# Quantifying Income Inequality by the Gini Index

**The Lorenz Curve for a Country**

Rank households in the country by yearly income, in increasing order.
Let $x$ = income percentile, expressed as a decimal. For example $x = 0.75$ would correspond to the household income that is larger than the household income of 75% of the families in the country
Observe $0 \leq x \leq 1$.

Now add up the household incomes for all the households at the $x^{th}$ percentile or lower.
Also add up the household incomes for all the households in the country.
Divide to get a ratio. The value of the ratio depends on $x$, so it is a function that we could call $f(x)$.

$$f(x) = \frac{sum\ of\ household\ incomes\ for\ all\ households\ at\ the\ x^{th}\ percentile\ or\ lower}{sum\ of\ household\ incomes\ for\ all\ households}$$

Observe $0 \leq f(x) \leq 1$

Make a graph of $f(x) - vs - x$. The resulting graph is called the *Lorenz Curve* for the country.

We will now consider examples of three small island countries. Each country has only five households. Their incomes are shown in the table below. We will make Lorenz Curves for the three countries.

| **Island Country 1** | **Island Country 2** | **Island Country 3** |
|---|---|---|
| Household a: \$10k/year | Household a: \$20k/year | Household a: \$58k/year |
| Household b: \$20k/year | Household b: \$40k/year | Household b: \$59k/year |
| Household c: \$30k/year | Household c: \$60k/year | Household c: \$60k/year |
| Household d: \$40k/year | Household d: \$80k/year | Household d: \$61k/year |
| Household e: \$200/year | Household e: \$100k/year | Household e: \$62/year |

We can compute the data for the Lorenz Curve for each country. Call their curves *f(x)*, *g(x)*, and *h(x)*.

| **Data for Lorenz Curve *f(x)* for Island Country 1** | | | |
|---|---|---|---|
| Income Percentile $x$ | Households at $x^{th}$ Percentile or Lower | Sum of Incomes of Those Households | $f(x)$ |
| .2 | a | \$10k | 10/300 = 0.03 |
| .4 | a,b | \$10k + \$20k = \$30k | 30/300 = 0.10 |
| .6 | a,b,c | \$10k + \$20k + \$30k = \$60k | 60/300 = 0.2 |
| .8 | a,b,c,d | \$10k + \$20k + \$30k + \$40k = \$100k | 100/300 = 0.33 |
| 1 | a,b,c,d,e | \$10k + \$20k + \$30k + \$40k + \$200k = \$300k | 300/300 = 1.00 |

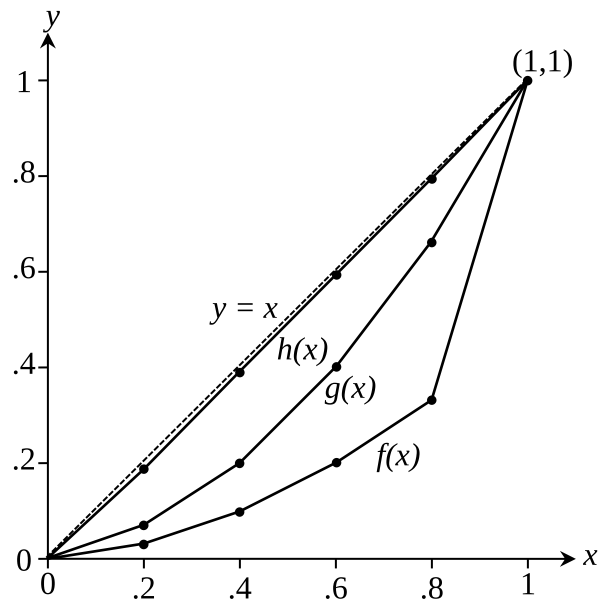
.

