



Standard Guide for Sampling Design¹

This standard is issued under the fixed designation E1402; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ϵ) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This guide defines terms and introduces basic methods for probability sampling of discrete populations, areas, and bulk materials. It provides an overview of common probability sampling methods employed by users of ASTM standards.

1.2 Sampling may be done for the purpose of estimation, of comparison between parts of a sampled population, or for acceptance of lots. Sampling is also used for the purpose of auditing information obtained from complete enumeration of the population.

1.3 No system of units is specified in this standard.

1.4 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety, health, and environmental practices and determine the applicability of regulatory limitations prior to use.*

1.5 *This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.*

2. Referenced Documents

2.1 ASTM Standards:²

- D7430 Practice for Mechanical Sampling of Coal
- E105 Guide for Probability Sampling of Materials
- E122 Practice for Calculating Sample Size to Estimate, With Specified Precision, the Average for a Characteristic of a Lot or Process
- E141 Practice for Acceptance of Evidence Based on the Results of Probability Sampling
- E456 Terminology Relating to Quality and Statistics

¹ This guide is under the jurisdiction of ASTM Committee E11 on Quality and Statistics and is the direct responsibility of Subcommittee E11.10 on Sampling / Statistics.

Current edition approved Nov. 1, 2023. Published November 2023. Originally approved in 2008. Last previous edition approved in 2018 as E1402 – 13 (2018). DOI: 10.1520/E1402-13R23.

² For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

3. Terminology

3.1 *Definitions*—For a more extensive list of statistical terms, refer to Terminology E456.

3.1.1 *area sampling, n*—probability sampling in which a map, rather than a tabulation of sampling units, serves as the sampling frame.

3.1.1.1 *Discussion*—Area sampling units are segments of land area and are listed by addresses on the frame prior to their actual delineation on the ground so that only the randomly selected ones need to be exactly identified.

3.1.2 *bulk sampling, n*—sampling to prepare a portion of a mass of material that is representative of the whole.

3.1.3 *cluster sampling, n*—sampling in which the sampling unit consists of a group of subunits, all of which are measured for sampled clusters.

3.1.4 *frame, n*—a list, compiled for sampling purposes, which designates all of the sampling units (items or groups) of a population or universe to be considered in a specific study.

3.1.5 *multi-stage sampling, n*—sampling in which the sample is selected by stages, the sampling units at each stage being selected from subunits of the larger sampling units chosen at the previous stage.

3.1.5.1 *Discussion*—The sampling unit for the first stage is the primary sampling unit. In multi-stage sampling, this unit is further subdivided. The second stage unit is called the secondary sampling unit. A third stage unit is called a tertiary sampling unit. The final sample is the set of all last stage sampling units that are obtained. As an example of sampling a lot of packaged product, the cartons of a lot could be the primary units, packages within the carton could be secondary units, and items within the packages could be the third-stage units.

3.1.6 *nested sampling, n*—same as *multi-stage sampling*.

3.1.7 *primary sampling unit, PSU, n*—the item, element, increment, segment or cluster selected at the first stage of the selection procedure from a population or universe.

3.1.8 *probability proportional to size sampling, PPS, n*—probability sampling in which the probabilities of selection of sampling units are proportional, or nearly proportional, to a quantity (the “size”) that is known for all sampling units.

3.1.9 *probability sample, n*—a sample in which the sampling units are selected by a chance process such that a

specified probability of selection can be attached to each possible sample that can be selected.

3.1.10 *proportional sampling, n*—a method of selection in stratified sampling such that the proportions of the sampling units (usually, PSUs) selected for the sample from each stratum are equal.

3.1.11 *quota sampling, n*—a method of selection similar to stratified sampling in which the numbers of units to be selected from each stratum is specified and the selection is done by trained enumerators but is not a probability sample.

3.1.12 *sampling fraction, f, n*—the ratio of the number of sampling units selected for the sample to the number of sampling units available.

