BENFORD'S LAW AND THE BIBLE

Denka Markova
1306 Wilson Street
Arkadelphia, AR 71923
denka_markova@hotmail.com

Linda Njoh
Box 5446
Arkadelphia, AR 71999
faith_lm@yahoo.com

Michael Lloyd, Ph.D. (mentor)
Box 7782
Henderson State University
lloydm@hsu.edu

Abstract

Benford's Law states that the distribution of the first digits in numbers obtained from real-life sources follows a particular non-uniform distribution. In this paper, we determined that each of the first five books of the Bible does not obey this law. However, some of their distributions resemble Benford's Law.

Introduction

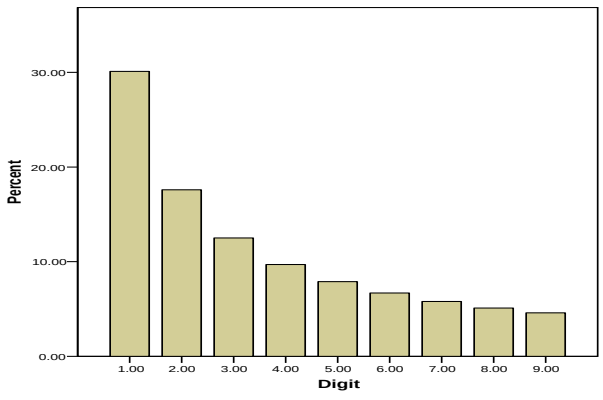
Benford's Law is also known as the first-digit law. It states that in list of numbers from many real life sources of data, the leading digit 1 occurs more often than the others. This digit occurs with a probability of 30%, which is more than the expected 11.1% if all the digits from 1 to 9 had the same probability. The larger the digit, the less likely it is to appear as the leading digit of a number. The probability of the first digit being d is $P[\text{first digit is } d] = \log_{10}(1+1/d)$.

Probabilities for Each Digit according to Benford's Law

Digit	1	2	3	4	5	6	7	8	9
Probability of Obtaining for 1st Digit (%)	30.1	17.6	12.5	9.7	7.9	6.7	5.8	5.1	4.6

Graph of the Benford's Law Probabilities

If we graph the probabilities we will get the following graph whose shape is similar to an exponential decay curve.



This remarkable theorem applies to figures obtained from the natural world and also from numbers taken from electric bills, newspaper articles, street addresses, population counts, death rates, and so on. It can be used for ferreting out suspicious frauds, embezzlers, tax evaders and even computer bugs.

Data

We only had time to consider the first five books of the Bible: Genesis, Exodus, Leviticus, Numbers, and Deuteronomy. Recall that these are collectively called the Books of Moses or the Pentateuch. We searched for numbers using an online version of the New International Edition of the Bible on the site http://www.biblegateway.com. We used the "Match ANY" option in its search engine and used the following keywords:

> one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fifteen, twenty, thirty, forty, fifty, hundred, thousand, first, second, third, fifth, ninth, twelfth

Note that we did not have to include numbers like sixty or sixty-five since the six or the five will catch them. The search engine found 1847 cases, which are too many to list this paper. However, a sample of this data is shown in the table here. The variables were book, the chapter:verse, the number in the verse, and the first digit.

Sample of Data Set

	Book	Verse	Number	First Digit
1	GEN	1:05	1	1
2	GEN	1:08	2	2
4	GEN	1:13	3	3
6	GEN	1:19	4	4
12	DEUT	2:07	40	4
13	DEUT	2:14	38	3
16	DEUT	3:04	60	6
17	DEUT	3:08	2	2
18	DEUT	3:11	13	1

Distributions

The next three figures summarize the occurrence of numbers in each book and the frequency of first digits in the entire the Pentateuch. Note that Numbers really has more numbers than any of the other books in the Pentateuch.

For the entire Pentateuch, notice that the percentage of 1's is 15 percentage points higher than what is predicted by Benford's Law. Visually, this distribution resembles Benford's Law, but there are too many ones and sevens.

