CONVOLUTIONS OF LONG-TAILED AND SUBEXPONENTIAL DISTRIBUTIONS

SERGEY FOSS,* ** *Heriot-Watt University and Sobolev Institute of Mathematics*

DMITRY KORSHUNOV,*** *Sobolev Institute of Mathematics*

STAN ZACHARY,* **** *Heriot-Watt University*

Abstract

Convolutions of long-tailed and subexponential distributions play a major role in the analysis of many stochastic systems. We study these convolutions, proving some important new results through a simple and coherent approach, and also showing that the standard properties of such convolutions follow as easy consequences.

Keywords: Long-tailed distribution; subexponential distribution

2000 Mathematics Subject Classification: Primary 60E05
                            Secondary 60F10; 60G70

1. Introduction

Heavy-tailed distributions play a major role in the analysis of many stochastic systems. For example, they are frequently necessary to accurately model inputs to computer and communications networks, they are an essential component of the description of many risk processes, and they occur naturally in models of epidemiological spread.

Since the inputs to such systems are frequently cumulative in their effects, the analysis of the corresponding models typically features convolutions of such heavy-tailed distributions. The properties of such convolutions depend on their satisfying certain regularity conditions. From the point of view of applications, practically all such distributions may be considered to be long tailed, and indeed to possess the stronger property of subexponentiality (see below for definitions).

In this paper we study convolutions of long-tailed and subexponential distributions (probability measures), and (in passing) more general finite measures, on the real line. Our aim is to prove some important new results, and to do so through a simple, coherent, and systematic approach. It turns out that all the standard properties of such convolutions are then obtained as easy consequences of these results. Thus, we also hope to provide further insight into these properties, and to dispel some of the mystery which still seems to surround the phenomenon of subexponentiality in particular.

Our approach is based on a simple decomposition for such convolutions, and on the concept of 'h-insensitivity' for a long-tailed distribution or measure with respect to some (slowly) increasing function h. This novel approach and the basic, and very simple, new results we

require are given in Section 2. In Section 3 we study convolutions of long-tailed distributions. The key results here are Theorems 1 and 2, which give conditions under which a random shifting preserves tail equivalence; in the remainder of this section we show how other (mostly known) results follow quickly and easily from our approach, and provide some generalisations. In Section 4 we similarly study convolutions of subexponential distributions. The main results here—Theorems 5 and 6—are new, as is Corollary 6; again some classical results are immediate consequences. Finally, in Section 5 we consider closure properties for the class of subexponential distributions. Theorem 7 gives a new necessary and sufficient condition for the convolution of two subexponential distributions to be subexponential, together with a simple demonstration of the equivalence of some existing conditions.

Occasionally, we take a few lines to reprove something from the literature. This enables us to give a self-contained treatment of our subject.

Good introductions to the current state of knowledge on long-tailed and subexponential distributions may be found in [1], [2], [11], and [15].

2. Basic results

We are primarily concerned with probability distributions. However, we find it convenient to also work with more general finite measures, allowing these to be added, convoluted, etc., as usual. As we shall discuss further, our later results for distributions translate easily into this more general setting, where they then provide additional insight.

Recall that, for any two measures F and G and any two nonnegative numbers p and q, their *mixture* $pF + qG$ is also a finite measure defined by $(pF + qG)(B) = pF(B) + qG(B)$ for any Borel set B.

Recall also that a (finite) measure F on $\mathbb{R}$ is *long tailed* if and only if $\overline{F}(x) > 0$ for all x and, for all constants a,

$$\overline{F}(x + a) = \overline{F}(x) + o(\overline{F}(x)) \quad \text{as } x \to \infty, \tag{1}$$

where the *tail function* $\overline{F}$ of the measure F is given by $\overline{F}(x) = F(x, \infty)$. It is easy to see that it is sufficient that relation (1) holds for some $a \neq 0$—an observation due to Landau [13].

We also note that the class of long-tailed measures has the following readily verified closure property (see Proposition 1.3.6(iii) of [5]), which we henceforth use without comment: if measures $F_1, \ldots, F_n$ are long tailed and if the measure F is such that $\overline{F}(x) \sim \sum_{k=1}^n c_k \overline{F_k}(x)$ as $x \to \infty$ for some $c_1, \ldots, c_k > 0$, then F is also long tailed. (Here and throughout, we use '$\sim$' to mean that the ratio of the quantities on either side of this symbol converges to 1. Thus, for example, relation (1) may be written as $\overline{F}(x + a) \sim \overline{F}(x)$ as $x \to \infty$; we further frequently omit, especially in the proofs, the qualifier 'as $x \to \infty$', as all our limits will be of this form.)

We shall make frequent use of the following construction: given any long-tailed measure F, we may choose a positive function h on $\mathbb{R}^+ = [0, \infty)$ such that $h(x) \to \infty$ as $x \to \infty$ and additionally h is increasing sufficiently slowly that

$$\overline{F}(x \pm h(x)) \sim \overline{F}(x) \quad \text{as } x \to \infty.$$

For example, we may choose a sequence x_n increasing to ∞ such that, for all n,

$$|\overline{F}(x \pm n) - \overline{F}(x)| \leq \frac{\overline{F}(x)}{n} \quad \text{for all } x > x_n,$$

and then set $h(x) = n$ for $x \in (x_n, x_{n+1}]$. (The introduction of such a function h will allow us to avoid the general messiness of repeatedly taking limits first as x tends to ∞ and then as a further

constant a—essentially that featuring in (1)—tends to ∞.) Given a long-tailed measure F and a function h satisfying the above conditions, we shall say that F is *h-insensitive*.

