J. Aust. Math. Soc. **78** (2005), 339–371

ASYMPTOTIC EXPANSIONS OF CONVOLUTIONS OF REGULARLY VARYING DISTRIBUTIONS

PHILIPPE BARBE and WILLIAM P. MCCORMICK

(Received 6 January 2001; revised 28 January 2004)

Communicated by V. T. Stefanov

Abstract

In this paper we derive precise tail-area approximations for the sum of an arbitrary finite number of independent heavy-tailed random variables. In order to achieve second-order asymptotics, a mild regularity condition is imposed on the class of distribution functions with regularly varying tails.

Higher-order asymptotics are also obtained when considering a semiparametric subclass of distribution functions with regularly varying tails. These semiparametric subclasses are shown to be closed under convolutions and a convolution algebra is constructed to evaluate the parameters of a convolution from the parameters of the constituent distributions in the convolution. A Maple code is presented which does this task.

2000 *Mathematics subject classification*: primary 62E17; secondary 62E20, 60G70.
Keywords and phrases: convolution, convolution algebra, tail area, asymptotic expansion, regular variation, heavy tail.

1. Introduction

In this paper we make a precise study of tail-area calculations for sums of independent random variables with heavy tails. The need for such evaluations occurs in a variety of situations. Statistical applications include p-value computations as well as construction of confidence intervals. In the nonheavy tailed case, a good reference for such statistical applications is Field and Ronchetti [9]. In insurance risk, tail-areas provide a measure of liability against large claims. The monograph by Embrechts *et al.* [7] provides extensive coverage of this topic. The recent monograph by Willmot and Lin [14] is another reference. In queueing theory, tail-areas model the likelihood of long waits. In portfolio analysis, attention has been given to a quantity referred to

© 2005 Australian Mathematical Society 1446-7887/05 $A2.00 + 0.00

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

as tail risk. Advocates of this financial theory can apply the results in this paper to assess the effect of diversification in a portfolio on this type of risk.

Let F denote a distribution with regularly varying tails. Recall $\overline{F} = 1 - F$ is regularly varying at ∞ with index $-\alpha$ if

$$\lim_{t\to\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = x^{-\alpha}, \quad x > 0.$$

We write $\overline{F} \in R_{-\alpha}$ to signify the above relation holds. We also consider d.f.'s with regularly varying left tails. F is such a distribution provided $F(-t) \in R_{-\alpha}$, that is,

$$\lim_{t\to\infty} \frac{F(-tx)}{F(-t)} = x^{-\alpha}, \quad x > 0.$$

Let us also introduce the following notation to indicate two functions are of the same asymptotic order at ∞. We write $f(t) \asymp g(t)$ as t tends to infinity to signify that

$$0 < \liminf_{t\to\infty} \frac{g(t)}{f(t)} \le \limsup_{t\to\infty} \frac{g(t)}{f(t)} < \infty.$$

Regular variation has the useful property that it is inherited by convolutions. Thus, for example, as shown in Feller [8, page 278], for distribution functions F_i, $i = 1, 2$ with $\overline{F}_i(x) = x^{-\alpha} L_i(x)$ where $L_i \in R_0$, $i = 1, 2$,

$$\overline{F_1 * F_2}(x) \sim x^{-\alpha}\big(L_1(x) + L_2(x)\big) \quad \text{as} \quad x \to \infty.$$

In the case of $F_1 = F_2 \in R_{-\alpha}$, the result gives

$$\lim_{t\to\infty} \frac{\overline{F^{*2}}(t)}{\overline{F}(t)} = 2.$$

Thus, the result not only tells us that the property of regular variation is preserved under convolution but also gives the exact magnitude of increase to the tail-area as a result of convolution. This asymptotic doubling of the tail under convolution is itself an important property and distributions with this property are referred to as subexponential. An extensive literature exists describing properties of this class and we refer to [2, Appendix 4] for the basic results and references.

In Section 2 we introduce a subclass of distribution functions with regularly varying tails for which the second-order behaviour of tail areas under convolution may be determined. In a related paper Geluk *et al.* [10] address similar problems to ours. In that paper they derive their results under an assumption of second-order regular variation. Also very much related to the present paper, Borovkov and Borovkov [3] introduce an asymptotic differentiability condition. This condition appears to

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

be more natural than that used by Geluk *et al.* [10]. Equipped with this condition, Borovkov and Borovkov [3] obtained asymptotic expansions for the tail of a sum of i.i.d. r.v.'s. The striking aspect of these expansions is that they are valid uniformly in the number of summands. Here, our results are obtained under the first order asymptotic differentiability assumption in Borovkov and Borovkov's [3] terminology, which we refer to as asymptotic smoothness. A distinctive feature of this property is that it is preserved under convolutions, thus allowing us to deduce second-order results for arbitrary convolution powers.

In Section 3 we examine higher-order expansions for the tail-areas of arbitrary convolution products. This section is strongly related to the work of Borovkov and Borovkov [3], but our perspective is radically different. Our goal is to develop effective computational tools to obtain tail expansion for weighted sums, while Borovkov and Borovkov [3] focus on the uniformity with respect to the number of i.i.d. r.v.'s involved in a nonweighted sum. In order to provide a clean theory for constructing such expansions, we require that the distributions being convolved belong to a class having a representation of the form $\overline{F}(t) = t^{-\alpha} P(1/t) + o(t^{-m})$, where $P(x) = \sum_{0 \le k \le m} c_k x^k$ for some real constants c_k, nonnegative integer m and positive real α. For example, this applies to the Student distribution with α degrees of freedom. This class is closed under convolutions and we give a formula for computing the coefficients of the associated tail-area polynomial to the convolution power or product. In the case of a weighted sum of i.i.d. random variables with underlying distribution in this class, this procedure is easily done on a computer and we provide the Maple code to achieve this goal. The methodology is quite broadly applicable, and we indicate what adjustments are required to apply the method for a distribution function outside our class but having a finite series representation in some basis other than $t^{-\alpha-k}$, $k = 0, 1, \ldots, m$ described here.

A number of authors for a variety of applications have derived first-order asymptotic results for tail-areas under convolutions of heavy-tailed distributions. We mention Embrechts and Goldie [6] for an application in transient renewal theory, Cohen [5] for the stationary distribution of waiting time in a stable queue, Greenwood [11] and Resnick [12, Section 6] for randomly stopped sums, Resnick [13] for weighted sums, and Bingham *et al.* [2, page 406] for an application to a supercritical branching process.

Finally, we describe the outline of the paper. Section 2 provides the definition of asymptotic smoothness and our general theorem on second-order expansion of tail-areas. Proofs of the results in this section as well as additional supporting lemmas are presented in Section 5. Section 3 presents our algebraic calculus on tail-areas of convolutions. Since the proofs of the statements in this section are short and so as to make the section nearly self-contained, proofs of the propositions are given following their statement. Section 4 addresses the question of the accuracy of these asymptotic

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

expansion and their use in statistics.

2. Convolutions of asymptotically smooth distributions

Recall that a function f is regularly varying at infinity with index ρ if

$$\lim_{t\to\infty} \frac{f(tx)}{f(t)} = x^{\rho}, \quad \text{for all } x > 0.$$

We denote by R_ρ the set of all such measurable functions. Karamata's representation theorem (see, for instance, Bingham *et al.* [2, page 12]) asserts that these functions are those of the form $x^{\rho}c(x)\exp\left(\int_1^x u^{-1}\epsilon(u)\,du\right)$, where $c(x)$ tends to $c \in (0,\infty)$ and $\epsilon(x)$ tends to 0 as x tends to infinity. In particular, the function $c(x)$ is in R_0.

We now introduce a class of functions of main concern for the paper. This class is related to the function satisfying condition $[D_1]$ of Borovkov and Borovkov [3].

DEFINITION. We say that a function f is *asymptotically smooth* with index $-\alpha$ if

$$\lim_{\delta\to 0}\limsup_{t\to\infty}\sup_{0<|x|\le\delta}\left|\frac{f\big(t(1-x)\big)-f(t)}{xf(t)}-\alpha\right| = 0.$$

Asymptotically smooth functions of index α are related to regularly varying ones with the same index. To make this assertion precise, let us introduce the upper and lower densities of a function f,

$$\dot{f}_U(t) = \limsup_{x\to 0}\frac{f(t+x)-f(t)}{x}, \qquad \dot{f}_L(t) = \liminf_{x\to 0}\frac{f(t+x)-f(t)}{x}.$$

PROPOSITION 2.1. (i) *Let f be an asymptotically smooth function with index $-\alpha$. Then $t\,\dot{f}_U(t)/f(t)$ and $t\,\dot{f}_L(t)/f(t)$ have limit $-\alpha$ as t tends to infinity. Moreover, if f is differentiable, then f is in $R_{-\alpha}$.*

(ii) *Let f be regularly varying. Its asymptotic smoothness is equivalent to that of $c(\cdot)$ in the Karamata representation of f, with index 0 for the function $c(\cdot)$. Moreover, if f has an ultimately monotone derivative, then f is asymptotically smooth.*

The first part of statement (ii) implies that if the slowly varying function $x^{\alpha}f(x)$ can be normalized, that is, the function $c(\cdot)$ in its Karamata representation can be choosen constant (see, for example, Bingham *et al.* [2, page 15]), then f is asymptotically smooth. Note the Karamata representation is not unique, nor does every regularly varying function have a normalized form; for instance, a normalized slowly varying function must be differentiable. Moreover, (ii) provides that if f is asymptotically smooth, then any choice of $c(\cdot)$ in the Karamata representation of f is asymptotically

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

smooth with index 0. In other words, although the function $c(\cdot)$ is not unique, asymptotic smoothness places restrictions on its choice. The first statement also shows that asymptotic smoothness is quite analogous to smooth variation (see [2, Section 1.8]); the main distinction is that smooth variation implies differentiability while asymptotic smoothness is only a form of asymptotic differentiability.

For the applications that we have in mind, the last statement of (ii) is the most important. It tells us that if $\overline{F}$ is regularly varying and has an ultimately monotone density, then it is asymptotically smooth. For instance, this implies that the Student, the log-gamma and the Pareto distributions are asymptotically smooth.

For a tail-area function $\overline{F}$, asymptotic smoothness has an interpretation in term of random variables. Indeed, if X has an asymptotically smooth tail-area function $\overline{F}$, then

$$\lim_{\delta\to 0}\limsup_{t\to\infty}\sup_{0\le\epsilon\le\delta}\left|\epsilon^{-1}P\{t\le X\le t+\epsilon t\mid X\ge t\}-\alpha\right|=0.$$

This implies that, when it exists, the hazard rate of X at t is equivalent to α/t as t tends to infinity. Notice that if the hazard rate exists, then F has a density. Then, Proposition 2.1 (i) implies that $\overline{F}$ is regularly varying with index $-\alpha$.

DEFINITION. A distribution function F is *right tail dominant* if

$$\lim_{t\to\infty}\frac{F(-t\delta)}{\overline{F}(t)}=0\quad\text{for any positive }\delta.$$

For any distribution function F, denote the truncated mean by

$$\mu_F(t)=\int_{-t}^{t}x\,dF(x),\quad t\ge 0.$$

If F has a finite absolute moment of order 1,

$$\nu_F=\int_{-\infty}^{\infty}|x|\,dF(x)=\int_0^{\infty}(1-F(x)+F(-x))\,dx,$$

the truncated mean $\mu_F(t)$ has a limit as t tends to infinity. This limit is the usual mean

$$\mu_F=\int_{-\infty}^{\infty}x\,dF(x)=\int_0^{\infty}(1-F(x)-F(-x))\,dx.$$

We can now state a first tail-area expansion for convolutions.

