Design of parametric insurance via machine learning and optimal combination with traditional insurance

joint work with A. Heranval (Inrae Avignon), D. Nkameni (Detralytics and Crest)

Excellence Chair CARE - Allianz France, Fondation du Risque, Ensae

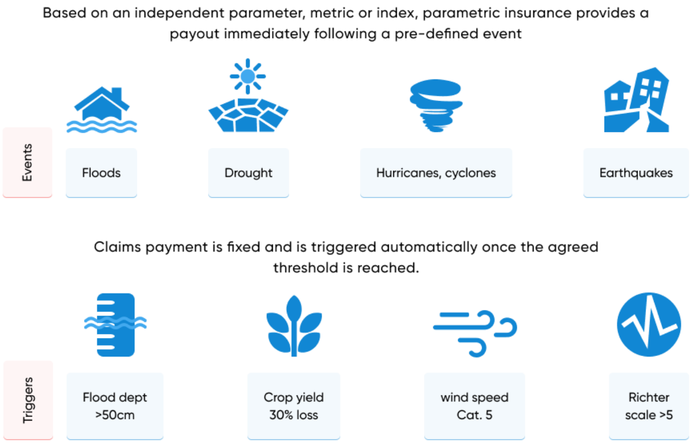
Olivier Lopez

Crest, Ensae Institut Polytechnique de Paris

June 2025

Institut Louis Bachelier

Parametric insurance in climate risk



Source : EDGE.

The cost of a cyber incident

The cost can be (roughly) decomposed into :

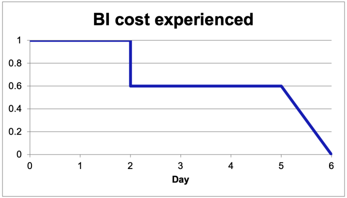
- cost that are directly related to the data exposed ;
- business interruption ;
- damage to reputation.

Example of Business interruption (Lloyd's and AIR, 2018, "Cloud down" report) :

- duration of the service interruption
- activation (or not) of a back-up plan
- progressive return to normal after restoration of the service.

The estimated cost is proportional to the duration (with a different factor depending on the different phases).

Figure 11: (C)BI cost experienced as a function of time



Institut Louis Bachelier

ENSAE IP PARIS

Parametric insurance

A policyholder experiences a loss Y.

"Traditional insurance" :

- the policyholder receives a compensation Y (minus the deductible) from the insurance company ;
- this usually requires expertise (and time).

Parametric insurance (or "index-based") :

- the payment is based on an index $f(\mathbf{W})$, where $\mathbf{W}$ are covariates
- $\mathbf{W}$ are observed just after the claim
- $f(\mathbf{W}) \leqslant Y$ (otherwise it is not insurance anymore) or at least with high probability
- no need for expertise, fast payment

Parametric insurance products for cyber : essentially related to the time of business interruption and/or volume of data exposed.

Institut Louis Bachelier

ENSAE IP PARIS

Outline

Institut Louis Bachelier

ENSAE

IP PARIS

Outline

Institut Louis Bachelier

Utility theory and index conception

Utility theory can be used to formalize the needs of the policyholder.

In the following, we consider a utility function which represents the satisfaction of the policyholder, which is a function

$$U : \mathbb{R} \to \mathbb{R},$$

with U strictly concave.

Consider a compensation Z paid to the policyholder.

The fortune of the policyholder after the claim is :

$$u - Y + Z - (1+\theta)E[Z],$$

where u is the initial fortune.

Here, $(1+\theta)E[Z]$ is the insurance premium (pure premium + a loading factor $\theta > 0$).

The idea is to find the product Z^* that maximizes

$$E\left[U\left(u - Y + Z - (1+\theta)E[Z]\right)\right],$$

over a set of possible products Z.

