Subject for this video:

Quantifying Income Concentration with the Gini Index

Reading:

- **General:** Section 6.1 The Area Between Curves
- **More Specifically:** Pages 392 - 394 Example 8

Homework: H85: The Gini Index (6.1#83,85)

Recall the theorem introduced in the video for Homework H82

Theorem about the Area between Curves for a Simple Region.

If $top(x)$ and $bottom(x)$ are continuous functions

and $bottom(x) \leq top(x)$ on the interval $[a, b]$,

then the region bounded by the four curves

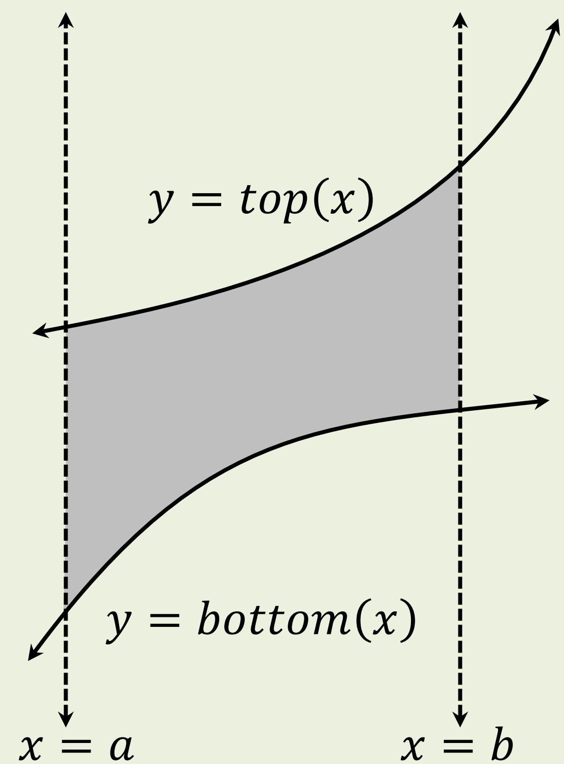
$$y = top(x)$$

$$y = bottom(x)$$

$$x = a$$

$$x = b$$

is what we in this course will call a ***simple region***.



The ***Area Between Curves*** (*unsigned* area) for this region is given by the definite integral

$$USA = \int_a^b top(x) - bottom(x)dx$$

Quantifying Income Inequality by the Gini Index

The degree of income inequality in the United States is a common subject in our society. One frequently hears mentioned the so-called *one percent*. It is generally known that this refers to the relatively small group of people whose income is greater than 99% of the people in country. The discussion about the one percent tends to be about the idea that most of the income (and wealth) in our country is concentrated in that small group. But beyond that vague idea, discussions of income concentration in the news are mostly not very quantitative.

There is a quantitative measure of the income concentration that is based on measuring the area between curves. The quantitative measure is called the *Gini Index*; it is the subject for this video.

The Gini Index is computed using a definite integral involving a function called a *Lorenz Curve*. In order to understand the Gini Index, we must first learn about that curve.

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 greater than or equal to 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.

[Example 1]

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

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: $\$200k/year$	Household e: $\$100k/year$	Household e: $\$62k/year$

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

Data for Lorenz Curve $f(x)$ for Island Country 1

Income Percentile x	Households at Percentile x or Lower	Sum of Incomes of Those Households	$f(x)$ $f(0) = 0$
0.2	a	$\$10k$	$\frac{10}{300} = 0.03$
0.4	a, b	$\$10k + \$20k = \$30k$	$\frac{30}{300} = 0.10$
0.6	a, b, c	$\$10k + \$20k + \$30k = \$60k$	$\frac{60}{300} = 0.20$
0.8	a, b, c, d	$\$10k + \$20k + \$30k + \$40k = \$100k$	$\frac{100}{300} = 0.33$
1.0	a, b, c, d, e	$\$10k + \$20k + \$30k + \$40k + \$200k = \$300k$	$\frac{300}{300} = 1.00$

.

