MATHÉMATIQUES ET SCIENCES HUMAINES

Problèmes d'enseignement. Applications pratiques des lois de probabilité 3

Mathématiques et sciences humaines, tome 26 (1969), p. 45-49
<http://www.numdam.org/item?id=MSH_1969__26__45_0>

M.S.H. (7e année, nº 26, 1969, p. 45-49)

PROBLÈMES D'ENSEIGNEMENT

APPLICATIONS PRATIQUES DES LOIS DE PROBABILITÉ 3

par

B. LECLERC

LOI DE PARETO — *ÉCONOMIE*

HAYAKAWA Miyoji. — "The application of Pareto's law of income to Japanese Data", *Econometrica*, Vol. 19, nº 2, April 1951, p. 174-183.

Modèle.

La distribution des revenus supérieurs à la valeur modale de la distribution totale suit approximativement la loi de Pareto :

$$N = A\,X^{-\alpha}.$$

N est le nombre d'individus ayant un revenu supérieur ou égal à X. A, α sont des constantes à déterminer empiriquement. Application à des données japonaises : revenus de la population totale dans 108 des 274 communes de Hokkaido en 1932 — résultats complets d'une enquête fiscale. Le tableau détaillé des résultats est donné.

Estimation.

A partir de $\log_{10} N = \log_{10} A - \alpha \log_{10} X$, on calcule pour la distribution tronquée (revenus au-dessus de 1 000 yens, c'est-à-dire revenus imposables) :

$$\alpha = 1{,}55837 \qquad \log A = 10{,}8448$$

pour la distribution entière :

$$\alpha = 1{,}40774 \qquad \log A = 9{,}96036.$$

L'auteur note que le fait que α se trouve aux environs de 1,5 est classique, stable dans le temps et dans l'espace. Les constantes sont calculées également pour les années 1926, 1930, 1934.

Pas de test, mais

Discussion du modèle.

La moyenne est égale à 656 yens ; le mode à 363 yens ; l'écart-type à 4 424 yens ; le coefficient d'asymétrie $\frac{\sqrt[3]{\mu_3}}{\sqrt{\mu_2}} = 7{,}72423.$

Ce dernier est très prononcé. L'auteur essaie l'adaptation à une courbe de Gram-Charlier de type B, mais la rejette à cause de l'asymétrie trop élevée.

Enfin l'auteur trouve bonne l'adaptation à la loi de densité :

$$\Phi (Z) = \frac{a}{z^{\mu+1}} \frac{1}{e^{\frac{b}{z}} - 1} \qquad \mu > 0$$

où $z = X - c$ et $\mu > O$, proposée par H. T. Davis (voir fiche correspondante dans *M.S.H.*, n° 24).

Il donne $\frac{a}{b} = 64 \quad 196 \quad 200 \quad b = 407$

$c = 100$ yens $\qquad \mu = 2{,}5$

estimés à partir de la distribution empirique.

RÉFÉRENCES BIBLIOGRAPHIQUES

JOHNSON N. O. — "The Pareto Law", *Review of Economic Statistics*, 19, Feb. 37, p. 20-26.

DAVIS H. T. — *The Analysis of Economic Time-Series.* Cowles Commission Monograph 6, Bloomington, Ind. Principia Press, p. 409-418.

CHAMPERNOWNE D. G. — "The Theory of Income Distribution", *Econometrika*, Vol. 5, Oct. 1937, p. 379-381.

LOI DE POISSON — *DISTRIBUTION DES ACCIDENTS DISTRIBUTIONS SPATIALES*

FELLER William. — "The Poisson distribution" et "Observations fitting the Poisson distribution", par. 6 et 7 du chap. VI de *An introduction to probability theory and its applications*, 2nd edition, Vol. 1, New York, John Wiley and Sons, 1957.

Modèle.

