The Sensitivity of the Asymptotic Variance of Performance Measures with Respect to Skewness and Kurtosis

Philippe Bertrand[a] and Costin Protopopescu[b]

[a] *Greqam, University of Aix-Marseille 2*
2, rue de la charité
13236 Marseille cedex 02, France
philippe.bertrand@univmed.fr
[b] *Greqam, University of Aix-Marseille 2*
2, rue de la charité
13236 Marseille cedex 02, France
costin.protopopescu@univmed.fr

ABSTRACT

Performance measures such as the Sharpe ratio and the information ratio are estimation subject to estimation error. Lo (2002) derives the explicit expressions for the statistical distribution of the Sharpe ratio. Bertrand and Protopopescu (2007) have extended his work to the bivariate case which corresponds to the Information ratio.

In the present paper, we analyze the effects of skewness and kurtosis of portfolio and benchmark returns on the precision of the estimation of the Sharpe ratio and of the information ratio.

We show that these effects are in line with what decision theory suggests about preferences of investors about skewness and kurtosis. Moreover, these effects are significant and can disturb the performance evaluation process if they are neglected.

JEL classification: G11, G12, C10

Keywords: Sharpe ratio; Information ratio; Asymptotic distribution; Skewness; Kurtosis; Comparative static

I.   INTRODUCTION

Investment management industry provides investors with performance measures based on some capital market theory. Two of the most popular performance measures are the Sharpe (1966,1994) ratio and the information ratio.

The common practice in the money managers industry is to impose a limit on the volatility of the deviation of the active portfolio from the benchmark, namely on the tracking error volatility (TEV). The pioneer of this approach is Roll (1992). This setup leads naturally to the use of information ratios (IR), as a performance measure, defined as the ratios of the portfolio excess return over his benchmark to its TEV.

In a recent paper, Lo (2002) derives the explicit expressions for the statistical distribution of the Sharpe ratio using the standard asymptotic theory under several sets of assumptions for the return-generating process.

Bertrand and Protopopescu (2007) (hereafter, BP (2007)) have extended his work to the information ratio (IR), assuming that each return generating process is i.i.d. while allowing however for cross-correlation between the returns.

In the present paper, we analyze the effects of skewness and kurtosis of portfolio and benchmark returns on the precision of the estimation of the Sharpe ratio and of the information ratio.

In a first section, we recall a result from Lo (2002) and gives another proof for the derivation of the asymptotic variance of the Sharpe ratio statistics, SR, when the returns are supposed to be i.i.d. Then, we recall briefly two results from BP (2007).

In a second section, we analyze the effects of skewness and kurtosis on the variance of the Sharpe ratio and of the information ratio. We show that these effects are in line with what decision theory suggests about preferences of investors about skewness and kurtosis. Moreover, these effects are significant and can disturb the performance evaluation process if they are neglected.

II.   ASYMPTOTIC VARIANCE OF PERFORMANCE MEASURES

A.   The Sharpe Ratio

Let R_t denote the one-period simple return of a portfolio. Its mean and variance, μ and σ^2, are given by:

$$\mu \equiv E(R_t) \text{ and } \sigma^2 \equiv Var(R_t)$$

Recall that the Sharpe ratio (SR) is defined as the ratio of the excess expected return (relative to the risk-free rate, r) to the standard deviation of return:

$$SR = \frac{\mu - r}{\sigma}$$

As recall by Lo (2002), μ and σ are the population moments of the distribution of R_t. As such, they are unobservable and must be estimated statistically using historical data and are, therefore, subject to estimation error. He then derives explicit expressions for the statistical distribution of the Sharpe ratio using standard asymptotic theory under several sets of assumptions for the return-generating process.

In particular, Lo has established that the standard error of the Sharpe ratio (SR) in the i.i.d. case[1] is given by:

$$\sigma_N\left(\hat{SR}\right) = \sqrt{\left(1 + \frac{1}{2}SR^2\right)/T}$$ (1)

In the general i.i.d. case, we can establish the following result[2] :

Theorem 1: *The variance of the Sharpe ratio statistics, SR, when the returns are supposed to be i.i.d. is*:

$$\sigma^2(\hat{SR}) = 1/T\,[1 - SR\,\mu^3/\sigma^3 + 1/2\,SR^2\,\mu^4 - \sigma^4/2\sigma^4\,]$$
$$= 1/T\,[1 - SR \cdot SK + 1/4\,SR^2(K-1)]$$ (2)

Proof: see appendix A.

