# APPLICATION OF DIGITAL ANALYSIS IN THE AUDIT

Paul Ashcroft, Assistant Professor
Nicholls State University
Department of Accounting and Business Law
P. O. Box 2015
Thibodaux, LA 70310
Telephone: 985-448-4215
Fax: 985-448-4922
Email: acbl-paa@nicholls.edu

Benjamin Bae, Assistant Professor.
Department of Accounting
School of Business
Virginia Commonwealth University
Richmond, VA 23284
Phone: (804) 828-7137
FAX: (804) 828-1719
Email: brbae@vcu.edu

Jeanell Norvell
Business Consultant
Montgomery, Alabama

March 27, 2002

Please address all correspondence to Benjamin Bae at the above address.

**Topic area: Auditing**

# APPLICATION OF DIGITAL ANALYSIS IN THE AUDIT


## ABSTRACT

Performing adequate procedures to detect errors, misstatements, and potential fraud in client's financial statements is extremely important in the work of external auditors. A relatively new technique known as digital analysis can significantly improve the auditor's effectiveness and efficiency in discovering such events. Digital analysis is based on mathematical probabilities of the occurrence of individual numbers in a population of dollar amounts. These probabilities are collectively known as Benford's Law and were developed from actual data. Significant advancements in computer technology in recent years have provided auditors with the tools necessary to use digital analysis to audit large amounts of data. This paper explains how digital analysis combined with the benchmarks established by Benford's Law can be used to increase the likelihood that an auditor will detect inefficiencies, fraud, misstatements, and errors. Examples of fraud and error discovered with digital analysis are illustrated.



**Key Words**: digital analysis, fraud, Benford's Law

## I. Introduction

Auditors typically perform various analytical procedures such as ratio analysis and trend analysis in order to discover unusual or unexpected fluctuations. By identifying unexpected patterns in financial items or other data, the auditor can then increase substantive testing of those areas and perform a more effective audit. The reason is that unusual changes in data indicate a greater chance of misstatement. This paper discusses the technique of digital analysis, a relatively new and very efficient analytical procedure that improves the ability of the auditor to identify potentially misstated, fraudulent, and erroneous items. Digital analysis combined with the mathematical foundation of Benford's Law provides auditors a method to thoroughly and efficiently evaluate large amounts of data.

Businesses have significantly increased the use of electronic data processing in the last ten to twenty years, and as such computers and networks now have a vital role in generating and recording the information in company's accounting records and financial statements. The efficiency of processing transactions and other information with computers has resulted in substantial increases in the amount of data available for auditors to review.

The large volumes of data that companies generate today with computer processing present auditors with important challenges. One challenge is that a greater volume of transactions means that it may be more difficult for the auditor to uncover misstatements or fraud. Another challenge is that auditors must test and evaluate larger quantities of data, which can be overwhelming. For these reasons, it has become increasingly important for auditors to employ cost-effective tools for selecting and analyzing audit evidence while maintaining a high-quality level in performing the audit work. Digital analysis is one such tool. Digital analysis combined with benchmarks set by Benford's Law is being used by at least two of the Big Five

auditing firms and by many Fortune 500 companies to increase the likelihood of discovering the presence of fraud, processing inefficiencies, bias, and errors (Lanza 2000).

**Digital Analysis and Benford's Law**

Digital analysis involves examining the digits in each individual amount in a set of numbers to determine if the frequency of the digits is reasonable and is thus likely to be correct. If the digits occur more frequently or less frequently than expected, the amounts are much more likely to be false or erroneous. Digital analysis is an analytical procedure that examines the patterns followed by specific digits in a group of amounts or numbers. The purpose of digital analysis is to detect abnormal occurrences of individual digits, combinations of digits, or of specific numbers.

The underlying concept of digital analysis is that abnormal patterns in a set of numbers may point to the existence of fraud, inefficiency or error in the data being reviewed. For auditors to effectively use digital analysis they need a benchmark indicating the frequencies that certain digits and combinations of digits are expected to occur. That is, after the auditor determines the frequencies that certain digits occur in an actual dataset of a client a standard is needed to evaluate if the observed frequencies are reasonable. Benford's Law provides this benchmark of comparison by defining the expected frequencies and thus allows the auditor to interpret their actual results (Nigrini 1998).

