

**Mémoire présenté devant l'Institut de Science Financière
et d'Assurances
pour l'obtention du diplôme du Master Actuariat
et l'admission à l'Institut des Actuaires**

Par : Jean-François Le Louarne

Titre : Détection robuste des changements d'intensité

Confidentialité : Non Oui (Durée : 1 an 2 ans)

Les signataires s'engagent à respecter la confidentialité ci-dessus

*Membres présents du jury de l'Institut
des Actuaires :*
Catherine Pigeon
Eric Ralaimiadana

Entreprise :
Nom : AFI-ESCA
Signature :

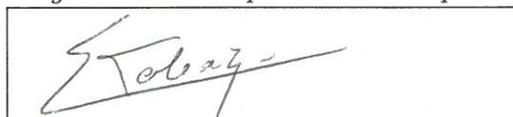
*Membres présents du Jury du Master
Actuariat de l'ISFA :*
Stéphane Loisel

Directeur de Mémoire en entreprise :
Nom : Elie Toledano
Signature ;



*Autorisation de publication et de mise en ligne sur un site de diffusion de
documents actuariels (après expiration de l'éventuel délai de confidentialité)*

Signature du responsable entreprise



Signature du candidat



Détection robuste des changements d'intensité

Application à l'assurance emprunteur

Jean-François Le Louarne

16 février 2023

Table des matières

Remerciements	1
Note de Synthèse	2
0.1 Introduction	2
0.2 Détection des changements d'intensité dans les Processus de Poisson	2
0.2.1 Problématique	2
0.2.2 Critère pour définir la meilleure approche	3
0.2.3 Processus CUSUM pour les processus de Poisson	3
0.3 Considérations pratiques	4
0.4 Résultats	7
0.5 Conclusions	9
Synthesis	11
0.6 Introduction	11
0.7 Detection of intensity changes in Poisson processes	11
0.7.1 Problem	11
0.7.2 Criterion for determining the best approach.	11
0.7.3 CUSUM process for Poisson processes	12
0.8 Practical considerations	13
0.9 Results	16
0.10 Conclusions	18
Introduction	20
1 Contexte de l'assurance emprunteur	22
1.1 Définition de l'assurance emprunteur	22
1.2 Types de contrats	22
1.3 Garanties les plus courantes	24
1.4 Franchises, carences, exclusions, plafonds, limites d'âge	24
1.5 Options	26
1.6 Catégories ministérielles et participation aux bénéfiques	26
1.7 Le marché	26
1.8 Evolutions	27
1.9 La guerre des prix	29
2 Changement d'intensité dans un processus de Poisson	32
2.1 Processus de Poisson	32
2.2 Poisson disorder Problem	34
2.2.1 Quel est le problème?	34
2.2.2 Formulation mathématique	35

2.2.3	Critère de Lorden	35
3	Algorithmes de détections des ruptures	37
3.1	Détecteurs de dépassement de limites et cartes de contrôle de Shewhart . .	38
3.2	Cartes de contrôle à moyenne mobile géométrique	40
3.3	Cartes de contrôle à moyenne mobile finie	41
3.4	Algorithmes à dérivée filtrée	42
3.5	Algorithme CUSUM	43
3.5.1	Principe élémentaire	43
3.5.2	L'algorithme CUSUM comme répétition du test séquentiel du rapport de probabilité	44
3.5.3	CUSUM pour un processus de Poisson	45
3.5.4	Optimalité du processus CUSUM	47
4	Mise en œuvre pratique	49
4.1	L'intensité avant le changement	49
4.1.1	Régression de Poisson	49
4.1.2	D'autres modèles	52
4.2	Détermination de ρ ampleur du changement à détecter	52
4.3	Choix du seuil d'alerte	52
4.3.1	La magie du changement de temps	52
4.3.2	Par calcul direct	53
4.3.3	Par simulations	60
4.3.4	Comparaison des deux méthodes	60
4.3.5	Obtenir un seuil d'alerte	60
5	Applications à la surveillance de la mortalité	63
5.1	Intensité de référence	63
5.1.1	Données	63
5.1.2	Modélisation	66
5.2	Détection des changements d'intensité	68
5.2.1	Mise en Œuvre	68
5.2.2	Résultats	69
6	Applications à la garantie ITT	85
6.1	Intensité de référence	86
6.1.1	Données	86
6.1.2	Modélisation	87
6.2	Détection des changements d'intensité	89
6.2.1	Mise en Œuvre	89
6.2.2	Résultats	89
7	Chasse-trapes	95
7.1	Biais cognitifs	95
7.1.1	La courbe qui monte (ou qui descend)	95
7.1.2	La courbe qui stagne	95
7.1.3	La courbe qui reste au-dessus du seuil d'alerte	95
7.2	Erreurs de modèles	96
7.3	Qualité de données	96
7.4	Le problème de l'exposition	97

7.5	Influence des sinistres survenus mais non déclarés	98
	Conclusion	99
	Bibliographie	106
A	Modèles	108
A.1	Modèles étudiés pour le décès	108
A.2	Modèles étudiés pour l'incapacité temporaire de travail	111
B	Choix d'une métrique pour la régression	113
C	Éléments de code	115
C.1	Fonction de coût pour la validation croisée	115
C.2	Fonctions génériques	116
C.3	Exemple d'utilisation	125

Abstract

Loan insurance, which protects creditors against the possible insolvency of their debtors, represents a very small share of the insurance market in France.

Traditionally dominated by the creditors themselves, this market showed nevertheless a high profitability, due to the high level of premium linked to too scarce competition, and a higher growth rate than the set of all other insurances.

Over the years, several regulatory texts have tried to change the situation and open the market, with the main consequence being an increasingly fierce race to lower prices. In addition, the latest law, which facilitates the possibilities for policyholders to change insurers and abolish medical questionnaires for a significant part of the contracts of insurance of real estate loans, could upset the claims experience and increase cancellation rates. Insurers operating in this market must therefore, more than ever, equip themselves with precise tools to monitor changes, whether structural or conjunctural, that could affect their portfolios.

In this context, the CUSUM process, in its version adapted to counting processes, because it allows for the monitoring of the intensities of point processes, and offers the possibility of setting up an alert system, could give an advantage to those who use it.

This paper presents, after a quick overview of loan insurance, the mathematical basis of the CUSUM process, gives elements for its concrete implementation and proposes, through two case studies, a reflection on the use that can be made of it by insurers.

Keywords : Loan insurance, Poisson process, Poisson disorder problem, CUSUM, Markov chain, Death, Work incapacity.

Résumé

L'assurance emprunteur, qui protège les créanciers contre l'éventuelle insolvabilité de leurs débiteurs, représente une toute petite part du marché de l'assurance en France. Traditionnellement dominé par les créanciers eux-mêmes, ce marché présentait néanmoins une forte rentabilité, du fait des tarifs élevés liés à une concurrence trop faible, et une croissance plus importante que celle de l'ensemble des autres assurances.

Au fil des ans, plusieurs textes réglementaires ont tenté de changer la donne et de forcer à une plus grande ouverture du marché, avec pour conséquence principale une course de plus en plus acharnée à la baisse des prix. Par ailleurs la dernière loi en date, qui facilite les possibilités, pour les assurés, de changer d'assureur, et supprime les questionnaires médicaux pour une partie non négligeable des contrats d'assurance de prêts immobiliers, pourrait bouleverser la sinistralité et accroître les taux de résiliations.

Les assureurs qui œuvrent sur ce marché doivent donc plus que jamais se doter d'outils précis pour suivre les changements, structurels ou conjoncturels, qui pourraient affecter leurs portefeuilles. Dans ce cadre, le processus CUSUM, dans sa version adaptée aux processus de comptage, parce qu'il permet un suivi des intensités des processus ponctuels, et offre la possibilité de mise en place d'un système d'alerte, pourrait donner un avantage à ceux qui l'utilisent.

Ce mémoire présente, après un rapide tour d'horizon de l'assurance emprunteur, les bases mathématiques du processus CUSUM, donne des éléments pour sa mise en œuvre concrète, et propose, à travers deux études de cas, une réflexion sur l'usage qui peut en être fait par les assureurs.

Mots-clés : Assurance emprunteur, Processus de Poisson, Poisson disorder problem, CUSUM, chaîne de Markov, Décès, ITT.

Remerciements

Ce travail n'aurait pas vu le jour si Juliano Merolle, qui m'a initié au domaine de l'assurance emprunteur, ne m'avait pas raconté son expérience en formation continue à l'ISFA, et s'il n'avait pas appuyé ma démarche lorsque j'ai souhaité à mon tour suivre cette formation. C'est donc à lui que s'adressent mes premiers remerciements.

J'adresse aussi des remerciements à Elie Toledano, qui a pris la décision de m'accorder le soutien de l'entreprise, alors que je n'y travaillais que depuis à peine une année.

Ma première rencontre avec le sujet de ce mémoire a eu lieu lors d'une présentation par Stéphane Loisel dans le cadre d'un cours de théorie de la ruine (théorie pour laquelle des processus du type de ceux étudiés ici interviennent, bien que ce ne soit pas abordé dans ce mémoire). De plus, alors que je lui faisais part de mon souhait d'essayer de mettre en œuvre ces méthodes et par ailleurs de trouver un sujet de mémoire, Stéphane Loisel s'est spontanément proposé pour suivre mon travail sur ce sujet particulier. Je lui suis reconnaissant de cette présentation, de cette spontanéité, mais aussi des conseils précieux qu'il m'a apportés ensuite pour l'élaboration de ce document.

Les échanges avec mes collègues de l'actuariat, du service gestion des sinistres, du service juridique ou encore du développement commercial sur les sujets abordés dans ce mémoire, ainsi que les relectures effectuées par certains d'entre eux, ont aussi fortement contribué à l'amélioration du présent document.

Je remercie en outre Romain Gauchon qui m'a fait l'honneur de relire ce travail et dont la vigilance m'a permis de corriger quelques imprécisions.

Ce mémoire doit aussi beaucoup à Morgann qui m'a aidé à m'y consacrer pleinement.

Note de Synthèse

0.1 Introduction

Le marché de l'assurance emprunteur, après des années de règne sans partage des prêteurs, s'ouvre lentement à la concurrence, avec pour premier effet notable une évolution à la baisse des prix. Dans ce contexte, une nouvelle loi, la **loi Lemoine**, vient bouleverser la donne, par la suppression des questionnaires médicaux pour une partie des prêts et par la possibilité de résiliation à tout moment.

Les assureurs se doivent donc d'être de plus en plus réactifs, pour répondre à la concurrence, aux évolutions réglementaires et à celles de la sinistralité. Pour ce dernier point, le processus **CUSUM** (pour **CUmulative SUM control chart**) proposé par E. S. Page dans les années 50 pourrait être un outil extrêmement utile.

L'étude qui suit vise à évaluer la pertinence de son utilisation, et présente deux études de cas qui permettent de soulever quelques questions importantes.

0.2 Détection des changements d'intensité dans les Processus de Poisson

0.2.1 Problématique

La présente formulation du problème suit l'article [El Karoui et al., 2017]

Soit un processus de comptage $(N_t)_{t \geq 0}$, et $\tilde{\lambda} = (\tilde{\lambda}_t)_{t \geq 0}$ son intensité.

Le processus de comptage N_t est observable séquentiellement à travers la filtration $\mathcal{F}_t = \sigma\{N_s, 0 \leq s \leq t\}$.

Le processus subit un changement d'intensité de λ_t à $\rho\lambda_t$ à un temps θ inconnu. Autrement dit :

$$\tilde{\lambda}_t = \lambda_t \mathbb{1}_{\{t \leq \theta\}} + \rho\lambda_t \mathbb{1}_{\{t > \theta\}}, \rho \neq 1.$$

L'intensité cumulée du processus est notée $\Lambda_t = \int_0^t \lambda_s ds$

Le problème est de trouver une règle pour détecter θ le plus rapidement possible, avec un nombre limité de fausses alarmes sans hypothèse émise sur la distribution de la variable aléatoire θ (donc dans cadre non bayésien).

Tant que le changement n'a pas eu lieu, le processus $(N_t)_{t \geq 0}$ est considéré sous la mesure de probabilité \mathbb{P} , associée à l'intensité $(\lambda_t)_{t \geq 0}$.

Une probabilité équivalente, \mathbb{P}_θ , décrit le cas où le changement d'intensité a lieu en θ . Ainsi $\mathbb{P}_\infty = \mathbb{P}$ (i.e. le changement n'a jamais lieu).

0.2.2 Critère pour définir la meilleure approche

Le critère présenté ci-dessous est de type **minimax**. Il s'agit de minimiser le délai avant la détection dans la pire situation possible, tout en limitant le nombre de fausses alertes.

Une version modifiée du critère de Lorden

Les auteurs de [El Karoui et al., 2017] proposent, comme ceux de [Gandy et al., 2010] dans le même esprit, de s'attacher au nombre moyen d'évènements entre le changement et sa détection, et au nombre moyen d'évènements avant une fausse alarme :

- Le nombre moyen d'évènements entre le changement en θ et la détection à un temps τ :

$$\mathbb{E} \left[(N_\tau - N_\theta)^+ | \mathcal{F}_\theta \right].$$

- Nombre moyen d'évènements avant une fausse alarme :

$$\mathbb{E}_\infty [N_\tau].$$

Le pire cas s'écrira alors :

$$C(\tau) = \sup_{\theta \in [0, \infty]} \operatorname{ess\,sup}_\omega \mathbb{E} \left[(N_\tau - N_\theta)^+ | \mathcal{F}_\theta \right].$$

Ce critère a le mérite d'être indifférent aux procédures de changement de temps. Et le problème à résoudre est ici de trouver τ solution de

$$\inf_T \{C(T) | \mathbb{E}_\infty [N_T] \geq \pi\}.$$

0.2.3 Processus CUSUM pour les processus de Poisson

Principe élémentaire

Si une série d'observations est soumise à une loi statistique de paramètre θ_0 avant qu'un changement intervienne et que la nouvelle loi des observations soit de paramètre θ_1 , le logarithme du rapport de vraisemblance pour l'observation i s'écrira :

$$s_i = \ln \frac{p_{\theta_1}(y_i)}{p_{\theta_0}(y_i)}.$$

L'espérance de cette quantité sera négative sous \mathbb{P}_{θ_0} et positive sous \mathbb{P}_{θ_1} . La somme des logarithmes des rapports de vraisemblance est notée :

$$S_k = \sum_{i=1}^k s_i.$$

S_k montre une tendance à la baisse avant le changement, et une tendance à la hausse ensuite. L'information pertinente se trouve donc dans la différence entre la valeur du logarithme du rapport de vraisemblance et le minimum qu'il ait pris jusqu'alors :

$$g_k = S_k - m_k \geq h,$$

avec

$$m_k = \min_{1 \leq j \leq k} S_j.$$

LPSR pour le processus de Poisson

Le processus séquentiel du logarithme du rapport de vraisemblance (LPSR pour Log Sequential Probability Ratio) sous \mathbb{P}_θ et \mathbb{P}_∞ , noté U , avec l'information disponible en t_n s'écrira pour tout $t > 0$

$$U_t = \ln(\rho)N_t + (1 - \rho)\Lambda_t.$$

Notation pratique

A un changement d'échelle près, il est assez courant d'écrire

$$U_t = N_t + \frac{(1 - \rho)}{\ln(\rho)}\Lambda_t.$$

C'est la notation qui sera retenue dans la suite.

CUSUM pour le processus de Poisson

Cas où $\rho > 1$

Dans ce cas, la tendance de U devrait être à la baisse si $t < \theta$ et à la hausse sinon. Il convient donc de s'intéresser au processus CUSUM suivant :

$$V_t = U_t - \min_{s \leq t} \{U_s\}.$$

Cas où $\rho < 1$

Dans ce cas, il s'agit de détecter une baisse de l'intensité du processus. Plutôt que de se baser sur U , la méthode de détection utilisera X défini par :

$$X_t = -U_t,$$

et le processus CUSUM Y donné par :

$$Y_t = X_t - \min_{s \leq t} \{X_s\}.$$

Optimalité du processus CUSUM

C'est dans [El Karoui et al., 2017] que se trouve, dans un cadre non bayésien, la démonstration de l'optimalité de la décision basée sur le processus CUSUM dans le cas des processus de Poisson, pour le critère de Lorden modifié.

0.3 Considérations pratiques

L'intensité avant le changement : régression de Poisson

Pour calculer les valeurs du processus CUSUM, il est nécessaire de connaître le paramètre θ_0 . Dans ce travail, l'estimation est obtenue par une régression de Poisson (ce qui n'est qu'une approche possible parmi d'autres).

Le modèle est :

$$\mathbb{E}(Y|X) = \lambda_0 e^{\beta'X + \ln(E)}. \quad (1)$$

Le coefficient associé au logarithme de l'exposition est 1 (ce qui signifie simplement que le nombre moyen d'évènements observés varie comme l'exposition, toutes choses étant égales par ailleurs).

Détermination de ρ ampleur du changement à détecter

Le niveau de changement à détecter est déterminé par des considérations pragmatiques d'appétence au risque. Les questions qui peuvent se poser sont par exemple : À partir de quelle baisse de l'intensité de la sinistralité est-il possible d'envisager une baisse de tarif ? A partir de quelle hausse faut-il considérer que le tarif devient trop bas ? Ou encore : à partir de quel niveau de changement de l'intensité des sinistres l'impact devient-il non négligeable dans les projections nécessaires au pilier I de Solvabilité II ?

Choix du seuil d'alerte

Après avoir fait le choix de ρ , il est nécessaire de déterminer la fréquence acceptable de fausses alertes (cf. [sous-section 2.2.3](#)), faute de quoi il ne serait pas possible de déterminer le niveau du processus CUSUM au-delà duquel lancer une alerte.

Tout processus de Poisson, homogène ou non, sous des hypothèses assez générales, peut se ramener à un processus d'intensité 1 par changement de temps.

Grâce à cette observation, le seul nombre d'évènements observés suffira à déterminer le niveau atteint par le processus CUSUM.

L'objectif est donc de déterminer le nombre moyen d'évènements, dans l'hypothèse où il n'y a pas de changement dans l'intensité du processus étudié, avant que le seuil d'alerte soit franchi, c'est à dire $\mathbb{E}_\infty [N_\tau]$.

L'utilisation d'une chaîne de Markov est proposée dans une annexe de l'article [[Gandy et al., 2010](#)] pour caractériser les dépassements progressifs de niveaux intermédiaires entre 0 et le seuil d'alerte. Dans la suite, un processus d'intensité constante λ est considéré, et \mathbb{P}_λ désigne la probabilité associée. De plus, quel que soit λ , c'est au processus CUSUM défini pour un changement d'intensité proportionnel de proportion ρ que ce qui suit s'applique.

Cas $\rho > 1$

Pour un seuil d'alerte c , posons :

$$a = \frac{\ln(\rho)}{1 - \rho},$$

$$k = \lceil c - 1 \rceil.$$

C'est à dire que k est le plus petit entier supérieur ou égal à $c - 1$. Notons de plus $\xi = a(c - k)$, et $\eta = a - \xi$.

Une chaîne de Markov M à k états transients et un état absorbant est définie de la façon suivante. On pose $\sigma_0 = 0$, puis

$$\sigma_n = \inf\{t > \sigma_{n-1} : U_t \in \{c - k, \dots, c - 1\}\},$$

et

$$M_n = \begin{cases} \infty, & \sigma_n \geq \tau, \\ \lceil U_{\sigma_n} \rceil, & \sigma_n < \tau, \end{cases}$$

où $\tau = \inf\{t | U_t > c\}$.

La probabilité d'observer un nombre $j > 0$ d'évènements pendant la durée t , pour un processus de Poisson d'intensité constante λ est notée :

$$p_j(t) = \frac{(\lambda t)^j}{j!} e^{-\lambda t}. \quad (2)$$

Notant $a_j(\xi, \eta)$ pour la probabilité de transition de l'état 1 à l'état j , il vient :

$$\begin{aligned} a_j(\xi, \eta) &= p_0(\xi)p_{j-1}(\eta) + \sum_{l=1}^{l=j} p_l(\xi)p_{j-l}(\eta) \\ &= p_j(\xi + \eta) + p_0(\xi)(p_{j-1}(\eta) - p_j(\eta)). \end{aligned}$$

Et quelques réflexions supplémentaires conduisent à la matrice de transition (pour les états transients) de M :

$$Q = Q(\xi, \eta) = \begin{pmatrix} a_1(\xi, \eta) & a_2(\xi, \eta) & \cdots & a_k(\xi, \eta) \\ p_0(\xi + \eta) & p_1(\xi + \eta) & \cdots & p_{k-1}(\xi + \eta) \\ 0 & p_0(\xi + \eta) & \ddots & p_{k-2}(\xi + \eta) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & p_0(\xi + \eta) & p_1(\xi + \eta) \end{pmatrix}$$

Distribution de M_1

Notons maintenant ν le nombre minimal de sauts nécessaires pour que U puisse passer du niveau U_0 au niveau c , c'est à dire $\nu = \lceil c - U_0 \rceil$.

Si $U_0 > c - k$, dans ce cas la distribution de M_1 sera :

$$q = (0, \dots, 0, p_0(t_0), \dots, p_{\nu-1}(t_0)).$$

Par ailleurs, si $U_0 < c - k$, il vient $q = (a_1(aU_0, \eta), \dots, a_k(aU_0, \eta))$, qui peut se réécrire :

$$q = (1, 0, \dots, 0)Q(aU_0, \eta).$$

Nombre moyens de sauts pour franchir le niveau c

Le nombre moyen de sauts avant le dépassement du seuil c est donné par :

$$\mathbb{E}_\lambda(N_\tau) = \nu + q(I - Q)^{-1}\mathbf{1},$$

où $\mathbf{1}$ est le vecteur dont tous les termes sont égaux à 1.

Cas $\rho < 1$

Les définitions de a, ξ, η, ν, p_j sont les mêmes que dans le cas précédent.

La chaîne de Markov qui va être utilisée est la suivante (cf. 4.3) : On pose $\sigma_0 = 0$, puis

$$\sigma_n = \inf\{t > \sigma_{n-1} : U_t \in \{0, c - k, \dots, c - 1\}\},$$

et

$$M_n = \begin{cases} \infty, & \sigma_n \geq \tau, \\ \lceil Y_{\sigma_n} \rceil, & \sigma_n < \tau. \end{cases}$$

Posons alors, pour $j \geq 2$,

$$\begin{aligned} b_j(\xi, \eta) &= p_{j-1}(\eta + \xi) + p_0(\xi) \sum_{l=j}^{l=\infty} p_l(\eta) \\ &= p_{j-1}(\eta + \xi) + p_0(\xi) \left(1 - \sum_{l=0}^{l=j-1} p_l(\eta)\right). \end{aligned}$$

La matrice de transition (pour les états transients) de M est :

$$Q = Q(\xi, \eta) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 - \sum_{j=0}^{j=k-1} p_j(\eta) & p_0(\eta) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 1 - \sum_{j=0}^{j=k-1} p_j(\eta) & p_{k-1}(\eta) & \ddots & p_0(\eta) \end{pmatrix} * \begin{pmatrix} 1 - \sum_{j=0}^{j=0} p_j(\xi) & p_0(\xi) & \cdots & 0 \\ \vdots & p_1(\xi) & \ddots & \vdots \\ \vdots & \vdots & \ddots & p_0(\xi) \\ 1 - \sum_{j=0}^{j=k} p_j(\xi) & p_k(\xi) & \ddots & p_1(\xi) \end{pmatrix}$$

Distribution de M_1

La distribution de la chaîne dans les états transients est donc :

$$q = e_{k+2-\nu} Q(\max(a - t_0 - \xi), \min(a - t_0)),$$

où $e_{k+2-\nu}$ est le vecteur dont tous les termes sont nuls sauf celui qui se trouve à la position $k + 2 - \nu$, qui vaut 1.

Nombre moyen de sauts pour franchir le niveau c

Le nombre moyen de sauts pour atteindre le niveau c est maintenant :

$$\mathbb{E}_\lambda(N_\tau) = qQ^{\nu-1}(I - Q)^{-1}\mathbf{1}.$$

Obtenir un seuil d'alerte

La chaîne Markov nous donne, pour un seuil donné, le nombre moyen d'évènements avant une fausse alerte, et non le contraire. Pour répondre à la question, il suffit donc de définir quel est le nombre moyen d'évènements qui devraient, en dehors d'un changement, arriver en x années, puis, par exemple avec une fonction d'optimisation, retrouver le seuil c correspondant.

0.4 Résultats

Pour tester la méthode dans un contexte réel, le taux de référence est modélisé à partir d'observations sur 2015-2018. Puis les processus CUSUM sont étudiés à partir du 1^{er} janvier 2019.

L'application de ces principes amène à des graphiques tels que [1](#) (pour le décès) et [2](#) (pour l'ITT). Le choix du seuil d'alerte (des tentatives sont présentées avec des seuils définis à partir d'une fausse alerte tous les trois ans ou bien tous les vingt ans) est mis en question. L'évolution de l'exposition totale s'avère aussi un éventuel sujet dans la définition de ces seuils.

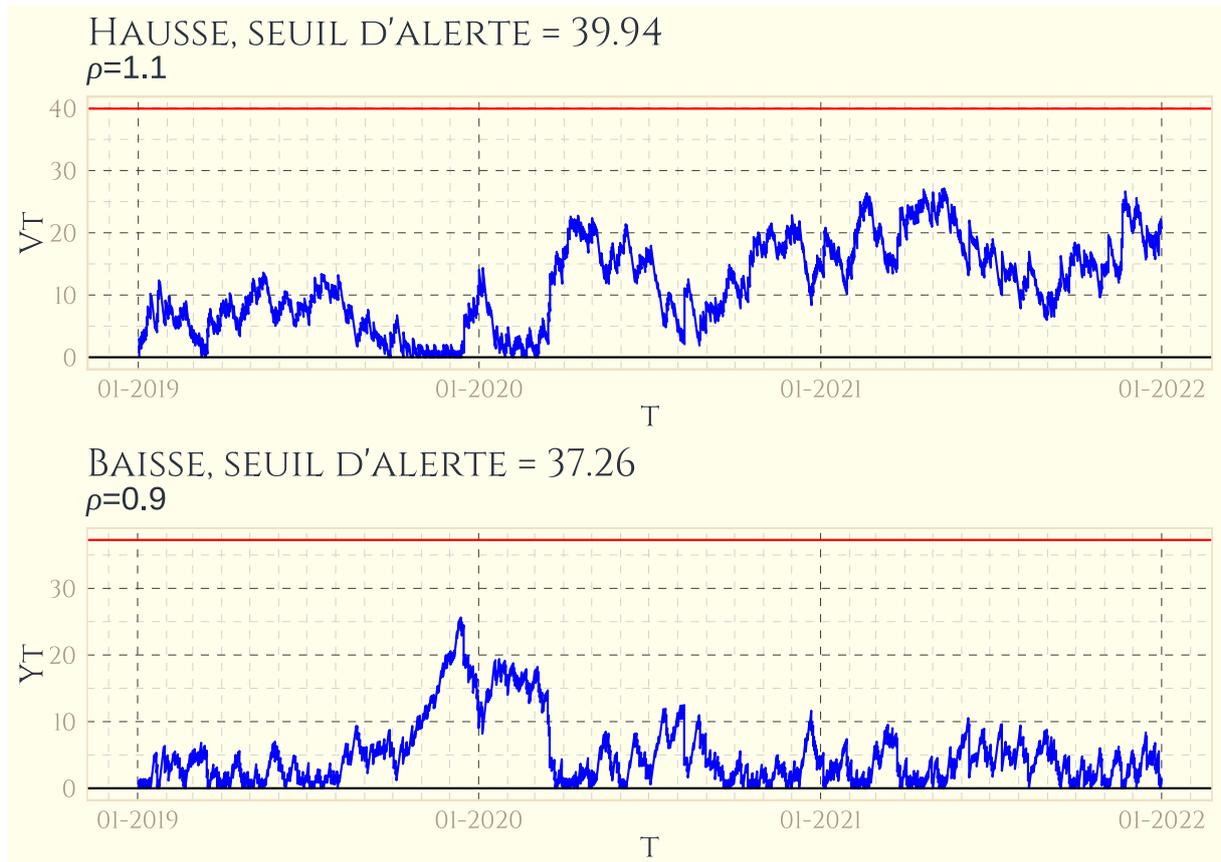


FIGURE 1 – Évolution du processus CUSUM pour le décès (population générale) depuis le 1er janvier 2019, seuil d’alerte fixé pour une fausse alerte tous les vingt ans

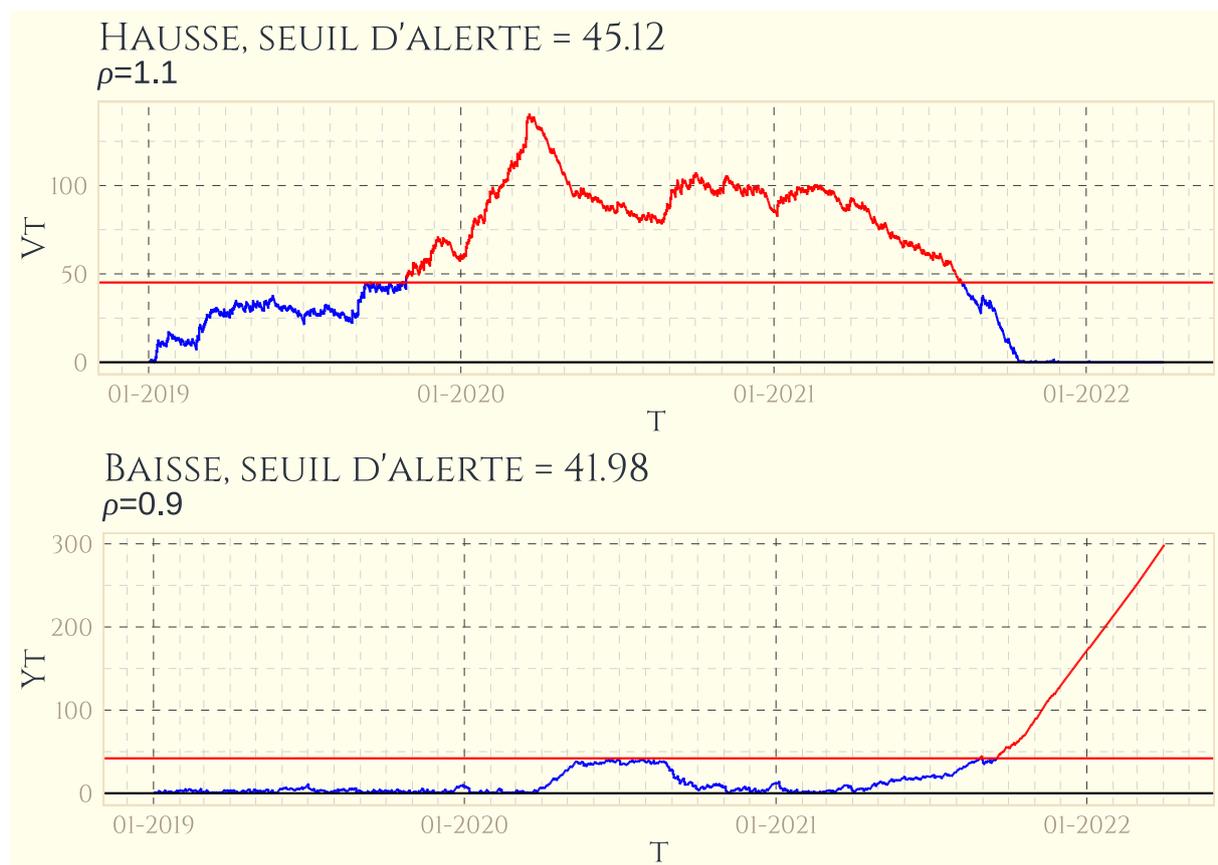


FIGURE 2 – Évolution du processus CUSUM pour l'ITT (population générale) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans

Dans les deux cas, le processus CUSUM peut être adapté pour des sous-populations, ce qui peut donner quelques indications. Pour le décès, une croissante soudaine du CUSUM pour la détection d'un changement à la hausse a lieu sur l'année 2021 pour les plus âgés (75-84 ans).

Dans le cas des arrêts de travail (garantie ITT), une hausse assez généralisée semble avoir lieu dès début 2019.

0.5 Conclusions

La méthode étudiée est relativement aisée à mettre en œuvre.

Le niveau de changement dont la détection pourrait nécessiter une alerte relève des choix de l'assureur.

Le nombre de maximal de fausses alertes qui détermine le seuil d'alerte doit être choisi sur la base d'un arbitrage entre le risque de ne pas percevoir un changement et le risque qu'à force de trop crier au loup les alertes ne soient plus considérées comme pertinentes.

La modélisation du taux de référence doit être menée avec le plus de soin de possible, et la qualité des données, pour cette modélisation comme pour la construction ultérieure du processus CUSUM doit être la meilleure possible.

L'évolution des courbes CUSUM est liée aux caractéristiques du logarithme des log-vraisemblances, et leur étude comme leur présentation doivent être assorties de précautions.

La durée du changement d'intensité, s'il y en a un, peu conduire le CUSUM à des niveaux élevés, qu'il faut donc se garder d'interpréter comme des signes d'un changement

de plus grande ampleur encore.

Enfin, malgré la tendance naturelle à vouloir interpréter les courbes, les causes réelles des alertes (ou même de l'aspect des courbes) ne peuvent être déduites de ces seules courbes. Leur utilité est d'attirer l'attention, et le plus vite possible, sur des changements qui auraient lieu, mais n'est pas de les expliquer. En cas d'alertes, d'autres études doivent permettre d'en déterminer les raisons et les éventuelles décisions à prendre (par exemple : révision des tarifs, des niveaux de garanties, des exclusions, des traités de réassurance, des projections dans le cadre de Solvabilité II).

Synthesis

0.6 Introduction

The loan insurance market, after years of undivided reign by the lenders, is slowly opening up to competition, with the first notable effect being a downward trend in prices. In this context, a new law, the **loi Lemoine**, has upset the situation, by abolishing medical questionnaires for a part of the loans and by the possibility of cancellation at any time.

Insurers must therefore be more and more reactive in order to respond to competition, regulatory changes and changes in the claims experience. For this last point, the process **CUSUM** (for **CUMulative SUM control chart**) proposed by E. S. Page in the 1950s could be an extremely useful tool.

The following work aims to study the relevance of its use, and presents two case studies that raise some important questions.

0.7 Detection of intensity changes in Poisson processes

0.7.1 Problem

This formulation of the problem follows the article [El Karoui et al., 2017]

Let $(N_t)_{t \geq 0}$ be a counting process, and $\tilde{\lambda} = (\tilde{\lambda}_t)_{t \geq 0}$ his intensity.

The counting process N_t can be observed sequentially through the filtration $\mathcal{F}_t = \sigma\{N_s, 0 \leq s \leq t\}$.

The process undergoes a change from λ_t to $\rho\lambda_t$ at an unknown time θ . That is to say :

$$\tilde{\lambda}_t = \lambda_t \mathbb{1}_{\{t \leq \theta\}} + \rho \lambda_t \mathbb{1}_{\{t > \theta\}}, \rho \neq 1.$$

The cumulative intensity of the process is $\Lambda_t = \int_0^t \lambda_s ds$.

The problem is to find a rule to detect θ as quickly as possible, with a limited number of false alarms without hypotheses on the distribution of the random variable θ (thus in a non Bayesian framework).

As long as the change has not taken place, the process $(N_t)_{t \geq 0}$ is considered under the probability measure \mathbb{P} , associated with the intensity $(\lambda_t)_{t \geq 0}$. An equivalent probability, \mathbb{P}_θ , is the probability measure that describes the case where the change in intensity occurs in θ . Thus $\mathbb{P}_\infty = \mathbb{P}$ (i.e. the change never takes place).

0.7.2 Criterion for determining the best approach.

The criterion presented below are of type **minimax**. The goal is to minimize the delay before detection in the worst possible situation, while limiting the number of false alarms.

A modified version of Lorden's criterion The authors of [El Karoui et al., 2017] propose, like those of [Gandy et al., 2010] in the same spirit, to focus on the average number of events between the change and its detection, and the average number of events before a false alarm:

- The average number of events between the change at θ and its detection at time τ :

$$\mathbb{E} \left[(N_\tau - N_\theta)^+ | \mathcal{F}_\theta \right].$$

- Average number of events before a false alarm:

$$\mathbb{E}_\infty [N_\tau].$$

The worst case will then be written:

$$C(\tau) = \sup_{\theta \in [0, \infty]} \operatorname{ess\,sup}_\omega \mathbb{E} \left[(N_\tau - N_\theta)^+ | \mathcal{F}_\theta \right].$$

This criterion has the merit of being indifferent to the time change procedures. And the problem to solve here is to find τ solution of

$$\inf_T \{C(T) | \mathbb{E}_\infty [N_T] \geq \pi\}.$$

0.7.3 CUSUM process for Poisson processes

Basic principle

Given a series of observations subject to a statistical law of parameter θ_0 before a change occurs and the new law of the observations be of parameter θ_1 , the log likelihood ratio for observation i will be written:

$$s_i = \ln \frac{p_{\theta_1}(y_i)}{p_{\theta_0}(y_i)}.$$

The expectation of this quantity will be negative under \mathbb{P}_{θ_0} and positive under \mathbb{P}_{θ_1} . The sum of the log likelihood ratios is denoted by:

$$S_k = \sum_{i=1}^k s_i.$$

S_k shows a negative trend before the change, and a positive trend afterwards. The relevant information is therefore found in the difference between the value of the log likelihood ratio and the minimum it has taken so far:

$$g_k = S_k - m_k \geq h,$$

avec

$$m_k = \min_{1 \leq j \leq k} S_j.$$

LPSR for Poisson process

The Log Sequential Probability Ratio(LPSR) under \mathbb{P}_θ and \mathbb{P}_∞ , noted U , with available information at time t_n will be written $t > 0$

$$U_t = \ln(\rho)N_t + (1 - \rho)\Lambda_t.$$

Practical notation

With one change of scale, it is quite common to write

$$U_t = N_t + \frac{(1 - \rho)}{\ln(\rho)} \Lambda_t.$$

This is the notation that will be used in the following.

CUSUM for Poisson process

When $\rho > 1$

In this case, the trend of U should go downward if $t < \theta$ and upward otherwise. It is therefore appropriate to look at the following CUSUM process:

$$V_t = U_t - \min_{s \leq t} \{U_s\}.$$

When $\rho < 1$

In this case, the aim is to detect a decrease in the intensity of the process. Instead of using U , the detection method will use X defined by:

$$X_t = -U_t,$$

and the CUSUM Y process given by:

$$Y_t = X_t - \min_{s \leq t} \{X_s\}.$$

Optimality of the CUSUM process

[El Karoui et al., 2017] gives, in a non-Bayesian framework, the demonstration of the optimality of the decision based on the CUSUM process in the case of Poisson processes, for the modified Lorden's criteria.

0.8 Practical considerations

Intensity before the change: Poisson regression

It's necessary to know the parameter θ_0 (i.e before change) to compute the CUSUM process. In this work, the estimate is obtained by a Poisson regression (which is only one possible approach among others).

The model is:

$$\mathbb{E}(Y|X) = \lambda_0 e^{\beta' X + \ln(E)}.$$

The coefficient associated with the logarithm of exposure is 1 (which simply means that the average number of observed events varies with exposure, all else being equal).

Determination of ρ magnitude of change to detect

The level of change to be detected is determined by pragmatic considerations of risk appetite. Questions that may arise are, for example: At what level of loss intensity is it possible to consider lowering the premium rates? At what point should we consider that the premium rates are too low? Or else: at what level of change in claims intensity does the impact become non-negligible in the projections required by Pillar I of Solvency II?

Calculation of the alert threshold

Once the choice of ρ has been made, it is necessary to determine the acceptable frequency of false alarms (cf. [subsection 2.2.3](#)), otherwise it would not be possible to determine the level of the CUSUM process beyond which to launch an alert.

Any Poisson process, homogeneous or not, under fairly general assumptions, can be reduced to a process of intensity 1 per time change.

Thanks to this observation, the number of observed events alone will be sufficient to determine the level reached by the CUSUM process.

The objective is therefore to determine the average number of events, under the hypothesis that there is no change in the intensity of the process studied, before the alert threshold is crossed, i.e. $\mathbb{E}_\infty [N_\tau]$.

An appendix of the article proposes the use of a Markov chain to characterize the progressive exceedances of intermediate levels between 0 and the alert threshold. In the following, a constant intensity process λ is considered, and \mathbb{P}_λ denotes the associated probability. Moreover, whatever λ is, it is to the CUSUM process defined for a proportional change of intensity of proportion ρ that what follows applies.

When $\rho > 1$

For an alert threshold c , let :

$$a = \frac{\ln(\rho)}{1 - \rho},$$

$$k = \lceil c - 1 \rceil.$$

That is to say, k is the ceiling of $c - 1$. Let's also note $\xi = a(c - k)$, and $\eta = a - \xi$.

A Markov chain M with k transients states and one absorbing state is defined in the following way : let $\sigma_0 = 0$, and

$$\sigma_n = \inf\{t > \sigma_{n-1} : U_t \in \{c - k, \dots, c - 1\}\},$$

and then

$$M_n = \begin{cases} \infty, & \sigma_n \geq \tau, \\ \lceil U_{\sigma_n} \rceil, & \sigma_n < \tau. \end{cases}$$

where $\tau = \inf\{t | U_t > c\}$.

The probability of observing a number $j > 0$ of events during the duration t , for a Poisson process of constant intensity λ is noted :

$$p_j(t) = \frac{(\lambda t)^j}{j!} e^{-\lambda t}, \tag{3}$$

Noting $a_j(\xi, \eta)$ for the transition probability from state 1 to state j , it follows:

$$\begin{aligned} a_j(\xi, \eta) &= p_0(\xi)p_{j-1}(\eta) + \sum_{l=1}^{l=j} p_l(\xi)p_{j-l}(\eta) \\ &= p_j(\xi + \eta) + p_0(\xi)(p_{j-1}(\eta) - p_j(\eta)). \end{aligned}$$

And a few more thoughts lead to the transition matrix (for transient states) of M :

$$Q = Q(\xi, \eta) = \begin{pmatrix} a_1(\xi, \eta) & a_2(\xi, \eta) & \cdots & a_k(\xi, \eta) \\ p_0(\xi + \eta) & p_1(\xi + \eta) & \cdots & p_{k-1}(\xi + \eta) \\ 0 & p_0(\xi + \eta) & \ddots & p_{k-2}(\xi + \eta) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & p_0(\xi + \eta) & p_1(\xi + \eta) \end{pmatrix}.$$

Distribution of M_1

Let the minimal number of jumps needed for U to hit the threshold c be ν , i.e $\nu = \lceil c - U_0 \rceil$.

