Exponential Family Techniques for the Lognormal Left Tail

Søren Asmussen[1], Jens Ledet Jensen[1] and Leonardo Rojas-Nandayapa[2]

[1]Department of Mathematics, Aarhus University
[2]School of Mathematics and Physics, University of Queensland

Abstract

Let X be lognormal(μ, σ^2) with density $f(x)$, let $\theta > 0$ and define $L(\theta) = \mathbb{E}e^{-\theta X}$. We study properties of the exponentially tilted density (Esscher transform) $f_\theta(x) = e^{-\theta x} f(x)/L(\theta)$, in particular its moments, its asymptotic form as $\theta \to \infty$ and asymptotics for the saddlepoint $\theta(x)$ determined by $\mathbb{E}[Xe^{-\theta X}]/L(\theta) = x$. The asymptotic formulas involve the Lambert W function. The established relations are used to provide two different numerical methods for evaluating the left tail probability of the sum of lognormals $S_n = X_1 + \cdots + X_n$: a saddlepoint approximation and an exponential tilting importance sampling estimator. For the latter we demonstrate logarithmic efficiency. Numerical examples for the cdf $F_n(x)$ and the pdf $f_n(x)$ of S_n are given in a range of values of σ^2, n, x motivated by portfolio Value-at-Risk calculations.

Keywords: Cramér function, Esscher transform, exponential change of measure, importance sampling, Lambert W function, Laplace method, Laplace transform, Lognormal distribution, outage probability, rare–event simulation, saddlepoint approximation, VaR.

MSC: 60E05, 60E10, 90-04

1 Introduction

The lognormal distribution arises in a wide variety of disciplines such as engineering, economics, insurance, finance, and across the sciences (Aitchison and Brown, 1957; Crow and Shimizu, 1988; Dufresne, 2009; Johnson et al., 1994; Limpert et al., 2001). Therefore, it is natural that sums of lognormals come up in a number of contexts. A basic example in finance is the Black–Scholes model, which assumes that security prices can be modeled as independent lognormals, and hence the value of a portfolio with n securities can be conveniently modeled as a sum of lognormals. Another example occurs in the valuation of arithmetic Asian options where the payoff depends on the finite sum of correlated lognormals (Dufresne, 2004; Milevsky and Posner, 1998). In insurance, individual claim sizes are often modeled as independent lognormals, so the total claim amount in a certain period is a random sum of lognormals (Thorin and Wikstad, 1977). A further example occurs in telecommunications, where the inverse of the signal-to-noise ratio (a measure of performance in wireless systems) can be modeled as a sum of iid lognormals (Gubner, 2006).

However, the distribution of a sum of n lognormals S_n is not available in explicit form, and its numerical approximation is considered to be a challenging problem. In consequence, a number of methods for its evaluation have been developed over several decades, but these can rarely deliver arbitrary precisions on the whole support of the distribution, particularly in the tails. The latter case is of key relevance in certain applications which often require evaluation of tail probabilities at a very high precision.

When considering lognormals sums, the literature has so far concentrated on the right tail (with the exception of the recent paper by Gulisashvili and Tankov, 2014). In this paper, our object of study is rather the left tail and certain mathematical problems that naturally come up in this context. To be precise, let Y_i be normal(μ_i, σ_i^2) (we don't at the moment specify the dependence structure), $X_i = \mathrm{e}^{Y_i}$ and $S_n = X_1 + \cdots + X_n$. We then want to compute $\mathbb{P}(S_n \leq z)$ in situations where this probability is small.

A main application is VaR calculations in the finance industry that have become mandatory following the treatise Basel II (2004). The VaR is an important measure of market risk defined as an appropriate $(1 - \alpha)$-quantile of the distribution of the loss, and Basel II asks for calculation of the VaR for α as small as 0.03% (the values depend on the type of business). For a careful explanation of these matters and references, see for instance Duelman (2010), Ch. 1 of McNeil et al. (2015) and Embrechts et al. (2014). For a specific example relevant for this paper, let a portfolio be based on n assets with upcoming prices $X_1, \ldots, X_n$. In case of a short position, the potential loss then corresponds to a large value of S_n so that for the VaR one needs the right tails. With a long position, a loss is caused by a small value so that one gets into calculations in the left tail, precisely the problem of this paper. A second example is in wireless systems where the outage capacity is defined as the probability that the inverse of the signal-to-noise ratio operates in a range below certain thresholds (cf. Slimane (2001); Navidpour et al. (2004)). For relevance of calculations of small values in the left tail, see Barakat (1976); Beaulieu et al. (1995); Beaulieu and Xie (2004).

The problem of approximating the distribution of a sum of iid lognormals has a long history. The classical approach is to approximate the distribution of this sum with another lognormal distribution. This goes back at least to Fenton (1960) and it is nowadays known as the *Fenton–Wilkinson method* as according to Marlow (1967) this approximation was already used by Wilkinson in 1934. However, the Fenton–Wilkinson method, being a central limit type result, can deliver rather inaccurate approximations of the distribution of the lognormal sum when the number of summands is rather small, or when the dispersion parameter is too high—in particular in the tail regions. Another topic which has been much studied recently is approximations and simulation algorithms for right tail probabilities $\mathbb{P}(S_n \geq y)$ under heavy-tailed assumptions and allowing for dependence, see in particular Asmussen et al. (2011); Asmussen and Rojas-Nandayapa (2008); Blanchet and Rojas-

Nandayapa (2011); Foss and Richards (2010); Mitra and Resnick (2009). For further literature surveys, see Gulisashvili and Tankov (2014).

Our approach is to use saddlepoint approximations and a closely related simulation algorithm based on the same exponential change of measure. This requires iid assumptions, so we assume that $\mu_i \equiv \mu$, $\sigma_i^2 \equiv \sigma^2$. Since μ is just a scaling factor, we will assume $\mu = 0$. The saddlepoint approximation occurs in various (closely related) forms, but all involve the function $\kappa(\theta) = \log L(\theta)$ and its two first derivatives $\kappa'(\theta)$ and $\kappa''(\theta)$, where $L(\theta)$ is the Laplace transform

$$L(\theta) \;=\; \mathbb{E}\mathrm{e}^{-\theta X_i} \;=\; \int_0^\infty \mathrm{e}^{-\theta x} f(x)\,\mathrm{d}x \quad \text{with} \quad f(x) \;=\; \frac{1}{x\sigma\sqrt{2\pi}}\mathrm{e}^{-\log^2 x/2\sigma^2}$$

(note that since the right tail of the lognormal distribution is heavy, these quantities are only defined for $\theta \geq 0$). Formally, $\kappa(\theta)$ is the cumulant transform of the random variable $-X_i$. Define the exponentially tilted density $f_\theta(x)$ (Esscher transform) by

$$f_\theta(x) = \mathrm{e}^{-\theta x} f(x)/L(\theta), \quad x > 0, \tag{1}$$

and let its corresponding cdf be F_θ with expectation operator $\mathbb{E}_\theta$. Then

$$\kappa'(\theta) = \mathbb{E}_\theta[-X_i] = -\mathbb{E}_\theta X_i\,, \quad \kappa''(\theta) = \mathrm{Var}_\theta\, X_i \tag{2}$$

and one can connect the distribution of S_n (corresponding to $\theta = 0$) to the $\mathbb{P}_\theta$-distribution by means of the likelihood ratio identity

$$\mathbb{P}(S_n \in A) \;=\; \mathbb{E}_\theta\big[\exp\{\theta S_n + n\kappa(\theta)\};\, S_n \in A\big]\,.$$

The construction of the saddlepoint approximation requires the *saddlepoint* $\theta(x)$ being the solution of the equation

$$\kappa'\big(\theta(x)\big) = -\mathbb{E}_{\theta(x)}[X_i] = -x, \tag{3}$$

and the tilted measure $\mathbb{P}_\theta$ with $\theta = \theta(x)$. This choice of θ means that $\mathbb{E}_\theta S_n = nx$ so that the $\mathbb{P}_\theta$-distribution is centered around nx and central limit expansions apply. For a short exposition of the implementation of this program in its simplest form, see p. 355, Asmussen (2003).

