Risk-parameter estimation in volatility models[1]

Christian Francq & Jean-Michel Zakoïan

CREST (CNRS) et Université Lille III (EQUIPPE)
BP 60149
59653 Villeneuve d'Ascq cedex

Résumé. Nous introduisons le concept de paramètre de risque dans un modèle de volatilité conditionnelle de la forme $\epsilon_t = \sigma_t(\theta_0)\eta_t$ et nous proposons plusieurs estimateurs de ce paramètre. Pour une mesure de risque donnée r, le paramètre de risque conditionnel est une fonction du coefficient θ_0 de la volatilité et du risque $r(\eta_t)$ du processus des innovations. Une méthode en deux étapes permet d'estimer successivement ces quantités. Nous proposons également une méthode alternative en une seule étape, qui est fondée sur une reparamétrisation du modèle et sur l'utilisation d'un estimateur du quasi-maximum de vraisemblance non gaussien. La théorie asymptotique, établie pour des mesures de risque générales, comprenant notamment la valeur à risque (VaR) ou l'ES (Expected Shortfall), permet de quantifier le risque d'estimation. Dans le cas GARCH standard, la comparaison asymptotique montre la supériorité de la méthode en une étape quand les innovations sont à queues lourdes. Les résultats théoriques sont illustrés sur des séries de rendements d'indices boursiers.

Mots-clés. Expected Shortfall, QMV non gaussien, Regression quantile, Risque d'estimation, Valeur à Risque.

Abstract. This paper introduces the concept of risk parameter in conditional volatility models of the form $\epsilon_t = \sigma_t(\theta_0)\eta_t$ and develops statistical procedures to estimate this parameter. For a given risk measure r, the risk parameter is expressed as a function of the volatility coefficients θ_0 and the risk, $r(\eta_t)$, of the innovation process. A two-step method can be used to successively estimate these quantities. An alternative one-step approach, relying on a reparameterization of the model and the use of a non Gaussian QML, is proposed. The asymptotic results, established for general risk measures, the Value-at-Risk (VaR) and the Expected Shortfall (ES), allow to quantify the estimation risk in conditional risk measurement. Asymptotic comparisons, in the case of standard GARCH models, suggest a superiority of the one-step method when the innovations are heavy-tailed. An empirical illustration for stock market indices is proposed.

Keywords. Estimation risk, Expected Shortfall, Non Gaussian QML, Quantile Régression, Value-at-Risk.

[1]Ce travail bénéficie du soutien de l'Agence Nationale de la Recherche (ANR) (Projet ECONOM&RISK, ANR 2010 blanc 1804 03)

1 Risk parameter in volatility models

Most conditional volatility models are of the form

$$\begin{cases} \epsilon_t = \sigma_t \eta_t \\ \sigma_t = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \theta_0) \end{cases} \tag{1}$$

where (η_t) is a sequence of iid random variables, η_t being independent of $\{\epsilon_u, u < t\}$, $\theta_0 \in \mathbb{R}^m$ is a parameter belonging to a parameter space Θ, and $\sigma : \mathbb{R}^\infty \times \Theta \to (0, \infty)$. When $E\eta_t = 0$ and $E\eta_t^2 = 1$, the variable σ_t^2 is generally referred to as the volatility of ϵ_t. A leading model, the most widely used among practitioners, is the GARCH(1,1) model defined by

$$\sigma_t^2 = \omega_0 + \alpha_0 \epsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2, \tag{2}$$

where $\theta_0 = (\omega_0, \alpha_0, \beta_0)' \in (0, \infty) \times [0, \infty) \times [0, 1)$. For this model we have $\sigma_t^2 = \sum_{i=1}^\infty \beta_0^{i-1}(\omega_0 + \alpha_0 \epsilon_{t-i}^2)$, which is of the form (1).

It is assumed throughout that there exists a function H such that for any $\theta \in \Theta$, for any $K > 0$, and any sequence $(x_i)_i$

$$K\sigma(x_1, x_2, \dots; \theta) = \sigma(x_1, x_2, \dots; \theta^*), \quad \text{where } \theta^* = H(\theta, K).$$

Most conditional volatility models are such that for $K \geq 1$, $\theta^* \geq \theta$ componentwise. For instance, in the GARCH(1,1) case we have $\theta^* = (K^2\omega, K^2\alpha, \beta)'$ with standard notation. The parameter θ_0 can thus be interpreted as a *volatility parameter* in the sense that the larger θ_0 the larger the volatility.

