Lecture 9: Distributions to Excite You: Power Law and Benford's Law

1.**Power law. "Frequent events small/ rare events big."**

A power law is an empirical relation between Y, usually the frequency of an event – # of times it occurs -- and S, which measures size of the event: $\mathbf{Y= B\ S^{-a}}$ with a fixed power a, and constant B. The term power refers to the relation between Y and S via powers: a could be 2 so relation is 1/quadratic. Taking ln, power law is a straight line between ln Y and ln S: $\ln Y = \ln B - a \ln S$. Differentiating we have $d\ln Y/d\ln S = - a$ (constant elasticity relation). Note you can also write relation as $\ln S = -1/a \ln Y -1/a \ln B$.

Big events rare/small events frequent is not vacuous. Normal distribution has most observations around the mean and is symmetric wrt size: large and small and big values have equal representation – human height fits normal. Uniform distribution has equal likelihood of events by size– throwing dice 6 comes up as often as 1.

Another way to describe power law is by relating rank of object by size, with rank as measure of where you are on a distribution function of objects. You are 1st, 2nd, 3rd … nth in ordering. Rank cities by population and you have a power law relation known as **ZIPF'S LAW** (named after George Zipf (Harvard, Human Behavior and the Principle of Least Effort (1949)) between population size/ frequency S to the **rank R** of the object: $\mathbf{S = BR^{-b}}$, where rank goes from 1 for biggest to n for nth on list. Note b = 1/a in power law form. Krugman claimed that Zipf law holds for US cities with -1 coefficient When a is 1, the power law is: Size times Rank = constant.: SR= B or $\log S = \log B - \log R$

Eliminate the scaling factor B by division of the cities by rank over the next lower rank gives

Size (City N+ 1)/Size (City N) = (N+1)/N	**Size (City N) /Size (City 1)**
City Size (2)/City Size (1) = ½	City Size (2)/City Size (1) = ½
City Size (3)/City Size (2) = 2/3	City Size (3)/City Size (1) = 1/3
City Size (4)/City Size (3) = 3/4	City Size (4)/City Size (1) = 1/4
City Size (5)/City Size (4) = 4/5	City Size (5)/City Size (1) = 1/5

US Cities in 2015	pop millions	Ratio of largest	Predicted Ratio
1. NYC	8.6	1.0	1.0
2 LA	4.0	0.47	0.5
3, Chicago	2.7	0.31	0.33
4. Houston	2.3	0.27	0.25
5/6 Philadelphia/Phoenix	1.6	0.19	0.20
22/23 Washington/Boston	0.7	0.08	0.04

But neither urban area of China's cities in 2010 (excluding HK) nor UK cities fit

	pop millions	Ratio to Larger	Predicted Ratio		pop millions	Ratio to Larger	Predicted Ratio
1. Shanghai	22.3	1.0	1.0	London	7,200	1.0	1.0
2 Beijing	19.3	0.87	0.50	Birmingham	992	0.14	0.50
3 Guangzhou	11.1	0.50	0.33	Leeds	720	0.10	0.33
4. Tianjin	11.1	0.50	0.25	Glasgow	560	0.08	0.25
5. Shenzhen	10.4	0.47	0.20	Sheffiel	512	0.07	0.20

We can also look at the power law in terms of the number of cities with different sizes in bins,

US		China		United Kingdom	
>8 million	1	> 12 million	2	> 7.2	1
3 -8 million	1	9-12 million	4	3.6 to 7.2	0
2 -3 million	2	6- 9 million	6	1.8 to 3.6	0
1 to 2 million	5	3.0 - 6	19	.9 to 1.8	1
.37 to 1.0	40	2 -3	16	.45 to .9	5
<.45	134	1 – 2	88	<.45	60

Cities definition depends on political boundaries. SMSAs in US looks quite different.

Pareto's Law is economics' famous power law where size is income. Pareto law says that upper tail of income distribution fits power law with coefficient that produces thicker tail than normal distribution. But Pareto wrote this as function of **cumulative distribution,** not as frequency:

$P(S > s) = s^{-k}$ – the probability that people have incomes above s – is a power function of the value of s. The cumulative distribution is $1 - s^{-k}$, which is the proportion of people below s in the ranking of income. The density $P(S = s)$ is the derivative of cumulative distribution**:** $\mathbf{(k)s^{-(k-1)}}$ (recall calculus $dx^n = nx^{n-1}$)

Pareto's estimated k led him to claim that 20% always hold 80% of wealth. WJ Reed claims power laws at both ends of distribution (http://phys.ubbcluj.ro/~zneda/edu/mc/pareto.pdf) but hard to measure income at lower tail. For upper tail check Forbes' top billionaires 2019. You can find billionaires list for other years and see how they have fared. How did Forbes estimate the wealth of these folk? Estimates for China from Hurun Foundation differ from Forbes. Could # billionaires better measure of inequality in a country than widely used Gini coefficients?

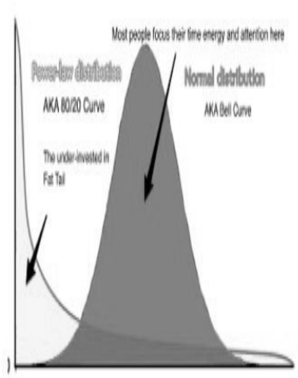
No.	Name	Net worth (USD)	Age	Nationality	Source(s) of wealth
1 –	Jeff Bezos	$131 billion ▲	55	United States	Amazon
2 –	Bill Gates	$96.5 billion ▲	63	United States	Microsoft
3 –	Warren Buffett	$82.5 billion ▼	88	United States	Berkshire Hathaway
4 –	Bernard Arnault	$76 billion ▲	70	France	LVMH
5 ▲	Carlos Slim	$64 billion ▼	79	Mexico	América Móvil, Grupo Carso
6 –	Amancio Ortega	$62.7 billion ▼	82	Spain	Inditex, Zara
7 ▲	Larry Ellison	$62.5 billion ▲	74	United States	Oracle Corporation
8 ▼	Mark Zuckerberg	$62.3 billion ▼	34	United States	Facebook
9 ▲	Michael Bloomberg	$55.5 billion ▲	77	United States	Bloomberg L.P.
10 ▲	Larry Page	$50.8 billion ▲	45	United States	Alphabet Inc.

