COLLOQUE ASTIN 1963 DE TRIESTE

Rapport introductif sur le 2e thème, partie a

P. THYRION

Bruxelles, Belgique

Le ,,Contrôle actuariel dans la comptabilité des branches élémentaires'' constitue le 2e thème de ce colloque, tel que l'annonce le programme officiel. Ce thème comprend une 1ère partie intitulée: ,,Problèmes relatifs aux sinistres'', qui est le sujet de cette séance.

Les notes préliminaires de notre Président, parues dans notre dernier bulletin, suggéraient d'orienter les travaux vers les applications pratiques de l'important travail théorique déjà réalisé dans le cadre d'Astin. Plus particulièrement en ce qui concerne les sinistres dans les assurances de responsabilité civile, ces notes attiraient l'attention sur une certaine analogie que la longue durée de règlement des sinistres pouvait créer entre l'estimation des réserves pour sinistres à régler et l'évaluation des réserves mathématiques en branche Vie et elles soulevaient dès lors assez naturellement la question des méthodes d'estimation collective. En fait, les communications qui ont été rattachées à cette 1ère partie du 2e thème interprètent le sujet de manière extrêmement large. Ne voyez dans cette réflexion initiale nul reproche aux auteurs, mais la simple constatation d'un fait dont mon rapport introductif doit tenir compte: il ne sera pas une synthèse dans les directions proposées par notre Président.

Une partie des communications se rattache au fond à la théorie du risque. Pour mieux dégager l'apport de ces notes, je me propose de le dépouiller autant que possible de son aspect purement mathématique et de le situer dans l'évolution des idées et des recherches en la matière.

J'indiquerai ensuite l'apport des communications au sujet de l'estimation des réserves pour sinistres à régler.

Enfin, je résumerai les travaux qui se rapportent à la tarification du risque automobile.

https://doi.org/10.1017/S0515036100010564 Published online by Cambridge University Press

La théorie du risque repose essentiellement sur l'étude de la variable aléatoire „coût total des sinistres survenus pendant un intervalle de temps déterminé pour un risque ou un ensemble de risques". Cette variable dépend elle-même de la variable „nombre de sinistres survenus pendant cet intervalle de temps" et de la variable „coût d'un sinistre si l'on sait qu'il est survenu". Autrement dit, il s'agit d'étudier une variable qui est la somme d'un nombre aléatoire de variables elles-mêmes aléatoires.

L'évolution dans le temps de la variable „coût total des sinistres" $X(t)$ se représente donc graphiquement par un escalier doublement aléatoire, par le nombre de ses marches et par la hauteur de chacune

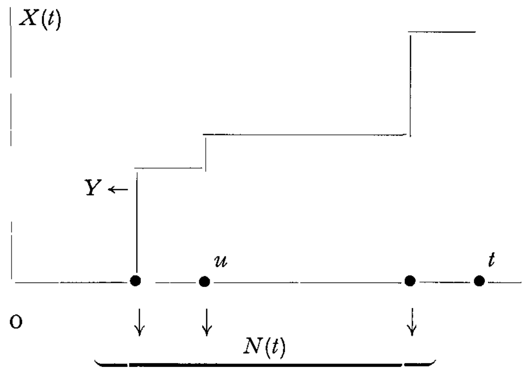

Fig. 1

d'elles. C'est un schéma fort général, se prêtant à des applications dans bien d'autres domaines que les assurances: diverses théories développées pour les besoins de la recherche opérationnelle y font appel, qu'il s'agisse de problèmes de stockage ou de files d'attente par exemple.

