



Available online at <http://scik.org>

J. Math. Comput. Sci. 11 (2021), No. 3, 3729-3746

<https://doi.org/10.28919/jmcs/5762>

ISSN: 1927-5307

MEASURING THE IMPACT OF THE STOCK MARKET INDEX RETURN ON STOCKS RETURN USING THE STOCHASTIC APPROXIMATION

ALI LABRIJI*, ABDELKRIM BENNAR, EL HOUSSINE LABRIJI, MOSTAFA RACHIK

Department of Mathematics and Computer Sciences, University Hassan II, Casablanca, Morocco

Copyright © 2021 the author(s). This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract: Value at Risk (VAR) is a risk measure frequently used in market finance. This tool gives an idea of the losses that may occur to a financial asset (share or option), but does not predict when these losses may occur.

The objective of our work is to propose a method that allows us to know the state of the economy necessary for Microsoft's performance to suffer extreme losses with alpha probability. This represents a significant asset for investors to consider before trading on the stock market. To answer this problem, we will try to explain the VAR using the ROBBINS-MONRO-JOSEPH procedure for the estimation of a percentile for binary variables, as a function of a systemic risk factor. We will also propose a method for applying the procedure without using real-time experiments. We are going to estimate the different parameters of the process, based on the history of the data available for statistical modelling, then analyse the results of the convergence of the process over the iterations, and finally see the impact of adding a random element to the binary variable y_n of the process on convergence. The final results are satisfying and seem to be in line with reality, but there is room for improvement, either by increasing the number of iterations needed to refine convergence, whereas the stochastic approximation aims to obtain a good estimate in a minimum number of iterations, or by applying other more recent procedures that show better results following simulations.

*Corresponding author

E-mail address: alilabriji@gmail.com

Received March 23, 2021

Keywords: stochastic approximation; percentile estimation; value at risk; risk factor; statistical modelisation.

2010 AMS Subject Classification: 60H30, 65C60, 62G32, 91G70.

1. INTRODUCTION

The main arbitrage in market finance is to know the right time to sell or buy financial products, while minimising losses and maximising gains. One of the first attempts to address this arbitrage was made by Markowitz (1952) [1] with his approach to minimizing the variance of a portfolio of financial assets. Since then, several approaches have been adopted, but one of the most widely used methods for managing financial product risk remains Value At Risk (VAR). Thus, the Basel Committee has even imposed its use on financial institutions [2].

VaR is the maximum level of loss that a financial asset can sustain, with a confidence level $\alpha \in (0,1)$. There are several methods for calculating VaR. However, the most commonly used are: historical VaR which is based on past returns, and assumes that future returns are independent and have the same distribution as past returns. The second method is the parametric method. In this case it is assumed that the returns on the financial asset follow a known distribution, and we deduce the VAR according to the percentile of this law which is consistent with the confidence level set. The third method uses Monte Carlo simulation, which, by setting a given time horizon, generates returns according to a specific distribution, and then calculates the VaR on the basis of the simulated returns. Typically, the return on assets is assumed to follow the following differential equation

$$R_i = \mu\Delta t + \sigma\varepsilon\sqrt{\Delta t} \quad (1)$$

With:

R_i the return on financial assets after i periods,

μ the average of financial asset prices,

σ the standard deviation of the price of the financial asset,

ε a random number generated according to the standard normal law,

and Δt the process step.

The frequently used generalisation is Wiener's process, which is more suitable for distributions with a thick tail, and which incorporates time jumps. To know more about the VaR see [3].

However, we do not know when these losses might occur, and therefore we have difficulty protecting ourselves against this risk of loss. In this paper, we will introduce a method for estimating the impact of a given exogenous risk factor on the return on the financial asset of interest, using a sequential stochastic approximation algorithm of the Robbins and Monro (1951) [4] type.

Robbins and Monro have shown that $x_n \rightarrow \theta$, the percentile value α of an unknown distribution, in probability if x_n is generated according to the following process.

$$x_{n+1} = x_n - a_n(y_n - \alpha) \quad (2)$$

with

$$\sum_{n=1}^{\infty} a_n = \infty \text{ et } \sum_{n=1}^{\infty} a_n^2 < \infty \quad (3)$$

and

$$y_n = \begin{cases} 1 & \text{if } v_n > s \\ 0 & \text{Else} \end{cases} \quad (4)$$

- v_n The value of a latent variable at iteration n
- S A threshold linked to the latent variable

Therefore, we will try to estimate the value θ of the α percentile that will cause losses equivalent to the VaR of the financial asset of interest. To do so, we will use the optimal Robbins-Monro procedure proposed by V. Roshan Joseph (2004) [5].

