Extremal behavior and risk indicators for a class of regularly varying stochastic processes

Charles Tillier [1] & Olivier Wintenberger [2]

[1] *Université Paris Nanterre, MODAL'X. 200 avenue de la république, 92001, Nanterre. Email: c.tillier@univ-paris10.fr*

[2] *Université Pierre et Marie Curie, LSTA. 4 place Jussieu, 75005 Paris. Email: olivier.wintenberger@upmc.fr*

Résumé. Dans le cadre de l'évaluation des risques à horizon de temps fini, le problème peut souvent se ramener à l'étude d'un vecteur $C(N) = (C_1, \dots, C_N)$, de taille aléatoire N, dont les composantes sont des variables aléatoires issues du produit d'une matrice $A(N)$, de dimension aléatoire $N \times N$ par un vecteur $X(N)$ de taille aléatoire N. Dans ce contexte, l'objectif est alors de définir la propriété de variation régulière du vecteur $C(N)$, d'étudier son comportement extrémal, notamment par l'intermédiaire de sa queue de distribution et dans le même temps, de développer des indicateurs de risque. Dans de nombreuses applications telles que l'assurance, la finance ou encore en hydrologie, les indicateurs de risque peuvent être analysés à partir du comportement extrémal de $\|C(N)\|$, pour une norme $\| \cdot \|$ convenablement choisie. On propose dans ce travail de généraliser le lemme de Breiman, qui permet d'obtenir un équivalent de la queue de distribution de $\|C(N)\|$, qui, selon la norme et la matrice considérées, donne lieu à une estimation asymptotique d'indicateurs de risque tels que la probabilité de ruine à horizon fini ou le temps passé au dessus d'un seuil pour de tels processus. Afin de mettre en avant notre méthode, on applique notre résultat principal à un processus "Shot Noise".

Mots-clés. Théorie de la ruine, variation régulière multivariée, indicateurs de risque, Lemme de Breiman, processus stochastiques, valeurs extrêmes, événements rares.

Abstract. When assessing risks on a finite-time horizon, the problem can often be reduced to the study of a random sequence $C(N) = (C_1, \dots, C_N)$ of random length N, where $C(N)$ comes from the product of a matrix $A(N)$ of random size $N \times N$ and a random sequence $X(N)$ of random length N. Our aim is to build a regular variation framework for such random sequences of random length, to study their spectral properties and, subsequently, to develop risk measures. In several applications, many risk indicators can be expressed from the extremal behavior of $\|C(N)\|$, for some norm $\| \cdot \|$. We propose a generalization of Breiman's Lemma that gives way to a tail estimate of $\|C(N)\|$. To illustrate the applicability of our method, we apply our main result for a class of stochastic processes - the Shot Noise Processes - to provide risk indicators such as the tail index and the ruin probability on a finite-time horizon.

Keywords. Ruin theory, multivariate regular variation, risk indicators, Breiman's Lemma, asymptotic properties, stochastic processes, extremes, rare events.

1 Motivation

Risk analyses play a leading role within many fields such as dietary risk, hydrology, nuclear security, finance and insurance and is more and more present in the applications of various probability tools and statistical methods; see [3], [1], [4] or [5]. Besides, risk theory typically deals with the probability of occurrence of rare events which are functions of heavy-tailed random variables, for example, sums or products of regularly varying random variables; see [7] and [10]. For instance, non-life insurance mathematics and dietary risk management both deal with a particular kind of *Shot Noise Processes* (*SNP*) $\{S(t)\}_{t\geq 0}$, defined as

