Suora (direct) ja epäsuora (indirect) estimaattori

(Lehtonen and Veijanen 2009)

A *direct* estimator uses values of the variable of interest only from the time period of interest and only from units in the domain of interest (Federal Committee on Statistical Methodology, 1993).

A Horvitz-Thompson type estimator

$$\hat{t}_d = \sum_{k \in s_d} y_k / \pi_k$$

provides a simple example of direct estimator.

In model-assisted estimation, direct estimators are constructed by using models fitted separately in each domain; an example is a model

$$Y_k = \mathbf{x}_k' \boldsymbol{\beta}_d + \varepsilon_k, \; k \in U_d,$$

with domain specific auxiliary *x*-data and a vector of regression coefficients $\boldsymbol{\beta}_d$, $d = 1,...,D$.

A direct domain estimator can still incorporate auxiliary data outside the domain of interest. This is relevant if accurate population data about the auxiliary *x*-variables are only available at a higher aggregate level.

An *indirect* domain estimator uses values of the variable of interest from a domain and/or time period other than the domain and time period of interest (Federal Committee on Statistical Methodology, 1993).

For example, if a linear model

$$Y_k = \mathbf{x}_k' \boldsymbol{\beta} + \varepsilon_k, \; k \in U,$$

with a common vector $\boldsymbol{\beta}$ is used as an assisting model, the resulting domain estimator will be indirect.

In general, indirect estimators are attempting to “borrow strength” from other domains and/or in a temporal dimension. The concept of “borrowing strength” is often used in model-based small area estimation (e.g. Rao, 2003).

Indirect model-assisted estimators for domains are discussed in the literature (e.g. Estevao and Särndal, 1999, Lehtonen, Särndal and Veijanen, 2003, 2005, and Hidiroglou and Patak, 2004).

Estevao and Särndal (2004) have argued in favour of direct estimators in the context of design-based estimation for domains.