Application of Benford's Law in Information Systems Auditing

Zdravko Krakar
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
Zdravko.Krakar@foi.hr

Mario Žgela
Croatian National Bank
Trg hrvatskih velikana 3, 10000 Zagreb, Croatia
Mario.Zgela@hnb.hr

Abstract*. Information systems auditing activities are obligatory for today's modern CIOs. There are numerous useful methods that can be conducted in such audits. One of incoming methods is application of so-called Benford's Law. In this paper we examine ways of application of this law in examination of certain number set with aim to make a conclusion if number set conforms to Benford's Law. As an examination target we used foreign payment system messages which are issued between foreign and domestic business entities i.e. business and central banks. We chose sample of 1.745.311 transactions and conducted examination for first, second and first two digits. Results prove adequate potential of this law in audit practice.*

Keywords. Benford's Law, Information systems auditing, foreign payment messages, data auditing

1 Introduction

Auditing of business activities are indispensable practice in all business systems. However, number of transactions enormously grow because of rising complexity of modern businesses which is caused by their mutual relationships. Consequently, amounts of data stored in databases exponentially grow. This growth leads to more complicated audits because it is harder to identify possible flaws, mistakes and frauds. This can easily be proved by blatant examples (Barings Bank in Singapore/UK, Allied Irish Bank, in UK Enron in USA, Societe Generale in France, Riječka banka in Croatia etc.) Those are the reasons of lately intensified development of business and especially information systems auditing. New auditing methods are developed which are based on use of information-communication technology, especially CAAT (Computer Assisted Auditing Techniques) tools.

Aim of this conference paper is to perform data audit of foreign payments i.e. foreign payment messages issued or received by Croatian commercial banks and central bank. Since these data are confidential, no information on any single payment message, any single bank or any single payer will be presented. Data that are included into in this examination are from period 1st of February to 1st of May 2008 which totalled in 1.745.311 payment messages.

Doubtlessly, exchange of various goods between Croatia and foreign economies is in constant focus of Croatian economists. Lot of economy analytics considers that information on this exchange shows status of Croatian economy. Changes in economy are often described and explained by changes in foreign exchange and values exported and/or exported. Not bearing in mind considerable dependency of national and world economies on these parameters, it is especially interesting to answer the question if it is possible to note whether there are some irregularities or deviations in creation of payment messages i.e. payment amounts.

Amounts in original currency, euros and Croatian kunas will be analyzed for certain groups of transactions (e.g. received payments for imported goods) as well as for all payment transactions.

Examination of conformity with Benford's law according to first digit, first two digits and second digit i.e. examination of frequencies of certain first, first two and second digits in payment amounts is included. Intuitively, it may be assumed that payment amounts are uniformly distributed which means that appearance of each digit on leading (most important, most valuable), first two and second positions are equal.

2 Research Method

Benford's Law was used as a primary data auditing method in above mentioned setting. Benford's Law defines expected digit frequencies in certain number sets. Also, it is often called „First Digit Law", „First Digit Phenomenon" and „Leading Digit Phenomenon". American astronomer Simon Newcomb was first who found out that numbers more frequently begin with smaller digits than with greater digits. Newcomb noticed that pages in logarithm tables were dirtier at the start, i. e. more used, and progressively cleaner as approaching to the end. He concluded that numbers more often begin with digit 1 than with any other digit, and in addition, that probability of each following digit (up to 9) at the most significant position in number progressively decreases.

Frank Benford gathered more than 20.000 observations from different sources (geographical area, population, river areas, physical constants etc.). He analysed frequencies of first digits for each number set. After he summarized all individual analyses he concluded that probability of first digit being 1 is 0,30103 which equals $\log_{10}2$, probability of first digit being 2 is 0,17609 which equals $\log_{10}3/2$ etc. Benford's Law of first digit i.e. probability V of appearance of digit z on leftmost position in number can be expressed by following formula:

$$V(z1)=\log_{10}(1+1/z1),\ z1\in\{1,2,...,9\} \quad (1)$$

Probabilities of each digit on the most significant position in number are shown in Table 1.

