数理解析研究所講究録
第1818巻 2012年 91-112

Quantitative Operational Risk Management: Properties of Operational Value at Risk (OpVaR)

Takashi Kato
Division of Mathematical Science for Social Systems, Graduate School of Engineering Science
Osaka University

1 Introduction

Basel II (International Convergence of Capital Measurement and Capital Standards: A Revised Framework) was published in 2004, and in it, operational risk was added as a new risk category. The Basel Committee on Banking Supervision [2] defines operational risk as "the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events. This definition includes legal risk, but excludes strategic and reputational risk" (see also McNeil et al. [19]).

For measuring the capital charge for operational risk, banks may choose from three approaches: the basic indicator approach (BIA), the standardized approach (SA), and the advanced measurement approach (AMA). While BIA and SA provide explicit formulas, AMA does not specify a model for quantifying risk amount (risk capital). Hence, banks adopting AMA must construct their own quantitative risk models and conduct periodic verification.

Basel II states that "a bank must be able to demonstrate that its approach captures potentially severe 'tail' loss events", as well as that "a bank must demonstrate that its operational risk measure meets a soundness standard comparable to a one year holding period and a 99.9th percentile confidence interval" [2]. The value-at-risk (VaR) with a confidence level of 0.999 is a typical risk measure, and therefore we adopt such operational VaR (abbreviated as OpVaR).

In this paper, we focus the following two topics:

(i) Analytical methods for calculating OpVaR

(ii) Asymptotic behavior of OpVaR

Item (i) is important from a practical rather than theoretical viewpoint. In fact, many banks adopt the so-called loss distributional approach (LDA) and calculate OpVaR by using Monte Carlo (MC) simulations. However, MC simulations are not optimal in terms of computation speed and robustness. We point out the problems associated with MC and introduce alternative analytical methods for calculating OpVaR in the LDA model.

Item (ii) is a theoretical issue. It is well known that distributions of operational risk amount are characterized by fat tails. In this study, we show the asymptotic behavior of the difference between the VaRs $\mathrm{VaR}_\alpha(L+S)$ and $\mathrm{VaR}_\alpha(L)$ (α denotes the confidence level of VaR) for heavy-tailed random variables L and S with $\alpha \to 1(=100\%)$ as an application to the sensitivity analysis of quantitative operational risk management in the framework of AMA. Here, the variable L denotes the loss amount of the current risk profile, and S indicates the loss amount caused by an additional loss factor. We obtain different results depending on the magnitude relation between

the thicknesses of the tails of L and S. In particular, if the tail of S is sufficiently thinner than that of L, then the difference between prior and posterior risk amounts is asymptotically equivalent to the expected loss of S, i.e., $\mathrm{VaR}_\alpha(L+S) - \mathrm{VaR}_\alpha(L) \sim \mathrm{E}[S], \quad \alpha \to 1$.

2 Analytical methods for calculating OpVaR

2.1 LDA Model

As mentioned above, banks who adopt AMA can use an arbitrary model to estimate OpVaR. In actuality, however, most banks choose the LDA, in this section, we briefly introduce the definition of LDA.

Let L be a random variable that denotes the total loss amount in one year. In the LDA framework, the distribution of L is constructed on the basis of the following two distributions:

- Loss severity distribution $\mu \in \mathcal{P}([0,\infty))$,
- Loss frequency distribution $\nu \in \mathcal{P}(\mathbb{Z}_+)$.

Here, we use $\mathcal{P}(D)$ to denote the set of all probability distributions defined on the space D, where $\mathbb{Z}_+ = \{0,1,2,\ldots\}$. Let N be a random variable distributed by ν and $(L_k)_k$ be a sequence of identically distributed random variables distributed by μ. We regard N as the number of loss events in a year and L_k as the amount of the kth loss. Then, the total loss amount L is given by

$$L = \sum_{k=1}^{N} L_k = L_1 + \cdots + L_N. \tag{2.1}$$

We assume that all random variables $N, L_1, L_2, \ldots$ are independent.

This model is closely related to the so-called Cramer-Lundberg model, which is widely used in actuarial science. We note that L becomes the compound Poisson model when ν is the Poisson distribution. In operational risk management, the LDA model is applied to each event type or business line. For instance, if a bank has M event types, then the total loss amount L^i of the ith event type is given by (2.1) with a severity distribution μ_k and a frequency distribution ν_k. Then, the "bank-wide" total loss amount is given by the sum of $L^1, \ldots, L^M$. However, for simplicity, we ignore the characteristics of different event types and business lines in this paper since our purpose is to calculate the VaR and OpVaR

$$\mathrm{VaR}_\alpha(L) = \inf\{x \in \mathbb{R} \ ; \ P(L \le x) \ge \alpha\} \tag{2.2}$$

of L as defined in (2.1) with a confidence level $\alpha = 0.999$.

A characteristic property of distributions of operational losses is a fat tail. It is well known that operational loss distributions are strongly affected by loss events with low frequency and substantial loss amount. Thus, we should set μ as a heavy-tailed distribution to capture such events. Typical examples of such distributions are the log-normal distribution (LND) and the

generalized Pareto distribution (GPD):

$$\begin{aligned}\mathrm{LND}(\gamma,\sigma)((-\infty,x]) &= \int_0^x \frac{1}{\sqrt{2\pi}\sigma y}\exp\left(-\frac{(\log y-\gamma)^2}{2\sigma^2}\right)dy \ (x\geq 0),\ \ 0\ (x<0),\\ \mathrm{GPD}(\xi,\beta)((-\infty,x]) &= 1-\left(1+\frac{\xi}{\beta}x\right)^{-1/\xi} \quad (x\geq 0),\ \ 0\ (x<0).\end{aligned}$$

Here, $\gamma,\beta>0$ are location parameters and $\sigma,\xi>0$ are shape parameters. Larger values of σ and ξ yield a fatter tail of the severity distribution. In fact, GPD is a representative fat-tailed distribution and plays an essential role in extreme value theory (EVT). For details on EVT, see [11].

2.2 MC Simulation

The most widely used method for calculating (2.2) is MC simulation. The algorithm for calculating OpVaR by MC simulation is as follows:

1. Generate a (pseudo)random variable N distributed by ν.
2. Generate i.i.d. (pseudo)random variables $L_1,\dots,L_N$ distributed by μ.
3. Put $L=L_1+\cdots+L_N$.
4. Repeat steps 1–3 m times (m is the number of simulation iterations). Then, we obtain m independent copies of L.
5. Sort the variables $L^1,\dots,L^m$ such that $L^{(1)}\leq\cdots\leq L^{(m)}$.
6. Set $\mathrm{VaR}_\alpha^{\mathrm{MC}}(L)=L^{([m\alpha])}$, where $[x]$ is the largest integer not greater than x.

The estimator $\mathrm{VaR}_\alpha^{\mathrm{MC}}(L)$ is an order statistic of $\mathrm{VaR}_\alpha(L)$. Letting $m\to\infty$, $\mathrm{VaR}_\alpha^{\mathrm{MC}}(L)$ converges to $\mathrm{VaR}_\alpha(L)$ under some technical conditions.

MC is an extremely useful method which is widely adopted in practice because of its ease of implementation and wide applicability. However, it is known to suffer from several problems, one of which is that it requires a large number of simulation iterations (i.e., m should be very large). The reason is that convergence of $\mathrm{VaR}_\alpha^{\mathrm{MC}}(L)$ is generally very slow. Another problem is that the estimated value $\mathrm{VaR}_\alpha^{\mathrm{MC}}(L)$ is unstable for some specific distributions. In the following subsections, we examine these problems numerically.

2.2.1 Test 1: Simulation Iterations and Computation Time

Until the end of Section 2, we adopt the Poisson distribution with an intensity λ as the frequency distribution ν :

$$\nu(\{k\})=\mathrm{Poi}(\lambda)(\{k\})=\frac{\lambda^k e^{-\lambda}}{k!},\quad k\in\mathbb{Z}_+.$$

We investigate the number of simulation iterations m required for estimating OpVaR with a tolerance lower than 1%. It is difficult to obtain a strict number, so we calculate m approximately

by using the Pritsker method with a 95% confidence level. We iterate the above MC algorithm while increasing m until

$$\frac{\mathrm{VaR}_\alpha^{\mathrm{MC,upper}}(L) - \mathrm{VaR}_\alpha^{\mathrm{MC,lower}}(L)}{\mathrm{VaR}_\alpha^{\mathrm{MC}}(L)} < 0.01,$$

is satisfied, where $\mathrm{VaR}_\alpha^{\mathrm{MC,upper}}(L)$ and $\mathrm{VaR}_\alpha^{\mathrm{MC,lower}}(L)$ are order statistics satisfying

$$P(\mathrm{VaR}_\alpha^{\mathrm{MC,lower}}(L) \leq \mathrm{VaR}_\alpha(L) < \mathrm{VaR}_\alpha^{\mathrm{MC,upper}}(L)) \simeq 0.95.$$

For an explanation of how to obtain the values of $\mathrm{VaR}_\alpha^{\mathrm{MC,upper}}(L)$ and $\mathrm{VaR}_\alpha^{\mathrm{MC,lower}}(L)$, please refer to [26]. Note again that the estimated value of m obtained by the above procedure is not exact. Nevertheless, it is obvious that a substantial number of simulation iterations are necessary to obtain the estimator of OpVaR with tolerable accuracy. Figure 1 gives the number of simulation iterations in the case of $\mu = \mathrm{LND}(\gamma, \sigma)$. We see that many iterations are necessary when σ is large and λ is small. Figure 2 corresponds to the case of $\mu = \mathrm{GPD}(\xi, \beta)$. This implies that the number of iterations is similarly large when ξ is large, but in contrast to the case of LND, the number of iterations is unaffected by the value of λ. This difference is caused by the tail probability function of LND is rapidly varying while that of GPD are regularly varying (see [4] and [11] for details). Here, we remark that the location parameters γ and β do not influence the number of simulation iterations.

