Article

Not peer-reviewed version

Extreme Behavior of Competing Risks with Random Sample Size

Long Bai , Kaihao Hu , Conghua Wen , Zhongquan Tan , Chengxiu Ling *

Posted Date: 9 July 2024

doi: 10.20944/preprints202407.0661.v1

Keywords: extreme value theory; competing risks; random sample size

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 July 2024 doi:10.20944/preprints202407.0661.v1

Article

Extreme Behavior of Competing Risks with Random Sample Size

Long Bai [1,†], Kaihao Hu [23,†], Conghua Wen[1], Zhongquan Tan[4] and Chengxiu Ling [25,*]

1 Department of Financial and Actuarial Mathematics, School of Mathematics and Physics, Xi'an Jiaotong-Liverpool University, Suzhou, 215123
2 Academy of Pharmacy, Xi'an Jiaotong-Liverpool University, Suzhou, 215123, China
3 Department of Mathematical Sciences, University of Liverpool, Liverpool, L69 3BX, UK
4 College of Data Science, Jiaxing University, 314001, China
5 The Key Lab of Jiangsu Higher Education Institutions (under Construction), XJTLU, Suzhou, 215123, China
* Correspondence: chengxiu.ling@xjtlu.edu.cn; Tel.: +86-512-8188-9023
† These authors contributed equally to this work.

Abstract: This paper examines the impact of random sample sizes on the extreme value theory of competing risks, a significant area in finance and environmental science. We capture limit distributions of two extreme types under random sampling sizes, known as accelerated mixed l-max and p-max stable type distributions. The study presents results for both maxima and minima in competing risks scenarios, addressing cases of independent and non-independent random sample sizes. Numerical examples validate our theoretical findings, demonstrating the applicability of our approach to various random sample size distributions, including time-shifted Poisson or binomial, geometric, and negative binomial distributions.

Keywords: extreme value theory; competing risks; random sample size

---

1. Introduction

Extreme Value Theory (EVT) is dedicated to modeling extreme events within a sequence of a large number of independent and identically distributed (i.i.d.) random variables. Its applications are diverse, spanning fields such as finance, insurance, environmental science, and engineering [? ?]. Let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. random variables with common distribution function (d.f.) F, and denote by $M_n = \max(X_1, X_2, \ldots, X_n)$ the sample maxima. The risk $X \sim F$ is called to be in the max-domain attraction of G, if there exist some normalization constants $a_n > 0, b_n \in \mathbb{R}$ and a non-degenerate d.f. G such that (with $\xrightarrow{d}$ convergence in distribution)

$$\mathbb{P}\left(a_n\left(M_n - b_n\right) \le x\right) \xrightarrow{d} G\left(x\right) \quad \text{as } n \to \infty. \tag{1.1}$$

The limit distribution G is the so-called generalized extreme value distribution (GEV), which is of the sample l-type (namely, x can be replaced with $ax + b$ for some $a > 0, b \in \mathbb{R}$) as

$$G(x; \gamma, \mu, \sigma) = \exp\left(-\left(1 + \gamma\frac{x-\mu}{\sigma}\right)_+^{-1/\gamma}\right). \tag{1.2}$$

We denote this by $F \in D_l(G)$. Here the three parameters $\gamma, \mu \in \mathbb{R}, \sigma > 0$ are called the shape, location, and scale parameters, respectively. In addition, the tail behavior of the potential risk X is well classified into Fréchet, Weibull, and Gumbel domains, corresponding to the cases with $\gamma >, =, < 0$, respectively [?].

Given the wide applications of EVT, many extensive studies of limit theory alike Eq.(**??**) have been conducted. ?] extended first the limit distribution under linear normalization in Eq.(**??**) to

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 July 2024 doi:10.20944/preprints202407.0661.v1

the power limit laws $D_p(H)$, i.e., there exist some power normalization constants $\alpha_n, \beta_n > 0$ and a non-degenerate d.f. H such that

$$\mathbb{P}\left\{\alpha_n |M_n|^{\beta_n} \operatorname{sign}(M_n) \le x\right\} \xrightarrow{d} H(x) \tag{1.3}$$

with the sign function sign(x) equal 1, -1 and 0 for x being positive, negative and zero, respectively. It is well-known that H is of p-max stable distributions composed of six types of limits, which can be rewritten uniformly as below [?]. For some $\mu, \sigma > 0$ and $\gamma \in \mathbb{R}$ (recall G is the GEV defined in Eq.(??)),

$$H(x;\gamma,\mu,\sigma) = \begin{cases} G(\log x;\gamma,\mu,\sigma), & \text{if the support is included in } (0,\infty), \\ G(-\log(-x);\gamma,-\mu,\sigma), & \text{otherwise.} \end{cases} \tag{1.4}$$

In what follows, we denote this by $F \in D_p(H)$.

Recently, ?] and ?] investigated the limit behavior of extremes under linear and power normalization in the scenario of competing risks, with the practical consideration of aggregating multiple sources. Namely, the studied sample maxima M_n is actually obtained from k heterogeneous subsamples $X_{j,i}, i = 1,\ldots,n_j$ from source/population $X_j \sim F_j, j = 1,\ldots,k$. This considerate modeling in the big data era is desirable due to the complexity of real applications [? ?]. The limit theory of M_n obtained for the k multiple sources is the so-called limit theory of max of max since

$$M_n = \max_{1\le j\le k} M_{j,n_j} \quad \text{with } M_{j,n_j} = \max_{1\le i\le n_j} X_{j,i}. \tag{1.5}$$

Clearly, the obtained limit laws of Eq.(??) extending the classical extreme value theory given in Eqs.(??) and (??) are the so-called accelerated l-max stable and accelerated p-max stable distributions, see ? , Theorem 2.1] and ? , Theorem 2.1]. Note that the key condition in determining accelerated limit theory is the interplay of the sample length and the tail behavior among the multiple competing risks. A natural question is how the extreme law varies in the uncertainty of the sample size involved. This is very common in environmental and financial fields, for instance, the extreme claim size of ν_n claims over a n-day period and the extreme daily precipitation within a ν_n duration of wet period [? ?]. This paper aims to study the limit theory under both linear and power normalization in the framework of competing risks with random sample size.

Many authors refined the extreme limit theory under linear normalization with random sample size for two different cases:

Case I) with independent random sample size. The basic risks $X_1,\ldots,X_n$ and sample size index ν_n are supposed to be independent and ν_n/n converges weakly to a non-degenerate distribution function [?];

Case II) with non-independent random sample size. There exists a positive-valued variable V such that ν_n/n converges to V in probability, allowing the interrelation of the basic risk and sample size index ν_n [?].