2. PRELIMINARIES

2.1 Optimal Robbins-Monro procedure

Let $M(x)$ be a distribution function of an unknown law and assume that $\dot{M}(\theta)$ is known with θ , such that $M(\theta) = \alpha$ with α the order of the percentile that we are trying to estimate. The experiment starts with a value x_1 close to the value of θ of the percentile of order α that we are trying to estimate, based on a priori knowledge. Therefore, we can choose an a priori distribution for θ with $E(\Theta) = x_1$ and $Var(\Theta) = \tau_1^2 < \infty$, where τ_1 represents uncertainty

of estimation of θ . Joseph proposes the following modified Robbins-Monro procedure

$$x_{n+1} = x_n - a_n(y_n - b_n) \quad (5)$$

Note that even if x_1 is fixed, x_2, \dots, x_n are random, given their dependencies on past data.

Let be

$$Z_n = x_n - \theta \quad (6)$$

Therefore

$$y_n | Z_n \sim \text{Ber}\{M(Z_n)\} \quad (7)$$

$$Z_{n+1} = Z_n - a_n(y_n - b_n) \quad (8)$$

$$E(Z_1) = 0 \text{ et } E(Z_1^2) = \tau_1^2 \quad (9)$$

Joseph found the sequences $\{a_n\}$ and $\{b_n\}$ such as $Z_n \rightarrow 0$ in probability, assuming that they verify the following condition

$$\sum_{n=1}^{\infty} a_n |b_n - \alpha| \sum_{j=1}^{n-1} a_j^2 < \infty \quad (10)$$

He has shown that:

$$a_n = \frac{c_n}{\beta b_n(1 - b_n)} \quad (11)$$

With:

$$c_n = \frac{v_n}{(1 + v_n)^{\frac{1}{2}}} \phi \left\{ \frac{\Phi^{-1}(\alpha)}{(1 + v_n)^{\frac{1}{2}}} \right\} \quad (12)$$

$$b_n = \Phi \left\{ \frac{\Phi^{-1}(\alpha)}{(1 + v_n)^{\frac{1}{2}}} \right\} \quad (13)$$

$$v_{n+1} = v_n - \frac{c_n^2}{b_n(1 - b_n)} \quad (14)$$

With Φ the distribution function of the standard normal law, ϕ its density function,

$\beta = \dot{M}(\theta)/\phi\{\Phi^{-1}(\alpha)\}$, and $v_1 = \beta^2 \tau_1^2$, so once v_1 is determined, we can start the process.

Moreover, the confidence interval at the level $(1 - \gamma)$ for each value of de x_n is

$[x_n \pm \Phi^{-1}(\gamma/2)\tau_n]$, with $\tau_n = v_n^{1/2}/\beta$

2.2 Problematic

It should be noted that this is not the first application of stochastic approximation to the VAR, see [6]-[7], but where the cited work focuses on modeling the VAR, we seek earlier to explain it.

Let VaR_q the Value at Risk of a financial asset with a $q \in (0,1)$ level, we define y_n such as :

$$y_n = \begin{cases} 1 & \text{if } r_n > VaR_q \\ 0 & \text{Else} \end{cases} \quad (15)$$

With r_n the return on the financial asset we want to study.

Let X be an exogenous factor, correlated to the performance of the financial asset we wish to study, and which is capable of explaining the losses or gains of the financial asset.

So, we can use Robbins-Monro-Joseph's algorithm to estimate the value x_n such that x_n has a probability α that $y_n = 1$ and thus a probability $1 - \alpha$ of causing losses greater than or equal to the threshold of VaR_q , with α the percentile of the unknown density function that we want to estimate

Let $f(x_n)$ be a function such that $r_n = f(x_n) + e$ with e a noise with a mean of zero. We can therefore write y_n in the following form

$$y_n = \begin{cases} 1 & \text{if } f(x_n) > VaR_q \\ 0 & \text{Else} \end{cases} \quad (16)$$

2.3 Model estimation

The most well-known basic model in the literature is the Capital Asset Pricing Model (CAMP). Created by Sharpe, in the 1960s [8] and based on studies by Markowitz in 1952 (modern portfolio theory). This tool describes the relationship between the risk of a financial asset and the expected return on that asset, and can be formulated as follows

$$\bar{r}_a = r_f + \beta(\bar{r}_m - r_f) \quad (17)$$

With

- \bar{r}_a The expected return on financial assets
- r_f The risk-free rate
- \bar{r}_m The expected market profitability

Therefore, the equation can be written as follows.

$$\bar{r}_a = \beta \bar{r}_m + c \quad (18)$$

With β and c , respectively, the coefficient and constant estimated by linear regression. Indeed, the stock market indicator summarizes several information, and gives us an idea about the trend of the global evolution of the economy, as well as of finance.

