Gini Coefficient

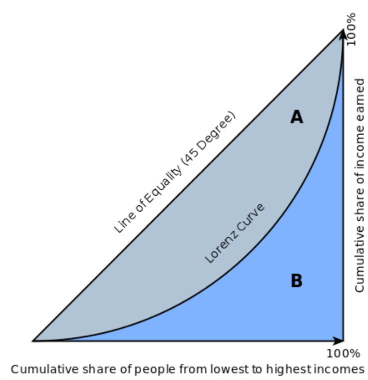
The Gini coefficient (sometimes expressed as a Gini ratio or a normalized Gini index) is a measure of statistical dispersion intended to represent the income or wealth distribution of a nation's residents, and is the most commonly used measure of inequality. It was developed by the Italian statistician and sociologist Corrado Gini and published in his 1912 paper *Variability and Mutability* (Italian: *Variabilità e mutabilità*).[1][2]

The Gini coefficient measures the inequality among values of a frequency distribution (for example, levels of income). A Gini coefficient of zero expresses perfect equality, where all values are the same (for example, where everyone has the same income). A Gini coefficient of 1 (or 100%) expresses maximal inequality among values (e.g., for a large number of people, where only one person has all the income or consumption, and all others have none, the Gini coefficient will be very nearly one).[3][4] However, a value greater than one may occur if some persons represent negative contribution to the total (for example, having negative income or wealth). For larger groups, values close to or above 1 are very unlikely in practice. Given the normalization of both the cumulative population and the cumulative share of income used to calculate the Gini coefficient, the measure is not overly sensitive to the specifics of the income distribution, but rather only on how incomes vary relative to the other members of a population. The exception to this is in the redistribution of wealth resulting in a minimum income for all people. When the population is sorted, if their income distribution were to approximate a well-known function, then some representative values could be calculated.

There are some issues in interpreting a Gini coefficient. The same value may result from many different distribution curves. The demographic structure should be taken into account. Countries with an aging population, or with a baby boom, experience an increasing pre-tax Gini coefficient even if real income distribution for working adults remains constant. Scholars have devised over a dozen variants of the Gini coefficient.[12][13][14]

1 Definition

The graph shows that the Gini coefficient is equal to the area marked *A* divided by the sum of the areas marked *A* and *B*, that is, Gini = $A / (A + B)$. It is also equal to $2A$ and to $1 - 2B$ due to the fact that $A + B = 0.5$ (since the axes scale from 0 to 1).



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*). It is also equal to 2*A* and to 1 - 2*B* due to the fact that *A* + *B* = 0.5 (since the axes scale from 0 to 1).

If all people have non-negative income (or wealth, as the case may be), the Gini coefficient can theoretically range from 0 (complete equality) to 1 (complete inequality); it is sometimes expressed as a percentage ranging between 0 and 100. In practice, both extreme values are not quite reached. If negative values are possible (such as the negative wealth of people with debts), then the Gini coefficient could theoretically be more than 1. Normally the mean (or total) is assumed positive, which rules out a Gini coefficient less than zero.

An alternative approach would be to consider the Gini coefficient as half of the relative mean absolute difference, which is a mathematical equivalence.[15] The mean absolute difference is the average absolute difference of all pairs of items of the population, and the relative mean absolute difference is the mean absolute difference divided by the average, to normalize for scale. if x_i is the wealth or income of person *i*, and there are *n* persons, then the Gini coefficient *G* is given by:

$$\mathrm{Gini} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\left|x_i - x_j\right|}{2\sum_{i=1}^{n}\sum_{j=1}^{n} x_j} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\left|x_i - x_j\right|}{2n\sum_{i=1}^{n} x_j}$$

When the income (or wealth) distribution is given as a continuous probability distribution function *p(x)*, where *p(x)dx* is the fraction of the population with income *x* to *x+dx*, then the Gini coefficient is again half of the relative mean absolute difference:

$$\mathrm{Gini} = \frac{1}{2\mu}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x)\, p(y)|x - y|dxdy$$

here μ is the mean of the distribution $\mu = \int_{-\infty}^{\infty} f(x)dx$ and the lower limits of integration may be replaced by zero when all incomes are positive.

