Journal of Econometrics 172 (2013) 235–247

Contents lists available at SciVerse ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

The method of simulated quantiles

Yves Dominicy *, David Veredas [1]

Université libre de Bruxelles, Solvay Brussels School of Economics and Management, ECARES. 50, Av Roosevelt CP139, B1050 Brussels, Belgium

ARTICLE INFO

Article history:
Available online 23 August 2012

JEL classification:
C32
G14
E44

Keywords:
Quantiles
Simulation
Matching
Inference

ABSTRACT

We introduce the Method of Simulated Quantiles, or MSQ, an indirect inference method based on quantile matching that is useful for situations where the density function does not have a closed form and/or moments do not exist. Functions of theoretical quantiles, which depend on the parameters of the assumed probability law, are matched with the sample counterparts, which depend on the observations. Since the theoretical quantiles may not be available analytically, the optimization is based on simulations. We illustrate the method with the estimation of α-stable distributions. A thorough Monte Carlo study and an illustration to 22 financial indexes show the usefulness of MSQ.

1. Introduction

Estimation of the parameters of an econometric or economic parametric model is a first order concern. In the case that we know the probability law that governs the random variables, Maximum Likelihood (ML henceforth) is the benchmark technique. If we relax the assumption of knowledge of the distribution but we still have knowledge of the moments, the Generalized Method of Moments (GMM henceforth) becomes the benchmark technique. However, there are models that cannot be easily estimated with ML or GMM: stochastic volatilities, models with stochastic regimes switches, or involving expected utilities to name a few. For those models the likelihood function may not be available analytically (or can be difficult to estimate), and/or the moments may not exist.

To circumvent these estimation difficulties, numerous estimation methods based on simulations have been developed. Gouriéroux and Monfort (1996) and Hajivassiliou and Ruud (1994) introduce Simulated ML (SML), similar to ML except that simulated probabilities are used instead of the exact probabilities. McFadden (1989), Pakes and Pollard (1989), and Duffie and Singleton (1993) independently introduced the Method of Simulated Moments (MSM), which is based on matching sample moments and theoretical moments that are generated by simulations. Gouriéroux et al. (1993) propose Indirect Inference (IndInf), a method based on estimating indirectly the parameters of the model of interest through matching the parameters of an auxiliary model. The Efficient Method of Moments (EMM) of Gallant and Tauchen (1996) is based on the same idea.

In this article we introduce the Method of Simulated Quantiles (MSQ henceforth). Since it is based on quantiles, it is a moment-free method. And since it is based on simulations, we do not need closed form expressions of any function that represents the probability law of the process. Also, due to the robustness of quantiles, MSQ is appropriate when data shows unusually large observations that do not follow the same process as the rest of the observations. In a nutshell, MSQ is based on a vector of functions of quantiles. These functions can be either computed from data (the sample functions) or from the distribution (the theoretical functions). The estimated parameters are those that minimize a quadratic distance between both. Since the theoretical functions of quantiles may not have a closed form expression, we rely on simulations. MSQ is therefore an application of the IndInf principle of Gouriéroux et al. (1993), where the vector of functions of quantiles stands for the vector of auxiliary parameters. Throughout the article, we exemplify the method with the α-stable distribution. Different fonts are used in Sections 1 and 2 to make a clearcut distinction between the theory and the example, which ends with a white square.

Example. Let X_t be a random variable distributed following an α-stable distribution that is represented as $X_t \sim S_\alpha(\sigma, \beta, \mu)$. The parameter $\alpha \in (0, 2]$, often denoted as tail index, measures the thickness of the tails and governs the existence of moments: $E[X_t^p] < \infty, \forall p < \alpha$. Asymmetry is captured by $\beta \in [-1, 1]$. The dispersion parameter $\sigma \in \mathbb{R}^+$ expands or contracts the distribution, and the location parameter $\mu \in \mathbb{R}$ controls the location of the distribution. The α-stable distributions possess the property of sta-

* Corresponding author. Tel.: +32 0 26504502; fax: +32 26504475.
E-mail addresses: yves.dominicy@ulb.ac.be (Y. Dominicy), david.veredas@ulb.ac.be (D. Veredas).
[1] Tel.: +32 0 26504218; fax: +32 0 26504475.

