An Evaluation of the Negative Binomial Method for Tests of Controls

Dr. John P. Wendell, Accountancy, University of Hawaii at Manoa
Dr. Terry Gregson, Accountancy, University of Hawaii at Manoa

Abstract

Auditors often use statistical sampling techniques to test controls. Dhavale (1991) presents a new statistical technique based on the negative binomial distribution for this purpose. This article examines in detail the properties of the new method. The article also provides the computational details necessary to apply the new method. The article concludes that the new method may be useful to auditors in some circumstances.

Introduction

Auditors commonly use statistical sampling methods to test controls while performing a financial statement audit. The most commonly used method for this purpose is to evaluate a simple random sample using the binomial distribution. Dhavale (1991) presents a new method for this purpose based upon evaluating a cluster sample using the negative binomial distribution. This article examines the computational details and properties of this negative binomial method (NBM) from the perspective of established statistical sampling principles.

Statistical Sampling

Statistical sampling methods involve two steps: (1) the selection of the sample and (2) the evaluation of the sample. Dhavale (1991) considers the use of three different sample selection methods for tests of controls. In *simple random sampling* every item in the population has an equal chance of being included in the sample. The sample items are selected independently from each other. In *stratified random sampling* the population is divided into groups called strata and each item within a strata has an equal chance of being included in the sample. Items in different strata will have a different probability of being included in the sample. A stratified sampling plan ensures that a predetermined number of items are included in the sample from each strata. In *cluster sampling* items in a population are divided into groups called clusters. A simple random sample of clusters is then taken and every item in the cluster is included in the sample.[1] Cochran (1977) provides a very comprehensive discussion of these sample selection methods and Arkin (1984) discusses the methods from an auditing perspective.

Sample evaluation procedures are designed to be used with particular sample selection procedures. A sample evaluation procedure should not be used with a selection procedure for which it was not designed. This is because classical statistical inference is based on sampling distributions. A sampling distribution is a probability distribution for observing different sample outcomes in repeated samples taken from a particular population. The probability of a particular sample outcome will depend upon the sampling procedure used. Different sample selection methods usually will not have the same sampling distribution. It cannot, therefore, be expected that a sample evaluation method that is used with a particular sample selection method will work with another sample selection method. For example, a sample evaluation method based on simple random sampling should not be used with a stratified random sample or a cluster sample.[2]

Sampling With a Variable Error Rate

Dhavale (1991) takes the position that cluster sampling allows for the calculation of more accurate bounds on the error rate in a population than simple or stratified random sampling when the error rate varies between clusters. It is true that cluster samples will provide a better estimate of variability of the population error rate between clusters than either simple or stratified random samples (Dhavale 1991, footnote 4). However, it is incorrect to assert that an accurate estimate of the variability between clusters is necessary to accurately put an upper bound on the error rate in the population.

The variability between clusters is a characteristic of the population and not necessarily of the sampling distribution. For simple and stratified random sampling the variability between the clusters has no effect whatsoever on the sampling distribution of the number of errors included in a sample. In simple random sampling the sampling distribution is completely determined by the number of errors in the population. The order in which the errors occur in the population makes no difference because the sample is selected at random. If sampling is with replacement the binomial distribution provides an exact model of this sampling distribution and can be used to accurately calculate an upper confidence bound on the error rate in a population.[3] If sampling is without replacement the hypergeometric distribution provides an exact model for the sampling distribution and the binomial will provide a close approximation.[4]

Under stratified random sampling the sampling distribution is determined completely by the number of errors in each strata in the population. Variability between clusters within each strata is irrelevant to the sampling distribution. Wendell and Schmee (1996) provide an exact method for evaluating stratified samples.

