Hamed Taherdoost

Chapter 10. Sampling Techniques and Population

This chapter examines the principles and procedures involved in sampling within research, highlighting its essential role in linking populations to empirical investigations. Sampling serves as a fundamental step in research design, providing a practical means of studying large populations by selecting representative subsets. Understanding the relationship between a sample and its population is essential for ensuring the validity and reliability of research findings. The chapter begins by distinguishing between probability and non-probability sampling methods. Probability sampling relies on random selection, offering each member of the population a known chance of inclusion. This approach allows for statistical inference and generalization of results. In contrast, non-probability sampling involves selection based on criteria such as convenience, judgment, or voluntary participation, which may limit generalizability but can be practical in exploratory or qualitative research. Each method is examined in terms of its theoretical foundations, advantages, and limitations, providing readers with the context needed to select appropriate sampling strategies for specific research goals. Determining the appropriate sample size is a critical focus of this chapter. Several factors influence sample size decisions, including the population size, research objectives, acceptable levels of error, and the desired precision of estimates. In addition, the chapter addresses sources of error and bias that can affect sampling outcomes. Sampling error, coverage error, non-response bias, and measurement error are discussed, along with strategies for minimizing their impact. These strategies include careful frame construction, pilot testing, stratification, and systematic selection procedures. Practical considerations are integrated throughout the discussion, with attention to fieldwork logistics, ethical responsibilities, and the use of technological tools in sampling. Constructing a sampling frame, managing data collection, and ensuring participant confidentiality are examined in the context of real-world examples drawn from business research, ecological studies, social sciences, and medical investigations. These examples illustrate how theoretical principles can be applied effectively to diverse research settings. The chapter also emphasizes the iterative and reflective nature of sampling design. Developing, testing, and implementing a sample plan requires careful evaluation and methodological rigor, combining methodological rigor with practical constraints. Sampling is not a purely mechanical process; it is a deliberate design decision that directly affects the accuracy, reliability, and ethical integrity of research results. By integrating theoretical knowledge with applied examples, this chapter equips readers to construct scientifically sound and practically feasible sampling strategies.

10.1. Introduction

This chapter addresses the fundamental concepts and practical applications of sampling in research, emphasizing its role in connecting the target population to empirical data collection for inference. The quality of a sampling approach directly influences the usefulness of study findings, the precision of calculated estimates, and the validity of conclusions regarding causality or generalizability. Even when

estimators are designed to be unbiased, improper sampling can produce misleading results, as skewed or unrepresentative samples fail to reflect the population accurately.

The primary aim of this chapter is to equip readers with both theoretical understanding and practical skills for designing, testing, and implementing sampling procedures across a variety of domains. A wide range of methodologies is discussed, including probability-based approaches that support statistical inference and non-probability-based approaches that are more feasible in resource-constrained settings but require careful interpretation. The chapter emphasizes the importance of informed decision-making in selecting sampling techniques, considering factors such as research objectives, available resources, acceptable levels of uncertainty, and ethical responsibilities. Practical examples, guidance on sample sizes, and tests for common data collection issues are provided to connect methodological principles with practical application.

The focus is on sampling for descriptive estimates, hypothesis testing, and observational research. Advanced spatial point procedures and specialized adaptive clinical trial designs are beyond the scope of this chapter. Core topics include fundamental sample size calculations, principal sampling techniques, random, systematic, stratified, cluster, quota, convenience, targeted, and snowball sampling, and methods to measure and reduce sampling error and bias. Practical application issues such as frame construction, non-response handling, grading, and field planning are illustrated with real-world examples. Cross-disciplinary case studies demonstrate how consistent sampling principles can be applied in diverse research settings, adapting to different objectives and constraints.

To ensure clarity, the chapter defines key terms. The population is the complete set of units from which inferences are drawn, while the target population specifies who is included and excluded. Coverage error occurs when the sampling frame does not accurately represent the target population. The sampling frame is the list or procedure used to select sample units, and a sample unit refers to an individual item chosen, such as a person, family, plot, or machine. The sample is the group actually observed. Probabilistic sampling allows the use of unbiased estimators and variance estimation, whereas non-probability sampling offers limited formal generalizability because selection probabilities are unpredictable or uncontrolled. These definitions form the conceptual foundation for all sampling decisions discussed throughout the chapter.

The chapter also clarifies the distinction between sampling error and non-sampling error. Sampling error refers to discrepancies that arise when a sample is examined instead of the full population and can be reduced through improved design or larger sample sizes. Non-sampling errors, including coverage, measurement, processing, and non-response errors, often account for the majority of total errors in practice and are not mitigated simply by increasing sample size. Bias occurs when an estimator consistently differs from the true population parameter, potentially resulting from inadequate frame construction, flawed selection methods, or inaccurate measurement instruments.

Probability sampling is particularly valuable for formal inference because it has a clear theoretical foundation. When the probability of each unit being selected is known, estimators can be unbiased or have a known bias, and variance estimators enable the construction of confidence intervals and hypothesis tests with predictable coverage properties. Design features such as stratification and clustering affect variance structures and need to be accounted for using weights,

design-based variance formulas, or model-based adjustments. Non-probability samples, by contrast, are suitable for exploratory research, qualitative studies, or situations where probability sampling is impractical or costly. However, these samples require explicit documentation of selection criteria and cannot support definitive population-level conclusions.

Practical guidance is provided for selecting sampling procedures and understanding their limitations. Probability sampling is appropriate when precise population estimates are needed or when findings may be subject to regulatory review or scientific evaluation. Stratification improves accuracy for specific subgroups, while cluster sampling reduces costs when populations are geographically dispersed. Non-probability designs can be useful for hypothesis testing, rapid evaluations, or hard-to-list populations, provided that selection processes are clearly documented, benchmarks are used to assess representativeness, and conclusions are cautiously framed. Throughout the chapter, trade-offs among cost, time, precision, and validity are emphasized to enable informed decision-making in research design (Figure 35).

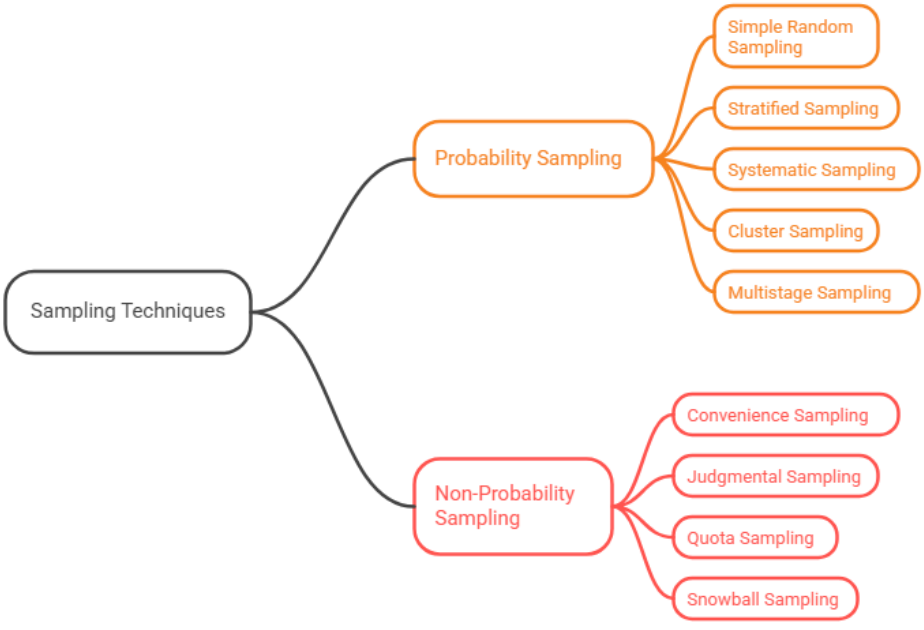


Figure 35. Sampling techniques in research. Source: Figure by author.

