COLEMAN FUNG
RISK MANAGEMENT
RESEARCH CENTER

Working Paper # 2012 - 07

Improving the Asmussen- Kroese Type Simulation Estimators

Samim Ghamami,University of Southern California,Los Angeles
Sheldon M. Ross,University of Southern California,Los Angeles

May 25, 2012

**University of California
Berkeley**

Improving the Asmussen-Kroese Type Simulation Estimators *

Samim Ghamami† and Sheldon M. Ross‡

May 25, 2012

Abstract

Asmussen-Kroese [1] Monte Carlo estimators of $P(S_n > u)$ and $P(S_N > u)$ are known to work well in rare event settings when S_n is the sum of n i.i.d. heavy-tailed random variables, and N is a non-negative 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 non-negative. We also apply our ideas to estimate the quantity $E[(S_N - u)^+]$.

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

2010 Mathematics Subject Classification: Primary 65C05 ; Secondary 91G20

1 Introduction

We consider the well-known problem of efficient Monte Carlo estimation of $P(S_n > u)$ and $P(S_N > u)$, where $S_n = X_1 + ... + X_n$, X_i's are non-negative i.i.d. heavy-tailed random variables, and N is a non-negative integer-valued random variable that is independent of X_i's. The estimation of these probabilities has applications in insurance risk, financial mathematics, and queueing theory, and their efficient Monte Carlo estimation has been subject of extensive research in the last decade, (see [1], [4], and the references there).

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

$$Z_1 \equiv nP(S_n > u, X_n = M_n | X_1, ..., X_{n-1}) = n\bar{F}(M_{n-1} \vee (u - S_{n-1})),$$

*Department of Industrial and Systems Engineering, University of Southern California, Los Angeles, CA 90089.
This material is based upon work supported by the U. S. Army Research Laboratory and the U. S. Army Research Office under grant number W911NF-11-1-0115.
†Email: ghamami@usc.edu
‡Email: smross@usc.edu

$$Z_2 \equiv NP(S_N > u, X_n = M_n | N, X_1, ..., X_{N-1}) = N\bar{F}(M_{N-1} \vee (u - S_{N-1})),$$

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.

When estimating $P(S_N > u)$, 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 non-negativity to improve their estimator of $P(S_n > u)$. The improved estimator has a smaller variance and requires less data simulation. In section 2.2 we present an improved estimator of $P(S_N > u)$. Whereas the numerical work we have carried out indicated only a small improvement in variance when using the proposed estimator of $P(S_n > u)$ as opposed to Z_1, the improvement was much greater when using our proposed estimator of $P(S_N > u)$ versus Z_2. One part of the reason for the greater improvement in this latter case is that the Asmussen-Kroese estimator of $P(S_n > u)$ 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 $X_1 = X_2 = ... = X_{n-1} = u/n$. Consequently,

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

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 preceding also gives some intuition about why the Asmussen-Kroese estimator can be so good, namely the right hand side of the preceding equation is often small, and so its variance is too.)

The paper [4] uses the ideas underlying Asmussen-Kroese estimators to estimate the stop-loss transform identities, $E[((S_N - u)^+)^k]$, $k = 1, 2$, (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 $E[(S_N - u)^+]$.

2 Efficient Estimators of $P(S_n > u)$ and $P(S_N > u)$

2.1 An Improved Variation of the Asmussen-Kroese Estimator

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

$$R = \min\{n - 1, \min\{j \geq 1 : M_j + S_j > u\}\}. \quad (1)$$

Conditioning on $R, X_1, ..., X_R$ gives the estimator

$$\mathcal{E} = nP(S_n > u, X_n = M_n | R, X_1, ..., X_R) = \begin{cases} \frac{n}{n-R}(1 - F^{n-R}(M_R)) & \text{if } R < n-1 \\ \\ n\bar{F}(M_{n-1} \vee (u - S_{n-1})) & \text{if } R = n-1 \end{cases} \quad (2)$$

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

Numerical Examples Table 1 gives numerical results for the (standard) Weibull distribution with $\bar{F}(x) = e^{-x^\beta}$ based on 10^5 simulation runs (and using MATLAB). The last column is the estimate of $P(S_n > u)$. $\beta = .25, .5, .75$ have also been considered in the numerical examples of [1]. Values of u have been chosen so that the order of $P(S_n > u)$ varies from 10^{-1} to 10^{-4}.

