Solutions Problem Set 1B Social Welfare

Professor: Marcelo Neri TA: Tiago Bonomo

March 24, 2017

** Exercise 1

Define formally what is a dual distribution of a random variable X with mean µ and distribution such that the value of some measure of inequality is M.

Solution

A dual distribution is a distribution with the following characteristics:

- $X = 0$ with probability U_t and $X = \frac{\mu}{(1-U_t)}$ with probability $(1 - U_t)$. That is, the distribution has the original mean for any U_t.
- The inequality measure value is also equal to M, once we ajust the value of U_t.

** Exercise 2

Draw a Lorenz curve for a distribution suggested by the definition of a dual distribution.

Solution

** Exercise 3

Show that the dispersion measure coefficient of variation attends the Pigou-Dalton condition.

Solution

First, remember that the coefficient of variation is the ratio of the standard deviation to the mean of a certain variable. To show that it attends the Pigou-Dalton condition is equivalent of showing that it increases with a regressive transfer keeping the mean constant. As a transfer from one individual to another keeps constant the mean of the income distribution, we only need to show that the variance increases with a regressive transfer. Consider a population with size

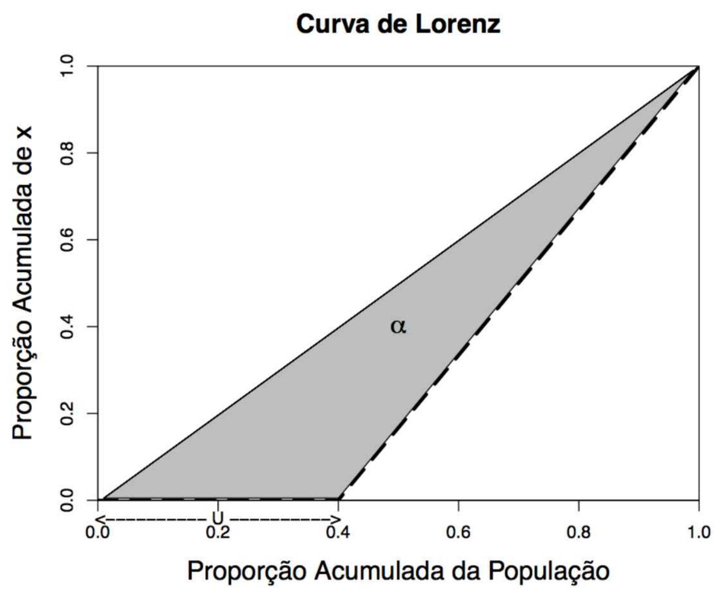

Figure 1: Lorenz Curve

N and x_i as the income of individual i. Let μ be the mean of the distribution, σ_o^2 the variance before the regressive transfer and σ_o^2 the variance after the transfer. Without loss of generality, we assume that $x_j > x_h$ and $x_j > x_h$. Let $\tau > 0$ be the value of the transfer. We have that

$$\sigma_o^2 = \frac{1}{N}(\sum_{i=1}^{N} x_i^2) - \mu^2$$

$$\sigma_1^2 = \frac{1}{N}\left[\sum_{i\neq h,j}^{N} x_i^2 + (x_h - \tau)^2 + (x_j + \tau)^2\right] - \mu^2$$

We have that

$$\sigma_1^2 - \sigma_0^2 = \frac{1}{N}\left[(x_h - \tau)^2 + (x_j + \tau)^2 - x_h^2 - x_j^2\right] = \frac{1}{N}\left[(x_h^2 - 2\tau x_h + \tau^2 + x_j^2 + 2\tau x_j + \tau^2 - x_h^2 - x_j^2\right]$$

$$\Rightarrow \sigma_1^2 - \sigma_0^2 = \frac{2\tau}{N}(\tau + x_j - x_h) > 0$$

because we assumed that $\tau > 0$ and $x_j > x_h$, we have that the variance always increases with a regressive transfer keeping the mean constant. Therefore, the coefficient of variation also increases and we can conclude that it attends the Pigou-Dalton condition.

Exercise 4

According to the general formula of a inequality measure

$S = \frac{1}{\varepsilon(1-\varepsilon)}\left[1 - \frac{1}{N}\sum_{i=1}^{N}(\frac{x_i}{\mu})^{1-\varepsilon}\right]$

Tell which measure of inequality it represents when $\varepsilon = -1, \varepsilon = 0$ and $\varepsilon = 1$

Solution

- $\varepsilon = -1$

We have that

$S = \frac{1}{2}\left[\frac{1}{N}\sum_{i=1}^{N}(\frac{x_i}{\mu})^2 - 1\right]$

$2S = \frac{\frac{1}{N}\sum_{i=1}^{N} x_i^2 - \mu^2}{\mu^2} = CV^2 \implies S = \frac{CV^2}{2}$

That is, when $\varepsilon = -1$ the measure is half of the square of the coefficient of variation.

