INTERNATIONAL JOURNAL OF BUSINESS, 15(1), 2010 ISSN: 1083−4346

The Statistics of the Information Ratio

Philippe Bertrand[a] and Costin Protopopescu[b]

[a] *GREQAM, University of Aix-Marseille 2 and Euromed Management, France*
philippe.bertrand@univmed.fr
[b] *GREQAM, University of Aix-Marseille 2 and Euromed Management, France*
costin.protopopescu@univmed.fr

ABTRACT

In a recent paper, Lo (2002) derives the asymptotic distribution of the Sharpe ratio under several sets of assumptions for the return-generating process. In this paper, we extend his work to the information ratio (IR), the ratio of the excess return of a portfolio over his benchmark to its tracking-error volatility. We assume that each return generating process is i.i.d., allowing however for cross-correlation.

First, given the cross-dependency between the portfolio and the benchmark returns, we derive the analytic expression of the asymptotic variance of the IR and we show explicitly how the higher order covariance influence the precision of the variance estimation. On the other hand we study the partial derivatives of the asymptotic variance of the IR with respect to the different moments of the returns.

Second, we conduct some simulations in order to highlight the behavior of the IR's asymptotic variance.

JEL classification: *G11, G12, G13*

Key words: *Information ratio; Asymptotic distribution*

I. INTRODUCTION

The common practice in the investment management 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), who noted that the active portfolio had not only a systematically higher risk than the benchmark, but also a beta greater than one and it is not optimal. These problems have been further pointed out by Bertrand, Prigent and Sobotka (2001), Jorion (2003) and Bertrand (2010), who have studied several alternatives of portfolio optimization program.

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. Bertrand (2005, 2010) proves, by defining the IR as a forward-looking measure, that the information ratio is constant across all TEV efficient portfolios. Thus, it is the appropriate measure for the risk adjusted performance of an active portfolio with a TEV constraint against a benchmark. Indeed, it ranks all the TEV efficient portfolios in the same way.

Traditionally, the IR has been computed as an ex-post performance measure. As such, it is commonly used to compare investment managers. But, the IR must be estimated in a certain way and on a given sample. Thus, estimation errors arise, raising the following question: how accurately is the IR computed?

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. Since Sharpe ratios must be estimated, they are also subject to estimation error.

In this paper, we have extended his work to the information ratio (IR). We assume that each return generating process is i.i.d., allowing however for cross-correlation.

First of all, given the cross-dependency between the portfolio and the benchmark returns, we derive the analytic expression of the asymptotic variance of the IR and we show explicitly how the higher order covariance influence the precision of the variance estimation. Then, we derive the analytic expression for the asymptotic variance of the IR in the Gaussian i.i.d. case. Finally, we conduct some simulations in order to highlight the behavior of the asymptotic variance of the IR.

II. ASYMPTOTIC VARIANCE OF THE IMFORMATION RATIO

Recall that Lo (2002) has established that the standard error for the Sharpe ratio (SR) in the i.i.d. case estimator is given by:

$$\sigma(\mathrm{SR}) = \sqrt{\left(1 + \tfrac{1}{2}\mathrm{SR}^2\right)/\,\mathrm{T}}$$

where the Sharpe ratio is define as:

$$\mathrm{SR} = \frac{\mu_P - R_f}{\sigma(R_P)}$$

The information Ratio is defined as:

$$IR \equiv \frac{\mu_P - \mu_B}{\sigma(R_P - R_B)}$$

Let R_P and R_B denote the one-period return of an active portfolio and of its associated benchmark. The risk less interest rate is denoted R_f. The mean returns are denoted μ_p and μ_B, and the variances σ_p and σ_B. The term $\sigma(R_p - R_B)$ is the tracking error volatility and is defined as[1]:

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

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.

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:

$$IR \equiv \frac{\mu_P - \mu_B}{\sqrt{\sigma_P^2 + \sigma_B^2 - 2\rho\sigma_P\sigma_B}} \tag{1}$$

According to Equation (1), 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, $\mu_P - \mu_B$. 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 $\sigma_P = \sigma_B$ as the correlation coefficient, ρ, tends towards unity. When the correlation coefficient is lower than one, the maximum is reached for $\sigma_P < \sigma_B$ (or $\sigma_B < \sigma_P$). Smaller is the value of the correlation coefficient, bigger is the spread between σ_P and σ_B (for which the IR is maximum). In other words, less absolute risk compensates less correlation. In the hypothetical case where ρ would be less or equal to zero, the IR would be maximum for $\sigma_P = 0$ or $\sigma_B = 0$.

