Capturing the intrinsic uncertainty of the VaR: Spectrum representation of a saddlepoint approximation for an estimator of the VaR

Dominique Guegan, Bertrand Hassani, Kehan Li

UNIVERSITÉ PARIS 1 PANTHÉON SORBONNE

Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers

Capturing the intrinsic uncertainty of the VaR: Spectrum representation of a saddlepoint approximation for an estimator of the VaR

Dominique GUEGAN, Bertrand HASSANI, Kehan LI

2016.34R

Version révisée

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
http://centredeconomiesorbonne.univ-paris1.fr/
ISSN : 1955-611X

Capturing the intrinsic uncertainty of the VaR: Spectrum representation of a saddlepoint approximation for an estimator of the VaR

Dominique Guégan · Bertrand Hassani · Kehan Li

Received: date / Accepted: date

Abstract Though risk measurement is the core of most regulatory document, in both financial and insurance industries, risk managers and regulators pay little attention to the random behaviour of risk measures. To address this uncertainty, we provide a novel way to build a robust parametric confidence interval (CI) of Value-at-Risk (VaR) for different lengths of samples. We compute this CI from a saddlepoint approximation of the distribution of an estimator of VaR. Based on the CI, we create a spectrum representation that represents an area that we use to define a risk measure. We apply this methodology to risk management and stress testing providing an indicator of threats caused by events uncaptured in the traditional VaR methodology which can lead to dramatic failures.

Keywords Financial regulation · Value-at-Risk · Order statistic · Uncertainty · Saddlepoint approximation · Stress testing
JEL classification C14 · D81 · G28 · G32

1 Introduction

In both the financial and insurance industries, the computation of the risk is an important activity. Most of the work is done using a conventional risk measure,

D. Guégan
Université Paris 1 Panthéon-Sorbonne, CES UMR 8174, 106 bd l'Hopital 75013, Paris, France.
Labex Refi & Associate researcher to IPAG.
E-mail: dominique.guegan@univ-paris1.fr

B. Hassani
Grupo Santander; Université Paris 1 Panthéon-Sorbonne, CES UMR 8174 & Labex Refi.
E-mail: bertrand.hassani@gmail.com

K. Li (✉)
Université Paris 1 Panthéon-Sorbonne, CES UMR 8174, 106 bd l'Hopital 75013, Paris, France.
Labex Refi.
Tel.: +33-614844745
Fax: +33-144078336
E-mail: kehanleex@gmail.com

say the Value-at-Risk (VaR), introduced by J.P. Morgan (Morgan 1997 [11]).

In this paper we investigate the way that is used to obtain a value of the risk associated with a bank or an insurance company through the concept of Value-at-Risk. Indeed, given a risk factor X characterized by a distribution function F_θ, the risk associated with this risk factor, for a given p, $0 < p < 1$, is not more that the quantile ψ_p :

$$P_{F_\theta}[X > \psi_p] = 1 - p \tag{1}$$

computed from the distribution law F_θ. In practice F_θ is not known, and in a common way we estimate ψ_p: given an information set on the risk factor X, $\{X_1, \cdots, X_n\}$, we rank them and obtain $X_{(1)} \leq \cdots \leq X_{(n)}$. We define $X_{(m)}$, where $m = np$ if np is an integer and $m = [np] + 1$ otherwise [1], and $X_{(m)}$ is a consistent estimator of ψ_p (Serfling 2001 [15]).

We are interested in investigating the properties of this random variable (r.v.) $X_{(m)}$ to provide a new way of risk measuring based on an area built on the confidence interval(CI) associated with ψ_p, in order to determine an alert indicator for bank and insurance risk managers in presence of large risky losses. To attain this objective we need to know the distribution law of $X_{(m)}$.

The investigation of the properties of the r.v. $X_{(m)}$ is not new. There have been good results on the asymptotic distribution of the r.v., obtained by Smirnov (1949) [16], Rao (2002) [14] and Zhu and Zhou (2009) [17] that we recall and extend later. Empirical works in risk management have also been carried out. Most of the works which investigate the properties of the estimate of the risk measure VaR assume that the underlying distribution F_θ is known (Gaussian or Student-t), and use a Gaussian approximation for the distribution of the $X_{(m)}$ (Jorion 1996 [8]); other papers use techniques based on Bootstrap to obtain the distribution of $X_{(m)}$ (Pritsker 1997 [13]; Christorffersen and Gonçalves 2005 [4]). All these approaches introduce a bias in the building of the CI associated with these different techniques. When the CI depends on simulations or Bootstrap, instability is introduced; when the CI depends on an asymptotic approximation, the bias is introduced by the computation of the variance which depends on the original distribution F_θ, and the sample size n.

In this paper we proceed in a different way, we assume that the unknown distribution law F_θ is unknown and verifies smooth conditions. Then, considering the results of Smirnov (1949) [16] and Zhu and Zhou (2009) [17] which provide the asymptotic distributions of $X_{(m)}$ [2], we derive two approximations of the distribution of $X_{(m)}$ using parametric and nonparametric estimates of F_θ. These two approaches allow the CI associated with $X_{(m)}$ to be determined directly. The main difference with the previous cited works comes from the role of the sample size n.

[1] $[x]$ denotes the largest integer less than or equal to x. $X_{(m)}$ is also called the mth order statistic, which is a fundamental tool in nonparametric statistics.

[2] The asymptotic distribution of the maxima $X_{(n)} = max(X_1, ..., X_n)$, which is a special case of $X_{(m)}$, has been discussed (see Fisher-Tippett theorem) and implemented in risk measurement and stress testing, see Embretch et al. (1997) [5] and Longin (2000) [10]. However, in this paper, we discuss general $X_{(m)}$.

