Teachers Corner

A Simple Explanation of Benford's Law

R. M. FEWSTER

Benford's Law, also known as the first-digit law, has long been seen as a tantalizing and mysterious law of nature. Attempts to explain it range from the supernatural to the measure-theoretical, and applications range from fraud detection to computer disk space allocation. Publications on the topic have escalated in recent years, largely covering investigation of the law in different data sources, applications in fraud and computer science, and new probability theorems. The underlying reason why Benford's Law occurs is, however, elusive. Many researchers have verified for themselves that the law is widely obeyed, but have also noted that the popular explanations are not completely satisfying. In this article we do nothing rigorous, but provide a simple, intuitive explanation of why and when the law applies. It is intended that the explanation should be accessible to school students and anyone with a basic knowledge of probability density curves and logarithms.

KEY WORDS: First-digit phenomenon; Law of anomalous numbers

1. THE IMPORTANCE OF BEING BENFORD

How good a statistical sleuth are you? Table 1 shows a list of land areas of world states and territories. That is, one of the columns does. The other column consists of faked data created with a random number generator (followed by a little massaging to make the numbers look better). Can you tell which column is which?

If you have heard of Benford's Law, you will know at a glance which column is correct. If you haven't, you probably won't have a clue—in which case you would be well-advised not to try faking any random data in the near future, especially on your tax return.

Benford's Law has intrigued scientists and laypeople for over a century. The story began in 1881, when the American astronomer Simon Newcomb noticed that books of logarithm tables always seemed grubby on the early pages, and clean toward the back. For some reason, people seemed to look up numbers beginning with the digits 1 and 2 far more often than they looked up numbers beginning with the digits 8 and 9. Indeed, it seemed that numbers beginning with 1 and 2 actually *occurred* more often in nature than numbers beginning with 8 and 9.

Newcomb went so far as to sketch out mathematically the law he expected first digits to follow (Newcomb 1881). A randomly selected number should begin with the digit 1 about 30% of the time: more precisely, the proportion should be 0.301, or $\log_{10}(2)$. The frequency of numbers with leading digit 2 should be about 18% (obtained from $\log_{10}(3/2)$), those with leading digit 3 should be about 12% (from $\log_{10}(4/3)$), and so on until the frequency of 8's should be 5.1% and that of 9's should be 4.6%. Overall, for a number chosen at random from those sought in the logarithm tables of 1881, Newcomb suggested that the probability of the first or leading digit being d should be

$$P(\text{leading digit} = d) = \log_{10}\left(\frac{d+1}{d}\right), \quad d = 1, 2, \ldots, 9.$$

With an apparent sleight-of-hand reminiscent of Fermat, Newcomb simply stated that his logarithmic rule was 'evident'. It is counter-intuitive enough to think that the digits should be anything but uniform, but this obscure expression for their frequencies has been the last straw for many a baffled onlooker. Exactly how evident could it have been that the logarithm books were precisely 6.58 times grubbier on page 1 than on page 9? Perhaps by way of stunned silence—or perhaps because Newcomb's observation seemed so obvious in 1881—nothing further was said on the topic for 57 years.

In 1938, Frank Benford, a physicist with the General Electric Company, assembled over 20,000 numbers from sources as diverse as *Readers' Digest* articles, street addresses of *American Men of Science*, atomic weights, population sizes, drainage rates of rivers, and physical constants (Benford 1938). His data showed that leading digits from a wide range of sources showed an uncanny adherence to the logarithmic rule that Newcomb had penned, apparently unnoticed, decades earlier. Benford gave the law its name and a certain mystique, but no convincing explanation. Indeed, he decried the phenomenon as belonging to 'outlaw' and 'anomalous' numbers, and reassured us that the dapper digits of orderly data such as atomic weights and specific heats would be free of such absurdities. Benford's Law was the province of wild data—population sizes, addresses of American men of science, and figures in *Readers' Digest* articles. Over a third began with the digit 1, whereas fewer than one in 20 began with the digit 9.

R. M. Fewster is with the Department of Statistics, University of Auckland, Auckland, New Zealand. (E-mail: *r.fewster@auckland.ac.nz*). The author is most grateful to the referees and editors for their constructive comments on this manuscript. Special thanks to Eric Sampson and Janet Wallace for editorial assistance, and to LottoStrategies.com for permission to use the Powerball jackpot data. Population sizes of Californian congressional districts and cities are available from the US Census Bureau, *www.census.gov*. A list of world populations is available from Wikipedia at *en.wikipedia.org/wiki/List_of_countries_by_population*.