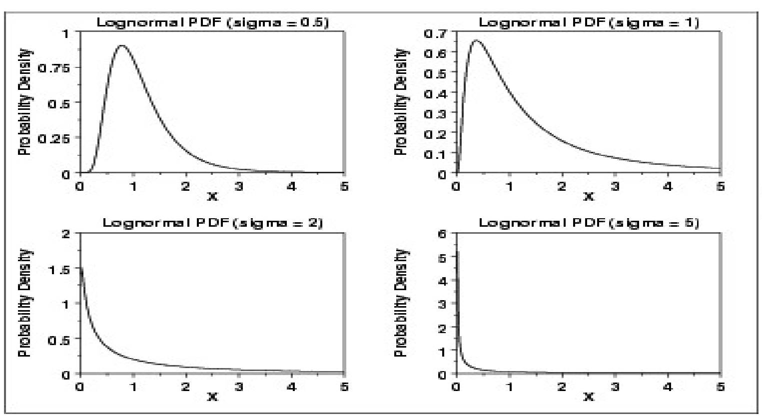
Number and combined net worth of billionaires by year[50]

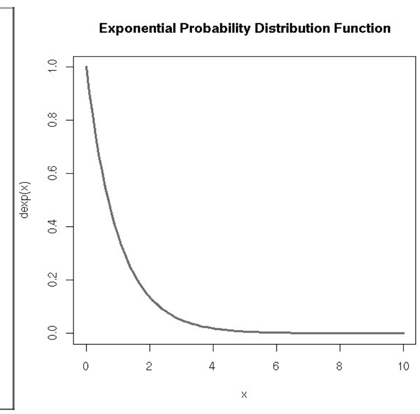
Year	Number of billionaires	Group's combined net worth
2019	2,153	$8.7 trillion
2018	2,208	$9.1 trillion
2017	2,043	$7.7 trillion
2016	1,810	$6.5 trillion
2015[7]	1,826	$7.1 trillion
2014[51]	1,645	$6.4 trillion
2013[52]	1,426	$5.4 trillion
2012	1,226	$4.6 trillion
2011	1,210	$4.5 trillion
2010	1,011	$3.6 trillion
2009	793	$2.4 trillion
2008	1,125	$4.4 trillion
2007	946	$3.5 trillion
2006	793	$2.6 trillion
2005	691	$2.2 trillion
2004	587	$1.9 trillion
2003	476	$1.4 trillion
2002	497	$1.5 trillion
2001	538	$1.8 trillion
2000	470	$898 billion

Source: *Forbes*.[7][51][50][52]

Power laws have THICK tails that could be so thick that the distribution has NO second moment/variance. In such distributions (a < 3 for $\mathbf{Y = S^{-a}}$) you can calculate an empirical variance but theoretical variance is infinite. Power laws are not the only way to represent data with a long tail. The log-normal becomes more power-law like as the standard deviation increases. Exponential can also have a long tail. But ln normal and exponential have thinner tails than power law. Can add another parameter to thicken the tail -- stretched exponential.







Power laws or power law type relations are found in all sorts of data. See AClauset, C Rohilla Shalizi, M. E. J. Newman http://arxiv.org/abs/0706.1062 – a paper which has the most mentions of power law in any paper: the number of power law mentions in a paper follows as power law!.**A**nything that fits lots of phenomenon deserves attention.

Power laws are SCALE INVARIANT: distribution does not change with units/scale of variable. Take $\mathbf{Y = S^{-a}}$. Measuring S in λS units affects the constant but not a :$\mathbf{Y = \lambda^{-a} S^{-a}}$ so $\ln Y = -a \ln \lambda - a \ln S$. Note scale invariance deserves attention since it raises possibility that you can learn about the rare events from frequent events.

Kauffman, Bak, Gell-Mann argue that power laws are the signature of complex adaptive systems that lead to a state of self-organized criticality.

KEY FACT: PARETO , ZIPEF AND POWER LAW ARE THE SAME

Lada Adamic shows that Pareto and Zipf are alternative cumulative distribution representations of the same Power law with independent and dependent variables reversed and key parameters transformations of each .

The Pareto distribution: $P(S > s) = s^{-k}$ – the probability that people have incomes above s is a power function of the value of s. The cumulative distribution is $1- s^{-k}$ is the proportion of people below s in the ranking of income (ie cumulative distribution is position/rank in a distribution). The frequency/density is $P(S = s)$ **$ks^{-(k+1)}$.**

Pareto, Zipf and power laws are the same fat-tailed distribution, shown in different forms.

The power law linking frequency to size of objects is $Y= BS^{-a}$ so this is just Pareto with $a = k+ 1$. Since rank is a position that reflects where the object is in the distribution, **Zipf relates size to rank.** When a city of size s has rank R, then there are R cities with size >s. Rewrite Zipf in cumulative frequency/rank form as $R = B^{1/b} s^{-1/b}$. Divide by the total number of cities T to get $R/T = (B^{1/b}/T)$ $\mathbf{s^{-1/b}}$. R/T is the proportion of cities with size > s. Thus Zipf proportion with size > s is $(B^{1/b}/T)s^{-1/b}$. But the Pareto proportion with size> S is s^{-k}. So $k = 1/b$. The **Zipf coefficient is 1/ Pareto coefficient**. Also, note that density for Zipf is derivative of $(B^{1/b}/T)s^{-1/b} = -1/b(B^{1/b}/T)\, s^{-1/b-1}$

Summary Table

Distribution	“dependent” Measure	Right hand side measure	Coefficient for density
Power law	Density	Size S^{-a}	-a
Pareto	Upper tail cumulative	s^{-k} cum is $1-s^{-k}$	
	Density	s-(k+1)	- (k+1)
Zipf	Size	Rank/upper tail R^{-b}	
	Rank/upper tail	Size $s^{-1/b}$	
	Density	$s^{-1/b-1}$	- (1/b +1)

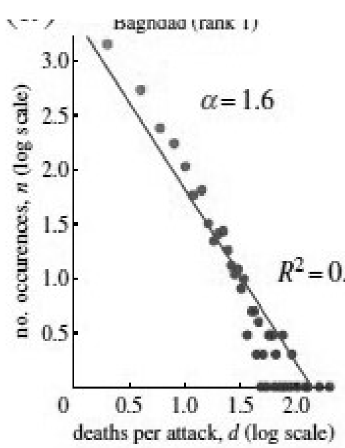
Thus all three forms represent the same **power law coefficient for density a = k +1= 1+ 1/b**

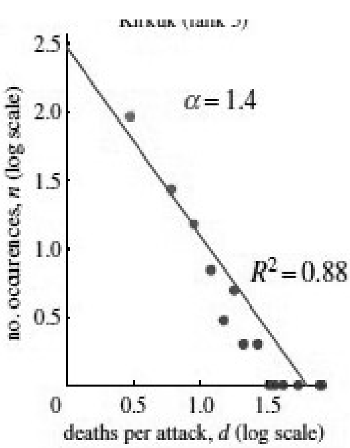
Interesting Power Laws:

1)Global terrorism follows a power law 10 February 2005 arXiv.org/abs/physics/0502014.

