranking information is measured with $\rho = cor(X, Y)$, or equivalently with variance of ϵ (τ^2).

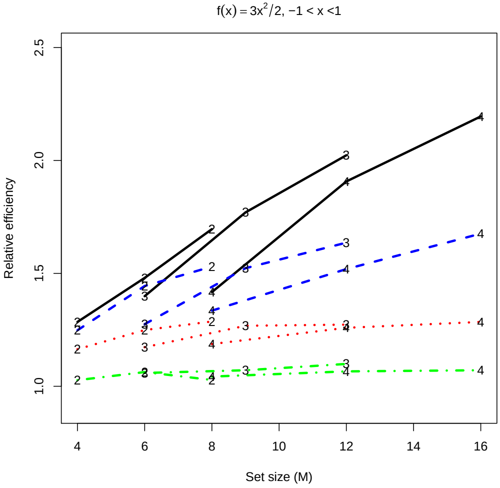
Relative efficiency of $\bar{X}_{PROS}$ with respect to $\bar{X}_{RSS}$



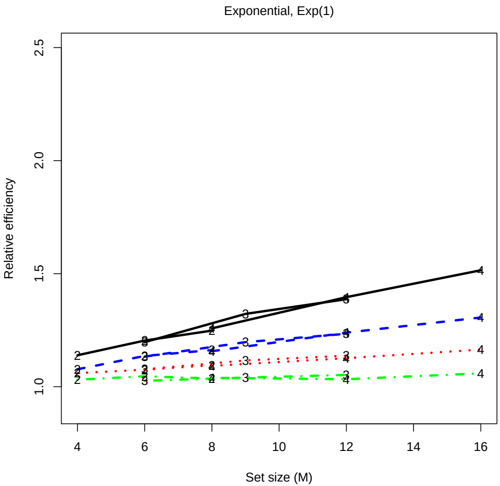
Relative efficiency of $\bar{X}_{PROS}$ with respect to $\bar{X}_{RSS}$



Relative efficiency of $\bar{X}_{PROS}$ with respect to $\bar{X}_{RSS}$



Relative efficiency of $\bar{X}_{PROS}$ with respect to $\bar{X}_{RSS}$



- Simple random sample estimator: Sample variance, S^2, based on a simple random sample.
- Stokes (1980) estimator: Sample variance based on a ranked set sample.
- MOSW (MacEachern, Ozturk, Stark and Wolfe, 2002) estimator:

$$
\begin{aligned}
\hat{\sigma}^2 &= \frac{1}{2n(n-1)k^2}\sum_{i\neq j}^{n}\sum_{h=1}^{k}(Y_{[h]i} - Y_{[h]j})^2 \\
&\quad + \frac{1}{2n^2k^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{h\neq h'}^{k}(Y_{[h]i} - Y_{[h']j})^2, \\
&= A + B
\end{aligned}
$$

- The first term represents within judgment class variation.
- The second term represents between judgment class variation.

Assume that $m_{h,i} = m$, $h = 1, \cdots, n_i$, $i = 1, \cdots, K$, with $n_i \equiv n$.

- Let
$$A = \frac{1}{2K^2M^2} \sum_{i=1}^{K} \sum_{j=1}^{K} \sum_{r \neq h}^{n} (mX_{[s_{r,r,i}]} - mX_{s_{[h,h,i]}})^2.$$
- Let
$$B = \frac{1}{2K(K-1)M^2} \sum_{i \neq j}^{K} \sum_{r=1}^{n} (mX_{[s_{r,r,i}]} - mX_{[s_{r,r,i}]})^2.$$
- The new estimator is defined as $T^2 = A + B$.
- T^2 is unbiased.

