Power-Law Size Distributions

Last updated: 2024/09/16, 22:10:34 EDT

Principles of Complex Systems, Vols. 1, 2, & 3D
CSYS/MATH 6701, 6713, & a pretend number, 2024–2025

Prof. Peter Sheridan Dodds

Computational Story Lab | Vermont Complex Systems Center
Santa Fe Institute | University of Vermont

Outline

Our Intuition

Definition

Examples

Wild vs. Mild

CCDFs

Size rankings and Zipf's law

Size ranking ⇔ CCDF

References

Two of the many things we struggle with cognitively:

1. Probability.
   - Ex. The Monty Hall Problem. ⌷
   - Ex. Daughter/Son born on Tuesday. ⌷
     (see next two slides; Wikipedia entry here ⌷.)
2. Logarithmic scales.

On counting and logarithms:

- Listen to Radiolab's 2009 piece: "Numbers." ⌷.
- Later: Benford's Law ⌷.

Also to be enjoyed: The Dunning-Kruger effect ⌷[1]

---

[1] 2000 Ig Nobel winners ⌷

The PoCSverse
Power-Law Size
Distributions
1 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Homo probabilisticus?

The set up:

- A parent has two children.

Simple probability question:

- What is the probability that both children are girls?

The next set up:

- A parent has two children.
- We know one of them is a girl.

The next probabilistic poser:

- What is the probability that both children are girls?

The PoCSverse
Power-Law Size
Distributions
2 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Try this one:

- A parent has two children.
- We know one of them is a girl born on a Tuesday.

Simple question #3:

- What is the probability that both children are girls?

Last:

- A parent has two children.
- We know one of them is a girl born on December 31.

And ...

- What is the probability that both children are girls?

The PoCSverse
Power-Law Size
Distributions
3 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Let's test our collective intuition:

$$\text{Money} \equiv \text{Belief}$$

Two questions about wealth distribution in the United States:

1. Please estimate the percentage of all wealth owned by individuals when grouped into quintiles.
2. Please estimate what you believe each quintile should own, ideally.
3. Extremes: 100, 0, 0, 0, 0 and 20, 20, 20, 20, 20.

The PoCSverse
Power-Law Size
Distributions
4 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

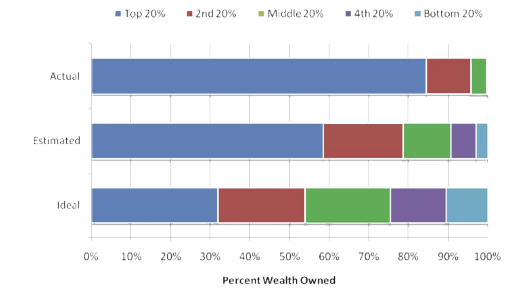## Wealth distribution in the United States: [13]

Fig. 2. The actual United States wealth distribution plotted against the estimated and ideal distributions across all respondents. Because of their small percentage share of total wealth, both the "4th 20%" value (0.2%) and the "Bottom 20%" value (0.1%) are not visible in the "Actual" distribution.

"Building a better America—One wealth quintile at a time"
Norton and Ariely, 2011. [13]
But: Fraud. ⌷

The PoCSverse
Power-Law Size
Distributions
5 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

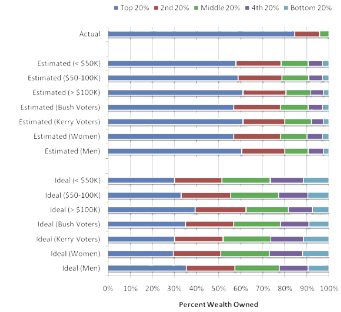## Wealth distribution in the United States: [13]

Fig. 3. The actual United States wealth distribution plotted against the estimated and ideal distributions of respondents of different income levels, political affiliations, and genders. Because of their small percentage share of total wealth, both the "4th 20%" value (0.2%) and the "Bottom 20%" value (0.1%) are not visible in the "Actual" distribution.

- A highly watched video based on this research is here. ⌷

The PoCSverse
Power-Law Size
Distributions
6 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

The PoCSverse
Power-Law Size
Distributions
7 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

The sizes of many systems' elements appear to obey an inverse power-law size distribution:

$$P(\text{size} = x) \sim c\,x^{-\gamma}$$

where $\quad 0 < x_{\min} < x < x_{\max} \quad$ and $\quad \gamma > 1$.

- $x_{\min}$ = lower cutoff, $x_{\max}$ = upper cutoff
- Negative linear relationship in log-log space:

$$\log_{10} P(x) = \log_{10} c - \gamma \log_{10} x$$

- We use base 10 because we are good people.

The PoCSverse
Power-Law Size
Distributions
9 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
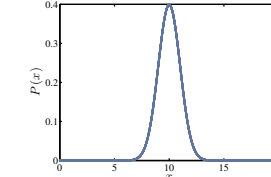Size ranking ⇔
CCDF
References

The PoCSverse
Power-Law Size
Distributions
10 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Size distributions:

Usually, only the tail of the distribution obeys a power law:

$$P(x) \sim c\,x^{-\gamma} \text{ for } x \text{ large.}$$

- Still use term 'power-law size distribution.'
- Other terms:
  - **Fat-tailed** distributions.
  - **Heavy-tailed** distributions.

Beware:

- Inverse power laws aren't the only ones:
  lognormals [↗], Weibull distributions [↗], …

The PoCSverse
Power-Law Size
Distributions
11 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Jonathan Harris's (not quite dead) Wordcount: [↗]

A word frequency distribution explorer:

The PoCSverse
Power-Law Size
Distributions
14 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
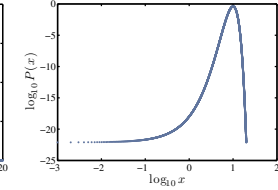References

The statistics of surprise—words:

First—a Gaussian example:

$$P(x)\mathrm{d}x = \frac{1}{\sqrt{2\pi}\sigma}e^{-(x-\mu)^2/2\sigma^2}\,\mathrm{d}x$$

linear:

log-log

mean $\mu = 10$, variance $\sigma^2 = 1$.

- Activity: Sketch $P(x) \sim x^{-1}$ for $x = 1$ to $x = 10^7$.

The PoCSverse
Power-Law Size
Distributions
17 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Size distributions:

Many systems have discrete sizes k:

- Word frequency
- Node degree in networks: # friends, # hyperlinks, etc.
- # citations for articles, court decisions, etc.

$$P(k) \sim c\,k^{-\gamma}$$

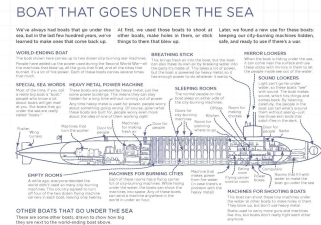$$\text{where } k_{\min} \leq k \leq k_{\max}$$

- Obvious fail for $k = 0$.
- Again, typically a description of distribution's tail.
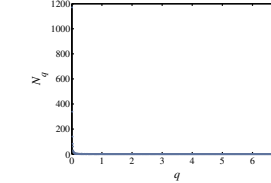
The PoCSverse
Power-Law Size
Distributions
12 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

"Thing Explainer: Complicated Stuff in Simple Words" [a] [↗]
by Randall Munroe (2015). [11]

Up goer five [↗]

The PoCSverse
Power-Law Size
Distributions
15 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
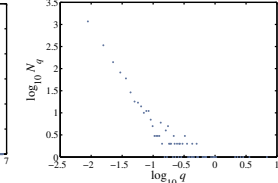Size ranking ⇔
CCDF
References

The statistics of surprise—words:

Raw 'probability' (binned) for Brown Corpus:

linear:

log-log

- q_w = normalized frequency of occurrence of word w (%).
- N_q = number of distinct words that have a normalized frequency of occurrence q.
- e.g, $q_{\text{the}} \simeq 6.9\%$, $N_{q_{\text{the}}} = 1$.