Clauset and Young analyze database with of more than 19,900 terrorist events in 187 countries between 1968 and 2004 (https://en.wikipedia.org/wiki/MIPT_Terrorism_Knowledge_Base) where at least one person was killed or injured in some 7,088 of these events. They found that the probability of an event with a severity of x or higher was proportional to $x^{-\alpha}$, where α close to two and that the distributions did not fit other "heavy-tailed" distributions like a log-normal curve. This *“means”* that extreme events like September 11 are not "outliers" but part of the overall pattern of terrorist attacks that is "scale invariant" with frequency and severity of terrorist attacks - number of deaths plus the number of injuries - related by a power law. For more recent data see http://www.rand.org/nsrd/projects/terrorism-incidents.html; has 40,000 incidents and http://www.start.umd.edu/gtd/, global terrorism report > 190,000 incidents.

"Unfortunately, the implications of the scale invariance are almost all negative," Clauset and Young told *PhysicsWeb*. "... because the scaling parameter is less than two, the size of the largest terrorist attack to date will only grow with time. *If we assume* that the scaling relationship and the frequency of events do not change in the future, we can expect to see another attack at least as severe as September 11 within the next seven years."





Guo W. 2019 Commonstatistical patterns in urban terrorism. R. Soc.open sci. 6: 190645 tBy examining over 30 000 geotagged terrorism acts over 7000 cities worldwide from 2002 to today, the results show the following. All cities experience attacks A that are uncorrelated to the population and separated by a time interval t that is negative exponentially distributed with a **death-toll per attack that follows a power law distribution**. The prediction parameters yield a high confidence of explaining up to 87% of the variations in frequency and 89% in the death-toll data. These findings show that the aggregate statistical behaviour of terror attacks are seemingly random and memoryless for all global cities

2) Citations to Academic Papers – Golosovsky M, Solomon S. Runaway events dominate the heavy tail of citations distributions. Eur. Phys. J. Special topics. 2012; 205: 303–311

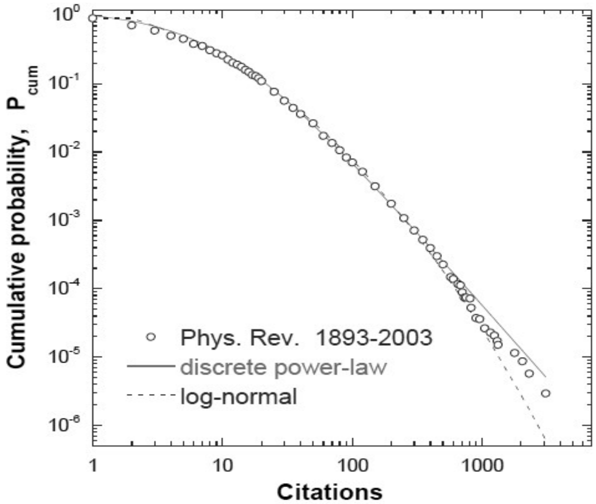


Fig. 1. Cumulative probability distribution (cdf) of citations to 353,268 papers published in Physical Review journals during 1893-2003 and cited by 2003. Only PR to PR citations were counted. The data were adapted from Ref. [13]. The continuous red line shows a fit with the discrete-power-law cdf (Eq.1) with $\gamma = 3.15, w = 10.2$. The dashed blue line shows a fit with the log-normal cdf (Eq.3) with $\mu = 1.15, \sigma = 1.42$.

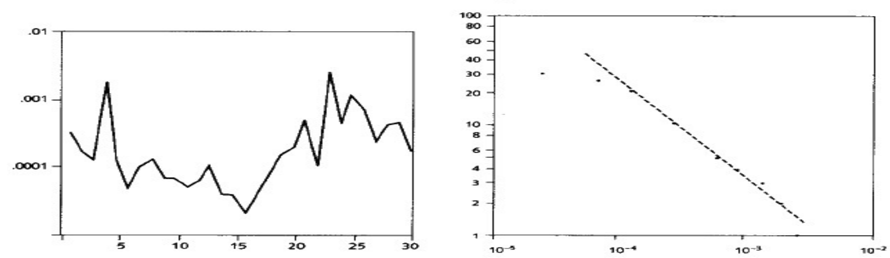
3)-- Shares in Facebook posts – similar power law tail from Tsallis-Pareto.--
https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0179656

RESEARCH ARTICLE

Science and Facebook: The same popularity law!