Note further that, given a finite collection of long-tailed measures $F_1, \ldots, F_n$, we may choose a function h on $\mathbb{R}^+$ such that each F_i is h-insensitive. For example, for each i, we may choose h_i such that F_i is h_i-insensitive, and then define h by $h(x) = \min_i h_i(x)$.

The tail function of the convolution of any two measures F and G is given by

$$\overline{F * G}(x) = \int_{-\infty}^{\infty} \overline{F}(x - y) G(\mathrm{d}y) = \int_{-\infty}^{\infty} \overline{G}(x - y) F(\mathrm{d}y).$$

For any measure F and any Borel set B, we denote by F_B the measure given by the restriction of F to B, that is, $F_B(A) = F(A \cap B)$ for all Borel sets A. We also use the shorthands $F_{\leq h} = F_{(-\infty, h]}$ and $F_{>h} = F_{(h, \infty)}$. Finally, we define $m_F = F(\mathbb{R})$ to be the total mass associated with the measure F.

Now let h be any positive function on $\mathbb{R}^+$. Then the tail function of the convolution of any two measures F and G possesses the following decomposition: for $x \geq 0$,

$$\overline{F * G}(x) = \overline{F_{\leq h} * G}(x) + \overline{F_{>h} * G}(x),$$ (2)

with upper estimate

$$\overline{F * G}(x) \leq \overline{F_{\leq h} * G}(x) + \overline{F * G_{\leq h}}(x) + \overline{F_{>h} * G_{>h}}(x),$$ (3)

where in (2) and (3) h stands for $h(x)$, so that, for example, $F_{\leq h}(B) = F_{\leq h(x)}(B) = F(B \cap (-\infty, h(x)])$, again for any Borel set B. If, in addition, $h(x) \leq x/2$ then

$$\overline{F * G}(x) = \overline{F_{\leq h} * G}(x) + \overline{F * G_{\leq h}}(x) + \overline{F_{>h} * G_{>h}}(x),$$ (4)

because in this case $\overline{F_{\leq h} * G}(x) = \overline{F_{\leq h} * G_{>h}}(x)$ and $\overline{F * G_{\leq h}}(x) = \overline{F_{>h} * G_{\leq h}}(x)$. Note that

$$\overline{F_{\leq h} * G}(x) = \int_{-\infty}^{h(x)} \overline{G}(x - y) F(\mathrm{d}y),$$ (5)

$$\overline{F_{>h} * G}(x) = \int_{-\infty}^{\infty} \overline{F}(\max(h(x), x - y)) G(\mathrm{d}y),$$ (6)

while $\overline{F_{>h} * G_{>h}}$ is symmetric in F and G, and

$$\overline{F_{>h} * G_{>h}}(x) = \int_{h(x)}^{\infty} \overline{F}(\max(h(x), x - y)) G(\mathrm{d}y) = \int_{h(x)}^{\infty} \overline{G}(\max(h(x), x - y)) F(\mathrm{d}y). \ (7)$$

Note also that if, on some probability space with probability measure P, ξ and η are independent random variables with respective distributions F and G (with $m_F = m_G = 1$), then, for $x \geq 0$,

$$\overline{F_{\leq h} * G}(x) = \mathrm{P}(\xi + \eta > x, \ \xi \leq h(x)),$$

$$\overline{F_{>h} * G_{>h}}(x) = \mathrm{P}(\xi + \eta > x, \ \xi > h(x), \ \eta > h(x)).$$

The following four lemmas are the keys to everything that follows.

Lemma 1. *Suppose that the measure G is long tailed and that h is such that G is h-insensitive. Then, for any measure F,*

$$\overline{F_{\leq h} * G}(x) \sim m_F \overline{G}(x) \qquad \text{as } x \to \infty.$$

Proof. It follows from (5) that $\overline{F_{\leq h} * G}(x) \leq m_F \overline{G}(x - h(x))$. On the other hand,

$$\begin{aligned}
\overline{F_{\leq h} * G}(x) &\geq \overline{F_{[-h,h]} * G}(x) \\
&\geq F[-h(x), h(x)] \overline{G}(x + h(x)) \\
&\sim m_F \overline{G}(x + h(x)) \quad \text{as } x \to \infty,
\end{aligned}$$

where the last equivalence follows since $h(x) \to \infty$ as $x \to \infty$. The required result now follows from the h-insensitivity of G.

We now prove a version of Lemma 1 which is symmetric in F and G, and will allow us to get many important results for convolutions—see the further discussion below.

Lemma 2. *Suppose that measures F and G are such that $m_F G + m_G F$ is long tailed and that h is such that $m_F G + m_G F$ is h-insensitive. Then, as $x \to \infty$,*

$$\overline{F_{\leq h} * G}(x) + \overline{F * G_{\leq h}}(x) \sim m_F \overline{G}(x) + m_G \overline{F}(x).$$

Proof. It follows from (5) that

$$\overline{F_{\leq h} * G}(x) + \overline{F * G_{\leq h}}(x) \leq m_F \overline{G}(x - h(x)) + m_G \overline{F}(x - h(x)).$$

On the other hand,

$$\begin{aligned}
\overline{F_{\leq h} * G}(x) + \overline{F * G_{\leq h}}(x) &\geq \overline{F_{[-h,h]} * G}(x) + \overline{F * G_{[-h,h]}}(x) \\
&\geq F[-h(x), h(x)] \overline{G}(x + h(x)) + G[-h(x), h(x)] \overline{F}(x + h(x)) \\
&\sim m_F \overline{G}(x + h(x)) + m_G \overline{F}(x + h(x)) \quad \text{as } x \to \infty,
\end{aligned}$$

where the last equivalence follows since $h(x) \to \infty$ as $x \to \infty$. The required result now follows from the h-insensitivity of $m_F G + m_G F$.

