



PERIYAR
MANIAMMAI

INSTITUTE OF SCIENCE & TECHNOLOGY

(Deemed to be University)

Established Under Sec. 3 of UGC Act, 1956 • NAAC Accredited

think • innovate • transform

Name : Aadhilafemine AJ.

Reg.no :121011012715.

Subject : data science.

Dept :B.TECH(CSE) with
specialization in data science.

What Is a Lorenz Curve?

A Lorenz curve, developed by American economist Max Lorenz in 1905, is a graphical representation of income equality or wealth inequality. The graph plots percentiles of the population on the horizontal axis according to income or wealth and plots cumulative income or wealth on the vertical axis.

KEY TAKEAWAYS

- A Lorenz curve is a graphical representation of the distribution of income or wealth within a population.
- Lorenz curves graph percentiles of the population against cumulative income or wealth of people at or below that percentile.
- Lorenz curves, along with their derivative statistics, are widely used to measure inequality across a population.
- The Lorenz curve is a central piece in calculating the Gini coefficient, a mathematical representation of inequality levels.
- Because Lorenz curves are mathematical estimates based on fitting a continuous curve to incomplete and discontinuous data, they may be imperfect measures of true inequality.

Understanding the Lorenz Curve

In practice, a Lorenz curve is usually a mathematical function estimated from an incomplete set of observations of income or wealth. The Lorenz curve is often accompanied by a straight diagonal line with a slope of 1, which represents perfect equality in income or wealth distribution; the Lorenz curve lies beneath it, showing the observed or estimated distribution.

While the Lorenz curve is most often used to represent economic equality, it can also demonstrate unequal distribution in any system. The farther the curve is from the baseline, represented by the straight diagonal line, the higher the level of inequality. In economics, the Lorenz curve denotes inequality in the distribution of either wealth or income. These are not synonymous since it is possible to have either high earnings but zero or negative net worth, or low earnings but a large net worth.

A Lorenz curve usually starts with an empirical measurement of wealth or income distribution across a population based on data such as tax returns, which report income for a large portion of the population. A graph of the data may be used directly as a Lorenz curve, or economists and

statisticians may fit a curve that represents a continuous function to fill in any gaps in the observed data.

The closer the Lorenz curve is to the line of equality, the less inequality exists. The further the Lorenz curve is to the line of equality, the more inequality exists.

Components of the Lorenz Curve

The United States Federal Reserve collected net worth statistics from U.S. households. It then graphically depicted the inequality distribution of wealth. The latest data collected by the Survey of Consumer Finances is from 2019. Per the Federal Reserve, the underlying data demonstrates that the bottom 50% of households hold just 1.5% of overall household wealth.

There are several important components to understand when analyzing a Lorenz curve:

- The *x-axis* is often denoted as the percentile. In the graph above, the *x-axis* is the percentile of net worth ranking compared to other U.S. households.
- The *y-axis* is often denoted as the cumulative percentage of occurrences. In the graph above, the percentages represent the cumulative amount of net worth of households.
- The *line of equality* is demonstrated by a 45-degree, upward-sloping line. In the graph above, it is denoted as the dashed line.
- The *Lorenz curve* is demonstrated often by an upward-sloping but often exponentially rising curve. In the graph above, it is denoted as a solid line.
- The Gini coefficient (discussed below) is the gap between the line of equality and the Lorenz curve.

The Lorenz Curve and the Gini Coefficient

The Gini coefficient is used to express the extent of inequality in a single figure. It most often ranges from 0 (or 0%) to 1 (or 100%). Complete equality, in which every individual has the exact same income or wealth, corresponds to a coefficient of 0. Plotted as a Lorenz curve, complete equality would be a straight diagonal line with a slope of 1 (the area between this curve and itself is 0, so the Gini coefficient is 0).

A coefficient of 1 means that one person earns all of the income or holds all of the wealth. In theory, the Gini coefficient can exceed 100% in

extreme situations. For example, when handling negative wealth or income, the figure can theoretically be higher than 1; in that case, the Lorenz curve would dip below the horizontal axis.

The Gini coefficient is equal to the area below the line of perfect equality minus the area below the Lorenz curve, divided by the area below the line of perfect equality. In the graph above, the Gini coefficient is the area below the dashed line but above the solid line. The Gini coefficient is used to measure the extent of inequality. It can also be used to compare two different nations or countries to see which has more inequality.

For data sets demonstrating inequality, it may be difficult to properly analyze certain sections of data. For example, the graph above has a very long left-skewed tail. This makes it difficult to examine changes in equality for this section of data.

