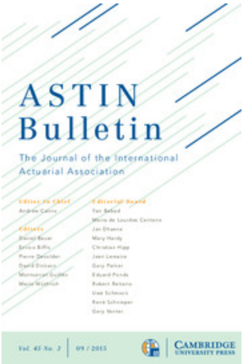

COMPARISON OF APPROXIMATIONS FOR COMPOUND POISSON PROCESSES

Raffaello Seri and Christine Choirat

ASTIN Bulletin / Volume 45 / Issue 03 / September 2015, pp 601 - 637
DOI: 10.1017/asb.2015.7, Published online: 25 June 2015

COMPARISON OF APPROXIMATIONS FOR COMPOUND POISSON PROCESSES

BY

RAFFAELLO SERI AND CHRISTINE CHOIRAT

ABSTRACT

In this paper, we compare the error in several approximation methods for the cumulative aggregate claim distribution customarily used in the collective model of insurance theory. In this model, it is usually supposed that a portfolio is at risk for a time period of length t. The occurrences of the claims are governed by a Poisson process of intensity μ so that the number of claims in $[0, t]$ is a Poisson random variable with parameter $\lambda = \mu t$. Each single claim is an independent replication of the random variable X, representing the claim severity. The aggregate claim or total claim amount process in $[0, t]$ is represented by the random sum of N independent replications of X, whose cumulative distribution function (cdf) is the object of study. Due to its computational complexity, several approximation methods for this cdf have been proposed. In this paper, we consider 15 approximations put forward in the literature that only use information on the lower order moments of the involved distributions. For each approximation, we consider the difference between the true distribution and the approximating one and we propose to use expansions of this difference related to Edgeworth series to measure their accuracy as $\lambda = \mu t$ diverges to infinity. Using these expansions, several statements concerning the quality of approximations for the distribution of the aggregate claim process can find theoretical support. Other statements can be disproved on the same grounds. Finally, we investigate numerically the accuracy of the proposed formulas.

KEYWORDS

Collective risk theory, compound Poisson process, Edgeworth series.

1. INTRODUCTION

The aim of this paper is to provide a comparison of the error in several approximation methods for the cumulative aggregate claim distribution customarily used in the collective model of insurance theory. In this model, it is usually

Astin Bulletin 45(3), 601-637. doi: 10.1017/asb.2015.7

supposed that a portfolio is at risk for a time period of length t. The claims take place according to a Poisson process of intensity μ so that the number of claims in $[0, t]$ is a Poisson random variable with parameter $\lambda = \mu t$. Each single claim is a random variable X_i for $i = 1, \ldots, N$ with a common distribution. We consider the random sum $S_N = \sum_{i=1}^{N} X_i$, i.e. a compound Poisson process representing the aggregate claim or total claim amount process in $[0, t]$. The object of study is therefore the cumulative distribution function (cdf) $\mathbb{P}\{S_N \leq x\}$ (see, e.g., Philipson, 1968).

Since the computation of the function $\mathbb{P}\{S_N \leq x\}$ can be quite complex, some approximations have been introduced. Several of these methods, such as the normal, NP2, NP3, etc., originated from the central limit theorem (see below) and have been introduced as approximations to the whole distribution of the standardized random variable $S_N^\star \triangleq \frac{S_N - \mathbb{E}S_N}{\sqrt{\mathbb{V}(S_N)}}$. As these approximations generated many discussions in the literature as to their relative merits, we will consider the accuracy of all the methods with respect to the whole standardized distribution $F(x) \triangleq \mathbb{P}\{S_N^\star \leq x\}$. However, as Brockett (1983) (see also, e.g., Mikosch, 2006, Section 3.3.4 and Pesonen, 1969a, p. 28) correctly states, what matters in some cases is the tail $\mathbb{P}\{S_N > x\}$ for large x, i.e. a large deviation probability so that the application of a normal approximation based on the central limit theorem may not yield an accurate result. However, as this would take us too far, methods approximating only the right tail of the distribution of S_N (i.e. $\mathbb{P}\{S_N > x\}$ for large x) are not considered here. Moreover, in this paper, we only consider approximations that use information on the lower order moments of the involved distributions. This requirement rules out the Esscher series approximation, since it requires the knowledge (at least approximate) of the characteristic function of the random variable. It also rules out methods of exact computation based on a preliminary discretization of the claim distribution, such as Panjer recursion and the FFT, since they do not only use information on lower order moments. These methods will be analyzed in companion papers. As often happens in the literature, we do not consider here the problems introduced by the computation of the parameters from a set of data.

In this paper, we consider the normal, Edgeworth, NP2, NP2a, Adjusted NP2, NP3, Wilson–Hilferty, Haldane A and B, lognormal, Gamma, translated Gamma, Bowers Gamma, inverse Gaussian and Gamma–IG approximations. For these 15 approximations put forward in the literature, we consider the difference between the true distribution and the approximating one and we propose to use expansions of this difference related to the Edgeworth series to measure their accuracy as $\lambda \to \infty$. Using these expansions, several statements concerning the quality of approximations for the distribution of the aggregate claim process can find theoretical support. Other statements can be disproved on the same grounds. With respect to related uniform bounds similar to the famous Berry–Esséen one, the present approach has three definite advantages. First, it holds also when the accuracy of the method is not uniform throughout the support (as happens, e.g., for what we call, following Pentikäinen, 1977, the NP2a approximation). Second, it provides a different measure of the quality of the approximation in different points of the real line. Third, the expansions are additive, in the sense that the expansion for $F_1(x) - F_2(x)$ can be obtained summing up the expansions for $F_1(x) - F_3(x)$ and $F_3(x) - F_2(x)$ (this property will be occasionally used in the proofs).

Our approach passes through a reappraisal of the Edgeworth expansion. Indeed, the literature on the approximation of the distribution of compound Poisson processes has often dismissed without appeal the Edgeworth expansion both as an approximation method and as a check of adequacy. The first dismissal is quite justified since, as the literature has widely documented and as we will show below, far better approximations exist that are not characterized by an increased computational complexity. The second dismissal is, up to our comprehension of the matter, difficult to justify and peculiar to insurance theory. Indeed, it has often been stressed (see, e.g., Kauppi and Ojantakanen, 1969, p. 222, Beard *et al.*, 1990, p. 108) that the Edgeworth expansion as an infinite series is divergent when the number of terms in the expansion is increased without limit.[1] Nevertheless, the approximation error of the truncated Edgeworth series is uniform throughout the domain of the function and is of an higher order than the last included term. Therefore, we will avoid the problems due to the divergence of the Edgeworth expansion stopping our series at the term of order $\lambda^{-3/2}$. From this point of view, our approach is similar to the one described in Cramér (1946, p. 229) and used in Hall (1983). Summing up, we make ours the following affirmation by DasGupta:

> It is potentially risky to use the Edgeworth expansion to approximate very small tail probabilities. However, it [the Edgeworth expansion] has succeeded in predicting with astounding accuracy the performance of various types of procedures. Instances of these success stories can be seen in the bootstrap literature ... (DasGupta, 2008, p. 218)

The results of this paper concern approximation methods that have been mainly introduced before the development of personal computers. Nevertheless, even if these methods are sometimes quite out-dated and are not always very reliable,[2] we believe that it is important to tell a definitive word concerning their relative performances, especially when the topic has spawned a lot of discussion, such as in the case of the relative merits of NP2 and Gamma approximations or of the supposed independence of the Gamma approximation on the skewness parameter. There is another sense in which our analysis may be of interest to the readers. Some of these methods were introduced as truncations of potentially exact series methods for the computation of the cumulative aggregate claim distribution. The first and simplest case is the whole range of Edgeworth expansions, starting from the normal approximation of Section 3 (with error $O(\lambda^{-1/2})$), the Edgeworth expansion of Section 4 (with error $O(\lambda^{-1})$), and the one that Cramér (1955) called Normal II approximation (with error $O(\lambda^{-3/2})$).

A second example is constituted by the NP2 and NP3 approximations (see respectively Sections 5 and 8) and by the NP4 approximation considered in Bowers (1967) that are given by the inversion of inverse Cornish–Fisher expansions of varying degrees. Finally, the foremost example from this point of view is constituted by the Bowers Gamma approximations that can be seen as truncations of the numerical inversion of Laplace transforms in terms of Laguerre polynomials: the starting point is the Gamma approximation of Section 13, the addition of one term yields the approximation in Heilmann (1988, p. 129) (that we do not consider explicitly here), while the one in Section 15 derives from the addition of two further terms. The transformation of this approximation into an exact method was proposed and dismissed in the actuarial literature (see the discussion in Pfenninger, 1974; Seal, 1975/76; Taylor, 1977). These examples show that several results contained in the present paper can be seen as critiques of truncated series methods (see Heilmann, 1988, Section 3.3) for low orders of the truncation parameter. As an example, our analysis of the Gamma and Bowers Gamma approximations shows that little can be expected from the corresponding exact method as concerns the rate of convergence.

The organization of the paper is as follows. In Section 2, we first provide our main tool, i.e. an Edgeworth expansion for the distribution of compound Poisson processes whose rigorous proof appears to be new. Using this tool, we analyze 15 approximations characterized by the fact that they only use information on the lower order moments of the involved distributions. We therefore consider the normal (Section 3), Edgeworth (Section 4), NP2 (Section 5), NP2a (Section 6), Adjusted NP2 (Section 7), NP3 (Section 8), Wilson–Hilferty (Section 9), Haldane A (Section 10) and B (Section 11), lognormal (Section 12), Gamma (Section 13), translated Gamma (Section 14), Bowers Gamma (Section 15), inverse Gaussian (Section 16) and Gamma–IG (Section 17) approximations. For each approximation, we provide expansions of the error in powers of λ. The methods employed are as disparate as Edgeworth series for compound Poisson processes and for sums of iid random variables, Taylor expansions, and Lagrange's inversion formula. In Section 18, we investigate numerically the accuracy of the proposed formulas. Proofs are gathered together in Appendix A.

2. Mathematical preliminaries

We briefly recall the framework where the claims take place according to a Poisson process of intensity μ and the number of claims in $[0, t]$ is a Poisson random variable with parameter $\lambda = \mu t$. Each single claim is a random variable X_i for $i = 1, \ldots, N$ that is an independent copy of a random variable X. We consider the random sum $S_N = \sum_{i=1}^{N} X_i$ and the normalized sum $S_N^\star \triangleq \frac{S_N - \mathbb{E} S_N}{\sqrt{\mathbb{V}(S_N)}}$. The approximation of $F(x) = \mathbb{P}\left\{S_N^\star \leq x\right\}$ is the object of the paper.

We denote as μ_i the i−th noncentral moment of X. Then the aggregate claim process has moments:

$$\begin{aligned}
\mathbb{E}S_N &= \mu_1 \cdot \lambda \\
\mathbb{V}(S_N) &= \mu_2 \cdot \lambda \\
\mathbb{E}S_N^2 &= \left(\mu_2 + \mu_1^2\lambda\right) \cdot \lambda \\
\mathbb{E}S_N^3 &= \left\{\mu_3 + 3\mu_2\mu_1\lambda + \mu_1^3\lambda^2\right\} \cdot \lambda \\
\kappa_3(S_N) &= \mu_3 \cdot \lambda \\
\kappa_4(S_N) &= \mu_4 \cdot \lambda \\
\kappa_5(S_N) &= \mu_5 \cdot \lambda.
\end{aligned}$$

Moreover, we will write $\gamma_1 \triangleq \frac{\mu_3}{\mu_2^{3/2}\lambda^{1/2}}$ (skewness index), $\gamma_2 \triangleq \frac{\mu_4}{\mu_2^2\lambda}$ (kurtosis index) and $\gamma_3 \triangleq \frac{\mu_5}{\mu_2^{5/2}\lambda^{3/2}}$.

We start our investigation with an Edgeworth expansion of the distribution $F(x)$ for large values of the Poisson parameter λ. Unfortunately, not many results are available in the literature in this direction. Cramér (1955, p. 32) provides a formal Edgeworth expansion up to terms of order λ^{-1}; the result is essentially the same reported in Beard *et al.* (1990, Appendix B). Marsh (1973) considers rigorous Edgeworth expansions of arbitrary order for a class of random sums slightly more general than compound Poisson processes. However, his centering is different from ours, since he (at least implicitly) uses $\frac{S_N - N\mu_1}{\sqrt{\lambda}}$. von Chossy and Rappl (1983) consider expansions of arbitrary order under an additional (restrictive) condition on X. Hipp (1985) considers compound Poisson processes in which the random variable X is replaced by a random vector; moreover he allows for dependence between the waiting times and the vector so that his results are extremely general but quite difficult to use. If a dependence between inter-claim times and claim intensity were to be supposed, his results would be the ideal starting point. Having in mind applications to the bootstrap, Babu *et al.* (2003) consider one-term asymptotic expansions for the more difficult case of studentized means, i.e. $\frac{S_N - \lambda\mu_1}{\sqrt{\sum_{i=1}^N X_i^2}}$.

The next theorem shows that the formal Edgeworth expansion found in the literature can be made rigorous following the lines of Marsh (1973). A more general version of the same theorem is contained in Appendix A.

Theorem 1. *Consider a compound Poisson process with intensity μ observed over $[0, t]$, and let $\lambda = \mu t$. Let Φ be the cumulative distribution function and ϕ the probability density function of a standard normal random variable. Let μ_j be the j-th noncentral moment of the random variable X, and $\varphi(s)$ its characteristic function. Suppose that $\mu_5 < +\infty$ and* $\limsup_{|s|\to\infty} |\varphi(s)| < 1$. *Then, if*

$F(x) = \mathbb{P}\{\frac{S_N - \mathbb{E}S_N}{\sqrt{\mathbb{V}(S_N)}} \leq x\}$:

$$\begin{aligned} F(x) &= \Phi(x) + \phi(x) \cdot \left(1 - x^2\right) \cdot \frac{\gamma_1}{6} + \phi(x) \cdot x \cdot \\ &\quad \times \left\{\left(3 - x^2\right) \cdot \frac{\gamma_2}{24} + \left(10x^2 - x^4 - 15\right) \cdot \frac{\gamma_1^2}{72}\right\} \\ &\quad + \phi(x) \cdot \left\{\left(-x^4 + 6x^2 - 3\right) \cdot \frac{\gamma_3}{120} + \left(-x^6 + 15x^4 - 45x^2 + 15\right) \cdot \frac{\gamma_1 \gamma_2}{144} \right. \\ &\quad \left. + \left(-x^8 + 28x^6 - 210x^4 + 420x^2 - 105\right) \cdot \frac{\gamma_1^3}{1296}\right\} + o\left(\lambda^{-3/2}\right) \end{aligned}$$

where the remainder term is uniform.

Remark 2. (i) *The result in Marsh (1973) for the compound Poisson process can be recovered replacing, in Theorem 1,* X_i *with* $X_i - \mathbb{E}X_i$*, so that* $S_N = \sum_{i=1}^{N} X_i - N \cdot \mathbb{E}X_i$ *and* $\mathbb{E}S_N = 0$. *In that case,* μ_j *has to be replaced in the statement by the corresponding central moment* μ'_j *for* $j = 2, 3, 4$.