The limit theorems with random sample size were further extended for sample minima [?], extreme order statistics under power normalization [? ?], stationary Gaussian process [?], stationary chi-process [?], and recent contributions on multivariate extreme behavior [?]. This paper will further consider the limit behavior of extremes (both minima and maxima) under linear and power normalization in the competing risk scenario, extending those accelerated l-max and p-max stable limit distributions when the sample size sequence $\{\nu_n, n \ge 1\}$ satisfies conditions indicated in Cases I) and II). The theoretical results will be illustrated by numerical studies with typical examples such as ν_n are time-shifted Poisson, (negative) Binomial distributions, which have extensive applications in insurance and hydrology [? ?].

The remainder of the paper is organized as follows. Section ?? presents the main results for maxima of maxima under both linear and power normalization with sample sizes. Extensional

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 July 2024 doi:10.20944/preprints202407.0661.v1

results for competing minima and typical examples are discussed in Section ??. Numerical studies are conducted to illustrate our theoretical findings in Section ??. The proofs of all theoretical results are deferred to the Appendix.

2. Main Results

Notation. Recall that the max of max defined in Eq.(**??**), is generated from k independent samples of size n_j's from risk $X_j \sim F_j$. Let $\nu_{n_j}, 1 \le j \le k$ be mutually independent, positive integer-valued variables, standing for the random sample size, which is independent of the basic risks $X_j \sim F_j$. Similar to Eq.(**??**), we write

$$M_{\nu_n} = \max_{1\le j\le k} M_{j,\nu_{n_j}} \quad \text{with} \quad M_{j,\nu_{n_j}} = \max_{1\le i\le \nu_{n_j}} X_{j,i}. \tag{2.1}$$

Here $\nu_n := \sum_{1\le j\le k} \nu_{n_j}$ and $n = \sum_{j=1}^k n_j$. Throughout this paper, for any risk X following a cumulative distribution function (cdf) F, we write $\underline{F}(x) := 1 - F(-x-0)$, standing for the cdf of $-X$. Further, all limits are taken as $\min_{1\le j\le k} n_j \to \infty$.
To simplify the notation, in what follows, we consider competing risks from two sources, namely with $k = 2$. We will present below the limit behavior of M_{ν_n} for Cases I) and II) in Section **??** and **??**, respectively.

2.1. Limit theorem for Case I) with independent sample size

In this section, we present our main results on the limit behavior of competing risks under linear and power normalization in Theorems **??** and **??**, respectively. Basically, we focus on the following random sample size scenario: Assume that there exist independent non-degenerate distributed V_j, $1 \le j \le k$ such that

$$\frac{\nu_{n_j}}{n_j} \xrightarrow{d} V_j. \tag{2.2}$$

Condition (**??**) is commonly used for the limit behavior of extremes with random sample size [**? ? ?**]. We refer to **?**] for relevant examples, see also Examples **??** to **??**.

Limit behavior of M_{ν_n} under linear normalization. Clearly, for $F_j \in D_l(G_j)$, there exist $a_{j,n_j} > 0, b_{j,n_j} \in \mathbb{R}$ such that M_{j,n_j} satisfies Eq.(**??**) as $n_j \to \infty$. We will show in Theorem **??** below that, under condition (**??**), the limit theorem for competing extremes M_{ν_n} holds for an accelerated mixed GEV distribution with

$$L_j(x) = \int_0^\infty (G_j(x))^z d\mathbb{P}\left\{V_j \le z\right\}. \tag{2.3}$$

Theorem 1. *Let M_{ν_n} be given by Eq.(**??**) with the basic risks $X_j \sim F_j, j = 1,2$ and random sample sizes ν_1, ν_2 mutually independent. Suppose that condition (**??**) and $F_j \in D_l(G_j)$ holds with Eq.(**??**) for M_{j,n_j} and $a_{j,n_j}, b_{j,n_j}, j = 1,2$. If there exist two constants $a \in [0,\infty]$ and $b \in \mathbb{R}$ such that*

$$a_n := \frac{a_{1,n_1}}{a_{2,n_2}} \to a, \quad b_n := a_{1,n_1}(b_{2,n_2} - b_{1,n_1}) \to b \tag{2.4}$$

as $\min(n_1, n_2) \to \infty$.

*(i). If Eq.(**??**) holds for $a > 0$ and $b < \infty$, then*

$$\mathbb{P}\left(a_{2,n_2}(M_{\nu_n} - b_{2,n_2}) \le x\right) \xrightarrow{d} L_1(ax+b)L_2(x).$$

*(ii). If Eq.(**??**) holds for $a = 0$ and $b = \infty$, then*

$$\mathbb{P}\left(a_{2,n_2}(M_{\nu_n} - b_{2,n_2}) \le x\right) \xrightarrow{d} L_2(x).$$

*Here $L_j, j = 1,2$ are given by Eq.(**??**).*

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 July 2024 doi:10.20944/preprints202407.0661.v1

Remark 1. *a) Theorem* **??** *is reduced to Theorem 2.1 by* **?** *] if V_j's are degenerate at one, the limit theorem for competing maxima with determinant sample size.*
b) In addition, the two results in (i) and (ii) correspond to the cases for two competing risks being comparable tails and balanced sampling process and the dominated case, respectively.
c) Theorem **??** *extends Theorem 6.2.2 in* **?** *] for a non-competing risk scenario, where the extremes are from one single source. In general, the accelerated mixed H distributions family is a larger class including those of form in Eq.(**??**).*

Limit behavior of M_{ν_n} under power normalization. Clearly, for $F_j \in D_p(H_j)$, there exist $\alpha_{j,n_j}, \beta_{j,n_j} > 0$ such that M_{j,n_j} satisfies Eq.(**??**) as $n_j \to \infty$. We will show in Theorem **??** below that, under condition (**??**), the limit theorem for competing extremes M_{ν_n} holds for an accelerated mixed H distribution given below

$$P_j(x) = \int_0^\infty (H_j(x))^z d\mathbb{P}\left\{V_j \le z\right\}. \tag{2.5}$$

Theorem 2. *Let M_{ν_n} be given by Eq.(**??**) with the basic risks $X_j \sim F_j, j = 1,2$ and random sample sizes ν_1, ν_2 mutually independent. Suppose that condition (**??**) and $F_j \in D_p(H_j)$ holds with Eq.(**??**) for M_{j,n_j} and $\alpha_{j,n_j}, \beta_{j,n_j} > 0, j = 1,2$. If there exist two non-negative constants α and β such that*

$$\alpha_n := \alpha_{1,n_1}\left(\frac{1}{\alpha_{2,n_2}}\right)^{\frac{\beta_{1,n_1}}{\beta_{2,n_2}}} \to \alpha, \quad \beta_n := \frac{\beta_{1,n_1}}{\beta_{2,n_2}} \to \beta. \tag{2.6}$$

*as $\min(n_1, n_2) \to \infty$. The following claims hold for $P_j, j = 1,2$, the mixed H_j distributions defined in Eq.(**??**).*

