MATHÉMATIQUES ET SCIENCES HUMAINES

M. EYTAN

Quelques propriétés de la loi de Pareto et leurs incidences en sciences humaines (I)

Mathématiques et sciences humaines, tome 8 (1964), p. 21-26

<http://www.numdam.org/item?id=MSH_1964__8__21_0>

M. EYTAN

QUELQUES PROPRIETES DE LA LOI DE PARETO ET LEURS INCIDENCES EN SCIENCES HUMAINES (1)

Parmi les modèles mathématiques utilisés en Sciences Humaines, on fait très largement appel aux modèles probabilistes. Il s'agit souvent de phénomènes que l'on peut se représenter comme entièrement caractérisés par la donnée d'un nombre, valeur d'une certaine variable aléatoire X. Tout un chacun sait combien l'hypothèse selon laquelle X suit une loi "normale" (c'est-à-dire de Laplace-Gauss) est (ou du moins était jusqu'à ces dernières années) répandue.

Ce postulat est critiquable pour plusieurs raisons:

1) il semble arbitraire,
2) la loi de Laplace-Gauss, dont la fonction de répartition est donnée par

$$P\ (X < x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-m)^2}{2\sigma^2}}\, dt$$

est très fortement concentrée autour de la moyenne m, les valeurs de X qui s'en éloignent assez peu sont très peu probables. Par exemple on sait que

$$P\ (|X - m| \geqslant 5\sigma) < 6 \cdot 10^{-7}$$

Il est donc "probable" que X ne prendra "jamais" une valeur s'éloignant de la moyenne de plus de 5σ. Plus précisément on a une probabilité négligeable d'observer une valeur tant soit peu éloignée (c'est-à-dire de plus de 5σ) de m. Or il est d'observation courante de constater que certains phénomènes expérimentaux ne présentent pas cette caractéristique: par exemple la répartition des villes des Etats-Unis selon leur population.

Enfin la loi de Laplace-Gauss a une variance

$$E\ (X - m)^2 = \sigma^2 \ \underline{\text{finie}}.$$

Cette circonstance aussi peut être contredite par l'expérience, dans le sens suivant: lorsqu'on augmente la taille de l'échantillon, la variance expérimentale augmente de plus en plus, sans borne. Par exemple l'étude de certaines valeurs en bourse présente ce caractère de variance expérimentale croissante sans borne.

Qu'à cela ne tienne, dira-t-on: si les caractéristiques ci-dessus sont criti-

(1) Les propriétés exposées ci-dessous sont bien connues des probabilistes; leur utilisation en Sciences Humaines est due à B. Mandelbrot.

quables, choisissons une autre loi, et choisissons-la bien. La loi de Pareto par exemple: les grandes valeurs de la variable sont relativement fréquentes (du moins en comparaison avec la loi de Laplace-Gauss), et la variance est infinie, quoique la moyenne soit finie.

Contrairement à la loi de Laplace-Gauss, elle s'intéresse davantage aux probabilités des valeurs éloignées de l'origine. Sa fonction de répartition est donnée par:

$$P\ (X < x) = \begin{cases} 1 - Cx^{-\alpha} & \text{pour } x \geqslant x_0 > 0 \\ 0 & \text{pour } x < x_0 \end{cases}$$

où α est un paramètre (en général on impose la condition $1 < \alpha < 2$), C une constante > 0.

Sous cette forme, elle a été étudiée par Pareto en 1897 (1) pour représenter la distribution des revenus, le nombre d'individus ayant un revenu supérieur ou égal à x étant donné par $G(x) = Cx^{-\alpha}$.

