Flex Conference (Physical / Digital)

International Conference on Mathematical Applications (ICMATAP - 25)

7th - 8th March 2025, Paris, France

ANALYSIS OF SKEWNESS IN GINI INDEX ESTIMATION AND ITS IMPACT ON CONFIDENCE INTERVALS

Encarnación Álvarez Verdejo*, Pablo J. Moya-Fernández, Raúl Amor Pulido, Juan Francisco Muñoz Rosas

University of Granada **encarniav@ugr.es*

Objective

To evaluate the main methodologies for constructing confidence intervals (CI) for the Gini Index in the presence of skewness.

State of research

Definition: Let Y a non-negative continuous random variable, the Gini index can be defined as:

$$G = \frac{1}{2\mu_Y} \int_0^\infty \int_0^\infty |x - y| \, dF_Y(x) dF_Y(y)$$

where $\mu_Y = E[Y] = \int_0^\infty y dF_Y(y)$ is the mean of Y and $F_Y(y)$ is the distribution function of Y.

- G is a measure of inequality that takes values between 0 y 1; 0 being a perfectly equal situation and 1 the opposite.

The definition of the estimator and the CI can be consulted though the QR.

Methodology

A Monte Carlo simulation study is carried out to evaluate the CIs of the Gini Index based on its empirical Coverage Rate (CR). The following scenarios are analysed:

- Infinite populations with Gamma (low skewness) and Pareto (high skewness).
- Medium sample sizes and high sample sizes.
- Low (0.1), medium (0.3) and high (0.5) Gini Index.

Conclusions

- In scenarios with low asymmetry, the CIs generally have values close to the optimal CR. The OLS method performs worse.
- In scenarios with high asymmetry, the CIs have worse CR values. The t-Jackknife method is recommended.

References

Muñoz, J.F., Pavía, J.M. & Álvarez-Verdejo, E. (2023). giniVarCI: Gini Indices, Variances and Confidence Intervals for Finite and Infinite Populations. R packages version 0.0.1-3. https://CRAN.R-project.org/package=giniVarCI

Results