Preface

Stephen J. Mildenhall

2026-10-22 19:57 commit b5ca906+

This book reproduces published figures, tables and results using `aggregate`. Every number in it is computed by the library, never transcribed. A chapter is recomputed when its text changes, and every chapter is recomputed whenever the library or any other installed package changes, so all the numbers in a given copy come from the one environment described below. A match therefore demonstrates the correctness of the library version named here, down to the git commit of its source.

```
Exception reporting mode: Minimal
Rendered 2026-10-22 19:57:30 GMT Daylight Time (UTC+0100)
----------------------------------------------------------
Python           | 3.14.7 on Windows-11-10.0.26200-SP0
aggregate        | 1.7.1, source at git commit 0036744
greater_tables   | 8.1.0
```

A computational library must be correct. `aggregate` has an extensive set of *internal* correctness checks embodied in its validation system: these check for a match between the first three theoretical and empirical moments. However, a published number computed by somebody independent using a different method is a good *external* check. We work through results from twenty eight papers. Most of them confirm the library. Several of them corrected it, or located a limit in it that was not previously known. Nine of them turn out to contain an error, an internal inconsistency or a misprint, which is not a criticism of the authors so much as evidence that the check runs in both directions.

Solving external problems tests `aggregate` in a second way: it shows the library is flexible enough to solve problems posed *by other people*! Obviously, it can solve problems I think are important. The library's success on external problems bodes well when *you* come to use it.

The first chapter is the pith: five papers, each with the DecL program that reproduces the published exhibit and the answer set beside the paper's own numbers. It highlights how a simple DecL statement can answer an externally set question.

> ⚠ AI Generated Work!
>
> The mechanical work in this book was conducted by Claude Code Opus 5 model with high-level human review. In a sense, the work checks itself, but the reader should be aware it was essentially auto generated by AI.

1 The structure of each chapter

Each chapter follows the same four-part structure: the context and the problem the paper is solving, the paper's own solution and the gold standard it publishes, the `aggregate` reproduction, and

comments. A simple reproduction can be very short. Something unusual, or unfamiliar to actuaries, needs more context.

When reviewing the code mentally discount the Python scaffolding and focus on the DecL nugget that does the work. Since the reproductions have to recreate the paper's work, they include a lot of code that is conceptually irrelevant other than helping to display correctness.

Jin et al. (2016) turned out not to be reproducible because of `aggregate` limitations, as the chapter on the paper records. It models a common shock count mode for bivariate frequencies.

2 Three categories of paper

The papers fall into three categories.

A **model paper** poses an insurance structure with explicit parameters. Reproduction means building it in DecL and matching the published numbers, and the paper is the authority. Venter (1983), Mack and Fackler (2003), Mata and Verheyen (2005), Bodoff (2017), Homer and Clark (2003) and Bear and Nemlick (1990) are of this kind, and they make up the first part.

A **method paper** proposes a competing numerical technique for the aggregate distribution. Here `aggregate` is the exact ground truth against which the paper's own approximation is measured, so the roles reverse: the paper is the thing being checked. Whether reproduction is possible at all depends on whether the paper published concrete input parameters and tabulated output. Several rare-event papers publish neither, reporting only sample counts and CPU seconds, and that is a large part of why the later parts read the way they do.

A **theory paper** is a third kind, and the two Denuit, Dhaene, et al. (2025) and Denuit, Ortega-Jiménez, et al. (2025) chapters are the instances. These prove results about the conditional expectation given the sum, $\mathsf{E}[X_i \mid S = s]$, and illustrate them with figures rather than tables. There is no insurance contract in either, and no compound distribution: each agent carries one claim with certainty, so what is under test is the allocation kernel and the grid it runs on. Success here means reproducing a *shape*, and it is checkable only because both papers state something numerical about the shape, a turning point in one case and a monotonicity failure in the other. Reproducing a figure is weaker evidence than reproducing a table, so these chapters lean on constraints the model must satisfy whatever the picture looks like: that shares add up, and that identically distributed agents get identical shares.

