

Modèle de provisionnement des sinistres sur données détaillées en assurance de responsabilité civile des risques industriels dans l'environnement de Solvabilité II

Jean-Michel VAUTRIN

10/07/2025

SOMMAIRE

- 1 • Contexte et Enjeux
- 2 • Modèles sur données agrégés utilisés
- 3 • Architecture du modèle sur données détaillées
- 4 • Résultats
- 5 • Perspectives et Développements
- 6 • Conclusion
- Q&A - Remerciements

CONTEXTE ET ENJEUX BUSINESS

- Principe fondamental: « L'assureur doit faire face à ses engagements à tout instant »
- Entreprise : HDI Global SE, succursale pour la France
- Direction : Direction Technique
- Portefeuille: Contrats RCG de clients industriels, dit « grands risques », activités hétérogènes, branche à développement long

- Profil des sinistres clos (période 1992 – 2009):

Montants en €	Nombre de sinistres clos	% Charge Ultime Totale
Sans paiements	12,791	0%
0 - 10k	12,372	7%
10k - 100k	2,380	22%
100k - 500k	374	24%
500k - 1m	50	11%
1m - 5m	25	15%
5m - 10m	3	6%
Plus de 10m	1	14%
Total	27,996	100.0%

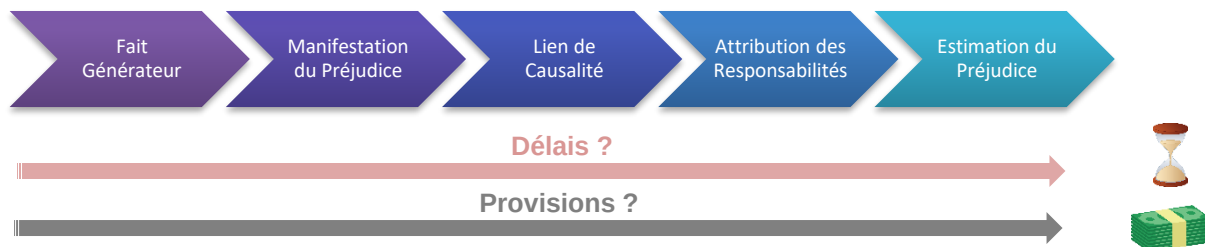
1.6 % des dossiers clos
 =
 70% de la charge totale

- Challenge: Comment estimer le provisionnement des sinistres sur ce portefeuille?

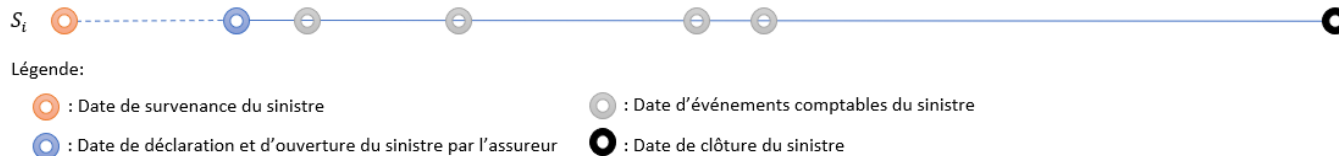
CONTEXTE CONTRACTUEL

- Couverture RCG: protège l'assuré contre un éventail de réclamations potentielles de la part de tiers découlant de préjudices corporels, matériels, ou immatériels.

- Étapes clés d'un sinistre RCG:

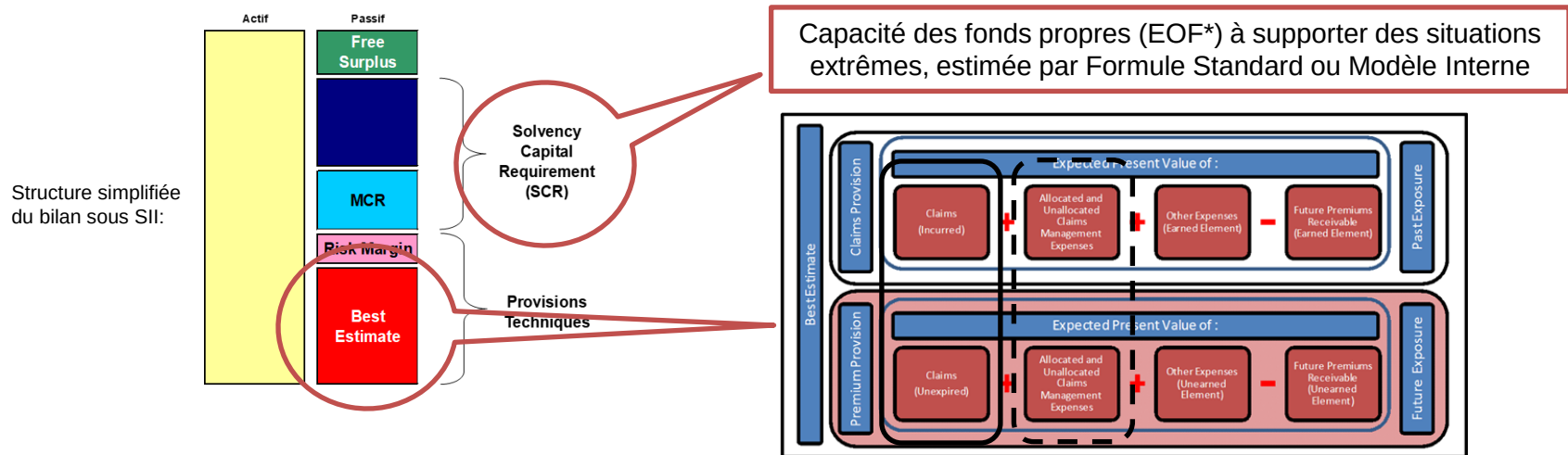


- ...et dans les comptes de l'assureur:



CONTEXTE RÉGLEMENTAIRE

- Solvabilité II : Standardisation des exigences quantitatives, qualitatives et de reporting des risques des assureurs de l'UE.



L'application des formules de SII sont simplifiées au seul risque de provisionnement de cette branche d'activité.

- French GAAP : Taxe sur boni de liquidation ? Justification de certaines positions auprès de l'administration fiscale?

Modèles basés sur les triangles de liquidations des sinistres utilisés

- Forme générale:

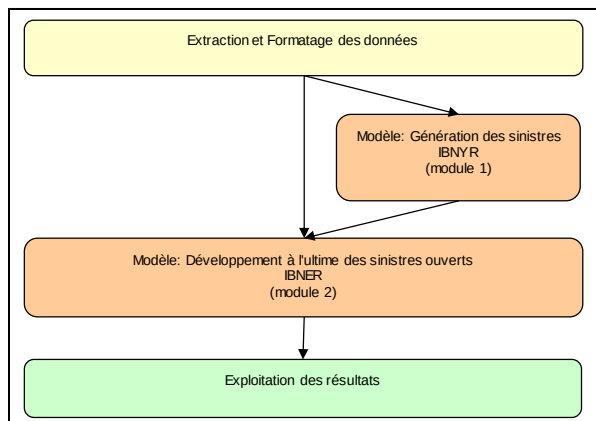
$$C = \begin{pmatrix} C_{1,1} & C_{1,2} & \cdots & C_{1,n-1} & C_{1,n} \\ C_{2,1} & C_{2,2} & \cdots & C_{2,n-1} & \\ \vdots & \vdots & \ddots & & \\ C_{n-1,1} & C_{n-1,2} & & & \\ C_{n,1} & & & & \end{pmatrix}$$

- $C_{i,k}$ sont appliquées aux Paiements et aux Charge Totales, sans distinction Indemnitaire / Frais.
- Approche déterministe classique: Chain-Ladder
- 2 modèles stochastiques sur triangles agrégés seront utilisés:
- Méthode de Mack: La distribution des provisions $\hat{R}$ est estimée par ajustement d'une loi Log-Normale (méthode des moments).
- Par GLM à Bootstrap: basé sur les triangles, paiements et charges totales, des incréments de facteurs de développement des valeurs cumulées:

$$\hat{f}_k = \frac{\sum_{i=1}^{n-k} C_{i,k} \hat{f}_{i,k}}{\sum_{i=1}^{n-k} C_{i,k}}$$

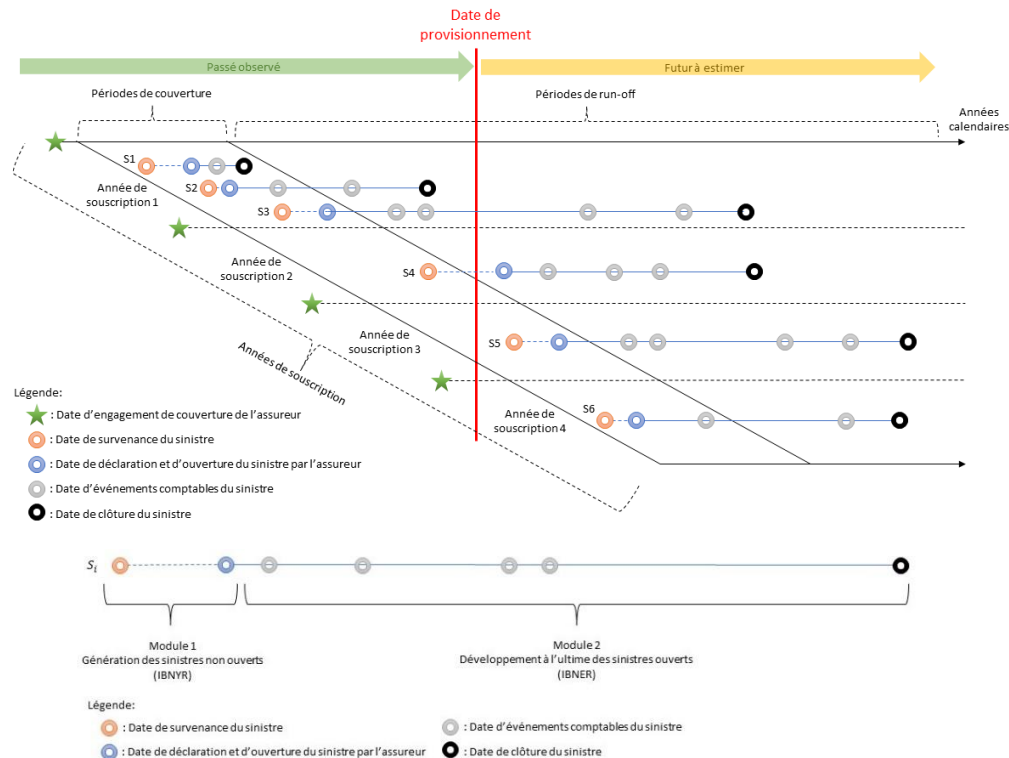
- Variable réponse: Loi de Poisson – Fonction lien: logarithme – Variable de pondération: $C_{i,k}$
- Simulation de 10,000 itérations des résidus de Pearson pour définir la distribution de $\hat{R}$.