(ii) *The difference with respect to the classical Edgeworth expansion is that in that case* μ_j *is replaced by* μ'_j *for* $j = 2, 3$ *and* μ_4 *is replaced by* $\mu'_4 - 3\left(\mu'_2\right)^2$. *Additional terms can be obtained from the classical Edgeworth expansion replacing the cumulants appearing in that case with the noncentral moments.*

Table 1 provides a summary of the results for the 15 approximation methods considered below. The column "Moment" contains the number of moments whose knowledge is required for the computation of the approximation. The column "Order" displays the order of the polynomial representing the approximation error in terms of powers of λ. It is an indicator of the lowest order term of the Edgeworth expansion of $F(x)$ that is not annihilated by the approximation; as an example, the fact that the NP2 approximation has λ^{-1} implies that the term of order $\lambda^{-1/2}$ in $F(x)$ is annihilated by the approximation while the term of order λ^{-1} is not. The column "Leading term" reproduces, for the leading term of the approximation error, the term associated with the highest powers of x (note that, in order to save space, we do not display the normal density $\phi(x)$ multiplying each term). The choice of this measure has a double purpose. First of all, it shows in a compact way which moments of the distribution enter the formula. As an example, the term of order $x^3\lambda^{-1}$ of the NP2 approximation annihilates when $3\gamma_2 = 4\gamma_1^2$. Second, it provides, through the power of x, an indicator of the instability of the tail. The latter point deserves a more thorough explanation. Despite the fact that the Edgeworth series fail to converge or to adequately describe the probability tail for values of x of size $\lambda^{1/6}$ or smaller (see Hall, 1992, p. 325), the presence in the expansion of a higher order polynomial seems to be associated with higher volatility in the tails (see below for several references illustrating this fact). Therefore, provided x is considered as

TABLE 1

SUMMARY OF THE DIFFERENT APPROXIMATIONS.

Method	Moment	Order	Leading term	Domain	LE value
Normal	2	$\lambda^{-1/2}$	$-x^2\frac{\gamma_1}{6}$	$\mathbb{R}$	$-\infty$
Edgeworth	3	λ^{-1}	$-x^3\frac{\gamma_2}{24}$	$\mathbb{R}$	$-\infty$
NP2	3	λ^{-1}	$-\frac{x^3}{6}\left(\frac{\gamma_2}{4}-\frac{\gamma_1^2}{3}\right)$	$\left[\lambda\left(\mu_1-\frac{3\mu_2^2}{2\mu_3}\right)-\frac{\mu_3}{6\mu_2},+\infty\right)$	$\Phi\left(-\frac{3}{\gamma_1}\right)$
NP2a	3	λ^{-1}	$-\frac{x^3}{3}\left(\frac{\gamma_2}{8}-\frac{\gamma_1^2}{3}\right)$	$\mathbb{R}$	$-\infty$
Adjusted NP2	3	λ^{-1}	$-\frac{x^3}{6}\left(\frac{\gamma_2}{4}-\frac{\gamma_1^2}{3}\right)$	comp.	comp.
NP3	4	$\lambda^{-3/2}$	$-\frac{x^4}{3}\left(\frac{\gamma_3}{40}-\frac{\gamma_1\gamma_2}{8}+\frac{\gamma_1^3}{9}\right)$	comp.	comp.
Wilson-Hilferty	3	λ^{-1}	$-\frac{x^3}{12}\left(\frac{\gamma_2}{2}-\frac{7\gamma_1^2}{9}\right)$	$\left[\lambda\mu_1-\lambda\frac{2\mu_2^2}{\mu_3},+\infty\right)$	$\Phi\left(\frac{\gamma_1}{6}-\frac{6}{\gamma_1}\right)$
Haldane A	3	λ^{-1}	$-\frac{x^3}{6}\left(\frac{\gamma_2}{4}+\frac{\mu_3}{3\lambda\mu_1\mu_2}-\frac{5\gamma_1^2}{9}\right)$	$[0,+\infty)$	comp.
Haldane B	4	$\lambda^{-3/2}$	$-\frac{x^4}{8}\left(\frac{\gamma_3}{15}+\frac{\gamma_1\gamma_2}{36}+\frac{\gamma_1^3}{27}-\frac{\gamma_2^2}{8\gamma_1}\right)$	$\left[\lambda\frac{20\mu_1\mu_3^2-9\mu_1\mu_2\mu_4-12\mu_2^2\mu_3}{20\mu_3^2-9\mu_2\mu_4},+\infty\right)$	comp.
Lognormal	2	$\lambda^{-1/2}$	$-\frac{x^2}{2}\left(\frac{\gamma_1}{3}-\frac{\mu_2^{1/2}}{\lambda^{1/2}\mu_1}\right)$	$[0,+\infty)$	0
Gamma	2	$\lambda^{-1/2}$	$-\frac{x^2}{3}\left(\frac{\gamma_1}{2}-\frac{\mu_2^{1/2}}{\lambda^{1/2}\mu_1}\right)$	$[0,+\infty)$	0
Tr. Gamma	3	λ^{-1}	$-\frac{x^3}{8}\left(\frac{\gamma_2}{3}-\frac{\gamma_1^2}{2}\right)$	$\left[\lambda\left(\mu_1-\frac{2\mu_2^2}{\mu_3}\right),+\infty\right)$	0
Bowers Gamma	5	λ^{-1}	$-\frac{x^5}{18}\left(\frac{\mu_2}{\mu_1^2\lambda}+\frac{\gamma_1^2}{4}-\frac{\mu_3}{\mu_1\mu_2\lambda}\right)$	$[0,+\infty)$	0
Inv. Gaussian	3	λ^{-1}	$-\frac{x^3}{72}\left(3\gamma_2-5\gamma_1^2\right)$	$[\lambda(\mu_1-3\mu_2),+\infty)$	0
Gamma–IG	4	$\lambda^{-3/2}$	$-\frac{x^4}{3}\left(\frac{\gamma_3}{40}-\frac{2\gamma_2\gamma_1}{15}+\frac{\gamma_1^3}{8}\right)$	$\left[\lambda\left(\mu_1-\mu_2\max\left\{\frac{2\mu_2}{\mu_3},3\right\}\right),+\infty\right)$	0

fixed (or slowly increasing), this is a useful indicator of the reliability of the approximation in the tails. Supposing that the domain of the random variable X is $[0, +\infty)$, the column "Domain" indicates the domain on which the approximation to the distribution of S_N (and not of $S_N^\star$) is defined, while the column "LE value" provides the value assumed by the approximation at the left endpoint, i.e. LE (provided this is finite). The word "comp.", standing for "complicated", labels those cases in which the left endpoint and/or the value assumed by the function at this point are difficult to manage and uninformative.

3. Normal approximation

The normal approximation is based on an application of Rényi's version of Anscombe central limit theorem (see Gut, 2005, Theorem 3.2). Cramér (1955, p. 30) calls it "Normal I" approximation and credits F. Lundberg with the introduction of this approximation in the first years of the 20th century, as well as with the first criticisms concerning its inadequacy, especially in the tails. In this case, the error is trivially:

$$\begin{aligned} F(x) - \Phi(x) &= -\phi(x) \cdot \left(x^2 - 1\right) \cdot \frac{\gamma_1}{6} - \phi(x) \cdot x \cdot \\ &\quad \times \left\{ \left(x^2 - 3\right) \cdot \frac{\gamma_2}{24} + \left(x^4 - 10x^2 + 15\right) \cdot \frac{\gamma_1^2}{72} \right\} \\ &\quad -\phi(x) \cdot \left\{ \left(x^4 - 6x^2 + 3\right) \cdot \frac{\gamma_3}{120} + \left(x^6 - 15x^4 + 45x^2 - 15\right) \cdot \frac{\gamma_1 \gamma_2}{144} \right. \\ &\quad \left. + \left(x^8 - 28x^6 + 210x^4 - 420x^2 + 105\right) \cdot \frac{\gamma_1^3}{1296} \right\} + o\left(\lambda^{-3/2}\right). \end{aligned}$$

This error holds uniformly for all x.

Several criticisms have been raised concerning this approximation (see, among others, Cramér, 1955, p. 30, Kauppi and Ojantakanen, 1969, p. 214, Papush *et al.*, 2001, p. 181, Beard *et al.*, 1990, p. 105). The above expansion confirms that, for large γ_1, the distribution is particularly badly approximated. As concerns the approximation accuracy of the tail, especially in relative terms, see Section 4.

4. Edgeworth approximation

Up to the authors' knowledge, this approximation first appears in Cramér (1955) where the version including terms up to order λ^{-1} is called "Normal II" approximation. It is considered, among others, in Bohman and Esscher (1963a) where it is correctly attributed to Cramér (1955) and is called "N-method". In the literature, also the Gram-Charlier A expansion has been discussed (see, e.g., Heilmann, 1988, Section 3.3.2): this is simply the Edgeworth series arranged in a different way (according to the order of the derivatives of Φ and not of the powers of $\lambda^{-1/2}$). However, it is well-known that its approximation properties are inferior to the Edgeworth series and it is usually used only for introductory purposes. Therefore, we do not consider it here. In the case of the Edgeworth approximation, the error is:

$$\begin{aligned}
&F(x) - \Phi(x) + \phi(x) \cdot \left(x^2 - 1\right) \cdot \frac{\gamma_1}{6} \\
&= -\phi(x) \cdot x \cdot \left\{ \left(x^2 - 3\right) \cdot \frac{\gamma_2}{24} - \left(x^4 - 10x^2 + 15\right) \cdot \frac{\gamma_1^2}{72} \right\} \\
&\quad -\phi(x) \cdot \left\{ \left(x^4 - 6x^2 + 3\right) \cdot \frac{\gamma_3}{120} + \left(x^6 - 15x^4 + 45x^2 - 15\right) \cdot \frac{\gamma_1 \gamma_2}{144} \right. \\
&\quad \left. + \left(x^8 - 28x^6 + 210x^4 - 420x^2 + 105\right) \cdot \frac{\gamma_1^3}{1296} \right\} + o\left(\lambda^{-3/2}\right).
\end{aligned}$$

The same remarks concerning the normal approximation are also applicable, to a lesser extent, to this approximation (see, e.g., Bohman and Esscher, 1963a, pp. 205–206, Kauppi and Ojantakanen, 1969, p. 221, Pesonen, 1969a, p. 32, Beard *et al.*, 1990, p. 108). Most of these criticisms concern the fact that this approximation only describes accurately a neighborhood of the mean of the distribution. However, the absolute error term in a truncated Edgeworth series (i.e. obtained stopping the sum after a finite number of terms) is uniform so that the error in the approximation of the tail is not expected to be larger than the error in the center of the distribution. Even more can be said: for the classical Edgeworth expansion, the bound is uniformly decreasing in x so that $o\left(\lambda^{-1}\right)$ can be replaced by $o[\lambda^{-1} \cdot (1 + |x|)^4]$. Therefore, the error in the tail is in general smaller than in the center of the distribution. What is true, on the other hand, is that the decrease in the absolute error for large $|x|$ is not sufficient to make the relative error small (see, e.g., Kolassa, 2006, Section 3.6) so that far from the center of the distribution the error can even be larger than 100% of the true value (for a heuristic reasoning, see Pesonen, 1969a, p. 28). By the way, the same thing can be said of the following methods, as no result on relative approximation is known to the authors. However, due to the absence of the term in the fourth power of x in (some of) the following methods, the approximation will be generally better in absolute (and relative) terms.

5. NP2 APPROXIMATION

The NP2 approximation is usually credited to Kauppi and Ojantakanen. Indeed, it was first found by Loimaranta (see Kauppi and Ojantakanen, 1969, p. 219) and described in Kauppi and Ojantakanen (1969). The first uses of the formula seem to be due to Bowers (1967) (who therefore predated Kauppi and Ojantakanen, 1969), Hovinen (1969, pp. 229–230), Pesonen (1969a, p. 31) and Pesonen (1969b), all acknowledging priority to a communication of Kauppi and Ojantakanen at the ASTIN Colloquium in Arnhem in 1966. In this case, we have:

$$\begin{aligned} &F(x) - \Phi\left(\frac{3}{\gamma_1}\left[\sqrt{1+\frac{\gamma_1^2}{9}+\frac{2\gamma_1}{3}\cdot x}-1\right]\right) \\ &\quad = -\phi(x)\cdot x\cdot\left\{\left(x^2-3\right)\cdot\frac{\gamma_2}{24}+\left(-2x^2+5\right)\cdot\frac{\gamma_1^2}{36}\right\} \\ &\qquad -\phi(x)\cdot\left\{\left(x^4-6x^2+3\right)\cdot\frac{\gamma_3}{120}+\left(x^6-15x^4+45x^2-15\right)\cdot\frac{\gamma_1\gamma_2}{144}\right. \\ &\qquad \left. +\left(-6x^6+75x^4-184x^2+49\right)\cdot\frac{\gamma_1^3}{648}\right\}+o\left(\lambda^{-3/2}\right). \end{aligned}$$

This shows that the formula provides a better approximation for the tail when both γ_1 and γ_2 are small, or more generally when $\frac{\gamma_1^2}{3}-\frac{\gamma_2}{4}$ is near to zero (see Kauppi and Ojantakanen, 1969, p. 226, Pesonen, 1969a, p. 32, Berger, 1972, p. 92, Bühlmann, 1974, p. 131, Pentikäinen, 1977, p. 285, Chaubey *et al.*, 1998, p. 230).

Already in Kauppi and Ojantakanen (1969, p. 221 and pp. 224–226) (see also Beard *et al.*, 1990, p. 112), it was realized that the NP2 approximation is far better than the Edgeworth one. This is due to the fact that the polynomial associated with γ_1^2 is $\frac{2x^2-5}{36}$ for the NP approximation and $\frac{-x^4+10x^2-15}{72}$ for the Edgeworth one. Clearly the second one gives a much larger error when x is large.

Our results also allow one to prove the statement of Gendron and Crepeau (1989, p. 257) according to which this approximation usually underestimates the exact value of $F(x,t)$. Indeed, the authors consider mainly tail approximations (the graphs represent probabilities larger than 0.5). In this situation, the NP2 approximation underestimates $F(x,t)$, while the contrary is true for small values of x.

However, the main drawback of the NP2 approximation (and also of the adjusted NP2 and of the NP3) concerns the fact that it is defined only where the square root term is positive (see Hardy, 2004 for a statement, and Beard *et al.*, 1990, pp. 116–117 for corrections to this fact).

