



Benford's Law, Forensic Mathematics, and Mathematics in Politics

Jackson Howell



An Observation in 1881

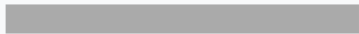
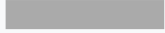
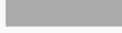
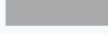
Simon Newcomb: “That the ten digits do not occur with equal frequency must be evident to any one making use of logarithmic tables, and noticing how much faster the first pages wear than the last ones”

N	0	1	2	3	4	5	6	7	8	9	Differences.
100	000000	0434	0868	1301	1734	2166	2598	3029	3461	3891	433 430 425 420
101	4321	4751	5181	5609	6038	6466	6894	7321	7748	8174	1 44 43 42
102	8600	9026	9451	9876	*300	*724	1147	1570	1993	2415	2 85 86 85 84
103	012837	3259	3680	4100	4521	4940	5360	5779	6197	6616	3 121 120 118 116
104	7033	7451	7868	8284	8700	9116	9532	9947	*361	*775	4 174 173 170 168
105	021189	1603	2016	2428	2841	3252	3664	4075	4486	4896	5 218 215 213 210
106	5306	5715	6125	6533	6942	7350	7757	8164	8571	8978	6 261 258 255 252
107	9384	9789	*195	*600	1004	1408	1812	2216	2619	3021	7 305 301 298 294
108	033424	3820	4227	4628	5029	5430	5830	6230	6629	7028	8 348 344 340 336
109	7426	7825	8223	8620	9017	9414	9811	*207	*602	*998	9 392 387 383 378
110	041383	1787	2182	2576	2960	3352	3755	4148	4540	4932	415 410 405 400
111	5323	5714	6105	6495	6885	7275	7664	8053	8442	8830	1 42 41 41 40
112	9218	9606	9993	*380	*766	1153	1538	1924	2309	2694	2 83 82 81 80
113	053078	3403	3846	4230	4613	4996	5378	5760	6142	6524	3 125 123 122 120
114	6905	7286	7666	8046	8426	8805	9185	9563	9949	*370	4 166 164 162 160

The Pattern

- For many real-life random data sets, the leading digit is likely to be small
- In 1938, Frank Benford specified a probability function for leading digits
- Data are said to follow Benford's Law if they exhibit this pattern

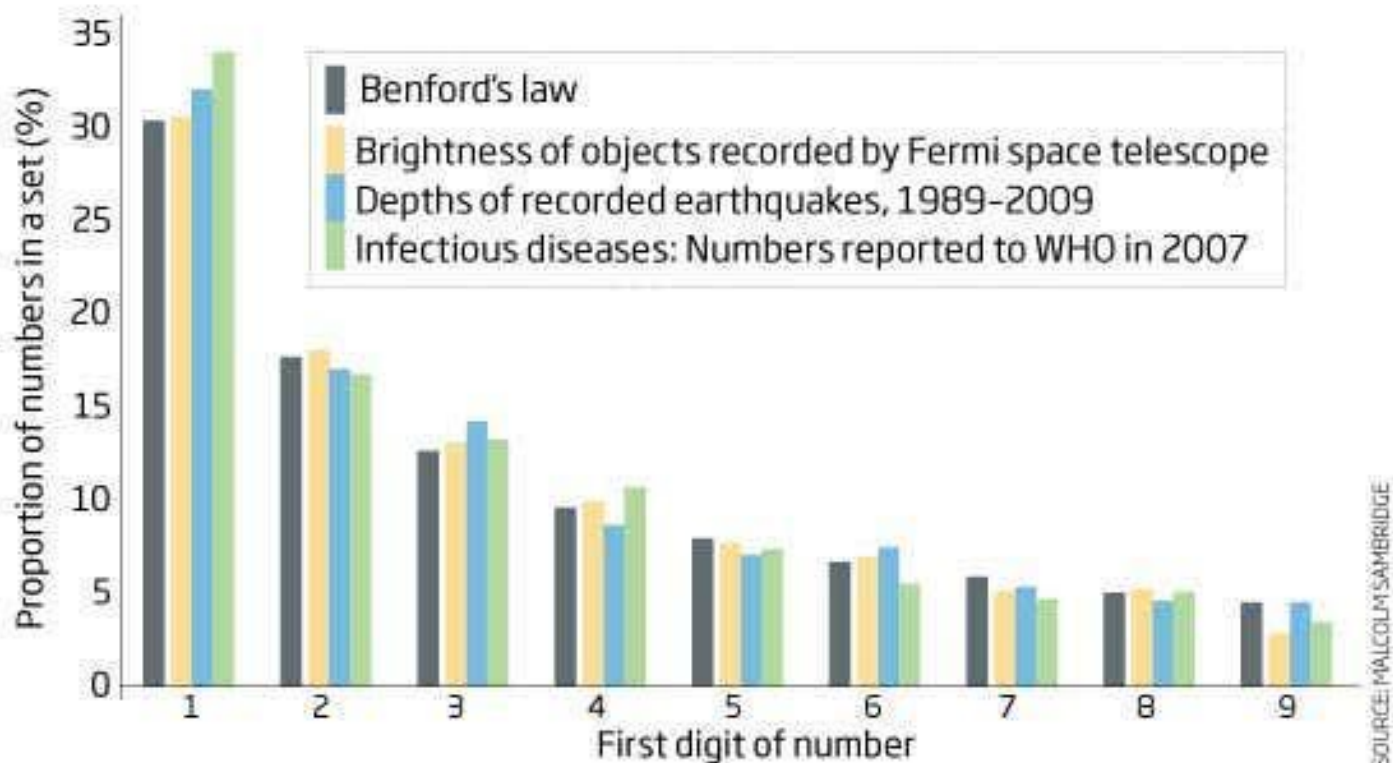
$$P(d) = \log_{10}(d + 1) - \log_{10}(d)$$

