

PANJER RECURSION VERSUS FFT FOR COMPOUND DISTRIBUTIONS

PAUL EMBRECHTS¹ AND MARCO FREI^{2,*}

ABSTRACT. Numerical evaluation of compound distributions is an important task in insurance mathematics and quantitative risk management (QRM). In practice, both recursive methods as well as transform based techniques are widely used. We give a survey of these tools, point out the respective merits and provide some numerical examples.

1. INTRODUCTION

The study of numerical methods to evaluate the distribution of compound sums $S_N = Y_1 + \dots + Y_N$ is one of the pillars of classical risk theory and non-life insurance. In the recent years these tools have been extensively used in other areas as well, particularly in credit risk modeling, for example as part of the widely used CreditRisk+ framework, and in operational risk (OpRisk) modeling; see for example Moscadelli (2004) or Frachot et al. (2001) for typical loss distribution approach (LDA) based models.

Before the advent of the personal computer, actuaries were reliant on relatively crude approximation techniques to treat compound distributions. Such approximations are either based on a central limit result (these are, for example, the normal approximation, the Edgeworth expansion and the saddle-point methods) — others are simply based on some ad-hoc reasoning that often lacks rigorous theoretical justification (which leads to approximations such as the shifted gamma, shifted F or inverse Gaussian approximation). After all, these methods are limited to qualitative inquiries and cannot be considered viable QRM tools; the results are just too unpredictable, and reasonable error bounds cannot readily be obtained. Nevertheless, approximations of this kind are still covered in most modern text books on risk theory. However, taking into account the giant leaps in processing power during the recent decades, we are mainly interested in “exact” procedures. The algorithmic foundations to these were laid down in the early 1980s. Recursive techniques were

¹DEPARTMENT OF MATHEMATICS, ETH ZURICH, 8092 ZURICH, SWITZERLAND

²SEMINAR FOR STATISTICS, ETH ZURICH, 8092 ZURICH, SWITZERLAND

* *Corresponding author*

E-mail addresses: embrechts@math.ethz.ch, frei@stat.math.ethz.ch.

Date: October, 2007; this version: July, 2010.

introduced by Panjer (1981); techniques using Fourier inversion were pioneered by Heckman and Meyers (1983) and Feilmeier and Bertram (1987) — making extensive use of the fast Fourier transform (FFT), which has been popularized by Cooley and Tukey (1965).

During the last 20 years, there have been various notable publications on the subject. Panjer and Wang (1993) examined the stability of the Panjer recursion; Willmot (1988), Sundt (1999) and Hess et al. (2002) have given generalizations of the Panjer recursion; Grübel and Hermesmeier (1999, 2000) have investigated the propagation of discretization errors through compounding and established an improved FFT based procedure using an exponential change of measure. The latter contribution is quite substantial since it essentially eliminates the so called aliasing error, which is *the* fundamental deficit that arises through the use of the discrete Fourier transform.

The rest of this paper is organized as follows: In Section 2 we will summarize the methods and discuss various practical aspects; in Section 3 we will give numerical examples that illustrate some relevant issues; and the final section contains some closing remarks.

2. THE METHODS

To begin with, we introduce some notation. Throughout, all random variables (rv's) are defined on some fixed non-atomic probability space (Ω, \mathcal{F}, P) . Let N be a rv with values in \mathbb{N}_0 and distribution Q . Let $\{Y_i\}_{i \in \mathbb{N}}$ be an iid sequence of non-negative rv's independent of N with common distribution (function) F ; F is called the (claim) severity distribution, Q is referred to as the frequency distribution. Put

$$S_N := \sum_{j=1}^N Y_j$$

with the obvious interpretation $\sum_{j=1}^0 Y_j = 0$. The distribution of S_N is called the compound distribution of F under Q , denoted by $Q \vee F$. Formally, we have

$$(1) \quad Q \vee F(x) = \sum_{j=0}^{\infty} F^{*j}(x) P(N = j),$$

where $F^{*j}(x) = \int_0^{\infty} F^{*(j-1)}(x-y)F(dy)$ and $F^{*0} = I_{[0, \infty)}$. Since the convolution powers F^{*j} cannot be tracked analytically for most of the models used in practice, we are reliant on numerical techniques.

