# DISTRIBUTIONS DES SINISTRES INCENDIE SELON LEUR COÛT

GIOVANNA FERRARA

Italie

## 1. INTRODUCTION

On peut penser que la dimension des sinistres est la résultante d'un très grand nombre de causes indépendantes à effets positifs.

Soient

$$\xi_1, \xi_2, \ldots, \xi_n$$

les variables aléatoires indépendantes qui représentent les facteurs élementaires exerçant leur action dans l'ordre indiqué par les indices.

Si $X_\nu$ est l'ampleur du sinistre dû aux facteurs $\xi_1, \xi_2, \ldots, \xi_\nu$ nous pouvons supposer que l'augmentation causée par $\xi_{\nu+1}$ soit proportionelle à $\xi_{\nu+1}$ et à une certaine fonction, $g(X_\nu)$ de la dimension $X_\nu$.

C'est à dire, nous supposons qu'il y ait les relations:

$$X_{\nu+1} = X_\nu + \xi_{\nu+1}\, g(X_\nu)$$

où

$$\nu = 0, 1, \ldots, n-1$$

et

$$X_0 = 0; \quad g(X_0) = 1$$

Il s'ensuit que

$$\xi_1 + \xi_2 + \ldots + \xi_n = \sum_{0}^{n-1} \frac{X_{\nu+1} - X_\nu}{g(X_\nu)} \tag{1}$$

Si chaque facteur n'apporte qu'un faible contribut à l'augmentation du sinistre, nous pouvons alors poser approximativement aussi:

$$\xi_1 + \xi_2 + \ldots + \xi_n = \int_{X_0}^{X} \frac{dt}{g(t)} \tag{2}$$

https://doi.org/10.1017/S0515036100008217 Published online by Cambridge University Press

Comme $X$ varie de $X_0$ à $X_n$ (et que $X = X_n$ désigne la dimension extrême du sinistre) le second membre de la (2) décrit une variable aléatoire, fonction de $X$.

Puisque par hypothèse $\xi_1, \xi_2, \ldots, \xi_n$ sont des variables aléatoires indépendante, si $n$ est suffisament grand, d'après le théorème de la limite des probabilités il s'insuit que la variable aléatoire au second membre de la (2), est, à la limite, distribuée suivant la loi normale. Si l'on pose $g(t) = t - c$, la variable aléatoire normale qui décrit la dimension du sinistre pourra être posée dans la forme $\ln(X - c)$.

Si $Z^*$ est une variable aléatoire normale réduite (0,1), entre $Z^*$ et $X$ il y aura la relation

$$Z^* = \frac{\ln(X - c) - m}{\sigma} \tag{3}$$

C'est à dire en posant $\alpha = 1/\sigma$; $\beta = - m/\sigma$

$$Z^* = \alpha \ln(X - c) + \beta \tag{4}$$

d'où

$$X = c + \exp \frac{Z^* - \beta}{\alpha} \tag{5}$$

La fonction cumulative de la variable aléatoire $X$ sera pour $x > c$

$$\begin{aligned} F(x) = \text{Prob}\,\{X \leqslant x\} = \text{Prob}\left\{c + \exp \frac{Z^* - \beta}{\alpha} \leqslant x\right\} = \\ = \text{Prob}\,\{Z^* \leqslant \alpha \ln(x - c) + \beta\} \end{aligned} \tag{6}$$

c'est à dire

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha \ln(x-c) + \beta} \exp\left\{-\frac{t^2}{2}\right\} dt \tag{7}$$

et la fonction de densité

$$f(x) = \frac{1}{\sqrt{2\pi}} \frac{\alpha}{x - c} \exp - \frac{1}{2} \{\alpha \ln(x - c) + \beta\}^2 \tag{8}$$

Pour les développements successifs il est préférable de considérer la variable normale $Z\left(0; \frac{1}{\sqrt{2}}\right)$ liée à $Z^*$ par la relation

$$Z^* = \frac{Z}{\sqrt{2}}$$

c'est à dire

$$Z = a \ln(x - c) + b \tag{4 bis}$$

où $$a = \frac{\alpha}{\sqrt{2}}; \qquad b = \frac{\beta}{\sqrt{2}}$$

La fonction de densité de $Z$ est:

$$f(z) = \frac{1}{\sqrt{\pi}} \exp(-z^2)$$

et celle de $X$

$$f(x) = \frac{1}{\sqrt{\pi}} \cdot \frac{a}{x - c} \exp - \{a \ln(x - c) + b\}^2 \tag{8 bis}$$

