# Do Populations Conform to the Law of Anomalous Numbers?

Frédéric SANDRON*

The “first significant digit” of a number is its leftmost non-zero digit. For example, the first significant digit of the number 325 is 3 and the first significant digit of 0.8732 is 8. It might be expected that the first significant digits of any given series of numbers, or of a set of numbers measuring any given phenomenon, are randomly distributed. Nothing of the sort: in most series found in the real world, figure 1 appears more often than figure 2, which in turn appears more often than figure 3, and so on. The purpose of this note is to illustrate this rule, known as Benford’s law, using data for the populations of all world countries, and to show its underlying logic, which in this particular case, relies on the pattern of population growth.

## I. Benford’s law

In 1881, the mathematician and astronomer Simon Newcomb noticed that the first volumes of the tables of logarithms in the library of his institution were more worn than subsequent ones. This meant that there were more frequent consultations for numbers starting with 1 or 2 than for numbers starting with 8 or 9. In 1938, the engineer Benford made the same observation about the same tables of logarithms, independently of Newcomb’s work. In his article, Benford compiled numerous sets of data, from physical constants to baseball results and various number series he found in newspapers. The average number of appearances of the first significant digits followed a logarithmic law already proposed by Newcomb (Bedford, 1938):

$$F_d = \log_{10}(1 + 1/d)$$

— where $F_d$ is the frequency of appearance of the first significant digit $d$.

Benford’s law predicts that first significant digits will be distributed as described in Table 1.

*Institut de recherche pour le développement, Laboratoire Population-Environnement, Université de Provence.

Translated by Sarah R. Hayford.

TABLE 1.– PREDICTED FREQUENCY OF APPEARANCE OF THE FIRST SIGNIFICANT DIGIT ACCORDING TO BENFORD'S LAW

| First significant digit | Predicted frequency |
|---|---|
| 1 | 0.301 |
| 2 | 0.176 |
| 3 | 0.125 |
| 4 | 0.097 |
| 5 | 0.079 |
| 6 | 0.067 |
| 7 | 0.058 |
| 8 | 0.051 |
| 9 | 0.046 |

The sum of probabilities can be shown to be equal to 1:

$$\sum_{i=1}^{i=9} F_i = \sum_{i=1}^{i=9} \log_{10}(1 + 1/i) = \sum_{i=1}^{i=9} \log_{10}[(i+1)/i] = \log_{10} \prod_{i=1}^{i=9} [(i+1)/i]$$

$$= \log_{10}\left(\frac{2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8 \times 9 \times 10}{1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8 \times 9}\right)$$

$$= \log_{10}(10) = 1$$

In practice, it is difficult to apply this law universally; some numerical series do not conform to it. Hill (1995) demonstrated, however, that for these series, random samples taken from the complete series did follow Benford's law. Next, we will see how the law works for the current populations of different countries worldwide.

## II. The distribution of world populations

INED regularly publishes a list of country-specific statistics worldwide. Here, we will use the 1997 list, which gives information on 198 countries or geopolitical entities. Figure 1 compares the observed distribution with the distribution predicted by Benford's law. The difference between the distributions is small, confirming that the distribution of population among countries does in fact follow Benford's law.

If we perform the same type of calculations for surface areas and population densities, the results are similar (Figures 2 and 3).

## III. An explanation for population sizes

First, it must be made clear that there is no bias resulting from rounding the population size of small countries to 1 as a first significant digit, the smallest number in the table being 0.02 (million).

The explanation lies elsewhere. To work it out, one needs to look at the distribution of population sizes for a population with constant growth rate, for example 2% annually. Table 2 shows that at a constant annual rate of growth, population sizes with first significant digit 1 are more common than others, and follow Benford's law. If we extend these calculations over the long term, the observed frequency of first significant digits follows Benford's law perfectly; this is also true with other constant growth rates.