The PoCSverse
Power-Law Size
Distributions
18 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Word rank and frequency:

Brown Corpus [↗] ($\sim 10^6$ words):

rank	word	% q	rank	word	% q
1.	the	6.8872	1945.	apply	0.0055
2.	of	3.5839	1946.	vital	0.0055
3.	and	2.8401	1947.	September	0.0055
4.	to	2.5744	1948.	review	0.0055
5.	a	2.2996	1949.	wage	0.0055
6.	in	2.1010	1950.	motor	0.0055
7.	that	1.0428	1951.	fifteen	0.0055
8.	is	0.9943	1952.	regarded	0.0055
9.	was	0.9661	1953.	draw	0.0055
10.	he	0.9392	1954.	wheel	0.0055
11.	for	0.9340	1955.	organized	0.0055
12.	it	0.8623	1956.	vision	0.0055
13.	with	0.7176	1957.	wild	0.0055
14.	as	0.7137	1958.	Palmer	0.0055
15.	his	0.6886	1959.	intensity	0.0055

Later: Connect rankings and size distributions.

The PoCSverse
Power-Law Size
Distributions
13 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

The long tail of knowledge:

Cited by — VIEW ALL

	All	Since 2016
Citations	432350	165872
h-index	155	104
i10-index	369	277

Take a scrolling voyage
to the citational abyss,
starting at the surface with
the lonely, giant citaceans,
moving down
to the legion of strange,
sometimes misplaced,
unloved creatures,
that dwell in
Kahneman's Google Scholar page [↗]

- Papers are the events, size is the number of citations
- Natural to order by size or publication date.

The PoCSverse
Power-Law Size
Distributions
16 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
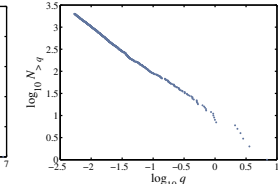Zipf's law
Size ranking ⇔
CCDF
References

The statistics of surprise—words:

Complementary Cumulative Distribution (for frequency or probability) $N_{\geq q}$:

linear:
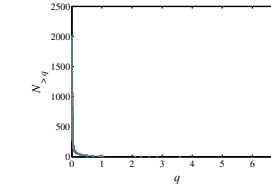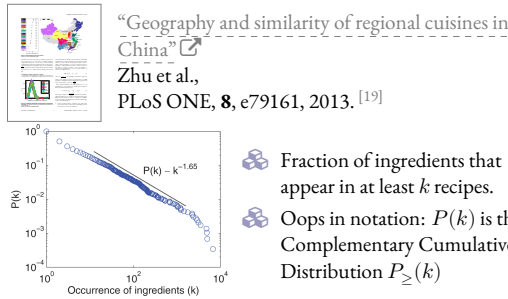
log-log

- Also known as the 'Exceedance Probability.'

The PoCSverse
Power-Law Size
Distributions
19 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

My, what big words you have ...

The PoCSverse
Power-Law Size
Distributions
20 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

How many words do you know?

🐝 Test ⬦ capitalizes on word frequency following a heavily skewed frequency distribution with a decaying power-law tail.

🐝 This Man Can Pronounce Every Word in the Dictionary ⬦ (story here ⬦)

🐝 Best of Dr. Bailly ⬦

The statistics of surprise:

The PoCSverse
Power-Law Size
Distributions
21 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
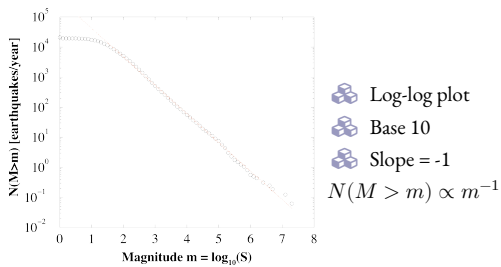Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Gutenberg-Richter law ⬦

🐝 Log-log plot

🐝 Base 10

🐝 Slope = -1
$N(M > m) \propto m^{-1}$

🐝 From both the very awkwardly similar Christensen et al. and Bak et al.:
"Unified scaling law for earthquakes" [4, 1]

The statistics of surprise:

The PoCSverse
Power-Law Size
Distributions
22 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
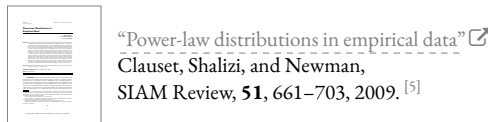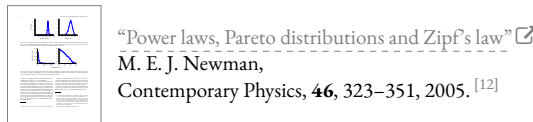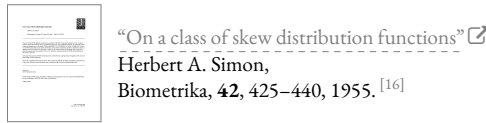Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

From: "Quake Moves Japan Closer to U.S. and Alters Earth's Spin" ⬦ by Kenneth Chang, March 13, 2011, NYT:

'What is perhaps most surprising about the Japan earthquake is how misleading history can be. In the past 300 years, no earthquake nearly that large—nothing larger than magnitude eight—had struck in the Japan subduction zone. That, in turn, led to assumptions about how large a tsunami might strike the coast.'

"'It did them a giant disservice," said Dr. Stein of the geological survey. That is not the first time that the earthquake potential of a fault has been underestimated. Most geophysicists did not think the Sumatra fault could generate a magnitude 9.1 earthquake, ...'

The PoCSverse
Power-Law Size
Distributions
23 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

"Geography and similarity of regional cuisines in China" ⬦
Zhu et al.,
PLoS ONE, **8**, e79161, 2013. [19]

🐝 Fraction of ingredients that appear in at least k recipes.