6. NP2A APPROXIMATION

This approximation uses $\Phi(x-\frac{\lambda^{-1/2}\mu_3}{6\mu_2^{3/2}}\cdot(x^2-1))$. This is known in Statistics as the first-order "normalizing" Cornish–Fisher expansion (see, e.g., Hill and

Davis, 1968). It seems to us that the first appearance in actuarial sciences is due to Pentikäinen (1977), as a simpler version of the NP2 approximation. We get:

$$\begin{aligned}
&F(x) - \Phi\left(x - \frac{\gamma_1}{6}\cdot\left(x^2-1\right)\right) \\
&\quad = -\phi(x)\cdot x\cdot\left\{\left(x^2-3\right)\cdot\frac{\gamma_2}{24} + \left(-4x^2+7\right)\cdot\frac{\gamma_1^2}{36}\right\} \\
&\qquad -\phi(x)\cdot\left\{\left(x^4-6x^2+3\right)\cdot\frac{\gamma_3}{120} + \left(x^6-15x^4+45x^2-15\right)\cdot\frac{\gamma_1\gamma_2}{144}\right. \\
&\qquad \left. + \left(-6x^6+51x^4-104x^2+26\right)\cdot\frac{\gamma_1^3}{324}\right\} + o\left(\lambda^{-3/2}\right).
\end{aligned}$$

The only difference with respect to the NP2 approximation is that the term $-2x^2+5$ is here replaced by $-4x^2+7$. This means that the NP2 approximation is slightly better than the NP2a approximation in the tail, but this difference is smaller when γ_1^2 is small (see Pentikäinen, 1977, p. 281). However, this approximation is much worse than the previous one for large x: when $\gamma_1 > 0$ and $x \to \infty$, $F(x)$ converges to 1 while $\Phi\left(x - \frac{\gamma_1}{6}\cdot\left(x^2-1\right)\right)$ converges to 0. Indeed, a well-known fact in Statistics is that inverse Cornish–Fisher expansions are "far more accurate" than normalizing Cornish–Fisher expansions and give "very good approximation at the right tail" (see Lee and Lee, 1992, pp. 448–449). Since the NP2 approximation is obtained through the inversion of an inverse Cornish-Fisher expansion (the procedure is described in Lee and Lee, 1992, p. 449, quoting the 1984 edition of Beard *et al.*, 1990 as a reference) and the NP2a is a normalizing expansion, the different behavior in the extreme tail comes as no surprise.

7. ADJUSTED NP2 APPROXIMATION

The adjusted NP2 approximation was introduced by Ramsay (1991). It is based on the computation of b_0 that is the unique root of equation $\gamma_1 = 6b - 4b^3$ lying in the interval $[0, 1/\sqrt{2}]$, and $a_0 = \sqrt{1-2b_0^2}$. Then, the approximation is $\Phi\left(-\frac{a_0}{2b_0} + \sqrt{1+\frac{1}{b_0}\cdot x+\frac{a_0^2}{4b_0^2}}\right)$ and the approximation error turns out to be:

$$\begin{aligned}
&F(x) - \Phi\left(-\frac{a_0}{2b_0} + \sqrt{1+\frac{1}{b_0}\cdot x+\frac{a_0^2}{4b_0^2}}\right) = -\phi(x)\cdot x\left(x^2-3\right)\cdot\left\{\frac{\gamma_2}{24} - \frac{\gamma_1^2}{18}\right\} \\
&\quad -\phi(x)\cdot\left\{\left(x^4-6x^2+3\right)\cdot\frac{\gamma_3}{120} + \left(x^6-15x^4+45x^2-15\right)\cdot\frac{\gamma_1\gamma_2}{144}\right. \\
&\quad \left. + \left(-x^6+13x^4-33x^2+9\right)\cdot\frac{\gamma_1^3}{108}\right\} + o\left(\lambda^{-3/2}\right).
\end{aligned}$$

Note that even if the NP2 and adjusted NP2 methods yield similar approximation errors (see Ramsay, 1991, p. 150), the graphs in Section 18 show that the two errors can have quite different behaviors in practice.

8. NP3 APPROXIMATION

This approximation was proposed in Kauppi and Ojantakanen (1969), but note that already Bowers (1967) used an NP4 approximation (crediting Kauppi and Ojantakanen with it). The approximation is also used in Pesonen (1969b) and Pentikäinen (1977). Consider the cubic (in y) equation $(\frac{\gamma_2}{24}-\frac{\gamma_1^2}{18})\cdot y^3+\frac{\gamma_1}{6}\cdot y^2+(1-\frac{\gamma_2}{8}+\frac{5\gamma_1^2}{36})\cdot y-\left(x+\frac{\gamma_1}{6}\right)=0$. Let $y(x)$ be the value, function of x, solving it. As this equation has in general 3 solutions, it is necessary to select the appropriate root. Following the same method of proof outlined in Appendix A, a rule of thumb is to choose y as the root that is nearest to the development $y=x-\frac{\gamma_1}{6}\cdot(x^2-1)-\frac{\gamma_2}{24}\cdot(x^3-3x)+\frac{\gamma_1^2}{36}\cdot(4x^3-7x)+O(\gamma_1^3)$. Then:

$$F(x)-\Phi(y(x))=-\phi(x)\cdot\left\{\left(x^4-6x^2+3\right)\cdot\frac{\gamma_3}{120}+\left(-x^4+5x^2-2\right)\cdot\frac{\gamma_1\gamma_2}{24}\right.$$
$$\left.+\left(12x^4-53x^2+17\right)\cdot\frac{\gamma_1^3}{324}\right\}+o\left(\lambda^{-3/2}\right).$$

The complexity of the procedure has induced someone to remove this approximation from the list of simple procedures (Seal, 1977, p. 214).

The increase in precision due to the NP3 approximation has been subject to doubt: some authors have suggested that the NP2 and NP3 approximations are so similar that the latter does not deserve particular consideration (see, e.g., Berger, 1972, p. 92), while others have found that the NP3 formula improves over the NP2 one but can, nevertheless, worsen in an irregular way especially in the tails (see, e.g., Beard *et al.*, 1990, p. 116). Our expansion shows that the NP2 and NP3 approximations annihilate respectively terms up to order $\lambda^{-1/2}$ and λ^{-1} so that the latter really improves upon the former. The reason for which the behavior of the NP3 approximation can worsen in the tails is to be found in the fact that the first-order approximation of the error $F(x)-\Phi(y(x))$ behaves as $\phi(x)\,x^4\cdot(-\frac{\gamma_3}{120}+\frac{\gamma_1\gamma_2}{24}-\frac{\gamma_1^3}{27})$ (for $x=o(\lambda^{\frac{1}{6}})$) and this can be quite large: even if $\phi(x)$ decreases exponentially fast while x^4 increases only polynomially, the resulting approximation can be inaccurate for extreme probabilities.

9. WILSON–HILFERTY APPROXIMATION

The Wilson–Hilferty approximation was proposed in 1931 (Wilson and Hilferty, 1931) as an improvement over Fisher's approximation (Fisher, 1928, pp. 96–97) of the χ^2 distribution. Both transformations were originally limited to the χ^2 case. It was Pentikäinen (1987, p. 22), in his study of the Haldane approximations (see below), who extended it to the approximation of the aggregate claim distribution. It can be seen as a version of Haldane A method in which $h = \frac{1}{3}$, namely the value taken by $1 - \frac{\mu_1 \mu_3}{3\mu_2^2}$ for a χ^2. As the Haldane methods (see below), it is better seen as a transformation of S_N yielding an approximately normal distribution. The distance between the two distributions can be analyzed applying the delta method for the Edgeworth expansion (see Skovgaard, 1981) and provides a uniform expansion. However, this would lead to a formula that is not comparable with the other formulas of the paper. Therefore, following Pentikäinen (1987), we invert the transformation in order to provide a modification of the normal quantile with improved properties.

The method approximates $F(x)$ through $\Phi([(1 + \frac{x\gamma_1}{2})^{1/3} - (1 - \frac{\gamma_1^2}{36})] \cdot \frac{6}{\gamma_1})$:

$$
\begin{aligned}
&F(x) - \Phi\left(\left[\left(1 + \frac{x\gamma_1}{2}\right)^{1/3} - \left(1 - \frac{\gamma_1^2}{36}\right)\right] \cdot \frac{6}{\gamma_1}\right) \\
&= -\phi(x) \cdot x\left(x^2 - 3\right) \cdot \left\{\frac{\gamma_2}{24} - \frac{7\gamma_1^2}{108}\right\} \\
&\quad -\phi(x) \cdot \left\{\left(x^4 - 6x^2 + 3\right) \cdot \frac{\gamma_3}{120} + \left(x^6 - 15x^4 + 45x^2 - 15\right) \cdot \frac{\gamma_1\gamma_2}{144}\right. \\
&\quad \left. + \left(-7x^6 + 87x^4 - 208x^2 + 52\right) \cdot \frac{\gamma_1^3}{648}\right\} + o\left(\lambda^{-3/2}\right)
\end{aligned}
$$

The limits and the advantages of the formula have correctly been identified in the literature, where it is often unfavorably compared to the Haldane method presented below (see Pentikäinen, 1987, p. 30, Hardy, 2004).

10. Haldane A approximation

The Haldane A approximation was proposed in Haldane (1938) as a generalization of the Wilson–Hilferty transformation. Pentikäinen (1987) provided an adaptation to the computation of the aggregate claim distribution. It is based on the transformation $(\frac{S_N}{\mathbb{E}S_N})^h$ where h is chosen in such a way as to annihilate the leading term in the skewness index of $(\frac{S_N}{\mathbb{E}S_N})^h$. In practice, we are led to consider:

$$
S_N = \frac{\left(\frac{S_N}{\mu_1\lambda}\right)^h - 1 + \frac{h(h-1)}{2}\left[1 - \frac{(2-h)(1-3h)}{4}\frac{\mu_2}{\mu_1^2}\lambda^{-1}\right]\frac{\mu_2}{\mu_1^2}\lambda^{-1}}{h\frac{\sqrt{\mu_2}}{\mu_1}\lambda^{-1/2}\sqrt{1 - \frac{(1-h)(1-3h)}{2}\frac{\mu_2}{\mu_1^2}\lambda^{-1}}}
$$

where $h \triangleq 1 - \frac{\gamma_1 \mu_1 \lambda^{1/2}}{3\sqrt{\mu_2}} = 1 - \frac{\mu_1 \mu_3}{3\mu_2^2}$ (the case $h = 0$ has to be covered separately, see Pentikäinen, 1987, p. 21, Hardy, 2004). S'_N has a distribution that can be approximated by a normal one. We stick to the formulation used by Pentikäinen (1987, pp. 22–23) for Haldane B approximation. Haldane A approximation uses $\Phi(\frac{(1+bx)^h - m_y}{\sigma_y})$ where

$$m_y = 1 - \frac{1}{2} c\,(b-c)\left[1 + \frac{1}{4}(2b-c)(3c-b)\right]$$

$$\sigma_y = |c|\sqrt{1 + \frac{1}{2}(b-c)(3c-b)}.$$

For Haldane A method $b = \frac{\mu_2^{1/2}}{\mu_1}\lambda^{-1/2}$, $c = \frac{\mu_2^{1/2}}{\mu_1}\lambda^{-1/2} - \frac{\gamma_1}{3}$ and $h = \frac{c}{b} = 1 - \frac{\mu_1\mu_3}{3\mu_2^2}$. We have:

$$\begin{aligned}
&F(x) - \Phi\left(\frac{(1+bx)^h - m_y}{\sigma_y}\right) \\
&= -\phi(x)\cdot x\cdot\left(x^2-3\right)\cdot\left(\frac{\gamma_2}{24} + \frac{\mu_3}{18\mu_1\mu_2\lambda} - \frac{5\gamma_1^2}{54}\right) \\
&\quad -\phi(x)\cdot\Bigg\{\left(x^4 - 6x^2 + 3\right)\cdot\frac{\gamma_3}{120} + \left(x^6 - 15x^4 + 45x^2 - 15\right)\cdot\frac{\gamma_1\gamma_2}{144} \\
&\quad + \left(-5x^6 + 54x^4 - 113x^2 + 26\right)\cdot\frac{\gamma_1^3}{324} \\
&\quad + \left(2x^6 - 11x^4 + 12x^2 - 9\right)\cdot\frac{\mu_3^2}{216\mu_1\mu_2^{5/2}\lambda^{3/2}} \\
&\quad + \left(-x^4 + 3\right)\cdot\frac{\mu_3}{36\mu_1^2\mu_2^{1/2}\lambda^{3/2}}\Bigg\} + o\left(\lambda^{-3/2}\right).
\end{aligned}$$

The performance of the method is quite good, especially in the right tail (see, e.g., Hardy, 2004).

11. Haldane B approximation

The method was introduced by Haldane (1938) and is based on the transformation $(\frac{S_N + g - \mathbb{E}S_N}{g})^h$ where h and g are chosen in such a way to annihilate the leading terms in the skewness and in the kurtosis indices. It was adapted to the present case by Pentikäinen (1987). For the Haldane B method, taking $b = \frac{5}{3}\gamma_1 - \frac{3\gamma_2}{4\gamma_1}$,

$c = \frac{4}{3}\gamma_1 - \frac{3\gamma_2}{4\gamma_1}$ and $h = \frac{c}{b} = \frac{16\gamma_1^2 - 9\gamma_2}{20\gamma_1^2 - 9\gamma_2}$, we get:

$$\begin{aligned} F(x) - \Phi\left(\frac{(1+bx)^h - m_y}{\sigma_y}\right) &\\ = -\phi(x)\cdot\Bigg\{\left(x^4 - 6x^2 + 3\right)\cdot\frac{\gamma_3}{120} + \left(x^4 + 78x^2 - 81\right)\cdot\frac{\gamma_1\gamma_2}{288} &\\ + \left(6x^4 - 332x^2 + 314\right)\cdot\frac{\gamma_1^3}{1296} + \left(-x^4 + 3\right)\cdot\frac{\gamma_2^2}{64\gamma_1}\Bigg\} + o\left(\lambda^{-3/2}\right). \end{aligned}$$

Already Haldane remarked that in some cases this approximation is not better than the previous Haldane A approximation. In Pentikäinen (1987, p. 29), it is remarked that the Haldane B approximation gives remarkably precise results when the skewness γ_1 is small, but deteriorates rapidly as γ_1 increases.

12. Lognormal approximation

The lognormal distribution has a long history as a claim distribution (see, e.g., Papush *et al.*, 2001 for references), but it is quite difficult to reconstruct its history as a distribution for the aggregate claim amount. A reference is Heilmann (1988, p. 124). It approximates $\frac{S_N - \mathbb{E}S_N}{\sqrt{\mathbb{V}(S_N)}}$ through $\frac{\mathcal{LN}(m,s^2) - e^{m+\frac{s^2}{2}}}{\sqrt{e^{2m+s^2}(e^{s^2}-1)}}$ with $m = \ln(\mu_1^2\lambda^2) - \frac{1}{2}\ln\left(\mu_1^2\lambda^2 + \mu_2\lambda\right)$ and $s^2 = \ln(1 + \frac{\mu_2}{\mu_1^2\lambda})$, yielding:

$$\begin{aligned} F(x) - \Phi\left\{\frac{\ln^{\frac{1}{2}}\left(1 + \frac{\mu_2}{\mu_1^2}\lambda^{-1}\right)}{2} + \frac{\ln\left(1 + x\cdot\frac{\sqrt{\mu_2}}{\mu_1}\lambda^{-\frac{1}{2}}\right)}{\ln^{\frac{1}{2}}\left(1 + \frac{\mu_2}{\mu_1^2}\lambda^{-1}\right)}\right\} &\\ = -\phi(x)\cdot\left(x^2 - 1\right)\cdot\left\{\frac{\gamma_1}{6} - \frac{\mu_2^{1/2}}{2\mu_1\lambda^{1/2}}\right\} &\\ -\phi(x)\cdot\Bigg\{\left(x^3 - 3x\right)\cdot\frac{\gamma_2}{24} + \left(x^5 - 10x^3 + 15x\right)\cdot\frac{\gamma_1^2}{72} &\\ + \left(-3x^5 + 14x^3 + 3x\right)\cdot\frac{\mu_2}{24\mu_1^2\lambda}\Bigg\} &\\ -\phi(x)\cdot\Bigg\{\left(x^4 - 6x^2 + 3\right)\cdot\frac{\gamma_3}{120} + \left(x^6 - 15x^4 + 45x^2 - 15\right)\cdot\frac{\gamma_1\gamma_2}{144} & \end{aligned}$$

$$+\left(x^8-28x^6+210x^4-420x^2+105\right)\cdot\frac{\gamma_1^3}{1296}$$
$$\left.+\left(-x^8+12x^6-20x^4-8x^2-7\right)\cdot\frac{\mu_2^{3/2}}{48\mu_1^3\lambda^{3/2}}\right\}+o\left(\lambda^{-3/2}\right).$$

This approximation does not even annihilate the $\lambda^{-1/2}$ term.

