Journal de la société statistique de Paris

M. L. Dufrénoy
J. Dufrénoy

La distribution des biens et des richesses obéit-elle à une loi naturelle ?

Journal de la société statistique de Paris, tome 106 (1965), p. 255-272
<http://www.numdam.org/item?id=JSFS_1965__106__255_0>

V

LA DISTRIBUTION DES BIENS ET DES RICHESSES OBÉIT-ELLE A UNE LOI NATURELLE?

SOMMAIRE

INTRODUCTION

La nature est un code qui nous cache et nous livre à la fois la réalité...

En ces termes A. Molitor (1) interprète l'œuvre de Paul Claudel en tant que tentative de révélation de « cette unité secrète au sein de la variété », de « cet enchaînement à l'infini de nombres et de signes qui est le chiffre même de la création (au sens où l'on dit : un chiffre diplomatique) ».

Le signe ou la figure est avant tout l'expression d'un rapport.

L'objet de notre étude a été de rassembler l'information pouvant guider le choix vers le modèle le plus vraisemblable, permettant d'exprimer un « rapport » entre la « dimension » d'une exploitation agricole (Surface Agricole Utile, ou SAU, quantifiée en unités de surface du pays : Ha, ou acre,...) et la fréquence des exploitations dans chaque classe de SAU.

Nous nous proposons de rechercher quel est le modèle le plus vraisemblable de représentation de cette forme de « biens » que représente le sol utilisé en tant que « ressource » en vue de la production agricole.

Quel modèle mathématique peut être considéré comme le plus vraisemblable parmi tous ceux qui peuvent être « imaginés » pour représenter les distributions de fréquences

1. A. Molitor, Aspects de Paul Claudel, 191-192, Paris, 1945.

observées, soit pour l'ensemble d'un pays, comme la France, soit pour telle ou telle région, administrative, ou géographique?

Deux concepts peuvent être invoqués pour guider le choix entre les divers modèles.

CONCEPT TECHNOLOGIQUE

Une exploitation, telle qu'elle est définie aujourd'hui par sa surface, est considérée comme résultant d'un processus de fragmentation; on invoque le modèle mathématique s'appliquant, en granulométrie, à la distribution des résultats de la fragmentation, en particules de tailles au plus égale à x_i, d'une masse originairement indivise d'un matériau; le modèle le plus vraisemblable est celui de la distribution log normale.

CONCEPT IDÉOLOGIQUE

Une exploitation agricole met en œuvre, comme source de revenus, trois ressources principales : sol (ressource quantifiée en unités de surface de SAU) capital, et main-d'œuvre (quantifiée en heures de travail humain, les ressources mécaniques adéquates étant mises au service de cette main-d'œuvre). Selon ce concept, doit exister une forte corrélation entre « SAU » et « revenu »; il convient donc de comparer la distribution des fréquences de « SAU » avec la distribution des fréquences de niveaux de revenus : le modèle le plus vraisemblable étant encore celui de la distribution log normale.

« FRAGMENTATION » ET « GRANULOMÉTRIE »

A une époque donnée et en une région donnée, la distribution des exploitations agricoles par classes de SAU révèle le résultat d'un processus de fragmentation (1) d'une ressource initialement indivise : la surface du sol.

Les techniques expérimentales de la granulométrie permettent de comparer des « distributions observées » de particules par classes de fréquence, chacune définie par une limite inférieure et une limite supérieure des tailles des particules : le modèle le plus vraisemblable pour représenter ces distributions observées est le modèle de distribution log normale, ainsi qu'il est attesté notamment par des articles parus dans trois numéros de *The Analyst* (2), (3), (4). Le sol cultivable, au sens agronomique, représente une population de particules, chacune provenant de la fragmentation du matériau primitivement indivis représenté par la roche mère : sauf modifications par apports de sédiments ou par élimination (érosion) de certains constituants, un sol révèle à l'analyse granulométrique (5) une distribution log normale des particules par classes de fréquences, chaque classe de fréquences, étant séparée successivement des autres par tamisage ou toute autre technique de triage par dimensions : en effet, si une distribution est log normale quant à la distribution

1. G. MEDICI, V. SORBI et A. CASTRATARO, Polverizzazione e frammentatazione della proprietà fondiaria in Italia. *Istituto Naz. di Eco. Agraria*, Milano, 1963.
2. A. D. WILSON, Sampling of Rock powders. *The Analyst*, 89 : 18-30, janv. 1964.
3. Particle Size, Measurement, interpretation and application, by R. R. IRANI et CLAYTON F. CALLES. Review by D. G. Beech, *The Analyst*, 89, 151, 2, fev. 1964.
4. Classification of methods, Society for Analytical Chemistry, Particle Size Analysis Sub. Committee. *The Analyst*, march 1963.
5. Esther PERRY, Les sols de Californie, *C. R. Acad. Agricult.*, 50 : nº 1, 95-98, 1964.

de dimensions linéaires, elle est aussi log normale quant à la distribution des surfaces des particules, ou de leur volume ou de leur masse (1).

Une classification par tailles peut s'effectuer par la technique du « tamisage ». Chaque tamis, avec orifices de dimension x_i retient toutes les particules pour lesquelles x_i représente la limite inférieure de classe, et laisse passer toutes les particules de taille inférieure à x_i (ou de taille au plus égale à x_i).

Cette classification par tailles, telle qu'elle est utilisée en granulométrie, s'applique à toute classification, par tailles, des diverses parties résultant de la fragmentation de ce qui avait été un tout, et notamment de l'allocation à des exploitations d'une partie de la ressource sol.

LA « FORTUNE MORALE » ET LA « FORTUNE PHYSIQUE »

De Bernouilli, fondant la Théorie de l'*utilité* de la monnaie, faisait correspondre la « fortune morale » au logarithme de la « fortune physique ».

Pour une exploitation agricole, la Surface Agricole Utile (SAU) fournit une estimation de la « fortune physique ».

La distribution observée des exploitations par classes de SAU est conforme à l'hypothèse de la distribution log normale. Ce modèle de distribution log normale peut être utilisé pour interpréter le concept de « fortune morale » imaginé il y a deux siècles par Bernouilli, et pour réinterpréter la théorie de l'*utilité* du capital, du point de vue de l'interconvertibilité des diverses ressources que doit mettre en œuvre une exploitation agricole : ressource sol (exprimée en hectares de SAU) ressource capital monétaire, ressource main-d'œuvre (exprimée en unités de travail humain, c'est-à-dire en heures par semaine, ou heures par hectares de SAU) et enfin, cette ressource dont Bernouilli ne pouvait, il y a deux siècles, soupçonner l'importance, ressource qui met à profit tous les moyens de production inventés, découverts et rendus utilisables par les progrès technologiques.

LA FORTUNE MORALE CORRESPOND AU LOGARITHME DE LA FORTUNE PHYSIQUE

Le modèle de distribution log normale

La relation fonctionnelle entre « fortune physique » et « fortune morale » au sens de Bernouilli peut être calculée en faisant intervenir la VMP ou « valeur marginale de productivité » d'une ressource.

Parmi les diverses ressources investies en vue d'une production, il convient de choisir celle qui n'est disponible qu'en quantité relativement insuffisante : la quantité de cette ressource est choisie comme variable indépendante (X_1) : on étudie expérimentalement l'effet (y_1) de la variation de (X_1), chacune des autres ressources (X_i) étant fixée à un certain niveau (x_i) : selon le modèle de la distribution log normale, conforme à la notion de fortune morale, ce niveau pourra être fixé à la moyenne géométrique (x_{ig}).

