

Using Benford's Law for Fraud Detection in Accounting Practices

Arben Asllani (Corresponding author)

College of Business, University of Tennessee at Chattanooga

615 McCallie Avenue, Chattanooga, TN 37403 USA

Tel: 1-423-425-4412 E-mail: beni-asllani@utc.edu

Manjola Naco

Accounting Department, Tirana University, Tirana, Albania

Tel: 1-355-4237-55-76 E-mail: manjola_naco@yahoo.com

Received: June 8, 2014 Accepted: September 5, 2014 Published: October 1, 2014

doi:10.5296/jsss.v2i1.6395 URL: <http://dx.doi.org/10.5296/jsss.v2i1.6395>

Abstract

This paper offers a template that can be used by practitioners and accounting auditors to identify fraud in accounting practices. The proposed tool is based on the principles of Benford's Law. An actual example from a local Albanian hospital is used to demonstrate the use of the proposed methodology. The results of the investigation lead to some important findings and demonstrate the efficiency of the approach. The paper concludes with several practical recommendations and suggestions which should be considered to avoid any potential weaknesses of the proposed methodology.

Keywords: Fraud detection, Benford's law, Data analytics, Healthcare accounting practices

1. Introduction

Data analytics (DA) is often used to help identify and control fraud (Han, 2005), (Kotsiantis & Tzelepis, 2006), (Wegener & Rüping, 2010). DA consists of methods of gathering data, and analytical tools such as visualization programs, statistical methods, and data mining algorithms (Negash, 2004). The information gathered by DA is used to provide decision makers with accurate, timely, well-presented information. DA tools can be used to organize and analyze data from disparate sources to provide meaningful information for decision makers (Davenport & Harris, 2007) (Negash, 2004). The purpose of the paper is to demonstrate the use of Benford's Law as an analytical tool for fraud detection. A framework of the tool is provided as a practical approach and several recommendations for using the tool are provided.

The paper initially presents a brief overview of several auditing practices to detect and prevent fraud according to methods presented by Rezaee & Riley (2009) for preventing, detecting, and controlling incidents of financial reporting fraud. The next section discusses Benford's Law and presents the theoretical and mathematical formulation of the law. The model is elaborated using data from the accounting practices of a local hospital in a major city in Albania. A separate section discusses the results and demonstrates any identified anomalies and detected fraud if any. Finally, conclusions, recommendations, and implications for future use of the approach are discussed.

2. Accounting Practices for Fraud Detection

Financial fraud is becoming an increasingly serious problem to the global financial system due to the increasing number of factors that determine the occurrence of fraud, such as the globalization markets, merging and division of companies, constantly growing competition, political and economic factors. As Rezaee (2005) states, not only researchers but also practitioners have a more important task – to prevent frauds from occurring; therefore, it is necessary to analyze thoroughly the reasons for their appearance and possible techniques for execution and hiding. Effective detecting accounting fraud has always been an important but complex task for accounting professionals (Yue, Wu, Wang, & Chu, 2007). The auditing tools are generally considered to be costly, time consuming, and inefficient (Buddhakulsomsiri & Parthanadee, 2008).

The detection of accounting fraud using traditional internal audit procedures is a difficult or sometimes an impossible task (Fanning, Cogger, & Srivastava, 1995). First, the internal auditors usually lack the required knowledge concerning the characteristics of accounting fraud. Second, as the fraudulent manipulation of accounting data is so infrequent, most of the auditors lack the experience and expertise needed to detect and prevent frauds. Finally, the other concern is that people are intentionally trying to deceive the internal or external auditors (Fanning & Cogger, 1998).

Traditional fraud detection typically begins with an indication or anomaly that something isn't right after management, auditors or fraud examiners investigate with additional research, computer queries, or interviews. In order to develop his expectations, the auditor employs

analytical review techniques, which allow for the estimation of account balances without examining relevant individual transactions (Kirkos, Spathis, & Manolopoulos, 2007). This approach can be viewed as an inductive method: it begins with anomalies brought to someone's attention and continues by researching additional events and data until it is determined that fraud is causing the anomaly of indicators.

Deloitte and Touche conducted internal research into the use of statistical methods to focus on abnormal data, Statistical Techniques for Analytical Review (STAR), which identified significant fluctuations in data that warranted further information (Albrecht, Albrecht, & et.al, 2001). A research on audit techniques by Blocher and Willingham, was focus on using computers to conduct analytical procedures. In their research it is used regression modeling for trending and analysis of data (Blocher & Willingham, 1988).