13. Gamma approximation

This approximation was introduced in Bartlett (1965), where the author proposed to use a (non-translated) Gamma random variable whose first two moments agreed with the compound Poisson process. Bartlett's paper had a certain impact since also Bowers (1966), Thompson (1969) and Beekman (1969) considered related problems in the light of Bartlett approximation, but its use declined after the revival of the translated Gamma distribution (see below). Despite the improvement generally constituted by the latter, the non-translated two parameter Gamma is sometimes still used (see, e.g., Papush *et al.*, 2001), mainly because it has the same domain as the approximated random variable.

The idea is to approximate the centered and normalized sum using the random variable $\frac{\mu_2^{1/2}}{2\mu_1\lambda^{1/2}}(\chi^2_{\frac{2\lambda\mu_1^2}{\mu_2}}-\frac{2\lambda\mu_1^2}{\mu_2})$, where χ_n^2 is a shortcut for $\chi_n^2\sim\Gamma(\frac{n}{2},\frac{1}{2})$ with positive real n:

$$F(x)-\mathbb{P}\left\{\frac{\mu_2^{1/2}}{2\mu_1\lambda^{1/2}}\left(\chi^2_{\frac{2\lambda\mu_1^2}{\mu_2}}-\frac{2\lambda\mu_1^2}{\mu_2}\right)\le x\right\}$$
$$=-\phi(x)\cdot\left(x^2-1\right)\cdot\left\{\frac{\gamma_1}{6}-\frac{\mu_2^{1/2}}{3\mu_1\lambda^{1/2}}\right\}-\phi(x)\cdot\left\{\left(x^3-3x\right)\cdot\left(\frac{\gamma_2}{24}-\frac{\mu_2}{4\mu_1^2\lambda}\right)\right.$$
$$\left.+\left(x^5-10x^3+15x\right)\cdot\left(\frac{\gamma_1^2}{72}-\frac{\mu_2}{18\mu_1^2\lambda}\right)\right\}$$
$$-\phi(x)\cdot\left\{\left(x^4-6x^2+3\right)\cdot\left(\frac{\gamma_3}{120}-\frac{\mu_2^{3/2}}{5\mu_1^3\lambda^{3/2}}\right)\right.$$
$$+\left(x^6-15x^4+45x^2-15\right)\cdot\left(\frac{\gamma_1\gamma_2}{144}-\frac{\mu_2^{3/2}}{12\mu_1^3\lambda^{3/2}}\right)$$
$$\left.+\left(x^8-28x^6+210x^4-420x^2+105\right)\cdot\left(\frac{\gamma_1^3}{1296}-\frac{\mu_2^{3/2}}{162\mu_1^3\lambda^{3/2}}\right)\right\}+o\left(\lambda^{-3/2}\right).$$

14. Translated Gamma approximation

This approximation was introduced by Bohman and Esscher (1963a,b) under the name "G-method". Independently, in the discussion of Bartlett (1965) (where the Gamma approximation was introduced, see above), Jones remarked that in a related situation (i.e. in the individual risk model) Taylor had already used in 1952 a translated Gamma distribution whose first three moments were equal to the corresponding moments of the true distribution. Thus, he proposed to do the same in the collective risk model situation considered by Bartlett. Despite the interest raised by Bartlett's paper, Jones' improvement received little or no attention. The translated Gamma approximation was at last revived and popularized by Seal (1977).

The idea is to approximate the centered and normalized sum using the random variable $\frac{\gamma_1}{4}(\chi^2_{\frac{8}{\gamma_1^2}} - \frac{8}{\gamma_1^2})$:

$$
\begin{aligned}
F(x) - \mathbb{P}\left\{\frac{\gamma_1}{4}\left(\chi^2_{\frac{8}{\gamma_1^2}} - \frac{8}{\gamma_1^2}\right) \le x\right\}
&= -\phi(x)\cdot\left(x^3 - 3x\right)\cdot\left(\frac{\gamma_2}{24} - \frac{\gamma_1^2}{16}\right) - \phi(x)\cdot\left\{\left(x^4 - 6x^2 + 3\right)\cdot\left(\frac{\gamma_3}{120} - \frac{\gamma_1^3}{40}\right)\right. \\
&\quad \left. + \left(x^6 - 15x^4 + 45x^2 - 15\right)\cdot\left(\frac{\gamma_1\gamma_2}{144} - \frac{\gamma_1^3}{96}\right)\right\} + o\left(\lambda^{-3/2}\right).
\end{aligned}
$$

The approximation has proved to be very accurate, especially in the right tail (see Bohman and Esscher, 1963a, p. 207, Papush *et al.*, 2001, p. 181), while in the left tail it may give positive probability to negative losses (see Papush *et al.*, 2001, p. 178, Hardy, 2004).

The comparison between the NP2 and the Gamma approximations generated a discussion in the literature (see, e.g., Gendron and Crepeau, 1989, p. 255) as to which method was better, with Seal (1977, pp. 214–215) as a partisan of the Gamma approximation and Pentikäinen (1977, p. 285) of the NP2 one. Our results show that there is little to choose from the point of view of the approximation error. Perhaps the main difference is that the Gamma approximation is uniform (in the sense that the error is uniformly of order λ^{-1} over the whole line) while as concerns the NP there is no proof of this kind of result.

Starting from Seal (1977, p. 215), another topic that raised a certain debate is the supposed independence of the error of the Gamma approximation from γ_1 (with, once more, Pentikäinen, 1977, p. 285 on the opposite side, see Gendron and Crepeau, 1989, p. 257). However, this claim is not supported by our derivations, since the remainder term depends on γ_1.

As concerns the dependence on x, the presence of a systematic pattern in $F(x) - \mathbb{P}\{\frac{\gamma_1}{4}(\chi^2_{\frac{8}{\gamma_1^2}} - \frac{8}{\gamma_1^2}) \leq x\}$ has already been remarked in empirical studies (see Gendron and Crepeau, 1989, p. 257, Papush *et al.*, 2001, p. 181) and is confirmed by our formula. This can be seen also in the graphs of Section 18. Note that the graph of the Gamma approximation is quite similar to the one in Buckley and Eagleson (1988, p. 155) and in Choirat and Seri (2013, p. 2147).

15. Bowers Gamma approximation

This kind of approximation has a longer history in Statistics than in Insurance (see Seal, 1975/76 for an historical account). In the latter field, this approximation has been introduced in Bowers (1966), as a Laguerre expansion around the (non-translated) Gamma approximation given above, much in the same way that the Edgeworth series is an expansion in Hermite polynomials around the normal distribution (see Heilmann, 1988, Section 3.3 for a unified presentation). Despite the fact that the series can be truncated at any desired order (provided the moments exist, see Hardy, 2004), Bowers mainly considers the following case, matching the first five moments of the original distribution. Pfenninger (1974), apparently unaware of Bowers' paper, has extended the treatment to higher order expansions. The method has been criticized in Seal (1975/76), where the full history of the method in Statistics is recalled and in Taylor (1977). Heilmann (1988, p. 129) considers an approximation similar to the one below, with $B = C = 0$.

Let $\alpha = \frac{\lambda\mu_1^2}{\mu_2}$. Then, the approximation is:[3]

$$\begin{aligned}
F_B(x) &= \Gamma\left(\alpha + \sqrt{\alpha}x, \alpha\right) - A \cdot \left(\alpha + \sqrt{\alpha}x\right)^{\alpha} e^{-\left(\alpha+\sqrt{\alpha}x\right)} \\
&\quad \cdot \left\{ \frac{1}{\Gamma(\alpha+1)} - \frac{2\left(\alpha + \sqrt{\alpha}x\right)}{\Gamma(\alpha+2)} + \frac{\left(\alpha + \sqrt{\alpha}x\right)^2}{\Gamma(\alpha+3)} \right\} \\
&\quad + B \cdot \left(\alpha + \sqrt{\alpha}x\right)^{\alpha} e^{-\left(\alpha+\sqrt{\alpha}x\right)} \\
&\quad \cdot \left\{ \frac{1}{\Gamma(\alpha+1)} - \frac{3\left(\alpha + \sqrt{\alpha}x\right)}{\Gamma(\alpha+2)} + \frac{3\left(\alpha + \sqrt{\alpha}x\right)^2}{\Gamma(\alpha+3)} - \frac{\left(\alpha + \sqrt{\alpha}x\right)^3}{\Gamma(\alpha+4)} \right\} \\
&\quad - C \cdot \left(\alpha + \sqrt{\alpha}x\right)^{\alpha} e^{-\left(\alpha+\sqrt{\alpha}x\right)} \\
&\quad \cdot \left\{ \frac{1}{\Gamma(\alpha+1)} - \frac{4\left(\alpha+\sqrt{\alpha}x\right)}{\Gamma(\alpha+2)} + \frac{6\left(\alpha + \sqrt{\alpha}x\right)^2}{\Gamma(\alpha+3)} - \frac{4\left(\alpha + \sqrt{\alpha}x\right)^3}{\Gamma(\alpha+4)} + \frac{\left(\alpha + \sqrt{\alpha}x\right)^4}{\Gamma(\alpha+5)} \right\}
\end{aligned}$$

where:

$$A = \frac{\left(\mu_1\mu_3 - 2\mu_2^2\right)\mu_1^2\lambda}{6\mu_2^3}$$

$$B = \frac{\left(\mu_1^2\mu_4 - 12\mu_2\mu_1\mu_3 + 18\mu_2^3\right)\mu_1^2\lambda}{24\mu_2^4}$$

$$C = \frac{\left(\mu_1^3\mu_5 - 20\mu_2\mu_1^2\mu_4 + 120\mu_2^2\mu_1\mu_3 - 144\mu_2^4\right)\mu_1^2\lambda}{120\mu_2^5}.$$

The method yields:

$$\begin{aligned}
&F(x) - F_B(x) \\
&= -\phi(x)\cdot\left(x^5 - 10x^3 + 15x\right)\cdot\left(\frac{\mu_2}{18\mu_1^2\lambda} + \frac{\gamma_1^2}{72} - \frac{\mu_3}{18\mu_1\mu_2\lambda}\right) \\
&\quad -\phi(x)\cdot\left\{\left(x^6 - 15x^4 + 45x^2 - 15\right)\cdot\left(\frac{\gamma_1\gamma_2}{144} - \frac{\mu_4}{72\mu_1\mu_2^{3/2}\lambda^{3/2}}\right)\right. \\
&\quad + \left(x^8 - 28x^6 + 210x^4 - 420x^2 + 105\right)\cdot\frac{\gamma_1}{1296} \\
&\quad + \left(-2x^8 + 47x^6 - 285x^4 + 435x^2 - 75\right)\cdot\frac{\mu_3}{216\mu_1^2\mu_2^{1/2}\lambda^{3/2}} \\
&\quad \left. + \left(4x^8 - 85x^6 + 435x^4 - 465x^2 + 15\right)\cdot\frac{\mu_2^{3/2}}{324\mu_1^3\lambda^{3/2}}\right\} + o\left(\lambda^{-3/2}\right).
\end{aligned}$$

As we will see in Section 18, the fit of this method is quite poor despite its complexity (see Seal, 1975/76, p. 133, Hardy, 2004). The order of approximation of the method is the same of the translated Gamma method, but its fit in the right tail is worse, since the approximation error depends on the fourth and fifth powers of x while it depends only on the third power for the translated Gamma method.

16. Inverse Gaussian approximation

We use the notation IG (m, b) as parameterized in Chaubey *et al.* (1998). If x_0 is the left endpoint of the shifted inverse Gaussian distribution (not explicitly considered in the original paper), its mean is $m + x_0$, its variance mb and its skewness $3\sqrt{\frac{b}{m}}$. We match the first three moments taking $m = \frac{3\mu_2^2\lambda}{\mu_3}$, $b = \frac{\mu_3}{3\mu_2}$ and $x_0 = \mu_1\lambda - m$. The skewness of the approximating distribution is $\frac{\mu_3}{\mu_2^{3/2}\lambda^{1/2}}$, and the kurtosis is $\frac{15b}{m} = \frac{5\mu_3^2}{3\lambda\mu_2^3}$. Therefore, we get:

$$\begin{aligned}
&F(x) - \mathbb{P}\left\{\frac{\mathrm{IG}(m,b) - m}{\sqrt{mb}} \leq x\right\} \\
&= -\phi(x) \cdot \left(x^3 - 3x\right) \cdot \frac{3\gamma_2 - 5\gamma_1^2}{72} - \phi(x) \cdot \left\{\left(x^4 - 6x^2 + 3\right) \cdot \left(\frac{\gamma_3}{120} - \frac{7\gamma_1^3}{216}\right)\right. \\
&\quad \left. + \left(x^6 - 15x^4 + 45x^2 - 15\right) \cdot \left(\frac{\gamma_1\gamma_2}{144} - \frac{5\gamma_1^3}{432}\right)\right\} + o\left(\lambda^{-3/2}\right).
\end{aligned}$$

Our computations show that, according to the second-order Edgeworth expansion, the inverse Gaussian and the Gamma approximations are of comparable accuracy. The leading polynomial appearing in both expansions is $x\left(3 - x^2\right)$, while the coefficient is $\frac{\gamma_2}{24} - \frac{5\gamma_1^2}{72}$ for the IG and $\frac{\gamma_2}{24} - \frac{\gamma_1^2}{16}$ for the Gamma. Some tedious algebra shows that the IG yields a smaller absolute error when $\gamma_1^2 < \frac{12}{19} \cdot \gamma_2$, while in the reverse case the contrary is true. This does not rule out the possibility that the term of order $\lambda^{-3/2}$ reverses the order. This virtual equivalence between Gamma and IG approximations has already been acknowledged in the literature (see Chaubey *et al.*, 1998, p. 230).

17. Gamma–IG approximation

A further approximation, proposed in Chaubey *et al.* (1998), is obtained as a linear combination of the Gamma and the IG approximation given above. The idea is to use:

$$\begin{aligned}
&F(x) - w\mathbb{P}\left\{\frac{\gamma_1}{4}\left(\chi^2_{\frac{8}{\gamma_1^2}} - \frac{8}{\gamma_1^2}\right) \leq x\right\} - (1-w)\,\mathbb{P}\left\{\frac{\mathrm{IG}(m,b) - m}{\sqrt{mb}} \leq x\right\} \\
&\quad = -\phi(x) \cdot \left(x^4 - 6x^2 + 3\right) \cdot \left(\frac{\gamma_1^3}{24} + \frac{\gamma_3}{120} - \frac{2\gamma_2\gamma_1}{45}\right) + o\left(\lambda^{-3/2}\right)
\end{aligned}$$

where $w = \frac{\gamma_2 - \frac{5\gamma_1^2}{3}}{\frac{3\gamma_1^2}{2} - \frac{5\gamma_1^2}{3}} = \frac{10\gamma_1^2 - 6\gamma_2}{\gamma_1^2}$. The error is uniformly $o\left(\lambda^{-1}\right)$. This proves that the accuracy of both Gamma and IG approximations is uniformly improved by taking a particular linear combination of the two (see Chaubey *et al.*, 1998, p. 230).

