Coding Analytics

Lecture 7 – Fraud Detection

Module Leader: Dr Anthony Wood
Email: A.P.Wood@exeter.ac.uk
Location: 1.18a Streatham Court
Office Hours: Friday 10am - 12pm

Course Overview - Content

- Syllabus Plan
    - Introduction to Python
    - Handling data with the ‘pandas’ software library
    - Statistical Analysis
    - Visualisation
    - Machine Learning*
    - Textual Analysis*
    - Fraud Detection*
    - Social Network Analysis*

* Topics examinable within the assignment

Lecture outline

- Introduction to Benford law
- Tests of Conformity
- Academic Application
- Python analysis

Fraud Detection

Introduction to Benford Law

Fraud Detection

What is Fraud?

“Wrongful or criminal deception intended to result in financial or personal gain, or to deprive a victim of a legal right” (OED)

What is Financial Fraud?

Financial fraud can be broadly defined as an “intentional act of deception involving financial transactions for purpose of personal gain”. (OED)

Fraud is a crime, and can violate civil and criminal law. Punishment will vary depending on jurisdiction and the severity of the crime.

Fraud Detection

Fraud is a multi-billion-dollar business and it is increasing every year.

The PwC global economic crime and fraud survey of 2020 found that 47 percent of the 5000 companies they surveyed had experienced fraud of some kind. This is an increase from the PwC 2016 study in which slightly more than a third of organizations surveyed (36%) had experienced economic crime.

PwC survey can be found here

The total cost of these crimes?

“An eye-watering US$42 billion. That’s cash taken straight off companies’ bottom line. And 13% of those who’d experienced a fraud said they’d lost US$50 million-plus” (PwC 2020).

Worryingly, *“Too many organisations are failing to respond effectively. Only 56% conducted an investigation into their worst incident. And barely one-third reported it to the board”* (PwC 2020).

Fraud Detection

Forensic accounting

"The use of accounting skills to investigate fraud or embezzlement and to analyse financial information for use in legal proceedings" (OED).

- Forensic accountants collect, analyse, interpret and summarize complex financial data sets.
- They may be employed by insurance companies, banks, police forces, government agencies or public accounting firms.
- Forensic accountants compile financial evidence, develop computer applications to manage the information collected and communicate their findings in the form of reports or presentations.

They employ many data analysis tools, this lecture focuses on one such tool, **Benford Law**

Fraud Detection

Benford Law - the rule of leading digits

What is a leading digit?

The leading digit (or first digit) of any number is the first digit represented by a number other than zero.

Examples

Number	Leading Digit
2	2
245	2
0.254	2
7346	7
0.1098	1
0.00592	5

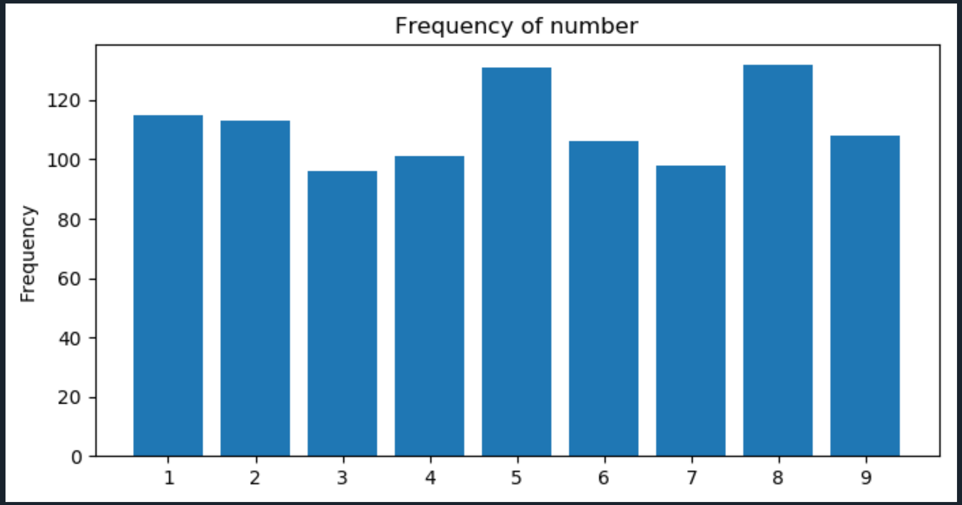
Fraud Detection

There are 9 possible outcomes (not including 0) in base 10 arithmetic.

1,2,3,4,5,6,7,8, and 9.

If we were to randomly select a digit then the probability of each digit occurring would be approximately 1 in 9 (or 11%).

Here I randomly generate 1000 numbers between 1 and 9 in python and plot the frequency of occurrence. (expected occurrence = 111)



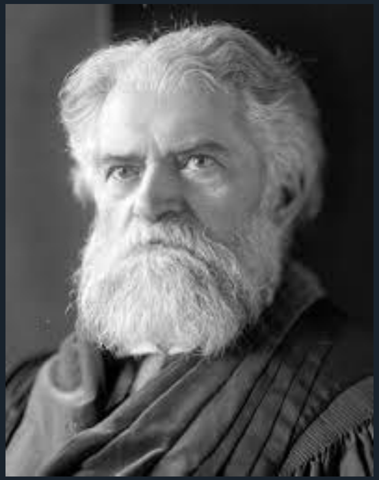
The rule simply states that each digit has an equal probability of occurring, given a random choice.

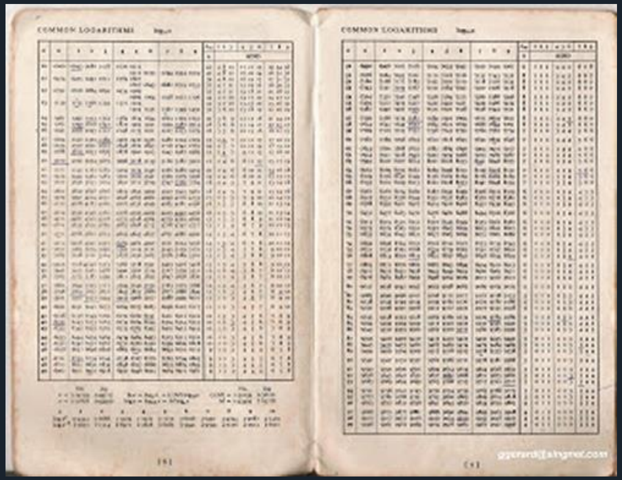
Fraud Detection

In reality, many (or most) naturally occurring number sequences and data sets DO NOT follow this rule.