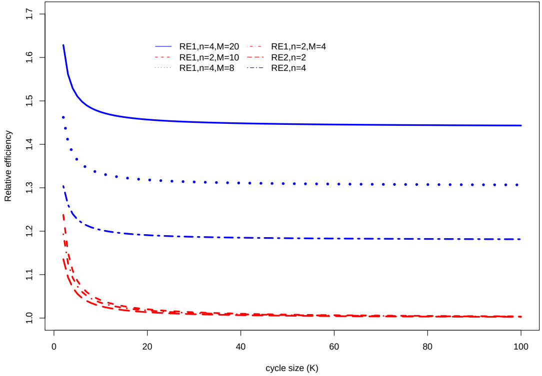
Table: Relative efficiency of T_K^2 with respect $\hat{\sigma}^2$ and S^2 for several distributions, $RE_1 = var(\hat{\sigma}^2)/var(T_K^2)$, $RE_2 = var(S^2)/var(\hat{\sigma}^2)$.

			$U(0,1)$		$Expon$		$N(0,1)$		$G(5,1)$	
M	K	n	RE_1	RE_2	RE_1	RE_2	RE_1	RE_2	RE_1	RE_2
4	2	2	1.130	1.463	1.013	1.066	1.051	1.193	1.033	1.135
4	5	2	1.045	1.187	1.005	1.049	1.016	1.073	1.011	1.065
10	2	2	1.248	1.615	1.021	1.075	1.090	1.238	1.057	1.161
10	5	2	1.081	1.228	1.007	1.050	1.028	1.086	1.018	1.072
6	2	3	1.204	1.738	1.022	1.124	1.085	1.322	1.054	1.229
6	5	3	1.156	1.441	1.017	1.105	1.068	1.212	1.043	1.164
15	2	3	1.509	2.178	1.037	1.141	1.169	1.424	1.101	1.283
15	5	3	1.409	1.757	1.029	1.117	1.138	1.292	1.081	1.207
8	2	4	1.286	2.023	1.033	1.182	1.121	1.462	1.077	1.326
8	5	4	1.257	1.732	1.030	1.165	1.111	1.364	1.070	1.269
20	2	4	1.791	2.817	1.058	1.210	1.249	1.629	1.147	1.413
20	5	4	1.726	2.378	1.052	1.190	1.231	1.510	1.136	1.347

Summary:

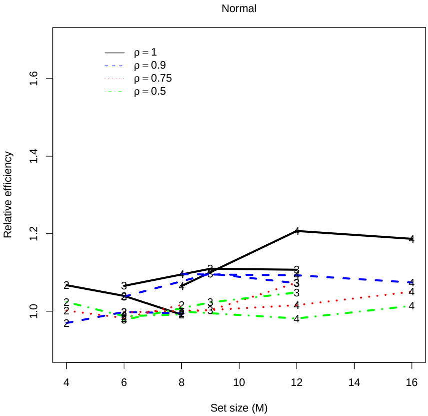
- Efficiency of T_K^2 is higher than the efficiencies of $\hat{\sigma}^2$ and S^2.
- Efficiency increases with set size M when $n > 2$.
- Efficiency decreases for a fixed set size M and $n = 2$ when K goes from 2 to 5.
- For skewed distributions, proposed variance estimator appears to have the same efficiency as the RSS variance estimator of MacEachern, Ozturk, Stark and Wolfe(2002).

Relative efficiency of T_K^2 and $\hat{\sigma}^2$ with respect to S^2 ($RE_1 = var(\hat{\sigma}^2)/var(T_K^2)$ and $RE_2 = var(S^2)/var(\hat{\sigma}^2)$).

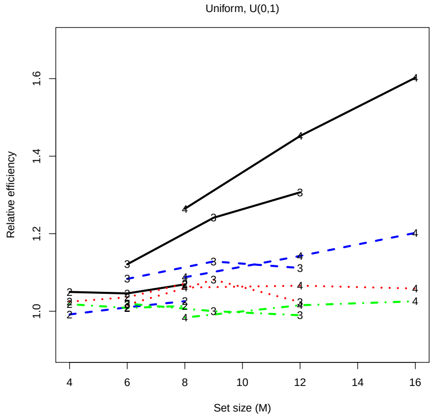


When $n = 2$, T_K^2 is asymptotically equivalent to $\hat{\sigma}^2$ even for large M.

