# Some Efficient Sampling Designs Based on Partial Order Sets

Bardia Panahbehagh and Rainer Bruggemann

Department of Mathematics, Kharazmi University, Iran
Leibniz Institute of Freshwater Ecology and Inland Fisheries, Germany

12th Partial Order Workshop, University of Neuchatel
October 26-27, 2018

# Outline; Sampling Designs and Partial Order Theory

# Outline; Sampling Designs and Partial Order Theory

- **Part 1:** Improve sampling designs based on ranks of data (RSS).

# Outline; Sampling Designs and Partial Order Theory

- **Part 1:** Improve sampling designs based on ranks of data (RSS).

- **Part 2:** An economic version of RSS (VSR).

# Outline; Sampling Designs and Partial Order Theory

- **Part 1:** Improve sampling designs based on ranks of data (RSS).
- **Part 2:** An economic version of RSS (VSR).
- **Part 3:** Multivariate VSR.

# Outline; Sampling Designs and Partial Order Theory

- **Part 1:** Improve sampling designs based on ranks of data (RSS).
- **Part 2:** An economic version of RSS (VSR).
- **Part 3:** Multivariate VSR.
- **Part 4:** Multivariate variable sampling based on Partial Order Theory.

# Outline; Sampling Designs and Partial Order Theory

- **Part 1:** Improve sampling designs based on ranks of data (RSS).
- **Part 2:** An economic version of RSS (VSR).
- **Part 3:** Multivariate VSR.
- **Part 4:** Multivariate variable sampling based on Partial Order Theory.
- **Part 5:** Simulations.

# What are we looking for in Sampling Theory?

# What are we looking for in Sampling Theory?

- Efficient Sampling Designs


  - Low Cost
  - High Precision

# Assumptions and Notations

- Suppose that $X$ is the variable of interest with $E(X) = \mu$.

# Assumptions and Notations

- Suppose that $X$ is the variable of interest with $E(X) = \mu$.

- The goal is to estimate $\mu$ and the variance of the estimator.

# Assumptions and Notations

- Suppose that $X$ is the variable of interest with $E(X) = \mu$.
- The goal is to estimate $\mu$ and the variance of the estimator.
- Suppose further that $Y$ is an auxiliary variable (used for ranking) and suppose that $Y$ has reasonable correlation with the main variable $X$.

# Improve sampling designs based on ranks of data (RSS)

# Part 1

# Simple Random Sampling with $n = 3$ (sample size), X is height.

# Simple Random Sampling with $n = 3$ (sample size), X is height.

# Ranked Set Sampling (RSS) with $n = 3$.

# Ranked Set Sampling (RSS) with $n = 3$.

# Ranked Set Sampling (RSS)

Ranked Set Sampling of Size $m$:

- Select $m$ sets of size $m$ based on SRS

# Ranked Set Sampling (RSS)

Ranked Set Sampling of Size $m$:

- Select $m$ sets of size $m$ based on SRS
- Sort each set, based on an auxiliary variables

# Ranked Set Sampling (RSS)

Ranked Set Sampling of Size $m$:

- Select $m$ sets of size $m$ based on SRS
- Sort each set, based on an auxiliary variables
- Full measure the main diagonal of the matrix as the final sample

# Ranked Set Sampling (RSS)

Ranked Set Sampling of Size $m$:

- Select $m$ sets of size $m$ based on SRS
- Sort each set, based on an auxiliary variables
- Full measure the main diagonal of the matrix as the final sample

| | | Min | Second Min | $\cdots$ | Max |
|---|---|---|---|---|---|
| | 1 | $X_{[1]1}$ | $X_{[2]1}$ | $\cdots$ | $X_{[m]1}$ |
| Sets | 2 | $X_{[1]2}$ | $X_{[2]2}$ | $\cdots$ | $X_{[m]2}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| | m | $X_{[1]m}$ | $X_{[2]m}$ | $\cdots$ | $X_{[m]m}$ |

# Ranked Set Sampling (RSS)

Ranked Set Sampling of Size $m$:

- Select $m$ sets of size $m$ based on SRS
- Sort each set, based on an auxiliary variables
- Full measure the main diagonal of the matrix as the final sample

| | | Min | Second Min | $\cdots$ | Max |
|---|---|---|---|---|---|
| | 1 | $X_{[1]1}$ | $X_{[2]1}$ | $\cdots$ | $X_{[m]1}$ |
| Sets | 2 | $X_{[1]2}$ | $X_{[2]2}$ | $\cdots$ | $X_{[m]2}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| | m | $X_{[1]m}$ | $X_{[2]m}$ | $\cdots$ | $X_{[m]m}$ |