L'auteur présente le modèle poissonnien :

$$p\,(k\,;\lambda) = e^{-\lambda}\frac{\lambda^k}{k!}$$

où k est un entier positif ou nul et λ est la moyenne théorique. L'auteur observe que ce modèle s'applique non seulement à des distributions de nombres d'événements aléatoires par période de temps, mais aussi de la même façon, à la distribution du nombre de points aléatoires dans un domaine spatial, quel que soit la forme de ce domaine. Il en donne quelques exemples.

Estimation.

Identification de la moyenne théorique λ et de la moyenne empirique.

Test.

Test du χ^2. Pour chaque application, l'auteur ne se fixe pas de seuil a priori, mais donne a posteriori la probabilité p pour la variable du χ^2 correspondante d'être supérieure à la quantité-test.

Applications.

a) Rutherford, Chadwick et Ellis ont observé une substance radioactive pendant 2 608 intervalles de temps de 7,5 secondes chacun. L'ajustement à la loi de Poisson de la distribution du nombre de désintégrations par unité de temps donne $p = 0{,}17$.

b) R. D. Clarke (1946) a étudié une distribution spatiale : celle des impacts des bombes sur Londres durant la seconde guerre mondiale. La surface de la ville a été divisée en 576 parties de 0,25 km^2 chacune. L'ajustement à la loi de Poisson est étonnamment bon : $p = 0{,}88$.

c) D. G. Catcheside, D. E. Lea et J. M. Thoday ajustent la loi de Poisson au nombre de modifications chromosiques par cellule, provoquées par des rayons X. Au cours de l'expérience, on trouve des valeurs de p entre 0,05 et 0,95, bien réparties sur le segment (0,1).

d) Une étude sur 267 numéros de téléphone a été effectuée par F. Thorndike : la distribution du nombre d'erreurs par numéro au cours d'une période de temps fixée s'ajuste très bien à la loi de Poisson. Il n'en est pas de même pour d'autres statistiques téléphoniques pour lesquelles il y a une dépendance évidente entre les événements.

e) Comptage de bactéries sur une assiette de Petri : la surface circulaire visible au microscope est approximativement divisée en petits carrés égaux. Huit observations de la distribution du nombre de bactéries par carré sont ajustées à la loi de Poisson. Dans le moins bon cas, on a $p = 0{,}26$. Les sept autres ajustements donnent p compris entre 0,53 et 0,97 (T. Matuszewsky, J. Supinska et J. Neyman, 1936).

L'auteur donne les tableaux numériques correspondants à ces diverses applications.

RÉFÉRENCES BIBLIOGRAPHIQUES

RUTHERFORD, CHADWICK et ELLIS. — *Radiations from radioactive substances*, Cambridge, p. 172. 1920.

CRAMER H. — *Mathematical methods of statistics*, Upsala and Princeton, p. 436. 1945.

CLARKE R. D. — "An application of the Poisson distribution", *J. of the Institute of Actuaries*, Vol. 72, p. 48. 1946.

CATCHESIDE D. G., LEA D. E. et THODAY J. M. — "Types of chromosome structural change induced by the irradiation of tradescantia microspores", *Journal of Genetics*, Vol. 47, p. 113-136. 1945-46.

THORNDIKE F. — "Applications of Poisson's probability summation", *The Bill System Technical Journal*, Vol. 5, p. 604-624. 1926.

NEYMAN J. — *Lectures and conferences on mathematical statistics*, Dept. of Agriculture, Washington. 1938.

MATUSZEWSKY T., SUSPINKA J. et NEYMAN J. — In *Zentralblatt fur Bakteriologie, Parasitenkunde und Infektions Krankheiten*, II. Abt., Vol. 95. 1936.

LOIS DE PARETO, DE ZIPF ET DE YULE
LOI BINOMINALE NÉGATIVE — *ÉCONOMIE*
LOI DE POISSON — *LITTÉRATURE SCIENTIFIQUE*

KENDALL M. G. — "Natural Law in the Social Sciences", *J. of the Royal Statistical Society*, Série A (General), 1961, Part I, p. 1-16.

