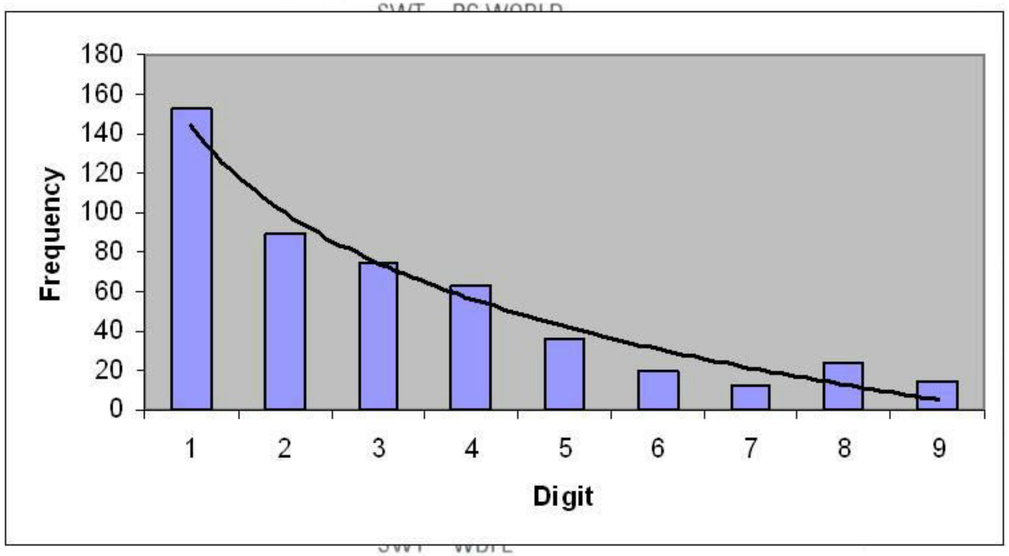
THEOREM OF THE DAY

Benford's Law *Let X be a union of random samples chosen without bias from a collection of different distributions on* $\mathbb{R}$. *Then, for* $d \in \{1, \ldots, 9\}$, *for x chosen uniformly at random from X,*

$$\mathbb{P}(\textit{1st significant digit of } x = d) = \log_{10}(1 + 1/d).$$

Your Bank Account details

Date	Payment type and details		Paid out	Paid in	Balance
21 Dec 04		BALANCE BROUGHT FORWARD			4,309.95
22 Dec 04	SWT	ASDA STORES 5871			
		OLD KENT ROAD	96.78		4,213.17
23 Dec 04	SWT	SAINSBURYS PLC			
		NEW CROSS 031	89.55		4,123.62
24 Dec 04	SWT	BOOKS ETC			
		LUTON	12.98		
	SWT	SOUTHERN NEW X GT			
		RAIL TICKET	21.40		
			99.93		
			70.00		3,919.31
			85.76		
			7.25		
			20.80		
			9.99		
			21.20		
			26.49		
		EDINBURGH	20.34		



The graph shows the number of times each digit from 1 to 9 appears as the first significant digit in the 'Paid out' column of a year's worth of bank statements. On the statement page shown, '£12.98' is the only time 1 occurs as first significant digit. Over many pages, however, Benford's law asserts itself. Note that the law does not hold for *all* collections of numbers. The 'Balance', for example, reduces regularly and cyclically from pay-day to pay-day and will not necessarily obey the law.

The example illustrates the increasing use of Benford's law in detecting anomalous or fraudulent data. Suppose I were to 'invent' a whole lot of expenses on a tax return, say. Then it is likely that the first digits of these fictitious expenses would differ significantly from the distribution shown in the graph.

A law observed by S. Newcomb in 1881 and confirmed empirically by F. Benford in 1938. It is often stated as applying to 'naturally occuring numbers' (e.g. physical constants) or 'numbers having dimension' (e.g. physical measurements). A sound probability-theoretic explanation of the law was finally provided by Theodore Hill in 1996. He showed that suitably combining data from different data sets (say, food expenses, travel expenses, mortage payments, etc) will produce precisely the special kind of data set which obeys Benford's Law: a 'scale-invariant' distribution.

Web link: projecteuclid.org/euclid.ps/1311860830

Further reading: *Benford's Law: Theory and Applications* by Steven J. Miller (ed.), Princeton University Press, 2015.