*(i). If condition (**??**) holds with two positive constants α and β, then*

$$\mathbb{P}\left\{\alpha_{2,n_2}\left|M_{\nu_n}\right|^{\beta_{2,n_2}} \operatorname{sign}\left(M_{\nu_n}\right) \le x\right\} \xrightarrow{d} P_1\left(\alpha|x|^\beta \operatorname{sign}(x)\right)P_2(x). \tag{2.7}$$

(ii). The following limit distribution holds

$$\mathbb{P}\left\{\alpha_{2,n_2}\left|M_{\nu_n}\right|^{\beta_{2,n_2}} \operatorname{sign}\left(M_{\nu_n}\right) \le x\right\} \xrightarrow{d} P_2(x)$$

provided that one of the following four conditions is satisfied (notation: $x_1^* := \inf\{x : H_1(x) < 1\}$, the right endpoint of H_1)

a). When H_2 is one of the same p-types of $G(\log x; \gamma, \mu, \sigma)$, and H_1 is one of the same p-types of $G(-\log(-x); \gamma, -\mu, \sigma)$.
*b). When H_2 is one of the same p-types of $G(\log x; \gamma, \mu, \sigma)$, and H_1 is one of the same p-types of $G(\log x; \gamma, \mu, \sigma)$ for $\gamma \ge 0$. In addition, Eq.(**??**) holds with $\alpha = \infty$ and $0 \le \beta < \infty$.*
*c). When H_2 is one of the same p-types of $G(\log x; \gamma, \mu, \sigma)$, and H_1 is the same type of $G(\log x; \gamma, \mu, \sigma)$ for $\gamma < 0$. In addition, Eq.(**??**) holds with $x_1^* \le \alpha < \infty$ and $\beta = 0$ or $\alpha = \infty$ and $0 \le \beta < \infty$.*
*d). When both H_1 and H_2 are one of the same p-types of $G(-\log(-x); \gamma, -\mu, \sigma)$. In addition, Eq.(**??**) holds with $0 \le \alpha \le -x_1^*$ and $\beta = 0$ or $\alpha = 0$ and $0 \le \beta < \infty$.*

Remark 2. *a) Theorem* **??** *is reduced to Theorem 2.1 by* **?** *] if V_j's are degenerate at one, which means the asymptotically almost randomlessness of ν_n, the limit theorem for competing maxima with determinant sample size. This situation happens in practice, e.g., ν_n follows a shifted Poisson df with mean λ_n such that $\lambda_n \sim n$. For more examples, see Example* **??** *below and Remark 2.2 by* **?** *].*
*b) In addition, the two results in i) and ii) correspond to the two different cases with $\alpha\beta > 0$ and $\alpha\beta = 0, \infty$ in condition (**??**), illustrating the limit behavior of two competing risks with comparable tails and balanced sampling process and the dominated case, respectively.*
c) Theorem **??** *extends Theorem 2.1 by* **?** *] for a non-competing risk scenario, where the extremes are from one*

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 July 2024 doi:10.20944/preprints202407.0661.v1

*single source. In general, the accelerated mixed-GEV distributions family is a larger class including those of form in Eq.(**??**).*

2.2. Limit theorem for Case II) with non-independent sample size

In this section, we focus on Case II), relaxing the independent condition between the basic risk and random sample size. On the other hand, we need to strengthen the convergence in distribution as the convergence in probability, as stated below. Assume that there exist positive random variables V_j, $j = 1, 2$ such that (notation: $\stackrel{p}{\to}$ stands for convergence in probability)

$$\nu_{n_j}/n_j \stackrel{p}{\to} V_j, \quad j = 1, 2. \tag{2.8}$$

Theorem 3. *Let M_{ν_n} be given by Eq.(**??**) from two independent pairs of basic risk and sample size $(X_j, \nu_j), j = 1, 2$. Suppose that conditions (**??**) and (**??**) hold for $X_j \sim F_j \in D_l(G_j)$ with Eq.(**??**) satisfied for M_{j,n_j} and $a_{j,n_j}, b_{j,n_j}, j = 1, 2$. Then the claim in Theorem **??** holds.*

Theorem 4. *Let M_{ν_n} be given by Eq.(**??**) from two independent pairs of basic risk and sample size $(X_j, \nu_j), j = 1, 2$. Suppose that conditions (**??**) and (**??**) hold for $X_j \sim F_j \in D_p(H_j)$ with Eq.(**??**) satisfied for M_{j,n_j} and $\alpha_{j,n_j}, \beta_{j,n_j}, j = 1, 2$. Then the claim in Theorem **??** holds.*

3. Discussion

In this section, we first extend our results for competing minima risks in Section **??**, and then present typical examples of random sizes with specific mixed extreme distributions in Section **??**.

3.1. Extreme Limit Theory for Competing Minima Risks

In some practical applications, such as the lifetime in reliability analysis or race time of athletes in physical studies, extreme minima plays an important role. As we will see in Corollaries **??** and **??** below, analytical claims follow for competing risks with random sample size in terms of minima of minima. Essentially, noting that the right tail behavior of $X_j \sim F_j$ is demonstrated by its sample maxima M_{j,n_j}, the left tail behavior of $-X_j \sim \underline{F}_j(x) := 1 - F_j(-x - 0)$ can be shown by the sample minima $\underline{m}_{j,n_j} = \min(-X_{j,1}, \ldots, -X_{j,n_j})$ since (cf. see **?** , Theorem 1.8.3] and **?**])

$$\underline{m}_n = \min_{1\le j\le k} \underline{m}_{j,n_j} = -\max_{1\le j\le k} M_{j,n_j} = -M_n. \tag{3.1}$$

Noting that, the condition that $F_j \in D_l(G_j)$, i.e., there exist $a_{j,n_j} > 0, b_{j,n_j} \in \mathbb{R}$ such that M_{j,n_j} satisfies Eq.(**??**) as $n_j \to \infty$, is equivalent that

$$\mathbb{P}\left(a_{j,n_j}(\underline{m}_{j,n_j} + b_{j,n_j}) \le x\right) \stackrel{d}{\to} \underline{G}_j(x), \tag{3.2}$$

where G_j is of the same l-type of GEV distribution given in Eq.(**??**).

Corollary 1. *Suppose the same conditions as for Theorems **??** or **??** are satisfied.*

*(i). If Eq.(**??**) holds for $a > 0$ and $b < \infty$, then*

$$\mathbb{P}\left(a_{2,n_2}(\underline{m}_{\nu_n} + b_{2,n_2}) \le x\right) \stackrel{d}{\to} 1 - L_1\big(-(ax - b)\big)L_2(-x).$$

*(ii). If Eq.(**??**) holds for $a = 0$ and $b = \infty$, then*

$$\mathbb{P}\left(a_{2,n_2}(\underline{m}_{\nu_n} + b_{2,n_2}) \le x\right) \stackrel{d}{\to} \underline{L}_2(x).$$

Here $L_j, j = 1, 2$ are given by Eq.(??).

