When Sequence Counts

Benford's Law Can Help You Detect Fraud

BY MICHAEL RHOADS

Fraudsters sometimes remain undetected for long periods of time, even years, causing organizations to lose millions of dollars. How do they get away with it?

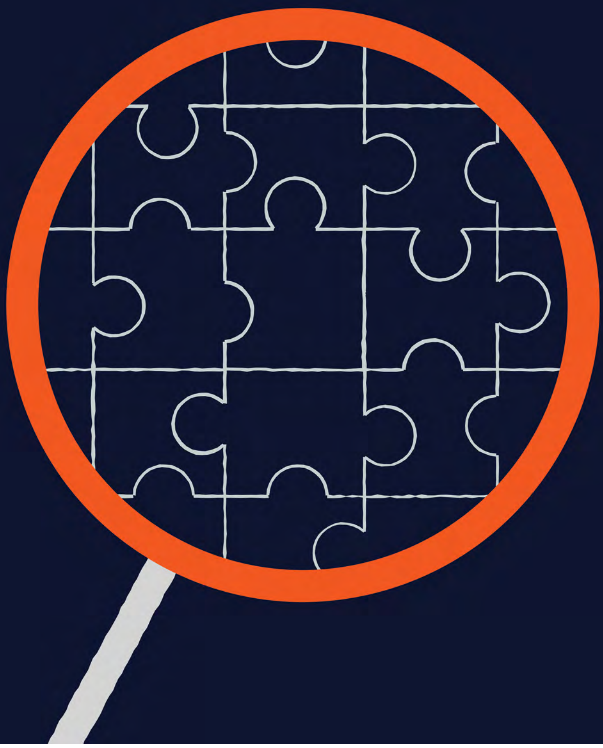
Internal controls are critical for fraud prevention, but they are not perfect. There are many ways they can fail, including poor judgment, human error, collusion, management override of internal controls, lack of segregation of duties, and unforeseeable circumstances. As a result, organizations should have additional internal controls in place for fraud detection. One of those additional controls is something you might not have heard of—Benford's Law.

A BRIEF HISTORY

In 1881, mathematician Simon Newcomb came across this data analytics principle while flipping through the pages of a book of logarithmic tables. He noticed that the pages in the beginning of the book were dirtier than the pages at the end and deduced that this was because his colleagues, who shared the library, preferred quantities beginning with the number one.

This leads us to Benford's Law, explained by physicist Frank Benford in a 1938 paper about "expected distribution of significant digits in a diverse set of naturally occurring datasets and how this can be used for anomaly or fraud detection in scientific or technical publications."[1]

Benford's Law maintains that the numeral "1" will be the leading digit in a set of data numbers 30.1 percent of the time; "2" will be the leading digit 17.6 percent of the time; and each subsequent numeral, 3 through 9, will be the leading digit with decreasing frequency. Keep in mind that Benford's Law works best with large data sets—the larger the better—and to conform with the law, the data set you use must contain data in which each number, 1 through 9, has an equal chance of being the leading digit.[2] This expected occurrence of leading digits is illustrated in Exhibit 1, in which the resulting curve closely resembles a steep water slide.



However fraudsters fabricate their numbers, including adding them manually, they probably won't be able to produce results that adhere closely to a Benford's curve.

USING THE LAW

Using any version of Microsoft Excel, government finance officers can count the leading digits contained in any large data set of natural numbers, such as accounts payable invoice dollar amounts, chart the findings, and compare the results to Benford's Curve to determine if that data set follows the expectations set forth by Benford's Law.

In 1972, economist Hal Varian suggested that the law could be used to detect possible fraud in lists of socioeconomic data submitted in support of public planning decisions. "Based on the plausible assumption that people who fabricate figures tend to distribute their digits fairly uniformly, a simple comparison of first-digit frequency distribution from the data with the expected distribution according to Benford's Law ought to show up any anomalous results."[3]

In a similar example, Benford's Law was used to show that the macroeconomic data the Greek government reported to the European Union before entering the eurozone was probably fraudulent (although that analysis happened years after the country joined).[4]