Les trois premiers moments par rapport à 0 de la variable aléatoire $X$ en fonction de $Z$ seront:

$$\mu_1 = c + \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \exp\left(\frac{z - b}{a}\right) \exp(-z^2)\, dz \tag{9}$$

$$= c + \exp\left(\frac{1 - 4ab}{4a^2}\right)$$

$$\mu_2 = c^2 + (\mu_1 - c)^2 \exp\left(\frac{1}{2a^2}\right) + 2c(\mu_1 - c) \tag{10}$$

$$\mu_3 = c^3 + (\mu_1 - c)^3 \exp\left(\frac{3}{2a^2}\right) + 3c^2(\mu_1 - c) + 3c(\mu_1 - c)^2 \exp\left(\frac{1}{2a^2}\right) \tag{11}$$

Pour l'écart absolu moyen par rapport à la moyenne, on aura:

$$S = \int_{-\infty}^{+\infty} | x - \mu_1 | f(x)\, dx = 2 \int_{-\infty}^{\mu_1} (\mu_1 - x) f(x)\, dx \tag{12}$$

puisque pour $x = \mu_1$; $z = 1/4a$

$$S = \frac{2}{\sqrt{\pi}} \int_{-\infty}^{+1/4a} \exp \frac{1 - 4ab}{4a^2} - \exp \frac{z - b}{a} \cdot dz \tag{13}$$

d'où vient:

$$S = \frac{2(\mu_1 - c)}{\sqrt{\pi}} \int_{-1/4a}^{+1/4a} \exp(-t^2)\, dt$$

rappelant la fonction connue:

$$\Theta(\lambda) = \frac{1}{\sqrt{\pi}} \int_{-\lambda}^{+\lambda} \exp(-t^2)\, dt$$

la relation précédente devient:

$$S = 2(\mu_1 - c)\, \Theta\left(\frac{1}{4a}\right) \tag{14}$$

Le coefficient de variation suivant l'écart absolu moyen est donc:

$$\frac{S}{\mu_1} = \frac{2(\mu_1 - c)}{\mu_1}\, \Theta\left(\frac{1}{4a}\right) \tag{15}$$

## 2. Ajustement des données italiennes incendie à la loi log-normale

D'après recherches faites sur les données du „Concordato Italiano Incendi Rischi Industriali" pour la période 1963-1965, il est possible de tirer la distribution des sinistres selon leur coût soit pour l'ensemble des risques assurés soit pour les principaux regroupements des industries effectués suivant l'activité exercée et suivant le capital assuré.

Ces distributions présentent quelques particularités permettant de supposer que l'ampleur des sinistres suit la loi log-normale.

Il s'agit en fait de distributions unimodales, dissymétriques, étalées vers la droite: le 50% des sinistres est en général concentré dans les deux premières classes; tandis qu'ils diminuent progressivement dans les autres.

La fonction d'ajustement est donc donnée par la (8 bis).

Pour déterminer la valeur des paramètres, on a consideré la variable $Z$ liée à $X$ par la relation (4 bis).

Puisque $Z$ est une variable aléatoire normale, les fréquences cumulées, c'est à dire les fréquences des valeurs de $Z$ inférieures à $z_i$, sont exprimées par

$$\frac{1}{\sqrt{\pi}} \int_{-\infty}^{z_i} \exp\,(-z^2)\,dz$$

Désignant par $P(x_i)$ les fréquences cumulées de $X$ (c'est à dire les fréquences des valeurs de $X$ inférieure a $x_i$) on peut poser

$$P(x_i) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{z_i} \exp\,(-z^2)\,dz \tag{16}$$

Après avoir établi les valeurs $z_i$, on a considéré les relations

$$z_i = a \ln(x_i - c) + b \tag{17}$$

Le paramètre $c$ qui se présente en forme non linéaire a été déterminé par tâtonnements, en prenant comme valeur de départ la quantité $c$ d'après l'égalité:

$$\frac{z_3 - z_2}{z_2 - z_1} = \frac{\ln(x_3 - c) - \ln(x_2 - c)}{\ln(x_2 - c) - \ln(x_1 - c)} \tag{18}$$

concernant les trois premières classes de sinistres.

Les paramètres $a$ et $b$ ont été déterminés par la méthode des moindres carrés.