2 Calculation

In some cases, this equation can be applied to calculate the Gini coefficient without direct reference to the Lorenz curve. For example, (taking *y* to mean the income or wealth of a person or household):

For a population uniform on the values y_i, *i* = 1 to *n*, and $y_{(i)}$, *i* = 1 to *n*, be the order statistics. The Gini index is

$$\mathrm{Gini} = \frac{1}{n}\left(n + 1 - 2\left(\frac{\sum_{i=1}^{n}(n+1-i)y_{(i)}}{\sum_{=1}^{n} y_{(i)}}\right)\right) = \frac{2\sum_{=1}^{n} i \times y_{(i)}}{n\sum_{=1}^{n} y_{(i)}} - \frac{n+1}{n}.$$

Gini Coefficient

Since the Gini coefficient is half the relative mean absolute difference, it can also be calculated using formulas for the relative mean absolute difference. For a random sample S consisting of values $y_{(i)}$, i = 1 to n, be the order statistics of y_i:

$$\mathbf{Gini}(\boldsymbol{S}) = \frac{\mathbf{1}}{\boldsymbol{n}-\mathbf{1}}\left(\boldsymbol{n}+\mathbf{1}-\mathbf{2}\left(\frac{\sum_{i=1}^{n}(\boldsymbol{n}+\mathbf{1}-\boldsymbol{i})\boldsymbol{y}_{(i)}}{\sum_{=1}^{n}\boldsymbol{y}_{(i)}}\right)\right)$$
$$= \mathbf{1}-\frac{\mathbf{2}}{\boldsymbol{n}-\mathbf{1}}\left(\boldsymbol{n}-\frac{\sum_{i=1}^{n}(\boldsymbol{n}+\mathbf{1}-\boldsymbol{i})\boldsymbol{y}_{(i)}}{\sum_{=1}^{n}\boldsymbol{y}_{(i)}}\right)$$

Gini(S) is a consistent estimator of the population Gini coefficient, but is not, in general, unbiased. There does not exist a sample statistic that is in general an unbiased estimator of the population Gini coefficient.

Discrete probability distribution

For a discrete probability distribution with probability mass function $f(y_i)$, i = 1 to n, where $f(y_i)$ is the fraction of the population with income or wealth y_i >0, the Gini coefficient is:

$$\text{Gini} = \frac{1}{2\mu}\sum_{i=1}^{n}\sum_{j=1}^{n} f(y_i)f(y_j)\,|y_i - y_j|$$

, where $\mu = \sum_{i=1}^{n} y_i f(y_i)$.

Continuous probability distribution

When the population is large, the income distribution may be represented by a continuous probability density function f(x) where f(x) dx is the fraction of the population with wealth or income in the interval dx about x. If F(x) is the cumulative distribution function for f(x), then the Lorenz curve L(F) may then be represented as a function parametric in L(x) and F(x) and the value of B can be found by integration: $\mathrm{B} = \int_0^1 L(F)\,dF$.

The Gini coefficient can also be calculated directly from the cumulative distribution function of the distribution F(y). Defining μ as the mean of the distribution, and specifying that F(y) is zero for all negative values, the Gini coefficient is given by:

$$\text{Gini} = 1 - \frac{1}{\mu}\int_0^\infty (1-F(y))^2 dy = \frac{1}{\mu}\int_0^\infty F(y)(1-F(y))\,dy$$

The latter result comes from integration by parts. (Note that this formula can be applied when there are negative values if the integration is taken from minus infinity to plus infinity.)
The Gini coefficient may be expressed in terms of the quantile function $Q(F)$ (inverse of the cumulative distribution function: Q(F(x))=x)

$$\text{Gini} = \frac{1}{2\mu}\int_0^1\int_0^1 |Q(F_1) - Q(F_2)|\,dF_1 dF_2.$$

For some functional forms, the Gini index can be calculated explicitly. For example, if y follows a lognormal distribution with the standard deviation of logs equal to σ, then $G = \operatorname{erf}\left(\frac{\sigma}{2}\right)$ where *erf* is the error function (since $G = 2\Phi\left(\frac{\sigma}{\sqrt{2}}\right) - 1$, where Φ is the cumulative standard normal distribution)**.**[**17**] In the table below, some examples are shown. The Dirac delta distribution represents the case where everyone has the same wealth (or income); it implies that there are no variations at all between incomes.

