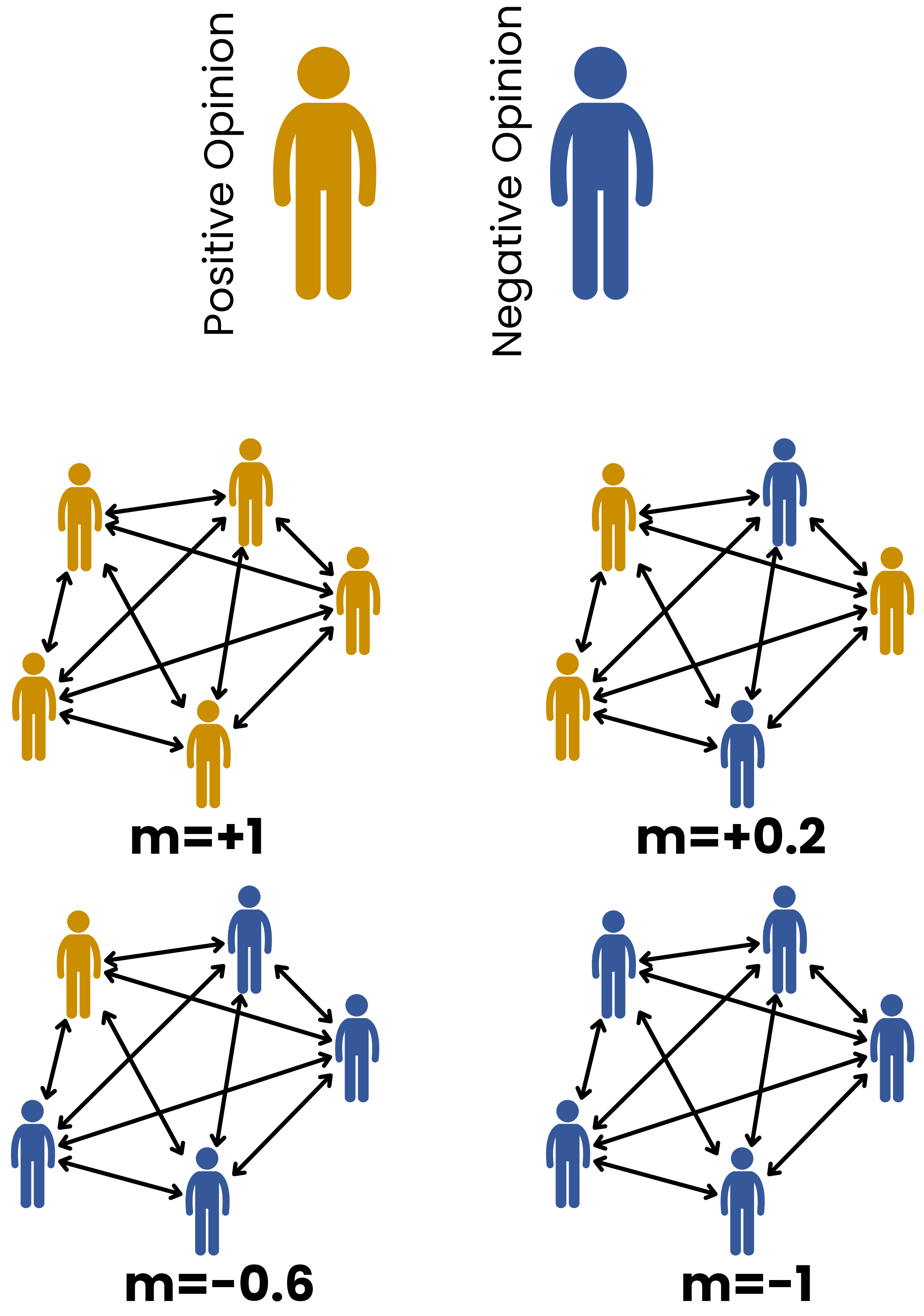
UNIVERSITÄT KONSTANZ

Power Laws

Computational Modelling of Social Systems

Giordano De Marzo
Max Pellert

Universität Konstanz



Recap

Opinion Dynamics

Study of how opinions are formed and evolve in a group of individuals

Voter Model

Model for binary (or discrete) opinions. Has a strong tendency toward consensus.

Bounded Confidence

Model with continuous opinions. Agents do not interact with very different peers.

Recommendation Algorithms

Opinion dynamics can be used to better understand the downsides of recommendation algorithms.

Outline

1. The Paretian World
2. Power Law Probability Distributions
3. Zipf’s Law and Heaps’ Law
4. Modelling the Cryptocurrency Market

The Paretian World

The Gaussian World

We are accustomed to think in term of gaussians and average values:

- height
- weight
- speed
- performances

In the Gaussian world there are no surprises:

- a small sample is enough for knowing everything
- the future is hardly surprising

The Paretian World

However many relevant phenomena are characterized by extreme events (Pareto distribution):

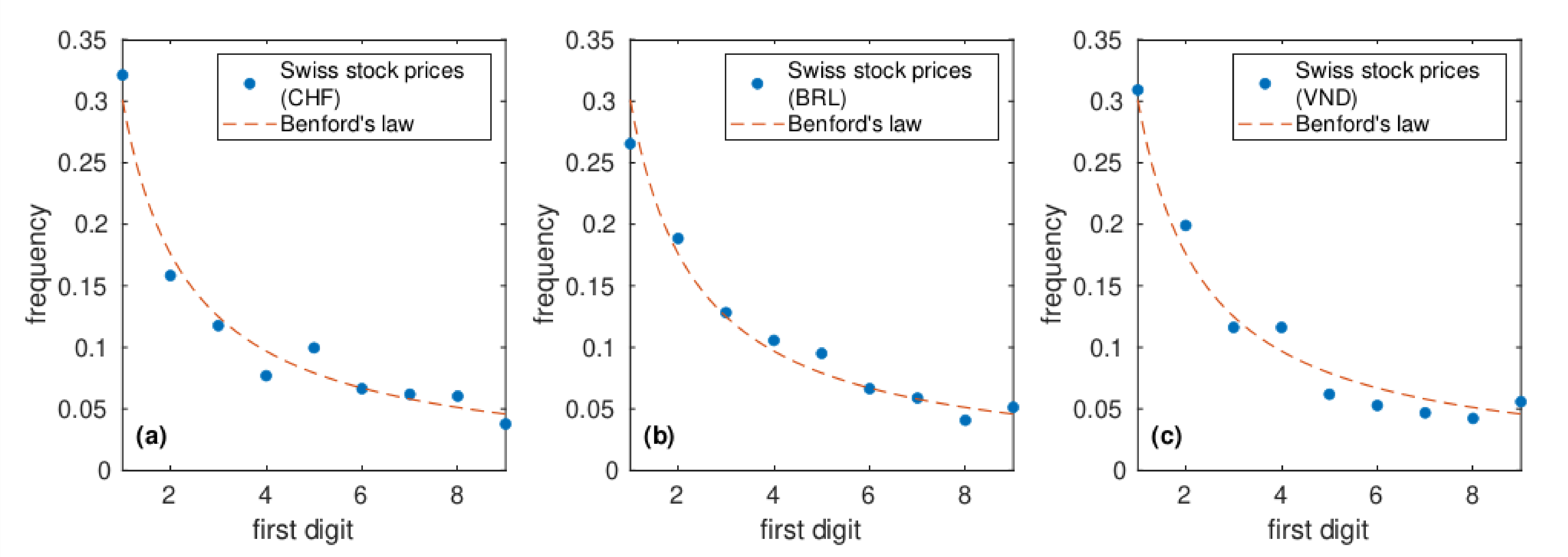
- financial crises
- wars
- pandemics
- natural disasters

The Paretian world is full of surprises and strange properties:

- a large sample is not enough for knowing everything
- the future is surprising

Benford’s Law

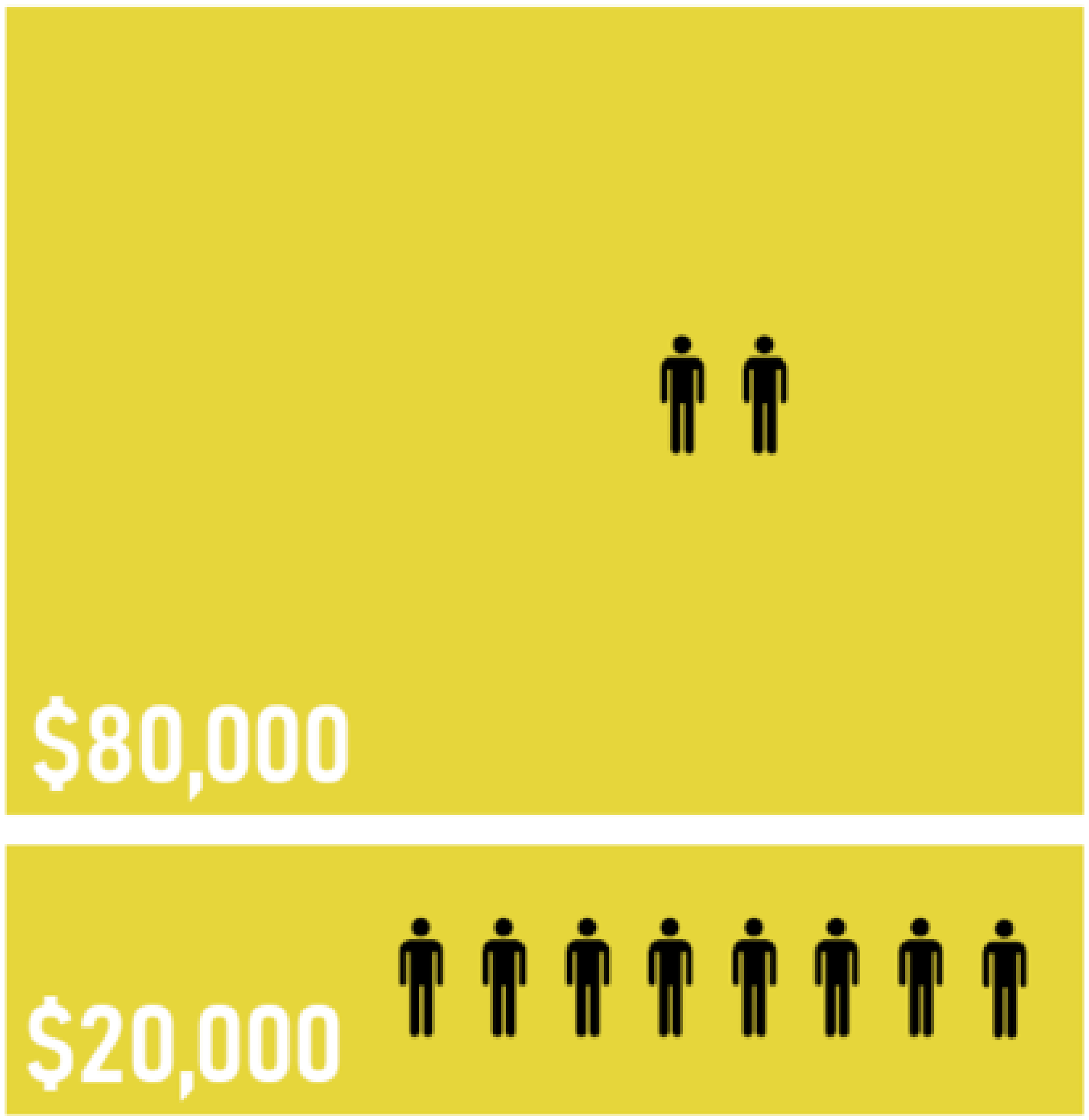
Consider the first digit of the Swiss stock market, one would expect all the digits be be equally common, but the 1 is much more common than the 9. This is the same if we change currency in which we express the prices. The same phenomenon is observed in many other datasets.



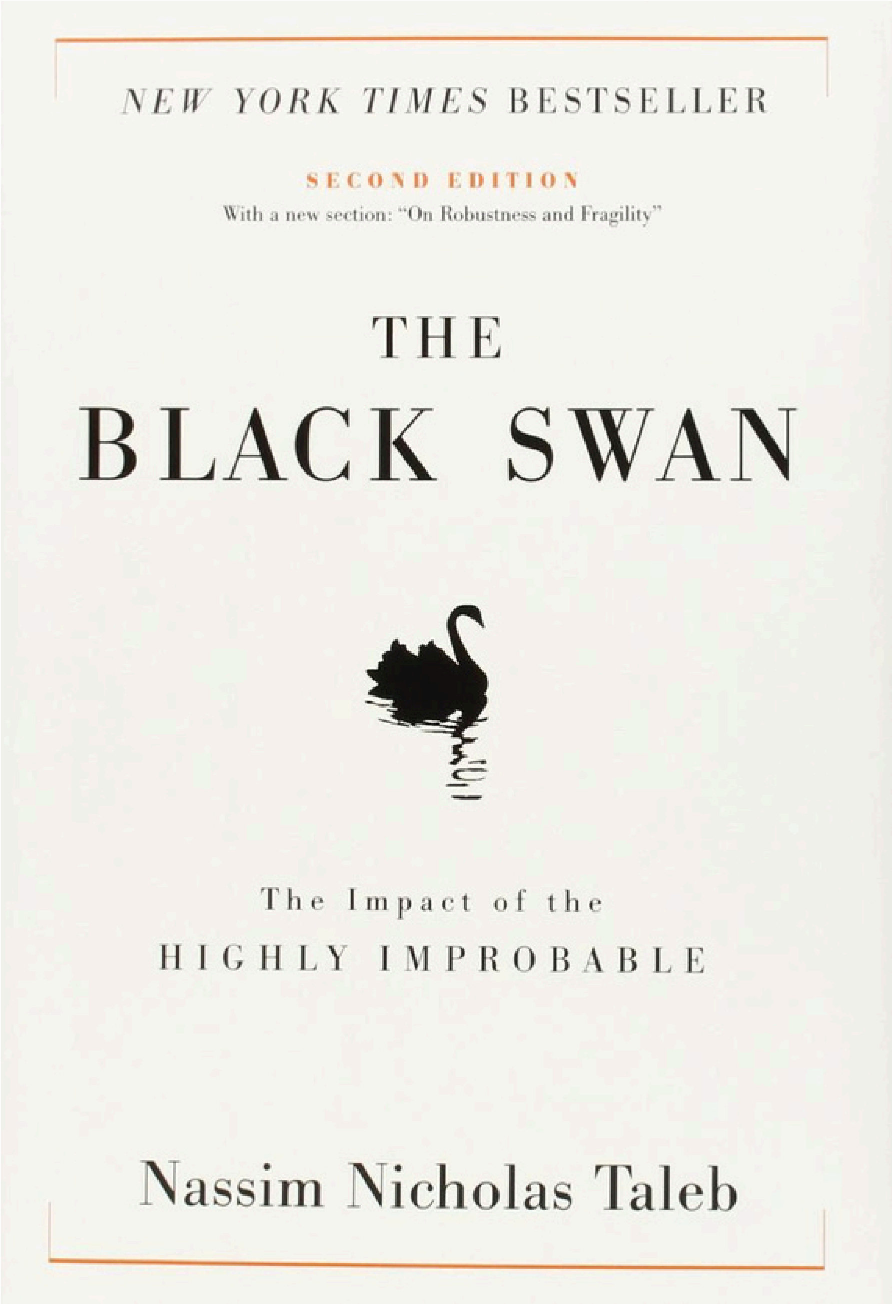
80-20 Rule

The 80-20 Rule or Pareto Principle states that around 80% of the resources are concentrated in 20% of the actors in the system:

- richest 20% people control around 80% of the total wealth
- around 20% of patience incur 80% of total healthcare expenses
- in computing fixing 20% top reported bugs eliminates 80% of errors and crashes



Black Swans



The Pareto world is characterized by extreme events called Blacks Swans. A Black Swan event:

- is unpredictable and unexpected
- has very relevant consequences, either positive or negative
- is often explained a posteriori

Examples of Black Swans are

- WWI (caused way more casualties than previous wars)
- 11/9 Terrorist attacks (caused way more casualties than previous attacks)
- Lionel Messi (scored way more goals in LaLiga than previous record holder)

Why are Power Laws Important?