d	$P(d)$	Relative size of $P(d)$
1	30.1%	
2	17.6%	
3	12.5%	
4	9.7%	
5	7.9%	
6	6.7%	
7	5.8%	
8	5.1%	
9	4.6%	

Natural law?

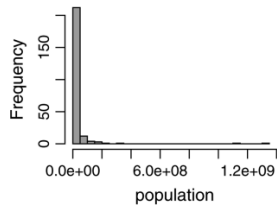
©NewScientist

Under Benford's law, the first digits of the numbers in a set have a non-random distribution. Such sets are being discovered throughout nature

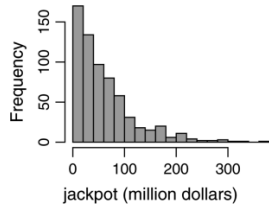


To Be, or Not to Be (Benford)

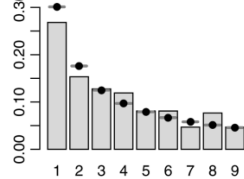
Population of world states



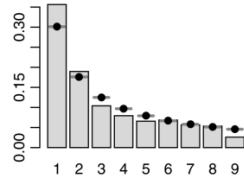
Powerball jackpots



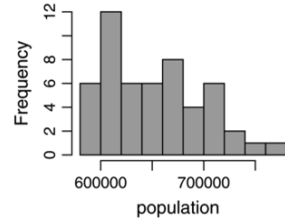
Leading digits ($p=0.59$; $n=235$)



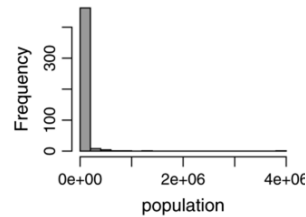
Leading digits ($p=0.01$; $n=654$)



