DHĀRANA
INTERNATIONAL JOURNAL OF BUSINESS FROM BHARATIYA VIDYA BHAVAN'S
M. P. BIRLA INSTITUTE OF MANAGEMENT, BENGALURU
Vol.12, #1 (2018) pp 14-19 ISSN 0974-0082

Application of Statistical Sampling to Audit and Control

T. Srivenkataramana*

Abstract

The traditional literature applying statistical sampling to auditing sometimes overlooks the special structure of audit populations. Much of the literature is based on techniques developed for sample surveys. Of late there is an increasing awareness to take note of the unique environment in which audit sampling takes place and to incorporate all available auxiliary information to improve the precision of estimators. The present paper begins with a brief historical review and then focuses on the special nature of audit populations. This is followed by the description of a class of auxiliary information estimators and the occasional problem caused by situations of low frequency of errors but with large magnitudes. Next, monetary unit sampling is reviewed and key unit sampling is proposed as an alternative when the former may not apply. An outline of a Bayesian formulation to use prior information is provided. Finally, guidelines are provided for a choice of procedure enumerating the major factors to be considered.

Keywords: *Audit and control, Key unit sampling, Monetary unit sampling.*

1. INTRODUCTION

The development of statistical methods in the field of accounting and auditing is instructive of the difficulties in the acceptance of statistical ideas and of problems in adapting standard statistical methodology to special situations. In recent years the accounting profession has turned to statistical techniques to aid in the analysis of financial data. Specifically, statistical sampling has increasing popularity among internal and external auditors. Auditors are expected to ensure effective internal control by creating value to their function. Since complete enumeration is not possible because of vast data, frequently, samples are employed to make this review.

Sampling of accounts is generally concerned with quantitative characteristics. When the auditor examines a set of N accounts, he typically knows the book value $Y_1, Y_2, \ldots\ldots Y_N$ in the population. Denote the total book value by Y. The auditor ascertains, after examining an item, the correct value for this item

** Former Chairman and Professor Emeritus, Department of Statistics, Bangalore University, Bangalore 560 056. Former Director of Canara Bank School of Management Studies, Central College, Bangalore 560 001*

(called the audit value Xi). Clearly if the book amount is correct then Xi =Yi. Thus the purpose is to measure the error: Di = (Yi – Xi).

The objective of sampling here is to infer about the accuracy of the book value. That is whether the total book amount Y is reasonably close to the total audit amount:

$$X = \sum_{1}^{N} Xi$$

2. ATTRIBUTES VS VARIABLES

The first suggestions to use sampling in auditing came in the 1930s (eg.Carman, 1933). Several papers followed in the 1940s. Neter (1949) clearly explained the advantages of these methods. The initial focus was on sampling for attributes. This posed relatively few difficulties. Attributes sampling plans are adequate if the auditor is concerned only with detection of errors and determination of frequency of errors with any prescribed level of confidence variable sampling is necessary. But serious problems arise when we apply standard sampling theory for the case of variables in auditing. For example, if a simple random sample of n items is selected and their audit amounts determined then an unbiased estimator of X is:

$$\hat{X}1 = \frac{N}{n} \sum_{1}^{n} Xi = N\bar{X}$$

Where Xi is the audit amount of the ith sample item and $\bar{x}$ is the sample mean audit amount. This estimator tends to be very imprecise in view of the large variability in many accounting populations. Though stratification by book value reduces the sampling variability there are still several limitations. Comparison of $\hat{X}1$ with the known Y may reveal substantial total error even though no errors were found in the sample.

3. A CLASS OF AUXILIARY INFORMATION ESTIMATORS

Kaplan (1973) has investigated special cases of the following general type of estimator: $\hat{X}_2 = \mathrm{N}\,[\bar{X} + Z\,(\bar{Y} - \bar{y})]$.