Structures de développement d'un sinistre individuel et Conséquence sur l'architecture du modèle détaillé

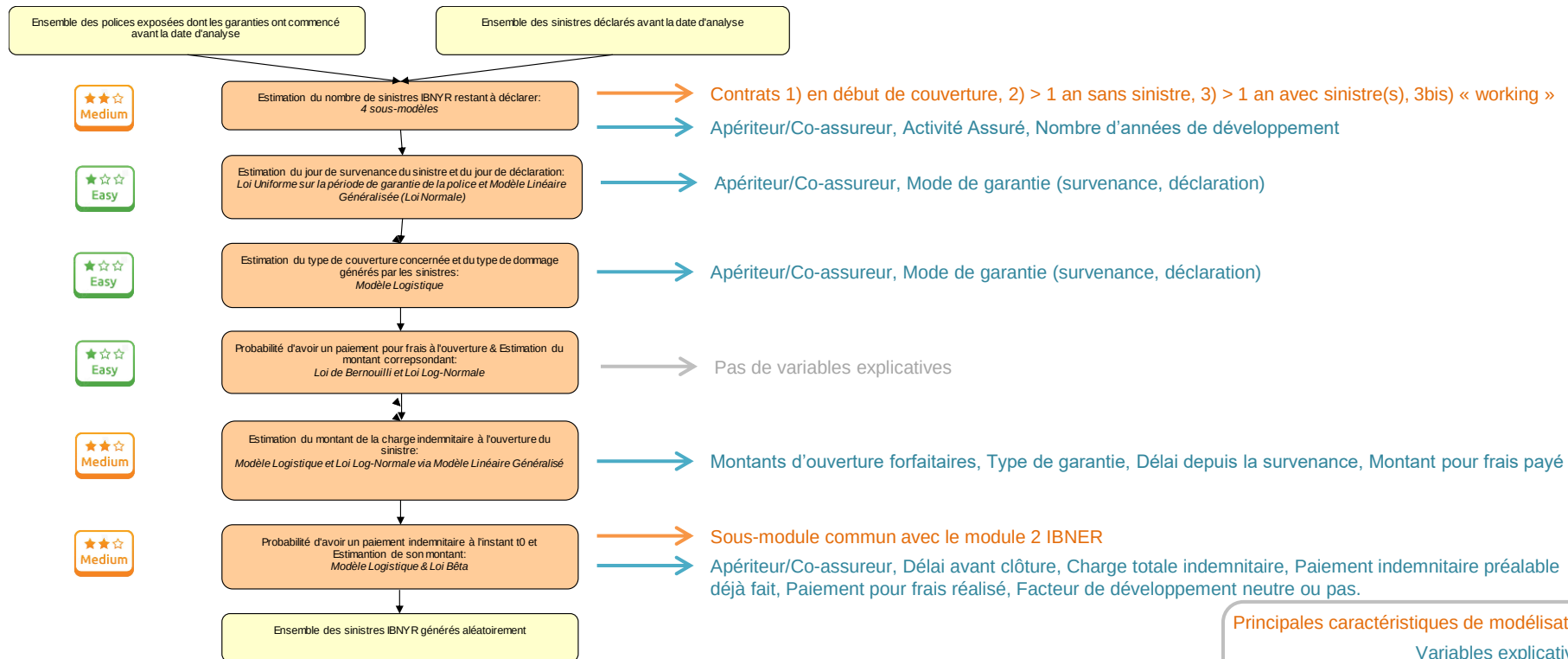


Basé sur le système de gestion des contrats, incluant l'intégralité

- des informations qualitatives et quantitatives des contrats
- des transactions comptables des sinistres