10.2. Probability Sampling

Probability sampling is a method in which each unit in the population frame has a known, non-zero probability of selection. This characteristic allows researchers to generate estimators and formal assessments of error, providing a framework for accurate inference. One of the primary advantages of probability sampling is that selection probabilities can be defined, monitored, and applied to calculate estimators and confidence intervals that are compatible with the study design. In practical terms, probability sampling supports more precise generalizations, enables transparent reporting of error margins, and ensures that selection procedures are repeatable.

The simplest form of probability sampling is random sampling, in which every possible sample of a given size has an equal chance of being selected. This approach produces unbiased estimates of population means and proportions, and the calculation of variance is straightforward. Variance decreases as the sample size increases, particularly when the sample represents a small proportion of the population. Simple random sampling is conceptually easy to implement when a complete sampling frame is available, but it becomes costly and impractical for large, heterogeneous, or hard-to-access populations [1,2].

Systematic sampling selects units at regular intervals from an ordered frame, beginning with a randomly chosen starting point. This approach is relatively simple to implement and can approximate simple random sampling when the ordering of units is unrelated to the variable of interest. However, if periodic patterns in the list correspond to outcomes, systematic sampling may introduce bias where random sampling would not [1,2].

Stratified sampling involves dividing the population into homogeneous groups, or strata, and drawing random samples from each stratum, often with unequal allocations. Stratification reduces variance when strata are internally similar but differ significantly from one another, reflecting important differences within the population. It also enables precise estimation for small but important subgroups. In practice, stratification can reduce the total number of samples needed or allow researchers to focus on critical subpopulations [1,3]. Strata should be chosen by using both statistical and substantive judgment, based on characteristics closely related to key outcomes or administrative divisions that facilitate research. Over stratification with numerous small strata may complicate implementation and analysis, producing empty or underfilled strata and unstable variance estimates. Decisions regarding stratification should weigh accuracy improvements against costs and frame limitations.

Cluster sampling organizes units into natural groups, such as households, schools, or geographic areas, and selects entire clusters rather than individual units. Sampling within clusters is often conducted to reduce logistical costs, as visiting clustered units is more efficient than sampling dispersed units. The statistical trade-off is increased in variance because units within clusters tend to be correlated. Multi-stage sampling extends this approach by selecting clusters, then subclusters, and finally individual units, while preserving representativeness across sampling stages. This method is commonly employed in national surveys where complete frames for individuals are unavailable or prohibitively expensive. Maintaining detailed records of selection probabilities at each stage allows calculation of overall inclusion probabilities and the derivation of weights for analysis.

Unequal selection probabilities are sometimes employed to improve efficiency or ensure adequate representation of subgroups. Probability-proportional-to-size (PPS) sampling is a common example, where larger or more variable units have a higher chance of selection. This method allows oversampling of rare or policy-relevant groups, but analysts need to apply weights to account for unequal probabilities, as simple equal-weight estimators would otherwise be biased. Design weights, calculated as the inverse of selection probabilities, are required for producing accurate population estimates, particularly when response rates vary across units. Proper weighting often includes base weights, adjustments for non-response, and calibration against external population totals to minimize bias and improve precision.

Complex probability designs require specialized variance estimation techniques, as standard formulas may not account for clustering, stratification, or unequal weights.

Common methods include Taylor linearization, replication approaches such as the jackknife and bootstrap, and design-based variance formulas with known inclusion probabilities [2,4]. Stratification and weighting are closely linked: when sampling is unequal across strata, final weights need to reflect both selection probabilities and post-sampling adjustments. Purposeful oversampling of minority strata, followed by weighted analysis, ensures estimates remain representative. Transparency in reporting should include unweighted sample sizes, weighted statistics, and effective sample sizes accounting for design effects.

Design effects indicate how complex sampling alters variance relative to simple random sampling. They are influenced by factors such as the degree of clustering, intra-class correlation, and unequal weighting. Design effects are important for planning sample sizes, performing power calculations, and managing research budgets. When using regression or other model-based analyses, probability sampling can be integrated by specifying whether inference is design-based, relying on random mechanisms, or model-based, relying on assumptions about the data. Hybrid approaches, combining design weights with model verification, are often the most effective for reliable results.

Field procedures need to preserve randomization, using random number generators, standardized selection protocols, and monitoring for deviations. Substitutions, operator judgment, or convenience-based selections compromise the integrity of probability sampling. Quality control, audits, and spot checks are required in large-scale applications.

Methods for addressing non-response include follow-ups, weight adjustments based on response likelihood, and imputation of missing data. Sampling plans with unequal probabilities require careful balancing of efficiency and cost, with simulations or pilot studies helpful in evaluating trade-offs. Calibration weighting, which aligns sample estimates with known population statistics, can further reduce bias and adjust for frame coverage errors, but it relies on selecting appropriate control variables available across the target population [2,5].

Maintaining scientific integrity in probability sampling requires full disclosure of the sampling frame, selection procedures, response rates, weighting formulas, and variance estimation methods. Transparent reporting allows secondary analysts to assess robustness, replicate findings, or integrate probability samples with other data sources (Figure 36). Proper documentation is an essential aspect of planning, implementation, and interpretation in applied research.

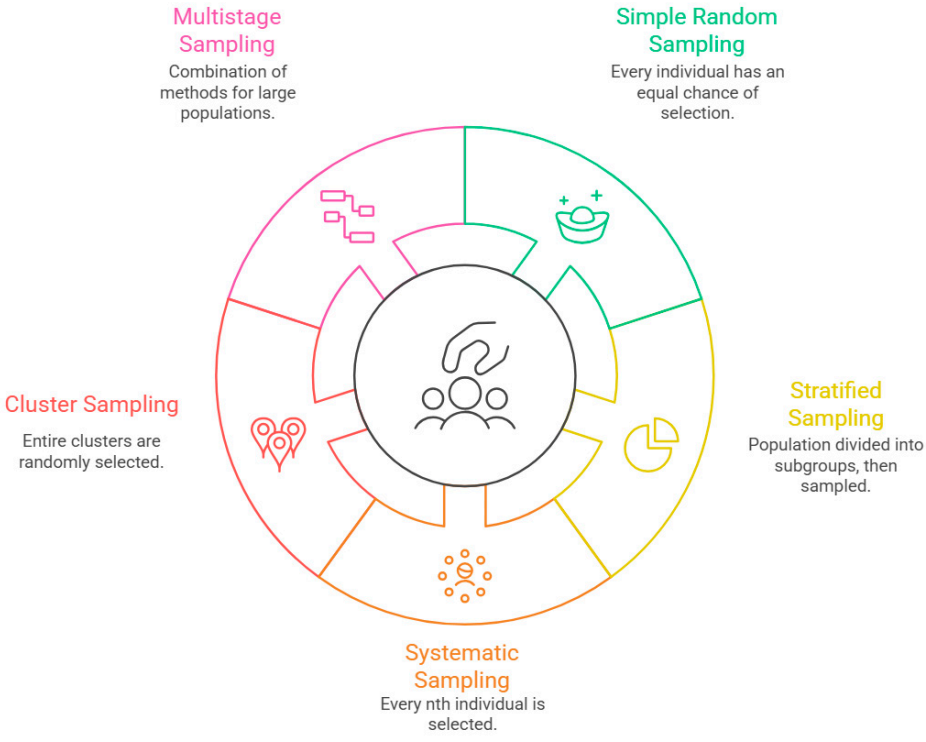


Figure 36. Probability sampling methods. Source: Figure by author.