$$S(t) = \sum_{i=1}^{N(t)} X_i h_i(t, T_i), \qquad t \geq 0, \tag{1}$$

where $(h_i)_{i\geq 1}$ are independent and identically distributed (*i.i.d.*) non-negative measurable random functions called "*shock functions*", which are independent of the shocks $(X_i)_{i\geq 1}$. The "shock arrivals" $(T_i)_{i\geq 1}$ are random variables (*r.v.*'s) independent of $(X_i)_{i\geq 1}$ such that for all $i \geq 1$, $T_i = \sum_{1\leq k\leq i} \Delta T_k$, where $(\Delta T_k)_{k\geq 1}$ is a sequence of positive *r.v.*'s called "inter-arrivals" and $N(t) = \#\{i : T_i \leq t\}$. In this setup, we define the ruin probability on a finite-time horizon of such *SNP* as the probability that the supremum of S exceeds a threshold on a time window $[0, T]$, for a given $T > 0$, which necessarily happens on the embedded chain - the process $\{S(t)\}_{t\geq 0}$ sampled at the shock arrivals $T_1, T_2 \ldots$ - and it is enough to study the discrete-time random sequence $S(N(T)) := (S(T_1), S(T_2), \ldots, S(T_{N(T)}))'$, which is of random length $N(T)$. Then, instead of dealing with the extremal behavior of $\{S(t)\}_{t\leq T}$, we only need to understand the extremal behavior of $\|S(N(T))\|_\infty$. We go further and point out that many risk measures in non-life insurance mathematics and in dietary risk assessment can be analyzed from the tail behavior of $\|C(N)\|$ where $C(N) = (C_1, \ldots, C_N)'$ is a random sequence of random length N and $\|\cdot\|$ is a norm such that

$$\|\cdot\|_\infty \leq \|\cdot\| \leq \|\cdot\|_1. \tag{2}$$

Thus we consider discrete-time processes $C(N) = (C_1, \ldots, C_N)$ where for all $1 \leq i \leq N$, $C_i \in \mathbb{R}^+$ and N is an integer-valued *r.v.*, independent of the C_i's. We are interested in the case where the C_i's are regularly varying *r.v.*'s and we restrict ourselves to the process $C(N)$ which can be written in the form

$$C(N) = A(N)X(N), \tag{3}$$

where $X(N) = (X_1, \ldots, X_N)'$ is a random length sequence with identically distributed marginals, which are not necessarily independent and $A(N)$ is a random matrix of random size $N \times N$ independent of the X_i's. However, $X(N)$ and $A(N)$ are still dependent through N, which determines their dimensions. Our main objectives are to define regular variation properties for a random length sequence of random variables and to study its spectral properties in order to develop risk measures. As it will become clear later, the randomness of the size N of the vector $C(N)$ makes it difficult to use the common definition of multivariate regular variation in terms of vague convergence; see [6]. We tackle the problem using the notion of $\mathbb{M}$-convergence introduced recently in [9] and the use of a norm satisfying (2) that allows to build regular variation via polar coordinates. A main difference with the finite-dimensional case is that the choice of the norm matters as it determines the infinite-dimensional space to consider; see [2], [12] and [13].

2 Framework and assumptions

We require the following conditions. For $\alpha > 0$ and $\epsilon > 0$, we assume

(H0) Length: N is a positive integer-valued *r.v.* such that $\mathbb{E}[N] > 0$ and $\mathbb{E}[N^{1+\alpha+\epsilon}]$.

(H1) Regular variation: $(X_i)_{i\geq 1}$ are identically distributed *r.v*'s with mean γ, cumulative distribution function (*c.d.f.*) F_X, integrated tail distribution $\overline{F_X^I}$, such that the survival function $\overline{F}_X = 1 - F_X$ is regularly varying with index $\alpha > 0$, denoted $X \in RV_{-\alpha}$.

(H2) Uniform asymptotic independence: For $i, j \geq 1$, we assume

$$\sup_{i\neq j}\left|\frac{\mathbb{P}(X_i > x, X_j > x)}{\mathbb{P}(X_1 > x)}\right| \underset{x\to\infty}{\longrightarrow} 0.$$

(H3) Regularity of the norm: The norm $\|\cdot\|$ satisfies $\|\cdot\|_\infty \leq \|\cdot\| \leq \|\cdot\|_1$.

(H4) Tail condition on the matrix $A(N)$**:** The random entries $(a_{i,j})_{1\leq i,j\leq N}$ of $A(N)$ are independent of the $(X_i)_{1\leq i\leq N}$. Moreover, we assume that $\mathbb{E}[\|A(N)\|^{\alpha+\epsilon}N^{1+\alpha+\epsilon}] < \infty$, where $\|\cdot\|$ also denotes the corresponding induced norm on the space of N-by-N matrices.

