Professor Aida TOMA, PhD
Department of Applied Mathematics
The Bucharest Academy of Economic Studies
ISMMA, Romanian Academy
E-mail: aida_toma@yahoo.com

ROBUST ESTIMATIONS FOR FINANCIAL RETURNS: AN APPROACH BASED ON PSEUDODISTANCE MINIMIZATION

Abstract. *We consider a new robust estimation method based on pseudodistances minimization. Unlike other existing methods of this type, such as minimum Hellinger distance estimation, the proposed approach avoid the use of nonparametric density estimation and associated complications such as bandwidth selection. The proposed class of estimators is indexed by a single parameter α which controls the trade-off between robustness and efficiency. The method is applied to expected return and volatility estimation of financial asset returns under normality. Empirical results on simulated or financial data prove the performance of this method.*

Keywords: pseudodistance, robust estimation, influence function, financial returns.

JEL Classification: C02, C13, G10.

1 Introduction

Many decision-making and asset pricing models in finance rely on assumptions on the stochastic model underlying the data. The normal model is one of the most used. However, for a typical sample of financial returns, the empirical distribution can deviate from normality in various degrees. Therefore, some sophisticated representations have been proposed in literature (see Bawens et al. (2006), Zhao (2008)). In the meantime, a valid alternative is to retain a simple stochastic model and to use an estimation method stable to possible deviations from the assumed model. In the present paper we consider this approach by using a new robust estimation method for financial returns.

The notion which stands at the basis of this method is the pseudodistance between probability measures. Like divergences, the pseudodistances are information measures and generalizations of distances between probability measures, they not satisfying the triangle inequality. The information measures, and particularly the divergences, are widely used in statistical inference (see Pardo (2006) and the references herein) and are also useful in buiding different models with applications in system management, in allocation problems, or in transportation problems (see Purcaru et al. (2009), Purcaru and Verboncu (2010), Purcaru (2011)).

The robust estimation method which we consider was introduced by Broniatowski and Vajda (2009) and consists in minimization of an empirical version of a pseudodistance between the assumed model and the true model by using the empirical measure pertaining to the sample. This method can be applied to any parametric model, but in the present paper we focus on the normal location-scale model. The method is appealing, since it conciliate robustness and efficiency, usually requiring distinct techniques. The behavior of the estimator depends on a tuning positive parameter α which controls the trade-off between the two properties. When the data are consistent with normality and $\alpha \to 0$, the estimation method corresponds to the maximum likelihood method (MLE) which is known to have full asymptotic efficiency at the model. When $\alpha > 0$, the estimator gains robustness, while keeping high efficiency.

The outline of the paper is as follows: in Section 2 we present the estimation method and its asymptotic properties. The performance in finite samples of the method is studied in Section 3 through Monte Carlo simulations, for both contaminated and noncontaminated samples. Finally, the estimation method is applied to monthly log-returns of the Standard & Poor's 500 stock index.

2 Minimum pseudodistance estimators

2.1 Context and definition

The minimum divergence estimators and related methods have received a considerable attention in recent years. Among others, Bouzebda and Keziou (2010), Karagrigoriou and Mattheou (2010), Mattheou and Karagrigoriou (2010), Toma (2007), Toma (2009), Toma and Leoni-Aubin (2010), Toma and Broniatowski (2011) have developed divergence based estimation and testing methods and proved advantages of using these methods. In the same line of research, Broniatowski and Vajda (2009) introduced a family of pseudodistances between probability measures which they used to define minimum pseudodistance estimators for general parametric models. This new family of estimators is indexed by a positive tuning parameter α.

For two probability measures P and Q, admitting densities p, respectively q with respect to some dominating σ-finite measure λ, these pseudodistances are defined through

$$R_\alpha(P,Q) := \frac{1}{\alpha+1}\ln\int p^\alpha \mathrm{d}P + \frac{1}{\alpha(\alpha+1)}\ln\int q^\alpha \mathrm{d}Q - \frac{1}{\alpha}\ln\int p^\alpha \mathrm{d}Q \qquad (2.1)$$

for $\alpha > 0$ and satisfy the limit relation

$$R_\alpha(P,Q) \to R_0(P,Q) := \int \ln\frac{q}{p}\mathrm{d}Q \quad \text{for} \quad \alpha \downarrow 0. \qquad (2.2)$$

Note that $R_0(P,Q)$ coincides with the modified Kullback-Leibler divergence.

Let $\mathcal{P}$ be a parametric model with parameter space $\Theta \subset \mathbb{R}^d$ and assume that every probability measure P_θ in $\mathcal{P}$ has a density p_θ with respect to the Lebesgue measure.