The first person to notice that the fraction of numbers starting with digit D was not 1/9th but followed a logarithmic distribution was Simon Newcomb.

Newcomb (1881) noted that log table pages were grubbier and more worn out around the number 1 than the numbers 8 or 9. The frequency probabilities for each leading digit in his log book were shown as being…





Dig.	First Digit.
0	. . .
1	0.3010
2	0.1761
3	0.1249
4	0.0969
5	0.0792
6	0.0669
7	0.0580
8	0.0512
9	0.0458

…which (much) later became known as Benford law.

Fraud Detection

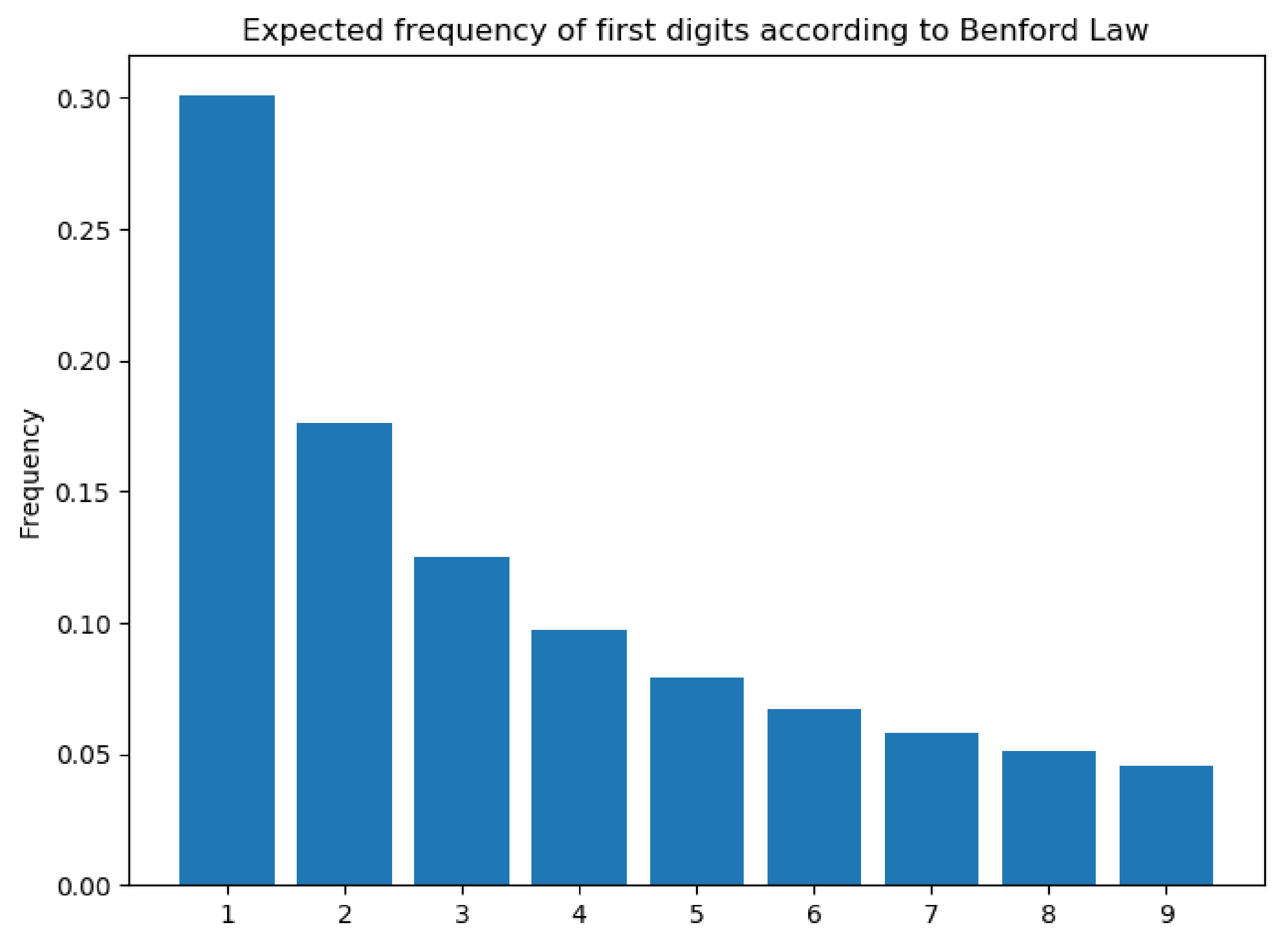
So why is it called Benford Law?

Frank Benford, a physicist at General Electric, rediscovered the pattern in 1930s and wrote a paper in 1938 that showed that distribution of first digits of 20,229 sets of numbers from the areas of rivers to physical constants and death rates followed the law:

$$Frequency\ of\ leading\ digit\ d = log_{10}(1 + 1/d)$$

Leading first digit, d	1	2	3	4	5	6	7	8	9
Occurrence probability, P_d	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046

Fraud Detection



Examples Found (Benford, 1932)

TABLE I

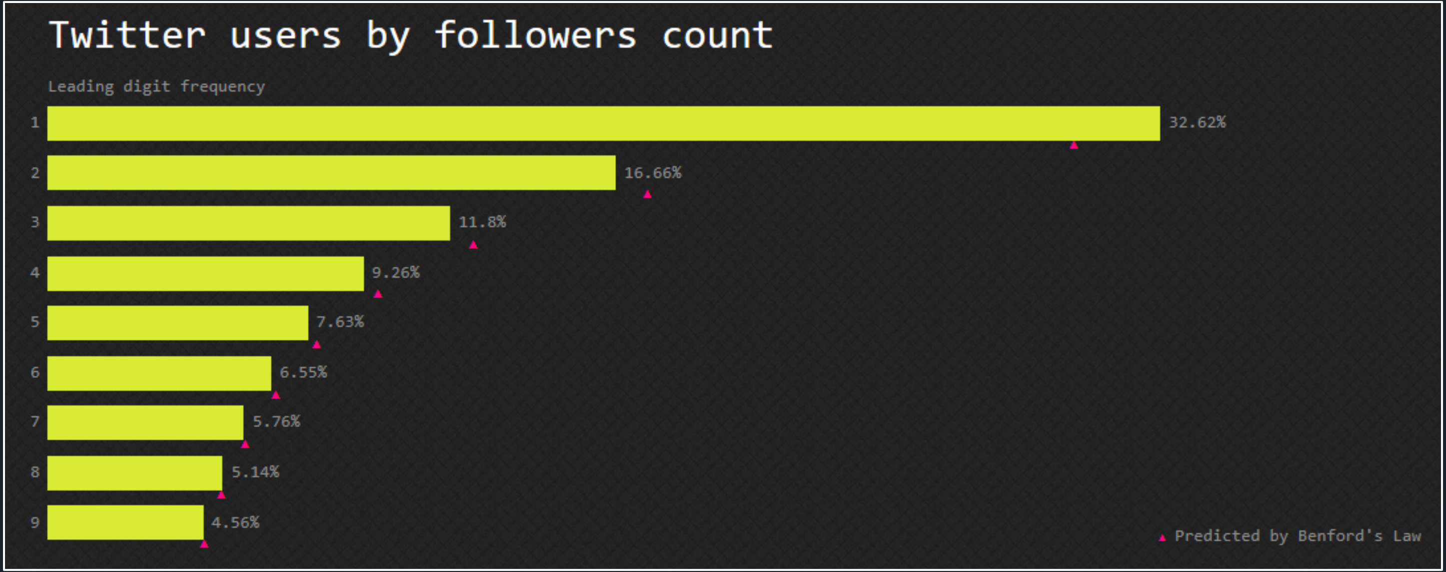
PERCENTAGE OF TIMES THE NATURAL NUMBERS 1 TO 9 ARE USED AS FIRST DIGITS IN NUMBERS, AS DETERMINED BY 20,229 OBSERVATIONS

