

Ranked set sampling

Environmental monitoring and assessment typically requires observational data, as opposed to data obtained from controlled experiments. This is true whether we are assessing the extent of soil contamination at a one-acre site, or some measure of forest resources over the Pacific Northwest region of the US. Obtaining such data requires identification of sample units to represent the population of concern, followed by selection of particular units to quantify the characteristic(s) of interest. Sample units are generally the smallest possible units of measurement, such as plots, soil cores, individuals, etc., while typical characteristics of interest include **biomass**, chemical concentrations or 'head counts'.

Typically the most expensive part of this process is laboratory analysis, while identification of potential sample units is a comparatively simple matter. We can therefore achieve great observational economy if we are able to identify a large number of sample units to represent the population of interest, yet only have to quantify a carefully selected subsample.

This potential for observational economy was recognized for estimating mean pasture and forage yields in the early 1950s when McIntyre [18] proposed a method, later coined *ranked set sampling* (RSS) by Halls and Dell [9], and currently under active investigation in various quarters.

As a simple introduction to the concept of RSS, consider the following example.

Suppose we wish to estimate the mean height of students at a university from a random sample of three students. Furthermore, in order to acknowledge the inherent uncertainty, we need to present this estimate as a confidence interval within which we expect the true population mean to lie with desired confidence.

The simplest way to obtain our sample is to randomly select three students from the university's population, then measure their heights. While the arithmetic average of the three heights is an unbiased point estimate of the population mean, the associated confidence interval can be very large, reflecting the high degree of uncertainty with estimating a large population mean from only three measurements. This is because we have no control over which individuals of the population enter the sample. For example, we may happen to choose two very short

people and one very tall; or we may choose three very tall people. The only way to overcome such a problem with simple random sampling (SRS – see **Sampling, environmental**) is to increase the sample size.

On the other hand, we may obtain a ranked set sample. To do this, we may randomly invite three students to breakfast and visually rank them with respect to height. We then select the student we believe is shortest and actually measure his or her height. Repeating this process at lunch, we then select the middle ranked person and, as such, select the tallest ranked person at dinner. The resulting measurements of student heights constitute a ranked set sample. As with the SRS measurements, the arithmetic average of the RSS measurements provides an unbiased point estimate of the population mean; however, the associated confidence interval can potentially be much smaller than that obtained with SRS measurements, thus reflecting decreased uncertainty. This encouraging feature results because measurements obtained through RSS are likely to be more regularly spaced than those obtained through SRS, and therefore are more representative of the population. Amazingly, the RSS procedure induces stratification of the whole population at the sample level; in effect, we are randomly sampling from the subpopulations of predominantly short, medium and tall students without having to construct the subpopulation strata. Each subpopulation has its own distribution, as visualized in Figure 1, where we see how the parent population gets partitioned into subpopulations.

McIntyre's proposal does not appear to have been applied for over a decade, after which **forestry** and range researchers continued to discover the effectiveness of RSS (see [5], [8], [9], [11] and [17]).

Theoretical investigations by Dell and Clutter [7] showed that, regardless of ranking errors, the RSS estimator of a population mean is unbiased and at least as precise as the SRS estimator with the same number of quantifications. David and Levine [6] investigated the case where ranking is done by a numerical covariate.

Furthermore, RSS also provides more precise estimators of the variance [24], the cumulative **distribution function** [26], and at times the Pearson correlation coefficient [25]. For an annotated bibliography with a historical perspective, see [14].

2 Ranked set sampling

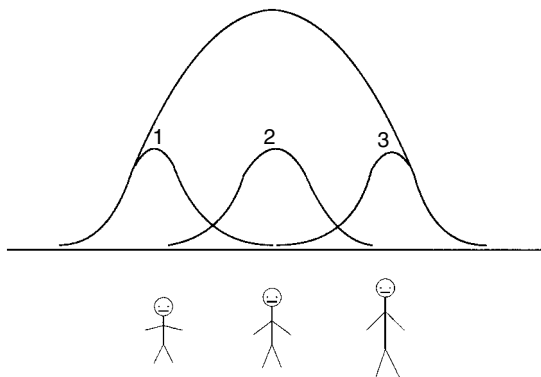


Figure 1 Frequency distributions of heights of different ranks superimposed on population frequency distribution of all heights – a schematic diagram

What is RSS?

Description

As mentioned above, to create ranked sets we must partition the selected first phase sample into sets of equal size. In order to plan an RSS design, we must therefore choose a set size that is typically small, around three or four, to minimize ranking error. Call this set size m , where m is the number of sample units allocated to each set. Now proceed as follows.