  DOI: 10.1198/tast.2009.0005

Table 1. One of the columns gives the land area of political states and territories in km^2. The other column contains faked data, generated with a random number generator.

State/Territory	Real or Faked Area (km^2)	
Afghanistan	645,807	796,467
Albania	28,748	9,943
Algeria	2,381,741	3,168,262
American Samoa	197	301
Andorra	464	577
Anguilla	96	82
Antigua and Barbuda	442	949
Argentina	2,777,409	4,021,545
Armenia	29,743	54,159
Aruba	193	367
Australia	7,682,557	6,563,132
Austria	83,858	64,154
Azerbaijan	86,530	71,661
Bahamas	13,962	9,125
Bahrain	694	755
Bangladesh	142,615	347,722
Barbados	431	818
Belgium	30,518	47,123
Belize	22,965	20,648
Benin	112,620	97,768
...	...	...

By this time, you don't need me to tell you that it is the first column in Table 1 that is correct. According to Benford's law, leading digits as far-fetched as those in the second column would be a one-in-a-thousand rarity. In this way, the law is used to detect fraudulent data in applications as diverse as election campaign finance and toxic gas emissions (e.g., Cho and Gaines 2007). Yet where is the explanation? Why should 101 Dalmatians be so much more likely than 57 Heinz Varieties? How did a grubby logarithm book evoke a logarithmic probability lore governing its own pages? (For the record, the number of Heinz Varieties currently stands at the highly anomalous 1,100: Benford wins again!)

Many writers have come to the conclusion that Benford's law is a mysterious law of nature, for which a true explanation lies with the gods. It is commonly accepted that the first rigorous explanation of the law was due to Hill (1995), who provided a measure-theoretical proof that data drawn from a random mix of different distributions—rather like Frank Benford's original 20,000—will ultimately converge to Benford's law. Although rigorous, this gives us little insight into why Benford's law applies, or when. Should we expect, for example, that the street addresses of American men of science would qualify?

Attempts at intuitive explanations have centered on ideas of scale-invariance and base-invariance. The scale-invariance argument says that, if there is a universal law of nature that governs the distribution of leading digits, then it should not depend upon the units in which the numbers are measured. If we were to convert the areas in Table 1 from km^2 to square miles—a simple scaling—the same distribution of leading digits should result. It can be shown that Benford's law follows automatically: if there is a universal law of nature, it must be Benford's. Similarly, any universal law should apply whether it is being observed by humans, with 10 fingers apiece, or by ducks with six toes. There should be nothing special about the base 10 number system. Again, it can be shown that a universal pattern in the leading digits that applies in all number bases is forced to be an obvious generalization of Benford's law.

These explanations are not wholly satisfactory, because they do not explain why a universal law of nature should arise in the first place. And while modern applications of Benford's Law flourish—for example in deciding the allocation of computer disk space and detecting fraud—popular insight into why and when numbers should be Benford is lacking. There is, however, a simple and intuitive explanation accessible to anyone who has a basic knowledge of probability density curves and logarithms. The simple explanation has been seen by previous authors, but seems only to have been published using technical language and tools—for example, as an application of Poincaré's Theorem in circular statistics (Mardia and Jupp 2000), or in terms of digital signal processing with Fourier transforms (Smith 2007). The aim of this note is to give a simple, graphical explanation of why and when the law holds, in language accessible to school students and laypeople.

2. LAW OF THE STRIPEY HAT

Benford's law hinges on the simple observation that, if a hat is covered evenly in black and white stripes, then about half of the hat will be black. More generally, think of a two-dimensional hat-shaped piece of cardboard (Fig. 1). If the black stripes cover proportion p of the 'rim', they will cover approximately proportion p of the whole hat. The approximation will tend to improve as the hat-shape contains more stripes: a large number of thinner stripes is more likely to average out any asymmetries in the hat-shape than a smaller number of fatter stripes. On average, as we slide the stripes randomly to left or right, the striped area will cover proportion p of the total area of the hat-shape.

A short step relates this to Benford's law. We will show that, for any positive number X, the leading digit of X is 1 precisely when $\log_{10}(X)$ is between n and $n + 0.301$ for some integer n. Now think of X as a random number drawn from some probability distribution. The hat in Figure 1(b) represents the probability density curve of $\log_{10}(X)$. The numbers X with leading digit 1 correspond to the stripes on the hat, with each stripe covering the interval from n to $n + 0.301$ for some integer n. There is one stripe for every integer n included within the distribution of $\log_{10}(X)$, so the total number of stripes is the total number of integers spanned by $\log_{10}(X)$. The stripes cover about 0.301 of the 'rim' of the hat, so they will capture about 0.301 of the total probability of X. The chance of X having leading digit 1 will be almost a third.

