STRATUM Population Estimators and Standard Error Equations for Simple Random Street Segment Sampling:

Sample size (number of street segments) = n

Sample size of k-th subpopulation = n_k

Value measured at street segment i is y_i (with $i = 1,2,\ldots,n$)

Value measured at street segment i within subpopulation k: y_{k_i} (with $i = 1,2,\ldots,n_k$)

Population size (number of street segments) = N

Citywide sample mean = $\bar{y} = \frac{1}{n}\left(y_1 + y_2 + \cdots + y_n\right)$ Equation 1

Estimate of the citywide population total = $\hat{\tau} = N\bar{y}$ Equation 2

Citywide sample variance = $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2$ Equation 3

Estimated variance of the citywide sample mean = $\hat{v}(\bar{y}) = \left(\frac{N-n}{N}\right)\frac{s^2}{n}$ Equation 4

Standard error (se) of the citywide sample mean = $\sqrt{\hat{v}(\bar{y})}$ Equation 5

Estimated variance of the estimate of the citywide total = $\hat{v}(\hat{\tau}) = N^2\hat{v}(\bar{y})$ Equation 6

Standard error (se) of the estimated citywide total = $\sqrt{\hat{v}(\hat{\tau})}$ Equation 7

Estimate of the subpopulation (zone) total = $\hat{\tau}_k = \frac{N_k}{n_k}\sum_{i=1}^{n_k} y_{ki}$ Equation 8

Subpopulation (zone) sample variance = $s_k^2 = \frac{1}{n_k - 1}\sum_{i=1}^{n_k}\left(y_{ki} - \bar{y}_k\right)^2$ Equation 9

Estimated variance of the estimate of the subpopulation (zone) total = $\hat{v}(\hat{\tau}_k) = N_k^2\left(\frac{N_k - n_k}{N_k n_k}\right)s_k^2$ Equation 10

Standard error (se) of the estimated subpopulation (zone) total = $\sqrt{\hat{v}(\hat{\tau}_k)}$ Equation 11

Example (1) citywide calculations:
The city of Evergreen possesses 100 total street segments (N). For purposes of a STRATUM analysis, a 4% random sample of street segments were inventoried for municipal street trees. The number of trees in the 4 (n) sample units (y_1, y_2, y_3, y_4) were 11,9,12, and 7, respectively.

The mean number of trees per segment (using Equation 1) is

$$\bar{y} = \left(\frac{11+9+12+7}{4} \right) = 9.75$$

The sample variance (using Equation 3) is

$$s^2 = \frac{(11-9.75)^2 + (9-9.75)^2 + (12-9.75)^2 + (7-9.75)^2}{4-1} = 4.9166$$

The estimated variance of the sample mean (using Equation 4) is

$$\hat{v}(\bar{y}) = \left(\frac{100-4}{100} \right) \frac{4.9166}{4} = 1.1799$$

so that the estimated standard error (using Equation 5) is $\sqrt{1.1799} = 1.086$.

An estimate of the total number of trees in the city (using Equation 2) is

$$\hat{\tau} = 100(9.75) = 975$$

The estimated variance associated with the estimate of the total (using Equation 6) is

$$\hat{v}(\hat{\tau}) = 100^2(1.1799) = 11799$$

giving an estimated citywide standard error (using Equation 7) of $\sqrt{11799} = 108.62$. The city of Evergreen tree population, therefore, is 975 ±109 (estimate ±standard error).

Example (2) zone calculations:

Following Example 1, the city of Evergreen possesses 100 total street segments (N). For purposes of a STRATUM analysis, a 4% random sample of street segments were inventoried for municipal street trees. The number of trees in the 4 (n) sample units (y_1, y_2, y_3, y_4) were 11,9,12, and 7, respectively. The city is divided into two management zones (k = 1,2), where $k = 1$ includes sample units y_1 and y_2 and $k = 2$ includes sample units y_3 and y_4. The total number of segments per zone ($k_{1,2}$) is 40 and 60, respectively.

An estimate of the total number of trees in the management zone 1 (using Equation 8) is

$$\hat{\tau}_k = \frac{40}{2}(11+9) = 400$$

The sample variance of the sampling units associated with zone 1 (using Equation 9) is

$$s_k^2 = \frac{1}{2-1}\left((11-10)^2 + (9-10)^2\right) = 1$$

The estimated variance associated with the estimate of the zone 1 total (using Equation 10) is

$$\hat{v}(\hat{\tau}_k) = 40^2\left(\frac{40-2}{40(2)} \right)(1) = 760$$

giving an estimated zone 1 standard error (using Equation 11) of $\sqrt{760} = 27.57$. The zone 1 tree population, therefore, is 400 ±28 (estimate $\pm$ standard error).

Notice that the city of Evergreen zone 2 tree population would equal 570 ±147, where Example 2 equations were used. The citywide population estimation is, therefore, additive (zone1 [400] and zone 2 [570] sum to a total of 970). However, the standard error citywide does not equal the sum of the standard error by zone (28+147≠109).