Benford's Law

David Groce
Lyncean Group
March 23, 2005

What do these have in common?

- SAIC's 2004 Annual Report
- Bill Clinton's 1977 to 1992 Tax Returns
- Monte Carlo results from Bill Scott
- Compound Interest
- Fibonacci Numbers
- NIST (old NBS) Table of Physical Constants
- Populations of California & Colorado Counties

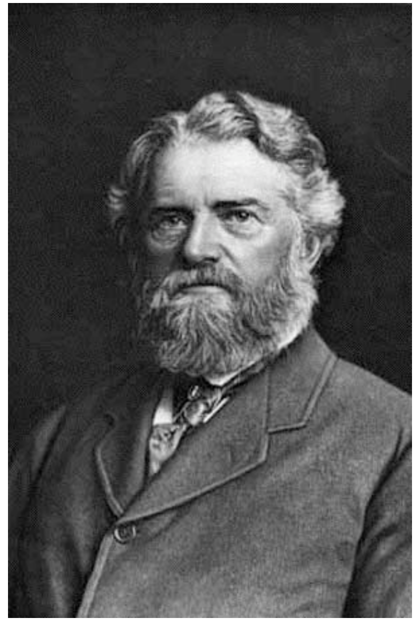
ALL Obey Benford's Law!!!

What is Benford’s Law?

Let’s go back to 1881

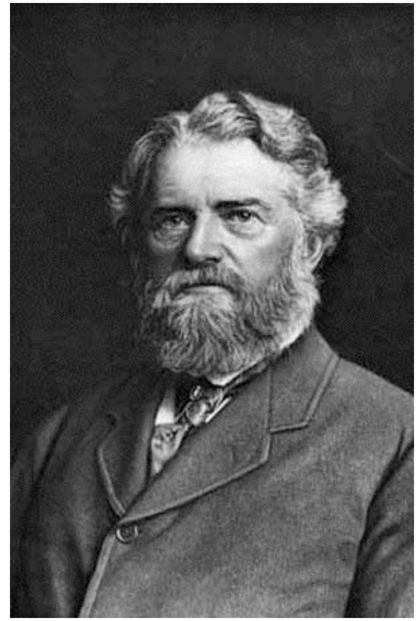
1881 Simon Newcomb

1835—1909



- "… had risen from rags to intellectual riches."
- Computer in National Almanac Office
- Studied Phrenology
- Director, National Almanac Office
- Professor Math & Astronomy, Johns Hopkins
- Most famous American astronomer of the times
- Measured speed of light with Michaelson
- Economist
  - Developed Quantity Theory of Money MV = PT
- President of most of the Scientific Societies
- First Bruce Medalist (1898)

Simon Newcomb

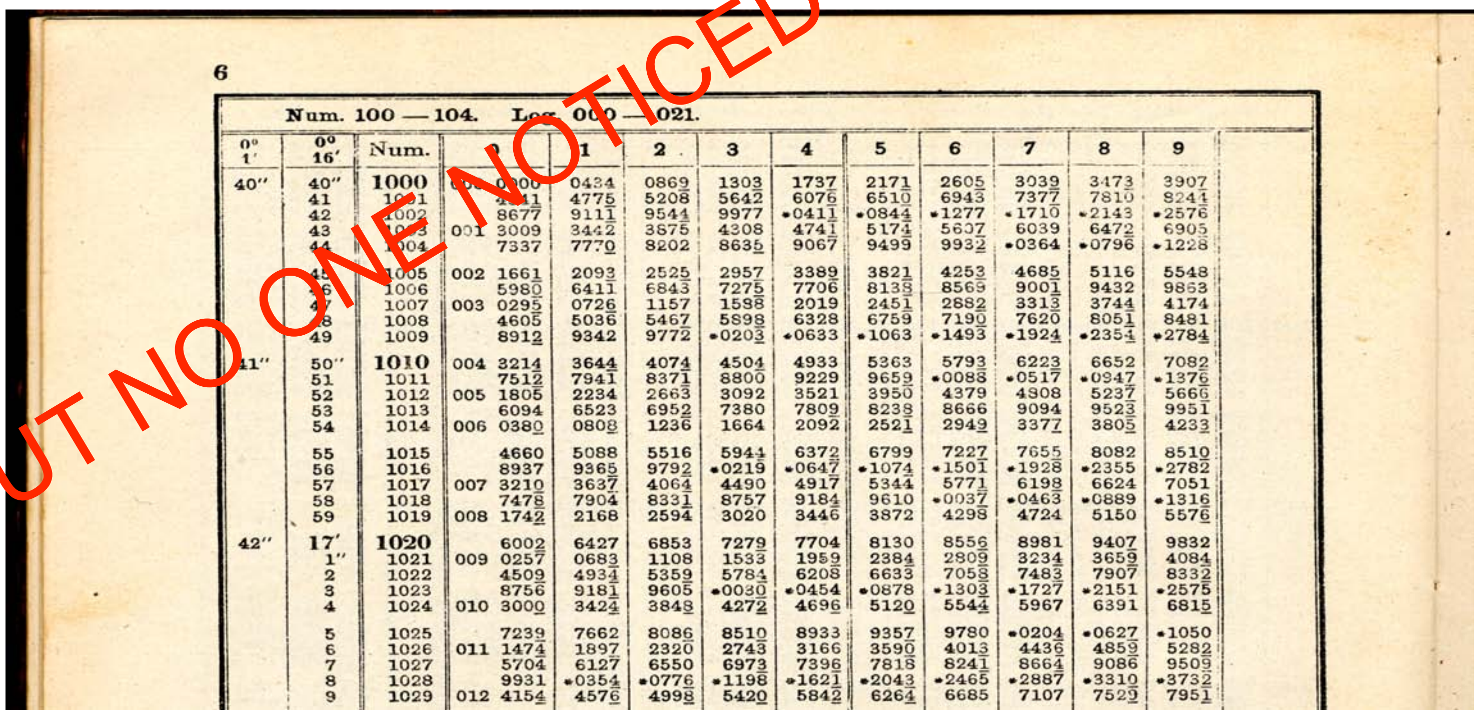


1) Noticed first pages of log tables dirtier than last pages
2) Concluded that numbers beginning with 1, 2 & 3s were looked up more often than those beginning with 7, 8 & 9s
3) Conjectured: probability of initial digits in a number followed

   $P(d) = \log_{10}(1 + 1/d)$; where: $d = 1,2,3,4,5,6,7,8,9$

   and $\sum P(d) = 1$
4) Published: Amer J Math 4, 1881 (pp 39-40)

BUT NO ONE NOTICED HIS PAPER!



6

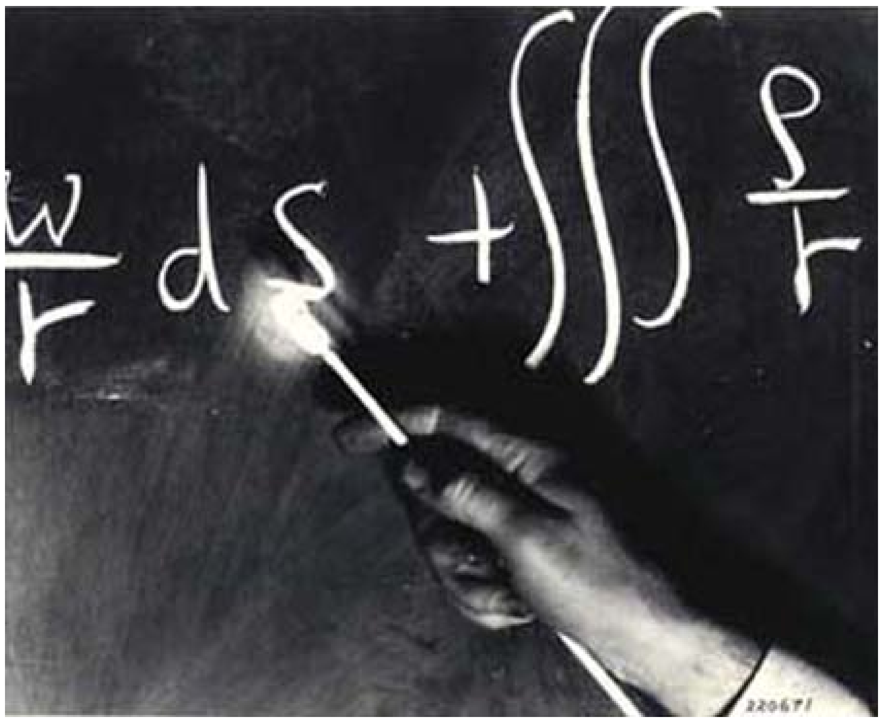
Num. 100 — 104. Log. 000 — 021.

0° 1′	0° 16′	Num.	0	1	2	3	4	5	6	7	8	9
40″	40″	1000	[illegible] 0000	0434	0869	1303	1737	2171	2605	3039	3473	3907
	41	[illegible]	[illegible]	4775	5208	5642	6076	6510	6943	7377	7810	8244
	42	1002	8677	9111	9544	9977	•0411	•0844	•1277	•1710	•2143	•2576
	43	[illegible]	001 3009	3442	3875	4308	4741	5174	5607	6039	6472	6905
	44	1004	7337	7770	8202	8635	9067	9499	9932	•0364	•0796	•1228
	45	1005	002 1661	2093	2525	2957	3389	3821	4253	4685	5116	5548
	46	1006	5980	6411	6843	7275	7706	8138	8569	9001	9432	9863
	47	1007	003 0295	0726	1157	1588	2019	2451	2882	3313	3744	4174
	48	1008	4605	5036	5467	5898	6328	6759	7190	7620	8051	8481
	49	1009	8912	9342	9772	•0203	•0633	•1063	•1493	•1924	•2354	•2784
41″	50″	1010	004 3214	3644	4074	4504	4933	5363	5793	6223	6652	7082
	51	1011	7512	7941	8371	8800	9229	9659	•0088	•0517	•0947	•1376
	52	1012	005 1805	2234	2663	3092	3521	3950	4379	4808	5237	5666
	53	1013	6094	6523	6952	7380	7809	8238	8666	9094	9523	9951
	54	1014	006 0380	0808	1236	1664	2092	2521	2949	3377	3805	4233
	55	1015	4660	5088	5516	5944	6372	6799	7227	7655	8082	8510
	56	1016	8937	9365	9792	•0219	•0647	•1074	•1501	•1928	•2355	•2782
	57	1017	007 3210	3637	4064	4490	4917	5344	5771	6198	6624	7051
	58	1018	7478	7904	8331	8757	9184	9610	•0037	•0463	•0889	•1316
	59	1019	008 1742	2168	2594	3020	3446	3872	4298	4724	5150	5576
42″	17′	1020	6002	6427	6853	7279	7704	8130	8556	8981	9407	9832
	1″	1021	009 0257	0683	1108	1533	1959	2384	2809	3234	3659	4084
	2	1022	4509	4934	5359	5784	6208	6633	7058	7483	7907	8332
	3	1023	8756	9181	9605	•0030	•0454	•0878	•1303	•1727	•2151	•2575
	4	1024	010 3000	3424	3848	4272	4696	5120	5544	5967	6391	6815
	5	1025	7239	7662	8086	8510	8933	9357	9780	•0204	•0627	•1050
	6	1026	011 1474	1897	2320	2743	3166	3590	4013	4436	4859	5282
	7	1027	5704	6127	6550	6973	7396	7818	8241	8664	9086	9509
	8	1028	9931	•0354	•0776	•1198	•1621	•2043	•2465	•2887	•3310	•3732
	9	1029	012 4154	4576	4998	5420	5842	6264	6685	7107	7529	7951

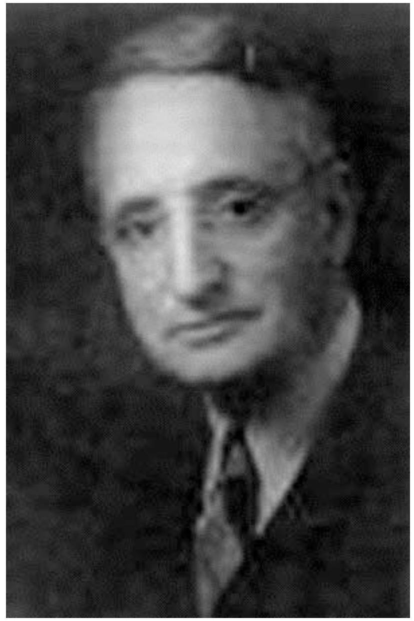
1938 Frank Benford

1883—1948

- Physicist at GE, Schenectady, NY
- Optics expert; high interest in math
- WWI: theory & design of searchlights
- Many patents



1938 Frank Benford



1) Noticed first pages of log tables dirtier than last pages
2) Concluded that numbers beginning with 1, 2 & 3s were looked up more often than those beginning with 7, 8 & 9s
3) Had not seen Newcomb's 1881 paper in Amer J of Math
4) Conjectured: probability of initial digits in a number followed

   $P(d) = \log_{10}(1 + 1/d)$; where: $d = 1,2,3,4,5,6,7,8,9$

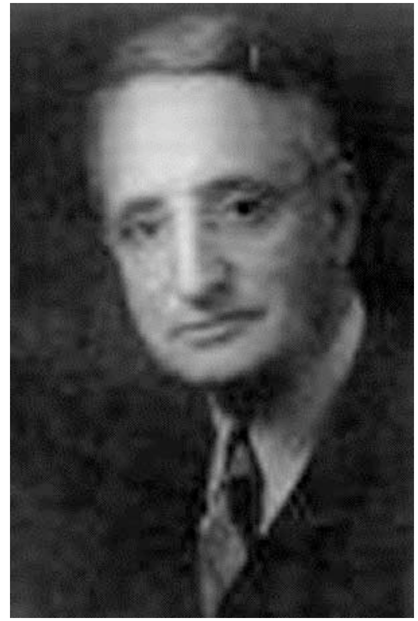
   and $\sum P(d) = 1$
5) But, he investigated 20 different sets of natural numbers involving over 20,000 samples (1929 to 1934 when GE on 1/2 time)