Understanding Power Laws is crucial since they are the standard in the Paretian World

- Gaussian distributions can not be applied to the Paretian World
- Power Laws help us understand why extreme events occur
- They help us dealing with phenomena that are too large for our minds to fully capture
- They help us not underestimating small but non negligible probabilities

Using Gaussians instead of Power Laws may is very risky

- 2008 Financial Crisis
- Covid-19 Pandemics
- Ukraine War

Power Law Probability Distributions

Pareto or Power Law Distributions

Let us consider a series of events with sizes S_1, S_2 … etc.
We say that these event follow a Pareto or Power Law distribution if the probability P(S) of observing an event with size S is of the form

$$P(S) = \frac{c}{S^\alpha}$$

In this expression

- c is a normalization constant to ensure the probability to sum to one
- α is the power law exponent or scaling exponent

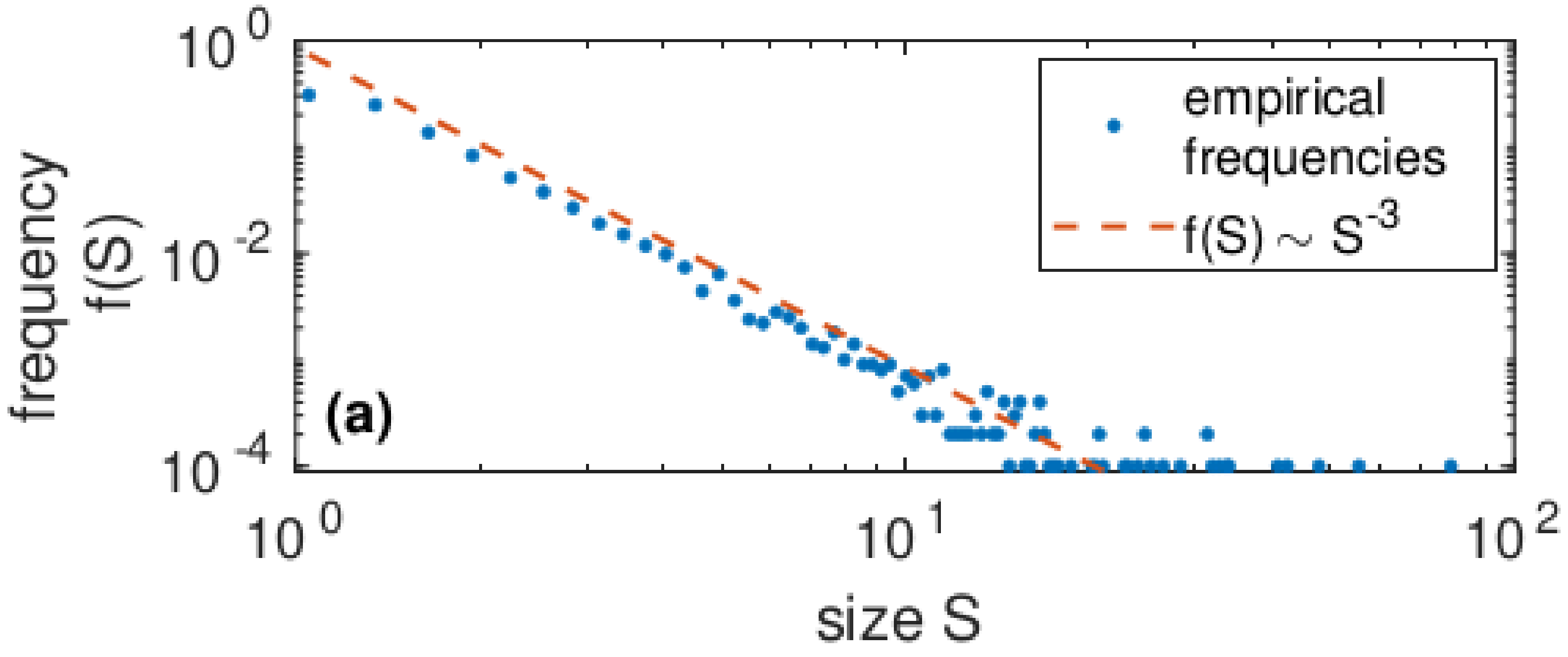
The power law shows a much slower decay with respect to a Gaussian

$$P(S) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

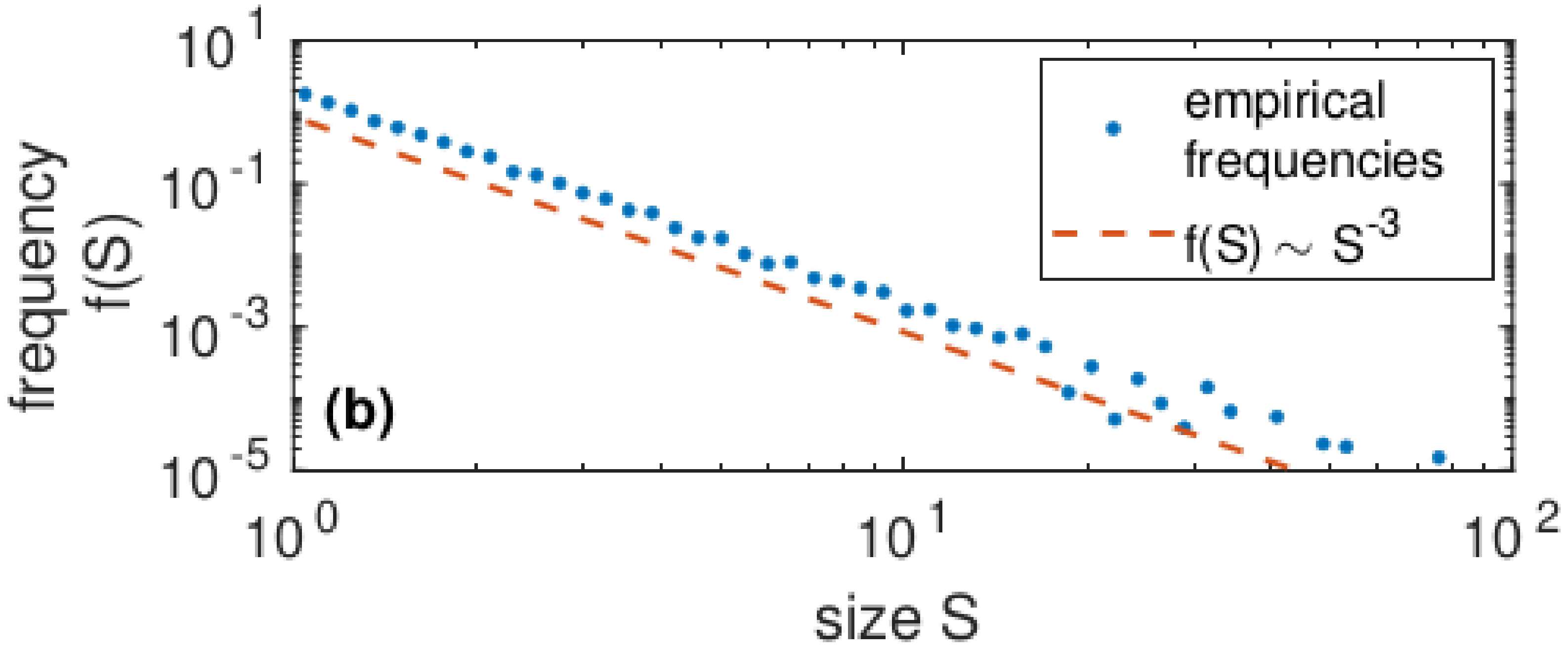
Visualizing Power Laws

Given a set of sizes we can determine their distribution performing an histogram. If the data follow a power law distribution the histogram will look like a straight line using a double logarithmic scale. In order to obtain better plots it is important to use logarithmic binning.

