STATISTICA, anno LXVI, n. 3, 2006

THE STOCHASTIC INTERPRETATION OF THE DAGUM PERSONAL INCOME DISTRIBUTION: A TALE

L. Fattorini, A. Lemmi

In November 1975, one of the authors was involved by C. Scala in a research project aimed at the preparation of an exhaustive handbook on the probability density functions of wide applicability in scientific works (Scala, 1981). To this end, three years of intense bibliographic research was carried out in important journal of economics, finance, biology, forest science, ecology, medicine and other sciences. One year later, in Novembre 1976, while searching in meteorological journals, a family of distributions referred to as the *kappa(3) distribution* (Mielke, 1973 and Mielke and Johnson, 1973) was encountered for the first time. Mielke and Johnson, (1973, p. 701) emphasized that that the kappa(3) distribution "*is relatively easy to work with, possesses a simple closed form for both its density and cumulative distribution function and appears to be appropriate in applications involving precipitation data*".

The kappa(3) distribution's density function reported in these works was of the form

$$f(x)=\frac{\kappa\theta}{\delta}\frac{\left(\frac{x}{\delta}\right)^{\theta-1}}{\left\{\kappa+\left(\frac{x}{\delta}\right)^{\kappa\theta}\right\}^{1/\kappa+1}} \quad ; \quad x>0 \tag{1}$$

where $\kappa,\theta>0$ are shape parameters and $\delta>0$ is the scale parameter. It is at once apparent that (1) is completely equivalent to the density function proposed one year later by Dagum (see Dagum, 1977) for income distributions.

Bryson (1974) recognized the kappa(3) distribution as a heavy-tailed distribution. Owing to its heavy-tailedness and its marked positive skewness, even if unaware of the Dagum (1977) work, the authors immediately recognized the kappa(3) distribution as a suitable candidate for describing income distributions over the whole range $(0,\infty)$. Accordingly, in the beginning of 1978, the reliability of the kappa(3) model was checked on a wide variety of personal income frequency distributions. More specifically, the following income distributions were considered: Italy 1967(1)1976, U.S.A. 1954(1)1957, Sweden 1960(1)1962, 1965,

1966. The kappa(3) distribution was also compared with the log-normal and Champernowne distributions, which, at the moment, were very familiar in income studies. All the distributions adopted in the study were fitted by means of a quantile-quantile plot fitting technique.

The results of the comparison, subsequently reported in Fattorini and Lemmi (1979), turned out to be very encouraging. Indeed, while the superiority of the kappa(3) distribution with respect to log-normal distribution appeared quite trivial, owing to the presence of a unique shape parameter in the log-normal model, what appeared to be more relevant was the comparability of the kappa(3) performance with respect to performance of the Champernowne model. Indeed, the kappa(3) model proved to never be worse than the Champernowne model, a widely applied and deeply analysed family of distributions proposed by Champernowne (1953) with the specific aim of describing income data.

Like the kappa(3) distribution, the Champernowne distribution involves two shape parameters, say $\lambda \geq -1$ and $\theta > 0$, with density function

$$f(x) \propto \frac{\left(\dfrac{x}{\delta}\right)^{\theta-1}}{\left\{1+2\lambda\left(\dfrac{x}{\delta}\right)^{\theta}+\left(\dfrac{x}{\delta}\right)^{2\theta}\right\}} \quad ; \quad x>0 \tag{2}$$

where $\delta > 0$ is the scale parameter. It is worth noting that the normalising constant in (2) attains three different forms in accordance with $|\lambda| < 1$, $\lambda = 1$ and $\lambda > 1$.

As already mentioned, the properties of the Champernowne distribution were deeply investigated. Champernowne (1952, 1953) presented a considerable body of evidence in support of (2), emphasizing that the distribution has paretian tails at both ends, which probably constituted one of the main attraction of the model. Moreover, the rationale for the Champernowne's model had been recently illustrated in a seminal article by Ord (1975), which derived (2) as the equilibrium distribution of a diffusion process.

Thus, in order to provide a comparable body of evidence for the kappa(3) distribution, the authors (who in the beginning of 1979 were still unaware of the Dagum works) proved that (1) has a paretian tail at its right end (see Fattorini and Lemmi, 1979) and also attempted a derivation of (1) as the equilibrium distribution of a diffusion process.