Benford’s Samples

Details not known

		First Digit									
Col.	Title	1	2	3	4	5	6	7	8	9	Samples
A	Rivers, Area	31.0	16.4	10.7	11.3	7.2	8.6	5.5	4.2	5.1	335
B	Population	33.9	20.4	14.2	8.1	7.2	6.2	4.1	3.7	2.2	3259
C	Constants	41.3	14.4	4.8	8.6	10.6	5.8	1.0	2.9	10.6	104
D	Newspapers	30.0	18.0	12.0	10.0	8.0	6.0	6.0	5.0	5.0	100
E	Specific Heat	24.0	18.4	16.2	14.6	10.6	4.1	3.2	4.8	4.1	1389
F	Pressure	29.6	18.3	12.8	9.8	8.3	6.4	5.7	4.4	4.7	703
G	H.P. Lost	30.0	18.4	11.9	10.8	8.1	7.0	5.1	5.1	3.6	690
H	Mol. Wgt.	26.7	25.2	15.4	10.8	6.7	5.1	4.1	2.8	3.2	1800
I	Drainage	27.1	23.9	13.8	12.6	8.2	5.0	5.0	2.5	1.9	159
J	Atomic Wgt.	47.2	18.7	5.5	4.4	6.6	4.4	3.3	4.4	5.5	91
K	n^{-1}, $\sqrt{n}$	25.7	20.3	9.7	6.8	6.6	6.8	7.2	8.0	8.9	5000
L	Design	26.8	14.8	14.3	7.5	8.3	8.4	7.0	7.3	5.6	560
M	*Reader's Digest*	33.4	18.5	12.4	7.5	7.1	6.5	5.5	4.9	4.2	308
N	Cost Data	32.4	18.8	10.1	10.1	9.8	5.5	4.7	5.5	3.1	741
O	X-Ray Volts	27.9	17.5	14.4	9.0	8.1	7.4	5.1	5.8	4.8	707
P	Am. League	32.7	17.6	12.6	9.8	7.4	6.4	4.9	5.6	3.0	1458
Q	Blackbody	31.0	17.3	14.1	8.7	6.6	7.0	5.2	4.7	5.4	1165
R	Addresses	28.9	19.2	12.6	8.8	8.5	6.4	5.6	5.0	5.0	342
S	n^1, $n^2 \cdots n!$	25.3	16.0	12.0	10.0	8.5	8.8	6.8	7.1	5.5	900
T	Death Rate	27.0	18.6	15.7	9.4	6.7	6.5	7.2	4.8	4.1	418
	Average	30.6	18.5	12.4	9.4	8.0	6.4	5.1	4.9	4.7	1011
	Probable Error	± 0.8	± 0.4	± 0.4	± 0.3	± 0.2	± 0.2	± 0.2	± 0.3		

1938 Frank Benford



1) Noticed first pages of log tables dirtier than last pages
2) Concluded that numbers beginning with 1, 2 & 3s were looked up more often than those beginning with 7, 8 & 9s
3) Had not seen Newcomb's 1881 paper in Amer J of Math
4) Conjectured: probability of initial digits in a number followed

   $P(d) = \log_{10}(1 + 1/d)$; where: $d = 1,2,3,4,5,6,7,8,9$

   and $\sum P(d) = 1$
5) But, he investigated 20 different sets of natural numbers involving over 20,000 samples (1929 to 1934 when GE on 1/2 time)

6) Published: Proc of the Amer Philosophical Soc, 78 no. 4, March, 1938, (pp. 551-571).
7) Paper was noticed (located just after the seminal paper on electron scattering by Bethe and Rose).
8) Lo, it was a beautiful "law" that was called Benford's Law.

Issue: What kind of numbers?

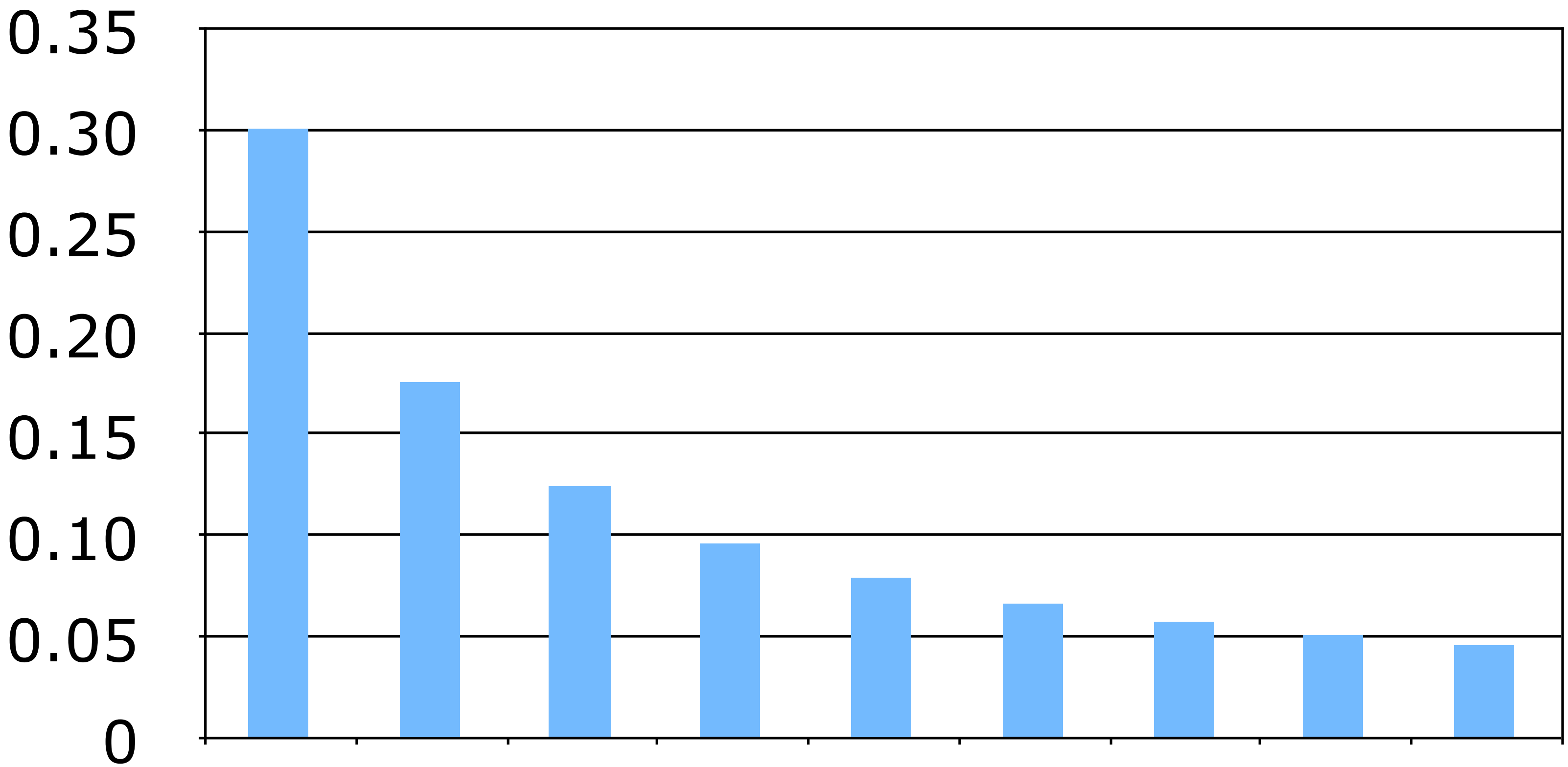
California

(58 counties, 155,959 sq. mi. land; pop. 34,501,130)

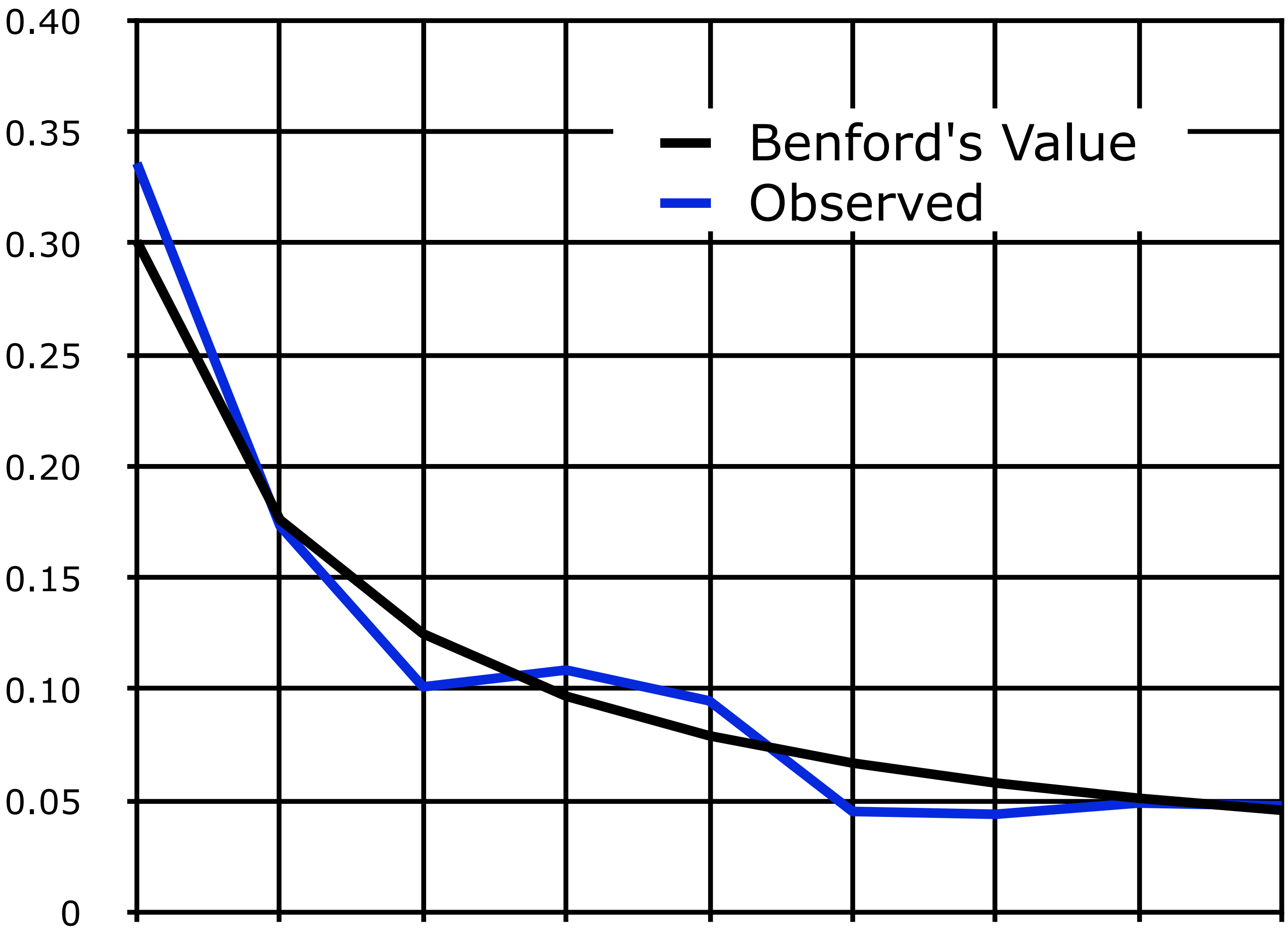
County	County seat or courthouse	2001 Pop.	2000 Pop.	Land area sq.mi.
Alameda	Oakland	1,458,420	1,443,741	738
Alpine	Markleeville	1,192	1,208	739
Amador	Jackson	36,269	35,100	593
Butte	Oroville	205,973	203,171	1,639
Calaveras	San Andreas	42,005	40,554	1,020
Colusa	Colusa	19,277	18,804	1,151
Contra Costa	Martinez	975,532	948,816	720
Del Norte	Crescent City	27,554	27,507	1,008
El Dorado	Placerville	162,586	156,299	1,711
Fresno	Fresno	815,734	799,407	5,963
Glenn	Willows	26,558	26,453	1,315
Humboldt	Eureka	126,468	126,518	3,572
mperial	El Centro	145,744	142,361	4,175
Inyo	Independence	17,944	17,945	10,203
Kern	Bakersfield	676,367	661,645	8,141
Kings	Hanford	132,119	129,461	1,391

Issue: What kind of Law?

A law is a mathematical statement which always holds true. Whereas "laws" in physics are generally experimental observations backed up by theoretical underpinning, laws in mathematics are generally theorems which can formally be proven true under the stated conditions. However, the term is also sometimes used in the sense of an empirical observation, e.g., Benford's law.
Wolfram MathMorld



Example: SAIC 2004 Report



Issue: How good is the fit?

A Quick Review of Statistics 1

Poisson Distribution (for enumeration, digits ≥ 0)
with large numbers approaches normal distr.

for a mean of m, $P(d) = m^d e^{-m} / d!$
standard deviation = $\sqrt{m}$

A Quick Review of Statistics 2

The statistic, Chi Squared $\chi 2$, is summed over i by:

$$\chi 2 = \sum (F_i - f_i)^2/F_i$$

where: F_i are the theoretical values

f_i are the observed values

i are the observation categories

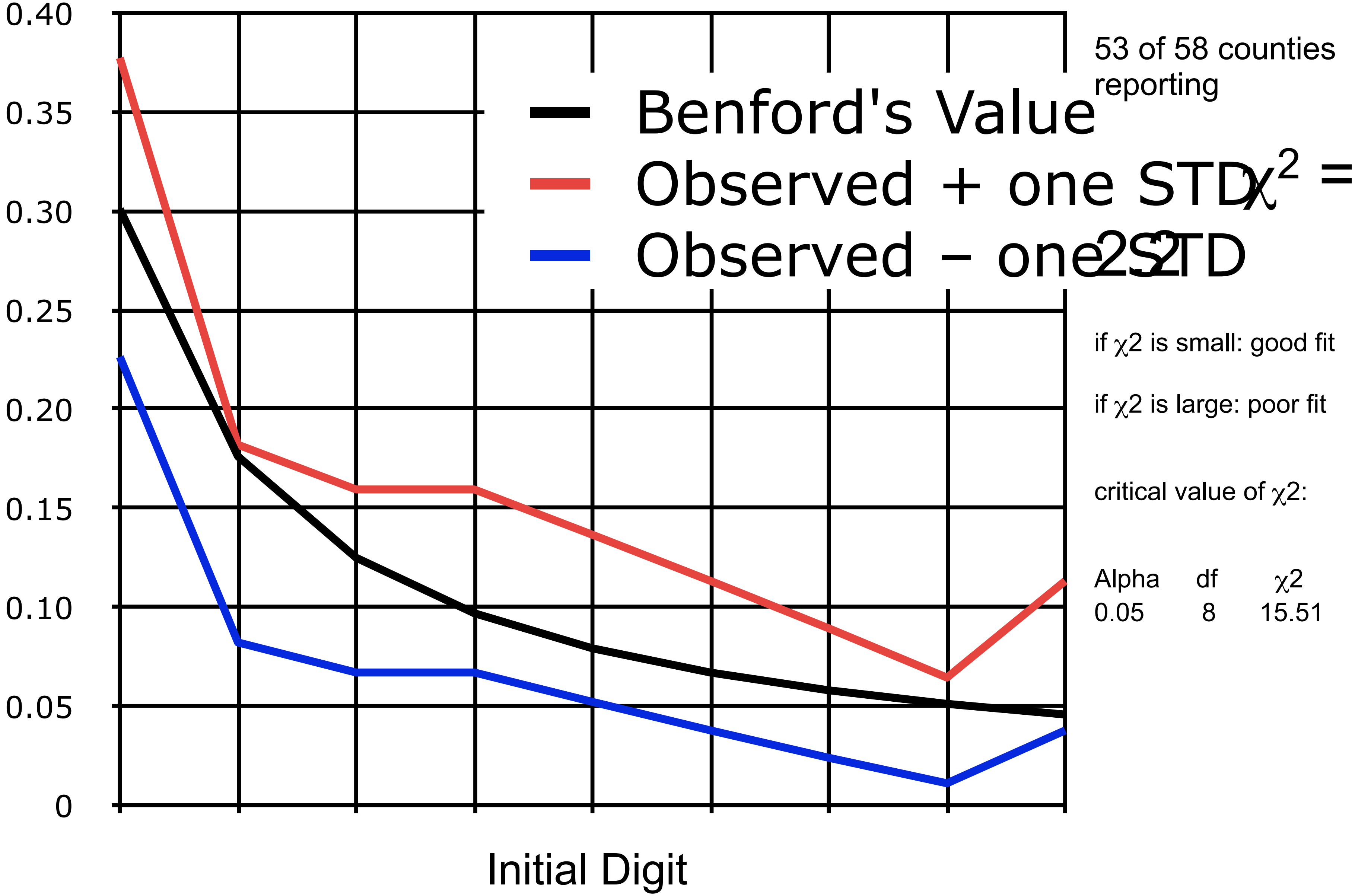
if $\chi 2$ is small, fit (hypothesis) is good

