Zeitschrift: Mitteilungen / Vereinigung Schweizerischer Versicherungsmathematiker = Bulletin / Association des Actuaires Suisses = Bulletin / Association of Swiss Actuaries

Herausgeber: Vereinigung Schweizerischer Versicherungsmathematiker

Band: 51 (1951)

Artikel: Étude statistique du risque mathématique dans l'assurance collective sur la vie

Autor: Hagstroem, K.-G.

DOI: https://doi.org/10.5169/seals-555005

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 02.07.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

Etude statistique du risque mathématique dans l'assurance collective sur la vie

Par *K.-G. Hagstroem*, Stockholm

I. Introduction

Dans ce qui suit, nous présenterons une série de résultats émanant de recherches faites sur un portefeuille d'assurances de groupes en cas de décès et d'invalidité. Ces résultats sont destinés à illustrer les questions de risque qui sont caractéristiques pour de telles assurances. Les données statistiques utilisées dans nos recherches nous ont été fournies par la «Bâloise» à laquelle nous tenons à exprimer ici notre gratitude pour son obligeance. Une communication préalable sur une statistique analogue d'une compagnie américaine sera publiée ailleurs.

Après quelques préliminaires sur le mode de présenter une statistique de cette espèce — vu qu'il s'agit d'une question stochastique, en empruntant le terme général de *Jacques Bernoulli* —, nous allons donner la statistique dans un état assez rudimentaire, par des tables et à l'aide de graphiques. Quant à l'utilisation finale du matériel pour la théorie abstraite du risque mathématique, nous la laisserons à des recherches futures. Nous ajouterons quelques mots sur «l'assurance des grands écarts», qui serait bien désirable, mais dont la création présente beaucoup d'obstacles, dignes d'être profondément scrutés.

Le problème fondamental de la théorie du risque mathématique peut s'énoncer de la manière suivante. Soit η la variation de la réserve mathématique *calculée* pour une période d'observation quelconque (dans le cas le plus simple: le nombre calculé de décès) et soit ξ la valeur correspondante *observée*. En supposant qu'on a passablement réussi à prévoir *en moyenne* la quantité ξ, quelle sera la *distribution* à conjecturer pour la variable ξ autour de la valeur probable η, ou bien quel est *l'erreur probable* du pronostic stochastique de la variation de la

réserve? Il est évident que la connaissance de la distribution rendrait possible la formation de réserves appropriées à couvrir le risque mathématique de l'industrie de l'assurance, sans que l'on eût besoin de recourir à des procédés mécaniques et grossiers pour la formation d'un fonds de sûreté suffisant. Nous nous restreignons ici au risque mathématique proprement dit, causé par le hasard de la mortalité et de la morbidité, en écartant la question de l'intérêt. Or les considérations de cette espèce donnant lieu à des réserves de risque assez modiques, si l'on applique les théorèmes de la théorie classique des probabilités, on se demande si cette théorie est suffisante ou s'il faut la compliquer en acceptant par exemple d'autres lois combinatoires que celles adoptées dans la théorie classique. Le seul moyen d'étudier à fond ces questions, c'est de multiplier les expériences et de faire des statistiques étendues. L'assurance collective présente l'avantage de fournir des séries de groupes naturelles où l'on peut étudier simultanément la correspondance des résultats réels avec les pronostics. Il y aura donc un intérêt général à amasser beaucoup de statistiques de cette branche et de les présenter d'une manière accessible. C'est ce que nous avons cherché à faire dans ce travail.

II. La fonction statistique à étudier

En observant deux «variables statistiques», disons ξ et η (dans ce qui suit elles seront: le nombre réel de décès et le nombre calculé correspondant), on étudie l'une en «fonction statistique» de l'autre, en donnant *la fonction de corrélation* $F(x, y)$, telle que la différentielle

$$F(x, y)\, dx\, dy$$

exprime la probabilité pour que, simultanément, ξ et η se trouvent dans les intervalles $\xi = x \ldots x + dx$, $\eta = y \ldots y + dy$. La généralisation aux cas discontinus, partiellement ou totalement, au moyen d'intégrales de *Stieltjes*, n'offre aucune difficulté sérieuse. D'une façon générale on peut, de deux manières différentes, séparer la différentielle en deux facteurs, à savoir

$$F(x, y)\, dx\, dy = p(x)\, dx \quad \alpha(x, y)\, dy,$$

ou bien, selon l'autre des axes de coordonnées,

$$F(x, y)\, dx\, dy = q(y)\, dy \quad \beta(y, x)\, dx,$$

où $\alpha(x, y)$ et $\beta(y, x)$ sont des fonctions «génératrices» de η par rapport à ξ ou de ξ par rapport à η, ou bien, les «probabilités conditionnelles» respectives. La fonction $\alpha(x, y)$ est la distribution de η *sous la condition que* ξ *ait pris la valeur* x *ou une valeur dans l'intervalle* $x \ldots x + dx$.

Dans les recherches théoriques, on peut souvent construire une fonction schématique ou à espérer, en ayant recours à $q(y)$ et $\beta(y, x)$, par exemple. Ayant calculé ainsi la fonction $F(x, y)$, on peut désormais calculer $p(x)$ et $\alpha(x, y)$. On a par exemple

$$p(x) = \int F(x, y)\, dy,$$

l'intégrale étant, comme toujours lorsque les bornes ne sont pas annotées, étendue à l'espace entier des variables d'intégration, où les fonctions à intégrer resteront définies. D'une manière générale on a supposé que

$$\int\int F(x, y)\, dx\, dy = 1,$$

et l'on aura aussi $\int p(x)\, dx = 1$. Une relation correspondante aura lieu pour $q(y)$.

La fonction bien connue de *Bravais*

$$F(x, y) = \frac{\sqrt{\delta}}{\pi} \exp - (\alpha x^2 - 2\beta xy + \gamma y^2)$$

avec $\delta = \alpha\gamma - \beta^2 > 0$ représente une approximation première de la corrélation observée dans la nature, de la même manière que la cloche gaussienne, ou disons plus simplement *la gaussienne*, représente l'approximation quasi-universelle de la distribution des observations d'une constante naturelle. En posant

$$\alpha = \frac{1}{2\sigma_1^2}, \quad \beta = \frac{r}{2\sigma_1\sigma_2}, \quad \gamma = \frac{1}{2\sigma_2^2},$$

on exprime la propriété fondamentale de la forme quadratique d'être définie et positive en disant que $-1 < r < +1$, r étant le fameux «coefficient de corrélation» et l'on obtient

$$F(x, y) = \frac{\sqrt{1 - r^2}}{2\pi\sigma_1\sigma_2} \exp - \frac{1}{2}\left\{\left(\frac{x}{\sigma_1}\right)^2 - 2r\frac{x}{\sigma_1}\frac{y}{\sigma_2} + \left(\frac{y}{\sigma_2}\right)^2\right\},$$

ce qui exprime que l'on ne restreint pas la généralité d'une manière essentielle en prenant $\sigma_1 = \sigma_2 = \sigma$, c'est-à-dire en étudiant le cas symétrique. Cela équivaut à des transformations homothétiques. En posant

$$u = \frac{\sigma_1 \sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \left(\frac{x}{\sigma_1} + \frac{y}{\sigma_2} \right), \; v = \frac{\sigma_1 \sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \left(\frac{x}{\sigma_1} - \frac{y}{\sigma_2} \right), \; \left| \frac{d(x, y)}{d(u, v)} \right| = \frac{\sigma_1^2 + \sigma_2^2}{2 \sigma_1 \sigma_2},$$

on trouve cependant que deux variables statistiques, corrélées par la fonction de *Bravais*, peuvent toujours être exprimées comme des combinaisons linéaires de deux variables *indépendantes*, le coefficient de uv s'annulant dans la nouvelle forme qui, par conséquent, se décompose en deux carrés. On voit immédiatement que dans le cas de la fonction de *Bravais* les «distributions marginales» $p(x)$ et $q(y)$ ainsi que les fonctions génératrices $\alpha(x, y)$ et $\beta(y, x)$, seront des gaussiennes très simples, p. ex.