These results indicate that the MC method requires a huge number of simulation iterations, especially when the values of the shape parameters σ and ξ are large, that is, when μ is fat-tailed. In this case, long computation time is required: if we conduct a calculation with $\xi = 3$ and $\lambda = 500$, as in Figure 2, the computation would require more than 2 days to complete on a typical personal computer.

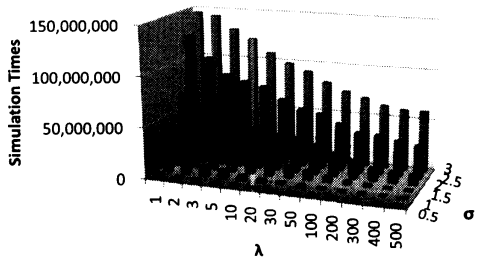

Figure 1: The number of simulation iterations required for estimating the error of OpVaR with tolerance $< 1\%$ with $\mu = \mathrm{LND}(\gamma, \sigma)$ and $\nu = \mathrm{Poi}(\lambda)$.

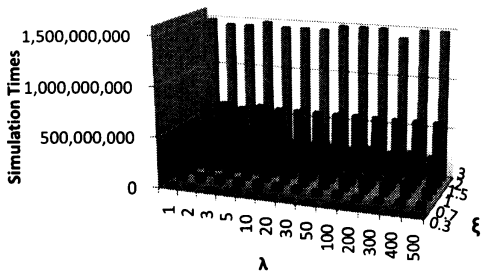

Figure 2: The number of simulation iterations required for the estimation of the error of OpVaR with tolerance $< 1\%$ with $\mu = \mathrm{GPD}(\xi, \beta)$ and $\nu = \mathrm{Poi}(\lambda)$.

Although the computation can be accelerated by employing parallel computing techniques,

such as multithread programming and general-purpose computing on graphics processing units (GPGPU), this entails high development cost.

2.2.2 Test 2: Accuracy for Specific Distributions

In this test, we calculate $\mathrm{VaR}_\alpha^{\mathrm{MC}}(L)$ in the case where μ is derived empirically. We generate 500 sets of dummy loss data ("virtual" realized data) and 15 sets of scenario data based on $\mathrm{GPD}(1.2, 100,000)$. Then, we set μ as the empirical distribution defined by these data. Figure 3 shows the tail probability function of the severity distribution on a log-log scale. We also take $\nu = \mathrm{Poisson}(500)$.

We conduct 100 calculations of $\mathrm{VaR}_{0.999}^{\mathrm{MC}}(L)$ (with $1,000,000$ iterations) as OpVaR. Figure 4 shows a histogram of 100 estimations of OpVaR. Clearly, the distribution of OpVaRs is bi-polarized. This can be explained by using Theorem 3.12 in [5]: $\mathrm{VaR}_{0.999}(L)$ is approximated by $\hat{v} \equiv \mathrm{VaR}_{0.9998}(L_1)$. Here, $\hat{v}$ is the 99.9998th percentile point of μ. In fact, the empirical distribution μ has a large point mass at the 99.99988th percentile point. This is near $\hat{v}$, and thus the value of the simulated OpVaR is sensitive to the above loss data.

This indicates that the MC method may result in serious estimation error in calculating OpVaR. Note that the severity distribution used in this numerical experiment is not far from a standard distribution, and such a phenomenon is conceivable in practice.

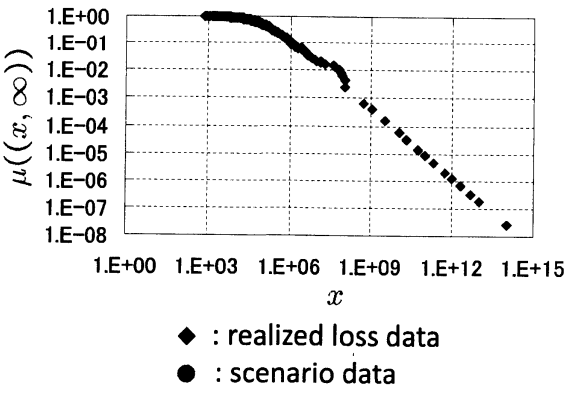

Figure 3: Log-log graph of the tail probability function $\mu((x,\infty)) = P(L_1 > x)$.

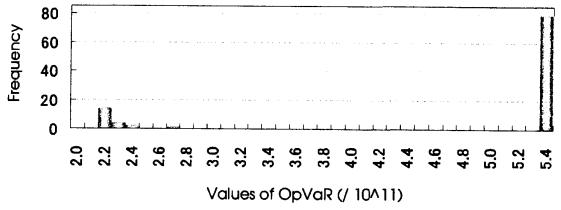

Figure 4: Histogram of 100 estimations of OpVaR.

2.3 Analytical Computation Methods

In Section 2.2, we point out some of the problems associated with the MC method. In this section, we introduce some alternative methods for calculating OpVaR.

There have been studies on probabilistic approximations for the LDA model. In particular, a closed-form approximation (Theorem 3.12 in [5]) is well-known and widely used for rough

calculation of OpVaR:

$$\mathrm{VaR}_\alpha^{\mathrm{EVT}}(L) = \mathrm{VaR}_{1-(1-\alpha)/\lambda}(L_1).$$

If μ is subexponential, then $\mathrm{VaR}_\alpha^{\mathrm{EVT}}(L)$ converges to $\mathrm{VaR}_\alpha(L)$ as $\alpha \to 1$. Thus, $\mathrm{VaR}_\alpha^{\mathrm{EVT}}(L)$ approximates $\mathrm{VaR}_\alpha(L)$ when α is close to 1. Usually, we can calculate $\mathrm{VaR}_\alpha^{\mathrm{EVT}}(L)$ almost instantly because we assign the explicit form of μ in most cases. This approximation method also plays an important role in EVT, and in Section 3, we present some results for the asymptotic behavior of OpVaRs with $\alpha \to 1$ which shows a close correspondence to the result of [5].

Although this approximation method can provide an extremely fast way for calculating OpVaR, it is susceptible to approximation errors. Here, we give some methods for direct calculation of OpVaR without mathematical approximation by using techniques from numerical analysis. Although there still remain numerical errors, we use direct evaluation methods to avoid theoretical approximation errors.

2.3.1 Direct Approach

Let us recall the definition of the total loss amount L (2.1). Since we assume that $(L_k)_k$ and N are mutually independent and that $\nu = \mathrm{Poi}(\lambda)$ for some $\lambda > 0$, Kac's theorem implies that the characteristic function φ_L can be written as

$$\varphi_L(\xi) = \mathrm{E}[e^{\sqrt{-1}\xi L}] = \exp(\lambda(\varphi(\xi) - 1)),$$

where φ is the characteristic function of μ. Moreover, using Lévy's inversion formula, we obtain

$$P(a < L < b) + \frac{1}{2}P(L = a \text{ or } L = b) = \lim_{T\to\infty} \frac{1}{2\pi}\int_{-T}^{T} \frac{\exp(-ia\xi) - \exp(-ib\xi)}{i\xi}\varphi_L(\xi)d\xi$$

for each $a < b$. Since L is non-negative, by substituting $a = -x$ and $b = x$ into the above equality, we obtain the following Fourier inversion formula

$$F_L(x) = \frac{2}{\pi}\int_0^\infty \frac{\mathrm{Re}\varphi_L(\xi)}{\xi}\sin(x\xi)d\xi + \frac{1}{2}P(L = x), \quad x \geq 0. \tag{2.3}$$

If μ has no point mass, then the second term on the right-hand side of (2.3) vanishes. Otherwise (e.g., in case μ is an empirical distribution), we adopt an approximation such as $P(L = x) \simeq P(\text{a single event with a loss amount of } \$x \text{ occurs})$. Using (2.3), we can calculate OpVaR ($= \mathrm{VaR}_{0.999}(L)$) by solving the root-finding problem $F_L(\mathrm{OpVaR}) - 0.999 = 0$. In our numerical experiments presented in the next section, we search for a solution of this problem on the interval $[0.3 \times \mathrm{VaR}_{0.999}^{\mathrm{EVT}}, 3 \times \mathrm{VaR}_{0.999}^{\mathrm{EVT}}]$ by using Brent's method (if $\mathrm{VaR}_{0.999}(L)$ is not in this interval, we expand the search interval).