This expression allows some interesting comments:
- First, notice that for a Gaussian distribution (SK=0 and K=3), the expression (2) reduces to the expression (1).
- The way skewness and kurtosis entered in the expression (2) is in direct link with what decision theory suggests about preferences about the third and fourth moments of a risky outcome. Indeed, the asymptotic variance of the Sharpe ratio, $\sigma^2(SR)$, is increasing in kurtosis and decreasing in skewness. More precisely, portfolio with returns that are positively (resp. negatively) skewed has a $\sigma^2(SR)$ that is lower (resp. higher) than that of the Gaussian case. Recalling that investors are supposed to exhibit preferences for skewness and aversion to kurtosis, we may conclude that the variance of the Sharpe ratio will be all the more small as portfolio returns will be preferred (in terms of the third and fourth moments) by an investor. Thus, for a given Sharpe ratio level, the portfolio that is preferred by an investor will be also that for which the Sharpe ratio will be estimated with the most precision.

B.   The Information Ratio

In this section, we recall two results from BP (2007). They consider the statistics of the information ratio in the general i.i.d. case and finally in the Gaussian i.i.d. case, where closed analytical formula is obtained.

Let R_P and R_B denote the one-period return of an active portfolio and of its associated benchmark. The mean returns are denoted by μ_P and μ_B, and the variances by σ_P and σ_B. The term $\sigma(R_P\text{-}R_B)$ is the tracking error volatility and is defined as[3] :

$$\sigma^2(R_P\text{-}R_B) \equiv \sigma_{TE}^2 \equiv \sigma_P^2 + \sigma_B^2 - 2\sigma_{PB} = \sigma_P^2 + \sigma_B^2 - 2\rho\sigma_P\sigma_B$$

The quantities μ_P, μ_B, σ_P, σ_B and σ_{PB} (or ρ) are the population moments of the joint distribution of R_P and R_B. These are unobservable and must be estimated using historical data.

The information Ratio is defined as:

$$IR \equiv \frac{\mu_P - \mu_B}{\sqrt{\sigma_P^2 + \sigma_B^2 - 2\rho\sigma_P\sigma_B}}$$

Given a sample of historical returns $((R_{P1},R_{B1})$, $(R_{P2},R_{B2}),..., (R_{PT},R_{BT}))$, the estimator of the IR is given by:

$$\hat{IR} = \frac{\hat{\mu}_P - \hat{\mu}_B}{\sqrt{\hat{\sigma}_P^2 + \hat{\sigma}_B^2 - 2\hat{\rho}\hat{\sigma}_P\hat{\sigma}_B}} \tag{3}$$

According to the equation (3), the estimator of the IR is positively and linearly related to the estimator of the excess mean returns of the portfolio relative to the benchmark, $\hat{\mu}_P - \hat{\mu}_B$ It is also an increasing function of the correlation coefficient, ρ.

Ceteris Paribus, the estimator of the IR as a function of the estimator of the portfolio or the benchmark variance reaches a maximum. This maximum is reached for $\hat{\sigma}_P = \hat{\sigma}_B$ as the correlation coefficient, $\hat{\rho}$, tends towards unity. When the correlation coefficient is lower than one, the maximum is reached for $\hat{\sigma}_P < \hat{\sigma}_B$ (or $\hat{\sigma}_P > \hat{\sigma}_B$). Smaller is the value of the correlation coefficient, bigger is the spread between $\hat{\sigma}_P$ and $\hat{\sigma}_B$ (for which the IR is maximum). In other words, less absolute risk compensates less correlation. In the hypothetical case where $\hat{\rho}$ would be less or equal to zero, the estimator of the IR would be maximum for $\hat{\sigma}_P = 0$ or $\hat{\sigma}_B = 0$.

BP (2007) has established the following results.