To show that the stripes correspond to numbers X with leading digit 1, first write X in 'scientific notation'. This is the unique format $X = r \times 10^n$, where r is a real number with $1 \leq r < 10$, and n is an integer. For example, if $X = 124$, then $X = 1.24 \times 10^2$, or if $X = 76$ then $X = 7.6 \times 10^1$. The leading digit of X is the same as the leading digit of r, but whereas we need to consider many intervals for X (e.g., X = 1–1.999; 10–19.999; 100–199.999), we only need to consider one interval for r: the leading digit of X is 1 precisely when $1 \leq r < 2$.

We can isolate r by taking logs to base 10:

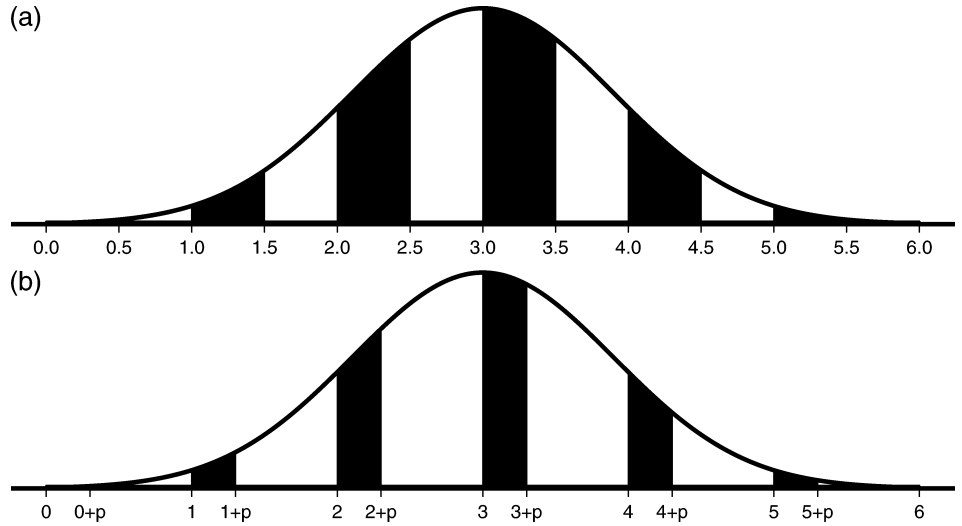

Figure 1. (a) A hat with alternating black and white stripes of equal width will be approximately half black and half white. (b) If the black stripes cover proportion p of the rim of the hat, then approximately proportion p of the hat will be colored black. The p shown in (b) is $p = 0.301$. An arbitrary number-line is featured along the rim of the "hat." The black stripes correspond to the numbers X with leading digit 1, if the hat-shape represents the probability density curve of $\log_{10}(X)$.

$$\log_{10}(X) = \log_{10}(r \times 10^n) = \log_{10}(r) + n.$$

The leading digit of X is 1 precisely when $1 \le r < 2$, which is the same as saying that $0 \le \log_{10}(r) < \log_{10}(2) = 0.301$. Thus X will have leading digit 1 precisely when $\log_{10}(X)$ lies between $n + 0$ and $n + 0.301$ for some integer n: in a set of evenly spaced "stripes" of equal width 0.301. For example, when $X = 124$, we find $\log_{10}(X) = 2.093$. Sure enough, 2.093 lies between 2 and 2.301, and the leading digit of 124 is 1. On the other hand, when $X = 76$, then $\log_{10}(X) = 1.881$, which lies outside the interval from 1–1.301; happily, the leading digit of $X = 76$ is not 1.

We have shown that the numbers X with leading digit 1 satisfy:

$$0 \le \log_{10}(X) < 0.301; \quad 1 \le \log_{10}(X) < 1.301;$$
$$2 \le \log_{10}(X) < 2.301; \quad \ldots$$

The X-values with leading digit 1 lie in regularly-spaced stripes on the $\log_{10}$ scale, where each stripe has width 0.301. If the cardboard hat in Figure 1 is the probability density function of $\log_{10}(X)$, then the probability that X has leading digit 1 is precisely the probability that $\log_{10}(X)$ falls within the black stripes on Figure 1(b). The stripes cover proportion $p = 0.301$ of the rim of the hat, so they cover approximately proportion $p = 0.301$ of the area of the hat. For probability density curves, areas correspond to probabilities. So the probability that X has leading digit 1 will be somewhere close to $p = 0.301$—just as Benford discovered.