🐝 Oops in notation: $P(k)$ is the Complementary Cumulative Distribution $P_\geq(k)$

The PoCSverse
Power-Law Size
Distributions
24 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

"On a class of skew distribution functions" ⬦
Herbert A. Simon,
Biometrika, **42**, 425–440, 1955. [16]

"Power laws, Pareto distributions and Zipf's law" ⬦
M. E. J. Newman,
Contemporary Physics, **46**, 323–351, 2005. [12]

"Power-law distributions in empirical data" ⬦
Clauset, Shalizi, and Newman,
SIAM Review, **51**, 661–703, 2009. [5]

The PoCSverse
Power-Law Size
Distributions
25 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
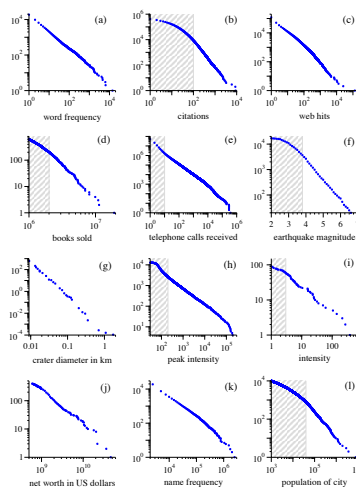Size ranking ⇔
CCDF
References

FIG. 4 Cumulative distributions or "rank/frequency plots" of twelve quantities reputed to follow power laws. The distributions were computed as described in Appendix A. Data in the shaded regions were excluded from the calculations of the exponents in Table I. Source references for the data are given in the text. (a) Numbers of occurrences of words in the novel *Moby Dick* by Herman Melville. (b) Numbers of citations to scientific papers published in 1981, from time of publication until June 1997. (c) Numbers of hits on web sites by 60 000 users of the America Online Internet service for the day of 1 December 1997. (d) Numbers of copies of bestselling books sold in the US between 1895 and 1965. (e) Number of calls received by AT&T telephone customers in the US for a single day. (f) Magnitude of earthquakes in California between January 1910 and May 1992. Magnitude is proportional to the logarithm of the maximum amplitude of the earthquake, and hence the distribution obeys a power law even though the horizontal axis is linear. (g) Diameter of craters on the moon. Vertical axis is measured per square kilometre. (h) Peak gamma-ray intensity of solar flares in counts per second, measured from Earth orbit between February 1980 and November 1989. (i) Intensity of wars from 1816 to 1980, measured as battle deaths per 10 000 of the population of the participating countries. (j) Aggregate net worth in dollars of the richest individuals in the US in October 2003. (k) Frequency of occurrence of family names in the US in the year 1990. (l) Population of US cities in the year 2000.

Size distributions:

The PoCSverse
Power-Law Size
Distributions
26 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Some examples:

🐝 Earthquake magnitude (Gutenberg-Richter law ⬦): [9, 1]
$P(M) \propto M^{-2}$

🐝 # war deaths: [15] $P(d) \propto d^{-1.8}$

🐝 Sizes of forest fires [8]

🐝 Sizes of cities: [16] $P(n) \propto n^{-2.1}$

🐝 # links to and from websites [2]

🐝 Note: Exponents range in error

Size distributions:

The PoCSverse
Power-Law Size
Distributions
27 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

More examples:

🐝 # citations to papers: [6, 14] $P(k) \propto k^{-3}$.

🐝 Individual wealth (maybe): $P(W) \propto W^{-2}$.

🐝 Distributions of tree trunk diameters: $P(d) \propto d^{-2}$.

🐝 The gravitational force at a random point in the universe: [10] $P(F) \propto F^{-5/2}$. (See the Holtsmark distribution ⬦ and stable distributions ⬦.)

🐝 Diameter of moon craters: [12] $P(d) \propto d^{-3}$.

🐝 Word frequency: [16] e.g., $P(k) \propto k^{-2.2}$ (variable).

🐝 # religious adherents in cults: [5] $P(k) \propto k^{-1.8 \pm 0.1}$.

🐝 # sightings of birds per species (North American Breeding Bird Survey for 2003): [5] $P(k) \propto k^{-2.1 \pm 0.1}$.

🐝 # species per genus: [18, 16, 5] $P(k) \propto k^{-2.4 \pm 0.2}$.

Table 3 from Clauset, Shalizi, and Newman [5]:

The PoCSverse
Power-Law Size
Distributions
28 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Basic parameters of the data sets described in section 6, along with their power-law fits and the corresponding p-values (statistically significant values are denoted in **bold**).

Quantity	n	$\langle x \rangle$	σ	x_{max}	$\hat{x}_{min}$	$\hat{\alpha}$	n_{tail}	p
count of word use	18 855	11.14	148.33	14 086	7 ± 2	1.95(2)	2958 ± 987	0.49
protein interaction degree	1846	2.34	3.05	56	5 ± 2	3.1(3)	204 ± 263	0.31
metabolic degree	1641	5.68	17.81	468	4 ± 1	2.8(1)	748 ± 136	0.00
Internet degree	22 688	5.63	37.83	2583	21 ± 9	2.12(9)	770 ± 1124	0.29
telephone calls received	51 360 423	3.88	179.09	375 746	120 ± 49	2.09(1)	$102 592 \pm 210 147$	0.63
intensity of wars	115	15.70	49.97	382	2.1 ± 3.5	1.7(2)	70 ± 14	0.20
terrorist attack severity	9101	4.35	31.58	2749	12 ± 4	2.4(2)	547 ± 1663	0.68
HTTP size (kilobytes)	226 386	7.36	57.94	10 971	36.25 ± 22.74	2.48(5)	6794 ± 2232	0.00
species per genus	509	5.59	6.94	56	4 ± 2	2.4(2)	233 ± 138	0.10
bird species sightings	591	3384.36	10 952.34	138 705	6679 ± 2463	2.1(2)	66 ± 41	0.55
blackouts ($\times 10^3$)	211	253.87	610.31	7500	230 ± 90	2.3(3)	59 ± 35	0.62
sales of books ($\times 10^3$)	633	1986.67	1396.60	19 077	2400 ± 430	3.7(3)	139 ± 115	0.66
population of cities ($\times 10^3$)	19 447	9.00	77.83	8009	52.46 ± 11.88	2.37(8)	580 ± 177	0.76
email address books size	4581	12.45	21.49	333	57 ± 21	3.5(6)	196 ± 449	0.16
forest fire size (acres)	203 785	0.90	20.99	4121	6324 ± 3487	2.2(3)	521 ± 6801	0.05
solar flare intensity	12 773	689.41	6520.59	231 300	323 ± 89	1.79(2)	1711 ± 384	1.00
quake intensity ($\times 10^3$)	19 302	24.54	563.83	63 096	0.794 ± 80.198	1.64(4)	$11 697 \pm 2159$	0.00
religious followers ($\times 10^6$)	103	27.36	136.64	1050	3.85 ± 1.60	1.8(1)	39 ± 26	0.42
freq. of surnames ($\times 10^3$)	2753	50.59	113.99	2502	111.92 ± 40.67	2.5(2)	239 ± 215	0.20
net worth (mil. USD)	400	2388.69	4167.35	46 000	900 ± 364	2.3(1)	302 ± 77	0.00
citations to papers	415 229	16.17	44.02	8904	160 ± 35	3.16(6)	3455 ± 1859	0.20
papers authored	401 445	7.21	16.52	1416	133 ± 13	4.3(1)	988 ± 377	0.90
hits to web sites	119 724	9.83	392.52	129 641	2 ± 13	1.81(8)	$50 981 \pm 16 898$	0.00
links to web sites	241 428 853	9.15	106 871.65	1 199 466	3684 ± 151	2.336(5)	$28 986 \pm 1560$	0.00

🐝 We'll explore various exponent measurement techniques in assignments.

power-law size distributions

Gaussians versus power-law size distributions:

- 🎲 Mediocristan versus Extremistan
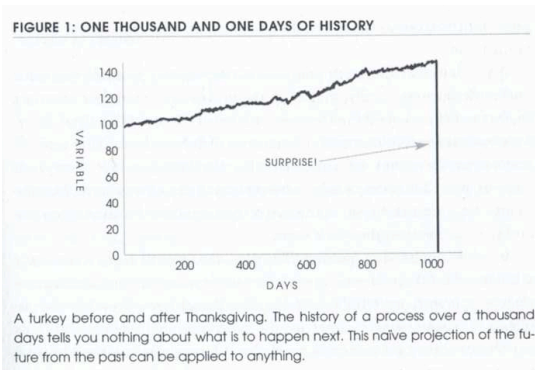- 🎲 Mild versus Wild (Mandelbrot)
- 🎲 Example: Height versus wealth.