- Step 1: randomly select m^2 sample units from the population.
- Step 2: allocate the m^2 selected units as randomly as possible into m sets, each of size m .
- Step 3: without yet knowing any values for the variable of interest, rank the units within each set based on a perception of relative values for this variable. This may be based on personal judgment or done with measurements of a covariate that is correlated with the variable of interest.
- Step 4: choose a sample for actual analysis by including the smallest ranked unit in the first set, then the second smallest ranked unit in the second set, continuing in this fashion until the largest ranked unit is selected in the last set.
- Step 5: repeat steps 1 through 4 for r cycles until the desired sample size, $n = mr$, is obtained for analysis.

As an illustration, consider the set size $m = 3$ with $r = 4$ cycles. This situation is illustrated in

| Cycle | Rank | | |
|-------|------|---|---|
| | 1 | 2 | 3 |
| 1 | ⊙ | . | . |
| | . | ⊙ | . |
| | . | . | ⊙ |
| 2 | ⊙ | . | . |
| | . | ⊙ | . |
| | . | . | ⊙ |
| 3 | ⊙ | . | . |
| | . | ⊙ | . |
| | . | . | ⊙ |
| 4 | ⊙ | . | . |
| | . | ⊙ | . |
| | . | . | ⊙ |

Figure 2 A ranked set sample design with set size $m = 3$ and number of sampling cycles $r = 4$. Although 36 sample units have been selected from the population, only the 12 circled units are actually included in the final sample for quantitative analysis

Figure 2, where each row denotes a judgment-ordered sample within a cycle, and the units selected for quantitative analysis are circled. Note that 36 units have been randomly selected in four cycles; however, only 12 units are actually analyzed to obtain the ranked set sample of measurements.

Obtaining a sample in this manner maintains the unbiasedness of SRS; however, by incorporating ‘outside’ information about the sample units, we are able to contribute a structure to the sample that increases its representativeness of the true underlying population.

If we quantified the same number of sample units, $mr = 12$, by a simple random sample, then we have no control over which units enter the sample. Perhaps all the 12 units would come from the lower end of the range, or perhaps most would be clustered at the low end while one or two units would come from the middle or upper range. With SRS, the only way to increase the prospect of covering the full range of possible values is to increase the sample size. With RSS, however, we increase the representativeness with a fixed number of sample units, thus saving considerably on quantification costs.

With the ranked set sample thus obtained, it can be shown that unbiased estimators of several important population parameters can be calculated, including the mean and, in case of more than one sampling cycle, the variance.

Ranking Criteria

A real key to success lies with step 3 in the above procedure – ranking. This may be based on visual inspection or other expert opinion about the sample units. For example, a field-seasoned range scientist or forester may readily be able to rank three or four quadrats of grass with respect to overall volume or mass. Meanwhile, a hazardous waste site inspector may be able to reliably rank areas of soil with respect to concentrations of a toxic contaminant, based on features like surface staining, discoloration, or the appearance of stressed vegetation.

On the other hand, if another characteristic is available that is highly correlated with the characteristic of interest but costs much less to obtain, then we may rank by the values of such a ‘covariate’. For example, reflectance intensity of near-infrared electromagnetic radiation, as recorded in a remotely sensed digital image (see **Remote sensing**), is directly proportional to vegetation concentration on the ground. Another example might be to measure total organic halides (TOX) in soil in order to rank soil sampling units with respect to the concentration of volatile organic solvents (see **Soil surveys**). As an indicator variable, TOX is much less expensive to measure than specific organic compounds.

Robustness of the Procedure

Several questions now arise.

1. What if the distribution of sample measurements is skewed, symmetric, or essentially unknown?
2. What if the sample units are not randomly allocated into sets?
3. How does error in ranking affect results?

First of all, while independent (random) and identically distributed sample measurements obtained through perfect ranking may lead to *optimum* performance of RSS, no matter how much these desirable characteristics are deviated from, the sampling efficiency will never be worse than with SRS using the same number of quantifications. In fact, when efficiency is expressed as the relative precision (RP) such that

$$RP = \frac{\text{variance of sample average with SRS}}{\text{variance of sample average with RSS}} \quad (1)$$

it can be shown that the bounds of this RP are

$$1 \leq RP \leq \frac{m+1}{2} \quad (2)$$

where m is the set size. Since RP cannot be less than one, the RSS protocol cannot be worse than the SRS protocol.

