Noname manuscript No.
(will be inserted by the editor)

Stochastic models for risk estimation in volatile markets: A survey

Stoyan V. Stoyanov · Borjana Racheva - Iotova · Svetlozar T. Rachev · Frank J. Fabozzi

Received: date / Accepted: date

Abstract The problem of portfolio risk estimation in volatile markets requires employing fat-tailed models for financial instrument returns combined with copula functions to capture asymmetries in dependence and a true downside risk measure for risk estimation. In this survey, we discuss how these three essential components can be combined together in a Monte Carlo based framework for risk estimation and risk budgeting with the average value-at-risk measure (AVaR). We consider in detail the questions of AVaR calculation and estimation and also stochastic stability of AVaR when combined with heavy-tailed scenarios.

Keywords fat-tailed distributions · stable distributions · downside risk · average value-at-risk · conditional value-at-risk · risk budgeting

1 Introduction

The two main conventional approaches to modeling asset returns are based either on a historical or a normal (Gaussian) distribution for returns. Neither approach adequately captures unusual behavior of asset prices and returns. The historical model is bounded

Stoyan V. Stoyanov
FinAnalytica Inc., USA
University of Karlsruhe, Germany, KIT
E-mail: stoyan.stoyanov@finanalytica.com

Borjana Racheva-Iotova
FinAnalytica Inc., USA
E-mail: borjana.racheva@finanalytica.com

Svetlozar T. Rachev
University of Karlsruhe, Germany, KIT
University of California Santa Barbara, USA
and FinAnalytica Inc., USA
E-mail: zari.rachev@statistik.uni-karlsruhe.de

Frank J. Fabozzi
135 Prospect Street, New Haven, CT 06511 USA
Telephone: (203) 432-2421
E-mail: frank.fabozzi@yale.edu

by the extent of the available observations and the normal distribution model inherently cannot produce extreme returns.

There are many studies exploring the non-normality of assets returns and suggesting alternative approaches. Among the well known candidates are Student's t distribution, generalized hyperbolic distributions (see Hurst et al. (1997), Bibby and Sorensen (2003), and Platen and Rendek (2007)) and stable Paretian distributions (see Rachev and Mittnik (2000)). At least some of their forms are subordinated normal models and thus provide a very practical and tractable framework. Rachev et al. (2005*b*) provide an introduction to heavy-tailed models in finance.

In order to have a reliable risk model, not only do we need realistic assumptions for the returns distributions but also a true downside risk measure. For this matter, we employ the average value-at-risk (AVaR) risk measure, see Rachev et al. (2007). In this survey, we discuss calculation and estimation of AVaR and also a risk budgeting framework allowing one to compute marginal and percentage contribution to AVaR with fat-tailed scenarios. Finally, we discuss the stochastic stability of the sample AVaR estimator and how we can improve it. The last part of the discussion can be used for construction of confidence intervals for Monte Carlo based AVaR.

2 Heavy-tailed and asymmetric models for assets returns

Specifying properly the distribution of assets returns is vital for risk management and optimal asset allocation. A failure may lead to significant underestimation of portfolio risk and, consequently, to wrong decisions.

The distributional modeling of financial variables has several dimensions. First, there should be a realistic model for the returns of each financial variable considered separately. That is, we should employ realistic one-dimensional models. Second, the model should capture properly the dependence between the one-dimensional variables. Therefore, we need a true multivariate model with the above two building blocks correctly specified.

2.1 One-dimensional models

The cornerstone theories in finance such as mean-variance model for portfolio selection and asset pricing models that have been developed rest upon the assumption that asset returns follow a normal distribution. Yet, there is little, if any, credible empirical evidence that supports this assumption for financial assets traded in most markets throughout the world. Moreover, the evidence is clear that financial return series are heavy-tailed and, possibly, skewed. Fortunately, several papers have analyzed the consequences of relaxing the normality assumption and developed generalizations of prevalent concepts in financial theory that can accommodate heavy-tailed returns (see Rachev and Mittnik (2000) and Rachev (2003) and references therein).

Mandelbrot (1963) strongly rejected normality as a distributional model for asset returns, conjecturing that financial return processes behave like non-Gaussian stable processes. To distinguish between Gaussian and non-Gaussian stable distributions, the latter are commonly referred to as ”stable Paretian” distributions or “Levy stable” distributions.

While there have been several studies in the 1960s that have extended Mandelbrot's investigation of financial return processes, probably, the most notable is Fama (1963, 1965). Fama's work and others led to a consolidation of the stable Paretian hypothesis. In the 1970s, however, closer empirical scrutiny of the "stability" of fitted stable Paretian distributions also produced evidence that was not consistent with the stable Paretian hypothesis. Specifically, it was often reported that fitted characteristic exponents (or tail-indices) did not remain constant under temporal aggregation. Partly in response to these empirical "inconsistencies," various alternatives to the stable law were proposed in the literature, including fat-tailed distributions being only in the domain of attraction of a stable Paretian law, finite mixtures of normal distributions, the Student's t distribution, and the hyperbolic distribution, see Bibby and Sorensen (2003).

Recent attacks on Mandelbrot's stable Paretian hypothesis focus on the claim that empirical asset return distributions are not as heavy-tailed as the non-Gaussian stable law suggests. Studies that come to such conclusions are typically based on tail-index estimates obtained with the Hill estimator. Because sample sizes beyond 100,000 are required to obtain reasonably accurate estimates, the Hill estimator is highly unreliable for testing the stable hypothesis. More importantly, Mandelbrot's stable Paretian hypothesis is interpreted too narrowly, if one focuses solely on the marginal distribution of return processes. The hypothesis involves more than simply fitting marginal asset return distributions. Stable Paretian laws describe the fundamental "building blocks" (e.g., innovations) that drive asset return processes. In addition to describing these "building blocks," a complete model should be rich enough to encompass relevant stylized facts, such as

- non-Gaussian, heavy-tailed and skewed distributions
- volatility clustering (ARCH-effects)
- temporal dependence of the tail behavior
- short- and long-range dependence

An attractive feature of stable models — not shared by other distributional models — is that they allow us to generalize Gaussian-based financial theories and, thus, to build a coherent and more general framework for financial modeling (see Mittnik and Rachev (1999)). The generalizations are only possible because of specific probabilistic properties that are unique to (Gaussian and non-Gaussian) stable laws, namely, the stability property, the Central Limit Theorem, and the Invariance Principle for stable processes. Detailed accounts of properties of stable distributed random variables can be found in Samorodnitsky and Taqqu (1994) and Janicki and Weron (1994).