Theorem 2: *The variance of the information ratio statistics, IR, is given by the following quadratic form defined by the semi-positive definite asymptotic cross-variance matrix Σ of the simple empirical moments[4]:*

$$\sigma^2(\hat{IR}) = \frac{1}{T} \sum_{\substack{i \in I \\ 1 \le i+j \le 2}} \sum_{j \in I} \frac{\partial \varphi}{\partial M_i} \frac{\partial \varphi}{\partial M_j} \Sigma_{i,j}(M), \ \ i,j \in I = \{0,1,2\}$$

Proof: see BP (2007).

As we can see in the appendix *B*, the variance of the information ratio statistics is a non linear function of the 14 cross-moments of the returns.

　　　BP (2007) have specified the result of the preceding theorem in the special case of i.i.d. normal distributed returns for the portfolio and the benchmark.

Theorem 3: *The variance of the information ratio statistics, IR, when the returns are supposed to be i.i.d. normal is:*

$$\sigma_N(\hat{IR}) = \sqrt{\left(1 + \frac{1}{2} IR^2\right)/T}$$

Proof : see BP (2007).

Notice that under the Gaussian hypothesis, the expression of the variance of the information ratio have a closed form with respect to *IR*.

III.　　ANALYSIS OF THE EFFECTS OF SKEWNESS AND KURTOSIS ON ASYMPTOTIC VARIANCE OF PERFORMANCE MEASURE

A.　　The Sharpe Ratio

Taking into account non-normality in the returns, especially skewness and kurtosis, can have significant effect on the standard error of the Sharpe ratio estimator. The tables below contain the percentage variation in the standard error of the Sharpe ratio compared to normality (bold number at SK=0 and K=3) given by[5] :

$$\sigma(\hat{SR})/\sigma_{\mathcal{N}}(\hat{SR})-1$$

The Table 1 highlights, for a Sharpe ratio of 0.25, the effects of asymmetries and extreme events on σ (SR). For example, ceteris paribus, a skewness of -1 leads to an increase of 11.5 % of σ (SR) compared to its value under the Gaussian assumption. In the same way, ceteris paribus, a kurtosis of 10 leads to an increase of 5.2 % of σ (SR).

　　　For example, the "Gabelli Asset Fund" exhibits over a 5 years sample, a skewness of -1.41 and a kurtosis of 4.87 for an estimated Sharpe ratio of 0.313. This leads to an increase in σ (SR) of 21.0% from $\sigma_{\mathcal{N}}$(SR)=0.132 to σ (SR)=0.160, T=60.

Table 1
Variation of the standard error of the Sharpe ratio estimator (SR=0.25), %.

	Skewness							
Kurtosis	**-5**	**-4**	**-3**	**-2**	**-1**	**0**	**1**	**2**
0	47.2	38.7	29.7	20.0	9.4	-2.3	-15.6	-31.5
3	48.7	40.3	31.4	21.9	11.5	**0.0**	-13.0	-28.2
5	49.7	41.4	32.6	23.1	12.8	1.5	-11.2	-26.1
10	52.3	44.1	35.4	26.1	16.1	5.2	-7.1	-21.2
15	54.7	46.7	38.2	29.1	19.3	8.7	-3.1	-16.5
20	57.2	49.2	40.9	32.0	22.5	12.1	0.8	-12.1
25	59.5	51.8	43.5	34.8	25.5	15.5	4.4	-7.9
30	61.9	54.2	46.2	37.6	28.5	18.7	8.0	-3.9
35	64.2	56.7	48.7	40.3	31.4	21.9	11.5	0.0
40	66.5	59.1	51.3	43.0	34.3	24.9	14.8	3.7
45	68.8	61.4	53.7	45.6	37.1	27.9	18.1	7.3

Table 2
Variation of the standard error of the Sharpe ratio estimator (SR=0.5), %.

