BENFORD'S LAW: A DETAILED UNDERSTANDING AND ITS APPLICATION IN DETECTING ACCOUNTING SCAMS

M Aishvariya Valliammai[1] **Syeda Suphia Fathima[2]** **Bindiya C Mastikatte[3]**
B.Com F&A, Department of Professional Studies, Christ University, Karnataka, India

Asra Khan[4]
Assistant Process Associate, McKinsey & Co.

Abstract
This paper seeks to explain the definition and formula of Benford's Law and provides the statistical test used to check for its compliance. The law has also been used to identify the accounting fraud on the basis of the quarterly reports of two companies. The two companies on which such model has been applied include Satyam Computer Services Limited and Reliance Industries Limited. This paper concludes with an understanding of the efficiency of the law in identifying frauds in the financial statements by having a comparative analysis of the results of both the companies.
Key words: Benford's Law, statistical tests, accounting fraud, comparative analysis

1. INTRODUCTION

Numbers are an inescapable part of everyday life which can be used for many purposes such as counting, measuring, reporting, accounting, mathematics, labelling, ordering, and code. However, in today's time, numbers are increasingly being 'misused' rather than being 'used' to influence the reported figures of the financial statements in such a way that instead of showing the actual financial standing of the company, they delineate what the management wants to convey to the stakeholders. That being said, it's high time that corporates start implementing proactive measures that can prevent, detect and investigate potential acts of accounting fraud. One such technique is the Benford's Law (also known as the first digit law or Benford's distribution) which is a simple observation of the frequency of occurrence of the leading digits of numbers found in the real world data-sets which suggests that smaller digits have a greater probability of being the leading digit of a number as opposed to the larger digits. The Benford's law extensively helps in detecting fraud in financial statements by simply looking at the frequency of occurrence of digits. A very high repetition of a particular digit as the leading digit of multiple transactions doesn't necessarily indicate a possibility of prevalence of a fraud, rather, it gives the auditor an indication to probe deeper into investigating the matter to check for fraud.

2. REVIEW OF LITERATURE

(Cleary & Thibodeau, May 2005) This study ensures that the application of digital analysis using Benford's Law holds good for a fraud detection process. However, it does not apply when data has the same magnitude. The Chi-squared goodness of fit tests is standard in many statistical packages and the output could be easily adapted and tested. Thus the prudent approach would be to begin the analysis stage with an overall analysis using a chi-squared test, and then follow up with a "digit-by-digit" analysis if there is an indication of possible fraud in the overall analysis.

(Johnson & Weggenmann, November 3 2013) This articles states that the data in the first digit location were analyzed using Audit Command Language (ACL) Benford's law programming for identifying biased data. It also suggests that greater the number of data distribution

embedded in a number, the greater the effectiveness of Benford's law in identifying the anomalies in that data. Even though repeating studies in the private sector by analyzing account level data is inviting, these applications of Benford's law should be left to the practicing auditor and fraud examiners.

(Geyer) the data sets obeying Benford's law applies for 30.1% of numbers have 1 as first digit whereas this percentage falls to 4.6% for numbers having 9 as first digits. The fundamental property of the law is the scale invariance that is if data set is multiplied by a non-null constant, the new data set will also obey the same law. Thus if data sets of shares valued in Euros obey Benford's law , data sets valued in dollars or yens will have the same property. This is a problem in cases of fraud by systematic under or overvaluation.

(Tiscini) A common aspect characterizing the main scandals is the relevance of corporate reporting failure, as shown by the willingness of corporate managers to inflate financial results, either by overstating revenues or by understating costs or to divert the company funds to private uses. Financial and accounting fraud prevention is not always possible. Due to the difficulty to prevent accounting fraud, the mechanisms of internal/external control which are strictly linked to corporate governance systems, must be very effective and play anyway a primary role to detect any kind of possible fraud.

(Durtschi, Hillison, & Pacini, January 2004) Digital analysis can increase an auditor's ability to detect fraud. If distributions are selected at random and random samples are taken from each of this distribution, then the significant digit frequency of combined sampling will converge to Benford's Distribution, even though the individual distributions may not closely follow the law. Certain sets of accounting numbers often appear to closely follow Binomial Distribution, example: accounts receivables, cost of goods sold. Auditors often analyze payment amounts to test for duplicate payments.

3. RESEARCH DESIGN

STATEMENT OF PROBLEM

Detecting fraud or material misrepresentation in financial statements has always been a difficult task for an auditor or an accounting practitioner. In spite of there being multiple reforms in the field of accounting in terms of issuance of various accounting standards, audit compliance procedures and stricter corporate governance rules, yet most corporates manage to find their way out to manipulate their financials to meet their own personal motives. Thus, in this paper, by using Benford's law, a tool for identifying the possibility of fraud in the financial statements is provided to give a sense of security to the investors with regards to the quality of corporate financial reporting.

