J. Appl. Prob. **49,** 1188–1193 (2012)
Printed in England

IMPROVING THE ASMUSSEN–KROESE-TYPE SIMULATION ESTIMATORS

SAMIM GHAMAMI,* ** AND
SHELDON M. ROSS,* *** *University of Southern California*

Abstract

The Asmussen–Kroese Monte Carlo estimators of $\mathrm{P}(S_n > u)$ and $\mathrm{P}(S_N > u)$ are known to work well in rare event settings, where S_n is the sum of independent, identically distributed heavy-tailed random variables $X_1, \ldots, X_n$ and N is a nonnegative, integer-valued random variable independent of the X_i. In this paper we show how to improve the Asmussen–Kroese estimators of both probabilities when the X_i are nonnegative. We also apply our ideas to estimate the quantity $\mathrm{E}[(S_N - u)^+]$.

Keywords: Heavy-tailed random variable; rare event; efficient Monte Carlo estimation; variance reduction; conditioning; stratification; control variate; stop-loss transform

2010 Mathematics Subject Classification: Primary 65C05
Secondary 91G20

1. Introduction

We consider the well-known problem of finding efficient Monte Carlo estimations of $\mathrm{P}(S_n > u)$ and $\mathrm{P}(S_N > u)$, where $S_n = X_1 + \cdots + X_n$, the X_i are nonnegative, independent, identically distributed heavy-tailed random variables, and N is a nonnegative, integer-valued random variable that is independent of the X_i. The estimation of these probabilities has applications in insurance risk, financial mathematics, and queueing theory, and their efficient Monte Carlo estimation has been the subject of extensive research in the last decade (see [1], [4], and the references therein).

In this paper we are interested in improving the Asmussen–Kroese estimators of $\mathrm{P}(S_n > u)$ and $\mathrm{P}(S_N > u)$ introduced in [1]. Let F denote the common distribution of X_i, and let $M_n = \max(X_1, \ldots, X_n)$. The Asmussen–Kroese estimators of $\mathrm{P}(S_n > u)$ and $\mathrm{P}(S_N > u)$ are respectively given by

$$Z_1 \equiv n\,\mathrm{P}(S_n > u,\ X_n = M_n \mid X_1, \ldots, X_{n-1}) = n\bar{F}(M_{n-1} \vee (u - S_{n-1}))$$

and

$$Z_2 \equiv N\,\mathrm{P}(S_N > u,\ X_n = M_n \mid N, X_1, \ldots, X_{N-1}) = N\bar{F}(M_{N-1} \vee (u - S_{N-1})),$$

where $a \vee b = \max(a, b)$ and $\bar{F}(x) = 1 - F(x)$. These estimators appear to perform very well for subexponential distributions in the rare event setting (see [2] for different classes of heavy-tailed distributions, including subexponential distributions).

Received 23 February 2012; revision received 28 May 2012.
Accepted by Onno Boxma, Coordinating Editor.
* Postal address: Department of Industrial and Systems Engineering, University of Southern California, Los Angeles, CA 90089, USA.
This material is based upon work supported by the US Army Research Laboratory and the US Army Research Office under grant number W911NF-11-1-0115.
** Email address: ghamami@usc.edu
*** Email address: smross@usc.edu

When estimating $\mathrm{P}(S_N > u)$, Asmussen and Kroese [1] suggested further variance reduction by either stratifying over values of N or by using N as a control variate, and they concluded that both variance reduction methods perform roughly the same.

Whereas the Asmussen–Kroese estimators do not require that the X_i be nonnegative, we will make that assumption in this paper. In Section 2.1 we show how to use the nonnegativity to improve their estimator of $\mathrm{P}(S_n > u)$. The improved estimator has a smaller variance and requires less data simulation. In Section 2.2 we present an improved estimator of $\mathrm{P}(S_N > u)$. Whereas the numerical work we have carried out indicated only a small improvement in variance when using the proposed estimator of $\mathrm{P}(S_n > u)$ as opposed to Z_1, the improvement was much greater when using our proposed estimator of $\mathrm{P}(S_N > u)$ versus Z_2. One part of the reason for the greater improvement in this latter case is that the Asmussen–Kroese estimator of $\mathrm{P}(S_n > u)$ can be a poor estimator when n is large. To understand why, note that the minimum of $\max(M_{n-1}, u - S_{n-1})$ occurs when $X_1 = X_2 = \cdots = X_{n-1} = u/n$. Consequently,

$$\max_{X_1,\dots,X_{n-1}} Z_1 = n\bar{F}\left(\frac{u}{n}\right).$$

Thus, for large n, the estimator can be large, and need not even be preferable to the raw simulation estimator $\mathbf{1}_{\{S_n>u\}}$. (The above also gives some intuition about why the Asmussen–Kroese estimator can be so good, namely the right-hand side of the above equation is often small, and so its variance is too.)