Now we define the notion of *conditional risk parameter*. Let r denote a risk measure, that is, a mapping from the set of the real random variables to $\mathbb{R}$. Assume that r is nonnegative, positively homogenous[2] and law-invariant [3]. Then the risk of ϵ_t conditional on $\{\epsilon_u, u < t\}$ is given by

$$r_{t-1}(\epsilon_t) = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \theta_0) r(\eta_t). \tag{3}$$

Now, assuming $r(\eta_t) \neq 0$, let $\eta_t^* = \eta_t / r(\eta_t)$ and let $\theta_0^* = H(\theta_0, r(\eta_t))$. The model can be reparameterized as

$$\begin{cases} \epsilon_t = \sigma_t^* \eta_t^*, \qquad r(\eta_t^*) = 1, \\ \sigma_t^* = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \theta_0^*). \end{cases}$$

Because the conditional risk of ϵ_t is now simply

$$r_{t-1}(\epsilon_t) = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \theta_0^*),$$

θ_0^* will be called the conditional risk parameter.

[2]that is, $r(\lambda X) = \lambda r(X)$ for any risk variable X and any $\lambda > 0$.

[3]that is, the risk $r(X)$ of any risk variable X only depends on the distribution of X.

Example 1 (VaR parameter) An important example is the VaR, which is the most standard risk measure used in the current regulations. For a continuous risk variable X with quantile function F_X^{-1}, the VaR at level α, with $\alpha \in (0,1)$, is given by $r(X) = -F_X^{-1}(\alpha)$. The *conditional* VaR of the process (ϵ_t) at risk level $\alpha \in (0,1)$, denoted by $\mathrm{VaR}_t(\alpha)$, is defined by

$$P_{t-1}[\epsilon_t < -\mathrm{VaR}_t(\alpha)] = \alpha,$$

where P_{t-1} denotes the historical distribution conditional on $\{\epsilon_u, u < t\}$. When (ϵ_t) satisfies (1), the theoretical VaR is then given by

$$\mathrm{VaR}_t(\alpha) = -\sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \theta_0) F_\eta^{-1}(\alpha)$$

where F_η is the probability distribution function of η_t. Let α be small enough so that $F_\eta^{-1}(\alpha) < 0$. Thus (3) holds with $r_{t-1}(\epsilon_t) = \mathrm{VaR}_t(\alpha)$ and $r(\eta_t) = -F_\eta^{-1}(\alpha)$. Now suppose that the volatility model is the GARCH(1,1) model (2). Then the risk parameter at level α is given by $\theta_0^* = (K^2\omega_0, K^2\alpha_0, \beta_0)'$ with $K = -F_\eta^{-1}(\alpha)$. This coefficient takes into account the dynamics of the GARCH process through the volatility parameters, but also the lower tail of the innovations distribution.

Example 2 (Expected Shortfall parameter) Another popular measure of financial risk is the expected shortfall (ES). One reason for its attractiveness is that, in contrast to the VaR, the ES satisfies the sub-additivity property. For a continuous risk variable X such that $E(X^-) < \infty$, the ES at level $\alpha \in (0,1)$ is given by $r(X) = -E[X \mid X \leq F_X^{-1}(\alpha)]$. The *conditional* ES of the process (ϵ_t) at risk level α, denoted by $\mathrm{ES}_t(\alpha)$, is defined by

$$\mathrm{ES}_t(\alpha) = -E_{t-1}[\epsilon_t \mid \epsilon_t < -\mathrm{VaR}_t(\alpha)],$$

where E_{t-1} denotes the expectation conditional on $\{\epsilon_u, u < t\}$. When (ϵ_t) satisfies (1), the theoretical ES is then given by

$$\mathrm{ES}_t(\alpha) = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \theta_0)\mathrm{ES}_\eta(\alpha),$$

where $\mathrm{ES}_\eta(\alpha)$ is the ES at level α of η_t, which is of the form (3).

Example 3 (VaR and ES parameters for two GARCH(1,1)) For the sake of illustration, consider two GARCH(1,1) models with, respectively, standard Gaussian and standardized Student(4) innovations. The volatility parameter, as displayed in Table 1, is larger for the Gaussian-innovation model than for the Student-innovation model. In contrast, the VaR parameter at level 1% is slightly larger for the second model. In other words, the first model is more volatile but less risky than the second one for the VaR at 1%. The difference between the two models is even more pronounced when ES-parameters at the level 1% are considered. In particular, the coefficient α_0^*, measuring the impact of a large squared return on the risk of the next period, is 1.5 larger in the model with student errors than in the conditionally Gaussian model.