Table 1. Probabilities of each digit on the first position in number according to Benford's Law

digit (z1)	probability V(z1)
1	0,30103
2	0,17609
3	0,12494
4	0,09691
5	0,07918
6	0,06695
7	0,05799
8	0,05115
9	0,04576

It was explained that Benford's Law also applies to other significant digits in numbers. Furthermore, if second digit is observed, probability V of appearance of digit z on 2nd position in number is:

$$V(z2)=\log_{10}(1+1/1z2),\ z2\in\{1,2,...,9\} \quad (2)$$

Also, Benford's Law defines following rules of appearance of first two (V(z1z2)) and first three (V(z1z2z3)) digits in number:

$$V(z1z2)=\log_{10}(1+1/z1z2),\ z1z2\in\{10,11,...,99\} \quad (3)$$

$$V(z1z2z3)=\log_{10}(1+1/z1z2z3),\ z1z2z3\in\{100,101,...,999\} \quad (4)$$

This law is based on assumption that number set sorted ascending forms geometric series. Intuitive explanation of Benford's Law is pretty clear. If company with early turnover of 10.000 kunas is observed, first digit is 1. Digit 1 will stay on first position of turnover amount until turnover rise for 100%, which is 20.000 kunas. After this, only rise of 50% is needed in order to change first digit from 2 to digit 3. It is clear that early company's turnover will have digit 1 the most of the time because the most time is needed to change first digit from 1 to 2.

In [Nigrini] and [Nigrini & Mittermaier] criteria are set for number series to conform to Benford's Law:

1. Number series must describe values of same or similar phenomenon. E.g. lake area, heights of mountains, total yearly revenue of companies, total daily turnover on stock exchange.
2. Number series should not have defined minimal and maximal values. If minimal commission on foreign currency exchange in exchange office is 3 kunas, then set of commission values will not fit to Benford's Law, because large number of commission values will have digit 3 as a first digit. Digit 0 is allowed minimum.
3. Number series should not comprise of so called assigned numbers. These numbers are assigned to various phenomena instead of description, and their important attribute is that there is no sense to perform mathematical operations on these numbers. Examples are citizens identification numbers, bank account numbers, telephone numbers, numbers on car registration plates etc.
4. This law does not apply to numbers which creation is influenced by psychological factors, like prices in supermarket or ATM cash withdrawals.

Very important feature of Benford's Law is invariance. If certain number set fits Benford's Law, then the set will follow the law independently on measurement unit in which it is expressed. Consequently, if all numbers in a set that conforms to Benford's Law are multiplied by a constant, then new set will also conform to the law. For example, if the law is followed by set of total yearly companies' turnover, then the law will be followed independently on currency in which turnovers are expressed. Invariance rule also holds for reciprocal number sets. For example, if the law is followed by set of prices in kunas per stock, it will hold for numbers of stocks per kuna.

3 Status of knowledge about the problem

There is very extensive literature on various fields of usage of Benford's Law. Also, there are numerous works carried out on application of Benford’s Law in information systems auditing.

In [Nigrini and Mittermaier] it was shown in details how data auditing based on Benford’s Law was executed in accounting department. Audit observed 28.736 invoices authorized for payment by the accounts payable system of a business segment. After performing first digit test it was noted that the largest deviation is for the first digit 1, where the actual proportion exceeds expected proportion by 1,3 percent. However, with the results of Z-statistics it was concluded that the first digit test conformed to properties of Benford’s Law. Second digit test shown that digit 0 and digit 5 on the second position in number exceeded those of Benford’s Law. It was concluded, since invoices are the selling prices of goods and services, it is not surprising there are excess second digit 0 and 5. It was mentioned that separate digital analysis tests were performed on 1,6 million invoices of US Industrial conglomerate. Performed test shown nearly identical deviations so auditors can expect the same biases when analyzing invoices. First two digits test (FTD) showed significant excess of 5 two digits combinations which could be result of fraud, errors or processing inefficiencies. As an extension, number duplication test was performed. Authors were focused on 50 most frequent dollar values and additionally observed values that had excess in FTD test. Among other conclusions it was proved that some duplication of values were resulting in inefficient way of invoice processing.