3.1.13 *sampling unit, n*—an item, group of items, or segment of material that can be selected as part of a probability sampling plan.

3.1.13.1 *Discussion*—The full collection of sampling units listed on a frame serves to describe the sampled population of a probability sampling plan.

3.1.14 *sampling with replacement, n*—probability sampling in which a selected unit is replaced after any step in selection so that this sampling unit is available for selection again at the next step of selection, or at any other succeeding step of the sample selection procedure.

3.1.15 *sampling without replacement, n*—probability sampling in which a selected sampling unit is set aside and cannot be selected at a later step of selection.

3.1.15.1 *Discussion*—Most samplings, including simple random sampling and stratified random sampling, are conducted by sampling without replacement.

3.1.16 *simple random sample, n*—(without replacement) probability sample of n sampling units from a population of N units selected in such a way that each of the $\frac{N!}{n!(N-n)!}$ subsets of n units is equally probable – (with replacement) a probability sample of n sampling units from a population of N units selected in such a way that, in order of selection, each of the N^n ordered sequences of units from the population is equally probable.

3.1.17 *stratified sampling, n*—sampling in which the population to be sampled is first divided into mutually exclusive subsets or strata, and independent samples taken within each stratum.

3.1.18 *systematic sampling, n*—a sampling procedure in which evenly spaced sampling units are selected.

3.2 Definitions of Terms Specific to This Standard:

3.2.1 *address, n*—(sampling) a unique label or instructions attached to a sampling unit by which it can be located and measured.

3.2.2 *area segment, n*—(area sampling) final sampling unit for area sampling, the delimited area from which a characteristic can be measured.

3.2.3 *composite sample, n*—(bulk sampling) sample prepared by aggregating increments of sampled material.

3.2.4 *increment, n*—(bulk sampling) individual portion of material collected by a single operation of a sampling device.

3.3 Symbols:

N	= number of units in the population to be sampled.
n	= number of units in the sample.
Y_i	= quantity value for the i -th unit in the population.
y_i	= quantity observed for i -th sampling unit.
$\bar{y}$	= average quantity for the population.
$\bar{y}$	= average of the observations in the sample.
X_i	= value of an auxiliary variable for the i -th unit in the population.
x_i	= value of an auxiliary variable for the i -th sampling unit.
P	= population proportion of units having an attribute of interest.
p	= sample proportion.
f	= sampling fraction.
s	= sample standard deviation of the observations in the sample.
s^2	= sample variance of the observations in the sample.
$SE(\bar{y})$	= standard error of an estimated mean $\bar{y}$.

4. Significance and Use

4.1 This guide describes the principal types of sampling designs and provides formulas for estimating population means and standard errors of the estimates. Practice E105 provides principles for designing probability sampling plans in relation to the objectives of study, costs, and practical constraints. Practice E122 aids in specifying the required sample size. Practice E141 describes conditions to ensure validity of the results of sampling. Further description of the designs and formulas in this guide, and beyond it, can be found in textbooks (1-10).³

4.2 Sampling, both discrete and bulk, is a clerical and physical operation. It generally involves training enumerators and technicians to use maps, directories and stop watches so as to locate designated sampling units. Once a sampling unit is located at its address, discrete sampling and area sampling enumeration proceeds to a measurement. For bulk sampling, material is extracted into a composite.

4.3 A sampling plan consists of instructions telling how to list addresses and how to select the addresses to be measured or extracted. A frame is a listing of addresses each of which is indexed by a single integer or by an n -tuple (several integer) number. The sampled population consists of all addresses in the frame that can actually be selected and measured. It is sometimes different from a targeted population that the user would have preferred to be covered.

4.4 A selection scheme designates which indexes constitute the sample. If certified random numbers completely control the selection scheme the sample is called a probability sample. Certified random numbers are those generated either from a table (for example, Ref (11)) that has been tested for equal digit frequencies and for serial independence, from a computer

³ The boldface numbers in parentheses refer to a list of references at the end of this standard.