

Projet de recherche sur :

La loi de Pareto et ses applications en SHS

Réponse à l'AAP Agorantic

« Culture, Patrimoine, Sociétés numériques »

Année 2015

Porteurs : Delphine Blanke, Philippe Martin

Chercheurs : Luc Dinh The, Daniel Gourion, François Moriconi-Ebrard, Mounir Redjimi.

Laboratoires :

- Laboratoire de mathématique EA 2151,
- UMR ESPACE 7300 du CNRS,
- IUT de statistique.

Wilfried Pareto est connu pour son cours d'économie tout autant que pour sa loi des 80 – 20 qui traduit, très didactiquement, l'idée d'une forte inégalité : ainsi 80 % des possédants disposent-ils de 20 % des richesses mondiales alors que 20 % des plus riches détiennent eux, 80 % des avoirs planétaires, mais 1 % des plus fortunés possède plus de 50 % des richesses, etc. C'est donc bien d'une hiérarchie dont rendent compte ces observations empiriques.

Ce type de hiérarchie se rencontre non seulement en économie (niveau de capital, revenus, PNB par habitant, etc.), mais aussi dans de très nombreux autres domaines, comme en particulier dans le fonctionnement du web (réseaux, flux, etc.) et en géographie (réseaux urbains, incendies de forêt, etc.).

Plus techniquement les séries statistiques représentant ces fortes hiérarchies peuvent être traduites empiriquement dans un graphique bilogarithmique - qui croise une série expérimentale et une série théorique (le rang, une fréquence, etc.) - par un alignement de points, auquel peut être ajusté un modèle de puissance dont l'exposant, dit de Pareto, est le seul paramètre statistique. Ce sont ainsi des lois qui n'ont ni moyenne ni écart-type si l'exposant de Pareto est inférieur ou égal à 1. En géographie, ce type d'approche (sous une forme discrète) est connu sous le nom de loi de Zipf, laquelle a été réinterprétée par B. Mandelbrot comme étant l'expression d'une fractalité, d'une organisation en échelle.

Il apparaît toutefois expérimentalement que cette modélisation à un paramètre n'est pas suffisante, et cela même si on envisage la loi de Pareto comme une loi L-stable. Très souvent lorsque les catalogues de données sont importants, une courbure apparaît sur la modélisation

effectuée, ce qui traduit un défaut de modélisation avec la loi de Pareto en exhibant une atténuation de la hiérarchie attendue

L'ubiquité de ces structures statistiques « contraintes » que l'on retrouve sur des données tant informationnelles que matérielles a conduit les statisticiens, en particulier œuvrant pour des assureurs (problème du coût des catastrophes), à ne considérer que le point de vue asymptotique, solution partielle et partielle pouvant être discutable pour l'interprétation des résultats obtenus.

Il convient donc de reprendre la tentative qu'avait faite B. Mandelbrot de modéliser cette « courbure » à partir d'exemples littéraires (dénombrement de mots) et en s'appuyant sur la théorie de l'information de Shannon, afin :

-1- de disposer d'une modélisation statistique robuste dont les paramètres seront intrinsèquement compris et explicités ;

et -2- d'accéder à une compréhension de ces structures ubiquistes qui ne soit pas contrainte par le type de substrats ou de milieux dans lesquels elles se manifestent.

Ces questions de construction formelle d'un modèle parétien statistique « contraint » et d'explicitation quel que soit le domaine des paramètres du modèle, restent encore aujourd'hui largement ouvertes malgré leur place fondamentale dans la compréhension de certains phénomènes ou fonctionnements.

Pour essayer d'aller plus loin, nous avons réuni une équipe de mathématiciens, de statisticiens et de géographes de l'université d'Avignon permettant de traiter tant les aspects les plus formels que les questions les plus statistiques, à partir de données dont la géographie dispose, mais pas seulement.

