A statistical sampling plan, designed especially for auditors, that would help reduce nonsampling error by organizing available data into a quantifiable form for decision-making

RELATING STATISTICAL SAMPLING TO AUDIT OBJECTIVES

BY ROBERT K. ELLIOTT AND JOHN R. ROGERS

For many years independent auditors have been encouraged to use statistical sampling procedures in conducting audit tests. The advantages have been pointed out in many books and articles on the subject. Most auditors trained in the last decade have been exposed to statistical sampling during their formal training and encouraged to use it. The American Institute of Certified Public Accountants has encouraged its use. Notwithstanding all of this encouragement, use of statistical techniques is still not widespread within the profession. Although many factors contribute to the lagging implementation of statistical techniques, two stand out as most important:

1. Statistical techniques have generally been cumbersome to apply manually.
2. Statistical techniques have not been adequately related to common audit objectives.

With the general availability of computers in auditing, the cumbersome aspects of statistical sampling procedures need no longer be done manually. The computer will handle the intricacies of sample size calculation, sample selection and evaluation effortlessly and without mechanical error. Computerization of statistical techniques will virtually eliminate the first major deterrent to the use of statistical sampling in auditing.

However, the second major deterrent must still be eliminated before statistical sampling can be widely implemented. The purpose of this article is to suggest an approach to relating statistical techniques to audit objectives.

AUDIT OBJECTIVES

The overall objective of an auditor in carrying out the attest function is to make reasonably certain that financial statements examined by him are materially correct.[1] He therefore wishes to find material errors in the financial statements if they exist. To do this he employs numerous procedures, one of the most important being the examination of documentary evidence. Auditors long ago abandoned any attempt to examine documentary evidence for each and every item represented in the financial statements. The resulting partial (or sample) examination *must* introduce an element of uncertainty into his conclusions, for he must draw inferences from incomplete information. This is true whether or not the auditor uses statistical sampling procedures.

The uncertainty resulting from partial examination may frustrate the auditor's desire for materially correct financial statements in two distinct ways:

1. Financial statements may be correct, but he may erroneously conclude they are incorrect and either qualify his opinion or unnecessarily insist upon adjustments.
2. Financial statements may be materially in error, but he may erroneously conclude that they are correct and give an unqualified opinion.

Although financial statements may be in error by any amount (from zero on up), it is only necessary to control audit risks at the points of zero error and material error (that is, exactly the amount specified by the auditor as material) in order to meet the overall audit objective. Therefore, for the remainder of this article, only these two magnitudes of error will be considered. The level of risk at all other magnitudes can be observed in Figure 1, page 51.

[1] "Materially correct" in this context means "fairly presented in conformity with generally accepted accounting principles."

For convenience, the risk of rejecting perfectly correct financial statements will be referred to as the α (alpha) risk, and the risk of accepting financial statements in error by exactly a material amount will be referred to as the β (beta) risk.

These two risks exist even when the auditor uses judgment sampling procedures, but they cannot in any way be quantified. The result is that, in some cases, auditors do less work than necessary to control these risks and, in other cases, more work than necessary. When too little work is done, the auditor is not effectively carrying out the attest function. When too much work is done, the auditor is carrying out the attest function effectively, but at an excessive cost.

USE OF STATISTICAL TECHNIQUES

Statistical sampling procedures are ideally suited to measure the α and β risks inherent in partial (or test) examinations. If the auditor will state what α and β risks are justified, and what amount he considers material for a given account, a statistical test can be designed to meet these criteria. The ability to measure, in this case, means there is an ability to control.

When the auditor performs a documentary examination, he may have either or both of two objectives:

1. To establish the effectiveness of systems and procedures, in order to plan the type, extent and timing of other audit procedures.
2. To establish the material correctness of a financial statement amount.

Any relevant test of material correctness must be stated in *dollar* terms. Statistically, it must be a test of *variables*. An attribute test, conversely, can express conclusions only in terms of rates of occurrence, not dollars.

An attribute test can indicate, for example, that the error rate in certain documents is probably no more than a certain percentage. While this may be of some use to the auditor, it is not a conclusion that can be measured in terms of financial statement impact. A low error rate does not necessarily mean that the financial statements are materially correct, nor does a high error rate mean the opposite. The principal usefulness of the technique is to assist the auditor in making qualitative judgments about the operation of internal controls and accounting systems. These judgments may then be used in planning the type, extent and timing of other audit procedures, but they cannot normally be used for direct audit conclusions concerning fair presentation of the financial statements.

Although a statistical attribute test may be of some use in examining systems and controls, it is by no means necessary. Generally, in examining systems the auditor seeks a qualitative conclusion (for example, the controls over cash disbursements are good). Seldom does he need precise measurement of rates of occurrence, such as would be provided by an attribute test. Neither is he overly concerned with α and β risk levels.

