The Lorenz Curve and Gini Index

Name **Student ID**

The **Lorenz curve** is the curve that represents the income (or wealth) distribution of a society. Consider a graph where the x-axis represents the percentage of the population and the y-axis the percentage of income. The point (x, y) is on the Lorenz curve, $y = L(x)$, if the bottom $x\%$ of the population have $y\%$ of total income of the society.
By definition, we know that every Lorenz curve has the following properties:

- $L(0) = 0$, $L(100) = 100$.
- $y = L(x)$ is non-decreasing.

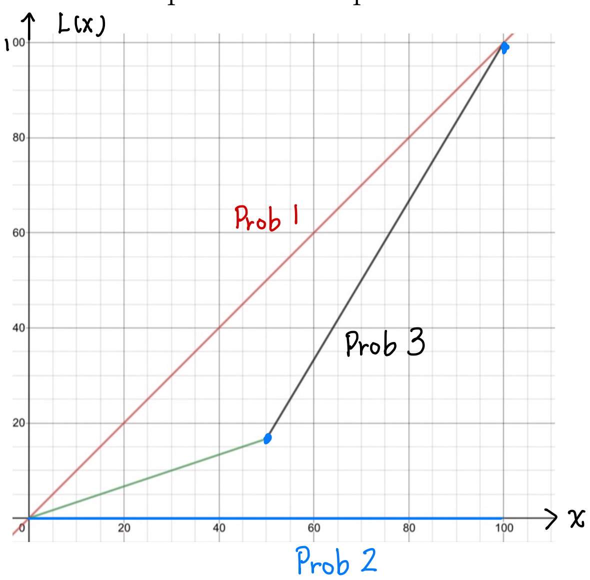
In this worksheet, you will explore some examples and investigate more interesting properties about Lorenz curves.

1. Suppose that everyone has the same income. What percentage of income would the bottom $x\%$ of the population have? Write down the Lorenz curve equation, $L(x)$, for this perfectly equal distribution. Suppose that everyone has income I, and the total population is P_0. Then total income of the society is $I \cdot P_0$. The bottom $x\%$ has income $I \cdot \frac{x}{100} \cdot P_0$ which is the $x\%$ of the total income. This means that $L(x) = x$ for $0 \le x \le 100$.

2. Suppose that one person has all the income and everyone else has none. Write down the Lorenz curve equation, $L(x)$, for this perfectly unequal distribution.
For $x < 100$, the bottom $x\%$ has zero income which is 0% of total income. Hence $L(x) = 0$ for $0 \le x < 100$. $L(100) = 100$ by definition. Thus $L(x) = \begin{cases} 0, & \text{for } 0 \le x < 100 \\ 100, & \text{for } x = 100 \end{cases}$

3. Suppose that the top 50% and the bottom 50% of the population each have the same amount of income (within each group), and the top 50% have income which is 5 times the bottom's income. Write down the Lorenz curve equation.
Suppose that the total population is P_0 and each of the bottom 50% earns I_0. Then each of the top 50% earns $5I_0$. The total income of the society is $P_0 \times \frac{1}{2} \times I_0 + P_0 \times \frac{1}{2} \times 5I_0$ $= 3P_0I_0$. For $0 \le x \le 50$, the bottom $x\%$ has income $P_0 \times \frac{x}{100} \times I_0$ which is $\frac{P_0 \frac{x}{100} I_0}{3P_0I_0} \times 100 = \frac{x}{3}$ % of total income. For $x > 50$, the bottom $x\%$ has income $P_0 \cdot \frac{1}{2} \cdot I_0 + P_0 \cdot \frac{x-50}{100} 5I_0$ $= P_0I_0\left(\frac{x}{20} - 2\right)$ which is $\left(\frac{5}{3}x - \frac{200}{3}\right)\%$ of total income. Hence $L(x) = \begin{cases} \frac{x}{3}, & \text{for } 0 \le x \le 50 \\ \frac{5}{3}x - \frac{200}{3}, & \text{for } x > 50 \end{cases}$

4. Draw the Lorenz curves of problem 1 to problem 3 in the same figure.



Suppose that the total population is P_0 and the total income is I_0.

5. Show that every Lorenz curve is under the Lorenz curve of the perfectly equal distribution.

We want to show that $L(x) \le x$ for $0 \le x \le 100$. Suppose that $L(x_0) > x_0$ for some x_0. Thus the bottom x_0% earns $L(x_0)$% $> x_0$% of total income and the rest (richer) $100-x$% earns $100 - L(x_0)$% $< 100 - x_0$% of total income. This is impossible since the average income of the x_0% bottom is $\frac{L(x_0)}{x_0} \cdot \frac{I_0}{P_0} > \frac{x_0}{x_0} \cdot \frac{I_0}{P_0} = \frac{I_0}{P_0}$ but the average income of the top $100-x$% is $\frac{100-L(x_0)}{100-x_0} \cdot \frac{I_0}{P_0} < \frac{100-x_0}{100-x_0} \cdot \frac{I_0}{P_0} = \frac{I_0}{P_0}$. Thus there is no such x_0 with $L(x_0) > x_0$.

