FORMULATION BAYESIENNE DU PROBLEME DES VALEURS EXTREMES EN RELATION À LA RÉASSURANCE EN „EXCEDENT DE SINISTRES"

DARIO FÜRST
Rome, Italie

En matière de détermination de la prime de réassurance en „excédent de sinistres", nous nous trouvons devant le problème de la distribution des valeurs extrêmes, duquel nous avons déjà une littérature vaste et bien connue, plus particulièrement en ce qui concerne l'allure asymptotique de la distribution qui nous intéresse. Dans cette note on veut indiquer, dans ses grandes lignes, un procédé largement dégagé des résultats de la théorie mentionnée, qui permet un examen de l'influence réciproque entre les hypothèses et la signification des résultats expérimentaux qui est dans l'esprit du procédé par induction.

Pour fixer l'attention sur l'allure asymptotique de la distribution, sans nous occuper de sa forme d'ensemble, nous n'examinerons que la distribution des sinistres qui dépassent une certaine valeur ξ. En particulier, nous pouvons choisir pour cette valeur la limite assignée pour une certaine forme de réassurance; d'ailleurs, on peut penser à un choix arbitraire, ce qui se fera également dans cette note, où elle sera supposée indéterminée à priori, afin que l'on puisse la choisir sur la base des résultats obtenus.

Soit donc ξ l'extrême inferieur des sinistres, exprimés en termes monétaires, dont la distribution est à l'étude; et soit $F(x;\xi) = P(X \leqslant x \mid X > \xi)$ la fonction de répartition correspondante. En pratique, nous pourrons assumer comme $F(x;\xi)$ une fonction, opportunément simple, qui respecte nos hypothèses dans un voisinage droit suffisamment grand de ξ. Cette précision sera sous-entendu par la suite, de sorte que, lorsque nous considérerons le simple exemple concret exposé plus bas, nous devrons tenir compte du fait que la foncion $F(x; \xi)$ considérée n'est qu'une approximation, valide dans un voisinage droit de ξ, de la distribution effective. Notons maintenant que la fonction $F(x; \xi)$ devra être

https://doi.org/10.1017/S0515036100010448 Published online by Cambridge University Press

considérée dépendante d'un vecteur aléatoire $\Theta \equiv (\theta_1, \theta_2, \ldots, \theta_s)$ ayant une distribution assignée dans l'espace à s dimensions dans lequel il est défini, et ayant une densité $\varphi(\theta_1, \theta_2, \ldots, \theta_s)$. Comme simplification de discours, nous supposerons par la suite que $s = 1$, c'est-à-dire que F dépende d'un seul paramètre θ, et pour mettre ce fait en évidence nous écrirons $F_\theta(x; \xi)$ au lieu de $F(x; \xi)$; à la rigueur, $F_\theta(x; \xi)$ est la distribution des x plus grands de ξ subordonnément à l'hypothèse $\Theta = \theta$ (ou mieux, pour l'hypothèse de continuité absolue faite sur la distribution du paramètre, $\theta \leqslant \Theta < \theta + d\theta$) et dès lors résulterait de plus claire interprétation le symbole $F(x; \xi \mid \theta)$ (et respectivement $f(x; \xi \mid \theta)$ pour la densité $F'_x(x; \xi \mid \theta)$ que nous supposerons existante); nous nous orienterons indifféremment d'une façon ou de l'autre, suivant les raisons d'opportunité. D'après la terminologie de Raiffa-Schlaifer [1]), toute valeur du paramètre θ représente un „état du monde" potentiel que l'opérateur (celui qui doit prendre les décisions) ne peut prédire avec certitude.

