USING BENFORD'S LAW TO DETECT FRAUD

(NO. 02-5410)

VII. THE DIGIT TESTS

The five digit tests developed by Mark Nigrini mentioned earlier are described in detail as follows. They are listed in the order in which they would normally be performed.

- The first digit test
- The second digit test
- The first two digits test
- The first three digits test
- The last two digits test

The first and second digit tests are high-level tests of reasonableness and are used to determine whether the dataset appears reasonable. If the first and second digit tests indicate that the dataset is significantly different than Benford's Law, the first two digits and first three digits tests will be performed to select audit targets. The last two digits test is used to detect rounding.

The *first-order tests* include the first digit test, the second digit test, and the first two digits test.

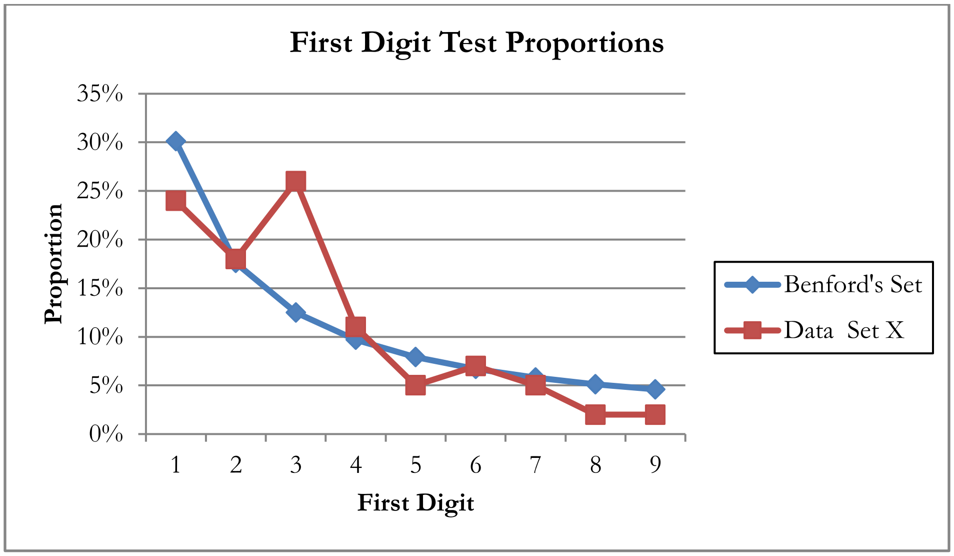
The first-order tests are usually run on the positive numbers or the negative numbers, but not both in the same analysis. This is because the incentive to manipulate is opposite for positive and negative numbers. For example, if a company has a positive net income, management is looking to manipulate this figure and will likely want to inflate it. On the other hand, if the company is reporting a net loss, management might want this amount to be as close to zero as possible.

First Digit Test

The *first digit test* compares the actual first digit frequency distribution of a dataset with that developed by Benford. It is an extremely high-level test and will only identify obvious anomalies (i.e., it will only point you in the right direction). It should not be used to select targets for sampling, as the sample size will be too large. For example, consider the following graphical comparison of a Benford set (first digit frequencies conform to Benford's Law) and another sample dataset (Dataset X). Assume Dataset X represents the first digit frequency of 10,000 vendor invoices.

First Digit	Benford's Set	Dataset X	Deviation
1	30.1%	24%	0.06
2	17.6%	18%	0.00
3	12.5%	26%	-0.14
4	9.7%	11%	-0.01

First Digit	Benford's Set	Dataset X	Deviation
5	7.9%	5%	0.03
6	6.7%	7%	0.00
7	5.8%	5%	0.01
8	5.1%	2%	0.03
9	4.6%	2%	0.03



As shown above, the first digit frequency distribution of Dataset X does not conform to Benford's Law based on the graph alone. However, as can be seen from the graph, significant anomalies occur with the first digit 3—invoice amounts beginning with 3 appear 26% of the time, as opposed to the 12.5% figure calculated by Benford. Other anomalies occur in which the actual frequency is less than the expected frequency. However, fraud examiners are concerned with the overuse of digits, because fraudsters, when inventing numbers, tend to overuse certain digit patterns. The digits that occur fewer times than Benford's Law predicted (1, 5, 7, 8, and 9) result primarily from the overuse of 3.