3 Conclusions: correctness, flexibility, errors, and exactness

The reproductions exercise confirms `aggregate`'s correctness and flexibility. Venter (1983)'s Exhibit 3 reproduces in all three columns, including the transformed gamma, which is `scipy.stats.gengamma` under a different name. Bruno et al. (2006)'s Tables 6, 7 and 8 come back exact at every printed decimal. Bodoff (2017)'s Exhibit 4 reproduces cell for cell. Homer and Clark (2003) reproduces every published cell of sections 2 and 4, and 136 of the 147 cells of section 3. Bear and Nemlick (1990)'s Table I reproduces in both of its reproducible columns, six adjustable features priced by six DecL clauses. Pittarello et al. (2024)'s accuracy table measures where the grid ends rather than how finely it is spaced, and on a long enough grid the error falls as h^2, in the opposite direction to the one the paper reports. Gollini and Rougier (2016)'s 32,060 event hurricane loss table, where the authors' Panjer recursion stopped at a \$100 rounding after 4,651 seconds, is computed unrounded in under a second, and the moment bound they recommend turns out to be four to ten times the exact tail probability.

The exercise uncovered small errors in six papers. Mata and Verheyen (2005)'s Table 1 limited expected values are not the exact lognormal ones, running 0.04% high at $250,000 and drifting to 0.24% low at $5M, while the closed form matches to eight figures. Homer and Clark (2003)'s Table 3.3 does not sum to one, carrying 99.97% of the mass. Asmussen et al. (2016) has a row where its two published saddlepoint columns and its Monte Carlo column disagree, and `aggregate`, being neither a saddlepoint expansion nor a sampler, can determine which one to trust. Denuit, Dhaene, et al. (2025)'s Figure 1 cannot hold at its right hand edge: agent 3's loss is capped at 30, so adding up forces agents 1 and 2 above 5 at a pool loss of 40, and their panels are drawn on an axis whose top tick is 4. Pittarello et al. (2024)'s Table 3 prints the same reference value in all four rows for four and five dimensions, a distribution function that would have to be flat from 10 to 10,000; Arbenz et al. (2011), where the values come from, has the right ones. Gollini and Rougier (2016) report Monte Carlo samples of 10^5 in the text, while their own design calculation, their summary and the script that drew their figures all use 10^4.

Finally, it reveals the FFT computational resolution floor. Ben Rached et al. (2024)'s Theorem 3 predicts that FFT convolution error is bounded absolutely rather than relatively, so below roughly 10^{-13} the output is noise regardless of grid refinement. `aggregate` sits on that predicted floor to within a factor of a few. Wilson and Keich (2016) shows the same failure on adversarial inputs and, more usefully, shows that it is not always detectable: a wrong answer can come back positive and plausible. Wilson and Keich (2017) exhibits it on a benign input and demonstrates the repair. Together they turned a vague caveat into a measured property, and explain why the far tail now has documented limits rather than optimistic ones.

Computational accuracy is discussed further in two other recent blog posts. Dice of Dice: The FFT Meets Exact Arithmetic compares FFT computed probabilities with an exact arithmetic probability calculation, and Dice of Dice: Exact Finite FFTs discusses using number-theoretic transforms which give exact rational number answers with no floating point noise. The latter distinguishes between the Fourier method which is exact and different numerical implementations. That method now ships in the library as the `nttagg` kind, and the Wilson and Keich (2016) and Wilson and Keich (2017) chapters use it: their exact reference values, previously hand algebra in one case and trusted naive convolution in the other, are computed in-library as exact rational numbers, and the float pipeline's error against them is itself computed exactly, per bucket. An optional compiled engine, the Rust add-in `aggregate-ntt` working over 64-bit primes, extends the reach of the exact method; this book runs the pure-Python reference path.

4 The chapters

Table 1 and Table 2 list the chapters in book order, split at the part boundary between the actuarial chapters and the numerical ones: the paper each chapter reproduces, where it was published and how often it has been cited, what the chapter targets, and how it came out.