D'après les valeurs établies suivant la (17), par l'emploi de la (16) on a trouvé les valeurs théoriques $P^*(x_i)$ et $f^*(x_i)$.

Pour apprécier l'ajustement on a calculé l'écart moyen entre les valeurs théoriques et les valeurs observées, c'est à dire

$$I = \sum_{1}^{n}{}_i \, | f(x_i) - f^*(x_i) |$$

et l'indice $\chi^2$.

De plus la moyenne et le coefficient de variation selon l'écart absolu moyen des distributions observées ont été comparés à ceux de l'application des relations (9) et (15).

Nous reportons en appendice les phases du déroulement d'ajustement concernant la categorie des industries métallurgiques.

## 3. Analyse des résultats obtenus

A) *Catégories d'industries selon leur activité*

L'ajustement à la loi log-normale des distributions des sinistres incendie suivant leur coût peut être considéré en général satisfaisant.

Les valeurs de $\chi^2$ sont pour toutes les catégories inférieures à la valeur théorique tabulée au seuil de probabilité de 10%.

L'écart moyen entre les valeurs observées et les valeurs ajustées est uniquement élevé pour les industries des combustibles (23%). De plus dans cette distribution la moyenne théorique s'éloigne dans une certaine mesure de la moyenne observée et il en est de même pour le coefficient de variation selon l'écart absolu moyen.

Il faut noter cependant que les sinistres de cette catégorie ne sont qu'une modeste partie de l'ensemble des sinistres (seulement le 1,3%) et c'est pourquoi la moyenne relevée a un degré de précision très bas.

En ce qui concerne les industries de l'alimentation et du cuir les ajustements peuvent être acceptables bien que les valeurs effectives de la moyenne et du coefficient de variation s'éloignent des valeurs théoriques.

B) *Catégories d'industries selon le capital assuré*

Si l'on regroupe les données des sinistres suivant le capital assuré, les distributions des sinistres selon leur coût présentent un ajustement moins satisfaisant à la loi log-normale.

Les valeurs de $\chi^2$ sont élevées excepté la classe des risques avec un capital supérieur à 1.000 millions.

L'écart moyen entre les valeurs théoriques et les valeurs observées est cependant relativament bas; les moyennes et les coefficients de variation sont suffisamment proches.

https://doi.org/10.1017/S0515036100008217 Published online by Cambridge University Press

## Appendice

*Tableau I*

Distribution (ramenée à un total de 10.000) des sinistres incendie selon le coût par catégories d'industries

Données du „Concordato Italiano Incendi — Rischi Industriali (1963/1965)"

| Montant des si nistres (en 1.000 L.) | Catégories d'industries | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Alimentat. | Papier | Céramiq. | Chimiq. | Cumbust. | Cuir | Bois | Métallurg. | Textiles | Ensemble |
| De 0 à 250 | 4.446 | 4.346 | 4.635 | 3.125 | 2.208 | 5.286 | 4.194 | 4.971 | 5.123 | 4.663 |
| de 250 à 500 | 1.477 | 1.231 | 1.685 | 1.504 | 1.688 | 1.499 | 1.300 | 1.602 | 1.802 | 1.579 |
| de 500 à 1.000 | 1.380 | 1.423 | 1.236 | 1.641 | 1.169 | 1.065 | 1.161 | 1.214 | 1.077 | 1.222 |
| de 1.000 à 2.000 | 979 | 846 | 787 | 1.328 | 649 | 730 | 1.144 | 934 | 662 | 891 |
| de 2.000 à 4.000 | 530 | 462 | 702 | 937 | 1.299 | 454 | 971 | 481 | 493 | 598 |
| de 4.000 à 8.000 | 385 | 692 | 393 | 547 | 1.299 | 355 | 607 | 374 | 394 | 429 |
| de 8.000 à 16.000 | 337 | 346 | 225 | 371 | 519 | 197 | 225 | 180 | 175 | 231 |
| de 16.000 à 32.000 | 161 | 346 | 225 | 254 | 519 | 158 | 225 | 151 | 169 | 192 |
| de 32.000 à 64.000 | 177 | 192 | 84 | 117 | 260 | 158 | 69 | 72 | 123 | 117 |
| plus de 64.000 | 128 | 116 | 28 | 176 | 390 | 98 | 104 | 21 | 52 | 78 |

https://doi.org/10.1017/S0515036100008217 Published online by Cambridge University Press