Because Dataset X consists of 10,000 invoices, fraud investigators would need to examine approximately 2,600 invoices (10,000 x .26) to see first-digit frequency distribution for Dataset X in the graph. Given the fact that this requires a substantial amount of time, this test should not be used to select the audit sample. Rather, it should only be used as a benchmark.

Second Digit Test

The *second digit test* is also a high-level test designed to test conformity or reasonableness. Remember that expected second digit proportions are less skewed than expected first digit proportions. Because this test results in a large sample selection, it should not be used to select audit samples. However, it can be used to quickly identify potential problems in a dataset, especially if one assesses conformity using the Z-statistic as discussed in the previous chapter.

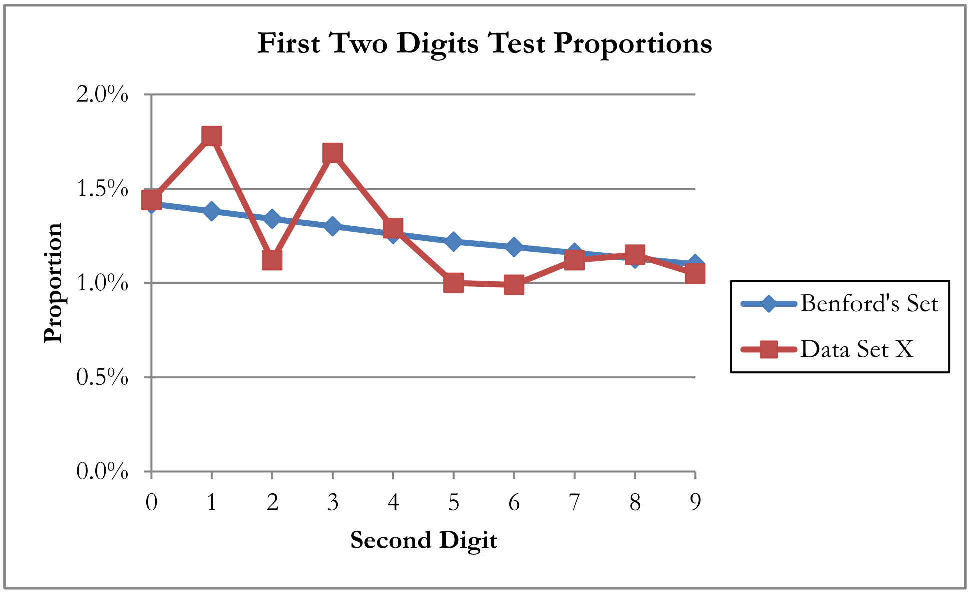
First Two Digits Test

The *first two digits test* combines the previous two tests and identifies manifested deviations that warrant further review. To that end, it can be used to select efficient audit samples for testing. For example, consider the sample data in the first digit test above. An abundance of invoices beginning with 3 were detected. In fact, based upon that review, approximately 2,600 invoices need to be examined. However, using the first two digits test, it is apparent that not all the invoices need to be examined (see the following table and chart).

Instead, only those invoices beginning with the first two digits 31 and 33 need to be examined. As seen in the chart, these are the first two digits whose actual frequencies differ the most from their expected frequencies (-.40 and -.39, respectively). Therefore, if numbers beginning with 31 or 33 were focused on, 347 (178 + 169) invoices need to be reviewed. This was calculated by multiplying 1.78% (actual frequency percentage for the first two digits 31) by 10,000 (number of total invoices) and adding that to 1.69% (actual frequency percentage for the first two digits 33) times 10,000 (number of total invoices). This test results in a required audit sample of more than 2,000 fewer invoices—certainly a more efficient and focused sample.

