MATHÉMATIQUES ET SCIENCES HUMAINES

M. PETRUSZEWYCZ

Loi de Pareto ou loi log-normale : un choix difficile

Mathématiques et sciences humaines, tome 39 (1972), p. 37-52

<http://www.numdam.org/item?id=MSH_1972__39__37_0>

Math. Sci. hum. (10e année, n° 39, 1972, p. 37-52)

LOI DE PARETO OU LOI LOG-NORMALE : UN CHOIX DIFFICILE

par

M. PETRUSZEWYCZ [1]

INTRODUCTION

Présentant [2] un article de R. E. Quandt: "Statistical discrimination among alternative hypotheses and some economic regularities" (*Journal of regional science*, vol. 5, n° 2, 1964, pp. 1-23), B. Leclerc [4] écrivait: « De nombreuses distributions de grandeurs socio-économiques positives possèdent une propriété commune: très dissymétriques, elles s'étirent vers la droite. C'est par exemple le cas de la distribution des revenus dans une région donnée à une époque fixée, ou des tailles des villes dans une aire géographique donnée. Ces distributions ont été ajustées à diverses lois, les plus courantes étant les lois de Pareto et la loi log-normale. »

A une certaine époque, on a même fait une querelle de la confrontation de ces deux lois, querelle sans conclusion à ce jour, mais tombée dans l'oubli, sans doute parce que la loi de Pareto paraît elle-même oubliée. On en citera pour preuve le fait que des publications très averties comme *Études et conjoncture* n'offrent à leurs lecteurs qu'un graphique d'ajustement log-normal pour rendre compte de la distribution des revenus imposés en France en 1962, par exemple.

Sans vouloir ranimer une querelle stérile, cet article se propose de réagir contre cet état de choses et de montrer qu'il est parfois difficile, au vu d'une distribution, de décider de la loi qui permettrait d'en rendre compte au mieux. Quandt faisait « une tentative pour trouver un instrument statistique de discrimination entre des hypothèses concurrentes ». Cet auteur traitait, au moyen d'instruments statistiques classiques ou originaux, quatre séries de données concrètes confrontées à neuf lois de distribution parmi lesquelles la loi de Pareto et la loi log-normale; il faut signaler que la loi de Pareto à trois paramètres remportait l'avantage. Le propos de cet article, beaucoup plus modeste (on ne considérera que ces deux derniers types de distribution) est cependant légèrement différent car le problème sera abordé, dans une première partie, sur un plan théorique, les ajustements à des données concrètes n'étant étudiés qu'en deuxième partie.

I. UN EXERCICE D'AJUSTEMENTS A PARTIR DE DONNÉES ARTIFICIELLES

On va successivement se donner une distribution selon la loi de Pareto, puis une distribution selon la

1. Centre de Mathématique Sociale, EPHE.

2. Je remercie M. G. Th. Guilbaud pour m'avoir suggéré d'entreprendre cette étude et Mme Carcassonne pour les conseils qu'elle m'a dispensés.

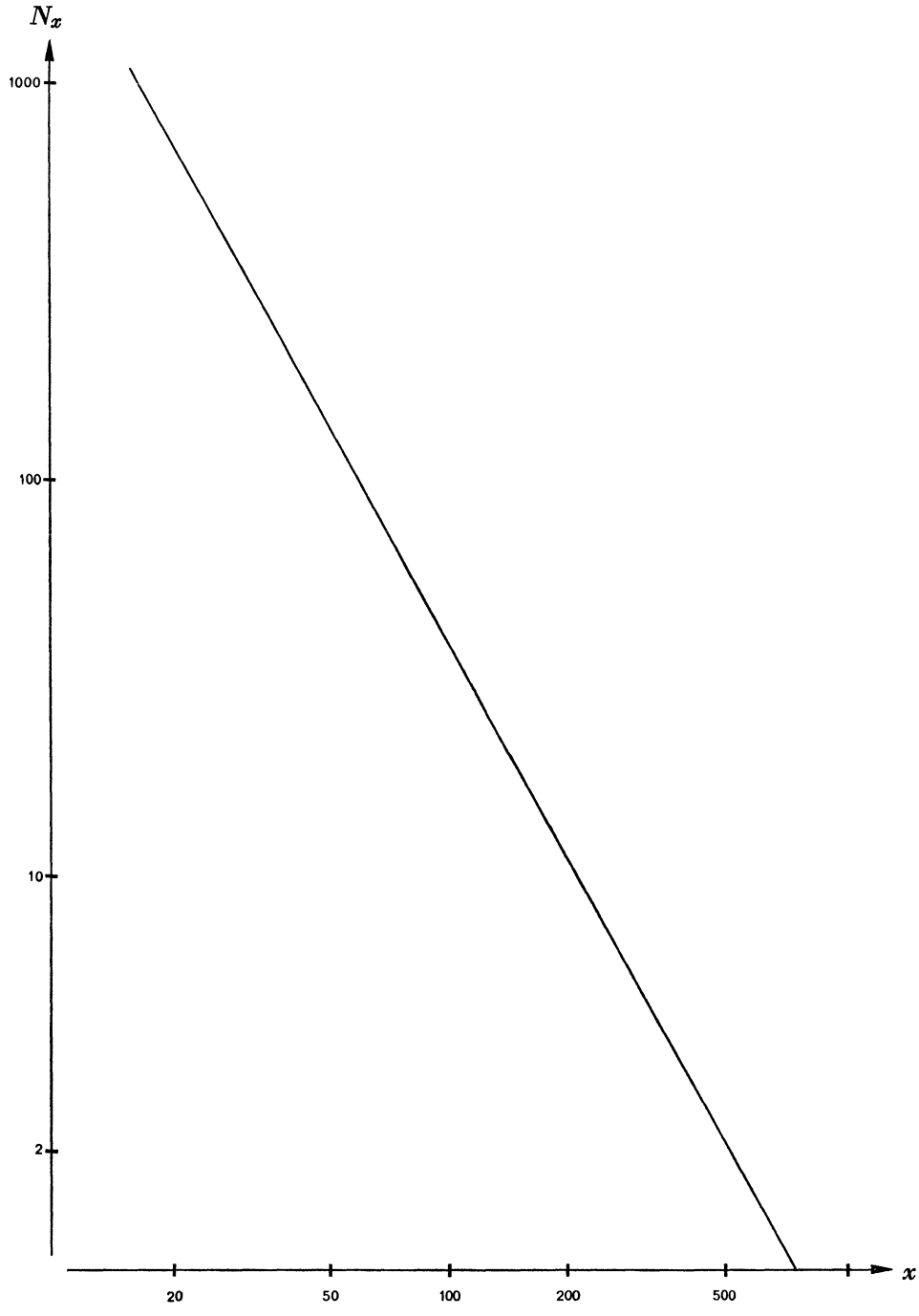

Figure 1

Distribution théorique parétienne Log $N_x = -1{,}8$ Log x

loi log-normale et essayer d'ajuster les données ainsi obtenues au moyen de l'autre loi. On rappellera d'abord que la loi de Pareto a pour fonction cumulative:

$$F(x) = 1 - \left(\frac{x_o}{x}\right)^{\alpha} \quad \text{si} \quad x > x_o$$

et que sa densité est:

$$f(x) = \frac{\alpha}{x_0}\left(\frac{x_0}{x}\right)^{\alpha+1}.$$

Elle possède donc deux paramètres x_0 et α, dit coefficient de Pareto. Pour la loi log-normale, on a:

— la fonction cumulative:

$$F(x) = \pi\left(\frac{\text{Log } x - m}{\sigma}\right)$$

— et la densité:

$$f(x) = \frac{1}{\sigma x \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\text{Log } x - m}{\sigma}\right)^2\right].$$

Elle a pour paramètres m et σ. Il ne faut pas confondre σ avec l'écart-type de la variable x qui s'exprime, en fonction de ces deux paramètres par la formule:

$$(\text{variance})^{1/2} = e^{m+\sigma^2} \times (1 - e^{-\sigma^2})^{1/2}.$$

1.1. Ajustement par une loi log-normale de données parétiennes

Sur du papier fonctionnel bi-logarithmique, on trace une droite de pente $\alpha = -1{,}8$, valeur du paramètre fréquente dans la littérature. Cette droite représente la courbe cumulative d'une loi de Pareto; son équation est:

$$\text{Log } N_x = -1{,}8 \text{ Log } x$$

où N_x représente l'effectif des individus recevant un revenu supérieur ou égal à x (le caractère distribué ou taille) (Fig. 1). Par lecture directe sur le graphique on relève, en face des effectifs en ordonnées, les valeurs correspondantes de x: ce sont les données de base, présentées au Tableau 1.