Data for Lorenz Curve $g(x)$ for Island Country 2

Income Percentile x	Households at Percentile x or Lower	Sum of Incomes of Those Households	$g(x)$
0.2	a	$\$20k$	$\frac{20}{300} = 0.06$
0.4	a, b	$\$20k + \$40k = \$60k$	$\frac{60}{300} = 0.20$
0.6	a, b, c	$\$20k + \$40k + \$60k = \$120k$	$\frac{120}{300} = 0.40$
0.8	a, b, c, d	$\$20k + \$40k + \$60k + \$80k = \$200k$	$\frac{200}{300} = 0.67$
1.0	a, b, c, d, e	$\$20k + \$40k + \$60k + \$80k + \$100k = \$300k$	$\frac{300}{300} = 1.00$

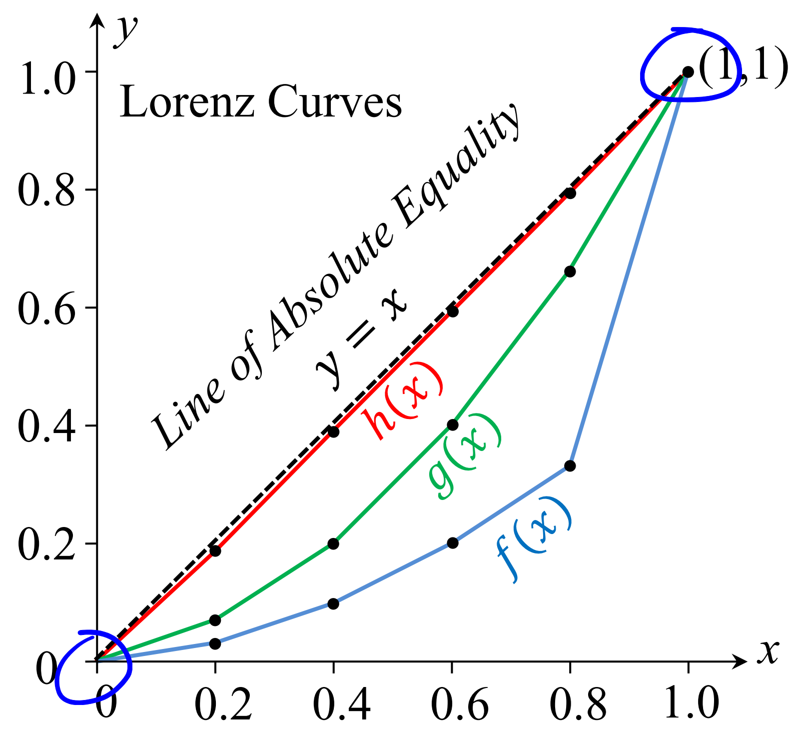
.

Data for Lorenz Curve $h(x)$ for Island Country 3

Income Percentile x	Households at Percentile x or Lower	Sum of Incomes of Those Households	$h(x)$
0.2	a	$\$58k$	$\frac{58}{300} = 0.19$
0.4	a, b	$\$58k + \$59k = \$117k$	$\frac{117}{300} = 0.39$
0.6	a, b, c	$\$58k + \$59k + \$60k = \$177k$	$\frac{177}{300} = 0.59$
0.8	a, b, c, d	$\$58k + \$59k + \$60k + \$61k = \$238k$	$\frac{238}{300} = 0.79$
1.0	a, b, c, d, e	$\$58k + \$59k + \$60k + \$61k + \$62k = \$300k$	$\frac{300}{300} = 1.00$

.

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:

