Estimation non paramétrique de la fonction de Gerber–Shiu dans le modèle de Cramér–Lundberg

Florian Dussap

Colloque Jeunes Probabilistes et Statisticiens – 2021

The compound Poisson risk model [Asmussen and Albrecher, 2010]

Let $(U_t)_{t\geqslant 0}$ be the reserve process of an insurance company. In the compound Poisson risk model, this process is given by:

$$U_t = u + ct - \sum_{i=1}^{N_t} X_i,$$

where:

- $u \geqslant 0$ is the initial reserve,
- $c > 0$ is the premium rate,
- the claim number process $(N_t)_{t\geqslant 0}$ is a homogeneous Poisson process with intensity λ,
- the individual claim sizes $(X_i)_{i\geqslant 1}$ are positive, i.i.d. with density f and mean μ, independent of $(N_t)_{t\geqslant 0}$.

The Gerber–Shiu function

We denote the time of ruin by $\tau \coloneqq \inf\{t \geqslant 0 \mid U_t < 0\} \in \mathbb{R}_+ \cup \{\infty\}$.

Assumption (Safety Loading Condition)

A1 We assume that $c > \lambda\mu$. Introducing the parameter $\theta \coloneqq \frac{\lambda\mu}{c}$, the previous condition is equivalent to $\theta < 1$.

Under the SLC, we have $\mathbb{P}[\tau < \infty] < 1$.

The Gerber–Shiu function, also called the *Expected Discounted Penalty Function (EDPF)*, is defined as:

$$\phi(u) \coloneqq \mathbb{E}\left[\mathrm{e}^{-\delta\tau} w(U_{\tau^-}, |U_\tau|)\,\mathbf{1}_{\{\tau<\infty\}} \,\middle|\, U_0 = u\right],$$

where $\delta \geqslant 0$ is a discounting force of interest, and $w \colon \mathbb{R}_+^2 \to \mathbb{R}_+$ is a penalty function.

Observations and goal

We assume that c is known but the parameters (λ, μ, f) of the compound Poisson process are not. We suppose we have access to a trajectory of the reserve process $(U_t)_{t \in [0,T]}$ on a time interval $[0, T]$, on which we observe:

$$N_T \text{ and } X_1, \dots, X_{N_T}.$$

Goal

We want to estimate the Gerber–Shiu function from the observations $(N_T, X_1, \dots, X_{N_T})$ with c known but (λ, μ, f) unknown.

Key result

Theorem ([Gerber and Shiu, 1998])

Under Assumption A1, the EDPF satisfies the equation:

$$\phi = \phi * g + h,$$

with:

$$g(x) := \frac{\lambda}{c} \int_x^{+\infty} \mathrm{e}^{-\rho_\delta (y-x)} f(y)\,\mathrm{d}y,$$

$$h(u) := \frac{\lambda}{c} \int_u^{+\infty} \mathrm{e}^{-\rho_\delta (x-u)} \left(\int_x^{+\infty} w(x, y-x) f(y)\,\mathrm{d}y \right) \mathrm{d}x,$$

and ρ_δ the non-negative solution of the Lundberg equation:

$$cs - \lambda\big(1 - \mathcal{L}f(s)\big) = \delta.$$

When $\delta = 0$, we have $\rho_\delta = 0$ as well.

Estimation strategy in a nutshell

1. We have:

$$g(x) = \frac{\lambda}{c}\,\mathbb{E}\Big[\mathrm{e}^{-\rho_\delta (X-x)}\mathbf{1}_{\{X>x\}}\Big],$$
$$h(u) = \frac{\lambda}{c}\,\mathbb{E}\left[\left(\int_u^X \mathrm{e}^{-\rho_\delta (x-u)} w(x, X-x)\,\mathrm{d}x\right)\mathbf{1}_{\{X>u\}}\right],$$

   with ρ_δ solution of $cs - \lambda\big(1 - \mathbb{E}\big[\mathrm{e}^{-sX}\big]\big) = \delta$. These quantities can be estimated from the observations.

2. Once we have estimated g and h, we solve the equation $\phi = \phi * g + h$ to estimate ϕ.

Following the work of [Zhang and Su, 2018], we estimate g and h by projection on the Laguerre basis.