Income Distribution function	PDF(x)	Gini Coefficient
Dirac delta function	$\delta(x - x_0), x_0 > 0$	0
Uniform distribution	$\begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & otherwise \end{cases}$	$\frac{b-a}{3(b+a)}$
Exponential distribution	$\lambda e^{-x\lambda}, \quad x > 0$	$1/2$
Log-normal distribution	$\frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(ln(x)-\mu)^2}{\sigma^2}}$	$\operatorname{erf}\left(\frac{\sigma}{2}\right)$
Pareto distribution	$\begin{cases} \frac{\alpha k^\alpha}{x^{\alpha+1\alpha}} & x \ge k \\ 0 & x < k \end{cases}$	$\begin{cases} 1 & 0 < \alpha < 1 \\ \frac{1}{2\alpha - 1} & \alpha > 1 \end{cases}$
Chi-squared distribution	$\frac{2^{-k/2} e^{-x/2} x^{k/2-1}}{\Gamma(k/2)}$	$\frac{2\Gamma\left(\frac{1+k}{2}\right)}{k\Gamma(k/2)\sqrt{\pi}}$
Gamma distribution	$\frac{2^{-k/\theta} e^{-x/2} \theta^{-k}}{\Gamma(k)}$	$\frac{\Gamma\left(\frac{1+2k}{2}\right)}{k\Gamma(k)\sqrt{\pi}}$
Weibull distribution	$\frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^{\beta-1}}$	$1 - 2^{-1/k}$
Beta distribution	$\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \quad \beta)}$	$\left(\frac{2}{\alpha}\right)\frac{B(\alpha+\beta, \quad \alpha+\beta)}{B(\alpha, \quad \alpha)B(\beta, \quad \beta)}$

Other approaches

Sometimes the entire Lorenz curve is not known, and only values at certain intervals are given. In that case, the Gini coefficient can be approximated by using various techniques for interpolating the missing values of the Lorenz curve. If (X_k, Y_k) are the known points on the Lorenz curve, with the X_k indexed in increasing order (X_k – **1** < X_k), so that:

- X_k is the cumulated proportion of the population variable, for k = **0,...**,n, with X_0 = **0**, X_n = **1.**
- Y_k is the cumulated proportion of the income variable, for k = **0,...**,n, with Y_0 = **0**, Yn = **1.**

- Y_k should be indexed in non-decreasing order ($Y_k > Y_k - \mathbf{1}$)

If the Lorenz curve is approximated on each interval as a line between consecutive points, then the area B can be approximated with trapezoids and:

$$G_1 = 1 - \sum_{k=1}^{n} (X_k - X_{k-1})\,(Y_k + Y_{k-1})$$

is the resulting approximation for G. More accurate results can be obtained using other methods to approximate the area B, such as approximating the Lorenz curve with a quadratic function across pairs of intervals, or building an appropriately smooth approximation to the underlying distribution function that matches the known data. If the population mean and boundary values for each interval are also known, these can also often be used to improve the accuracy of the approximation.

The Gini coefficient calculated from a sample is a statistic and its standard error, or confidence intervals for the population Gini coefficient, should be reported. These can be calculated using bootstrap techniques but those proposed have been mathematically complicated and computationally onerous even in an era of fast computers. Ogwang (**2000**) made the process more efficient by setting up a "trick regression model" in which respective income variables in the sample are ranked with the lowest income being allocated rank **1.** The model then expresses the rank (dependent variable) as the sum of a constant A and a normal error term whose variance is inversely proportional to y_k;

$$\mathrm{k} = \mathrm{A} + \mathrm{N}(0, s^2/y_k).$$

Ogwang showed that G can be expressed as a function of the weighted least squares estimate of the constant A and that this can be used to speed up the calculation of the jackknife estimate for the standard error. Giles (**2004**) argued that the standard error of the estimate of A can be used to derive that of the estimate of G directly without using a jackknife at all. This method only requires the use of ordinary least squares regression after ordering the sample data. The results compare favorably with the estimates from the jackknife with agreement improving with increasing sample size. [**18**]