For the sake of simplicity, we define 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 1 in R_B is defined as:

$$m_{kl} = \frac{1}{T}\sum_{i=1}^{T} R_{Pi}^{k}.R_{Bi}^{1}$$

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

$$m_{k0} = \frac{1}{T}\sum_{i=1}^{T} R_{Pi}^{k} \quad \text{and} \quad m_{0k} = \frac{1}{T}\sum_{i=1}^{T} R_{Bi}^{k}$$

With this notation, the expression for the IR becomes:

$$IR \equiv \frac{m_{10} - m_{01}}{\sqrt{m_{20} - m_{10}^2 + m_{02} - m_{01}^2 - 2(m_{11} - m_{10}.m_{01})}} \tag{2}$$

Thus, the IR may be considered as a function φ of $M = (m_{10}, m_{01}, m_{20}, m_{02}, m_{11})$:

$$IR = \varphi(M)$$

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:

$$\sqrt{T}(M - M)\ N(0_{R^5}, \sum(M)) \tag{3}$$

Where the elements of the (5x5) symmetric matrix $\Sigma(M)$ are given by $Cov(m_{i,j}, m_{k,l}) = m_{i+k,j+1} - m_{i,j}m_{k,l}$:

$$\begin{pmatrix} m_{20} - m_{10} \cdot m_{10} & & & & \\ m_{11} - m_{10} \cdot m_{01} & m_{02} - m_{01} \cdot m_{01} & & & \\ m_{30} - m_{10} \cdot m_{20} & m_{21} - m_{20} \cdot m_{01} & m_{40} - m_{20} \cdot m_{20} & & \\ m_{12} - m_{10} \cdot m_{02} & m_{03} - m_{01} \cdot m_{02} & m_{22} - m_{20} \cdot m_{02} & m_{04} - m_{02} \cdot m_{02} & \\ m_{21} - m_{10} \cdot m_{11} & m_{12} - m_{01} \cdot m_{11} & m_{31} - m_{20} \cdot m_{11} & m_{13} - m_{02} \cdot m_{11} & m_{22} - m_{11} \cdot m_{11} \end{pmatrix}$$

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. We establish the following result.

Theorem 1: The asymptotic variance of the information ratio statistics, IR, is given by the following quadratic form*:*

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

Proof: see appendix.

Remark: The analytic expression of $\sigma^2(IR)$ as a function of the components of M is cumbersome for two reasons. The first one is connected with the expression of the partial derivatives of φ.

The second one refers to the asymptotic variance $\sigma^2(\text{IR})$ which depends on the asymptotic variance of the empirical moments, i.e. the 5x5 symmetric matrix Σ, containing all the cross-moments of order less than 4 between the elements of the vector M. Given the fact that in practice the basic analysis of the financial data is made using the moments of order lower than 2, it is useful for us to focus on the comparative analysis of the mean, of the standard-deviation and of the linear correlation coefficient between the portfolio and the benchmark returns. The other cross-moments are kept constant and equal to their empirical counterparts obtained from the data.
We can specify the result of the preceding theorem in the special case of i.i.d. normal distributed returns for the portfolio and the benchmark.

Theorem 2: The asymptotic variance of the information ratio statistics, IR, when the returns are supposed to be i.i.d. normal is:

$$\sigma^2(\text{IR}) = \left(1 + \frac{1}{2}\text{IR}^2\right) / \text{T}$$

Proof: see the appendix.

This result extends the result of Lo to the case of the information ratio. Therefore, our work is a mixed analysis. First, it uses the analytical formulas deduced in the theorem stated before. Secondly, we use simulated data to replace the cross-moments. We have chosen to use simulated data because it is always more logical to do a theoretical analysis for a flexible GDP. Notice that the use of simulated data has a different connotation than in the pure Monte Carlo works (see Sherer (2006) for example).

Before starting our analysis, we anticipate the fact that the numerical analysis based on the analytical expression given by the theorem for the asymptotic variance has a great potential to reveal some interesting facts about the behavior of the standard deviation $\sigma(\text{IR})$. Note that it is in general difficult to highlight these facts on the sole basis of an empirical study because of the multiple nonlinear dependence on the crossed moments of order higher than 2.

III. THE SAMPLING

We simulated the data generating process of portfolio and benchmark returns as a bivariate sample of normal random variables:

$$N\left(\begin{pmatrix} 01.3 \\ 0.10 \end{pmatrix}, \begin{pmatrix} 0.04 & 0.0256 \\ 0.0256 & 0.0256 \end{pmatrix}\right)$$

Which ensures a linear coefficient $\rho = 0.85$. Thus, the theoretical level for the information ratio is equal to 0.2835, value considered, for example by Grinold and Kahn (1995) as a “good” level. According to BARRA, the average annual information ratio estimated on 300 funds is 0.12 after taking into account the management fees and 0.36 before.

Recall that $\sigma^2(\mathrm{IR})$ depends on all the cross-moments of order less than 4, so the use of simulation is required in order to provide fixed values for the (cross-) moments of order 3 and 4 coherent with our choice made above for the (cross-) moments of order less than 2.