Variations of the Basic Protocol

Unequal Allocation of Sample Units

The performance of RSS decreases as the underlying distribution of the characteristic of interest becomes increasingly skewed. McIntyre [18] originally suggested that this problem may be overcome by allocating sample units into ranks in proportion to the standard deviation of each rank. This is the same approach as used in stratified random sampling, known as *Neyman allocation*, and would indeed be optimal if we had reliable prior estimates of the rank standard deviations. An example of unequal allocation is displayed in Figure 3. Here we have the same set size, $m = 3$, and sample size, $n = 12$, as in the earlier example of equal allocation; however, the number of sampling cycles is adjusted so as to yield the desired unequal allocation of samples.

Unequal allocation can actually increase the performance of RSS above and beyond that achievable with standard equal allocation; however, if not properly applied, the performance of RSS can be worse than the performance of SRS. Actually, the bounds on RP with unequal allocation become

$$0 \leq RP \leq m \quad (3)$$

indicating that, with appropriate unequal allocation, the RP may even increase to a level of m , and not just $(m+1)/2$ as in the case with equal allocation.

Although an optimal RSS design would allocate samples into ranks in direct proportion to the rank standard deviations, we rarely know the standard deviations beforehand. We do know, however, that the distributions of many environmental and ecological variables are skewed towards the right, meaning that while most values are clustered around a median, a few much larger values are usually present. This skewness can actually be exploited to increase the precision beyond that obtained with

4 Ranked set sampling

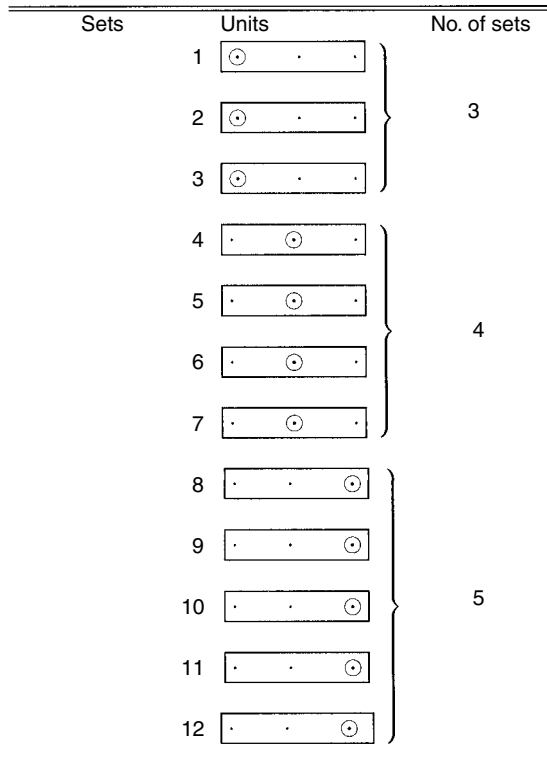


Figure 3 RSS with unequal allocation: circles indicate sample units chosen for quantification

RSS under equal allocation, because standard deviations usually tend to increase with increasing rank values for right-skewed distributions. With some idea of the degree of skewness, Kaur et al. [13] have devised a rule-of-thumb for allocating sample units into ranks that performs closely to the optimal Neyman allocation. Therefore, distributions of many environmental and ecological variables may actually lend themselves well to being estimated with very high precision relative to that obtainable through SRS.

Combining with Line-intercept Sampling

A common field sampling method for ecological assessments is to include sample units that one encounters along a line (transect) that is randomly selected within a two-dimensional area of interest (see **Line-transect sampling**). Units are typically members of a plant or animal species.

Often the number of sample units identified is too numerous to select every one for quantification, especially if measurements are destructive, such as with cutting vegetation for weighing. If the initially identified sample units are treated as a larger first phase sample, $n' = m^2r$, then the RSS protocol can be applied to select a smaller subsample, $n = mr$, for actual quantification. For example, consider a single sampling cycle when the set size m equals three for estimating the **biomass** of shrubs in a given area. A line transect for such a situation may be visualized as in Figure 4.

Such an RSS-based line intercept sample has been found to produce more precise, and still unbiased, estimators of the population mean, size, total and cover, compared with the SRS-based line-intercept sample [19].