Group	Title	First Digit 1	2	3	4	5	6	7	8	9	Count
A	Rivers, Area	31.0	16.4	10.7	11.3	7.2	8.6	5.5	4.2	5.1	335
B	Population	33.9	20.4	14.2	8.1	7.2	6.2	4.1	3.7	2.2	3259
C	Constants	41.3	14.4	4.8	8.6	10.6	5.8	1.0	2.9	10.6	104
D	Newspapers	30.0	18.0	12.0	10.0	8.0	6.0	6.0	5.0	5.0	100
E	Spec. Heat	24.0	18.4	16.2	14.6	10.6	4.1	3.2	4.8	4.1	1389
F	Pressure	29.6	18.3	12.8	9.8	8.3	6.4	5.7	4.4	4.7	703
G	H.P. Lost	30.0	18.4	11.9	10.8	8.1	7.0	5.1	5.1	3.6	690
H	Mol. Wgt.	26.7	25.2	15.4	10.8	6.7	5.1	4.1	2.8	3.2	1800
I	Drainage	27.1	23.9	13.8	12.6	8.2	5.0	5.0	2.5	1.9	159
J	Atomic Wgt.	47.2	18.7	5.5	4.4	6.6	4.4	3.3	4.4	5.5	91
K	$n^{-1}, \sqrt{n}, \cdots$	25.7	20.3	9.7	6.8	6.6	6.8	7.2	8.0	8.9	5000
L	Design	26.8	14.8	14.3	7.5	8.3	8.4	7.0	7.3	5.6	560
M	*Digest*	33.4	18.5	12.4	7.5	7.1	6.5	5.5	4.9	4.2	308
N	Cost Data	32.4	18.8	10.1	10.1	9.8	5.5	4.7	5.5	3.1	741
O	X-Ray Volts	27.9	17.5	14.4	9.0	8.1	7.4	5.1	5.8	4.8	707
P	Am. League	32.7	17.6	12.6	9.8	7.4	6.4	4.9	5.6	3.0	1458
Q	Black Body	31.0	17.3	14.1	8.7	6.6	7.0	5.2	4.7	5.4	1165
R	Addresses	28.9	19.2	12.6	8.8	8.5	6.4	5.6	5.0	5.0	342
S	$n^1, n^2 \cdots n!$	25.3	16.0	12.0	10.0	8.5	8.8	6.8	7.1	5.5	900
T	Death Rate	27.0	18.6	15.7	9.4	6.7	6.5	7.2	4.8	4.1	418
Average.......		30.6	18.5	12.4	9.4	8.0	6.4	5.1	4.9	4.7	1011
Probable Error		±0.8	±0.4	±0.4	±0.3	±0.2	±0.2	±0.2	±0.2	±0.3	—

Fraud Detection

Other examples can be found in:

- Stock data
- Accounting data
- Numbers on the internet
- Speed of light
- Gravitational force
- Internet traffic
- Survey and response data
- Ebay bids



This website has some visual examples of various widely-used large data sets.

Fraud Detection

How to use

- Benford Law can be used to analyse data and identify red flags (identify suspicious data).
- By analysing numbers within annual reports (for example) we can use Benford Law to determine whether the numbers follow Benford Law or not.
- If the numbers do not follow Benford Law then it is likely that the numbers have been fabricated.
- Humans do not do a very good job in replicating this distribution when fraudulently doctoring numbers.
- (see for example my paper on academic fraud and the case of Professor Hunton).
- Benford Law will not apply if the data is truncated or limited in any way (e.g. surveys - rate this module 1-5)

