Rapporto n. 199

**Interval estimation for the Sharpe Ratio when returns are not i.i.d. with special emphasis on the GARC(1,1) process with symmetric innovations**

Lucio De Capitani

Novembre 2010

**Dipartimento di Metodi Quantitativi per le Scienze Economiche ed Aziendali**
Università degli Studi di Milano Bicocca
Via Bicocca degli Arcimboldi 8 - 20126 Milano - Italia
Tel +39/02/64483102/3 - Fax +39/2/64483105
Segreteria di redazione: Andrea Bertolini

# Interval estimation for the Sharpe Ratio when returns are not i.i.d. with special emphasis on the GARCH(1,1) process with symmetric innovations

**Lucio De Capitani**

**Abstract** In this paper, assuming that returns follows a stationary and ergodic stochastic process, the asymptotic distribution of the natural estimator of the Sharpe Ratio is explicitly given. This distribution is used in order to define an approximated confidence interval for the Sharpe ratio. Particular attention is devoted to the case of the GARCH(1,1) process. In this latter case, a simulation study is performed in order to evaluate the minimum sample size for reaching a good coverage accuracy of the asymptotic confidence intervals.

**Keywords** Sharpe ratio · stationary and ergodic process · generalized method of moments · GARCH(1,1) process

## 1 Introduction

The Sharpe Ratio (Sharpe [17], [18], [19]) is probably the best known and applied financial performance measure. It is used to measure the risk-adjusted performance of a financial asset and to compare different portfolios of financial activities. The Sharpe Ratio is defined using the standard deviation of returns (interpreted as a risk measure) and the expected excess return (interpreted as reward measure) to determine the reward per unit of risk. In more detail, let $X$ be the random variable describing the log-returns of a risky financial activity ad let $\xi$ be the (log) *risk-free rate* of return. Let $\mu$ and $\sigma$ be the expected value and the standard deviation of $X$, respectively. The Sharpe Ratio is defined as:

$$\psi = \frac{\mu - \xi}{\sigma} \quad . \tag{1}$$

The Sharpe Ratio of a particular financial activity is usually estimated using a time series of log-returns[1]. In particular, let $X_1, ..., X_n$ be a time series of returns and let $\bar{X}$

L. De Capitani
Department of Quantitative Methods for Business and Economic Sciences. - University of Milan-Bicocca. - Piazza dell'Ateneo Nuovo, 1, 20126 MILANO
Tel.: +39 0264483186
Fax: +39 0264483105
E-mail: l.decapitani@campus.unimib.it

[1] To ease the exposition, in the following we will refer to the log-returns simply as returns.

and $S^2$ be the sample mean and the unbiased sample variance, respectively. Usually, in order to estimate $\psi$, the natural estimator $\widehat{\Psi} = \frac{\bar{X}-\xi}{S}$ is used. In literature, there are several results concerning the features of $\widehat{\Psi}$. For example, Miller and Gher [13] studied its bias in the case of independent and identically normally distributed returns (i.i.d.-normal returns). Later Jobson and Korkie [10] derived the asymptotic distribution of $\widehat{\Psi}$, under the assumption of i.i.d.-normal returns:

$$\sqrt{n}(\widehat{\Psi} - \psi) \overset{a}{\sim} \mathcal{N}\left(0; 1 + \frac{1}{2}\psi^2\right) \quad . \tag{2}$$

The distribution (2) was used by Jobson and Korkie [10] to define asymptotic critical regions for testing hypotheses on $\psi$ and it can be used to build, by analytical inversion, an approximated confidence interval for $\psi$. In De Capitani and Zenga [7], it is shown that the coverage accuracy of the aforementioned asymptotic confidence interval is good stating from the relatively small sample size of 50.

However, the results concerning the i.i.d.-normal returns are of little practical interest because empirical evidence shows that the stock market returns cannot be considered normally distributed and serially independent. This observation motivates the seminal works of Lo (2002) who studied the asymptotic behavior of $\widehat{\Psi}$ firstly removing the normality assumption and secondly removing the independence hypotheses. Under the i.i.d. hypotheses he obtained that

$$\sqrt{n}(\widehat{\Psi} - \psi) \xrightarrow{d} \mathcal{N}(0; V) \tag{3}$$

where

$$V = 1 + \frac{1}{2}\psi^2 - \gamma_1\psi + \frac{\gamma_2}{4}\psi^2$$

and $\gamma_1$ and $\gamma_2$ stand for the third standardized central moment (usually interpreted as an indicator of skewness) and for the fourth standardized central moment minus 3 (commonly interpreted as a kurtosis indicator), respectively:

$$\gamma_1 = \frac{\mu_3}{\sigma^3} \qquad \text{and} \qquad \gamma_2 = \left(\frac{\mu_4}{\sigma^4} - 3\right) , \qquad \mu_k = E[(X-\mu)^k] \ .$$

The distribution (3) can be used to define an asymptotic confidence interval for $\Psi$. In De Capitani and Zenga [7] the coverage accuracy of this last confidence interval is studied. In detail, simulation shows that in finite samples the actual coverage of the confidence interval for $\psi$ is lower than the nominal confidence level. Furthermore, it is shown that the actual coverage depends strongly on the tails of the distribution of $X$ (the fatter the tails the worst is the coverage accuracy) and on the true value of the ratio (the higher the value of $\psi$, the lower is the simulated coverage). As a consequence, the minimum sample size necessary to assure a good coverage accuracy of the confidence interval based on (3) increases from 50 (as in the normal case) to 400 and, when the $\psi$ is high and the tail of the distribution on $X$ are fat, 400 is no longer sufficient[2].

In order to determine the asymptotic distribution of $\widehat{\Psi}$ in the general context of not-i.i.d. returns, Lo [12] observed that $\bar{X}$ and $S^2$ are GMM estimators of $\mu$ and $\sigma^2$, respectively. Assuming that all the regularity condition necessary to assure the

[2] The impact of the tails and asymmetry of the distribution of $X$ on the coverage accuracy are discussed in details in De Capitani and Zenga [7].

asymptotic properties of the GMM estimators hold, it is possible to conclude that (see, e.g., Hall, [8] or Hansen, [9]):

$$\sqrt{n}\begin{bmatrix}\bar{X}-\mu\\ S^2-\sigma^2\end{bmatrix}\overset{d}{\to}\mathbf{NB}\left(\mathbf{0},\mathbf{\Sigma}_{GMM}\right) \tag{4}$$

where **NB** stands for "Bivariate Normal". Starting from this last result and applying the Delta Method (see Sen and Singer, [16]), Lo [12] concluded that

$$\sqrt{n}(\widehat{\Psi}-\psi)\overset{d}{\to}\mathcal{N}\left(0;V_{GMM}\right) \quad .$$

However, in Lo [12] the expression of $\mathbf{\Sigma}_{GMM}$ and $V_{GMM}$ are not explicitly given. In the aforementioned paper it is only observed that, for practical purposes the asymptotic variance-covariance matrix $\mathbf{\Sigma}_{GMM}$, and then the variance $V_{GMM}$, can be consistently estimated using the well known Newey-West estimator (see Newey and West, [14]). Christie [5] tries to bridge this latter gap but its approach was incorrect. In detail, he explicitly derived the asymptotic distribution of $\widehat{\Psi}$ using the asymptotic theory of the GMM estimator in the context of i.i.d. observation interpreting the results as they were obtained through the general asymptotic theory of the GMM estimators. Furthermore, the asymptotic variance $V_{GMM}$ he obtained was not written parsimoniously. Later, Opdike [15] rationalized the expression given in Christie [5] finding, not surprisingly, that $V_{GMM}$ is the same as $V$. However, Opdicke (2007) interpreted this equivalence as an interesting result rather than pointing out the error of Christie [5].

In order to avoid further confusion, in this work we re-examine the problem of determining the asymptotic distribution of $\widehat{\Psi}$ in the general case of not-i.i.d. returns. In detail, the paper is organized as follows. In Section 2 the GMM and the general assumptions assuring the validity of the asymptotic properties of GMM estimators are recalled. In Section 3 the general result concerning the large sample properties of the GMM estimators is particularized in order to determine the asymptotic joint distribution of $\bar{X}$ and $S^2$. We think this analysis is useful in order to highlight the real assumptions to be made in order to assure the asymptotic normality of $\widehat{\Psi}$. Further, an asymptotic confidence interval for $\psi$ is introduced. In Section 4 the results of Section 3 are specialized to the case of the GARCH(1,1) process with symmetric innovations. In Section 5 we describe the simulation study performed in order to asses the coverage accuracy of the large sample confidence interval for $\psi$. Moreover, in this section we discuss the computational details. In Section 6 the results of the simulations are given and discussed. Finally, Section 7 is devoted to conclusions.

## 2 Recalling the GMM

First of all, we briefly recall the Generalized Method of Moments (GMM) and the large sample properties of the GMM estimators.

The GMM was introduced by Hansen [9] and, nowadays, it is commonly used, especially by the econometricians. As shown in Hall [8], the GMM is substantially a generalization of the minimum chi-square estimation method and many other classical estimation procedures (such as the method of moments, ordinary least square, maximum likelihood and instrumental variables) fall into the GMM framework. A great advantage of using the GMM is that the estimators obtained through this method are, under certain regularity conditions, consistent and asymptotically normally distributed. Further, these properties hold not only when the random variables used to define the estimators are i.i.d. but also when they come from a strictly stationary and ergodic stochastic process. Here we recall the definition of the GMM estimators and we report a theorem stating their large sample properties. In doing that, we follows the formalization of Hall [8] considering only the case of univariate stochastic process and constant weighting matrix.

