Stat Methods Appl (2009) 18:205–220
DOI 10.1007/s10260-007-0091-3

ORIGINAL ARTICLE

Some developments on the log-Dagum distribution

Filippo Domma · Pier Francesco Perri

Accepted: 17 December 2007 / Published online: 9 January 2008

Abstract Skewed and fat-tailed distributions frequently occur in many applications. Models proposed to deal with skewness and kurtosis may be difficult to treat because the density function cannot usually be written in a closed form and the moments might not exist. The log-Dagum distribution is a flexible and simple model obtained by a logarithmic transformation of the Dagum random variable. In this paper, some characteristics of the model are illustrated and the estimation of the parameters is considered. An application is given with the purpose of modeling kurtosis and skewness that mark the financial return distribution.

Keywords Skewness · Kurtosis · ML estimates · Fisher information matrix · Value-at-Risk

Mathematics Subject Classification (2000) 60E05 (Primary) · 62F12 (Secondary)

1 Introduction

The practice of obtaining new random variables from transformations of well-known distributions is commonly employed in Statistics. The underlying idea is to find new distributions which can emphasize or mitigate some aspects of the original model in order to provide a better fitting to data.

F. Domma (✉) · P. F. Perri
Department of Economics and Statistics,
University of Calabria, Via P. Bucci (Cubo 0C),
87036 Arcavacata di Rende (CS), Italy
e-mail: f.domma@unical.it

P. F. Perri
e-mail: pierfrancesco.perri@unical.it

In this paper, the main characteristics of the log-Dagum random variable, obtained by a logarithmic transformation of the Dagum model (Dagum 1977, 1980), are analyzed. The log-Dagum random variable seems to be a flexible parametric family as it takes values on the real line and the shape of the distribution is always leptokurtic. Moreover, it may be both symmetric and asymmetric (positive or negative). Therefore, it may be useful in modeling skewed and leptokurtic distributions which frequently occur in many fields such as hydrology, seismology, telecommunications, Web traffic, economics, insurance and finance. Currently, a wide variety of parametric models have been considered to deal with asymmetric and heavy-tailed distributions. Amongst these, choices are usually made as a matter of analytical and numerical tractability. The log-Dagum distribution does not particularly suffer from analytical and computational limitations and its cumulative distribution function, quantiles, mode and moments are given in a closed form. Moreover, computational aspects concerned with the maximum likelihood (ML) estimates do not involve great complexity. These reasons ought to make the log-Dagum distribution a competitive model for data marked by skewness and kurtosis.

The present paper is organized as follows. Section 2 introduces the genesis and the main features of the log-Dagum model, and points out that the model is leptokurtic and with skewness and kurtosis governed only by one parameter. Section 3 provides some results concerning the ML estimates of the log-Dagum parameters and quantiles. In Sect. 4, a simulation study is carried out in order to appraise the performance of the ML estimates of the model parameters for finite sample size. Section 5 reports our application of the log-Dagum distribution to a financial context to model daily returns and to calculate the value-at-Risk (VaR) of some stocks showing a different degree of skewness and kurtosis. A comparison with the α-Stable distribution is then carried out. In the conclusive part of the paper, some final considerations are made.

2 Genesis and properties of the log-Dagum distribution

In this section, we outline the genesis of the log-Dagum model, moving from some considerations on survival data model building. We also investigate its shape characteristics, showing that the distribution is always leptokurtic.

The Dagum model (Dagum 1977, 1980) has been successfully used in studies on income and wage distribution as well as in those on wealth distribution. Characteristics and properties of the model have been extensively analyzed by several authors. Nevertheless, in their monograph Kleiber and Kotz (2003, p. 215) argued that "...*the hazard rate and the mean excess function of the Dagum distribution have not been investigated in the statistical literature*". As a matter of fact, Domma (2002) studied the Dagum hazard function proving that it is very flexible and suited for describing different situations. Specifically, with a proper combination of the parameters, the hazard function is either always decreasing, or it shows a maximum (upside down bathtub), or both a minimum and a maximum (bathtub and then upside down bathtub). This nice characteristic ought to make the Dagum distribution particularly suitable for modeling survival data. On this issue, we observe that it is a very common practice in survival analysis to consider the observed heterogeneity by means the so-called accelerated failure time model (AFTM), in which a logarithmic transformation of survival time is adopted. The possibility of using an AFTM with Dagum-distributed survival time, requires a preliminary study of the main properties and characteristics of the random variable that we call log-Dagum. Therefore, in this work, focus will not be placed on the study of the log-Dagum model in the context of survival analysis, but rather only on the exploration of its statistical properties.

A positive random variable Y is Dagum-distributed if its cumulative distribution function (CDF) is given by:

$$F_Y(y;\beta,\lambda,\delta)=\left(1+\lambda y^{-\delta}\right)^{-\beta},$$

where $\lambda>0$ is a scale parameter and $\beta>0$, $\delta>0$ are two shape parameters. The logarithmic transformation of Y, $X=\ln Y$, has the following CDF

$$F_X(x;\beta,\lambda,\delta)=F_Y\left(e^x;\beta,\lambda,\delta\right)=\left(1+\lambda e^{-\delta x}\right)^{-\beta}, \tag{1}$$

where, unlike the Dagum model, $x\in\mathbb{R}$, $\beta>0$ is a shape parameter, $\lambda>0$ influences only the location while $\delta>0$ is a scale parameter. The probability density function (PDF) is:

$$f_X(x;\beta,\lambda,\delta)=\beta\lambda\delta e^{-\delta x}\left(1+\lambda e^{-\delta x}\right)^{-\beta-1}.$$

Throughout this paper, the log-Dagum model will be denoted with $LDa(\beta,\lambda,\delta)$.