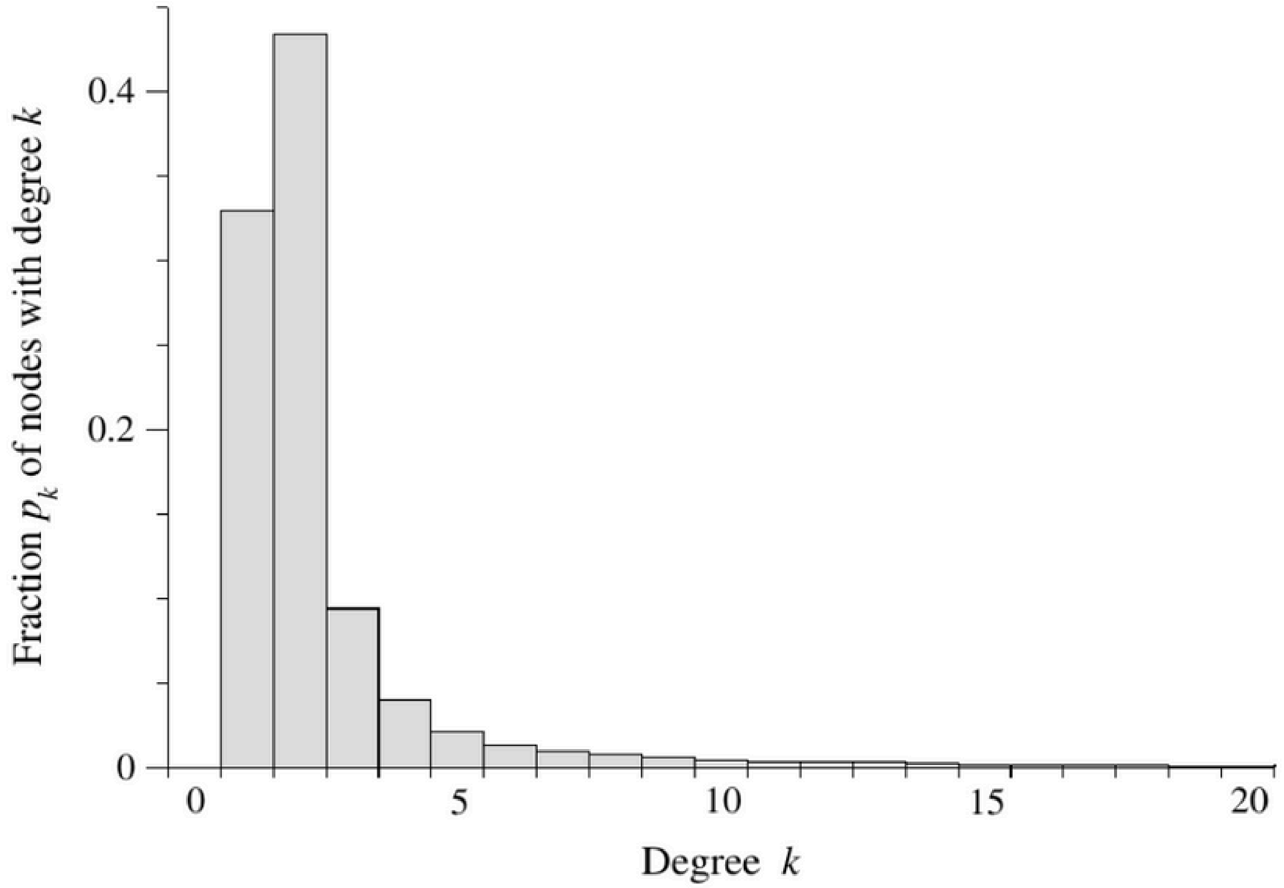
Normal Binning

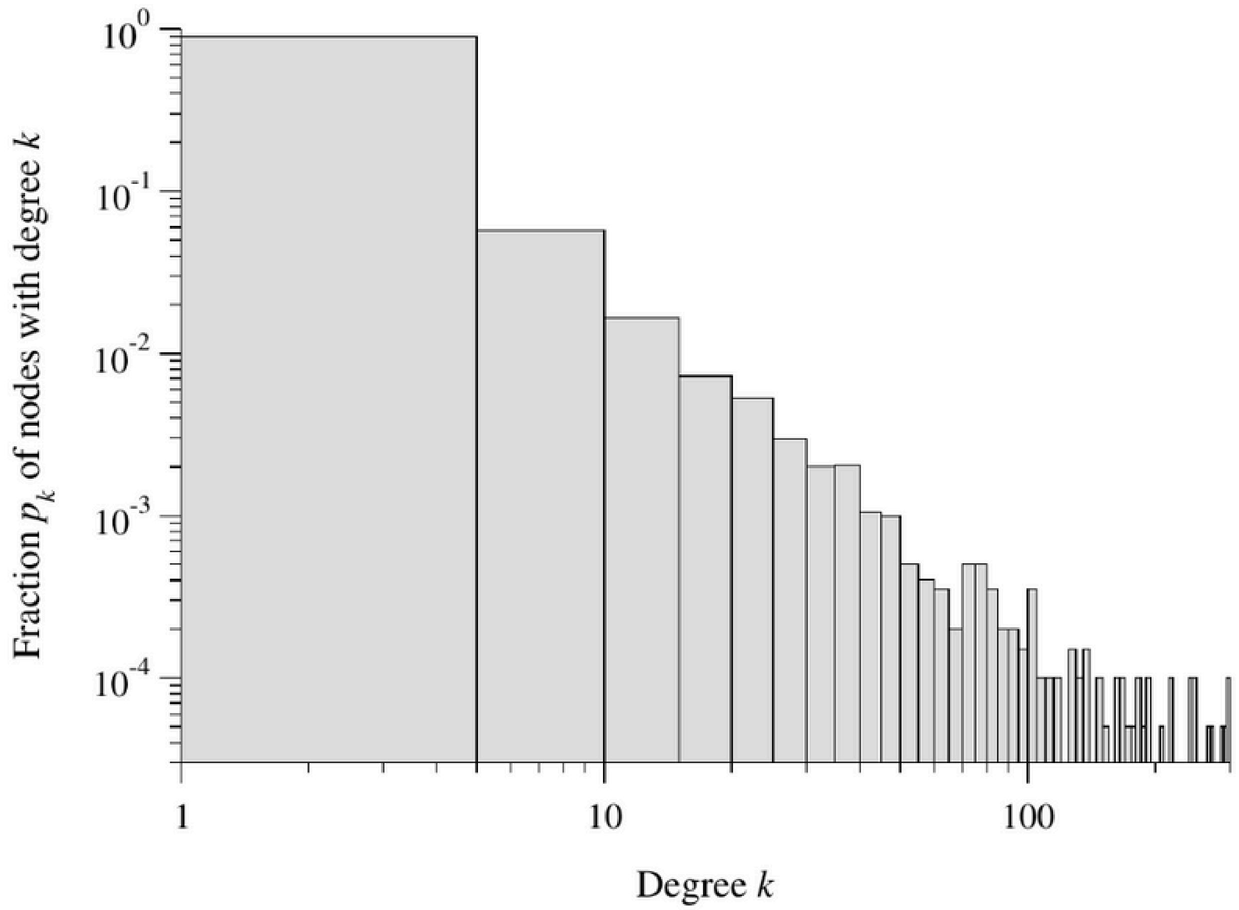


Logarithmic Binning



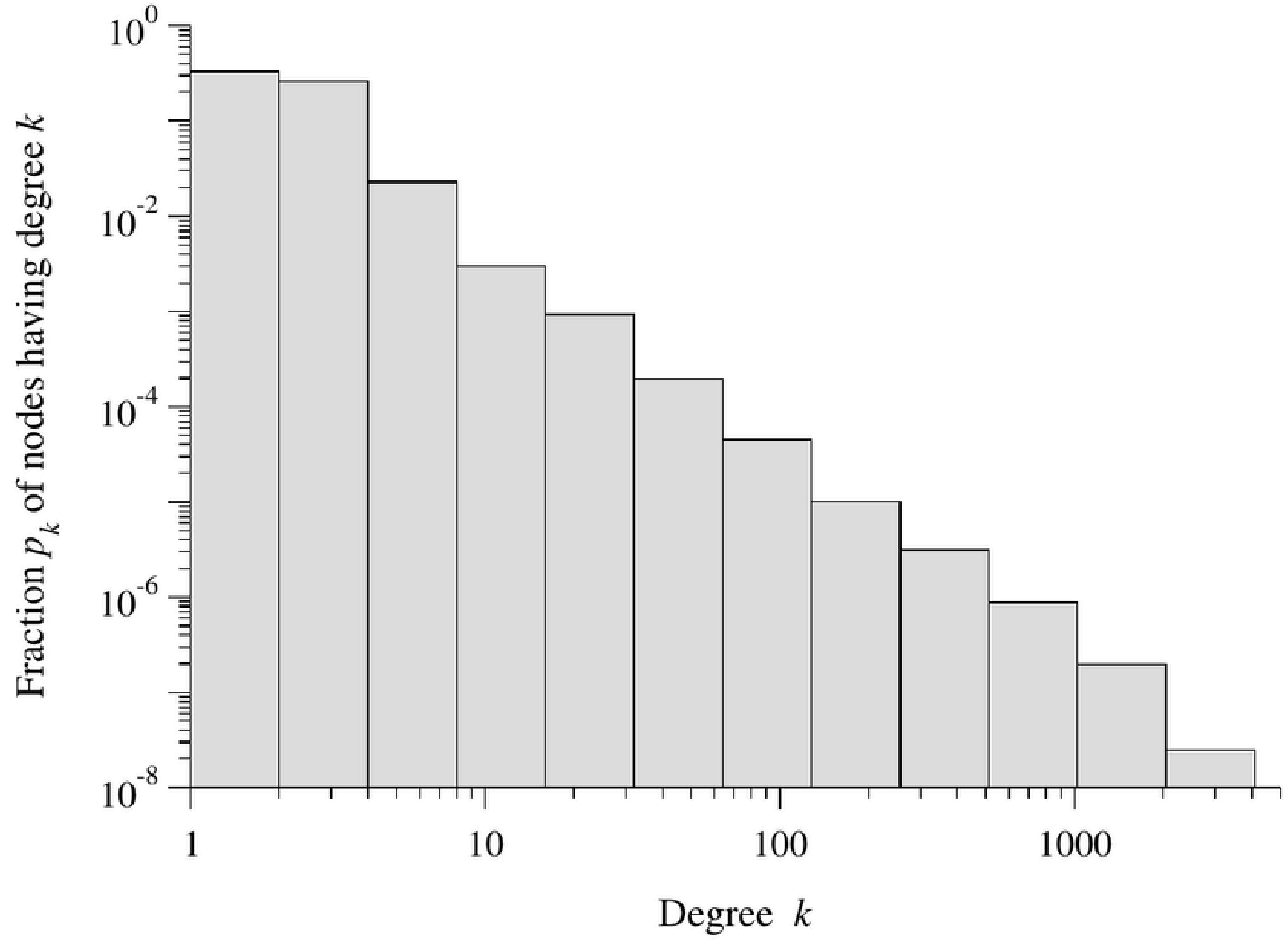
Linear **Binning**





Two histograms of the same distribution. The second one has log-transformed x and y axes and the same bins. Bins are all of the same width in linear scale, but appear different in log scale

Logarithmic Binning

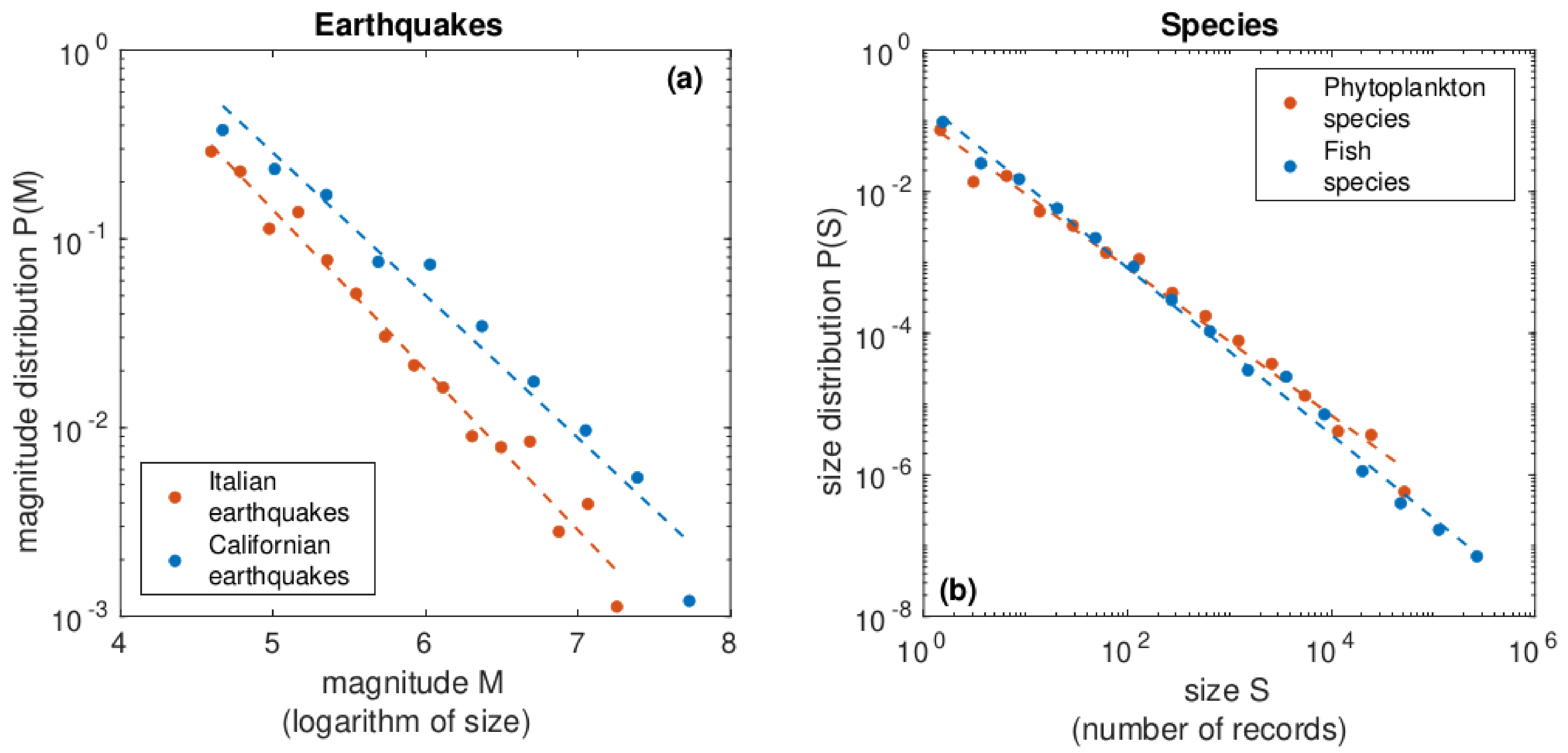


Same log-log histogram but with logarithmic binning: the width of a bin is a multiple of the one on the left. Bin heights are divided by their width. Bins now look all the same in log scale.

Power Laws in Natural Systems

Power Law probability distributions are observed in countless natural systems:

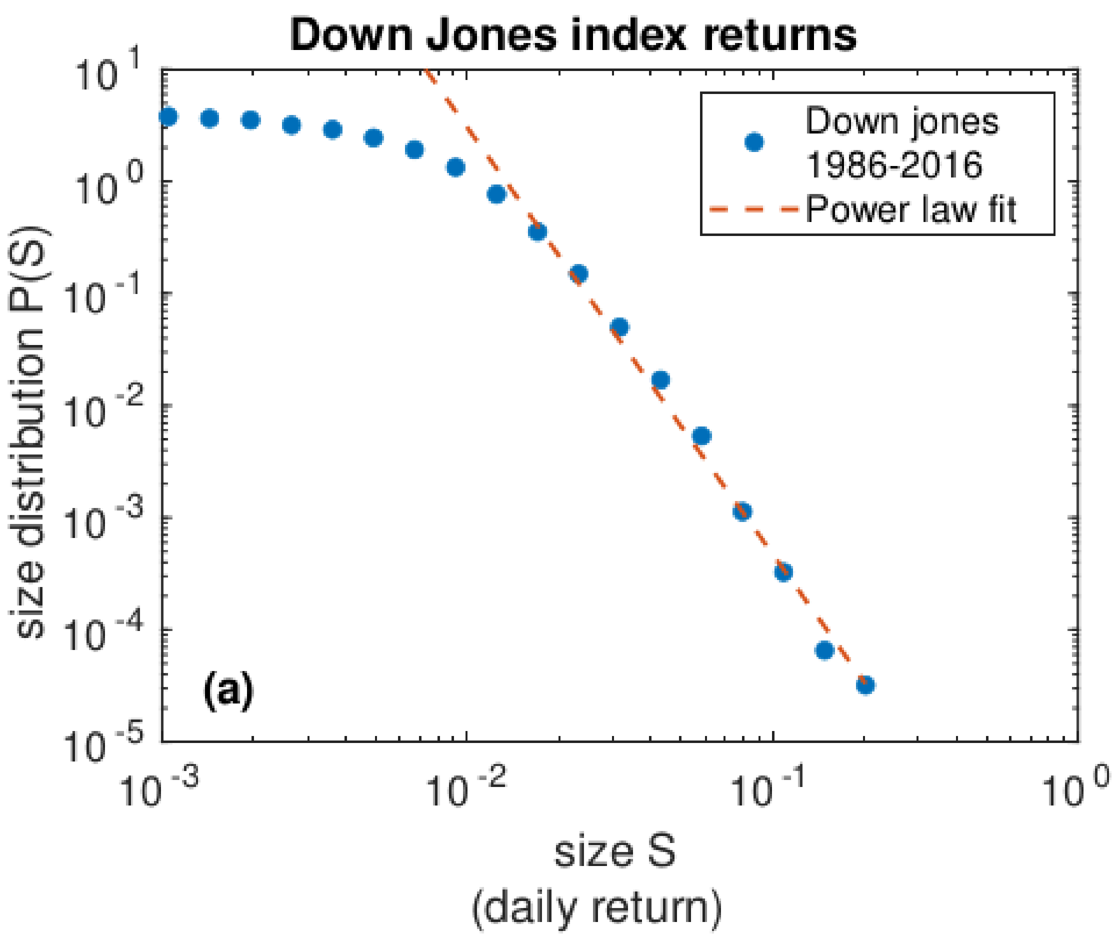
- earthquakes (Gutenberg–Richter law)
- species abundance
- pandemics
- solar flares

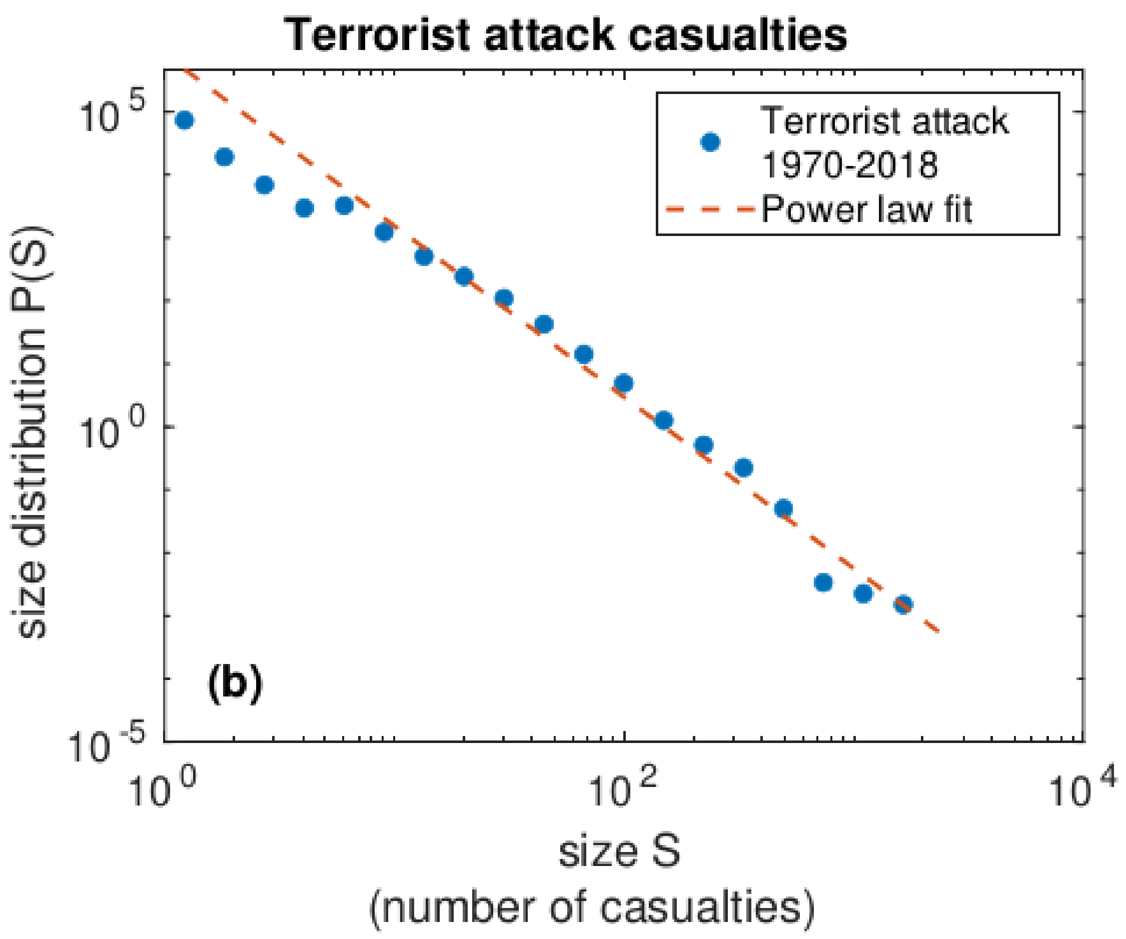


Power Laws in Social Systems

Also social systems are dominated by power laws:

- stocks returns
- wars and terrorist attacks
- people popularity in social systems
- wealth distribution





Scale-free Property

A Power Law probability distribution is of the form

$$P(S) = \frac{c}{S^{\alpha}}$$

As a consequence if we multiply all sizes by a constant factor K, the shape of the distribution does not change

$$P(K \cdot S) = K^{-\alpha} P(S)$$

For this reason we say that power laws are scale free. They have not a typical scale like a Gaussian.

