

Investigation on income inequality of the Islands

STAT3003 Midterm A - Survey Plan

ZHENG Weijia (William, 1155124322)

April 6, 2021

1 Introduction and Notations

Income inequality continues to receive world-wide attention. The most famous method to measure it to use the Gini Index proposed by Gini. Gini index gives a number inside the interval $[0, 1]$, the more the index close to 1, the worse the inequality problem is.

Assume we will investigate the situation of income inequality for the scale of a town (e.g., Hofn, Vardo and Takazaki are all towns).

Then denote the population of the town as n , note that this is given.

Denote the total amount of wealth of the town as M , which is also given by consulting the Hall of the town.

Denote the average amount of wealth as $\mu = \frac{M}{n}$.

For the i -th individual (i ranges from 1 to n), denote the his individual amount of wealth as m_i , note that $\sum_{i=1}^n m_i = M$.

Then if we follow the definition of Gini index, we need to calculate

$$G = \frac{1}{2n^2\mu} \sum_{j=1}^n \sum_{i=1}^n |m_i - m_j|.$$

But this is troublesome because it is hard to estimate the G of a town properly from a sample of smaller size. Hence we switch to another method of measuring the inequality.

2 Mathematical Model

As said above, hindered by the complex way to calculate the Gini index out and noted that the Islands provide the information of the total amount of a town and every individual's. We can then understand the concept of "income euquality" from the respective of the difference between "every person's ability (as a possibility) to gain wealth".

For the i -th person, we regard the ratio between his own wealth and the total wealth, i.e., $p_i = \frac{m_i}{M}$ of the town as his "possibility of gaining a unit of wealth".

Note that $\sum_{i=1}^n p_i = 1$, hence the p_i forms a proper probability distribution. Note that the entropy

$$H(p_1, \dots, p_n) = - \sum_{i=1}^n p_i \log_n p_i$$

actually measures the inequality of the town economy. Hence define it as the " H index" of the economy.

When the town economy is more fair, i.e., everyone has the same probability of earning money: $\frac{1}{n}$, then

$$H(p_1, \dots, p_n) = - \sum_{i=1}^n p_i \log_n p_i = - \sum_{i=1}^n \frac{1}{n} \log_n \frac{1}{n} = 1.$$

When the economy is extremely unequal, i.e., all the money go to one person, then the entropy will be 0.

The following is a simple example of the application of this Mathematical model. Consider a economy with only 3 people, each income is 10, 20, 15. And another economy with 5 people with each income 3, 20, 16, 40. The first economy has $H_1 = 0.9656$ and the second economy with $H_2 = 0.8223$, which conforms our intuition.

Note that the H index is a τ -type metric, so all the knowledge for estimating τ can be applied properly.

3 Sampling Method

4 Preliminary Survey Result