SOURCES OF DATA

Secondary data has been taken for the purpose of this research paper. The quarterly financials of the companies have been obtained from: www.moneycontrol.com

HYPOTHESIS

Since we're already aware that Satyam Computers is a fraudulent company and we are proving it with help of effectiveness of Benford's law, the hypothesis shall be as follows:
Null hypothesis (H_0): The financial data consists of manipulative data
Alternate hypothesis (H_1): The financial data does not consist of manipulative data

DATA ANALYSIS TOOLS

1. Benford's Law

The following formula put forth by Benford is considered for computation of expected frequencies:

For first digit of the number:
Probability (D1=d1) = log(1+(1/d1)); d1= (1,2, 3…9)

For second digit of the number:
Probability (D1D2 = d1d2) = log(1+(1/ d1d2))
Probability (D2= d2 | D1= d1) = log(1+(1/ d1d2))/log(1+(1/ d1))
Where D1 represents the first digit of a number, D2 represents the second digit of a number, etc.

2. Chi-square test of goodness of fit

Since this test is used to compare the expected and actual frequencies, it is best suitable for this research. It is a one tailed (upper tailed) test with a significance value of 5%.

OBJECTIVE

To obtain a detailed understanding of Benford's Law and its utilization in detecting fraud by actual application in the financials of a company noted for scam in comparison with a company having clean financial statements.

LIMITATIONS OF THE STUDY

1. The financials of the company are required for a long time period to get accurate results.
2. The study does not perform any other method apart from Benford's Law to detect accounting fraud in the financial statements.
3. This law cannot be used to detect fraud when very few transactions that have been manipulated.
4. This law determines if there exists a fraud but does not quantify to what extent and which period exactly the fraud is done.

4. DATA ANALYSIS AND INTERPRETATIONS

i. Benford's Law - With the application of Benford's law formula the following figure of expected frequencies can be obtained

Figure 1:

Expected Frequencies Based on Benford's Law

Digit	1st place	2nd place	3rd place	4th place
0		.11968	.10178	.10018
1	.30103	.11389	.10138	.10014
2	.17609	.19882	.10097	.10010
3	.12494	.10433	.10057	.10006
4	.09691	.10031	.10018	.10002
5	.07918	.09668	.09979	.09998
6	.06695	.09337	.09940	.09994
7	.05799	.09035	.09902	.09990
8	.05115	.08757	.09864	.09986
9	.04576	.08500	.09827	.09982

Source: Nigrini, 1996.

ii. Chi-square test of goodness of fit - The financials of a company needs to be obtained such as profits, debts, cash flow etc. (in this case income from operations). Also, the data size needs

to be large. Subsequently, the first digits of respective financials should be listed down and a frequency table for the same should be prepared. These will be the actual frequencies that need to be compared with the expected. For testing if the frequencies are in accordance with Benford's, Chi Square Test of goodness of fit test is conducted. The formula is,

$$X^2 = \sum \frac{(\text{observed - expected})^2}{\text{expected}}$$

SOURCE: (Mohan, 2010)

Following is the quarterly financials of Reliance Industries and Satyam Computer Services:

RELIANCE INDUSTRIES

Table 1:

Quarterly	Total Income (Amount in crores)	First Digit	First Digit(clubbed to get expected frequency)	Quarterly	Total Income (Amount in Crores)	First Digit	First Digit(clubbed to get expected frequency)
Mar '00	6594	6	6	Jun '09	32055	3	3
Jun '00	6824	6	6	Sep '09	46848	4	4
Sep '00	8394	8	7	Dec '09	56856	5	5
Dec '00	6555	6	6	Mar '10	57570	5	5
Mar '01	5900	5	5	Jun '10	58228	5	5
Jun '01	6390	6	6	Sep '10	57479	5	5
Sep '01	6234	6	6	Dec '10	59789	5	5
Dec '01	5766	5	5	Mar '11	72674	7	7
Mar '02	6642	6	6	Jun '11	81018	8	7
Jun '02	11593	1	1	Sep '11	78569	7	7
Sep '02	12507	1	1	Dec '11	85135	8	7
Dec '02	12033	1	1	Mar '12	85182	8	7
Mar '03	13962	1	1	Jun '12	91875	9	7
Jun '03	13509	1	1	Sep '12	90335	9	7
Sep '03	13802	1	1	Dec '12	93886	9	7
Dec '03	13622	1	1	Mar '13	84198	8	7
Mar '04	15314	1	1	Jun '13	87645	8	7
Jun '04	15746	1	1	Sep '13	103758	1	1
Sep '04	17864	1	1	Dec '13	103521	1	1
Dec '04	19714	1	1	Mar '14	95193	9	7
Mar '05	19840	1	1	Jun '14	96351	9	7
Jun '05	17784	1	1	Sep '14	96486	9	7
Sep '05	22893	2	2	Dec '14	80196	8	7
Dec '05	19899	1	1	Mar '15	56043	5	5
Mar '06	26448	2	2	Jun '15	65817	6	6
Jun '06	26166	2	2	Sep '15	64515	6	6
Sep '06	29550	2	2	Dec '15	56567	5	5
Dec '06	27771	2	2	Mar '16	49957	4	4
Mar '07	27399	2	2	Jun '16	53496	5	5
Jun '07	29493	2	2	Sep '16	59577	5	5