Taking in our case the example of Microsoft as well as the Standard and Poor's (S&P) stock market index. We have taken the closing prices of the Microsoft and S&P from 16/02/2015 to 07/02/2020 and calculated the weekly return with the following formula. $r_n = \ln\left(\frac{p_n}{p_{n-1}}\right)$.

Below is Table 1 describing the two returns.

Table 1. Description of the returns studied

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Return of S&P	261	0.002	0.018	-0.073	-0.006	0.012	0.047
Return of Microsoft	261	0.005	0.028	-0.094	-0.010	0.021	0.140

The two yields have a correlation of 0.7414856. It can also be seen from the graph below, which represents the Microsoft Return versus the S&P Return, that there is indeed a linear relationship between the two returns, visible through the linear regression line. However, we also note the existence of outliers which may affect the quality of our model (Figure 1).

MEASURING THE IMPACT OF THE STOCK MARKET INDEX RETURN

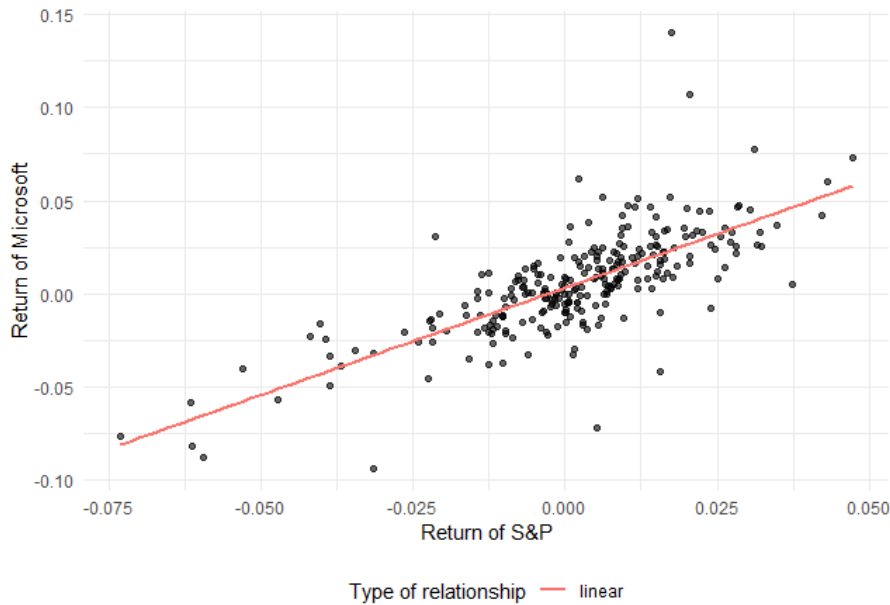


Figure 1. Adjusting Microsoft's performance to S&P's performance by a linear line

A linear regression model is a statistical model that estimates a straight line using the ordinary least squares method. At first glance, the fit is acceptable. However, it is noticeable that in some regions, it would be preferable to use a polynomial with a degree greater than 1. The most commonly used method to solve this local fitting problem is the LOESS (LOcally weighted Scatterplot Smoother) method. Originally proposed by Cleveland (1979) [9] and further developed by Cleveland and Devlin (1988) [10], it specifically refers to a locally weighted polynomial regression. At each point in the database, a low-level polynomial is fitted to a sample, based on the values of explanatory variables close to the point whose response is estimated. The polynomial is adjusted using weighted least squares, giving more weight to points close to the point whose response is estimated and less weight to points further away. The value of the regression function for the point is then obtained by evaluating the local polynomial using the explanatory variable values for that data point. The LOESS adjustment is achieved once the regression function values have been calculated for each of the n points in our database (Figure 2).