Indeed, the speed of convergence of the asymptotic distribution obtained with Zhu and Zhou (2009) [17] using saddlepoint approach is quicker than the approximation obtained by Smirnov (1949) [16], then it is possible to build a robust CI for $\psi(p)$ even with small samples, which is a key-point for risk managers. The building of the CI under general assumptions on F_θ (we can work for instance with heavy tail distributions) and different values for n provides a way to build a new spectrum risk measure which risk managers could easily use to control their risks. Our approach does not need simulations, boostrapping or specific assumptions on the distribution on F_θ. A simulation experiment compares the both approaches of Smirnov (1949) [16] and Zhu and Zhou (2009) [17]. We exhibit the good performance in the fitting of the asymptotic distribution with the saddlepoint approach. In particular this method works well with p closing to 0 or 1, asymmetric distribution, fat-tailed behaviour and small samples. Finally, we use this new spectrum risk measure as an alert indicator for risk managers in case of extreme events: an empirical example illustrates this novel and interesting point.

This paper is organized as follows. Section 2 describes the Smirnov (1949) [16] and Zhu and Zhou's (2009) [17] approximations for the asymptotic distribution of $X_{(m)}$. We provide a comparison of the performance of these two approximations by simulations. Section 3 computes the CI of $\psi(p)$ using these two approximations based on different data sets, and a stress testing application is provided. Section 4 concludes.

2 Asymptotic distribution of $X_{(m)}$

Consider a r.v. X (for example the return of a portfolio, the return of a risk factor or an operational loss), with a cumulative distribution function (cdf) F_θ (f_θ is the associated probability density function (pdf) and θ are the parameters). Let $X_1, ..., X_n$ be an independent and identically distributed sample (i.i.d) [3] from X. It is well known that the distribution of the order statistic $X_{(m)}$ is:

$$P(X_{(m)} \leq x) = \sum_{k=0}^{m-1} C_n^k F_\theta(x)^{n-k}(1 - F_\theta(x))^k. \tag{2}$$

Nevertheless, to realise the objective of this paper which is to build a CI around the VaR, we use asymptotic approximations of the distribution of $X_{(m)}$ provided by Smirnov (1949) [16] and Zhu and Zhou (2009) [17]. Indeed, the expression (2) is complicated to use when F_θ is unknown and m is large enough.

Theorem 1 (Asymptotic normality approximation (Smirnov, 1949))
Given a r.v. X with a continuous and differentiable cdf F_θ and a density f_θ strictly positive at $F_\theta^{-1}(p)$, then

$$\sqrt{n}\frac{X_{(m)} - F_\theta^{-1}(p)}{\sqrt{V}} \rightarrow_{(d)} N(0,1), \quad as \quad n \longrightarrow \infty \tag{3}$$

[3] Or if they are not, we assume that we can transform them to an i.i.d set by filtering

where $\to_{(d)}$ *means convergence in distribution,* $V = \frac{p(1-p)}{f_\theta(F_\theta^{-1}(p))^2}$*.* $N(0,1)$ *represents the standard Gaussian distribution.*

The expression (3) depends on the values of $\frac{1}{f_\theta(F_\theta^{-1}(p))}$, which are unknown in most cases. Therefore we need to estimate it. We can use a nonparametric approach, for instance the Siddiqui-Bloch-Gastwirth estimator:

$$S_{r,n} = (\frac{n}{2r})(X_{(m+r)} - X_{(m-r)}) \tag{4}$$

Hall and Sheather (1988) [7] give an Edgeworth expansion approximation for the distribution of the studentized r.v. $\sqrt{n}\frac{X_{(m)}-F_\theta^{-1}(p)}{p^{1/2}(1-p)^{1/2}S_{r,n}}$. Kaplan (2015) [9] proposes a test for the optimal choice of a smoothing parameter r, which is crucial for this approximation.

Instead of using such a nonparametric estimator, we can also fit a panel of distributions using $X_1, ..., X_n$ and compute the estimators of θ, denoted $\hat{\theta}$ by maximum likelihood, then $F_{\hat{\theta}}$ and $f_{\hat{\theta}}$ are the estimators of F_θ and f_θ. By plugging $F_{\hat{\theta}}$ and $f_{\hat{\theta}}$ in expression (3), we provide the following corollary:

Corollary 1 (Plug-in AN approximation)
Given a r.v. X *whose cdf* F_θ *and density* f_θ *are continuous functions with respect to (w.r.t)* θ*, and* $\hat{\theta}$ *is a consistent estimator of* θ [4]*, then*

$$\sqrt{n}\frac{X_{(m)} - F_\theta^{-1}(p)}{\sqrt{\widehat{V}}} \to_{(d)} N(0,1), \quad as \quad n \longrightarrow \infty \tag{5}$$

where $\widehat{V} = \frac{p(1-p)}{[f_{\hat{\theta}}(F_{\hat{\theta}}^{-1}(p))]^2}$*.*

The proof is presented in Appendix A.