However it has since been argued that this is dependent on the model's assumptions about the error distributions (Ogwang 2004) and the independence of error terms (Reza & Gastwirth 2006) and that these assumptions are often not valid for real data sets. It may therefore be better to stick with jackknife methods such as those proposed by Yitzhaki (1991) and Karagiannis and Kovacevic (2000). The debate continues. [citation needed]

Guillermina Jasso [19] and Angus Deaton[20] independently proposed the following formula for the Gini coefficient:

$$\mathrm{G} = \frac{N+1}{N-1} - \frac{2}{N(N-1)\mu}\left(\sum\nolimits_{i=1}^{n} P_i X_i\right)$$

where μ is mean income of the population, Pi is the income rank P of person i, with income X, such that the richest person receives a rank of 1 and the poorest a rank of N. This effectively gives higher weight to poorer people in the income distribution, which allows the Gini to meet the Transfer Principle. Note that the Jasso-Deaton formula rescales the coefficient so that its value is 1 if all the X_i are zero except one. Note however Allison's reply on the need to divide by N^2 instead. [**21**]

Generalized inequality indices

The Gini coefficient and other standard inequality indices reduce to a common form. Perfect equality—the absence of inequality—exists when and only when the inequality ratio, $r_j = x_j/\bar{x}$, equals **1** for all j units in some population (for example, there is perfect income equality when everyone's income x_j equals the mean income $\bar{x}$, so that $r_j = 1$ for everyone). Measures of inequality, then, are measures of the average deviations of the $r_j = 1$ from 1; the greater the average deviation, the greater the inequality. Based on these observations the inequality indices have this common form:

$$\text{Inequality} = \sum_j p_j f(r_j)$$

, where p_j weights the units by their population share, and $f(r_j)$ is a function of the deviation of each unit's r_j from **1**, the point of equality. The insight of this generalized inequality index is that inequality indices differ because they employ different functions of the distance of the inequality ratios (the r_j) from **1.**

Features of Gini coefficient

The Gini coefficient has features that make it useful as a measure of dispersion in a population, and inequalities in particular. [22] It is a ratio analysis method making it easier to interpret. It also avoids references to a statistical average or position unrepresentative of most of the population, such as per capita income or gross domestic product. For a given time interval, Gini coefficient can therefore be used to compare diverse countries and different regions or groups within a country; for example states, counties, urban versus rural areas, gender and ethnic groups. [citation needed] Gini coefficients can be used to compare income distribution over time, thus it is possible to see if inequality is increasing or decreasing independent of absolute incomes. [citation needed]
Other useful features of the Gini coefficient include: [48] [citation needed] [49]

- Anonymity: it does not matter who the high and low earners are.
- Scale independence: the Gini coefficient does not consider the size of the economy, the way it is measured, or whether it is a rich or poor country on average.
- Population independence: it does not matter how large the population of the country is.
- Transfer principle: if income (less than the difference), is transferred from a rich person to a poor person the resulting distribution is more equal.

Limitations of Gini coefficient

Gini Coefficient

The Gini coefficient is a relative measure. Its proper use and interpretation is controversial.[**51**] It is possible for the Gini coefficient of a developing country to rise (due to increasing inequality of income) while the number of people in absolute poverty decreases.[**52**] This is because the Gini coefficient measures relative, not absolute, wealth. Changing income inequality, measured by Gini coefficients, can be due to structural changes in a society such as growing population (baby booms, aging populations, increased divorce rates, extended family households splitting into nuclear families, emigration, immigration) and income mobility.[**53**] Gini coefficients are simple, and this simplicity can lead to oversights and can confuse the comparison of different populations; for example, while both Bangladesh (per capita income of $**1**,**693**) and the Netherlands (per capita income of $**42**,**183**) had an income Gini coefficient of **0.31** in **2010**,[**54**] the quality of life, economic opportunity and absolute income in these countries are very different, i**.e.** countries may have identical Gini coefficients, but differ greatly in wealth. Basic necessities may be available to all in a developed economy, while in an undeveloped economy with the same Gini coefficient, basic necessities may be unavailable to most or unequally available, due to lower absolute wealth.