	Skewness							
Kurtosis	**-5**	**-4**	**-3**	**-2**	**-1**	**0**	**1**	**2**
0	74.8	61.6	47.2	31.2	13.0	-8.7	-37.6	NA
3	79.5	66.7	52.8	37.4	20.2	**0.0**	-25.5	-66.7
5	82.6	70.0	56.3	41.4	24.7	5.4	-18.4	-52.9
10	90.0	78.0	65.0	50.9	35.4	17.9	-2.8	-29.3
15	97.2	85.6	73.2	59.9	45.3	29.1	10.6	-11.8
20	104.1	92.9	81.0	68.3	54.6	39.4	22.5	2.7
25	110.8	100.0	88.6	76.4	63.3	49.1	33.3	15.5
30	117.3	106.8	95.8	84.1	71.6	58.1	43.4	26.9
35	123.6	113.4	102.8	91.5	79.5	66.7	52.8	37.4
40	129.7	119.8	109.5	98.6	87.1	74.8	61.6	47.2
45	135.7	126.1	116.0	105.5	94.4	82.6	70.0	56.3

Table 3
Variation of the standard error of the Sharpe ratio estimator (SR=0.75), %.

	Skewness							
Kurtosis	**-5**	**-4**	**-3**	**-2**	**-1**	**0**	**1**	**2**
0	89.7	73.6	55.8	35.7	12.1	-18.1	-70.8	NA
3	98.2	82.8	66.0	47.3	25.9	**0.0**	-35.6	NA
5	103.6	88.7	72.5	54.6	34.3	10.4	-20.4	-77.9
10	116.7	102.7	87.7	71.4	53.4	33.0	8.8	-22.7
15	129.0	115.8	101.8	86.8	70.4	52.2	31.6	7.1
20	140.7	128.2	115.0	100.9	85.8	69.3	51.0	30.2
25	151.8	139.9	127.4	114.1	100.0	84.8	68.2	49.8
30	162.5	151.1	139.2	126.6	113.3	99.1	83.8	67.1
35	172.7	161.8	150.4	138.4	125.8	112.4	98.2	82.8
40	182.6	172.1	161.1	149.6	137.6	125.0	111.6	97.2
45	192.2	182.0	171.4	160.4	148.9	136.8	124.2	110.7

We can see on the Tables 2 and 3 that the effect is increasing with the value of SR. The three tables confirm that $\sigma(SR)$ is more sensitive to the skewness than to the kurtosis as the expression (2) had already shown it, as long as SR is smaller than one.

B.   The Information Ratio

Recall from theorem 2 that the asymptotic variance of the information ratio, $\sigma^2(\hat{IR})$, can not be expressed like a function of the IR. Thus, we cannot provide tables as simple as those of the preceding section. Nevertheless, we can prove the following result:

Proposition 1: *The partial derivatives of the asymptotic variance (and standard error) of the information ratio, $\sigma^2(IR)$, with respect to the skewness and to the kurtosis of the portfolio and of the benchmark returns are given by:*

$$\frac{\partial \sigma_{IR}}{\partial SK_P} = -\frac{\sigma_P^3\left(IR + IR^3\right)}{2T\sigma_{TE}^2\sigma_{IR}} \text{ and } \frac{\partial \sigma_{IR}}{\partial SK_B} = -\frac{\sigma_B^3\left(IR + IR^3\right)}{2T\sigma_{TE}^2\sigma_{IR}}$$

$$\frac{\partial \sigma_{IR}}{\partial K_P} = \frac{\sigma_P^4 IR^2}{2T\sigma_{TE}^3\sigma_{IR}} \text{ and } \frac{\partial \sigma_{IR}}{\partial K_B} = \frac{\sigma_B^4 IR^2}{2T\sigma_{TE}^3\sigma_{IR}}$$

Proof : First, we can prove by direct computation the following:

$$\frac{\partial \sigma_{IR}^2}{\partial SK_P} = -\frac{\sigma_P^3\left(IR + IR^3\right)}{T\sigma_{TE}^2} \text{ and } \frac{\partial \sigma_{IR}^2}{\partial SK_B} = -\frac{\sigma_B^3\left(IR + IR^3\right)}{T\sigma_{TE}^2}$$

$$\frac{\partial \sigma_{IR}^2}{\partial K_P} = \frac{\sigma_P^4 IR^2}{2\sigma_{TE}^3} \text{ and } \frac{\partial \sigma_{IR}^2}{\partial K_B} = \frac{\sigma_B^4 IR^2}{T\sigma_{TE}^3}$$

