Chapter 11
Sampling

Introduction to sampling. Sample selection and how to take the sample. The various sample selection techniques: Random numbers, interval, stratified, and cluster. Haphazard and judgment sampling defined. Sample sizes. Statistical sampling theory simplified. Various sampling plans: sampling for attributes, stop-and-go sampling, discovery sampling. Variables sampling and the variables sampling techniques: mean-per unit, difference estimation, and ratio estimation. Formula for use in variables sampling. Dollar Unit Sampling (DUS), also called Monetary Unit Sampling. When judgment sampling is appropriate. Contrasting statistical and nonstatistical sampling. Evaluating sample results. Population proportions in known and unknown populations. Summary of sampling rules. Various selection techniques and sampling plans and when to use them.

Introduction

Sampling is the process of applying audit procedures to less than an entire population to draw conclusions about the totality. Sampling theory assumes that a demonstrated quality of a representative sample can be extrapolated to the population.

Sampling, in essence, is the process of learning about a great deal by looking at a little. At the same time, by sampling the auditor must accept the risk that the sample selected does not truly represent the population; that is, that the characteristics projected from the sample are not the same as those that would be found if the entire population — or a larger sample — were audited.

Sampling is not an end in itself; it is only a means to an end. The sample and the sample results are merely raw data — data that must be weighed and sifted. That data must be analyzed for materiality, reasons, causes, and actual or potential effect. The sample is but the first step to an informed audit opinion.

With the increasing use of information technology, the auditor must decide whether sampling is the most efficient and effective way to obtain evidence. With data warehouses and information retrieval approaches, it may be more efficient to perform computer-assisted tests on the whole population. Given the movement toward enterprise-wide software and other integrated software, an auditor may decide not to sample from a population but to audit 100 percent of the population.

Generalized audit software and other techniques may allow an auditor to more efficiently audit 100 percent of a population than to select a sample and audit the sample. This is especially the case where the entity has eliminated manual documents.

This chapter will explain how the auditor decides: (1) which sampling approach to use; (2) how many units (or sample items) to select; (3) how to select the units; and (4) how to evaluate the results in terms of the objectives of the audit procedures.

Sampling terminology often varies in textbooks and in certification examinations. Accordingly, the CIA examination questions at the end of this chapter were used as written, without attempting to establish consistent terminology. But the reader needs to be aware of the variations to be able to deal effectively with sampling terminology.

This book presents a survey of key sampling techniques and plans. For a more detailed discussion of specific sampling methods, see the supplementary reading list at the end of the chapter.

Sample Selection

When auditors select samples, they may take at least two paths. The first leads to the directed sample; the second leads to the random sample.

The directed or purposeful sample is used when auditors suspect serious errors or manipulation and want either to obtain evidence to support their suspicions or to find as many of the suspected items as they can. This process has nothing to do with statistical sampling. It is pure detective work. And the better a sleuth the auditor is the more useful his or her sample will be. But auditors may not draw conclusions about a population from a directed sample. Such a conclusion would be completely unwarranted, because the sample did not represent the population.

The random sample seeks to represent, as closely as possible, the population from which it was drawn. When auditors take a random sample, they are trying to take a picture, in miniature, of the great mass of records or data that make up the population from which the sample is selected. The larger the sample, the more closely it depicts the population. In audit argot, the sample is then termed "representative."

Statistical sampling permits internal auditors to measure sampling risk — that is the risk that a sample will not represent the population. To measure that risk statistically the sample selection must be random. Random selection means that every item in the population has an equal chance of being selected.

Nonstatistical sampling does not permit the auditor to measure the sampling risk objectively, since each population item did not have an equal chance of being selected. Nevertheless, nonstatistical sampling can be of value in such sampling plans as purposive (directed) or other forms of judgmental sampling.

It is possible, of course, for an auditor to select items at random without attempting to draw statistical inferences about the entire population. But by using random selection the auditor could reject any imputation of bias and, also, feel more comfortable that the sample selected is more likely to represent the values of the population.

Taking the Sample

There are certain rules for taking a representative sample. Here are three fundamental principles of selection that apply in any sampling procedure:

1. Know your population because audit conclusions may be based only on the sample taken from that population.

2. Define the sampling unit in terms of the audit objectives.

3. Let every item (sampling unit) in the population have an equal (or known) chance of being selected.

The mass of data, records, or documents from which the auditor selects a sample is variously referred to as population, universe, and field. They all mean the same thing. They also imply something that is central to good sampling: Know what you are testing. Clearly define the audit objective. Establish the population to be tested. Identify the sampling unit to be examined. The initial questions must be: "What are the objectives of my audit?" "What is the population I want to test?" "What are the sampling units I need to examine?" These are significant questions, having to do as much with good audit practice as with statistical sampling.

If the three principles above are violated, the tests are open to serious questions on technical grounds, and conclusions are without objective support. Consider, for example, the auditor who wishes to estimate the annual fuel consumption of the organization's automotive fleet. The fleet is made up of the following 750 pieces of automotive equipment:

50	12-wheel trucks
100	pickup trucks
100	forklift trucks
200	passenger cars
300	motor scooters

Let us assume in this example that the fleet has been well maintained and that there are no "gas guzzlers" within each group. Since the motor scooters were the most numerous, the auditor decided to select a sample of 50 motor scooters, examine fuel consumption for these 50, and project the finding to all 750 pieces of equipment. Stated this way, the results would be clearly ludicrous. Yet auditors fall into similar traps in their sampling every day. For example, in their audits they might:

- Select a sample of invoices for all those paid in July and use that test to form an opinion of all invoices paid during the year.
- Select a sample of travel vouchers of domestic travel only and project their findings to a population that includes foreign travel as well.
- Select a sample of purchase orders from a population that excludes all orders under $5,000 and express opinions on all purchase orders issued — from $1 on up.
- Sample inventory records in one tool crib and express an opinion on the records of tool cribs in all the many locations within the organization.

Each of these opinions and projections — unless properly qualified — is without support and just plain wrong. In the case of the motor vehicles, the auditor either should have sampled from all the various types of equipment or should have restricted the audit opinion solely to the scooters sampled.

Hence, auditors must always remember to define the population and the sampling unit in terms of the audit objectives. For example, the population of the 750 vehicles can have a different meaning, depending on the audit objective. If the objective were to determine whether all vehicles are maintained regularly — and let us assume that every item of equipment, be it scooter or truck, must be periodically maintained — the population assumes a different character than if the audit objective were to estimate fuel consumption. When the population or the sampling unit is improperly defined in the light of the audit objective, the result is bad sampling and bad auditing.

When the population and sampling unit are properly defined, the whole audit thrust and approach improve. A good technique is to plot the population *before* sampling to identify subpopulations or strata.

The principles on which scientific sampling is based operate only if the sample is selected at random. There are several ways in which random selection can be accomplished. Each method has its advantages and disadvantages. The three common methods are referred to as: (1) random number sampling; (2) interval (systematic) sampling; and (3) stratified random sampling. In addition, three nonrandom sampling techniques are discussed: cluster sampling, haphazard sampling, and judgment sampling.

Random Number Sampling

Random number sampling is generally considered the most likely to result in a representative sample. It makes use of computer algorithms or tables of digits that have been scientifically "randomized." The algorithms or tables provide substantially complete assurance that every item in a population has an equal chance of being selected. Many such tables have been compiled.

In general, the tables are easy to use if some simple rules are followed:

- Enter the tables by opening them at random and, with eyes averted, place a pencil point on the page. Start the number selection with the digit closest to the pencil point.
- Use as many digits in a line as there are digits in the reference numbers of the documents being selected. For example, if receiving memos being tested have a maximum of seven digits in their numbers, use seven adjoining numbers from the tables.
- Once a starting point has been selected, proceed through the tables in a predetermined order — down the columns or across the rows — without deviation, because deviation implies personal bias.
- If an applicable number does not appear, continue on to the next.

The numbers selected can then be arranged in numerical order for ease in locating the documents bearing those numbers, assuming the documents themselves are filed in numerical sequence.

Random number sampling can sometimes be difficult to use. Documents may not be numbered or may resist ready identification, or the identification numbers assigned may be quite long; thus it would be an almost impossible task to match random digits with the identification numbers.

The selection job is simplified, however, if the population of items is listed on a printout. Then, the random number tables could be used to select a page and a line. For example, assume a printout of 250 pages with 52 lines to a page. The auditor might want to use the first three digits of a column of random digits to identify a page and the last two digits to identify a line.

Random number software can also be used. Most such software gives the reader the option to place the selected items in numerical order. An auditor can also use the random number function available in most spreadsheets.

Interval Sampling

Some auditors use interval sampling — sometimes referred to as systematic sampling.

Interval sampling simply means selecting items at intervals. It is a relatively simple method but, in using it, the auditor must remember basic selection principles: Because the audit opinion may properly be based only on the population sampled, no items should be missing from the population. Also, because every item must have an equal chance of being selected, the first item in the selection process must be picked at random. Finally, because no pattern in the population should affect the selection, the auditor should make two or more passes through the population, each with a random start.

The concept of interval sampling is simple. As an example, suppose the auditor wants to draw a sample of 40 items from a population of 2,000 unnumbered documents. Divide the population of 2,000 by the sample size of 40. The sample interval is thus 50. Start the selection with a random number of say, 15, which is less than 50, and the sample items would be the fifteenth item, the sixty-fifth items, the one-hundred and fifteenth item, the one-hundred and sixth-fifth item, and so forth — every fiftieth item until 40 items are selected.

But the auditor must always be concerned with bias; for example, say every tenth item has a different characteristic from the others. This might happen if every tenth name on a payroll represented a supervisor. Then it would be necessary to make multiple starts — at random, of course — and hence multiple passes through the population.

To accomplish three passes through the population, take the following steps:

- Multiply the interval of 50 by 3 to get 150.

- From the random number tables select three random starting points, all less than 150. They are, say, 39, 66, and 91 — all under 150. The selection would then be:

First item	Second item	Third item	Etc.
39	189	339	(Last item + 150)
66	216	366	(Last item + 150)
91	241	391	(Last item + 150)

Because the total sample involves three series of selections, it will be slightly larger than a sample with only one series. Interval sampling is the simplest selection technique to use and, if it is used with care, can provide adequate assurance that the sample has been selected at random.

Stratified Random Sampling

In every population we should look for wide variations in size — the amount or characteristics of the items making up the population. When we see wide variations, we should consider stratification. Stratified sampling arranges the population so as to provide greater sampling efficiency. Properly used, stratified sampling will result in a smaller variance within a given sample than simple random sampling.

In stratified sampling we separate the population into two or more strata — in effect, separate populations — and then take samples from each. Auditors have always used the principles of stratification. Usually, they set aside the largest or more expensive or most significant items in a population for complete examination and then select a sample from the remainder.