Hartinger and Kortschak [4] used the ideas underlying the Asmussen–Kroese estimators to estimate the stop-loss transform identities, $\mathrm{E}[((S_N - u)^+)^k]$, $k = 1, 2$ (stop-loss transforms have applications in the pricing of stop-loss reinsurance contracts and the valuation of catastrophe risk bonds). In Section 2.3 we give an improved estimator of $\mathrm{E}[(S_N - u)^+]$.

2. Efficient estimators of $\mathrm{P}(S_n > u)$ and $\mathrm{P}(S_N > u)$

2.1. An improved variation of the Asmussen–Kroese estimator

Our improved variation of the Asmussen–Kroese estimator of $\mathrm{P}(S_n > u)$ is derived by further conditioning on the first time that the current maximum plus the current sum exceed u. More specifically, let

$$R = \min(n-1, \min(j \ge 1 \colon M_j + S_j > u)).$$

Conditioning on $R, X_1, \dots, X_R$ gives the estimator

$$\mathcal{E} = n\,\mathrm{P}(S_n > u,\ X_n = M_n \mid R, X_1, \dots, X_R) = \begin{cases} \dfrac{n}{n-R}(1 - F^{n-R}(M_R)) & \text{if } R < n-1, \\ n\bar{F}(M_{n-1} \vee (u - S_{n-1})) & \text{if } R = n-1. \end{cases}$$

To derive the last equality above when $R < n-1$, note that, for the event $\{S_n > u,\ X_n = M_n\}$ to occur conditional on $R, X_1, \dots, X_R$, the maximum of $X_{R+1}, \dots, X_n$ should be larger than M_R and X_n must be the maximum of $X_{R+1}, \dots, X_n$. Because $\mathcal{E} = \mathrm{E}[Z_1 \mid R, X_1, \dots, X_R]$, it has a smaller variance than Z_1 as well as requires less data simulation.

2.1.1. *Numerical examples.* In Table 1 we present numerical results for the (standard) Weibull distribution with $\bar{F}(x) = \mathrm{e}^{-x^{\beta}}$ based on 10^5 simulation runs (performed in MATLAB®). In the last column of Table 1 we record the estimates of $\mathrm{P}(S_n > u)$. The values $\beta = 0.25, 0.5, 0.75$ have also been considered in the numerical examples of [1]. Values of u have been chosen so

https://doi.org/10.1239/jap/1354716666 Published online by Cambridge University Press

TABLE 1: Numerical results for the Weibull distribution with $\bar{F}(x) = \mathrm{e}^{-x^{\beta}}$.

β	n	u	$\operatorname{var}(Z_1)$	$\operatorname{var}(\mathcal{E})$	$\mathrm{P}(S_n > u)$
0.50	10	32.609	0.0121	0.0119	0.1466
0.50	10	72.583	1.26×10^{-4}	1.24×10^{-4}	86×10^{-4}
0.75	20	28.104	0.0803	0.0790	0.2490
0.75	20	43.85	0.0013	0.0012	0.0108
0.25	5	234.210	8.44×10^{-4}	8.34×10^{-4}	0.1099
0.25	10	7.1962×10^3	5.7×10^{-8}	5.6×10^{-8}	0.0011

that the order of $\mathrm{P}(S_n > u)$ varies from 10^{-1} to 10^{-4}. As can be seen from the results, the variance improvement in this case is very marginal.

Remark. Asmussen and Kroese [1, Theorem 3.1] showed that, for the Weibull case, their estimator, Z_1, is *polynomial time* for $\beta < 0.585$. That is, as a function of u, $\operatorname{var}(Z_1)/\operatorname{E}[Z_1]^{2-\varepsilon}$ is bounded in u for any $\varepsilon > 0$ when $\beta < 0.585$. Also, their numerical results, both for Z_1 and Z_2, showed performance degradation for β outside of this critical range.