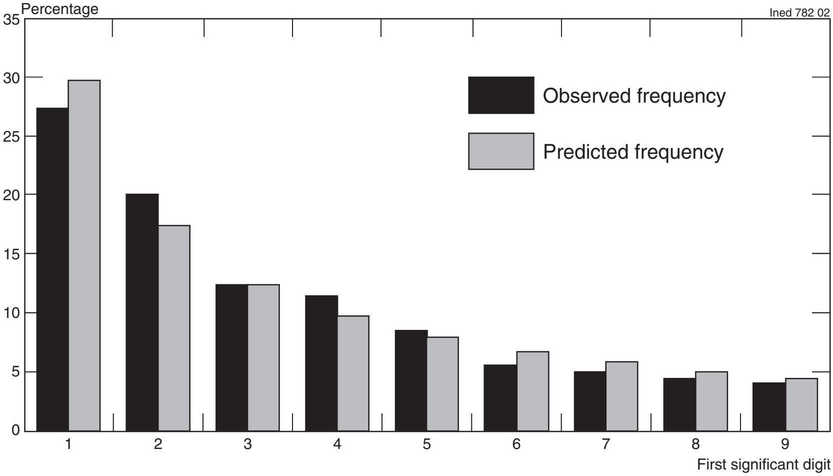


Figure 1.–Observed frequency of the first significant digit of the 1997 population size for 198 countries and predicted frequency according to Benford's law

*Source*: INED 1997.

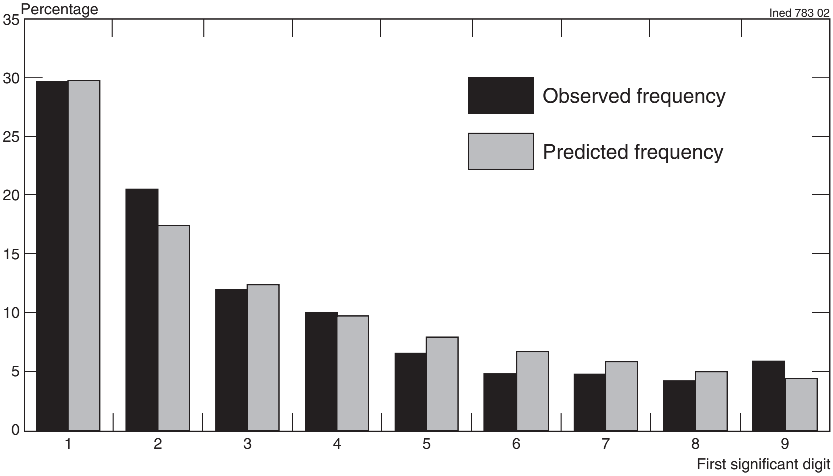


Figure 2.–Observed frequency of the first significant digit of the 1997 surface area for 198 countries and predicted frequency according to Benford's law

*Source*: INED 1997.

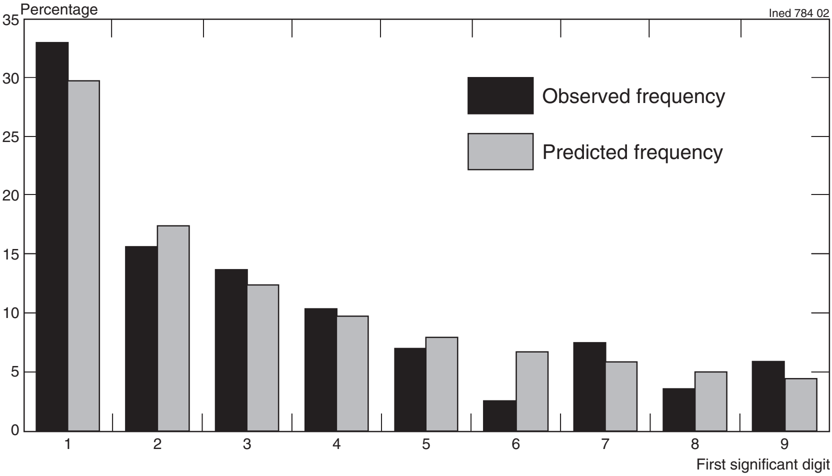


Figure 3.–Observed frequency of the first significant digit of the 1997 population density for 198 countries and predicted frequency according to Benford's law

*Source*: INED 1997.