We observe that the expression of $F_X(x;\beta,\lambda,\delta)$ may also be obtained in different ways. For instance, the log-Dagum density may be considered as a reparameterization of the type I generalized logistic distribution (see, e.g., Johnson et al. 1995, p. 140) and, consequently, a special case of the exponential generalized beta distribution of second type (McDonald and Xu 1995).

From (1) it is easy to verify that the mode of the distribution always exists, and that it is unique and given by $m=\delta^{-1}\ln(\lambda\beta)$. Instead, we remember that the Dagum distribution is unimodal for $\beta\delta>1$ and zeromodal for $\beta\delta\leq 1$.

Solving with respect to x the equation $F_X(x;\beta,\lambda,\delta)=p$, with $p\in(0,1)$, we obtain the simple and closed expression for the pth quantile of $LDa(\beta,\lambda,\delta)$:

$$x_p=\frac{1}{\delta}\ln\left(\frac{\lambda}{p^{-1/\beta}-1}\right). \tag{2}$$

As for the shape of $LDa(\beta,\lambda,\delta)$, Domma (2001) showed that β is a direct indicator of asymmetry. In particular, unlike the Dagum model which shows only positive skewness, $LDa(\beta,\lambda,\delta)$ is negative skewed for $\beta\in(0,1)$; for $\beta>1$, it is positive skewed and symmetric for $\beta=1$. Moreover, in Domma (2004) it is also shown that the kurtosis of the distribution is influenced only by β. As follows, through the moments of the distribution, we will prove that $LDa(\beta,\lambda,\delta)$ is always leptokurtic, whatever the value of β is.

The moment generating function of $LDa(\beta, \lambda, \delta)$ is equal to the moment of order t of the Dagum random variable:

$$m_X(t) = E\left[e^{tX}\right] = E\left[Y^t\right] = \beta\lambda^{\frac{1}{\delta}} B\left(\beta + \frac{t}{\delta}, 1 - \frac{t}{\delta}\right), \quad \delta > t,$$

where $B(.,.)$ is the Beta mathematical function. Nevertheless, in order to calculate the moments of $LDa(\beta, \lambda, \delta)$, it is more convenient to use the cumulant generating function, $\ln[m_X(t)]$, from which the rth cumulant is given by:

$$K_r(t) = \left\{\frac{\partial^r \ln[m_X(t)]}{\partial t^r}\right\}_{t=0}.$$

Using the cumulant generating function, it is possible to prove that, for any $\beta, \lambda, \delta > 0$, the first four moments of $LDa(\beta, \lambda, \delta)$ are:

$$\begin{aligned}
E(X) &= \delta^{-1}[\ln\lambda + \Psi(\beta) - \Psi(1)] \\
E\left(X^2\right) &= \delta^{-2}\left\{\left[\Psi'(\beta) + \Psi'(1)\right] + [\ln\lambda + \Psi(\beta) - \Psi(1)]^2\right\} \\
E\left(X^3\right) &= \delta^{-3}\left\{\left[\Psi''(\beta) - \Psi''(1)\right] + [\ln\lambda + \Psi(\beta) - \Psi(1)]^3\right. \\
&\quad \left. + 3[\ln\lambda + \Psi(\beta) - \Psi(1)]\left[\Psi'(\beta) + \Psi'(1)\right]\right\} \\
E\left(X^4\right) &= \delta^{-4}\left\{\left[\Psi'''(\beta) + \Psi'''(1)\right] + 3\left[\Psi'(\beta) + \Psi'(1)\right]^2\right. \\
&\quad +4[\ln\lambda + \Psi(\beta) - \Psi(1)]\left[\Psi''(\beta) - \Psi''(1)\right] \\
&\quad +6[\ln\lambda + \Psi(\beta) - \Psi(1)]^2\left[\Psi'(\beta) + \Psi'(1)\right] \\
&\quad \left. + [\ln\lambda + \Psi(\beta) - \Psi(1)]^4\right\},
\end{aligned}$$

where $\Psi(.)$, $\Psi'(.)$, $\Psi''(.)$ and $\Psi'''(.)$ are the respective digamma, trigamma, tetragamma and pentagamma functions. Let Z be the standardized log-Dagum distribution. Then, after some algebra, we obtain the standardized third and fourth moments:

$$E\left(Z^3\right) = \frac{\left[\Psi''(\beta) - \Psi''(1)\right]}{\delta_\beta^3} \quad \text{and} \quad E\left(Z^4\right) = \frac{\left[\Psi'''(\beta) + \Psi'''(1)\right]}{\delta_\beta^4} + 3,$$

where $\delta_\beta = \sqrt{\Psi'(\beta) + \Psi'(1)}$.

We highlight that the standardized fourth moment is always greater than 3 since, for any $\beta > 0$, $\Psi'''(\beta) > 0$. Consequently, the log-Dagum distribution turns out to be always leptokurtic and, thus, it is potentially suitable to model data which exhibit fat tails.

3 Maximum likelihood estimates

In this section, we aim at providing the elements to estimate the log-Dagum model parameters using the ML method. We illustrate the elements of the Fisher information matrix which are useful to implement the Fisher-scoring method and to compute the asymptotic variance–covariance matrix of the ML estimators. Moreover, since many applications may find it useful to determine the quantiles of a distribution, we also employ the ML method to estimate $LDa(\beta, \lambda, \delta)$ quantiles.

3.1 Parameter estimation

Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ be a realization of the random sample $\mathbf{X} = (X_1, X_2, \ldots, X_n)$, where $X_1, X_2, \ldots, X_n$ are i.i.d. random variables according to $LDa\,(\beta, \lambda, \delta)$. Then, the log-likelihood function of $\mathbf{x}$ is:

$$\ell(\beta, \lambda, \delta; \mathbf{x}) = n \ln(\beta\lambda\delta) - \delta \sum_{i=1}^{n} x_i - (\beta + 1) \sum_{i=1}^{n} \ln\left(1 + \lambda e^{-\delta x_i}\right).$$

Maximization of $\ell(\beta, \lambda, \delta; \mathbf{x})$ does not admit any explicit solution. Therefore, the ML estimates $\hat{\boldsymbol{\theta}}_n = (\hat{\beta}_n, \hat{\lambda}_n, \hat{\delta}_n)$ can be obtained only by means of numerical procedures such as the Fisher-scoring method.