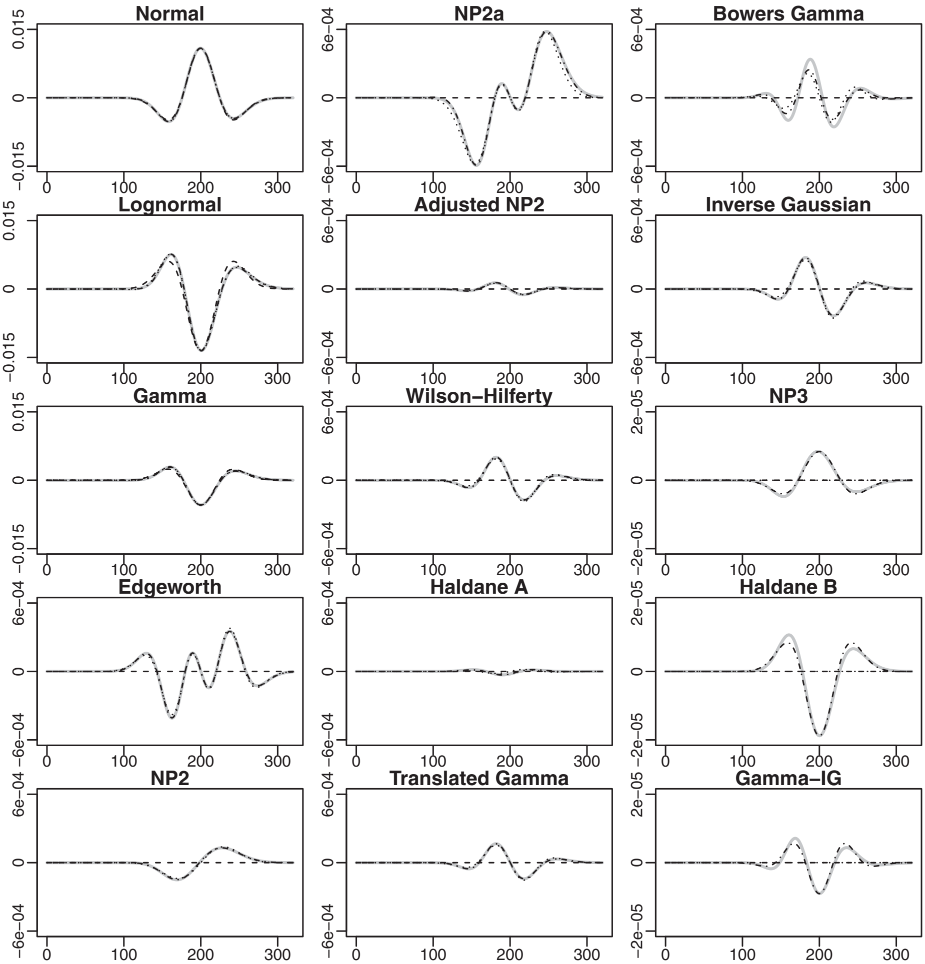

FIGURE 1: Approximation errors for $\lambda = 100$ and shape parameter equal to 2.

18. COMPUTATIONS

In this section, we reproduce a limited computational study to show the relevance of the previous results. Our aim is not to provide an empirical guidance for which method is better in which situation (for this see Reijnen *et al.*, 2005), but only to assess whether and when the theoretical formulas provided above offer a reliable picture of the true distribution.

We consider four alternative situations in which the random variable X is a Gamma random variable with scale parameter 1 and shape parameter equal to 2 or 0.1. The results of the analysis are plotted in Figures 1–4. The 15 approximations have been arranged in order of increasing theoretical precision from

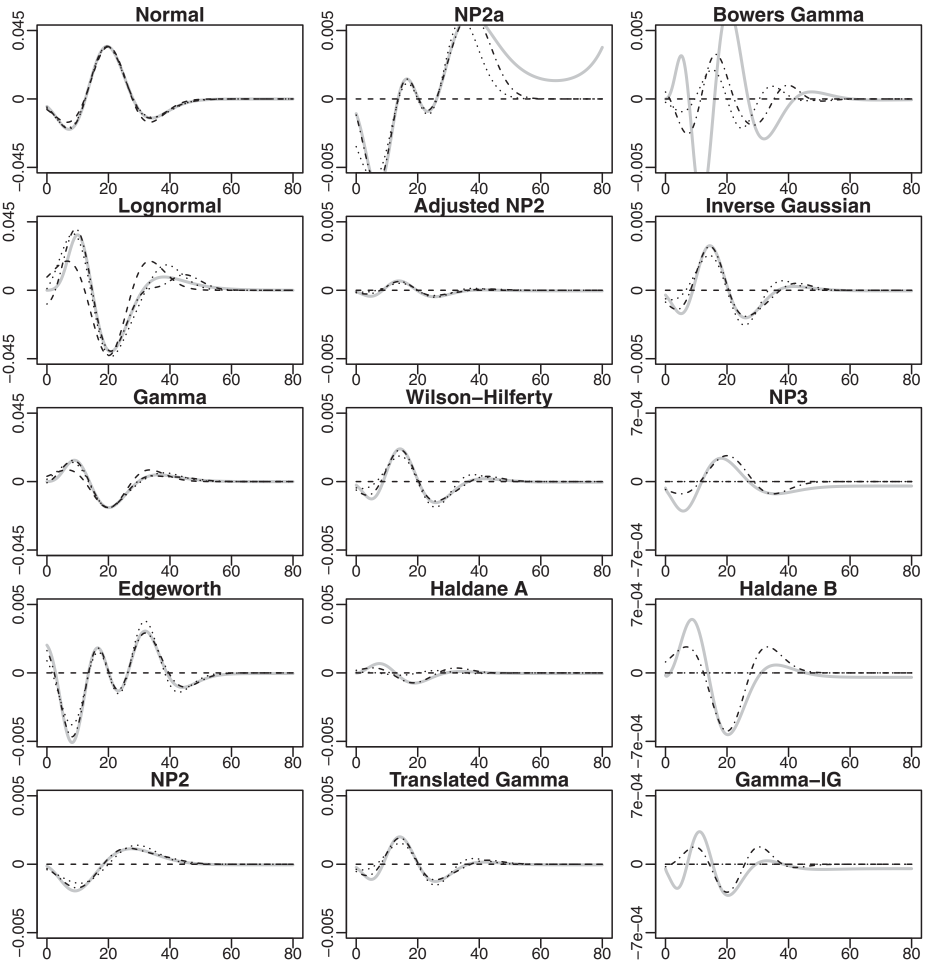

FIGURE 2: Approximation errors for $\lambda = 10$ and shape parameter equal to 2.

the left to the right and from top to down. Therefore, we first display the normal, lognormal and Gamma approximations, characterized by an error of order $\lambda^{-1/2}$. Then, we represent the Edgeworth, NP2, NP2a, Adjusted NP2, Wilson–Hilferty, Haldane A, translated Gamma, Bowers Gamma and inverse Gaussian approximations, all annihilating terms up to order $\lambda^{-1/2}$. At last, we display the NP3, Haldane B and Gamma–IG approximations, whose error is of order $\lambda^{-3/2}$. While in the text we considered the normalized cdf, namely $\mathbb{P}\left\{S_N^\star \leq x\right\}$, here we transpose the results to the nonnormalized cdf $\mathbb{P}\{S_N \leq x\}$. On the y-axis, we represent, for each value x, the difference between the true cdf $\mathbb{P}\{S_N \leq x\}$ and the approximate one (in gray), the approximation of this difference up to order $\lambda^{-1/2}$ (in dashed line), up to order λ^{-1} (in dotted line) and up to order $\lambda^{-3/2}$ (in dash-dot line). To increase comparability among approximations, the range of

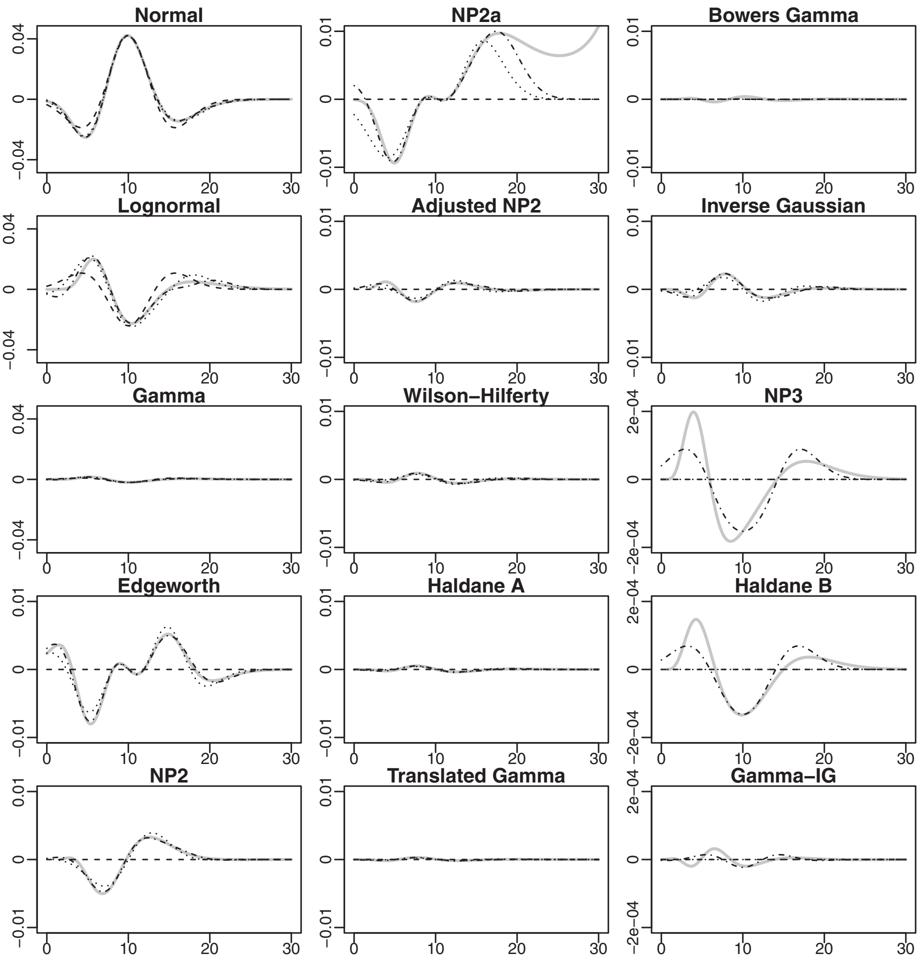

FIGURE 3: Approximation errors for $\lambda = 100$ and shape parameter equal to 0.1.

the y axis is the same for all the approximations with the same order of accuracy. The curve $x \mapsto \mathbb{P}\{S_N \le x\}$ is computed through the method described in Appendix A, with an absolute error of at most $1 \cdot 10^{-15}$. We have also estimated $x \mapsto \mathbb{P}\{S_N \le x\}$ through the empirical cdf based on n points with n ranging from $n = 1 \cdot 10^{10}$ to $n = 1.2 \cdot 10^{11}$, but the results were indistinguishable (apart from the cases in which the empirical approximation had some fluctuations).

Here are the parameters used in the computations, together with the values taken by the indexes γ_1, γ_2 and γ_3:

- $\lambda = 100$, shape parameter 2 (Figure 1): $\gamma_1 = 0.1632993$, $\gamma_2 = 0.03333333$, $\gamma_3 = 0.008164966$;

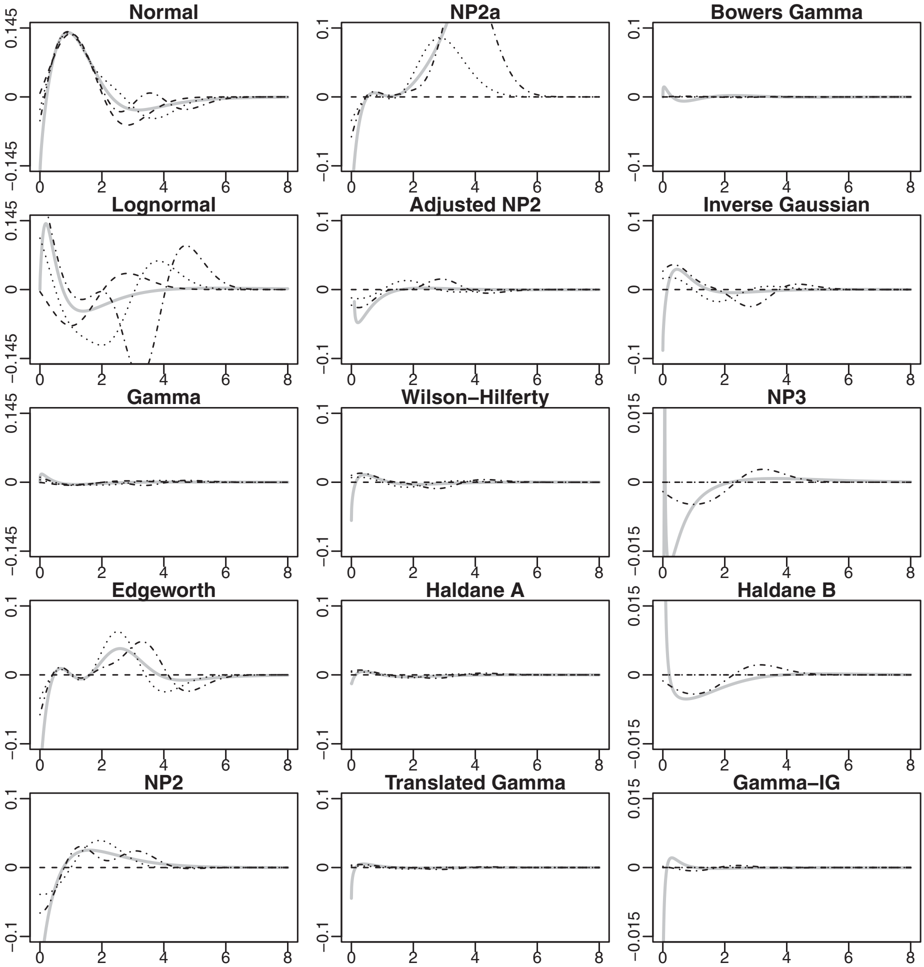

FIGURE 4: Approximation errors for λ = 10 and shape parameter equal to 0.1.

- $\lambda = 10$, shape parameter 2 (Figure 2): $\gamma_1 = 0.5163978$, $\gamma_2 = 0.3333333$, $\gamma_3 = 0.2581989$;
- $\lambda = 100$, shape parameter 0.1 (Figure 3): $\gamma_1 = 0.6331738$, $\gamma_2 = 0.5918182$, $\gamma_3 = 0.7316036$;
- $\lambda = 10$, shape parameter 0.1 (Figure 4): $\gamma_1 = 2.002271$, $\gamma_2 = 5.918182$, $\gamma_3 = 23.13534$.