By the way, we have two possibilities for dealing with the simulation. First, one can try to simulate some symmetric distribution (student like or other), but in this case all the moments will depend in a too rigid manner on a single parameter (degree of freedom), so the flexibility of the family to generate a wide class of samples with variable high moments is very poor.

Second, we can start with a bivariate Gaussian DGP and deal with a non-Gaussian DGP as we restrict the moments of order higher than 3 to remain constant while forcing some of the moments of order smaller than 2 to vary. Indeed, given the recurrence law between the cross-moments in a bivariate Gaussian distribution, the variation of the moments of order lower than 2 implies the change of the moments of order higher than 3 as shown in the formula given in the proof of Theorem 2.

For the generated sample, we obtain the following estimated counterparts of the cross-moments:

$$
\begin{array}{llll}
 & & & m_{13} = 0.00410 \\
 & & m_{21} = 0.012698 & m_{31} = 0.00653 \\
m_{10} = 0.12760 & m_{02} = 0.035624 & m_{12} = 0.010080 & m_{22} = 0.00502 \\
m_{01} = 0.09986 & m_{11} = 0.040084 & m_{30} = 0.017468 & m_{40} = 0.00917 \\
 & m_{20} = 0.056175 & m_{03} = 0.008728 & m_{04} = 0.00364
\end{array}
$$

Throughout the next sections we will use the following estimated spreads, obtained from the simulated data:

$$\hat{\delta}_{\mu} \mathrm{def} = \mu_P - \mu_B = m_{10} - m_{01} = 0.0277$$

$$\hat{\delta}_{\sigma^2} \mathrm{def} = \sigma_P^2 - \sigma_B^2 = m_{20} - m_{10}^2 - m_{02} + m_{01}^2 = 0.01424$$

IV. NUMERICAL SENSITIVITY ANALYSIS OF THE IR'S ASYMPTOTIC STANDARD DEVIATION

The aim of this section is to study the behavior of the asymptotic variance of the IR with respect to the excess return of the active portfolio, the excess volatility of the active portfolio over its benchmark as well as with respect to the linear correlation between the portfolio and the benchmark returns.

A. The Behavior of the Information Ratio

First, we analyze the behavior of the IR as a function of ρ (the linear correlation coefficient), by applying an additive variation on $\hat{\delta}_{\mu}$ (the difference of the first order moments) and by keeping the second order moments equal to the empirical value obtained from the simulated data sample. As our aim is to make a comparative analysis, we have chosen to make our variation using a multiplicative scale factor K which acts as a magnitude factor on the empirical spread found from the data. More precisely, we try to give an answer to the following question: what happens to the IR if we keep the same variances, but we change the empirical mean of the active portfolio by a scale factor K:

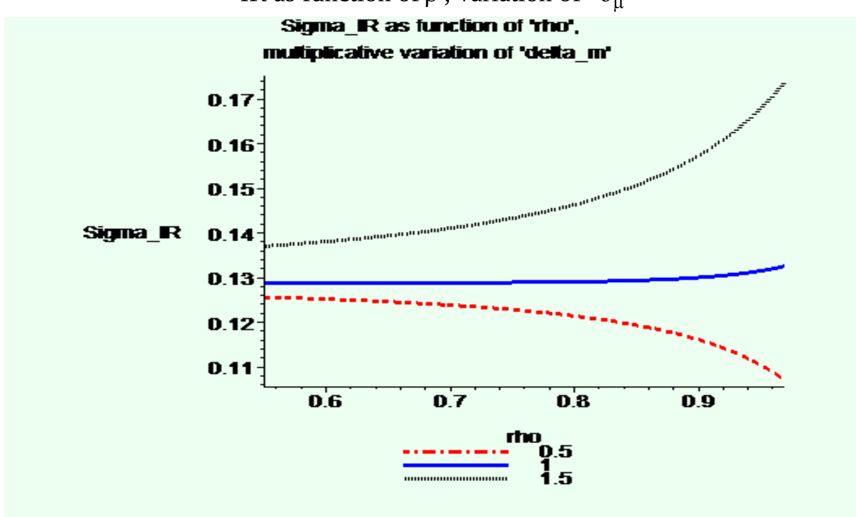
$$\mu_P - \mu_B = K \cdot \hat{\delta}_\mu$$
$$K \in \{0.5, 1, 1.5\}$$

We note by μ_P the new hypothetical empirical mean in order to differentiate it from μ_P the empirical mean found from the data.

