Characterizations of Income Distributions and the Moment Problem of Order Statistics

Christian Kleiber
Universität Dortmund, Fachbereich Statistik
Vogelpothsweg 78
44221 Dortmund, Germany
kleiber@statistik.uni-dortmund.de

Samuel Kotz
George Washington University, Department of Engineering Management
2130 H Street, NW
Washington, DC 20052, U.S.A.
kotz@seas.gwu.edu

1. Introduction

A basic feature of empirical income distributions is that they are well approximated by models with polynomially decreasing tails, i.e. $1 - F(x) \sim x^{-\alpha}$, for some $\alpha > 0$. Such models cannot be characterized in terms of their moments, since only a few of the moments exist. Recently Aaberge (2000), viewing the Lorenz curve as a distribution function, has pointed out that any income distribution with a finite mean is characterized up to a scale by the moments of its 'Lorenz curve distribution'. These moments turn out to be affine functions of Kakwani's (1980) generalized Gini coefficients. We show that income distributions can be characterized directly in terms of these generalized Gini coefficients.

2. A characterization of income distributions

Consider a sample of size n from a distribution with the c.d.f. F and define the order statistics $X_{k:n}$ in the ascending order by $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$. The moment problem of order statistics inquires to what extent the c.d.f. F is uniquely determined by (a subset of) the first moments of all of its order statistics $\{E(X_{k:n}) \mid k = 1, 2, \ldots, n; n = 1, 2, 3, \ldots\}$.

A finite mean of the parent distribution assures the existence of the first moment of any order statistic, hence characterizations in terms of the moments of order statistics are of interest for heavy-tailed distributions of the Pareto type. The basic characterization result is as follows:

Lemma 1: *Let $E|X| < \infty$. Then F is uniquely determined by the sequence $\{E(X_{1:n}) \mid n = 1, 2, 3, \ldots\}$.*

An income distribution has the property that its c.d.f. F is supported on the nonnegative halfline. All income distributions with $F \in \mathcal{L} := \{F | 0 < \int x\,dF(x) < \infty\}$ admit a Gini coefficient $G = 1 - 2\int_0^1 L(p)dp$, where $L(\cdot)$ denotes the Lorenz curve of F. Kakwani (1980) proposed a one-parameter family of generalized Gini coefficients,

$$G_n = 1 - n(n-1)\int_0^1 L(p)(1-p)^{n-2}\,dp.$$

From Arnold and Laguna (1977) the traditional Gini coefficient, i.e. G_2, may be written as $G_2 = 1 - E(X_{1:2})/E(X)$, which generalizes to $G_n = 1 - E(X_{1:n})/E(X)$. The following characterization is now immediate via Lemma 1:

Theorem 2: *Any $F \in \mathcal{L}$ is characterized up to a scale by its sequence of generalized Gini indices.*

Aaberge (2000) recently obtained a similar result using a different approach. Considering the Lorenz curve as a distribution function, he observes that this 'Lorenz curve distribution' has bounded support and is therefore characterized by the sequence of its moments. These moments turn out to be affine functions of the generalized Gini coefficients presented above. Consequently, our Theorem 2 also follows from his result.

Further refinements of Theorem 2 are possible via Müntz's theorem (see, e.g., Huang 1989). An example is

Theorem 3: *Any $F \in \mathcal{L}$ is determined up to a scale by any subsequence of generalized Gini indices $\{G_{n_j}\}$ for which Müntz's condition $\sum_{j=1}^{\infty} n_j^{-1} = \infty$ holds.*

REFERENCES

Aaberge, R. (2000). Characterizations of Lorenz curves and income distributions. *Social Choice and Welfare*, **17**, 639-653.

Arnold, B.C., and Laguna, L. (1977). *On generalized Pareto distributions with applications to income data*, International Studies in Economics no. 10, Dept. of Economics, Iowa State University, Ames, Iowa.

Huang, J.S. (1989). Moment problem of order statistics: A review. *International Statistical Review*, **57**, 59-66.

Kakwani, N. (1980). On a class of poverty measures. *Econometrica*, **48**, 437-446.

RESUME

All income distributions with a finite mean can be characterized in terms of the sequence of their generalized Gini coefficients. This is connected with the moment problem of order statistics.