(H5) The matrix $A(N)$ **is not null**.

We first need to define regular variation for any $X(N) = (X_1, X_2, \ldots, X_N)'$ when N is an integer-valued random variable. We work on the space $c_{\|\cdot\|}$, defined below.

Definition 1. *The space* $c_{\|\cdot\|}$ *is the completion of* c_{00} *in* $(\mathbb{R}^+)^{\mathbb{N}}$ *equipped with the convergence in the sequential norm* $\|\cdot\|$.

In the following, we denote by $\{\mathbf{0}\} = (0, 0, \ldots)$ the null element in $c_{\|\cdot\|}$, $e_j = (0, \ldots, 0, 1, 0, \ldots) \in c_{\|\cdot\|}$ the canonical basis of $c_{\|\cdot\|}$ and $S(\infty)$, the unit sphere over $c_{\|\cdot\|}$ defined as $S(\infty) = \{X \in c_{\|\cdot\|} : \|X\| = 1\}$. As $c_{\|\cdot\|}$ is a Banach space, the notion of weak convergence holds on $c_{\|\cdot\|}$ and one can also define regular variation as in [6].

Proposition 2. *A sequence of random elements* $X(N) = (X_1, X_2, \ldots, X_N)' \in c_{\|\cdot\|}\backslash\{\mathbf{0}\}$ *for* N *satisfying* ***(H0)*** *is regularly varying if the random variable* $\|X(N)\|$ *is regularly varying and*

$$\mathcal{L}\left(\|X(N)\|^{-1}X(N) \mid \|X(N)\| > x\right) \underset{x\to\infty}{\longrightarrow} \mathcal{L}(\Theta(N)),$$

for some random element $\Theta(N) \in S(\infty)$*. The distribution of* $\Theta(N)$ *is the spectral measure of* $X(N)$*.*

It means that the regular variation of $X(N)$ is completely characterized by the tail index α of $\|X(N)\|$ and the spectral measure of X. Notice that $\Theta(N) \in S(\infty)$ is an infinite-dimensional sequence $\Theta(N) = (\Theta_1(N), \Theta_2(N), \ldots)$. The following proposition is relevant to prove Theorem 6. It is a first example of such regularly varying random sequence of random length under **(H3)**. Besides, it is an extension of Lemma A6 in [14].

Proposition 3. *Let $X(N) = (X_1, \dots, X_N) \in c_{\|\cdot\|}$ such that **(H0)-(H3)** hold. Then we have*

$$\lim_{x\to\infty} \frac{\mathbb{P}(\|X(N)\| > x)}{\mathbb{P}(X_1 > x)} = \mathbb{E}[N] > 0. \tag{4}$$

We conclude this section by characterizing the spectral measure of $X(N)$. Note that when $N = n$ for a fixed $n \geq 1$, it follows that for all $1 \leq j \leq n$, $\mathbb{P}(\Theta(n) = e_j) = n^{-1}$.

Proposition 4. *If **(H0)-(H3)** hold then $X(N) = (X_1, \dots, X_N) \in c_{\|\cdot\|}\backslash\{\mathbf{0}\}$ is regularly varying in the sense of Proposition 2 and its spectral measure is characterized by*

$$\mathbb{P}(\Theta(N) = e_j) = \frac{\mathbb{P}(N \geq j)}{\mathbb{E}[N]}, \qquad j \geq 1.$$

3 Main results

We generalize the previous approach to sequences in $c_{\|\cdot\|}$ defined by the matrix product (3).

Theorem 5. *Let $C(N)$ be a random length sequence defined as in* (3)*. Assume that **(H0)-(H4)** hold. Then, we have*

$$\lim_{x\to\infty} \frac{\mathbb{P}(\|C(N)\| > x)}{\mathbb{P}(X_1 > x)} = \mathbb{E}\left[\sum_{k=1}^{N} \|A_k(N)\|^\alpha\right]. \tag{5}$$

Notice that Theorem 5 holds if $A(N)$ does not necessarily satisfy **(H5)** and then we allow that $\mathbb{P}(\|C(N)\| > x)/\mathbb{P}(X_1 > x) \to 0$ when x goes to infinity. Under the additional assumption **(H5)**, we are now ready to prove that $C(N)$ is regularly varying.