Plus généralement on appellera loi asymptotiquement parétienne une loi dont la fonction de répartition H(x) vérifie:

$$\lim_{x\to\infty} \frac{H(x)}{F(x)} = 1, \quad \text{noté } \text{"}H(x) \sim F(x) \quad \text{quand } x \to \infty\text{"}$$

$$(\text{où } F(x) = \begin{cases} 1 - Cx^{-\alpha} & \text{pour } x \geqslant x_0 > 0) \\ 0 & \text{ailleurs} \end{cases}$$

Cependant, nous dira-t-on, pourquoi la loi de Pareto serait-elle digne d'intérêt de préférence à une autre, n'est-elle pas tout aussi arbitraire que la loi de Laplace-Gauss? Certes il y a d'autres raisons de s'intéresser à la loi de Pareto, des raisons "profondes" c'est-à-dire théoriques et non pas empiriques. Son choix n'est pas arbitraire, pas plus d'ailleurs que celui de la loi de Laplace-Gauss pour décrire certains phénomènes dans les Sciences Humaines.

Venons-en donc aux résultats de la théorie. Soient X_1 et X_2 des variables aléatoires <u>indépendantes</u> suivant la même loi $\mathcal{L}$.

On dit que <u>la loi $\mathcal{L}$ est stable</u> si, quels que soient les nombres $a_1 > 0$, b_1, $a_2 > 0$, b_2, il existe des nombres $a > 0$, b et une variable aléatoire X suivant la même loi $\mathcal{L}$ que X_1 et X_2, tels que:

$$a_1 X_1 + b_1 + a_2 X_2 + b_2 = a X + b.$$

Autrement dit, on ne sort pas du domaine de la loi $\mathcal{L}$ en y procédant à des additions de variables aléatoires indépendantes définies à l'origine et à l'échelle (positive) près.

Définissons maintenant une généralisation, en un certain sens, de la notion de loi stable. Pour cela soit donnée une suite $X_1, X_2, \ldots, X_n \ldots$ de variables aléatoires indépendantes suivant toutes la même loi $\mathcal{L}$.

(1) Pareto: Cours d'Economie Politique, Travaux de Droit, d'Economie et de Sociologie n° 26, nouvelle édition par Bousquet et Busino - Librairie Droz-Genève (1964).

Considérons la variable aléatoire:

$$S_n = \frac{X_1 + X_2 + \ldots + X_n}{\lambda_n} - \mu_n \qquad (\text{où } \lambda_n \longrightarrow \infty,\ \frac{\lambda_{n+1}}{\lambda_n} \rightarrow 1)$$

et soit $\mathcal{L}'_n$ sa loi. Si la suite $\mathcal{L}'_n$ converge (1) vers une loi $\mathcal{L}'$ quand $n \longrightarrow \infty$, on dit que $\mathcal{L}'$ est une loi-limite.

Théorème 1 (2): Pour que la loi $\mathcal{L}$ soit une loi-limite, il faut et il suffit qu'elle soit stable.

Il faut donc, si nous voulons déterminer la classe des lois-limites, chercher quelles sont les lois stables. Or on les connait toutes (3):

Théorème 2: les lois stables sont de la forme (4) $F(x;\ \alpha,\ \beta,\ \gamma,\ c)$ où $0 < \alpha \leqslant 2$, β, γ, c sont des paramètres. Elles ont les propriétés suivantes:

Pour $\alpha \neq 1$ et $\alpha \neq 2$ elles sont asymptotiquement parétiennes

$$1 - F(x;\ \alpha,\ \beta,\ \gamma,\ c) = P(X \geqslant x) \sim \frac{c}{\alpha}\ x^{-\alpha}$$

Pour $\alpha = 1$ ce sont des lois du type

$$P(X < x) = \frac{1}{\pi} \int_{-\infty}^{x} \frac{a}{a^2 + (t-b)^2}\ dt \qquad (a > 0)$$

dite loi de Cauchy, ne possédant ni moyenne ni variance finie.

Pour $\alpha = 2$ ce sont les lois de Laplace-Gauss

$$P(X < x) = \frac{1}{\sqrt{2\pi}.\sigma} \int_{-\infty}^{x} e^{-\frac{(t-m)^2}{2\sigma^2}}\ dt$$

Corollaire (5): la loi de Laplace-Gauss est la seule loi stable à variance finie non nulle.