Synthétiser les résultats obtenus dans l'étude de cette variable $X(t)$, revient au fond à passer en revue les hypothèses qui ont été faites successivement sur les variables $N(t)$ et Y qui la déterminent. Ces hypothèses sont d'ordre général ou bien sont des particularisations des lois de probabilité des variables $N(t)$ et Y. Comme dans tout développement scientifique, elles ont permis de passer petit à petit du simple au composé. Elles ont été faites un peu sans doute sous l'effet du penchant irrépressible des mathématiciens à „généra-

https://doi.org/10.1017/S0515036100010564 Published online by Cambridge University Press

liser des résultats précédents" selon la formule consacrée ou bien à prospecter de nouveaux filons; mais il serait injuste de ne pas reconnaître qu'elles ont été influencées surtout par le désir de créer des schémas se pliant de manière suffisante à une réalité souvent très complexe et pouvant ainsi être utilisés pour résoudre des problèmes pratiques.

Dans la question qui nous occupe, on peut tout d'abord supposer que les variables $N(t)$ et Y sont stochastiquement indépendantes et que les variables Y successives sont elles-mêmes stochastiquement indépendantes et identiques, c'est-à-dire notamment qu'elles ne dépendent pas de t.

Dès lors, au point de vue du calcul des probabilités, il s'agit d'étudier la fonction de répartition de la variable $X(t)$, soit

$$F(x\,;t) = \sum_{n=0}^{\infty} P(n\,;t)\,.\,S^{n*}(x)$$

où $P(n\,;t)$ est la loi de probabilité de la variable $N(t)$, nombre de sinistres pendant l'intervalle de temps (o, t);

$S(y)$ est la fonction de répartition de la variable Y, coût d'un sinistre si on sait qu'il est survenu;

$S^{n*}(y)$ est la convoluée d'ordre n de $S(y)$ avec elle-même ($S^{0*}(y) = 1$).

On peut dire aussi qu'il s'agit d'étudier la fonction caractéristique (f. c.)

$\psi(z\,;t)$ de $X(t)$ donnée par

$$\psi(z\,;t) = \sum_{n=0}^{\infty} P(n\,;t).\ \zeta^n$$

si ζ est la f.c. de Y.

Sous cette dernière forme — et au point de vue de la théorie des f.c. — il s'agit d'un problème de transformation linéaire de f.c.

Si l'on ne veut faire aucune hypothèse sur la forme analytique des lois de probabilité des variables $N(t)$ et Y, il est clair que le schéma indiqué ne peut fournir que les moments des divers ordres de la variable $X(t)$ en fonction des moments de $N(t)$ et de Y. Les moments de $X(t)$ s'obtiendront aisément en dérivant successivement sa f.c. et en utilisant la relation bien connue entre les dérivées d'une f.c. au point o et les moments correspondants de la variable aléatoire. On aura ainsi.

$$E_X = E_N \,.\, E_Y$$

$$E_{X^2} = E^2{}_Y \,(E_{N^2} - E_N) + E_N \,.\, E_{Y^2}$$

$$\text{D'où } \sigma^2{}_X = E^2{}_Y \,.\, \sigma^2{}_N + E_N \,.\, \sigma^2{}_Y$$

Ceci est en bref la position adoptée et les résultats obtenus par H. Bühlmann. En effet, si l'on considère que la variable $X\ (t)$ concerne un risque, le moment et la variance de la variable $X_m\ (t)$, moyenne des M variables $X\ (t)$ concernant M risques indépendants et identiques seront respectivement

$$E_{Xm} = E_X,$$

$$\sigma^2{}_{Xm} = \frac{\sigma^2{}_X}{M}$$

Et l'on obtiendrait de même les moments d'ordre supérieur.

Je doute que le sentiment de malaise que H. Bühlmann dit éprouver à l'égard d'hypothèses supplémentaires soit partagé par tous. Je ne suis également pas convaincu que ces hypothèses supplémentaires soient faites uniquement par raison de convenance mathématique. On peut penser que certains problèmes pratiques exigent plus que la détermination des moments de la variable. Quoi qu'il en soit, le point de vue de H. Bühlmann a le mérite d'attirer l'attention sur la nécessité d'analyser ,,honnêtement'' les hypothèses que l'on utilise. Mais il me semble qu'il peut être utile d'en poser, à condition que l'on n'oublie pas dans les applications que ce ne sont que des hypothèses qu'il faut tester avant d'en déduire des résultats pratiques. Procéder de cette manière n'est pas une infirmité propre à la science actuarielle.

