Domains:

Large domains: (Classes, Areas)

Sample is large enough to make reliable estimates

Small domains:

Sample is too small (or, non-existent) to make reliable estimates

Population : (K sub-populations)

$$Y_{11} \quad Y_{12} \cdots \quad Y_{1N_1}$$

$$Y_{21} \quad Y_{22} \cdots \quad Y_{2N_2}$$

$$\vdots$$

$$Y_{K1} \quad Y_{K2} \cdots \quad Y_{KN_K}$$

Samples:

$$Y_{11} \quad Y_{12} \cdots \quad Y_{1n_1}$$

$$Y_{21} \quad Y_{22} \cdots \quad Y_{2n_2}$$

$$\vdots$$

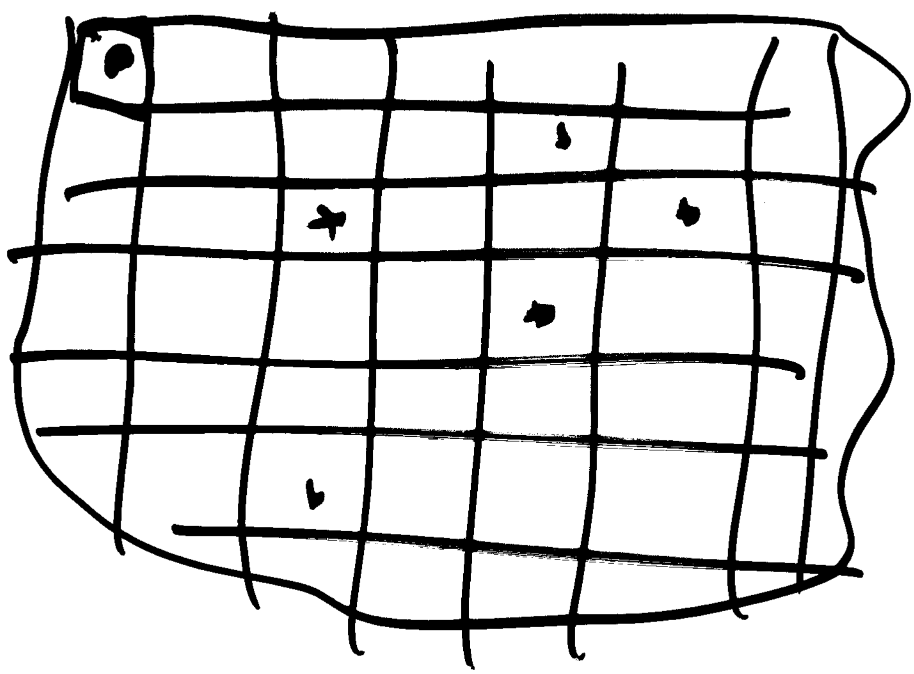
$$Y_{K1} \quad Y_{K2} \cdots \quad Y_{Kn_K}$$

Assumption : Simple random samples.

Population characteristic : $\bar{Y}_1$
(Average income) $\bar{Y}_2$
$\vdots$
$\bar{Y}_K$

Sample estimates : $\bar{y}_1$
(Direct estimates) $\vdots$
$\bar{y}_K$

Reliability : (S.E.) σ_1^2/n_1
σ_2^2/n_2
$\vdots$
σ_K^2/n_K



Y_i : Truth

X_i : Classification according to the satellite image

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \quad (\text{Model})$$

Additional infn: (a) Model structure
(b) Estimated coeff.

Two possibilities

1. Large domain

   Use regression fit from the data in the domain only.

2. Small domain

   Assume common coefficients across the small areas

P_{it} : # of households
in segmento 'i'
at time 't'.

$p_{it} \in (0, 1)$

$$\theta_{it} = \log \frac{p_{it}}{1-p_{it}} \qquad (-\infty, \infty)$$

Allows 'normality' assumption.

N_{it} : # children attending school

<u>$\log N_{it}$</u> : To make it "normal".