10.3. Non-Probability Sampling

Non-probability sampling refers to a group of approaches in which units are selected without known or controlled probabilities of inclusion. Because selection probabilities are not defined, these methods do not support traditional design-based inference to a clearly specified target population. Instead, non-probability sampling relies on researcher judgment, accessibility, or social connections to assemble samples that serve specific research purposes. These approaches are commonly used in exploratory studies, hypothesis development, qualitative research, and investigations involving populations that are difficult to enumerate or reach using formal sampling frames. The primary trade-off of non-probability sampling is faster, less costly access to respondents at the expense of representativeness [4,6].

The simplest and most frequently used non-probability method is convenience sampling, in which participants are selected because they are readily available and willing to participate. Typical examples include students recruited from classes, online volunteers responding to open invitations, or individuals approached in public locations. Convenience samples are widely used in pilot studies, instrument development, and early stages of research, where the main objective is to test procedures, refine measurement tools, or identify obvious implementation problems. Their primary advantage lies in efficiency, as data can be collected quickly and at low cost. However, from a theoretical perspective, convenience sampling is often criticized because participants may differ systematically from the intended target population on variables related to study outcomes, leading to selection bias and limited external validity. Despite these limitations, convenience samples can still provide

useful evidence about relationships among variables, particularly when the focus is on internal consistency or theoretical mechanisms rather than population-level estimation, provided that assumptions are clearly stated and evaluated.

Purposive sampling involves the intentional selection of units based on predefined characteristics that are directly relevant to the research question. This approach is especially common in qualitative research, where analytical depth and contextual understanding are prioritized over statistical representativeness. For example, a study examining professional judgment may deliberately recruit individuals with specific roles, qualifications, or levels of experience. When the objective is analytical rather than statistical generalization, purposive sampling provides a coherent strategy for selecting information-rich cases. Well-established purposive techniques include maximum variation sampling, which captures a wide range of perspectives; homogeneous sampling, which focuses on a narrowly defined group; and criterion sampling, which restricts participation to units that meet clearly specified conditions. These approaches allow sampling decisions to be closely aligned with theoretical aims and research design.

Snowball sampling extends purposive selection by using social networks to identify additional participants through referrals. This method is particularly useful when studying hidden, stigmatized, or hard-to-reach populations for which no reliable sampling frame exists, such as individuals involved in informal economic activities, sensitive health conditions, or marginalized communities. Researchers begin with an initial set of participants, often referred to as seeds, and ask them to refer others who meet the study criteria. Recruitment continues through successive waves until sufficient coverage or conceptual saturation is reached. While snowball sampling can improve access and reduce non-response by leveraging trust within networks, it also introduces dependency and bias, as well-connected individuals are more likely to be recruited. Respondent-driven sampling (RDS) was developed to address some of these limitations by formalizing recruitment processes and using reported network sizes to weight estimates, although it relies on strong assumptions about network structure and recruitment behavior that are difficult to verify in practice [7,8].

Quota sampling represents another widely used non-probability approach that combines ease of implementation with controlled sample composition. In quota sampling, researchers set predetermined quotas for selected characteristics such as age, gender, or geographic location, and fieldworkers recruit participants until all quota cells are filled. This approach produces samples that resemble the marginal distributions of the controlled variables and are commonly used in market research and opinion polling. While quota sampling improves upon pure convenience sampling by imposing structure and ensuring representation of key groups, it does not guarantee representativeness beyond the specified quotas. If important predictors of outcomes are excluded from quota controls, or if selection within quota cells is biased, resulting estimates may still be distorted. Post-stratification or adjustment to external benchmarks can reduce some of these limitations, but such corrections depend on reliable auxiliary data and cannot replace probability-based inclusion mechanisms.

Several strategies are available for evaluating and strengthening non-probability samples. One common approach is triangulation, which involves comparing results across multiple non-probability samples or integrating non-probability data with probability-based sources to assess consistency. Detailed documentation of recruitment procedures, inclusion criteria, and fieldwork practices allows readers to judge the plausibility of claims and the likely direction of bias. Sensitivity analyses

that vary assumptions about recruitment processes, non-response, or inclusion rules further demonstrate how conclusions depend on aspects of the sampling process that cannot be directly verified. Together, these practices support more transparent and cautious interpretation of findings derived from non-probability data.

Model-based adjustments may be used to extend inferences from non-probability samples to broader populations, although they require strong assumptions and high-quality auxiliary data. Propensity score weighting uses additional variables to estimate participation probabilities and assign weights that align the sample with a reference population. Outcome modeling, by contrast, estimates relationships between predictors and outcomes to generate population-level predictions. Both approaches depend on the availability of variables that are strongly related to both participation and outcomes. If such variables are poorly chosen or incomplete, adjustments may increase variance or introduce additional bias [9]. Because rich secondary data are not always available, researchers need to balance methodological ambition with practical constraints and clearly communicate the limitations of adjusted estimates.

Ethical and practical considerations play a significant role in the design and implementation of non-probability sampling. Recruitment through social networks or convenience settings may raise concerns about privacy, informed consent, and power relationships, particularly when participants are vulnerable or dependent on gatekeepers. Researchers need to balance the potential social value of the research against risks to participants, often by involving community members in the design of recruitment strategies and consent processes. These considerations are especially important in studies involving stigmatized or marginalized populations.

Non-probability sampling is particularly appropriate when research objectives focus on understanding processes, systems, or experiences, or when time and resource constraints limit the feasibility of probability sampling. Purposive and snowball approaches are commonly used for rapid needs assessments, policy testing, and in-depth examination of lived experiences. In such contexts, consistent and practically useful insights may be more relevant than precise population-level estimates. Nevertheless, analysts need to avoid broad generalizations and clearly state how findings relate to the specific study group and recruitment method.

Reporting standards for non-probability studies should parallel those for probability-based research, including clear descriptions of the target population, recruitment universe, inclusion and exclusion criteria, recruitment procedures, response patterns, and any post-collection adjustments.

Overall, non-probability sampling represents a set of practical tools rather than a universally applicable approach. When applied with careful planning, transparent reporting, and appropriate diagnostic testing, these methods provide access to forms of knowledge that probability sampling alone may not capture. Researchers need to ensure that sampling strategies align with study objectives, clearly communicate uncertainty, and apply available adjustments and sensitivity analyses to support defensible conclusions (Figure 37).

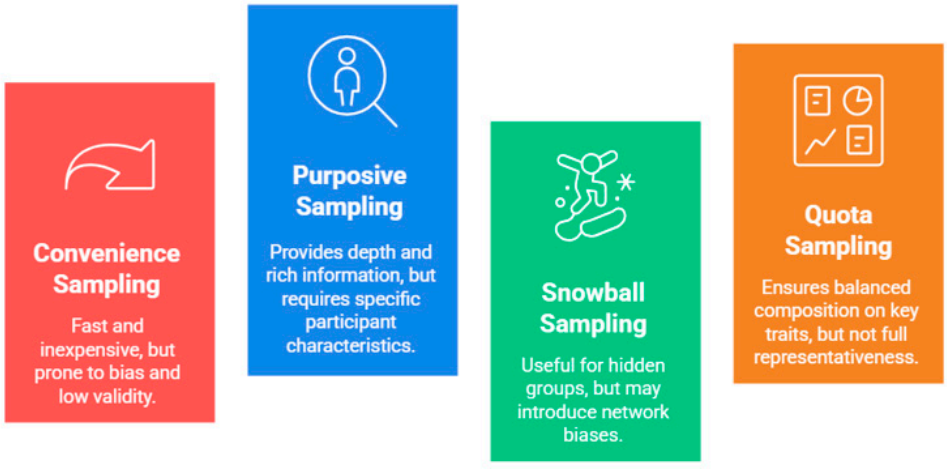


Figure 37. Non-probability sampling methods. Source: Figure by author.