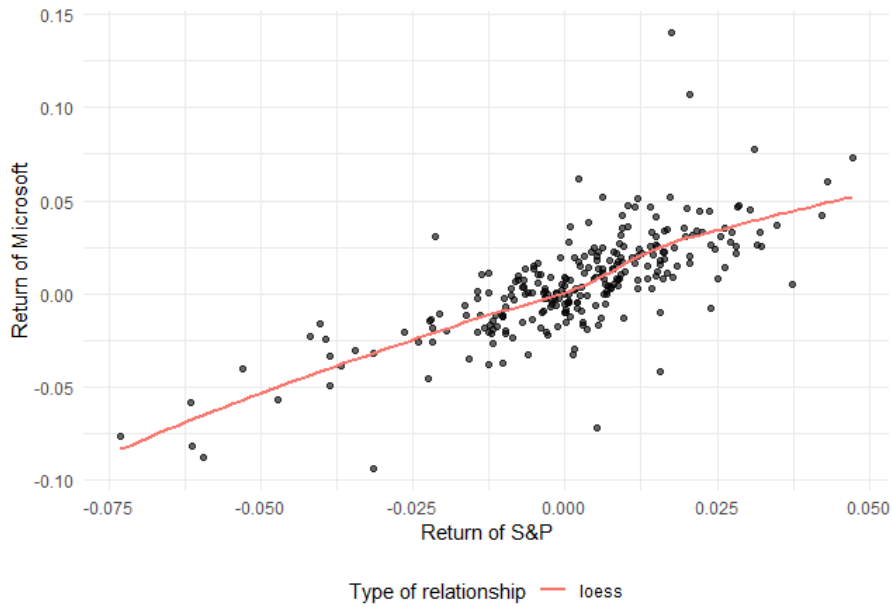


Figure 2. Adjusting Microsoft's performance to S&P's performance using a LOESS model

We can see that there is a slight parabola near zero yields as well as towards low yields. So the relationship is not perfectly linear. (Figure 2)



Figure 3. Comparison of adjustments

MEASURING THE IMPACT OF THE STOCK MARKET INDEX RETURN

If only Microsoft's negative returns are selected (Figure 3), the relationship between the two returns is no longer linear, but becomes parabolic. An ordinary least squares adjustment will push us to impute the data that we consider aberrant, with a linear view, whereas the relationship that describes the links between the two variables tends towards a quadratic relationship. From this, we retain the LOESS model.

The second element necessary for our approximation is to have the β , as $\beta = \dot{M}(\theta)/\phi\{\Phi^{-1}(\alpha)\}$, as we assume that $\dot{M}(\theta)$ is known.

So the first step is to have an approximation of $M(x)$. Since we have a history, we can define the variable y as follows

$$y = \begin{cases} 1 & \text{if } r_m > VaR_{95\%} \\ 0 & \text{Else} \end{cases} \quad (20)$$

With.

- r_m Microsoft's performance
- $VaR_{95\%}$ Value at risk at the 95% level which is -3.73% in our case, using the historical method.

Therefore, we can express the relationship between the values of de $r_{S\&P}$ and y trough the logistic regression.

A relationship which is formulated as follows.

$$P(y = 1|r_{S\&P}) = \frac{1}{1 + \exp(-(\alpha r_{S\&P} + c))} \quad (21)$$

With α , et c , the parameters estimated using logistic regression.

The result of the estimation of the model coefficients is as follows (Table 2).

Table 2. Result of the maximum likelihood estimation of the parameters of the logistic regression

Results	
Dependent variable.	
y	
<i>Returne of S&P</i>	98.786*** (18.608)
<i>Intercept</i>	4.062*** (0.527)
Observations	261
Log Likelihood	-30.904
Akaike Inf. Crit.	65.807
Note.	*p<0.1; **p<0.05; ***p<0.01

Thus .

$$P(y = 1|r_{S\&P}) = \frac{1}{1 + \exp(-(98.786r_{S\&P} + 4.062))} \quad (22)$$

In addition, one can also approximate $\dot{M}(r_{S\&P})$ by the function

$$m(r_{S\&P}) = \frac{98.786 \exp(-(98.786r_{S\&P} + 4.062))}{(1 + \exp(-(98.786r_{S\&P} + 4.062)))^2} \quad (23)$$

Therefore, one can approximate $\dot{M}(\theta)$ by $m(\theta)$. Furthermore, to have an approximation of θ associated with each α , we can use the solution of the following equation

$$\frac{1}{1 + \exp(-(98.786r_{S\&P} + 4.062))} - \alpha = 0 \quad (24)$$

Moreover, even if this is only an approximation of the actual value of $\dot{M}(\theta)$ Joseph has demonstrated that the process converges towards the target value of θ , regardless of the value of $\dot{M}(\theta)$ provided [4]. However, an estimate far from the real value should cause a slowing down of the convergence. In our case the slowdown should be minimal, as our estimates are based on available data.

