THIELE CENTRE
FOR APPLIED MATHEMATICS IN NATURAL SCIENCE

On error rates in rare event simulation with heavy tails

Søren Asmussen and Dominik Kortschak

12

Research Report

No. 01 | May 2012

On error rates in rare event simulation with heavy tails

Søren Asmussen[1] and Dominik Kortschak[2]

[1]Department of Mathematical Sciences, Aarhus University, asmus@imf.au.dk
[2]Institut de Science Financière et d'Assurances, Université Lyon 1, kortschakdominik@gmail.com

Abstract

For estimating $\mathbb{P}(S_n > x)$ by simulation where $S_k = Y_1 + \cdots + Y_k$ with $Y_1, \ldots, Y_n$ are heavy-tailed with distribution F, (Asmussen and Kroese 2006) suggested the estimator $n\,\overline{F}\big(M_{n-1} \vee (x - S_{n-1})\big)$ where $M_k = \max(Y_1, \ldots, Y_k)$. The estimator has shown to perform excellently in practice and has also nice theoretical properties. In particular, (Hartinger and Kortschak 2009) showed that the relative error goes to 0 as $x \to \infty$. We identify here the exact rate of decay and propose some related estimators with even faster rates.

1 Introduction

This paper is concerned with the efficient simulation of

$$z \;=\; z(x) \;=\; \mathbb{P}(S_n > x)\,,$$

where $Y_1, \ldots, Y_n$ are i.i.d. with a common subexponential distribution F, $S_n = Y_1 + \cdots + Y_n$ and x is large so that z is small (for background on subexponential distributions, see, e.g., Embrechts et al. (1997), (Asmussen and Albrecher, 2010, X.1), or Foss et al. (2011)). Recall from the outset the standard fact (or definition of subexponentiality!) that $z \sim n\overline{F}(x)$ as $n \to \infty$ where $\overline{F}(x) = 1 - F(x)$ is the tail.

This problem has a long history. As is traditional in the literature, we denote by a simulation estimator a r.v. $Z = Z(x)$ that can be generated by simulation and is unbiased, $\mathbb{E}Z = z$. The usual performance measure is the relative error $e(x) = (\mathrm{Var}\ Z)^{1/2}/z$. The relative error is bounded if $\limsup_{x\to\infty} e(x) < \infty$, and the estimator Z is logarithmically efficient if $\limsup_{x\to\infty} z(x)^{\epsilon} e(x) < \infty$ for all $\epsilon > 0$.

Efficient estimators have long been known with light tails (see e.g. (Asmussen and Glynn, 2007, VI.2), Bucklew (1990), Heidelberger (1995), Juneja and Shahabuddin (2006) for surveys), and are typically based on ideas from large deviations theory implemented via exponential change of measure. The heavy tailed case is more recent. In Asmussen et al. (2000), some of the difficulties in a literal translation of the light-tailed ideas are explained. However, Asmussen and Binswanger (1997) gave the first logarithmically efficient estimator for $\mathbb{P}(S > x)$ using a conditional Monte Carlo idea. The idea was further improved in Asmussen and Kroese (2006), which as of today stands as a model of an efficient and at the same time easily implementable algorithm, and is also at the core of this paper. The idea is to combine an exchangeability argument with the conditional Monte Carlo idea. More precisely (assuming existence of densities to exclude multiple maxima) one has

$$z \;=\; n\mathbb{P}(S > x,\, M_n = Y_n)\,.$$

where $M_k = \max\bigl(Y_1, \ldots, Y_k\bigr)$. An unbiased simulation estimator of z based on simulated values $Y_1, \ldots, Y_n$ is therefore the conditional expectation

$$Z_{\mathrm{AK}} \;=\; n\overline{F}\bigl(M_{n-1} \vee (x - S_{n-1})\bigr)$$

of this expression given $Y_1, \ldots, Y_{n-1}$, where $S_{n-1} = Y_1 + \cdots + Y_{n-1}$.This estimator, baptized the Asmussen-Kroese estimator by the simulation community, is shown in Asmussen and Kroese (2006) to have bounded relative error and in Hartinger and Kortschak (2009) to have vanishing relative error ($e(x) \to 0$), though the argument for this is rather implicit and no quantitative rates are given.

The contribution of this note is two-fold: to compute the exact error rate of Z_{AK} in the regularly varying case; and to produce another estimator with a better rate in some cases. Both aspects combine with ideas of higher order subexponential methodology (cf. Remark 1). For subexponential distributions with a lighter tail than regular variation like the Weibull, a corresponding theory is developed in Asmussen and Kortschak (2012) and summarized in part in Section 5.

For the regularly varying case, our main result is the following:

Theorem 1. *Assume $f(x) = \alpha L(x)/x^{\alpha+1}$. If $\alpha > 2$ or, more generally, $\mathbb{E}[Y^2] < \infty$ then*

$$\mathrm{Var}\ Z_{\mathrm{AK}} \sim n^2\,\mathrm{Var}[S_{n-1}]f(x)^2 = n^2(n-1)\,\mathrm{Var}[Y_1]f(x)^2.$$

If $\alpha = 2$ and $\mathbb{E}[Y^2] = \infty$ then

$$\mathrm{Var}\ Z_{\mathrm{AK}} \sim 2n^2(n-1)f(x)^2 \int_0^x y\overline{F}(y)\,\mathrm{d}y.$$