NOTE: The following chart assumes the first digit of the multi-digit number was 3.

Second Digit	Benford's Set	Dataset X	Deviation
0	1.42%	1.44%	-.02
1	1.38%	1.78%	-.40
2	1.34%	1.12%	.22
3	1.30%	1.69%	-.39
4	1.26%	1.29%	-.03
5	1.22%	1.00%	.22
6	1.19%	0.99%	.10
7	1.16%	1.12%	.04
8	1.13%	1.15%	-.02
9	1.10%	1.05%	.05



First Three Digits Test

The *first three digits test* is a highly focused test that is also used to select audit samples. While the first two digits test tends to indicate broad categories of abnormality, such as payments made just below an authorized limit, the first three digits test tends to identify unusual amounts that have been duplicated. An example has not been provided in graphical format here because there are 900 possible three-digit combinations and such a graph is too large to include on this page. However, this test will most likely produce a sample that is smaller than that determined by the first two digits test. Therefore, in the example, the audit sample is less than 347.

Both the first two and first three digits tests tend to identify overused digit patterns that indicate fraud, represent erroneous inputs, or suggest the duplicate processing of the same invoice on multiple occasions.

Nigrini also warns that it is imperative to delete all data records fewer than ten, as these numbers are usually immaterial and might skew the results.

Additionally, Nigrini states that all negative numbers (credit memos) should be removed from the dataset when investigating potentially fraudulent invoices because they represent the inverse of an invoice amount.

Last Two Digits Test

The last two digits test is used to identify fabricated and rounded numbers. This test is especially handy because it might be all the fraud examiner needs to select audit targets in populations smaller than 10,000. Because the expected proportion of all possible last two digit combinations is .01, it is easy to identify abnormalities via a graph. This test is especially useful if financial statement figures have been rounded, thereby suggesting that the figures are estimates rather than actual amounts.

Because this test results in small and efficient sample sizes, it can be used to identify patterns that might not be evident when using the previous four tests.

Applying the Z-Statistic to the Digit Tests

As previously stated, the first and second digit tests are high-level tests of reasonableness and can be used to judge general conformity to a Benford curve, but they should not be used to select a sample. If these tests indicate the data is significantly different than a Benford curve, then fraud examiners should continue with the first two digits and first three digits tests to select audit targets. The last two digits test is used to detect artificially rounded numbers.

But how is it judged whether a set of data conforms to Benford's Law? Sometimes the dataset lends itself to using professional judgment to assess conformity. That is, a few digits will have significantly larger deviations from the expected proportion than others, and no calculations are necessary. However, sometimes it is not as clear which digits are statistically significant. In these instances, fraud examiners should use an empirical method to determine which digits to focus on.

The previous chapter discussed the Z-statistic, which is used to assess the conformity of a dataset to Benford's Law. The Z-statistic can be applied to any of Nigrini's digit tests listed above. Recall that the formula to calculate the Z-statistic looks like this:

$$Z = \frac{|AP - EP| - \left(\frac{1}{2N}\right)}{\sqrt{\frac{EP(1-EP)}{N}}}$$

Remember that the (1/2N) term is a continuity correction term and is only used when it is smaller than the first term in the numerator.

When using the Z-statistic, fraud examiners can select a level of significance (5% was the example used here) that represents how much variance is acceptable. Another way of saying this is, if a 5% level of significance was selected, then 95% of the data must conform to a Benford curve for the data to be acceptable. Recall that the cutoff level for a 5% level of significance is 1.96.