THEOREM 2.2. *Let F and G be two distribution functions such that $\overline{F}\in R_{-\alpha}$ and $\overline{G}\in R_{-\beta}$ with $\alpha\wedge\beta\ge 1$. Assume F and G are asymptotically smooth and right-tail dominant, with $\int_{-\infty}^{0}x\,dF(x)$ and $\int_{-\infty}^{0}x\,dG(x)$ both finite. Then, as t tends to infinity,*

$$1-F*G(t)=\overline{F}(t)+\overline{G}(t)+\frac{1}{t}\left(\alpha\overline{F}(t)\mu_G(t)+\beta\overline{G}(t)\mu_F(t)\right)(1+o(1)).$$

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

In particular, if ν_F and ν_G are finite, the conclusion of Theorem 2.2 yields

$$1 - F * G(t) = \overline{F}(t) + \overline{G}(t) + \frac{1}{t}\left(\alpha\overline{F}(t)\mu_G + \beta\overline{G}(t)\mu_F\right)(1 + o(1))$$

REMARK. If $\mu_F = \mu_G = 0$, the statement of Theorem 2.2 reads

$$1 - F * G(t) = \overline{F}(t) + \overline{G}(t) + o\left(\frac{\overline{F}(t) + \overline{G}(t)}{t}\right) \quad \text{as } t \to \infty.$$

REMARK. It is quite striking that though asymptotic smoothness is not a second order type condition on tails, it yields a two-term expansion on the tail of the convolution. In that aspect, asymptotic smoothness is very much like differentiability: convolution regularizes.

The following result is useful in extending the result in Theorem 2.2 to higher-order convolutions. It asserts that within the class of distribution functions with regularly varying tails, the properties of asymptotic smoothness and right tail dominance are preserved under convolution.

THEOREM 2.3. *Under the assumptions of Theorem* 2.2, $F * G$ *is also asymptotically smooth and right tail dominant.*

REMARK. When F and G are identical, the last assertion in Theorem 2.2 follows from [3]. It is noticeable that Borovkov and Borovkov [3] give a more direct proof of this last assertion, but do not prove that their condition $[D_k]$ is stable by convolution. Theorem 2.3 asserts that their condition $[D_1]$ is stable by convolution. Similar remarks apply to the other results of the current paper.

This stability under convolution has the following interesting consequence. It gives a two-term expansion for the tail of sums of n independent and identically distributed random variables.

PROPOSITION 2.4. *Under the assumptions of Theorem* 2.2, *we have*

$$\lim_{t\to\infty} \frac{t}{\overline{F}(t)\mu_F(t)}\left(1 - F^{*n}(t) - n\overline{F}(t)\right) = \alpha n(n-1).$$

REMARK. As before, if $\mu_F = 0$, Proposition 2.4 reads

$$1 - F^{*n}(t) - n\overline{F}(t) = o\left(\frac{\overline{F}(t)}{t}\right).$$

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

In particular, if ν_F is finite, the conclusion of Proposition 2.4 reads

$$\lim_{t\to\infty} \frac{t}{\overline{F}(t)} \left(1 - F^{*n}(t) - n\overline{F}(t)\right) = \alpha n(n-1)\mu_F.$$

The previous results were obtained under the assumption that $\alpha \geq 1$ or $\alpha \wedge \beta \geq 1$. We now present the analogous results for the case $\alpha \vee \beta < 1$. First for $\alpha \vee \beta < 1$, define the quantity

$$I(\alpha, \beta) = \int_0^{1/2} \left((1-y)^{-\alpha} - 1\right)\beta y^{-\beta-1}\, dy.$$

THEOREM 2.5. *If F and G are asymptotically smooth distribution functions supported on $\mathbb{R}^+$ with $\overline{F} \in R_{-\alpha}$ and $\overline{G} \in R_{-\beta}$ where $\alpha \vee \beta < 1$, then*

$$\lim_{t\to\infty} \frac{1 - F * G(t) - \overline{F}(t) - \overline{G}(t)}{\overline{F}(t)\overline{G}(t)} = I(\alpha, \beta) + I(\beta, \alpha) + 2^{(\alpha+\beta)} - 2^{\alpha} - 2^{\beta}.$$

The next result shows that in the $\alpha \vee \beta < 1$ case, asymptotic smoothness is still preserved by convolution.

THEOREM 2.6. *Under the assumptions of Theorem 2.5, $F * G$ is supported on $\mathbb{R}^+$ and is asymptotically smooth.*

We deduce from Theorem 2.6 the following two-term expansion.

PROPOSITION 2.7. *Under the assumptions of Theorem 2.5,*

$$\lim_{t\to\infty} \frac{1 - F^{*n}(t) - n\overline{F}(t)}{\overline{F}(t)^2} = n(n-1)(I(\alpha, \alpha) + 2^{2\alpha-1} - 2^{\alpha}).$$

We now present an application of the above results which illustrates the difference between our result and those of Geluk *et al.* [10]. Consider the c.d.f.

$$F(t) = 1 - \frac{1}{t^{\alpha}}\left(1 + \frac{1}{\log t}\right) \quad \text{if } t \geq t_0,$$

where $t_0 > 1$ is such that $F(t_0) = 0$. Its density f is ultimately decreasing, since $f''(t) \sim -\alpha(\alpha+1)/t^{\alpha+2}$ as t tends to infinity. Consequently, $\overline{F}$ is asymptotically smooth, thanks to Proposition 2.1 (ii). When $\alpha > 1$, Theorem 2.2 yields

$$\overline{F^{*2}}(t) = 2\overline{F}(t) + 2\alpha\overline{F}(t)\frac{\mu_F}{t} + o(t^{-\alpha-1}) \tag{2.1}$$

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

as t tends to infinity. On the other hand, one can check that $\overline{F}(t)$ is second-order regularly varying, with auxiliary function $a(t) = (\log t)^{-2}$, since

$$\lim_{t\to\infty} (\log t)^2 \left(\frac{\overline{F}(tx)}{\overline{F}(t)} - x^{-\alpha} \right) = -x^{-\alpha} \log x.$$

Theorem 2.1 (ii) of Geluk *et al.* gives, for $\alpha > 1$, $\overline{F^{*2}}(t) = 2\overline{F}(t) + o(a(t)\overline{F}(t))$, which is a coarser estimate than (2.1). We will see in Section 3.4 how to obtain even sharper estimates than (2.1).

3. An algebraic calculus on tails of convolutions

The purpose of this section is to explain how variations on our assumptions can lead to results in the same spirit. The key idea is the obvious fact that convolution is a bilinear operation. This suggests that algebraic arguments may be useful. In particular, if one wants to prove a result on how a bilinear operator acts, then one can first derive this result on a basis and extend its scope by bilinearity. To be more specific, let us rewrite the convolution in a more synthetic way. To a c.d.f. F we associate an operator T_F acting on the space of functions which are integrable with respect to the measure dF by

$$T_F h(t) = \int_{-\infty}^{t/2} h(t-x)\, dF(x).$$

Moreover, we define the multiplication operator M_c by $M_c h(t) = h(t/c)$. Then, Lemma 5.1 asserts that

$$\overline{F * G} = T_F(\overline{G}) + T_G(\overline{F}) + M_2(\overline{F}\ \overline{G}). \tag{3.1}$$

We can now make our algebraic construction.

3.1. Some elementary distributions on $\mathbb{R}^+$ In this subsection, we focus on some very specific distributions. Namely, we assume that F and G are concentrated on $\mathbb{R}^+$ and satisfy $\overline{F}(t) = t^{-\alpha}(1 + o(t^{-a}))$ and $\overline{G}(t) = t^{-\beta}(1 + o(t^{-b}))$ as t tends to infinity, where $a, b > 0$. Those distributions with the $o(\cdot)$ term replaced by a monomial times $1 + o(1)$ are known in the statistical literature on extreme-value theory as the Hall and Weisman [11] model—but this fact is rather irrelevant to what follows, and should not prevent the reader from proceeding. For such distributions,

$$T_F\overline{G}(t) = \int_0^{t/2} t^{-\beta} \left(1 - \frac{x}{t}\right)^{-\beta} dF(x) + o(t^{-b-\beta}).$$

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

Let ϵ be a positive number. Then, there exists a positive δ such that for any u in $[0, \delta]$,

$$\left|(1-u)^{-\beta} - 1 - \sum_{1\le k\le m} \frac{u^k}{k!}\frac{\Gamma(\beta+k)}{\Gamma(\beta)}\right| \le \epsilon u^m.$$

Thus

$$\int_0^{\delta t}\left|t^{-\beta}(1-x/t)^{-\beta} - t^{-\beta}\sum_{1\le k\le m}\left(\frac{x}{t}\right)^k \frac{1}{k!}\frac{\Gamma(\beta+k)}{\Gamma(\beta)}\right| dF(x) \le \frac{\epsilon}{t^{\beta+m}} E_F X^m.$$

Moreover,

$$\int_{\delta t}^{t/2} \overline{G}(t-x)\,dF(x) \le \overline{G}\big(t(1-\delta)\big)\overline{F}(\delta t) = O\big(\overline{F}\,\overline{G}(t)\big),$$

and for any k less than α, an integration by parts show that $\int_{\delta t}^{\infty} x^k\,dF(x) = O(t^{k-\alpha})$. Hence, provided that $m < \alpha$,

$$T_F\overline{G}(t) = \sum_{0\le k\le m} t^{-\beta-k}\frac{\Gamma(\beta+k)}{k!\Gamma(\beta)}\mu_{k,F} + O(t^{-b-\beta} \vee t^{-\beta-m}).$$

Permuting F and G and using formula (3.1), we obtain

$$\begin{aligned} 1 - F * G(t) = \overline{F}(t) + \overline{G}(t) &+ t^{-\alpha}\sum_{1\le k\le n}\frac{1}{t^k}\frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}\frac{\mu_{k,G}}{k!} \\ &+ t^{-\beta}\sum_{1\le k\le m}\frac{1}{t^k}\frac{\Gamma(\beta+k)}{\Gamma(\beta)}\frac{\mu_{k,F}}{k!} \\ &+ o(t^{-\alpha-a} \vee t^{-\beta-b} \vee t^{-\alpha-n} \vee t^{-\beta-m}) \end{aligned}$$

whenever $m < \alpha$ and $n < \beta$.

There are two important features in this formula. First, the expansion of $\overline{F * G}$ is in powers of t of the form $t^{-\alpha-k}$ or $t^{-\beta-k}$. Second, the error is in $1/t$ at the power $\min(\alpha+a, \beta+b, \alpha+n, \beta+m)$. Two conclusions can be drawn.

First, all the terms in the expansion of $\overline{F * G}$ are meaningful for example if $\alpha+n = \beta+m < \min(\alpha+a, \beta+b)$.

The second conclusion is that if one wants to have a class of distributions with tails which have asymptotic expansions in powers of t, and which is stable by convolution, then the choice is very limited. Here we will only study those in $t^{-\alpha-k}$, $k \in \mathbb{N}$. This class is suitable to handle the classical distributions, yields a theory which is neat and easy to use, and which we believe is sufficient for most applications. We will expound this theory in Sections 3.2–3.4.

There is no insight gained in exploring tails with any arbitrary series expansion in power of t; nor is there in adding powers of $\log t$ or other specific functions.

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

3.2. The class $\mathscr{P}^+_{\alpha,m}$ Let us denote $\mathscr{P}^+_{\alpha,m}$ the set of all c.d.f.'s F on the positive half-line, such that $\overline{F}(t) = t^{-\alpha}P_{F,m}(1/t) + o(t^{-\alpha-m})$, as $t \to \infty$, where $P_{F,m}$ is a polynomial of degree at most m. In this context, $\overline{F}(t) \asymp t^{-\alpha}$ as t tends to infinity if and only if $P_{F,m}$ does not vanish at the origin. More generally, if 0 is a root of degree k of $P_{F,m}$, then $\overline{F}(t) \asymp t^{-\alpha-k}$ as t tends to infinity.