If $U_0 > c - k$, then the distribution of M_1 will be :

$$q = (0, \dots, 0, p_0(t_0), \dots, p_{\nu-1}(t_0)).$$

Otherwise , if $U_0 < c - k$, then $q = (a_1(aU_0, \eta), \dots, a_k(aU_0, \eta))$, which can be rewritten as:

$$q = (1, 0, \dots, 0)Q(aU_0, \eta).$$

Average number of jumps to reach the level c

The average number of jumps before the threshold c is reached is given by :

$$\mathbb{E}_\lambda(N_\tau) = \nu + q(I - Q)^{-1}\underline{1}.$$

where $\underline{1}$ is the vector whose all terms are equal to 1.

When $\rho < 1$

Définitions of a, ξ, η, ν, p_j are the same.

The Markov chain to study is here given by (cf. 4.3) : Let $\sigma_0 = 0$, and

$$\sigma_n = \inf\{t > \sigma_{n-1} : U_t \in \{0, c - k, \dots, c - 1\}\},$$

and

$$M_n = \begin{cases} \infty, & \sigma_n \geq \tau, \\ \lceil Y_{\sigma_n} \rceil, & \sigma_n < \tau. \end{cases}$$

Then let, for $j \geq 2$:

$$\begin{aligned} b_j(\xi, \eta) &= p_{j-1}(\eta + \xi) + p_0(\xi) \sum_{l=j}^{l=\infty} p_l(\eta) \\ &= p_{j-1}(\eta + \xi) + p_0(\xi) \left(1 - \sum_{l=0}^{l=j-1} p_l(\eta)\right). \end{aligned}$$

The transition matrix (for transient states) of M is :

$$Q = Q(\xi, \eta) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 - \sum_{j=0}^{j=0} p_j(\eta) & p_0(\eta) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 1 - \sum_{j=0}^{j=k-1} p_j(\eta) & p_{k-1}(\eta) & \ddots & p_0(\eta) \end{pmatrix} * \begin{pmatrix} 1 - \sum_{j=0}^{j=0} p_j(\xi) & p_0(\xi) & \cdots & 0 \\ \vdots & p_1(\xi) & \ddots & \vdots \\ \vdots & \vdots & \ddots & p_0(\xi) \\ 1 - \sum_{j=0}^{j=k} p_j(\xi) & p_k(\xi) & \ddots & p_1(\xi) \end{pmatrix}$$

Distribution of M_1

The distribution of the chain in the transient states is therefore :

$$q = e_{k+2-\nu} Q(\max(a - t_0 - \xi), \min(a - t_0)),$$

where $e_{k+2-\nu}$ is the vector whose all terms are zero except the one indexed by $k + 2 - \nu$, which is 1.

Average number of jumps to reach level c

The average number of jumps to reach level c is now :

$$\mathbb{E}_\lambda(N_\tau) = qQ^{\nu-1}(I - Q)^{-1}\mathbf{1}.$$

Getting an alert threshold

The Markov chain gives us, for a given threshold, the average number of events before a false alarm, and not the opposite. To answer the question, it is therefore sufficient to define what is the average number of events that should, apart from a change, occur in x years, and then, for example with an optimisation function, find the corresponding c threshold.

0.9 Results

To test the method in a real context, the reference rate is modelled from observations over 2015-2018. Then the CUSUM processes are studied from 1 January 2019.

The application of these principles leads to graphs such as 3 (for death) and 4 (for temporary work incapacity). The choice of the alert threshold (attempts are presented with thresholds defined from a false alarm every three years or every twenty years) is questioned. The evolution of the total exposure is also a possible issue in the definition of these thresholds.



Figure 3 – Evolution of the CUSUM process for death (general population) since 1st January 2019, threshold for a false alarm every 20 years

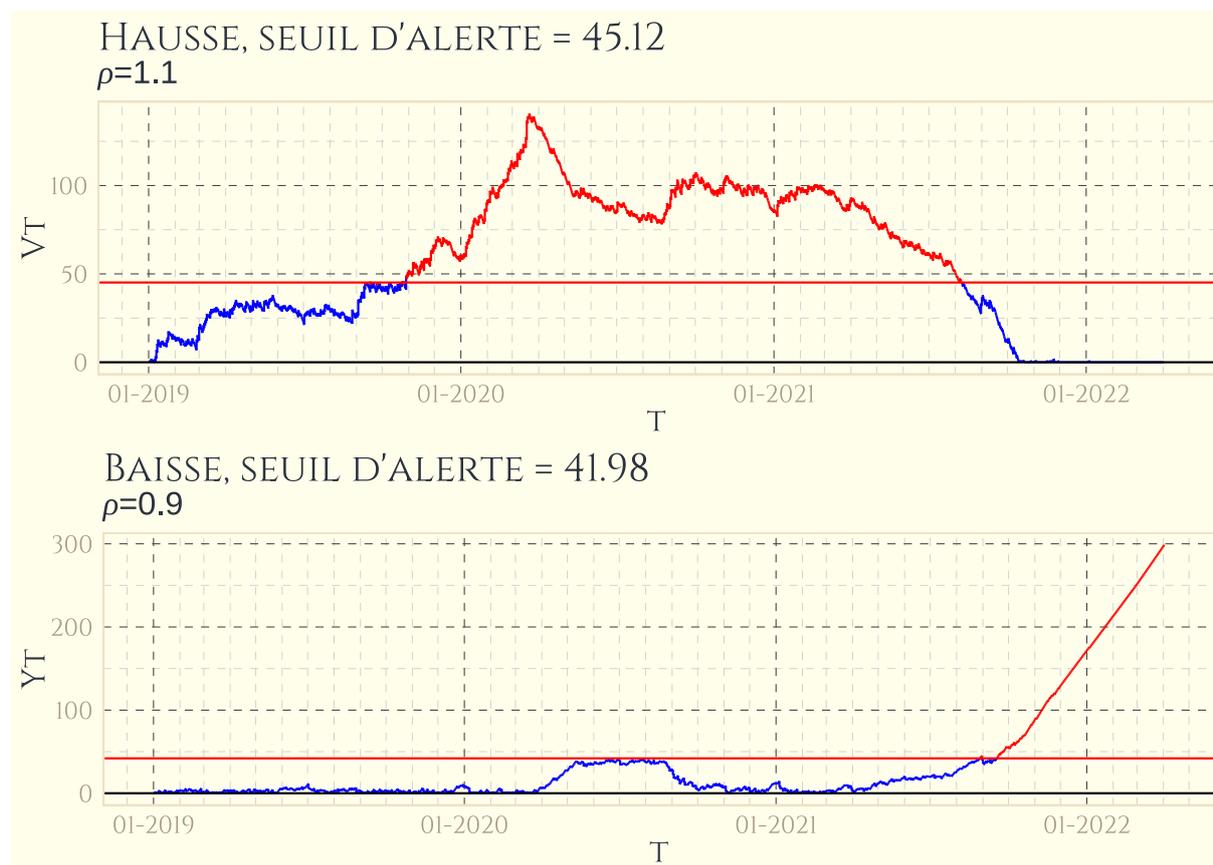


Figure 4 – Evolution of the CUSUM process for temporary work incapacity (general population) since 1st January 2019, alert threshold set for one false alarm every twenty years

In both cases, the CUSUM process can be adapted for sub-populations, which may give some indication. For death, a sudden increase in the CUSUM for detecting an upward change occurs in the year 2021 for the oldest age group (75-84 years).

In the case of work stoppages (ITT coverage), a fairly generalised increase seems to take place from the beginning of 2019.

0.10 Conclusions

The method studied is relatively easy to implement.

The level of change that might require an alert is a matter of choice for the insurer.

The maximum number of false alarms that determines the alert threshold must be chosen on the basis of a trade-off between the risk of not perceiving a change and the risk that, after too many pointless alerts, the alerts are no longer considered relevant.

The modelling of the reference rate must be carried out as carefully as possible, and the quality of the data, both for this modelling and for the subsequent construction of the CUSUM process, must be as high as possible.

The evolution of the CUSUM curves is related to the characteristics of the log-likelihoods, and care must be taken in their study and presentation.

The duration of the change in intensity, if any, can lead to high levels of CUSUM, which should not be interpreted as signs of a larger change.

Finally, despite the natural tendency to interpret the curves, the real causes of the alerts (or even the appearance of the curves) cannot be deduced from the curves alone. Their usefulness is to draw attention, as quickly as possible, to changes that may have occurred, but not to explain them. In the event of alerts, further studies must be carried out to determine the reasons for them.

Introduction

Le premier cas donné par Christian Morel en exemple dans le premier volume de son ouvrage *Les décisions absurdes* [Morel, 2002], est celui du vol British Midlands 92 du 8 janvier 1989 entre Londres et Belfast. Le Boeing 737-400 s'est écrasé à la suite d'un problème moteur, et le crash a fait 47 morts. Pourtant ce n'est pas sur la défaillance d'un moteur que l'auteur souhaite attirer notre attention, mais sur la décision qui a été prise par les pilotes d'arrêter le seul réacteur qui fonctionnait encore, décision qui a précipité l'aéroplane dans un talus, alors que s'ils avaient maintenu en fonction ce réacteur les pilotes auraient pu tenter un atterrissage avec peu de dégâts supplémentaires. Et le motif de cette décision, que tout semble en effet désigner comme absurde, est lié à l'habitude qu'avaient les pilotes de modèles d'avions plus anciens. Sur ces anciens modèles, l'air conditionné fonctionnait avec le seul moteur droit, et une odeur de fumée dans le cockpit signifiait donc que ce moteur était en feu. Si des passagers signalaient des fumées aperçues par les hublots, l'absence d'odeurs de fumée signifiait au contraire que l'incendie provenait du réacteur gauche. Leur connaissance de ces éléments a mené à la décision, pour eux rationnelle, puisqu'il fallait éteindre le moteur en feu et qu'une odeur était sensible dans le cockpit, de couper le moteur droit. Il se trouve que sur les Boeing 747-400, le constructeur a fait fonctionner l'air conditionné avec les deux moteurs, et que les pilotes, qui ignoraient cette évolution, ont éteint le seul des deux réacteurs qui marchait correctement. Ignorer un changement peut donc être fatal.

Dans le domaine de l'assurance, comme dans celui du pilotage d'avions de ligne, l'ignorance des changements peut s'avérer sinon fatale, du moins fortement dommageable, et prendre le plus vite possible la bonne décision est toujours un enjeu majeur. C'est pourquoi les assureurs ont toujours intérêt à se doter des moyens les plus efficaces pour détecter le moindre changement, quel qu'en soit la nature, et quelles qu'en soient les causes. Il n'est probablement pas possible de s'attaquer de front au *changement*, dans toute sa généralité. Mais se concentrer sur certains types de changements, qui partagent quelque chose de commun dans leurs natures, peut mener à des méthodes concrètes et applicables pour une classe assez vaste de phénomènes. C'est le cas par exemple des changements qui interviennent dans les processus de comptage. Dès que des événements peuvent être comptés, ils peuvent potentiellement être étudiés par le prisme des processus stochastiques. Et il se trouve justement que les changements dans ces processus font depuis longtemps l'objet d'une attention particulière, dans des domaines très variés et à l'origine éloignés du monde de l'assurance.

Dès 1931, un chercheur des Bell Telephone Laboratories, Walter A. Shewart, met au point les premières **cartes de contrôles**, destinées à identifier au mieux, dans une série d'évènements, ce qui relève d'un aléa naturel de ce qui tient à un aléa accidentel. Ceci a donné naissance à la **Maîtrise statistique des processus**. Popularisées au Japon après la seconde guerre mondiale par un disciple de Shewart, les succès de ces méthodes dans ce pays ont ensuite rapidement intéressé les pays occidentaux. Dans les années 50, un

autre chercheur, E. S. Page jette les bases de l'étude des changements dans les processus stochastiques. Il propose notamment le **CUMulative SUM control chart**, très répandu depuis et connu sous le nom de **CUSUM** [Page, 1954]. Non content de donner une méthode de surveillance et permettre la mise en place d'un système d'alerte, ce processus présente aussi l'avantage d'avoir de bonnes propriétés pour la gestion des fausses alertes. L'étendue des applications peut être appréhendée avec un peu de recherche sur internet : le CUSUM est utilisé en contrôle de qualité dans les processus industriels, dans des études économiques et même dans des études historiques ou sociologiques, dans la lutte contre les attaques informatiques, dans la surveillance du trafic aérien, dans les recherches sur le changement climatique, sans compter des applications militaires.

Dans le domaine de l'assurance, le CUSUM, appliqué aux processus ponctuels, fait l'objet d'une attention particulière depuis quelques années. De façon très synthétique, le but est d'étudier, à partir d'une chronique d'évènements, l'intensité du processus de comptage des ces évènements, et de chercher à savoir si cette intensité change au cours du temps, avec une variation d'un niveau prédéterminé. Par exemple, il s'agit de répondre à une question du type : l'intensité des décès observés sur un portefeuille de passif connaît-elle une hausse de 10%, et le cas échéant, puis-je détecter cette modification le plus vite possible et alerter sur ce changement tout en limitant le nombre de fausses alertes ?

L'objet du présent mémoire est de présenter le processus CUSUM pour les processus de Poisson, à la fois à travers ses bases mathématiques mais aussi avec des éléments plus intuitifs pour mieux appréhender son comportement, de proposer une démarche pratique de mise en œuvre, d'étudier la pertinence de son utilisation dans le cadre du suivi d'un portefeuille d'assurance emprunteur, et de partager avec ceux qui voudraient par la suite se servir de cette méthode quelques réflexions sur les erreurs potentielles dans son utilisation ou les interprétations erronées qui pourraient être faites de ses résultats.

Chapitre 1

Contexte de l'assurance emprunteur

1.1 Définition de l'assurance emprunteur

Selon le glossaire du Comité Consultatif du Secteur Financier (CCSF), l'assurance emprunteur est définie comme une **assurance temporaire souscrite à l'occasion de la mise en place d'un crédit, qui garantit la prise en charge de tout ou partie du crédit en cas de survenance de certains événements (par exemple en cas de décès de l'emprunteur, en cas de perte totale et irréversible d'autonomie (PTIA), invalidité permanente, incapacité temporaire de travail, et éventuellement la perte d'emploi)** [CCSF, 2021b].

Le bilan produit en 2020 par le cabinet Actelior pour le CCSF est plus précis : **le contrat d'assurance emprunteur consiste en la couverture d'une ou plusieurs personnes physiques (assuré, coassuré voire caution) contre un risque potentiel d'insolvabilité lié à un événement (décès, arrêt de travail, invalidité ou, plus rarement perte d'emploi) venant potentiellement réduire leurs capacités de remboursement d'un prêt souscrit auprès d'un organisme prêteur. L'assurance emprunteur représente donc un mécanisme permettant la couverture de l'emprunteur mais aussi de l'organisme prêteur, en général bénéficiaire des garanties** [CCSF, 2021a].

Aucune loi n'impose à l'emprunteur de souscrire ce type d'assurance, cependant les organismes prêteurs conditionnent le plus souvent (et même quasi-systématiquement¹ pour les prêts immobiliers) l'octroi du prêt à cette souscription, et imposent des garanties minimales.

1.2 Types de contrats

L'inspection générale des finances, dans un rapport de 2013 sur l'assurance emprunteur [Inspection Générale des Finances, 2013], distingue trois types de contrats.

Les contrats dits de groupe sont les contrats traditionnellement proposés par les établissements de crédits à leurs emprunteurs. Ils sont conçus et assurés par des assureurs partenaires des établissements de crédit, souvent entités du même groupe que ces derniers. Ce sont des contrats collectifs à adhésion facultative, autrefois à tarif unique pour tous,

1. Dans certains cas, l'emprunteur peut donner par lui-même des garanties suffisantes (via une hypothèque, un nantissement ou encore une caution apportée par un tiers)

avec donc une mutualisation totale du risque, et dont la prime est basée sur le capital emprunté.

Les contrats dits individuels ou encore alternatifs, apparus plus récemment, sont tarifés différemment pour chaque assuré, mais restent au sens légal des contrats collectifs. Plus segmentés que les précédents, leur tarification prend en compte l'écoulement des capitaux restant dus, l'âge de l'assuré, la tabagie, l'état de santé ou la catégorie socio-professionnelle. Ils s'adressent donc principalement à des profils de risque plus faible pour lesquels les contrats de groupe pourraient être sensiblement plus coûteux. Ils sont alternatifs en ce sens qu'ils sont proposés par les banques en alternative à leur propres contrats de groupe.

Les contrats dits individuels purs, ne ressortent plus du cadre légal des contrats collectifs. Comme pour les contrats alternatifs, leur tarification prend en compte l'écoulement des capitaux restants dus, l'âge de l'assuré, la tabagie, l'état de santé ou la catégorie socio-professionnelle. Le rapport produit en 2013 par l'Inspection Générale des Finances note à leur propos : "**Ils représentent une fraction extrêmement marginale de l'ensemble des contrats et n'ont, de ce fait, pas fait l'objet d'une étude spécifique par la mission.**" [[Inspection Générale des Finances, 2013](#)]

Du point de vue du code des assurances, c'est l'article L141-1 (Modifié par Loi n°2006-1770 du 30 décembre 2006 - art. 65 (V) JORF 31 décembre 2006 en vigueur le 1er octobre 2007) qui précise la notion de contrat de groupe : "**Est un contrat d'assurance de groupe le contrat souscrit par une personne morale ou un chef d'entreprise en vue de l'adhésion d'un ensemble de personnes répondant à des conditions définies au contrat, pour la couverture des risques dépendant de la durée de la vie humaine, des risques portant atteinte à l'intégrité physique de la personne ou liés à la maternité, des risques d'incapacité de travail ou d'invalidité ou du risque de chômage. Les adhérents doivent avoir un lien de même nature avec le souscripteur.**" [[Code des Assurances, 2006](#)]

C'est donc bien le fait que l'assuré adhère, ainsi que d'autres assurés liés comme lui au souscripteur, à un contrat dont il n'est pas souscripteur, qui définit le contrat de groupe, et non le fait que ce contrat soit proposé par l'établissement de crédit. Quant aux contrats d'assurance emprunteur qui ne seraient pas des contrats de groupe, la situation décrite par l'IGF en 2013 prévaut toujours : les contrats individuels (un seul souscripteur qui est en général l'assuré) restent très minoritaires. Ils sont cependant très majoritaires chez AFI-ESCA.

De son côté, le bilan du CCSF [[CCSF, 2021a](#)] ne prête pas attention au caractère collectif ou non du contrat, et distingue les **contrats de groupe**, les **contrats alternatifs internes**, proposés par les organismes de crédit en alternative à leur propres contrats de groupe, et les **contrats alternatifs externes**, assurés par une compagnie qui n'est pas partenaire du créancier. Cette façon de considérer les choses est, comme nous allons le voir par la suite, plus appropriée pour la compréhension de l'évolution du marché de l'assurance emprunteur.

1.3 Garanties les plus courantes

Les garanties contre le décès et la perte totale et irréversible d'autonomie (PTIA) sont systématiquement exigées.

Très souvent demandées, les garanties couvrant l'incapacité temporaire de travail (ITT) concernent des périodes pendant lesquelles l'assuré n'est pas en capacité de travailler, périodes généralement inférieures à trois ans, en lien avec les prises en charge de la sécurité sociale. Cependant, les conditions générales ou particulières des polices peuvent ne pas faire référence à cette durée maximale, et il est donc potentiellement possible d'être couvert par l'assurance emprunteur au titre d'une incapacité de travail sur une période plus longue. L'assureur prend en charge le paiement des mensualités des prêts (dans la limite de la quotité assurée), soit de façon forfaitaire -i.e l'assureur paie toute la mensualité -, soit de façon indemnitaire - l'assurance couvre la mensualité du prêt à proportion de la perte de revenu liée à l'ITT -, et parfois exonère l'assuré du paiement des cotisations d'assurance pendant cette période.

Les couvertures contre l'invalidité sont elles aussi fréquemment exigées. L'invalidité permanente totale (IPT) caractérise la situation d'un assuré qui a perdu au moins 66% de sa capacité de travail². La plupart des assureurs prennent alors en charge le paiement des mensualités des prêts, et quelques-uns remboursent le capital restant dû. Même si une rémission est en théorie exclue, l'avantage pour les premiers est que la prise en charge peut être arrêtée lorsque l'assuré arrive à l'âge de la retraite. L'invalidité permanente partielle (IPP) est déclarée lorsque l'assuré a perdu entre 33% et 66% de sa capacité de travail³. Certains produits d'assurance emprunteur prévoient alors la prise en charge par l'assureur de 50% des mensualités (pour la quotité assurée), quand d'autres utilisent un prorata qui prend en compte le taux d'invalidité.

Le contrat d'assurance peut aussi contenir une garantie contre le chômage, mais celle-ci est rarement exigée, et rarement souscrite. Les couvertures proposées sont diverses selon les assureurs, avec en général de fortes limites dans le temps comme dans les prestations.

1.4 Franchises, carences, exclusions, plafonds, limites d'âge

Plusieurs éléments viennent encadrer les risques pris par l'assureur.

En assurance emprunteur, la notion de **franchise**, contrairement à ce qui est courant dans d'autres types d'assurances [[Wikipedia, 2021b](#)] se réfère essentiellement au délai de franchise. Il s'agit d'un délai avant que la prise en charge par l'assureur soit effective. La prise en charge d'une ITT, par exemple, n'a communément lieu que lorsque l'interruption de travail dépasse trois mois. Il s'agit en ce cas d'une franchise de 90 jours. Cette disposition

2. Établi par un médecin expert, ce taux, de 66% est actuellement un standard de place, notamment depuis la parution des critères du CCSF. D'anciens contrats peuvent avoir une définition différente de l'IPT, parfois sans même notion de taux d'invalidité. Pour être précis, il ne s'agit pas seulement de la capacité de travail, le taux peut tenir compte à la fois de l'incapacité professionnelle et de l'incapacité fonctionnelle.

3. Même remarque à propos du taux.

évite à l'assureur la prise en charge et la gestion d'un très grand nombre de petits sinistres.

La franchise peut être **absolue** ou **relative**. Dans le premier cas, l'assuré n'est pas couvert par la garantie pendant la période de franchise, et si l'ITT dure au-delà de trois mois, la prestation ne portera que sur la durée de l'interruption de travail courue depuis la fin de la franchise. Dans le second cas, à l'issue de la période de franchise, la prestation couvrira l'arrêt de travail dès le premier jour.

Le **délai de carence** désigne quant à lui une période pendant laquelle l'assuré n'est pas couvert. Ce délai court à partir de la date d'effet du contrat d'assurance. Si un sinistre intervient pendant cette période, et quelle que soit sa durée ultérieure, il ne sera pas pris en charge par l'assureur. Dans le cas d'une garantie perte d'emploi, un délai de carence de six mois est tout à fait explicable, puisqu'il permet d'éviter une antisélection évidente.

Les **exclusions** sont des clauses du contrat d'assurance qui viennent exclure de la portée des garanties certains cas particuliers.

Au sens des articles L112-4 et L113-1 du code des assurances ([Code des Assurances, 1994] et [Code des Assurances, 1981]), elles doivent être **formelles et limitées** (ce qui signifie qu'elles ne peuvent pas donner lieu à interprétation, doivent se référer à des critères précis et à des hypothèses limitativement énumérées [Daloz Actualités, 2021]), explicites et mentionnées en caractères très apparents.

Les exclusions les plus fréquentes portent sur le suicide de l'assuré au cours des 12 premiers mois de couverture, les faits volontaires de l'assuré ou tout autre personne ayant intérêt à la garantie, les guerres civiles ou étrangères, insurrections, émeutes et autres dès lors que l'assuré y prend une part active (sauf s'il agit dans l'exercice de sa profession), les suites de catastrophes nucléaires, les conséquences de la conduite sous l'emprise de l'alcool.

En ce qui concerne les garanties complémentaires (ITT, IPT, IPP), l'exclusion la plus commune portent sur les affections ostéoarticulaires et périarticulaires de la colonne vertébrale et les affections psychiatriques. Ces exclusions sont résumées par les praticiens sous l'appellation **exclusions dos-psy**, ou encore via le sigle **MNO** pour **maladies non objectivables** (c'est à dire qu'il s'agit de pathologie que les praticiens peuvent difficilement quantifier).

Les **clauses de limitations** ont pour objet de limiter la portée des garanties.

Les assureurs peuvent ainsi, dans les conditions générales des contrats qu'ils proposent, **plafonner** les prestations de certaines garanties. Ceci protège l'assureur contre la réalisation d'un risque très rare (car les contrats portant sur de très gros capitaux sont rares) mais très pénalisant.

Les **limites d'âge** sont liées à deux aspects différents. D'abord un assuré qui est retraité n'a plus à craindre de perte de revenus suite à une incapacité temporaire ou permanente de travail. Les garanties ITT, IPT, et IPP s'arrêtent ainsi généralement à 65 ans ou bien au départ en retraite. D'autre part, au-delà d'un certain âge, les probabilités de décès ou de PTIA sont tellement élevées que l'assurance deviendrait trop chère, parfois plus chère que le montant de l'emprunt lui-même. Une limite d'âge est donc fixée, ce qui permet tout de même à un assuré âgé de souscrire à une assurance qui le couvrira jusqu'à ce qu'il atteigne cette limite.

1.5 Options

Si, par tous les mécanismes évoqués en 1.4, les assureurs peuvent limiter la portée de leur engagement, ils peuvent aussi proposer à leurs assurés d'étendre la couverture de certaines garanties, en contrepartie d'une augmentation de la prime que ceux-ci devront payer, ou de la réduire en échange d'une diminution de la prime, par le biais d'options. Selon les cas et les assureurs, l'assuré peut ainsi choisir de réduire ou d'allonger les franchises, de **racheter des exclusions** (c'est à dire de payer un surplus de prime pour que certaines exclusions ne jouent pas), ou de s'assurer au-delà des limites d'âge (par exemple pour les salariés qui exerceraient une activité professionnelle au-delà des 65 ans auxquels se limiteraient par défaut leurs garanties complémentaires).

1.6 Catégories ministérielles et participation aux bénéfices

Pour la partie décès, l'assurance emprunteur est similaire à un contrat temporaire en cas de décès, et relève de la catégorie **3** (*Contrats individuels d'assurance temporaire décès (y compris groupes ouverts)*) pour les contrats individuels, et de la catégorie **6** (*Contrats collectifs d'assurance en cas de décès*) pour les contrats collectifs (i.e. contrats de groupe).

En ce qui concerne les garanties liées à l'incapacité ou l'invalidité, les catégories ministérielles correspondantes sont : la catégorie **20** (*Dommages corporels (contrats individuels) (y compris garanties accessoires aux contrats d'assurance vie individuels)*) et la catégorie **21** (*Dommages corporels (contrats collectifs) (y compris garanties accessoires aux contrats d'assurance vie collectifs)*).

La prépondérance historique des contrats de groupe dans l'assurance emprunteur n'est peut-être pas anodine. En effet, selon **une ancienne rédaction de l'article A. 331-3** [[Assemblée Nationale, 2012](#)] du code des assurances, les contrats collectifs ne rentraient pas dans les dispositifs de participation aux bénéfices. Cette rédaction a été déclarée illégale par le Conseil d'État le 23 juillet 2012 [[Conseil d'Etat, 2012](#)].

1.7 Le marché

D'après la Fédération Française des Assurances⁴, les cotisations pour l'assurance emprunteur représentent 9,8 milliards d'euros en 2019, avec une progression constante depuis au moins 2005. Au regard du marché de l'assurance, passé de 208 milliards d'euros de cotisations en 2015 à 227 milliards d'euros en 2019, ou même du seul marché de l'assurance de personnes, passé sur la même période de 156 milliards d'euros à 169 milliards d'euros, l'assurance emprunteur est donc un domaine d'ampleur plutôt modeste.

Pour autant, le dynamisme de l'assurance emprunteur est remarquable, et sur ces seules cinq années, la croissance du montant total de cotisations fut de plus de 13,4%, quand le marché de l'assurance de personne progressait de 8,4%, et celui des assurances en général, de 9,2%

Toujours selon la FFA, 70% de ces cotisations concernent une assurance de crédit immobilier, là où les crédits à la consommation contribuent à hauteur de 21% ou 22%. Le

4. Devenue France Assurance pendant la rédaction de ce mémoire.

reste des cotisations est lié à l'assurance de prêts professionnels.

Comme le montre le graphique 1.1, ce sont les cotisations sur les garanties décès qui en constituent la majeure partie.

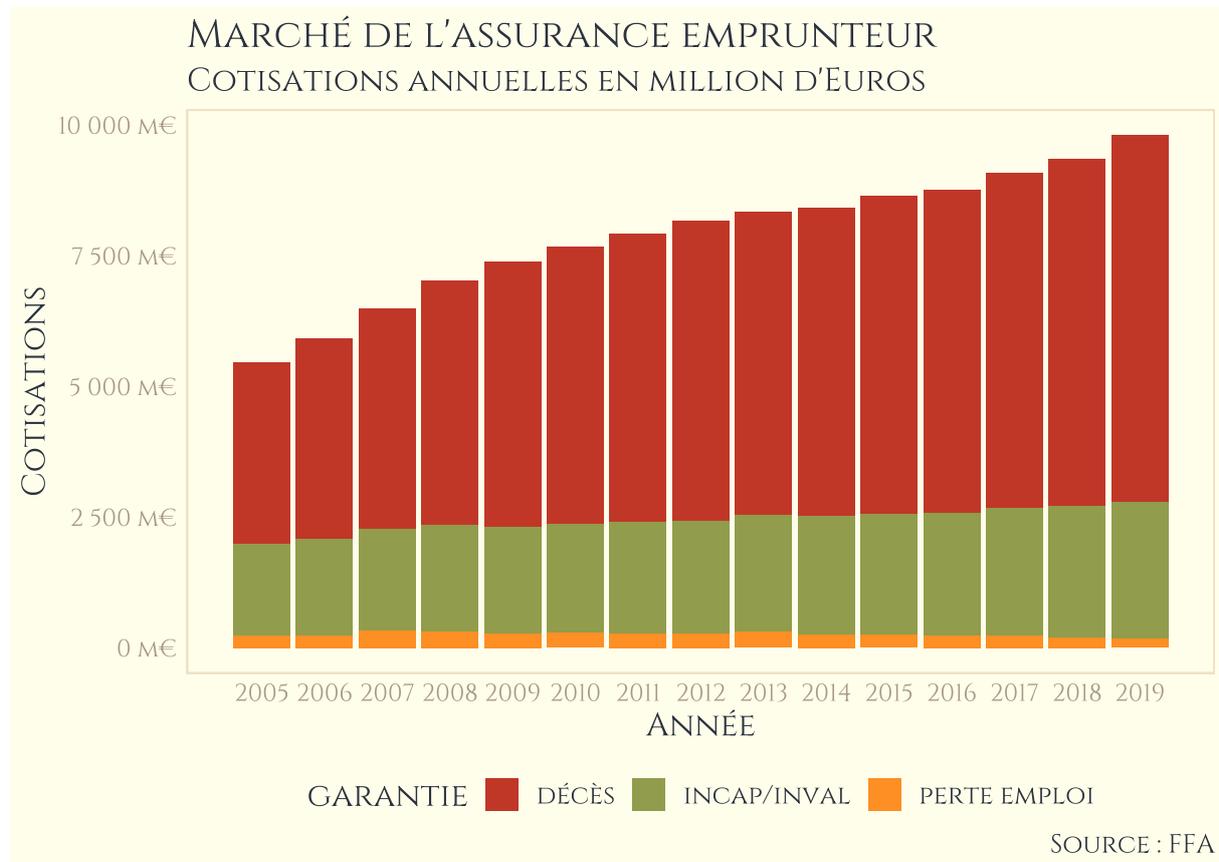


FIGURE 1.1 – Marché de l'assurance emprunteur

1.8 Evolutions

Au cours des deux dernières décennies, plusieurs évolutions notables ont eu lieu.

Comme évoqué plus haut, en 2007 les résultats obtenus par les assureurs via les garanties vies des contrats collectifs d'assurance emprunteur sont entrés dans les comptes de participation. L'évolution ultérieure du marché, présentée au paragraphe précédent, montre que ceci n'a pas eu d'effet sur l'attractivité du marché.

L'assurance emprunteur, exigée par les organismes de crédit, leur a permis de se tailler la part du lion sur ce marché. Que ce soit par simplicité ou par manque d'information, l'emprunteur choisit encore aujourd'hui très souvent l'assurance proposée par la banque de laquelle il obtient son prêt. Ce phénomène a depuis longtemps été désigné comme un problème de distorsion de concurrence, avec pour conséquence des prix qui restent élevés, puisque les créditeurs ne subissaient pratiquement pas d'incitation à la baisse de leurs tarifs dès lors que leur assurance était souscrite presque systématiquement.

L'article L312-9 du code de la consommation, dans sa version en vigueur du 27 Juillet 1993 au 1er Septembre 2010 [[Code de la Consommation, 1993](#)], permettait au prêteur d'imposer la souscription de son contrat groupe.

Loi Lagarde

Un premier coup à ce système a été porté en 2010, par la loi 2010-737 du 1er juillet 2010 portant réforme du crédit à la consommation, dite **loi Lagarde**. Cette loi instaure entre autre la liberté pour l'assuré de choisir un autre assureur que celui promu par le prêteur, dès lors que les garanties sont identiques. Il s'agit alors de ce qui est souvent appelé délégation d'assurance. Cette loi précise aussi que le prêteur ne peut pas revenir sur les conditions d'attribution du prêt en cas de délégation d'assurance, et prévient ainsi la tentation pour le prêteur de récupérer une partie du gain espéré sur l'assurance.

Loi Hamon

Quelques années après l'entrée en vigueur de la loi Lagarde, force était de reconnaître que son objectif n'était pas atteint (voir la figure 1.2). L'idée sous-jacente à la partie de la loi n° 2014-344 du 17 mars 2014 relative à la consommation, dite **Loi Hamon**, qui s'intéresse à l'assureur emprunteur est qu'au moment où les emprunteurs négocient leur prêts, ils n'ont pas nécessairement le temps ou les moyens de se consacrer à l'étude des assurances qu'ils pourraient souscrire, d'où en pratique la permanence de la prévalence des assurances apportées par les prêteurs, malgré la loi Lagarde. La loi Hamon instaure donc la possibilité pour l'emprunteur de pouvoir changer d'assurance emprunteur, sous réserve que le niveau des garanties reste équivalent⁵, pendant les 12 premiers mois qui suivent l'obtention de son prêt, et ceci sans que le prêteur ne puisse à cette occasion modifier les conditions du prêt ou demander des frais supplémentaires.

Amendement Bourquin

Quelques années après la loi Hamon, le marché de l'assurance emprunteur restait encore essentiellement tenu par les prêteurs. La loi sobrement nommée **LOI n° 2017-203 du 21 février 2017 ratifiant les ordonnances n° 2016-301 du 14 mars 2016 relative à la partie législative du code de la consommation et n° 2016-351 du 25 mars 2016 sur les contrats de crédit aux consommateurs relatifs aux biens immobiliers à usage d'habitation et simplifiant le dispositif de mise en œuvre des obligations en matière de conformité et de sécurité des produits et services** entérine ce qui est retenu comme l'amendement Bourquin, et qui permet, depuis le premier janvier 2018, un changement d'assureur au-delà de la première année, mais uniquement aux dates anniversaires des contrats. Ce dispositif a été complété par la suite par une précision sur la date à retenir. Le changement d'assureur ne peut être refusé par le prêteur, sous réserve de l'équivalence des garanties et d'un préavis de deux mois avant la date anniversaire.

Loi Lemoine

Réclamée depuis longtemps par des associations de consommateurs, notamment Que Choisir, une résiliation infra-annuelle a fini par passer les fourches caudines de l'Assemblée Nationale et du Sénat, avec une probabilité presque nulle d'être retoquée par le Conseil Constitutionnel. La **loi n° 2022-270 du 28 février 2022 pour un accès plus juste,**

5. Le CCSF a proposé une liste de critères pour préciser la notion d'équivalence, et ces critères sont reconnus par les acteurs du marché.

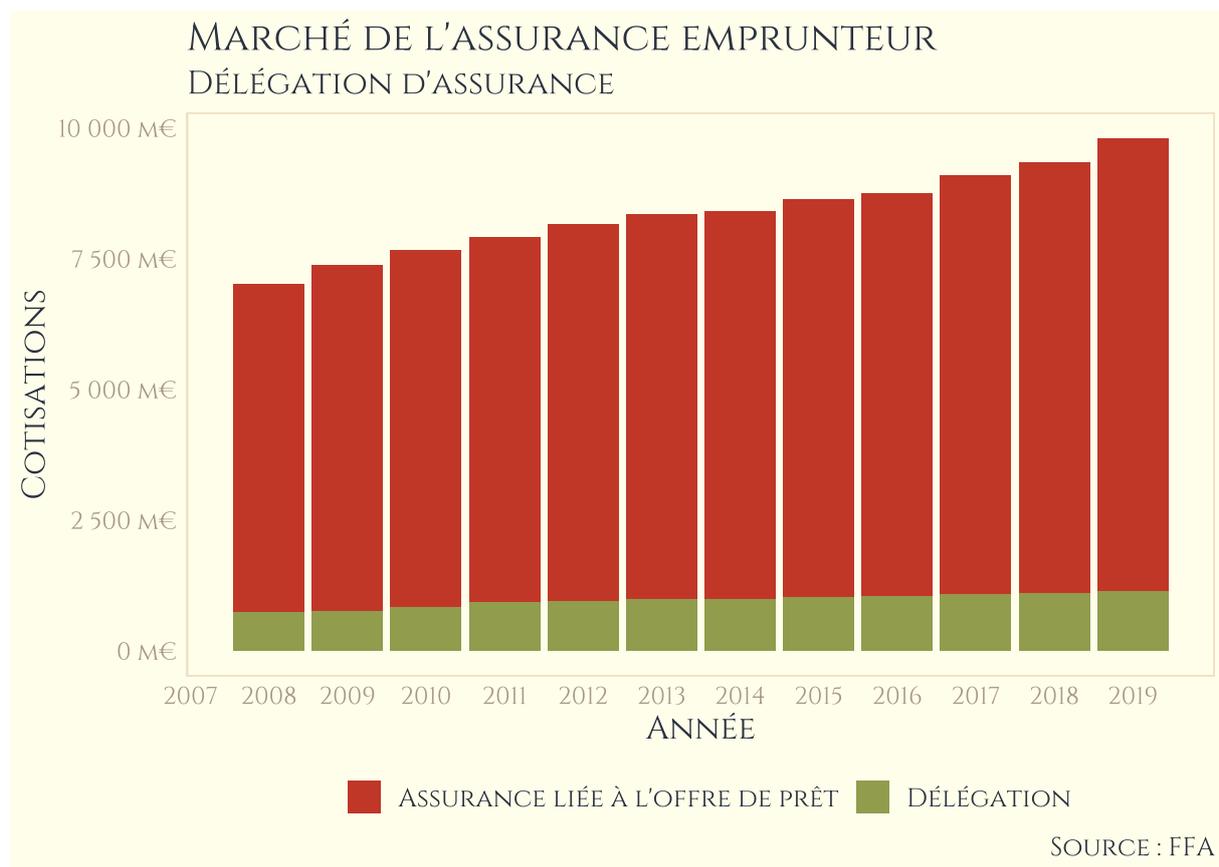


FIGURE 1.2 – Poids de la délégation dans le marché de l'assurance emprunteur

plus simple et plus transparent au marché de l'assurance emprunteur, plus connue sous le diminutif **loi Lemoine**, donne la possibilité d'une résiliation à tout moment.

Si cette évolution, attendue depuis longtemps et qui, à l'instar des précédentes, est supposée favoriser une meilleure concurrence entre les assureurs, la loi Lemoine présente aussi deux autres évolutions importantes.

La première, qui concerne la réduction des délais dans le cadre du droit à l'oubli, devrait de fait faciliter l'emprunt après rémission de certaines catégories de maladies.

L'autre mesure d'importance, qui a éclipsé dans la profession les inquiétudes qu'auraient légitimement pu soulever la résiliation infra-annuelle, porte sur la suppression des questionnaires médicaux pour assurés dont le cumul des montants assurés des prêts immobiliers est inférieur à 200 000 €, si l'âge de l'assuré au terme du remboursement ne dépasse pas 60 ans.

1.9 La guerre des prix

D'un marché très rentable et peu risqué à un marché moins rentable et plus risqué

Les évolutions réglementaires n'ont pas eu, jusqu'à présent, l'effet escompté sur la part de la délégation d'assurance dans l'ensemble du marché de l'assurance emprunteur, même si le rapport du CCSF (cf. [CCSF, 2021a]) mentionne une croissance plus rapide chez les délégataires que dans le reste du marché.

Pour autant, bien que les grands prêteurs aient continué à vendre, à des prix nettement

plus élevés que les délégataires, des assurances pour les prêts qu'ils consentaient, les acteurs alternatifs, s'ils ont parfois tenté de se distinguer par la qualité de la couverture proposée, ont naturellement misé sur des tarifs plus attractifs.

Le rapport du CCSF mentionne, depuis 2010, des baisses de 20% à 41% pour tous les assurés, avec une baisse moyenne de 29% pour les garanties DC/PTIA et ITT-IPT, et de 33% sur les seules garanties décès. Mais ces chiffres sont suivis d'un commentaire qui apporte un éclairage intéressant sur le positionnement des banques vis des évolutions tarifaires : **Notons que cette analyse porte uniquement sur des contrats alternatifs externes, les informations transmises concernant les contrats alternatifs bancaires n'étant pas suffisamment représentatives, elles n'ont pas été reprises dans cette partie de l'étude. Néanmoins, les représentants du secteur bancaire ont précisé que, ces contrats se positionnant en concurrence directe avec les contrats alternatifs externes, leur niveau de prix et leurs évolutions sont comparables au regard du choix des consommateurs se portant sur les uns ou sur les autres [CCSF, 2021a].**

Cette évolution générale des prix, très favorable aux assurés, n'a été possible que parce les tarifs étaient longtemps restés très élevés, et la rentabilité des assurances emprunteurs très forte, pour les assureurs comme pour leurs intermédiaires. Pour mémoire, le rapport de l'Inspection Générale des Finances de 2013 [[Inspection Générale des Finances, 2013](#)] note que [...] **le coût des sinistres représente moins de 50% de la prime payée par les emprunteurs et que les commissions perçues par les distributeurs représentent jusqu'à 44% de cette même prime.** Et, dans le même rapport, un tableau précise, pour les contrats alternatifs, que, en moyenne, 45% des primes servaient à couvrir les sinistres, 32% à la rémunération de l'assureur, et 22% à celle des distributeurs. Pour les contrats bancaires de groupe, ces chiffres étaient respectivement de 40%, 15% et 44%.