10.4. Comparing Probability and Non-Probability Approaches

Before comparing probability and non-probability sampling approaches, it is important to clarify what each method offers and the conditions under which it operates. Probability sampling designs are characterized by known inclusion probabilities, which allow researchers to construct unbiased estimators and quantify uncertainty using design-based principles. In contrast, non-probability designs do not provide known selection probabilities and therefore cannot rely on design-based inference alone. Instead, they depend on assumptions, statistical models, or external benchmarks to support broader interpretation. This distinction has important implications for how samples are constructed, how findings are interpreted, and how results are communicated to academic, policy, or professional audiences.

One of the primary strengths of probability sampling is its ability to support clear, design-based conclusions about a defined target population. When sampling designs are correctly implemented and accounted for in analysis, confidence intervals and hypothesis tests have interpretable coverage properties that reflect the sampling process. This supports transparent and reproducible analytical decisions. The principal limitation of probability sampling lies in its practical demands. Developing complete sampling frames, implementing randomized selection procedures, and achieving acceptable response rates often require substantial time, financial resources, and administrative effort, which may exceed the capacity of many research projects.

Non-probability sampling offers practical advantages in terms of speed, cost, and accessibility, particularly in studies involving hard-to-reach populations or contexts where no reliable sampling frame exists. These approaches enable researchers to gather data quickly and focus on specific groups or issues of interest. However, the main challenge with non-probability samples is that observed patterns may reflect both genuine relationships and systematic selection effects. Because selection mechanisms are not fully observable or controllable, it becomes more difficult to separate substantive findings from artifacts of recruitment, making causal or population-level statements more uncertain.

Precision and bias operate differently across the two families of sampling methods. In probability designs, precision is influenced by sample size and design

features such as stratification, clustering, and unequal selection probabilities, while bias is minimized through randomization and careful frame construction. In non-probability designs, standard errors may appear small, giving an impression of precision, but such measures often fail to capture systematic bias introduced through self-selection or convenience-based recruitment. As a result, formal indicators of variability may underestimate true uncertainty in non-probability samples.

Method selection is frequently shaped by how well a sampling approach performs in applied settings. Large national or government-sponsored studies typically rely on probability sampling because they require defensible population estimates and procedures that can be replicated and audited. Non-probability samples are more commonly used in pilot research, qualitative studies, and rapid needs assessments, where the priority is timely, focused information rather than formal representation. In these contexts, the ability to gather relevant insights efficiently may outweigh the need for population-level inference.

Cost considerations further differentiate between the two approaches and can be divided into direct costs, secondary costs, and opportunity costs. Probability sampling often increases direct expenses related to survey administration, travel, and follow-up, while reducing uncertainty about external validity. Non-probability approaches tend to reduce initial costs but may incur secondary costs through additional modeling, validation, or repeated data collection to support broader claims. Limited sample sizes may restrict the range of questions that can be addressed.

Ethical and practical considerations also influence sampling decisions. Probability sampling may impose greater burden on respondents through repeated contact attempts or intrusive frame construction, raising concerns about respondent fatigue and privacy. These concerns need to be weighed against the public value of producing reliable and generalizable data. Non-probability recruitment methods may appear less intrusive, but they can introduce ethical challenges related to privacy, referral pressure, unequal exposure to participation requests, or bias in network-based recruitment. Careful design and ethical oversight are therefore required for both approaches.

The choice between probability and non-probability sampling should be guided by the research question and the nature of the conclusions sought. When the objective is to estimate prevalence, totals, or rates within a population, or when findings will inform policy decisions, probability sampling is generally the preferred option. Non-probability sampling may be appropriate when the aim is to explore processes, test instruments, or obtain rapid feedback from a specific group, provided that limitations are clearly acknowledged.

Probability sampling allows researchers to oversample small or rare subgroups using known selection probabilities and apply weights to restore representativeness. Non-probability approaches also facilitate oversampling of uncommon groups, but without known inclusion probabilities, such samples require careful modeling or benchmarking to avoid misleading subgroup conclusions. Hybrid designs that combine probability-based core samples with targeted non-probability oversamples offer a practical compromise in some settings. By integrating these data sources through appropriate weighting and modeling, researchers can obtain population-level estimates while also gaining detailed insights into specific subgroups. For such designs to be effective, it is necessary to clearly identify the source of each observation and ensure that analytical methods account for the different selection mechanisms involved.

Probability and non-probability approaches also differ in their reliance on statistical modeling. Probability sampling supports design-based estimators that remain valid under a wide range of model misspecifications, as inference is grounded in the randomization process. Non-probability sampling, by contrast, requires stronger modeling assumptions to support generalization, including assumptions about selection mechanisms, relationships between auxiliary variables and outcomes, and the stability of these relationships across contexts.

Regardless of the approach, researchers should establish a clear strategy for identifying, assessing, and communicating these risks before data collection begins.

When non-probability designs or hybrid approaches are used, diagnostics and validation procedures become essential. Comparing sample distributions to reliable external benchmarks, testing the stability of estimates under alternative weighting schemes, and using administrative or auxiliary data for consistency checks can help identify potential biases and strengthen confidence in results (Table 6). Together, these practices improve interpretation and responsible use of findings derived from both probability and non-probability sampling methods.

Table 6. Comparison of probability and non-probability sampling approaches.

Dimension	Probability Sampling	Non-Probability Sampling
Core Principle	Known, non-zero inclusion probabilities; selection driven by randomization.	Unknown inclusion probabilities; selection driven by convenience, networks, or researcher judgment.
Inference Logic	Enables design-based inference with unbiased estimators and measurable uncertainty.	Requires model-based inference, assumptions, or external benchmarks to support generalization.
Strength of Conclusions	Supports clear, population-wide conclusions, valid confidence intervals, and hypothesis tests.	Conclusions may be ambiguous because patterns may arise from both selection bias and real phenomena.
Bias and Precision	Bias controlled through frame construction, randomization, stratification, and clustering; precision driven by sample size and design.	Small standard errors may mask systematic bias caused by selection processes; precision does not guarantee accuracy.
Cost and Practicality	Higher cost and longer timelines due to frame development, random selection, and response-rate management.	Lower cost, faster, and more accessible, especially when no sampling frame exists.

Table 6. *Cont.*

Dimension	Probability Sampling	Non-Probability Sampling
Common Use Cases	Large-scale federal surveys, national statistics, and policy-relevant studies require authoritative estimates.	Pilot studies, exploratory research, rapid assessments, hard-to-reach or undefined populations.
Ethical and Practical Concerns	May impose burdens on respondents (multiple contacts, formal sampling frames).	Easier recruitment but risks privacy concerns, biased referrals, or uneven exposure to survey requests.
Suitability by Research Goal	Best for estimating prevalence, totals, or providing evidence for policy or population decisions.	Best for understanding processes, testing tools, or gathering immediate insights from specific groups.
Subgroup Analysis	Allows planned oversampling with known weights to recover representativeness.	Easier oversampling of rare groups but requires careful modeling to avoid misleading subgroup conclusions.
Hybrid Designs	Provides the population-based component of mixed designs.	Supplements probability samples for cost-effective subgroup insights; requires proper integration and weighting.
Model Dependence	Less model-dependent; robust to misspecification due to randomization.	More model-dependent; requires assumptions about selection, covariate relationships, and transportability.
Risk Profile	Lower inferential risk; higher resource risk (time, money, logistics).	Higher inferential risk; mitigated through diagnostics, reweighting checks, and external validation.
Quality Assurance	Validated largely through design consistency and response-rate management.	Requires diagnostics, including margin comparisons, reweighting sensitivity tests, and administrative data checks.