Before we discuss convolutions in the class $\mathscr{P}^+_{\alpha,m}$, let us make a few remarks. First, $\mathscr{P}^+_{\alpha,m}$ contains $\mathscr{P}^+_{\alpha,m+k}$ for any positive integer k. Moreover, $\mathscr{P}^+_{\alpha+k,m}$ is a subset of $\mathscr{P}^+_{\alpha,m+k}$. Thus a c.d.f. F may be viewed as variously belonging to these sets, as soon as it belongs to one of them. In practice, one should work with α as large as possible; maximal α occurring when the polynomial $P_{F,m}$ does not vanish at the origin.

One way we can represent the polynomial $P_{F,m}$ is through the vector of its coefficients. If $P_{F,m}(x) = \sum_{0\le i\le m} p_{F,i}x^i$, we can think of $P_{F,m}$ as the vector $p_F = (p_{F,0}, \dots, p_{F,m})^T$ in $\mathbb{R}^{m+1}$. Now, if a random variable X has c.d.f. F belonging to $\mathscr{P}^+_{\alpha,m}$, and r is a positive real number, then rX has c.d.f. $F_r = F(\cdot/r)$. The c.d.f. F_r belongs to $\mathscr{P}^+_{\alpha,m}$ as well, and $P_{F_r,m}(x) = r^\alpha P_{F,m}(rx)$. Thus, changing scale amounts to multiplying the vector p_F by $r^\alpha\mathscr{S}_r$, where $\mathscr{S}_r$ is the diagonal matrix

$$\mathscr{S}_r = \begin{pmatrix} 1 & & & \\ & r & & 0 \\ & & r^2 & \\ 0 & & & \ddots \\ & & & & r^m \end{pmatrix}.$$

We denote by $\mu_{k,F} = \int x^k\,dF(x)$ the k-th moment of F, which is finite whenever $k < \alpha$. Define the $(m+1)\times(m+1)$ lower triangular matrix A_F, whose (i,j) coefficient is given by

$$(A_F)_{i,j} = \begin{cases} \dfrac{\Gamma(\alpha+i)}{\Gamma(\alpha+j)}\dfrac{\mu_{i-j,F}}{(i-j)!} & \text{if } 0 \le j \le i \le m; \\ 0 & \text{otherwise.} \end{cases}$$

Note that we label the rows and columns of A_F from 0 to m (and not from 1 to $m+1$ as would be customary).

All diagonal elements of A_F are equal to 1. In particular, the determinant of A_F is 1, and A_F is invertible. Using the binomial identity expressing $(x+y)^k$, we see that

$$\mu_{k,F*G} = \sum_{0\le i\le k}\binom{k}{i}\mu_{k-i,F}\mu_{i,G}.$$

Equipped with this relation, one checks that $A_{F*G} = A_FA_G = A_GA_F$. Moreover, to calculate A_{F_r} is easy, since $\mu_{k,F_r} = r^k\mu_{k,F}$. It follows that $A_{F_r} = \mathscr{S}_rA_F\mathscr{S}_r^{-1}$.

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

The following result is the key to our tail calculus. It shows that the class $\mathscr{P}^+_{\alpha,m}$ is stable by convolution (which should be rather obvious given what precedes), and it tells us how to calculate the polynomial associated with the expansion of $F * G$. When F and G are identical, it follows from Borovkov and Borovkov [3]; but we will see that our algebraic viewpoint seems much more suitable for practical computation and extension to different classes of distributions.

PROPOSITION 3.1. *If F, G are in $\mathscr{P}^+_{\alpha,m}$ with $m < \alpha$, then $p_{F*G} = A_F p_G + A_G p_F$.*

The proof of Proposition 3.1 is given here because of its simplicity, importance and because the last sentence of the proof illustrates a key argument.

PROOF. When both $P_{F,m}$ and $P_{G,m}$ are monomials,

$$\overline{F}(t) = t^{-\alpha-i} + o(t^{-\alpha-m}) = t^{-\alpha-i}(1 + o(t^{i-m})),$$

and the expansion for $\overline{G}$ is obtained from that of $\overline{F}$ by substituting j for i. The conclusion follows from the formula for $1 - F * G(t)$ obtained in Section 3.1. In general, the conclusion follows from linearity of the operators T_F and T_G or, analogously, by bilinearity of the convolution. □

Let us make a comment on the proof. In general, if one has an expansion not in powers of $t^{-\alpha-k}$ as we assume but in other functions such as $t^{-\alpha-(k/2)}$ or $t^{-\alpha-k}(\log t)^{\beta+l}$, one could still use a linearity argument. Thus the problem of approximating tails of a sum can always be broken into simpler problems where the tail is given by a simple function plus a $o(\cdot)$ term. Everything works in the same way provided these simple functions (here the monomials of degree $-\alpha - k$) are reasonably well behaved. In any such case, provided one introduces the right basis, one can transform the problem of expanding the tail of the convolution into linear algebraic manipulations as we did, eventually even on an infinite-dimensional space.

In the same spirit, we could perfectly well replace the $o(\cdot)$ error terms in our simple functions by an explicit error bound. This would lead to error bound on the convolution in exactly the same way.

3.3. Distributions on $\mathbb{R}$. The class $\mathscr{P}_{\alpha,m}$ Now that we know how to manipulate distributions on the positive half line, we can move on to distributions on the whole line. The key ingredient is again the bilinearity of the convolution. We can decompose any c.d.f. F as a mixture of a c.d.f. supported on the positive half-line and a c.d.f. supported on the negative half-line. More precisely, if F is given, set

$$F_- = \min\left(\frac{F}{F(0)}, 1\right) \quad \text{and} \quad F_+ = \max\left(\frac{F - F(0)}{1 - F(0)}, 0\right).$$

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

If $F(0)$ vanishes then F is supported on $\mathbb{R}^+$, and we set F_- to be whatever distribution we like; while if $F(0) = 1$, then we set F_+ to be whatever distribution we like. With these notations, $F = F(0)F_- + (1 - F(0))F_+$.

However, to deal with convolution, it is easier to change notations, writing

$$F = (1-u)F_- + uF_+ \quad \text{and} \quad G = (1-v)G_- + vG_+.$$

Observe that

$$\begin{aligned}\overline{F * G} = uv\overline{F_+ * G_+} + (1-u)v\overline{F_- * G_+} + (1-v)u\overline{F_+ * G_-} \\ + (1-u)(1-v)\overline{F_- * G_-}.\end{aligned}$$

The distribution $F_- * G_-$ is that of the sum of two independent random variables concentrated on the negative half-line. Thus, $\overline{F_- * G_-}(t) = 0$ for any positive t. Consequently, on the positive half-line,

$$\overline{F * G} = uv\overline{F_+ * G_+} + (1-u)v\overline{F_- * G_+} + (1-v)u\overline{F_+ * G_-}. \tag{3.2}$$

From the preceding sections, we know how to derive an asymptotic expansion for the term $\overline{F_+ * G_+}$. For the other two terms, notice that

$$\overline{F_- * G_+}(t) = \int_{-\infty}^{0} \overline{G}_+(t-x)\,dF_-(x),$$

and we can proceed similarly to what we have done in Section 3.1. As a consequence, let us define the class $\mathscr{P}_{\alpha,m}$.

DEFINITION. We say that F belongs to $\mathscr{P}_{\alpha,m}$ if both $1 - F_-(-t)$ and $F_+(t)$ belongs to $\mathscr{P}^+_{\alpha,m}$.

Thus, if F is in $\mathscr{P}_{\alpha,m}$, then $F(-t) = t^{-\alpha}P^-_{F,m}(1/t) + o(t^{-\alpha-m})$ as t tends to infinity for some polynomial $P^-_{F,m}$. We denote by $P^+_{F,m}$ the polynomial associated with the upper tail of F. Notice that $P^-_F = (1-u)P^-_{F^-}$ and $P^+_F = uP^+_{F^+}$.

We can now extend our tail calculus to the class $\mathscr{P}_{\alpha,m}$.

First, if X has distribution F, then write F_r the distribution of rX. We write p^+_F (respectively p^-_F) the vector associated to P^+_F (respectively P^-_F). Clearly,

$$p^+_{F_r} = \begin{cases} r^\alpha \mathscr{S}_r p^+_F & \text{if } r > 0; \\ (-r)^\alpha \mathscr{S}_{-r} p^-_F & \text{if } r < 0. \end{cases}$$

Thus, we now know how to compute the tail of a multiple of a given random variable whose distribution is in $\mathscr{P}_{\alpha,m}$. The next result tells us, as far as the tail is concerned, how to calculate the polynomial associated with the addition of such variables, when they are independent.

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

THEOREM 3.2. *If F, G are in $\mathscr{P}_{m,\alpha}$, with $m < \alpha$, then $p^+_{F*G} = A_F p^+_G + A_G p^+_F$.*

PROOF. Given (3.2), the upper tail of $F * G$ is given by the polynomial

$$p^+_{F*G} = uvp^+_{F_+*G_+} + (1-u)vp^+_{F_-*G^+} + (1-v)up^+_{F_+*G_-}.$$

Since $p^+_{F_-} = p^+_{G_-} = 0$, we have, applying the same arguments as in Section 2, $p^+_{F_-*G_+} = A_{F_-}\, p^+_{G_+}$. Thus,

$$\begin{aligned} p^+_{F*G} &= uvp^+_{F_+*G_+} + (1-u)vA_{F_-}\, p^+_{G^+} + (1-v)uA_{G_-}\, p^+_{F^+} \\ &= (uA_{F_+} + (1-u)A_{F_-})vp^+_{G_+} + (vA_{G_+} + (1-v)A_{G_-})up^+_{F_+}. \end{aligned}$$

Since $A_F = uA_{F_+} + (1-u)A_{F_-}$ and $vp^+_{G_+} = p^+_G$ and the like with the roles of F and G reversed, the result follows. □

This completes our calculus. It allows one to calculate the tail of weighted sums of random variables in the class $\mathscr{P}_{\alpha,m}$ only by formal linear algebra.

To make this clear, let us derive a neat formula for the tail expression of $\sum_{1\le i\le n} X_i$ when the X_i's are independent and X_i has c.d.f. F_i belonging to $\mathscr{P}_{\alpha,m}$.

THEOREM 3.3. *If $X_1, \ldots, X_n$ are independent, respectively with c.d.f. F_i belonging to $\mathscr{P}_{\alpha,m}$, then $S_n = X_1 + \cdots + X_n$ has a distribution which belongs to $\mathscr{P}_{\alpha,m}$, and whose polynomial is given by*

$$p^+_{S_n} = \sum_{1\le i\le n} \left(\prod_{\substack{1\le j\le n \\ j\ne i}} A_{F_j} \right) p^+_{F_i} = \left(\prod_{1\le i\le n} A_{F_i} \right) \sum_{1\le i\le n} A^{-1}_{F_i} p^+_{F_i}.$$

PROOF. The first formula follows from Theorem 3.2 when $n = 2$. For n larger, we proceed by induction, using that $A_{F*G} = A_F A_G$. The second formula follows from the first one, using the commutativity and the invertibility of the matrices A_{F_i}. □

To apply our calculus, there are a few useful observations which we now relate. Define the lower triangular matrix

$$R_F := \begin{cases} \dfrac{\mu_{i-j,F}}{(i-j)!} & \text{if } 0 \le j \le i \le m; \\ 0 & \text{otherwise.} \end{cases}$$

Again, the binomial identity yields $R_{F*G} = R_F R_G$. Let $\mathscr{G}$ be the diagonal matrix

$$\mathscr{G} = \operatorname{diag}\left(\Gamma(\alpha), \Gamma(\alpha+1), \ldots, \Gamma(\alpha+m)\right).$$

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

Then $A_F = \mathscr{G} R_F \mathscr{G}^{-1}$. Thus, the computation of A_F and multiplication of such matrices can be done with the R_F matrices.