3 • Architecture du modèle sur données détaillées: Module 1 IBNYR

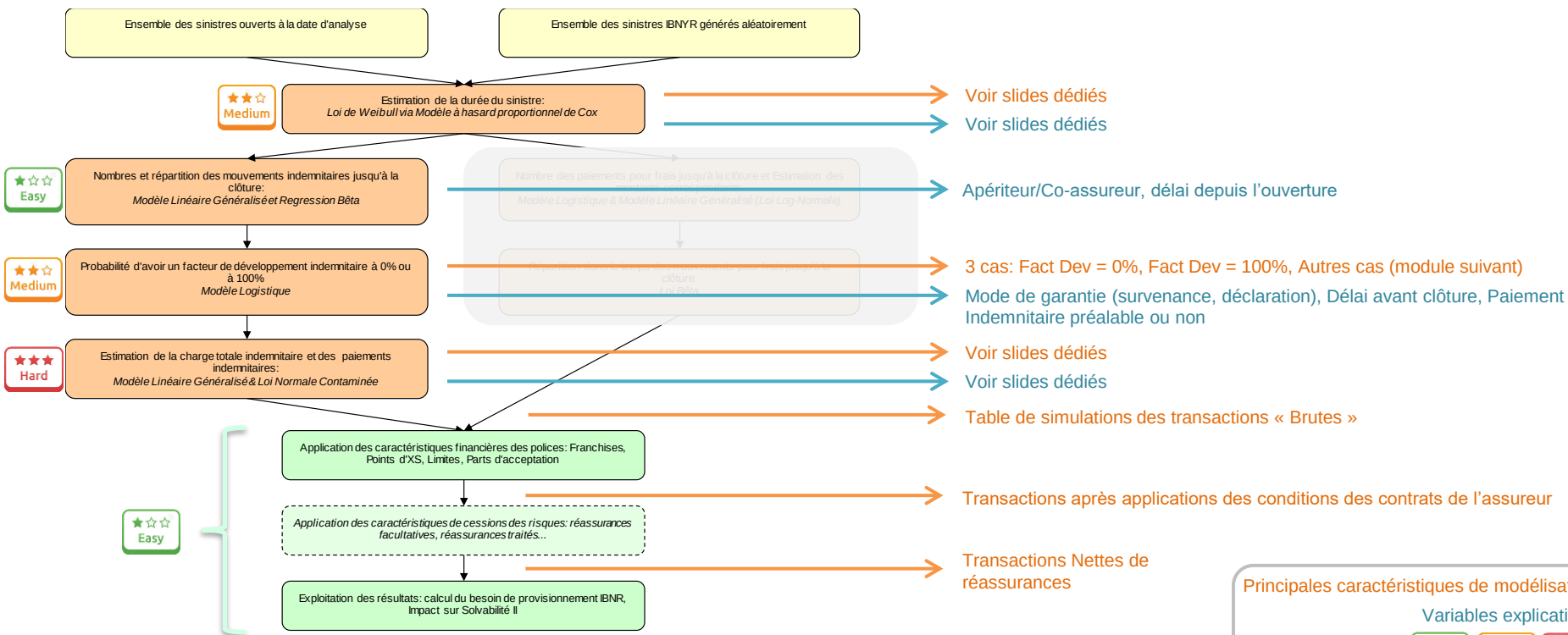


Principales caractéristiques de modélisation

Variables explicatives

Difficulté:



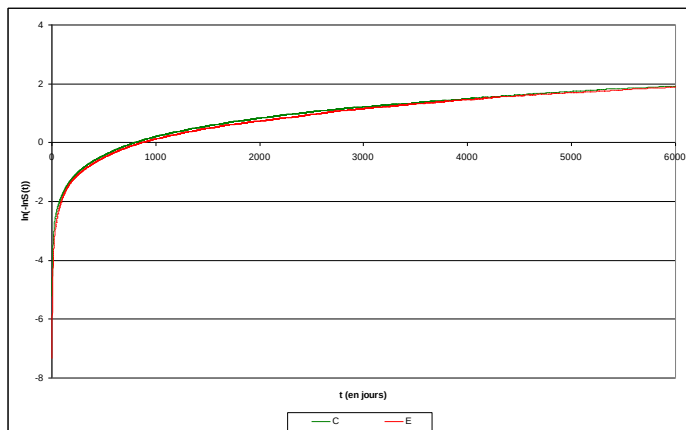


Application du modèle à hasard proportionnel de Cox

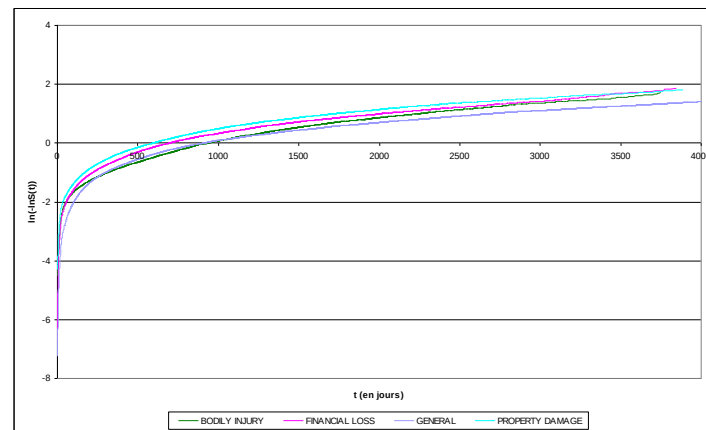
- Modèle Linéaire Généralisé: $g(Y|X) = X\beta$
- Modèle de Cox: avec fonction du risque instantané $h(t|X) = h_0(t)e^{\beta'X}$ et fonction de Survie $S(t/X) = S_0(t)e^{\beta'X}$
- Propriété de proportionnalité dans le temps: $\frac{h(t/X)}{h_0(t)} = e^{\beta'X}$
Le risque d'un individu est proportionnel au risque de base $h_0(t)$ et pas du temps.
- Structure du modèle retenu:
 - Variable de stratification: Apériteur/Co-assureur: 1 fonction du risque de base estimée par classe de stratification $h_{Apé,0}(t)$ et $h_{Coass,0}(t)$
 - Variables explicatives: Type de dommage, Mode de garantie (survenance, déclaration)
- Validation de l'hypothèse de proportionnalité par 2 méthodes:
 - Graphiquement avec par classe de variables explicatives, les courbes $t = \ln(-\ln\hat{S}(t))$ doivent être parallèles.
 - Par régression des co-variables dépendant du temps de chaque variable explicative: doivent être non-significatives.