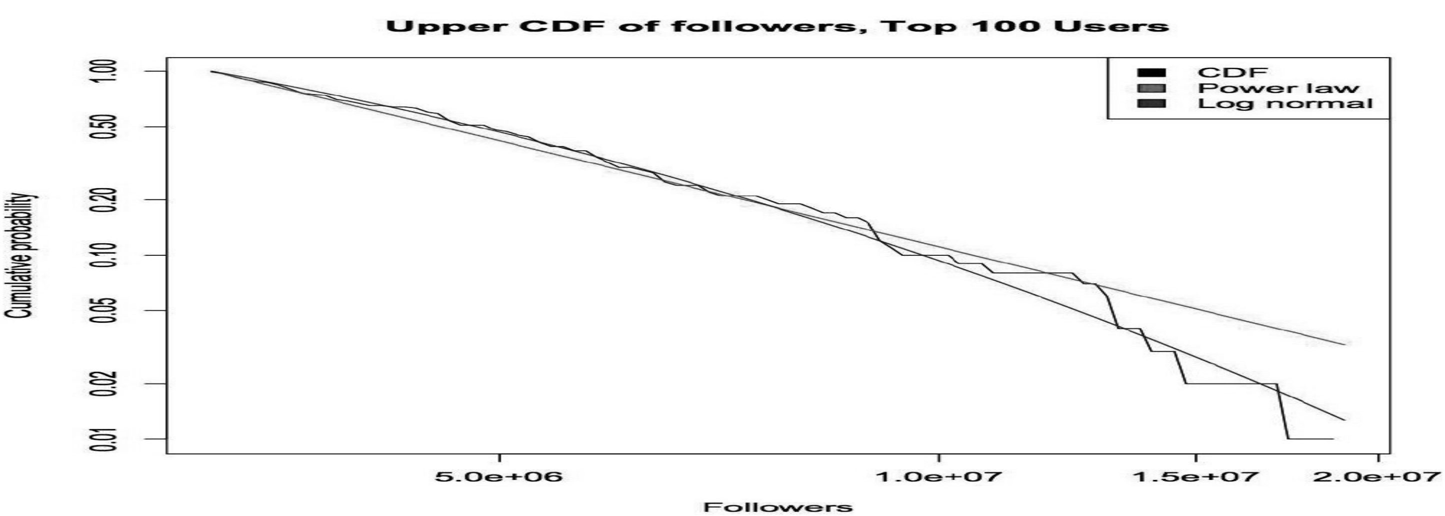
4) Mandelbrot: fat tail of changes in commodity/security prices: more months with small variation than with large variation so that transforms of variables fits a log-log curve. Is this true of all price variation? How about wages?

Power Law Relationships – Cotton Prices



Mandelbrot's (1963) analysis of monthly variations in cotton prices during a 30 month period. The left plot shows the month by month changes. Note how they vary; lots of small changes, and fewer large changes. The right logarithmic graph shows the same data is a power-law distribution, indicating the cotton

5)Does Twitter follow Power Law? –



6) Clauset "Trends and fluctuations in the severity of interstate wars"Science Advances 2018;4: February 2018
"Richardson's original analysis of interstate wars from 1820 to 1945 (32) ... argued that war sizes followed a precise pattern, called a power-law distribution, in which the probability that a war kills x people is Pr(x) º x−a , where a > 1 is … "scaling" parameter and x ≥ xmin > 0. He also argued that the timing of wars followed a simple Poisson process, implying a constant annual hazard rate and an exponential distribution for the time between wars (24, 25). Although Richardson's analysis would not be considered statistically rigorous today but these patterns...(still hold).

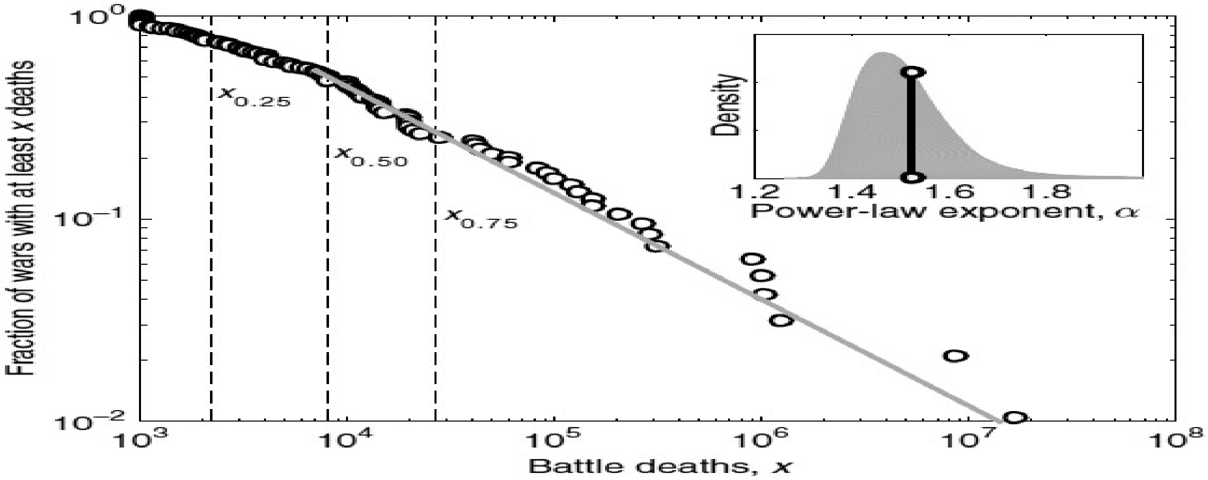
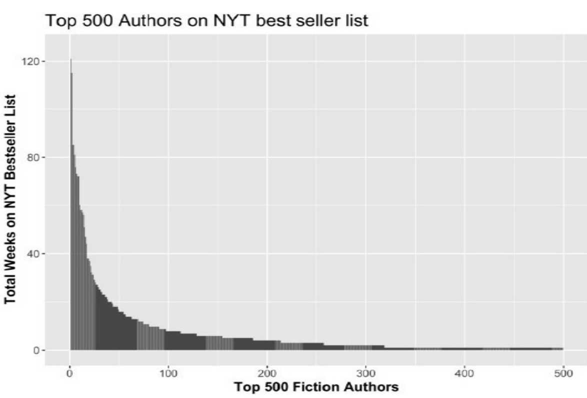
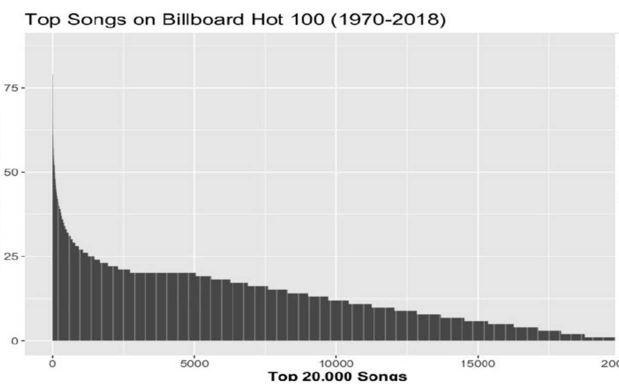


Fig. 2. Interstate wars sizes, 1823–2003. The maximum likelihood power-law model of the largest-severity wars (solid line, $\hat{\alpha} = 1.53 \pm 0.07$ for $x = \hat{x}_{min} = 7061$) is a plausible data-generating process of the empirical severities (Monte Carlo, $p_{KS} = 0.78 \pm 0.03$). For reference, distribution quartiles are marked by vertical dashed lines. Inset: Bootstrap distribution of maximum likelihood parameters $\Pr(\hat{\alpha})$, with the empirical value (black line).

