USING DIGITAL ANALYSIS TO DETECT FRAUD

Review of the DATAS® Statistical Analysis Tool

Richard B. Lanza

Digits (more commonly known as numbers) have played a key role in business and have contributed as much to our well being as has the wheel, fire, and agriculture. Although starting from a simple place in society, numbers have grown more and more difficult to assess as compendious databases have been growing exponentially in the 20th and now the 21st century. Such growth has made fraud that much harder to identify without sophisticated analytical tools.

This review discusses one such analytical tool, entitled Digital Analysis®, or more aptly named, Digital Analysis Tests and Statistics (DATAS®). This new audit technology, as used by two Big Five Firms and numerous Fortune 500 companies, helps auditors to be more effective and efficient by presenting various high level analysis along with the ability to "drill down" deeper as needed. With such a comprehensive tool, fraud is a usual target for the analysis.

DATAS® is based on the theory that there are expected frequencies or occurrences of digits in a list of numbers. This theory, otherwise known as Benford's Law, was based on the work completed by Frank Benford, a General Electric Research Laboratories physicist who conducted extensive studies on digit frequencies in tabulated data. Since his death, numerous studies and papers (over 150) have been written supporting and lending credibility to this theory, proving it to be the one true theory of numbers.

But why should forensic accountants care? Forensic accountants can use this mathematical principle to perform powerful analytics. It works by assigning an expected frequency to each number in a population, and then highlighting for review any numbers that fall outside of these expected frequencies. Since it is easily understood, the information can be readily presented to management. Auditors should also concern themselves with this analysis as it provides not only evidence of fraud but also indications of process inefficiencies and errors.

In short, DATAS® identifies process inefficiencies, errors, and fraud by searching for abnormal:

- Digit and number patterns
- Round number occurrences
- Duplications of numbers

Richard B. Lanza, CPA, is Senior Manager of the Program Management Office at the American Institute of Certified Public Accountants (AICPA), currently working on web, intranet, and application development project management. He has written numerous articles and a book on audit technology. The author's opinions expressed in this article are his own and do not necessarily represent policies or positions of the AICPA. Official positions of the AICPA are determined through certain specific committee procedures, due process, and deliberation.
Digit and Number Patterns

To understand how DATAS® would compare the frequencies of the first digits of numbers in a population to that expected under Benford's Law, assume there were five numbers in your population (200, 2050, 2200, 2532, and 1677). All but one of these numbers (1677) start with the digit "2." Benford's Law states that in a population of random numbers, the expected percentage of numbers starting with a "2" is roughly 19%. In this case, there is a deviation from Benford's Law of over 50%. Therefore, it is highly probable that these numbers have been invented (as they were in the above example) or there is some other systematic abnormality in the data set. (Note, for example, that such a finding could also indicate that there is an abnormal number of digits due to set numbers – e.g., retail prices of $19.99 leading to a high probability of the digit "1" in the overall population.)

Tax cheats, check forgers, and embezzlers are human; and hence, simply can't consciously produce truly random numbers. Police and auditors have depended on this human quirk for years, and DATAS® has proven itself capable of pinpointing this habit.

Let's consider DATAS® in a true example. The "First Digits" graph shows a payment history file (over 50,000 payments). Of all 9 digits, the digits "1" and "2" exceed Benford's Law (middle black line in the graph), yet only the "1" digit extends far beyond the law.

Although this may be interesting, the auditor cannot conclude there are inefficiencies or fraud in the payment history. This is because the First Digit test is only meant to provide a high level reasonableness check as to the expected frequencies in the database. Instead, the auditor must dig deeper into the data, which can be effectuated by running:

- The first two digit test
- The first three digit test
- Reports that extract key spikes in the first two and three digit tests
- Reports that summarize the key first two and three digits by subset (e.g., vendor)

In summary, Benford's Law is a proven law of numbers that can be used to identify fraud in large data sets by detecting potentially invented numbers. A non-inclusive list of examples of data sets where Benford's Law might be used includes:
• Investment sales/purchases
• Check register
• Sales history/Price history
• 401 contributions
• Inventory unit costs
• Expense accounts
• Wire transfer information
• Life insurance policy values
• Bad debt expenses
• Asset/liability accounts
• Bad debt expenses

Round Number Tests

Another feature of DATAS® is its ability to focus not only on the first few digits in a number, but also on the last two. These digits would generally represent the cents portion of an amount (e.g., in the number 123.45, "45" would be the last two digits). For example, the following graph illustrates the differences between the actual frequencies and Benford's Law expected frequencies of the last two digits of a sales history file.

It is quite evident that there is a high proportion of numbers ending with "00" or no cents (33% of the sales history), but there are also many "25," "50" and "75" digits which could also be viewed as round digits. This could be cause for alarm to the auditor considering it could be concluded that such numbers may have been invented. This conclusion follows the theory that humans are not random and will select numbers that they are accustomed to or could easily create. In this particular case, no irregularity existed since it was common in the industry for many of the sales prices to be in round numbers.

Using the Last Digits test looks for round numbers, but does not consider the entire number. Rather, it focuses on only that portion that extends to the right of the decimal point. DATAS® can also be used to review entire numbers for roundness. Historically speaking, the first accounting application of Benford's Law looked for round numbers. The application was based on the premise that when corporate net incomes were just below psychological boundaries, managers would tend to round these numbers up. For example, numbers such as $298,000 would be rounded to $300,000. The belief was that the latter number seems larger, although in percentage terms, it is only marginally higher.