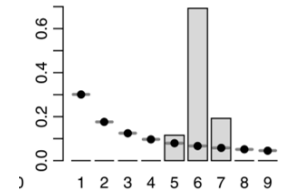
Congressional districts in CA



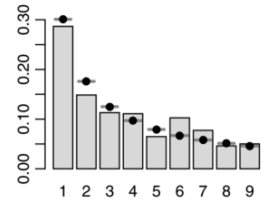
Cities in CA



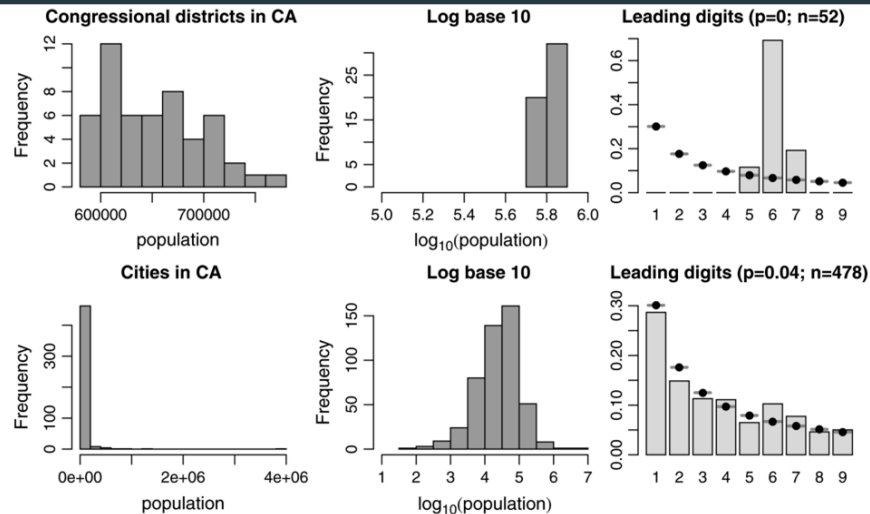
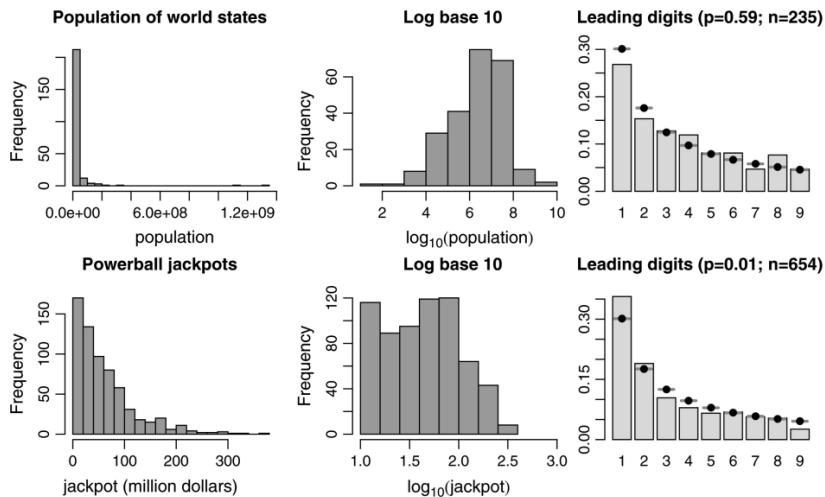
Leading digits ($p=0$; $n=52$)



Leading digits ($p=0.04$; $n=478$)



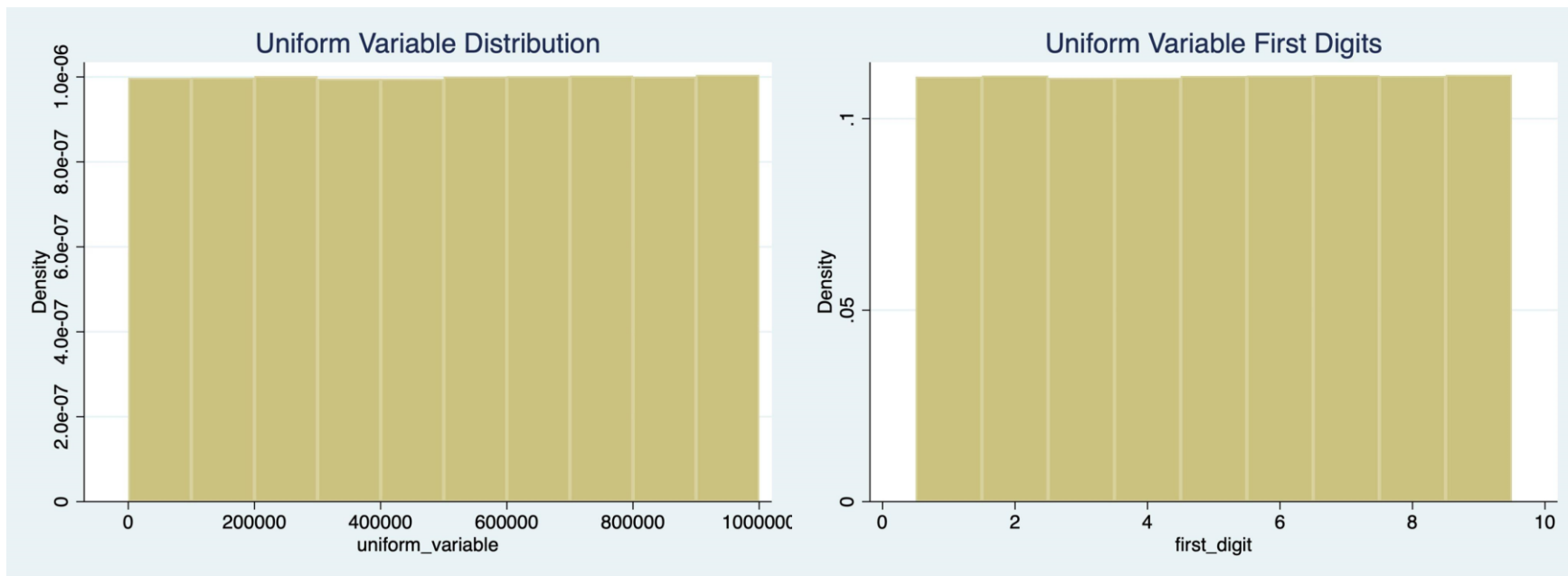
To Be, or Not to Be (Benford)