Different income distributions with the same Gini coefficient

Even when the total income of a population is the same, in certain situations two countries with different income distributions can have the same Gini index (e.g. cases when income Lorenz Curves cross).[22] Table A illustrates one such situation. Both countries have a Gini coefficient of 0.2, but the average income distributions for household groups are different. As another example, in a population where the lowest 50% of individuals have no income and the other 50% have equal income, the Gini coefficient is 0.5; whereas for another population where the lowest 75% of people have 25% of income and the top 25% have 75% of the income, the Gini index is also 0.5. Economies with similar incomes and Gini coefficients can have very different income distributions. Bellù and Liberati claim that to rank income inequality between two different populations based on their Gini indices is sometimes not possible, or misleading.[55]

Table A. Different income distributions
with the same Gini Index [22]

Household Group	Country A Annual Income ($)	Country B Annual Income ($)
1	$20,000	$9,000
2	$30,000	$40,000
3	$40,000	$48,000
4	$50,000	$48,000
5	$60,000	$55,000
Total Income	$200,000	$200,000
Country's Gini	**0.2**	**0.2**

Extreme wealth inequality, yet low income Gini coefficient

Table B. Same Income Distributions (Different Gini Index)			
Household number	**Country A Annual Income ($)**	**Household combined number**	**Country A combined Annual Income ($)**
1	$20,000		
2	$30,000	1 & 2	$50,000
3	$40,000		
4	$50,000	3 & 4	$90,000
5	$60,000		
6	$70,000	5 & 6	$130,000
7	$80,000		
8	$90,000	7 & 8	$170,000
9	$120,000		
10	$150,000	9 & 10	$270,000
Total Income	$710,000		$710,000
Country's Gini	**0.303**		**0.293**

A Gini index does not contain information about absolute national or personal incomes. Populations can have very low income Gini indices, yet simultaneously very high wealth Gini index. By measuring inequality in income, the Gini ignores the differential efficiency of use of household income. By ignoring wealth (except as it contributes to income) the Gini can create the appearance of inequality when the people compared are at different stages in their life. Wealthy countries such as Sweden can show a low Gini coefficient for disposable income of **0.31** thereby appearing equal, yet have very high Gini coefficient for wealth of **0.79** to **0.86** thereby suggesting an extremely unequal wealth distribution in its society. [**56**][**57**] These factors are not assessed in income-based Gini.

Small sample bias – sparsely populated regions more likely to have low Gini coefficient

Gini index has a downward-bias for small populations.[**58**] Counties or states or countries with small populations and less diverse economies will tend to report small Gini coefficients. For economically diverse large population groups, a much higher coefficient is expected than for each of its regions. Taking world economy as one, and income distribution for all human beings, for example, different scholars estimate global Gini index to range between **0.61** and **0.68.** [**10**][**11**] As with other inequality coefficients, the Gini coefficient is influenced by the granularity of the measurements. For example, five **20**% quantiles (low granularity) will usually yield a lower Gini coefficient than twenty **5**% quantiles (high granularity) for the same distribution. Philippe Monfort has shown that using inconsistent or unspecified granularity limits the usefulness of Gini coefficient measurements. [**59**]

Gini Coefficient

The Gini coefficient measure gives different results when applied to individuals instead of households, for the same economy and same income distributions. If household data is used, the measured value of income Gini depends on how the household is defined. When different populations are not measured with consistent definitions, comparison is not meaningful.