It may sometimes be desirable to allocate the population to many strata so as to reduce the number of items needed to obtain a representative sample of the population. As we shall see, it is variability in the population, not its size, that causes sharp increases in the number of samples needed to give a good picture of the population.

Obviously, if the population were composed of identical items, a sample of only one of them would be representative of the whole. For example, if we wished to estimate fuel consumption for a fleet of 1,000 cars, and each automotive unit in the fleet was exactly the same as the others, all we would need to do is study the consumption of one unit and multiply by 1,000. We would have fairly good assurance that the projection would be a pretty reliable indicator of the true condition. If, however, the fleet were made up of tiny scooters, huge trailer trucks, and many different types of units in between, we would have to select samples from each type; in other words, we would have to stratify the population.

In real-life situations, the quality of the population usually varies widely. The more the quality or character of the individual items differs in the characteristic under study, the greater the number of items we must select to obtain a fair representation of the population. We are seeking to obtain a good picture of the population through our sample. The picture tends to become distorted by unusual items or wide swings of variability. Usually the only way to get that picture is through stratification.

Stratification, then, helps the auditor in two important ways: It controls distortion and it permits smaller sample sizes.

Just how to stratify, how many strata to develop, and what items to group together, call for audit judgment. It can be done with available software programs; however, in most cases sound audit judgment will suffice. Any reasonable stratification is better than none.

Once the population has been stratified, the sample items can be selected through random number sampling or interval sampling, depending on the circumstances.

Cluster Sampling

Cluster sampling, also called block sampling, acknowledges the shortness of life. Sometimes documents or records to be sampled are so scattered or dispersed that it is too time consuming and costly to use simple random number sampling. In such cases, auditors can use the technique of cluster sampling — even though it is less efficient than random number sampling or interval sampling in typical audit situations.

Cluster sampling is what the name implies. Clusters of items are selected at random, and then the clusters either are examined in their entirety or are themselves sampled. The latter method is referred to as multistage sampling. So long as each selection is at random — first the clusters and then, if necessary, the items within the clusters — no rules are violated; each item has been afforded an equal chance of being selected.

Clusters may be natural — that is, all the documents in a file cabinet drawer or in a bundle of records. Or the clusters may be artificial; the auditor may decide that each half-inch group of cards represents a cluster at, say, 10-inch intervals. Cluster sampling can be used to select:

- A sample of tool cribs and then a sample of tools in each crib.
- A sample of stockrooms and a sample of the inventory records for each of the rooms.
- A sample of file drawers and all or some of the documents filed in those drawers.
- A sample of months, weeks, or days, and a sample of the documents processed during the sample periods.

Haphazard Sampling

Haphazard sampling involves selecting items without being concerned about bias toward any feature of the sampling unit. For example, in the haphazard sampling of invoices, the auditor would take whatever invoices are readily at hand and disregard in the selection such information as invoice number, customer, location of sale, dollar amount of sale, number of items in the sale, or any other feature. Haphazard sampling takes the easy approach rather than a reasoned one.

Judgment Sampling

Judgment sampling involves selecting sample items based on the auditor's personal reasoning or suspicions. For example, an auditor may examine only those invoices processed by accounts payable clerks newly assigned to their jobs. Or the auditor might concentrate on invoices for products or services delivered directly to using departments without clearing through the normal receiving process. Further discussion of judgment sampling will be found later in this chapter.

Sample Sizes

Sample sizes can be determined judgmentally or statistically. The decision depends on the audit objective. In many audit situations, a large audit sample or a statistically determined sample size is unnecessary. Often, after a preliminary survey, auditors may be so impressed with the quality of a control system and of the management of the activity that they will be content with a review of handpicked items to assure themselves that the system is actually in operation.

In such circumstances, how many items are examined? If the system is being used to process three separate types of transactions, one sample of each of the three types is walked through the system, touching each of the control points. It can then be declared that the system does, indeed, have the purported control points and that they are, in fact, operating. Since the system satisfied us that aberrant items would be detected and corrected automatically, further sampling may be wasteful.

Suppose, however, that we want to feel reasonably sure in our minds that the system is working with substantial effectiveness. What is the smallest sample we can take to give us that assurance? That is hard to say without knowing the system and the quality of the population; but we should not place undue reliance on a sample of under 30 items. Only at 30 may the sample begin to adopt the characteristics of the population. In many situations, a statistical sample of 30 or 40 items will give sufficient assurance that the system is working with reasonable effectiveness. Auditors should be aware, however, that they are counting on the system to detect errors rather than on their small sample.

What if auditors want to be able to objectively measure the reliability of their sample results? Now they are making a quantum jump. Now they must be willing to step into statistical sampling. With an open mind and a few basic concepts, they will soon find that they can grasp this important subject.

Statistical Sampling Theory

Auditors must understand that when they deal with sampling, they are seeking a reliable estimate, not an exact answer. For instance, let us say we examine 100 items out of a population of 1,000 — a sample of one tenth of the population. Let us assume that we have found five errors. May we then multiply the number of errors by 10 and say with certainty that the population contains 50 errors? No! We may say, if the selection was at random, that we have a certain degree of confidence that our estimate, our projection, comes within a certain range or tolerance — that is, plus or minus some determinable percentage.

This brings us to two concepts that are pivotal to an understanding of scientific sampling: confidence level and precision.

Confidence level, also called reliability level, is the degree to which we are justified in believing that the estimate based on a sample drawn at random will fall within a specified range. A confidence level is usually expressed as a percentage. For instance, a confidence level of 95 percent means that there are 95 chances out of 100 that the sample results will not vary from the true characteristics of the whole population by more than a certain specified amount; there are 5 chances out of 100 that they will.

The confidence level for a sample can never be 100 percent. For that degree of confidence, the auditor will have to examine the entire population.

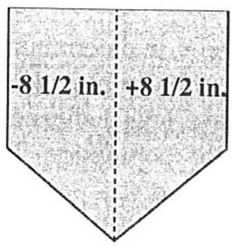
Precision, also called confidence interval, is the range within which the estimate of the population characteristic will fall at the stipulated confidence level. Precision, being a range or tolerance, is usually expressed as a plus-or-minus percentage, such as ±2 percent, or as an amount, such as ±$5,000.

Thus the estimate obtained from a sample may permit us to say that we are 95 percent confident that the value of a population is:

X dollars ± Y dollars.

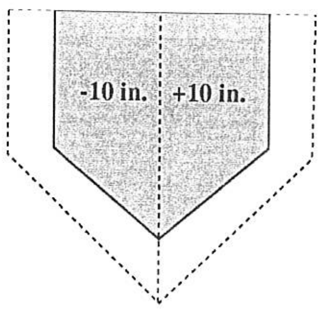
Confidence level and precision are inescapably related. Each has an effect on the other. The meaning of sample results cannot be understood without understanding that relationship. A simple example will go much further in showing this relationship than any definitions.

Assume that a baseball pitcher is given 100 balls and is asked to throw them over the center of home plate. Home plate is 17 inches wide and looks like this:



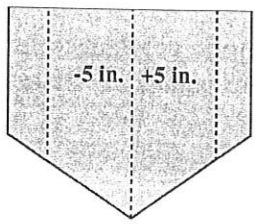
Let us assume that the pitcher can regularly get 95 out of the 100 balls over the plate. It could then be said that he or she has a 95 percent chance (confidence level) of getting a ball over the centerline ±8 1/2 inches (precision).

But then let us say that we widen the plate to 20 inches, like this:



Now, with a broader range, the confidence level improves. Most likely the pitcher could get 98 or 99 balls over the center of the plate ±10 inches.

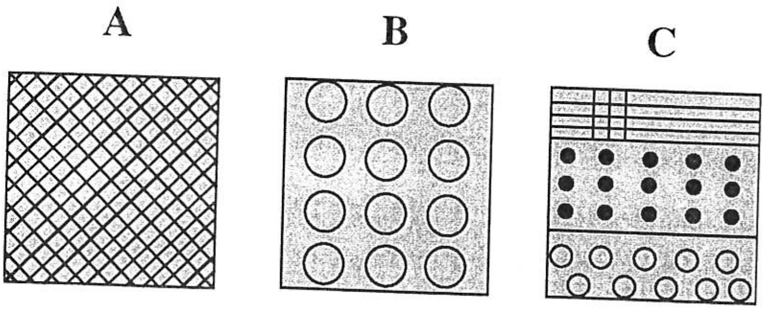
Pursuing the analogy further, let us reduce the size of the plate to 10 inches:



Now perhaps the pitcher can get only 80 balls over the center of the plate ±5 inches. The analogy is rough, but it points out the relationship between confidence level and precision. A change in one must change the other.

Another concept that must be understood in determining sample size is variability. Variability determines how large the sample must be to provide the acceptable level of confidence and range of precision. Many auditors have long believed that the size of the sample must bear a direct relationship to the size of the population. In earlier days, the percentage usually employed was 10 percent. Under that theory, a population of 100 would call for a sample of 10, and a population of one million would call for a sample of 100,000. In truth 10 percent of 100 is too small and 10 percent of a million is far too large.

Another simple example will bring the picture into sharper focus. Assume three lengths of cloth, each with a different design. The owner wishes to take a swatch from each so as to match them at the store and purchase more cloth of the same designs. This is how the lengths of cloth look:



It becomes immediately apparent that although all three lengths are the same size, different sample sizes will have to be taken from each to obtain a good representation of the total design — in effect, the population. From A, a very small sample will be representative. From B, a much larger sample would be necessary to represent the total design. From C, unless a sample is taken from each of the four design patterns — in effect, stratification — a very large sample would be needed.

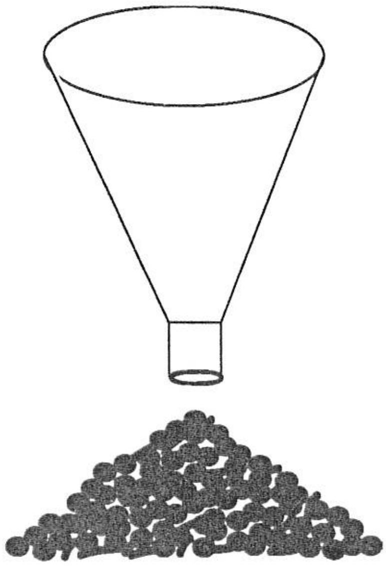
So sample size depends on four factors: confidence level, precision, population variability, and population size. The first two are under our control. But deciding how much reliability is needed from the sample calls for sound judgment based on the audit objectives and the nature of the associated system of internal control.

Although we may sometimes reduce variability by stratifying the population, variability is outside our control once the population to be sampled has been defined. It is part of the nature and character of the population we must deal with. That is why it is so important to "know your population."

Once the confidence level and precision have been decided, and once the variability of the population and its size have been measured, we are in a good position to determine how large a sample we will need to give us sample results which are sufficiently reliable for our purpose. In other words, we will be able to predict how close to the true population values (determined precisely only by examining the entire population) are the values of our sample.