Table 1: VaR and ES parameters at the 1% level for GARCH(1,1) models

Errors distribution	$\eta_t \sim \mathcal{N}(0,1)$	$\eta_t \sim \frac{1}{\sqrt{2}}St_4$
Volatility parameter	$(1, 0.05, 0.9)$	$(1, 0.04, 0.9)$
VaR parameter	$(5.41, 0.27, 0.9)$	$(7.01, 0.28, 0.9)$
ES parameter	$(7.10, 0.36, 0.9)$	$(13.63, 0.55, 0.9)$

2 Estimating the conditional risk parameter

We developed an estimation procedure for a general conditional risk measure. For simplicity, we concentrate here on the VaR.

The VaR parameter of a GARCH(1,1) has been defined in Example 1. To extend the definition of the conditional VaR at level α, we first need to reparameterize Model (1). If α is not too large (more precisely $\alpha < P(\eta_0 > 0)$), from $P[\eta_t < F^{-1}(\alpha)] = \alpha$ we deduce $P[\eta_t^* < -1] = \alpha$ where $\eta_t^* = -\eta_t/F^{-1}(\alpha)$. Letting $\theta_0^* = H(\theta_0, -F^{-1}(\alpha))$, the model can be reparameterized as

$$\begin{cases} \epsilon_t = \sigma_t^* \eta_t^*, \quad P[\eta_t^* < -1] = \alpha, \\ \sigma_t^* = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \theta_0^*). \end{cases} \tag{4}$$

The theoretical VaR is now given by

$$\mathrm{VaR}_t(\alpha) = \sigma_t(\theta_0^*), \tag{5}$$

where θ_0^* is the *VaR parameter* at level α.

Two estimators of that parameter can be considered.

2.1 Two-step VaR estimation

First consider the usual approach for estimating the VaR in Model (1) under the identifiability condition

$$E\eta_t^2 = 1. \tag{6}$$

This approach involves two steps. In a first step, the model is estimated by the standard QMLE and, in a second step, the theoretical quantile $\xi_\alpha := F_\eta^{-1}(\alpha)$ is estimated using the estimated rescaled innovations. More precisely, let $\hat{\theta}_n$ denote the Gaussian QMLE of θ_0 in Model (1) under the constraint (6), let

$$\hat{\eta}_t = \frac{\epsilon_t}{\tilde{\sigma}_t(\hat{\theta}_n)},$$

and let $\xi_{n,\alpha}$ denote the empirical α-quantile of $\hat{\eta}_1, \dots, \hat{\eta}_n$.

When $\xi_{n,\alpha} < 0$, an estimator of the VaR at level α is then given by

$$\widetilde{\mathrm{VaR}}_t(\alpha) = -\tilde{\sigma}_t(\hat{\theta}_n)\xi_{n,\alpha} = \tilde{\sigma}_t\{H(\hat{\theta}_n, -\xi_{n,\alpha})\}.$$

If the distribution of η_t is assumed to be symmetric, another estimator of $\mathrm{VaR}_t(\alpha)$ is

$$\widetilde{\widetilde{\mathrm{VaR}}}_t(\alpha) = \tilde{\sigma}_t(\hat{\theta}_n)\tilde{\xi}_{n,1-2\alpha} = \tilde{\sigma}_t\{H(\hat{\theta}_n, \tilde{\xi}_{n,1-2\alpha})\} \tag{7}$$

where $\tilde{\xi}_{n,1-2\alpha}$ is the empirical $(1-2\alpha)$-quantile of $|\hat{\eta}_1|, \dots, |\hat{\eta}_n|$.

2.2 One-step VaR estimation

Define a QMLE of θ_0^* by

$$\hat{\theta}_n^* = \arg\max_{\theta\in\Theta} \sum_{t=1}^{n} \log \frac{1}{\tilde{\sigma}_t(\theta)} h_\alpha\left(\frac{\epsilon_t}{\tilde{\sigma}_t(\theta)}\right) \tag{8}$$

where

$$h(x) = h_\alpha(x) = \lambda\alpha(1-2\alpha)|x|^{2\lambda\alpha-1}\{|x|^{-\lambda}\mathbf{1}_{\{|x|>1\}} + \mathbf{1}_{\{|x|\le 1\}}\} \tag{9}$$

for some positive constant λ (which does not matter). It can be shown that, when the distribution of η_t is symmetric and under some regularity conditions, a one-step consistent estimator of the VaR parameter, not requiring any estimation of the quantile function of the innovations η_t, is given by

$$\widehat{\mathrm{VaR}}_t(\alpha) = \tilde{\sigma}_t(\hat{\theta}_n^*). \tag{10}$$