Tableau 1
Données des figures 1 et 2

N_x	%	x	T_3	T_5
1 000	100	16,95	0,45	0,001
950	95	17,40	0,90	0,45
900	90	17,90	1,30	0,95
500	50	24,80	8,30	7,85
100	10	60,10	43,60	43,15
50	5	80		73,05
10	1	220	203,50	203,05
1	0,1	790	773,50	773,05

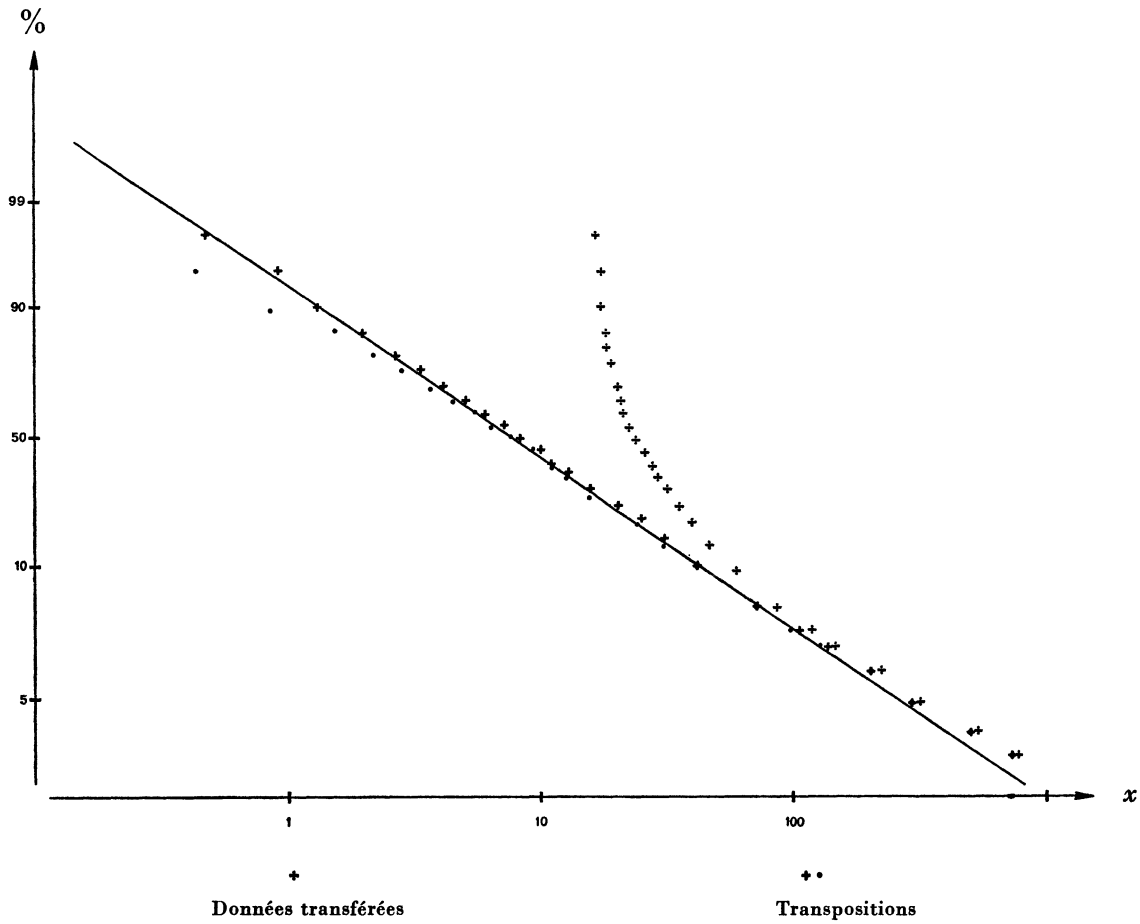

Figure 2

Ajustement Log-normal des données théoriques parétiennes avec $\alpha = -1{,}8$
(première méthode)

L'interprétation est la suivante: 1 000 individus ont un revenu supérieur ou égal à 16,95; 500 individus ont un revenu supérieur ou égal à 24,80; 5 individus ont un revenu supérieur ou égal à 90, etc. Ces deux premières colonnes du Tableau 1 vont permettre d'établir un deuxième graphique. En effet, si cette distribution était présentée à un statisticien comme ayant une origine empirique, il pourrait être tenté de l'ajuster au moyen d'une loi log-normale. La courbe cumulative d'une loi log-normale a pour image une droite en coordonnées gausso-logarithmiques dite *droite de Henri.* Conservant les graduations logarithmiques en abcisse (le caractère distribué), on transforme les effectifs correspondants en pourcentage de l'effectif total (Tableau 1, colonne 2) pour les porter en ordonnées (Fig. 2) lorsque celles-ci ont un sens: on ne peut par exemple porter la valeur 100% qui n'existe plus. La courbe initiale est fortement incurvée vers la droite. Il faut ici faire remarquer que, traditionnellement, les courbes cumulatives log-normales n'ont pas une pente négative mais positive, parce que l'on considère les unités statistiques pour lesquelles le caractère distribué est inférieur ou égal à un certain niveau. On a choisi d'adopter la courbe cumulative complémentaire par souci d'homogénéité avec les courbes cumulatives parétiennes.

Telle qu'elle est, cette courbe cumulative n'est de toute évidence pas une droite de Henri. On va donc considérer une nouvelle variable dite log-normale généralisée en recherchant un changement d'origine qui, par le jeu même de la transformation logarithmique, « redresse » la courbe cumulative.

On ne s'intéresse plus au logarithme de x mais au logarithme de $(x - x_0)$. L'équation de la droite de Henri n'est plus $u_{F(x)} = \frac{\log x - m}{\sigma}$, mais:

$$u_{F(x)} = \frac{\log (x - x_0) - m}{\sigma}.$$

Ce faisant, on tente un ajustement des données par une loi log-normale généralisée à trois paramètres: x_0, m et σ. Comment déterminer le paramètre x_0 ? Sa valeur maximum est celle qui rendrait négatif le nombre dont on prend le logarithme et même, pour ne pas nous embarrasser de valeurs nulles, on prendra pour x_0 la valeur maximum 16,949. On a essayé plusieurs transpositions dont certaines sont données au Tableau 1, deux des courbes cumulatives correspondantes étant présentées sur la Figure 2. Si la courbe initiale T était incurvée vers la droite, on voit que la transposée T_5 est au contraire, incurvée vers la gauche: il existe donc un point d'inflexion. La plus longue tangente en ce point à la courbe sera la meilleure droite de Henri que l'on puisse obtenir. C'est pour la transposée T_3, c'est-à-dire pour la variable $(x - 16{,}5)$ qu'on peut choisir graphiquement cette plus longue tangente.

Une remarque à propos de la valeur $x_0 = 16{,}5$: ce paramètre est une estimation du revenu minimum quand on traite des données de distribution de revenus. Cette estimation peut être très grossière en raison du peu d'influence qu'elle a sur celle des paramètres fondamentaux m et surtout σ qui peut être considéré comme un indicateur d'inégalité. Cette valeur critique est souvent connue en dehors de toute considération statistique, par exemple, lorsqu'il s'agit de revenus imposables. Ici on peut constater directement sur la Figure 1 que ce revenu minimum est réellement de 16,9.