Source: Table by author.

10.5. Sample Size Determination

Sample size determination complements data collection decisions by translating research objectives into quantitative requirements for precision and inference. Researchers need to first establish whether the study aims to estimate population parameters with a specified level of precision, test hypotheses with adequate statistical

power, or address both objectives concurrently. Precision-based estimation focuses on how accurately population quantities can be measured, whereas power-based testing emphasizes the ability to detect effects of a given magnitude. Because these objectives impose different constraints on the number of observations required, clarity at this stage is required for aligning methodological choices with research goals, resource availability, and interpretability of results.

Precision is commonly operationalized through the width of confidence intervals and is governed by the standard error of the estimate. Under simple random sampling, the standard error increases with population variability and decreases with the square root of the sample size, implying diminishing returns as sample size grows. In practical terms, achieving noticeably narrower confidence intervals often requires a substantial increase in observations, such as quadrupling the sample size to halve the margin of error. This relationship has direct implications for budgeting and fieldwork planning, particularly when more complex sampling designs introduce additional variance through clustering, stratification, or unequal selection probabilities, which need to be incorporated into sample size calculations [1,2].

Statistical power addresses a different aspect of sample adequacy by focusing on the likelihood of detecting a true effect at a chosen significance level. Power calculations require researchers to specify the smallest effect size of substantive interest, the acceptable probability of Type I error, and assumptions about outcome variability. These elements jointly determine the minimum sample size required to support meaningful hypothesis testing. Studies that fail to consider power adequately risk producing inconclusive findings, not because effects are absent, but because the design lacks sufficient sensitivity. As a result, power-based planning is particularly important for ensuring that empirical tests provide reliable and interpretable evidence.

Effect size is a key parameter underlying both precision- and power-based planning, as it reflects the magnitude of the relationship or difference the study seeks to estimate or detect. Smaller effect sizes require larger samples to be identified with acceptable certainty, while larger effects may be detectable with fewer observations. Selecting an appropriate effect size involves both scientific judgment and practical reasoning, drawing on theory, prior empirical findings, and the relevance of potential outcomes for policy or decision-making. Researchers need to balance the desire to detect subtle but meaningful effects with the constraints imposed by time, cost, and data collection feasibility.

Sample size calculations differ depending on whether the outcome of interest is a mean, a proportion, or a comparative measure such as a difference between groups or a correlation coefficient. For estimating means, calculations typically depend on the desired margin of error, an estimate of the population standard deviation, and the chosen confidence level. In applied research, standard deviation estimates are often derived from prior studies, pilot data, or expert knowledge. Overly optimistic assumptions about variability can result in insufficiently precise estimates, whereas overly conservative assumptions may lead to unrequired increases in cost. Similar considerations apply to proportions, where variance is maximized at 0.5. When the true proportion is unknown, this value is commonly used to ensure adequate sample size, although informed estimates may be substituted when justified and clearly documented.

For hypothesis testing involving group comparisons or associations, sample size planning incorporates expected variance within groups, allocation ratios, and desired levels of power and significance. Unequal allocation may be appropriate when sampling costs differ across groups or when certain subpopulations are rare, provided that the analytical approach can accommodate such imbalances. An alternative planning framework focuses on confidence intervals rather than hypothesis testing, specifying an acceptable margin of error around an estimate and determining the sample size required to achieve it. This approach is particularly suitable for descriptive and prevalence studies, where the primary concern is estimation accuracy rather than statistical significance.

Adjustments to sample size are often required to account for population structure and design features. When the sample represents a substantial proportion of the target population, finite population correction reduces variance and permits smaller samples, particularly in studies of bounded populations such as organizations or schools. Design effects further modify required sample sizes by capturing the impact of clustering, stratification, and weighting on variance. Clustering generally increases required sample sizes due to intracluster correlation, with stronger effects observed for larger clusters and higher within-cluster similarity. Stratification, by contrast, can reduce required sample sizes when strata are internally homogeneous, while also supporting reliable subgroup analysis through controlled oversampling and appropriate weighting [1,3].

Finally, realistic sample size determination needs to account for non-response and broader practical constraints. Anticipated refusals and noncontacts increase the number of units that need to be approached to achieve the desired analytic sample, often requiring inflation of initial targets based on expected response rates. More advanced planning models address differential non-response across subgroups and incorporate targeted follow-up strategies to preserve representativeness. Throughout this process, researchers need to balance statistical requirements with financial, logistical, and ethical considerations. Pilot studies, sequential designs, and adaptive re-estimation strategies offer structured ways to manage uncertainty, improve variance estimates, and adjust sample sizes during data collection without compromising inferential integrity, provided that these procedures are clearly specified in advance (Figure 38).

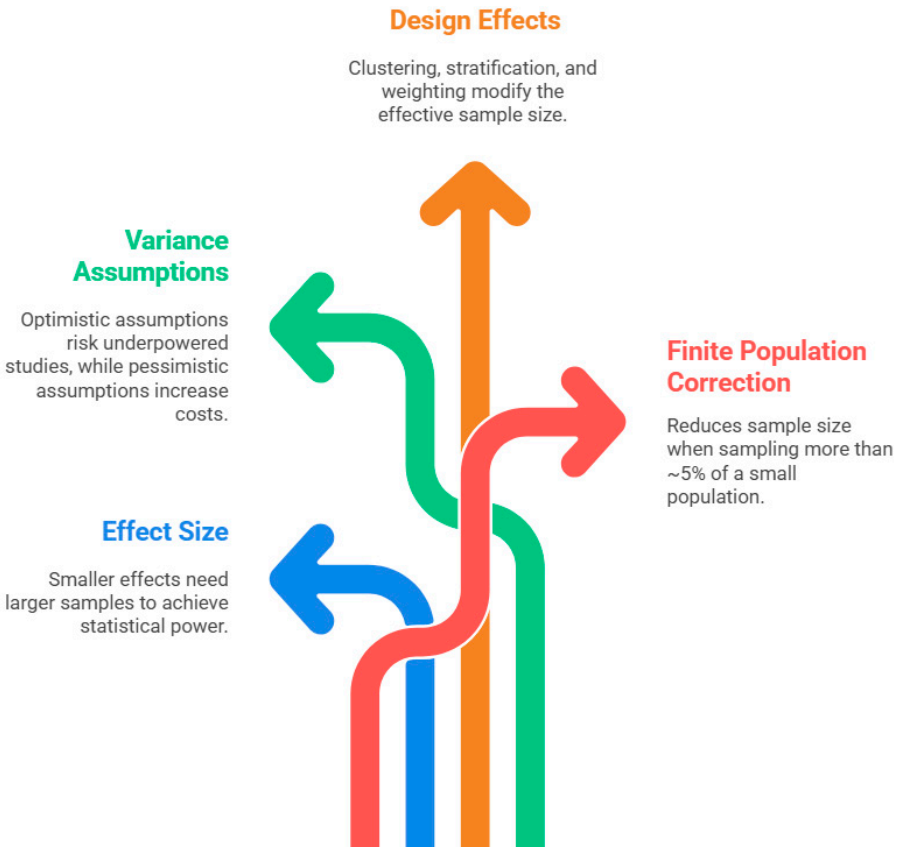


Figure 38. Four key factors shaping sample size: design effects, variance assumptions, effect size, and population correction. Source: Figure by author.