In a recent paper, Guégan et al. (2015) [6] show that the AN obtained in the previous theorem depends strongly on the properties of the distribution F_θ (asymmetric and fat-tailed behaviour) and on the size n. Consequently, using it to build build a CI for $\psi(p)$ created a potential bias that was difficult to anticipate, making the new methodology proposed in this paper not as reliable as we expected for proposing a new risk measure. It is the reason why we now introduce another way to obtain the asymptotic distribution of $X_{(m)}$ which is free of the behaviour of the distribution F_θ, and which does not depend so much on the sample size n. Based on the integral representation of the binomial distribution and Barndorff-Nielsen formula (Barndorff-Nielsen 1991 [2]; Ma 1998 [12]), Zhu and Zhou (2009) [17] derives an asymptotic Saddlepoint approximation (SP) for the distribution of $X_{(m)}$:

Theorem 2 (Saddlepoint approximation (Zhu and Zhou, 2009))
Given a r.v. X *with cdf* F_θ*, for* $\forall \epsilon > 0$ *and* $\forall x$ *in the domain of* F_θ*, we assume* $\epsilon < p < 1-\epsilon$*. Then, if* $r_0 = \frac{m}{n}$*,* $F_\theta(x) = t$ *and* $t \neq p$*, we have:*

[4] Asymptotically consistent estimator means $\hat{\theta} \to_{(P)} \theta$, where $\to_{(P)}$ represents convergence in probability

$$X_{(m)} \rightarrow_{(d)} \Psi(\sqrt{n}\omega^\sharp) \quad as \quad n \longrightarrow \infty \tag{6}$$

where the convergence speed is $O(\frac{1}{n})$ uniformly w.r.t x. Φ denotes the cdf of standard Gaussian distribution ($N(0,1)$)

$$\begin{aligned} &\Psi(\sqrt{n}\omega^\sharp) = 1 - \Phi(\sqrt{n}\omega^\sharp) && \omega^\sharp = \omega + \frac{1}{n\omega} ln \frac{1}{\psi(-\omega)} \\ &\psi(-\omega) = \frac{\omega(t-1)}{t-r_0}(\frac{r_0}{1-r_0})^{\frac{1}{2}} && \omega = -\sqrt{2h(t)} sign(t-r_0) \\ &h(t) = r_0 ln \frac{r_0}{t} + (1-r_0) ln \frac{1-r_0}{t} \end{aligned} \tag{7}$$

where $sign(x) = 1$ if $x \geq 0$ and $sign(x) = -1$ otherwise. For $t = p$

$$P(X_{(m)} \leq x) \approx \frac{1}{2} + \frac{1}{\sqrt{2\pi n}} \frac{1+r_0}{3r_0} (\frac{r_0}{1-r_0})^{\frac{1}{2}} \quad as \quad n \longrightarrow \infty. \tag{8}$$

The proof of Theorem 2 is provided by Zhu and Zhou (2009) [17]. In order to use the result of this theorem to build a CI of $F_\theta^{-1}(p)$, we do a linear transformation in the following way: for $\forall x$ fixed,

$$X_{(m)} - F_\theta^{-1}(p) \rightarrow_{(d)} \Psi(\sqrt{n}\omega^\sharp) \quad as \quad n \longrightarrow \infty \tag{9}$$

where we substitute t in Theorem 2 by $F_\theta(x + F_\theta^{-1}(p))$. By plugging $F_{\hat{\theta}}$ and $f_{\hat{\theta}}$ in expression (7), we introduce the following corollary:

Corollary 2 (Plug-in and linear transformed SP approximation)
Assume F_θ and f_θ are continuous functions with respect to (w.r.t) θ, and $\hat{\theta}$ is an consistent estimator of θ. For $\forall \epsilon > 0$ and $\forall x$ in the domain of F_θ, we assume $\epsilon < p < 1 - \epsilon$. Let $r_0 = \frac{m}{n}$, $F_{\hat{\theta}}(x + F_{\hat{\theta}}^{-1}(p)) = t$, then for $t \neq p$

$$X_{(m)} - F_\theta^{-1}(p) \rightarrow_{(d)} \Psi(\sqrt{n}\widehat{\omega}^\sharp) \quad as \quad n \longrightarrow \infty \tag{10}$$

where the convergence speed is $O(\frac{1}{n})$ uniformly w.r.t x.

$$\begin{aligned} &\Psi(\sqrt{n}\widehat{\omega}^\sharp) = 1 - \Phi(\sqrt{n}\widehat{\omega}^\sharp) && \widehat{\omega}^\sharp = \widehat{\omega} + \frac{1}{n\widehat{\omega}} ln \frac{1}{\psi(-\widehat{\omega})} \\ &\psi(-\widehat{\omega}) = \frac{\widehat{\omega}(t-1)}{t-r_0}(\frac{r_0}{1-r_0})^{\frac{1}{2}} && \widehat{\omega} = -\sqrt{2h(t)} sign(t-r_0) \\ &h(t) = r_0 ln \frac{r_0}{t} + (1-r_0) ln \frac{1-r_0}{t}. \end{aligned} \tag{11}$$

For $t = p$

$$P(X_{(m)} \leq x) \approx \frac{1}{2} + \frac{1}{\sqrt{2\pi n}} \frac{1+r_0}{3r_0} (\frac{r_0}{1-r_0})^{\frac{1}{2}}. \tag{12}$$

The proof of Corollary 2 is similar to that of Corollary 1. Comparing Theorems 1 and 2, we find that the convergence speed of the SP approximation ($O(\frac{1}{n})$) is faster than the speed of the AN approximation ($O(\frac{1}{\sqrt{n}})$). It means the SP approximation can be more accurate, especially when we have a small sample. It appears that the use of SP approach is more robust if we wish to approximate the distribution of $X_{(m)}$.

We now compare the performance of the convergence obtained in Theorem 1 and Theorem 2 by simulation, by checking different sample sizes and distributions for F_θ. We use for F_θ, a $N(0,1)$ distribution, a NIG_0 (Normal-inverse Gaussian, Barndorff-Nielsen (1978) [1]) distribution [5] and a GEV (Generalized extreme value, Embretch et al. (1997) [5]) distribution [6]. These distributions belong respectively to elliptical distribution family, Generalised hyperbolic distribution family and extreme value distribution family. These different families of distributions mean that the main features of the financial data sets can be identified.

Our simulation plan is as follows: for different n and p [7], we simulate $n * 1000$ random numbers from one of the previous distributions. Then given $0 < p < 1$, we take the realizations of the $X_{(m)}$ for every n points to have 1000 realizations. We compute the K-S statistic [8] and A-D statistic[9] between the AN (or SP) approximation and the empirical cdf (ecdf) of $X_{(m)}$. The results for K-S statistic are provided in Table (2), and for A-D statistic in Table (3).