Sep '07	33402	3	3	Dec '16	61806	6	6
Dec '07	35880	3	3	Mar '17	67146	6	6
Mar '08	38697	3	3	Jun '17	64217	6	6
Jun '08	41579	4	4	Sep '17	68532	6	6
Sep '08	44787	4	4	Dec '17	73256	7	7
Dec '08	31563	3	3	Mar '18	84037	8	7
Mar '09	28362	2	2	Jun '18	91159	9	7

SATYAM COMPUTER SERVICES

Table 2:

Quarterly	Income from operations	First digit	First Digit(clubbed to get expected frequencies)	Quarterly	Income from operations	First Digit	First Digit (clubbed to get expected frequencies
Jun '97	29.6	2	2	Sep '04	848.1	8	6
Sep '97	37.95	3	3	Dec '04	891.26	8	6
Dec '97	57.58	5	4	Mar '05	953.36	9	6
Mar '98	64.54	6	5	Jun '05	1034.43	1	1
Jun '98	78.86	7	5	Sep '05	1117.27	1	1
Sep '98	91.85	9	6	Dec '05	1222.63	1	1
Dec '98	100.42	1	1	Mar '06	1259.97	1	1
Mar '99	107	1	1	Jun '06	1386.86	1	1
Jun '99	120.79	1	1	Sep '06	1537.71	1	1
Sep '99	157.04	1	1	Dec '06	1594.87	1	1
Dec '99	177.63	1	1	Mar '07	1709.03	1	1
Mar '00	205.72	2	2	Jun '07	1759.08	1	1
Jun '00	233.38	2	2	Sep '07	1948.24	1	1
Sep '00	272.49	2	2	Dec '07	2110.58	2	2
Dec '00	327.57	3	3	Mar '08	2319.38	2	2
Mar '01	386.55	3	3	Jun '08	2526.9	2	2
Jun '01	411.9	4	4	Sep '08	2700.52	2	2
Sep '01	426.63	4	4	Jun '10	1154.1	1	1
Dec '01	435.77	4	4	Sep '10	1150.8	1	1
Mar '02	457.64	4	4	Dec '10	1193.5	1	1
Jun '02	463.82	4	4	Mar '11	1277.7	1	1
Sep '02	499.14	4	4	Jun '11	1329.39	1	1
Dec '02	522.26	5	4	Sep '11	1478.8	1	1
Mar '03	538.43	5	4	Dec '11	1605.48	1	1
Jun '03	559.65	5	4	Mar '12	1550.54	1	1
Sep '03	598.49	5	4	Jun '12	1738.06	1	1
Dec '03	662.7	6	5	Sep '12	1780.87	1	1
Mar '04	720.71	7	5	Dec '12	1782.44	1	1
Jun '04	771.5	7	5	Mar '13	1787.48	1	1

Using the above data the following frequencies were obtained for Reliance Industries:

Table 3:

First Digit	1	2	3	4	5	6	7	8	9
Observed	16	8	5	4	11	12	3	8	7
Expected	22.3	13.0	9.2	7.2	5.9	5.0	4.3	3.8	3.4

For the application of Benford's law and to test the data using Chi Square Test of Goodness of Fit, the data size requirement is large. As per the condition of Chi Square test, no expected frequencies should be lesser than 5. Due to lack of availability of large amount of data, the expected frequencies are pooled together to meet this condition (Mohan, 2010) For example: In case of Reliance Industries, the first digits 7, 8 and 9 have expected frequencies as 4.3, 3.8 and 3.4 respectively which is lesser than 5. Hence, digits 8 and 9 are clubbed to 7 i.e. assuming 8 and 9 also as 7, resulting in an increase of expected frequency 11.5 i.e. the sum of 4.3, 3.8 and 3.4. The same process is done for Satyam as well where digits 4 & 5, 6 & 7 and 8 & 9 are clubbed.