Soit maintenant $X_1,\ X_2,\ \ldots,\ X_n,\ \ldots$ une suite de variables aléatoires indé-

(1) Il s'agit de convergence en loi c.a.d. de la convergence des fonctions de répartition aux points de continuité de la fonction de répartition limite.

(2) (SRV) p. 162.

(3) (AVA) p. 198; (SRV) § 34; (PDI) pp. 89-90).

(4) Leur fonction caractéristique f(t) est de la forme: $f(t) = \exp\left(i\gamma t - c|t|^{\alpha}\left(1 + i\beta \frac{t}{|t|}\,\omega(t,\ \alpha)\right)\right)$

où γ est réel (c'est la moyenne), $c \geqslant 0$, $|\beta| \leqslant 1$, $0 < \alpha \leqslant 2$, $\omega(t,\ \alpha) = \begin{cases} \operatorname{tg} \frac{\pi}{2}\alpha & \text{si } \alpha \neq 1 \\ \frac{2}{\pi} \log |t| & \text{si } \alpha = 1 \end{cases}$

(5) (AVA) p. 97.

pendantes suivant la même loi $\mathcal{L}$; intéressons nous à la valeur extrême (1)

$$M_n = \max (X_1, X_2, \ldots, X_n)$$

Théorème 3 (Fréchet): les lois-limites de la valeur extrême sont de la forme (à un changement d'origine et d'échelle près):

$$F(x; \alpha) = \begin{cases} e^{-x^{-\alpha}} & x \geqslant 0 \\ 0 & x < 0 \end{cases}$$

c'est-à-dire asymptotiquement parétiennes.

La suite $X_1, X_2, \ldots, X_n, \ldots$ de variables aléatoires suivant toujours la même loi $\mathcal{L}$, on dira que la loi $\mathcal{L}$ appartient au domaine d'attraction de la loi de Laplace-Gauss si la loi de

$$S_n = \frac{X_1 + X_2 + \ldots + X_n}{\lambda_n} - \mu_n$$

tend vers la loi de Laplace-Gauss lorsque $n \longrightarrow \infty$

Théorème 4 (2): pour que la loi $\mathcal{L}$ appartienne au domaine d'attraction de la loi de Laplace-Gauss centrée et réduite

$$P(X < x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

il faut et il suffit qu'elle ait une variance finie $\sigma^2 = E(X_i - M)^2 < \infty$, et alors nécessairement $\lambda_n = \sigma\sqrt{n}$, $\mu_n = nM$.

Définissant le domaine d'attraction d'une loi asymptotiquement parétienne de façon analogue, on a le résultat suivant (3):

Théorème 5: pour qu'une loi $\mathcal{L}$ appartienne au domaine d'attraction de la loi asymptotiquement parétienne, il faut et il suffit qu'elle soit elle-même asymptotiquement parétienne et alors $\lambda_n = a\, n^{1/\alpha}$ (μ_n quelconque).

Revenons maintenant aux modèles probabilistes, et supposons que l'on puisse considérer comme acquis, tout au moins en première approximation, que la variable aléatoire représentative du phénomène est une moyenne d'un très grand nombre de variables aléatoires indépendantes suivant la même loi. Une telle hypothèse peut être valable par exemple dans la théorie des erreurs d'observation ou dans celle

(1) (VE), (SE)
(2) (AVA) p. 110; (SRV) p. 181.
(3) (SRV) p. 181 et 182..