The family of minimum pseudodistance estimators of the unknown parameter θ_0 is obtained by replacing the hypothetical probability measure P_{θ_0} in the pseudodistances $R_\alpha(P_\theta, P_{\theta_0})$ by the empirical measure P_n pertaining to the sample and then minimizing $R_\alpha(P_\theta, P_{\theta_0})$ with respect to θ. When the true distribution of the observations is associated to a measure P, the target parameter is the point $\theta_0 \in \Theta$ corresponding to the measure P_θ closest to P according to the pseudodistance $R_\alpha(P_\theta, P)$. In this parametric framework, the density p_{θ_0} can be interpreted as the projection of the true density p on the parametric family. Of course, if P is a member of the family then $p_\theta = p_{\theta_0}$.

Eliminating the terms that not involve θ, the minimum pseudodistance estimator (in symbol $\min R_\alpha$ estimator) is defined by

$$\widehat{\theta}_n(\alpha) = \begin{cases} \arg\inf_\theta \left[\frac{1}{1+\alpha}\ln\left(\int p_\theta^\alpha \mathrm{d}P_\theta\right) - \frac{1}{\alpha}\ln(\int p_\theta^\alpha \mathrm{d}P_n)\right] & \text{if } \alpha > 0 \\ \arg\inf_\theta -\int \ln p_\theta \mathrm{d}P_n & \text{if } \alpha = 0. \end{cases} \tag{2.3}$$

or equivalently as

$$\widehat{\theta}_n(\alpha) = \begin{cases} \arg\sup_\theta C_\alpha(\theta)^{-1}\frac{1}{n}\sum_{i=1}^n p_\theta^\alpha(X_i) & \text{if } \alpha > 0 \\ \arg\sup_\theta \frac{1}{n}\sum_{i=1}^n \ln p_\theta(X_i) & \text{if } \alpha = 0 \end{cases} \tag{2.4}$$

where $C_\alpha(\theta) = \left(\int p_\theta^{1+\alpha}\mathrm{d}\lambda\right)^{\alpha/(1+\alpha)}$.

For $\alpha > 0$, the $\min R_\alpha$ estimator entails solving the estimating equation

$$\sum_{i=1}^n p_\theta^{\alpha-1}(X_i)\dot{p}_\theta(X_i) - c_\alpha(\theta)p_\theta^\alpha(X_i) = 0,$$

where $c_\alpha(\theta) = \frac{\int p_\theta^\alpha \dot{p}_\theta d\lambda}{\int p_\theta^{\alpha+1} d\lambda}$, while for $\alpha = 0$ the estimating equation is

$$\sum_{i=1}^n \frac{\dot{p}_\theta(X_i)}{p_\theta(X_i)} = 0. \tag{2.5}$$

Note that (2.5) is the equation defining the classical maximum likelihood estimator (MLE).

In the present paper our interest is on the normal location-scale model. When $\mathcal{P}$ is the normal location-scale model, denoting $\theta_0 := (\mu_0, \sigma_0)$ the parameter of interest, the definition of a $\min R_\alpha$ estimator of θ_0 corresponding to a positive α becomes

$$\widehat{\theta}_n(\alpha) = \arg\sup_{\mu,\sigma} \frac{1}{n} \left(\frac{\sqrt{\alpha+1}}{\sigma\sqrt{2\pi}} \right)^{\alpha/(\alpha+1)} \sum_{i=1}^{n} \exp\left(-\frac{\alpha}{2} \left(\frac{X_i - \mu}{\sigma} \right)^2 \right) \tag{2.6}$$

where $\widehat{\theta}_n(\alpha) = (\widehat{\mu}_n(\alpha), \widehat{\sigma}_n(\alpha))$. In the rest of the paper we are most interested by the case positive α, since the case $\alpha = 0$ leads to the well known classical maximum likelihood procedure.

2.2 Robustness properties

The tuning parameter α which indexes the class of estimators balances infinitesimal robustness and asymptotic efficiency.

The robustness measure that we use in the present paper is the influence function of the statistical functional corresponding to the estimator. Recall that, a map T defined on a set of probability measures and parameter space valued is a statistical functional corresponding to an estimator $\widehat{\theta}_n$ of the parameter θ, if $\widehat{\theta}_n = T(P_n)$, where P_n is the empirical measure associated to the sample. As it is known, the influence function of T at P_θ is defined by

$$\mathrm{IF}(x; T, P_\theta) := \left. \frac{\partial T(\widetilde{P}_{\varepsilon x})}{\partial \varepsilon} \right|_{\varepsilon=0}$$

where $\widetilde{P}_{\varepsilon x} := (1-\varepsilon)P_\theta + \varepsilon\delta_x$, δ_x being the Dirac measure putting all mass at x. The influence function measures the standardized effect of an infinitezimal contamination in a point x on the asymptotic value of the estimator. Whenever the influence function is bounded with respect to x the corresponding estimator is called robust.