When the auditor needs quantifiable conclusions, though, he will be best served by a variables test (usually one in which the results are stated in dollars). Therefore, the remainder of this article will concern itself solely with variables tests designed to establish the material correctness of financial statement amounts.

PUBLISHED LITERATURE

Surprisingly, there does not appear to be any published literature which deals explicitly with the use of statistical techniques to measure and control the auditor's α and β risks with respect to variables testing. The AICPA programed instruction series (*An Auditor's Approach to Statistical Sampling*, Volumes I through IV) and all known textbooks and articles known to us on audit uses of variables sampling refer to and use variables *estimation* techniques. They propose that the auditor sample a population at a given confidence and precision to estimate the true value.

Generally, the auditor is advised to relate confidence to the level of internal control and precision to materiality, but he is not given guidance in placing numerical values on confidence and precision. Furthermore, he is not given any advice on what to do with the statistical results. These are correctly asserted to be auditing, not statistical, matters. However, without some assistance in un-

ROBERT K. ELLIOTT, *CPA, is manager in the executive office of Peat, Marwick, Mitchell & Co. in New York City. He is a member of the AICPA and the Pennsylvania Institute of CPAs. His article, "Aspects of Behavioral Accounting" was published in Statements in Quotes in the April 1971* JOURNAL. *Mr. Elliott received his bachelor's degree from Harvard College and his M.B.A. degree from Rutgers University.* JOHN R. ROGERS, *CPA, is a partner of Peat, Marwick, Mitchell & Co., in New York City. He serves on the Institute's committee on statistical sampling and is a member of the New York State Society of CPAs and the Canadian Institute of Chartered Accountants. Mr. Rogers received his bachelor's degree from Dartmouth College and his M.B.A. degree from Amos Tuck School of Business Administration.*

derstanding how statistical inferences can be related to auditing objectives, most auditors are reluctant to use statistical techniques. Even when they are used, an auditor is not able explicitly to control his α and β risks. That is, an auditor who has statistically tested $10 million in accounts receivable and has 95 per cent confidence that the true value is between $9 million and $9,500,000 has some useful audit information, but, based on the standard approach, he cannot come to any conclusion as to what α and β risks he has incurred.

In the absence of guidance, many auditors revert to rules of thumb in applying statistical techniques. The following approach is typical: let precision (P) equal the amount considered material (M) for an account; let the confidence level (CL) be 95 per cent if internal control is poor, or less (say 80 per cent) if internal control is excellent; if the statistical confidence interval[2] includes the book value, accept it as correct, otherwise reject the book value.

If the auditor consistently follows such a plan, with its predefined decision rule, it is possible to compute the α and β risks he is assuming. If the book value is correct, there is a probability of 1 – CL that the sample estimate plus or minus the precision will not include the book value, resulting in rejection of the book value. Thus α equals 1 – CL. On the other hand, if the book value is in error by an amount equal to the predetermined measure of materiality (M), one half of all sample estimates plus or minus the precision will include the book value, resulting in accepting the book value. Thus β equals .50. In fact, β equals approximately .50 no matter what confidence level the auditor uses.[3]

Use of such a plan may not be optimal for auditing. The auditor uses a lower confidence level when internal control is good. But a lower confidence level implies a higher α risk. Thus the better the controls and the more likely the book value is to be correct, the higher is the auditor's risk of rejecting the book value. Conversely, the auditor's risk of accepting a value which is in error by the amount predefined as material remains a high 50 per cent whether such an error condition is likely (poor control) or unlikely (good control).

A more logical objective of the auditor would be to adjust his β risk for different internal control conditions, not his α risk. Thus, some auditors have compensated by simply letting precision equal one half the amount considered material. (This adjustment generally results in quadrupling the sample size.) This plan has α and β risks that can be calculated also. In fact:

$$a = 1 - CL$$
$$\beta = \alpha/2.$$

While this adjusted approach at least reduces β to a lower level, it still leaves the auditor more likely to reject a book value when the controls are judged to be good. Also, α and β are always fixed in a two-to-one relationship, which may not be desirable. And, finally, in the adjusted approach, use of a confidence level of 95 per cent implies a β of .025, which seems unnecessarily low for typical audit circumstances. If the auditor could justifiably accept a higher value, he could be satisfied with smaller sample sizes.

Variables estimation cannot be used to limit α and β risks without the existence of a predefined decision rule for interpreting sample results. The commonly used decision rules may not result in the best α and β risk levels for auditing.

HYPOTHESIS TESTING

The auditor's objective can be better achieved by statistical hypothesis testing. A statistical hypothesis test is specifically designed to discriminate between two hypotheses with precisely defined α and β risks. In auditing, there fortunately exist two very convenient alternate hypotheses, which permit decisions highly relevant for auditing use. These hypotheses are as follows:

H_0: The financial statement amount is correct.

H_1: The financial statement amount is materially in error.

Now α is the risk of rejecting H_0 if H_0 is true, and β is the risk of accepting H_0 if H_1 is true.