Under the usual regularity conditions, the well-known asymptotic properties of the ML method ensure that $\sqrt{n}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right) \xrightarrow{d} N\,(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$, where $\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = [\mathbf{I}\,(\boldsymbol{\theta})]^{-1}$ is the asymptotic variance–covariance matrix and $\mathbf{I}\,(\boldsymbol{\theta})$ is the Fisher information matrix in a single observation, whose elements are:

$$i_{\beta\beta} = \frac{1}{\beta^2}, \quad i_{\beta\lambda} = \frac{1}{\lambda(\beta+1)}, \quad i_{\beta\delta} = -\frac{\ln\lambda + \Psi(\beta) - \Psi(2)}{\delta(\beta+1)}$$

$$i_{\lambda\lambda} = \frac{\beta}{\lambda^2(\beta+2)}, \quad i_{\lambda\delta} = -\beta\frac{\ln\lambda + \Psi(\beta+1) - \Psi(2)}{\lambda\delta(\beta+2)}$$

$$i_{\delta\delta} = \frac{1}{\delta^2}\left(1 + A_{1,\delta} + A_{2,\delta}\right),$$

where

$$\begin{aligned} A_{1,\delta} &= \frac{\beta}{\beta+2}\Big\{[\ln\lambda + \Psi(\beta+1) - \Psi(2)]^2 + \Psi'(\beta+1) - 2\Psi'(\beta+3) + \Psi'(2) \\ &\quad + 2\,[\Psi(\beta+1) - \Psi(\beta+3)]\,[\Psi(2) - \Psi(\beta+3)]\Big\} \\ A_{2,\delta} &= 2\beta(\beta+1) \\ &\quad \times \left[\frac{\Psi(\beta+3) - \Psi(1)}{(\beta+2)^2} - \frac{\Psi'(\beta+3)}{\beta+2} - \frac{\Psi(\beta+2) - \Psi(1)}{(\beta+1)^2} + \frac{\Psi'(\beta+2)}{\beta+1}\right]. \end{aligned}$$

The mathematical developments, needed to derive these elements, are given in the Appendix.

Since the log-Dagum model can alternatively be obtained through a proper reparameterization of the type I generalized logistic distribution, it is worth pointing out some issues, regarding the ML estimates, which have been debated in the literature. Focusing on the generalized logistic distribution, Zelterman (1987) proved that the likelihood function becomes unbounded at some points on the edge of the parameter space and, consequently, the global ML estimator does not exist. Despite the fact that there are no ML estimators, the information matrix is non-singular and may be used to analyze the behavior of asymptotically efficient estimators. Fortunately, Abberger and Heiler (2000) and Abberger (2002) derived local maximum likelihood estimators and showed that likelihood equations have a root which is consistent and asymptotically normally distributed. The problem has been further investigated by Shao (2002).

3.2 Quantile estimation

In several situations, like those involving finance and economic studies, it may be useful to determine the pth quantile of a distribution. We have seen that for $LDa(\beta, \lambda, \delta)$ the pth quantile is given by (2). Now, if the population parameters are unknown, the *invariance principle* (Zehna 1966), provides us with the ML estimator of x_p given by:

$$\hat{x}_p = \frac{1}{\hat{\delta}_n} \ln\left(\frac{\hat{\lambda}_n}{p^{-1/\hat{\beta}_n} - 1}\right), \tag{3}$$

where $\hat{\beta}_n$, $\hat{\lambda}_n$, $\hat{\delta}_n$ are the ML estimators of β, λ and δ based on a sample of size n. Moreover, under the usual regularity conditions, the well-known asymptotic properties of the ML method ensure that:

$$\sqrt{n}\left(\hat{x}_p - x_p\right) \xrightarrow{d} N\left(0, \sigma_{\boldsymbol{\theta},p}\right), \tag{4}$$

where $\sigma_{\boldsymbol{\theta},p} = \mathbf{g}\left[\mathbf{I}\left(\boldsymbol{\theta}\right)\right]^{-1} \mathbf{g}'$ and

$$\begin{aligned}\mathbf{g} &= \left[\frac{\partial x_p}{\partial \beta}, \frac{\partial x_p}{\partial \lambda}, \frac{\partial x_p}{\partial \delta}\right] \\ &= \left[\frac{-p^{-1/\beta}\ln p}{\delta\beta^2\left(p^{-1/\beta} - 1\right)}, \frac{1}{\lambda\delta}, -\frac{1}{\delta^2}\ln\left(\frac{\lambda}{p^{-1/\beta} - 1}\right)\right].\end{aligned}$$

Even though the asymptotic variance in (4) is unknown, it can be estimated through the ML estimator of $\boldsymbol{\theta}$. In this case, the asymptotic result defined in (4) is still valid and

$$\sqrt{n}\left(\hat{x}_p - x_p\right) \xrightarrow{d} N\left(0, \hat{\sigma}_{\hat{\boldsymbol{\theta}}_n,p}\right).$$

This result makes the construction of the $100 \times (1-\alpha)\%$ confidence interval for the quantile x_p particularly easy. It is given by:

$$CI(x_p) = \left[\hat{x}_p - z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{\sigma}_{\hat{\boldsymbol{\theta}}_n,p}}{n}}, \quad \hat{x}_p + z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{\sigma}_{\hat{\boldsymbol{\theta}}_n,p}}{n}} \right],$$

where z_γ denotes the $(1-\gamma)$th quantile of the standard Normal distribution.