Choix entre les modèles de distribution; « log normale » ou « normale »

Parmi les divers modèles mathématiques permettant de représenter une distribution observée, il convient de choisir « le plus vraisemblable »; dans le cas de distribution de

1. A. Guinet, Contribution aux techniques d'appréciation du Calibrage. *Ibid.*, 98-99.

fréquences, le modèle le plus vraisemblable est celui qui permet de prévoir, pour chaque classe, une « fréquence espérée » aussi voisine que possible de la « fréquence observée »; le « test » classique est celui du Chi²; le modèle le plus vraisemblable est celui qui minimise la somme des Chi². Il est d'ailleurs beaucoup plus simple de calculer le critérium d'information 2i, qui, peut être calculé, non seulement quant aux fréquences comme le Chi², mais aussi quant aux pourcentages.

Aux méthodes d'estimation communes à toutes les distributions définies par deux paramètres (μ et σ), c'est-à-dire, méthodes de la « vraisemblance maxima », des « moments », ou des « quantiles », s'ajoute l'utilisation de tables de valeurs de $N\left(\frac{\sigma}{2}\ 0{,}1\right)$ publiées par J. Aitchison (*The log normal distribution*, Cambridge Univ. Press, 1957). En fait, l'estimation algébrique des paramètres, théoriquement simple, pour les distributions qui n'ont été soumises ni à groupement arbitraire ni à censure, devient difficile pour les distributions observées telles qu'elles ont été le plus souvent fournies. C'est alors la méthode graphique, permettant d'apprécier les irrégularités éventuelles d'alignement des points sur la droite de régression, qui fournit la meilleure information quant au choix du modèle : il suffit de comparer les distributions des points, en utilisant la même échelle de probabilité normale pour les pourcentages cumulés, l'échelle des abscisses étant arithmétique pour le modèle « normal », logarithmique pour le modèle log normal. Le choix peut être difficile lorsque l'étendue des échelles d'abscisses est si courte que soit peu marquée la différence entre les deux échelles; il convient alors de construire une échelle log dont chaque « cycle » soit suffisamment étalé.

Représentation graphique d'une distribution de fréquences $F(x_i)$ selon le modèle log normal.

Chaque valeur de (x_i) définit une valeur supérieure de classe, par exemple 5, 10, 20, 50... 1 000 ha. La fréquence dans chaque classe : 0 à 5, puis 5 à 10, puis 10 à 20... est $(f x_i)$ telle que pour N classes la somme soit représentée symboliquement par $\sum_{i=1}^{i=N} (f x_i)$.

On calcule les fréquences cumulées, c'est-à-dire qu'on ajoute successivement à la fréquence dans la première classe avec $i = 1$ (classe 0 à 5), la fréquence dans la 2e classe (5 à 10)... puis la fréquence dans la classe x_{N-1} à x_N; la somme est évidemment égale à $\sum_{i=1}^{i=N} (f x_i)$.

On pose $\sum_{i=1}^{i=N} (f x_i) = 100$ et on calcule les pourcentages cumulés, correspondant chacun successivement, à l'une des fréquences cumulées; (ce calcul de pourcentages se fait rapidement avec une règle à calcul).

Utilisant le papier log « normal », on porte en abscisses chaque limite supérieure de classe (x_i) sur échelle logarithmique, et en ordonnée, sur échelle de probabilité normale, le pourcentage cumulé; on définit ainsi, par la position d'un point, pour chaque limite de classe (x_i), le niveau sur l'échelle de probabilité normale.

Par ces points on trace la droite de régression, qui coupe le niveau 50 % (centre de symétrie de l'échelle de probabilité normale) à l'aplomb de la valeur qui, sur l'échelle d'abscisses, correspond à la médiane de la distribution, c'est-à-dire à la SAU telle que 50 % des exploitations ont une surface inférieure (ou au plus égale à cette médiane) et 50 % une surface au moins égale.

Selon le modèle de la distribution log normale, la médiane correspond au « quantile », $(Q_{50\,\%})$ dont la valeur numérique est e^{μ} et correspond à la base des log népériens ($e = 2\,718...$)

élevée à la puissance μ, lorsque μ représente la moyenne de la distribution, soit $\mu = \log Q_{50\,\%}$.

Cette valeur de $Q_{50\,\%}$ ou son log définit le paramètre de position de la distribution log normale.

Il reste à caractériser, par un 2e paramètre, la pente de la droite, en déterminant, sur le graphique, le quantile $Q_{16\,\%}$, lequel correspond, sur l'échelle log des abscisses, à l'aplomb de l'intersection de la droite avec le niveau 16 %; symétriquement, on peut estimer, sur l'échelle des abscisses, la valeur du $Q_{84\,\%}$; ces deux quantiles sont en effet distribués symétriquement, sur l'échelle des abscisses par rapport au Q_{50} puisque

$$Q_{16\,\%} = e^{\mu - \sigma} \qquad Q_{50\,\%} = e^{\mu} \qquad Q_{84\,\%} = e^{\mu + \sigma}$$

On désigne par les lettres grecques mu (μ) et sigma (σ) les valeurs « théoriques » des deux paramètres : de position et de dispersion.

On estime le paramètre de position en lisant sur le graphique l'abscisse correspondant à $Q_{50\,\%}$ et on estime la moyenne $m = \log Q_{50\,\%}$.

On estime la variance s correspondant à la valeur théorique par ce simple calcul de proportions :

$$\sigma = \log \left[\frac{1}{2}\left(\frac{Q_{50\,\%}}{Q_{16\,\%}} + \frac{Q_{84\,\%}}{Q_{50\,\%}}\right)\right]$$

Avantages théoriques et pratiques de la distribution log normale. La distribution est définie par 2 paramètres, faciles à estimer graphiquement, par la position et la pente de la droite de régression. Les distributions de fréquence observées quel que soit le groupement par classe, et quel que soit le niveau où elles ont été censurées, pour les classes inférieures ou les classes supérieures, sont immédiatement utilisables, sans autre calcul que celui des fréquences cumulé et leur transformation en pourcentages cumuléss.