Dans un contexte où les prix sont régulièrement revus à la baisse, et où cette tendance est notamment exacerbée par les nouveaux distributeurs que sont les comparateurs en ligne, et même si le marché est parti d'une situation où les contrats étaient extrêmement rentables, cette course aux tarifs les plus bas n'est pas sans danger. Pour l'assurance des prêts immobiliers, notamment, qui porte sur des temps longs (plusieurs années en général, et parfois jusqu'à plus de deux décennies), proposer des assurances avec un tarif très bas peut devenir dangereux. Les comptes de résultats actuels, comme les analyses de type *S/P* qui sont produites sur des portefeuilles déjà un peu anciens reflètent encore en grande partie la rentabilité des prêts souscrits il y a quelques années. Il est donc, encore plus qu'autrefois, nécessaire de se doter d'outils de suivi très fins pour mesurer l'évolution de la sinistralité, et de pouvoir adapter rapidement l'offre si besoin.

Les effets de la loi Lemoine

D'un autre côté, la loi Lemoine, avec la suppression du questionnaire médical, va certainement mettre un terme à la baisse des prix sur la catégorie de prêts concernés par cette suppression. Mais cela ne signifie pas que la baisse de la rentabilité va ralentir, puisque, justement, cette suppression devrait conduire à une hausse importante du nombre et de la charge des sinistres. Face à cette situation, les assureurs sont quelque peu dans l'incertitude : faute d'expérience, nul ne peut finement estimer quel sera le poids réel de cette suppression. Pour autant, afin de mettre sur le marché un produit pour cette catégorie de contrats, il est indispensable d'émettre des hypothèses, et de prendre position. Là encore, un suivi très réactif sera nécessaire, pour détecter au plus vite si les hypothèses sous-jacentes au tarif proposé sont en adéquation avec la sinistralité subie.

D'autre part, les nouvelles possibilités de changement d'assurance à tout moment

pourraient aussi bousculer le marché, et peut-être relancer de plus belle la guerre des prix. Un suivi des résiliations sera tout aussi impératif, pour savoir si l'offre commerciale reste concurrentielle, et si les éventuelles stratégies de défense de portefeuille sont efficaces.

Chapitre 2

Changement d'intensité dans un processus de Poisson

2.1 Processus de Poisson

Processus de Poisson homogène

Définition 1 (Processus de comptage). *Un processus de comptage est une suite de variables aléatoires réelles indexée par le temps, $(N_t)_{t \geq 0}$ telle que :*

- $N_0 = 0$,
- $\forall t \geq 0, N_t \in \mathbb{N}$,
- la fonction $t \mapsto N_t$ est croissante.

Définition 2 (Processus de Poisson). *Un processus de Poisson est un processus de comptage $N = (N_t)_{t \geq 0}$ tel que :*

- Les accroissements du processus N sont indépendants :
 $\forall k \in \mathbb{N}, \forall t_0 \leq t_1 \leq \dots \leq t_k$,
 $N_{t_k} - N_{t_{k-1}}, N_{t_{k-1}} - N_{t_{k-2}}, \dots, N_{t_1} - N_{t_0}$ sont des variables aléatoires indépendantes,
- $\forall (s, t) \in \mathbb{R}_*^+$, $N_{s+t} - N_t$ suit une loi de Poisson de paramètre λt .
 λ est appelé l'intensité du processus.

Les propriétés dans la définition du processus de Poisson garantissent la stationnarité des accroissements du processus, c'est à dire que : $\forall k \in \mathbb{N}, \forall t_0 \leq t_1 \leq \dots \leq t_k$, et $\forall \tau \in \mathbb{R}_*^+$ la variable aléatoire multivariée $(N_{t_k} - N_{t_{k-1}}, N_{t_{k-1}} - N_{t_{k-2}}, \dots, N_{t_1} - N_{t_0})$ est de même loi que $(N_{t_k+\tau} - N_{t_{k-1}+\tau}, N_{t_{k-1}+\tau} - N_{t_{k-2}+\tau}, \dots, N_{t_1+\tau} - N_{t_0+\tau})$.

Dans la définition précédente, l'intensité du processus de Poisson est constante, égale à λ . Cependant, pour la plupart des phénomènes observés, cette hypothèse est trop restrictive. Les évolutions de l'environnement, les épisodes épidémiques ou pandémiques, ou même simplement l'évolution de la taille de la population observée, modifieront certainement le nombre de décès observés sur une population donnée sur un laps de temps donné. Pour une approche globale de ce problème, le cadre des processus de Poisson non homogènes est plus adapté. Dans ce contexte, l'intensité λ n'est plus une constante, mais une fonction du temps $t \mapsto \lambda_t$.

Processus de Poisson non homogène

Définition 3 (Processus de Poisson non homogène). *Un processus de Poisson non homogène est un processus de comptage $N = (N_t)_{t \geq 0}$ tel que :*

- Les accroissements du processus N sont indépendants
- $\forall t \geq 0$,

$$\lim_{h \rightarrow 0^+} \frac{\mathbb{P}[N_{t+h} - N_t \geq 2]}{\mathbb{P}[N_{t+h} - N_t = 1]} = 0,$$

- Il existe une fonction à valeur positive $\lambda : t \mapsto \lambda_t$ telle que $\forall t \geq 0$:

$$\lim_{h \rightarrow 0^+} \frac{1 - \mathbb{P}[N_{t+h} - N_t = 0]}{h} = \lambda_t.$$

Lorsque la fonction λ est constante, nous retrouvons le processus homogène, et puisque N_t est de loi de Poisson de paramètre λt , nous avons pour l'espérance dans le cas homogène :

$$\mathbb{E}(N_t) = \lambda t = \int_0^t \lambda ds.$$

Dans le cas non homogène, la fonction moyenne est définie par :

$$\Lambda_t = \int_0^t \lambda_s ds.$$

Il est alors possible de montrer que $\forall t \geq 0, N_t \sim \mathcal{P}(\Lambda_t)$ et

$$\mathbb{E}(N_t) = \Lambda_t.$$

Une remarque utile dans l'étude des processus stochastiques est qu'il est parfois plus pratique ou plus fructueux d'étudier un autre processus qui caractérise complètement le processus que l'on souhaite étudier. C'est le cas par exemple de la transformée de Laplace du processus de Poisson non homogène qui est donnée par :

$$\mathbb{E} \left[e^{\alpha N_t} \right] = e^{\Lambda_t (e^\alpha - 1)}.$$

Processus de Cox

Dans le cas du processus de Poisson non homogène, la fonction $t \mapsto \lambda_t$ est déterministe. Il existe une extension de ce type de processus au cas où elle est elle aussi un processus aléatoire.

Définition 4 (Processus doublement stochastique, ou processus de Cox). *Un processus de Cox est un processus de Poisson non homogène dont l'intensité est aléatoire*

Ainsi, conditionnellement à une réalisation du processus λ , le processus de Cox est un processus de Poisson non homogène.

Dans ce cas, sous certaines hypothèses (cf. [El Karoui et al., 2017] p.4), Λ_t est appelée \mathbb{P} -intensité cumulative du processus N si $M_t = N_t - \Lambda_t$ est une martingale.

La martingale de Wald, qui est utilisée pour généraliser la transformée de Laplace, s'écrit $\mathcal{E}_t = \exp(\alpha N_t - \Lambda_t (e^\alpha - 1))$. Les auteurs de [El Karoui et al., 2017] préfèrent l'écrire, pour $\eta \geq 0$,

$$\mathcal{E}_t^\eta = \exp(\log(\eta) N_t - (\eta - 1) \Lambda_t) = \eta^{U_t^\eta},$$

où

$$U_t^\eta = N_t - \beta(\eta) \Lambda_t,$$

et

$$\beta(\eta) = \frac{\eta - 1}{\ln(\eta)}.$$

2.2 Poisson disorder Problem

2.2.1 Quel est le problème ?

Un cas particulier de processus de Poisson non homogène est celui où λ_t décrit l'intensité **habituelle** du processus, celle qui est supposée prévaloir dans des conditions connues et courantes, puis, sous l'effet d'un changement soudain, est modifiée, à la hausse ou à la baisse. Autrement dit, le processus observé pourrait être modélisé par la concaténation de deux processus de Poisson qui se succéderaient dans le temps. Cette situation pourrait servir par exemple à modéliser l'effet d'un changement soudain de l'environnement qui conduirait à une augmentation ou une baisse de la mortalité observée. En assurance, il n'est pas rare de modéliser l'occurrence d'évènements liés à un risque par un processus de Poisson, et pour l'assureur, s'assurer que l'intensité du processus ainsi modélisé est bien celle qui prévalait dans ses hypothèses peut s'avérer crucial.

Réagir trop tard à une hausse de sinistralité conduirait à avoir en portefeuille des risques sous tarifés, et serait donc en premier lieu dangereux pour la rentabilité de l'entreprise, et peut-être même, si le choc est plus extrême, pour sa capacité à honorer ses engagements. De manière moins directe, mais peut-être plus immédiate, la hausse de sinistralité une fois constatée, son introduction dans les hypothèses de base des simulations imposées dans le cadre de Solvabilité II pourrait conduire à une baisse non négligeable des taux de couverture des capitaux de solvabilité requis.

A contrario, être le dernier assureur à s'apercevoir que la sinistralité sur un risque donné connaît une baisse importante, ou même n'être seulement pas le premier à s'en rendre compte, risque de conduire l'assureur à voir sa part de marché diminuer, puisque les concurrents mieux avertis auront déjà pu baisser leurs tarifs en conséquence.

Pour autant, si rien ne s'oppose à la séparation d'une série d'évènements observés avec leurs temps de survenance en deux séries, et à la modélisation de ces deux séries par deux processus de Poisson, puis à comparer leurs intensités respectives, comment choisirait-on le moment de la séparation ? De plus, même si tous les temps envisageables pour l'apparition du changement d'intensité étaient testés, le problème d'un changement très récent resterait entier : au-delà de ce temps, les observations pourraient n'être plus assez nombreuses pour que l'estimation du paramètre λ soit suffisamment crédible.

Il faut chercher une autre approche, moins naïve, qui permette de déterminer le temps auquel à lieu le changement d'intensité, et de le détecter aussi, et même surtout, aussitôt que possible.

Change point detection

Il n'est pas inutile de noter que ce problème fait partie d'une catégorie bien plus vaste de problèmes, qui sont ceux qui, plus généralement, traitent des changements abrupts dans les caractéristiques de processus stochastiques, et plus particulièrement de la détection du moment ou un changement s'est produit [Wikipedia, 2021a]. Et, comme le remarquent Michèle Basseville et Igor V. Nikiforov dans leur ouvrage *Detection of Abrupt Changes : Theory and Application* [Basseville and Nikiforov, 1993], les domaines d'applications sont vastes : détection et diagnostics de pannes, maintenance conditionnelle des processus industriels, sécurité des systèmes complexes (avions, bateaux, fusées, centrales nucléaires, procédés technologiques chimiques, etc.), contrôles de qualité, prédiction de catastrophes naturelles, surveillance en biomédecine pour n'en citer que quelques-uns.

Par ailleurs, étant donnée la sensibilité des processus concernés, deux autres points

méritent d'être soulignés. D'abord, les changements dont la détection fait l'objet de ces recherches ne sont pas nécessairement de grande ampleur. De plus, compte tenu des conséquences désastreuses que pourrait avoir une mauvaise décision, que pourrait prendre par exemple le gestionnaire d'une centrale nucléaire, due à une détection erronée d'un changement de régime, les méthodes élaborées au fil des décennies passées ont toujours été assorties d'une attention particulière aux problématiques de fausses alarmes. Sur ce point, le problème du désordre poissonnien ne fait pas exception.

Deux façons d'aborder le problème se distinguent, selon que les observations sur lesquelles porte l'étude soient entièrement disponibles ou bien qu'elles continuent à être observées. Dans le second cas, l'analyse sera menée de façon séquentielle : les nouvelles données disponibles seront intégrées au fur et à mesure dans les procédures de détection.

2.2.2 Formulation mathématique

La présente formulation du problème suit l'article [El Karoui et al., 2017].

Soit un processus de comptage $(N_t)_{t \geq 0}$, et $\tilde{\lambda} = (\tilde{\lambda}_t)_{t \geq 0}$ son intensité.

Le processus de comptage N_t est observable séquentiellement à travers la filtration $\mathcal{F}_t = \sigma\{N_s, 0 \leq s \leq t\}$.

Le processus subit un changement d'intensité de λ_t à $\rho\lambda_t$ à un temps θ inconnu. Autrement dit :

$$\tilde{\lambda}_t = \lambda_t \mathbb{1}_{\{t \leq \theta\}} + \rho\lambda_t \mathbb{1}_{\{t > \theta\}}, \rho \neq 1.$$

L'intensité cumulée du processus avant changement est notée $\Lambda_t = \int_0^t \lambda_s ds$.

Le problème est de trouver une règle pour détecter θ **le plus rapidement possible, avec un nombre limité de fausses alertes.**

Lorsque des hypothèses sont émises sur la distribution de la variable aléatoire θ , le problème entre dans un cadre bayésien (voir des références dans [El Karoui et al., 2017]). Si des résultats sont obtenus dans ce cadre, ils restent très sensibles au choix de la distribution a priori sur θ , et par ailleurs, dans une situation réelle, rares sont les cas où une telle distribution peut être déterminée avec un minimum de pertinence.

Tant que le changement n'a pas eu lieu, le processus $(N_t)_{t \geq 0}$ est considéré sous la mesure de probabilité \mathbb{P} , associée à l'intensité $(\lambda_t)_{t \geq 0}$.

Une probabilité équivalente, \mathbb{P}_θ , décrit le cas où le changement d'intensité a lieu en θ . Ainsi $\mathbb{P}_\infty = \mathbb{P}$ (i.e. le changement n'a jamais lieu).

Le rapport de la vraisemblance des observations jusqu'au temps T sous \mathbb{P}_θ et de celle sous \mathbb{P} est noté $\mathcal{E}_{0,T}^\theta$. Ainsi pour tout $t < \theta$, $\mathcal{E}_{0,t}^\theta = 1$.

2.2.3 Critère de Lorden

Comme souvent, une proposition en vue d'obtenir une amélioration ne peut être discutée que si l'amélioration elle-même est bien définie. S'il fallait par exemple admettre que la meilleure procédure de détection du changement d'intensité était celle qui détectait un changement le plus tôt possible, sans autres considérations, il suffirait de fournir la procédure qui stipule que, quoi qu'il arrive, le changement est détecté au temps $t = 0$. Cette procédure aurait le mérite d'être la meilleure au sens précédent, et en outre serait extrêmement simple à mettre en place, et n'en serait pas moins complètement inutile en pratique, puisque qu'elle fournirait à tout instant une fausse alerte. Il faut donc préciser ce qui est attendu en terme de détection, en prenant en compte à la fois les temps de détection et le nombre probable de fausses alertes.

L'idée de Lorden

Gary Lorden a proposé en 1971 [Lorden, 1971], dans le cadre d'une détection robuste (i.e. sans distribution à priori sur θ), un critère qui prend en compte :

- le délai moyen entre le changement en θ et la détection à un temps τ :

$$\mathbb{E} [(\tau - \theta)^+ | \mathcal{F}_\theta],$$

- le temps moyen avant une fausse alerte (espérance sous \mathbb{P}_∞) :

$$\mathbb{E}_\infty [\tau].$$

L'idée de Lorden est de trouver une règle qui minimisera le délai avant la détection dans le pire des cas possible, sous la contrainte d'avoir un temps assez long, s'il n'y a pas de changement, avant une fausse alarme. Le pire cas possible est à comprendre comme le pire délai sur tous les θ possibles, d'où le pire retard de détection :

$$C_{lor}(\tau) = \sup_{\theta \in [0, \infty]} \text{ess sup}_\omega \mathbb{E} [(\tau - \theta)^+ | \mathcal{F}_\theta].$$

Où le supremum essentiel ess sup désigne la plus petite valeur (éventuellement infinie) telle que $\mathbb{E} [(\tau - \theta)^+ | \mathcal{F}_\theta]$ ne peut dépasser cette valeur sauf sur un ensemble de probabilité nulle.

Si π est le temps minimal sans fausse alarme requis lorsqu'il n'y a pas de changement, le problème à résoudre est donc de minimiser le pire délai de détection sous cette contrainte, c'est à dire trouver la règle qui détecte un changement au temps τ solution de

$$\inf_T \{C_{lor}(T) | \mathbb{E}_\infty [T] \geq \pi\}.$$

Une version modifiée du critère de Lorden

L'approche précédente a notamment du sens si tous les intervalles de temps sont susceptibles d'accueillir le même nombre d'évènements observés, et donc en particulier les cas où l'intensité du processus avant changement est constante. Elle est peut-être un peu moins adaptée pour des processus non homogènes, voire des processus où l'intensité serait elle-même aléatoire. C'est pourquoi, les auteurs de [El Karoui et al., 2017] proposent, comme ceux de [Gandy et al., 2010] dans le même esprit, de s'attacher au nombre moyen d'évènements entre le changement et sa détection, et au nombre moyen d'évènements avant une fausse alarme :

- Le nombre moyen d'évènements entre le changement en θ et la détection à un temps τ :

$$\mathbb{E} [(N_\tau - N_\theta)^+ | \mathcal{F}_\theta],$$

- Nombre moyen d'évènements avant une fausse alarme :

$$\mathbb{E}_\infty [N_\tau].$$

Le pire cas s'écrira alors :

$$C(\tau) = \sup_{\theta \in [0, \infty]} \text{ess sup}_\omega \mathbb{E} [(N_\tau - N_\theta)^+ | \mathcal{F}_\theta].$$

Ce critère a le mérite d'être indifférent aux procédures de changement de temps. Et le problème à résoudre est ici de trouver τ solution de

$$\inf_T \{C(T) | \mathbb{E}_\infty [N_T] \geq \pi\}.$$

Où π désigne cette fois un nombre d'évènements.

Chapitre 3

Algorithmes de détections des ruptures

Comme évoqué plus haut, le problème du désordre poissonnien est un cas particulier de la thématique plus générale de la détection des ruptures (Change Point Detection). Avant d'aborder la réponse proposée à ce problème particulier dans [El Karoui et al., 2017], il n'est pas inutile de donner quelques idées générales habituellement utilisées dans ce cadre plus large. La suite de ce chapitre s'inspire notamment du chapitre *Change Detection Algorithms* de [Basseville and Nikiforov, 1993]. Il faut prendre garde qu'ici θ désigne un paramètre des lois de probabilités en jeu, et non, comme précédemment, le temps auquel a lieu le changement.

Considérons une séquence de variables aléatoires indépendantes $(y_k)_k$ dotées d'une densité de probabilité $p_\theta(y)$ dépendant d'un seul paramètre θ . Avant un temps inconnu t_0 auquel a lieu la rupture, le paramètre θ est égal à θ_0 , et après t_0 , il est égal à $\theta_1 \neq \theta_0$. Le paramètre θ_0 est supposé connu, et le but est d'établir un algorithme séquentiel qui détecte la rupture.

L'outil principal utilisé dans ces algorithmes est le logarithme du rapport de vraisemblance, défini par :

$$s(y) = \ln \frac{p_{\theta_1}(y)}{p_{\theta_0}(y)}.$$

La propriété importante de cette quantité est la suivante : soient \mathbb{E}_{θ_0} et \mathbb{E}_{θ_1} les espérances sous les probabilités de densités respectives p_{θ_0} et p_{θ_1} , alors :

$$\mathbb{E}_{\theta_0}(s) < 0 \text{ et } \mathbb{E}_{\theta_1}(s) > 0.$$

En d'autres termes, un changement du paramètre θ peut être repéré par un changement de signe de la moyenne du logarithme du rapport de vraisemblance. La figure 3.1 présente le comportement typique de l'évolution du ratio de vraisemblance, au fur et à mesure que de nouvelles observations arrivent. Ici, la moyenne est le paramètre qui change. Elle vaut $\theta_0 = 1,5$ jusqu'à la 500^{ème} observation, puis $\theta_1 = 1,8$ ensuite. Au-delà de la 500^{ème} observation, la vraisemblance de θ_1 devient plus importante, et le logarithme du rapport de vraisemblance commence à croître. Il est intéressant de noter que le processus représenté n'atteint pas 0 après 500 réalisations de la variable aléatoire de moyenne θ_1 . Ceci permet de supputer que pour détecter un changement dans une série d'observations, ce n'est pas vraiment le niveau atteint par le processus des rapports de vraisemblance qui sera le plus utile. L'écart entre le point le plus bas atteint jusqu'à présent et le point présent pourrait être un meilleur sujet d'étude. Autrement dit, si le processus montre une tendance

croissante, plus longtemps dure cette tendance, plus nous devrions être convaincus que le changement a eu lieu.

Par ailleurs, l'information de Kullback est définie par $K(\theta_1, \theta_0) = \mathbb{E}_{\theta_1}(s)$, et

$$\mathbb{E}_{\theta_1}(s) - \mathbb{E}_{\theta_0}(s) = K(\theta_1, \theta_0) + K(\theta_0, \theta_1) > 0.$$

Et la détectabilité d'un changement sur θ peut aussi être étudiée à partir de l'information de Kullback entre les deux modèles, avant et après le changement.

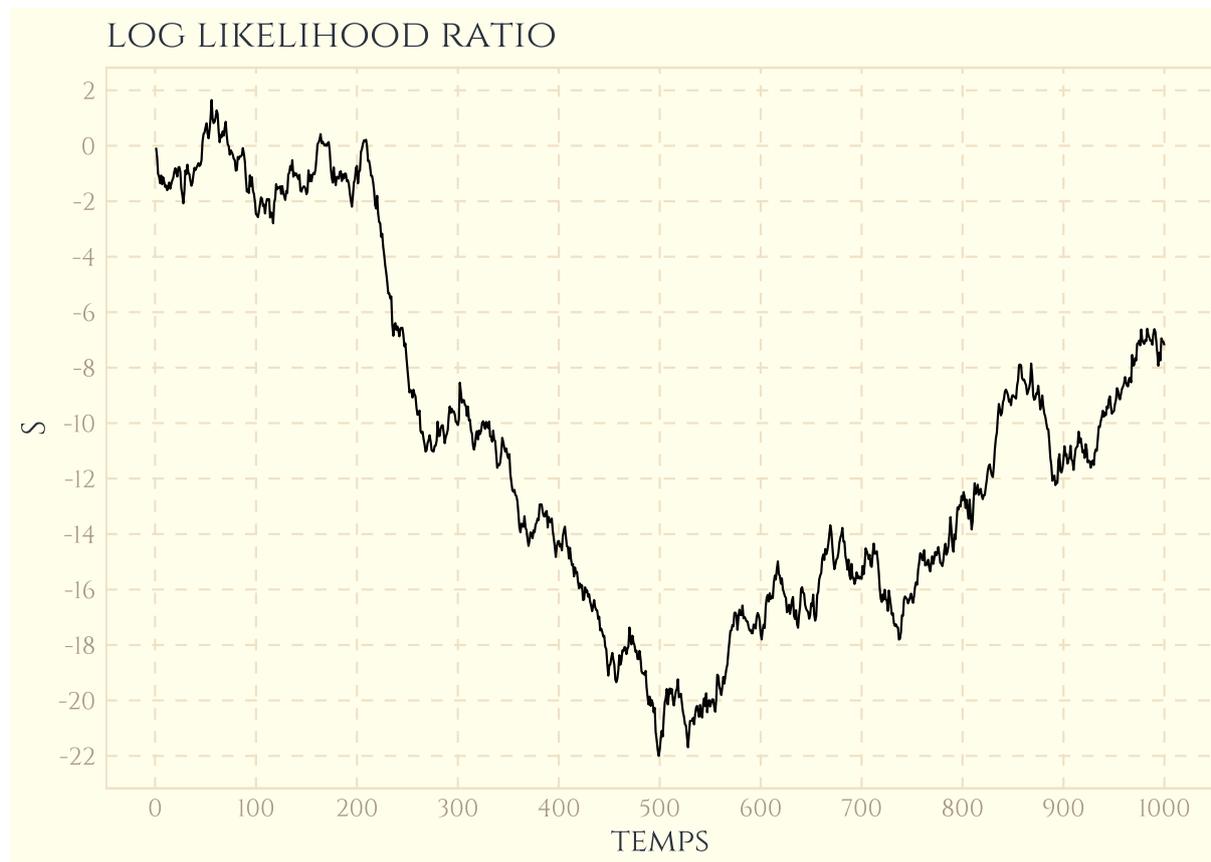


FIGURE 3.1 – Processus séquentiel du logarithme du rapport de vraisemblance (LPSR pour *Log Sequential Probability Ratio*) dans le cas d'une suite de variables aléatoires gaussiennes de moyenne 1,5 jusqu'à la 500^{ème} occurrence, et de moyenne 1,8 ensuite (ici le temps est le nombre d'observations)

3.1 Détecteurs de dépassement de limites et cartes de contrôle de Shewhart

L'idée de base du contrôle de qualité continu est la suivante : un échantillon de taille N est sélectionné, puis un test statistique est appliqué à cet échantillon pour choisir l'une des deux hypothèses suivantes :

$$H_0 : \theta = \theta_0,$$

$$H_1 : \theta = \theta_1.$$

Tant que le test privilégie l'hypothèse H_0 , l'échantillonnage est reconduit et suivi du test. La procédure s'arrête à la première occurrence de la décision en faveur de H_1 . Notons alors :

$$S_j^k = \sum_{i=j}^k s_i,$$

avec

$$s_i = \ln \frac{p_{\theta_1}(y_i)}{p_{\theta_0}(y_i)}$$

le logarithme du rapport de vraisemblance pour les observations de y_j à y_k .

Une conséquence du lemme de Neyman-Pearson (voir [Basseville and Nikiforov, 1993]) est que la règle de décision optimale d est donnée par

$$d = \begin{cases} 0 & \text{si } S_1^N < h; H_0 \text{ est choisie,} \\ 1 & \text{si } S_1^N \geq h; H_1 \text{ est choisie.} \end{cases}$$

où h est un seuil convenablement choisi. La somme S_1^N est appelée fonction de décision. La décision est prise à l'aide d'une règle d'arrêt (*stopping rule*), qui dans ce cas est définie par :

$$t_a = N.\min\{K : d_K = 1\}.$$

où d_K est la règle de décision associée à l'échantillon K (de taille N), et t_a le temps de l'alarme (défini ici en termes de nombre d'observations).

L'application de ces principes de contrôle continu au cas gaussien, de moyenne μ_0 et de variance σ^2 , où le paramètre θ est la moyenne, correspond aux cartes de contrôle de Shewhart. Ces *cartes* sont en réalité des graphiques où sont reportées les valeurs successives de la moyenne empirique

$$\bar{y}(K) = \frac{1}{N} \sum_{i=N(K-1)+1}^{NK} y_i,$$

et la ligne qui matérialise la limite supérieure

$$\mu_0 + \kappa \frac{\sigma}{\sqrt{N}},$$

où κ et N sont des paramètres à choisir convenablement. Une alarme est lancée dès qu'un point $\bar{y}(K)$ est placé au-dessus de la limite supérieure. L'idée se généralise à une borne inférieure, et, pour éviter un trop grand nombre de fausses alarmes, certains auteurs suggèrent de ne déclencher d'alarme qu'après un certain nombre de franchissements.

La figure 3.2 présente la carte de Shewhart pour les mêmes données que dans la figure 3.1. Le changement est détecté un peu avant la 600^{ème} observation. L'exemple montre aussi l'occurrence d'une fausse alerte à la baisse avant le changement réel. Les éléments pour la détermination des seuils d'alerte (lignes pointillées rouges) peuvent s'envisager avec des idées similaires à celles discutées au chapitre 4.



FIGURE 3.2 – Carte de Shewart

3.2 Cartes de contrôle à moyenne mobile géométrique

L'idée de l'algorithme à moyenne mobile géométrique (GMA algorithm pour *geometric moving average algorithm*) est d'utiliser le comportement du logarithme du rapport de vraisemblance, tout en donnant plus de poids aux observations les plus récentes. La fonction de décision peut alors prendre la forme :

$$g_k = \sum_{i=0}^{\infty} \gamma_i \ln \frac{p_{\theta_1}(y_{k-i})}{p_{\theta_0}(y_{k-i})},$$

$$= \sum_{i=0}^{\infty} \gamma_i s_{k-i}.$$

où les poids γ_i sont exponentiels, c'est à dire de la forme $\gamma_i = \alpha(1 - \alpha)^i$, $0 < \alpha \leq 1$, et où l'on a posé $y_{k-i} = 0$ si $k_i < 0$. Le coefficient α agit comme un facteur d'oubli. La fonction de décision peut être réécrite sous forme récursive :

$$g_k = (1 - \alpha)g_{k-1} + \alpha s_k, \text{ avec } g_0 = 0.$$

Le temps de l'alarme est donné par la règle d'arrêt suivante :

$$t_a = \min\{k : g_k \geq h\},$$

où h est un seuil convenablement choisi.



FIGURE 3.3 – Carte de contrôle à moyenne mobile géométrique

3.3 Cartes de contrôle à moyenne mobile finie

Le principe est ici très similaire à celui du GMA, à cette exception près que le nombre d'observations utilisées dans la fonction de décision est limité, et que les poids ne sont plus supposés être exponentiels. Pour le reste, l'algorithme de décision ainsi obtenu (FMA control charts, pour *Finite Moving Average control charts*) est identique au précédent :

$$g_k = \sum_{i=0}^{N-1} \gamma_i \ln \frac{p_{\theta_1}(y_{k-i})}{p_{\theta_0}(y_{k-i})},$$

où les poids γ_i sont n'importe quels poids d'un filtre causal¹, et la règle de décision est toujours :

$$t_a = \min\{k : g_k \geq h\}.$$

1. C'est à dire un filtre tel que pour toute entrée e nulle pour tout $t < T$, la réponse s est nulle pour tout $t < T$. Autrement dit, les observations futures n'interviennent pas dans la réponse.



FIGURE 3.4 – Carte de contrôle à moyenne mobile finie

3.4 Algorithmes à dérivée filtrée

Dans le cas d'un changement de moyenne pour une séquence de variables aléatoires gaussiennes, les algorithmes à dérivées filtrées (FDA pour *Filtered Derivative Algorithms*), tirent parti du fait que, si il n'y a pas de bruit sur la séquence, la modification de la moyenne se reflétera par une grande valeur absolue de la dérivée discrète dans la séquence des logarithmes des rapports de vraisemblance. La sensibilité de la dérivée discrète à la présence de bruit dans les observations impose cependant l'utilisation d'un filtre. De plus, en partie à cause de ce lissage, plusieurs alarmes pourraient être déclenchées au voisinage de t_0 . Pour plus de robustesse dans la détection, il est par exemple préférable de déclencher une alarme après un certain nombre η de dépassements du seuil h convenablement choisi. Nous aurons de nouveau :

$$g_k = \sum_{i=0}^{N-1} \gamma_i \ln \frac{p_{\theta_1}(y_{k-i})}{p_{\theta_0}(y_{k-i})},$$

où les poids γ_i sont n'importe quels poids d'un filtre causal. La dérivée discrète est alors :

$$\nabla g_k = g_k - g_{k-1}.$$

et la règle de décision est donnée par

$$t_a = \min\left\{k : \sum_{i=0}^{N-1} \mathbb{1}_{\{\nabla g_{k-i} \geq h\}} \geq \eta\right\}.$$

La figure 3.5 présente un exemple du graphique obtenu avec un filtre avec trois poids constants $\delta_1 = \delta_2 = \delta_3 = 1$. Ici, l'utilisation du paramètre η pour ne déclencher une alerte

qu'après plusieurs dépassements du seuil h permettrait d'éviter une fausse alerte avant que le changement ait eu lieu.

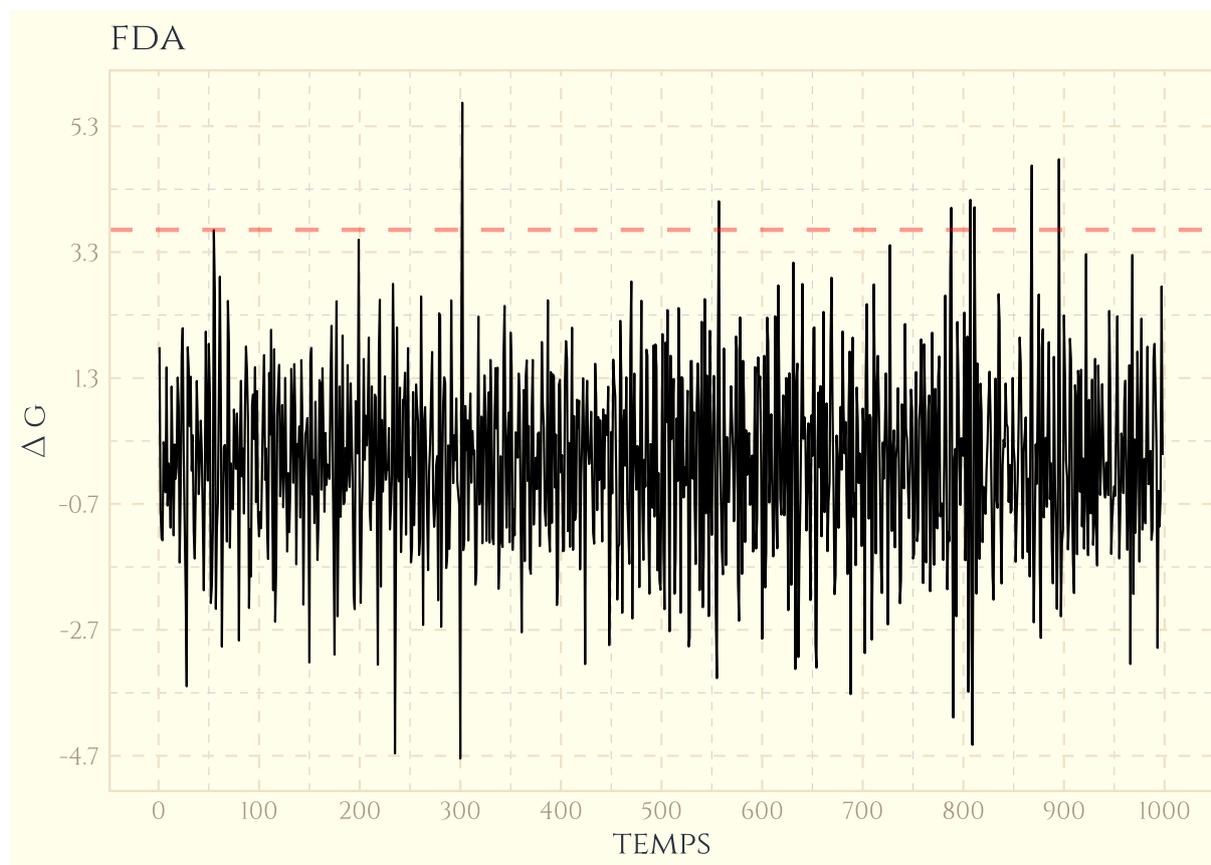


FIGURE 3.5 – Algorithme à dérivée filtrée

3.5 Algorithme CUSUM

3.5.1 Principe élémentaire

Comme mentionné précédemment, le comportement typique du logarithme du rapport de vraisemblance S_k montre une tendance décroissante avant le changement, et une tendance croissante ensuite. L'information pertinente se trouve donc dans la différence entre la valeur du logarithme du rapport de vraisemblance et le minimum qu'il ait pris jusqu'alors :

$$g_k = S_k - m_k \geq h,$$

avec

$$S_k = \sum_{i=1}^k s_i,$$

$$s_i = \ln \frac{p_{\theta_1}(y_i)}{p_{\theta_0}(y_i)},$$

et

$$m_k = \min_{1 \leq j \leq k} S_j.$$

La règle de décision est alors donnée par :

$$t_a = \min\{k : g_k \geq h\},$$

qui s'écrit aussi

$$t_a = \min\{k : S_k \geq m_k + h\}.$$

Il apparaît donc que cette règle de décision est simplement une comparaison entre les sommes cumulées S_k et un seuil adaptatif $m_k + h$. Par le biais de m_k , le seuil n'est pas seulement potentiellement modifié à chaque nouvelle observation, mais il conserve la mémoire de l'information contenue dans les observations passées.



FIGURE 3.6 – Le processus séquentiel du logarithme des rapports de vraisemblance (LPSR) S_k en rouge et le processus CUSUM $S_k - m_k$ en bleu

3.5.2 L'algorithme CUSUM comme répétition du test séquentiel du rapport de probabilité

Page [Page, 1954] a suggéré l'utilisation répétée du test à deux hypothèses :

$$H_0 : \theta = \theta_0,$$

$$H_1 : \theta = \theta_1.$$

à l'aide du test séquentiel de rapport de probabilité (SPRT pour *sequential probability ratio test*). Le SPRT est défini par le couple (d, T) où d est la règle de décision et T le

temps d'arrêt auquel la décision est prise de façon similaire à la définition de la règle de Neyman-Pearson à l'aide de la règle de décision d :

$$d = \begin{cases} 0 & \text{si } S_1^T \leq -\varepsilon, \\ 1 & \text{si } S_1^T \geq h. \end{cases}$$

où T est le temps de sortie :

$$T = T_{-\varepsilon, h} = \min\{k : (S_1^k \geq h) \cup (S_1^k \leq -\varepsilon)\},$$

et $\varepsilon \geq 0$ et $h \geq 0$ sont des seuils convenablement choisis.

La procédure est donc la suivante :

- accepter H_0 si $S_1^T \leq -\varepsilon$,
- accepter H_1 si $S_1^T \geq h$,
- continuer à observer et à tester si $-\varepsilon < S_1^T < h$.

Ce que suggère Page est de réitérer cette procédure chaque fois que H_0 est décidée. A la suite de l'intuition de Page, il a été montré que la valeur optimale pour ε est 0 (voir références dans [Basseville and Nikiforov, 1993]). En ce cas, la règle de décision du test répété peut être exprimée de façon récursive :

$$g_k = \begin{cases} g_{k-1} + \ln \frac{p_{\theta_1}(y_k)}{p_{\theta_0}(y_k)} & \text{si } g_{k-1} + \ln \frac{p_{\theta_1}(y_k)}{p_{\theta_0}(y_k)} > 0, \\ 0 & \text{si } g_{k-1} + \ln \frac{p_{\theta_1}(y_k)}{p_{\theta_0}(y_k)} \leq 0. \end{cases}$$

avec $g_0 = 0$. Ceci peut encore s'écrire de façon condensée : $g_k = (g_{k-1} + s_k)^+$ La règle d'arrêt et le temps d'alarme sont alors :

$$t_a = \min\{k : g_k \geq h\}.$$

Il est aussi possible d'écrire, avec les notations précédentes :

$$g_k = (S_{k-N_k+1}^k)^+,$$

où

$$N_k = N_{k-1} \cdot \mathbb{1}_{g_{k-1} > 0} + 1.$$

Ici N_k est le nombre d'observations depuis le redémarrage du SPRT. Et g_k correspond aussi à l'observation des données via une fenêtre glissante de taille aléatoire.

3.5.3 CUSUM pour un processus de Poisson

Dans cette partie sont reprises les notations du chapitre 2.

LPSR pour le processus de Poisson

Pour $n + 1$ sauts du processus observés aux temps t_0, \dots, t_n , et d'amplitudes $k_{t_i} = N_{t_i} - N_{t_{i-1}}$, en notant $\delta_i = \Lambda_{t_{i-1}} - \Lambda_{t_i}$ nous avons sous \mathbb{P}_∞ ,

$$N_{t_{i-1}} - N_{t_i} = k_{t_i} \sim \mathcal{P}(\delta_i),$$

et sous \mathbb{P}_θ :

$$N_{t_{i-1}} - N_{t_i} = k_{t_i} \sim \mathcal{P}(\rho\delta_i).$$

Le logarithme du rapport des vraisemblances sous \mathbb{P}_θ et \mathbb{P}_∞ , noté U , avec l'information disponible en t_n s'écrira donc :

$$\begin{aligned} U_{t_n} &= \ln \left\{ \frac{\prod_{i=1}^n \frac{(\rho\delta_{t_i})^{k_{t_i}}}{k_{t_i}!} e^{-\rho\delta_{t_i}}}{\prod_{i=1}^n \frac{\delta_{t_i}^{k_{t_i}}}{k_{t_i}!} e^{-\delta_{t_i}}} \right\} \\ &= \ln \left(\prod_{i=1}^n \rho^{k_{t_i}} e^{\delta_{t_i}(1-\rho)} \right) \\ &= \sum_{i=1}^{i=n} \ln(\rho)k_{t_i} + (1 - \rho)\delta_{t_i}. \end{aligned}$$

Et puisque $\sum_{i=1}^{i=n} k_{t_i} = N_{t_n}$ et $\sum_{i=1}^{i=n} \delta_{t_i} = \Lambda_{t_n}$, il vient donc :

$$U_{t_n} = \ln(\rho)N_{t_n} + (1 - \rho)\Lambda_{t_n}.$$

Pour un temps t tel que $t > t_n$, et pour lequel aucun évènement n'est survenu entre t_n et t , sous \mathbb{P}_∞ :

$$\mathbb{P}_\infty(N_t - N_{t_n} = 0) = e^{\Lambda_{t_n} - \Lambda_t},$$

et sous \mathbb{P}_θ :

$$\mathbb{P}_\theta(N_t - N_{t_n} = 0) = e^{\rho\Lambda_{t_n} - \rho\Lambda_t}.$$

D'où

$$U_t = \ln(\rho)N_{t_n} + (1 - \rho)\Lambda_t,$$

et puisqu'il n'y a pas eu d'évènement entre t_n et t , $N_t = N_{t_n}$, et nous pouvons écrire pour tout $t > 0$

$$U_t = \ln(\rho)N_t + (1 - \rho)\Lambda_t.$$

Notation pratique

A un changement d'échelle près, il est assez courant d'écrire

$$U_t = N_t + \frac{(1 - \rho)}{\ln(\rho)}\Lambda_t.$$

Ceci permet d'avoir un processus affine de pente proportionnelle à $(1 - \rho)/\ln(\rho)$ en dehors des sauts, et avec des sauts d'une unité. Le comportement est pour le reste identique, et, sous réserve de s'assurer de la cohérence avec cette notation, l'utilisation qui peut être faite de ce processus est identique. C'est la notation qui sera retenue dans la suite.