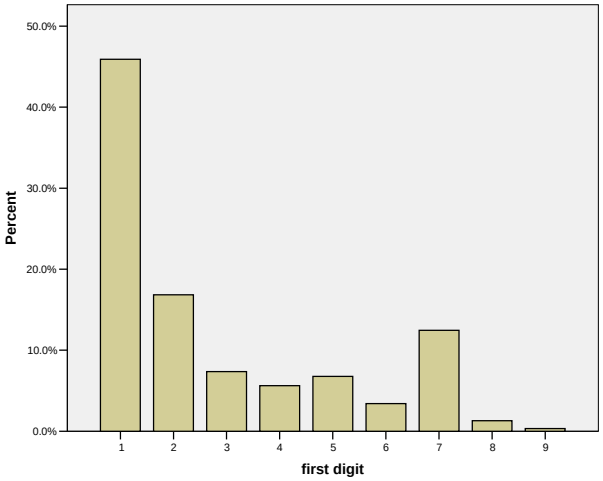
Distribution of Numbers in the Pentateuch

Books	Frequency	Percent
Genesis	376	20.4
Exodus	472	25.6
Leviticus	257	13.9
Numbers	556	30.1
Deuteronomy	186	10.1
Total	**1847**	**100.0**

Distribution of First Digit in the Pentateuch

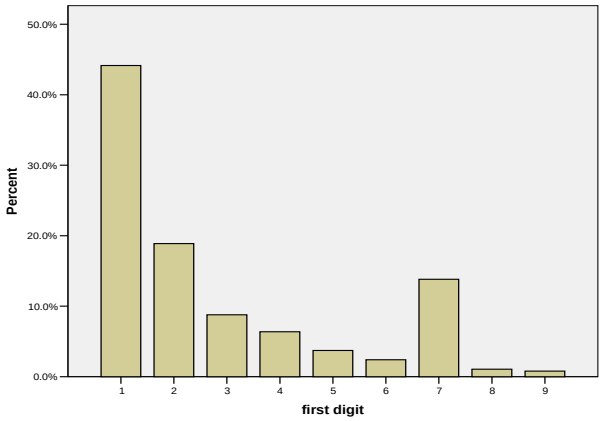
First Digit	Frequency	Percent
1	848	45.9
2	311	16.8
3	136	7.4
4	104	5.6
5	125	6.8
6	63	3.4
7	230	12.5
8	24	1.3
9	6	.3
Total	**1847**	**100.0**

Distribution in the Pentateuch

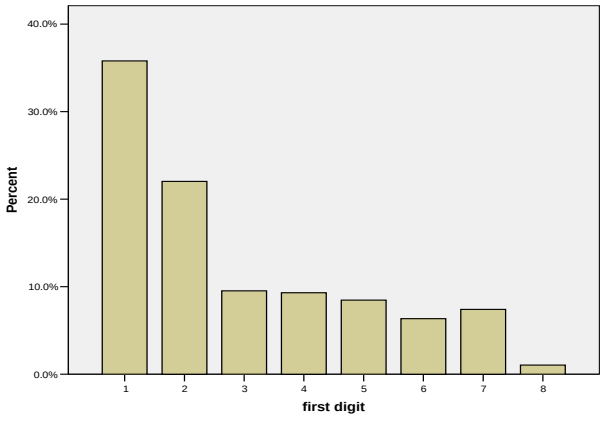


We will now examine the distribution of the first digits in each book separately. Each book has too many ones and Genesis and Exodus have too many sevens. Exodus appears to be closest to having the Benford's Law distribution, although it has no nines.