10.6. Sampling Errors and Biases

Sampling error refers to the random variation that arises when researchers observe a subset of a population rather than the population in its entirety. This form of error is inherent in all sample-based studies and reflects the fact that different samples drawn from the same population will yield different results, even when the sampling design is correctly implemented. Sampling distributions provide a theoretical framework for understanding this variation and demonstrate that sampling error decreases as sample size increases, following well-established mathematical relationships. This principle underpins the construction of confidence intervals, the planning of adequate sample sizes, and the interpretation of observed differences. A critical distinction needs to be made between random sampling error, which diminishes with larger samples, and systematic bias, which persists regardless of sample size and therefore requires different corrective strategies.

In contrast, non-sampling error encompasses all sources of inaccuracy that are not attributable to random sampling variation and often represents the dominant source of total error in applied research. These errors include coverage problems, non-response, measurement inaccuracies, data processing mistakes, and respondent misreporting. Unlike sampling errors, non-sampling errors are not reduced simply

by increasing the number of observations, as they arise from behavioral factors, institutional constraints, and operational decisions. Their magnitude and direction are therefore more difficult to quantify using theory alone. In practice, the most effective safeguards against non-sampling error involve careful study design, pilot testing, interviewer training, quality control procedures, and transparent documentation throughout the research process.

Coverage error occurs when the sampling frame does not align fully with the defined target population, resulting in missing, duplicated, or misclassified units. In such cases, estimates pertain to the frame rather than the population of interest, and bias arises when excluded or overrepresented units differ systematically on key study variables. Coverage problems frequently stem from outdated lists, incomplete administrative records, or the exclusion of populations that are difficult to enumerate, such as highly mobile individuals, institutionalized populations, or those without reliable internet access. Addressing coverage error may involve updating or combining frames, using mixed-frame designs, or applying post-survey adjustments when reliable auxiliary data are available.

The construction and evaluation of sampling frames therefore play a central role in minimizing coverage-related bias. Well-documented frames clearly specify their sources, reference periods, inclusion criteria, and known limitations. Analytical evaluation requires assessing whether the characteristics of the frame correspond to those of the target population for outcomes of interest. When no single frame is sufficient, mixed-frame approaches, such as combining address-based samples with online panels or dual-frame telephone surveys, can improve coverage. However, these strategies introduce additional complexity in weighting and variance estimation, which needs to be explicitly addressed during both design and analysis.

Selection bias arises when the mechanism determining inclusion in the sample is correlated with the outcome being measured, leading to systematic deviations from true population values. From a theoretical perspective, such bias undermines the validity of design-based inference because selection probabilities no longer fully govern sample composition. In practice, selection bias may occur during frame construction, through deviations from prescribed sampling procedures, or when substitutions are made for convenience during fieldwork. Preventing selection bias requires strict adherence to randomization rules, clear protocols for handling noncontact, and strong oversight of field operations to discourage discretionary substitutions.

A related but distinct concern is self-selection bias, which occurs when individuals choose to participate based on characteristics that are relevant to study outcomes. This form of bias is common in opt-in surveys, volunteer panels, and recruitment methods that rely on advertising or open invitations. Self-selection disrupts the link between the sampling mechanism and the target population, and simple weighting adjustments are often insufficient unless rich auxiliary data are available. Mitigation strategies include the use of experimental designs when feasible, model-based adjustments informed by external data, and cautious interpretation that avoids extending conclusions beyond the observed sample.

Non-response bias occurs when sampled units fail to provide data and nonrespondents differ systematically from respondents with respect to key variables. The extent of this bias depends on the strength of the relationship between response propensity and the outcome of interest, and even modest correlations can produce substantial distortions. Non-response may occur at the unit level, when individuals decline to participate entirely, or at the item level, when specific questions are skipped. Both forms reduce effective sample size and can compromise inference if not properly addressed.

Strategies for addressing non-response range from design-based planning to post-collection adjustments. At the planning stage, expected non-response can be accommodated by inflating initial contact targets based on anticipated response rates. More advanced approaches include response propensity modeling, weighting adjustments using auxiliary variables, and multiple imputations for item non-response. These methods rely on the assumption that variables used in adjustments are related to both participation and outcomes, which underscores the importance of collecting relevant auxiliary information and evaluating model assumptions carefully.

Measurement error represents another important source of bias that is often intertwined with sampling decisions. The mode of data collection, question wording, survey context, and respondent burden can all influence responses in systematic ways. From a theoretical standpoint, measurement error affects estimates by increasing variance or attenuating associations between variables. Differences between interviewer-administered and self-administered modes, as well as device-related effects in online surveys, illustrate how data collection methods interact with respondent behavior. Paradata, such as response times or device types, can provide useful insights into these mode-related differences.

Interviewer effects further contribute to measurement error when interviewers vary in tone, probing behavior, or adherence to protocols. In analytical models, interviewers may be treated as random effects, increasing variance and potentially introducing bias if interviewer characteristics correlate with respondent traits. In practice, interviewer-related errors can be reduced through standardized training, scripted instruments, supervision, performance monitoring, and validation checks such as audio reviews or reinterviews.

Identifying potential biases requires systematic diagnostic analysis. Descriptive comparisons between sample distributions and known population benchmarks, examination of response rates across subgroups, and analysis of missing data patterns can reveal areas of concern. Integrating routine diagnostics into the data collection and processing workflow allows researchers to respond promptly through targeted follow-up or analytical adjustments.

Assessing the magnitude of bias often requires external data or validation studies. Conceptually, total bias can be decomposed into coverage bias, selection bias, measurement error, and processing error [4]. When linked data or validation subsamples are available, these components can be estimated or bounded under explicit assumptions. Small validation studies using high-quality reference measures or administrative records are particularly valuable for evaluating measurement error and non-response bias, informing both correction strategies and sensitivity analyses.

Bias mitigation needs to be addressed across all stages of research, including design, fieldwork, and analysis. Design-stage measures include selecting appropriate frames, oversampling underrepresented groups, and planning weighting strategies

in advance. Fieldwork mitigation focuses on maximizing response through multiple contact attempts, mixed contact modes, ethically appropriate incentives, and real-time monitoring of response patterns, with resources directed toward groups with the greatest risk of bias. Analytical mitigation involves weighting, calibration, model-based adjustments, and imputation, each of which depends on assumptions that n to be examined and reported transparently.

Sensitivity analysis provides a structured way to assess the robustness of findings under alternative assumptions. Rather than presenting a single adjusted estimate, researchers can show how results vary across plausible scenarios involving non-response, coverage gaps, or measurement error. Such transparency supports more informed interpretation and reduces the risk of overconfidence in uncertain estimates (Table 7).

Documentation is both a practical and ethical requirement in studies affected by sampling error and bias. Detailed records of sampling procedures, frame limitations, contact attempts, response rates, weighting methods, and diagnostic findings enable replication, external review, and cumulative learning. In applied settings, such documentation also supports accountability and informed decision-making. Because multiple sources of error often interact, reducing bias requires planning across all stages of the study that anticipates common combinations of errors and prioritizes interventions with the greatest potential impact on data quality (Figure 39).

Table 7. Summary of sampling and non-sampling errors: types, causes, impacts, and mitigation strategies in survey research.