THE
BLACK SWAN

The Impact of the
HIGHLY IMPROBABLE

Nassim Nicholas Taleb

- 🎲 See "The Black Swan" by Nassim Taleb. [17]
- 🎲 Terrible if successful framing: Black swans are not that surprising …

The PoCSverse
Power-Law Size
Distributions
29 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Size distributions:

Power-law size distributions are sometimes called Pareto distributions 🔗 after Italian scholar Vilfredo Pareto. 🔗

- 🎲 Pareto noted wealth in Italy was distributed unevenly (80/20 rule; misleading, see later).
- 🎲 Term used especially by practitioners of the Dismal Science 🔗.

The PoCSverse
Power-Law Size
Distributions
32 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

"The horror, the horror …" 🔗

Moments:

- 🎲 All moments depend only on cutoffs.
- 🎲 No internal scale that dominates/matters.
- 🎲 Compare to a Gaussian, exponential, etc.

For many real size distributions: $2 < \gamma < 3$

- 🎲 mean is finite (depends on lower cutoff)
- 🎲 σ^2 = variance is 'infinite' (depends on upper cutoff)
- 🎲 Width of distribution is 'infinite'
- 🎲 If $\gamma > 3$, distribution is less terrifying and may be easily confused with other kinds of distributions.

Insert assignment question 🔗

The PoCSverse
Power-Law Size
Distributions
35 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Turkeys …

FIGURE 1: ONE THOUSAND AND ONE DAYS OF HISTORY

A turkey before and after Thanksgiving. The history of a process over a thousand days tells you nothing about what is to happen next. This naïve projection of the future from the past can be applied to anything.

From "The Black Swan" [17]

The PoCSverse
Power-Law Size
Distributions
30 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Devilish power-law size distribution details:

Exhibit A:

- 🎲 Given $P(x) = cx^{-\gamma}$ with $0 < x_{\min} < x < x_{\max}$, the mean is ($\gamma \neq 2$):

$$\langle x \rangle = \frac{c}{2-\gamma}\left(x_{\max}^{2-\gamma} - x_{\min}^{2-\gamma}\right).$$

- 🎲 Mean 'blows up' with upper cutoff if $\gamma < 2$.
- 🎲 Mean depends on lower cutoff if $\gamma > 2$.
- 🎲 $\gamma < 2$: Typical sample is large.
- 🎲 $\gamma > 2$: Typical sample is small.

Insert assignment question 🔗

The PoCSverse
Power-Law Size
Distributions
33 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Moments

Standard deviation is a mathematical convenience:

- 🎲 Variance is nice analytically …
- 🎲 Another measure of distribution width:

$$\text{Mean average deviation (MAD)} = \langle |x - \langle x \rangle| \rangle$$

- 🎲 For a pure power law with $2 < \gamma < 3$:

$$\langle |x - \langle x \rangle| \rangle \text{ is finite.}$$

- 🎲 But MAD is mildly unpleasant analytically …
- 🎲 We still speak of infinite 'width' if $\gamma < 3$.

The PoCSverse
Power-Law Size
Distributions
36 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Taleb's table [17]

Mediocristan/Extremistan

- 🎲 Most typical member is mediocre/Most typical is either giant or tiny
- 🎲 Winners get a small segment/Winner take almost all effects
- 🎲 When you observe for a while, you know what's going on/It takes a very long time to figure out what's going on
- 🎲 Prediction is easy/Prediction is hard
- 🎲 History crawls/History makes jumps
- 🎲 Tyranny of the collective/Tyranny of the rare and accidental

---

Moments:

- 🎲 If $n \neq \gamma - 1$:

$$\langle x^n \rangle = \int_{x_{\min}}^{x_{\max}} x^n P(x)\, \mathrm{d}x = \frac{c}{n-\gamma+1}\left(x_{\max}^{n-\gamma+1} - x_{\min}^{n-\gamma+1}\right).$$

where $c = \dfrac{\gamma-1}{a^{-(\gamma-1)} - b^{-(\gamma-1)}}$.

- 🎲 Because both $n - \gamma + 1$ and $\left(x_{\max}^{n-\gamma+1} - x_{\min}^{n-\gamma+1}\right)$ are either negative or positive, we can write:

$$\langle x^n \rangle = \frac{c}{|n-\gamma+1|}\left|x_{\max}^{n-\gamma+1} - x_{\min}^{n-\gamma+1}\right|.$$

- 🎲 If $n = \gamma - 1$:

$$\langle x^n \rangle = c\frac{x_{\max}}{x_{\min}}.$$

The PoCSverse
Power-Law Size
Distributions
31 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

How sample sizes grow …

Given $P(x) \sim cx^{-\gamma}$:

- 🎲 We can show that after n samples, we expect the largest sample to be[2]

$$x_1 \gtrsim c' n^{1/(\gamma-1)}$$

- 🎲 Sampling from a finite-variance distribution gives a much slower growth with n.
- 🎲 e.g., for $P(x) = \lambda e^{-\lambda x}$, we find