Where Z is either a constant or a function of the sample values (Yi x i), i = 1,2,.......n, drawn by simple random sampling. All estimators of the form (3) are consistent in the trivial sense that if the entire populations is audited, $\hat{X}2 = \mathrm{X}$ so that the estimate coincides with the true value. If z is any constant the estimator given by (3) is unbiased. It is easily verified that by setting Z = 0 is (3) we obtain the mean per unit estimator and when $\mathrm{Z} = \bar{Y}/\bar{X}$ we have the classical ratio estimator. Finally, the regression estimator occurs when Z represents the sample regression co-efficient. In this set up, for statistical inference one has to appeal to the central limit theorem and hope that the sample size is large enough for the sample mean to be drawn from a normal distribution. Since the sample standard deviation is also to be estimated from the data, the t-distribution is generally used for statistical tests. Thus, if we wish to construct a confidence interval about $\hat{X}_2$ with confidence co-efficient (1 - ∝), the interval is given by:

$$\hat{X}^2 \pm \mathrm{tk}\,(\propto)\, \mathrm{S}\widehat{Xn}$$

Where, tk (∝) is the value of Student's t with (k-2) degrees of freedom that is exceeded with probability ∝ and $\mathrm{S}\hat{X}$ is the standard error of $\hat{X}_2$.

4. FURTHER USE OF AUXILIARY INFORMATION

Since the book amount Yi are known for all items in the population this information can be gainfully incorporated into estimation strategy. For example, consider the difference estimator. Under simple random sampling one may estimate the total error.

D = (Y- X) by

$$\hat{D}1 = \frac{N}{n}\sum_{1}^{n} d1 = N\bar{d}$$

Where di = (yi – xi) is the difference found in the ith sample item and $\bar{d}$ is the sample mean difference. Alternatively, one may use the ratio estimator.

$$\hat{D}2 = Y\,(\overline{d/\,y}\,)$$

Where $\bar{y}$ is the sample mean for y. These estimators can also be used along with stratification of the population units. Another alternative is to use pps sampling with book value as size measure. Then an unbiased estimator D is:

$$\widehat{D3} = \frac{1}{n} \sum_{i-1}^{n} di\,/(y_i\,/y)$$

$$= \frac{y}{n} \sum_{i-1}^{n} di\,/y_i$$

Assuming with replacement sampling for illustration.

Estimators such as the difference, ratio and regression estimators with simple or stratified random sampling or $\widehat{D3}$ have several serious limitations in the very common situation where the frequency of errors is low, often only 1 or 2%, although potentially great in financial impact, and the variability of the data is estimated from the sample. Where no errors are detected in the sample, that is all sample di = 0, the estimated standard error for these estimators is zero. This may be seen from the estimated variance formula for the difference estimator.

$$V(D) = \frac{N(N-n)}{n(n-1)} \sum_{i=1}^{n} (d_i - \bar{d})^2$$

An estimated standard error of zero leads to the unwarranted conclusion that all the book values in the population are correct! With traditional schemes and estimators this problem is frequently encountered in auditing.

MONETARY UNIT SAMPLING (MUS)

The most appropriate form of sampling will depend, among other things, on the relative liability of items of various sizes to be in error and on the relative costs of checking them. Often, the items with larger values will be relatively more important. Anderson and Teitlebaum (1973) have suggested the use of an individual monetary unit such as rupee or dollar as the sampling unit. The auditor would still examine the account to which the sample rupee belongs but then would prorate the total error for the account to each rupee. Thus, if the jth population rupee (j=1,2…..Y) belongs to the ith (i=1….N) account in the population, its audit value is $\sum_j^1 Xi \,/\, \mathbf{Yi}$ and the error in the jth rupee is Dj = Di / Yi. A simple random sample of n rupee amounts is then selected. If the selections are with replacement the process is equivalent to a PPSWR scheme.

While Neter, John suggest multinomial approach by dividing the population into mixtures of populations which potentially may not serve the purpose where the quality of governance is high. In situations where quality of governance is perceived to be low, binomial approach could serve by dividing the population into two of no error and of potential error. Given the cost of auditing, population over statement error (as has been done by Fienberg, Neter & Leitch) may be attempted in spite of computational complexities.

5. NEEDLE IN HAYSTACK PROBLEM

The situation of most of the accounting units being error free, but with the possibility of a few large errors has been described as above in the literature. MUS tackles this problem effectively only if these errors are associated with large book values. Then the large needle is chopped up into a lot of small needles which occur with sufficient frequency that atleast some of them have a chance of being detected by the sample. On the other hand, if the needles were in audit units with small account balances then MUS might not improve situation. For example, consider the error function p(y) which yields the error rate as a function of the book value of each audit unit. Let

$P_1 (y) = C_1$

$P_2 (y) = C_2 y$

$P_3 (y) = C_3 y\ (1000 - y)^2$

Here MUS will be effective only for the case of P_2 (y). In this setting, we propose below an alternative scheme.