Note that special cases under which $m_F G + m_G F$ is long tailed are (a) F and G are both long tailed, and (b) F is long tailed and $\overline{G}(x) = o(\overline{F}(x))$ as $x \to \infty$. There are also other possibilities. For instance, the following is an example where F and G are probability distributions such that F is long tailed and $F + G$ is long tailed while G does not satisfy (a) or (b). Take $\alpha > 0$, and let $\overline{F}(x) = x^{-\alpha}$ for $x \geq 1$ and $\overline{F}(x) = 1$ for $x \leq 1$. Let $\overline{G}(x) = 1$ for $x \leq 1$, and construct $\overline{G}$ on $(1, \infty)$ inductively on the intervals $[x_n, x_{n+1}]$, with $x_1 = 1$. For $n = 1, 2, \ldots$, let $y_n = \min\{x > x_n : 1 + \ln(x/x_n) = 2^n\}$ and then, for $x \in [x_n, y_n)$, let $\overline{G}(x) = \overline{F}(x)/(1 + \ln(x/x_n))$. Furthermore, let $\overline{G}(y_n) = \overline{F}(y_n) 2^{-n-1}$ (this means that $\overline{G}(y_n) = \overline{G}(y_n-)/2$). Finally, let $x_{n+1} = \min\{x > y_n : \overline{F}(x) = \overline{G}(y_n)\}$ and then let $\overline{G}(x) = \overline{G}(y_n)$ for $x \in [y_n, x_{n+1}]$. As n increases, the jumps of $\overline{G}$ at the points y_n become negligible with respect to $\overline{F}$, so $F + G$ is long tailed. On the other hand, these jumps are not negligible with respect to $\overline{G}$ itself, and G cannot be long tailed. Also, $\overline{G}(x_n) = \overline{F}(x_n)$, and so condition (a) is violated.

Lemma 2 is important because in many applications, while we may naturally be concerned primarily with long-tailed measures (e.g. service times in queueing models), we nevertheless

require to make small corrections arising from other input measures whose tails are relatively lighter (e.g. those of interarrival times); Lemma 2 is then necessary in order to obtain asymptotic results such as Theorem 5, below, ensuring that such lighter tails indeed make a negligible contribution. In the case where we are solely concerned with long-tailed measures, there are obvious simplifications to the proofs of our main results.

Finally, the following two simple lemmas will be useful when we come to consider subexponential measures.

Lemma 3. *Let h be any positive function on $\mathbb{R}^+$ such that $h(x) \to \infty$. Then, for any measures F_1, F_2, and G on $\mathbb{R}$,*

$$\limsup_{x \to \infty} \frac{\overline{(F_1)_{>h} * G}(x)}{\overline{(F_2)_{>h} * G}(x)} \leq \limsup_{x \to \infty} \frac{\overline{F_1}(x)}{\overline{F_2}(x)}.$$

In particular, in the case where the limit of the ratio $\overline{F_1}(x)/\overline{F_2}(x)$ exists, we have

$$\lim_{x \to \infty} \frac{\overline{(F_1)_{>h} * G}(x)}{\overline{(F_2)_{>h} * G}(x)} = \lim_{x \to \infty} \frac{\overline{F_1}(x)}{\overline{F_2}(x)}.$$

Proof. The results are immediate from (6) and from the first of the integral representations in (7) on noting that $\max(h(x), x - y) \to \infty$ as $x \to \infty$ uniformly in all $y \in \mathbb{R}$.

Taking into account the symmetry of $\overline{F_{>h} * G_{>h}}$ in F and G (see (7)), we also obtain the following result.

Lemma 4. *Let h be any positive function on $\mathbb{R}^+$ such that $h(x) \to \infty$. Then, for any measures F_1, F_2, G_1, and G_2 on $\mathbb{R}$,*

$$\limsup_{x \to \infty} \frac{\overline{(F_1)_{>h} * (G_1)_{>h}}(x)}{\overline{(F_2)_{>h} * (G_2)_{>h}}(x)} \leq \limsup_{x \to \infty} \frac{\overline{F_1}(x)}{\overline{F_2}(x)} \limsup_{x \to \infty} \frac{\overline{G_1}(x)}{\overline{G_2}(x)}.$$

In particular, in the case where the limits of the ratios $\overline{F_1}(x)/\overline{F_2}(x)$ and $\overline{G_1}(x)/\overline{G_2}(x)$ exist, we have

$$\lim_{x \to \infty} \frac{\overline{(F_1)_{>h} * (G_1)_{>h}}(x)}{\overline{(F_2)_{>h} * (G_2)_{>h}}(x)} = \lim_{x \to \infty} \frac{\overline{F_1}(x)}{\overline{F_2}(x)} \lim_{x \to \infty} \frac{\overline{G_1}(x)}{\overline{G_2}(x)}.$$

3. Convolutions of long-tailed distributions

For the remainder of this paper, we specialise to *distributions*, i.e. to probability measures on $\mathbb{R}$ each of total mass 1. This simplifies the statements and proofs of our results, enabling us to dispense with the constants m_F, etc., and is in line with most applications. Of course, our results may nevertheless be translated to the case of more general finite measures by renormalising the latter, applying the results, and re-expressing the conclusions in terms of the original measures.

We denote by $\mathcal{L}$ the class of long-tailed distributions on $\mathbb{R}$. We shall say that two distributions F_1 and F_2 on $\mathbb{R}$ are *tail equivalent* if and only if $\overline{F_1}(x) \sim \overline{F_2}(x)$ as $x \to \infty$. The following two theorems, which provide conditions under which a random shifting preserves tail equivalence, turn out to be of key importance in studying convolutions where at least one of the distributions involved belongs to the class $\mathcal{L}$.