Nous allons donc parcourir rapidement quelques hypothèses supplémentaires sur la nature des lois de probabilité qui ont été faites successivement pour affiner le schéma fondamental.

La première hypothèse faite sur la variable ,,nombre de sinistres'', a été d'identifier sa loi de probabilité à la loi de Poisson.

$$P\ (n\ ;\ t) = e^{-\lambda t}\ \frac{(\lambda\, t)^n}{n!}\ (\lambda \text{ constante})$$

A côté de son caractère spécifique, la théorie collective du risque offre un large développement analytique de ce type de processus stochastique discontinu que certains auteurs appellent loi de Poisson non élémentaire et d'autres, loi de Poisson généralisée. Il est bien évident que les résultats analytiques acquis dans l'étude de ce schéma sont valables aussi bien en dehors du concept ,,théorie collective du risque'' qu'à l'intérieur de celui-ci.

Une hypothèse plus nuancée sur la variable ,,nombre de sinistres'' est l'hypothèse d'hétérogénéité dont nous verrons une application plus loin. Elle consiste à supposer que le risque considéré fait partie d'un ensemble qui se subdivise en groupes homogènes à l'intérieur de chacun desquels la loi de probabilité de $N(t)$ est une loi de Poisson, où l'intensité du processus λ varie de groupe en groupe. Au point de vue du calcul des probabilités cela revient à considérer que λ est une variable aléatoire dont la fonction de répartition $U(\lambda)$ définit la structure de l'ensemble hétérogène en groupes homogènes. La variable $N(t)$ obéit alors à une loi de Poisson composée au sens restreint.

$$P(n\,;t) = \int_0^\infty e^{-\lambda t}\,\frac{(\lambda\,t)^n}{n!}\;dU(\lambda)$$

On peut en outre supposer que la fonction $U(\lambda)$ dépend du temps et on a alors une loi de Poisson composée au sens large, selon la dénomination de C. Philipson.

Les lois de probabilité précédentes concernant la variable $N(t)$ impliquent qu'il ne peut se produire qu'un seul sinistre pendant un intervalle de temps élémentaire $(t, t + dt)$.

Un pas supplémentaire a naturellement été franchi à cet égard: c'est l'étude de schémas où plusieurs sinistres peuvent survenir simultanément et où $N(t)$ obéit à ce que l'on peut appeler de manière imagée une distribution ,,par grappes'', dans laquelle la loi $P(n\,;t)$ conjugue deux lois

— la loi de probabilité du nombre de grappes ou lots de sinistres,
— la loi de probabilité du nombre de sinistres dans une grappe.

Le bon sens permet de pressentir et la théorie confirme que la fonction de répartition d'un processus stochastique discontinu $X(t)$ dont la loi de probabilité du nombre de sinistres est une distribution par grappes, est — sous des conditions fort générales —

équivalente à la fonction de répartition d'un processus où la loi de probabilité du nombre de sinistres serait simplement celle du nombre de grappes et ce moyennant une transformation adéquate de la loi de probabilité du coût d'un sinistre. Cette propriété permet d'appliquer à des schémas apparemment compliqués des développements mathématiques établis pour des schémas plus simples. Ainsi quand la loi du nombre de sinistres est une loi de Poisson par grappes, l'étude du schéma peut se conduire comme si la loi du nombre de sinistres était une loi de Poisson pure et l'on peut donc transposer les développements mathématiques valables pour ce dernier type de schéma — quand les conditions requises sont réalisées bien entendu.

Nous avons supposé jusqu'ici que la variable Y, coût d'un sinistre, ne dépendait pas du temps. On peut assouplir cette hypothèse et considérer le cas où la loi de probabilité de Y dépend de l'instant u où survient le sinistre, u étant compris dans l'intervalle $(0, t)$ où l'on étudie le processus.