CUSUM pour le processus de Poisson

Cas où $\rho > 1$

Dans ce cas, la tendance de U devrait être à la baisse si $t < \theta$ et à la hausse sinon. Il convient donc de s'intéresser au processus CUSUM suivant :

$$V_t = U_t - \min_{s \leq t} \{U_s\}.$$

Cas où $\rho < 1$

Dans ce cas, il s'agit de détecter une baisse de l'intensité du processus. Plutôt que de se baser sur U , la méthode de détection utilisera X défini par :

$$X_t = -U_t,$$

et le processus CUSUM Y donné par :

$$Y_t = X_t - \min_{s \leq t} \{X_s\}.$$

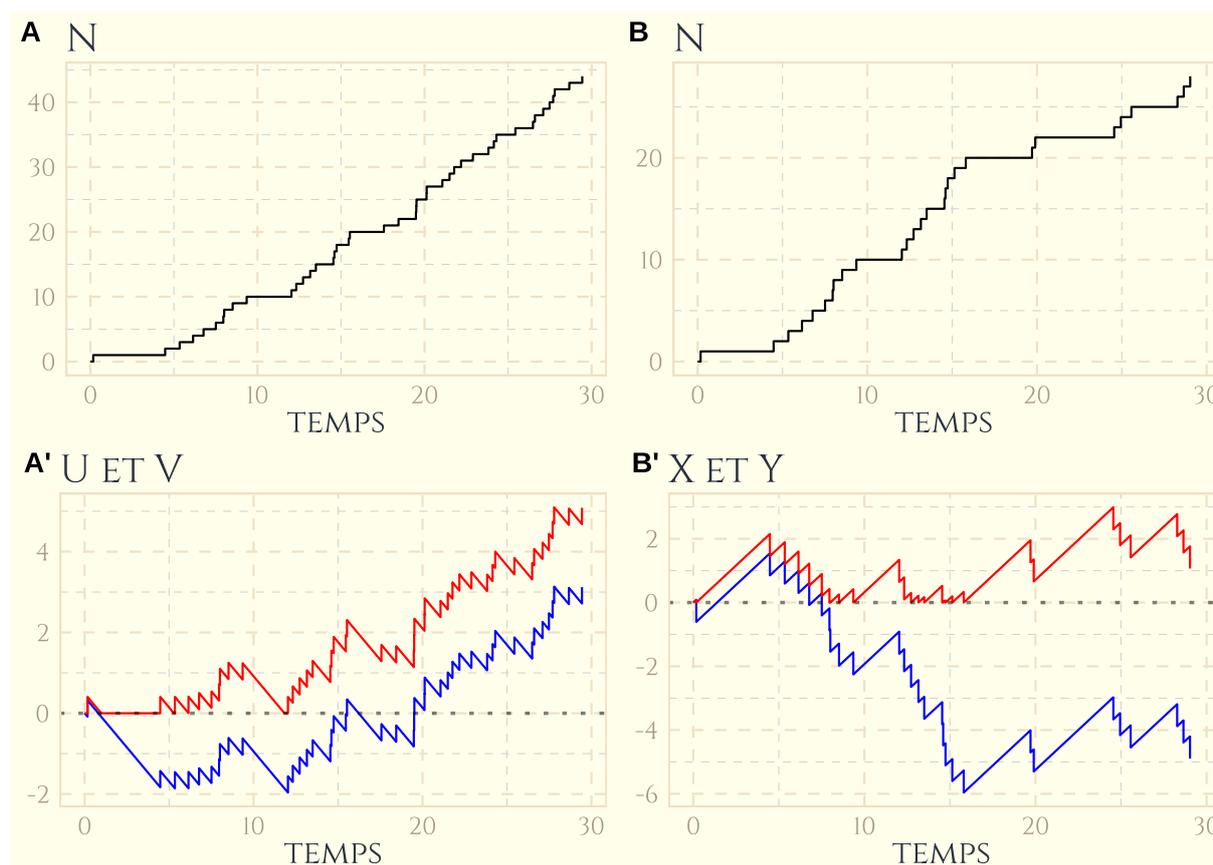


FIGURE 3.7 – A : Processus de Poisson d'intensité 1 jusqu'au temps 15 et d'intensité 1,5 ensuite, A' : processus U (en bleu) et V (en rouge) associés ; B : Processus de Poisson d'intensité 1 jusqu'au temps 15 et d'intensité 0,5 ensuite, B' : processus X (en bleu) et Y (en rouge) associés

3.5.4 Optimalité du processus CUSUM

Il a été montré (voir références dans l'introduction de [El Karoui et al., 2017]), dans le cas des processus de Wiener en particulier, que la décision à l'aide du processus CUSUM est optimale pour le critère exposé en 2.2.3 :

$$\inf_T \{C_{lor}(T) | E[T] \geq \pi\}.$$

Pour les processus de Poisson, le problème a été étudié dans le cadre bayésien (c'est à dire dans le cas où une distribution à priori est donnée pour le temps auquel intervient le changement).

C'est dans [El Karoui et al., 2017] que se trouve, dans un cadre non bayésien cette fois, la démonstration de l'optimalité de la décision basée sur le processus CUSUM dans le cas des processus de Poisson, pour le critère :

$$\inf_T \{C(\tau) | \mathbb{E}[N_\tau] \geq \pi\},$$

où

$$C(\tau) = \sup_{\theta \in [0, \infty]} \operatorname{ess\,sup}_\omega \mathbb{E}[(N_\tau - N_\theta)^+ | \mathcal{F}_\theta].$$

Chapitre 4

Mise en œuvre pratique

La présentation faite jusqu'à présent ne dit pas grand-chose de la mise en œuvre pratique des méthodes de détections, en particulier dans le cas poissonnien. Le présent chapitre vise à donner quelques éclaircissements à ce sujet.

4.1 L'intensité avant le changement

Pour toutes les méthodes présentées au chapitre 3, il est nécessaire de connaître le paramètre θ_0 .

Dans l'exposé du problème du désordre poissonnien (cf. 2.2), il est question d'un processus avec une intensité initiale $t \mapsto \lambda_t$, et de la détection d'un changement qui aurait pu altérer cette intensité. Dans ce cas, il est nécessaire de connaître cette intensité initiale avant de pouvoir chercher à mettre en œuvre une de ces méthodes.

Lorsque l'on connaît à priori, ou que l'on peut retrouver à partir d'hypothèses connues ou supposées telles, l'intensité de base, la question est de fait résolue. Faute de quoi, une étude préalable est requise.

Pour étudier les survenances d'un type de sinistres, il est indispensable de prendre en compte les éléments qui influencent ce risque. Par exemple, dans le cadre de l'étude de l'intensité sous-jacente à une série de décès observés sur une période donnée, nul ne songerait à ignorer l'exposition, l'âge des individus observés, leur sexe, ou encore leur propension à la tabagie. L'approche par la régression poissonnienne, sur laquelle suivent quelques détails, n'est qu'une approche possible.

4.1.1 Régression de Poisson

Lorsque l'on prend pour hypothèse que, pour un jeu de modalités des variables disponibles, le risque est constant, un modèle poissonnien peut convenir. Il suppose justement que l'intensité du processus, conditionnellement à ces covariables, est constante.

Modèle

Soit N_t un processus de Poisson non homogène d'intensité $\lambda(t)$, alors conditionnellement à un vecteur de covariables X , pour une exposition d'une unité de temps :

$$\lambda(t|X) = \lambda_0 e^{\beta' X}, \quad (4.1)$$

où β est un vecteur de coefficients qu'il s'agit d'estimer à partir d'observations et β' sa transposée. Et comme ici le risque est supposé constant (toujours conditionnellement aux

covariables), pour une exposition de t unités de temps il vient :

$$\int_0^t \lambda(s|X) ds = \int_0^t \lambda_0 e^{\beta' X} ds = t \lambda_0 e^{\beta' X} = \lambda_0 e^{\beta' X + \ln(t)}.$$

Étant donnée une série d'observations du nombre d'évènements étudiés, $Y = (y_i)$, avec les expositions respectives $E = (e_i)$ données avec les covariables $X = (x_i)$, le modèle est donc :

$$\mathbb{E}(Y|X) = \lambda_0 e^{\beta' X + \ln(E)}. \quad (4.2)$$

Ici nous pouvons noter que le coefficient associé au logarithme de l'exposition est 1 (ce qui signifie simplement que le nombre moyen d'évènements observés varie comme l'exposition, toutes choses étant égales par ailleurs). Une variable explicative pour laquelle le coefficient de régression est forcé à 1 est appelée *offset*.

Estimation

Puisque nous avons

$$\mathbb{P}(Y = y_i | x_i) = \frac{\lambda_0^{y_i} e^{(\beta' x_i + \ln(e_i)) y_i}}{e^{\lambda_0 e^{\beta' x_i + \ln(e_i)}} y_i!},$$

la log-vraisemblance s'écrira :

$$\mathcal{L}(Y|X) = \sum_{i=1}^{i=n} y_i \ln(\lambda_0) + y_i (\beta' x_i + \ln(e_i)) - \log(y_i!) - \lambda_0 e^{\beta' x_i + \ln(e_i)}. \quad (4.3)$$

Comme il n'existe pas, en toute généralité, de formule fermée pour trouver le maximum de cette log-vraisemblance, des méthodes numériques, type Newton-Raphson sont utilisées par les logiciels.

En Pratique

Regroupements

Il peut être intéressant, pour des raisons de temps de calcul, de regrouper les observations selon leurs covariables. Mais comme les dates d'entrées et de sorties de l'étude ne sont pas les mêmes, pour un même regroupement d'observations qui partagent les mêmes modalités, les effectifs (et donc l'exposition du groupe) varient à chaque entrée ou sortie.

De plus, en pratique, il arrive souvent que les covariables varient au cours du temps. C'est notamment le cas dès que l'âge des assurés entre en jeu et que la période d'observation dépasse quelques mois.

Une réponse peut être de découper toute la période d'observation en intervalles définis par chacune des dates d'entrée et de sortie d'observation, ou tout autre évènement qui viendrait modifier les covariables, de sorte que **dans chaque intervalle l'exposition soit constante, et que pour chaque individu, les covariables restent les mêmes**. Ensuite, pour chaque intervalle, les individus peuvent être regroupés selon les modalités de leurs variables explicatives, et leurs expositions sommées pour obtenir l'exposition totale.

La régression s'effectue alors sur les groupes d'individus plutôt que directement sur chaque individu.

Indépendance

Chaque observation devient alors un ensemble d'observations et les expositions sont calculées pour chaque intervalle (la durée totale d'observation est découpée en N intervalles). Pour autant, la log-vraisemblance telle qu'écrite en 4.1.1 suppose évidemment l'indépendance des observations, ce qui est pleinement en contradiction avec le découpage et les regroupements qui viennent d'être proposés au paragraphe précédent.

Il s'avère que dans le cas où les évènements observés ne peuvent arriver qu'une fois pour un individu, et où l'individu n'est plus exposé ensuite au risque étudié (ce qui est donc le cas lorsque que l'on étudie les décès, ou plus généralement la première survenance d'un évènement particulier pour un individu), l'estimation par maximum de vraisemblance est insensible, dans le cadre du modèle de Poisson, à la dépendance due à ces regroupements.

Pour approfondir ce sujet, posons y_{ij} le nombre d'évènements observés sur l'intervalle de temps $j \in 1, \dots, N$ pour l'individu i dont les covariables sur l'intervalle j sont données par le vecteur x_{ij} , et l'exposition sur ce même intervalle est e_{ij} . Nous avons avec cette approche :

$$\mathbb{P}(Y = y_{ij} | x_{ij}) = \frac{\lambda_0^{y_{ij}} e^{(\beta' x_{ij} + \ln(e_{ij})) y_{ij}}}{e^{\lambda_0 e^{\beta' x_{ij} + \ln(e_{ij})}} y_{ij}!}.$$

Nous pouvons cependant remarquer que, puisque l'évènement étudié ne peut arriver qu'une fois pour un individu, y_{ij} est nul sauf peut-être pour une seule valeur de j .

Dans le cas où les covariables pour chaque individu sont constantes, c'est à dire $x_{ij} = x_i$ pour tout intervalle j où l'individu i est présent, nous pouvons donc réécrire la contribution de l'individu i à la log-vraisemblance de la façon suivante, si nous notons N le nombre d'intervalles, avec les conventions que si l'individu i n'est pas présent sur un intervalle j , alors $y_{ij} = 0$, $e_{ij} = 0$ et $x_{ij} = 0$:

$$\begin{aligned} \mathcal{L}_i &= y_i \ln(\lambda_0) + y_i (\beta' x_i + \ln(e_i)) - \log(y_i!) - \lambda_0 e^{\beta' x_i + \ln(e_i)} \\ &= \sum_{j=1}^{j=N} y_{ij} \ln(\lambda_0) + y_{ij} (\beta' x_{ij} + \ln(e_{ij})) - \log(y_{ij}!) - \lambda_0 e^{\beta' x_{ij} + \ln(e_{ij})}. \end{aligned}$$

Et pour estimer β par maximum de vraisemblance, il suffira de maximiser

$$\sum_{i=1}^{i=n} \left\{ \sum_{j=1}^{j=N} y_{ij} (\beta' x_i) - \lambda_0 e^{\beta' x_i + \ln(e_{ij})} \right\} = \sum_{j=1}^{j=N} \left\{ \beta' x_i \sum_{i=1}^{i=n} y_{ij} - \lambda_0 e^{\beta' x_{ij}} \sum_{i=1}^{i=n} e_{ij} \right\}.$$

Et avec un peu d'écriture, il est possible de montrer que l'expression à maximiser serait la même avec l'hypothèse d'indépendance des pseudo-observations issues des regroupements. Ceci permet, pour l'estimation de β , de s'affranchir de la préoccupation de cette indépendance. De plus, la log-vraisemblance obtenue sera équivalente, à une constante près, à celle du modèle sans regroupement. Les valeurs prédites, et toute analyse ultérieure d'adéquation du modèle basée sur les ratios et les dérivées de vraisemblance, resteront similaires. A contrario, les déviations produites par le modèle obtenu sur les données agrégées ne seront pas les mêmes que celles qui auraient été obtenues avec des données non agrégées, et par ailleurs tout ce qui utilise le nombre d'observations ne sera plus directement utilisable (cf. [Rodriguez, 2016]).

Les déviations obtenues sur les données agrégées peuvent cependant être utilisées pour comparer différents modèles entre eux. C'est ce que nous utiliserons par la suite, sur des échantillons de tests, via une validation croisée, pour choisir nos modèles.

4.1.2 D'autres modèles

D'autres modèles sont sans conteste possibles pour obtenir des intensités sur des processus de comptage.

4.2 Détermination de ρ ampleur du changement à détecter

Le niveau de changement à détecter est déterminé par des considérations pragmatiques d'appétence au risque. Les questions qui peuvent se poser sont par exemple : à partir de quelle baisse de l'intensité de la sinistralité est-il possible d'envisager une baisse de tarif ? À partir de quelle hausse faut-il considérer que le tarif devient trop bas ? Ou encore : à partir de quel niveau de changement de l'intensité des sinistres l'impact devient-il non négligeable dans les projections nécessaires au pilier I de Solvabilité II ?

4.3 Choix du seuil d'alerte

Après avoir fait le choix de ρ , il est nécessaire de déterminer la fréquence acceptable de fausses alertes (cf. 2.2.3), faute de quoi il ne serait pas possible de déterminer le niveau du processus CUSUM au-delà duquel lancer une alerte.

4.3.1 La magie du changement de temps

L'intuition première à propos du processus CUSUM est que les niveaux qu'il atteint localement (que ce soit les minima ou les maxima) sont fortement dépendants de l'intensité du processus sous-jacent. Ce qui pourrait rendre complexe la détermination d'un niveau d'alerte, notamment si l'intensité est variable au cours du temps. Par exemple, dans le cas où l'exposition au risque varie avec le temps, comment estimer un niveau d'alerte valable pour les mois où les semaines à venir ? Ou bien faudrait-il réestimer le niveau d'alerte à chaque consultation des valeurs du processus ? Et de fait, cette intuition est pleinement fondée si l'on considère le temps qu'il faut pour que le processus atteigne un certain niveau.

Par contre, il en va tout autrement si, au lieu de considérer le temps écoulé, ce sont les nombres d'évènements qui sont pris en compte, comme l'ont remarqué les auteurs de [El Karoui et al., 2017]. En ce cas, seul le nombre d'évènements observés suffira à déterminer le niveau atteint par le processus CUSUM. Ceci s'explique par le fait que tout processus de Poisson, homogène ou non, sous des hypothèses assez générales, peut se ramener à un processus d'intensité 1 par changement de temps.

Plus généralement, le résultat est le suivant : pour un processus de comptage (N_t) d'intensité cumulée Λ_t , sous réserve que Λ_t soit strictement croissant et tende vers l'infini quand t tend vers l'infini, alors le processus $N_{\Lambda_t^{-1}}$ est un processus de Poisson homogène d'intensité 1, où

$$\Lambda_t^{-1} = \inf_{s>0} \{s : \Lambda_s > t\}.$$

La figure 4.1 présente en gris les niveaux d'un processus de Poisson N_t d'intensité $\lambda = 0.8$, et du processus CUSUM V associé (pour $\rho = 1.5$) et en bleu le processus $N_{\frac{t}{\lambda}}$ d'intensité 1. Ce qu'il importe de noter ici, c'est que les niveaux auxquels arrive le processus CUSUM avant et après changements de temps, s'ils ne sont pas atteints aux mêmes temps, le sont toujours après le même nombre d'évènements observés. Ainsi, le nombre moyen

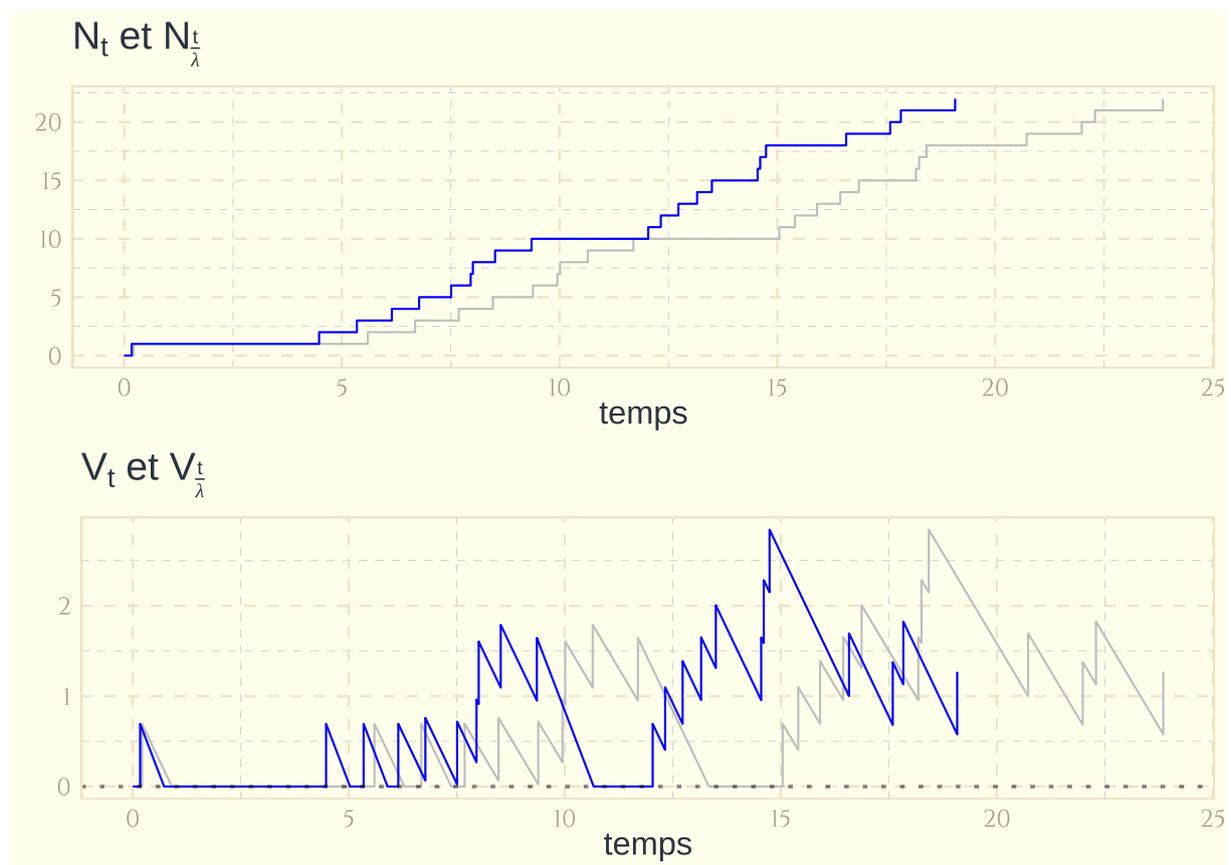


FIGURE 4.1 – Effet du changement de temps sur un processus de Poisson N et sur le processus CUSUM V associé

d'évènements avant que le processus avant changement de temps ne dépasse un certain seuil $\pi > 0$ est le même que pour un processus de Poisson d'intensité 1.

Une conséquence utile : un test d'adéquation pour les modélisations

Outre la simplification des calculs des seuils d'alerte pour le processus CUSUM, ce résultat fournit aussi un moyen de tester l'adéquation de la modélisation choisie pour un processus de comptage. En effet, si une modélisation d'un processus de comptage N_t donne une intensité estimée $\tilde{\Lambda}_t$, le processus $N_{\tilde{\Lambda}_t^{-1}}$ doit être d'autant plus proche d'un processus de Poisson homogène d'intensité 1 que $\tilde{\Lambda}_t$ est une bonne estimation de Λ_t (Cf. [Daley and Vere-Jones, 2003] page 262).

4.3.2 Par calcul direct

L'objectif est donc de déterminer le nombre moyen d'évènements, dans l'hypothèse où il n'y a pas de changement dans l'intensité du processus étudié, avant que le seuil d'alerte soit franchi, c'est à dire $\mathbb{E}_\infty [N_\tau]$.

La méthode proposée ici est issue d'une annexe de l'article [Gandy et al., 2010] et se base sur l'utilisation d'une chaîne de Markov pour caractériser les dépassements progressifs de niveaux intermédiaires entre 0 et le seuil d'alerte. Dans la suite, un processus d'intensité constante λ est considéré, et \mathbb{P}_λ désigne la probabilité associée. De plus, quel que soit λ , c'est au processus CUSUM défini pour un changement d'intensité proportionnel de proportion ρ que ce qui suit s'applique.

Cas $\rho > 1$

Pour un seuil d'alerte c , l'idée est d'utiliser les niveaux successifs $c - k, c - k + 1, \dots, c - 1$, où k est le plus grand entier tel que $c - k > 0$. Plus précisément, posons :

$$a = \frac{\ln(\rho)}{1 - \rho},$$

a est donc le temps qu'il faut au processus U , lorsqu'il n'y a pas d'évènement, pour descendre d'une unité, et

$$k = \lceil c - 1 \rceil.$$

C'est à dire que k est le plus petit entier supérieur ou égal à $c - 1$.

Une chaîne de Markov M à k états transients et un état absorbant est définie de la façon suivante. On pose $\sigma_0 = 0$, puis

$$\sigma_n = \inf\{t > \sigma_{n-1} : U_t \in \{c - k, \dots, c - 1\}\}.$$

et

$$M_n = \begin{cases} \infty, & \sigma_n \geq \tau, \\ \lceil U_{\sigma_n} \rceil, & \sigma_n < \tau. \end{cases}$$

où $\tau = \inf\{t | U_t > c\}$.

La figure 4.2 donne une idée des chemins possibles pour U , et les ronds bleus représentent les états de la chaîne de Markov.

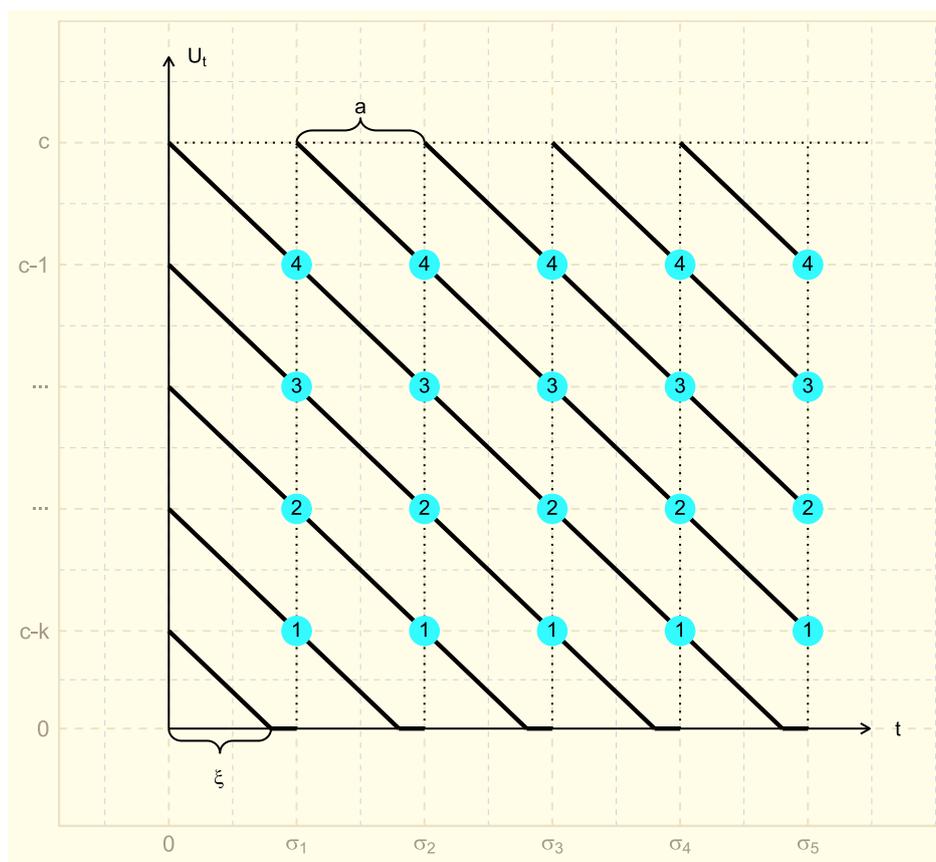


FIGURE 4.2 – Parcours possibles pour U_t

Notons de plus $\xi = a(c - k)$, le temps nécessaire pour que le processus U descende de $c - k$ unités en l'absence d'évènements, et $\eta = a - \xi$.

La probabilité d'observer un nombre $j > 0$ d'évènements pendant la durée t , pour un processus de poisson d'intensité constante λ est notée :

$$p_j(t) = \frac{(\lambda t)^j}{j!} e^{-\lambda t}. \tag{4.4}$$

Etude des probabilités de transitions

Lorsque, pour un entier n donné, M est dans l'état 1 (i.e. $M_n = \lceil U_{\sigma_n} \rceil = c - k$), alors tant qu'il n'y a pas de nouvel évènement, la chaîne de Markov reste dans cet état. Pour observer M_{n+1} et que la chaîne soit toujours dans le même état (i.e. $M_{n+1} = \lceil U_{\sigma_n} \rceil = c - k$), deux cas sont possibles :

- un évènement et un seul a lieu entre σ_n et $\sigma_n + a$.
- il n'y a pas d'évènement avant que U atteigne le niveau 0, puis U reste à ce niveau pour une durée indéterminée, et enfin un évènement se produit (U atteint alors le niveau 1), qui n'est pas suivi d'un autre évènement avant que U redescende jusqu'au niveau $c - k$ (c'est à dire sur une durée de η).

Si l'on note $a_1(\eta, \xi)$ la probabilité de transition de l'état 1 à l'état 1, ces remarques conduisent donc à :

$$\begin{aligned} a_1(\eta, \xi) &= p_1(\xi + \eta) + \int_0^\infty p_0(\xi)p_1(s)p_0(\eta)ds \\ &= p_1(\xi + \eta) + p_0(\xi)p_0(\eta) \int_0^\infty p_1(s)ds \\ &= p_1(\xi + \eta) + p_0(\xi)p_0(\eta) \left(\int_0^\infty \lambda s e^{-\lambda s} ds \right) \\ &= p_1(\xi + \eta) + p_0(\xi)p_0(\eta). \end{aligned}$$

où une intégration par partie montre que l'intégrale est égale à 1 (ou bien, plus directement, la probabilité qu'il existe un temps t tel qu'un évènement ait lieu est de 1).

Avec le même type de raisonnement, et la notation $a_j(\xi, \eta)$ pour la probabilité de transition de l'état 1 à l'état j , il vient :

$$\begin{aligned} a_j(\xi, \eta) &= p_0(\xi)p_{j-1}(\eta) + \sum_{l=1}^{l=j} p_l(\xi)p_{j-l}(\eta) \\ &= p_j(\xi + \eta) + p_0(\xi)(p_{j-1}(\eta) - p_j(\eta)). \end{aligned}$$

Enfin, si la chaîne est dans un état $i > 1$ (i.e. si $M_n = c - k + i$), alors

- sans évènements entre σ_n et $\sigma_n + a$, elle arrivera en $n + 1$ à l'état $i - 1$,
- s'il y a un seul évènement entre σ_n et $\sigma_n + a$ elle reviendra en $n + 1$ au même état i ,
- et selon le nombre d'évènements entre σ_n et $\sigma_n + a$, s'il y en a au moins deux, elle passera aux états suivants, ou bien, si $U_{\sigma_{n+1}} > c$ à l'état absorbant.

La matrice de transition (pour les états transients) de M est donc :

$$Q = Q(\xi, \eta) = \begin{pmatrix} a_1(\xi, \eta) & a_2(\xi, \eta) & \cdots & a_k(\xi, \eta) \\ p_0(\xi + \eta) & p_1(\xi + \eta) & \cdots & p_{k-1}(\xi + \eta) \\ 0 & p_0(\xi + \eta) & \ddots & p_{k-2}(\xi + \eta) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & p_0(\xi + \eta) & p_1(\xi + \eta) \end{pmatrix}.$$

Distribution de M_1

Notons maintenant ν le nombre minimal de sauts nécessaires pour que U puisse passer du niveau U_0 au niveau c , c'est à dire $\nu = \lceil c - U_0 \rceil$.

Si $U_0 > c - k$, le processus peut, s'il n'y a pas de nouvel évènement pendant une durée $t_0 = a(U_0 - (c - \nu))$ atteindre le niveau de type $\lceil c - i \rceil$ immédiatement inférieur à U_0 , ou bien selon le nombre de sauts pendant la durée t_0 atteindre des niveaux supérieurs. La distribution de M_1 sera donc dans ce cas :

$$q = (0, \dots, 0, p_0(t_0), \dots, p_{\nu-1}(t_0)).$$

Par ailleurs, si $U_0 < c - k$, aU_0 est le temps nécessaire, s'il n'y a pas d'évènements, pour que U descende jusqu'à 0, et par un raisonnement similaire à l'étude des transitions depuis l'état 1, il vient $q = (a_1(aU_0, \eta), \dots, a_k(aU_0, \eta))$, qui peut se réécrire :

$$q = (1, 0, \dots, 0)Q(aU_0, \eta).$$

Nombre moyens de sauts pour franchir le niveau c

Avec τ le temps d'arrêt défini par $\tau = \inf\{t | U_t > c\}$, il faut ν sauts au minimum pour que U atteigne c .

Si, après ν sauts le processus U n'a pas atteint le niveau c , il est nécessaire qu'il soit passé par au moins un des niveaux $(c - k, \dots, c - 1)$. Autrement dit, $M_1 < \infty$. Et de même, si, quel que soit $j \geq 1$, après $\nu + j$ sauts, le processus U n'a pas atteint le niveau c , alors nécessairement il est repassé au moins $j + 1$ fois par certains des niveaux $(c - k, \dots, c - 1)$, c'est à dire que $M_{j+1} < \infty$. Réciproquement, si, pour $j \geq 0$, $M_{j+1} = \infty$, il est certain que $N_\tau < \nu + j$. En d'autres termes, les évènements $\{N_\tau > \nu + j\}$ et $\{M_{j+1} < \infty\}$ sont équivalents.

Par ailleurs, la distribution de M_{j+1} dans les états transients nous est donnée par celle de M_1 multipliée $j - 1$ fois par la matrice de transition, nous avons :

$$\mathbb{P}_\lambda[M_{j+1} < \infty] = qQ^{j-1}\underline{1},$$

où $\underline{1}$ est le vecteur dont tous les termes sont égaux à 1.

Il ne reste plus qu'à utiliser une astuce classique sur les espérances pour les variables aléatoires à valeurs discrètes, appliquée à $N_\tau - \nu$, pour retrouver le résultat donné par [Gandy et al., 2010] :

$$\begin{aligned} \mathbb{E}_\lambda(N_\tau - \nu) &= \sum_{j=0}^{j=\infty} \mathbb{P}[N_\tau - \nu > j] \\ &= \sum_{j=1}^{j=\infty} \mathbb{P}[M_{j+1} < \infty] \\ &= \sum_{j=1}^{j=\infty} qQ^{j-1}\underline{1} \\ &= q(I - Q)^{-1}\underline{1}. \end{aligned}$$

Et donc le nombre moyen de sauts avant le dépassement du seuil c est donné par :

$$\mathbb{E}_\lambda(N_\tau) = \nu + q(I - Q)^{-1}\underline{1}.$$

Cas $\rho < 1$

Les définitions de a, ξ, η, ν, p_j sont les mêmes que dans le cas précédent. Ici ν représentera le plus petit entier $n > 0$ tel que $\mathbb{P}(Y_n = \infty) > 0$, et ξ la durée nécessaire pour que le processus Y , en l'absence de saut, monte de $c - k$. La chaîne de Markov qui va être utilisée est la suivante (cf. 4.3) : On pose $\sigma_0 = 0$, puis

$$\sigma_n = \inf\{t > \sigma_{n-1} : U_t \in \{0, c - k, \dots, c - 1\}\},$$

et

$$M_n = \begin{cases} \infty, & \sigma_n \geq \tau, \\ \lceil Y_{\sigma_n} \rceil, & \sigma_n < \tau \end{cases}.$$

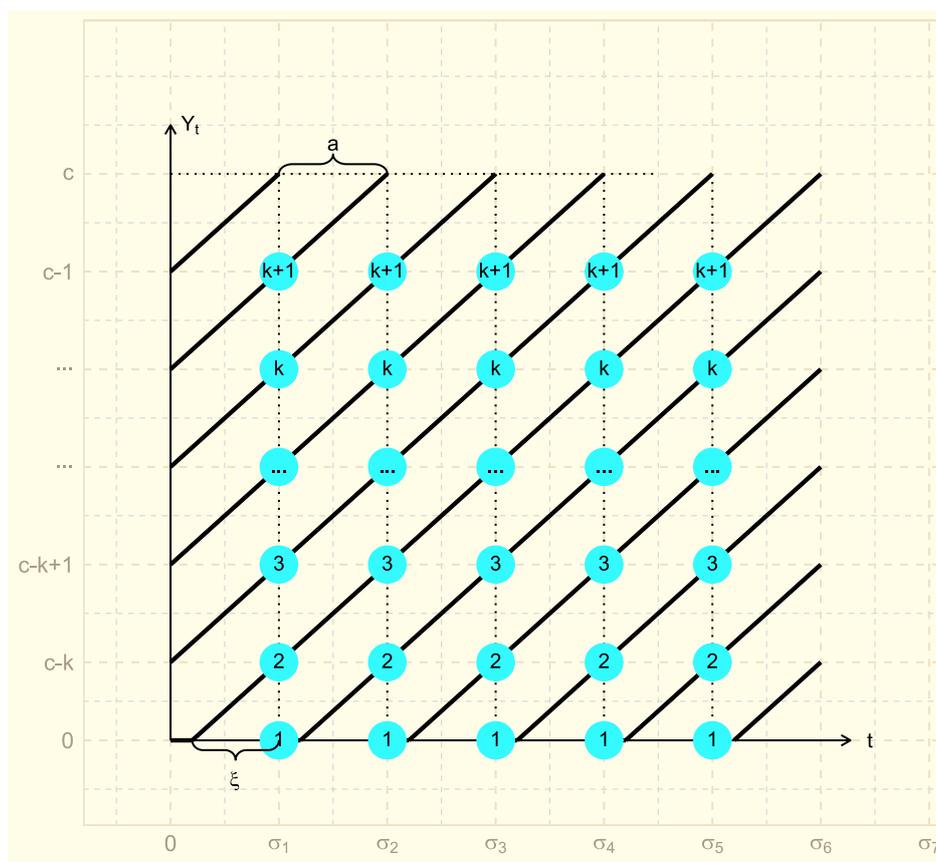


FIGURE 4.3 – Parcours possibles pour Y_t

Étude des probabilités de transitions

Ici chaque saut fait descendre Y de 1. Depuis l'état 1 (pour lequel $Y_{\sigma_n} = 0$), quel que soit le nombre de sauts, s'il y en a au moins 1 sur une durée plus petite que ξ , la chaîne de Markov ne peut que rester dans cet état. S'il n'y a aucun saut pendant une durée au moins égale à ξ alors la chaîne de Markov passe à l'état 2. Elle ne peut en aucun cas atteindre un autre état, donc les probabilités de transition depuis l'état 1 sont données par :

$$\begin{aligned} & (\mathbb{P}_\lambda(M_{n+1} = 1 | M_n = 0), \dots, \mathbb{P}_\lambda(M_{n+1} = k + 1 | M_n = 0)) \\ & = (1 - p_0(\xi), p_0(\xi), 0, \dots, 0). \end{aligned}$$

Si la chaîne de Markov est dans un état $M_n = j > 1$, elle peut, en l'absence de saut, atteindre l'état $j + 1$ (ou bien l'état absorbant ∞ si $j = k + 1$). Si un saut a lieu entre σ_n et $\sigma_n + a$, alors $M_{n+1} = j$, la chaîne de Markov a transitité vers le même état, et selon le nombre de sauts, s'il y en a plus de 1, elle atteindra un des états $i < j$. Plus précisément, s'il y a $j - i$ sauts (avec $i < j$) entre σ_n et $\sigma_n + a$, la chaîne atteint l'état $j - i + 1$, et s'il y a j sauts ou plus, elle atteindra l'état 2 s'il y a une durée ξ sans saut (i.e. le temps qu'il faut au processus X pour remonter de $c - k$), et l'état 1 sinon.

Posons alors, pour $j \geq 2$,

$$\begin{aligned} b_j(\xi, \eta) &= p_{j-1}(\eta + \xi) + p_0(\xi) \sum_{l=j}^{l=\infty} p_l(\eta) \\ &= p_{j-1}(\eta + \xi) + p_0(\xi) \left(1 - \sum_{l=0}^{l=j-1} p_l(\eta)\right). \end{aligned}$$

La matrice de transition (pour les états transients) de M est donc :

$$Q(\xi, \eta) = \begin{pmatrix} 1 - p_0(\xi) & p_0(\xi) & 0 & 0 & \cdots & 0 \\ 1 - p_0(\xi + \eta) - b_2(\xi, \eta) & b_2(\xi, \eta) & p_0(\xi + \eta) & 0 & \cdots & 0 \\ 1 - \sum_{l=0}^{l=1} p_l(\xi + \eta) - b_3(\xi, \eta) & b_3(\xi, \eta) & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & p_0(\xi + \eta) \\ 1 - \sum_{l=0}^{l=k-1} p_l(\xi + \eta) - b_k(\xi, \eta) & b_k(\xi, \eta) & p_k(\xi + \eta) & \cdots & p_2(\xi + \eta) & p_1(\xi + \eta) \end{pmatrix}.$$

Cette matrice peut se réécrire sous la forme, plus lisible et plus pratique pour l'implémentation, proposée par [Gandy et al., 2010] :

$$Q = Q(\xi, \eta) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 - \sum_{j=0}^{j=0} p_j(\eta) & p_0(\eta) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 1 - \sum_{j=0}^{j=k-1} p_j(\eta) & p_{k-1}(\eta) & \ddots & p_0(\eta) \end{pmatrix} \begin{pmatrix} 1 - \sum_{j=0}^{j=0} p_j(\xi) & p_0(\xi) & \cdots & 0 \\ \vdots & p_1(\xi) & \ddots & \vdots \\ \vdots & \vdots & \ddots & p_0(\xi) \\ 1 - \sum_{j=0}^{j=k} p_j(\xi) & p_k(\xi) & \cdots & p_1(\xi) \end{pmatrix}.$$

Distribution de M_1

Ici $t_0 = a(Y_0 - (c - \nu))$ le temps nécessaire depuis le niveau initial Y_0 pour que, s'il n'y a pas de saut, le processus Y atteigne l'état $c - \nu$ de la chaîne de Markov, la façon d'appréhender les cas possibles est par ailleurs similaire à ce qui précède, et la distribution de la chaîne dans les états transients est donc :

$$q = e_{k+2-\nu} Q(\max(a - t_0 - \xi), \min(a - t_0)),$$

où $e_{k+2-\nu}$ est le vecteur dont tous les termes sont nuls sauf celui qui se trouve à la position $k + 2 - \nu$, qui vaut 1.

Nombre moyen de sauts pour franchir le niveau c

Le nombre moyen de sauts pour atteindre le niveau c est maintenant accessible avec un

calcul d'esprit assez similaire à ce qui a été fait pour la cas $\rho > 1$:

$$\begin{aligned}
 \mathbb{E}_\lambda(N_\tau) &= \sum_{j=0}^{j=\infty} \mathbb{P}[N_\tau > j] \\
 &= \sum_{j=0}^{j=\infty} \mathbb{P}[M_{\nu+j} < \infty] \\
 &= \sum_{j=0}^{j=\infty} qQ^{\nu+j-1}\underline{1} \\
 &= qQ^{\nu-1}(I - Q)^{-1}\underline{1}.
 \end{aligned}$$

Nombre moyen de sauts avant une fausse alerte et nombre moyen de sauts avant la détection d'un changement d'intensité du processus

Ce qui précède s'applique tout aussi bien au calcul du nombre moyen de sauts avant une fausse alerte, s'il n'y a pas de changement, qu'au calcul du nombre moyen de sauts entre le moment où le changement se produit, et le moment où l'alerte est déclenchée.

Tout d'abord, en vertu des considérations sur le changement de temps (cf. 4.3.1), le nombre moyen de sauts pour franchir le niveau c s'il n'y a pas de changement est le même que pour un processus de Poisson d'intensité constante égale à 1. Il suffit donc de prendre les résultats précédents avec $\lambda = 1$ pour obtenir $\mathbb{E}_\infty(N_\tau)$.

D'autre part, dans l'hypothèse où le changement s'est produit (i.e. après θ , ou encore sous \mathbb{P}_θ), l'intensité du processus de poisson N n'est plus λ_t mais $\rho\lambda_t$, et de plus :

$$\begin{aligned}
 (\Lambda_t)^{-1} &= \text{inf}_{s>0} \{s : \Lambda_s > t\} \\
 &= \text{inf}_{s>0} \{s : \rho\Lambda_s > \rho t\} \\
 &= (\rho\Lambda_{\rho t})^{-1}.
 \end{aligned}$$

Et il vient donc :

$$N_{\Lambda_t^{-1}} = N_{(\rho\Lambda_{\rho t})^{-1}}.$$

C'est à dire que le changement de temps qui permet de transformer le processus avant changement d'intensité en processus de Poisson d'intensité constante 1, s'il est appliqué au processus après changement d'intensité, transforme ce dernier en processus d'intensité constante ρ .

Pour autant, les pentes des processus U ou Y ne sont pas modifiées. Il suffit donc, dans les réflexions précédentes sur les chaînes de Markov, de changer uniquement les probabilités de saut, sans rien changer d'autre, pour obtenir l'espérance du nombre moyen de sauts avant franchissement du seuil d'alerte c après le changement. Et donc, les résultats précédents appliqués avec $\lambda = \rho$ permettent d'obtenir $\mathbb{E}_\theta(N_\tau)$ (à ceci près qu'il faut admettre que l'on ne connaît pas exactement la position du processus au moment du changement).