If $\alpha < 2$ then $\mathrm{Var}\ Z_{\mathrm{AK}} \sim n^2(n-1)k_\alpha\overline{F}(x)^3$ *where*

$$\begin{aligned} k_\alpha &= \Bigl(2^\alpha + \frac{1}{3}2^{3\alpha} - 2^{2\alpha} + \alpha\int_0^{1/2}\bigl((1-y)^{-\alpha} - 1\bigr)^2 y^{-\alpha-1}dy\Bigr) \\ &= \alpha\int_0^\infty \bigl[((1-y)\vee y)^{-\alpha} - 1\bigr]^2 y^{-\alpha-1}\,\mathrm{d}y\,. \end{aligned}$$

Remark 1. A main idea of higher order subexponential methodology is the Taylor expansion

$$\overline{F}(x - S_{n-1}) \;= \overline{F}(u) + f(x)S_{n-1} + \cdots \qquad (1.1)$$

which easily leads to the refinement

$$\mathbb{P}(S_n > x) \;=\; n\overline{F}(x) + nf(x)\mathbb{E}S_{n-1} + \cdots$$

at least in the regularly varying case, cf. Omey and Willekens (1987), Baltrūnas and Omey (1998) and Barbe and McCormick (2009). Technically, the Taylor expansion is only useful for moderate S_{n-1}, and large values have to be shown to be negligible by a separate argument; this also is the case in the present paper. One may note that (1.1) is only useful for heavy-tailed distributions where typically $\overline{F}(x) \gg f(x) \gg f'(x) \gg \cdots$ – for light-tailed distributions like the exponential typically $\overline{F}(x), f(x), f'(x), \ldots$ have the same magnitude. □

Remark 2. The rates for Var Z_{AK} in Theorem 1 have to be compared with the bounded relative error rate $L(x)^2/x^{2\alpha}$. For $\alpha > 2$, one sees an improvement to $L(x)^2/x^{2\alpha+2}$, for $\alpha < 2$ to $L(x)^3/x^{3\alpha}$. In Section 3, we exhibit an estimator improving this rate for $\alpha > 2$ and in Section 4 one for $1 < \alpha < 2$.

The feature of vanishing relative error is quite unusual. The few further examples we know of are Blanchet and Glynn (2008) and Dupuis et al. (2007) in the setting of dynamic importance sampling, though it should be remarked that the algorithms there are much more complicated than those of this paper and that the rate results are not very explicit. □

Remark 3. In applications to ruin theory and the M/G/1 queue, the number n of terms in S_n is an independent r.v. With some effort, our theory can be refined to this case, but we will not give the details here. □

2 Proof of theorem 1

Proof of Theorem 1. We will use the notation $S_{(n-2)} = S_{n-1} - M_{n-1}$ and $A_{n,x} = \big\{M_{n-1} \le x/2(n-1)\big\}$. Then the the Asmussen-Kroese estimator can be written as $Z_{\mathrm{AK}} = n(X_1 + X_2)$ where

$$X_1 \;=\; I(A^c_{n,x})\overline{F}\big(M_{n-1} \vee (x - S_{n-1})\big)\,, \quad X_2 \;=\; I(A_{n,x})\overline{F}\big(M_{n-1} \vee (x - S_{n-1})\big)\,.$$

Recall that the density of M_{n-1} is $(n-1)f(y)F(y)^{n-2}$ and hence the tail is $\sim (n-1)\overline{F}(y)$. We have

$$\begin{aligned}\mathbb{E}[X_1^k] &= \mathbb{E}\big[\overline{F}(M_{n-1} \vee x - S_{n-1})^k;\, S_{(n-2)} > \sqrt{x},\, A^c_{n,x}\big] \\ &\quad + \mathbb{E}\big[\overline{F}(M_{n-1} \vee x - S_{n-1})^k;\, S_{(n-2)} \le \sqrt{x},\, A^c_{n,x}\big].\end{aligned}$$

The first term is a $\mathrm{O}\big(\overline{F}(x)^{k+1}\overline{F}(\sqrt{x})\big)$, since we can bound $\overline{F}(\cdot)$ by $\mathrm{O}\big(\overline{F}(x)\big)$ and the event $S_{(n-2)} > \sqrt{x}$, $A_{n,x}$ has probability $\mathrm{O}\big(\overline{F}(x)\overline{F}(\sqrt{x})\big)$ since it occurs only if at least one in the i.i.d. sample $Y_1, \ldots, Y_{n-1}$ that is not the maximum exceeds $\sqrt{x}/(n-2)$ and another exceeds $x/2$.

For the second term, note that here $M_{n-1} \le S_{n-1} \le M_{n-1} + \sqrt{x}$. Hence an upper bound is

$$\begin{aligned}
&\mathbb{E}\big[\overline{F}\big(M_{n-1} \vee (x - M_{n-1} - \sqrt{x})\big)^k;\ A_{n,x}^c\big] \\
&\quad = \int_{x/(2(n-1))}^{(x-\sqrt{x})/2} \overline{F}(x - \sqrt{x} - y)^k (n-1) f(y) F(y)^{n-2}\,\mathrm{d}y \\
&\qquad + \int_{(x-\sqrt{x})/2}^{\infty} \overline{F}(y)^k (n-1) f(y) F(y)^{n-2}\,\mathrm{d}y \\
&\quad \sim (n-1)\Big(\alpha \int_{1/(2(n-1))}^{1/2} y^{-\alpha-1}(1-y)^{-k\alpha}\,\mathrm{d}y + \frac{1}{k+1} 2^{\alpha(k+1)}\Big)\overline{F}(x)^{k+1}
\end{aligned}$$

where we used

$$\begin{aligned}
&\int_{x/(2(n-1))}^{(x-\sqrt{x})/2} \overline{F}(x - \sqrt{x} - y)^k f(y) F(y)^{n-2}\,\mathrm{d}y \\
&\quad \sim \int_{x/(2(n-1))}^{(x-\sqrt{x})/2} \overline{F}(x - \sqrt{x} - y)^k f(y)\,\mathrm{d}y \\
&\quad = \int_{1/(2(n-1))}^{(1-1/\sqrt{x})/2} \overline{F}(x - \sqrt{x} - xz)^k f(xz) x\,\mathrm{d}z \\
&\quad \sim x\overline{F}(x)^k f(x) \int_{1/(2(n-1))}^{(1-1/\sqrt{x})/2} (1 - 1/\sqrt{x} - z)^{-\alpha k} z^{-\alpha-1}\,\mathrm{d}z \\
&\quad \sim \alpha\overline{F}(x)^{k+1} \int_{1/(2(n-1))}^{1/2} (1-z)^{-\alpha k} z^{-\alpha-1}\,\mathrm{d}z.
\end{aligned}$$