The preceding argument explains not only why the proportion of numbers beginning with 1 should be close to a third, but when it should be. The more stripes we have, the closer to 0.301 we should expect our proportion to be, because local asymmetries in the hat shape sampled by one stripe will be better balanced by results from other stripes. Therefore, the more stripes there are, the more 'Benford' our data should look. We can't increase the number of stripes by making them closer together, because they have to be located at the intervals [0, 0.301), [1,1.301), [2, 2.301), and so on. The only way to get more stripes is to make the hat wider. This means the distribution of $\log_{10}(X)$ should cover a larger range: in other words, the distribution of X should span several orders of magnitude. If X can take values from $1–10^6$, then $\log_{10}(X)$ will span six integers, giving six stripes as in Figure 1. This will usually be enough to make the distribution look convincingly 'Benford'.

We can conclude that data from *any* distribution will tend to be 'Benford', as long as the distribution spans several integers on the $\log_{10}$ scale—several orders of magnitude on the original scale—and as long as the distribution is reasonably smooth. Clearly, we could cheat the Benford property by deliberately punching a dent in the hat at every black stripe. We look more at this in Section 4. However, the dent-punching requires deliberate manipulation: any distribution arising in nature that is reasonably smooth and covers several orders of magnitude is almost guaranteed to obey Benford's law. Example distributions such as areas of world states (ranging from 0.4 km^2 for Vatican City to 1.7×10^7 km^2 for Russia), or world populations (from 50 for the Pitcairn Islands to 1.3×10^9 for China), explain why Benford's law provides such a reliable party trick whenever an atlas is at hand.

A more complete explanation of Benford's law runs like this. Write $X = r \times 10^n$ for $1 \le r < 10$, $r \in \mathbb{R}$, and $n \in \mathbb{Z}$. For $d \in \{1,2,\ldots,9\}$, the leading digit of X is d if and only if $d \le r < d + 1$, or equivalently, $\log_{10}(d) \le \log_{10}(r) < \log_{10}(d + 1)$. This corresponds to sampling strips of width $\log_{10}(d + 1) - \log_{10}(d) = \log_{10}\{(d + 1)/d\}$ at integer spacing across the distribution of $\log_{10}(X)$. If the distribution of $\log_{10}(X)$ is reasonably smooth and spans several integers, then the area covered by these strips will be approximately the same as the proportion of the interval they cover, namely $\log_{10}\{(d+ 1)/d\}$. We therefore expect that the probability of obtaining leading digit 1 is about $\log_{10}(2/1) = 0.301$; the probability of leading digit 2 is about $\log_{10}(3/2) = 0.176$; and so on. Benford's law amounts to the observation that a

sample of areas taken systematically along the rim of a hat-shape will give a good estimate of the total area of the hat.

3. BENFORD OR NOT?

Figure 2 shows a selection of real data distributions that are candidates for Benfordness. The first is a Benford classic: populations of world states and territories. The histogram of populations (X) has extreme skew, covering nine orders of magnitude from 10^1 to 10^9. By contrast, the histogram of log-populations ($\log_{10}(X)$) is much more symmetric. It is this second histogram that contains the clues for the Benford detective—it is the hat shape from Section 2. The second histogram is reasonably smooth and spans nine integers from 1–9, clues that will clinch the detective's case. The barplot in the third column shows the evidence. The bars give the proportions of each leading digit in the sample of 235 world states, whereas the horizontal lines give the Benford predictions. The world states pass the Benford test with flying colors. The minor deviations of the bars from the predictions do not approach statistical significance.

The second example is more of a challenge for our detective. It shows the size of United States Powerball jackpots, in millions of dollars, twice-weekly from June 2002 to September 2008. The data are available from *www.lottostrategies.com*. At first glance the distribution seems a reasonable candidate: it ranges widely from \$10 million to \$365 million. The difference is certainly a lot of money—enough to launch a small fleet of satellites around the moon—but in the land of Benford it is merely small change. Only the orders of magnitude count, and the range from $\$1 \times 10^7$ to $\$3.65 \times 10^8$ gives us a diminutive hat spanning only 1.5 integers on the log scale of the second histogram. This means only one or two stripes per digit, and no guarantee of Benfordness. On the other hand, the shape of the