7)P**ower Law in Popular Media – Michael Tauberg**





Debating Value of power laws

1–YES Parsimonious way to describe data from distributions with a scale free structure that allows us to use distribution at one scale to predict distribution at another and possibly learn about mechanism behind few big events from larger sample of smaller events.

NO Curve-fitting is not understanding. Power law is ONE parametization of a distribution where big events are rarer than small events. Other heavy tailed curves may fit better and evidence that ln-normal works well until upper tail and then the Pareto fits, but may itself fail at extreme, suggests viewing power law as descriptive of part of distribution.

NO. Estimating a in power law is dicey because so few tail observations so power law lives on shaky grounds. Mandelbrot adds two parameters and gets a better fit for city power laws so that the first city would have size B/ (n + const)^t. But adding parameters to fit data does not lead to good generalization. And many are not scale free , so no reason to trust extrapolations.

YES scale-freeness is intended as an idealized model, not something that precisely captures real-world networks. Many of the most important properties of scale-free networks, they say, also hold for a broader class called "heavy-tailed networks" to which many real-world networks may belong (these are networks that have significantly more highly connected hubs than a random network has, but don't necessarily obey a strict power law).

2—YES Power laws can lead to good policy. If small number cause most of problem or most of gains, focus attention on that group. Allows us to escape from details to focus on underlying dynameics This motivates the Per

Bak/Kauffman efforts to develop Self-Organized Criticality to account for power laws.

NO General all-encompassing rules rarely tell us what to do. The only way to figure out which mechanism causes a particular outcome is to study the details of the system. Extinctions fit power laws but cause is exogenous (asteroid crash) rather than some rule of life, while asteroid crashes do not explain power law in city sizes and income.

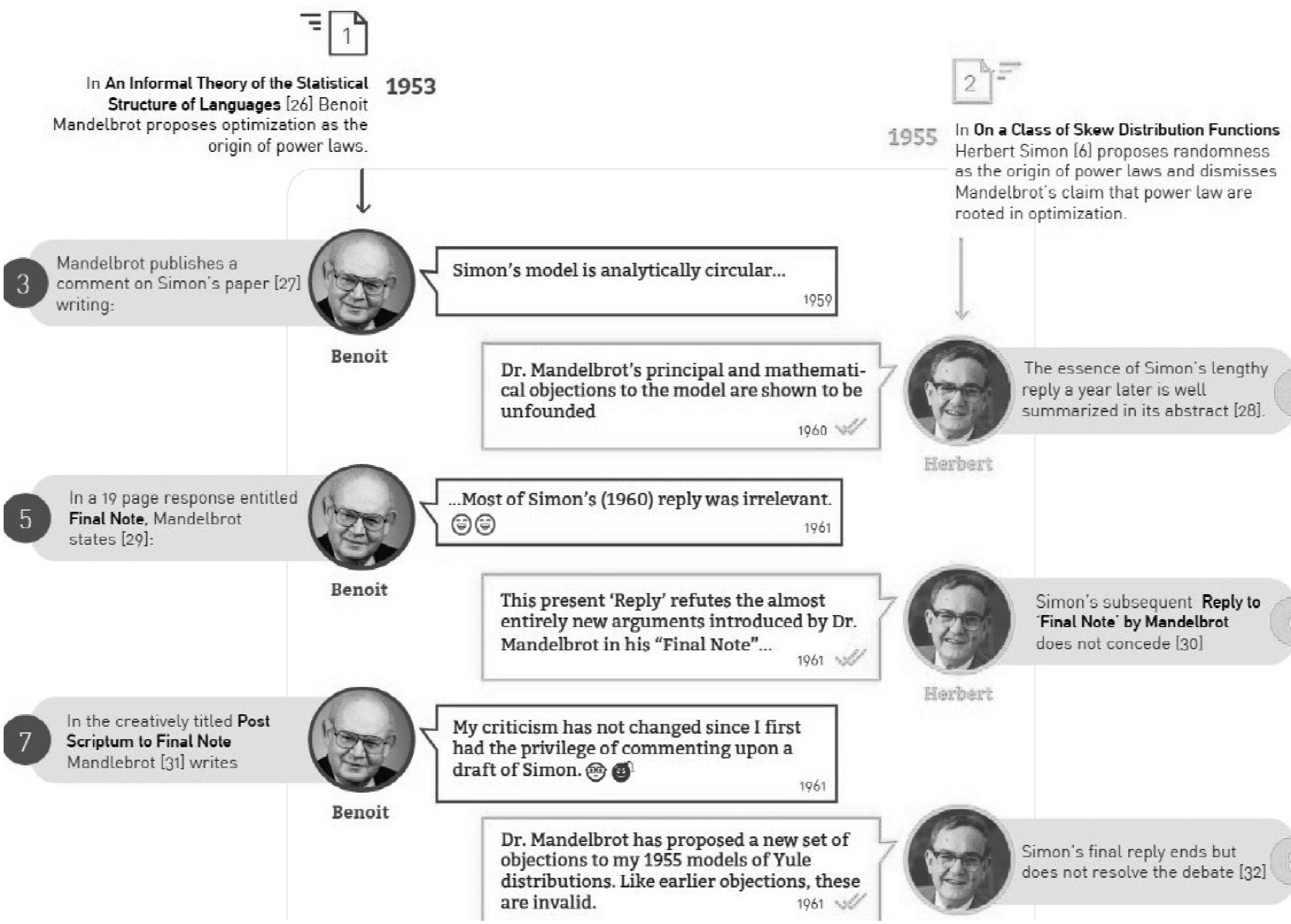
3. NO Even if there are power laws, parameters may change. So you could have policies that would alter the parameters of the power curve. Olson and Associates found that a scaling relation for the variability of Italian currency fluctuations – lots of small swings, few large swings – had smaller coefficients when Italy was part of the European Monetary System in the 1980s - 1990s than when it was out. Pareto parameters can differ across countries, time.