Recently, information on Benford's Law and Digital Analysis has surfaced in the literature. Nigrini (1999) which discusses the use of digital analysis in accounts payable data, estimations in the general ledger, inventories, payments, and refunds. However, despite the initial interest in technology-based fraud-detection techniques, little work has been done (other than digital analysis) in recent literature. A detailed description of auditing and statistical analysis as fraud detection tools is offered by (Copeland, 2012).

The use of statistical tools has significantly increased as the availability of electronic accounting records has amplified due to the advent of enterprise wide information technology applications. Statistical methods have become even more popular as advanced statistical tools and data mining applications such as neural networks and fuzzy logic are developed.

Statistical methods are very efficient to pinpoint suspicious areas of fraud. On the other side, when using statistical tools, there is always a risk of identifying legitimate claims as fraudulent (Type I error) or missing to identify fraudulent claims (Type II error). In addition, these tools require advanced expertise on fraud detection and advanced knowledge in statistical methods (Major & Riedinger, 2002) in general, and data mining tools (Ortega, 2006) in particular.

Statistical considerations indicate that there is a far greater chance of making a Type I error if the Benford's Law analysis is completed on a "digit-by-digit" basis, as compared to the "test-by-test" basis typically employed by statisticians. The proposed methodology in this paper, "digit by digit" Benford's Law approach, can be used as an effective tool to assist auditors in targeting the potential areas of fraud and as a result can save time and reduce the auditing costs. The proposed methodology applies Benford's Law (Nigrini, 1992), (Nigrini, M, Mittermaier, L, 1997) via an Excel-based template that is simple to be used by practitioners.

3. Benford's Law and Fraud Detection

Benford's Law is considered as an effective and simple tool in the hands of the auditors for fraud detection (Durtschi, 2004). The fact that approximately 150 articles regarding this law are published during the last five decades indicates its application to real cases. Benford's Law is based on the unique observation that certain digits appear more frequently than others

in data sets.

For example, it has been observed that more than 30% of numbers begin with the digit one. Benford (1938) collected more than 20,000 observations from such diverse data sets as areas of rivers, atomic weights of elements, and numbers appearing in magazine articles. The law suggests the following two formulas to estimate the probability that a number has any particular non-zero first digit (1) or zero or non-zero second digit (2):

$$P(D_1 = d_1) = \log\left(1 + \left(\frac{1}{d_1}\right)\right); \quad d_1=(1,2, 3\dots9) \quad (1)$$

$$P(D_2 = d_2) = \sum_{d_1=1}^9 \log\left(1 + \left(\frac{1}{d_1 d_2}\right)\right); \quad d_2=(1,2, 3\dots0) \quad (2)$$

Where:

D1 represents the first digit of a number

D2 represents the second digit of a number

The results of such a formula are shown in Table 1. The table indicates that when data follows Benford's Law there is almost a 30.1% chance that number 1 is in the first position and almost 11.96% that number one is in the second position. Likewise, there is almost a 17.6% chance that number 2 is in the first position and almost 11.38% that number one is in the second position, and so on. If the actual distribution of the digits from a data set of accounting transactions does not follow these distributions, then there is reason to believe that data is manipulated by human intervention, and as such, these data sets must be further investigated for potential fraud.

Table 1. Expected frequencies based on benford's law

Digit	1st place	2nd place
0		.11968
1	.30103	.11389
2	.17609	.19882
3	.12494	.10433
4	.09691	.10031
5	.07918	.09668
6	.06695	.09337
7	.05799	.09035
8	.05115	.08757
9	.04576	.08500

The mathematical proof of Benford's Law is beyond the scope of this article. However, an

intuitive explanation is offered below. Consider the purchasing price of a medical appliance to be 1,000,000 lek. This value has to double in size before the first digit is a “2,” in other words it needs to grow 100%. For the first digit to become “3,” it only needs to grow 50%. To be a “4” the price of the medical equipment must only grow 33% and so on. Therefore, in many accounts payable transaction data, the first digit one is much further from two than eight is from nine. Thus, the observed finding is that for these distributions, smaller values of the first significant digits are much more likely than larger values.