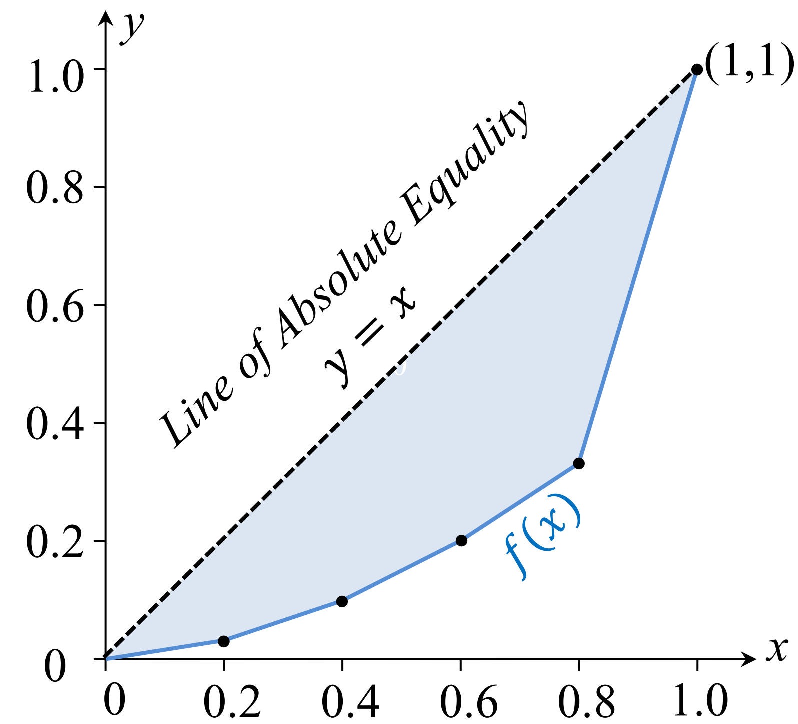
- All three of the curves go through the points (0,0) and (1,1)
- When household income is very concentrated, very unequally distributed, the Lorenz curve is very bowed; When household income is less concentrated, more equally distributed, the Lorenz curve is less bowed and stays close to the line $y = x$, which is shown as a dotted line.

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

End of [Example 1]

The Gini Index

It is clear that countries with a high concentration of household income among the top income earners 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 concentration among nations. One straightforward way to quantify the income concentration is to simply measure the area between the Lorenz Curve $f(x)$ and the Line of Absolute Equality $y = x$.



Notice that the blue region is a *simple region*, whose *top* curve is the *Line of Absolute Equality* and whose *bottom* curve is the *Lorenz Curve*. That is, $top(x) = x$ and $bottom(x) = f(x)$.

So the area of the blue region will be

$$USA = \int_0^1 top(x) - bottom(x)\, dx = \int_0^1 x - f(x)\, dx$$

That is the idea behind the *Gini Index*

Definition of the Gini Index of Income Concentration

The ***Gini Index (GI)*** for a country is defined to be twice the area of the region bounded by the *Lorenz Curve* for the country and the *Line of Absolute Equality*.

$$Gini\ Index = GI = 2 \cdot \int_0^1 x - f(x)\, dx$$

Remark: The 2 is a scale factor put in so that the Gini Index *(GI)* will be a number $0 \leq GI \leq 1$.

[Example 1] **(a)** Find the *Gini Index* if the *Lorenz Curve* is $f(x) = x^2$.

Solution:

$$GI = 2 \cdot \int_0^1 x - f(x)\, dx$$

$$= 2 \cdot \int_0^1 x - x^2\, dx$$

$n=1$ $n=2$

$$\underset{FTC}{=} 2\left(\int x - x^2 dx\right)\Big|_0^1$$

power rule for integrals, with $n=1$ and $n=2$

$$= 2\left(\frac{x^2}{2} - \frac{x^3}{3} + C\right)\Big|_0^1$$

$$= 2 \cdot \left(\left(\frac{(1)^2}{2} - \frac{(1)^3}{3} + \cancel{C}\right) - \left(\frac{(0)^2}{2} - \frac{(0)^3}{3} + \cancel{C}\right)\right)$$

$$= 2 \cdot \left(\frac{1}{2} - \frac{1}{3}\right)$$

$$= 2 \cdot \left(\frac{1}{6}\right)$$

$$= \frac{1}{3}$$

(b) Find the *Gini Index* if the *Lorenz Curve* is $g(x) = x^3$.