In [Coderre] it was explained that FTD test may be used in contract audit. A comparison of the actual contract amounts to expected FTD frequencies according to Benford’s Law may highlight a higher than expected occurrence of contracts with amounts starting just bellow sole sourcing limit. Further investigation of these contracts may identify that individuals deciding on sole source contracting is directing them to friends or relatives.

In [Opaska] it is described how Benford’s Law principles were used in accounts payable disbursements audit in one federal agency. First digit test shows that leading digit 1 occurs with a substantially higher frequency than its predicted recurrence under Benford's law. Second step that was conducted was number duplication test for all values with digit 1 at leftmost position. It shown excess of one single value when compared to Benford’s Law frequencies.

However, we did not manage to find any literature which will indicate that research on usage of Benford's Law in auditing of foreign payment messages was carried out.

4 Results

We set the following hypotheses:
H1: Foreign payment messages amounts follow Benford’s Law.
H2: Benford’s Law can be used for auditing of foreign payment messages

Our research included payment messages in Croatian banking system in period between February 1st 2008 and May 1st 2008. Among others, the following are the most interesting data that were available:

- date
- amount (in original currency)
- amount in kunas
- amount in EUR
- code of payment purpose
- bank code
- type of payment

Total number of data rows of payment messages in observed period was 1.745.311.

We used Chi-square (χ^2) test in order to evaluate conformance of payment amounts with Benford's Law. With that test we wanted to show if eventual deviation of frequencies of observed phenomena from Benford's Law frequencies, are incidental or not. We set two null hypotheses and tested significance level on 5%.

Firstly we conducted first digit analysis of payment amounts in original currency. The results are shown in Table 2.

Table 2: First digit (d) analysis of amounts in original currency for all payments

D	Fo	fs	fs-fo	Po	χ^2
1	525.392	565.119	39.727	7,56	3.003,94
2	307.335	318.481	11.146	3,63	404,26
3	218.057	211.775	-6.282	-2,88	181,00
4	169.138	151.429	-17.709	-10,47	1.854,24
5	138.196	163.018	24.822	17,96	4.458,34
6	116.843	102.665	-14.178	-12,13	1.720,43
7	101.214	92.205	-9.009	-8,90	801,91
8	89.277	73.562	-15.715	-17,60	2.766,30
9	79.861	67.057	-12.804	-16,03	2.052,90
	Total:	1.745.311	POP	10,80	17.243,32

Explanation of symbols used:

fo – expected frequency – number of observations expected according to Benford's Law

fs – actual frequency

fs-fo – difference between actual and expected frequency

Po – percentage of deviation of actual from expected frequency

POP – average deviation from percentages of deviation (sum of absolute values Po divided by number of frequency categories i.e. 9)

Average deviation from percentages of deviation (POP) is used in [Nigrini & Mittermaier, 59] for intuitive explanation if certain number set conforms to Benford's Law. This measure does not have defined limit values i.e. range in which it can be stated whether deviation of value sets is significant or not. Furthermore, by means of average deviation from percentages of deviation it can not be judged if payment amounts conform to Benford's Law.

According to 8 degrees of freedom and testing on significance level of 5%, in order to confirm our first hypothesis, value of χ^2 should be less than 15,507. Since it is not the case, we have to reject assumption that number set of foreign payment amounts in original currencies conform to Benford's Law. It can easily be noted that significant (positive) deviations exist for digits 1 and 5 i. e. there is notable surplus of payment amounts in original currency with digits 1 and 5. It is not correct to immediately state that irregularities exist because of fraudulent behaviour or errors. For correct conclusion concerning this phenomenon additional data from business entities (companies, state organizations and agencies, persons) should be investigated. It may be assumed business entities pay their invoices and get paid in amounts which are significantly more often results of rounding and psychology of payer than exact calculations. It is clear from first digit analysis that banks are paying and receiving payment in amounts with first digits 1 and 5 more often than expected according to Benford’s Law. In order to perform more detailed and precise audit, more tests have to be executed. That will be performed with first two digits test (FTD) and second digit test of all payment amounts.