In Tables (2) and (3), the results provided by SP approximation are always smaller than those provided by AN approximation, for all the three distributions whatever n and p are. Consequently, the SP approximation always performs the AN approximation. More precisely, when p is closer to 0 or 1 (like $p = 0.005$ or $p = 0.995$), the difference of accuracy between these two approximations is more obvious. When n is small, the SP approximation provides robust approximation but the AN does not.

To illustrate the previous results, we exhibit some graphs showing the accuracy of the SP approximation for small samples ($n = 241$ [10]). Given $p = 0.975$ with $N(0,1)$, on Figure (1) we plot: the solid line, which is the ecdf of $X_{(m)}$ (it is a benchmark). The dot-dash line is the AN approximation obtained from Theorem 1. The dash line is the SP approximation obtained from Theorem 2. We observe that the SP approximation is nearly on the ecdf. But the AN approximation is far

[5] The tail parameter is equal to 0.3250, skewness parameter is equal to $5.9248e-04$, location parameter is equal to $-1.6125e-04$ and scale parameter is equal to 0.0972.

[6] The shape parameter is equal to 0.8876698, scale parameter is equal to 2049.7625278 and location parameter is equal to 245.7930751.

[7] Here, $n = 11, 121, 241, 501, 1001, 10001, 30001$ and $p = 0.05, 0.01, 0.005, 0.95, 0.99, 0.995$.

[8] For an empirical cdf $F_n(x)$ and a cdf F: the Kolmogorov–Smirnov statistic is:

$$DKS_n = sup_x \mid F_n(x) - F(x) \mid, \tag{13}$$

$F_n(x)$ is the empirical cdf defined as $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i \leq x\}}$, where $1_{\{X_i \leq x\}} = 1$ if $X_i \leq x$ and 0 otherwise.

[9] the Anderson–Darling statistic is:

$$DAD_n = n \int_{-\infty}^{\infty} \frac{(F_n(x) - F(x))^2}{F(x)(1 - F(x))} dF(x) \tag{14}$$

[10] 241 is around the number of one year trading days of a stock market.

away from the ecdf.

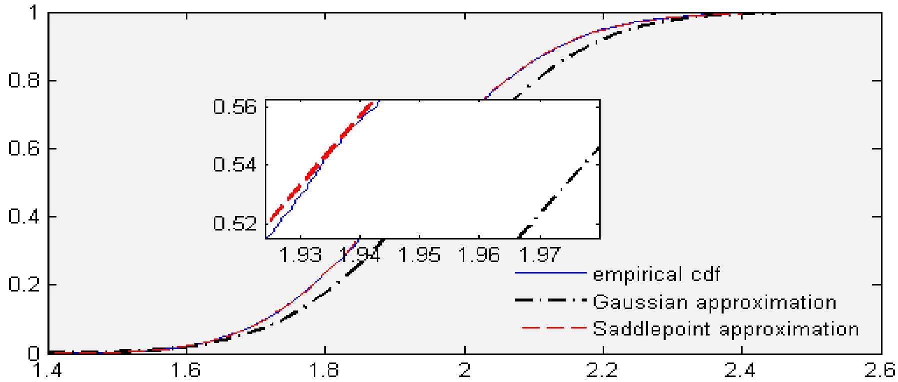

Fig. 1 Given $p = 0.975$ and $n = 241$, with $N(0, 1)$, we provide the solid line is the ecdf of $X_{(m)}$. The dot-dash line is the AN approximation from Theorem 1. The dash line is the SP approximation from Theorem 2.

Using a NIG_0 for F_θ and $p = 0.01$ (left tail), we plot the ecdf of $X_{(m)}$, the AN and SP approximations in the left graph of Figure (2). In the right graph we use a GEV distribution and $p = 0.995$ (right tail), to plot the ecdf of $X_{(m)}$, the AN and SP approximations. In Figure (2), the solid line is the ecdf. The dot-dash line is the AN approximation. The dash line is the SP approximation.

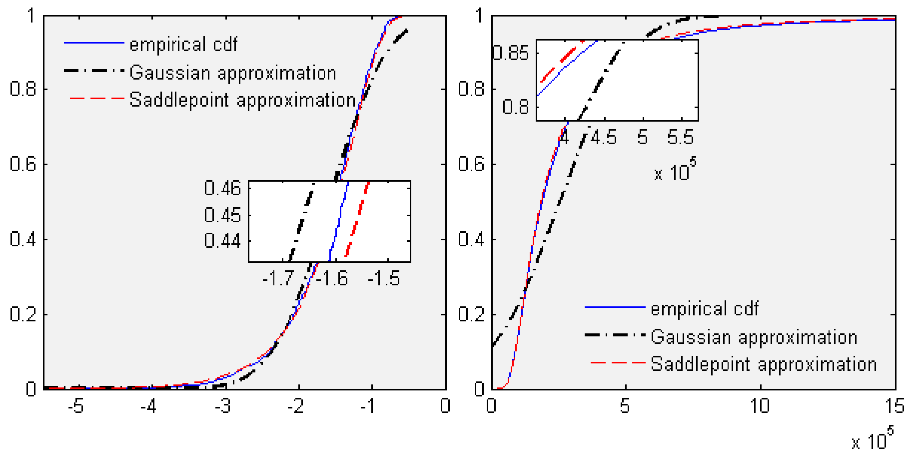

Fig. 2 With NIG_0 and $p = 0.01$, we plot the ecdf of $X_{(m)}$, the AN and SP approximation in the left graph of Figure 2. In the right graph with GEV and $p = 0.995$, we plot the ecdf of $X_{(m)}$, the AN and SP approximation. We provide that the solid line is the ecdf. The dot-dash line is the AN approximation. The dash line is the SP approximation.