Table 1: The chapters on actuarial models, distortion pricing and risk sharing, in book order, with each paper's citation count, the chapter's target, and the outcome.

Part	Paper	Year	Journal	Citations	Target	Outcome
Actuarial models	Venter (1983)	1983	Proceedings of the Casualty Actuarial Society	109	Exhibit 3, three methods over 34 rows	Reproduced. The bucket size selects which of the two exact columns you land on
	Mack and Fackler (2003)	2003	Blätter der DGVFM	10	Table 1, and the threshold above which Riebesell is exact	Reproduced. All 26 cells within 8e-4, and the threshold is 5.186 $\mathsf{E}[X]$ against the paper's "about 5"
	Mata and Verheyen (2005)	2005	CAS Forum	10	Tables 1 to 14, trend and exposure adjustment by layer	Reproduced, and a discrepancy found in the paper's Table 1
	Bodoff (2017)	2017	Variance	0	Exhibit 4, excess of policy limits	Reproduced. Two unstated model choices inferred and confirmed by fit
	Homer and Clark (2003)	2003	Proceedings of the Casualty Actuarial Society	14	All three worked examples of the bivariate FFT	Reproduced. Every construction is a shipped DecL form
	Bear and Nemlick (1990)	1990	Proceedings of the Casualty Actuarial Society	10	Table I, six treaties with adjustable features	Reproduced. The lognormal column exactly, the collective risk column to the printed digit in five of six
Distortion pricing	Wang (2000)	2000	Journal of Risk and Insurance	676	Table 1, risk load by layer under two distortions	Reproduced to 0.31%, and the basic limit expected loss is misprinted
	Wang (2002)	2002	ASTIN Bulletin	268	Table 2, weather derivative calls off 22 observations	Reproduced. Prices to under half a cent
Risk sharing	Denuit (2019)	2019	ASTIN Bulletin	63	Figure 1 and the shares at $s = 1$, four participant pool	Reproduced. The horizontal line resolves a table that contradicts its own text
	Denuit et al. (2022)	2022	ASTIN Bulletin	18	Example 4.5, Figure 6, mortality credits at three pool sizes	Reproduced. Shares add up exactly at every pool size
	Denuit, Dhaene, et al. (2025)	2025	Insurance: Mathematics and Economics	5	Figure 1, the comonotonic repair of a conditional mean allocation	Reproduced, and the published figure cannot hold at its right hand edge
	Denuit, Ortega-Jiménez, et al. (2025)	2025	ASTIN Bulletin	2	Figures 2, 5 and 6, conditional expectation given the sum	Reproduced. The turning point lands at 64.46, which is what settles the index convention

Table 2: The chapters on exact computation, rare events and the paper that did not reproduce, in book order, with each paper's citation count, the chapter's target, and the outcome.