For fixed $\alpha > 0$, the statistical functional corresponding to the $\min R_\alpha$ estimator defined by (2.4) is

$$T_\alpha(Q) = \arg\sup_{\theta} \frac{\int p_\theta^\alpha \mathrm{d}Q}{C_\alpha(\theta)} \tag{2.7}$$

and the corresponding influence function is

$$\mathrm{IF}(x; T_\alpha, P_\theta) = \left. \frac{\partial}{\partial \varepsilon} T_\alpha(\widetilde{P}_{\varepsilon x}) \right|_{\varepsilon=0} = M_\alpha(\theta)^{-1} [p_\theta^{\alpha-1}(x) \dot{p}_\theta(x) - c_\alpha(\theta) p_\theta^\alpha(x)] \tag{2.8}$$

where

$$M_\alpha(\theta) = \int p_\theta^{\alpha-1} \dot{p}_\theta \dot{p}_\theta^t \mathrm{d}\lambda - \frac{\int p_\theta^\alpha \dot{p}_\theta \mathrm{d}\lambda (\int p_\theta^\alpha \dot{p}_\theta \mathrm{d}\lambda)^t}{\int p_\theta^{\alpha+1} \mathrm{d}\lambda} \tag{2.9}$$

$\dot{p}_\theta$ being the derivative of p_θ with respect to θ.

Assuming that $M_\alpha(\theta)$ is finite, $\mathrm{IF}(x; T_\alpha, P_\theta)$ is a bounded function of x whenever $p_\theta^{\alpha-1}(x)\dot{p}_\theta(x)$ and $p_\theta^\alpha(x)$ are bounded. For example, this is true for any $\alpha > 0$ in the normal location-scale model, unlike other minimum divergence procedures, such as those based on the Hellinger distance.

Indeed, the influence function of the statistical functional T_α corresponding to the estimator (2.6) is given by

$$\begin{aligned}
\mathrm{IF}_1(x; T_\alpha, P_\theta) &= (\alpha+1)^{3/2}(x-\mu)\exp\left(-\frac{\alpha}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) \\
\mathrm{IF}_2(x; T_\alpha, P_\theta) &= \frac{\sigma(\alpha+1)^{5/2}}{2}\left[\left(\frac{x-\mu}{\sigma}\right)^2 - \frac{1}{\alpha+1}\right]\exp\left(-\frac{\alpha}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)
\end{aligned}$$

where $\mathrm{IF}_1(x; T_\alpha, P_\theta)$ represents the influence function of the mean estimator $\widehat{\mu}_n(\alpha)$ and $\mathrm{IF}_2(x; T_\alpha, P_\theta)$ the influence function of the estimator $\widehat{\sigma}_n(\alpha)$ of the standard deviation σ. Note that the above influence functions correspond to redescending estimators, meaning that the IFs approach to zero as $|x| \to \infty$. Therefore large outliers will have a very reduced influence on the estimations.

2.3 Asymptotic properties and choice of α

For fixed α, the solution of the optimization problem (2.4) is an M-estimator and the asymptotic distribution can be derived from existing theory (Hampel et al (1986), van der Vaart (1989), Broniatowski et al. (2011)).

Suppose that $X_1, \ldots, X_n$ are i.i.d. random variables with a distribution associated to the measure P. Under standard regularity conditions it hold:

1. There exists a sequence $\widehat{\theta}_n(\alpha)$ which converges to θ_0 in probability as $n \to \infty$.

2. For any consistent sequence $\widehat{\theta}_n(\alpha)$, $\sqrt{n}(\widehat{\theta}_n(\alpha) - \theta_0)$ converges in distribution to a multivariate normal with mean vector zero and the covariance matrix

$$V_\alpha(\theta_0, P) = J_\alpha(\theta_0)^{-1}K_\alpha(\theta_0)J_\alpha(\theta_0)^{-1} \tag{2.10}$$

where

$$\begin{aligned}
J_\alpha(\theta) &= -H_\theta\left[\int h(\theta, x)p(x)dx\right] \\
K_\alpha(\theta) &= \int \frac{\partial}{\partial\theta}h(\theta, x)\left[\frac{\partial}{\partial\theta}h(\theta, x)\right]^t p(x)dx
\end{aligned}$$

with $h(\theta, x) = \frac{p_\theta^\alpha(x)}{C_\alpha(\theta)}$, when $\alpha > 0$, H_θ denoting the second derivative (the Hessian matrix).

Table 1. MSE in the case of noncontaminated data

	$\alpha = \alpha^*$	MLE	$\alpha = 0.5$	$\alpha = 1$
$n = 50$	0.0392	0.0314	0.0403	0.0631
$n = 100$	0.0149	0.0147	0.0186	0.0273
$n = 500$	0.0029	0.0029	0.0037	0.0053

In the particular case $\alpha = 0$, the asymptotic covariance matrix $V_0(\theta_0, P)$ of the MLE is given by the formula (2.10) with

$$\begin{aligned} J_0(\theta) &= -\int \frac{\ddot{p}_\theta(x)}{p_\theta(x)} p(x)dx + \int \frac{\dot{p}_\theta(x)\dot{p}_\theta(x)^t}{p_\theta^2(x)} p(x)dx \\ K_0(\theta) &= \int \frac{\dot{p}_\theta(x)\dot{p}_\theta(x)^t}{p_\theta^2(x)} p(x)dx \end{aligned}$$

$\ddot{p}_\theta$ being the second derivative of p_θ with respect to θ. The above expressions follow directly from Theorems 5.14 and 5.41 in van der Vaart (1998). If $\alpha = 0$ and the model is correctly specified, i.e. $p(x) = p_{\theta_0}(x)$, it can be seen that $K_\alpha(\theta_0) = J_\alpha(\theta_0)$ and $V_\alpha(\theta_0)$ is just the inverse of the Fisher information matrix.