The class of the stable distributions is defined by means of their characteristic functions. With very few exceptions, no closed-form expressions are known for their densities and cumulative distribution functions (c.d.f.). A random variable X is said to have a stable distribution if its characteristic function $\varphi_X(t) = Ee^{itX}$ has the following form

$$\varphi_X(t) = \begin{cases} \exp\{-\sigma^\alpha |t|^\alpha (1 - i\beta \frac{t}{|t|} \tan(\frac{\pi\alpha}{2})) + i\mu t\}, & \alpha \neq 1 \\ \exp\{-\sigma |t| (1 + i\beta \frac{2}{\pi} \frac{t}{|t|} \ln(|t|)) + i\mu t\}, & \alpha = 1 \end{cases} \tag{1}$$

where $\frac{t}{|t|} = 0$ if $t = 0$. The formula in (1) implies that they are described by four parameters: α, called the index of stability, which determines the tail weight or density's kurtosis with $0 < \alpha \leq 2$, β, called the skewness parameter, which determines the density's skewness with $-1 \leq \beta \leq 1$, $\sigma > 0$ which is a scale parameter, and $\mu \in \mathbb{R}$ which is a location parameter. Stable distributions allow for skewed distributions when $\beta \neq 0$ and when $\beta = 0$, the distribution is symmetric around μ. Stable Paretian laws have fat tails, meaning that extreme events have high probability relative to the normal distribution when $\alpha < 2$. The Gaussian distribution is a stable distribution with $\alpha = 2$. (For more details on the properties of stable distributions, see Samorodnitsky, Taqqu (1994).) Of the four parameters, α and β are most important as they identify two fundamental properties that are atypical of the normal distribution — heavy tails and asymmetry.

Rachev, Stoyanov, Biglova and Fabozzi (2006) consider the daily return distribution of 382 U.S. stocks in the framework of two probability models — the homoskedastic independent, identical distributed model and the conditional heteroskedastic ARMA-GARCH model. In both models, the Gaussian hypothesis is strongly rejected in favor of the stable Paretian hypothesis which better explains the tails and the central part of the return distribution. The companies in the study are the constituents of the S&P 500 with complete history in the 12-year time period from January 1, 1992 to December 12, 2003. The estimated parameters suggest significant heavy-tail and asymmetry in the residual which cannot be accounted for by the normal distribution.

Even though there is much empirical evidence in favor of the stable hypothesis, it is a theoretical fact that stable distributions with $\alpha < 2$ have an infinite second moment. Thus, if we model the return distribution of a stock with such a model, we assume it has an infinite volatility. This property creates problems in derivatives pricing models and, in order to avoid it, modifications to stable distributions have been proposed such as smoothly truncated stable laws, see Rachev et al. (2005*a*). More general models in this direction applied to option pricing include tempered stable distributions, see Kim et al. (2008).

2.2 Multivariate models

For the purposes of portfolio risk estimation, constructing one-dimensional models for financial instruments is incomplete. Failure to account for the dependencies between financial instruments may be fatal for the analysis.

There are two ways to build a complete multivariate model. It is possible to hypothesize a multivariate distribution directly (i.e., the dependence between stock returns as well as their one-dimensional behavior). Assumptions of this type include the multivariate normal, the multivariate Student's *t*, the more general elliptical family, the multivariate stable, etc. Sometimes, in analyzing dependence, an explicit assumption is not made, for instance, the covariance matrix is very often relied on. While an explicit multivariate assumption is not present, it should be kept in mind that this is consistent with the mutivariate normal hypothesis. More importantly, the covariance matrix can describe only linear dependencies and this is a basic limitation.

In the last decade, a second approach has become popular. One can specify separately the one-dimensional hypotheses and the dependence structure through a function called copula. This is a more general and more appealing method because one is free to choose separately different parametric models for the stand-alone variables and a parametric copula function. For more information, see Embrechts et al. (2002) and Embrechts et al. (2003).

From a mathematical viewpoint, a copula function C is nothing more than a probability distribution function on the d-dimensional hypercube

$$C(u_1, \ldots, u_d), \quad u_i \in [0, 1], \quad i = 1, d$$

where $C(u_i) = u_i, \ i = 1, d$. It is known that for any multivariate cumulative distribution function:

$$F(x_1, \ldots, x_d) = P(X_1 \leq x_1, \ldots, X_d \leq x_d)$$

there exists a copula C such that

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d))$$

where the $F_i(x_i)$ are the marginal distributions of $F(x_1, \ldots, x_d)$, and conversely for any copula C the right-hand-side of the above equation defines a multivariate distribution function $F(x_1, \ldots, x_d)$. See, for example, Bradley and Taqqu (2003), Sklar (1996), and Embrechts et al. (2003).

The main idea behind the use of copulas is that one can first specify the marginal distributions in whatever way makes sense (e.g. fitting marginal distribution models to risk factor data, and then specify a copula C to capture the multivariate dependency structure in the best suited manner).

A possible approach for choosing a flexible copula model is to adopt the copula of a parametric multivariate distribution. In this way, the copula itself will have a parametric form. There are many multivariate laws mentioned in the literature, which can be used for this purpose. One such example is the Gaussian copula, i.e. the copula of the multivariate normal distribution. It is easy to work with but it has one major drawback: It implies that extreme events are asymptotically independent. Thus, the probability of joint occurrence of large in absolute value negative returns of two stocks is significantly underestimated. An alternative to the Gaussian copula is the Student's t copula (i.e., the copula of the multivariate Student's t distribution). It models better the probability of joint extreme events but it has the disadvantage that it is symmetric. Thus, the probability of joint occurrence of very large returns is the same as the probability of joint occurrence of very small returns. This deficiency is not present in the skewed Student's t copula which we believe is a much more realistic model of dependency. This is the copula of the multivariate skewed Student's t distribution defined by means of the following stochastic representation,

$$X = \mu + \gamma W + Z\sqrt{W}$$

where $W \in IG(\nu/2, \nu/2)$, i.e., W is inverse gamma distributed, Z is multivariate normal random variable, $Z \in N_d(0, \Sigma)$, W and Z are independent, and the constants μ and γ are such that the sign of a given component of γ controls the asymmetry of the corresponding component of X and μ is a location parameter contributing to the mean of X. The skewed Student's t copula has the following features which make it a flexible and attractive model:

- it has a parametric form which makes the copula an attractive model in higher dimensions
- the underlying stochastic representation facilitates scenario generation from the copula

– it can describe tail dependence, if present in the data
– it can describe asymmetric dependence, if present in the data

For additional information about the skewed Student's t copula and a case study for the German equity market, see Sun et al. (2008).