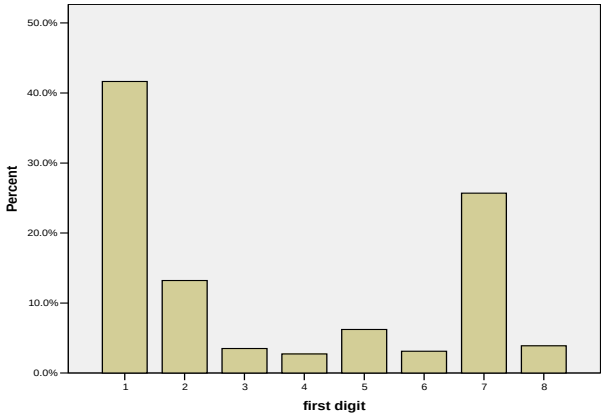
Distribution in Genesis



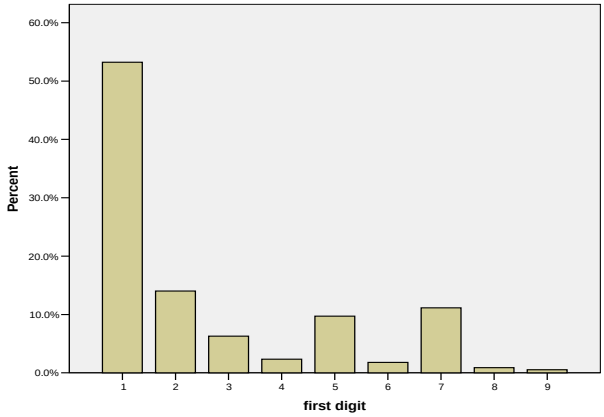
Distribution in Exodus



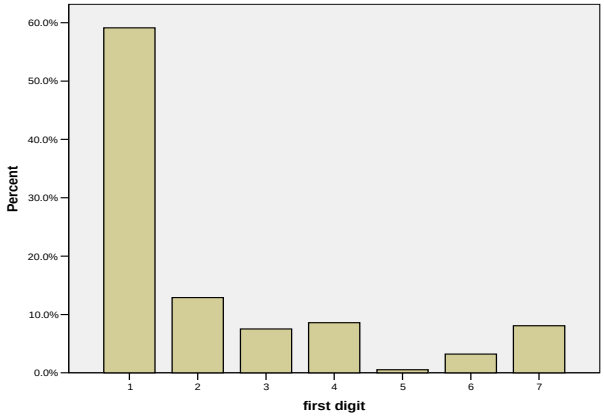
Distribution in Leviticus



Distribution in Numbers



Distribution in Deuteronomy



Leviticus also has no nines and Deuteronomy has neither eights nor nines. Assuming that books sample the first digit randomly and that they obey Benford's Law, the probability that these digits are omitted is very small. For example, the probability that Leviticus has no nines is $(1-0.046)^{257} = 6\times10^{-6}$.

Chi-Square Goodness-of-Fit Test

The chi-squared goodness-of-fit test will be performed separately on each book in the Pentateuch using the hypotheses:

H_0: The distribution of the first digits fits Benford's law.
H_a: The distribution of the first digits does not fit Benford's law.

The p-value for each chi-squared statistic was 0.000 which is highly significant. Therefore, we fail to accept the null hypothesis. That is, we have strong evidence that the distribution of the first digits in each of the books of Moses do not fit Benford's Law. Note that Exodus comes closest fitting Benford's Law since it has the smallest chi-squared value.

Chi-Square Test Statistics and Degrees of Freedom

Book	Test Statistic	df
Genesis	117.8	8
Exodus	28.8	7
Leviticus	214.5	7
Deuteronomy	67.0	6
Numbers	240.0	8

Chi-Square Test for Homogeneity

We used this test to determine whether the distributions of the first digits in each of the books were homogeneous. The test hypotheses are:

H_0: The distributions in each book are the same.
H_a: There are at least two books which have different distributions.

The Chi-square test statistic is 195.2 with degrees of freedom 32. Since the p-value was 0.000, the data are significant and therefore we are failing to accept the null hypothesis. That is, the distributions for all the books are not the same.

Procedural Errors

There are several mistakes that we made and correcting them may change the previous statistical test results.

- Add *pair* to our Boolean for the search
- Account for numbers being duplicated when referring to the same object.
- Eliminate some of the ones because they are not being used in a quantitative sense. For example, Genesis 15:2 reads “But Abram said, ‘O Sovereign LORD, what can you give me since I remain childless and the one who will inherit my estate is Eliezer of Damascus?’” The word "one" here is not being used in a numerical sense, but merely means "the person".

Further Research

- Add another variable *level*, which will record whether a number is being used in an ordinal or ratio sense.
- Analyze the entire Bible.
- Redo the chi-squared goodness-of-fit tests excluding the sevens.

References

- "Probability and Statistical Inference", by Robert Hogg and Elliot Tanis, 7th Edition (c)2005 by Prentice Hall
- "Statistics" by Richard Weimer. (c)1993 Wm. C. Brown
- BIBLE.COM, http://www.bible.com/
- BibleGateway.com, http://www.biblegateway.com/
- MathPages, http://www.mathpages.com/
- York University Department of Mathematics and Statistics, http://www.math.yorku.ca