Noting that $\alpha_n|\underline{m}_n|^{\beta_n}\text{sign}(\underline{m}_n) = -\alpha_n|M_n|^{\beta_n}\text{sign}(M_n)$, the following corollary holds for the power normalized minima of minima.

Corollary 2. *Suppose the same conditions as for Theorems ?? or ?? are satisfied. The following claims hold for $\underline{P}_j(x), j = 1, 2$ with P_j the mixed H_j distributions defined in Eq.(??).*

(i). If condition (??) holds with two positive constants α and β, then

$$\mathbb{P}\left\{\alpha_{2,n_2}\left|\underline{m}_{\nu_n}\right|^{\beta_{2,n_2}}\text{sign}\left(\underline{m}_{\nu_n}\right) \le x\right\} \xrightarrow{d} 1 - P_1\big(-\alpha|x|^\beta\text{sign}(x)\big)P_2(-x).$$

(ii). The following limit distribution holds

$$\mathbb{P}\left\{\alpha_{2,n_2}\left|\underline{m}_{\nu_n}\right|^{\beta_{2,n_2}}\text{sign}\left(\underline{m}_{\nu_n}\right) \le x\right\} \xrightarrow{d} \underline{P}_2(x)$$

provided that one of the conditions a)∼ d) in Theorem ?? holds.

Remark 3. *a) Recalling that G and H given by Eqs.(??) and (??) are the so-called l-max stable and p-max stable, we call L and P the mixed accelerated l-max stable and the mixed accelerated p-max stable distributions if they can be written as a product of L_j and P_j, respectively.*
b) Recalling that the $\underline{G}$ and $\underline{H}$ are the so-called l-min stable and p-min stable distributions [? , Corollary 1] if G and H are given by Eqs.(??) and (??), we call $\underline{L}$ and $\underline{P}$ the mixed accelerated l-min stable and the mixed accelerated p-min stable distributions if they can be written as a product of $\underline{L}_j$ and $\underline{P}_j$, respectively.

3.2. Examples

Below, we will give three examples to illustrate our main results obtained in Theorems ?? and ??. In particular, we considered that random sample size follows respectively time-shifted version of Poisson or binomial distribution, geometric and negative binomial distributions with relevant parameters satisfying certain average stable conditions [?], see Examples ?? ∼ ??.

Example 1 (Time-shifted binomial/Poisson distributed random sample size)**.** *Let ν_n follow a time-shifted binomial distribution with probability mass function (pmf) given as*

$$\mathbb{P}\{\nu_n = k+m\} = \binom{l_n}{k} p_n^k q_n^{l_n-k}, \quad k+m = m, m+1, \ldots, l_n+m.$$

If $l_n p_n \to 1$, then ν_n/n converges in probability to one. Similarly, for a time-shifted Poisson distributed $\nu_n \stackrel{d}{=} m + \text{Poisson}(\lambda_n)$ with $\lambda_n/n \to 1$, then ν_n/n converges in probability to 1 [? , Lemmas 4.3]. For the random sample size aforementioned, the claims of Theorems ?? and ?? follow as the reduced determinant random size cases, see Remarks ??(a) and ??(a).

Example 2 (Time-shifted geometric distributed sample size)**.** *In the case of linear normalization, with G being one of the three l-types distribution, say $G(x;\gamma,\mu,\sigma)$ specified in Eq.(??). Suppose that the random sample size ν_{n_j} follows a geometric distribution with mean n_j. We have $\nu_{n_j} \approx n_j V$ in distribution with a random scale V following a standard exponential distribution. Consequently, Theorem ?? holds with an accelerated mixed l-max stable distribution L, the product of mixed l-max stable distributions of form L as below.*

$$\begin{aligned} L(x;\gamma,\mu,\sigma) &= \int_0^\infty (G(x;\gamma,\mu,\sigma))^z d(1-\mathrm{e}^{-z}) \\ &= \int_0^\infty \exp\left(-z\left(\left(1+\gamma\frac{x-\mu}{\sigma}\right)_+^{-1/\gamma}+1\right)\right)dz = \frac{1}{1+\left(1+\gamma\frac{x-\mu}{\sigma}\right)_+^{-1/\gamma}}, \end{aligned}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 July 2024 doi:10.20944/preprints202407.0661.v1

which is taken as its limit $1/[1+\exp\left(-(x-\mu)/\sigma\right), x>\mu$ *for* $\gamma=0$.
*Similarly, for the power normalization case, recalling H is specified in Eq.(**??**), as the p-max type of limit distributions, Theorem **??** follows with an accelerated mixed p-max stable distribution P, the product of mixed p-max stable distribution P as below (cf.* **?** *, Example 2.1]).*

$$\begin{aligned}
P(x;\gamma,\mu,\sigma) &= \int_0^\infty (H(x;\gamma,\mu,\sigma))^z d(1-\mathrm{e}^{-z}) \\
&= \begin{cases} \int_0^\infty G^z(\log x;\gamma,\mu,\sigma)\mathrm{e}^{-z}dz, & \textit{if the support is included in } (0,\infty), \\ \int_0^\infty G^z(-\log(-x);\gamma,-\mu,\sigma)\mathrm{e}^{-z}dz, & \textit{otherwise} \end{cases} \\
&= \begin{cases} \dfrac{1}{1+\left(1+\gamma\dfrac{\log x-\mu}{\sigma}\right)_+^{-1/\gamma}}, & \textit{if the support is included in } (0,\infty), \\ \dfrac{1}{1+\left(1+\gamma\dfrac{-\log(-x)+\mu}{\sigma}\right)_+^{-1/\gamma}}, & \textit{otherwise.} \end{cases}
\end{aligned}$$

Example 3 (Time-shifted negative binomial distributed sample size)**.** *As an extension of m-shifted geometric distributions, we consider time-shifted negative binomial distributed sample size* ν_n *with* $r\geq 1$ *given by*

$$\mathbb{P}\left\{\nu_n=k\right\}=\binom{-r}{k-rm}p_n^r[-(1-p_n)]^{k-rm},\quad k=rm,rm+1,\ldots.$$

It follows by Lemma 4.1 by **?** *] that, as* $np_n\to 1$*, we have* ν_n/n *converges in distribution to V, a gamma random variable with shape parameter r and scale parameter 1, i.e., the cdf of V is given by*

$$F_V(z)=\mathbb{P}\left\{V\leq z\right\}=\int_0^z\frac{1}{\Gamma(r)}t^{r-1}\exp(-t)dt,\quad z>0,$$

where $\Gamma(\cdot)$ *denotes the gamma function. It follows by Theorem* **??** *that*

$$\begin{aligned}
L(x;\gamma,\mu,\sigma) &= \int_0^\infty (G(x;\gamma,\mu,\sigma))^z dF_V(z) \\
&= \int_0^\infty \frac{1}{\Gamma(r)}z^{r-1}\exp\left(-z\left(\left(1+\gamma\frac{x-\mu}{\sigma}\right)_+^{-1/\gamma}+1\right)\right)dz \\
&= \left[1+\left(1+\gamma\frac{x-\mu}{\sigma}\right)_+^{-1/\gamma}\right]^{-r}.
\end{aligned}$$