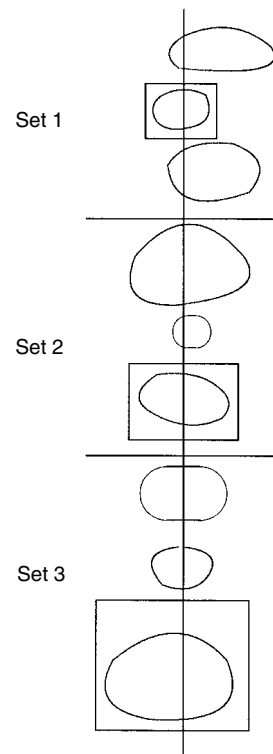


Figure 4 Aerial view of a line transect intercepting shrubs. For set size $m = 3$, nine shrubs are partitioned into three sets of three. Using apparent shrub size for ranking with respect to biomass, the shrubs taken for analyses include the smallest ranked in the first set, the second smallest ranked in the second set, and the largest ranked in the third set

Improved Compositing of Samples

Consider a situation that calls for **composite sampling**. If our primary objective is **classification** of the individual samples used to form the composites and/or identification of those individual samples that constitute an upper percentile with respect to the characteristic of interest, then we will need to retest certain individual samples. Since the purpose of composite sampling is to minimize the number of analytical tests required, we obviously want to minimize the extent of retesting individual samples. Maximizing the homogeneity within the composites will minimize the necessary number of retests. Therefore, it is desired to form composites from individual sample units that are as much alike as possible.

As pointed out by Patil et al. [21], one can increase the chances of obtaining maximum homogeneity within composites by forming composites from samples identified to be in the same rank as conceptualized in Figure 5. The RSS protocol can thus be combined with composite sampling to achieve even greater observational economy than composite sampling alone.

Geographic Information Systems (GIS) and RSS

With the availability of computerized GIS, ranking prospective sample locations across a landscape

may be done rapidly prior to expensive field visits, thus allowing RSS to be applied to large-scale surveys to obtain a more precise estimate at reduced cost. If prospective locations are selected at random from across a region and allocated to a set, then each location can be referenced to data layers in a GIS and, based on a derived ranking index, each member of the set can be ranked relative to each other. This merger of GIS and RSS has been recommended by Johnson and Myers [12] and by Myers et al. [20].

Rapid Damage Assessment

Following a catastrophic event such as flooding or fire (*see Natural disasters*), those in charge of management and planning of natural or cultural resources need rapid assessments of the spatial extent and magnitude of damage. The authors cited above recommend that the combination of RSS and GIS can result in rapid mobilization of available information to design a very efficient field sampling strategy.

For more information on RSS, see the special issue of *Environmental and Ecological Statistics* (Vol. 6, 1999), and in particular the articles by Ross and Stokes [23], Kvam and Tiwari [15], Yu et al. [27], Lavine [16], Barnett [2], Aragon et al. [1], Patil et al. [22], Hartlaub and Wolfe [10], Baretto and Barnett [3] and Chen [4].

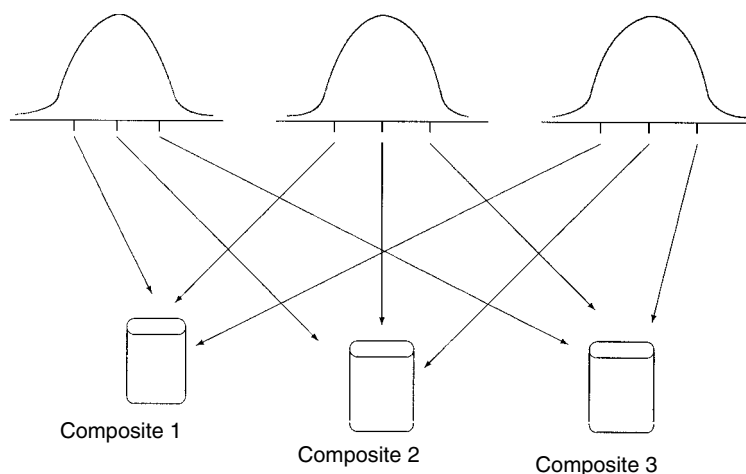


Figure 5 Formation of three composites from three ranked set samples, each with a set size of three. Homogeneity is maximized (variability is minimized) within each composite by forming composites from equally ranked samples