To work with either the Panjer recursion or an FFT based algorithm, the severities must be concentrated on a lattice $h\mathbb{N}_0 = \{0, h, 2h, 3h, \dots\}$, where h is some

strictly positive constant. In applications, however, severity distributions are usually taken from the continuous distributions, and thus an initial arithmetization is required. We choose a suitably small bandwidth $h > 0$ and replace F with a distribution $F_h = \{f_{h,j}\}_{j \in \mathbb{N}_0}$ on $h\mathbb{N}_0$. A very intuitive arithmetization design is the *rounding method*: The severities are rounded to the closest integer multiple of h , that is

$$(2) \quad f_{h,j} := F\left(jh + \frac{h}{2}\right) - F\left(jh - \frac{h}{2}\right).$$

It is clear that F_h converges to F weakly as $h \rightarrow 0$, and in turn: $Q \vee F_h \rightarrow Q \vee F$ weakly as $h \rightarrow 0$. A-priori we do not know how $Q \vee F_h$ relates to $Q \vee F$, and we can not give any error bounds. In this regard, the so called forward and backward differences are useful. The forward difference is given by

$$f_{h,j}^+ := F((j+1)h) - F(jh)$$

and, likewise, the backward difference is given by

$$f_{h,j}^- := F(jh) - F((j-1)h).$$

It is clear that $Q \vee F_h^+$ gives an upper bound for $Q \vee F$, and analogously $Q \vee F_h^-$ yields a lower bound. Moreover, $Q \vee F_h^\pm$ converges weakly to $Q \vee F$ as $h \rightarrow 0$. In theory, we thus can sandwich the true compound distribution with any given precision.

Notice that both the forward/backward differences and the rounding method do not conserve any moments of the original distribution. In this light, Gerber (1982) suggests a procedure that locally matches the first k moments. Practically interesting is only the case $k = 1$; for $k \geq 2$ the procedure is not well defined, potentially leading to negative probability mass on certain lattice points. The moment matching method is much more involved than the rounding method in terms of implementation; we need to calculate the limited expected values $\int_0^{kh} xF(dx)$. Apart from that, the gain is rather modest; moment matching only pays off for large bandwidths, and after all, the rounding method is to be preferred. This is further reinforced by the work of Gröbel and Hermesmeier (2000): if the severity distribution is absolutely continuous with a sufficiently smooth density, the quantity $f_{h,j}/h$, an approximation for the density of F , and the corresponding approximate compound density can be quadratically extrapolated.

Assume for now that F is arithmetic (or has been suitably arithmetized). We will use the notation

$$q_k = P(N = k), \quad f_k = P(Y_i = k), \quad g_k = P(S_N = k).$$

Observe that the discrete version of (1) is given by

$$g_n = \sum_{j=0}^{\infty} q_j f_n^{*j}$$

where

$$f_n^{*j} := \begin{cases} 1 & \text{if } j = 0 \text{ and } n = 0 \\ 0 & \text{if } j = 0 \text{ and } n \in \mathbb{N} \setminus \{0\} \\ \sum_{i=0}^n f_{n-i}^{*(j-1)} f_i & \text{if } j \geq 1 \end{cases}$$

This is messy and numerically expensive.