The last element to be fixed is the $VarR_q$ to define y_n , in our case, as well as the definition of y , the historical VAR at the 95% level is -3.73%. So

$$y_n = \begin{cases} 1 & \text{if } f(x_n) > -3.73\% \\ 0 & \text{Else} \end{cases} \quad (25)$$

The definition of y_n causes the approximation to return the value of x_n , will be greater than the VAR, with a given probability α . Therefore, it returns the value of x_n which will be lower than the VAR with a probability $1-\alpha$.

3. MAIN RESULTS

3.1 Simulation

In our case we take x_1 as the solution to the equation

$$\frac{1}{1 + \exp(-(98.786r_{S\&P} + 4.062))} - \alpha = 0 \quad (26)$$

And $\tau_1 = 0.026$ the standard deviation of $r_{S\&P}$ values when the VAR occurred. And after 100 iterations, we get the following results. The graph below (Figure 4) shows the different values of x_n as a function of the different possible alpha values.

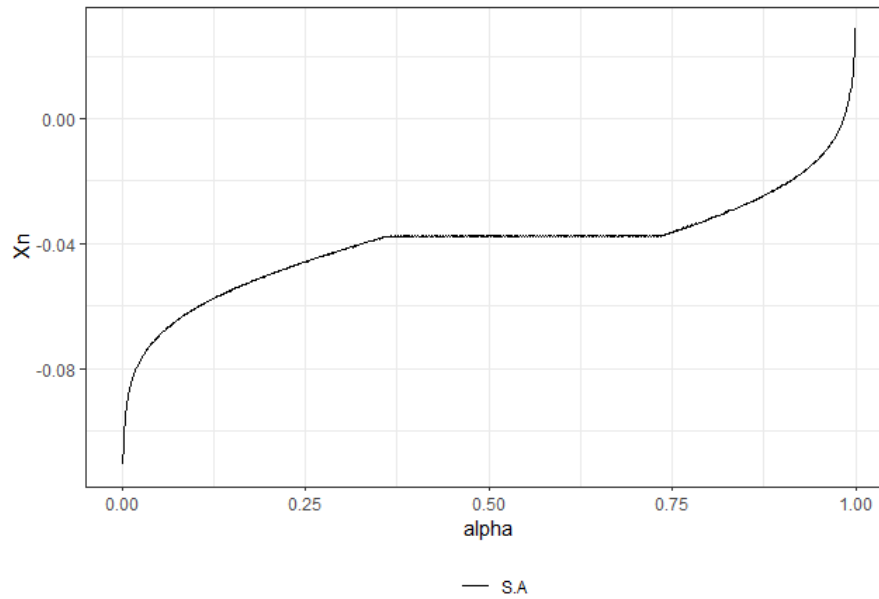


Figure 4. Result of the stochastic approximation as a function of each alpha

The black curve is the result of the S.A. We can see that the S.A. has converged directly to the LOESS model solution and this, independently of the value of α . This is due to our definition of y_n which only takes into account the systemic risk factor that is represented by the S&P return, whereas it is the intrinsic risk that causes volatility.

To overcome this problem, we define $y_n(e_n)$ as follows

$$y_n(e_n) = \begin{cases} 1 & \text{if } f(x_n) + e_n > -3.73\% \\ 0 & \text{Else} \end{cases} \quad (27)$$

With e_n a randomly simulated value according to the distribution of the residuals of the LOESS regression. The next step therefore is to determine the error distribution of our model.

The density function estimated from the residuals can be seen below (Figure 5).