6 Ranked set sampling

References

- [1] Aragon, M.E.D., Patil, G.P. & Taillie, C. (1999). A performance indicator for ranked set sampling using ranking error probability matrix, *Environmental and Ecological Statistics* **6**, 75–89.
- [2] Barnett, V. (1999). Ranked set sample design for environmental investigations, *Environmental and Ecological Statistics* **6**, 59–74.
- [3] Barreto, M.C.M. & Barnett, V. (1999). Best linear unbiased estimators for the simple linear regression model using ranked set sampling, *Environmental and Ecological Statistics* **6**, 119–133.
- [4] Chen, Z. (1999). Density estimation using ranked-set sampling data, *Environmental and Ecological Statistics* **6**, 135–146.
- [5] Cobby, J.M., Ridout, M.S., Bassett, P.J. & Large, R.V. (1985). An investigation into the use of ranked set sampling on grass and grass-clover swards, *Grass and Forage Science* **40**, 257–263.
- [6] David, H.A. & Levine, D.N. (1972). Ranked set sampling in the presence of judgment error, *Biometrics* **28**, 553–555.
- [7] Dell, T.R. & Clutter, J.L. (1972). Ranked set sampling theory with order statistics background, *Biometrics* **28**, 545–553.
- [8] Evans, M.J. (1967). Application of ranked set sampling to regeneration surveys in areas direct-seeded to longleaf pine. Master's Thesis, School of Forestry and Wildlife Management, Louisiana State University, Baton Rouge.
- [9] Halls, L.K. & Dell, T.R. (1966). Trial of ranked set sampling for forage yields, *Forest Science* **12**, 22–26.
- [10] Hartlaub, B.A. & Wolfe, D.A. (1999). Distribution-free ranked-set sample procedures for umbrella alternatives in the m -sample setting, *Environmental and Ecological Statistics* **6**, 105–118.
- [11] Jewiss, O.R. (1981). Shoot development and number, in *Sward Measurement Handbook*, J. Hodgson, R.D. Baker, A. Davies, A.S. Laidlaw & J.D. Leaver, eds, The British Grassland Society, Hurley, pp. 93–114.
- [12] Johnson, G.D. & Myers, W.L. (1993). Potential of ranked-set sampling for disaster assessment. Paper presented at IUFRO S4.02 Conference on Inventory and Management Techniques in the Context of Catastrophic Events, June.
- [13] Kaur, A., Patil, G.P. & Taillie, C. (1994). Unequal allocation model for ranked set sampling with skew distributions, Technical Report 94-0930, Center for Statistical Ecology and Environmental Statistics, Department of Statistics, Pennsylvania State University, University Park.
- [14] Kaur, A., Patil, G.P., Sinha, A.K. & Taillie, C. (1995). Ranked set sampling: an annotated bibliography, *Environmental and Ecological Statistics* **2**, 25–54.
- [15] Kvam, P.H. & Tiwari, R.C. (1999). Bayes estimation of a distribution function using ranked set samples, *Environmental and Ecological Statistics* **6**, 11–22.
- [16] Lavine, M. (1999). The 'Bayesics' of ranked set sampling, *Environmental and Ecological Statistics* **6**, 47–57.
- [17] Martin, W.L., Sharik, T.L., Oderwald, R.G. & Smith, D.W. (1980). Evaluation of ranked set sampling for estimating shrub phytomass in Appalachian oak forests, Publication Number FWS-4-80, School of Forestry and Wildlife Resources, Virginia Polytechnic Institute and State University, Blacksburg.
- [18] McIntyre, G.A. (1952). A method for unbiased selective sampling, using ranked sets, *Australian Journal of Agricultural Research* **3**, 385–390.
- [19] Muttalak, H.A. & McDonald, L.L. (1992). Ranked set sampling and the line-intercept method: a more efficient procedure, *Biomedical Journal* **3**, 329–346.
- [20] Myers, W., Johnson, G.D. & Patil, G.P. (1994). Rapid mobilization of spatial/temporal information in the context of natural catastrophes, in *1994 Proceedings of the Section on Statistical Graphics*, American Statistical Association, Alexandria, pp. 25–31.
- [21] Patil, G.P., Sinha, A.K. & Taillie, C. (1994). Ranked set sampling, in *Handbook of Statistics, Environmental Statistics*, Vol. 12, G.P. Patil & C.R. Rao, eds, North-Holland, Amsterdam.
- [22] Patil, G.P., Sinha, A.K. & Taillie, C. (1999). Ranked set sampling: a bibliography, *Environmental and Ecological Statistics* **6**, 91–98.
- [23] Ross, N.P. & Stokes, L. (1999). Editorial: special issue on statistical design and analysis with ranked set samples, *Environmental and Ecological Statistics* **6**, 5–9.
- [24] Stokes, S.L. (1980). Estimation of variance using judgment ordered ranked set samples, *Biometrics* **36**, 35–42.
- [25] Stokes, S.L. (1980). Inferences on the correlation coefficient in bivariate normal populations from ranked set samples, *Journal of the American Statistical Association* **75**, 989–995.
- [26] Stokes, S.L. & Sager, T.W. (1988). Characterization of a ranked set sample with application to estimating distribution functions, *Journal of the American Statistical Association* **83**, 374–381.
- [27] Yu, P.L.H., Lam, K. & Sinha, B.K. (1999). Estimation of normal variance based on balanced and unbalanced ranked set samples, *Environmental and Ecological Statistics* **6**, 23–46.

(See also **Multistage design; Ranks**)

G.P. PATIL