Common Simplification for Describing Benfordness

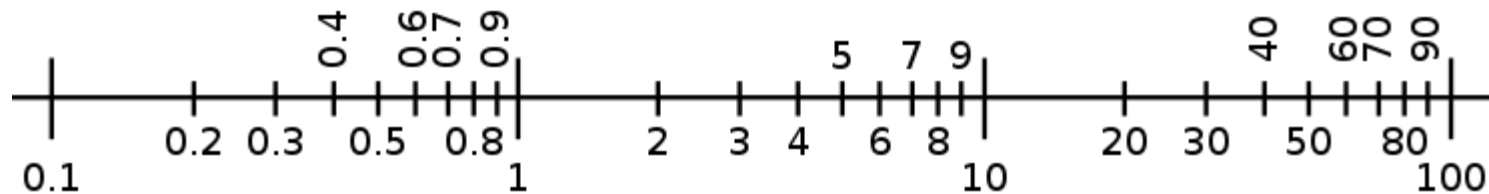
If data are spread out over many orders of magnitude,
they will conform to Benford's Law

Non-BL-Conforming Distributions



Distance Interpretation

$$P(d) = \log_{10}(d + 1) - \log_{10}(d)$$

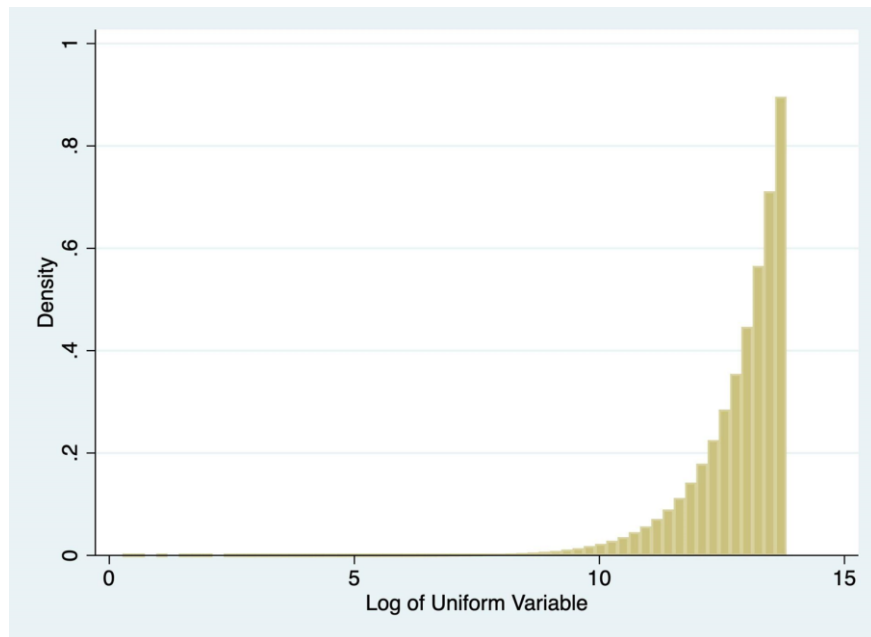




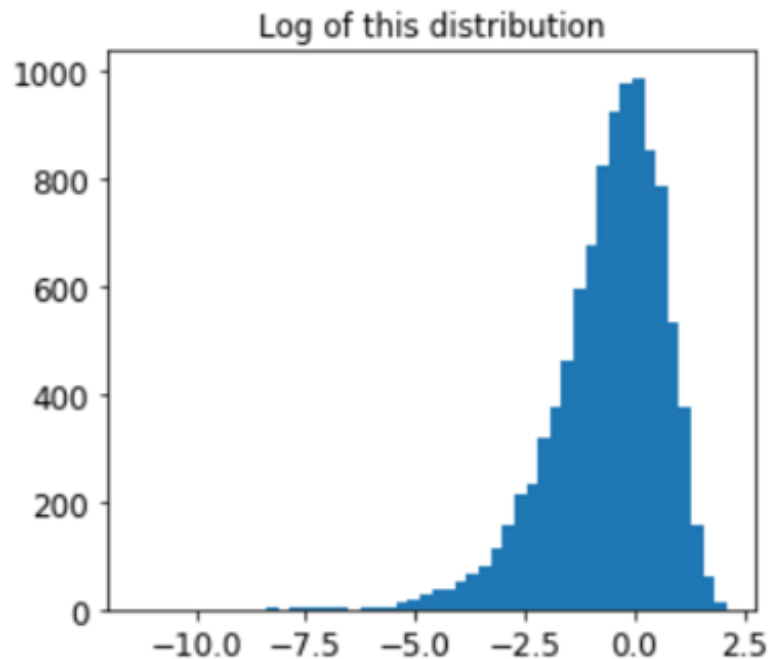
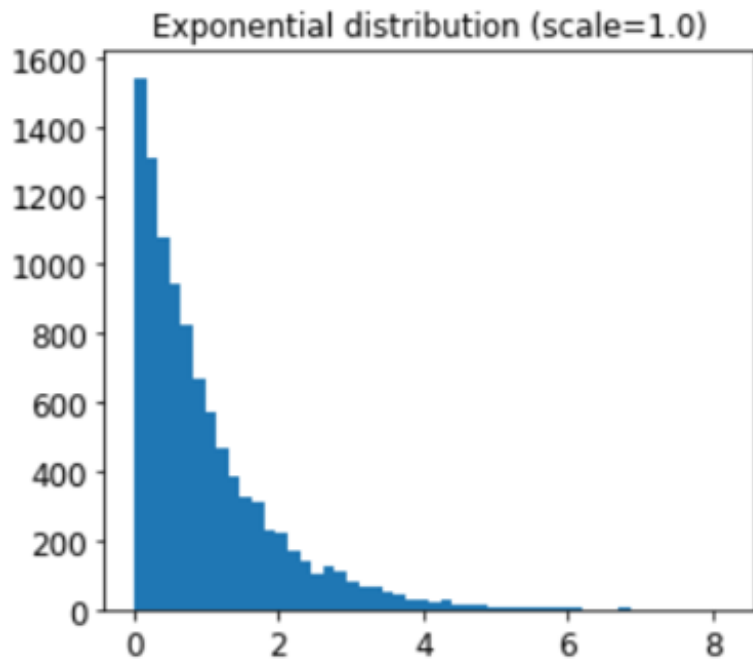
In general, data that form
a “broad” distribution on a
log-linear scale comply to
Benford’s Law