Note that the values of γ_1 as computed from Table 1 in Seal (1977) are 0.811501, 1.213861, 1.215901, 2.009974, 2.613988, 3.504514 and 3.838496, while those in Table 1 of Pentikäinen (1977) are 0.0671, 0.1570, 0.2122, 0.3879, 0.4543, 0.5570, 0.7749, 0.8115, 1.2139, 1.2159, 1.5286, 1.7615, 1.8564, 2.7318, 3.4504 and 3.8385 (some of them are repeated in the two lists). Therefore, the proposed illustrations should reproduce quite faithfully most of the skewness values encountered in practice.

A first obvious remark concerns the fact that the expansions we provide are much better the larger λ, because the $o\left(\lambda^{-3/2}\right)$ remainder term appearing in the above formulas is smaller. The entity of the first neglected term of the Edgeworth expansion is larger the larger are γ_1 and γ_2: in the present illustration, this happens when the shape parameter is small. In particular when γ_1 and γ_2 are very large our expansions are not very precise. This is particularly evident for the Bowers Gamma approximation; this seems to be due to the complexity of the approximating function.

Moreover, our expansions are more reliable for the methods that match only the lower orders of the Edgeworth approximations so that the NP3, Haldane B and Gamma–IG curves are quite roughly approximated by our expansions even for large λ. However, also in these cases the message is quite clear, since the Edgeworth expansions provide the correct magnitude of the error especially in the center and in the right tail of the distribution.

A striking fact that was not evident from the formulas (but is reliably reproduced by their graphs) is the difference between NP2, NP2a and Adjusted NP2: NP2 and Adjusted NP2, despite the similarity in the formulas, have often very different behaviors and NP2a is reliable only for very small values of γ_1 and not too large values of x. The (non-translated) Gamma approximation is particularly bad, but the translated Gamma approximation provides a much better fit.

A fact that was expected from the formulas is the similarity between the translated Gamma and the IG approximations, as is the extreme precision of the Gamma–IG approximation.

19. Conclusions

In this paper, we have compared several methods for the approximation of compound Poisson distributions. The error of these methods has been studied using series related to Edgeworth expansions. This allows us to formulate some general rules for the approximation of these distributions.

The best approximation results are obtained using the Gamma–IG, Haldane B and NP3 approximations that require knowledge or estimation of four moments of the original claim distribution. The numerical computations suggest that the Gamma–IG method has an advantage over the other two methods.

Slightly worse results are obtained through the Edgeworth, NP2, NP2a, adjusted NP2, Wilson–Hilferty, Haldane A, translated Gamma and inverse Gaussian approximations using three moments of the original distribution. Among them, the NP2a approximation has a tail behavior that is different from the one of the original distribution and should be avoided as far as possible. The Bowers Gamma achieves the same order of accuracy but requires five moments; moreover, its performance is quite unreliable. As a general rule, we confirm the better performance of the NP2 approximation with respect to the Edgeworth one, as well as the qualitative equivalence of the NP2, Gamma and inverse Gaussian approximations.

At the bottom of the scale, the normal, lognormal and Gamma approximations require only two moments but perform quite poorly, except in exceedingly large samples.

All in all, infinitely divisible distributions matching three or more moments of the original distribution, namely Gamma–IG, translated Gamma, inverse Gaussian, perform well with respect to similar approximations.

Despite these general instructions provide a quite clear ranking of the various methods, their performance can still be volatile. It is therefore advisable that, whenever possible, the researcher interested in the computation of a compound Poisson distribution uses exact methods, whose feasibility has been steadily increasing in the last decades thanks to the progress in computing power.

Notes

1. Conditions for convergence have been worked out in the classical case by Cramér (1946, p. 223) and Feller (1971, p. 542), and are indeed very restrictive.
2. However, note that also exact computation methods, such as the FFT and Panjer recursion, are based on a preliminary estimation of the claim distribution (see, e.g., Lee and Lin, 2010 for references) and on its discretization. The effects of these two steps are not completely understood, despite some important steps have been done (see Grübel and Hermesmeier, 1999, 2000; Embrechts and Frei, 2009).
3. Note that formula (7) in Hardy (2004) contains a slip since the sign before C should be a minus.

References

Abramowitz, M. and Stegun, I.A. (1964) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards Applied Mathematics Series, vol. 55. Washington, DC: U.S. Government Printing Office.

Babu, G.J., Singh, K. and Yang, Y. (2003) Edgeworth expansions for compound Poisson processes and the bootstrap. *Annals of the Institute of Statistical Mathematics*, **55**, 83–94.

Bartlett, D.K. (1965) Excess ratio distribution in risk theory. *Transactions of the Society of Actuaries*, **XVII**, 435–463.

Beard, R.E., Pentikäinen, T. and Pesonen, E. (1990) *Risk Theory: The Stochastic Basis of Insurance*. London: Chapman and Hall.

Beekman, J.A. (1969) A ruin function approximation. *Transactions of the Society of Actuaries*, **XXI**, 41–48.

Berger, G. (1972) Integration of the normal power approximation. *ASTIN Bulletin*, **7**, 90–96.

Bohman, H. and Esscher, F. (1963a) Studies in risk theory with numerical illustrations concerning distribution functions and stop-loss premiums. *Skandinavisk Aktuarietidskrift*, **46**, 173–225.

Bohman, H. and Esscher, F. (1963b) Studies in risk theory with numerical illustrations concerning distribution functions and stop-loss premiums. *Skandinavisk Aktuarietidskrift*, **46**, 1–40.

Bowers, N.L. (1966) Expansion of probability density functions as a sum of gamma densities with applications to risk theory. *Transactions of the Society of Actuaries*, **XVIII**, 125–147.

Bowers, N.L. (1967) An approximation to the distribution of annuity costs. *Transactions of the Society of Actuaries*, **XIX**, 295–309.

BROCKETT, P.L. (1983) On the misuse of the central limit theorem in some risk calculations. *The Journal of Risk and Insurance*, **50**, 727–731.

BUCKLEY, M.J. and EAGLESON, G.K. (1988) An approximation to the distribution of quadratic forms in normal random variables. *Australian Journal of Statistics*, **30A**, 150–159.

BÜHLMANN, H. (1974) Book review of "Two stochastic processes, by J. A. Beekman. Published by Almqvist & Wiksell, Stockholm." *ASTIN Bulletin*, **8**, 131–132.

CHAUBEY, Y.P., GARRIDO, J. and TRUDEAU, S. (1998) On the computation of aggregate claims distributions: some new approximations. *Insurance: Mathematics and Economics*, **23**, 215–230.

CHOIRAT, C. and SERI, R. (2013) Computational aspects of Cui-Freeden statistics for equidistribution on the sphere. *Mathematics of Computation*, **82**, 2137–2156.

CRAMÉR, H. (1946) *Mathematical Models of Statistics*. Princeton, NJ: Princeton University Press.

CRAMÉR, H. (1955) *Collective Risk Theory: A Survey of the Theory from the Point of View of the Theory of Stochastic Processes*. Reprinted from the Jubilee Volume of Försäkringsaktiebolaget Skandia. Stockholm: Skandia Insurance Company.

DASGUPTA, A. (2008) *Asymptotic Theory of Statistics and Probability.* Springer Texts in Statistics. New York, NY: Springer.

EMBRECHTS, P. and FREI, M. (2009) Panjer recursion versus FFT for compound distributions. *Mathematical Methods of Operations Research*, **69**, 497–508.

FELLER, W. (1971) *An Introduction to Probability Theory and its Applications. Vol II. 2nd ed.* Wiley Series in Probability and Mathematical Statistics. New York, NY: John Wiley and Sons, Inc.

FISHER, R.A. (1928) *Statistical Methods for Research Workers,* 2nd ed. Biological methods and manuals No. 5. Edinburgh: Oliver and Boyd.

GENDRON, M. and CREPEAU, H. (1989) On the computation of the aggregate claim distribution when individual claims are inverse Gaussian. *Insurance: Mathematics and Economics*, **8**, 251–258.

GRÜBEL, R. and HERMESMEIER, R. (1999) Computation of compound distributions I: Aliasing errors and exponential tilting. *ASTIN Bulletin*, **29**, 197–214.

GRÜBEL, R. and HERMESMEIER, R. (2000) Computation of compound distributions II: Discretization errors and Richardson extrapolation. *ASTIN Bulletin*, **30**, 309–331.

GUT, A. (2005) *Probability: A Graduate Course.* Springer Texts in Statistics. New York, NY: Springer.

HALDANE, J.B.S. (1938) The approximate normalization of a class of frequency distributions. *Biometrika*, **29**, 392–404.

HALL, P. (1983) Chi squared approximations to the distribution of a sum of independent random variables. *Annals of Probability*, **11**, 1028–1036.

HALL, P. (1992) *The Bootstrap and Edgeworth Expansion*. Springer Series in Statistics. New York: Springer-Verlag.

HARDY, M.R. (2004) Approximating the aggregate claims distribution. In *Encyclopedia of Actuarial Science*, pp. 76–86. New York, NY: John Wiley and Sons.

HEILMANN, W.-R. (1988) *Fundamentals of Risk Theory*. Karlsruhe: Verlag Versicherungs wirtschaft.

HILL, G.W. and DAVIS, A.W. (1968) Generalized asymptotic expansions of Cornish-Fisher type. *Annals of Mathematical Statistics*, **39**, 1264–1273.

HIPP, C. (1985) Asymptotic expansions in the central limit theorem for compound and Markov processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, **69**, 361–385.

HOVINEN, E. (1969) Procedures and basic statistics to be used in magnitude control of equalisation reserves in Finland. *ASTIN Bulletin*, **5**, 227–238.

KAUPPI, L. and OJANTAKANEN, P. (1969) Approximations of the generalised Poisson function. *ASTIN Bulletin*, **5**, 213–228.

KOLASSA, J.E. (2006) *Series Approximation Methods in Statistics*, 3rd ed. Lecture Notes in Statistics, vol. 88. New York, NY: Springer.

LEE, S.C.K. and LIN, X.S. (2010) Modeling and evaluating insurance losses via mixtures of Erlang distributions. *North American Actuarial Journal*, **14**, 107–130.

LEE, Y.-S. and LEE, M.C. (1992) On the derivation and computation of the Cornish-Fisher expansion. *Australian Journal of Statistics*, **34**, 443–450.

MARSH, L.M. (1973) An expansion for the distribution function of a random sum. *Journal of Applied Probability*, **10**, 869–874.
MIKOSCH, T. (2006) *Non-Life Insurance Mathematics: An Introduction with Stochastic Processes*. Berlin: Springer Verlag.
OLVER, F.W.J., LOZIER, D.W., BOISVERT, R.F. and CLARK, C.W. ed. (2010) *NIST Handbook of Mathematical Functions*. U.S. Department of Commerce National Institute of Standards and Technology, Washington, DC.
PAPUSH, D.E., PATRIK, G.S. and PODGAITS, F. (2001) Approximations of the aggregate loss distribution. *Casualty Actuarial Society Forum*, **Winter**, 175–186.
PENTIKÄINEN, T. (1977) On the approximation of the total amount of claims. *ASTIN Bulletin*, **9**, 281–289.
PENTIKÄINEN, T. (1987) Approximative evaluation of the distribution function of aggregate claims. *ASTIN Bulletin*, **17**, 15–39.
PESONEN, E. (1969a) Computation of the Poisson process by means of divergent series. *Skandinavisk Aktuarietidskrift*, **Suppl. 3**, 25–33.
PESONEN, E. (1969b) NP-approximation of risk processes. *Skandinavisk Aktuarietidskrift*, **Suppl. 3**, 63–69.
PFENNINGER, F. (1974) Eine neue Methode zur Berechnung der Ruinwahrscheinlichkeit mittels Laguerre-Entwicklung. *Blätter der Deutschen Gesellschaft für Versicherungsmathematik*, **11**, 491–532.
PHILIPSON, C. (1968) A review of the collective theory of risk. *Skandinavisk Aktuarietidskrift*, **1–2**, 45–68.
RAMSAY, C.M. (1991) A note of the normal power approximation. *ASTIN Bulletin*, **21**, 147–150.
REIJNEN, R., ALBERS, W. and KALLENBERG, W.C.M. (2005) Approximations for stop-loss reinsurance premiums. *Insurance: Mathematics and Economics*, **36**, 237–250.
SEAL, H.L. (1975/76) A note on the use of Laguerre polynomials in the inversion of Laplace transforms. *Blätter der Deutschen Gesellschaft für Versicherungsmathematik*, **12**, 131–134.
SEAL, H.L. (1977) Approximations to risk theory's $F(x, t)$ by means of the gamma distribution. *ASTIN Bulletin*, **9**, 213–218.
SKOVGAARD, I.M. (1981) Transformation of an Edgeworth expansion by a sequence of smooth functions. *Scandinavian Journal of Statistics*, **8**, 207–217.
TAYLOR, G.C. (1977) Concerning the use of Laguerre polynomials for inversion of Laplace transforms in risk theory. *Blätter der Deutschen Gesellschaft für Versicherungsmathematik*, **13**, 85–92.
THOMPSON, P. (1969) The standard deviation of excess losses. *Transactions of the Society of Actuaries*, **XIX**, 267–272.
TRICOMI, F. G. and ERDÉLYI, A. (1951) The asymptotic expansion of a ratio of gamma functions. *Pacific Journal of Mathematics*, **1**, 133–142.
VON CHOSSY, R. and RAPPL, G. (1983) Some approximation methods for the distribution of random sums. *Insurance: Mathematics and Economics*, **2**, 251–270.
WHITTAKER, E.T. and WATSON, G.N. (1996) *A Course of Modern Analysis*. Cambridge Mathematical Library. Reprint of the fourth (1927) edition. Cambridge: Cambridge University Press.
WILSON, E.B. and HILFERTY, M.M. (1931) The distribution of chi-square. *Proceedings of the National Academy of Sciences*, **17**, 684–688.

RAFFAELLO SERI (Corresponding author)
Dipartimento di Economia,
Università degli Studi dell'Insubria,
Via Monte Generoso 71, 21100 Varese, Italy
E-Mail: raffaello.seri@uninsubria.it

CHRISTINE CHOIRAT
IQSS Harvard University, 1737 Cambridge Street,
CGIS Knafel Building, Cambridge, MA 02138, USA
E-Mail: cchoirat@iq.harvard.edu

APPENDIX A. PROOFS

A.1. Edgeworth expansion for compound Poisson random variables

For the following result we need the definition of the Hermite polynomials $\{He_n(x)\}$, i.e. the set of polynomials orthogonal with respect to the weighting function $e^{-\frac{x^2}{2}}$:

$$\int_{\mathbb{R}} He_m(x)\, He_n(x)\, e^{-\frac{x^2}{2}}\, \mathrm{d}x = \sqrt{2\pi}\,\delta_{mn} n!,$$

where δ_{mn} is Kronecker's delta, i.e. the function taking value 1 if $m = n$ and 0 otherwise. In order to avoid possible confusions, it is important to remark that we are considering here the probabilistic Hermite polynomials $\{He_n(x)\}$, and not the ones customarily used in mathematics or physics, often denoted as $\{H_n(x)\}$ (these notations are used, e.g., in Abramowitz and Stegun, 1964, Table 22.2, p. 775 or in Olver *et al.*, 2010, Table 18.3.1, p. 439). The first Hermite polynomials are $He_0(x) = 1$, $He_1(x) = x$, $He_2(x) = x^2 - 1$, $He_3(x) = x^3 - 3x$ and $He_4(x) = x^4 - 6x^2 + 3$. For a longer list, see Olver *et al.* (2010, 18.5.19, p. 443) and Abramowitz and Stegun (1964, Table 22.12).