Supposons maintenant que les sinistres successifs qui se produisent dans le temps soient indépendants subordonnément à une hypothèse θ (c'est-à-dire $\theta \leqslant \Theta \leqslant \theta + d\theta$) et que, dans la période expérimentale prise en considération, soient observés r sinistres $x_1 > x_2 > \ldots > x_r$ excédant la valeur ξ. On aura, en vertu du théorème de Bayes,

$$\varphi(\theta \mid x_1, x_2, \ldots, x_r; \xi) = K\varphi(\theta)\, f(x_1, x_2, \ldots, x_r; \xi \mid \theta) \qquad (1)$$

c'est-à-dire

$$\varphi(\theta \mid x_1, x_2, \ldots, x_r; \xi) = K\varphi(\theta)\, \prod_{i=1}^{r} f(x_i; \xi \mid \theta) \qquad (1a)$$

où par K on indique la constante de normalisation. C'est à peine nécessaire de rappeler que ξ est un paramètre arbitraire, dont le choix est dicté par des raisons d'opportunité. Nous reviendrons sur ces raisons sous peu. Une autre observation évidente est que les $F(x; \xi \mid \theta)$, $f(x; \xi \mid \theta)$ sont des fonctions de ξ seulement et en tant que fonction de la différence $x - \xi$, de façon qu'un changement de variables simple et secondaire aurait permis de simplifier les expressions qui figurent en (1) et (1') et, plus sensiblement, celles que

[1]) H. Raiffa et R. Schlaifer *Applied Statistical Decision Theory*, Boston, 1961.

l'on obtient concrètement dans les cas particuliers: ceci aurait toutefois mis en ombre la dépendance de ξ qui sera l'objet des considérations qui s'ensuivent. La dépendance en question intéresse expressément les (1) et (1') et tout particulièrement le facteur de vraisemblance („likelihood") $\Pi_{1}^{r} f(x_i; \xi \mid \theta)$; en effet, non seulement il s'altère, avec la croissance avec continuité de ξ, de manière constante dans tous les points de continuité de la $f(x; . \mid .)$ mais, il subit en général[1]) un saut égal à $[1 - f(x_r; \xi \mid \theta)]\, \Pi_{1}^{r-1} f(x_i; \xi \mid \theta)$ pour $\xi = x_r$, et analoguement dans les points $x_{r-1}, \ldots, x_1$. Il en découle la possibilité de cette étude de la fonction prise comme approximation de F — c'est-à-dire si et combien elle est adéquate — en relation à l' *r-ième valeur extrême*, qui est dans les finalités de la recherche.

Considérons un cas particulier, mais évidemment intéressant pour les applications effectives. Soit

$$F_{\theta}(x; \xi) = 1 - e^{-\theta(x-\xi)} \tag{2}$$

et laissons pour l'instant non précisée la distribution de θ.

(1') devient, dans ce cas

$$\varphi(\theta \mid x_1, x_2, \ldots, x_r; \xi) = K\varphi(\theta)\, \theta^r e^{-\theta \sum_{1}^{r} {}_i (x_i - \xi)}. \tag{3}$$

Une première, grossière utilisation de (3), aux fins de l'évaluation de θ, s'obtient en déterminant la valeur de θ qui rend maximum le facteur de vraisemblance

$$g(\theta, \xi) = \theta^r e^{-\theta \sum_{1}^{r} {}_i (x_i - \xi)}. \tag{4}$$

On a

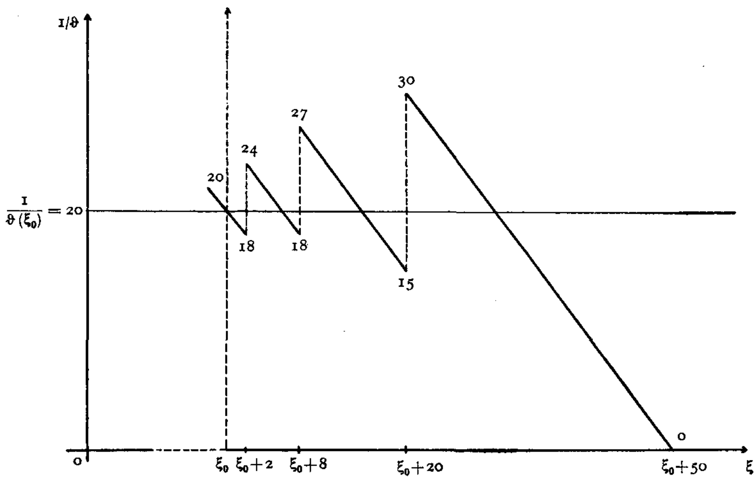
$$\frac{\partial}{\partial\theta} g(\theta, \xi) = [r - \theta \sum_{1}^{r} {}_i (x_i - \xi)]\, e^{-\theta \sum_{1}^{r} {}_i (x_i - \theta)} \tag{5}$$