This belief can be tested using DATAS® by first executing the Round Number Profile report (seen below as executed on the sales history file)
Table 1

<table>
<thead>
<tr>
<th>Multiples of</th>
<th>Actual Count</th>
<th>Actual Proportion</th>
<th>Expected Proportion</th>
<th>Over Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>10215</td>
<td>0.2592</td>
<td>0.1000</td>
<td>0.1592</td>
</tr>
<tr>
<td>25</td>
<td>7645</td>
<td>0.1940</td>
<td>0.0400</td>
<td>0.1540</td>
</tr>
<tr>
<td>100</td>
<td>4668</td>
<td>0.1184</td>
<td>0.0100</td>
<td>0.1084</td>
</tr>
<tr>
<td>1000</td>
<td>1525</td>
<td>0.0387</td>
<td>0.0010</td>
<td>0.0377</td>
</tr>
</tbody>
</table>

Total Obs = 39,413

The profile report counts (see actual count column) the frequency of numbers that are evenly divisible by 10, 25, 100, and 1,000. As can be seen, 26% of the sales history (or 10,215 sales out of 39,413) is evenly divisible by $10, which is expected considering the inherently round numbers detected in the last digits test above. Although expected, an additional report was executed to ensure there were no estimations or irregularities hidden by the numerous round numbers. This additional report listed all customers who were comprised solely of round numbers (their proportion of round numbers was 100%):

Table 2

<table>
<thead>
<tr>
<th>Customer</th>
<th>Total invoices</th>
<th>Total round payments</th>
<th>Individual check paid</th>
<th>Number of payments</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>27,000.00</td>
<td>27,000.00</td>
<td>9,000.00</td>
<td>3.00</td>
</tr>
<tr>
<td>B</td>
<td>58,000.00</td>
<td>58,000.00</td>
<td>58,000.00</td>
<td>1.00</td>
</tr>
<tr>
<td>C</td>
<td>200,000.00</td>
<td>200,000.00</td>
<td>200,000.00</td>
<td>10.00</td>
</tr>
<tr>
<td>D</td>
<td>2,000.00</td>
<td>2,000.00</td>
<td>2,000.00</td>
<td>1.00</td>
</tr>
<tr>
<td>E</td>
<td>10,000.00</td>
<td>10,000.00</td>
<td>10,000.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

When this report was researched, it was determined that customers A, B, D, and E were comprised of legitimate round numbers and properly supported by invoices and shipping documents. Customer C, on the other hand, did not have any invoices since it was based solely on estimates the sales representative was making about the expected level of sales from this customer. This irregularity increased the salesperson's commission (the motive for such estimation) and was not immediately detected due to poor oversight in the sales department. Needless to say, the employee was terminated and the sales entries were reversed at the conclusion of the audit.

Thus, DATAS® has an ability to not only provide a high level viewpoint for estimating trends, but it is also able to hone in on information for review that is most probably an exception. For example, in the above case, 39,413 sales were easily viewed for roundness in a one page graph or profile report yet, when the auditor needed to review detail information, only 5 customers (made of 17 sales) were isolated as most probable for review.

Abnormal Duplication of Numbers

There are many common duplication reports that auditors run, but the most basic is a check on accounts payable payments where the invoice, vendor, and amount are the same. From a fraud perspective, an employee could be paying a vendor on multiple occasions and receiving compensation for assisting in the deceit. DATAS® is able to perform such simple analysis, but it also has the ability to go further by providing advanced programs.

One of these programs will isolate those vendors made solely of duplicate payments. Going back to the theory that humans are not random, it is extremely probable that an employee fabricating payment amounts would select the same amount and quite probably the same vendor when effectuating their scheme. The below is from an actual audit where vendor B was found to be fraudulently created by the employee and used solely for the purpose of fabricated payments. The remaining vendors were either office rental or equipment lease payments made in monthly or quarterly intervals, respectively.
### Table 3

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Total Payments</th>
<th>Individual Payment</th>
<th># of Payments</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$84,000</td>
<td>$7,000</td>
<td>12</td>
<td>Rental</td>
</tr>
<tr>
<td>B</td>
<td>$12,000</td>
<td>$4,000</td>
<td>3</td>
<td>Fraud</td>
</tr>
<tr>
<td>C</td>
<td>$168,000</td>
<td>$14,000</td>
<td>12</td>
<td>Rental</td>
</tr>
<tr>
<td>D</td>
<td>$18,000</td>
<td>$4,500</td>
<td>4</td>
<td>Lease</td>
</tr>
</tbody>
</table>

Many auditors focus solely on duplicate payments when utilizing the duplicate search option of their data extraction/analysis software, due to the high payback, fraud or otherwise, that the analysis produces. It is important to recognize, however, that there are many other areas to look for abnormal duplication. For instance, a customer made solely of duplicate sale invoices could signal an invented customer or other estimation practice (to improve sales figures or corresponding sales commissions). Parts in inventory that have the same unit cost or extended cost may be erroneously double counted in inventory, requiring a restatement by the auditor. The applications for this test are endless and should be used in all areas where the auditor is concerned about overstatement.

**Conclusion**

As company databases grow and become more complex, fraud and other accounting irregularities are becoming more difficult to detect without sophisticated analytical tools. DATAS\textsuperscript{®} is one such tool. Its stated goal is to "provide a view of your entire audit population in easy to assimilate graphs so as to assess, from an analytical perspective, whether the population possesses abnormalities." These abnormalities either present themselves as duplications of digits, random numbers, or as duplicate amounts. Although powerful, DATAS\textsuperscript{®} is extremely easy to use. It requires no programming experience, utilizes menus to guide the user through the audit, and requires minimal setup time. Additional information on DATAS\textsuperscript{®} may be obtained by visiting the www.digitalanalysisonline.com website, and selecting the product section.