Panjer recursion. If the frequency distribution $Q = \{q_k\}_{k \in \mathbb{N}_0}$ satisfies

$$(3) \quad q_k = \left(a + \frac{b}{k}\right) q_{k-1}, \quad k \geq 1,$$

for some $a, b \in \mathbb{R}$, then the compound distribution $g_n = P(S_N = n)$ satisfies the recursion

$$(4) \quad g_n = \frac{1}{1 - af_0} \sum_{j=1}^n \left(a + \frac{bj}{n}\right) f_j g_{n-j}, \quad n \geq 1$$

with initial condition

$$(5) \quad g_0 = \mathcal{P}_N(f_0),$$

where $\mathcal{P}_N(t) = \sum_{j=0}^{\infty} q_j t^j$ denotes the probability generating function (pgf) of N . The recurrence relation (4) is referred to as the Panjer recursion. Observe that the initial value (5) simplifies to $g_0 = q_0$ for $f_0 = 0$. Sundt and Jewell (1981) showed that (3) holds precisely for the binomial, negative binomial and Poisson distribution (and, as a matter of course, for the degenerate distribution $q_0 = 1$).

The Panjer recursion is very easy to implement in practice and numerically cheaper than brute-force convolution, which asymptotically requires $O(n^3)$ operations to obtain g_0, \dots, g_n compared to the $O(n^2)$ complexity in the Panjer case. As recently shown by Hipp (2003), recursion (4) can be simplified if the severity distribution is of phase-type (or more generally if the pgf is rational). The simplified recursion speeds up the computation considerably, reducing the operation count to $O(n)$. Of course, in a practical situation the severities will not be of phase-type mostly and thus must be approximated accordingly. Since the phase-type distributions are uniformly dense in the space of arithmetic distributions (with respect to the Kolmogorov metric), this approach is useful, at least in the small claims case; however, phase-type distributions are not suitable as approximations for typical heavy-tailed distributions because any distribution with a rational pgf is light-tailed in the sense that its Laplace-Stieltjes transform exists in a neighborhood

of 0. For a review of the use of phase-type distributions in insurance risk theory, see Bladt (2005).

Regarding robustness of the Panjer recursion, Panjer and Wang (1993) established strong stability in the negative binomial and Poisson case: The accumulated rounding error throughout recursion (4) grows linearly with a slope bounded by 1. Moreover, if r digits are used, the number of valid digits $\#(r, n)$ for g_n is bounded by (at a 99% significance level):

$$\#(r, n) \geq r + \left\lfloor \log_{10} \frac{4}{3} - \frac{1}{2} \log_{10}(n+1) \right\rfloor,$$

where \log_{10} denotes the decadic logarithm and $\lfloor \cdot \rfloor$ the floor function; if, for example, the number of lattice points is increased by a factor of 100, the same precision can be retained with an additional two digits. The compound binomial case is somewhat pathological; serious error blow-up can happen, since terms with mixed signs crop up throughout the recursion. The practical implications, however, are debatable: Besides that the binomial distribution is the least interesting distribution in the Panjer class, instabilities arise primarily if the frequency is heavily underdispersed or if the severities exhibit substantial negative skewness — this is rather uncommon in typical applications.

A further computational problem is potential underflow¹ during the initialization (5) of the Panjer recursion; this typically happens for frequencies with a large mean. Naively, one would rescale g_0 with some constant c , carry out the recursion and descale afterwards. It is obvious that this kind of linear scaling introduces new obstacles; cg_n will eventually lead to overflow. Panjer and Willmot (1986) suggest two alternative procedures to prevent underflow and circumvent the overflow issues associated to linear scaling: These are exponential scaling and portfolio splitting; both methods are equally applicable to the compound Poisson and compound negative binomial case. We will briefly explain how to decompose a portfolio in the compound Poisson case; it is well known that

$$(6) \quad (Pois(\lambda/m) \vee F)^{*m} = Pois(\lambda) \vee F,$$

and for a suitably large m , $Pois(\lambda/m) \vee F$ can be evaluated without underflow; the m -fold convolution is then carried out in brute-force fashion (or via some FFT based fast convolution algorithm). Compound mixed Poisson distributions (such as the compound negative binomial) can be decomposed similarly. After all, in an OpRisk context, where we deal with low to moderate frequencies and (extremely)

¹Underflow or overflow occurs if the computer's floating point arithmetic produces a number outside the range of representable numbers; in this case either 0 or a formal constant $\pm\infty$ is output.

heavy-tailed severities, we will usually not be concerned with underflow; see Aue and Kalkbrenner (2006) for a discussion on typical OpRisk data.