The advantage of this framework is that it is scale free, so K=0.5 represents a variation of the means which is only 50% of the simulated value found from the sample, while K=1.5 represents a variation of the means which is 50% bigger than the simulated value found from the sample. As we have seen in the former section, the IR is increasing with respect to ρ. However the graph will inform us about the convexity of the curve which will be different in the two cases presented below (the sensitivity with respect to the mean variation and the sensitivity with respect to the variance variation, respectively).

Notice that the IR given by expression (2.2) is not directly written as a function of ρ (contrary to the expression (2.1)). It is a function of m_{11} and of the fixed standard deviation of portfolio and benchmark returns. So we used the reparametrisation in order to obtain the linear correlation as direct argument for IR.

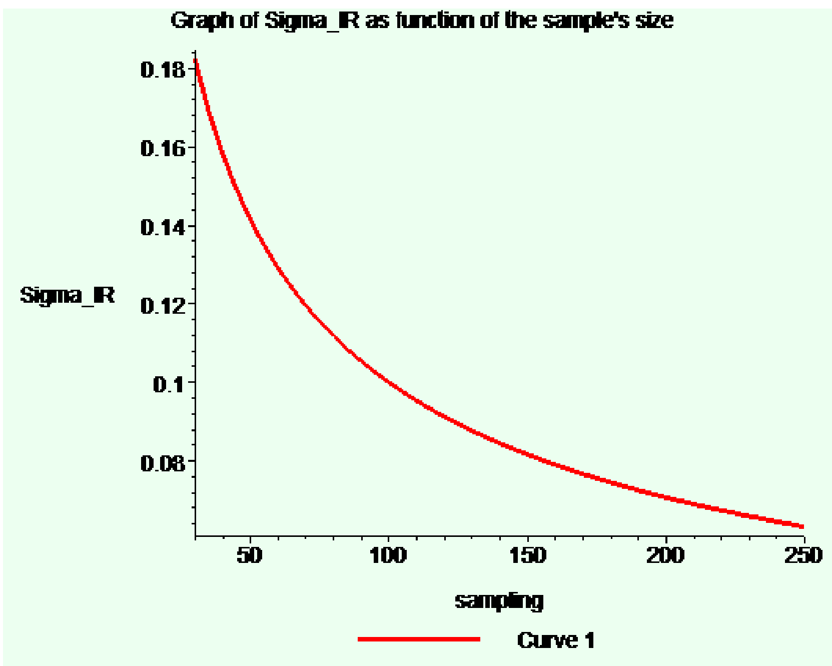
Figure 1
IR as function of ρ, variation of δ_μ



B. The Behavior of the Asymptotic Standard Deviation of the Information Ratio

In what follows, we are interested to show the effect of the sample size on the $\sigma(\text{IR})$ so we plot $\sigma(\text{IR})$ as a function of T. As expected, the curve is decreasing and convex. We notice that by using weekly data rather than monthly on a five years sample, the asymptotic standard deviation of the IR decreases by half. This fact can have important practical issues because it allows the analysts to judge the precision of his estimation for the information ratio statistic.

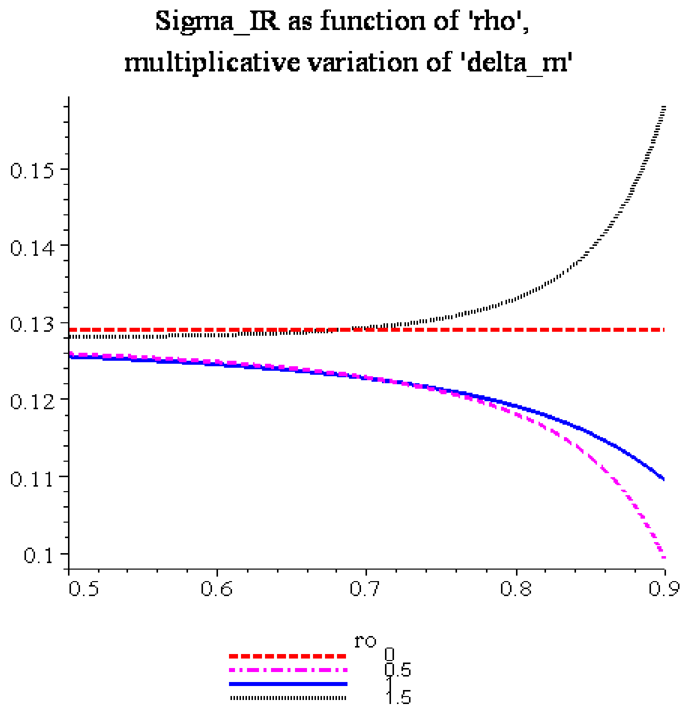
Figure 2
$\sigma(\text{IR})$ as function of the sample size sigma



In the next graph, we analyze the evolution of the asymptotic standard deviation of the IR as a function of ρ, by applying a multiplicative scale factor for $\hat{\delta}_\mu$.

