

Mathématiques appliquées à l'assurance des risques numériques

Sébastien Farkas

Allianz France

//

Doctorat réalisé à :

Sorbonne Université

Ecole Doctorale Sciences Mathématiques de Paris-Centre

Laboratoire de Probabilités, Statistique et Modélisation

Sous la direction d'Olivier Lopez et Caroline Hillairet

23 avril 2025

Le numérique au sein des entreprises et ses risques

Définition de Didier Dubasque (2019) : *"Le numérique représente toutes les applications qui utilisent un langage binaire [...]. Il envisage à la fois les outils, les contenus et les usages."*

Sondage Eurobarometer (2021)

Le sondage réalisé sur 12 863 petites et moyennes entreprises européennes montre :

- la place incontournable du numérique pour les entreprises sondées :
 - 71% ont un site internet,
 - 38% utilisent un stockage en ligne,
 - 30% proposent un système de paiement en ligne à leurs clients,
- le contact fréquent avec une criminalité numérique variée au cours de l'année 2021 :
 - tentative d'hameçonnage (11%),
 - tentative de piratage de leurs comptes bancaires (4%),
 - attaque par un rançongiciel (4%),
 - attaque par déni de service (3%).

Les acteurs malveillants et leurs stratégies

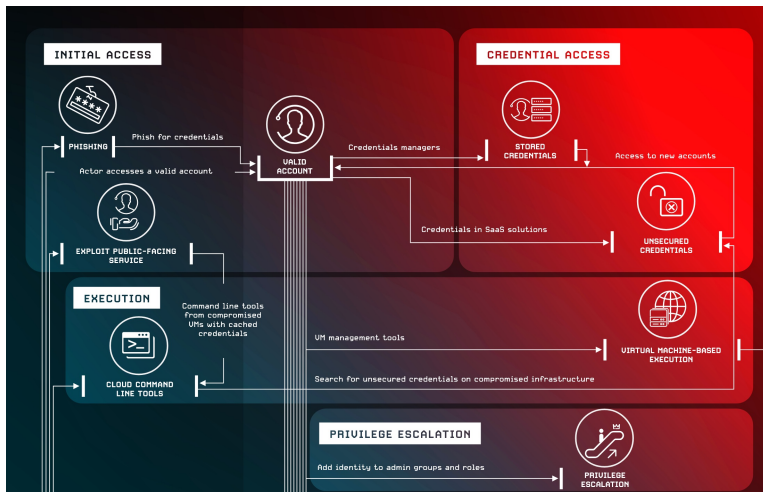
Les équipes de défense numérique font face à trois types d'adversaires aux motivations différentes :

- "Nation-State" : espionnage, sabotage, perturbation, destruction
- "eCrime" : gains financiers, retours sur investissements élevés
- "Hacktivist" : nuire à des cibles opposées à leurs causes

Qui peuvent mettre en place plusieurs types d'intrusions :

- "Zero-click" : intrusion basée sur l'exploitation d'une ou plusieurs vulnérabilités
- "One-click & more" : intrusion nécessitant qu'un utilisateur réalise au moins une action compromettante (clic sur un lien, ouverture d'un fichier, partage d'informations sensibles, réalisation d'un virement,...)

Les schémas d'attaques : de l'intrusion à l'impact



Infographie extraite du Crowdstrike 2025 Global Threat Report, basée sur MITRE ATT&CK

Les schémas d'attaques : de l'intrusion à l'impact



Infographie extraite du Crowdstrike 2025 Global Threat Report, basée sur MITRE ATT&CK

Les solutions de défense proactives et leurs fournisseurs

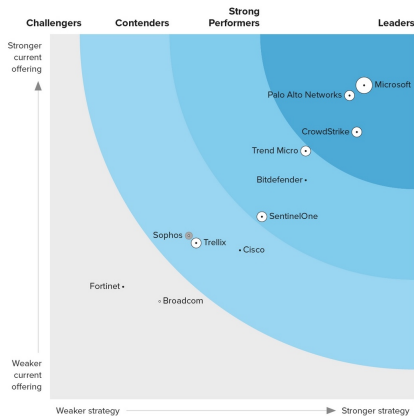
Plusieurs solutions complémentaires visent à prévenir les attaques sur différents maillons.

Mobile Threat Defense Solutions



Forrester Wave

Extended Detection and Response Platforms



Forrester Wave

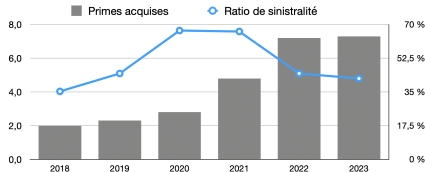
La sécurisation à l'échelle d'une organisation

Les Responsables de la Sécurité des Systèmes d'Information (RSSI) sont avant tout des spécialistes de la gestion des risques à l'échelle d'une organisation. Leurs enjeux peuvent être segmentés de deux manières selon les méthodologies de l'ANSSI :

- par thématique :
 - Gouvernance et interactions entre équipes mobilisées
 - Processus et outillage
 - Communication de crise et relations externes
 - Détection et réponse à incidents
 - Continuité d'activité et reconstruction
- par temporalité :
 - Se préparer à affronter une crise cyber
 - Réagir - Phase 1 : Alerter, mobiliser et endiguer
 - Réagir - Phase 2 : Maintenir la confiance et comprendre l'attaque
 - Réagir - Phase 3 : Relancer les activités métiers et durcir les systèmes d'information
 - Réagir - Phase 4 : Tirer les leçons de la crise et capitaliser

L'assurance des risques numériques et ses enjeux mathématiques

- Aux États-Unis, le marché est rentable, en croissance et très concurrentiel (NAIC) malgré une volatilité des coûts : médiane à 140K\$ et moyenne à 630K\$ (IAIS).



- La dépendance des primes en fonction des profils de risques repose davantage sur des avis d'experts que sur une modélisation statistique des pertes conditionnelles.
- Dans un contexte de bases de données de sinistres en cours de constitution, comment peut-on contribuer à améliorer la maîtrise des engagements des assureurs ?