We observe that in both graphs of Figure (2), the dash line is always closer to the solid line than the dot-dash line. That means for fat-tailed F_θ, the SP approximation is still more accurate than the AN approximation. In particular, the $X_{(m)}$ realisations simulated from a GEV distribution is asymmetric ($skewness = 9.8760$) and leptokurtic ($kurtosis = 176.0057$). Apparently, the AN approximation cannot fit these behaviours accurately, since it is symmetric and its tail is thin. Consequently, it is reasonable to use SP approximation in this case.

From Figures (1) and (2), we observe that no matter if we consider the left or right tail $X_{(m)}$ (p is close to 0 or 1), no matter if the F_θ is symmetric or asymmetric, fat-tailed or thin-tailed, the SP approximation is always more accurate. Besides, unlike the symmetric, thin-tailed AN approximation, the SP approximation means the asymmetric and fat-tailed behaviour of the distribution of $X_{(m)}$ can be captured. Consequently, we suggest that risk managers and regulators use the SP approximation to model the uncertainty in $X_{(m)}$, especially when the sample size is small (for example one year daily data), when the data set has a fat-tailed and asymmetric distribution and when p is close to 0 or 1 (like 0.001 or 0.999).

Given another confidence level $0 < q < 1$, based on the results of Corollaries 1 and 2, we build the CI around the true VaR ψ_p:

$$\left[X_{(m)} - z_{1-\frac{q}{2}}\hat{\sigma}, \quad X_{(m)} - z_{\frac{q}{2}}\hat{\sigma}\right] \tag{15}$$

$$\left[X_{(m)} - Z_{1-\frac{q}{2}}, \quad X_{(m)} - Z_{\frac{q}{2}}\right] \tag{16}$$

Using Corollary 1 we get a symmetric CI in expression (15), where $z_{\frac{q}{2}}$ is a quantile of N(0,1) and $\hat{\sigma}$ is a standard deviation from $\widehat{V}$. Using Corollary 2, we get a CI in expression (16), where $Z_q = \Psi_q^{-1}(\sqrt{n}\widehat{\omega}^\sharp)$. Ψ and $\widehat{\omega}^\sharp$ are provided in expressions (10) and (11). This bound Z_q will be generally obtained numerically [11]. The CI from Corollary 2 is either symmetric or asymmetric. For both CI (15) and (16), we recall that F_θ has to be estimated first. In Guégan et al. (2015) [6] we introduce the Spectrum Stress VaR (SSVaR) using the CI (15), denoted AN-SSVaR, which is an area delineated by the lower and upper bounds of CI, given two sequences $\{0 \le p_i \le 1\}_{i=1,\dots,k}$ and $\{0 \le q_j \le 1\}_{j=1,\dots,k}$, p_i corresponds to the confidence level of the VaR and q_j corresponds to the confidence level for the CI of the VaR. In this paper, we extend this new risk spectrum using the CI (16) from the saddlepoint approximation. We call it SP-SSVaR. Its bounds are delineated by the lower and upper bounds of CI (16). In the next section, we exhibit a practical application showing how we can use these CI for risk management.

3 Application: Using the CI for stress testing considering

Assessing the SP approximation of the distribution of $X_{(m)}$ in practice, we consider a fictive financial institution. This one holds two market portfolios (that is, the same stock components and weights as a benchmark index of a stock market):

[11] The code can be provided under request.

the Standard Poor's 500 (S&P 500) in the U.S. and the Hang Seng Index (HSI) in China. The S&P 500 represents a developed stock market in the U.S.; the HSI represents an emerging market in Asia. We consider the daily returns computed using daily closing prices of the portfolio. More precisely, four data sets are considered for different periods including three financial crises: Black Monday (1987), the Asian financial crisis (1997) and the global financial crisis (2007-08) [12]. In Table (1), we provide the first four empirical moments and the number of observations of these four data sets.

We build the two risk measures AN-SSVaR and SP-SSVaR, using data sets SP1 and HSI1. The distributions characterizing these data sets are asymmetric and leptokurtic, so we fit a NIG distribution on them [13]. Given $q = 0.01$ and $n = 252, 244$, we compute the bounds of AN-SSVaR and SP-SSVaR using the CI defined respectively in (15) and (16) with $0.001 \leq p \leq 0.05$. To compare the AN-SSVaR and the SP-SSVaR, we compute the ecdf of SP2 and HSI2 for a stress testing application (BCBS 2005 [3]).

			Empirical moments		
	points	mean	variance	skewness	kurtosis
SP1 (03/01/2008-31/12/2008)	252	-0.0015	0.0007	0.1841	6.8849
SP2 (03/01/1987-31/12/1987)	252	0.0003	0.0004	-4.0440	45.5834
HSI1 (03/01/1997-31/12/1997)	244	-0.0005	0.0006	0.7616	18.7190
HSI2 (03/01/1987-31/12/1987)	245	0.0000	0.0008	-6.7209	78.8165

Table 1 In this table, we provide the first four empirical moments of these four data sets and the number of observations.

In Figure (3), we use the data set SP1 to build the SSVaR. The dash curves are the bounds of AN-SSVaR. The solid curves are the bounds of SP-SSVaR. The dash-dot curve is the ecdf of the data set SP2. First, we observe that the lower bound of the AN-SSVaR is always smaller than the lower bound of the SP-SSVaR which means that with this former CI we accept higher risks. Indeed for risks between $0 < p < 0.003$, if we use as a risk measure the AN-SSVaR, we consider that the value of the risks provided with the ecdf line is acceptable; on the another hand if we consider as a risk measure the SP-SSVaR for $0 < p < 0.003$, it can be seen from the ecdf line that an alert is done. Second, when $p < 0.003$ the upper bound of the AN-SSVaR is outside the SP-SSVaR. It comes from the structural symmetric behaviour of the CI (15) which creates an inappropriate shape for this upper bound when p is smaller than 0.003.

