Società Italiana di Statistica

ì

ATTI DELLA XLIV RIUNIONE SCIENTIFICA (2008)

XLIV Riunione Scientifica
(Arcavacata di Rende, 25-27 giugno 2008)

ISBN 978-88-6129-228-4
Editore: CLEUP – Padova

Francesca Greselin (Università di Milano-Bicocca), Leo Pasquazzi (Università di Milano-Bicocca)
Asymptotic Confidence Intervals for a New Inequality Measure

Asymptotic Confidence Intervals for a New Inequality Measure(*)

Intervalli di confidenza asintotici per una nuova misura di Ineguaglianza

Francesca Greselin, Leo Pasquazzi
Dip. Met. Quant. per le Scienze Economiche e Aziendali, Università di Milano Bicocca
e-mail: francesca.greselin@unimib.it

Keywords: Zenga's inequality measure, Gini's index, bootstrap confidence intervals

1. Introduction

Recently Zenga (2007) introduced a new inequality measure based on ratios between lower and upper group means. Differently from traditional inequality measures (i.e. Gini, Bonferroni, Theil), which may be viewed as a unique ratio of two means, the new synthetic inequality measure is defined as a mean of several ratios. In this paper the performance of asymptotic confidence intervals for Gini's measure and for the new measure is tested. Several types of confidence intervals are considered: the normal, the percentile, the BCa and the t-bootstrap. Definitions of these confidence intervals may be found in Davison and Hinkley (1997). While the underlying asymptotic theory for Gini's measure is well established, formal proofs for Zenga's index are still missing in literature. Nevertheless, also in view of our simulation results, asymptotic properties similar to those of Gini's index can be expected to hold also for Zenga's new inequality measure.

2. Simulation results

This section describes the results of a simulation study about asymptotic confidence intervals for Gini's and Zenga's inequality functionals defined respectively by

$$G(F) = \frac{\int_0^\infty \int_0^\infty |x-y| dF(x) dF(y)}{2\int_0^\infty x dF(x)}, \qquad \text{and} \qquad Z(F) = 1 - \int_0^\infty \frac{\frac{\int_0^x y dF(y)}{F(x)}}{\frac{\int_x^\infty y dF(y)}{1-F(x)}} dF(x).$$

We drew 10000 samples of size 100, 200 and 400 from the Dagum distribution with probability density function given by

$$f(x) = \lambda\beta\theta x^{-(\theta+1)} \left(1+\lambda x^{-\theta}\right)^{-(\beta+1)}, \qquad (\lambda, \beta, \theta > 0,\ x > 0).$$

The parameters were set to their ML estimates ($\beta = 1.055$, $\theta = 3.095$ and $\lambda = 1$) obtained on the Italian expenses distribution as given by the Banca d'Italia survey on Household Income and Wealth in 2002 (8001 households). This yields $G = 0.3193$ and $Z = 0.6505$. The bootstrap distributions have been obtained by taking 9999 resamples from each sample and by computing the functionals or their standardized version (for the t-bootstrap confidence intervals) on each resample. As suggested by Efron (1987)

(*) This research was partially supported by FAR Università Milano-Bicocca

we approximated the acceleration constant a needed for the BCa confidence intervals by $\widehat{a} = \sum_{i=1}^{n} h(\widehat{F}; X_i)^3 / [6(\sum_{i=1}^{n} h(\widehat{F}; X_i)^2)^{3/2}]$, where $h(\widehat{F}; X_i)$ are the empirical influence values obtained by numerical differentiation. Finally, we used $\widehat{\sigma}^2 = \sum_{i=1}^{n} h(\widehat{F}; X_i)^2/n$ as variance estimate for the t-bootstrap confidence interval.

Table 1: *Coverage accuracy and average size (only for the 95% nominal confidence level) of the asymptotic confidence intervals*

			Confidence intervals for Gini's index							
			normal					**percentile**		
n	***0.9000***	***0.9500***	***0.9750***	***0.9900***	***0.95 av.sz.***	***0.9000***	***0.9500***	***0.9750***	***0.9900***	***0.95 av.sz.***
100	0.8228	0.8842	0.9196	0.9468	0.1114	0.7986	0.8571	0.8933	0.9269	0.1095
200	0.8418	0.9005	0.9331	0.9592	0.0850	0.8271	0.8880	0.9183	0.9489	0.0840
400	0.8643	0.9194	0.9482	0.9689	0.0641	0.8539	0.9108	0.9421	0.9637	0.0636
			BCa					**t-bootstrap**		
100	0.8334	0.8959	0.9347	0.9593	0.1162	0.8642	0.9226	0.9524	0.9734	0.1512
200	0.8487	0.9020	0.9416	0.9670	0.0896	0.8699	0.9232	0.9575	0.9767	0.1064
400	0.8602	0.9181	0.9496	0.9733	0.0676	0.8780	0.9338	0.9608	0.9810	0.0763
			Confidence intervals for Zenga's index							
			normal					**percentile**		
n	***0.9000***	***0.9500***	***0.9750***	***0.9900***	***0.95 av.sz.***	***0.9000***	***0.9500***	***0.9750***	***0.9900***	***0.95 av.sz.***
100	0.8407	0.9032	0.9418	0.9662	0.1272	0.8157	0.8743	0.9115	0.9410	0.1258
200	0.8522	0.9102	0.9476	0.9715	0.0945	0.8313	0.8915	0.9235	0.9554	0.0938
400	0.8641	0.9236	0.9560	0.9758	0.0702	0.8569	0.9092	0.9417	0.9670	0.0698
			BCa					**t-bootstrap**		
100	0.8472	0.9075	0.9449	0.9694	0.1262	0.8647	0.9276	0.9576	0.9778	0.1491
200	0.8523	0.9098	0.9463	0.9728	0.0952	0.8696	0.9268	0.9613	0.9807	0.1081
400	0.8633	0.9212	0.9529	0.9762	0.0713	0.8769	0.9344	0.9634	0.9829	0.0786

The results reported in Table 1 show that the confidence intervals suffer from undercoverage. Nevertheless, it appears that the coverage accuracy of the t-bootstrap confidence intervals is reasonably close to the nominal confidence level for a sample size of 400. If large confidence levels are required, the number of bootstrap resamples has to be increased beyond 9999 in order to get meaningful BCa confidence intervals. Indeed, 825 (out of 10000) BCa confidence intervals for Gini's index, with samples of size 400 and 99% nominal confidence level, were based on the 100th percentile of the bootstrap distribution (for Zenga's index this figure is 659). Besides this, it is worth noting that the confidence intervals for Zenga's index are slightly larger - as they estimate a greater quantity - but have a better coverage probability than those referred to Gini's index. One might also conjecture that this is due to the peculiar structure of the new index, which is a mean of several ratios, instead of being a unique ratio of two means, as Gini's measure and other well known inequality measures are. Further research is underway in order to establish whether in some real situations the new index should be preferred to Gini's index.

References

Davison A.C., Hinkley D.V. (1997) *Bootstrap Methods and their Application*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge.

Efron B. (1987) Better Bootstrap confidence intervals (with discussion), *Journal of the American Statistical Association*, 82, 171-200.

Zenga M. (2007) Inequality curve and inequality index based on the ratios between lower and upper arithmetic means, *Statistica & Applicazioni*, 1, 3-27.