Table des matières

- 1 Classification des risques à partir de valeurs extrêmes
- 2 Estimation de copules à partir de données censurées
- 3 Estimation de phénomènes d'accumulation auto-excités

Théorie des valeurs extrêmes : paramètre de forme γ

La théorie des valeurs extrêmes repose sur deux résultats fondamentaux :

- le théorème de Fisher–Tippett–Gnedenko, cf Théorème 1.1.1,
- le théorème de Pickands–Balkema–De Haan, cf Théorème 1.1.2.

Proposition 1.1.4 cf Section 5.3.1 de Haan et Ferreira (2006)

Si Y une variable aléatoire appartient au domaine d'attraction de Fréchet, alors ses moments d'ordres strictement inférieurs à $1/\gamma$ sont finis et ceux d'ordres supérieurs à $1/\gamma$ sont infinis,

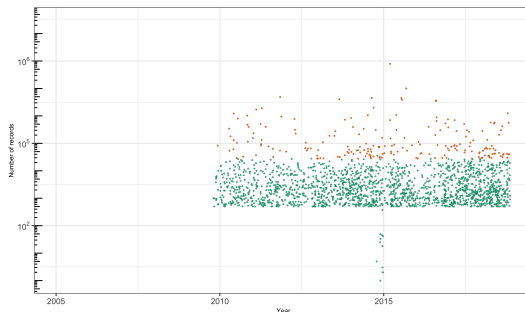
$$\forall 0 < \alpha < 1/\gamma, \mathbb{E}[|Y|^\alpha] < +\infty \text{ et } \forall \alpha > 1/\gamma, \mathbb{E}[|Y|^\alpha] = +\infty.$$

En particulier :

- si $\gamma > \frac{1}{2}$, $\mathbb{V}[Y] = +\infty$
- si $\gamma > 1$, $\mathbb{E}[Y] = +\infty$

Théorie des valeurs extrêmes : Distribution Pareto Généralisée (GPD)

- Choix empirique du seuil u pour déterminer les excès, représentés en rouge :



- Estimation des paramètres de la loi GPD par maximum de pseudo-vraisemblance :
 - σ : le paramètre d'échelle, strictement positif,
 - γ : le paramètre de forme, strictement positif dans le cadre considéré du domaine de Fréchet.

Théorie des valeurs extrêmes : analyse conditionnelle des excès

Objectif : estimation des paramètres conditionnels de la loi GPD $\theta_0(x) = (\sigma_0(x), \gamma_0(x))$.

Hypothèse 1.1.5

Il existe une fonction γ_0 strictement positive sur $\mathcal{X}$ vérifiant :

$$\forall x \in \mathcal{X}, \forall y > 0, \lim_{t \rightarrow \infty} \frac{\bar{F}(ty | x)}{\bar{F}(t | x)} = y^{-1/\gamma_0(x)}.$$

Corollaire 1.1.6

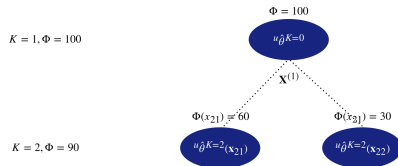
Si l'hypothèse 1.1.5 est vérifiée, alors il existe des fonctions $\sigma_0(x)$ et $\gamma_0(x)$, strictement positives sur $\mathcal{X}$ telles que :

$$\lim_{u \rightarrow \infty} \sup_{x \in \mathcal{X}} \sup_{z > 0} |\bar{F}_u(z | X = x) - \bar{H}_{(\sigma_0(x), \gamma_0(x))}(z)| = 0.$$

Méthodes existantes : Beirlant & Goegebeur (2003), Beirlant & Goegebeur (2004) et Chavez-Demoulin, Embrechts & Hofert (2015).

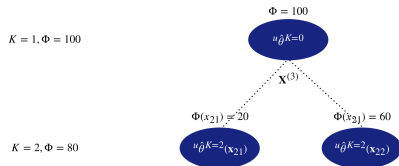
Classification : introduction à l'algorithme CART

$$\hat{m} = \arg \min_{m \in \mathcal{M}} \Phi(Y, m(X)) \text{ avec } \Phi(Y, m(X)) = \frac{1}{n} \sum_{i=1}^n \phi(Y_i, m(X_i))$$



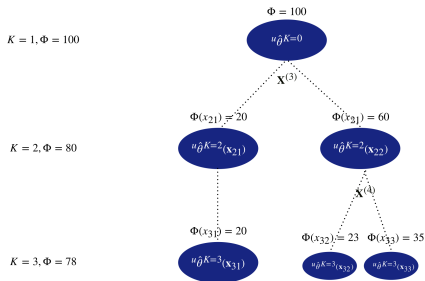
Classification : introduction à l'algorithme CART

$$\hat{m} = \arg \min_{m \in \mathcal{M}} \Phi(Y, m(X)) \text{ avec } \Phi(Y, m(X)) = \frac{1}{n} \sum_{i=1}^n \phi(Y_i, m(X_i))$$



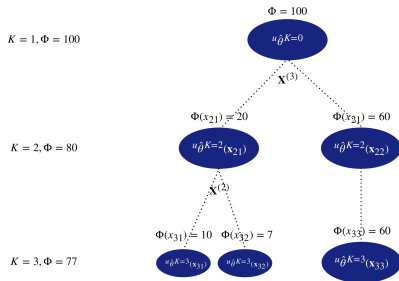
Classification : introduction à l'algorithme CART

$$\hat{m} = \arg \min_{m \in \mathcal{M}} \Phi(Y, m(X)) \text{ avec } \Phi(Y, m(X)) = \frac{1}{n} \sum_{i=1}^n \phi(Y_i, m(X_i))$$



Classification : introduction à l'algorithme CART

$$\hat{m} = \arg \min_{m \in \mathcal{M}} \Phi(Y, m(X)) \text{ avec } \Phi(Y, m(X)) = \frac{1}{n} \sum_{i=1}^n \phi(Y_i, m(X_i))$$



Classification : introduction à l'algorithme CART

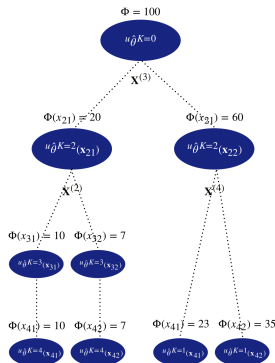
$$\hat{m} = \arg \min_{m \in \mathcal{M}} \Phi(Y, m(X)) \text{ avec } \Phi(Y, m(X)) = \frac{1}{n} \sum_{i=1}^n \phi(Y_i, m(X_i))$$