Given three parameters—M (the amount considered material), α and β—the auditor can design a suitable hypothesis test for auditing a financial statement amount.

It is easy to convert the hypothesis test into an equivalent test stated in terms of confidence and precision. The decision rule is that the book value will be accepted if it is included in the confidence interval, but rejected otherwise. The conversion is expressed as follows:

$$CL = 1 - \alpha$$
$$P = M/(1 + Z_\beta / Z_{\alpha/2})^4.$$

[2] The statistical confidence interval is the range of values around the statistical point estimate which is expected to contain the true value with the specified confidence.

[3] For any confidence level over 80 per cent, $\beta = .50$. If the confidence level is less than 80 per cent, β is slightly less. For example, if the confidence level is 70 per cent, $\beta = .48$, and if the confidence level is 60 per cent, $\beta = .45$.

[4] Z_X is the normal table value which includes an area of .5 – X. For example, $Z_{.05} = 1.645$. This formula is an approximation, but it is more than 99.9 per cent accurate as long as α does not exceed .1 and β does not exceed .5.

Thus, if the auditor wants to have no more than a 5 per cent chance of rejecting a given account balance if it is correct, and no more than a 5 per cent chance of accepting the balance if it is in error by $1,000 or more, he should use a confidence level of 95 per cent ($CL = 1 - .05$) and a precision of $544 ($P = 1000/(1 + 1.645/1.96)$).

The advantage of the hypothesis test approach is that it permits the auditor to recognize and control the α and β risk levels. It also provides a correct decision rule for using the statistical results. And, finally, it is stated in the most relevant audit terms.

Use of a hypothesis test, however, requires some guidelines in selecting the parameters to be used. The following three headings deal with selection of α, β and M parameters.

DETERMINATION OF ACCEPTABLE α RISK

The α risk is the probability the auditor will reject correct financial statement balances. However, if the auditor's tests point to rejection, he will usually investigate further to ascertain and correct the causes. There is little likelihood of actually committing such an error. For statistical purposes, therefore, the risk can be considered as the risk that the auditor will unnecessarily be forced to perform follow-up work when an account balance is erroneously rejected. The costs associated with this risk are only those of this unnecessary audit work. There is a theoretical optimum value for α (between 0 and 1) which will minimize total cost, but it is not practically possible to solve for this value. It is probable that the optimum value is relatively low, say .1 or less. Since sample sizes are substantially larger when α is less than .05, a value in the range .05 to .10 appears most reasonable.[5] The most practical approach would be for the auditor to select an acceptable α level and then use it for all statistical tests as a matter of policy.

DETERMINATION OF ACCEPTABLE β RISK

The β risk is of critical importance to the auditor; in fact, minimization of the overall β risk is the reason for the existence of the public accounting profession. Therefore, when a given audit is complete, the auditor wants to have a great deal of assurance that he has not given an unqualified opinion on (that is, accepted) materially incorrect financial statements. If his only source of reliance were his statistical tests, it is clear that a very low β risk would be required. However, there are many other factors in the typical audit. Two of the most critical are discussed below.

Internal control. The first factor the auditor must consider when selecting a β risk for a specific statistical test is the degree of risk that a material error could occur in the area being examined. This risk is inversely related to the quality of internal control in the given area, as good controls will tend to prevent the occurrence of errors and will help detect and correct those which nevertheless occur. When controls are good, the auditor's risk is lower, and he can consequently test at a higher β level.

In judging controls, the auditor need only concern himself with those specific controls to be relied upon in increasing his β risk. Once he decides to rely upon controls, however, he must test them for compliance and effectiveness. Conversely, there is obviously no need for a compliance test when controls are evaluated as nonexistent, because, in this case, no reliance will be placed on the controls when selecting a β level. Furthermore, since the purpose of compliance testing is to permit restriction of other work, it would not be logical to spend more effort on the compliance test than can be saved in the test of financial statement amounts. The auditor should evaluate, before making this decision, which approach will be more efficient. He should remember the purpose of the compliance test and omit it when it is not required or justified.

Even the most effective system of internal control will not prevent deliberate override of the controls by management personnel. Therefore, the auditor must consider the risk of material error through management override of the system of internal control. Although it is impossible to determine with certainty those cases in which management has overridden the internal controls, it should generally be possible to evaluate this risk through consideration of such factors as the type of organization being audited, the susceptibility of the area being examined to misstatement, the requirement for management judgment in determining the amounts in the records and prior experience in auditing the financial statements of the client. Note that the evaluation is not intended to assess the probability that management *is* overriding the controls, but merely whether the area being examined presents any significant potential for override.

For example, if an auditor were testing revenue and expense accounts of a nonprofit organization, he would usually assess the risk of override as very low. If he were testing the allowance for doubtful accounts in a commercial enterprise, however, he might assess the risk as significant, even though he believed that the management of the particular client would not in fact override the internal controls. On the other hand, if he were checking computation of withholding taxes on payroll checks for this second client, he would probably assess the

---

[5] For reasons given under the heading "Audit Adjustments Resulting From Statistical Sampling" (page 53), a value of .05 is recommended.

risk as low, for there would be little potential gain to management in overriding controls in this area.

