Wednesday, 19 June

Practice 3: Concentration index and Slope index of inequalities

Duration: 2h

In this practical session we aim to explore means of quantifying the degree of income-related inequality in a specific health variable, focusing in two indicators: the concentration index and the slope index of inequality.

# Concentration index (CIX)

The CIX is defined with reference to the concentration curve, using an approach analogous to the Gini coefficient. It graphs on the x-axis the cumulative percentage of the sample, ranking individuals according to socioeconomic position. The cumulative percentage of the health variable is plotted on the y-axis.

1. Using Microsoft Excel®

Before using the Stata codes, we will compute the CIX and its standard error (SE) from grouped data. For this purpose, we are going to use an Excel spreadsheet available among the files previously provided. Please, make sure that you have the file ‘concentration\_index.xls’.

The concentration index is easily computed in a spreadsheet program using the following formula[[1]](#footnote-1):

**C = (p1L2-p2L1) + (p2L3-p3L2) + … + (pT-1LT-pTLT-1)**

Where **p** is the cumulative percent of the sample ranked by economic status – we will use the wealth index quintiles as indicator of economic status; **L(p)** is the corresponding concentration curve ordinate; and **T** is the number of socioeconomic groups.

The file has three sheets. In the first one, we have the question: ‘Do you have standard deviations for the group means?’. It will help you decide which one to use. If you do not have the SD of the estimates, use the sheet 'no std devs'; if you do have it, use sheet 'have std devs'.

We will estimate the CIX for ‘stunting’ (height-for-age < -2SD), using the wealth quintiles dataset (wiq DHS + MICS). For Albania (MICS 2005), as an example, we have the following estimates:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| country | year | source | wiqn | stunt\_r | stunt\_se | stunt\_N | stunt\_pop |
| Albania | 2005 | MICS | Q1 | 0.351764 | 0.031577 | 214 | 243.9873 |
| Albania | 2005 | MICS | Q2 | 0.280018 | 0.036474 | 211 | 231.5362 |
| Albania | 2005 | MICS | Q3 | 0.283129 | 0.036225 | 193 | 192.7294 |
| Albania | 2005 | MICS | Q4 | 0.219511 | 0.030918 | 239 | 228.683 |
| Albania | 2005 | MICS | Q5 | 0.181064 | 0.02602 | 222 | 183.4465 |

We will use the sheet ‘no std devs’, in which you will find the box below.

|  |  |  |
| --- | --- | --- |
| **ENTER DATA BELOW** | | |
| Quintile | # persons | quintile |
| per quintile | means |
| 1 |  |  |
| 2 |  |  |
| 3 |  |  |
| 4 |  |  |
| 5 |  |  |

wiq

stunt\_r

stunt\_pop

We have to enter the group sizes and estimates. No other data is needed. Note that the other cells are locked and they will be calculated automatically based on the data entered. If you select a locked cell, you will be able to visualize the hidden formula in the toolbar.

We also highlight that we use the expanded population equivalent instead of the absolute sample size, as it already considers the sample design through the adjustment for the sample weights.

1. C:\Users\Equidade 2\AppData\Local\Microsoft\Windows\Temporary Internet Files\Content.IE5\T07QTV27\MC900405972[1].wmfSelect the variables *stunt\_pop* and *stunt\_r* for Albania (MICS 2005) in your datasheet.
2. Copy the values and paste in the sheet ‘no std devs’ of the file ‘concentration\_index.xls’**.** The column ‘# persons per quintile’ is equals to the *stunt\_pop* in our dataset, as the ‘quintile means’ is equal to the *stunt\_r* column.
3. Take note of the value of concentration index estimated. What does it mean?

2. Using Stata codes

We can estimate the CIX using the Stata command **cixr**, which uses the following syntax:

**cixr <socioeconomic position> <health variable> [if] [in], [graph]**

Where the wealth index quintiles are used as indicators of *socioeconomic position*; and *health variable* indicates any intervention or outcome. You can also include conditions using the [if] option and/or draw a concentration curve using the [graph] option.

Let’s estimate the CIX of stunting for Albania (MICS 2005), similarly to what we did using Excel. In this case, we will write the syntax as follows:

First we need to take into account that variable wiq is in a string format and we need to generate a numeric one:

**sort country year wiq**

**by country year: gen wiqn= \_n**

Then, type

**cixr wiqn stunt\_r in 1/5, graph**

Note: You must have the ado file **cixr**in the folder ‘c:\ado\personal’ of your computer. The variable *wiqn* must be ordinal and ordered so that low is poor/worst and high is rich/best. We use the option ‘in 1/5’ to indicate that we want to include only the lines from 1 to 5, in which we have the data for Albania (MICS 2005).