$$p(x) = \frac{1}{\sqrt{\pi}} \sqrt{\frac{\delta}{\gamma}} \, e^{-\frac{\delta}{\gamma} x^2} = \frac{\sqrt{1 - r^2}}{\sigma_1 \sqrt{2\pi}} \exp - \frac{1 - r^2}{2 \sigma_1^2} x^2,$$

ce que nous pouvons exprimer de façon assez commode par

$$\xi = 0 \left(\text{disp.} \; \frac{\sigma_1}{\sqrt{1 - r^2}} \right),$$

et en outre

$$\alpha(x, y) = \frac{\sqrt{\gamma}}{\sqrt{\pi}} \, e^{-\gamma \left(y - \frac{\beta}{\gamma} x \right)^2} = \frac{1}{\sigma_2 \sqrt{2\pi}} \exp - \frac{1}{2 \sigma_2^2} \left(y - \frac{\sigma_2}{\sigma_1} r x \right)^2,$$

ce que nous désignons par la notation

$$\eta = r \frac{\sigma_2}{\sigma_1} \xi \; (\text{disp.} \; \sigma_2).$$

Ces généralités, qui montrent que les génératrices dans notre cas sont des gaussiennes avec une dispersion *constante*, justifient notre jugement que la fonction de *Bravais* est un instrument trop simple pour des recherches un peu au-dessus des pures estimations. Dans le même sens où l'on peut dire que *la dispersion* est une mesure assez grossière d'une distribution, ne donnant qu'une première estimation de la condensation, la corrélation de *Bravais* ne donne qu'une image assez vague de la réalité. Ce qui ne nous empêchera pas d'ailleurs d'étudier avec un grand intérêt les indications que peuvent donner, d'un point de vue philo-

sophique, les valeurs des constantes d'une fonction de *Bravais*, adoptée comme première approximation des observations sur deux variables ξ et η [1]).

III. La fonction de corrélation basée sur la distribution de *Poisson*

Lorsqu'il s'agit de comparer le nombre calculé de sinistres, disons η, et le nombre observé, ξ, on sait que les recherches de *Jacques Bernoulli* et d'autres ont justifié l'hypothèse que la dispersion des valeurs observées autour de la vraie valeur augmente dans la proportion de la racine carrée de la valeur calculée. Il ne peut donc pas être question de représenter la fonction de corrélation de ξ et de η par une fonction de *Bravais* où, comme nous l'avons constaté, cette dispersion resterait constante. Seulement dans un voisinage assez petit de l'origine, cette approximation grossière rendrait des services, et c'est précisément ce qui n'est pas intéressant du point de vue de la théorie du risque mathématique. Dans l'étude de cette théorie il faut donc abandonner la fonction de *Bravais*.

Or si l'on étudie exclusivement le jeu instantané du hasard autour du nombre calculé de sinistres, le schème simple de *Bernoulli* ou bien, si le taux d'accident n'est pas grand, celui dit de *Poisson*, peut être admis comme point de départ. La dernière distribution, facilement maniable, est exprimée par la fonction

$$P(x) = \frac{y^x}{x!} e^{-y}, \quad \sum_{x=0}^{\infty} P(x) = 1,$$

où y est le nombre de sinistres à craindre, x le nombre observé. Les propriétés simples de cette fonction sont bien connues; on a par exemple

$$\sum_{x=0}^{\infty} x(x-1) \ldots (x-n+1) P(x) = y^n,$$

et, en posant $M_n = \sum_{x=0}^{\infty} x^n P(x)$,

$$M_1 = y, \; M_2 = y^2 + y, \; M_3 = y^3 + 3y^2 + y, \ldots,$$

$$\sum_{x=0}^{\infty} (x-y)^n P(x) = y.$$

[1]) Voir Der Begriff der statistischen Funktion, Svenska Aktuarief. Tidskr. 1919, 1; Bemerkungen zur Theorie der statistischen Funktion, Ibidem, 1919, 204; Connaissance et statistique, Dialectica, Neuchâtel, 1949, 153.

On trouve donc spécialement pour la dispersion quadratique

$$\left[\sum_{x=0}^{\infty}(x-y)^2\frac{y^x}{x!}e^{-y}\right]^{\frac{1}{2}}=\sqrt{y},$$

ce qui exprime la propriété citée plus haut.

Il est évident que la fonction $P(x)$ doit jouer ici le rôle d'une fonction génératrice $\beta(y,x)$ et qu'il suffira de connaître, en outre, la fonction $q(y)$, pour pouvoir construire la fonction de corrélation $F(x,y)$. En premier lieu, admettons pour $q(y)$ une simple exponentielle décroissante, ce qui peut aussi inclure le cas d'une fréquence quasi-constante selon l'axe des y. Posons donc

$$q(y)\,dy=\varepsilon\,e^{-\varepsilon y}\,dy=e^{-\varepsilon y}\,d(\varepsilon y).$$

La constante est donnée par la condition $\int q(y)\,dy=1$. On peut simplement garder le caractère discontinu de $P(x)$ si l'on pose

$$\beta(y,x)\,dx=d\sum_{0}^{x}\frac{y^x}{x!}e^{-y}$$

et si l'on convient par exemple de construire l'échelle en plaçant les sauts dans les demi-entiers ou bien dans les entiers eux-mêmes, en admettant la continuité de l'un seul des deux côtés, p. ex. d'en haut. Au moins préalablement, ce sera un procédé plus avantageux que l'interpolation analytique à l'aide de la formule, valable pour des x positifs,

$$\beta(y,x)\,dx=\frac{y^x}{\Gamma(x+1)}e^{-y}\,dx.$$

On a donc, en écrivant de manière simplifiée, mais facile à comprendre,

$$F(x,y)\,dx\,dy=e^{-\varepsilon y}\,d(\varepsilon y)\,\frac{y^x}{x!}e^{-y}.$$

Pour obtenir $p(x)$, il faut intégrer par rapport à y et on aura

$$p(x)=\varepsilon:(1+\varepsilon)^{x+1}.$$

L'autre fonction génératrice $\alpha(x,y)$ s'obtient en divisant $F(x,y)$ par $p(x)$, ce qui donne une autre décomposition de la différentielle fondamentale

$$F(x,y)\,dx\,dy=\frac{\varepsilon}{(1+\varepsilon)^{x+1}}\,\frac{[(1+\varepsilon)y]^x}{\Gamma(x+1)}\,e^{-(1+\varepsilon)y}\,d[(1+\varepsilon)y],$$

et la génératrice devient

$$\alpha(x, y) = \frac{[(1+\varepsilon)y]^x}{\Gamma(x+1)} e^{-(1+\varepsilon)y}(1+\varepsilon).$$

On trouve pour les moments de cette distribution

$$M_n = \int_0^\infty y^n \alpha(x, y)\, dy = (x+1)(x+2)\ldots(x+n) : (1+\varepsilon)^n.$$

La moyenne devient $M_1 = (x+1) : (1+\varepsilon)$ et le carré de la dispersion se calcule ainsi:

$$\int_0^\infty (y - M_1)^2 \alpha(x, y)\, dy = (x+1)(x+2) : (1+\varepsilon)^2 - M_1^2 = (x+1) : (1+\varepsilon)^2,$$

d'où l'on conclut que la courbe donnant les dispersions quadratiques devient

$$y = \frac{x+1 \pm \sqrt{x+1}}{1+\varepsilon}.$$

La courbe correspondante dans la direction des x est, d'autre part,

$$x = y \pm \sqrt{y}.$$

Il s'agit de deux paraboles dont on trouve la construction dans la figure 1 et dont les équations sont

$$[y(1+\varepsilon) - (x+1)]^2 - (x+1) = 0, \quad (x-y)^2 - y = 0.$$

Le résultat que nous avons trouvé nous semble assez remarquable: la distribution $\alpha(x, y)$ garde la qualité caractéristique que nous avons admise pour $\beta(y, x)$, à savoir que la dispersion quadratique croît comme la racine carrée. Mais la «droite de régression», c'est-à-dire le lieu des M_1, devient, tout comme dans le cas simpliste de la fonction de *Bravais*, plus ou moins inclinée, suivant que varie le nombre ε. Pour $\varepsilon = 1$, on trouve la construction dans la figure.