Solution:

$$GI = 2 \cdot \int_0^1 x - g(x)\,dx$$

$$= 2 \cdot \int_0^1 x - x^3\,dx$$

n = 1 n = 3

$$\underset{FTC}{=} 2\left(\int x - x^3 dx\right)\Big|_0^1$$

$$= 2\left(\frac{x^2}{2} - \frac{x^4}{4} + C\right)\Big|_0^1$$

$$= 2 \cdot \left(\left(\frac{(1)^2}{2} - \frac{(1)^4}{4} + \cancel{C}\right) - \left(\frac{(0)^2}{2} - \frac{(0)^4}{4} + \cancel{C}\right)\right)$$

$$= 2 \cdot \left(\frac{1}{2} - \frac{1}{4}\right)$$

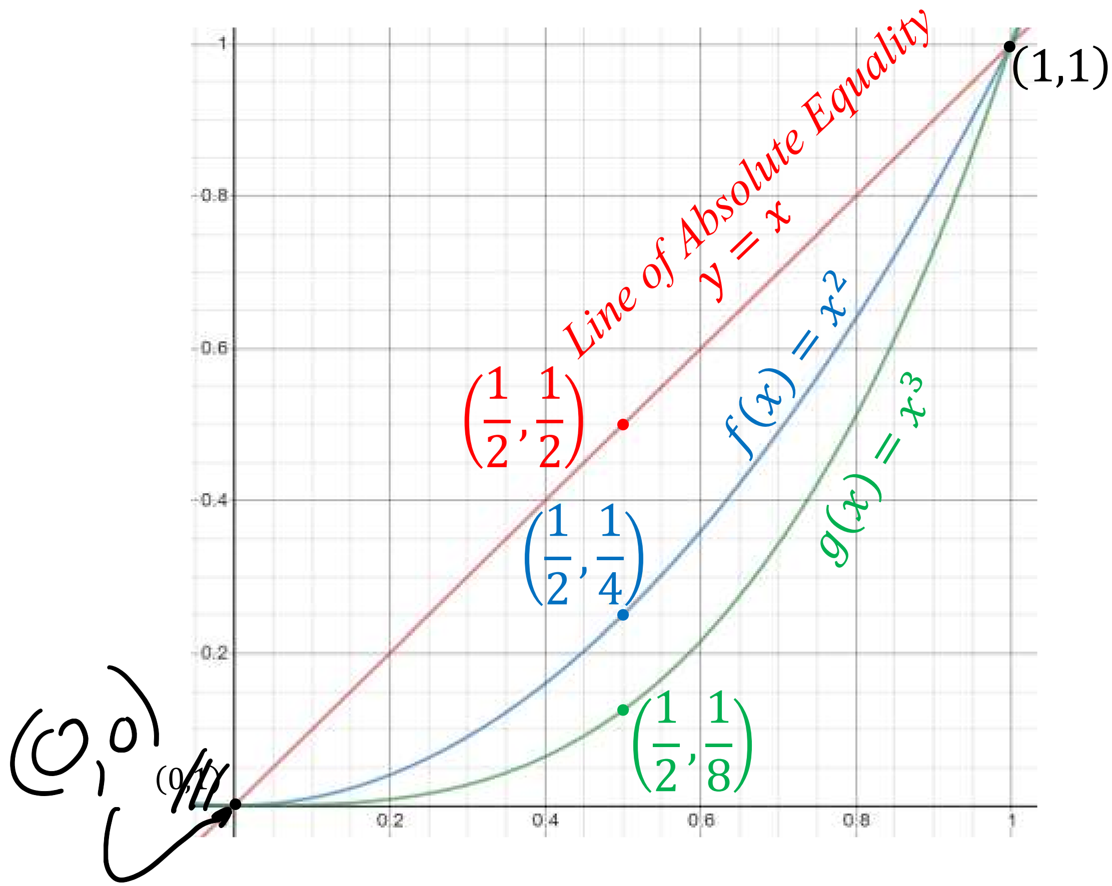
$$= 2 \cdot \left(\frac{1}{4}\right)$$

$$= \frac{1}{2}$$

(c) Graph the two Lorentz curves along with the Line of Absolute Equality.

Solution:

Here is a graph from Desmos, with annotations.



Observations:

- All the graphs go through $(0,0)$ and $(1,1)$
- The graph of $g(x)$ bows away from the line of absolute equality more than $f(x)$ bows.
- Function $g(x)$ has a larger Gini Index: $\frac{1}{2} > \frac{1}{3}$.

End of [Example 1]