$K = 1, \Phi = 100$

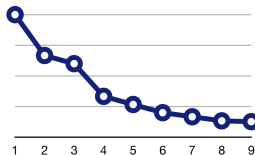
$K = 2, \Phi = 80$

$K = 3, \Phi = 77$

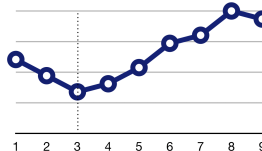
$K = 4, \Phi = 75$



Échantillon d'entraînement

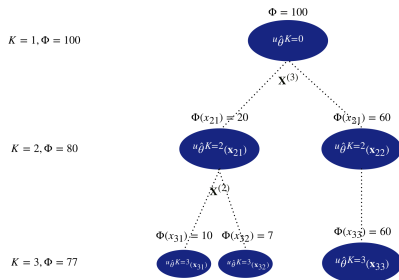


Validation croisée

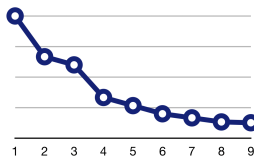


Classification : introduction à l'algorithme CART

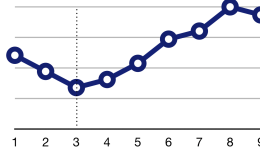
$$\hat{m} = \arg \min_{m \in \mathcal{M}} \Phi(Y, m(X)) + \lambda K \text{ avec } \Phi(Y, m(X)) = \frac{1}{n} \sum_{i=1}^n \phi(Y_i, m(X_i))$$



Échantillon d'entraînement

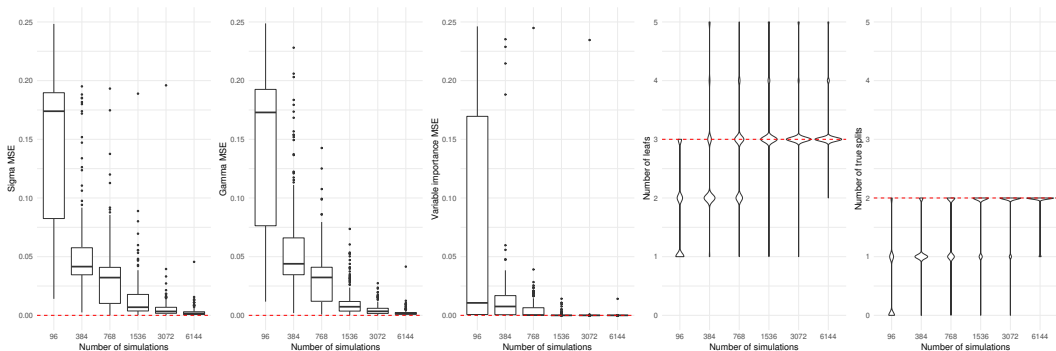


Validation croisée



CART & GPD : résultats de la simulation numérique

Soit $(\sigma_1, \sigma_2, \sigma_3) = (\gamma_1, \gamma_2, \gamma_3) = (0.5, 1, 1.5)$.



CART & GPD : résultats

Théorème 1.3.1

Sous les conditions précisées au théorème 1 de la partie B, nous obtenons l'existence de constantes positives C_1 , C_2 et C_3 telles que $\forall t > K(\log k_n)k_n^{-1}$,

$$\begin{aligned} & \mathbb{P} \left(\sup_{u_{\min} \leq u \leq u_{\max}} |{}^u \hat{\theta}^K(x) - {}^u \theta^{*K}(x)|_2^2 \geq t \right) \\ & \leq 2 \left(\exp \left(-\frac{C_1 k_n t}{K(\log k_n)^2} \right) + \exp \left(-\frac{C_2 k_n t^{1/2}}{K^{1/2} \log k_n} \right) \right) + \frac{C_3 K}{k_n^{5/2} t^{3/2}}, \end{aligned}$$

Nous majorons ainsi l'espérance de l'erreur stochastique pour une constante positive C_4 :

$$\mathbb{E} \left[\sup_{u_{\min} \leq u \leq u_{\max}} |{}^u \hat{\theta}^K(x) - {}^u \theta^{*K}(x)|_2^2 \right]^{1/2} \leq C_4^{1/2} \frac{K^{1/2} \log k_n}{k_n^{1/2}}.$$

CART & GPD : résultats

Théorème 1.3.2

Parmi les conditions précisées au théorème 3 de l'annexe B, nous supposons en particulier que la constante de pénalisation du nombre de feuilles λ , vérifie l'inégalité suivante :

$$0 < c_2(\log k_n)^{1/2} k_n^{-1/2} \leq \lambda \leq \mathfrak{D} - 2c_2(n)^{1/2} k_n^{-1/2},$$

avec $c_2 > 0$ et $\mathfrak{D} = \inf_u \inf_{K < K^*(u)} \Delta L(u\theta^{*K}(x), u\hat{\theta}^{K^*}(x))$.

Sous l'ensemble des conditions, nous obtenons pour les constantes positives $\mathcal{C}_1$, $\mathcal{C}_2$ et $\mathcal{C}_3$ du Théorème 1.3.1 les majorations suivantes :

$$\begin{aligned} \forall K > K^*(u), \mathbb{P}(\hat{K}(u) = K) &\leq 2 \exp \left(-\frac{C_1 k_n \lambda^2 (K - K^*(u))^2}{\log k_n^2} \right) \\ &+ 2 \exp \left(-\frac{C_2 k_n \lambda (K - K^*(u))}{\log k_n} \right) \\ &+ \frac{C_3}{k_n^{5/2} \lambda^3 (K - K^*(u))^3}, \end{aligned}$$

CART & GPD : résultats

Théorème 1.3.2

$$\begin{aligned} \forall K < K^*(u), \mathbb{P}(\hat{K}(u) = K) \leq & 4 \exp \left(- \frac{C_1 k_n \{\mathfrak{D} - \lambda(K^*(u) - K)\}^2}{(\log k_n)^2} \right) \\ & + 4 \exp \left(- \frac{C_2 k_n \{\mathfrak{D} - \lambda(K^*(u) - K)\}}{\log k_n} \right) \\ & + \frac{2C_3}{k_n^{5/2} \{\mathfrak{D} - \lambda(K^*(u) - K)\}^3}, \end{aligned}$$