Dans la théorie collective du risque, il a été démontré que lorsque la loi de probabilité du nombre de sinistres est une loi de Poisson, l'on peut considérer que la variable Y ne dépend pas de u et la remplacer par une variable moyenne $\overline{Y}\ (t)$, dont la f.c. est donnée par

$$\bar{\zeta}_t = \frac{1}{t} \int_0^t \zeta_u \cdot du, \; \zeta_u \text{ étant la f.c. de } Y(u).$$

C. Philipson a étendu cette propriété successivement à diverses lois de Poisson composées pour en arriver finalement à apercevoir qu'elle reposait essentiellement sur une relation commune à toutes les lois de Poisson composées au sens restreint. Dès lors il nous apporte, en annexe à sa note principale, un résultat qui englobe et généralise les résultats antérieurs, en ce sens que la transformation indiquée ci-dessus peut être faite dans tous les cas où la loi de probabilité du nombre de sinistres est une loi de Poisson composée au sens restreint; donc, dans ce cas, la f.c. de la variable $X\ (t)$ peut toujours s'écrire sous la forme canonique

$$\psi\ (\zeta\ ; t) = P\ [0\ ; t\ (1 - \bar{\zeta}_t)].$$

Jusqulà présent le modèle classique de la théorie du risque impliquait que le coût d'un sinistre était bien déterminé au moment où

https://doi.org/10.1017/S0515036100010564 Published online by Cambridge University Press

ce sinistre survenait. Il n'était pas tenu compte de l'influence sur ce coût des événements se produisant pendant la période qui suivait la survenance du sinistre, c'est-à-dire pendant la période où les paiements relatifs à ce sinistre étaient faits.

Ceci se traduit analytiquement par le fait que la variable Y ne dépend pas du tout du temps, ou bien comme nous venons de le dire qu'elle dépend de l'instant u de survenance du sinistre. C. Philipson s'est attaché à étendre ce modèle. Nous n'avons pas attendu la présente note de C. Philipson pour apprendre qu'il ne recule pas devant la complexité d'un schéma. Je dois donc me borner à dégager l'idée maîtresse de son travail à cet égard.

Dans le processus généralisé du risque qu'il envisage, la survenance d'un sinistre entraîne une série de paiements se succédant dans le temps à partir de l'instant de survenance du sinistre. On distingue donc un processus primaire qui régit le nombre de sinistres survenus pendant un certain intervalle de temps et un processus secondaire qui régit, pour chaque sinistre survenu, le nombre et la grandeur des divers paiements. Le processus généralisé est constitué par le montant total des paiements faits pendant un certain intervalle de temps et afférents aux sinistres survenus pendant un autre intervalle de temps. On trouvera donc en principe dans ce processus généralisé deux mesures de temps. Je dirais volontiers que le nombre total des paiements pour l'ensemble des sinistres du processus généralisé répond à une distribution par grappes échelonnées dans le temps, qui est évidemment plus compliquée qu'une distribution par grappes instantanées analogue à celles que j'ai évoquées tout à l'heure. Moyennant certaines modifications aux processus primaire et secondaire tels que définis ci-dessus et en utilisant la propriété que j'ai commentée plus haut, C. Philipson ramène le processus généralisé à un schéma mathématique qu'il a étudié précédemment; dans ce schéma la loi de probabilité de la variable „nombre de paiements" obéit, dans sa forme fondamentale, à une distribution par grappes où la loi du nombre de grappes et la loi du nombre d'éléments dans une grappe sont toutes deux des lois de Poisson composées.