For the above reasons, it is not necessary to use cluster sampling when there is a variable error rate in order to achieve accurate bounds. However, there are many good reasons why an auditor would want to use cluster sampling for tests of controls in some cases. Here are just a few, see Cochran (1977) and Arkin (1984) for others: (1) The sampling frame may not allow easy selection of a simple or stratified random sample. (2) Cluster sampling may cost much less per unit to sample. This can make cluster sampling more efficient even though it often requires larger sample sizes. (3) The variability of the error rate may be of interest to the auditor. Widely fluctuating error rates between clusters of transactions could indicate fraud or other serious problems with the accounting system. This fluctuation in error rate could be missed with a simple or stratified random sample.

The standard methods for evaluating cluster samples for tests of controls rely on a normal approximation (Cochran 1977) that may not be reliable with the small error rates and moderate sample sizes encountered in auditing. The NBM is an alternative to the standard methods.

How to Calculate the NBM Bounds

Dhavale (1991) does not provide the details necessary to calculate the NBM bounds. The method that is presented here is based on the method used by the software provided to us by Dhavale (Dhavale 1991, footnote 6). This method calculates bounds that agree with the bounds presented in the tables of Dhavale (1991).

The NBM calculated an upper bound (when a bound could be calculated) that was greater than or equal to the population error rate 81.8%, 93.3%, 96.7%, and 100% of the time at a nominal confidence level of 80%, 90%, 95% and 99% respectively. This indicates that the NBM is slightly conservative at these confidence levels. Many more simulations at different values of m and a and different sample and cluster sizes must be done before the properties of the NBM bound can be considered to be well known when the population has a compound Poisson distribution. Simulations also need to be done with other plausible distributions to determine to what extent the NBM is robust to deviations of the population from the compound Poisson distribution.

Other Considerations for Using the NBM

Counterintuitive Behavior

The upper bound on the error rate calculated using the NBM can decrease as additional errors are found in the sample and increase when fewer errors are found. This counterintuitive property of the NBM is possible because the calculation of the bound depends upon both $\hat{m}$ and $\hat{a}$. In some cases a change in $\hat{a}$ will result in a larger bound even when $\hat{m}$ decreases. For example, consider the following scenario: (1) The tolerable rate of non-compliance is 10%; (2) The planned risk for overreliance on controls is 5%; (3) A sample of 25 clusters of ten vouchers each is tested for compliance with the following result: (3a) 23 clusters contain vouchers in compliance with controls; (3b) One cluster contains six vouchers in non-compliance; and (3c) One cluster has nine vouchers in compliance and one voucher that was not tested for compliance because the client was unable to locate it. In accordance with AU § 350.40 (AICPA, 1996) the missing voucher is considered in non-compliance.

The upper bound for this sample using the NBM is an error rate of 8.8%. On the basis of this bound the auditor decides to rely on internal controls and reduces substantive testing. Several weeks later the client finds the voucher behind a file cabinet drawer and gives it to the auditor. The voucher is found to be in compliance. The upper bound using the NBM is now 10.4% and the auditor can no longer rely on the controls. The auditor is now faced with having to tell the client that substantive testing will have to be increased because the voucher was found to be *in* compliance with controls.

Sample Size

The NBM requires a larger sample size than simple random sampling. The NBM may still be more efficient if the cost per unit for the cluster sample is low enough; but, in situations where the per unit sampling cost is similar for all sampling plans other methods are probably more efficient.

Dhavale (1991, footnote 5) advocates determining the sample size given the cluster size by first estimating m and a and then calculating the sample size using a formula based on a normal approximation to the negative binomial. The normal approximation is not necessary. Instead, equation 4 can be used directly to determine N. Consider the following example: (1) The planned risk of overreliance is 10%; (2) The tolerable rate of non-compliance is 10%; (3) The cluster size is 10 and the auditor expects that the actual error rate is 5%. This results in a value for $\hat{m}$ of .5 (10 × .05); and (4) Based on past experience or a pilot sample $\hat{a}$ is set to 0.3.

Under these circumstances X/NK in equation (4) will be lower than 10% only when $N \geq 17$. This makes the required sample size 170 (17 × 10). In contrast, the binomial sample size (for a simple random sample) under the same conditions is 80 (Guy and Carmichael 1986, Table 1, page 52). The NBM would be cost effective in these circumstances if the cost of auditing one item in a cluster was less than 47% (80 ÷ 170) of the cost of auditing one item in the simple random sample.