Pour T_3, la courbe cumulative relative au pourcentage de la population compris entre 10% et 98% se confond avec la droite de Henri tracée à travers le nuage de points. Que peut-on en déduire ?

1.1.1. Si on prend pour référence Aitchison et Brown [1], on constate que la Figure 41, p. 33, est sensiblement comparable à la Figure 2 présentée ici: l'ajustement est graphiquement satisfaisant entre 20% et 95% de la population. Par ailleurs, ces auteurs indiquent les transpositions comme une procédure pour estimer x_0 et un moyen licite de recherche de la droite de Henri. Ils qualifient la procédure comme étant « plus un art qu'une science», mais l'admettent en première analyse, à condition qu'elle soit appliquée à plusieurs échantillons. (On ne présente ici qu'un seul de la dizaine d'essais effectués en variant les critères de sélection des points de la droite de Pareto transférés sur le papier log-normal.)

Si on tronque la distribution à 10% et 90%, le critère IV d'Aitchison et Brown (dans quelle mesure la fonction ajustée approxime les données) est parfaitement satisfait et on peut estimer que l'ajustement est licite en ce qui concerne le test graphique. Cela signifie, toujours d'après ces auteurs, que l'hypothèse d'une loi log-normale est assez solide pour être testée au moyen par exemple, d'un χ^2. En effet, on est en présence d'une loi à trois paramètres et on ne peut plus utiliser les tests propres à la loi log-normale à deux paramètres proposés et tabulés par Geary et Pearson. Par ailleurs, on ne peut estimer graphiquement m et σ, mais il faut utiliser les estimateurs tabulés par Hald, et pour les moments recourir aux tables publiées par Pearson et Lee. Les références à tous ces travaux figurent dans [1].

1.1.2. Si on se réfère à l'ouvrage de G. Galot [2], on quitte le domaine de l'art pour celui de la technique statistique. L'auteur conseille de procéder comme si on ignorait la valeur de x_0. On trace la courbe cumulative puis on cherche « par tâtonnements une droite [...] telle que l'écart horizontal entre courbe cumulative et droite soit approximativement constant lorsqu'il est mesuré en x. On doit alors vérifier que la droite verticale d'abcisse x_0 est approximativement asymptote» (Fig. 3).

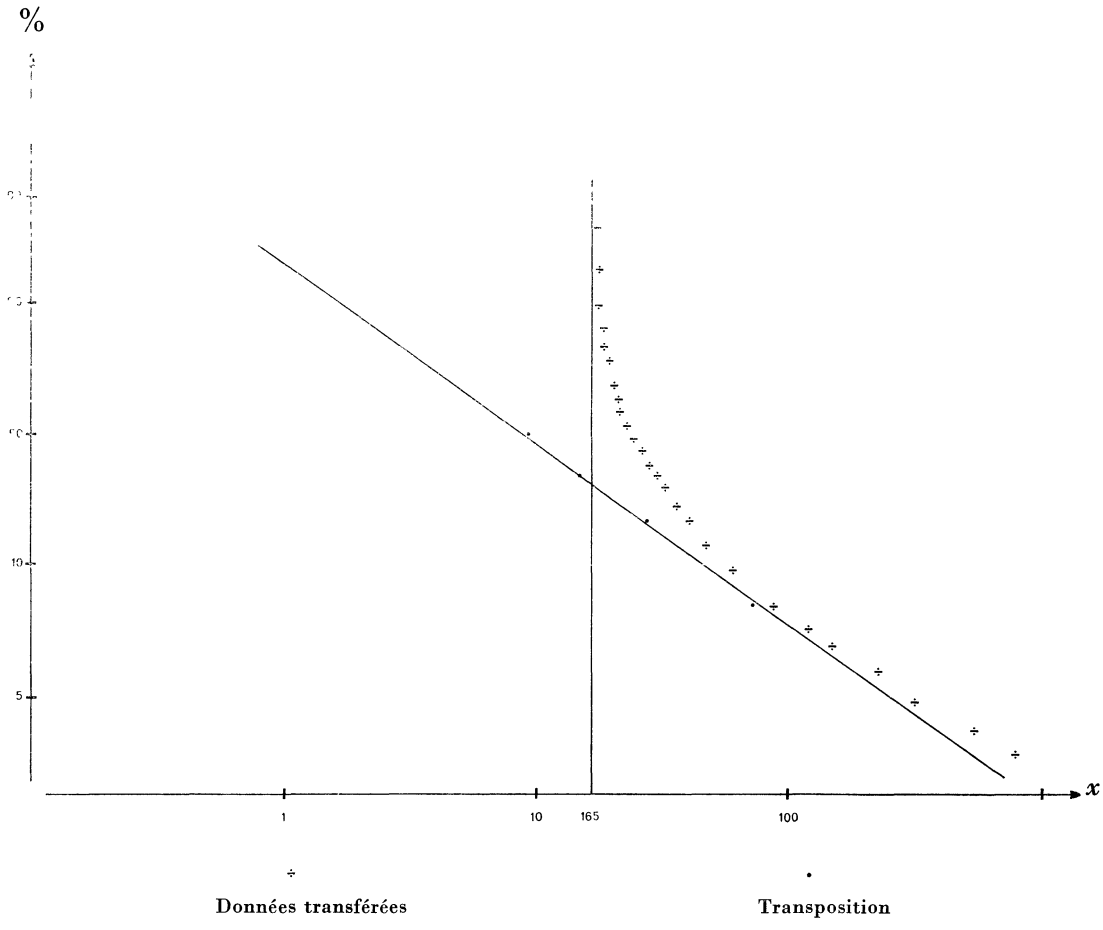

Figure 3

Ajustement log-normal des données théoriques parétiennes (deuxième méthode)

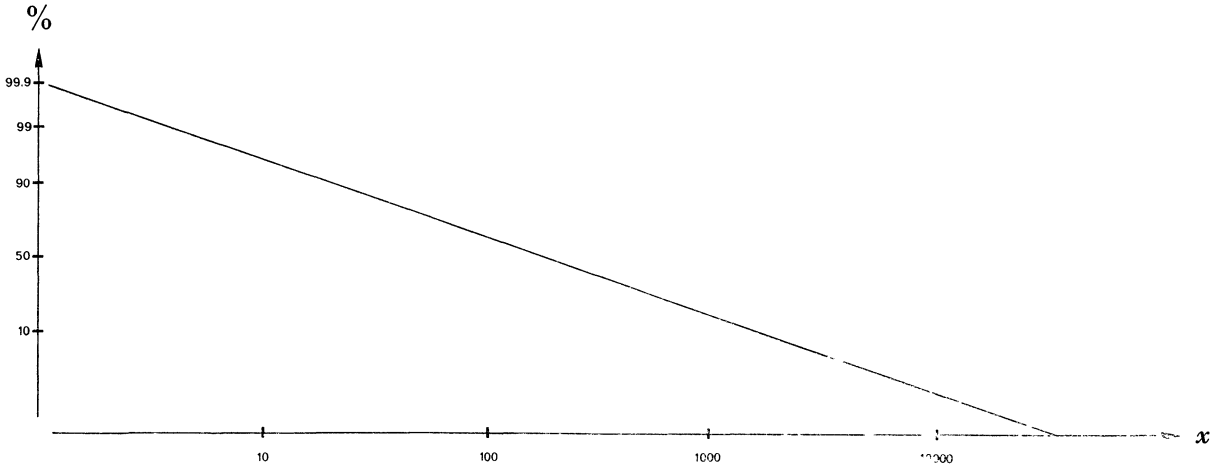

Figure 4

Distribution théorique log-normale ($m = 174$; $\sigma = 2{,}58$)

On aboutit par cette procédure, à une droite de Henri de pente légèrement plus relevée que celle de la droite choisie par la première méthode. On obtient la même valeur pour $x_0 = 16{,}5$ et la droite d'abcisse x_0 est effectivement asymptotique à la branche verticale de la courbe cumulative (branche verticale parce qu'on a cumulé par rapport aux valeurs supérieures ou égales au niveau considéré du caractère distribué; chez l'auteur qui cumule dans l'autre sens, il s'agit de la branche horizontale). Le fait d'avoir relevé la pente de la droite de Henri amène à tronquer la distribution un peu plus bas que dans le cas précédent, à 90%. Mais on peut ici aussi admettre, après troncature, le caractère licite de l'ajustement log-normal.