The benchmarks provided by Benford's Law were developed from the work of Frank Benford, who was employed as a physicist at GE Research Laboratories in the 1920's. Mr. Benford frequently used books containing tables of logarithmic numbers. He noticed that the beginning pages of his logbooks were more worn than those at the end. He realized that he was using the first few pages more often than the other pages and considered that perhaps this was

simply because in most situations there would be more numbers with low first digits than numbers with high first digits (Nigrini 1999a). A first digit is the initial number in an amount – for example, the first digit of 3,258 is 3.

Benford then examined 20,229 different numbers and noted the actual frequencies of each number 1 to 9 being the first digit. He obtained his set of numbers from various sources that included geographic, scientific, and demographic data, as well as all of the numbers in one issue of *Reader's Digest*. This analysis revealed that the number 1 was the first digit about 30% of the time, which was the most often of all numbers. The number 2 occurred next most often, at about 18%. Each successive number occurred less often than the prior one and the number 9 had a frequency of only about 5%. Benford then estimated the expected frequencies of the second through fourth digits by using integral calculus and principles from physics. He also used integral calculus to determine the expected frequencies of the various digit combinations (Nigrini 1999a). An example of a digit combination is the probability of the first two digits in a number being a 1 and a 7 in that order. These frequencies for digits and digit combinations became known as Benford's Law.

Benford's Law gives the auditor a benchmark or standard against which actual client data can be tested for possible errors, fraud, and other misstatements. Thus the auditor identifies abnormal digits and numbers in the client's data by noting which items did not follow the normal frequency level that Benford's Law specifies should occur. Table 1 presents the frequency benchmarks established by Benford's Law for the first four digits.

Insert Table 1 about here

It is important to note that not every set of numbers follows the expected frequencies of Benford's Law. The auditor must know the type of data being examined before deciding to use Benford's Law in assessing the reasonableness of the data. The data must meet the following requirements for Benford's Law to apply:

1. The numbers must be random rather than have assigned or predetermined values. Items such as checking account numbers, drivers license numbers, and Social Security numbers do not conform to Benford's Law.
2. The data do not have set maximum values (such as a limit on the contributions to a pension plan) or minimum values (such as a minimum balance to maintain a savings account).
3. Numbers should not occur at specified break points, such as a standard mailing rate of $12.00 for all packages up to two pounds.
4. The numbers describe the size of items. Good examples would be sales figures, payroll amounts, and payments by customers. Examples of items not qualifying are zip codes, vendor numbers, and insurance policy numbers.

These criteria simply mean that some populations of numbers have logical reasons for certain digits occurring more frequently than the predictions of Benford's Law. For example, if the minimum balance to maintain a savings account is $500, then the first digit (5) or the first two digits (50) will likely occur more frequently in the population of all savings account balances than Benford's Law predicts.

Sets of numbers that meet the four requirements above can be tested using digital analysis based on Benford's Law. Financial data such as corporate sales revenues and cash disbursements, income taxes and stock exchange data follow the digit patterns included in

Benford's Law (Nigrini 1999a). If the numbers in these and similar financial databases are valid and accurate, then it is reasonable to expect that the digits will occur in approximately the predicted frequencies. Data that does not follow the predicted percentages likely include fraudulent, misstated, or otherwise erroneous amounts. The larger the differences between the actual frequencies of digits in real data are from the expected frequencies, the greater the probability that the amounts contain errors.

Audit software programs are available to efficiently perform digital analysis. The software compares the amounts in the actual dataset against the benchmarks established by Benford's Law and identifies for further analysis all items that are beyond the benchmarks. The comparison identifies specific patterns in digits and numbers, abnormal duplications of numbers, and the occurrence of round numbers (Lanza 2000). Digital analysis is very useful in discovering fraud and misstatements because numbers that have been altered, falsified, or contain errors are not usually chosen randomly. Thus, numbers that are invented or have been misstated will probably not conform to Benford's Law. Digital analysis is an effective technique because in attempting to perpetrate and conceal fraud or other misstatements, people often repeat the use of the same numbers rather than selecting purely random amounts. When non-random numbers are used, the actual frequencies of individual digits and digit combinations will likely differ from the expected frequencies stated in Benford's Law. The greater the frequencies of the actual data deviate from the established benchmark the more likely there are fraud, processing inefficiencies, bias, and/or errors present.

Using computer software increases the auditor's chance of detecting fraud or misstatements by significantly reducing the time needed to review data and since 100% of the suspect data is examined rather than selecting samples of items from a population. Before

choosing datasets to test with digital analysis, it is wise to perform preliminary analytical review procedures such as ratio analysis, trend analysis, and comparison to industry averages. These preliminary analytical review procedures will identify components of the population that are likely to contain errors or misstatements and are most appropriate for testing with digital analysis. In other words, it may not be cost effective to apply digital analysis to all the data populations present in a client's files. Preliminary analytical procedures greatly support the auditor's professional judgment in deciding which datasets should be tested with digital analysis.