The application of saddlepoint approximations to the lognormal left tail seems to have appeared for the first time in the third author's 2008 Dissertation (Rojas-Nandayapa, 2008), but in a more incomplete and preliminary form than the one presented here. A first difficulty is that $L(\theta)$ is not explicitly available for the lognormal distribution. However, approximations with error rates were recently given in the companion paper Asmussen et al. (2014b) (see also Laub et al., 2016)). The result is in terms of the Lambert W function $\mathcal{W}(a)$ (Corless et al., 1996), defined as the unique solution of $\mathcal{W}(a)\mathrm{e}^{\mathcal{W}(a)} = a$ for $a > 0$. The expression for the Laplace transform $L(\theta)$ from Asmussen et al. (2014b) is the case $k = 0$ in Proposition 1 below, the general case being the expectation $\mathbb{E}[X^k \mathrm{e}^{-\theta X}]$. Note that the Lambert W function is convenient for numerical computations since it is implemented in many software packages.

The paper is organized as follows. In Section 2, we study the exponential family $(F_\theta)_{\theta\geq 0}$. We give approximations to the derivatives of the Laplace transform, an approximation to the saddlepoint $\theta(x)$, and discuss various approximations to the tilted density f_θ. The first important application of our results, namely the saddlepoint approximation for $\mathbb{P}(S_n \leq nx)$, is given in Section 3. The second is a Monte Carlo estimator for $\mathbb{P}(S_n \leq nx)$ given in Section 4.2. It follows a classical route (VI.2, Asmussen and Glynn, 2007) by attempting importance sampling where the importance distribution is $F_{\theta(x)}$. The implementation faces the difficulty that neither $\theta(x)$ nor $L\big(\theta(x)\big)$ are explicitly known. The importance sampling algorithm requires simulation from F_θ, and we suggest an acceptance-rejection (A-R) for this with a Gamma proposal. The Appendix contains a proof that the importance sampling proposed in Section 4.2 has a certain asymptotical efficiency property.

2 The exponential family generated by the lognormal distribution

We let F be the cdf of X and adopt the notation $X \sim \mathrm{LN}\,(0,\sigma^2)$. For convenience, we write f_n and F_n for the pdf and cdf of S_n, respectively.

The exponential tilting scheme in the introduction is often also referred to as the *Esscher transform*. Note that since $\kappa(\theta)$ is well-defined for all $\theta > 0$, the saddlepoint $\theta(x)$ exists for all $0 < x \leq \mathbb{E}X$ (the relevant case for our left tail problem) and large deviation results can be used. The latter are based on the *Legendre–Fenchel transform* defined as the convex conjugate of $\kappa(\theta)$.

We first consider ways of evaluating and approximating derivatives of the Laplace transform given through

$$L_k(\theta) = \mathbb{E}[X^k \mathrm{e}^{-\theta X}] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-h_k(y)}\mathrm{d}y, \quad \text{with } h_k(y) = -ky + \theta \mathrm{e}^y + \frac{y^2}{2\sigma^2}. \tag{4}$$

Define $w_k(\theta) = \mathcal{W}(\theta\sigma^2\mathrm{e}^{k\sigma^2})$, $\sigma_k(\theta)^2 = \sigma^2/(1+w_k(\theta))$ as well as:

$$\begin{aligned}
L_a(k,\theta) &= \frac{\sigma_k(\theta)}{\sigma} \exp\Big\{ -\frac{1}{2\sigma^2} w_k(\theta)^2 - \frac{1}{\sigma^2} w_k(\theta) + \tfrac{1}{2}k^2\sigma^2 \Big\}, & (5)\\
H_k(z;\theta) &= \big(e^{z\sigma_k(\theta)} - 1 - z\sigma_k(\theta)\big) w_k(\theta)/\sigma^2 - z^2\sigma_k(\theta)^2/(2\sigma^2), \\
I_k(\theta) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\{-H_k(z;\theta)\}\mathrm{d}z, & (6)\\
J_k(\theta) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\big\{ -H_0(z;\theta) + k\sigma_0(\theta)z - \tfrac{1}{2}\sigma_0(\theta)^2k^2 \big\}\,\mathrm{d}z.
\end{aligned}$$

The following proposition extends Proposition 2.1 of Asmussen et al. (2014b). To understand the orders of the different terms one should keep in mind that $w_k(\theta)$ is asymptotically of order $\log(\theta)$ for $\theta \to \infty$. We also use the fact that $w_k(0) = 0$.

Proposition 1. *Let $X \sim \mathrm{LN}\,(0,\sigma^2)$, $k \in \mathbb{N}^+$ and $\theta \geq 0$. Then*

$$L_k(\theta) = L_a(k,\theta) I_k(\theta) \ \textit{and}\ L_k(\theta) = L_a(0,\theta) \exp\big\{ -kw_0(\theta) + \tfrac{1}{2}\sigma_0(\theta)^2k^2 \big\} J_k(\theta). \tag{7}$$

Moreover, $I_k(0) = J_k(0) = 1$ and with $\lambda_k(m;\theta) = w_k(\theta)\sigma_k(\theta)^m/(\sigma^2 m!)$ we have for $\theta \to \infty$ the expansions

$$I_k(\theta) = 1 - 3\lambda_k(4;\theta) + \frac{15}{2}\lambda_k(3;\theta)^2 + \mathcal{O}\big(\sigma_k(\theta)^4\big),$$
$$J_k(\theta) = 1 - 3(1+4k)\lambda_0(4;\theta) + \frac{15}{2}\lambda_0(3;\theta)^2 + \mathcal{O}\big(\sigma_0(\theta)^4\big).$$

This proposition's proof employs Laplace's approximation as used in the proof of Proposition 2.1 in the companion paper Asmussen et al. (2014b). We give here only a brief sketch of the proof. For the first result involving $L_a(k,\theta)$, the function $h_k(y)$ is expanded around its minimizer $y_k(\theta)$ given as the solution to $\theta \mathrm{e}^y + y/\sigma^2 - k = 0$, that is, $y_k(\theta) = k\sigma^2 - w_k(\theta)$. Then the exponential part of $L_a(k,\theta)$ is simply $-h_k\big(y_k(\theta)\big)$, and the exponential part of the integrand in $I_k(\theta)$ is $-\big\{h_k\big(y_k(\theta) + \sigma_k(\theta)z\big) - h_k\big(y_k(\theta)\big)\big\}$. Expansion of the latter gives $-\frac{1}{2}z^2 - \lambda_k(3;\theta)z^3 - \lambda_k(4;\theta)z^4 + \mathcal{O}\big(\sigma_k(\theta)^3|z|^5\big)$, and expanding the exponential of the last three terms gives the result in the proposition for $I_k(\theta)$. For the alternative formula with $J_k(\theta)$ we expand $h_k(y)$ around $y_0(\theta)$. Then the exponential part of $L_a(0,\theta)$ together with $-kw_0(\theta)$ is simply $-h(y_0(\theta))$, and the exponential part of the integrand in $J_k(\theta)$ is $-\big\{h_k\big(y_0(\theta)+\sigma_0(\theta)z\big) - h_k\big(y_0(\theta)\big)\big\} - \sigma_0(\theta)^2k^2/2$. Expanding the latter we get $-\frac{1}{2}\big(z - k\sigma_0(\theta)\big)^2 - \lambda_0(3;\theta)z^3 - \lambda_0(4;\theta)z^4 + \mathcal{O}\big(\sigma_0(\theta)^3|z|^5\big)$, which leads to the result in the proposition.