Au cours de son allocution présidentielle à la Royal Statistical Society, l'auteur présente quelques lois observées par divers chercheurs dans le domaine des sciences sociales. Il rappelle d'abord la loi de Pareto : le nombre y de personnes ayant le revenu x est donné par :

$$y = \frac{A}{x^{1+\alpha}}$$

où A et α sont des constantes. Cette loi est au moins une bonne approximation dans tous les pays où des statistiques précises ont pu être établies.

I. — LOI DE ZIPF.

La loi de Pareto est un cas particulier de la loi de Zipf, selon lequel on observe fréquemment dans les Sciences Sociales des distributions du type

$$y = \frac{A}{x^p}$$

où y est le nombre d'individus pour lesquels on observe une valeur $\geqslant x$ d'une caractéristique mesurable. La constante p est souvent proche de l'unité, auquel cas on a : $xy \simeq A$, pour tout x.

L'auteur souligne le grand intérêt des travaux de Zipf, mais critique le fondement de la loi sur le « principe du moindre effort », appliqué à toutes les activités humaines.

Un exemple, dû à Simon et Bonini est donné : débit de lingots, en millions de tonnes par an, au 1 - 1 - 54, de dix grands groupes sidérurgiques américains (U.S. Steel, Bethleem, etc.), classées par ordre décroissant : le produit du rang des aciéries par leur débit varie de 38,7 à 22,5, irrégulièrement, avec une légère tendance à la décroissance. La loi de Zipf de paramètre unité est donc une approximation un peu grossière.

II. — LOI DE YULE.

Pour l'auteur, la loi de Zipf est une approximation de la loi de Yule (parétienne au sens large), pour laquelle la probabilité de la valeur x est donnée par :

$$f(x) = K\,B\,(x, p+1) = \frac{K\,\Gamma\,(x)\,\Gamma\,(p+1)}{\Gamma\,(x+p+1)}$$

et pour p entier :

$$f(x) = \frac{K\,p\,!}{(x+p)\,(x+p-1)\,\ldots\,x}$$

où p est un paramètre $\geqslant 1$, k une constante et $x = 0, 1, 2, \ldots$

B est la fonction « beta » et Γ la fonction « gamma ».

L'auteur en donne plusieurs exemples. Le mode d'estimation n'est pas précisé. Le text du χ^2 est utilisé pour contrôler la qualité de l'ajustement.

1) Le premier volume d'une bibliographie de Statistique théorique et Probabilité, par M. G. Kendall et Miss Allison Doig, contient environ 9 000 références pour les années 1950 à 1958 incluses. L'auteur en extrait deux échantillons (lettres A et B d'une part, N à R d'autre part), auxquels il en adjoint un troisième (lettres N à R, années 1940 à 1949), et ajuste à la loi de Yule la distribution du nombre de références par revue scientifique. On obtient des quantités-tests pour lesquelles il y a respectivement des probabilités 0,33, 0,32 et 0,87 d'être dépassées par les variables du χ^2 correspondantes.

2) A la fin du second tome de "The Advanced Theory of Statistics", l'auteur donne 1 866 références sur 632 auteurs (livres et certains comptes rendus exclus). La distribution du nombre de références par auteur, ajustée à la loi de Yule, donne une quantité-test ayant la probabilité 0,06 d'être dépassée par une variable du χ^2 à 10 d.l. L'auteur le juge « surprenamment bon, considérant la nature du matériel ».

III. — LOI BINOMIALE NÉGATIVE.

Ehrenberg (1959) a obtenu d'importants échantillons de consommateurs privés des données sur leur quantité d'achats en certains produits vendus par quantités unitaires constantes (conserves de légumes, nourriture pour chats, café, savon, etc.). Il a montré que les distributions de nombre d'unités achetées par période de temps fixée par un consommateur donné s'ajustaient remarquablement à la loi binomiale négative. L'auteur applique un test du χ^2 à des données de Ehrenberg et trouve une probabilité 0,95. (L'analyse de l'article de Ehrenberg sera publiée ultérieurement dans cette rubrique.)