Second, some straightforward algebraic manipulations lead to the desired result.
        Remark that the effect of skewness and kurtosis in the portfolio returns and in the benchmark returns are almost the same one except for the standard deviation raised to the power 3 or 4.
        Also notice that the asymptotic variance of the information ratio is decreasing in the skewness and increasing in the kurtosis as in the case of $\sigma(SR)$.
   We illustrate the sensitivity of $\sigma(IR)$ with respect to skewness and kurtosis on the data of the Gabelli Asset Funds with the SP500 as benchmark.

Figure 1
Effect of Skewness on $\sigma(IR)$
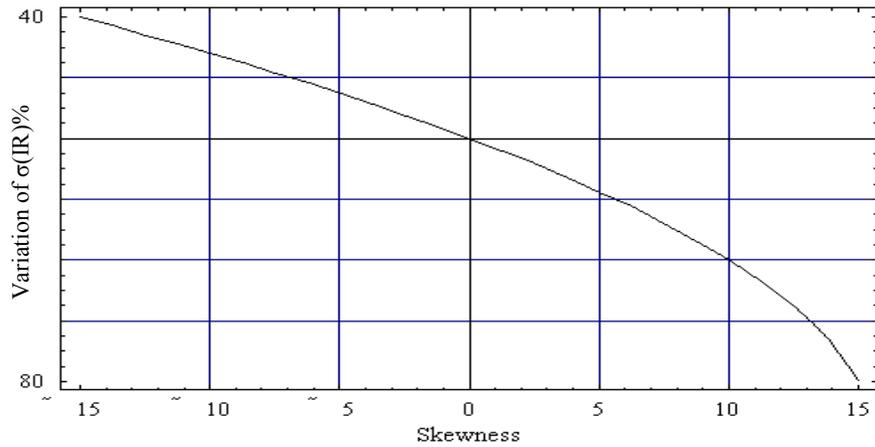
We check on Figure 1 that a variation of the skewness of the fund with respect to the Gaussian case (SK=0) can have a significant effect on the precision of the estimation of $\sigma(IR)$. The effect of the variation of the kurtosis, as shown in Figure 2, is weak.

Figure 2
Effect of Kurtosis on $\sigma(IR)$
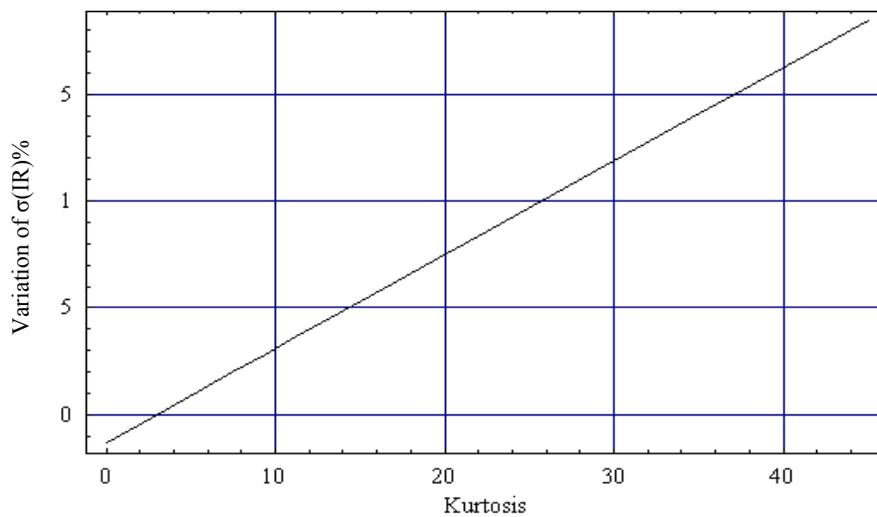
IV.   CONCLUSION

In this paper, we have analyzed the effects of skewness and kurtosis of portfolio and benchmark returns on the precision of the estimation of two of the most popular performance measures: the Sharpe ratio and the information ratio.