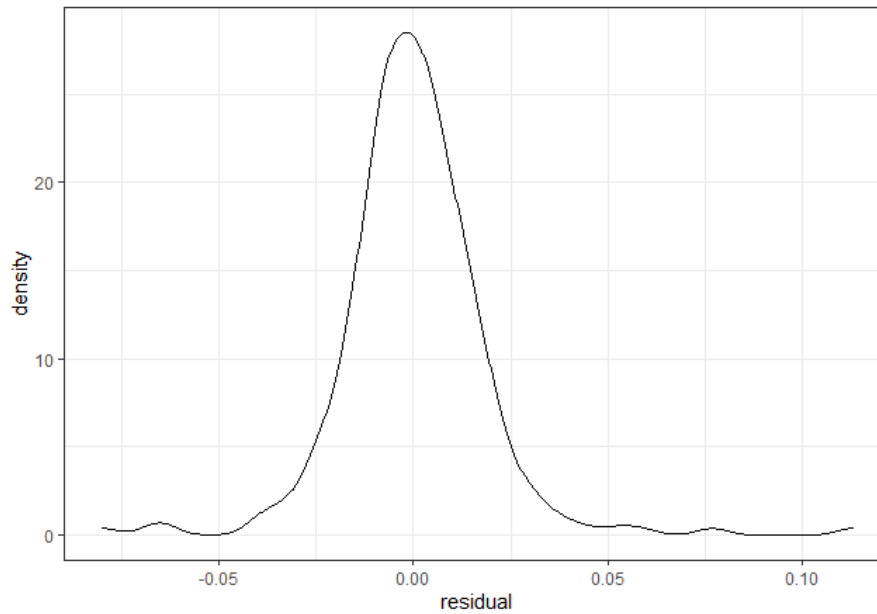


Figure 5. distribution of regression residuals

It can be seen that the curve is closer to a Student distribution than the normal distribution.

Furthermore, the result of Student's Kolmogorov Smirnov test for standardised residuals using a scale=0.01, location=0 with 3 degrees of freedom showed that the test statistic $D = 0.026157$, p-value = 0.9941. Therefore, we cannot reject the hypothesis $H_0 =$ "the residuals follow a Student distribution". Thus, we can see below (Figure 6) the result of the stochastic approximation by adding a noise for the calculation of y_n after 100 iterations.

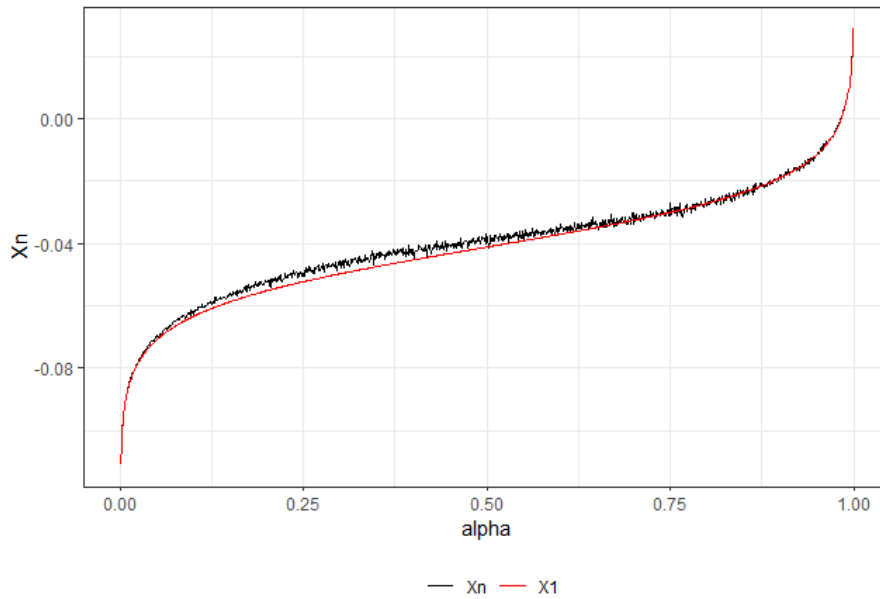


Figure 6. Result of stochastic approximation with noisy forecasts

Firstly, we notice that the results of the approximation are more homogeneous. Indeed, there was an abrupt variation in the results of the first stochastic approximation, whereas in the second approximation a sigmoid is obtained as the result of the S.A. Moreover, there was a convergence of the values of x_n towards their target values for alpha lower than 0.6, whereas the rest of the x_n remained stable. Thus, our estimate of x_1 was quite close to the target value for these alphas. The second point is that the values of x_n as a function of alpha reflect reality as well. Indeed, we can see below (Figure 7) the graph of Microsoft's return versus S&P's return, but only in the case where Microsoft's return suffers losses below VaR.