Alternative asymptotic results for $LDa(\beta, \lambda, \delta)$ quantiles may be provided by the theory of order statistics. For $p \in (0, 1)$, let be $r = [np] + 1$, where $[np]$ is the integer part of np. Then, $X_{r:n}$ represents the pth sample quantile. Since the log-Dagum CDF is absolutely continuous with PDF which is positive and continuous in the pth quantile, then, as $n \to \infty$,

$$\sqrt{n}\left(X_{r:n} - x_p\right) \xrightarrow{d} N\left(0, \xi_{\boldsymbol{\theta},p}\right)$$

with $\xi_{\boldsymbol{\theta},p} = \dfrac{p(1-p)}{f_X^2\left(x_p; \beta, \lambda, \delta\right)}$.

This result can be used to obtain approximate confidence intervals for x_p either the form of $f_X(x; \beta, \lambda, \delta)$ is completely specified around x_p or a good estimator of $f_X\left(x_p; \beta, \lambda, \delta\right)$ is available (Arnold et al. 1992, p. 224).

4 Simulation study

In this section, we intend to investigate the behavior of the ML estimates for LDa (β, λ, δ) parameters in a finite sample size context. For this purpose, we carry out a simulation study based on different log-Dagum distributions obtained by varying the parameter values. For each combination of parameters supposedly known, k random samples of n i.i.d. observations are independently generated, and the ML estimates are computed for each sample. The simulation can be summarized in the following steps:

1. choose the values, β_0, λ_0 and δ_0, for $LDa(\beta, \lambda, \delta)$ parameters;
2. choose the sample size n;
3. compute the elements of the asymptotic variance–covariance matrix, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = [\mathbf{I}(\boldsymbol{\theta})]^{-1}$, $\boldsymbol{\theta}=\boldsymbol{\theta}_0 = (\beta_0, \lambda_0, \delta_0)$;
4. draw k random samples of size n from $LDa(\beta_0, \lambda_0, \delta_0)$;
5. compute the ML estimates $\hat{\boldsymbol{\theta}}_n = (\hat{\beta}_n, \hat{\lambda}_n, \hat{\delta}_n)$ of $\boldsymbol{\theta}_0$ for each of the k samples. Furthermore, provide a consistent estimate of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ by means of $\left[\mathbf{I}(\hat{\boldsymbol{\theta}}_n)\right]^{-1}$;
6. for each of the ML estimates obtained at the previous step, compute the empirical bias and empirical mean square error (MSE) as:

$$b(\hat{\theta}_n) = \mu(\hat{\theta}_n) - \theta_0 \quad \text{and} \quad MSE(\hat{\theta}_n) = k^{-1} \sum_{i=1}^{k} \left(\hat{\theta}_n^{(i)} - \theta_0\right)^2,$$

where $\mu(\hat{\theta}_n) = k^{-1} \sum_{i=1}^{k} \hat{\theta}_n^{(i)}$, $\hat{\theta}_n = \hat{\beta}_n, \hat{\lambda}_n, \hat{\delta}_n$.

As far as the first step is concerned, we observe that the model parameters ought to be chosen in such a way as to provide a wide spectrum of different situations that might occur with real data. For the purpose of saving space, we have selected in the analysis the values $\beta_0 = 0.5, 1, 2.5$; $\delta_0 = 70, 140, 200$ and $\lambda_0 = 1$. It is worth pointing out that we have also considered many other different values for the parameters. Even so, the results are not significantly different from those obtained with the aforementioned values. The parameter values have been selected on the basis of a preliminary study in which we have fitted the log-Dagum model to real data from different fields of applications. The values, considered for the simulation, fall in the range of the ML estimates obtained for the considered data set. As for the number of replications, we have fixed $k = 5{,}000$. Finally, we have independently selected samples with size 250 and 1,000. Having specified the population, the number of samples to be drawn and the sample size, we have computed the ML estimates for the population parameters by means of the Fisher-scoring procedure.

Results for each combination of the parameter values and for each sample size are shown in Table 1. In general, we observe that both the bias and the MSE of the ML estimates are quite small and decrease as the sample size increases. Moreover, the contribution of bias to the empirical MSE is negligible. Estimates provided by $\hat{\delta}_n$ are very close to the corresponding population parameter, δ_0, even if the sample size is small. On the contrary, only for $\beta = 2.5$ and $n = 250$ the bias and the MSE seem to be higher than the ones obtained for the other situations. Moreover, the bias turns out to be positive for the parameter β. Conclusions are still valid if we consider the estimates of the asymptotic variance. Therefore, we can state that the ML method provides efficient and unbiased estimates for the log-Dagum parameters for finite sample size.

We have presently restricted our brief analysis only to the ML parameter estimates in order to keep the length of the paper to a minimum. Indeed, a more extensive simulation study has been carried out to additionally investigate the behavior of the estimators for $LDa(\beta, \lambda, \delta)$ quantiles. Briefly, we have found that estimates based on order statistics approach are less efficient than those provided by the ML method. Further details of an exhaustive simulation are shown in Domma and Perri (2005).

5 An application to VaR estimation

In the field of financial applications, the need for modeling skewed and heavy-tailed distributions has become particularly relevant. In fact, the increasing growth of trading activities and frequent trading loss of well-known financial institutions have led financial regulators to favor quantitative techniques which evaluate the possible loss that can be faced by financial traders or institutions. Focusing on the *market risk*, a well-known measure to quantify and control risk is the VaR, defined as the maximal loss that a financial position can incur with a given probability over a given period. From a statistical viewpoint it can be simply interpreted as the quantile of the financial return distribution (profit/loss distribution).