Ajoutons quelques curiosités. Soit $q(y)$ une fonction arbitraire, pour laquelle $q(y) = q(-y)$ et $\int q(y)\, dy = 1$, d'où il s'ensuit que $\int_0^\infty q(y)\, dy = \frac{1}{2}$. La différentielle de corrélation sera pour $x > 0$, $y > 0$

$$F(x, y)\, dx\, dy = q(y)\, dy \frac{y^x}{x!} e^{-y}.$$

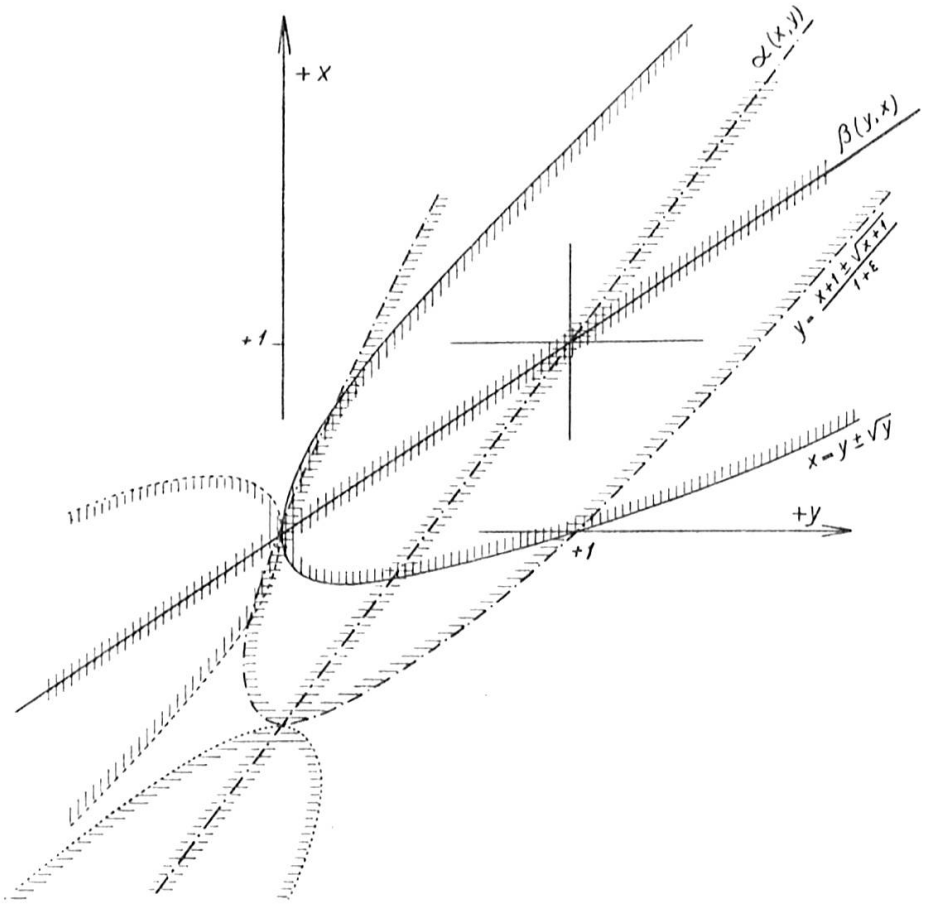

Figure 1. Régressions et dispersions dans le cas d'une distribution de Poisson et d'une fréquence exponentielle en y. (On a posé $\varepsilon = 1$)

Dans le troisième quadrant du plan, $x < 0$, $y < 0$, elle sera donnée par réflexion linéaire à travers l'origine, tandis qu'elle sera nulle dans les deuxième et quatrième quadrants. Pour un y quelconque, on aura, en sommant par rapport à x seulement de 0 à ∞,

$$\int F(x,y)\,dx\,dy = q(y)\,dy\,,\ \int y\,F\,dx\,dy = y\,q(y)\,dy\,,$$

$$\int y^2 F\,dx\,dy = y^2 q(y)\,dy\,,\ \int x\,F\,dx\,dy = y\,q(y)\,dy\,,$$

$$\int x^2 F\,dx\,dy = (y^2 + y)\,q(y)\,dy\,,\ \int xy\,F\,dx\,dy = y^2 q(y)\,dy\,.$$

On trouvera les moments en sommant ces différentielles par rapport à y de $-\infty$ à $+\infty$. Mais puisque l'on a pris $\int y\,q(y)\,dy = 0$, on aura tout simplement

$$M_{00} = 1\,,\ M_{10} = M_{01} = 0\,,\ M_{20} = M_{11} = M_{02} = \int y^2 q(y)\,dy\,.$$

Par conséquent $r^2 = M_{11}^2 : M_{20} M_{02} = 1$, et on voit que dans ce cas, un peu singulier mais d'une certaine généralité, le coefficient de corrélation est égal à 1 sans que l'image de la fonction statistique se réduise à une couche à une seule dimension (une droite pesante).

D'autre part, on peut se demander quelle valeur prendra le coefficient r dans le cas asymétrique que nous venons de traîter. Formons donc ce coefficient dans deux cas particuliers: (1) $q(y)$ étant $= \frac{1}{n}$ dans le cas des n entiers $1, 2, \ldots, n$ et nul ailleurs, (2) $q(y)$ étant $= \varepsilon \exp - \varepsilon y$. On trouve dans le premier cas $r^2 = (n-1):(n+5)$, dans le second cas $r^2 = 1 : (1+\varepsilon)$. Le nombre r tendra vers 1 lorsque $n \to \infty$ ou $\varepsilon \to 0$.

IV. La distribution de *Pareto*

Il est bien connu que la distribution des revenus s'exprime régulièrement par la loi de *Pareto*, la probabilité pour que le revenu z d'un citoyen soit au moins égal à z étant donnée par

$$\left(\frac{z}{z_0}\right)^{-k},$$

où z_0 représente un revenu minimum (choisi presque arbitrairement) et k est *l'indice d'égalité*, variant normalement entre 1,3 et 2. Le nivellement des revenus causé par la dernière guerre mondiale a amené des valeurs de k inconnues autrefois, croissant de la valeur quasi-normale 1,5 jusqu'à 2,4 (pour la Suède en 1947) [1]). Parmi les essais de donner une interpretation adéquate de la loi en question citons la définition suivante. En désignant par le *degré d'impossibilité* d'un événement le nombre $\ln \frac{1}{p}$, où p signifie la probabilité, et par la *valeur morale* ou *subjective* d'un revenu z le nombre $\ln z$, introduit par *Daniel Bernoulli* en 1731 (1738), la loi de *Pareto* s'exprime en disant que *le degré d'impossibilité pour que la valeur morale du revenu d'un citoyen atteigne ou surpasse une certaine limite est proportionnel à cette limite.* Le facteur de

---

[1]) Voir p. ex. *La loi de Pareto et la réassurance*, Skand. Akt. Tidskr. 1925, p. 65, *Inkomstinflation och inkomstnivellering i Sverige*, Skand. Bankens Kvartalsskr. 1949, Avril, *Den socialekonomiska generalplanen*, Gjallarhornet 1945 (nr 19). Je vais démontrer ailleurs que sous certaines hypothèses la somme d'un nombre de variables statistiques aux distributions parétoënnes aura une distribution de la même espèce, d'une manière asymptotique.

proportionnalité, c'est l'indice d'égalité k. Evidemment, si la distribution des revenus privés d'un pays est représentée par une loi de *Pareto*, on doit espérer la même fonction de distribution pour les capitaux assurés, en négligeant en première approximation la possibilité d'une sélection spéciale.

Tâchons donc de choisir pour $q(y)\,dy$ une fonction de *Pareto*, en posant

$$q(y)\,dy = d\left\{-\left(\frac{y}{y_0}\right)^{-k}\right\} = k\left(\frac{y}{y_0}\right)^{-k-1} d\,\frac{y}{y_0}.$$

Nous prenons $y_0 < 1$ (arbitrairement petit mais > 0) et $k > 1$, ce qui donne $\int q(y)\,dy = 1$. On obtient

$$p(x) = \int F(x,y)\,dy = \frac{k\,y_0^k}{\Gamma(x+1)} \int_0^\infty y^x\,e^{-y}\,y^{-k-1}\,dy = k\,y_0^k \frac{\Gamma(x-k)}{\Gamma(x+1)},$$

donc

$$\alpha(x,y)\,dy = \frac{1}{\Gamma(x-k)}\,y^{x-k-1}\,e^{-y}\,dy.$$

Pour les moments de $\alpha(x,y)$ on obtient donc

$$M_n = \int \alpha(x,y)\,y^n\,dy = \frac{\Gamma(x-k+n)}{\Gamma(x-k)},$$

$$M_1 = x-k+1,\ \int (y-M_1)^2\,\alpha(x,y)\,dy = M_2 - M_1^2 = x-k+1.$$

Dans ce cas, la droite de régression correspondante devient

$$y = x-k+1,$$

et la courbe donnant les dispersions quadratiques devient la parabole

$$y = x-k+1 \pm \sqrt{x-k+1}$$

L'inclinaison de la droite de régression reste constante, mais la position de la parabole avance dans la direction des x positifs lorsque l'indice k augmente. Dans ce cas aussi, on peut cependant constater que la dispersion garde son caractère général d'augmenter à la mesure de la racine carrée.