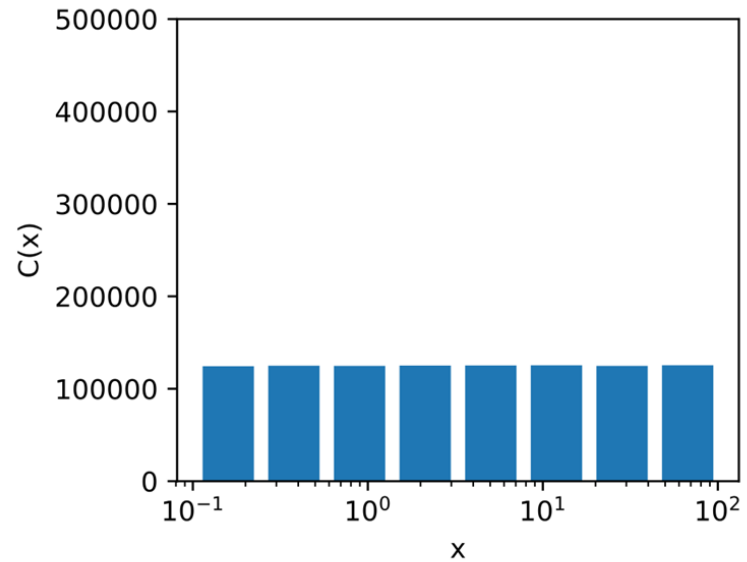
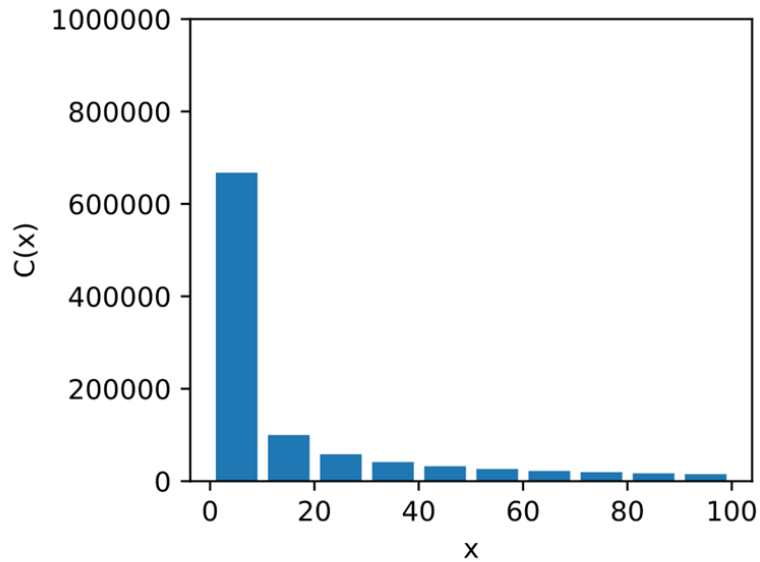
Logarithm of the Uniform Variable



Benfordy Distributions: Exponential



Canonical Benford Distribution: Reciprocal





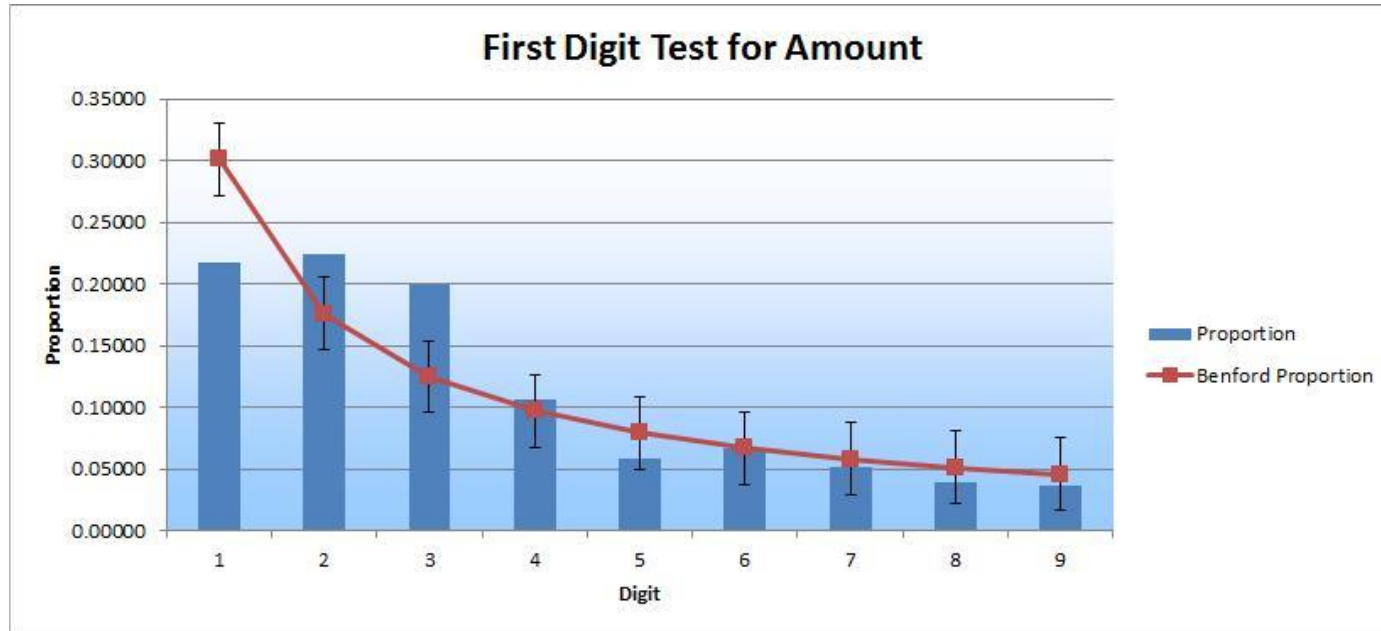
“There is currently no simple intuitive argument to explain the appearance of Benford’s Law in the wide array of contexts in which it has been observed, including statistics, number theory, dynamical systems, and real world data.” (Berger and Hill 2020)