Some items that are most appropriate for testing with digital analysis are sales, purchases and accounts payable, payments by check, employee reimbursable expenses such as travel and meals, inventory unit costs, and searching for duplicate payments. Applying digital analysis to these and similar items allows the auditor to efficiently test for fraud, processing inefficiencies (e.g. abnormal duplication of invoice amounts in accounts payable), bias (e.g. excessive expenses or purchases just below cut off limits), and errors (e.g. overstatement of inventory costs).

## II. Digital Analysis Application and Examples

### Overview of Expected Frequencies

The expected frequency distribution of the first digit in a set of numbers is provided by the formula:

$$\text{Expected First Digit Frequency} = \log 10(1 + 1/\text{First Digit})$$

Remember, the first digit can only be between 1 and 9. Thus, the expected frequency of the first digit being a 1 is log 10(1 + 1/1) = log 10(2) = 0.3010 and for the second digit being a 2 is log 10

(1 + 1/2) = log 10(1.5) = 0.1761 and so forth. These are the frequencies given in Table1 for the first digit. The frequencies for the second through forth digits were estimated by Frank Benford based on integral calculus and physics.

Another useful analysis is to evaluate the frequency of the first two digits in a set of numbers. The following formula provides the expected frequency of the first two digits:

Expected First Two Digits Frequency = log 10(1 + 1/First Two Digits)

For example, the first-two digits of 1,958 are 19. The expected frequency of 19 being the first two digits in a set of numbers is log 10(1 + 1/19) = log 10(1 + 0.0526) = 0.0223. The expected frequencies start at 0.0414 for 10 and gradually decrease to 0.0044 for 99 being the first two digits.

**Performing Digit Tests**

The initial step in testing a set of numbers is to perform the first-digit and second-digit tests. These tests compare the actual frequencies of the individual first digit of the data and the individual second digit of the data to the established benchmarks shown in Table 1. It is important to note that these tests are meant merely to test the reasonableness of the data. No final conclusions should be drawn from the first-digit analysis or the second-digit analysis. In other words, an auditor cannot assume that there is a problem just because the data fails to meet the first-digit or the second-digit test. Instead, significant anomalies revealed by these tests indicate the need for further examination by the auditor. Equally important to understand is that if the data passes the reasonableness of the first-digit and the second-digit analysis, this does not mean that all the numbers are definitely correct. It only means any errors or misstatements are not significant enough to disrupt the digit patterns, and that examining a normal-sized sample

with regular statistical sampling and substantive testing procedures should be sufficient for testing such data (Nigrini 1998).

If the first and second digits of the data do not reasonably fit the expected frequencies of Benford's Law, then the next step is to perform the first two-digit test. The first two combined digits of each number in the chosen dataset are compared to the expected frequencies given by the log formula for the first-two digits stated above. Items that are outside of the expected frequencies are identified for further analysis. For example, if the auditor found that 0.0215 of purchases began with the two-digit combination of 99, the auditor would likely test a much larger sample of purchase transactions that began with 99 than they would have otherwise. This is because the observed frequency of 0.0215 (or 215 out of 10,000 items) is much higher than the expected frequency of 0.0044 (or only 44 items out of 10,000). Here, the auditor might discover questionable purchases that are made below an established minimum (such as $1,000 or $10,000) that would require management approval.

Just as the combined first two digits can be tested for anomalies the last two digits can be tested as well. The actual frequency of each last-two-digits combination is compared to its expected frequency to perform this analysis. Since the possible last-two digit combinations are 00 to 99 consecutively, Benford's Law states that the expected proportion for each of the possible 100 last-two digit combinations is .01 (Nigrini 1999b).

In performing each of these digit tests, it is important for the auditor to continually apply professional judgment and determine other necessary analytical procedures needed. Auditors should look carefully at the abnormal data and examine supporting evidence to determine why the discrepancies are present.