Theorem 1. *Suppose that F_1, F_2, and G are distributions on $\mathbb{R}$ such that F_1 and F_2 are tail equivalent. If $G \in \mathcal{L}$ then the distributions $F_1 * G$ and $F_2 * G$ are tail equivalent.*

Proof. Let the function h on $\mathbb{R}^+$ be such that G is h-insensitive. We use decomposition (2). It follows from Lemma 3 that

$$\overline{(F_2)_{>h} * G}(x) \sim \overline{(F_1)_{>h} * G}(x).$$

Furthermore, by Lemma 1,

$$\overline{(F_2)_{\le h} * G}(x) \sim \overline{G}(x) \sim \overline{(F_1)_{\le h} * G}(x),$$

so that the conclusion of the theorem now follows from (2). $\square$

The next theorem generalises Lemma 2.4(ii) of [7], where the case F_1, F_2, G_1, $G_2 \in \mathcal{L}$ was considered.

Theorem 2. *Suppose that F_1, F_2, G_1, and G_2 are distributions on $\mathbb{R}$ such that $\overline{F}_1(x) \sim \overline{F}_2(x)$ and $\overline{G}_1(x) \sim \overline{G}_2(x)$ as $x \to \infty$. If the measure $F_1 + G_1$ is long tailed then the distributions $F_1 * G_1$ and $F_2 * G_2$ are tail equivalent.*

Proof. The conditions imply that $F_2 + G_2$ is long tailed. Let the function h on $\mathbb{R}^+$ be such that $h(x) \le x/2$ and both $F_1 + G_1$ and $F_2 + G_2$ are h-insensitive. We use decomposition (4). By Lemma 4,

$$\begin{aligned}
\overline{(F_2)_{>h} * (G_2)_{>h}}(x) &= \overline{(F_1)_{>h} * (G_1)_{>h}}(x) + o(\overline{(F_1)_{>h} * (G_1)_{>h}}(x)) \\
&= \overline{(F_1)_{>h} * (G_1)_{>h}}(x) + o(\overline{F_1 * G_1}(x)).
\end{aligned} \tag{8}$$

Furthermore, by Lemma 2,

$$\begin{aligned}
\overline{(F_1)_{\le h} * G_1}&(x) + \overline{F_1 * (G_1)_{\le h}}(x) \\
&= \overline{G}_1(x) + \overline{F}_1(x) + o(\overline{G}_1(x) + \overline{F}_1(x)) \\
&= \overline{G}_2(x) + \overline{F}_2(x) + o(\overline{G}_1(x) + \overline{F}_1(x)) \\
&= \overline{(F_2)_{\le h} * G_2}(x) + \overline{F_2 * (G_2)_{\le h}}(x) + o(\overline{G}_1(x) + \overline{F}_1(x)).
\end{aligned}$$

It further follows from Lemma 2 and from (4) that $\overline{F}_1(x) = O(\overline{F_1 * G_1}(x))$ and $\overline{G}_1(x) = O(\overline{F_1 * G_1}(x))$. Thus, we obtain

$$\overline{(F_1)_{\le h} * G_1}(x) + \overline{F_1 * (G_1)_{\le h}}(x) = \overline{(F_2)_{\le h} * G_2}(x) + \overline{F_2 * (G_2)_{\le h}}(x) + o(\overline{F_1 * G_1}(x)). \tag{9}$$

Hence, again from decomposition (4), together with (8) and (9),

$$\overline{F_2 * G_2}(x) = \overline{F_1 * G_1}(x) + o(\overline{F_1 * G_1}(x)) \quad \text{as } x \to \infty,$$

implying that $F_1 * G_1$ and $F_2 * G_2$ are tail equivalent, as required. $\square$

Theorems 1 and 2 have a number of important corollaries, the first of which is well known from [8], but of which we may now give a very simple proof.

Corollary 1. *The class $\mathcal{L}$ of long-tailed distributions is closed under convolutions.*

Proof. Suppose that F, $G \in \mathcal{L}$. Fix $y > 0$. Define the distribution F_y to be equal to F shifted by $-y$, that is, $\overline{F}_y(x) = \overline{F}(x + y)$. Then $F_y * G$ is equal to $F * G$ shifted by $-y$. Since $F \in \mathcal{L}$, it follows that F and F_y are tail equivalent, and since $G \in \mathcal{L}$ also, it follows from Theorem 1 that $F * G$ and $F_y * G$ are tail equivalent. Hence, $\overline{F * G}(x) \sim \overline{F * G}(x + y)$, implying that $F * G \in \mathcal{L}$. $\square$

We also have the following, and new, generalisation of Corollary 1, the proof of which is identical, except that we must appeal to the (slightly less straightforward) Theorem 2 in place of Theorem 1.

Corollary 2. *Suppose that the distributions F and G are such that $F \in \mathcal{L}$ and $F + G$ is long tailed. Then $F * G \in \mathcal{L}$.*

We remark also that a further special case of Corollary 2 arises when we have $F \in \mathcal{L}$ and G is such that $\overline{G}(x) = o(\overline{F}(x))$ as $x \to \infty$, so that here again $F * G \in \mathcal{L}$.

We give two further general, and known, theorems for convolutions of long-tailed distributions.