Part	Paper	Year	Journal	Citations	Target	Outcome
Exact computation	Gerber (1982)	1982	Insurance: Mathematics and Economics	69	Tables 1 to 3, the exact stop loss premium columns	Reproduced. All 28 values inside the printed precision
	Robertson (1992)	1992	Proceedings of the Casualty Actuarial Society	25	Table 4, the piecewise uniform algorithm	Partial. Exact for the paper, $O(h)$ for a lattice, so the printed precision is out of reach
	Embrechts et al. (1993)	1993	Statistica Neerlandica	42	Figure 1, discretization error and Richardson extrapolation	Reproduced. The $O(h^2)$ rate measured at 3.998
	Grübel and Hermesmeier (1999)	1999	ASTIN Bulletin	50	Tables 1 and 2, the L1 sums, the Theorem 2 bound	Reproduced, including the bound at 0.4641
	Pittarello et al. (2024)	2024	Annals of Actuarial Science	1	Section 3.5 examples, Tables 1 to 3	Reproduced. Table 1 is a truncation study, Table 2 re-timed on one machine, Table 3's reference values misprinted
	Embrechts and Frei (2009)	2009	Mathematical Methods of Operations Research	94	Table 1, three arithmetization designs	Reproduced exactly, all twelve values
	Schaller and Temnov (2008)	2008	Computational Methods in Applied Mathematics	23	Table 4.3, operational risk quantiles	Reproduced to 2.2e-4, all twelve distinct values
	Bruno et al. (2006)	2006	Applied Mathematics and Computation	8	Tables 6, 7 and 8	Reproduced exact at six decimals. 92,832 convolutions in seconds against the paper's 1,950
	Gollini and Rougier (2016)	2016	The Journal of Operational Risk	1	Figures 1 to 3 and 5 to 7, bounds on a 32,060 event hurricane ELT	Reproduced. The exact answer the authors ruled out takes under a second, and the moment bound is 4 to 10 times high
	Wilson and Keich (2016)	2016	Computational Statistics and Data Analysis	8	Table 2, the FFT-C columns	Reproduced. The finding is detectability, not error size. The Section 1 example is now computed exactly in-library
	Wilson and Keich (2017)	2017	Journal of Computational and Graphical Statistics	5	The worst case, and its repair	Reproduced to the digit and the repair demonstrated, both now certified against exact arithmetic
Rare events	Asmussen et al. (2016)	2016	Scandinavian Journal of Statistics	36	Twelve tables, CDF and PDF	Reproduced from five build calls, and one bad row adjudicated
	Ben Amar et al. (2023)	2023	Statistics and Computing	7	The seven probabilities the paper actually quotes	Reproduced. The whole left tail curve costs what one estimate costs them
	Ben Rached et al. (2024)	2024	Statistics and Computing	0	Table 6, and Theorem 3's floor	Reproduced, and the floor confirmed on the library itself
	Liu et al. (2026)	2026	Operations Research Letters	0	Tables 1, 2 and C.1	Three of four reference points reproduced. The fourth is below the floor
	Commentary				The rare-event group as a whole	Where these methods earn their keep, and where they do not
Did not reproduce	Jin et al. (2016)	2016	Scandinavian Actuarial Journal	14	Nothing	Not a reproduction. The evidence, and the feature that would change it

Counts are from OpenAlex, retrieved 2026-08-08 (for Pittarello et al. (2024), 2026-10-11, and for Gollini and Rougier (2016), 2026-10-21), except for the Casualty Actuarial Society (CAS) literature. The CAS papers are not indexed and their counts come from a manual Google Scholar lookup. The counts are a snapshot and do not auto-update on rebuild. They should be read relative to the year of publication, and they say how widely each result has been taken up, which is a different question from whether it reproduces.

The two most cited papers here by an order of magnitude, Wang (2000) and Wang (2002), are the two whose reproductions are the shortest and least eventful. The papers that changed the library, Ben Rached et al. (2024), Wilson and Keich (2016) and Wilson and Keich (2017), have 0, 8 and 5 citations between them. Thus, being useful to an implementation and being cited are close to independent!

The final chapter is reference rather than reproduction: it records where each published severity curve shipped in `library.agg` came from, what the source paper uses it for, and every transcription decision behind the DecL.

5 Reading and rebuilding

Each chapter stands alone and they can read independently. Two pairs reward being read together: Homer and Clark (2003) and Jin et al. (2016) sit on opposite sides of the same feature boundary, dependence through a joint severity under a shared count against dependence through a common shock on the counts, and Wilson and Keich (2016) and Wilson and Keich (2017) are companion papers on the same failure, the first on adversarial inputs and the second on ordinary ones.

Rendering requires a Python environment with `aggregate` and its documentation extras installed, since every chapter executes. The two working documents in this folder, `README.md` and `plan-examples.md`, are notes rather than chapters and are deliberately outside the book: the first is a folder index and the second is the candidate assessment, which carries the tier ratings for the papers that were considered and rejected.

