Pacific Training on Sampling Methods for Producing Core Data Items for Agricultural and Rural Statistics

13-17 August, Suva, Fiji

Module 2: Review of Basics of Sampling Methods
Session 2.3: Systematic Sampling

By
Chris Ryan
Statistician (ESCAP Pacific Office)

UNITED NATIONS SIAP Statistical Institute for Asia and the Pacific

FAO FIAT PANIS

UNITED NATIONS ESCAP Economic and Social Commission for Asia and the Pacific

Topics Covered

- **Different Types of Systematic Sampling**
  - **Linear systematic sampling**
    - **With N/n = Integer**
    - **Without N/n = integer**
  - **Circular systematic sampling**

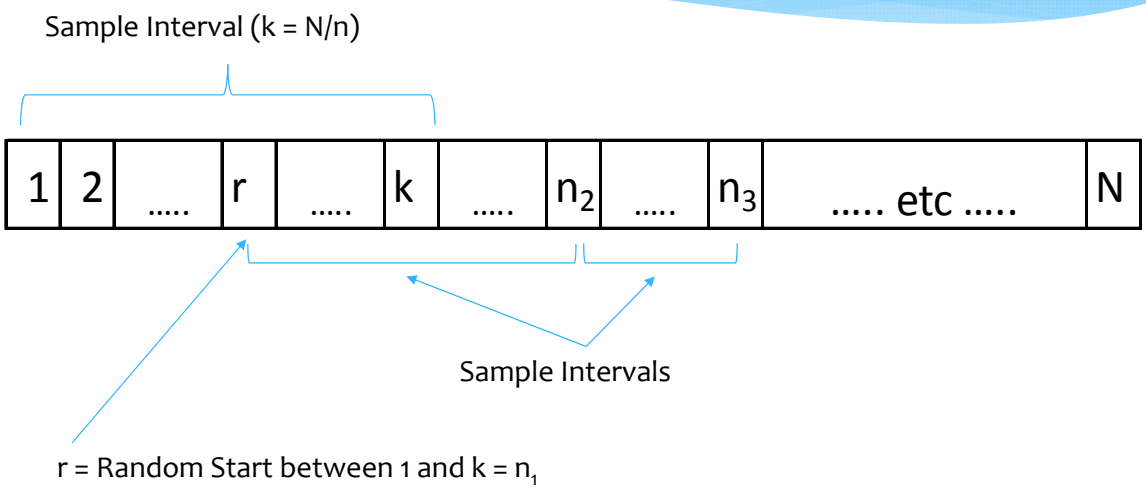
Linear systematic sampling

* Systematic Sampling (SYS), like SRS, involves selecting ***n*** sampling units from a population of ***N*** units
* Instead of randomly choosing the ***n*** units in the sample, a skip pattern is run through a list (frame) of the ***N*** units to select the sample
* The *skip* or *sampling interval*, ***k*** = ***N***/***n***

Linear systematic sampling: Selection process

1) Form a **sequential list** of population units
2) Decide on a sample size *n* and compute the skip (***sampling interval***), ***k*** = ***N***/***n***
3) Choose a random number, ***r*** (***random start***) between 1 and ***k*** (inclusive)
4) Add "***k***" to selected random number to select the second unit and continue to add "***k***" repeatedly to previously selected unit number to select the remainder of the sample

Linear systematic sampling: Selection process



Linear systematic sampling

* Previous example assumed that k = N/n is an integer
* Question: What if k = N/n is NOT an integer?
  * Solution 1: Work with decimal places and round
  * Solution 2: Circular sampling

Example – working with decimals and rounding

N	32	
n	7	
skip	4.571429	
R.Start	2.636695	
Sel1	2.636695	3
Sel2	7.208123	8
Sel3	11.77955	12
Sel4	16.35098	17
Sel5	20.92241	21
Sel6	25.49384	26
Sel7	30.06527	31

=32/7

=RAND()*4.571429

Same as the random start

=roundup(2.636695,0)

Circular sampling

1) Determine the interval ***k*** – rounding down to the integer nearest to ***N***/***n***
   (If ***N*** = 15 and ***n*** = 4, then ***k*** is taken as 3 and not 4)

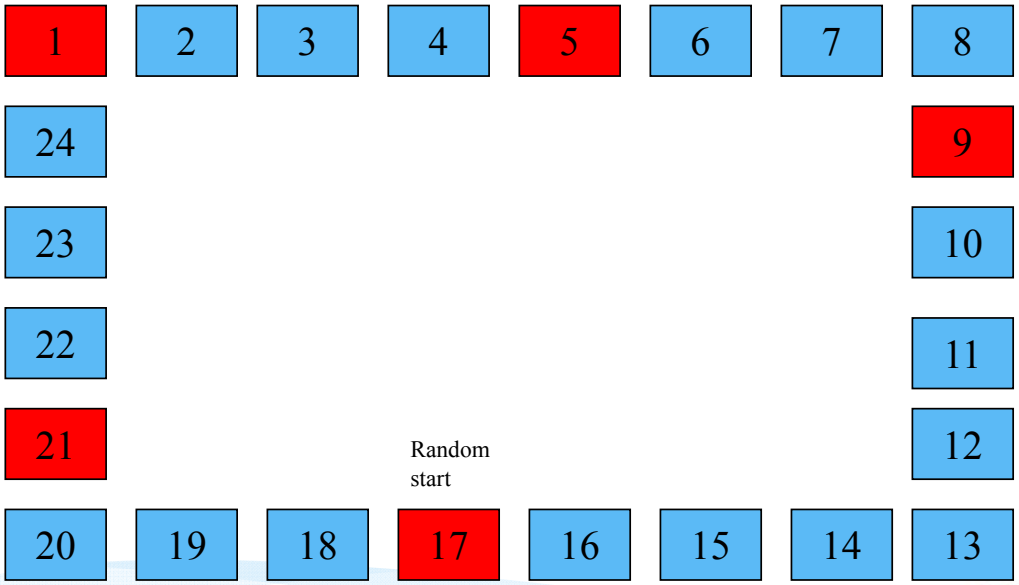
2) Take a random start between 1 and ***N***

3) Skip through the circle by ***k*** units each time to select the next unit until ***n*** units are selected

4) Thus there could be ***N*** possible distinct samples instead of ***k***

Circular sampling illustration

Population = 24, Sample = 5, Skip = Int(24/5=4.6) = 4



Estimation with Systematic Sampling

The weight for a systematic sample is the same as Simple Random Sampling

Estimate of a total

$$\hat{Y} = N \times \sum_{i=1}^{n} \frac{y_i}{n} \quad = \sum_{i=1}^{n} \frac{N}{n} y_i \quad = \sum_{i=1}^{n} w y_i$$

Estimate of a mean

$$\hat{\bar{Y}} = \sum_{i=1}^{n} \frac{y_i}{n}$$

Estimation with Systematic Sampling (cont)

Estimate of a proportion

$$y_i = \begin{cases} 1 & i^{th} \text{ sample unit has characteristic} \\ 0 & \text{otherwise} \end{cases}$$

$$\hat{Y} = \frac{N}{n}\sum_{i=1}^{n} y_i = \hat{N}_c$$

$$\hat{\bar{Y}} = \frac{1}{n}\sum_{i=1}^{n} y_i = \hat{P}_c$$