if $\chi 2$ is large, fit (hypothesis) is poor

critical value of $\chi 2$:

	significance	df	$\chi 2$
"statistically significant"	0.05	8	15.51

Davis Recall Signatures



53 of 58 counties reporting

$\chi^2 = 22$

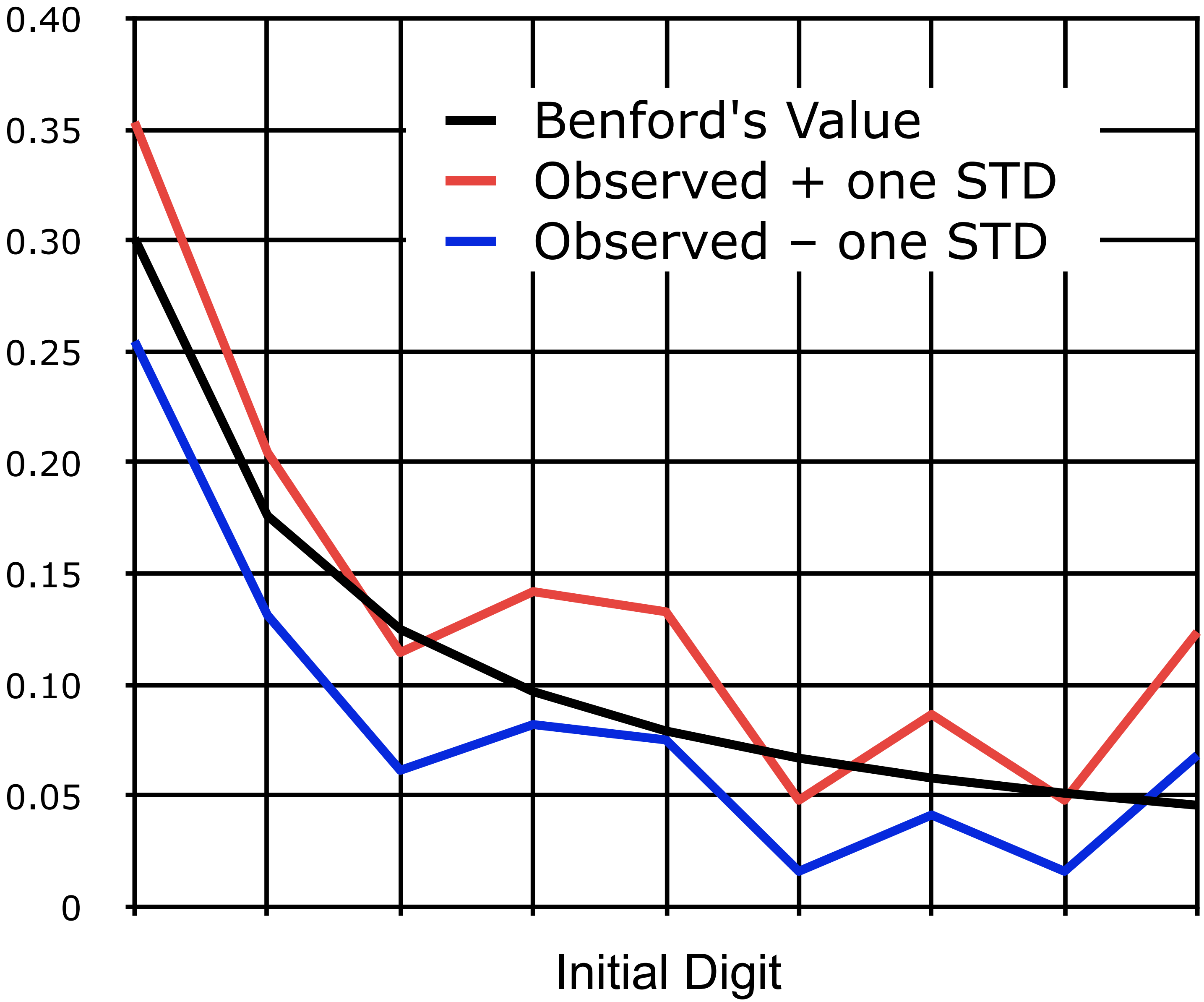
if $\chi2$ is small: good fit

if $\chi2$ is large: poor fit

critical value of $\chi2$:

Alpha	df	$\chi2$
0.05	8	15.51

2000 Population CA&CO counties



Benford:

$\chi^2 = 13$

Random:

$\chi^2 = 63$

Exponential:

$\chi^2 = 20$

if χ2 is small: good fit

if χ2 is large: poor fit

critical value of χ2:

Alpha	df	χ2
0.05	8	15.51

Fibonacci Numbers

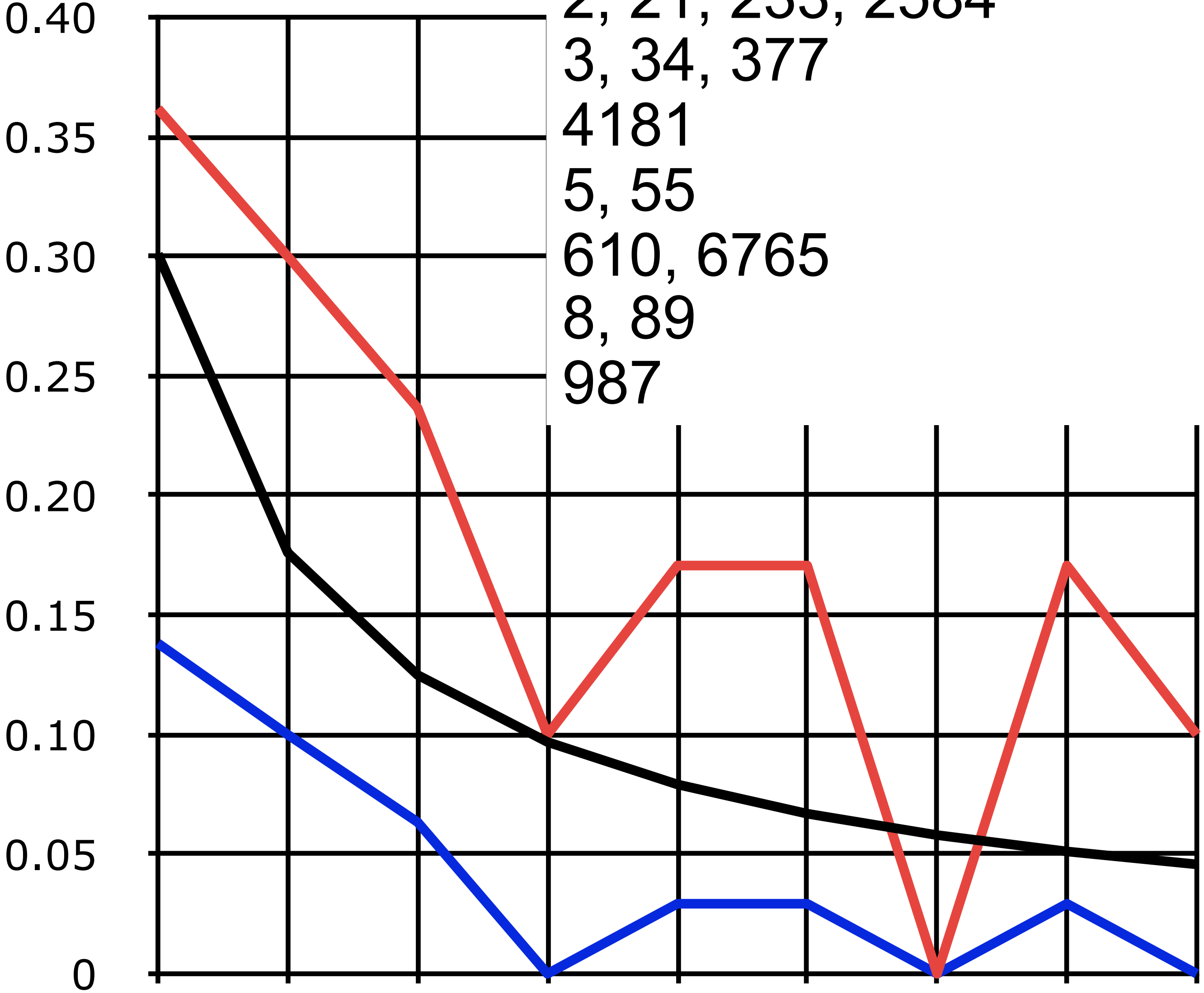
$F_n = F_{n-2} + F_{n-1}$

(1, 1 start): 1,1,2,3,5,8,13,21,34,55, ...
(2, 1 start)*: 2,1,3,4,7,11,18,29,47, ...

Limit F_n/F_{n-1} = Phi = $(1 + \sqrt{5})/2$
= 1.61803 (Golden Ratio)

First 20 Fibonacci Numbers

1, 1, 13, 144, 1397
2, 21, 233, 2584
3, 34, 377
4181
5, 55
610, 6765
8, 89
987



Benford:

$\chi^2 = 3.3$

Random:

$\chi^2 = 8.8$

Exponential:

$\chi^2 =$ --

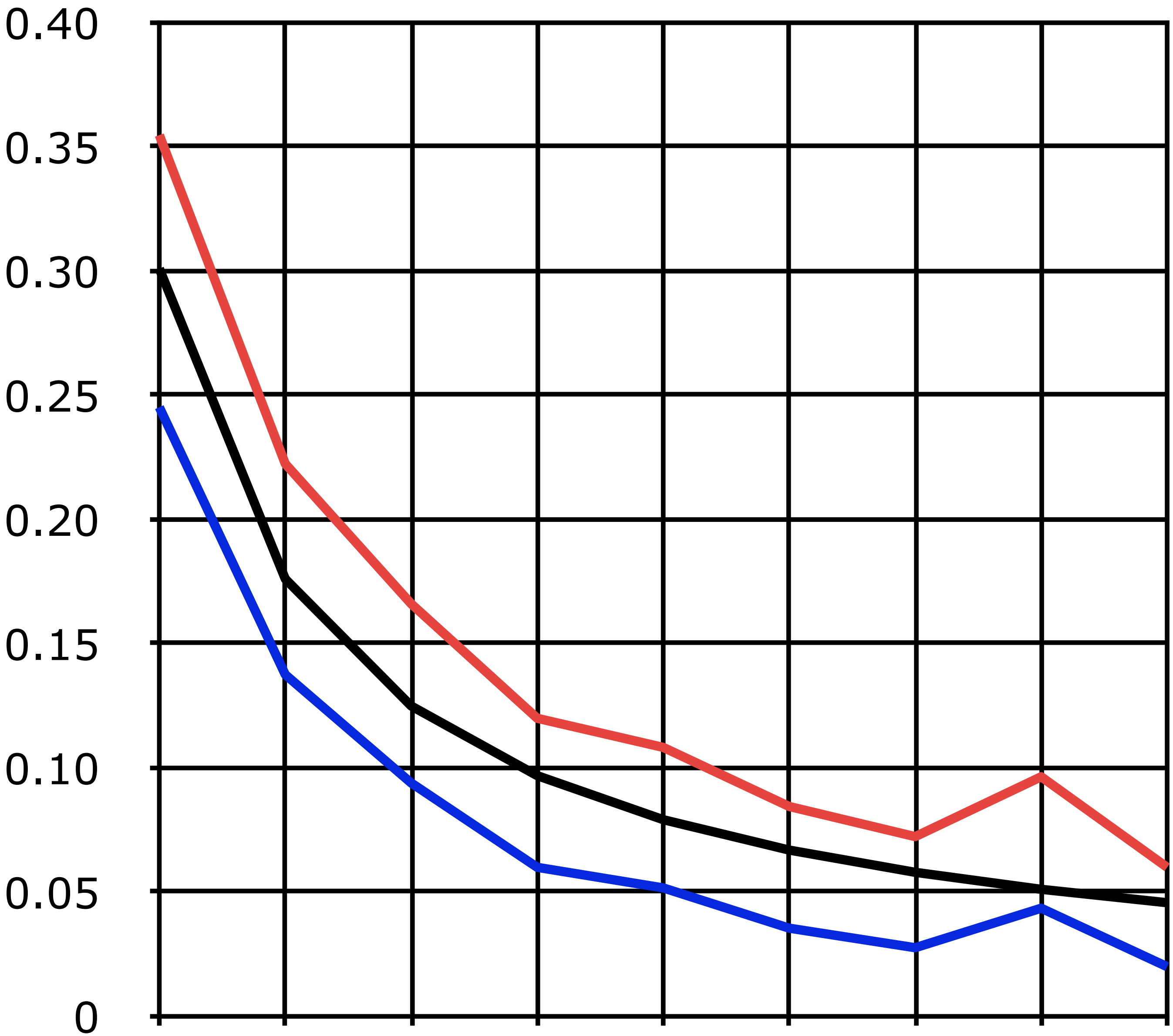
if $\chi 2$ is small: good fit

if $\chi 2$ is large: poor fit

critical value of $\chi 2$:

Alpha	df	$\chi 2$
0.05	8	15.51

First 100 Fibonacci Numbers



Benford:

$\chi^2 = 1.0$

Random:

$\chi^2 = 50$

Exponential:

$\chi^2 = 4.8$

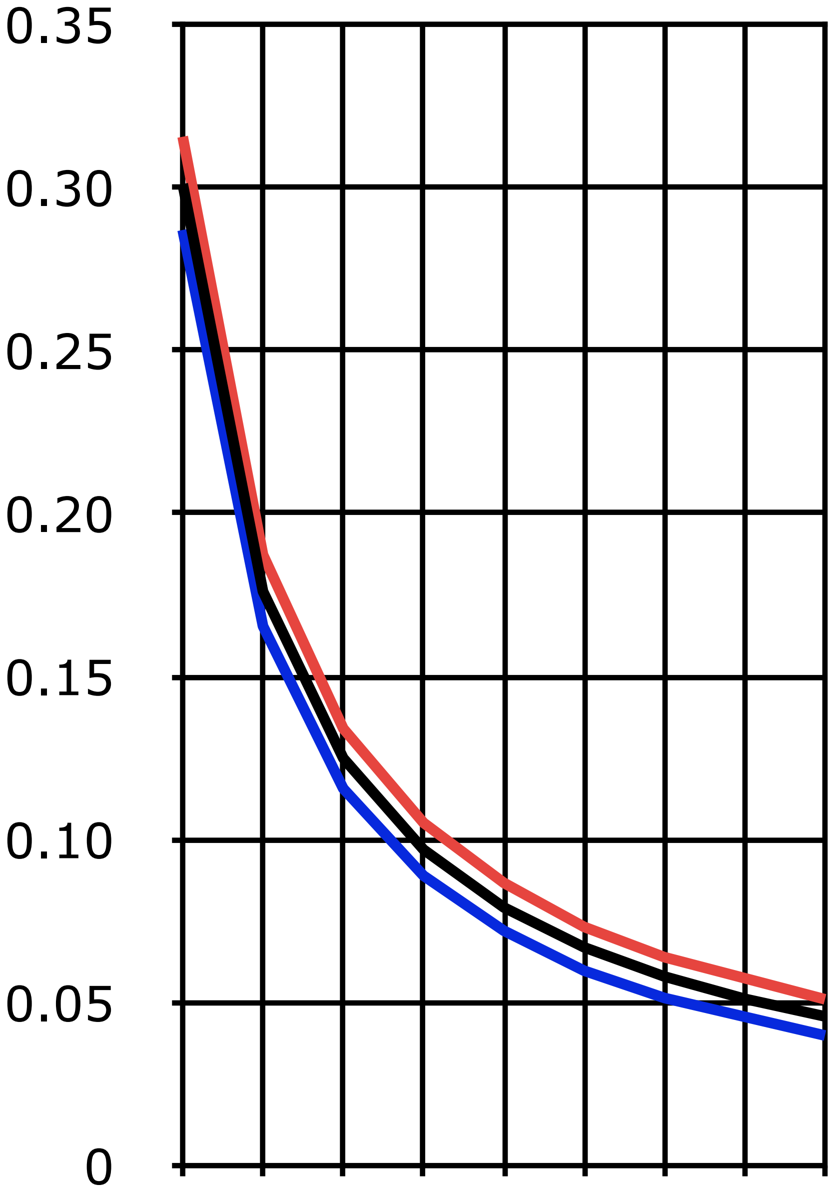
if $\chi 2$ is small: good fit

if $\chi 2$ is large: poor fit

critical value of $\chi 2$:

Alpha	df	$\chi 2$
0.05	8	15.51

First 1,474* Fibonacci Numbers



Benford:

$\chi^2 = .02$

Random:

$\chi^2 = 723$

Exponential:

$\chi^2 = 53$

2,000,000 #s:

$\chi^2 = 2\text{x}10^{-4}$

Lucas (2,1):

$\chi^2 = .06$

"Groce #s" (random):

$\chi^2 = .05$

Divide each by π:

$\chi^2 = .04$

Multiply each by e:

$\chi^2 = .06$

Multiply by random #s:

$\chi^2 = 5.2$

Multiply twice by rand #s:

$\chi^2 = 3.9$

Each term (739) squared:

$\chi^2 = .10$

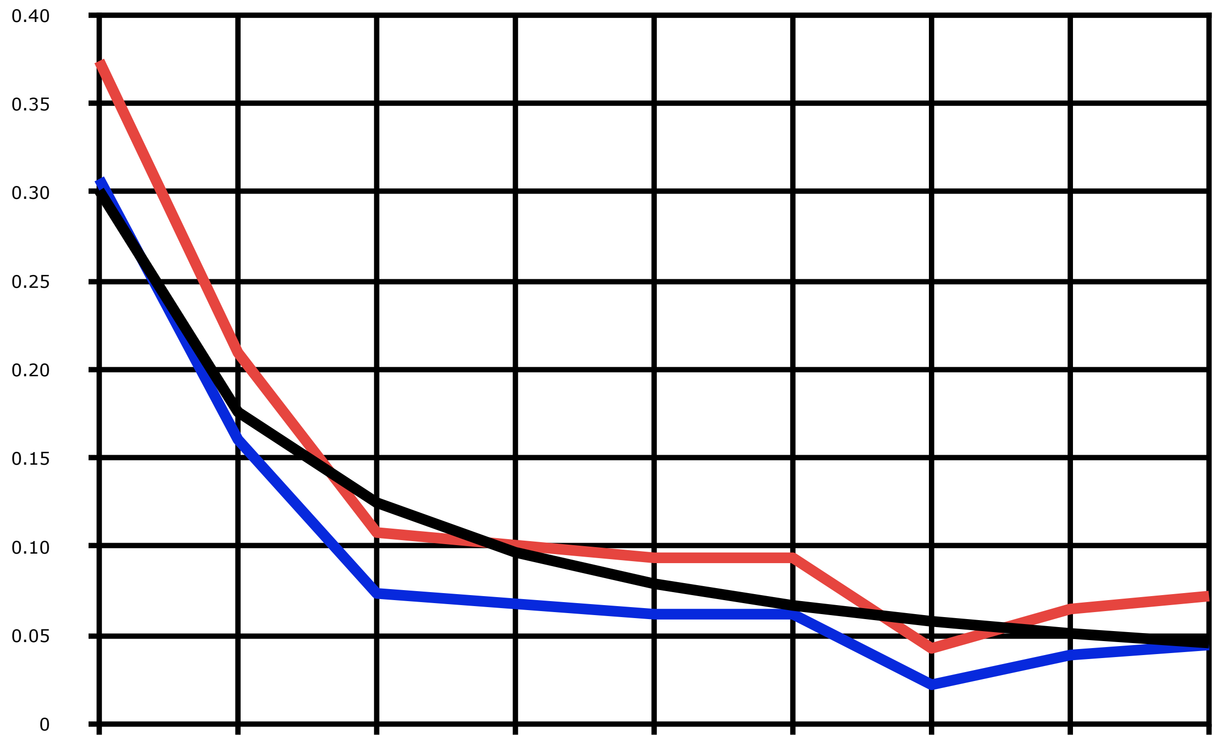
Observations

Any Benford distribution when multiplied (or divided) by any real number gives another Benford distribution.

A Benford distribution with dimensions (units) is scale invariant (miles to km, US$ to any currency, etc.).

A Benford distribution is numerical base invariant (decimal to octal, hexadecimal, etc., but not binary.

309 NIST Physical Constants



Benford:
$\chi^2 = 10$

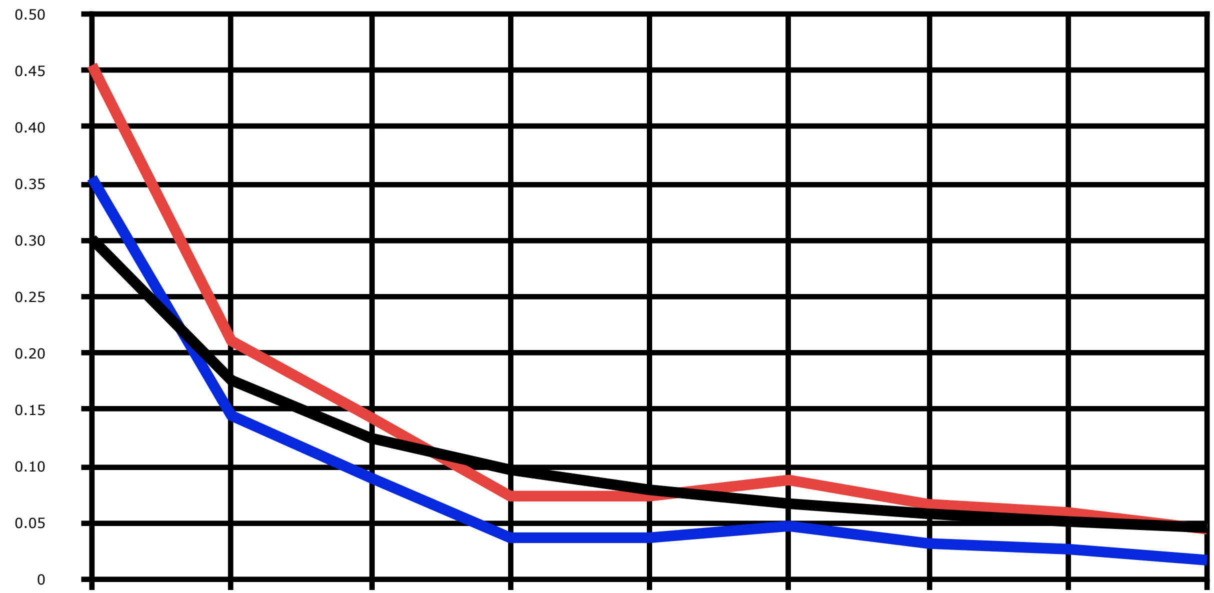
Random:
$\chi^2 = 205$

Exponential:
$\chi^2 = 35$

Reciprocals:
$\chi^2 = 8.2$

speed of light in vacuum	299792458	m s^-1
magn. constant	12.566 370 614e-7	N A^-2
electric constant	8.854187817e-12	F m^-1
characteristic impedance of vacuum	376.730313461	ohm
Newtonian constant of gravitation	6.6742e-11	m^3 kg^-1 s^-2
Newtonian constant of gravitation over h-bar c	6.7087e-39	(GeV/c^2)^-2
Planck constant	6.626 0693e-34	J s
Planck constant in eV s	4.13566743	eV s
Planck constant over 2 pi times c in MeV fm	197.326 968	MeV fm

164 Mathematical Constants



Benford:
$\chi^2 = 11$

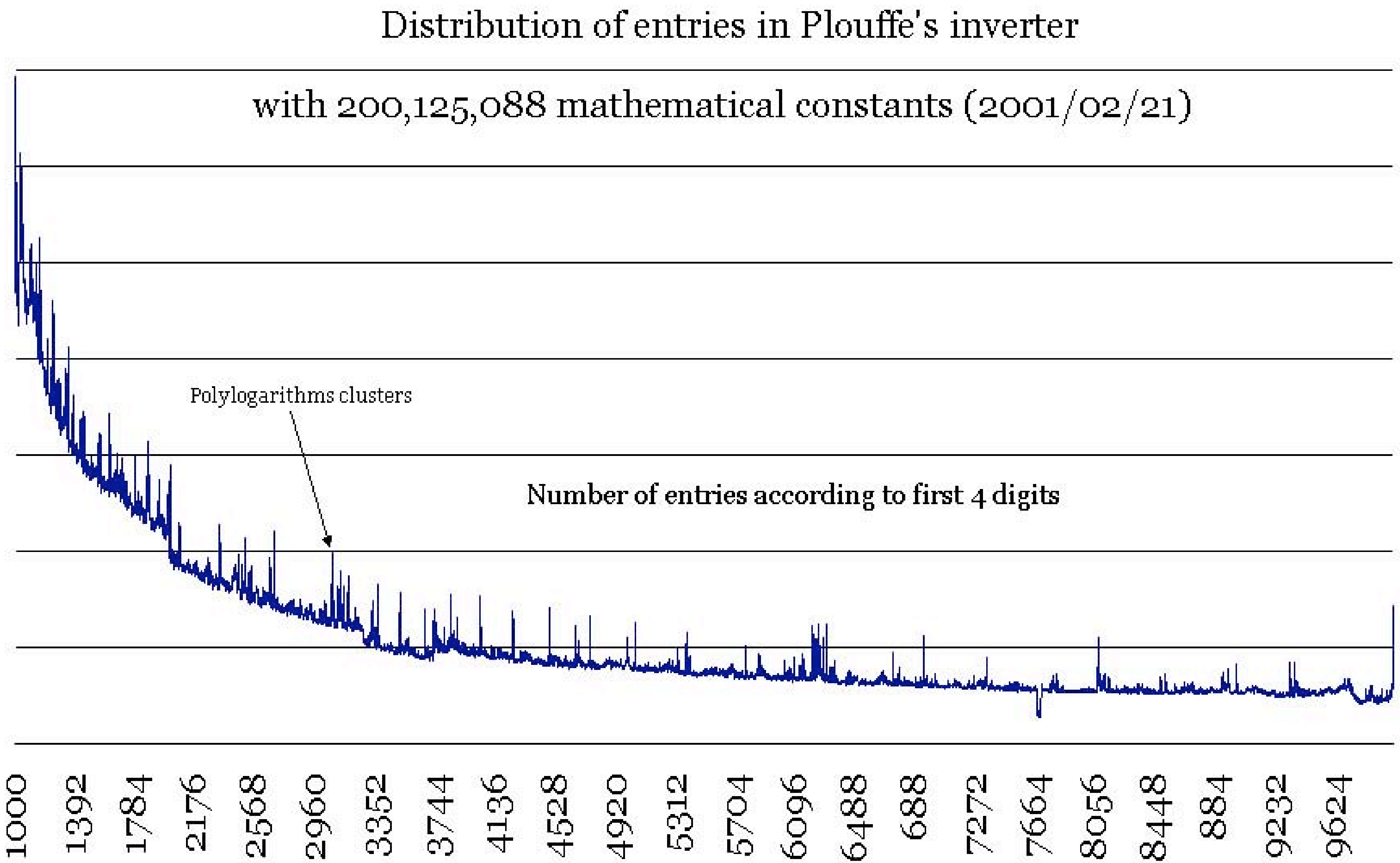
Random:
$\chi^2 = 167$

Exponential:
$\chi^2 = 22$

Base 8:
$\chi^2 = 9.1$

2.302566137	a Flajolet constant
1.29621066	A Salem number of degree 38
0.847213085	A.G.M. of (1,sqrt(2)/2)
0.258819404	-Ai'(0) or 3**(-1/3)/Gamma(1/3)
0.355028054	Ai(0) or 3**(-2/3)/Gamma(2/3)
0.463647609	arctan(1/2)
0.147583618	arctan(1/2)/Pi
0.373955814	Artin constant
1.456074949	Backhouse constant
0.280169499	Bernstein constant
0.697774658	BesselI(1,2)/BesselI(0,2)

Entries in Plouffe's Inverter



http://pi.lacim.uqam.ca/eng/

Why does the Benford Distribution work?

Let’s go back to 1961

1961 Roger Pinkham

Stevens Institute of Technology

- Statistics and Probability
- Numerical Methods
- Vision Research

Prior papers dismissed Bedford's Law

Pinkham proposed that if there is such a law, it must be scale invariant.

He proved Benford's Law is scale invariant under multiplication.

And the Benford Distribution is the only invariant distribution.