Table 1 Empirical bias and empirical MSE (in brackets) of the ML estimators for $LDa(\beta, \lambda, \delta)$ parameters

β	δ	n	$\hat{\beta}_n$	$\hat{\lambda}_n$	$\hat{\delta}_n$	β	δ	n	$\hat{\beta}_n$	$\hat{\lambda}_n$	$\hat{\delta}_n$
0.5	70	250	0.0080	0.0095	2.24E-09			1,000	0.0050	0.0034	−1.12E-11
			(0.0025)	(0.0474)	(3.88E-15)				(0.0041)	(0.0124)	(9.83E-19)
		1,000	0.0019	0.0003	−4.51E-11	1	200	250	0.0254	0.0047	1.44E-11
			(0.0006)	(0.0114)	(5.74E-16)				(0.0192)	(0.0508)	(7.46E-19)
0.5	140	250	0.0076	0.0072	1.66E-10			1,000	0.0071	−0.0021	2.81E-12
			(0.0024)	(0.0472)	(5.55E-17)				(0.0041)	(0.0116)	(1.10E-19)
		1,000	0.0015	0.0018	−5.88E-12	2.5	70	250	0.2115	0.0192	5.38E-11
			(0.0006)	(0.0113)	(8.40E-18)				(0.5990)	(0.2616)	(1.71E-17)
0.5	200	250	0.0069	0.0085	4.11E-11			1,000	0.0463	−0.0023	−2.03E-11
			(0.0024)	(0.0460)	(5.85E-18)				(0.0883)	(0.0217)	(9.90E-19)
		1,000	0.0018	0.0018	−6.76E-12	2.5	140	250	0.2298	0.0182	1.94E-11
			(0.0006)	(0.0113)	(8.89E-19)				(0.6548)	(0.3079)	(1.98E-19)
1	70	250	0.0277	0.0003	2.28E-10			1,000	0.0508	−0.0033	−5.40E-13
			(0.0199)	(0.0500)	(4.90E-16)				(0.0897)	(0.0226)	(1.39E-20)
		1,000	0.0061	0.0006	1.80E-11	2.5	200	250	0.2201	0.0149	3.44E-12
			(0.0043)	(0.0125)	(6.90E-17)				(0.6563)	(0.2297)	(2.23E-20)
1	140	250	0.0254	0.0021	−4.88E-11			1,000	0.0414	0.0008	−7.67E-13
			(0.0194)	(0.0504)	(6.76E-18)				(0.0891)	(0.0225)	(1.62E-21)

Let P_t denote the value of a financial asset (a stock, an exchange rate or market index) at the time index t, $t = 1, 2, \ldots, T$. Then, the (log)return between dates t and $t + k$ is defined as $R_t(k) = \ln(P_{t+k}/P_t)$.

In this section we intend to estimate VaR for a *long* and *short* financial trading position by using the log-Dagum distribution. Suppose that at the time t we are interested in the risk of a *long* financial position. Then, for a fixed probability level p, VaR is defined as the quantity which satisfies $p = Pr[R_t(k) \leq VaR]$. Since the holder of a *long* position suffers a loss when $R_t(k) < 0$, VaR typically assumes a negative value when p is small. The opposite situation occurs for traders with a *short* position. In this case, the holder suffers a loss when the value of the asset increases (i.e. $R_t(k) > 0$) because he/she would have to buy the asset at a higher price than the one he/she obtained when selling it. Therefore, for a *short* position, VaR at level p is defined through $p = Pr[R_t(k) \geq VaR]$ and, for small p, it typically assumes a positive value. The previous definitions show that VaR is concerned with the tail behavior of the return distribution. For a *long* position, the lower (left) tail is important, while for a *short* position, the upper (right) tail has to be investigated. As a matter of fact, the return distribution is unknown even though many empirical studies showed that it diverges sensibly from the Normal one, being asymmetric and with fat tails (see, e.g., Cont 2001).

In the previous sections, we have described some of the main features concerning skewness, kurtosis and quantiles of $LDa(\beta, \lambda, \delta)$. Now, an empirical analysis is car-

Table 2 Descriptive statistics for the daily returns

	Generali	Unipol	Telecom	Tiscali
Minimum	−0.0845	−0.0150	−0.1692	−0.1787
Maximum	0.0884	0.0921	0.1411	0.3064
Mean	−0.0005	−0.0001	−0.0006	−0.0008
Standard deviation	0.0192	0.0133	0.0280	0.0386
Skewness	−0.0351	−0.1043	−0.2324	1.2079
Kurtosis	5.0013	18.331	5.9827	11.005
JP test p-value	0.0000	0.0000	0.0000	0.0000

ried out to explore the capability of the model in predicting *long* and *short* VaR for daily trading positions. The study is based on the daily returns of four Italian stocks: Generali (from 8 January 1999 to 7 January 2004), Unipol (from 18 June 1998 to 19 May 2005), Telecom (from 8 January 1999 to 7 January 2004) and Tiscali (from 27 October 1999 to 19 May 2005). Descriptive statistics, given in Table 2, point out that the returns are negatively or positively skewed with a rather high level of kurtosis for Tiscali and Unipol stocks. The assumption of normality to describe the data seems to be inadequate as is also confirmed by the Jarque–Bera test (JB test).

In order to capture the asymmetry and tail heaviness, we have fitted the log-Dagum model to the observed returns and we have evaluated model performance, first by using the Akaike's Information Criterion, and then by comparing VaR estimates with those provided under Normal and α-Stable distributions. The use of the Normal distribution is mainly motivated by the need to introduce a comparative term as a limit case. In so doing, the way competitive models can describe departures from normality can be better appreciated. The α-Stable distribution is widely and successfully used in finance (see, e.g., McCulloch 1996; Rachev and Mittnik 2000). There are several reasons for the popularity of this distribution. First, the family of the α-Stable laws is fairly flexible, given that it is characterized by four parameters. It represents a generalization of the Normal distribution and allows both skewness and tail heaviness regulation. In fact, in addition to a location and scale parameter, the distribution is characterized by a shape parameter or tail index, $\alpha \in (0, 2]$, and a skewness parameter, $\beta \in [-1, 1]$. If $\alpha = 2$, the distribution coincides with the Normal one. A particular attractive aspect of α-Stable models is that they allow us to generalize Normal-based concepts and theories, providing a more general framework for modeling financial data. Generalization is possible because the α-Stable models show specific and unique probabilistic properties, namely, *domains of attraction*, *stability property*, *Central Limit Theorem* and *invariance principle* (see, e.g., Rachev and Mittnik 2000).