Let $\{X_t\}_{t\in\mathbb{N}}$ be a strictly stationary discrete time stochastic process with state space $V \subseteq \mathbb{R}$ and let $\theta \in \Theta$ be the $(k \times 1)$ vector of parameters to be estimated. The true value of $\theta$ is denoted by $\theta_0$. Consider the vector-valued function $\mathbf{f} : V \times \Theta \to \mathbb{R}^q$ and suppose that $E[\mathbf{f}(X_t, \theta)] = \mathbf{0}$ if and only if $\theta = \theta_0$. Roughly speaking, the identity $E[\mathbf{f}(X_t, \theta_0)] = \mathbf{0}$ represent the set of constraints, commonly referred to as "*orthogonality conditions*" or "*population moment conditions*", useful to perform the estimation procedure. In detail, let $(X_1, ..., X_n)$ be a time series of length $n$ from the stochastic process $\{X_t\}_{t\in\mathbb{N}}$. On the basis of the random vector $(X_1, ..., X_n)$, the sample average of $\mathbf{f}$ can be introduced:

$$\mathbf{g}_n(\theta) = \frac{1}{n}\sum_{t=1}^{n} \mathbf{f}(X_t, \theta) \quad .$$

The GMM estimator $\hat{\theta}$ of $\theta_0$ is defined as $\hat{\theta} = \arg\min_\theta Q_n(\theta)$ where $Q_n(\theta) = \mathbf{g}_n(\theta)'\mathbf{W}\mathbf{g}_n(\theta)$ and $\mathbf{W}$ is a symmetric positive definite $(q \times q)$ matrix. For a given realization $(x_1, ..., x_n)$ of the random vector $(X_1, ..., X_n)$, the quadratic form $Q_n(\theta)$ coincide with a non-standard euclidean distance between the vector $\mathbf{g}_n(\theta)$ and the origin $\mathbf{0}$. Then, the GMM estimator $\hat{\theta}$ is defined so as to make the sample average $\mathbf{g}_n(\theta)$ as close as possible to the theoretical value of $\mathbf{0}$ prescribed by the population moment conditions.

The GMM estimators possess the desirable large sample properties formalized in the following theorem.

**Theorem 1** *Assume that the following conditions hold:*

1. *the process $\{X_t\}_{t\in\mathbb{N}}$ is ergodic*[3]*;*
2. *$\Theta$ is a compact set and $\theta_0$ is an interior point of $\Theta$;*
3. *the vector-function $\mathbf{f}$ satisfy the following regularity conditions: (i) $\mathbf{f}$ is continuous on $\Theta$ for all $x_t \in V$; (ii) the derivatives matrix $\partial\mathbf{f}(x_t, \theta)/\partial\theta'$ exists and is continuous on $\Theta$ for each $x_t \in V$;*
4. *the vector-function $\mathbf{f}$ and the random variable $X_t$ satisfy: (i) $E[\mathbf{f}(X_t, \theta)]$ exists finite for every $\theta \in \Theta$ and is continuous on $\Theta$; (ii) $E[\partial\mathbf{f}(X_t, \theta)/\partial\theta']$ exists finite and is continuous in some neighborhood of radius $\epsilon$, say $N_\epsilon$, of $\theta_0$; (iii) $E[sup_{\theta\in\Theta}\mathbf{f}(X_t, \theta)'\mathbf{f}(X_t, \theta)] < \infty$; (iv) $E[\mathbf{f}(X_t, \theta)\mathbf{f}(X_t, \theta)'] < \infty$ for all $\theta \in \Theta$;*
5. *the matrix $G_n(\theta) = n^{-1}\sum_{i=1}^{n} \partial f(X_t, \theta)/\partial\theta'$ converge uniformly to $E[\partial\mathbf{f}(X_t, \theta)/\partial\theta']$ in a neighborhood $N_\epsilon$ of $\theta_0$. That is:*

$$\exists\, \epsilon > 0 \;\; : \;\; \sup_{\theta\in N_\epsilon} \left\| G_n(\theta) - E\left[\partial\mathbf{f}(\mathbf{X}_t; \theta)/\partial\theta'\right] \right\| \xrightarrow{p} 0$$

*where, for a given square matrix $\mathbf{A}$, $||\mathbf{A}|| = \sqrt{tr(\mathbf{A}'\mathbf{A})}$.*

[3] For a formal definition of ergodicity, see, e.g., Karlin and Taylor [11].

6. *the variance-covariance matrix* $\mathbf{S} = \lim_{n\to\infty} Var[\sqrt{n}\mathbf{g}_n(\theta_0)]$ *exists finite and definite positive;*

*Then, the GMM estimator* $\hat{\theta}$ *is consistent and asymptotically normally distributed:*

– $\hat{\theta} \xrightarrow{p} \theta$;
– $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{H})$ *where* $\mathbf{H} = (\mathbf{G}_0'\mathbf{W}\mathbf{G}_0)^{-1}\mathbf{G}_0'\mathbf{W}\mathbf{S}\mathbf{W}\mathbf{G}_0(\mathbf{G}_0'\mathbf{W}\mathbf{G}_0)^{-1}$ *and* $\mathbf{G}_0 = E[\partial\mathbf{f}(X_t, \theta_0)/\partial\theta']$.

## 3 Asymptotic distribution of $\widehat{\Psi}$ and confidence interval for $\psi$ when returns are not i.i.d.

In this section, following the trace of Lo [12], we specialize theorem 1 to the context of the estimation of the vector of parameter $\theta = [\mu, \sigma^2]'$ where $\mu = E[X_t]$ and $\sigma^2 = Var[X_t]$ for all $t \in \mathbb{N}$. From now on, we interpret $\{X_t\}_{t\in\mathbb{N}}$ as the process of the returns of a financial activity and we assume that $\mathbb{R}$ is the state space of the process. Further, we suppose that $\{X_t\}_{t\in\mathbb{N}}$ is strictly stationary, ergodic and

$$E[X_t^r] < \infty \qquad \text{for} \qquad r = 1, 2, 3, 4\ . \tag{5}$$

In order to estimate $\theta$ it is natural to choose the method of moments estimator $\hat{\theta} = [\bar{X}, S_b^2]$ where $S_b^2$ denote the (biased) sample variance. The estimator $\hat{\theta}$ can be seen as the GMM estimator of $\theta$ defined by the population moment conditions

$$E\begin{bmatrix} X_t - \mu_0 \\ (X_t - \mu_0)^2 - \sigma_0^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

induced by the function

$$\begin{aligned} &\mathbf{f} : \mathbb{R} \times \Theta \to \mathbb{R}^2 \\ &(x; \mu, \sigma^2) \mapsto \begin{bmatrix} x - \mu \\ (x - \mu)^2 - \sigma^2 \end{bmatrix} \end{aligned}$$

and choosing $\mathbf{W}$ as the $2 \times 2$ identity matrix.

Now, we analyze the assumptions of Theorem 1 in the context just introduced. First of all, observe that assumption *1.* is satisfied thanks to the hypothesis made at the beginning of this section.

Assumption *2.* require some additional observations. In detail, the most general parametric space for $\theta$ is $\mathbb{R} \times \mathbb{R}^+$. This last parametric space is not compact because it is not limited. However, it is reasonable to assume for $\mu$ and $\sigma$ a lower and an upper bound. In particular, we can assume that $\mu \in [-\mu^*, \mu^*]$ where $\mu^*$ is large enough to cover all the value that an expected return of a financial activity can likely assume. Analogously, we can suppose that $\sigma^2 \in [\sigma_-^2, \sigma_+^2]$ where $\sigma_-^2 > 0$ and $\sigma_+^2$ are, respectively, small and large enough to include all the reasonable values for the standard deviation of returns. Then, in the following, we assume that the parametric space is given by $\Theta = [-\mu^*, \mu^*] \times [\sigma_-^2, \sigma_+^2]$, which is compact, and we assume that $(\mu_0, \sigma_0^2) \in [-\mu^*, \mu^*] \times [\sigma_-^2, \sigma_+^2]$.

In regards to assumption *3.*, the function $\mathbf{f}$ is clearly continuous on $\Theta$ for all $x \in \mathbb{R}$. Further, the elements of the derivatives matrix

$$\frac{\partial\mathbf{f}(x, \theta)}{\partial\theta'} = \begin{bmatrix} -1 & 0 \\ -2(x - \mu) & -1 \end{bmatrix}$$

exist and are continuous on $\Theta$ for all $x \in \mathbb{R}$ as it is required.
Concerning assumption *4.*, we first observe that the elements of the matrices

$$E[\mathbf{f}(X_t,\theta)] = \begin{bmatrix} \mu_0 - \mu \\ \sigma_0^2 - \sigma^2 \end{bmatrix} \qquad \text{and} \qquad E\left[\frac{\partial \mathbf{f}(x,\theta)}{\partial \theta^T}\right] = \begin{bmatrix} -1 & 0 \\ \mu - \mu_0 & -1 \end{bmatrix}$$

are finite and continuous functions on $\Theta$. Then, condition *4.(i)* and *4.(ii)* are fulfilled. In order to verify condition *4.(iii)*, observe that

$$\mathbf{f}(x_t,\theta)'\mathbf{f}(X_t,\theta) = (x_t-\mu)^4 + (x_t-\mu)^2 - 2\sigma^2(x_t-\mu)^2 + \sigma^4 \quad .$$

Consequently

$$\sup_{\theta\in\Theta} \; \mathbf{f}(X_t,\theta)^T\mathbf{f}(X_t,\theta) \le M(x_t)^4 + M(x_t)^2 - 2\sigma_-^2 m(x_t)^2 + \sigma_+^4$$

where $M(x_t) = \max[|x_t+\mu^*|, |x_t-\mu^*|]$ and $m(x_t) = \min[|x_t+\mu^*|, |x_t-\mu^*|]$. Let $f_{X_t}$ denotes the density of $X_t$. Thanks to assumption (5), we have that

$$\begin{aligned} E[M(x_t)^h] &= \int_{-\infty}^{0} (x-\mu^*)^h f_{X_t}(x)\,dx + \int_{0}^{\infty} (x+\mu^*)^h f_{X_t}(x)\,dx \\ &< E[(X_t-\mu^*)^h] + E[(X_t+\mu^*)^h] < \infty \qquad h = 2,4 \quad . \end{aligned}$$