The results of Proposition 1 immediately lead to an approximation of the mean and variance of the exponentially tilted measure. These are denoted by $\mathbb{E}_\theta$ and Var_θ, respectively. Note that, although the results below are for $\theta \to \infty$, the approximations are actually exact for $\theta = 0$ as well.

Corollary 2. *Let $X \sim \mathrm{LN}\,(0,\sigma^2)$. Then as $\theta \to \infty$*

$$\mathbb{E}_\theta[X] = \exp\big\{-w_0(\theta) + \tfrac{1}{2}\sigma_0(\theta)^2\big\}\big(1 + \mathcal{O}(\sigma_0(\theta)^2)\big), \tag{8}$$
$$\mathrm{Var}_\theta[X] = \exp\big\{-2w_0(\theta) + \sigma_0(\theta)^2\big\}\big(e^{\sigma_0(\theta)^2} - 1\big)\big(1 + \mathcal{O}(\sigma_0(\theta)^2)\big). \tag{9}$$

Proof. Simply use $\mathbb{E}_\theta[X] = L_1(\theta)/L_0(\theta)$ and $\mathrm{Var}_\theta[X] = L_2(\theta)/L_0(\theta) - (L_1(\theta)/L_0(\theta))^2$ together with the second part of (7). □

The Laplace approximation in Proposition 1 and Corollary 2 corresponds to replacing the distribution of $Y_\theta = \log(X_\theta)$, $X_\theta \sim F_\theta$, by a normal distribution, $Y_\theta \sim N(-w_0(\theta), \sigma_0(\theta)^2)$. This is equivalent to a lognormal approximation for the tilted measure F_θ, $F_\theta \approx \mathrm{LN}(-w_0(\theta), \sigma_0(\theta)^2)$. For $\theta = 0$ we have the correct lognormal distribution, and Proposition 3 below shows that this is a correct interpretation in the limit $\theta \to \infty$. The proposition shows that the limiting centered and scaled tilted density f_θ is a standard normal density. It follows from this that the lognormal approximation becomes exact as $\theta \to \infty$. We will use the result in the following sections.

Proposition 3. *Write the tilted density $f_\theta(x)$ as $\exp\big\{-m(x) - \kappa(\theta)\big\}/\sqrt{2\pi\sigma^2}$ with $m(x) = \log(x) + (\log(x))^2/(2\sigma^2) + \theta x$. Furthermore, let $w = w_0(\theta)$ and define $m_0(u) = m\big(e^{-w}(1 + \sigma u/\sqrt{w})\big) - m(\mathrm{e}^{-w})$. Then, as $\theta \to \infty$,*

$$m_0(u) = \tfrac{1}{2}u^2 + \mathcal{O}\big((|u| + |u|^3)\sigma/\sqrt{w}\big) \quad \textit{for} \quad |u|^3/\sqrt{w} \le 1,$$

and for θ sufficiently large $|m_0'(u)| > (\sqrt{w}/\sigma)^{1/6}/2$ for $|u| > (\sqrt{w}/\sigma)^{1/6}$ with the sign of $m_0'(u)$ being that of u. The properties of m_0 imply that the centered and scaled density f_θ converges to a standard normal density, and moments of f_θ converge as well.

Proof. We first note that the lognormal density $f(x)$ is logconcave for $x < \mathrm{e}^{1-\sigma^2}$ since

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\log(f(x)) = -\frac{1}{x^2\sigma^2}\big(-\log(x)+\sigma^2-1\big) < 0 \quad \text{for} \quad x < \mathrm{e}^{1-\sigma^2}. \tag{10}$$

We rewrite $m_0(u)$ as

$$m_0(u) = \log\Big(1+\frac{\sigma}{\sqrt{w}}u\Big) + \frac{1}{2\sigma^2}\Big\{\Big[-w+\log\Big(1+\frac{\sigma}{\sqrt{w}}u\Big)\Big]^2 - w^2\Big\} + \frac{\sqrt{w}}{\sigma}u.$$

Taylor expanding $\log(1+\sigma u/\sqrt{w})$ we obtain the first result of the proposition. Next, we find the derivative of $m_0(u)$:

$$m_0'(u) \;=\; \frac{\sigma/\sqrt{w}}{1+\sigma u/\sqrt{w}} + \frac{\sqrt{w}}{\sigma} + \frac{1}{\sigma^2}\big[-w+\log(1+\sigma u/\sqrt{w})\big]\frac{\sigma/\sqrt{w}}{1+\sigma u/\sqrt{w}}.$$

For $u > (\sqrt{w}/\sigma)^{1/6}$ we get the bound

$$m_0' > \frac{\sqrt{w}}{\sigma}\Big(1-\frac{1}{1+\sigma u/\sqrt{w}}\Big) = \frac{(\sqrt{w}/\sigma)^{1/6}}{1+(\sqrt{w}/\sigma)^{-5/6}} > \tfrac{1}{2}(\sqrt{w}/\sigma)^{1/6},$$

as long as $\sigma/\sqrt{w} < 1$, which is true for $\theta \to \infty$. For $u < -(\sqrt{w}/\sigma)^{1/6}$ we have from the logconcavity that $m_0'(u) < m_0'\big(-(\sqrt{w}/\sigma)^{1/6}\big)$. For the latter we find $m_0'\big(-(\sqrt{w}/\sigma)^{1/6}\big) \sim -(\sqrt{w}/\sigma)^{1/6}$ as $\theta \to \infty$. □

2.1 The saddlepoint $\theta(x)$

The result in Corollary 2 leads in a natural way to an approximation of the saddlepoint $\theta(x)$, the latter being the solution of the equation $-\kappa'(\theta) = L_1(\theta)/L_0(\theta) = x$. We simply let the approximation $\tilde\theta(x)$ be the solution of $\exp\big\{-w_0(\theta)+\frac{1}{2}\sigma_0(\theta)^2\big\} = x$. This gives the equation $-w_0(\theta)+\frac{1}{2}\sigma^2/(1+w_0(\theta)) = \log(x)$, which leads to a quadratic equation in $w_0(\theta)$. Since $w_0(\theta) \geq 0$, and using the definition of $w_0(\theta)$, we find with $\gamma(x) = \frac{1}{2}\big(-1-\log x+\sqrt{(1-\log x)^2+2\sigma^2}\,\big)$ that

$$w_0(\tilde\theta(x)) = \gamma(x) \quad \text{or} \quad \tilde\theta(x) = \gamma(x)\mathrm{e}^{\gamma(x)}/\sigma^2. \tag{11}$$

The following proposition states the quality of this approximation.