Truncated Power Laws

Most real life systems have intrinsic lower and upper limits:

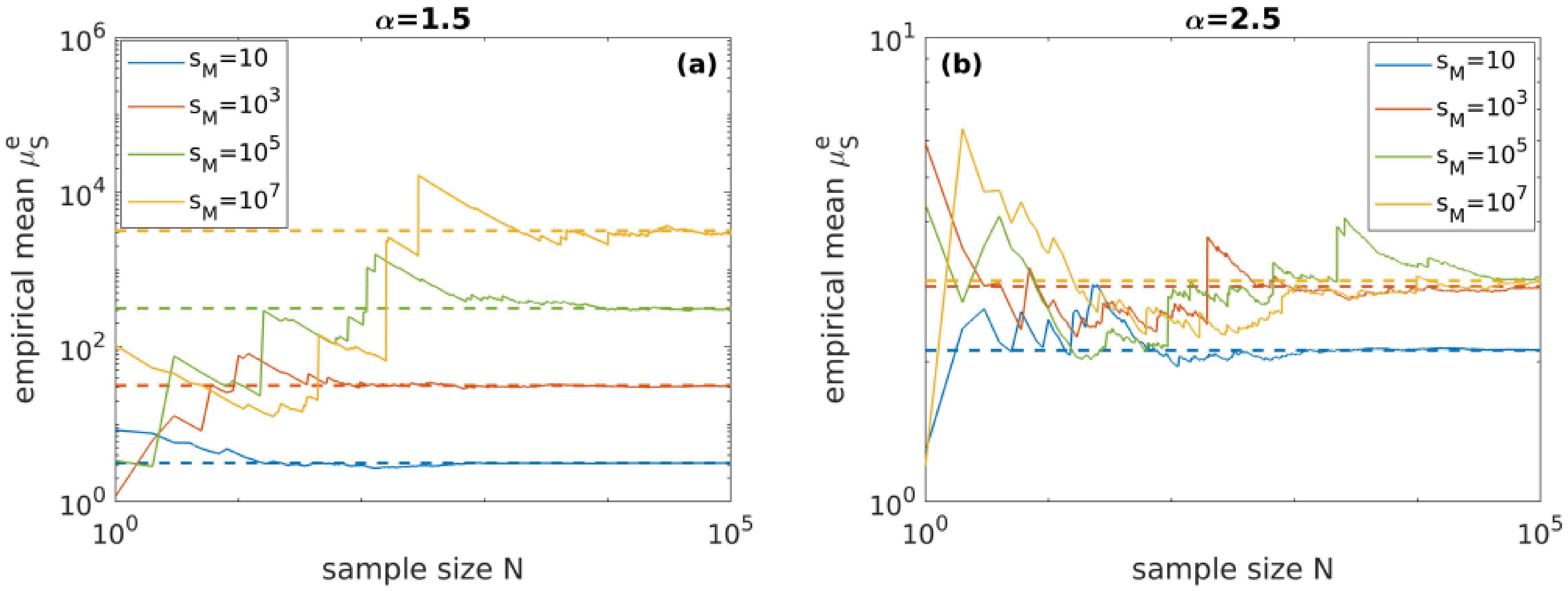
- a pandemic can't kill less than one person and more than the world population
- a building can't be less tall than a couple of meters
- an animal specie can't have less than an individual

For this reason all power law distributions have a lower cutoff s_{min} and may also have an upper cutoff s_{max}. The expression for the probability becomes

$$P(S) = \begin{cases} 0 \text{ if } S < s_{min} \\ \frac{c}{S^{\alpha}} \text{ if } s_{min} < S < s_{max} \\ 0 \text{ if } S > s_{max} \end{cases}$$

Diverging Moments

For α<2 the mean value of a power law is infinite, while for α<3 the variance is infinite. We can see this by computing the mean value as function of the sample size and gradually increasing the upper cutoff s_{max}. For α=1.5 the average keep growing every time we increase s_{max}.



Additive Growth

In additive growth processes the size of an object grows by the addition of single units to it. The object may also decrease in size, but just of a single unit.

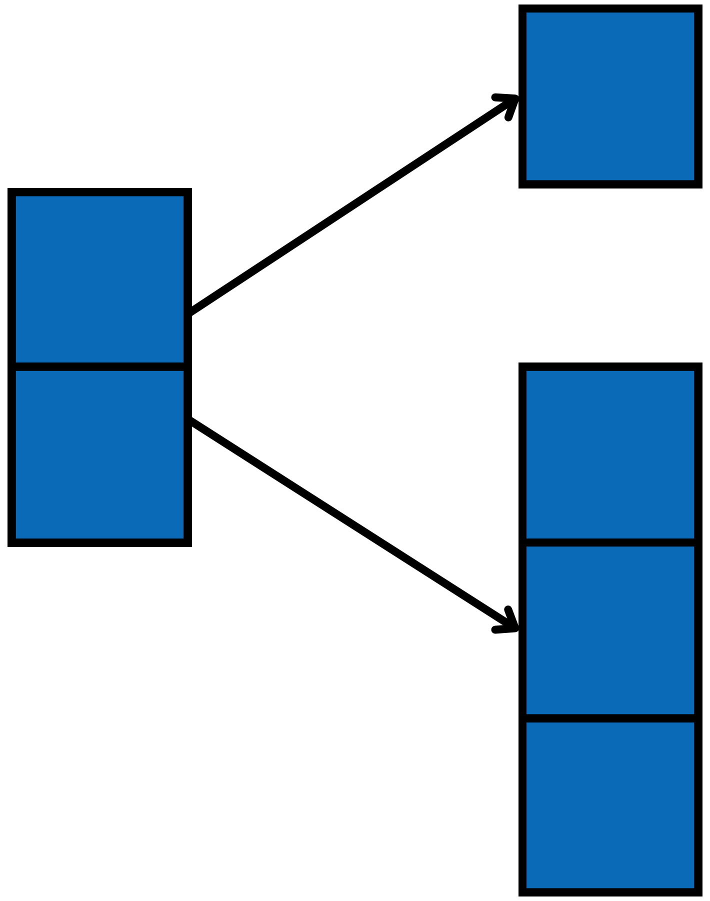
We can describe the process by a simple stochastic equation (random walk)

$$S(t+1) = S(t) + \eta \cdot r_t$$

Here

- S(t) is the size at time t
- η is a numerica constant
- r_t is a random variable

In our case $r_t = \pm 1$ with prob 0.5



Multiplicative Growth

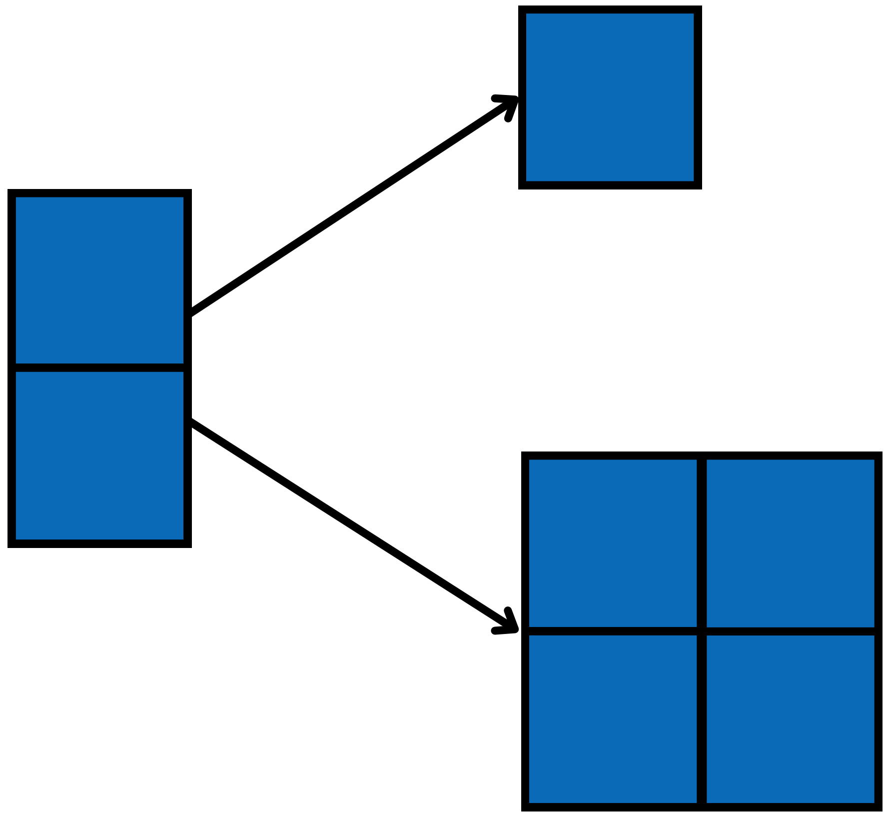
In multiplicative growth processes the size of an object grows (or decrease) in relative terms (percentage). It can either double or halve its size.

We can describe the process by a simple stochastic equation (logarithmic random walk)

$$S(t+1) = \eta \cdot r_t \cdot S(t)$$

If we take the logarithm of both sizes we get a standard random walk for the variable y=log(S)

$$y(t+1) = y(t) + \log(\eta \cdot r_t)$$

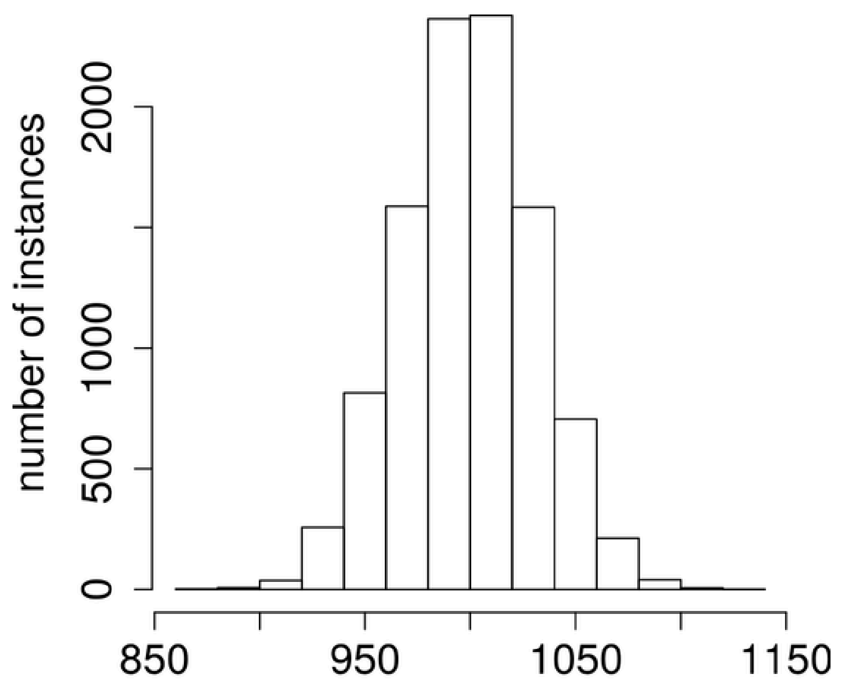


Growth Processes Distribution

While the additive growth process produces a Gaussian distribution of sizes, the multiplicative growth process generates a log-normal distribution. It is not a power law, but it is much more tailed than a Gaussian.

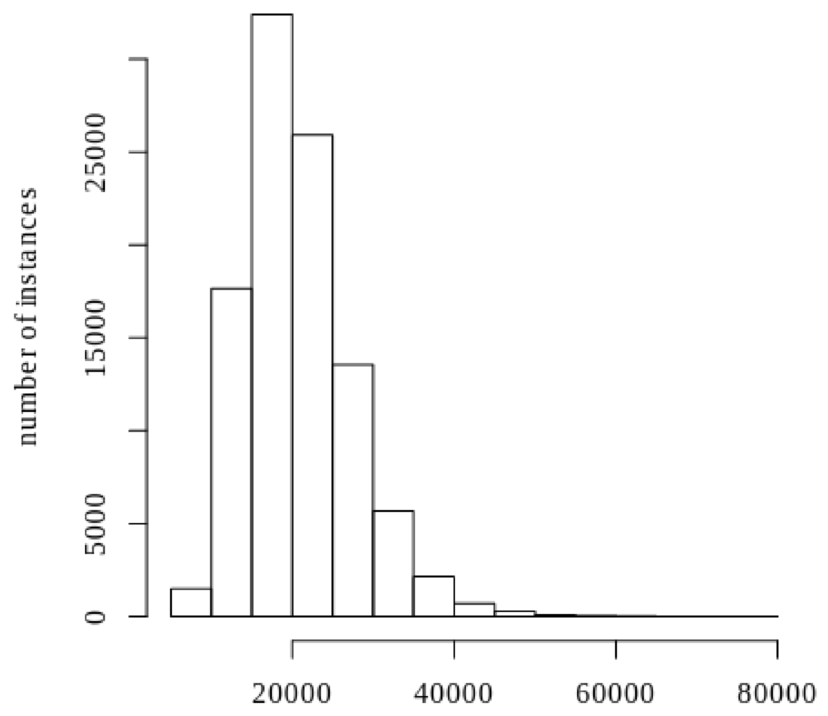
Gaussian (Normal)

$$P(S) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$



Lognormal

$$P(S) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left[-\frac{\log(x)^2}{2\sigma^2}\right]$$



From Lognormals to Power Laws

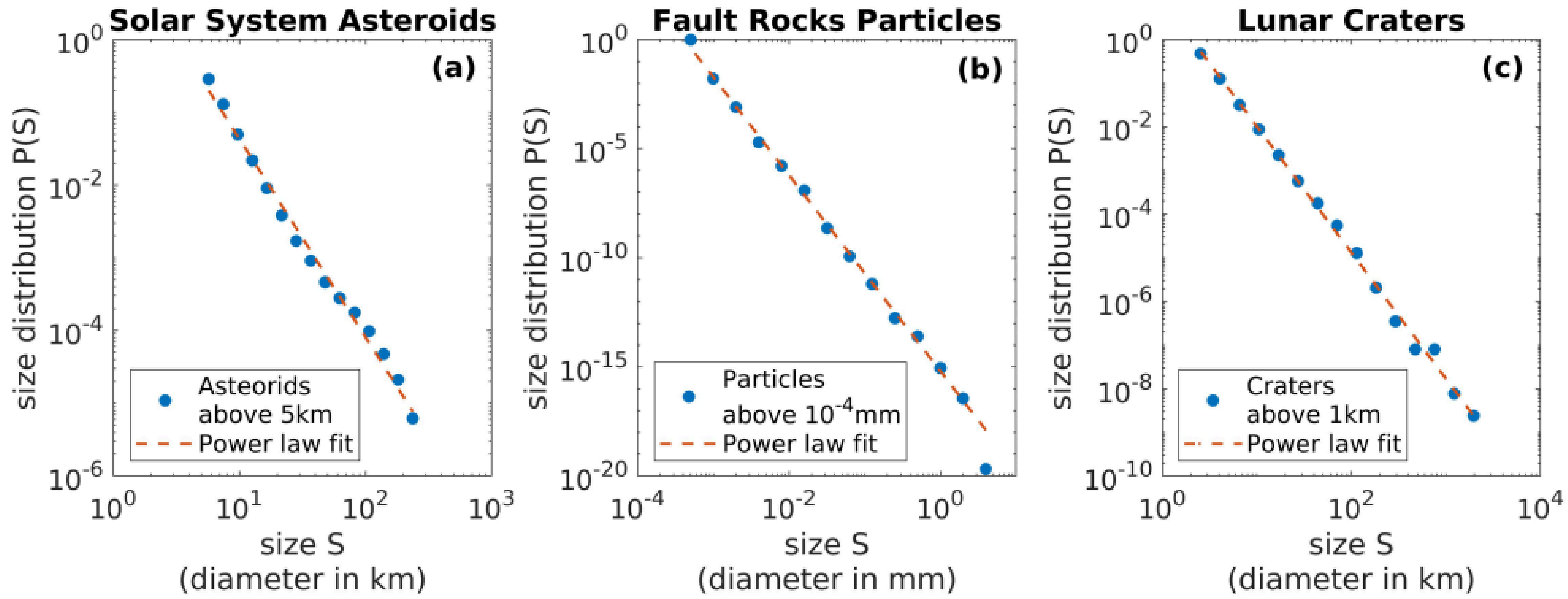
Even if the multiplicative growth produces lognormals, it is sufficient to slightly modify it to get power law distributions

- adding heterogeneous ages
  - we start with just one element in the system
  - we add new elements following a poisson process
- adding a reflecting barrier
  - we make elements never get smaller than a given minimal size (lower cutoff of the process)

Both these modifications result in an asymptotic power law distribution of elements sizes.