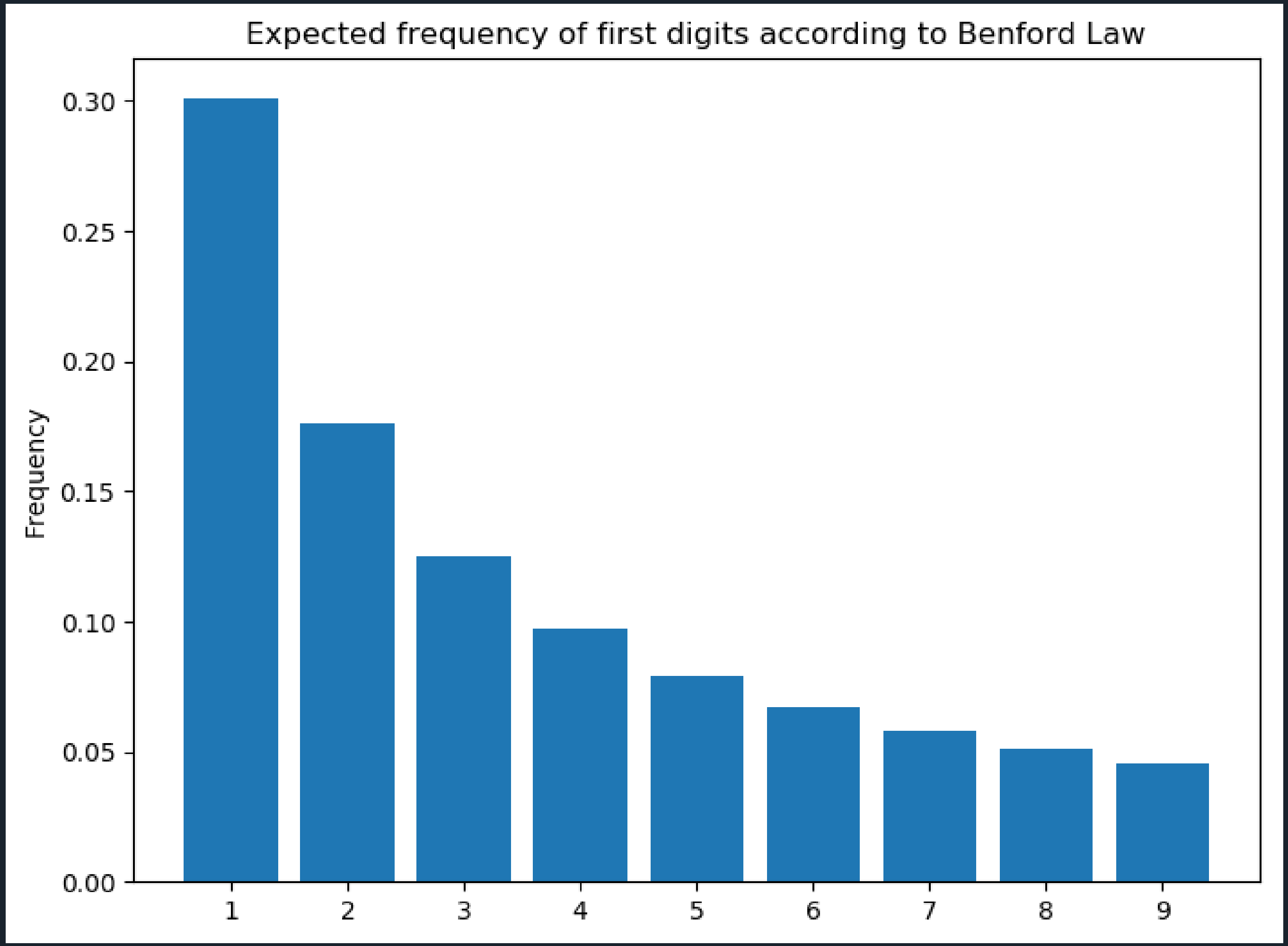
How do we determine if a particular data set conforms to Benford Law or not? – see next section!

Further information, examples and guidance

Post any questions to ELE!!

End of Section

Fraud Detection



Fraud Detection

Measuring conformity to Benford Law

(1) Mean Absolute Deviation (MAD)
(2) Kolmogorov-Smirnoff test (KS)
(3) Chi-Square test

These test provide a single statistic regarding the whole sample, from which we can observe whether the data conforms to Benford Law, or not.

Fraud Detection

The MAD score

A common measure of conformity to BL is the mean absolute deviation (*MAD*) score (Nigrini, 2012). The *MAD* score is defined as the mean of the absolute value of the difference between the frequency of each first digit within the sample, and the frequency as determined by BL.

$$\frac{\sum_{i=1}^{K}|AF - EF|}{K}$$

Where *AF* is the actual frequency of the leading digit observed, *EF* is the expected frequency as determined by BL, and *K* is the number of leading digit bins (equal to 9 for the first leading digit).

Fraud Detection

The MAD score (Example)

$$MAD = \frac{\sum_{i=1}^{K}|AF - EF|}{K}$$

Digit	Actual Frequency AF (Sample)	Expected frequency EF (Benford)	|AF-EF|
1	0.287	0.301	0.014
2	0.185	0.176	0.009
3	0.113	0.125	0.012
4	0.089	0.097	0.008
5	0.074	0.079	0.005
6	0.071	0.067	0.004
7	0.068	0.058	0.010
8	0.059	0.051	0.008
9	0.054	0.046	0.008
		MAD	**0.009 (rounded up)**

Fraud Detection

MAD - Rules and Regions

The MAD score does not provide a statistical test of significance, but provides a guideline as to the level of conformity.

A generally accepted rule is defined below.

- 0.0000 to 0.006: Close Conformity
- 0.006 to 0.012: Acceptable Conformity
- 0.012 to 0.015: Marginally Acceptable Conformity
- Above 0.015: Nonconformity

Therefore our MAD score of 0.009 would put us in the “Acceptable Conformity” region.

Fraud Detection

The KS score

The second measure of conformity is the Kolmogorov-Smirnov (*KS*) statistic which is defined as the maximum cumulative deviation from the theoretical distribution of BL, for leading digits 1-9 and is calculated as follows:

$$KS = \mathrm{Max}(|AF_1 - EF_1|, |(AF_1+AF_2)-(EF_1+EF_2)|, \ldots, |(AF_1+AF_2+..+AF_9)-(EF_1+EF_2+..+EF_9)|$$

The advantage that *KS* has over *MAD* is that we can empirically test conformity to Benford Law using critical values of the Kolmogorov-Smirnov test statistic.

At 5% significance, the critical value is $1.36/\sqrt{P}$ assuming $P>35$, where P is the total number (*Pool*) of first digits used in the calculation of *KS*.

At 1% significance, the critical value is $1.63/\sqrt{P}$

When *KS* is greater than the critical value then we can infer that the distribution does not follow Benford Law.

Fraud Detection

The KS score (Example)

$$KS = \mathrm{Max}(|AF_1 - EF_1|, |(AF_1+AF_2)-(EF_1+EF_2)|, .., |(AF_1+AF_2+..+AF_9)-(EF_1+EF_2+..+EF_9)|$$

Digit	**Actual Frequency AF (Sample)**	**Expected frequency EF (Benford)**	**AF1+…Afn (a)**	**EF1+…EFn (b)**	**|a-b|**
1	0.287	0.301	0.287	0.301	0.014
2	0.185	0.176	0.472	0.477	0.005
3	0.113	0.125	0.585	0.602	0.017
4	0.089	0.097	0.674	0.699	0.025
5	0.074	0.079	0.748	0.778	**0.030**
6	0.071	0.067	0.819	0.845	0.026
7	0.068	0.058	0.887	0.903	0.016
8	0.059	0.051	0.946	0.954	0.008
9	0.054	0.046	1	1	0
				KS (MAX)	**0.030**

Fraud Detection

KS Critical value and inference

Sample (Pool) size $P = 2426$

Critical value (at 5% significance) $= 1.36/\sqrt{P} = 1.36/\sqrt{2426} = 0.027$

Inference:
When *KS* is greater than the critical value then we can infer that the distribution does **not** follow Benford Law.