Demographers use the shortcut of dividing 70 by the growth rate (in %) of a population to find the population's doubling time. This result can be derived as follows:

— the pace of the growth is given by $\frac{P(T)}{P(0)} = e^{rT}$ ;

— where $P(u)$ = population size at time $u$ ;

— $r$ = growth rate.

If the population size doubles between 0 and time $T_2$, then $\frac{P(T_2)}{P(0)} = e^{rT_2} = 2$ ;

— then $rT_2 = \ln\left[\frac{P(T_2)}{P(0)}\right] = \ln 2$ ;

— and $T_2 = \frac{\ln 2}{r} \approx \frac{0.693}{r} \approx \frac{0.70}{r} \approx \frac{70}{r'}$ where $r'$ is the constant growth rate in %.

To grow from size 100 to size 200, it takes about 35 units of time at a 2% growth rate.

To move from size 200 to size 300, it only takes 21 units of time:

$$T = \frac{1}{0.02}\ln\left[\frac{300}{200}\right] = \frac{\ln 1.5}{0.02}$$

It takes 15 units of time to grow from 300 to 400:

$$T = \frac{1}{0.02}\ln\left[\frac{400}{300}\right] \approx \frac{\ln 1.333}{0.02}$$

and only 5 units from 800 to 900:

$$T = \frac{1}{0.02}\ln\left[\frac{900}{800}\right] = \frac{\ln 1.125}{0.02}$$

The number of time periods necessary for a population to grow from size ($d$)00 to ($d$+1)00 is given by the formula:

$$T = \left(\frac{d+1}{d}\right)/r = \frac{1}{M}\log_{10}\left(\frac{d+1}{d}\right)/r = \frac{1}{M}F_d/r$$

— with $\ln(x) = \frac{1}{M}\log_{10}(x)$

— where $M$ is a constant

— and $F_d$ the frequency of appearance of the first significant digit $d$.

Thus, this formula is determined by the distribution of first significant digits, which is precisely the distribution specified by Benford's law. Over a long period, the first significant digit of the population size of any given country is therefore more often 1 than 2, 2 than 3, and so on, up to 9.

To look for a cross-sectional version of this longitudinal regularity for countries is equivalent to drawing a random sample among a set of 198 series that conform to Bedford's law, assuming that the following two hypotheses hold, which can be reasonably assumed.

*Hypothesis 1: the population sizes of different countries are independent;*

*Hypothesis 2: the timing of the onset of the demographic transition in any given country is independent of the first significant digit of its population size.*

## Conclusion

The explanation given here for the conformity of population sizes to Bedford's law does not negate the existence of other data sets, for example the areas of countries, that follow Benford's law but do not fit this longitudinal argument. Certain advances in probability theory have led to a better understanding of the principal aspects of Benford's law, but a completely satisfactory explanation still remains to be developed (Hill, 1999).