First two digits test shown enormously significant deviation of all first two digits divisible by 10. Deviation ranges from 53% for digits 90 to almost 430% for digits 50. In Figure 1 percentages of deviation of actual from expected frequency are shown for two digit combinations (10 - 99). Significant deviations can be easily spotted. Also, it is obvious that there are positive deviations for some digits divisible by 5 and not by 10 (15, 25, 35, 65, 75).

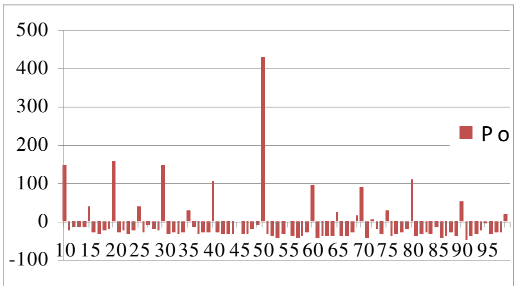

Figure 1. Frequency deviation from Benford’s Law for first two digits for all payments in original currency

In absolute value, there were 179.205 amounts in original currencies beginning with digits 10 which is 148% more than frequency according to Benford’s Law. Also, this is more than 10% of all amounts i.e. observed transactions. According to Benford’s Law, only 72.243 transaction amounts should begin with digits 10.

Deviations of numbers starting with digits 50 are the largest. In absolute value, there were 79.325 amounts in original currencies beginning with digits 50 which is almost 430% more than frequency according to Benford’s Law. However, this is only 4,5% of all amounts i.e. observed transactions. According to Benford’s Law, only 15.010 transaction amounts should begin with digits 50.

We examined this characteristic in more details by searching for rounded numbers starting with 10 (e.g. 10, 100, 1.000, 10.000 etc.). It was shown that there are 127.593 amounts in original currency which are members of set {10, 100, 1.000, 10.000, 100.000, 1.000.000, 10.000.000} which is more than 7% off all transactions.

Also, concerning rounded numbers starting with 50 search shown that there are 70.066 amounts in original currency which are members of set {50, 500, 5.000, 50.000, 500.000, 5.000.000, 50.000.000} which is 4% of all transactions.

Further, we searched for different type of payments and discovered that important percentage of rounded transactions and transactions starting with digits 10 or 50 are money transfers deposits to Croatia and to abroad as well as exchange of foreign currencies in Croatia. So, number set of amounts in exchange of currencies (buy or sell activities of kunas or foreign currencies) has 30.448 more amounts beginning with 10 than it should be expected according to Benford’s Law which is 353% difference. Also, there are 25.142 more occurrences of amounts of same type starting with digits 50 which is 14 times more than Benford's Law distribution for those digits (1.792). These deviations can be explained by psychology or custom in exchanging (buying or selling) rounded amounts or amounts starting with digits 10 or even more with digits 50. Also transfers of money from abroad to Croatia and from Croatia to abroad related with worker’s payments to their families, transfers connected with migration abroad or to Croatia, payments for memberships in organizations and associations and transfers of government funds from abroad that begin with digits 10 are also significantly higher than expected: 92% more than Benford’s Law distribution. Deposits and loans are expressed in rounded numbers much more than expected by Benford’s Law or uniform distributions. Deviation is almost 105%.

As it is already stated, these deviations can be explained by custom in number creation and influence of psychology when inventing numbers. For example, it is more usual to exchange (sell) 100 EUR for kunas then to exchange 98 EUR or 112 EUR. Or, Croatian workers temporarily or permanently working abroad will more often than it is expected by Benford's Law (and uniform) distribution send money amounts in rounded numbers or numbers beginning with digits 10 or 50 (e.g. 500 or 1.000 EUR) to their families residing in Croatia. Similar is with deposits or loans taken/given to/from abroad. This is why we noticed significant deviations in first digit test and first two digits test.

To furtherly clarify these conclusions, we performed second digit test. Second digit test shown that digit 0 on second position in amounts appears significantly more often than it should according to Benford's Law - for more than 150%. Digit 5 also appears significantly more on second position. It is noted that positive deviation is larger than 22%. This finding shows amounts are more often rounded on 0s and 5s than expected according to Benford's Law. It may be observed that business entities pay their amounts (or get paid), transfer or exchange their money in rounded numbers more often than it could be expected according to Benford's Law. This is of course possible, and not immediate proof of errors, misstatements or similar reasons.

