Some reasons to consider Stratified Sampling

(Adapted from Cochran 1977)

1. If information is desired for certain subdivisions of the population, it is advisable to treat each subdivision as a "population" in its own right.

2. Administrative convenience may dictate the use of stratification. For example, the agency conducting the experiment may have field offices, each of which can supervise the experiment for a part of the whole population.

3. Sampling problems may differ in different parts of the population. With human populations, people living in institutions (e.g., hotels, hospitals, prisons) are often placed in a different stratum from people living in ordinary homes because a different approach to the sampling is appropriate for the two situations.

4. Stratification may produce a gain in precision in the estimates of characteristics (parameters) of the whole population. It may be possible to divide a heterogeneous population into subpopulations, each of which is internally homogeneous. This is suggested by the name strata, with its implication of a division into layers. If each stratum is homogeneous, in that the measurements vary little from one unit to another, a precise estimate of any stratum mean can be obtained from a small sample in that stratum. These estimates can then be combined into precise estimates of characteristics of the whole population.

Approximate dfs using Satterthwaite's (1946) Method

If sample sizes are small and assuming normally distributed populations, a $100(1-\alpha)\%$ CI for μ is

$$\bar{y}_{st} \pm t \sqrt{\widehat{Var}(\bar{y}_{st})}$$

Q: df = ?

By definition, $t = \dfrac{Z}{\sqrt{\chi^2/df}}$

In our case, $t = \dfrac{\bar{y}_{st}-\mu}{\sqrt{\widehat{Var}(\bar{y}_{st})}} = \dfrac{\dfrac{\bar{y}_{st}-\mu}{\sqrt{Var(\bar{y}_{st})}}}{\sqrt{\dfrac{\widehat{Var}(\bar{y}_{st})}{Var(\bar{y}_{st})}}}$

Objective is to find df such that $\dfrac{\widehat{Var}(\bar{y}_{st})}{Var(\bar{y}_{st})}$ is χ^2/df.

Note that

$$E\left(\frac{\chi^2}{df}\right) = \frac{1}{df}E(\chi^2) = \frac{1}{df}\cdot df = 1$$

$$Var\left(\frac{\chi^2}{df}\right) = \frac{1}{(df)^2}Var(\chi^2) = \frac{1}{(df)^2}\,2\,df = \frac{2}{df}$$

and

$$\frac{(n_h-1)S_h^2}{\sigma_h^2} \sim \chi^2_{n_h-1}.$$

Satterthwaite's (1946) approximation for df is obtained by solving

$$Var\left(\frac{\widehat{Var}(\bar{y}_{st})}{Var(\bar{y}_{st})}\right) = \frac{2}{df}$$

It follows that, when replacing σ_h^2 with S_h^2, we get

$$df = \frac{\left(\sum_{h=1}^{L} a_h S_h^2\right)^2}{\sum_{h=1}^{L} \frac{(a_h S_h^2)^2}{n_h-1}}, \quad \text{where} \quad a_h = \frac{N_h(N_h-n_h)}{n_h}.$$