Nigrini & Mittermaier (1997) show the logic of using Benford’s Law to help discover unusual patterns in accounting transaction activities. It is very likely that an individual making fraudulent entries will enter the same amount or similar amounts many times. In that case, the resulting variation of first and second digits from the Benford’s Law probability distribution may lead the auditor to discover the fraudulent transactions. Nigrini also outlined several practical applications where a fraud auditor could effectively employ Benford’s Law: accounts payable data, general ledger estimations, duplicate payments, and customer refunds (Nigrini, 1999). Other applications of Benford’s Law include detecting vendor kickbacks, detecting fictitious vendors, and detecting overstated divisional performance (Tapp & Burg, 2001).

3.1 Using Benford’s Law for hospital auditing

This paper demonstrates how an accountant or auditor may use Benford’s Law as an aid tool to detect fraud. Data from a local hospital in a major city in central Albania are used to illustrate the technique. This is one of the largest hospitals in Albania and operates in a region with over 300,000 people. The hospital has 420 beds and serves about 12,000 patients per year. The hospital is among the few if not the only hospital in Albania that uses modern accounting practices such as activity based costing (ABC).

The data used in this paper are expressed in the Albanian currency. It is noted that Benford’s Law applies to data in any currency. Pinkham (1961) demonstrated that multiplying all values in the Benford set by any constant yields another Benford set. It is worth noting that any indication of potential fraud identified by the proposed methodology must be followed and confirmed by on-site audits or investigations. Authors of the paper are aware of potential limitations of the methodology and discuss those concerns in the discussion section of the paper.

Table 2 represents an overall structure of accounting expenses for the hospital according to each department. As shown, almost half of the costs of the hospital occur in outpatient service activities. This cost structure is typical for hospitals in Albania. Since there is a lack of healthcare distribution in rural areas, people seek medical help from hospitals for a wide variety of medical conditions.

Table 3 shows a list of accounting expenses according to each cost activity. It is surprising to note that the most expensive activities of the hospital are salaries (311,902,778 or 25% of total cost) and maintenance (276,244,832 or 22% of total cost). Other significant expenses occur in water usage, blood transfusion, social security, and communication (phone, fax, and

Internet). A significant cost occurs in the “others” group (152,614,397 or 12% of total cost). From an auditor’s perspective this group probably needs further investigation to ensure that each item in this category is justifiable.

Table 2. Distribution of annual cost by department

Department	Total Cost
Pathology	56472551
Infections	14746100
Surgery	113431323
Maternity	110308176
Special Units	6,255,357
Pediatric Units	45,192,596
Children Rehab	10,779,850
Outpatient Services	624051309
Dispensary	158081461
Emergency	33,251,957
Blood analysis	4,884,638
Laboratory	10,620,012
Physiotherapy	2,058,609
Pharmacy	3,437,648
Kitchen	5,094,451
Anatomy	3,386,061
Administration	25,874,899
Records	2,510,026
Inpatient Services	5,952,760
Laundry	3,156,052
Warehouse	1,766,134

Table 3. Distribution of annual cost by activity

Activity	Total Cost
Salaries	311902778
Social Sec	82936661
Blood Transfusion	112526356
Food	13811361
Laundry	1,738,192
Clothing	2,004,182
Office Supplies	3,138,241
Travel/Per Diem	1,786,613
Daily Consumption	4,046,756
Gas/Oil	35,263,987
Water	111558560
Electricity	20,363,641
Phone/Fax/Internet	87,259,342
Security	24,116,071
Maintenance	276244832
Others	152614397
Total	1241311970

4. Excel Based Methodology

This section introduces the methodology in general. Next section applies the proposed methodology in the hospital settings. Considering a given data set under investigation, the auditor can use the following steps to identify potential areas of accounting fraud:

Step 1: *Identify the first and second digit in each number of the data set.*

This step can be easily implemented using the following Excel formulas:

$$D_1 = \text{left}(X, 1) \quad (3)$$

$$D_2 = \text{right}(\text{left}(X, 2), 1) \quad (4)$$

Where D_1 and D_2 are respectively the digits in the first and second position and X is the cell that contains a number in the data set. For a more thorough investigation, the auditor can identify the digit in the k^{th} position using the formula:

$$D_k = \text{right}(\text{left}(X, k), 1) \quad (5)$$

Step 2: *Count how many times each digit appears in the first and second position of the numbers in the data set.*

A simple *countif* function can be used in this case. Specifically, the number of times that $n = \{0, 1, 2, 3...9\}$ appears in the first or second position of the data set can be calculated respectively:

$$C_1(n) = \text{countif}(D_1=n) \quad (6)$$

$$C_2(n) = \text{countif}(D_2=n) \quad (7)$$

Step 3: Calculate the percentage that each digit appears in the first and second position of the numbers in the data set.