Further, $E[m(x_t)^2] < E[M(x_t)^2] < \infty$ and then condition *4.(iii)* follows. Finally, condition *4(iv)* is a direct consequence of (5).
Regarding condition *5.*, observe that

$$G_n(\theta) = \begin{bmatrix} -1 & 0 \\ -2(\bar{X}-\mu) & -1 \end{bmatrix}$$

and, consequently,

$$\left\| G_n(\theta) - E\left[\frac{\partial \mathbf{f}(X_t;\theta)}{\partial \theta^T}\right] \right\| = |2(\bar{X}-\mu_0)| \quad .$$

Condition *5.* now follows from the ergodicity of the process $\{X_t\}_{t\in\mathbb{N}}$ which assures that $\bar{X} \stackrel{qc}{\to} \mu_0$.
Finally, we analyze the condition *6.* First of all, we observe that $\mathbf{S}$ is the asymptotic variance-covariance matrix of the random variables

$$\sqrt{n}(\bar{X}-\mu_0) \qquad \text{and} \qquad \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu_0)^2 - \sigma_0^2\right) \quad .$$

It is well known (see, e.g., Brockwell and Davis, [4]) that

$$Var[\sqrt{n}(\bar{X}-\mu_0)] = n\,Var[\bar{X}] = \sigma_0^2\left[1 + 2\sum_{i=1}^{n}\left(1-\frac{i}{n}\right)\rho_i\right]$$

where $\rho_i$ denotes the correlation between $X_1$ and $X_{i+1}$.

The variance of the random variable

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu_0)^2 - \sigma_0^2\right) \quad .$$

in the case of not-i.i.d. observation can be derived as follows:

$$Var\left[\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu_0)^2-\sigma_0^2\right)\right]=\frac{1}{n}Var\left[\sum_{i=1}^{n}(X_i-\mu_0)^2\right]$$
$$=\frac{1}{n}\sum_{i=1}^{n}V\left[(X_i-\mu_0)^2\right]+\frac{2}{n}\sum_{i=1}^{n}\sum_{j=i+1}^{n}Cov\left[(X_i-\mu_0)^2;(X_j-\mu_0)^2\right] \quad . \qquad (6)$$

Thanks to the strict stationarity of $\{X_t\}_{t\in\mathbb{N}}$ we have that $Var\left[(X_i-\mu_0)^2\right]=(\mu_4-\sigma_0^4)$ for all $i\in\mathbb{N}$ and

$$Cov\left[(X_i-\mu_0)^2;(X_j-\mu_0)^2\right]=Cov\left[(X_{i+h}-\mu_0)^2;(X_{j+h}-\mu_0)^2\right]$$

for all $i,j\in\mathbb{N}$ and for all $h\in\mathbb{Z}$. Consequently

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=i+1}^{n}Cov\left[(X_i-\mu_0)^2;(X_j-\mu_0)^2\right] &= \sum_{i=1}^{n-1}(n-i)Cov\left[(X_1-\mu_0)^2;(X_{i+1}-\mu_0)^2\right]\\ &= \sum_{i=1}^{n-1}(n-i)\left(\mu_i^{2,2}-\sigma_0^4\right)\end{aligned}$$

where $\mu_i^{h,k}=E\left[(X_1-\mu_0)^h(X_{i+1}-\mu_0)^k\right]$. In conclusion:

$$Var\left[\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu_0)^2-\sigma_0^2\right)\right]=(\mu_4-\sigma_0^4)+2\sum_{i=1}^{n}\left(1-\frac{i}{n}\right)(\mu_i^{2,2}-\sigma_0^4) \quad .$$

Finally, we derive the covariance between

$$\sqrt{n}(\bar{X}-\mu_0) \qquad \text{and} \qquad \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu_0)^2-\sigma_0^2\right) \quad .$$

We have that:

$$\begin{aligned}&Cov\left[\sqrt{n}(\bar{X}-\mu_0);\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu_0)^2-\sigma_0^2\right)\right]=\\ &=nE\left[\frac{1}{n}\left(\sum_{i=1}^{n}(X_i-\mu_0)\right)\frac{1}{n}\left(\sum_{j=1}^{n}(X_j-\mu_0)^2\right)\right]\\ &=\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left[(X_i-\mu_0)(X_j-\mu_0)^2\right]\\ &=\frac{1}{n}\sum_{i=1}^{n}E\left[(X_i-\mu_0)^3\right]+\frac{1}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}E\left[(X_i-\mu_0)(X_j-\mu_0)^2\right]+\\ &\quad+\frac{1}{n}\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left[(X_i-\mu_0)(X_j-\mu_0)^2\right] \quad .\end{aligned}$$

From the strict stationarity of the process $\{X_t\}_{t\in\mathbb{N}}$, it follows that $E\left[(X_i-\mu_0)^3\right]=\mu_3$ for all $i\in\mathbb{N}$ and

$$Cov\left[(X_i-\mu_0);(X_j-\mu_0)^2\right]=Cov\left[(X_{i+h}-\mu_0);(X_{j+h}-\mu_0)^2\right]$$

for all $i,j\in\mathbb{N}$ and for all $h\in\mathbb{Z}$. Consequently

$$\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}E\left[(X_i-\mu_0)(X_j-\mu_0)^2\right]=\sum_{i=1}^{n-1}(n-i)\mu_i^{1,2}\quad ;$$

$$\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left[(X_i-\mu_0)(X_j-\mu_0)^2\right]=\sum_{i=1}^{n-1}(n-i)\mu_i^{2,1}\quad .$$

Then, the covariance results:

$$Cov\left[\sqrt{n}(\bar{X}-\mu_0);\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu_0)^2-\sigma_0^2\right)\right]=\mu_3+\sum_{i=1}^{n}\left(1-\frac{i}{n}\right)\left(\mu_i^{1,2}+\mu_i^{2,1}\right)$$

From the results just given it follows that

$$\mathbf{S}=\begin{bmatrix}\sigma_0^2 & \mu_3\\ \mu_3 & \mu_4-\sigma_0^4\end{bmatrix}+$$
$$+\lim_{n\to\infty}\begin{bmatrix}2\sigma_0^2\sum_{i=1}^{n}\left(1-\frac{i}{n}\right)\rho_i & \sum_{i=1}^{n}\left(1-\frac{i}{n}\right)\left(\mu_i^{1,2}+\mu_i^{2,1}\right)\\ \sum_{i=1}^{n}\left(1-\frac{i}{n}\right)\left(\mu_i^{1,2}+\mu_i^{2,1}\right) & 2\sum_{i=1}^{n}\left(1-\frac{i}{n}\right)(\mu_i^{2,2}-\sigma_0^4)\end{bmatrix}.$$

As a consequence, condition *6.* is satisfied if

$$\begin{aligned}&\lim_{n\to\infty}\sum_{i=1}^{n}\left(1-\frac{i}{n}\right)\rho_i<\infty\;;\\ &\lim_{n\to\infty}\sum_{i=1}^{n}\left(1-\frac{i}{n}\right)\left(\mu_i^{1,2}+\mu_i^{2,1}\right)<\infty\;;\\ &\lim_{n\to\infty}\sum_{i=1}^{n}\left(1-\frac{i}{n}\right)(\mu_i^{2,2}-\sigma_0^4)<\infty\;.\end{aligned}\tag{7}$$

Observe that conditions (7) are stronger than the ergodicity assumption. In detail, ergodicity assure that the following condition holds (see Davidson, [6]):

$$\begin{aligned}&\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\left(1-\frac{i}{n}\right)\rho_i=0\\ &\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\left(1-\frac{i}{n}\right)\left(\mu_i^{1,2}+\mu_i^{2,1}\right)=0\quad .\\ &\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\left(1-\frac{i}{n}\right)(\mu_i^{2,2}-\sigma_0^4)=0\end{aligned}\tag{8}$$

Conditions (8) do not assure the convergence of the series (7). For example, it can happen that (8) are verified even if the terms $\rho_i$, $\mu_i^{1,2}$ and $(\mu_i^{2,2} - \sigma_0^4)$ do not converge to 0 as $i$ increase.

The expression of $\mathbf{S}$ become easier if it is assumed that

$$\begin{aligned} &\sum_{i=1}^{\infty} |\rho_i| < \infty \ ; \\ &\sum_{i=1}^{\infty} \left|\mu_i^{1,2}\right| + \sum_{i=1}^{\infty} \left|\mu_i^{2,1}\right| < \infty \ ; \\ &\sum_{i=1}^{\infty} \left|\mu_i^{2,2} - \sigma_0^4\right| < \infty \ . \end{aligned} \tag{9}$$

In this case, we have that

$$\begin{aligned} &\lim_{n\to\infty} \sum_{i=1}^{n} \left(1 - \frac{i}{n}\right) \rho_i = \sum_{i=1}^{\infty} \rho_i \ ; \\ &\lim_{n\to\infty} \sum_{i=1}^{n} \left(1 - \frac{i}{n}\right) \mu_i^{1,2} = \sum_{i=1}^{\infty} \mu_i^{1,2} + \sum_{i=1}^{\infty} \mu_i^{2,1} \ ; \\ &\lim_{n\to\infty} \sum_{i=1}^{n} \left(1 - \frac{i}{n}\right) (\mu_i^{2,2} - \sigma_0^4) = \sum_{i=1}^{\infty} (\mu_i^{2,2} - \sigma_0^4) \ ; \end{aligned} \tag{10}$$

and $\mathbf{S}$ coincide with

$$\mathbf{S}' = \begin{bmatrix} \sigma_0^2 \left(1 + 2\sum_{i=1}^{\infty} \rho_i\right) & \mu_3 + \sum_{i=1}^{\infty} \mu_i^{1,2} + \sum_{i=1}^{\infty} \mu_i^{2,1} \\ \\ \mu_3 + \sum_{i=1}^{\infty} \mu_i^{1,2} + \sum_{i=1}^{\infty} \mu_i^{2,1} & (\mu_4 - \sigma_0^4) + 2\sum_{i=1}^{\infty} (\mu_i^{2,2} - \sigma_0^4) \end{bmatrix} .$$

As is well known, the estimators $S^2$ and $S_b^2$ are asymptotically equivalent. Then $\mathbf{S}$ is the asymptotic variance-covariance matrix also for the random vector

$$\sqrt{n} \begin{bmatrix} \bar{X} - \mu_0 \\ S^2 - \sigma_0^2 \end{bmatrix} \ .$$

It is now possible to specialize Theorem 1 as follows.