Forensic Use of Benford's Law

- Various types of accounting data (particularly company expenses) are found to follow Benford's Law
- Some accountants have begun to use first-digit analysis as a **first-sweep** to find suspicious data

Example from Pete Miller, CPA



Two-Digit Example of Benford's Law

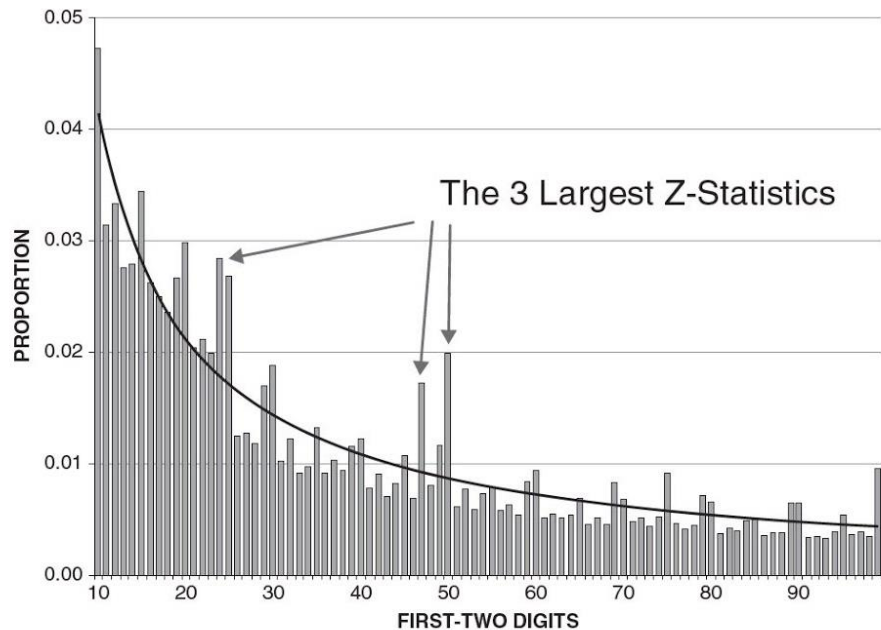


FIGURE 4.6 The First-Two Digits of the Purchasing Card Transactions



Benford's Law does not
prove anything, it only flags
unexpected behavior!