Nous majorons ainsi l'espérance de l'erreur stochastique pour une constante positive C_5 dépendante de ${}^u\hat{\theta}^{K^*}(x)$.

$$\mathbb{E} \left[|{}^u\hat{\theta}^{\hat{K}}(x) - {}^u\hat{\theta}^{K^*}(x)|_2^2 \right]^{1/2} \leq \frac{C_5^{1/2} K^*(u)^{1/2} \log k_n}{k_n^{1/2}}.$$

Table des matières

- 1 Classification des risques à partir de valeurs extrêmes
- 2 Estimation de copules à partir de données censurées
- 3 Estimation de phénomènes d'accumulation auto-excités

Censure : estimateur de Kaplan-Meier (1958)

- T : le délai de traitement d'un sinistre, inconnu pour les dossiers ouverts,
- C : la censure de T ,
- Y : le délai observé de traitement d'un sinistre $Y = \min(T, C)$,
- δ : l'indicatrice de clôture d'un sinistre, $\delta = \mathbb{1}_{\{Y=T\}}$,

$$\hat{S}^T(y) = \prod_{Y_i \leq y} \left(1 - \frac{\delta_i}{\sum_{k=1}^n \mathbb{1}_{Y_k \geq Y_i}} \right).$$

Représentation en somme de variables pondérées : Van der Laan et Robins (2003)

$$\hat{S}^T(y) = \sum_{i=1}^n \hat{W}_{in} \mathbb{1}_{Y_i \leq y} \text{ avec } \hat{W}_{in} = \frac{\delta_i}{n \hat{S}^C(Y_i^-)}.$$

Censure : résultats sur l'estimateur de Kaplan-Meier

Théorème 2.1.5 (Théorème de Akritas 2000)

Supposons T et C indépendantes, de même support et telles que $\mathbb{P}(T = C) = 0$. Pour toute fonction ϕ vérifiant la condition suivante :

$$\mathbb{E} \left[\frac{\phi(T)^2}{\hat{S}^C(T)} \right] < \infty,$$

nous obtenons une décomposition en sommes i.i.d de l'expression pondérée :

$$\sum_{i=1}^n \hat{W}_{i,n} \phi(Y_i) = \sum_{i=1}^n W_{in}^* \phi(Y_i) + \frac{1}{n} \sum_{i=1}^n \xi_i(\phi) + o_P(n^{-1/2}), \text{ avec } \mathbb{E}[\xi_i(\phi)] = 0.$$

Ce résultat permet notamment d'obtenir des résultats de type central limite, cf Stute 1995.

Dépendance : introduction au théorème de Sklar (1959)

Principe

Le théorème de Sklar montre que la loi jointe F d'un vecteur aléatoire L de dimension d se caractérise par deux composantes indépendantes :

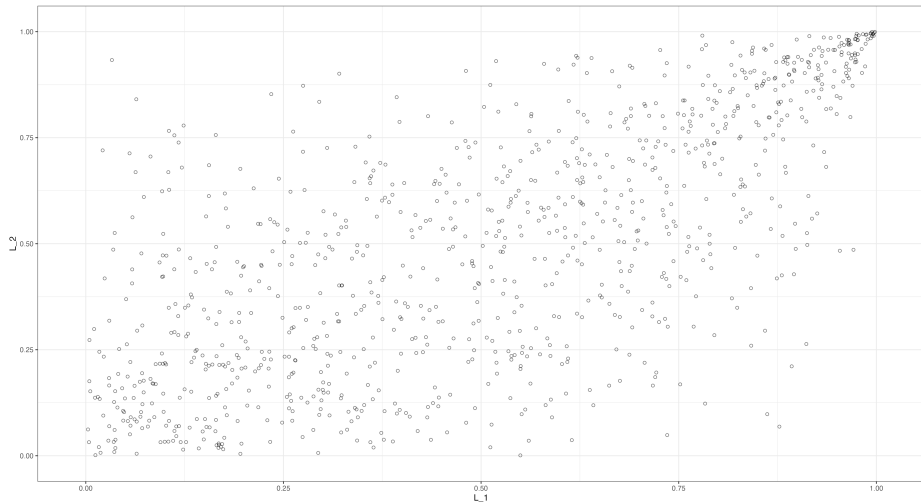
- ses d lois marginales,
 - sa structure de dépendance entre les marginales, appelée copule.
-
- La copule est unique lorsque les lois des marginales sont absolument continues. Dans ce cas, la copule s'exprime de la manière suivante :

$$\mathfrak{C}(u_1, \dots, u_d) = F \left(F_1^{(-1)}(u), \dots, F_d^{(-1)}(u) \right).$$

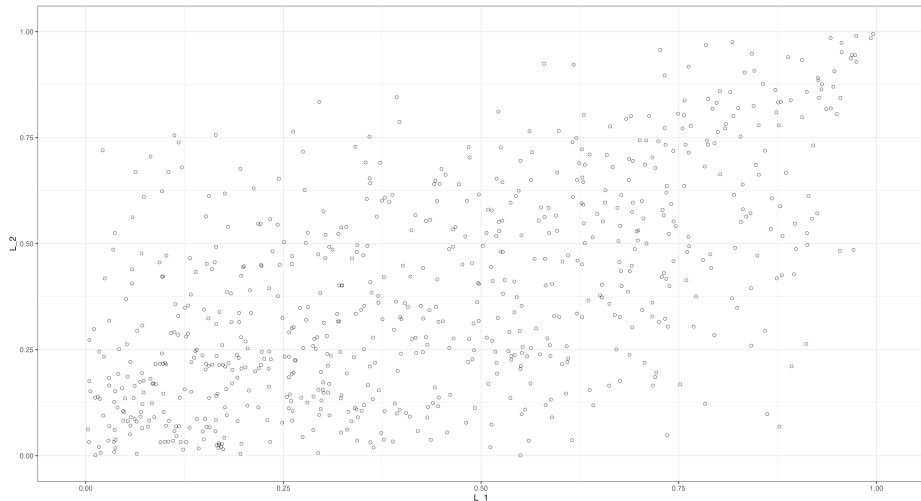
- Empiriquement, la copule s'estime sur des pseudo-observations $\hat{U}_i \in [0; 1]^d$:

$$\forall j \in \{1, \dots, d\}, \forall i \in \{1, \dots, n\}, \hat{U}_i^{(j)} = \hat{F}^{(j)}(L_i^{(j)}), \text{ avec } \hat{F}^{(j)}(l) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{L_i^{(j)} \leq l}.$$