3 AVaR calculation and estimation

Value-at-risk (VaR) at a confidence level $1-\epsilon$ is defined as the negative of the ϵ-quantile of the return distribution, $VaR_\epsilon(X) = -F^{-1}(\epsilon)$, where F^{-1} is the inverse distribution function. It the has been widely adopted as a risk measure. However, it is not very informative which we illustrate in the following example. Suppose that X and Y are two random variables describing the return distribution of two financial instruments. If at a given confidence level $VaR_\epsilon(X) = VaR_\epsilon(Y) = q_\epsilon$, can we state that the two financial instruments are equally risky? The answer is negative because while we know that losses larger than q_ϵ for both financial instruments will occur with the same probability ϵ, we are not sure about the magnitude of these losses. A risk measure which captures this information is average value-at-risk (AVaR)[1] which computes the average VaR provided that the it is larger than the VaR at the corresponding confidence level,

$$AVaR_\epsilon(X) := \frac{1}{\epsilon}\int_0^\epsilon VaR_p(X)dp. \tag{2}$$

In the example above, the two financial instruments may have one and the same VaR and different average losses provided that the loss is larger than that VaR. Thus, even if $VaR_\epsilon(X) = VaR_\epsilon(Y)$, the AVaRs may differ, $AVaR_\epsilon(X) \neq AVaR_\epsilon(Y)$. Such an example is illustrated in Figure 1 in which the two distribution functions intersect at $\epsilon = 0.05$ which means that the VaRs at 95% confidence level are equal. However, the tail of X is heavier resulting in larger losses on average provided that the losses exceed the VaR at 95% confidence level. Therefore, $AVaR_{0.05}(X) > AVaR_{0.05}(Y)$. Figure 1 also shows a geometric interpretation of $AVaR_\epsilon$ — it is a number such that if we draw a rectangle with height equal to ϵ and width equal to $AVaR_\epsilon$ as shown on the figure, then the area of the rectangle equals the corresponding shaded area. In the discussion in this section, we assume that the return distribution has no point masses. Under this assumption, AVaR equals the following conditional expectation,[2]

$$AVaR_\epsilon(X) = -E(X|X < -VaR_\epsilon(X)).$$

Not only does AVaR have a sound practical meaning, but it also has good mathematical properties. AVaR satisfies the coherency axioms introduced in Artzner et al. (1998). As a result, AVaR is a sub-additive, convex function of portfolio weights which can be readily optimized.

[1] In the literature, AVaR is also called *conditional value-at-risk* (see, Rockafellar and Uryasev (2002)) or *expected shortfall* but we will use AVaR as it best describes the quantity it refers to.

[2] The median version of AVaR, $MVaR_\epsilon(X) = -\text{med}(X|X < VaR_\epsilon(X))$ is infact equal to $VaR_{\epsilon/2}(X)$, see Rachev et al. (2008). Therefore, if risk managers prefer using VaR as a measure for risk, it is advisable that they use $VaR_{\epsilon/2}(X)$ for a specified tail probability ϵ because this is going to provide them with a robust measure of the average losses above $VaR_\epsilon(X)$.

Apart from the definition in (2), AVaR can be represented through a minimization formula,

$$AVaR_\epsilon(X) = \min_{\theta\in\mathbb{R}} \left(\theta + \frac{1}{\epsilon} E(-X-\theta)_+\right) \tag{3}$$

where $(x)_+ = \max(x, 0)$ and X describes the return distribution. It turns out that this formula has an important application in optimal portfolio problems based on AVaR as a risk measure. Equation (3) was first studied by Pflug (2000). A proof that equation (2) is indeed the AVaR can be found in Rockafellar and Uryasev (2002).

The formula in (3) also allows for a geometric interpretation. If we re-write it as

$$AVaR_\epsilon(X) = \frac{1}{\epsilon} \min_{\theta\in\mathbb{R}} \left(\epsilon\theta + E(-X-\theta)_+\right), \tag{4}$$

we can find a similarity with (2). In Figure 1, we interpreted the integral in (2) as the shaded area. In a similar way, we can find an area corresponding to the objective function in the minimization problem in (4). A little algebra shows that the expectation $E(-X-\theta)_+$ equals the area closed between the graph of the c.d.f. and a line parallel to the vertical axis passing through the point $(-\theta, 0)$. The product $\epsilon \times \theta$ equals the area of a rectangle with sides equal to ϵ and θ. This area is added to $E(-X-\theta)_+$. Figure 2 shows the two areas together. The shaded areas on the two plots equal $\epsilon \times AVaR_\epsilon(X)$. The top plot shows the case in which $-\theta < -VaR_\epsilon(X)$. The total area corresponding to the objective in the minimization formula, $\epsilon\theta + E(-X-\theta)_+$, equals the shaded area plus the marked area. If $-\theta > -VaR_\epsilon(X)$, then we obtain a similar case shown on the bottom plot. By varying θ, the total area changes but it always remains larger than the shaded area unless $\theta = VaR_\epsilon(X)$. Thus, when $\theta = VaR_\epsilon(X)$ the minimum area is attained which equals exactly $\epsilon \times AVaR_\epsilon(X)$. Therefore, the minimization formula in (3) calculates AVaR. For additional geometric interpretations, see Rachev et al. (2008).

3.1 Closed-form expressions

For some continuous distributions, it is possible to calculate explicitly the AVaR through the definition. We provide closed-form expressions for the normal distribution and Student's t distribution.

1. *The Normal distribution*
   Suppose that X is distributed according to a normal distribution with standard deviation σ_X and mathematical expectation EX. The AVaR of X at tail probability ϵ equals

$$AVaR_\epsilon(X) = \frac{\sigma_X}{\epsilon\sqrt{2\pi}} \exp\left(-\frac{(VaR_\epsilon(Y))^2}{2}\right) - EX \tag{5}$$

   where Y has the standard normal distribution, $Y \in N(0, 1)$.

2. *The Student's t distribution*
   Suppose that X has Student's t distribution with ν degrees of freedom, $X \in t(\nu)$. The AVaR of X at tail probability ϵ equals

$$AVaR_\epsilon(X) = \begin{cases} \dfrac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \dfrac{\sqrt{\nu}}{(\nu-1)\epsilon\sqrt{\pi}} \left(1 + \dfrac{(VaR_\epsilon(X))^2}{\nu}\right)^{\frac{1-\nu}{2}} & , \ \nu > 1 \\ \infty & , \ \nu = 1 \end{cases}$$

Note that equation (5) can be represented in a more compact way,

$$AVaR_\epsilon(X) = \sigma_X C_\epsilon - EX, \tag{6}$$

where C_ϵ is a constant which depends only on the tail probability ϵ. Therefore, the AVaR of the normal distribution has the same structure as the normal VaR — the difference between the properly scaled standard deviation and the mathematical expectation. In effect, similar to the normal VaR, the normal AVaR properties are dictated by the standard deviation. Even though AVaR is focused on the extreme losses only, due to the limitations of the normal assumption, it is symmetric. Exactly the same conclusion holds for the AVaR of Student's t distribution. The true merits of AVaR become apparent if the underlying distributional model is skewed.