Category	Definition	Examples/ Types	Causes	Impact on Results	Mitigation/ Solutions
Sampling Error	Random differences due to studying a subset rather than the whole population.	N/A (inherent to sampling)	Small sample size; natural variability	Produces variability across samples; affects confidence intervals and hypothesis tests	Increase sample size, employ a proper sampling design, and understand variance and standard errors.
Non-sampling Error	Non-sampling errors often exceed sampling errors and persist regardless of sample size; non-sampling error often exceeds sampling error.	Coverage error, selection bias, self-selection bias, non-response bias, item non-response, measurement error, processing error	Incomplete frames, participant behavior, survey administration flaws, and institutional limitations	Biased estimates may misrepresent the population	Careful study design, piloting, quality control, documentation, diagnostics, weighting, and imputation
Coverage Error	When the sampling frame does not match the target population.	Missing, duplicated, or misrepresented units	Outdated lists, excluded populations, and administrative data gaps	Estimates rely on frames, not the true population; this introduces bias	Refine frames, combine multiple frames, and make post-survey adjustments
Selection Bias	The sampling process itself causes systematic differences.	N/A	Non-random selection, substitution of units	Misleading sample estimates; invalid design-based inference	Strict randomization, avoid substitutions, consistent field protocols

Table 7. *Cont.*

Category	Definition	Examples/ Types	Causes	Impact on Results	Mitigation/ Solutions
Self-Selection Bias	Bias due to voluntary participation in relation to the variables studied.	Online panels, volunteer programs, and advertising recruitment	Participants opt in based on interest or relevance	Breaks the link between the sample and population, undermines weighting	Additional data for adjustment, experimental designs, cautious interpretation
Non-response Bias	Bias from nonparticipation or partial participation.	Unit non-response, item non-response	Low response rates, selective participation	Reduces effective sample size; distorts population estimates	Follow-ups, response propensity modeling, weighting adjustments, and multiple imputation

Source: Table by author.



Figure 39. Five common sources of bias in data collection. Source: Figure by author.

10.7. Practical Implementation Issues

Applying sampling methods in real-world research begins with the construction of a sampling frame, defined as the operational list of units from which the sample is drawn. In theory, this frame should correspond exactly to the target population; however, such alignment is uncommon in practice. Researchers therefore need to evaluate whether the frame is complete, current, and free from duplication, while also identifying known limitations. For instance, household surveys that rely on address lists may systematically exclude individuals who move frequently, leading to coverage errors. Recognizing these deficiencies at the outset is essential, as unexamined frame problems can introduce bias that persists throughout the study if not addressed through design adjustments or analytical corrections.

Evaluating the quality of a sampling frame requires both statistical assessment and substantive knowledge of the population being studied. Statistical comparisons between frame distributions and trusted population benchmarks help identify undercoverage or overcoverage of key groups, while subject-matter expertise clarifies whether omitted groups are relevant to the research objectives. Frames designed for general population studies may be insufficient when detailed subgroup analysis is required, particularly for minority or rare populations. To improve coverage, researchers often combine multiple frame sources, a practice that is common in large-scale surveys. Dual-frame telephone studies, which incorporate both landline and mobile phone lists, illustrate how coverage can be expanded to reduce bias. While this approach improves representativeness, it also increases methodological

complexity, as inclusion probabilities differ across frames and some units may appear in more than one source. As a result, accurate recordkeeping is required to identify each unit's origin and manage overlap appropriately. Researchers need to weigh the benefits of broader coverage against the additional cost, logistical effort, and analytical complexity introduced by multiple frames.

Once a sampling frame has been established, attention shifts to fieldwork, which involves contacting sampled units, securing participation, and collecting data in accordance with the design. Although sampling theory assumes strict adherence to selection rules, field conditions frequently introduce deviations. Interviewers may encounter unavailable respondents, restricted access to locations, or pressure to substitute easier-to-reach units. If such deviations are not closely monitored, they can compromise randomization and introduce selection bias. Effective fieldwork therefore requires clear protocols, supervision, and mechanisms for identifying and correcting departures from the planned design.

Quality control during data collection plays a central role in preserving the integrity of the sampling process. Interviewers need to be trained to apply selection criteria consistently, record all contact attempts accurately, and document refusals without making unauthorized substitutions. Supervisors are responsible for monitoring compliance and reinforcing the importance of randomization as the foundation of valid inference. In practice, technologies such as GPS tracking, call logs, and audio recordings support quality control by enabling verification of interviewer activity and early detection of procedural problems.

Fieldwork logistics further influence how sampling designs are implemented. Scheduling interviews, allocating personnel, and organizing travel routes are required to control costs and maintain feasibility, particularly in large-scale studies. Clustering respondents geographically can reduce travel expenses, but it also increases design effects that need to be addressed analytically. Researchers therefore need to balance operational efficiency against statistical precision, ensuring that logistical decisions are aligned with the assumptions underlying variance estimation and weighting procedures.

Ethical considerations are integral to the practical application of sampling methods. Informed consent is both a legal and ethical requirement, intended to ensure that participation is voluntary and based on adequate understanding of the study. While theoretical frameworks emphasize fully informed and autonomous decision-making, consent procedures need to be adapted to the research context. Written consent may be appropriate in some settings, while verbal consent may be more practical in others. Researchers need to ensure that consent processes are accessible to individuals with different literacy levels and that participation is free from coercion.

Ethical sampling also requires special attention to vulnerable or marginalized populations. Recruitment strategies should minimize the risk of harm, stigmatization, or unintended disclosure of personal information. For example, using social networks to recruit participants may inadvertently reveal sensitive affiliations. Ethical frameworks such as the Belmont Report emphasize respect for people, beneficence, and justice, guiding researchers to design sampling procedures that protect participants while still meeting research objectives [10]. These principles should inform decisions about whom to sample, how often to make contact, and how to communicate study purposes.

Informed consent considerations also affect sampling intensity and follow-up strategies. Repeated contact attempts with rare or hard-to-reach groups may be perceived as intrusive, even when statistically justified. Researchers therefore need to balance the need for accurate estimates with respect for participants' time, privacy, and safety. Incentives can encourage participation, but they should be proportionate and structured to avoid undue influence. Transparency regarding oversampling practices and study goals helps build trust and supports ethical engagement with participants.

Technological developments have significantly reshaped sampling implementation, particularly through the use of online panels. These panels allow researchers to access large numbers of respondents quickly and to target recruitment based on demographic or behavioral characteristics. Although online panels are generally non-probability samples, they can be adjusted using weighting and modeling techniques to approximate population distributions. Their speed and lower cost make them attractive for market research and opinion studies, although limitations in representativeness need to be acknowledged.

Assessing the quality of online samples presents specific challenges. Individuals who join panels may differ systematically from the broader population, and not all biases can be corrected through weighting alone. Researchers need to evaluate whether available auxiliary data are sufficient for calibration and whether resulting estimates are appropriate for the intended use. While theoretical debates continue regarding the validity of online panels, practical experience indicates that they can provide useful information when their limitations are clearly understood and transparently reported.

Technology also supports field operations and quality assurance across sampling modes. Mobile data collection reduces processing errors by eliminating manual data entry, while GPS verification confirms interviewer presence at sampled locations. Real-time monitoring systems enable supervisors to track response rates and adjust field strategies as needed. Beyond panels, digital recruitment methods such as social media advertising enable access to populations that are difficult to reach through traditional sampling. Digital sampling methods also raise important questions about privacy and consent. Online recruitment often involves tracking user behavior or targeting individuals based on personal data. Researchers need to ensure that participants understand how they were identified, what data are being collected, and how those data will be used. Consent procedures should be adapted to digital contexts, emphasizing clarity, transparency, and accessibility. Maintaining public trust requires careful attention to data protection and ethical accountability.

Overall, implementing sampling methods in practice involves the integration of statistical principles, operational planning, ethical responsibility, and technological tools. Decisions made in one area often influence outcomes in others, requiring researchers to consider these factors together. Effective sampling designs are those that are theoretically sound, operationally feasible, ethically defensible, and transparent in their limitations, allowing researchers to produce credible and informative evidence in real-world settings.