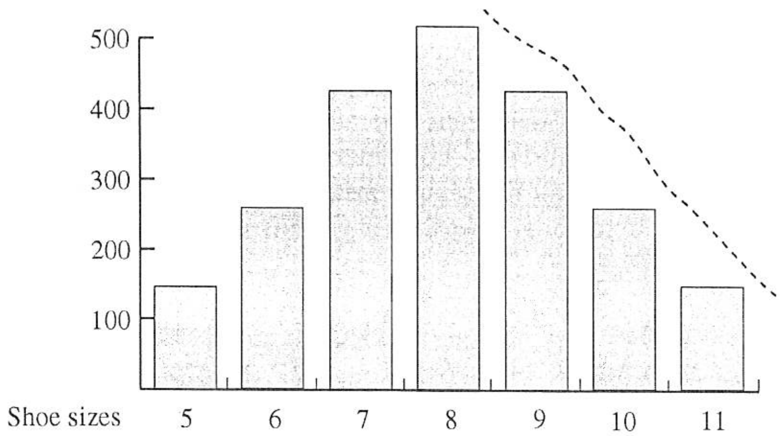
This ability to predict is based on principles that have been developed mathematically. The principle, referred to as the Central Limit Theorem, roughly states that the measurements of the values of many similar objects — when arrayed according to value or size — tend to take the shape of a bell, also known as the normal curve. And if one were to select from any population, at random, any infinite number of samples of the same size of about 30 or more units, the frequency distribution of the means (averages) of all those samples would inevitably take the shape of a bell-shaped curve — no matter how the values of the population were distributed. Further, the mean of all the sample means would be the same as the population mean. This latter concept is important because it permits a prediction of population values based on sample values.

To visualize — so as to show how natural that shape is — assume a huge funnel, closed at the bottom and filled with gravel, suspended over a flat surface. As the gravel is released from the funnel, it will invariably assume a shape somewhat like:



Viewed in silhouette, the pile of sand seems to have a bell-shaped curve. This shape seems natural to the viewer. Any other shape under the same circumstances might seem unnatural.

That is how the measurements of a great number of objects could be pictured. Assume the measurement of the shoe size of 2,000 men selected at random, plotted on a graph. The results could be:



Again, we see the frequency of the measurements tending to take the form of a bell-shaped curve. The shape of the curve will depict the variability of the population. A high, narrow curve will illustrate little variability; a flat, wide curve will illustrate great variability.

Variability within a population can be measured. Just as coal can be weighed by the ton, speed can be measured by miles per hour, length can be measured by yards, and time can be measured by hours, so can variability be measured by what is known as the standard deviation. This formidable term means simply, "the measure of the variability of a particular population or of a sample from that population."

Let us break it down into its components. The standard deviation, technically, is the square root of the average of the squared deviations from the mean. It is portrayed in the following formula for a determination from a statistical sample:

$$s = \sqrt{\frac{\Sigma(x-x)^2}{n-1}}$$

where:

s	=	standard deviation
Σ	=	the sum of
x	=	each observation — the characteristic or value of each sample item
x	=	the average (arithmetical mean} of the sample item values
n	=	the size of the sample

To determine the standard deviation:

- Obtain the mean (average) of the sample items
- Subtract the mean from each item
- Square the results and sum them
- Divide that sum by the number of sample items — minus 1
- Extract that square root

The result is the standard deviation of the sample or the measure of variability of the sample — hence, an estimate of the variability of the population.

Let us determine the standard deviation, through this method, of two groups of numbers. Each has the same mean, but different variability. The resulting standard deviations will show how this measure portrays variability, whereas the mean or average does not:

A	Sample Values		Mean†		Difference	Difference Squared
	17	-	20	=	-3	9
	20	-	20	=	0	0
	23	-	20	=	+3	9
						18

$$s = \sqrt{\frac{18}{2}}$$

$$s = \sqrt{9}$$

$$s = 3$$

*Sample size of 3 minus 1 = 2.

B

Sample Values		Mean†		Difference	Difference Squared
11	-	20	=	-9	81
20	-	20	=	0	0
29	-	20	=	+9	81
					162

$$s = \sqrt{\frac{162}{2}}$$

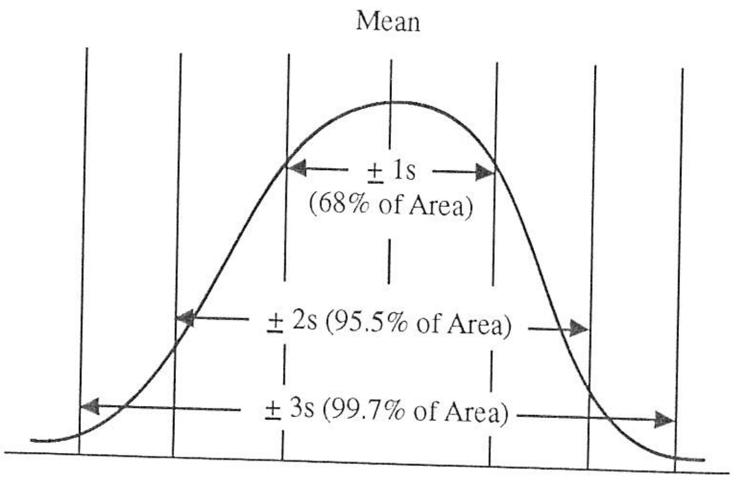
$$s = \sqrt{81}$$

$$s = 9$$

*Sample size of 3 minus 1 = 2.

Although both groups of numbers have a common mean, the standard deviation of the one with the greater variability is three times that with the smaller variability.

The relationship between the bell curve, or normal distribution, and the standard deviation is an interesting one. It has been determined that in any normal distribution, the mean of the distribution, plus or minus one standard deviation, includes about 68 percent of the area under the normal curve; the mean plus or minus two deviations includes 95.5 percent of the area; and the mean plus or minus three standard deviations includes about 99.7 percent of the area. The relationships between standard deviations and the curve can be shown as:



Unlike fixed units of measurement, such as an inch or a minute, the standard deviation will be different for each sample or each population because the standard deviation is the measure of variability of individual samples or populations. But regardless of the nature of the population, if we select at random a large number of samples of the same size, the distribution of the means of all those samples approximates a normal curve, and the average of the sample averages equals the average of the population.

Now let us see how this helps us in a sampling problem. Based on the principles of the normal curve, we could say that any item selected at random would fall — 68 percent of the time — within the range measured by the sample mean, plus or minus one standard deviation. Let us put it another way: Assuming a sample of invoices, "normally distributed," with an arithmetic mean of $100 and a standard deviation of $10, we could say that 68 percent of the sample units will fall within the value of $100 plus or minus one standard deviation of $10 — from $90 to $110 at a confidence level of 68 percent.

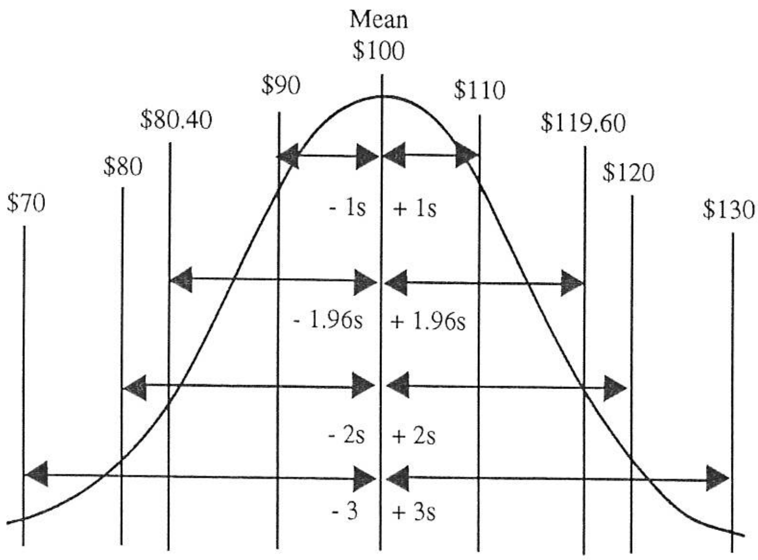
If we wished to increase our confidence level to 95.5 percent, we must now be satisfied with a wider range — plus or minus two standard deviations. Thus:

$80 to $120 at a confidence level of 95.5 percent.

If we wished to increase our confidence level still further to 99.7 percent, we would have to be satisfied with a still wider range — plus or minus three standard deviations. Thus:

$70 to $130 at a confidence level of 99.7 percent.

Or if we wished a confidence level of 95 percent — plus or minus 1.96 standard deviations — our range would be $80.40 to $119.60. On the normal curve, the results just enumerated could be pictured as:



Recapitulating, then, the sample size is largely dependent on the confidence level and precision stipulated for the sample results and the variability found in the population.

One other factor affects sample size: the audit objective — what the auditors seek to determine by their tests. The objective may call for any one of several plans:

- Sampling for attributes. Determining "how many."
- Stop-and-go sampling. An economical sampling approach.
- Discovery sampling. Seeking to find a single suspected item.
- Variable sampling. Determining "how much."
- Dollar unit sampling. A combination of attribute and variable sampling.
- Judgment. When sample results are not to be attributed to the entire population.

Each of these plans will be discussed in the following pages.

Sampling for Attributes

Sampling for attributes calls for yes-or-no, right-or-wrong answers. It is usually applied to testing systems of internal control. It is concerned with estimating the number of errors or other characteristics in a population. It can provide an estimate of the number of engineering drawings received late by production people, but it will not give an estimate of how late — that is a function of variables sampling. It can provide an estimate of the number of purchase orders issued to sole sources, but it will not give an estimate of their value — that too is the function of variables sampling.

Determining sample sizes is relatively easy when tables are used, and the method of using tables is set forth in the sampling literature. Briefly stated, the auditor first determines the population size, the desired confidence level, the desired precision, and the expected error rate.

The principle of the expected error rate was touched on in the discussion of variability — the more variable the items in a population, the larger the sample size needed. Obviously, there is no great variability in a situation where the characteristic is either "yes" or "no," — that is, binary. If all the characteristics were "yes," it would take a test of only one to then predict that all the other items in the population were also "yes." But the more "no's" — up to 50 percent of the items — that are sprinkled throughout the population, the larger the sample needed to obtain a good representation of that population.

Estimating the error rate calls for judgment, but methods are available to give reasonably sound basis for an estimate:

- Examine a pilot of about 25 or 30 items.
- Review prior working papers for past experience.
- Discuss the estimated number of errors being encountered with knowledgeable people.
- Estimate the percentage of error it would take to automatically alert management that something is wrong.

Once the necessary decisions are made, we can proceed to the tables without difficulty, finding the one that shows the population size, confidence level, precision, and expected error rate in which we are interested.