# RSS of Size $n_\bullet = nm \Longrightarrow n$ times a RSS of size $m$

# RSS of Size $n_\bullet = nm \Longrightarrow n$ times a RSS of size $m$

| Cycle | Set | Ranks 1 | 2 | $\cdots$ | m |
|---|---|---|---|---|---|
| 1 | 1 | $X_{[1]11}$ | $X_{[2]11}$ | $\cdots$ | $X_{[m]11}$ |
| | 2 | $X_{[1]21}$ | $X_{[2]21}$ | $\cdots$ | $X_{[m]21}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| | m | $X_{[1]m1}$ | $X_{[2]m1}$ | $\cdots$ | $X_{[m]m1}$ |
| 2 | 1 | $X_{[1]12}$ | $X_{[2]12}$ | $\cdots$ | $X_{[m]12}$ |
| | 2 | $X_{[1]22}$ | $X_{[2]22}$ | $\cdots$ | $X_{[m]22}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| | m | $X_{[1]m2}$ | $X_{[2]m2}$ | $\cdots$ | $X_{[m]m2}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| n | 1 | $X_{[1]1n}$ | $X_{[2]1n}$ | $\cdots$ | $X_{[m]1n}$ |
| | 2 | $X_{[1]2n}$ | $X_{[2]2n}$ | $\cdots$ | $X_{[m]2n}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| | m | $X_{[1]mn}$ | $X_{[2]mn}$ | $\cdots$ | $X_{[m]mn}$ |

# RSS of Size $n_\bullet = nm \Longrightarrow n$ times a RSS of size $m$

| Cycle | Set | Ranks 1 | 2 | $\cdots$ | m |
|---|---|---|---|---|---|
| 1 | 1 | $X_{[1]11}$ | $X_{[2]11}$ | $\cdots$ | $X_{[m]11}$ |
| | 2 | $X_{[1]21}$ | $X_{[2]21}$ | $\cdots$ | $X_{[m]21}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| | m | $X_{[1]m1}$ | $X_{[2]m1}$ | $\cdots$ | $X_{[m]m1}$ |
| 2 | 1 | $X_{[1]12}$ | $X_{[2]12}$ | $\cdots$ | $X_{[m]12}$ |
| | 2 | $X_{[1]22}$ | $X_{[2]22}$ | $\cdots$ | $X_{[m]22}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| | m | $X_{[1]m2}$ | $X_{[2]m2}$ | $\cdots$ | $X_{[m]m2}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| n | 1 | $X_{[1]1n}$ | $X_{[2]1n}$ | $\cdots$ | $X_{[m]1n}$ |
| | 2 | $X_{[1]2n}$ | $X_{[2]2n}$ | $\cdots$ | $X_{[m]2n}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| | m | $X_{[1]mn}$ | $X_{[2]mn}$ | $\cdots$ | $X_{[m]mn}$ |

# An economic version of RSS (VSR)

# Part 2

# Ranked Set Sampling (RSS)

Just a unit from each set?!

| | | Min | Second Min | $\cdots$ | Max |
|---|---|---|---|---|---|
| | 1 | $X_{[1]1}$ | $X_{[2]1}$ | $\cdots$ | $X_{[m]1}$ |
| Sets | 2 | $X_{[1]2}$ | $X_{[2]2}$ | $\cdots$ | $X_{[m]2}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| | m | $X_{[1]m}$ | $X_{[2]m}$ | $\cdots$ | $X_{[m]m}$ |

# Virtual Stratified Sampling Using RSS (VSR) Panahbehagh and Bruggemann (2017)

VSR of size $n_\bullet = \sum_{h=1}^{m} n_h$

- Select $K$ sets of size $m$ based on SRS
- Sort each set, based on an auxiliary variables to make $m$ strata
- Take a sample of size $n_h$ from stratum (rank) h.

| | | stratum 1 | stratum 2 | $\cdots$ | stratum m |
|---|---|---|---|---|---|
| | 1 | $X_{[1]1}$ | $X_{[2]1}$ | $\cdots$ | $X_{[m]1}$ |
| Sets | 2 | $X_{[1]2}$ | $X_{[2]2}$ | $\cdots$ | $X_{[m]2}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| | $K$ | $X_{[1]K}$ | $X_{[2]K}$ | $\cdots$ | $X_{[m]K}$ |

# Virtual Stratified Sampling Using RSS (VSR)

- Example: VSR of size $n_\bullet = 12$, $m = 4$, $n_h = n_\bullet / m = 3$

# Virtual Stratified Sampling Using RSS (VSR)

- Example: VSR of size $n_\bullet = 12$, $m = 4$, $n_h = n_\bullet / m = 3$
- $K > 3$ (Number of sets)

# Virtual Stratified Sampling Using RSS (VSR)

- Example: VSR of size $n_\bullet = 12$, $m = 4$, $n_h = n_\bullet/m = 3$
- $K > 3$ (Number of sets) for example $K = 5$

# Virtual Stratified Sampling Using RSS (VSR)

- Example: VSR of size $n_\bullet = 12$, $m = 4$, $n_h = n_\bullet/m = 3$
- $K > 3$ (Number of sets) for example $K = 5$

| | | stratum 1 | stratum 2 | stratum 3 | stratum 4 |
|---|---|---|---|---|---|
| | 1 | $X_{[1]1}$ | $X_{[2]1}$ | $X_{[3]1}$ | $X_{[4]1}$ |
| Sets | 2 | $X_{[1]2}$ | $X_{[2]2}$ | $X_{[3]2}$ | $X_{[4]2}$ |
| | 3 | $X_{[1]3}$ | $X_{[2]3}$ | $X_{[3]3}$ | $X_{[4]3}$ |
| | 4 | $X_{[1]4}$ | $X_{[2]4}$ | $X_{[3]4}$ | $X_{[4]4}$ |
| | 5 | $X_{[1]5}$ | $X_{[2]5}$ | $X_{[3]5}$ | $X_{[4]5}$ |