**Theorem 2** *Let $\{X_t\}_{t\in\mathbb{N}}$ be a strictly stationary and ergodic stochastic processes with $E[X_1^r] < \infty$ for $r = 1, 2, 3, 4$. Let $\mu_0$ e $\sigma_0^2$ denote the mean and the variance of the process, respectively, and assume that $\left(\mu_0, \sigma_0^2\right) \in [-\mu^* -; \mu_*] \times [\sigma_-^2 ; \sigma_+^2]$ where $\mu^* < \infty$ and $\sigma_+^2 < \infty$ are arbitrarily large and $\sigma_-^2 > 0$ is arbitrarily small. If conditions (7) hold, then*

$$\begin{bmatrix} \bar{X} \\ S^2 \end{bmatrix} \xrightarrow{p} \begin{bmatrix} \mu_0 \\ \sigma_0^2 \end{bmatrix} \qquad \textit{and} \qquad \sqrt{n} \begin{bmatrix} \bar{X} - \mu_0 \\ S^2 - \sigma_0^2 \end{bmatrix} \overset{a}{\sim} \mathcal{N}(\mathbf{0}, \mathbf{S}) \ .$$

*If conditions (9) hold, then*

$$\begin{bmatrix} \bar{X} \\ S^2 \end{bmatrix} \xrightarrow{p} \begin{bmatrix} \mu_0 \\ \sigma_0^2 \end{bmatrix} \qquad \textit{and} \qquad \sqrt{n} \begin{bmatrix} \bar{X} - \mu_0 \\ S^2 - \sigma_0^2 \end{bmatrix} \overset{a}{\sim} \mathcal{N}\left(\mathbf{0}, \mathbf{S}'\right) \ .$$

By applying the Delta Method, the following corollary is obtained.

**Corollary 1** *Under the assumptions of Theorem 2, the estimator* $\widehat{\Psi}$ *is asymptotically normally distributed. In detail,* $\sqrt{n}(\widehat{\Psi} - \psi) \overset{a}{\sim} \mathcal{N}(0, V)$ *where*

1. *if conditions (7) hold then*

$$\begin{aligned} V &= \begin{bmatrix} \dfrac{1}{\sigma_0} \\ \\ -\dfrac{\mu_0 - \xi}{2\sigma_0^3} \end{bmatrix}' \mathbf{S} \begin{bmatrix} \dfrac{1}{\sigma_0} \\ \\ -\dfrac{\mu_0 - \xi}{2\sigma_0^3} \end{bmatrix} \\ &= 1 + \frac{1}{2}\psi^2 - \gamma_1\psi + \gamma_2\frac{\psi^2}{4} + 2\lim_{n\to\infty}\sum_{i=1}^{n}\left(1 - \frac{i}{n}\right)\rho_i + \\ &\quad -\psi\lim_{n\to\infty}\sum_{i=1}^{n}\left(1 - \frac{i}{n}\right)\left(\frac{\mu_i^{1,2} + \mu_i^{2,1}}{\sigma_0^3}\right) + \frac{\psi^2}{2}\lim_{n\to\infty}\sum_{i=1}^{n}\left(1 - \frac{i}{n}\right)\left(\frac{\mu_i^{2,2}}{\sigma_0^4} - 1\right) \; . \end{aligned}$$

2. *if conditions (9) hold then*

$$\begin{aligned} V' &= \begin{bmatrix} \dfrac{1}{\sigma_0} \\ \\ -\dfrac{\mu_0 - \xi}{2\sigma_0^3} \end{bmatrix}' \mathbf{S}' \begin{bmatrix} \dfrac{1}{\sigma_0} \\ \\ -\dfrac{\mu_0 - \xi}{2\sigma_0^3} \end{bmatrix} \\ &= 1 + \frac{1}{2}\psi^2 - \gamma_1\psi + \gamma_2\frac{\psi^2}{4} + 2\sum_{i=1}^{\infty}\rho_i - \psi\sum_{i=1}^{\infty}\left(\frac{\mu_i^{1,2} + \mu_i^{2,1}}{\sigma_0^3}\right) + \frac{\psi^2}{2}\sum_{i=1}^{\infty}\left(\frac{\mu_i^{2,2}}{\sigma_0^4} - 1\right) \; . \end{aligned}$$

From the above expressions it can be observed that the asymptotic variance $V$ is equal to the variance obtained by Lo [12] in the i.i.d. case plus three terms that reflect the time dependence in the process $\{X_t\}_{t\in\mathbb{N}}$. As suggested by Lo [12], under some additional regularity conditions, the variances $V$ and $V'$ can be consistently estimated starting from the Newey and West estimator of the variance-covariance matrix $\mathbf{S}$ (which is the same as those of the variance-covariance matrix $\mathbf{S}'$). In detail, in the present context this last estimator is given by

$$\widehat{\mathbf{S}}_{NW} = \begin{bmatrix} S_b^2 & \hat{\mu}_3 \\ \hat{\mu}_3 & \hat{\mu}_4 - S_b^4 \end{bmatrix} + \begin{bmatrix} s_1 & s_3 \\ s_3 & s_2 \end{bmatrix}$$

where

$$s_1 = \sum_{j=1}^{m}\left(1 - \frac{j}{m}\right)\left[\frac{2}{n}\sum_{i=j+1}^{n}(X_i - \bar{X})(X_{i-j} - \bar{X})\right] \; ,$$

$$s_2 = \sum_{j=1}^{m}\left(1 - \frac{j}{m}\right)\left\{\frac{2}{n}\sum_{i=j+1}^{n}\left[(X_i - \bar{X})^2 - S_b^2\right]\left[(X_{i-j} - \bar{X})^2 - S_b^2\right]\right\} \; ,$$

$$\begin{aligned} s_3 = \sum_{j=1}^{m}\left(1 - \frac{j}{m}\right)&\left\{\frac{1}{n}\sum_{i=j+1}^{n}(X_i - \bar{X})\left[(X_{i-j} - \bar{X})^2 - S_b^2\right]\right. \\ &\left. + \sum_{i=j+1}^{n}\left[(X_i - \bar{X})^2 - S_b^2\right](X_{i-j} - \bar{X})\right\} \; , \end{aligned}$$

and $m$ is a function of the sample size $n$ so that $m = O(n^{1/4})$ to assure consistency. Then, the Newey and West estimator of the variances $V$ and $V'$ is given by:

$$\widehat{V}_{NW} = \begin{bmatrix} \dfrac{1}{S} \\ \\ -\dfrac{\bar{X}-\xi}{2S^3} \end{bmatrix}' \widehat{\mathbf{S}}_{NW} \begin{bmatrix} \dfrac{1}{S} \\ \\ -\dfrac{\bar{X}-\xi}{2S^3} \end{bmatrix} .$$

Concluding, the asymptotic confidence interval for $\psi$ under the general and more realistic assumption of Theorem 2 is:

$$\left( \widehat{\Psi} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\widehat{V}_{NW}}{n}} \;;\; \widehat{\Psi} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\widehat{V}_{NW}}{n}} \right) . \quad (11)$$

## 4 Asymptotic distribution of $\widehat{\Psi}$ and confidence interval for $\psi$ when returns follow a GARCH(1,1) process with symmetric innovations

In this section we specialize the results of theorem 2 and corollary 1 to the case of the GARCH(1,1) process. To ease the notation, in the remainder of this section we denote the true value of the expectation and standard deviation of the process $\{X_t\}_{t\in\mathbb{N}}$ simply by $\mu$ and $\sigma$, omitting the subscript 0.

The GARCH(1,1) model (see Bollerslev, [3]) is particularly useful in order to explain the volatility clustering in financial time series and it is the most used parametric model in the financial literature. The definition of the GARCH(1,1) process is the following:

$$\begin{aligned} &X_t - \mu = \sigma_t \epsilon_t \\ &\sigma_t^2 = \alpha_0 + \alpha_1 (X_{t-1} - \mu)^2 + \beta \sigma_{t-1}^2 \end{aligned}$$

where $\alpha_0 > 0$, $\alpha_1 \geq 0$, $\beta \geq 0$ and the random variables $\epsilon_t$, $(t = 1, 2, ...)$, referred to as the innovations of the process, are i.i.d with $E[\epsilon_1] = 0$ and $E[\epsilon_1^2] = 1$. In addition, it is common to assume that the innovations are normally distributed. However, in this paper we consider the more general case of innovations with symmetric distribution and finite fourth moment $E[\epsilon_t^4] = h_2 < \infty$. In the following, we give the expressions of $\sigma^2$, $\mu_3$, $\mu_4$, $\rho_i$, $\mu_i^{1,2}$, $\mu_i^{2,1}$, and $\mu_i^{2,2}$ in terms of the parameters $\alpha_0$, $\alpha_1$, and $\beta$, $h_2$ of the GARCH(1,1) process. In doing that, we use the following notation: $Y_t = (X_t - \mu)$, $\gamma = (\alpha_1 + \beta)$, $d = (1 - \gamma^2 - (h_2 - 1)\alpha_1^2)$. Firstly, it is well known that

$$\begin{aligned} &E[X_t] = \mu \\ &E[Y_t^2] = \sigma^2 = \begin{cases} \frac{\alpha_0}{1-\gamma} & \gamma < 1 \\ \text{do not exists} & \gamma \geq 1 \end{cases} \\ &E[Y_t^3] = \mu_3 = 0 && (12) \\ &E[Y_t^4] = \mu_4 = \begin{cases} \frac{h_2 \sigma^4}{d}(1-\gamma^2) & d > 0 \\ \\ \text{do not exists} & d \leq 0 \end{cases} \\ &E[Y_t Y_{t-j}] = 0 = \rho_j \qquad j = 1, 2, ... && (13) \end{aligned}$$

In the following, according to the assumptions of the previous sections, we suppose that $\mu_4 < \infty$. Observe that the assumption of existence of the fourth moment of $X_t$ imply that $d > 0$ and $\gamma < 1$.