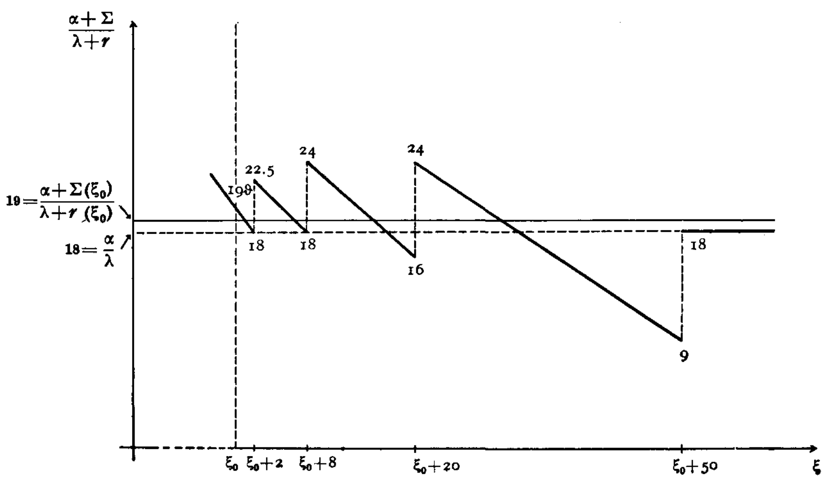
et donc on a le maximum de g, pour ξ assigné, lorsque

$$\frac{1}{\theta} = \frac{1}{r} \sum_{1}^{r} {}_i (x_i - \xi) = \frac{1}{r} \sum_{i}^{r} {}_i\, x_i - \xi. \tag{6}$$

En parole: le maximum du facteur $g(\theta, \xi)$ se vérifie lorsque le réciproque de θ est égal à la moyenne des excédences observées. Il

[1]) C'est-a-dire: sauf le cas $f(x_r; \xi \mid \theta) = 1$.





https://doi.org/10.1017/S0515036100010448 Published online by Cambridge University Press

est presque superflu de dire que, assumer comme valeur de θ celle qui rend maximum $g(\theta, \xi)$, signifie faire abstraction de nos informations et opinions précédentes, synthétisées dans la distribution de densité $\varphi(\theta)$. La fonction $[\theta(\xi)]^{-1}$ (v. fig. 1) a comme graphique une brisée (si on le pense complété dans les points de saut avec les segments verticaux correspondants); si nous indiquons par $\bar{x}_h$ la moyenne des valeurs $x_1, x_2, \ldots x_h$, on trouve facilement que la valeur du saut, en un point x_h, est indépendante de ξ, pourvu que $\xi < x_h$, et est donnée précisément par

