

The Gini Index

LA Math Circle
High School I

Based on notes by Robert Brown, and a handout by Dillon Zhi

February 7, 2016

Lorenz curve and Gini index

Definition : The Lorenz Curve is a graph used to depict the distribution of wealth within a society. Both axes go from 0 to 100%; the x -axis is percentage of the population, the y -axis is percentage of the total wealth, and the curve consists of points (x, y) where the bottom $x\%$ of the population has $y\%$ of the wealth. Typically the Lorenz curve is also drawn the the “equal distribution line” $y = x$.

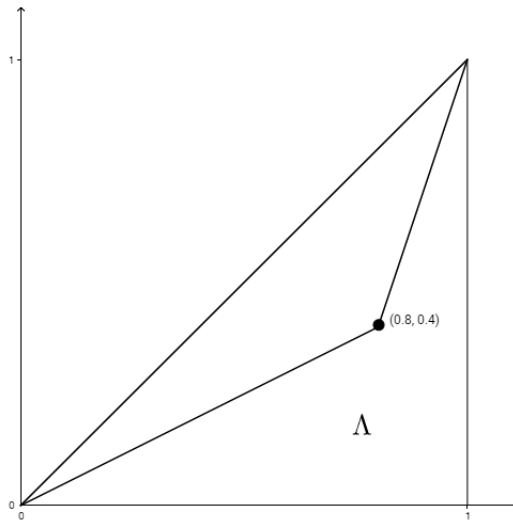
1. Why do the points $(0, 0)$ and $(1, 1)$ lie on the Lorenz curve?
2. Explain why the line segment from $(0, 0)$ to $(1, 1)$ should be called the “equal distribution line”.
3. Why does the Lorenz curve lie below the equal distribution line? For instance, why can't $(.25, .75)$ lie on the Lorenz curve?

Definition : The Gini coefficient is a way of attaching a single number to the inequality that can be observed from the Lorenz curve: if the area between the equal distribution line and the Lorenz curve is Γ , and the area between the Lorenz curve and the axes is Λ , then the Gini coefficient is $\frac{\Gamma}{\Gamma+\Lambda}$

4. What distribution of income would make the area of Λ equal $\frac{1}{2}$? What distribution would make the area of Γ equal $\frac{1}{2}$?

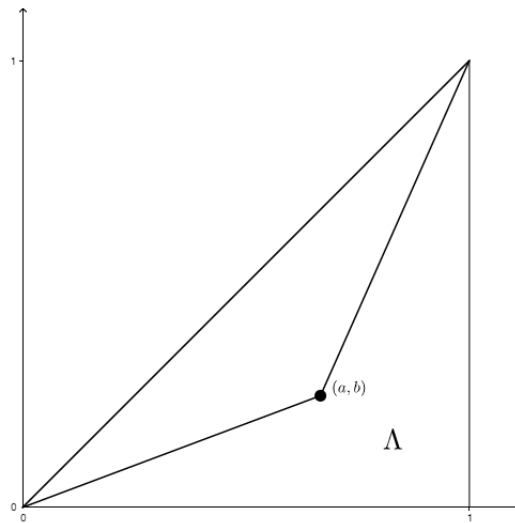
5. Given the Lorenz curve consisting of two line segments in the figure, calculate the area of Λ and use that information to compute the Gini index.

Figure for Problem 5



6. Calculate the area of Λ in the figure and then write a general formula for this “one-point estimate” of the Gini index in terms of a and b .

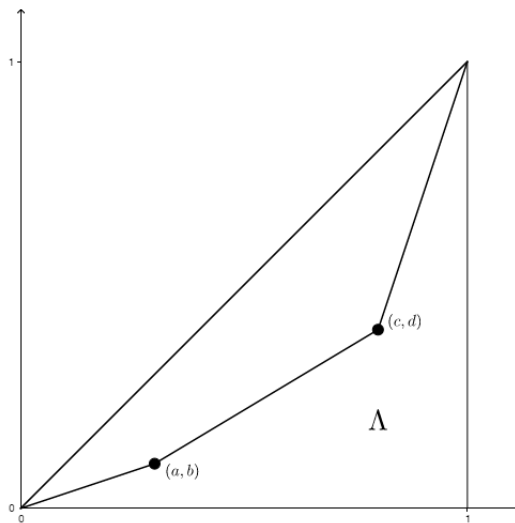
Figure for Problem 6



7. If we base our calculation of the Gini index on a one-point estimate, will that estimate be larger or smaller than the actual value of the Gini index? Why?

