The Long and the Short of Benford's Law

Arthur T. Benjamin

Quick! If you multiply a five-digit number by an eight-digit number, how many digits will the product have? The answer is 12 or 13 digits. (I'm sorry to report that despite efforts to incorporate more material on mental estimation into the precollege math curriculum, very few students or teachers can quickly answer this question. But that's not the point of the article.) In general, the product of an m-digit integer M and n-digit integer N has $m+n$ digits (the *long* product) or $m+n-1$ digits (the *short* product). Here's the one-sentence proof: If $10^{m-1} \le M < 10^m$ and $10^{n-1} \le N < 10^n$, then $10^{m+n-2} \le MN < 10^{m+n}$.

We can often determine which of these lengths is obtained by looking at the first digits of M and N. If the product of the first digits is 10 or more, we get the long product, if the product is 4 or less, we get the short product, and if the product is between 5 and 9, we must examine more digits.

Now, if M and N are chosen at random, are we more likely to get a short product or a long one? The answer depends on what we mean by "at random." As we'll see, two very natural ways to choose M and N lead to different conclusions. Specifically, we show that if M and N are randomly chosen from a uniform distribution, then the probability of a short product is pretty low, but if the numbers being multiplied are generated by Benford's law, then the probability of getting a short product is essentially 1/2.

Uniform Distribution

Let's first examine the case in which M and N are randomly chosen integers between 1 and 9. If all 81 possible choices of M and N are equally likely, then the probability that their product is a one-digit number is

$$\frac{9+4+3+2+1+1+1+1+1}{81} = \frac{23}{81} \approx 0.276.$$

Otherwise, it is a two-digit number. Similarly, if M is a random one-digit number and N is a random two-digit number, then the probability that the product is below 100 is

$$\frac{90+40+24+15+10+7+5+3+2}{810} = \frac{196}{810} \approx 0.242.$$

When M and N are random two-digit numbers, the probability that their product is a three-digit number is $1{,}490 / 8{,}100 \approx 0.184$.

In general, we can write $M = X \cdot 10^{m-1}$ and $N = Y \cdot 10^{n-1}$, where X and Y are real numbers in the interval $[1,10)$ with $m-1$ and $n-1$ digits, respectively, after the decimal point. The product is short if and only if $XY < 10$.

As the number of digits in m and n gets larger, we approximate X and Y with continuous random variables. Specifically, when X and Y are independent uniform random variables chosen from the interval $[1,10)$, then

$$P(XY < 10) = \int_1^{10} \int_1^{10/x} \frac{1}{81}\, dy\, dx = \frac{10 \ln 10 - 9}{81} \approx 0.173.$$

So, the probability of a short product is indeed quite low when all digits are equally likely.

Benford Distribution

In the real world (and even in the mathematical world), it is often the case that the numbers we encounter are much more likely start with small digits. For example, in a large finite collection of street addresses, electricity bills, lengths of rivers (measured in *any* units), distances to stars, powers of 2, and Fibonacci numbers, we find that well over half of the entries begin with digits 1, 2, or 3, and they vastly outnumber those entries that begin with 7, 8, and 9.

The formal statement of this observation is known as *Benford's law*. It states that the probability that the leading digit is d is $\log(d+1) - \log d$ (where the logarithm is base-10). Table 1 shows the leading digit probabilities. For more about Benford's law see Arno Berger and Theodore P. Hill's *An Introduction to Benford's Law* (Princeton University Press, 2015) or Steven J. Miller's edited volume *Benford's Law: Theory and Applications* (Princeton University Press, 2015).

So, now suppose that M and N are generated by Benford's law. As before, let $M = X \cdot 10^{m-1}$ and $N = Y \cdot 10^{n-1}$ with continuous random variables X and Y over the interval [1,10), but now X and Y are independent and identically distributed with probability density function

$$f(x) = \frac{1}{x \ln 10}.$$

This is consistent with Benford's law because the probability that X begins with digit d is

$$\int_d^{d+1} \frac{dx}{x \ln 10} = \frac{\ln(d+1) - \ln(d)}{\ln 10} = \log(d+1) - \log d.$$

Consequently, the probability of a short product is

$$\begin{aligned} P(XY < 10) &= \int_1^{10} \int_1^{10/x} \frac{1}{xy(\ln 10)^2} \, dy \, dx \\ &= \frac{1}{(\ln 10)^2} \int_1^{10} \frac{1}{x} (\ln(10) - \ln x) dx \\ &= \frac{1}{(\ln 10)^2} \left[\ln x \ln 10 - \frac{(\ln x)^2}{2} \right]_1^{10} \\ &= \frac{1}{2}, \end{aligned}$$

as desired.

This result has another quick derivation. It's easy to show that for X and Y described previously, $\log X$ and $\log Y$ are independent uniform random variables over the interval [0,1). Consequently, $P(XY < 10) = P(\log X + \log Y < 1) = 1/2$, by symmetry.

Table 1. Benford probabilities.

Leading digit	Probability (rounded)
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 result is even more surprising when we consider the fact that the product of two independent Benford random variables is still a Benford random variable (see Berger and Hill). For example, the probability that the product begins with 1 is still $\log 2 \approx 0.301$. Nevertheless, the probability that the product is a short number beginning with 1 is $(\log 2)^2 / 2 \approx 0.045$, and the probability that it is a long number beginning with 1 is about 0.256. In short, even though the product of Benford is Benford, if we compute the conditional probability assuming we have a short (or long) product, the resulting distribution is no longer Benford!

The situation for division is even simpler than multiplication. When we divide an m-digit number by an n-digit number in which $m \geq n$, the integer part of the quotient will have either $m - n$ or $m - n + 1$ digits. Specifically, if $M = X \cdot 10^{m-1}$ and $N = Y \cdot 10^{n-1}$, where X and Y are in the interval [1,10), then the integer part M / N has the shorter length if and only if $X < Y$. Thus, if X and Y are independent and identically distributed random variables chosen from *any* continuous distribution, we obtain a short quotient with probability 1/2.

Arthur T. Benjamin is the Smallwood Family Professor of Mathematics at Harvey Mudd College and president of the Fibonacci Association. He enjoys playing backgammon, racing calculators, writing parodies, and performing magic.

10.1080/10724117.2019.1568086