Deininger and Squire (**1996**) show that income Gini coefficient based on individual income, rather than household income, are different. For United States, for example, they find that individual income-based Gini coefficient was **0.35**, while for France they report individual income-based Gini index to be **0.43.** According to their individual focused method, in the **108** countries they studied, South Africa had the world's highest Gini coefficient at 0.62, Malaysia had Asia's highest Gini coefficient at **0.5**, Brazil the highest at **0.57** in Latin America and Caribbean region, and Turkey the highest at **0.5** in OECD countries. [**60**]

Gini coefficient is unable to discern the effects of structural changes in populations [53]

Expanding on the importance of life-span measures, the Gini coefficient as a point-estimate of equality at a certain time, ignores life-span changes in income. Typically, increases in the proportion of young or old members of a society will drive apparent changes in equality, simply because people generally have lower incomes and wealth when they are young than when they are old. Because of this, factors such as age distribution within a population and mobility within income classes can create the appearance of inequality when none exist taking into account demographic effects. Thus a given economy may have a higher Gini coefficient at any one point in time compared to another, while the Gini coefficient calculated over individuals' lifetime income is actually lower than the apparently more equal (at a given point in time) economy's. [**14**] Essentially, what matters is not just inequality in any particular year, but the composition of the distribution over time.

Table C. Household Money income		
(Distributions and Gini Index)		
Income bracket	**% of Population**	**% of Population**
(in 2010 adjusted dollars)	**1979**	**2010**
Under $15,000	14.60%	13.70%
$15,000 – $24,999	11.90%	12.00%
$25,000 – $34,999	12.10%	10.90%
$35,000 – $49,999	15.40%	13.90%
$50,000 – $74,999	22.10%	17.70%
$75,000 – $99,999	12.40%	11.40%
$100,000 – $149,999	8.30%	12.10%
$150,000 – $199,999	2.00%	4.50%
$200,000 and over	1.20%	3.90%
Total Households	80,776,000	118,682,000
United States' Gini	**0.404**	**0.469**

on pre-tax basis		

Kwok claims income Gini coefficient for Hong Kong has been high (**0.434** in **2010**[**54**]), in part because of structural changes in its population. Over recent decades, Hong Kong has witnessed increasing numbers of small households, elderly households and elderly living alone. The combined income is now split into more households. Many old people are living separately from their children in Hong Kong. These social changes have caused substantial changes in household income distribution. Income Gini coefficient, claims Kwok, does not discern these structural changes in its society.[**53**] Household money income distribution for the United States, summarized in Table C of this section, confirms that this issue is not limited to just Hong Kong. According to the US Census Bureau, between **1979** and **2010**, the population of United States experienced structural changes in overall households, the income for all income brackets increased in inflation-adjusted terms, household income distributions shifted into higher income brackets *over time, while the income Gini coefficient increased.*

Another limitation of Gini coefficient is that it is not a proper measure of egalitarianism, as it is only measures income dispersion. For example, if two equally egalitarian countries pursue different immigration policies, the country accepting a higher proportion of low-income or impoverished migrants will report a higher Gini coefficient and therefore may appear to exhibit more income inequality.

Inability to value benefits and income from informal economy affects Gini coefficient accuracy

Some countries distribute benefits that are difficult to value. Countries that provide subsidized housing, medical care, education or other such services are difficult to value objectively, as it depends on quality and extent of the benefit. In absence of free markets, valuing these income transfers as household income is subjective. The theoretical model of Gini coefficient is limited to accepting correct or incorrect subjective assumptions.

In subsistence-driven and informal economies, people may have significant income in other forms than money, for example through subsistence farming or bartering. These income tend to accrue to the segment of population that is below-poverty line or very poor, in emerging and transitional economy countries such as those in sub-Saharan Africa, Latin America, Asia and Eastern Europe. Informal economy accounts for over half of global employment and as much as **90** per cent of employment in some of the poorer sub-Saharan countries with high official Gini inequality coefficients. Schneider et al., in their **2010** study of **162** countries, [**63**] report about **31.2**%, or about $**20** trillion, of world's GDP is informal. In developing countries, the informal economy predominates for all income brackets except for the richer, urban upper income bracket populations. Even in developed economies, between 8% (United States) to 27% (Italy) of each nation's GDP is informal, and resulting informal income predominates as a livelihood activity for those in the lowest income brackets.[**64**] The value and distribution of the incomes from informal or underground economy is difficult to quantify, making true income Gini coefficients estimates difficult.[**65**][**66**] Different assumptions and quantifications of these incomes will yield different Gini coefficients.[**67**][**68**][**69**]

Gini has some mathematical limitations as well. It is not additive and different sets of people cannot be averaged to obtain the Gini coefficient of all the people in the sets.