The key point of the derivation (Fattorini and Lemmi, 1979) was the intersection between the Champernowne and the kappa(3) families of distributions, i.e. the distribution with density function

$$f(x) = \frac{\theta}{\delta}\frac{\left(\dfrac{x}{\delta}\right)^{\theta-1}}{\left\{1+\left(\dfrac{x}{\delta}\right)^{\theta}\right\}^{2}} \quad ; \quad x>0 \tag{3}$$

which may be viewed as a member of the kappa(3) family with $\kappa = 1$ or a member of the Champernowne family with $\lambda = 1$. This distribution is usually referred to as *log-logistic distribution* since its logarithmic transform $Y = \ln X$ has a logistic distribution.

Ord (1975) derived (3) as the equilibrium distribution of a diffusion process in which the logarithmic transform has an infinitesimal mean and variance respectively given by

$$b(y,t) = -\frac{\sigma^2}{2\beta}\{1 + h\ e^{-\zeta t}\},\ a(y,t) = \sigma^2\{1 + c(t)e^{-(y-\alpha)/\beta}\}$$

where $c(t) = e^{\rho e^{-\zeta t}}$, $h = (2\rho\zeta\beta^2)/\sigma^2$ and σ^2 and ζ are positive constants, in such a way that the solution of the forward Kolmogorov equation

$$\frac{1}{2}\frac{\partial^2}{\partial y^2}\{a(y,t)f(y,t)\} - \frac{\partial}{\partial y}\{b(y,t)f(y,t)\} = \frac{\partial}{\partial t}f(y,t)$$

gives rise to

$$f(y,t) = \frac{c(t)}{\beta}\frac{e^{-c(t)(y-\alpha)/\beta}}{\{1 + e^{-c(t)(y-\alpha)/\beta}\}^2} \quad, -\infty < y < \infty$$

Accordingly, since $c(t) \to 1$ as t increases, the equilibrium distribution turns out to be

$$f(y) = \frac{1}{\beta}\frac{e^{-(y-\alpha)/\beta}}{\{1 + e^{-(y-\alpha)/\beta}\}^2} \quad, -\infty < y < \infty \qquad (4)$$

i.e. a logistic distribution with location parameter $\alpha = \ln\delta$ and scale parameter $\beta = 1/\theta$, which, in turn, corresponds to a log-logistic distribution of type (3) for the income variable $X = e^Y$.

However, the generalization of this result to the whole kappa(3) family ($\kappa > 0$) did not appear so obvious. In this framework, an effective contribution was provided by L. Pandolfi, a mathematician at the University of Florence, who suggested dropping the time variable on the forward Kolmogorov equation by letting $t \to \infty$, in such a way as to directly find the stable solution, if it existed.

As $t \to \infty$, the forward Kolmogorov equation reduces to

$$\frac{1}{2}\frac{\partial}{\partial y}\{a(y)f(y)\} - b(y)f(y) = \cos t \qquad (5)$$

in such a way that the solution of (5) for

$$b(y)=\lim_{t\to\infty} b(y,t)=-\frac{\sigma^2}{2\beta}\ ,\ a(y)=\lim_{t\to\infty} a(y,t)=\sigma^2\{1+e^{-(y-\alpha)/\beta}\} \tag{6}$$

directly gives (4) as the equilibrium distribution. Then, it was quite easy to find a generalization of the infinitesimal mean and variance of type (6), in such a way that the solution of (5) gave rise to the kappa(3) distribution for $X=e^Y$. Indeed, it is at once apparent that for

$$b(y)=-\frac{\sigma^2}{2\beta}\{1-(1-\kappa)e^{-\kappa(y-\alpha)/\beta}\}\ ,\ a(y)=\sigma^2\{1+\kappa e^{-(y-\alpha)/\beta}\} \tag{7}$$

the solution of (5) gives rise to

$$f(y)=\frac{\kappa}{\beta}\frac{e^{-\kappa(y-\alpha)/\beta}}{\{1+\kappa e^{-\kappa(y-\alpha)/\beta}\}^{1/\kappa+1}}$$

which, in turn, corresponds to the kappa(3) family as the equilibrium distribution for $X=e^Y$.