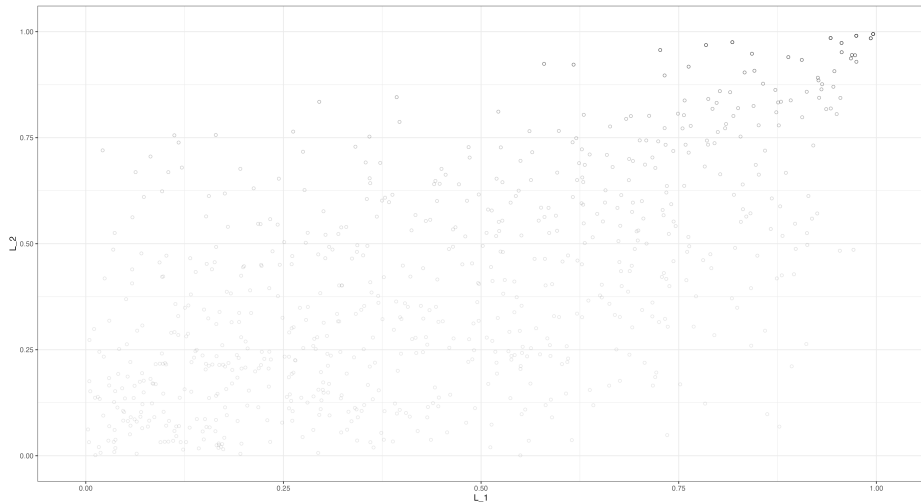
Dépendance & Censure : introduction



Dépendance & Censure : introduction



Dépendance & Censure : introduction



Dépendance & Censure : cadre et objectifs

- C est supposée indépendante de (T, L, X) et avec un support identique à celui de T ,
- La structure de dépendances des coûts marginaux est supposée paramétrique, appartenant à une famille $\mathcal{C} = \{\mathfrak{C}_\theta : \theta \in \Theta\}$, avec $\Theta \subset \mathbb{R}^m$ et nous considérons deux cas de figure :
 - soit une dépendance simple, l'objectif est alors θ_0 et l'erreur d'estimation est égale $\hat{\theta} - \theta_0$,
 - soit une dépendance conditionnelle, l'objectif est alors $\theta_0(x)$ et on distingue deux erreurs :
 - l'erreur d'estimation entre $\hat{\theta}(x)$ et $\theta_h^*(x)$,
 - l'erreur d'approximation entre $\theta_h^*(x)$ et $\theta_0(x)$.

$N(\theta)$	$=$	$E [\log \mathfrak{c}_\theta(U)]$	θ_0	$=$	$\arg \max_\theta N(\theta)$
$N_n^*(\theta)$	$=$	$\sum_{i=1}^n W_{i,n}^* \log \mathfrak{c}_\theta(U_i^*)$	θ^*	$=$	$\arg \max_\theta N_n^*(\theta)$
$\hat{N}_n(\theta)$	$=$	$\sum_{i=1}^n \hat{W}_{i,n} \log \mathfrak{c}_\theta(\hat{U}_i)$	$\hat{\theta}$	$=$	$\arg \max_\theta \hat{N}_n(\theta)$
$M(\theta, x)$	$=$	$E [\log \mathfrak{c}_\theta(U) X = x] f_X(x)$	$\theta(x)$	$=$	$\arg \max_\theta M(\theta, x)$
$M_n^*(\theta, x)$	$=$	$\frac{1}{h^{d+1}} \sum_{i=1}^n W_{i,n}^* K\left(\frac{x_i - x}{h}\right) \log \mathfrak{c}_{(\theta, x)}(U_i^*)$	$\theta_h^*(x)$	$=$	$\arg \max_\theta M_n^*(\theta, x)$
$\hat{M}_n(\theta, x)$	$=$	$\frac{1}{h^{d+1}} \sum_{i=1}^n \hat{W}_{i,n} K\left(\frac{x_i - x}{h}\right) \log \mathfrak{c}_{(\theta, x)}(\hat{U}_i)$	$\hat{\theta}(x)$	$=$	$\arg \max_\theta \hat{M}_n^*(\theta, x)$

Dépendance & Censure : résultat pour le modèle simplifié

Théorème 2.3.1

Pour le modèle simplifié et dans les conditions précisées dans notre contribution, y compris la définition de certains ξ_i^ϕ , nous obtenons la décomposition i.i.d suivante de l'erreur globale :

$$(\hat{\theta} - \theta_0) = -\Sigma^{-1} \left[\sum_{i=1}^n W_{in}^* \phi(U_i, Y_i, X_i) + \frac{1}{n} \sum_{i=1}^n \xi_i^\phi + \sum_{k=1}^d \Lambda_i^{(k)} \right] + o_P(n^{-1/2}), \text{ avec}$$

$$\begin{cases} \Sigma &= E \left[\nabla_{\theta}^2 \log c_{\theta_0}(U) \right] \text{ supposée inversible,} \\ \mathbb{E}[\xi_i^\phi] &= 0, \\ \mathbb{E}[\Lambda_i^{(k)}] &= 0. \end{cases}$$

Ainsi, pour une certaine matrice de covariance V , nous obtenons la convergence en loi :

$$\sqrt{n} \left(\hat{\theta} - \theta_0 \right) \Rightarrow \mathcal{N} \left(0, \Sigma^{-1} V \Sigma \right),$$

Dépendance & Censure : résultat pour le modèle semi-paramétrique

Théorème 2.3.2

Pour le modèle semi-paramétrique, avec K pour noyau, et dans les conditions précisées dans notre contribution, nous obtenons la décomposition i.i.d suivante de l'erreur stochastique :

$$\hat{\theta}(x) - \theta_h^*(x) = -\Sigma(x)^{-1} \left\{ \frac{1}{h^{d+1}} \sum_{i=1}^n W_{i,n}^* K\left(\frac{x_i - x}{h}\right) \Phi(U_i) \right\} + o_P(n^{-1/2} h^{(d+1)/2}),$$

avec $\Sigma(x) = \left(E \left[\Phi_{\theta(x)}^{(j,k)}(U) | x = x \right] \right)_{j,k}$ supposée inversible.