In our analysis, we have first computed the ML estimates of the log-Dagum parameters by fitting the model to the selected stock daily returns. The results are given in Table 3, together with some indices computed for the estimated models. Subsequently, considering both the lower and upper tail of $LDa(\beta, \lambda, \delta)$, and assuming that p takes value in the interval $[0.001, 0.1]$, we have estimated *long* and *short* VaR

Table 3 ML estimates for $LDa(\beta, \lambda, \delta)$

	Generali	Unipol	Telecom	Tiscali
$\hat{\beta}_n$	0.9522	0.9242	0.9083	1.3439
$\hat{\lambda}_n$	1.0225	1.0576	1.1162	0.5917
$\hat{\delta}_n$	99.586	170.60	69.155	48.050
Mean	−0.0006	−0.0005	−0.0008	−0.0014
Standard deviation	0.0110	0.0066	0.0166	0.0148
Skewness	−0.0541	−0.0879	−0.1078	0.2894
Kurtosis	4.2305	4.2530	4.2674	4.1568

Table 4 AIC for the log-Dagum, α-Stable and Normal distributions

	Generali	Unipol	Telecom	Tiscali
log-Dagum	**−3241.6**	−5360.3	**−2765.1**	−2719.1
α-Stable	−3230.5	**−5505.1**	−2760.3	**−2756.4**
Normal	−3198.9	−5092.9	−2722.3	−2581.9

using the expression defined in (3). We have used the STABLE program, downloadable at http://academic2.american.edu/~jpnolan, to estimate the α-Stable quantiles.

In order to evaluate the overall goodness-of-fit of the log-Dagum, α-Stable and Normal distributions, in Table 4 we have provided the Akaike's Information Criterion value, $AIC = -2\ell(\hat{\boldsymbol{\theta}}_n) + 2k$, where $\ell(\hat{\boldsymbol{\theta}}_n)$ is the maximum log-likelihood value and k the number of parameters to be estimated. We note that the log-Dagum model performs better than the α-Stable for Generali and Telecom return series which are marked by a moderate kurtosis. This outcome seems to be quite consistent with the conclusions that can be drawn by comparing empirical and estimation results shown in Tables 2 and 3, respectively.

Additionally, the performance of the three distributions in estimating VaR has been assessed by carrying out a simple *backtesting* procedure based on the in-sample computation of the *failure rate*, $\hat{p}$, i.e. the proportion of times returns exceed (in absolute value) the predicted VaR. The results for *long* and *short* VaR computation are summarized in Fig. 1, which shows the behavior of $\hat{p}$ under the log-Dagum, Normal and α-Stable distributions.

For a model which perfectly fits the tails of the return empirical distribution, the points $(p, \hat{p})$ should lie on a straight line because we expect the proportion of observations falling further on the VaR to be equal to p. The models we have fitted to the data show departures from the straight line: the failure rate may be greater or smaller than p. Values of $\hat{p}$ greater than p indicate that the VaR underestimates the actual loss risk, while for $\hat{p} < p$, the model appears conservative, leading to an overestimation of the risk.

In general, we observe a poor fitting for the Normal distribution. The log-Dagum and α-Stable distributions better fit the data since their plots are closer to the straight line.