The evaluation of this risk is not based on the character of the management; it is not a moral judgment and no shadow is cast upon the management by an evaluation that there is a significant risk of override.

Whenever the auditor has assessed the risk of override as significant, he should limit his reliance on internal control, whether or not he has any evidence or reason to believe that management has overridden the controls. Therefore, he should test at a lower β level, or more accurately, he should not increase his β level in reliance upon internal control.

Other audit procedures. The other factor the auditor must consider is the nature and effectiveness of other auditing procedures he may be applying in the areas under examination. If he is performing analytical reviews of the ratios and trends and/or additional detail audit procedures, his reliance on the statistical test is obviously less than it would be in the absence of these other procedures. He can therefore use a higher β level. It is extremely important to note, however, that in increasing his β level in reliance upon other procedures, the auditor should evaluate very carefully any unusual conditions revealed by any of the tests performed. He cannot find an unusual condition in one test and then ignore it because his other tests fail to reveal it. The failure of any single test to reveal a condition of interest is not positive indication that it does not exist; therefore unusual circumstances revealed in any test require further investigation regardless of the outcome of other tests.

For convenience, we can consider two types of other auditing procedures and classify them as to whether they are significantly effective or only moderately effective. A significantly effective additional test would be a test with a relatively high probability of discovering material error conditions which exist (e.g., most detail tests), while a moderately effective test would have a fair probability of discovering material errors (e.g., many well-designed analytical tests). For example, assume the audit test in question were a price test of inventory. The following additional tests would ordinarily be significantly effective: comparison of carrying prices with subsequent sales prices as adjusted to exclude normal gross profit, testing to published price quotations, confirmations with vendors, detailed appraisals, etc. The following tests would ordinarily be moderately effective: analysis of gross profit ratios by product lines, discussions with knowledgeable and reasonably disinterested persons, analysis of standard cost system and variances, etc.

EXHIBIT 1

Selection of β Level for Hypothesis Test

Reliance assigned to internal control

If there is a significant risk that management could override the controls in effect over the area being examined, enter 0.

Otherwise, evaluate the internal controls in effect over the area being examined. ______

If the controls are	*Enter*
Excellent	4
Good	3
Fair	2
Poor	1
Nonexistent	0

Reliance assigned to other audit procedures

Evaluate the other audit procedures which might detect material errors of the type being tested for by the statistical test.

For each significantly effective additional test allow 2 points and for each moderately effective additional test allow 1 point. Enter the total (not to exceed 4 points). ______

Total ______

If the total above is	*Use this β*
0	.05
1	.10
2	.15
3	.30
4	.50
5	.50*
6-8	.50**

*In view of these conditions, the auditor may wish to consider increasing the effectiveness of other auditing procedures and omitting the statistical test.

