International Labour Organization

ILO-IPEC Interactive Sampling Tools No. 2

Allocation of Sample Size among domains or strata

Version 1

August 2014

International Programme on the Elimination of Child Labour (IPEC)

Fundamental Principles and Rights at Work (FPRW) Branch
Governance and Tripartism Department

ILO-IPEC

ILO-IPEC Interactive Sampling Tools No. 2 – Allocation of Sample Size among domains or strata / International Labour Office, International Programme on the Elimination of Child Labour (IPEC) - Geneva: ILO, 2014

ACKNOWLEDGEMENTS

This publication was elaborated by Mr. Farhad Mehran, consultant, for ILO-IPEC and coordinated by Mr. Federico Blanco Allais from IPEC Geneva Office.

Funding for this ILO publication was provided by the United States Department of Labor (Projects GLO/13/21/USA & GLO/10/55/USA).

This publication does not necessarily reflect the views or policies of the United States Department of Labor, nor does mention of trade names, commercial products, or organizations imply endorsement by the United States Government.

Visit our website: www.ilo.org/ipec

Available in electronic PDF format only.

Photocomposed by ILO-IPEC Geneva.

1. Introduction

This document describes the use of the template "Allocation" of the SIMPOC Interactive Sampling Tools. The template assists the user to examine sample allocations among domains or strata, allowing for different allocation procedures, including equal allocation, proportional allocation, the compromise square-root allocation and general allocation where the allocation parameter and the minimum sample size in each stratum can be specified. More strata or domains can be inserted in the template as needed.

The template is divided into three parts: Input values, Output values and Intermediary calculations. The contents and use of each part is described in turn below.

2. Input values

There are two types of input values: input parameters and input data. The input parameters are few and specified manually in the top rows of the template, while the input data are generally numerous and transferred from another file in the first three columns of the template.

- **Input parameters**

Sample allocation is based on three input parameters:

n = Total sample size of the survey in terms of number of households.

H = Number of domains or strata among which the total sample size should be allocated.

α = Optional parameter specifying the method of allocation. The default value is 1 corresponding to proportional allocation; 0 corresponds to equal allocation; and 0.5 to square-root accolcation.

- **Input data**

The input data are in the form of three columns:

Col A h = List of the domains or strata names

Col B Zh = The allocation variable. Vector of domains or strata values to be used for allocation. For example, the number of households in the domain or stratum according to the sampling frame. There should be as many values as there are domains or strata as specified by the input parameter H.

Col C Zo = The minimum allocation for each domain or stratum to be used in the general allocation method with the input parameter α. The minimum values may be the same for all domains or strata or may differ from one domain or stratum to another.

3. Output values

The output values consist of four allocations presented in four columns:

Col E Equal allocation

$$n_h = \frac{n}{H}$$

Col F Proportional allocation

$$n_h = \frac{n \times Z_h}{\sum_h Z_h}$$

Col G Square-root allocation

$$n_h = \frac{n \times \sqrt{Z_h}}{\sum_h \sqrt{Z_h}}$$

Col H General allocation

$$n_h = \frac{\min(Z_{oh}, k \times Z_h^{\alpha})}{\sum_h Z_h^{\alpha}}$$

where Z_{oh} is the minimum allocation in domain or stratum specified in Col C, the value of α is the input parameter specifying the allocation method and k is a proportionality factor determined as part of the intermediary calculations.

4. Intermediary calculations

The intermediary calculations use the input values to derive the output values.

Col J Calculation of the square-root of the allocation variable Z_h for each of the domains or strata

$$\sqrt{Z_h}$$

Col K Calculation of the power function of the allocation variable specified by the input parameter α, for each of the domains or strata

$$Z_h^{\alpha}$$

Cols L to O. Calculation of the proportionality factor k for the general allocation method of Col H. The calculation is based on the first three iterations of the Newton approximation for solving the equation: Find k such that the following equation holds:

$$\sum_h \min(Z_{oh}, k \times Z_h^{\alpha}) = n$$

Newton's iterative approximation for solving the above equation starts from an initial value of k and gives successively a new value based on its previous value, using the formulae

$$k = k - \frac{\sum_h f(k) - n}{\sum_h f'(k)}$$

where the initial value is given by

$$k_o = \frac{n}{\sum_h Z_h^{\alpha}}$$

and the function f(k) and its derivative for a given value of k are

$$f(k) = \min(Z_{o_h}, k \times Z_h^{\alpha})$$

$$f'(k) = (Z_h^{\alpha}/2) \times [1 - sign(k \times Z_h^{\alpha} - Z_{o_h})]$$

where sign is the sign function, sign(x) = 1 if x is positive, sign(x) = -1 if x is negative, and sign(x) = 0 if x is zero.

The successive calculated values of k for three iterations are given in row 6 of the intermediary calculations block.