TABLE 2.– POPULATION SIZE OVER 172 TIME PERIODS WITH A 2% GROWTH RATE*

| Periods | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 1- 36 | 37-56 | 57-71 | 72-82 | 83-91 | 92-99 | 100-106 | 107-111 | 112-117 | 118-152 | 153-172 |
| 100.00 | 203.99 | 303.12 | 407.95 | 507.24 | 606.20 | 710.26 | 815.86 | 900.78 | 1014.43 | 2028.74 |
| 102.00 | 208.07 | 309.18 | 416.11 | 517.39 | 618.32 | 724.46 | 832.18 | 918.80 | 1034.71 | 2069.31 |
| 104.04 | 212.23 | 315.36 | 424.44 | 527.73 | 630.69 | 738.95 | 848.83 | 937.17 | 1055.41 | 2110.70 |
| 106.12 | 216.47 | 321.67 | 432.93 | 538.29 | 643.30 | 753.73 | 865.80 | 955.92 | 1076.52 | 2152.91 |
| 108.24 | 220.80 | 328.10 | 441.58 | 549.05 | 656.17 | 768.81 | 883.12 | 975.03 | 1098.05 | 2195.97 |
| 110.41 | 225.22 | 334.67 | 450.42 | 560.03 | 669.29 | 784.18 | | 994.53 | 1120.01 | 2239.89 |
| 112.62 | 229.72 | 341.36 | 459.42 | 571.24 | 682.68 | 799.87 | | | 1142.41 | 2284.69 |
| 114.87 | 234.32 | 348.19 | 468.61 | 582.66 | 696.33 | | | | 1165.26 | 2330.38 |
| 117.17 | 239.01 | 355.15 | 477.98 | 594.31 | | | | | 1188.56 | 2376.99 |
| 119.51 | 243.79 | 362.25 | 487.54 | | | | | | 1212.33 | 2424.53 |
| 121.90 | 248.66 | 369.50 | 497.29 | | | | | | 1236.58 | 2473.02 |
| 124.34 | 253.63 | 376.89 | | | | | | | 1261.31 | 2522.48 |
| 126.82 | 258.71 | 384.43 | | | | | | | 1286.54 | 2572.93 |
| 129.36 | 263.88 | 392.11 | | | | | | | 1312.27 | 2624.39 |
| 131.95 | 269.16 | 399.96 | | | | | | | 1338.51 | 2676.88 |
| 134.59 | 274.54 | | | | | | | | 1365.28 | 2730.42 |
| 137.28 | 280.03 | | | | | | | | 1392.59 | 2785.02 |
| 140.02 | 285.63 | | | | | | | | 1420.44 | 2840.72 |
| 142.82 | 291.35 | | | | | | | | 1448.85 | 2897.54 |
| 145.68 | 297.17 | | | | | | | | 1477.83 | 2955.49 |
| 148.59 | | | | | | | | | 1507.38 | |
| 151.57 | | | | | | | | | 1537.53 | |
| 154.60 | | | | | | | | | 1568.28 | |
| 157.69 | | | | | | | | | 1599.65 | |
| 160.84 | | | | | | | | | 1631.64 | |
| 164.06 | | | | | | | | | 1664.27 | |
| 167.34 | | | | | | | | | 1697.56 | |
| 170.69 | | | | | | | | | 1731.51 | |
| 174.10 | | | | | | | | | 1766.14 | |
| 177.58 | | | | | | | | | 1801.46 | |
| 181.14 | | | | | | | | | 1837.49 | |
| 184.76 | | | | | | | | | 1874.24 | |
| 188.45 | | | | | | | | | 1911.73 | |
| 192.22 | | | | | | | | | 1949.96 | |
| 196.07 | | | | | | | | | 1988.96 | |
| 199.99 | | | | | | | | | | |

*The table should be read sequentially by column.

**Acknowledgements**: The author wishes to thank Éva Lelièvre and Etienne van de Walle, as well as the anonymous referees, for their valuable comments.

## REFERENCES

BENFORD F., 1938, "The law of anomalous numbers", *Proceedings of the American Philosophical Society*, 78(4), pp. 551-572.

INED, 1997, "Tous les pays du monde", *Population et Sociétés*, 326, July-Aug.

HILL T., 1995, "A statistical derivation of the significant-digit law", *Statistical Science*, 10(4), pp. 354-363.

HILL T., 1999, "Le premier chiffre significatif fait sa loi", *La Recherche*, 316, January, pp. 72-75.

NEWCOMB S., 1881, "Note on the frequency of use of the different digits in natural numbers", *Am. J. Mathematics*, 4, pp. 39-40.