As indicated below, (8) and (9) show the formulas that can be used to calculate the respective percentages for the first and second position:

$$P_1(n) = C_1(n) / (C_1(0) + C_1(1) + C_1(2) + \dots + C_1(9)) \quad (8)$$

$$P_2(n) = C_2(n) / (C_2(0) + C_2(1) + C_2(2) + \dots + C_2(9)) \quad (9)$$

Step 4: Compare the actual distribution of the digits in the data set to the distribution indicated in Table 1.

Advanced statistical tests such as Z-test or t-test can be performed for this comparison. However, graphical comparisons of actual distributions to the Benford's Law distribution are sufficient when the methodology is used to simply locate potential areas of fraud, which will later be undergoing a case-by-case investigation by the auditor.

4.2 Implementation of the Methodology

The above methodology is applied to an overall data set that consists of all activity costs in the hospital. Figures 1 and 2 demonstrate the distribution of the digits in the first and second position, respectively. Figure 3 indicates that numbers 1 and 2 appear less often in the first digit than the distribution suggested by Benford's Law. This is a red-flag for the auditor; it indicates a potential manipulation of data, probably a tendency to roundup numbers. This finding can be further supported by the evidence that there is a significant higher than usual percentage of digit 0 in the second position.

These two findings demonstrate the power of the proposed methodology. Many of the fraudulent activities can be found by chance. A recent survey indicates that "34.2 percent of frauds were detected through tips, 25.4 percent by accident, and 20.2 percent through internal audits" (Hancox, 2013). Using the Benford's Law methodology the auditors can scan large amount of data sets from many organizations and potentially eliminate unnecessary internal audits. The auditors can only focus on those organizations which have recorded data set which indicate potential fraud.