0.030 (KS) is greater than 0.027 (CV)

Therefore our data <u>DOES NOT</u> follow Benford Law in this example (at the 5% significance level).

Fraud Detection

The Chi-Square test

The third measure of conformity is the Chi-Square test of significance.

The Chi-Square test compares the counts for each digit to the expected count derived from Benford Law and is calculated as follows:

$$Chi - Square\ test\ statistic\ \chi^2 = \sum_1^9 \frac{(AC - EC)^2}{EC}$$

where *AC* is the actual count and *EC* is the expected count

When *the* Chi-square test statistic is greater than the critical value then we can infer that the distribution does not follow Benford Law.

Fraud Detection

The Chi-Square test (Example)

$$Chi - Square\ test\ statistic\ \chi^2 = \sum_1^9 \frac{(AC - EC)^2}{EC}$$

Digit	Actual Count	Expected* Count	AC - EC	(AC - EC)^2	(AC-EC)^2 / EC
1	697	730	-33	1104	1.511
2	449	427	22	485	1.136
3	267	303	-36	1314	4.333
4	216	235	-19	373	1.586
5	180	192	-12	136	0.708
6	174	163	11	131	0.807
7	166	141	25	640	4.546
8	144	124	20	411	3.322
9	133	112	21	458	4.105
				SUM (Chi-Square)	**22.057**

* Pool size is 2426

Expected count for digit 1 = 2426*0.301 ≈ 730

Fraud Detection

Chi-Square Critical value and inference

Critical value (at 5% significance) = 15.51
(8 degrees of freedom)

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21

Inference:
When *Chi-square test statistic* is greater than the critical value then we can infer that the distribution does not follow Benford Law.

22.057 is greater than 15.51
Therefore our data <u>DOES NOT</u> follow Benford Law according to this test (at the 5% significance level).

Fraud Detection

Which test to use?

Our results

MAD Score - data conforms to Benford Law
KS Score - data does not conform to Benford Law
Chi-Square - data does not conform to Benford Law!

In the Accounting literature, the MAD score is a popular choice. The KS score is also being used more frequently in recent years, given that there is no “grey” area of assessment associated with the MAD score.

In practice it is better to perform several tests for confirmatory purposes rather than relying on any one single measure.

note that other measures exist beyond these three, but these are most prevalent in the literature to date.

Fraud Detection

Benford Law - beyond the first digit

This lecture details first digit conformity to Benford Law.

There are also tests for the 2nd digit, 3rd digit… nth digit, following the general rule

$$\boldsymbol{Frequency\ of\ digit\ d\ as\ the\ nth\ digit} = \sum_{k=10^{n-2}}^{10^{n-1}-1} \boldsymbol{log_{10}(1 + \frac{1}{10k+d})}$$

e.g. the probability of the number 2 occurring as the second digit

$$log_{10}\left(1+\frac{1}{12}\right) + log_{10}\left(1+\frac{1}{22}\right) + \cdots log_{10}\left(1+\frac{1}{92}\right) = 0.109$$

Fraud Detection

Full table of occurrences

Digit	0	1	2	3	4	5	6	7	8	9
1st	N/A	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046
2nd	0.120	0.114	0.109	0.104	0.100	0.097	0.093	0.090	0.088	0.085
3rd	0.102	0.101	0.101	0.101	0.100	0.099	0.099	0.099	0.099	0.098
4th	0.100	0.100	0.100	0.100	0.100	0.099	0.099	0.099	0.099	0.099
5th	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100

The distribution quickly approaches a uniform distribution as the nth digit increases.

Fraud Detection

Benford Law - beyond the first digit

One final extension to the law is the possibility to calculate the probability of encountering any number starting with a string of digits n of any length (discarding leading zeros), using the formula:

$$log_{10}(1 + \frac{1}{n})$$

For example, the probability of the occurrence of 524 at the beginning of a number (excluding leading zeros) is..

$$log_{10}\left(1 + \frac{1}{524}\right) = 0.000828$$

Further information, examples and guidance

Post any questions to ELE!!

End of Section

Fraud Detection

Academic Application

Fraud Detection

Amiram, Bozanic and Rouen (2015)

Key Findings

- Conformity to Benford Law is an indication of financial statement data quality.
- Restated financial statements more closely conform to Benford Law than mis-stated versions in the same firm-year.
- As divergence from Benford Law increases, earnings persistence decreases.
- Benford Law can be used as a predictive indicator for financial mis-statements.

Fraud Detection

Amiram, Bozanic and Rouen (2015)