It turns out that it is possible to arrive at formulae for the AVaR of stable distributions and skewed Student's t distributions. The expressions are more complicated even though they are suitable for numerical work. They involve numerical integration but this is not a severe restriction because tools are available in many software packages and the integrands are nicely behaved functions. The calculations for the case of stable distributions can be found in Stoyanov et al. (2006). In this section, we only provide the result.

Suppose that the random variable X has a stable distribution with tail exponent α, skewness parameter β, scale parameter σ, and location parameter μ, $X \in S_\alpha(\sigma, \beta, \mu)$. If $\alpha \leq 1$, then $AVaR_\epsilon(X) = \infty$. The reason is that stable distributions with $\alpha \leq 1$ have infinite mathematical expectation and the AVaR is unbounded.

If $\alpha > 1$ and $VaR_\epsilon(X) \neq 0$, then the AVaR can be represented as

$$AVaR_\epsilon(X) = \sigma A_{\epsilon,\alpha,\beta} - \mu$$

where the term $A_{\epsilon,\alpha,\beta}$ does not depend on the scale and the location parameters. In fact, this representation is a consequence of the positive homogeneity and the invariance property of AVaR. Concerning the term $A_{\epsilon,\alpha,\beta}$,

$$A_{\epsilon,\alpha,\beta} = \frac{\alpha}{1-\alpha} \frac{|VaR_\epsilon(X)|}{\pi\epsilon} \int_{-\overline{\theta}_0}^{\pi/2} g(\theta) \exp\left(-|VaR_\epsilon(X)|^{\frac{\alpha}{\alpha-1}} v(\theta)\right) d\theta$$

where

$$g(\theta) = \frac{\sin(\alpha(\overline{\theta}_0 + \theta) - 2\theta)}{\sin\alpha(\overline{\theta}_0 + \theta)} - \frac{\alpha\cos^2\theta}{\sin^2\alpha(\overline{\theta}_0 + \theta)},$$

$$v(\theta) = \left(\cos\alpha\overline{\theta}_0\right)^{\frac{1}{\alpha-1}} \left(\frac{\cos\theta}{\sin\alpha(\overline{\theta}_0 + \theta)}\right)^{\frac{\alpha}{\alpha-1}} \frac{\cos(\alpha\overline{\theta}_0 + (\alpha-1)\theta)}{\cos\theta},$$

in which $\overline{\theta}_0 = \frac{1}{\alpha}\arctan\left(\overline{\beta}\tan\frac{\pi\alpha}{2}\right)$, $\overline{\beta} = -\text{sign}(VaR_\epsilon(X))\beta$, and $VaR_\epsilon(X)$ is the VaR of the stable distribution at tail probability ϵ.

If $VaR_\epsilon(X) = 0$, then the AVaR admits a very simple expression,

$$AVaR_\epsilon(X) = \frac{2\Gamma\left(\frac{\alpha-1}{\alpha}\right)}{(\pi - 2\theta_0)} \frac{\cos\theta_0}{(\cos\alpha\theta_0)^{1/\alpha}}.$$

in which $\Gamma(x)$ is the gamma function and $\theta_0 = \frac{1}{\alpha}\arctan(\beta\tan\frac{\pi\alpha}{2})$.

A similar result for skewed Student's t distribution is given in Dokov et al. (2008).

3.2 Estimating AVaR from a sample

Suppose that we have a sample of observed portfolio returns and we are not aware of their distribution. Provided that we do not impose any distributional model, the AVaR of portfolio return can be estimated from the sample of observed portfolio returns. Denote the observed portfolio returns by $r_1, r_2, \ldots, r_n$ at time instants $t_1, t_2, \ldots, t_n$. The numbers in the sample are given in order of observation. Denote the sorted sample by $r_{(1)} \leq r_{(2)} \leq, \ldots, \leq r_{(n)}$. Thus, $r_{(1)}$ equals the smallest observed portfolio return and $r_{(n)}$ is the largest. The AVaR of portfolio returns at tail probability ϵ is estimated according to the formula[3]

$$\widehat{AVaR}_\epsilon(r) = -\frac{1}{\epsilon}\left(\frac{1}{n}\sum_{k=1}^{\lceil n\epsilon\rceil-1} r_{(k)} + \left(\epsilon - \frac{\lceil n\epsilon\rceil - 1}{n}\right) r_{(\lceil n\epsilon\rceil)}\right) \tag{7}$$

where the notation $\lceil x\rceil$ stands for the smallest integer larger than x. The "hat" above AVaR denotes that the number calculated by equation (7) is an estimate of the true value because it is based on a sample.

Besides formula (7), there is another method for calculation of AVaR. It is based on the minimization formula (3) in which we replace the mathematical expectation by the sample average,

$$\widehat{AVaR}_\epsilon(r) = \min_{\theta\in\mathbb{R}}\left(\theta + \frac{1}{n\epsilon}\sum_{i=1}^{n}\max(-r_i - \theta, 0)\right). \tag{8}$$

Even though it is not obvious, equations (7) and (8) are completely equivalent.

The minimization formula in equation (8) is appealing because it can be calculated through the methods of linear programming. It can be restated as a linear optimization problem by introducing auxiliary variables $d_1, \ldots, d_n$, one for each observation in the sample,

$$\begin{array}{ll} \min\limits_{\theta,d} & \theta + \dfrac{1}{n\epsilon}\sum\limits_{k=1}^{n} d_k \\ \text{subject to} & -r_k - \theta \leq d_k,\ k = 1, n \\ & d_k \geq 0,\ k = 1, n \\ & \theta \in \mathbb{R}. \end{array} \tag{9}$$

The linear problem (9) is obtained from (8) through standard methods in mathematical programming. We briefly demonstrate the equivalence between them. Let us fix the value of θ to θ^*. Then the following choice of the auxiliary variables yields

[3] This formula is a simple consequence of the definition of AVaR for discrete distributions. A detailed derivation is provided by Rockafellar and Uryasev (2002).

the minimum in (9). If $-r_k - \theta^* < 0$, then $d_k = 0$. Conversely, if it turns out that $-r_k - \theta^* \geq 0$, then $-r_k - \theta^* = d_k$. In this way, the sum in the objective function becomes equal to the sum of maxima in equation (8).