# Virtual Stratified Sampling Using RSS (VSR)

- Example: VSR of size $n_\bullet = 12$, $m = 4$, $n_h = n_\bullet / m = 3$
- $K > 3$ (Number of sets) for example $K = 5$

| | | stratum 1 | stratum 2 | stratum 3 | stratum 4 |
|---|---|---|---|---|---|
| | 1 | $X_{[1]1}$ | $X_{[2]1}$ | $X_{[3]1}$ | $X_{[4]1}$ |
| Sets | 2 | $X_{[1]2}$ | $X_{[2]2}$ | $X_{[3]2}$ | $X_{[4]2}$ |
| | 3 | $X_{[1]3}$ | $X_{[2]3}$ | $X_{[3]3}$ | $X_{[4]3}$ |
| | 4 | $X_{[1]4}$ | $X_{[2]4}$ | $X_{[3]4}$ | $X_{[4]4}$ |
| | 5 | $X_{[1]4}$ | $X_{[2]5}$ | $X_{[3]5}$ | $X_{[4]5}$ |

# Virtual Stratified Sampling Using RSS (VSR)

- Example: VSR of size $n_\bullet = 12$, $m = 4$, $n_h = n_\bullet/m = 3$
- $K > 3$ (Number of sets) for example $K = 5$

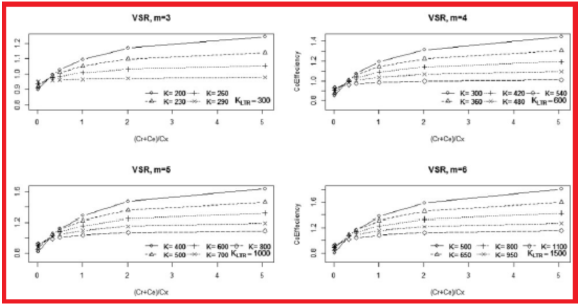
| | | stratum 1 | stratum 2 | stratum 3 | stratum 4 |
|---|---|---|---|---|---|
| | 1 | $X_{[1]1}$ | $X_{[2]1}$ | $X_{[3]1}$ | $X_{[4]1}$ |
| Sets | 2 | $X_{[1]2}$ | $X_{[2]2}$ | $X_{[3]2}$ | $X_{[4]2}$ |
| | 3 | $X_{[1]3}$ | $X_{[2]3}$ | $X_{[3]3}$ | $X_{[4]3}$ |
| | 4 | $X_{[1]4}$ | $X_{[2]4}$ | $X_{[3]4}$ | $X_{[4]4}$ |
| | 5 | $X_{[1]4}$ | $X_{[2]5}$ | $X_{[3]5}$ | $X_{[4]5}$ |

- For RSS with $n_\bullet = 12$ and $m = 4$ we need 12 sets!

# Virtual Stratified Sampling Using RSS (VSR)

With considering Cost and Precision simultaneously, VSR can be more efficient than RSS and SRS.

(Panahbehagh and Bruggemann (2017))



# Multivariate VSR

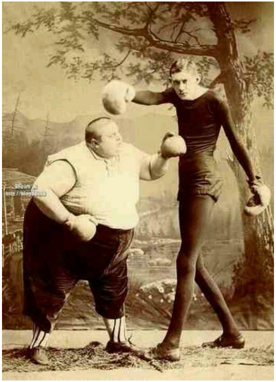
# Part 3

# Multivariate VSR

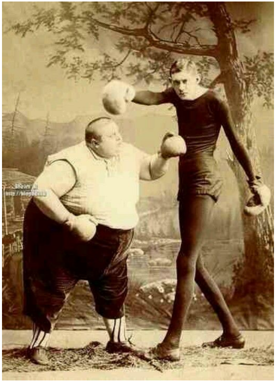
**Univariate $\Longrightarrow$ Multivariate**

# How to sort the data based on multivariate variable?!

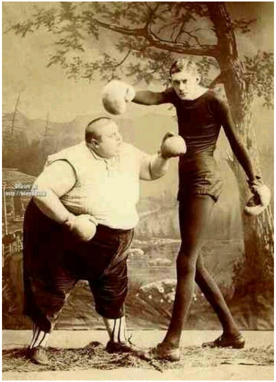
# How to sort the data based on multivariate variable?!



# How to sort the data based on multivariate variable?!

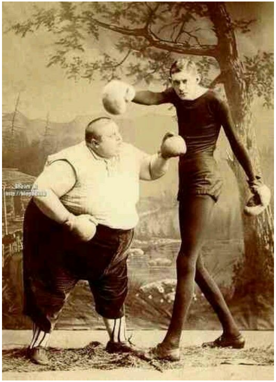


# How to sort the data based on multivariate variable?!



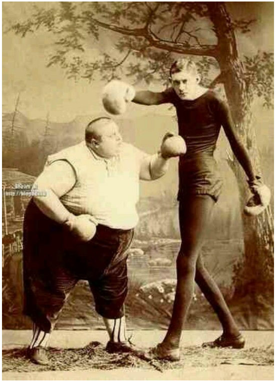
- Height? $\surd$

# How to sort the data based on multivariate variable?!



- Height? √
- Weight? √

# How to sort the data based on multivariate variable?!



- Height? √
- Weight? √
- Height and Weight? ?!?!