6. KEY UNIT SAMPLING

Suppose the auditor knows, based on previous audits, that the errors are greater among units with certain characteristics, other than higher book balance. Thus, if the rate of errors is known to higher in Location II, one may stratify the population of location and draw a larger sample from Location II. A logical extension of this is key unit sampling. Here the rupee units are replaced by key units and sampling may be done with probability proportional to key units.

As an example, the key units may be the monetary units where the majority of errors are suspected to exist, and audit units themselves in the other cases. If the population of vouchers can be assumed to have a logical arrangement, one operationally convenient method for sample selection is to apply PPS systematic sampling where the key units are the measures of size. Let Z_1, Z_2…. Zn denote these measures. We cumulate the measures of size of the units and assign them the ranges 1 to Z_1, Z_1 + 1 to $Z_1 + Z_2$ and so on. In order to select a sample of n units, a random number r is drawn between 1 and k = Z/n where

$$Z = \sum_{1}^{N} Zi$$

Then the units in the sample are those in whose range lie the random number r and all other numbers r + k, r + 2k,…… obtained by adding k successively to r. if there is any unit whose measure of size is k or larger, it is removed beforehand from the selection procedure and is taken into the sample with certainty. Under this scheme the probability that a unit is included in the sample is $\pi_{ij} = \mathrm{nm_{ij}} / \mathrm{Z}$

Key Unit Sampling has distinct advantages in detecting the needle in the haystack where there is previous information suggesting that errors follow a certain pattern, other than occurring in the larger audit units. It is also suggested as being beneficial when the audit units are not quantified in terms of rupees (currency terms).

7. BAYESIAN FORMULATION & OTHER APPROACHES

Whenever apriroi information is not made available of non-zero accounts, Bayesian approach can be used. Felix & Grimlund (1984) assumed normal distribution of error amounts which is incompatible in situations of non-normality. Cox & Snell (1979) have proposed Bayesian bound approach for MUS by making assumptions of gamma distribution for population error rate 'p' & inverse gamma distribution for assessing mean taint error 'M'. The upper bound has been of MUS under Bayes formulation has been attempted by Moors & Jenssens (1989). Observed number of errors and mean value of errors in money terms is expected to follow Poisson variable in its discrete form.

For a given prior model and a given population this upper bound is a random variable, the outcome of which depends only on the sampling results. The authors describe a general theoretical method to derive the probability distribution of the random variable. However, this theoretical approach is essentially applicable only when the population distribution can be represented by a relatively simple function. For complicated population distributions, as are likely to occur in practical situations, simulation remains an indispensable tool.

8. CHOICE OF PROCEDURE

Loebbecke and Neter (1975) point out that the choice of an appropriate statistical audit sampling procedure depends upon audit objectives, the audit environment and audit procedures. Under certain conditions of each of these factors, the different estimators may perform poorly. Based on an extensive empirical study the authors present a decision flowchart to aid the auditor in selecting the procedures.

When the auditor utilizes a statistical sampling approach as a means of gathering evidence to meet one or a set of audit objectives a careful choice must be made of the procedure. A number of factors must be considered if the results are to ultimately satisfy the audit objectives at hand.

These include the constraints of three types as follows:

1. Audit Objectives:

- Attributes versus variables
- Estimation versus testing
- Combined attributes and variables versus variables alone

Environmental Factors:

- Skewness of book values
- Error rate
- Error magnitude
- Error direction and Computer availability

Characteristics of Audit Procedures:

- Ability to enlarge sample
- Nature of sample frame
- Bias of audit procedure
- Availability of corroborative audit procedures and
- Isolates versus simultaneous audit procedures

Audit Objectives:

The audit objectives have a significant bearing on which statistic is of greatest interest and therefore, on whether the sampling procedure concerns an attribute or a variable. For example, in compliance auditing, the concern is, with the degree to which the accounting control procedures are being applied as prescribed.

