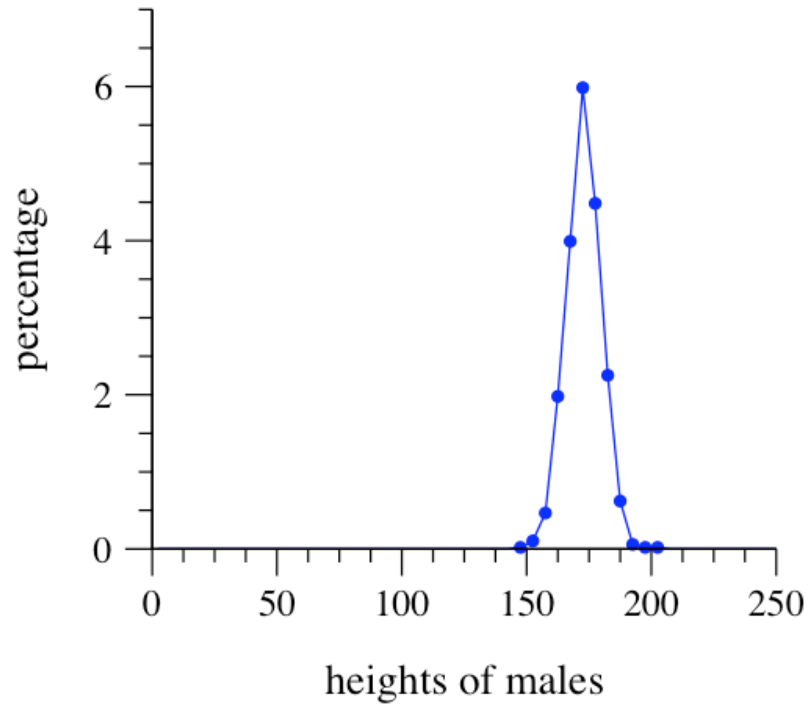
Scale

- Many of the things that scientists measure have a typical scale - a typical value around which individual measurements are centered.
- As an example consider the heights of human beings.