In the case of the univariate normal model, it holds

$$h(\theta, x) = \left(\frac{\sqrt{\alpha+1}}{\sigma\sqrt{2\pi}}\right)^{\alpha/(\alpha+1)} \exp\left(-\frac{\alpha}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right). \tag{2.11}$$

Different values of α correspond to estimators with different robustness and efficiency levels. One approach for choosing α is to select the estimator with the largest empirical efficiency among all $\min R_\alpha$ estimators with $\alpha \leq 1$. (We do not take $\alpha > 1$ since the corresponding estimators have unacceptable low asymptotic efficiency.) In this sense, we consider the ratio

$$\Lambda(\alpha, \theta_0, P) = V_0(\theta_0, P)V_\alpha^{-1}(\theta_0, P), \tag{2.12}$$

where $V_\alpha(\theta_0, P)$ is given by (2.10) (see Ferrari and Young (2010) for a similar formula). Since θ_0 and P are unknown, we consider a grid of escort parameters $A = \{\alpha_1, \ldots, \alpha_r\}$ and compute the estimates $\widehat{\theta}_n(\alpha_1), \ldots, \widehat{\theta}_n(\alpha_r)$. Then we choose α^* defined by

$$\alpha^* := \arg\max_{\alpha \in A} \operatorname{tr}\{\widehat{\Lambda}(\alpha, \widehat{\theta}_n(\alpha), P_n)\}, \tag{2.13}$$

where $\widehat{\Lambda}(\alpha, \widehat{\theta}_n(\alpha), P_n)$ is obtained from $\Lambda(\alpha, \theta_0, P)$ by replacing θ_0 with the estimate $\widehat{\theta}_n(\alpha)$ and P with the empirical measure P_n associated to the sample.

3 Monte Carlo simulations

We performed a Monte Carlo simulation study in order to evaluate the performance of the proposed estimators. We investigated the efficiency and robustness for various levels and type of contamination. For a sample size n we generated B samples containing about $(1-\varepsilon)n$ observations from $\mathcal{N}(\mu_0, \sigma_0)$, while the proportion εn comes from the contaminating distribution $\mathcal{N}(\mu_c, \sigma_c)$. We fixed $\mu_0 = 0$, $\sigma_0 = 1$ and generated samples of size $n = 50, 100, 500$, with contamination $\varepsilon = 5\%, 10\%, 20\%$, for each configuration the number of replications being $B = 1000$.

We considered the $\min R_\alpha$ estimator with α selected by the formula (2.13), the MLE, as well as the $\min R_\alpha$ estimators with $\alpha = 0.5$, respectively with $\alpha = 1$.

Table 2 presents the mean squared errors of these estimators, when $\mu_c = 2, 4, 6, 8$ and $\sigma_c = \sigma_0$, while Table 3 presents the mean squared errors of the same estimators when $\mu_c = \mu_0$ and $\sigma_c^2 = 2, 4, 6, 8$. The results in these tables are obtained for $n = 50$. Similar results are given in Tables 4,5,6,7 for $n = 100, 500$.

As it can be seen, in the presence of contamination, the $\min R_\alpha$ estimators outperform the MLE. Few exceptions are for $\sigma_c^2 = 2$, situation in which the contaminating distribution is very close to the true one. However, the mean square errors of the MLE and of the $\min R_\alpha$ estimator with $\alpha = \alpha^*$ are very close in this particular case. Generally, for contaminated data, the best results are provided by the $\min R_\alpha$ estimator with $\alpha = 0.5$. In the meantime, the choice $\alpha = \alpha^*$ leads to estimations which are close to the best ones in most cases. Also, when the data are not contaminated, the $\min R_\alpha$ estimator with $\alpha = \alpha^*$ is very close to the MLE which provides the best results in this case. Such results are given in Table 1.

Thus, the proposed procedure for choosing α leads to estimations which tend to be close to the best ones in each case. Usually in practice we do not dispose of any information, whether the data are contaminated or not. Therefore the $\min R_\alpha$ procedure with α^* selected could be preferred, being an adaptive one, flexible for each situation in part.