SAU ET NIVEAU NUTRITIONNEL DES EXPLOITANTS

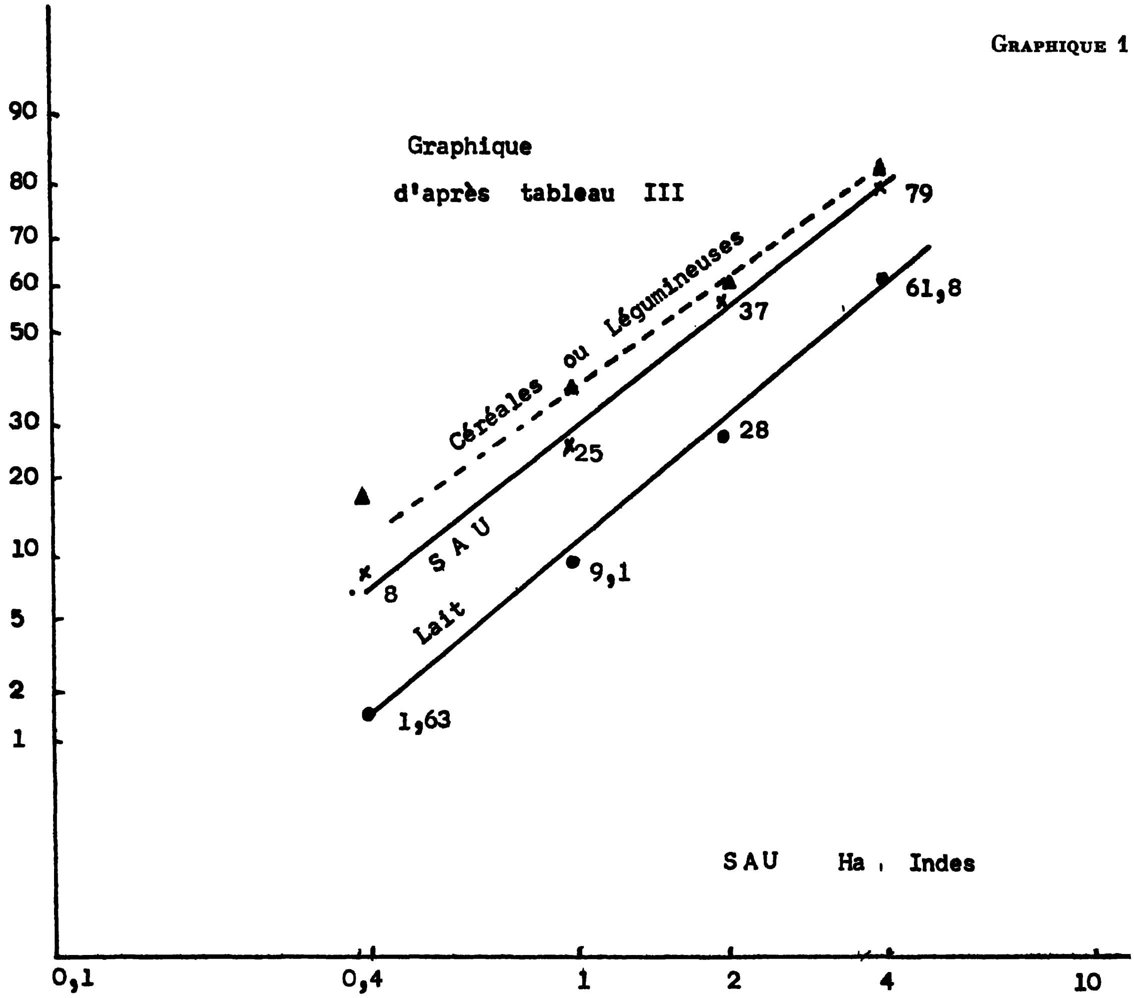
La corrélation entre SAU et niveau nutritionnel peut être illustrée par l'étude de 320 exploitations aux Indes (*Indian J. Agric. Economics 18*, nº 1, p. 323, 1963). Le tableau indique : les limites supérieures (x ha) de classe de SAU; les % cumulés d'exploitation disposant au plus de x ha; il indique en outre les % cumulés de la distribution des sources de calories pour diverses ressources nutritionnelles, pour chaque % cumulé d'exploitation disposant d'au plus x ha.

TABLEAU III

x	%	Céréales	Légumineuses	Lait	Total
0,4	8	17	16,3	1,63	16
1	25	36	37,2	9,1	34
2	51	57	57,5	28	54
4	83	79	79	61,8	75,5

Les 320 familles étudiées disposent en moyenne de 2 615 calories; celles de la classe inférieure ne disposant que de 2 157 et celles de classe supérieure (SAU supérieure à 4 ha) de 3 000, mais l'importance relative des sources de ces calories varie considérablement : 50 % des familles, avec des exploitations de moins de 1,6 ha, participent pour moins de 25 % à la consommation du lait : les 16 familles disposant de plus de 4 ha consomment 60 % du lait.

Graphique 1

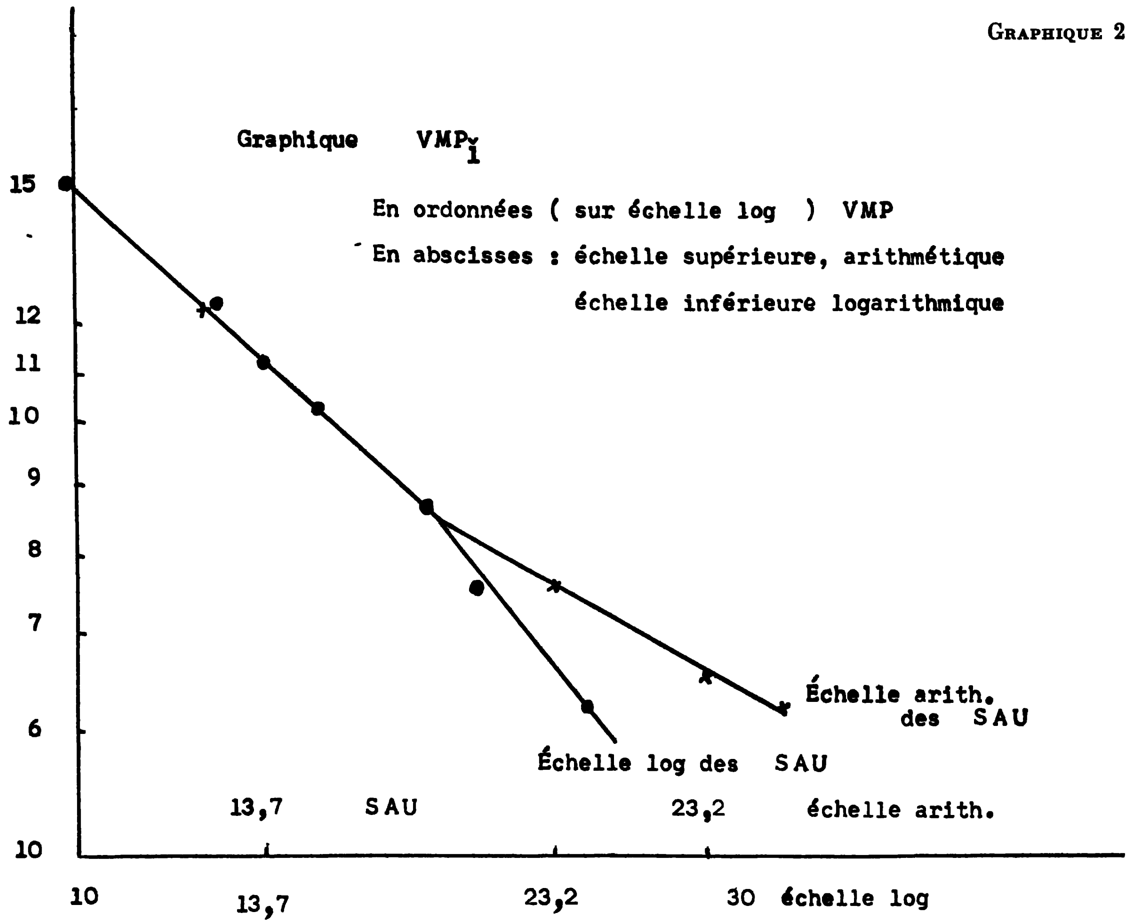


SAU ET VMP

Exemple : K. Izumi, à l'École d'Agriculture de l'Université de Tottori, au Japon, a étudié 31 exploitations agricoles, dont les SAU se distribuent entre 10 et 23,2 tans (c'est-à-dire entre 0,1 et 2,32 ha) autour de la moyenne géométrique 13,7 tans : Exprimée en milliers de « yens » la valeur marginale de la productivité (VMP) d'un tan varie selon la position de ce tan par rapport à la moyenne géométrique : chacune des autres ressources étant fixée à la moyenne géométrique, le tan de rang « 13,7 », (celui qui fait passer la SAU de 12,7 à 13,7) a une VMP de 11.06; la VMP atteint 15.15 pour le 10e tan (qui fait passer la SAU de 9,5 à 10,5) et n'est plus que 6,5 pour le tan de rang 23,2, qui fait passer la SAU de 22,2 à 23,2 tans (2,22 à 2,32 ha)

La distribution log normale des SAU s'équilibrant autour de la moyenne géométrique 1,37 ha, il paraît logique d'estimer la VMP en fonction du nombre des unités (de 10 ares) « qui manquent pour atteindre 1,37 ha », ou « en excès de 1,37 ha ». En portant en abscisses sur échelle log les valeurs de x et en ordonnées sur échelle de probabilité normale, les « pourcentages cumulés » des VMP par classes, « on détermine une série de points par lesquels on

GRAPHIQUE 2



peut faire passer une droite de régression, ce qui suggère que la distribution des VMP est elle aussi, conforme à la distribution log normale (Graphique VMP_2).