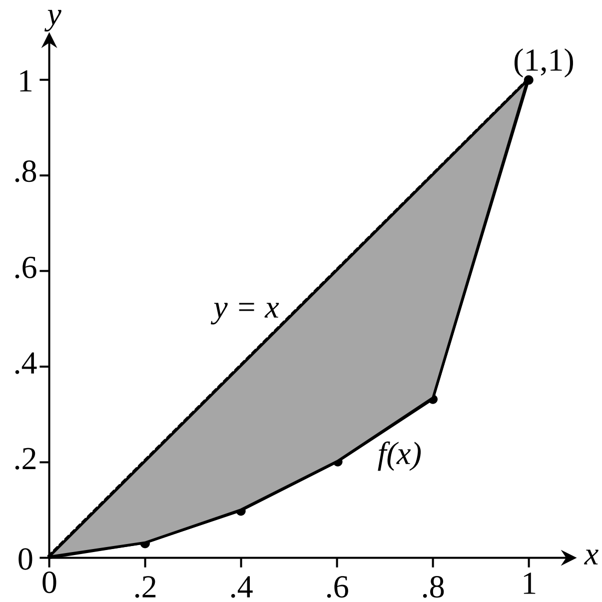
| **Data for Lorenz Curve *g(x)* for Island Country 2** | | | |
|---|---|---|---|
| Income Percentile $x$ | Households at $x^{th}$ Percentile or Lower | Sum of Incomes of Those Households | $g(x)$ |
| .2 | a | \$20k | 20/300 = 0.06 |
| .4 | a,b | \$20k + \$40k = \$60k | 60/300 = 0.20 |
| .6 | a,b,c | \$20k + \$40k + \$60k = \$120k | 120/300 = 0.4 |
| .8 | a,b,c,d | \$20k + \$40k + \$60k + \$80k = \$200k | 200/300 = 0.67 |
| 1 | a,b,c,d,e | \$20k + \$40k + \$60k + \$80k + \$100k = \$300k | 300/300 = 1.00 |

.

| Data for Lorenz Curve *h(x)* for Island Country 3 | | | |
|---|---|---|---|
| Income Percentile $x$ | Households at $x^{th}$ Percentile or Lower | Sum of Incomes of Those Households | $h(x)$ |
| .2 | a | \$58k | 58/300 = 0.19 |
| .4 | a,b | \$58k + \$59k = \$117k | 117/300 = 0.39 |
| .6 | a,b,c | \$58k + \$59k + \$60k = \$177k | 177/300 = 0.59 |
| .8 | a,b,c,d | \$58k + \$59k + \$60k + \$61k = \$238k | 238/300 = 0.79 |
| 1 | a,b,c,d,e | \$58k + \$59k + \$60k + \$61k + \$62k = \$300k | 300/300 = 1.00 |

Now make plots of $f(x)$ and $g(x)$ and $h(x)$ to get the Lorenz Curves for the countries.




Notice two things about the curves:

- All three of the curves go through the points (0,0) and (1,1)
- When household income is very unequally distributed (as in Country 1), the Lorenz curve (curve *f(x)*) is very bowed; When household income is nearly equally distributed (as in Country 3), the Lorenz curve (curve *h(x)*) is not very bowed and stays very close to the line *y* = *x*, which is shown as a dotted line.

The line $y = x$ is called the *Line of Absolute Equality*.

**The Gini Index**

It is clear that countries with a very unequal distribution of household income will have Lorenz curves *f(x)* that deviate a lot from the Line of Absolute Equality, *y* = *x*. We would like to quantify the deviation, so that we can compare income inequality among nations. One straightforward way to quantify the income inequality is to simply measure the area between the Lorenz Curve *f(x)* and the Line of Absolute Equality *y* = *x*. (See the figure at right, above). That is the idea behind the *Gini Index*:

The *Gini Index (GI)* for a country is defined to be twice the area of the region bounded by the Lorenz Curve $f(x)$ for the country and the Line of Absolute Equality $y = x$. (The 2 is a scale factor put in so that the Gini Index *(GI)* will be a number between 0 and 1.)

$$Gini\ Index = GI = 2 \cdot \int_{x=0}^{x=1} x - f(x)\, dx$$