Then, our main task here is to calculate the first term on the right-hand side of (2.3). This is an oscillatory integral, where the integrand ξ oscillates near 0. Now, we present a method to avoid the difficulties associated with calculating this integral.

We rewrite the integral in (2.3) as

$$\frac{2}{\pi}\int_0^\infty \frac{\mathrm{Re}\varphi_L(t/x)}{t}\sin t dt = \sum_{k=0}^{\infty}(-1)^k a_k, \tag{2.4}$$

where

$$a_k = \frac{2}{\pi}\int_0^\pi \frac{\mathrm{Re}\varphi_L((k\pi+t)/x)}{k\pi+t}\sin t dt. \tag{2.5}$$

If we know the values of $(a_k)_k$, then we can calculate $F_L(x)$ by summing the terms. Let us denote by $c(\xi)$ (resp. $d(\xi)$) the real part (resp. the imaginary part) of $\varphi(\xi)$. Then we can rewrite (2.5) as

$$a_k = \frac{2}{\pi}\int_0^\pi \frac{\exp(\lambda(c(k\pi+t)-1))\cos(\lambda d(k\pi+t))}{k\pi+t}\sin t dt. \tag{2.6}$$

We can calculate this integral quickly by using the Takahashi–Mori double exponential (DE) formula ([28], [29]) when $c(\xi)$ and $d(\xi)$ are known and analytic on $(0,\pi)$. For the case of $k=0$, we omit the integration on $(0,10^{-8})$ because the integrand is unstable near $t=0$.

If the severity function μ has a density function f, then we can write $c(\zeta)$ and $d(\zeta)$ as

$$c(\xi) = \int_0^\infty f(t)\cos(t\xi)dt, \quad d(\xi) = \int_0^\infty f(t)\sin(t\xi)dt.$$

We can calculate these oscillatory integrals numerically by using the Ooura–Mori DE formula ([23], [24]).

Note that we can also calculate a_k numerically when μ is an empirical distribution such as

$$\mu = \sum_{k=1}^n p_k \delta_{c_k}, \tag{2.7}$$

where δ_x denotes the Dirac measure. Indeed, we have

$$c(\xi) = \sum_{k=1}^n p_k \exp(\cos(\xi c_k)), \quad d(\xi) = \sum_{k=1}^n p_k \exp(\sin(\xi c_k)).$$

Upon substituting these terms into (2.6), we can apply the Takahashi–Mori DE formula.

Moreover, in evaluating the right-hand side of (2.4), we apply Wynn's ε-algorithm to accelerate the convergence of the sum. We define the double-indexed sequence $(\varepsilon_{r,k})_{k\geq 0, r\geq -1}$ by

$$\varepsilon_{-1,k} = 0, \qquad \varepsilon_{0,k} = \sum_{l=0}^k (-1)^l a_l, \qquad \varepsilon_{r+1,k} = \varepsilon_{r-1,k+1} + \frac{1}{\varepsilon_{r,k+1} - \varepsilon_{r,k}}.$$

Then, for a fixed even number r, the convergence of $(\varepsilon_{r,k})_k$ becomes faster than the original series $(\varepsilon_{0,k})_k$. Please refer to [31] for more details. We stop updating of the sequence $(\varepsilon_{r,k})_k$ when $|\varepsilon_{r,k} - \varepsilon_{r,k-1}| < \delta$ for a small constant $\delta > 0$.

2.3.2 Other Methods

In Section 2.3.1, we introduced an analytical computation method using the inverse Fourier transform. Similarly, Luo and Shevchenko [18] developed an inverse Fourier transform approach known as direct numerical integration (DNI).

On the other hand, there have been studies on other calculation methods based on the discrete Fourier transform. Some of the most representative methods using this approach are

the Panjer recursion and Fast Fourier Transform (FFT) with tilting. Algorithms for these methods can be found in [10] and [27].

Ishitani and Sato [20] have constructed another computation method using wavelet transform, which is similar to the method in Section 2.3.1 and also uses the DE formula proposed by Ooura and Mori [23] and the ε-algorithm proposed by Wynn [31].

When μ is an empirical distribution, such as that in (2.7), we can calculate the distribution function $F_L(x) = P(L \leq x) = F^{(n)}(x)$ by a simple convolution method:

$$\begin{aligned} F_L^{(0)}(x) &= 1_{(0,\infty)}(x), \\ F_L^{(k)}(x) &= \sum_{j=0}^{\infty} F_L^{(k-1)}(x - jc_k) \times \frac{e^{-\lambda p_k}(\lambda p_k)^j}{j!}, \quad k = 1, \ldots, n. \end{aligned} \tag{2.8}$$

This inductive calculation easily gives the value of OpVaR.

There have also been studies on the use of the importance sampling (IS) method for accelerating the convergence speed in MC simulations. Although a certain amount of skill is required for the effective implementation of IS, it is certainly a promising option.

2.3.3 Numerical Experiments: Comparison of Accuracy and Computation Time

Recently, a comparison of the precisions of DNI, the Panjer recursion and FFT was studied by Shevchenko [27]. The results indicated that FFT is fast and accurate for relatively small λ, and that DNI is effective for large λ. Similarly to [27], in this section we examine the accuracies and computation times of the analytical methods introduced above.

As a measure of the accuracy of the methods, we compute the relative error (RE) defined as

$$\mathrm{RE} = \frac{\widehat{\mathrm{OpVaR}} - \mathrm{OpVaR}}{\mathrm{OpVaR}},$$

where OpVaR denotes the actual value of$\mathrm{VaR}_{0.999}(L)$ and $\widehat{\mathrm{OpVaR}}$ denotes the estimated value of OpVaR calculated by the method under test. Since it is difficult to obtain an exact value for OpVaR, we use $\mathrm{VaR}_{0.999}^{\mathrm{MC}}(L)$ with 10 billion simulation iterations ($m = 10^{10}$) as OpVaR. Note here that 10^{10} iterations are not sufficient for estimating OpVaR with an error that is negligible compared with the approximation errors of analytical methods, especially in the case of μ = GPD. However, conducting MC simulations with larger m on a standard personal computer is unrealistic.

Our parameter settings are introduced below. We always assume that the frequency distribution ν is the Poisson distribution. The intensity parameter λ is set to be in the interval $[1, 10^4]$. For the severity distribution μ, we test the following four patterns:

LND $\mu = \mathrm{LND}(5, 2)$

GPD $\mu = \mathrm{GPD}(2, 10)$

EMP μ is the empirical distribution used in Test 2 in Section 2.2.2

LIN-EMP A linearly interpolated version of EMP: $\mu((-\infty, x]) = \sum_{l=1}^{k} p_l + \frac{x-c_k}{c_{k+1}-c_k} \times p_{k+1}$ for $x \geq 0$, where k is the largest integer such that $\sum_{l=1}^{k} c_l \leq x$, and (c_k, p_k) is the kth pair of the loss data (loss amount and frequency) used in Test 2 in Section 2.2.2

We verify the accuracies and computation times of the following methods:

Direct The direct approach introduced in Section 2.3.1 with a tolerance of 10^{-8} as used in the DE formulas and Brent's method, and with $r = 8$ and $\delta = 10^{-8}$ for the ε-algorithm.

Panjer The Panjer recursion given in [10] with a number of partitions $M = 2^{17}$ and a step size $h = \mathrm{VaR}_{0.999}^{\mathrm{EVT}}/(3M)$

FFT An FFT-based method given in [10] with the same M, h combination as in Panjer and a tilting parameter $\theta = 20/M$, as suggested in [10]

In the case of EMP, we examine the following additional method:

Convol The convolution method as introduced in (2.8)

In the following numerical calculations, computation times are quoted for a standard personal computer with a 3.33-GHz Intel® Core™ i7 X980 CPU and 6.00 GB of RAM.

Figures 5–8 show the REs for each method. We can see that the REs for most methods are less than 0.2%. Here, the REs in Figure 6 are rather similar. This implies that there are non-negligible simulation errors in the OpVaRs themselves because of the fat tail property of GPD. Therefore, the accuracies of Direct, FFT and Panjer are higher than the ones shown in Figure 6.

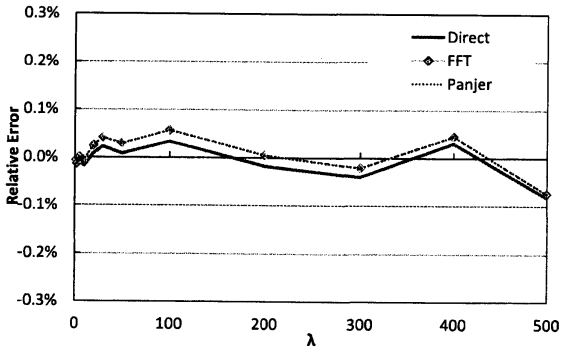

Figure 5: Relative error in the case of LND.