Proposition 4. *For $x \to 0$ we have $\tilde\theta(x) \sim (-\log x)/(x\sigma^2)$ and*

$$\mathbb{E}_{\tilde\theta(x)}[X] = x\Big(1+\mathcal{O}\Big(\frac{1}{|\log(x)|}\Big)\Big), \qquad \theta(x) = \tilde\theta(x)\Big(1+\mathcal{O}\Big(\frac{1}{|\log(x)|}\Big)\Big).$$

Proof. Below we write $\bar{\theta}$ for $\theta(x)$ and $\tilde{\theta}$ for $\tilde{\theta}(x)$. We first note that $\gamma(x) \sim -\log(x) + \mathcal{O}(1/|\log(x)|)$ for $x \to 0$, which gives that $\tilde{\theta} = \gamma(x)e^{\gamma(x)}/\sigma^2 \sim (-\log x)/(x\sigma^2)$. From the definition of $\tilde{\theta}$ and (8) we have $\mathbb{E}_{\tilde{\theta}}[X] = x(1+\mathcal{O}(1/w_0(\tilde{\theta})\big)$, and the first displayed result follows from $w_0(\tilde{\theta}) = \gamma(x) \sim -\log(x)$. Using that $\mathbb{E}_{\bar{\theta}}[X] = x$ we rewrite the first result as $\mathbb{E}_{\bar{\theta}}[X]/\mathbb{E}_{\tilde{\theta}}[X] = 1 + \mathcal{O}\big(1/w_0(\tilde{\theta})\big)$. From a Taylor expansion of $\mathbb{E}_\theta[X]$ around $\tilde{\theta}$ we find $\mathbb{E}_{\bar{\theta}}[X]/\mathbb{E}_{\tilde{\theta}}[X]$ to first order as

$$1 + \frac{\mathrm{Var}_{\tilde{\theta}}[X]}{\mathbb{E}_{\tilde{\theta}}[X]}\tilde{\theta}\Big(1 - \frac{\bar{\theta}}{\tilde{\theta}}\Big) \approx 1 + \frac{\frac{\sigma^2}{w_0(\tilde{\theta})}e^{-2w_0(\tilde{\theta})}}{e^{-w_0(\tilde{\theta})}}\frac{w_0(\tilde{\theta})}{\sigma^2}e^{w_0(\tilde{\theta})}\Big(1 - \frac{\bar{\theta}}{\tilde{\theta}}\Big) \approx 1 + \Big(1 - \frac{\bar{\theta}}{\tilde{\theta}}\Big),$$

and comparing this with $1 + \mathcal{O}\big(1/w_0(\tilde{\theta})\big)$ we conclude that $1 - \bar{\theta}/\tilde{\theta} = \mathcal{O}\big(1/w_0(\tilde{\theta})\big)$ or $\bar{\theta} = \tilde{\theta}\big(1 + \mathcal{O}(1/|\log(x)|\big)$. $\square$

In Sections 3 and 4.2 we will employ the results of this section to construct a saddlepoint approximation and a Monte Carlo estimator of the left tail probability of a sum of lognormal random variables. In particular, the asymptotic results derived above will be useful to show that when the approximation $\tilde{\theta}(x)$ is used as the tilting parameter of an exponential change of measure, the Monte Carlo estimator remains asymptotically efficient as $x \to 0$.

3 Saddlepoint approximation in the left tail of a lognormal sum

Daniels' saddlepoint method produces an approximation of the density function of a sum of iid random variables which is valid asymptotically on the number of summands. The first and second order approximations are embodied in the formula

$$f_n(nx) \approx \big\{2\pi n\kappa''(\theta(x))\big\}^{-1/2}\exp\big\{-n\kappa^\dagger\big\}\big(1 + \tfrac{1}{n}\big[\zeta_4/8 - 5\zeta_3^2/24\big]\big),$$

where $\kappa^\dagger = -\{\kappa\big(\theta(x)\big) + x\theta(x)\}$ is the convex conjugate of κ evaluated at $-x$, and $\zeta_k = \kappa^{(k)}(\theta(x))/\kappa''(\theta(x))^{k/2}$, is the standardized cumulant.

The corresponding saddlepoint approximation for the cdf (cf. Jensen, 1995, Chapter 2) is given by

$$F_n(nx) = \frac{1}{\lambda_n}\exp\big\{n\kappa^\dagger\big\}\Big\{B_0 + \frac{\zeta_3}{6\sqrt{n}}B_3 + \frac{\zeta_4}{24n}B_4 + \frac{\zeta_3^2}{72n}B_6\Big\},$$

where $\lambda_n = \theta(x)\sqrt{n\kappa''(\theta(x))}$ and

$$B_0 = \lambda_n e^{\lambda_n^2/2}\Phi(-\lambda_n), \qquad B_3 = -\big\{\lambda_n^3 B_0 - (\lambda_n^3 - \lambda_n)/\sqrt{2\pi}\big\},$$
$$B_4 = \lambda_n^4 B_0 - \big(\lambda_n^4 - \lambda_n^2\big)/\sqrt{2\pi}, \qquad B_6 = \lambda_n^6 B_0 - \big(\lambda_n^6 - \lambda_n^4 + 3\lambda_n^2\big)/\sqrt{2\pi}.$$

General results for the saddlepoint approximation state that for a fixed x the relative error is $\mathcal{O}(1/n)$ for the first order approximation and $\mathcal{O}(1/n^2)$ for the second order approximation. More can be said, however, for the case of a lognormal sum. It is simple to see that the density $f(x)$ is logconcave for $x < e^{1-\sigma^2}$, see (10), and according to (Jensen, 1995, section 6.2) we have that the saddlepoint approximations have the stated relative errors uniformly for x in a region around zero. Furthermore, the convergence of the tilted density as $\theta \to \infty$ outlined in Proposition 3 implies that the saddlepoint approximation become exact in the same limit.

To calculate the saddlepoint approximation we need to find the Laplace transform and its derivatives numerically. We want to implement the integration in such a way that the relative accuracy of the integration is of the same order irrespective of the argument θ. For $k = 0, 1, 2, 3, 4$ we want to evaluate the integral $L_k(\theta)$ from (4). This leads to the integral $I_k(\theta)$ from (6). Instead of the scale $\sigma_k(\theta)$ chosen there, we consider another scale τ and the integral

$$I_k(\theta) = \frac{1}{\sqrt{2\pi}} \frac{\tau}{\sigma_k(\theta)} \int_{-\infty}^{\infty} \exp\bigl\{-\tilde{h}(z)\bigr\} \mathrm{d}z, \quad \tilde{h}(z) = \frac{w_k(\theta)}{\sigma^2}\bigl(\mathrm{e}^{\tau z} - 1 - \tau z\bigr) + \frac{\tau^2}{2\sigma^2} z^2.$$

Since $\tilde{h}''(z) = \tau^2\bigl(w_k(\theta)\mathrm{e}^{\tau z} + 1\bigr)/\sigma^2 > 0$ we see that $\tilde{h}$ is convex. Choosing the scale τ such that $2\tilde{h}(-\tau) = 1$ we obtain that $2\tilde{h}(z)$ is a convex function bounded between 0 and 1 for $-1 < z < 0$, is above $-z$ for $z < -1$ and with $\tilde{h}(z) \geq \tilde{h}(-z)$ for $z > 0$. In this way the precision of the numerical integration of $\exp\bigl\{-\tilde{h}(z)\bigr\}$ will be of the same order irrespective of the value of w and σ^2. In practice we can take τ as $\tau = \sigma_k(\theta)$ as long as $\sigma_k(\theta) \leq c_0$ and $\tau = \{w_k(\theta)^2 + 2w_k(\theta) + \sigma^2\}^{1/2} - w_k(\theta)$ otherwise, where c_0 is an arbitrary constant. Unless σ^2 is large we can use $\tau = \sigma_k(\theta)$ for all θ.