du revenu d'un individu (beaucoup de facteurs contribuant au revenu qui est considéré comme leur moyenne), ou d'un groupe d'individus. Nous pouvons supposer alors qu'en fait on observe la variable "apparente" S_n résultant des n variables "latentes" X_i par la formule:

$$S_n = \frac{X_1 + X_2 + \dots + X_n}{\lambda_n} - \mu_n$$

Pour qu'il y ait une certaine régularité, il faut que la loi de S_n tende vers une limite (sinon nos observations oscilleraient de façon erratique avec n). Mais d'après le théorème 2 une telle limite ne peut être qu'une loi stable c'est-à-dire ou bien gaussienne (variance et moyenne finies), ou bien de Cauchy (ni variance ni moyenne finies) ou bien asymptotiquement parétienne. Eliminant le cas particulier de la loi de Cauchy, nous voyons donc que nous sommes conduits à postuler que le modèle doit fournir des observations suivant soit une loi gaussienne (variance finie), soit une loi asymptotiquement parétienne (variance non finie). Mais dans le cas de variance non finie, le théorème 5 nous indique que si la loi limite est asymptotiquement parétienne, la loi des X_i l'est aussi.

Ainsi donc, à condition de bien pondérer les moyennes que l'on étudie, un phénomène pourra être postulé suivre la loi de Laplace-Gauss si la variance est finie, la loi asymptotiquement parétienne sinon; dans ce dernier cas ce ne sont pas seulement les moyennes mais encore les constituants primitifs qui seront asymptotiquement parétiens. Tout cela sous réserve que l'on puisse postuler une <u>aggrégation se traduisant par une moyenne pondérée</u> (et l'<u>indépendance</u> des constituants).

Mais il y a encore mieux: supposons que l'on procède à une opération de maximisation sur les variables aléatoires du modèle (par exemple pour les revenus un individu pourra se comporter comme s'il avait le loisir de remplir plusieurs fonctions, et il choisit toujours celle qui lui rapporte le plus). Alors d'après le théorème 3 la loi-limite ne pourra être qu'asymptotiquement parétienne.

Remarquons en passant que lorsqu'on observe un phénomène aussi mal défini que ceux qu'il est usuel d'observer en Sciences Humaines, la stabilité est indispensable pour une autre raison encore: on ne sait jamais si l'on traite le phénomène global (par exemple revenu total) ou seulement une partie isolée (par exemple des parties isolables du revenu) du phénomène. Il doit donc exister une invariance par addition pour que les diverses définitions possibles ne donnent pas lieu à des lois différentes (1).

Disons enfin quelques mots sur les comportements différents suivant qu'une variable est gaussienne ou asymptotiquement parétienne. On peut montrer (2) que dans le cas de la loi de Laplace-Gauss

$$\frac{M_n}{S_n} \equiv \frac{\max(X_1, X_2, \dots, X_n)}{S_n} \longrightarrow 0 \text{ quand } n \longrightarrow \infty$$

et réciproquement. Donc dans le cas asymptotiquement parétien le rapport $\frac{M_n}{S_n}$ ne tend

(1) (PRF) p. 525.
(2) (PRF) p. 523.

pas vers 0 : si a priori les X_i contribuent de façon identique à S_n (et pour une part négligeable à lim S_n), il n'en est pas de même a postériori, car alors le plus grand des X_i n'est plus négligeable, l'emportant sur les autres très largement.

BIBLIOGRAPHIE

(SRV) Gnedenko & Kolmogorof: Limit Distributions for Sums of Independent Random Variables, Addison-Wesley, 1954.

(AVA) P. Levy: Théorie de l'addition des variables aléatoires, Gauthier-Villars, 1954.

(PDI) B. Mandelbrot: In International Economic Review, Vol. I, n°2, May 1960.

(PRF) B. Mandelbrot: In Econometrica, Vol. 29, n°4, (October 1961).

(PIM) B. Mandelbrot: In the Quaterly Journal of Economics, Vol. LXXVI, February 1962.

(VE) Geffroy: Contributions à la théorie des valeurs extrêmes, Publ. ISUP, 1958.

(SE) Gumbel: Statistics of Extremes, Columbia Univ. Press, 1958.

- - - - - - - - - - - - - -