*Tableau II*

Distribution (ramenée à un total de 10.000) des sinistres incendie selon le coût et le capital assuré

| Montant des sinistres (en 1.000 L.) | Industries avec un capital assuré | | | |
|---|---|---|---|---|
| | de 0 à 100 millions | de 100 à 250 millions | de 250 à 1.000 millions | plus de 1.000 millions |
| De 0 à 250 | 5.663 | 4.773 | 4.527 | 3.450 |
| de 250 à 500 | 1.344 | 1.709 | 1.638 | 1.649 |
| de 500 à 1.000 | 1.021 | 1.223 | 1.259 | 1.429 |
| de 1.000 à 2.000 | 759 | 788 | 953 | 1.109 |
| de 2.000 à 4.000 | 411 | 593 | 721 | 707 |
| de 4.000 à 8.000 | 361 | 391 | 320 | 669 |
| de 8.000 à 16.000 | 162 | 208 | 189 | 387 |
| de 16.000 à 32.000 | 180 | 120 | 197 | 289 |
| de 32.000 à 64.000 | 87 | 145 | 116 | 121 |
| plus de 64.000 | 12 | 50 | 80 | 190 |

*Tableau III*

Industrie métallurgique: ajustement de la distribution observée à la loi log-normale

| Extremité supérieure de classe | $\mathrm{Log}_{10}$ $(x_i - 70)$ | Valeurs observées | | | | Valeurs ajustées | | | | $f(x_i) - f^*(x_i)$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Effectifs: sinistres | $f(x_i)$ | $P(x_i)$ | $z(x_i)$ | $z^*(x_i)$ | $P^*(x_i)$ | $f^*(x_i)$ | $n^*(x_i)$ | |
| (1) = $x_i$ | (2) | (3) = $n(x_i)$ | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| 250 | 2,25527 | 692 | 0,4971 | 0,4971 | —0,005 | —0,009 | 0,4949 | 0,4949 | 689 | 0,0022 |
| 500 | 2,63347 | 223 | 0,1602 | 0,6573 | 0,286 | 0,278 | 0,6529 | 0,1580 | 220 | 0,0022 |
| 1.000 | 2,96848 | 169 | 0,1214 | 0,7787 | 0,543 | 0,553 | 0,7745 | 0,1216 | 169 | —0,0002 |
| 2.000 | 3,28556 | 130 | 0,0934 | 0,8721 | 0,804 | 0,774 | 0,8632 | 0,0887 | 123 | 0,0047 |
| 4.000 | 3,59439 | 67 | 0,0481 | 0,9202 | 0,995 | 1,008 | 0,9230 | 0,0598 | 83 | —0,0117 |
| 8.000 | 3,89927 | 52 | 0,0374 | 0,9576 | 1,219 | 1,239 | 0,9602 | 0,0372 | 52 | 0,0002 |
| 16.000 | 4,20222 | 25 | 0,0180 | 0,9756 | 1,393 | 1,469 | 0,9812 | 0,0210 | 29 | —0,0030 |
| 32.000 | 4,50420 | 21 | 0,0151 | 0,9907 | 1,664 | 1,669 | 0,9919 | 0,0107 | 15 | 0,0044 |
| 64.000 | 4,80570 | 10 | 0,0072 | 0,9979 | 2,020 | 1,928 | 0,9968 | 0,0049 | 7 | 0,0023 |
| ∞ | — | 3 | 0,0021 | 1,0000 | — | — | 1,0000 | 0,0032 | 4 | —0,0011 |