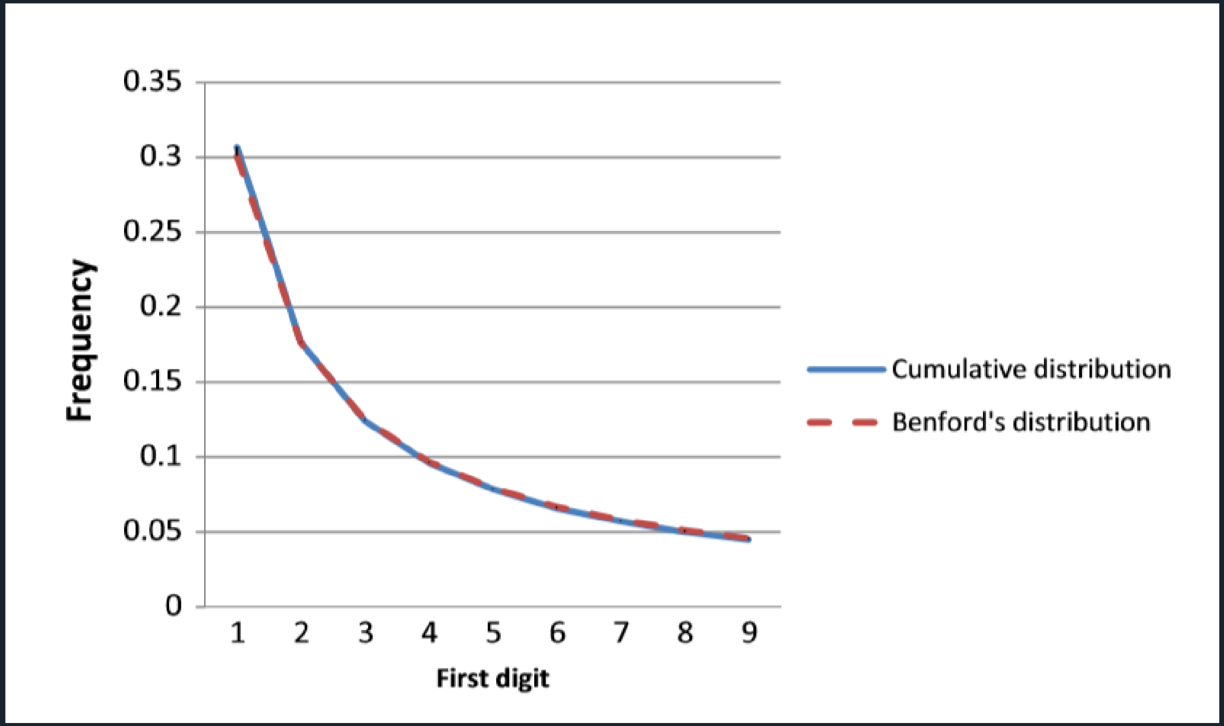
Data

- 43,332 firm-years of data for US companies from the period 2001-2011
- All Compustat variables that appear on the Balance Sheet, Income Statement, and Statement of Cashflow.
- Companies with less than 100 items, and those with negative total assets are removed.

Fraud Detection

Amiram, Bozanic and Rouen (2015)

On Average financial statements conform to Benford Law – Both for the sample as a whole, and when aggregated by year or industry.



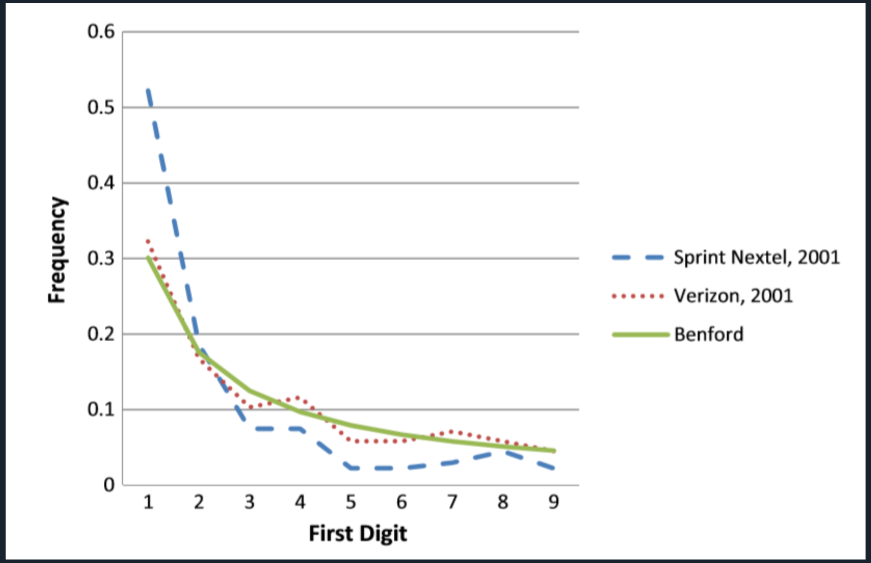
Fraud Detection

Amiram, Bozanic and Rouen (2015)

Individual firms vary in their adherence to Benford Law.

Here the company Sprint Nexal does NOT conform to BL with a KS score of 0.224 and a MAD score of 0.052

Verizon on the other hand DOES conform, with a KS score of 0.056 and a MAD score of 0.017.



- Note that whereas the KS score can statistically test whether the distribution conforms or not (comparing the statistic to a critical value), The MAD score cannot. The MAD score for Verizon is 0.017 which according to the general rule would put it in the non-conformity region.
- The Authors note that because of the lack of a statistical test for the MAD score, the KS method is preferred.
- The mean MAD score for the whole sample is 0.0009, and the mean KS score is 0.0296.

Fraud Detection

Amiram, Bozanic and Rouen (2015)

Firms conforming to Benford Law.

Table 3 Firm-year conformity to Benford's Law

Firm-years conforming	Percent conforming
Panel A: Number of firm-years conforming to Benford's Law	
37,104	85.63

Financial statement	Firm-years conforming	Percent conforming
Panel B: Number of firm-years conforming to Benford's Law by financial statement		
Balance sheet	39,274	90.64
Income statement	34,138	78.78
Cash flow statement	42,259	97.52

Note how the income statement numbers are less likely to conform

Fraud Detection

Restatements.

Sometimes a firm may restate (update, correct, amend) its annual reporting figures, sometime after the initial report has been published.

This may be for various reasons including:

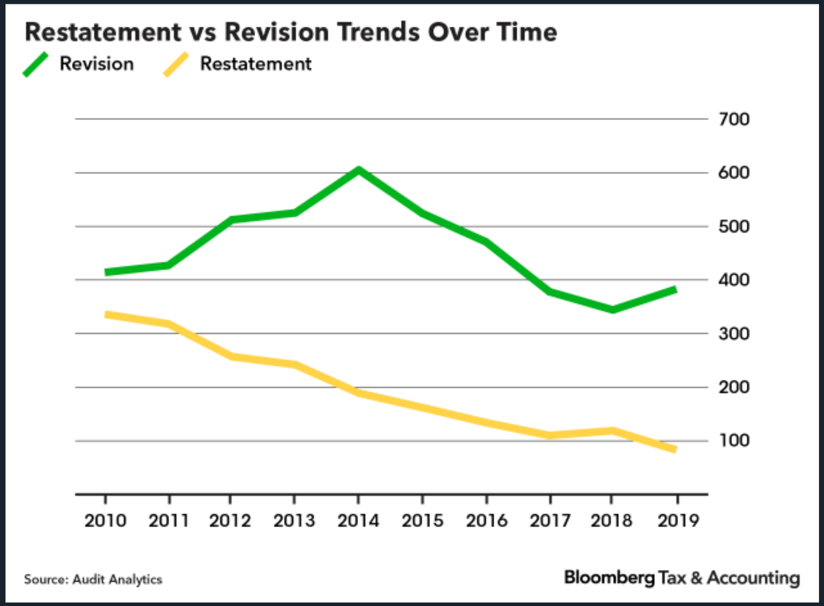
- The original statement contained a “material” inaccuracy
- Non-compliance with Accounting principles
- Administrative error
- Misrepresentation
- Fraud