**Uncovering Processing Inefficiencies**

In addition to indicating data that likely contain misstatements, fraud or errors, digital analysis is useful for identifying areas of inefficient transaction processing. For example, 50,000 actual payments were examined by using digital analysis. Any of the digits 1 to 9 could appear as the first number in each payment, as there were no restrictions on the payment amount. The first-digit test revealed that the numbers 1 and 2 each appeared more often than the expected frequencies of Benford's Law. This indicated that a problem may exist and that further analysis was necessary. The payments were then examined by using the first two-digit test. The first two-digit test indicated to the auditors that the numbers 10, 18, and 25 substantially exceeded the established benchmarks. Next, the auditors selected all items in the population of payments that began with 10,18,and 25. The following schedule summarizes these subsets of payments (Lanza 2000):

| Amounts starting with "10, 18, & 25" | Number of Payments | Reason |
|---|---|---|
| 100 | 1,872 | Employee advances |
| 1,000 | 1,220 | Royalty payments/ advances |
| 1,800 | 717 | Royalty payments |
| 250 | 645 | Employee advances |

The auditors examined the invoices related to the advances and royalty payments and discovered that the company was issuing many more checks than necessary. Very many payments began with the digits of 10, 18, and 25 because the company often issued small advances to employees and made weekly royalty payments to software vendors despite the fact that contracts with vendors allowed for monthly payments. Reducing the number of payments would improve processing efficiencies and save the company money. Far fewer checks could be

issued by providing company credit cards to employees who traveled regularly and by paying all software vendors monthly. When the first two-digits for low-value numbers (such as 10 to 40) occur more often than expected, then processing inefficiencies are likely to be present.

**Uncovering Fraud**

A key part of every audit is searching for fraudulent items that are potentially included in the client's financial statements. Discovering fraud that has been created by client management or employees is often a very difficult task for the auditor. One reason for this is that the auditor usually samples evidence rather than testing entire populations of available data. Another reason is that fraud often consists of misstatements or thefts of small dollar amounts that are perpetrated on a regular basis and may occur hundreds or even thousands of times in a year and thus would not stand out as glaringly fraudulent as would one or two very large misstatements discovered by the auditors.

Digital analysis significantly increases the auditor's effectiveness and efficiency in searching for fraud. Digital analysis improves audit efficiency by allowing the auditor to easily identify groups of transactions that are more likely than average to be fraudulent or misstated. This means that the auditor can test large amounts of data, even entire populations, using digital analysis to identify sets of transactions or amounts to be more closely examined for possible fraud or misstatement. Thus, rather than initially sampling a limited number of transactions at random and testing them for fraud, the auditor can more efficiently and effectively examine possible fraud by using digital analysis to separate out items in the population that have an above-normal probability of being erroneous or fraudulent.

Auditors investigating the possibility of fraud used the first two-digit method and the benchmarks of Benford's Law to analyze the dollar amounts of contracts issued by a client who issued thousands of contracts each month. They found that the first two-digit combination of 49 appeared in the contract amounts much more often than expected. The client required a bidding process for all contracts over $50,000 but those under $50,000 could be negotiated with only one supplier. The auditors discovered that the cause of the higher than expected proportion of contracts beginning with 49 was that a contract administrator was avoiding the controls by establishing contract values between $49,000 and $49,999. The contract administrator was perpetrating fraud by entering into these contracts with an enterprise owned by his wife (Coderre 1999).

Examining the last two-digits in amounts is another effective digital analysis technique to find fraud. The last two-digit test revealed fraud perpetrated by a restaurant by examining daily sales amounts of an entire year. The restaurant owner created false amounts of each day's sales to intentionally understate revenues and thus avoid sales and income taxes (Nigrini 1999b). The last two digits of the restaurant's daily sales were compared to the benchmarks established by Benford's Law. The last two digits of 40 occurred 5.6 percent more often than the benchmark and the last two digits of 87 were over the benchmark by 2.5 percent. The owner was subconsciously repeating these numbers in the fraudulent data.

**Uncovering Bias**

Amounts that are biased are items that are incorrectly stated in a similar pattern on a regular basis. In other words, the recorded amounts are not totally falsified but are rather slightly

to moderately skewed from the actual amounts. Reimbursements are an excellent item to test for bias, as claims often tend to be higher than actual costs incurred.

In auditing employee reimbursements, an auditor detected abnormally frequent first-two digit occurrences of the numbers 95, 99, and 10. Closer examination revealed that employees were excessively claiming breakfast expenses of $9.50, $9.90 and $10.00. The company policy stated that meals of $10.00 or less did not need a voucher for reimbursement. The employees biased the actual values to avoid turning in vouchers. The auditor also informed management that any increase in the dollar amount required for vouchers might result in an increase in the expenses claimed (Nigrini 1998).