2.2. Efficient estimator of $\mathrm{P}(S_N > u)$

As mentioned in Section 1, Asmussen and Kroese [1] separately combined their estimator Z_2 of $\mathrm{P}(S_N > u)$ with stratification and with using N as a control, and concluded that the empirical performance of both methods is similar. However, it is not difficult to see that the two variance reduction techniques can be combined. Using the independence of N and the X_i, the stratification identity can be written as

$$\mathrm{P}(S_N > u) = \sum_{n=1}^{l} \mathrm{P}(S_n > u) P_n + \mathrm{P}(S_N > u \mid N > l)\tilde{P}_l,$$

where $P_n = \mathrm{P}(N = n)$ and l is chosen so that $\tilde{P}_l = \mathrm{P}(N > l)$ is small. Given that the P_n and $\tilde{P}_l$ are analytically computable, recall that standard stratification returns an estimate of $\mathrm{P}(S_N > u)$ by estimating the preceding conditional probabilities. The Monte Carlo estimate of the conditional probability above associated with the last truncated stratum, $\mathrm{P}(S_N > u \mid N > l)$, will be based on the sampled value of N given that it exceeds l. We suggest that the Monte Carlo estimate of $\mathrm{P}(S_N > u \mid N > l)$ be improved by using N, conditional on it exceeding l, as a control variate. So, stratification and the control variate method can be combined and need not be considered separately. Moreover, instead of standard stratification we suggest using 'single simulation run stratification' which returns an estimate of $\mathrm{P}(S_N > u)$ after each run. More specifically, at the beginning of a given run, we generate N, given that it exceeds l. Then, we estimate all the conditional probabilities based on the Monte Carlo realizations of X_i, $i = 1, \dots, l, n_l$, where n_l denotes a sampled value of N, given that it exceeds l. (Although, in contrast to standard stratification, our Monte Carlo estimates of the quantities $\mathrm{P}(S_n > u)$, $n \geq 1$, obtained in a run are positively correlated, we feel that the savings in time in using the same data to estimate them more than compensates.) The final improvement we present is based on our earlier observation that the variance of Z_1 becomes large for large values of n. Thus, we suggest deviating from the Asmussen–Kroese estimator and instead estimating $\mathrm{P}(S_n > u)$ by the estimator $\mathrm{P}(S_n > u \mid X_1, \dots, X_{n-1}) = \bar{F}(u - S_{n-1})$ whenever n

TABLE 2: Numerical results for Weibull random variables with tail $\bar{F}(x) = \mathrm{e}^{-x^\beta}$ and geometric N with parameter p.

β	p	u	$\mathrm{var}(Z_2)$	$\mathrm{var}(Z_2^c)$	$\mathrm{var}(\tilde{\mathcal{E}})$	$\mathrm{P}(S_N > u)$
0.50	0.25	32.533	0.0083	0.0046	2.17×10^{-4}	0.0316
0.50	0.10	130.1325	0.0017	0.0014	1.3×10^{-5}	0.0039
0.75	0.50	3.04	0.0646	0.0216	0.0014	0.1353
0.75	0.15	63.361	4.626×10^{-4}	3.564×10^{-4}	4.3×10^{-7}	5.234×10^{-4}
0.25	0.10	409.99	0.0397	0.0144	0.00145	0.1337
0.25	0.30	10233	1.68×10^{-8}	1.07×10^{-8}	9.5×10^{-11}	1.027×10^{-4}

is such that $n\bar{F}(u/n) > 1$. Hence, with $\tilde{n} = \min(n : n\bar{F}(u/n) > 1)$, we propose the following estimator of $\mathrm{P}(S_N > u)$:

$$\tilde{\mathcal{E}} = \begin{cases} \displaystyle\sum_{n=1}^{\tilde{n}-1} \mathcal{E}_n P_n + \sum_{n=\tilde{n}}^{l} \bar{F}(u - S_{n-1}) P_n \\ \qquad + (\bar{F}(u - S_{n_l-1}) + c_1(N_l - \mathrm{E}[N_l]))\tilde{P}_l & \text{if } \tilde{n} \le l, \\ \displaystyle\sum_{n=1}^{l} \mathcal{E}_n P_n + (\bar{F}(u - S_{n_l-1}) + c_1(N_l - \mathrm{E}[N_l]))\tilde{P}_l & \text{if } l < \tilde{n} \le n_l, \\ \displaystyle\sum_{n=1}^{l} \mathcal{E}_n P_n + (\mathcal{E}_{n_l} + c_2(N_l - \mathrm{E}[N_l]))\tilde{P}_l & \text{if } n_l < \tilde{n}. \end{cases}$$