We also audited payments related to buying or selling of goods abroad or in Croatia. We conducted first digit analysis of payment amounts in original currency. The results are shown in Table 3.

Table 3. First digit (d) analysis of amounts in original currency payments for goods

d	Fo	Fs	fs-fo	Po	χ^2
1	74.123	74.540	417	0,56	2,34
2	43.359	43.656	297	0,68	2,03
3	30.764	30.558	-206	-0,67	1,38
4	23.862	23.797	-65	-0,27	0,18
5	19.497	21.429	1.932	9,91	191,45
6	16.484	15.428	-1.056	-6,41	67,70
7	14.279	13.464	-815	-5,71	46,57
8	12.595	12.067	-528	-4,20	22,17
9	11.267	11.293	26	0,23	0,06
	Total:	246.232	POP	3,18	333,89

Deviations for first digits 1,2,3,4 and 9 are not significant. There is significant surplus of amounts starting with digit 5 (almost 10%) and deficit of amounts starting with digits 6, 7 and 8. Deficits can be explained by surplus for digit 5, i. e. digit 5 surplus caused deficits in some other digits. Since value for χ^2 (333,89) is still above cut off value (15,507), we must to conclude that this number subset does not follow Benford's Law. However, it can easily be noticed that deviations are significantly lesser than with previously mentioned types of messages.

First two digits test shown significant deviation of all first two digits divisible by 10. Deviation ranges from 26% for digits 90 to 135% for digits 50. In Figure 2 percentages of deviation of actual from expected frequency are shown for two digit combinations (10 - 99). Significant deviations can be easily spotted. Also, it is obvious that there are positive deviations for some digits divisible by 5 and not by 10 (15, 25, 75). Also, there is negative deviation for digits 85.

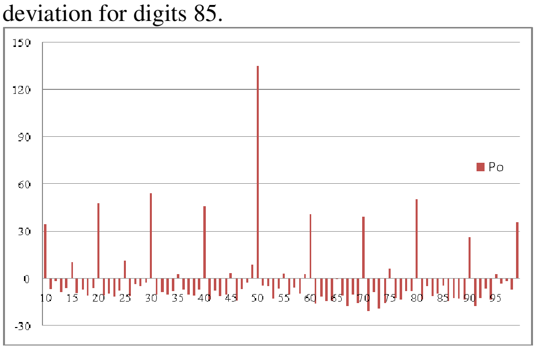

Figure 2: Frequency deviation from Benford's Law for first two digits of original currency payments for goods

Total amounts on invoices that correspond to value of contracts for goods may not be deviated in the same manner, i.e. only single payment amounts may be deviated according to contract terms. Business entities may agree to pay amounts in more than one rate and to apply payment method which will more often than expected according to Benford's Law produce rounded numbers or numbers starting with multiples of 10 for payments. To audit this and proof this assumption, we would have to choose certain number of business entities and to check their contracts and defined payment methods. However, this is beyond the scope of this paper and also beyond the capabilities of authors. Identification of business entities is done in each payment transaction for goods, so this is not impossible task for state authorities.

However, for purpose of audit it is possible to identify individual payer or payee and then perform Benford's Law tests on his/her received or ordered payments. If tests against Benford's Law significantly deviate, auditor may perform some additional tests only for certain digits which deviate from Benford's Law frequencies in order to check against fraudulent or erroneous activity:

- rounded amounts, amounts starting with multiples of 10 or 100
- counting frequency for each amount – searching for multiple occurrences (number duplication) of the same amounts (especially in short period of time)
- rounded and duplicated amounts on very same day
- rounded and duplicated amounts for the same type of payments in specific period of time
- rounded and duplicated amounts in specific or different banks
- amounts slightly below certain threshold – this can be done to check if payer or payee wanted to evade entering certain tax limits or additional checks by authorities (for

example, if person transfers amount above 100.000 kunas)

- amounts slightly greater than certain threshold – sometimes business entities with performance (money spent, invested, bought etc.) above certain limits may get special status in their relationships with state authorities, banks or other business entities
- multiple payments of same type paid same day (or in short time interval) with sum above certain threshold

Of course, for audit purpose it may be desirable to perform all these tests for all payments and then make conclusion. Then it is possible to choose among certain business entities which have the largest number of rounded and duplicated values or values slightly lesser or greater than certain threshold and perform detailed audit.