4 Real data example

We apply our method to 263 monthly observations of the log-returns of the Standard & Poor's 500 from December 1987 to October 2009. Figure 1 shows a normal quantile plot for these data. While the bulk of the data follows normality fairly closely, there are values in the tails which depart from normality in various degrees.

In Table 8 we give the estimates of the mean and of the standard deviation computed with min R_α corresponding to α^* and to $\alpha \in \{0, 0.5, 1\}$. A visual rep

Table 2. MSE when $n = 50$ and $\mu_c = 2, 4, 6, 8$

	$\alpha = \alpha^*$	MLE	$\alpha = 0.5$	$\alpha = 1$
		$\varepsilon = 5\%$		
$\mu_c = 2$	0.0422	0.0389	0.0412	0.0617
$\mu_c = 4$	0.0731	0.1172	0.0451	0.0694
$\mu_c = 6$	0.0510	0.3344	0.0416	0.0674
$\mu_c = 8$	0.0413	0.7612	0.0417	0.0641
		$\varepsilon = 10\%$		
$\mu_c = 2$	0.0982	0.0957	0.0690	0.0822
$\mu_c = 4$	0.3550	0.4906	0.0644	0.0710
$\mu_c = 6$	0.1432	1.4644	0.0471	0.0696
$\mu_c = 8$	0.0467	3.1122	0.0418	0.0652
		$\varepsilon = 20\%$		
$\mu_c = 2$	0.2661	0.2684	0.1944	0.1691
$\mu_c = 4$	1.0282	1.4360	0.3305	0.1091
$\mu_c = 6$	0.3385	3.9702	0.0606	0.0777
$\mu_c = 8$	0.0633	7.9597	0.0475	0.0795

Table 3. MSE when $n = 50$ and $\sigma_c^2 = 2, 4, 6, 8$

	$\alpha = \alpha^*$	MLE	$\alpha = 0.5$	$\alpha = 1$
		$\varepsilon = 5\%$		
$\sigma_c^2 = 2$	0.0366	0.0302	0.0378	0.0618
$\sigma_c^2 = 4$	0.0404	0.0379	0.0413	0.0616
$\sigma_c^2 = 6$	0.0400	0.0467	0.0392	0.0615
$\sigma_c^2 = 8$	0.0454	0.0571	0.0420	0.0658
		$\varepsilon = 10\%$		
$\sigma_c^2 = 2$	0.0376	0.0336	0.0405	0.0611
$\sigma_c^2 = 4$	0.0570	0.0613	0.0475	0.0676
$\sigma_c^2 = 6$	0.0648	0.0945	0.0481	0.0691
$\sigma_c^2 = 8$	0.0652	0.1375	0.0481	0.0655
		$\varepsilon = 20\%$		
$\sigma_c^2 = 2$	0.0484	0.0431	0.0481	0.0704
$\sigma_c^2 = 4$	0.0858	0.1069	0.0624	0.0786
$\sigma_c^2 = 6$	0.1137	0.2017	0.0681	0.0787
$\sigma_c^2 = 8$	0.1250	0.3473	0.0718	0.0824

Table 4. MSE when $n = 100$ and $\mu_c = 2, 4, 6, 8$

	$\alpha = \alpha^*$	MLE	$\alpha = 0.5$	$\alpha = 1$
		$\varepsilon = 5\%$		
$\mu_c = 2$	0.0310	0.0323	0.0254	0.0329
$\mu_c = 4$	0.0715	0.1485	0.0241	0.0317
$\mu_c = 6$	0.0202	0.4758	0.0200	0.0290
$\mu_c = 8$	0.0181	1.0836	0.0198	0.0303
		$\varepsilon = 10\%$		
$\mu_c = 2$	0.0787	0.0807	0.0501	0.0475
$\mu_c = 4$	0.3979	0.4737	0.0381	0.0363
$\mu_c = 6$	0.0795	1.4245	0.0220	0.0316
$\mu_c = 8$	0.0199	3.0675	0.0216	0.0309
		$\varepsilon = 20\%$		
$\mu_c = 2$	0.2517	0.2532	0.1805	0.1391
$\mu_c = 4$	1.2165	1.4123	0.2982	0.0494
$\mu_c = 6$	0.1744	3.9703	0.0266	0.0357
$\mu_c = 8$	0.0239	7.9887	0.0238	0.0356

Table 5. MSE when $n = 100$ and $\sigma_c^2 = 2, 4, 6, 8$

	$\alpha = \alpha^*$	MLE	$\alpha = 0.5$	$\alpha = 1$
		$\varepsilon = 5\%$		
$\sigma_c^2 = 2$	0.0163	0.0160	0.0197	0.0284
$\sigma_c^2 = 4$	0.0205	0.0230	0.0222	0.0310
$\sigma_c^2 = 6$	0.0219	0.0331	0.0216	0.0309
$\sigma_c^2 = 8$	0.0223	0.0476	0.0212	0.0296
		$\varepsilon = 10\%$		
$\sigma_c^2 = 2$	0.0171	0.0173	0.0204	0.0298
$\sigma_c^2 = 4$	0.0268	0.0374	0.0236	0.0311
$\sigma_c^2 = 6$	0.0352	0.0696	0.0266	0.0332
$\sigma_c^2 = 8$	0.0366	0.1139	0.0265	0.0332
		$\varepsilon = 20\%$		
$\sigma_c^2 = 2$	0.0254	0.0264	0.0254	0.0339
$\sigma_c^2 = 4$	0.0621	0.0884	0.0407	0.0432
$\sigma_c^2 = 6$	0.0881	0.1944	0.0470	0.0445
$\sigma_c^2 = 8$	0.0892	0.3155	0.0492	0.0438