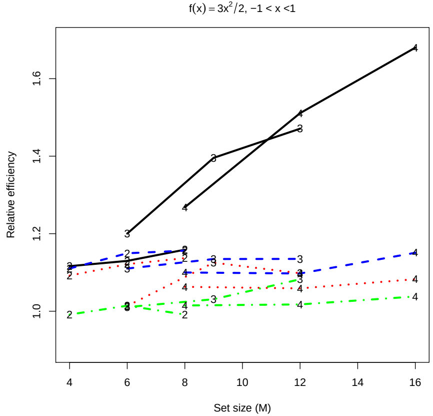
Relative efficiency of T_K^2 with respect to $\hat{\sigma}^2$, $RE = var(\hat{\sigma}^2)/var(T_K^2)$. The number of subsets (n) are given on each line.



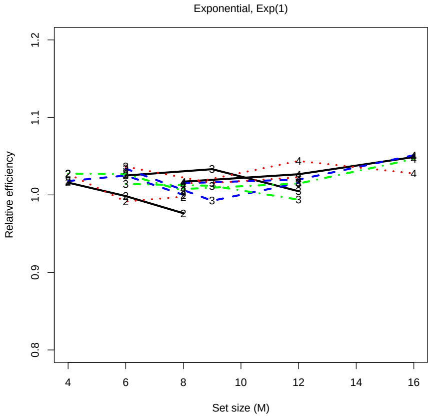
Relative efficiency of T_K^2 with respect to $\hat{\sigma}^2$, $RE = var(\hat{\sigma}^2)/var(T_K^2)$. The number of subsets (n) are given on each line.



Relative efficiency of T_K^2 with respect to $\hat{\sigma}^2$, $RE = var(\hat{\sigma}^2)/var(T_K^2)$. The number of subsets (n) are given on each line.



Relative efficiency of T_K^2 with respect to $\hat{\sigma}^2$, $RE = var(\hat{\sigma}^2)/var(T_K^2)$. The number of subsets (n) are given on each line.



Judgment subsetting error

- Treat judgment subset classes as a factor in one-way anova model.
- Compute MST and MSE.
- The proposed variance can be written as

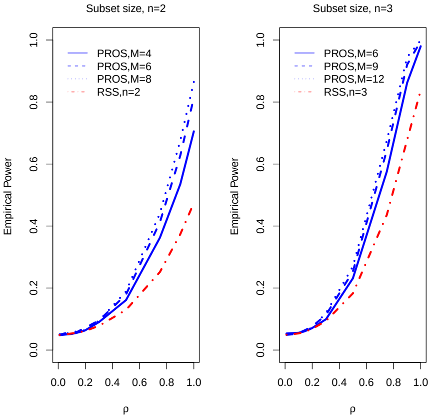
$$T_K^2 = \frac{1}{nK}\{(n-1)MST + (nK - n + 1)MSE\}.$$

- We then use

$$U_K = \frac{MST}{MSE}$$

  to test if the judgment subsetting process is effective to create partially ordered subsets.
- Large values of U_K indicates that judgment subsetting is effective.

Empirical power of U_K



Two sample Problem

- Let F and H be the CDF of X and Y samples with unique medians θ_1 and θ_2, respectively.
- Let $X_{s_{[r_i]i}}$, $i = 1, \ldots, N_1$ and $Y_{s_{[t_j]j}}$, $j = 1, \ldots, N_2$ be PROS samples from distribution F and H with respective set and cycle sizes, M_1,K_1, and M_2,K_2.
- We wish to test $H_o : \Delta = \theta_2 - \theta_1 = 0$ against $H_a : \Delta \neq 0$.
- We use the test statistics

$$\bar{T}^* = \frac{1}{K_1 M_1 K_2 M_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} m_{[r_i]i} q_{[t_j]j} ((I(X_{s_{[r_i]i}} \leq Y_{s_{[t_j]j}}) - \tau_{r_i t_j}),$$

where $m_{[r_i]i}$ and $q_{[t_j]j}$ are the number of unranked units in each subsets of X and Y samples, respectively, and $\tau_{r_i t_j} = E(I(X_{s_{[r_i]i}} \leq Y_{s_{[t_j]j}}))$.