# How to sort the data based on multivariate variable?!

# How to sort the data based on multivariate variable?!

1 Using just one of the variables (MVSR).

# How to sort the data based on multivariate variable?!

1. Using just one of the variables (MVSR).

2. Using a linear combination of the variables.

# How to sort the data based on multivariate variable?!

1. Using just one of the variables (MVSR).

2. Using a linear combination of the variables.

3. Using many initial sample.

# Multivariate VSR

- $\mathbf{X} \sim f_{\boldsymbol{\mu}}$, where $\mathbf{X} = (X^1, X^2, ..., X^R)$ ($1, 2, ..., R$ are indexes and not powers!)

$$E(\mathbf{X}) = \boldsymbol{\mu} = (\mu^1, \mu^2, ..., \mu^R).$$

- Main aim is to estimate $\boldsymbol{\mu}$.

- Our strategy is the same as VSR and just we sort the set units only based on $X^1$.

- Virtual strata, using conventional RSS, sorted based on $X^1$

| stratum 1 | stratum 2 | $\cdots$ | stratum m |
|---|---|---|---|
| $\mathbf{X}_{[1]1}$ | $\mathbf{X}_{[2]1}$ | $\cdots$ | $\mathbf{X}_{[m]1}$ |
| $\mathbf{X}_{[1]2}$ | $\mathbf{X}_{[2]2}$ | $\cdots$ | $\mathbf{X}_{[m]2}$ |
| $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| $\mathbf{X}_{[1]K}$ | $\mathbf{X}_{[2]K}$ | $\cdots$ | $\mathbf{X}_{[m]K}$ |

- Virtual strata, using conventional RSS, sorted based on $X^1$

| stratum 1 | stratum 2 | $\cdots$ | stratum m |
|---|---|---|---|
| $\mathbf{X}_{[1]1}$ | $\mathbf{X}_{[2]1}$ | $\cdots$ | $\mathbf{X}_{[m]1}$ |
| $\mathbf{X}_{[1]2}$ | $\mathbf{X}_{[2]2}$ | $\cdots$ | $\mathbf{X}_{[m]2}$ |
| $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| $\mathbf{X}_{[1]K}$ | $\mathbf{X}_{[2]K}$ | $\cdots$ | $\mathbf{X}_{[m]K}$ |

- Now we can take a sample like VSR, a SRS from each stratum.

- Virtual strata, using conventional RSS, sorted based on $X^1$

| stratum 1 | stratum 2 | $\cdots$ | stratum m |
|---|---|---|---|
| $\mathbf{X}_{[1]1}$ | $\mathbf{X}_{[2]1}$ | $\cdots$ | $\mathbf{X}_{[m]1}$ |
| $\mathbf{X}_{[1]2}$ | $\mathbf{X}_{[2]2}$ | $\cdots$ | $\mathbf{X}_{[m]2}$ |
| $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| $\mathbf{X}_{[1]K}$ | $\mathbf{X}_{[2]K}$ | $\cdots$ | $\mathbf{X}_{[m]K}$ |

- Now we can take a sample like VSR, a SRS from each stratum.

| stratum 1 | stratum 2 | $\cdots$ | stratum m |
|---|---|---|---|
| $\mathbf{X}_{[1]1}$ | $\mathbf{X}_{[2]1}$ | $\cdots$ | $\mathbf{X}_{[m]1}$ |
| $\mathbf{X}_{[1]2}$ | $\mathbf{X}_{[2]2}$ | $\cdots$ | $\mathbf{X}_{[m]2}$ |
| $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| $\mathbf{X}_{[1]K}$ | $\mathbf{X}_{[2]K}$ | $\cdots$ | $\mathbf{X}_{[m]K}$ |

| stratum 1 | stratum 2 | $\cdots$ | stratum m |
|---|---|---|---|
| $\mathbf{X}_{[1]1}$ | $\mathbf{X}_{[2]1}$ | $\cdots$ | $\mathbf{X}_{[m]1}$ |
| $\mathbf{X}_{[1]2}$ | $\mathbf{X}_{[2]2}$ | $\cdots$ | $\mathbf{X}_{[m]2}$ |
| $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| $\mathbf{X}_{[1]K}$ | $\mathbf{X}_{[2]K}$ | $\cdots$ | $\mathbf{X}_{[m]K}$ |

$$\widehat{\mu}_V^j = \frac{1}{m}\sum_{h=1}^{m} \bar{X}_{[h]}^j, \quad \bar{X}_{[h]}^j = \frac{1}{n}\sum_{i \epsilon s_h} X_{[h]i}^j$$

## Theorem

*In MVSR, $\widehat{\mu}_V^j$ is an unbiased estimator for $\mu^j$ and*

$$V(\widehat{\mu}_V^j) = \frac{1}{nm}(\sigma_j^2 - \frac{(1-\frac{n}{K})}{m}\sum_{h=1}^{m}(\mu_{[h]}^j - \mu^j)^2)$$

*and if $X_i^j = \mu^j + \rho_{1j}\frac{\sigma_j}{\sigma_1}(X_i^1 - \mu^1) + \varepsilon_i$ where $\varepsilon \perp X^1$, then*

$$V(\widehat{\mu}_V^j) = \frac{1}{nm}(\sigma_j^2 - \frac{(1-\frac{n}{K})}{m}\rho_{1j}^2\sum_{h=1}^{m}(\mu_{(h)}^j - \mu^j)^2)$$

*and at the end an unbiased estimator for $V(\widehat{\mu}_V^j)$ is*

$$\frac{K-1}{m(mK-1)}\sum_{h=1}^{m}\frac{1}{n(n-1)}\sum_{i\epsilon s_h}(X_{[h]i}^j - \bar{X}_{[h]}^j)^2 + \frac{1}{m(mK-1)}\sum_{h=1}^{m}(\bar{X}_{[h]}^j - \widehat{\mu}_V^j)^2$$

## Multivariate variable sampling based on Partial Order Theory

# Part 4

# Linear Extensions

Linear extensions are different projections of the partial order into a complete order that respect all the relations in the partial order set.

