EXPLORING THE COPULA APPROACH FOR THE ANALYSIS OF FINANCIAL DURATIONS

Giovanni De Luca[1], Giorgia Rivieccio[1] e Paola Zuccolotto[2]

[1] Dipartimento di Statistica e Matematica per la Ricerca Economica
Università degli Studi di Napoli *Parthenope*
(e-mail: giovanni.deluca@uniparthenope.it, giorgia.rivieccio@uniparthenope.it)

[2] Dipartimento di Metodi Quantitativi
Università degli Studi di Brescia
(e-mail: zuk@unibs.it)

ABSTRACT: The object of the paper is the comparison of two approaches for the analysis of financial durations. The former is the parametric approach popularized by Engle & Russell (1998) and is implemented using the exponential, the Weibull, the Burr and the Pareto density functions. The latter makes use of bivariate and trivariate copula functions.

KEYWORDS: Financial durations, autoregressive conditional durations, copula function

1 Introduction

The seminal work of Engle & Russell (1998) has opened the interest in ultra-high frequency financial data, also known as tick-by-tick data. They are irregularly spaced time series which enable the researcher to investigate the process of transaction of a financial asset traded on a financial market. The Autoregressive Conditional Duration (ACD) model is now a consolidated statistical tool to describe the time between the occurrence of two market events.

In their pioneering work, Engle and Russell made use of exponential and Weibull random variable for the durations. Many authors tried to extend the analysis relying on more elaborated densities, such as the Burr or the Pareto.

In this paper we want to compare the traditional parametric approach with a semi-parametric approach based on the copula function. In particular we will make comparisons using bivariate and trivariate copulas.

The paper is structured as follows. In Section 2 the ACD modelling is presented. Section 3 briefly reviews the concept of copula function. In Section 4 the data analysis is carried out.

2 ACD models

The class of ACD models is aimed at modelling the durations between two markets events, such as price changes or bid-ask spreads. Let X_i be the duration between two observations at times t_{i-1} and t_i. Engle & Russell (1998) proposed the ACD (q,p)

model

$$\begin{array}{rcl} X_i & = & \phi(t_i)\Psi_i\varepsilon_i \\ \Psi_i & = & \omega+\sum_{j=1}^{q}\alpha_j x_{i-j}+\sum_{j=1}^{p}\beta_j\Psi_{i-j}. \end{array} \quad (1)$$

where $\phi(t_i)$ is a deterministic daily seasonal component and $x_i = X_i/\phi(t_i)$ is the seasonally adjusted duration at time t_i.

Assuming ε_i identically and independently distributed with $E(\varepsilon_i) = 1$, it is easy to see that $E(x_i|\mathcal{F}_{i-1}) = \Psi_i$, where $\mathcal{F}_{i-1}$ is the information at time t_{i-1}. As a result, Ψ_i can be interpreted as the i-th expected (deseasonalized) duration conditionally on the information at time t_{i-1}.

Some restrictions on the parameters of (1) are to be encountered.

In order to estimate the model using parametric methods, a distributional assumption on ε_i is needed. The most popular ones are the exponential, the Weibull (Engle & Russell, 1998), the Burr (Grammig & Maurer, 2000) and the Pareto (DeLuca & Zuccolotto, 2003). When $q = p = 1$ the popular ACD(1,1) model is obtained. It is a parsimonious model which adequately fits durations in most cases. The model (1) becomes

$$\begin{array}{rcl} x_i & = & \Psi_i\varepsilon_i \\ \Psi_i & = & \omega+\alpha x_{i-1}+\beta\Psi_{i-1}. \end{array}$$

3 Copula functions

A copula is a multivariate distribution function H of random variables $X_1,\ldots,X_n$ with standard uniform marginal distributions $F_1,\ldots,F_n$ defined on the unit n-cube $[0,1]^n$ with the following properties:

1. the range of the copula $C(u_1,\ldots,u_n)$ is the unit interval $[0,1]$;
2. $C(u_1,\ldots,u_n) = 0$ if any $u_i = 0$ for $i = 1,2,\ldots,n$;
3. $C(u_1,1\ldots,1) = C(1,u_2,\ldots,1) = C(1,1,\ldots,u_n) = u_i$ for all $u_i \in [0,1]$;

where $F_i = P(X_i \leq x_i) = u_i$ is uniform in $[0,1]$ for all $i = 1,2,\ldots,n$.

One of the most important copula based theorem is the Sklar's theorem, which justifies the role of copula as dependence function. Let H be a joint distribution function with marginal distribution functions $F_1,F_2,\ldots,F_n$, then there exists a copula function C, such that:

$$C(F_1,\ldots,F_n) = H(x_1,\ldots,x_n).$$

Therefore, the joint distribution is splitted into two components: the unconditional marginal distributions and the dependence structure given by the copula. The copula tool allows to describe the whole dependence structure that characterizes the relationships among the variables. For this reason, the copula function has recently become a very significant quantitative technique to handle many financial time series analysis characterized by a remarkable temporal dependence.

In literature there are two main families of copulas, the Elliptical copulas which are copula functions of elliptical distributions and the Archimedean ones, based on the definition of a generator function $\Phi : I \in R^+$, continuous, decreasing, convex and such that $\Phi(1) = 0$. For a detailed description, see Nelsen (1999).

