

INNOVATION SURVEY - SAMPLING ERROR (using E. Raulin Formula)

Variable: X (Innovation expenditure; Number of innovative enterprises: where X=1 when the enterprise is innovative, and X=0 otherwise; etc.)

The **estimator** in domain m is:

$$\hat{X}_m = \sum_h \sum_j F_{hj} * X_{hj} * Z_{hjm}$$

where,

j = questionnaire

h = initial strata (NUTSlevel2 + NACE2digit + size)

F = weighting factor

Z = random variable with the following values: 1 when $j \in m$, and 0 otherwise

The **estimated variance** for \hat{X}_m is:

$$\hat{V}(\hat{X}_m) = \sum_h \sum_j F_{hj} (F_{hj} - 1) * (X_{hj} * Z_{hjm} - \hat{X}_{hm})^2$$

where

$$\hat{X}_{hm} = \frac{\sum_j F_{hj} * X_{hj} * Z_{hjm}}{\sum_j F_{hj}} ; \text{ (enterprises with weighting factor=1 are not included)}$$

The **coefficient of variation** (%) is:

$$CV(\hat{X}_m) = \frac{\sqrt{\hat{V}(\hat{X}_m)}}{\hat{X}_m} * 100 ; \text{ (ALL enterprises included in } \hat{X}_m \text{)}$$

When calculating the estimation of a ratio, it is used a **ratio estimator**:

$$\hat{R} = \frac{\hat{X}}{\hat{Y}}$$

where:

$$\hat{V}(\hat{R}) = \frac{1}{\hat{Y}^2} (\hat{V}(\hat{X}) + \hat{R}^2 \hat{V}(\hat{Y}) - 2\hat{R}Cov(\hat{X}, \hat{Y}))$$

and:

$$\text{Cov}(\hat{X}, \hat{Y}) = \sum_h \sum_j F_{hj} (F_{hj} - 1) * \left(X_{hj} * Z_{hjm} - \hat{X}_{hm} \right) \left(Y_{hj} * Z_{hjm} - \hat{Y}_{hm} \right)$$