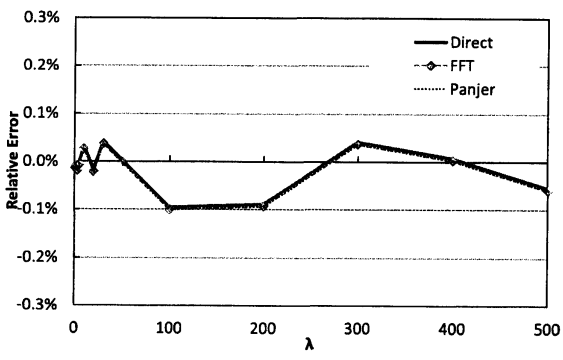

Figure 6: Relative error in the case of GPD.

Table 1 presents the computation times for each method. We can see that FFT is rather fast in each case. On the other hand, Direct is sufficiently fast in the cases of LND and GPD, whereas its computation speed is somewhat lower in the case of LIN-EMP and even lower in the case of EMP.

From the above results, we see that FFT is one of the most adaptive methods for calculating OpVaR. As an additional experiment, we investigate the accuracies of the methods with large

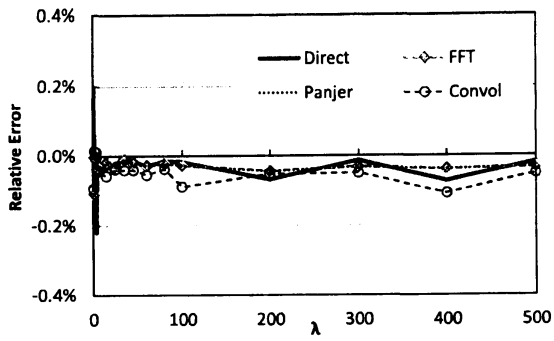

Figure 7: Relative error in the case of EMP.

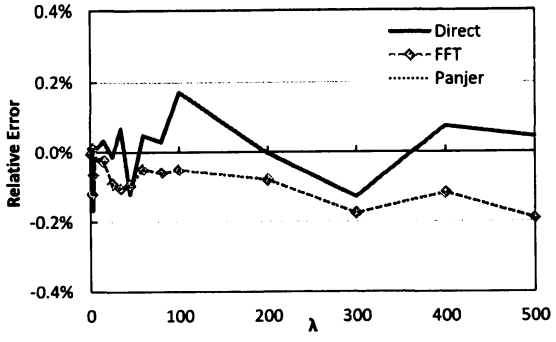

Figure 8: Relative error in the case of LIN-EMP.

Table 1: Computation times (s)

	LND			GPD			EMP				LIN-EMP		
λ	Direct	FFT	Panjer	Direct	FFT	Panjer	Direct	Convol	FFT	Panjer	Direct	FFT	Panjer
1	0.18	0.05	6.93	0.72	0.03	6.85	119.83	5.77	0.13	3.53	24.72	0.031	0.374
5	0.21	0.03	6.61	0.69	0.03	6.94	105.71	7.11	0.13	1.53	12.74	0.031	0.39
10	0.20	0.03	6.82	0.73	0.03	6.88	94.84	8.47	0.12	0.73	9.69	0.047	0.39
50	0.23	0.03	7.10	0.93	0.06	7.32	87.48	11.16	0.14	1.55	7.27	0.03	0.4
100	0.22	0.03	7.00	0.87	0.03	6.93	91.27	13.84	0.14	1.03	5.53	0.031	0.39
200	0.30	0.03	6.60	0.98	0.03	6.76	151.79	15.21	0.14	1.34	6.04	0.031	0.406
300	0.33	0.03	6.77	1.13	0.03	6.88	38.20	17.89	0.14	1.54	5.72	0.032	0.405
400	0.40	0.03	7.05	1.15	0.03	6.83	119.10	19.25	0.13	1.54	9.12	0.031	0.406
500	0.37	0.08	6.86	1.18	0.03	7.19	88.91	20.63	0.14	1.33	4.55	0.031	0.405

λ. We set $\nu = \mathrm{Poi}(\lambda)$ with $\lambda = 1, 10^3, 10^4$ and $\mathrm{GPD}(1, 1)$. Here, we also perform a comparison with the wavelet transform approach (abbreviated as Wavelet) presented in [20].

The results are shown in Tables 2–3. We can see that the accuracies of FFT and Panjer are somewhat lower when λ is large, whereas the errors for Direct and Wavelet are still less than 0.1% even when $\lambda = 10^4$. This phenomenon is consistent with the results presented in [27]. Obviously, the accuracies of FFT and Panjer can be improved by increasing the number of partitions M, which entails a longer computation time. Figure 9 shows a comparison of the REs and computation times for Direct, FFT and Wavelet, where the number of partitions M for FFT is varied between 2^{17} and 2^{24}. We can see that Direct is the fastest and most accurate method in this case.

2.4 Concluding Remarks

As we have seen in Section 2.2, the MC method is slow and not robust in the calculation of VaRs of fat-tailed distributions, such as OpVaRs. The numerical results in the preceding section imply that FFT is fast and highly accurate in many cases (including many realistic situations) but becomes slower when λ is large, in which case Direct is more effective than FFT.

Table 2: Relative error in the case of $\mu = \mathrm{GPD}(1,1)$ and $\nu = \mathrm{Poi}(\lambda)$.

λ	Direct	FFT	Panjer	Wavelet
10	-0.02%	-0.02%	-0.02%	-0.03%
1,000	-0.03%	-0.17%	-0.17%	-0.01%
10,000	-0.06%	-0.41%	-0.41%	-0.07%

Table 3: Computation times (s) in the case of $\mu = \mathrm{GPD}(1,1)$ and $\nu = \mathrm{Poi}(\lambda)$.

λ	Direct	FFT	Panjer	Wavelet
10	0.38	0.03	5.88	43.88
1,000	0.62	0.03	5.86	42.93
10,000	0.66	0.03	5.89	42.33

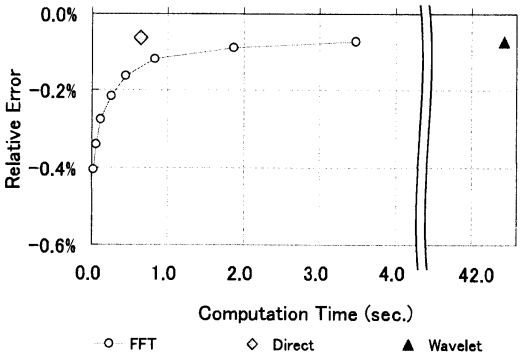

Figure 9: Comparison of relative errors and computation times in the case of $\mu = \mathrm{GPD}(1,1)$ and $\nu = \mathrm{Poi}(10,000)$. Here, M is taken as $2^{17}, 2^{18}, \ldots, 2^{24}$.

On the other hand, to our knowledge, MC is still the de facto standard method in banks, especially in Japan. Although it is true that the estimated value $\mathrm{VaR}_\alpha^{\mathrm{MC}}(L)$ converges to the actual value of OpVaR with increasing the number of simulation iterations, the cost of increasing the accuracy of the MC method without any theoretical improvements, such as IS, is rather high. Although we could not find the most adaptive method for all cases in practice, the methods introduced in this paper are sufficiently fast and accurate, and appear to be more reliable than the simple MC method. Thus, it might be of use to study the improvement of analytical methods and the application of the IS method in calculating OpVaRs.

3 Asymptotic behavior of OpVaR and Sensitivity Analysis

In the above section, we considered some methods for calculation of OpVaRs. Another notable issue for banks adopting AMA is the verification of their own models since AMA requires not only the calculation of OpVaR, but also the verification of the adequacy and robustness of the model used.

As pointed out in McNeil et al. [19], whereas everyone agrees on the importance of understanding operational risk, it is a controversial issue how far one should (or can) quantify such risks. Since empirical studies find that the distribution of operational loss has a fat tail (see Moscadelli [22],) this requires capturing the tail of the loss distribution. Meanwhile, estimating

the tail of an operational loss distribution is often difficult due to the fact that the accumulated historical data are insufficient, there are various kind of factors of operational loss, and so on. Thus we need sufficient verification for the appropriateness and robustness of the model in quantitative operational risk management.

One of the verification approaches for a risk model is sensitivity analysis (or behaviour analysis). There are a few interpretations for the word "sensitivity analysis". In this paper, we use this word to mean the relevance of a change of the risk amount with changing input information (for instance, added/deleted loss data or changing model parameters). There is also an advantage in using sensitivity analysis not only to validate the accuracy of a risk model but also to decide on the most effective policy with regard to the variable factors. This examination of how the variation in the output of a model can be apportioned to different sources of variations of risk will give an incentive to business improvement. Moreover, sensitivity analysis is also meaningful for a scenario analysis. Basel II claims not only to use historical internal/external data and BEICFs (Business Environment and Internal Control Factors) as input information, but also to use scenario analyses to evaluate low frequency and high severity loss events which cannot be captured by empirical data. As noted above, to quantify operational risk we need to estimate the tail of the loss distribution, so it is important to recognize the impact of our scenarios on the risk amount.