Cette association entre disciplines formelles et empiriques est relativement originale et s'inscrit dans un mouvement plus vaste de décloisonnement comme le projet Math in Terre a pu y encourager.

Dans la mesure où trois disciplines au moins (et deux laboratoires, deux UFR et l'IUT de l'UAPV) seront associées pour aborder des questions de structures, de territoire et de méthodologie, notre projet ne peut s'inscrire que transversalement dans les axes de recherche d'Agorantic.

À l'issue de cette phase, si comme nous le suspectons, la question posée est profonde, nous envisageons de la développer, au travers de partenariats extérieurs (sciences formelles et SHS) et d'une réponse à un AAP national.

Budget 6000 euros dont :

2000 euros pour de la documentation (articles, livres, base de données...) et un soutien informatique et 4000 euros pour le fonctionnement du groupe, pour des invitations des missions (colloques...).

En fonction du financement, nous prévoyons d'organiser au moins un séminaire avec des intervenants extérieurs. Par ailleurs nous envisageons, en fin de période, quelques actions de communication (participation à des colloques, des publications, etc.).

Financement Agorantic demandé : 6000 euros

Éléments bibliographiques

- Adler R. J., Feldman R. E., Taqqu M. S. (Dir), 1998, *A practical guide to heavy tails. Statistical techniques and applications*. Birkhäuser, Boston, 533 p.
- Feuerverger A., Hall P., 1999, Estimating a tail exponent by modelling departure from a Pareto distribution. *The annals of statistics*, vol.27, n°2, p.760-781.
- Forriez M., Martin Ph., 2009, Structures hiérarchiques en géographie : des modèles linéaires aux modèles non linéaires (lois de puissance et corrections log-périodiques), in Foltête, J.-Ch. (Dir.), 8° Rencontre de ThéoQuant, Besançon, THEMA, 10 p. <http://thema.univ-fcomte.fr/theoq/pdf/2007/TQ2007%20ARTICLE%2051.pdf>
- Gan Li, Dong Li, Song S, 2006, Is the zipf law spurious in explaining city-size distributions? *Economics letters* 92, p.256-262.
- Luckstead J., Devadoss S., 2014, Do the world's largest cities follow Zipf's and Gibrat's laws? *Economics letters* 125, p.182-186.
- Mahanti A, et al., 2013, A tale of the tails: Power-laws in internet measurements, *Network, IEEE*, Vol.27, Issue 1, p.59-64.
- Martin Ph., 2013, Statistique dynamique d'une série parétienne d'observations approche méthodologique. 11° Rencontre de ThéoQuant, Besançon, livre des résumés, p.209-210. <http://thema.univ-fcomte.fr/theoq/pdf/resumes/TQ2013%20RESUMES.pdf>
- Martin Ph., Redjimi M., 2014, Évolution du réseau de peuplement algérien depuis l'Indépendance Approches parétienne et fractale parabolique. *Aux frontières de l'urbain : Petites villes du monde*, Collection : Actes Avignon - ISBN : 978-2-9105-4509-1, p.374-395.
- Mandelbrot B., 1997, *Fractales, hasard et finance*. Coll.: Champ, Flammarion éditeur, Paris, 246 p.
- Newman M.E.J., Power laws, Pareto distributions and Zipf's law. arXiv:cond-mat/041204v3 [cond-mat.stat-mech] 29 May 2006, 28 p.
- Pumain D., 2006, (Dir), *Hierarchy in natural and social sciences*. Methodos séries, vol.3, Springer éditeur, Berlin, 243 p.
- Saichev A., Malevergne Y, Sornette D., 2010, *Theory of Zipf's law and beyond*. Springer Verlag, Berlin, 169 p.
- Zajdwenweber D., 2009, *Économie des extrêmes. Krachs, catastrophe et inégalité*. Coll. Champs essais, Flammarion, Paris, 236 p.
- Zipf G.K., 1949, *Human behaviour and principle of least effort*, Cambridge, MA: Addison-Wesley.