- $\varepsilon = 0$

In this case, the value of S is not defined. We have to use L'Hôpital to calculate the limit when $\varepsilon \to 0$. Note that

$limS = \frac{1}{1-2\varepsilon}\frac{1}{n}\sum(\frac{x_i}{\mu})^{1-\varepsilon} ln(\frac{x_i}{\mu})$

When $\varepsilon \to 0$, we have that

$limS = \frac{1}{n}\sum(\frac{x_i}{\mu}) ln(\frac{x_i}{\mu}) = T$

That is, when $\varepsilon \to 0$, the measure approaches the Theil-T.

- $\varepsilon = 1$

Once again we will have to use L'Hôpital. When $\varepsilon \to 1$, we have that

$limS = -\frac{1}{n}\sum ln(\frac{x_i}{\mu}) = L$

That is, when $\varepsilon \to 1$, the measure approaches the Theil-L.

Exercise 5

Conceptually discuss the objectives, advantages and limitations of the following empirical techniques:

- Mincerian (or income) equation (Coefficients, R^2)
- Differences in Differences Analysis (compare the bivariate with the multivariate)

Solution

- Mincerian equation

The coefficients of a traditional Mincerian income equation are the return to education (or educational premium) and the return to experience. The interpretation of the coefficients is what is the impact in terms of variation on the wage (in %) of an increase in the years of schooling or experience. The R^2 gives us the power of the independent variables used in the equation to explain the variation in the dependent variable, that is, the log of the wage. The limitation is that we usually have econometric problems which bias the estimatives of the coefficients. The main problems are omitted variables (usually ability, which tends to overestimate the coefficient of the return) and measurement error (which bias the estimative towards zero - attenuation bias).

- Differences in Differences

The differences in differences analysis (DD) requires longitudinal data before and after the treatment. In the bivariate analysis, the comparison is between two groups (for example, whites and blacks) and the DD estimative gives us the average variation of the dependent variable (for example, wages) for the treated group compared to the control group. That is, it gives us the treatment effect with all regular hypothesis. In the multivariate analysis, we can consider many groups (for example, whites, blacks and asians) but only one will be the control group. In this way, we will have a DD estimative for each of the treatment groups. To capture a causal effect of the treatment in a given dependent variable, it is necessary that all the groups considered have the same tren in the absence of the treatment.

Exercise 6

Econometric Interpretation

Using the regression below, discuss the level and the evolution of the differentials in income by education in Brazil between 2001 and 2009. How to interpret the two coefficients in bold?

Solution

First, we can see that the coefficient ANO2009 is equal to -0.1652, which means that the returns to education have fallen 16.52% from 2001 to 2009 (ANO2001 is the baseline).

The coefficient EDUCA03 corresponds to the returns to education for the individuals with 0 to 3 years of schooling for the years of 2001 and 2009 pooled. Following the same logic, EDUCA48 is corresponds to the returns for the individuals with 4 to 8 years of scholing, EDUCA812 for individuals with 8 to 12 years and EDUCA12 for individuals with 12 or more years (which is the baseline). Note that the negative coefficients were expected, that is, less educated people earn less than more educated, and the relation is monotonic as we can see looking at the coefficients. Analysing the evolution of the returns for different levels of schooling (looking at the interaction between levels of schooling and the years 2001 and 2009), we can see that, despite earning less, the increases in the returns for the less educated were higher than the increases for the more educated (or the decreases were lower). Note that this relationship is also monotonic, with the coefficient ANOEDUCA032009 being the highest, followed by ANOEDUCA482009, ANOEDUCA8122009 and finally ANOEDUCA122009.

Exercise 7

Define and compare the uses of the following conceps using capsular formulas and graphs:
a) Lorenz Curve and the Generalized Lorenz Curve
b) Lorenz Curve and Concentration Curve
c) Gini index and Concentration Ratio
d) Absolute and relative concepts of inequality

Solution

a) The Lorenz Curves allow us to compare different income distributions in terms of inequality if we have Lorenz dominance (see exercise 2.5 in Problem Set I). On the other hand, Lorenz curves imply welfare dominance only when one compares distributions with the same mean. Shorrocks (1983) and Kakwani (1984) developed a criterion to compare distributions with different means in terms of welfare using the concept of the Generalized Lorenz Curve. The Generalized Lorenz Curve is a modification to the Lorenz Curve in which the accumulated fraction of incomes up to each fraction of the population is multiplied by the average income of the distribution. Because of this multiplication, the generalized curve brings information about the form and level of the distribution, or the joint first two moments of the distribution such as the income distribution curve and its congeners of basic statistics. The Generalized Lorenz Curve is represented by a function $L(\mu, P) = \mu L(P)$. If the Generalized Lorenz Curve of a distribution A is always above the Generalized Lorenz Curve of a distribution B, we can say unequivocally that distribution A has a superior welfare than distribution B (see more in handout 3).

b) The Concentration Curves are a representation that bears similiarities with the Lorenz Curve. However, while the latter refers to the distribution of a single variable throughout the population, the former are constructed from the distribution of two variables in the population. In fact, the Lorenz Curve is a particular case of the Concentration Curves.

c) The Concentration Ratio is the Gini of when the Concentration Curve is the Lorenz Curve, that is, the Concentration Curve of itself. While the Gini ranges from 0 to 1, the Concentration Rate ranges from -1 to +1, where -1 is completely pro-poor and +1 completely pro-rich.

d) First of all, remember that almost all of the inequality measures we say are relative, in the sense that if everybody's income increases by the same percentage, inequality remains unchanged. The problem is that an equal percentage increase for all corresponds to absolute gains that may be extremely unequal. For example, a person who has an income one hundred times higher that another one will also have absolute gains that are one hundred times greater. So the question of why we use relative measures or why they are better arises.