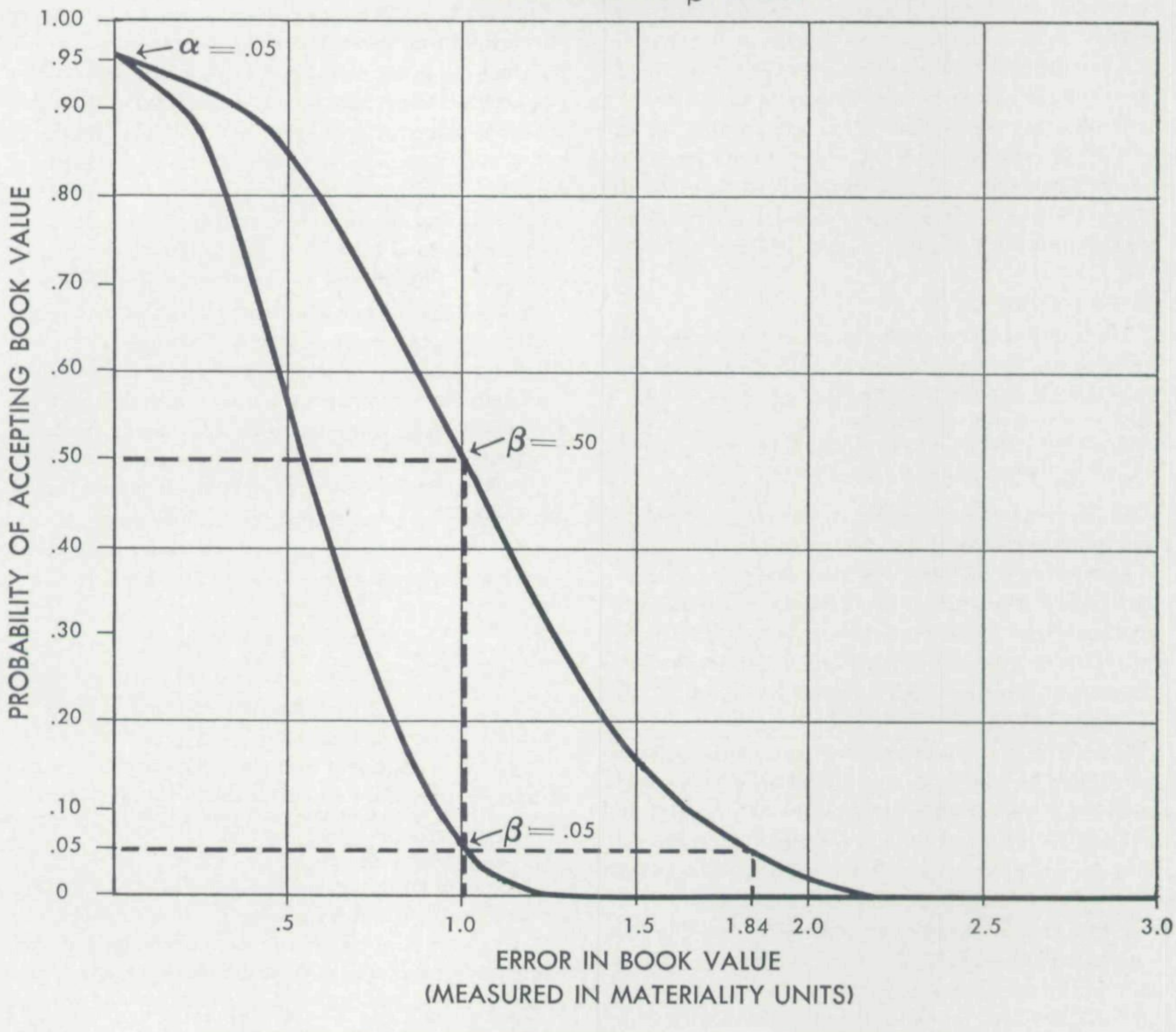
**In view of these conditions, the auditor may wish to consider omitting the statistical test.

Evaluation. After all of these factors have been considered, it is necessary to select a β level. It is clear that when the auditor has nothing else to rely upon, he should use a low β level. In view of the relative importance of any single financial statement amount in an audit and the fact that in most cases there will be no material error to discover, a level of .05 seems reasonable and proper. On the other hand, if internal controls are excellent and other audit procedures are especially strong, there is probably no need to conduct a statistical test at all; this is equivalent to a β of 1.0. However, once the auditor decides to conduct a statistical test, he would probably not want to bother with any test which had less than an even chance of discovering a material error. Therefore, the upper limit on β could reasonably be set at .50, giving a range of .05 to .50.

The next question is how to select a β level in this

FIGURE 1

ACCEPTANCE CHARACTERISTICS
OF HIGHEST AND LOWEST
RECOMMENDED β LEVELS



range. By combining all of the relevant factors mathematically and making some conservative estimates of values, it is possible to compute reasonable β values, which will result in an overall β risk of .05 or less, when considering the statistical test in conjunction with various combinations of other relevant factors. This has been done, and the results are set forth in Exhibit 1, opposite. Although the resulting levels are computed values, it is important to note that they appear quite reasonable when related to the conditions giving rise to them.

In order to demonstrate the value of increasing β to the justifiable maximum, some relative sample sizes follow, based on a certain population with fixed α and M values (the same sample size *relationship* would hold in all cases):

β	*Sample Size*
.05	340
.10	275
.15	235
.30	160
.50	100

The probabilities of accepting book values in error by various amounts for the highest and lowest recommended β levels (assuming a fixed α of .05) are given in Figure 1, above, (which is based on a normal distribution of sample means). It is worth noting that even when the auditor chooses a β of .50, he still has only a small chance (.05) of accepting as correct a book value in error by 1.84 times his measure of materiality and a still smaller chance if the error is greater than that).

The explicit consideration of both of the above factors in the determination of β may go beyond current practice, which seems to concentrate principally upon the evaluation of internal control (notwithstanding the provisions of the 1964 statement of the statistical sampling committee of the American Institute of CPAs, which also relates the level of statistical testing to the other audit procedures). It is clear, however, that other factors may be of equal importance with internal control in many cases. Evaluation of these factors requires audit judgment, but these types of judgment are required for all audit tests, not just statistical ones.[6]

DETERMINATION OF M

The third parameter the auditor must specify is the measure of materiality (M). That measure defines the alternative hypothesis:

H_0: True value = book value.
H_1: True value = book value $\pm$ M.

Some auditors may object to the requirement to specify M *in advance*, but it is conceptually similar to evaluating errors *after* their discovery as to their materiality. The auditor must have some standards for doing this, so he should be able to apply these standards in advance and specify a maximum dollar amount of errors which he would be willing to accept.

Most auditors would welcome some guidance in specifying M, yet there are no generally accepted criteria for quantifying the measure of materiality. Although it is impossible to prescribe definitive criteria in this area, it is at least possible to suggest some general limitations and considerations.