Table 6. MSE when $n = 500$ and $\mu_c = 2, 4, 6, 8$

	$\alpha = \alpha^*$	MLE	$\alpha = 0.5$	$\alpha = 1$
		$\varepsilon = 5\%$		
$\mu_c = 2$	0.0191	0.0200	0.0100	0.0090
$\mu_c = 4$	0.0648	0.1397	0.0057	0.0058
$\mu_c = 6$	0.0038	0.4657	0.0041	0.0057
$\mu_c = 8$	0.0035	1.0749	0.0041	0.0059
		$\varepsilon = 10\%$		
$\mu_c = 2$	0.0663	0.0680	0.0336	0.0225
$\mu_c = 4$	0.4642	0.4658	0.0160	0.0071
$\mu_c = 6$	0.0041	1.4286	0.0043	0.0059
$\mu_c = 8$	0.0036	3.0768	0.0041	0.0058
		$\varepsilon = 20\%$		
$\mu_c = 2$	0.2375	0.2380	0.1617	0.1129
$\mu_c = 4$	1.4089	1.4104	0.2646	0.0111
$\mu_c = 6$	0.0049	3.9557	0.0052	0.0068
$\mu_c = 8$	0.0043	8.0141	0.0049	0.0069

Table 7. MSE when $n = 500$ and $\sigma_c^2 = 2, 4, 6, 8$

	$\alpha = \alpha^*$	MLE	$\alpha = 0.5$	$\alpha = 1$
		$\varepsilon = 5\%$		
$\sigma_c^2 = 2$	0.0035	0.0036	0.0040	0.0057
$\sigma_c^2 = 4$	0.0059	0.0085	0.0047	0.0059
$\sigma_c^2 = 6$	0.0069	0.0160	0.0050	0.0061
$\sigma_c^2 = 8$	0.0075	0.0272	0.0052	0.0064
		$\varepsilon = 10\%$		
$\sigma_c^2 = 2$	0.0053	0.0056	0.0052	0.0066
$\sigma_c^2 = 4$	0.0135	0.0226	0.0080	0.0080
$\sigma_c^2 = 6$	0.0163	0.0505	0.0082	0.0074
$\sigma_c^2 = 8$	0.0169	0.0896	0.0089	0.0079
		$\varepsilon = 20\%$		
$\sigma_c^2 = 2$	0.0113	0.0123	0.0091	0.0098
$\sigma_c^2 = 4$	0.0439	0.0712	0.0207	0.0145
$\sigma_c^2 = 6$	0.0611	0.1670	0.0273	0.0164
$\sigma_c^2 = 8$	0.0579	0.2917	0.0295	0.0155

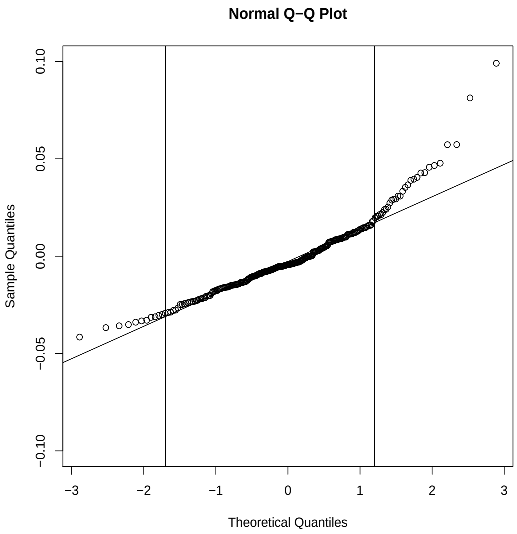

Figure 1. QQ-plot of monthly log-returns of the S&P 500 stock index

Table 8. Expected return and volatility estimates for monthly S&P 500

	$\alpha = \alpha^*$	MLE	$\alpha = 0.5$	$\alpha = 1$
μ	0.0048	0.0025	0.0045	0.0051
σ	0.0157	0.0185	0.0161	0.0151

resentation of this is given in Figure 2 where the normal densities $\mathcal{N}(\widehat{\mu}_n, \widehat{\sigma}_n)$ are superimposed on a histogram of the data. Excepting the case when MLE is used, all the normal densities fit the main body of the data quite well. In Figure 3 we can see the influence of the observations on the min R_α estimates with $\alpha = 1$ (in the left hand side the influence on the mean and in the right hand side, the influence on the standard deviation). Extreme observations, both positive and negative, which would affect the final estimates have nearly zero influence on the estimates.