Figure 3

$\sigma(\mathrm{IR})$ as function of ρ, multiplicative variation of δ_μ ($\sigma_p = 20\%$)



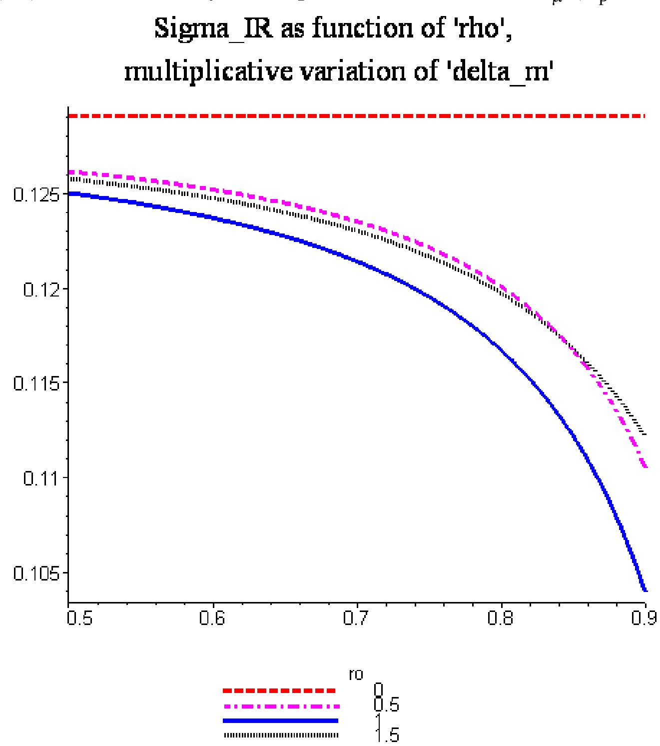
For low levels of the correlation coefficient, the asymptotic standard deviation of the IR is not very sensitive to the excess return of the portfolio with respect to the benchmark. As the value of the correlation coefficient increases and becomes more realistic, it is no truer: the gap between the two extremes curves increases. What is more surprising is that for some values of K (i.e. K=1.5) the curve is increasing while for others (i.e. K=0.5) it is decreasing. As the portfolio return is set equal to the benchmark return, K=0, the asymptotic standard deviation of the IR is constant but not equal to zero although the IR itself is null. Thus, a rising correlation coefficient can increases, as well as decreases, the precision of the estimation of the IR, depending on the portfolio excess return over the benchmark.

For example, ceteris paribus, if $\rho = 0.9$ (resp. 0.95), for K-1.5 (that is $\mu_P = 14.50\%$) , $\sigma(\mathrm{IR}) = 0.1582$ (resp. 0.236) whereas for K=0.5 (that is $\mu_P = 11.50\%$), $\sigma(\mathrm{IR}) = 0.099$ (resp. 0.0314).

In the following graph, we will explore further the relation between the asymptotic standard deviation of the IR and the correlation coefficient for an increase in the volatility of the portfolio, $\sigma_P = 22\%$.

Figure 4

$\sigma(\text{IR})$ as function of ρ, multiplicative variation of δ_μ ($\sigma_p = 22\%$)



We still note a low sensitivity to the portfolio excess return when the coefficient of correlation takes relatively low values. As the coefficient of correlation increases, the gap between the different curves widens. But, the curves are displayed in a different order. We will reconsider this last point in the comment of graph 6.

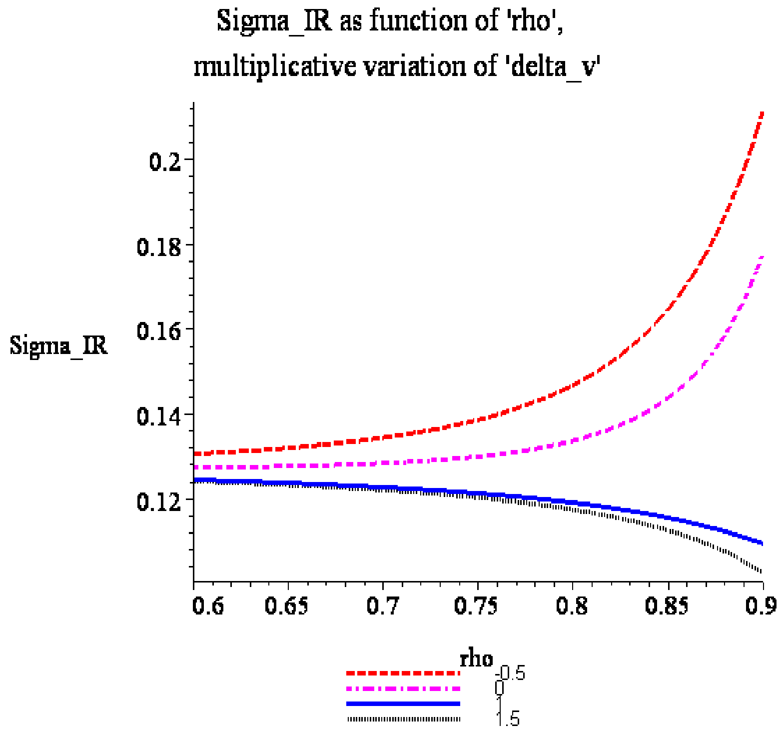
For example, ceteris paribus, if $\rho - 0.9$ (resp. 0.95), for K=1 (that is $\mu_P = 13\%$ and IR=0.2344), $\sigma(\text{IR}) = 0.10$ (resp. 0.070) whereas for K=0 (that is $\mu_P = \mu_B = 10\%$ and IR=0), $\sigma(\text{IR}) = 0.129$ (resp. 0.129). Thus, it is the portfolio that exhibits the smallest performance, which has also the smallest precision of the estimation of the IR. Thus, the assessment of the portfolio performance should be aware of that phenomenon.