The following is the result of Chi Square Test for Reliance Industries:

First_digit_Clubbed_to_get_expected_frequency

	Observed N	Expected N	Residual
1	16	22.3	-6.3
2	8	13.0	-5.0
3	5	9.2	-4.2
4	4	7.2	-3.2
5	11	5.9	5.1
6	12	5.0	7.0
7	18	11.5	6.5
Total	74		

Test Statistics

	First_digit_Clubbed_to_get_expected_frequency
Chi-Square	25.330[a]
df	6
Asymp. Sig.	.000

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 5.0.

Since the significance value is 0.000 is less than 0.05. Also, the chi square value 25.330 is greater than critical value k = 12.59. Hence we reject the null hypothesis and thereby concluding, the financials of Reliance Industries does not contain manipulations.

The following is the result of Chi Square Goodness of Fit test for Satyam Computer Services:

First_digit_Clubbed_to_get_expected_frequency

	Observed N	Expected N	Residual
1.0	27	17.5	9.5
2.0	8	10.2	-2.2
3.0	3	7.2	-4.2
4.0	11	10.2	.8
5.0	5	7.3	2.3
6.0	4	5.7	-1.7
Total	58		

Test Statistics

	First_digit_clubbed_to_get_expected_frequency
Chi-Square	9.377[a]
df	5
Asymp. Sig.	.095

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 5.7.

The significance value of the test i.e. 'p' value is .095 which is greater than significance value 0.05 and the chi square value of 9.377 is lesser than critical value 'k' = 11.07. Hence, we fail to reject the null hypothesis ultimately concluding that the financial data of Satyam Computer Services has been manipulated.
Similarly, frequencies for the second and third digit can also be computed depending on the availability of data.

5. FINDINGS AND SUGGESTIONS

Through this research, it can be agreed upon that Benford's law is one of the techniques to detect frauds. In the process, certain characteristics of Benford's law came to notice such as:

1. Applicable to large amount of data only.
2. Any number having 0 frequencies in actual observations will not procure appropriate results.
3. The slightest change in actual frequencies will get incorrect results catching the eye of the person evaluating it.
4. Particular events might lead to hike or fall in the financials, this might be detected as a fraud as per Benford's Law.

Furthermore, other statistical tests such as Z test and RMSE index may also be used for testing the accuracy of Benford's law.

6. CONCLUSION

Through this study, it is seen that the law has been applied on the frequencies of the actual financials which play a vital role in efficiently determining as to whether or not the financials consist of material misstatements which provides scope for the auditor to have random checks on the financials by taking the first digits instead of performing seemingly lengthy and redundant activities in identifying financial frauds by other means like physical verification, obtaining confirmations and clarifications from employees, which is very time consuming. Through using Benford's law, we can conveniently figure out if there exists any fraud misrepresentation in the financial statements through keeping a check on the frequency distribution of financial data which helps in improving investor's confidence in the quality of financial reporting since it is computed mechanically.

7. BIBLIOGRAPHY

Cleary, R., & Thibodeau, J. C. (May 2005). Applying Digital Analysis Using Benford's Law to Detect Fraud: The Dangers of Type I Error. Auditing: A Journal of Practise And Theory, 77-81.

Durtschi, C., Hillison, W., & Pacini, C. (January 2004). The Effective Use of Benford's Law to Assist in Detecting Accounting Data. Journal of Forensic Accounting, 17-34.

Fraud Statistics Every Business Should Know. (n.d.). Retrieved September 16, 2018, from www.quickbooks.in: https://quickbooks.intuit.com/r/trends-stats/fraud-statistics-every-business-should-know/

Geyer, D. (n.d.). Cheating Behavior and the Benford's Law. Nntes Graduate School of Management.

Johnson, G. G., & Weggenmann, J. (November 3 2013). Exploratory Research Applying Benford's Law to Selected Balance in the Financial Statement of State Government. Academy of Accounting and Financial Studies Journal, 31-44.

Kurger, P. S., & Yadavalli, Y. S. (August 2017). The Power Of One: Benford's Law. South African Journal Of Industrial Engineering , 1-13.

Mohan, R. (2010). A Text Book Of Statistics. Udupi: Benaka Books.

Tiscini, R. (n.d.). The relation between accounting frauds and corporate governance systems: an analysis of recent scandals. Luiss Guido Carli University.