We summarize the attractive properties of AVaR below:

- AVaR gives an informed view of losses beyond VaR.
- AVaR is a convex function of portfolio weights, and is therefore attractive to optimize portfolios (see Rockafellar and Uryasev (2002)).
- AVaR is sub-additive and satisfies a set of intuitively appealing coherent risk measure properties (see Artzner et al. (1998)).
- AVaR is a form of expected loss (i.e., a conditional expected loss) and is a very convenient form for use in scenario-based portfolio optimization. It is also quite a natural risk-adjustment to expected return (see Rachev, Martin, Racheva-Iotova and Stoyanov (2006)).

Even though AVaR is not widely adopted, we expect it to become an accepted risk measure as portfolio and risk managers become more familiar with its attractive properties. For portfolio optimization, we recommend the use of heavy-tailed distributions and AVaR, and limiting the use of historical, normal or stable VaR to required regulatory reporting purposes only. Finally, organizations should consider the advantages of AVaR with heavy-tailed distributions for risk assessment purposes and non-regulatory reporting purposes.

4 Risk budgeting with AVaR

The concept of AVaR allows for scenario-based risk decomposition which is a concept similar to the standard deviation based percentage contribution to risk (PCTR). The practical issue is to identify the contribution of each position to portfolio risk and since ETL is a tail risk measure, percentage contribution to ETL allows one to build a framework for tail risk budgeting. The approach largely depends on one of the coherence axioms given Artzner et al. (1998), which is the positive homogeneity property

$$AVaR_\epsilon(aX) = aAVaR_\epsilon(X), \quad a > 0.$$

Euler's formula is valid for such functions. According to it, the risk measure can be expressed in terms of a weighted average of the partial derivatives with respect to portfolio weights,

$$AVaR_\epsilon(w'X) = \sum_i w_i \frac{\partial AVaR_\epsilon(w'X)}{\partial w_i}$$

where w is a vector of weights, X is a random vector describing the multivariate return of all financial instruments in the portfolio, and $w'X$ is the portfolio return. The left hand-side of the equation equals total portfolio risk and if we divide both sides by it, we obtain the needed tail risk decomposition,

$$\begin{aligned} 1 &= \sum_i \frac{w_i}{AVaR_\epsilon(w'X)} \frac{\partial AVaR_\epsilon(w'X)}{\partial w_i} \\ &= \sum_i p_i. \end{aligned} \tag{10}$$

In order to compute the percentage contribution to risk of the i-th position, the i-th summand p_i in (10), we have to calculate first the partial derivative. It turns out that the derivative can be expressed as a conditional expectation,

$$\frac{\partial AVaR_\epsilon(w'X)}{\partial w_i} = -E(X_i|w'X < -VaR_\epsilon(w'X)).$$

when X is an absolutely continuous random variable, see Zhang and Rachev (2006) and the references therein. The conditional expectation can be computed through the Monte Carlo method.

4.1 Identifying risk diversifiers and risk contributors

Exploiting a link between percentage contribution to AVaR and the global minimum AVaR portfolio, we can arrive at a rule for identifying AVaR diversifiers and AVaR contributors in long-only portfolios.

Consider the global minimum portfolio AVaR problem,

$$\begin{aligned}&\min_w AVaR_\epsilon(w'X)\\ &\text{s.t.}\\ &w'e = 1.\end{aligned}$$

where e is a vector of ones and the condition $w'e = 1$ means that the sum of all weights should be equal to 1. Due to the convexity property of AVaR, this problem has a unique minimum which can be obtained through the standard first-order optimality conditions for constrained optimization problems. The solution of the optimization problem is the global minimum AVaR portfolio which should satisfy the conditions,

$$\nabla AVaR_\epsilon(w'X) = \lambda e, \tag{11}$$

where $\nabla AVaR_\epsilon(w'X)$ is the gradient of AVaR computed at the optimal solution and λ is the Lagrange multiplier. The Lagrange multiplier can be explicitly computed,

$$\lambda = w'\nabla AVaR_\epsilon(w'X) = AVaR_\epsilon(w'X),$$

in which we make use of Euler's formula. Equation (11) implies that the partial derivatives of the global minimum AVaR portfolio are all equal and the percentage contribution to AVaR of the i-th position equals the weight of this position. On the basis of this observation, a simple rule can be designed identifying risk diversifiers and risk contributors for long-only portfolios:

- If $p_i > w_i$, then the i-th position is a risk contributor.
- If $p_i < w_i$, then the i-th position is a risk diversifier.
- If $p_i = w_i$, then the i-th position is neither a contributor nor a diversifier.

The rationale is that if we increase marginally the weights of the diversifiers and decrease marginally the weights of the contributors, making sure that the weights sum up to one, the AVaR of the new portfolio is marginally improved. In case $p_i = w_i$ for all i, then we are holding the global minimum AVaR portfolio and the risk cannot be marginally reduced.

The analysis presented above is valid for long-only portfolios because negative weights change the sign of the percentage contribution statistics p_i and identifying contributors and diversifiers by comparing to w_i may be problematic. For arbitrary portfolios however, we can use another rule for marginal rebalancing which is based solely on the partial derivatives of AVaR,

– Compute the partial derivatives of AVaR for a given portfolio w.
– Sort the portfolio positions by the derivatives in decreasing order.
– The position on top is a risk contributor and the position at the bottom is a risk diversifier.
– In order to improve marginally portfolio risk, decrease the weight of the position on top by a small amount and increase the weight of the position at the bottom making sure all weights sum up to one.

If the partial derivative of the position on top equals the partial derivative of the position at the bottom, then we are holding the global minimum AVaR portfolio and no marginal improvement of risk is possible.

5 AVaR variability

In practice, computing portfolio AVaR is done through the Monte Carlo method. We hypothesize a parametric model for the multivariate distribution of financial instruments returns, we fit the model, and then we generate a large number of scenarios. From the generated scenarios, we compute scenarios for portfolio return. Employing formula (7), we calculate portfolio AVaR at a specified tail probability ϵ.

We can regard the generated scenarios as a sample from the fitted model and thus the computed AVaR in the end appears as an estimate of the true AVaR. The larger the sample, the closer the estimated AVaR is to the true value. If we regenerate the scenarios, the portfolio AVaR number will change and it will fluctuate around the true value. Figure 3 illustrates this phenomenon for the standard normal distribution with $\epsilon = 0.01$, for which the true value $AVaR_{0.01}(X) = 2.665$. This stochastic variability is an issue inherent in the Monte Carlo method and cannot be avoided. In this context, the Monte Carlo method can be viewed as a numerical method of computing portfolio AVaR when the hypothesized multivariate model does not allow portfolio AVaR to be computed analytically. In this section, we discuss the asymptotic distribution of the estimator in (7) which we can use to determine approximately the variance of (7) when the number of scenarios is large.