Similarly, Theorem **??** *follows with the accelerated mixed p-max stable distributions, the product of form P given below.*

$$P(x;\gamma,\mu,\sigma)=\begin{cases}\left[1+\left(1+\gamma\frac{\log x-\mu}{\sigma}\right)_+^{-1/\gamma}\right]^{-r}, & \textit{if the support is included in } (0,\infty), \\ \left[1+\left(1+\gamma\frac{-\log(-x)+\mu}{\sigma}\right)_+^{-1/\gamma}\right]^{-r}, & \textit{otherwise.}\end{cases}$$

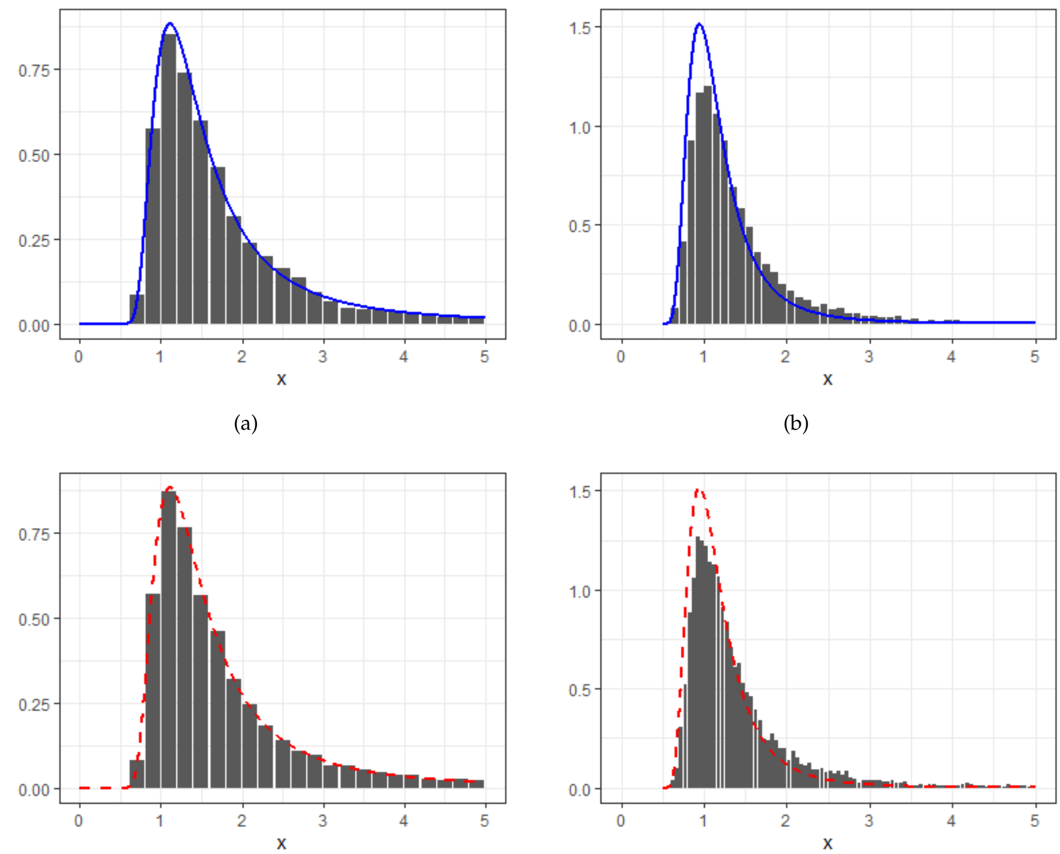
4. Numerical Studies

We will conduct a Monte Carlo simulation to illustrate Theorems **??** and **??** with m-shifted random sample size given in Examples **??** $\sim$ **??**. In what follows, we take the shift parameter $m=5$ in all time-shifted random sample size distributions, and the basic risks X_1, X_2 from Pareto distributions with

 doi:10.20944/preprints202407.0661.v1

parameters $\alpha_1, \alpha_2 > 0$[1] and the random sample sizes ν_{n_1}, ν_{n_2} are supposed to be mutually independent. In addition, the repeated time is taken as $R = 10,000$. We will illustrate our main results specified in Theorems **??** with the three examples given in Section **??** above.

1. Comparison of Pareto competing extremes with determinant sample size and Poisson distributed random sample size. In Figure **??**, we will demonstrate that the competing extremes with



(a) (b) (c) (d)

Figure 1. Distribution approximation of linear normalized $M_{\nu_n} = \max(M_{1,\nu_{n_1}}, M_{2,\nu_{n_2}})$ (a, b) and $M_n = \max(M_{1,n_1}, M_{2,n_2})$ (c,d) with both M_{j,n_j}'s from Pareto(α_j) and Poisson distributed sample size ν_{n_j} with mean n_j. Here $(\alpha_1, \alpha_2) = (2, 4)$ and $n_1 = 100, n_2 = n_1^c$ with $c = 2, 2.5$ in (b,d) and (a,c) by $\Phi(x;\alpha_1) * \Phi(x;\alpha_2)$ and $\Phi(x;\alpha_2)$, respectively.

Poisson distributed sample size are similar to the case with nonrandom sample size case. Let ν_{n_j} follow m-shifted Poisson with mean parameters $1/n_j, j = 1, 2$. We then generate competing Pareto extremes with basic risks following Pareto(α_j) with $\alpha_j > 0, j = 1, 2$. It follows from Theorems **??**, **??** and Example **??** together with Example 4.6 by **?**] that (recall $\Phi(x;\alpha) = \exp\left(-x^{-\alpha}\right), x > 0$ the Fréchet distribution)

1. For $n_2 = n_1$ or $n_2 = n_1^c$ with $c > \alpha_2/\alpha_1$, we have

$$n_2^{-1/\alpha_2} M_{\nu_n} \xrightarrow{d} \Phi(x;\alpha_2), \quad n_2^{-1} M_{\nu_n}^{\alpha_2} \xrightarrow{d} \Phi(x;1).$$

2. For $n_2 = n_1^c$ with $c = \alpha_2/\alpha_1$, we have

$$n_2^{-1/\alpha_2} M_{\nu_n} \xrightarrow{d} \Phi(x;\alpha_1)\Phi(x;\alpha_2), \quad n_2^{-1} M_{\nu_n}^{\alpha_2} \xrightarrow{d} \Phi(x;\alpha_1/\alpha_2)\Phi(x;1).$$

---

[1] The cdf of Pareto(α) is given as $\mathbb{P}\{X \le x\} = 1 - x^{-\alpha}, x > 1$.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 July 2024 doi:10.20944/preprints202407.0661.v1

Noting that the power normalized extremes will behave similarly to the linear normalized ones up to a power transformation. We show only the behavior of linear normalization for the numerical studies below.