Finally, we compute estimates for the annual expected returns and volatilities. The estimates are obtained by the min R_α method with optimally chosen α, respectively with $\alpha \in \{0, 0.5, 1\}$. The results are presented in Figure 4. The min R_α estimators with $\alpha = 0.5, 1$ provide results that are less affected by the presence of atypical observations in the sample. Also, the choice of α^* allows for a flexible treatment of the periods characterized by high (low) volatilities and by the presence of anomalous data.

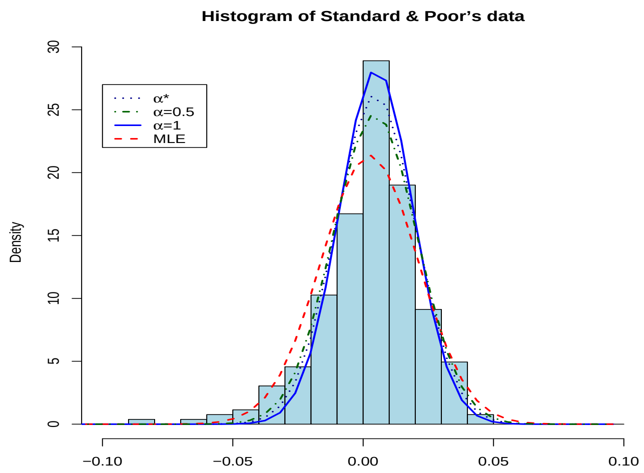

Figure 2. Histogram of the S&P 500 data with normal fits

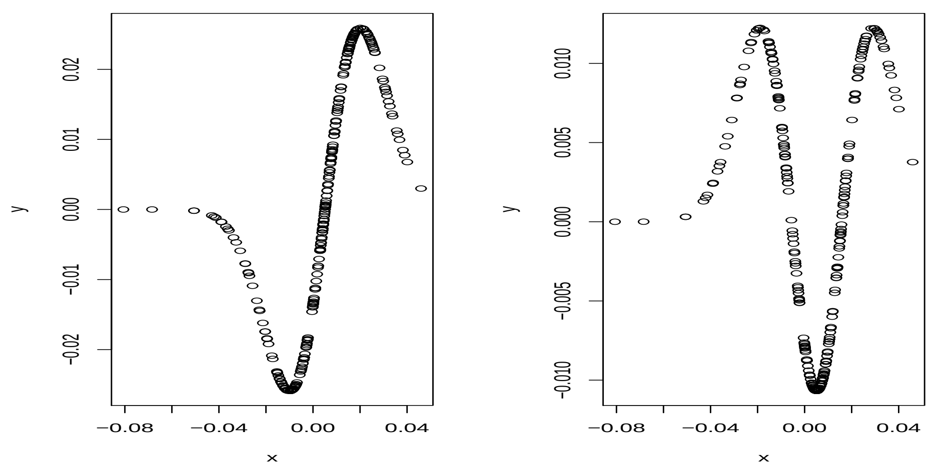

Figure 3. The influence of the observations on the expected return (left hand side) and on the volatility (right hand side) estimates