Before considering the measure of materiality for a given audit test, it is necessary to examine materiality for financial statements as a whole. (This will be designated M_0, or overall materiality, and is considered to be the smallest amount by which the financial statements could be in error and still require a qualified auditor's opinion.) M_0 should not be too large, for then financial statements would be insufficiently precise. Conversely, it should not be too small either. Nearly every amount in financial statements includes elements of estimation and uncertainty. In the aggregate, this uncertainty may be quite large. It would not be very logical to concentrate auditing effort on minor refinements in individual accounts which will be swamped by the aggregate uncertainty in the financial statements. Furthermore, sample sizes increase very rapidly as materiality decreases; this mandates use of the largest M_0 consistent with overall audit objectives.

Readers of financial statements are usually most interested in net income (and such related factors as funds generated). It is therefore probable that the auditor will begin his determination of materiality with net income. Since any changes in income will be partially offset by changes in such profit-related charges[7] as income taxes and profit-sharing expense, these should also be considered in determining materiality. (Only rarely would a statistically located difference result in a permanent timing difference for tax purposes.) Therefore, if the auditor considers the measure of materiality to be 5 per cent of net income, he would set M_0 as follows:

$$M_0 = \frac{5\% \text{ of net income}}{1 - \text{marginal rate of income determined charges}^{8}}.$$

Naturally, the resulting figure is subject to the auditor's judgment. In some cases, he may wish to lower or raise the figure, based on other considerations. For example, if the company is in or near a break-even situation, a fixed percentage of income would result in an impractically small measure of materiality; in such a case, the auditor might consider using the company's average net income or a normal return on investment instead of net income. Also, when there are extraordinary items on the income statement, it may often be appropriate to base materiality on income before extraordinary items.

Now, assuming the auditor is satisfied that he has a reasonable value for M_0, the question arises as to

[6] Statement on Auditing Procedure No. 33, Chapter 6, Paragraph 9 states: "The amount and kinds of evidential matter required to support an informed opinion are matters for the auditor to determine in the exercise of his professional judgment after a careful study of the circumstances in the particular case. In making such decisions, he should consider the nature of the item under examination; the materiality of possible errors and irregularities; the degree of risk involved, which is dependent on the adequacy of the internal control and the susceptibility of the given item to conversion, manipulation, or misstatement; and the kinds and competence of evidential matter available."

[7] "Income" and "charges" would be replaced by "loss" and "credits" when appropriate.

[8] The "marginal rate of income determined charges" is the fraction of the next dollar of income which would be taken by income-determined charges such as income taxes and profit-sharing expense. Assume a company's federal tax rate on additional income is 48 per cent and there is a profit-sharing plan which requires 10 per cent of pretax income to be distributed to employees. If there are no other income determined charges, the "marginal rate of income determined charges" is 53.2 per cent (.48 + (1 − .48) (.1)).

how to allocate this figure among the various financial statement accounts. Obviously, he should not set M for any statistical test equal to (or larger than) M_0, for then there would be no margin for error in any other account in the financial statements.

Assume that the materiality (or amount of uncertainty) can be quantified for the audit of every balance sheet account (this is the case if all balance sheet accounts are statistically audited). Let the materiality of the *i*th account be M_i. Also assume the same α and β are used in all tests, and the sampling error in all accounts is independent. Then the following relationship would be required in order for the total sampling error to be less than M_0 with the desired probability:

$$\sqrt{\Sigma M_i^2} \leq M_0.$$

How valid are these assumptions? If the auditor uses a fixed α risk and a statistical β designed to reduce his overall β risk to a predefined level, the α and β risks for all statistically audited accounts *will* be approximately equal. If an independent random sample is drawn for each account audited, then the sampling error for each account *will* be independent. (The fact that there are biases in the data such as deliberate overstatement of assets and understatement of liabilities will not cause any dependence in the sampling error for each account.)

The one assumption which fails is the quantifiability of sampling error in all accounts, because almost certainly never would all accounts be audited statistically. However, the auditor should estimate the total uncertainty remaining in the accounts not audited statistically. He should review the potential problem areas and estimate the probable outside limit of possible error for all these accounts. This total must always be less than M_0 for otherwise he would not have done sufficient work to give an unqualified opinion. Let this estimated total amount be M_{est}. The auditor would normally have a direction in mind for this possible error; therefore, it cannot be considered independent of the statistical sampling errors. If the materiality of the *i*th *statistically* audited account is M_i, the statistical tests must then meet the following criterion:

$$\sqrt{\Sigma M_i^2} \leq M_0 - M_{est}\,.$$ [9]

In determining M_{est}, the auditor should consider the nature of the accounts not statistically audited. If, for example, they were capital stock, additional paid-in capital and long-term debt, not much allowance for error would ordinarily be assumed, because each of these accounts could theoretically be audited to achieve a highly accurate result. On the other hand, if the accounts not audited statistically included receivables, inventories or accounts payable, a rather larger allowance for error would ordinarily be assumed.