War over Interpreting Power Law Patterns: Results from Opimizing Behavior vs Random linkages (Barabasi)

LUCK OR REASON: AN ANCIENT FIGHT

The tension between randomness and optimization, two apparently antagonistic explanations for power laws, is by no means new: In the 1960s Herbert Simon and Benoit Mandelbrot have engaged in a fierce public dispute over this very topic. Simon proposed that preferential attachment is responsible for the power-law nature of word frequencies. Mandelbrot fiercely defended an optimization-based framework. The debate spanned seven papers and two years and is one of the

In the context of networks today the argument titled in Simon's favor: The power laws observed in complex networks appear to be driven by randomness and preferential attachment. Yet, the optimization-based ideas proposed by Mandelbrot play an important role in explaining the origins of preferential attachment. So at the end they were both right.



2. Benford's Law – a useful clue to uncovering fraud

I.

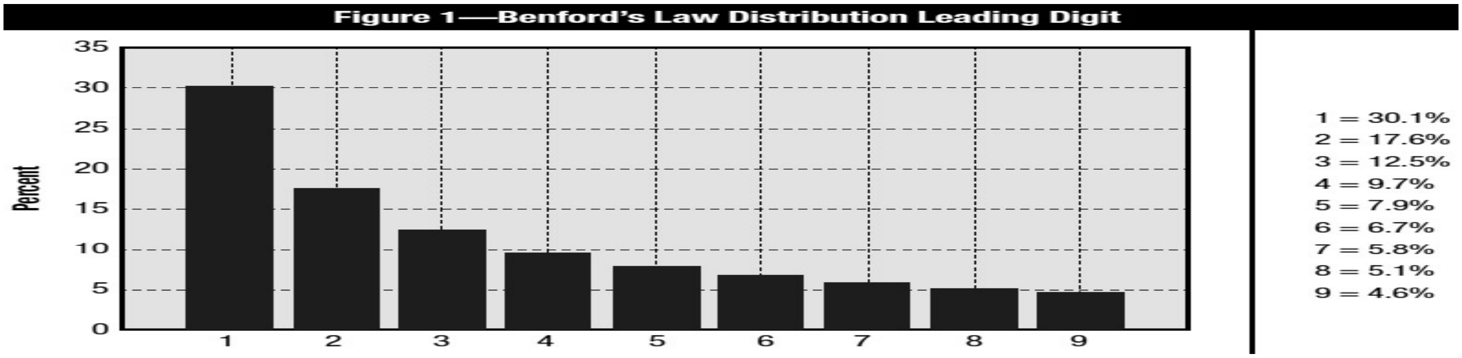
There are 9 digits (and 0) in our base 10 arithmetic. Randomly select a digit, exclusive of zero, and what are the odds the digit is 2, 7, 4, …? Since there are nine digits, you might guess 1/9 or about 11%

Here is the distribution of leading digits from different statistical series (http://testingbenfordslaw.com/):

Digit	Twitter followers	Distance of stars	UK govt spending	Spanish cities pop	Google books one-grams
1	32.62%	30.00%	29.10%	31.07% 1.	28.32%
2	16.66%	14.67%	17.50%	18.02%	16.45%
3	11.80%	12.00%	12.20%	12.42% 2.	13.24%
4	9.26%	10.33%	9.60%	9.18%	10.66%
5	7.63%	9.33%	8.60%	7.95% 3.	7.88%
6	6.55%	5.67%	7.30%	6.57% 4.	7.20%
7	5.76%	7.00%	6.10%	5/36%	5.98%
8	5.14%	7.00%	5.60%	4.95%	5.11%
9	4.56%	4.00%	4.60%	4.47%	5.16%
# records	38,670,514	300	190,379	8114	2055
min	1	4	1	5	303
max	4,706,631	3000	999,994	3,255,944	13,598,879,452
Magnitude	6	3	5	6	10

1.What is Benford's Law of leading digits? That the digits in any set of numbers follows a logarithmic pattern, with the first digit D having the frequency of $\mathbf{\log_{10}(1+ 1/D)}$ not 1/9. Simon Newcomb's 1881 paper "Note on the frequency of use of the different digits in natural numbers". American Journal of Mathematics. 4(1/4): 39–40) said the probability of N being the first digit of a number was equal to log(N + 1) − log(N). There are logarithmic probabilities for other digits as well! And similar laws for other bases. Wikipedia article is superb on the law.

Figure 1—Benford's Law Distribution Leading Digit



Newcomb discovered the law in 1881 when he noticed that log table pages were grubbier around the number 1 than numbers 8 or 9 and fit the formula. But it is called Benford's Law after Frank Benford, a GE physicist rediscovered the pattern and wrote a paper in 1938 that showed that distribution of first digits of 20,229 sets of numbers from areas of rivers to physical constants and death rates followed the law (*The law of anomalous numbers. Proceedings of the American Philosophical Society 78, 551-572).* In 1961 Roger Pinkham, Rutgers mathematician, claimed that any general law of digits must be **scale invariant** (independent of units – ie log) and that it had to be Benford's law, but Pinkham implicitly assumed that there exists a scale-invariant probability distribution on the positive real numbers, which is not so. In 1996 Feller's classic text <u>An Introduction to Probability</u> claimed that "regularity and large spread implies Benford's Law" but this was wrong.

In the 1995 Ted Hill, of West Point, later Ga Tech, proved that data resulting from **a mix of factors** will follow Benford's Law, and is "absorbing" in the sense that it causes products of numbers and other distributions that incorporate it to obey the law as well. Benford is base and scale invariance. Hill, T.P (1995) "A statistical derivation of the significant digit law" Statistical Science , 10(4), 354-363.   Hill and Berger write that there is "No **Simple** Explanation In Sight For [the] Mathematical Gem" because it arises from very different processes, sequences, product of Random variables, mixtures of data sets, but there are ways to understand. Hill, a Vietnam veteran with an amazing career in Army, academics, tells about his wild life in math in a memoir **From Beast to Berkeley**, and inadvertently got into recent controversy for math paper on evolution.