Application du modèle à hasard proportionnel de Cox

- Validation de l'hypothèse de proportionnalité:



Paramètre		DDL	Valeur estimée	Erreur	Khi-2	Pr > Khi-2
KZ_VERSFALL	C	1	0.29477	0.02551	133.4909	<.0001
LOG_TIME*KZ_VERSFALL	C	1	-0.03548	0.0041	74.7646	<.0001

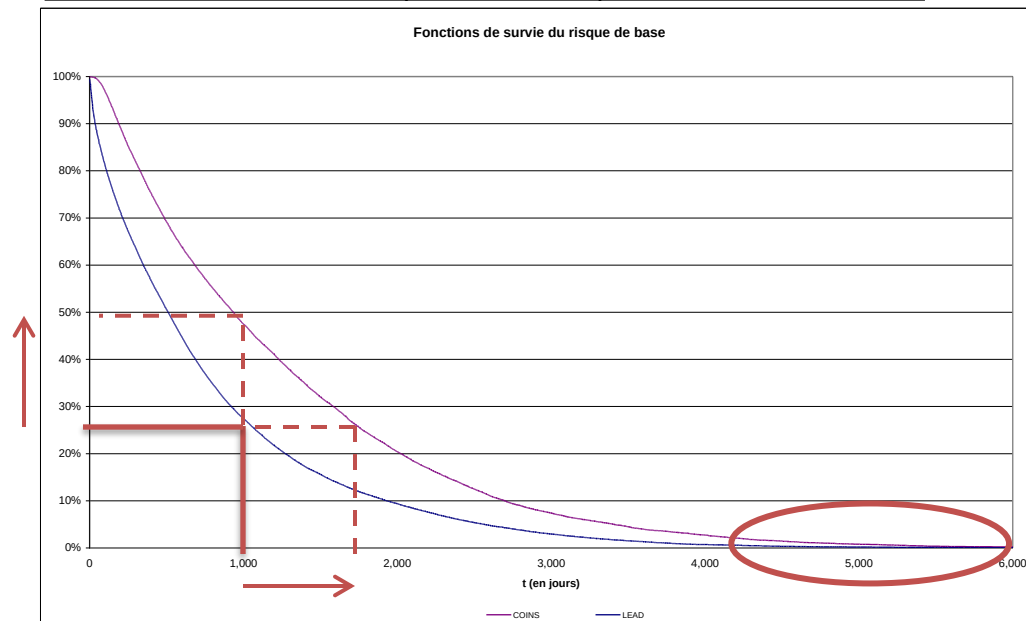


Paramètre		DDL	Valeur estimée	Erreur	Khi-2	Pr > Khi-2
MAIN_CAUSE_BEZ_40	Corporel	1	0.21073	0.04707	20.043	<.0001
MAIN_CAUSE_BEZ_40	Perte Financière	1	0.42023	0.03821	120.9252	<.0001
MAIN_CAUSE_BEZ_40	Matériel	1	0.93228	0.02931	1011.49	<.0001
LOG_TIME*MAIN_CAUSE	Corporel	1	-0.02077	0.00772	7.2356	0.0071
LOG_TIME*MAIN_CAUSE	Perte Financière	1	-0.02299	0.00637	13.0188	0.0003
LOG_TIME*MAIN_CAUSE	Matériel	1	-0.0872	0.00504	298.9637	<.0001