MEASURING THE IMPACT OF THE STOCK MARKET INDEX RETURN

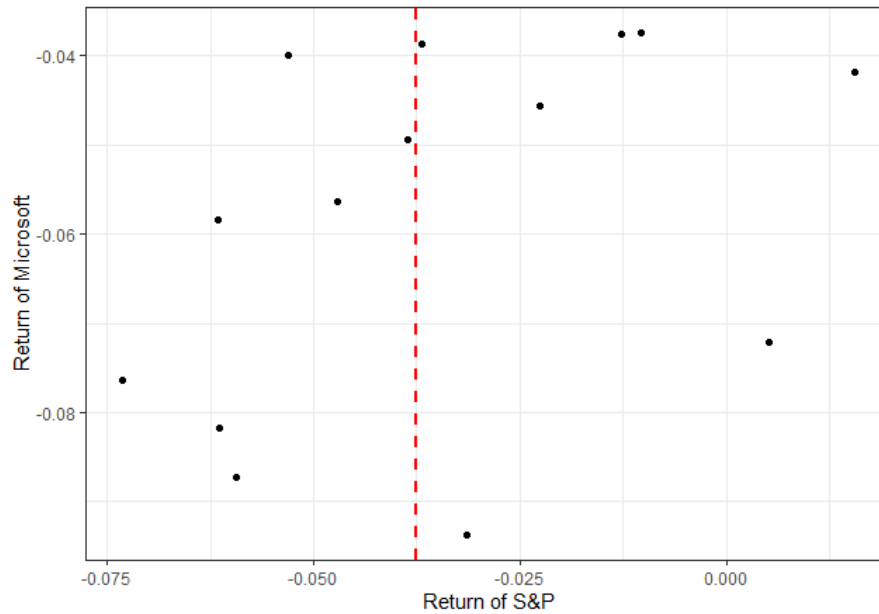


Figure 7 . Microsoft's performance versus S&P's performance for extreme losses

It can therefore be clearly seen that the results of the stochastic approximation are consistent with Microsoft's historical performance. Indeed, for Alpha close to 1, the S.A. clearly indicates a low but positive return at a $1-\alpha$ chance of causing losses exceeding the VAR threshold at Microsoft, and the same is true for Alpha close to zero. The red lines represent the result of the S.A. for alpha equal to 0.5. It should be noted that 50% of the occurrences of Microsoft's losses had an associated S&P return greater than the estimate for alpha equal to 0.5.

3.2 Discussion

The estimation method chosen is the method proposed by Joseph, the latter had demonstrated that his algorithm gives better results than the classical Robbins and Monro algorithm even for extreme alpha values, and this for a low number of iterations but with a high root mean square error for extreme alpha values, but still lower than the basic algorithm. We cannot verify this finding because the target value of our process is unknown. However, 100 iterations are not enough to converge, as can be seen in the graph below (Figure 8) after 1000 iterations.

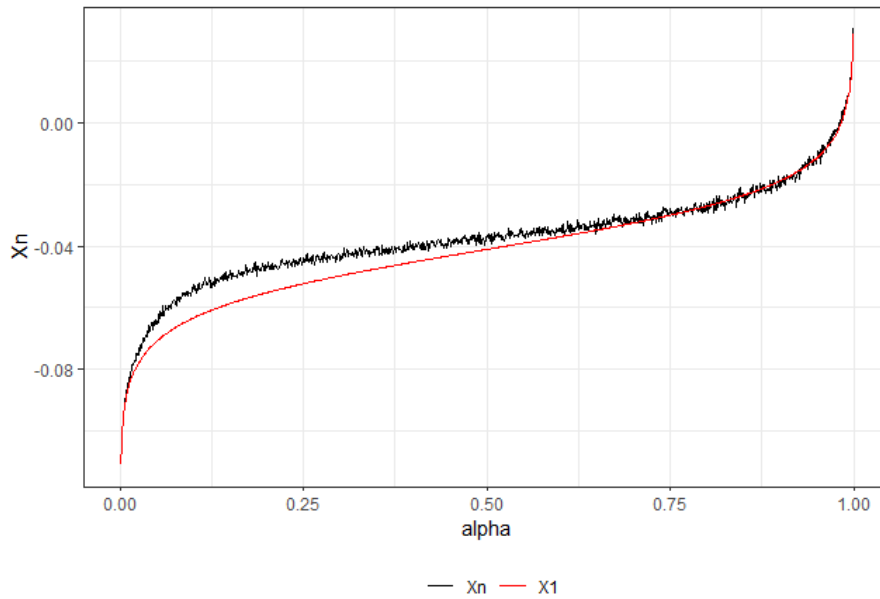


Figure 8. Results of the stochastic approximation after 1000 iterations.

