Gini Index by racial groups in the United States

Rebecca Sherwood
Mentor: Dr. Susan D'Agostino

Southern New Hampshire University

Abstract

Using regression analysis and calculus, economists measure the income concentration of a population by way of a metric known as a Gini index. A Gini index is a number ranging from 0 to 1, where 0 indicates perfect income equality in the measured population and 1 indicates perfect inequality in the measured population. In order to compute a Gini index, we draw a graph that plots the bottom x percent of income earners in a population against the percent of the income (from all available income) that the group has earned. Unlike other Gini index studies, we compute a Gini index in which we omit the top 5% of income earners in each population. As a result, our study considers income distribution in the bulk of a population without the skewed effect of high-income outliers.

In addition, our study further differs from other Gini index studies because we compute Gini indexes for subsets of the United States populations by race including: Whites, non-Hispanic Whites, Blacks, Asians, and Hispanics. All of our data hails from the United States Census Bureau. Some questions that this study answers are: how successful are individual racial groups in the United States in achieving income equality? Are some racial groups faring better than others in achieving income equality? By understanding how race affects income distribution, we may inform public policy experts who are in a position to affect positive change in regard to income equality.

Procedure

- Obtain data from the United States Census Bureau
  - Racial groups (White, White non-Hispanic, Hispanic, Black, Asian, and all Races) and the specific income earned

- Build regression models on Minitab with obtained data
  - Conduct a nested F-test to determine whether a quadratic or a cubic equation is the best fit model.
  - Hypothesis: for a quadratic equation
    $H_0 = \beta_2 = 0$
    $H_a = \beta_2 \neq 0$
  - Compare the F value given fro each regression model with $F_{.05}$. If $F > F_{.05}$ then we reject H_0. Then we can see if the quadratic equation is the best fit model.

- Use the best fit model to calculate the Gini Index using calculus.
  -For example, here is the best fit model for All Races :
  $y = -.2355 + 2.110x - 4.245x^2 + 3.565x^3$
  - we put this equation in to our equation for Gini Index
  $2\int^1_0 [\, x - (-.2355 + 2.110x - 4.245x^2 + 3.565x^3)]dx$

- We now have one Gini Index. We must repeat the steps for all of the categories of races. Then we can compare the results.

Results

Race	Gini Index
Hispanic	.3906
White, Not Hispanic	.3979
White	.399
Asian	.4022
All Races	.4085
Black	.4254

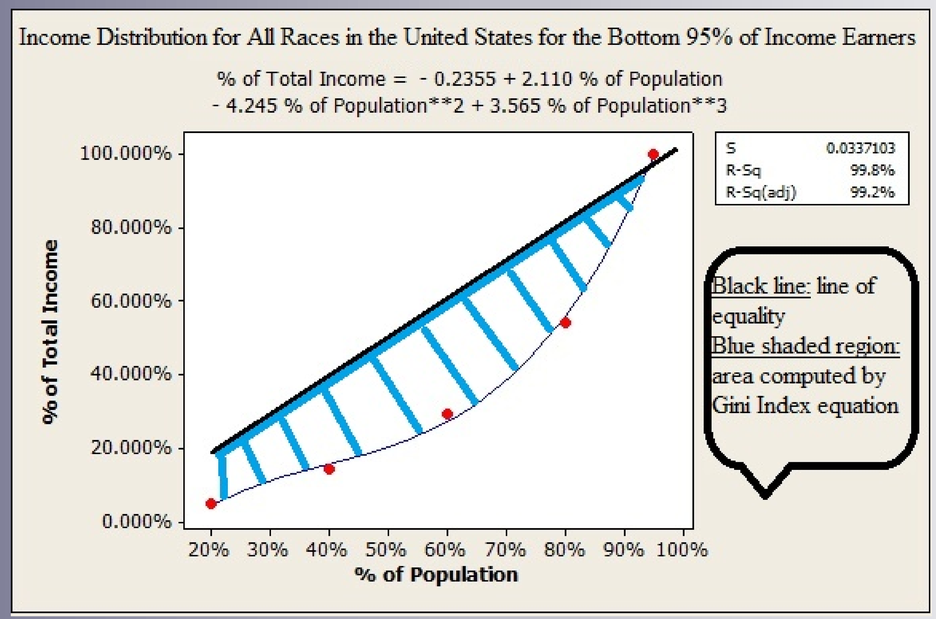
This information is telling us that the Hispanic population in the United States as of 2013 has the most equally distributed income with a Gini Index of .3906. The Black population in the United States as of 2013 has the most income distribution inequality with a Gini Index of .4254.

Big Idea

- Use regression analysis to create and test equations for each racial group.
- Compute Gini Indexes using calculus and the equation for Gini Indexes
  - $2\int^1_0 [\, x - f(x)]dx$
- Analyze what the Gini Indexes computed actually mean
  -We are looking at a number from 0 to 1, where
    -0 is complete income equality
    -1 is complete income inequality

-Equality is a line y=x, the Gini Index computes the area between this line of equality and the equation found.
  -Therefore, the less area under the line of equality the closer the Gini Index is to 0

-Omit the top 5% of income earners so 95% of the population will earn 100% of the income



Further Questions

-How does income range affect income quality?

-What is more important the average level of income earned or the distribution equality?

-How would having more data effect my results? For example if I could break down the income ranges in to smaller brackets what would that do for the computed Gini Indexes?

References

Barnett, R. A., Ziegler, M. R., & Byleen, K. E. (2011). Additional Integration Topics. In *Calculus for Bussiness, Economics, Life Sciences, and Social Sciences* (12th ed., p. 416). Boston, MA: Prentice Hall.

Mendenhall, W., & Sincich, T. (2012). *A second course in statistics: Regression analysis* (7th ed.). Boston, MA: Prentice Hall.

United States Census Bureau. (n.d.). Retrieved from https://www.census.gov/hhes/www/income/data/historical/household/

* Work in Progress