The non-Gaussian QML estimator defined by (8)-(9) is also related to estimators introduced in the quantile regression literature. Letting $\rho_\alpha(u) = u(\alpha - \mathbf{1}_{\{u\le 0\}})$, it can be seen that

$$\hat{\theta}_n^* \quad = \quad \arg\min_{\theta\in\Theta} \frac{1}{n}\sum_{t=1}^{n} \rho_{1-2\alpha}\left\{\log\left(\frac{|\epsilon_t|}{\tilde{\sigma}_t(\theta)}\right)\right\}.$$

To interpret this expression, note that the first equation in Model (4) can be equivalently written as

$$\log|\epsilon_t| = \log\sigma_t^* + \log|\eta_t^*|, \qquad P[\log|\eta_0^*| < 0] = 1 - 2\alpha \tag{11}$$

under the assumption of a symmetric distribution for η_0^*. Model (11) resembling a quantile regression model, it is not surprising to get an estimator of the form (11). An important difference with the quantile regression or autoregression, however, is that $\tilde{\sigma}_t(\theta)$ is not assumed to be a linear combination of explanatory variables, or past observables.

Under some assumptions, and assuming that the density f^* of η_0^* is continuous at 1 and satisfies $f^*(1) > 0$ and $M = \sup_{x\in\mathbb{R}} |x| f^*(x) < \infty$, there exists a sequence of local minimizers $\hat{\theta}_n^*$ of the criterion defined in (11) satisfying

$$\sqrt{n}(\hat{\theta}_n^* - \theta_0^*) \xrightarrow{d} \mathcal{N}\left(0, \Xi_\alpha := \frac{2\alpha(1-2\alpha)}{4f^{*2}(1)} J_\alpha^{-1}\right),$$

where $J_\alpha = ED_t(\theta_0^*)D_t'(\theta_0^*)$ and $D_t(\theta) = \sigma_t^{-1}(\theta)\partial\sigma_t(\theta)/\partial\theta$.

Let $\widehat{\Xi}_\alpha$ denote a consistent estimator of the asymptotic variance Ξ_α. The delta method thus suggests a $(1-\alpha_0)\%$ confidence interval for $\mathrm{VaR}_t(\alpha)$ whose bounds are

$$\tilde{\sigma}_t(\hat{\theta}_n^*) \pm \frac{\Phi^{-1}_{1-\alpha_0/2}}{\sqrt{n}} \left\{ \frac{\partial\tilde{\sigma}_t(\hat{\theta}_n^*)}{\partial\theta'} \widehat{\Xi}_\alpha \frac{\partial\tilde{\sigma}_t(\hat{\theta}_n^*)}{\partial\theta} \right\}^{1/2}, \tag{12}$$

where $\Phi^{-1}_{\alpha_0}$ denotes the α_0-quantile of the standard gaussian distribution. Drawing such confidence intervals allows to underline that the VaR evaluation is subject to estimation risk. Even when the model is correctly specified, the market risk, as measured by the theoretical VaR defined by (5), is not known with exactness, but is likely to belong to the confidence interval (12).

2.3 Comparison of the one-step and two-step estimators

A comparison of the VaR estimators $\widetilde{\widehat{\mathrm{VaR}}}_t(\alpha)$ and $\widehat{\mathrm{VaR}}_t(\alpha)$ defined in (7) and (10) can then be based on the asymptotic accuracies of the estimators $\hat{\theta}_n^*$ and $H(\hat{\theta}_n, \tilde{\xi}_{n,1-2\alpha})$ of θ_0^*. For standard GARCH models with a symmetric density f for η_t, we have

$$\mathrm{Var}_{as}\{\sqrt{n}(\hat{\theta}_n^* - \theta_{0,\alpha})\} \preceq \mathrm{Var}_{as}\left\{\sqrt{n}\left(H(\hat{\theta}_n, \tilde{\xi}_{n,1-2\alpha}) - \theta_{0,\alpha}\right)\right\} \quad \text{iff} \quad \Delta_\alpha \leq 0.$$

where

$$\Delta_\alpha = \frac{2\alpha(1-2\alpha)}{\xi_\alpha^2 f^2(\xi_\alpha)} - (E\eta_t^4 - 1).$$

Interestingly, comparing the asymptotic variance matrices of the estimators amounts to determining the sign of a real coefficient, which solely depends on the distribution of η_t. None of the methods is superior in every situation. If the fourth-order moment is large, *a fortiori* if it does not exist, the one-step estimator will be better. Conversely, for distributions admitting moments at any order (such as the Gaussian) the two-step estimator may be superior.

Reference

[1] Francq, C. and Zakoïan, J-M. (2012), Risk-parameter estimation in volatility models, *MPRA Paper No. 41713.*