A sampling plan is nonstatistical when it fails to meet at least one of the criteria of a statistical sampling plan (Apostolou & Alleman, 1991a). Yet, there are similarities between the two types of plans. Both require judgment on the part of the analyst. Audit procedures will be the same for either type of plan, and both are permissible by the standards of the Institute of Internal Auditors (IIA, 2013). The major difference between them is that sampling risks cannot be measured in a nonstatistical sampling plan.

There are circumstances in which nonstatistical sampling is the more appropriate of the two. Nonstatistical sampling may be used when results are needed to confirm a condition, or when judgment is more important than random sampling, such as in reviewing records, searching for unusual relationships, or interviewing client personnel (Guy, 1981a). Nonstatistical sampling is also correct when there is no sampling—the entire lot is tested. This seems to be an oxymoron; a sampling procedure for a non-sampled population. Yet, sampling plans bring a formality and direction necessary to any analysis. Consider that even if you intend to test the entire lot, it may be that some of the evidence is trivial or irrelevant and will be abandoned, in which case the balance becomes a sample.
The analyst should understand that a nonstatistical plan does not mean that statistical methods are not used. On the contrary, they are used. Rather, it means that sampling error cannot be measured in a nonstatistical plan, so conclusions based on such statistics are easily challenged. But statistical methods, sampling error or not, can be used to provide direction to corroborating evidence.

Although nonstatistical methods cannot measure the sampling risks, it is possible to derive important conclusions from the results of a sample, using information derived from the control parameters. This is possible if the test is a Bernouilli trial—a random experiment in which there are only two mutually exclusive outcomes and in which the probability of a given event is constant. Within the context of this book, the mutually exclusive events are conformity and nonconformity. The constant probability of concern is the rejection or deviation rate, that is, a control that is nonconforming to requirements (Grant & Leavenworth, 1988). The rate has a binomial distribution about which limiting probabilities can be estimated, such as upper deviation rates, which is another reason for using statistical rules whenever possible in a nonstatistical sampling plan.

F.1 Sampling Format

The sampling format for a nonstatistical plan is the same as for a statistical plan and consists of framing the dimensions of the sampling: its objectives, scope, and basis. The primary interest here is how statistical procedures might fit into a nonstatistical analysis, but a holistic approach to the question increases the confidence level of a nonstatistical plan. Qualitative assessments may be used in lieu of the quantitative approach used in statistics, with necessary precautions and supporting evidence.

The format will focus on those characteristics of internal control deemed essential to acceptable performance. The nonstatistical sampling plan most relevant to this book is an attribute plan and the same seven-step strategy suggested in Appendix E for statistical sampling is used in a nonstatistical audit:

- Determine the objectives of the test
- Define the attribute and deviation conditions
- Define the population
- Determine the method of sample selection
- Determine the sample size
- Perform audit procedures on the sample
- Evaluate the results and state conclusions

F.1.1 Frame of the Sampling Plan

A sampling plan is framed by its purpose, scope, and basis. The purpose of the plan is, whether statistical or nonstatistical, within the context of this book, to assess the performance of internal controls.
The scope of the analysis is the set of functions or processes to be assessed. In the case of assessing the operations of a corporation, an approach can be general, in which all activities are analyzed, or specific, in which the focus is on a subset of the activities. The subset may be selected because of an inquiry into an area of special concern, or perhaps because of an identified set of consistent problems. But even if the analysis is general, the activities to be assessed must be specified if the analysis is to be bounded.

The basis of the analysis is the set of applicable standards, rules, or regulations from which the assessment is evaluated. These requirements will be mapped into each activity according to their applicability. This is called a vertical audit. If you are going to analyze a purchasing function of the company, then requirements applicable to the purchasing function are appropriate and the requirements applicable to, say, design reviews are not.

F.1.2 Attribute and Deviation Conditions

The attribute and deviation conditions are the same as those for statistical sampling. An internal control is deemed either conforming or nonconforming to requirements. The conditions of interest are deviation rates and are estimated as described in Chapter 14: in terms of major, moderate, or minor risks to effective control. This is a subjective decision tempered by the period of time in which the nonconformity exists. A single event may be simply regarded as an outlier, but risk increases with time and an extended period of nonconformity is always a major risk.

F.1.3 The Population

The population will be a system or subsystem of internal controls of the operations under study. The sample population will probably, but not necessarily, be a subset of the total. It is possible that the entire lot will be tested, in which case the study is no longer a sample, but statistical sampling procedures are nevertheless continued in order to enhance the credibility and validity of the results.