We have first recall some results from Lo (2002) and BP (2007) on the asymptotic variance of the performance measure estimator.

We have then studied the sensitivity of these asymptotic variances on the skewness and on the kurtosis. The magnitude of the effect is not negligible and can disturb the performance evaluation process. Recall that most of the mutual funds exhibit negative skewness and excess kurtosis (greater than 3).Thus, it is significant to make the adjustment to take into account skewness and kurtosis.

ENDNOTES

1. In fact, this result holds **only in the Gaussian case**.
2. Note that Lo (2002) establish such a result as a particular case of his GMM estimator when returns are i.i.d.
3. The covariance (resp. the correlation coefficient) between the return of the portfolio and the return of the benchmark is denoted by $\sigma_{PB} \equiv Cov(R_P, R_B)$ (resp. ρ).
4. See also appendix B.
5. Remark that the size of the sample doesn't matter in the computation as it enters in the numerator as well as in the denominator.

REFERENCES

Bertrand, P.H. and Protopopescu, C., 2007, "The Statistics of the Information Ratio", in revision for the *European Journal of Finance*.

Lo, Andrew, 2002, "The Statistics of Sharpe Ratios", *Financial Analyst Journals*, *vol*.58, 4, pp. 36-52.

Roll, R., 1992, "A Mean/Variance Analysis of Tracking Error", *The Journal of Portfolio Management*, vol. 18, 4 Summer:13-22.

Sharpe, W.F., 1966, "Mutual Fund Performance", *Journal of Business,* 39, 119–138.

Sharpe, W.F., 1994, "The Sharpe ratio", *Journal of Portfolio Management,* 21(1), 49–58.

Appendix A
Proof of Theorem 1

Recall that

$$\hat{SR} = \frac{\hat{\mu} - r}{\hat{\sigma}} = \varphi(\hat{\theta})$$

where $\hat{\theta}' = (\hat{\mu}, \hat{\sigma})$.

First, as the returns are supposed to be *i.i.d.*, the central limit theorem shows that the asymptotic distribution of $\hat{\theta}$ is given by:

$$\sqrt{T}(\hat{\theta} - \theta) \sim \mathcal{N}(0, \Sigma(\theta)) \tag{4}$$

with

$$\Sigma(\theta) = \begin{bmatrix} \sigma^2 & \mu_3 \\ \mu_3 & \mu_3 - \sigma^4 \end{bmatrix}$$

where $\mu_k = E[(X-\mu_1)^k]$ denoted the k-th centered moment of X.

Given the first-order Taylor approximation for the function $\varphi(\cdot)$:

$$\varphi(\hat{\theta}) = \varphi(\theta_0) + \frac{1}{1!} \frac{\partial \varphi}{\partial \theta}(\theta_0)(\hat{\theta} - \theta_0) + R_2(\hat{\theta}, \theta_0)$$

where the rest $\mathbf{R}_2(\hat{\theta}, \theta_0)$ of the second order verifies:

$$P\lim \sqrt{T} R_2(\hat{\theta}, \theta_0) = 0 \text{ for } T \to \infty$$

So, we have:

$$\sqrt{T}(\varphi(\hat{\theta}) - \varphi(\theta_0)) = \sqrt{T} \frac{\partial \varphi}{\partial \theta}(\theta_0)(\hat{\theta} - \theta_0) + \sqrt{T} R_2(\hat{\theta}, \theta_0)$$

Given the asymptotic property of the rest, one can write also:

$$\sqrt{T}(\varphi(\hat{\theta}) - \varphi(\theta_0)) = \sqrt{T} \frac{\partial \varphi}{\partial \theta}(\theta_0)(\hat{\theta} - \theta_0) + 0_P(1)$$

By using the classical result of a law of large numbers (5), this equality allows us to deduce the following asymptotic convergence in law:

$$\sqrt{T}(\varphi(\hat{\theta}) - \varphi(\theta_0)) \sim \mathcal{N}(0, \Omega)$$

where, $\Omega \equiv \dfrac{\partial \varphi}{\partial \theta} \Sigma(\theta) \dfrac{\partial \varphi}{\partial \theta'}$ and $\dfrac{\partial \varphi}{\partial \theta} = \begin{bmatrix} 1/\sigma \\ -\dfrac{\mu - r}{2\sigma^3} \end{bmatrix}$