Next, observe that $\mathscr{G}$ and $\mathscr{S}_r$ being both diagonal, they commute. In particular,

$$A_{F_r} = \mathscr{S}_r \mathscr{G} R_F \mathscr{G}^{-1} \mathscr{S}_r^{-1} = \mathscr{G} \mathscr{S}_r R_F \mathscr{S}_r^{-1} \mathscr{G}^{-1} = \mathscr{G} R_{F_r} \mathscr{G}^{-1}.$$

Moreover, $\mathscr{S}_r \mathscr{S}_s = \mathscr{S}_{sr}$.

3.4. Examples

EXAMPLE 1. Tail approximation for the mean. Let $X_1, \ldots, X_n$ be n independent random variables, identically distributed. Assume that their common c.d.f. belongs to $\mathscr{P}_{\alpha,m}$. Let $S_n = X_1 + \cdots + X_n$ be their sum, and $\overline{X}_n = S_n/n$ be their mean. We are seeking for an expansion of $P\{\overline{X}_n \geq t\} = P\{S_n \geq nt\}$. Applying Theorem 3.3, we deduce $p_{S_n}^+ = n A_{F^{*(n-1)}} p_F^+$. Thus, for $0 \leq k \leq m$,

$$(p_{S_n}^+)_k = n \sum_{0 \leq i \leq k} \frac{\Gamma(\alpha + k)}{\Gamma(\alpha + i)} \frac{E(S_{n-1}^{k-i})}{(k-i)!} (p_F^+)_i.$$

Then, for $m < \alpha$,

$$P\{\overline{X}_n \geq t\} = \frac{1}{n^\alpha t^\alpha} \sum_{0 \leq k \leq m} (p_{S_n}^+)_k \frac{1}{(nt)^k} + o(t^{-\alpha-m}) \quad \text{as } t \to \infty.$$

A quite explicit formula! To make this expansion even more explicit, we need to tell how to calculate the moments of S_{n-1}. We will address a more general problem in the next example. In the case of a symmetric distribution about 0, the first terms in such expansion are simple. By virtue of symmetry, the odd moments of S_{n-1} vanish and

$$q_2 = \frac{1}{2!} E S_{n-1}^2 = (n-1)\frac{\mu_2}{2},$$
$$q_4 = \frac{1}{4!} E S_{n-1}^4 = \frac{n-1}{24}(\mu_4 - 3\mu_2^2) + \frac{(n-1)^2}{8}\mu_2^2.$$

Consequently, whenever F is a symmetric c.d.f. belonging to $\mathscr{P}_{\alpha,m}$ for some $6 \leq m < \alpha$, we write $P_F(x) = \sum_{0 \leq i \leq m} p_i x^i$ its tail polynomial, and we have

$$\begin{aligned} P\{\overline{X}_n \geq t\} = \frac{1}{n^{\alpha-1} t^\alpha} \Bigg(& p_0 + \frac{(\alpha+1)\alpha q_2 p_0 + p_2}{n^2 t^2} + \frac{(\alpha+2)(\alpha+1) q_2 p_1 + p_3}{n^3 t^3} \\ & + \frac{(\alpha+3)(\alpha+2)(\alpha+1)\alpha q_4 p_0 + (\alpha+3)(\alpha+2) q_2 p_2 + p_4}{n^4 t^4} \\ & + \frac{(\alpha+4)(\alpha+3)(\alpha+2)(\alpha+1) q_4 p_1 + (\alpha+4)(\alpha+3) q_2 p_3 + p_5}{n^5 t^5} \Bigg) \\ & + O(t^{-\alpha-6}). \end{aligned}$$

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

Notice that if F belongs to $\mathscr{P}_{\alpha,m}$ for some $\alpha \leq 6$, this expression is still valid, provided we truncate it to the last term for which q_i is well defined.

Various formulas of this type can be derived. However, in many applications, it is not clear that it makes too much sense to seek for a very long expansion. Indeed, one should not forget that a probabilistic model never fits perfectly, and that the lack of fit may induce more important errors than the hundredth term of an asymptotic expansion. With regard to this point, we are reminded of Box's admonition, 'all models are wrong; some models are useful', Box [4].

EXAMPLE 2. Let us now turn to the more general problem of approximating the tail of a positively weighted sum. Thus, let $X = (X_1, \ldots, X_n)$ be a random vector with independent and identically distributed components. Again, assume that their common c.d.f. belongs to $\mathscr{P}_{\alpha,m}$. Let $c = (c_1, \ldots, c_n)$ be a vector of nonnegative constants. Denote by $W_n = \sum_{1\leq i\leq n} c_i X_i = \langle c, X\rangle$ the weighted sum of the X_i's. Let us explain how to obtain an approximation of $P\{W_n \geq t\}$ up to $o(t^{-\alpha-m})$ for $m < \alpha$.

It is convenient to define

$$C_p = \sum_{1\leq i\leq n} c_i^p.$$

Applying Theorem 3.3, and using the remark that follows its statement, we need to calculate the matrix

$$\mathscr{A} = A_{F_{c_1}*\cdots*F_{c_n}} \sum_{1\leq i\leq n} A_{F_{c_i}}^{-1} c_i^\alpha \mathscr{S}_{c_i} = \mathscr{G} R_{F_{c_1}*\cdots*F_{c_n}} \sum_{1\leq i\leq n} c_i^\alpha \mathscr{S}_{c_i} R_F^{-1} \mathscr{G}^{-1}.$$

The formula

$$\sum_{1\leq i\leq n} c_i^\alpha \mathscr{S}_{c_i} = \operatorname{diag}(C_\alpha, C_{\alpha+1}, \ldots, C_{\alpha+m}) \tag{3.3}$$

is useful for an actual computation.

We now need to explain how to calculate $R_{F_{c_1}*\cdots*F_{c_n}}$. It is plain that this reduces to a calculation of the moments $E(W_n^k)$ for any integer k. The trick is to derive a formal expression using much more integrable random variables and a computer algebra system. Let indeed Y be a random vector with independent and identically distributed components having a moment-generating function finite in the neighborhood of the origin. Set $V_n = \langle c, Y\rangle$. Then,

$$\frac{E(V_n^k)}{k!} = \frac{1}{k!}\frac{d^k}{dt^k} E\exp(tV_n)\bigg|_{t=0} = \frac{1}{k!}\frac{d^k}{dt^k}\exp\left(\sum_{1\leq i\leq n}\log E(e^{tc_iY_i})\right)\bigg|_{t=0}.$$

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

The following Maple code gives $E(V_n^k)/k!$ as a function of the C_i's and the moments of Y_1. Once the formal expression is obtained, we substitute the moments of X_1 for that of Y_1. This gives us the moments of W_n expressed solely in terms of the C_i's and the moments of X_1.

```
mu(0):=1:  # We represent EY_i^k by mu(k).  Thus mu(0)=1.
# We now write E exp(tc_iY_i) as a polynomial P1 plus a
# remainder term of order t^{-m-1}.
for k from 0 by 1 to m do:
  Q[k]:=t^k*c^k*mu(k)/k!:
od:
P1:=add(Q[k],k=0..m):
# We expand log E exp(tc_iY_i) as a polynomial P2 plus a
# remainder term.
P2:=convert(series(ln(P1),t=0,m+1),polynom);
# now, summing over i corresponds to substitute c_i^k
# with C_k.  We write C(k) for C_k.
P3:=add(C(k)*coeff(P2,c,k),k=0..m):
# We then take the Taylor series of
# exp(∑_{1≤i≤n} log E(exp(tc_iY_i))).
P4:=convert(series(exp(P3),t=0..m),polynom):
# We pull out Q[k] = E(W_n^k)/k!.
for k from 0 by 1 to m do;
  Q[k]:=coeff(P4,t,k);
od;
```

The expressions `Q[k]` contain $E(W_n^k)/k!$ as a function of $\mu_i = EX_1^i$ (that is, `mu(i)` in the code) and the C_k (that is, `C(k)` in the code). For instance, the first ones are

$$EW_n^0 = 1, \quad EW_n = C_1\mu_1, \quad \frac{EW_n^2}{2!} = \frac{C_2}{2}(\mu_2 - \mu_1^2) + \frac{C_1^2}{2}\mu_1^2,$$
$$\frac{EW_n^3}{3!} = \frac{C_3}{6}(\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3) - \frac{C_1C_2}{2}\mu_1(\mu_2 - \mu_1^2) + \frac{1}{6}C_1^3\mu_1^3.$$

We can now build the matrices R_{W_n} and R_F associated with W_n and F, as well as the matrix $\mathscr{G}$ and the matrix defined in (3.3) (denoted `mat33` in the code).

```
RW:=evalm(array(1..m+1,1..m+1,identity)):
RF:=evalm(array(1..m+1,1..m+1,identity)):
G:=evalm(array(1..m+1,1..m+1,identity)):
mat34:=evalm(array(1..m+1,1..m+1,identity)):
for i from 1 by 1 to m+1 do:
  for j from 1 by 1 to i do:
```

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

```
    RW[i,j]:=Q[i-j]:
    RF[i,j]:=mu(i-j)/(i-j)!:
  od;
  G[i,i]:=Gamma(alpha+i-1):
  mat33[i,i]:=C(alpha+i-1):
od:
# We can now calculate the final matrix 𝒜.
AA:=evalm(G &* RW &* mat33 &* RF^(-1) &* G^(-1));
```

For instance, we can obtain the 3×3 matrix $\mathscr{A}$ needed to write a 3-term expansion. Since $\mathscr{A}$ is lower triangular we just need to give its entries with row index larger than the column index. Writing $\sigma^2 = \mu_2 - \mu_1^2$ for the variance of the distribution, we compute

$$A_{1,1} = C_\alpha, \quad A_{2,1} = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)}(C_1 C_\alpha - C_{\alpha+1})\mu_1, \quad A_{2,2} = C_{\alpha+1},$$
$$A_{3,1} = \frac{\Gamma(\alpha+2)}{2\Gamma(\alpha)}\big(\sigma^2(C_2 C_\alpha - C_{\alpha+2}) + \mu_1^2(C_1^2 C_\alpha - 2C_1 C_{\alpha+1} + C_{\alpha+2})\big),$$
$$A_{3,2} = \frac{\Gamma(\alpha+2)}{\Gamma(\alpha+1)}(C_1 C_{\alpha+1} - C_{\alpha+2})\mu_1, \quad A_{3,3} = C_{\alpha+2}.$$

It is needless to say that though these coefficients have been produced by a computer, the asymptotic expansion that they provide has been proved rigorously. Though the proof of the validity of this expansion was computer-assisted, it is nonetheless a formal proof.

EXAMPLE 3. Motivated by some applications in finance, Geluk, Peng and De Vries [10] make a point that if X and Y are independent from a Student distribution with 3 degrees of freedom, then $(X + 7^{1/3}Y)/2$ has a tail which coincides with that of X at the first order, and has a second order which is larger than that of X. Without making any claim that the whole thing is relevant to money making, this raises an amusing problem. Consider n random variables $X_1, \ldots, X_n$ independent and identically distributed from a Student distribution with parameter α (which may not be integer). What are all the linear combinations $c_1 X_1 + \cdots + c_n X_n$ with all c_i's positive, which have the same first order tail as X_1? This is a trivial matter to settle, because (see, for example, Feller [8, Section VIII.8], or use the result of the previous example), as t tends to infinity

$$P\left\{\sum_{1 \le i \le n} c_i X_i \ge t\right\} \sim \sum_{1 \le i \le n} P\{c_i X_i \ge t\} \sim \sum_{1 \le i \le n} c_i^\alpha P\{X_1 \ge t\}.$$

Thus, writing $c = (c_1, \ldots, c_n)$ and $X = (X_1, \ldots, X_n)$, the tail of $\langle c, X\rangle$ matches that of X_1 at the first order if c has unit ℓ_α-norm. In particular, for $\alpha = 3$ and $n = 2$, we see that $c = (1/2, 7^{1/3}/2)$ is such a vector.