Benford Law: $\text{Prob}\,(D_1 = d_1) = \log_{10}\left(1 + d_1^{-1}\right)$ for all $d_1 = 1, 2, \ldots, 9;$

This is a probability distribution because PROBABILITY OF "D" is $\log_{10}(1 + 1/D)$ and the sum of the probabilities for the nine values = 1 bcs $\log_{10}(1+1/1) + \log_{10}(1+1/2) + \log_{10}(1+1/3) + \ldots + \log_{10}(1+1/9) = \log_{10}$ (2 3/2 4/3 ... 9/8 10/9)= $\log_{10} 10$

But this is not all. There is a formula for the 2nd digit as well!

$P(D_2 = d) = \sum \log_{10}(1+(10k+d)^{-1})$ k = 1 to 9.

$$\text{Prob}\,(D_2 = 1) = \sum_{j=1}^{9} \log_{10}\left(1 + \frac{1}{10j+1}\right) = \log_{10} \frac{6029312}{4638501} = 0.1138\ldots,$$

whereas, given that the first digit equals 1, the (conditional) probability that st the second digit equals 1 as well is

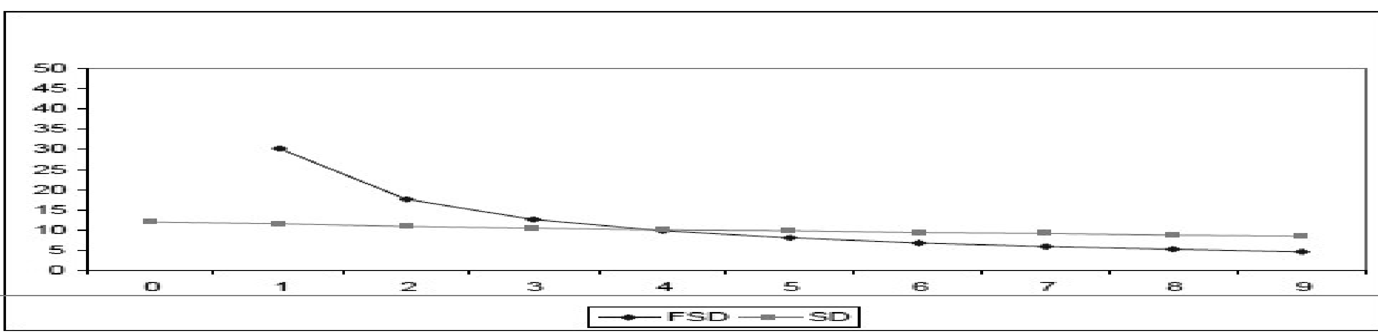
$$\text{Prob}\,(D_2 = 1 | D_1 = 1) = \frac{\log_{10} 12 - \log_{10} 11}{\log_{10} 2} = 0.1255\ldots.$$

The law for a set of digits is:

$$\text{Prob}((D_1, D_2, \ldots, D_n) = (d_1, d_2, \ldots, d_n)) = \log_{10}\left(1 + \left(\sum_{j=1}^{n} 10^{n-j} d_j\right)^{-1}\right)$$

But as digits go up, the probabilities → 1/9 so the first digit deviation from 1/9 is much greater than other digits:

Figure I. Benford's Law First and Second Digits Frequencies



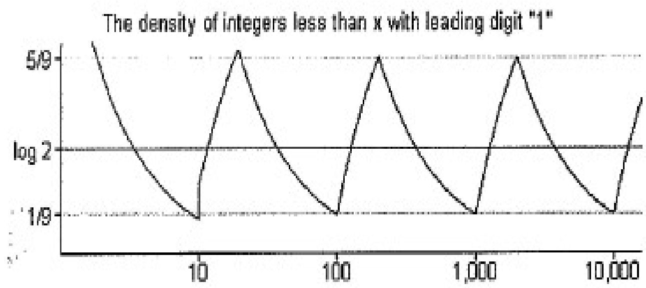
If first digits have different distribution, watch out! Accounting/business deviations often reflect judicious rounding, but some may indicate manipulation – fake numbers.

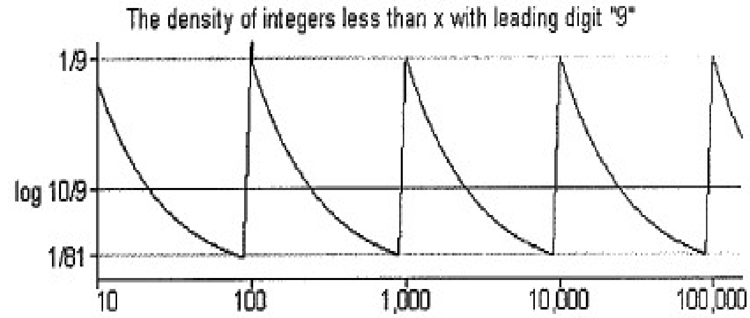
2.Why does it work? The magic of log growth and stopping rules --scale and base invariance.

Take an index that starts at 1000 and has 20% growth per year

1000	2074	3583	1 occurrence
1200	2449	4300	1 occurrence
1440	2986	5160	1 occurrence
1738			
4 occurrences of 1	3 occurrences of 2		

It keeps growing at 20% so you get 6192, 7430, 8916 and then NO DIGIT WITH 9, 10670... Alternatively, think of addresses on a street, which go from 1, 2, ...9. On street with 4 houses, you get addresses 1,2,3, 4. If you have a random stopping rule, always start with 1 and get 1 but may not get other digits. If you take the density of integers less than or equal to x you get a saw-toothed pattern.