Primary interest is in whether or not the procedures are correctly applied and in the extent of cases in which they are not correctly applied. Thus the study characteristic is an attribute. While in substantive auditing the objective pertains to a monetary magnitude and the characteristic of interest is a variable. Combined attributes – variables (CAV) procedures have been formulated to obtain an upper bound on the total monetary error. This upper bound is initially estimated based on attributes and the estimated bound is then modified by variables data to obtain a tighter bound.

When the sample results are assessed by means of a confidence interval which indicates a range within which the population characteristic can be expected to lie an estimation approach is said to be used. On the other hand, when the sample results are assessed by means of a decision rule which leads to one or more alternative decisions a testing approach is applied. There is generally a direct link between these two approaches and the distinction is only in the uses. In the testing approach a decision is made based on the sample while in the estimation approach, information about the magnitude of the study variate is obtained without leading directly to a decision.

Environmental Factors:

There are atleast four important environmental characteristics which affect the choice of a sampling procedure for variables or combined attributes – variables: skewness of the population, error rate, and magnitude of errors and direction of errors. These factors affect the behavior of a sample statistic in two main ways: the precision of the estimator and reliability. If the auditor knows in advance the exact nature of these environmental factors, the problem of concern would be solely that of finding a sampling procedure which needs the stated audit objectives and provides adequate precision at an acceptable confidence level. As an example, the auditor dealing with a highly positively skew population and an estimator which is known to be inefficient in such a situation may examine all items on the right tail.

Characteristics of Audit Procedures

There are several situations where the auditor may like to enlarge the sample for instance, in acceptance samplings the testing may be done in two or more stages. Or if the population turns out to be highly skew the auditor may wish to enlarge the sample so that use of normal distribution for constructing a confidence interval is appropriate. Also the nature of sampling frame has an important effect on the choice of the procedure. This determination of optimal strata boundaries based on the book values can be made for stratified random sampling of accounts. An audit procedure is said to be biased when the inherent errors are persistent in their effect. The errors may be introduced by persons other than the auditor too. For example, a customer may not respond at all. Availability of corroborative procedures may be of help when the distributional assumptions are not met fully. Finally, it is to be noted that the auditor may use several audit procedures simultaneously in order to obtain evidence about a set of objectives in an audit area. An illustration is provided by the examination of inventory balances for verification of accurate quantities, correct costs and proper extensions. The situation is quite complex when several audit procedures are used simultaneously. If one could model this situation, taking into account the areas of overlap which exist between the procedures as well as the areas of uniqueness, it might be feasible to develop a test for deciding whether or not the book value is reasonable, such that the test embodies known risks of making incorrect decisions.

CONCLUSION

Audit and control present a special area of application of statistical sampling. Due to the mixture nature of audit populations, direct use of traditional sampling may not yield reliable results. Novel techniques like key unit sampling which suitably incorporate corroborative evidence will be helpful in this context. The choice of the sampling procedure must also consider the objectives, and the audit environment. Auditing forms an area of increasing interest to the statistical samplers.

References:

- Anderson, R and Teitlebaum, A.D (1973). Dollar-Unit Sampling C.A.Magazine, 102(4), 30-39.
- Carman, L.A. (1933). The Efficacy of Tests, American Accountant, 18, 360-66.

- Cos, D.R. and Snell, E.J. (1979). On Sampling and the Estimation of Rare Errors. Biometrika, 66, 125-32.
- Felix, W.I and Grimlund, R.A (1984). Dollar Unit Sampling for Accounts Receivable and Inventory. Accounting Review, 59, 218-41.
- Fienberg, S.,Neter, J and Leitch, R.A (1977). Estimating the Total over – statement Error in Accounting Populations. Journal of American Statistical Association, 72, 295-302.
- Kaplan, R.S (1973), Statistical Sampling in Auditing with Auxiliary Information Estimators. Journal of Accounting Research, 11(2), 238-58.
- Loebbecke, J.K and Neter, J (1975). Considerations in choosing Statistical Sampling procedures in Auditing. Journal of Accounting Research, 13(1), 38-97.
- Moors, J J A and Janssens, MJBT (1989). Exact distributions of Bayesian Cox-Snell Bounds in Auditing. Journal of Accounting Research, 27(1), 135-44.
- Neter.J (1949). An Investigation of the usefulness of Statistical Sampling Methods in Auditing. Journal of Accountancy, 87, 390-98.