Fragmentation Processes

Multiplicative growth can also describe fragmentation processes. In this case there is no growth, but fragmentation, but the idea is similar. At each time step the size of the object is multiplied by a random number (smaller than one).



Zipf's Law and Heaps' Law

Zipf's Law

Power law distributions often manifest under different forms. Zipf's law is one of the most know of these

- we denote by S(k) the size of the k-th largest element
- S(1) is the size of the largest element
- k is called the rank

The systems shows Zipf's law if it holds

$$S(k) \approx \frac{S(1)}{k^{\gamma}}$$

Some people call Zipf's law only the special case $\gamma=1$ and generalized Zipf's law the other exponents.

Examples of Zipf's Law

Language

- The system is a long book
- The size S of a world is its frequency in the book
- S(1) is the frequency of the most common word

Ex: English

1. "The" S(1)=6.5%
2. "Of" S(2)=2.8%
3. "To" S(2)=2.4%

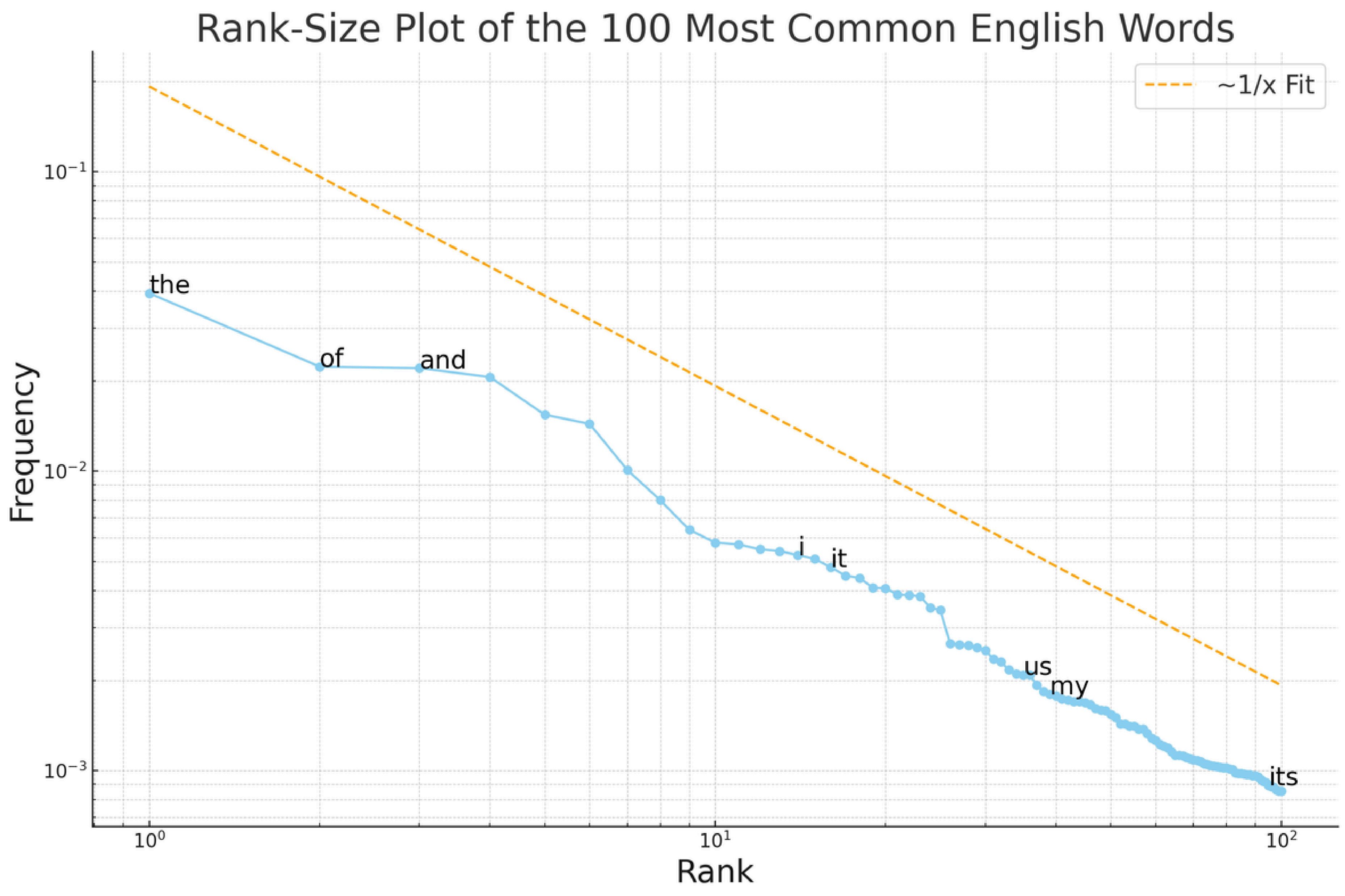
Cities

- The system is country
- The size S of a city is its population
- S(1) is the population of the largest city

Ex: Italy

1. "Roma" S(1)=2.6M
2. "Milano" S(2)=1.2M
3. "Napoli" S(2)=0.9M

The Rank-Size Plot



Zipf's law is typically visualized using the rank size plot:

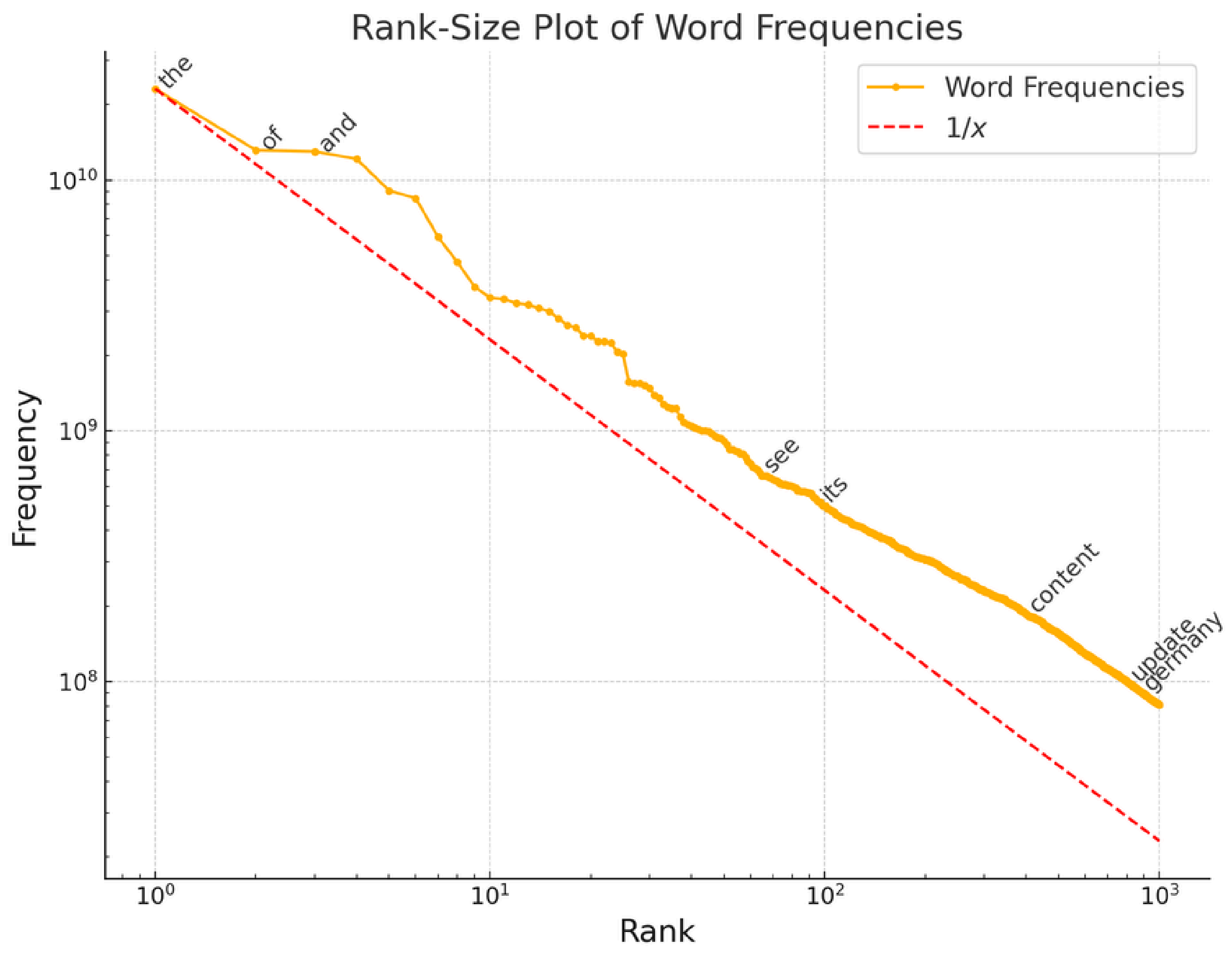
- first order the sizes from the largest to the smallest
- then plot the ordered sizes vs their rank in the sequence

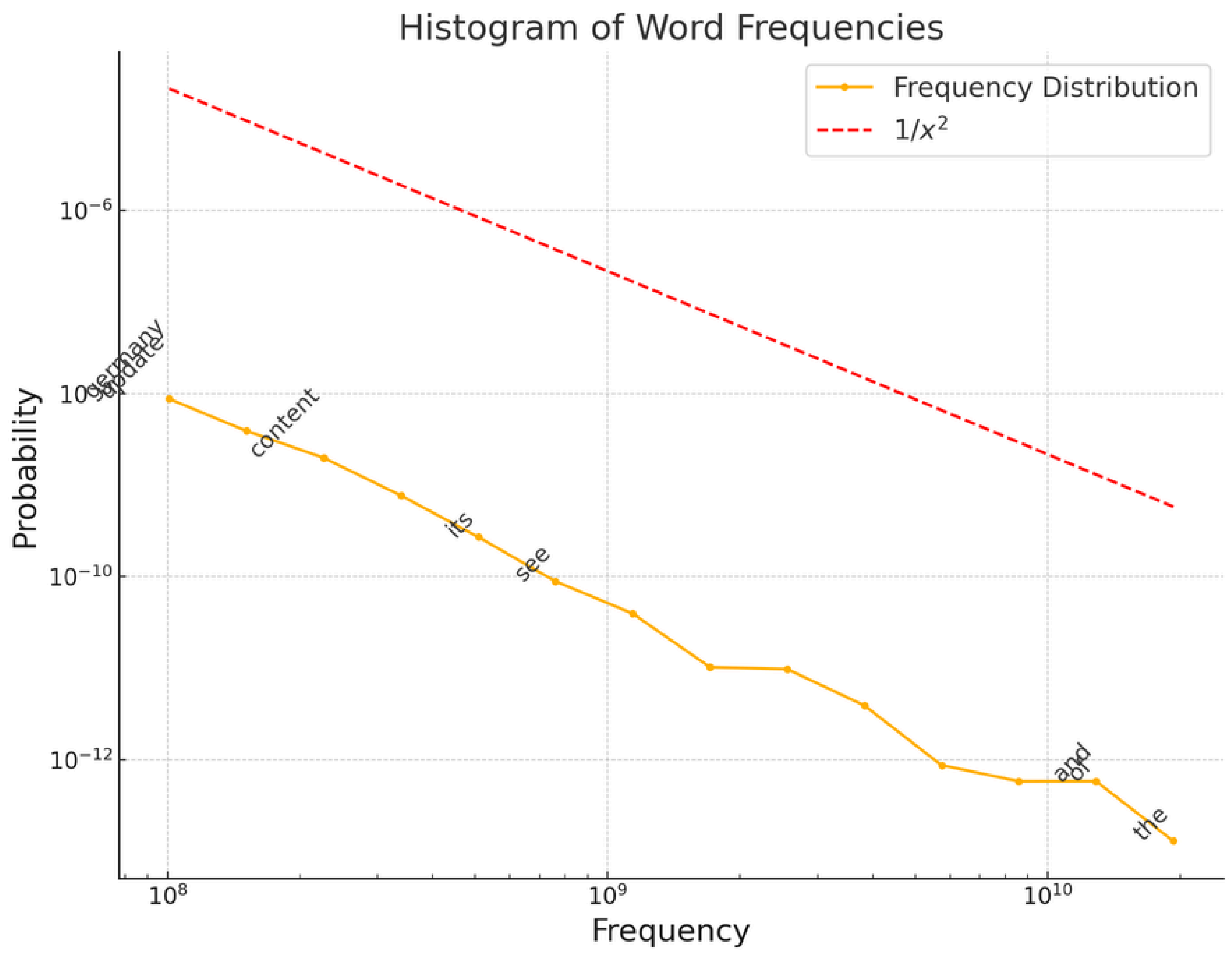
Using a log-log scale the rank-size plot is a straight line if the system shows Zipf's law.