Theorem 3. *Suppose that $F_1, \ldots, F_n \in \mathcal{L}$. Then*

$$\overline{F_1 * \cdots * F_n}(x) \geq (1 + o(1)) \sum_{k=1}^{n} \overline{F_k}(x) \quad as \ x \to \infty.$$

Proof. It is sufficient to prove the result for the case in which $n = 2$, the general result following by induction (since, from Corollary 2, the class $\mathcal{L}$ is closed under convolutions). Let the function h be such that $h(x) \leq x/2$, and both F_1 and F_2 are h-insensitive. The required result is now immediate from the inequality

$$\overline{F_1 * F_2}(x) \geq \overline{(F_1)_{\leq h} * F_2}(x) + \overline{F_1 * (F_2)_{\leq h}}(x)$$

and Lemma 1.

In particular, we have the following corollary (where, as usual, F^{*n} denotes the n-fold convolution of F with itself).

Corollary 3. *Suppose that $F \in \mathcal{L}$. Then, for any $n \geq 2$,*

$$\liminf_{x \to \infty} \frac{\overline{F^{*n}}(x)}{\overline{F}(x)} \geq n.$$

Theorem 4. *Suppose that $F \in \mathcal{L}$. Then, for any distribution G on $\mathbb{R}$,*

$$\liminf_{x \to \infty} \frac{\overline{F * G}(x)}{\overline{F}(x)} \geq 1.$$

If, furthermore, the function h on $\mathbb{R}^+$ is such that $F \in \mathcal{L}$ is h-insensitive and $\overline{G}(h(x)) = o(\overline{F}(x))$ as $x \to \infty$, then $\overline{F * G}(x) \sim \overline{F}(x)$ as $x \to \infty$. In particular, this conclusion holds for $F \in \mathcal{L}$ and any distribution G such that $\overline{G}(a) = 0$ for some a.

Proof. Let the function h be such that F is h-insensitive. We use decomposition (2) with F and G interchanged. From Lemma 1,

$$\overline{F * G_{\leq h}}(x) \sim \overline{F}(x), \tag{10}$$

and so the first result is immediate. For the second result, note that, under the given additional condition and for h as above,

$$\overline{F * G_{>h}}(x) \leq \overline{G}(h(x)) = o(\overline{F}(x)),$$

so that the required result again follows on using (2) and (10).

4. Convolutions of subexponential distributions on $\mathbb{R}$

We recall that a distribution F on the positive real line $\mathbb{R}^+$ is *subexponential* if and only if $\overline{F}(x) > 0$ for all x and

$$\overline{F * F}(x) = 2\overline{F}(x) + o(\overline{F}(x)) \quad \text{as } x \to \infty; \tag{11}$$

this notion goes back to Chistyakov [6]. By way of interpretation, suppose that, on some probability space with probability measure P, ξ_1 and ξ_2 are independent random variables with common distribution F. Then, since $P(\max(\xi_1, \xi_2) > x) = 2\overline{F}(x) + o(\overline{F}(x))$, it follows that the subexponentiality of F is equivalent to the condition that $P(\xi_1 + \xi_2 > x) \sim P(\max(\xi_1, \xi_2) > x)$ as $x \to \infty$, i.e. that, for large x, the only significant way in which $\xi_1 + \xi_2$ can exceed x is that either ξ_1 or ξ_2 should itself exceed x. This is the well-known 'principle of a single big jump' for sums of subexponentially distributed random variables.

It is also well known [6] (see also [4]) that if F on $\mathbb{R}^+$ is subexponential then $F \in \mathcal{L}$. (To see this, note that, for any distribution F on $\mathbb{R}^+$ such that $\overline{F}(x) > 0$ for all x, for any $a > 0$ and all $x \ge a$,

$$\overline{F * F}(x) = \overline{F * F_{[0,x-a]}}(x) + \overline{F * F_{(x-a,x]}}(x) + \overline{F * F_{(x,\infty)}}(x)$$
$$\ge \overline{F}(x)F(x-a) + \overline{F}(a)F(x-a,x] + \overline{F}(x);$$

since $F(x-a) \to 1$ and $\overline{F}(a) > 0$, it follows that, under condition (11), we have $F(x-a,x] = o(\overline{F}(x))$ as $x \to \infty$, and so $F \in \mathcal{L}$, as required.)

Following recent practice, we extend the definition of subexponentiality to distributions on the entire real line by saying that a distribution F on $\mathbb{R}$ is subexponential if and only if $F \in \mathcal{L}$ and (11) holds (the latter condition no longer being sufficient to ensure that $F \in \mathcal{L}$). We write $\mathcal{S}$ for the class of subexponential distributions on $\mathbb{R}$.

We shall make repeated use of Lemma 5, below, which follows from the upper estimate (3), Lemma 1 (or Lemma 2), and (11).

Lemma 5. ([3].) *If $F \in \mathcal{L}$ then the following statements are equivalent:*

(i) $F \in \mathcal{S}$;

(ii) *for every function h such that $h(x) \to \infty$, $\overline{F_{>h} * F_{>h}}(x) = o(\overline{F}(x))$ as $x \to \infty$;*

(iii) *for some function h such that $h(x) \to \infty$ and F is h-insensitive, $\overline{F_{>h} * F_{>h}}(x) = o(\overline{F}(x))$ as $x \to \infty$.*

The following statement provides the foundation for our results on convolutions of subexponential distributions.

