The generalized lognormal distribution as an income distribution

Christian Kleiber
Universität Dortmund, Fachbereich Statistik
Vogelpothsweg 78
D-44227 Dortmund, Germany
kleiber@statistik.uni-dortmund.de

and

Stefanie Scheid
Max Planck Institute for Molecular Genetics, Dept. of Computational Molecular Biology
Ihnestraße 63-73
D-14195 Berlin, Germany
stefanie.scheid@molgen.mpg.de

1. Introduction

Over the last 100 years, a large number of distributions has been proposed for the modeling of size phenomena, notably the size distribution of personal incomes (see, for example, Kleiber and Kotz, 2003). The most widely known of these models are the Pareto and the lognormal distributions; however, both are known to be not flexible enough for the modeling of income data. Here we consider a generalization of the lognormal distribution for which the available literature is mainly in Italian.

2. A generalized lognormal distribution

Vianelli (1982, 1983) proposed a three-parameter generalized lognormal distribution with PDF

$$f(x) = \frac{1}{2xr^{1/r}\sigma_r\Gamma(1+1/r)} \exp\left\{-\frac{1}{r\sigma_r^r}|\log x - \mu|^r\right\}, \quad 0 < x < \infty. \tag{1}$$

Here e^μ is a scale parameter and σ_r, r are shape parameters. The distribution can be considered a generalization of the lognormal distribution, which is obtained for $r = 2$, and also a generalization of the log-Laplace distribution, which is obtained for $r = 1$.

If we consider the random variable $Y = \log X$, for X following the distribution (1), we obtain a distribution with the PDF

$$f(y) = \frac{1}{2r^{1/r}\sigma_r\Gamma(1+1/r)} \exp\left\{-\frac{1}{r\sigma_r^r}|y - \mu|^r\right\}, \quad -\infty < y < \infty, \tag{2}$$

where now $-\infty < \mu < \infty$ is the location parameter, $\sigma_r = [E|Y-\mu|^r]^{1/r}$ is the scale parameter, and $r > 0$ is the shape parameter. This distribution is perhaps best known as the exponential power distribution; it has been used in robustness studies and also as a prior distribution in various Bayesian models (e.g., Box and Tiao, 1973).

This connection permits us to derive a number of properties of the generalized lognormal distribution from known results for the exponential power distribution. For example, random number generation is possible via a mixture representation provided by Devroye (1986). Of course, ML estimation is completely analogous for both models. In addition, we present a number of functionals which are of interest in connection with income data, including Theil-type entropy measures and related notions.

3. Applications

Brunazzo and Pollastri (1986) applied the generalized lognormal distribution to the Italian incomes for the year 1948, obtaining a shape parameter $r = 1.45$, halfway between the lognormal and log-Laplace special cases. Here we consider German household incomes from the 1993 German income and expenditure survey (sample size: 40,230). It turns out that the generalized lognormal distribution improves upon the classical lognormal for these data, with an estimate of the additional parameter r in the vicinity of 1.9 (Scheid, 2001). However, nonparametric goodness of fit tests still reject the model, a feature not uncommon with samples of this size.

REFERENCES

Box, G.E.P., and Tiao, G. (1973). *Bayesian Inference in Statistical Analysis.* Reading, MA: Addison-Wesley.

Brunazzo, A., and Pollastri, A. (1986). Proposta di una nuova distribuzione: la lognormale generalizzata. In: *Scritti in Onore di Francesco Brambilla*, Vol. 1, Milan: Ed. Bocconi Comunicazioni, pp. 57-83.

Devroye, L. (1986). *Non-Uniform Random Variate Generation.* Berlin and New York: Springer.

Kleiber, C., and Kotz, S. (2003). *Statistical Size Distributions in Economics and Actuarial Sciences.* Hoboken, NJ: John Wiley.

Scheid, S. (2001). Die verallgemeinerte Lognormalverteilung. Diploma thesis, Dept. of Statistics, Universität Dortmund, Germany.

Vianelli, S. (1982). Sulle curve lognormali di ordine r quali famiglie di distribuzioni di errori di proporzione. *Statistica*, 42, 155-176.

Vianelli, S. (1983). The family of normal and lognormal distributions of order r. *Metron*, 41, 3-10.

RÉSUMÉ

This paper presents a number of properties of a generalized lognormal distribution which has been suggested as a model for economic size distributions in the Italian language statistical literature.