Using the relationship between distribution and density function, it is possible to derive the conditional copula

$$C(u_n|u_1,\ldots,u_{n-1}) = \frac{\partial^{n-1}C(u_1,\ldots,u_n)/\partial u_1\ldots\partial u_{n-1}}{\partial^{n-1}C(u_1,\ldots,u_{n-1})/\partial u_1\ldots\partial u_{n-1}},$$

and the copula density $c(u_1,\ldots,u_n)$, the n-th derivative of the copula function with respect to $u_1,\ldots,u_n$, which depends on a parameter α.

The estimation of α can be carried out through a two-steps method without assumptions on the parametric form for the marginals. In the former the empirical distribution function is computed to estimate the marginal distribution functions, in the latter the copula parameter is evaluated via maximum likelihood method.

4 Data analysis

The empirical analysis has focused on duration data of the transactions involving a price change of the Italian stock Comit in the month of February 2000. The total number of obervations is 8221. We removed the daily seasonal component after estimating it using a cubic spline with nodes set at each hour.

For the parametric approach, we estimated by maximum likelihood method the ACD models ($p = q = 1$) under the assumptions of exponential, Weibull, Burr and Pareto distributions. Common features of the models is the average of the residuals $\hat{\varepsilon}_i = x_i/\widehat{\Psi_i}$ strictly close to one and the variance far from unity which confirms the importance of departing from the exponential distribution. On the other hand, the semiparametric approach was carried out using two Archimedean copulas, the Clayton and the Frank, for the dependence between x_i and x_{i-1}, first, and then among x_i, x_{i-1} and x_{i-2}. In order to compare the two approaches we use the density forecast evaluation. The probability integral transforms of the one-step-ahead forecasts of the durations are independently and uniformly distributed in (0,1) if the model is correctly specified. The eight histograms are reported in Figures 1 and 2. In particular the copula functions appear to have a better performance.

References

DELUCA, G, & ZUCCOLOTTO, P. 2003. Finite and infinite mixture for financial durations. *Metron*, **61**, 431–455.

ENGLE, R F, & RUSSELL, J R. 1998. Autoregressive Conditional Duration: a new model for irregularly spaced transaction data. *Econometrica*, **66**, 1127–1162.

GRAMMIG, J, & MAURER, K. 2000. Non-monotonic hazard functions and the autoregressive conditional duration model. *Econometrics Journal*, **3**, 16–38.

NELSEN, R B. 1999. *An Introduction to Copulas, Lectures Notes in Statistics*. New York: Springer.

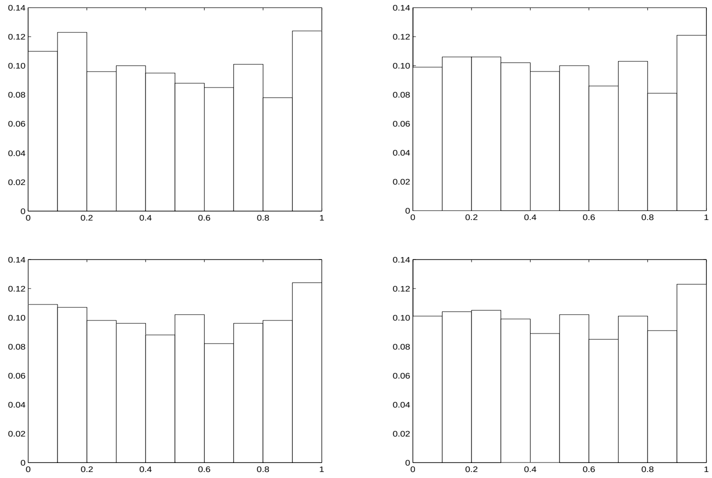

Figura 1. *Histogram of one-step-ahead forecasts for exponential (top-left), Weibull (top-right), Burr (bottom-left) and Pareto (bottom-right).*

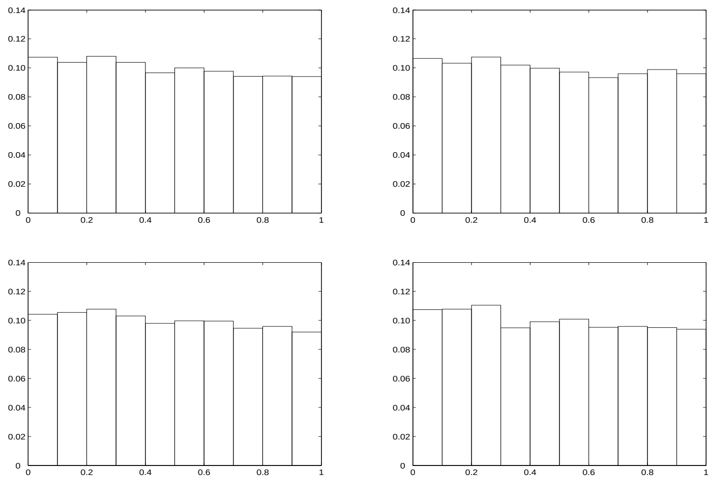

Figura 2. *Histogram of one-step-ahead forecasts for bivariate Clayton copula (top-left), bivariate Frank copula (top-right), trivariate Clayton copula (bottom-left) and trivariate Frank copula (bottom-right).*