1.2. Ajustement par une loi de Pareto de données log-normales

1.2.1. *Distribution log-normale de paramètres* $m = 1{,}74$ $\sigma = 2{,}58$.

La distribution est représentée par sa courbe cumulative ou droite de Henri sur papier gausso-logarithmique (Fig. 4). Par lecture directe sur le graphique, on a retenu les données figurant au Tableau 2. Celles-ci sont alors reportées sur du papier fonctionnel bi-logarithmique où la courbe cumulative d'une distribution de Pareto est une droite de pente négative dont l'équation est Log $N_x = -\alpha$ Log x.

Tableau 2

Données des figures 3 et 4

%	N_x	x	$x + 100$	$x + 200$
99	1 000 990	3,9	103,9	203,9
90	900	26	126	226
54	540	148	248	348
8	80	1 760	1 860	1 960
0,1	1	27 500	27 600	27 700

On n'obtient pas une droite avec les données initiales (Fig. 5), mais on peut, là aussi, essayer de « redresser » en considérant non plus la variable x, mais une variable $(x + x_0)$. Lorsqu'on fait cela, on ajuste les données au moyen d'une loi de Pareto à trois paramètres: x, x_0 et α. En prenant pour $x_0 = 200$, on obtient un alignement des premières valeurs et on trace la droite d'ajustement de pente $\alpha = -1{,}16$; celle-ci se sépare très nettement des données représentant à peu près les 2,5% de la population pour lesquels les valeurs de x sont les plus élevées. Les différences constatées entre valeurs « théoriques » et valeurs « observées » correspondantes, si ces dernières étaient d'origine empirique, pourraient amener à

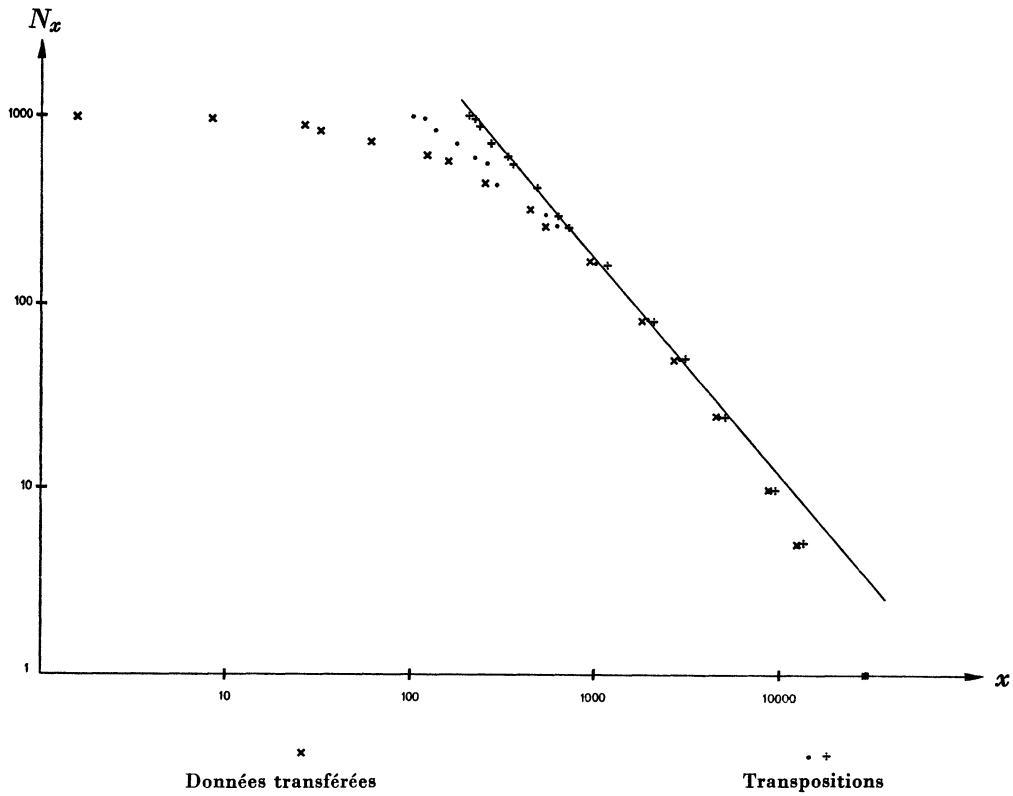

Figure 5

Ajustement parétien des données théoriques log-normales

effectuer un test de χ^2 pour cet ajustement ; la valeur obtenue 6,75 (avec 7 degrés de liberté au seuil 5%) ne permettrait pas de rejeter l'hypothèse que l'on soit en présence d'une loi de Pareto, à trois paramètres, ceci sans troncature.

On a pu, dans chaque cas, ajuster les données issues d'une des deux lois étudiées au moyen d'une droite d'ajustement figurant l'autre loi. Il faut cependant noter une dissymétrie: si une distribution de Pareto à deux paramètres est graphiquement ajustée de façon très satisfaisante, après troncature, par une loi log-normale à trois paramètres, inversement une distribution log-normale à deux paramètres n'est que médiocrement ajustée au moyen d'une loi de Pareto à trois paramètres mais sans troncature et sans que l'on soit cependant en mesure, si l'on était en droit d'appliquer les tests statistiques habituels, de rejeter l'hypothèse.

Par ailleurs, on sait l'importance qu'a la valeur 2 pour le coefficient α quand on rattache la loi de Pareto aux types des lois stables de P. Lévy (*cf.* [5]), et il ressort de la formule du moment non centré d'ordre r qu'α ne peut être inférieur à 1, pour que la distribution ait une moyenne. On est donc limité dans son choix lorsqu'on se « donne » une distribution théorique parétienne. Il n'en va pas de même dans l'autre sens où l'on peut faire varier les deux paramètres autant que l'on voudra (si l'on en fixe un, faire varier l'autre revient à se donner des étendues différentes); on doit veiller à se donner des distributions convenables pour que, après transfert et transposition on obtienne une *droite de Pareto* acceptable.

Une dernière remarque: α et σ étant tous deux des indicateurs d'inégalité, on pouvait songer à les comparer par le biais d'un autre indicateur d'inégalité. En effet, il existe une relation bien attestée entre α et la mesure de Lorenz: $L = \dfrac{1}{2\,\alpha - 1}$ et d'autre part Aitchison et Brown ont donné pour des valeurs de σ cette mesure de Lorenz (Table A.1, p. 154 [1]). Cependant, cette mesure de Lorenz est un mauvais critère d'ajustement, puisque deux distributions très différentes relevant de la même loi,

peuvent avoir même mesure de Lorenz. Aussi, est-ce intentionnellement que sont présentées une distribution de Pareto et une distribution log-normale de concentrations très différentes.

II. DONNÉES EMPIRIQUES

On va présenter dans ce paragraphe, deux exemples de données concrètes et chacune des séries sera ajustée successivement au moyen d'une loi de Pareto et d'une loi log-normale. Ceci pour souligner qu'il est regrettable de ne pratiquer systématiquement qu'un seul de ces ajustements: que ce soit parce qu'on se laisse tenter par le degré de commodité de calculs ultérieurs ou même lorsque le cadre conceptuel paraît fournir des raisons valables de penser immédiatement à l'une plutôt qu'à l'autre des distributions comme apportant « des informations sur la structure du processus qui conduit à l'obtention des observations» ([2], p. 179).