[12] The data sets are: the daily return of S&P 500 over the period from 03/01/2008 to 31/12/2008 with 252 observations (denoted SP1); the daily return of S&P 500 over the period from 03/01/1987 to 31/12/1987 with 252 observations (denoted SP2); the daily return of HSI over the period from 03/01/1997 to 31/12/1997 with 244 observations (denoted HSI1); the daily return of HSI over the period from 03/01/1987 to 31/12/1987 with 245 observations (denoted HSI2); All the data sets have been obtained from Bloomberg.

[13] We use the moment method to get the estimates of NIG parameters. The estimates are provided in Appendix B

In Figure (4), we change the data set SP1 by HSI1 and SP2 by HSI2. We plot the graph again and the results are similar to those of Figure (3). We observe that the lower bound of the AN-SSVaR is always smaller than the lower bound of the SP-SSVaR which means that with this former CI we accept higher risks. Indeed for risks between $0 < p < 0.007$, if we use as a risk measure the AN-SSVaR, we consider that the value of the risks provided with the ecdf line is acceptable; on the another hand if we consider as a risk measure the SP-SSVaR for $0 < p < 0.007$, it can be seen from the ecdf line that an alert is done.

Consequently, it appears that as an alert indicator, the SP-SSVaR is more sensitive than the AN-SSVaR. It means the random behaviour of $X_{(m)}$ from different sample sizes can be incorporated robustly. Thus, risk managers can use it to evaluate a capital buffer to cover the risk embedded in the data set and the uncertainty of risk measure. In fact, by using the appropriate distribution of the VaR, it is possible to obtain more accurate confidence intervals for the VaR. Using the lower bounds of the confidence interval would lead to the creation of a more realistic Lower-VaR as this one would be more resilient to the occurrence of an extreme incident.

4 Conclusion

In this paper, to incorporate the random behaviour of $X_{(m)}$, we suggest using the SSVaR considering a saddlepoint approximation as an efficient risk measure. In particular, when we consider a small sample characterized by a fat-tailed distribution, and when we consider tail $X_{(m)}$, the random behaviour of $X_{(m)}$ affects the decision of risk managers. Indeed, theoretically the SP method can approximate the distribution of $X_{(m)}$ more accurately than the AN method. Consequently it provides more robust CI to build the SSVaR. To understand the performance of SP approximation comprehensively, we compared SP and AN approximations quantitatively and exhibited the difference graphically. The results of the simulation experiment were consistent with the theoretical result. Finally, we built the SSVaR from AN and SP approximations as alert indicators, and performed a stress testing application with data sets from different stock markets including three financial crises either in developed or in emerging markets. We found that compared to the alert indicator from AN approximations, the alert indicator from SP approximations is more sensitive to abnormal events. Therefore, in practice, we suggest risk managers and regulators use the SSVaR considering a SP approximation to incorporate the uncertainty of the VaR, particularly for those data sets in which the random behaviour of $X_{(m)}$ is not negligible, in order to control their risk.

5 Acknowledgments

This work was achieved through the Laboratory of Excellence on Financial Regulation (Labex ReFi) supported by PRES heSam under the reference ANR10LABX0095.

It benefited from a French government support managed by the National Research Agency (ANR) within the project Investissements d'Avenir Paris Nouveaux Mondes (investments for the future Paris New Worlds) under the reference ANR11IDEX000602.

References

1. Barndorff-Nielsen, O. (1978) Hyperbolic distributions and distributions on hyperbolae. Scandinavian Journal of Statistics, 5(3), 151-157.
2. Barndorff-Nielsen, O. (1991) Modified signed log likelihood ratio. Biometrika, 78(3), 557-563.
3. Basel Committee on Banking Supervision (2005) Amendment to the Capital Accord to incorporate market risks. Working Paper. http://www.bis.org/publ/bcbs119.pdf.
4. Christoffersen, P., Gonçalve, S. (2005) Estimation risk in financial risk management. Journal of Risk, 7(3), 1-28.
5. Embrechts, P., Kluppelberg, C., Mikosch T. (1997) Modelling Extremal Events for Insurance and Finance. Springer-Verlag Berlin Heidelberg.
6. Guégan, D., Hassani, B., Li K. (2015) The Spectral Stress VaR (SS-VaR). Centre d'Économie de la Sorbonne Working Paper. ftp://mse.univ-paris1.fr/pub/mse/CES2015/15052.pdf.
7. Hall, P., Sheather, S.J. (1988) On the Distribution of a Studentized Quantile. Journal of the Royal Statistical Society: Series B, 50(3), 381-391.
8. Jorion, P. (1996) Risk2: Measuring the risk in value at risk. Financial Analysts Journal, 52(6), 47-56.
9. Kaplan, D.M. (2015) Improved quantile inference via fixed-smoothing asymptotics and Edgeworth expansion. Journal of Econometrics, 185(1), 20-32.
10. Longin, F.M. (2015) From value at risk to stress testing: The extreme value approach. Journal of Banking & Finance, 24(7), 1097-1130.
11. Morgan, J.P. (1997) RiskMetrics Technical Documents (Fourth Edition). New York: J.P. Morgan Co.
12. Ma, C., Robinson J. (1998) Saddlepoint Approximation for Sample and Bootstrap Quantiles. Australian & New Zealand Journal of Statistics, 40(4), 479-486.
13. Pritsker, M. (1997) Evaluating Value at Risk Methodologies: Accuracy versus Computational Time. Journal of Financial Services Research, 12(2-3), 201-242.
14. Rao, C.R. (2002) Linear statistical inference and its applications. John Wiley & Sons, Hoboken, New Jersey.
15. Serfling, C.J. (2009) Approximation Theorems of Mathematical Statistics. John Wiley & Sons, Hoboken, New Jersey.
16. Smirnov, N.V. (1949) Limit Distributions for the Terms of a Variational Series. Trudy Mat. Inst. Steklov, 25.
17. Zhu, J., Zhou, W. (2009) Saddlepoint Approximation for Sample Quantiles with Some Applications. Communications in Statistics - Theory and Methods, 38(13), 2241-2250.