8. Calculate the area of Λ in the figure and then write a general formula for this “two-point estimate” of the Gini index in terms of a , b , c , and d .

Figure for Problem 8



9. Use the answer to Problem 8 to make a two-point estimate of the Gini index if $(a, b) = (.8, .4)$ and $(c, d) = (.99, .8)$.

10. Consider a country with n people. Let the incomes of all the people in the country be y_1, \dots, y_n , sorted in ascending order, and let $Y = y_1 + \dots + y_n$.

a. Why is the Gini index

$$G = \frac{2}{n} \sum_{k=1}^n \left(\frac{k}{n} - \frac{y_1 + \dots + y_k}{Y} \right)?$$

b. Show that

$$G = \frac{1}{n} \sum_k (2k - (n + 1)) \frac{y_k}{Y}.$$

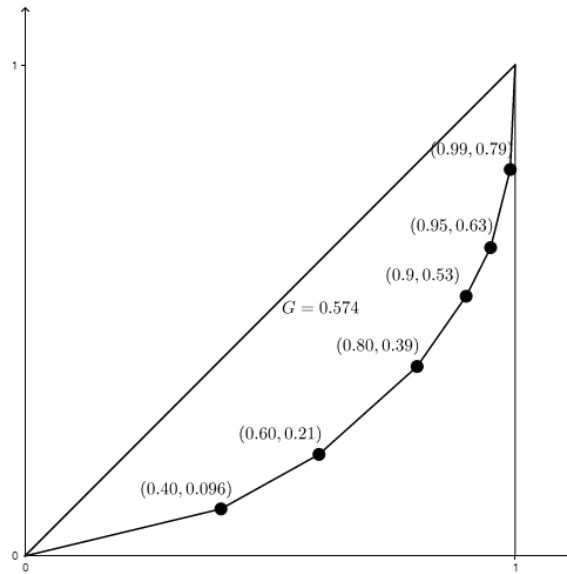
c. Show that

$$G = \frac{\sum_{i,j=1}^n |y_i - y_j|}{2nY},$$

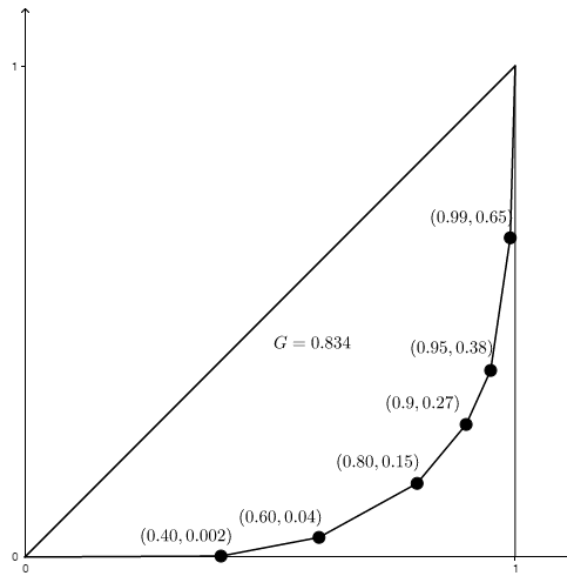
even if y_1, \dots, y_n are not necessarily sorted in ascending order.

d. The city of Villestown has four people, Alice, Bob, Carl, and Danielle. Their annual incomes are \$20000, \$40000, \$60000, are \$80000, respectively. Find the Gini index for Villestown.

Lorenz curve for income distribution in the US, 2006¹



Lorenz curve for wealth distribution in the US, 2007²



¹Wolff, Edward N. (2010). Recent Trends in Household Wealth in the United States: Rising Debt and the Middle-Class Squeeze—an Update to 2007. Levy Economics Institute. http://www.levyinstitute.org/pubs/wp_589.pdf

²ibid.