Here $\mathcal{E}_n$ refers to our improved variation of the Asmussen–Kroese estimator when $N = n$, and c_k, $k = 1, 2$, are coefficients of the control variate which can be specified optimally and estimated based on the simulation (see [3] or [5]).

2.2.1. *Numerical examples.* In Table 2 we compare the empirical performance of our proposed estimator of $\mathrm{P}(S_N > u)$, $\tilde{\mathcal{E}}$, with that of the Asmussen–Kroese estimator,

$$Z_2 = N\bar{F}(M_{N-1} \vee (u - S_{N-1})),$$

and the control variate method,

$$Z_2^c = N\bar{F}(M_{N-1} \vee (u - S_{N-1})) + c(N - \mathrm{E}[N]),$$

where c is the simulation-based estimate of the optimal coefficient of the control variable. In the numerical examples below the X_i are (standard) Weibull random variables with tail distribution $\bar{F}(x) = \mathrm{e}^{-x^\beta}$, and, similar to the numerical studies of [1], N is a geometric random variable with 'success probability' p. In the last column of Table 2 we record the estimates of $\mathrm{P}(S_N > u)$ based on $\tilde{\mathcal{E}}$ and 10^5 simulation runs.

For this numerical example, the computing time of our proposed estimator, $\tilde{\mathcal{E}}$, is on average 4 to 4.5 times the computing time of the estimator Z_2^c, which is the Asmussen–Kroese estimator combined with a control variate. However, the results of Table 2 indicate that substantial variance reduction is gained using our estimator.

Remark. Most of the variance reduction gained in the above situation is due to the single-run stratification with control idea, with a much smaller amount due to the changed estimator when $n > \tilde{n}$. (This is not too surprising because the improved estimators occurred when n was large and so these estimators were given the small weight $\mathrm{P}(N = n)$.) For instance, if we had not changed the estimator then the variances of our estimator in the first five of the six cases considered would be 2.37×10^{-4}, 1.39×10^{-5}, 0.0017, 4.6×10^{-7}, and 0.0017.

2.3. Efficient estimator of $\mathrm{E}[(S_N - u)^+]$

Hartinger and Kortschak [4] gave an Asmussen–Kroese-type estimator of $\theta = \mathrm{E}[(S_n - u)^+]$. Using the fact that $\theta = n\,\mathrm{E}[(S_n - u)^+ \mathbf{1}_{\{X_n = M_n\}}]$, they proposed estimating θ by

$$n\,\mathrm{E}[(S_n - u)^+ \mathbf{1}_{\{X_n = M_n\}} \mid \boldsymbol{X}_{n-1}] = n(\mathrm{E}[X_n \mid X_n > a] + (S_{n-1} - u))\bar{F}(a),$$

where $\boldsymbol{X}_{n-1} \equiv (X_1, \dots, X_{n-1})$ and $a = \max(M_{n-1}, u - S_{n-1})$.

To obtain an improved estimator, as before, let $R = \min(n-1, \min(j \geq 1 \colon M_j + S_j > u))$. With $R^* = \{R, X_1, \dots, X_R\}$, the improved estimator is $n\,\mathrm{E}[(S_n - u)^+ \mathbf{1}_{\{X_n = M_n\}} \mid R^*]$. When $R = n - 1$, the two estimators are equal. Now consider simulation runs with $R < n - 1$. Let $A \equiv \{$at least one of $X_{R+1}, \dots, X_n$ exceeds $M_R\}$ and $M \equiv \max(X_{R+1}, \dots, X_n)$. With the notation E_{R^*} and P_{R^*} for the conditional expectation and conditional probability given R^*, we have

