LECTURE 14: **10.2 Simple Two-Stage Cluster Sampling When Clusters Have Different Sizes**

As before, m clusters are chosen from among M clusters by drawing an SRS. Then within each sampled cluster i of size N_i, we take an SRS of n_i elements, where n_i is such that n_i/N_i is as close as possible to the predetermined second-stage sampling fraction f_2.

Example. Suppose a population has 10 clusters of sizes 10, 10, 20, 10, 10, 20, 20, 20, 30, 20, respectively. We want to draw an SRS of 3 clusters, and within each sampled cluster take an SRS of 20% of the elements.

99 42 03 29 12 68 48 53 93 44 25 45 97 42 76 22 09 93 66 51 98 47 12 77 35
05 05 39 74 43 77 34 60 56 59 19 27 48 17 61 48 57 38 49 15 44 30 22 15 41
33 22 54 56 22 21 26 53 38 88 61 87 91 07 71 89 93 55 83 02 75 54 99 13 44
87 50 47 05 94 40 67 58 16 01 62 62 02 21 16 94 30 56 16 42 09 47 25 29 59
37 32 53 94 59 30 40 24 52 51 95 89 56 08 47 08 06 27 93 81 04 83 91 51 24
82 97 54 21 50 32 81 34 28 03 72 92 20 52 60 71 28 81 13 51 69 35 14 66 52
20 94 11 12 90 22 55 55 41 34 01 01 62 68 94 24 21 54

Estimation of Population Characteristics

The **population total** X is estimated by

$$x'_{clu} = \left(\frac{M}{m}\right) \sum_{i=1}^{m} \left(\frac{N_i}{n_i}\right) \sum_{j=1}^{n_i} x_{ij} = \left(\frac{M}{m}\right) \sum_{i=1}^{m} \left(\frac{N_i}{n_i}\right) x_i \,.$$

Note that before $N_i = \bar{N}, n_i = \bar{n}, \bar{n}/\bar{N} = f_2, m/M = f_1, f = f_1 f_2$, and $x'_{clu} = x/f = \sum_{i=1}^{m} x_i/f$.

A $100(1-\alpha)\%$ CI for X is

$$x'_{clu} \pm z_{1-\alpha/2} \frac{M}{\sqrt{m}\, f_2} \sqrt{\frac{\sum_{i=1}^{m}(x_i - \bar{x})^2}{m-1}} \sqrt{\frac{N-n}{N}} \,.$$

The **population mean per cluster** $\bar{X}$ is estimated by $\bar{x}_{clu} = x'_{clu}/M$. A $100(1-\alpha)\%$ CI for $\bar{X}$ is

$$\bar{x}_{clu} \pm z_{1-\alpha/2} \frac{1}{\sqrt{m}\, f_2} \sqrt{\frac{\sum_{i=1}^{m}(x_i - \bar{x})^2}{m-1}} \sqrt{\frac{N-n}{N}} \,.$$

The **population mean per element** $\bar{\bar{X}}$ is estimated by $\bar{\bar{x}}_{clu} = x'_{clu}/N$. A $100(1-\alpha)\%$ CI for $\bar{\bar{X}}$ is

$$\bar{\bar{x}}_{clu} \pm z_{1-\alpha/2} \frac{1}{\bar{N}\sqrt{m}\, f_2} \sqrt{\frac{\sum_{i=1}^{m}(x_i - \bar{x})^2}{m-1}} \sqrt{\frac{N-n}{N}} \,.$$

Example (pg 302). $N_1 = 4,288$, $n_1 = 429$, $x_1 = 5$, $N_2 = 638$, $n_2 = 64$, $x_2 = 7$, $N_3 = 2,154$, $n_3 = 215$, $x_3 = 3, M = 10, N = 50,056$.

SOLUTION. $f_2 = 0.1, \bar{x} = 5, n = 708,$

$$\frac{\sum_{i=1}^{m}(x_i - \bar{x})^2}{m-1} = 4,$$

$$x'_{clu} = 499.38,\ \widehat{SE}(x'_{clu}) = 114.65,$$

A 95% CI for X is $[274.67, 724.09]$.