6. If a Lorenz curve is differentiable, the slope of the Lorenz curve can be used to compare income. From the definition of derivatives, explain the meaning of $\frac{L'(x_1)}{L'(x_2)}$ where x_1, x_2 are any points between 0 and 100. Show that the Lorenz curve is concave upward.

$\frac{L(x_1+h) - L(x_1)}{h} \times \frac{I_0}{P_0}$ = the total income of people between the bottom x_1% and x_1+h%.

As $h \to 0$, we get $L'(x_1) \cdot \frac{I_0}{P_0}$ = the income of the x_1-th percentile person.

Hence $\frac{L'(x_1)}{L'(x_2)} = \frac{\text{the income of the } x_1\text{-th percentile person}}{\text{the income of the } x_2\text{-th percentile person}}$.

If $x_1 > x_2$, the x_1-th percentile earns more than the x_2-th percentile. Hence $\frac{L'(x_1)}{L'(x_2)} \ge 1$ i.e. $L'(x_1) \ge L'(x_2)$ if $x_1 > x_2$. Thus $L(x)$ is concave upward.

The **Gini index** is the ratio of the area between the line of perfect equality and the observed Lorenz curve to the area between the line of perfect equality and the line of perfect inequality.

- The Gini index measures how *far* the Lorenz curve is from perfectly equality.
- The Gini index is between 0 and 1. The higher the index, the more unequal the society is.

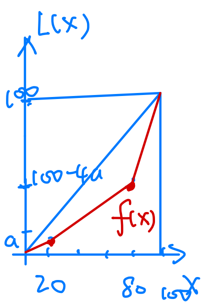
1. Compute the Gini index for the Lorenz curve of problem 3.

The area between Lorenz curves of Prob 3 and perfect equality is $\frac{5000}{3}$

The area between the perfect equality and perfect inequality is 5000.

Hence the Gini index is $\frac{5000}{3} / 5000 = \frac{1}{3}$.

2. There are other indices that describe income inequality. For example, the **20:20 Ratio** compares how much richer the top 20% of the population are compared to the bottom 20% of the population. Suppose the richest 20% earn 4 times the poorest 20%, what is the smallest possible value for the Gini index? Construct an example where the 20:20 Ratio is 4 but the Gini index is greater than 0.5.



If the bottom 20% earn a% of total income, then the top 20% earn $4a$% of total income. Thus $L(20) = a$, $L(80) = 100 - 4a$.

Because $L(x)$ is concave upward, $L(x)$ is below the piecewisely linear function

$$f(x) = \begin{cases} \frac{a}{20}x, & \text{for } 0 \le x \le 20 \\ a + \frac{100-5a}{60}(x-20), & \text{for } 20 < x \le 80 \\ 100-4a + \frac{a}{5}(x-80), & \text{for } 80 < x \le 100. \end{cases}$$

which passes points $(0,0)$, $(20,a)$, $(80, 100-4a)$, and $(100,100)$.

The area between $y=x$ and $y=f(x)$ is $\boxed{120a}$. Hence the area between $y=x$ and $y=L(x)$ is greater than or equal to $120a$.

However, there are constraints on the value "a" $\because$ $L(x)$ is concave upward,

$\because$ The average slope of $L(x)$ between $x \in [0, 20]$

$\leq$ the average slope of $L(x)$ between $[20, 80]$

$\leq$ the average slope of $L(x)$ between $[80, 100]$.

i.e. $$\frac{a}{20} \leq \frac{100-5a}{60} \leq \frac{4a}{20}$$

$$\Rightarrow \quad \frac{100}{17} \leq a \leq \frac{25}{2}.$$

Hence the area between $y=x$ and $y=L(x) \geq 120a \geq 120 \times \frac{100}{17}$.

The Gini index $\geq \dfrac{\frac{12000}{17}}{5000} = \dfrac{12}{85} \approx 0.142$.

The smallest Gini index occurs when $a = \frac{100}{17}$, and the Lorenz curve is the piecewisely linear function

$$L(x) = \begin{cases} \frac{5}{17}x, & 0 \leq x \leq 20 \\ \frac{20}{17}x - \frac{300}{17}, & 20 < x \leq 100. \end{cases}$$

It seems that for 20:20 Ratio $= 4$, the Gini index can not be greater than 0.5. But we can construct an example (with $a = \frac{25}{2}$)

$$L(x) = \begin{cases} \frac{5}{8}x & \text{for } 0 \leq x \leq 80. \\ \frac{5}{2}x - 150, & \text{for } 80 < x \leq 100 \end{cases}$$

with the Gini index 0.3.

It's an interesting question whether 0.3 is the maximum Gini index.