As to the economic interpretation of these results, they appeared to be quite similar to the conclusions achieved by Ord (1975) for the Champernowne model. Indeed, for large values of *t*, relations (7) allow for an income-dependent, eventually negative drift, for the logarithm of relative increments, while their infinitesimal variance is greatest at the lower end of the income scale. As pointed out by Ord (1975, p. 157) "*This does not seem unreasonable, as proportionate changes are likely to be greater in this region*".

Such research results were discussed with Prof. Camilo Dagum during the Amsterdam ISI Session, starting a long scientific co-operation in the field of income distribution, economic inequality and poverty measures (Dagum *et al.*, 1992; Dagum and Lemmi, 1995a; Dagum and Lemmi, 1995b). Moreover, Camilo Dagum had an authoritative role in the organization of several scientific events at the University of Siena and, in particular the 1991 International Conference on: Personal Income Distribution, Inequality and Poverty and the recent (2005) International Conference for celebrating Max O. Lorenz and Corrado Gini.

Dipartimento di Metodi Quantitativi L. FATTORINI
Università di Siena A. LEMMI

REFERENCES

M.C. BRYSON (1974), *Heavy-tailed distributions: properties and tests*, "Technometrics", 16, pp. 61-68.

D.G. CHAMPERNOWNE (1952), *The graduation of income distribution*, "Econometrica", 20, pp. 591-615.

D.G. CHAMPERNOWNE (1953), *A model for income distribution*, "Economic Journal", 63, pp. 318-351.

C. DAGUM (1977), *A new model of personal income distribution: specification and estimation*, "Economie Appliquée", 30, pp. 413-437.

C. DAGUM, R. GAMBASSI, A. LEMMI (1992), *New approaches in measuring poverty*; in "Proceedings of the International Scientific Conference on Poverty Measurement for Economies in Transition in Eastern European Countries", Polish Statistical Association, Warsaw, pp. 201-225.

C. DAGUM, A. LEMMI eds (1995a), *Income distribution, social welfare, inequality and poverty*. Researches on Income Inequality, Slotjee (ed.), Volume VI , Jai Press, Greenwich, CT (USA).

C. DAGUM, A. LEMMI (1995b), *Income distribution, social welfare, inequality and poverty: Preface*, in "Income Distribution, Social Welfare, Inequality and Poverty", in Dagum, Lemmi (eds), Researches on Income Inequality, D. J. Slotjee (ed.), Volume VI, Jai Press, Greenwich, CT (USA).

L. FATTORINI, A. LEMMI (1979), *Proposta di un modello alternativo per l'analisi della distribuzione personale del reddito*, "Atti delle Giornate di Lavoro Airo", 1979, pp. 89-117 (in Italian).

P.W. MIELKE (1973), *Another family of distributions for describing and analysing precipitation data*, "Journal of Applied Meteorology", 12, pp. 275-280.

P.W. MIELKE, E.S. JOHNSON (1973), *Three-parameter kappa distribution maximum likelihood estimates and likelihood ratio tests*, "Monthly Weather Review", 101, pp. 701-707.

J.K. ORD (1975), *Statistical models for personal income distributions*, in "Statistical Distributions in Scientific Work", vol 2, Patil, Kotz and Ord ed., Reidel, Dordrecht, pp. 151-158.

C. SCALA (1981), *Funzioni di densità di probabilità: atlante descrittivo*, Scuola di Specializzazione in Discipline Bancarie, Università degli Studi di Siena, Siena.

RIASSUNTO

Un'interpretazione stocastica della distribuzione personale del reddito di Dagum

Questa breve nota ricorda gli avvenimenti e gli elementi teorici di base che, circa trenta anni fa, portarono a definire il modello di Dagum a tre parametri per la distribuzione personale del reddito come la distribuzione di equilibrio di un processo stocastico di diffusione. La maggior parte di questo contributo tratta della derivazione di una soluzione stabile per la cosiddetta equazione "in avanti" di Kolmogorov. Tutto ciò ha fondamenti statistico-economici nei contributi di Sylos Labini all'analisi della stratificazione sociale.

SUMMARY

The stochastic interpretation of the Dagum personal income distribution: a tale

The present notes delineates the circumstances and the basic theoretical steps which led, about thirty years ago, to the derivation of the Dagum income distribution as the equilibrium distribution of a diffusion process. Most of theoretical issues of this note deal with the derivation of the stable solution for the forward Kolmogorov equation. Such a derivation has socio-economic fundaments from the contributions of Sylos Labini on social stratification.