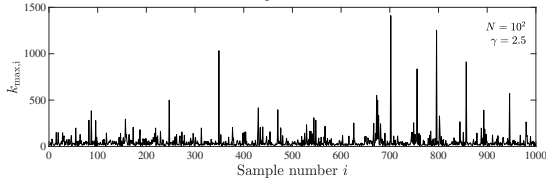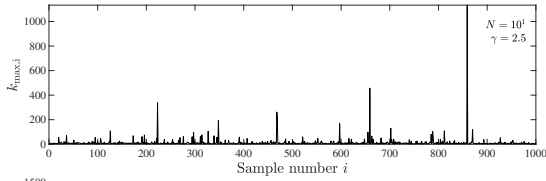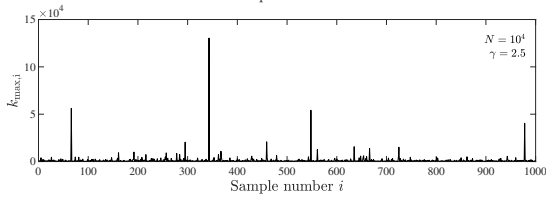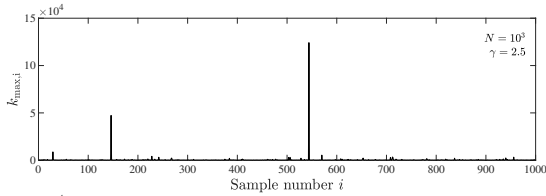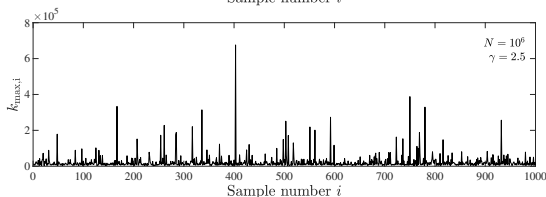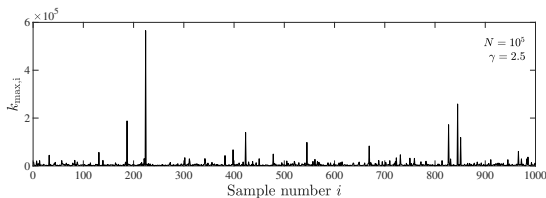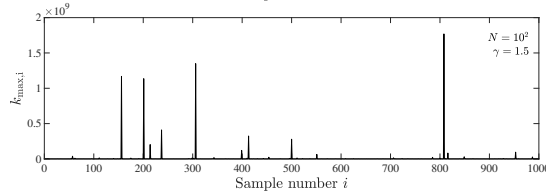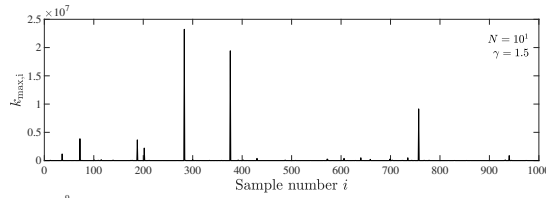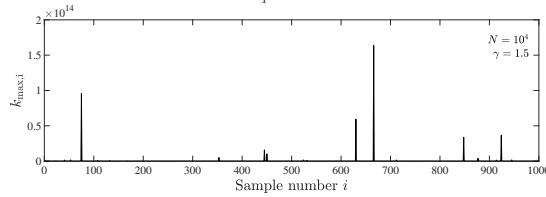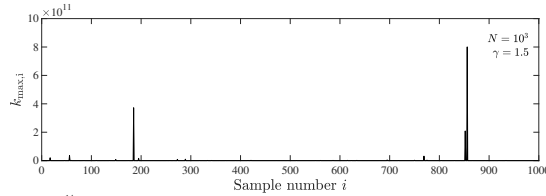$$x_1 \gtrsim \frac{1}{\lambda}\ln n.$$

Insert assignment question 🔗

---

[2]Later, we see that the largest sample grows as n^α where α is the size-ranking exponent

The PoCSverse
Power-Law Size
Distributions
34 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

The PoCSverse
Power-Law Size
Distributions
38 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References
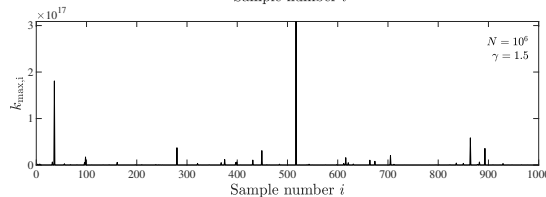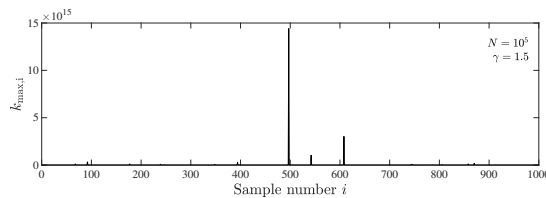
$\gamma = 5/2$, maxima of N samples, $n = 1000$ sets of samples:

$\gamma = 3/2$, maxima of N samples, $n = 1000$ sets of samples:

The PoCSverse
Power-Law Size
Distributions
41 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Scaling of expected largest value as a function of sample size N:

The PoCSverse
Power-Law Size
Distributions
44 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References
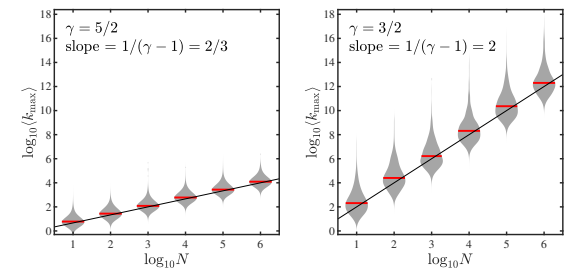
Fit for $\gamma = 5/2$:[3] $k_{\max} \sim N^{0.660 \pm 0.066}$ (sublinear)

Fit for $\gamma = 3/2$: $k_{\max} \sim N^{2.063 \pm 0.215}$ (superlinear)

---

[3] 95% confidence interval

Back to understanding the 80/20 rule:

Imagine a population of n people with variable x for individual wealth.

Define $N(x) = cx^{-\gamma}$ as the distribution of wealth x.

Must have $\int_{x_{\min}}^{\infty} dx\, N(x) = n$.

Total wealth W in the system:
$W = \int_{x_{\min}}^{\infty} dx\; xN(x)$.

Imagine that the bottom $100\,\theta_{\text{pop}}$ percent of the population holds $100\,\theta_{\text{wealth}}$ percent of the wealth.

Find γ depends on θ_{pop} and θ_{wealth} as

$$\gamma = 1 + \frac{\ln \frac{1}{(1-\theta_{\text{pop}})}}{\ln \frac{1}{(1-\theta_{\text{pop}})} - \ln \frac{1}{(1-\theta_{\text{wealth}})}}. \qquad (1)$$

Pleasant detail: $x_{\min}$ does not matter.
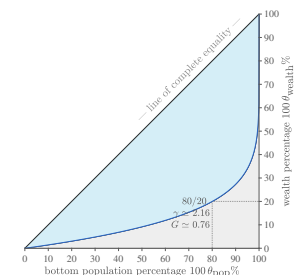
Insert assignment question ⬀

The PoCSverse
Power-Law Size
Distributions
45 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

$\gamma = 5/2$, maxima of N samples, $n = 1000$ sets of samples:

$\gamma = 3/2$, maxima of N samples, $n = 1000$ sets of samples:

The PoCSverse
Power-Law Size
Distributions
39 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

The PoCSverse
Power-Law Size
Distributions
42 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

80/20, γ, and the Gini coefficient G:

Gini coefficient ⬀:
Ratio of blue shape's area to triangle's area.
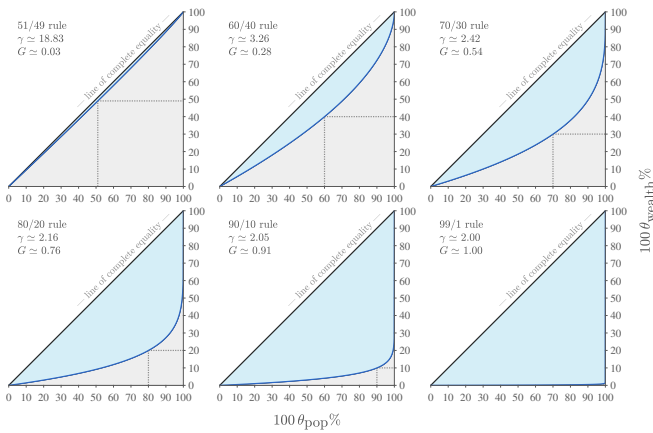$0 \leq G \leq 1$
Blue line is the "Lorenz curve."