$$\begin{aligned}
\mathrm{E}_{R^*}[(S_n - u)^+ \mathbf{1}_{\{X_n = M_n\}}] &= \mathrm{E}_{R^*}\biggl[\biggl(\sum_{i=R+1}^{n} X_i + S_R - u\biggr)^+ \mathbf{1}_{\{X_n = M_n\}} \biggm| A\biggr]\mathrm{P}_{R^*}(A) \\
&= \biggl(\mathrm{E}_{R^*}\biggl[\sum_{i=R+1}^{n} X_i \mathbf{1}_{\{X_n = M\}} \biggm| A\biggr] + \frac{S_R - u}{n - R}\biggr)\mathrm{P}_{R^*}(A) \\
&= \frac{1}{n - R}\biggl(\mathrm{E}_{R^*}\biggl[\sum_{i=R+1}^{n} X_i \biggm| A\biggr] + S_R - u\biggr)\mathrm{P}_{R^*}(A), \qquad (1)
\end{aligned}$$

where $\mathrm{P}_{R^*}(A) = 1 - F^{n-R}(M_R)$. To calculate $\mathrm{E}_{R^*}[\sum_{i=R+1}^{n} X_i \mid A]\,\mathrm{P}_{R^*}(A)$, let A^{c} denote the complement of the event A, and use the fact that

$$\mathrm{E}_{R^*}\biggl[\sum_{i=R+1}^{n} X_i\biggr] = \mathrm{E}_{R^*}\biggl[\sum_{i=R+1}^{n} X_i \biggm| A\biggr]\mathrm{P}_{R^*}(A) + \mathrm{E}_{R^*}\biggl[\sum_{i=R+1}^{n} X_i \biggm| A^{\mathrm{c}}\biggr]\mathrm{P}_{R^*}(A^{\mathrm{c}}),$$

which yields

$$\begin{aligned}
\mathrm{E}_{R^*}\biggl[\sum_{i=R+1}^{n} X_i \biggm| A\biggr]\mathrm{P}_{R^*}(A) &= \mathrm{E}_{R^*}\biggl[\sum_{i=R+1}^{n} X_i\biggr] - \mathrm{E}_{R^*}\biggl[\sum_{i=R+1}^{n} X_i \biggm| A^{\mathrm{c}}\biggr]\mathrm{P}_{R^*}(A^{\mathrm{c}}) \\
&= (n - R)(\mathrm{E}[X] - \mathrm{E}[X \mid X < M_R]\,\mathrm{P}_{R^*}(A^{\mathrm{c}})). \qquad (2)
\end{aligned}$$

Using (2) in (1), our proposed estimator of θ becomes

$$\hat{\theta} = \begin{cases} n(\mathrm{E}[X] - F^{n-R}(M_R)\,\mathrm{E}[X \mid X < M_R] + \dfrac{S_R - u}{n - R}(1 - F^{n-R}(M_R)) & \text{if } R < n - 1, \\ n(\mathrm{E}[X_n \mid X_n > a] + (S_{n-1} - u))\bar{F}(a) & \text{if } R = n - 1, \end{cases}$$

where $a = \max(M_{n-1}, u - S_{n-1})$.

To estimate $\mathrm{E}[(S_N - u)^+]$, we suggest using the stratification identity

$$\mathrm{E}[(S_N - u)^+] = \sum_{n=1}^{l} \mathrm{E}[(S_N - u)^+ \mid N = n] P_n + \mathrm{E}[(S_N - u)^+ \mid N > l] \tilde{P}_l$$

and a single-run stratification estimator with a control variate as in Section 2.2. Our approach could also be used to estimate $\mathrm{E}[((S_N - u)^+)^2]$, the second stop-loss transform identity considered in [4].

References

[1] Asmussen, S. and Kroese, D. P. (2006). Improved algorithms for rare event simulation with heavy tails. *Adv. Appl. Prob.* **38,** 545–558.

[2] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events*. Springer, Berlin.

[3] Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering*. Springer, New York.

[4] Hartinger, J. and Kortschak, D. (2009). On the efficiency of the Asmussen-Kroese-estimator and its application to stop-loss transforms. *Blätter DGVFM* **30,** 363–377.

[5] Ross, S. M. (2006). *Simulation*, 4th edn. Academic Press, San Diego, CA.