A comparison of semiparametric density estimation methods for multivariate risk management

Marco Bee
Department of Economics, University of Trento

Abstract Elliptical distributions are effective models of financial returns. Parametric estimation has been developed for cases obtained with specific density generators. A more general approach uses semi-parametric techniques. In this paper we propose to estimate the density generator by means of the Maximum Entropy method.

Key words: Maximum entropy; Semiparametric Estimation; Elliptical Distributions; Risk Measures.

1 Introduction

Probability distributions of financial log-returns are usually non-standard. This is caused by a number of stylized facts commonly observed in financial data, such as dependence in time, non constant volatility and leptokurtosis; see McNeil et al. (2005, Sect. 4.1.1). At the portfolio level, the distribution may be thought of as the combined effect of the marginal distributions and the dependence structure. This distribution is typically non-standard as well, because also the dependence structure of financial variables presents specific features, such as tail dependence. Thus, even models incorporating non-standard marginal distributions may fail to adequately represent the joint distribution. Two possible ways out are distributions based on copulas (possibly with tail dependence) and elliptical distributions. In the following we focus on the latter.

2 The model and the estimation methodology

Elliptical distributions include some models that have proved to be effective for multivariate financial data, such as the normal variance-mean mixtures class, which contains as a special case the family of generalized hyperbolic distributions (McNeil et al., 2005, p. 94). The density of an r-dimensional elliptical random vector $\boldsymbol{X}$ with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ is given by

$$f(\boldsymbol{x}) = \frac{1}{(\det \boldsymbol{\Sigma})^{1/2}} g(Q), \tag{1}$$

where $Q = (\boldsymbol{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu})$ and g is a real function called density generator.

Parametric estimation techniques have been developed for cases corresponding to specific density generators. More generally, the semi-parametric approach introduced by Bingham et al. (2003) uses the median estimator for $\boldsymbol{\mu}$ and an LS-estimator for $\boldsymbol{\Sigma}$, and estimates g nonparametrically by means of a kernel density approach. However, the application of the latter method is not straightforward, as the choice of the bandwidth requires to set several parameters. Moreover, the estimated generator is evaluated at a finite number of points. In order to have a continuous generator not affected by subjective decisions of the analyst, we propose to use the Maximum Entropy (ME) approach for estimating the density generator.

The ME distribution is "uniquely determined as the one which is maximally noncommittal with respect to missing information, and it agrees with what is known, but expresses maximum uncertainty with respect to all other matters" (Jaynes, 1957). The ME density is given by

$$f(x) = \exp\left\{ - \sum_{i=0}^{k} \lambda_i g_i(x) \right\}, \tag{2}$$

where k is the number of moment constraints and g_is are the functional forms of the so-called "characterizing moments". In most cases, they are the arithmetic or logarithmic moments, corresponding respectively to $g(x) = x$ and $g(x) = \log(x)$. The $k+1$ parameters are the Lagrange multipliers of the maximization problem

$$\max_f W = \int f(x) \log(f(x)) dx, \tag{3}$$

under the constraints $\int g_i(x) f(x) dx = \hat{\mu}_i$ $(i = 0, 1, \ldots, k)$, where f is a density, W is the Shannon entropy associated to f and $\hat{\mu}_i$ is the sample counterpart of the i-th characterizing moment. It can be shown that the solution, called ME density, takes the form (2).

The problem (3) cannot be solved analytically for $k \geq 2$. However, Wu (2003) has proposed an algorithm that imposes the constraints one at a time, from the lowest to the highest moment, with the result that the maximization can be easily carried out with standard Newton-Raphson. As for the choice of k, a likelihood ratio test can be used. The maximized log-likelihood is equal to $-n \sum_{i=0}^{k} \lambda_i \hat{\mu}_i$, where n is the sample size. Hence, the test of $H_0 : k = s$ $(s = 1, 2, \ldots)$ against $H_1 : k = s+1$ is given by $\text{llr} = -2n(\sum_{i=0}^{s+1} \hat{\lambda}_i \hat{\mu}_i - \sum_{i=0}^{s} \hat{\lambda}_i \hat{\mu}_i)$; from standard limiting theory, its asymptotic distribution is χ^2_1. Thus, the model-selection procedure is based on the following steps: (a) estimate sequentially the ME density with $s = 1, 2, \ldots$, (b) perform the test for each value of s, (c) stop at the first value of s (s_0, say) such that the hypothesis $s = s_0$ cannot be rejected and conclude that the optimal value of k (from now on denoted by k^*) is equal to s_0.

3 Application

It can be shown (Bingham et al., 2003) that the Value at Risk (VaR) at coverage level α for (1) is given by:

$$VaR_\alpha = \boldsymbol{x}'\boldsymbol{\mu} - h(\alpha)\sqrt{\boldsymbol{x}'\boldsymbol{\Sigma}\boldsymbol{x}},$$

where $h(\alpha)$ is the $1 - 2\alpha$ quantile of $R_r B$, with $B^2 \sim \text{Beta}(1/2, (r-1)/2)$ and

$$f_{R_r^2}(u) = (\pi^{r/2}/\Gamma(r/2))u^{r/2-1}g(u). \quad (4)$$

Moreover, B is independent from R_r. It can also be shown that $f_{R_r^2}$ is the density of the r.v. $(\boldsymbol{X} - \boldsymbol{\mu})'\boldsymbol{\Sigma}(\boldsymbol{X} - \boldsymbol{\mu})$, so that we use observations $(\boldsymbol{x_i} - \hat{\boldsymbol{\mu}})'\hat{\boldsymbol{\Sigma}}(\boldsymbol{x_i} - \hat{\boldsymbol{\mu}})$ for estimating $f_{R_r^2}$ via the ME method and invert (4) to obtain an estimate of the generator. For comparison purposes, we also estimate $f_{R_r^2}$ by means of a kernel density (KD), as proposed by Bingham et al. (2003). In this case, setting all the inputs is rather involved (Bingham et al., 2003, Sect. 4.2).