- An example with $R = 2$ and $m = 5$

| | $X^1$ | $X^2$ |
|---|---|---|
| a | 0 | 1 |
| b | 2 | 1 |
| c | 1 | 2 |
| d | 3 | 3 |
| e | 0 | 4 |

| LE1 | LE2 | LE3 | LE4 | LE5 | LE6 | LE7 | LE8 |
|---|---|---|---|---|---|---|---|
| d | d | d | e | d | d | d | e |
| c | c | e | d | b | b | e | d |
| b | e | c | c | c | e | b | b |
| e | b | b | b | e | c | c | c |
| a | a | a | a | a | a | a | a |

# Linear Extensions

- Two Designs to Rank Based on Partial Order Sets

I) **CPOR**; An example of **CPOR** with $R = 2$ and $m = 5$

| | $X^1$ | $X^2$ |
|---|---|---|
| a | 0 | 1 |
| b | 2 | 1 |
| c | 1 | 2 |
| d | 3 | 3 |
| e | 0 | 4 |

| LE1 | LE2 | LE3 | LE4 | LE5 | LE6 | LE7 | LE8 |
|---|---|---|---|---|---|---|---|
| d | d | d | e | d | d | d | e |
| c | c | e | d | b | b | e | d |
| b | e | c | c | c | e | b | b |
| e | b | b | b | e | c | c | c |
| a | a | a | a | a | a | a | a |

| | mean height | rounded height |
|---|---|---|
| a | 1 | 1 |
| b | 2.875 | 3 |
| c | 2.875 | 3 |
| d | 4.75 | 5 |
| e | 3.5 | 4 |

| strata | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| | a | | b | e | d |
| | | | c | | |

# CPOR

I) **CPOR**;

| stratum 1 | stratum 2 | $\cdots$ | stratum m |
|---|---|---|---|
| $\mathbf{X}_{\{1\}1}$ | $\mathbf{X}_{\{2\}1}$ | $\cdots$ | $\mathbf{X}_{\{m\}1}$ |
| $\mathbf{X}_{\{1\}2}$ | $\mathbf{X}_{\{2\}2}$ | $\cdots$ | $\mathbf{X}_{\{m\}2}$ |
| $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| $\vdots$ | $\vdots$ | $\cdots$ | $\mathbf{X}_{\{m\}K_m}$ |
| $\mathbf{X}_{\{1\}K_1}$ | $\vdots$ | | |
| | $\mathbf{X}_{\{2\}K_2}$ | | |

# CPOR

Theorem

*In CPOR, $\widehat{\mu}_P^j$ is an unbiased estimator for $\mu^j$, where*

$$\widehat{\mu}_P^j = \sum_{h=1}^{m} W_h \bar{X}_{\{h\}}^j, \quad W_h = \frac{K_h}{Km}; \;\; \bar{X}_{\{h\}}^j = \frac{1}{n_h} \sum_{i \epsilon s_h} X_{\{h\}i}^j$$

$$Var(\widehat{\mu}_P^j)??!!$$

# Linear Extensions

- Two Designs to Rank Based on Partial Order Sets

II) **RPOR**; An example of **RPOR** with $R = 2$ and $m = 5$

| | $X^1$ | $X^2$ |
|---|---|---|
| a | 0 | 1 |
| b | 2 | 1 |
| c | 1 | 2 |
| d | 3 | 3 |
| e | 0 | 4 |

| LE1 | LE2 | LE3 | LE4 | LE5 | LE6 | LE7 | LE8 |
|---|---|---|---|---|---|---|---|
| d | d | d | e | d | d | d | e |
| c | c | e | d | b | b | e | d |
| b | e | c | c | c | e | b | b |
| e | b | b | b | e | c | c | c |
| a | a | a | a | a | a | a | a |

| strata | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| | a | c | b | e | d |

# RPOR

Theorem

*In RPOR,* $\widehat{\mu}_R^j$ *is an unbiased estimator for* $\mu^j$ *where*

$$\widehat{\mu}_R^j = \frac{1}{m}\sum_{h=1}^{m}\bar{X}_{[h\}}^j; \quad \bar{X} = \frac{1}{n}\sum_{i\in s_h} X_{[h\}i}^j$$

*with variance*

$$V(\widehat{\mu}_R^j) = \frac{\sigma_j^2}{Km} + \frac{1}{m^2}\sum_{h=1}^{m}\frac{1-\frac{n}{K}}{n}E_M(\frac{1}{Q}\sum_{q=1}^{Q}S_{[h\}qjK}^2).$$