Laguerre basis decomposition

The Laguerre functions $(\psi_k)_{k\in\mathbb{N}}$ are defined as:

$$\forall x \in \mathbb{R}_+, \quad \psi_k(x) := \sqrt{2}\, L_k(2x)\, \mathrm{e}^{-x}, \quad L_k(x) := \sum_{j=0}^{k} \binom{k}{j} \frac{(-x)^j}{j!}.$$

The Laguerre functions form a basis of $\mathrm{L}^2(\mathbb{R}_+)$. We decompose ϕ, g and h on this basis:

$$\phi = \sum_{k=0}^{+\infty} a_k\, \psi_k, \qquad g = \sum_{k=0}^{+\infty} b_k\, \psi_k, \qquad h = \sum_{k=0}^{+\infty} c_k\, \psi_k,$$

$$a_k = \langle \phi, \psi_k \rangle, \qquad b_k = \langle g, \psi_k \rangle, \qquad c_k = \langle h, \psi_k \rangle.$$

Estimation of g and h

Assumption

A2 $\int_0^{+\infty}(1+x)\int_x^{\infty} w(x, y-x) f(y)\,\mathrm{d}y\,\mathrm{d}x$ is finite. ($\implies h \in \mathrm{L}^2(\mathbb{R}_+)$)

A3 Let $W(X) := \int_0^X \left(\int_u^X w(x, X-x)\,\mathrm{d}x\right)^2 \mathrm{d}u$. If $\delta = 0$ we assume that $\mathbb{E}[W(X)]$ is finite, if $\delta > 0$ we assume that $\mathbb{E}[W(X)^2]$ is finite.

The coefficients of g and h are given by:

$$b_k = \frac{\lambda}{c}\,\mathbb{E}\left[\int_0^X \mathrm{e}^{-\rho_\delta(X-x)}\psi_k(x)\,\mathrm{d}x\right],$$

$$c_k = \frac{\lambda}{c}\,\mathbb{E}\left[\int_0^X \left(\int_u^X \mathrm{e}^{-\rho_\delta(x-u)} w(x, X-x)\,\mathrm{d}x\right)\psi_k(u)\,\mathrm{d}u\right].$$

We estimate the coefficients of g and h by empirical means:

$$\hat{b}_k = \frac{1}{cT} \sum_{i=1}^{N_T} \int_0^{X_i} \mathrm{e}^{-\hat{\rho}_\delta (X_i - x)}\, \psi_k(x)\, \mathrm{d}x,$$

$$\hat{c}_k = \frac{1}{cT} \sum_{i=1}^{N_T} \int_0^{X_i} \left(\int_u^{X_i} \mathrm{e}^{-\hat{\rho}_\delta (x-u)}\, w(x, X_i - x)\, \mathrm{d}x \right) \psi_k(u)\, \mathrm{d}u,$$

with $\hat{\rho}_\delta$ the non-negative solution of the empirical Lundberg equation:

$$cs - \frac{N_T}{T} \left(1 - \frac{1}{N_T} \sum_{i=1}^{N_T} \mathrm{e}^{-sX_i} \right) = \delta.$$

For $m \in \mathbb{N}_+$, the projection estimators of g and h are:

$$\hat{g}_m := \sum_{k=0}^{m-1} \hat{b}_k\, \psi_k, \qquad \hat{h}_m := \sum_{k=0}^{m-1} \hat{c}_k\, \psi_k.$$

Bias-variance decomposition of the MISE

We quantify the quality of an estimator by its Mean Integrated Squared Error (MISE):

$$\mathbb{E}\|g - \hat{g}_m\|_{\mathsf{L}^2}^2.$$

The MISE of an estimator can be decomposed as the sum of a bias term and a variance term:

$$\begin{aligned}\mathbb{E}\|g - \hat{g}_m\|_{\mathsf{L}^2}^2 &= \left\|g - \Pi_{S_m}(g)\right\|_{\mathsf{L}^2}^2 + \mathbb{E}\left\|\hat{g}_m - \Pi_{S_m}(g)\right\|_{\mathsf{L}^2}^2 \\ &= \mathrm{dist}_{\mathsf{L}^2}^2(g, S_m) + \sum_{k=0}^{m-1} \mathbb{E}\Big[(\hat{b}_k - b_k)^2\Big],\end{aligned}$$

with $S_m := \mathrm{Span}(\psi_0, \dots, \psi_{m-1})$.

Proposition

Under Assumptions A1, A2 and A3, if $\delta = 0$ then it holds:

$$\mathbb{E}\|g - \hat{g}_m\|^2_{\mathrm{L}^2} \leqslant \mathrm{dist}^2_{\mathrm{L}^2}(g, S_m) + \frac{\lambda}{c^2 T}\mathbb{E}[X],$$
$$\mathbb{E}\|h - \hat{h}_m\|^2_{\mathrm{L}^2} \leqslant \mathrm{dist}^2_{\mathrm{L}^2}(h, S_m) + \frac{\lambda}{c^2 T}\mathbb{E}[W(X)],$$

and if $\delta > 0$ then it holds:

$$\mathbb{E}\|g - \hat{g}_m\|^2_{\mathrm{L}^2} \leqslant \mathrm{dist}^2_{\mathrm{L}^2}(g, S_m) + \frac{C(\lambda)}{c^2 T}\left(\mathbb{E}[X] + \frac{\mathbb{E}\left[X^2\right]^{\frac{1}{2}}}{(1-\theta)^2\delta^2}\right),$$

$$\mathbb{E}\|h - \hat{h}_m\|^2_{\mathrm{L}^2} \leqslant \mathrm{dist}^2_{\mathrm{L}^2}(h, S_m) + \frac{C(\lambda)}{c^2 T}\left(\mathbb{E}[W(X)] + \frac{\mathbb{E}\left[W(X)^2\right]^{\frac{1}{2}}}{(1-\theta)^2\delta^2}\right),$$

where $C(\lambda) \asymp \lambda^2$.