La faible étendue des variations de la variable indépendante (x tans) rend difficile le choix du modèle la plus vraisemblable pour représenter la relation fonctionnelle entre (x) et la variable dépendante (y Yens).

Le modèle de la distribution log normale paraît logique, mais on obtient dès alignements de points selon une droite de régression en portant en abscisses x et en ordonnées $(1/y)$ (transformation d'hyperbole en droite) ou $\log y$ (transformation de courbe exponentielle).

Enfin on peut envisager la fonction $\log y = K \log (x)$ puisqu'on peut calculer l'équation polynomiale

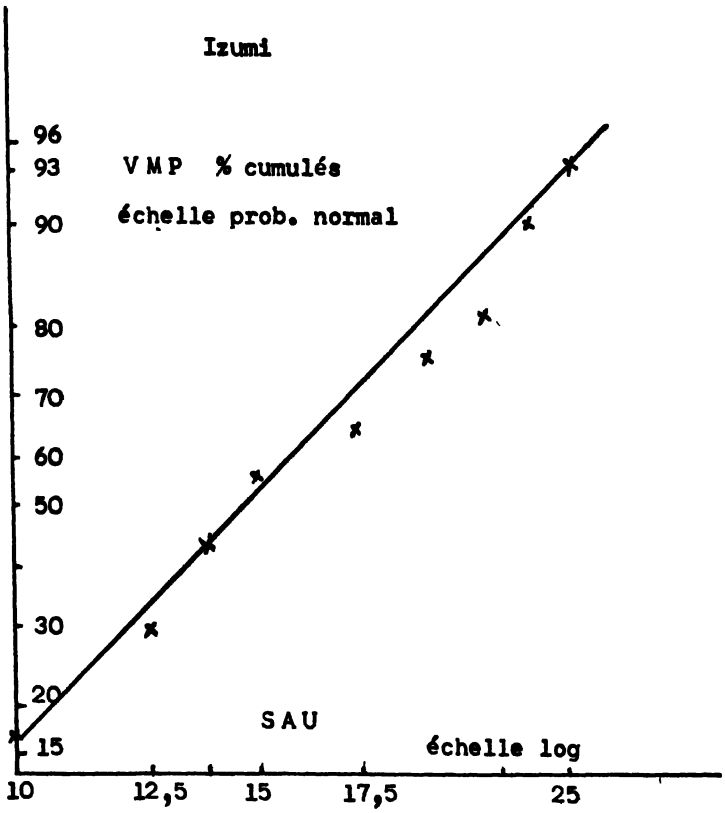
$$y = 9{,}945 - 0{,}86\ Z_1 - 0{,}146\ Z_2$$

MODÈLE MATHÉMATIQUE LE PLUS VRAISEMBLABLE DE REPRÉSENTATION DES DISTRIBUTIONS DE FRÉQUENCES DE « SAU »

La « propriété foncière individuelle » a pour origine la « fragmentation » d'une ressource originairement indivise : le sol.

Un matériau originairement indivis étant soumis à fragmentation, les fragments se répartissent par classes de tailles selon le modèle de la distribution log normale : c'est

GRAPHIQUE 3



ce modèle qui peut être adopté comme le plus vraisemblable *a priori* pour la distribution des SAU ou des revenus (1).

Signification de la distribution log normale (1)

Premier exemple : Région de petites et moyennes exploitations : la Bresse louhannaise.

Répartissons les SAU par classes telles que la limite supérieure de chaque classe soit 5, puis (5) (2); (5) (3); (5) (n) les limites supérieures des classes forment les termes d'une progression arithmétique (tableau IV).

TABLEAU IV

Bresse Louhannaise

Limites			%	% cumulés
inférieure		supérieure		
1	0	5	16	16
2	5	10	30	46
3	10	15	33	79
4	15	20	8	87
5	20	25	7	94
6	25	30	3	97
7	30		3	100

1. Les « distributions observées » des SAU dans diverses régions de la France, et dont il sera fait état dans ce qui suit, nous ont été communiquées par des ingénieurs D. P. E., ingenieurs de C. E. T. A., ou ingénieurs de Centre de Gestion : M. ESCANDE, pour le Sud-Ouest; M. MAERTENS pour la Brenne, M. MICONNET pour la Bresse; M. LEMAIRE pour le Gâtinais, M. M. VIVIER pour le Calvados.

En face de chaque classe inscrivons :

1° le pourcentage des exploitations dans cette classe;

2° le pourcentage cumulé, c'est-à-dire le pourcentage des exploitations de SAU au plus égale à la limite supérieure x_i.

Classes de SAU

La délimitation des « limites » selon les termes d'une progression arithmétique présente ces inconvénients :

1° la moyenne arithmétique des limites inférieures des 7 classes est 15 : cette limite est atteinte par 79 % des exploitations : la distribution est donc fortement dissymétrique.

2° les pourcentages sont très inégalement distribués par classe : 16 % des exploitations sont attribuées à la classe inférieure (moins de 5) c'est-à-dire presqu'autant qu'à l'ensemble des 3 classes 15 à 30.

Deuxième exemple : Le tableau V indique : colonne (1), les limites supérieures (x) de classes de SAU, en hectares, pour 2 536 exploitations agricoles de la Brenne : colonne (2) les logarithmes naturels, ou $\log_e$ de ces limites de classes, soit $\log_e x$; on constate que les $\log_e x$ sont approximativement les termes d'une progression arithmétique, ce qui indique que les (x) sont approximativement les termes d'une progression géométrique; colonne (3) fréquences des exploitations par classes de SAU; colonne (4) fréquences cumulées; colonne (5) pourcentages cumulés.