Now we derive the expression of $\mu_j^{1,2}$. By the smoothing property of the conditional expectation (see Billingsley, [2]) we have that

$$\begin{aligned}
E[Y_t^2 Y_{t-j}] = \mu_j^{1,2} &= E[Y_{t-j} E[Y_t^2 | X_{t-1}, X_{t-2}, ...]] \\
&= E[Y_{t-j} (\alpha_0 + \alpha_1 Y_{t-1}^2 + \beta \sigma_{t-1}^2) E[\epsilon_t^2]] \\
&= (\alpha_1 + \beta) E[Y_{t-j} Y_{t-1}^2] \\
&= \gamma E[Y_{t-j} Y_{t-1}^2] \ .
\end{aligned}$$

Iterating the above procedure, we obtain that

$$E[Y_t^2 Y_{t-j}] = \gamma^j E[Y_t^3] = 0 \qquad \forall j = 1, 2, ...; \ . \tag{14}$$

Concerning to the product moment $\mu_j^{2,1}$ we have that

$$\begin{aligned}
E[Y_t Y_{t-j}^2] = \mu_j^{2,1} &= E[Y_{t-j}^2 E[Y_t | X_{t-1}, X_{t-2}, ...]] \\
&= E[Y_{t-j}^2 (\alpha_0 + \alpha_1 Y_{t-1}^2 + \beta \sigma_{t-1}^2)^{1/2} E[\epsilon_t]] \\
&= 0 \qquad \forall j = 1, 2, ...
\end{aligned} \tag{15}$$

Finally we compute the product moment $\mu_j^{2,2}$.

$$\begin{aligned}
E[Y_t^2 Y_{t-j}^2] = \mu_j^{2,2} &= E[Y_{t-j}^2 (\alpha_0 + \alpha_1 Y_{t-1}^2 + \beta \sigma_{t-1}^2) E[\epsilon_t^2]] \\
&= \Big( \alpha_0 E[Y_{t-j}^2] + \alpha_1 E[Y_{t-j}^2 Y_{t-1}^2] + \beta E[Y_{t-j}^2 \sigma_{t-1}^2] \Big) \\
&= \left( \frac{\alpha_0^2}{1 - (\alpha_1 + \beta)} \right) + (\alpha_1 + \beta) E[Y_{t-j}^2 Y_{t-1}^2] \\
&= \left( \frac{\alpha_0^2}{1 - \gamma} \right) + \gamma E[Y_{t-j}^2 Y_{t-1}^2] \ .
\end{aligned}$$

Iterating the above procedure we obtain:

$$\begin{aligned}
E[Y_t^2 Y_{t-j}^2] &= \left( \frac{\alpha_0^2}{1 - \gamma} \right) \sum_{i=0}^{j-2} \gamma^i + \gamma^{j-1} E[Y_t^2 Y_{t-1}^2] \\
&= \left( \frac{\alpha_0}{1 - \gamma} \right)^2 \left( 1 - \gamma^{j-1} \right) + \gamma^{j-1} E[Y_t^2 Y_{t-1}^2] \\
&= \sigma^4 + \gamma^{j-1} \left( E[Y_t^2 Y_{t-1}^2] - \sigma^4 \right) \ .
\end{aligned}$$

Focusing on $E[Y_t^2 Y_{t-1}^2]$, it turns out that

$$\begin{aligned}
E[Y_t^2 Y_{t-1}^2] &= \Big( \alpha_0 E[Y_{t-1}^2] + \alpha_1 E[Y_{t-1}^4] + \beta E[Y_{t-1}^2 \sigma_{t-1}^2] \Big) \\
&= \left( \alpha_0 \frac{\alpha_0}{1 - \gamma} + \left( \alpha_1 + \frac{1}{h_2} \beta \right) E[Y_{t-1}^4] \right) \\
&= \sigma^4 (1 - \gamma) + \left( \alpha_1 + \frac{1}{h_2} \beta \right) \frac{h_2 \sigma^4}{d} (1 - \gamma^2) \\
&= \sigma^4 + \sigma^4 \frac{\alpha_1 (h_2 - 1)(1 - \beta\gamma)}{d} \ .
\end{aligned}$$

Consequently,

$$\mu_j^{2,2} = \sigma^4 \left[1 + \gamma^{j-1} \left( \frac{\alpha_1(h_2-1)(1-\beta\gamma)}{d} \right)\right] \qquad \forall j = 1, 2, ... \tag{16}$$

It is now possible to derive the expression of $V$ in terms of the parameters of the GARCH model. In detail, from expression (12) it follows that the third standardized central moment $\gamma_1$ equals 0 while, from (13), we have that

$$\lim_{n\to\infty} \sum_{i=1}^{n} \left(1 - \frac{i}{n}\right) \rho_i = \sum_{j=1}^{\infty} \rho_j = 0 \ .$$

Analogously, from (14) and (15), it follows that

$$\lim_{n\to\infty} \sum_{i=1}^{n} \left(1 - \frac{i}{n}\right) \left( \frac{\mu_i^{1,2} + \mu_i^{2,1}}{\sigma^3} \right) = \sum_{j=1}^{\infty} \left( \frac{\mu_j^{1,2} + \mu_j^{2,1}}{\sigma^3} \right) = 0 \ .$$

From (16), noting that $(h_2 - 1) > 0$, $(1 - \beta\gamma) > 0$, and $0 \geq \gamma > 1$ it follows that

$$\begin{aligned} \sum_{j=1}^{\infty} \left| \frac{\mu_j^{2,2}}{\sigma^4} - 1 \right| &= \frac{\alpha_1(h_2-1)(1-\beta\gamma)}{d} \sum_{j=1}^{\infty} \gamma^{j-1} \\ &= \frac{\alpha_1(h_2-1)(1-\beta\gamma)}{d(1-\gamma)} < \infty \ . \end{aligned}$$

Then

$$\lim_{n\to\infty} \sum_{i=1}^{n} \left(1 - \frac{i}{n}\right) \left( \frac{\mu_i^{2,2}}{\sigma^4} - 1 \right) = \sum_{j=1}^{\infty} \left( \frac{\mu_j^{2,2}}{\sigma^4} - 1 \right) = \frac{\alpha_1(h_2-1)(1-\beta\gamma)}{d(1-\gamma)} \ .$$

Consequently, the asymptotic variance of $\sqrt{(}\widehat{\Psi} - \psi)$ under the GARCH assumption is

$$\begin{aligned} V_{GARCH} &= 1 + \frac{1}{2}\psi^2 + \gamma_2 \frac{\psi^2}{4} + \frac{\psi^2}{2} \sum_{i=1}^{\infty} \left( \frac{\mu_i^{2,2}}{\sigma^4} - 1 \right) \\ &= 1 + \frac{\psi^2}{4} \left[ \frac{(h_2-1)(1+\gamma)(1-\beta)^2}{d(1-\gamma)} \right] \qquad (17) \\ &= 1 + \frac{(\mu-\xi)^2(1-\gamma)}{4\alpha_0} \left[ \frac{(h_2-1)(1+\gamma)(1-\beta)^2}{d(1-\gamma)} \right] \ . \qquad (18) \end{aligned}$$

The variance $V_{GARCH}$ defined above can be consistently estimated in several ways, for example using the Maximum Likelihood Estimation (MLE) method (once a particular distribution for the innovations is fixed) or the Quasi Maximum Likelihood Estimation (QMLE) method. A particular estimator will be introduced and described later. At the moment, we simply observe that if $\widehat{V}_{GARCH}$ is consistent for $V_{GARCH}$, then the following large sample $(1-\alpha)$-confidence interval for $\psi$ can be introduced:

$$\left( \widehat{\Psi} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\widehat{V}_{GARCH}}{n}} \ ; \ \widehat{\Psi} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\widehat{V}_{GARCH}}{n}} \right) \ . \tag{19}$$

## 5 Design of the simulation study and computational details

In order to evaluate the effective coverage and the length of the asymptotic confidence intervals (11) and (19) we make a simulation study organized as follows.

- **Data generating process**. We suppose that the daily returns follow the $GARCH(1,1)$ model with parameters $\alpha_0 = 0.001$, $\alpha_1 = 0.1$, $\beta = 0.8$. As regard the parameter $\mu$, we consider the three different values 0.0049, 0.0249, and 0.0499. Finally, concerning the value of the daily risk free rate, we fix $\xi = 0.00068$ (which corresponds to an annual rate of about 2.5% ). The three values of the Sharpe Ratio associated to the different combination of parameters settings are: $\psi = 0.05$ when $\mu$ is small; $\psi = 0.25$ for the intermediate value of $\mu$; $\psi = 0.5$ when $\mu$ is large.
- **Distribution of the innovations**. As regard the distribution of the innovations $\{\epsilon_t\}_{t\in\mathbb{N}}$ we consider the standard Normal distribution ($h_2 = 3$), the Laplace distribution with unit variance and mean zero ($h_2 = 6$) and the Student's $t$ with 5 degrees of freedom rescaled by the factor $(3/5)^{0.5}$ (in this way, the resulting distribution has unit variance and $h_2 = 9$).
- **Sample sizes**: $n = 50; 100; 200; 400; 800; 1600$.
- **Nominal coverages**: $(1-\alpha) = 0.9; 0.95; 0.975; 0.99$.
- **Iteration**: $5 \times 10^4$

In all the above scenarios, the actual coverage and the mean length of the confidence intervals (19) and (11) are computed. We used the particular estimators of the variance of $\widehat{\Psi}$ described below.