Good (1965) suggestion re: English random number generation

Multi-Digit Benford's Law

Newcomb & Benford both proposed initial multi-digit probabilities:

$P(d_1d_2) = \log_{10}(1 + 1/[d_1d_2])$

d_1d_2 = 10 thru 99

$\sum P(d_1d_2) = 1.00$

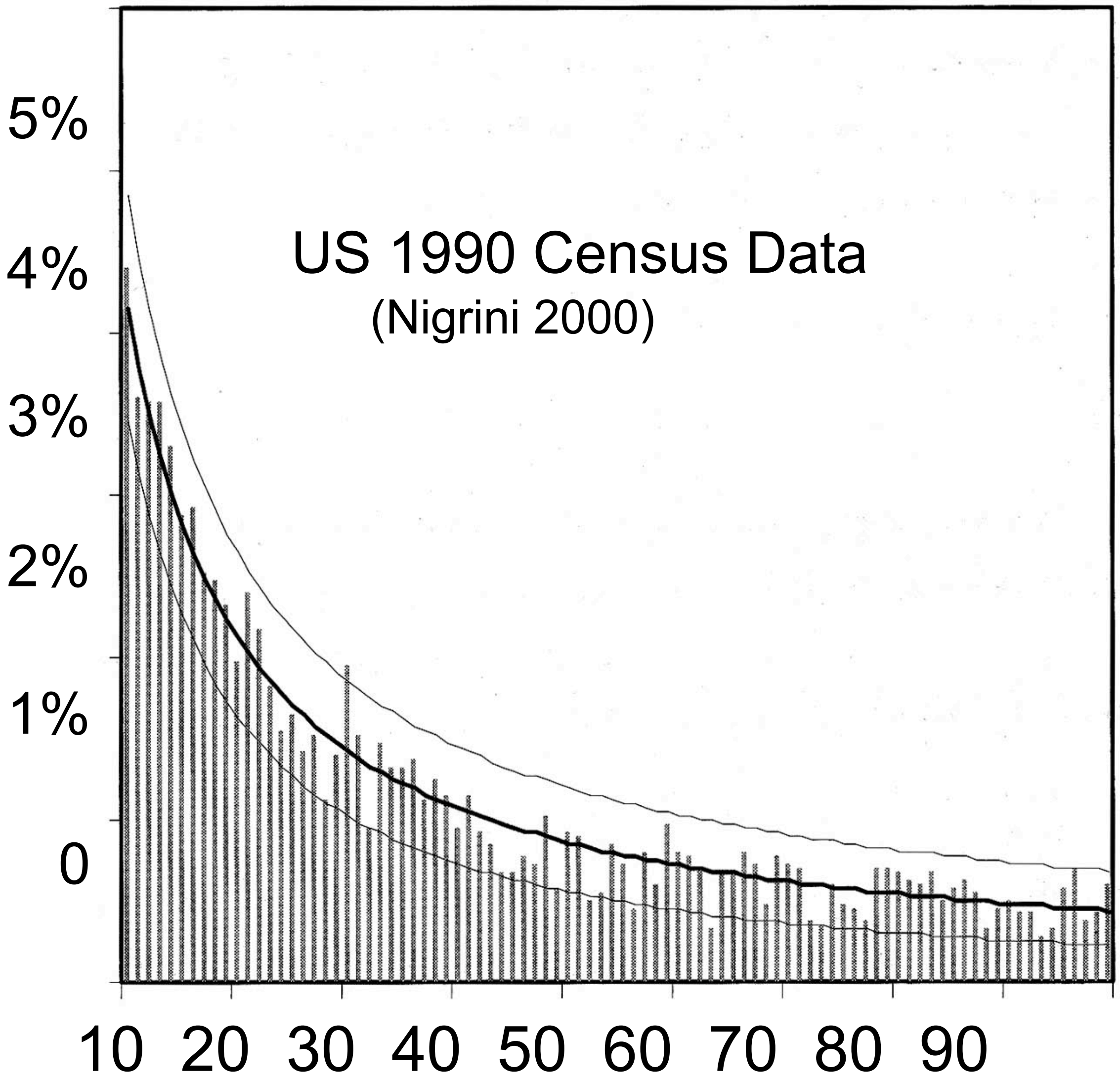
P(10) = 4.1%

P(25) = 1.7%

P(50) = 0.86%

P(75) = 0.58%

P(99) = 0.44%



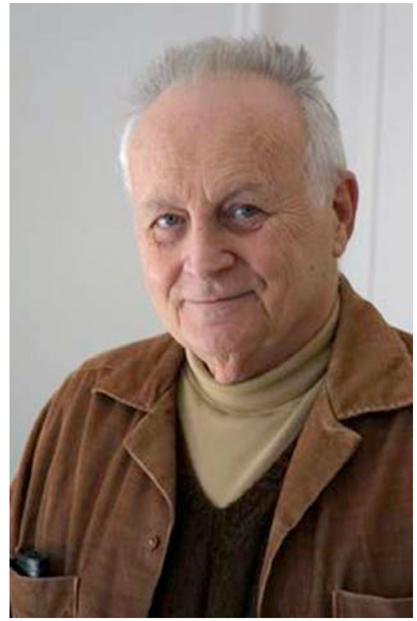
Middle Digit Probabilities

Digit	First	Second	Third	Fourth
0	--	0.120	0.102	0.100
1	0.301	0.114	0.101	0.100
2	0.176	0.109	0.101	0.100
3	0.125	0.104	0.101	0.100
4	0.097	0.100	0.100	0.100
5	0.079	0.097	0.100	0.100
6	0.067	0.093	0.099	0.100
7	0.058	0.090	0.099	0.100
8	0.051	0.088	0.099	0.100
9	0.046	0.085	0.098	0.100
totals:	**1.000**	**1.000**	**1.000**	**1.000**

Nigrini 1996

1969 Ralph Raimi

1924 — University of Rochester



- Amateur wine maker
- Interest in academic dishonesty
- Math and Physics
- Several papers on initial digits
- Scientific American, Dec. 1969, pp. 109-120
  - Non-mathematical review of Benford's
  - Intuitive explanations of distribution

1992 Mark Nigrini

1924 — Southern Methodist Univ.

WAS ONLY GUY TO MAKE $$$$$ FROM BENFORD'S LAW

- Ph.D. thesis (1992):

  Detection of Income Tax Invasion by Benford
- Assistant Professor of Accounting
- Analyzed Bill Clinton's 1977-92 tax returns
- Proprietary (big $$$) fraud codes for accountants

  Single and multiple digits, middle digits, round-off

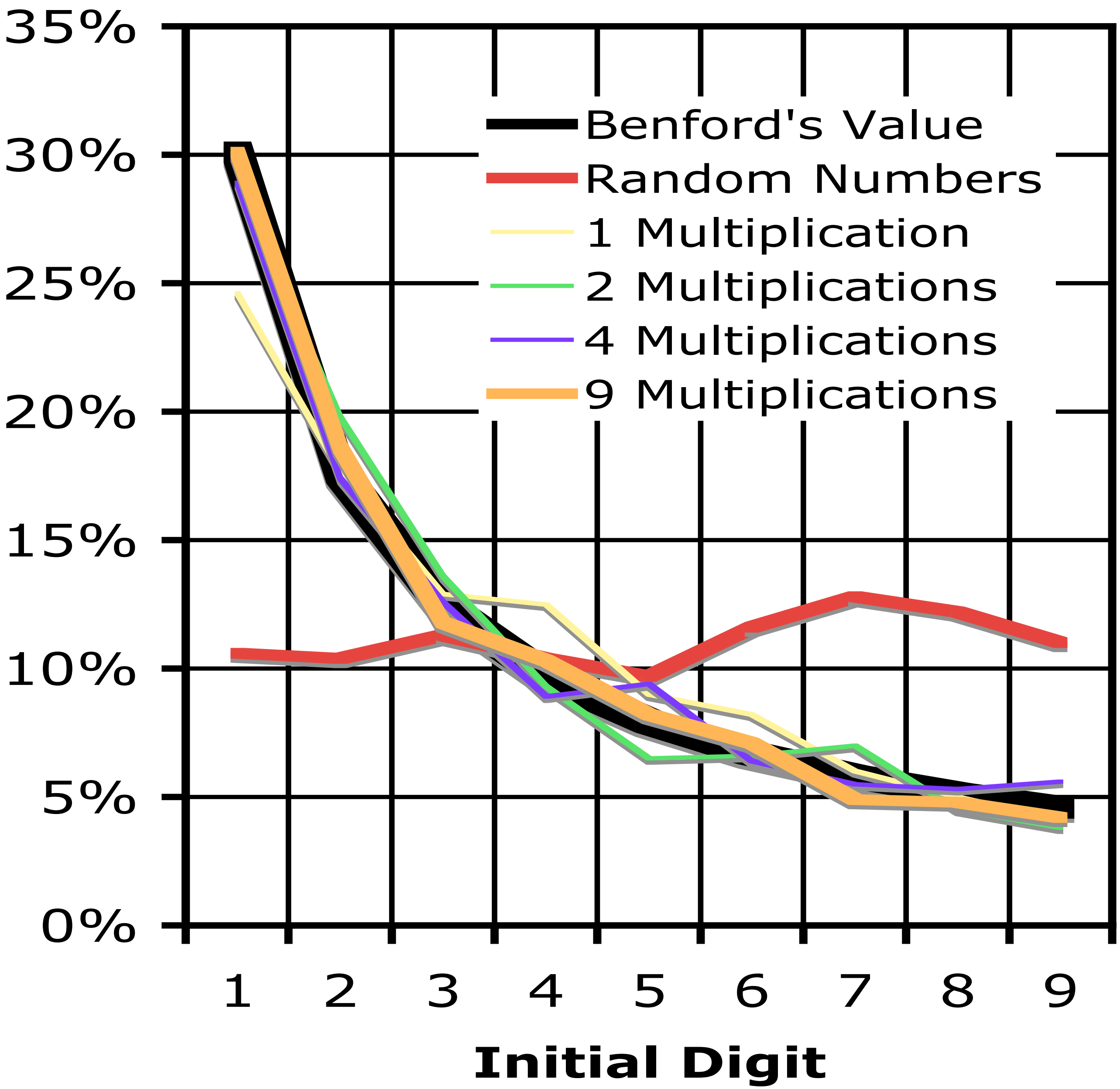
  Seminars ($150+)
- Book (2000; 32¢ per page):

  Digital Analysis Using Benford's Law

  Many cases of accounting fraud studies

Multiplication of Random Numbers

Two random #s multiplied give a negative log distribution with a “soft” infinity at zero and zero at one. (Scott, 2005)



Random #s:

$\chi^2 = 470$

1 Multiplication:

$\chi^2 = 24$

2 Multiplications:

$\chi^2 = 11$

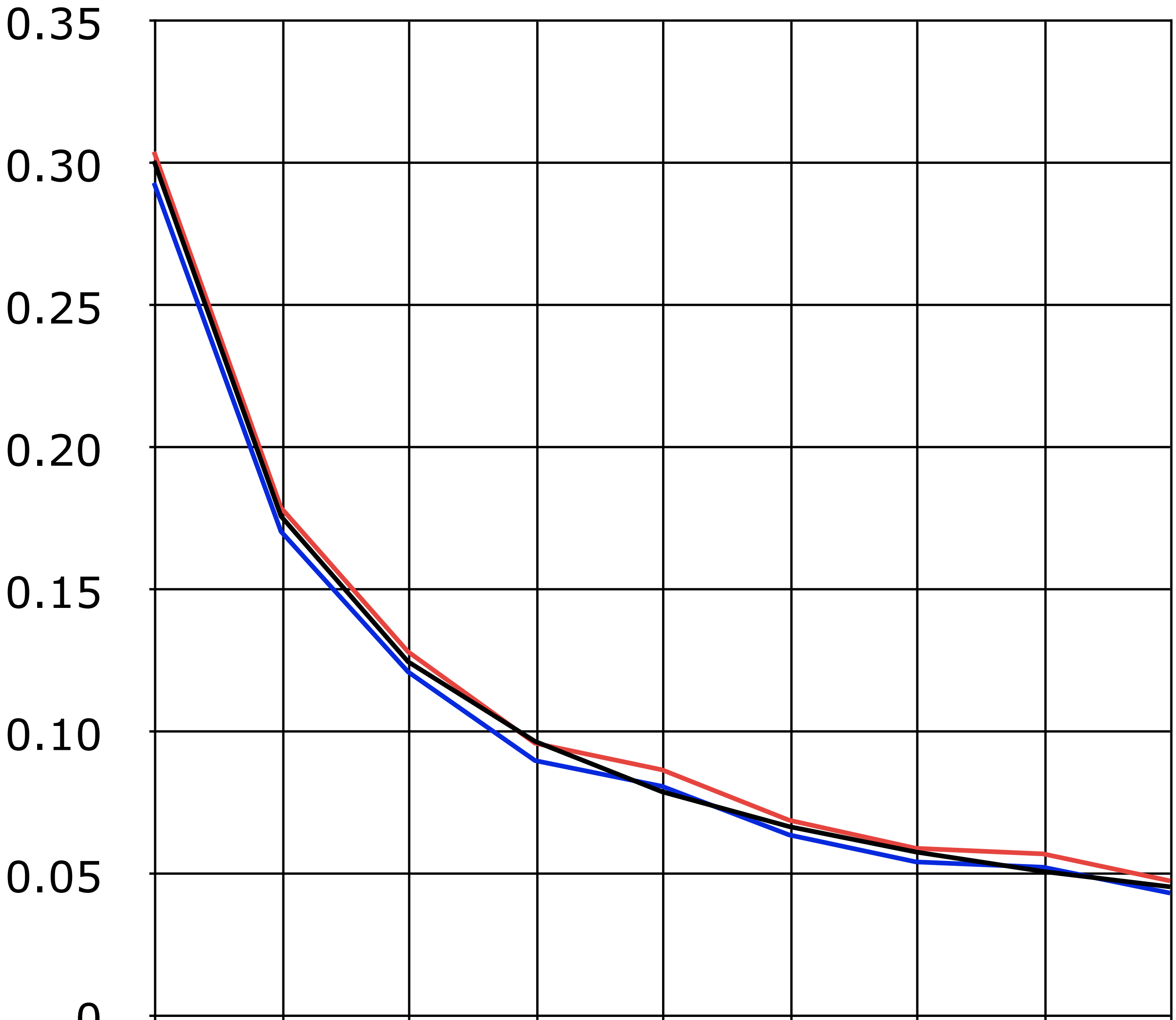
4 Multiplications:

$\chi^2 = 6.6$

9 Multiplications:

$\chi^2 = 3.7$

10,000 Random Exponentials



Benford:

$\chi^2 = 7.7$

Random:

$\chi^2 = 4760$

Exponential:

$\chi^2 = 370$

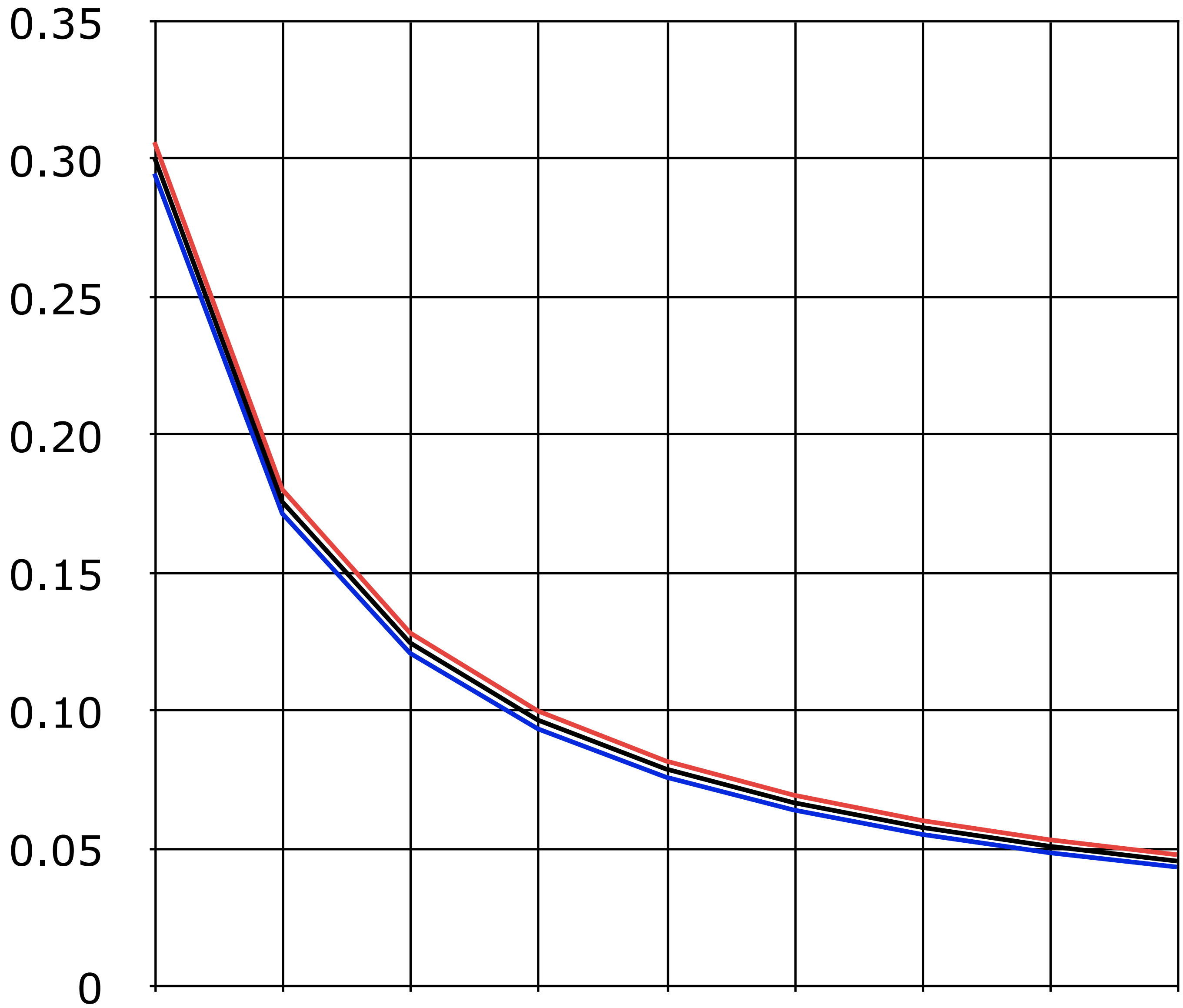
if χ2 is small: good fit

if χ2 is large: poor fit

critical value of χ2:

Alpha	df	χ2
0.05	8	15.51

9,210 Uniform Exponentials



Benford:

χ^2 = .004

Random:

χ^2 = 4500

Exponential:

χ^2 = 328

if χ2 is small: good fit

if χ2 is large: poor fit

critical value of χ2:

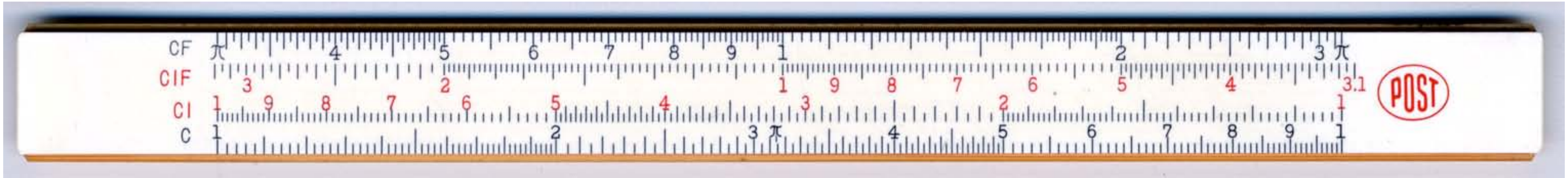
Alpha	df	χ2
0.05	8	15.51

Dow Jones Intuitive Explanation

- DJIA took 1983 to 1999 to go from 1,000 to 10,000
- Implied compounded appreciation rate of near 15%

Dow Jones Avg.	Theor. quarters	Actual quarters
1k to 2k	19	16
2k to 3k	11	20
3k to 4k	8	12
4k to 5k	6	4
5k to 6k	5	3
6k to 7k	4.4	2
7k to 8k	3.8	2.5
8k to 9k	3.2	2.5
9k to 10k	2.6	1
Chi-Squared	0.05	14

Slide Rule Intuitive Explanation



C-scale: $\log_{10}$ from 1 to 10

If select random points on C-scale and consider initial digits:

$P(1)$ = (length of log scale from 1 to 2)/(length from 1 to 10)

$= (\log 2 - \log 1)/(\log 10 - \log 1)$

$= \log 2 = 0.301$

$P(d) = [\log(d+1) - \log(d)]/ (\log 10 - \log 1)$

$= \log(1 + 1/d)$ Benford's Law

1996 Ted Hill

Georgia Institute of Technology

- Math Professor
- Stochastics
- Probability theory
- Fair division
- Lottery strategies

- Hill's Theorem:
  1) Showed that random sampling from random distributions will converge on the Benford distr.
  2) Benford is the distribution of all distributions

  American Scientist, 1998, Vol. 13, No. 6, pp. 358-363

Why do some distributions conform to Benford's Law but others do not?

- Feller (1966) proves that the empirical distribution of any first digit in observational data follows Benford's Law.
- Hill's Theorem (1998): Benford is the distribution of all distributions. Real data can be a complex mix of many distributions.
- Boyle (1994) proved that:
  - The Benford distr. is the limiting distribution when random variables are repeatedly multiplied, divided, or raised to integer powers.
  - Once achieved, the Benford distribution persists under all further multiplication, division, and raising to integer powers.
- Nigrini (2000): "We should never assume, dogmatically, that the Law *must* apply."
- Limitations
  - Numbers with a natural or social origin, with dimensions, from calculations
  - No built-in min or max
  - Not assigned numbers (SSNs, car licenses, lottery numbers)
  - Numbers should not unduly duplicated
  - Four or more digits?
  - Sets should have 100+ terms (digit-9 < 5%)

Practical Applications 1

- S. S. Goudsmit (electron spin) in 1978:

  … To a physicist Simon Newcomb's explanation of the first-digit phenomenon is more than sufficient "for all practical purposes." Of course here the expression "for all practical purposes" has no meaning. There are no practical purposes, unless you consider betting for money on the first digit frequencies with gullible colleagues a practical use.

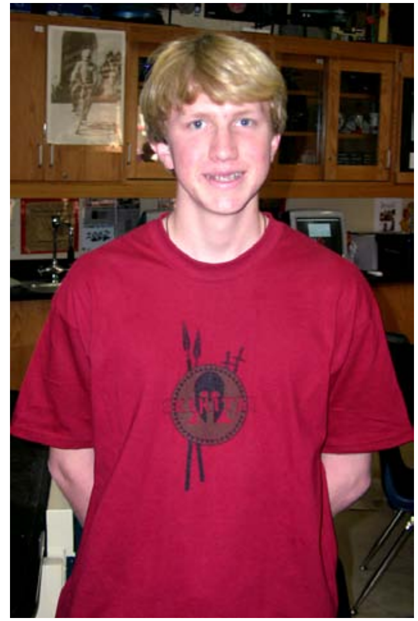
- Ian Stewart (1993) writes of betting on first digits at a trade fair in England.

Practical Applications 2

- Errors in data sets.
  - NYSE daily volumes of buy-sell transactions (Benford distr.)
  - NASDAQ daily volume of each buy and each sell transaction (Benford distr.)
- Fraud detection: accounting, pay roll, tax returns (Nigrini, 2000)
  - Reported interest received: excess lower first digits.
  - Reported interest paid: excess higher first digits.
- Prices, sales figures
- Forecasts from data (Varian, 1972)
- Computer round off errors (Tsao, 1974)
- Failure rates and MTTF tables (Becker, 1982)
- Numerical structure of the Quran.
- Mathematician Peter Schatte at the Bergakademie Technical University, Freiberg, has come up with rules that optimize computer data storage, by allocating disk space according to the proportions dictated by Benford's law.

2005 Alex Trahan

Sophomore, La Jolla High School



- Science team member of 50±
- Freshman science award
- Academic varsity letter
- Academic League
- Benford Science Fair math project

  "When does Randomness End"

  How fast do random numbers and functions statistically converge to a Benford distribution?

  Sweepstakes Runner-up
  First Place
  Will go to State Science Fair
  Three Society Awards

Has also made $$ from Benford's Law

References

1. "Benford's Law" in Google gets 9,920 references.
2. Benford's Law now has a journal:

   Frequencies The Journal of Size Law Applications (pub 2001)
3. Benford's Law has hit the big-time newspapers:

   Wall Street Journal: L. Berton, *He's got their numbers: scholar uses math to foil financial fraud*, p. B1, 10 July 1995

   New York Times: Malcolm W. Brown, *Following Benford's Law, or Looking Out for No. 1*, 4 August 1998
4. Journal articles:

   Simon Newcomb, Amer J of Math, vol 4, 1881 (pp. 39–40)

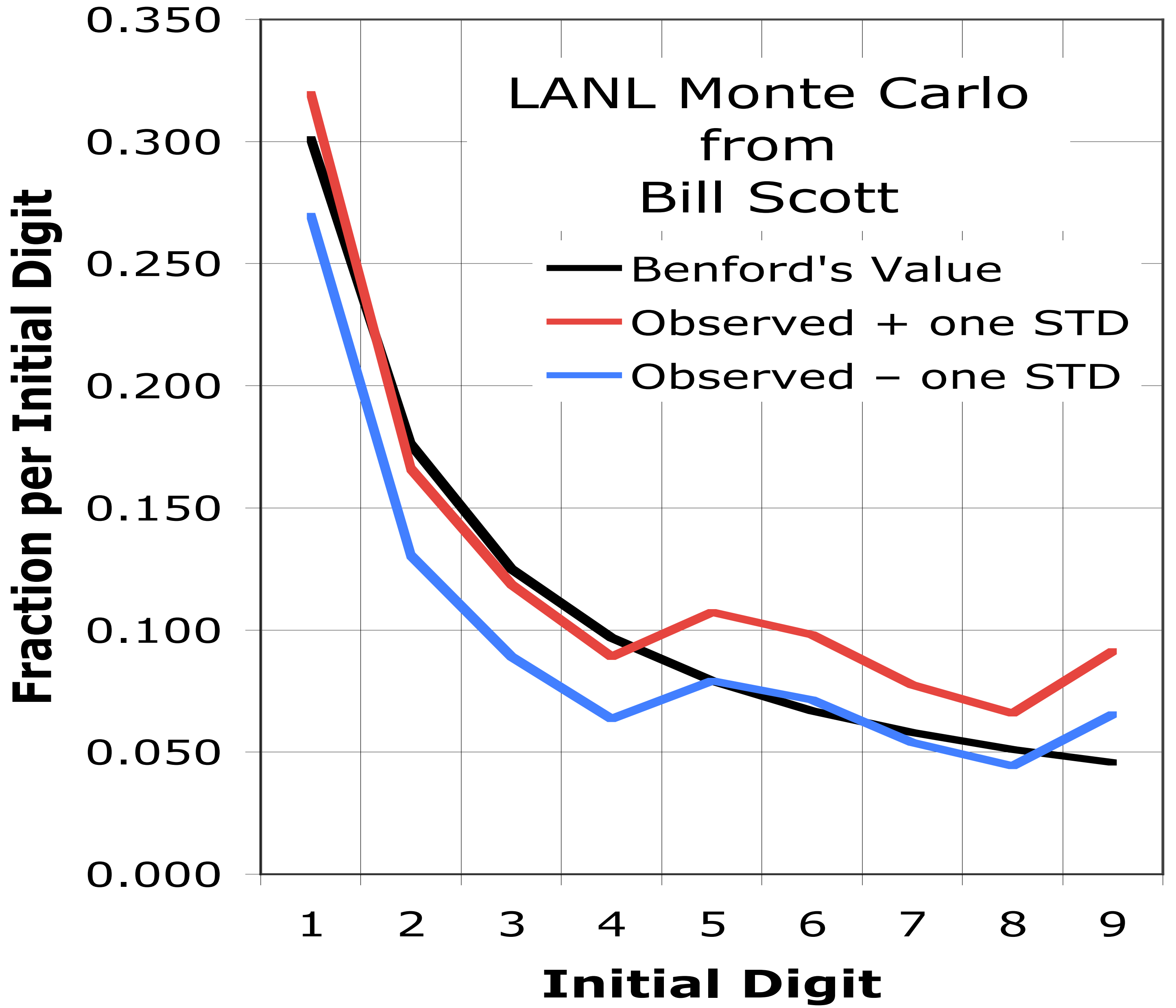
   Frank Benford, Proc of the Amer Philosophical Soc, vol 78, no 4, March, 1938 (pp 551–571)
5. Science Magazines:

   Ralph Raimi, Sci American, December 1969 (pp 109–119)

   T. P. Hill, Amer Scientist, vol 86, July-August 1998 (pp 358–363)
6. Book:

   March J. Nigrini, Ph.D., *Digital Analysis Using Benford's Law*, Global Audit Publications, Vamcouver, 2000 (ISBN 1-894497-09)

LANL Monte Carlo from Bill Scott



Benford:

$\chi^2 = 21$

Random:

$\chi^2 = 185$

Exponential:

$\chi^2 = 46$

if $\chi 2$ is small: good fit

if $\chi 2$ is large: poor fit

critical value of $\chi 2$:

Alpha	df	$\chi 2$
0.05	8	15.51

Benford TidBits

- Nigrini (2000) claims for a Benford distr. the sums of all numbers with each initial digit are equal. Close but no cigar; ±10%.

Chi-Squared vs. Significance

Goodness of Fit

Null Hypothesis: Observed distr. Is a random sample of the theoretical distr.

Significance (alpha): Risk of rejecting Null Hypothesis when it is true (Type I Error). Statisticians consider alpha of 0.05 to be statistically significant (1 chance in 20 of rejecting a true Null Hypothesis).