Asymptotic results

- Let $N_t = K_1 n_1 + K_2 n_2$, $\lambda = \lim_{N_1 \to \infty} \frac{K_1 M_1}{N_t}$,and $\eta_i = \lim_{N_1 \to \infty} \frac{N_i}{K_i M_i}$, where $0 < \lambda < 1$. For large values of N_t, under the null hypothesis, the conditional distribution of $\sqrt{N_t} T^*$ for a given design G converges to a normal distribution with mean 0, and variance $\sigma^2 = \frac{\eta_1}{\lambda}\zeta_{0,1} + \frac{\eta_2}{1-\lambda}\zeta_{1,0}$, where

$$\zeta_{0,1} = \lim_{N_1 \to \infty} \sum_{i=1}^{N_1} \frac{m^2_{[r_i]i}}{K_1 M_1} \left\{ E_{\bar{F}_{[r_i]}}(H^2(y)) - (E_{\bar{F}_{[r_i]}}(H(y)))^2 \right\}$$

$$\zeta_{1,0} = \lim_{N_2 \to \infty} \sum_{j=1}^{N_2} \frac{q^2_{[t_j]j}}{K_2 M_2} \left\{ E_{\bar{H}_{[t_j]}}(F^2(y)) - (E_{\bar{H}_{[t_j]}}(F(y)))^2 \right\},$$

where $\bar{F}_{[r_i]}(y) = P(X_{[s_{r_i},r_i,i]} \leq y)$ and $\bar{H}_{[r_i]}(y) = P(Y_{[s_{t_j},t_j,j]} \leq y)$

- Under G^{**} design and H_0, $\zeta_{0,1}$ and $\zeta_{1,0}$ simplify to:

$$\zeta_{0,1} = \frac{m}{3} - \frac{m^2}{M_1} \sum_{r=1}^{M_1/m} \left(\int F(z) d\bar{F}_{[r]}(z)\right)^2 = \frac{m}{3} - \frac{m^2}{M_1}\tau_1^2$$

$$\zeta_{0,1} = \frac{q}{3} - \frac{q^2}{M_2} \sum_{t=1}^{M_2/q} \left(\int F(z) d\bar{F}_{[t]}(z)\right)^2 = \frac{q}{3}\frac{q^2}{M_2}\tau_2^2.$$

- Under perfect ranking, $\zeta_{0,1}$ and $\zeta_{1,0}$ are asymptotically distribution free, and equal to :

$$\zeta_{0,1} = \frac{m}{3} - \frac{m^2}{4M_1(M_1+1)^2} \sum_{r=1}^{M_1/m} (2rm - m + 1)^2$$

$$\zeta_{1,0} = \frac{q}{3} - \frac{q^2}{4M_2(M_2+1)^2} \sum_{t=1}^{M_2/q} (2tq - q + 1)^2.$$

Pittman Asymptotic Relative efficiency Results

- Simple random sampling (SRS)

$$eff(SRS) = 3\left(\int f^2(x)dx\right)^2.$$

- Rank set sampling (RSS). Let $n = M/m$, and then

$$eff(RSS) = \frac{3(n+1)}{2}\left(\int f^2(x)dx\right)^2.$$

- Partially Ranked-Ordered set sampling.Let $M_1 = M_2 = M$, $K_1 = K_2$, $m = q$, then

$$eff(PROS) = \frac{3\int f^2(x)dx)^2}{4 - \frac{3}{n(M+1)^2}\sum_{r=1}^{n}(2rm - m + 1)^2}.$$

Table: Pittman Asymptotic Relative efficiency(ARE)

M	m	n	ARE(RSS,SRS)	ARE(PROS,SRS)	ARE(PROS,RSS)
6	2	3	2	2.88	1.44
6	3	2	1.5	2.23	1.48
8	2	4	2.5	3.85	1.54
8	4	2	1.5	2.45	1.63

