

Sample Size Calculation: general guidance, calculations, documentation, and tools development

Prepared by: Mahmoud Elkasabi PhD, RTI International

Sample size calculation is a critical component of probability surveys, with direct implications for both the survey budget and results. Overestimating the required sample size can place an unnecessary burden on the survey budget, while underestimating it can have severe consequences for the reliability of the results, particularly if the precision falls below acceptable thresholds. Although the methods for calculating sample size are well-documented, not all statisticians possess the necessary experience or tools for accurate sample size determination. This challenge arises from the numerous factors involved in the calculation, including the survey's primary objective (e.g., estimating point values versus tracking changes over time or across domains), the desired level of precision, budget constraints, variable costs, and expected response rates at various stages.

The main objectives of this project are to:

1. Review techniques for sample size calculation used in various surveys conducted by GASTAT.
2. Develop a standardized framework for sample size calculation that can be applied across projects.
3. Prepare a technical report to document the proposed framework.
4. Create tools to support implementation, such as R packages, Shiny applications, SAS programs, or Stata macros.

Additionally, the sampling designs of different surveys will be revised to ensure compatibility with the proposed sample size calculation framework.