P. R. Krishnaiah and C. R. Rao, eds., *Handbook of Statistics, Vol. 6*
© Elsevier Science Publishers B.V. (1988) 147–185

Systematic Sampling with Illustrative Examples

M. N. Murthy and T. J. Rao

1. The basic procedure

Systematic sampling, in its simplest and commonly used form, is selection of every kth unit from a finite population of N units, k being an integer nearest to the inverse of the sampling fraction aimed at. Here k is termed the *sampling interval*, and if the first unit from 1 to k is selected at random, the resulting sample is said to be a systematic sample with a *random start*. The sample size is given by $n = (N/k)$, if N is an integral multiple of k with $N = nk$, and the sample size is a random variable taking the values $[N/k]$ and $[N/k] + 1$ with probabilities $[N/k] + 1 - N/k$ and $N/k - [N/k]$ respectively leading to the expected sample size (N/k), if N is not an integral multiple of k.

More specifically, the procedure consists of serially numbering the units in the population, $\{U_i\}$, $i = 1, 2, \ldots, N$, selecting a random start (r) from 1 to k, which is an integer nearest to the inverse of the sampling fraction, and then selecting the units

$$\{U_{r+jk}\}, \quad j = 0, 1, 2, \ldots, l_r - 1, \tag{1}$$

where l_r is an integer such that

$$(l_r - 1)k \leqslant N - r < l_r k .$$

In other words, starting from the rth unit, r being a random number from 1 to k, the sampling interval, every kth unit is selected till the entire population is covered in the process. This procedure is termed *linear systematic sampling*, as the selection is done in a linear direction, and abbreviated to lss. If N is a multiple of k with $N = nk$, then the systematic sample with random start consists of units $\{U_{r+jk}\}$, $j = 0, 1, 2, \ldots, (n-1)$.

It can be easily seen that in this sampling scheme, the arrangement of units has an important influence on its efficiency, as any change in the arrangement will change the composition of the sample possible through this scheme. For instance if $k \geqslant 2$, which is generally the case, no two consecutive units can occur in the