Computational methods for sums of random variables

Patrick J. Laub

Abstract

A typical problem in the field of rare-event estimation is to find the probability $\mathbb{P}(S > \gamma)$ where $S := X_1 + \cdots + X_d$ for a fixed $d \in \mathbb{N}_+$ and where the $\gamma \in \mathbb{R}$ is large or increasing. In applications we often wish to understand the behaviour of a combination of random factors. Hence the random variable S is ubiquitous in real-world modeling problems. It can model, for example, aggregate risk or portfolio value for holding d risky assets [4, 6], the aggregate losses for d insurance policy claims [1, 3], and the combined signal interference from d wireless transmission sources [2]. Probabilities of this form are used to understand how a system would behave under extreme scenarios such as a market crash, a power surge, or a natural disaster. One is typically interested in, not just the quantity $\mathbb{P}(S > \gamma)$, but the behaviour of the summands when the extreme event $\{S > \gamma\}$ occurs.

This probability is available in a closed form for only a few basic cases, when the density of S (which is a d-fold convolution) has a known form; see [5]. For example, when the summands are independent and identically distributed (iid), then it is sometimes simple to calculate (for exponential, gamma, normal, binomial, geometric, or negative binomial summands) and sometimes intractable (for lognormal, Weibull, Cauchy, Laplace, Beta, or Chi-squared summands). However, requiring the assumption of independence (let alone iid-ness) of the summands is a stifling restriction when modeling real-world events; a notorious example would be the partial blame of the 2008–09 global financial crisis on mathematicians' inappropriate use of a simplistic dependence model (the Gaussian copula) [7].

This thesis outlines methods for approximating quantities related to sums of random variables. Two chapters consider the use of *orthogonal polynomial expansions* in approximating probability density functions; one focuses on sums of correlated lognormal random variables, and the other on random sums which are used in insurance. We also introduce an *importance sampling estimator* for the survival function of a sum distribution which uses knowledge of the asymptotic form of the sum. We also give the results of an *asymptotic analysis* of the Laplace transform for the sum of lognormal random variables. A related problem, of estimating the probability of the maximum of a random vector exceeding a large threshold, is also considered.

References

[1] Søren Asmussen and Hansjörg Albrecher. *Ruin probabilities*, volume 14 of *Advanced Series on Statistical Science and Applied Probability*. World Scientific Publishing Co Pte Ltd, River Edge, NJ, 2 edition, 2010. Advanced series on statistical science & applied probability; v. 14.

[2] Carlo Fischione, Fabio Graziosi, and Fortunato Santucci. Approximation for a sum of on-off lognormal processes with wireless applications. *IEEE Transactions on Communications*, 55(10):1984–1993, 2007.

[3] Stuart A Klugman, Harry H Panjer, and Gordon E Willmot. *Loss models: from data to decisions*, volume 715. John Wiley & Sons, 2012.

[4] Alexander J McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, 2nd edition, 2015.

[5] Saralees Nadarajah. A review of results on sums of random variables. *Acta Applicandae Mathematicae*, 103(2):131–140, 2008.

[6] Ludger Rüschendorf. *Mathematical Risk Analysis*. Springer, 2013.

[7] F. Salmon. Recipe for disaster: The formula that killed Wall Street, 2009. Online 23/02/2009.