Interlude: Laguerre deconvolution [Mabon, 2017]

The Laguerre functions satisfy the relation:

$$\forall j,k \in \mathbb{N}, \quad \psi_j * \psi_k = 2^{-\frac{1}{2}}(\psi_{j+k} - \psi_{j+k+1}).$$

Using this relation, one can show that if f and g are two functions on $\mathbb{R}_+$ then their Laguerre coefficients satisfy:

$$c(f*g) = c(f) * \Delta(g), \quad \Delta_k(g) := \begin{cases} 2^{-\frac{1}{2}}\left(c_k(g) - c_{k-1}(g)\right) & : k \geqslant 1, \\ 2^{-\frac{1}{2}}\, c_0(g) & : k = 0. \end{cases}$$

If $\mathbf{c}_m(f)$ denotes the vector of the first m coefficients of f, we have:

$$\mathbf{c}_m(f*g) = \mathbf{G}_m \times \mathbf{c}_m(f), \quad \mathbf{G}_m := \begin{bmatrix} \Delta_0 & 0 & 0 & 0 & 0 \\ \Delta_1 & \Delta_0 & 0 & 0 & 0 \\ \Delta_2 & \Delta_1 & \Delta_0 & 0 & 0 \\ \cdots & \cdots & \cdots & \Delta_0 & 0 \\ \Delta_{m-1} & \Delta_{m-2} & \cdots & \cdots & \Delta_0 \end{bmatrix}.$$

Laguerre deconvolution estimator

If we use the convolution property of the Laguerre functions in the equation $\phi = \phi * g + h$, we obtain the following relation between the coefficients of ϕ, g and h:

$$\mathbf{c}_m = \mathbf{A}_m \times \mathbf{a}_m \iff \mathbf{a}_m = \mathbf{A}_m^{-1} \times \mathbf{c}_m,$$

with $\mathbf{A}_m := \mathbf{Id}_m - \mathbf{G}_m$.

Assumption

A4 $(b_{k+1} - b_k)_{k\in\mathbb{N}} \in \ell^1(\mathbb{N})$.

Lemma

Under Assumption A1 and A4, we have $\|\mathbf{A}_m^{-1}\|_{\mathrm{op}} \leqslant \frac{2}{1-\|g\|_{\mathsf{L}^1}} \leqslant \frac{2}{1-\theta}$.

Definition

For $\theta_0 < 1$ a truncation parameter, we estimate ϕ by:

$$\hat{\phi}_m := \sum_{k=0}^{m-1} \hat{a}_k \, \psi_k, \quad \hat{\mathbf{a}}_m := \widetilde{\mathbf{A}}_m^{-1} \times \hat{\mathbf{c}}_m, \quad \widetilde{\mathbf{A}}_m^{-1} := \hat{\mathbf{A}}_m^{-1} \mathbf{1}_{\left\{\|\hat{\mathbf{A}}_m^{-1}\|_{\mathrm{op}} \leqslant \frac{2}{1-\theta_0}\right\}}.$$

Proposition

Under Assumptions A1, A2, A3, and A4, if $\theta < \theta_0$ then it holds:

$$\mathbb{E}\|\phi - \hat{\phi}_m\|_{\mathsf{L}^2}^2 \leqslant \mathrm{dist}_{\mathsf{L}^2}^2(\phi, S_m) + C\,\frac{m}{T},$$

where C is a constant depending on λ, c, θ, $\mathbb{E}[X]$, $\mathbb{E}[W(X)]$ and $\theta_0 - \theta$; and also δ, $\mathbb{E}[X^2]$, $\mathbb{E}[W(X)^2]$ if $\delta > 0$.

