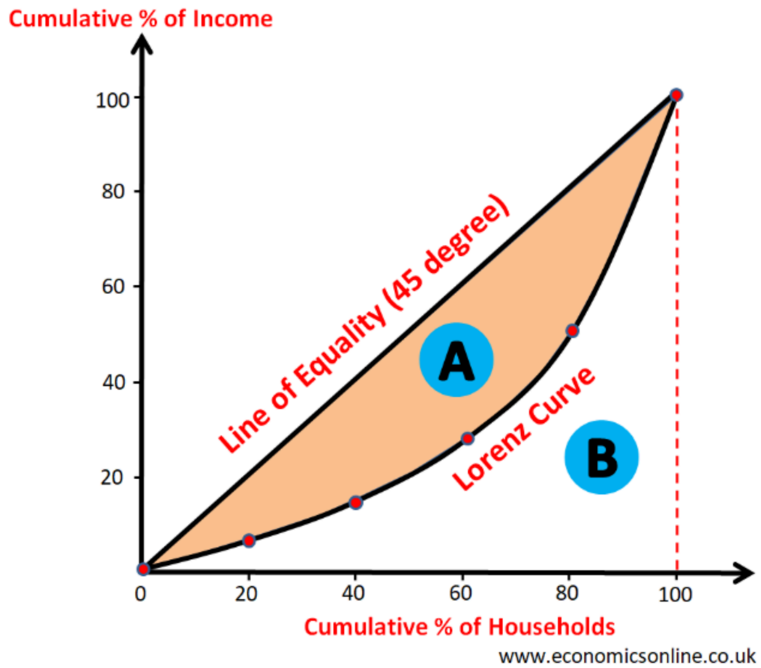
How Effectively Does the GINI Coefficient Measure Inequality?

What does it actually mean for a society to be “unequal”? People use this term to describe the world so effortlessly, and usually never provide any evidence suggesting this claim; they usually just complain about how troubling life is. However, what if I told you there was actually a measure to calculate equality.

Picture this, there’s two countries, with the exact same GDP. In one of the countries, every single person earns the exact same amount; however, in the other country, one person earns all the money, despite the fact that you are working. Intuitively, you would assume country one is the better option; however you missed out on the fact that you could be working a full-time contract and you would still earn the same as a newborn baby. Mathematically, we can represent this in a format called the Gini Coefficient, where 0 represents complete equality and 1 represents maximum inequality



The Lorenz curve can showcase this in a way. On the x-axis, we have the number of households, on the y-axis, we can see how the total GDP is split within the country. The line of equality showcases that 50% of the population would earn exactly 50% of the income, however, the Lorenz curve is below the line of equality, for example if we were to interpolate the first 50% of households here, they only make up for roughly 20% of the country’s entire GDP. The greater the area is between the line of equality and the Lorenz curve, the more inequality it shows. One key thing we can visually see is that the rich keep getting richer, whilst people on lower incomes struggle more. The area here shows the GINI coefficient. Therefore, by analysing the shape and the position of the Lorenz curve, it becomes possible to both visualise and quantify differences in income distribution.

So far, whilst reading this essay, you might have thought about the fact that we don't need a society where everyone is equal, after all, that is what communism is. On the other hand, we should comprehend the fact that people should get paid for how hard they work and the importance they have on society. We can measure this impact using the double summation formula.

$$G = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} |x_i - x_j|}{2n^2\mu}$$

- Xi is the income of person i
- Xj is the income for person j
- n is the number of people
- Mod of xi - xj is the absolute difference between two incomes
- μ is the mean income

This formula measures how different everyone's income is from one another, the reason for the double summation means you can compare every person with one another. The absolute difference lets us see how large the gap is; it doesn't matter if it is positive or negative, it depends on the magnitude of the gap.

For the sum of i and j = 1, we would: choose person i, compare their income with j and repeat this for all individuals, in theory this would look like:

| x1-x1 | + | x1 - x2 | + … + | x1 - xn |

then

| x2 - x1 | + | x2 - x2 | + … | x2 - xn |

this means that there are n^2 comparisons, imagine there are 4 people the comparisons would be

1v1, 1v2, 1v3, 1v4, 2v1, 2v2 … 4v4 so 4^2 comparisons so 16 comparisons

The denominator contains 2 to ensure the GINI coefficient lies between 0 <= G <= 1, as we are double counting, for example 1v2 is the same as 2v1, this is to remove that. We also divide by the mean ensuring the Gini coefficient stays unchanged, this is a property called scale invariance. Governments and international organisations, like the World Bank, allow countries to track trends in inequality. Policymakers can simulate changes, as they can suppose a government subsidises low income citizens, also potentially add tax brackets, where very low income workers don’t pay tax at all, it can also compare city vs rural salaries and measure wealth concentration.

The double summation formula reveals that the GINI coefficient is fundamentally based on pairwise income comparisons, measuring the inequality as the average absolute difference between the incomes of all possible pairs of individuals in a population, normalised by the mean income.

$$G = \frac{1}{2\mu} \int_0^\infty \int_0^\infty |x - y| f(x) f(y) \, dx \, dy$$

Another way is through the continuous distribution version, which can be represented on the Lorenz curve. Pick two independent values from the curve “X” and “Y”, however this isn’t the best to represent it, as it depends of the scale of income, meaning that if someone were to get a payrise, it assumes everyone gets a pay rise, which is not the case, so it’s not a fair measure of inequality.