A côté de cette extension du processus classique du risque, H. Ammeter en amorce une autre modification que l'on pourrait appeler une amputation. Sa note part de la considération que les sinistres les plus élevés peuvent d'une part être éliminés par la réassurance et d'autre part être une source d'irrégularités dans les résultats statistiques; il est donc intéressant d'étudier la loi de distribution du coût total des sinistres après élimination des plus élevés en faisant appel à la théorie des valeurs extrêmes. En fait, la présente note ne considère que le cas du sinistre le plus élevé; elle suppose également que la loi de probabilité du nombre de sinistres est la loi de Poisson et elle donne dans ce cas les formules de la loi de répartition, de la grandeur moyenne et de la variance de la variable „coût total amputé". On peut observer d'ailleurs que ces formules s'établiraient aussi aisément si la loi de probabilité du nombre de sinistres était une loi quelconque. H. Ammeter applique cette théorie au cas où la loi de distribution du coût d'un sinistre est une loi de Pareto. Il constate ainsi, ce qui me paraît conforme au bon sens, que le domaine d'existence des moments de la loi „amputée" est plus large que celui des moments correspondants de la loi „totale". Il est clair, comme l'auteur l'annonce, que le travail ainsi entamé peut se poursuivre dans le cas où l'on exclut plus d'un gros sinistre.

* * *

Abordons maintenant la question de l'estimation des réserves pour sinistres à régler.

On peut la raccrocher — en théorie du moins — à ce qui vient d'être dit sur la variable „coût total des paiements". Il suffit d'exprimer, à un moment donné, la réserve pour sinistres à régler, pour les sinistres survenus pendant un intervalle de temps déterminé, par la différence entre le coût total des paiements jusqu'à un instant suffisamment éloigné pour que tous les sinistres soient entièrement réglés et ce même coût total à l'instant considéré.

C'est ce que fait C. Philipson qui traite la question comme une illustration de l'applicabilité de la théorie développée dans sa note, en introduisant par la même occasion des formules inspirées de la technique de l'assurance sur la vie. Si l'on admet que le déroulement des paiements dans le temps est un phénomène stationnaire, on aboutit à la méthode tout à fait intuitive qui consiste à évaluer la réserve pour sinistres à régler en multipliant les paiements déjà effectués réellement à l'instant de l'estimation par un coefficient

https://doi.org/10.1017/S0515036100010564 Published online by Cambridge University Press

lié à la durée de temps écoulée depuis l'origine de la période de survenance des sinistres. C'est précisément cette question de stationnarité qui peut se révéler délicate dans la pratique pour certains risques.

La communication de C. P. Welten sur l'assurance individuelle, en Hollande, des frais de traitement dans un sanatorium pour cause de tuberculose, relève essentiellement de la 2e partie de ce thème. Je me bornerai donc à signaler qu'il nous donne notamment la formule de calcul des réserves pour sinistres à régler. Cette formule se base sur le groupement des malades en traitement au moment du calcul, selon la durée du traitement déjà écoulée, arrondie au mois supérieur, et sans tenir compte de leur âge. Elle implique la détermination statistique du nombre moyen de jours pendant lesquels un malade, qui est au sanatorium à la fin du Te mois après son admission, restera encore en traitement; ce nombre moyen est lui-même lié à la probabilité mensuelle de fin de traitement par guérison ou décès et s'obtient par un calcul analogue à celui de l'espérance de vie.

* * *

Dans le schéma général de la théorie du risque tel que nous l'avons rappelé, on peut faire abstraction de la variable „coût d'un sinistre" et porter son attention spécialement sur la variable „nombre de sinistres". C'est ce qui est fait dans deux travaux concernant le risque automobile qui reste toujours d'actualité.