Zipf‘s Law and **Power Law** Distributions

<u>Given a power law distribution with exponent α, then elements extracted from this distribution follow Zipf’s law with exponent γ and it holds</u>

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





Zipf-Mandelbrot Law

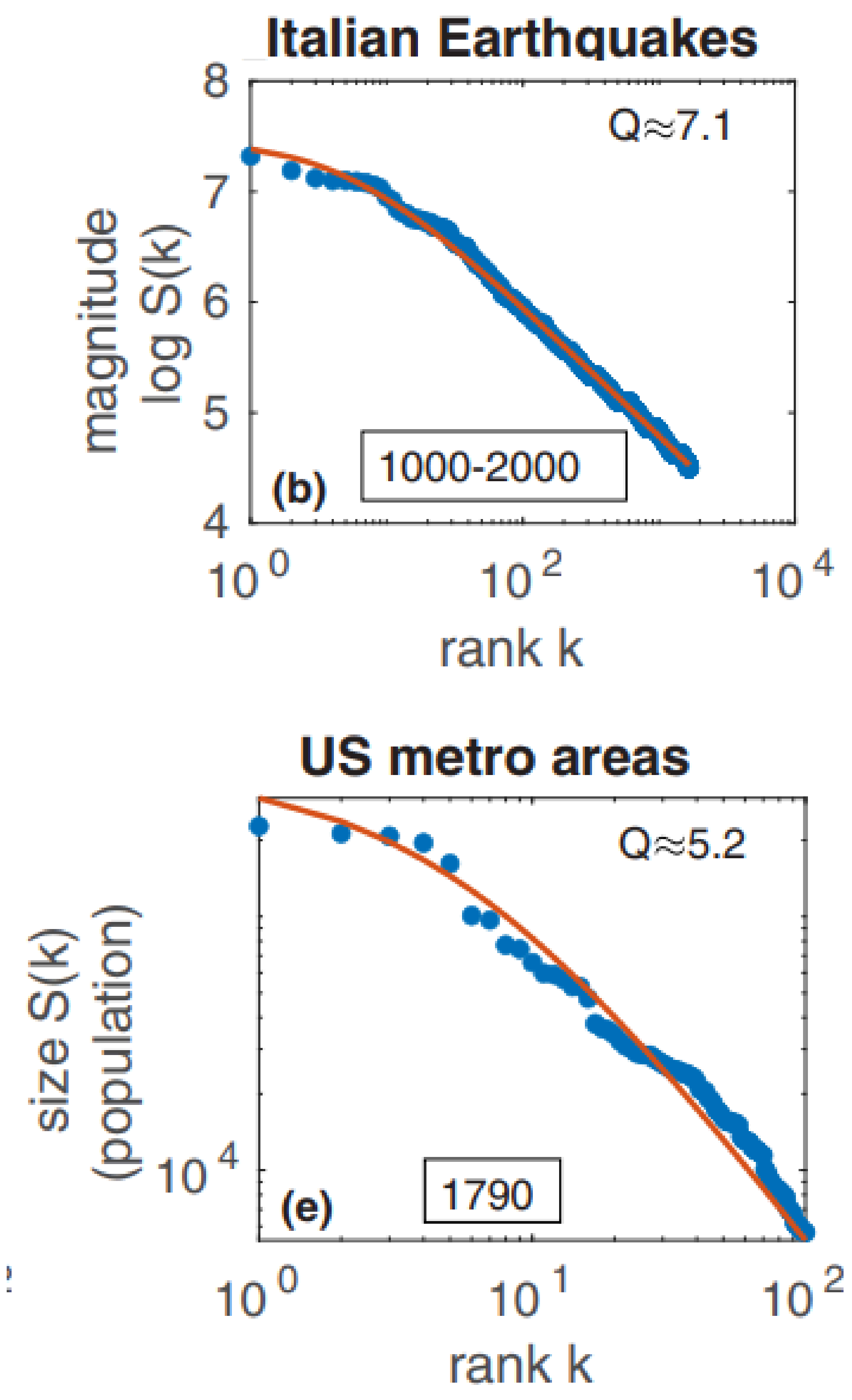
Many systems show deviations from Zipf's law at low ranks. Largest objects are smaller than they should be. This is described by Zipf-Mandelbrot law

$$S(k) = \frac{\bar{S}}{(k+Q)^{\gamma}}$$

Q is the deviation parameter, the larger is Q the large the deviations. Zipf-Mandelbrot law is related to truncated power laws

$$\bar{S} = N^{\gamma} s_m,$$

$$Q = N\left(\frac{s_m}{s_M}\right)^{\frac{1}{\gamma}}.$$



Heaps’ Law and Novelties

Most systems are non-stationary and grow over time. Heaps’ Law describes how the number of distinct elements in a system grows as function of the system size t

$$N(t) \sim t^{\beta}$$

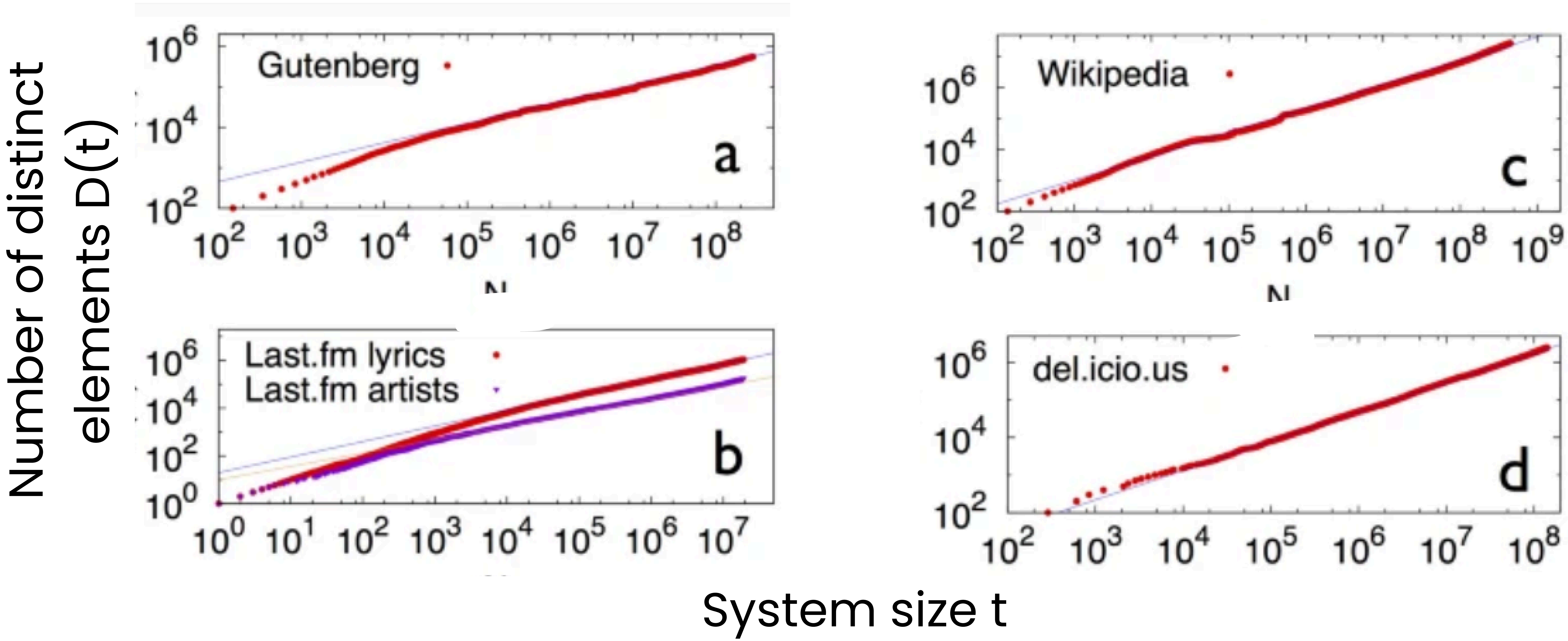
Here

- t is the total system size
  - total urban population of a country
  - total number of words in a book
- N(t) is the number of distinct element
  - number of cities in a country
  - number of distinct words in a book

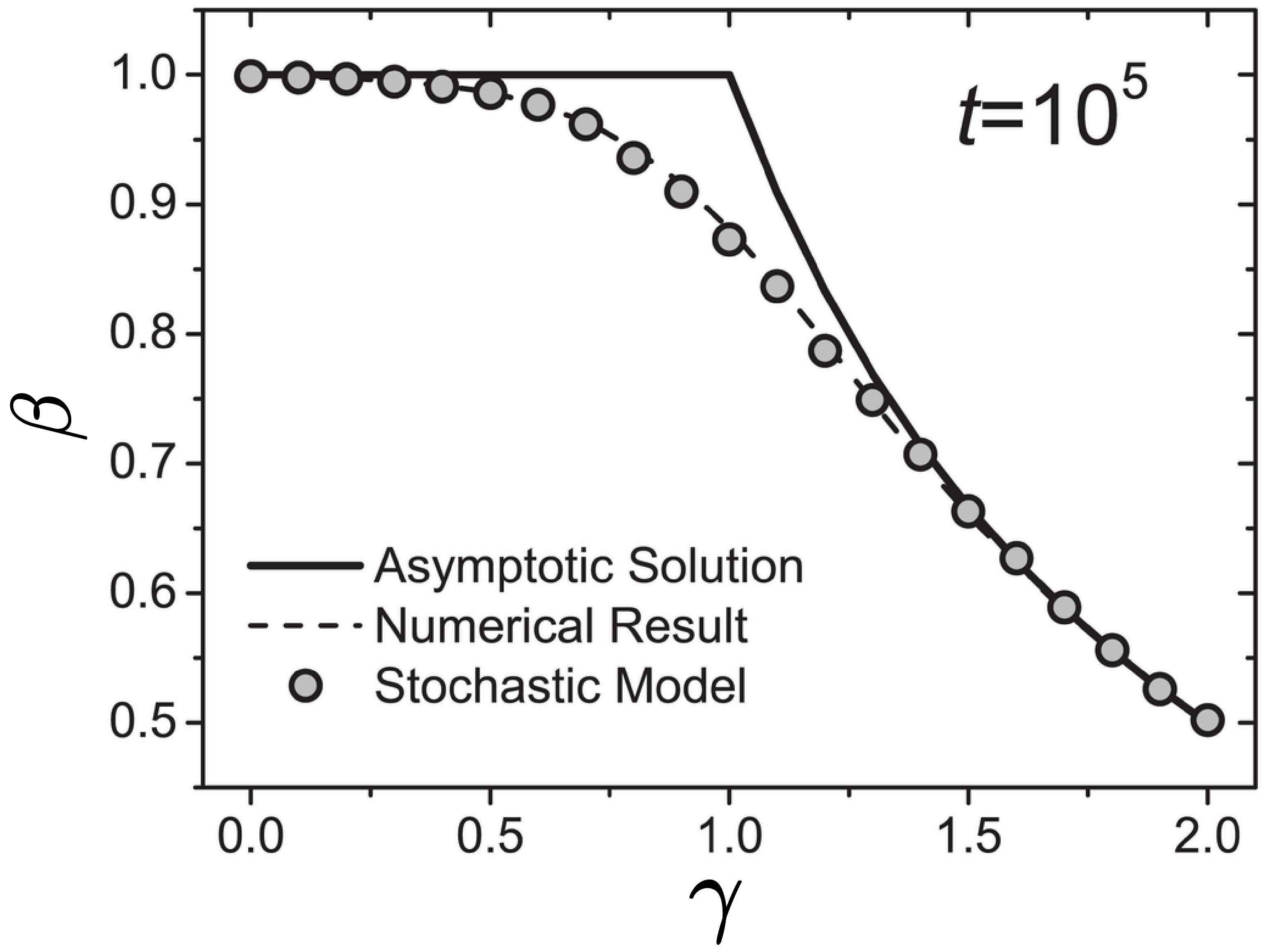
Note that the growth can be at most linear.

Examples of Heaps‘ Law

Heaps’ Law is as ubiquitous as power law distributions and Zipf’s Law.



Heaps‘ and Zipf‘s Exponents



Heaps’ and Zipf’s law are related. In particular any system that shows Zipf’s law stably over time will also present Heaps’ law.

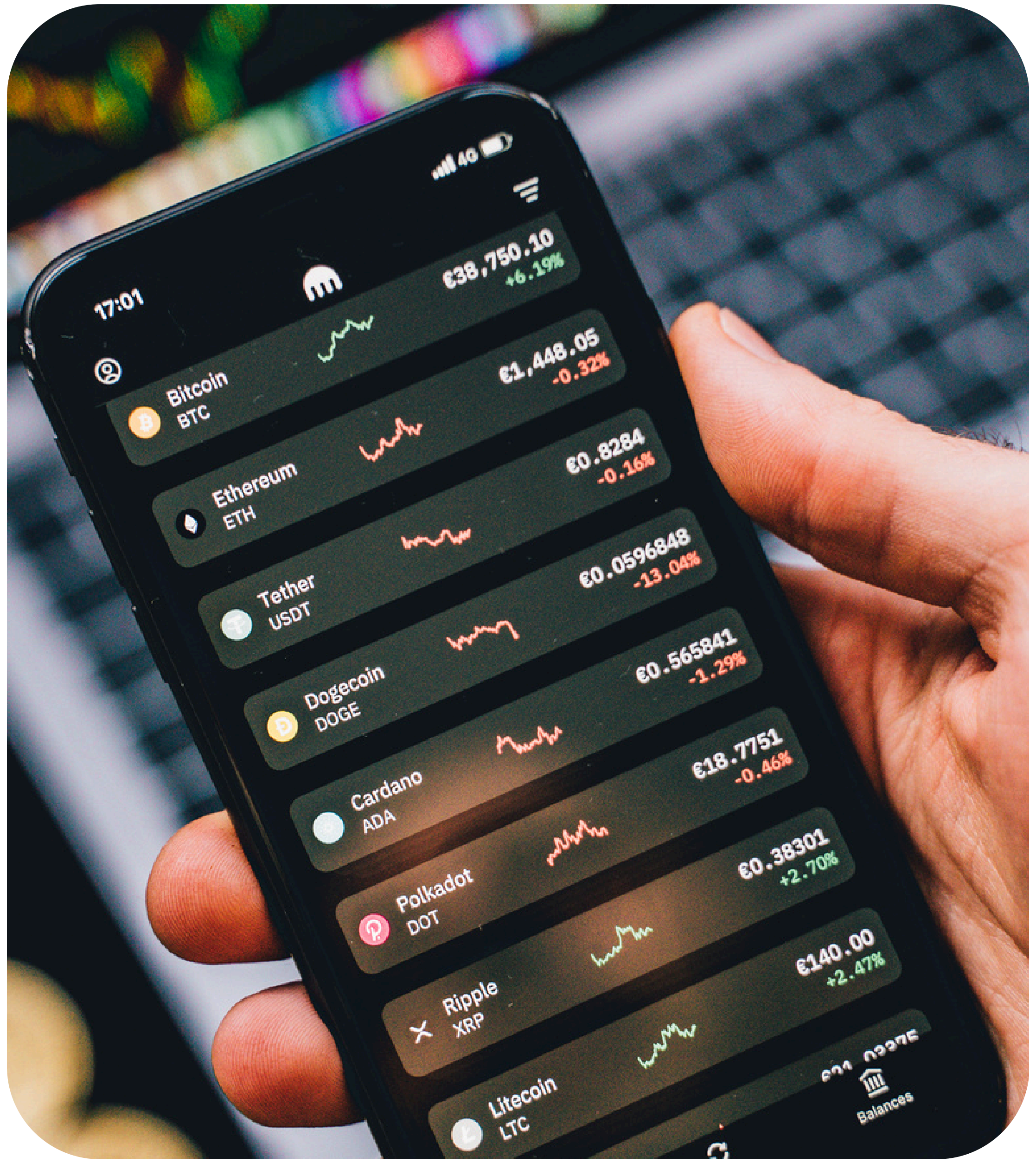
The exponents of the two laws are related by a simple equation

$$\beta = \begin{cases} 1 \text{ if } \gamma < 1 \\ \frac{1}{\gamma} \text{ if } \gamma > 1 \end{cases}$$

Note, however, that in finite systems there are deviations from this exact relation.