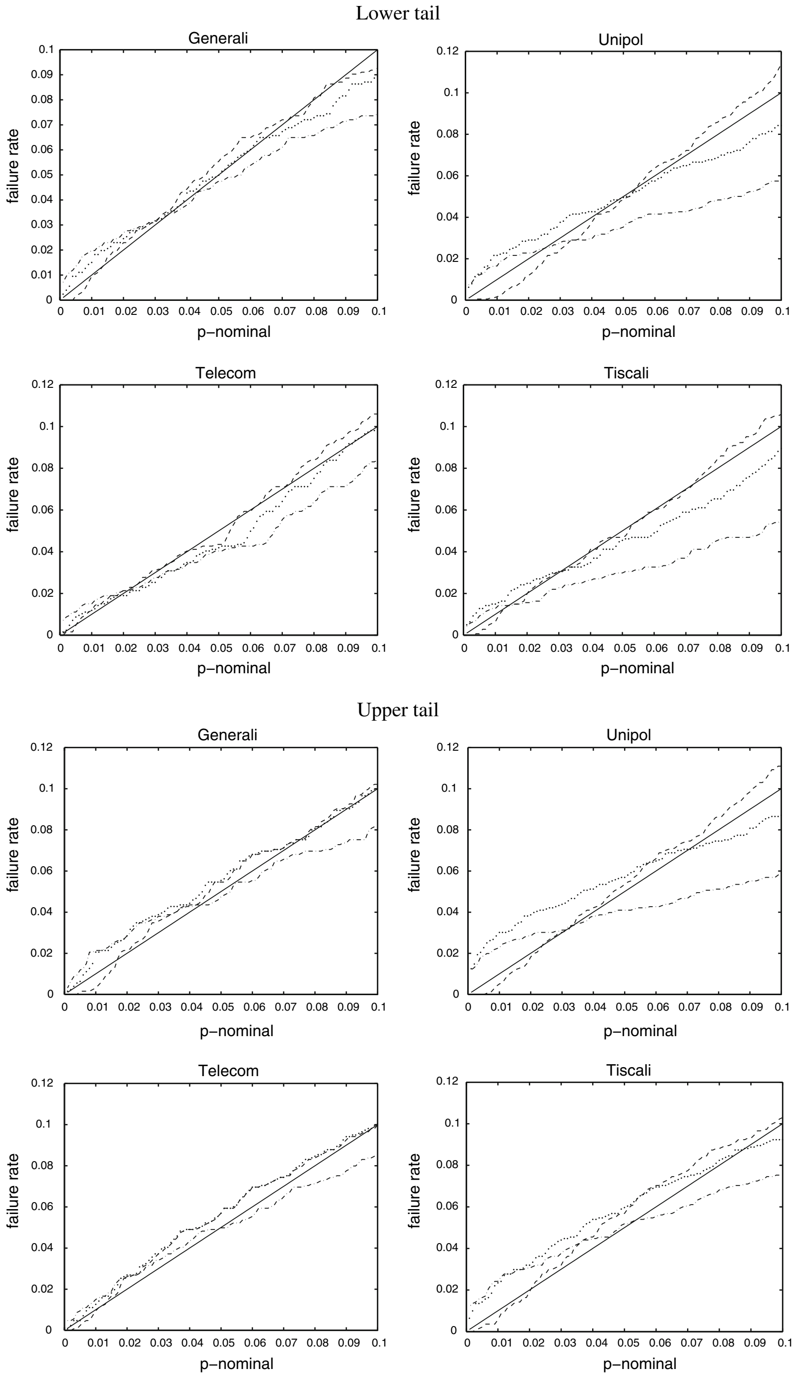

Fig. 1 Behavior of the failure rate for *long* and *short* VaR under the log-Dagum (.), Normal (−.) and α-Stable (− −) distributions

Moreover, the two competitive models seem to have a similar behavior for Generali and Telecom, which are series marked by a moderate degree of kurtosis. In particular, we observe that, for upper tail observations, the plots of the log-Dagum and α-Stable distributions are indistinguishable. The α-Stable distribution performs better than the log-Dagum for Unipol and Tiscali lower tail (VaR prediction for a long position) and, except for values of p nearly falling in the range (0.055, 0.08), also for Unipol upper tail (VaR prediction for a short position). On the other hand, the log-Dagum model should be preferred for Tiscali upper tail when $p \in (0.06, 0.1)$. However, values of p which make the log-Dagum model suitable for VaR estimation are evident also for Generali and Telecom lower tails.

6 Concluding remarks

In this paper, we have described some shape characteristics of the log-Dagum model and analyzed the performance of the ML estimates. In order to investigate the capability of the model in describing data in which skewness and fat tails occur, an application to VaR estimation has been carried out by considering four Italian stocks. VaR, estimated under the log-Dagum model, has been compared with that obtained by using the Normal and α-Stable distributions. If compared with the α-Stable distribution, the log-Dagum model seems to provide a good representation of data which show a moderate kurtosis. To better appreciate the potentiality of the distribution, it is worth underlining that the α-Stable distribution requires quite complex numerical procedures to estimate parameters and quantiles. Furthermore, except in few special cases, the density function cannot be expressed in a closed form and the moments may not exist. In spite of its simplicity, the log-Dagum model allows us to overcome these possible drawbacks.

In addition to employing the log-Dagum distribution to model data which exhibit skewness and kurtosis, we think that the preliminary study, performed in this paper, has been necessary in the light of a possible use of the distribution for survival data to build non-Gaussian regression models. We believe that the investigation of this topic may certainly be of interest for future researches.

Acknowledgments The authors wish to thank the anonymous Associate Editor and the referee for their useful comments in improving the original version of the paper.

Appendix

In order to derive the elements of the Fisher information matrix, we preliminarily obtain, for $q > 0$, the following expressions:

$$I_1(q+1) = \int_0^1 z^q (1-z) \ln z dz = B(q+1, 2) \left[\Psi(q+1) - \Psi(q+3)\right]$$

$$I_2(q+1) = \int_0^1 z^q (1-z)\ln(1-z)dz = B(2, q+1)\,[\Psi(2) - \Psi(q+3)]$$

$$I_3(q+1) = \int_0^1 z^q (1-z)\,(\ln z)^2\, dz$$
$$= B(q+1, 2)\left\{[\Psi(q+1) - \Psi(q+3)]^2 + \Psi'(q+1) - \Psi'(q+3)\right\}$$

$$I_4(q+1) = \int_0^1 z^q (1-z)\,[\ln(1-z)]^2\, dz$$
$$= B(2, q+1)\left\{[\Psi(2) - \Psi(q+3)]^2 + \Psi'(2) - \Psi'(q+3)\right\}$$

$$I_5(q+1) = \int_0^1 z^q (1-z)\ln z\ \ln(1-z)dz = I(q+1) - I(q+2)$$

where

$$I(q+1) = \int_0^1 z^q \ln z \ln(1-z)dz = \frac{1}{q+1}\left\{\frac{1}{q+1}\,[\Psi(q+2) - \Psi(1)] - \Psi'(q+2)\right\}.$$

Considering the following expectation

$$E_{j,h,k} = E\left[\frac{X^j e^{-h\delta X}}{\left(1+\lambda e^{-\delta X}\right)^k}\right] = \beta\lambda\delta \int_{-\infty}^{+\infty} x^j e^{-\delta(h+1)x}\left[1+\lambda e^{-\delta x}\right]^{-\beta-1-k} dx$$

and putting $w = e^x$, we have:

$$E_{j,h,k} = \beta\lambda\delta \int_0^{+\infty} (\ln w)^j\, w^{-\delta(h+1)-1}\left[1+\lambda w^{-\delta}\right]^{-\beta-1-k} dw.$$

Moreover, setting $z = (1+\lambda w^{-\delta})^{-1}$, after some algebra, we obtain:

$$E_{j,h,k} = \frac{\beta}{\lambda^h \delta^j}\int_0^1 [\ln\lambda + \ln z - \ln(1-z)]^j\, z^{\beta+k-h-1}(1-z)^h dw.$$