Equal Cluster Size

The NBM method as formulated can

that have a solve function. It can also be solved with many scientific calculators that have a solve function. It can also be solved reasonably quickly by hand using trial and error methods. A good first guess for $\hat{a}$ is

$$\hat{a} \cong \frac{\hat{m}^2}{\hat{\sigma}^2 - \hat{m}} \tag{8}$$

In the example, equation (8) gives an initial estimate of $\hat{a}$ of approximately 0.25. Substituting 0.25 for x in equation (7) yields a value of -1.327. A negative value indicates that the estimate was too high so a lower estimate would be tried next, perhaps 0.125. This gives a value of 6.774, indicating that 0.125 is too small. It is now known that $\hat{a}$ is between 0.125 and 0.25. By repeating this process $\hat{a}$ will eventually be found to be 0.2305.[5]

Step 3 - Calculate the bound from $\hat{m}$, $\hat{a}$, N, and K using the following substeps:

Sub-step a - Find the minimum positive integer value of X for which equation (4a) holds. This can be easily done by starting with $X = 0$ and increasing X one integer at a time until the first value of X for which the inequality holds is found. In the example with $X = 18$ the calculated risk is 0.052 and with $X = 19$ it is 0.038, so 19 is the minimum value for X for which the inequality holds.

Sub-step b - Finally, substitute X, N, and K into equation (4). For the example, this yields $19/(40 \times 10) = 0.0475$. This agrees with the bound given in Dhavale (1991, Table 3).

End Notes

1 In multi-stage sampling a simple random sample of items in a cluster is included in the sample instead of all of the items in a cluster.

2 To illustrate, consider this example: in a population of invoices there are 100 batches (clusters) of 20 invoices each, a total of 2,000 invoices. Four clusters have incorrect extensions, i.e. each invoice in four clusters is in error, all other clusters are correctly extended. The population error rate is 4%, which is also the cluster error rate. The auditor selects at random without replacement four *clusters* (80 invoices). There is an 85% chance that the sample will not contain an error. The binomial upper bound for a *simple random sample* of 80 and a confidence level of 95% is 3.7%. Thus, the 95% upper bound will provide coverage 15% of the time, i.e. it is actually a 15% confidence bound and not a 95% bound.

3 Sampling with replacement means that an item in the population can be included in a sample more than once. Sampling without replacement means that an item in the population can only be included in a sample once.

4 See Arkin (1984) for a detailed discussion of the appropriateness of the binomial as an approximation to the hypergeometric when sampling without replacement with various population and sample sizes.

5 There are more efficient search strategies than merely guessing. For a detailed discussion of various iterative methods that could be applied to this problem see Traub (1982).

References

1. American Institute of Certified Public Accountants, *Codification of Statements on Auditing Standards,* AICPA, New York, 1996.
2. Anscombe, F. J., "Sampling Theory of the Negative Binomial and Logarithmic Series Distributions,'' *Biometrika,* Vol. 39, pp. 358-382, 1950.
3. Arkin, H., *Handbook of Sampling for Auditing and Accounting* third ed., McGraw-Hill, New York, 1984.
4. Cochran, W. G., *Sampling Techniques* third ed., Wiley, 1977).
5. Dhavale, D. G., "A Note on a Variable Error Rate Model in Compliance Testing,'' *Auditing: A Journal of Practice and Theory* (Spring 1991): 159-166.

6. Guy, D. M. and D. R., Carmichael, *Audit Sampling* second ed., Wiley, New York, 1986.
7. Traub, J. F., *Iterative Methods* second ed., Chelsea, New York, 1982.
8. Wendell, J. P. and J. Schmee, "Exact Inference for Proportions from a Stratified Finite Population," *Journal of the American Statistical Association,* Vol. 91, No. 434, pp. 825-830, June 1996.