V. Les phases d'une assurance

On a étudié [1]) les phases successives qui peuvent être discernées dans une assurance sur la vie, embrassant un nombre de personnes. La transition d'une phase à une autre est, en général, accompagnée d'une augmentation ou d'une diminution brusque de la réserve mathématique de l'assurance, et c'est surtout ce changement discontinu de la réserve qui caractérise l'assurance. Il est la conséquence d'un changement analogue dans le revenu de l'assuré, qui justement doit être compensé par les stipulations de la police d'assurance.

Dans la manière de traiter la statistique et la comptabilité en matière d'assurance, on tient souvent séparées des phases des contrats: p. ex. les rentes viagères différées avant et après la date de pension, les rentes de survie avant et après le décès du mari, etc. — phases qui dans certains cas seulement correspondent avec les phases théoriques. Dans la théorie du risque mathématique, il faut séparer en outre certaines des phases en périodes différentes, caractérisées par le signe du montant de risque, c'est-à-dire selon les cas où le sinistre entraîne une augmentation ou une diminution de la réserve. On peut classifier les contrats d'après le signe du changement calculé de la réserve, et on obtient de cette manière deux différents domaines, dont chacun est conforme au cas simple que nous venons d'étudier, en parlant seulement des cas de décès, calculés et observés. Dans ce cas plus général, le résultat a un caractère continu. Nous avons vu que les deux cas, celui des variables continues et celui des variables discontinues, peuvent bien être assimilés l'un à l'autre au moyen de l'intégrale de *Stieltjes*.

La pratique actuarielle introduit souvent une séparation plus détaillée des parties de l'assurance, en considérant p. ex. une assurance de pension comme l'ensemble de trois contrats, à savoir: 1° la pension de retraite, 2° la pension d'invalidité, 3° la protection de famille. Certes, on peut trouver peu naturel que la même personne assurée figure comme membre de plusieurs collectives différentes. Mais l'existence de tels cas ne doit pas en général fausser l'image de l'ensemble des risques courus. D'autre part, l'erreur, s'il y en a, doit exagérer le nombre de déviations considérables. On peut remarquer que l'exagération a une signification objective, représentant les conditions d'une affaire où

[1]) Voir *The General Life Assurance*, Skand. Akt. Tidskr. 1941, p. 137, *La riserva prospettiva dell' assicurazione generale sulla vita e la misura del contenuto assicurativo del contratto*, Giorn. dell' Ist. Ital. degli Attuari, 1941, p. 103.

toutes les parties discernées seraient réellement portées par différents assureurs, ce qui peut bien s'établir.

VI. Résultats calculés et observés

La comparaison des résultats espérés et des observations suppose que l'on a calculé ces premiers pour chaque cas observé, donc lorsqu'il s'agit d'une assurance collective pour chaque groupe et pour chaque exercice considéré. Quand l'assurance est du type le plus simple, tel qu'on la pratique souvent aux Etats-Unis, il faut seulement calculer le nombre probable de décès, et on peut directement comparer ce nombre au nombre enregistré. Si l'on veut que la somme des nombres calculés soit égale à celle des décès observés, il suffit d'établir simplement une proportion. Mais dans la plupart des contrats d'assurance collective, il s'agit de formes mixtes plus compliquées. Remarquons seulement que les limites des différentes phases qu'il faut séparer dépendent du changement de la mortalité, effectué par la réduction ordinaire dont nous venons de parler.

Envisageons le cas d'une assurance mixte, p. ex. une assurance de pension complète. D'une manière générale, l'économie d'une assurance mixte peut être interprétée par un système de relations thieléennes, disons pour une phase P_m par la relation

$$\frac{dV_m(t)}{dt} = \delta\, V_m(t) + p(t) - r(t) - \mu_{ma}\, dV_{ma} - \mu_{mb}\, dV_{mb} - \ldots,$$

où μ_{mc} représentent les intensités de transition de la phase P_m aux phases suivantes P_c qui sont possibles, $dV_{mc} = V_c - V_m$ l'augmentation de la réserve lors du changement de P_m à P_c, V_m la réserve mathématique dans la phase P_m, $p(t)$ la prime par unité de temps, $r(t)$ la rente à payer par unité de temps — $p(t)$ s'annulant en général dans ce cas —, δ l'intensité d'intérêt (que nous avons renoncé à considérer comme fonction du temps). Pour intégrer définitivement un tel système il faut introduire un ensemble de «discumulateurs», comprenant les intensités d'intérêt δ et celles des transitions μ_{mc}. Mais on peut approximativement intégrer la relation pour $V_m(t)$ sur un petit intervalle, disons un an, et l'on obtient alors

$$\sum_c \int_0^1 \mu_{mc}\, dV_{mc}\, dt = V_m(t)\left(1 + \frac{\delta}{2}\right) - V_m(t+1)\left(1 - \frac{\delta}{2}\right) + p\left(t + \frac{1}{2}\right) - r\left(t + \frac{1}{2}\right).$$

Pour la phase initiale P_0 des actifs, on a $r = 0$ et il faut regarder les transitions

01 = l'arrivée de l'âge de retraite,
02 = cessation des primes,
03 = invalidité,
04 = décès de l'assuré actif.

Ecartant les deux premières qui n'ont pas d'intérêt ici, on peut écrire avec $\mu_{04} = \mu_{x+t}$ et $\mu_{03} = \nu_{x+t}$ (x représentant pour l'instant l'âge)

$$\int_0^1 \mu_{x+t}\, dV_{04}\, dt + \int_0^1 \nu_{x+t}\, dV_{03}\, dt = V_0(t)\left(1 + \frac{\delta}{2}\right) - V_0(t+1)\left(1 - \frac{\delta}{2}\right) + p.$$

Il est facile de calculer, par un simple enregistrement,

$$\int_0^1 \nu_{x+t}\, dV_{03}\, dt = \sum (p + j)\, \nu(x + \tfrac{1}{2})\, f(x + \tfrac{1}{2})\, z),$$

où p est la prime, j la rente d'invalidité assurée, $f(x + \frac{1}{2})\, z)$ la valeur de la rente viagère, payable en cas d'activité, de l'âge $x + \frac{1}{2}$ jusqu'à l'âge z de retraite. Donc, dans ce cas on obtient par une simple soustraction, en calculant toujours au même temps les deux réserves $V(t)$ et $V(t+1)$ (pour la même date, bien entendu), l'expression importante

$$\int_0^1 \mu_{x+t}\, dV_{04}\, dt,$$

c'est-à-dire la croissance calculée de la réserve pour les décès probables. Ici on peut donc séparer les risques. En général c'est plus difficile. Pour les questions de la théorie du risque mathématique, il n'est pas nécessaire de séparer les risques différents. Mais il faut le faire si l'on veut approfondir la recherche.

Dans la pratique de la «Bâloise», on a introduit un calcul très détaillé des changements probables de réserves, séparé suivant les composantes. Il me semble que le système est d'une perfection presque parfaite. Si l'on veut contrebalancer les résultats particuliers d'un même groupe pour une certaine année, il me semble qu'il faut toujours tenir séparés les résultats des différents sexes et ceux qui se rapportent aux pensionnaires retraités et aux membres de famille des décédés. Bien que l'arrangement ne soit pas tout à fait rationnel, j'ai adopté la division suivante:

Assurés du sexe masculin:

1° Assurances de pension de retraite (42), d'invalidité (50), de veuvage (41), d'orphelins (60), assurances de capitaux et d'annuités au décès (11, 04) et en cas de vie (20) ou d'invalidité (30), rentes d'invalidité actuelles (80).

2° Pensions de retraite actuelles (72).