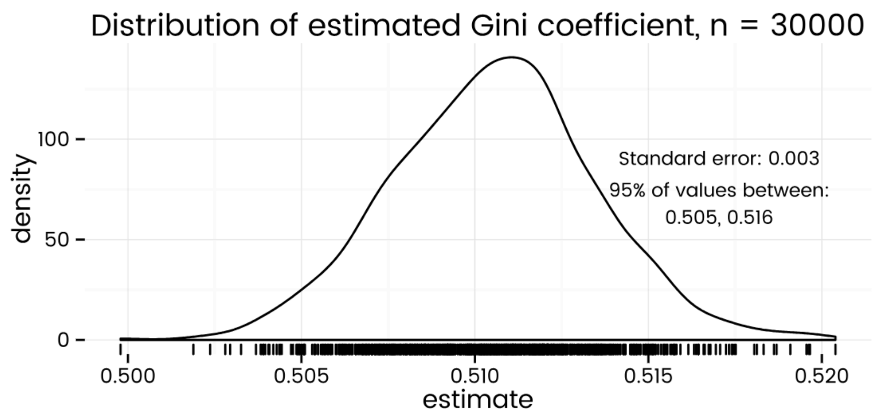
Consequently, we divide by ½ Mu, as by dividing it by Mu makes it relative to the actual population size, but we divide it by 2 Mu, as we’re comparing 2 people, therefore it shows the maximum possible difference between these two people, as we may have instances where one person has everything, meanwhile another has 0. In addition, the reason we use 2 is so that it lies between 0 and 1, where the Gini coefficient is measured in, if we were to use something else, the scales would break. To put it simply, the numerator is “how far apart people are” and the denominator is “how far could they possibly be”.

The reason for the double integral is because we are comparing two variables, hence why we integrate is with respect to both x and y. However, the reason we multiply f(x) with f(y) is stemmed from the probability theory, where f(x) is picking value x and f(y) is probability of picking value y, they are independent, so we multiply them, this gives us a greater chance of picking a pair near (x,y). Therefore in short, the double integral averages pairs of people.

Overall, the continuous distribution model does show spread between two variables, however it has drawbacks, as firstly, it is very sensitive to values near the middle, however, it can underweight extreme top end high income earners, therefore we could get GINI coefficient values for very rich elites and moderately unequal middle values to be the same. In addition, it requires the mean of the distribution and variance to be finite, as suppose Mu was infinite, GINI would become meaningless, as it would automatically show the value to be 0, and furthermore we would want the variance to be infinite, as this shows on the Lorenz curve that the distribution has extreme outliers, and in the real world, this is very common, as some people are substantially below the poverty line, meanwhile others are very rich. Therefore, it excludes heavy-tailed models.

Additionally, the GINI coefficient has a lot of advantages, for example, it provides a very simplistic scale, which allows comparisons between countries, this makes it very clear to us from an analytic point of view and easy to understand. It also has scale independence and is representative to a countries' population, for example, if it quickly rises from 0.3 to 0.4, it could lead to government intervention. In addition, it considers the entire income distribution, imagine it being modelled as a normal distribution, rather than just focusing on the top 10% wealthiest and poorest people in each country, so it's every country in relation to itself.

On the contrary, the GINI coefficient has as many cons as it does have pros, firstly, countries have completely different policies to one another, for example in the UK you can have access to free healthcare, strong welfare support and affordable housing but in Niger this is not the case, with this in mind, despite the income inequality for both countries being similar, as their GINI coefficients are about around 0.34, why is that people don't choose to live in Niger when inequality is the same there as it is in the UK? Moreover, GINI also ignored the strength of a countries' currency, a country could have low inequality but still have low purchasing power, if their currency is weak, which prevents them from buying as many oil reserves, they would struggle to import foreign goods, making inhabitants limited to choice, which in theory, should actually increase inequality. Finally, going back to the point about the GINI coefficient looking like a normal distribution, this would mean it's bell shaped meaning it would look very symmetrical, this doesn't show us where the inequality actually is, whether it could be between upper and middle class or middle and lower class, so if they government were to intervene by adding tax brackets, it may cause more inequality if they were to tax the lower class more, and the rich may not be affected that much by it. GINI only gives the overall inequality.



To add to my previous point earlier of the GINI coefficient being simple to understand, this truth disguises the complex reality that some people from lower economic backgrounds are born “inequal”. The model doesn’t show the types of inequality, whether it be earned or inherited, it may be used to hide the regional disparities. Stephen J Gould once said “I am, somehow, less interested in the weight and convolution of Einstein’s brain than in the near certainty that people of equal talent have lived and died in the cotton fields and sweatshops”. Overall, the simplification can obscure important ethical nuances.

In conclusion, the GINI coefficient’s strength lies in its mathematical simplicity by having a scale that only spans from 0 to 1, the integration used to find the area under the Lorenz Curve and the line of inequality reflects how each individual’s share of income contributes to overall inequality. However, this does not paint the whole picture due to the lack of third party information we have whilst doing these calculations, extreme poverty or wealth may not be fully demonstrated. Ultimately, the GINI coefficient demonstrates how mathematical abstraction can simplify social realities, providing a useful but inherently partial lens on economic inequality.

Bibliography:

https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/families/methodologies/theginicoefficient

https://www.economicsonline.co.uk/definitions/thelorenzcurve.html/

https://onlinelibrary.wiley.com/doi/10.1111/roiw.12256

https://pubmed.ncbi.nlm.nih.gov/25733724/

https://academic.oup.com/pnasnexus/article/3/7/pgae221/7687932

YouTube videos:

https://youtu.be/BwSB__Ugo1s?si=O4kgGLSupvoYdzsM

https://youtu.be/UcyOiVuzAQY?si=WAh_oIS6fZJRaBQy

https://youtu.be/jqN7pIJrjNo?si=vnFo9aUeLMMIFRFX