Before proceeding to a more formal result, let us check what intuition may suggest. If we look at equation (7), we notice that the leading term is the average of the smallest observations in the sample. The fact that we average observations reminds of the central limit theorem (CLT) and the fact that we average the smallest observations in the sample suggests that the variability should be influenced by the behavior of the left tail of the portfolio return distribution. Basically, a result based on CLT would state that the distribution of the AVaR estimator becomes more and more normal as we increase the sample size. Applicability of CLT however depends on certain conditions such as finite variance which guarantee certain regularity of the random numbers. If this regularity is not present, the smallest numbers in a sample may vary quite a lot as they are not naturally bounded in any respect. Therefore, for heavy-tailed distributions we can expect that CLT may not hold and the distribution of the estimator in such cases may not be normal at all.

The formal result in Stoyanov and Rachev (2008*b*) confirms these expectations. Taking advantage of the generalized CLT, we can demonstrate that

Theorem 1 *Suppose that X is random variable with distribution function $F(x)$ which satisfies the following conditions*

– $x^\alpha F(x) = L(x)$ *is slowly varying at infinity, i.e.* $\lim_{x\to\infty} L(tx)/L(x) = 1$, $\forall t > 0$.
– $\int_{-\infty}^{0} x dF(x) < \infty$
– $F(x)$ *is differentiable at* $x = q_\epsilon$ *where* q_ϵ *is the* ϵ*-quantile of* X.

Then, there exist c_n $n = 1, 2, ...,$ *such that for any* $0 < \epsilon < 1$,

$$c_n^{-1}\left(\widehat{AVaR}_\epsilon(X) - AVaR_\epsilon(X)\right) \xrightarrow{w} S_{\alpha^*}(1, 1, 0) \tag{12}$$

in which $\xrightarrow{w}$ *denotes weak limit,* $1 < \alpha^* = \min(\alpha, 2)$, *and* $c_n = n^{1/\alpha^*} L_0(n)/\epsilon$ *where* L_0 *is a function slowly varying at infinity and* $\widehat{AVaR}_\epsilon(X)$ *is computed from a sample of independent copies of* X *according to equation* (7).

This theorem implies that the limit distribution of the AVaR estimator in (7) is necessarily a stable distribution totally skewed to the left. In the context of the theorem, we can think of X as a random variable describing portfolio return. If the index α governing the left tail of X is $\alpha \geq 2$, then the above result reduces to the classical CLT as in this case $\alpha^* = 2$ and the limit distribution is normal. This case is considered in detail in Stoyanov and Rachev (2008*a*).

5.1 Stable distributed returns

In this section, we consider the case in which portfolio return distribution is a stable law with parameter $1 < \alpha < 2$. Under this assumption, $\alpha^* = \alpha$ and thus the limit distribution of the AVaR estimator is also a stable law with parameter α. That is, we are not in the case of the classical CLT. We carry out a Monte Carlo study in order to see for how many scenarios the limit distribution in the theorem is sufficiently close to the real distribution of the estimator. We choose $X \in S_{1.5}(0.7, 1, 0)$ and we generate 2,000 samples from the corresponding distribution the size of which equals $n = 250$; $1,000$; $10,000$; and $100,000$. Figure 4 shows the density of the random variable in the left part of equation (12) and how it approaches the limit distribution, i.e. the right part of (12), as the number of scenarios increases.

5.2 Student's t distribution

If we assume that portfolio return has a Student's t distribution, then the degrees of freedom parameter ν determines the tail thickness. In the notation of the theorem, $\alpha = \nu$ and, therefore, if $\nu \geq 2$, then the asymptotic distribution of the AVaR estimator is normal. Even though it is normal, the tail thickness influences the convergence rate which means that the minimum number of scenarios that guarantee the limit distribution can be accepted as an approximate model decrease as ν increases. The rationale is that the higher ν is, the more regular the random variable is.

The first two columns of Table 1 show how the number of scenarios changes when $\epsilon = 0.01$ and $\epsilon = 0.05$ which correspond to the two standard choices of 95% and 99% for the confidence level. The numbers in the table are calculated by generating 2,000 samples of a given size which we use to approximate the distribution of the left part of equation (12). We use the Kolmogorov distance in order to check if the hypothesis of a normal distribution can be accepted. We notice that the minimum number of scenarios increases from about 10,000 when $\nu = 25$ to 70,000 when $\nu = 3$ for $\epsilon = 0.01$. For additional details, see Stoyanov and Rachev (2008*a*).

5.3 The effect of tail truncation

In the introduction, we noted that one way to deal with the issue that a heavy-tailed model may have an infinite volatility is to apply a tail truncation. This means that we truncate the tails of the distribution very far away from the center, e.g. at 0.1% and 99.9% quantiles. This is not so artificial as it may seem. Stock exchanges usually have regulations according to which trading stops if a given market index drops too much. Essentially, this regulation does not allow panick in the market to result in arbitrarily large losses. If the truncation is done far away from the center, the general shape of the distribution is preserved and the descriptive power of the model does not deteriorate.

The implementation of a tail truncation method increases dramatically the stochastic stability of the AVaR estimator in two ways. First, if the portfolio return distribution is such that the limit law of the AVaR estimator is stable, after tail truncation the limit law becomes normal. However, the convergence rate to the normal distribution depends on how far away from the center we truncate — the deeper we go into the tails, the slower the convergence rate. Second, if the limit law of the AVaR estimator is normal, tail truncation increases the convergence rate. The last two columns of Table 1 illustrate this observation when portfolio returns are assumed to have Student's t distribution for a truncation threshold equal to the 0.1% quantile. We notice that the minimum number of scenarios drops from 70,000 to 12,000 when $\nu = 3$ and $\epsilon = 0.01$. Stoyanov and Rachev (2008*a*) provide additional examples and details.

6 Summary

In this paper, we provided a review of the essential components of a model for portfolio risk estimation in volatile markets. First, the model has to be capable of describing well the marginal distribution phenomena of the returns series such as fat-tails, skewness, and clustering of volatility. Second, the model has to capture the dependence structure which can be done through a copula function. Finally, the risk model has to incorporate a true risk measure. We considered the AVaR risk measure which has a practical meaning and appealing properties. It allows for building a risk budgeting framework based on Monte Carlo scenarios produced from a fat-tailed probabilistic model. We discussed the stochastic stability of the sample AVaR estimator with fat-tailed scenarios and a tail-truncation method as a way to increase stochastic stability.

Acknowledgements Prof Rachev gratefully acknowledges research support by grants from Division of Mathematical, Life and Physical Sciences, College of Letters and Science, University of California, Santa Barbara, the Deutschen Forschungsgemeinschaft and the Deutscher Akademischer Austausch Dienst.