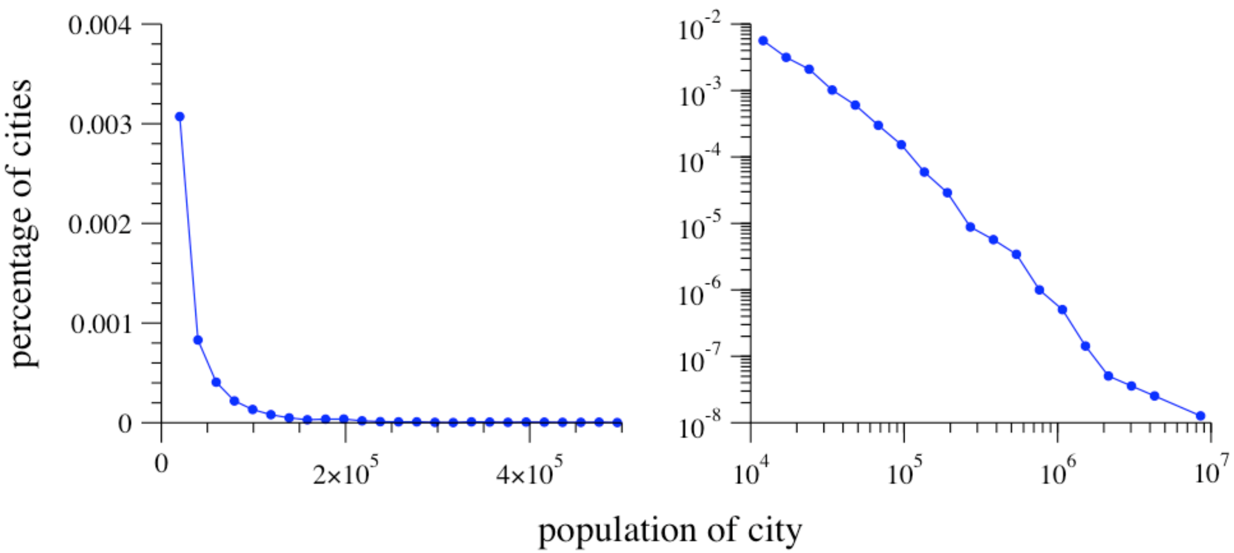
The max/min ratio

is 272/57=4.8



Scale free variables

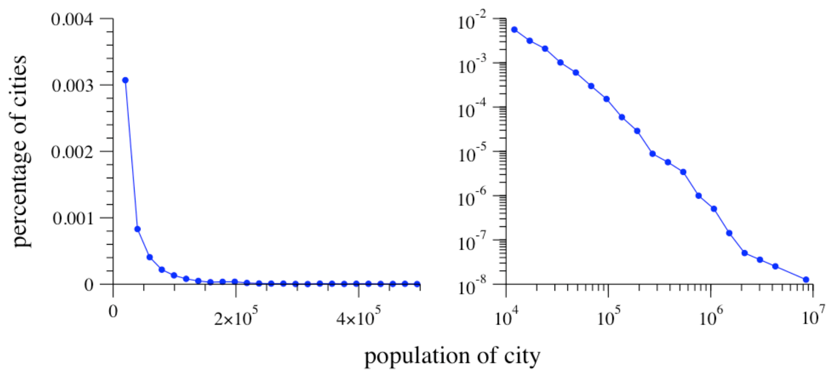
- There are other quantities that do not seem to have a typical scale.
- As an example consider the sizes of towns and cities. The max/min ratio is of the order of 150,000 (It may depends on the definition of town)



- Note that all quantities must have a finite maximum size.

Scale free variables

- When we plot the histogram of city size in log-log scale we get a straight line (Auerbach 1913, Zipf 1949)



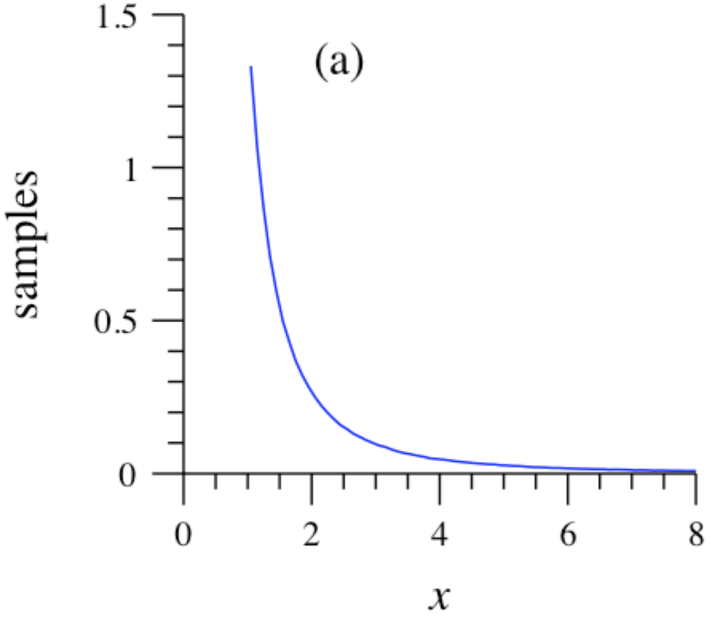
- A straight line in log-log scale implies a functional dependence of the probability density p(x) of a city with size x of the form

