Efficient Simulation of Ruin Probabilities When Claims are Mixtures of Heavy and Light Tails

Eleni Vatamidou

jointed work with Hansjörg Albrecher, Martin Bladt

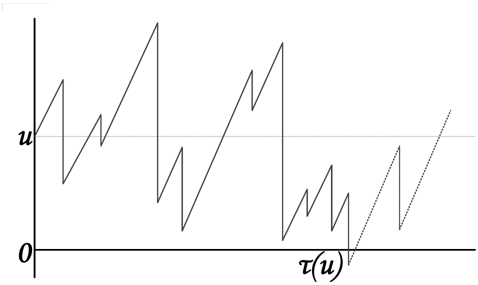
AUEB Stats Seminars – November 12, 2020

UNIL | Université de Lausanne

Insurance company with:

- 2 opposing cash flows
  - incoming cash premiums
  - outgoing claims, and
- an initial capital.

Question of interest: What is the probability of bankruptcy?



- λ: Poisson arrival rate ($N(t)$)
- U_k: i.i.d. claim sizes (G)
- $c = 1$: premium rate
- u: initial capital.

Risk reserve process

$$R(t) = u + t - \sum_{k=1}^{N(t)} U_k.$$

Ruin probability for the Cramér-Lundberg risk model

- λ: Poisson arrival rate ($N(t)$)
- U_k: i.i.d. claim sizes (G)
- $c = 1$: premium rate
- u: initial capital.

Risk reserve process

$$R(t) = u + t - \sum_{k=1}^{N(t)} U_k.$$

Claim surplus: $M = \sup_{0 \leq t < \infty} \left(u - R(t)\right)$

Ruin probability

$$\psi(u) = \mathbb{P}(M > u).$$

Pollaczek-Khinchine (PK) formula

PK formula

If $\rho = \lambda \mathbb{E} U < 1$ (safety loading condition) then

$$\psi(u) = 1 - (1-\rho) \sum_{k=0}^{\infty} \rho^k (G^e)^{*k}(u),$$

where $G^e(u) = \int_0^u \big(1 - G(x)\big)\, dx / \mathbb{E} U$ (stationary excess claim size distribution).

Pollaczek-Khinchine (PK) formula

PK formula

If $\rho = \lambda \mathbb{E} U < 1$ (safety loading condition) then

$$\psi(u) = 1 - (1-\rho) \sum_{k=0}^{\infty} \rho^k (G^e)^{*k}(u),$$

where $G^e(u) = \int_0^u \big(1 - G(x)\big) dx / \mathbb{E} U$ (stationary excess claim size distribution).

Ways to calculate the ruin probability:

- closed-form solutions (e.g. algebraic, analytic/Laplace transforms)
- approximations (e.g. PH, asymptotic/tail probabilities)
- simulations

Pollaczek-Khinchine (PK) formula

PK formula

If $\rho = \lambda \mathbb{E} U < 1$ (safety loading condition) then

$$\psi(u) = 1 - (1-\rho) \sum_{k=0}^{\infty} \rho^k (G^e)^{*k}(u),$$

where $G^e(u) = \int_0^u \left(1 - G(x)\right) dx / \mathbb{E} U$ (stationary excess claim size distribution).

Ways to calculate the ruin probability:

- closed-form solutions (e.g. algebraic, analytic/Laplace transforms)
- approximations (e.g. PH, asymptotic/tail probabilities)
- simulations

Corrected phase-type approximations

Combining the best of 2 worlds

Phase-type approximations	Asymptotic approximations
high accuracy computationally tractable error bounds	correct tail behavior

Corrected phase-type approximations

Combining the best of 2 worlds

Phase-type approximations	Asymptotic approximations
high accuracy computationally tractable error bounds	correct tail behavior

Fitting distributions to data

1. higher order statistics for tail probabilities (r.v. C)
2. remaining data set phase-type distribution (r.v. B)

Claim size distribution

$$G(x) = (1-\epsilon)F_p(x) + \epsilon F_h(x), \quad x \geq 0.$$

New series expansion for the ruin probability

Theorem

We have

$$\psi(u) = \frac{1-\rho}{1-\rho^\bullet}\,\psi^\bullet(u) + \frac{1-\rho}{1-\rho^\bullet}\sum_{k=1}^{\infty}\left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^k \mathscr{A}_k(u),$$

where $\mathscr{A}_k(u) = \mathbb{P}(M_0^\bullet + M_1^\bullet + \cdots + M_k^\bullet + C_1^e + \cdots + C_k^e > u)$ *and* $M_k^\bullet \stackrel{\mathfrak{D}}{=} M^\bullet$*. This expansion converges for all values of* u*.*

