The Pareto principle is a widely applied heuristic that states that 80% of effects result from 20% of the causes (for this reason the Pareto principle is also known as the 80-20 rule or the law of the vital few). The name Pareto comes from economist Vilfredo Pareto who observed in 1906 that 80% of the land in Italy was owned by 20% of the population. Since that time the rule has been found to apply very widely.

- In business: 80% of sales come from 20% of clients.
- In stock markets: 80% of the growth of the market comes from 20% of the stocks.
- In computer programming: 80% of crashes are caused by 20% of the bugs.
- In health care: 80% of expenditures are on 20% of the patients.

Exercise 1. The Pareto principle applies to wealth distribution in the country of Extremistan, where 20% of the people control 80% of the wealth. The Pareto principle also applies to just the richest 20% of people in Extremistan: 20% of the richest 20% of the population control 80% of the wealth of the richest 20%.

a) The richest 4% of the population controls what percentage of all wealth in Extremistan?

b) Keep applying the Pareto principle to find the number p so that about 41% of the wealth is controlled by the richest p% of the population.

c) The Pareto principle also works for the poorest people. The poorest 51.2% control what percent of the wealth?

Exercise 2. In another country, Mediocristan, wealth is normally distributed with $\mu = 6$ and $\sigma = 1$ (units are 10,000 Mediocristani dollars). The proportion of wealth controlled by the poorest $100p$ percent of the population is given by

$$p - \frac{1}{6}\left(\sqrt{\frac{2}{\pi}}\right) e^{-\left(\frac{z(p)}{2}\right)}$$

where $z(p)$ is the $100p^{\text{th}}$ percentile of the standard normal distribution.

a) What percent of the wealth is controlled by the poorest 50% in Mediocristan?

b) What percent of the wealth is controlled by the richest 20% in Mediocristan?

As Vilfredo Pareto observed, wealth distributions in the real world are much more like Extremistan than Mediocristan (the names Mediocristan and Extremistan come from Nassim Nicholas Taleb's very interesting book *The Black Swan*). The probability distributions that capture this behavior are known as Pareto distributions. The pdf of Pareto a distribution with parameters $\alpha > 0$ and $\beta > 0$ is

$$f(x; \alpha, \beta) = \begin{cases} \frac{\alpha(\beta^\alpha)}{x^{\alpha+1}} & \text{if } x \geq \beta \\ 0 & \text{otherwise} \end{cases}$$

Exercise 3. Let X be the wealth of a random person. X is a Pareto random variable with parameters $\alpha > 1$ and $\beta = 1$.

a) Find the cdf for X. $F(x) = \Bigg\{$

b) Calculate $\mathrm{E}(X)$.

The *Lorenz curve* of our random variable X is the graph of the function

$$L(x) = \frac{\int_1^{F^{-1}(x)} tf(t)dt}{\mathrm{E}(X)}$$

where f is the pdf of X, F^{-1} is the inverse function of the cdf of X, and the domain of the Lorenz curve is $[0, 1]$. $L(x)$ is the proportion of wealth controlled by the poorest $100x$ percent of the population.

Exercise 4.

a) Find $F^{-1}(x)$.

b) Use part (a) to find a shorter expression for $L(x)$ (one that doesn't involve an integral).

c) Calculate $L(0.8)$ when $\alpha = \log_4 5$. Explain what this means (without resorting to fancy mathematics).

Note that different values of α in the previous problem correspond to different versions of the Pareto principle. The value $\alpha = \log_4 5$ happens to give the 80-20 law.

Exercise 5. Use a computer or calculator to sketch the Lorenz curve for $\alpha = \frac{10}{9}$, $\alpha = 2$, and $\alpha = 10$. What happens to the Lorenz curve as α goes to ∞?

The *Gini coefficient* is a measure of wealth inequality based on the Lorenz curve. The Gini coefficient $1 - 2B$ where B is the area under the Lorenz curve (between 0 and 1). A Gini coefficient of 1 corresponds to $\alpha = 1$ and represents all the wealth being owned by one person. A Gini coefficient of 0 corresponds to $\alpha = \infty$ and represents the completely equal distribution of wealth. A list of countries and their Gini coefficients can be found on Wikipedia: `http://en.wikipedia.org/wiki/List_of_countries_by_income_equality`. Wikipedia also has a list of US states by Gini coefficient: `http://en.wikipedia.org/wiki/List_of_U.S._states_by_Gini_coefficient`. For example, Washington has a Gini coefficient of 0.439 while New York is the highest at 0.502.