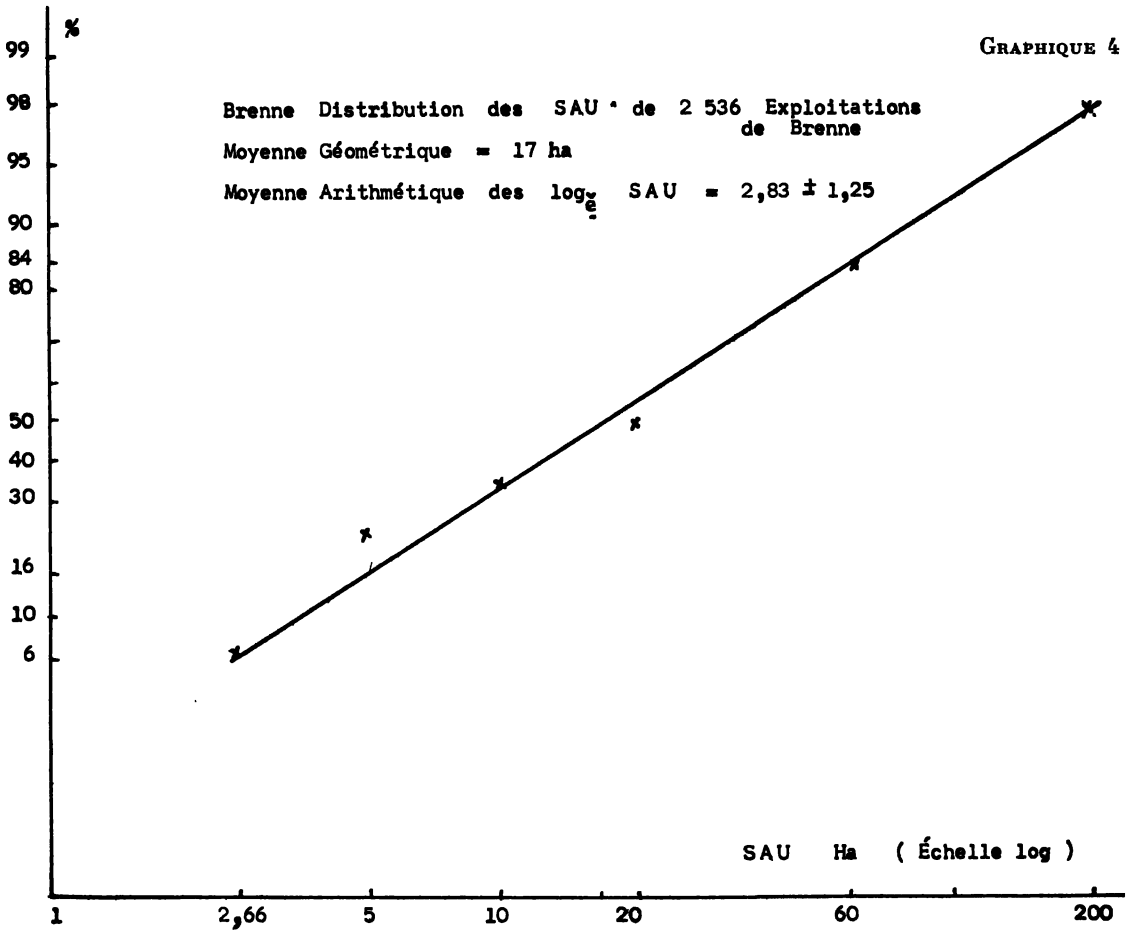
TABLEAU V

SAU Limites de classes		Fréquences		% cumulés
inférieure x	superieure $\log_e x$	par classe	cumulées	
⩽ 2,66	1	150	150	6
2,67 5	1,6	374	524	20,8
5,1 10	2,3	303	827	32,6
10,1 20	3	396	1 223	48,2
20,1 60	4	929	2 153	84,9
60,1 200	5,3	335	2 508	98,3
200 et au-dessus		28	2 536	100,0

D'après la colonne de % cumulés du tableau ci-dessus, la limite supérieure (20) est atteinte par près de 50 % des exploitations appartenant aux trois classes de « petites exploitations de 5 à 20 ha, et est dépassée par un peu plus de 50 % des exploitations (des classes de SAU excédant 20 ha).

Portons en abscisses sur échelle logarithmique, de gauche à droite, les limites supérieures de classes (2,66; 5; 20; 60; 200) de la colonne (x) du tableau : portons de bas en haut sur échelle de probabilité normale, en ordonnées, les % cumulés du tableau : nous déterminons 6 points par lesquels nous pouvons faire passer une droite de régression qui intercepte le niveau 50 % à l'aplomb de la SAU correspondant à $x = 17$: le paramètre de position ($x = 17$) de la distribution des 2 536 exploitations, correspond à la « moyenne géométrique, et à la médiane »; la pente de la droite de régression, évaluée par l'angle de cette droite avec l'horizontale est d'autant plus faible que la dispersion des SAU est plus grande : le graphique 4 suggère que 16 % des exploitations ont une SAU inférieure à 5 ha, et 16% une surface supérieure à 60 ha : sur échelle log des SAU, la moyenne géométrique qui correspond à la

GRAPHIQUE 4



valeur de la médiane qui est celle du quantile $Q_{50\,\%} = 17$ se situe au milieu de l'intervalle $Q_{16\,\%} = 5$ et $Q_{84\,\%} = 60$.

Les tables des log naturels ou $\log_e$, indiquent :

$\log_e 17 = 2\ 833$; $\log_e 5 = 1\ 609$ et $\log_e 60 = 4\ 094$, on peut donc écrire $\log_e 17 - \log_e 5 = \log_e 60 - \log_e 17 = 1{,}25$. La droite de régression est définie par les deux paramètres : de position ($2{,}83 = \log_e 17$) et de dispersion ($\pm\ 1{,}25$).

On désigne par les lettres grecques μ et σ les valeurs théoriques des paramètres de position et de dispersion de la distribution mathématique la plus vraisemblable pour représenter une distribution « observée ».

Selon le modèle de la distribution log normale le paramètre de position correspond à $e^{\mu} = Q_{50\,\%}$ où $Q_{50\,\%}$ représente le « Quantile 50 % » c'est-à-dire le centre de symétrie de la distribution, telle que 50 % des variantes soient inférieures à e^{μ}.

Par rapport à ce centre de symétrie on définit les quantiles « inférieur » $Q_{16\,\%} = e^{\mu-\sigma}$ et « supérieur » $Q_{84\,\%} = e^{\mu+\sigma}$ situés par rapport à $Q_{50\,\%}$ symétriquement aux distances $\pm\ \sigma$.

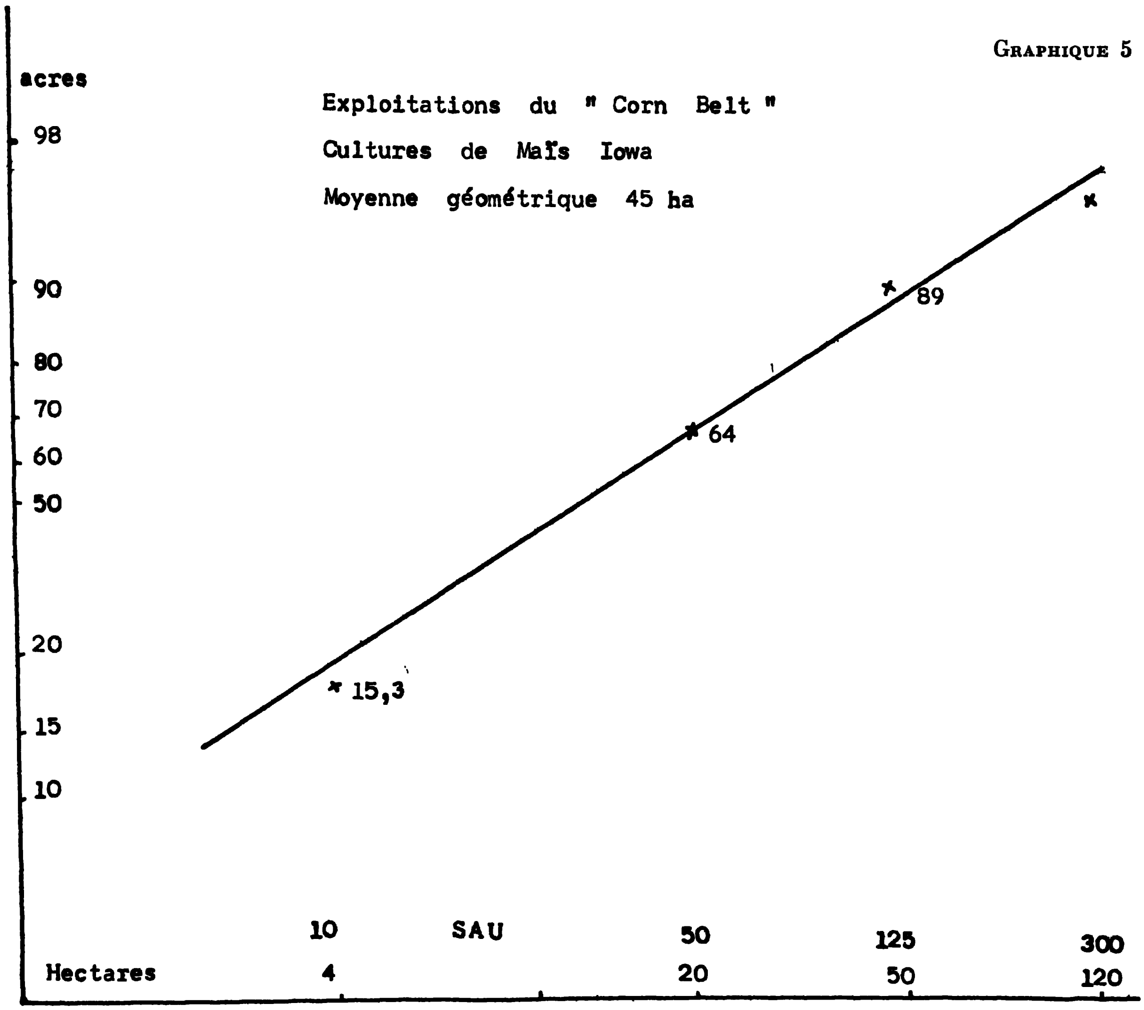
Chacune de ces 3 relations définit un point sur le graphique où l'on porte en abscisses les limites de classes (x_i) et en ordonnées la fréquence ($f(x_i)$) dans la classe (0 à x_i).