Demand in parametric insurance

Two possible contracts :

- Traditional insurance, compensation Y but with delay, seen as $\exp(-\tau)Y$ from the policyholder with $\tau \geqslant 0$;
- Parametric insurance : $\phi(\mathbf{W})$ (main example we consider : $\phi(\mathbf{W}) = \beta E[Y|\mathbf{W}]$ with $\beta < 1$).

Utility : all policyholders have an exponential utility function, that is, for $\alpha > 0$,

$$U_\alpha(x) = -\alpha \exp(-\alpha x).$$

A measure μ describes the distribution of the "risk aversion" α.

Demand

If N is the total size of the target population, the number of policyholders chosing the parametric product is $n \approx N\mu(A)$, where

$$A = \{\alpha \in \mathbb{R} : E\left[U_\alpha(\phi(\mathbf{W}) - Y - \pi_\phi) - U_\alpha(Y[\exp(-\tau) - 1] - \pi_Y)\right] > 0\},$$

where π_ϕ (resp. π_Y) is the price of the parametric contract (resp. traditional one).

Institut Louis Bachelier

ENSAE

IP PARIS

Outline

Institut Louis Bachelier

ENSAE IP PARIS

Viability of the contract

To achieve viability, we need the portfolio to be large enough, in the sense that

$$\mathbb{P}\left(\sum_{i=1}^{n}\phi(\mathbf{W}_i) - n\pi_\phi > 0\right) \leqslant \varepsilon \quad (\mathcal{C}_\varepsilon),$$

for some ε close to zero, assuming that the policyholders are i.i.d.

Recall that $\pi_\phi = (1+\theta_\phi)E[\phi(\mathbf{W})]$.

Gaussian approximation (based on the Central Limit Theorem) : this works if

$$\theta_\phi \geqslant \frac{S^{-1}(\varepsilon)\sigma_\phi}{n^{1/2}E[\phi(\mathbf{W})]},$$

where S is the survival function of a $\mathcal{N}(0,1)$ variable, and σ_ϕ^2 is the variance of $\phi(\mathbf{W})$.

Increasing θ_ϕ will allow to become viable for smaller n, but will reduce the demand, hence will have effect on n.

Institut Louis Bachelier

ENSAE IP PARIS

Alternative vision of the viability

In the paper (not discussed here), we also consider the viability in terms of resilience against "shocks".

The idea is to introduce the possibility of catastrophic events that break the independence between policyholders.

We consider the simple model where the loss of the company is

$$\sum_{i=1}^{n} \phi(\mathbf{W}_i) + \delta Z,$$

where Z is a Generalized Pareto distribution with scale parameter proportional to n, and δ a Bernoulli random variable.

On the other hand, the premium is not modified, and does not take Z into account.

Institut Louis Bachelier

ENSAE IP PARIS

Conditions 1/2

Recall that we consider the particular case where $\phi(\mathbf{W}) = \beta E[Y|\mathbf{W}]$.

θ_Y is the loading factor of the traditional insurance contract, θ_ϕ the loading factor of the parametric one.

Let

$$m_Y(\alpha|\mathbf{w}) = \frac{\log E\left[\exp(\alpha Y)|\mathbf{W} = \mathbf{w}\right]}{\alpha}.$$

Proposition

Let

$$\eta(\alpha) = 1 - \beta + \theta_Y - \left\{\sup_{\mathbf{w}} \frac{m_Y(\alpha|\mathbf{w}) - \phi(\mathbf{w})}{E[Y]}\right\}.$$

If

$$0 < \theta_\phi \leqslant \frac{\eta(\alpha)}{\beta},$$

the policyholder with risk aversion α prefers the parametric product.