The sampling procedures are where the choice of nonstatistical and semistatistical activity comes into play and are the same for both statistical and nonstatistical plans. Even though you do not use probability laws in a nonstatistical sampling plan, using statistical procedures enhances objectivity. In keeping with statistical discipline where appropriate, records for inquiry should be done randomly even if the sample size is too small to make assumptions about the distribution of the greater population.

A population is a set of objects or persons possessing the characteristic(s) that are being studied. A representative sample is a subset of the population, each member of which possesses the characteristics that represent the entire population. A simple analogy is that of a sample representing a population of
oranges. The sample should not contain apples. However, this analogy is too simple; differences among members can be quite subtle and is greatly influenced by the objectives and characteristics of the study. For example, suppose the effect of a certain medication is under study. Depending upon the study objective and the nature of the effect, samples taken from a population of those using the medication may not be representative if some members of the sample were young adults and some were senior citizens.

Sometimes a representative sample is called a random sample, but the two are not the same. A random sample is one taken from a lot in which each member of the lot has an equal probability of being selected. Thus, a sample is selected from a lot in such way that it is homogeneous, representative, and random.

F.1.4 Nonstatistical Sample Selection

The selection of samples is made either by haphazard, blocking, judgment, or stratified techniques. Haphazard selection is a selection made without bias, and this is best done through randomness if possible. You must be satisfied that the sample is representative of the population. Defined probability concepts are not used, nor is statistical inference.

Block selection means that you select a homogeneous subgroup, say all the job orders for a given day or week. To ensure fair representation, many blocks should be selected. Randomness can be introduced by randomly selecting the days or weeks for which you will examine the job orders. For example, suppose you want to examine job orders over the last 30 days. Then using block selection, you choose all the job orders for several randomly selected sets of, say, 4 or 5 days, depending upon your time available.

Judgment means that samples are taken from areas that you believe have high control risk, either due to the nature of the process or due to past audit results. Shipyards provide a good example. A typical US Navy ship repair contract will contain over 5000 work items, but only about one-tenth of those work items will be on the critical path. These are called controlling work items, and their status provides a good estimate of repair progress. Once you have selected a set of controlling work items or similar set of high-risk items for examination, then you should randomly select from that pool. This is where the notion of materiality comes into play. It is a departure from randomness but may be needed for effective selection.

In regard to materiality, there are three definitions that are pertinent to the sampling of evidence: (i) Materiality is a measure of the effect that an item of information will have on the accuracy or validity of a judgment; (ii) Materiality is the threshold of importance of a given process on the performance of operations; (iii) “Information is material if its omission or misstatement could influence the economic decisions of users taken on the basis of financial statements” (IASB, 2001).
The first definition refers to a judgmental error on the part of an analyst vis-à-vis the conformity of a control. The second refers to the importance of a given control relative to the operations in question. The third definition refers to the financial importance of the control. Investors purchase shares in a company based in part on its price/earnings ratio. Thus, materiality is tied to that value of a company’s estimated worth that could influence investors. But materiality also reflects corporate operations because financial statements are about the costs of doing business. They include all the costs associated with operations, which itself is a function of the quality of its performance. Thus, all three definitions come together in the judgment of the analyst of a system of controls of operations.

F.1.5 Sample Size

In statistical sampling, the choice of sample size was influenced by five parameters: the expected system deviation rate (SDR); the acceptable deviation rate (ADR), the desired Alpha and Beta error probabilities, and the desired confidence level. An alpha error (type I) occurs when the analyst assesses a control as ineffective, when in reality the control is effective, with little nonconforming output. A beta error (type II) occurs with an assessment by the analyst of a control as effective, when in reality the control is ineffective.

Thus, Alpha and Beta errors are made in the evaluation and analyses of sample results and will happen in nonstatistical audits as well. However, these errors cannot be measured in nonstatistical sampling because not all of the conditions required of statistical inference are present. Nevertheless, you need to know, to a “reasonable” degree, whether a control is effective or not.

A nonstatistical plan, properly conducted, will yield a valid proportion of deviations to sample size and you are free to preselect an ADR. It is the interval between these two rates that cannot be accurately measured, but you can still make a good estimate of control effectiveness. There is good agreement among audit experts on the assignment of numbers to a qualitative assessment of a control. Table F.1 compares the value system of several experts, such as Kelly (1986) and Guy (1981b).

Table F.1 is read in this way. Using qualitative judgment, you select a desired confidence level, say “substantial.” Then the table provides the ADR that can reflect this level. Then armed with this rate and an expected SDR, you can select sample size from Table E.1.

The levels of the ADRs of Table F.1 may seem unusually large to people in operations, but they must be understood in context. They apply to the deviation rate of controls, and not to the deviation rate of key characteristics of a product or service. Table E.1 does not provide sample sizes for rates greater than 10%, but they can be determined from Equation E.1 if occasion demands.
However, Guy (1981c) advises against using ADRs greater than 10% if the purpose of the inquiry is about the reliance of internal controls, as opposed to some other of their characteristics. As reliance of a control is a major issue in systems operation, then 10% can be considered an upper limit for ADRs for purposes of forensic systems examination.