Concerning the estimators of the asymptotic variance $V_{GARCH}$, in a preliminary simulation study we investigate the properties of the estimator obtained plugging into expression (18) the MLE estimators $\hat{\alpha}_0^*$, $\hat{\alpha}_1^*$, $\hat{\beta}^*$ and $\hat{\mu}^*$ of the parameters $\alpha_0$, $\alpha_1$, $\beta$ and $\mu$:

$$\widehat{V}^*_{GARCH} = 1 + \frac{(\hat{\mu}^* - \xi)^2(1-\hat{\gamma}^*)}{4\hat{\alpha_0}^*}\left[\frac{(h_2-1)(1+\hat{\gamma}^*)(1-\hat{\beta}^*)^2}{d(1-\hat{\gamma}^*)}\right] . \tag{20}$$

where $\hat{\gamma}^* = (\hat{\alpha}_1^* + \hat{\beta}^*)$ and $\hat{d}^* = (1 - (\hat{\gamma}^*)^2 - (h_2-1)(\hat{\alpha}_1^*)^2)$. The MLE of the GARCH parameters are obtained by numerical maximization of the following log-likelihood function

- Gaussian innovations:

$$\mathcal{L}^G(\alpha_0, \alpha_1, \beta, \mu|\mathbf{x}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\left[\log(\sigma_t^2) + \frac{(x_t-\mu)^2}{\sigma_t^2}\right]$$

  where $\mathbf{x} = (x_1, ..., x_n)$,
- Laplace innovations:

$$\mathcal{L}^L(\alpha_0, \alpha_1, \beta, \mu|\mathbf{x}) = -n\log(2\sqrt{2}) - \frac{1}{2}\sum_{t=1}^{n}\left[\log\left(\sigma_t^2\right) + \sqrt{2}\frac{|x_t-\mu|}{\sigma_t}\right]$$

- Student's $t$ innovations (5 degrees of freedom):

$$\mathcal{L}^t(\alpha_0, \alpha_1, \beta, \mu|\mathbf{x}) = -n\log\left(\frac{3\sqrt{3}}{8}\pi\right) - \frac{1}{2}\sum_{t=1}^{n}\left[\log(\sigma_t^2) + 6\log\left(1 + \frac{(x_t-\mu)^2}{3\sigma_t^2}\right)\right].$$

The preliminary simulation study shows that the estimator (20) is not adequate because it often happens that the estimators obtained starting from a simulated time series of returns are not coherent with the assumption of the existence of the fourth moment of the process $\{X_t\}_{t\in\mathbb{N}}$, i.e. the estimates does not satisfy the condition $d > 0$. In these cases the value of $V_{GARCH}$ and the confidence interval (19) do not exist. In order to avoid this problem we consider the Constrained Maximum Likelihood Estimators (CMLE) of the parameters of the GARCH model. In detail, the just mentioned estimators are defined as the solution of the following optimization:

$$\begin{cases} \max\limits_{\mu,\alpha_0,\alpha,\beta} \mathcal{L}^{\bullet}(\mu,\alpha_0,\alpha_1,\beta|\mathbf{x}) \\ \\ \text{sub} \quad d > 0 \end{cases} \tag{21}$$

Let $\hat{\alpha_0}$, $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\mu}$ denote the CMLE estimators just defined. A further estimators of $V_{GARCH}$ can be obtained by replacing the MLE in expression (20) by the CMLE. In a further preliminary simulation study, it turn out that the just introduced estimator improves those provided by expression (20) but sometimes gives huge values of $\hat{v}_{GARCH}$. A simple investigation shows that this fact happens when the estimated value of $\alpha_0$ is too small. In order to avoid this further problem we adopt the following "hybrid" estimator of $V_{GARCH}$:

$$\widehat{V}_{GARCH} = 1 + \frac{\widehat{\Psi}^2}{4}\left[\frac{(h_2-1)(1+\hat{\gamma})(1-\hat{\beta})^2}{d(1-\hat{\gamma})}\right] \quad . \tag{22}$$

Concerning the Newey-West estimator $\widehat{V}_{NW}$, for the sake of simplicity, we do not determine the value of the bandwidth by the automatic selection procedure proposed in Newey and West [14] or Andrews and Monahan [1] but we set $m = 5 \times n^{1/4}$.

All the simulation are performed in Matlab. The latter program was chosen because it possesses the very powerful optimization function `fmincon` which was used in order to solve the constrained optimization problem (21).

## 6 Results

The simulated actual coverage of the confidence intervals (11) and (19) are given in Table 1 and Table 2, respectively. As it was easy to guess, simulations show that the fatter the tails of the innovations and, consequently, of the distribution of the returns, the worst the coverage accuracy of the large sample confidence intervals. Concerning the confidence intervals based on the Newey and West variance estimators it can be observed that the greater the true value of the Sharpe Ratio, the lower the actual coverage of the asymptotic confidence intervals. Moreover, the actual coverage of the confidence intervals (11) approaches the nominal level from the low. In our opinion, the actual coverage of the large sample confidence interval (11) is sufficiently close to its nominal value only when the sample size is 1600 and the true value of the Sharpe Ratio is 0.05. In the other cases an even larger sample size is necessary to reach a good coverage accuracy. Concerning the coverage accuracy of the confidence interval (19), the actual coverage is not always decreasing in the true value of the Sharpe Ratio and it does not always converge to the nominal confidence level from the low. For example, the actual coverage tends to be greater than the nominal coverage in the case of the

| $\psi$ | 0.05 | | | | 0.25 | | | | 0.50 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | 0.90 | 0.95 | 0.975 | 0.99 | 0.90 | 0.95 | 0.975 | 0.99 | 0.90 | 0.95 | 0.975 | 0.99 |
| **Gaussian Innovations** | | | | | | | | | | | | |
| 50 | 0.7928 | 0.8600 | 0.9025 | 0.9368 | 0.7800 | 0.8474 | 0.8908 | 0.9282 | 0.7521 | 0.8225 | 0.8711 | 0.9108 |
| 100 | 0.8356 | 0.8956 | 0.9299 | 0.9594 | 0.8249 | 0.8861 | 0.9267 | 0.9561 | 0.7985 | 0.8652 | 0.9072 | 0.9432 |
| 200 | 0.8581 | 0.9170 | 0.9495 | 0.9729 | 0.8526 | 0.9101 | 0.9446 | 0.9699 | 0.8297 | 0.8929 | 0.9305 | 0.9604 |
| 400 | 0.8772 | 0.9325 | 0.9615 | 0.9812 | 0.8678 | 0.9258 | 0.9569 | 0.9783 | 0.8497 | 0.9104 | 0.9460 | 0.9711 |
| 800 | 0.8824 | 0.9373 | 0.9654 | 0.9842 | 0.8796 | 0.9343 | 0.9641 | 0.9823 | 0.8645 | 0.9228 | 0.9550 | 0.9780 |
| 1600 | **0.8895** | **0.9415** | **0.9692** | **0.9868** | 0.8852 | 0.9382 | 0.9662 | 0.9848 | 0.8738 | 0.9300 | 0.9628 | 0.9826 |
| **Laplace Innovations** | | | | | | | | | | | | |
| 50 | 0.7785 | 0.8447 | 0.8859 | 0.9235 | 0.7522 | 0.8210 | 0.8681 | 0.9081 | 0.6990 | 0.7722 | 0.8243 | 0.8708 |
| 100 | 0.8221 | 0.8850 | 0.9249 | 0.9540 | 0.7947 | 0.8612 | 0.9049 | 0.9391 | 0.7456 | 0.8198 | 0.8670 | 0.9093 |
| 200 | 0.8515 | 0.9109 | 0.9446 | 0.9692 | 0.8299 | 0.8913 | 0.9293 | 0.9591 | 0.7842 | 0.8560 | 0.9006 | 0.9374 |
| 400 | 0.8716 | 0.9275 | 0.9585 | 0.9796 | 0.8492 | 0.9094 | 0.9447 | 0.9695 | 0.8159 | 0.8822 | 0.9237 | 0.9534 |
| 800 | 0.8834 | 0.9365 | 0.9642 | 0.9830 | 0.8639 | 0.9200 | 0.9533 | 0.9776 | 0.8349 | 0.8991 | 0.9371 | 0.9662 |
| 1600 | **0.8886** | **0.9418** | **0.9683** | **0.9852** | 0.8740 | 0.9300 | 0.9608 | 0.9812 | 0.8526 | 0.9145 | 0.9494 | 0.9740 |
| **Student's $t$ Innovations** | | | | | | | | | | | | |
| 50 | 0.7789 | 0.8454 | 0.8896 | 0.9259 | 0.7559 | 0.8229 | 0.8704 | 0.9102 | 0.6974 | 0.7714 | 0.8250 | 0.8716 |
| 100 | 0.8243 | 0.8883 | 0.9256 | 0.9542 | 0.7926 | 0.8615 | 0.9044 | 0.9395 | 0.7422 | 0.8159 | 0.8663 | 0.9087 |
| 200 | 0.8513 | 0.9095 | 0.9444 | 0.9691 | 0.8242 | 0.8878 | 0.9279 | 0.9573 | 0.7745 | 0.8448 | 0.8914 | 0.9300 |
| 400 | 0.8679 | 0.9253 | 0.9563 | 0.9776 | 0.8436 | 0.9054 | 0.9423 | 0.9700 | 0.7975 | 0.8684 | 0.9122 | 0.9479 |
| 800 | 0.8802 | 0.9351 | 0.9651 | 0.9835 | 0.8581 | 0.9175 | 0.9518 | 0.9746 | 0.8233 | 0.8887 | 0.9288 | 0.9589 |
| 1600 | **0.8886** | **0.9419** | **0.9679** | **0.9858** | 0.8696 | 0.9274 | 0.9589 | 0.9799 | 0.8370 | 0.9021 | 0.9398 | 0.9673 |