*where* $q = 1, 2, ..., Q$ *are all the possible combinations of LEs,and*

$$\widehat{V}(\widehat{\mu}_R^j) = \frac{1}{nm(Km-1)}[\sum_{h=1}^{m}\sum_{i\epsilon s_{[h\}}}(X_{[h\}i}^j - \widehat{\mu}_R^j)^2 + (K-n)\sum_{h=1}^{m}s_{[h\}j}^2].$$

# Simulations

**Part 5**

# How to compare designs?

To evaluate and compare the efficiency of the designs, we calculate

$$\text{Efficiency}(\widehat{\mu}_.) = \frac{V(\bar{y})}{\text{MSE}(\widehat{\mu}_.)}$$

where

- $\bar{y}$ is the sample mean of a simple random sample
- $\widehat{\mu}_.$ stands for
  - $\widehat{\mu}_V$ (MVSR design)
  - $\widehat{\mu}_C$ (CPOR design)
  - $\widehat{\mu}_R$ (RPOR design)
- MSE indicates mean square error

# Data for the simulations

- Bivariate Normal Distribution.
- Data of a Medical Flower; Chamomile (Panahbehagh and Bruggemann 2017).
- Data of Chemical Pollution.

# Bivariate Normal, $\mu^1 = \mu^2 = 0$, $\sigma_1 = \sigma_2 = 1$

| $m$ | $K$ | $n$ | $\rho$ | $\widehat{\mu}_V^1$ | $\widehat{\mu}_V^2$ | $\widehat{\mu}_C^1$ | $\widehat{\mu}_C^2$ | $\widehat{\mu}_R^1$ | $\widehat{\mu}_R^2$ |
|---|---|---|---|---|---|---|---|---|---|
| 3 | 12 | 4 | 0.3 | 1.49 | 1.00 | 1.16 | 1.12 | 1.12 | 1.11 |
| | | | 0.5 | 1.45 | 1.05 | 1.20 | 1.21 | 1.17 | 1.18 |
| | | | 0.7 | 1.47 | 1.17 | 1.27 | 1.30 | 1.26 | 1.26 |
| | | | 0.9 | 1.49 | 1.35 | 1.41 | 1.42 | 1.39 | 1.41 |
| | | 6 | 0.3 | 1.31 | 1.01 | 1.13 | 1.10 | 1.10 | 1.07 |
| | | | 0.5 | 1.30 | 1.05 | 1.16 | 1.14 | 1.13 | 1.11 |
| | | | 0.7 | 1.33 | 1.13 | 1.23 | 1.23 | 1.19 | 1.20 |
| | | | 0.9 | 1.32 | 1.23 | 1.31 | 1.31 | 1.29 | 1.27 |

# Data of a medical flower; Chamomile

$X^1$ : Flower dry weight (F)

$Y^1$ : Flower height; $\rho(X^1, Y^1) = 0.78$

$X^2$ : Essence (E)

$Y^2$ : Number of petals; $\rho(X^2, Y^2) = 0.71$

- Also $\rho(Y^1, Y^2) = 0.77$

# Data of a medical flower; Chamomile

| K | m | n | $\widehat{\mu}_V^1$(F) | $\widehat{\mu}_V^2$(E) | $\widehat{\mu}_C^1$(F) | $\widehat{\mu}_C^2$(E) | $\widehat{\mu}_R^1$(F) | $\widehat{\mu}_R^2$(E) |
|---|---|---|---|---|---|---|---|---|
| 5 | 3 | 2 | 1.40 | 1.11 | 1.32 | 1.17 | 1.28 | 1.14 |
| | | 3 | 1.23 | 1.07 | 1.18 | 1.06 | 1.18 | 1.07 |
| | | 4 | 1.10 | 1.04 | 1.09 | 1.04 | 1.09 | 1.05 |
| | 5 | 2 | 1.63 | 1.18 | 1.45 | 1.24 | 1.45 | 1.23 |
| | | 3 | 1.35 | 1.10 | 1.26 | 1.11 | 1.24 | 1.13 |
| | | 4 | 1.14 | 1.05 | 1.11 | 1.06 | 1.11 | 1.06 |
| | 7 | 2 | 1.77 | 1.19 | 1.55 | 1.27 | 1.53 | 1.27 |
| | | 3 | 1.40 | 1.10 | 1.29 | 1.12 | 1.30 | 1.14 |
| | | 4 | 1.17 | 1.06 | 1.13 | 1.07 | 1.13 | 1.07 |
| 7 | 3 | 3 | 1.36 | 1.09 | 1.26 | 1.15 | 1.25 | 1.13 |
| | | 5 | 1.15 | 1.06 | 1.13 | 1.07 | 1.12 | 1.08 |
| | | 6 | 1.09 | 1.03 | 1.07 | 1.03 | 1.06 | 1.03 |
| | 5 | 3 | 1.58 | 1.16 | 1.43 | 1.22 | 1.40 | 1.20 |
| | | 5 | 1.23 | 1.07 | 1.18 | 1.08 | 1.17 | 1.09 |
| | | 6 | 1.10 | 1.03 | 1.08 | 1.04 | 1.08 | 1.04 |
| | 7 | 3 | 1.71 | 1.19 | 1.53 | 1.25 | 1.51 | 1.24 |
| | | 5 | 1.26 | 1.08 | 1.22 | 1.09 | 1.20 | 1.11 |
| | | 6 | 1.12 | 1.04 | 1.09 | 1.05 | 1.09 | 1.05 |
| | Ave. | | 1.32 | 1.09 | 1.24 | 1.12 | 1.23 | 1.12 |