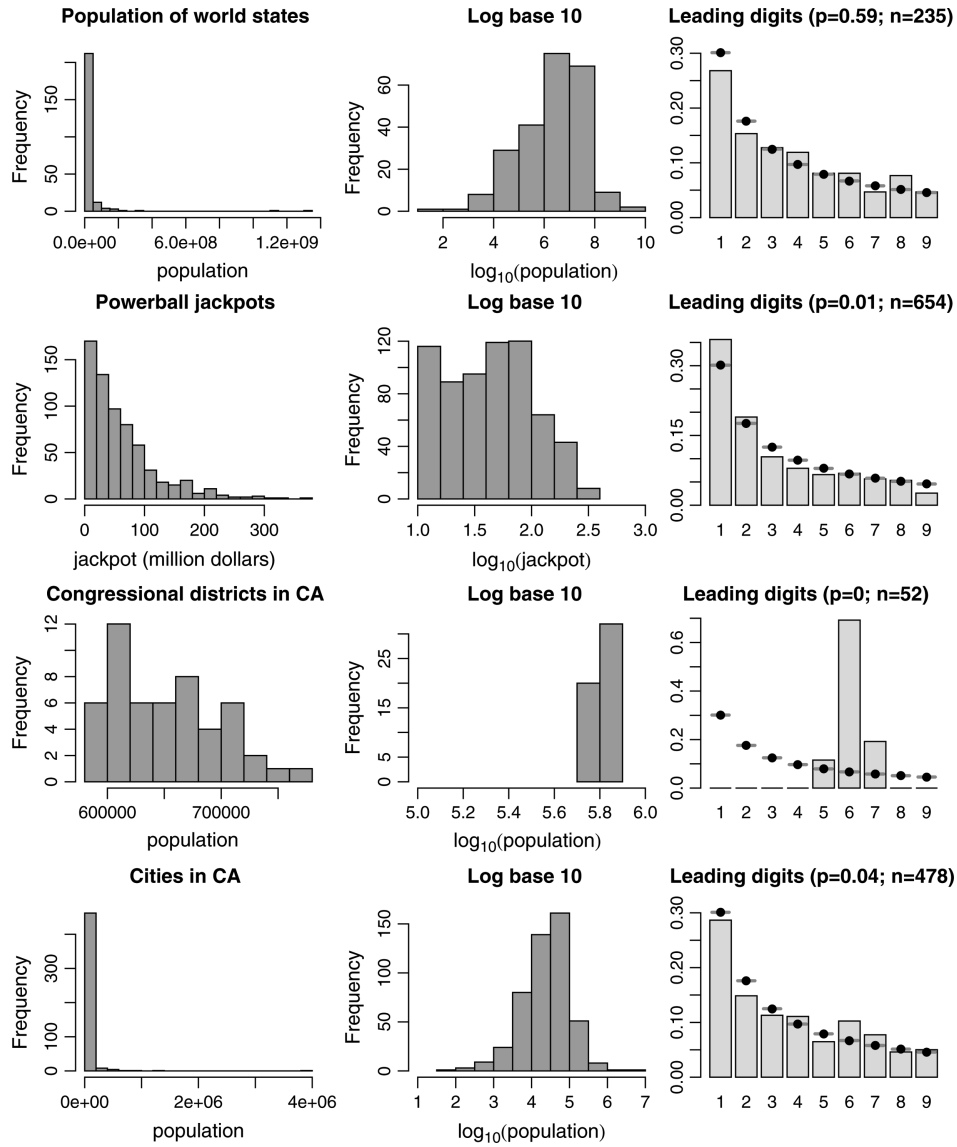

Figure 2. Benford detection for four real datasets. Each row features one dataset. Shown are the histogram of the original data, the histogram of the log-transformed data, and a barplot of the leading digit proportions. Heights of the bars give the proportions of the leading digits in the dataset, whereas the Benford proportions are superimposed with horizontal lines marked with plotted points. Also marked on the barplots are the sample size, n, and the p-value from a chi-squared test for Benfordness with 8 degrees of freedom.

hat is rather rectangular—more like a top hat than a Napoleonic hat—so one or two stripes might actually be enough to capture the Benford proportions. The only recourse is to look at the facts of the case in the third column. The fit to Benford is surprisingly good given the clues, but a well-equipped statistician would be able to spot the difference. In fact, Powerball jackpots form a sort of 'super-Benford' distribution: there are too many 1's (36% instead of 30%) and not enough 9's (2.6% instead of 4.6%).

The third distribution shows population sizes of congressional districts in California. Deliberately chosen to be as uniform as possible, to be fair to the voters, the populations range from 583,000–773,000. Being a confined and deliberately manipulated dataset, its prospects for Benfordness are hopeless. The log-scale hat covers only a fraction of an integer, and the leading digits are all 5, 6, or 7. No dice for Benford here.