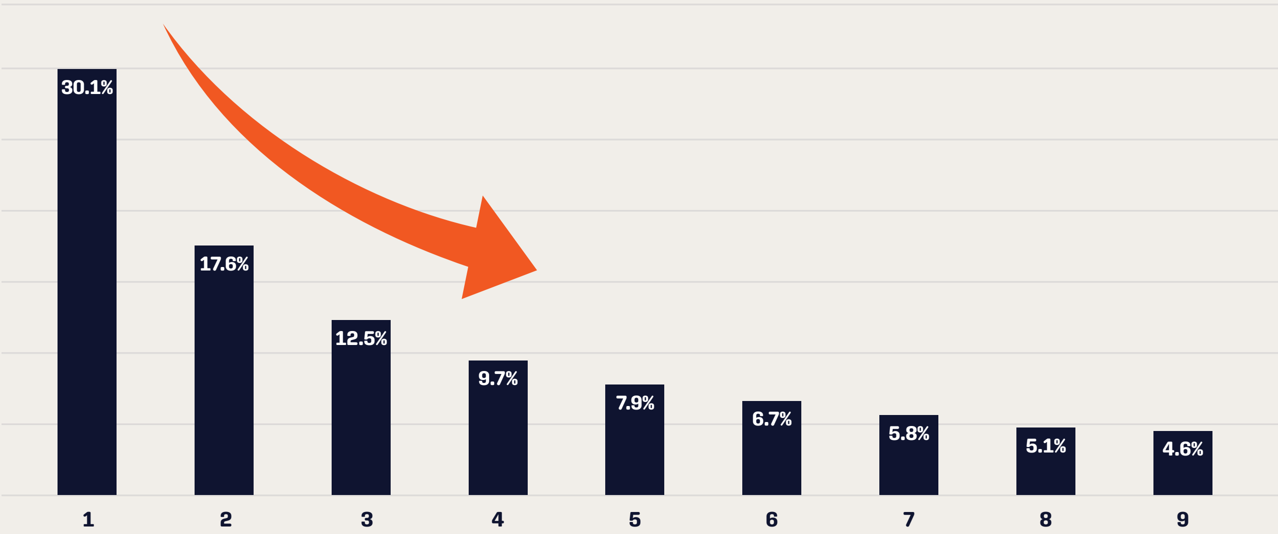
Of course, the law doesn't work for every application. Looking at accounting data, we'd expect the law to apply when the mean is greater than the median and the skew is positive; for numbers that result from mathematical combination of numbers (like quantity × price); and for transaction-level data (for example, disbursements and sales).[5] Distributions that would not be expected to obey Benford's Law include situations in which numbers are assigned sequentially (as with check and invoice numbers) or in which numbers are influenced by human thought (as in prices set by psychological thresholds, like $1.99); accounts with a large number of firm-specific numbers (such as accounts set up to record $100 refunds); or a built-in minimum or maximum; or distributions that don't span an order of magnitude of numbers.

Fraud examiners use Benford's Law tests on natural numbers such as payment amounts.[6] Fraudsters are unlikely to submit fake invoices in small amounts, like $100 or $200; they're more likely to try for bigger amounts, like $800 or $900. Doing this often enough upsets the natural order of the way numbers should occur. If, for example you run a Benford's Law test on your monthly payments and find that the first digit is 9 in 35 percent of the payments, this is an anomaly, since Benford's Law says that 9 should be the first digit only 4.6 percent of the time.

I have analyzed the invoice dollar amounts in my data set according to Bedford's Law and found that they conformed to the curve. As a result, I can be more confident that certain types of fraud aren't occurring in Accounts Payable. If a deviation from the curve were found, I would need to review the invoices in more detail to determine if fraud exists.

EXHIBIT 1 | **THE BENFORD CURVE**

Benford's Law: Percentage of time digits 1 through 9 are expected to occur in the first position



Data doesn't always match the curve and because of this, minor, statistically insignificant variations don't indicate fraud. If you are interested in learning more or want tips to implement Benford's Law for your organization or department, "Using Excel and Benford's Law to Detect Fraud," an article in the Journal of Accountancy, offers an excellent primer.[7]

The article also explains that however fraudsters fabricate their numbers, including adding them manually, they probably won't be able to produce results that adhere closely to a Benford's curve. A person who's producing numbers mentally is likely to repeat certain patterns—overusing 1, 3, and 4 to produce false data, for example, and underusing 6 and 8. Those anomalies would produce an erratic bar chart.

Benford's Law analysis isn't definitive, of course. This is an analytical tool that may help you determine whether further investigation is warranted. If you analyze your data and don't produce Benford's curve, consider whether one of the following issues may be the cause:

- The data may not be suitable for this kind of analysis.
- Ensure that the data set is valid.[8]
- Consider whether reliable controls are in place to detect or prevent improprieties.
- Reconsider the source of the data, and if it came from an outside source, find out how the source verified the information.

I urge all government finance officers to set aside time to apply Benford's Law in their organizations. You can start small by setting aside one day a year to apply Benford's Law to Accounts Payable invoice dollar amounts using Excel. Consider how Benford's Law could be applied to any other data your organization produces. The result might reveal fraud and reduce ongoing risk.

Michael Rhoads*, CPA, is the internal control director for the Auditor of State in Indiana. If you'd like to discuss Benford's Law in more detail, he can be reached at mrhoads@auditor.in.gov.*

[1] Tirthajyoti Sarkar, "What Is Benford's Law, and Why Is It Important for Data Science?" TowardsDataScience.com, October 25, 2018.

[2] J. Carlton Collins, CPA, "Using Excel and Benford's Law to detect fraud," *Journal of Accountancy*, April 1, 2017.

[3] Hal Varian, "Benford's Law, Letters to the Editor, *The American Statistician*, Volume 26, 1972.

[4] William Goodman, "The promises and pitfalls of Benford's Law," *Significance*, Royal Statistical Society, June 2016.

[5] Cindy Durtschi, W. Hillison, Carl Pacini, "The effective use of Benford's Law to assist in detecting fraud in accounting data," *Journal of Forensic Accounting*, Volume 5, 2004.

[6] John Gill, "What Is Benford's Law and Why Do Fraud Examiners Use It?" *AFCE Insights*, May 16, 2019.

[7] J. Carlton Collins, CPA, "Using Excel and Benford's Law to detect fraud," *Journal of Accountancy*, April 1, 2017.

[8] See AU-C Section 520, Analytical Procedures, for guidance on how to conduct an analytical review. Also, if the data set involves inventory, consider physical inspecting a sample; see AU-C Section 501, *Audit Evidence—Specific Considerations for Selected Items*, for guidance.