To see the saw-tooth distribution

count numbers with digit 1 along the real line

	digits	ratio to all counted digits
1	1	1 of 1
2	1,2	1/2
3	1,2,3	1/3
….		
9	**1,...9**	**1/9**
10		2/10
11		3/11
….		
19		11/19 ≈ 5/9
20		11/20
…		
99		**11/99= 1/9**
100		12/100
199		111/199 ≈ 5/9
999		111/999 = 1/9

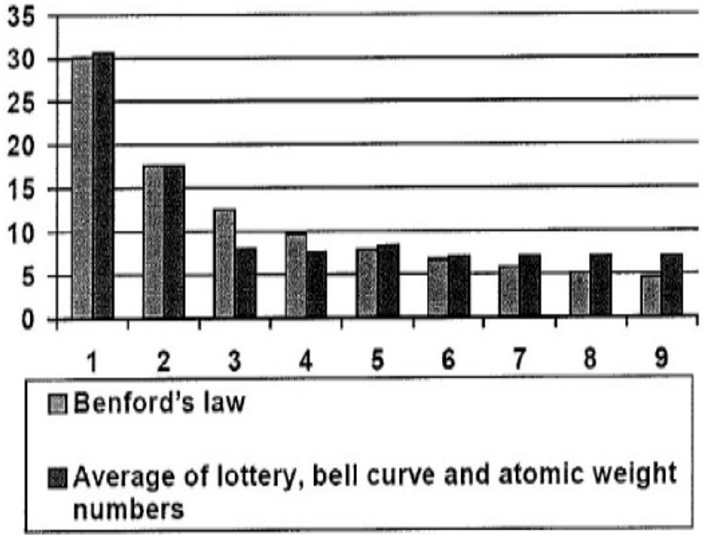
count numbers with LD 9 along the real line

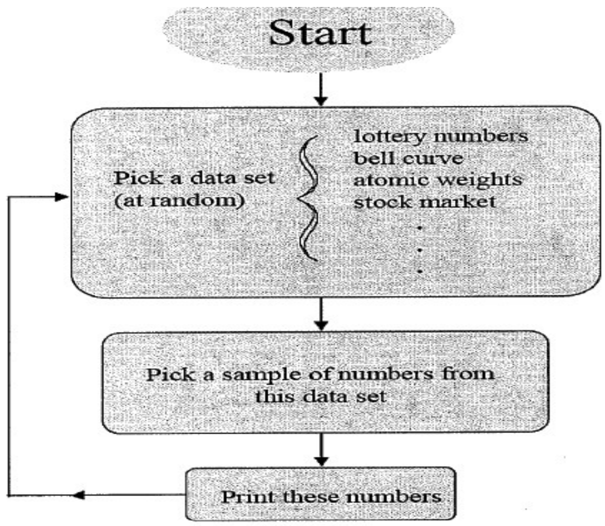
	digits	ratio to all digits
	1	0/1
	2	0/2
	1,...9	**1/9**
	10	1/10
	89	1/89 ≈ 1/81 ?? (.0112->.0123)
	90	2/90
	99	**11/99**
	...	
899		11/899 ≈ 1/81 (.0122)
999		111/999 = 1/9
	8999	111/8999≈ 1/81 (.0123)

Hill's theorem: if we repeatedly pick random entries from random distributions, the result tends towards the distribution of Benford's law --that combinations of distributions tend towards the distribution predicted by Benford''s law *even when the original distributions do not* [Hill1996]. This because you are picking across many scales and so this is basically a scale free distribution –> Benford distribution.

Hill example: you pick lottery numbers from newspaper – uniform; You pick numbers of weights of people – normal distribution; you pick number of automobile accidents – exponential??
Take the % of lead digits of each and average them and you will get Benford!

	First Digits								
	1	2	3	4	5	6	7	8	9
Benford's Law	30.1%	17.6	12.5	9.7	7.9	6.7	5.8	5.1	4.6
Lottery	11%	11	11	11	11	11	11	11	11
Bell Curve	40%	13	8	8	7	7	6	6	5
Atomic Wts.	41%	28	5	4	7	3	4	4	5





Log Distribution (Benford's Law)

The "random samples from random distributions" theorem says that if distributions are selected at random (in a neutral way), and samples are taken from each of these distributions, then the resulting data set will have digital frequencies approaching Benford's Law (see Figure 5).

Benford's law is legally admissible as evidence in criminal cases at the federal, state and local levels. Major accounting firms use computer-assisted audit tools to check data since deviations from Benford's Law should "cause an analyst to question the validity, accuracy, or the completeness of numbers". (Nigrini, Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection. Hoboken, NJ: Wiley, 2012)

ISACA (**Information Systems Audit and Control Association)** tells auditors: "if the audit objective is to detect fraud in the disbursements cycle, the IT auditor **could** use Benford's Law to measure the … leading digits in disbursements compared to the digits' probability. Not work well if sample is small (<100) or subject to special rules such as thresholds and cutoffs. For instance, if a bank's policy is to refer loans at or above US $50,000 to a loan committee, looking just below that approval threshold gives a loan officer the potential to discover loan frauds. " (Simkin, Mark G. "Using Spreadsheets and Benford's Law to Test Accounting Data," ISACA Journal, V1, 2010.

Tracking the London Interbank Offered Rate – Libor Rate -- average interest rate calculated by British Banker's Association from rates by major London banks. **(**www.escholarship.org/uc/item/2p33x7dk)-- Abrantes-Metz, Villas-Boas, and George " use Benford **second digit** reference distribution to track the daily Libor over ...2005-2008 (and find)... Why 2nd digit ? First digits do not span the nine digit space. Find libor rates depart significantly from Benford.... This raises potential concerns relative to the unbiased nature of the signals coming from the sixteen banks from which the Libor is computed and the usefulness of the Libor as a major economic indicator. integrity of prices." In fact, **Authorities discovered that banks were falsely reporting their rates to profit from trades, or to give impression that they were more creditworthy than they were. Libor underpins approximately $350 trillion in derivatives.**

Scientific fraud in 20 falsified anesthesia papers : (Anaesthesist. 2012 Jun;61(6):543-9) papers known to be falsified by an author were investigated for irregularities with respect to Benford's law using the $\chi(2)$-test and the Z-test. In an analysis of each paper 17 out of 20 studies differed significantly from the expected value for the first digit and 18 out of 20 studies varied significantly from the expected value of the second digit...

Fact and Fiction in EU-Governmental Economic Data German Economic Review Volume 12, Issue 3, pages 243–255, August 2011 To detect manipulations or fraud in accounting data, auditors have successfully used Benford's law as part of their fraud detection processes... In the European Union (EU), there is pressure to comply with the Stability and Growth Pact criteria. Therefore … overnments might try to make their economic situation seem better.,, we use a Benford test to investigate the quality of macroeconomic data relevant to the deficit criteria reported to Eurostat by EU member states,,,D**ata ...by Greece shows the greatest deviation from Benford's law among all euro states.**