References

Artzner, P., F. Delbaen, J.-M. Eber and D. Heath (1998), 'Coherent measures of risk', *Math. Fin.* **6**, 203–228.

Bibby, B. M. and M. Sorensen (2003), 'Hyperbolic processes in finance', *Handbook of heavy-tailed distributions in finance, S. Rachev (ed)* pp. 212–248.

Bradley, B. and M. S. Taqqu (2003), 'Financial risk and heavy tails', *In Handbook of Heavy-Tailed Distributions in Finance, S. T. Rachev, ed. Elsevier, Amsterdam* pp. 35–103.

Dokov, S., S. Stoyanov and S. Rachev (2008), 'Computing var and avar of skewed t distribution', *Journal of Applied Functional Analysis* **3**, 189–209.

Embrechts, P., A. McNeil and D. Straumann (2002), 'Correlation and dependence in risk management: properties and pitfalls', *In Risk management: Value at risk and beyond, edited by M. Dempster, published by Cambridge University Press, Cambridge* **10, (3)**, 341–352.

Embrechts, P., F. Lindskog and A. McNeil (2003), 'Modelling dependence with copulas and applications to risk management', *In: Handbook of heavy-tailed distributions in finance, ed. S. Rachev* pp. 329–384.

Fama, E. (1963), 'Mandelbrot and the stable paretian hypothesis', *Journal of Business* **36**, 420–429.

Fama, E. (1965), 'The behavior of stock market pricess', *Journal of Business* **38**, 34–105.

Hurst, S. H., E. Platen and S. Rachev (1997), 'Subordinated market index models: A comparison', *Financial Engineering and the Japanese Markets* **4**, 97–124.

Janicki, A. and A. Weron (1994), *Simulation and chaotic behavior of alpha-stable stochastic processes*, Marcel Dekker, New York.

Kim, Y., S. Rachev, M. Bianchi and F. Fabozzi (2008), 'Financial market models with levy processes and time-varying volatility', *Journal of Banking and Finance* **32,7**, 1363–1378.

Mandelbrot, B. (1963), 'The variation of certain speculative prices', *Journal of Business* **26**, 394–419.

Mittnik, S. and S. Rachev (1999), 'Stable non-gaussian models in finance and econometrics', *Mathematical and Computer Modeeling* **10-12**.

Pflug, G. (2000), 'Some remarks on the value-at-risk and the conditional value-at-risk', *In: Uryasev, S. (Ed.), Probabilistic Constrained Optimization: Methodology and Applications. Kluwer Academic Publishers, Dordrecht* .

Platen, E. and R. Rendek (2007), 'Empirical evidence on student-t log-returns of diversified world stock indices', *Research Paper Series 194, Quantitative Finance Research Centre, University of Technology, Sydney* .

Rachev, S. (2003), *Handbook of heavy-tailed distributions in finance*, editor, Rachev, Elsevier/North-Holland.

Rachev, S., D. Martin, B. Racheva-Iotova and S. Stoyanov (2006), 'Stable etl optimal portfolios and extreme risk management', *forthcoming in Decisions in Banking and Finance, Springer/Physika, 2007* .

Rachev, S., D. Martin, B. Racheva-Iotova and S. Stoyanov (2007), 'Stable etl optimal portfolios and extreme risk management', *in Risk Assessment: Decisions in Banking and Finance, G. Bol, S. Rachev and R. Wuerth (Editors), Springer - Physika Verlag* .

Rachev, S., F. Fabozzi and C. Menn (2005*a*), *Fat-tails and skewed asset return distributions*, Wiley, Finance.

Rachev, S., F. Fabozzi and C. Menn (2005*b*), *Fat tails and skewed asset returns distributions*, Wiley, Finance.

Rachev, S., S. Stoyanov, A. Biglova and F. Fabozzi (2006), 'An empirical examination of daily stock return distributions for u.s. stocks', *in Data Analysis and Decision Making, Springer* .

Rachev, S. T., Stoyan V. Stoyanov and F. J. Fabozzi (2008), *Advanced stochastic models, risk assessment, and portfolio optimization: The ideal risk, uncertainty, and performance measures*, Wiley, Finance.

Rachev, S.T. and S. Mittnik (2000), *Stable Paretian Models in Finance*, John Wiley & Sons, Series in Financial Economics.

Rockafellar, R. T. and S. Uryasev (2002), 'Conditional value-at-risk for general loss distributions', *Journal of Banking and Finance* **26, (7)**, 1443–1471.

Samorodnitsky, G. and M.S. Taqqu (1994), *Stable Non-Gaussian Random Processes*, Chapman & Hall, New York, London.

Sklar, A. (1996), 'Random variables, distribution functions and copulas — a personal look backward and forward.', *In: Rüschendorf, L., Schweizer, B., Taylor, M. D. (Eds.), Distributions with fixed marginals and related topics. Institute of Mathematical Statistics, Hayward, CA* pp. 1–14.

Stoyanov, S., G. Samorodnitsky, S. Rachev and S. Ortobelli (2006), 'Computing the portfolio conditional value-at-risk in the α-stable case', *Probability and Mathematical Statistics* **26**, 1–22.

Stoyanov, S. and S. Rachev (2008*a*), 'Asymptotic distribution of the sample average value-at-risk', *Journal of Computational Analysis and Applications* .

Stoyanov, S. and S. Rachev (2008*b*), 'Asymptotic distribution of the sample average value-at-risk in the case of heavy-tailed returns'.

Sun, W., S. Rachev, S. Stoyanov and F. Fabozzi (2008), 'Multivariate skewed student's t copula in the analysis of nonlinear and asymmetric dependence in the german equity market', *The Berkeley Electronic Press* .

Zhang, Y. and S. Rachev (2006), 'Risk attribution and portfolio performance measurement', *Journal of Applied Functional Analysis* **4, (1)**, 373–402.

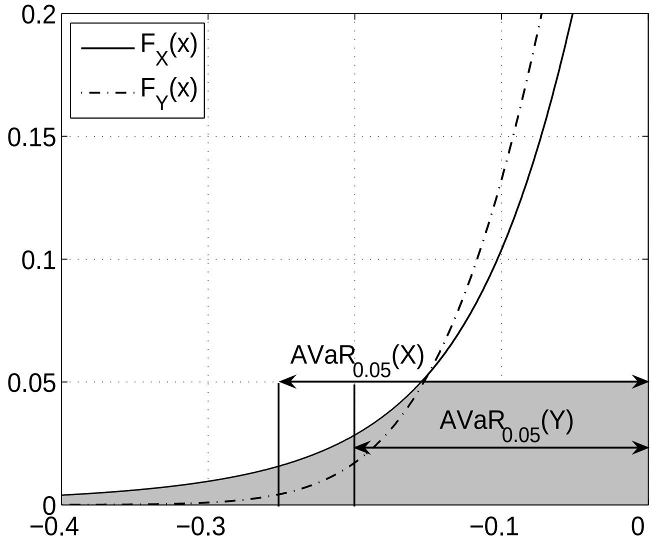

Fig. 1 Even though the VaRs at $\epsilon = 0.05$ are equal, $AVaR_{0.05}(X) > AVaR_{0.05}(Y)$.

