Assessment of Random Samples versus Systematic Transects
Exercise 13

The objective of this exercise is to consider the efficiency of a random sampling design compared to a sampling design based on systematic transects. The exercise utilizes a single data set collected from grassland monitored during the grazing period. The objective of the sampling was to periodically monitor the remaining leaf length /stubble height as the grazing period proceeded. The data set is from mid-season.

Sampling was conducted using a random sampling design which located fifteen $0.1m^2$ quadrats within a key area. Within each quadrat, 3 forage plants were measured for leaf height to the nearest cm and averaged to yield a single leaf height value for each quadrat. For the transect example, the data set was partitioned into a systematic transect design so that a comparison could be made of the two techniques without changing the mean or the total variation contained in the original data set. The transect length was 25m with $0.1m^2$ quadrats placed at 5m intervals. The 3 transects were randomly located within the key area.

Sample Unit	Random Sample	Systematic Transect	Transect Unit Value
1	8	8	
2	8	8	
3	9	9	8
4	7	7	
5	8	8	
6	6	6	
7	5	5	
8	4	4	5
9	5	5	
10	5	5	
11	7	7	
12	6	6	
13	7	7	6.4
14	6	6	
15	6	6	

Data summary from the two sampling designs:

Attribute	Random (R) Design	Transect (T) Design	Contrast
Sample Size	15	3	T design sacrificed 12 df
Mean/Range	6.4 / 4 - 9	6.4 / 5 - 8	Same M, T less range
Std. Dev.	1.36	1.22	T Std. Dev 11% smaller
Coef. Of Variation (%)	21	18.9	T CV is 2 % units smaller
Conf. Interval (90%)	5.8 – 7.1	4.4 – 8.5	T CI is 3.3 times larger
Adeq. Sample Size 10% 0f Mean, 90% Conf.	14	31	T needs 2.2 times more samples (transects)

Discussion:

A **random sampling design** is described as one in which each potential sample has an equal opportunity for selection. In this case, our sampling boundary was defined by the key area, the population being sampled was all forage species located within the key area and our quadrat size was selected because it gave us an efficient platform (Exercise 12) for sampling leaf height within the population. Each quadrat placement was done randomly in a way that achieved interspersion throughout the population and an initial sampling of 15 quadrats was used as a first step toward achieving sample adequacy. The random sampling platform gives us an opportunity to estimate mean leaf height of forage species within the key area and the level of confidence that we should have in the mean estimate.

A **systematic transect design** is described as one in which each potential sample in the population does not have an equal chance of selection. In our example, several 25m transects have been placed randomly within the key area but the quadrats placed systematically along the transect which had a specific direction and distance between the quadrats. A fully systematic data set can be used to estimate a mean but cannot yield information about the level of confidence you have in the mean. In our example, the random data set resides in the transect values not the quadrats values.

Comparison One:

The previous discussion sets the stage for evaluating the efficiencies of each design. Given our discussion we can see that the sample size for the random data set is 15 and the random transect data set is 3. This yields 14 and 2 degrees of freedom and t-table values (90%) of 1.76 and 2.92, respectively. The loss of 12 degrees of freedom using the transect design results in a 66% increase in the t-value which translates into a loss of sensitivity in subsequent confidence assessments.

Comparison Two:

The mean estimates from each design are the same but the data range is reduced for the transect data. This reduction occurs because the transect value is a composite mean from 5 quadrats. A transect design will typically reduce data set variation when comparing a random design with a transect design.

Comparison Three:

A similar reduction can be seen between the transect CV and random CV. However this benefit is of little consequence when compared to the impact of lost degrees of freedom. Both the 90% confidence interval and the required sample size (10% of Mean, 90% confidence) show a low confidence in the transect mean and a dramatic increase in the number of transects that would be required to reach a level of confidence comparable to the random design.

Comparison Four:

We can assume that the amount of time required to measure leaf length in a quadrat is the same regardless of the design. Furthermore we can assume that the amount of time required to travel between random points is greater than the amount of time required to travel the 5 m distance between systematic quadrats. However, it should also be obvious that a significantly larger amount of time will be required to measure the 155 systematic quadrats compared to 14 random quadrats to achieve a similar level of confidence.

This exercise has raised the question of why transect data collection is used. If you are only interested in obtaining a mean estimate, the compositing of quadrats within a transect and the averaging of a few transects can give you adequate data for day to day management decisions. However, when you need to have scientific, regulatory, or litigation questions answered that require a measure of the adequacy of your sampling or the confidence that can be placed in your mean estimate, the random design will be the better choice.