Ainsi, en notant $S(x) = E[\Phi(U) | x = x]$ et $S(x)'$ sa transposée, nous obtenons la convergence en loi :

$$n^{1/2} h^{(d+1)/2} \left\{ \hat{\theta}(x) - \theta_h^*(x) \right\} \Rightarrow \mathcal{N} \left(0, \Sigma(x)^{-1} S(x) S(x)' \Sigma(x)^{-1} \right).$$

Table des matières

- 1 Classification des risques à partir de valeurs extrêmes
- 2 Estimation de copules à partir de données censurées
- 3 Estimation de phénomènes d'accumulation auto-excités

Processus de comptage : introduction

Un processus ponctuel N défini sur $\mathbb{R}^+$ est un processus de comptage si presque sûrement ses trajectoires sont croissantes par saut d'amplitude 1, continues à droite et nulles à l'instant 0.

Autrement dit, le processus de comptage N saute unitairement aux temps de réalisations, disjoints, d'événements d'intérêts.

Pour un horizon $T > 0$, l'observation de N sur $[0; T]$ se caractérise ainsi par ses temps de sauts $(\tau_i)_{i \geq 1}$ inférieurs à T et le temps T correspondant à la fin de l'observation : $((\tau_i)_{i \geq 1}, T)$.

$$\forall t \leq T, N(t) = \sum_{i \geq 1} \mathbb{1}_{\tau_i \leq t}.$$

Processus de comptage : compensateur, intensité et estimation

Compensateur

Soit N un processus ponctuel. On appelle compensateur de N un processus croissant, continu à droite et prévisible Λ tel que $N - \Lambda$ soit une martingale locale.

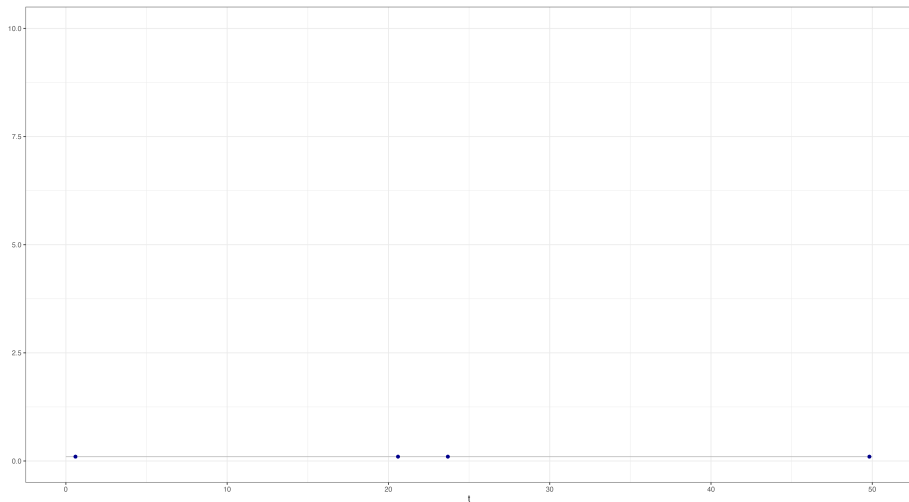
Intensité

Soit N un processus ponctuel et Λ son compensateur. Si presque sûrement toute trajectoire de Λ est absolument continue, alors on note $\lambda(t) = \frac{d\Lambda(t)}{dt}$ l'intensité de N .

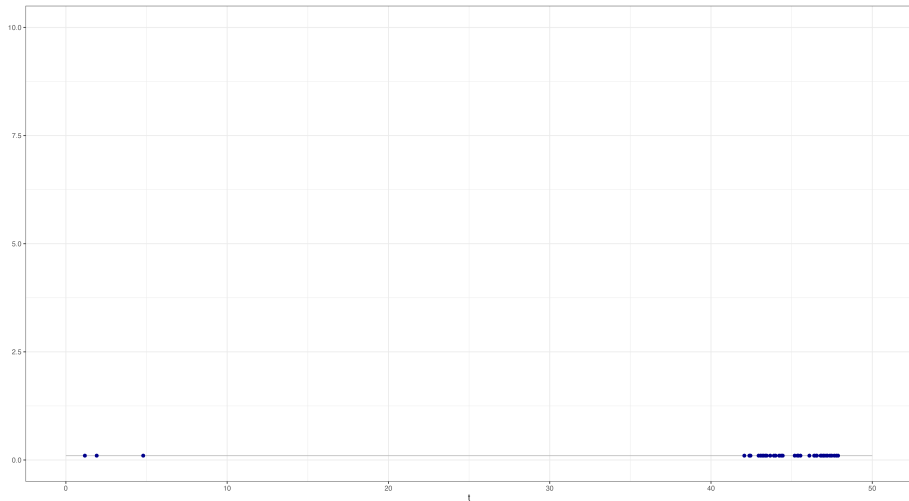
Vraisemblance d'observer dans l'intervalle $[0; T]$ une suite d'événements aux temps (τ_i) :

$$\mathcal{L}\left(\left((\tau_i)_{i \geq 1}, T\right), N\right) = \prod_{i=1}^{N(T)} \left(\lim_{t \nearrow \tau_i} \lambda(t) \right) e^{-\Lambda(T)}.$$