Application of these guidelines to determine M is admittedly not easy, but will at least result in a rational determination of M, which will meet overall audit objectives at a reasonable cost.

AUDIT ADJUSTMENTS RESULTING FROM STATISTICAL SAMPLING

The hypothesis test approach to statistical auditing results in a binary decision: accept/reject. If the decision is *accept*, the auditor is finished with the account (if his assumptions about other auditing procedures and internal controls, on which he relied in raising β, are sustained). He accepts the balance as correct within the originally specified risk levels. But if the decision is *reject*, the auditor cannot merely qualify his opinion. He needs to know what amount he can accept as materially correct and then have the client adjust the books to that amount.

In general, the same sample data used to conduct the hypothesis test can be used to make a dollar-value estimate of an account balance. Given the sample data and the amount specified by the auditor as material, it is possible to calculate the confidence that any specified value of the account balance is materially correct.

Under the heading "Determination of Acceptable β Risk" (page 49), it was proposed that the auditor assume no more than a 5 per cent probability of accepting materially erroneous account balances. The statistical β used may have been more than .05 because of reliance upon internal control or other audit procedures. Once the hypothesis test has rejected the account balance (because of errors in the audited balance), the auditor has an indication that whatever he relied upon was ineffective. He should therefore adjust his statistical risks so he has no more than a 5 per cent chance that the estimated book value is materially in error.

One case of estimation is illustrated in Figure 2,

[9] An example will illustrate the procedure. Assume the following:

Accounts to be audited statistically: receivables and inventory

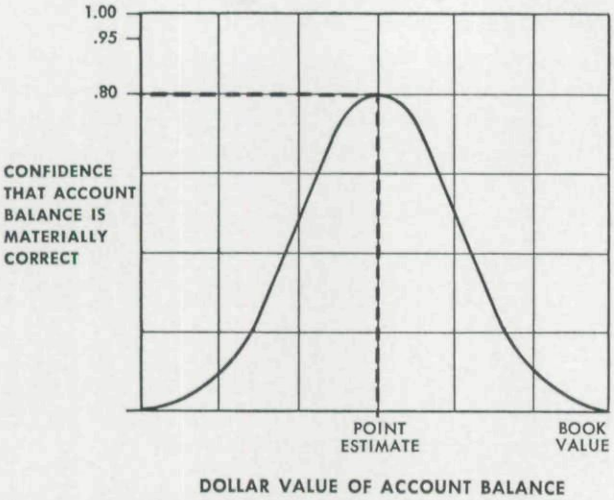
M_0: \$1,000,000 — M_{est}: \$500,000.

Then $M_{receivables}$ of \$300,000 and $M_{inventory}$ of \$400,000 would be a reasonable choice, for

$$\sqrt{300{,}000^2 + 400{,}000^2} \leq 1{,}000{,}000 - 500{,}000.$$

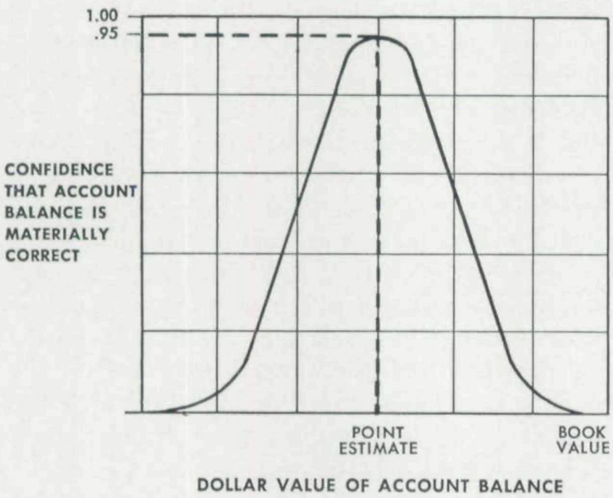
Note: There is an optimum solution to this allocation problem given the variability and relative audit costs of the accounts to be statistically sampled, but that is beyond the scope of this article.

FIGURE 2



Note: figures 2, 3 and 4 are read as follows: for any value on the horizontal axis, the confidence that the true value is within ± M of that point is read off the vertical axis. These figures should not be confused with normal curves.

FIGURE 3



above. In this case, the point estimate[10] is the auditor's best estimate of the correct value, but, because of inadequate sample size, he can be only 80 per cent confident that if the books are adjusted to the estimate, there would be no material error in the account balance. Thus the sample evidence is sufficient to convince the auditor that the book value is wrong, but insufficient to say what is right at a high level of confidence. In this case, the sample is too small to make an estimate at 95 per cent confidence.