Theorem 6. *Let $C(N)$ be a random length sequence defined as in* (3)*. If **(H0)-(H5)** hold, then $C(N)$ is regularly varying and its spectral measure is given by*

$$\mathbb{P}\left(\|C(N)\|^{-1}C(N) \in \cdot \mid \|C(N)\| > x\right) \underset{x\to\infty}{\longrightarrow} \frac{\mathbb{E}\left[\sum_{k=1}^{N} \|A_k(N)\|^\alpha \mathbb{1}_{\|A_k(N)\|^{-1}A_k(N)\in\cdot}\right]}{\mathbb{E}\left[\sum_{k=1}^{N} \|A_k(N)\|^\alpha\right]}. \tag{6}$$

Although the characterization is common for any norm such that **(H3)** holds, the result essentially depends on the choice of the norm. Despite this remark, it is noteworthy that the spectral measure can be described in a unified way even if it belongs to different spaces, regarding the choice of the norm.

4 Applications

We finally provide quite a few examples to see the applicability of our method. We assume in this section that $\alpha \geq 1$, for all $i \geq 1$, $t > 0$, there exists $\epsilon > 0$ such that $\mathbb{E}[h_i^{\alpha+\epsilon}(t, T_i)] < \infty$ and $N := N(t)$ is an Poisson process with intensity function (resp. cumulative intensity) $\lambda(t)$ (resp. $m(t)$). The computation of the constant $\mathbb{E}[\sum_{k=1}^{N} \|A_k(N)\|^\alpha]$ for different norms and different matrices $A(N)$ allows to develop various risk measures. Altough $C(N)$ defined as in (3) covers a wide family of processes, we only deal here with a particular class of *Shot Noise Processes* defined as in (1). We first apply our main result to determine their extremal behavior as a corollary of Theorem 5.

Corollary 7. *Under* **(H1)**-**(H2)**, *assume that the random functions $h_j(T,\cdot)$'s are i.i.d., independent of the T_j's and integrable of order $p > \alpha$, then we have*

$$\lim_{x\to\infty} \frac{\mathbb{P}(S(T) > x)}{\mathbb{P}(X_1 > x)} = m(T)\,\mathbb{E}[h_1^\alpha(T, V_1)],$$

where V_1 admits the density $\lambda(t)/m(T)$, $0 \le t \le T$.

This corollary plays a leading role to determine the following risk indicators. Besides, we recover the recent results of [14] and [8] on the tail behavior of $\{S(T)\}_{T\geq 0}$. We now deal with finite-time ruin probability ψ of (1), which we define as the probability that $S(t)$ exceeds some given threshold $x \in \mathbb{R}^+$ on a period $[0, T]$, *i.e.* $\psi(x, T) = \mathbb{P}(\sup_{0\le t\le T} S(t) > x)$.

Corollary 8. *Assume that the conditions of Corollary 7 hold. If $h_j(\cdot, T)$ is a non-increasing function for any $T > 0$, then,*

$$\lim_{x\to\infty} \frac{\psi(x, T)}{\mathbb{P}(X_1 > x)} = m(T)\mathbb{E}[h_1^\alpha(V_1, V_1)].$$

Notice that if $h_j(\cdot, T)$ is a non-decreasing function for any $T > 0$, then the maximum of the *SNP* is achieved at time T and the ruin probability can be computed thanks to Corollary 7. We finally propose two indicators to supplement the information given by the ruin probability and the tail behavior. The ruin probability permits to know if the process has exceeded the threshold but provides no information about the exceedances themselves or about the duration of the exceedances. To fill the gap, we first bear our interest on the Integrated Expected Severity (IES(x)), which deals with the average of the cumulated exceedances when the process is over the high threshold x on a time window $[0, T]$. It is defined for all $T > 0$ by

$$IES(x) := \int_0^T ES(x)dt, \quad x > 0.$$

and then we focus on the "Expected Time Over a threshold" (*ETOT(x)*), which provides information about how long the process stays, in average, above a threshold x, knowing that it has already reached it. It is defined for all $T > 0$ and $x > 0$ by

$$ETOT(x) := \mathbb{E}\left[\int_0^T \mathbb{1}_{\{S(t)\in]x,\infty[\}} dt \mid \max_{0\le t\le T} S(t) > x\right].$$