Arbenz, P., Embrechts, P., & Puccetti, G. (2011). The AEP algorithm for the fast computation of the distribution of the sum of dependent random variables. *Bernoulli*, *17*(2), 562–591. https://doi.org/10.3150/10-BEJ284

Asmussen, S., Jensen, J. L., & Rojas-Nandayapa, L. (2016). Exponential Family Techniques for the Lognormal Left Tail. *Scandinavian Journal of Statistics*, *43*(3), 774–787. https://doi.org/10.1111/sjos.12203

Bear, R. A., & Nemlick, K. J. (1990). Pricing the impact of adjustable features and loss sharing provisions of reinsurance treaties. *Proceedings of the Casualty Actuarial Society*, *77*(147), 86–87. https://doi.org/10.1016/0167-6687(93)91078-9

Ben Rached, N., Hoel, H., & Meo, J. V. (2024). A fast and accurate numerical method for the left tail of sums of independent random variables. *Statistics and Computing*. https://doi.org/10.1007/s11222-024-10514-5

Bodoff, N. M. (2017). An Actuarial Model of Excess of Policy Limits Losses. *Variance*, *11*. https://doi.org/10.66573/001c.116059

Bruno, M. G., Camerini, E., Manna, A., & Tomassetti, A. (2006). A new method for evaluating the distribution of aggregate claims. *Applied Mathematics and Computation*. https://doi.org/10.1016/j.amc.2005.09.038

Denuit, M. M., Dhaene, J., Ghossoub, M., & Robert, C. Y. (2025). Comonotonicity and Pareto optimality, with application to collaborative insurance. *Insurance: Mathematics and Economics*, *120*, 1–16. https://doi.org/10.1016/j.insmatheco.2024.11.001

Denuit, M. M., Ortega-Jiménez, P., & Robert, C.-Y. (2025). Conditional expectations given the sum of independent random variables with regularly varying densities. *ASTIN Bulletin*. https://doi.org/10.1017/asb.2025.11

Gollini, I., & Rougier, J. (2016). Rapidly bounding the exceedance probabilities of high aggregate losses. *The Journal of Operational Risk*, 1–33. https://doi.org/10.21314/JOP.2016.179

Homer, D. L., & Clark, D. R. (2003). Insurance Applications of Bivariate Distributions. *Proceedings of the Casualty Actuarial Society*, *90*(iid), 274–307. http://www.casact.org/pubs/proceed/proceed03/03274.pdf

Jin, T., Provost, S. B., & Ren, J. (2016). Moment-based density approximations for aggregate losses. *Scandinavian Actuarial Journal*, *2016*(3), 216–245. https://doi.org/10.1080/03461238.2014.921640

Mack, T., & Fackler, M. (2003). Exposure-rating in liability reinsurance. *Blätter der DGVFM*, *26*(2), 229–238.

Mata, A. J., & Verheyen, M. A. (2005). An Improved Method for Experience Rating Excess of Loss Treaties using Exposure Rating Techniques. *CAS Forum*.

Pittarello, G., Luini, E., & Marchione, M. M. (2024). GEMAct: a Python package for non-life (re) insurance modeling. *Annals of Actuarial Science*, 1–37. https://doi.org/10.1017/S1748499524000022

Venter, G. G. (1983). Transformed beta and gamma distributions and aggregate losses. *Proceedings of the Casualty Actuarial Society*, *70*(133), 289–308.

Wang, S. S. (2000). A class of distorion operators for pricing financial and insurance risks. *Journal of Risk and Insurance*, *67*(1), 15–36.

Wang, S. S. (2002). A Universal Framework for Pricing Financial and Insurance Risks. *ASTIN Bulletin*, *32*(2), 213–234. https://doi.org/10.2143/AST.32.2.1027

Wilson, H., & Keich, U. (2016). Accurate pairwise convolutions of non-negative vectors via FFT. *Computational Statistics and Data Analysis*, *101*, 300–315. https://doi.org/10.1016/j.csda.2016.03.010

Wilson, H., & Keich, U. (2017). Accurate small tail probabilities of sums of iid lattice-valued random variables via FFT. *Journal of Computational and Graphical Statistics*, *26*(1), 223–229.