Introduction to R Programming
Slide Set 6: Application 1 - Income Inequality

Maria Ptashkina

Barcelona GSE ITFD

September 2021

Learning Objectives

- Content
  - Draw Lorenz curves
  - Calculate and interpret the Gini coefficient
- R-specific learning objectives
  - Load data in R
  - Data wrangling
  - Plotting

R Projects

- Create an R project by going to `File → Create New Project`
- You can already create a folder for it, and then select an existing directory
- Give a name to your project, and you will see a file with `.Rproj` extension appearing in your working directory
- Now when your `RStudio` session is running through the project file (`.Rproj`), the current working directory points to the root folder where that `.Rproj` file is saved, making your working directory a relative path ▸ R Projects

Housekeeping and Loading Data

- To start working in R script, it's a good practice to clean your workspace

- If you realize you need more packages - install them once, and call the libraries in the beginning of the script

- Go to the Globalinc website ▸Data and download the Excel file containing the data by clicking 'xlsx'

- Save it in your working directory

- Import the data into R

- Inspect the structure of the data: each row shows data for a different country-year combination

Lorenz Curves

- Lorenz curve is a way to visualize the income distribution of a country
- It shows the entire population lined up along the horizontal axis from the poorest to the richest
- The height of the curve at any point on the vertical axis indicates the fraction of total income received by the fraction of the population given by that point on the horizontal axis
- To draw Lorenz curves, we need to calculate the cumulative share of total income owned by each decile (these will be the vertical axis values)
- For example, if `Decile 1` has $1/10$ of total income and `Decile 2` has $2/10$ of total income, the cumulative income share of `Decile 2` is $3/10$

Task 1: Calculate Cumulative Shares

- Choose two countries (US and China)
- Choose years 1980 and 2014, as the basis for your Lorenz curves
- Calculate the total income for each country-year combination using the mean income and the population size
- Assume that there is only one person in each decile, in other words the total income is 10 times the mean income (this simplification works because, by definition, each decile has exactly the same number of people)
- Calculate the cumulative income shares for China in 1980 and 2014, the US in 1980 and 2014

Task 2: Plot the Lorenz Curves

- Plot the four vectors and add the perfect equality line
- Use base `R` plot tools or `ggplot`
- Compare the distribution of income across time for each country
- Compare the distribution of income across countries for each year

Task 3: Calculate the Gini Coefficients

- The Gini coefficient is calculated by dividing the area between the Lorenz curve and the perfect equality line, by the total area underneath the perfect equality line
- It ranges from 0 (complete equality) to 1 (complete inequality)
- Intuitively, the further away the Lorenz curve is from the perfect equality line, the more unequal the income distribution is, and the higher the Gini coefficient will be
- Calculate the Gini coefficient for each of your Lorenz curves (of course someone has already created a package for that!)
- Label each Lorenz curve with its corresponding Gini coefficient

Task 4: Calculate All Gini Coefficients

- For each country-year combination calculate a Gini coefficient
- Clearly, do not copy-paste the code 4,799 times, but use a `for` loop to go through all elements of the vector
- Which countries are the most equal (Gini coefficient < 0.2)?
- Which countries are the most unequal (Gini coefficient > 0.75)?
- Select a small list of countries, to plot their Gini coefficients over time
- Is inequality increasing or decresing over time?

- Select ten countries
- Plot a bar chart of Gini coefficients for 1980 and separately for 2014
- Order the bars
- Plot a bar chart of both 1980 and 2014 coefficients for selected countries on the same bar chart

References and Resources

- Doing Economics ▸ Tutorial