Advantages and Disadvantages of the Lorenz Curve

Advantages of the Lorenz Curve

A Lorenz curve gives more detailed information about the exact distribution of wealth or income across a population than summary statistics such as the Gini coefficient or the Lorenz asymmetry coefficient. Because a Lorenz curve visually displays the distribution across each percentile (or other unit breakdowns), it can show precisely at which income (or wealth) percentiles the observed distribution varies from the line of equality and by how much.

The underlying data for a Lorenz curve is necessary to calculate the Gini coefficient, a primary means of mathematically measuring inequality. Assuming sufficient data can be collected, a Lorenz curve can be prepared for any geographical region (i.e. by country, by state, by city, etc.). This may be especially important for governments relying on a Lorenz curve to better understand how to most appropriately assess tax brackets.

Though the Lorenz curve often displays sensitive financial information, it also maintains the anonymity of the underlying surveyed individuals. In addition, the Lorenz curve can be layered on top of itself over time, demonstrating how public policy has changed the level of inequality at various points along the Lorenz curve.

Disadvantages of the Lorenz Curve

Because 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. This is also true because it may be too expensive or demand too many resources to collect enough data points to fill all gaps.

Most of the points along the curve are just guesses based on the shape of the curve that best fits the observed data points. Therefore, 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 a best-fit curve, and these may represent sources of substantial error between the Lorenz curve and the actual distribution.

The Lorenz curve (or Gini coefficient) may also be insufficient to clearly demonstrate the extent of change of inequality over time. Consider a situation where two Gini coefficients are calculated to be the exact same. It is possible for the slope of the Lorenz curve to be radically different, yet the underlying aggregated inequality level is calculated to be the same.

Pros

- Visually depicts inequality across a population in a manner easy to understand and analyze
- Is used to help calculate the Gini coefficient, a primary mathematical mean of calculating inequality
- May assist governments in making public policy changes or impacting tax bracket ranges based on income
- Maintains anonymity of surveyed individuals
- May be compiled to show how the curve has changed over time

Cons

- Sample data may not appropriately reflect the overall population, therefore displaying an incorrect Lorenz curve
- May require extensive data collection to adequately fill in the entire curve
- May require estimation or preparer inference as to the curve to draw
- May mislead analysts due to the varying shapes and sizes of the Gini coefficient area; different areas may be equal in size yet vary in appearance

Lorenz Curve Example

The curve above shows a continuous Lorenz curve that has been fitted to data that describes income distribution in Brazil as of 2020. The data set is also compared to a straight diagonal line representing perfect equality.

At the 55th income percentile, the value of the Lorenz curve is 22.39%. This means that the Lorenz curve estimates that the bottom 55% of the population takes in 22.39% of the nation's total income. If Brazil were a perfectly equal society, the bottom 55% would earn 55% of the total.

Elsewhere, we can see that the 99th percentile corresponds to 89.32% in cumulative income. This means that the top 1% takes in 11.68% of Brazil's income.

To find the approximate Gini coefficient, subtract the area beneath the Lorenz curve (around 0.25) from the area beneath the line of perfect equality (0.5 by definition). Divide the result by the area beneath the line of perfect equality, which yields a coefficient of around 0.5 or 50%. According to the world bank, Brazil's Gini coefficient was 48.9 in 2020.

Why Is the Lorenz Curve Important?

The Lorenz curve is important because it represents one of the best and simplest ways to illustrate the level of economic inequality in society. As the Lorenz curve moves away from the baseline, the underlying data suggests that the unequal distribution keeps increasing.

How Does the Lorenz Curve Measure Inequality?

The Lorenz curve is a graphical representation of the distribution of income or wealth in a society. Basically, the farther the curve moves from the baseline, represented by the straight diagonal line, the higher the level of inequality.

Who Uses Lorenz Curves?

Government agencies are especially interested in Lorenz curves, especially for net worth and income distributions within their country. Lorenz curves inform governments of how public policy is working (or not working). It may also be an indicator of how a government should establish its tax brackets based on gaps or ranges of income.

How Is a Lorenz Curve Calculated?

After collecting data from a relatively large sample size, a Lorenz curve is then fitted to best demonstrate the distribution of that data set. Because each set of information is different, there is no single universal curve formula, and the nonlinear regression will often look very different for different data sets.

The Bottom Line

A common method of visually analyzing inequality distributions is to use a Lorenz curve. The curve graphically depicts what a standard distribution of equality would look like. Then, the graph layers on the actual distribution of a given data set. The Lorenz curve is used to calculate the Gini coefficient to mathematically analyze how unequal a dataset is, and it is used to visually exam inequality levels as well.