But the available tables may not always fit the auditor's specifications. Doing so would take an enormous number of tables, and the auditors should never change their specifications merely to fit the tables at hand. That would be like the tail wagging the dog.

A simple formula can solve the dilemma. It adjusts to any confidence level, precision, and error rate. It also adjusts to any population by applying a simple supplemental formula. It uses the standard deviation factors — here called Z factors (often referred to as t factors). The values include:

Confidence Level	Z Factors
99	2.5758
95.5	2.0000
95	1.9600
90	1.6449
85	1.4395
80	1.2816

The formula has two parts, and both are simple. Neither requires the extraction of square roots. The first formula is used to obtain the sample for an infinite population. The second adjusts the formula for the population being considered.

The first formula (the normal approximation of a binomial distribution) is:

$$n_{(e)} = \frac{Z^2(p)(1-p)}{A^2}$$

where:

$n_{(e)}$ = First estimate of sample size
Z = Standard deviation factor
p = Occurrence rate
A = Desired precision

The second formula uses the first estimate of sample size and adjusts it to fit the population:

$$n_{(f)} = \frac{n_{(e)}}{1 + (n_{(e)}/N)}$$

where:

$n_{(f)}$ = Final sample size
$n_{(e)}$ = First estimate of sample size
N = Population

Using these formulas, here is the computation for the sample size where the population (N) is 1,000, the desired precision (A) is ±2 percent, the desired confidence level (Z) is 95 percent, and the estimated error rate (p) is not to exceed five percent:

1. $$n_{(e)} = \frac{1.96^2 \text{ x } .05 \text{ x } (1-.05)}{.02^2}$$

$$= \frac{3.8416 \text{ x } .05 \text{ x } (.95)}{.0004}$$

$$= \frac{.182476}{.0004}$$

$$= 456$$

2. $$n_{(f)} = \frac{n_{(e)}}{1 + (n_{(e)}/N)}$$

$$= \frac{456}{1 + (456/1000)}$$

$$= \frac{456}{1.456}$$

$$= 313$$

The second formula, representing a "finite population correction factor," can generally be omitted if the first sample estimate is less than five percent of the universe size. The second formula, in such cases, has little effect on the sample size — merely making it slightly smaller.

If we examine the sample of 313 items and find that there is indeed an error rate of five percent, then we can declare that we are 95 percent sure that our population of 1000 contains 950 error-free items, plus or minus two percent of 1000, or plus or minus 20 items. In other words, the number of satisfactory items in the population can be estimated to be anywhere from 930 to 970 — at a 95 percent confidence level.

If it turned out that the error rate in our sample is actually somewhat higher, say six or seven percent, we would recompute the formula, using the new error rate, and determine how many more items we would have to examine to obtain an estimate with the required confidence level and precision. If the actual error rate is much higher, say 10 percent or more, the sampling could well be finished — since 10 percent of 313 items would be sufficient to indicate a real problem that needs correcting.

Stop-and-Go Sampling

Stop-and-go sampling was devised to permit audit decisions, with appropriate reliability, that are based on relatively small samples. It is used in testing for attributes. It applies, by and large, to fairly "clean" populations — those in which auditors want to do as little sampling as possible. Based on their knowledge of the system, the auditors may conclude that the population is relatively error-free; but they want statistical support without extensive testing.

So if they can examine a small sample and find few or no errors, they will have a measurable assurance that they can discontinue their tests and accept the reasonable accuracy of the population.

Stop-and-go sampling tables and software are available. In a stop-and-go sampling table as illustrated below, the first column shows the size of the sample. The second column shows the number of errors brought to light by the auditor's analysis of the sample. The other columns, headed by various possible maximum error rates, show different levels of probability (the number of times out of 100) that the true error rates in the population will be less than the indicated maximum error rates in the headings.

A simple example can show how stop-and-go sampling works. Let us say that we are interested in whether receiving memorandums bear evidence that the materials received have been inspected. We shall assume that all items must be inspected. The absence of an inspector's stamp indicates no inspection.

Employees appear to be well trained. Supervisors watch the operations carefully. The manager periodically checks completed receiving memos to see that the rules are being followed. Clearly, under such a system, extensive testing would be wasteful. We take a sample of 50 items out of a population of 20,000. We stipulate that we would be satisfied if we have adequate assurance that the population has an error rate no higher than five percent.

Probability that error rate in universe of size over 2,000 is less than:

Sample Size	# of Errors Found	1%	2%	3%	4%	5%	6%	...
50	0	39.50	63.58	78.19	87.01	92.31	95.47	...
	1	8.94	26.42	44.47	59.95	72.06	81.00	...
	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...
100	0	63.40	86.74	95.25	98.31	99.41	99.80	...
	1	26.42	59.67	80.54	91.28	96.29	98.48	...
...	...	...	...	...	...	...	...	...

Let us say that we examine the 50 items and find no errors. The condition is therefore as shown in the first line of the table: a 92.31 percent assurance (probability) that the population contains no more than a five percent error rate. If we are satisfied with that assurance, we could discontinue our tests.

But let us say that we found one error in our sample of 50. Now, the condition is as in the second line: 72.06 percent assurance that the population contains no more than a five percent error rate. If this does not satisfy us, we might take a sample of 50 more receiving memos — a total of 100. Assume that we find no more errors. The condition is then as in the second line for a sample size of 100: 96.29 percent assurance. If we consider that to be adequate — taking into account the excellent system of internal control — we may discontinue our test. Otherwise, we will take additional samples.

If errors keep showing up in a small sample size, the auditor should not continue with stop-and-go sampling for sample sizes. At that point, the auditor will want to obtain an estimate of the error rate in the population within a plus-or-minus range of precision at an appropriate confidence level.

Discovery Sampling

Discovery sampling is used when the auditor is examining populations where the existence of fraud or gross error is suspected. Such populations might include fictitious employees on the payroll, duplicate payments, unauthorized shipments of goods, or nonexistent collateral for loans.

We would not try to express an opinion on the population as a whole. We are trying, through sampling, to find at least one item with a particular characteristic — assuming a stipulated number of those items in the population. The stipulation is significant. The population would require examination, item by item, until the one such item was found. There would be no other choice if the item of interest were a single unique unit. But if we are willing to specify some limited, assumed number of items, we may use discovery sampling to obtain a measurable assurance that we will find at least one of that number, if the actual quantity in the population is equal to or greater than the assumed quantity.

A sample of part of a table follows.

Probability, in percent, of finding at least one error if total number of errors in universe is as indicated:

Total errors in universe size of 10,000

Sample Size	1	2	3	4	...	50	75	...
5	0.1	0.1	0.1	0.2	...	2.5	3.7	...
10	0.1	0.2	0.3	0.4	...	4.9	7.3	...
...	...	...	...	...	...	...	...	...
600	6.0	11.6	16.9	21.9	...	95.5	99.1	...
...	...	...	...	...	...	...	...	...

Here is how the tables are used: We are examining the organization records on conflict of interest. Every employee is required to complete a record designed to disclose any such conflict. There are 10,000 employees on the rolls and we want to be 95 percent certain that we would locate at least one instance of impropriety — no record, wrong record, or an uninvestigated record of potential conflict.

By consulting the previous table for a population size of 10,000, we see that if we were to stipulate 50 errors in the population, we would have to examine a sample of 600 items to be 95.5 percent sure that our sample would include one of the erroneous items. If our sample contained none of the errors, we would be 95.5 percent sure that the population included fewer than 50 erroneous items.

How many erroneous items to stipulate is a matter of judgment, taking into account the seriousness of the errors under consideration. The only alternative to discovery sampling is the examination of each item until an example has been found or the entire population has been examined. The auditor will have to evaluate the impact of the undiscovered errors in the population.

Variables Sampling

This form of sampling is sometimes called dollar estimation, since it usually deals with dollar values. It can also be used for any other kind of values, like time periods or weights.

Variables sampling can be used to obtain estimates based on a sample of the value of inventories, the value of disallowances of travel vouchers, the value of aged accounts receivable, and the like. Computing sample sizes and sample results is simplified by the use of tables. Such tables can be found in the sampling literature previously mentioned. Here we shall make use of an alternative set of tables that allows a certain degree of flexibility often desirable in computing sample sizes.

To establish suitable sample sizes, we must determine, as we did for attributes sampling, the population size, the desired confidence level, and the desired precision. Instead of an expected error rate, however, we must determine the standard deviation.

As we pointed out in attributes sampling, the expected error rate is a measure of variability that is needed to work out the formulas for sample sizes. In variables sampling, the standard deviation does the same thing.

Accordingly, for the first step in establishing the sample size needed to provide the desired degree of reliability, we must estimate the standard deviation of the population. We do that by determining the standard deviation of the sample. Just as we may take a pilot sample in attributes sampling to estimate the error rate, so we must take a pilot sample in variables sampling to determine the standard deviation.

The sample for that purpose should not be fewer than 100 units. Fewer than that number may not be representative of the population and, therefore, may not give a correct reading. The sample must be drawn at random, and even though it is termed a pilot sample, all the items selected can be used with the additional sample items that may be required to achieve desired sample reliability.

In dealing with so many numbers, it is best to use a simplified computation in determining the standard deviation. This computation does not require subtracting the mean from the value of each sample item. The results are the same as those obtained from the formula described earlier in this chapter. The simpler formula (easier to compute, even though it looks more complicated) is:

$$s = \sqrt{\frac{\Sigma(x^2) - (\Sigma x)^2/n}{n - 1}}$$

where:

s = Standard deviation of the sample
Σ = Sum of
x = Value of each sample item
n = Sample size

Using the figures in example 2 shown previously in determining the standard deviation:

	x = Sample Values	$\Sigma(x^2)$ = Sum of Squared Sample Values	(Σx) = Sum of Sample Values
	11	121	11
	20	400	20
	29	841	29
Totals		**1362**	**60**

and substituting these values in the formula, we have:

$$s = \sqrt{\frac{1362 - \frac{60^2}{3}}{3 - 1}}$$

$$= \sqrt{\frac{1362 - 1200}{2}}$$

$$= \sqrt{81}$$

$$= 9$$

Variables Sampling Techniques

There are three classical variables sampling plans used to calculate the population estimate: mean-per-unit, difference estimation, and ratio estimation. These three techniques will be demonstrated through the use of the following example:

The accounts receivable balance of $5,000,000 is composed of 4,000 customer accounts. The auditor selects a sample of 200 accounts with a book value of $250,000. After applying audit procedures, the audit value of the sample is determined to be $265,000.

1. *Mean-per-unit.* By this method, the auditor calculates the average audit value ($265,000 ÷ 200 = $1,325) and multiplies this mean-per-unit value by the population units ($1,325 x 4,000 = $5,300,000) to estimate the correct accounts receivable balance. This method is also referred to as simple extension.