Et ceci se généralise même sans autres difficultés à toute autre intensité. Autrement dit, lorsque l'on cherche à détecter le passage de l'intensité λ_t d'un processus à l'intensité $\rho\lambda_t$, nous pouvons connaître tout aussi bien $\mathbb{E}_\infty(N_\tau)$ et $\mathbb{E}_\theta(N_\tau)$, mais aussi $\mathbb{E}_{\theta_x}(N_\tau)$ où \mathbb{P}_{θ_x} désigne les probabilités de sauts pour n'importe quelle intensité proportionnelle à l'intensité initiale.

Si par exemple l'intensité du processus réellement observé après changement est $\alpha\lambda_t$, le nombre moyen de sauts du processus CUSUM utilisé pour détecter un changement d'intensité de λ_t à $\rho\lambda_t$ est obtenu par applications des résultats précédents avec $\lambda = \alpha$.

4.3.3 Par simulations

Principe

Il est aussi envisageable, pour un seuil c , d'obtenir π tel que $\mathbb{E}_\infty [N_{\tau_c}] \geq \pi$ par simulations d'un grand nombre de trajectoires sous \mathbb{P}_∞ d'un processus de Poisson d'intensité 1.

De même, pour les mêmes raisons que celles exposées précédemment, pour c fixé, $\mathbb{E}_\theta [N_{\tau_m}]$ s'estimera facilement par la simulation d'un grand nombre de trajectoires d'un processus de Poisson d'intensité ρ .

4.3.4 Comparaison des deux méthodes

Pour la comparaison des deux méthodes, la simulation porte sur 2000 trajectoires de 5000 évènements chacune. Les figures 4.4 et 4.5 présentent les nombres moyens d'évènements avant une fausse alerte, obtenus par simulations et via la formule exacte, pour différents seuils. Par exemple, pour $\rho = 1, 5$ et un seuil d'alerte de $c = 7, 5$, il faut s'attendre à une fausse alerte après 200 évènements (209, 22 pour être précis).

Ce que montrent ces graphiques, c'est que plus le nombre moyen d'évènements avant une fausse alerte devient important au regard du nombre d'évènements simulés, plus la moyenne obtenue par simulation s'éloigne de la moyenne théorique. Il est bien sûr envisageable d'accroître le nombre d'évènements simulés, mais d'une part, les temps de calculs s'allongent vite avec le nombre d'évènements simulés, et d'autre part si le nombre moyen attendu n'est pas connu, il est difficile de savoir combien d'évènements il faudra prendre en compte (et si on le connaît, la simulation ne se justifierait plus si ce n'était que pour le retrouver). Ceci plaide fortement en faveur de l'utilisation de la formule exacte.

4.3.5 Obtenir un seuil d'alerte

Les deux méthodes ci-dessus donnent en réalité un nombre moyen d'évènements, sous l'hypothèse qu'il n'y a pas de changement, avant franchissement d'un seuil prédéfini. En pratique la question se pose plutôt de façon contraire : *quel est le seuil d'alerte à retenir pour ne pas avoir plus d'une fausse alerte, en moyenne, tous les x années ?*

Mais la chaîne de Markov nous donne, pour un seuil donné, le nombre moyen d'évènements avant une fausse alerte, et non le contraire. Pour répondre à la question, il suffit donc de définir quel est le nombre moyen d'évènements qui devraient, en dehors d'un changement, arriver en x années, puis, par exemple avec une fonction d'optimisation, retrouver le seuil c correspondant (voir par exemple le code proposé en l'annexe C.2).

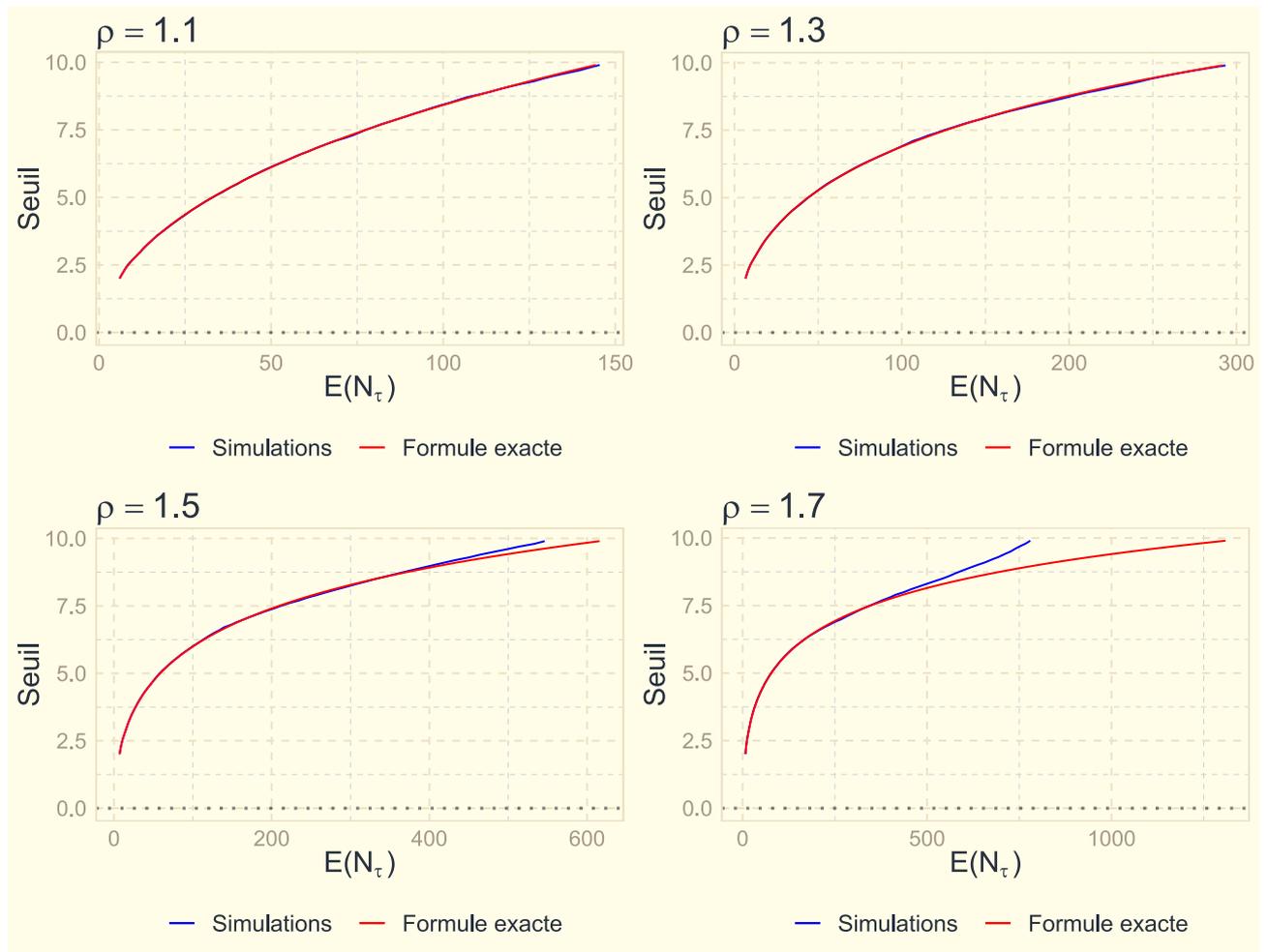


FIGURE 4.4 – Comparaisons des espérances obtenues par simulations et par la formule exacte pour quelques valeurs de $\rho > 1$

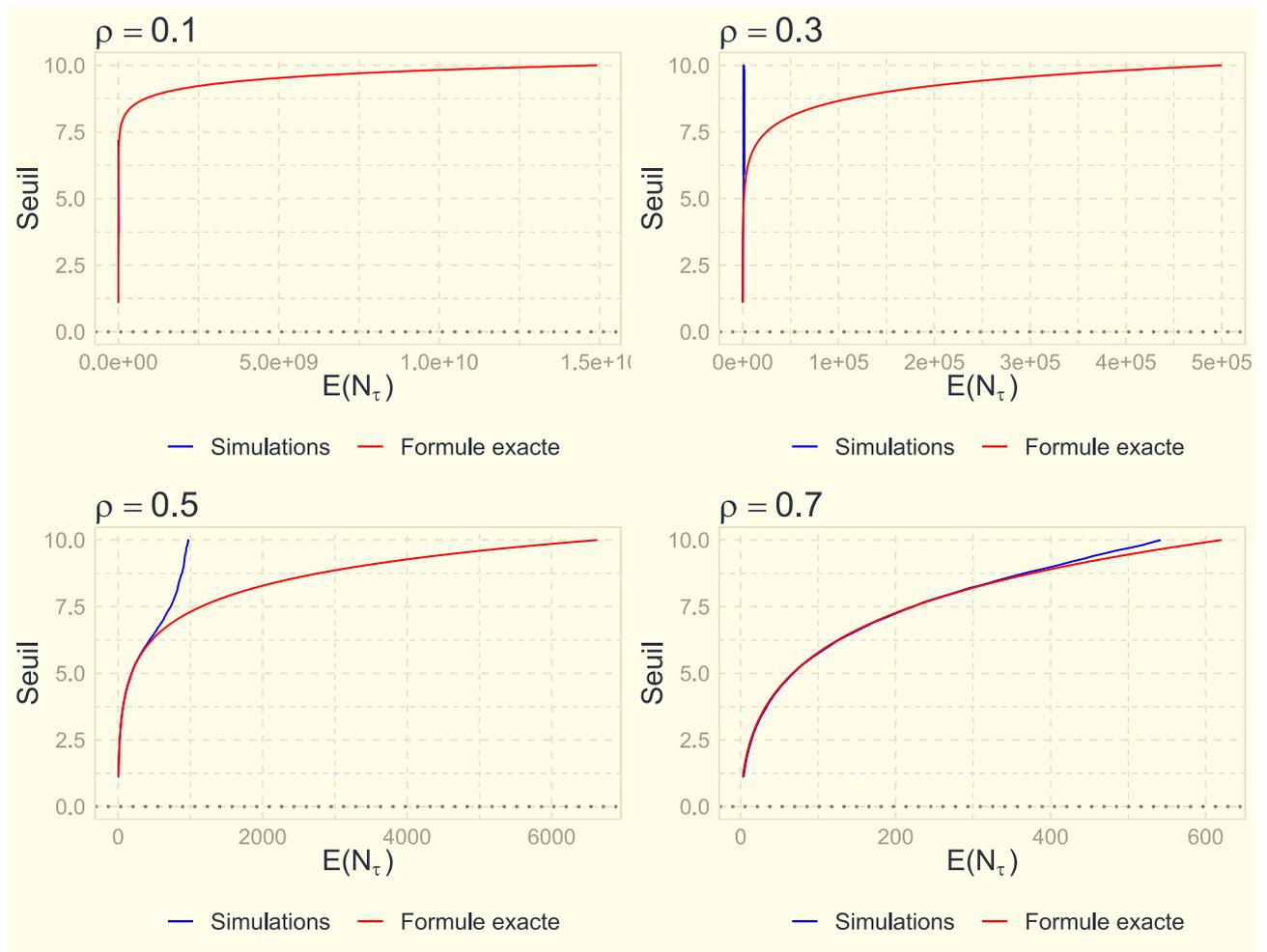


FIGURE 4.5 – Comparaisons des espérances obtenues par simulations et par la formule exacte pour quelques valeurs de $\rho < 1$

Chapitre 5

Applications à la surveillance de la mortalité

La finalité de ce mémoire est la mise en œuvre d'une surveillance des changements qui pourraient intervenir dans l'intensité d'un processus quelconque, avec l'objectif de se doter d'un outil qui permette d'alerter dès qu'un changement potentiel est détecté, et ceci le plus tôt possible. Pour étudier la pertinence d'un tel dispositif, les données disponibles remontent à quelques années. Mais bien évidemment, si l'intensité de base du processus étudié portait sur toutes ces données jusqu'à aujourd'hui, cette méthode ne pourrait guère nous apporter plus d'indications que nous n'en avons déjà. Il est par exemple déjà bien connu qu'une fameuse crise sanitaire s'est abattue sur le monde depuis la fin de l'année 2019. Si notre outil ne nous permettait que de le confirmer deux ans après, il ne serait pas forcément le plus utile. Mais pour autant, il n'est pas possible de seulement se contenter de mettre en place cette nouvelle approche, et d'attendre que survienne un événement qui viendrait chambouler l'intensité des processus sous surveillance. Il faut donc prendre pour référence une date passée, et tenter d'imaginer ce qui se serait passé dans les semaines et mois suivants si le dispositif avait été fonctionnel à cette date.

Dans la suite, il est donc supposé que nous nous trouvons au premier janvier 2019 - ou plutôt quelques mois plus tard, pour éviter une sous-estimation de la sinistralité qui serait due aux sinistres connus plus tardivement - , et que nous disposons des données d'expositions et de sinistres sur les quatre années précédentes. Ces données servent de référence pour établir une intensité de base.

Ensuite, nous devons imaginer que le processus CUSUM est mis en œuvre pour détecter un changement.

5.1 Intensité de référence

5.1.1 Données

Les données observées sont issues de plusieurs systèmes de gestion. Certaines proviennent de succursales (Système 3) et (Système 2), et d'autres de contrats gérés par des partenaires (Système 5 et Système 1), mais la plus grande quantité de données provient des contrats gérés directement par AFI-ESCA en France (Système 4).¹

1. *Système 2, Système 3 et Système 4* correspondent aux logiciels de gestion de contrats propres à AFI-ESCA. Par abus de langage, ceci désigne aussi bien les interfaces homme-machine que les bases de données associées aux logiciels.

Expositions

Le tableau 5.1 présente les expositions sur les quatre années de référence. L'exposition pour un contrat est calculée, pour chaque année, comme le prorata temporis de sa présence dans l'année en question. Ainsi, un contrat souscrit en 2007 et toujours actif au 31/12/2015 aura une exposition de 1 pour 2015. Si ce contrat s'arrête (quelle qu'en soit la raison) au 30/06/2016, son exposition sera de 0,5 au titre de 2016, puis de 0 pour 2017 et 2018.

Il est utile de noter que l'exposition totale augmente, ce qui est le signe d'un portefeuille en croissance. Cette information sera à prendre en compte lorsque que seront définis les seuils d'alerte (voir 5.2.1).

	Système de Gestion	2015	2016	2017	2018
1	Système 1	0	0	1	440
2	Système 2	16 074	21 532	29 516	37 679
3	Système 3	1 510	2 803	4 102	5 469
4	Système 4	159 373	169 463	170 328	173 964
5	Système 5	5 264	9 301	9 477	9 242
6	TOTAL	182 222	203 101	213 425	226 796

TABLE 5.1 – Exposition par année d'observation et sources de données

Nombre de décès

Les nombres de décès constatés chaque année sont répertoriés dans le tableau 5.2. Ces nombres sont en croissance, ce qui semble à priori cohérent avec la croissance des expositions (à priori seulement, car sans précision sur l'évolution de la composition du portefeuille d'assurés, cette cohérence apparente pourrait masquer par exemple une drastique chute de la mortalité âge par âge sur un portefeuille dont l'âge moyen augmenterait très vite).

	Système de Gestion	2015	2016	2017	2018
1	Système 1	0	0	0	1
2	Système 2	38	57	61	51
3	Système 3	1	2	2	2
4	Système 4	258	262	281	307
5	Système 5	1	2	3	2
6	TOTAL	298	323	347	363

TABLE 5.2 – Nombre de décès par année d'observation et sources de données

Taux Bruts

Les taux bruts présentés dans le tableau 5.3 sont calculés comme le rapport entre un nombre de décès annuel et l'exposition correspondante. Sans surprise, les taux sont plus stables si le portefeuille étudié est plus important. Pour l'ensemble, les taux bruts annuels semblent relativement stables autour de 1.6 ‰. Là encore, cette stabilité apparente pourrait masquer des évolutions contradictoires dans la composition du portefeuille.

	Système de Gestion	2015	2016	2017	2018
1	Système 1	0,000 ‰	0,000 ‰	0,000 ‰	2,269 ‰
2	Système 2	2,364 ‰	2,647 ‰	2,067 ‰	1,354 ‰
3	Système 3	0,662 ‰	0,713 ‰	0,488 ‰	0,366 ‰
4	Système 4	1,619 ‰	1,546 ‰	1,650 ‰	1,765 ‰
5	Système 5	0,190 ‰	0,215 ‰	0,317 ‰	0,216 ‰
6	TOTAL	1,635 ‰	1,590 ‰	1,626 ‰	1,601 ‰

TABLE 5.3 – Taux de décès bruts par année d'observation et sources de données

Composition du portefeuille

Les remarques faites aux deux paragraphes précédents incitent à étudier quelques éléments sur la constitution du portefeuille 5.4.

L'âge moyen (calculé comme la moyenne pondérée par l'exposition) est en très légère hausse.

Les assurés **surprimés** sont des assurés qui, soit parce qu'ils souffrent de pathologies particulières, soit parce qu'ils exercent des activités dangereuses, se voient demander un supplément de prime (une *surprime*). La baisse constatée de la part d'assurés surprimés au cours des quatre années observées peut trouver plusieurs explications : amélioration générale de la santé de la population, ou bien sévérité accrue des critères de sélection médicale qui conduirait à augmenter les refus d'assurer, ou encore au contraire, allègement des critères qui contribuent à imposer une surprime. Une autre hypothèse envisageable peut être celle de la position nouvelle d'un assureur concurrent vis à vis de ces risques particuliers. Assez probablement, l'évolution peut tout simplement être liée à une plus grande diversification du portefeuille d'AFI-ESCA au fur et à mesure que ce dernier croît.

L'étiquette **VLR** (acronyme pour *Very Low Risk*) s'applique à des assurés cadres ou fonctionnaires de catégorie A ou chefs d'entreprise, et qui, dans l'exercice de leurs professions, n'ont pas d'activités dangereuses (pas de travail à plus de 20 mètres de hauteur, pas de contact avec des produits explosifs ou inflammables, pas de travail manuel), ni de déplacements de plus 30 000 kilomètres annuellement. Les assurés qui relèvent de cette catégorie sont connus pour présenter un risque moins élevé que ceux qui n'en relèvent pas. L'évolution observée de la part des VLR dans le portefeuille montre une nette croissance. Les raisons n'en sont pas faciles à déterminer. Il peut aussi bien s'agir d'une évolution générale de la population française, ou bien de la frange de la population française qui accède à la propriété (ce qui, socialement, peut être le contraire complet du premier cas), ou encore un effet lié à la différenciation tarifaire qu'apporte AFI-ESCA à cette catégorie d'assurés, et à l'affinage de cette différenciation au cours du temps.

Sur les quatre années observées, la part des femmes est en légère baisse (de 0,9 points). Ici encore, toute tentative d'interprétation serait hasardeuse.

La chute un peu plus marquée de la part des fumeurs peut, elle, être mise en regard de l'évolution de la tabagie en France. Le document [[Santé Publique France, 2021](#)] montre dans sa figure 1 une chute de l'ordre de trois points de la prévalence du tabagisme quotidien parmi les 18-75 ans en France, tant chez les hommes que les femmes. Cette prévalence passe, tous sexes confondus, selon ce document, de 29,4% en 2015 à 25,4%. Même si l'évolution de la tabagie semble suivre une tendance similaire, il est notable que le portefeuille des assurés qui fait l'objet de cette étude présente une prévalence inférieure de plus de dix points à celle de la population française.

	2015	2016	2017	2018
1 Age Moyen	43,7	43,8	44,2	44,4
2 Proportion d'assurés classés VLR	23,0 %	24,5 %	24,9 %	25,6 %
3 Proportion d'assurés surprimés	20,0 %	19,1 %	18,4 %	17,6 %
4 Proportion de femmes	46,8 %	46,6 %	46,2 %	45,9 %
5 Proportion de fumeurs	15,0 %	14,0 %	13,2 %	12,6 %

TABLE 5.4 – Quelques éléments sur la composition du portefeuille

5.1.2 Modélisation

Le principe retenu pour la modélisation est celui décrit en [sous-section 4.1.1](#). Chaque observation correspond donc à une période très brève (presque toujours un seul jour) et à un n-uplet de valeurs possibles pour les variables explicatives autres que l'exposition. Les variables explicatives prises en compte sont les suivantes :

- Age (en millésime)
- Genre
- Tabagie
- Caractère VLR
- La présence ou non de surprime
- L'ancienneté du contrat
- Un indicateur d'âge avant ou après 60 ans
- La saison²
- L'exposition totale pour le n-uplet de modalités des autres variables explicatives

La variable à expliquer est le nombre des décès observés sur chaque période, et pour chaque n-uplet des modalités des variables explicatives.

Le choix de l'âge en millésime peut être discuté. Cependant il correspond à l'âge utilisé pour la tarification des contrats, et tant que la méthode de calcul reste la même dans la modélisation des taux de décès et dans l'utilisation qui en sera faite ensuite via le processus CUSUM, c'est une donnée cohérente.

Une autre approximation porte sur les surprimes : quel qu'en soit le niveau, ou le type (il existe des surprimes pour raisons médicales, d'autres liées à l'exercice d'une profession risquée, et d'autres encore à la pratique d'un sport à risque), l'indication sera la même. Mais prendre en compte l'ensemble des différents types et niveaux de surprimes pourrait aussi conduire à un sur-ajustement du modèle. Le principe de parcimonie et la significativité statistique de l'effet de cette variable binaire plaident en faveur du choix qui a été fait.

Enfin, la question se pose de l'utilisation de l'information sur la source de la donnée. Si, sans autres éléments - qui pourraient porter par exemple sur une sélection médicale différente - , il ne semble pas opportun de distinguer les assurés français selon le gestionnaire, la question se pose tout de même pour les assurés étrangers. En réalité, dans les modèles testés, ce critère n'est jamais apparu statistiquement significatif, et cette donnée a été écartée assez vite lors du travail sur la modélisation.

2. Saison approximative : hiver = décembre, janvier, février ; printemps = mars, avril mai ; été = juin, juillet, août et automne = septembre, octobre, novembre.

Sélection du modèle

Puisqu'il n'est pas possible d'utiliser directement l'étude des résidus pour contribuer à la validation du modèle, ou bien de se fier aux critères habituels pour départager les modèles entre eux, le critère retenu sera celui de la somme des déviations produites par les prédictions du modèle sur des ensembles de test, dans le cadre d'une validation croisée (sur 10 sous-ensembles) - concrètement, à partir des données disponibles sur 2015-2018.

La déviance de Poisson (voir l'annexe B à ce propos) est donnée par :

$$D = 2 * \sum_i \left[y_i * \log\left(\frac{y_i}{\hat{y}_i}\right) - (y_i - \hat{y}_i) \right],$$

où y_i est l'observation, \hat{y}_i la valeur prédite par le modèle.

Plusieurs modélisations sont mises en œuvre, avec la présence ou l'absence de certaines des données et quelques interactions entre ces données (Voir l'annexe A pour un aperçu de quelques-uns de ces modèles)

Modèle retenu

Avec une application stricte de la préférence pour le modèle qui aboutirait à la plus petite déviance sur l'échantillon de test, le modèle retenu serait le modèle 4 (voir A.1) . Mais au vu du gain en déviance, par souci de parcimonie le modèle présenté en table 5.5 est préféré (voir section A.1).

L'équation 4.1 nous permet d'ores et déjà de tirer quelques informations intéressantes, indépendamment de l'utilisation qui sera faite du modèle pour la détection des changements d'intensité.

En effet, l'exponentielle de chaque coefficient estimé par le modèle donne une idée de l'effet relatif, toute chose étant égale par ailleurs, sur l'intensité observée, de la variation d'une variable explicative. Ainsi, selon ce modèle, l'intensité des sinistres décès est, toute chose égale par ailleurs, 2,15 fois supérieure chez les fumeurs que chez les non-fumeurs, et 1,86 fois supérieure chez les hommes que chez les femmes, les assurés meurent 1,18 fois plus en hiver qu'en automne (ici il faut remarquer que l'automne n'étant pas présente dans la liste des coefficients, c'est elle qui est la modalité de référence, c'est à dire qui correspond au coefficient nommé *intercept*), et chaque année d'ancienneté d'un contrat apporte un léger surplus de sinistralité³.

3. L'idée selon laquelle la sélection médicale joue sur la sinistralité surtout les trois premières années a été testée via une variable binaire pour les anciennetés de plus de trois ans - dans un modèle sans la variable ancienneté -, mais elle n'a pas donné d'amélioration par rapport au présent modèle.

<i>Dependent variable :</i>	n.sinistre				
	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-12.2679	0.2092	-58.63	0.0000	***
age	0.1030	0.0029	35.30	0.0000	***
genreH	0.6193	0.0416	14.88	0.0000	***
est_fumeur	0.7653	0.0514	14.89	0.0000	***
anciennete	0.0192	0.0047	4.08	0.0000	***
avec_surprime	0.1952	0.0438	4.46	0.0000	***
age.60avant60	-0.2025	0.0725	-2.79	0.0052	**
VLR	-0.4457	0.0550	-8.11	0.0000	***
saisonete	-0.0862	0.0563	-1.53	0.1255	
saisonhiver	0.1703	0.0532	3.20	0.0014	**
saisonprintemps	0.1103	0.0539	2.05	0.0405	***
Observations	28 654				
Log Likelihood	- 5936.730				
Akaike Inf. Crit.	11 889.460				

Note : *p<0.1; **p<0.05; ***p<0.01

TABLE 5.5 – Modèle retenu pour le décès

5.2 Détection des changements d'intensité

5.2.1 Mise en Œuvre

Une fois le modèle qui donne l'intensité de référence estimé, il est possible de mettre en œuvre le suivi via le processus CUSUM tel que décrit en 3.5.3. Les prédictions selon le modèle permettent d'obtenir, en fonction des variables explicatives observées période après période, l'intensité cumulée théorique Λ_t , et les sinistres survenus donnent directement N_t .

Les seuils d'alerte ont d'abord été estimés sur la base d'une fausse alerte en moyenne tous les trois ans, ce qui concrètement revient à considérer le nombre moyen de sinistres sur trois ans pendant la période de référence. Ce nombre moyen une fois déterminé, les seuils (à la hausse et à la baisse) sont obtenus grâce à la chaîne de Markov décrite en 4.3.2. Sur le conseil de Stéphane Loisel, les seuils ont été ensuite adaptés pour une fausse alerte tous les vingt ans, car la première approche conduirait à des alertes trop fréquentes en pratique pour que les décideurs puissent prendre des mesures particulières avec assez de certitude sur la nature de l'alerte.

Il n'est pas impossible que l'intensité des décès puisse connaître des fluctuations mineures (en fonction par exemple de l'évolution de la sélection médicale, ou de l'évolution des proportions d'assurés fumeurs et non-fumeurs dans le portefeuille). Cependant, une variation de 10 points, sans autres facteurs exogènes, semble assez peu probable, et par ailleurs, une telle fluctuation pourrait s'avérer suffisamment impactante sur les résultats de l'assureur pour qu'elle nécessite une attention particulière, voire des évolutions des produits (soit tarifaires, soit au niveau de la couverture). Pour cette étude, c'est ce niveau de variation qui est surveillé, sans que cela signifie que ce serait ce niveau qui serait retenu si une telle surveillance était mise en œuvre de façon régulière.

5.2.2 Résultats

Population globale

Sur l'ensemble du portefeuille étudié, avec un seuil d'alerte fixé à partir d'une fausse alerte tous les trois ans, un premier dépassement du seuil d'alerte pour une hausse de plus de 10% est visible courant février 2021 (figure 5.1, sur cette figure comme sur les suivantes, les premiers jours de l'année sont repérés par des lignes discontinues verticales noires, et les premiers jours de chaque mois -sauf janvier- par des lignes discontinues verticales grises). Ce premier dépassement semble de courte durée, et pourrait être considéré comme une fausse alerte. Il est suivi d'un second dépassement sur avril-mai de la même année, et d'une troisième période de dépassement en novembre. Pour une baisse de l'intensité de 10%, un seul dépassement est observé, en décembre 2019. Il est donc extrêmement difficile de déterminer, sur la seule base de ces graphiques, si des changements d'intensité de l'ampleur étudiée ont eu lieu. Pourtant, il serait tentant de mettre l'absence de retour au niveau 0 du processus CUSUM pour la hausse depuis mars 2020 en regard des courbes de mortalités liées à la COVID (voir figure 5.3) selon les données du site **Our World In Data** [Ritchie et al., 2020].

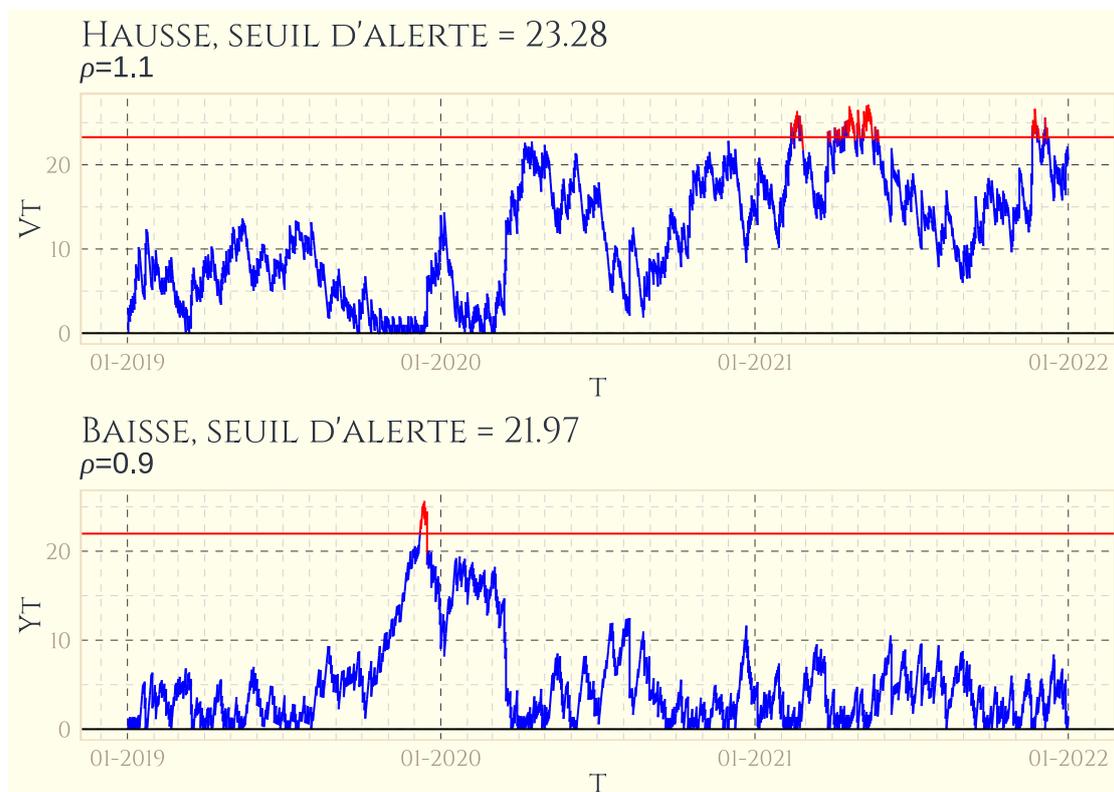


FIGURE 5.1 – Évolution du processus CUSUM pour le décès (population générale) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les trois ans

Lorsque le seuil d'alerte est basé sur une fausse alerte tous les vingt ans, les courbes sont les mêmes (voir figure 5.2), mais le seuil d'alerte est nettement plus élevé. Sur la base de ce niveau, le dispositif n'aurait à aucun moment déclenché d'alerte, ni à la hausse, ni à la baisse, depuis le premier janvier 2019.

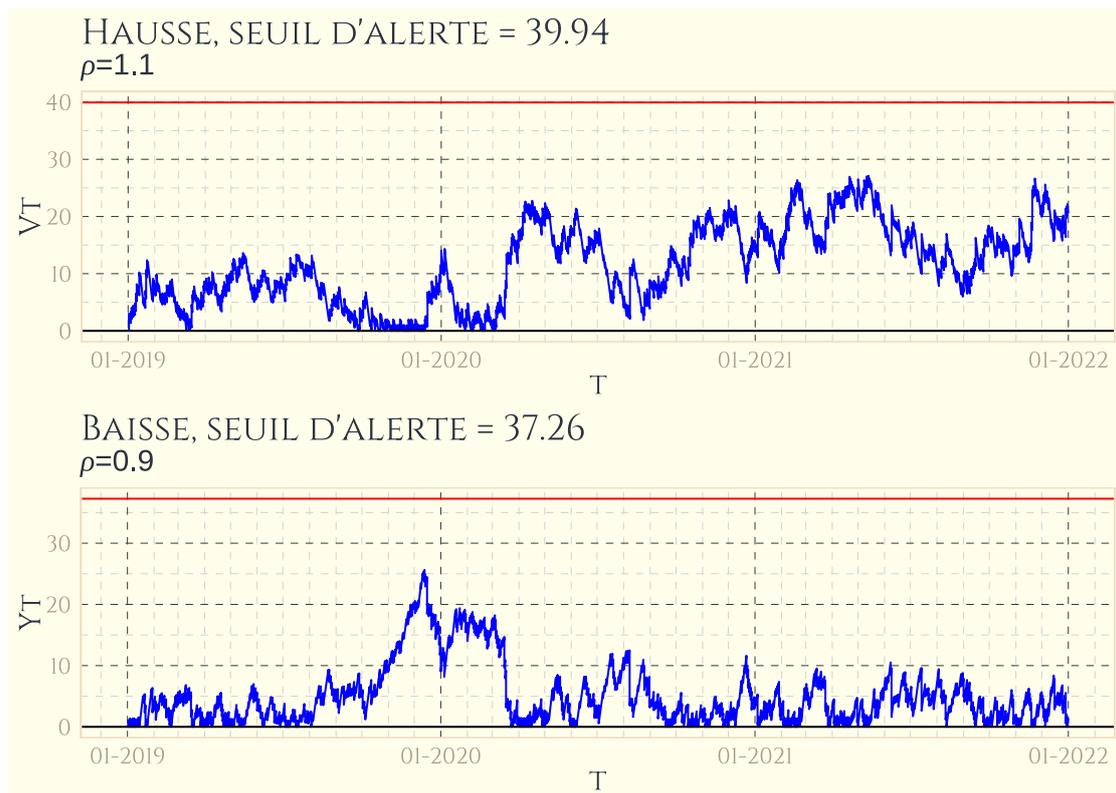


FIGURE 5.2 – Évolution du processus CUSUM pour le décès (population générale) depuis le 1er janvier 2019, seuil d’alerte fixé pour une fausse alerte tous les vingt ans

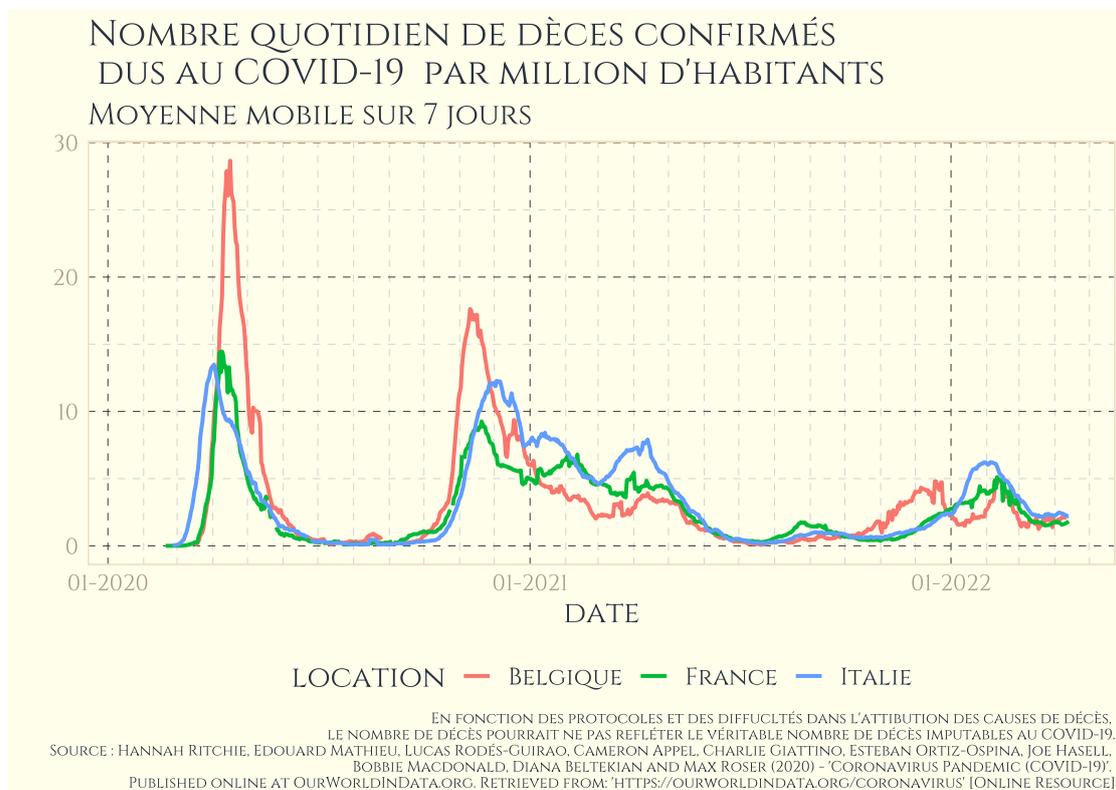


FIGURE 5.3 – Sinistralité liée au coronavirus en France, Italie et Belgique

Par sous-groupes

Sur les assurés en France

La notion d'assurés en France mérite quelques précisions. Il s'agit des assurés dont les contrats sont gérés soit dans la base Système 4, soit par les partenaires (Système 1 et Système 5). Ceci ne signifie pas pour autant que tous ces assurés résident en France. AFI-ESCA propose en effet des contrats d'assurance à certains résidents étrangers. Cependant, ces cas restent à ce jour fortement minoritaires dans l'ensemble.

Le graphique 5.4, basé sur une fausse alerte tous les trois ans en moyenne, lance une alerte à la baisse en décembre 2019, qui pourrait bien être une fausse alerte, et plusieurs alertes à la hausse, en mai, juin et juillet 2020, puis le seuil d'alerte semble franchi durablement depuis février 2021. Si cette fois encore il faut s'interroger sur la pertinence d'un seuil établi à partir de cette fréquence de fausses alertes, la figure 5.5, construite sur une fausse alerte tous les vingt ans, ne montre plus aucun dépassement de seuil avant octobre 2021. Le processus franchit alors le seuil et ne repasse pas en deçà jusqu'à la fin de l'étude. Ce dépassement semble nécessairement plus crédible que les dépassements obtenus à partir d'une fréquence moindre de fausses alertes. Pour autant, rien n'interdit encore de penser qu'il pourrait s'agir soit d'une *vraie* fausse alerte ou encore qu'un changement ait réellement eu lieu à la hausse dans l'intensité des sinistres, mais pas forcément d'une ampleur telle que celle pour laquelle le seuil d'alerte a été défini.

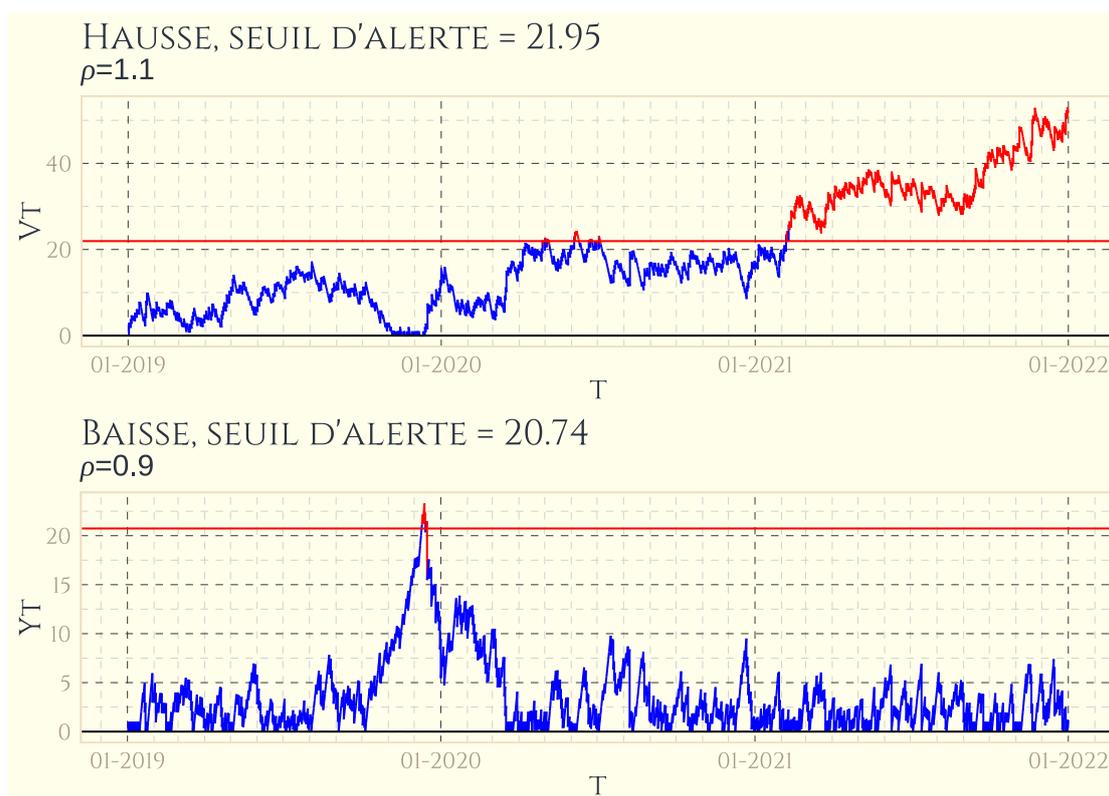


FIGURE 5.4 – Évolution du processus CUSUM pour le décès (assurés en France) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les trois ans

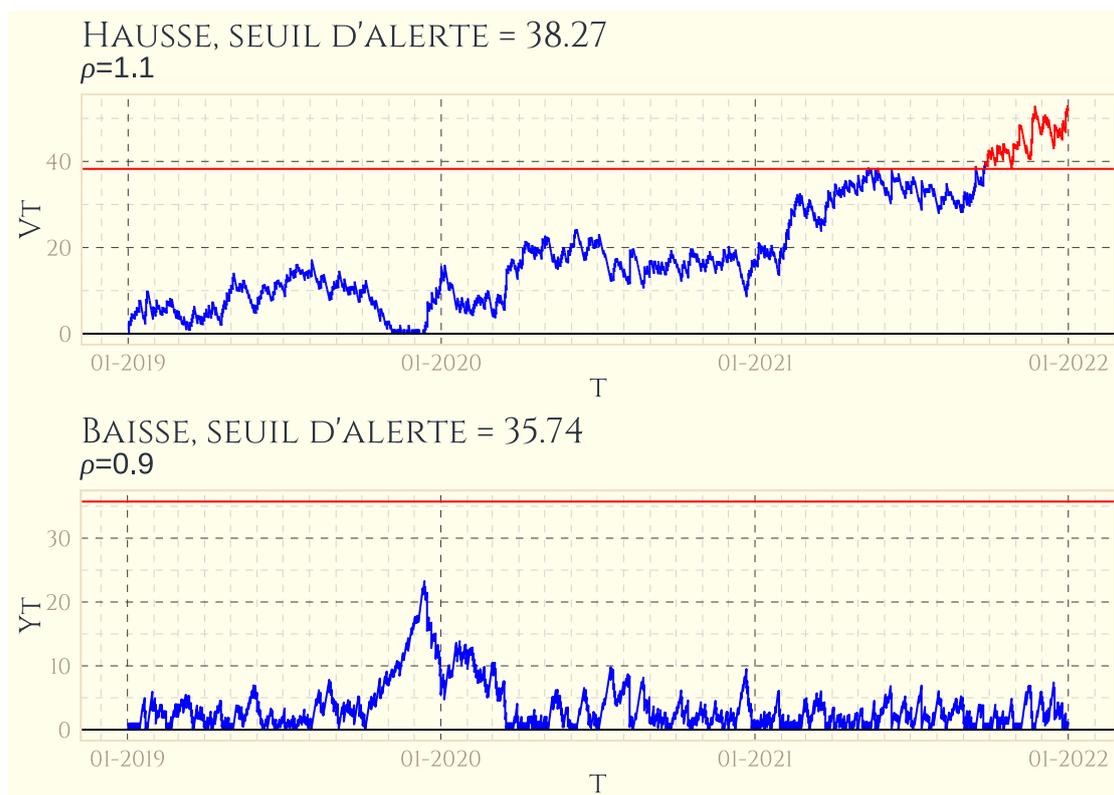


FIGURE 5.5 – Évolution du processus CUSUM pour le décès (assurés en France) depuis le 1er janvier 2019, seuil d’alerte fixé pour une fausse alerte tous les vingt ans

Selon le genre

Le graphique 5.6 présente le processus CUSUM depuis 2019 sur la seule population féminine. La première chose à noter ici est le niveau du seuil d’alerte, lui aussi basé sur une fausse alerte tous les trois ans. Il est à 15,08, quand pour la population générale (cf 5.1) il est à 23,28. En effet, le nombre moyen de sinistres en 3 ans n’est pas le même s’il s’agit de la population générale, ou s’il s’agit d’une partie de cette population. Il est donc essentiel de recalculer un seuil adapté à chaque fois qu’un sous-groupe doit être surveillé. Ici, avec un seuil d’alerte fixé pour une fausse alerte tous les 3 ans, une alerte à la hausse est déclenchée courant septembre 2020, et le seuil d’alerte reste franchi presque constamment depuis. A contrario, pour la population masculine (figure 5.8) un unique dépassement très bref a lieu courant mars 2020. Dans les deux cas, des franchissements de seuils à la baisse de très courte durée sont aussi visibles.