In Figure **??**, we take $\alpha_1 = 2, \alpha_2 = 4$ and $n_1 = 100, n_2 = n_1^c$ with $c = 2, 2.5$ to show the above two cases. Overall, the competing Pareto extremes are well fitted by the accelerated GEV distribution for the non-randomized sample size, where the latter is slightly better than the randomized sample size cases. Further, the accelerated GEV approximation (Figure **??** (a, c)) is relatively closer to the empirical competing extremes than the dominated case.

2. Comparison of Pareto competing extremes with Geometric distributed and negative Binomial distributed sample size. We consider the max of maxima M_{ν_n} with both basic risks $X_j \sim Pareto(\alpha_j)$ with random sample size ν_{n_j} following m-shifted negative Binomial distribution with probability $1/n_j$ and $r \geq 1$. It follows by Example 4.6 by **?**], Theorem **??**(a, b) and Example **??** that, with $\widetilde{L}(x;\alpha) = [1 + x^{-\alpha}]^{-r}$

1. For $n_2 = n_1$ or $n_2 = n_1^c$ with $c > \alpha_2/\alpha_1$, we have $n_2^{-1/\alpha_2} M_{\nu_n} \xrightarrow{d} \widetilde{L}(x;\alpha_2)$.
2. For $n_2 = n_1^c$ with $c = \alpha_2/\alpha_1$, we have $n_2^{-1/\alpha_2} M_{\nu_n} \xrightarrow{d} \widetilde{L}(x;\alpha_1)\widetilde{L}(x;\alpha_2)$.

Thus, its density function is given by

$$l(x;\alpha_1,\alpha_2) = \begin{cases} \frac{d\widetilde{L}(x;\alpha_2)}{dx} = \frac{r\alpha_2}{x^{\alpha_2+1}(1+x^{-\alpha_2})^{r+1}}, \\ \frac{d\widetilde{L}(x;\alpha_1)\widetilde{L}(x;\alpha_2)}{dx} = \frac{r\alpha_1 x^{-\alpha_1-1}}{(1+x^{-\alpha_1})^{r+1}(1+x^{-\alpha_2})^r} + \frac{r\alpha_2 x^{-\alpha_2-1}}{(1+x^{-\alpha_2})^{r+1}(1+x^{-\alpha_1})^r}, \end{cases} \tag{4.1}$$

In Figure **??**, we set $n_1 = 100, n_2 = n_1^c$ with $c = 2, 2.5$ in (a, c) and (b, d), respectively. Meanwhile, the random sample size follows a 5-shifted negative Binomial distribution with $r = 1$ in (a, b) (namely geometric distribution), and $r = 2$ in (c, d), and successful probability $1/n_j, j = 1, 2$. Meanwhile, we take $\alpha_1 = 2, \alpha_2 = 4$ in the Pareto basic risks. Consequently, the sub-maxima are completely competing when $n_2 = n_1^2$, resulting in the accelerated mixed extreme limit distributions as shown in Figure **??** (a, c). In contrast, the dominated limit behavior is given in Figure **??** (b, d) as $n_2 = n_1^{2.5}$.
In general, our theoretical density curve given by Eq.(**??**) approximates the histogram very well (Figure **??**). Further, we see that the approximation with geometric distributed random size is slightly better than the negative binomial case. In addition, the approximation for the dominated case (Figure **??** (d)) is slightly better than the accelerated case when negative Binomial random size applies.

 doi:10.20944/preprints202407.0661.v1

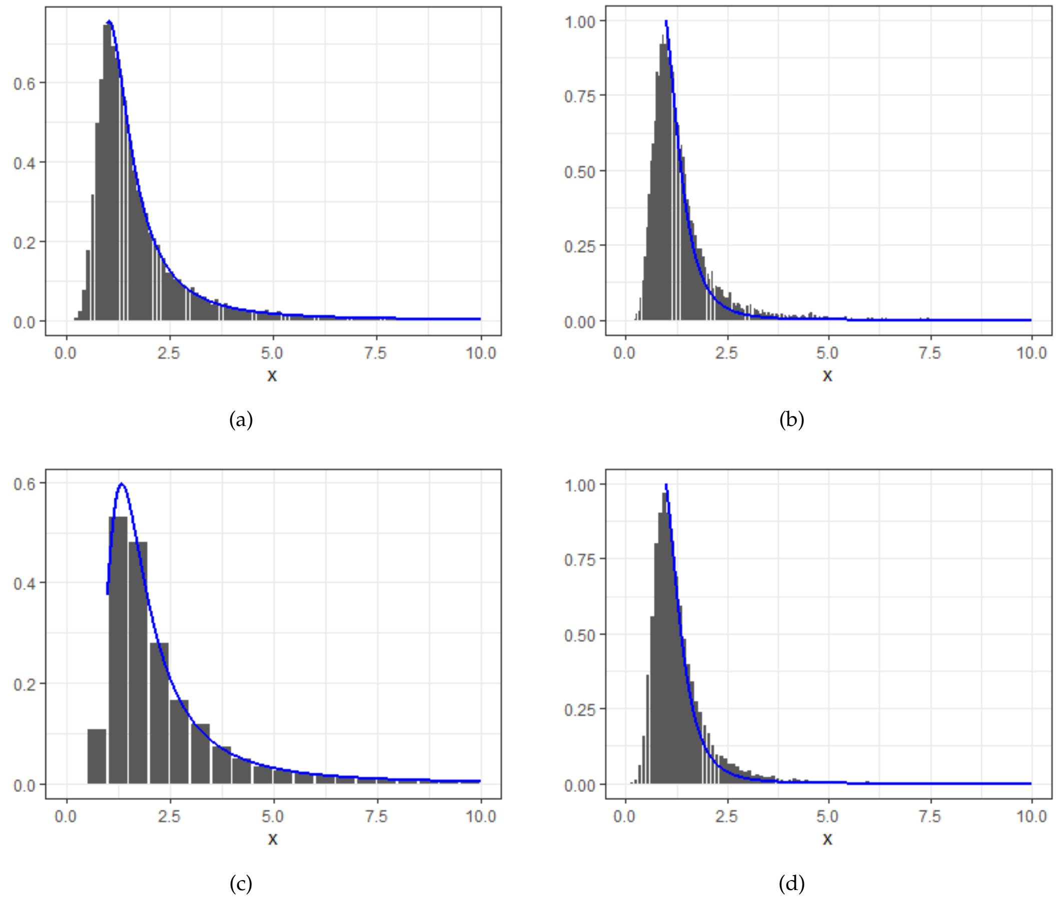


Figure 2. Distribution approximation of linear normalized $M_{\nu_n} = \max(M_{1,\nu_{n_1}}, M_{2,\nu_{n_2}})$ with both M_{j,n_j}'s from Pareto(α_j). The random sample size follows a negative Binomial distribution with $r = 1$ (the geometric distribution) (a,b) and $r = 2$ (c, d) and successful probability $1/n_j$. Here $(\alpha_1, \alpha_2) = (2, 4)$ and $n_1 = 100, n_2 = n_1^c$ with $c = 2, 2.5$ in (a, c) and (b, d) with pdf curves of $\widetilde{L}(x;\alpha_1)\widetilde{L}(x;\alpha_2)$ and $\widetilde{L}(x;\alpha_2)$, respectively.