Assurés du sexe féminin:

3° Assurances de pension de retraite (47), d'invalidité (55), d'orphelins (60), assurances de capitaux sur le décès (16) et sur le cas de vie (25) ou d'invalidité (35), rentes d'invalidité actuelles (85).

4° Pensions de retraite actuelles (77).

Veuves et orphelins:

5° Rentes de survie actuelles (71, 90).

Les numéros figurant dans ce tableau sont les chiffres de code adoptés par la «Bâloise» elle-même. En soldant les résultats qui se rapportent au même groupe et au même exercice, ce qui est fait dans une des recherches suivantes, j'ai négligé le fait que les assurances de protection familiale pour les retraités restent dans le même détail que celles qui se rapportent aux actifs. La petite erreur qui en résulte n'a aucune importance. Sauf dans les cas particuliers 2°, 4° et 5°, on fera donc l'addition algébrique de tous les résultats d'un même groupe, en séparant seulement les deux sexes.

VII. La théorie perfectionnée du risque mathématique et la réassurance des grands écarts

En traitant le schéma théorique, nous nous sommes limités à une étude du cas particulier représenté, en ce qui concerne le nombre possible de sinistres, par la distribution de *Poisson*. Or il faut avouer que c'est là une spécialisation bien restreinte. Elle ne regarde que des pertes d'une même grandeur, tandis qu'il faudra toujours prendre égard à la distribution réelle des capitaux assurés; elle néglige la possibilité d'une autre fréquence des cas multiples que celle supposée par l'hypo-

thèse d'indépendance de *Bernoulli*. Mais elle nous a semblé suffisante lorsqu'il ne s'agit que d'esquisser le fond, sur lequel se dessinerait la statistique à présenter. Pour le moment nous resterons dans ce schématisme imparfait, sans aborder les questions plus profondes qui ont été traitées avec tant de succès par MM. *Filip Lundberg*, *Harald Cramér* et d'autres.

Le problème pratique qui nous semble particulièrement attrayant est celui d'une construction satisfaisante de la «*réassurance des grands écarts*». Sans aucun doute, elle représenterait, si on pouvait la réaliser, la forme la plus parfaite de réassurance. Supposons cependant que l'on ait trouvé la probabilité $\beta(y, x)\,dx$ pour que le payement soit entre x et $x + dx$ sous la forme d'une gaussienne $f_s(v)\,dv$, où $v = x - y$ est l'écart du payement réel x sur le payement calculé y,

$$f_s(v)\,dv = \frac{dv}{s\sqrt{2\pi}} \exp - \frac{v^2}{2s^2}.$$

On peut alors calculer la prime r d'une réassurance de tout écart $> ls = U$, les écarts plus petits que U étant couverts par un fonds d'ajustement de cette grandeur, appartenant à l'assureur primaire lui-même. On obtient

$$\frac{r}{s} = f_1(l) - l(1 - Q(l))$$

avec

$$Q(l) = \int_{-\infty}^{l} f_1(z)\,dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{l} e^{-z^2/2}\,dz.$$

Supposons que l'assureur primaire puisse, pendant quelques années, disposer de 40 % de la prime de risque (brute) comme profit, dont 20 % sont mis à côté comme fonds d'ajustement U et 20 % sont employés pour couvrir la prime de réassurance et pour bénéfice aux assurés. Posons en outre $s = 25$ % de la prime de risque. Si la somme des primes de risque rentrant chaque année est prise pour unité, le développement de l'affaire se calcule comme il suit.

An	Fonds d'ajustement U	En fraction de la dispersion		Calcul de la prime de réassurance des grands écarts			Fonds de bénéfices avant | après le payement de la prime r	
		$l = U/s$	$f_1(l)$	$Q(l)$	r/s	r		
1	0,2	0,8	0,28969	0,78814	0,12020	0,03000	0,200	0,170
2	0,4	1,6	0,11092	0,94520	0,02324	0,00581	0,370	0,364
3	0,6	2,4	0,02239	0,99180	0,00271	0,00068	0,564	0,563
4	0,8	3,2	0,00239	0,99931	0,00017	0,00004	0,763	0,763
5	1,0	4,0	0,00013	0,99997	0,00001	0,00000	0,963	0,963

Par ce tabelau il est clairement visible qu'un réassureur ne peut éprouver qu'un intérêt assez médiocre pour une telle réassurance, si elle ne peut être autrement organisée. En effet, en échange d'une responsabilité assez sévère au commencement de l'activité, le réassureur n'obtient qu'une participation très limitée, grâce à la qualité de l'assurance de se dissiper automatiquement dans peu d'années. Elle ne serait donc guère praticable que si l'on réussissait à y introduire des clauses plus favorables pour le réassureur. On peut réfléchir à une limitation inférieure de la responsabilité.

VIII. Présentation de la statistique

Lorsqu'on veut comparer la réalité, représentée par une statistique de résultats calculés et observés à un schéma théorique — plus ou moins perfectionné — tel que nous avons examiné, il semble tout d'abord indispensable de faire quelques transformations des données statistiques. Désignons pour le moment les quantités réelles par des astérisques, donc ξ^*, η^*, x^*, y^*, F^*, etc., on a comme point de départ une table de corrélation réelle, où la masse totale est représentée par un nombre N de cas traités ou de points, chacun étant mesuré par $\frac{1}{N}$ de la masse totale. La base de calcul étant toujours plus ou moins arbitraire, la somme calculée de changements de réserve (ou le nombre total calculé de sinistres) ne correspondra pas aux sommes observées. Pour adapter la réalité au schéma théorique, il faudrait donc en premier lieu réduire les données afin de rétablir cette correspondance. Or, comme nous avons déjà remarqué plus haut, la question d'une telle réduction, très simple dans le cas d'une assurance de décès collective à une somme

constante, peut être plus délicate dans les cas plus généraux. On peut faire *une* réduction pour les cas aux changements calculés *positifs* et une autre (peut-être en sens opposé) pour les cas aux changements *négatifs*. On peut aussi éviter cette réduction et chercher à faire une altération correspondante dans le schéma théorique, qui serait alors plus compliqué. Pour une comparaison préalable il peut suffire de se figurer un changement proportionnel des y^*, c'est-à-dire d'introduire une nouvelle unité dans la direction des y^*, de façon que la droite de régression soit transformée dans la diagonale $y^* = x^*$ du premier quadrant de coordonnées. Sans faire expressément ce changement — ce qui altérerait les statistiques primaires d'une manière non désirable — on peut constater que la recherche de la variation de la dispersion en x^* avec la grandeur de y^* (ou *vice versa*) ne doit pas être dissimulée si l'on omet la réduction. Dans une statistique plus étendue, on pourrait pousser plus loin cet examen: dans une statistique restreinte il faut se borner à une estimation grossière.

Nous préférons donc, en présentant la statistique dans quelques tables et graphiques qui sont aisément compréhensibles par eux-mêmes, nous borner à une recherche préliminaire et assez grossière de la variation de la dispersion. Cependant, nous avons poursuivi nos recherches théoriques assez loin, à ce qu'il nous semble, pour être persuadés du résultat probable suivant. Si le schéma basé sur la distribution de *Poisson* est approximativement valable, la dispersion dans toutes les deux directions des x^* et des y^* varie proportionnellement à la racine carrée de la coordonnée correspondante.

La fonction de corrélation obtenue, $F^*(x^*, y^*)$ — nous omettons dans les tables et graphiques les astérisques — présente une couche discontinue sur l'axe des y^*, qui figure comme une ligne singulière. C'est ce qui n'est pas conforme aux conditions dans un schéma stochastique ordinaire. L'attraction exercée par cette ligne singulière doit dépendre d'un minimum de changement de réserve dans un groupe.