Pour les mêmes courbes, avec un seuil d’alerte déterminé pour une fausse alerte tous les 20 ans, tant chez les hommes que chez les femmes, plus aucun dépassement n’a lieu.

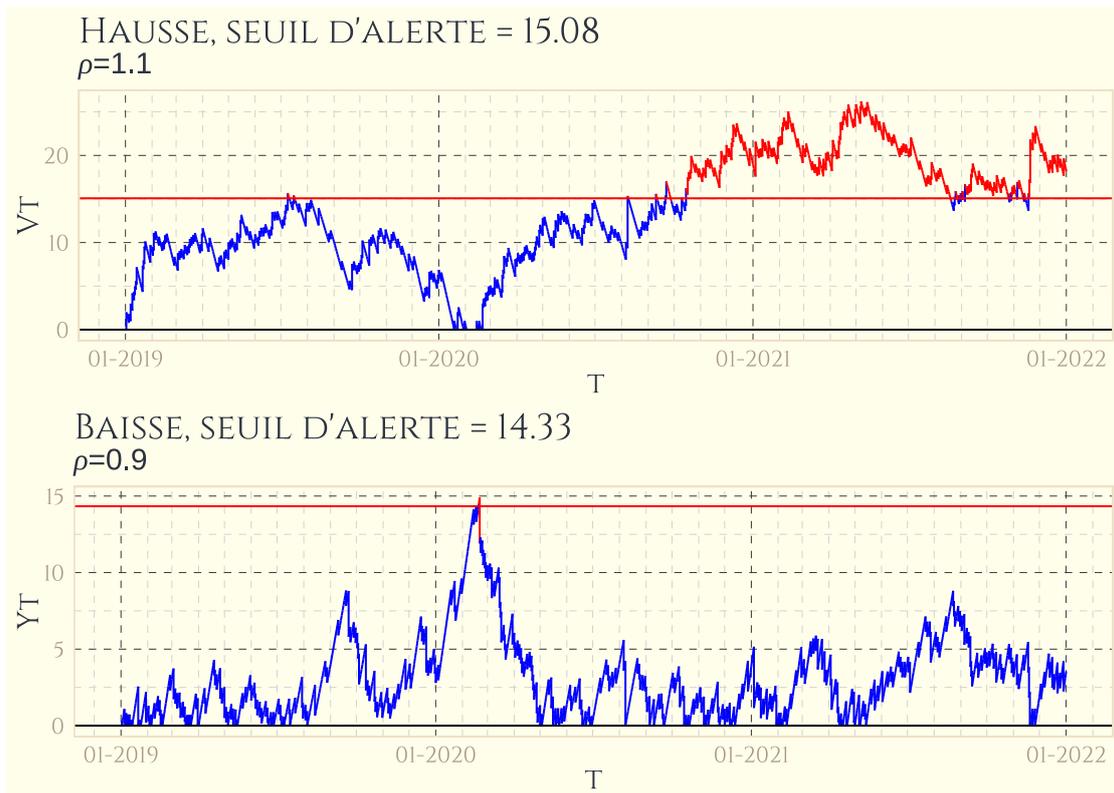


FIGURE 5.6 – Évolution du processus CUSUM pour le décès de la population féminine depuis le 1er janvier 2019, seuil d’alerte fixé pour une fausse alerte tous les trois ans



FIGURE 5.7 – Évolution du processus CUSUM pour le décès de la population féminine depuis le 1er janvier 2019, seuil d’alerte fixé pour une fausse alerte tous les vingt ans

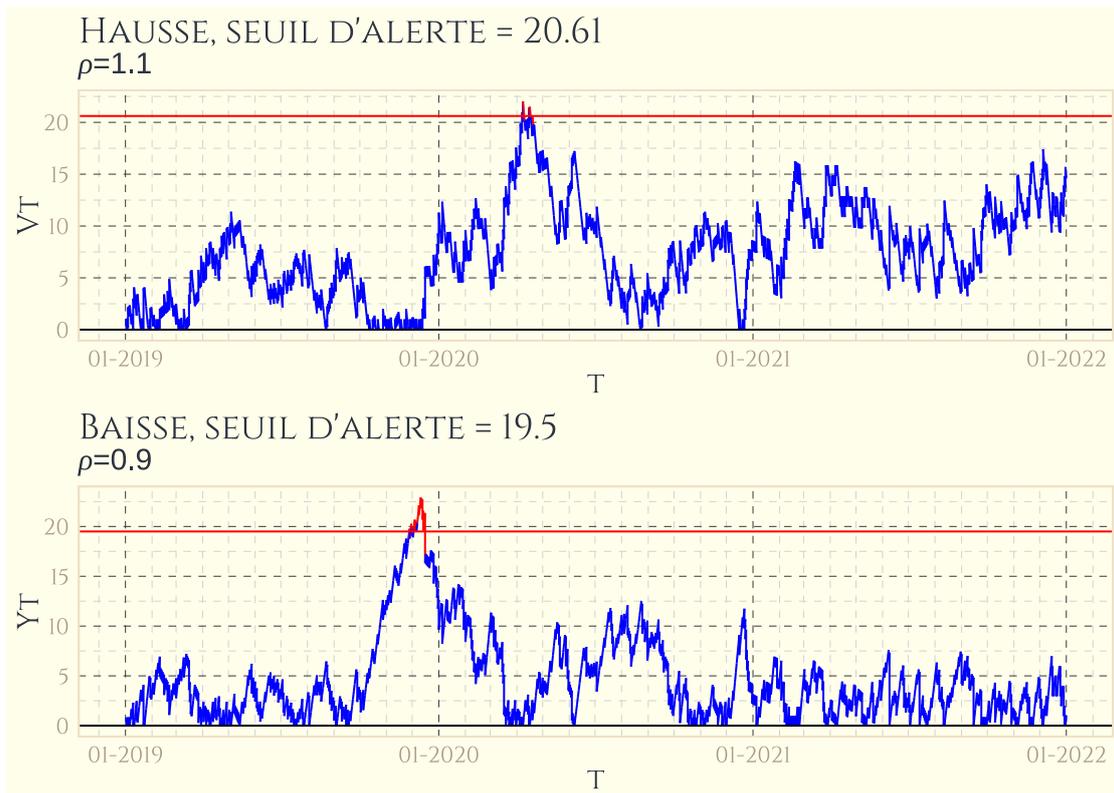


FIGURE 5.8 – Évolution du processus CUSUM pour le décès de la population masculine depuis le 1er janvier 2019, seuil d’alerte fixé pour une fausse alerte tous les trois ans



FIGURE 5.9 – Évolution du processus CUSUM pour le décès de la population masculine depuis le 1er janvier 2019, seuil d’alerte fixé pour une fausse alerte tous les vingt ans

Selon la tabagie

Ici encore, le choix du nombre moyen de fausses alertes qui dicte le seuil d'alerte s'avère crucial. Pour une fausse alerte en moyenne tous les trois ans, des alertes à la hausse surviennent en mai 2020 puis à partir de novembre de la même année, et jusqu'en juin 2021 pour les non-fumeurs (cf. figure 5.10), mais il n'y a plus du tout d'alerte si c'est une fausse alerte en moyenne tous les vingt ans qui est retenue (cf. figure 5.12).

La situation pour les fumeurs appelle le même commentaire (figures 5.11 et 5.13). L'aspect moins *agité* des courbes pour ces derniers est lié au fait que les fumeurs sont nettement moins nombreux, et qu'il y a donc beaucoup moins d'évènements observés sur la même période.

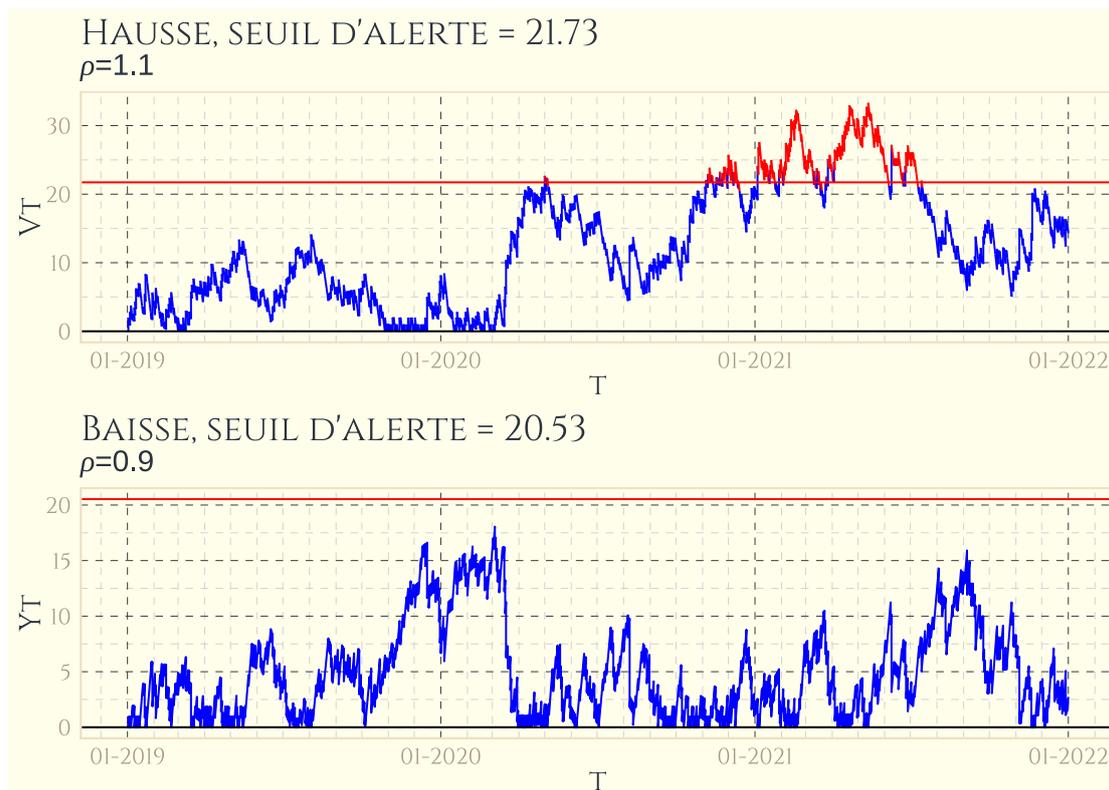


FIGURE 5.10 – Évolution du processus CUSUM pour le décès des non-fumeurs depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les trois ans



FIGURE 5.11 – Évolution du processus CUSUM pour le décès des fumeurs depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les trois ans



FIGURE 5.12 – Évolution du processus CUSUM pour le décès des non-fumeurs depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans



FIGURE 5.13 – Évolution du processus CUSUM pour le décès des fumeurs depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans

Selon l'âge

Pour éviter de répéter à nouveau la remarque sur le choix des seuils d'alertes, ne sont plus présentés ici que les seuils d'alerte basés sur une fausse alerte en moyenne tous les vingt ans.

Le premier graphique montre les processus CUSUM pour une détection à la hausse et à la baisse de la sinistralité décès chez les 35-54 ans. Si le suivi avait été mis en œuvre de début 2019 à fin 2021, aucune alerte n'aurait retenti (figure 5.14)



FIGURE 5.14 – Évolution du processus CUSUM pour le décès des 35-54 ans depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans

La situation est tout à fait similaire chez les 55-64 ans (figure 5.15).

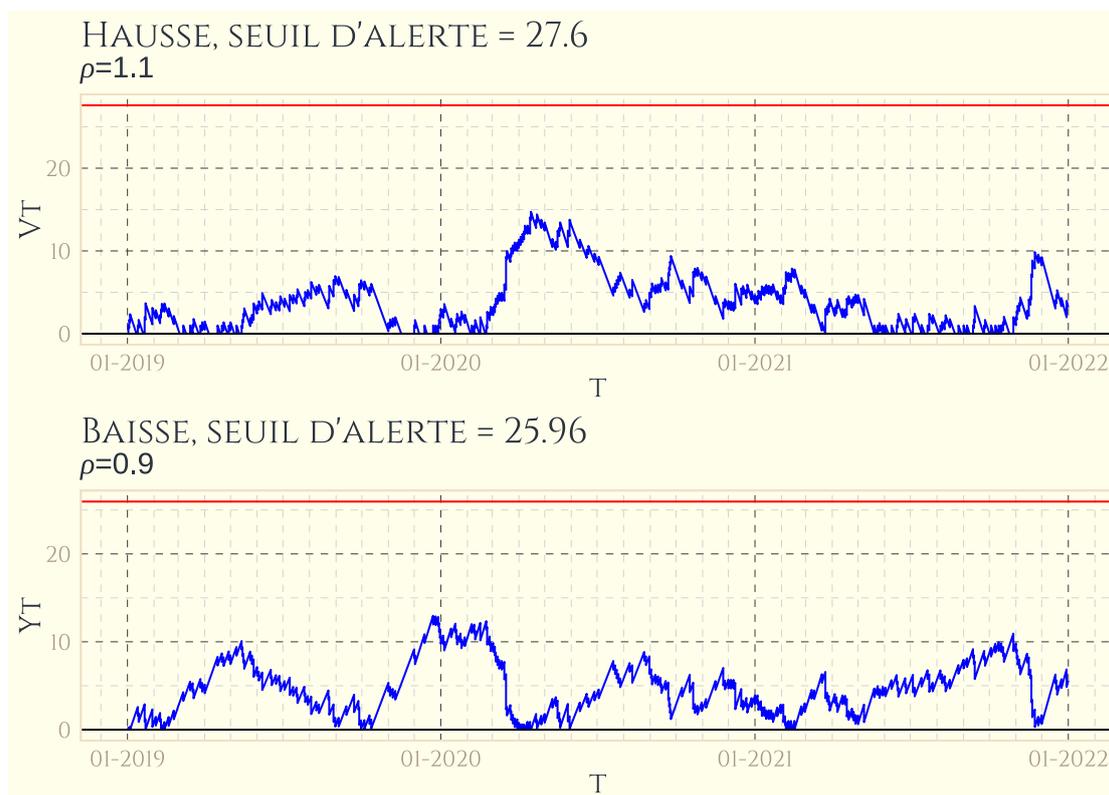


FIGURE 5.15 – Évolution du processus CUSUM pour le décès des 55-64 ans depuis le 1er janvier 2019, seuil d’alerte fixé pour une fausse alerte tous les vingt ans

En ce qui concerne les 65-74 ans (figure 5.16), à nouveau il n’y a pas de franchissement du seuil d’alerte sur toute la période. Pour autant, il se passe quelque chose de différent : outre que la courbe CUSUM pour la détection d’un changement à la hausse s’approche nettement du seuil d’alerte fin 2022, entre début 2019 et fin 2022, elle ne revient plus au niveau 0, contrairement à ce qui se passe pour les 35-54 ans ou les 55-64 ans. Ceci pourrait laisser penser que, peut-être, un changement a eu lieu ... et suggère qu’il faudrait étudier, de façon plus générale, le nombre d’évènements moyens sans retour à 0 du processus CUSUM. Cela invite aussi à considérer que les résultats de la méthode mise en œuvre pourraient être suivis simultanément avec différents niveaux de changement (et pas seulement 10% à la hausse ou à la baisse comme ici).

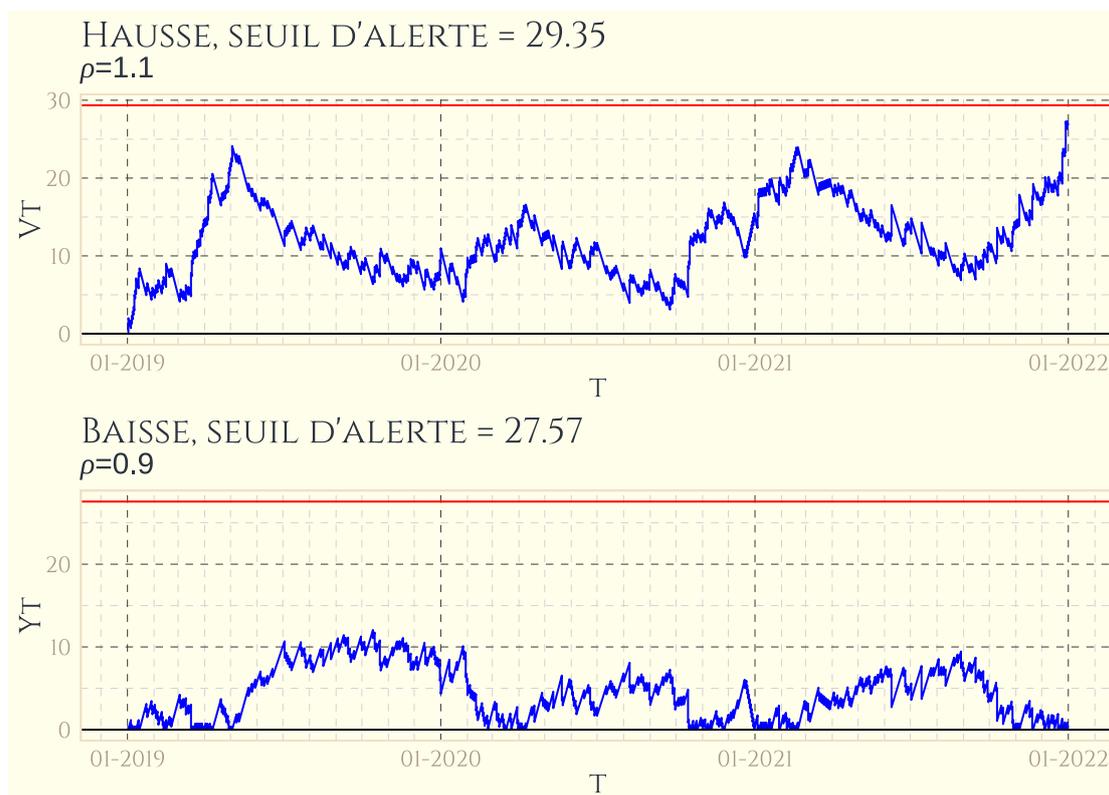


FIGURE 5.16 – Évolution du processus CUSUM pour le décès des 65-74 ans depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans

Pour les 75-84 ans, une première alerte à la hausse retentit en avril 2021, puis une autre en septembre de la même année. Entre les deux alertes, la courbe reste très proche du seuil d'alerte. Le dernier contact de la courbe avec le niveau 0 date de fin janvier ou début février 2021. L'aspect du processus entre début 2019 et ce dernier contact avec le niveau 0 exclut donc que la hausse constatée ensuite soit liée à une sous-estimation, par le modèle qui donne l'intensité de référence, du niveau de sinistres pour cette tranche d'âge particulière. Ici, une étude particulière des sinistres pour cette tranche d'âge sera nécessaire pour tenter de comprendre ce qui aurait pu provoquer une telle hausse sur l'année 2021.

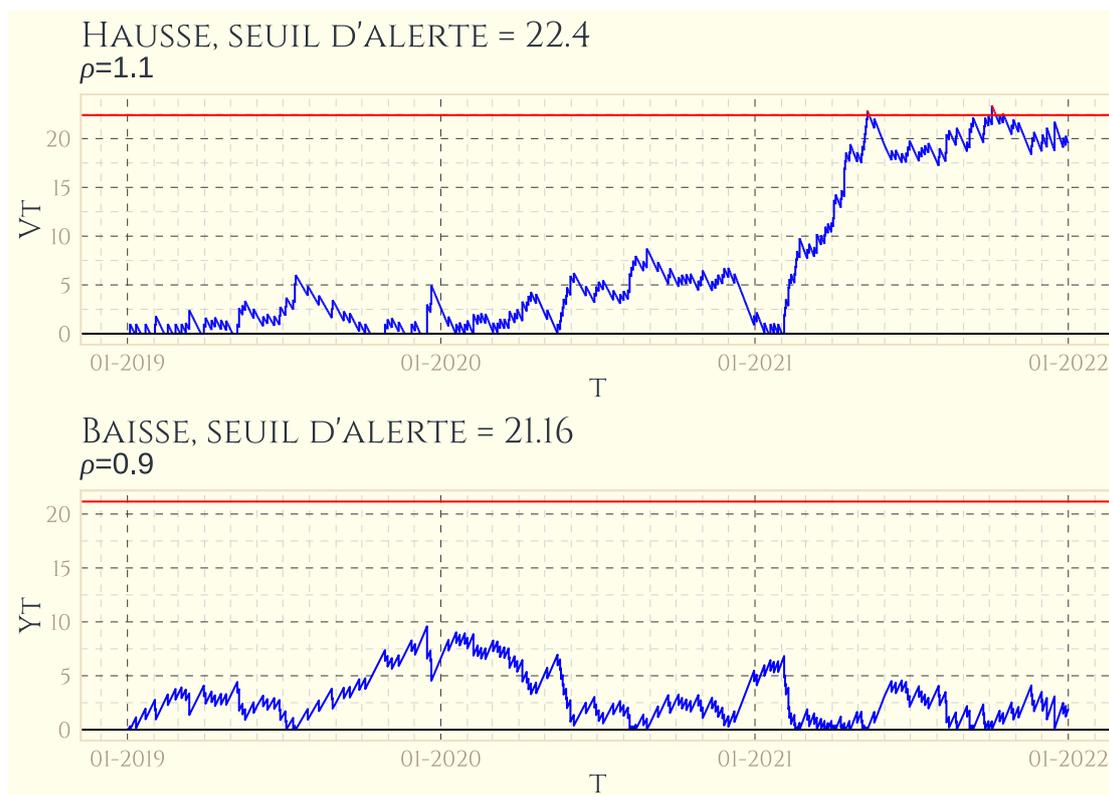


FIGURE 5.17 – Évolution du processus CUSUM pour le décès des 75-84 ans depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans

Selon le caractère VLR

En ce qui concerne les assurés caractérisés comme VLR, pour un seuil d'alerte fixé sur la base d'une fausse alerte tous les vingt ans, la surveillance ne conduit à aucune alarme entre 2019 et 2022 (cf 5.18).

Pour les assurés non VLR, en revanche, le seuil d'alerte est franchi de début mars à mi-mai 2022. Il pourrait tout aussi bien s'agir d'une fausse alerte, que d'une alerte liée à une modification passagère du niveau de sinistralité. L'évolution ultérieure de la courbe ne permet pas à elle seule de conclure.



FIGURE 5.18 – Évolution du processus CUSUM pour le décès des VLR depuis le 1er janvier 2019, seuil d’alerte fixé pour une fausse alerte tous les vingt ans

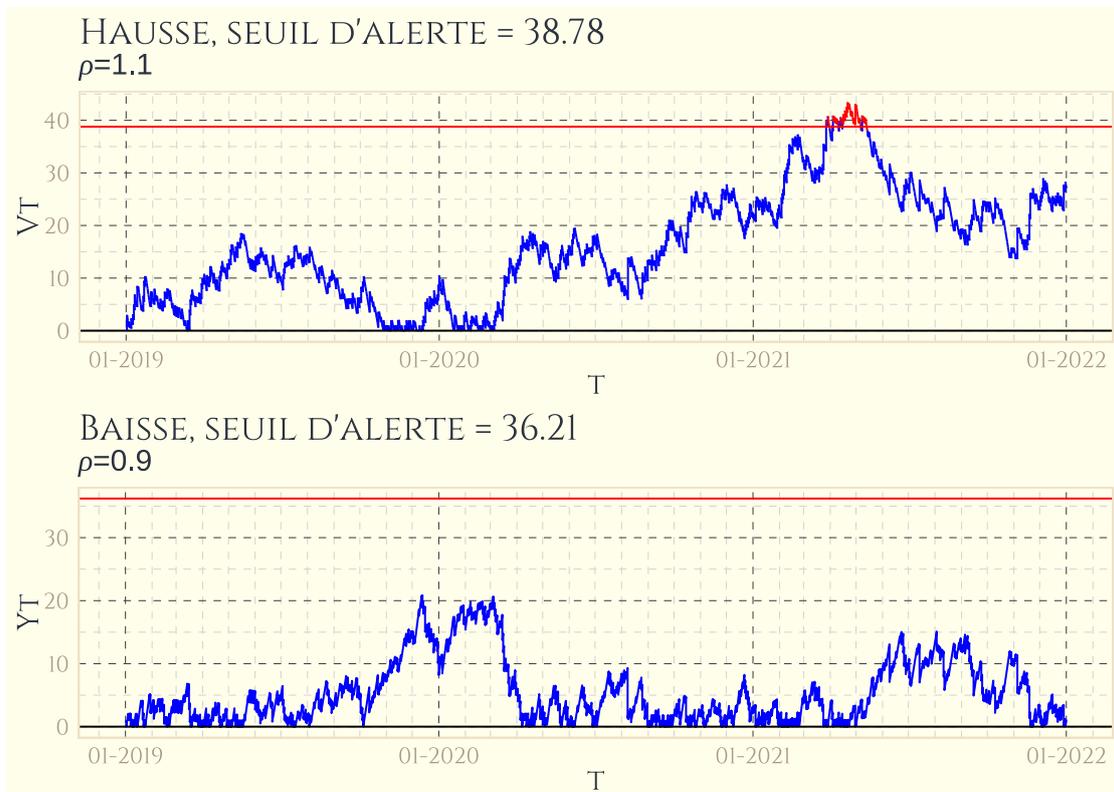


FIGURE 5.19 – Évolution du processus CUSUM pour le décès des non VLR depuis le 1er janvier 2019, seuil d’alerte fixé pour une fausse alerte tous les vingt ans

Selon la présence de surprime

Il semble légitime de se demander si les alertes remarquées jusqu'à présent seraient liées à une hausse plus particulière de la sinistralité chez les assurés surprimés, dont une partie représente les assurés les plus fragiles, et une autre les assurés dont les activités sont reconnues comme plus à risque. L'observation des figures 5.21 et 5.20 à cet égard montre en effet une différence notable entre les deux sous-populations. Rien de notable ne semble se produire du côté des non surprimés, à part peut-être une tendance à la baisse de la sinistralité qui aurait eu lieu entre juillet et novembre 2019, mais insuffisante pour déclencher une alerte. Le CUSUM pour les non surprimés amorce lui une phase de croissance courant juin 2020, jusqu'à heurter le seuil d'alerte courant 2021. A nouveau, en dehors de l'alerte sonnée, le processus n'aide pas à la compréhension des causes de cette situation. Pourrait-il par exemple s'agir de conséquences psychologiques des confinements subis à cause de la crise sanitaire ? En ce cas, les assurés surprimés seraient-ils aussi les plus sensibles psychologiquement ? Il ne semble pas, s'il faut se fier à la seconde courbe de la figure 5.21 que les confinements et précautions sanitaires importantes eussent entraîné dans cette population une baisse notable des décès, et donc il est au plus admissible de dire que la hausse n'est probablement pas due à un effet de rattrapage qui serait lié à un relâchement de ces précautions.

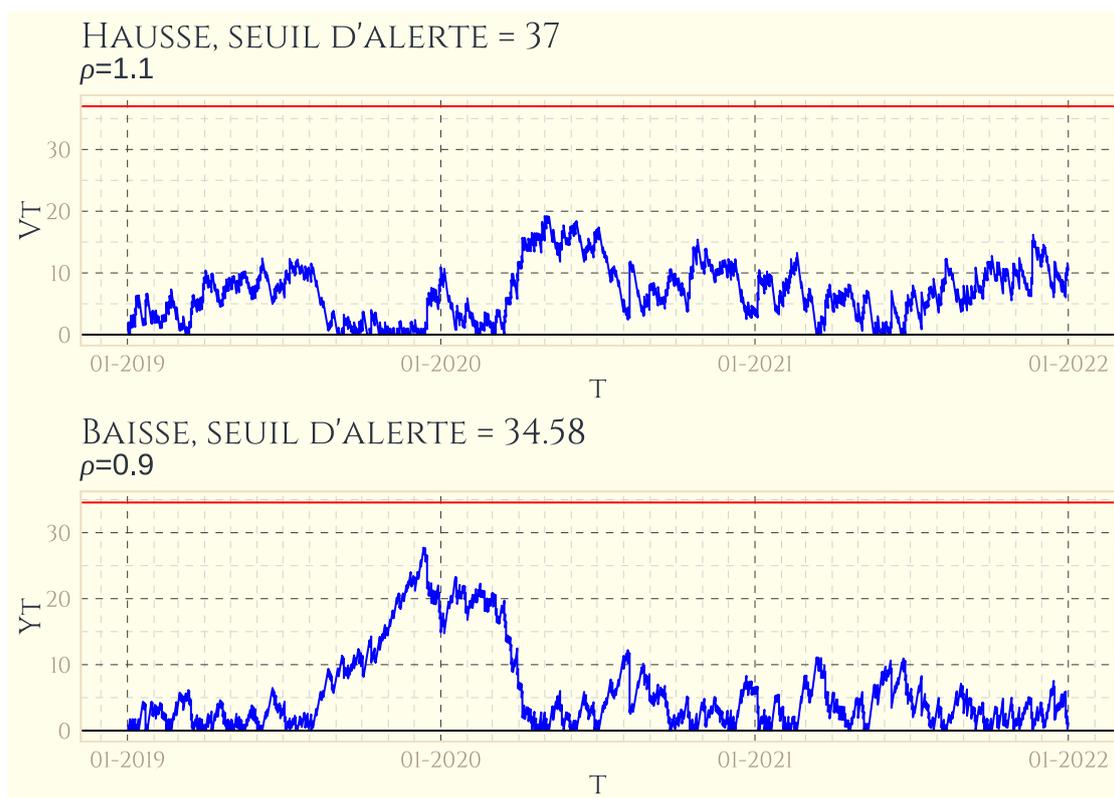


FIGURE 5.20 – Évolution du processus CUSUM pour le décès des non surprimés depuis le 1er janvier 2019



FIGURE 5.21 – Évolution du processus CUSUM pour le décès des assurés soumis à surprime depuis le 1er janvier 2019

Chapitre 6

Applications à la garantie ITT

L'application de la méthode de détection des changements d'intensité telle que décrite jusqu'à présent pourrait s'appliquer à des événements répétés, tels que l'entrée en ITT. Cependant, le choix est ici de ne s'intéresser qu'à la première ITT connue pour chaque assuré, et seulement si cette ITT a donné lieu à une prestation.

Ce choix est discutable pour plusieurs raisons. D'abord parce qu'en réalité le nombre d'assurés avec plus d'une ITT d'une durée qui dépasse la franchise, et donc qui donne lieu à prestation, n'est pas complètement négligeable. Sur la période étudiée, près de 9% des assurés qui ont subi une première interruption de travail ayant donné lieu à prestation en ont eu une deuxième. De plus, il ne faudrait pas écarter la possibilité qu'une évolution de la sinistralité sur les arrêts de travail porte justement sur une augmentation ou une diminution des nouveaux arrêts de travail pour les assurés qui auraient déjà eu un premier arrêt.

Cependant, dans les données disponibles, notamment celles qui remontent à plus de deux ou trois ans, pour des raisons d'évolution des technologies utilisées pour la gestion des sinistres et la conservation des données, les durées réelles des ITT ne sont pas particulièrement fiables ou complètes. Et par ailleurs, une hausse ou une baisse de la sinistralité sur une première ITT est déjà une information utile en soi pour comprendre la sinistralité globale liée à cette garantie. Il n'en reste pas moins qu'une étude plus complète serait souhaitable si la méthode étudiée devait être utilisée de façon usuelle par un assureur ¹.

Contrairement au cas de la garantie décès, ici ne sont considérés que les arrêts de travail qui ont donné lieu à une prestation. En ce qui concerne le décès, la garantie est de fait présente pour tous les contrats, et les rares cas où le décès ne donne pas lieu à une prestation sont très particuliers (liés à des exclusions ou bien à de fausses déclarations). Pour la garantie ITT, la situation est très différente, car nombre de sinistres déclarés ne donnent pas lieu à prestation du fait de la période de franchise. Cette franchise, en général de 90 jours, exclut de facto la prise en charge de l'assureur (et donc aussi le coût de l'assurance pour l'assuré) pour les arrêts de courtes durées. Les exclusions sont un peu plus larges (et certaines même rachetables, voir 1.5), et ne pas exclure de l'analyse les sinistres non payés du fait de ces exclusions pourrait conduire à surestimer fortement le risque réellement encouru.

1. Et il ne faut pas perdre de vue que la méthode présentée ici ne porte que sur les survenances des sinistres. Pour l'ITT, la sinistralité ne peut être étudiée complètement qu'avec une étude des durées de prises en charge et des coûts.

6.1 Intensité de référence

6.1.1 Données

Les données pour l'étude sur la garantie ITT concernent un portefeuille de contrats d'assurance emprunteur observés entre 2015 et 2018. Comme pour les données utilisées pour l'étude sur le décès, les assurés sont regroupés sur la base de leur nom, prénom, date de naissance et selon le gestionnaire des contrats. Déjà discutable dans le cas de l'étude du décès, la présence d'assurés étrangers résidents à l'étranger est exclue ici.

Expositions

Le tableau 6.1 présente les expositions selon le gestionnaire. Ici aussi, le portefeuille d'assurés est en croissance sur toute la période (croissance de 13% entre 2015 et 2018).

Par comparaison avec les lignes 1, 4 et 5 du tableau 5.1, il apparaît que l'exposition liée à la garantie ITT est présente pour 70% des expositions au décès, de façon relativement stable d'une année sur l'autre.

Le principe de calcul des expositions est le même que pour le décès, sur la base de la période de couverture par la garantie ITT. A partir de l'occurrence de la première ITT, l'assuré n'est plus compté dans les expositions.

	Système de Gestion	2015	2016	2017	2018
1	Système 1	0	0	0	358
2	Système 4	107 932	115 996	116 618	119 215
3	Système 5	4 414	7 705	7 651	7 341
4	TOTAL	112 346	123 702	124 270	126 915

TABLE 6.1 – Exposition par année d'observation

Nombre d'arrêts de travail

Le nombre d'arrêts de travail retenu est présenté dans le tableau 6.2. Comme l'exposition, il est en croissance constante, mais avec une croissance plus marquée que l'exposition (23% pour le nombre de sinistres contre 13% pour l'exposition).

	Système de Gestion	2015	2016	2017	2018
1	Système 1	0	0	0	1
2	Système 4	471	588	560	573
3	Système 5	9	23	19	20
4	TOTAL	480	611	579	594

TABLE 6.2 – Nombre de première ITT avec prestation par année d'observation

Taux bruts

Au global, le taux de sinistralité brut, défini comme le rapport annuel entre le nombre de sinistres et l'exposition, varie entre 4,272‰ et 4,939‰. L'année 2016 apparaît comme la plus sinistrée.

	Système de Gestion	2015	2016	2017	2018
1	Système 1	0,000 ‰	0,000 ‰	0,000 ‰	2,787 ‰
2	Système 4	4,364 ‰	5,069 ‰	4,802 ‰	4,806 ‰
3	Système 5	2,039 ‰	2,985 ‰	2,483 ‰	2,724 ‰
4	TOTAL	4,272 ‰	4,939 ‰	4,659 ‰	4,680 ‰

TABLE 6.3 – Taux bruts par année d'observation

Composition du portefeuille

Le tableau 6.4 présente quelques statistiques descriptives sur le portefeuille étudié. Cette sous-population de la population étudiée dans le cas du décès présente quelques traits similaires. L'âge moyen est en très légère hausse au cours des quatre années, mais inférieur de presque quatre ans à celui des garanties décès (ce qui à vrai dire n'est pas mystérieux, puisque l'ITT couvre les arrêts de travail, sauf en cas de souscription d'une extension de garantie, jusqu'à 65 ans). Le poids des assurés classés VLR est ici aussi en hausse, mais avec quatre à cinq points de plus. L'extension de garantie jusqu'à 70 ans est très peu souscrite, au contraire de l'**option plus** (rachat des exclusions *dos-psy*, voir 1.4), qui est souscrite par 10% des assurés en 2015, taux en croissance constante jusqu'à plus de 20% en 2018. La proportion d'assurés surprimés est de l'ordre de 22%, avec une légère tendance à la baisse. La proportion des femmes est elle aussi en légère baisse, quand la proportion de fumeurs connaît une baisse de deux points sur la période.

	2015	2016	2017	2018
1 Age Moyen	40,1	40,4	41	41,5
2 Proportion d'assurés classés VLR	27 %	29 %	30,1 %	31,5 %
3 Proportion d'assurés souscripteurs d'une extension de garantie jusqu'à 70 ans	0 %	0 %	0,1 %	0,1 %
4 Proportion d'assurés souscripteurs de l'option +	10 %	13,1 %	16,3 %	20,8 %
5 Proportion d'assurés surprimés	22,5 %	21,5 %	21,4 %	21 %
6 Proportion de femmes	48,1 %	48,1 %	48 %	47,9 %
7 Proportion de fumeurs	17,6 %	16,5 %	16,1 %	15,6 %

TABLE 6.4 – Statistiques descriptives du portefeuille des assurés en ITT

6.1.2 Modélisation

Comme dans le cas précédent, le principe retenu pour la modélisation est celui décrit en 4.1.1 et, de façon similaire, chaque observation correspond à une période (presque toujours un seul jour) et à un n-uplet de valeurs possibles pour les variables explicatives autres que l'exposition. Les variables explicatives utilisées sont celles-ci :

- Age (en millésime)
- Genre
- Tabagie
- Caractère VLR
- La présence ou non de surprime
- L'ancienneté du contrat
- Un indicateur d'ancienneté inférieure ou égale à 3 ans (new/old) ²

2. La proximité avec la sélection médicale est supposée être un facteur d'atténuation du risque, au moins pendant les premières années, d'où la présence de cet indicateur.

- Un indicateur d'âge avant ou après 60 ans
- La souscription d'une franchise de 180 jours (CARENCE_180)
- La souscription d'une franchise de 30 jours (CARENCE_30)
- La souscription de l'option plus (PLUS)
- La saison³
- l'exposition totale pour le n-uplet de modalités des autres variables explicatives

La variable à expliquer est le nombre des survenances d'arrêts de travail (ayant donné lieu à prestation) observées sur chaque période, et pour chaque n-uplet des modalités des variables explicatives.

Le modèle retenu est présenté dans le tableau 6.5 :

<i>Dependent variable :</i>	n.sinistre				
	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-7.1606	0.1812	-39.51	0.0000	***
age	0.0291	0.0027	10.87	0.0000	***
genreH	-1.3092	0.1672	-7.83	0.0000	***
est_fumeur	0.2529	0.0407	6.22	0.0000	***
avec_surprime	0.5320	0.0356	14.94	0.0000	***
CARENCE_180	-0.7553	0.2243	-3.37	0.0008	***
CARENCE_30	0.9855	0.1008	9.78	0.0000	***
PLUS	0.1958	0.0394	4.97	0.0000	***
age.60avant60	0.9368	0.1185	7.90	0.0000	***
VLR	-0.7098	0.0404	-17.56	0.0000	***
saisonete	-0.2476	0.0455	-5.44	0.0000	***
saisonhiver	-0.1909	0.0440	-4.34	0.0000	***
saisonprintemps	-0.1259	0.0436	-2.89	0.0039	**
anciennete	-0.0544	0.0050	-10.81	0.0000	***
age :genreH	0.0229	0.0038	6.09	0.0000	***
Observations	319 403				
Log Likelihood	-15 835,78				
Akaike Inf. Crit.	31 701,56				

Note : *p<0.1; **p<0.05; ***p<0.01

TABLE 6.5 – Modèle retenu pour l'ITT

Comme précédemment, la nature du modèle nous autorise quelques remarques. Tout d'abord, toutes choses étant égales par ailleurs, les hommes ont un risque d'avoir une ITT réduit de près de 73% par rapport aux femmes. Cependant, chaque année supplémentaire (mais toujours pour la même ancienneté de contrat) ajoute, après 60 ans, 0,23 fois plus de risque aux hommes qu'aux femmes. Avant 60 ans, il faut tenir compte du facteur 0,9368 dans le modèle, c'est à dire que toutes choses étant égales par ailleurs, un assuré aura, s'il a moins de 60 ans, 2,55 fois plus de chances d'avoir un arrêt de travail pris en charge qu'un autre assuré ayant un an de moins. Sans surprise, la souscription d'une franchise de 180

3. Saison approximative : hiver = décembre, janvier, février ; printemps = mars, avril mai ; été = juin, juillet, août, et automne = septembre, octobre, novembre

jours (qui permet à l'assuré de payer moins cher) fait diminuer le risque d'un sinistre pris en charge (l'intensité moyenne est multipliée par 0,47 par rapport à l'intensité de base), quand celle d'une franchise de 30 jours, plus onéreuse pour l'assuré que la franchise de base qui est de 90 jours, multiplie le risque d'un sinistre pris en charge par presque 2,7. La souscription de l'option plus, selon ce modèle, multiplierait le risque par 1,22. Et la saison où le risque d'ITT est le plus fort est l'automne (puisque, pour toutes les autres saisons, le coefficient est négatif, et donc son exponentielle inférieure à 1).