What about naturally-occurring populations in California? The problem with the congressional districts is that they are deliberately chosen to be within certain population bounds. If, on the other hand, we look at the populations of all Californian cities, we have a natural, unfettered distribution—one that is a real candidate for Benfordness. The population sizes of Californian cities in 2003 are shown in the fourth distribution. They range from a population of 94 for the tiny city of Vernon, to 3.9 million for the vast metropolis of Los Angeles. Once again, the original data are highly skewed, but the log-transformed data are smooth and symmetric. The Benford detective is quietly confident: the log-histogram hat spans over 4.5 integers, and the distribution is smooth. The third plot shows that the leading digits are distinctively Benford. Interestingly, the departures from Benfordness in this case are weakly statistically significant—for some unexplained reason, there are too many Californian cities with populations beginning with a 6!

4. MAXIMIZING BADNESS-OF-BENFORD

We mentioned that Benford's law can be cheated by punching judiciously-chosen dents in the Benford hat. Intuitively, it seems clear that Benford's law will *have* to hold as the distribution of X becomes wide and smooth enough. Here we investigate how far from the Benford fold a distribution can stray, and how this changes as the distribution becomes wider and smoother. For this, we need ways of controlling the width and smoothness of a distribution, and we also need some measure of 'badness-of-Benford' so we can find the worst-case scenario for a distribution of given width and smoothness.

For distribution width, we focus on the number of integers spanned by the support of $\log_{10}(X)$. If the support of $\log_{10}(X)$ spans a single integer—e.g., if X takes values between 1 and 10, or between 10 and 100—then there is only one stripe on the Benford hat, and it will be quite easy to cheat Benford's law. When the support of $\log_{10}(X)$ ranges over s integers—e.g., X can take values between 1 and 10^s—there will be s stripes on the Benford hat. Cheating Benford's law when s is large will require a more deliberate sabotage operation, and one that is less likely to arise in nature.

For our badness-of-Benford measure, we recommend launching a sabotage job on the statistician's favorite goodness of fit tool: the Pearson chi-squared test statistic. For $d = 1, 2, \ldots, 9$, define the Benford digit probabilities as $b_d = \log_{10}\{(d+1)/d\}$, so b_d is the probability according to Benford's law that the leading digit is d. For a random variable X, let $\pi_x(x)$ be the probability density function (p.d.f.) of $\log_{10}(X)$, defined on the interval $[0, s]$ for some integer s. The probability that X has leading digit d is:

$$p_d = P(X \text{ has leading digit } d) = \sum_{n=0}^{s-1} \int_{n+\log_{10}(d)}^{n+\log_{10}(d+1)} \pi_X(x)\, dx. \quad (1)$$

By analogy with the Pearson chi-squared statistic, we define the badness-of-Benford for the distribution to be

$$B_X = \sum_{d=1}^{9} \frac{(p_d - b_d)^2}{b_d}.$$

It is easily shown that B_X dominates the expected Pearson chi-squared statistic that would be obtained if a sample of data were drawn from X. If we choose π_x to maximize B_X, we will have the best chance of gaining a significant chi-squared test result against the null hypothesis that the data follow Benford's law. If the chi-squared test is our yardstick of Benfordness, then B_X is our best sabotage tool.

Finally, we need some way of controlling the smoothness of X. If we allow sufficient flexibility in the p.d.f. π_X, we will always be able to enact a perfect sabotage: simply bring out the statistical scimitar and hack holes in π_X where the leading digit is 1, while drawing out peaks where the leading digit is 9. Natural distributions don't look like this, however. Instead it is fair to require $\pi_X(x)$ to be reasonably smooth over its support interval $[0, s]$. We can do this by maximizing a *penalized* badness-of-Benford criterion—by subtracting a term from B_X that drags it down if π_X is too rough or wiggly.

A useful penalty term for downweighting wiggly functions is $\lambda \int_0^s \pi_X''(x)^2 dx$, where λ is called the smoothing parameter (Hastie and Tibshirani 1990), and is chosen at our discretion. By integrating the square of the second derivative of π_X over its support interval, we penalize excessive curvature or wiggliness. Instead of simply choosing π_X to maximize B_X, we choose it to maximize

$$B_X - \lambda \int_0^s \pi_X''(x)^2 dx. \quad (2)$$

Choosing a large value of λ will force the p.d.f. π_X to be very smooth, whereas a small value of λ will allow π_X to be more flexible.

Our Benford sabotage job is formulated as finding π_X to maximize the expression (2) out of all possible p.d.f.s π that are continuous and have continuous first and second derivatives on the interval $[0, s]$. Remarkably, such a sabotage job has a unique solution that can be written down in closed form. The maximizing π_X is a cubic spline, which means it is a piecewise cubic. If we approximate the integrals in (1) using the trapezium rule over a fine grid, the knots or joins in the piecewise cubic are placed at the same grid-points. Technical details of how to calculate the cubic spline are given in Fewster and Patenaude (2008). For technical reasons, a few constraints are

necessary at the endpoints. We choose to constrain $\pi_X(x)$ such that $\pi_X(0) = \pi_X(s) = 0$, and ordain that π_X should be linear (perhaps flat) at $x = 0$ and $x = s$.