# Chemical Pollution; Pb, Cd, Zn and S

- First Case; high correlation
  $X^1$ : Pb
  $X^2$ : Zn;
  $$\rho(X^1, X^2) = 0.60$$
- Second Case; low correlation
  $X^1$ : Cd
  $X^2$ : S;
  $$\rho(X^1, X^2) = 0.05$$

# $\rho(Pb, Zn) = 0.60$ and $\rho(Cd, S) = 0.05$

| m | K | n | $\widehat{\mu}_V^1$ Pb | $\widehat{\mu}_V^2$ Zn | $\widehat{\mu}_C^1$ Pb | $\widehat{\mu}_C^2$ Zn | $\widehat{\mu}_R^1$ Pb | $\widehat{\mu}_R^2$ Zn | $\widehat{\mu}_V^1$ Cd | $\widehat{\mu}_V^2$ S | $\widehat{\mu}_C^1$ Cd | $\widehat{\mu}_C^2$ S | $\widehat{\mu}_R^1$ Cd | $\widehat{\mu}_R^2$ S |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 5 | 2 | 1.32 | 1.11 | 1.13 | 1.21 | 1.12 | 1.15 | 1.36 | 1.01 | 1.13 | 1.09 | 1.12 | 1.09 |
| | | 4 | 1.11 | 1.02 | 1.06 | 1.03 | 1.05 | 1.03 | 1.11 | 1.01 | 1.05 | 1.00 | 1.05 | 1.03 |
| | 7 | 2 | 1.41 | 1.11 | 1.15 | 1.16 | 1.10 | 1.12 | 1.41 | 0.99 | 1.16 | 1.08 | 1.11 | 1.07 |
| | | 4 | 1.25 | 1.08 | 1.11 | 1.10 | 1.09 | 1.10 | 1.16 | 1.01 | 1.00 | 1.01 | 1.03 | 1.06 |
| | 10 | 2 | 1.59 | 1.16 | 1.24 | 1.25 | 1.16 | 1.17 | 1.53 | 1.04 | 1.22 | 1.12 | 1.19 | 1.11 |
| | | 4 | 1.38 | 1.11 | 1.15 | 1.17 | 1.12 | 1.14 | 1.29 | 1.02 | 1.16 | 1.09 | 1.07 | 1.07 |
| 5 | 5 | 2 | 1.61 | 1.21 | 1.23 | 1.27 | 1.18 | 1.23 | 1.53 | 1.02 | 1.19 | 1.12 | 1.15 | 1.10 |
| | | 4 | 1.15 | 1.04 | 1.05 | 1.05 | 1.04 | 1.06 | 1.13 | 1.01 | 1.05 | 1.02 | 1.04 | 1.02 |
| | 7 | 2 | 1.69 | 1.21 | 1.23 | 1.30 | 1.21 | 1.26 | 1.62 | 1.00 | 1.23 | 1.11 | 1.18 | 1.08 |
| | | 4 | 1.35 | 1.14 | 1.14 | 1.15 | 1.13 | 1.16 | 1.33 | 1.00 | 1.12 | 1.04 | 1.09 | 1.04 |
| | 10 | 2 | 1.93 | 1.28 | 1.31 | 1.37 | 1.24 | 1.34 | 1.78 | 0.99 | 1.28 | 1.13 | 1.21 | 1.12 |
| | | 4 | 1.56 | 1.21 | 1.20 | 1.28 | 1.17 | 1.26 | 1.47 | 1.03 | 1.15 | 1.12 | 1.13 | 1.11 |
| 7 | 5 | 2 | 1.69 | 1.21 | 1.26 | 1.31 | 1.21 | 1.28 | 1.63 | 1.04 | 1.26 | 1.18 | 1.21 | 1.19 |
| | | 4 | 1.16 | 1.10 | 1.09 | 1.10 | 1.07 | 1.12 | 1.15 | 0.99 | 1.06 | 1.04 | 1.06 | 1.01 |
| | 7 | 2 | 1.90 | 1.28 | 1.27 | 1.35 | 1.29 | 1.32 | 1.85 | 1.03 | 1.30 | 1.12 | 1.28 | 1.09 |
| | | 4 | 1.45 | 1.19 | 1.19 | 1.19 | 1.17 | 1.20 | 1.37 | 1.02 | 1.17 | 1.07 | 1.17 | 1.07 |
| | 10 | 2 | 2.20 | 1.36 | 1.40 | 1.49 | 1.39 | 1.49 | 1.99 | 1.02 | 1.40 | 1.19 | 1.31 | 1.15 |
| | | 4 | 1.66 | 1.30 | 1.27 | 1.36 | 1.25 | 1.34 | 1.62 | 1.02 | 1.25 | 1.13 | 1.24 | 1.12 |
| | A. | | 1.52 | 1.17 | 1.19 | 1.23 | 1.17 | 1.21 | 1.46 | 1.01 | 1.18 | 1.09 | 1.15 | 1.09 |

# Why Ranking Based on Partial Order?

# Why Ranking Based on Partial Order?

- To consider all the variables simultaneously.