This leads to the following expression for the variance for the Sharpe ratio estimator:

$$\sigma^2(\hat{SR}) = 1/T\,[1\text{-}SR\,\mu^3/\sigma^3 + 1/2\,SR^2\,\mu^4\text{-}\sigma^4/2\sigma^4\,]$$
$$= 1/T\,[1\text{-}SR{\cdot}SK + 1/4\,SR^2(K\text{-}1)]$$

where SK denotes the skewness and K the kurtosis.

Appendix B
Statistics of the *IR*

We recall in this section the elements that entered in the theorem 2 of BP (2007). For the sake of simplicity, we have defined all the moments that entered in the IR expression in terms of the non-centered moments. For a sample of size T, the estimator of the non-centered cross-moments of order k in R_P and of order l in R_B is defined as:

$$\hat{m}_{kl} = \frac{1}{T} \sum_{i=1}^{T} R_{Pi}^{k} R_{Bi}^{l}$$

In particular, the estimator of the non-centered moments of order k of the variables R_P and R_B is defined as:

$$\hat{m}_{k0} = \frac{1}{T} \sum_{i=1}^{T} R_{Pi}^{k} \text{ and } \hat{m}_{0l} = \frac{1}{T} \sum_{i=1}^{T} R_{Bi}^{l}$$

With this notation, the expression for the *IR* becomes:

$$\hat{IR} \equiv \frac{\hat{m}_{10}\text{-}\hat{m}_{01}}{\sqrt{\hat{m}_{20}\text{-}\hat{m}_{10}^{2} + \hat{m}_{02}\text{-}\hat{m}_{01}^{2}\text{-}2(\hat{m}_{11}\text{-}\hat{m}_{10}{\cdot}\hat{m}_{01})}} \qquad (6)$$

Thus, the *IR* may be considered as a function φ of $\hat{M} = \left(\hat{m}_{10}, \hat{m}_{01}, \hat{m}_{20}, \hat{m}_{02}, \hat{m}_{11} \right)$:

$$\hat{IR} = \varphi(\hat{M})$$

First, as the returns are supposed to be i.i.d., the central limit theorem shows that the asymptotic distribution of M is given by:

$$\sqrt{T}\,(\,\hat{M} - M)\, \sim \mathcal{N}(0_R{}^5, \Sigma(M)) \tag{7}$$

where the elements of the (5×5) symmetric matrix $\Sigma(M)$ are:

$$(\,T \times \mathrm{Cov}(\hat{m}_{i,j}, \hat{m}_{k,l}) = m_{i+k,\,j+l} - m_{i,j}.m_{k,l}\,,\ 0 \le i,j,k,l \le 2$$

If the centered moments had been used as variables in the function φ, then the associated covariance matrix would have had a much more complicated form.

Note that, here, contrary to the case studied by Lo (2002), the non-diagonal elements are not equal to zero.

The partial derivatives of the function φ are given by:

$$\frac{\partial \varphi}{\partial m_{10}} = -\frac{\partial \varphi}{\partial m_{01}} = \frac{m_{20} + m_{02} - 2m_{11}}{\left(m_{20} - m_{10}^2 + m_{02} - m_{01}^2 - 2\left(m_{11} - m_{10}m_{01}\right)\right)^{3/2}}$$

$$\frac{\partial \varphi}{\partial m_{20}} = \frac{\partial \varphi}{\partial m_{02}} = -\frac{1}{2}\,\frac{m_{10} - m_{01}}{\left(m_{20} - m_{10}^2 + m_{02} - m_{01}^2 - 2\left(m_{11} - m_{10}m_{01}\right)\right)^{3/2}}$$

$$\frac{\partial \varphi}{\partial m_{11}} = \frac{m_{10} - m_{01}}{\left(m_{20} - m_{10}^2 + m_{02} - m_{01}^2 - 2\left(m_{11} - m_{10}m_{01}\right)\right)^{3/2}}$$