0304-4076/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2012.08.010

bility, which plays an important role in the illustration: Linear combinations of *i.i.d.* α-stable random variables with the same α are also α-stable distributed, i.e. if $X_{i,j} \sim S_\alpha(\sigma_j, \beta_j, \mu_j)$ for $j = 1, \ldots, J$, then $\sum_{j=1}^{J} \omega_j X_{i,j} \sim S_\alpha(\sigma, \beta, \mu)$.

The probability density functions (pdf) of the α-stable distribution does not have a closed form. Since it is a complicated integral, even difficult to evaluate numerically, estimation by ML has been often not considered in applied work (though the theoretical properties of the ML estimator exist, DuMouchel (1973) and the actual estimation has been performed by Nolan (2001)). However, the characteristic function (CF hereafter) has a manageable closed form[2]:

$$E[\exp\{i\theta X_t\}] = \begin{cases} \exp\left\{-\sigma^\alpha|\theta|^\alpha\left(1 - i\beta(\text{sign}\theta)\tan\frac{\pi\alpha}{2}(|\sigma\theta|^{1-\alpha} - 1)\right) + i\mu\theta\right\} & \text{if } \alpha \neq 1 \\ \exp\left\{-\sigma|\theta|\left(1 + i\beta\frac{2}{\pi}(\text{sign}\theta)\ln(\sigma|\theta|)\right) + i\mu\theta\right\} & \text{if } \alpha = 1. \end{cases}$$

All the methods based on the CF match the theoretical and sample counterparts, but in different ways.[3] A problem inherent to these methods is the choice of the grid of frequencies at which to evaluate the CF. While Fielitz and Rozelle (1981) recommend, on the basis of Monte Carlo results, matching only a few frequencies, others, like Feuerverger and McDunnough (1981), recommend using as many frequencies as possible. However, in the latter case, Carrasco and Florens (2002) have shown that, even asymptotically, matching a continuum of moment conditions introduces a fundamental singularity problem.

An alternative is the use of simulation-based methods. From a Bayesian perspective Buckle (1995), Qiou and Ravishanker (1998), and Lombardi (2007) use Monte Carlo Markov Chain methods. From a frequentist perspective, and since random numbers from α-stable distributions can be obtained straightforwardly, IndInf and EMM are appealing, as has been shown by Garcia et al. (2011) and Lombardi and Calzolari (2008). They both use a skewed-t distribution as auxiliary model.

Finally, Fama and Roll (1971) and McCulloch (1986) propose using functions of quantiles. Four specific functions of quantiles are constructed to capture the same features as those captured by α, β, σ and μ. Since the pdf does not have a closed form, so do the cumulative distribution function and the quantiles. Estimation has to be done either by simulation or by tabulation. They opt for the latter. Fama and Roll (1971) and McCulloch (1986) estimate the parameters by calibrating the value of the sample functions of quantiles with tabulated values of the theoretical quantiles. This is a fast way to estimate the parameters, since it avoids optimization, but the theoretical properties remain unclear and the extension to the case of linear combinations of α-stable random variables is not possible (since the tail index has to be the same for all the random variables, estimation has to be done jointly). □

MSQ combines the simulation-based and the quantile-based methods. It is broader than Fama and Roll (1971) and McCulloch (1986), which were designed for the estimation of the α-stable distribution. In fact, MSQ is very general as it applies to any model and distribution. Moreover, it does not make any assumptions on the functional forms for the functions of quantiles. A second advantage is that the method is not based on tabulations but on simulations. This allows a larger flexibility and accuracy. Indeed, tabulation requires interpolation if the sample functions of quantiles are not exactly equal to the tabulated theoretical functions of quantiles. Third, we provide an asymptotic theory that shows the consistency, asymptotic normality and the asymptotic variance–covariance matrix of the estimated parameters.