Application du modèle à hasard proportionnel de Cox

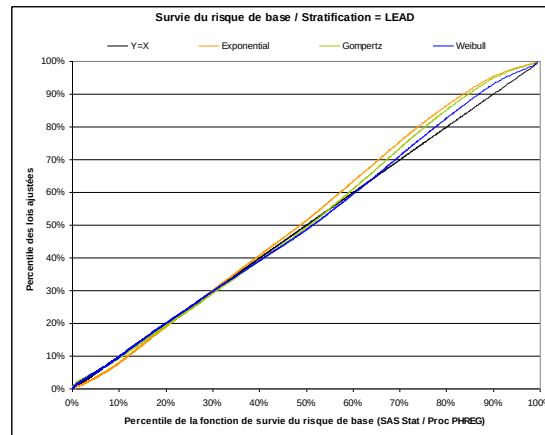
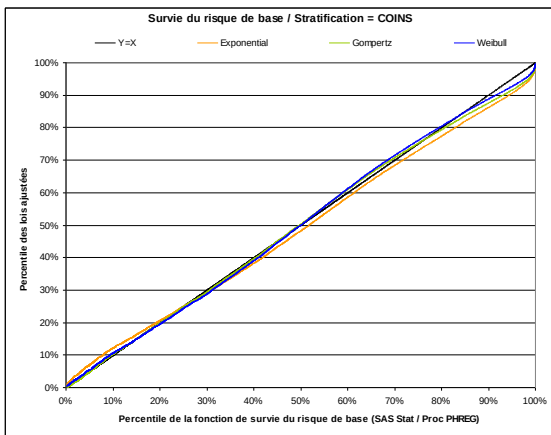
- Fonctions de survie des risques de base Apéritif et Co-assurance:



- 1,000 jours après la survenance d'un sinistre,
Apéritif: 27% des dossiers encore ouverts
Co-assureur: 50% des dossiers encore ouverts
- 700 jours supplémentaires environ (presque 2 ans) nécessaires à la Co-assureur pour atteindre les 27% de dossiers encore ouverts
- Différence limitée sur les dossiers les plus long et probablement les plus complexes.

Application du modèle à hasard proportionnel de Cox

- Ajustement paramétrique de la fonction de survie du risque de base: Exponentiel vs Gompertz vs Weibull*



- Sélection graphique: Ajustement par la loi de Weibull

- Fonction de survie: $S_0(t) = e^{-\lambda t^\nu}$

Fonction de hasard: $h_0(t) = \lambda \nu t^{\nu-1}$

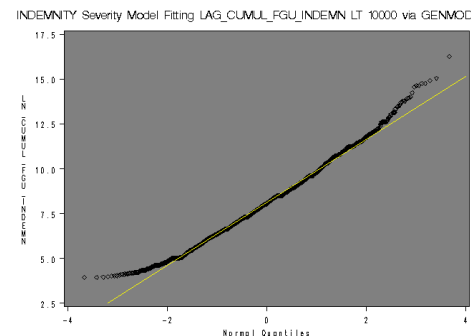
- Simulation de $t + t_0$ par $t + t_0 = \left(t_0^\nu - \frac{\ln(U)}{\lambda e^{\beta'X}}\right)^{\frac{1}{\nu}}$ avec U loi Uniforme.

3 • Architecture du modèle sur données détaillées: Exemple 2 – Module IBNER 2 Modèle Charge Totale Indemnitaire

Estimation de la charge totale indemnitaire

- Rappel: suit un premier modèle logistique qui permet d'estimer les masses de Dirac formées par la probabilité d'avoir un facteur de développement neutre à 1 ou à 0 supprimant la charge indemnitaire, en distinguant les cas où il y a eu au moins un paiement indemnitaire préalable..
- Décomposition en 3 cas:
- 1^{er} Cas: Charge Indemnitaire précédant le mouvement considéré est < 10,000€
- Argument: Pour les faibles montants indemnitaires, le régulateur sinistre réalise une réévaluation du montant plutôt qu'une translation de la valeur précédente.
- Ajustement par Loi Log-Normale:

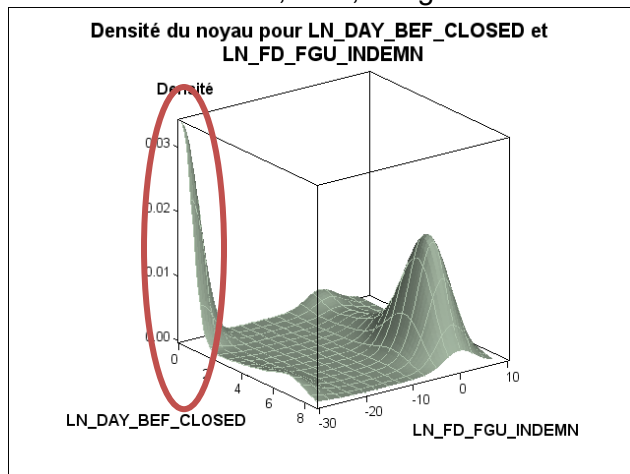
Paramètre	Symbole	Valeur estimée
Mean	Mu	8.1149
Std Dev	Sigma	1.7589



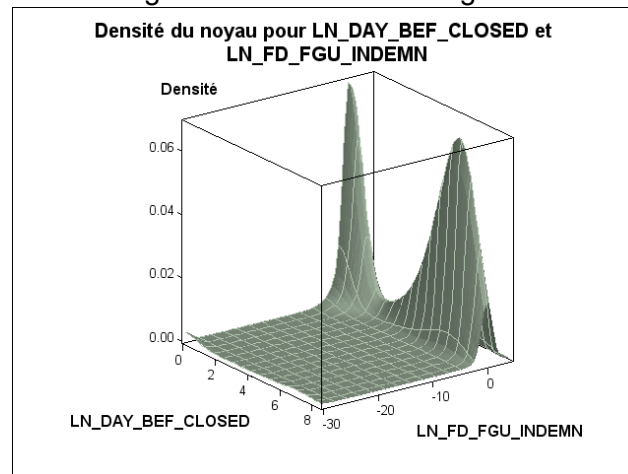
3 • Architecture du modèle sur données détaillées: Exemple 2 – Module IBNER 2 Modèle Charge Totale Indemnitaire