	$N(0,1)$		NIG_0			GEV	
	AN	SP	AN	SP		AN	SP
$p = 0.05$					$p = 0.95$		
$n = 11$	0.2183	0.0361	0.1933	0.0260	$n = 241$	0.0847	0.0265
$n = 121$	0.0767	0.0192	0.0876	0.0358	$n = 501$	0.0503	0.0265
$n = 241$	0.0509	0.0274	0.0621	0.0197	$n = 1001$	0.0756	0.0449
$n = 1001$	0.0435	0.0248	0.0276	0.0228	$n = 10001$	0.0371	0.0363
$n = 10001$	0.0322	0.0259	0.0208	0.0207	$n = 30001$	0.0368	0.0327
$p = 0.01$					$p = 0.99$		
$n = 11$	0.0757	0.0211	0.1963	0.0191	$n = 241$	0.1155	0.0205
$n = 121$	0.0847	0.0318	0.1072	0.0222	$n = 501$	0.1409	0.0370
$n = 241$	0.0948	0.0330	0.0564	0.0265	$n = 1001$	0.0988	0.0266
$n = 1001$	0.0456	0.0191	0.0546	0.0213	$n = 10001$	0.0459	0.0238
$n = 10001$	0.0204	0.0195	0.0227	0.0187	$n = 30001$	0.0327	0.0238
$p = 0.005$					$p = 0.995$		
$n = 11$	0.3930	0.0245	0.3959	0.0138	$n = 241$	0.1584	0.0278
$n = 121$	0.1711	0.0213	0.1418	0.0278	$n = 501$	0.1213	0.0253
$n = 241$	0.0653	0.0225	0.0744	0.0266	$n = 1001$	0.1461	0.0311
$n = 1001$	0.1071	0.0296	0.0976	0.0308	$n = 10001$	0.0515	0.0149
$n = 10001$	0.0367	0.0139	0.0359	0.0163	$n = 30001$	0.0392	0.0227

Table 2 For $N(0,1)$, NIG_0 and GEV, for $n = 11, 121, 241, 501, 1001, 10001, 30001$ and $p = 0.05, 0.01, 0.005, 0.95, 0.99, 0.995$, we compute the K-S statistic between the AN (or SP) approximation and the ecdf of $X_{(m)}$.

	$N(0,1)$		NIG_0			GEV	
	AN	SP	AN	SP		AN	SP
$p = 0.05$					$p = 0.95$		
$n = 11$	46.5478	0.3701	45.4794	0.5787	$n = 241$	4.1887	0.3820
$n = 121$	8.1637	0.4610	5.6806	0.3378	$n = 501$	2.9653	0.2765
$n = 241$	2.6894	0.2122	1.8976	1.0557	$n = 1001$	1.7426	0.3486
$n = 1001$	1.7430	0.3728	0.6408	0.7989	$n = 10001$	0.6837	0.6725
$n = 10001$	0.4598	0.2603	0.2827	0.1647	$n = 30001$	0.7047	0.5905
$p = 0.01$					$p = 0.99$		
$n = 11$	5.9214	0.1463	35.3727	0.3105	$n = 241$	16.2251	0.5430
$n = 121$	9.5424	0.8465	9.8631	0.2396	$n = 501$	17.6468	0.6223
$n = 241$	7.2525	0.3538	5.1074	0.2027	$n = 1001$	5.1793	0.6358
$n = 1001$	2.2158	0.4727	1.0760	0.4856	$n = 10001$	1.1022	0.0926
$n = 10001$	0.1742	0.1876	0.5491	0.2616	$n = 30001$	0.9970	0.4869
$p = 0.005$					$p = 0.995$		
$n = 11$	181.0830	0.2873	180.4938	0.1900	$n = 241$	43.3808	0.4520
$n = 121$	27.1190	0.3710	27.9483	0.1365	$n = 501$	20.2579	2.0985
$n = 241$	4.6398	0.1650	10.7707	0.5890	$n = 1001$	17.7949	1.2266
$n = 1001$	10.0369	0.2362	9.6725	0.4381	$n = 10001$	5.4230	1.6109
$n = 10001$	1.1836	0.2042	1.8859	0.3874	$n = 30001$	2.1239	0.9011

Table 3 For $N(0,1)$, NIG_0 and GEV, for $n = 11, 121, 241, 501, 1001, 10001, 30001$ and $p = 0.05, 0.01, 0.005, 0.95, 0.99, 0.995$, we compute the A-D statistic between the AN (or SP) approximation and the ecdf of $X_{(m)}$.