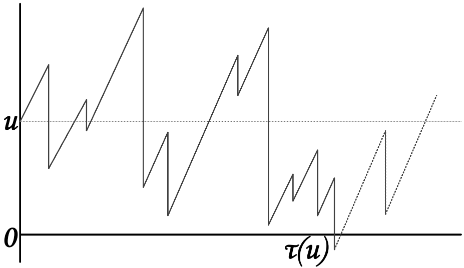
New series expansion for the ruin probability

Theorem

We have

$$\psi(u) = \frac{1-\rho}{1-\rho^\bullet}\,\psi^\bullet(u) + \frac{1-\rho}{1-\rho^\bullet}\sum_{k=1}^{\infty}\left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^k \mathscr{A}_k(u),$$

where $\mathscr{A}_k(u) = \mathbb{P}(M_0^\bullet + M_1^\bullet + \cdots + M_k^\bullet + C_1^e + \cdots + C_k^e > u)$ *and* $M_k^\bullet \overset{\mathfrak{D}}{=} M^\bullet$. *This expansion converges for all values of* u.



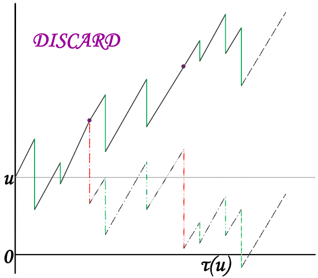
New series expansion for the ruin probability

Theorem

We have

$$\psi(u) = \frac{1-\rho}{1-\rho^\bullet}\,\psi^\bullet(u) + \frac{1-\rho}{1-\rho^\bullet}\sum_{k=1}^{\infty}\left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^k \mathscr{A}_k(u),$$

where $\mathscr{A}_k(u) = \mathbb{P}(M_0^\bullet + M_1^\bullet + \cdots + M_k^\bullet + C_1^e + \cdots + C_k^e > u)$ *and* $M_k^\bullet \overset{\mathfrak{D}}{=} M^\bullet$. *This expansion converges for all values of* u.



Question: can this series expansion improve simulations?

$$\psi(u) = \underbrace{\frac{1-\rho}{1-\rho^\bullet}\,\psi^\bullet(u) + \frac{1-\rho}{1-\rho^\bullet}\,\frac{\epsilon\theta}{1-\rho^\bullet}\,\mathbb{P}(M_0^\bullet + M_1^\bullet + C_1^e > u)}_{\textit{explicit}}$$

$$+ \frac{1-\rho}{1-\rho^\bullet}\sum_{k=2}^{\infty}\left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^k \mathscr{A}_k(u).$$

Question: can this series expansion improve simulations?

$$\psi(u) = \underbrace{\frac{1-\rho}{1-\rho^\bullet}\psi^\bullet(u) + \frac{1-\rho}{1-\rho^\bullet}\frac{\epsilon\theta}{1-\rho^\bullet}\mathbb{P}(M_0^\bullet + M_1^\bullet + C_1^e > u)}_{explicit}$$

$$+ \frac{1-\rho}{1-\rho^\bullet}\sum_{k=2}^{\infty}\left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^k \mathscr{A}_k(u).$$

$$\varphi(u) := \frac{1-\rho}{1-\rho^\bullet}\sum_{k=2}^{\infty}\left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^k \mathscr{A}_k(u) = \left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^2 \mathbb{E}\mathscr{A}_{N+2}(u)$$

$$= \left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^2 \mathbb{P}(M_0^\bullet + M_1^\bullet + \cdots + M_{N+2}^\bullet + C_1^e + \cdots + C_{N+2}^e > u),$$

Question: can this series expansion improve simulations?