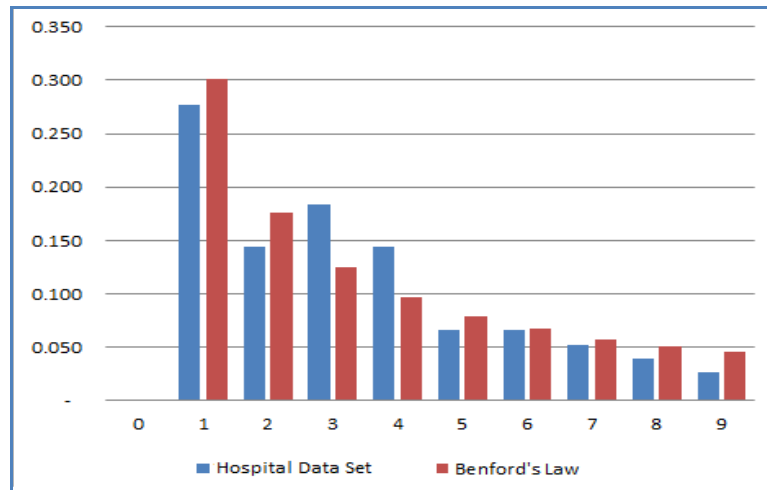


Figure 1. Distribution of the first digit in the activity costs data set

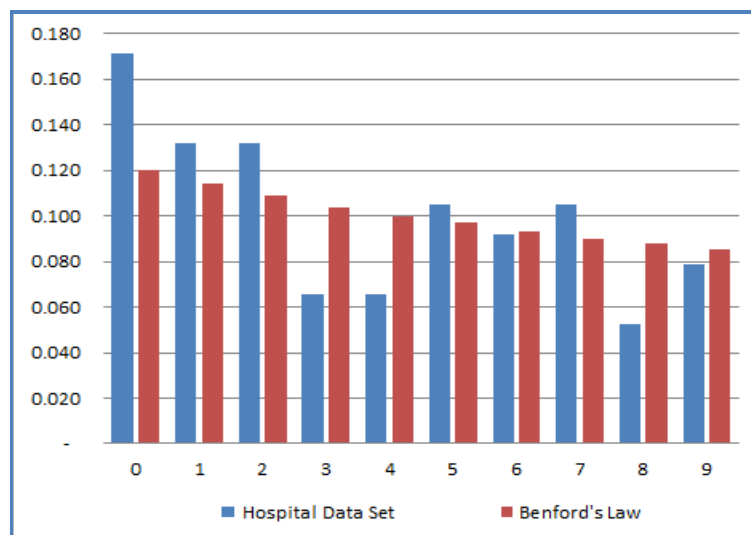


Figure 2. Distribution of the second digit in the activity costs data set

Once an auditor learns that there is a potential fraud in a certain organization, the auditor can further investigate the data set to pinpoint the fraudulent activities. For example, once the proposed methodology indicates anomalies in the overall data set of accounting records in the hospital, the auditor can further explore the data set by selecting specific activities. This paper selected the following activities for further investigation: salary, security, maintenance, and others. It is found, as shown in Figure 3, that while salary and security seem to follow Benford’s Law of the distribution of the first digit in the data set, that is not the case for maintenance and others. Considering that the last two categories are significant, as noted earlier, the auditor investigates these two categories more closely. Similarly, Figure 4 displays the distribution of the digits in the second position for numbers in the same data set. In this display, the others group shows significant deviation from the Benford’s Law distribution.

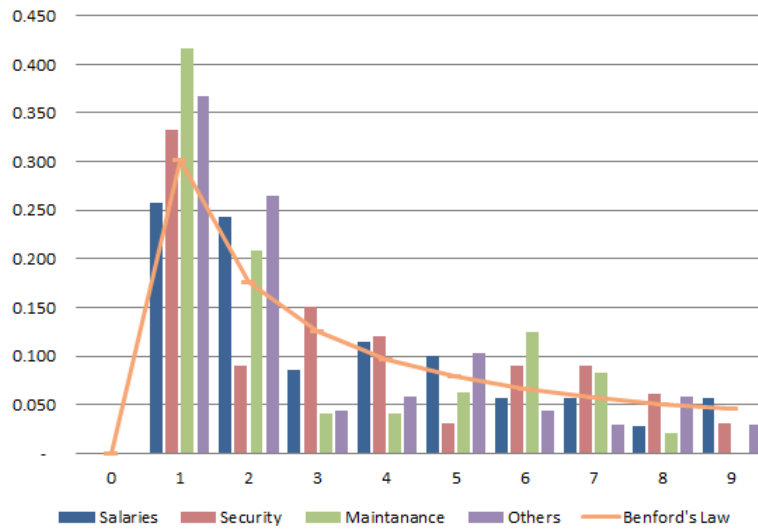


Figure 3. Distribution of the first digit in selected activities

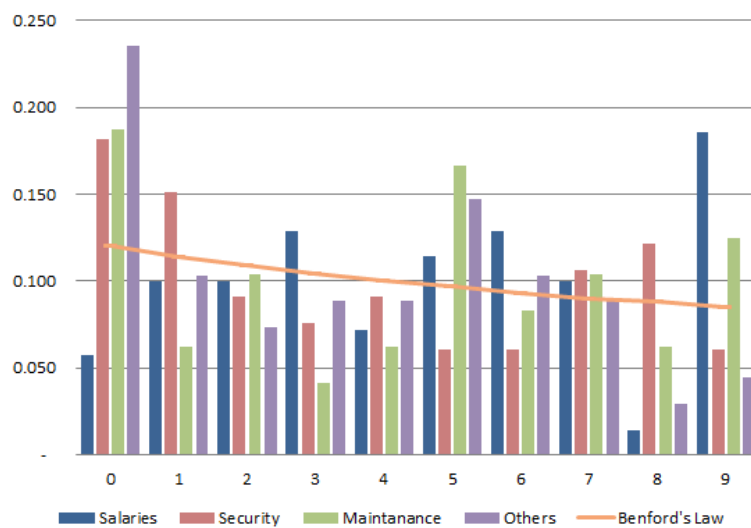


Figure 4. Distribution of the second digit in selected activities

The same analysis can be applied when the cost is classified by department. Figure 5 shows the distribution of the digits in the first position of numbers in a data set created by a selected group of hospital units. As shown, there is no indication of fraud in the emergency department as the data set related to this department seems to follow Benford's Law. That is not the case for other hospital units such as pathology, dispensary, and inpatient services. An onsite auditing should be focused in these three departments. Similarly, Figure 6 displays the distribution of the digits in the second position of numbers in the same data set. In this display, inpatient services unit continues to indicate irregularities. Also, other departments show warning signs such as emergency and special units.

The auditor may further decide to drill down in sub-categories. For example, Figure 7 shows

the distribution of the digits in the first position of numbers in a data set created by several hospital subunits within the pediatric unit. At this level, there is another sign of potential fraud: missing digits in a certain category. For example, there is no expense that starts with the number 3 in the emergency department. Or, there are no expenses that start with 3s or 5s in the department of reanimation. Finally, Figure 8 displays the distribution of the digits in the second position of numbers in the same data set.