Lemma 6. *Suppose that $F \in \mathcal{S}$ and that the function h is such that $h(x) \to \infty$ as $x \to \infty$. Let distributions G_1 and G_2 be such that, for $i = 1, 2$, we have $\overline{G_i}(x) = O(\overline{F}(x))$ as $x \to \infty$. Then $\overline{(G_1)_{>h} * (G_2)_{>h}}(x) = o(\overline{F}(x))$ as $x \to \infty$.*

Proof. Since $\overline{G_i}(x) = O(\overline{F}(x))$, it follows from Lemma 4 that

$$\limsup_{x \to \infty} \frac{\overline{(G_1)_{>h} * (G_2)_{>h}}(x)}{\overline{F_{>h} * F_{>h}}(x)} \le \limsup_{x \to \infty} \frac{\overline{G_1}(x)}{\overline{F}(x)} \limsup_{x \to \infty} \frac{\overline{G_2}(x)}{\overline{F}(x)} < \infty.$$

Hence, since $F \in \mathcal{S}$, it follows from Lemma 5 that

$$\overline{(G_1)_{>h} * (G_2)_{>h}}(x) = O(\overline{F_{>h} * F_{>h}}(x)) = o(\overline{F}(x)).$$

The required result follows.

We shall say that distributions F and G on $\mathbb{R}$ are *weakly tail equivalent* if both $\overline{F}(x) = O(\overline{G}(x))$ and $\overline{G}(x) = O(\overline{F}(x))$ as $x \to \infty$. We now have the following corollary to Lemma 6.

Corollary 4. ([12].) *Suppose that* $F \in \mathscr{S}$, *that* $G \in \mathscr{L}$, *and that* F *and* G *are weakly tail equivalent. Then* $G \in \mathscr{S}$.

Proof. Choose the function h on $\mathbb{R}^+$ so that both F and G are h-insensitive. Then, from Lemma 6 and the given weak tail equivalence, $\overline{G_{>h} * G_{>h}}(x) = o(\overline{G}(x))$, and so it follows from Lemma 5 that $G \in \mathscr{S}$.

Remark 1. It follows in particular, as is well known, that subexponentiality is a *tail property* of a distribution, i.e. for any given x_0, it depends only on the restriction of the distribution to the right of x_0. (Indeed, this result also follows from Theorem 2: suppose that distributions F_1 and F_2 on $\mathbb{R}$ are tail equivalent and that $F_1 \in \mathscr{S}$. Since $F_1 \in \mathscr{L}$, we also have $F_2 \in \mathscr{L}$; hence, from Theorem 2, on identifying F_i with G_i for $i = 1, 2$ and on using (11), we have $F_2 \in \mathscr{S}$.) Thus, a distribution F on $\mathbb{R}$ is also subexponential if and only if the distribution F^+ on $\mathbb{R}^+$, given by $\overline{F^+}(x) = \overline{F}(x)$ for $x \geq 0$ and $\overline{F^+}(x) = 0$ for $x < 0$, is subexponential; this provides an alternative definition of subexponentiality on $\mathbb{R}$.

We now have the following theorem.

Theorem 5. *Let (a reference distribution)* $F \in \mathscr{S}$. *Suppose that distributions* $G_1, \ldots, G_n$ *are such that, for each* k, *the measure* $F + G_k$ *is long tailed and* $\overline{G_k}(x) = O(\overline{F}(x))$ *as* $x \to \infty$. *Then*

$$\overline{G_1 * \cdots * G_n}(x) = \sum_{i=1}^{n} \overline{G_i}(x) + o(\overline{F}(x)) \quad \text{as } x \to \infty. \tag{12}$$

Corollary 5. *Let* $F \in \mathscr{S}$ *and so* $F \in \mathscr{L}$. *Suppose that distributions* $G_1, \ldots, G_n$ *are such that, individually for each* k, *either* (i) $G_k \in \mathscr{L}$ *and* $\overline{G_k}(x) = O(\overline{F}(x))$ *as* $x \to \infty$ *or* (ii) $\overline{G_k}(x) = o(\overline{F}(x))$ *as* $x \to \infty$. *Then (12) holds.*

The latter corollary was proved in [10] for the case in which $n = 2$, $G_1 = F$, and $\overline{G_2}(x) = o(\overline{F}(x))$.

Proof of Theorem 5. Note first that it follows from the conditions of the theorem that, for each i and any constant a,

$$\overline{F}(x+a) + \overline{G_i}(x+a) = \overline{F}(x) + \overline{G_i}(x) + o(\overline{F}(x) + \overline{G_i}(x))$$
$$= \overline{F}(x) + \overline{G_i}(x) + o(\overline{F}(x)).$$

Hence, from the representation $F + \sum_{i=1}^{k} G_i = \sum_{i=1}^{k}(F + G_i) - (k-1)F$ and since F is also long tailed, for each k and any constant a,

$$\overline{F}(x+a) + \sum_{i=1}^{k} \overline{G_i}(x+a) = \overline{F}(x) + \sum_{i=1}^{k} \overline{G_i}(x) + o(\overline{F}(x)),$$

and so the measure $F + \sum_{i=1}^{k} G_i$ is also long tailed. Note also that, for each k, we have $\sum_{i=1}^{k} \overline{G_i}(x) = O(\overline{F}(x))$. It now follows that it is sufficient to prove the theorem for the case in which $n = 2$, the general result then following by induction.

Let the function h on $\mathbb{R}^+$ be such that $h(x) \leq x/2$, and all F, $F + G_1$, and $F + G_2$ are h-insensitive. It then follows from Lemma 1 that, as $x \to \infty$,

$$
\begin{aligned}
\overline{G_1 * (G_2)_{\leq h}}(x) &= \overline{(G_1 + F) * (G_2)_{\leq h}}(x) - \overline{F * (G_2)_{\leq h}}(x) \\
&= \overline{G_1 + F}(x) - \overline{F}(x) + o(\overline{G_1}(x) + \overline{F}(x)) \\
&= \overline{G_1}(x) + o(\overline{F}(x)),
\end{aligned} \tag{13}
$$

and, similarly,

$$
\overline{(G_1)_{\leq h} * G_2}(x) = \overline{G_2}(x) + o(\overline{F}(x)). \tag{14}
$$

Furthermore, from Lemma 6,

$$
\overline{(G_1)_{>h} * (G_2)_{>h}}(x) = o(\overline{F}(x)). \tag{15}
$$

The required result (12) now follows from decomposition (4) and from (13)–(15). □

The following result strengthens the conditions of Theorem 5 to provide a sufficient condition for the convolution obtained there to be subexponential.