Modelling the Cryptocurrency Market

Cryptocurrencies



The crypto market has experienced a huge growth since the introduction of Bitcoin in 2008

- more than 1 trillion euros of market cap
- more than 20000 different cryptocurrencies
- Institutional investors

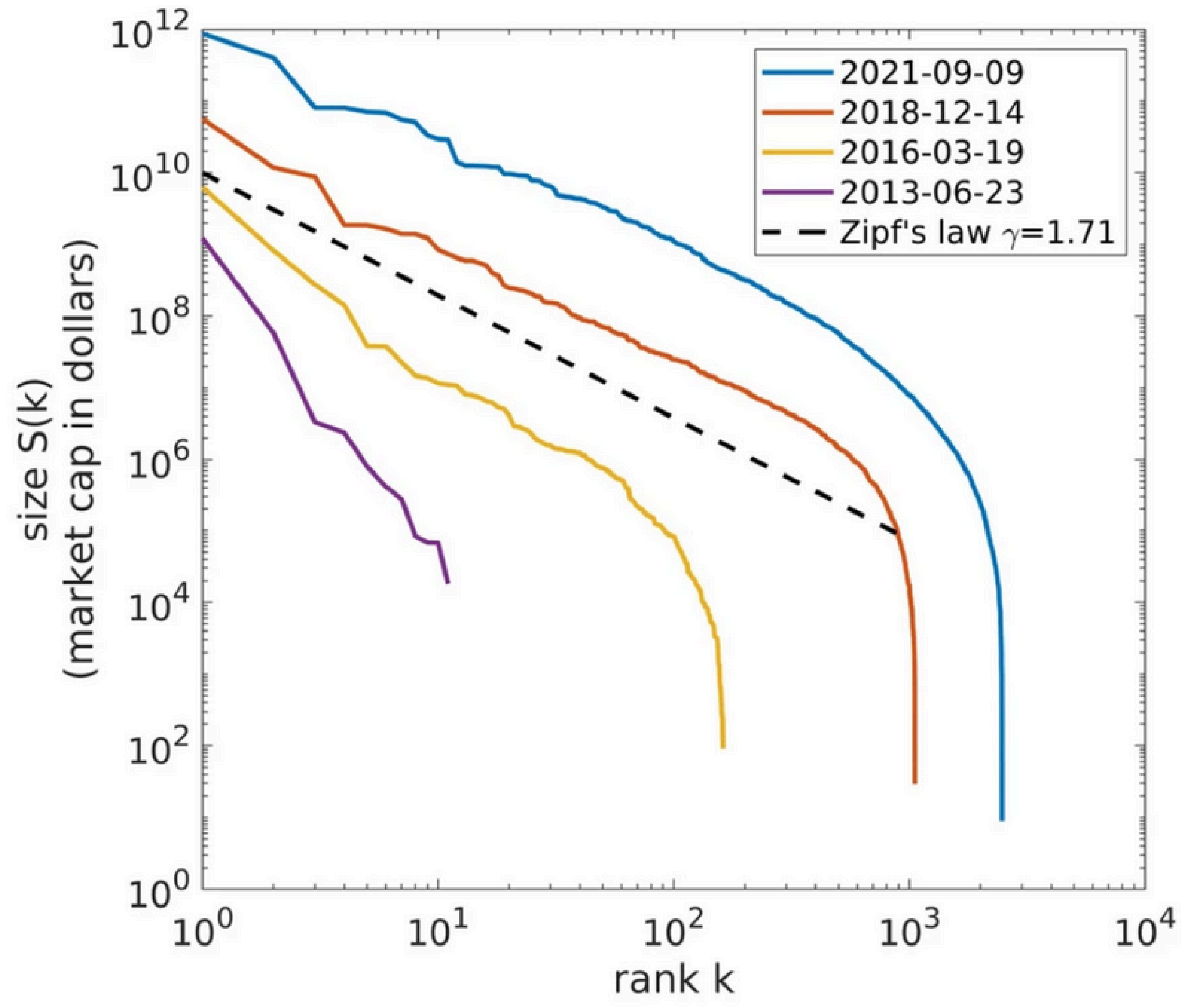
Cryptocurrencies are no more just a curiosity

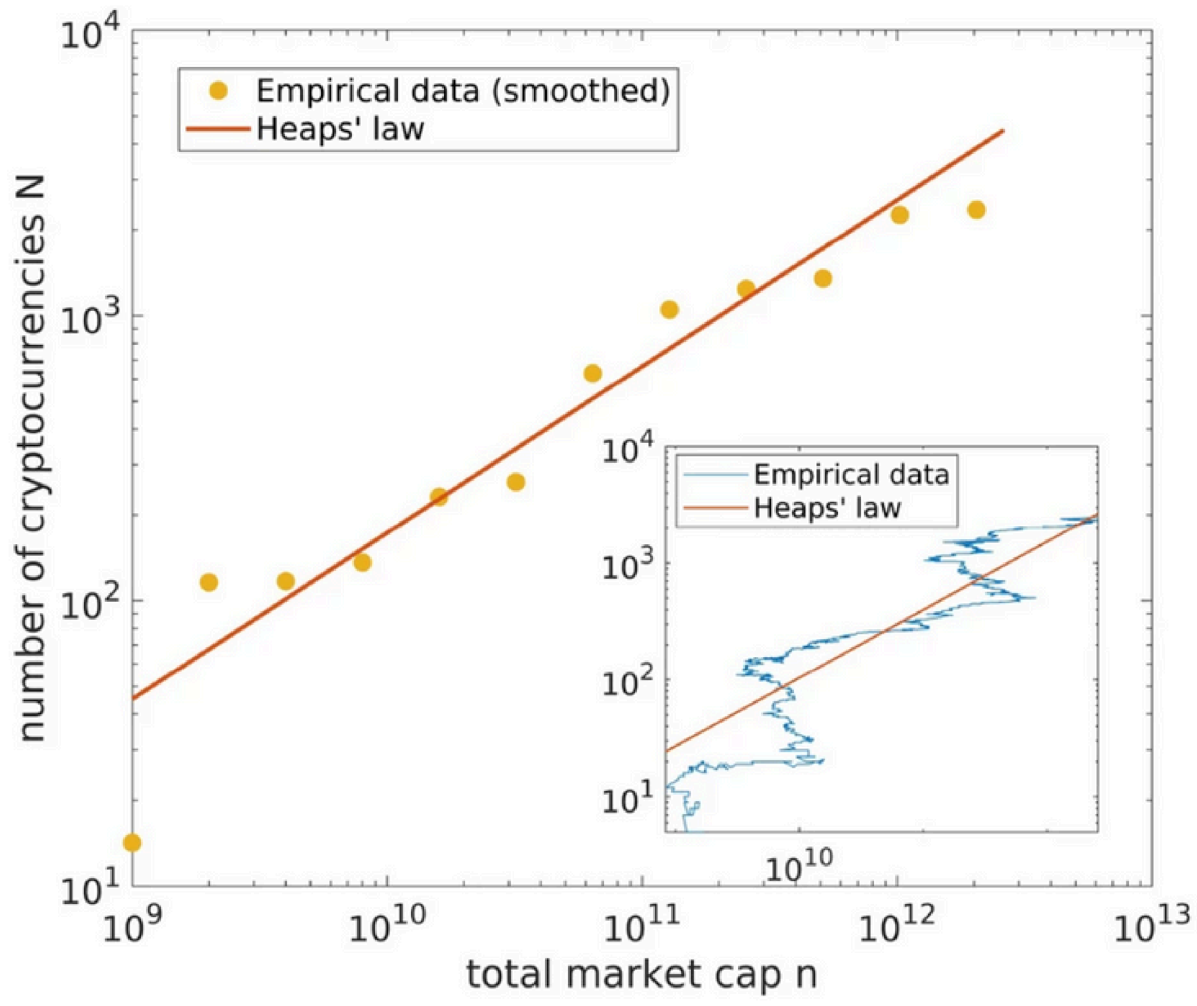
If you don't believe it or don't get it, I don't have the time to try to convince you, sorry. — Satoshi Nakamoto

Power Laws in the Crypto Market

The crypto market shows both Heaps' and Zipf's law

- **Zipf's Law:** sizes are the market caps of cryptocurrencies. ɣ=1.71
- **Heaps' Law:** relates the number of crypto to the total market cap. β=0.58

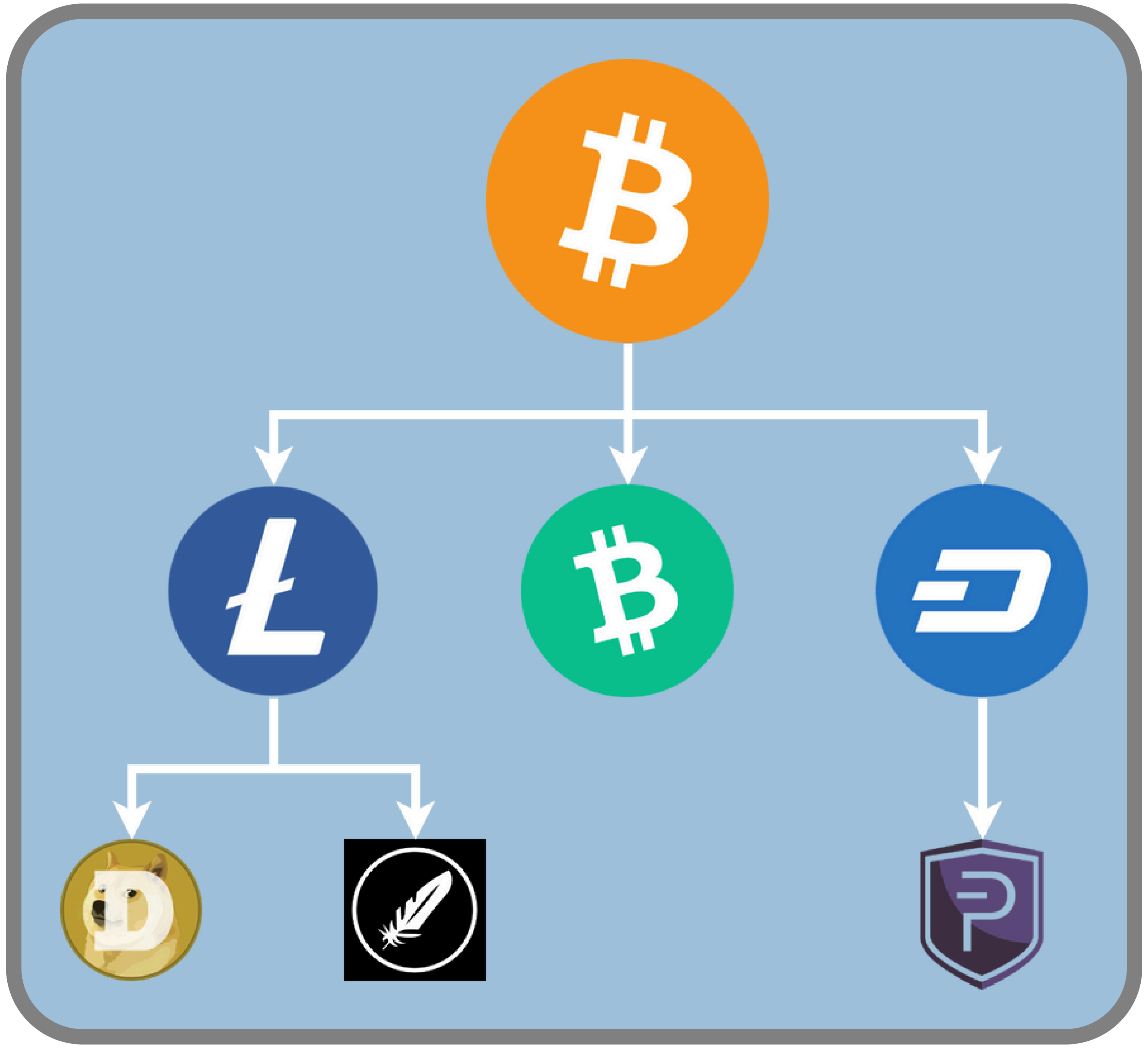




Forking a Crypto

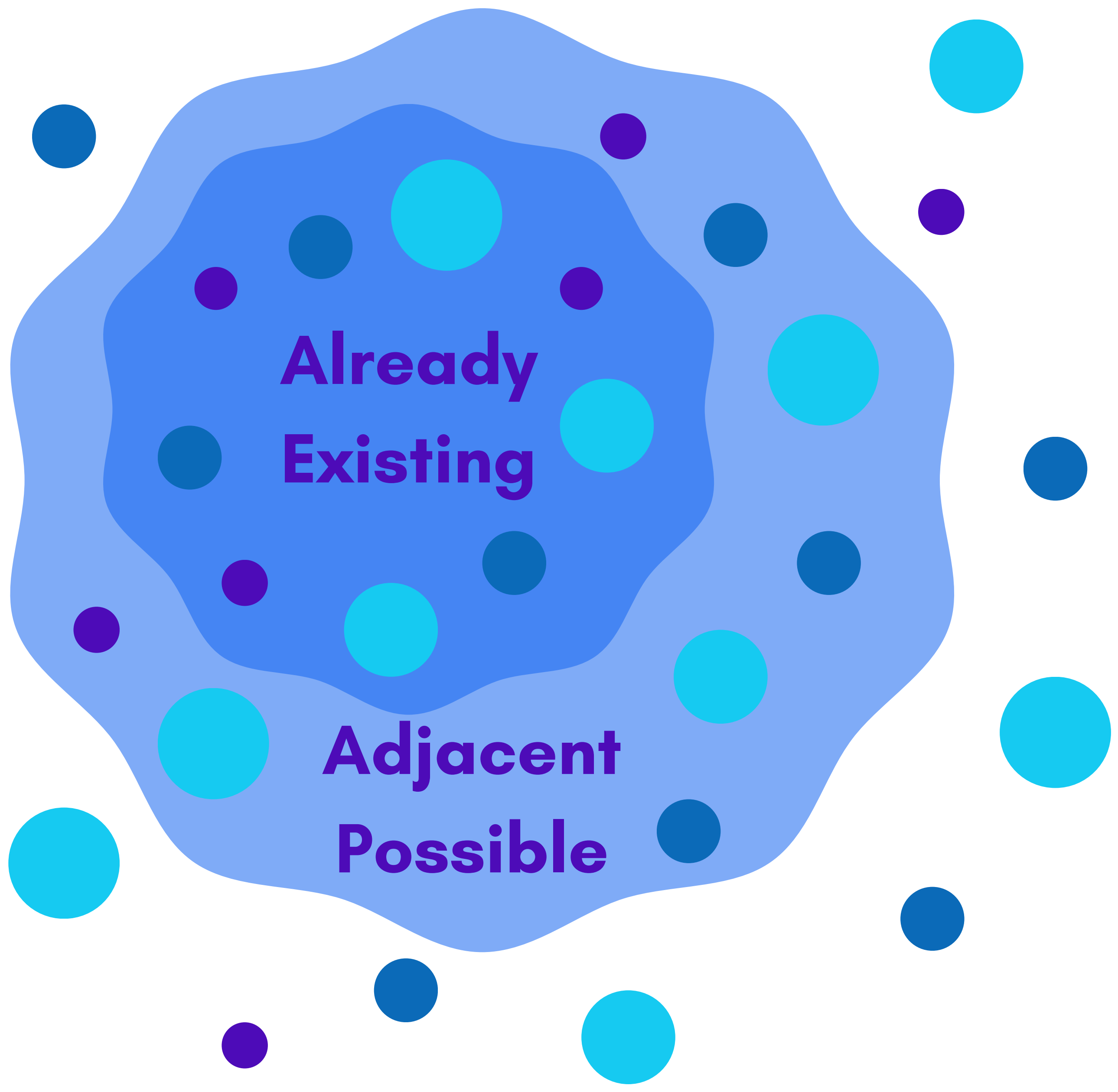
The cryptocurrency ecosystem is innovation driven. Technologies introduced in a project are reused and improved in new cryptocurrencies

- Blockchain forking (Ethereum Classic)
- Software forking (Litecoin)
- DeFi platforms (Ethereum)



Every time a new crypto is created, it triggers the potential birth of novel different cryptos.