In this large section we study the sensitivity analysis for the operational risk model from a theoretical viewpoint. In particular, we mainly consider the case of adding loss factors. Let L be a random variable which represents the loss amount with respect to the present risk profile and let S be a random variable of the loss amount caused by an additional loss factor found by a minute investigation or brought about by expanded business operation. In a practical sensitivity analysis it is also important to consider the statistical effect (the estimation error of parameters, etc.) for validating an actual risk model, but such an effect should be treated separately. We focus on the change from a prior risk amount $\rho(L)$ to a posterior risk amount $\rho(L+S)$, where ρ is a risk measure.

We mainly treat the case where the tails of the loss distributions are regularly varying. We use VaR at the confidence level α as our risk measure ρ and we study the asymptotic behaviour of VaR as $\alpha \to 1$. Our framework is mathematically similar to the study of Böcker and Klüppelberg [6]. They regard L and S as loss amount variables of separate categories (cells) and study the asymptotic behaviour of an aggregated loss amount $\mathrm{VaR}_\alpha(L+S)$ as $\alpha \to 1$ (in addition, a similar study, adopting an expected shortfall (or conditional VaR), is found in Biagini and Ulmer [3]). In contrast, our purpose is to estimate a more precise difference between $\mathrm{VaR}_\alpha(L)$ and $\mathrm{VaR}_\alpha(L+S)$ and we obtain different results according to the magnitude relationship of the thicknesses of the tails of L and S.

The rest of this large section is organized as follows. In Section 3.1 we introduce the framework of our model and some notation. In Section 3.2 we give rough estimations of the asymptotic behaviour of the risk amount $\mathrm{VaR}_\alpha(L+S)$. Our main results are in Section 3.3 and we present a finer estimation of the difference between $\mathrm{VaR}_\alpha(L)$ and $\mathrm{VaR}_\alpha(L+S)$. Section 3.3.1 treats the case where L and S are independent. Section 3.3.2 includes a tiny generalization of the

results in Section 3.3.1 and we give some results when L and S are not independent. One of these results is related to the study of risk capital decomposition and we study these relations in Section 3.6. In Section 3.4 we present numerical examples of our results. Section 3.5 presents some conclusions. For the proofs of our results, see [17].

3.1 Settings

We always study a given probability space $(\Omega, \mathcal{F}, P)$. We recall the definition of the α-quantile (Value at Risk): for a random variable X and $\alpha \in (0,1)$, put

$$\mathrm{VaR}_\alpha(X) = \inf\{x \in \mathbb{R} \ ; \ F_X(x) \geq \alpha\},$$

where $F_X(x) = P(X \leq x)$ is the distribution function of X.

We denote by $\mathcal{R}_k$ the set of regularly varying functions with index $k \in \mathbb{R}$, that is, $f \in \mathcal{R}_k$ if and only if $\lim_{x\to\infty} f(tx)/f(x) = t^k$ for any $t > 0$. When $k = 0$, a function $f \in \mathcal{R}_0$ is called slowly varying. For the details of regular variation and slow variation, see Bingham et al. [4] and Embrechts et al. [11]. For a random variable X, we also say $X \in \mathcal{R}_k$ when the tail probability function $\bar{F}_X(x) = 1 - F_X(x) = P(X > x)$ is in $\mathcal{R}_k$. We mainly treat the case of $k < 0$. In this case, the mth moment of $X \in \mathcal{R}_k$ is infinite for $m > -k$. As examples of heavy-tailed distributions which have regularly varying tails, the generalized Pareto distribution (GPD) and the g-h distribution (see Degen et al. [7], Dutta and Perry [9]) are well-known and are widely used in quantitative operational risk management. In particular, GPD plays an important role in extreme value theory (EVT), and it can approximate the excess distributions over a high threshold of all the commonly used continuous distributions. See Embrechts et al. [11] and McNeil et al. [19] for details.

Let L and S be non-negative random variables and assume $L \in \mathcal{R}_{-\beta}$ and $S \in \mathcal{R}_{-\gamma}$ for some $\beta, \gamma > 0$. We call β (respectively, γ) the tail index of L (respectively, S). A tail index represents the thickness of a tail probability. For example, the relation $\beta < \gamma$ means that the tail of L is fatter than S.

We regard L as the total loss amount of a present risk profile. In the framework of LDA, L is given as (2.1). If we consider a multivariate model, L is given by $L = \sum_{k=1}^{d} L_k$, where L_k is the loss amount variable of the kth operational risk cell ($k = 1, \ldots, d$). We are aware of such formulations, but we do not limit ourselves to such situations in our settings.

The random variable S means an additional loss amount. We will consider the total loss amount variable $L + S$ as a new risk profile. As mentioned above, our interest is in how a prior risk amount $\mathrm{VaR}_\alpha(L)$ changes to a posterior one $\mathrm{VaR}_\alpha(L + S)$.

3.2 Basic Results of Asymptotic Behaviour of $\mathrm{VaR}_\alpha(L + S)$

First we give a rough estimations of $\mathrm{VaR}_\alpha(L + S)$. We introduce the following condition.

[A] A joint distribution of (L, S) satisfies the negligible joint tail condition when

$$\frac{P(L > x, S > x)}{\bar{F}_L(x) + \bar{F}_S(x)} \longrightarrow 0, \quad x \to \infty. \tag{3.1}$$

Then we have the following proposition.

Proposition 1 *Under condition* $[A]$ *it holds that*

(i) *If* $\beta < \gamma$, *then* $\mathrm{VaR}_\alpha(L+S) \sim \mathrm{VaR}_\alpha(L)$,

(ii) *If* $\beta = \gamma$, *then* $\mathrm{VaR}_\alpha(L+S) \sim \mathrm{VaR}_{1-(1-\alpha)/2}(U)$,

(iii) *If* $\beta > \gamma$, *then* $\mathrm{VaR}_\alpha(L+S) \sim \mathrm{VaR}_\alpha(S)$

as $\alpha \to 1$, *where the notation* $f(x) \sim g(x)$, $x \to a$ *denotes* $\lim_{x\to a} f(x)/g(x) = 1$ *and* U *is a random variable whose distribution function is given by* $F_U(x) = (F_X(x) + F_Y(x))/2$.

These results are easily obtained and not novel. In particular, when L and S are independent, this proposition is a special case of Theorem 3.12 in Böcker and Klüppelberg [6] (in the framework of LDA).

In contrast with Theorem 3.12 in Böcker and Klüppelberg [6], which implies an estimate for $\mathrm{VaR}_\alpha(L+S)$ as "an aggregation of L and S", we review the implications of Proposition 1 from the viewpoint of sensitivity analysis. Proposition 1 implies that when α is close to 1, the posterior risk amount is determined nearly equally by either risk amount of L or S showing fatter tail. On the other hand, when the thicknesses of the tails is the same (i.e., $\beta = \gamma$,) the posterior risk amount $\mathrm{VaR}_\alpha(L+S)$ is given by the VaR of the random variable U and is influenced by both L and S even if α is close to 1. The random variable U is the variable determined by a fair coin flipping. The risk amount of U is alternated by the toss of coin (head-L and tail-S).

3.3 Main Results

3.3.1 Independent Case

In this section we present a finer estimation of the difference between $\mathrm{VaR}_\alpha(L+S)$ and $\mathrm{VaR}_\alpha(L)$ than Proposition 1 when L and S are independent. The assumption of independence implies the loss events are caused independently by the factors L or S. In this case condition $[A]$ is satisfied. We prepare additional conditions.

[B] There is some $x_0 > 0$ such that F_L has a positive, non-increasing density function f_L on $[x_0, \infty)$, i.e., $F_L(x) = F_L(x_0) + \int_{x_0}^{x} f_L(y)dy$, $x \geq x_0$.

[C] The function $x^{\gamma-\beta}\bar{F}_S(x)/\bar{F}_L(x)$ converges to some real number k as $x \to \infty$.

[D] The same assertion of $[B]$ holds by replacing L with S.

We remark that condition $[B]$ (respectively, $[D]$) and the monotone density theorem (Theorem 1.7.2 in Bingham et al. [4]) imply $f_L \in \mathcal{R}_{-\beta-1}$ (respectively, $f_S \in \mathcal{R}_{-\gamma-1}$).

The condition $[C]$ seems a little strict: this implies that $\mathcal{L}_L$ and (a constant multiple of) $\mathcal{L}_S$ are asymptotically equivalent, where $\mathcal{L}_L(x) = x^\beta \bar{F}_L(x)$ and $\mathcal{L}_S(x) = x^\gamma \bar{F}_S(x)$ are slowly varying functions. However, since the Pickands–Balkema–de Haan theorem (see [1] and [25])

imples that $\bar{F}_L$ and $\bar{F}_S$ are approximated by GPD, the asymptotic equivalence of $\mathcal{L}_L$ and $\mathcal{L}_S$ "approximately" holds.

Our main theorem is the following.

Theorem 1 *The following assertions hold as* $\alpha \to 1$.