4 Simulation

In this section we discuss an optimal importance sampling estimator based on a classical exponential change of measure. For implementing such an estimator we require an algorithm for generating a sample from F_θ. Also, we use the previous approximation $\tilde{\theta}(x)$ for the saddlepoint.

4.1 Random variate generation from F_θ

We first consider the problem of generating a random variable from the tilted density f_θ from (1).

The obvious naïve choice is acceptance-rejection (A-R; Asmussen and Glynn, 2007, II.2), simulating Z from f and rejecting with probaility $\mathrm{e}^{-\theta Z}$. This choice produces a very simple algorithm for generating from f_θ and the method is exact even when we do not have an explicit expression for $\kappa(\theta)$. Ideally, we would like to have an acceptance probability p close to 1, but in our case $p = \mathrm{e}^{\kappa(\theta)}$, so as the value of θ increases, the probability of acceptance diminishes, and hence the expected number of rejection steps goes to infinity. In consequence, the naïve estimator is very inefficient for large values of θ.

As noted in Proposition 3, if X_θ is a random variable with density f_θ, the variable $U = (X_\theta - \mathrm{e}^{-w})\sqrt{w}\mathrm{e}^{w}/\sigma$, $w = w_0(\theta)$, has a standard normal distribution in the limit $\theta \to \infty$. However, the limiting normal distribution cannot be used as a proposal for an A-R algorithm because the right tail is lighter than that of X_θ. Similarly, the lognormal approximation is not applicable because the left tail of $\big(X_\theta+w_0(\theta)\big)/\sigma_0(\theta)$ is lighter.

What we know, however, is that the right tail of X_θ is trivially lighter than $\mathrm{e}^{-\theta x}$. This points to the possibility of using a gamma proposal $Z \sim \mathrm{Gamma}(\lambda,\theta)$. For a given $\lambda > 0$ we rewrite the tilted density as

$$f_\theta(x) = \frac{\mathrm{e}^{-\kappa(\theta)+\frac{1}{2}\lambda^2\sigma^2}}{\sqrt{2\pi\sigma^2}} x^{\lambda-1}\exp\Big\{-\theta x - \frac{1}{2\sigma^2}(\lambda\sigma^2+\log x)^2\Big\}. \tag{12}$$

We choose λ such that $\lambda\sigma^2 + \log(\mathbb{E}[Z]) = 0$. Solving for λ we obtain $0 = \lambda\sigma^2 + \log(\lambda/\theta)$ or $\lambda = w_0(\theta)/\sigma^2$. This gives the following A-R algorithm.

Algorithm 5.

1. Simulate $U \sim \mathrm{U}(0,1)$ and $Z \sim \mathrm{Gamma}(w_0(\theta)/\sigma^2,\theta)$.
2. If $U > \exp\big\{-\big(w_0(\theta)+\log Z\big)^2/(2\sigma^2)\big\}$ repeat. Else, return $X_\theta = Z$.

Writing $w = w_0(\theta)$, using (12) and $\theta = w_0(\theta)e^{w_0(\theta)}/\sigma^2$, we find the acceptance probability as

$$\mathbb{E}[\mathrm{e}^{-\frac{1}{2\sigma^2}(w+\log Z)^2}] = R(\theta)e^{\kappa(\theta)} \quad \text{with} \quad R(\theta) = \sqrt{2\pi\sigma^2}\frac{(w/\sigma^2)^{w/\sigma^2}\mathrm{e}^{\frac{1}{2}w^2/\sigma^2}}{\Gamma(w/\sigma^2)}.$$

As $\theta \to \infty$ we have $\kappa(\theta) \sim -w^2/(2\sigma^2) - w/\sigma^2 - \log(1+w)/2$, and from $\Gamma(x) \sim x^{x-1/2}e^{-x}\sqrt{2\pi}$ for large x we therefore find that the acceptance probability $R(\theta)e^{\kappa(\theta)} \to 1$. Since $\Gamma(x) \to \infty$ for $x \to 0$ we have that the factor $R(\theta)$ is below one for small θ. Thus, it seems natural to choose between the two algorithms according to which has the highest acceptance probability.

4.2 Efficient Monte Carlo for left tails of lognormal sums

In this section we develop an *asymptotically efficient Monte Carlo estimator* $\widehat{\alpha}_n(x)$, for the left tail probability of a lognormal sum $\alpha_n(x) = \mathbb{P}(S_n \le nx)$ which may be small either because x is small or because n is large.

As is standard concepts in rare event simulation (VI.1 Asmussen and Glynn, 2007), we say that a Monte Carlo estimator $\widehat{\alpha}_n(x)$ has *bounded relative error* as $x \to \infty$ if $\limsup_{x\to 0}\mathrm{Var}(\widehat{\alpha}_n(x))/\alpha_n^{2-\epsilon}(x) = 0$, for all $\epsilon > 0$, or is *logarithmically efficient* if $\limsup_{x\to 0}\mathrm{Var}(\widehat{\alpha}_n(x))/\alpha_n^2(x) < \infty$. Bounded relative error implies that the number of replications required to estimate $\alpha_n(x)$ with certain fixed relative precision remains bounded as $x \to 0$ and logarithmic efficiency that it grows at rate of order at most $|\log(\alpha_n(x))|$ which is only marginally weaker in practice. This is to be compared with the much cruder rate $\alpha_n(x)^{-1/2}$ obtained using naive simulation.

The probabilty of the event $(S_n \le x)$ may be small either because x is small or because n is large, so one could alternatively study $\alpha_n(x)^{-1/2}$ in the limit $n \to \infty$ instead of $x \to 0$. In fact, for light right tails it is then standard to apply importance sampling (V.1, Asmussen and Glynn, 2007). Noting that the lognormal density is log-concave for small x (see (10)), the following analogue of Theorem 2.10, Chapter VI in Asmussen and Glynn (2007) follows immediately from the proof in *loc. cit.*:

Theorem 6. *Consider $X_1, \ldots, X_n \sim F_{\theta(x)}$, where $\theta(x)$ is the saddlepoint from* (3)*, and set $S_n = X_1 + \cdots + X_n$. Define $\beta_n(x) = \mathrm{e}^{\theta(x)S_n} L(\theta(x))^n \, \mathbb{I}\{S_n < nx\}$. Then $\beta_n(x)$ is a logarithmically efficient and unbiased estimator of $\alpha_n(x)$ as $n \to \infty$.*

Unfortunately, this algorithm requires the value of the Laplace transform $L(\theta)$ and the saddlepoint $\theta(\cdot)$. Instead, we consider an alternative estimator based on the approximation (11) to the saddlepoint and an unbiased estimator of the Laplace transform. This alternative estimator is unbiased and logarithmically efficient as $x \to 0$. It uses the unbiased estimator of the Laplace transform suggested in Asmussen et al. (2014b) and given by $\widehat{\mathcal{L}}(\theta) = (\sigma/\sigma_0(\theta))L_a(0,\theta)V$ with $V = \exp\{-(\mathrm{e}^Y - 1 - Y)w_0(\theta)/\sigma^2\}$ where $Y \sim \mathrm{N}(0, \sigma^2)$.

Algorithm 7.

1. Use the approximation $\tilde{\theta} = \tilde{\theta}(x)$ to the saddlepoint given in (11).
2. Obtain n independent unbiased estimates $\widehat{\mathcal{L}}_i(\tilde{\theta}) = (\sigma/\sigma_0(\tilde{\theta}))L_a(0,\tilde{\theta})V_i$ of the Laplace transform and set $\widehat{\mathcal{L}}^*(\tilde{\theta}) = \prod_{i=1}^n \widehat{\mathcal{L}}_i(\tilde{\theta})$.
3. Simulate $X_1, \ldots, X_n \sim F_{\tilde{\theta}}$ and set $S_n = X_1 + \cdots + X_n$.
4. Return $\widehat{\alpha}_n(x) = \mathrm{e}^{\tilde{\theta}S_n} \widehat{\mathcal{L}}^*(\tilde{\theta}) \, \mathbb{I}\{S_n < nx\}$.

