

## More on Stratified Random Sampling (StRS)

### Selecting the total sample size for estimation of population means or totals

Using the same strategy from the last chapter, we can set  $2\sqrt{V(\bar{y}_{st})} = B$  and solve for  $n$  to get the total sample size required, but now there are two additional considerations. First, we substitute the sample fpc  $(N - n)/N$  for the population fpc  $(N - n)/(N - 1)$  to simplify the calculation, and second, to account for the unknown  $n_i$  terms, we substitute  $n_i = na_i$  for sample allocation fractions  $a_i$  that we will determine later. With these two simplifications, we can solve to obtain:

$$n = \frac{\sum_{i=1}^L N_i^2 \sigma_i^2 / a_i}{N^2 D + \sum_{i=1}^L N_i \sigma_i^2},$$

where  $D = B^2/4$  when estimating  $\mu$ , and  $D = B^2/4N^2$  when estimating  $\tau$ .

### Allocation among strata:

It can be shown (Cochran, 1977, Chapter 5, pg. 96) that to minimize the cost of a sample for fixed variance, the optimal allocation of samples among strata is to make the stratum size proportional to:  $N_i \sigma_i / \sqrt{c_i}$ , or in other words,

$$n_i = n \left( \frac{N_i \sigma_i / \sqrt{c_i}}{\sum_{k=1}^L N_k \sigma_k / \sqrt{c_k}} \right).$$

Also, substituting back into the total sample size formula above, we get:

$$n = \frac{\left( \sum_{k=1}^L N_k \sigma_k / \sqrt{c_k} \right) \left( \sum_{i=1}^L N_i \sigma_i \sqrt{c_i} \right)}{N^2 D + \sum_{i=1}^L N_i \sigma_i^2}$$

These two equations are then the ones to use to calculate total sample size and the optimal allocation of stratum sample sizes in stratified random sampling. If all costs are assumed equal, the simplified expressions are referred to as *Neyman allocation*. If both the costs and the variances are assumed equal, the simplified equations are then called *proportional allocation*.

**Reference:**

Cochran, W.G. 1977. Sampling Techniques. New York: John Wiley & Sons