(i) *If* $\beta + 1 < \gamma$, *then* $\mathrm{VaR}_\alpha(L+S) - \mathrm{VaR}_\alpha(L) \sim \mathrm{E}[S]$ *under* $[B]$.

(ii) *If* $\beta < \gamma \leq \beta + 1$, *then* $\mathrm{VaR}_\alpha(L+S) - \mathrm{VaR}_\alpha(L) \sim \frac{k}{\beta}\mathrm{VaR}_\alpha(L)^{\beta+1-\gamma}$ *under* $[B]$ *and* $[C]$.

(iii) *If* $\beta = \gamma$, *then* $\mathrm{VaR}_\alpha(L+S) \sim (1+k)^{1/\beta}\mathrm{VaR}_\alpha(L)$ *under* $[C]$.

(iv) *If* $\gamma < \beta \leq \gamma + 1$, *then* $\mathrm{VaR}_\alpha(L+S) - \mathrm{VaR}_\alpha(S) \sim \frac{1}{k\gamma}\mathrm{VaR}_\alpha(S)^{\gamma+1-\beta}$ *under* $[C]$ *and* $[D]$.

(v) *If* $\gamma + 1 < \beta$, *then* $\mathrm{VaR}_\alpha(L+S) - \mathrm{VaR}_\alpha(S) \sim \mathrm{E}[L]$ *under* $[C]$ *and* $[D]$.

The assertions of Theorem 1 are divided into five cases according to the magnitude relationship between β and γ. In particular, when $\beta < \gamma$, we get different results depending on whether γ is greater than $\beta + 1$ or not. The assertion (i) implies that if the tail probability of S is sufficiently thinner than that of L, then the effect of a supplement of S is limited to the expected loss (EL) of S. In fact, we can also get a similar result to the assertion (i), which we introduce in Section 3.6, when the moment of S is very small. These results indicate that if an additional loss amount S is not so large, we may not have to be nervous about the effect of a tail event which is raised by S.

The assertion (ii) implies that when $\gamma \leq \beta + 1$, the difference of a risk amount cannot be approximated by EL even if $\gamma > 1$. Let $l > 0$ and $p \in (0,1)$ be such that $P(S > l) = p$ and l is large enough (or, equivalently, p is small enough) that $\mathrm{VaR}_{1-p}(L) \geq \mathrm{VaR}_{1-p}(S) = l$. Then we can interpret the assertion (ii) formally as

$$\mathrm{VaR}_\alpha(L+S) - \mathrm{VaR}_\alpha(L) \approx \frac{1}{\beta}\left(\frac{l}{\mathrm{VaR}_{1-p}(L)}\right)^\gamma \mathrm{VaR}_\alpha(L) \leq \frac{1}{\beta}\left(\frac{l}{\mathrm{VaR}_{1-p}(L)}\right)^\beta \mathrm{VaR}_\alpha(L). \quad (3.2)$$

Thus it is enough to provide an amount of the right hand side of (3.2) for an additional risk capital. So, in this case, the information of the pair (l,p) (and detailed information about the tail of L) enables us to estimate the difference conservatively.

When the tail of S has the same thickness as that of L, we have the assertion (iii). In this case we see that by a supplement of S, the risk amount is multiplied by $(1+k)^{1/\beta}$. The slower is the decay speed of $\bar{F}_S(x)$, which means the fatter the tail amount variable becomes with an additional loss, the larger is the multiplier $(1+k)^{1/\beta}$. Moreover, if k is small, we have the following approximation,

$$\mathrm{VaR}_\alpha(L+S) - \mathrm{VaR}_\alpha(L) \sim \frac{k + o(k)}{\beta}\mathrm{VaR}_\alpha(L), \quad (3.3)$$

where $o(\cdot)$ is the Landau symbol (little o): $\lim_{k\to 0} o(k)/k = 0$. The relation (3.3) has the same form as assertion (ii), and in this case we have a similar implication as (3.2) by letting $\alpha = 1-p$ and $k = (l/\mathrm{VaR}_{1-p}(L))^\beta$.

The assertions (iv)–(v) are restated consequences of the assertions (i)–(ii). In these cases, $\mathrm{VaR}_\alpha(L)$ is too much smaller than $\mathrm{VaR}_\alpha(L+S)$ and $\mathrm{VaR}_\alpha(S)$, so we need to compare $\mathrm{VaR}_\alpha(L+S)$ with $\mathrm{VaR}_\alpha(S)$. In estimating the posterior risk amount $\mathrm{VaR}_\alpha(L+S)$, the effect of the tail index γ of S is significant. We remark that we can replace $\mathrm{VaR}_\alpha(S)$ with $k^{1/\gamma}\mathrm{VaR}_\alpha(L)^{\beta/\gamma}$ when either $x^\beta F_L(x)$ or $x^\gamma F_S(x)$ converges to some positive number (see [17]).

By Theorem 1, we see that the smaller is the tail index γ, the more precise is the information which we need about the tail of S.

3.3.2 Consideration of Dependency Structure

In the previous section we assumed that L and S were independent, since they were caused by different loss factors. However, huge losses often happen due to multiple simultaneous loss events. Thus it is important to prepare a risk capital considering a dependency structure between loss factors. Basel II states that "scenario analysis should be used to assess the impact of deviations from the correlation assumptions embedded in the bank's operational risk measurement framework, in particular, to evaluate potential losses arising from multiple simultaneous operational risk loss events" in paragraph 675 of Basel Committee on Banking Supervision [2].

In this section we consider the case where L and S are not necessarily independent, and present generalizations of Theorem 1(i)–(ii). Let $L \in \mathcal{R}_{-\beta}$ and $S \in \mathcal{R}_{-\gamma}$ be random variables for some $\beta, \gamma > 0$. We only consider the case of $\beta < \gamma$. By using the fact that $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ is a standard Borel space and Theorem 5.3.19 in Karatzas and Shreve [16], we see that there is a regular conditional probability distribution p (respectively, q) : $[0,\infty) \times \Omega \longrightarrow [0,1]$ with respect to $\sigma(L,S)$ given S (respectively, L). We define the function $F_L(x|S=s)$ by $F_L(x|S=s) = p(s, \{L \le x\})$. We see that the function $F_L(x|S=s)$ satisfies

$$\int_B F_L(x|S=s)F_S(ds) = P(L \le x, S \in B)$$

for each Borel subset $B \subset [0,\infty)$.

We prepare the following conditions.

[E] There is some $x_0 > 0$ such that $F_L(\cdot|S=s)$ has a positive, non-increasing and continuous density function $f_L(\cdot|S=s)$ on $[x_0,\infty)$ for $P(S \in \cdot)$-almost all s.

[F] It holds that

$$\operatorname*{ess\,sup}_{s\ge 0} \sup_{t\in K} \left| \frac{f_L(tx|S=s)}{f_L(x|S=s)} - t^{-\beta-1} \right| \longrightarrow 0, \quad x \to \infty \tag{3.4}$$

for any compact set $K \subset (0,1]$ and

$$\int_{[0,\infty)} s^\eta \frac{f_L(x|S=s)}{f_L(x)} F_S(ds) \le C, \quad x \ge x_0 \tag{3.5}$$

for some constants $C > 0$ and $\eta > \gamma - \beta$, where ess sup is the L^∞-norm under the measure $P(S \in \cdot)$.

We notice that the condition $[E]$ includes the condition $[B]$. Under these conditions we have $P(L > x, S > x) \leq Cx^{-\eta}\bar{F}_L(x)$ and then the condition $[A]$ is also satisfied.

Let $\mathrm{E}[\cdot|L = x]$ be the expectation under the probability measure $q(x, \cdot)$. Under the condition [E], we see that for each $\varphi \in L^1([0,\infty), P(S \in \cdot))$

$$\mathrm{E}[\varphi(S)|L = x] = \int_{[0,\infty)} \varphi(s)\frac{f_L(x|S = s)}{f_L(x)}F_S(ds) \tag{3.6}$$

for $P(L \in \cdot)$-almost all $x \geq x_0$. We do not distinguish the left hand side and the right hand side of (3.6). The left hand side of (3.5) is regarded as $\mathrm{E}[S^\eta|L = x]$.

The conditions [E] and [F] seem to be a little strong, but we can construct a non-trivial example. Please refer to [17] for details.

Now we present the following theorem.

Theorem 2 *Assume* $[E]$ *and* $[F]$. *If* $\beta + 1 < \gamma$, *then*

$$\mathrm{VaR}_\alpha(L + S) - \mathrm{VaR}_\alpha(L) \sim \mathrm{E}[S|L = \mathrm{VaR}_\alpha(L)], \quad \alpha \to 1. \tag{3.7}$$

The relation (3.7) gives a similar indication of (5.12) in Tasche [30]. The right hand side of (3.7) has the same form as the so-called component VaR:

$$\mathrm{E}[S|L + S = \mathrm{VaR}_\alpha(L + S)] = \frac{\partial}{\partial\varepsilon}\mathrm{VaR}_\alpha(L + \varepsilon S)\Big|_{\varepsilon=1} \tag{3.8}$$

under some suitable mathematical assumptions. In Section 3.6 we study the details. We can replace the right hand side of (3.7) with (3.8) by a few modifications of our assumptions:

[E'] The same condition as $[E]$ holds by replacing L with $L + S$.

