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

1. Wealth distribution in the country of Extremistan follows the Pareto principle rule strictly: the richest 20% have 80% of the wealth. Moreover, the Pareto principle applies to the richest 20%: the top 20% of this group controls 80% of the wealth of the group. This means that the richest 4% (0.2×0.2) of Extremistanis have 64% (0.8×0.8) of the wealth.

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

b) The Pareto principle also applies to the poorest 80% of Extremistanis (who have 20% of the wealth); the poorest 64% have only what percent of the wealth?

2. In another country, 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 $100p$ percent of Mediocristanis is

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

where $z(p)$ is the $100p^{\text{th}}$percentile of the standard normal distribution. Compare the percentage of Mediocristan's wealth controlled by the poorest 64%, the richest 20%, and the richest 4% of Mediocristanis to the wealth of their counterparts in Extremistan.

The *Lorenz curve* of a wealth distribution is the curve describing the proportion of wealth owned by the bottom $100p\%$ of the population. If we think of wealth as a random variable X, then the Lorenz curve is

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

where $x(p)$ is the $100p^{\text{th}}$percentile of X.

In Mediocristan X is normally distributed with mean 6 and standard deviation 1, so it follows that $L(p) = \dfrac{1}{6}\displaystyle\int_{-\infty}^{x(p)} x \, \mathrm{n}(x; 6, 1)dx$.

Making the substitution $z = x - 6$ gives $L(p) = \dfrac{1}{6}\displaystyle\int_{-\infty}^{z(p)} (z+6)\left(\frac{1}{\sqrt{2\pi}}\right) e^{-\frac{1}{2}z^2} dz$ which we evaluate to find the function in equation 1. The graph of the function $L(p)$ in equation 1) is the Lorenz curve for Mediocristan.

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

$$f(x; \alpha, \beta) = \begin{cases} \alpha\beta^{\alpha}x^{-(\alpha+1)} & x > \beta \\ 0 & x \leq \beta \end{cases}$$

3. Suppose that X has a Pareto distribution with parameters $\alpha > 1$ and $\beta = 1$. Find a formula for the Lorenz curve $L(p)$ and graph the Lorenz curve for different values of α. Calculate $L(0.8)$ for your different values of α and explain what these numbers mean.

4. Calculate $\lim_{\alpha\to\infty} L(p)$ and $\lim_{\alpha\to 1^+} L(p)$ and explain what these mean.

The Gini coefficient of a distribution is defined to be $1 - 2B$ where B is the area under the Lorenz curve (over the interval $[0, 1]$). Our 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. Wikipedia's list of countries ordered by Gini coefficient is a good summary: `http://en.wikipedia.org/wiki/List_of_countries_by_income_equality`.