INCOME INEQUALITY

In 1906 Italian economist Vilfredo Pareto observed that the wealthiest 20% of Italian land-owners controlled 80% of the land. Since Pareto first stated this 80-20 rule, it has been found to apply (approximately) to many different situations: about 80% of sales come from 20% of customers, about 80% of stock market growth comes from 20% of stocks, about 80% of health care spending is on 20% of patients, and so on. This rule has come to be known as the **Pareto principle** (or the 80-20 rule).

1. Suppose wealth distribution in the country of Extremistan[1] follows the Pareto principle strictly: the richest 20% have 80% of the wealth. Moreover, among the richest 20% of the population the Pareto principle still applies: the richest 20% of this group have 80% of the wealth of this group (that's 80% of the 80% of all wealth this group controls). This means that the richest 4% (0.2×0.2) of all Extremistanis have 64% (0.8×0.8) of all the wealth.

a) Apply the Pareto principle again to find the percentage of all wealth owned by the richest 0.8%.

b) The Pareto principle also applies to the poorest 80% of Extremistanis (who have 20% of all the wealth). What percent of the wealth do the poorest 64% have?

2. In Mediocristan, wealth is normally distributed with a mean of 6 and a standard deviation of 1 (units are 10,000 Mediocristani dollars). The proportion of wealth in the hands of the poorest $100p\%$ of Mediocristanis is then

$$L(p) = p - \frac{1}{6}\left(\frac{1}{\sqrt{2\pi}}\right) e^{-\frac{1}{2}[z(p)]^2} \tag{1}$$

where $z(p)$ is the $100p^{\text{th}}$ percentile of the standard normal distribution (`qnorm(p)` in **R**). Note that this means that the proportion of wealth owned by the richest $100(1-p)\%$ is then $1 - L(p)$. Compare the percentage of Mediocristan's wealth controlled by the poorest 80%, poorest 64%, richest 20%, and richest 4% of Mediocristanis to the wealth of their counterparts in Extremistan.

Date: April 29, 2024.

[1]the names Extremistan and Mediocristan come from Nassim Nicholas Taleb's book, **The Black Swan**

Definition. The **Lorenz curve** of a wealth distribution is the curve describing the proportion of wealth owned by the poorest $100p\%$ of the population. If the random variable X is the wealth of a random person in the population, then the Lorenz curve is the graph of

$$L(p) = \frac{\int_{-\infty}^{x(p)} xf(x)dx}{E(X)}$$

where $f(x)$ is the PDF of X and $x(p)$ is the $100p^{\text{th}}$ percentile of X. The function in $L(p)$ in the previous problem (equation 1) is the Lorenz curve for Mediocristan.

As Vilfredo Pareto observed, the real world is much more like Extremistan than Mediocristan. The probability distributions that capture this kind of behavior **Pareto distributions**.

Definition. The PDF of a Pareto distribution with parameter $\alpha > 1$ is

$$f(x) = \begin{cases} \alpha x^{-(\alpha+1)} & \text{if } x > 1 \\ 0 & \text{elsewhere} \end{cases}$$

Note that different values of α give different versions of the Pareto principle: the 80-20 rule corresponds to a parameter of $\alpha \approx 1.16096$.

3. The goal of the next problem is to calculate the Lorenz curve for a Pareto distribution. To do this, you'll first need to find a formula for $x(p)$, the $100p^{\text{th}}$ percentile. Do this by solving the following equation for $x(p)$:

$$p = \int_1^{x(p)} \alpha x^{-(\alpha+1)} dx$$

4. Suppose that income in some country follows a Pareto distribution with parameter $\alpha > 1$.

a) Find a formula for the Lorenz curve $L(p)$ (this will depend on α).

b) Graph the Lorenz curve for $\alpha = 1.1$ and $\alpha = 1.5$ on a single set of axes (I recommend using graphing software such as Desmos; remember p is a proportion, which means the domain for $L(p)$ is the interval $[0, 1]$).

c) Calculate $L(0.8)$ for both values of α and explain what these numbers mean (e.g. if $\alpha = 1.1$, the poorest 80% have ...).

Definition. The **Gini coefficient** of a distribution is defined to be $1 - 2B$ where B is the area under the Lorenz curve: $B = \int_0^1 L(p)dp$.

The Pareto distribution has a Gini coefficient that approaches 1 as α approaches 1 from the right; this corresponds to all the wealth in Extremistan being in the hands of just one person (**maximum inequality**). The Gini coefficient approaches 0 as α increases to ∞; this corresponds to a completely equal distribution of wealth (**minimum inequality**). People (at places like the World Bank and the CIA) actually care about Gini coefficients and use them as a way of measuring income inequality. Wikipedia's list of countries ordered by Gini coefficient is a good summary: `http://en.wikipedia.org/wiki/List_of_countries_by_income_equality`.

5. Each Gini coefficient corresponds to a different value of the parameter α in the Pareto distribution. And each value of α corresponds to a different version of the Pareto principle. For example, the World Bank estimate of the Gini coefficient for US family income in 2016 was 0.411. This corresponds to a parameter of $\alpha \approx 1.7165$ and a 65-35 rule (approximately). Choose your favorite country (other than the US) and look up the the Gini coefficient for income for your country. The calculate the corresponding Pareto parameter α and Use this to find the proportion of the country's wealth owned by the poorest 80%, $L(0.8)$.

Challenge. Find the version of the 80-20 rule that holds for your country: is it an 80-20 rule, a 65-35 rule, or something else?