Stratified Sampling for Population Mean Estimation with Multiple Allocation Strategies

Pratap Mangalam

December 16, 2025

Abstract

This paper investigates stratified sampling for estimating the population mean using proportional, Neyman, and optimised allocation strategies. Sampling bias, cost, and time constraints are incorporated, and the design is implemented using an interactive R Shiny application.

1 Introduction

Stratified sampling divides the population into homogeneous strata and samples independently within each stratum to improve estimator precision.

2 Problem Statement

Determine the required sample size and stratum-wise allocation for population mean estimation under bias, cost, and time constraints.

3 Allocation Strategies

3.1 Proportional Allocation

$$n_h = n\frac{N_h}{N}$$

3.2 Neyman Allocation

$$n_h = n\frac{N_h S_h}{\sum N_h S_h}$$

3.3 Optimised Allocation

$$n_h = n\frac{N_h S_h}{\sqrt{c_h + t_h}}$$

4 Sample Size Formula

$$n = \frac{z^2 \sum W_h^2 S_h^2}{d^2 - b^2}$$

5 Experimental Design

Simple random sampling is applied independently within each stratum. Design plots illustrate population units and sampled units across strata.

6 Implementation

An R Shiny application allows users to select allocation strategies and dynamically visualize stratified sampling designs.

7 Conclusion

Stratified sampling significantly improves efficiency. Neyman and optimised allocations outperform proportional allocation when variances and survey costs vary across strata.