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

What is more difficult a priori (but is straightforward with our calculus) is to find among all those vectors, all those which either match the second term of the tail expansion of $\langle c, X\rangle$ and X_1 or those which maximize or minimize the second term of the tail expansion of $\langle c, X\rangle$. For this, assume that $\alpha > 2$.

To use our calculus we first need to show that the Student distribution belongs to $\mathscr{P}_{\alpha,m}$, and calculate its associated polynomial. Define the beta integral

$$B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\,dx,$$

and set $K_\alpha = \left(\sqrt{\alpha}B(1/2,\alpha/2)\right)^{-1}$. The tail-area of the Student distribution with parameter α is

$$\overline{S}_\alpha(t) = K_\alpha \int_t^\infty \left(1+\frac{x^2}{\alpha}\right)^{-(\alpha+1)/2} dx.$$

It is easy to prove that S_α is in $\mathscr{P}_{\alpha,m}$ for any m. Indeed,

$$\overline{S}_\alpha(t) = K_\alpha \int_t^\infty \left(\frac{\alpha}{x^2}\right)^{(\alpha+1)/2}\left(1+\frac{\alpha}{x^2}\right)^{-(\alpha+1)/2} dx.$$

Using the Taylor expansion of $(1+u)^{-(\alpha+1)/2}$, we obtain the tail-polynomial of the Student distribution,

$$P_{S_\alpha,m}(x) = K_\alpha \alpha^{(\alpha+1)/2} \sum_{0\le k\le m/2} \frac{(-\alpha)^k}{k!}\frac{\Gamma((\alpha+1+2k)/2)}{\Gamma((\alpha+1)/2)}\frac{x^{2k}}{\alpha+2k}.$$

We can now proceed. Let us write $p = p_{S_\alpha,2}$ the vector corresponding to the polynomial of degree 2 in the two-term expansion of $\overline{S}_\alpha$. Thus,

$$p = K_\alpha \alpha^{(\alpha+1)/2}\begin{pmatrix} 1/\alpha \\ 0 \\ -\alpha(\alpha+1)/(2(\alpha+2)) \end{pmatrix}.$$

We can either use the result of the previous section or make an ad hoc derivation of the matrix $\mathscr{A}$. Because the Student distribution is symmetric, the calculation of the matrix $\mathscr{A}$ by hand is quite easy. For any real number a, it is convenient for what follows to introduce the matrix

$$J(a) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a & 0 & 1 \end{pmatrix}.$$

One can easily verify the identity $J(a)J(b) = J(a+b)$.

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

Next, because the Student distribution is centered, $\mu_{1,S_\alpha} = 0$. This gives

$$\mathscr{S}_r R_{S_\alpha} \mathscr{S}_r^{-1} = J(r^2 \mu_{2,S_\alpha}/2).$$

Let us write $p_{\langle c,X\rangle}$ the polynomial involved in the upper tail expansion of the distribution of $\langle c, X\rangle$. Define $b_i = (C_2 - c_i^2)\mu_{2,S_\alpha}/2$. Applying Theorem 3.3, we obtain

$$p_{\langle c,X\rangle} = \mathscr{G} \sum_{1\le i\le n} J\left(\sum_{\substack{1\le j\le n\\ j\ne i}} c_j^2 \frac{\mu_{2,S_\alpha}}{2}\right) c_i^\alpha \mathscr{S}_{c_i} \mathscr{G}^{-1} p_{S_\alpha}^+ = \mathscr{G} \sum_{1\le i\le n} J(b_i) c_i^\alpha \mathscr{S}_{c_i} \mathscr{G}^{-1} p_{S_\alpha}^+.$$

Since

$$J(a)\mathscr{S}_r = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r & 0 \\ a & 0 & r^2 \end{pmatrix},$$

$$\begin{aligned} p_{\langle c,X\rangle} &= \mathscr{G} \sum_{1\le i\le n} \begin{pmatrix} c_i^\alpha & 0 & 0 \\ 0 & c_i^{\alpha+1} & 0 \\ b_i c_i^\alpha & 0 & c_i^{\alpha+2} \end{pmatrix} \mathscr{G}^{-1} p_{S_\alpha}^+ \\ &= K_\alpha \alpha^{(\alpha+1)/2} \begin{pmatrix} C_\alpha/\alpha \\ 0 \\ (\alpha+1)(C_2 C_\alpha - C_{\alpha+2})\dfrac{\mu_{2,S_\alpha}}{2} - \dfrac{\alpha(\alpha+1)}{2(\alpha+2)} C_{\alpha+2} \end{pmatrix}. \end{aligned}$$

Since $\mu_{2,S_\alpha} = \alpha/(\alpha-2)$, we deduce

$$\begin{aligned} P\{\langle c, X\rangle \ge t\} = \frac{K_\alpha \alpha^{(\alpha+1)/2}}{t^\alpha} \Bigg(&\frac{C_\alpha}{\alpha} \\ &+ \frac{1}{t^2}\left((\alpha+1)(C_2 C_\alpha - C_{\alpha+2})\frac{\alpha}{2(\alpha-2)} - \frac{\alpha(\alpha+1)}{2(\alpha+2)} C_{\alpha+2}\right)\Bigg) \\ &+ O(t^{-\alpha-4}). \end{aligned}$$

And we hope that the reader enjoys the ease with which this expansion has been obtained!

If one wants to match this two-term expansion to that of $\overline{S}_\alpha$, we must have $C_\alpha = 1$ to match the first term, and

$$(\alpha+1)(C_2 C_\alpha - C_{\alpha+2})\mu_{2,S_\alpha} - \frac{\alpha(\alpha+1)}{\alpha+2} C_{\alpha+2} = -\frac{\alpha(\alpha+1)}{\alpha+2}.$$

With the constraint coming from the matching first term, $C_\alpha = 1$, this equation defines an $(n-2)$-dimensional manifold. We can write it as all the points $(c_1, \ldots, c_n)$

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

providing a solution to

$$\begin{cases} 2\alpha C_{\alpha+2} - (\alpha+2)C_2 = \alpha - 2; \\ C_\alpha = 1. \end{cases}$$

Thus, there are infinitely many such c's as soon as $n > 2$. Now, it is also plain, for a given vector c, to decide if the second term of the expansion of $\langle c, X\rangle$ is larger or smaller than that of X_1.

Let us now turn to the question of optimizing the second-order term in the tail expansion of $\langle c, X\rangle$ subject to matching the first-order term with that of the tail expansion of X_1, that is $C_\alpha = 1$.

First, if $C_\alpha = 1$, then all the c_i's are between 0 and 1. Thus, the map $r \in [0, \infty) \mapsto C_r$ is decreasing; it takes values between 0 and 1 on $[\alpha, \infty)$. It follows that $C_2 \geq C_{\alpha+2}$ whenever $C_\alpha = 1$. Thus, using $C_\alpha = 1$, the coefficient of the second-order term satisfies

$$(\alpha+1)(C_2C_\alpha - C_{\alpha+2})\frac{\alpha}{2(\alpha-2)} - \frac{\alpha(\alpha+1)}{2(\alpha+2)}C_{\alpha+2} \geq -\frac{\alpha(\alpha+1)}{2(\alpha+2)},$$

with equality if and only if one of the c_i's is 1. Thus, if the first-order term of the tail expansion of $\langle c, X\rangle$ matches that of X_1, then the second term is always larger, and equal if and only if $\langle c, X\rangle = X_i$ for some i. Thus, if the first term matches, averaging alway increases the second term!

Finally, to obtain the maximum value of the second term under the constraint $C_\alpha = 1$, it can be shown that this occurs by taking all the c_i's equal to $n^{-1/\alpha}$ yielding maximal value

$$\frac{\alpha(\alpha+1)}{2(\alpha^2-4)}\big((\alpha+2)n - 2\alpha\big)n^{-2/\alpha}.$$

4. Accuracy

One of the main applications of an approximation for a tail probability is to derive approximate confidence intervals. Below, we show that the main line of research that has been carried out by numerous authors, including ourselves, is mainly of theoretical interest, and should not be used indiscriminately in applications.

We make our point by looking at the sample mean of independent and identically distributed Pareto random variables. Thus, consider $X_1, \ldots, X_n$ independent, with $P\{X_i \geq t\} = t^{-\alpha}, t \geq 1$. Let $\overline{X}_n = n^{-1}(X_1 + \cdots + X_n)$ be their mean. Sections 2, 3.3 and 3.4 provide asymptotic expansions for the tail of $\overline{X}_n$. In particular, set

$$f_1(t) = n^{1-\alpha}t^{-\alpha} \quad \text{and} \quad f_2(t) = \frac{\alpha^2}{\alpha-1}\frac{n-1}{n}\frac{1}{t}.$$

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

When $\alpha \geq 3$, Proposition 2.4 or the first example in Section 3.4 show that

$$P\{\overline{X}_n \geq t\} = f_1(t)(1 + f_2(t)) + o(t^{-\alpha-1}), \tag{4.1}$$

as t tends to infinity. When n is large, it would be natural to rely on the central limit theorem, possibly with an Edgeworth correction (see Beirlant *et al.* [1, page 118]) to estimate the upper tail of $\overline{X}_n$. We consider as an example $n = 20$, a value low for a safe application of the central limit theorem, but high in terms of the complexity of the tail distribution of a twenty-fold convolution.

To show how inaccurate expansion (4.1) is, let us give a lower bound on a restricted part of the tail-area.

PROPOSITION 4.1. *For any positive integer n and any* $t < 2^{1/\alpha}$,

$$P\{\overline{X}_n \geq t\} \geq 1 - \frac{1}{12n(t^{-\alpha} - 0.5)^2}.$$

This lower bound looks bad, and is indeed poor! But the asymptotic expansions give a so poor approximation that even this terrible lower bound will do for our purpose. Since its proof is rather irrelevant to our point, it is deferred to the end of Section 5.

Now assume that we are willing to use the approximation given by f_1 to find a quantile of order 99% for the sample mean. This gives us $t = 100^{1/\alpha} n^{(1/\alpha)-1}$. When $\alpha \geq 3$, notice that $100^{1/\alpha} \leq 5$ and $5n^{(1/\alpha)-1} \leq 2^{1/\alpha}$ for any useful values of α and n. Thus, we can use the lower bound from Proposition 4.1,

$$1 - \frac{1}{12n(n^{\alpha-1}/100 - 0.5)^2} \approx 1 - \frac{10,000}{12n^{2\alpha-1}}.$$

For $\alpha \geq 3$, the number $n^{2\alpha-1}$ is large, if not gigantic, as soon as $n \geq 6$ say, making the lower bound larger than $1/2$. Thus, if one wants to use $[0, t]$ as a 99% one-sided prediction interval for $\overline{X}_n$ with $t = 100^{1/\alpha} n^{(1/\alpha)-1} = f_1^{-1}(0.01)$, the fraction of time the interval includes $\overline{X}_n$ would be near 0 for n as large as 20—a complete failure!