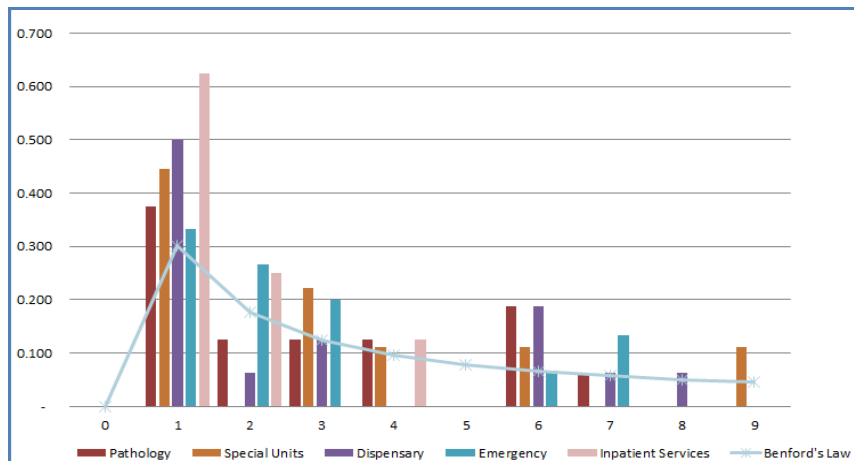


Figure 5. Distribution of the first digit in selected hospital units

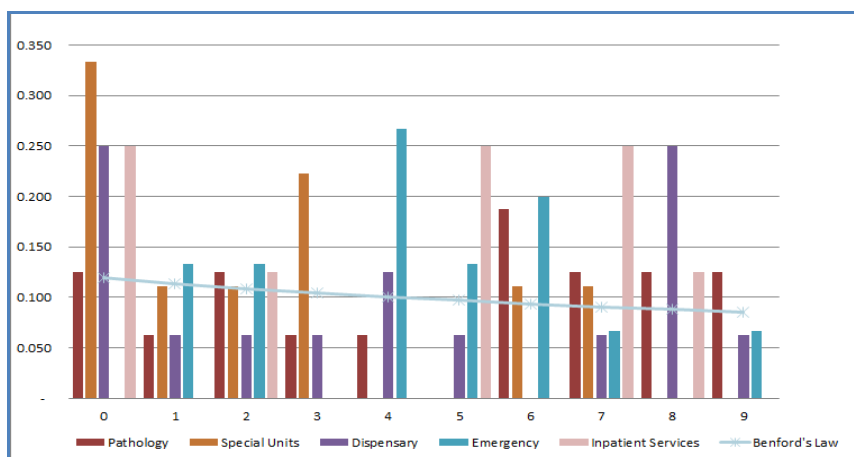


Figure 6. Distribution of the second digit in selected hospital units

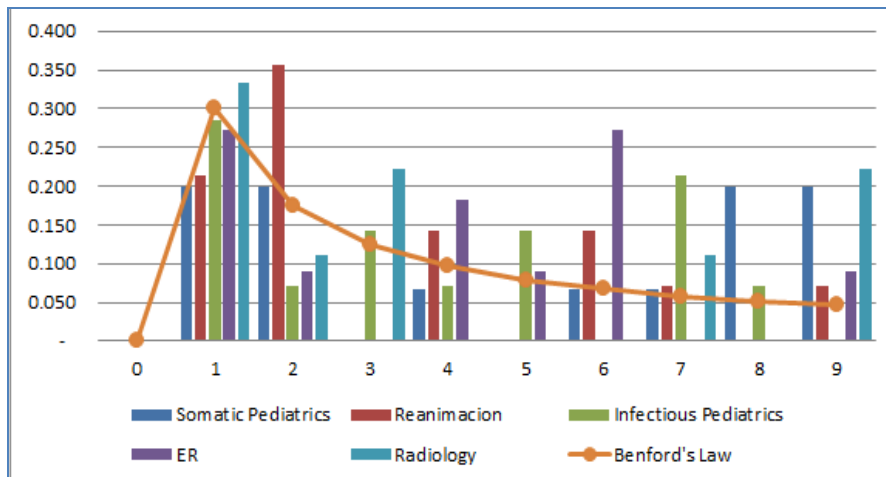


Figure 7. Distribution of the first digit in selected hospital sub-units

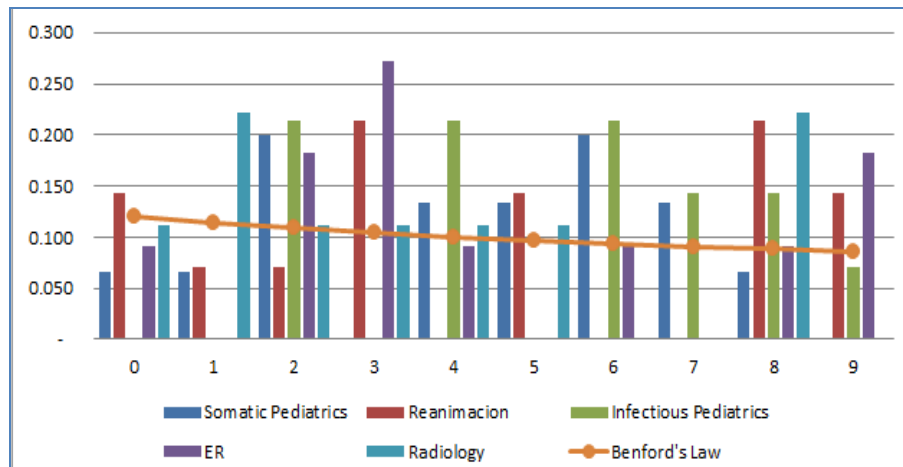


Figure 7. Distribution of the second digit in selected hospital sub-units

4.2 Results

This implementation of the proposed methodology in hospital demonstrates the efficiency of the template. First, the template is easy to use and can be implemented by any auditor who has a basic understanding of Microsoft® Excel. Second, the methodology can be used to efficiently identify potential organizations which may be performing fraudulent activities. For example, the authors of the paper had access to cost's data set from seven hospitals in Albania. When implementing the methodology to the overall data sets, the authors identified two hospitals whose data sets did not follow Benford's distribution and the hospital presented in this paper is one of them. This indicates that the methodology not only can eliminate unnecessary internal audits, but also can be used to apply a more due care methodology instead of leaving fraud be discovered by accident.

Third, the implementation of the methodology shows drill down capability. The paper demonstrated that by investigating selected activities, one can identify areas where numbers

are rounded up, or certain digits are completely missing. Specifically, the methodology indicated potential fraudulent activities in items classified as maintenance and others. Also, focusing on a specific department of the hospital, the methodology indicated fraudulent transactions in pathology, dispensary, and inpatient services. Finally, the Benford's Law was able to identify missing digits in the first and second position, especially in the department of emergency and reanimación.

5. Discussion and Conclusions

The detection of fraud in financial statements is a difficult task, which requires more than just using standard auditing procedures. Therefore, auditors need new tools and techniques to simplify auditing tasks and help them in detecting such fraudulent financial statements.