Omitting $\sqrt{x}$ gives the same lower bound, so

$$\mathbb{E}[X_1^k] \sim (n-1)\Big(\alpha \int_{1/(2(n-1))}^{1/2} y^{-\alpha-1}(1-y)^{-k\alpha}\,\mathrm{d}y + \frac{1}{k+1} 2^{\alpha(k+1)}\Big)\overline{F}(x)^{k+1}.$$

Next consider X_2. If $M_{n-1} \le x/2(n-1)$, then $S_{n-1} \le x/2$ and so

$$\begin{aligned}
X_2 &= I(A_{n,x})\overline{F}(x - S_{n-1}) \\
&= \overline{F}(x) - I(A_{n,x}^c)\overline{F}(x) + I(A_{n,x}) S_{n-1} f(\Xi) = \overline{F}(x) - X_{2,1} + X_{2,2}
\end{aligned} \tag{2.1}$$

where $x - S_{n-1} \le \Xi \le x$. Now

$$\mathbb{E}[X_{2,1}^k] \;=\; \overline{F}(x)^k \mathbb{P}(A_{n,x}^c) \;\sim\; (n-1)2^\alpha (n-1)^\alpha \overline{F}(x)^{k+1}.$$

To evaluate $\mathbb{E}[X_{2,2}^k]$, we split the expectation into $S_{n-1} \le \epsilon x$ and $S_{n-1} \ge \epsilon x$, for some $1/(2(n-1)) > \epsilon > 0$. When $S_{n-1} \le \epsilon x$, we have $f(x) \le f(\Xi) \le f\big((1-\epsilon)x\big)$. This and monotone convergence gives

$$\lim_{\epsilon\to 0} \lim_{x\to\infty} \frac{\mathbb{E}\big[S_{n-1}^k f(\Xi)^k;\ S_{n-1} \le \epsilon x\big]}{f(x)^k \mathbb{E}\big[S_{n-1}^k;\ S_{n-1} \le \epsilon x\big]} = 1\,.$$

Further by (Bingham et al., 1989, Proposition 1.5.8 and 1.5.9a) and partial integration

$$\mathbb{E}\big[S_{n-1}^k;\, S_{n-1} \le \epsilon x\big] = k\int_0^{\epsilon x} y^{k-1}\overline{F}_{S_{n-1}}(y)\mathrm{d}y - (\epsilon x)^k\overline{F}_{S_{n-1}}(\epsilon x)$$

$$\sim \begin{cases} \mathbb{E}S_{n-1}^k & \mathbb{E}Y_1^k < \infty, \\ k\int_0^x y^{k-1}\overline{F}_{S_{n-1}}(y)\mathrm{d}y & \alpha = k, \\ \epsilon^{k-\alpha}\frac{\alpha}{k-\alpha}x^k\overline{F}_{S_{n-1}}(x) & \alpha < k. \end{cases}$$

If $\epsilon x \le S_{n-1} \le x/2$ then it holds uniformly in S_{n-1} that

$$\lim_{x\to\infty}\frac{S_{n-1}f(\Xi)}{\overline{F}(x)} = \lim_{x\to\infty}\frac{\overline{F}(x - S_{n-1}) - \overline{F}(x)}{\overline{F}(x)} = (1 - S_{n-1}/x)^{-\alpha} - 1\,,$$

and hence

$$\mathbb{E}\big[S_{n-1}^k f(\Xi)^k;\, S_{n-1} > \epsilon x,\, M_{n-1} \le x/(2(n-1))\big]$$
$$\sim \overline{F}(x)^k\mathbb{E}\big[\big((1 - S_{n-1}/x)^{-\alpha} - 1\big)^k;\, S_{n-1} > \epsilon x,\, M_{n-1} \le x/(2(n-1))\big]\,.$$

As above we can split the expectation into $S_{(n-2)} \le \sqrt{x}$ and $S_{(n-2)} > \sqrt{x}$, such that we can prove that for $x \to \infty$

$$\mathbb{E}\big[\big((1 - S_{n-1}/x)^{-\alpha} - 1\big)^k;\, S_{n-1} > \epsilon x,\, M_{n-1} \le x/(2(n-1))\big]$$
$$\sim \mathbb{E}\big[\big((1 - M_{n-1}/x)^{-\alpha} - 1\big)^k;\, \epsilon x < M_{n-1} \le x/(2(n-1))\big]$$
$$\sim (n-1)\int_{\epsilon x}^{x/(2(n-1))}\big((1 - y/x)^{-\alpha} - 1\big)^k f(y)\,\mathrm{d}y$$
$$\sim (n-1)x\int_{\epsilon}^{1/(2(n-1))}\big((1 - y)^{-\alpha} - 1\big)^k f(yx)\,\mathrm{d}y$$
$$\sim (n-1)xf(x)\int_{\epsilon}^{1/(2(n-1))}\big((1 - y)^{-\alpha} - 1\big)^k y^{-\alpha-1}\,\mathrm{d}y$$
$$\sim \alpha(n-1)\overline{F}(x)\int_{\epsilon}^{1/(2(n-1))}\big((1 - y)^{-\alpha} - 1\big)^k y^{-\alpha-1}\,\mathrm{d}y\,.$$