Null Distribution of rank-sum test

- We reject$H_0 : \theta_1 - \theta_2 = 0$ against $H_A : \theta_1 - \theta_2 \neq 0$ for extreme values of T^*.
- Limiting distribution of $\sqrt{N_t} T^* / \sqrt{\sigma}$ is normal.
- The asymptotic variance $\sigma^2 = \frac{\eta_1}{\lambda}\zeta_{0,1} + \frac{\eta_2}{1-\lambda}\zeta_{1,0}$ is not distribution free in general. It is distribution free under perfect ranking,

$$\zeta_{0,1} = \frac{m}{3} - \frac{m^2}{M_1} \sum_{r=1}^{M_1/m} (\int F(z) d\bar{F}_{[r]}(z))^2 = \frac{m}{3} - \frac{m^2}{M_1}\tau_1$$

$$\begin{aligned}
\zeta_{0,1} &= \frac{q}{3} - \frac{q^2}{M_2} \sum_{t=1}^{M_2/q} (\int F(z) d\bar{F}_{[t]}(z))^2 = \frac{q}{3} - \frac{q^2}{M_2}\tau_2, \\
\tau_1 &= \sum_{r}^{M_1/m} \tau_{.r}^2, \tau_2 = \sum_{r=1}^{M_2/q} \tau_{.r}^2
\end{aligned}$$

- Under imperfect ranking σ must be estimated.

Estimation of σ

- Let $\hat{\theta}_1$ and $\hat{\theta}_2$ be the sample medians of X- and Y-samples, respectively.
- We first center the X-and Y-sample data by subtracting the sample medians from each observations

$$X^*_{s_{[r]i}} = X_{s_{[r]i}} - \hat{\theta}_1, \, Y^*_{s_{[t]j}} = Y_{s_{[t]j}} - \hat{\theta}_2,$$

- The centered observations are pooled to create

$$\boldsymbol{Z}^*_{[r]} = (X^*_{s_{[r]1}}, \cdots X^*_{s_{[r]K_1}}, Y^*_{s_{[r]1}}, \cdots Y^*_{s_{[r]K_2}}), \quad r = 1, \cdots, n.$$

- Let $\hat{F}(t)$ be the empirical CDF of F based on $\boldsymbol{Z}^*_{[r]}$, $r = 1, \cdots, n$

$$\hat{F}(t) = \frac{1}{nK_1 + nK_2} \sum_{u=1}^{n} \sum_{i=1}^{K_1+K_2} I(Z^*_{[u]i} \leq t).$$

- An unbiased estimator of $\tau_{.r}$ is then given by

$$\hat{\tau}_{.r} = \frac{1}{K_1 + K_2} \sum_{i=1}^{K_1+K_2} \{\hat{F}(Z_{[r]i}) - \frac{K_1/2 + K_2/2}{nK_1 + nK_2}\}, r = 1, \cdots, n.$$

Null distribution of T^*

- Assume that $m = q$, then $\tau_1 = \tau_2$,

$$\hat{\zeta}_{0,1} = \hat{\zeta}_{1,0} = \frac{m}{3} - \frac{m^2}{M_1}\hat{\tau}_{.r}^2$$

- The estimate of the σ^2 is given by

$$\hat{\sigma}^2 = \frac{\eta_1}{\lambda}\hat{\zeta}_{0,1} + \frac{\eta_2}{1-\lambda}\hat{\zeta}_{1,0}$$

- The limiting distribution of $T^*/\hat{\sigma}$ can be approximated by Studen's t-distributions with degrees of freedom $N_t - 1$

Confidence Interval for Δ

- Let $D_{(1)} < D_{(2)} < \cdots < D_{(N_1 N_2)}$ be the ordered differences $X_{s_{[r_i]i}} - Y_{s_{[t_j]j}}$, $i = 1, \cdots, N_1$, $j = 1, \cdots, N_2$.
- A $(1-\alpha)100\%$ confidence interval for Δ is given by $(D_{(k^*+1)}, D_{(N_1 N_2 - k^*)})$, where

$$P(R(0) \leq k^*) = \alpha/2$$

  and $R(0) = N_1 N_2 \bar{T}(0) + N_1 N_2/2$.
- We then use continuity corrected Student's t-approximation to compute k^*

$$P(R(0) \leq k^* + 1/2) \approx P(t_{N-2} \leq \frac{\sqrt{N}\{(k^* + 1/2)/(N_1 N_2) - 1/2\}}{\hat{\sigma}}).$$