6.2 Détection des changements d'intensité

6.2.1 Mise en Œuvre

Tous les résultats qui sont présentés dans la suite se réfèrent à un seuil d'alerte conçu pour une fausse alerte en moyenne tous les vingt ans.

Pour le reste, la mise en œuvre est strictement identique à celle présentée dans le chapitre relatif à l'étude des décès. Là encore, la détection porte sur un changement à la hausse ou bien à la baisse de 10% par rapport à l'intensité de base (intensité modélisée par les observations sur la période 2015-2018). Ce choix est bien sûr discutable et, là aussi, une surveillance simultanée de plusieurs niveaux pourrait être envisagée.

6.2.2 Résultats

Population globale

Les graphiques 6.1 présentent deux aspects particulièrement notables : un dépassement prolongé du seuil d'alerte à la hausse, et un dépassement très marqué et radical du seuil d'alerte à la baisse en fin d'année 2021.

Le dépassement à la baisse est sans conteste de peu d'intérêt. En effet, avec une franchise, dans la plupart des cas, de trois mois, les trois derniers mois de toute analyse de ce type ne peuvent que présenter une baisse d'intensité : il n'y a plus aucun sinistre pris en charge depuis trois mois (ou plus précisément presque plus, puisque quelques assurés ont souscrit une option pour que la franchise soit portée à 30 jours) ! D'une manière plus générale, les sinistres déjà survenus mais non encore connus viennent aussi ajouter à cette baisse apparente mais en réalité inexistante de la sinistralité (voir 7.5).

Le dépassement à la hausse, observé de manière franche à partir de novembre 2019, mais avec des atteintes du seuil d'alerte dès le mois d'août de la même année est certainement plus probant. Si ce système d'alerte avait été en place depuis le premier janvier 2019, avec ce niveau d'alerte, le dépassement du seuil aurait pu être constaté au moins trois mois après, soit à partir de novembre ou décembre 2019. Les premières questions qui se seraient alors posées, une fois l'alerte sonnée, sont *S'agit-il d'une fausse alerte ?* et *Quelles causes pourraient expliquer une hausse de la sinistralité à ce moment-là ?* En décembre 2019, la pandémie de COVID-19 n'avait pas encore commencé, et à cette époque nul n'aurait songé à lui attribuer cette alerte. Aujourd'hui encore, il paraît peu probable que ce dépassement de seuil soit lié au SARS-COV2. Le premier patient officiel en France est en effet une personne rentrée de Chine le 22 janvier 2020. Une hypothèse émise par le service sinistre d'AFI-ESCA⁴, suggère qu'une hausse des arrêts de travail pourrait avoir été une conséquence du mouvement des Gilets Jaunes. Ce mouvement a pris son essor

4. Lors d'un échange informel en avril 2022 autour des courbes présentées dans ce mémoire. Cette hypothèse, émise sur le mode *et si c'était ça ?* n'est pas étayée à ce jour.

en octobre 2018, et *les heurts* (c'est à dire en l'occurrence l'extrême violence de part et d'autre) entre manifestants et forces de l'ordre ont provoqué 11 décès, 1 944 blessés au sein des forces de l'ordre et 28 500 blessés civils (voir [Wikipedia, 2022]). Une étude plus approfondie des motifs des arrêts de travail pris en charge par l'assureur serait nécessaire pour conforter ou invalider cette hypothèse, et c'est certainement le lancement de ce type d'étude qui aurait été la première décision si l'alarme avait retenti en décembre 2019 pour un dépassement de seuil d'alerte courant août de la même année. Puis se serait éventuellement posée la question de la poursuite dans la durée d'une telle hausse de sinistralité, et donc d'éventuelles modifications de tarif, de sélection médicale ou de niveau de garantie.

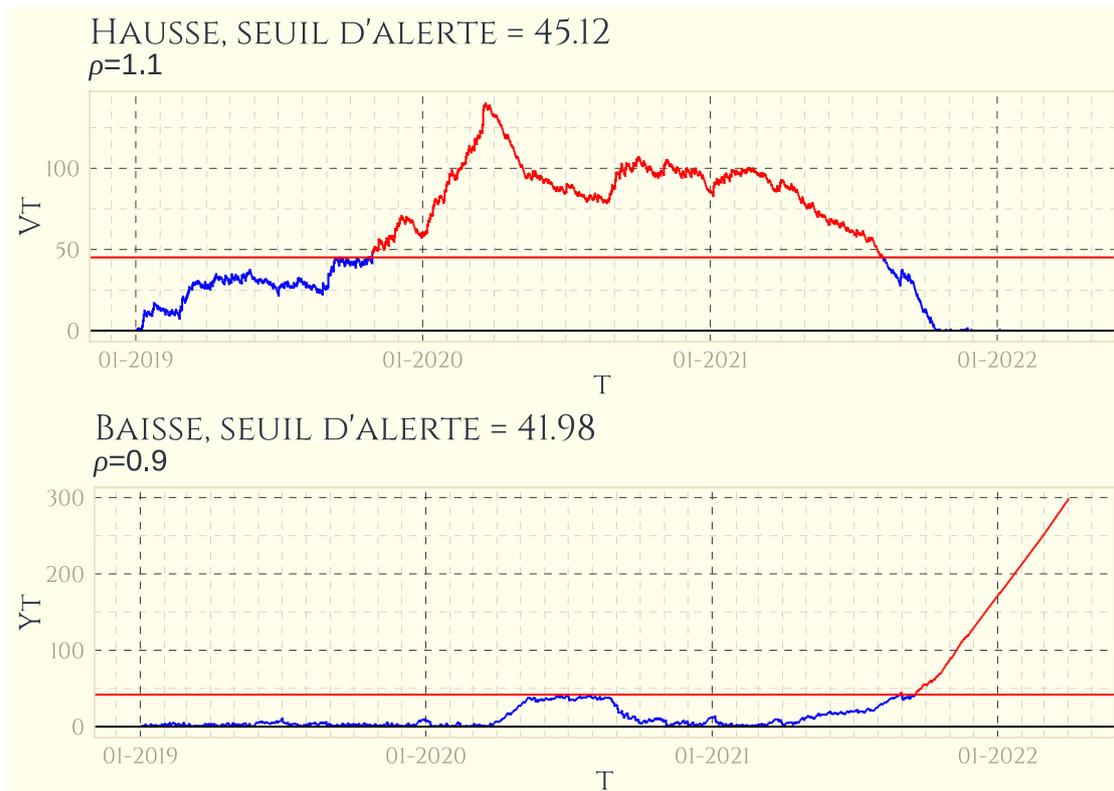


FIGURE 6.1 – Évolution du processus CUSUM pour l'ITT (population générale) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans

L'observation conjointe des processus CUSUM pour des détections de changements à la hausse et à la baisse appelle aussi un autre commentaire. En effet, la courbe du CUSUM qui sert à détecter un changement à la baisse frôle le seuil d'alerte, alors même que celui qui sert à détecter un changement à la hausse est toujours largement au-dessus du seuil d'alerte pour la hausse. Ceci peut paraître contre-intuitif, mais en réalité rien n'interdit qu'une alerte à la baisse retentisse alors que le processus pour la hausse est au-dessus de son seuil d'alerte, et vice-versa. En effet, lorsqu'il y a une hausse de l'intensité, et que le nouveau niveau de sinistralité se maintient sur la durée, le processus V_T aura tendance à croître de façon continue, même si l'intensité se stabilise mais qu'elle reste à un niveau suffisamment supérieur à l'intensité de référence. Selon la durée de cette phase de croissance de la courbe, le niveau atteint peut être très élevé, à tel point même que si une période de baisse marquée de l'intensité (baisse par rapport à l'intensité de référence), a lieu, l'alarme à la baisse peut retentir avant que le processus à la hausse soit repassé en deçà de son seuil d'alerte. D'où cette remarque, que ne doit pas négliger celui qui tente d'interpréter

les courbes obtenues avec cette méthode, que, sans même parler de fausse alerte, *être au-dessus de la courbe à l'instant t* ne signifie pas forcément que l'intensité est à cet instant t de ρ fois l'intensité de référence, et que cette situation peut être la conséquence d'une hausse passée mais qui a duré suffisamment pour qu'un retour à l'intensité de référence n'ait pas encore permis de repasser sous le seuil d'alerte. A cet égard, la question du point de départ se pose. Ne faudrait-il pas, en cas d'alerte en t , faire repartir le processus à 0 ? A vrai dire, pas nécessairement, car ce qui compte, une fois l'alerte sonnée, c'est soit que la courbe continue son ascension, auquel cas il faut penser que le nouveau taux de sinistralité se maintient durablement, soit au contraire que le CUSUM entame un retour vers des niveaux plus bas, signe que la nouvelle intensité détectée ne s'est pas maintenue sur le long terme. Ici par exemple, juste avant la fin février le processus pour la détection des changements à la hausse amorce une décroissance qui se poursuit jusqu'en avril, et pendant ce laps de temps le processus qui détecte les changements à la baisse décolle et va frôler son seuil d'alerte.

Plutôt que chercher à voir un effet de la pandémie dans la hausse détectée, il est concevable qu'une hausse préalable, déjà entamée dès 2019, des arrêts de travail longs soit à l'origine de l'alerte fin 2019, et que le premier confinement ait au contraire entraîné une baisse des survenances.

Par sous-groupes

Les processus CUSUM pour plusieurs sous-groupes ont été testés. Dans la plupart des cas, les courbes sont très similaires à celles de la population générale, avec parfois un décalage du moment de l'alerte. Ci-dessous ne sont présentés que les cas où une différence remarquable apparaît par rapport à la population générale pour au moins un des sous-groupes considérés.

Selon la présence de surprimes

Les graphiques obtenus pour les seuls assurés sans surprime sont très ressemblants à ceux de la population générale (figure 6.2) et appellent donc les mêmes commentaires et les mêmes interrogations.

En ce qui concerne les assurés surprimés, les courbes (figure 6.3) sont extrêmement différentes. Ici il n'y a pas d'alerte à la hausse en 2019, et un seul et bref épisode de franchissement du seuil à la hausse entre février et mars 2020. Si l'hypothèse d'un lien entre le mouvement des Gilets Jaunes et la hausse détectée fin 2019 dans la population générale s'avérait crédible, il faudrait se demander pourquoi les assurés surprimés auraient échappés au phénomène général. Serait-ce parce que ces assurés, en plus mauvaise santé, auraient moins participé à ce mouvement ? C'est oublier que les surprimes ne sont pas toutes médicales, certaines sont sportives, d'autres professionnelles. Là encore, seule une étude détaillée des causes de sinistres pourrait apporter des éléments de réponse. De même, pour ces assurés surprimés, il serait tentant d'interpréter le franchissement du seuil comme lié à la diffusion en France du SARS-COV2, et cette fois l'adéquation calendaire avec le premier pic de mortalité liée au SARS-COV2 en France semble plus forte.

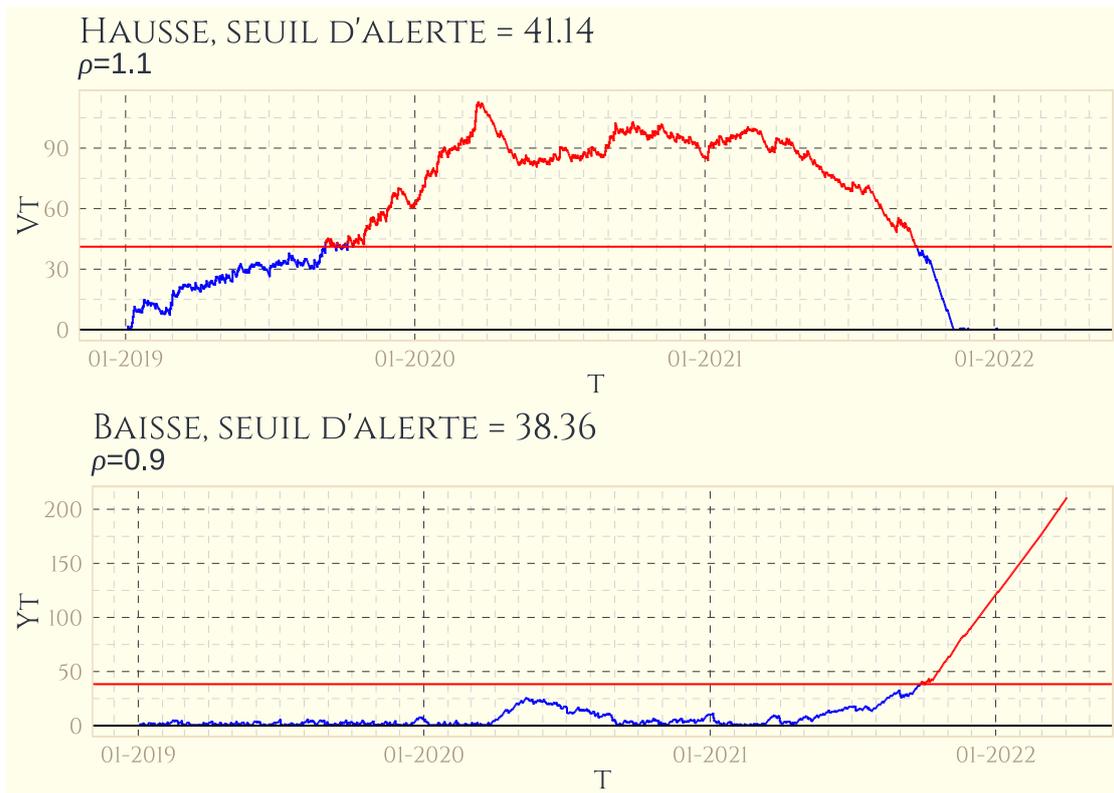


FIGURE 6.2 – Évolution du processus CUSUM pour l'ITT (assurés sans surprime) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans



FIGURE 6.3 – Évolution du processus CUSUM pour l'ITT (assurés avec surprime) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans

Selon la tabagie

Les courbes sur la partie des assurés qui se sont déclarés non-fumeurs sont très voisines de celles des non surprimés et de celle de la population générale (cf figure 6.4).

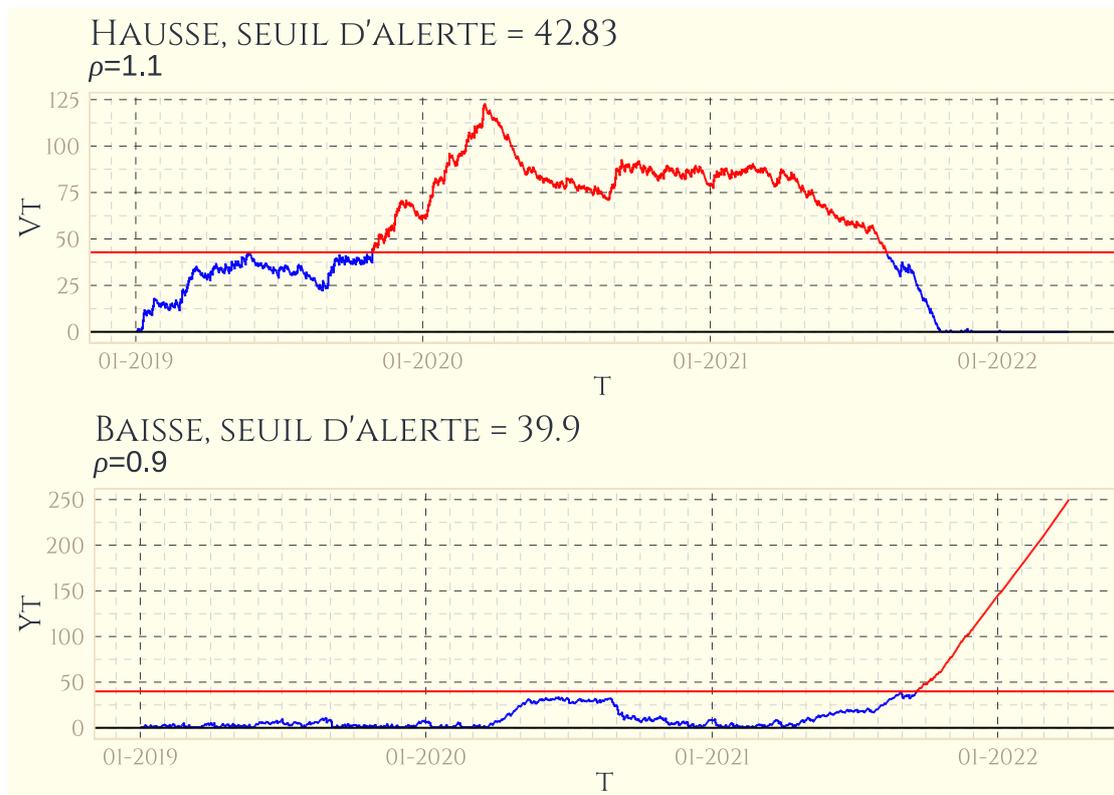


FIGURE 6.4 – Évolution du processus CUSUM pour l'ITT (assurés non-fumeurs) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans

Au contraire, chez les fumeurs, le dispositif de détection mis en œuvre n'aurait pas retenti une seule fois entre 2019 et 2022 (à l'exclusion du dépassement à la baisse sans intérêt en fin d'année 2022). Il serait éminemment hasardeux de lier ce fait à la théorie, vraisemblablement sans fondement scientifique, qui a couru un temps selon laquelle les fumeurs auraient été moins sensibles au SARS-COV2 que les non-fumeurs.

Par contre, il est possible ici de s'interroger sur le choix du niveau d'alerte. En effet, les seuils sont fixés pour une fausse alerte tous les vingt ans en moyenne, et le nombre moyen de sinistres sur vingt ans chez les fumeurs est estimé sur la base de ce qui a été observé sur 2015-2018. Mais il a été noté plus haut (cf. 6.1.1) que la proportion de fumeurs était en baisse. Selon les chiffres fournis, l'exposition pour les fumeurs passerait de 20 000 en 2017 ($124\,270.8 * 16\%$) à 19 799 en 2018. A supposer que la baisse se soit poursuivie, le nombre moyen de sinistres sur vingt ans, même sans changement dans l'intensité des sinistres, pourrait en réalité être plus bas, et donc les seuils d'alerte un peu trop hauts.

D'autre part, en dehors de l'absence d'alerte, l'aspect du CUSUM pour la détection d'un changement à la hausse est un peu différente de celle de la population des non-fumeurs. La courbe ici revient régulièrement au niveau 0 jusqu'en juillet 2019, alors qu'elle n'atteint pas une seule fois ce niveau chez les non-fumeurs entre janvier 2019 et fin 2021 (et encore ce retour à 0 est peut-être trop récent pour être interprétable). Tout se passe comme si la hausse suspectée des arrêts longs sur 2019 n'avait touché les fumeurs qu'à partir du milieu de l'année. Il peut très bien s'agir d'un simple effet de fluctuations statistiques naturelles.



FIGURE 6.5 – Évolution du processus CUSUM pour l'ITT (assurés fumeurs) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans

Chapitre 7

Chausse-trapes

7.1 Biais cognitifs

7.1.1 La courbe qui monte (ou qui descend)

Comme l'outil étudié dans ce mémoire produit un résultat graphique, l'esprit humain (en particulier l'esprit d'un actuaire ou d'un dirigeant d'entreprise d'assurance) a tendance à interpréter les courbes qui montent comme le signe d'une croissance continue de ce qu'il cherche à analyser. Ici par exemple, où le processus CUSUM est utilisé pour étudier une hausse de la sinistralité, nos habitudes tendent à nous faire interpréter une courbe qui semble ne pas cesser de croître comme le signe que la situation ne fait qu'empirer, et que la dégradation du niveau de sinistralité ne fait que s'accroître.

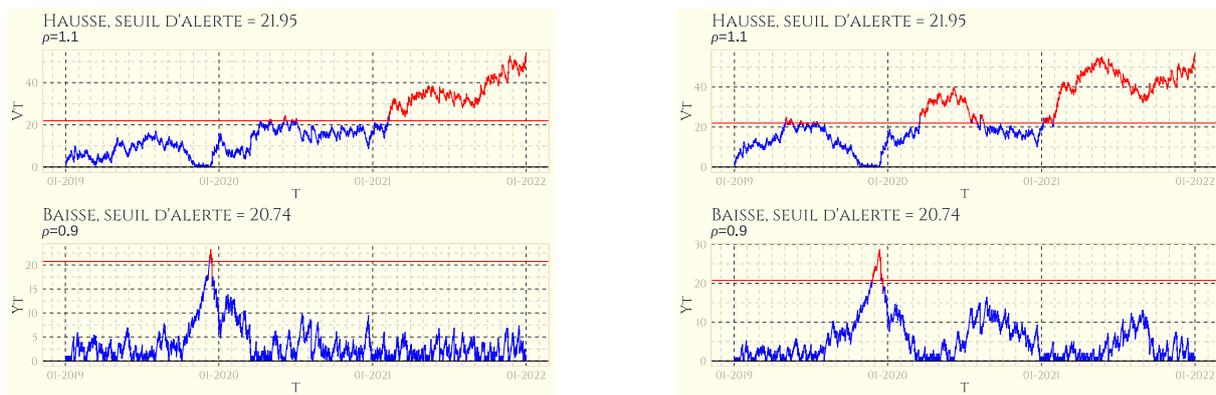
Cependant, la construction du processus (voir chapitre 3) est telle que, si l'intensité du processus de Poisson observé évolue à la hausse (i.e. ρ est supérieur à 1) mais qu'il n'y a plus d'évolution de cette intensité après ce changement, la courbe obtenue continuera de monter. C'est par exemple le cas figure 3.7 : après le changement, la courbe continue à croître. Pour autant, cela ne signifie en rien que l'intensité du processus observé est de plus en plus élevée.

7.1.2 La courbe qui stagne

De façon tout à fait similaire, une stagnation de la courbe pourrait nous faire croire à une stabilisation de l'intensité qui fait l'objet de notre attention. Ainsi, un *plateau* pourrait facilement, mais à tort, être interprété comme un retour à l'intensité initiale. En réalité, toujours par construction, si le processus revient à l'intensité de référence, la courbe obtenue doit nécessairement régulièrement revenir au niveau 0. Dans le cas contraire, si un plateau est observé sans retours réguliers au niveau 0, cela signifie que l'intensité réelle se situe quelque part entre l'intensité de référence et cette même intensité multipliée par ρ .

7.1.3 La courbe qui reste au-dessus du seuil d'alerte

Dans le même ordre d'idée, il convient de se prémunir contre la tentation d'interpréter une courbe CUSUM au-dessus du seuil d'alerte comme le signe que l'intensité est au niveau de celle qui doit déclencher l'alerte. Un exemple concret est détaillé en 6.2.2.



Modèle avec saisonnalité

Modèle sans saisonnalité

FIGURE 7.1 – Evolution du processus CUSUM pour le décès (assurés en France) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les trois ans

7.2 Erreurs de modèles

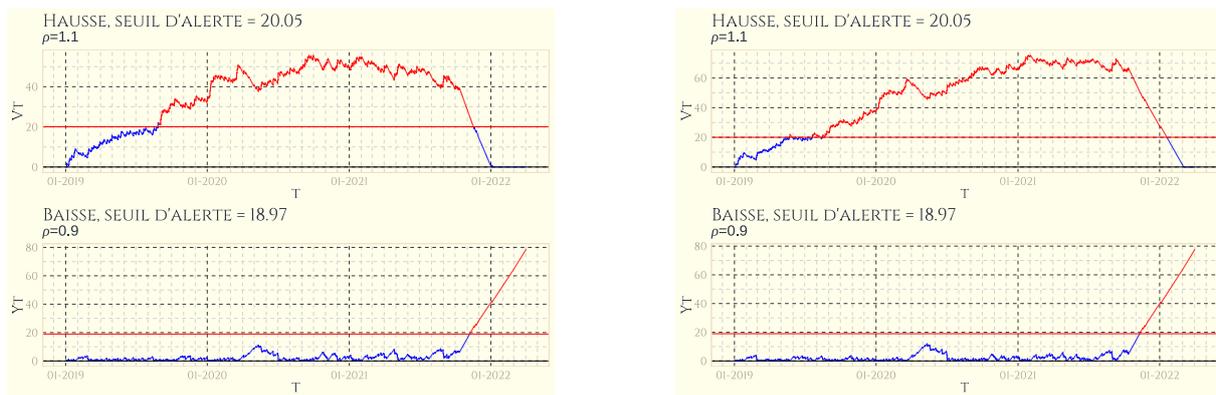
Avant de décréter que la méthode étudiée est valable et utilisable avec usufruit par les assureurs, il est impératif de s'assurer qu'elle est, d'une certaine façon, robuste. En particulier, il est important de vérifier que les analyses obtenues ne sont pas trop sensibles aux choix de modélisations qui président aux calculs des intensités théoriques sous \mathbb{P}_∞ . Sans prétendre répondre de façon définitive à cette question, tester quelques modèles plus ou moins judicieux peut déjà apporter quelques éléments de réflexion à ce sujet.

A titre d'exemple, la figure 7.1 présente le graphique obtenu à partir d'un modèle qui ne prend pas en compte la saison dans la modélisation de la sinistralité sur 2015-2018, en comparaison au graphique déjà présenté en 5.2.2, basé sur le modèle qui prend en compte la saisonnalité. Les différences d'aspects entre les deux courbes à la hausse et les deux courbes à la baisse sont déjà très significatives de la sensibilité de cette méthode aux choix de modélisations. Les alertes qui seraient obtenues dans ces deux cas sont suffisamment différentes pour que l'on puisse dès à présent insister sur l'importance d'une modélisation la plus pertinente possible. Dans la mise en œuvre pratique de la méthode présentée dans ce mémoire, la modélisation doit donc prendre une place cruciale et faire l'objet d'une attention particulière.

7.3 Qualité de données

De même que la sensibilité au modèle étudiée dans la section précédente, la sensibilité de la méthode à la qualité des données ne peut pas être négligée. Même si la qualité des données est une préoccupation constante dans tout ce qui concerne l'assurance, la complexité et la diversité des systèmes informatiques, les évolutions historiques de ces systèmes et des caractéristiques des contrats, et la contribution humaine inévitable à la présence d'erreurs - contribution qu'il faut reconnaître non pour la dénoncer mais pour la comprendre et la traiter au mieux - conduisent et vraisemblablement continueront de conduire à la nécessité d'une circonspection soutenue quant à cette qualité.

Comme précédemment, il n'est pas question ici de prétendre apporter une réponse définitive à ce problème dans le cadre de la méthode étudiée, mais seulement d'apporter un premier éclairage à ce sujet. A titre d'exemple, les graphiques de la figure 7.2 présentent le processus CUSUM pour les seuls assurés VLR, d'abord sur les données de l'étude, puis



CUSUM sur les données étudiées

CUSUM sur les données modifiées

FIGURE 7.2 – Evolution du processus CUSUM pour l'ITT pour les assurés classés VLR, depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans

sur un jeu de données modifiées. Ici, la modification a consisté à changer le caractère VLR pour 5% des pseudo-observations. Si les courbes avant et après modifications présentent un évident air de famille, les données modifiées sonnent une alarme en avril 2019, alarme qui n'existe pas avec les données originales. Or dans les deux cas, le modèle qui sert de référence est basé sur des données non modifiées. Vu la constatation faite en 7.2, un modèle ajusté sur des données elles aussi entachées d'erreurs aurait pu conduire à des écarts plus inquiétants encore.

7.4 Le problème de l'exposition

Le choix d'un critère de Lorden modifié (cf. 2.2.3), qui permet d'utiliser un nombre de sinistres moyens plutôt qu'une durée comme base pour le calcul des seuils d'alerte, et qui autorise les simplifications apportées par le principe du changement de temps (cf. 4.3.1), est aussi un point clé dans la démonstration de l'optimalité du processus CUSUM (voir [El Karoui et al., 2017]) pour la détection des changements d'intensité dans les processus de Poisson. Cependant, si le choix est de définir un seuil d'alerte tel qu'il n'y ait pas plus d'une fausse alerte en moyenne tous les vingt ans, et que le nombre de sinistres correspondant est estimé sur une période passée (2015-2018 dans les deux études de cas proposées dans ce mémoire), rien ne dit que ce nombre de sinistres correspondra toujours à une fausse alerte tous les vingt ans dans les années qui suivent. Il suffit d'imaginer le cas extrême où l'exposition doublerait, toute chose étant égale par ailleurs, pour comprendre qu'alors le seuil d'alerte ne correspondrait plus à une fausse alerte tous les vingt ans, mais à une fausse alerte tous les dix ans.

Ce problème du choix d'un seuil d'alerte adéquat pour un portefeuille dont l'exposition varie sera par exemple particulièrement sensible pour des études par générations de contrats (études qui pourraient être pertinentes par exemple pour évaluer les effets d'une modification des conditions générales par rapport à des générations de produits antérieures). Si une étude doit être faite pour les contrats souscrits en 2015, la première année l'exposition sera très probablement inférieure à l'exposition totale de la seconde, car les contrats auront pris effet tout au long de l'année 2015, puis, les années suivantes l'exposition ne fera que décroître. Une solution à envisager pourrait être un rééchantillonnage des données pour

obtenir une exposition constante (il serait préférable en ce cas d'étudier la variabilité des courbes obtenues et d'étudier des intervalles pour les temps d'alertes sous l'effet des fluctuations d'échantillonnage).

7.5 Influence des sinistres survenus mais non déclarés

Le graphique 6.1 montre dans les derniers mois un net et franc franchissement du seuil d'alerte à la baisse. Cependant, dans la plupart des cas, l'ITT est assortie d'une franchise de quatre-vingt-dix jours, soit environ trois mois. De plus ne sont pris en compte dans l'étude sur l'ITT que les sinistres réglés, donc qui ont nécessairement déjà duré au moins trois mois. Vu du premier janvier 2022, le dernier sinistre qui peut être pris en compte remonte à trois mois, et toute interprétation de la courbe au-delà du premier octobre 2021 serait donc erronée.

Plus généralement, les sinistres survenus mais non encore connus au moment de l'analyse peuvent toujours conduire à un franchissement par le processus CUSUM du seuil pour la détection d'une baisse. Il convient donc dans toute analyse de ce type d'avoir, autant que faire se peut, une idée des délais entre la survenance des sinistres et leurs déclarations.

Tenter de palier à cet inconvénient par une estimation des sinistres survenus mais non connus selon les méthodes habituelles serait une mauvaise idée dans ce contexte. En effet, ces méthodes utilisent la connaissance du même phénomène dans le passé. Ces estimations supposent donc que tout se passe comme avant, alors même que le but est de détecter que quelque chose ne se passe plus comme avant.

A contrario, une alerte à la hausse peut être prise en considération, même si elle est très proche du moment présent. Les sinistres survenus mais non encore connus à ce moment ne pourraient à la rigueur que faire sonner l'alerte plus tôt. Mais, toujours pour les mêmes raisons, une absence d'alerte à la hausse dans les derniers temps (à estimer à l'aune de ce qui est connu sur le délai - qui peut être différent pour chaque type de sinistre - généralement observé entre la survenance d'un sinistre et sa déclaration) ne peut pas être d'emblée considérée comme un signe que la sinistralité n'a pas connu, dans les derniers temps, une augmentation. Il se pourrait juste qu'il n'y ait pas encore assez d'information pour le découvrir.

Conclusion

D'un point de vue technique, ce travail montre que la mise en œuvre d'un suivi de sinistralité (ou de tout autre processus de comptage) et d'un système d'alerte basés sur le processus CUSUM est tout à fait réalisable. Une fois passés les quelques écueils particuliers (récupération des données avec une bonne qualité, calculs des expositions, et modélisation pertinente pour l'intensité de base) qui pourraient ralentir la (ou les) première(s) mise(s) en œuvre, il ne devrait pas y avoir de blocage particulier. Prendre un peu de temps au préalable pour bien comprendre les fondements mathématiques (qui restent relativement simples tant qu'il ne s'agit pas de comprendre la démonstration de l'optimalité du processus vis-à-vis du critère de Lorden modifié) est évidemment souhaitable.

La question des cas où l'usage de cette méthode, telle que présentée ici, serait utile appelle une réponse assez simple : à priori, tous les cas où des événements sont comptés peuvent faire l'objet d'un tel système de surveillance. Les enjeux d'exposition, d'exhaustivité et de qualité des données viendront peut-être tout de même limiter le champ d'application. Mais, pour une compagnie d'assurance, les applications en actuariat peuvent être nombreuses - par exemple pour un suivi des rachats, sujet non abordé ici mais évidemment présent à l'esprit de l'auteur alors que la loi Lemoine est sur le point d'entrer en vigueur -, et sans aucun doute en dehors des problématiques purement actuarielles, d'autres sujets pourraient se prêter à des suivis de ce type. Les pannes informatiques, les réclamations des assurés, les demandes d'avenants, les erreurs de saisies, les appels téléphoniques (un exemple avec prise en compte de la saisonnalité est étudié dans [Laub et al., 2020]), et sans doute d'autres événements pourraient mériter la mise en œuvre d'un tel suivi.

Une fois évacués les problèmes purement techniques, restent d'importantes questions. Tout d'abord, celle du niveau de changement à surveiller. Les 10% utilisés dans les deux études de cas présentées dans ce document sont tout à fait discutables. Pour beaucoup d'assureurs, une hausse de 10% de la sinistralité serait quelque chose de l'ordre de la catastrophe. D'un autre côté, une alerte signalant une hausse de cette ampleur ne signifie pas que ce nouveau niveau de sinistralité se prolonge. Par ailleurs, il est courant d'appliquer aux tarifs un chargement technique de quelques pourcents au titre la prudence et pour s'assurer de pouvoir soutenir les fluctuations que connaîtra naturellement la sinistralité même si son intensité ne change pas. La tentation de proposer une surveillance simultanée pour plusieurs niveaux de changement est forte, mais il faut alors se demander quel sera, parmi tous ces niveaux observés, celui qui déclenchera une alerte digne d'être signalée, et qui entraînera des études supplémentaires, et peut-être une révision des produits ou une renégociation des traités de réassurance.

Ensuite vient le problème du niveau d'alerte. Il semble, à la lumière du premier cas présenté ici, que partir d'une fausse alerte tous les trois ans expose à des alertes trop fréquentes. Mais avec une fausse alerte tous les vingt ans, n'existe-t-il pas de risque de ne

pas détecter, ou de détecter trop tard, un changement qui aurait réellement eu lieu, parce que le seuil d'alerte serait trop haut ?

Reste aussi ouverte la question de l'interprétation qui peut être faite des courbes obtenues et des alertes déclenchées. Les deux études de cas ont montré au moins que l'envie, voire le besoin, très humain, d'explications entraîne facilement l'observateur vers des hypothèses qui risqueraient fort de prendre force de vérité pour lui, parce que *c'est tellement évident !*, avant même que des analyses particulières sur les événements étudiés puissent venir étayer ou non ces hypothèses. La force du biais de confirmation est tel que cet écueil n'est pas à négliger. La plus élémentaire précaution qui peut être prise à ce sujet est assez simple : il s'agit de s'assurer que plusieurs personnes contribuent à l'analyse des courbes. Si une personne seule est chargée de ce travail et d'en tirer des conclusions, le risque d'une erreur d'interprétation ou qu'une mauvaise intuition amène une mauvaise conclusion sera nécessairement plus élevé.

Enfin, à supposer que toutes les questions précédentes aient été résolues de façon satisfaisante, que se passera-t-il lorsqu'une alerte retentira ? Comment s'assurer qu'il ne s'agit pas d'une fausse alerte ? L'étude du tracé du processus CUSUM pourrait peut-être apporter quelques indications à ce sujet : la courbe est-elle régulièrement revenue au niveau 0 avant que le niveau d'alerte soit atteint ? Ou bien au contraire cela fait-il longtemps que le niveau 0 n'a plus été atteint ? Une analyse des **temps de vol**¹, c'est à dire du temps (ou bien du nombre de sinistres) moyen entre deux passages par le niveau 0, sous l'hypothèse d'une absence de changement, pourrait-elle aider à répondre à cette question (cela ne semble pas évident, puisque que le niveau atteint lors d'une précédente hausse est liée à la durée de cette hausse) ? Faudrait-il prendre en compte les alertes dans les projections pour le pilier I de Solvabilité, ou bien pour l'ORSA ? Et si oui, comment ?

Ce qui semble le plus raisonnable est de considérer que, quel que soit le type d'alerte, le processus CUSUM ne dit rien au-delà de : *il y a peut-être une hausse ou une baisse de l'intensité*, et qu'aucune interprétation ne devrait être faite, ni aucune décision ne devrait être prise suite à ce constat sans qu'une étude plus aboutie, et si possible soumise aux jugements critiques de plusieurs personnes, n'apporte des éléments d'explications indubitables.

D'un autre côté, une absence d'alerte signifierait-elle que tout va bien ? En réalité, ce calme apparent pourrait tout simplement être dû à un niveau d'alerte trop élevé, par exemple parce que l'exposition globale du portefeuille baisse, et que l'on n'a pas pris soin de tenir compte de cette baisse dans l'estimation du nombre moyen de sinistres avant une fausse alerte sous l'hypothèse d'une absence de changement.

Enfin, la méthode étudiée ici ne porte que sur les occurrences de sinistres, et, du point de vue de l'assureur, un compte de résultat ne se bâtit pas seulement sur des occurrences, mais aussi sur les coûts des sinistres. Bien que complètement laissé de côté dans cette étude, un suivi de l'évolution des coûts des sinistres ne peut absolument pas être négligé en général, et encore moins dans un contexte tel que celui de la loi Lemoine.

A ce stade, les questions soulevées ajoutées à l'investissement nécessaire pour une maîtrise de ce type d'analyse par plusieurs personnes au sein d'une compagnie d'assurance pourraient rebuter ces entreprises et les dissuader d'utiliser cette méthode. Cependant ce n'est pas un hasard si la littérature sur le CUSUM est si vaste, et si les domaines d'applications sont si nombreux. Le CUSUM, parce qu'il est le moyen de détecter le plus rapidement des variations dans l'intensité des processus de comptage, est nécessairement

1. Par analogie avec le vocabulaire utilisé par les amateurs de la conjoncture de Syracuse

un outil qui peut apporter une plus-value aux assureurs. Il s'agit seulement de l'utiliser de façon raisonnée, en lien avec les autres analyses que les assureurs utilisent déjà. Par ailleurs, les analyses de type CUSUM existent aussi dans des cadres multivariés, et les recherches autour du CUSUM, et plus généralement sur la détection des changements, invitent à penser que ces sujets prendront à l'avenir de plus en plus de place dans le domaine de l'assurance. De plus, le CUSUM permet de focaliser l'étude sur des sous-populations, et donc, peut-être, d'éviter le crash suite à la décision d'éteindre le moteur qui marche.