Finally, we would like to mention that there are several generalizations of the Panjer recursion, which allow for more general frequency distributions. For example, modifying condition (3) to

$$(7) \quad q_k = \left(a + \frac{b}{k}\right) q_{k-1}, \quad k \geq l + 1,$$

with initial values $q_0 = q_1 = \dots = q_{l-1} = 0$, leads to the Panjer recursion of order l :

$$g_n = q_l f_n^{*l} + \sum_{j=1}^n \left(a + \frac{b_j}{n}\right) g_{n-j} f_j, \quad \forall n \geq l,$$

provided that $f_0 = 0$. The class given by (7) includes, inter alia, the l -truncations of the negative binomial and Poisson distributions, which are interesting in a reinsurance context. An overview of higher order Panjer recursions is given by Hess et al. (2002); furthermore, Sundt (1992) discusses recursions of the kind

$$q_n = \sum_{j=1}^k \left(a_j + \frac{b_j}{n}\right) q_{n-1}, \quad n \geq 1.$$

For multivariate versions we refer to Sundt (1999), Hesselager (1996) or Vernic (1999).

An algorithm based on FFT. Observe that the characteristic function of a compound sum S_N can be expressed in terms of the characteristic function of the severity Y_1 :

$$(8) \quad \phi_{S_N}(t) := E[e^{itS_N}] = \mathcal{P}_N(\phi_{Y_1}(t)),$$

where \mathcal{P}_N denotes the pgf of N . Applying the inverse Laplace transform to the RHS of (8), the distribution of S_N can be obtained. Heckman and Meyers (1983) give explicit formulas in case that the distribution function of Y_1 is piecewise linear or piecewise constant. However, we will resort directly to the discrete Fourier transform (DFT).

Set $f = (f_0, f_1, \dots, f_{M-1}) \in \mathbb{R}^M$ for some truncation point M . The DFT

$$\hat{f} = (\hat{f}_0, \dots, \hat{f}_{M-1})$$

is defined by

$$(9) \quad \hat{f}_j = \sum_{k=0}^{M-1} f_k e^{i2\pi jk/M}, \quad j = 0, 1, \dots, M-1.$$

Given \widehat{f} , the original sequence f can be reconstructed by

$$(10) \quad f_j = \frac{1}{M} \sum_{k=0}^{M-1} \widehat{f}_k e^{-i2\pi jk/M}.$$

If M is a power of 2, (9) and (10) can be computed efficiently via one of the various FFT algorithms.

With regards to (8), one takes the inverse DFT of

$$(11) \quad \widehat{g^M} := \mathcal{P}_N(\widehat{f}).$$

as an approximation for $g = (g_0, \dots, g_{M-1})$. If $\sum_{j=0}^{M-1} f_j = 1$, i.e. if there is no truncation error, formula (11) evaluates the compound distribution exactly on the cyclic group $\mathbb{Z}/M\mathbb{Z}$. Compound mass which lies at M and beyond will be “wrapped around” and erroneously appears in the range $0, \dots, M-1$. For severities with considerable tail mass, the truncation and wrap-around error (the so called aliasing error) can become quite an issue. If we are interested in the compound probabilities on $\{0, 1, \dots, n\}$, we possibly have to take M much larger than n to obtain reasonable values. Grübel and Hermesmeier (1999) quantified the aliasing error: If the frequency distribution has finite variance, then

$$(12) \quad \begin{aligned} & \sum_{n=0}^{M-1} |g_n^M - g_n| \\ &= \sum_{n=M}^{\infty} g_n - \mathcal{P}'_N(1) \sum_{n=M}^{\infty} f_n + O\left(\left(\sum_{n=M}^{\infty} f_n\right)^2\right). \end{aligned}$$