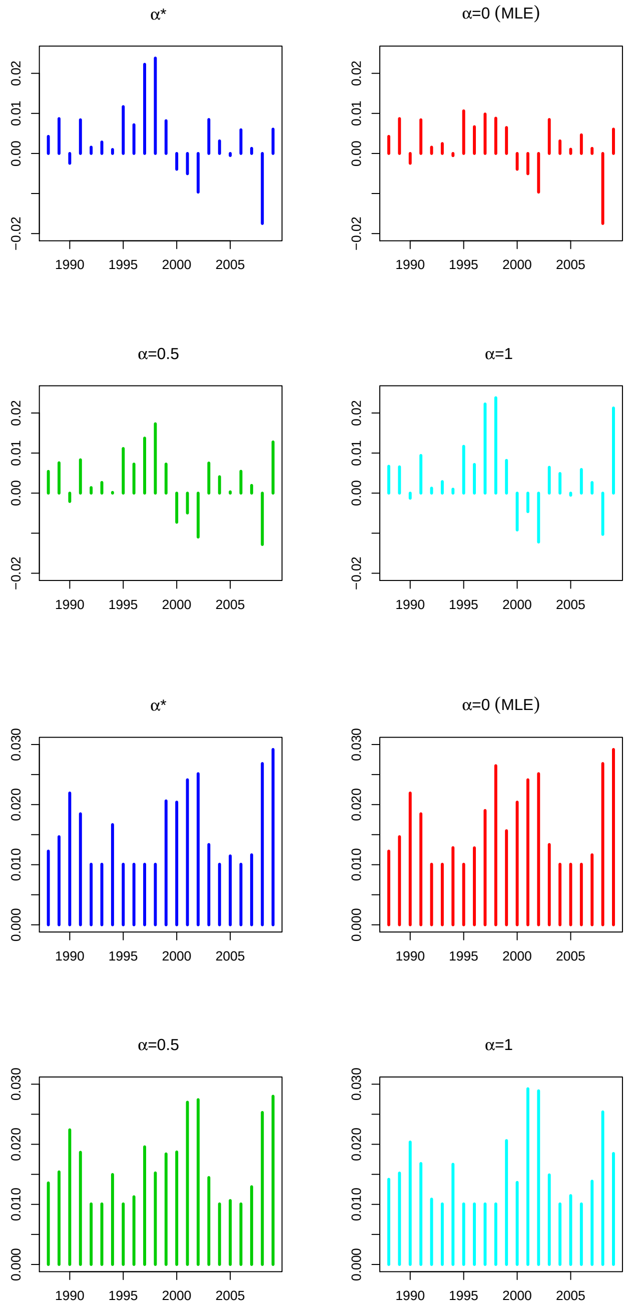

Figure 4. Annual estimates of expected return and volatility

Acknowledgements

This paper is supported by the Sectorial Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number SOP HRD/89/1.5/S/62988.

REFERENCES

[1] **Bauwens, L.; Laurent, S.; Rombouts, J.V.K. (2006),** ***Multivariate Garch Models: a Survey***, Journal of Applied Econometrics 21, 79–109;

[2] **Bouzebda, S.; Keziou, A. (2010),** ***New Estimates and Tests of Independence in Semiparametric Copula Models***, Kybernetica 46, 178–201;

[3] **Bouzebda, S.; Keziou, A. (2010),** ***A New Test Procedure of Independence in Copula Models via*** χ^2***-Divergence***, Commun. Statist. Theory Methods 39, 1–20;

[4] **Broniatowski, M.; Vajda, I. (2009),** ***Several Applications of Divergence Criteria in Continuous Families***, arXiv:0911.0937v1;

[5] **Broniatowski, M.; Toma, A.; Vajda, I. (2011),** ***Decomposable Pseudodistances and Applications in Statistical Estimation***, arXiv:1104.1541v1;

[6] **Ferrari, D.; Yang, Y. (2010),** ***Maximum LQ-Likelihood Estimation***, The Annals of Statistics 38, 753–783;

[7] **Hampel, F.; Ronchetti, E.; Rousseeuw, P.J.; Stahel, W. (1986),** ***Robust Statistics: the Approach Based on Influence Functions***, Wiley;

[8] **Karagrigoriou, A.; Mattheou, K. (2010)** ***On Distributional Properties and Goodness-of-Fit Tests for Generalized Measures of Divergence***, Comm. Statist. Theory Methods 39, 472–482;

[9] **Mattheou, K.; Karagrigoriou, A. (2010)** ***A New Family of Divergence Measures for Tests of Fit***, Aust. N. Z. J. Statist. 52, 187–200;

[10] **Pardo, L. (2006),** ***Statistical Inference Based on Divergence Measures***, Chapmann & Hall;

[11] **Purcaru, I.; Beganu, G.; Verboncu, A.I. (2009),** ***Interdependence and Diversity Measures for a Joint Ecosystem with Applications in System Management***, Economic Computation and Economic Cybernetics Studies and Research 1, 91–100;

[12] **Purcaru, I.; Verboncu, I. (2010),** ***Considerations on the Information and Diversity Measures***, Economic Computation and Economic Cybernetics Studies and Research 2, 5–18;

[13] **Purcaru, I. (2011),** ***Diversity and Solving a Three-Dimensional Allocation Problem***, Economic Computation and Economic Cybernetics Studies and Research 3, 59–70;

[14] **Toma, A. (2007)** ***Minimum Hellinger Distance Estimations for Some Multivariate Distributions: Influence Functions and Breakdown Point Results***, C.R. Acad. Sci. Paris, Ser. I 345, 353–358;
[15] **Toma, A. (2009),** ***Optimal Robust M-estimators using Divergences***, Statistics & Probability Letters 79, 1–5;
[16] **Toma, A.; Leoni-Aubin, S. (2010),** ***Robust Tests based on Dual Divergence Estimators and Saddlepoint Approximations***, Journal of Multivariate Analysis 101, 1143–1155;
[17] **Toma, A.; Broniatowski, M. (2011),** ***Dual Divergence Estimators and Tests: Robustness Results***, Journal of Multivariate Analysis 102, 20–36;
[18] **van der Vaart, A.W. (1998),** ***Asymptotic Statistics***, Cambridge Series in Statistical and Probabilistic Mathematics;
[19] **Zhao, Z. (2008),** ***Parametric and Nonparametric Models and Methods in Financial Econometrics***, Statistical Surveys 2, 1–42.