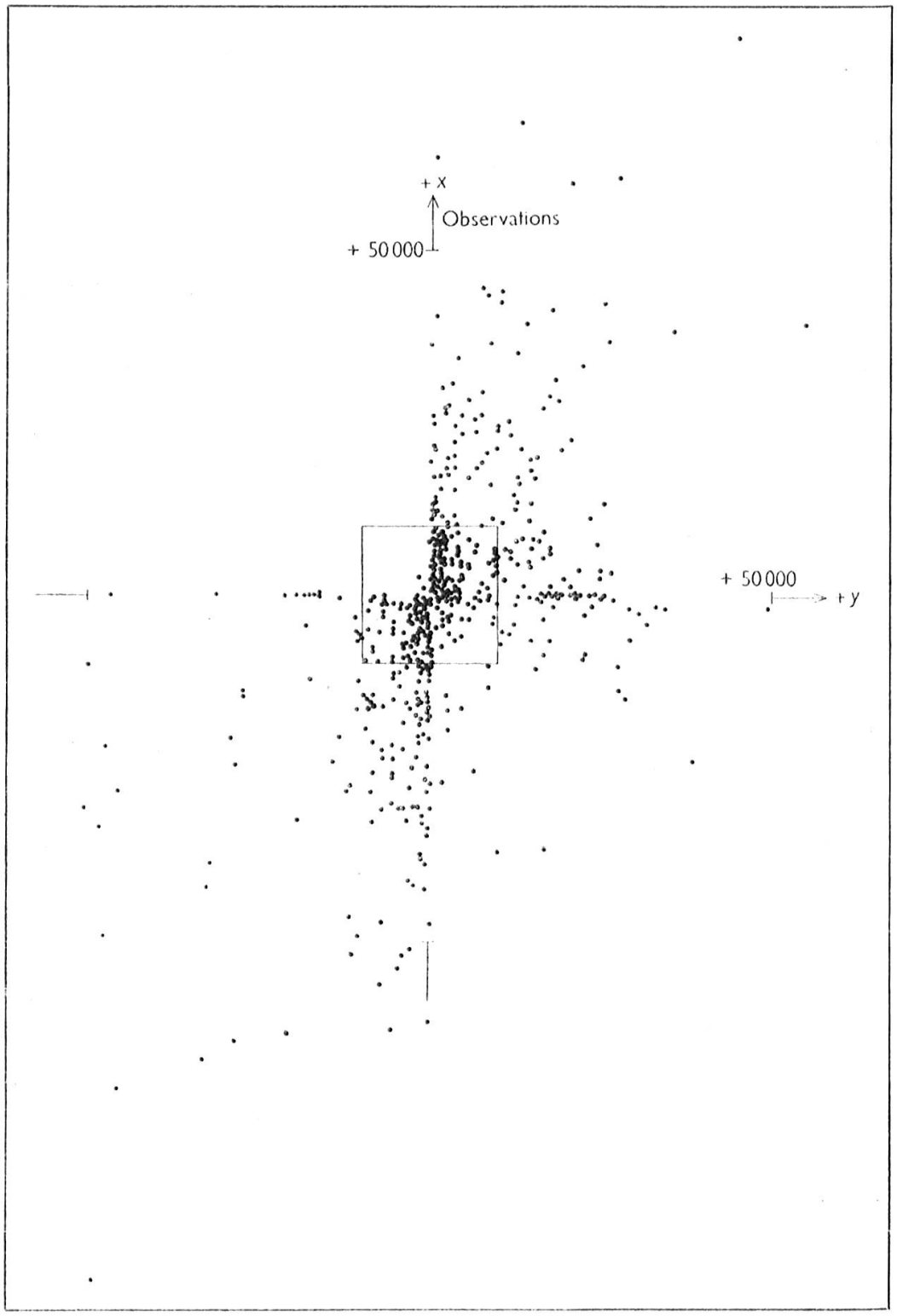

Figure 2

Image macroscopique de la corrélation entre observations x et calcul y

Le carré central est donné en plus grande échelle dans la figure 3

L'axe des y ($x = 0$) est représenté dans la figure 4

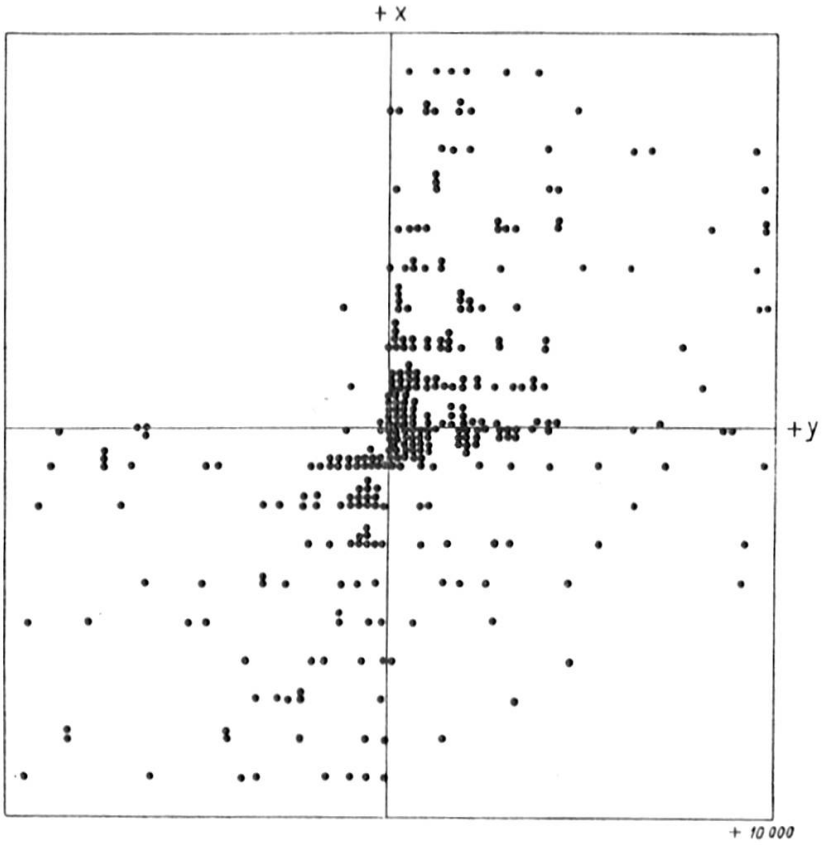

Figure 3. Image aggrandie du centre de la figure 2
— 10 000 < x < + 10 000, — 10 000 < y < + 10 000

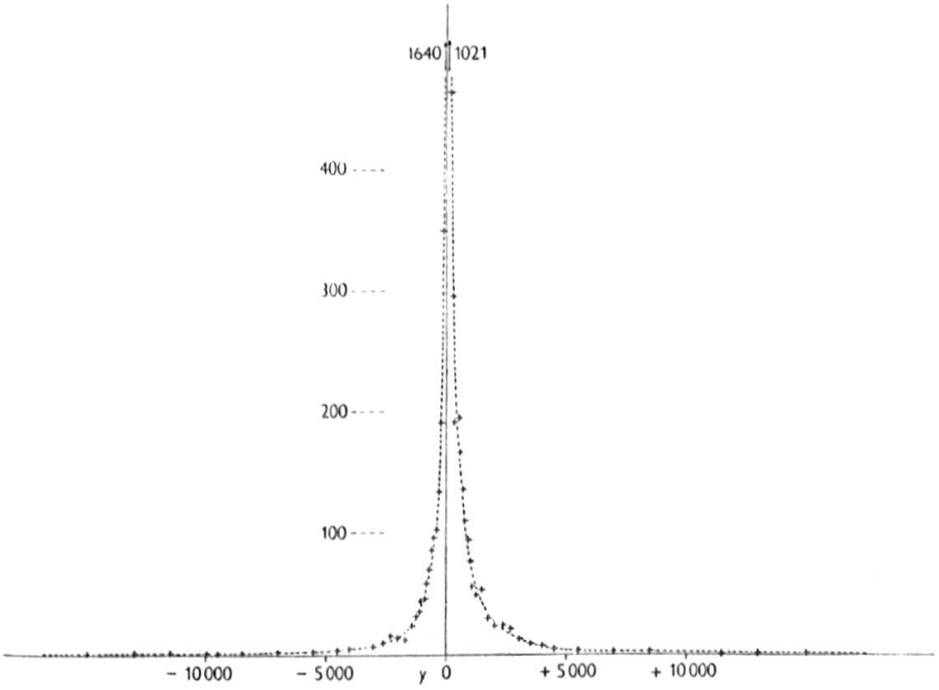

Figure 4. Distribution de masse stochastique sur la droite d'accroissement observé de réserves disparaissant, x = 0

La hauteur de la courbe représente le nombre de cas observés, correspondant aux points dans les figures 2 et 3

Table de corrélation
entre changements de la réserve mathématique observés (x) et calculés (y)
{l'axe des y (changements s'annulant) exclu}