$$\frac{1}{h}(\bar{x}_{h-1} - x_h). \tag{7}$$

Les segments compris entre deux valeurs successives x_{h-1}, x_h ont inclinaison 1. Le graphique a été obtenu en supposant, de quelque façon que ce soit, initialment fixée une valeur $\xi = \xi_0$, et en imaginant que l'on a observé les quatre valeurs $x_1 - \xi_0 = 50$, $x_2 - \xi_0 = 20$; $x_3 - \xi_0 = 8$, $x_4 - \xi_0 = 2$. Il faut noter la signification particulière du paramètre $1/\theta(\xi)$: c'est simplement la valeur moyenne des excédents de sinistres $x - \xi$ dans l'hypothèse θ. Implicitement, nous avons démontré (chose évidente, d'ailleurs) que les dimensions de θ sont celles de montant $^{-1}$; il s'ensuit en particulier qu'il n'y a pas de sens à se demander si dans un problème donné, ou exemple numérique, etc., la valeur de θ est adéquate, ou trop petite, et ainsi de suite, jusqu'à ce que soit précisée l'unité de mesure monétaire, dont θ dépend essentiellement (directement proportionelle à ladite unité). Il est peut-être opportun de signaler que le paramètre $1/\theta(\xi)$ fournit aussi, à moins du facteur $\log n$ (constant, si on suppose fixé n) la *valeur caractéristique maximum*, c'est-à-dire l'excédence $x - \xi$ qui, sur un échantillon de n observations, est dépassée (ou égalée) en moyenne une seule fois. On a en effet, par définition de valeur caractéristique maximum u_n

$$1 - e^{-\theta u_n} = 1 - 1/n \tag{8}$$

dont

$$u_n = \frac{1}{\theta} \log n.$$

Sur le graphique de la fig. 1 peut donc être étudié directement la variation soit de $M(x - \xi)$, soit de $u_n = u_n(\xi)$, au cas où l'on

https://doi.org/10.1017/S0515036100010448 Published online by Cambridge University Press

convient de prendre comme valeur de θ (ξ) la valeur fournie par le critère de la vraisemblance maximum. Nous mettrons en évidence par la suite — en nous servant de ce cas particulier — la signification pratique que prend ce critère, comme on le sait et ceci même dans le développement bayesien. Observons seulement que, même en se servant de ce critère simpliste, il est possible de répondre avec une certaine assurance aux questions a), b), b''), b''') qui seront mentionnées à la fin de cette note, pourvu que le nombre d'observations soit „suffisamment" grand; nous nous réservons d'éclaircir sous peu même la signification de „suffisamment".

En revenant au développement strictement bayesien donné à notre problème, en relation à la définition de la fonction φ (θ), nous estimons opportun de relever une observation de Raiffa et Schlaifer que, pour clarté comme pour concision, nous ne saurions exprimer mieux que ces Auteurs [1]:

„An obvious way of finding a specific distribution . . . [dans l'espace des états possibles] . . . is to start by selecting some *family* of distributions defined by a mathematical formula containing a certain number of adjustable parameters and then to select the specific member of this family which meets the decision maker's quantitative specifications by giving the proper numerical values to these parameters".

Et encore:

„The fact that the decision maker cannot specify every detail of his prior distribution by direct assessment means that there will usually be considerable latitude in the choice of the family of distributions to be used . . . even though the selection of a particular member *within* the chosen family will usually be wholly determined by the decision maker's expressed beliefs or betting odds".

Dans notre cas, pour des raisons qui apparaîtront bientôt évidentes [2], il convient de prendre comme φ (θ) la distribution Gamma$_1$ de paramètres α, λ:

$$\varphi\,(\theta)\,;\,\alpha,\lambda) = K\theta^{\lambda-1}\,e^{-\alpha\theta} \qquad (\theta > 0) \tag{9}$$

[1]) Voir Raiffa-Schlaifer, *Oeuv. cit.*, pag. 43.

[2]) Pour une tractation générale de l'argument, voir la monographie de Raiffa-Schlaifer maintes fois citée, Chap. 3.