Next, we study the evolution of the standard deviation of the IR as a function of ρ when we apply a multiplicative scale factor for the variation for the variance:

$$\sigma_P^2 = \sigma_B^2 + K \cdot \hat{\delta}_{\sigma^2}$$
$$K \in \{-0.5, 0, 1, 1.5\}$$

Figure 5

$\sigma(\text{IR})$ as function of ρ, multiplicative variation of δ_{σ^2}



The same qualitative comments as for the last two graphs apply to this graph. The effect of the correlation coefficient across volatility spread begins to take effect for relatively high level of the correlation coefficient. The order of the curves may be reversed according to the portfolio excess return.

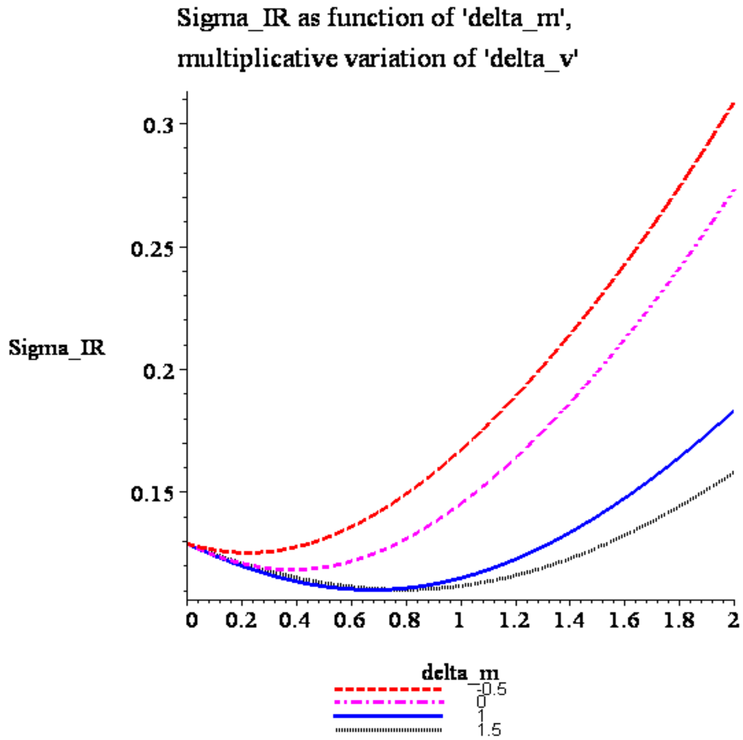
It is astonishing to note that it is the less risky portfolio for which the IR can be estimated with the least precision.

In the next graph, we study the sensitivity of $\sigma(\text{IR})$ with respect to $\hat{\delta}_{\mu}$, for the estimated value of ρ and for the four values of $\hat{\delta}_{\sigma^2}$. This can be described as a cross sensitivity with respect to both variations, in mean and in variance.

One can observe that there exists a curve parameterized by the volatility gap that minimizes $\sigma(\text{IR})$. Here, it is the curve for which $\hat{\delta}_{\sigma^2}$ =1.21, that is $\sigma_P = 20.74\%$. The minimum is reached for $\hat{\delta}_{\mu}$ =0.73, $\mu_P = 12.19\%$.