The product of n independent copies of an unbiased estimate $\widehat{\mathcal{L}}(\tilde{\theta})$ is needed because $\widehat{\mathcal{L}}(\tilde{\theta})^n$ is not an unbiased estimate of $L(\tilde{\theta})^n$. We next state the properties of the proposed algorithm; the proof is given in Appendix A, where also the logarithmic efficiency of an alternative estimator $\widehat{\beta}_n(x)$ is proved.

Proposition 8. *Let $\widehat{\alpha}_n(x)$ be defined as in Algorithm 7(4). Then $\widehat{\alpha}_n(x)$ is an unbiased and logarithmically efficient estimator of $\alpha_n(x)$ as $x \to 0$.*

Importance sampling can also be used to estimate the density of a lognormal sum via simulation. Following Example V.4.3 (p. 146) of Asmussen and Glynn (2007), slightly extended, we first note that the conditional density at nx of S_n given $S_{n,-i} = X_1 + \cdots + X_{i-1} + X_{i+1} + \cdots + X_n = S_n - X_i$ is $f(nx - S_{n,-i})$. Hence an unbiased estimator of $f_n(nx)$ is $\sum_1^n f(nx - S_{n,-i})/n$. To avoid the problem that many $S_{n,-i}$ will exceed x so that $f(nx - S_{n,-i}) = 0$, we simulate the X_j from $F_{\tilde{\theta}(x)}$ and return the estimator

$$\widehat{f}_n(nx) \;=\; \frac{1}{n}\sum_{i=1}^n f(nx - S_{n,-i}) \exp\{\tilde{\theta}(x)S_{n,-i} + (n-1)\kappa\big(\tilde{\theta}(x)\big)\}\,. \tag{13}$$

In Gulisashvili and Tankov (2014), an importance sampling estimator for $F_n(z)$ is suggested and it is written that a parallel estimator for $f_n(z)$ can be constructed in the same way. Nevertheless, we do not follow the details for the construction of that estimator of $f_n(z)$.

5 Numerical examples

In our numerical experiments, we have taken parameter values that we consider realistic from the point of view of financial applications. A yearly volatility of order 0.25 is often argued to be typical. We have considered periods of lengths one year, one quarter, one month and one week, corresponding to $\sigma = 0.25$, $\sigma = 0.25/\sqrt{4} = 0.125$, $\sigma = 0.25/\sqrt{12} = 0.072$, $\sigma = 0.25/\sqrt{52} = 0.035$ resp. (note that the value $\sigma = 0.25$ is also argued to be particularly relevant in the optical context of Barakat (1976)) The number of assets in portfolios is often large, even in the thousands; the values we have chosen are $n = 4, 16, 64, 256$.

For each combination of n and σ we have conducted several empirical analyses. In all numerical experiments involving simulation we have employed $R = 100,000$ replications. The complete set of numerical results can be found in Asmussen et al. (2014a). Here we show a few numerical illustrations.

We present and discuss an example with $n = 16$ and $\sigma = 0.125$. We consider the approximation $\tilde{\theta}(x)$ given in (11) to the saddlepoint $\theta(x)$. The overall result is given in Proposition 4. Table 1 gives $\tilde{\theta}(x)$, $\theta(x)$. The relative error of the mean under the tilted measure corresponding to $\tilde{\theta}(x)$ as an approximation to x is less than one percent (numbers not shown). Furthermore, when using $\tilde{\theta}(x)$ as the initial value in a Newton–Raphson search for $\theta(x)$, in all cases considered at most four iterations are needed to find $\theta(x)$ to accuracy 10^{-10}.

Next we verify the approximations for the cdf and pdf of the lognormal sum. We have thereby considered a portfolio of n assets with next-period values $Y_1, \dots, Y_n$ assumed iid lognormal(μ, σ^2), such that a loss corresponds to a small value x of $S_n = Y_1 + \ldots + Y_n$. When choosing x, we have had in mind the recommended VaR values 0.99%–0.997% of Basel II (2004) and have chosen $\mathbb{P}(S_n \le nx)$ to cover the interval 0.0001–0.0100.

We have proposed two types of approximations: saddlepoint approximations and Monte Carlo estimators. Thus, in Table 1 we included the saddlepoint approximation based on our formulas in Section 3, and Monte Carlo estimators (MC) based on our algorithms in Section 4. The last is based on the proposed importance sampling estimator where the importance distribution is selected from the exponential family. The general estimator for the cdf of the lognormal sum has the form $\widehat{F}_n(nx) = L(\theta)^n \, \mathrm{e}^{\theta S_n} \, \mathbb{I}\{S_n < nx\}$, where $S_n = X_1 + \cdots + X_n$ and $X_1, \dots, X_n$ is a sample from the exponentially tilted distribution F_θ. Similarly, the MC estimator of the pdf of the lognormal sum has the form (13). The parameter θ defining the distribution is selected to be equal to the saddlepoint $\theta(\cdot)$ evaluated at x.

In Table 1, the solution $\theta(x)$ obtained by using Newton-Raphson is the one used for obtaining the saddlepoint approximations and MC estimators. In the cases considered, the saddlepoint approximation agrees with the results from the Monte Carlo simulations. The column headed L_{app} of Table 1 indicates the relative error that one would introduce by replacing the Laplace transform in $\widehat{F}_n(nx)$ with its approximation $L_a(0, \tilde{\theta})$. For $n = 16$ the relative errors are $(1 + \epsilon)^n - 1$, where ϵ is the entry in the table.

Acknowledgements: LRN is supported by ARC grant DE130100819.

Table 1: Approximation of the cdf and pdf of a lognormal sum with $n = 16$ and $\sigma = 0.125$. *MC/Saddle* is the Monte-Carlo estimator divided by the saddlepoint approximation *Saddle*. The entry L_{app} is the relative error $L(\tilde{\theta}(x))/L_a(0, \tilde{\theta}(x)) - 1$.

x	$\theta(x)$	$\tilde{\theta}(x)$	Saddle-cdf	MC/Saddle	L_{app}	Saddle-pdf	MC/Saddle
0.70	33.13	33.33	$1.761 \cdot 10^{-31}$	0.992 ± 0.070	$2.12 \cdot 10^{-4}$	$5.873 \cdot 10^{-30}$	0.997 ± 0.009
0.80	18.36	18.48	$9.807 \cdot 10^{-14}$	1.001 ± 0.017	$2.04 \cdot 10^{-4}$	$1.829 \cdot 10^{-12}$	1.002 ± 0.009
0.85	12.74	12.83	$3.031 \cdot 10^{-8}$	0.991 ± 0.015	$1.83 \cdot 10^{-4}$	$3.975 \cdot 10^{-7}$	0.998 ± 0.009
0.90	7.99	8.05	$1.632 \cdot 10^{-4}$	0.995 ± 0.060	$1.48 \cdot 10^{-4}$	$1.388 \cdot 10^{-3}$	1.004 ± 0.009
0.91	7.13	7.18	$5.956 \cdot 10^{-4}$	0.994 ± 0.012	$1.38 \cdot 10^{-4}$	$4.577 \cdot 10^{-3}$	1.001 ± 0.009
0.92	6.30	6.34	$1.912 \cdot 10^{-3}$	1.011 ± 0.011	$1.28 \cdot 10^{-4}$	$1.319 \cdot 10^{-2}$	0.999 ± 0.008
0.93	5.49	5.53	$5.424 \cdot 10^{-3}$	1.001 ± 0.010	$1.17 \cdot 10^{-4}$	$3.332 \cdot 10^{-2}$	0.998 ± 0.009
0.94	4.71	4.74	$1.368 \cdot 10^{-2}$	0.997 ± 0.010	$1.06 \cdot 10^{-4}$	$7.416 \cdot 10^{-2}$	1.000 ± 0.009
0.95	3.95	3.98	$3.081 \cdot 10^{-2}$	0.992 ± 0.009	$9.29 \cdot 10^{-5}$	$1.460 \cdot 10^{-1}$	0.998 ± 0.009
0.98	1.82	1.83	$1.901 \cdot 10^{-1}$	1.005 ± 0.007	$4.92 \cdot 10^{-5}$	$5.520 \cdot 10^{-1}$	0.997 ± 0.009