Figure 3 shows a selection of optimal sabotage jobs for $s = 6$, 4, and 2. Each row corresponds to one value of *s:* the number of orders of magnitude spanned by the distribution of *X*. For each *s,* the first panel shows how our Benford-cheating ability deteriorates as the distribution becomes smoother. For small λ (wiggly distributions, or hats with many dents), we can arrange for the Benford test to be failed as often as we like—in Figure 3, this corresponds to $P(\bar{B}) \simeq 1$, where $P(\bar{B})$ is the probability that a sample of size 100 drawn from *X* will fail the Benford chi-squared test. (In a convenient abuse of terminology, we say that a test is 'failed' if it returns a *p*-value of 0.05 or less.) Examples of such distributions are shown in the second column of Figure 3. As λ increases and our dent-punching is brought under stricter control, Benford's law rapidly triumphs. Even the energetic sabotage jobs shown in the third column of Figure 3 will show a significant difference from Benford's law for only about a quarter of samples. By the time our dent-punching has subsided to finger-tapping in the fourth column of Figure 3, only about 5% of samples will fail the chi-squared test: exactly the proportion that would be expected if Benford's law holds exactly. And remember these are our best possible sabotage jobs! It is very difficult for a distribution *not* to look reasonably Benford in a sample of size 100, once it spans about $s = 4$ orders of magnitude or more.

In summarizing the explanations for Benford's law current in 1976, Ralph Raimi wrote, "... any phone company can print a directory violating Benford's law. What remains tantalizing is the notion that there is still some unexplained measure in the universe which says that the probability of such violations is small" (Raimi 1976). He suggested that the logarithmic rule originally set down by Simon Newcomb in 1881 was an inspired guess—pulled out of a hat with a magician's flourish and the muttered incantation, "It is evident that ...". We have seen that the tantalizing universal law is not so inexplicable, and it really is evident to anyone who has met probability density curves and logarithms. It's as easy as painting stripes on the magician's hat.

It is not surprising that Benford's law is everywhere. Even Frank Benford was a victim of his own law. In his 1938 paper, he collected a total of 20,229 observations from 20 vastly different datasets, a study "as wide as time and energy permitted". His sample sizes ranged from 91 atomic weights, to