Here $(\,\cdot\,)^k$ is of order y^k at $y = 0$. Since ϵ was arbitrary, it follows that if $\mathbb{E}[Y^2] < \infty$

$$\mathbb{E}[X_{2,2}^k] \sim f(x)^k\mathbb{E}[S_{n-1}^k]\,.$$

If $\alpha = 2$ and $\mathbb{E}[Y^2] = \infty$, (Bingham et al., 1989, Proposition 1.5.9a) yields

$$\mathbb{E}[X_{2,2}^2] \sim 2f(x)^2(n-1)\int_0^x y\overline{F}(y)\,\mathrm{d}y\,.$$

If $\alpha < 2$,

$$\mathbb{E}[X_{2,2}^2] \sim \alpha(n-1)\overline{F}(x)^3\int_0^{1/(2(n-1))}\big((1 - y)^{-\alpha} - 1\big)^2 y^{-\alpha-1}\,\mathrm{d}y.$$

Further if $\mathbb{E}[Y_1] < \infty$, then $\mathbb{E}[X_{2,2}] \sim f(x)\mathbb{E}[S_{n-1}]$. If $\alpha = 1$ and $\mathbb{E}[Y_1] = \infty$, then

$$\mathbb{E}[X_{2,2}] \sim (n-1)f(x)\int_0^x \overline{F}(y)\,\mathrm{d}y\,.$$

If $\alpha < 1$, then

$$\mathbb{E}[X_{2,2}] \sim \alpha(n-1)\int_0^{1/(2(n-1))} \big((1-y)^{-\alpha}-1\big)\,y^{-\alpha-1}\,\mathrm{d}y\,\overline{F}(x)^2.$$

Now recall the formula

$$\mathrm{Var}[X+Y] \;=\; \mathrm{Var}[X]+\mathrm{Var}[Y]+2\,\mathrm{Cov}[X,Y]\,.$$

Using

$$\mathbb{E}[X_1X_2]=\mathbb{E}[X_{2,1}X_{2,2}]=0\,,\quad \mathbb{E}[X_2]\sim\overline{F}(x)\,,\quad \mathbb{E}[X_{2,1}]\mathbb{E}[X_{2,2}]=\mathrm{o}(\mathbb{E}[X_{2,2}^2])\,,$$

we get

$$\frac{1}{n^2}\,\mathrm{Var}[Z_{\mathrm{AK}}]=\mathrm{Var}[X_1]+\mathrm{Var}[X_{2,1}]+\mathrm{Var}[X_{2,2}]-2\overline{F}(x)\mathbb{E}[X_1].$$

Collecting all terms we get that for $\alpha \geq 2$ $\mathrm{Var}[X_{2,2}]$ dominates the other terms asymptotically and for $\alpha < 2$ we get

$$\begin{aligned}
\frac{1}{n^2}\,\mathrm{Var}\;Z_{\mathrm{AK}} \sim (n-1)\bigg(&\alpha\int_{1/(2(n-1))}^{1/2} y^{-\alpha-1}(1-y)^{-2\alpha}dy+\tfrac{1}{3}2^{3\alpha}+(2(n-1))^{\alpha}\\
&-2\alpha\int_{1/(2(n-1))}^{1/2} y^{-\alpha-1}(1-y)^{-\alpha}dy-2^{2\alpha}\\
&+\alpha\int_0^{1/(2(n-1))}\big((1-y)^{-\alpha}-1\big)^2y^{-\alpha-1}dy\bigg)\overline{F}(x)^3\\
=(n-1)\bigg(&\alpha\int_{1/(2(n-1))}^{1/2} y^{-\alpha-1}\left((1-y)^{-2\alpha}-2(1-y)^{-\alpha}\right)dy+\tfrac{1}{3}2^{3\alpha}\\
&+(2(n-1))^{\alpha}-2^{2\alpha}+\alpha\int_0^{1/(2(n-1))}\big((1-y)^{-\alpha}-1\big)^2y^{-\alpha-1}dy\bigg)\overline{F}(x)^3\\
=(n-1)\bigg(&-\alpha\int_{1/(2(n-1))}^{1/2} y^{-\alpha-1}dy+\tfrac{1}{3}2^{3\alpha}+(2(n-1))^{\alpha}-2^{2\alpha}\\
&+\alpha\int_0^{1/2}\big((1-y)^{-\alpha}-1\big)^2y^{-\alpha-1}dy\bigg)\overline{F}(x)^3\\
=(n-1)\bigg(&2^{\alpha}+\frac{1}{3}2^{3\alpha}-2^{2\alpha}+\alpha\int_0^{1/2}\big((1-y)^{-\alpha}-1\big)^2y^{-\alpha-1}dy\bigg)\overline{F}(x)^3.
\end{aligned}$$

□

Remark 4. In the case $\alpha = 2$, $\mathbb{E}[Y^2] = \infty$, the above discussion leads to asking for more explicit growth rates of

$$I(x) \;=\; \int_0^x y\overline{F}(y)\,\mathrm{d}y \;=\; \int_0^x \frac{L(y)}{y}\,\mathrm{d}y\,.$$

An obvious conjecture is that the order is $L(x)\log x$. But this turns out not be true for all $L(x)$. For a more detailed analysis, assume that $L(x) \sim c\exp\{\int_a^x e(t)/t\,\mathrm{d}t\}$ where $e(t)$ is differentiable and $a, c > 0$. Further we assume that $I(x) \to \infty$. If there exists a $\kappa \neq -1$ with

$$\lim_{t\to\infty} \frac{e'(t)t}{e(t)^2} \to -\kappa$$

then a partial integration argument yields $I(x) \sim L(x)/(\kappa+1)e(x)$. For example, if $L(x) \sim (\log x)^\beta$ we can choose $e(x) = \beta/\log x$ and $\kappa = 1/\beta$, so that hence $I(x) \sim L(x)\log(x)/\beta + 1$. Another example is $L(x) \sim \exp\{(\log\log x)^\gamma\}$ where $\gamma > 0$. Then $e(x) = \gamma(\log\log x)^{\gamma-1}/\log x$, $\kappa = 0$ and $I(x) \sim L(x\log x/\gamma(\log\log x)^{\gamma-1}$.