Alternatives to Gini coefficient

Given the limitations of Gini coefficient, other statistical methods are used in combination or as an alternative measure of population dispersity. For example, ***entropy measures*** are frequently used (e.g. the Theil Index, the Atkinson index and the generalized entropy index). These measures attempt to compare the distribution of resources by intelligent agents in the market with a maximum entropy random distribution, which would occur if these agents acted like non-intelligent particles in a closed system following the laws of statistical physics.

Relation to other statistical measures

The Gini coefficient closely related to the AUC (Area Under receiver operating characteristic Curve) measure of performance. [**70**] The relation follows the formula $\mathrm{AUC} = (\mathrm{Gini} + 1)/2$. Gini coefficient is also closely related to Mann–Whitney U (c statistics).

The Gini index is also related to Pietra index—both of which are a measure of statistical heterogeneity and are derived from Lorenz curve and the diagonal line.[**71**][**72**]

In certain fields such as ecology, inverse Simpson's index $1/\lambda$ is used to quantify diversity, and this should not be confused with the Simpson index λ . These indicators are related to Gini. The inverse Simpson index increases with diversity, unlike Simpson index and Gini coefficient which decrease with diversity. The Simpson index is in the range [0, 1], where 0 means maximum and 1 means minimum diversity (or heterogeneity). Since diversity indices typically increase with increasing heterogeneity, Simpson index is often transformed into inverse Simpson, or using the complement $1 - \lambda$, known as Gini-Simpson Index. [73]

Other uses

Although the Gini coefficient is most popular in economics, it can in theory be applied in any field of science that studies a distribution. For example, in ecology the Gini coefficient has been used as a measure of biodiversity, where the cumulative proportion of species is plotted against cumulative proportion of individuals.[**74**] In health, it has been used as a measure of the inequality of health related quality of life in a population.[**75**] In education, it has been used as a measure of the inequality of universities.[**76**] In chemistry it has been used to express the selectivity of protein kinase inhibitors against a panel of kinases.[**77**] In engineering, it has been used to evaluate the fairness achieved by Internet routers in scheduling packet transmissions from different flows of traffic.[**78**]
The Gini coefficient is sometimes used for the measurement of the discriminatory power of rating systems in credit risk management.[**79**]
A **2005** study accessed US census data to measure home computer ownership and used the Gini coefficient to measure inequalities amongst whites and African Americans. Results indicated that although decreasing overall, home computer ownership inequality is substantially smaller among white households. [**80**]

A **2016** peer-reviewed study titled Employing the Gini coefficient to measure participation inequality in treatment-focused Digital Health Social Networks [**81**] illustrated that the Gini coefficient was helpful and accurate in measuring shifts in inequality, however as a standalone metric it failed to incorporate overall network size.
The discriminatory power refers to a credit risk model's ability to differentiate between defaulting and non-defaulting clients. The formula G_1, in calculation section above, may be used for the final model and also at individual model factor level, to quantify the discriminatory power of individual factors. It is related to accuracy ratio in population assessment models.

References (Omitted and see Wikipedia)

