I. Introduction to SRS

A. **<u>Simple Random Sampling</u>** - a sample of size n is drawn from a population of size N in such a way that each of the T = $\binom{N}{n}$ possible samples of n elements has the same $\frac{1}{\binom{N}{n}}$ probability of selection (sampling without replacement). SRS is element sampling, since elementary units serve as sampling units. Each element has an $\frac{n}{N}$ chance of being selected.

B. <u>How to draw a SRS</u>.
1. Assign each element in the population a number from 1 to N.
2. Pick n of the N numbers by some random process, using a random number table (Table A1, page 538) or a computer random number generator.
3. Take the population elements corresponding to the numbers randomly chosen.

C. <u>Finite Population Correction Factors</u>

When sampling from a finite population, $\left(\frac{N-n}{N-1}\right)$ is the **finite population correction factor (fpc), used to adjust standard errors. The fpc reduces the <u>standard error of an estimate</u> as n approaches N.**

II. <u>Estimation of Population Mean, Total, and Proportion</u>
Let $x_1, x_2, \ldots, x_n$ be a SRS from a population of size N.
A. <u>Estimation of the Population Mean $\overline{X}$</u> (or μ_X)

1. $\hat{\mu}_X = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x}{n}$ (3.3)

2. $E(\bar{x}) = \overline{X}$ (3.4)

3. $VAR(\bar{x}) = \left(\frac{\sigma_X^2}{n}\right)\left(\frac{N-n}{N-1}\right)$ (3.5)

4. $SE(\bar{x}) = \frac{\sigma_X}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ (3.6)

5. $V(\bar{x}) = \frac{\sigma_X}{\overline{X}\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{V_X}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ (3.7)

6. $\hat{VAR}(\bar{x}) = \frac{\hat{\sigma}_X^2}{n}\left(\frac{N-n}{N-1}\right) = \left(\frac{(N-1)}{N}s_x^2\right)\left(\frac{1}{n}\right)\left(\frac{N-n}{N-1}\right)$

$= \frac{s_x^2}{n}\left(\frac{N-n}{N}\right)$ (3.10)

7. <u>100(1-α) % Confidence Interval for $\overline{X}$</u>

$$\bar{x} \pm Z_{1-\frac{\alpha}{2}}\frac{s_x}{\sqrt{n}}\sqrt{\frac{N-n}{N}} \quad (3.10)$$

<u>Example: Confidence Interval for $\overline{X}$ (Table 3.4, Page 62).</u>

$N = 25,\ n = 9,\ \sum_{i=1}^{9} x_i = 44,\ \sum_{i=1}^{9} x_i^2 = 312,\ \bar{x} = \frac{44}{9} = 4.89$

$$s_x^2 = \frac{\sum_{i=1}^{9} x_i^2 - \frac{\left(\sum_{i=1}^{9} x_i\right)^2}{9}}{8} = \frac{312 - \frac{(44)^2}{9}}{8} = \frac{312 - 215.11}{8} = 12.111$$

95% Confidence Interval for $\overline{X} \Rightarrow 1-\alpha = .05,\ \frac{\alpha}{2} = .025,\ Z_{.975} = 1.96$

$\bar{x} \pm Z_{1-\frac{\alpha}{2}}\frac{s_x}{\sqrt{n}}\sqrt{\frac{N-n}{N}} = 4.89 \pm 1.96\left(\frac{3.48}{3}\right)\sqrt{\frac{16}{25}} = 4.89 \pm 1.96(0.928) = 4.89 \pm 1.82$
$= (3.07, 6.71)$
We are 95% sure that the average number of household visits per physician is enclosed by the interval (3.07, 6.71).

B. Estimation of the Population Total $X = N\overline{X}$

1. $\hat{X} = N\hat{\mu}_X = N\bar{x} = \frac{N\sum_{i=1}^{n} x_i}{n} = \left(\frac{N}{n}\right)x = x'$ (3.2)

2. E(x') = X (3.2)

3. VAR(x') = VAR($N\bar{x}$) = N^2 VAR($\bar{x}$) = $\left(\frac{N^2\sigma_X^2}{n}\right)\left(\frac{N-n}{N-1}\right)$ (3.2)

4. SE(x') = $\frac{N\sigma_X}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ (3.2)

5. V(x') = $\frac{N}{\sqrt{n}}\frac{\sigma_X}{X}\sqrt{\frac{N-n}{N-1}} = \frac{\sigma_X}{\overline{X}\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{V_X}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ (3.6)

6. $V\hat{A}R(x') = \frac{N^2}{n}\hat{\sigma}_X^2\left(\frac{N-n}{N-1}\right) = \frac{N^2}{n}\left(\frac{(N-1)}{N}s_x^2\right)\left(\frac{N-n}{N-1}\right)$

$$= N^2\frac{s_x^2}{n}\left(\frac{N-n}{N}\right) \quad (3.9)$$

7. <u>100(1- α) % Confidence Interval for X</u>

$$x' \pm Z_{1-\frac{\alpha}{2}} N \frac{s_x}{\sqrt{n}}\sqrt{\frac{N-n}{N}} \quad (3.9)$$

<u>Example: Confidence Interval for X (Table 3.4, Page 62).</u>

N = 25, n = 9, $x = 44, s_x = \sqrt{12.111} = 3.48, x' = \left(\frac{25}{9}\right)(44) = 122.22$

95% Confidence Interval for X:

$x' \pm Z_{1-\frac{\alpha}{2}} N \frac{s_x}{\sqrt{n}}\sqrt{\frac{N-n}{N}}$ = $122.22 \pm 1.96(25)\left(\frac{3.48}{3}\right)\sqrt{\frac{16}{25}}$ = 122.22 ± 45.47 = (76.75, 167.69)

We are 95% sure that the total number of household visits made by physicians in the population is enclosed by the interval (76.75, 167.69).