Theorem 3. *Consider a compound Poisson random variable with parameter* λ. *Let* μ_j *be the* j*-th noncentral moment of the random variable* X, *and* $\varphi(t)$ *its characteristic function. Suppose that* $\mu_s < +\infty$ *with* $s \geq 3$, *and* $\limsup_{|t|\to\infty} |\varphi(t)| < 1$. *Then, if* $F(x) = \mathbb{P}\{\frac{S_N - \mathbb{E}S_N}{\sqrt{\mathbb{V}(S_N)}} \leq x\}$:

$$F(x) = \Phi(x) + \sum_{\nu=1}^{s-2} \lambda^{-\frac{\nu}{2}} \cdot \sum_{\{k_m\}_\nu} \phi(x)\, He_{\nu+2k-1}(x) \cdot \prod_{m=1}^{\nu} \frac{1}{k_m!} \left(\frac{\mu_{m+2}}{(m+2)!\mu_2^{\frac{m+2}{2}}} \right)^{k_m} + o\left(\lambda^{-\frac{s-2}{2}}\right)$$

where $\sum_{\{k_m\}_\nu}$ *indicates that summation is performed over all non-negative integer solutions of* $k_1 + 2k_2 + \cdots + \nu k_\nu = \nu$, $k = k_1 + \cdots + k_\nu$, *and the remainder term is uniform.*

Proof of Theorem 3. 1. We start remarking that the characteristic function of $\frac{S_N - \mathbb{E}S_N}{\sqrt{\mathbb{V}(S_N)}}$ is $\omega(t\lambda^{-1/2}) \triangleq \exp\{\lambda \cdot (\varphi(t \cdot \lambda^{-1/2}\mu_2^{-1/2}) - 1) - it\lambda^{1/2}\mu_1\mu_2^{-1/2}\}$. Define $\xi(t\lambda^{-1/2}) \triangleq t^{-2} \cdot \{\lambda \cdot (\varphi(t \cdot \lambda^{-1/2}\mu_2^{-1/2}) - 1) - it\lambda^{1/2}\mu_1\mu_2^{-1/2} + t^2/2\}$ so that $\ln \omega(t) = t^2\xi(t) - t^2/2$, and $\xi_s\left(t\lambda^{-1/2}\right) \triangleq \sum_{j=3}^{s+2} \frac{i^j}{j!} \cdot t^{j-2} \cdot \lambda^{-j/2+1}\mu_j\mu_2^{-j/2}$. We approximate $\omega\left(t\lambda^{-1/2}\right)$ through:

$$\omega_s\left(t\lambda^{-1/2}\right) \triangleq e^{-\frac{1}{2}t^2} \cdot \sum_{k=0}^{s} \frac{1}{k!} \cdot \left\{\xi_s\left(t\lambda^{-1/2}\right)\right\}^k = e^{-\frac{1}{2}t^2} \cdot \sum_{k=0}^{s} \frac{1}{k!} \cdot \left\{ \sum_{j=3}^{s+2} \frac{i^j}{j!} \cdot t^j \lambda^{-\frac{j}{2}+1} \mu_j \mu_2^{-\frac{j}{2}} \right\}^k .$$

The distance $|\omega(t) - \omega_s(t)|$ can be bounded using Lemma 3 in Marsh (1973):

$$\begin{aligned}
&\left|\omega\left(t\lambda^{-1/2}\right) - \omega_s\left(t\lambda^{-1/2}\right)\right| \\
&= e^{-\frac{1}{2}t^2} \cdot \left| e^{t^2\xi\left(t\lambda^{-1/2}\right)} - \sum_{k=0}^{s} \frac{1}{k!} \cdot \left\{ \sum_{j=3}^{s+2} \frac{i^j}{j!} \cdot t^j \lambda^{-\frac{j}{2}+1} \mu_j \mu_2^{-\frac{j}{2}} \right\}^k \right| \\
&\leq e^{-\frac{1}{2}t^2} e^c \cdot \left\{ \left| t^2\xi\left(t\lambda^{-1/2}\right) - t^2\xi_s\left(t\lambda^{-1/2}\right)\right| + \frac{\left|t^2\xi_s\left(t\lambda^{-1/2}\right)\right|^{s+1}}{(s+1)!} \right\}. \qquad \text{(A.1)}
\end{aligned}$$

Now, for any $\varepsilon > 0$ there is a $\delta > 0$ such that:

$$\left|\xi\left(t\lambda^{-1/2}\right) - \xi_s\left(t\lambda^{-1/2}\right)\right| \leq \varepsilon \left|t\lambda^{-1/2}\right|^s$$

when $|t| \leq \delta\lambda^{1/2}$. Moreover:

$$\begin{aligned}
&\left|t^2\xi_s\left(t\lambda^{-1/2}\right)\right| \leq t^2 \left|\xi\left(t\lambda^{-1/2}\right) - \xi_s\left(t\lambda^{-1/2}\right)\right| + t^2 \left|\xi\left(t\lambda^{-1/2}\right)\right| \\
&\leq \varepsilon |t|^{2+s} \lambda^{-s/2} + \lambda \cdot \left|\varphi\left(t \cdot \lambda^{-1/2}\mu_2^{-1/2}\right) - 1 - it\lambda^{-1/2}\mu_1\mu_2^{-1/2} + t^2\lambda^{-1}/2\right| \\
&\leq \varepsilon |t|^{2+s} \lambda^{-s/2} + \frac{1}{6} |t|^3 \lambda^{-1/2}\mu_3\mu_2^{-3/2}
\end{aligned}$$

when $|t| \leq \delta\lambda^{1/2}$. Therefore the term in braces in Equation (A.1) is majorized by a constant times $|t|^s \lambda^{-s/2}$ for $|t| \leq \delta\lambda^{1/2}$. Reasoning as in Marsh (1973), the result is obtained. ■

A.2. NP2 approximation

We expand $\sqrt{1 + \frac{\gamma_1^2}{9} + \frac{2\gamma_1}{3} \cdot x}$ in a Taylor series in powers of $\lambda^{-1/2}$, obtaining:

$$\begin{aligned}
\frac{3}{\gamma_1}\left[\sqrt{1 + \frac{\gamma_1^2}{9} + \frac{2\gamma_1}{3} \cdot x} - 1\right] &= x + \left(1 - x^2\right) \cdot \frac{\gamma_1}{6} + x\left(x^2 - 1\right) \cdot \frac{\gamma_1^2}{18} \\
&\quad - \left(1 - 6x^2 + 5x^4\right) \cdot \frac{\gamma_1^3}{216} + O\left(\lambda^{-2}\right).
\end{aligned}$$

Then we do the same with $\Phi\left(\frac{3}{\gamma_1}\left[\sqrt{1 + \frac{\gamma_1^2}{9} + \frac{2\gamma_1}{3} \cdot x} - 1\right]\right)$:

$$\begin{aligned}
&\Phi\left(\frac{3}{\gamma_1}\left[\sqrt{1 + \frac{\gamma_1^2}{9} + \frac{2\gamma_1}{3} \cdot x} - 1\right]\right) \\
&= \Phi(x) + \phi(x)\left(1 - x^2\right) \cdot \frac{\gamma_1}{6} + \phi(x)\, x\left(-x^4 + 6x^2 - 5\right) \cdot \frac{\gamma_1^2}{72} \\
&\quad + \phi(x)\left(-x^8 + 16x^6 - 60x^4 + 52x^2 - 7\right) \cdot \frac{\gamma_1^3}{1296} + O\left(\lambda^{-2}\right).
\end{aligned}$$

A.3. NP2a approximation

A limited development in powers of $\lambda^{-1/2}$ yields:

$$\Phi\left(x - \frac{\gamma_1}{6} \cdot \left(x^2 - 1\right)\right) = \Phi(x) - \phi(x)\left(x^2 - 1\right) \cdot \frac{\gamma_1}{6} - \phi(x)\, x\left(x^2 - 1\right)^2 \cdot \frac{\gamma_1^2}{72}$$
$$-\phi(x)\left(x^2 - 1\right)^4 \cdot \frac{\gamma_1^3}{1296} + O\left(\lambda^{-2}\right).$$

A.4. Adjusted NP2 approximation

We first expand $\sqrt{1+\frac{1}{b_0}\cdot x+\frac{a_0^2}{4b_0^2}}$ around $b_0 = 0$, getting:

$$-\frac{a_0}{2b_0} + \sqrt{1 + \frac{1}{b_0} \cdot x + \frac{a_0^2}{4b_0^2}}$$
$$= x + \left(1 - x^2\right) \cdot b_0 + x\left(-1 + 2x^2\right) \cdot b_0^2 + x^2\left(3 - 5x^2\right) \cdot b_0^3 + O\left(b_0^4\right).$$

A further limited development yields:

$$\Phi\left(-\frac{a_0}{2b_0} + \sqrt{1 + \frac{1}{b_0} \cdot x + \frac{a_0^2}{4b_0^2}}\right)$$
$$= \Phi(x) + \phi(x)\left(1 - x^2\right) \cdot b_0 + \phi(x)\, x\left(-x^4 + 6x^2 - 3\right) \cdot \frac{b_0^2}{2}$$
$$+\phi(x)\left(-x^8 + 16x^6 - 54x^4 + 28x^2 - 1\right) \cdot \frac{b_0^3}{6} + O\left(b_0^4\right).$$

From Ramsay (1991, Equation (4)), we have $\gamma_1 = 6b_0 - 4b_0^3$ or $b_0 = \frac{1}{6}\gamma_1 + \frac{1}{324}\gamma_1^3 + O\left(\gamma_1^5\right)$ and:

$$\Phi\left(-\frac{a_0}{2b_0} + \sqrt{1 + \frac{1}{b_0} \cdot x + \frac{a_0^2}{4b_0^2}}\right)$$
$$= \Phi(x) + \phi(x)\left(1 - x^2\right) \cdot \frac{\gamma_1}{6} + \phi(x)\, x\left(-x^4 + 6x^2 - 3\right) \cdot \frac{\gamma_1^2}{72}$$
$$+\phi(x)\left(-x^8 + 16x^6 - 54x^4 + 24x^2 + 3\right) \cdot \frac{\gamma_1^3}{1296} + O\left(\lambda^{-2}\right).$$

A.5. NP3 approximation

In order to find the y that solves $x = y + \frac{\gamma_1}{6} \cdot (y^2 - 1) + \frac{\gamma_2}{24} \cdot (y^3 - 3y) - \frac{\gamma_1^2}{36} \cdot (2y^3 - 5y)$, we use Lagrange's inversion formula (see Whittaker and Watson, 1996, p. 133). In the previous source, we make the identifications $\zeta \triangleq y$, $a \triangleq x$, $t \triangleq \lambda^{-1/2}$, $f \triangleq \Phi$ and $\varphi(\zeta) \triangleq -\frac{\mu_3}{6\mu_2^{3/2}} \cdot$

$\left(\zeta^2-1\right)-\frac{\mu_4}{24\mu_2^2\lambda^{1/2}}\cdot\left(\zeta^3-3\zeta\right)+\frac{\mu_3^2}{36\mu_2^3\lambda^{1/2}}\cdot\left(2\zeta^3-5\zeta\right)$. At last, we get:

$$\begin{aligned}
\Phi(y) = {}& \Phi(x) - \phi(x)\left(x^2-1\right)\cdot\frac{\gamma_1}{6} - \phi(x)\left(x^3-3x\right)\cdot\frac{\gamma_2}{24} \\
& + \phi(x)\,x\left(-x^4+10x^2-15\right)\cdot\frac{\gamma_1^2}{72} \\
& - \phi(x)\left(x^6-9x^4+15x^2-3\right)\cdot\frac{\gamma_2\gamma_1}{144} \\
& + \phi(x)\left(-x^8+28x^6-162x^4+208x^2-37\right)\cdot\frac{\gamma_1^3}{1296} + o\left(\lambda^{-3/2}\right).
\end{aligned}$$

A.6. Wilson–Hilferty approximation

The proof simply proceeds expanding $\left(1+\frac{x\gamma_1}{2}\right)^{1/3}$ around $\gamma_1=0$:

$$\begin{aligned}
\left(1+\frac{x\gamma_1}{2}\right)^{1/3} &= 1 + x\frac{\gamma_1}{6} - x^2\frac{\gamma_1^2}{36} + x^3\frac{5\gamma_1^3}{648} - x^4\frac{5\gamma_1^4}{1944} + O\left(\gamma_1^5\right) \\
\left[\left(1+\frac{x\gamma_1}{2}\right)^{1/3} - \left(1-\frac{\gamma_1^2}{36}\right)\right]\frac{6}{\gamma_1} &= x - \left(x^2-1\right)\frac{\gamma_1}{6} + x^3\frac{5\gamma_1^2}{108} - x^4\frac{5\gamma_1^3}{324} + O\left(\gamma_1^4\right)
\end{aligned}$$

and $\Phi\left(x-\left(x^2-1\right)\frac{\gamma_1}{6}+x^3\frac{5\gamma_1^2}{108}-x^4\frac{5\gamma_1^3}{324}+O\left(\gamma_1^4\right)\right)$ around $\gamma_1=0$.