One can check in doing numerical computation that adding the term f_2 improves the accuracy, but not significantly compared to the error that is committed. The reason is the following. The second term is roughly like α/t. It is too small to balance the underestimation coming from f_1.

In the following a comparison is made between the true tail-area based on 100,000 simulations and the estimate of it based on our asymptotic expansions. We considered tail-areas for the sum of twenty Pareto variables with parameter $\alpha = 4.2$. We used a fifth-order expansion as shown in Section 3.4. We chose two different basis to

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

construct our expansion. For approximation denoted A_1 in Table 1, we used basis functions

$$(t-19)^{(-4.2-k)}, \quad k=0,1,2,3,4.$$

This shift amounts to shifting the underlying Pareto distribution to have zero left endpoint. For approximation A_2 in Table 1 we used the basis functions

$$t^{(-4.2-k)}, \quad k=0,1,2,3,4.$$

That is, approximation A_2 is the approximation discussed in Section 3.4. As the results show, behaviour of the estimates is sensitive to the basis. We tried a shift to zero mean in the underlying distribution as a natural step to consider. It did not lead to a better result. We also tried the first order approximation,

$$A_3(x) = 20P\{X_i > x\}.$$

The first-order approximation performed uniformly badly in the range considered. For example, as a measure of performance set

$$\text{RelError}_i(x) = |\text{true}(x) - A_i(x)|/A_i(x), \quad i=1,2,3,$$

where true(x) denotes the true tail area to the right of x. Then, at $x=32$ corresponding to a tail-area about 0.01, and at $x=42$ corresponding to tail-area around 10^{-4}, we have

TABLE 1.

approximation	A_1	A_2	A_3
RelError (32)	1.16	26	1,000
RelError (42)	0.07	2	50

5. Proofs

PROOF OF PROPOSITION 2.1. Suppose that f is asymptotically smooth with index α. Choose a sequence x_n converging to 0, so that

$$\lim_{n\to\infty} \frac{f(t(1-x_n)) - f(t)}{tx_n} = \dot{f}_U(t).$$

Now for any ϵ positive by asymptotic smoothness, we have for all $n \geq n_0(\epsilon)$ and $t \geq t_0(\epsilon)$ that

$$\frac{\alpha-\epsilon}{t} \leq \frac{f(t(1-x_n)) - f(t)}{tx_n f(t)} \leq \frac{\alpha+\epsilon}{t}.$$

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

Hence letting n tend to infinity in the above, we obtain for $t \geq t_0(\epsilon)$,

$$\alpha - \epsilon \leq \frac{t\dot{f}_U(t)}{f(t)} \leq \alpha + \epsilon,$$

which gives the result for f_U. An analogous argument shows the result for f_L.

Next, suppose that f has a derivative $\dot{f}$. Then $\dot{f}_U = \dot{f}_L = \dot{f}$. Set $h = \dot{f}/f$. Asymptotic smoothness of f implies $\lim_{t\to\infty} th(t) = \alpha$. Then by Lamperti's result [2, page 59] $f \in R_{-\alpha}$. This proves part (i) of Proposition 2.1.

Let us now prove part (ii). Assume that f is regularly varying with index $-\alpha$. Writing $\ell(x) = x^\alpha f(x)$,

$$\frac{f(t(1-x)) - f(t)}{xf(t)} = \frac{\ell(t(1-x)) - \ell(t)}{x\ell(t)} \frac{1}{(1-x)^\alpha} + \frac{1}{x}\left(\frac{1}{(1-x)^\alpha} - 1\right).$$

Since $\lim_{x\to 0}((1-x)^{-\alpha} - 1)/x = \alpha$, asymptotic smoothness may be rewritten as

$$\lim_{\delta\to 0} \limsup_{t\to\infty} \sup_{0<|x|\leq\delta} \left|\frac{\ell(t(1-x)) - \ell(t)}{x\ell(t)}\right| = 0.$$

By Karamata's representation theorem [2, page 12] we can write

$$\ell(t) = c(t)\exp\left(\int_1^t \frac{\epsilon(u)}{u}\,du\right),$$

where $c(t)$ has limit $c \in (0, \infty)$ and $\epsilon(t)$ tends to 0 as t tends to infinity. Thus,

$$\begin{aligned}\frac{\ell(t(1-x)) - \ell(t)}{x\ell(t)} &= \frac{1}{x}\left(\frac{c(t(1-x))}{c(t)}\exp\left(\int_t^{t(1-x)} \frac{\epsilon(u)}{u}\,du\right) - 1\right)\\ &= \frac{1}{x}\left(\frac{c(t(1-x))}{c(t)} - 1\right)\exp\left(\int_t^{t(1-x)} \frac{\epsilon(u)}{u}\,du\right)\\ &\quad + \frac{1}{x}\left(\exp\left(\int_t^{t(1-x)} \frac{\epsilon(u)}{u}\,du\right) - 1\right).\end{aligned}$$

Now observe for arbitrary positive η, t sufficiently large and arbitrary $|x|$ in $(0, \delta]$ with $\delta < 1$,

$$1 - \eta|x| \leq (1+|x|)^{-\eta} \leq \exp\left(\int_t^{t(1-x)} \frac{\epsilon(u)}{u}\,du\right) \leq (1+|x|)^\eta \leq 1 + \eta|x|.$$

Since $c(t) \to c > 0$, the asymptotic smoothness condition is equivalent to

$$\lim_{\delta\to 0} \limsup_{t\to\infty} \sup_{0<|x|\leq\delta} \left|\frac{c(t(1-x)) - c(t)}{x}\right| = 0.$$

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

This is the asymptotic smoothness of $c(\cdot)$ with index 0.

In particular, this last condition holds in the case ℓ is a normalized slowly varying function, that is, $c(t) = c$ is a constant function. For example, ℓ differentiable with $\epsilon(x) = x\ell'(x)/\ell(x)$ continuous and $o(1)$ at infinity suffice for ℓ to be a normalized slowly varying function. See [2, page 15].

To prove the second assertion of (ii), suppose f belongs to $R_{-\alpha}$ and has an ultimately monotone derivative $\dot{f}$. By [2, Theorem 1.7.2], if $f(x) \sim x^{-\alpha}\ell(x)$ as x tends to infinity, then $\dot{f}(x) \sim \alpha x^{-\alpha-1}\ell_1(x)$, where $\ell_1 \in R_0$. This implies asymptotic smoothness, for

$$\frac{f\big(t(1-x)\big) - f(t)}{xf(t)} = \frac{\int_{t(1-x)}^{t} \dot{f}(u)\,du}{xf(t)} = \frac{\alpha \int_{t(1-x)}^{t} u^{-\alpha-1}\ell_1(u)\,du}{xf(t)}$$
$$= \frac{\alpha \int_{1-x}^{1} t^{-\alpha}v^{-\alpha-1}\ell_1(tv)\,dv}{xf(t)},$$

where $\ell_1(u) = (1/\alpha)u^{\alpha+1}\dot{f}(u) \sim \ell(u)$ as u tends to infinity.

Let δ be a real number in $(0, 1)$. Theorem 1.7.2 in [2] asserts that

$$\lim_{t\to\infty} \frac{\ell_1(ut)}{\ell_1(t)} = 1,$$

uniformly in $\delta \le u \le 2$. Consequently, we have uniformly in $0 < |x| \le \delta$

$$\frac{\alpha \int_{1-x}^{1} t^{-\alpha}v^{-\alpha-1}\ell_1(vt)\,dv}{xf(t)} \sim \frac{t^{-\alpha}\ell_1(t)((1-x)^{-\alpha} - 1)}{xf(t)}$$
$$\sim \frac{(1-x)^{-\alpha} - 1}{x} \quad \text{as } t \to \infty.$$

It is clear that asymptotically smooth follows from this. □

We now turn to the proofs of the remaining results of Section 2. We begin with several preparatory lemmas.

LEMMA 5.1. *Equality* (3.1) *holds.*

PROOF. Write

$$1 - F * G(t) = \int_{-\infty}^{t/2} \overline{F}(t-x)\,dG(x) + \int_{t/2}^{\infty} \overline{F}(t-x)\,dG(x).$$

The lemma then follows from this by application of the identity

$$\int_{t/2}^{\infty} \overline{F}(t-x)\,dG(x) = \overline{F}(t/2)\overline{G}(t/2) + \int_{-\infty}^{t/2} \overline{G}(t-y)\,dF(y).$$ □

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

The next lemma provides the key estimate to prove Theorem 2.2.

LEMMA 5.2. *Under the hypothesis of Theorem* 2.2, *the following asymptotic equivalence holds as t tends to infinity*:

$$T_F\overline{G}(t) = \overline{G}(t) + \beta\frac{\overline{G}(t)}{t}\int_{-t}^{t} x\,dF(x)(1+o(1)).$$

PROOF. Let ϵ be a positive number. Since $\overline{G}$ is asymptotically smooth, there exists a positive δ less than $1/2$ such that for all large t

$$\sup_{0<|x|\le\delta}\left|\frac{\overline{G}(t(1-x))-\overline{G}(t)}{x\overline{G}(t)}-\beta\right|\le\epsilon.$$

We write $T_F\overline{G}(t)-\overline{G}(t)-\beta(\overline{G}(t)/t)\int_{-t}^{t}x\,dF(x)$ as a sum of five terms,

$$\begin{aligned}&\int_{-t\delta}^{t\delta}\big(\overline{G}(t-x)-\overline{G}(t)(1+\beta x/t)\big)\,dF(x)\\&\quad+\int_{-\infty}^{-t\delta}\overline{G}(t-x)\,dF(x)+\int_{t\delta}^{t/2}\overline{G}(t-x)\,dF(x)\\&\quad-\overline{G}(t)\big(F(-t\delta)+\overline{F}(t\delta)\big)\\&\quad-t^{-1}\overline{G}(t)\beta\left(\int_{-t}^{-\delta t}x\,dF(x)+\int_{\delta t}^{t}x\,dF(x)\right).\end{aligned}$$

The first term is bounded by

$$\epsilon t^{-1}\overline{G}(t)\int_{-t\delta}^{t\delta}|x|\,dF(x)\le\epsilon t^{-1}\overline{G}(t)\int_{-t}^{t}|x|\,dF(x).$$

The second term is less than $\overline{G}(t(1+\delta))F(-t\delta)$. Since $E_F(X_-)$ is finite, $F(-t\delta)=o(1/t)$ as t tends to infinity. Hence, the second term is $o(\overline{G}(t)/t)$. The third term is not more than $\overline{G}(t/2)\overline{F}(\delta t)$, which is $O(\overline{GF}(t))$. Similarly, the fourth term is at most $o(\overline{G}(t)/t)+O(\overline{G}(t)\overline{F}(t))$. The last term is $t^{-1}\overline{G}(t)\big(o(1)+\int_{t\delta}^{t}x\,dF(x)\big)$. The lemma follows if we can prove that

$$\overline{F}(t)=o\left(t^{-1}\int_{-t}^{t}|x|\,dF(x)\right)$$

and that

$$\int_{\delta t}^{t}x\,dF(x)=o\left(\int_{-t}^{t}|x|\,dF(x)\right)$$

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

as t tends to infinity. When ν_F is finite, this is plain. Thus, we assume that ν_F is infinite. In this case, we must have $\alpha = 1$. Since $E_F(X_-)$ is finite, we have

$$\int_{-t}^{t} x\,dF(x) \sim \mu_F(t) \sim \int_{1}^{t} x\,dF(x).$$

Proposition 1.5.9.a in [2] shows that the function $L(t) = \int_1^t x\,dF(x)$ is slowly varying. Consequently,

$$\int_{\delta t}^{t} x\,dF(x) = L(t)\left(1 - \frac{L(\delta t)}{L(t)}\right) = o(L(t)).$$