où la constante de normalisation K est donnée par

$$K = \alpha^{\lambda}/\Gamma(\lambda). \tag{9}$$

En vérité, on trouve que dans ce cas, en écrivant par simplicité Σ au lieu de $\sum_{1}^{r}{}_i (x_i - \xi)$ (2) devient

$$\varphi(\theta \mid x_1, x_2, \ldots, x_r; \xi) = K\theta^{\lambda+r-1}\, e^{-(\alpha+\Sigma)\theta} \quad ^{1)} \tag{10}$$

c'est-à-dire que la distribution de θ subordonnée au r observation $x_1, \ldots, x_r$ est encore la Gamma, avec paramètres $\alpha + \Sigma$ et $\lambda + r$ au lieu de α et λ. Remarquons maintenant que, une fois assigné la distribution (9) et précisé sur la base des connaissances et des opinions de l'operateur des paramètres α et λ, comme valeur de θ dans (2) on utilisera évidemment un indice opportun apte à synthétiser la distribution. Le choix peut être douteux, dans notre cas, entre moyenne et la mode; si toutefois on observe que la première est donnée par λ/α et la seconde par $(\lambda - 1)/\alpha$, il est évident qu'aussitôt α et λ sont suffisamment grands le choix devient indifférent; le fait que les expériences successives aient justement l'effet de faire croître les deux paramètres, comme on a observé plus haut, est essentiel.

Par consequent, une fois prise comme valeur de θ dans (2) sa valeur moyenne dans la distribution (9) (respectivement, après les experiences, dans la distribution (10)), nous aurons:

$$F_{\theta_0}(x; \xi) = 1 - e^{-\frac{\lambda}{a}(x-\xi)} \tag{11}$$

et après les r valeurs observées $x_1, x_2, \ldots, x_r$:

$$F_{\theta_r}(x; \xi) = 1 - e^{-\frac{\lambda+r}{a+\Sigma}(x+\xi)}, \tag{12}$$

avec signification évidente des symboles θ_0, θ_r.

Nous pouvons à ce point répéter, avec de légères variantes, les considérations faites plus haut à propos de la dépendance de ξ de la valeur θ, maximum de $g(\theta, \xi)$.

---

[1]) Continuons à indiquer par K la constante de normalisation, sa valeur étant déterminée sans ambiguité par sa signification, même si par là, dans le cas d'une même tractations, nous nous trouvons à indiquer par K les constantes diverses; dans le cas présent, évidemment $K = (\alpha + \Sigma)^{\lambda+r}/L(o + r)$.

https://doi.org/10.1017/S0515036100010448 Published online by Cambridge University Press

Nous avons en effet

$$\frac{1}{\theta_r(\xi)} = \frac{\alpha + \sum_1^r {}_i (x_i - \xi)}{\lambda + r(\xi)} = \frac{\alpha + \sum_1 {}_i a_i - r(\xi)\xi}{\lambda + r(\xi)} \tag{13}$$

où la fonction $r(\xi)$ [1]) est constante à intervalles, et passe de la valeur entière r à la valeur entière $r - 1$ lorsque ξ croit de $x_r - \varepsilon$ à $x_r + \varepsilon$; pour la clarté, avec évidente signification des symboles, nous pourrons écrire

$$\frac{1}{\theta_r(\xi)} = -A_\xi \xi + B_\xi. \tag{13'}$$

Dans chaque point de continuité (et donc de constance) des coefficients, c'est-à-dire pour $\xi \neq x_i$ ($i = 1, 2, \ldots, r$), (13') constitue l'équation d'une droite ayant inclinaison négative de valeur absolue $A_\xi < 1$, telle que $A_\xi \to 1$ lorsque $r(\xi) \to \infty$. Une étude détaillée des variations est facile, mais pour brévité nous ne nous y arrêtons pas. Observons seulement que si ξ varie dans un intervalle tel que Σ et r se maintiennent très grands par rapport à α et λ, de façon que α et λ puissent être négligées dans le rapport $(\lambda + r)/$ $/(\alpha + \Sigma)$, on se retrouve sans plus dans le cas considéré à propos de la vraisemblance maximum. Ici l'on voit clairement la signification et les limites du critère de la „likelihood" maximum que nous avons annoncés. Il est d'ailleurs aisé, suivant les buts que l'on se propose, de préciser quantitativement la vague affirmation „Σ et r très grands par rapport à α et λ".