Further reading

- *Amiel, Y.; Cowell, F. A. (1999). Thinking about Inequality. Cambridge. ISBN 0-521-46696-2.*
- *Anand, Sudhir (1983). Inequality and Poverty in Malaysia. New York: Oxford University Press. ISBN 0-19-520153-1.*
- *Brown, Malcolm (1994). "Using Gini-Style Indices to Evaluate the Spatial Patterns of Health Practitioners: Theoretical Considerations and an Application Based on Alberta Data". Social Science Medicine. **38** (9): 1243–1256. doi:10.1016/0277-9536(94)90189-9. PMID 8016689.*
- *Chakravarty, S. R. (1990). Ethical Social Index Numbers. New York: Springer-Verlag. ISBN 0-387-52274-3.*
- *Deaton, Angus (1997). Analysis of Household Surveys. Baltimore MD: Johns Hopkins University Press. ISBN 0-585-23787-5.*
- *Dixon, Philip M.; Weiner, Jacob; Mitchell-Olds, Thomas; Woodley, Robert (1987). "Bootstrapping the Gini coefficient of inequality". Ecology. Ecological Society of America. **68** (5): 1548–1551. doi:10.2307/1939238. JSTOR 1939238.*
- *Dorfman, Robert (1979). "A Formula for the Gini Coefficient". The Review of Economics and Statistics. The MIT Press. **61** (1): 146–149. doi:10.2307/1924845. JSTOR 1924845.*
- *Firebaugh, Glenn (2003). The New Geography of Global Income Inequality. Cambridge MA: Harvard University Press. ISBN 0-674-01067-1.*
- *Gastwirth, Joseph L. (1972). "The Estimation of the Lorenz Curve and Gini Index". The Review of Economics and Statistics. The MIT Press. **54** (3): 306–316. doi:10.2307/1937992. JSTOR 1937992.*
- *Giles, David (2004). "Calculating a Standard Error for the Gini Coefficient: Some Further Results". Oxford Bulletin of Economics and Statistics. **66** (3): 425–433. doi:10.1111/j.1468-0084.2004.00086.x.*
- *Gini, Corrado (1912). Variabilità e mutabilità.* Reprinted in *Pizetti, E.; Salvemini, T., eds. (1955). Memorie di metodologica statistica. Rome: Libreria Eredi Virgilio Veschi.*
- *Gini, Corrado (1921). "Measurement of Inequality of Incomes". The Economic Journal. Blackwell Publishing. **31** (121): 124–126. doi:10.2307/2223319. JSTOR 2223319.*
- *Giorgi, Giovanni Maria (1990). "Bibliographic portrait of the Gini concentration ratio" (PDF). Metron. **48**: 183–231.*

- *Karagiannis, E.; Kovacevic, M. (2000). "A Method to Calculate the Jackknife Variance Estimator for the Gini Coefficient". Oxford Bulletin of Economics and Statistics.* ***62****: 119–122. doi:10.1111/1468-0084.00163.*
- *Mills, Jeffrey A.; Zandvakili, Sourushe (1997). "Statistical Inference via Bootstrapping for Measures of Inequality". Journal of Applied Econometrics.* **12** *(2): 133–150. doi:10.1002/(SICI)1099-1255(199703)12:2<133::AID-JAE433>3.0.CO;2-H. JSTOR 2284908.*
- *Modarres, Reza; Gastwirth, Joseph L. (2006). "A Cautionary Note on Estimating the Standard Error of the Gini Index of Inequality". Oxford Bulletin of Economics and Statistics.* ***68*** *(3): 385–390. doi:10.1111/j.1468-0084.2006.00167.x.*
- *Morgan, James (1962). "The Anatomy of Income Distribution". The Review of Economics and Statistics. The MIT Press.* ***44*** *(3): 270–283. doi:10.2307/1926398. JSTOR 1926398.*
- *Ogwang, Tomson (2000). "A Convenient Method of Computing the Gini Index and its Standard Error". Oxford Bulletin of Economics and Statistics.* ***62****: 123–129. doi:10.1111/1468-0084.00164.*
- *Ogwang, Tomson (2004). "Calculating a Standard Error for the Gini Coefficient: Some Further Results: Reply". Oxford Bulletin of Economics and Statistics.* ***66*** *(3): 435–437. doi:10.1111/j.1468-0084.2004.00087.x.*
- *Xu, Kuan (January 2004). "How Has the Literature on Gini's Index Evolved in the Past 80 Years?" (PDF). Department of Economics, Dalhousie University. Retrieved 2006-06-01.* The Chinese version of this paper appears in *Xu, Kuan (2003). "How Has the Literature on Gini's Index Evolved in the Past 80 Years?". China Economic Quarterly.* **2***: 757–778.*
- *Yitzhaki, Shlomo (1991). "Calculating Jackknife Variance Estimators for Parameters of the Gini Method". Journal of Business and Economic Statistics. American Statistical Association.* ***9*** *(2): 235–239. doi:10.2307/1391792. JSTOR 1391792.*

Gini Coefficient