Sur une échelle log des abscisses les limites supérieures des classes sont définies par $\log x_i$; le paramètre de position correspond à la moyenne arithmétique des ($\log x_i$); ce paramètre de position ou centre de symétrie se situe sur l'échelle log des abscisses à la valeur qui correspond à l'estimation la plus vraisemblable de μ, la verticale élevée de cette valeur intercepte le niveau 50 % de l'échelle des ordonnées au point qui définit, sur la droite de régression, la position du $Q_{50\,\%}$. Les verticales élevées depuis l'échelle log des abscisses, à partir de $-\sigma$ et de $+\sigma$ coupent la droite de régression aux niveaux 16 % et 84 % et définissent les points correspondant à $Q_{16\,\%}$ et $Q_{84\,\%}$.

Les 3 points correspondant à $Q_{50\,\%}$ à $Q_{84\,\%}$ définissent chaque distribution log normale.

Qu'il s'agisse de pays comme le Japon, où l'unité de SAU est le tan (ou 10 ares) et où *K. Izumi* a étudié des exploitations dont les SAU se distribuent autour de $\mu = 13{,}7$ tans soit, 1,37 ha, ou d'une région des Indes, où 50 % des familles disposent de moins de 2 ha, ou d'une région d'exploitations familiales en France, avec des moyennes géométriques de distribution de SAU de l'ordre de 9 à 10 ha dans la Bresse, ou dans la Charente-Maritime ou dans la région viticole de la Gironde, ou, aux États-Unis, de régions de culture de maïs de l'État de Iowa (moy. géom. 45 ha) ou de cultures cotonnières en Louisiane, le modèle le plus vraisemblable pour chaque distribution est le modèle log normal; chaque distribution peut se représenter, sur papier log normal, par une droite, définie par la moyenne géométrique correspondant au $Q_{50\,\%}$ et par la position des $Q_{16\,\%}$ et $Q_{84\,\%}$.

Graphique 5



ÉTUDE DES STRUCTURES

La distribution des fréquences $f(x)$ d'exploitations agricoles selon la surface (x) de l'exploitation peut être évaluée en fonction de « surface géographique » ou de « SAU », ou de surface consacrée à une certaine culture principale » telle que « Vigne »; les valeurs observées de (x) sont groupées par classes de fréquences, chacune définie par une limite inférieure et une limite supérieure (M. Mussotte, *C. R. Ac. Agric.* p. 1275, 1963).

1° Théoriquement la limite inférieure de la classe des plus petites valeurs de x_i serait zéro, et dans les classifications figure généralement la classe des surfaces (x_i) inférieures à 10 hectares, en France ou à 10 acres aux États-Unis : exception faite de régions à cultures hautement spécialisées, maraîchères, florales ou viticoles, la classe des (x_i) inférieure à 10 ha peut comprendre une prédominance de manifestations d'activités qui ne sont que subsidiairement agricoles; il s'agit de « résidences » plutôt que d'exploitations à type rural.

2° La distribution des fréquences de (x_i) varie selon les régions pour des raisons historiques, par exemple du fait de la persistance de reste de grands domaines (latifondia).

3° La distribution des fréquences (f_{x_i}) dépend essentiellement de la nature des sols.

La comparaison des distributions (f_{x_i}) en fonction des facteurs géologiques et historiques sera proposée ici sous forme de droites de régression, tracées chacune à travers le nuage de points de coordonnées (x, f_x) où x représente le centre d'une classe, sur l'échelle logarithmique des abcisses, et (f_x) le résultat de la transformation en pourcentages cumulés, portés en ordonnées sur l'échelle de probabilité normale, des fréquences d'exploitations ayant au plus une surface x. Bien entendu on peut cumuler soit à partir de la classe des plus petites surfaces, soit à partir de la classe des plus grandes surfaces.

Dans les exemples qui suivent (sauf indication), nous avons cumulé à partir de la classe des surfaces inférieures à 10.

DISTRIBUTION DES SAU EN FRANCE

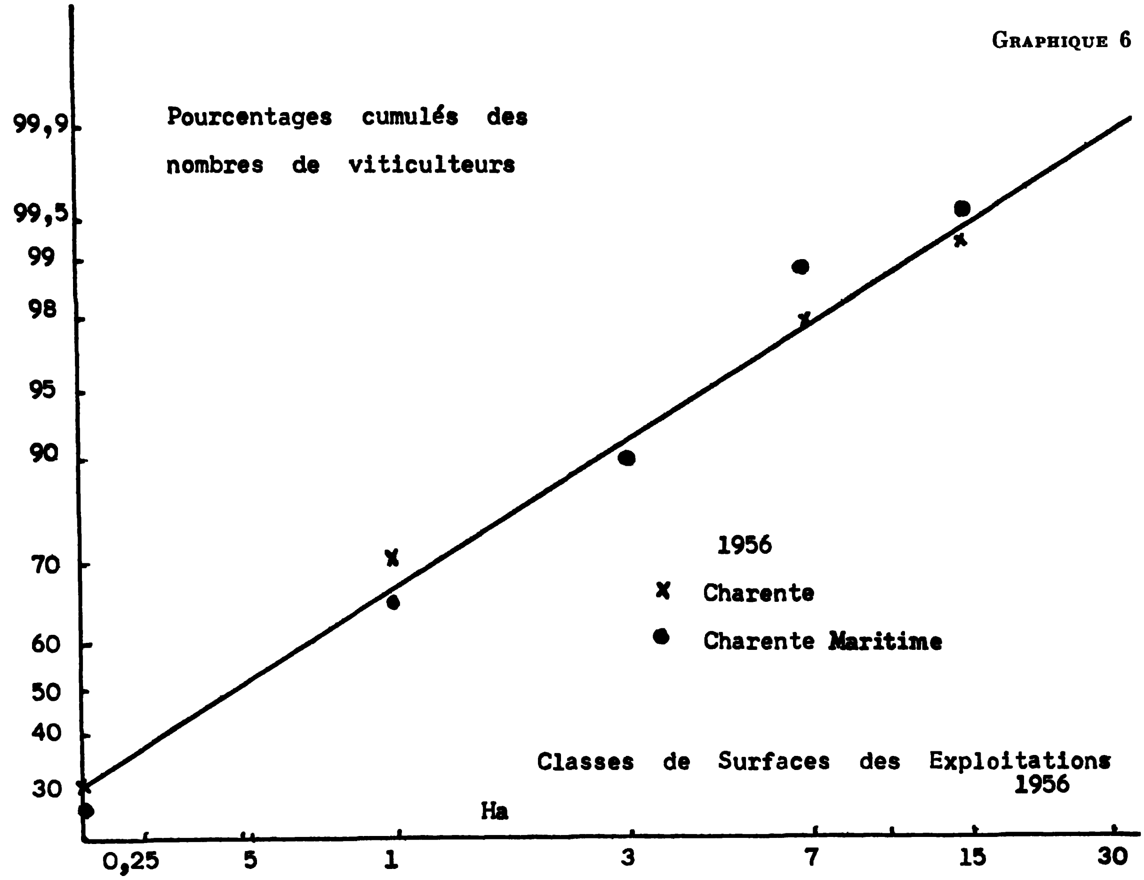
Le tableau I indique les pourcentages cumulés des exploitations de moins de (x) ha (1) pour l'ensemble de la France, d'après *France-Agriculture*, page 37, 1963, pour les départements de l'*Oise* (2), du *Maine-et-Loir* (3) et de *Charente-Maritime* (4) d'après le *Bulletin d'information* du C. N. E. M. A., juillet-août 1963; d'après une enquête effectuée en Charente-Maritime (utilisant les résultats du recencement de 1955) et enfin pour la Gironde.