Proposition 9. *Assume that the conditions of Corollary 7 hold. The IES is given by*

$$IES(x) \underset{x\to\infty}{\sim} \gamma \int_0^T \mathbb{E}[N(t)]\mathbb{E}\left[h_1^\alpha(t, V_1)\right] dt\, \overline{F_X^I}(x),$$

and the ETOT(x) is given by

$$ETOT(x) \underset{x\to\infty}{\sim} \frac{\int_0^T m(t)\mathbb{E}[h_1^\alpha(t, V_1)]dt}{m(T)\mathbb{E}[h_1^\alpha(V_1, V_1)]}.$$

To conclude, in many configurations, we can explicitly derive $\mathbb{E}\Big[\sum_{k=1}^N \|A_k(N)\|^\alpha\Big]$, especially with respect to $\|\cdot\|_1$ and $\|\cdot\|_\infty$ which provide interesting equivalents to obtain risk indicators. We used it to get equivalents of the tail behavior, the ruin probability, the ETOT and the IES but our result can also be applied to many other risk measures and to other stochastic processes regarding the choice of the norm and the matrix $A(N)$.

References

[1] Asmussen, S. (2003). *Applied Probabilities and Queues.* Springer-Verlag, New York.

[2] Basrak, B. and Segers, J. (2009). Regularly varying multivariate time series. *Stochastic Processes and their Applications.* Vol **119**, Issue 4, 1055–1080.

[3] Bertail, P., Clemencon, S. and Tressou, J. (2010). Statistical analysis of a dynamic model for food contaminant exposure with applications to dietary methylmercury contamination. *Journal of Biological Dynamics.* Vol. **4**, 212–234.

[4] Embrechts, P., Kluppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance.* Applications of Mathematics, Springer-Verlag.

[5] Grandell, J. (1991). *Aspects of risk theory.* Springer Series in Statistics: Probability and its Applications. Springer-Verlag, New York.

[6] Hult, H. and Lidskog, F. (2006). Regular variation for measures on metric spaces. *Publications de l'institut mathématique, Nouvelle série.* Vol **80**, No. 94, 121–140.

[7] Jessen, A. H. and Mikosch, T. (2006). Regularly varying functions. *Publications de l'institut mathématique.* Nouvelle serie, tome **80**(94), 171–192 DOI:10.2298/PIM0694171H.

[8] Li, X. and Wu, J. (2014). Asymptotic tail behavior of Poisson shot-noise processes with interdependence between shock and arrival time. *Statistics and Probability Letters.* Vol. **88**, 15–26.

[9] Lindskog, F., Resnick, S. I. and Roy, J. (2014). Regularly varying measures on metric spaces: hidden regular variation and hidden jumps. *Probability Survey.*

[10] Mikosch, T. (1999). Regular Variation, Subexponentiality and Their Applications in Probability Theory. *Eindhoven University of Technology* . Vol. **99**, No. 13 of EURANDOM report, ISSN 1389-2355.

[11] Mikosch, T. (2010). *Non life Insurance Mathematics*. Second edition. Springer-Verlag, Berlin, Heidelberg.

[12] Resnick, S.I. (2002). Hidden Regular Variation, Second Order Regular Variation and Asymptotic Independence. *Extremes*. Vol. **5**, Issue 4, 303–336.

[13] Resnick, S.I. (2004). On the Foundations of Multivariate Heavy-Tail Analysis. *Journal of applied probability*. Vol. **41**, 191–212.

[14] Weng, C., Zhang, Y. and Tang, K. S. (2013). Tail behavior of Poisson Shot Noise Processes under heavy-tailed shocks and Actuarial applications. *Methodology and Computing in Applied Probability*. Vol. **15**, Issue 3, 665–682.