These tests can be quite easily performed by use of computer assisted auditing tools. These tools have a number of predefined auditing methods to ease data examination and conclusions on auditee.

5 Conclusion

Basic objectives of this paper were to examine, by use of computer auditing technology

- if Benford's Law applies to foreign payment messages and

- if Benford's Law can be used for auditing of foreign payment messages

For purpose of this work, we collected all foreign payment messages that were issued or received by Croatian banks in period 1st February 2008 to 1st May 2008. Our observation included 1.745.311 data rows.

Results show that foreign payment messages, when analyzed without focusing on special types of messages or certain business entities, do not conform to Benford's Law. We conducted additional examinations on specific message types and we explained deviations from Benford's Law frequencies.

Also, Benford's Law was used in auditing of foreign payment messages. It was shown how tests on duplicates and rounded numbers can be used in further explanation of results found by Benford's Law analysis. It can be concluded that application of Benford's Law is very effective in auditing of information systems, specifically foreign payment system.

References

[1] Aggarwal R.: **Persistent Puzzles in International Finance and Economics**, The Economic and Social Review, Vol. 35, No. 3, 2004, pages 241-250.

[2] Coderre D. G.: **CAATs and Other BEASTs for Auditors**, Global Audit Publications, Vancouver, Canada, 2001.

[3] Coderre D. G.: **Fraud Detection – A Revealing Look at Fraud**, Ekaros Analytical Inc., Vancouver, Canada, 2003.

[4] DeCeuster M. J. K., Dhaene G., Schatteman T.: **On the hypothesis of psychological barriers in stock markets and Benford's law**, Journal of Empirical Finance 5(3), 1998., pages 263-279.

[5] Doucouliagos C.: **Number preference in Australian stock prices**, School Working Papers – Series 2003, Deakin University, 2003.

[6] Durtschi C., Hillison W., Pacini C.: **The Effective Use of Benford's Law to Assist in Detecting Fraud in Accounting Data**, Journal of Forensic Accounting 1524-5586/Vol. V, 2004., pages 17-34.

[7] Hill T. P.: **The First Digital Phenomenon**, American Scientist, v86 n4, 1998., pages 1-6

[8] Koedijk K. G., Stork P. A.: **Should we care? Psychological barriers in stock markets**, Economic Letters 44, 1994., pages 427-432.

[9] Krakar Z., Žgela M.: **Evaluation of Benford's Law application in stock prices and stock turnover**, Society and Technology 2008, 26th to 28th June 2008, Zadar, Croatia.

[10] Lanza R. B.: **Proactively Detecting Occupational Fraud Using Computer Audit Reports**, The IIA Research Foundation, Altamonte Springs, USA, 2003.

[11] Ley E.: **On the peculiar distribution of the U.S. Stock Indices Digits**, Resources for the Future, Washington DC, USA, 1994.

[12] Matthews R.: **The Power of One**, New Scientist vol. 163 issue 2194, 1999., pages 26-30.

[13] Nigrini M. J.: **Digital Analysis Using Benford's Law**, Global Audit Publications, Vancouver, Canada, 2000.

[14] Nigrini M. J., Mittermaier L. J.: **The Use of Benford's Law as an Aid in Analytical Procedures**, Auditing - A Journal of Practice & Theory, Vol. 16, No 2, 1997., pages 52-67.

[15] Opaska W: **Survey of Accounts Payable Disbursements**, available at http://www.fcc.gov/Bureaus/Inspector_General/Reports/surveymemo.txt, Accessed: 27th March 2008.

[16] Turnbull, C. S.: **Fraud Investigation Using IDEA**, Ekaros Analytical Inc., Vancouver, Canada, 2003.