The PoCSverse
Power-Law Size
Distributions
46 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

$\gamma = 5/2$, maxima of N samples, $n = 1000$ sets of samples:

$\gamma = 3/2$, maxima of N samples, $n = 1000$ sets of samples:

The PoCSverse
Power-Law Size
Distributions
40 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

The PoCSverse
Power-Law Size
Distributions
43 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

The top 1% owns 52.3%, the top 0.1% 38.4%, the top 0.01% 27.9%, the top 10^{-7}% 5.6%, ...

The PoCSverse
Power-Law Size
Distributions
50 of 71

Our Intuition

Definition

Examples

Wild vs. Mild

CCDFs

Size rankings and
Zipf's law

Size ranking ⇔
CCDF

References

51/49 rule
γ ≃ 18.83
G ≃ 0.03

60/40 rule
γ ≃ 3.26
G ≃ 0.28

70/30 rule
γ ≃ 2.42
G ≃ 0.54

80/20 rule
γ ≃ 2.16
G ≃ 0.76

90/10 rule
γ ≃ 2.05
G ≃ 0.91

99/1 rule
γ ≃ 2.00
G ≃ 1.00

99/1 rule:

$\gamma \simeq 2.002.$

$100\,\theta_{\text{pop}}$	$100\,\theta_{\text{wealth}}$	$100(1 - \theta_{\text{pop}})$	$100(1 - \theta_{\text{wealth}})$
20	0.05	80	99.95
50	0.15	50	99.85
80	0.35	20	99.65
$100 - 10^1$	0.50	10^1	99.50
99	**1**	**1**	**99**
$100 - 10^{-1}$	1.50	10^{-1}	98.50
$100 - 10^{-2}$	1.99	10^{-2}	98.01
$100 - 10^{-3}$	2.48	10^{-3}	97.52
$100 - 10^{-4}$	2.97	10^{-4}	97.03
$100 - 10^{-5}$	3.46	10^{-5}	96.54
$100 - 10^{-6}$	3.94	10^{-6}	96.06
$100 - 10^{-7}$	4.42	10^{-7}	95.58
$100 - 10^{-8}$	4.90	10^{-8}	95.10
$100 - 10^{-9}$	5.38	10^{-9}	94.62
$100 - 10^{-10}$	5.85	10^{-10}	94.15
$100 - 10^{-11}$	6.32	10^{-11}	93.68
$100 - 10^{-12}$	6.79	10^{-12}	93.21
$100 - 10^{-13}$	7.26	10^{-13}	92.74

Complementary Cumulative Distribution Function:

CCDF:

$$P_{\geq}(x) \propto x^{-(\gamma - 1)}$$
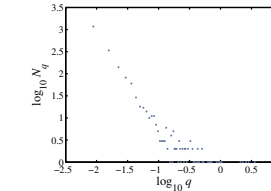
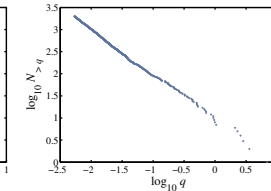- Use when tail of P follows a power law.
- Increases exponent by one.
- Useful in cleaning up data.

PDF:

CCDF:

The PoCSverse
Power-Law Size
Distributions
53 of 71

Our Intuition

Definition

Examples

Wild vs. Mild

CCDFs

Size rankings and
Zipf's law

Size ranking ⇔
CCDF

References

The PoCSverse
Power-Law Size
Distributions
48 of 71

Our Intuition

Definition

Examples

Wild vs. Mild

CCDFs

Size rankings and
Zipf's law

Size ranking ⇔
CCDF

References

The 51/49 rule:

$\gamma \simeq 18.8.$

$100\,\theta_{\text{pop}}$	$100\,\theta_{\text{wealth}}$	$100(1 - \theta_{\text{pop}})$	$100(1 - \theta_{\text{wealth}})$
20	18.99	80	81.01
51	**49**	**49**	**51**
80	78.11	20	21.89
90	88.62	10	11.38
99	98.71	1	1.29
$100 - 10^{-1}$	99.85	10^{-1}	0.15
$100 - 10^{-2}$	99.98	10^{-2}	0.02
$100 - 10^{-3}$	100.00	10^{-3}	0.00

Gini coefficent:

$$G = \begin{cases} 1 & \text{if } 1 < \gamma \leq 2, \\ \dfrac{1}{1 + 2(\gamma - 2)} & \text{if } \gamma > 2. \end{cases} \qquad (2)$$

Insert assignment question ↗

Complementary Cumulative Distribution Function:

- Same story for a discrete variable: $P(k) \sim ck^{-\gamma}$.

$$P_{\geq}(k) = P(k' \geq k)$$

$$= \sum_{k'=k}^{\infty} P(k)$$

$$\propto k^{-(\gamma - 1)}$$

- Use integrals to approximate sums.

The PoCSverse
Power-Law Size
Distributions
51 of 71

Our Intuition

Definition

Examples

Wild vs. Mild

CCDFs

Size rankings and
Zipf's law

Size ranking ⇔
CCDF

References

The PoCSverse
Power-Law Size
Distributions
54 of 71

Our Intuition

Definition

Examples

Wild vs. Mild

CCDFs

Size rankings and
Zipf's law

Size ranking ⇔
CCDF

References

The PoCSverse
Power-Law Size
Distributions
49 of 71

Our Intuition

Definition

Examples

Wild vs. Mild

CCDFs

Size rankings and
Zipf's law
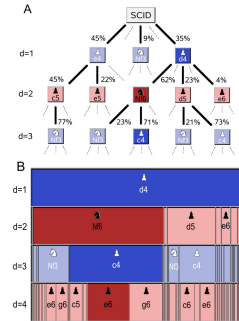
Size ranking ⇔
CCDF

References

80/20 rule:

$100\,\theta_{\text{pop}}$	$100\,\theta_{\text{wealth}}$	$100(1 - \theta_{\text{pop}})$	$100(1 - \theta_{\text{wealth}})$
20	3.05	80	96.95
50	9.16	50	90.84
80	**20**	**20**	**80**
90	27.33	10	72.67
99	47.19	1	52.81
$100 - 10^{-1}$	61.62	10^{-1}	38.38
$100 - 10^{-2}$	72.11	10^{-2}	27.89
$100 - 10^{-3}$	79.73	10^{-3}	20.27
$100 - 10^{-4}$	85.27	10^{-4}	14.73
$100 - 10^{-5}$	89.30	10^{-5}	10.70
$100 - 10^{-6}$	92.22	10^{-6}	7.78
$100 - 10^{-7}$	94.35	10^{-7}	5.65
$100 - 10^{-8}$	95.89	10^{-8}	4.11
$100 - 10^{-9}$	97.02	10^{-9}	2.98
$100 - 10^{-10}$	97.83	10^{-10}	2.17
$100 - 10^{-11}$	98.42	10^{-11}	1.58
$100 - 10^{-12}$	98.85	10^{-12}	1.15
$100 - 10^{-13}$	99.17	10^{-13}	0.83

$\gamma \simeq 2.16.$

Complementary Cumulative Distribution Function:

CCDF:

$$P_{\geq}(x) = P(x' \geq x) = 1 - P(x' < x)$$

$$= \int_{x'=x}^{\infty} P(x')\,\mathrm{d}x'$$

$$\propto \int_{x'=x}^{\infty} (x')^{-\gamma}\,\mathrm{d}x'$$

$$= \frac{1}{-\gamma + 1}(x')^{-\gamma+1}\Big|_{x'=x}^{\infty}$$

$$\propto x^{-(\gamma - 1)}$$

"Zipfian" size-rank plots

George Kingsley Zipf:

- Noted various rank distributions have power-law tails, often with exponent -1 (word frequency, city sizes, ...)
- Zipf's 1949 Magnum Opus ↗:

  "Human Behaviour and the Principle of Least-Effort" ⓐ ↗ by G. K. Zipf (1949). [20]

- We'll study Zipf's law in depth ...