2. *Difference Estimation.* Difference estimation may be appropriate when the following conditions are present:

   - There are sufficient errors in the population to generate a reliable sample estimate and the differences are not proportional to recorded book value. If the number of differences between the audit value and book value is small, a relatively large sample would be required to provide a representative difference between audit and book values.

   - The following information is available:
     BV = Book value of the population
     N = Number of units in the population
     UBV = Book value of each unit
     UAV = Audit value of each sample unit
     S = Size of sample

The following steps are followed in difference estimation:

- Determine the average "difference" between audit value and book value for the sample (UAV – UBV) ÷ S = average difference [($265,000 – $250,000) ÷ 200 = $75].
- Multiply the average difference by the number of population units to obtain a total difference estimate ($75 x 4,000 = $300,000).
- The difference is combined with the population book value to obtain an estimate of the population value ($300,000 + $5,000,000 = $5,300,000).

3. *Ratio Estimation.* Ratio estimation is similar to difference estimation and is appropriate when the differences are approximately proportional to the book values. The following steps are used to generate a ratio estimate of the population value.

- Determine the sample audit value ($265,000).
- Determine the sample book value ($250,000).
- Compute the "ratio" of the audit value to the book value ($265,000 ÷ $250,000 = 1.06).
- Multiply the ratio by the book value of the population (1.06 x $5,000,000 = $5,300,000).

These simplified examples illustrate the differences among the three methods. The following section provides a formula for the use of the more common method — mean-per-unit.

Variables Sampling— Formula

As in attributes sampling, fairly simple formulas can be used to determine sample sizes and precision (sampling error). Using the same example shown for the tables:

Population size	5,000
Population value	$500,000
Standard deviation of 200 items selected at random	$40
Desired confidence level	90%
Desired precision per unit of population	±$4

Here is the first formula for mean-per-unit:

$$n_{(e)} = \left(\frac{Zs}{A}\right)^2$$

where:

$n_{(e)}$ = First estimate of sample size
Z = Standard deviation factor
s = Standard deviation of the sample
A = Precision (sampling error)

The second is the same adjustment formula shown previously and uses the first estimate of sample size and adjusts it to fit the population:

$$n_{(f)} = \frac{n_{(e)}}{1 + (n_{(e)}/N)}$$

Using these formulas, here is the computation for the sample size:

1.

$$n_{(e)} = \left(\frac{1.645 \text{ x } 40}{4}\right)^2$$

$$= \left(\frac{65.8}{4}\right)^2$$

$$= 16.45^2$$

$$= 271$$

2.

$$n_{(f)} = \frac{271}{1 + (271/5{,}000)}$$

$$= \frac{271}{1.0542}$$

$$= 257$$

Also, as in attributes sampling, the second formula is generally not needed if the first sample is less than five percent of the population.

As a result of the computations, the internal auditor will select at random another 57 items to bring the total to 257. Let us assume that the examination of the sample of 257 items shows the standard deviation remains at $40. Let us also assume that the value of the sample items is as follows:

Book value	$27,000
Value determined by physical verification	$23,130

The average value of the inventory items is $90 ($23,130 ÷ 257). The estimated actual inventory would be $450,000 ($90 x 5,000). The next questions are: How reliable is that projection? What is the precision (sampling error) at a 90 percent confidence level? What is the range within which the estimated actual inventory value can be projected at the desired confidence level? A formula can provide that answer:

$$A = Z\frac{S}{\sqrt{n}}\left(\sqrt{1-\frac{n}{N}}\right)$$

where:

A	=	Precision
Z	=	Standard deviation factor
s	=	Standard deviation of the sample
n	=	Sample size
N	=	Population size

Substituting the values determined by the audit and by the desired confidence level, the value of the sampling error is reached as:

$$A = 1.645 \left(\frac{40}{\sqrt{257}}\right) \sqrt{1 - \frac{257}{5000}}$$

$$= 1.645 \left(\frac{40}{16.03}\right) \sqrt{1 - .0514}$$

$$= 1.645 \times 2.4953 \times .97396$$

$$= \$4.00$$

The precision of ±$4 for each unit computes to ±$20,000 for 5,000 units. The projected value of the entire inventory is therefore $450,000 plus or minus $20,000. In other words, at a confidence level of 90 percent, the actual value of the inventory would lie within a range of $430,000 to $470,000.

Dollar-unit Sampling (Monetary Unit Sampling)

The internal auditor has many statistical sampling models at his or her disposal. The attribute and variables sampling are often referred to as the classical statistical sampling models.

A relatively new form of dollar estimation has been gaining increased acceptance in recent years. It goes by many names and lacks common terminology. It has been referred to as probability proportional to size (PPS) — in essence, a sampling selection procedure. It is also referred to as dollar-unit sampling (DUS), which is a unique statistical approach based on a PPS selection process. Other terms for this method include Monetary Unit Sampling (MUS), cumulative monetary amount (CMA) sampling, combined attributes variables (CAV) sampling, monetary-unit sampling (MUS), and sampling proportionate to size (SPS). The differences and variations of the separate but closely related topics are beyond the scope of this text. Owing to their similarities, however, we will refer to the entire concept as dollar-unit sampling or DUS.

DUS is actually a hybrid that combines attribute and variable sampling methodology, as expressed in the term CAV. DUS relies on an attribute sampling approach (Poisson distribution) to express a conclusion in dollar amounts (variables) rather than as a deviation rate. DUS provides an alternative to stratification and combines the advantages of both attribute and variable sampling. It can be used for both compliance and substantive testing.

Classical sampling approaches designate physical items (vouchers, invoices, checks, etc.) as the sampling unit. Each of the physical items is likely to have a different value; hence, they are variable. Any selection of such "audit units" may include large amounts and small amounts. It is quite possible for significant items to be overlooked in the sample selection process unless considerable stratification is employed. DUS avoids this problem by defining units that do not vary. Each one has the same value because each one is one dollar; thus, the sample is composed of random dollars, not random items.

DUS avoids the difficulty of determining standard deviations of the sample by removing the variability of the sampling unit. DUS does not look at an account receivable population as totaling $5,000,000, for example, and comprising 2,000 customer accounts that range, say, from $1 to $100,000. Instead, it looks at the population as comprising 5,000,000 dollar bills, stretched out — dollar bill after dollar bill — for almost 500 miles.

In selecting a random sample of 60 dollars (see example that follows) from the population, each dollar must have an equal chance of being selected for inclusion in the sample. Although dollars are the units for sample selection, the auditor does not actually examine individual dollar bills. The auditor examines the item, account, transaction, or document associated with the dollar selected. The individual dollar selected is generally described as "hooking" the entire item with which it is associated. If one dollar is selected from an account receivable balance, it will not relinquish its companion dollars, say customer No. 5 (in the example that follows) with an account balance of $81,897, within which the dollar resides. The item snagged is termed the "logical sampling unit."

In applying DUS, the author is automatically stratifying the population for selection. The probability of any one item being selected for detailed verification is proportional to the size of the item — therefore the terms PPS and SPS. In the example given, the probability that customer account No. 9 ($10,000) is selected is 10 times greater than that of customer account No. 8 ($1,000). DUS provides an alternative to stratifying and is somewhat simpler. This characteristic, and its use where differences between audited and book amounts are rare, has popularized DUS in audit practice.

DUS has audit appeal, since large-value items are generally of greater concern than small-value items. DUS provides an estimate of the maximum amount of overstatement of a recorded amount with measurable levels of risk of making a decision error. It allows the auditor to make the following general conclusion:[1]

Based on the sample's evidence, I am x percent confident that the dollar amount of error in the account (related to the attribute) does not exceed $Z (where $Z depends on the sample's outcome).

DUS is appropriate where the audit objective is to detect overstatements and is not appropriate in testing understatements. Since basic audit philosophy is to audit assets for potential overstatement and liabilities for understatement, DUS is appropriate to test asset account balances. Frequent applications include the test of accounts receivable, investments, inventory, and fixed assets.

Before using the DUS model, the auditor should determine that the two assumptions incorporated in the model are valid for the population being tested:[2]

1. The error rate in the population should be small (less than 10 percent) and the population should contain 2,000 or more items. (The use of the Poisson probability distribution for evaluation of the sample requires this feature.)

2. The amount of error in any item of the population cannot be more than the reported book value of the item. That is, if the book value of a customer's balance is $100, the amount of error in the balance cannot exceed $100.

As noted, DUS is generally not appropriate in testing for understatement of liabilities since the more a balance is understated, the less is its chance of being included in the sample. Ironically, the presence of an understatement error limits the probability of its detection and adjustment. Also, unrecorded items obviously will not be selected by DUS. Zero and negative balances should be excluded from the population being tested because they have no probability of being selected for inclusion in the sample. These items are generally handled as a separate population for testing, using a classical selection process.

A few of the other advantages and disadvantages of DUS are:

Advantages

- It can be applied to a group of accounts, since the sampling units (dollars) are homogenous.
- It is an efficient model for establishing that a low error rate population is not materially misstated.
- It does not require consideration of the standard deviation to determine the sample size or to evaluate the results.
- It does not require the normal approximations required by variables sampling.
- It permits a statistically valid sample selection that includes more high-dollar balances for "dollar coverage" of the population being tested.

Disadvantages

- It requires that the population be cumulatively totaled to identify the random dollars selected. This disadvantage is generally offset with a software package.
- As the errors increase, it requires a larger sample size than that required when using classical statistics.
- It may overstate the allowance for sampling risk when errors are found and thereby cause the auditor to reject an acceptable book value.

Two methods are available for use in selecting sample items for DUS:

1. *Random DUS.* If this method were used in the following example, the auditor would select 60 random numbers from a random digit table or through the use of a random number generator. The next step would be to arrange the random numbers in ascending order and select the account balance whose cumulative amount included the number.

2. *Systematic DUS.* This is a more widely used method and is illustrated in the following example.

The more complex applications of DUS are beyond the scope of this text. A simplified step-by-step example is presented to provide a conceptual understanding that can serve as a foundation for more sophisticated applications.

Example: Ricky Roderick, a staff auditor of Petro, Inc., elects to use DUS to test for overstatement of the organization's accounts receivable book balance of $5,000,000. He has determined that the DUS assumptions are satisfied and that DUS is valid to achieve the audit objective. Roderick has decided that a five percent error in the book value is material. Five percent of $5,000,000 is $250,000. He also decides to use a 95 percent confidence level and expects no errors. Hence, the parameters are:

BV	(book value)	=	$5,000,000
N	(population units)	=	2,000
M	(maximum tolerable error)	=	$250,000
CL	(confidence level)	=	95%

Determine the reliability factor given a 95 percent confidence level and zero expected overstatement errors. An Error Factor (F) of 3.0 is available from a Poisson DUS Cumulative Evaluation Table (Table of Error Factors).