From equation (12) we deduce that the aliasing error is small if M is large and the tail of F (and thus likewise the tail of $Q \vee F$) decays fast. In this regard, the above authors suggest a tilting procedure to reduce the error: Fix some $\theta > 0$ and set

$$E_\theta f = (e^{-\theta j} f_j)_{j=0,1,\dots,M-1}.$$

The tail of $E_\theta f$ decays at an exponential rate and thus potentially much faster than f . The operator E_θ commutes with convolution and therefore we have

$$(13) \quad Q \vee F = E_{-\theta}(Q \vee E_\theta F).$$

In accordance with (11) and (13), this gives the following algorithm:

1. Choose a truncation point $M \in \mathbb{N}$ and a tilting parameter $\theta > 0$.
2. Set $f = (f_0, f_1, \dots, f_{M-1})$.
3. Tilt the sequence: $f \mapsto E_\theta f = (e^{-\theta j} f_j)_{j=0,1,\dots,M-1}$.
4. Calculate the DFT of the tilted sequence: $E_\theta f \mapsto \widehat{E_\theta f}$.
5. Take the inverse DFT of $\mathcal{P}_N(\widehat{E_\theta f})$ and untilt by applying $E_{-\theta}$.

The tilting is of much practical value and can reduce the aliasing error tremendously. In applications, the tilting parameter should be chosen as large as possible without producing under- or overflow; if we use common (64bit) double precision, a value $M\theta \approx 20$ seems reasonable in most cases. Care must also be taken for “large” frequencies; in this case, the evaluation of the pgf \mathcal{P}_N may lead to underflow as well. This problem can be solved with the portfolio splitting technique; see equation (6).

Compared to the Panjer recursion, FFT has two main advantages: It works with arbitrary frequency distributions and it is much more efficient. To obtain g_0, \dots, g_n , FFT takes essentially $O(n \log n)$ operations compared to the $O(n^2)$ operations for recursion. Bühlmann (1984) gives a detailed analysis in the compound Poisson case, concluding that FFT practically always beats Panjer recursion for $n \geq 256$. Of course, this figure is somewhat archaic in light of today’s processing powers, and both Panjer recursion and FFT take merely a split second to evaluate for $n = 256$ lattice points. In a realistic setup, with potentially heavy-tailed severities — say, we want to determine the 99.9%-quantile of a compound distribution with Pareto severities — several (ten) thousand knots are required; the Panjer recursion will be outperformed by far. This last example is typical for the calculation of regulatory capital for OpRisk; see Chapter 10 in McNeil et al. (2005).

3. EXAMPLES

The following examples will be carried out numerically within the R programming language. We will make use of the *actuar*-package freely available through CRAN. This package features the *discretize()* routine, which implements the arithmetization designs mentioned above, and incorporates some useful distributions (such as the Pareto, loggamma and Weibull).

Introductory Example. We start with a simplistic example that pools the essential concepts. Suppose we want to evaluate the distribution $Pois(\lambda) \vee Pareto(\alpha, \beta)$ on the range $[0, (M - 1)h]$, where the *Pareto*(α, β) distribution has density

$$f(x) = \frac{\alpha}{\beta} \left(1 + \frac{x}{\beta}\right)^{-(\alpha+1)} \quad x \geq 0, \alpha > 0, \beta > 0.$$

- Arithmetization of *Pareto*(α, β) with span h via the rounding method:

```
f <- discretize(ppareto(x,alpha,beta), from=0, to=M*h, by=h,
               method="rounding")
```

The difference designs can be obtained via “lower” and “upper”, the local moment matching method via “unbiased”.