On the other hand if $e(x)\log x \to \tau$ and $\tau \neq -1$, it can be shown that $I(x) \sim L(x)\log(x)/\tau + 1$. Again for $L(x) \sim (\log x)^\beta$ we can choose $e(x) = \beta/\log x$ and $\tau = \beta$ and find as above that $I(x) \sim L(x\log x/\gamma(\log\log x)^{\gamma-1}$. □

3 An improved estimator for $\alpha > 2$

In the case of finite second moment we get that the error is basically given by the variance of the term $X_{2,2}$ in (2.1). Since for large values of x $X_{2,2}$ is close to $S_{n-1}f(x)$ a natural idea is to use this approximation as a control variate which results in the estimator

$$Z \;=\; Z_{\mathrm{AK}} + n\left(\mathbb{E}S_{n-1} - S_{n-1}\right) f(x). \tag{3.1}$$

The next theorem shows that this estimator is indeed an improvement over Z_{AK}.

Theorem 2. *Assume* $f'(x) = -\alpha(\alpha-1)L(x)/x^{\alpha+2}$. *If* $\alpha > 4$ *or, more generally,* $\mathbb{E}[Y^4] < \infty$ *then*

$$\mathrm{Var}\ Z \sim \tfrac{1}{4}n^2\,\mathrm{Var}[S_{n-1}^2]f'(x)^2.$$

If $\alpha = 4$ *and* $\mathbb{E}[Y^4] = \infty$ *then*

$$\mathrm{Var}\ Z \sim n^2(n-1)f'(x)^2\int_0^x y^3\overline{F}(y)dy.$$

If $2 < \alpha < 4$ *then* $\mathrm{Var}\ Z \sim n^2(n-1)k_\alpha\overline{F}(x)^3$ *where*

$$k_\alpha = \alpha\int_0^\infty \big(((1-z)\vee z)^{-\alpha} - 1 - \alpha z\big)^2 z^{-\alpha-1}\mathrm{d}z\,.$$

Proof. The proof is a variation of the proof of Theorem 1. Define $A_{n,x} = \{M_{n-1} \le x/2(n-1)\}$. At first note that $\mathrm{Var}\ Z = n^2\,\mathrm{Var}\ Z_1$ where $Z_1 = \overline{F}(x - S_{n-1}\vee M_{n-1}) - S_{n-1}f(x)$. We will use

$$Z_1 = Z_1 I\left(A_{n,x}^c\right) + Z_1 I\left(A_{n,x}\right) = X_1 + X_2.$$

Similar to the proof of Theorem 1 we get that

$$\begin{aligned}\mathbb{E}X_1^k &= \mathbb{E}\big[\big(\overline{F}(x-S_{n-1}\vee M_{n-1})-S_{n-1}f(x)\big)^k;A_{n,x}^c\big]\\ &\sim \mathbb{E}\big[\big(\overline{F}(x-M_{n-1}\vee M_{n-1})-M_{n-1}f(x)\big)^k;A_{n,x}^c\big]\\ &= x(n-1)\int_{\frac{1}{2(n-1)}}^{\infty}\big[\overline{F}((x-xz)\vee xz)-xzf(x)\big]^k f(xz)F(xz)^{n-2}\mathrm{d}z\\ &\sim \alpha(n-1)\overline{F}(x)^{k+1}\int_{\frac{1}{2(n-1)}}^{\infty}\big[((1-z)\vee z)^{-\alpha}-\alpha z\big]^k z^{-\alpha-1}\mathrm{d}z.\end{aligned}$$

If $M_{n-1}\le x/2(n-1)$, then $S_{n-1}\le x/2$ and so

$$\begin{aligned}X_2 &= \overline{F}(x)-\overline{F}(x)I\left(A_{n,x}^c\right)+\big(\overline{F}(x-S_{n-1})-F(x)-S_{n-1}f(x)\big)\,I\left(A_{n,x}\right)\\ &= \overline{F}(x)-I(A_{n,x}^c)\overline{F}(x)-\tfrac{1}{2}I(A_{n,x})S_{n-1}^2f'(\Xi)=\overline{F}(x)-X_{2,1}+X_{2,2}.\end{aligned}$$

where $x-S_{n-1}\le\Xi\le x$.
Proceeding as in the proof of Theorem 1 we get that

$$\mathbb{E}[X_{2,1}^k] \;=\; \overline{F}(x)^k\mathbb{P}(A_{n,x}^c) \;\sim\; (n-1)2^{\alpha}(n-1)^{\alpha}\overline{F}(x)^{k+1}.$$

To evaluate $\mathbb{E}[X_{2,2}^k]$, we split the expectation into $S_{n-1}\le\epsilon x$ and $S_{n-1}\ge\epsilon x$, for some $1/(2(n-1))>\epsilon>0$. When $S_{n-1}\le\epsilon x$, we have $-f'(x)\le -f'(\Xi)\le -f'\big((1-\epsilon)x\big)$. This and monotone convergence gives

$$\lim_{\epsilon\to 0}\lim_{x\to\infty}\frac{\mathbb{E}\big[S_{n-1}^{2k}f'(\Xi)^k;\,S_{n-1}\le\epsilon x\big]}{f'(x)^k\mathbb{E}\big[S_{n-1}^{2k};\,S_{n-1}\le\epsilon x\big]}=1\,.$$