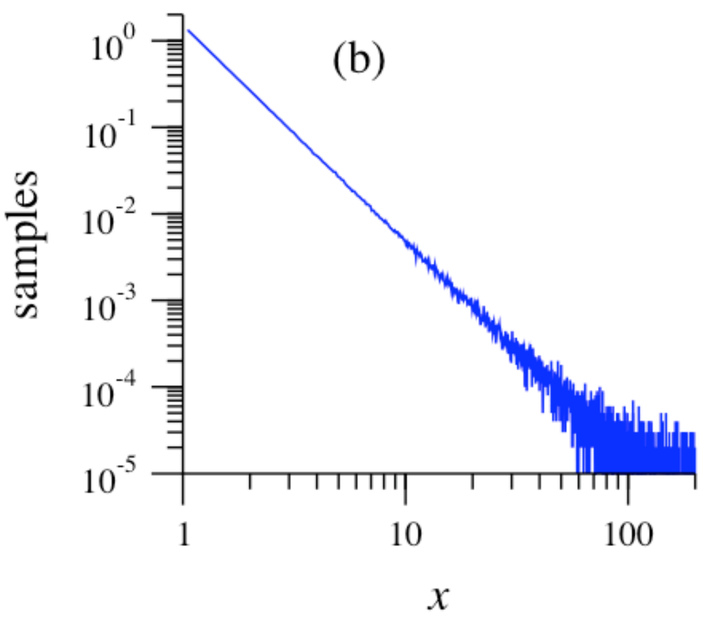
$$\boxed{p(x) = Cx^{-\alpha}}$$

- Distributions of this form are said to follow a **power law**

Measuring power laws

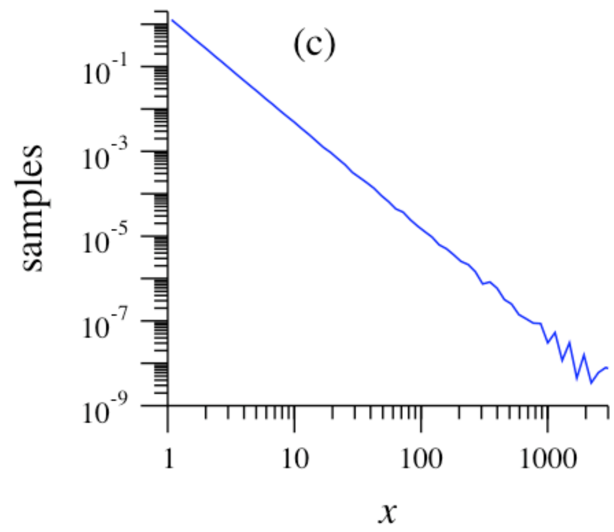
- Identifying power law behavior and measuring the exponent can be tricky.
- The quick and dirty procedure is to plot the histogram in a log log scale and fit a power law in some region.
- This procedure presents several problems due to the noise of sampling errors. These errors are large in the tail of the distribution, precisely where the power law is more likely to hold.
- Example: 1 million random numbers power law distributed with exponent α=2.5





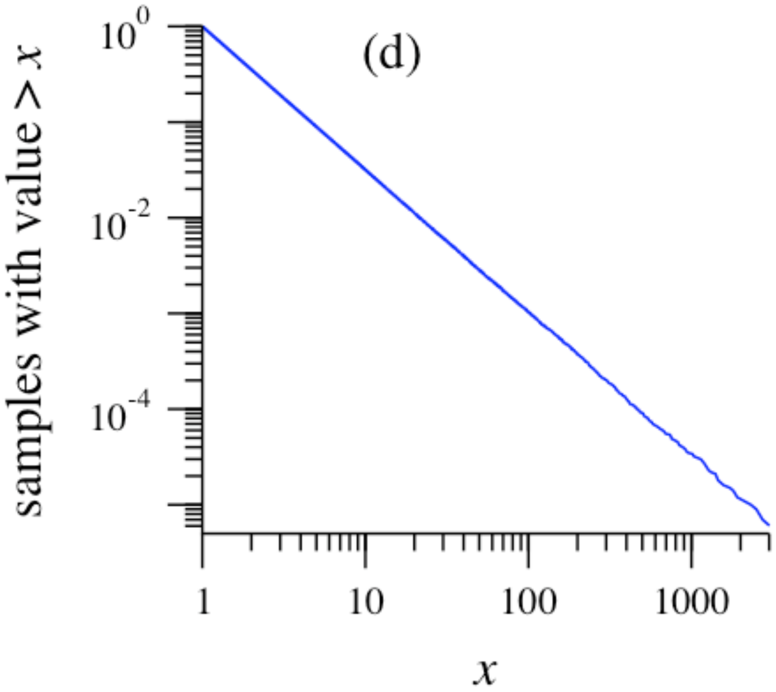
Logarithmic binning

- An improvement in the identification of power law is to compute the histogram by using bins of varying width.
- The number of samples in a bin of width Δx, should be divided by Δx to get a count per unit interval of x
- The most common binning is to create bins such that each is a fixed multiple wider than the one before it (E.g. [1;1.1],[1.1;1.3],[1.3:1.7],…).