Institut Louis Bachelier

ENSAE IP PARIS

Conditions 2/2

Proposition (Viability, i.i.d. case)

A sufficient condition for condition $(\mathcal{C}_\varepsilon)$ to hold is to have, for some α_0,

$$\eta(\alpha_0) \geqslant \frac{\sigma_\phi S^{-1}(\varepsilon)}{N^{1/2} E[Y] \mu([\alpha_0, \alpha_0 + h(\alpha_0, \tau)])^{1/2}},$$

where $h(\alpha_0, \tau)$ is some positive function, and N is the size of the total size of the target population, and if

$$0 < \theta_\phi \leqslant \frac{\eta(\alpha_0)}{\beta},$$

Institut Louis Bachelier

ENSAE IP PARIS

Outline

Institut Louis Bachelier

Hybrid product

The previous conditions naturally leads to a "hybrid" product mixing traditional insurance and parametic coverage.

The idea is to consider the following compensation

$$Z = \phi(\mathbf{W})\mathbf{1}_{\mathbf{W}\in\mathcal{W}} + Y\exp(-\tau)\mathbf{1}_{\mathbf{W}\in\mathcal{W}^c},$$

where

$$\mathcal{W} = \{m_Y(\alpha|\mathbf{w}) - \phi(\mathbf{w}) \leqslant \mathfrak{e}\},$$

for some $\mathfrak{e}$.

Institut Louis Bachelier

ENSAE IP PARIS

Determination of $\mathcal{W}$

For some α_0, use a model to estimate $\mathbf{w} \rightarrow E\left[\exp(\alpha_0 Y)|\mathbf{W} = \mathbf{w}\right]$.

Institut Louis Bachelier

ENSAE

IP PARIS

Determination of $\mathcal{W}$

For some α_0, use a model to estimate $\mathbf{w} \to E\left[\exp(\alpha_0 Y)|\mathbf{W} = \mathbf{w}\right]$.

Second model to estimate

$$\mathbf{w} \to E[Y|\mathbf{W}],$$

where we recall that, in this example $\phi(\mathbf{W}) = \beta E[Y|\mathbf{W}]$.

Institut Louis Bachelier

ENSAE IP PARIS

Determination of $\mathcal{W}$

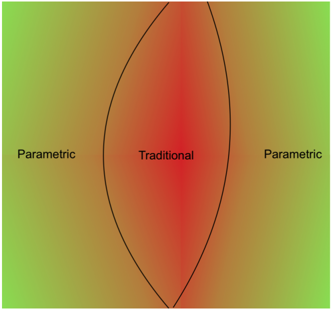
For some α_0, use a model to estimate $\mathbf{w} \to E\left[\exp(\alpha_0 Y)|\mathbf{W} = \mathbf{w}\right]$.

Second model to estimate

$$\mathbf{w} \to E[Y|\mathbf{W}],$$

where we recall that, in this example $\phi(\mathbf{W}) = \beta E[Y|\mathbf{W}]$.

Selection of a region $\mathcal{W}$ that matches the viability conditions.



Determination of $\mathcal{W}$

For some α_0, use a model to estimate $\mathbf{w} \to E\left[\exp(\alpha_0 Y)|\mathbf{W} = \mathbf{w}\right]$.

Second model to estimate

$$\mathbf{w} \to E[Y|\mathbf{W}],$$

where we recall that, in this example $\phi(\mathbf{W}) = \beta E[Y|\mathbf{W}]$.

Selection of a region $\mathcal{W}$ that matches the viability conditions.

These conditions are slightly modified, α prefers the hybrid product to full traditional insurance if

$$\frac{\mathfrak{e}}{E\left[Y|\mathcal{W}\right] p_{\mathfrak{e}}(\alpha)} \leqslant 1 - \beta + \theta_Y - \beta\theta,$$

where $p_{\mathfrak{e}}(\alpha) = \mathbb{P}(\mathbf{W} \in \mathcal{W})$.

Institut Louis Bachelier

ENSAE IP PARIS

Determination of $\mathcal{W}$

For some α_0, use a model to estimate $\mathbf{w} \to E\left[\exp(\alpha_0 Y)|\mathbf{W} = \mathbf{w}\right]$.