β	n	u	var(Z_1)	var$(\mathcal{E})$	$P(S_n > u)$
.5	10	32.609	.0121	.0119	.1466
.5	10	72.583	1.26×10^{-4}	1.24×10^{-4}	86×10^{-4}
.75	20	28.104	.0803	.0790	.2490
.75	20	43.85	.0013	.0012	.0108
.25	5	234.210	8.44×10^{-4}	8.34×10^{-4}	.1099
.25	10	7.1962×10^3	5.7×10^{-8}	5.6×10^{-8}	.0011

Table1: Numerical results for Weibull with $\bar{F}(x) = e^{-x^\beta}$

As can be seen, the variance improvement in this case is very marginal.

Remark Asmussen and Kroese [1] (Theorem 3.1) show that for the Weibull case their estimator, Z_1, is *polynomial time* for $\beta < .585$. That is, as a function of u, var$(Z_1)/E[Z_1]^{2-\epsilon}$ is bounded in u for any $\epsilon > 0$ when $\beta < .585$. Also, their numerical results, both for Z_1 and Z_2, show performance degradation for β outside of this critical range.

2.2 Efficient Estimator of $P(S_N > u)$

As mentioned in section 1, Asmussen and Kroese [1] separately combine their estimator Z_2 of $P(S_N > u)$ with stratification and with using N as a control and conclude 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

$$P(S_N > u) = \sum_{n=1}^{l} P(S_n > u)P_n + P(S_N > u|N > l)\tilde{P}_l, \tag{3}$$

where $P_n = P(N = n)$, and l is chosen so that $\tilde{P}_l = P(N > l)$ is small. Given that P_n's and $\tilde{P}_l$ are analytically computable, recall that standard stratification returns an estimate of $P(S_N > u)$ by estimating the preceding conditional probabilities. The Monte Carlo estimate of the conditional probability above associated with the last truncated stratum, $P(S_N > u|N > l)$, will be based on the sampled value of N given that it exceeds l. We suggest that the Monte Carlo estimate of $P(S_N > u|N > l)$ be improved by using N, conditional on it exceeding l, as a control variate. So, stratification and 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 $P(S_N > u)$ 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 $X_i, i = 1, \ldots, l, n_l$, 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 $P(S_n > u), n \geq 1$ 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.) Our final improvement idea 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 $P(S_n > u)$ by the estimator, $P(S_n > u|X_1, \ldots, X_{n-1}) = \bar{F}(u - S_{n-1})$ whenever n 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 $P(S_N > u)$:

$$\tilde{\mathcal{E}} = \begin{cases} \sum_{n=1}^{\tilde{n}-1} \mathcal{E}_n P_n + \sum_{n=\tilde{n}}^{l} \bar{F}(u - S_{n-1})P_n + (\bar{F}(u - S_{n_l-1}) + c_1(N_l - E[N_l]))\tilde{P}_l & \text{if } \tilde{n} \leq l \\ \sum_{n=1}^{l} \mathcal{E}_n P_n + (\bar{F}(u - S_{n_l-1}) + c_1(N_l - E[N_l]))\tilde{P}_l & \text{if } l < \tilde{n} \leq n_l \\ \sum_{n=1}^{l} \mathcal{E}_n P_n + (\mathcal{E}_{n_l} + c_2(N_l - E[N_l]))\tilde{P}_l & \text{if } n_l < \tilde{n} \end{cases} \tag{4}$$

where $\mathcal{E}_n$ 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 [5] or [3]).

Numerical Examples Table 2 compares the empirical performance of our proposed estimator of $P(S_N > u)$, $\tilde{\mathcal{E}}$, with that of Asmussen-Kroese estimator, $Z_2 = N\bar{F}(M_{N-1} \vee (u - S_{N-1}))$, and also its combination with the control variate method, $Z_2^c = N\bar{F}(M_{N-1} \vee (u - S_{N-1})) + c(N - E[N])$, where c is the simulation based estimate of the optimal coefficient of the control variable. In the numerical examples below X_i's are (standard) Weibull random variables with tail distribution $\bar{F}(x) = e^{-x^\beta}$, and, similar to the numerical studies of [1], N is a geometric random variable with "success probability" p. The last column is the estimate of $P(S_N > u)$ based on $\tilde{\mathcal{E}}$ and 10^5 simulation runs.