Nous indiquons ici brièvement les questions fondamentales et les situations d'ordre pratique qui, avec la formulation exposé ci-dessus, peuvent être étudiées et, respectivement, mises sous contrôle.

a) Interprété ξ comme limite des sinistres non recouverts de réassurance, c'est-à-dire dans l'hypothèse que soient réassurées seulement les excédences $x - \xi$, évaluer ξ sur la base des hypothèses

[1]) Par simplicité, nous indiquons ici et par la suite par $r(\xi)$ l'indice de la valeur minimum observée x_r qui excède ξ. Veuillez noter la légère différence de notations par rapport aux (6) et (7). La présente notation simplifiée est cohérente avec le fait que les valeurs expérimentales se comportent, aux fins de notre problème, comme si elles cessaient d'exister, lorsqu'on prend ξ plus grand qu'elles.

initiales et des résultats expérimentaux, de façon à ce que soient respectées les conditions données d'interprétation pratique aisée et immédiate: p. ex. que la valeur moyenne $1/\theta$ des excédences réassurées s'avère égale à une valeur assignée, ou bien la valeur caractéristique maximum relative à un nombre n d'excédence préfixé, etc. Ces considérations ont de l'intérêt quant aux deux points de vue opposés de la compagnie cédante et de la compagnie de réassurance.

b) Si les expériences successives [1]) conduisent à des „modifications" trop sensibles de la valeur θ, il faut en étudier les trois possibilités:

b′) la fonction choisie n'est pas adéquate;

b″) les hypothèses initiales étaient non-adhérentes à la réalité;

b‴) avec l'écoulement du temps (au sens macroscopique: dix ans, cinquante ans, etc.) et des circonstances — rappeler l'extrême facilité de tenir compte des variations du pouvoir d'achat de la monnaie, qui équivaut à un changement de l'unité de mesure — ont changé les distributions des sinistres (suivant l'entité) bien que la validité des *familles* de distributions considérées (respectivement pour f et pour φ) continue à subsister.

En ligne de principe, l'étude effectuée dans le but de distinguer les trois possibilités énoncées ci-dessus ne se présente pas comme prohibitif. En nous exprimant de façon grossière, nous pourions dire que dans le cas b″) devrait se produire une tendance de θ à se stabiliser autour d'une valeur sensiblement différente de λ/α, dans le cas b‴) une variation de θ lente et systématique, et dans le cas b′) une absence de ces types de „régularité". La possibilité b‴) nous semble d'un intérêt particulier et favoriserait une analyse particulière des résultats, en les distinguant et en les pondérant opportunément en fonction du temps, suivant la formulation de la „credibility theory" [2]).

[1]) Il est évident que toute expérience successive comporte une rénumération des valeurs expérimentales; ceci ne constitue pas une difficulté grave si, comme déjà dans la nature du problème, les valeurs excédentes sont rares.

[2]) Cette théorie développée et appliquée en Amérique par les „Casualty Actuaries" est presque inconnue en Europe. Dans le courant de cette année B. de Finetti en février et A. Mayerson en mai on traité l'argument au „Séminaire Actuariel" (Istituto Italiano degli Attuari). Je saisis cette occasion pour remercier le Prof. de Finetti pour avoir suggéré ce travail, développé dans l'ordre d'idées de ladite théorie.

A titre purement illustratif, nous croyons utile de compléter ces brèves considérations en complétant l'examen du cas numérique dont nous nous sommes servis pour tracer le graphique de la fig. 1.

Attribuons, dans ce but, à α et λ des valeurs numériques, à notre choix, bien entendu, étant donné le but purement d'exemple, mais choisies suivant des raisons évidentes d'opportunité. Faisons donc:

$$\alpha = 18$$

$$\lambda \equiv 1$$

et considérons l'allure de la fonction (13) (ou bien (13')).

On trouve aisément le graphique de la figure 2.

https://doi.org/10.1017/S0515036100010448 Published online by Cambridge University Press