Formulas for estimators and their variances in NFI
28.2.2014 K.T. Korhonen & Olli Salmensuu, point estimators

1. Brief description of the sampling

The estimators presented in this paper are valid for one phase systematic sampling design. It is assumed that the sampling unit is a cluster of 4 plots. The plots are nested rectangular or nested circular plots (to be confirmed later). A plot may fall in the border of two land use classes or two forest types, in this case plot has two parts. However, the plot center in the case of circular plots or corresponding reference point in the case of rectangular plots unambiguously falls in only one land use class or forest type.

Chapter 2 shows the estimators for one stratum. Chapter 3 presents the estimators over several stratums.

2. Estimators

<u>Area estimates</u>
Forest proportion (or proportion of category of interest) from the total land area is estimated using Formula (1):

$$\bar{p}^{(F)} = \frac{\sum^{n} p_i^{(F)}}{\sum^{n} m_i} \qquad (1)$$

where
$p_i^{(F)} = \sum_j^4 t_{ij} m_{ij}$, where t_{ij} = 1 if the plot center of plot j in cluster i is in forest (or in the category of interest; 0, otherwise
m_i= number of plot centers (in land if A is land area) in cluster i
n = number of clusters.

Forest area (or the area of the category of interest) is estimated as:

$$A^{(F)} = A \cdot \bar{p}^{(F)} \qquad (2)$$

Variance of the forest proportion estimate is obtained with Formula (3):

$$v(\bar{p}^{(F)}) = \frac{1}{(\sum^{n} m_i)^2} \cdot \frac{n}{n-1} \cdot \sum_{i=1}^{n} (p_i^{(F)} - \bar{p}^{(F)} m_i)^2 \qquad (3)$$

And for forest area (or any category of interest):

$$v\left(A^{(F)}\right) = A^2 \cdot v(\bar{p}^{(F)}) \qquad (4)$$

Mean biomass estimates

The mean biomass of growing stock in total area (land) is estimated in a manner that is similar to forest proportion estimation, using biomass (tons/ha) at the plot (part) in the variable *p*. **However, mean biomass in forests (or in a sub-category of forests) is often more interesting than mean biomass over the whole population. This can be estimated with the same Formula (1), but in this case only the plot parts in forest (or in the category of interest) are considered when calculating the sum in the nominator (sum of x_i), and plot centers that are in forest (or in the category of interest) are considered when calculating the denominator (sum of m_i)**:

$$\bar{x}^{(F)} = \frac{\bar{x}}{\bar{t}} \tag{5}$$

$$\bar{t} = \frac{\sum^{n} t_i}{\sum^{n} m_i} \tag{6}$$

$$\bar{x} = \frac{\sum^{n} x_i}{\sum^{n} m_i} \tag{7}$$

where

$t_i = \sum_j^4 t_{ij}$, number of plot centers in forest (or category of interest) in cluster i.

$x_i = \sum_j^4 x_{ij}$, sum of biomass (tons/ha) in plots (or parts of plot) that are forest (or category of interest) in cluster i (x_{ij} is set to 0 if the plot (part) is not in forest (or in category of interest)).

Note: when estimating mean biomass of forests, the biomass of single trees are multiplied by their diameter-specific expansion factors. **Even if a plot is only partly in forest, no correction for the plot size is done when calculating the x_i values.** The divided plots will be handled correctly by the fact that trees are measured also in the case that the plot center is outside forest and only part of plot is in forest. In this case the biomass of trees are summed for the nominator but nothing (or zero) is summed for the denominator.

The variance of mean biomass **of forests** is estimated with Formula 8:

$$v(\bar{x}^{(F)}) = \frac{1}{(\sum^n t_i)^2} \frac{n}{n-1} \sum^n \left(x_i - \bar{x}^{(F)} \cdot t_i\right)^2 \tag{8}$$

In other words, when calculating the variance of mean biomass in forests, consider only those clusters that have at least one sample point in forest for calculating the denominator part (sum of t_i). Correspondingly, only those clusters that have at least one part of plot in forest have effect on the sum of $(x_i - \bar{x}^{(F)} \cdot t_i)$ because both x_i and t_i are 0 if there is no plot part in forest in cluster i.

Total biomass or total volume estimates

The total biomass or total volume in forests (or in a category of interest) is the estimated as a product of the mean biomass and area estimate.

$$X^{(F)} = A^{(F)} \cdot \bar{x}^{(F)} \qquad (9)$$

The variance of total biomass in forest is estimated as:

$$v\left(X^{(F)}\right) = A^{(F)2} v\left(\bar{x}^{(F)}\right) + \bar{x}^{(F)2} v\left(A^{(F)}\right) \qquad (10)$$

Formula 10 in its' presented form estimates the variance of total biomass in absolute units (metric tons^2). Formula 10 can be presented also in the form that the relative standard error of total volume estimate is the square root of sum of squared relative standard error of mean biomass estimate and squared relative standard error of forest area estimate:

$$se\%(X^{(F)}) = \sqrt{se\%^2(\bar{x}^{(F)}) + se\%^2(A^{(F)})} \qquad (11)$$

where $se\%(\bar{x}^{(F)}) = \frac{100\sqrt{v(\bar{x}^{(F)})}}{\bar{x}^{(F)}}$ (relative standard error of forest area estimate)

and $se\%(A^{(F)}) = \frac{100\sqrt{v(A^{(F)})}}{A^{(F)}}$.

3. Summing over stratums

This Chapter presents the estimators in the case that country has been divided in two or more sub-regions, and sampling is done independently for each sub-region. Sampling design and/or intensity may vary by these sub-regions. In the following, these sub-regions are called stratums.

Results and variances are first calculated for each stratum sepraratELY as presented in Chapter 2. The sum of forest area (or area of any forest category) for the whole country (over sampling stratums) is estimated by summing the stratum wise estimates. The variance of total forest area estimates (or total area of any other category) is the sum of the stratum wise variances of that category. Correspondingly, the total biomass of forests or any category over the sampling stratums is estimated by summing the biomass estimates of each stratum. The variance of total biomass is the sum of the stratum wise total biomass estimate variances.

The mean biomass of forests in whole country over the stratums is estimated by dividing the total biomass estimate by the total forest area estimate. The variance of mean biomass is solved from Equation 10, resulting:

$$v(\bar{x}_{st}{}^{(F)}) = \frac{v(X_{st}{}^{(F)}) - \bar{x}_{st}{}^{(F)2} v(A_{st}{}^{(F)})}{A_{st}{}^{(F)2}} \qquad (12)$$

where the subscript st refers to the estimate summed over sampling stratums.