**Table 1** Actual coverage of the large sample confidence interval (11) for the Sharpe Ratio

Laplace innovations while it tends to be lower than the nominal coverage in the case of Student's $t$ innovations. In general, we retain that the confidence interval (19) has a better coverage accuracy than the confidence interval (11). In particular, in the most cases, a sample size of 50 can be considered sufficient in order to reach a good coverage accuracy when the innovations are Gaussian. In the case of the Laplace innovations, we observe a less uniform situation. In detail, when the value of the Sharpe ratio is small (0.05), a sample size of 800 is necessary. In the other cases, a sample size of 50 can be considered sufficient for the confidence intervals with nominal coverage 0.95, 0.975, and 0.99. Finally, when the innovations follows the Student's $t$ distribution, a good coverage accuracy for all the confidence levels is reached only in the case of small value of the Sharpe Ratio and sample size 1600 (a similar result was obtained with the Newey-West variance estimator).

| $\psi$ | 0.05 | | | | 0.25 | | | | 0.50 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | 0.90 | 0.95 | 0.975 | 0.99 | 0.90 | 0.95 | 0.975 | 0.99 | 0.90 | 0.95 | 0.975 | 0.99 |
| **Gaussian Innovations** | | | | | | | | | | | | |
| 50 | 0.9166 | 0.9595 | 0.9803 | 0.9919 | 0.9007 | 0.9500 | 0.9757 | 0.9903 | 0.8899 | 0.9401 | 0.9679 | 0.9852 |
| 100 | 0.9106 | 0.9565 | 0.9792 | 0.9922 | 0.9014 | 0.9501 | 0.9740 | 0.9897 | 0.8953 | 0.9430 | 0.9688 | 0.9854 |
| 200 | 0.9049 | 0.9540 | 0.9781 | 0.9918 | 0.9050 | 0.9520 | 0.9751 | 0.9895 | 0.8983 | 0.9467 | 0.9714 | 0.9869 |
| 400 | 0.9034 | 0.9537 | 0.9772 | 0.9911 | 0.9032 | 0.9528 | 0.9768 | 0.9898 | 0.9026 | 0.9476 | 0.9724 | 0.9876 |
| 800 | 0.8994 | 0.9506 | 0.9753 | 0.9907 | 0.9010 | 0.9507 | 0.9750 | 0.9898 | 0.9008 | 0.9501 | 0.9740 | 0.9887 |
| 1600 | 0.8984 | 0.9497 | 0.9747 | 0.9900 | 0.8990 | 0.9492 | 0.9741 | 0.9892 | 0.9003 | 0.9498 | 0.9751 | 0.9901 |
| **Laplace Innovations** | | | | | | | | | | | | |
| 50 | 0.9368 | 0.9721 | 0.9880 | 0.9966 | 0.9109 | 0.9624 | 0.9839 | 0.9943 | 0.8984 | 0.9444 | 0.9694 | 0.9863 |
| 100 | 0.9348 | 0.9730 | 0.9891 | 0.9966 | 0.9121 | 0.9528 | 0.9752 | 0.9896 | 0.9019 | 0.9450 | 0.9691 | 0.9858 |
| 200 | 0.9243 | 0.9676 | 0.9852 | 0.9951 | 0.9159 | 0.9559 | 0.9772 | 0.9903 | 0.9103 | 0.9510 | 0.9726 | 0.9877 |
| 400 | 0.9127 | 0.9614 | 0.9831 | 0.9942 | 0.9164 | 0.9577 | 0.9787 | 0.9909 | 0.9168 | 0.9553 | 0.9751 | 0.9886 |
| 800 | 0.9073 | 0.9546 | 0.9781 | 0.9919 | 0.9136 | 0.9566 | 0.9783 | 0.9909 | 0.9192 | 0.9580 | 0.9781 | 0.9912 |
| 1600 | 0.9013 | 0.9516 | 0.9758 | 0.9903 | 0.9097 | 0.9561 | 0.9779 | 0.9907 | 0.9163 | 0.9593 | 0.9787 | 0.9906 |
| **Student's $t$ Innovations** | | | | | | | | | | | | |
| 50 | 0.8773 | 0.9199 | 0.9387 | 0.9491 | 0.8004 | 0.8457 | 0.8662 | 0.8772 | 0.6962 | 0.7357 | 0.7544 | 0.7673 |
| 100 | 0.8825 | 0.9240 | 0.9420 | 0.9535 | 0.7645 | 0.8097 | 0.8341 | 0.8492 | 0.6944 | 0.7259 | 0.7419 | 0.7527 |
| 200 | 0.8904 | 0.9330 | 0.9523 | 0.9633 | 0.7605 | 0.7951 | 0.8149 | 0.8304 | 0.7332 | 0.7608 | 0.7752 | 0.7843 |
| 400 | 0.8938 | 0.9396 | 0.9607 | 0.9722 | 0.8011 | 0.8328 | 0.8479 | 0.8572 | 0.8042 | 0.8282 | 0.8412 | 0.8479 |
| 800 | 0.8988 | 0.9461 | 0.9676 | 0.9797 | 0.8658 | 0.8960 | 0.9101 | 0.9179 | 0.8728 | 0.8966 | 0.9079 | 0.9138 |
| 1600 | 0.9015 | 0.9509 | 0.9720 | 0.9859 | 0.9153 | 0.9465 | 0.9612 | 0.9687 | 0.9275 | 0.9528 | 0.9633 | 0.9693 |

**Table 2** Actual coverage of the large sample confidence interval (19) for the Sharpe Ratio

| $\psi$ | 0.05 | | | | 0.25 | | | | 0.50 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | 0.90 | 0.95 | 0.975 | 0.99 | 0.90 | 0.95 | 0.975 | 0.99 | 0.90 | 0.95 | 0.975 | 0.99 |
| **Gaussian Innovations** | | | | | | | | | | | | |
| 50 | 0.4896 | 0.5834 | 0.6672 | 0.7667 | 0.5395 | 0.6429 | 0.7352 | 0.8449 | 0.6582 | 0.7843 | 0.8969 | 1.0307 |
| 100 | 0.3373 | 0.4020 | 0.4597 | 0.5283 | 0.3689 | 0.4395 | 0.5026 | 0.5776 | 0.4459 | 0.5313 | 0.6076 | 0.6983 |
| 200 | 0.2355 | 0.2806 | 0.3209 | 0.3687 | 0.2553 | 0.3042 | 0.3479 | 0.3998 | 0.3052 | 0.3637 | 0.4159 | 0.4780 |
| 400 | 0.1656 | 0.1973 | 0.2257 | 0.2593 | 0.1779 | 0.2120 | 0.2425 | 0.2786 | 0.2106 | 0.2510 | 0.2870 | 0.3298 |
| 800 | 0.1168 | 0.1392 | 0.1592 | 0.1829 | 0.1249 | 0.1488 | 0.1702 | 0.1956 | 0.1470 | 0.1751 | 0.2003 | 0.2302 |
| 1600 | 0.0825 | 0.0983 | 0.1125 | 0.1292 | 0.0880 | 0.1048 | 0.1199 | 0.1378 | 0.1032 | 0.1230 | 0.1407 | 0.1616 |
| **Laplace Innovations** | | | | | | | | | | | | |
| 50 | 0.5423 | 0.6462 | 0.7390 | 0.8493 | 0.6928 | 0.8255 | 0.9441 | 1.0850 | 1.0148 | 1.2092 | 1.3828 | 1.5891 |
| 100 | 0.3611 | 0.4302 | 0.4920 | 0.5654 | 0.4682 | 0.5578 | 0.6379 | 0.7331 | 0.6896 | 0.8217 | 0.9397 | 1.0799 |
| 200 | 0.2452 | 0.2922 | 0.3341 | 0.3840 | 0.3189 | 0.3800 | 0.4345 | 0.4994 | 0.4688 | 0.5586 | 0.6388 | 0.7341 |
| 400 | 0.1696 | 0.2021 | 0.2311 | 0.2656 | 0.2171 | 0.2586 | 0.2958 | 0.3399 | 0.3153 | 0.3757 | 0.4296 | 0.4937 |
| 800 | 0.1185 | 0.1412 | 0.1615 | 0.1856 | 0.1477 | 0.1760 | 0.2013 | 0.2313 | 0.2103 | 0.2506 | 0.2866 | 0.3294 |
| 1600 | 0.0833 | 0.0992 | 0.1135 | 0.1304 | 0.1010 | 0.1203 | 0.1376 | 0.1581 | 0.1420 | 0.1692 | 0.1935 | 0.2223 |
| **Student's $t$ Innovations** | | | | | | | | | | | | |
| 50 | 0.4683 | 0.5580 | 0.6381 | 0.7333 | 0.4979 | 0.5933 | 0.6785 | 0.7798 | 0.6122 | 0.7295 | 0.8342 | 0.9587 |
| 100 | 0.3321 | 0.3957 | 0.4525 | 0.5201 | 0.3728 | 0.4442 | 0.5080 | 0.5838 | 0.5172 | 0.6163 | 0.7048 | 0.8099 |
| 200 | 0.2365 | 0.2819 | 0.3223 | 0.3704 | 0.2871 | 0.3421 | 0.3912 | 0.4496 | 0.4167 | 0.4965 | 0.5679 | 0.6526 |
| 400 | 0.1686 | 0.2008 | 0.2297 | 0.2640 | 0.2194 | 0.2615 | 0.2990 | 0.3436 | 0.3449 | 0.4110 | 0.4700 | 0.5401 |
| 800 | 0.1196 | 0.1425 | 0.1629 | 0.1872 | 0.1671 | 0.1992 | 0.2278 | 0.2617 | 0.2665 | 0.3175 | 0.3631 | 0.4173 |
| 1600 | 0.0844 | 0.1005 | 0.1150 | 0.1321 | 0.1203 | 0.1433 | 0.1639 | 0.1884 | 0.1920 | 0.2287 | 0.2616 | 0.3006 |