$$\psi(u) = \underbrace{\frac{1-\rho}{1-\rho^\bullet}\psi^\bullet(u) + \frac{1-\rho}{1-\rho^\bullet}\frac{\epsilon\theta}{1-\rho^\bullet}\mathbb{P}(M_0^\bullet + M_1^\bullet + C_1^e > u)}_{explicit}$$

$$+ \frac{1-\rho}{1-\rho^\bullet}\sum_{k=2}^{\infty}\left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^k \mathscr{A}_k(u).$$

$$\varphi(u) := \frac{1-\rho}{1-\rho^\bullet}\sum_{k=2}^{\infty}\left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^k \mathscr{A}_k(u) = \left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^2 \mathbb{E}\mathscr{A}_{N+2}(u)$$

$$= \left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^2 \mathbb{P}(M_0^\bullet + M_1^\bullet + \cdots + M_{N+2}^\bullet + C_1^e + \cdots + C_{N+2}^e > u),$$

Simulate:

$$V \stackrel{\mathfrak{D}}{=} M_0^\bullet + M_1^\bullet + C_1^e + \sum_{k=2}^{N+2}(M_k^\bullet + C_k^e), \quad N \sim Geom\left(\frac{1-\rho}{1-\rho^\bullet}\right).$$

Idea of the control variate techniques:

1. We must simulate a r.v. $Z(u)$ such that $\varphi(u) = \mathbb{E}Z(u)$.
2. We find another r.v. $W(u)$ that has a known expectation $\mathbb{E}W(u)$ and is strongly correlated with $Z(u)$.
3. We simulate them together, i.e. we take $\big(Z^{(i)}(u), W^{(i)}(u)\big)$, $i = 1, 2, \ldots, \kappa$, are independent copies of $\big(Z(u), W(u)\big)$.
4. We calculate

$$\hat{\varphi}_\kappa(u) := \hat{z}_\kappa(u) + \hat{\alpha}_\kappa\big(\hat{w}_\kappa(u) - \mathbb{E}W(u)\big),$$

where

$$\hat{z}_\kappa(u) = \frac{\sum_{i=1}^{\kappa} Z^{(i)}(u)}{\kappa}, \qquad \hat{w}_\kappa(u) = \frac{\sum_{i=1}^{\kappa} W^{(i)}(u)}{\kappa},$$

$$\hat{\alpha}_\kappa = -\frac{\sum_{i=1}^{\kappa}\big(Z^{(i)}(u) - \hat{z}_\kappa(u)\big)\big(W^{(i)}(u) - \hat{w}_\kappa(u)\big)}{\sum_{i=1}^{\kappa}\big(W^{(i)}(u) - \hat{w}_\kappa(u)\big)^2}.$$

Control variate: max of heavy tails

$$\varphi(u) = \left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^2 \mathbb{P}(\underbrace{M_0^\bullet + M_1^\bullet + \cdots + M_{N+2}^\bullet + C_1^e + \cdots + C_{N+2}^e}_{\stackrel{\mathfrak{D}}{=} V} > u)$$

Obviously: $Z(u) = \left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^2 \mathbb{1}_{\{V>u\}}$.

Control variate: max of heavy tails

$$\varphi(u) = \left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^2 \mathbb{P}(\underbrace{M_0^\bullet + M_1^\bullet + \cdots + M_{N+2}^\bullet + C_1^e + \cdots + C_{N+2}^e}_{\stackrel{\mathfrak{D}}{=} V} > u)$$

Obviously: $Z(u) = \left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^2 \mathbb{1}_{\{V>u\}}$.

We define: $V_n := \max\{C_1^e, \ldots, C_{N+2}^e\}\mathbb{1}_{\{N+2\leq n\}}$, for fixed n.

Control variate: $W_n(u) = \left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^2 \mathbb{1}_{\{V_n>u\}}$, for fixed n.

$$\varphi_n(u) = \mathbb{E}\,W_n(u) = \left(\frac{1-\rho}{1-\rho^\bullet}\right)\sum_{k=2}^{n}\left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^k \mathbb{P}\big(\max\{C_1^e, \ldots, C_k^e\} > u\big)$$

Alternative control variate: Conditional Monte Carlo

$$V \overset{\mathfrak{D}}{=} \underbrace{M_0^\bullet}_{=X_0^\star} + \underbrace{M_1^\bullet + C_1^e}_{=X_1} + \sum_{k=2}^{N+2} (\underbrace{M_k^\bullet + C_k^e}_{=X_k}) \overset{\mathfrak{D}}{=} X_0^\star + \sum_{k=1}^{N+2} X_k$$

If: $m_k := \max\{X_1, \ldots, X_k\}$, $\overline{F}_X$ is the c.c.d.f. of X_k's, and $S_\ell = \sum_{k=1}^{\ell} X_k$, $S_0 = 0$

Now: $Z^\star(u) = \left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^2 (N+2)\overline{F}_X\big(m_{N+1} \vee (u - X_0^\star - S_{N+1})\big)$.