Laguerre–Fourier estimator

Since $\phi = \phi * g + h$, we have $\mathcal{F}\phi = \frac{\mathcal{F}h}{1-\mathcal{F}g}$. We compute the coefficients of ϕ using Plancherel theorem:

$$a_k = \langle \phi, \psi_k \rangle = \frac{1}{2\pi} \langle \mathcal{F}\phi, \mathcal{F}\psi_k \rangle = \frac{1}{2\pi} \left\langle \frac{\mathcal{F}h}{1-\mathcal{F}g}, \mathcal{F}\psi_k \right\rangle.$$

Definition

For $\hat{g}$ and $\hat{h}$ two estimators of g and h, and for θ_0 a truncation parameter, we estimate ϕ by:

$$\hat{\phi}_{m_1,\hat{g},\hat{h}} := \sum_{k=0}^{m_1-1} \hat{a}_{k,\hat{g},\hat{h}} \psi_k, \quad \hat{a}_{k,\hat{g},\hat{h}} := \frac{1}{2\pi} \left\langle \frac{\mathcal{F}\hat{h}}{1-\widetilde{\mathcal{F}g}}, \mathcal{F}\psi_k \right\rangle,$$

$$\widetilde{\mathcal{F}g} := (\mathcal{F}\hat{g}) \mathbf{1}_{\{|\mathcal{F}\hat{g}| < \theta_0\}}.$$

Proposition

Under Assumption A1 and A2, if $\theta < \theta_0$ then it holds:

$$\|\phi - \hat{\phi}_{m_1,\hat{g},\hat{h}}\|_{\mathrm{L}^2}^2 \leqslant \mathrm{dist}_{\mathrm{L}^2}^2(\phi, S_{m_1}) + \frac{2}{(1-\theta_0)^2}\|h - \hat{h}\|_{\mathrm{L}^2}^2 + \frac{2\,\|h\|_{\mathrm{L}^1}^2}{(1-\theta_0)^2(1-\theta)^2}\left(1 + \frac{\|g\|_{\mathrm{L}^1}^2}{(\theta_0-\theta)^2}\right)\|g - \hat{g}\|_{\mathrm{L}^2}^2.$$

If we use the Laguerre projection estimators $\hat{g}_{m_2}$ and $\hat{h}_{m_3}$, we obtain the following result.

Corollary

Under Assumptions A1, A2 and A3, if $\theta < \theta_0$ then it holds:

$$\mathbb{E}\|\phi - \hat{\phi}_{m_1,m_2,m_3}\|_{\mathrm{L}^2}^2 \leqslant \mathrm{dist}_{\mathrm{L}^2}^2(\phi, S_{m_1}) + C\left(\mathrm{dist}_{\mathrm{L}^2}^2(g, S_{m_2}) + \mathrm{dist}_{\mathrm{L}^2}^2(h, S_{m_3}) + \frac{1}{T}\right).$$

Rates of convergence

Definition

For $s \in (0, +\infty)$, we define the Sobolev–Laguerre space as:

$$\mathrm{W}^s(\mathbb{R}_+) \coloneqq \left\{ v \in \mathrm{L}^2(\mathbb{R}_+) \,\middle|\, \sum_{k \in \mathbb{N}} \langle v, \psi_k \rangle^2 \, k^s < +\infty \right\}.$$

Theorem

We assume A1, A2, A3, A4, and we assume that $\theta < \theta_0$.

- ***Laguerre deconvolution:*** *If $\phi \in \mathrm{W}^s(\mathbb{R}_+)$, then choosing $m_{\text{opt}} \propto T^{\frac{1}{1+s}}$ yields:*

$$\mathbb{E}\|\phi - \hat{\phi}_{m_{\text{opt}}}\|_{\mathrm{L}^2}^2 \lesssim T^{-\frac{s}{1+s}}.$$

- ***Laguerre–Fourier:*** *If $\phi \in \mathrm{W}^{s_1}(\mathbb{R}_+)$, $g \in \mathrm{W}^{s_2}(\mathbb{R}_+)$ and $h \in \mathrm{W}^{s_3}(\mathbb{R}_+)$, then choosing $m_i > T^{\frac{1}{1+s_i}}$ for all $i \in \{1, 2, 3\}$ yields:*

$$\mathbb{E}\|\phi - \hat{\phi}_{m_1, m_2, m_3}\|_{\mathrm{L}^2}^2 \lesssim T^{-1}.$$

Conclusion

1. If ϕ belongs to a Sobolev–Laguerre space of any regularity, it is possible to estimate the EDPF with parametric rate.
2. The Laguerre deconvolution method fails to recover the parametric rate.
3. The absence of a bias-variance compromise raises questions about how to perform a model selection procedure.

References

Asmussen, S. and Albrecher, H. (2010).
Ruin probabilities, volume 14 of *Advanced series on statistical science and applied probability*.
World Scientific, Singapore ; New Jersey, 2nd edition.

Gerber, H. U. and Shiu, E. S. (1998).
On the Time Value of Ruin.
North American Actuarial Journal, 2(1):48–72.

Mabon, G. (2017).
Adaptive Deconvolution on the Non-negative Real Line: Adaptive deconvolution on R+.
Scandinavian Journal of Statistics, 44(3):707–740.

Zhang, Z. and Su, W. (2018).
A new efficient method for estimating the Gerber–Shiu function in the classical risk model.
Scandinavian Actuarial Journal, 2018(5):426–449.