TABLEAU I

x ha	1	2	3	4	4 *b*	5
Moins de 10	55,8				54,5	76,4
10 à 20	79,3	18,63	34	47	78,1	92,7
20 à 50	95,8	61,50	91	86	97,4	99,8
50 à 100	99,1	85,00				100,0
100-200	99,8					

La comparaison des colonnes 4 et 4 *b* confirme que les chiffres sont sujets au caution; néanmoins en cumulant les chiffres et en les transformant en pourcentages, on obtient l'information qui permet de comparer une distribution de fréquences observées à une distribution théorique, calculée selon un certain modèle mathématique; (modèle de distribution log normale) : exemple graphique 6.

GRAPHIQUE 6



Viticoles :

Répartitions des exploitations Cadastre viticole 1956

	Charente		Charente-Maritime	
x	$f(x)$ cumulés	%	$f(x)$ cumulés	%
0,25	9 477	30	9 298	26,4
0,99	23 597	76	24 757	69
6,99	28 225	90	32 456	90
14,99	30 403	99,3	35 641	99,5
29,99	30 537	99,8	35 689	99,9
plus de 30	30 569	100	35 697	100

50 % des exploitations ont moins de 5 ha

f_x nombres des viticulteurs dont l'exploitation a moins de x ha.
$f(x)$ cumulés transformées en pourcentages cumulés

Facteurs géographiques affectant les distributions de SAU. Pour une même spéculation végétale, et dans une même région, la distribution des SAU est affectée par la nature du sol; par exemple E. O. Heady (*l. c.*) a étudié comparativement les SAU et les fonctions de coût de production du maıs dans l'Est de l'État de Iowa, sur sol riche, à faible pente, où l'on peut cultıver maıs sur maıs, et dans l'Ouest, sur les collınes voisines du Mississipi, où les risques d'érosion exıgent l'ıntroduction de prairies.

Pour la même région de production de coton de Louisiane, les distributions de fréquence de SAU dıffèrent pour la plaine alluviale et pour les terrasses qui surplombent la vallée.

TABLEAU II

*Distributions des pourcentages cumulés des exploitations (« farms ») couvrant au plus (*x*) acres dans la région de culture du coton, en Louısiane (1959) sur « terrasses » et sur « alluvions »; on indique aussi les pourcentages cumulés de l'ensemble du sol couvert par les exploitations cultivant du coton sur terrasses et sur alluvıons* (Louisiana Rural Econ. 25 (1) feb. 1963).

x acres (1)	Terrasses			Alluvıons				
	Exploıtatıons		Sol		Exoloıtatıons		Sol	
0 à 99	81,5	81,5	[33,9	33,9	75,6	75,6	14,8	14,8
100 à 259	12,8	94,3	21,9	55,8	13,5	89,1	12,9	27,7
260 à 499	3,3	97,6	12,7	68,5	4,5	93,6	10,0	37,7
500 à 999	1,6	98,2	12 5	81,0	3 3	96,9	14,1	51,8
+ de 1 000	0,8	100,0	20,0	100,0	3,1	100,0	48,2	100,0
	100,0				100,0			

1 1 acre = 0,404 ha

TENDANCES

Dans un comté de l'État de Californie on note de 1939 à 1964 une augmentation relative des fréquences dans la classe de moins de 10 acres, comportant une prédominance des domaines utilisés à des fins non spécifiquement agricoles, mais on observe surtout une augmentation dans les classes de surfaces supérieures à 100 acres.

TABLEAU III

Colonne S_t, classes de surface (en acres) d'exploitations dans un Comté de Californie; fréquences d'exploitations par classes de surface, en 1939-1949 et 1954. Pourcentages cumulatifs d'exploitations d'au moins S_t acres.

S_t	Fréquences			Pourcentages cumulés	
	1939	1949	1954	Fréquence 1954	% 1954
inf à 10	75	134	125	125	7
10 à 29	301	354	332	457	25
30 à 49	278	342	293	750	40
50 à 69	132	164	146	896	50
70 à 99	191	214	170	1 066	59
100 à 139	106	136	150	1 216	67
140 à 179	100	115	131	1 347	69
180 à 219	37	47	56	1 403	77
200 à 259	35	39	37	1 440	80
260 à 499	114	119	121	1 561	86
500 à 999	63	75	94	1 655	92
+de 1 000	106	161	151	1 806	

TABLEAU IV

Surfaces inférieures a (x) acres, colonne 1, en 1954

x	F (x)	f (x) cumulés	% cumulés
10	125	125	7
30	332	457	25
50	293	750	40
70	146	896	50
100	170	1 066	59
140	150	1 216	67
180	131	1 347	69
220	56	1 403	77
260	37	1 440	80
500	121	1 561	86
1 000	94	1 655	92
	151	1 806	

ÉVOLUTION DES DISTRIBUTIONS DE SURFACE

Le tableau V indique de 1939 à 1954 une augmentation relative des fréquences de très petits domaines (moins de 10 acres, c'est-à-dire moins de 4 ha) : il ne s'agit plus « d'exploitations agricoles » mais de domaines acquis comme « résidences » par des non-agriculteurs, le plus souvent à des fins fiscales. En ce qui concerne les « propriétés agricoles » on constate, de 1939 à 1954 une augmentation des proportions relatives des domaines de plus de 100 acres, et surtout des domaines de plus de 1 000 acres.

Le tableau V indique l'évolution de la distribution des exploitations, par classes de surfaces (en acres) dans l'État de Iowa, de 1954 à 1959, la colonne 1 indique les classes de surfaces en acres; les colonnes 2 et 3 les fréquences *dans chaque classe* en 1954 et 1959, et la colonne 3, les pertes en %, dans les classes 10 à 179, et les gains en % dans les classes supérieures à 180 acres : le tableau est extrait du *Research bull.* 502, *(Iowa Agric. Expt. station, Ames, Iowa)* consacré à l'étude des modifications apportées par le regroupement *(Farm consolidation)* dans la distribution des ressources et des revenus.

TABLEAU V

Acres	1954	1959	%
10- 49	14 402	13 727	— 4,7
50- 69	4 338	3 912	— 9,8
70- 99	18 244	14 647	— 20,7
100-139	24 923	19 590	— 21,4
140-179	45 564	37 404	— 17,9
180-259	12 152	20 128	— 9,2
220-259	20 659	20 685	+ 0,1
260-499	29 960	34 342	+ 14,6
500-999	3 284	4 475	+ 36,3
+ 1 000	271	345	+ 27,3

L'étude des distributions de surface et l'évolution de ces distributions conduit à rechercher quel peut être, dans une région déterminée, l'optimum de la surface pour une exploitation.