Finally, formula (1.5.8) in [2] implies that $t\overline{F}(t) = o\big(\int_0^t x\,dF(x)\big)$ as t tends to infinity, which concludes the proof. The case $\mu_F = 0$ proceeds similarly. □

PROOF OF THEOREM 2.2. Combining Lemmas 5.1 and 5.2, we see that

$$\begin{aligned}\overline{F * G}(t) &= \overline{F}(t) + \overline{G}(t) + t^{-1}\left(\alpha\overline{F}(t)\mu_G(t) + \beta\overline{G}(t)\mu_F(t)\right)(1 + o(1)) \\ &\quad + M_2\left(\overline{F}\,\overline{G}(t)\right).\end{aligned}$$

To conclude the proof, regular variation and the steps in the proof of Lemma 5.2 show that $\overline{FG}(t/2) = o\big(\overline{F}(t)\mu_G(t) \wedge \overline{G}(t)\mu_F(t)\big)$. □

PROOF OF THEOREM 2.3. First, from Theorem 2.2 we infer that $\overline{F * G} \in R_{-(\alpha\wedge\beta)}$. Next using Lemma 5.1 we obtain

$$\begin{aligned}(5.1)\qquad &\overline{F * G}(t(1-x)) - \overline{F * G}(t) \\ &= \int_{-\infty}^{t(1-x)/2} \overline{F}(t(1-x) - y)\,dG(y) - \int_{-\infty}^{t/2} \overline{F}(t-y)\,dG(y) \\ &\quad + \int_{-\infty}^{t(1-x)/2} \overline{G}(t(1-x) - y)\,dF(y) - \int_{-\infty}^{t/2} \overline{G}(t-y)\,dF(y) \\ &\quad + \overline{F}(t(1-x)/2)\overline{G}(t(1-x)/2) - \overline{F}(t/2)\overline{G}(t/2).\end{aligned}$$

The right-hand side of (5.1) is the sum of the following five terms:

$$(5.2)\qquad T_1 = T_1(F, G) = \int_{-\infty}^{t/2} \left(\overline{F}(t(1-x) - y) - \overline{F}(t-y)\right) dG(y),$$

$$(5.3)\qquad T_2 = T_2(F, G) = \int_{t/2}^{t(1-x)/2} \overline{F}(t(1-x) - y)\,dG(y),$$

$T_3 = T_1(G, F)$, $T_4 = T_2(G, F)$ and

$$(5.4)\qquad T_5 = -\overline{F}(t/2)\overline{G}(t/2) + \overline{F}(t(1-x)/2)\overline{G}(t(1-x)/2).$$

We complete the proof of the theorem through a series of lemmas.

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

LEMMA 5.3. $\lim_{\delta\to 0}\limsup_{t\to\infty}\sup_{0<|x|\le\delta}\left|\frac{T_1+T_3}{x\overline{F*G}(t)}-\alpha\wedge\beta\right|=0.$

PROOF. We first consider T_1 in (5.2). For $0\le y\le t/2$ and $\epsilon>0$, we obtain using the asymptotic smoothness of F, that

$$(\alpha-\epsilon)\frac{|x|t}{t-y}\overline{F}(t-y)\le\overline{F}(t(1-x)-y)-\overline{F}(t-y)\le(\alpha+\epsilon)\frac{|x|t}{t-y}\overline{F}(t-y)$$

for all $0<|x|\le\delta$, for some positive δ and all sufficiently large t.

Therefore,

$$\begin{aligned}\left|T_1-\int_{-\infty}^{t/2}\frac{\alpha xt}{t-y}\overline{F}(t-y)\,dG(y)\right|&\le\epsilon|x|\int_{-\infty}^{t/2}\frac{t}{t-y}\overline{F}(t-y)\,dG(y)\\&\le 2\epsilon|x|\int_{-\infty}^{t/2}\overline{F}(t-x)\,dG(x).\end{aligned}$$

An analogous inequality holds for T_3 with the roles of F and G reversed. It follows that for some positive δ, for any $|x|\le\delta$ and t large

$$\begin{aligned}(5.5)\quad&\left|T_1+T_3-\alpha x\int_{-\infty}^{t/2}\frac{t}{t-y}\overline{F}(t-y)\,dG(y)-\beta x\int_{-\infty}^{t/2}\frac{t}{t-y}\overline{G}(t-y)\,dF(y)\right|\\&\le 2\epsilon|x|\left(\int_{-\infty}^{t/2}\overline{F}(t-y)\,dG(y)+\int_{-\infty}^{t/2}\overline{G}(t-y)\,dF(y)\right)\\&\le 2\epsilon|x|\overline{F*G}(t),\end{aligned}$$

where we used Lemma 5.1 in the last step.

Thus by (5.5) it is clear that Lemma 5.3 holds provided

$$\begin{aligned}(5.6)\qquad&\alpha\int_{-\infty}^{t/2}\frac{t}{t-y}\overline{F}(t-y)\,dG(y)+\beta\int_{-\infty}^{t/2}\frac{t}{t-y}\overline{G}(t-y)\,dF(y)\\&\sim(\alpha\wedge\beta)\overline{F*G}(t)\end{aligned}$$

as t tends to infinity. To that end notice that for any η positive,

$$\begin{aligned}(5.7)\qquad&0\le\int_{-\infty}^{-\eta t}\frac{t}{t-y}\overline{F}(t-y)\,dG(y)\le\frac{1}{1+\eta}\overline{F}(t[1+\eta])G(-\eta t)\\&=o(\overline{F*G}(t))\quad\text{as } t\to\infty,\end{aligned}$$

with a similar result with F and G reversed. Furthermore,

$$(5.8)\qquad 0\le\int_{\eta t}^{t/2}\frac{t}{t-y}\overline{F}(t-y)\,dG(y)\le 2\overline{F}(t/2)\overline{G}(\eta t)=o(\overline{F*G}(t)),$$

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

as t tends to infinity, and again one can reverse F and G in (5.8). Thus in view of (5.7) and (5.8), to show (5.6) it is enough to show that

$$\lim_{\eta\to 0}\limsup_{t\to\infty}\left|\frac{1}{(\alpha\wedge\beta)\overline{F*G}(t)}\left(\alpha\int_{-\eta t}^{\eta t}\frac{t}{t-y}\overline{F}(t-y)\,dG(y)\right.\right.$$
$$\left.\left.+\ \beta\int_{-\eta t}^{\eta t}\frac{t}{t-y}\overline{G}(t-y)\,dF(y)\right)-1\right|=0. \tag{5.9}$$

To establish (5.9) note that for $-\eta t\le y\le \eta t$, we have

$$\frac{1}{1+\eta}\le\frac{t}{t-y}\le\frac{1}{1-\eta}.$$

Hence.

$$\frac{1}{1+\eta}\left(\alpha\int_{-\eta t}^{\eta t}\overline{F}(t-y)\,dG(y)+\beta\int_{-\eta t}^{\eta t}\overline{G}(t-y)\,dF(y)\right)$$
$$\le\alpha\int_{-\eta t}^{\eta t}\frac{t}{t-y}\overline{F}(t-y)\,dG(y)+\beta\int_{-\eta t}^{\eta t}\frac{t}{t-y}\overline{G}(t-y)\,dF(y)$$
$$\le\frac{1}{1-\eta}\left(\alpha\int_{-\eta t}^{\eta t}\overline{F}(t-y)\,dG(y)+\beta\int_{-\eta t}^{\eta t}\overline{G}(t-y)\,dF(y)\right). \tag{5.10}$$

Next, note that by a straight forward argument, one can check that

$$\int_{-\eta t}^{\eta t}\overline{F}(t-y)\,dG(y)\sim\overline{F}(t)\quad\text{as } t\to\infty, \tag{5.11}$$

with, of course, the result being also true with F and G interchanged. It is clear that (5.9) now follows from (5.10), (5.11) and the fact that $\overline{F*G}(t)\sim\overline{F}(t)+\overline{G}(t)$ as t tends to infinity. Hence Lemma 5.3 holds. □

LEMMA 5.4. $\lim_{\delta\to 0}\limsup_{t\to\infty}\sup_{0<|x|\le\delta}\dfrac{|T_2|+|T_4|}{x\overline{F*G}(t)}=0.$

PROOF. For x positive, t large and arbitrary fixed positive ϵ,

$$|T_2|\le\overline{F}(t(1-2x)/2)(\beta+\epsilon)x\overline{G}(t/2),$$

where we used asymptotic smoothness of G. Thus for $0<x\le\delta<1/2$

$$\left|\frac{T_2}{x\overline{F*G}(t)}\right|\le\frac{\overline{F}\big(t(1-2x)/2\big)\overline{G}(t/2)(\beta+\epsilon)}{\overline{F*G}(t)}=o(1)\quad\text{as } t\to\infty,$$

where the asymptotic order holds uniformly in $0<x\le\delta<1/2$ by the Uniform Convergence Theorem for regular variation.

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

For $0 < -x \le \delta < 1/2$, we obtain by the same argument that

$$\left|\frac{T_2}{x\overline{F * G}(t)}\right| \le \frac{\overline{F}(t(1+2x)/2)\overline{G}(t/2)(\beta+\epsilon)}{\overline{F * G}(t)} = o(1) \quad \text{as } t \to \infty$$

uniformly in $x \in [-\delta, 0)$. Since the term $T_4 = T_2(G, F)$ this proves the lemma. □

LEMMA 5.5. $\lim_{\delta\to 0} \limsup_{t\to\infty} \sup_{0<|x|\le\delta} \left|\frac{T_5}{x\overline{F * G}(t)}\right| = 0.$

PROOF. Using asymptotic smoothness of F and G, we obtain for ϵ positive, $0 < |x| \le \delta$ and t large,

$$\begin{aligned}&\left|\overline{F}\,\overline{G}(t(1-x)/2) - \overline{F}\,\overline{G}(t/2)\right| \\ &\quad \le (\alpha+\epsilon)|x|\overline{F}(t/2)\overline{G}(t/4) + (\beta+\epsilon)|x|\overline{F}(t/2)\overline{G}(t/2).\end{aligned}$$

Hence

$$\left|\frac{\overline{F}\,\overline{G}(t(1-x)/2) - \overline{F}\,\overline{G}(t/2)}{x\overline{F * G}(t)}\right| = o(1) \quad \text{as } t \to \infty, \tag{5.12}$$

where the order holds uniformly in $0 < |x| \le \delta$. Thus the lemma holds by (5.4) and (5.12). □

Thus, by Lemmas 5.3–5.5, we obtain that $F * G$ is asymptotically smooth. Hence to finish the proof of Theorem 2.3, it remains to show that $F * G$ is right tail dominant, which can be seen by an easy argument. □

PROOF OF PROPOSITION 2.4. We first establish that

$$\mu_{F^{*k}}(t) \sim k\mu_F(t) \quad \text{as } t \to \infty. \tag{5.13}$$

To see this, first consider the case where ν_F is finite. Then (5.13) is clear since $\mu_{F^{*k}}(t) \sim kEX \sim k\mu_F(t)$ as t tends to infinity. Now assume that ν_F is infinite. We have that since $|E_F X_-| < \infty$,

$$\mu_{F^{*k}}(t) \sim \int_0^t x\,dF^{*k}(x) = -t\overline{F^{*k}}(t) + \int_0^t \overline{F^{*k}}(x)\,dx, \tag{5.14}$$

as t tends to infinity. Note in this case, $\alpha = 1$. Hence since $\overline{F^{*k}}(t) \sim k\overline{F}(t) \in R_{-1}$, we have by [2, (1.5.8)],

$$t\overline{F^{*k}}(t) = o\left(\int_1^t \overline{F^{*k}}(x)\,dx\right) \quad \text{as } t \to \infty. \tag{5.15}$$

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

Thus (5.13) holds in this case from (5.14) and (5.15) upon noting

$$\lim_{x\to\infty} \frac{\overline{F^{*k}}(x)}{k\overline{F}(x)} = 1.$$

Now, the proposition follows by induction, using Theorems 2.2 and 2.3. □

The following result provides the key estimate for the proof of Theorem 2.5.