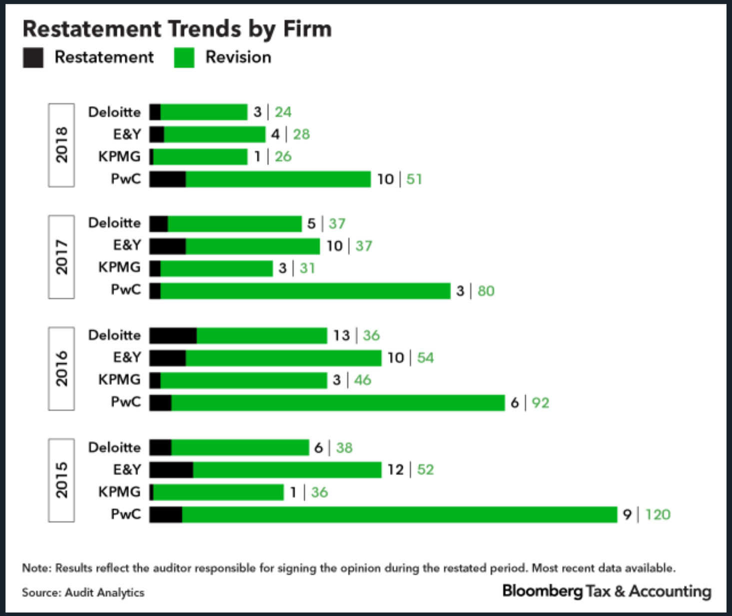
Fraud Detection

Restatements.

Bloomberg (2020) has noted a decline in restatements over the past 10 years citing an increase in audit quality since the 2008/9 financial crisis.

They also note that companies are more likely to report revisions in subsequent accounts rather than producing a full set of restated accounts for the same year.





Fraud Detection

Amiram et al (2015)

Misstated accounts and loss making firms are more likely to deviate from Benford Law.

Table 7 FSD_Score and Ex post measures of earnings manipulation

	FSD_Score	Number of firm-years	*t* statistic
RESTATED_NUMS = 0	0.0289	4935	5.36***
RESTATED_NUMS = 1	0.0280	4935	
LOSS = 0	0.0289	27,743	23.98***
LOSS = 1	0.0310	15,589	

Fraud Detection

Amiram et al (2015)

The authors cite that the firms which deviate from Benford Law tend to be:

- Smaller
- Younger
- More volatile
- Growing

Benford Law can be used to detect errors in Financial Reports

These errors may not be the direct result of deliberate fraudulent misrepresentation or fabrication, but can act as a “red flag” and identify those companies whose accounts may not be materially representative and warrant further inspection.

Further information, examples and guidance

Post any questions to ELE!!

End of Section

Fraud Detection

Python Analysis

Fraud Detection

Benford Law can be analysed in Python using the package benford_py.

Please see the documentation here for installation, documentation, and demonstration.

pip install benford_py from the anaconda prompt

benford_py allows us to easily calculate MAD, KS and the Chi-squared statistic from a given set of numbers.

Fraud Detection

Example

The file named “benford test data.csv” contains the total asset (*at*) values for 19114 companies.

Do the total asset values follow Benford Law?

To answer this we need the *benford_py* and *pandas* packages…

..and read in python the csv file (I like to also create a list of the numbers which I want to analyse).

```
import benford as bf
import pandas as pd


data = pd.read_csv("benford test data.csv")

mylist = data["at"].tolist()
```

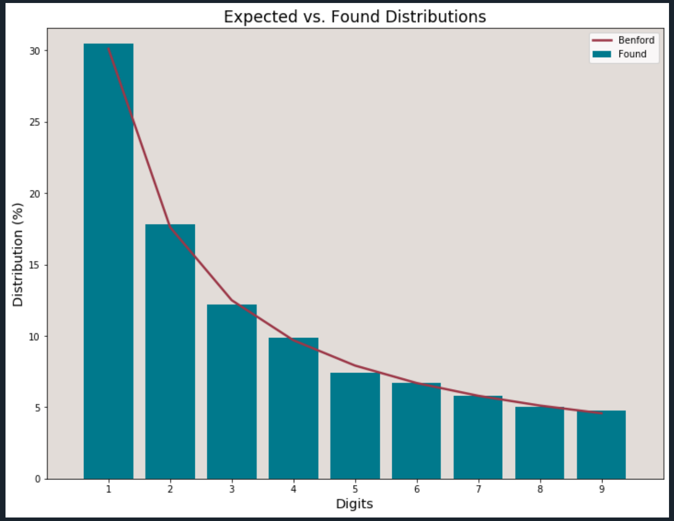
Fraud Detection

Example

Using the benford_py command “bf.first_digits” we can quickly count and plot the first digits of the total asset values.

```
f1d = bf.first_digits(mylist, digs=1, decimals="infer")
```

Initialized sequence with 19114 registries.
First_1_Dig
1 0.304855
2 0.178403
3 0.121952
4 0.098462
5 0.073977
6 0.066810
7 0.058073
8 0.050120
9 0.047347



<u>It looks like the data follows Benford Law!!</u>

the option “decimals” is set to “infer” (choose for each observation). I do this because not all observations have the same number of decimal places in this data. The default option is 2 decimals.