Table 1 The number of observations sufficient to accept the normal distribution as an approximate model when X has Student's t distribution for different values of ν and ϵ (Reproduced from Table 1 and Table 4 in Stoyanov and Rachev (2008*a*).)

	No tail truncation		With tail truncation	
ν	$\epsilon = 0.01$	$\epsilon = 0.05$	$\epsilon = 0.01$	$\epsilon = 0.05$
3	70000	17000	12000	4000
4	60000	9000	11500	3600
5	50000	7000	11000	3300
6	23000	4500	11000	3200
7	14000	4200	10500	3100
8	13000	4100	10000	3000
9	12000	4000	10000	3000
10	12000	3900	10000	3000
15	11000	3850	10000	2950
25	10000	3800	10000	2900
50	10000	3750	10000	2900
∞	10000	3300	10000	2900

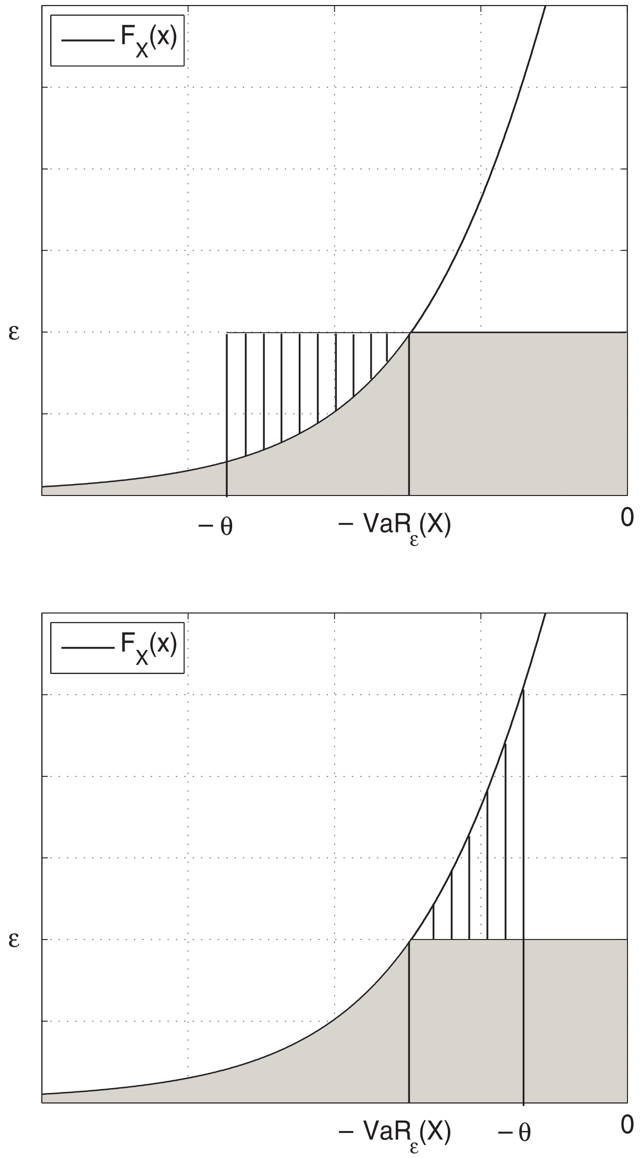

Fig. 2 The marked area is equal to zero if $\theta = VaR_{\epsilon}(X)$.

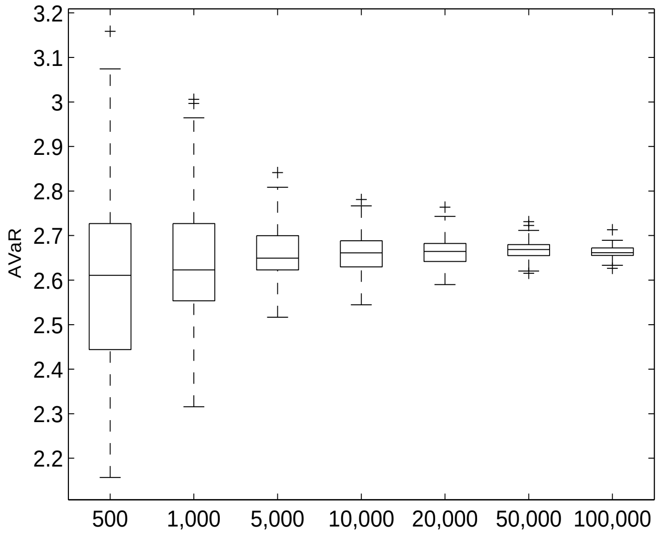

Fig. 3 Boxplot diagrams of the fluctuation of the AVaR at $\epsilon = 1\%$ of the standard normal distribution based on 100 independent samples (Reproduced from Figure 7.4 in Rachev et al. (2008))

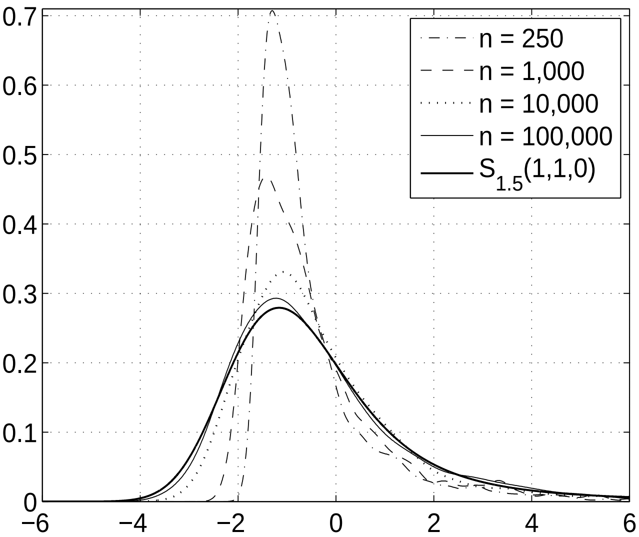

Fig. 4 The density of the sample AVaR as the number of scenarios increases together with the limit stable law, $X \in S_{1.5}(1, 0.7, 0)$ (Reproduced from Figure 2 in Stoyanov and Rachev (2008*b*))