Author Contributions: Conceptualization, L.B., K.H. and C.L.; Methodology, K.H. and C.L.; software, L.B.; validation, L.B., Z.T. and C.L.; formal analysis, L.B., K.H. and C.L.; investigation, K.H.; Writing, original draft preparation, C.L. and K.H.; Writing, review and editing, C.W., Z.T. and C.L.; Visualization, C.L.; Supervision, C.L.; Project administration, C.L.; Funding acquisition, C.L. All authors have read and agreed to the published version of the manuscript.

Funding: Long Bai is supported by National Natural Science Foundation of China Grant no. 11901469, Natural Science Foundation of the Jiangsu Higher Education Institutions of China grant no. 19KJB11002 and University Research Development Fund no. RDF-21-02-071. Chengxiu Ling is supported by the Research Development Fund [RDF1912017], and the Post-graduate Research Fund [PGRS2112022] at Xi'an Jiaotong-Liverpool University.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We thank the editors and all reviewers for their constructive suggestions and comments that greatly helped to improve the quality of this paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A Proofs of Theorems ?? ∼ ??

Proof of Theorem ??. It follows by the independence between the basic risks $X_j \sim F_j$ and the random sample size ν_j, and condition (**??**) holding for $M_{j,n_j}, a_{j,n_j}, b_{j,n_j}$ that

$$\mathbb{P}\left(a_{j,n_j}(M_{j,\nu_{n_j}} - b_{j,n_j}) \le x\right) \xrightarrow{d} L_j(x) = \int_0^\infty (G_j(x))^z d\mathbb{P}\left\{V_j \le z\right\}. \tag{A1}$$

Further, it follows by the mutual independence between (X_1, ν_1) and (X_2, ν_2) that

$$\begin{aligned} & \mathbb{P}\left\{a_{2,n_2}(M_{\nu_n} - b_{2,n_2}) \le x\right\} = \mathbb{P}\left\{\max(a_{2,n_2}(M_{1,\nu_{n_1}} - b_{2,n_2}), a_{2,n_2}(M_{2,\nu_{n_2}} - b_{2,n_2}))\right\} \\ = & \ \mathbb{P}\left\{a_{2,n_2}(M_{1,\nu_{n_1}} - b_{2,n_2}) \le x\right\}\mathbb{P}\left\{a_{2,n_2}(M_{2,\nu_{n_2}} - b_{2,n_2}) \le x\right\} \\ =: & \ I_n \cdot II_n. \end{aligned} \tag{A2}$$

The straightforward application of Eq.(**??**) gives

$$II_n \xrightarrow{d} L_2(x), \quad n_2 \to \infty. \tag{A3}$$

Similarly, we rewrite I_n as

$$\begin{aligned} I_n &= \mathbb{P}\left\{a_{2,n_2}(M_{1,\nu_{n_1}} - b_{2,n_2}) \le x\right\} \\ &= \mathbb{P}\left\{a_{1,n_1}(M_{1,\nu_{n_1}} - b_{1,n_1}) \le a_{1,n_1}\left(\frac{x}{a_{2,n_2}} + b_{2,n_2} - b_{1,n_1}\right)\right\}. \end{aligned}$$

The remaining proof of $I_n \xrightarrow{d} L_1(ax+b)$ follows from those for Theorem 2.1 in **?**] and Eq.(**??**).

We complete the proof of Theorem **??**.

Proof of Theorem ??. Firstly, we rewrite the left-hand side of Eq.(**??**) as follows.

$$\begin{aligned} & \mathbb{P}\left\{\alpha_{2,n_2}|M_{\nu_n}|^{\beta_{2,n_2}}\operatorname{sign}(M_{\nu_n}) \le x\right\} = \mathbb{P}\left\{M_{\nu_n} \le \left|\frac{x}{\alpha_{2,n_2}}\right|^{1/\beta_{2,n_2}}\operatorname{sign}(x)\right\} \\ = & \ \mathbb{P}\left\{\alpha_{2,n_2}\left|M_{1,\nu_{n_1}}\right|^{\beta_{2,n_2}}\operatorname{sign}(M_{1,\nu_{n_1}}) \le x\right\}\mathbb{P}\left\{\alpha_{2,n_2}\left|M_{2,\nu_{n_2}}\right|^{\beta_{2,n_2}}\operatorname{sign}(M_{2,\nu_{n_2}}) \le x\right\} \\ =: & \ I_n \cdot II_n. \end{aligned} \tag{A4}$$

It follows by Theorem 2.1 in **?**] that

$$II_n \xrightarrow{d} P^{(2)}(x), \quad n_2 \to \infty. \tag{A5}$$

Next, we show the limit of I_n. We rewrite I_n as

$$I_n = \mathbb{P}\left\{M_{1,\nu_{n_1}} \le \left(\frac{x_n}{\alpha_{1,n_1}}\right)^{1/\beta_{1,n_1}}\operatorname{sign}(x)\right\},$$

where $x_n = \alpha_{1,n_1}\left(\frac{|x|}{\alpha_{2,n_2}}\right)^{\beta_{1,n_1}/\beta_{2,n_2}} =: \alpha_n|x|^{\beta_n}$ with α_n, β_n given by Eq.(**??**). The remaining proof is similar to those for Theorem 2.1 by **?**]. We complete the proof of Theorem **??**.