Another example of bias occurred in one company's accounts payable data. In that case, the first two-digits of 24 appeared more frequently than expected. The auditors found that this was caused by employee claims for travel expenses biased to fall under the company's $25 voucher requirement (Nigrini 1999a).

**Uncovering Errors**

Errors are unintentional misstatements in recorded amounts. In addition to detecting fraud, auditors are responsible for performing procedures that are reasonably expected to discover errors in the client's financial statement amounts. Digital analysis is very useful in revealing errors.

A simple but effective method of digital analysis to discover errors is the number duplication test. The observed frequencies of the actual numbers in the selected dataset are noted, beginning with the number that occurs most often and ending with the number occurring the least frequently. Auditors usually focus on the items that occur abnormally often relative to

other numbers, especially odd numbers and round numbers, and on numbers that appeared more often than expected in the first two-digits test.

Most companies have many transactions involving inventory each year and thus inventory is quite susceptible to errors. In using the number duplication test to audit the inventory of an actual company, the auditors discovered several large odd numbers that occurred exactly twice. The auditors examined this situation in more detail and determined that inventory in three large sections of the client's warehouse was erroneously included twice on the inventory count sheets. This error would have materially overstated earnings had the auditors not found the problem (Nigrini 1998).

In another case, auditors examined the accounts payable of a large conglomerate by using digital analysis. The auditors found many large odd accounts payable amounts that occurred exactly twice in the number duplication test (Nigrini 1998). An in-depth examination by the auditors revealed that divisional databases were incorrectly merged. The invoices of one division were included twice, while the invoices of another division were omitted. Without discovering the error, the auditors would not have been able to select a sample to test the accounts payable of the omitted division.

## III. Conclusion

In this manuscript, we have discussed and illustrated how a relatively new audit technique, digital analysis, can be used to detect errors and frauds, thus leading to audit improvement. In fact, improving audit effectiveness and efficiency is very important to practicing auditors. As technology plays a key role in today's business environment, most of business transactions are processed and stored electronically. The tremendous growth of information technology provides opportunities and risks for auditors in performing their audit

tasks. The use of more advanced technology seems inevitable in audit area. Digital analysis along with Benford's Law may be a natural product of this IT era.

Benford's Law provides benchmarks that auditors can use in evaluating data to indicate items that are more likely than average to include fraud, misstatements, and other anomalies. Digital analysis combined with Benford's Law provides a productive method for processing large amounts of information and discovering such problems in reported amounts. This ability to efficiently process large amounts of data and detect anomalies increases the likelihood that an auditor will detect problems. Once the anomalies are located auditors use professional judgment and other analytical procedures to examine the anomalies and make determinations regarding whether fraud, processing inefficiencies, biases and errors are actually present. Consequently, the adoption of digital analysis technique as a basic audit tool in detecting errors, fraud and irregularities should greatly enhance auditor's performance, thus improving information quality of financial statements.

## REFERENCES


Coderre, David. 1999. "Fraud Detection Using Digital Analysis." The EDP Audit, Control and Security Newsletter (September).

Lanza, Richard. 2000. "Analytically Reviewing Digits for Fraud and Inefficiencies." CRC Press (June). <http://www.richlanza.com/infoarchive/articles/datas/digit.doc>

Nigrini, Mark . 1998. "Digital Analysis: a computer assisted data analysis technology for internal auditors." ITAudit (December).

_____. 1999a. "I've Got Your Number." Journal of Accountancy (May).

_____. 1999b. "Fraud Examination for Managers and Auditors." Working paper. <http://www.jsrsys.com/fema/new15sys.htm>

**TABLE 1**
**Expected Frequencies of Digits**

| Individual Digit | Position of Individual Digit in a Number | | | |
|---|---|---|---|---|
| | First | Second | Third | Fourth |
| 0 | not possible | .1197 | .1018 | .1002 |
| 1 | .3010 | .1139 | .1014 | .1001 |
| 2 | .1761 | .1088 | .1010 | .1001 |
| 3 | .1249 | .1043 | .1006 | .1001 |
| 4 | .0969 | .1003 | .1002 | .1000 |
| 5 | .0792 | .0967 | .0998 | .1000 |
| 6 | .0669 | .0934 | .0994 | .0999 |
| 7 | .0580 | .0904 | .0990 | .0999 |
| 8 | .0512 | .0876 | .0986 | .0999 |
| 9 | .0458 | .8500 | .0983 | .0998 |
| Total frequencies | 1.0000 | 1.0000 | 1.0000 | 1.0000 |