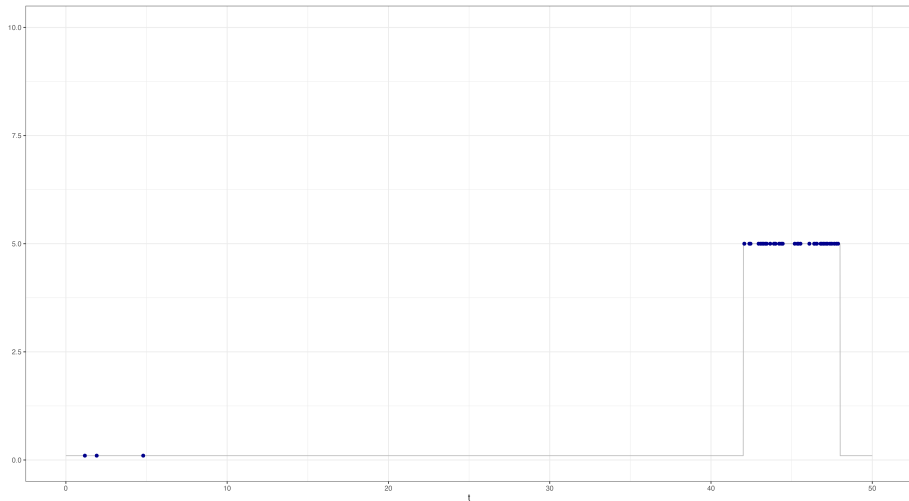
Processus de comptage : processus de Poisson homogène



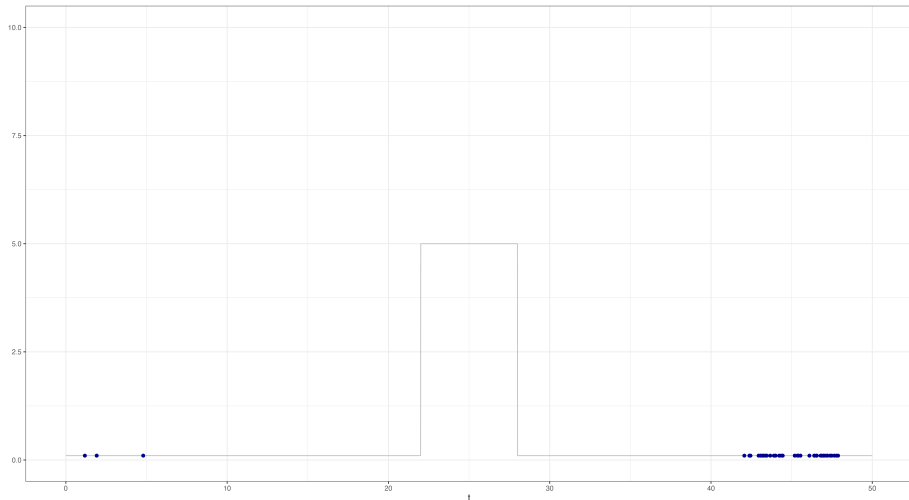
Processus de comptage : processus de Poisson homogène



Processus de comptage : processus de Poisson inhomogène



Processus de comptage : processus de Poisson inhomogène



Processus de comptage : processus de Hawkes (1971)

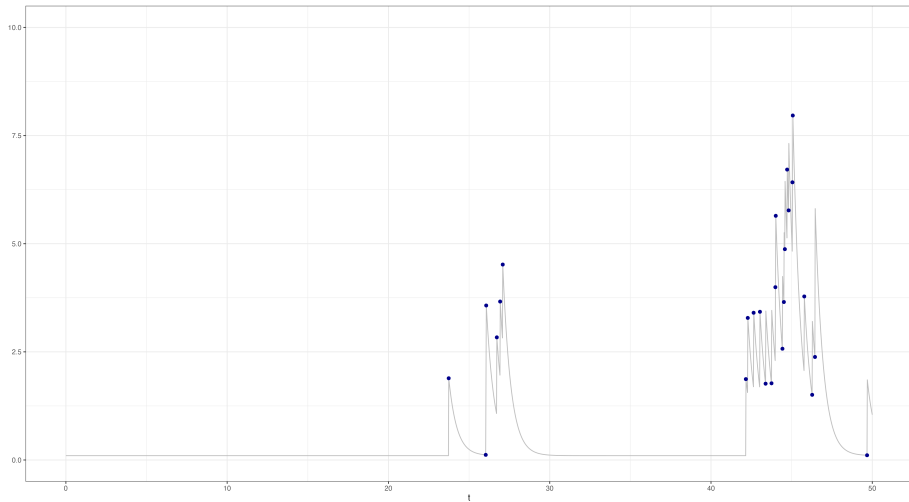
Définition

$$\lambda^{(\mu, h)}(t) = \mu + \int_0^t h(t-s) dN_s^{(\mu, h)}, \text{ avec}$$

- $\mu > 0$ une constante, assimilable à l'intensité d'un processus de Poisson homogène,
- h une fonction dite noyau, de $\mathbb{R}^+$ dans $\mathbb{R}^+$ traduisant le phénomène d'auto excitation,
- vérifiant $\|h\|_1 = \int_{\mathbb{R}^+} h(t) dt < 1$ pour être dans un cadre stable cf Bacry, Delattre et al. (2013)

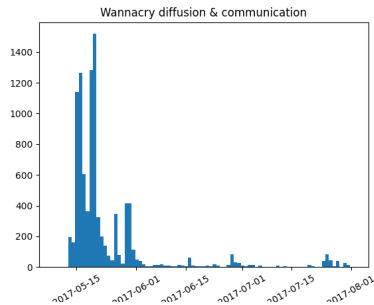
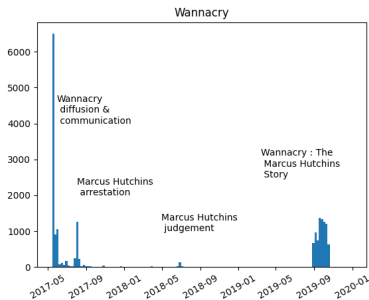
Nombreuses applications : Bacry, Mastromatteo & al. (2015) et Bonnet, Dion-Blanc & al. (2021).