References

Aitchison, I. and J. A. C. Brown (1957). *The Lognormal distribution with special reference to its uses in economics.* Cambridge: Cambridge University Press.

Asmussen, S. (2003). *Applied probability and queues.* New York: Springer-Verlag.

Asmussen, S., J. Blanchet, S. Juneja, and L. Rojas-Nandayapa (2011). Efficient simulation of tail probabilities of sums of correlated lognormals. *Annals of Operations Research 189*, 5–23.

Asmussen, S. and P. W. Glynn (2007). *Stochastic simulation: algorithms and analysis.* New York: Springer-Verlag.

Asmussen, S., J. L. Jensen, and L. Rojas-Nandayapa (2014a). Numerical examples accompanying the paper "Exponential Family Techniques for the Lognormal Left Tail". http://www.smp.uq.edu.au/people/LeonardoRojas-Nandayapa/NumericalExamples.pdf.

Asmussen, S., J. L. Jensen, and L. Rojas-Nandayapa (2014b). On the Laplace transform of the Lognormal distribution. *Methodology and Computing in Applied Probability*. DOI 10.1007/s11009-014-9430-7.

Asmussen, S. and L. Rojas-Nandayapa (2008). Asymptotics of sums of lognormal random variables with Gaussian copula. *Statistics and Probability Letters 78*, 2709–2714.

Barakat, R. (1976). Sums of independent lognormally distributed random variables. *Journal of the Optical Society of America 66*, 211–216.

Basel II (2004). *Basel II: revised international capital framework.* Bank for International Settlements.

Beaulieu, N. C., A. A. Abu-Dayya, and P. J. McLane (1995). Estimating the distribution of a sum of independent lognormal random variables. *IEEE Transactions on Communications 43*(12), 2869–2873.

Beaulieu, N. C. and Q. Xie (2004). An optimal lognormal approximation to lognormal sum distributions. *IEEE Transactions on Vehicular Technology 53*(2), 479–489.

Blanchet, J. H. and L. Rojas-Nandayapa (2011). Efficient simulation of tail probabilities of sums of dependent random variables. *Journal of Applied Probability 48A*, 147–164.

Corless, R. M., G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth (1996). On the Lambert W function. *Advances in Computational Mathematics 5*, 329–359.

Crow, E. L. and K. Shimizu (1988). *Lognormal distributions: theory and applications.* New York: Marcel Dekker Inc.

Duelman, K. (2010). Regulatory capital. *Encyclopedia of Quantitative Finance 17*, 1–14.

Dufresne, D. (2004). The log-normal approximation in financial and other computations. *Advances in Applied Probability*, 747–773.

Dufresne, D. (2009). Sums of lognormals. Technical report, Centre for Actuarial Sciences, University of Melbourne.

Embrechts, P., G. Puccetti, L. Rüschendorf, R. Wang, and A. Beleraj (2014). An academic response to basel 3.5. *Risks 2*, 25–48.

Fenton, L. (1960). The sum of log-normal probability distributions in scatter transmission systems. *IRE Transactions on Communications Systems 8*(1), 57–67.

Foss, S. and A. Richards (2010). On sums of conditionally independent subexponential random variables. *Mathematics of Operations Research 35*, 102–119.

Gubner, J. A. (2006). A new formula for lognormal characteristic functions. *IEEE Transactions on Vehicular Technology 55*(5), 1668–1671.

Gulisashvili, A. and P. Tankov (2014). Tail behavior of sums and differences of log-normal random variables. *arXiv:1309.3057*.

Jensen, J. L. (1995). *Saddlepoint approximations.* Oxford: Oxford Science Publications.

Johnson, N. L., S. Kotz, and N. Balakrishnan (1994). *Continuous univariate distributions*, Volume 1. New York: Wiley.

Laub, P., S. Asmussen, J. L. Jensen, and L. Rojas-Nandayapa (2016). Approximating the Laplace transform of the sum of dependent lognormals. In C. M. Goldie and A. Mijatovic (Eds.), *Festschrift for Nick Bingham*, Volume 48A of *Advances in Applied Probability.*

Limpert, E., W. A. Stahel, and M. Abbt (2001). Log-normal distributions across the sciences: Keys and clues. *BioScience 51*, 341–352.

Marlow, N. A. (1967). A normal limit theorem for power sums of normal random variables. *Bell System Technical Journal 46*, 2081–2089.

McNeil, A., R. Frey, and P. Embrechts (2015). *Quantitative Risk Management: Concepts, Techniques and Tools* (2nd ed.). New Jersey: Princeton University Press.

Milevsky, M. A. and S. E. Posner (1998). Asian options, the sum of lognormals, and the reciprocal gamma distribution. *The Journal of Financial and Quantitative Analysis 33*(3), 409–422.

Mitra, A. and S. I. Resnick (2009). Aggregation of risks and asymptotic independence. *Advances in Applied Probability 41*, 797–828.

Navidpour, S. M., M. Uysal, and J. Li (2004). BER performance of MIMO free-space optical links. In *Vehicular Technology Conference*. IEEE.

Rojas-Nandayapa, L. (2008). *Risk probabilities: asymptotics and simulation*. Ph. D. thesis, Aarhus University.

Slimane, B. (2001). Bounds on the distribution of a sum of independent lognormal random variables. *IEEE Transactions on Communications 49*(6), 975–978.

Thorin, O. and N. Wikstad (1977). Calculation of ruin probabilities when the claim distribution is lognormal. *Astin Bulletin 9*, 231–246.