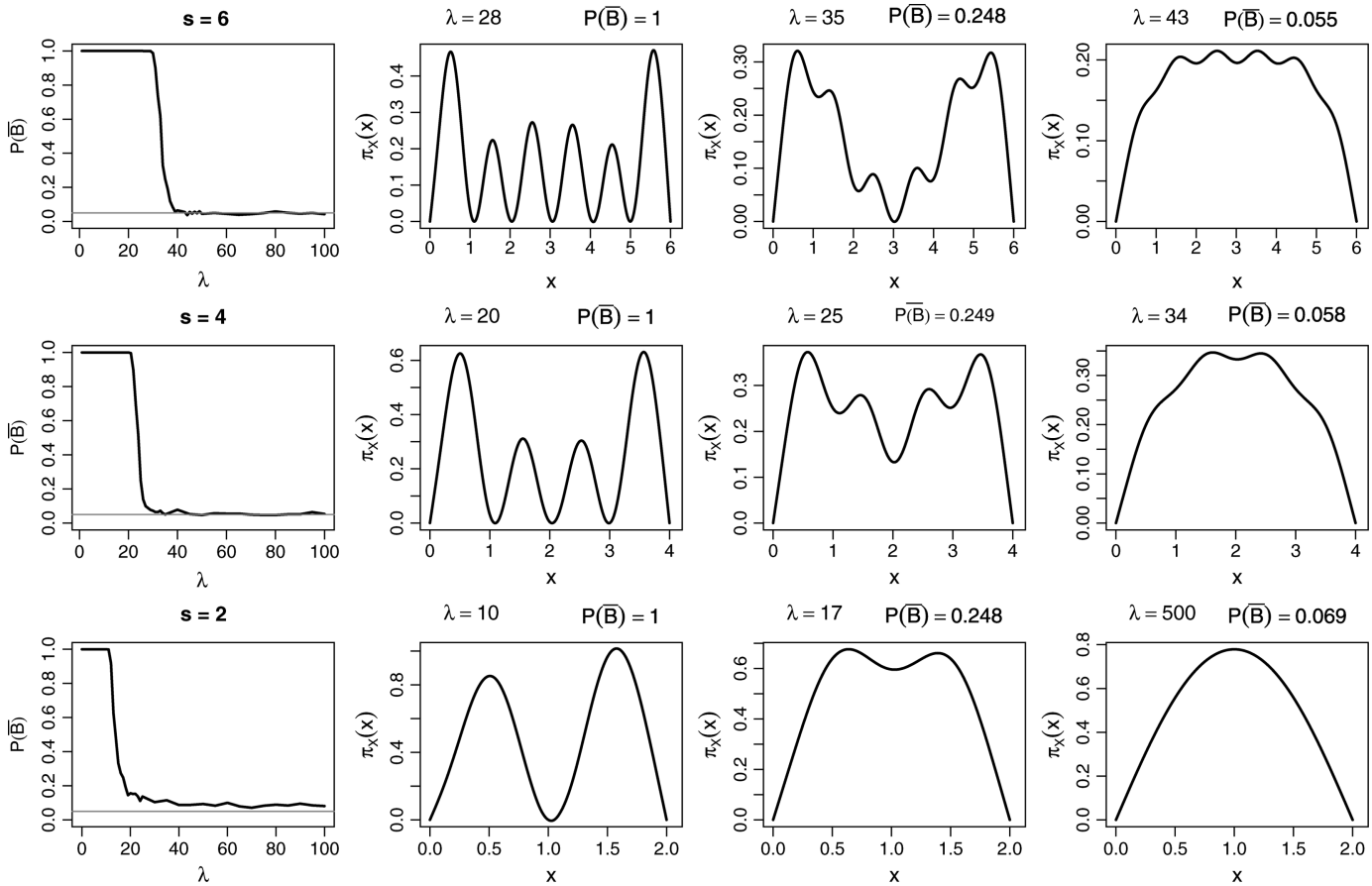

Figure 3. Worst-case Benford scenarios for a range of *s* and λ. The rows correspond to $s = 6$, 4, and 2 respectively. For each *s* and λ, the optimal badness-of-Benford density $\pi_x(x)$ is found by maximizing (2). We then draw 1,000 samples of size 100 from the distribution of *X* and conduct a chi-squared test with 8 degrees of freedom against the null hypothesis that the distribution of leading digits in the sample is Benford. The proportion of the 1,000 samples that 'fail' the chi-squared test is $P(\bar{B})$ in the first column, and it can be seen to plummet as the distributions reach a certain smoothness. The thin horizontal line shows $P(\bar{B}) = 0.05$, corresponding to a true Benford distribution. The other three columns show examples of the worst-case density $\pi_x(x)$ with increasing smoothness λ, for which $P(\bar{B})$ is respectively close to 1; close to 0.25; and close to 0.05, as it would be for a true Benford distribution. To create an integer scale for λ, the values shown are those from (2) multiplied by 8,000.

5,000 entries from a mathematical handbook. Funnily, of the 20 datasets that Benford collected, six of the sample sizes have leading digit 1. Notice anything strange about that?

[Received May 2008. Revised October 2008.]

REFERENCES

Benford, F. (1938), ''The Law of Anomalous Numbers,'' *Proceedings of the American Philosophical Society,* 78, 551–572.

Cho, W. K. T., and Gaines, B. J. (2007), ''Breaking the (Benford) Law: Statistical Fraud Detection in Campaign Finance,'' *The American Statistician,* 61, 218–223.

Fewster, R.M., and Patenaude, N. J. (2008), ''Cubic Splines for Estimating the Distribution of Residence Time Using Individual Resightings Data,'' in *Modeling Demographic Processes in Marked Populations,* eds. D. L. Thomson, E. G. Cooch, and M. J. Conroy. *Environmental and Ecological Statistics Series*, 3, 393–415.

Hastie, T. J., and Tibshirani, R. J. (1990), *Generalized Additive Models,* London: Chapman and Hall, p. 27.

Hill, T. P. (1995), ''A Statistical Derivation of the Significant-Digit Law,'' *Statistical Science,* 10, 354–363.

Mardia, K. V., and Jupp, P. E. (2000), *Directional Statistics,* Chichester: Wiley, Example 4.1.

Newcomb, S. (1881), ''Note on the Frequency of Use of the Different Digits in Natural Numbers,'' *American Journal of Mathematics,* 4, 39–40.

Raimi, R. A. (1976), ''The First Digit Problem,'' *The American Mathematical Monthly,* 83, 521–538.

Smith, S.W. (2007), ''Explaining Benford's Law,'' Chapter 34 in *The Scientist and Engineer's Guide to Digital Signal Processing*. Available at http://www.dspguide.com/.