LEMMA 5.6. *Under the hypothesis of Theorem* 2.5,

$$\lim_{t\to\infty} \int_0^{t/2} \left(\frac{\overline{F}(t-x)}{\overline{F}(t)} - 1 \right) \frac{1}{\overline{G}(t)}\, dG(x) = I(\alpha, \beta).$$

PROOF. First, observe that by the uniform convergence theorem for regularly varying functions, for $0 < \delta < 1/2$

$$\lim_{t\to\infty} \sup_{\delta\le y\le 1/2} \left| \frac{(\overline{F}(t(1-y))/\overline{F}(t)) - 1}{(1-y)^{-\alpha} - 1} - 1 \right| = 0.$$

Consequently,

$$\begin{aligned} \int_\delta^{1/2} \left(\frac{\overline{F}(t(1-y))}{\overline{F}(t)} - 1 \right) \frac{1}{\overline{G}(t)}\, dG(ty) \\ \sim \int_\delta^{1/2} \frac{(1-y)^{-\alpha} - 1}{\overline{G}(t)}\, dG(ty) \quad \text{as } t \to \infty. \end{aligned} \tag{5.16}$$

Next, for $0 < a < b$,

$$\int_a^b \frac{1}{\overline{G}(t)}\, dG(ty) = \frac{\overline{G}(ta^-) - \overline{G}(tb)}{\overline{G}(t)} \sim a^{-\beta} - b^{-\beta} \quad \text{as } t \to \infty.$$

Thus, the measures $\overline{G}(t)^{-1}\, dG(ty)$ converge vaguely on $(0, \infty)$ to $\beta y^{-\beta-1} dy$. This implies,

$$\int_\delta^{1/2} \frac{(1-y)^{-\alpha} - 1}{\overline{G}(t)}\, dG(ty) \sim \int_\delta^{1/2} ((1-y)^{-\alpha} - 1)\beta y^{-\beta-1}\, dy, \tag{5.17}$$

as t tends to infinity. Now, note that since $\alpha \vee \beta < 1$,

$$\lim_{\delta\to 0} \int_\delta^{1/2} \big((1-y)^{-\alpha} - 1\big)\beta y^{-\beta-1}\, dy = I(\alpha, \beta) < \infty. \tag{5.18}$$

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

Hence by (5.16)–(5.18) to complete the proof of the lemma, it remains to show

$$\lim_{\delta\to 0}\limsup_{t\to\infty}\int_0^\delta\left(\frac{\overline{F}(t(1-y))}{\overline{F}(t)}-1\right)\frac{1}{\overline{G}(t)}\,dG(ty)=0. \tag{5.19}$$

By asymptotic smoothness we have for ϵ positive and $0<y\le\delta$ with δ sufficiently small and t large

$$\begin{aligned}(\alpha-\epsilon)\int_0^\delta\frac{y}{\overline{G}(t)}\,dG(ty)&\le\int_0^\delta\left(\frac{\overline{F}\big(t(1-y)\big)}{\overline{F}(t)}-1\right)\frac{1}{\overline{G}(t)}\,dG(ty)\\&\le(\alpha+\epsilon)\int_0^\delta\frac{y}{\overline{G}(t)}\,dG(ty).\end{aligned} \tag{5.20}$$

Now observe that

$$\int_0^\delta\frac{y}{\overline{G}(t)}\,dG(ty)=\frac{1}{t\overline{G}(t)}\left(-t\delta\overline{G}(t\delta)+\int_0^{t\delta}\overline{G}(y)\,dy\right)\le\int_0^\delta\frac{\overline{G}(ty)}{\overline{G}(t)}\,dy. \tag{5.21}$$

By Potter's bounds [2, page 25], we have on $[\delta\overline{G}(t),\delta]$ that $\overline{G}(ty)/\overline{G}(t)\le cy^{-(\beta+\epsilon)}$ for all t large and arbitrary fixed positive ϵ and some constant c. Choose ϵ positive such that $\beta+\epsilon<1$. Then

$$\int_0^\delta\frac{\overline{G}(ty)}{\overline{G}(t)}\,dy=\int_{\delta\overline{G}(t)}^\delta\frac{\overline{G}(ty)}{\overline{G}(t)}\,dy+\int_0^{\delta\overline{G}(t)}\frac{\overline{G}(ty)}{\overline{G}(t)}\,dy\le c\int_0^\delta y^{-(\beta+\epsilon)}\,dy+\delta.$$

This upperbound tends to 0 as δ does. Thus, (5.19) follows by virtue of (5.20)–(5.21). □

PROOF OF THEOREM 2.5. Use Lemma 5.1 to obtain

$$\begin{aligned}&\frac{1-F*G(t)-\overline{F}(t)-\overline{G}(t)}{\overline{F}(t)\overline{G}(t)}\\&=\frac{1}{\overline{F}(t)\overline{G}(t)}\left(\int_0^{t/2}\overline{F}(t-x)\,dG(x)\right.\\&\quad\left.+\int_0^{t/2}\overline{G}(t-x)\,dF(x)+\overline{F}(t/2)\overline{G}(t/2)-\overline{F}(t)-\overline{G}(t)\right)\\&=\int_0^{t/2}\left(\frac{\overline{F}(t-x)}{\overline{F}(t)}-1\right)\frac{dG(x)}{\overline{G}(t)}+\int_0^{t/2}\left(\frac{\overline{G}(t-x)}{\overline{G}(t)}-1\right)\frac{dF(x)}{\overline{F}(t)}\\&\quad+\frac{\overline{F}(t/2)}{\overline{F}(t)}\frac{\overline{G}(t/2)}{\overline{G}(t)}-\frac{\overline{G}(t/2)}{\overline{G}(t)}-\frac{\overline{F}(t/2)}{\overline{F}(t)}.\end{aligned} \tag{5.22}$$

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

Taking limits in (5.22) and using Lemma 5.6, which is also valid with F and G interchanged, yields

$$\lim_{t\to\infty} \frac{1 - F * G(t) - \overline{F}(t) - \overline{G}(t)}{\overline{F}(t)\overline{G}(t)} = I(\alpha, \beta) + I(\beta, \alpha) + 2^{(\alpha+\beta)} - 2^{\alpha} - 2^{\beta},$$

proving the result. □

PROOF OF THEOREM 2.6. Argue as in the proof of Theorem 2.3 breaking

$$\overline{F * G}(t(1-x)) - \overline{F * G}(t)$$

into a sum $T_1 + \cdots + T_5$. The proof of Lemma 5.3 goes through up to and including (5.10), while Lemmas 5.4 and 5.5 still hold. Hence to finish the proof it suffices to show that $\int_{-\eta t}^{\eta t} \overline{F}(t-y)\, dG(y) \asymp \overline{F}(t)$ as $t \to \infty$. This is a consequence of the regular variation of $\overline{F}$, and, of course, F and G may be interchanged as well in this relation. □

PROOF OF PROPOSITION 2.7. Use induction and Theorems 2.5 and 2.6. □

PROOF OF PROPOSITION 4.1. If $U_1, \ldots, U_n$ are independent and uniformly distributed random variables over $[0, 1]$, then $U_i^{-1/\alpha}$ has a Pareto distribution with index α. Since the function $x \in (0, 1] \mapsto x^{-1/\alpha}$ is convex, Jensen's inequality gives

$$n^{-1}(U_1^{-1/\alpha} + \cdots + U_n^{-1/\alpha}) \geq \left(n^{-1}(U_1 + \cdots + U_n)\right)^{-1/\alpha}.$$

Consequently,

$$\begin{aligned} P\{\overline{X}_n \geq t\} &= P\{n^{-1}(U_1^{-1/\alpha} + \cdots + U_n^{-1/\alpha}) \geq t\} \\ &\geq P\left\{(n^{-1}(U_1 + \cdots + U_n))^{-1/\alpha} \geq t\right\} \\ &= 1 - P\{U_1 + \cdots + U_n \geq nt^{-\alpha}\}. \end{aligned}$$

Applying Tchebicheff's inequality, we obtain, in the range $t^{\alpha} < 2$,

$$\begin{aligned} P\{\overline{X}_n \geq t\} &\geq 1 - P\{U_1 + \cdots + U_n - (0.5n) \geq n(t^{-\alpha} - 0.5)\} \\ &\geq 1 - \frac{1}{12n(t^{-\alpha} - 0.5)^2}. \end{aligned}$$

This is the desired lower bound. □

NOTE ADDED IN PROOF. Regarding Proposition 2.1, it can be shown that the class of asymptotically smooth functions with index $-\alpha$ coincides with that of normalized regularly varying ones. It also coincides with regularly varying functions which are ultimately absolutely continuous and have a version of their Radon-Nikodým derivative regularly varying with index $-\alpha - 1$.

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press

Acknowledgements

Part of this work was done when the first author was visiting the statistics department at Yale University. He is very grateful to both his Yale colleagues and students for a wonderful year, and to the CNRS for giving so much freedom. We thank the referee for forcing us to shorten the too long original version of this paper.

References

[1] J. Beirlant, J. L. Teugels and P. Vynckier, *Practical analysis of extreme values* (Leuven University Press, Leuven, 1996).

[2] N. H. Bingham, C. M. Goldie and J. L. Teugels, *Regular variation* (Cambridge University Press, Cambridge, 1987).

[3] A. A. Borovkov and K. A. Borovkov, 'On probabilities of large deviations for random walks. I. Regularly varying distribution tails', *Theory Probab. Appl.* **46** (2001), 193–213.

[4] G. E. P. Box, 'Robustness is the strategy of scientific model building', in: *Statistics* (eds. R. L. Launer and G. N. Wilkinson) (Academic Press, 1979).

[5] J. W. Cohen, 'On the tail of the stationary waiting time distribution and limit theorems for the M/G/1 queue', *Ann. Inst. H. Poincaré Probab. Statist.* **8** (1972), 255–263.

[6] P. Embrechts and C. M. Goldie, 'On convolution tails', *Stochastic Process. Appl.* **13** (1982), 263–278.

[7] P. Embrechts, C. Klüppelberg and T. Mikosch, *Modeling extremal events* (Springer, New York,1997).

[8] W. Feller, *An introduction to probability theory and its applications* (Wiley, New York,1971).

[9] C. Field and E. Ronchetti, *Small sample asymptotics*, Lecture Notes Monograph Series 13 (Institute of Mathematical Statistics, Hayward, California, 1990).

[10] J. Geluk, L. Peng and G. de Vries, 'Convolutions of heavy tailed random variables and applications to portfolio diversification and MA(1) time series', *Adv. Appl. Probab.* **32** (2000), 1011–1026.

[11] P. Hall and I. Weissman, 'On the estimation of extreme tail probabilities', *Ann. Statist.* **25** (1997), 1311–1326.

[12] S. I. Resnick, 'Point processes, regular variation and weak convergence', *Adv. Appl. Probab.* **18** (1986), 66–138.

[13] ———, *Extreme values, regular variation, and point processes* (Springer, New York,1987).

[14] G. E. Willmot and X. S. Lin, *Lundberg approximations for compound distributions with insurances applications* (Springer, New York, 2001).

CNRS
90 rue de Vaugirard
75006 Paris
France

Department of Statistics
University of Georgia
Athens, GA 30602
USA
e-mail: bill@stat.uga.edu

https://doi.org/10.1017/S1446788700008570 Published online by Cambridge University Press