DUS Reliability Factors

Reliability Required	Reliability Factors
99%	4.605
95%	2.996 (rounded to 3.0)
90%	2.300

The sample size (n) is calculated as:

$$n = \frac{BV \times F}{M}$$

$$n = \frac{\$5{,}000{,}000 \times 3.0}{\$250{,}000}$$

$$n = 60$$

Alternative calculation:

$$n = \frac{\text{Error Factor}}{\text{Materiality Percentage}}$$

$$n = \frac{3}{5\%}$$

$$n = 60$$

Calculate the sample interval (I):

$$I = \frac{BV}{n}$$

$$I = \frac{\$5{,}000{,}000}{60}$$

$$I = \$83{,}333$$

Select the sample. The first dollar is selected through the use of a random number (say 1,000). The second number is determined by adding the random start to the skip interval. The third number is determined by adding the random start to two times the skip interval, and so on through the entire population.

Sample Units	Calculation	Unit Selected
#1	Random start	1,000
#2	1,000 + $83,333 x 1 (first skip interval)	84,333
#3	1,000 + $83,333 x 2 (second skip interval)	167,666
#60	1,000 + $83,333 x-59 (last skip interval)	4,916,647

The auditor must accumulate the cumulative dollar total of items in the population (therefore the term CMA). The worksheet for this DUS example is shown in Exhibit 11-1.

Evaluate sample results.

Assuming no errors, Roderick can state that he is 95 percent sure that the error rate for the population does not exceed $250,000. Alternatively stated, the auditor can say that he is 95 percent sure that the book value of $5,000,000 is not overstated by more than $250,000. The reader is referred to the February[3] 1982 issue of *The Internal Auditor* for the evaluation of sample results with errors present.

Given the appropriate situation, there are good reasons for the auditor to select DUS over classical statistical sampling. DUS is particularly efficient and effective when the auditor: (1) wants to maximize dollar coverage; (2) wishes to calculate the maximum dollar overstatement; and (3) expects the overstatement to be low. The auditor must exercise professional judgment to determine if the use of DUS is valid for the population being tested and if the model conclusions are consistent with the audit objectives.

Judgment Sampling

Judgment sampling usually receives poor notices whenever statisticians deign to speak of it. But auditors have used it from time immemorial and still find that it performs signal service when statistical sampling is neither needed nor warranted. Judgment sampling remains a significant part of the auditor's sample selection and evaluation procedures. But auditors should know when and how to use it.

Judgment sampling may be used to select examples of deficiencies to support the auditors' contention that the system is weak. They may make a purposive search for defective or improperly processed items to confirm their suspicions or support their position that the system is not capable of identifying improprieties. This is a valid use of judgment sampling. But it should not be used to estimate the number or value of such items in the total population. The auditors had not given every item in the population an equal chance of selection. The test was subjective, not objective. Hence the auditor definitely may not statistically extrapolate sample results to the entire population.

Judgment sampling can be used where it is known that the population has no variability. In an information system, for example, each item may be treated exactly the same by the system. The transactions would be either all wrong or all right. The examination of a single judgment sample will provide the auditor with adequate assurance of the propriety or impropriety of all the items the information system processed.

Exhibit 11-1
Worksheet for DUS Selection

Customer No. (Logical Sampling Unit)	Balance (Book Value)	Cumulative Total	Related Dollar Units (Sample Units)	Unit Selected (See Separate Calculations)	Logical Sampling Unit ("Hooked" Unit Value)
001	$ 1,798	$ 1,798	1 - 1,798	$ 1,000	$ 1,798
002	563	2,361	1,799 - 2,361		
003	974	3,335	2,363 - 3,335		
004	4,621	7,956	3,336 - 7,956		
005	81,897	89,853	7,957 - 89,853	84,333	81,897
006	421	90,274	89,854 - 90,274		
007	8,109	98,383	90,275 - 98,383		
008	1,000	99,383	98,384 - 99,383		
009	10,000	109,383	99,384 - 109,383		
010	67,705	177,088	109,384 - 177,088	167,666	67,705
.	.	.	. - .		
.	.	.	. - .		
.	.	.	. - .		
.	.	.	. - .		
2,000	5,974	5,000,000	4,994,025 - 5,000,000		
	$5,000,000 (100%)				**$2,897,427 (58%)**

NOTE: A sample of 60 units from 2,000 resulted in a 58% "dollar coverage."

Judgment sampling can provide auditors with some clues as to whether to proceed with a statistical sample. If they encounter a well-designed, well-controlled system, good management, well-trained employees, and a feedback mechanism that highlights errors, it would be extravagant to spend a great deal of time performing extensive transaction tests. A small sample — too small for stop-and-go sampling but nevertheless selected at random to obtain some reasonable representation of the population — might suffice.

If the auditors find no errors, they may be able to say that they see no basis for examining the population further or for suspecting any material error. They may not say that they have adequate assurance that the population is truly error-free or even reasonably error-free. They have no statistical basis for such a statement. But what they can say about the functioning of the system may be sufficient for their audit purposes.

Judgment sampling has its place, so long as the auditor is aware of its limitations. Where the audit objectives are fully met by a judgment sample, there would be no valid reason to insist on the discipline of added statistical support.

Statistical and Nonstatistical Sampling

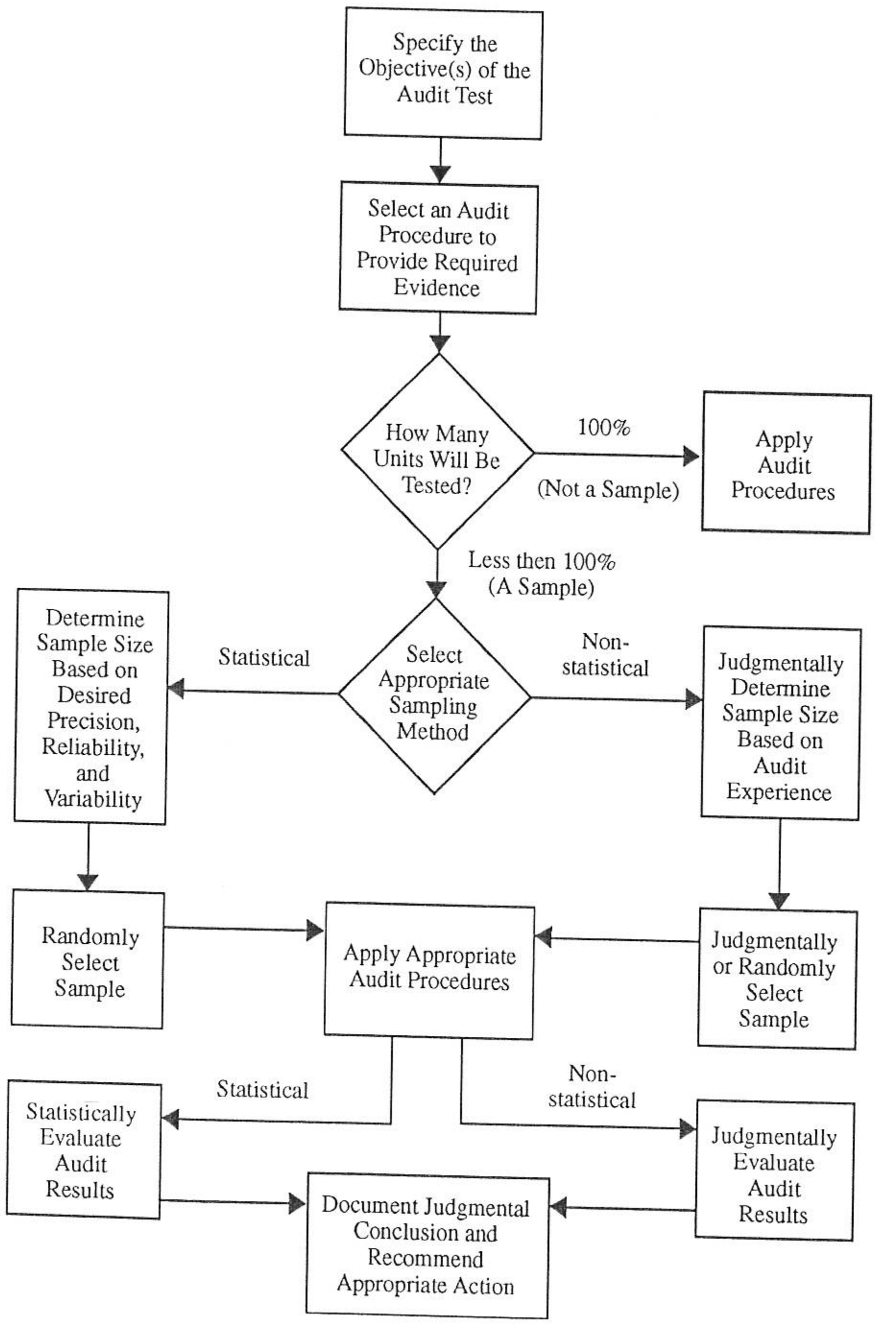
The internal auditor who elects to use audit sampling may select statistical sampling, nonstatistical sampling, or a combination of the two. The selection process is essentially the result of a cost-benefit decision and does not affect the audit procedures performed.

Two criteria are required for a sampling approach to be classified as statistical: (1) The sampling units must be randomly selected; (2) They must be quantitatively evaluated through the application of probability theory. The absence of either requirement defines the approach as nonstatistical (traditionally referred to as judgmental sampling). With nonstatistical sampling, the internal auditor determines the sample size and evaluates the results on the basis of subjective audit experience.

Statistical and nonstatistical sampling both project the characteristics of the sample to the population. Representativeness is generally required by both statistical and nonstatistical sampling unless directed sampling is used (directed sampling involves a search for specified items to prove a hypothesis). Every sample presents a risk that it does not reliably represent the population. Statistical sampling, however, uses the laws of probability to measure sampling risk. However, if the auditor randomly selects a sample but elects not to express the results in statistical terms, the sampling procedure is classified as nonstatistical.

As shown in Exhibit 11-2, there is no difference between statistical sampling and nonstatistical sampling in the execution of a sampling plan, nor does the approach affect the competence of the evidence obtained or the auditor's response to detected errors. Selection between statistical and nonstatistical sampling should be made after a careful evaluation of the advantages and disadvantages of each.

Exhibit 11-2
Statistical and Nonstatistical Sampling



A few of the advantages and disadvantages of both statistical and nonstatistical sampling are:

Statistical Sampling

Advantages

- Provides an opportunity to select the minimum sample size required to satisfy the objectives of the audit tests.
- Provides quantitative expression of the sample results.
- Provides a quantitative measure of sampling risk.
- Provides a measure of the sufficiency of the evidence gathered.
- Permits the auditor to explicitly specify a level of reliability (confidence) and a desired degree of precision (materiality).
- Is simple to apply with software.
- Provides a more defensible expression of the test results.
- Provides for more objective recommendations for management.