Chi-Sq	alpha	Chi-sq	alpha
15.5	0.05	2.7	0.95
13.4	0.10	2.2	0.975
10.2	0.25	1.7	0.99
7.3	0.50	1.3	0.995
6.2	0.75	0.9	0.999
3.5	0.90	0.5	0.9999

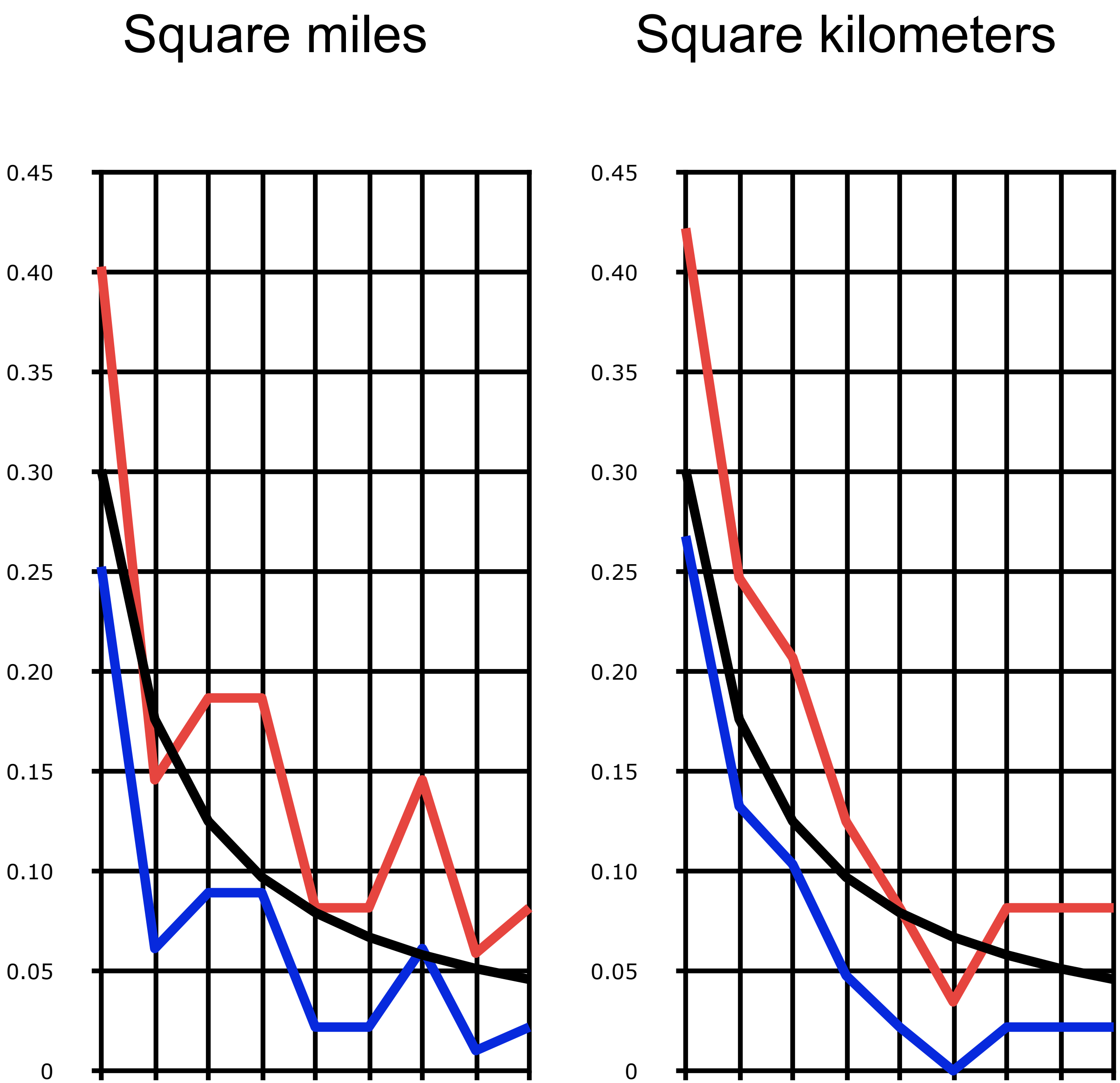
χ^2 for Interesting Sets of #s

	Sample #	Bedford	Random	Exp.
Fibonacci	20	3.3	8.8	--
Fibonacci	100	1.0	50	4.8
Fibonacci	200	0.7	98	8.5
Fibonacci	1,474 (1,1)	0.02	723	53
Lucas #s	1,474 (2,1)	0.06	723	53
"Groce #s"	1,474 (random)	0.05	716	54
Fibonacci	2,000,000	2×10^{-4}	--	--
Fibonacci	1,474 (div by π)	0.04	718	52
Fibonacci	1,474 (times e)	0.06	719	53
Fibonacci	Multiply each by random #s	5.2	700	56
Fibonacci	Multiply each twice by rand #s	3.9	693	71
Fibonacci	Square each term (739)	0.10	364	27
Bernoulli #s	32	4.0	14	--
Euler #s	30	11	25	12
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Addition?

- Addition of two random numbers [0 to 1] gives a triangular distribution with zero at both ends of interval (0 and 1) with a peak at 0.5. Scott (2005)
- Digit frequencies will be distorted by addition or subtraction unless the amount is small.

Area California Counties



58 Counties:

Square Miles:

$$\chi^2 = 6.1$$

Square Kilometers:

$$\chi^2 = 3.7$$

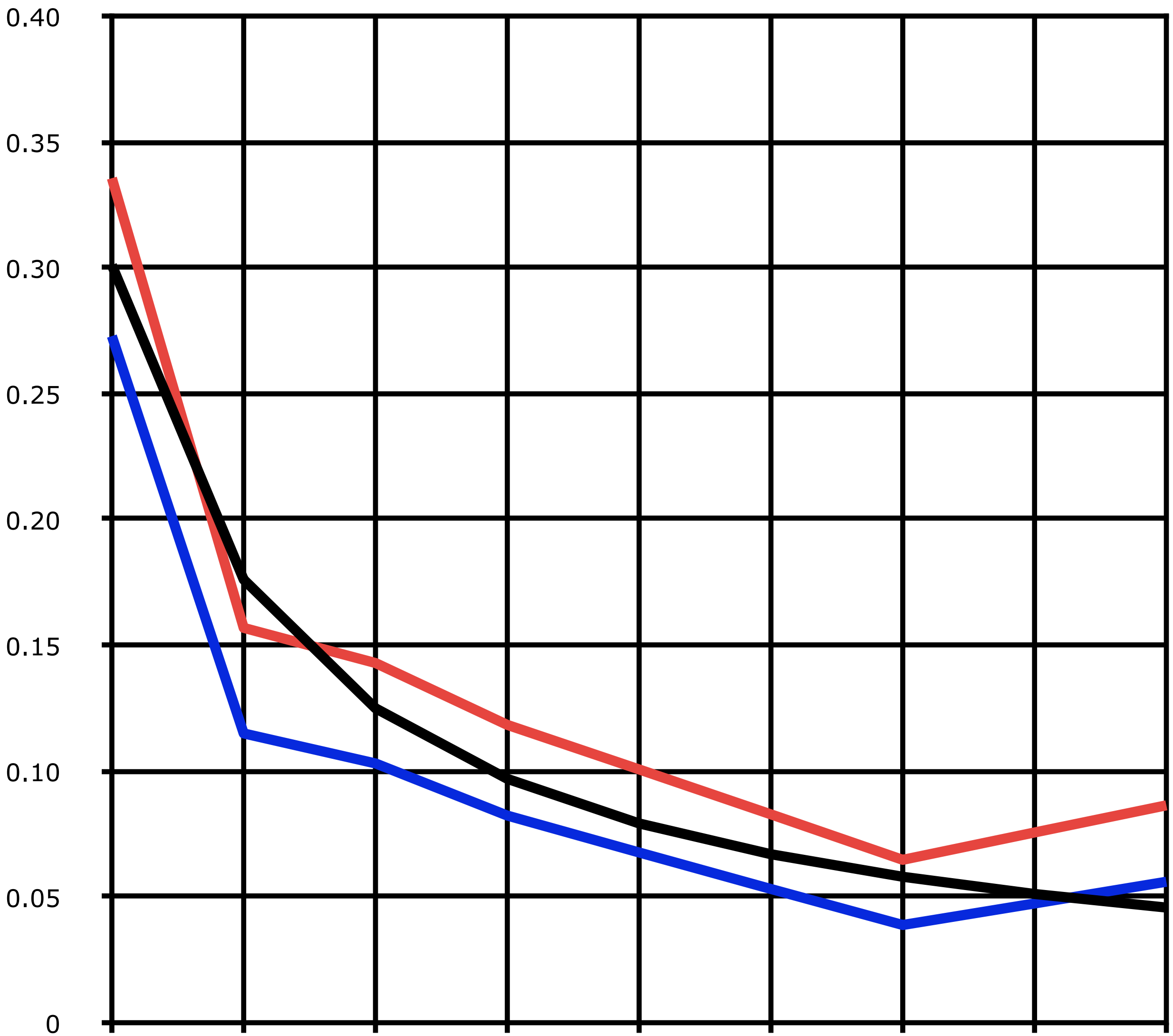
if $\chi 2$ is small: good fit

if $\chi 2$ is large: poor fit

critical value of $\chi 2$:

Alpha	df	$\chi 2$
0.05	8	15.51

Reciprocal Phys Constants



Benford:

$\chi^2 = 8.2$

Random:

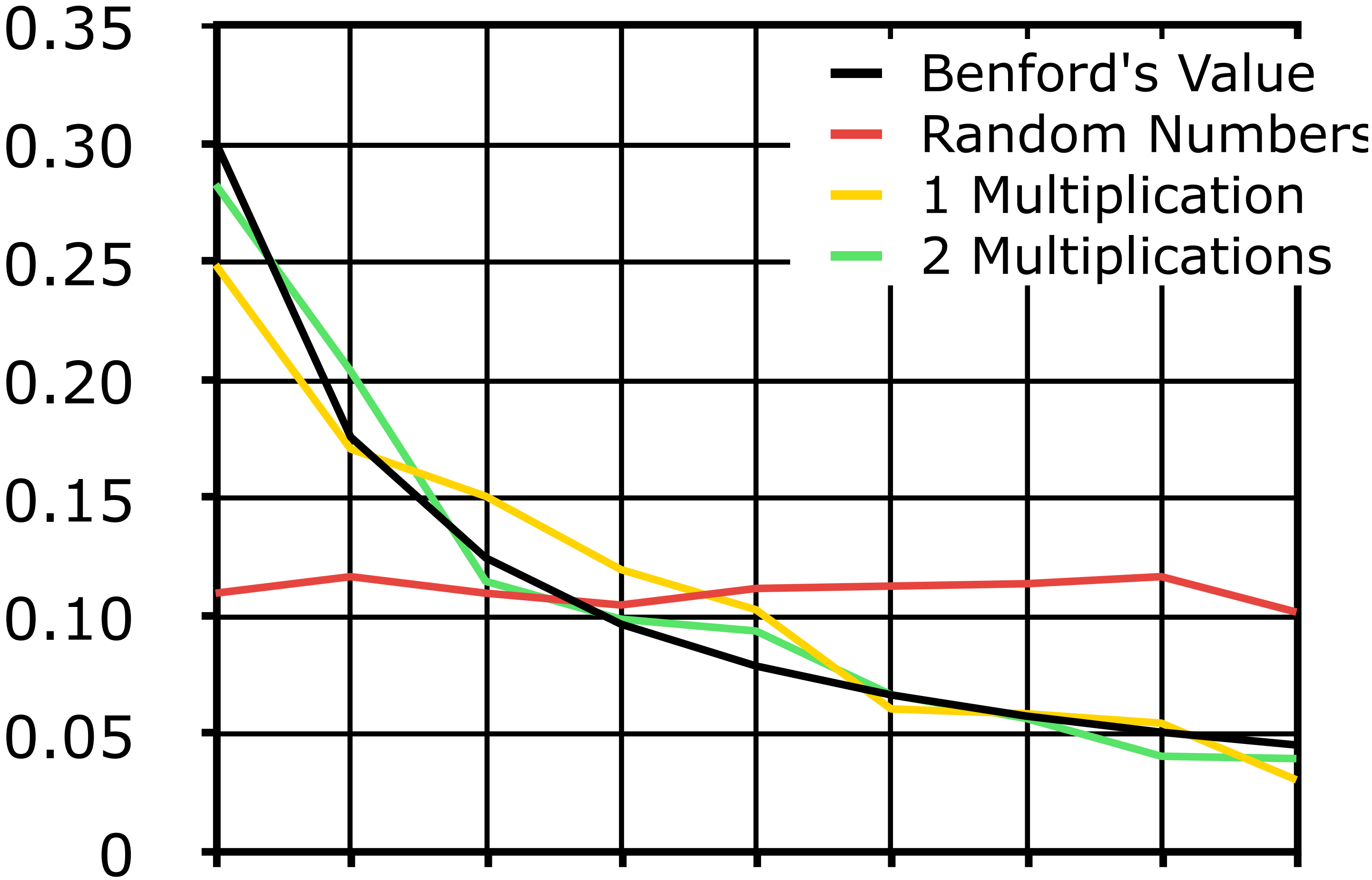
$\chi^2 = 134$

Exponential:

$\chi^2 = 26$

Multiplication of Random Numbers

Two random #s multiplied give a negative log distribution with a “soft” infinity at zero and zero at one. (Scott, 2005)



Random #s:

$\chi^2 = 397$

1 Multiplication:

$\chi^2 = 33$

2 Multiplications:

$\chi^2 = 12$

3 Multiplications:

$\chi^2 = 10$

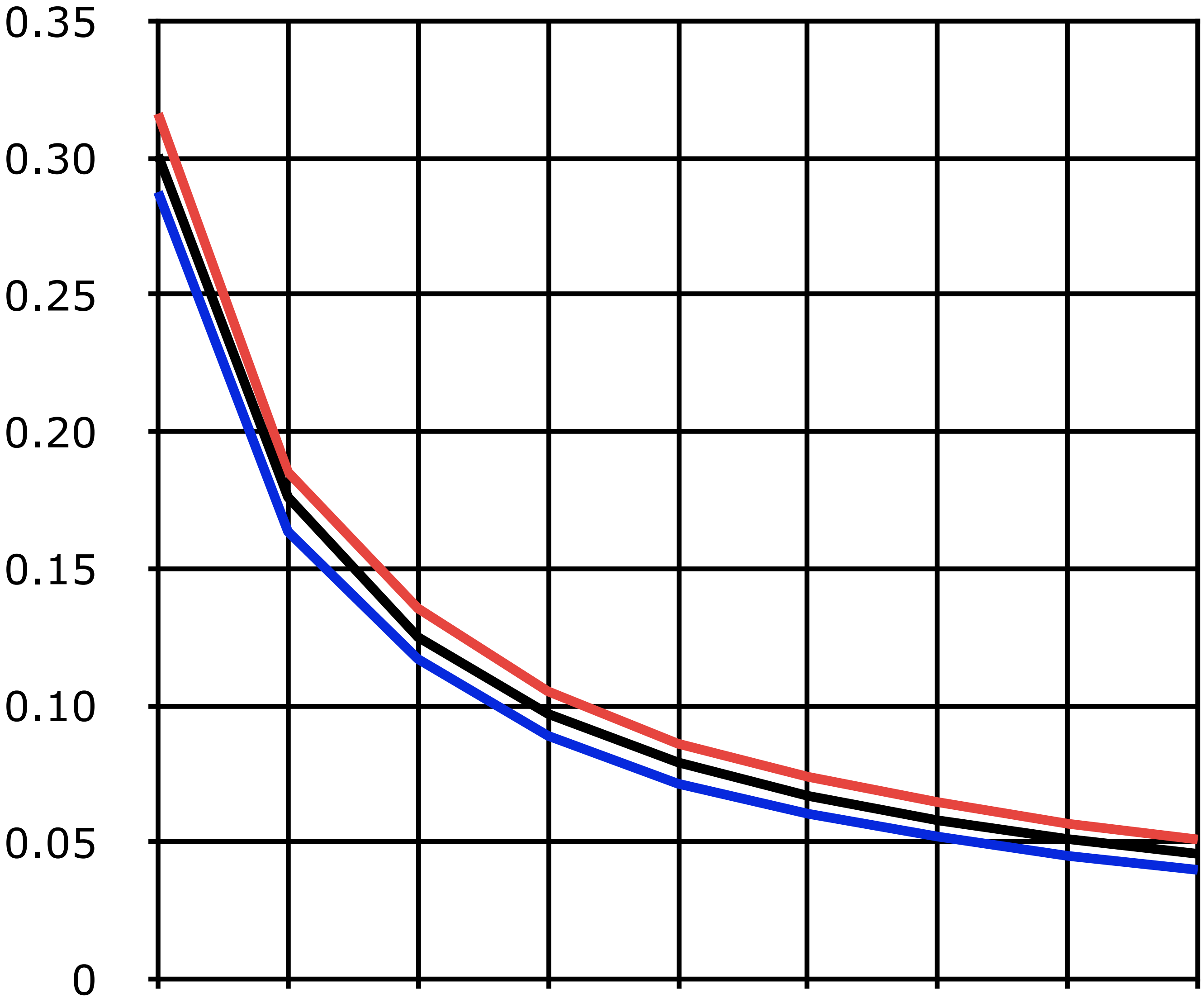
5 Multiplications:

$\chi^2 = 5.9$

9 Multiplications:

$\chi^2 = 5.5$

1,474 Lucas Numbers (2,1 start)



Benford:

$\chi^2 = .06$

Random:

$\chi^2 = 723$

Exponential:

$\chi^2 = 53$

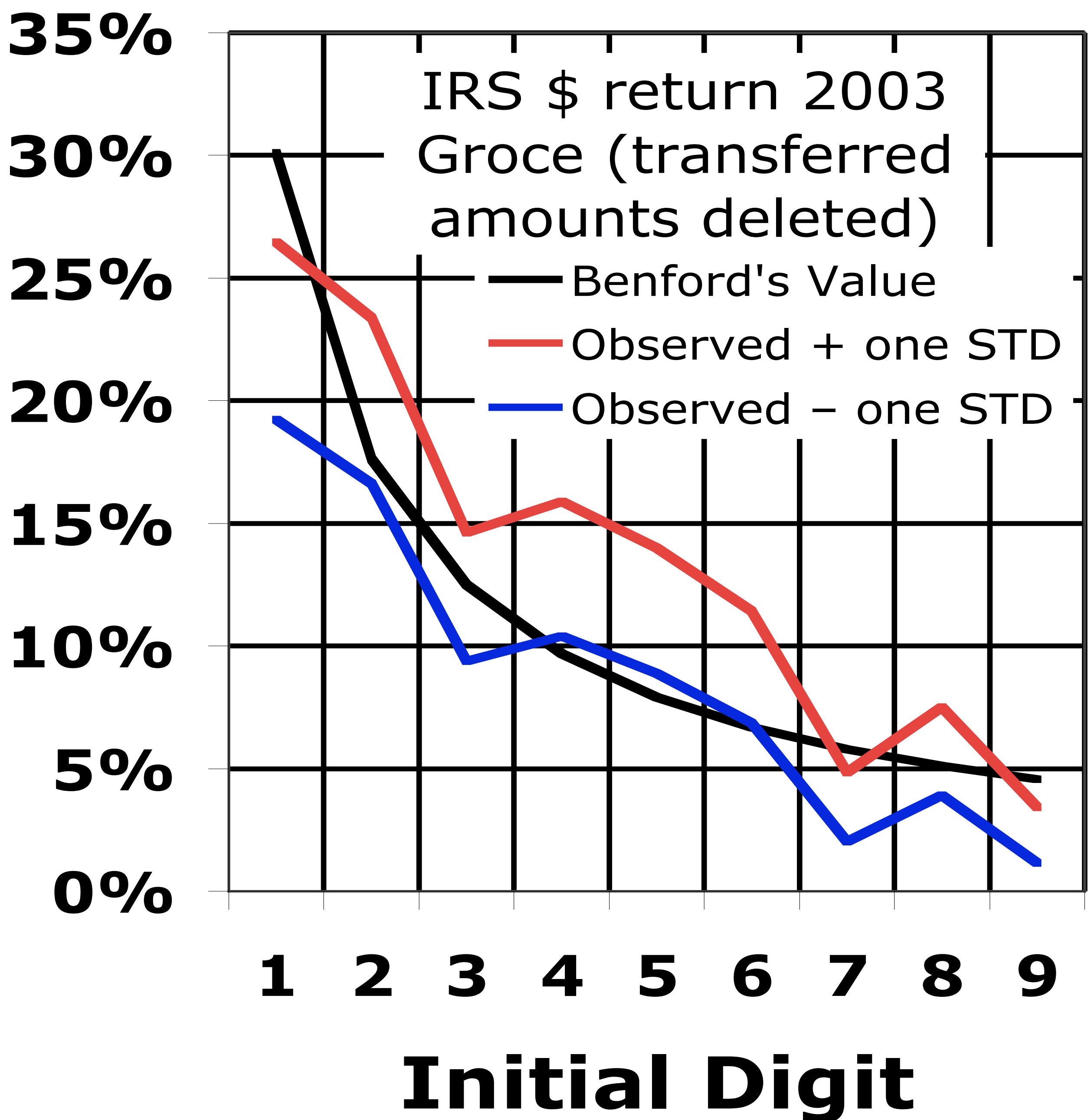
if $\chi 2$ is small: good fit

if $\chi 2$ is large: poor fit

critical value of $\chi 2$:

Alpha	df	$\chi 2$
0.05	8	15.51

Groce 2003 IRS Return



Benford:

$\chi^2 = 14$

Random:

$\chi^2 = 62$

Exponential:

$\chi^2 = 7$

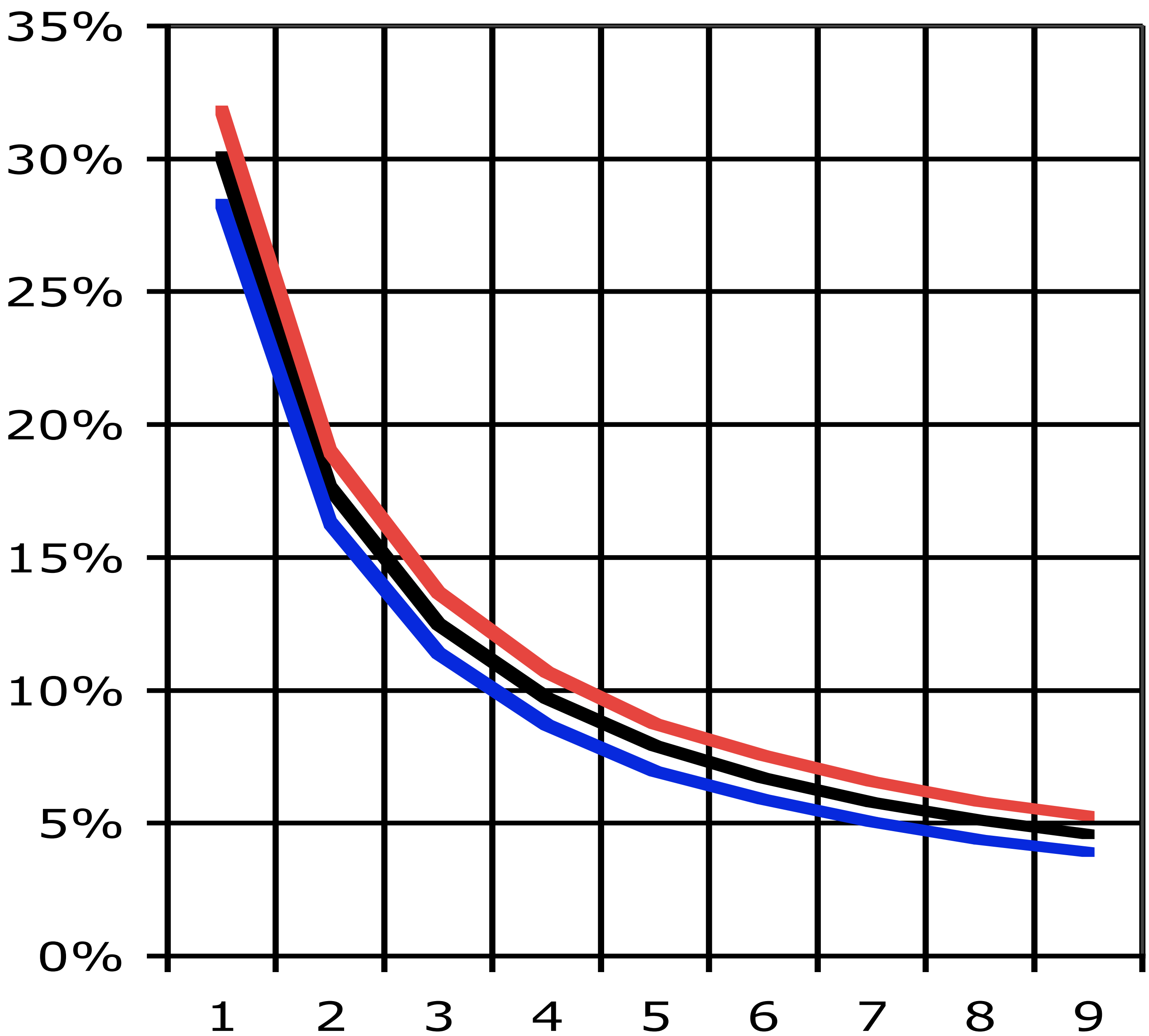
if $\chi 2$ is small: good fit

if $\chi 2$ is large: poor fit

critical value of $\chi 2$:

Alpha	df	$\chi 2$
0.05	8	15.51

Compound Interest ($100 to $1k in 981 periods)



Benford:

$\chi^2 = .01$

Random:

$\chi^2 = 479$

Exponential:

$\chi^2 = 35$

if $\chi 2$ is small: good fit

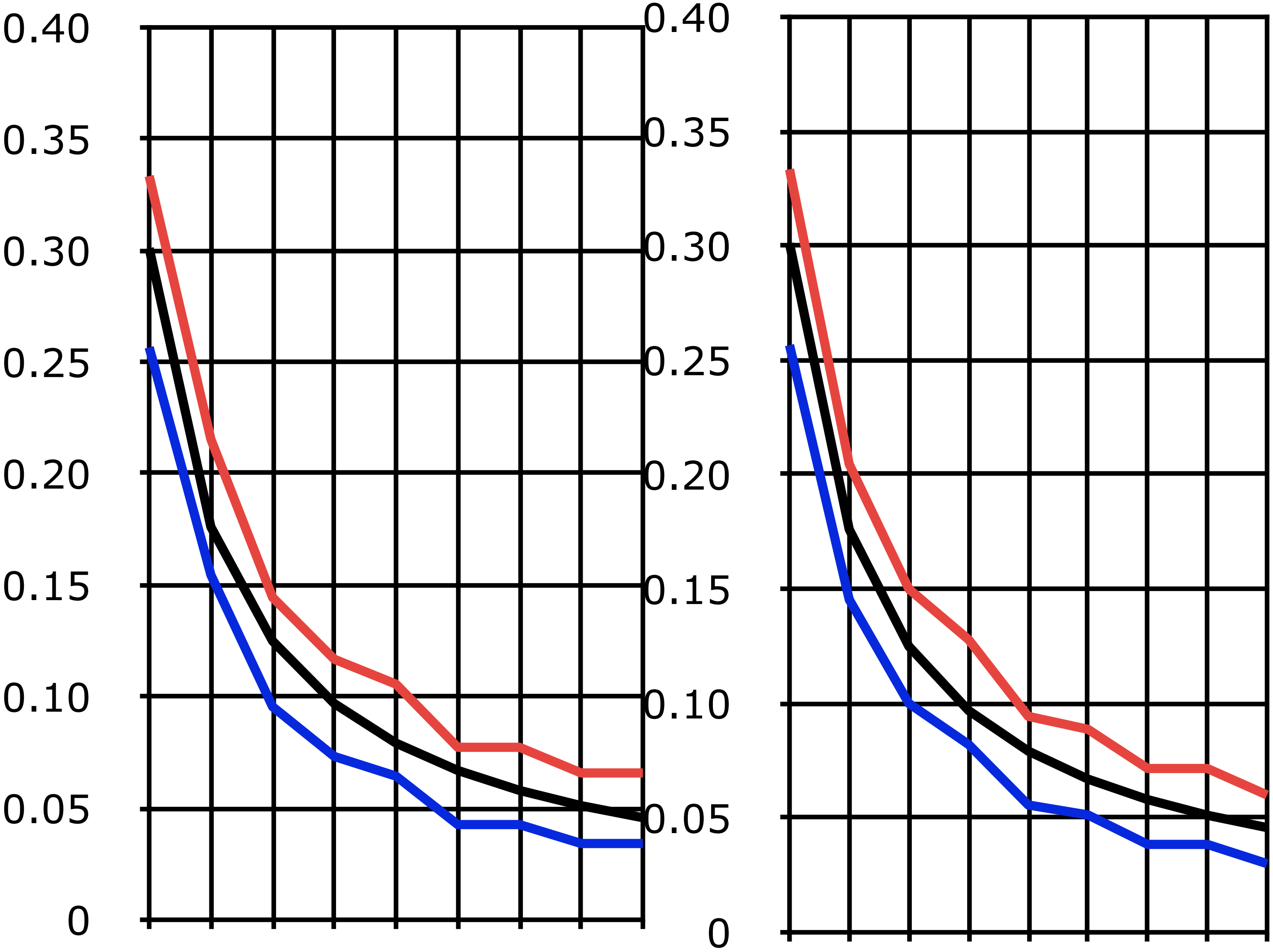
if $\chi 2$ is large: poor fit

critical value of $\chi 2$:

Alpha	df	$\chi 2$
0.05	8	15.51

First 200 Fibonacci Numbers

divide by π   multiply by e



Fibonacci:

$\chi^2 = 0.7$

Divided by π:

$\chi^2 = 0.3$

Multiplied by e:

$\chi^2 = 0.5$

if $\chi 2$ is small: good fit

if $\chi 2$ is large: poor fit

critical value of $\chi 2$:

Alpha	df	$\chi 2$
0.05	8	15.51