Further by (Bingham et al., 1989, Proposition 1.5.8 and 1.5.9a) and partial integration

$$\begin{aligned}\mathbb{E}\big[S_{n-1}^{2k};\,S_{n-1}\le\epsilon x\big] &= 2k\int_0^{\epsilon x}y^{2k-1}\overline{F}_{S_{n-1}}(y)\mathrm{d}y-(\epsilon x)^{2k}\overline{F}_{S_{n-1}}(\epsilon x)\\ &\sim\begin{cases}\mathbb{E}S_{n-1}^{2k} & \mathbb{E}Y_1^{2k}<\infty,\\ 2k\displaystyle\int_0^{x}y^{2k-1}\overline{F}_{S_{n-1}}(y)\mathrm{d}y & \alpha=2k,\\ \epsilon^{2k-\alpha}\dfrac{\alpha}{2k-\alpha}x^{2k}\overline{F}_{S_{n-1}}(x) & \alpha<2k.\end{cases}\end{aligned}$$

If $\epsilon x\le S_{n-1}\le x/2$ then it holds uniformly in S_{n-1} that

$$\frac{-S_{n-1}^2f'(\Xi)}{2\overline{F}(x)}\sim\frac{\overline{F}(x-S_{n-1})-\overline{F}(x)-S_{n-1}f(x)}{\overline{F}(x)}\sim(1-S_{n-1}/x)^{-\alpha}-1-\alpha\frac{S_{n-1}}{x}\,,$$

and hence

$$\begin{aligned}&\tfrac{1}{2}\mathbb{E}\big[S_{n-1}^{2k}(-f(\Xi))^k;\,S_{n-1}>\epsilon x,\,M_{n-1}\le x/(2(n-1))\big]\\ &\quad\sim\overline{F}(x)^k\mathbb{E}\big[\big(1-S_{n-1}/x\big)^{-\alpha}-1-\alpha S_{n-1}/x\big)^k;\,S_{n-1}>\epsilon x,\,M_{n-1}\le x/(2(n-1))\big]\\ &\quad\sim\overline{F}(x)^k\mathbb{E}\big[\big((1-M_{n-1}/x)^{-\alpha}-1-\alpha M_{n-1}/x\big)^k;\,\epsilon x<M_{n-1}\le x/(2(n-1))\big]\\ &\quad\sim\alpha(n-1)\overline{F}(x)^{k+1}\int_{\epsilon}^{1/(2(n-1))}\big((1-y)^{-\alpha}-1-\alpha y\big)^k\,y^{-\alpha-1}\,\mathrm{d}y\,.\end{aligned}$$

Here $(\,\cdot\,)^k$ is of order y^{2k} at $y = 0$. Since ϵ was arbitrary, it follows that if $\mathbb{E}[Y^4] < \infty$

$$\mathbb{E}[X_{2,2}^k] \sim \frac{1}{2^k}(-f'(x))^k \mathbb{E}[S_{n-1}^{2k}]\,.$$

If $\alpha = 4$ and $\mathbb{E}[Y^4] = \infty$, (Bingham et al., 1989, Proposition 1.5.9a) yields

$$\mathbb{E}[X_{2,2}^2] \sim (n-1)f'(x)^2 \int_0^x y^3 \overline{F}(y)\,\mathrm{d}y\,.$$

If $\alpha < 4$,

$$\mathbb{E}[X_{2,2}^2] \sim \alpha(n-1)\overline{F}(x)^3 \int_0^{1/(2(n-1))} \left((1-y)^{-\alpha} - 1 - \alpha y\right)^2 y^{-\alpha-1}\,\mathrm{d}y.$$

Further since $\mathbb{E}[Y_1^2] < \infty$, $\mathbb{E}[X_{2,2}] \sim -\frac{1}{2}f'(x)\mathbb{E}[S_{n-1}^2]$. Finally with the same collecting of terms as in the proof of Theorem 1 the Theorem follows. □

4 An improved estimator for $1 < \alpha < 2$

If $\mathbb{E}Y_1^2 = \infty$ then the estimator (3.1) has $\mathrm{Var}(Z) = \infty$ so this estimator is no improvement for $\alpha < 2$. Nevertheless it is an interesting question if we can improve on Z_{AK} also in this case. In this section we will consider the case of $1 < \alpha < 2$ where we have a finite mean but an infinite second moment. We will denote with $Y_{(1)} \le \cdots \le Y_{(n)}$ the order statistic of $Y_1, \ldots, Y_n$, with $M_k = \max_{1\le i\le k} Y_i$ and with $S_k = \sum_{i=1}^k Y_i$. At first note that

$$\mathbb{P}(S_n > x) = \mathbb{P}\Big(S_n > x, Y_{(n-1)} \le \frac{x}{2(n-1)}\Big) \tag{4.1}$$

$$+\,\mathbb{P}\Big(S_n > x, Y_{(n-1)} > \frac{x}{2(n-1)}\Big). \tag{4.2}$$

We will use separate estimators for (4.1) and (4.2). By conditioning on the two largest elements and using symmetry

$$\mathbb{P}\left(S_n > x, Y_{(n-1)} > \frac{x}{2(n-1)}\right) = n(n-1)\mathbb{P}\left(Y_1 > \frac{x}{2(n-1)}\right)^2 p_{n.x}$$

where

$$p_{n,x} = \mathbb{P}\Big(S_n > x, \min_{i\in\{n-1,n\}} Y_i \ge M_{n-2} \;\Big|\; \min_{i\in\{n-1,n\}} Y_i > \frac{x}{2(n-1)}\Big).$$