**Table 3** Average length of the large sample confidence interval (19).

| $\psi$ | 0.05 | | | | 0.25 | | | | 0.50 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | 0.90 | 0.95 | 0.975 | 0.99 | 0.90 | 0.95 | 0.975 | 0.99 | 0.90 | 0.95 | 0.975 | 0.99 |
| **Gaussian Innovations** | | | | | | | | | | | | |
| 50 | 0.3840 | 0.4576 | 0.5233 | 0.6013 | 0.3924 | 0.4676 | 0.5347 | 0.6145 | 0.4183 | 0.4984 | 0.5700 | 0.6551 |
| 100 | 0.2934 | 0.3497 | 0.3999 | 0.4595 | 0.3015 | 0.3593 | 0.4108 | 0.4721 | 0.3256 | 0.3879 | 0.4436 | 0.5098 |
| 200 | 0.2172 | 0.2588 | 0.2959 | 0.3401 | 0.2246 | 0.2676 | 0.3060 | 0.3516 | 0.2459 | 0.2931 | 0.3351 | 0.3852 |
| 400 | 0.1580 | 0.1883 | 0.2153 | 0.2474 | 0.1642 | 0.1957 | 0.2238 | 0.2571 | 0.1819 | 0.2167 | 0.2478 | 0.2848 |
| 800 | 0.1136 | 0.1354 | 0.1548 | 0.1779 | 0.1187 | 0.1415 | 0.1618 | 0.1859 | 0.1330 | 0.1585 | 0.1813 | 0.2083 |
| 1600 | 0.0812 | 0.0967 | 0.1106 | 0.1271 | 0.0851 | 0.1013 | 0.1159 | 0.1332 | 0.0962 | 0.1146 | 0.1311 | 0.1507 |
| **Laplace Innovations** | | | | | | | | | | | | |
| 50 | 0.3761 | 0.4481 | 0.5125 | 0.5889 | 0.3934 | 0.4688 | 0.5361 | 0.6161 | 0.4391 | 0.5232 | 0.5984 | 0.6876 |
| 100 | 0.2887 | 0.3440 | 0.3934 | 0.4521 | 0.3046 | 0.3629 | 0.4150 | 0.4770 | 0.3487 | 0.4154 | 0.4751 | 0.5460 |
| 200 | 0.2149 | 0.2560 | 0.2928 | 0.3365 | 0.2294 | 0.2733 | 0.3126 | 0.3592 | 0.2709 | 0.3228 | 0.3692 | 0.4243 |
| 400 | 0.1569 | 0.1869 | 0.2138 | 0.2456 | 0.1701 | 0.2027 | 0.2319 | 0.2664 | 0.2063 | 0.2459 | 0.2812 | 0.3231 |
| 800 | 0.1132 | 0.1349 | 0.1543 | 0.1773 | 0.1247 | 0.1486 | 0.1699 | 0.1952 | 0.1550 | 0.1847 | 0.2112 | 0.2427 |
| 1600 | 0.0812 | 0.0967 | 0.1106 | 0.1271 | 0.0905 | 0.1078 | 0.1233 | 0.1417 | 0.1150 | 0,1370 | 0,1567 | 0,1801 |
| **Student's $t$ Innovations** | | | | | | | | | | | | |
| 50 | 0.3779 | 0.4504 | 0.5150 | 0.5919 | 0.3915 | 0.4665 | 0.5335 | 0.6131 | 0.4338 | 0.5169 | 0.5912 | 0.6794 |
| 100 | 0.2893 | 0.3447 | 0.3942 | 0.4530 | 0.3038 | 0.3620 | 0.4140 | 0.4758 | 0.3456 | 0.4118 | 0.4710 | 0.5413 |
| 200 | 0.2145 | 0.2556 | 0.2923 | 0.3359 | 0.2287 | 0.2725 | 0.3116 | 0.3581 | 0.2692 | 0.3208 | 0.3668 | 0.4216 |
| 400 | 0.1566 | 0.1866 | 0.2134 | 0.2453 | 0.1703 | 0.2029 | 0.2320 | 0.2666 | 0.2070 | 0.2467 | 0.2821 | 0.3242 |
| 800 | 0.1131 | 0.1348 | 0.1542 | 0.1772 | 0.1254 | 0.1495 | 0.1709 | 0.1964 | 0.1577 | 0.1879 | 0.2149 | 0.2470 |
| 1600 | 0.0811 | 0.0966 | 0.1105 | 0.1269 | 0.0918 | 0.1094 | 0.1251 | 0.1438 | 0.1187 | 0.1415 | 0.1618 | 0.1859 |

**Table 4** Average length of the large sample confidence interval (11).

About the average length of the large sample confidence intervals we observe that the confidence interval (11) has a average length lower than the those of the confidence interval (19). The differences in the average length tend to decrease when the sample size increases.

## 7 Conclusions

In this paper we re-examine the problem of determining the asymptotic distribution of the natural estimator of the Sharpe Ratio under the general setting of not-i.i.d.returns. In detail we point out that the various and technical regularity conditions assuring the consistency and the asymptotic normality of the GMM estimators recalled in Lo [12] can be a lot simplified when the particular estimator under investigation is $\widehat{\Psi}$. In detail the regularity condition required for the validity of the large sample properties of $\widehat{\Psi}$ are the following: the stochastic process followed by the returns is strictly stationary, ergodic, with finite fourth moment and it satisfy the conditions (7). Moreover, we explicitly give the expression of the asymptotic variance of $\sqrt{n}(\widehat{\Psi} - \psi)$. Starting from that result, it is possible to define a large sample confidence interval for $\psi$ provided that a consistent estimator for the asymptotic variance is defined. As suggested by Lo [12], the asymptotic variance could be consistently estimated stating from the Newey-West estimators of the variance-covariance matrix of the random vector $(\bar{X}, S^2)$. As shown in Corollary 1, the asymptotic variance of $\sqrt{n}(\widehat{\Psi} - \psi)$ depends on infinite unknown parameters. This can lead to several estimation problems which could be avoided re-expressing the asymptotic variance in terms of a finite number of parameters once a "good" parametric model for the process is introduced. In particular, we suppose that the returns follow the GARCH(1,1) process with symmetric innovations and we obtain the expression of the asymptotic variance of $\sqrt{n}(\widehat{\Psi} - \psi)$ in terms of the four parameter $\alpha_0$, $\alpha_1$, $\beta$, and $\mu$.

In a simulation study we evaluate the performances of the general large sample confidence interval (11) and of the confidence interval derived under the GARCH assumption. The simulations show that the confidence interval based on the Newey-West estimator has a worst coverage accuracy than those based on the GARCH model. It is worthwhile to note that the over-performance of the GARCH-based confidence interval can be due to the fact that we really simulate a GARCH(1,1) process and, consequently, the confidence interval (19) does not suffer from misspecification error. However, the main object of the simulation study was to determine the minimum sample size assuring a good coverage accuracy of the large sample confidence intervals. Concerning this aspect, we observe that the general confidence intervals require a huge sample size in order to reach a good coverage accuracy: in the most cases 1600 observation are not sufficient. A slightly different conclusion can be made about the confidence interval based on GARCH assumption. The latter shows a very different behavior for the different models for the innovations considered. In detail, when the innovations are Gaussian, a sample size of 50 is generally sufficient to reach a good coverage accuracy while a sample size of 1600 turn out to be insufficient in the most cases when the innovations have the Student's $t$ distribution with 5 d.f. Concluding, the results obtained in the simulation study reveal that, although the asymptotic theory of GMM estimators allows to derive the expression of asymptotic confidence intervals for the Sharpe ratio also when returns are not i.i.d., these results should be applied with caution because the confidence intervals obtained could have an effective coverage very different from the nominal one.

## References

1. Andrews, D.W.K., Monahan, J.C., An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econornetrica*,**60**, 953-966, (1992).
2. Billingsley P., *Probability and Measure*, Wiley, (1995).
3. Bollerslev T., Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, **31**, 307-327, (1986).
4. Brockwell P.J., Davis R.A., *Time Series: Theory and Methods*. Springer Verlag, 2nd ed, (1991).
5. Christie S., Is the Sharpe Ratio Useful in Asset Allocation? *MAFC Research Papers No.31*, Applied Finance Centre, Macquarie University, (2005).
6. Davidson J., *Stochastic Limit Theory: An Introduction for Econometricians*, Oxford University Press, (1994).
7. De Capitani L., Zenga M., Point and interval estimation for some financial performance measures, *Submitted*, (2010)

8. Hall A.R., *Generalized Method of Moments*, Oxford University Press, (2005).
9. Hansen L.P., Large Sample Properties of Generalized Methods of Moments Estimators, *Econometrica*, **50**, 1029-1054, (1982).
10. Jobson, J., Korkie, B., Performance Hypothesis Testing with the Sharpe and Treynor Measure, *The Journal of Finance*, **36**, 889–908, (1981).
11. Karlin S., Taylor H. M., *A first course in stochastic processes*. Second edition, Academic Press, New York, (1975).
12. Lo A.W., The Statistics of Sharpe Ratio. *Financial Analyst Journal*, **58**, 36-52, (2002).
13. Miller, R., Gehr, A., Sample Bias and Sharpe's Performance Measure: A Note, *Journal of Financial and Quantitative Analysis*, (1978).
14. Newey W.K., West K.D., Automatic lag selection in covariance matrix estimation, *Review of Economic Studies*, **61**, 631-653, (1994).
15. Opdike J.D., Comparing Sharpe Ratios: So Where are the p-values?. *Journal of Asset Management*, **8**(5), (2007).
16. Sen P. K., Singer J. M., *Large Sample Methods in Statistics: An Introduction with Applications*, Chapman and Hall, London, (1993).
17. Sharpe, W.F., Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, *The Journal of Finance*, **19**, No. 3, 425-442, (1964).
18. Sharpe, W.F., Mutual fund Performance. *Journal of Business*, January, 119–138, (1966).
19. Sharpe, W.F., The Sharpe Ratio. *Journal of Portfolio Management*, Fall, 49–58, (1994).