Table des figures

1	Évolution du processus CUSUM pour le décès (population générale) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	8
2	Évolution du processus CUSUM pour l'ITT (population générale) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	9
3	Evolution of the CUSUM process for death (general population) since 1st January 2019, threshold for a false alarm every 20 years	17
4	Evolution of the CUSUM process for temporary work incapacity (general population) since 1st January 2019, alert threshold set for one false alarm every twenty years	18
1.1	Marché de l'assurance emprunteur	27
1.2	Poids de la délégation dans le marché de l'assurance emprunteur	29
3.1	Processus séquentiel du logarithme du rapport de vraisemblance (LPSR pour <i>Log Sequential Probability Ratio</i>) dans le cas d'une suite de variables aléatoires gaussiennes de moyenne 1,5 jusqu'à la 500 ^{ème} occurrence, et de moyenne 1,8 ensuite (ici le temps est le nombre d'observations)	38
3.2	Carte de Shewart	40
3.3	Carte de contrôle à moyenne mobile géométrique	41
3.4	Carte de contrôle à moyenne mobile finie	42
3.5	Algorithme à dérivée filtrée	43
3.6	Le processus séquentiel du logarithme des rapports de vraisemblance (LPSR) S_k en rouge et le processus CUSUM $S_k - m_k$ en bleu	44
3.7	A : Processus de Poisson d'intensité 1 jusqu'au temps 15 et d'intensité 1,5 ensuite, A' : processus U (en bleu) et V (en rouge) associés ; B : Processus de Poisson d'intensité 1 jusqu'au temps 15 et d'intensité 0,5 ensuite, B' : processus X (en bleu) et Y (en rouge) associés	47
4.1	Effet du changement de temps sur un processus de Poisson N et sur le processus CUSUM V associé	53
4.2	Parcours possibles pour U_t	54
4.3	Parcours possibles pour Y_t	57
4.4	Comparaisons des espérances obtenues par simulations et par la formule exacte pour quelques valeurs de $\rho > 1$	61
4.5	Comparaisons des espérances obtenues par simulations et par la formule exacte pour quelques valeurs de $\rho < 1$	62
5.1	Évolution du processus CUSUM pour le décès (population générale) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les trois ans	69
5.2	Évolution du processus CUSUM pour le décès (population générale) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	70

5.3	Sinistralité liée au coronavirus en France, Italie et Belgique	70
5.4	Évolution du processus CUSUM pour le décès (assurés en France) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les trois ans	71
5.5	Évolution du processus CUSUM pour le décès (assurés en France) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	72
5.6	Évolution du processus CUSUM pour le décès de la population féminine depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les trois ans	73
5.7	Évolution du processus CUSUM pour le décès de la population féminine depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	73
5.8	Évolution du processus CUSUM pour le décès de la population masculine depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les trois ans	74
5.9	Évolution du processus CUSUM pour le décès de la population masculine depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	74
5.10	Évolution du processus CUSUM pour le décès des non-fumeurs depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les trois ans	75
5.11	Évolution du processus CUSUM pour le décès des fumeurs depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les trois ans	76
5.12	Évolution du processus CUSUM pour le décès des non-fumeurs depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	76
5.13	Évolution du processus CUSUM pour le décès des fumeurs depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	77
5.14	Évolution du processus CUSUM pour le décès des 35-54 ans depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	78
5.15	Évolution du processus CUSUM pour le décès des 55-64 ans depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	79
5.16	Évolution du processus CUSUM pour le décès des 65-74 ans depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	80
5.17	Évolution du processus CUSUM pour le décès des 75-84 ans depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	81
5.18	Évolution du processus CUSUM pour le décès des VLR depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	82
5.19	Évolution du processus CUSUM pour le décès des non VLR depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	82
5.20	Évolution du processus CUSUM pour le décès des non surprimés depuis le 1er janvier 2019	83
5.21	Évolution du processus CUSUM pour le décès des assurés soumis à surprime depuis le 1er janvier 2019	84
6.1	Évolution du processus CUSUM pour l'ITT (population générale) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	90
6.2	Évolution du processus CUSUM pour l'ITT (assurés sans surprime) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	92
6.3	Évolution du processus CUSUM pour l'ITT (assurés avec surprime) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	92

6.4	Évolution du processus CUSUM pour l'ITT (assurés non-fumeurs) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	93
6.5	Évolution du processus CUSUM pour l'ITT (assurés fumeurs) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans . . .	94
7.1	Evolution du processus CUSUM pour le décès (assurés en France) depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les trois ans	96
7.2	Evolution du processus CUSUM pour l'ITT pour les assurés classés VLR, depuis le 1er janvier 2019, seuil d'alerte fixé pour une fausse alerte tous les vingt ans	97
B.1	Poids d'une erreur de prédiction pour une observation de valeur 1, selon la valeur prédite	114

Liste des tableaux

5.1	Exposition par année d'observation et sources de données	64
5.2	Nombre de décès par année d'observation et sources de données	64
5.3	Taux de décès bruts par année d'observation et sources de données	65
5.4	Quelques éléments sur la composition du portefeuille	66
5.5	Modèle retenu pour le décès	68
6.1	Exposition par année d'observation	86
6.2	Nombre de première ITT avec prestation par année d'observation	86
6.3	Taux bruts par année d'observation	87
6.4	Statistiques descriptives du portefeuille des assurés en ITT	87
6.5	Modèle retenu pour l'ITT	88
A.1	Validation croisée sur quelques modèles pour le décès	109
A.2	Quelques modèles étudiés	110
A.3	Validation croisée sur quelques modèles pour l'ITT	111
A.4	Results	112

Bibliographie

- [Assemblée Nationale, 2012] Assemblée Nationale (2012). Question publiée au journal : 25/09/2012 réponse publiée au journal : 12/03/2013 page : 2811. <https://questions.assemblee-nationale.fr/q14/14-5098QE.htm>.
- [Basseville and Nikiforov, 1993] Basseville, M. and Nikiforov, I. (1993). *Detection of Abrupt Changes : Theory and Application*. Prent.Hall.
- [CCSF, 2021a] CCSF (2021a). Bilan de l'assurance emprunteur. https://www.ccsfin.fr/sites/default/files/media/2020/11/11/2020_ccsf_rapport_bilan_assurance_emprunteur.pdf.
- [CCSF, 2021b] CCSF (2021b). Glossaire.
- [Code de la Consommation, 1993] Code de la Consommation (1993). Article l312-9, version en vigueur du 27 juillet 1993 au 01 septembre 2010. https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000006292448/1993-07-27/.
- [Code des Assurances, 1981] Code des Assurances (1981). Article l113-1, version en vigueur depuis le 08 janvier 1981. https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000006791984/.
- [Code des Assurances, 1994] Code des Assurances (1994). Article l112-4, version en vigueur depuis le 01 juillet 1994. https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000006791931/.
- [Code des Assurances, 2006] Code des Assurances (2006). Articles l141-1 à l141-7.
- [Conseil d'Etat, 2012] Conseil d'Etat (2012). <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000026229937>.
- [Daley and Vere-Jones, 2003] Daley, D. and Vere-Jones, D. (2003). *An introduction to the theory of point processes. Vol. I : Elementary theory and methods. 2nd ed*, volume Vol. 1. Springer.
- [Daloz Actualités, 2021] Daloz Actualités (2021). L'article l. 113-1 du code des assurances et les clauses d'exclusion non formelles sur la sellette. <https://www.daloz-actualite.fr/flash/l-article-l-113-1-du-code-des-assurances-et-clauses-d-exclusion-non-formelles-sur-sellette>.
- [El Karoui et al., 2017] El Karoui, N., Loisel, S., and Salhi, Y. (2017). Minimax optimality in robust detection of a disorder time in doubly-stochastic Poisson processes. *The Annals of Applied Probability*, 27(4) :2515 – 2538.
- [Gandy et al., 2010] Gandy, A., Kvaløy, J. T., Bottle, A., and Zhou, F. (2010). Risk-adjusted monitoring of time to event. *Biometrika*, 97(2) :375–388.
- [Inspection Générale des Finances, 2013] Inspection Générale des Finances (2013). Assurance-emprunteur. <https://www.igf.finances.gouv.fr/files/live/sites/igf/files/contributed/IGF%20internet/2.RapportsPublics/2013/2013-M-086.pdf>.

- [Laub et al., 2020] Laub, P. J., El Karoui, N., Loisel, S., and Salhi, Y. (2020). Quickest detection in practice in presence of seasonality : an illustration with call center data. In *Data analytics and Models for Insurance*. Economica.
- [Lorden, 1971] Lorden, G. (1971). Procedures for Reacting to a Change in Distribution. *The Annals of Mathematical Statistics*, 42(6) :1897 – 1908.
- [Morel, 2002] Morel, C. (2002). *Les décisions absurdes, I*. Gallimard.
- [Page, 1954] Page, E. S. (1954). Continuous Inspection Schemes. *Biometrika*, 41(1-2) :100–115.
- [Ritchie et al., 2020] Ritchie, H., Mathieu, E., Rodés-Guirao, L., Appel, C., Giattino, C., Ortiz-Ospina, E., Hasell, J., Macdonald, B., Beltekian, D., and Roser, M. (2020). Coronavirus pandemic (covid-19). *Our World in Data*. <https://ourworldindata.org/coronavirus>.
- [Rodriguez, 2016] Rodriguez, D. (2016). Lecture notes on generalized linear models. <https://data.princeton.edu/wws509/notes/>.
- [Sante Publique France, 2021] Sante Publique France (2021). https://www.santepubliquefrance.fr/content/download/358941/document_file/412205_spf00003046.pdf.
- [Wikipedia, 2021a] Wikipedia (2021a). Change detection. https://en.wikipedia.org/wiki/Change_detection.
- [Wikipedia, 2021b] Wikipedia (2021b). Franchise (assurance). [https://fr.wikipedia.org/wiki/Franchise_\(assurance\)](https://fr.wikipedia.org/wiki/Franchise_(assurance)).
- [Wikipedia, 2022] Wikipedia (2022). Mouvement des gilets jaunes. https://fr.wikipedia.org/wiki/Mouvement_des_Gilets_jaunes.

Annexe A

Modèles

Plusieurs modèles ont été élaborés à partir des données disponibles pour décrire les sinistralités liées au décès et aux arrêts de travail. Pour sélectionner sinon le meilleur modèle, du moins un modèle crédible (dont les variables explicatives sont statistiquement significatives) et efficace (parcimonieux et qui fait mieux qu'un modèle sans autre variable explicative que l'exposition), une validation croisée sur 10 sous-ensembles a été menée pour chaque modèle candidat.

Quelques autres types de modèles ont été testés (GAM, Régression négative binomiale, régression tweedie, Gradient Boosting avec le package R **xgboost**). Ces modèles n'ont pas donné d'amélioration importante des prédictions. Cependant, c'est le modèle GAM qui a fait émerger la variable binaire pour un âge avant ou après 60 ans, variable qui s'est avérée significative dans les deux cas.

A.1 Modèles étudiés pour le décès

Le tableau [A.1](#) recense les résultats obtenus par quelques modèles sur les données liées à la garantie décès. Le modèle 7 (offset uniquement) est estimé avec la seule exposition en offset. Les modèles 4 et 9 présentent les déviations moyennes (moyenne des déviations obtenues sur les échantillons de test de la validation croisée) les plus basses. Les résultats des modélisations sur l'ensemble des données, présentés dans le tableau [A.2](#) montrent cependant que ces modèles comportent des variables explicatives peu significatives. Au regard des faibles écarts sur les variances, un modèle qui ne contient que des variables dont l'effet est statistiquement significatif, le modèle 6, est donc préféré.

Les lecteurs particulièrement attentifs pourraient s'étonner de ce que le nombre d'observations dans le tableau [5.5](#) et dans le tableau [A.2](#) sont différents. C'est effectivement le cas, car pendant la phase de recherche des modèles, les regroupements étaient effectués sur plus de variables, et ceci conduisait donc à plus de pseudo-observations. Au demeurant, la comparaison des coefficients entre le modèle 6 du second tableau (modèle qui a été retenu) et ceux du premier de ces deux tableaux vient illustrer très concrètement l'analyse faite en [section 4.1.1](#), qui montre l'indifférence des coefficients obtenus par maximum de vraisemblance selon le modèle poissonnien aux regroupements utilisés. Il faut cependant bien prendre garde que, comme préconisé à la fin de cette section, la comparaison des modèles entre eux, afin de décider lequel retenir, doit se faire sur les mêmes pseudo-observations.

	modèles	déviante (Validation Croisée)
1	age + genre + est_fumeur + avec_surprime + age.60 + VLR + saison	1 345,28
2	age + genre + est_fumeur + anciennete + avec_surprime + genre * age + avec_surprime * age + avec_surprime * anciennete + age.60 + age.60 * est_fumeur + VLR + saison	1 344,17
3	age + genre + est_fumeur + anciennete + avec_surprime + genre * age + avec_surprime * age + avec_surprime * anciennete + age.60 + VLR + saison	1 343,80
4	age + genre + est_fumeur + anciennete + avec_surprime + genre * age + avec_surprime * anciennete + age.60 + VLR + saison	1 343,17
5	age + genre + est_fumeur + anciennete + avec_surprime + avec_surprime * anciennete + age.60 + VLR + saison	1 344,32
6	age + genre + est_fumeur + anciennete + avec_surprime + age.60 + VLR + saison	1 343,76
7	offset uniquement	1 935,37
8	age + genre + est_fumeur + avec_surprime + VLR + saison	1 345,33
9	age + genre + est_fumeur + avec_surprime + VLR + systeme_libelle + saison	1 343,42

TABLE A.1 – Validation croisée sur quelques modèles pour le décès

Avec une application stricte de la préférence pour le modèle qui aboutirait à la plus petite déviante sur l'échantillon de test, le modèle retenu serait celui présenté dans la table [5.5](#)

	Dependent variable : n.suiviste								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
age	0.105*** (0.003)	0.099*** (0.004)	0.099*** (0.004)	0.098*** (0.004)	0.103*** (0.003)	0.103*** (0.003)	0.111*** (0.002)	0.111*** (0.002)	0.110*** (0.002)
genreH	0.616*** (0.042)	0.150 (0.209)	0.148 (0.209)	0.169 (0.208)	0.618*** (0.042)	0.619*** (0.042)	0.615*** (0.042)	0.615*** (0.042)	0.622*** (0.042)
est_fumeur	0.787*** (0.051)	0.786*** (0.074)	0.764*** (0.051)	0.766*** (0.051)	0.765*** (0.051)	0.765*** (0.051)	0.785*** (0.051)	0.785*** (0.051)	0.770*** (0.052)
anciennete		0.015*** (0.006)	0.015*** (0.006)	0.015*** (0.006)	0.015*** (0.006)	0.019*** (0.005)		0.208*** (0.044)	0.198*** (0.045)
avec_surprime	0.212*** (0.044)	0.353 (0.229)	0.346 (0.228)	0.110 (0.075)	0.106 (0.075)	0.193*** (0.044)			
age.60avant60	-0.178** (0.072)	-0.187** (0.076)	-0.197*** (0.072)	-0.197*** (0.073)	-0.201*** (0.073)	-0.203*** (0.073)			
VLR	-0.472*** (0.055)	-0.443*** (0.055)	-0.443*** (0.055)	-0.444*** (0.055)	-0.444*** (0.055)	-0.446*** (0.055)		-0.475*** (0.055)	-0.486*** (0.056)
systeme_libelleSysteme 2									
systeme_libelleSysteme 3									
systeme_libelleSysteme 4									
systeme_libelleSysteme 5									
saisonete	-0.087 (0.056)	-0.086 (0.056)	-0.086 (0.056)	-0.086 (0.056)	-0.086 (0.056)	-0.086 (0.056)		-0.087 (0.056)	-0.088 (0.056)
saisonhiver	0.169*** (0.053)	0.171*** (0.053)	0.170*** (0.053)	0.170*** (0.053)	0.170*** (0.053)	0.170*** (0.053)		0.169*** (0.053)	0.168*** (0.053)
saisonprintemps	0.108** (0.054)	0.111** (0.054)	0.111** (0.054)	0.111** (0.054)	0.111** (0.054)	0.110** (0.054)		0.109** (0.054)	0.107** (0.054)
age :genreH		0.008** (0.003)	0.008** (0.003)	0.007** (0.003)					
age :avec_surprime		-0.004 (0.004)	-0.004 (0.004)						
anciennete :avec_surprime		0.017 (0.010)	0.017 (0.010)	0.015 (0.010)	0.015 (0.010)				
est_fumeur :age.60avant60		-0.043 (0.102)							
Constant	-12.286*** (0.209)	-12.015*** (0.252)	-12.008*** (0.251)	-11.961*** (0.247)	-12.255*** (0.209)	-12.268*** (0.209)	-6.370*** (0.019)	-12.723*** (0.113)	-13.367*** (0.422)
Observations	106.672	106.672	106.672	106.672	106.672	106.672	106.672	106.672	106.672
Log Likelihood	-9.300.378	-9.288.091	-9.288.181	-9.288.774	-9.291.173	-9.292.265	-12.261.280	-9.303.417	-9.293.698
Akaike Inf. Crit.	18.620.760	18.606.180	18.604.360	18.603.550	18.606.350	18.606.530	24.524.570	18.624.830	18.613.400

Note : ** p<0.01 ; *** p<0.001

TABLE A.2 – Quelques modèles étudiés

A.2 Modèles étudiés pour l'incapacité temporaire de travail

La démarche a été ici strictement similaire à celle pour le décès.

	modèles	déviante (Validation Croisée)
1	age + genre + est_fumeur + avec_surprime + CARENCE_180 + CARENCE_30 + PLUS + age.60 + VLR + mois + anciennete2	2 412,6
2	age + genre + est_fumeur + avec_surprime + CARENCE_180 + CARENCE_30 + PLUS + age.60 + VLR + mois + anciennete	2 410,23
3	age + genre + est_fumeur + avec_surprime + CARENCE_180 + CARENCE_30 + PLUS + age.60 + VLR + saison + anciennete2	2 410,37
4	age + genre + est_fumeur + avec_surprime + CARENCE_180 + CARENCE_30 + PLUS + age.60 + VLR + saison + anciennete	2 407,59
5	age + genre + est_fumeur + avec_surprime + CARENCE_180 + CARENCE_30 + PLUS + age.60 + VLR + saison + anciennete + anciennete2	2 406,81
6	age + genre + est_fumeur + anciennete + avec_surprime + age.60 + VLR	2 421,78
7	offset(log(exposition))	2 519,41
8	age + genre + est_fumeur + avec_surprime + CARENCE_180 + CARENCE_30 + PLUS + age.60 + VLR + saison + anciennete + anciennete2 + age * genre	2 403,5
9	age + genre + est_fumeur + avec_surprime + CARENCE_180 + CARENCE_30 + PLUS + age.60 + VLR + saison + anciennete + age * genre	2 403,47

TABLE A.3 – Validation croisée sur quelques modèles pour l'ITT

	Dependent variable :								
	h.sismistre								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
age	0.038*** (0.002)	0.040*** (0.002)	0.037*** (0.002)	0.039*** (0.002)	0.040*** (0.002)	0.038*** (0.002)	0.039*** (0.003)	0.029*** (0.003)	0.029*** (0.003)
genreH	-0.312*** (0.033)	-0.312*** (0.033)	-0.312*** (0.033)	-0.312*** (0.033)	-0.312*** (0.033)	-0.305*** (0.033)	-1.306*** (0.167)	-1.306*** (0.167)	-1.309*** (0.167)
est_fumeur	0.242*** (0.041)	0.247*** (0.041)	0.243*** (0.041)	0.248*** (0.041)	0.249*** (0.041)	0.244*** (0.041)	0.251*** (0.041)	0.251*** (0.041)	0.253*** (0.041)
avec_surprime	0.527*** (0.036)	0.533*** (0.036)	0.527*** (0.036)	0.533*** (0.036)	0.533*** (0.036)	0.536*** (0.036)	0.532*** (0.036)	0.532*** (0.036)	0.532*** (0.036)
CARENCE_180	-0.722*** (0.224)	-0.754*** (0.224)	-0.723*** (0.224)	-0.756*** (0.224)	-0.754*** (0.224)	-0.754*** (0.224)	-0.754*** (0.224)	-0.754*** (0.224)	-0.755*** (0.224)
CARENCE_30	0.994*** (0.101)	0.987*** (0.101)	0.993*** (0.101)	0.986*** (0.101)	0.984*** (0.101)	0.984*** (0.101)	0.984*** (0.101)	0.984*** (0.101)	0.986*** (0.101)
PLUS	0.214*** (0.039)	0.198*** (0.039)	0.212*** (0.039)	0.195*** (0.039)	0.188*** (0.039)	0.198*** (0.039)	0.189*** (0.039)	0.189*** (0.039)	0.196*** (0.039)
age.60avant60	0.961*** (0.118)	0.917*** (0.118)	0.960*** (0.118)	0.918*** (0.118)	0.924*** (0.118)	0.905*** (0.118)	0.944*** (0.119)	0.944*** (0.119)	0.937*** (0.119)
VLR	-0.690*** (0.040)	-0.706*** (0.040)	-0.691*** (0.040)	-0.707*** (0.040)	-0.708*** (0.040)	-0.655*** (0.040)	-0.710*** (0.040)	-0.710*** (0.040)	-0.710*** (0.040)
mois	-0.010** (0.005)	-0.010** (0.005)							
saisonete			-0.247*** (0.046)	-0.248*** (0.046)	-0.248*** (0.046)			-0.248*** (0.046)	-0.248*** (0.046)
saisonhiver			-0.190*** (0.044)	-0.191*** (0.044)	-0.191*** (0.044)			-0.191*** (0.044)	-0.191*** (0.044)
saisonprintemps			-0.124*** (0.044)	-0.126*** (0.044)	-0.126*** (0.044)			-0.126*** (0.044)	-0.126*** (0.044)
anciennete2old	-0.330*** (0.034)		-0.330*** (0.034)		-0.121** (0.045)			-0.119** (0.049)	
age :genreH								0.023*** (0.004)	0.023*** (0.004)
anciennete		-0.054*** (0.005)		-0.054*** (0.005)		-0.060*** (0.005)		-0.042** (0.007)	-0.054*** (0.005)
Constant	-7.653*** (0.170)	-7.661*** (0.170)	-7.577*** (0.170)	-7.583*** (0.170)	-7.583*** (0.170)	-7.583*** (0.166)	-5.470*** (0.016)	-7.152*** (0.181)	-7.161*** (0.181)
Observations	319,403	319,403	319,403	319,403	319,403	319,403	319,403	319,403	319,403
Log Likelihood	-15,883.250	-15,868.850	-15,869.050	-15,854.450	-15,851.340	-15,929.910	-16,428.180	-15,832.760	-15,835.780
Akaike Inf. Cht.	31,790.450	31,761.690	31,766.100	31,736.900	31,732.680	31,876.820	32,858.370	31,697.520	31,701.560

* p<0.1 ; ** p<0.05 ; *** p<0.01

TABLE A.4 – Results

Annexe B

Choix d'une métrique pour la régression

Le choix de la métrique est fondamental, puisqu'il préside à la qualité des résultats.

RMSE

C'est la racine carrée de la moyenne des erreurs de prédiction. Elle pénalise donc pareillement l'erreur qui sous-estime la valeur observée et celle qui la sur-estime.

Déviante de Poisson

La déviance est définie comme le double de l'écart entre la log-vraisemblance du modèle (i.e la log-vraisemblance maximisée) et la log-vraisemblance du modèle saturé (modèle qui posséderait autant de régresseurs que d'observation et qui donc prédirait exactement les observations, ici aussi il s'agit de la log-vraisemblance maximisée). A partir de l'équation 4.3, avec $\beta = (b_1, ..b_n)$, la log-vraisemblance du modèle saturé s'écrira :

$$\mathcal{L}_S(Y|X) = \sum_{i=1}^{i=n} y_i \ln(\lambda_0) + y_i (b_i + \ln(e_i)) - \log(y_i!) - \lambda_0 e^{b_i + \ln(e_i)}. \quad (\text{B.1})$$

Avec $\mu_i = \ln(b_i)$, le maximum de vraisemblance du modèle saturé est obtenu pour

$$\mu_i = \frac{y_i}{\lambda_0 e_i}.$$

D'autre part, pour le modèle a tester, la prédiction est donnée par

$$\hat{y}_i = \lambda_0 e^{\beta' x_i + \ln(e_i)}.$$

Et reportés dans B.1 et 4.3, ces éléments conduisent à la déviance :

$$D = 2 * \sum_i \left[y_i * \log\left(\frac{y_i}{\hat{y}_i}\right) - (y_i - \hat{y}_i) \right].$$

La figure B.1 illustre que la déviance pénalise plus fortement la sous-estimation que la surestimation, alors que la RMSE les pénalise de façon identique. Choisir la déviance plutôt que la RMSE comme critère pour départager les modèles conduit donc à plutôt privilégier les modèles qui auraient tendance à surestimer le risque à ceux qui auraient tendance à le sous-estimer.

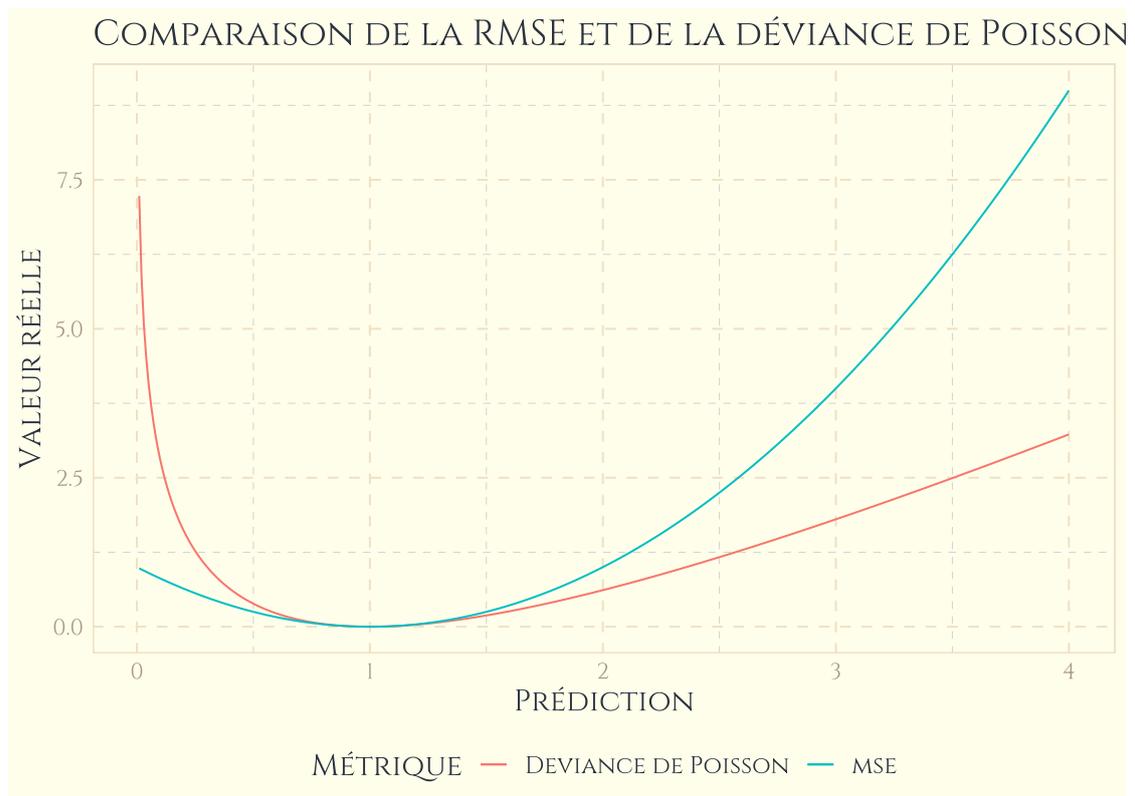


FIGURE B.1 – Poids d'une erreur de prédiction pour une observation de valeur 1, selon la valeur prédite

Annexe C

Eléments de code

L'extraction des données a nécessité de longues requêtes SQL alliées à une bonne connaissance des bases de données sources, et beaucoup de codes pour contrôler les données. La connaissance du SQL doit aujourd'hui faire partie du bagage des actuaires, et les particularités des structures de données sources seraient de peu d'intérêt ici. Il n'en reste pas moins que cette étape est fondamentale, et très chronophage.

Le calcul des expositions, après regroupement des données, presque jour par jour, peut aussi poser quelques problèmes, selon la quantité de données impliquées. Pour des jeux de données de la taille de ceux impliqués dans ce mémoire, une bonne connaissance du package **data.table**, et en particulier de la fonction **foverlaps** permet de s'en sortir avec des temps de calculs tout à fait raisonnables (sous réserve d'avoir une bonne machine).

Parmi les autres codes utilisés lors de l'élaboration de ce mémoire, ceux qui ont permis l'estimation des intensités de référence par régression de Poisson sont classiques et peuvent facilement être retrouvés sur internet. Le package **boot** a été utilisé pour la validation croisée, avec une fonction de coût particulière.

C.1 Fonction de coût pour la validation croisée

```
cout.deviance <- function(y, yhat) {  
  devyhat <- y * log(yhat) - yhat  
  devy <- y * log(y) - y  
  devy[y == 0] <- 0  
  return(2 * sum(devy - devyhat))  
}
```

C.2 Fonctions génériques

Le code ci-dessous présente la fonction `get.esperance.threshold`¹, qui effectue les calculs de niveau d'alerte en fonction de ρ et du nombre moyen de sinistres avant une fausse alerte (selon la méthode présentée en sous-section 4.3.2), et la fonction `get.cusum.graphic` qui, pour un processus donné avec son intensité théorique avant changement et des nombres de sinistres observés produit les graphiques à la hausse et à la baisse. Tous les détails sont dans les commentaires.

```
# Fonctions pour l'utilisation des Processus CUSUM
#
#
# J.LELOUARNE@afi-esca.com

# Appel (et éventuellement installation) des packages utilisés
for (package in c('data.table', 'lubridate', 'stringr', 'ggplot2',
  ↪ 'cowplot',
                    'Matrix', 'scales', 'latex2exp', 'expm', 'ggpubr'))
  ↪ {
  if (!require(package, character.only = TRUE, quietly = TRUE)) {
    install.packages(package, dependencies = TRUE)
    library(package, character.only = TRUE)
  }
}

##### -
# La fonction pour les seuils # ----
##### -

get.esperance.threshold <- function(pho, c, S0, lambda = 1 ){

  if (pho > 1){
    # Cas pho > 1 ----
    ##### -

    # a = temps minimal pour que R descende de 1 :
    a <- log(pho)/(pho -1)

    # k = nombre de seuils, espacés de 1, entre c (niveau limite)
    k <- ceiling(c-1)

    # xi = temps nécessaire pour descendre du niveau c-k à 0 :
    xi <- a*(c-k)
  }
}
```

1. Qui porte mal son nom, non seulement parce qu'il mélange le français et l'anglais, mais aussi parce qu'en réalité la fonction ne donne pas un seuil, mais renvoie un nombre moyen de sinistres et prend un seuil en paramètre. Il reste qu'elle peut servir effectivement à calculer les seuils pour un nombre de sinistre moyen donné.

```

# eta = temps restant après xi pour atteindre le temps a :
eta <- a - xi

# Nombre minimal de sauts pour atteindre le seuil c
nu <- ceiling(c-S0)

matrice.transition.hausse <- function(xi, eta){
  # Probabilités d'observer j évènements
  # pendant la durée xi
  # Pour j allant de 0 à k
  p.xi <- ((lambda*xi)^(0:k))*exp(-lambda*xi)/(factorial(0:k))

  # Probabilités d'observer j évènements
  # pendant la durée eta
  # Pour j allant de 0 à k
  p.eta <- ((lambda*eta)^(0:k))*exp(-lambda*eta)/(factorial(0:k))

  # Probabilités d'observer j évènements
  # pendant la durée eta + xi
  # Pour j allant de 0 à k
  p.eta.xi <- ((lambda*(eta + xi))^(0:k))*exp(-lambda*(eta +
  → xi))/(factorial(0:k))

  # Q = matrice de transition sur les états transitoires

  # Première ligne
  ligne.1 <- p.eta.xi[2:(k+1)] +
  → p.xi[1]*(p.eta[1:k]-p.eta[2:(k+1)])

  # k-1 lignes restantes
  Q <- matrix(pmax(rep(1:(k), k-1) - (rep(1:(k-1), each =
  → k)-1),0), nrow = k-1, byrow = TRUE)
  Q[which(Q>0)] <- p.eta.xi[Q[which(Q>0)]]

  # matrice complète
  Q <- rbind(ligne.1, Q)

  return(Q)
}

# q= Distribution de Y1
if (S0 > c-k){
  # t0 = temps nécessaire pour aller de S0 au niveau qui
  → caractérise l'état 1
  t0 <- a*(S0-(c-nu))

```

```

q <- c(rep(0, k-nu),
  → ((lambda*t0)^(0:(nu-1)))*exp(-lambda*t0)/(factorial(0:(nu-1))))

} else {
  q <- c(1, rep(0, k-1)) %*% matrice.transition.hausse(a*S0, eta)

}

Q <- matrice.transition.hausse(xi, eta)

esperance <- nu + q %*% solve(diag(nrow = k) - Q) %*% rep(1, k)
} else

{
  # Cas pho < 1 ----
  ##### -

  # a = temps minimal pour que R monte de 1 :
  a <- log(pho)/(pho -1)

  # k = nombre de seuils, espacés de 1, entre c (niveau limite)
  k <- ceiling(c-1)

  # xi = temps nécessaire pour monter du niveau 0 à c-k :
  xi <- a*(c-k)

  # eta = temps restant avant xi pour atteindre le temps a :
  eta <- a - xi

  # Nombre minimal de sauts pour atteindre le seuil c
  nu <- ceiling(c-S0)

matrice.transition.baisse <- fonction(p_eta, p_xi){
  # Probabilités d'observer j évènements
  # pendant la durée xi
  # Pour j allant de 0 à k
  p.xi <-
  → ((lambda*p_xi)^(0:k))*exp(-lambda*p_xi)/(factorial(0:k))

  # Probabilités d'observer j évènements
  # pendant la durée eta
  # Pour j allant de 0 à k

```

```

p.eta <-
  ↪ ((lambda*p_eta)^(0:k))*exp(-lambda*p_eta)/(factorial(0:k))

# Q = matrice de transition sur les états transitoires
# Q = Q1*Q2

# Q1
Q1 <- t(matrix(pmax(rep(1:(k), k) - (rep(1:(k), each = k)-1),0),
  ↪ nrow = k, byrow = TRUE))
Q1[which(Q1>0)] <- p.eta[Q1[which(Q1>0)]]
Q1 <- rbind(rep(0, k), Q1)
Q1 <- cbind(1-rowSums(Q1), Q1)

# Q2
Q2 <- t(matrix(pmax(rep(1:(k+1), k) - (rep(1:(k), each =
  ↪ k+1)-1),0), nrow = k, byrow = TRUE))
Q2[which(Q2>0)] <- p.xi[Q2[which(Q2>0)]]
Q2 <- cbind(1-rowSums(Q2), Q2)
Q2[k+1,1] <- Q2[k+1,1]-p.xi[1]

# matrice complète
Q <- Q1 %*% Q2
return(Q)
}

# q= Distribution de Y1
# t0 = temps nécessaire pour aller de S0 au niveau le plus
  ↪ proche
# s'il n'y a pas d'évènement
t0 <- a*(S0-(c-nu))

q <- c(rep(0, k-nu+1), 1, rep(0, nu-1)) %*%
  ↪ matrice.transition.baisse(p_eta = max(a-t0-xi,0), p_xi =
  ↪ min(a-t0, xi))

# Matrice de passage
Q <- matrice.transition.baisse(p_eta = eta, p_xi = xi)

# Nombre moyen d'évènements pour franchir le seuil :
esperance <- q %*% (Q %^(nu-1)) %*% solve(diag(nrow = k+1) - Q)
  ↪ %*% rep(1, k+1)

}

return(as.numeric(esperance))

```

```

}

##### -
# La fonction pour graphiques # ----
##### -

get.cusum.graphics <- function(pho = 1.1, pho.baisse = 0.9, S0 = 0,
  ↪ Nmax = 1000,
                                noms.graphiques = "Pas_de_nom",
                                ↪ tbl.process){

# Paramètres ----
##### -
# pho.baisse : Niveau changement à la baisse que l'on souhaite
  ↪ détecter au plus vite s'il survient
#           exemple : pho.baisse = 0.9 pour une baisse d'intensité
  ↪ de 10%

# pho : Niveau changement à la hausse que l'on souhaite détecter au
  ↪ plus vite s'il survient
#           exemple : pho = 1.1 pour une hausse d'intensité de
  ↪ 10%

# Nmax : nombre moyen d'événements avant franchissement du seuil
  ↪ d'alerte
#           s'il n'y a pas de changement (contrôle des fausses alertes)

# S0 : niveau du processus au départ (en général 0)

# noms.graphiques : nom du graphique produit
# Il sera enregistré dans le répertoire de travail de la session R

# tbl.process : data.table qui contient les expositions et sinistres
  ↪ :
#

# Date de calcul = dernière date présente ----
##### -

date.calcul <- tbl.process[, max(fin.période)]

# Seuils ----
##### -

# Calcul du seuil d'alerte à la hausse :
seuil <- optimize(interval = c(1,100),

```

```

        f = function(c) abs(get.esperance.threshold(pho =
        ↪ pho, c = c, S0=0)-Nmax),
        tol = 10^-15
    )[[1]]

# Calcul du seuil d'alerte à la baisse :
seuil.baisse <- optimize(interval = c(0.2,100),
        f = function(c)
        ↪ abs(get.esperance.threshold(pho =
        ↪ pho.baisse, c = c, S0=0)-Nmax),
        tol = 10^-15
    )[[1]]

# Processus -----
##### -
beta.cusum <-function(pho) (pho-1)/log(pho)

# vérification
tbl.process[, sum(lambda, na.rm = TRUE)] -
    ↪ tbl.process[,sum(n.sinistre)]

# Nt
tbl.process[, Nt := cumsum(n.sinistre)]

# Ut
b <- beta.cusum(pho)
b.baisse <- beta.cusum(pho.baisse)
tbl.process[, grand.lambda := cumsum(lambda)]
tbl.process[, Ut := Nt-b*grand.lambda]
tbl.process[, Ut.baisse := - Nt + b.baisse*grand.lambda ]

# ou bien (= juste un changement d'échelle) :
# tbl.process[, Ut := log(pho)*Nt+(1-pho)*grand.lambda]

tbl.process[, Ut.moins := shift(Nt, n=1, fill = 0)-b*grand.lambda] #
    ↪ Juste avant le saut ..
tbl.process[, Ut.baisse.moins := - shift(Nt, n = 1, fill = 0) +
    ↪ b.baisse*grand.lambda]

# Vt
tbl.process[, Vt := 0]
tbl.process[, min.prev.Ut := cummin(Ut.moins)] # Ici le minimum est
    ↪ forcément avant le saut ..
tbl.process[, Vt := Ut - min.prev.Ut]

```

```

tbl.process[, Vt.moins := pmax(Ut.moins - min.prev.Ut,0)] # Juste
  ↪ avant le saut ..
tbl.process[, pente := (Ut.moins - shift(Ut, n=1))/((debut.periode
  ↪ %--% fin.periode /days(1)) / 365.25) ]

# Yt
tbl.process[, Yt := 0]
tbl.process[, min.prev.Ut.baisse := cummin(Ut.baisse)] # Ici le min
  ↪ est forcément après le saut ..
tbl.process[, Yt := Ut.baisse - min.prev.Ut.baisse ]
tbl.process[, Yt.moins := pmax(Ut.baisse.moins -
  ↪ shift(min.prev.Ut.baisse, n= 1, fill = 0) ,0)] # Juste avant le
  ↪ saut ..

# Graphiques ----
##### -

# Table pour le graphique de contrôle des hausses
tbl.graph <- rbindlist(l = list(tbl.process[, .(fin.periode,
  ↪ debut.periode, Ut, Vt)],
                        tbl.process[, .(fin.periode,
  ↪ debut.periode, Ut = Ut.moins, Vt =
  ↪ Vt.moins)])
)))[order(fin.periode, Vt)]

# # Gestion des cas où V devrait arriver à 0 entre deux évènements
# # i.e potentiellement chaque fois que U atteint un minimum ..
# tbl.process[order(debut.periode), temps := cumsum(delta.t)]
# tbl.graph <- rbindlist(l = list(tbl.graph,
#                                tbl.process[order(temps)][,
#                                .(temps =
  ↪ shift(temps, n = 1, fill = 0) - shift(Vt, n= 1, fill =
  ↪ 0)/(1-pho),
#                                Vt =
  ↪ 0,
#                                Ut =
  ↪ shift(Ut, n = 1, fill = 0)- shift(Vt, n= 1, fill = 0),
#                                Ut.moins, min.prev.Ut
#                                )][Ut.moins == min.prev.Ut , .(temps, Ut, Vt)]
# ), use.names = TRUE)[order(temps, V)]

# Table pour le graphique de contrôle des baisses

```

```

tbl.graph.baisse <- rbindlist(1 = list(tbl.process[, .(fin.période,
  ↳ debut.période, Ut.baisse, Yt)],
                                tbl.process[, .(fin.période,
  ↳ debut.période, Ut.baisse =
  ↳ Ut.baisse.moins, Yt =
  ↳ Yt.moins)])
)))[order(fin.période, -Yt)]

tbl.graph.baisse[, alerte := fcase(Yt > seuil.baisse, "red", default =
  ↳ "darkblue")]
tbl.graph[, alerte := fcase(Vt > seuil, "red", default = "darkblue")]

graph_title.hausse <- paste0("Hausse, seuil d'alerte = ",
  ↳ round(seuil,2))
graph_trajectoires <-
  ggplot(data = tbl.graph[fin.période <= date.calcul]) +
  geom_line(aes(x = fin.période, y = Vt, color = Vt > seuil, group =
  ↳ 1)) +
  scale_colour_manual(values = setNames(c('red', 'blue'), c(T, F)),
  ↳ labels = NULL) +
  geom_hline(yintercept = seuil, col = 'red', size = 0.5) +
  geom_hline(yintercept = 0, col = 'black') +
  labs(title = graph_title.hausse, x="t", y = "Vt",
  ↳ subtitle = latex2exp::TeX(paste0("$\\rho = ", phi))
  ) +
  scale_x_date(date_labels = "%m-%Y", date_breaks = "1 year",
  ↳ limits = c(tbl.process[, min(debut.période)],
  ↳ tbl.process[, max(fin.période)]),
  ↳ date_minor_breaks = "1 month") +
  theme(legend.position="none",
  ↳ panel.grid.minor = element_line(color = "lightgrey",
  ↳ size = 0.1,
  ↳ linetype = "dashed"),
  ↳ panel.grid.major = element_line(color = "grey22",
  ↳ size = 0.1,
  ↳ linetype = 2))

graph_title.baisse <- paste0("Baisse, seuil d'alerte = ",
  ↳ round(seuil.baisse,2))

graph_trajectoires.baisse <-
  ggplot(data = tbl.graph.baisse[fin.période <= date.calcul,]) +

  geom_line(aes(x = fin.période, y = Yt, color = Yt > seuil.baisse,
  ↳ group = 1)) +
  scale_colour_manual(values = setNames(c('red', 'blue'), c(T, F)),
  ↳ labels = NULL) +
  geom_hline(yintercept = seuil.baisse, col = 'red', size = 0.5) +

```

```

geom_hline(yintercept = 0, col = 'black') +
labs(title = graph_title.baisse,
      x="t", y = "Yt",
      subtitle = latex2exp::TeX(paste0("$\\rho =", pho.baisse))
) +
scale_x_date(date_labels = "%m-%Y", date_breaks = "1 year",
             limits = c(tbl.process[, min(debut.periode)],
                       ↪ tbl.process[, max(fin.periode)]),
             date_minor_breaks = "1 month") +
theme(legend.position="none",
      panel.grid.minor = element_line(color = "lightgrey",
                                       size = 0.1,
                                       linetype = "dashed"),
      panel.grid.major = element_line(color = "grey22",
                                       size = 0.1,
                                       linetype = 2))

graph_trajectoires.cusum <- plot_grid(plotlist =
                                     list(graph_trajectoires,
                                          graph_trajectoires.baisse),
                                     ncol = 1)

# Sauvegarde des graphiques au format .SVG

# Les deux graphiques (hausse et baisse) avec des axes alignés
ggsave(paste0(noms.graphiques, "_", as.character(Sys.Date()),
             ↪ '%Y%m%d'), '.svg'),
       device = "svg", plot = graph_trajectoires.cusum,
       width = 210, height = 150, units = c("mm"))

# Hausse
ggsave(paste0(noms.graphiques, "_hausse_", as.character(Sys.Date()),
             ↪ '%Y%m%d'), '.svg'),
       device = "svg", plot = graph_trajectoires,
       width = 210, height = 150, units = c("mm"))

# Baisse
ggsave(paste0(noms.graphiques, "_baisse_", as.character(Sys.Date()),
             ↪ '%Y%m%d'), '.svg'),
       device = "svg", plot = graph_trajectoires.baisse,
       width = 210, height = 150, units = c("mm"))

# Renvois des objets ggplot en sortie de fonction
return(list(graph_trajectoires, graph_trajectoires.baisse,
           ↪ graph_trajectoires.cusum))
}

```

C.3 Exemple d'utilisation

A titre d'illustration, le code ci-dessous peut-être exécuté après avoir lancé dans une session R le code de la partie précédente.

```
tbl.poisson.process <- data.table(debut.période =
  ↪ as.Date('01/01/2020', '%d/%m/%Y') + 0:749,
                                fin.période = as.Date('01/01/2020',
  ↪ '%d/%m/%Y') + 1:750,
                                n.sinistre = c(rpois(375, 0.5),
  ↪ rpois(375, 0.5 * 1.1)),
                                lambda = rep(0.5, 750)
  )

date.calcul <- tbl.poisson.process[, max(fin.période)]
get.cusum.graphics(pho = 1.5, pho.baisse = 0.9, S0 = 0, Nmax = 500,
  noms.graphiques = "test_cusum_donnees_simulees",
  tbl.poisson.process)
```