AK control variate: $W^\star(u) = \left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^2 (N+2)\overline{F}_X(u)$

$$\varphi^\star(u) = \left(\frac{\epsilon\theta}{1-\rho^\bullet}\right)^2 \left(\frac{\epsilon\theta}{1-\rho} + 2\right)\overline{F}_X(u)$$

Numerical experiments

- Mixture claim size distribution
  - PH: $\overline{F_p}(u) = \overline{F_p^e}(u) = e^{-\mu u}$, and $\mu_B = 1/\mu$ $\underline{(\mu = 3)}$
  - HT: shifted Pareto with shape $a > 1$ and scale $b > 0$, i.e. $\overline{F_h}(u) = (1 + u/b)^{-a}$ and $\overline{F_h^e}(u) = (1 + u/b)^{-(a-1)}$, $u \geq 0$, with $\mu_C = b/(a-1)$ $\underline{(b = 1)}$
  - Perturbation parameter: $\epsilon \in \{0.1, 0.7\}$.
- Focus on $\rho \in \{0.9, 0.99, 0.999\}$.
- Order of $\varphi_n(u)$ equal to $n = 100$
- Number of simulations is $\kappa = 10,000$.

Figures for 1st control variate (max of heavy tails)

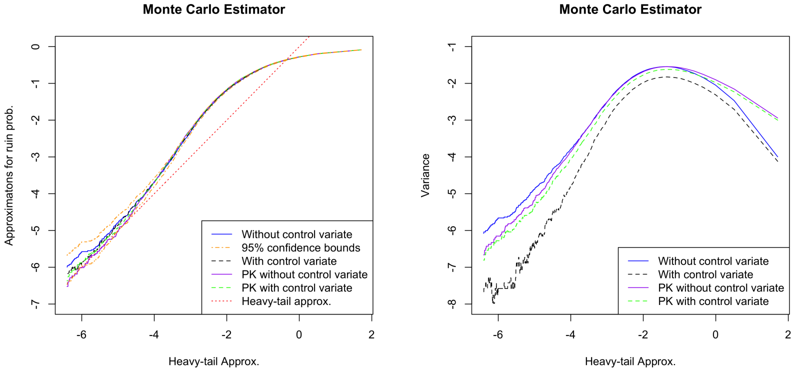


Figure: Plotted in a log-log scale. Model parameters: $a = 2$, $\epsilon = 0.1$, and $\rho = 0.99$.

Figures for 2nd control variate (AK estimator)

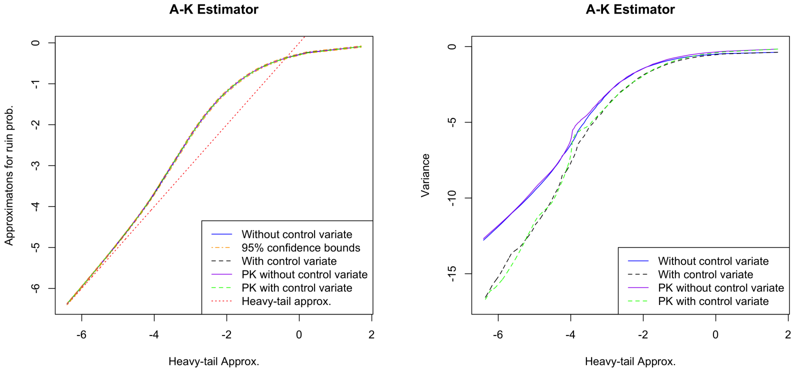


Figure: Plotted in a log-log scale. Model parameters: $a = 2$, $\epsilon = 0.1$, and $\rho = 0.99$.

Conclusions

- We introduced an alternative series expansion for the PK formula in the Cramér-Lundberg model for mixture claim sizes
- Significant improvement of simulation algorithms based on this series
- Proposed a control variate technique: fast and preferable in the heavy traffic regime
- Variance reduction is better with AK conditional Monte Carlo technique but the method is significantly slower
- For other mixtures that the 2nd term of the ruin probability cannot be evaluated, it can also be simulated
- Extension to the Sparre Andersen model,which also has a PK-type formula with respect to the ladder height distribution

Ευχαριστώ