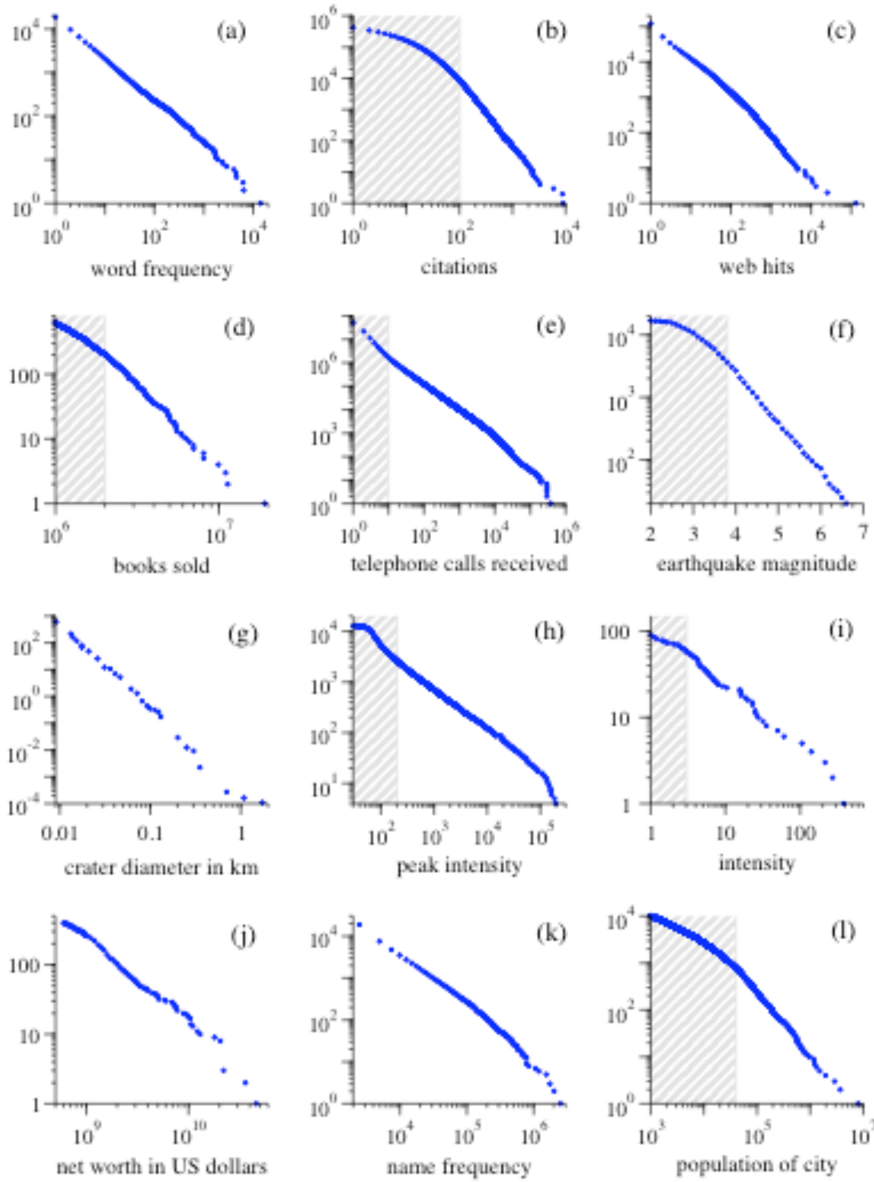


Cumulative distribution function

- Plotting the cumulative distribution function is equivalent to do a rank plot and to switch x and y axis
- It is the best way to plot power law distribution because
  - the noise is reduced
  - all the data are used