The PoCSverse
Power-Law Size
Distributions
52 of 71

Our Intuition

Definition

Examples

Wild vs. Mild

CCDFs

Size rankings and
Zipf's law

Size ranking ⇔
CCDF

References

The PoCSverse
Power-Law Size
Distributions
55 of 71

Our Intuition

Definition

Examples

Wild vs. Mild

CCDFs

Size rankings and
Zipf's law

Size ranking ⇔
CCDF

References

"Zipfian" size-rank plots

The PoCSverse
Power-Law Size
Distributions
56 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Zipf's way:

- Given a collection of entities, rank them by size, largest to smallest.
- S_r = the size of the rth ranked entity.
- General term: "Size ranking"
- $r = 1$ corresponds to the largest size.
- Example: S_1 could be the frequency of occurrence of the most common word in a text.
- Zipf's observation:

$$S_r \propto r^{-\alpha}$$

with α often close to 1.

Misrankings

The PoCSverse
Power-Law Size
Distributions
57 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

The "biggest" thing is rank #1, otherwise:

- "USA #195!"[4]
- "USA #195!"
- "USA #195!"
- "USA #195!"

More:

- Size distribution connects with '#1-is-biggest' 'size' ranking only
- Main form of ranking by decreasing 'size' is robust to low sampling of small 'size' entities (the tail 'fills in').

---

[4]As of August 2024 ↗. Not simple agreed upon by all.

Size ranking example:

The PoCSverse
Power-Law Size
Distributions
58 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References
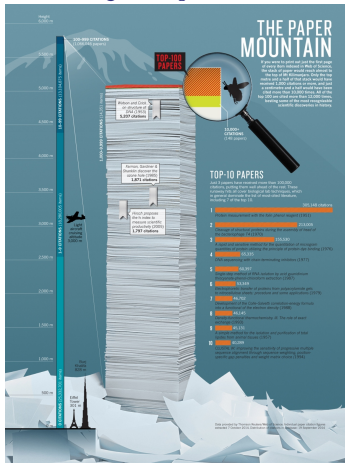
Nature (2014):
Most cited papers of all time ↗

Incel typology:

The PoCSverse
Power-Law Size
Distributions
59 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

"The incel lexicon: Deciphering the emergent cryptolect of a global misogynistic community" ↗
Gothard et al.,
, 2021. [7]

"Zipf's Law in the Popularity Distribution of Chess Openings" ↗
Blasius and Tönjes,
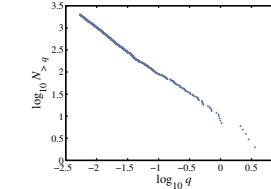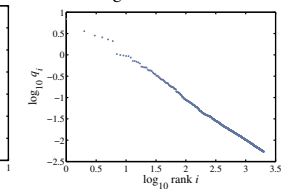Phys. Rev. Lett., **103**, 218701, 2009. [3]

- Examined all games of varying game depth d in a set of chess databases.
- n = popularity = how many times a specific game path appears in databases.
- $S(n; d)$ = number of depth d games with popularity n.
- Show "the frequencies of opening moves are distributed according to a power law with an exponent that increases linearly with the game depth, whereas the pooled distribution of all opening weights follows Zipf's law with universal exponent."
- Propose hierarchical fragmentation model that produces self-similar game trees.

The PoCSverse
Power-Law Size
Distributions
58 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

FIG. 1 (color online). (a) Schematic representation of the weighted game tree of chess based on the SCIDBASE [6] for the first three half moves. Each node indicates a state of the game. Possible game continuations are shown as solid lines together with the branching ratios r_d. Dotted lines symbolize other game continuations, which are not shown. (b) Alternative representation emphasizing the successive segmentation of the set of games, here indicated for games following a 1.d4 opening until the fourth half move $d = 4$. Each node σ is represented by a box of a size proportional to its frequency n_σ. In the subsequent half move these boxes split into subsets (indicated vertically below) according to the possible game continuations. Highlighted in (a) and (b) is a popular opening sequence 1.d4 Nf6 2.c4 e6 (Indian defense).

The PoCSverse
Power-Law Size
Distributions
61 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References
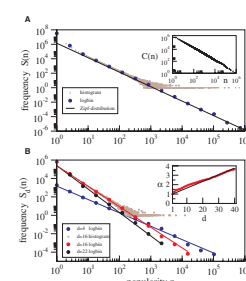
FIG. 2 (color online). (a) Histogram of weight frequencies $S(n)$ of openings up to $d = 40$ in the Scid database and with logarithmic binning. A straight line fit (not shown) yields an exponent of $\alpha = 2.05$ with a goodness of fit $R^2 > 0.9992$. For comparison, the Zipf distribution Eq. (8) with $\mu = 1$ is indicated as a solid line. Inset: number $C(n) = \sum_{m=n}^{\infty} S(m)$ of openings with a popularity $m > n$. $C(n)$ follows a power law with exponent $\alpha = 1.04$ ($R^2 = 0.994$). (b) Number $S_d(n)$ of openings of depth d with a given popularity n for $d = 16$ and histograms with logarithmic binning for $d = 4$, $d = 16$, and $d = 22$. Solid lines are regression lines to the logarithmically binned data ($R^2 > 0.99$ for $d < 35$). Inset: slope α_d of the regression line as a function of d and the analytical estimation Eq. (6) using $N = 1.4 \times 10^6$ and $\beta = 0$ (solid line).

Size distributions:

The PoCSverse
Power-Law Size
Distributions
62 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

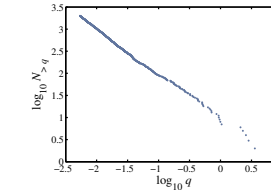### Brown Corpus (1,015,945 words):

CCDF:

Size ranking:

- The, of, and, to, a, ...= 'objects'
- 'Size' = word frequency
- Beep: (Important) CCDF and size-ranking plots are related ...

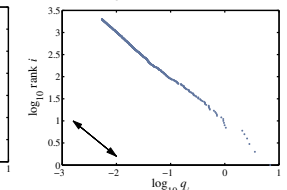---

Size distributions:

The PoCSverse
Power-Law Size
Distributions
63 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

Brown Corpus (1,015,945 words):

CCDF:

Size ranking:

- The, of, and, to, a, ...= 'objects'
- 'Size' = word frequency
- Beep: (Important) CCDF and size-ranking plots are related ...