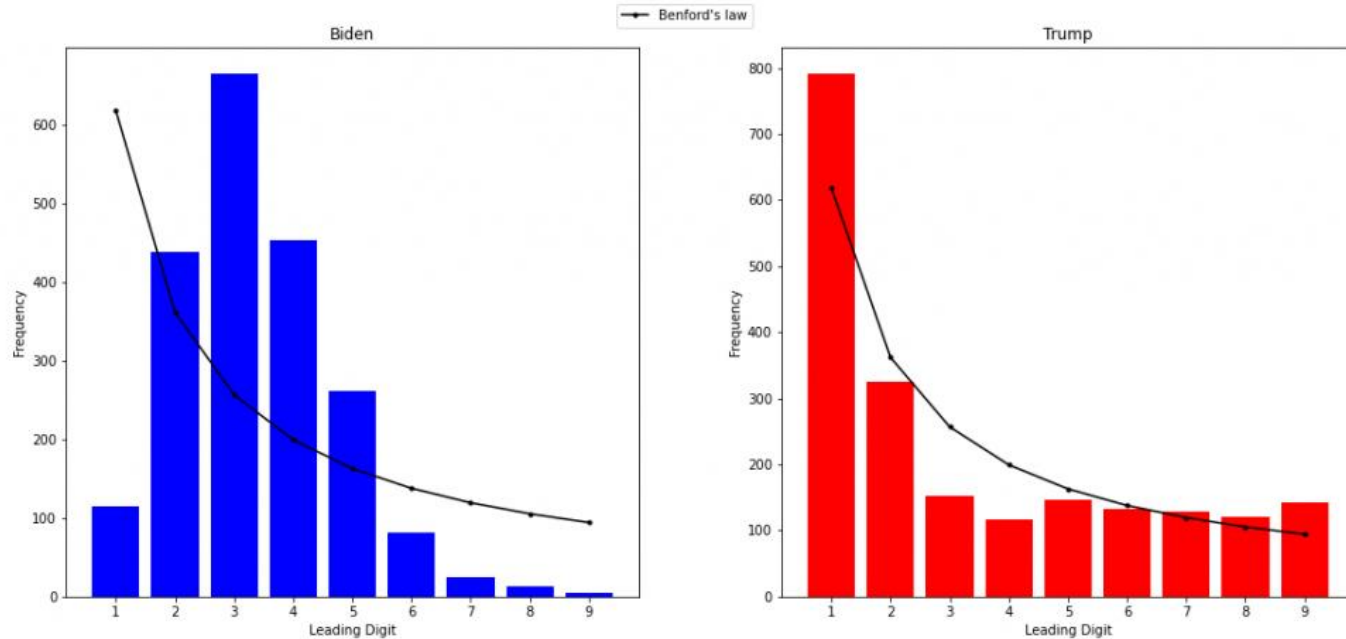




Is it the best way to flag this
behavior?

2020 Election: Evidence of Fraud

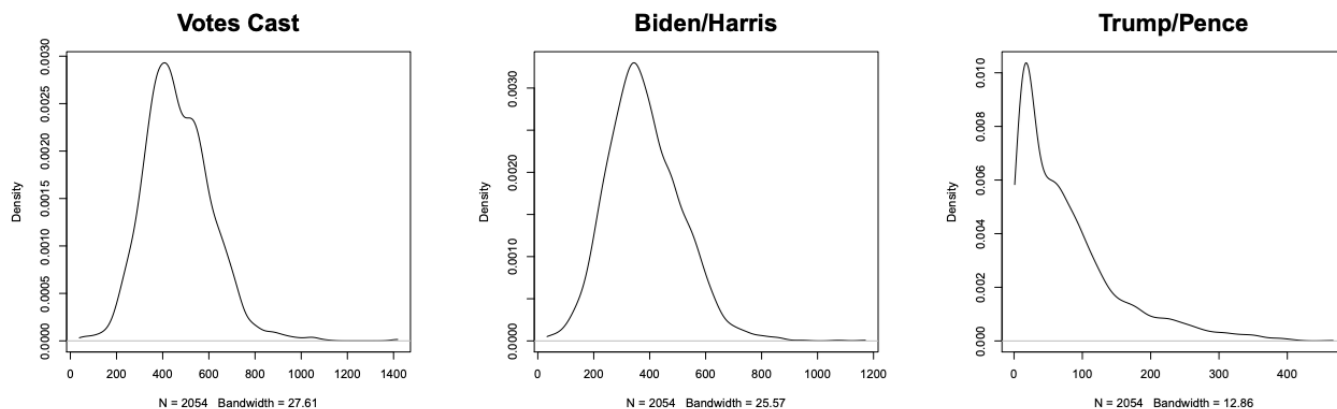
Frequency of Leading Digits in Chicago Precincts



Limitations of Benford's Law to Find Election Fraud

- Election precincts are similarly sized (not over many orders of magnitude)
- In Chicago, Biden won huge majorities of the precincts (average share was 85%), often over 100 votes
- Many Trump vote totals were 0, or single digits

Figure 1: Chicago 2020 Presidential Vote Distributions



Note: plots of empirical densities.

Benford's Law and Sore Losers

POLITICS

Larry Elder Announced He “Detected Fraud” in California Recall Vote Results When They Didn’t Yet Exist

BY BEN MATHIS-LILLEY SEPT 14, 2021 • 2:08 PM



“Statistical analyses used to detect fraud in elections held in 3rd-world nations (such as Russia, Venezuela, and Iran) have detected fraud in California resulting in Governor Gavin Newsom being reinstated as governor. The primary analytical tool used was Benford’s Law and can be readily reproduced.”

Discussion Questions

- Do you think Benford's Law is a compelling tool for forensic data analysis?
- What do you think of mathematics being used in this way in the mainstream?
- How should mathematicians communicate ideas to a broad, politically charged audience?
- Is there a growing class of folk/lay mathematicians?