Estimation via quantiles is a natural alternative to moment-based methods and tracks back to Aitchison and Brown (1957). They estimate a three-parameter log-normal distribution by matching quantiles. A similar result is also found in Bury (1975). Quantiles can also be used to construct functions that measure aspects of the probability distribution. Let q_τ denote the τ-th quantile of X_t for $\tau \in (0, 1)$. The median, $q_{0.50}$, is often used as an estimator of the location. The interquartile range, $q_{0.75}–q_{0.25}$ is a natural measure of dispersion. Bowley (1920) and Hinkley (1975) proposed the quartile skewness (known as the Bowley coefficient):

$$BC = \frac{(q_\tau - q_{0.5}) - (q_{0.5} - q_{1-\tau})}{(q_\tau - q_{1-\tau})}.$$

The smaller τ, the less sensitive to outliers, but the less information from the tails it uses. This measure of asymmetry is dispersion and location invariant, i.e. $\gamma(aX_t + b) = \gamma(X_t)$, where γ denotes the above measure. As far as measures for tail thickness are concerned, Moors (1988) proposed

$$Mo = \frac{(q_{0.875} - q_{0.625}) + (q_{0.375} - q_{0.125})}{(q_{0.750} - q_{0.250})}.$$

The two terms in the numerator are large if little probability mass is concentrated in the neighborhood of the first and third quartile. It is standardized by the interquartile range to guarantee invariance under linear transformations.[4]

The rest of the paper is organized as follows. In Section 2 we first introduce notation followed by MSQ. Each step of the presentation of the method is illustrated with our example. We also show the assumptions and the asymptotic distribution of the estimators. In Section 3 we report the results of a Monte Carlo study based on our example. We consider univariate and multidimensional estimation of α-stable distributions. Multidimensional should not be mistaken with multivariate. By multidimensional we mean joint estimation of univariate distributions that share the same tail index. For the univariate case, our method is compared with McCulloch (1986). In Section 4 we show an illustration to 22 world-wide market indexes, assumed to be distributed according to α-stable distributions. We first estimate the parameters independently. Then we estimate them jointly assuming a common tail index, which is needed for the construction of linear combinations, as we show in the last part of the section. Section 5 concludes. Proofs and other technicalities are relegated to the Appendix.

2. The method of simulated quantiles

Consider a random variable X that follows a distribution $\mathcal{D}(\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ denotes the vector of unknown parameters that are in

2 A different parametrization, among others, is given by

$$E[\exp\{i\theta X_t\}] = \begin{cases} \exp\left\{-\sigma^\alpha|\theta|^\alpha\left(1 - i\beta(\text{sign}\theta)\tan\frac{\pi\alpha}{2}\right) + i\mu\theta\right\} & \text{if } \alpha \neq 1 \\ \exp\left\{-\sigma|\theta|\left(1 + i\beta\frac{2}{\pi}(\text{sign}\theta)\ln|\theta|\right) + i\mu\theta\right\} & \text{if } \alpha = 1. \end{cases}$$

The one showed in the main body of the text is commonly used in numerical computations. For instance, Chambers et al. (1976) use it for simulation.

3 Since the sample CF is a random variable with complex values, one can think about (i) matching moments associated to real and imaginary components respectively (Press, 1972; Fielitz and Rozelle, 1981), (ii) minimizing a distance between the sample and the theoretical CF functions (Paulson et al., 1975; Feuerverger and McDunnough, 1981; Carrasco and Florens, 2002), (iii) performing a regression analysis between the real and imaginary parts of the sample and theoretical CF (Koutrouvelis, 1980), or (iv) using the fast Fourier transform to express the likelihood as a function of the CF (Chenyao et al., 1999).

4 About conditional quantile based kurtosis measures see Coroneo and Veredas (forthcoming).