Theorem 6. *Suppose again that the conditions of Theorem 5 hold, and that additionally G_1 satisfies the stronger condition that $G_1 \in \mathcal{L}$ and that G_1 is weakly tail equivalent to F. Then $G_1 * \cdots * G_n \in \mathcal{S}$, and additionally $G_1 * \cdots * G_n$ is weakly tail equivalent to F.*

Proof. It follows from Corollary 4 that $G_1 \in \mathcal{S}$. Furthermore, the weak tail equivalence of F and G_1 implies that, for each k, $\overline{G_k}(x) = O(\overline{G_1}(x))$. Hence, by Theorem 5 with $F = G_1$, the distribution $G_1 * G_2 * \cdots * G_n$ is long tailed and weakly tail equivalent to G_1 and so also to F. In particular, again by Corollary 4, $G_1 * \cdots * G_n \in \mathcal{S}$. □

We have the following two corollaries of Theorems 5 and 6. The first is new (the version with $G \in \mathcal{L}$ was proved in [8]), while the other is well known (and goes back to [9], where the case in which $n = 2$, $G_1 = G_2$ was considered; some particular results may be found in [14] and [16] (see also [3])).

Corollary 6. *Suppose that distributions F and G are such that $F \in \mathcal{S}$, that $F + G$ is long tailed, and that $\overline{G}(x) = O(\overline{F}(x))$ as $x \to \infty$. Then $F * G \in \mathcal{S}$ and*

$$
\overline{F * G}(x) = \overline{F}(x) + \overline{G}(x) + o(\overline{F}(x)) \quad \text{as } x \to \infty.
$$

Proof. The result follows from Theorems 5 and 6 in the case in which $n = 2$ with G_1 replaced by F and G_2 replaced by G. □

Corollary 7. *Suppose that $F \in \mathcal{S}$. Let $G_1, \ldots, G_n$ be distributions such that $\overline{G_i}(x)/\overline{F}(x) \to c_i$ as $x \to \infty$ for some constants $c_i \geq 0$, $i = 1, \ldots, n$. Then*

$$
\frac{\overline{G_1 * \cdots * G_n}(x)}{\overline{F}(x)} \to \sum_{i=1}^{n} c_i \quad \text{as } x \to \infty.
$$

If $c_1 + \cdots + c_n > 0$ then $G_1 * \cdots * G_n \in \mathcal{S}$.

Proof. The first statement of the corollary is immediate from Theorem 5. If $c_1 + \cdots + c_n > 0$, we may assume without loss of generality that $c_1 > 0$, so that the second statement follows from Theorem 6. □

5. Closure properties of subexponential distributions

It is well known that the class of *regularly varying* distributions, which is a subclass of the class $\mathcal{S}$ of subexponential distributions, is closed under convolution. (A distribution F is regularly varying if its right tail $\overline{F}$ is a regularly varying function. This means that the latter can be represented as $\overline{F}(x) = x^{-\alpha} l(x)$ for all positive x, where $l(x)$ is a *slowly varying* function, i.e. l is strictly positive and $l(cx) \sim l(x)$ as $x \to \infty$ for any positive constant c.) Indeed, if F and G are regularly varying, the result that $F * G$ is also regularly varying is straightforwardly obtained from Theorem 5 by taking the 'reference' distribution of that theorem to be $(F + G)/2$. It is also known that the class $\mathcal{S}$ does not possess this closure property. However, if distributions F and G belong to $\mathcal{S}$, then it follows from Corollary 6 that a sufficient condition for $F * G$ to belong to $\mathcal{S}$ is given by $\overline{G}(x) = O(\overline{F}(x))$ as $x \to \infty$. (Indeed, as the corollary shows, G may satisfy weaker conditions than that of being subexponential.) Furthermore, it follows that under this condition we have, for any function h such that both F and G are h-insensitive,

$$\overline{F_{>h} * G_{>h}}(x) = o(\overline{F}(x) + \overline{G}(x)) \quad \text{as } x \to \infty. \tag{16}$$

(See, for example, the proof of Theorem 5.) The following result is therefore not surprising: if $F, G \in \mathcal{S}$, condition (16) is *necessary and sufficient* for $F * G \in \mathcal{S}$. This is one of the results given by Theorem 7, below. The first three equivalences given by the theorem are known from [8]; the novelty of the result lies in the facts that each of them is equivalent to Theorem 7(iv) (condition (16)) and that we give new and short proofs of these equivalences.

Theorem 7. *Suppose that the distributions F and G on $\mathbb{R}$ are subexponential. Suppose also that p is any constant such that $0 < p < 1$, and that the function h on $\mathbb{R}^+$ is such that $h(x) \le x/2$ and both F and G are h-insensitive. Then the following conditions are equivalent:*

(i) $\overline{F * G}(x) \sim \overline{F}(x) + \overline{G}(x)$ *as $x \to \infty$;*

(ii) $F * G \in \mathcal{S}$;

(iii) *the mixture $pF + (1 - p)G \in \mathcal{S}$;*

(iv) $\overline{F_{>h} * G_{>h}}(x) = o(\overline{F}(x) + \overline{G}(x))$ *as $x \to \infty$.*

Proof. We show that each of the conditions (i)–(iii) is equivalent to condition (iv). First, since F and G are subexponential, and, hence, long tailed, it follows from decomposition (4) and Lemma 1 that

$$\overline{F * G}(x) = \overline{F * G_{\le h}}(x) + \overline{F_{\le h} * G}(x) + \overline{F_{>h} * G_{>h}}(x)$$
$$= \overline{F}(x) + o(\overline{F}(x)) + \overline{G}(x) + o(\overline{G}(x)) + \overline{F_{>h} * G_{>h}}(x). \tag{17}$$

Hence, conditions (i) and (iv) are equivalent.