2.1. Fréquence des mots dans le russe d'aujourd'hui

On a utilisé la « liste générale des mots d'après la fréquence de leur emploi» du *Dictionnaire des fréquences de mots dans le russe d'aujourd'hui*, de E. Steinfeldt, Moscou, Ed. du Progrès, dont on présente un résumé dans le Tableau 4. Les précautions prises par l'auteur dans la constitution de son corpus (400 000 mots relevés dans 350 textes d'œuvres littéraires, théâtrales, radiophoniques et journalistiques ; ont été retenus les 2 520 mots les plus usuels — la fréquence minimum étant de 14, maximum 14 576), permettent de considérer cette liste comme un échantillon du russe parlé et écrit contemporain. (Tableau 3.)

Tableau 3
Fréquence des mots dans le russe d'aujourd'hui

Rang	Fréquence
1	14 576
2	11 804
5	6 067
100	445
1 000	51
2 000	22
2 520	14

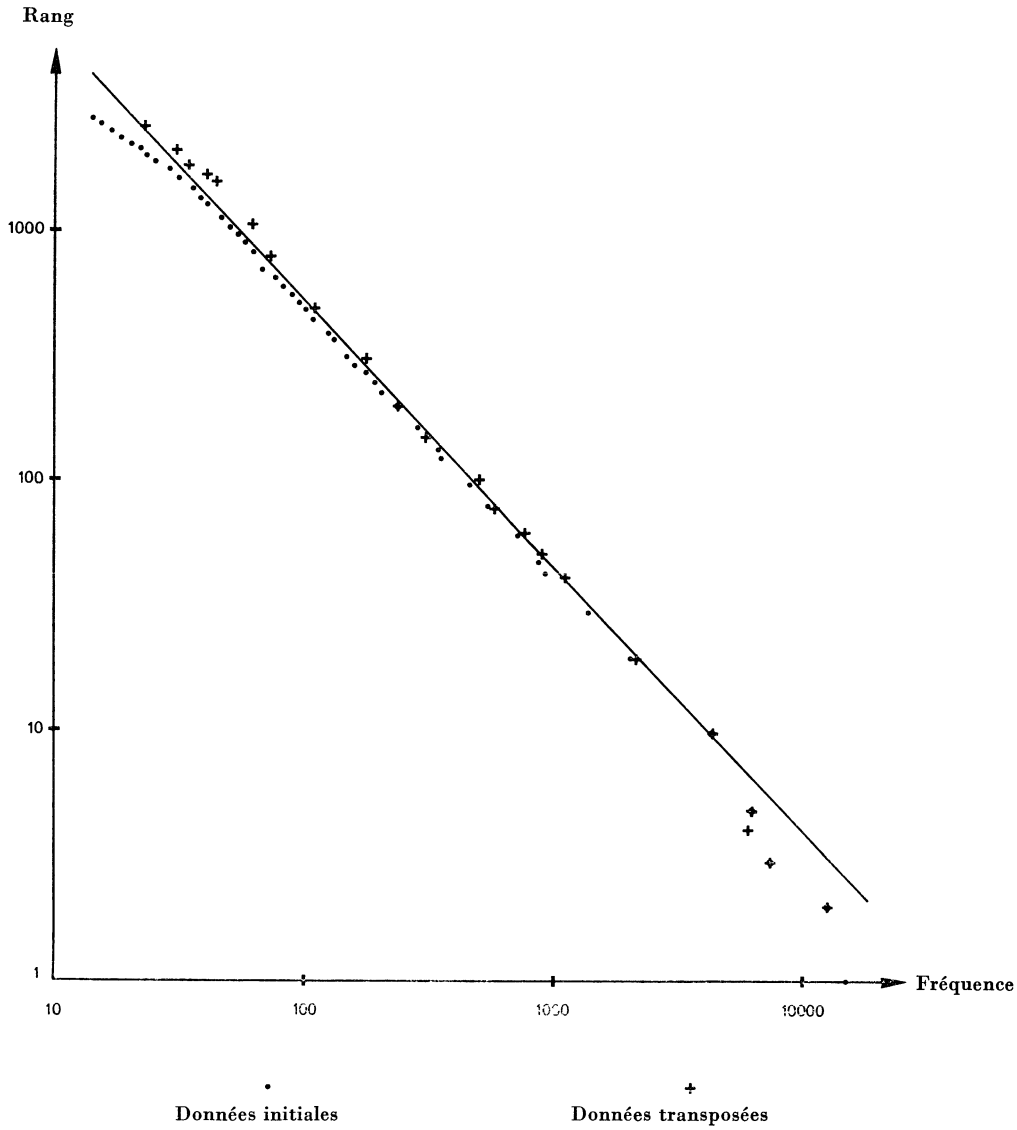

Figure 6

Fréquence des mots dans le russe contemporain

2.1.1. *Ajustement parétien* (Fig. 6). On obtient, avec une loi de Pareto à deux paramètres, une droite d'ajustement très « classique » en ce sens que les données réelles sont très bien ajustées en milieu de distribution, mais qu'elles se courbent aux deux extrémités. On adopte alors une loi à trois paramètres pour redresser ces courbures. L'ajustement graphique paraissant alors satisfaisant, on va appliquer les tests de χ^2 (on a admis que le corpus pouvait être considéré comme un échantillon) et de Kolmogorov (on ne considèrera pas celui-ci non plus dans une interprétation probabiliste rigoureuse impossible ici, mais simplement comme un indicateur de distance entre données et droite d'ajustement). Pour une critique très explicite de l'application de ces deux tests aux données de ce type on renvoie à l'article de R. E. Quandt dans [3]; on les a cependant appliqués car ils ont amené à rejeter les hypothèses, on aurait eu plus de scrupules à le faire dans le cas contraire.

Test de $\varkappa^2$ pour $\nu = 6$ au seuil 5%: $\chi^2 = 12{,}6$; valeur obtenue: $Q^2 = 53{,}20$.
Test de Kolmogorov: au seuil 5%, la borne $\varepsilon = 0{,}03$; valeur obtenue: $\Delta = 0{,}44$.
Ces tests amènent à rejeter l'hypothèse que la loi sous-jacente soit une loi de Pareto.