Changements observés x	Changements calculés en milliers, y											
	−50	−45	−40	−35	−31	−27	−23	−20	−17	−15	−13	−11
+ 43 400												
+ 24 700												
+ 14 400												
+ 7 800												
+ 4 900												
+ 2 450												
+ 990												
— 20	.	.	.	.	.	.	.	.	.	.	1	.
— 1 040	.	.	.	.	.	.	.	.	.	.	.	.
— 2 340	.	.	.	.	.	.	.	.	.	.	.	1
— 4 900	.	.	.	.	.	.	.	.	1	.	.	3
— 7 900	.	.	.	.	.	.	.	.	.	.	.	1
— 14 000	1	.	.	.	.	2	.	.	1	.	.	2
— 24 500	.	2	.	.	1	1	.	.	.	1	2	1
— 44 100	4	1	.	1	2	1	.	2	.	.	.	3
Total	5	3	0	1	3	4	0	2	2	1	3	11

Changements observés x	Changements calculés en milliers, y											
	−9	−7½	−6	−4½	−3	−2	−1	0	+1	+2	+3	+4½
+ 43 400	.	.	.	.	.	.	.	1	2	1	1	1
+ 24 700	.	.	.	.	.	.	.	2	2	2	2	3
+ 14 400	.	.	.	.	.	.	.	5	8	2	7	2
+ 7 800	.	.	.	.	.	.	.	3	5	7	1	3
+ 4 900	.	.	.	.	.	.	.	4	11	3	2	5
+ 2 450	.	.	.	.	.	.	1	12	14	2	2	3
+ 990	.	.	.	.	.	.	1	10	8	5	4	2
— 20	1	.	.	.	.	1	3	28	18	19	8	4
— 1 040	1	3	1	2	.	4	10	12	3	1	1	1
— 2 340	1	1	.	.	2	6	16	6	4	.	2	.
— 4 900	1	1	1	4	3	2	6	4	2	1	2	2
— 7 900	2	.	1	3	4	4	3	3	1	.	1	.
— 14 000	7	3	4	4	3	3	14	9	.	2	1	1
— 24 500	1	3	5	1	2	3	3	2	.	1	.	.
— 44 100	1	1	3	5	4	2	7	4	.	.	.	.
Total	15	12	15	19	18	25	64	105	78	46	34	27

Changements observés x	Changements calculés en milliers, y										
	+ 6	+ 7½	+ 9	+ 11	+ 13	+ 15	+ 17	+ 20	+ 23	+ 27	+ 31
+ 43 400	.	3	1	2	3	.	2	1	1	3	.
+ 24 700	3	2	3	2	1	.	2	3	.	.	.
+ 14 400	1	3	2	1	5	6	1	.	1	1	.
+ 7 800	1	1	1	3	1	3	1	.	1	.	.
+ 4 900	1	.	5	1	1	2	2	.	2	.	.
+ 2 450	.	1	2	1	.	.	.	1	1	1	.
+ 990	.	1	.	.	.	1	1	1	.	.	.
— 20	1	.	2	.	1	2	.	.	1	1	.
— 1 040	1	1	1	.	1	.	2	.	1	.	.
— 2 340	2	.	1	.	.	.	1	.	1	1	1
— 4 900	.	.	1	2	1	.	.	.	.	.	2
— 7 900	.	.	.	1	.	1	2	.	2	.	.
— 14 000	.	1	1	.	.	1	.	.	.	3	.
— 24 500	1	.	.	.	.	.	.	.	.	.	.
— 44 100	.	.	1	.	.	.	1	.	.	.	.
Total	11	13	21	13	14	16	15	6	11	10	3

Changements observés x	Changements calculés en milliers, y					Nombres totaux	Moyennes $\bar{y}$	Dispersions quadratiques moyennes
	+ 35	+ 40	+ 45	+ 50	+ 55			
+ 43 400	1	.	1	.	1	25	+ 15 716	13 900
+ 24 700	.	.	.	.	.	27	+ 8 059	6 050
+ 14 400	.	.	.	.	.	45	+ 7 065	6 850
+ 7 800	.	.	.	.	.	31	+ 6 108	6 050
+ 4 900	.	.	.	.	.	39	+ 5 959	6 350
+ 2 450	.	.	.	.	.	41	+ 3 565	6 200
+ 990	.	.	.	.	.	34	+ 2 978	4 800
— 20	.	.	.	.	.	92	+ 2 121	4 900
— 1 040	1	.	.	.	.	47	+ 1 555	7 900
— 2 340	1	.	.	1	.	48	+ 3 185	11 400
— 4 900	.	.	.	.	.	39	+ 298	9 400
— 7 900	.	.	.	.	.	29	+ 1 369	9 050
— 14 000	.	.	.	.	.	63	+ 2 939	10 900
— 24 500	.	1	.	.	.	31	— 7 542	12 900
— 44 100	.	.	.	.	.	43	— 11 469	16 800
Total	3	1	1	1	1	634	+ 2 411	
($\bar{x}$ = — 1818)								

IX. Résultats et interprétation

Dans la figure 5 on a donné une esquisse des résultats calculés des dispersions quadratiques dans le cas où l'on n'a pas encore soldé les nombres suivant les indications du n° VI. On voit immédiatement que l'allure des courbes de dispersions ne contredit point la vue que les dispersions vont en se reserrant, comparées aux variables elles-mêmes, dans la mesure de la racine carrée. C'est le résultat qu'on a attendu, et qui ne manque pas d'intérêt.

Lorsqu'il s'agit d'interpréter les résultats numériques d'après un schéma abstrait, des ambiguités se présentent. Nous avons déjà observé que la manière dont doit se faire la réduction à l'équivalence totale des résultats observés et calculés n'est pas évidente. Il serait à priori préférable d'introduire *deux* constantes de réduction, l'une pour les sommes de risque positives, l'autre pour les sommes négatives. D'autre part, on perd ainsi le caractère univoque de la réduction. Lorsqu'il s'agit d'une première orientation, on peut négliger la différence. En désignant par M^*_{rs} le moment de la distribution observée en x^* et y^*,

$$M^*_{rs} = \int\int x^{*r}\, y^{*s}\, F(x^*, y^*)\, dx^*\, dy^*,$$

il est évident que le coefficient de régression k défini par

$$\int\int (x^* - k\, y^*)^2 F(x^*, y^*)\, dx^*\, dy^* = minimum,$$

est donné par

$$k = M^*_{11} : M^*_{02},$$

sans qu'on ait besoin de faire auparavant la réduction qui annule les premiers moments. On obtient dans notre cas

$$k = 0{,}66\,.$$

Il s'ensuit que la droite $x^* = k\, y^*$, approximation grossière de la ligne de régression, peut-être transformée en

$$x^* = y\,,$$

si l'on pose $k\, y^* = y$. En même temps, l'autre ligne de régression tombera vers l'axe des x^*. Or nous avons vu que dans nos schèmes abstraits

Figure 5. Esquisse des courbes de régression avec leurs dispersions

une inclinaison arbitraire peut être obtenue en adoptant pour $q(y)$ une exponentielle ou une paretoënne convenable. La recherche de la variation des dispersions avec la distance de l'origine peut être effectuée en étudiant directement la distribution observée.

Supposons maintenant que la dispersion de la droite de régression en question soit donnée par l'expression

$$\sigma^* = \sqrt{c\,y^*},$$

c'est-à-dire

$$c\,y^* = \int (x^* - k\,y^*)^2\,\alpha(y^*, x^*)\,dx^*.$$

Pour obtenir la meilleure valeur de c, on doit multiplier par $q(y^*)\,dy^*$ et intégrer, ce qui donne

$$c \int\int y^* F\,dx^*\,dy^* = \int\int (x^* - k\,y^*)^2 F\,dx^*\,dy^*$$

ou bien

$$c\,M^*_{01} = (M^*_{02}\,M^*_{20} - M^*_{11})^2 : M^*_{02}.$$

Dans le cas d'une distribution s'étendant aux deux côtés de l'origine, on changera le moment figurant comme facteur de c en prenant la valeur absolue de y^* dans l'intégrale. Calculons c; il se monte à

$$c = 10{,}8 \quad \text{ou} \quad \sqrt{c} = 3{,}29.$$

Dans certains raisonnements, on peut s'intéresser au multiplicateur de la racine carré de y, c'est-à-dire *après* la transformation $y = k\,y^*$. Alors on obtient le multiplicateur modifié

$$\sqrt{\frac{c}{k}} = 4{,}04.$$

Les observations rudimentaires des dispersions quadratiques dans les distributions $\alpha(x^*, y^*)$ et $\beta(y^*, x^*)$ sont représentées dans le diagramme de la figure 6. On y voit aussi les approximations paraboliques obtenues par le calcul dont nous venons de donner la description. Dans la direction des x^*, on obtient pour le coefficient $\sqrt{c_1}$ de $\sqrt{x^*}$ la valeur 3,68, si la ligne singulière est incluse, mais la valeur 2,78, si elle est exclue. On voit d'ailleurs, par la dispersion des points autour des paraboles, qu'il ne s'agit que d'une estimation grossière.