First, relative income measures are conservative because they show no change in inequality in cases where absolute measures would show an increase (when all incomes go up by the same percentage) or a decrease (when they all go down by the same percentage). Inequality is a very important and inflammatory topic, so conservatism (in terms of measurement) is to be preferred.

Second, one of the disadvantages of absolute measures is that they are bound to increase with practically any increase in the mean: when incomes rise, the absolute distance between the rich, the middle class, and the poor becomes greater even if the relative gaps remain the same. Focus on absolute distances presents the disadvantage that practically every increase in the mean could be judged to be pro-inequality. We would lose the sharpness with which we can currently distinguish between pro-poor and pro-rich growth episodes.

Third, inequality and income growth are just two manifestations of the same phenomenon. Focus on the absolutes in growth, as in inequality, would lead us to nearly always find that growth in rich countries, however small in percentage terms, would be greater than growth in poor countries, however huge, so the logic of relativity that applies to growth should also apply to inequality.

Finally, a relative increase in income correlates with gains in utility if we believe that personal utility functions are logarithmic in income. In other words, one additional dollar will yield less utility, or seem less important, to a rich person than to a poor person. By this route too, we come to the conclusion that relative income changes are a more reasonable metric than absolute income changes.

Nevertheless, in the end of the day the choice of relative or absolute inequality mneasures is subjective and based on normative values. By no coincidence people who are in favor of absolute measures of inequality tend to favor more pure redistributive measures even when they imply a loss in the mean or a general welfare loss in a Pareto sense, common in the economics profession.

For more details, see Milanovic (2016), page 27, EXCURSUS 1.2.

Exercise 8

Using the Atkinson's (1970) approach, derive an inequality measure for the share of the 10% richest in income.

Solution

We could use the ratio of the mean income of the 10% richest to mean income of the population. This is related to the measure of equality (or shared prosperity) proposed in Kakwani, Neri and Vaz (2014), where they use the ratio of the mean income of the 40% poorest to the mean income of the total population as a measure of relative equality or one minus this ration of a measure of relative inquality. The indicator they propose follows Atkinson's approach and is defined over individual incomes. Therefore, if we want to derive an inequality measure for the share of the 10% richest in income using Atkinson's approach, we could use the same idea as in Kakwani, Neri and Vaz (2014). We should perhaps stress the fact that the ratio between the mean income of the bottom 40% to the mean income of the total population is a measure of equality so the same ratio in the case of the top 10% should be denoted as measure of inequality.

Exercise 9

Comment, using capsular formulas and graphs:

a) The most popular inequality measures, like the Gini and the Theil-T indexes, are not very useful to discuss redistributive anti-poverty measures.

b) The advantage of the J-divergence over the Theil-T is to allow decompositions betweeen and within groups and also to incorporate null incomes.

c) The advantage of the J-divergence over the Gini and the Theil-T indexes is to allow to decompose inequality in non negative income shares contributions.

Solution

ad

a) The sentence is true. The Gini and the Theil-T indexes are not very sensible to changes in the bottom of the distribution of income. In handout 19, we saw that the turning points of inequality (that is, the percentil for which a marginal income increase leads to an increase in inequality) is the percentile 75 for the Gini and 87 for the Theil-T, using data from PNAD. If we combine with data from Personal Income Tax (PIT), the turning points change to percentile 81 for the Gini and 95 for the Theil-T. That is, these measures are not pro-poor and therefore not very useful to discuss redistributive anti-poverty measures. Examples of more pro-poor measures are the one that combines the log of the Theil and the weights of the Gini and the one that uses explicitly the inequality measure derived directly from a poverty objective function (will be seen in the course).

b) The sentence is false. The J-divergence allows decompositions between and within groups but this is not an advantage over the Theil-T because the last also allows these decompositions. However, because the J-divergence is a combination of the Theil-T and the Theil-L indexes and the Theil-L doesn't incorporate null incomes (the index goes to infinity if we have null income for any individual), we have that the J-divergence doesn't incorporate null incomes. The Theil-T, in turn, does incorporate null incomes (see handouts 17 and 19).

c) The afirmation is true. J-divergence implies in shares (always non negative) for each income-bracket and individuals. That is, the J-divergence allows us to decompose inequality in non negative income shares contributions.