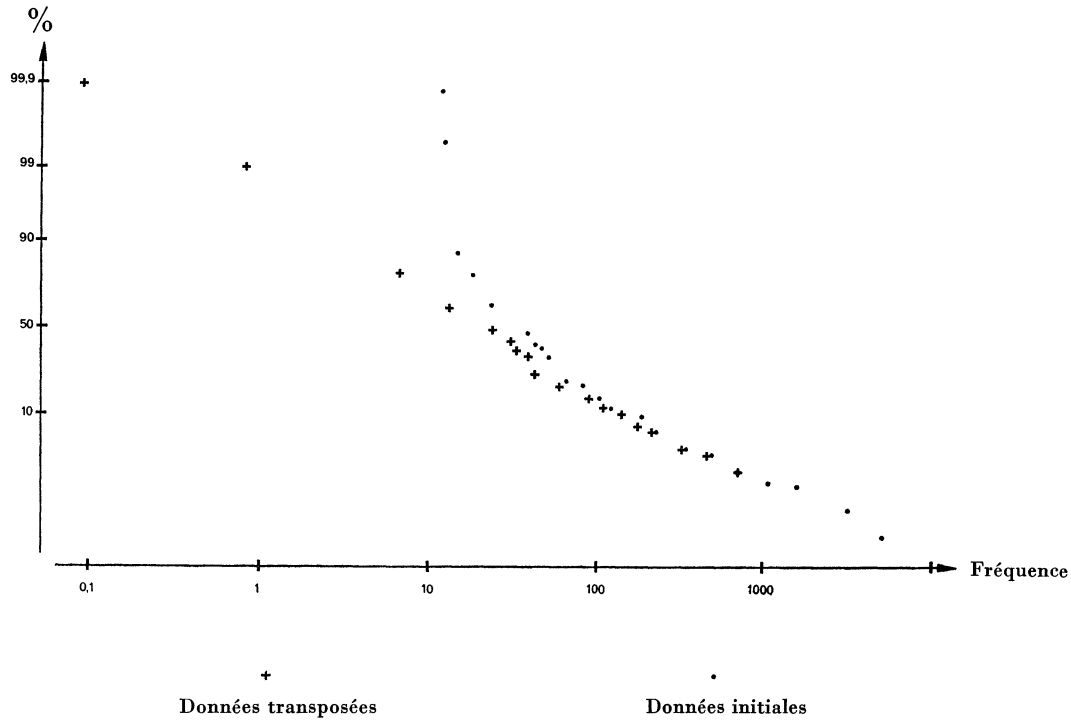

Figure 7

Ajustement log-normal des fréquences de mots dans le russe contemporain

2.1.2. *Ajustement log-normal*, à trois paramètres. On l'a essayé malgré les objections soulevées par B. Mandelbrot [5]. Ici aussi les tests amènent à rejeter l'hypothèse (Fig. 7).

Les auteurs qui étudient le type de données en question, on pourra s'en convaincre en lisant les notes de B. Leclerc [3], ne présentent pratiquement jamais de tests qui amèneraient dans la grande majorité des cas à rejeter l'hypothèse, ceci, pensons-nous, parce que le modèle sous-jacent ne convient pas. La plupart des auteurs ne présentent que des ajustements graphiques sans d'ailleurs parfois observer une grande rigueur dans leur application. On les voit alors adopter le plus fréquemment la loi log-normale, ce qui revient à s'installer dans la position confortable du statisticien qui dispose de tous les raffinements de l'analyse et d'un modèle. Et en effet, dans un cas comme celui-ci: ajustements graphiques acceptables, refus des deux hypothèses, pourquoi ne pas le faire ?

2.2. Concentration urbaine aux États-Unis

Il s'agit des villes de plus de 2 500 habitants, aux États-Unis en 1910 (Tableau 4).

De nombreux travaux démographiques ont mis en évidence le rapport existant entre la taille d'une agglomération et l'attraction qu'elle exerce sur la population d'une certaine zone, c'est-à-dire son accroissement. Le cadre conceptuel du théorème central limite paraissant tout à fait adéquat dans ce domaine, on tente en premier lieu un

2.2.1. *Ajustement log-normal* (Fig. 8): effectif des villes dont la population est supérieure ou égale à une limite donnée. On est amené à adopter une loi à trois paramètres avec $x_0 = 2\,400$, qui permet d'obtenir un ajustement graphique satisfaisant après troncature à 1% et 82%.

Tableau 4

Distribution des villes des États-Unis en 1910 selon leurs classes de population

Nombre de villes	Limite des classes de population
—	moins de 2 500 hab.
1 060	2 500 à 5 000
605	5 000 à 10 000
369	10 000 à 25 000
119	25 000 à 50 000
59	50 000 à 100 000
31	100 000 à 250 000
11	250 000 à 500 000
5	500 000 à 1.10^6
3	1.10^6 et plus
2 262	

Source. — Historical Statistics of the United States, Colonial Times to 1957, continuation to 1962 and revision, Statistical abstract supplements, Series A 195-209, US Bureau of Census.

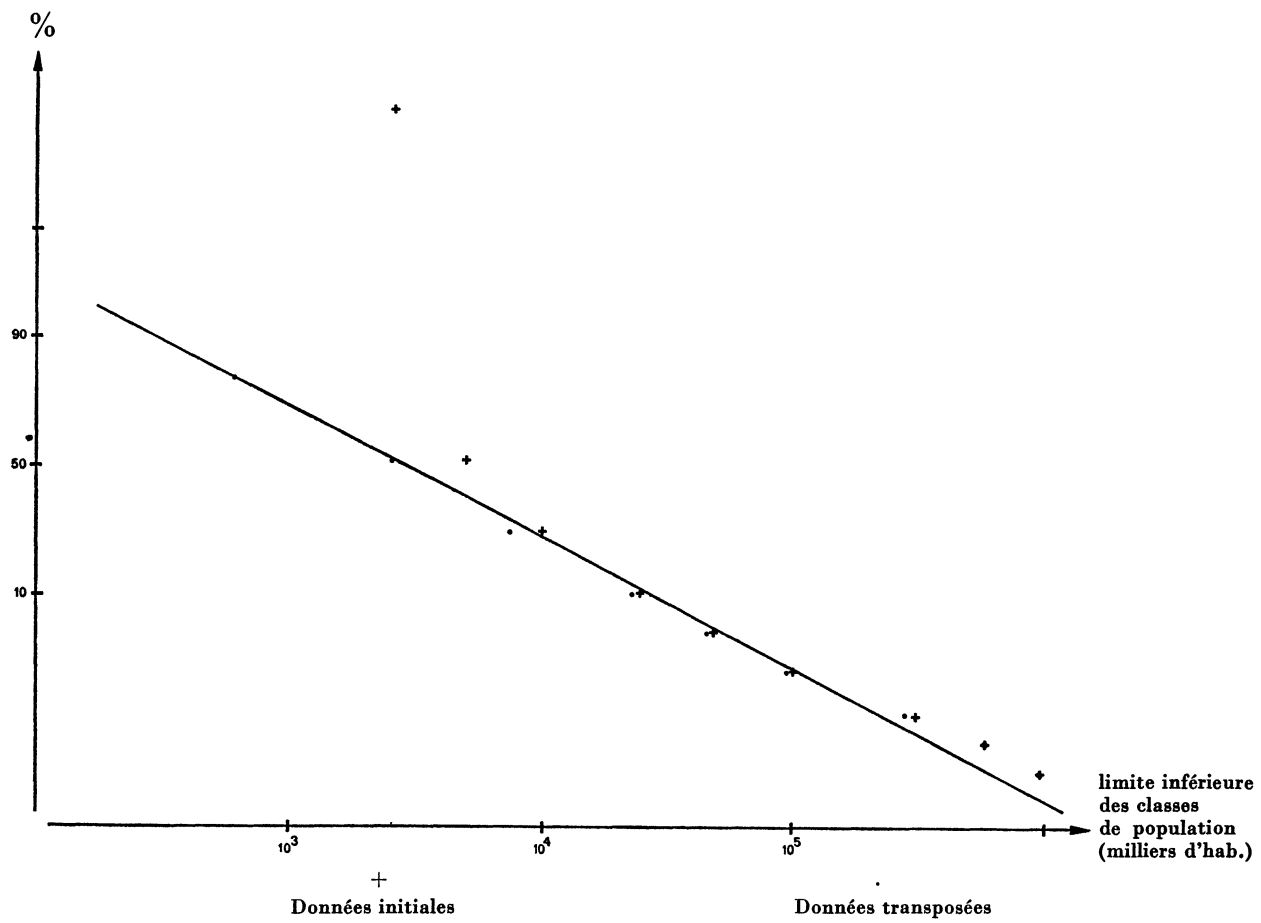

Figure 8

Répartition des communes aux USA 1910 (ajustement log-normal)

La liste des villes américaines ne pouvant être assimilée à un échantillon, on n'effectue pas le χ^2. On calcule la valeur de Δ de Kolmogorov, pour lequel on répète les réserves ci-dessus. Si on admettait l'application du test, on ne serait pas amené à rejeter l'hypothèse log-normale puisque, au seuil 5%, la borne est $\varepsilon = 0{,}02$ et qu'on a ici un $\Delta = 0{,}015$.