Estimation de la charge totale indemnitaire

- Pour les montants $\geq 10,000\text{€}$, 2 régimes de développement de la charge indemnitaire se distinguent nettement:



Cas **sans** paiement indemnitaire préalable



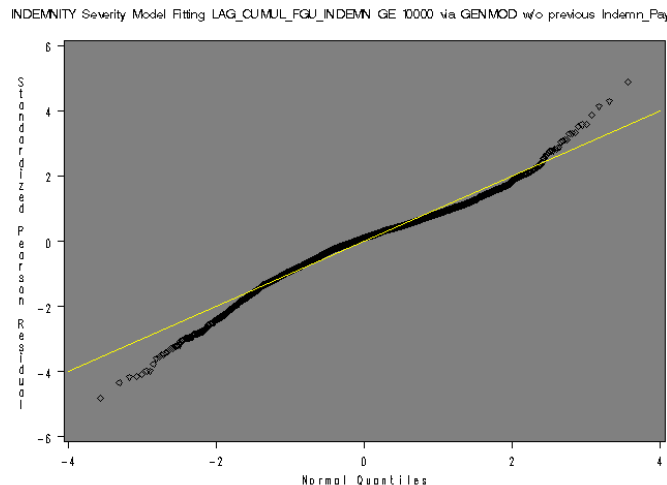
Cas **avec** paiement(s) indemnitaire(s) préalable

- Volatilités des facteurs de développements très différents dans les deux cas
- Dans le cas sans paiement indemnitaire préalable: 80% de chances de ne faire aucun paiement indemnitaire à la clôture

3 • Architecture du modèle sur données détaillées: Exemple 2 – Module IBNER 2 Modèle Charge Totale Indemnitaire

Estimation de la charge totale indemnitaire

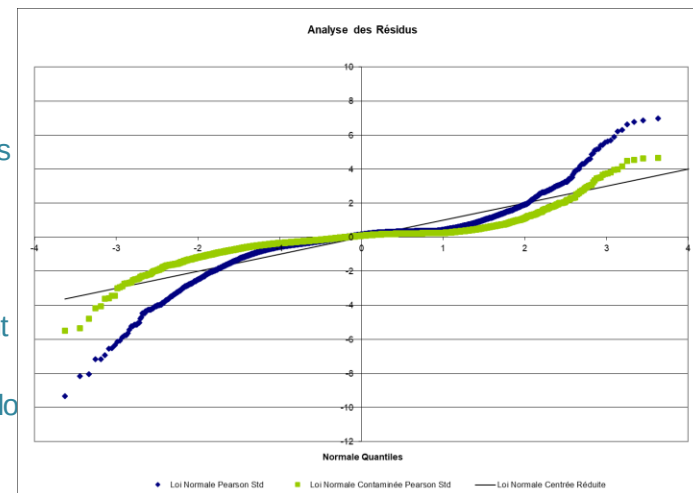
- 2ème Cas: montants $\geq 10,000\text{€}$ et **sans** paiement indemnitaire effectué précédemment
- Modèle GLM:
- Loi Normale
- Fonction lien: identité
- Variables explicatives: Nombre de Jour restant avant la clôture, variable de classes (0.5 avant la clôture, 1 à la clôture)
- Analyse des résidus standardisés de Pearson:



3 • Architecture du modèle sur données détaillées: Exemple 2 – Module IBNER 2 Modèle Charge Totale Indemnitaire

Estimation de la charge totale indemnitaire

- 3^{ème} Cas: montants $\geq 10,000\text{€}$ et **avec** paiement indemnitaire effectué précédemment
- Modélisation en 2 temps
- 1^{er} Modèle GLM pour estimer l'espérance μ_i
- Loi Normale (Proc GENMOD)
- Variables explicatives: Nombre de jours avant la clôture, ratio du cumul des paiements indemnitaires déjà effectué avant le mouvement considéré
- 2^{ème} Modèle GLM (Proc NLMIXED) pour estimer p , σ_1 et σ_2 (Gilbert et Chen, 2007)
- Loi Normale Contaminée CND (Loi mélange de 2 lois normales ayant la même espérance, mais différentes variances)
- Variables explicatives: ratio du cumul des paiements indemnitaires déjà effectué avant le mouvement considéré
- Densité de la loi CND: $f(y_i) = p\Phi\left(\frac{y-\mu_i}{\sigma_1}\right) + (1-p)\Phi\left(\frac{y-\mu_i}{\sigma_2}\right)$ où Φ est la densité de la loi normale centrée réduite.
- et $p = \frac{\exp(zp)}{1+\exp(zp)}$ avec $zp = \alpha X$ où X est la variable explicative
- Voir analyse des résidus standardisés de Pearson (GLM avec et sans CND):