- Compute g_0, \dots, g_{M-1} via FFT without tilting:

```
fhat <- fft(f, inverse=FALSE)
P <- exp(lambda*(fhat-1))
```



```
g <- 1/M*fft(P, inverse=TRUE)
```

Observe that the pgf of $N \sim Pois(\lambda)$ is given by

$$\mathcal{P}_N(t) = \exp(\lambda(t - 1))$$

and that `fft(·, inverse=TRUE)` returns the *unnormalized* inverse DFT, thus we divide by $1/M$.

- Tilting with $\theta > 0$ gives the procedure:

```
f <- exp(-theta*(0:(M-1)))*f
fhat <- fft(f, inverse=FALSE)
P <- exp(lambda*(fhat-1))
g <- exp(theta*(0:(M-1)))*(1/M*fft(P, inverse=TRUE))
```

- Since $Pois(\lambda)$ corresponds to $a = 0$ and $b = \lambda$ in (3), the Panjer recursion can be implemented as follows:

```
g <- vector(length=M)
g[1] <- exp(lambda*(f[1]-1))
for(n in 1:(M-1)){
  g[n+1] <- sum(lambda*(1:n)/n*f[(1:n)+1]*g[n-(1:n)+1])
}
```

Observe the index shift, since R allows solely strictly positive indices.

The vector g contains the probability masses on the set $\{0, h, 2h, \dots, (M-1)h\}$. Due to small rounding errors, the numerical values of g will most likely be complex numbers with non-vanishing imaginary parts. Of course, we only work with the real parts. Dividing the entries of g by h gives an approximation for the compound density on $(0, (M-1)h]$. Figure 1 displays these densities for $Pois(20) \vee Pareto(4, 3)$ with model parameters $h = 0.1$, $M = 2^8$ and $\theta = 0.01$. The wrapped-around mass

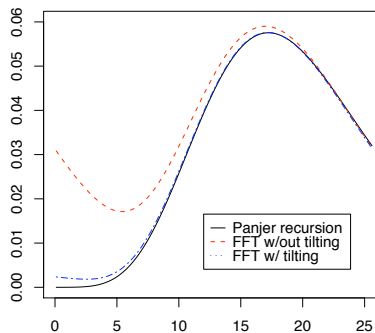


FIGURE 1. Approximate densities for $Pois(20) \vee Pareto(4, 3)$.

is clearly visible. For illustrative purposes the tilting parameter θ has been chosen

Design	$h = 1$	$h = 0.5$	$h = 0.1$	$h = 0.01$	Exact
Forward	58	70.0	81.9	84.78	
Rounding	84	84.5	85.1	85.11	85.11
Backward	124	103.0	88.4	85.43	

TABLE 1. 99.9%-quantile estimate for $Pois(50) \vee Exp(1)$ using different arithmetization designs.

quite small. According to our rule of thumb $\theta \approx 20/M$, we would use a much larger value, say, $\theta = 0.1$. In this case, however, the difference between the tilted FFT density and the Panjer density would be no longer visible in the plot; in fact, the absolute difference would be less than ≈ 0.0006 over the observed range.

Bandwidth choice. In practice, a suitably small bandwidth h must be chosen for the arithmetization of the severity distribution. Since there is no analytically tractable formula available for the discretization error, the choice is kind of an educated guess. As a simple ad-hoc method one might use the bounds obtained via the forward and backward differences, reducing h until the difference of the bounds is smaller than some margin of error which seems acceptable in the given context. However, this approach is somewhat dissatisfying. The forward/backward bounds are usually very pessimistic: approximations obtained via the rounding method are typically much closer to the exact value than to one of the bounds. An example is given in Table 1. In this regard, it is probably more advisable to successively reduce h and compute the corresponding compound distribution with the rounding method until the (relative) improvement is smaller than some threshold.