Processus de comptage : processus de Hawkes (1971)



Processus de comptage : processus de Hawkes (1971)

Exemple de la diffusion médiatique relative au rançongiciel Wannacry :



Accumulation & Hawkes : cadre de l'étude

Pour étudier la propagation d'impacts d'un logiciel malveillant et les besoins en assistance associés, nous supposons que la durée de rétablissement est constante, égale à δ . Nous introduisons plusieurs processus ponctuels :

- le processus de Hawkes correspondant au nombre de victimes impactées, noté $\mathcal{I}(t)$,
- le processus de comptage du nombre de victimes rétablies, noté $\mathcal{R}(t)$,
- le processus ponctuel du nombre de victimes nécessitant une assistance, noté $\mathcal{A}(t)$ et vérifiant $\mathcal{A}(t) = \mathcal{I}(t) - \mathcal{R}(t)$,

$$\left\{ \begin{array}{lcl} \mathcal{I}(t) & = & \sum_{\tau_i} \mathbb{1}_{\{\tau_i \leq t\}} \\ \mathcal{R}(t) & = & \sum_{\tau_i} \mathbb{1}_{\{\tau_i \leq t - \delta\}} \\ \mathcal{A}(t) & = & \sum_{\tau_i} \mathbb{1}_{\{t - \delta < \tau_i \leq t\}} \end{array} \right.$$

Accumulation & Hawkes : cadre de démonstration

- Décomposition du processus d'accumulation $\mathfrak{A}(t)$ entre son espérance et ses déviations, cf Jaisson and Rosenbaum (2015),

$$\mathfrak{A}(t) = \mathbb{E}[\mathfrak{A}(t)] + \mathfrak{M}(t),$$

- Étude de la norme infinie de $\mathfrak{M}(t)$ par disjonction de cas relative au compensateur du processus $\mathfrak{J}(t)$, noté $\Lambda(t)$, cf Guével (2021) et Reynaud-Bouret & Roy (2007),

$$\begin{aligned} \mathbb{P}(\sup_{0 \leq t \leq T} \mathfrak{M}(t) > x) &= \mathbb{P}(\{\sup_{0 \leq t \leq T} \mathfrak{M}(t) > x\} \cap \{\Lambda(T) < c_u^T T\}) \\ &+ \mathbb{P}(\{\sup_{0 \leq t \leq T} \mathfrak{M}(t) > x\} \cap \{\Lambda(T) \geq c_u^T T\}) \\ &\leq \mathbb{P}(\{\sup_{0 \leq t \leq T} \mathfrak{M}(t) > x\} \mid \{\Lambda(T) < c_u^T T\}) \mathbb{P}(\Lambda(T) < c_u^T T) \\ &+ \mathbb{P}(\Lambda(T) \geq c_u^T T) \\ &\leq \mathbb{P}(\{\sup_{0 \leq t \leq T} \mathfrak{M}(t) > x\} \mid \{\Lambda(T) < c_u^T T\}) \\ &+ \mathbb{P}(\Lambda(T) \geq c_u^T T). \end{aligned}$$

- Majoration de la probabilité de saturation des capacités d'assistance $\mathbb{P}(\sup_{0 \leq t \leq T} \mathfrak{A}(t) > x)$.

Accumulation & Hawkes : résultat

Hypothèse 3.3.1 : $\exists \Delta < +\infty$ tel que pour tout $t > \Delta$, $h(t) = 0$.

Hypothèse 3.3.2 : $\|h\|_\infty := \sup_{s \in [0; \Delta]} h(s) < +\infty$.

Théorème 3.3.3

Soit $N^{(\mu, h)}$ un processus de Hawkes linéaire auto-excitant, $\Lambda_{N^{(\mu, h)}}$ son compensateur et u un réel strictement positif. Si $\|h\|_1 < 1$ et si les hypothèses 3.3.1 et 3.3.2 sont vérifiées, alors il existe un $t_u > 0$, dépendant de $(u, \Delta, \|h\|_1)$, tel que pour tout $T \geq t_u$,

$$\mathbb{P} \left(\sup_{0 \leq t \leq T} \mathfrak{A}(t) > x \right) \leq \exp \left(-\mathfrak{c}_u^T T \mathfrak{H} I \left(\frac{x - \mu \delta (1 + \|H\|_1)}{\mathfrak{c}_u^T T \mathfrak{H}} \right) \right) + \left(3 + \frac{\mu e}{u + \log(T)} \right) e^{-u},$$

- avec $\|H\|_1 = \frac{\|h\|_1}{1 - \|h\|_1}$,
- $\mathfrak{H} = 1 + \|H\|_1$ de manière générale ou $\mathfrak{H} = 1 + \int_0^\delta H(u) du$ lorsque H est décroissante,
- $\mathfrak{c}_u^T = \mu + \|H\|_1 + f_T^{(\mu, h)}(u)$, pour une certaine fonction f précisée au théorème D.3.7.

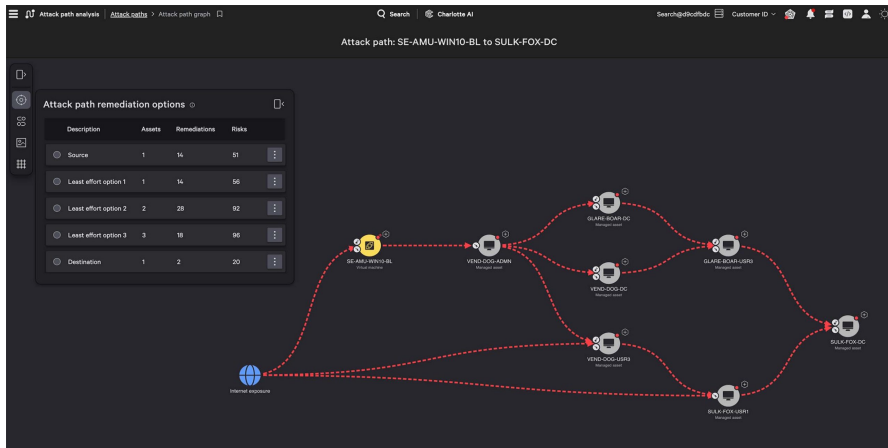
Conclusion théorique

Synthèse des contributions aux mathématiques appliquées à l'assurance des risques numériques :

- au premier chapitre, nous avons adapté l'algorithme d'arbre de régression au contexte de la classification de risques à partir de valeurs extrêmes,
- au second chapitre, nous avons ajusté l'estimation de copules dans un contexte de données censurées à droite,
- au troisième chapitre, nous avons proposé une majoration de l'accumulation d'événements issus d'un processus de Hawkes,

Au-delà des enjeux mathématiques, les pouvoirs publics, les organismes d'assurances et les acteurs de la sécurité numérique sont clés pour gérer les risques numériques, cf Hassler (2019).

Ouverture sur l'importance de la multidisciplinarité



Capture d'écran d'une des solutions de CrowdStrike

A votre disposition pour répondre à vos questions