Disadvantages

- Requires random sample selection that may be more costly and time consuming.
- May lead to problems in establishing a correlation between the selected random numbers and the units in the population if they are not appropriately organized.
- Might require additional training costs for staff members to use statistics or specialized software.

Nonstatistical Sampling

Advantages

- Allows the auditor to inject a subjective judgment in determining the sample size and selection process to audit items of greatest value and highest risk.
- May be designed so that it is equally effective and efficient as statistical sampling while being less costly.

Disadvantages

- Cannot draw objectively valid statistical inferences from the sample results.
- Cannot quantitatively measure and express sampling risk.
- Presents the risks of either over-auditing or under-auditing.
- May be inappropriate for inexperienced staff members, since the validity of the approach is based on experience with the sampling process.

The advantages of statistical sampling are well documented; however, nonstatistical sampling should not be viewed as less desirable. It may be more appropriate in many audit situations. With the increasing availability and use of audit software, however, the trend may be to use statistical sampling as a first choice and nonstatistical sampling only when statistical sampling is not cost beneficial.

Evaluating Sample Results

Internal auditors cannot content themselves with the mathematical results of their audit samples. True, these results will provide them with a measurable assurance that the sample is a facsimile, in miniature, of the audited population. The results will provide them with an objective estimate of the number of errors in the population or of the true value of the population. However, this is not necessarily what management needs.

When variances occur, management wants to know why they occurred. If the book value of an inventory is $500,000, and the auditor can demonstrate that the physical inventory represented by the books is only $450,000, management wants to know — or should want to know — where the $50,000 went.

Statistical sampling helps provide the auditors with assurance that they have found out what has happened; it cannot tell them why it happened. Thus when samples point to differences, auditors must determine whether the differences are material, how they happened, and what can be done to prevent their recurrence.

The audit objective of modern internal auditors transcends mere scorekeeping — the number of erroneous items they find in a sample. Their objective is to determine what the score means, whether it indicates a system failure, whether it points to poor supervision, whether it is highlighting adverse trends, or whether it hints of manipulation. So unless the sample results provide assurance of satisfactory conditions, the sample is merely a prologue to the real audit task.

A shrewd appraisal of the audit results and surrounding circumstances may give those results an entirely different character. Also, the appraisal may point to the direction the audit report should take. Here are some examples:

- An examination of 200 items discloses only one error. But that error represents a significant item of a material amount. The sample results may not portray a trend, but the individual matter is of sufficient materiality to require reporting to management without reference to the sample. In other words, management may be told that the control system is functioning adequately, although this one matter needs correcting.

- An examination of 100 items discloses 10 errors. The errors resulted from a control breakdown. The sample, taken from a population of 20,000, does not provide good statistical reliability. The auditor, however, feels that the job has been performed and proposes to test no further. The report to management should emphasize the control aspect, buttressed by the fact that 10 percent of the items were in error. There should be no implication that the error rate may reliably be projected to the entire population.

- An examination of 150 items discloses only three errors. But each of the errors is traced back to one clerk who has not been adequately trained. The auditors may then make a purposive test, in addition to their random sample, examining a substantial number of the transactions processed by that clerk to determine the seriousness of the deficiency. The results of the purposive test should not be combined with those of the random sample for the purpose of projecting results to the entire population.

- An auditor selects and examines an attributes sample of 796 items out of a population of 10,000 items and finds a 10 percent error rate. The projection of that rate to the total population carries a 95 percent confidence level with a precision of ±2 percent. This is extremely high reliability. The errors can be attributed to poor supervision. The auditor feels that the population should be purged of error. Management is not mathematically sophisticated and has no conception of the measurement of sample reliability. The auditor may recommend an operating review of the entire population, stating that in his or her opinion the population contains approximately 1,000 erroneous items. But the auditor would be well advised not to discuss confidence levels and precision, since this may do little more than complicate an issue that is better left simple.

Population Proportions

Auditors may have occasion to estimate the proportion of a population that possesses some property of interest. They are not concerned with error rates or sample variability; they merely wish to project the item or items of interest, found in a sample, to the entire population with some measurable degree of reliability.

For example, let us assume that certain purchases are made only after the receipt of competitive bids. Others are not. Still others are made from selected suppliers at the direction of the organization's customers. It may be significant to estimate for management — with adequate reliability — what the proportions are.

There are a variety of formulas that will provide the measure of reliability for such estimates. The use of the binomial probability distribution, or the hypergeometric probability distribution for a finite universe, would define the confidence limits with maximum accuracy. But confidence limits based on these distributions would involve computations so complex that they would not ordinarily be feasible without the use of software.

Fortunately, the formula for the standard deviation of the binomial distribution can readily be used by the auditor to determine the confidence limits with a reasonably acceptable degree of accuracy. We shall discuss here two formulas based on the standard deviation of the binomial distribution: The first applies when the size of the population is known. The second applies when it is not.

Known Populations

Let us assume a population of 40,000 purchase orders. Let us further assume that the auditor selected a random sample of 4,000 purchase orders — every tenth order with a random start — and distributed them to the three categories just enumerated. The results of the sample are:

A. Competitive bids	2,000	50%
B. No competitive bids	1,600	40%
C. Customer direction	400	10%
	4,000	100%

By simple projection, or "blow up" (multiplying each sample group by 10), the auditor estimates that the population contains 20,000 of A, 16,000 of B, and 4,000 of C. But how reliable is the estimate? In other words, what is the precision range for the estimate at a stipulated confidence level? The auditor can obtain that statement of reliability through the following formula:

$$A = ZN\sqrt{\frac{N - n}{N \times n}} \times \sqrt{p(1 - p)}$$

where:

A = Precision
Z = Normal deviate for the desire level of confidence
N = Population size
n = Sample size
p = Proportion of items of interest to sample

The computation of the range of precision, assuming a stipulated confidence level of 95.5 percent, is as follows:

Number in population	40,000
Number in sample	4,000
Z at 95.5% confidence level	2,000
Proportions of items with a particular characteristic	
A. 2,000/4,000	.50
B. 1,600/4,000	.40
C. 400/4,000	.10

First solve:

$$ZN \sqrt{\frac{N - n}{N \times n}}$$

$$2 \times 40{,}000 \sqrt{\frac{40{,}000 - 4{,}000}{40{,}000 \times 4{,}000}}$$

$$80{,}000 \sqrt{\frac{36{,}000}{160{,}000{,}000}}$$

$$80{,}000 \sqrt{.000225}$$

$$80{,}000 \times .015 = 1{,}200$$

Then solve, for each class:

$$\sqrt{p(1-p)} \times 1{,}200$$

Class	p	1 - p	p(1 - p)	p(1 - p)	x	1,200 = A Precision (±) at 95% Confidence Level
A	.50	.50	.25	.5000		600
B	.40	.60	.24	.4900		588
C	.10	.90	.09	.3000		360

Estimated Proportions and Reliability Statements

	Classification	Sample POs	Sample % of Total	Estimated POs in Population	Precision (±) at 95% Confidence Level POs	Precision (±) at 95% Confidence Level %
A.	Competitive bids	2,000	50	20,000	600	3.0
B.	No competitive bids	1,600	40	16,000	588	3.7
C.	Customer direction	400	10	4,000	360	9.0

It will be observed that the formula provides the best reliability when the item of interest represents a relatively high proportion of the sample.

Unknown Populations

Let us assume a large but unknown population of purchased tools. Auditors want to estimate with reasonable reliability how many cost $100 and over. Assume a sample of 400 tools with the division as:

Under $100	320
$100 and over	80
	400

Without an idea of population size, the auditors will be unable to estimate how many tools cost $100 or more. But they can estimate the proportion of items of interest. In this case, they can estimate that 20 percent of the population contains tools costing $100 or more. They will then seek to determine the reliability of that estimate to help decide whether they have taken a large enough sample for their purposes. The formula used to determine the precision of the estimated proportion is:

$$A = \pm Z \sqrt{\frac{p(1 - p)}{n}}$$

where:

- A = Precision
- p = Proportion
- n = sample size
- Z = Normal deviation for the desired level of confidence

The computation of the reliability statement for the estimate of tools at a 95 percent confidence level is:

$$A = \pm 1.96 \sqrt{\frac{.20\ (1 - .20)}{400}}$$

$$= \pm 1.96 \sqrt{\frac{.20 \times .80}{400}}$$

$$= \pm 1.96 \sqrt{\frac{.16}{400}}$$

$$= \pm 1.96 \sqrt{.0004}$$

$$= \pm 1.96 \times .02$$

$$= \pm .0392$$

Thus, the precision is .20 ±.0392, or approximately from 16 to 24 percent. In other words, at a 95 percent confidence level, between 16 and 24 percent of the tools cost $100 or more.

If the auditors want a more precise result, they might increase their sample size. Assuming that the sample size is doubled to 800 and the sample results remain the same, the precision would be computed:

$$A = \pm 1.96 \sqrt{\frac{.20\ (1 - .20)}{800}}$$

$$= \pm 1.96 \sqrt{\frac{.16}{800}}$$

$$= \pm 1.96 \sqrt{.0002}$$

$$= \pm 1.96 \times .01414$$

$$= \pm .0277$$

Now, the precision, at a 95 percent confidence level, is ±2.77 and the range is approximately from 17 to 23 percent. If this is sufficient reliability in the auditors' judgment, they may stop sampling.

Summary

In dealing with audit sampling, the auditor should keep these 10 commandments in mind:

1. Know the principles of scientific sampling, but use them only when they best fit the audit objectives.
2. Know the population, and base audit opinions only on the population sampled.
3. Let every item in the population have an equal chance of being selected.
4. Do not let personal bias affect the sample.
5. Do not permit patterns in the population to affect the randomness of the sample.
6. Do not draw conclusions about the entire population from the purposive or directed (judgment) sample, even though it does have its place.
7. Base estimates of maximum error rates on what is reasonable in the real world; try to determine at what point alarms would automatically go off.
8. Stratify wherever it would appear to reduce variability in the sample.

9. Do not set needlessly high reliability goals (confidence level and precision). Controls, supervision, feedback, self-correcting devices, and management awareness and surveillance should all be considered in seeking to reduce the extent of the audit tests.
10. Do not stop with the statistical results; know why the variances occurred.