Recherche de l'optimum

Le tableau VI indique : colonne 1 les classes de surfaces (en acres); colonne 3, la superficie totale en acres, pour les exploitations d'une classe; colonne 4 la SAU moyenne d'une exploitation dans cette classe; colonne 3, la surface labourable; 6 la surface pouvant être consacrée aux cultures sarclées.

En admettant que pour une exploitation d'une centaine d'acres un homme suffit, et qu'il faut 3 hommes dans une exploitation de 300 acres, on calcule (colonne 7) les nombres d'heures nécessaires par classe d'exploitation : la classe des surfaces de 10 acres est indiquée pour « mémoire », elle ne correspond pas à des exploitations agricoles, mais plutôt à des résidences.

TABLEAU VI

SAU (acres) 1	% cumulatifs 2	Total (acres) 3	Moyenne (acres) 4	Labour (acres) 5	sarclées (acres) 6	Heures nécessaires 7
0 à 9,9	15,3	3 584				
10 à 49,9	64,0	12 586	32	28,9	22,2	466
50 à 124,9	89,0	6 245	80	72,4	55,5	2 525
125 à 299,9	96,0	1 751	210	190,0	145,7	2 605
300 et +	100,0	747	635	574,4	440,7	6 890
		24 913				

Sur une exploitation de moins de 50 acres un homme n'utilise que 466 heures, ce qui laisse disponibles 576 heures (1).

1. Earl O. HEADY (Dept. of Agronomy Agric. and Home Economics, Expt. Station, Ames, Iowa). Production functions and methods of specifying optimum, *Research Bull.*, 518, july 1963.

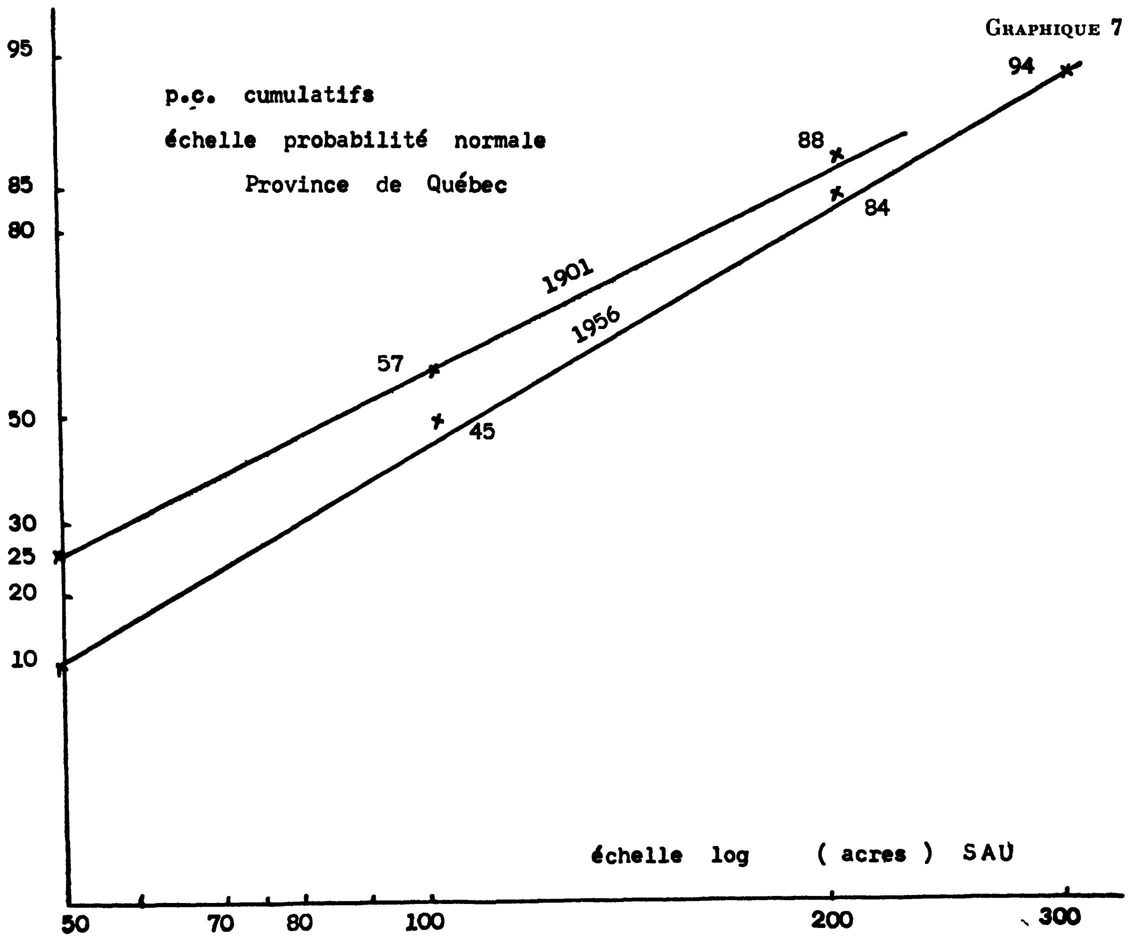
Moyenne arithmétique ou moyenne géométrique

Nous lisons dans le *Bulletin d'Information du C. N. E. E. M. A.* n° 72, p. 5, janv. 1964 : « L'Agriculture des États-Unis... est largement dimensionnée, puisque l'exploitation moyenne y est caractérisée par : 80 ha pour les exploitations maïsicoles, 250 ha pour les exploitations céréalières »; si la « moyenne » invoquée est la moyenne arithmétique, cette assertion laisse entendre qu'on a choisi le modèle « normal » pour représenter les distributions de fréquences de surface, pour chaque type de culture (maïs ou céréales telles que blé); selon ce modèle, la classe de la moyenne est aussi celle du mode et celle de la médiane. Il en est tout autrement pour la distribution log normale.

D'après E. O. Heady and R. D. Krenz (*Agric. Exp. Stat. Iowa State Univ., Ames, Res, Bull.* 504, may 1962) dans une région spécialement favorable à la culture du maïs, une exploitation doit pouvoir utiliser à plein temps au moins 2 travailleurs, fournissant de 4 000 à 4 500 heures de travail par an, avec un tracteur par unité de travailleur et par 2 000 heures de travail.

Le coût de production (en $) de la valeur (en $) d'une unité de production, suit une courbe qui descend, de 1 pour une très petite SAU, à 0,55 pour 40 ha et moins de 0,5 pour 80 ha; le moindre coût de production s'obtient entre 120 et 260 ha; l'optimum étant voisin de 120 dans les régions à climat moins favorable (150 jours de travail par an) et atteignant

Graphique 7



190 dans les régions où l'on peut prévoir 200 jours de travail; l'optimum de SAU étant d'ailleurs d'autant plus grand que l'équipement permet d'effectuer sans délai les travaux nécessaires : cependant quel que soit l'équipement, le coût de production tend à augmenter au dessus de 320 ha.

Dans la Province de Québec où en 1901 on dénombrait plus de 140 000 exploitations réparties selon une distribution log normale autour d'une moyenne géométrique de 80 ares (environ 33 ha), le nombre s'était réduit en 1956 à 122 617; la réduction porte surtout sur les exploitations de moins de 50 acres (20 ha), dont le pourcentage décroît de 24 % à 10 %; la moyenne géométrique passe de 80 à 105 acres, mais la distribution reste log normale (graphique 7).

M. L. Dufrénoy et J. Dufrénoy