Résultats à l'ultime

- Simulation de Monté-Carlo: 10,000 itérations

	 C-L/Mack Paielements	 GLM/Bootstrap Paielements	 C-L/Mack Charge Totale	 GLM/Bootstrap Charge Totale	 Modèle détaillé
Espérance	146	263	461	434	426
Ecart Type	162	188	43	105	107
Probabilité					
5.0%	105	33	417	305	268
10.0%	111	70	423	326	298
20.0%	120	116	431	355	333
50.0%	139	229	451	415	417
80.0%	166	381	482	499	511
90.0%	184	494	506	564	567
95.0%	201	611	532	631	620
99.0%	248	905	617	783	711
99.5%	278	1,010	659	851	763
99.8%	314	1,301	736	935	810

Valeur de référence pour le BE avant escompte

Similitude à l'ultime sur l'ensemble de la distribution de la charge totale entre le modèle GLM/Bootstrap de la Charge Totale et le modèle détaillé

- Rappel: le montant des provisions des dossiers sinistres à la date de reporting est de €379m
- Avant escompte, les 3 modèles de charges totales indiquent un besoin d'IBNR entre €47m (+15% des provisions existantes) et €82m (+21% des provisions existantes).
- Les deux modèles basés sur les triangles de paiements donnent des valeurs trop optimistes pour être raisonnables: entre 30% et 61% d'excédents de provisions!

Résultats à horizon 1 an

Formule Standard	Modèles Best Estimate				
	C-L/Mack Paielements	C-L/Mack Charge Totale	GLM/Bootstrap Paielements	GLM/Bootstrap Charge Totale	Modèle détaillé
(en million d'euros)					
SCR_{GTPL prem res}	69	171	107	162	159
TP_(nl,GTPL)			379		
MCR	31	49	48	49	49

Modèles Stochastiques	GLM/Bootstrap Paielements	GLM/Bootstrap Charge Totale	Modèle détaillé
Probabilité			
99.0%	459	443	344
99.5%	465	447	375
99.8%	472	452	415
99.9%	479	456	441
SCR_{GTPL prem res}	202	13	-51

Approche simplifiée: $SCR = VaR(99.5\%)_{1\text{ an}} - Exp_{ultime}$

- Formule Standard indique un SCR autour €160m pour le GLM/Bootstrap de Charge Totale et le modèle détaillé
- Modèle Stochastiques:
- GLM/Bootstrap de Charge Totale et le modèle détaillé, le SCR est très limité, même l'espérance à l'ultime du modèle détaillé est en excédent de sa VaR 99.5% à 1 an.
- le GLM/Bootstrap sur Paielements: effet de vase communicant entre l'espérance à l'ultime et le SCR.
- Questions:
- Impact de la différence de cadence de règlement entre le GLM/Bootstrap Charge Totale et le modèle détaillé ?
- Vision à 1 an pertinente pour ce portefeuille pour estimer sa solvabilité (hors effets de réassurance)?



Modèle Détaillé: Potentiels et Limites



- Granularité de la modélisation et des résultats: application précise des conditions du contrat et de la réassurance
- Estimation du Business Plan
- Aide à l'analyse de changement réglementaire ou de politique de souscription
- Support d'analyse et de décision opérationnel
- Management des données utilisées cruciales
- Maintenance et Validation de l'ensemble des sous-modèles
- Approche ad/hoc: coût/bénéfice à bien estimer dans le cadre d'un Modèle Interne (contrôles, validation, documentation, fluidité des échanges avec le régulateur)

Points clés de l'évolution de l'état de l'art depuis 2013

- Confirmation de l'intérêt des variables explicatives dans la précision du modèle par rapport à la perte d'information des modèles sur données agrégées
- Confirmation de l'amélioration de la précision de la compréhension du portefeuille en temps continu.
- Confirmation de l'intérêt d'un modèle de survie pour caractériser la durée des sinistres dans les comptes de l'assureur
- Nouveauté: Modèles non-paramétriques (type machine learning) ayant une meilleure capacité d'adaptation aux changements de comportements du portefeuille, quelle que soit son origine.
- Confirmation de la pertinence de l'approche ad-hoc pour bien adapter le modèle aux risques du portefeuille et ses caractéristiques opérationnelles.

Conclusion

- Conception d'un modèle sur données détaillées incluant la valeur ajoutée de variables explicatives issues des caractéristiques quantitatives et qualitatives des contrats et des sinistres.
- Les résultats du modèle détaillé et de l'approche GLM à bootstrap de la charge totale convergent
- Meilleures capacités du modèle détaillé les conditions des contrats et de la réassurance, parfois fait sur mesure, et donc très difficilement appréhendable par les méthodes utilisant des données agrégées.
- Nécessité d'approfondir la validation des sous-modèles du modèle détaillé
- La littérature écoulée depuis la conception du modèle détaillé confirme plusieurs choix optés, dont l'utilisation d'un modèle de survie ou encore d'être en temps continu.
- Potentiel d'amélioration important: utilisation de l'IA et d'approches non-paramétriques
- Le modèle détaillé: outil d'aide à la décision, aussi bien au niveau stratégique qu'opérationnel

MERCI DE VOTRE ATTENTION