In deciding which selection or sampling plan to use, the auditor should consider these applications:

Recommended Selection Technique	Recommended Sampling Plan
Random Numbers. Where each of the items in the population is or can readily be numbered or is included in lists or registers that are or can be numbered.	**Attributes.** To estimate the attributes or characteristics of a population — obtaining "yes or no" answers — with a measurable degree of reliability.
Interval. Where items are not or cannot readily be numbered, or where random number sampling would be excessively expensive. (Steps must be taken to avoid any bias that may be introduced by patterns in the population or by items missing from the population.)	**Stop-and-Go.** To estimate error rates or similar attributes from the smallest possible sample — discontinuing the sampling when a definitive answer is obtained.
Stratification. Where the population is composed of items that may vary considerably in value or in other characteristics of interest, and where sample size can be reduced by separating the population into groups of items with reasonably similar values or characteristics.	**Discovery.** To identify through sampling at least one suspected item — assuming some given number of such items in the population — and discontinuing sampling when at least one item is identified.
Cluster or Multistage. Where the population is so dispersed that random number or interval sampling would be burdensome. It must be remembered that there usually may be a loss of sample reliability when cluster or multistage sampling is used, as compared with random number or internal sampling, and that a larger sample size may usually be required to offset that loss.	**Variable.** To estimate the value of a population — dollars, weights, time spans, or other variable values — with a measurable degree of reliability. **Judgment.** To use samples for the purpose of obtaining information that may not be attributed to the entire population with measured reliability.

References

[1]Andrews Jr., W.T., and A.G. Mayper, "Dollar-Unit Sampling," *The Internal Auditor,* April 1983, 31.
[2]Guy, D.M., and D.R. Carmichael, *Audit Sampling: An Introduction to Statistical Sampling in Auditing,* 2nd Ed. (New York: John Wiley & Sons, 1986).
[3]Harwood, G.B., R.E. Schiffler, and J.M. Woods, Jr., "Putting Dollar-Unit Sampling to Work," *The Internal Auditor,* February 1982, 50-55.

Supplementary Readings

Apostolou, Barbara, and Francine Alleman, *Internal Audit Sampling* (Altamonte Springs, FL: The Institute of Internal Auditors, 1991).
Braun, Gary P., and Patricia M. Myers, "The Use of Statistical Sampling by Internal Auditors: Current Practices," *Internal Auditing*, September/October 2000, 13-16.
Calhoun, Charles, "International Guidance on Audit Sampling," *Internal Auditing Alert*, April 1998, 5-6.
Colbert, Janet, "Audit Sampling," *Internal Auditor*, February 2001, 27-29.
Hitzig, Neal B., "Audit Sampling: A Survey of Current Practice," *The CPA Journal*, July 1995, 54-57.
Sears, Brian P., *Internal Auditing Manual* (New York: RIA, 2003.) (See Chapter D5, "Testing and Sampling.")
Wendell, John P., "The Dual-Risk Method for Calculating Attribute Sampling Sizes," *Internal Auditing,* January/February 1998, 19-22.

Multiple-choice Questions

1. A sampling plan is needed to test for overstatement of a $3 million accounts payable book balance. The auditor determines that a $100,000 error is material and a 95 percent confidence level is appropriate. Based on these determinations, the sample size of 90 is needed. The sampling plan *most* likely to be used is:
   a. Stop-and-go sampling.
   b. Cluster sampling.
   c. Dollar unit sampling.
   d. Attributes sampling.

2. To use stratified sampling to evaluate a large, heterogeneous inventory, which of the following would least likely be used as criteria to classify inventory items into strata?
   a. Dollar values.
   b. Number of items.
   c. Turnover volume.
   d. Storage location.

3. Assuming no change in sample standard deviation, how would sample size and achieved precision be affected by a change in confidence level from 95.5 percent to 99.7 percent?
   a. Sample size would be smaller, but achieved precision would be larger.
   b. Sample size would be larger, but achieved precision would not change.
   c. Sample size would be smaller, but achieved precision would not change.
   d. Sample size would be larger, but achieved precision would be smaller.

4. An internal auditor is preparing to sample accounts receivable for overstatement. A statistical sampling method that automatically provides stratification when using systematic selection is:
   a. Attributes sampling.
   b. Ratio-estimation sampling.
   c. Dollar unit sampling.
   d. Mean-per-unit sampling.

5. In the audit of a health insurance claims processing department, a sample is taken to test for the presence of fictitious payees, though none are suspected. The most appropriate sampling plan would be:
   a. Attributes sampling.
   b. Discovery sampling.
   c. Variable sampling.
   d. Stop-and-go sampling.

6. An auditor wishes to sample 200 sales receipts from a population of 5,000 receipts issued during the last year. The receipts have preprinted serial numbers and are arranged in chronological (and thus serial number) order. The auditor randomly chooses a receipt from the first 25 receipts and then selects every twenty-fifth receipt thereafter. The sampling procedure described here is called:
   a. Systematic random sampling.
   b. Dollar unit sampling.
   c. Judgmental interval sampling.
   d. Variable sampling.

7. Each time an internal auditor draws a conclusion based on evidence drawn from a sample, an additional risk, sampling risk, is introduced. An example of sampling risk is:
   a. Projecting the results of sampling beyond the population tested.
   b. Using an improper audit procedure with a sample.
   c. Incorrectly applying an audit procedure to sample data.
   d. Drawing an erroneous conclusion from sample data.

8. A confidence level of 90 percent means that:
   a. The expected error rate is equal to 10 percent.
   b. The point estimate obtained is within 10 percent of the true population value.
   c. There are 90 chances out of 100 that the sample results will not vary from the true characteristics of the population by more than a specified amount.
   d. A larger sample size is required than if the desired confidence level were equal to 95 percent.

9. Internal auditing is conducting an operational audit of the organization's mail room activities to determine whether the use of express mail service is limited to cases of necessity. To test cost-effectiveness, the auditor selects the 100 most recent express mail transactions for review. A major limitation of such a sampling technique is that it:
   a. Does not allow a statistical generalization about all express-mail transactions.
   b. Results in a sample size that is too small to project the population.
   c. Does not evaluate existing controls in this area.
   d. Does not describe the population from which it is drawn.

10. A statistical sampling technique that will minimize sample size whenever a low rate of noncompliance is expected is:
   a. Ratio-estimation sampling.
   b. Difference-estimation sampling.
   c. Stratified mean-per-unit sampling.
   d. Stop-and-go sampling.

11. In a regional survey of suburban households to obtain data on television viewing habits, a statistical sample of suburban areas is first selected. Within the chosen areas, statistical samples of whole blocks are selected, and within the selected blocks, random samples of households are selected. This type of sample selection can best be described as:
   a. Attributes sampling.
   b. Stratified sampling.
   c. Cluster sampling.
   d. Interval sampling.

12. Using random numbers to select a sample:
   a. Is required for a variables sampling plan.
   b. Is likely to result in an unbiased sample.
   c. Results in a representative sample.
   d. Allows auditors to use smaller samples.

13. An auditor tested a population by examining 60 items selected judgmentally and found one error. The main limitation of the auditor's sample is the inability to:
    a. Quantify sampling risk.
    b. Quantify the acceptable error rate.
    c. Project the population's error rate.
    d. Determine whether the sample is random.

14. A sample from a population of over 10,000 bills of lading is needed to estimate an error rate. Since a sample size of 250 will satisfy precision and confidence level needs, a sampling interval of 40 is chosen. For ease of implementation, the auditor randomly selected a number between 1 and 40, and then selected each succeeding 40th item. Which of the following is true?
    a. The sample lacks randomness and will not be correct.
    b. Interval sampling is not an acceptable statistical method.
    c. If the population lacks bias, the sample is statistically valid.
    d. Interval sampling eliminates the use of auditor judgment.

15. The supervisor of claims processing for a health insurance firm selects all claims processed in the past two days by a particular employee for audit. From this sample, the supervisor can develop:
    a. An overall representative view of employee work for the year.
    b. A quantification of sampling error.
    c. Conclusions about the correctness of processing for the department.
    d. Understanding the details contained in the processing task.

16. To audit invoices paid over the past year, an auditor selects the two busiest months that account for 60 percent of invoices. Following a random start, every tenth invoice is chosen, yielding a sample of 116 invoices. This sample may not be valid because it is not a:
    a. Representative sample.
    b. Random sample.
    c. Large enough sample.
    d. None of the above — sample is valid.

17. An internal auditor with an international shipping organization needs to sample shipping records over the past six months. To do so, the auditor draws a random sample for ships operating in the Mediterranean and a separate sample for those operating in the North Atlantic. This method of sampling is called:
    a. Cluster sampling.
    b. Haphazard.
    c. Interval sampling.
    d. Stratified sampling.

18. By statistically projecting the population value based on the average value of sampled subsidiary accounts, the auditor has estimated the value of the total equipment account to be $2,800,000. This is an example of:
    a. Dollar unit sampling.
    b. Mean-per-unit sampling.
    c. Attributes sampling.
    d. Statistical difference estimation.

19. An inventory listing consisting of approximately 2,050 unnumbered items is arranged by category, with 10 items in each category. Within each category the most expensive (per unit) items are listed first. An auditor wants to use an interval sampling plan to select a representative sample of at least 100 items from the population. The best technique is to:
    a. Select a random number from 1 to 20 as the starting point and then select every 20th item, moving through the entire population.
    b. Select a random number from 1 to 15 as the starting point and then select every 15th item until he or she has 100 items.
    c. Select seven random digits from 1 to 135 as the starting points and then select every 135th item per pass, moving the entire population seven times.
    d. Select the 50 largest items (i.e., extensions with the highest dollar amounts); then excluding the 50 largest items already selected, select a random number from 1 to 37 as the starting point and select every 37th item, moving through the entire population.

20. Using organization policies to establish when approval is needed, an auditor has sampled accounts receivable balances exceeding $1,000 to determine whether the credit department is requiring a credit check for credit sales when appropriate. This is an example of:
    a. Dollar unit sampling.
    b. Mean-per-unit sampling.
    c. Attributes sampling.
    d. Variables sampling.

21. The probability that an estimate based on a random sample falls within a specified range is known as the:
    a. Error rate.
    b. Lower precision limit.
    c. Confidence level.
    d. Standard error of the mean.

22. One objective of an audit of the purchasing function is to determine the cost of late payment of invoices containing trade discounts. The appropriate population from which a sample would be drawn is the file of:
    a. Receiving reports.
    b. Purchase orders.
    c. Canceled checks.
    d. Paid vendor invoices.

23. A sample of 100 items was taken from a population of 5,000 items. The mean value was $200 and the standard deviation was $30. The computed confidence interval for a 95 percent confidence level (z = 1.96) is:
    a. $970,600 to $1,029,400.
    b. $706,000 to $1,294,000.
    c. $996,733 to $1,003,267.
    d. $997,060 to $1,002,940.

24. The range into which an estimate of a population characteristic is expected to fall at a stated confidence level is known as the:
    a. Precision.
    b. Measure of central tendency.
    c. Standard deviation.
    d. Sampling field.

25. An internal auditor suspects that the invoices from a small number of vendors contain serious errors and therefore limits the sample to only those vendors. A major disadvantage of selecting such directed sample of items to examine is the:
    a. Difficulty in obtaining sample items.
    b. Inability to quantify the sampling error related to the total population of vendor invoices.
    c. Absence of a normal distribution.
    d. Tendency to sample a greater number of units.