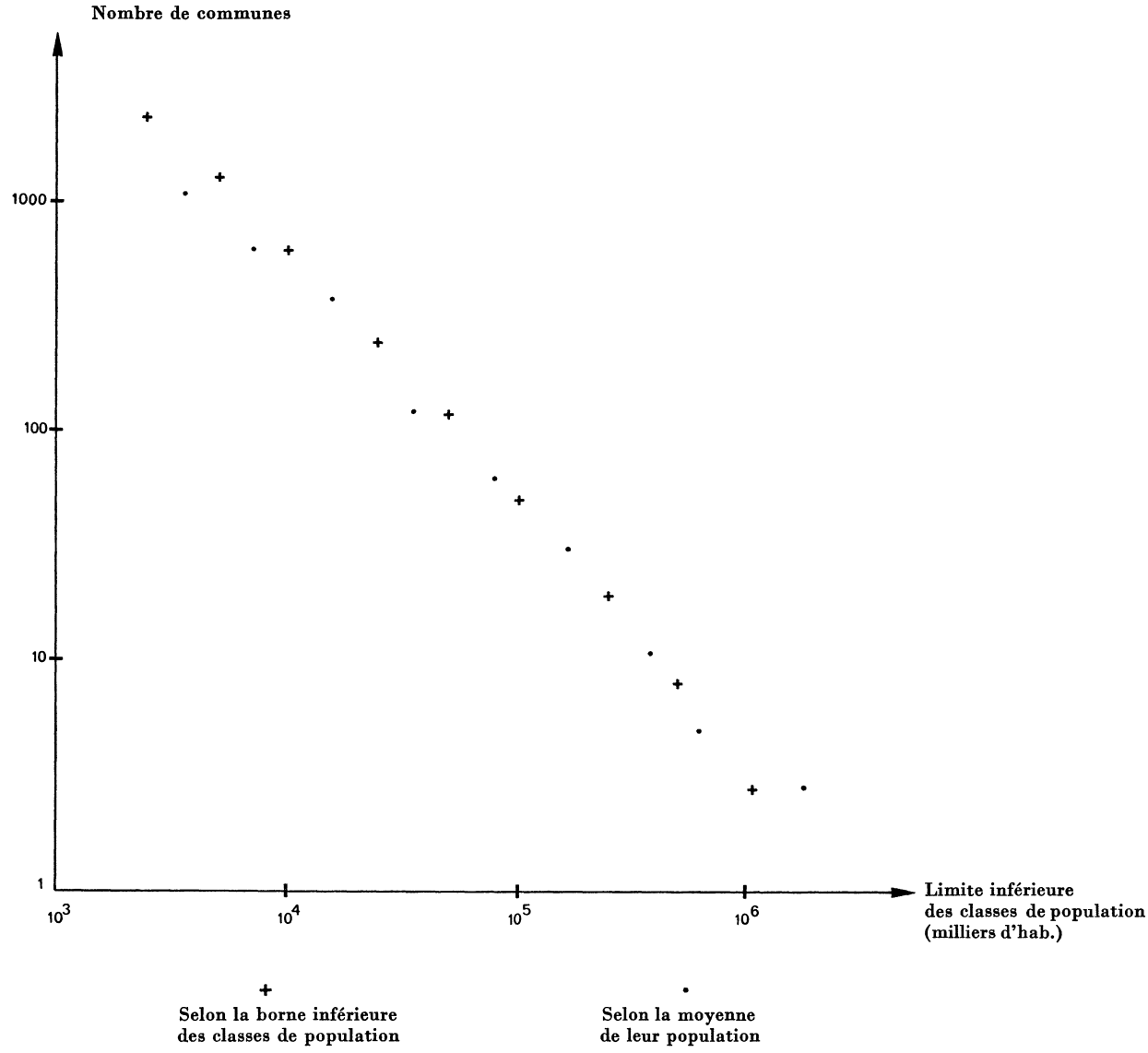

Figure 9

Répartition des communes aux USA 1910
(ajustement parétien)

2.2.2. *Ajustement parétien* (Fig. 9): effectif des villes dont la population est supérieure ou égale à une limite donnée. On obtient avec la loi de Pareto à deux paramètres, une très bonne droite d'ajustement de pente $\alpha = -1{,}06$. Seul le dernier point représentant l'effectif des villes de population supérieure ou égale à 1.10^6 s'en sépare.

Test de Kolmogorov: $\varepsilon = 0{,}02 \quad \Delta = 0{,}0006$ (au seuil 5%).

On n'est pas amené à rejeter l'hypothèse que la concentration urbaine suive un modèle parétien. Mais quel modèle ? N'a-t-on pas dit qu'il n'y en avait pas qui soit universellement accepté ?

2.2.3. *Un modèle de partage*

Dans ce cas précisément, on va pouvoir présenter un modèle déterministe car initialement, ces données ont été exploitées dans un cadre parétien. On sait que la loi de Pareto établit une relation entre deux progressions géométriques; les deux variables: le caractère étudié et les effectifs liés à chacun des niveaux repérés de celui-ci, variant en progression géométrique. On renvoie à ce sujet à G. Th. Guilbaud ([4], pp. 184 à 187). On peut considérer les variables chacune globalement comme une masse à fractionner dans des proportions différentes mais constantes: les raisons respectives de chacune des progressions géométriques. Un exemple très simple. On a une masse M de 100 000 unités que se partage une population P de 3 000 personnes. Ce partage se fait selon les modalités suivantes: le 1/10 de M est affecté aux 2/3 de P, c'est la couche la plus défavorisée et il reste 90 000 unités à partager entre 1 000 personnes. On réitère et on affecte 1/10 (1/10 de M) aux 2/3 (2/3 de P) et ainsi de suite (Fig. 10).