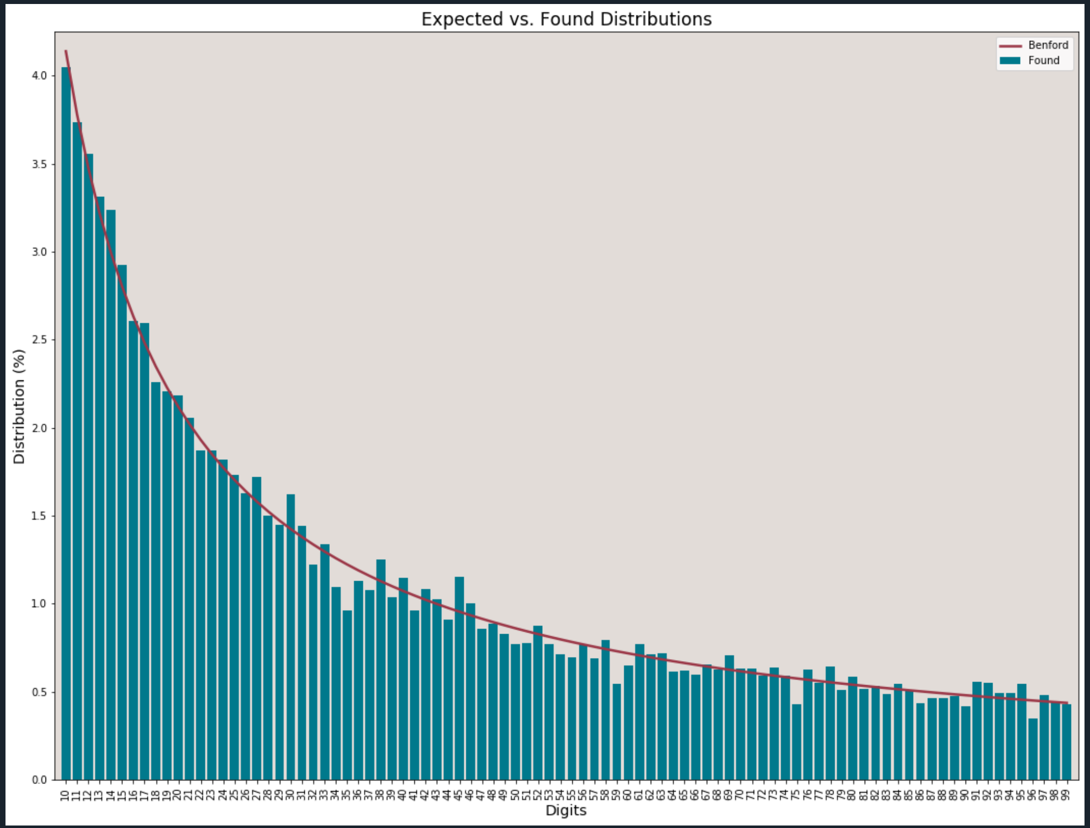
Fraud Detection

Example

We could also analyse the first 2 digits using the same method

```
f2d = bf.first_digits(mylist, digs=2, decimals="infer")
```

Again.. This looks like it follows Benford Law!!



Fraud Detection

Example

We can also perform some tests using additional options.

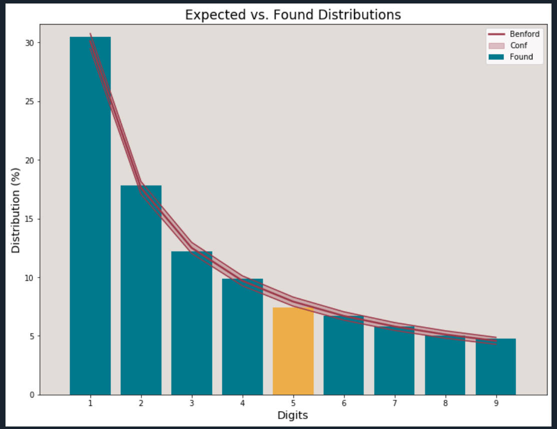
e.g. Calculate the MAD, KS and Chi-squared test statistics, along with their critical values (95% confidence).

```
f1d = bf.first_digits(mylist, digs=1, decimals="infer", MAD=True, chi_square=True, KS=True, confidence=95, show_plot=True)
```

```
Mean Absolute Deviation: 0.002080
MAD <= 0.006000: Close conformity.


The Chi-square statistic is 11.3477.
Critical Chi-square for this series: 15.507.

The Kolmogorov-Smirnov statistic is 0.0061.
Critical K-S for this series: 0.0098
```



Both test statistics are below their critical values and therefore the data conforms to Benford Law.

Fraud Detection

Example

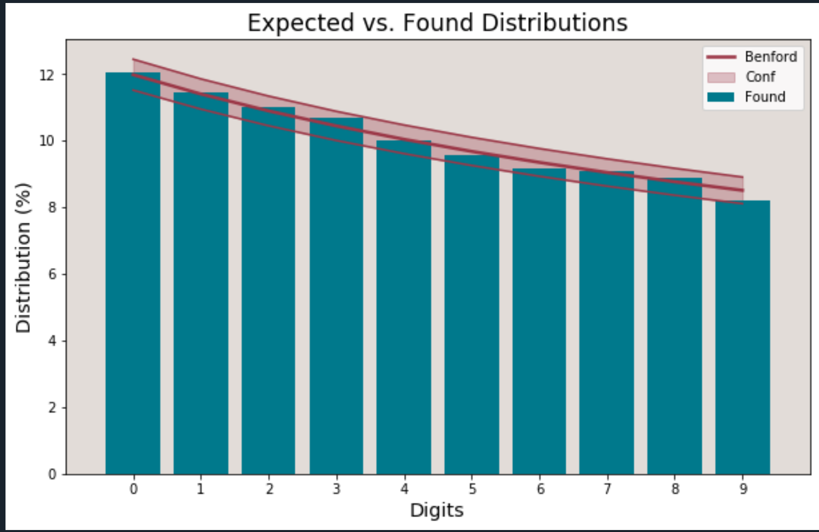
Does the second digit comply with Benford law?

```
sd = bf.second_digit(mylist, decimals="infer", MAD=True, chi_square=True, KS=True, confidence=95, show_plot=True)
```

```
Mean Absolute Deviation: 0.001271
MAD <= 0.008000: Close conformity.


The Chi-square statistic is 4.8216.
Critical Chi-square for this series: 16.919.

The Kolmogorov-Smirnov statistic is 0.0048.
Critical K-S for this series: 0.0099
```



Answer: Yes!

Fraud Detection

Summary

- Benford Law is the analysis of leading digits which should follow a particular distribution
- Benford Law can be applied to any number of data sets provided they are not confined or bounded.
- Humans do not do a very good job at replicating Benford Law
- MAD, KS and Chi-squared tests can be performed to identify “red flags” within data to point to potentially suspicious behaviour.

Fraud Detection

Useful references:

Benford, F. 1938. The law of anomalous numbers. Proceedings of the American Philosophical Society 78: 551–572.

Newcomb, S. 1881. Note on the Frequency of the Different Digits in Natural Numbers. American Journal of Mathematics 4(1): 39-40.

Hill, T.P. 1998. The first digital phenomenon: A century-old observation about an unexpected pattern in many numerical tables applies to the stock market, census statistics and accounting data. American Scientist 86(4): 358-363.

Amiram, D., Z. Bozanic, and E. Rouen. 2015. Financial statement errors: Evidence from the distributional properties of financial statement numbers. Review of Accounting Studies 29(4): 1540-1593.

Horton, Kumar & Wood. 2020. Detecting Academic Fraud Using Benford Law: The Case of Professor James Hunton. Forthcoming. Available on SSRN.

Nigrini, M. 2012. Benford's law: Applications for forensic accounting, auditing, and fraud detection. Hoboken, N.J.: Wiley.