A.7. Haldane A approximation

From the formulas it is evident that both b and c are $O\left(\lambda^{-1/2}\right)$, while $h=O(1)$. Therefore:

$$\begin{aligned}
(1+bx)^h = {}& 1 + hbx + \frac{h(h-1)}{2}b^2x^2 + \frac{h(h-1)(h-2)}{6}b^3x^3 \\
& + \frac{h(h-1)(h-2)(h-3)}{24}b^4x^4 + O\left(b^5\right) \\
m_y = {}& 1 - \frac{1}{2}b^2h(1-h) - \frac{1}{8}b^4h(1-h)(2-h)(3h-1) \\
\sigma_y^{-1} = {}& c^{-1}\left(1+\frac{b^2}{2}(1-h)(3h-1)\right)^{-1/2} \\
= {}& (hb)^{-1}\left\{1 - \frac{b^2}{4}(1-h)(3h-1) + \frac{3}{32}b^4(1-h)^2(3h-1)^2 + O\left(b^6\right)\right\}
\end{aligned}$$

and:

$$\begin{aligned}\frac{(1+bx)^h - m_y}{\sigma_y} &= x + b\frac{(h-1)}{2}\left(x^2-1\right) \\ &\quad + b^2\frac{(h-1)}{12}x\left[2\,(h-2)\,x^2 + 3\,(3h-1)\right] \\ &\quad + b^3\frac{(h-1)}{24}\left\{(h-2)\left[(h-3)\,x^4 + 3\,(3h-1)\right]\right. \\ &\quad \left. + 3\,(3h-1)\,(h-1)\left(x^2-1\right)\right\} + O\left(b^4\right).\end{aligned}$$

Expanding $\Phi(\frac{(1+bx)^h - m_y}{\sigma_y})$ around $\Phi(x)$ and replacing b, c and h with their expressions, we get the final result:

$$\begin{aligned}&\Phi\left(\frac{(1+bx)^h - m_y}{\sigma_y}\right) \\ &= \Phi(x) - \lambda^{-1/2}\phi(x)\left(x^2-1\right)\frac{\mu_3}{6\mu_2^{3/2}} \\ &\quad - \lambda^{-1}\phi(x)\,x\left\{\left(3x^4 - 10x^2 - 15\right)\frac{\mu_3^2}{216\mu_2^3} - \left(x^2-3\right)\frac{\mu_3}{18\mu_1\mu_2}\right\} \\ &\quad - \lambda^{-3/2}\phi(x)\left\{\left(x^8 - 8x^6 - 6x^4 + 32x^2 + 1\right)\frac{\mu_3^3}{1296\mu_2^{9/2}}\right. \\ &\quad \left. + \left(-2x^6 + 11x^4 - 12x^2 + 9\right)\frac{\mu_3^2}{216\mu_1\mu_2^{5/2}} + \left(x^4-3\right)\frac{\mu_3}{36\mu_1^2\mu_2^{1/2}}\right\} + O\left(\lambda^{-2}\right).\end{aligned}$$

A.8. Haldane B approximation

The method of proof is the same as for Haldane A approximation, but the values of b, c and h differ. For Haldane B approximation, we finally get:

$$\begin{aligned}&\Phi\left(\frac{(1+bx)^h - m_y}{\sigma_y}\right) \\ &= \Phi(x) - \phi(x)\left(x^2-1\right)\frac{\gamma_1}{6} - \phi(x)\,x\left\{\left(x^4 - 10x^2 + 15\right)\frac{\gamma_1^2}{72} + \left(x^2-3\right)\frac{\gamma_2}{24}\right\} \\ &\quad - \phi(x)\left\{\left(x^8 - 28x^6 + 204x^4 - 88x^2 - 209\right)\frac{\gamma_1^3}{1296}\right. \\ &\quad \left. + \left(2x^6 - 31x^4 + 12x^2 + 51\right)\frac{\gamma_1\gamma_2}{288} - \left(3 - x^4\right)\frac{\gamma_2^2}{64\gamma_1}\right\} + o\left(\lambda^{-3/2}\right).\end{aligned}$$

A.9. Lognormal approximation

The series expansion is obtained expanding around $\lambda^{-1} = 0$, first:

$$\begin{aligned}
&\frac{\ln^{\frac{1}{2}}\left(1+\frac{\mu_2}{\mu_1^2}\lambda^{-1}\right)}{2} + \frac{1}{\ln^{\frac{1}{2}}\left(1+\frac{\mu_2}{\mu_1^2}\lambda^{-1}\right)} \cdot \ln\left(1 + x \cdot \frac{\mu_2^{1/2}}{\mu_1}\lambda^{-1/2}\right) \\
&\quad = x - \left(x^2-1\right)\frac{\mu_2^{1/2}}{2\mu_1}\lambda^{-1/2} + \left(4x^3+3x\right)\frac{\mu_2}{12\mu_1^2}\lambda^{-1} \\
&\qquad - \left(2x^4+x^2+1\right)\frac{\mu_2^{3/2}}{8\mu_1^3}\lambda^{-3/2} + O\left(\lambda^{-2}\right)
\end{aligned}$$

and then:

$$\begin{aligned}
&\Phi\left(\frac{\ln^{\frac{1}{2}}\left(1+\frac{\mu_2}{\mu_1^2}\lambda^{-1}\right)}{2} + \frac{\ln\left(1 + x \cdot \frac{\mu_2^{1/2}}{\mu_1}\lambda^{-1/2}\right)}{\ln^{\frac{1}{2}}\left(1+\frac{\mu_2}{\mu_1^2}\lambda^{-1}\right)}\right) \\
&\quad = \Phi(x) - \lambda^{-1/2}\phi(x)\left(x^2-1\right)\frac{\mu_2^{1/2}}{2\mu_1} \\
&\qquad + \lambda^{-1}\phi(x)\left(-3x^5+14x^3+3x\right)\frac{\mu_2}{24\mu_1^2} \\
&\qquad + \lambda^{-3/2}\phi(x)\left(-x^8+12x^6-20x^4-8x^2-7\right)\frac{\mu_2^{3/2}}{48\mu_1^3} + O\left(\lambda^{-2}\right).
\end{aligned}$$

A.10. Gamma approximation

Note that $\chi_n^2 \sim \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$. On the other hand, from the Edgeworth expansion of a sum of independent and identically distributed random variables:

$$\begin{aligned}
&\mathbb{P}\left\{(2n)^{-1/2}\left(\chi_n^2 - n\right) \le x\right\} \\
&\quad = \Phi(x) - \phi(x)\left\{\frac{\kappa_3}{6}He_2(x) + \frac{\kappa_4}{24}He_3(x) + \frac{10\kappa_3^2}{720}He_5(x)\right. \\
&\qquad \left. + \frac{\kappa_5}{120}He_4(x) + \frac{35\kappa_3\kappa_4}{5040}He_6(x) + \frac{280\kappa_3^3}{362880}He_8(x)\right\} \\
&\quad = \Phi(x) - \phi(x)\left\{\frac{2^{1/2}}{3n^{1/2}}He_2(x) + \frac{1}{2n}He_3(x) + \frac{1}{9n}He_5(x)\right. \\
&\qquad \left. + \frac{2^{3/2}}{5n^{3/2}}He_4(x) + \frac{2^{1/2}}{6n^{3/2}}He_6(x) + \frac{2^{1/2}}{81n^{3/2}}He_8(x)\right\} \\
&\quad = \Phi(x) - \phi(x)\frac{2^{1/2}}{3n^{1/2}}\left(x^2-1\right) - \phi(x)\frac{1}{18n}x\left(2x^4-11x^2+3\right) \\
&\qquad - \frac{2^{1/2}}{810n^{3/2}}\phi(x)\left(10x^8-145x^6+399x^4-69x^2-3\right).
\end{aligned} \tag{A.2}$$

The result follows upon taking $n = \frac{2\lambda\mu_1^2}{\mu_2}$.

A.11. Translated Gamma approximation

The result follows from (A.2) upon taking $n = \frac{8}{\gamma_1^2}$.

A.12. Bowers Gamma approximation

Using the formulas in Tricomi and Erdélyi (1951), we get:

$$\begin{aligned}
&\frac{\left(\alpha+\sqrt{\alpha}x\right)\Gamma(\alpha)}{\Gamma(\alpha+1)} - 2\frac{\left(\alpha+\sqrt{\alpha}x\right)^2\Gamma(\alpha)}{\Gamma(\alpha+2)} + \frac{\left(\alpha+\sqrt{\alpha}x\right)^3\Gamma(\alpha)}{\Gamma(\alpha+3)} \\
&\quad = \alpha^{-1}\left(x^2-1\right) + \alpha^{-3/2}\left(x^3-5x\right) + \alpha^{-2}\left(-7x^2+5\right) \\
&\qquad +\alpha^{-5/2}\left(-3x^3+17x\right) + O\left(\alpha^{-3}\right) \\
&\frac{\left(\alpha+\sqrt{\alpha}x\right)\Gamma(\alpha)}{\Gamma(\alpha+1)} - 3\frac{\left(\alpha+\sqrt{\alpha}x\right)^2\Gamma(\alpha)}{\Gamma(\alpha+2)} + 3\frac{\left(\alpha+\sqrt{\alpha}x\right)^3\Gamma(\alpha)}{\Gamma(\alpha+3)} - \frac{\left(\alpha+\sqrt{\alpha}x\right)^4\Gamma(\alpha)}{\Gamma(\alpha+4)} \\
&\quad = \alpha^{-3/2}\left(-x^3+3x\right) + \alpha^{-2}\left(-x^4+12x^2-7\right) + \alpha^{-5/2}\left(15x^3-43x\right) + O\left(\alpha^{-3}\right) \\
&\frac{\left(\alpha+\sqrt{\alpha}x\right)\Gamma(\alpha)}{\Gamma(\alpha+1)} - 4\frac{\left(\alpha+\sqrt{\alpha}x\right)^2\Gamma(\alpha)}{\Gamma(\alpha+2)} + 6\frac{\left(\alpha+\sqrt{\alpha}x\right)^3\Gamma(\alpha)}{\Gamma(\alpha+3)} \\
&\qquad -4\frac{\left(\alpha+\sqrt{\alpha}x\right)^4\Gamma(\alpha)}{\Gamma(\alpha+4)} + \frac{\left(\alpha+\sqrt{\alpha}x\right)^5\Gamma(\alpha)}{\Gamma(\alpha+5)} \\
&\quad = \alpha^{-2}\left(x^4-6x^2+3\right) + \alpha^{-5/2}\left(x^5-22x^3+43x\right) + O\left(\alpha^{-3}\right).
\end{aligned}$$

On the other hand, consider the Gamma random variable Γ_α defined by the density $\frac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)}$. The density of the standardized version $\frac{\Gamma_\alpha-\alpha}{\sqrt{\alpha}}$ is $f_\alpha^\star(x) = \frac{\left(\alpha+\sqrt{\alpha}x\right)^{\alpha-1}e^{-\left(\alpha+\sqrt{\alpha}x\right)}}{\Gamma(\alpha)}\sqrt{\alpha}$. Now, $f_\alpha^\star$ admits an Edgeworth expansion:

$$\begin{aligned}
f_\alpha^\star(x) = \phi(x)\cdot\Bigg\{&1+\alpha^{-1/2}\left[\frac{He_2(x)}{3}\right] + \alpha^{-1}\left[\frac{He_3(x)}{4}+\frac{He_5(x)}{18}\right] \\
&+\alpha^{-3/2}\left[\frac{He_4(x)}{5}+\frac{He_6(x)}{12}+\frac{He_8(x)}{162}\right] + o\left(\alpha^{-3/2}\right)\Bigg\}.
\end{aligned}$$

This shows that the approximation can be expressed as $\Gamma\left(\alpha+\sqrt{\alpha}x,\alpha\right)$ plus $\phi(x)$ times a polynomial in $\alpha^{-1/2}$ and x. The difference $F(x)-\Gamma\left(\alpha+\sqrt{\alpha}x,\alpha\right)$ has already been computed for the Gamma method.

A.13. Inverse Gaussian approximation

Since the IG (m, b) distribution is infinitely divisible and therefore closed under addition, the following Edgeworth expansion holds:

$$\begin{aligned}
&\mathbb{P}\left\{\frac{\mathrm{IG}(m,b)-m}{\sqrt{mb}}\leq x\right\}\\
&=\Phi(x)-\phi(x)\left\{He_2(x)\frac{\kappa_3}{6}+He_3(x)\frac{\kappa_4}{24}+He_5(x)\frac{10\kappa_3^2}{720}+He_4(x)\frac{\kappa_5}{120}\right.\\
&\quad\left.+He_6(x)\frac{35\kappa_3\kappa_4}{5040}+He_8(x)\frac{280\kappa_3^3}{362880}\right\}+O\left(\lambda^{-2}\right)\\
&=\Phi(x)-\phi(x)\left\{He_2(x)\frac{\left(\frac{b}{m}\right)^{1/2}}{2}+He_3(x)\frac{5\frac{b}{m}}{8}+He_5(x)\frac{\frac{b}{m}}{8}\right.\\
&\quad\left.+He_4(x)\frac{7\left(\frac{b}{m}\right)^{3/2}}{8}+He_6(x)\frac{5\left(\frac{b}{m}\right)^{3/2}}{16}+He_8(x)\frac{\left(\frac{b}{m}\right)^{3/2}}{48}\right\}+O\left(\lambda^{-2}\right).
\end{aligned}$$

Replacing the values of m and b, we get:

$$\begin{aligned}
\mathbb{P}\left\{\frac{\mathrm{IG}(m,b)-m}{\sqrt{mb}}\leq x\right\}&=\Phi(x)-\phi(x)\left\{He_2(x)\frac{\gamma_1}{6}+He_3(x)\frac{5\gamma_1^2}{72}+He_5(x)\frac{\gamma_1^2}{72}\right.\\
&\quad\left.+He_4(x)\frac{7\gamma_1^3}{216}+He_6(x)\frac{5\gamma_1^3}{432}+He_8(x)\frac{\gamma_1^3}{1296}\right\}+O\left(\lambda^{-2}\right).
\end{aligned}$$

A.14. Gamma–IG approximation

From the formulas seen above, we get:

$$\begin{aligned}
&w\cdot\mathbb{P}\left\{\frac{\gamma_1}{4}\left(\chi^2_{\frac{8}{\gamma_1^2}}-\frac{8}{\gamma_1^2}\right)\leq x\right\}+(1-w)\cdot\mathbb{P}\left\{\frac{\mathrm{IG}(m,b)-m}{\sqrt{mb}}\leq x\right\}\\
&=w\cdot\left\{\Phi(x)+\phi(x)\left(1-x^2\right)\cdot\frac{\gamma_1}{6}+\phi(x)\,x\left(-2x^4+11x^2-3\right)\cdot\frac{\gamma_1^2}{144}+o\left(\lambda^{-3/2}\right)\right\}\\
&\quad+(1-w)\cdot\left\{\Phi(x)+\phi(x)\left(1-x^2\right)\cdot\frac{\gamma_1}{6}+\phi(x)\,x\left(-x^4+5x^2\right)\cdot\frac{\gamma_1^2}{72}+o\left(\lambda^{-3/2}\right)\right\}\\
&=\Phi(x)+\phi(x)\left(1-x^2\right)\cdot\frac{\gamma_1}{6}+\phi(x)\,x\left(-x^4+10x^2-15\right)\cdot\frac{\gamma_1^2}{72}\\
&\quad+\phi(x)\,x\left(3-x^2\right)\cdot\frac{\gamma_2}{24}+O\left(\lambda^{-2}\right).
\end{aligned}$$

This proves the result.

A.15. Exact computation of the cdf for the Gamma case

We want to compute the function $F(x) = \mathbb{P}\left\{\frac{S_N - \mathbb{E}S_N}{\sqrt{\mathbb{V}(S_N)}} \le x\right\}$ to a certain degree of accuracy. We have:

$$\mathbb{P}\{S_N \le y\} = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} \cdot \mathbb{P}\{S_k \le y\}$$

$$F(x) = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} \cdot \mathbb{P}\left\{S_k \le \mathbb{E}S_N + x\sqrt{\mathbb{V}(S_N)}\right\}.$$

We stop the sum at K, obtaining $F_K(x) = \sum_{k=0}^{K} \frac{e^{-\lambda}\lambda^k}{k!} \cdot \mathbb{P}\{S_k \le \mathbb{E}S_N + x\sqrt{\mathbb{V}(S_N)}\}$. The remaining error can be majorized as $\varepsilon_K \triangleq \sum_{k=K+1}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!}$, and the value K such that ε_K is smaller than a predetermined threshold can be obtained without effort from the cdf of the Poisson distribution. The computation is simplified by the fact that if $X_i \sim \Gamma(m, \theta)$, then $S_k \sim \Gamma(km, \theta)$.