Observe:

The PoCSverse
Power-Law Size
Distributions
64 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and
Zipf's law
Size ranking ⇔
CCDF
References

- $N P_{\geq}(x)$ = the number of objects with size at least x where N = total number of objects.
- If an object has size x_r, then $N P_{\geq}(x_r)$ is its rank r.
- So

$$x_r \propto r^{-\alpha} = (N P_{\geq}(x_r))^{-\alpha}$$

$$\propto x_r^{-(\gamma-1)(-\alpha)} \text{ since } P_{\geq}(x) \sim x^{-(\gamma-1)}.$$

We therefore have $1 = -(\gamma - 1)(-\alpha)$ or:

$$\boxed{\alpha = \frac{1}{\gamma - 1}}$$

- A rank distribution exponent of $\alpha = 1$ corresponds to a size distribution exponent $\gamma = 2$.

Nutshell for power-law size distributions and size-rank orderings:

- 🦠 Heavy-tailed distributions abound.
- 🦠 Some are power-law size distributions.
- 🦠 Continuous: $P(x) \sim x^{-\gamma}$, discrete: $P(k) \sim ck^{-\gamma}$
- 🦠 Mean 'blows up' with upper cutoff if $\gamma < 2$.
- 🦠 Mean depends on lower cutoff if $\gamma > 2$.
- 🦠 $\gamma < 2$: Typical sample is large.
- 🦠 $\gamma > 2$: Typical sample is small.
- 🦠 Complementary Cumulative Distribution Function (CCDF): $P(x) \propto x^{-(\gamma-1)}$ and $P_{\geq}(k) = k^{-(\gamma-1)}$
- 🦠 Size of largest sample from n samples grows as:

$$x_1 \gtrsim c'n^{1/(\gamma-1)}$$

More with the nutshelling:

- 🦠 Size rankings: Order types from "biggest" to "smallest" size S.
- 🦠 Widely observed: S_r is highly skewed.
- 🦠 When scaling is apparent:

$$S_r \propto r^{-\alpha}$$

- 🦠 Claim: α often close to 1. "Zipf's law":

$$S_r \propto r^{-1}.$$

- 🦠 Scalings of size distribution (γ) and size ranking (α) are connected:

$$\alpha = \frac{1}{\gamma - 1} \text{ and } \gamma = 1 + \frac{1}{\alpha}.$$

- 🦠 Danger Will Robinson point: $\gamma = 2 \Leftrightarrow \alpha = 1$.

The Don. 🔗

Extreme deviations in test cricket:

Histogram of Test career batting averages.
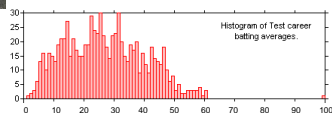
- 🦠 Don Bradman's batting average 🔗 = 166% next best.
- 🦠 That's pretty solid.
- 🦠 Later in the course: Understanding success— is the Mona Lisa like Don Bradman?

References I

[1] P. Bak, K. Christensen, L. Danon, and T. Scanlon. Unified scaling law for earthquakes. Phys. Rev. Lett., 88:178501, 2002. pdf 🔗

[2] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. Science, 286:509–511, 1999. pdf 🔗

[3] B. Blasius and R. Tönjes. Zipf's law in the popularity distribution of chess openings. Phys. Rev. Lett., 103:218701, 2009. pdf 🔗

[4] K. Christensen, L. Danon, T. Scanlon, and P. Bak. Unified scaling law for earthquakes. Proc. Natl. Acad. Sci., 99:2509–2513, 2002. pdf 🔗

References II

[5] A. Clauset, C. R. Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. SIAM Review, 51:661–703, 2009. pdf 🔗

[6] D. J. de Solla Price. Networks of scientific papers. Science, 149:510–515, 1965. pdf 🔗

[7] K. Gothard, D. R. Dewhurst, J. A. Minot, J. L. Adams, C. M. 5-Danforth, and P. S. Dodds. The incel lexicon: Deciphering the emergent cryptolect of a global misogynistic community, 2021. Available online at https://arxiv.org/abs/2105.12006. pdf 🔗

[8] P. Grassberger. Critical behaviour of the Drossel-Schwabl forest fire model. New Journal of Physics, 4:17.1–17.15, 2002. pdf 🔗

References III

[9] B. Gutenberg and C. F. Richter. Earthquake magnitude, intensity, energy, and acceleration. Bull. Seism. Soc. Am., 499:105–145, 1942. pdf 🔗

[10] J. Holtsmark. Über die verbreiterung von spektrallinien. Ann. Phys., 58:577–630, 1919. pdf 🔗

[11] R. Munroe. Thing Explainer: Complicated Stuff in Simple Words. Houghton Mifflin Harcourt, 2015.

[12] M. E. J. Newman. Power laws, Pareto distributions and Zipf's law. Contemporary Physics, 46:323–351, 2005. pdf 🔗

References IV

[13] M. I. Norton and D. Ariely. Building a better America—One wealth quintile at a time. Perspectives on Psychological Science, 6:9–12, 2011. pdf 🔗

[14] D. D. S. Price. A general theory of bibliometric and other cumulative advantage processes. Journal of the American Society for Information Science, pages 292–306, 1976. pdf 🔗

[15] L. F. Richardson. Variation of the frequency of fatal quarrels with magnitude. J. Amer. Stat. Assoc., 43:523–546, 1949.

[16] H. A. Simon. On a class of skew distribution functions. Biometrika, 42:425–440, 1955. pdf 🔗

References V

[17] N. N. Taleb. The Black Swan. Random House, New York, 2007.

[18] G. U. Yule. A mathematical theory of evolution, based on the conclusions of Dr J. C. Willis, F.R.S. Phil. Trans. B, 213:21–87, 1925. pdf 🔗

[19] Y.-X. Zhu, J. Huang, Z.-K. Zhang, Q.-M. Zhang, T. Zhou, and Y.-Y. Ahn. Geography and similarity of regional cuisines in China. PLoS ONE, 8:e79161, 2013. pdf 🔗

[20] G. K. Zipf. Human Behaviour and the Principle of Least-Effort. Addison-Wesley, Cambridge, MA, 1949.

The PoCSverse
Power-Law Size Distributions
65 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and Zipf's law
Size ranking ⇔ CCDF
References

The PoCSverse
Power-Law Size Distributions
68 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and Zipf's law
Size ranking ⇔ CCDF
References

The PoCSverse
Power-Law Size Distributions
71 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and Zipf's law
Size ranking ⇔ CCDF
References

The PoCSverse
Power-Law Size Distributions
66 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and Zipf's law
Size ranking ⇔ CCDF
References

The PoCSverse
Power-Law Size Distributions
69 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and Zipf's law
Size ranking ⇔ CCDF
References

The PoCSverse
Power-Law Size Distributions
72 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and Zipf's law
Size ranking ⇔ CCDF
References

The PoCSverse
Power-Law Size Distributions
67 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and Zipf's law
Size ranking ⇔ CCDF
References

The PoCSverse
Power-Law Size Distributions
70 of 71
Our Intuition
Definition
Examples
Wild vs. Mild
CCDFs
Size rankings and Zipf's law
Size ranking ⇔ CCDF
References