[F'] The relations (3.4) and (3.5) hold by replacing L with $L + S$ and by setting $K = [a, \infty)$ for any $a > 0$.

Indeed, our proof also works upon replacing $(L + S, L)$ with $(L, L + S)$.

3.4 Numerical Examples

In this section we confirm numerically our main results for typical examples in the standard LDA framework. Let L and S be given by the following compound Poisson variables: $L = L^1 + \cdots + L^N$, $S = S^1 + \cdots + S^{\tilde{N}}$, where $(L^i)_i, (S^i)_i, N, \tilde{N}$ are independent random variables and $(L^i)_i$, $(S^i)_i$ are each identically distributed. The variables N and $\tilde{N}$ mean the frequency of loss events, and the variables $(L^i)_i$ and $(S^i)_i$ mean the severity of each loss event. We assume that $N \sim \mathrm{Poi}(\lambda_L)$ and $\tilde{N} \sim \mathrm{Poi}(\lambda_S)$ for some $\lambda_L, \lambda_S > 0$. For severity, we assume that L^i follows $\mathrm{GPD}(\xi_L, \sigma_L)$ with $\xi_L = 2, \sigma_L = 10,000$ and set $\lambda_L = 10$. We also assume that S^i follows $\mathrm{GPD}(\xi_S, \sigma_S)$ and $\lambda_S = 10$. We set the parameters ξ_S and σ_S in each cases appropriately. We remark that $L \in \mathcal{R}_{-\beta}$ and $S \in \mathcal{R}_{-\gamma}$, where $\beta = 1/\xi_L$ and $\gamma = 1/\xi_S$. Moreover the condition $[C]$ is satisfied with

$$k = \frac{\lambda_S}{\lambda_L}(\sigma_S/\xi_S)^{1/\xi_S}(\sigma_L/\xi_L)^{-1/\xi_L}. \tag{3.9}$$

Here we apply the direct approach introduced in Section 2.3.1 to calculate VaRs numerically. Unless otherwise noted, we set $\alpha = 0.999$. Then the value of the prior risk amount $\mathrm{VaR}_\alpha(L)$ is 5.01×10^{11}.

3.4.1 The Case of $\beta + 1 < \gamma$

First we consider the case of Theorem 1(i). We set $\sigma_S = 10,000$. The result is given in Table 4, where

$$\Delta\mathrm{VaR} = \mathrm{VaR}_\alpha(L+S) - \mathrm{VaR}_\alpha(L), \quad \mathrm{Error} = \frac{\mathrm{Approx}}{\Delta\mathrm{VaR}} - 1 \tag{3.10}$$

and $\mathrm{Approx} = \mathrm{E}[S]$.

Although the absolute value of the error becomes a little large when $\gamma - \beta$ is near 1, the difference between the VaRs is accurately approximated by $\mathrm{E}[S]$.

Table 4: The case of $\beta + 1 < \gamma$.

ξ_S	$\gamma - \beta$	ΔVaR	Approx	Error
0.1	9.500	1,111,092	1,111,111	1.68E-05
0.2	4.500	1,249,995	1,250,000	4.26E-06
0.3	2.833	1,428,553	1,428,571	1.26E-05
0.4	2.000	1,666,647	1,666,667	1.21E-05
0.5	1.500	2,000,141	2,000,000	-7.05E-05

Table 5: The case of $\beta < \gamma \leq \beta + 1$.

ξ_S	$\gamma - \beta$	ΔVaR	Approx	Error
0.8	0.750	3.64E+06	3.14E+06	-1.36E-01
1.0	0.500	2.02E+08	2.00E+08	-8.38E-03
1.2	0.333	3.31E+09	3.30E+09	-1.73E-03
1.5	0.167	5.69E+10	5.63E+10	-9.98E-03
1.8	0.056	4.36E+11	3.81E+11	-1.26E-01

Table 6: The case of $\beta = \gamma$.

σ_S	ΔVaR	Approx	Error
100	1.05E+11	1.05E+11	-2.05E-07
1,000	3.67E+11	3.67E+11	-1.85E-07
10,000	1.50E+12	1.50E+12	-1.43E-07
100,000	8.17E+12	8.17E+12	-8.51E-08
1,000,000	6.01E+13	6.01E+13	-3.46E-08

Table 7: The case of $\beta > \gamma$.

ξ_S	$\mathrm{VaR}_\alpha(L+S)$	Approx	Error	c.f. $\mathrm{VaR}_\alpha(S)$
2.5	2.12E+12	1.52E+12	-2.82E-01	4.00E+11
3.0	4.64E+13	4.56E+13	-1.61E-02	3.34E+13
3.5	2.99E+15	2.99E+15	-3.04E-04	2.86E+15
4.0	2.52E+17	2.52E+17	-5.38E-06	2.50E+17
4.5	2.23E+19	2.23E+19	-2.09E-07	2.22E+19

3.4.2 The Case of $\beta < \gamma \leq \beta + 1$

This case corresponds to Theorem 1(ii). As in Section 3.4.1, we also set $\sigma_S = 10,000$. The result is given as Table 5, where $\mathrm{Approx} = k\mathrm{VaR}_\alpha(L)^{\beta+1-\gamma}/\beta$ and the error is the same as (3.10). We see that the accuracy is lower when $\gamma - \beta$ is close to 1 or 0. Even in these cases, the error approaches 0 by letting $\alpha \to \infty$ (see [17]).

3.4.3 The Case of $\beta = \gamma$

In this section we set $\xi_S = 2 (= \xi_L)$. We apply Theorem 1(iii). We compare the values of ΔVaR and $\mathrm{Approx} = ((1+k)^{\xi_L} - 1)\mathrm{VaR}_\alpha(L)$ in Table 6, where the error is the the same as (3.10). We see that they are very close.

3.4.4 The Case of $\beta > \gamma$

Finally we treat the case of Theorem 1(iv). We set $\sigma_S = 100$. In this case $\mathrm{VaR}_\alpha(L)$ is too much smaller than $\mathrm{VaR}_\alpha(L+S)$, so we compare the values of $\mathrm{VaR}_\alpha(L+S)$ and

$$\mathrm{Approx} = \mathrm{VaR}_\alpha(S) + \frac{1}{k\gamma}\mathrm{VaR}_\alpha(S)^{\gamma+1-\beta}.$$

The result is in Table 7. We see that the error ($= \mathrm{Approx}/\mathrm{VaR}_\alpha(L+S) - 1$) tends to become smaller when ξ_S is large.

Table 7 also indicates that the supplement of S has a quite significant effect on the risk amount when the distribution of S has a fat tail. For example, when $\xi_S = 3.0$, the value of $\mathrm{VaR}_\alpha(L+S)$ is more than 90 times $\mathrm{VaR}_\alpha(L)$ and is heavily influenced by the tail of S. We see that a little change of ξ_S may cause a huge impact on the risk model.

In our example we do not treat the case of Theorem 1(v), however we also have a similar implication in this case.

3.5 Concluding Remarks

We introduced the theoretical framework of sensitivity analysis for quantitative operational risk management. Concretely speaking, we investigated the impact on the risk amount (VaR) arising from adding the loss amount variable S to the present loss amount variable L when the tail probabilites of L and S are regularly varying ($L \in \mathcal{R}_{-\beta}, S \in \mathcal{R}_{-\gamma}$ for some $\beta, \gamma > 0$). The result depends on the magnitude relationship of β and γ. One of these implications is that we must pay more attention to the form of the tail of S when S has the fatter tail. Moreover, when $\gamma > \beta + 1$, the difference between the prior VaR and the posterior VaR is approximated by the component VaR of S (in particular by EL of S if L and S are independent).

We have mainly treated the case where L and S are independent except for a few cases in Section 3.3.2. As related study for dependent case, Böcker and Klüppelberg [5] invokes a Lévy copula to describe the dependency and gives an asymptotic estimate of Fréchet bounds of total VaR. To deepen and enhance our study in more general cases when L and S have a dependency structure is one of the directions of our future work.