Figure 3, above, illustrates a case in which the auditor can adjust to the point estimate with the required confidence. If α was orginally limited to .05 and β to .50, there will always be at least one such point determinable from the sample data.[11] If in addition, α is less than .05 and/or β is less than .50, there will be a range of values, all of which have at least a 95 per cent confidence of being materially correct. This case is illustrated in Figure 4, below. When this happens, there is some latitude for determining the size of adjustment to be proposed. The best estimate of the auditor is that the point estimate is correct, and this point will have the highest confidence of being materially correct. But no point in the range has less than a 95 per cent confidence of being materially correct. Therefore, any adjustment into this range can be accepted by the auditor without qualification.

COMPUTERIZATION OF STATISTICAL TESTING

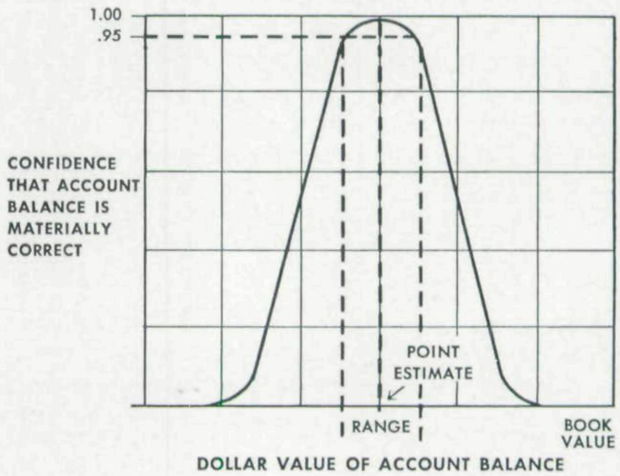
For a number of reasons, it is essential that the statistical auditing approach proposed in this article be computerized.

First, the use of a hypothesis test approach requires shifting parameters. The auditor's originally specified α, β and M, in conjunction with some knowledge of the variability of the population being audited, permit computation of a sample size. When this sample is drawn and audited, the estimated sampling error will always be somewhat different from that expected. This will require adjustment of one or more of the parameters (α, β and M) and, in some cases, additional sampling to bring them

[10] The statistical "point estimate" is the statistically determined estimate of the total value of the variable being estimated. For example, if a 10 per cent sample has a value of $1,000, the statistical point estimate of the total value is $10,000.

[11] This is the reason for recommending an α value of .05 under the "Determination of Acceptable α Risk" heading on page 49. By limiting α to .05 and β to .50, the auditor will have assurance of always being able to make a statistical estimate at a suitably high level of confidence without the need for sampling further.

FIGURE 4



within the required limits.[12] Adjustments of any of the parameters results in adjustment of the rejection point for the hypothesis test. While all of these factors can be dealt with manually, they are complicated enough to cause a high rate of nonsampling error and inconvenience in field work.

Second, the use of a hypothesis test approach will normally require greater precision than the typical audit approach; this means larger sample sizes. To offset the increase in sample sizes it is necessary to obtain the benefit of more efficient statistical procedures, such as optimal stratification and correlation analysis. The computer is invaluable for the more complex computations these procedures require.

And, finally, manual applications of statistical procedures have frequently resulted in various nonsampling errors, such as biased sample design, nonrandom sample selection, unwarranted substitutions in samples and the like. The computer can eliminate or vastly reduce all of these problems. Furthermore, many of the required calculations, if done manually, are so onerous as to discourage the use of statistical techniques.

The requirement for a computer should not be a hardship for any auditor. Most large CPA firms have generalized computer audit packages at their disposal, and all others have access to such a package through the American Institute of CPAs. All that is required is to incorporate the statistical mathematics into these generalized audit programs.

CONCLUSION

The recommended statistical sampling plan has been developed specifically to meet common audit objectives, and in a way which is convenient for the auditor to use. If the system is computerized, the auditor need not deal with formulas, tables, sample selection or unfamiliar statistical terminology.[13] All of this should go a long way toward reducing nonsampling error and auditcrs' reluctance to use statistical techniques, while providing a powerful audit tool.

This sampling plan (or any other) does not replace or reduce the need for audit judgment. If anything, more judgment is required because many audit judgments must now be clearly articulated, and this will nearly always result in their being more carefully considered. The auditor must exercise considerable judgment in all of the following areas: selection of cases in which he will rely on statistical procedures; evaluation of inputs such as materiality, internal control, etc.; evaluation of sample items; and the interpretation of results. All the statistical plan does is to organize the available information into a quantifiable form for decision-making.

[12] Since β and M are critical to the audit but α is not, the most logical approach would be to hold β and M at the specified values and let α vary within some reasonable range. For example, if the target value for α were .05, the test might be considered complete as long as β and M were met and α were no greater than .10. Otherwise, further sampling would be required.

[13] Note that once a hypothesis test with a fixed α is designed, the auditor's only inputs are his evaluations of materiality, strength of internal control and strength of other audit procedures. His output is a statistically supportable decision to accept the book value, or a statement of what book value would be acceptable. None of this requires any familiarity with statistical terminology or procedures.