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: Simple Random 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

* What is Simple Random Sampling (SRS)
* Advantage and Disadvantages of SRS
* What a good SRS looks like
* Producing Estimates for SRS

Simple Random Sampling (SRS)

* This is the most basic type of sampling
* In SRS, every unit in the sample population has the same chance of selection
  (NB: and only one chance of selection)
* Random numbers are usually generated and assigned to each sample population unit in order to select the sample
* To produce a random number between 0 and 1 in excel type **=rand()**
* To produce a random number between 0 and 100 in excel type **=rand()*100**

Simple Random Sampling (SRS)

How many different types of sample can we have?

* In total there are N!/[(N-n)!n!] different samples which can be selected from the population
  N – population size n - sample size

If N = 10, n = 4

* Number of different samples
  = (10 x 9 x 8 x 7 x 6 x 5 x 4 x 3 x 2 x 1)/[(6 x 5 x 4 x 3 x 2 x 1) x (4 x 3 x 2 x 1)]
  = 210

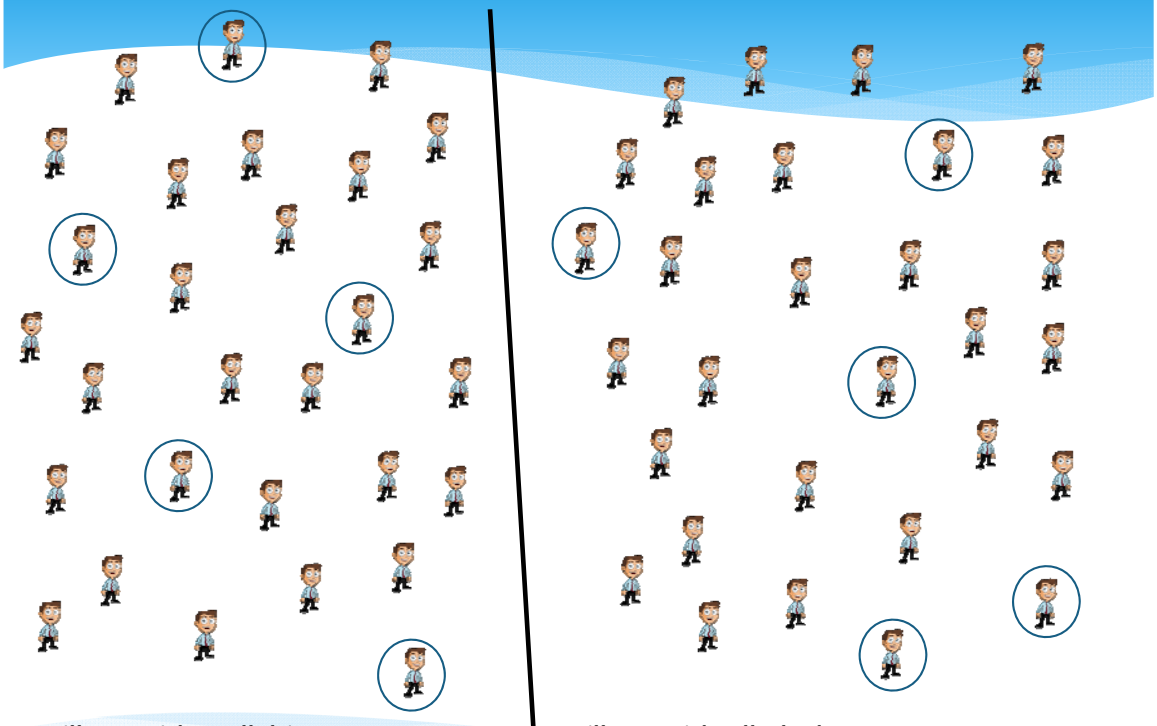
If N = 100, n = 30

* Number of different samples = 2.93723×10^{25}

Advantages & Disadvantages of SRS

* Advantages
  * It's simple – just need to generate a random number and use this for selection (provided you have a complete list of units)
  * Generally produces low SE's
* Disadvantages
  * Can be costly if the sample is well spread out geographically
  * Can't control the representativeness of the sample
    * Can't control the sample for sub-populations
    * Sample may be highly skewed to one area

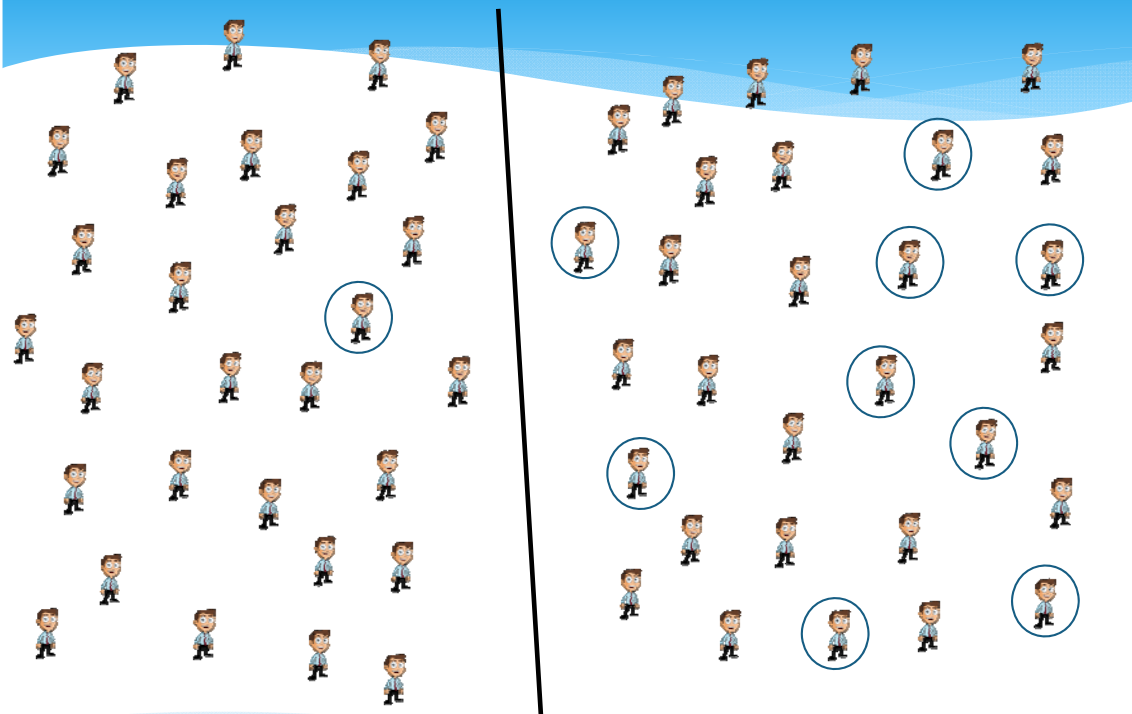
Sample nicely spread throughout the entire population



Village with Wallabies supporters

Village with All Blacks supporters

Sample skewed towards the population on the right



Village with Wallabies supporters

Village with All Blacks supporters

Producing estimates for a SRS

* The weight for a simple random sample can be calculated as the population (N) divided by the sample (n)

* That is, if the population is 200, and a simple random sample of 20 is selected, then the weight is:-
  * W = N/n = 200/20 = 10

Estimate of a total

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

Producing estimates for a SRS (cont)

Estimate of a mean

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

NB : This is the same as the formula for a total, without the N

Estimate of a proportion

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

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