- From the above expression, one can compute k^*

$$k^* = \frac{N_1 N_2}{2} + \frac{N_1 N_2 t_{N-2}(\alpha/2)\hat{\sigma}}{\sqrt{N}} - 1/2. \quad (1)$$

Empirical Evidence

Table: Calibrated(C) and Uncalibrated(UC) Type I error rates of rank-sum test. Simulation size is 5000.

K	M	n	ρ	normal		lognormal		$t(3)$	
				UC	C	UC	C	UC	C
5	6	3	1	0.042	0.063	0.048	0.061	0.048	0.063
5	6	3	0.9	0.096	0.067	0.156	0.057	0.116	0.063
5	6	3	0.75	0.140	0.056	0.197	0.060	0.163	0.058
5	6	3	0.5	0.195	0.060	0.228	0.057	0.199	0.055
5	6	3	0.3	0.226	0.056	0.233	0.058	0.222	0.048
5	12	3	1	0.043	0.066	0.048	0.061	0.050	0.065
5	12	3	0.9	0.120	0.063	0.220	0.058	0.175	0.063
5	12	3	0.75	0.215	0.057	0.286	0.063	0.244	0.058
5	12	3	0.5	0.277	0.055	0.306	0.057	0.295	0.056
5	12	3	0.3	0.310	0.049	0.306	0.059	0.307	0.051

Table: Calibrated(C) and uncalibrated(UC) coverage probabilities based on PROSS design. Simulation size is 5000. $RE = L^2_{RSS}/L^2_{PROSS}$, where L is the average length of simulated confidence intervals.

K	M	n	ρ	normal			lognormal		
				UC	C	RE	UC	C	RE
5	6	3	1.00	0.954	0.936	1.440	0.953	0.937	1.586
5	6	3	0.90	0.915	0.941	1.238	0.839	0.936	1.163
5	6	3	0.75	0.855	0.942	1.119	0.795	0.942	1.063
5	6	3	0.50	0.794	0.939	1.041	0.787	0.952	1.030
5	6	3	0.30	0.772	0.946	1.009	0.766	0.946	1.003
5	12	3	1.00	0.949	0.934	2.038	0.943	0.930	2.372
5	12	3	0.90	0.867	0.929	1.512	0.768	0.941	1.302
5	12	3	0.75	0.783	0.932	1.228	0.725	0.937	1.122
5	12	3	0.50	0.704	0.943	1.076	0.686	0.948	1.030
5	12	3	0.30	0.683	0.949	1.037	0.662	0.943	1.023

Stratified rank-sum test

- For small cycle sizes, test may not be distribution-free if there is ranking error. In this case, we construct an exact test.
- For each subset $s_{[i]}$, $i = 1, \cdots, n$. Let $X_{s_{[i]j}}$, $j = 1, \cdots, K_{1i}$, $i = 1, \cdots, n$ and $Y_{s_{[i]j}}$, $j = 1, \cdots, K_{2i}$, $i = 1, \cdots, n$ be PROSS samples from X and Y-populations, respectively
- Let $W_\alpha = \sum_{i=1}^n \alpha_i W_{ii}$ be a class of test statistics,

$$W_\alpha = \sum_{i=1}^{n} \alpha_i W_{ii}, \quad W_{ii} = \sum_{j=1}^{K_{1i}} \sum_{t=1}^{K_{2i}} I(X_{s_{[r_i]j}} < Y_{s_{[r_i]t}}),$$

- The null distribution of W_α is distribution-free regardles the quality of ranking information.
- Among all positive weights $\alpha_i > 0$, $i = 1, \cdots, n$, the weigh

$$\alpha_i = \frac{\int \bar{f}^2_{[r_i]}(y)dy}{1/\lambda_{1i} + 1/\lambda_{2i}}$$