Denote with $Z^c(x)$ the crude Monte Carlo estimator of $p_{n,x}$, i.e. we first simulate Y_{n-1}, Y_n conditioned to exceed $x/2(n-1)$ then the rest normal and use the estimator

$$Z^c(x) = I\left(S_n > x\right) I\left(\min(Y_{n-1}, Y_n) \ge M_{n-2}\right)$$

[note that conditioned r.v. generation is easy whenever inversion is available, cf. (Asmussen and Glynn, 2007, p. 39)]. The estimator

$$Z^b(x) = n(n-1)\overline{F}(x/(2(n-1)))^2 Z^c(x)$$

is an unbiased estimator for (4.2) and as $x \to \infty$

$$\begin{aligned}\operatorname{Var} Z^b(x) &= \operatorname{Var}\big[Z^s(x)\big]\,(n(n-1))^2\overline{F}(x/2(n-1))^4 \\ &\lesssim 2^{4\alpha} n^2 (n-1)^{4\alpha+2}\overline{F}(x)^4.\end{aligned}$$

So we have to find an estimator for (4.1). By conditioning on the largest element and using symmetry we get that

$$\mathbb{P}\Big(S_n > x, X_{(n-1)} \le \frac{x}{2(n-1)}\Big) = n\mathbb{P}\Big(S_n > x, M_{n-1} \le \frac{x}{2(n-1)}\Big).$$

Let $\widetilde{f}$ be an importance sampling density of the form $\widetilde{L}(x)/x^{\widetilde{\alpha}}$ with $\widetilde{\alpha} < 2\alpha - 2$ and $\widetilde{\mathbb{P}}, \widetilde{\mathbb{E}}$ the corresponding probability- and expectation operators. We will assume that $\widetilde{f}$ is bonded away from zero on finite intervals. We now combine the estimator which conditions on S_{n-1} with importance sampling and a control variate to get the estimator

$$Z^s(x) \;=\; \big(\overline{F}(x - S_{n-1}) - \overline{F}(x)\big) L I_{n,x} \;+\; \overline{F}(x)\mathbb{P}(I_{n,x})$$

where $L = \prod_{i=1}^{n-1} f(Y_i)/\widetilde{f}(Y_i)$ is the likelihood ratio and $I_{n,x}$ is the indicator of the event $M_{n-1} \le x/2(n-1)$.

Since $f(x) \sim \sup_{z \ge x} f(z)$ (e.g. (Bingham et al., 1989, Theorem 1.5.3)), we get by Taylor expansion

$$\widetilde{\mathbb{E}}\big[\big(\overline{F}(x - S_{n-1}) - \overline{F}(x)\big) L I_{n,x}\big] \;\lesssim f(x/2)\widetilde{\mathbb{E}}[S_{n-1} L I_{n,x}] \;\le\; f(x/2)\mathbb{E}S_{n-1}$$

and with the same arguments

$$\widetilde{\mathbb{E}}\big[\big(\overline{F}(x - S_{n-1}) - \overline{F}(x)\big) L I_{n,x}\big]^2 \;\lesssim f(x/2)^2\widetilde{\mathbb{E}}[S_{n-1}^2 L^2]\,.$$

We have that

$$\begin{aligned}\widetilde{\mathbb{E}}[S_{n-1}^2 L^2] = (n-1)\widetilde{\mathbb{E}}\Big[Y_1^2 \frac{f(Y_1)^2}{\widetilde{f}(Y_1)^2} \prod_{i=2}^{n} \frac{f(Y_i)^2}{\widetilde{f}(Y_i)^2}\Big] \\ + (n-1)(n-2)\widetilde{\mathbb{E}}\Big[Y_1 \frac{f(Y_1)^2}{\widetilde{f}(Y_1)^2} Y_2 \frac{f(Y_2)^2}{\widetilde{f}(Y_2)^2} \prod_{i=3}^{n} \frac{f(Y_i)^2}{\widetilde{f}(Y_i)^2}\Big].\end{aligned}$$

Since for $k = 0, 1, 2$

$$\widetilde{\mathbb{E}}\Big[Y^k \frac{f(Y_i)^2}{\widetilde{f}(Y_i)^2}\Big] \;=\; \int y^k \frac{f(y)^2}{\widetilde{f}(y)}\,\mathrm{d}y \;<\; \infty$$

(the integrand is regularly varying with index $k - 2\alpha - 2 + \widetilde{\alpha} + 1 < -1$), it follows by independence that $\widetilde{\mathbb{E}}[S_{n-1}^2 L^2] < \infty$ and hence $\widetilde{\operatorname{Var}}(Z^s(x)) \;=\; \mathrm{O}\big(f(x)^2\big)$. Thus we have shown

Theorem 3. *Assume $f(x) = \alpha L(x)/x^{\alpha+1}$. If $1 < \alpha < 2$ then the estimator $Z(x) = Z^b(x) + nZ^s(x)$ is an unbiased estimator for $\mathbb{P}(S_n > x)$ with*

$$\operatorname{Var}(Z(x)) \;=\; \mathrm{O}\big(f(x)^2\big)$$

Remark 5. Note that for a fixed number of simulations the confidence interval (of fixed level) for the estimator $Z(x)$ has a length of order $f(x)$. This is the same order as the error of the asymptotic approximation $\mathbb{P}(S_n > x) - n\overline{F}(x)$.