The alias effect. In situations where both Panjer recursion and FFT are equally applicable, and where we have to deal with a large number of lattice points, we would typically favor FFT because it is numerically cheaper. However, one may worry about the aliasing error. Yet, this worry is unsubstantiated, and with suitable tilting, the effect is negligible in practical circumstances. Consider, for example, the following two (extremely heavy-tailed) models:

- (a) $Pois(50) \vee Pareto(1.5, 0.5)$
- (b) $Pois(5) \vee Pareto(1, 0.1)$.

Model (a) has infinite variance, model (b) has infinite mean. We will compute the compound distribution function via Panjer recursion and FFT, and plot the absolute differences between the two distribution functions; this difference essentially arises from the aliasing since the Panjer recursion evaluates the distribution exactly up to discretization errors, and potential rounding effects are negligible. For reasonably large truncation points, the alias error is very small, as one can gather

from Figure 2. Tilting is particularly useful if one is only interested in the body of

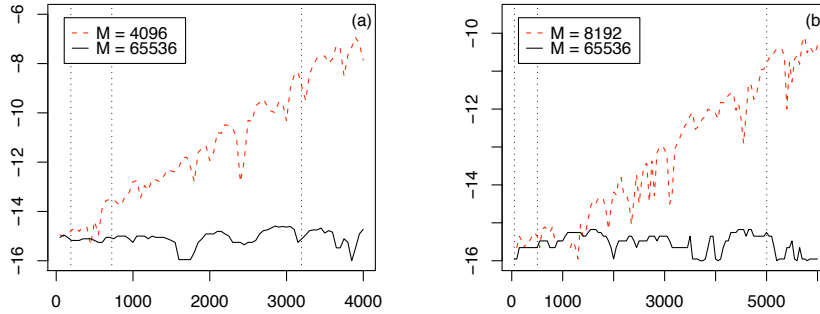


FIGURE 2. Logarithmic absolute aliasing errors of the approximate cumulative distribution function for $Pois(50) \vee Pareto(1.5, 0.5)$ (plot a) and $Pois(5) \vee Pareto(1, 0.1)$ (plot b) with different truncation points M . A span $h = 1$ was used. The dotted vertical lines indicate the 0.99, 0.999 and 0.9999 quantiles, respectively.

the distribution (say, below the 90%-quantile). Figure 3 displays the logarithmic absolute aliasing errors of the approximate cdf for a $Pois(20) \vee Pareto(1, 0.1)$ distribution — once with tilting (plot c), and once without (plot d). Without tilting, the approximation is useless in this example.

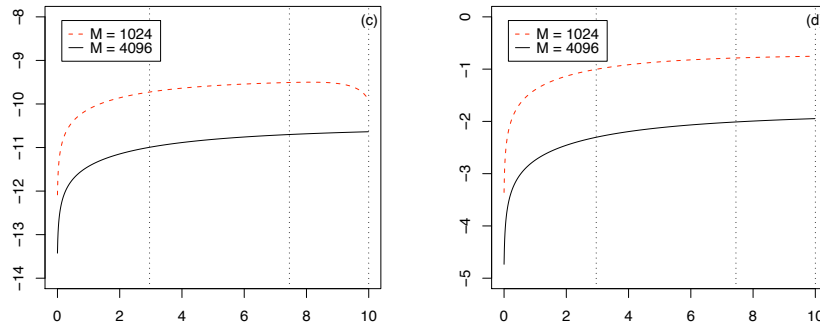


FIGURE 3. Logarithmic absolute aliasing errors of the approximate cumulative distribution function for a $Pois(20) \vee Pareto(1, 0.1)$ distribution with different truncation points M ; with tilting (plot c), without tilting (plot d). A span $h = 0.01$ was used. The dotted vertical lines indicate the 0.1, 0.5 and 0.9 quantiles, respectively.

4. CONCLUSION

The Panjer recursion is arguably the most widely used method to “exactly” evaluate compound distributions. However, FFT is a viable alternative: It can

be applied with arbitrary frequencies and offers a tremendous timing advantage for a large number of lattice points; moreover, the use of exponential tilting — which practically rules out any aliasing effects — facilitates applications (such as evaluation of the lower tail) that were thought to be an exclusive task for recursive procedures.