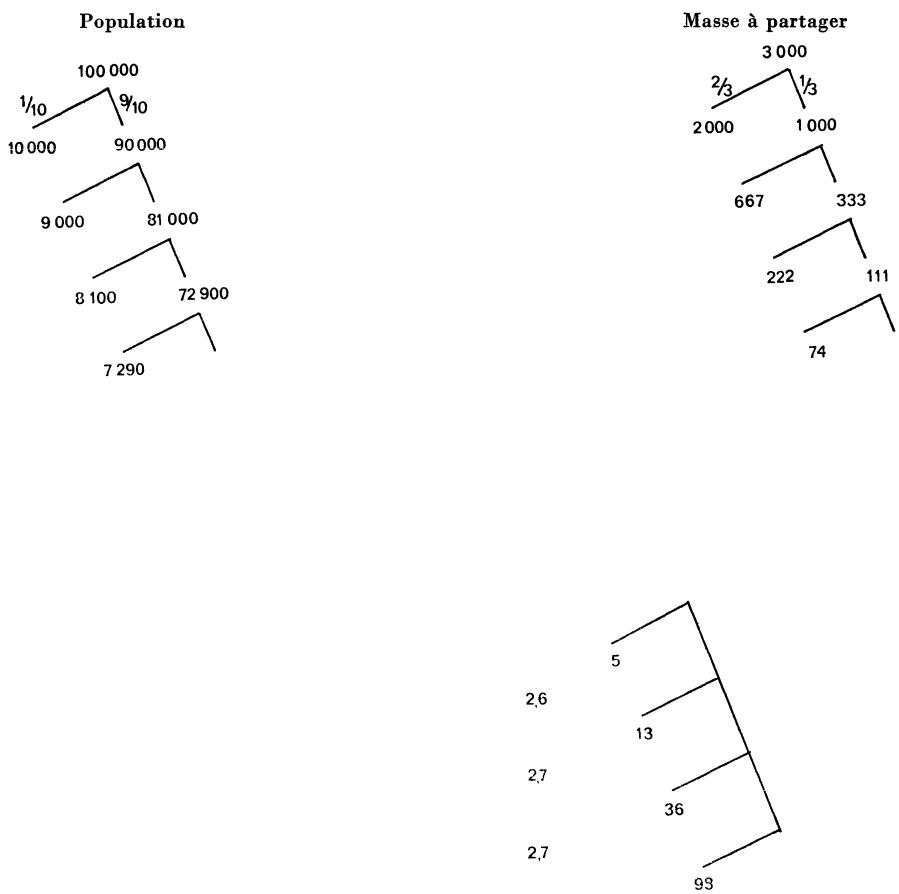

Figure 10
Schéma du processus de partage

On a alors, pour représenter le processus, trois arbres d'un type particulier mais connu des phyllotaxinomistes qui les appellent « plaviotropes »: du seul tronc se séparent des rameaux mais ceux-ci ne donnent jamais naissance à un autre rameau. Ces trois arbres peuvent être appelés les arbres des variables et celui qui établit la relation entre les deux premiers, l'arbre des parts ou arbre du revenu moyen par couche, si on étudie des revenus. La Figure 11 présente les trois arbres pour la concentration urbaine aux USA en 1910. Si on désigne par k une étape du processus, la moyenne dans la couche $(k, k + 1)$ s'écrit, si l'on désigne la population par P et le nombre des villes par N:

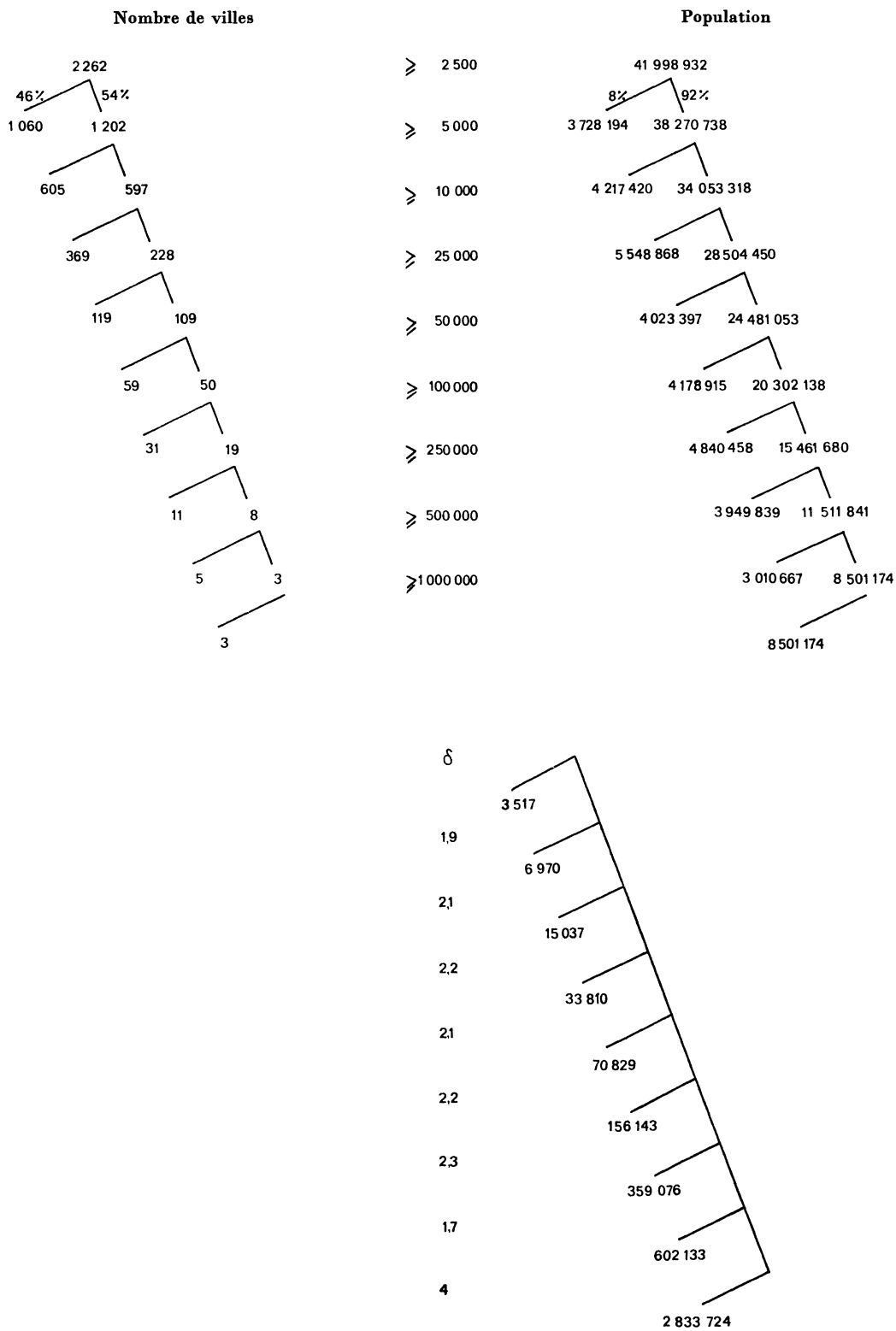

Figure 11

Les arbres des variables et l'arbre des parts

$$\frac{P_k - P_{k+1}}{N_k - N_{k+1}}$$

et on peut définir un indicateur d'inégalité δ en écrivant:

$$\frac{P_k}{N_k} = \delta \frac{P_k - P_{k+1}}{N_k - N_{k+1}}.$$

Selon les domaines d'application, ce modèle rencontre plus ou moins le sens commun; dans le domaine de la théorie de l'organisation on pourrait sans doute en trouver de bons exemples. S'il s'est avéré inapplicable à la concentration urbaine française, probablement en raison du long passé historique dont elle constitue le témoignage, il s'est avéré très satisfaisant quant à la stabilité des raisons et de δ pour la concentration américaine, comme on peut le voir sur la Figure 11. Cependant, le graphique de l'arbre des parts ne donne qu'un très médiocre ajustement parétien.

CONCLUSION

Pour rendre compte de distributions empiriques très dissymétriques vers la droite, le statisticien a le choix entre plusieurs lois, par exemple la loi de Pareto ou la loi log-normale. Ce choix est souvent difficile: on a essayé de montrer qu'on pouvait s'y tromper. Se référer au plan théorique serait une situation fort incommode: l'une des lois, la loi de Pareto, n'a pas de fondements probabilistes indiscutables et la loi log-normale en a de si puissants qu'il est difficile de considérer les contraintes impliquées comme étant vérifiées dans de nombreux cas concrets. On voudrait donc plaider que l'usage se répande d'effectuer les deux ajustements comme une pratique commode et rapide pour juger qualitativement une distribution empirique et que toute supposition sur le modèle sous-jacent n'apparaisse qu'une fois ce premier critère utilisé.

Les graphiques ont été calibrés et dessinés par J. Leconte.

BIBLIOGRAPHIE

[1] Aitchison, J. et Brown, J. A. C., *The log-normal distribution with special references to its uses in economics*, University of Cambridge, Monograph 5, Cambridge, Mass., Cambridge University Press, 1969, 5e ed.

[2] Calot, G., *Cours de statistique descriptive*, Paris, Dunod, Coll. Statistique et Programmes économiques, 1969.

[3] Guilbaud, G. Th., *Mathématiques*, Paris, Presses Universitaires de France, Coll. Thémis, 1963.

[4] Leclerc, B., "Applications pratiques des lois de probabilité (11)", *Math. Sci. hum.*, no 36, 1971, pp. 79-89.

L'ensemble de ces études sera publié sous le titre : *Distributions statistiques et lois de probabilités*, Paris, Mouton/Gauthier-Villars, à paraître en 1972.

[5] Mandelbrot, B., "On the theory of word frequencies and on related markovian models of discourse", in : R. Jakobson (ed.), *Structure of language and its mathematical aspects*, pp. 120-219, *Am. math. Soc.*, 1960.