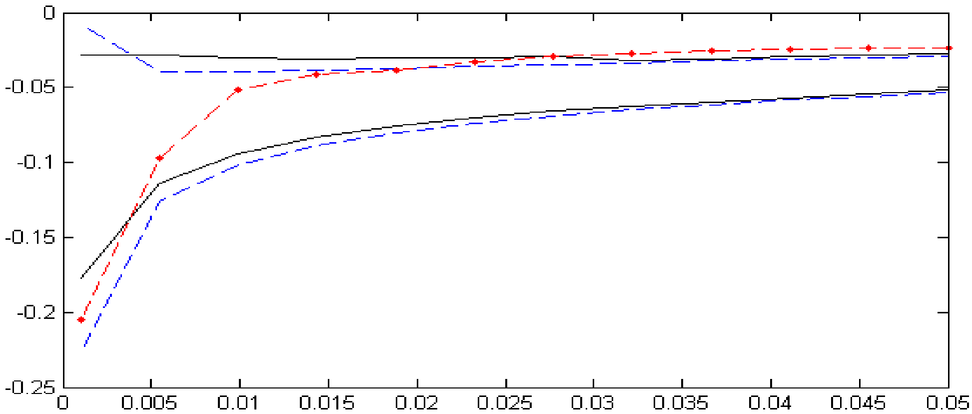

Fig. 3 In this figure, we use the data set SP1 to build the SSVaR. The dash curves are the bounds of AN-SSVaR. The solid curves are the bounds of SP-SSVaR. The dash-dot curve is the ecdf of the data set SP2. We observe that the lower bound of the AN-SSVaR is always smaller than the lower bound of the SP-SSVaR which means that with this former CI we accept higher risks. Indeed for risks between $0 < p < 0.003$, if we use as a risk measure the AN-SSVaR, we consider that the value of the risks provided with the ecdf line is acceptable; on the other hand if we consider as a risk measure the SP-SSVaR for $0 < p < 0.003$, it can be seen from the ecdf line that an alert is done.

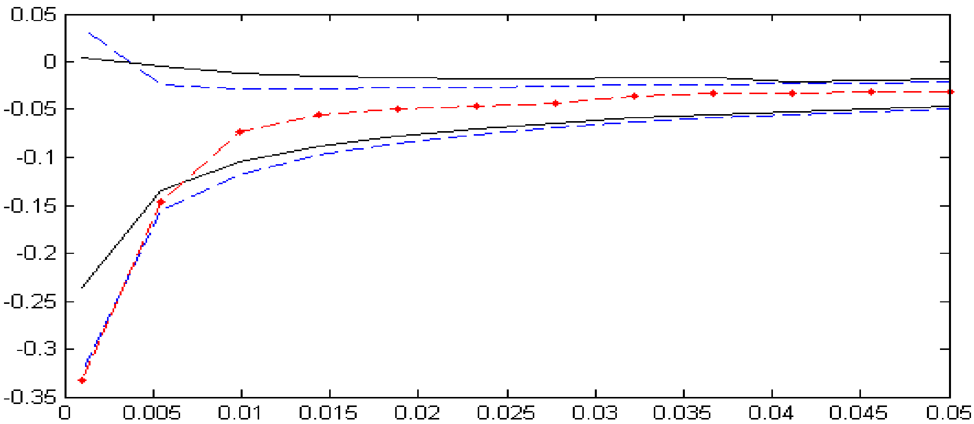

Fig. 4 In this figure, we use the data set HSI1 to build the SSVaR. The dash curves are the bounds of AN-SSVaR. The solid curves are the bounds of SP-SSVaR. The dash-dot curve is the ecdf of the data set HSI2. We observe that the lower bound of the AN-SSVaR is always smaller than the lower bound of the SP-SSVaR which means that with this former CI we accept higher risks. Indeed for risks between $0 < p < 0.007$, if we use as a risk measure the AN-SSVaR, we consider that the value of the risks provided with the ecdf line is acceptable; on the other hand if we consider as a risk measure the SP-SSVaR for $0 < p < 0.007$, it can be seen from the ecdf line that an alert is done.

A Proof of Corollary 1

At first, we introduce the Slutsky's theorem

Theorem 3 (Slutsky's theorem) *Let $\{X_n\}$, $\{Y_n\}$ be sequences of r.v., If $\{X_n\}$ converges in distribution ($\rightarrow_{(d)}$) to a r.v. X and $\{Y_n\}$ converges in probability to a constant c ($\rightarrow_{(p)}$),* *then*

$$X_nY_n \to_{(d)} cX. \tag{17}$$

To prove Corollary 1, we begin with

$$\begin{aligned} &\frac{X_{(m)} - F_\theta^{-1}(p)}{\sqrt{\widehat{V}}} \\ &= \sqrt{\frac{V}{\widehat{V}}}\frac{X_{(m)} - F_\theta^{-1}(p)}{\sqrt{V}} \\ &= \frac{f_{\hat{\theta}}(F_{\hat{\theta}}^{-1}(p))}{f_\theta(F_\theta^{-1}(p)}\frac{X_{(m)} - F_\theta^{-1}(p))}{\sqrt{V}}. \end{aligned} \tag{18}$$

Since convergence in probability is preserved under continuous transformations, from $\hat{\theta} \to_{(P)} \theta$ we have

$$f_{\hat{\theta}}(F_{\hat{\theta}}^{-1})(p) - f_\theta(F_\theta^{-1}(p)) \to_{(P)} 0. \tag{19}$$

From Theorem 1 we know that $\frac{X_{(m)} - F_\theta^{-1}(p))}{\sqrt{V}} \to_{(d)} N(0,1)$, then from Slutsky's theorem we have finally

$$\frac{X_{(m)} - F_\theta^{-1}(p)}{\sqrt{\widehat{V}}} \to_{(d)} N(0,1). \tag{20}$$

B Estimates of the NIG distribution for SP1 and HSI1

We fit the data sets SP1 and HSI1 with NIG distribution. The estimates are provided in Table (4).

	Fitted parameters (NIG)			
	tail	skewness	location	scale
SP1	23.7081	-7.3294	0.0015	0.0029
HSI1	17.7947	-9.0184	0.0029	0.0034

Table 4 We provide the fitted parameters (NIG) for data sets SP1 and HSI1