For example, assume the second digit test is being performed. Here is an example of the detailed output of a fictitious set of data (assume the first digit for this set of data is 4 and that "Expected Proportion" represents Benford's Law):

Digit	Count	Actual Proportion (AP)	Expected Proportion (EP)	Deviation	Z-Statistic
0	220	10.7%	12.0%	-1.3%	1.74
1	268	13.1%	11.4%	1.7%	**2.34**
2	260	12.7%	10.9%	1.8%	**2.55**
3	200	9.8%	10.4%	-0.6%	0.93
4	205	10.0%	10.0%	0.0%	0.01
5	199	9.7%	9.7%	0.0%	0.00
6	214	10.4%	9.3%	1.1%	1.73
7	165	8.0%	9.0%	-1.0%	1.47
8	140	6.8%	8.8%	-2.0%	**3.12**
9	180	8.8%	8.5%	0.3%	0.41
TOTAL	**2,051**	1	1		

If the digits that are statistically significant at 95% are chosen, then all numbers beginning with 41, 42, and 48 will be examined, as these do not conform to Benford's Law. This is determined because the Z-statistics for the second digits 1, 2, and 8 are greater than 1.96.

Recall from the previous chapter that the Z-statistic suffers from the excess power problem. As a dataset grows, the Z-statistic tolerates smaller and smaller deviations. To illustrate this, take the set of data above but multiply each count by 10 so that there is a dataset of 20,510 units rather than 2,051, and all the digits have the same proportion as before:

Digit	Count	Actual Proportion (AP)	Expected Proportion (EP)	Deviation	Z-Statistic
0	2,200	10.7%	12.0%	-1.3%	**5.60**
1	2,680	13.1%	11.4%	1.7%	**7.50**
2	2,600	12.7%	10.9%	1.8%	**8.15**
3	2,000	9.8%	10.4%	-0.6%	**3.03**
4	2,050	10.0%	10.0%	0.0%	0.02
5	1,990	9.7%	9.7%	0.0%	0.01
6	2,140	10.4%	9.3%	1.1%	**5.58**
7	1,650	8.0%	9.0%	-1.0%	**4.77**
8	1,400	6.8%	8.8%	-2.0%	**9.97**
9	1,800	8.8%	8.5%	0.3%	1.41
TOTAL	**20,510**	1	1		

The statistically significant digits now include 0, 1, 2, 3, 6, 7, and 8. Either this data does not conform to Benford's Law and the analysis should not be used, or the excess power problem is to blame. One solution to overcoming the excess power problem is to ignore the actual numerical value of the Z-statistics. In other words, only concentrate on the largest statistics without worrying about how large they are. In this instance, fraud examiners might choose to focus on just 1, 2, and 8 (the same digits from the previous analysis), which is far more reasonable.

Summary

This section discussed how Mark Nigrini pioneered a new application of an old mathematical phenomenon, Benford's Law, to assist auditors and investigators in identifying potentially fraudulent transactions. He showed that digital analysis is an easily applied, cost-efficient method of uncovering fraud and other errors.

Nigrini identified five tests that can be used either proactively or reactively to test for fraudulent transactions, inefficiencies, rounded numbers, and duplicate payments. These digit tests include:

1. The first digit test
2. The second digit test
3. The first two digits test

4. The first three digits test
5. The last two digits test

The first and second digit tests are high-level tests designed to assess overall conformity and detect obvious anomalies. Because they are so high-level, these tests should not be used to select an audit sample. The first two digits test, on the other hand, combines these two tests and identifies clear deviations that need to be investigated further. Therefore, it can be used to select efficient audit samples for testing.

The first three digits and last two digits tests are also used to select audit samples. While the first two digits test tends to indicate broad categories of abnormality, such as payments made just below an authorized limit, the first three digits test tends to identify unusual amounts that have been duplicated. The last two digits test is primarily used to identify rounded numbers.

This chapter also explained how to apply the Z-statistic analysis from the previous chapter to the digit tests. This analysis is used to determine a dataset's conformity with Benford's Law. When using the Z-statistic, be cautious about the excess power problem; sometimes a set is so large that the Z-statistic only tolerates minor fluctuations and a dataset might be inappropriately rejected.