The elements of the Fisher information matrix are functions of the expectations $E_{0,1,1}$, $E_{0,2,2}$, $E_{1,1,1}$, $E_{1,1,2}$ and $E_{2,1,2}$, which are particular cases of $E_{j,h,k}$. These expectations are given by:

$$E_{0,1,1} = \frac{1}{\lambda(\beta+1)}$$

$$E_{0,2,2} = \frac{2}{\lambda^2(\beta+1)(\beta+2)}$$

$$E_{1,1,1} = \frac{\beta}{\lambda\delta}\{B(\beta,2)\ln\lambda + I_1(\beta) - I_2(\beta)\} = \frac{\ln\lambda + \Psi(\beta) - \Psi(2)}{\lambda\delta(\beta+1)}$$

$$E_{1,1,2} = \frac{\beta}{\lambda\delta}[B(\beta+1,2)\ln\lambda + I_1(\beta+1) - I_2(\beta+1)]$$

$$= \beta\frac{\ln\lambda + \Psi(\beta+1) - \Psi(2)}{\lambda\delta(\beta+1)(\beta+2)}$$

$$E_{2,1,2} = \frac{\beta}{\lambda\delta^2}\Big\{(\ln\lambda)^2 B(\beta+1,2) + I_3(\beta+1) + I_4(\beta+1)$$

$$+ 2I_1(\beta+1)\ln\lambda - 2I_2(\beta+1)\ln\lambda - 2I_5(\beta+1)\Big\}$$

$$= \beta\frac{[\ln\lambda + \Psi(\beta+1) - \Psi(2)]^2 + 2[\Psi(\beta+1) - \Psi(\beta+3)][\Psi(2) - \Psi(\beta+3)]}{\lambda\delta^2(\beta+1)(\beta+2)}$$

$$+\frac{2\beta}{\lambda\delta^2}\left\{\frac{[\Psi(\beta+3) - \Psi(1)]}{(\beta+2)^2} - \frac{\Psi'(\beta+3)}{\beta+2} - \frac{[\Psi(\beta+2) - \Psi(1)]}{(\beta+1)^2} + \frac{\Psi'(\beta+2)}{\beta+1}\right\}.$$

Finally, the elements of the Fisher information matrix in a single observation are:

$$i_{\beta\beta} = -E\left[\frac{\partial^2\ell(\beta,\lambda,\delta;\mathbf{X})}{\partial\beta^2}\right] = \frac{1}{\beta^2}$$

$$i_{\lambda\lambda} = -E\left[\frac{\partial^2\ell(\beta,\lambda,\delta;\mathbf{X})}{\partial\lambda^2}\right] = \frac{1}{\lambda^2} - (\beta+1)E_{0,2,2}$$

$$i_{\delta\delta} = -E\left[\frac{\partial^2\ell(\beta,\lambda,\delta;\mathbf{X})}{\partial\delta^2}\right] = \frac{1}{\delta^2} + \lambda(\beta+1)E_{2,1,2}$$

$$i_{\beta\lambda} = -E\left[\frac{\partial^2\ell(\beta,\lambda,\delta;\mathbf{X})}{\partial\beta\partial\lambda}\right] = E_{0,1,1}$$

$$i_{\beta\delta} = -E\left[\frac{\partial^2\ell(\beta,\lambda,\delta;\mathbf{X})}{\partial\beta\partial\delta}\right] = -\lambda E_{1,1,1}$$

$$i_{\lambda\delta} = -E\left[\frac{\partial^2\ell(\beta,\lambda,\delta;\mathbf{X})}{\partial\lambda\partial\delta}\right] = -(\beta+1)E_{1,1,2}.$$

References

Abberger K (2002) ML-estimation in the location-scale-shape model of the generalized logistic distribution, CoFE Discussion Paper, No. 02-15. http://econpapers.repec.org/paper/knzcofedp/0215.htm

Abberger K, Heiler S (2000) Simultaneous estimation of parameters for a generalized distribution and application to time series models. All Stat Arch 84:41–49

Arnold BC, Balakrishnan N, Nagaraja HN (1992) A first course in order statistics. Wiley, New York

Cont R (2001) Empirical properties of asset returns: stylized facts and statistical issues. Quant Finance 1:223–236

Dagum C (1977) A new model for personal income distribution: specification and estimation. Economie Appliquée 30:413–437

Dagum C (1980) The generation and distribution of income, the Lorenz curve and the Gini ratio. Economie Appliquée 33:327–367

Domma F (2001) Asimmetrie puntuali e trasformazioni monotone. Quad Stat 3:145–164

Domma F (2002) L'andamento della hazard function nel modello di Dagum. Quad Stat 4:103–114

Domma F (2004) Kurtosis diagram for the Log-Dagum distribution. Statistica Applicazioni 2:3–23

Domma F, Perri PF (2005) Some developments on Log-Dagum distribution. Discussion Paper Series, no. 44, Dipartimento di Economia e Statistica, Università della Calabria. http://www.ecostat.unical.it

Johnson NL, Kotz S, Balakrishnan N (1995) Continuous univariate distributions, vol 2. Wiley, New York

Kleiber C, Kotz S (2003) Statistical size distribution in economics and actuarial sciences. Wiley, Hoboken

McCulloch JH (1996) Financial applications of stable distribution. In: Maddala GS, Rao CR (eds) Handbook of Statistics, vol 14. Elsevier, North-Holland, pp 393–425

McDonald JB, Xu YJ (1995) A generalization of the beta distribution with applications. J Econom 66:132–152

Rachev S, Mittnik S (2000) Stable paretian models in finance. Wiley, New York

Shao Q (2002) Maximum likelihood estimation for generalized logistic distributions. Commun Stat Theory Methods 31:1687–1700

Zehna PV (1966) Invariance of maximum likelihood estimation. Ann Math Stat 37:744–755

Zelterman D (1987) Parameter estimation in the generalized logistic distribution. Comput Stat Data Anal 5:177–187