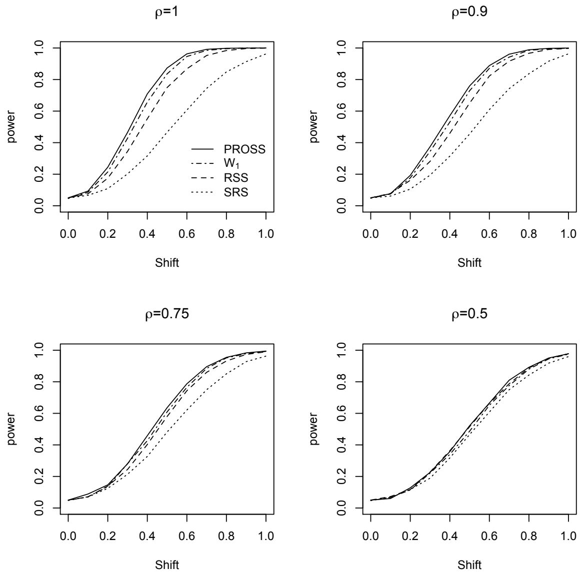
maximizes Pittman efficacy.

Asympotic Pitman relative efficiecny

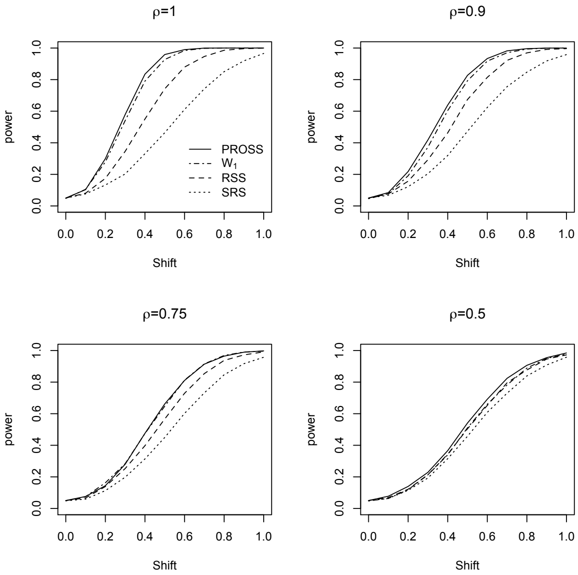
Table: Pittman Asymptotic Relative efficiency of stratified rank-sum test, $RE_1 = \frac{c^2_{Wo}}{c^2_{W1}}$, $RE_2 = \frac{c^2_{W1}}{c^2_{PROSS}}$, $RE_3 = \frac{c^2_{W1}}{c^2_{RSS}}$.

M	n	normal			t3		
		RE1	RE2	RE3	RE1	RE2	RE3
6	2	1.000	0.930	1.381	1.000	0.870	1.292
6	3	1.009	0.939	1.353	1.042	0.866	1.248
8	2	1.000	0.914	1.495	1.000	0.858	1.404
8	4	1.017	0.932	1.438	1.069	0.856	1.321
12	2	1.000	0.895	1.653	1.000	0.846	1.563
12	3	1.022	0.899	1.853	1.073	0.841	1.733
12	4	1.027	0.910	1.809	1.091	0.844	1.677
12	6	1.025	0.926	1.542	1.097	0.848	1.413
16	2	1.000	0.886	1.760	1.000	0.842	1.672
16	4	1.035	0.896	2.113	1.107	0.837	1.975
16	8	1.030	0.923	1.602	1.111	0.846	1.468

Simulation Results



Simulation Results



A few remarks

- We proposed a sampling scheme that generates data from partially ordered subsets.
- The new sampling scheme yields smaller ranking error than a standard ranked set sampling design.
- Even with the fixed designs, the proposed sampling scheme provides improvement for sample mean and variance estimators over the same estimators based on standard RSS scheme.
- The proposed sampling design can be expanded to other inferential procedures easily.