10.8. Sample Size Formulas

Determining appropriate sample sizes for statistical estimation or hypothesis testing requires the use of explicit formulas that relate the desired precision, confidence level, variability in the data, and other study parameters. Sample size

formulas differ depending on whether the objective is to estimate a population mean, to estimate a proportion, or to provide adequate power for detecting differences between groups. Understanding these formulas enables researchers to align sample size decisions with study objectives and statistical requirements.

For estimation of a population proportion, a commonly used formula is based on the normal approximation to the binomial distribution. In this context, the standard error of a sample proportion $\hat{p}$ is $\sqrt{p(1-p)/n}$, where p is the true population proportion and n is the sample size. When p is unknown, a conservative estimate of 0.5 is often used because it maximizes variance and leads to the largest sample size requirement. Under these assumptions, requiring a margin of error of B and using $Z_{1-\alpha/2}$ as the critical value for a chosen confidence level (for example, 1.96 for a 95 percent confidence level), the sample size formula becomes

$$n = \frac{Z_{1-\alpha/2}^2 \times p(1-p)}{B^2}.$$

When p is unknown and set to 0.5 as the most conservative estimate, this simplifies to

$$n = \frac{Z_{1-\alpha/2}^2}{4B^2}.$$

This form indicates that as the confidence level increases or the margin of error decreases, the required sample size increases proportionally to the square of the critical value and inversely with the square of the allowable error.

For estimating a mean, sample size is derived from the desired precision around the mean estimate. If σ denotes the population standard deviation and W denotes the desired total width of the confidence interval, then the margin of error is $W/2$, and the sample size can be obtained from

$$\frac{Z_{1-\alpha/2} \times \sigma}{\sqrt{n}} = \frac{W}{2}.$$

Solving this equation for n yields

$$n = \frac{4Z_{1-\alpha/2}^2 \times \sigma^2}{W^2}.$$

This formula emphasizes the relationship between variability, confidence level, and desired precision. When the population standard deviation is unknown, an estimate from prior studies or pilot data is used; otherwise, iterative procedures based on the t-distribution may be required, especially in small samples.

When research questions involve hypothesis testing, sample size calculations integrate both the significance level (Type I error rate α) and the statistical power (1–Type II error β). For example, in tests comparing two independent means, parameters include the expected difference in means Δ, the standard deviation σ, the pooled variance, and the allocation ratio between groups. Although the exact

formula depends on the distributional assumptions and design, a general expression for a two-sample mean comparison under normality is derived from

$$n = \frac{\left(Z_{1-\alpha/2} + Z_{1-\beta}\right)^2 2\sigma^2}{\Delta^2}.$$

Here, $Z_{1-\beta}$ corresponds to the critical value associated with desired power and $Z_{1-\alpha/2}$ to the confidence level. Larger expected differences or lower variability reduce the required sample size, while higher power or tighter significance thresholds increase it. Researchers often use software or iterative methods to refine this calculation, particularly when variances differ between groups or when using t-based tests rather than normal approximations.

For tests comparing two proportions, analogous formulas apply. Let p_1 and p_2 denote the proportions in the two groups and let $\hat{p}$ represent their pooled estimate. A simplified form of the required sample size per group for detecting a difference in proportions with specified α and power is given by

$$n = \frac{\left(Z_{1-\alpha/2}\sqrt{2\hat{p}(1-\hat{p})} + Z_{1-\beta}\sqrt{p_1(1-p_1) + p_2(1-p_2)}\right)^2}{(p_1 - p_2)^2}.$$

This expression incorporates both sampling variability and the magnitude of difference the study is designed to detect. Estimators for the pooled and individual variances appear in the numerator, and the squared effect size appears in the denominator.

In all these formulas, finite population correction may be applied when the sample constitutes a substantial fraction of the target population. The adjusted sample size n_{adj} can be calculated from an initial estimate n_0 as

$$n_{ndj} = \frac{n_o}{1 + \frac{n_0 - 1}{N}},$$

where N denotes the total population size. This adjustment reduces the required sample size when the population is small relative to planned sampling, reflecting the reduced variance associated with sampling a large proportion of the population.

Online calculators, such as those available at calculator.net or other sample size tools, implement these formulas and allow researchers to explore parameters such as confidence level, margin of error, expected variance, and effect size influence required sample sizes. These tools often provide iterative solutions to account for unknown parameters and can incorporate finite population correction automatically when population size is specified.

Sample size formulas provide a mathematical basis for determining the number of observations needed to achieve specified precision or power in research. By linking confidence levels, effect sizes, variance, and study goals, these expressions form the quantitative foundation for planning statistically sound sample designs and justify sample size choices in method sections of research reports.

Sampling plays an important role in empirical research by shaping how data are collected, analyzed, and interpreted. Sample size determination illustrates the continuous interaction between theoretical guidance and practical constraints. Decisions about sample size therefore involve balancing desired statistical quality with what is operationally feasible. Ethical considerations further emphasize that sampling is not solely a technical process. Each sampling decision involves interaction with individuals and communities, requiring respect for autonomy, privacy, and fairness. This is particularly relevant when studies involve vulnerable or marginalized groups, where recruitment strategies and follow-up procedures are needed to minimize burden and risk. Ethical principles guide researchers to balance the pursuit of reliable evidence with responsibility toward participants, reinforcing the need for informed consent, transparency, and proportional use of incentives.

Technological developments have expanded the range of available sampling tools, including online panels, digital recruitment methods, and real-time quality control systems. These innovations have increased efficiency and reduced costs, but they have also introduced new challenges related to representativeness, data protection, and consent. Experiences across disciplines show that no single sampling approach is suitable for all research contexts. Instead, effective sampling strategies are those that are adapted to the study's objectives, population, and constraints, while remaining explicit about their limitations.

In summary, effective sampling requires the integration of statistical reasoning, operational planning, ethical awareness, and methodological transparency. It involves applying established principles while remaining responsive to practical realities and evolving research environments. By carefully aligning sampling choices with research goals and clearly communicating assumptions and limitations, researchers can produce findings that are credible, useful, and appropriate for informing theory, policy, and practice.

References

1. Cochran, W.G. *Sampling Techniques*, 3rd ed.; Wiley: New York, NY, USA, 1977.
2. Lohr, S.L. *Sampling: Design and Analysis*, 3rd ed.; Chapman & Hall/CRC: Boca Raton, FL, USA, 2021.
3. Kish, L. *Survey Sampling*; Wiley: New York, NY, USA, 1965.
4. Groves, R.M.; Fowler, F.J.; Couper, M.P.; Lepkowski, J.M.; Singer, E.; Tourangeau, R. *Survey Methodology*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2009.
5. Deville, J.C.; Särndal, C.E. Calibration estimators in survey sampling. *J. Am. Stat. Assoc.* **1992**, *87*, 376–382. [CrossRef]
6. Kalton, G. *Introduction to Survey Sampling*; SAGE Publications: Thousand Oaks, CA, USA, 1983.
7. Heckathorn, D.D. Respondent-driven sampling: A new approach to the study of hidden populations. *Soc. Probl.* **1997**, *44*, 174–199. [CrossRef]
8. Heckathorn, D.D. Respondent-driven sampling II: Deriving valid population estimates from chain-referral samples of hidden populations. *Soc. Probl.* **2002**, *49*, 11–34. [CrossRef]

9. Rosenbaum, P.R.; Rubin, D.B. The central role of the propensity score in observational studies for causal effects. *Biometrika* **1983**, *70*, 41–55. [CrossRef]
10. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*; U.S. Department of Health, Education, and Welfare: Washington, DC, USA, 1979. Available online: https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/ (accessed on 24 July 2026).

© 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).