Adjacent Possible



By adjacent possible we mean all those things that are not yet existing, but that are just a step away from being realized

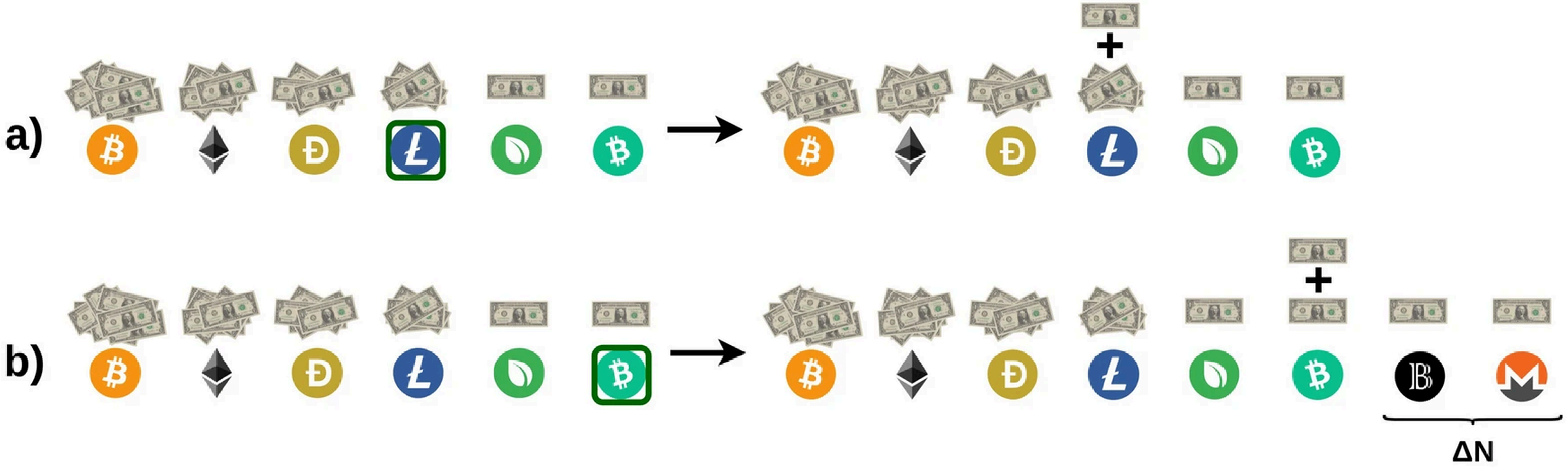
- a novel technology combining two existing technologies
- a land not yet urbanized but close to infrastructures

The idea is that whenever something moves from the Adjacent Possible into existence, the Adjacent Possible enlarges, since the new “invention” facilitates the occurrence of further novelties.

A Model for Cryptocurrencies

We combine a rich get richer mechanism with a triggering effect

- A unit of money is invested in a crypto selected with a probability proportional to its market cap (a)
- Whenever a crypto gets its first investment, ΔN new cryptocurrencies enter the system (b)



Analytical Results

The model is equivalente to the UMT and Heaps' and Zipf's exponent depend on the adjacent possible size ΔN

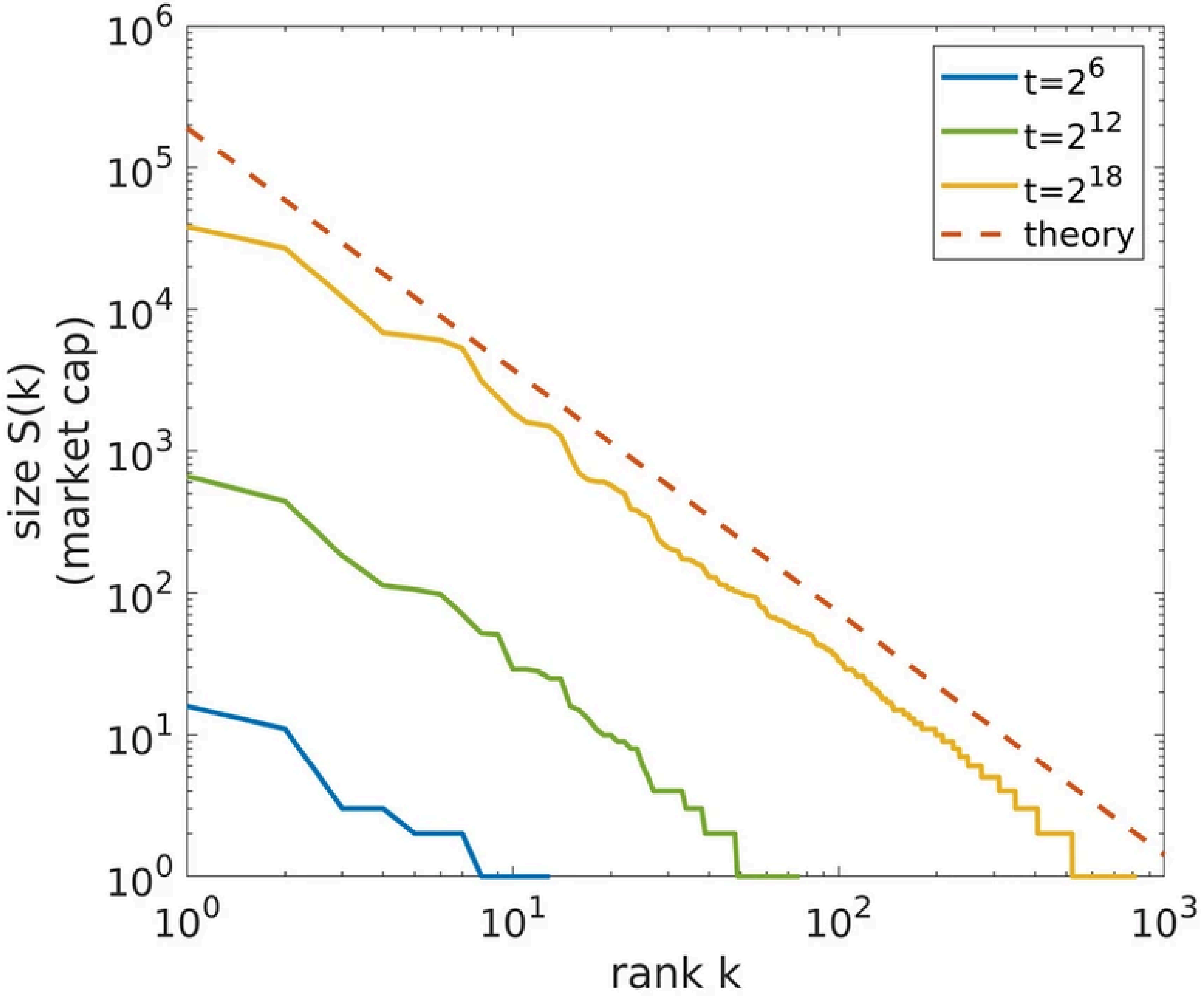
$$\gamma_{mod} = \frac{1}{\Delta N - 1} \qquad \beta_{mod} = \begin{cases} \Delta N - 1 \text{ if } \Delta N < 2 \\ 1 \quad \text{ if } \Delta N > 2 \end{cases}$$

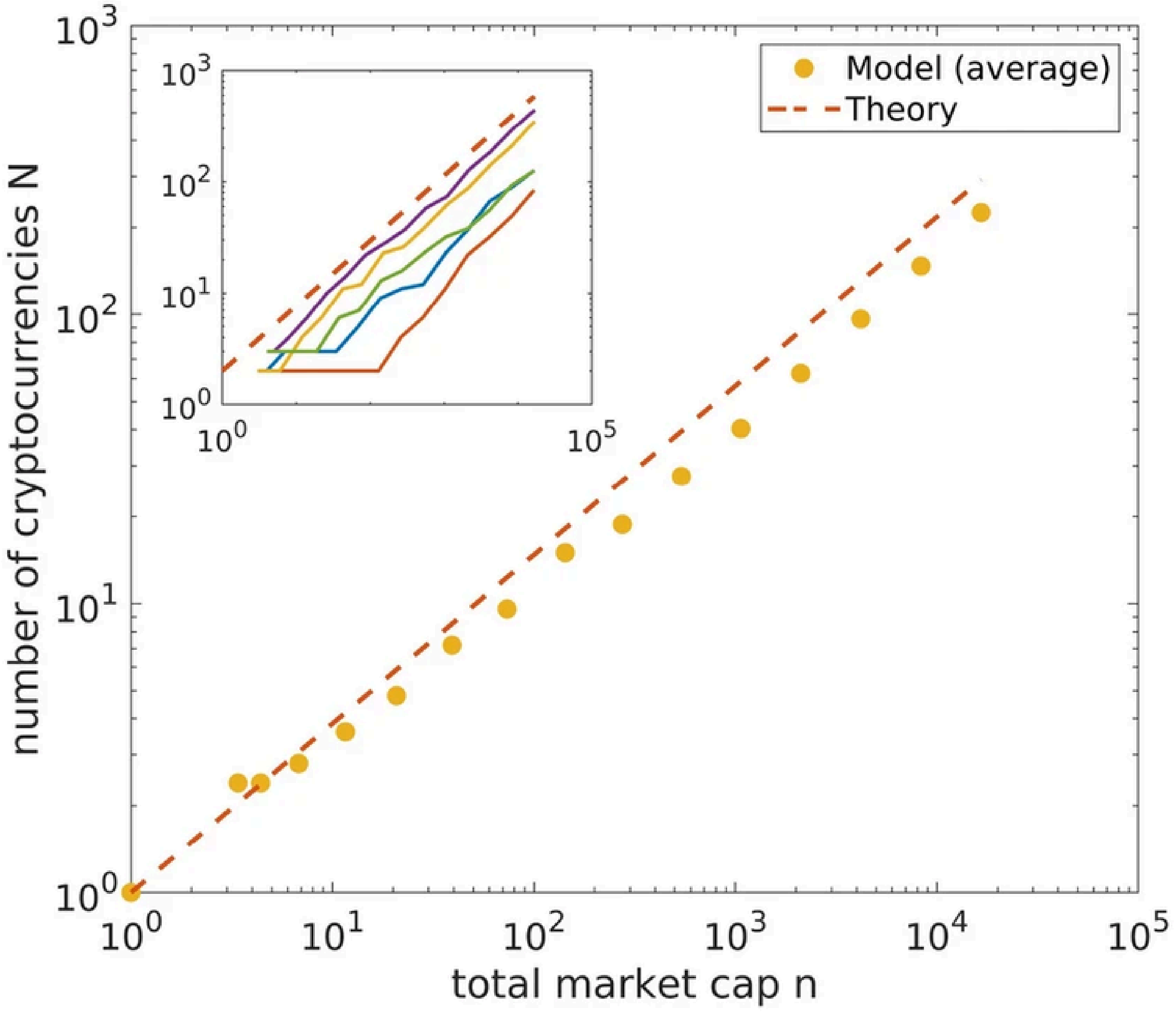
The regime of the market is determined by the adjacent possible size

- if ΔN<2 (each crypto generates less than two forks on average) then the number of cryptos grows sub-linearly
- if ΔN>2 (each crypto generates more than two forks on average) then the number of cryptos grows linearly

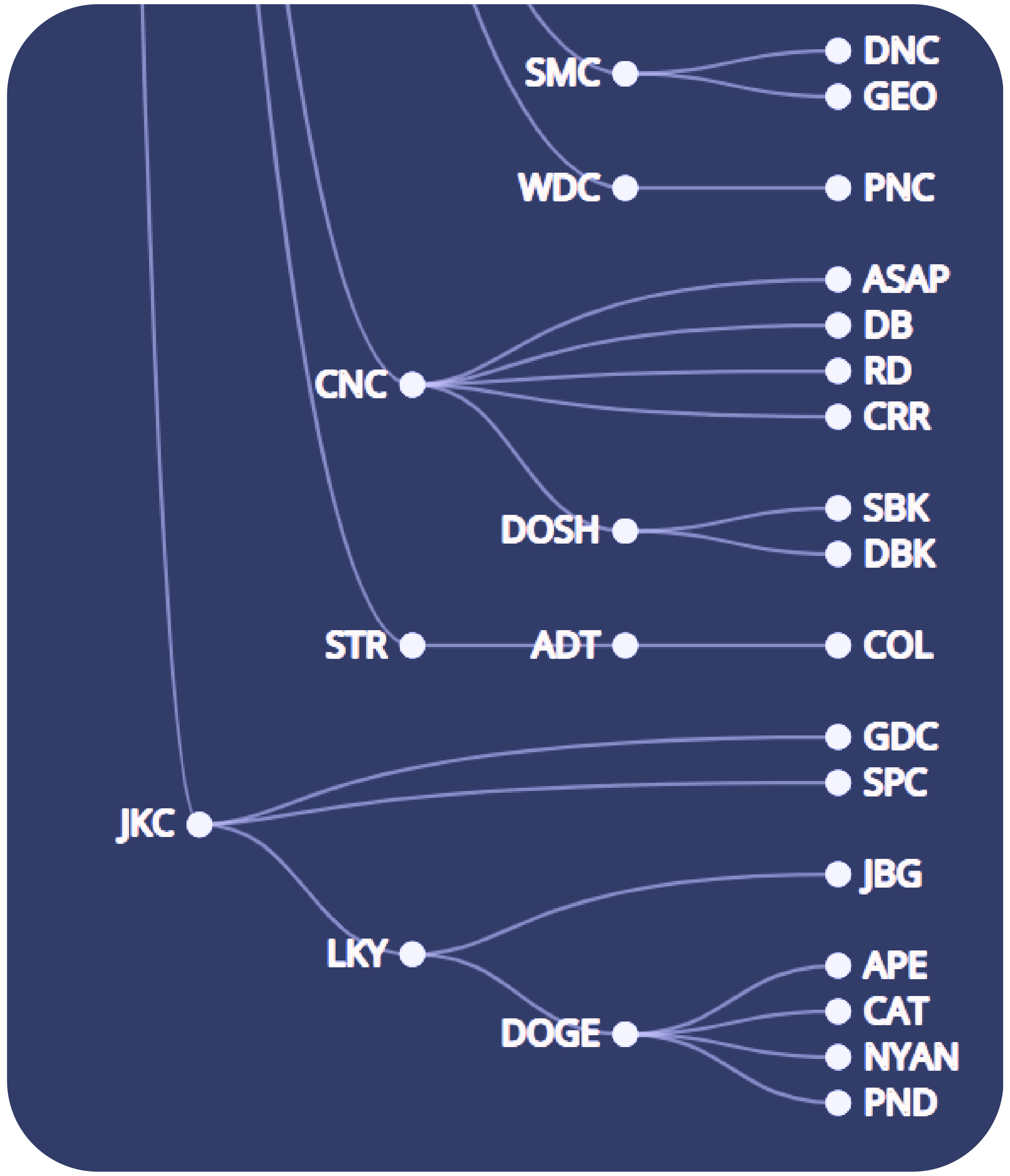
Fitting the Model

Fitting the model to the data we can determine how many cryptocurrencies are triggered. Since β=0.58 we get ΔN=1.58.





Forking Tree



The Forking Tree of Bitcoin contains all cryptocurrencies that originated from Bitcoin and its derivative

- the root of the tree is Bitcoin
- each node of the three is a cryptocurrency
- each crypto il linked to all the cryptos that where forked from it (either code or blockchain)

Testing the Model

We can test the model by measuring the size of the adjacent possible

- we look at all coins ever listed on coinmarketcap.com
  - these are the cryptos that got adopted and thus realized
  - all the cryptos never listed represent the never realized adjacent possible
- we use the Bitcoin forking tree to measure the number of cryptocurrencies that they originated
- this allows to measure the adjacent possible ΔN

From the data we obtain $\Delta N=1.57$, a value almost equal to the prediction of our model $\Delta N=1.58$.

Conclusions

The Paretian World

We generally think in terms of Gaussian distributions, but the world is dominated by events that are poorly describe by the average.

Power Law Probability Distributions

Power law probability distributions are characterized by a much slower decay with respect a Gaussian and they describe extreme events (wars, pandemics...)

Zipf’s Law and Heaps’ Law

Often power law distributions manifest in different shapes, Heaps’ and Zipf’s law are among the most know statistical regularities deriving from power laws.

Modelling the Cryptocurrency Market

The crypto market shows both Heaps’ and Zipf’s law and can be modeled using and adjacent possible based mechanism. Empirical data support this idea.

Quiz

- What is an example of non-power law distributed data?
- What is an example of power law distributed data?
- Do you think a pandemic is a rare event? What about a war or an earthquake?
- Do you think that WWII was a Black Swan?
- What is the rank 1 word for German?
- What are some systems driven by the adjacent possible mechanism?
- What are the consequence of ZIpf’s law for the crypto market?
- What are the limit of the model?