3.6 Appendix: The Effect of a Supplement of Small Loss Amount

In this section we treat another version of Theorem 1(i) and Theorem 2(i). We do not assume that the random variables are regularly varying but that the additional loss amount variable is very small. Let $L, \tilde{S}$ be non-negative random variables and let $\varepsilon > 0$. We define a random variable S_ε by $S_\varepsilon = \varepsilon\tilde{S}$. We regard L (respectively, $L + S_\varepsilon$) as the prior (respectively, posterior) loss amount variable and consider the limit of the difference between the prior and posterior VaR by taking $\varepsilon \to 0$. Instead of making assumptions of regular variation, we make "Assumption (S)" in Tasche [30]. Then Lemma 5.3 and Remark 5.4 in Tasche [30] imply

$$\lim_{\varepsilon\to 0}\frac{\mathrm{VaR}_\alpha(L+S_\varepsilon) - \mathrm{VaR}_\alpha(L)}{\varepsilon} = \frac{\partial}{\partial\varepsilon}\mathrm{VaR}_\alpha(L+\varepsilon\tilde{S})\Big|_{\varepsilon=0} = \mathrm{E}[\tilde{S}|L = \mathrm{VaR}_\alpha(L)]. \tag{3.11}$$

By (3.11), we have

$$\mathrm{VaR}_\alpha(L+S) - \mathrm{VaR}_\alpha(L) \;=\; \mathrm{E}[S|L = \mathrm{VaR}_\alpha(L)] + o(\varepsilon), \tag{3.12}$$

where we simply put $S = S_\varepsilon$. In particular, if L and S are independent, then

$$\mathrm{VaR}_\alpha(L+S) - \mathrm{VaR}_\alpha(L) \;=\; \mathrm{E}[S] + o(\varepsilon).$$

Thus the effect of a supplement of the additional loss amount variable S is approximated by its component VaR or EL. So the assertions of Theorem 1(i) and Theorem 2(i) also hold in this case.

The concept of the component VaR is related to the theory of risk capital decomposition (or risk capital allocation). Let us consider the case where L and S are loss amount variables and where the total loss amount variable is given by $T(w_1, w_2) = w_1 L + w_2 S$ with a portfolio $(w_1, w_2) \in \mathbb{R}^2$ such that $w_1 + w_2 = 1$. We try to calculate the risk contributions for the total risk capital $\rho(T(w_1, w_2))$, where ρ is a risk measure.

One of the ideas is to apply Euler's relation

$$\rho(T(w_1, w_2)) = w_1 \frac{\partial}{\partial w_1}\rho(T(w_1, w_2)) + w_2 \frac{\partial}{\partial w_2}\rho(T(w_1, w_2))$$

when ρ is linear homogeneous and $\rho(T(w_1, w_2))$ is differentiable with respect to w_1 and w_2. In particular we have

$$\rho(L+S) = \frac{\partial}{\partial u}\rho(uL+S)\Big|_{u=1} + \frac{\partial}{\partial u}\rho(L+uS)\Big|_{u=1} \tag{3.13}$$

and the second term in the right hand side of (3.13) is regarded as the risk contribution of S. As in early studies in the case of $\rho = \mathrm{VaR}_\alpha$, the same decomposition as (3.13) is obtained in Garman [12] and Hallerbach [13] and the risk contribution of S is called the component VaR. The consistency of the decomposition of (3.13) has been studied from several points of views (Denault [8], Kalkbrener [15], Tasche [30], and so on). In particular, Theorem 4.4 in Tasche [30] implies that the decomposition of (3.13) is "suitable for performance measurement" (Definition 4.2 of Tasche [30]). Although many studies assume that ρ is a coherent risk measure, the result of Tasche [30] also applies to the case of $\rho = \mathrm{VaR}_\alpha$.

Another approach towards calculating the risk contribution of S is to estimate the difference of the risk amounts $\rho(L+S) - \rho(L)$, which is called the marginal risk capital—see Merton and Perold [21]. (When $\rho = \mathrm{VaR}_\alpha$, it is called a marginal VaR.) This is intuitively intelligible, whereas an aggregation of marginal risk capitals is not equal to the total risk amount $\rho(L+S)$.

The relation (3.12) gives the equivalence between the marginal VaR and the component VaR when $S(= \varepsilon\tilde{S})$ is very small. Theorem 2(i) implies that the marginal VaR and the component VaR are also (asymptotically) equivalent when L and S have regulary varying tails and the tail of S is sufficiently thinner than that of L.

References

[1] A. Balkema and L. de Haan, Residual Life Time at Great Age, Ann. Prob. 2(1974), pp. 792–804.

[2] Basel Committee on Banking Supervision, International Convergence of Capital Measurement and Capital Standards: A Revised Framework, Bank of International Settlements, Available from http://www.bis.org/, 2004.

[3] F. Biagini and S. Ulmer, Asymptotics for Operational Risk Quantified with Expected Shortfall, Astin Bull., 39(2008), pp. 735–752.

[4] N. H. Bingham, C. M. Goldie and J. L. Teugels, Regular Variation, Cambridge University Press, 1987.

[5] C. Böcker and C. Klüppelberg, Operational VaR: A Closed-Form Approximation, Risk, 18-12(2005), pp. 90–93.

[6] C. Böcker and C. Klüppelberg, Multivariate Models for Operational Risk, Quant. Finance, 10-8(2010), pp. 855–869.

[7] M. Degen, P. Embrechts and D. D. Lambrigger, The Quantitative Modeling of Operational Risk: Between g-and-h and EVT, Astin Bull., 37-2(2006), pp. 265–291.

[8] M. Denault, Coherent Allocation of Risk Capital, Journal of Risk, 4–1(2001), pp. 1–34.

[9] K. Dutta and J. Perry, A Tale of Tails: An Empirical Analysis of Loss Distribution Models for Estimating Operational Risk Capital, Working papers of the Federal Reserve Bank of Boston, Available from http://www.bos.frb.org/, 2006. No. 06-13, 2006.

[10] P. Embrechts and M.Frei, Panjer Recursion versus FFT for Compound Distributions, Mathematical Methods of Operations Research, 69–3(2009), pp. 497–508.

[11] P. Embrechts, C. Klüppelberg and T. Mikosch, Modelling Extremal Events, Springer, Berlin, 2003.

[12] M. Garman, Taking VaR to Pieces, Risk, 10-10(1997), pp. 70–71.

[13] W. Hallerbach, Decomposing Portfolio Value-at-Risk : A General Analysis, Journal of Risk, 5-2(2003), pp. 1–18.

[14] J. Jang and J. H. Jho, Asymptotic Super(sub)additivity of Value-at-Risk of Regularly Varying Dependent Variables, Preprint, MacQuarie University, Sydney, 2007.

[15] M. Kalkbrener, An Axiomatic Approach to Capital Allocation, Mathematical Finance, 15-3(2005), pp. 425–437.

[16] I. Karatzas and S. E. Shreve, Brownian Motion and Stochastic Calculus 2nd. edition, Springer, New York, 1991.

[17] T. Kato, Theoretical Sensitivity Analysis for Quantitative Operational Risk Management, Intellectual Archive, 1–2(2012), pp. 99–115.

[18] X. Luo and P. V. Shevchenko, Computing Tails of Compound Distributions Using Direct Numerical Integration, The Journal of Computational Finance, 13–2(2009), pp. 73–111.

[19] A. J. McNeil, R. Frey and P. Embrechts, Quantitative Risk Management : Concepts, Techniques and Tools, Princeton University Press, 2005.

[20] K. Ishitani and K. Sato, Wavelet 変換を用いたオペレーショナルリスクの解析的評価方法, 第 36 回ジャフィー大会予稿集 (2012).

[21] R. C. Merton and A. F. Perold, Theory of Risk Capital in Financial Firms, Journal of Applied Corporate Finance, 5-1(1993), pp. 16–32.

[22] M. Moscadelli, The Modelling of Operational Risk: Experience with the Analysis of the Data Collected by the Basel Committee, Technical report of Banking Supervision Department, Banca d'Italia, 517, Available from http://www.bancaditalia.it/, 2004.

[23] T. Ooura and M. Mori, The Double Exponential Formula for Oscillatory Functions Over the Half Infinite Interval, J.Comput. Appl. Math. 38(1991), pp. 353–360.

[24] T. Ooura and M. Mori, A Robust Double Exponential Formula for Fourier Type Integrals, J.Comput. Appl. Math. 112(1999), pp. 229–241.

[25] J. Pickands, Statistical Inference Using Extreme Order Statistics, Ann. Stat. 3(1975), pp. 119–131.

[26] M. Pritsker, Evaluating Value at Risk Methodologies: Accuracy versus Computational Time, Journal of Financial Services Research, 12(1997), pp. 201–242.

[27] P. V. Shevchenko, Calculation of Aggregate Loss Distributions, The Journal of Operational Risk, 5–2(2010), pp. 3–40.

[28] H. Takahashi and M. Mori, Quadrature Formulas Obtained by Variable Transformation, Numer. Math., 21(1973), pp. 206–219.

[29] H. Takahashi and M. Mori, Double Exponential Formulas for Numerical Integration, Publ. RIMS Kyoto Univ. 9(1974), pp. 721–741.

[30] D. Tasche, Risk Contributions and Performance Measurement, Working Paper, Center for Mathematical Sciences, Munich University of Technology, Available from http://www-m4.ma.tum.de/, 2000.

[31] Wynn P., On a Device for Computing the $e_m(s_n)$ Tranformation, Mathematical Tables and Other Aids to Computation, 10(1956), pp. 91–96.

Division of Mathematical Science for Social Systems, Graduate School of Engineering Science
Osaka University, Toyonaka 560-8531, Japan
E-mail address: kato@sigmath.es.osaka-u.ac.jp

大阪大学・大学院基礎工学研究科　加藤 恭