As demonstrated above, Benford's Law can be used as an effective tool to target potential areas of concern in accounting data. The proposed approach should not be used as the only tool for accounting audit; instead it must be used as a complementary tool. Statistically speaking, when using the proposed methodology, an auditor would essentially be testing the following statistical hypothesis: digits in a data set are distributed according to Benford's Law. Rejecting this hypothesis could mean that one of the following three scenarios might have occurred:

- *Although the data really do follow Benford's Law, due to random chance this particular set of observations does not (Type I error).*
- *There is a reasonable explanation for a deviation of some particular first digit from the expected distribution (recurring similar such as monthly phone charges, rent expense based on a contractual lease, etc.).*
- *Some of the entries are fraudulent; Benford's Law is not likely to apply to measurements where human intervention has occurred.*

As shown, the presence of fraudulent activity is just one of the possible explanations. For an auditor, rejecting the hypothesis means that additional work needs to be completed in order to determine whether fraud is present or not. "Case by case" and on-site auditing approaches to find evidence of fraudulent entries have relative benefits. It makes it easy to control the probability of Type I error. On the other side, using the "digit by digit" Benford's Law approach increases the chances of a Type I error, but also increases the chances of finding actual fraudulent entries.

This paper offers a methodology that can be used by auditors to efficiently pinpoint potential areas of fraud and further explore fraudulent accounting practices. This methodology is based on Benford's Law and uses Excel-based functions. When used properly, the proposed approach can significantly save auditing time and other resources by allowing the auditor to focus only in more specific areas of accounting records to fulfill the fraud detection responsibility.