5 The weibull case

We finally give a brief survey of our results for the Weibull case $\overline{F}(x) = \mathrm{e}^{-x^\beta}$ with $0 < \beta < 1$ (related distributions, say modified by a power, are easily included, but for simplicity, we refrain from this). We refer to Asmussen and Kortschak (2012) for a more complete treatment. The density is $f(x) = \beta x^{\beta-1}\mathrm{e}^{-x^\beta}$ and $f'(x) = -p(x)\overline{F}(x)$ where $p(x) = \beta^2 x^{2(\beta-1)} + \beta(1-\beta)x^{\beta-2}$.

Theorem 4. *If* $0 < \beta < \log(3/2)/\log(2)$*, then the Asmussen-Kroese estimator's variance is asymptotically given by*

$$\mathrm{Var}(Z_{\mathrm{AK}}) \sim n^2\,\mathrm{Var}(S_{n-1})f(x)^2$$

Note that $\log(3/2)/\log(2) \approx 0.585$ is also found to be critical in Asmussen and Kroese (2006) as the threshold for logarithmic efficiency to hold.

As in Section 3 we can use the estimator defined in (3.1) to improve on Z_{AK}.

Theorem 5. *Assume that* $0 < \beta < \log(3/2)/\log(2) \approx 0.585$*. Then the estimator* Z *in* (3.1) *has vanishing relative error. More precisely,*

$$\mathrm{Var}(Z(x)) \sim \frac{n^2}{4}\,\mathrm{Var}(S_{n-1}^2)f'(x)^2.$$

The estimator Z in (3.1) has the form $Z_{\mathrm{AK}} + \alpha(S_{n-1} - \mathbb{E}S_{n-1})$, so it is a control variate estimator, using S_{n-1} as control for Z_{AK}. It is natural to ask whether the $\alpha = -nf(x)$ at least asymptotically coincides with the optimal

$$\alpha^* = -\mathrm{Cov}(Z_{\mathrm{AK}}, S_{n-1})/\mathrm{Var}(S_{n-1})$$

(cf. (Asmussen and Glynn, 2007, V.2)). The following lemma shows that this is the case:

Lemma 6. $\mathrm{Cov}(Z_{\mathrm{AK}}, S_{n-1}) = n\,\mathrm{Var}(S_{n-1})f(x) + \mathrm{o}\big(f(x)\big)$.

Acknowledgments

We thank Sandeep Juneja for discussions stimulating the present research. The second author was supported by the the MIRACCLE-GICC project and the Chaire d'excellence "Generali – Actuariat responsable: gestion des risques naturels et changements climatiques".

References

Asmussen, S., and H. Albrecher. 2010. *Ruin probabilities*. Second ed. Advanced Series on Statistical Science & Applied Probability, 14. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ.

Asmussen, S., and K. Binswanger. 1997. "Simulation of ruin probabilities for subexponential claims". *ASTIN Bulletin* 27 (2): 297–318.

Asmussen, S., K. Binswanger, and B. Højgaard. 2000. "Rare events simulation for heavy-tailed distributions". *Bernoulli* 6 (2): 303–322.

Asmussen, S., and P. W. Glynn. 2007. *Stochastic simulation: algorithms and analysis*, Volume 57 of *Stochastic Modelling and Applied Probability.* New York: Springer.

Asmussen, S., and D. Kortschak. 2012. "Error rates and improved algorithms for rare event simulation with heavy Weibull tails". Technical report, Working paper.

Asmussen, S., and D. P. Kroese. 2006. "Improved algorithms for rare event simulation with heavy tails". *Adv. in Appl. Probab.* 38 (2): 545–558.

Baltrūnas, A., and E. Omey. 1998. "The rate of convergence for subexponential distributions". *Liet. Mat. Rink.* 38 (1): 1–18.

Barbe, P., and W. P. McCormick. 2009. *Asymptotic expansions for infinite weighted convolutions of heavy tail distributions and applications.* American Mathematical Society.

Bingham, N. H., C. M. Goldie, and J. L. Teugels. 1989. *Regular variation*, Volume 27 of *Encyclopedia of Mathematics and its Applications.* Cambridge: Cambridge University Press.

Blanchet, J., and P. Glynn. 2008. "Efficient rare-event simulation for the maximum of heavy-tailed random walks". *Ann. Appl. Probab.* 18 (4): 1351–1378.

Bucklew, J. A. 1990. *Large deviation techniques in decision, simulation, and estimation.* Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. New York: John Wiley & Sons Inc. A Wiley-Interscience Publication.

Dupuis, P., K. Leder, and H. Wang. 2007. "Importance sampling for sums of random variables with regularly varying tails". *ACM TOMACS* 17 (3).

Embrechts, P., C. Klüppelberg, and T. Mikosch. 1997. *Modelling extremal events, for insurance and finance*, Volume 33 of *Applications of Mathematics (New York).* Berlin: Springer-Verlag.

Foss, S., D. Korshunov, and S. Zachary. 2011. *An introduction to heavy-tailed and subexponential distributions.* Springer Series in Operations Research and Financial Engineering. New York: Springer.

Hartinger, J., and D. Kortschak. 2009. "On the efficiency of the Asmussen-Kroese-estimators and its application to stop-loss transforms". *Blätter DGVFM* 30 (2): 363–377.

Heidelberger, P. 1995. "Fast simulation of rare events in queueing and reliability models". *ACM Trans. Model. Comput. Simul.* 5 (1): 43–85.

Juneja, S., and P. Shahabuddin. 2006. "Rare event simulation techniques: An introduction and recent advances". In *Handbook on Simulation (S. Henderson & B. Nelson, eds.)*, 291–350. North-Holland.

Omey, E., and E. Willekens. 1987. "Second-order behaviour of distributions subordinate to a distribution with finite mean". *Comm. Statist. Stochastic Models* 3 (3): 311–342.