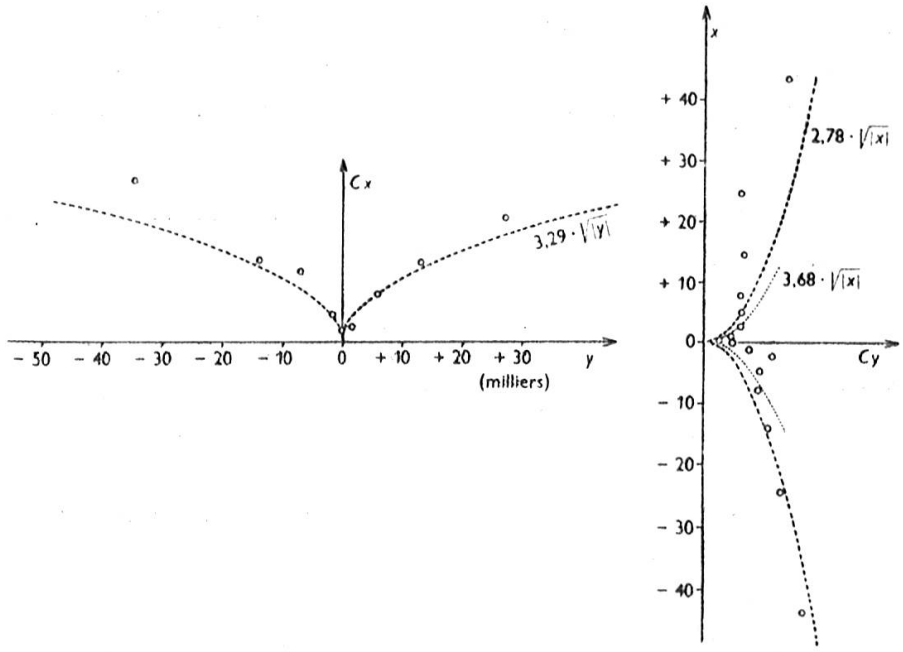

Figure 6. La variation des dispersions quadratiques

Les remarques que nous avons faites sur l'influence de la séparation des phases — plus ou moins propres — de l'assurance, nous a forcé à faire aussi un calcul supplémentaire, après avoir soldé les résultats d'après les conventions établies au n° VI. On a obtenu, en excluant ici d'une manière adéquate la ligne singulière $x = 0$, qui ne joue aucun rôle dans ce raisonnement, les valeurs suivantes.

Dispersions quadratiques dans $\beta(y, x)$ pour des valeurs différentes de y

y	Dispersion des valeurs non soldées	Dispersion des valeurs soldées
— 35,5	23,4	24,0
— 13,9	15,5	13,1
— 7	14,6	14,2
— 1,6	12,2	13,3
0	15,3	15,4
+ 1,6	9,0	10,0
+ 6	13,7	14,2
+ 13,2	16,6	13,4
+ 27	23,4	23,0

Les dispersions sont exprimées en milliers. Il est bien visible que les observations ne suffisent pas à mettre en évidence l'augmentation de la dispersion à laquelle on devrait s'attendre en soldant les résultats, ou, moins encore, de révéler une allure différente de la courbe quasi-parabolique étudiée.

Exprimons le résultat de notre investigation dans quelques points:

1º La distribution pratique contient une couche singulière sur l'axe $x = 0$, dont la masse relative dépendra probablement de l'unité choisie pour l'exercice et du *minimum* de changement causé par un seul cas d'assurance.

2º La dispersion semble varier, tout comme dans les modèles schématiques étudiés, dans la mesure de la racine carrée du changement croissant.

3º Tandis que, dans le cas le plus simple d'assurance collective sur la vie, on peut étudier directement le nombre de décès, il faut, dans le cas plus général, admettre une «unité» assez arbitraire, où la dispersion deviendra égale à un. Pour la détermination de cette unité, nous n'avons pas trouvé d'indications théoriques.

4º Le procédé qui consiste à faire l'addition algébrique des résultats qui se rapportent au même groupe, mais qui concernent différentes parties d'une même assurance, ne semble pas changer l'image générale.

5º Rien ne contredit l'hypothèse que la recherche de l'allure générale de la variation des dispersions peut être effectuée sur les droites $x =$ const., si cela s'avère plus commode pour la statistique.

(Les graphiques ont été mis au point pour l'impression avec l'aimable collaboration de M. Hans Zimmermann)

Table de la densité sur l'axe des y $x = 0$

(Aucun changement observé)

y	Nombre de cas par intervalle de cent unités	Nombre total de cas	y	Nombre de cas par intervalle de cent unités	Nombre total de cas
+ 25 000	0,053	4	— 400	102	102
+ 20 000	0,2	5	— 500	96	96
+ 17 500	0,24	6	— 600	85	85
+ 15 000	0,72	18	— 700	69	69
+ 13 000	0,47	7	— 800	58	58
+ 11 500	0,47	7	— 900	45	45
+ 10 000	1,47	22	— 1 000	44	44
+ 8 500	1,47	22	— 1 100	35	35
+ 7 000	2,0	30	— 1 200	31	31
+ 5 500	2,8	42	— 1 400	23,7	71
+ 4 500	4,6	23	— 1 700	12,7	38
+ 4 000	7,2	36	— 2 000	14,3	43
+ 3 500	8,0	40	— 2 300	14,7	44
+ 3 000	12,4	62	— 2 600	9,7	29
+ 2 600	20,7	62	— 3 000	6,4	32
+ 2 300	24,0	72	— 3 500	6,4	32
+ 2 000	24,0	72	— 4 000	4,6	23
+ 1 700	29,3	88	— 4 500	3,8	19
+ 1 400	54,3	163	— 5 500	2,0	30
+ 1 200	49	49	— 7 000	0,73	11
+ 1 100	55	55	— 8 500	0,40	6
+ 1 000	77	77	— 10 000	0,53	8
+ 900	94	94	— 11 500	0,27	4
+ 800	110	110	— 13 000	0,40	6
+ 700	137	137	— 15 000	0,20	5
+ 600	166	166	— 17 500	0,20	5
+ 500	194	194	— 20 000	0,08	2
+ 400	192	192	— 25 000	0,00	0
+ 300	296	296	— 32 500	0,01	1
+ 200	464	464	— 50 000		1
+ 100	1021	1021			
0	1640	1640	Nombre total de cas		6927
— 100	351	351			
— 200	191	191			
— 300	134	134			

Moyenne $\bar{y} = 280$

Dispersion quadratique moyenne = 2 440

Correction

concernant

l'Etude statistique du risque mathématique dans l'assurance collective sur la vie par *K.-G. Hagstroem,* publié dans le volume 51, 1951, p. 63.

Par une erreur de signe, on a malheureusement été amené à une conclusion fausse sur la page 71, en disant que dans le cas singulier considéré, on aurait le coefficient de corrélation $r = 1$ sans que l'image se réduisît à une droite pesante. On peut exprimer r à l'aide des intégrales $\int y\, q(y)\, dy$ et $\int y^2\, q(y)\, dy$ selon l'axe des y positifs, et l'on obtient comme toujours $r^2 < 1$.

Remarquons encore que les formules aux pages 65 et 66 seront mieux en conformité avec l'algèbre usuelle de la théorie de corrélation, si l'on remplace partout σ_1 et σ_2 par $\sigma_1 \sqrt{1-r^2}$ et $\sigma_2 \sqrt{1-r^2}$. On aurait alors $\xi = 0$ (disp. σ_1) et $\eta = r \dfrac{\sigma_2}{\sigma_1} \xi$ (disp. $\sigma_2 \sqrt{1-r^2}$).

K.-G. Hagstroem