IV. — LOI DE POISSON.

Deux exemples sont donnés d'ajustements à la loi de Poisson.

1) Données de Quincy Wright (1942) ajustées à la loi de Poisson (méthode non précisée) par Richardson (1944) sur le nombre de déclenchements de guerres par an de 1500 à 1931. Le test du χ^2 donne une probabilité 0,26.

2) Analyse par l'auteur de la distribution du nombre de déclenchements de grèves par semaine au Royaume-Uni, durant les années 1948 à 1959 (arrêts de travail majeurs annoncés comme tels dans le *Ministry of Labour Gazette*). On identifie les moyennes théorique et empirique. Le test du χ^2 donne une probabilité 0,86 que la quantité-test soit excédée par la variable aléatoire correspondante.

La qualité de cet ajustement est jugée « presque inquiétante » par l'auteur. Cependant l'ajustement à la loi de Poisson est moins bon lorsque l'on considère la distribution des déclenchements de grèves dans un secteur particulier de l'industrie.

RÉFÉRENCES BIBLIOGRAPHIQUES

BRADFORD S. C.. — *Documentation*, London, Crosby Lockwood. 1948.

CHAMPERNOWNE D. G. — "A model of income distribution", *Econ. J.*, 63, p. 318-351. 1943.

EHRENBERG A. S. C. — "The pattern of consumer purchases", *Appl. Statistics*, 8, p. 26-41. 1959.

ELDERTON Sir, William P. — "Crickett scores and some shew correlation distributions (an arithmetical study)", *J. of the Royal Stat. Soc.* 108, p. 1-11. 1941.

HART P. E. et PHELPS BROWN E. H. — "The sizes of trade unions : a study of the laws of aggregation", *Econ. J.*, 67, p. 1-15. 1957.

HART P. E. et PRAIS S. J. — "The analysis of business concentration", *J. of the Royal Stat. Soc.*, A, 119, p. 150-180. 1956.

KENDALL D. G. — "On some modes of population growth leading to R. A. Fisher's logarithmic series distributions", *Biometrika*, 35, p. 6-15. 1948.

KENDALL M. G. — "The bibliography of operationel research", *Op. Research Quarterly*, 11, p. 31-36. 1960.

LOTKA A. J. — "The frequency distribution of scientific productivity", *J. Wash. Acad. Sci.*, 12, p. 317-323. 1926.

PENROSE L. S. — "Elementary statistics of majority voting », *J. of the Royal Stat. Soc.*, 109, p. 53-57. 1946.

PHILIPS H. — *Memorandum on football pools. Evidence before royal commission on betting, lotteries and gaming, 25th day*, London, H. M. Stationnery Office. 1950.

RICHARDSON L. F. — "The distributions of wars in time", *J. of the Royal Stat. Soc.*, 107, p. 242-250. 1944.

SIMON H. A. — "On a class of skew distribution fonctions", *Biometrika*, 42, p. 425-440. 1955.

SIMON H. A. et BONINI C. P. — *The size distribution of business firms*, Reprint 20 of the graduate school of Business Administration, Carnegie Institute of Technology, Pittsburg.

SNYDER C. — *Capitalism the creator*, N.Y., Macmillan. 1940.

WRIGHT Q. — *A study of war*, Chicago, University of Chicago Press. 1942.

YULE G. U. — "A mathematical theory of evolution based on the conclusions of doctor Willis", *F. R. S. Phil. Transactions*, A, 213, p. 21-87. 1924.

YULE G. U. et GRENNWOOD M. — "An inquiry into the nature of frequency distributions representatived multiple happenings", *J. of the Royal Stat. Soc.*, 83, p. 255-279. 1920.

ZIPF G. K. — *Human behaviour and the principle of least effort*, Addison, Wesley Press. 1949.

A suivre.