$z^* = a\ \mathrm{Log}_{10}(x - 70) + b$

$a' = 0{,}7591; \quad b = -1{,}7207$

$I = \Sigma \mid f(x_i) - f^*(x_i) \mid = 0{,}0320$

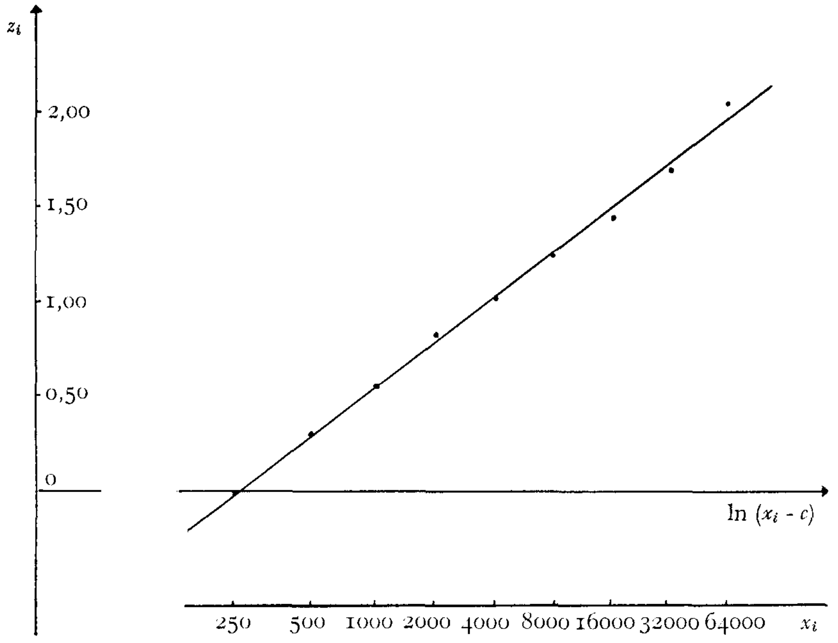
$$\chi^2 = \sum \frac{[n(x_i) - n^*(x_i)]^2}{n^*(x_i)} = 6{,}852; \quad \text{degrés de liberté} = 5$$

https://doi.org/10.1017/S0515036100008217 Published online by Cambridge University Press

*Tableau IV*

Ajustement des distributions du „Concordato Italiano Incendi" à la loi log-normale

| Industries | a | b | c | I | $\chi^2$ | degrés de liberté | Moyenne en millions de L | | Coefficient de variation | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | obser-vé | théoriq. | obser-vée | théoriq. |
| 1) Alimentation | 0,2655 | −1,3802 | 120 | 0,0850 | 8,582 | 6 | 4,100 | 6,390 | 1,50 | 1,60 |
| Papier | 0,2953 | −1,7279 | — | 0,1290 | 7,249 | 6 | 5,980 | 6,110 | 1,50 | 1,54 |
| Céramiques | 0,3432 | −1,9280 | — | 0,0994 | 8,087 | 6 | 2,430 | 2,300 | 1,38 | 1,39 |
| Chimiques | 0,3050 | −1,8278 | 120 | 0,1050 | 7,816 | 6 | 6,230 | 6,005 | 1,48 | 1,48 |
| Combustibles | 0,3169 | −2,2281 | — | 0,2332 | 6,955 | 5 | 8,540 | 13,630 | 1,32 | 1,47 |
| Cuir | 0,2402 | −1,0655 | 150 | 0,0478 | 3,348 | 6 | 3,080 | 6,550 | 1,52 | 1,68 |
| Bois | 0,3306 | −1,9569 | — | 0,0634 | 7,257 | 7 | 3,030 | 3,660 | 1,32 | 1,43 |
| Métallurgiques | 0,3297 | −1,7207 | 70 | 0,0320 | 6,852 | 5 | 2,070 | 1,910 | 1,40 | 1,38 |
| Textiles | 0,2635 | −1,1976 | 150 | 0,0298 | 9,379 | 6 | 3,000 | 3,980 | 1,53 | 1,57 |
| 2) Avec un capital assuré | | | | | | | | | | |
| de 0 à 100 mill. | 0,3329 | −1,7296 | — | 0,0600 | 29,359 | 7 | 1,740 | 1,720 | 1,40 | 1,42 |
| de 100 à 250 mill. | 0,2845 | −1,4107 | 130 | 0,0374 | 13,544 | 6 | 2,490 | 3,250 | 1,40 | 1,51 |
| de 250 à 1.000 mill. | 0,2832 | −1,4489 | 125 | 0,0593 | 13,174 | 6 | 3,430 | 3,890 | 1,48 | 1,53 |
| plus de 1.000 mill. | 0,2846 | −1,6539 | 125 | 0,0534 | 6,698 | 6 | 6,310 | 7,430 | 1,48 | 1,54 |
| Ensemble | 0,2838 | −1,4640 | 110 | 0,0258 | 14,328 | 6 | 3,360 | 3,870 | 1,48 | 1,53 |




INDUSTRIE METALLURGIQUE: Ajustement de la distribution des sinistres incendie selon le coût à la loi log-normale. Représentation grafique de la variable $Z$. $Z = a \ln(X - c) + b$; $a = 0{,}3297$ $b = -1{,}7107$; $c = 70$

Source: „Concordato Italiano Incendi — Rischi Industriali" Données de la période 1963-1965. Elaboration du Service Etudes des Assicurazioni Generali