# Why Ranking Based on Partial Order?

- To consider all the variables simultaneously.
- To avoid ambiguity of weighting variables.

# Why Ranking Based on Partial Order?

- To consider all the variables simultaneously.
- To avoid ambiguity of weighting variables.
- To avoid needing any extra initial sample.

# References

Al-Saleh, M. and Zheng, G. (2002) Estimation of bivariate characterstics using ranked set sampling. *Australian & New Zealand Jurnal of Statistics*, **44**, 221–232.

Bruggemann, R. and Carlsen, L. (2011) An Improved Estimation of Averaged Ranks of Partial Orders. *MATCH Comm.Math.Comput.Chem.* **65**, 383–414.

Brggemann, R., Kaune, A. and Voigt. K. (1996) Vergleichende kologische Bewertung von Regionen in Baden- Wrttemberg. Pages 455-467 in Landesanstalt fr Umweltschutz Baden-Wrttemberg, ed. 4.Statuskolloquium, Projekt "Angewandte kologie" Nr. 16. Przis-Druck Karlsruhe, Karlsruhe.

Bruggemann, R., Mucha, H.-J. and Bartel, H.-G. (2013) Ranking of Polluted Regions in South West Germany Based on a Multi-indicator System. *MATCH Commun. Math. Comput. Chem.*, **69**,433–462.

Bruggemann, R. and Patil, G. P. (2011) *Ranking and Prioritization with Multi- Indicator Systems, Introduction to Partial Order and Its Applications*, Springer, New York.

Bruggemann, R., Pudenz S., Voigt K., Kaune A., and Kreimes K. (1999) An algebraic/graphical tool to compare ecosystems with respect to their pollution. IV: Comparative regional analysis by Boolean arithmetics. *Chemosphere*, **38**, 2263–2279.

Bruggemann, R., Sorensen, P. B., Lerche, D. and Carlsen, L. (2004) Estimation of Averaged Ranks by a Local Partial Order Model. *J. Chem. Inf. Comp. Sc.* **44**, 618–625.

Bruggemann, R., Voigt, K., Kaune, A., Pudenz, S., Komossa, D., and Friedrich, J. (1998) Vergleichende kologische Bewertung von Regionen in Baden- Wrttemberg GSF-Bericht 20/98. GSF, Neuherberg.

Bruggemann, R., Welzl, G., and Voigt, K. (2003) Order Theoretical Tools for the Evaluation of Complex Regional Pollution Patterns. *J. Chem. Inf. Comp. Sc.*, **43**, 1771–1779.

# References


- Chen, Z., Bai, Z. and Sinha, B. (2004) *Ranked set sampling: theory and applications. Lecture Notes in Statistics*, Springer, New York.
- Chen, Z. and Shen, L. (2003) Two-layer ranked set sampling with concomitant variables. *Journal of Statistical Planning and Inference*, **115**, 45–57.
- David, H. A. and Nagaraja, H. N. (2003) *Order Statistic*, third ed. Wiley, New York.
- De Loof, K., De Baets, B. and De Meyer, H. (2011) Approximation of Average Ranks in Posets. *MATCH Commun. Math. Comput. Chem.*, **66**, 219–229.
- Environmental Protection Agency Baden-Wurttemberg, (1994) Signale aus der Natur 10 Jahre Okologisches Wirkungskataster Baden-Wurttemberg. Kraft Druck GmbH, Ettlingen
- McIntyre, G. A. (1952) A method of unbiased selective sampling. using ranked sets. *Australian Journal of Agricultural Research*, **3**, 385–390.
- Norris, R. C., Patil, G. P. and Sinha, A. K. (1995) Estimation of multiple characteristics by ranked set sampling methods. *Coenoses*, **10**, 95–111.
- Panahbehagh, B., Bruggemann R., Parvardeh, A., Salehi, M. and Sabzalian, M. R. (2017) An unbalanced ranked set sampling to get more than one sample from each set. *Journal of Survey Statistics and Methodology*, smx026, https://doi.org/10.1093/jssam/smx026
- Patil, G. P., Sinha, A. K. and Taillie, C. (1994) Ranked set sampling for multiple characteristics. *International Journal of Ecology and Environmental Sciences*, **20**, 94–109.
- Patil, G. P., Sinha, A. K. and Taillie, C. (1999) Ranked set sampling: A bibliography. *Environmental and Ecological Statistics*, **6**, 91–98.

# References


Patil, G. P., Sinha, A. K. and Taillie, C. (1994) *Ranked set sampling*, in Handbook of Statistics, Environmental Statistics, Vol. 12, G.P. Patil and C.R. Rao, eds, NorthHolland, Amsterdam.

Ridout, M. S. (2003) On ranked set sampling for multiple characterestics. *Environmental and Ecological Statistics*, **10**, 225–262.

Samawi, H. M. (1996) Stratified ranked set sample. *Pakistan Journal of Statistics*, **12**, 9–16.

Sarndal, C. E., Swensson, B. and Wretman, J. (1992), *Model Assisted Survey Sampling*, New York, Springer.'

Yang, S. S. (1977) General distribution theory of the concomitants of order statistics. *The Annals of Statistics*, **5**, 996–1002.

# Thanks For Your Attention

*brg_home@web.de*
*bardia.panah@gmail.com*
*panahbehagh@khu.ac.ir*