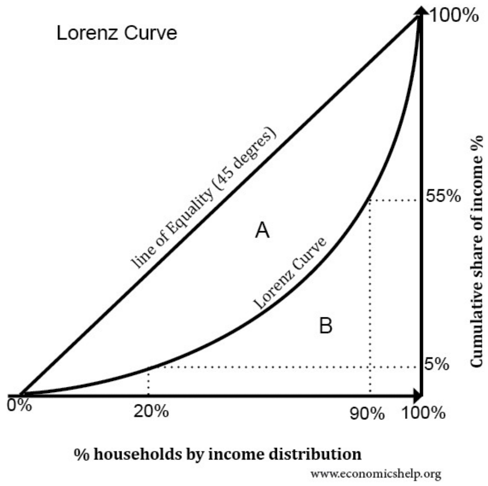
Axis 7: Lorenz curve and Gini Coefficient

1. Definition: The Lorenz curve is a way of showing the distribution of income (or wealth) within an economy. It was developed by Max O. Lorenz in 1905 for representing wealth distribution.

The Lorenz curve shows the cumulative share of income from different sections of the population.

If there was perfect equality – if everyone had the same salary – the poorest 20% of the population would gain 20% of the total income. The poorest 60% of the population would get 60% of the income.

2. Diagram of Lorenz curve



In this Lorenz curve, the poorest 20% of households have 5% of the nation's total income.

The poorest 90% of the population holds 55% of the total income. That means the richest 10% of income earners gain 45% of total income.

3. Explanation of Lorenz Curve

The Lorenz curve is represented by a straight diagonal line, which represents perfect equality in income or wealth distribution; the Lorenz curve lies beneath it, showing estimated distribution. The area that is between the straight line and the curved line is the Gini coefficient. The Gini Coefficient itself is expressed as a representation of the scalar measurement of inequality. In the Lorenz Curve, the Gini Coefficient is expressed as the ratio of the area under the straight line

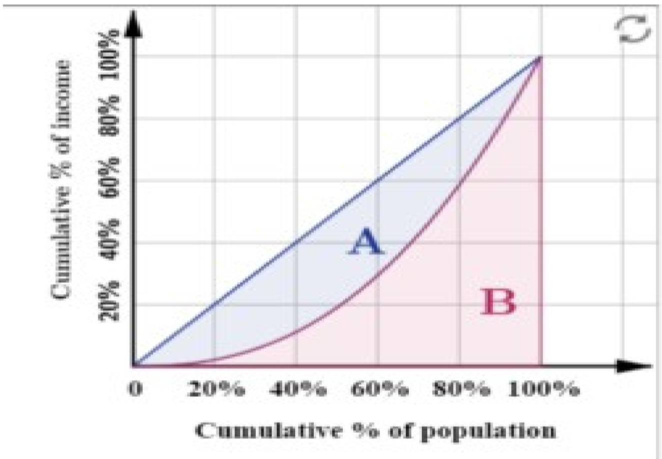
The Lorenz curve is used to represent economic inequality as well as unequal wealth distribution. The farther away the curved line is way from the straight diagonal line, the higher the level of inequality.

Constructing a Lorenz curve involves fitting a continuous function to some incomplete set of data, there is no guarantee that the values along a Lorenz curve (other than those actually observed in the data) actually correspond to the true distributions of income.

Most of the points along the curve are just guesses based on the shape of the curve that best fits the observed data points. So the shape of the Lorenz curve can be sensitive to the quality and sample size of the data and to the mathematical assumptions and judgments as to what constitutes the best fit curve, and these may represent sources of substantial error between the Lorenz curve and the actual distribution.

Solve previous years UPSC Mains Economics Questions by visiting the linked article

For better clarity, the Lorenz curve will be better elaborated upon through the image below:



The curve (B) is a graph showing the proportion of overall income or wealth assumed by the bottom x% of the people, although this is not rigorously true for a finite population. It is often used to represent income distribution, where it shows for the bottom x% of households, what percentage, represented by a straight line – A, (y%) of the total income they have. The percentage of households is plotted on the x-axis, the percentage of income on the y-axis. It can also be used to show the distribution of assets. In such use, many economists consider it to be a measure of social inequality.

4. The Gini coefficient

The Gini coefficient is an index for the degree of inequality in the distribution of income/wealth, used to estimate how far a country's wealth or income distribution deviates from an equal distribution.

The Gini coefficient is usually defined mathematically based on the Lorenz curve, which plots the proportion of the total income of the population (y-axis) that is cumulatively earned by the bottom x of the population (see diagram). The line at 45 degrees thus represents perfect equality of incomes. The Gini coefficient can then be thought of as the ratio of the area that lies between the line of equality and the Lorenz curve (marked A in the diagram) over the total area under the line of equality (marked A and B in the diagram); i.e., $G = A/(A + B)$. If there are no negative incomes, it is also equal to $2A$ and $1 - 2B$ due to the fact that $A + B = 0.5$.

There are several formulas for calculating the GINI coefficient. We will adopt one of the most used which is given in the explanatory note of the World Bank for calculating the inequality of distributions24, sometimes called the “BROWN formula”.
This formula is written:

$$G = 1 - \sum_{i=1}^{n} \left(X_i - X_{i-1} \right) \left(Y_i + Y_{i+1} \right)$$

where X is the cumulative share of the population, and Y is the cumulative share of the mass to to share out. In the case that interests us here, as the data is known individually this formula can be simplified to:

$$G = 1 - \frac{1}{n} \sum_{i=1}^{n} \left(Y_i + Y_{i+1} \right)$$

Where n represents the number of statistical units (the population).
We will see that these two formulas give identical results.