Pareto Distribution

Background

- Vilfredo Pareto (1897) presents a versatile functional relation that well describes wealth distribution across countries and centuries.
- Same concept is applied to several other fields and colloquially called *Pareto Principle*.
    - 80% of land owned by 20% of individuals (revenue $\sim$ products; sales $\sim$ clients; etc)
- Generally, it follows a *power law probability distribution*, where one measure varies constantly as an exponential of another, independently of initial values.
    - Example: if one increases the side length of a square by x, its area increases by x^2, independently of initial area of square.

Background

- Vilfredo Pareto (1897) presents a versatile functional relation that well describes wealth distribution across countries and centuries.
- Same concept is applied to several other fields and colloquially called *Pareto Principle*.
  - 80% of land owned by 20% of individuals (revenue $\sim$ products; sales $\sim$ clients; etc)
- Generally, it follows a *power law probability distribution*, where one measure varies constantly as an exponential of another, independently of initial values.
  - Example: if one increases the side length of a square by x, its area increases by x^2, independently of initial area of square.

- Latent sector errors in hard-drives failures
- Clusters of Bose–Einstein condensate near absolute zero

Functional Form

The **Pareto Distribution** is defined by

$$f(y, \underline{y}, \alpha) = \frac{\alpha \underline{y}^{\alpha}}{y^{\alpha+1}}, \quad 0 < \underline{y} < y$$

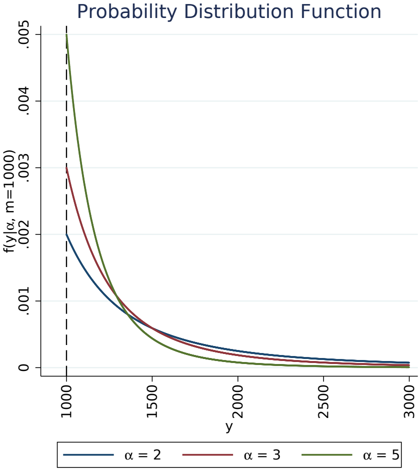
and

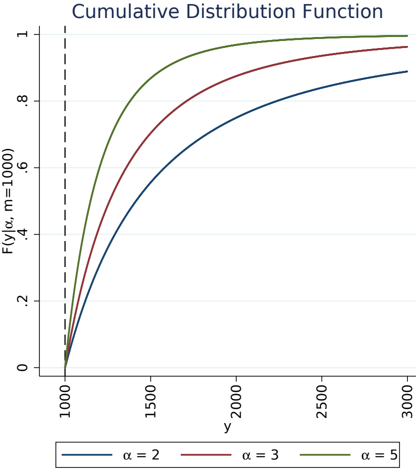
$$F(y, \underline{y}, \alpha) = 1 - \left(\frac{\underline{y}}{y} \right)^{\alpha}, \quad 0 < \underline{y} < y$$

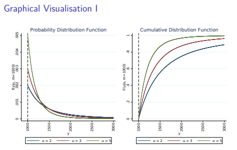
where:

- y: wealth measure
- α: Pareto's α (or *shape/slope parameter*)
- $\underline{y}$: lower bound (or *scale parameter* or *threshhold value*)

Graphical Visualisation I

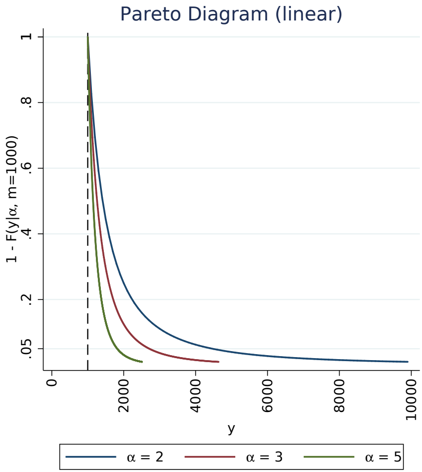


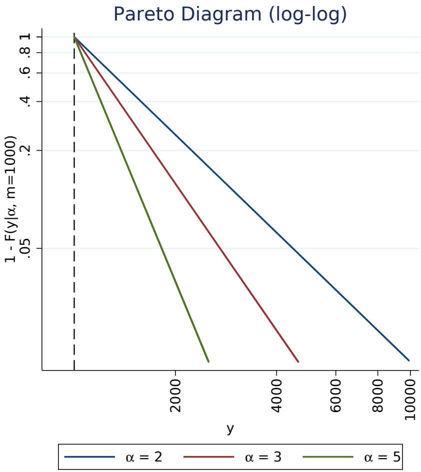




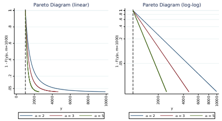
Here we can see how the Probability and Cumulative Distribution looks like.
- Only defined after lowerbound
- very heavy right tail
- very left dense

Graphical Visualisation II





Graphical Visualisation II



In order to better visualise this distribution TWO transformations are taken:

1. calculate 1 - CDF
2. plot on log-log scale (or log-log the data)

    .

    .

HERE: lower bound = 1000

α = Slope in loglog

Properties

Pareto's α:

- Sloppy interpretation: for a percentage increase in y, the proportion of *richer* individuals by α percents.
- Higher α values $\Rightarrow$ less inequality.[1]
- Several inequality indexes can be estimated based on α.
  - Example: Gini coefficient: $\frac{1}{2\alpha-1}$.

Possible problems:

- High flexibility on estimating the lower bound
- Sensibility of α due to choice of the lower bound $\underline{y}$

[1] for inequality measures satisfying the *Weak Transfers Principle* (Cowell 2011, p. 89).

Properties

Pareto's α:

- Sloppy interpretation: for a percentage increase in y, the proportion of *richer* individuals by α percents.
- Higher α values $\Rightarrow$ less inequality.[1]
- Several inequality indexes can be estimated based on α.
  - Example: Gini coefficient: $\frac{1}{2\alpha-1}$.

Possible problems:

- High flexibility on estimating the lower bound
- Sensibility of α due to choice of the lower bound $\underline{y}$

[1] for inequality measures satisfying the *Weak Transfers Principle* (Cowell 2011, p. 89).

- For narrow excerpts of the data other distributions are "just as good"