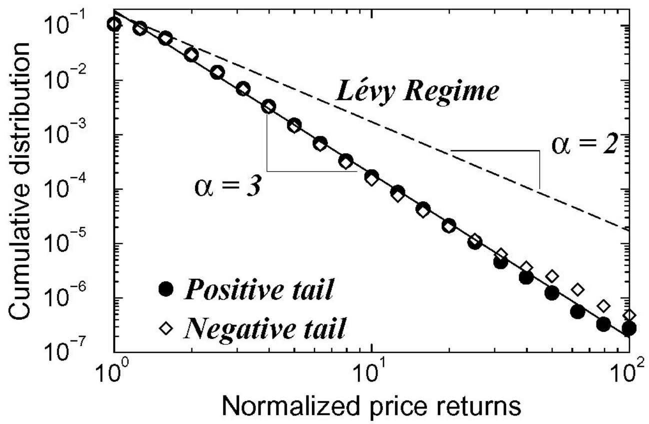


Examples of power laws distributions



Power laws distributions in Economics

Asset returns (Stanley 1998)



Firm size (Axtell 2001)

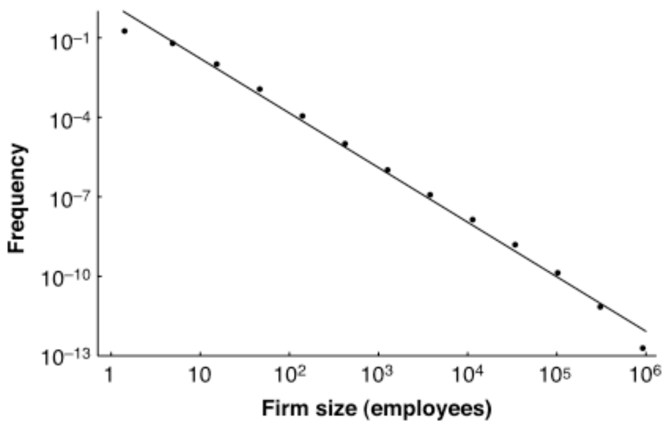


Fig. 1. Histogram of U.S. firm sizes, by employees. Data are for 1997 from the U.S. Census Bureau, tabulated in bins having width increasing in powers of three (*30*). The solid line is the OLS regression line through the data, and it has a slope of 2.059 (SE = 0.054; adjusted R^2 = 0.992), meaning that α = 1.059; maximum likelihood and nonparametric methods yield similar results. The data are slightly concave to the origin in log-log coordinates, reflecting finite size cutoffs at the limits of very small and very large firms.

Income (Fujiwara 2002)

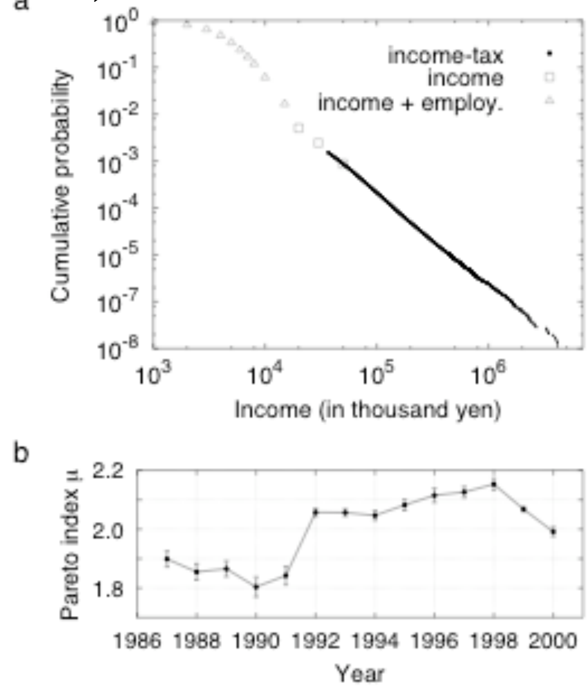


Figure 1: Personal income in Japan. a. Cumulative probability distribution of personal income from low to high income range in the year 2000. A data-point

Example

$$p(x) = e^{-x/10}$$

$$p(x) = \frac{1}{x^{1.5}}$$