β	p	u	$\text{var}(Z_2)$	$\text{var}(Z_2^c)$	$\text{var}(\tilde{\mathcal{E}})$	$P(S_N > u)$
.5	.25	32.533	.0083	.0046	2.17×10^{-4}	.0316
.5	.1	130.1325	.0017	.0014	1.3×10^{-5}	.0039
.75	.5	3.04	.0646	.0216	.0014	.1353
.75	.15	63.361	4.626×10^{-4}	3.564×10^{-4}	4.3×10^{-7}	5.234×10^{-4}
.25	.1	409.99	.0397	.0144	.00145	.1337
.25	.3	10233	1.68×10^{-8}	1.07×10^{-8}	9.5×10^{-11}	1.027×10^{-4}

Table2: Numerical results for Weibull F with tail e^{-x^β} and geometric N with parameter p

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 Asmussen-Kroese estimator combined with a control variate. However, Table 2 indicates that substantial variance reduction is gained using our estimator.

Remark: Most of the variance reduction of our method in the preceding 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 $P(N = n)$.) 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}, .0017, 4.6×10^{-7}, .0017.

2.3 Efficient Estimator of $E[(S_N - u)^+]$

The paper [4] gives an Asmussen-Kroese type estimator of $\theta = E[(S_n - u)^+]$. Using that $\theta = nE[(S_n - u)^+ \mathbf{1}_{\{X_n = M_n\}}]$, [4] proposes to estimate θ by

$$nE[(S_n - u)^+ \mathbf{1}_{\{X_n = M_n\}} | \mathbf{X}_{n-1}] = n \left(E[X_n | X_n > a] + (S_{n-1} - u) \right) \bar{F}(a), \tag{5}$$

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

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

$$E_{R^*}[(S_n - u)^+ \mathbf{1}_{\{X_n = M_n\}}] = E_{R^*}[(\sum_{i=R+1}^{n} X_i + S_R - u)^+ \mathbf{1}_{\{X_n = M_n\}} | A] P_{R^*}(A)$$

$$
\begin{aligned}
&= (E_{R^*}[\sum_{i=R+1}^{n} X_i \mathbf{1}_{\{X_n=M\}}|A] + \frac{S_R-u}{n-R}) P_{R^*}(A) \\
&= \frac{1}{n-R}\left(E_{R^*}[\sum_{i=R+1}^{n} X_i|A] + S_R - u\right) P_{R^*}(A), \qquad (6)
\end{aligned}
$$

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

$$
E_{R^*}[\sum_{i=R+1}^{n} X_i] = E_{R^*}[\sum_{i=R+1}^{n} X_i|A]P_{R^*}(A) + E_{R^*}[\sum_{i=R+1}^{n} X_i|A^c]P_{R^*}(A^c),
$$

which yields that

$$
\begin{aligned}
E_{R^*}[\sum_{i=R+1}^{n} X_i|A]P_{R^*}(A) &= E_{R^*}[\sum_{i=R+1}^{n} X_i] - E_{R^*}[\sum_{i=R+1}^{n} X_i|A^c]P_{R^*}(A^c) \\
&= (n-R)\left(E[X] - E[X|X<M_R]P_{R^*}(A^c)\right). \qquad (7)
\end{aligned}
$$

Using (7) in (6), our proposed estimator of θ becomes,

$$
\hat{\theta} = \begin{cases} n(E[X] - F^{n-R}(M_R)E[X|X<M_R] + \frac{(S_R-u)}{n-R}(1-F^{n-R}(M_R)) & \text{if } R < n-1 \\ n\left(E[X_n|X_n>a] + (S_{n-1}-u)\right)\bar{F}(a) & \text{if } R = n-1 \end{cases} \qquad (8)
$$

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

To estimate $E[(S_N-u)^+]$ we suggest using the following stratification identity

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

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

References

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

[2] Embrechts, P., Klüppelberg, C., Mikosch, T., "Modelling Extremal Events for Finance and Insurance", Springer, (1997)

[3] Glasserman, P., “Monte Carlo Methods in Financial Engineering”, Springer, (2004)

[4] Hartinger, J., and Kortschak, D., “On the efficiency of the Asmussen-Kroese-estimator and its application to stop-loss transforms”, *Blätter der DGVFM*, **30**(2), 363-377, (2009)

[5] Ross, S. M., “Simulation”, fourth ed., Academic Press, (2006)