Figure 6

$\sigma(\text{IR})$ as function of δ_μ, multiplicative variation of δ_{σ^2}



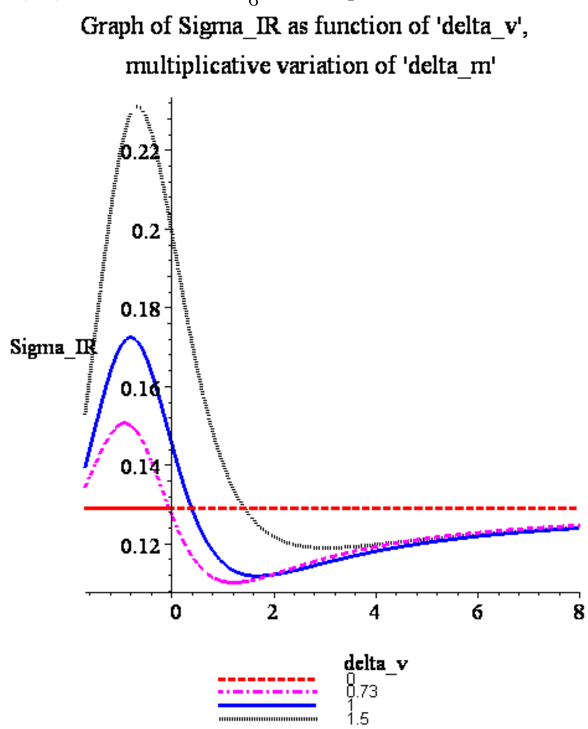
When the same analysis is conducted after exchanging the role of $\hat{\delta}_\mu$ and of $\hat{\delta}_{\sigma^2}$ respectively, we obtain a graph that is in line with the behavior of the IR as a function of the portfolio volatility. Namely, the asymptotic standard deviation of the IR, as the IR itself, reaches a maximum for a portfolio volatility less than the benchmark volatility (for $\rho < 1$). But here, there is also a minimum which is reached for a finite value $\hat{\delta}_{\sigma^2}$. Secondly, notice that the curve "$\hat{\delta}_\mu$ =0" plays the role of an asymptote for the other curves.

V. CONCLUSION

In this paper, we have extended the seminal work of Lo (2002)[2], to the information ratio IR, in the general i.i.d. case. One can see the IR as a bivariate "version" of the SR by allowing an analysis relative to a benchmark.

Figure 7

$\sigma(\text{IR})$ as function of δ_{σ^2}, multiplicative variation of δ_μ



First, we have established a result that gives the analytic expression for the asymptotic variance of the IR (theorem 1). This result highlights the effects of parameter variation on the precision of the IR estimate. Without that, one has to limit our self to the Monte Carlo simulations for every particular case (see Sherer (2004) for a bootstrapped analysis in the case of SR statistic). In addition, these simulations are not very reliable in the sense that they do not allow to deduce the global behavior of $\sigma(\text{IR})$, nor its sensitivity to the variation of the classical statistics of the portfolio returns : mean, variance and correlation.

In the Gaussian i.i.d. case, the former analytic expression for the asymptotic variance of the IR will depend only on the moments of order lower than two and the theorem 2 shows that $\sigma(\text{IR})$ is a nonlinear function of IR.

In the general i.i.d. case, the expression of $\sigma(\text{IR})$ is a complex function of the cross-moments of order less than 4. Given that $\sigma(\text{IR})$ is defined as a function of 14 cross moments, we have focused our numerical simulations on the effect of the marginal variation of the classical statistics: mean, variance and correlation. We have therefore studied the sensitivity of $\sigma(\text{IR})$ with respect to these moments of portfolio and benchmark returns.

The analysis has shown the key role played by the correlation coefficient. Indeed, $\sigma(IR)$ can be an increasing or decreasing function of the correlation coefficient, depending on $\hat{\delta}_{\mu}$ and on $\hat{\delta}_{\sigma^2}$. Besides, we have seen that $\sigma(IR)$ as a function of $\hat{\delta}_{\mu}$ reaches a minimum and as a function of $\hat{\delta}_{\sigma^2}$ reaches a maximum and a minimum.

Thus, the present work allowed us to highlight the importance of taking into account the estimated precision of the IR influenced by the individual characteristics of the portfolio and its benchmark.

Further work is needed in order to take into account the effect of skewness and kurtosis in the returns on the asymptotic variance of the IR.

ENDNOTES

1. 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) .
2. Recall that Lo has studied the asymptotic behavior of the Sharpe ratio

REFERENCES

Bertrand, P., 2005, “A Note on Portfolio Performance Attribution: Taking Risk into Account”, *Journal of Asset Management*, vol 5, 6, pp. 428-437.

Bertrand, P., 2010, “Another Look at Portfolio Optimization under Tracking-Error Constraint”, *Financial Analysts Journal*, vol. 66, no. 3 May/June.

Bertrand, P.., J.L. Prigent, and R Sobotka, 2001, “Optimisation de Portefeuille sous Contrainte de Variance de la Tracking-error”, *Banque & Marchés*, 54, pp. 19-28.

Grinold, Richard, and Ronald Kahn, (1995) *Active Portfolio Management*, Chicago, IL: Irwin.