If x_α is the $(1-2\alpha)$-quantile of $R_r B$, x_α^2 is the $(1-2\alpha)$-quantile of $Z = R_r^2 B^2$. Thus, if we were able to find the density of Z, we could compute the VaR explicitly. Unfortunately, both when g is the ME density and when it is KD, the integral cannot be solved explicitly. Thus, we resort to Monte Carlo simulation. For simulating R_r we use the inverse distribution function method. As the ME cdf is not available in closed form, we compute it by means of standard numerical integration techniques.

As for the parametric part of the model, the estimates are the same used by Bingham et al. (2003), i.e. $\hat{\mu} = \text{median}(x_i)$ and $\hat{\boldsymbol{\Sigma}}$ given by the LS-estimator. We implement the model on a four-dimensional portfolio consisting of the Dow Jones, Dax 30, CAC 40 and Nikkei 225 log-returns and a two-dimensional portfolio formed by the Nasdaq composite and Tecdax indexes. For the former, the time horizon is Jan. 1, 1997 to Jan, 31, 2012, whereas for the latter it is Jan. 1, 1998 to Jan, 31, 2012.

The generator must be monotonic, and, in general, neither the ME nor the KD estimate are. To this aim, we smooth them by means of a monotone regression method based on the Generalized Pooled-Adjacent-Violators Algorithm (Härdle, 1990, Sect. 8.1). Fig. 1 shows the estimated generators. In both cases, they are very close to each other, but the ME-based generator is smoother. VaR estimates are shown in Table 1. It is clear that the two methods give essentially the same results.

Two main conclusions can be drawn: the method is robust, as different ways of estimating the generator produce the same results, and the ME-based generator estimation method is preferable, as it requires essentially no input from the user.

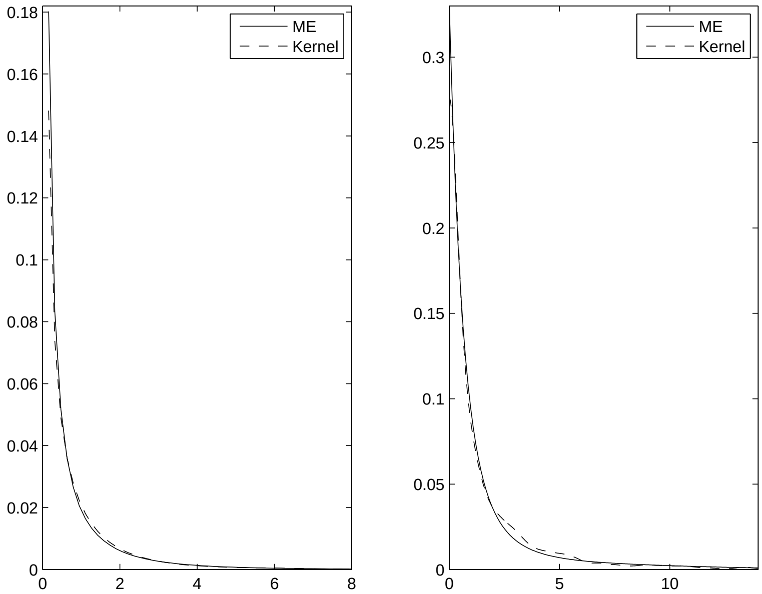

Fig. 1 The ME and KD generators in the four-dimensional (left panel) and two-dimensional (right panel) case. When $r = 2$, the optimal ME density has $k = 6$; when $r = 4$, it has $k = 5$.

	$\alpha = 0.05$				$\alpha = 0.025$				$\alpha = 0.01$			
	ME	KD	Norm.	Emp.	ME	KD	Norm.	Emp.	ME	KD	Norm.	Emp.
4-dim.	-112.92	-110.98	-119.90	-122.21	-143.06	-142.40	-142.04	-147.49	-200.60	-201.56	-170.01	-202.95
2-dim.	-45.87	-46.26	-49.00	-50.78	-63.39	-63.06	-60.67	-67.12	-86.62	-84.64	-69.61	-95.13

Table 1 VaR measures at different coverage levels for four- and two-dimensional portfolios.

References

Bingham, N., Kiesel, R., and Schmidt, R. (2003). A semi-parametric approach to risk management. *Quantitative Finance*, 3:426–441.

Härdle, W. (1990). *Applied nonparametric regression*. Cambridge University Press.

Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physics Review*, 106:620–630.

McNeil, A., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, Tools*. Princeton University Press.

Wu, X. (2003). Calculation of maximum entropy densities with application to income distribution. *Journal of Econometrics*, 115:347–354.