Handling incomplete information: Gini coefficient from coarse-grained data

Szabolcs Kelemen, Máté Józsa, Zoltán Néda
Babeş-Bolyai University, Cluj-Napoca, Romania

Acquiring income or wealth data at agent-level for large social systems is an experimentally challenging task [1, 2]. The absence of such exhaustive data poses significant difficulties for accurately characterizing the level of inequality in social systems, particularly when it comes to calculating the commonly used Gini coefficient [3]. To overcome this challenge, we introduce a novel methodology that enables a better computation of the Gini coefficient using only the average income or wealth and population at a lower hierarchical level, such as settlements or communes. Our approach is based on the assumption that a one-parameter wealth/income distribution is valid at the lower organizational level, and we validate this hypothesis using various data sources. These include exhaustive income data for Cluj county, Romania and average income data at the settlement level, as well as percentile and average income data for settlements from Hungary and the United States. Our methodology provides a compact analytical formulae that allows the calculation of the Gini coefficient from datasets where only statistical information, like population and measures of general well-being (i.e. average income or wealth), are available for the lower hierarchical level. Although, the proposed approach has its limitations, it represent a significant step forward in quantifying inequality in large social systems. In contrast to traditional Gini coefficient calculation methods based on statistical data, which can often lead to coarse-graining of inequalities at lower level, our methodology provides a more accurate estimation of the inequality level.

References

[1] N. Derzsy, Z. Neda, M. A. Santos, Income distribution patterns from a complete social security database, Physica A 391, 5611-5619 (2012).
[2] I. Gere, Sz. Kelemen, T.S. Biró, Z. Néda, Wealth Distribution in Villages. Transition From Socialism to Capitalism in View of Exhaustive Wealth Data and a Master Equation Approach, Front. Phys. 10, 827143 (2022).
[3] C. Gini, Colorado College Publication, On the Measure of Concentration with Special Reference to Income and Statistics, General Series 208, 73-79 (1936).