As in statistical sampling, in a nonstatistical plan you use knowledge of past system performance to estimate the control deviation rate in order to arrive at a sample size. Then having chosen the sample size and conducted the test, you evaluate the control based on the difference between the sample and the ADRs.

### F.1.6 The Effect of Sample Size on Beta Error

Many analysts consider Beta error as the more grave sampling error, as Alpha error tends to be self-correcting. If an analyst makes an Alpha error, the decision will be challenged because the analyst is challenging a control that the performer believes to be good. The performer will insist on additional testing with larger sample size. However, a Beta error will not be challenged; the analyst reports good news to the performer, although the good news is false. In addition, additional testing increases cost. Apostolou (1991b) classifies an Alpha Error as an efficiency indicator; a Beta error as an effectiveness indicator.

Nonstatistical sampling plans do not measure Beta error, but you know that it exists. You must make some effort to reduce it, and sample size has much to do with it. Sample size is affected by the error rate as shown in Table E.2, in which changes in sample size affect key system parameters and conversely.

For example, suppose that you want to choose a smaller sample size. By doing so, you increase the chance of Beta error—the risk of assessing a control as effective when, in fact, it is not, because sample size and Beta error are inversely related. As another example, if you want to increase the ADR, then you can decrease the sample size. But in doing so, you increase the chance of Beta error. The reason should be clear. A larger ADR offers a less effective
control, or a less accurate sample of a good control. But if you want to tighten the system assessment by decreasing the ADR, then a larger sample population is needed, again to err on the fail-safe side.

As a final example, if you expect a low SDR, then you can use a lower sample size, understanding that the lesser accuracy is unimportant for a very good system. But there, again, the chance of Beta error increases. On the other hand, if you expect a high system error rate, possibly above the acceptance rate, then you need a larger sample size to be sure of the margins. These considerations require judgment, experience, and knowledge of the system, but they enhance the validity of a nonstatistical plan.

F.1.7 Evaluating Sample Results

Because nonstatistical sampling does not provide a reliable estimate of sampling risk, the analyst must make a judgment of whether the difference between the ADR and the measured deviation rate is an adequate allowance for sampling error. For example, suppose the analyst expects a SDR of about 1% and will accept a deviation rate (ADR) of 7%.

From Table E.1, a sample of 83 is taken. Suppose further that in this sample, two errors are found. Then the measured deviation rate is \( \frac{2}{83} = 2.4\% \). The analyst must then decide whether the 4.6% difference between the measured rate and the ADR is sufficient to cover sampling errors.

Here is where the more you know about the process the better will be your judgment. A sample of \( \frac{2}{83} \) is about as likely to derive from a distribution centered at 5% as from a distribution centered at 1%. In this case, the measured deviation rate (2.4%) was more than double the expected rate (1%), so a test of the 95% confidence level is called for. A BETAINV test of \( \frac{2}{83} \) shows the result to have a 95% confidence that the true SDR is less than 7.4% (Apostolou, 1991c). As this limit is just over the ADR, the control should be deemed effective in order to avoid an unwinnable challenge in litigation.

Thus, the bounds and confidence levels of a test of controls are critical and must be justified, with generous allowance made for uncertainty. For example, if the confidence level were set at 90%, the upper limit of the SDR at \( \frac{2}{83} \) would have been 6.8%, thus below the ADR. If the expected SDR were 1.5% instead of 1%, the upper limit of the SDR at \( \frac{2}{109} \) would have been 5.7%, well below the ADR.

F.2 Nonstatistical Estimations

Sampling conclusions are always estimations, irrespective of the type of sampling plan that is used. Given that caution, statistical sampling plans offer a means to estimate the errors in comparing the sample to the population. Nonstatistical sampling plans cannot estimate measurement errors and require
extraordinary understanding of the processes being monitored and larger margins between ADRs and estimated deviation rates.

Nevertheless, numerical results from nonstatistical sampling are useful when framed to answer the right questions: (i) “What is the highest SDR that is likely to yield the sample?” (ii) “What is the smallest Beta risk that is likely, given the sample?” Still, forensic systems engineering is about preparing engineering data for trial. All the conclusions based on the analysis of evidence can be subject to challenge and rebuttal. Therefore, when using a nonstatistical sampling plan the benefit of doubt in close margins of deviation rates goes to the defense. Chapter 17 addresses this dilemma and recommends a forensic focus on systemic failure when statistical conclusions derived from the evidence contain controversial margins in probable sampling errors.

References