REFERENCES

- Aue, F. and Kalkbrenner, M. (2006). LDA at work: Deutsche Bank's approach to quantifying operational risk. *Journal of Operational Risk*, 1(4):49–93.
- Bladt, M. (2005). A review of phase-type distributions and their use in risk theory. *ASTIN Bulletin*, 35(1):145–167.
- Bühlmann, H. (1984). Numerical evaluation of the compound Poisson distribution: Recursion or fast Fourier transform. *Scandinavian Actuarial Journal*, 2:116–126.
- Cooley, J. W. and Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. *Math. Comput.*, 19:297–301.
- Credit Suisse Financial Products (1997). CreditRisk+: a credit risk management framework. Technical Document.
- Feilmeier, M. and Bertram, J. (1987). *Anwendung numerischer Methoden in der Risikotheorie*. Verlag Versicherungswirtschaft E.V., Karlsruhe.
- Frachot, A., Georges, P., and Roncalli, T. (2001). Loss distribution approach for operational risk. *Working Paper, Groupe de Recherche Opérationnelle, Crédit Lyonnais, France*.
- Gerber, H. U. (1982). On the numerical evaluation of the distribution of aggregate claims and its stop-loss premiums. *Insurance: Mathematics and Economics*, 1(1):13–18.
- Grübel, R. and Hermesmeier, R. (1999). Computation of compound distributions I: Aliasing errors and exponential tilting. *ASTIN Bulletin*, 29(2):197–214.
- Grübel, R. and Hermesmeier, R. (2000). Computation of compound distributions II: Discretization errors and Richardson extrapolation. *ASTIN Bulletin*, 30(2):309–331.
- Heckman, P. E. and Meyers, G. G. (1983). The calculation of aggregate loss distributions from claim severity and claim count distributions. *Proceedings of the Casualty Actuarial Society*, LXX:22–61.
- Hess, K. T., Liewald, A., and Schmidt, K. D. (2002). An extension of Panjer's recursion. *ASTIN Bulletin*, 32(2):283–297.
- Hesselager, O. (1996). Recursions for certain bivariate counting distributions and their compound distributions. *ASTIN Bulletin*, 26(1):35–52.
- Hipp, C. (2003). Speedy Panjer for phase-type claims. Preprint, Universität Karlsruhe.

- McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management. Concept, Techniques, Tools*. Princeton University Press, Princeton.
- Moscadelli, M. (2004). The modelling of operational risk: experience with the analysis of the data collected by the Basel Committee. *Temi di discussione del Servizio Studi, Number 517*.
- Panjer, H. H. (1981). Recursive evaluation of a family of compound distributions. *ASTIN Bulletin*, 12(1):22–26.
- Panjer, H. H. and Wang, S. (1993). On the stability of recursive formulas. *ASTIN Bulletin*, 23(2):227–258.
- Panjer, H. H. and Willmot, G. E. (1986). Computational aspects of recursive evaluation of compound distributions. *Insurance: Mathematics and Economics*, 5:113–116.
- Sundt, B. (1992). On some extensions of Panjer’s class of counting distributions. *ASTIN Bulletin*, 22(1):61–80.
- Sundt, B. (1999). On multivariate Panjer recursions. *ASTIN Bulletin*, 29(1):29–45.
- Sundt, B. and Jewell, W. S. (1981). Further results on recursive evaluation of compound distributions. *ASTIN Bulletin*, 12(1):27–39.
- Vernic, R. (1999). Recursive evaluation of some bivariate compound distributions. *ASTIN Bulletin*, 29(2):315–325.
- Willmot, G. E. (1988). Sundt and Jewell’s family of discrete distributions. *ASTIN Bulletin*, 18(1):17–29.