It can be seen that the values of the x_n estimate for alphas below 0.6 have been refined. This is due to the low speed of convergence that can be caused by the same value τ_1 independently of alpha. Several methods have been proposed to overcome this convergence speed problem, such as those of [11], [12] and [13]. They have shown that their improved versions return a lower MSE than Joseph's version for low iterations, especially for extreme alpha values. But, which are relatively heavier in terms of compilations, and require several iterations to calibrate particular parameters. Especially since in our case, we have to be more interested in the time needed for compilation than the number of iterations performed given the power of the compilers available. All the more so as adding iterations allows us to better control the effect of the noise we add to the Microsoft performance predictions associated with x_n , since after 100 iterations, two compilations can give quite different estimates for the extreme values of alpha, whereas after 1000 iterations for example the SA results are stable.

CONCLUSION

The results obtained represent a significant asset to be taken into consideration by investors before trading on the stock market. Therefore, sell when the return of the stock market index is dangerously close to the results obtained by the stochastic approximation, or buy if the return of the stock market index is far from our results and thus be sure not to suffer significant losses. Moreover, this work can be duplicated for any value return. Therefore, at time "t" we can judge the state of the economy and make our forecasts of the evolution of the stock market index. We can set a target positive return, for example, and thus know whether our forecasts will generate the targeted return, and thus buy the stock at a low price to be able to resell it at a more advantageous price. It would also be interesting to apply the stochastic approximation, but this time not to a single stock, but to a portfolio of stocks, trying to capture the covariance between the different stocks that make up the portfolio.

MATERIALS AND METHODS

All data used to conduct this study was obtained through Yahoo Finance. In addition, the theory discussed in this article has been transcribed on the programming language "R" with functions native to it. As for the graphs, they were drawn using the functions of the "Ggplot2" package of this language.

DATA AVAILABILITY

1. The Microsoft stock price data used to support the findings of this study can be accessed free of charge on « <https://finance.yahoo.com/quote/MSFT/history?p=MSFT> »
2. The S&P data used to support the findings of this study can be accessed free of charge on «<https://finance.yahoo.com/quote/%5EGSPC/history?p=%5EGSPC>»

CONFLICT OF INTERESTS

The author(s) declare that there is no conflict of interests.

REFERENCES

- [1] H. Markowitz, Portfolio Selection, *J. Finance*, 7 (1952), 77-91
- [2] Basel Committee on Banking Supervision, Revisions to the Basel II market risk framework, (2009).
- [3] P. Jorion, Value at risk: the new benchmark for managing financial risk, McGraw-Hill, New York, (2007).
- [4] H. Robbins, S. Monro, A stochastic approximation method, *Ann. Math. Stat.* 29 (1951), 373-405.
- [5] J. Roshan, Efficient Robbins--Monro procedure for binary data. *Biometrika*, 91 (2004), 461-470.
- [6] O.A. Bardou, N. Frikha, Computation of VaR and CVaR using stochastic approximations and unconstrained importance sampling. *Monte Carlo Meth. Appl.* 15 (2009), 173-210.
- [7] D. Barrera, S. Crépey, B. Diallo, et al. Stochastic approximation schemes for economic capital and risk margin computations, *ESAIM: Proc. Surv.* 65 (2019), 182-218.
- [8] W.F. Sharpe, Capital asset prices: a theory of market equilibrium under conditions of risk, *J. Finance*, 19 (1964), 425-442.
- [9] W. S. Cleveland, Robust locally weighted regression and smoothing scatterplots, *J. Amer. Stat. Assoc.* 74 (1979), 829-836.
- [10] W.S. Cleveland, S.J. Devlin, Locally weighted regression: an approach to regression analysis by local fitting, *J. Amer. Stat. Assoc.* 83 (1988), 596-610.
- [11] Y. Tian, C.J. Wu, A skewed version of the Robbins-Monro-Joseph procedure for binary response, *Stat. Sin.* 25 (2015), 1679-1689.
- [12] C.J. Wu, Y. Tian, Three-phase optimal design of sensitivity experiments, *J. Stat. Plan. Inference*, 149 (2014), 1-15.
- [13] A.G. Joseph, S. Bhatnagar, An adaptive and incremental approach to quantile estimation, in: 2019 IEEE 58th Conference on Decision and Control (CDC), IEEE, Nice, France, 2019: pp. 6025–6031.