Second model to estimate

$$\mathbf{w} \to E[Y|\mathbf{W}],$$

where we recall that, in this example $\phi(\mathbf{W}) = \beta E[Y|\mathbf{W}]$.

Selection of a region $\mathcal{W}$ that matches the viability conditions.

These conditions are slightly modified, α prefers the hybrid product to full traditional insurance if

$$\frac{\mathfrak{e}}{E\left[Y|\mathcal{W}\right] p_{\mathfrak{e}}(\alpha)} \leqslant 1 - \beta + \theta_Y - \beta\theta,$$

where $p_{\mathfrak{e}}(\alpha) = \mathbb{P}(\mathbf{W} \in \mathcal{W})$.

We expect all the policyholders with risk aversion between α_- (the minimum among the policyholder's population) and α_0 to chose the product.

Example of calibration

Illustration in cyber insurance.

LUCY study from AMRAE : describes each year the amount of premium collected on a (significant) part of the French cyber insurance market.

Average premium : $\pi_Y = 9,163$ euros in 2022.

Loading factor : rough estimation at 0.4 based on past data.

$N = 500$ corresponds approximately to the number of companies in the perimeter of the LUCY study in 2022.

Distribution of the risk aversion : shifted exponential,

$$d\mu(t) = \lambda \exp(-\lambda(t - \alpha_-)).$$

α_- corresponds to the minimum risk aversion to accept the loading factor $\theta_Y = 0.4$.

Determination of λ : We make the assumption that half of the population is ready to accept a 40% increase of the premium.

Institut Louis Bachelier

ENSAE IP PARIS

Database of cyber incidents

Synthetic database of cyber incidents.

Among available information : Y = the economic loss, $\mathbf{W}$ = cyber maturity of the victim, time of business interruption,...

Method 1 :

1) Regression tree to estimate $m_Y(\alpha)$.
2) Tree based estimation of ϕ based on **the same subdivision** as for the previous tree.

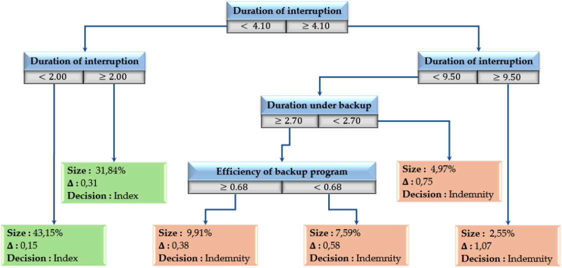
Method 2 :

1) Regression tree to estimate $m_Y(\alpha)$.
2) Independent estimation of ϕ within each leaf of the tree using neural networks.

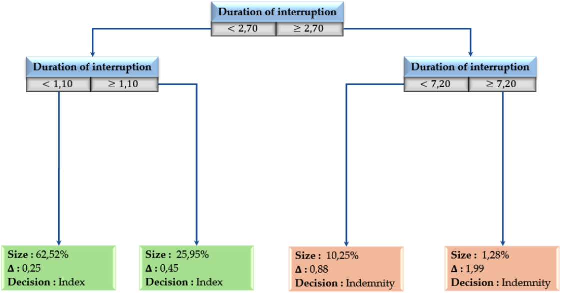
Institut Louis Bachelier

ENSAE IP PARIS

Database of cyber incidents



(a) Backup plan was at least partially successful ($\delta = 1$).



(b) Backup plan failed ($\delta = 0$).

Conclusion

Extension : a relatively similar framework can be used to optimize claim management in traditional insurance :

- after the claim, the idea is to propose to some policyholders (depending on the context of the claim) an immediate compensation ;
- focus the expert effort on some specific claims.

This requires to include a probability for the policyholder to accept or not the automatic compensation (which is expected to be, in average, less than the true loss).

Institut Louis Bachelier

ENSAE IP PARIS

Thank you for your attention!