Proof of Theorem ?? It follows by Theorem 6.2.1 of **?**] that, for the jth sample maxima $M_{j,\nu_{n_j}}$, when the constant sequences $a_{j,n_j} > 0, b_{j,n_j} \in \mathbb{R}$ such that (**??**) holds, and ν_{n_j} satisfies Eq.(**??**), we have the claim in Eq.(**??**). The remaining proofs follow by those for Theorem 2.1 by **?**].

 doi:10.20944/preprints202407.0661.v1

Proof of Theorem ?? We show first that, the claim follows for the jth sample maxima $M_{j,\nu_{n_j}}$ with the constant sequences $\alpha_{j,n_j}, \beta_{j,n_j} > 0$, i.e.,

$$\mathbb{P}\left\{\alpha_{j,n_j}\left|M_{\nu_{n_j}}\right|^{\beta_{j,n_j}}\operatorname{sign}\left(M_{\nu_{n_j}}\right) \le x\right\} \xrightarrow{d} P_j(x). \tag{A6}$$

Denote by $\{p_{n_j}(k),\, k \ge 0\}$ the probability mass function of ν_{n_j}. We have

$$p_{n_j}(k) \ge 0, \quad \sum_{k=0}^{\infty} p_{n_j}(k) = 1.$$

It follows by the total law of probability and the independence between basic risks and sample size that

$$\mathbb{P}\left\{\alpha_{j,n_j}|M_{j,\nu_{n_j}}|^{\beta_{j,n_j}}\operatorname{sign}(M_{j,\nu_{n_j}}) \le x\right\} = p_{n_j}(0) + \sum_{k=1}^{\infty} p_{n_j}(k)\left[\mathbb{P}\left\{\alpha_{j,n_j}|M_{j,n_j}|^{\beta_{j,n_j}}\operatorname{sign}(M_{j,n_j}) \le x\right\}\right]^k.$$

Since $\nu_{n_j} \xrightarrow{p} \infty$ as $n_j \to \infty$, we have $\lim_{n_j\to\infty} p_{n_j}(0) = 0$. Therefore,

$$\mathbb{P}\left\{\alpha_{j,n_j}|M_{j,\nu_{n_j}}|^{\beta_{j,n_j}}\operatorname{sign}(M_{j,\nu_{n_j}}) \le x\right\} \quad \sim \quad \mathbb{E}\left\{\exp\left[\left(\frac{\nu_{n_j}}{n_j}\right)n_j \log F_j(|x/\alpha_{j,n_j}|^{1/\beta_{j,n_j}}\operatorname{sign}(x))\right]\right\}. \tag{A7}$$

Noting that condition (??) implies that, there exists a sub-sequence $\nu_{n_j'}$ such that $\nu_{n_j'}/n_j' \xrightarrow{p} V_j$. It follows thus from Theorem 2.1 by ?] that, for every $s > 0$,

$$\lim_{n_j'\to\infty} \mathbb{E}\left\{\exp\left[-s\frac{\nu_{n_j'}}{n_j'}\right]\right\} = \int_0^\infty e^{-sz} d\mathbb{P}\left\{V_j \le z\right\}.$$

This together with condition (??) for $M_{j,n_j}, \alpha_{j,n_j}, \beta_{j,n_j}$ and Eq.(??) implies that

$$\lim_{n_j'\to\infty} \mathbb{E}\left\{\exp\left[\left(\frac{\nu_{n_j}}{n_j}\right)n_j \log F_j(|x/\alpha_{j,n_j}|^{1/\beta_{j,n_j}}\operatorname{sign}(x))\right]\right\} = \int_0^\infty \exp(z\log H_j(x))d\mathbb{P}\left\{V_j \le z\right\}.$$

Consequently, we obtained the claim (??).
Consequently, the claims follow by combining the obtained (??), Eq.(??) and the proof of Theorem 2.1 by ?].

References

. Abd Elgawad, M., Barakat, H., Qin, H., and Yan, T. (2017). Limit theory of bivariate dual generalized order statistics with random index. *Statistics*, 51(3):572–590.

. Barakat, H. and Nigm, E. (2002). Extreme order statistics under power normalization and random sample size. *Kuwait Journal of Science & Engineering*, 29(1):27–41.

. Beirlant, J. and Teugels, J. L. (1992). Limit distributions for compounded sums of extreme order statistics. *Journal of Applied Probability*, 29(3):557–574.

. Berman, S. M. (1962). Limiting distribution of the maximum term in sequences of dependent random variables. *The Annals of Mathematical Statistics*, 33(3):894–908.

. Cao, W. and Zhang, Z. (2021). New extreme value theory for maxima of maxima. *Statistical Theory and Related Fields*, 5(3):232–252.

. Cui, Q., Xu, Y., Zhang, Z., and Chan, V. (2021). Max-linear regression models with regularization. *Journal of Econometrics*, 222(1, Part B):579–600.

. Dorea, C. C. and GonÇalves, C. R. (1999). Asymptotic distribution of extremes of randomly indexed random variables. *Extremes*, 2(1):95–109.
. Embrechts, P., Kluppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events*. Stochastic Modelling and Applied Probability. Springer, Heidelberg.
. Galambos, J. (1978). *The Asymptotic Theory of Extreme Order Statistics*. Wiley series in probability and mathematical statistics. Wiley, New York.
. Grigelionis, B. (2004). On the extreme-value theory for stationary diffusions under power normalization. *Lithuanian Mathematical Journal*, 44:36–46.
. Hashorva, E., Padoan, S. A., and Rizzelli, S. (2021). Multivariate extremes over a random number of observations. *Scandinavian Journal of Statistics*, 48(3):845–880.
. Hu, K., Wang, K., Constantinescu, C., Zhang, Z., and Ling, C. (2023). Extreme limit theory of competing risks under power normalization. *arXiv: 2305.02742*.
. Korolev, V. and Gorshenin, A. (2020). Probability models and statistical tests for extreme precipitation based on generalized negative binomial distributions. *Mathematics*, 8(4):604.
. Leadbetter, M. R., Lindgren, G., and Rootzén, H. (1983). *Extremes and related properties of random sequences and processes*. Springer Science & Business Media.
. Nasri-Roudsari, D. (1999). Limit distributions of generalized order statistics under power normalization. *Communications in Statistics - Theory and Methods*, 28(6):1379–1389.
. Pantcheva, E. (1985). *Limit theorems for extreme order statistics under nonlinear normalization*. Springer Berlin Heidelberg.
. Peng, Z., Jiang, Q., and Nadarajah, S. (2012). Limiting distributions of extreme order statistics under power normalization and random index. *Stochastics*, 84(4):553–560.
. Peng, Z., Shuai, Y., and Nadarajah, S. (2013). On convergence of extremes under power normalization. *Extremes*, 16(3):285–301.
. Ribereau, P., Masiello, E., and Naveau, P. (2016). Skew generalized extreme value distribution: Probability-weighted moments estimation and application to block maxima procedure. *Communications in Statistics-Theory and Methods*, 45(17):5037–5052.
. Shi, P. and Valdez, E. A. (2014). Multivariate negative binomial models for insurance claim counts. *Insurance: Mathematics and Economics*, 55:18–29.
. Soliman, A. A. (2000). Bayes prediction in a Pareto lifetime model with random sample size. *Journal of the Royal Statistical Society. Series D*, 49(1):51–62.
. Tan, Z. and Wu, C. (2014). Limit laws for the maxima of stationary chi-processes under random index. *Test*, 23(4):769–786.
. Tan, Z. Q. (2014). The limit theorems for maxima of stationary Gaussian processes with random index. *Acta Mathematica Sinica*, 30(6):1021–1032.
. Zhang, Z. (2021). Five critical genes related to seven COVID-19 subtypes: A data science discovery. *Journal of Data Science*, 19(1):142–150.