Jorion, P., 2003, “Portfolio Optimization with Constraints on Tracking Error”, *Financial Analysts Journal*, vol 59, 5 pp. 70-82.

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

Sherer, B., 2004, “An Alternative Route to Performance Hypothesis Testing”, *Journal of Asset Management*, vol 5, 1, 5-12.

Willink, R., 2005, “Normal Moments and Hermite Polynomials”, *Statistics & Probability Letters*, vol 73, 271–275

APPENDIX

Proof of Theorem 1:
Given the first-order Taylor approximation for the function $\varphi(\cdot)$:

$$\varphi(M) = \varphi(M_0) + \frac{1}{1!}\frac{\partial\varphi}{\partial M}(M_0)\left(M - M_0\right) + R_2(M, M_0)$$

where the second order rest $R_2(M, M_0)$ verifies:

$$P\lim \sqrt{T} R_2(M, M_0) = 0 \text{ for } T \to \infty$$

So, we have:

$$\sqrt{T}(\varphi(M) - \varphi(M_0)) = \sqrt{T}\frac{\partial\varphi}{\partial M}(M_0)(M - M_0) + \sqrt{T} R_2(M, M_0)$$

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

$$\sqrt{T}(\varphi(M) - \varphi(M_0)) = \partial\varphi\partial M(M_0)\sqrt{T}(M - M_0) + O_P(1)$$

By using the classical result of a law of large numbers (3), this equality allows us to deduce the following asymptotic convergence in law:

$$\sqrt{T}(\varphi(M) - \varphi(M_0))N(O_{R^5}, \Omega)$$

where, $\Omega = \sum_{1\le i+j\le 2 i\in I}\sum_{j\in I}\partial\varphi\partial M_i\partial\varphi\partial M_j\Sigma_{i,j}(M)$ is a quadratic form obtained as a consequence of the asymptotic normality described in (3). Notice that indexes i and j are elements of the following set: $I = \{0,1,2\}$.

Finally, we obtain the following expression for the asymptotic variance of IR:

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

Partial derivatives of the function φ:

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

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

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

Proof of Theorem 2:
First, we need the expression of the non-centered moments of a normal random vector, Y. According to a result established by Willink (2005), we have:

$$\text{for } r_j 1, E(Y_1^1 \dots Y_K^k \dots Y_P^p) = \mu_k E(Y_1^1 \dots Y_P^p) + {}^{p}_{j=1}\, \sigma_{kj} r_j E(Y_1^1 \dots Y_j^{j-1} \dots Y_P^p)$$

where μ_k is the non-centered moments of the k-th elements of Y.
Thus, we obtain:

$$\begin{aligned}
m_{30} &= 3m_{20}m_{10} - 2m_{10}^3 \\
m_{40} &= 3m_{20}^2 - 2m_{10}^4 \\
m_{21} &= m_{10}^2 m_{01} + \sigma_{11}m_{01} + 2\sigma_{12}m_{10} \\
m_{31} &= m_{10}m_{21} + 2\sigma_{11}m_{11} + \sigma_{12}m_{20} \\
m_{22} &= m_{10}m_{12} + \sigma_{11}m_{02} + 2\sigma_{12}m_{11}
\end{aligned}$$

The moments m_{03}, m_{04}, m_{12} and m_{13} are obtained in the same way.

This leads to the following (5x5) symmetric matrix $\sum(M)$ which depends only on the moments of order one and two of R_P and R_b as well as the first cross-moment:

$$\begin{pmatrix}
m_{20} - m_{10}m_{10} & * & * & * & * \\
m_{11} - m_{10}m_{01} & m_{02} - m_{01}m_{01} & * & * & * \\
-2m_{10}^3 + 2m_{10}m_{20} & 2m_{10}(m_{11} - m_{10}m_{01}) & 2m_{20}^2 - 2m_{10}^4 & * & * \\
2m_{01}(m_{11} - m_{10} \cdot m_{01}) & -2m_{01}^3 + 2m_{01}m_{02} & 2(m_{11}^2 - m_{10}^2 m_{01}^2) & 2m_{02}^2 - 2m_{01}^4 & * \\
m_{10}m_{11} + m_{01}(m_{20} - 2m_{10}^2) & m_{01}m_{11} + m_{10}(m_{02} - 2m_{01}^2) & -2m_{01}m_{10}^3 + 2m_{11}m_{20} & -2m_{10}m_{01}^3 + 2m_{11}m_{02} & -2m_{10}^2 m_{01}^2 + m_{11}^2 + m_{02}m_{20}
\end{pmatrix}$$

Substituting this expression for $\sum(M)$ into $\sigma^2(IR)$ from theorem 1, we get after some algebra the desired result:

$$\sigma^2(IR) = \frac{1}{T}(1 + \frac{1}{2} IR^2).$$