Corresponding author: Leonardo Rojas-Nandayapa, School of Mathematics and Physics, The University of Queensland. email: l.rojas@uq.edu.au

A Appendix: Proof of Proposition 8

Let $\tilde{\theta} = \tilde{\theta}(x)$, $\tilde{\mu} = \mathbb{E}_{\tilde{\theta}}[X]$, $\tilde{\tau}^2 = \mathrm{Var}_{\tilde{\theta}}[X]$ and $Z_n = (S_n - n\tilde{\mu})/(\tilde{\tau}\sqrt{n})$, where S_n is based on a random sample from $F_{\tilde{\theta}}$. We want to estimate

$$\begin{aligned}\alpha_n(x) &= \mathbb{E}_{\tilde{\theta}}\big[L(\tilde{\theta})^n \mathrm{e}^{\tilde{\theta}S_n}\mathbb{I}\{S_n < nx\}\big] \\ &= \frac{\mathrm{e}^{n\tilde{\theta}x}}{\sqrt{n}\tilde{\theta}\tilde{\tau}} L(\tilde{\theta})^n \mathbb{E}_{\tilde{\theta}}[\sqrt{n}\tilde{\theta}\tilde{\tau}\mathrm{e}^{-\sqrt{n}\tilde{\theta}\tilde{\tau}(\xi - Z_n)}\mathbb{I}\{Z_n < \xi\}], \quad \xi = \sqrt{n}(x-\tilde{\mu})/\tilde{\tau}. \qquad (14)\end{aligned}$$

We know from Proposition 3 that as $x \to 0$, corresponding to $\tilde{\theta} \to \infty$, the distribution of Z_n approaches a standard normal distribution. Furthermore, $\tilde{\theta} \sim |\log x|/(\sigma^2 x)$, $\tilde{\mu} \sim x\big(1 + \mathcal{O}\big(1/|\log x|\big)\big)$ and $\tilde{\tau} \sim x\sigma/\sqrt{(|\log x|)}$. Since generally $\tilde{w} = w_0(\tilde{\theta}) \sim \log(\tilde{\theta})$, we have also $\tilde{w} \sim |\log x|$. We therefore find

$$\xi = \frac{\sqrt{n}(x - \tilde{\mu})}{\tilde{\tau}} = \mathcal{O}\big(\sqrt{n}/\sqrt{(|\log x|)}\,\big) \to 0,$$

and

$$\tilde{\theta}\tilde{\tau} \sim \frac{\sqrt{|\log x|}}{\sigma} \sim \frac{\sqrt{\tilde{w}}}{\sigma}, \qquad \tilde{\theta}x \sim \frac{|\log x|}{\sigma^2} \sim \frac{\tilde{w}}{\sigma^2}.$$

These findings show that the mean value in (14) tends to $1/\sqrt{2\pi}$ as $x \to 0$. The same type of argument also gives that

$$\begin{aligned}&\mathbb{E}_{\tilde{\theta}}\big[\big\{\sqrt{n}\tilde{\theta}\tilde{\tau}\mathrm{e}^{-\sqrt{n}\tilde{\theta}\tilde{\tau}(\xi - Z_n)}\mathbb{I}\{Z_n < \xi\}\big\}^2\big] \\ &\quad = \tfrac{1}{2}\sqrt{n}\tilde{\theta}\tilde{\tau}E_{\tilde{\theta}}\big[2\sqrt{n}\tilde{\theta}\tilde{\tau}\mathrm{e}^{-2\sqrt{n}\tilde{\theta}\tilde{\tau}(\xi - Z_n)}\mathbb{I}\{Z_n < \xi\}\big] \\ &\quad \sim \frac{1}{2\sqrt{2\pi}}\sqrt{n}\tilde{\theta}\tilde{\tau}. \qquad (15)\end{aligned}$$

Consider now the unbiased estimator

$$\hat{\beta}_n(x) = L(\tilde{\theta})^n \mathrm{e}^{\tilde{\theta} S_n} \mathbb{I}\{S_n < nx\},$$

where S_n is based on a sample from the tilted measure $F_{\tilde{\theta}}$. The above calculations show that

$$\begin{aligned}
\frac{\operatorname{Var}[\hat{\beta}_n(x)]}{\beta_n(x)^{2-\epsilon}} &= \mathcal{O}\Big(\frac{\big\{\mathrm{e}^{n\tilde{\theta}x} L(\tilde{\theta})^n/(\sqrt{n}\tilde{\theta}\tilde{\tau})\big\}^2 \sqrt{n}\tilde{\theta}\tilde{\tau}}{\big\{\mathrm{e}^{n\tilde{\theta}x} L(\tilde{\theta})^n/(\sqrt{n}\tilde{\theta}\tilde{\tau})\big\}^{2-\epsilon}} \Big) \\
&= \mathcal{O}\Big(\big\{\mathrm{e}^{\tilde{\theta}x} L(\tilde{\theta})\big\}^{n\epsilon} \big(\sqrt{n}\tilde{\theta}\tilde{\tau}\big)^{1-\epsilon} \Big) \\
&= \mathcal{O}\Big(\big\{(1+\tilde{w})^{-1/2} \exp[2\frac{\tilde{w}}{\sigma^2} - \frac{\tilde{w}^2}{2\sigma^2} - \frac{\tilde{w}}{\sigma^2}]\big\}^{n\epsilon} \big(\frac{n\tilde{w}}{\sigma^2}\big)^{(1-\epsilon)/2} \Big) \to 0.
\end{aligned}$$

This shows the logaritmic efficiency of the estimator $\hat{\beta}_n(x)$ as $x \to 0$.

Consider next the unbiased estimator

$$\hat{\alpha}_n(x) = \widehat{\mathcal{L}}^*(\tilde{\theta}) \mathrm{e}^{\tilde{\theta} S_n} \mathbb{I}\{S_n < nx\},$$

where $\widehat{\mathcal{L}}^*(\tilde{\theta}) = \prod_{i=1}^n \widehat{\mathcal{L}}_i(\tilde{\theta})$. Here $\widehat{\mathcal{L}}_i(\tilde{\theta})$, $i = 1, \ldots, n$, are independent and $\widehat{\mathcal{L}}_i(\tilde{\theta}) = (\sigma/\sigma_0(\tilde{\theta})) L_a(0, \tilde{\theta}) V_i$ with $\mathbb{E}[V_i^2] \leq 1$ by construction. Instead of (15) we have

$$\mathbb{E}_{\tilde{\theta}}\Big[\Big\{\Big(\prod_{i=1}^n V_i\Big) \sqrt{n}\tilde{\theta}\tilde{\tau} \mathrm{e}^{-\sqrt{n}\tilde{\theta}\tilde{\tau}(xi - Z_n)} \mathbb{I}\{Z_n < \xi\}\Big\}^2\Big] = \mathcal{O}\big(\sqrt{n}\tilde{\theta}\tilde{\tau}\big),$$

and

$$\begin{aligned}
\frac{\operatorname{Var}[\hat{\alpha}_n(x)]}{\alpha_n(x)^{2-\epsilon}} &= \mathcal{O}\Big(\frac{\big\{\mathrm{e}^{n\tilde{\theta}x} (\sigma/\sigma_0(\tilde{\theta}))^n L_a(0,\tilde{\theta})^n/(\sqrt{n}\tilde{\theta}\tilde{\tau})\big\}^2 \sqrt{n}\tilde{\theta}\tilde{\tau}}{\big\{\mathrm{e}^{n\tilde{\theta}x} L(\tilde{\theta})^n/(\sqrt{n}\tilde{\theta}\tilde{\tau})\big\}^{2-\epsilon}} \Big) \\
&= \mathcal{O}\Big((1+\tilde{w})^{n/2} \big\{\mathrm{e}^{\tilde{\theta}x} L(\tilde{\theta})\big\}^{n\epsilon} \big(\sqrt{n}\tilde{\theta}\tilde{\tau}\big)^{1-\epsilon} \Big) \\
&= \mathcal{O}\Big((1+\tilde{w})^{n/2} \big\{(1+\tilde{w})^{-1/2} \exp[2\frac{\tilde{w}}{\sigma^2} - \frac{\tilde{w}^2}{2\sigma^2} - \frac{\tilde{w}}{\sigma^2}]\big\}^{n\epsilon} \big(\frac{n\tilde{w}}{\sigma^2}\big)^{(1-\epsilon)/2} \Big) \to 0.
\end{aligned}$$

This shows the logaritmic efficiency of the estimator $\hat{\alpha}_n(x)$.