To show the equivalence of (ii) and (iv), observe first that the subexponentiality of F and G implies that

$$\overline{F^{*2}}(x) \sim 2\overline{F}(x), \qquad \overline{G^{*2}}(x) \sim 2\overline{G}(x), \tag{18}$$

and, thus, in particular, from Lemma 4, that

$$\overline{(F^{*2})_{>h} * (G^{*2})_{>h}}(x) \sim 4\overline{F_{>h} * G_{>h}}(x). \tag{19}$$

Furthermore, since $(F * G)^{*2} = F^{*2} * G^{*2}$ and since both F^{*2} and G^{*2} are h-insensitive, $\overline{(F * G)^{*2}}(x)$ may be estimated as in (17) with F^{*2} and G^{*2} replacing F and G. Hence, using also (18) and (19),

$$\overline{(F * G)^{*2}}(x) = (2 + o(1))(\overline{F}(x) + \overline{G}(x)) + (4 + o(1))\overline{F_{>h} * G_{>h}}(x). \qquad (20)$$

Now since the subexponentiality of F and G also implies, by Corollary 1, that $F * G \in \mathcal{L}$, condition (ii) is equivalent to the requirement that

$$\overline{(F * G)^{*2}}(x) = (2 + o(1))\overline{F * G}(x)$$
$$= (2 + o(1))(\overline{F}(x) + \overline{G}(x)) + (2 + o(1))\overline{F_{>h} * G_{>h}}(x), \qquad (21)$$

where (21) follows from (17). However, equalities (20) and (21) hold simultaneously if and only if $\overline{F_{>h} * G_{>h}}(x) = o(\overline{F}(x) + \overline{G}(x))$, i.e. if and only if condition (iv) holds.

Finally, to show the equivalence of (iii) and (iv), note first that $pF + (1 - p)G$ is h-insensitive. Hence, by Lemma 5, the subexponentiality of $pF + (1 - p)G$ is equivalent to

$$\overline{(pF + (1 - p)G)_{>h} * (pF + (1 - p)G)_{>h}}(x) = o(\overline{F}(x) + \overline{G}(x)).$$

The left-hand side is equal to

$$p^2\overline{F_{>h} * F_{>h}}(x) + (1 - p)^2\overline{G_{>h} * G_{>h}}(x) + 2p(1 - p)\overline{F_{>h} * G_{>h}}(x).$$

By the subexponentiality of F and G and again by Lemma 5, $\overline{F_{>h} * F_{>h}}(x) = o(\overline{F}(x))$ and $\overline{G_{>h} * G_{>h}}(x) = o(\overline{G}(x))$. The equivalence of (iii) and (iv) now follows.

Acknowledgements

References

[1] ASMUSSEN, S. (2000). *Ruin Probabilities*. World Scientific, River Edge, NJ.
[2] ASMUSSEN, S. (2003). *Applied Probability and Queues*, 2nd edn. Springer, New York.
[3] ASMUSSEN, A., FOSS, S. AND KORSHUNOV, D. (2003). Asymptotics for sums of random variables with local subexponential behaviour. *J. Theoret. Prob.* **16,** 489–518.
[4] ATHREYA, K. B. AND NEY, P. E. (1972). *Branching Processes*. Springer, New York.
[5] BINGHAM, N. H., GOLDIE, C. M. AND TEUGELS, J. L. (1987). *Regular Variation*. Cambridge University Press.
[6] CHISTYAKOV, V. P. (1964). A theorem on sums of independent positive random variables and its application to branching random processes. *Theory Prob. Appl.* **9,** 640–648.
[7] CLINE, D. B. H. (1987). Convolutions of distributions with exponential and subexponential tails. *J. Austral. Math. Soc.* **43,** 347–365.
[8] EMBRECHTS, P. AND GOLDIE, C. M. (1980). On closure and factorization properties of subexponential and related distributions. *J. Austral. Math. Soc.* **29,** 243–256.
[9] EMBRECHTS, P. AND GOLDIE, C. M. (1982). On convolution tails. *Stoch. Process. Appl.* **13,** 263–278.
[10] EMBRECHTS, P., GOLDIE, C. M. AND VERAVERBEKE, N. (1979). Subexponentiality and infinite divisibility. *Z. Wahrscheinlichkeitsth.* **49,** 335–347.
[11] EMBRECHTS, P., KLÜPPELBERG, C. AND MIKOSCH, T. (1997). *Modelling Extremal Events*. Springer, Berlin.
[12] KLÜPPELBERG, C. (1988). Subexponential distributions and integrated tails. *J. Appl. Prob.* **25,** 132–141.
[13] LANDAU, E. (1911). Sur les valeurs moyennes de certaines fonctions arithmétiques. *Bull. Acad. Roy. Belgique*, 443–472.
[14] PAKES, A. G. (1975). On the tails of waiting-time distributions. *J. Appl. Prob.* **12,** 555–564.
[15] ROLSKI, T., SCHMIDLI, H., SCHMIDT, V. AND TEUGELS, J. (1998). *Stochastic Processes for Insurance and Finance*. John Wiley, Chichester.
[16] TEUGELS, J. L. (1975). The class of subexponential distributions. *Ann. Prob.* **3,** 1000–1011.