1. C:\Users\Equidade 2\AppData\Local\Microsoft\Windows\Temporary Internet Files\Content.IE5\T07QTV27\MC900405972[1].wmfRun the **cixr** command using the syntax provided above.
2. Did you obtain the same estimate as using the Excel spreadsheet?
3. What does the graph show?
4. Now we aim to compare estimates of CIX for the same indicator across different countries/surveys. Estimate the cix of stunting prevalence for the Belarus (MICS 2005) and Vietnam (MICS 2010). How do you compare these estimates? What would you say about the income-related inequality in stunting among these countries?

# Slope index of inequalities (SII)

The SII is estimated through a different approach. It is typically derived through linear regression, but for prevalence indicators, it is more appropriate to use logistic regression. The independent variable in this regression are the midpoints of ranks obtained by ordering the sample by the explanatory variable when using grouped data. The ranks are scaled so that the values range from zero to one. Therefore, the SII represents the absolute difference in the fitted value of the health indicator between the top (score of 1) and the bottom (score of 0) of the socioeconomic scale.

We can estimate the SII using the Stata command **siilogit**, which uses the following syntax:

**siilogit** <*socioeconomic position*> <*health variable*> [if] [in], [nograph]

where the wealth quintiles are used as indicators of *socioeconomic position*; and *health variable* indicates any intervention or outcome. You can also include conditions using the [if] option and the command generates a graph that allows checking the linearity as a default. You can use the option [nograph] if the graph is not necessary.

Let’s estimate the SII of stunting for Albania (MICS 2005). In this case, we will write the command as follows:

**siilogit wiqn stunt\_r in 1/5**

Note: You must have the ado file **siilogit**in the folder ‘c:\ado\personal’ of your computer. The variable *wiqn* must be ordinal and ordered so that low is poor/worst and high is rich/best. We use the option ‘in 1/5’ to indicate that we want to include only the lines 1 to 5 from the dataset, in which we have the data for Albania (MICS 2005).

1. C:\Users\Equidade 2\AppData\Local\Microsoft\Windows\Temporary Internet Files\Content.IE5\T07QTV27\MC900405972[1].wmfRun the **siilogit** command using the syntax provided above.
2. How would you interpret the SII?
3. What does the graph show?
4. Now we aim to compare estimates of SII for the same indicator across different countries/surveys. Estimate the SII of stunting prevalence for the Belarus (MICS 2005) and Vietnam (MICS 2010). How do you compare these estimates? What would you say about the income-related inequality in stunting among these countries?

Now lets estimate both CIX and SII for different indicators for the same country.

1. C:\Users\Equidade 2\AppData\Local\Microsoft\Windows\Temporary Internet Files\Content.IE5\T07QTV27\MC900405972[1].wmfEstimate CIX and SII for measles vaccination and underweight using data from the Nigeria (DHS 2008).
2. Analyze the estimates and graphs. How would you interpret them? What would you say about the income-related inequality in both indicators?

**Templetes**



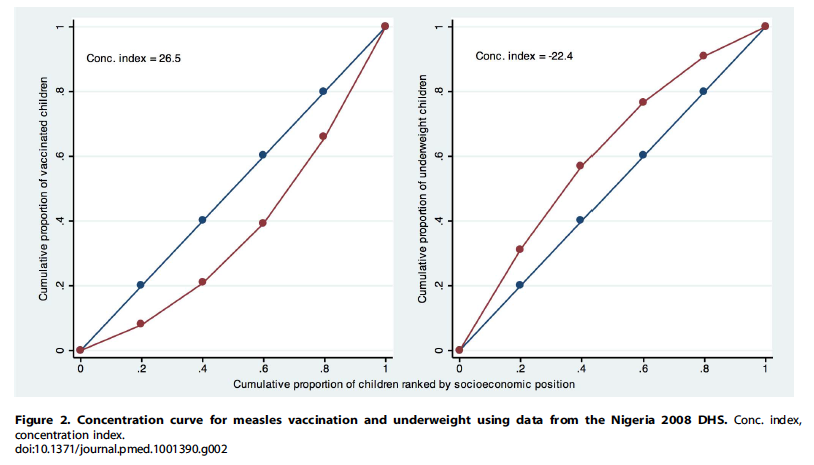
Albania (MICS 2005)



Belarus (MICS 2005)



Vietnam (MICS 2010)



1. N Kakwani, A Wagstaff & E van Doorslaer "Socioeconomic inequalities in health: measurement, computation and statistical inference", Journal of Econometrics 77 (1997) 87-103. [↑](#footnote-ref-1)