References

- Albrecht, C., Albrecht, W. S., & et al. (2001). *Can Auditors Detect Fraud: A review of the research evidence. The Journal of Forensic Accounting, 1*, 1-12
- Benford, F. (1938). *The Law of anomalous numbers*. Proceedings American Philosophical Society.
- Bloch, E., & Willingham, J. J. (1988). *Analytical review: A guide to analytical procedures* (2nd ed). Colorado: Shepard's/McGraw-Hill.
- Copeland, L. (2012). Applying Business Intelligence Concepts to Medicaid Claim Fraud Detection. *Journal of Information Systems Applied Research, 51-61*.
- Davenport, T. H., & Harris, J. G. (2007). *Competing on Analytics: New Science of Winning*. Boston: Harvard Business School Press.
- Durtschi, C. (2004). The Effective Use of Benford's Law to Assist in Detecting Fraud in Accounting Data. *Journal of Forensic Accounting, 5*, 17-34.
- Fanning, K., & Cogger, K. (1998). *Neural network detection of management fraud using published financial data. International Journal of Intelligent Systems in Accounting, Finance & Management, 7*, 21-24.
[http://dx.doi.org/10.1002/\(SICI\)1099-1174\(199803\)7:1<21::AID-ISAF138>3.0.CO;2-K](http://dx.doi.org/10.1002/(SICI)1099-1174(199803)7:1<21::AID-ISAF138>3.0.CO;2-K)
- Fanning, K., Cogger, K., & Srivastava, R. (1995). Detection of management fraud: a neural network approach.
- Han, K. (2005). *Data Mining: Concept and Techniques*. San Francisco: Morgan Kaufmann.
- Kirkos, E., Spathis, C., & Manolopoulos, Y. (2007). Data Mining Techniques for the detection of fraudulent financial statement. *Expert Systems with Applications, 32*, 995-1003.
<http://dx.doi.org/10.1016/j.eswa.2006.02.016>
- Kotsiantis, K., & Tzelepis, T. (2006). Forecasting Fraudulent Financial Statements Using Data Mining. *International Journal of Computational Intelligence, 104-110*.
- Major, J., Riedinger, D. (2002). A Hybrid Knowledge-Statistical-Based System for the Detection of Fraud. *The Journal of Risk and Insurance, 309-324*.
<http://dx.doi.org/10.1111/1539-6975.00025>
- Negash, S. (2004). Business Intelligence. *The Communications of the Association for Information Systems, 13*, 54-67.
- Nigrini, M., & Mittermaier, L. (1997). The Use of Benford's Law as an Aid in Analytical Procedures. *Auditing: A Journal of Practice & Theory, 16*, 52-67.
- Nigrini, M. (1992). *The Detection of Income Tax Evasion Through an Analysis of Digital Frequencies*. Cincinnati, OH: Ph.D. thesis. University of Cincinnati.
- Nigrini, M. (1999). Adding Value with Digital Analysis. *The Internal Auditor, 56*, 21-23.
- Ortega, P. (2006). A Medical Claim Fraud or Abuse Detection System Based on Data Mining: A Case Study in Chile. *Conference on Data Mining* (pp 224-231). Las-Vegas, Nevada.
- Pinkham, (1961). On the Distribution of First Significant Digit. *Ann Mathematical Statistics, 32*, 1223-1230. <http://dx.doi.org/10.1214/aoms/1177704862>
- Rezaee, Z. (2005). Causes, consequences and deterrence of financial statement fraud. *Critical*

- Perspective on Accounting*, 16, 277-298. [http://dx.doi.org/10.1016/S1045-2354\(03\)00072-8](http://dx.doi.org/10.1016/S1045-2354(03)00072-8)
- Rezaee, Z., & Riley, R. (2009). *Financial Statement Fraud: Prevention and Detection Second edition*. WILEY.
- Tapp, D., & Burg, D. (2001). Using Technology to Detect Fraud. *Pennsylvania CPA Journal*, 71, 20-23.
- Singlenton, T., Singlenton, A., Bologna, J., & Lindquist, R. (2006). *Fraud auditing and forensic accounting, Third edition*. Wiley.
- Wegener, D., & Ruping, S. (2010). On Integrating Data Mining into Business Processes. *13th International Conference, BIS 2010* (pp. pp183-194). Berlin, Germany: Springer.
- Yue, X., Wu, Y., Wang, Y., & Chu, C. (2007). *A review of data mining-based financial fraud detection research*. Interantional conference on wireless communication.

Copyright Disclaimer

Copyright reserved by the author(s).

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).