B. Dubois de Montreynaud propose une méthode de sondage pour l'étude rapide de l'influence de nouveaux critères sur le tarif „automobile". Cette méthode consiste à constituer l'échantillon par l'ensemble des véhicules sinistrés au cours d'une période fixée a priori; ce procédé ayant l'avantage notamment de faciliter l'interview du conducteur que son sinistre amènera à prendre contact avec sa compagnie. B. Dubois de Montreynaud part du concept de groupe homogène de véhicules c'est-à-dire d'ensemble de véhicules dont la variable „nombre de sinistres" obéit à la même loi de Poisson. On sait alors que la variable intervalle de temps entre deux sinistres consécutifs d'un même véhicule obéit à une loi de probabilité exponentielle dont le paramètre est la fréquence annuelle de sinistres si le temps est mesuré en années. La grandeur moyenne

https://doi.org/10.1017/S0515036100010564 Published online by Cambridge University Press

d'une telle variable est l'inverse de la fréquence. Dès lors si l'on observe pendant le temps T un groupe fermé homogène de véhicules ayant déjà subi un sinistre, on constatera que S sinistres frapperont ce groupe. Les véhicules affectés par ces sinistres constitueront l'échantillon qui sera tel que la moyenne des temps séparant chacun de ces S sinistres du sinistre précédent donnera une mesure de l'inverse de la fréquence. Le raisonnement s'étend à un ensemble hétérogène de groupes homogènes.

En pratique tout groupe ainsi observé ayant une ancienneté limitée, un tel échantillon contiendra des véhicules qui n'ont pas été sinistrés précédemment. L'auteur propose une méthode approchée de détermination de la date fictive du sinistre virtuel précédant la prise de garantie; cette méthode revient à supposer que l'entrée dans l'assurance des véhicules étudiés a été linéaire dans le temps.

Poursuivant ses travaux en la matière, P. Delaporte approfondit le problème de la tarification du risque automobile. Ainsi qu'il a déjà été constaté précédemment, la variable ,,nombre de sinistres'' pour des véhicules appartenant à la même classe de tarif définie par des critères a priori, n'obéit pas à une loi de Poisson mais bien à une loi de Poisson composée traduisant le caractère hétérogène de l'ensemble des risques ainsi qu'on l'a rappelé tout à l'heure. Dès lors on peut estimer progressivement le risque individuel en tenant compte du nombre de sinistres réellement survenus pendant la période d'observation.

Dès le colloque de La Baule, il avait été proposé d'estimer — dans ce cas-le risque individuel par son espérance mathématique a posteriori, c'est-à-dire liée à l'information déjà obtenue; ceci conduit à une correction progressive du tarif moyen initial par un système de bonus-malus qui maintient le niveau global théorique des recettes. P. Delaporte montre que c'est la méthode la plus efficace au sens de la variance minimum.

Il examine les aménagements que l'on peut apporter aux résultats généraux pour présenter commercialement un tel tarif à prime modelée sur le risque. Tout d'abord les coefficients correctifs ne dépendent que du nombre total de sinistres observés et non de leur répartition entre chacune des années de survenance: c'est ce que Polya avait appelé le phénomène de l' ,,influence globale''. Ensuite on ne peut évidemment multiplier trop le nombre de ces coefficients:

il faut donc les arrondir. On peut estimer qu'il y a une limite pratique aux majorations que l'on peut imposer: dès lors il faut recalculer le système théorique des coefficients en tenant compte de cette limite et d'une condition d'équilibre. P. Delaporte introduit aussi dans le calcul des coefficients, une probabilité de non-paiement de majorations cependant prévues au tarif. Il étudie également les problèmes de stratégie posés par la tarification du risque automobile.

Pour qui se souvient des discussions parfois passionnées que les divers aspects de la question du bonus ont allumées autrefois à nos colloques, il serait étonnant que la solution étudiée ici n'en soulève pas à nouveau. On doit en tout cas reconnaître que la tarification modelée sur le risque s'accorde fort bien avec les observations faites a posteriori sur la variable ,,nombre de sinistres''.

Nous voici au terme de cette revue. La diversité des sujets sur lesquels elle a porté m'autorise, je l'espère, à ne pas conclure par une synthèse, comme il se devrait. Mais je puis terminer en affirmant que les notes dont je n'ai pu exposer que les axes, ne manquent certainement pas de ,,substantifique moëlle'' et par conséquent d'aliment pour nos échanges de vues, quelle que soit la direction qu'ils emprunteront.