

TD : Interprétation de graphiques :

étude des inégalités de répartition des revenus des ménages à l'aide de graphiques. Les données, tableaux et graphiques proviennent de l'étude de l'INSEE : « les revenus et patrimoines des ménages, édition 2006 » d'après l'enquête 2003-2004.

I. Un graphique de base : le regroupement par classe des données

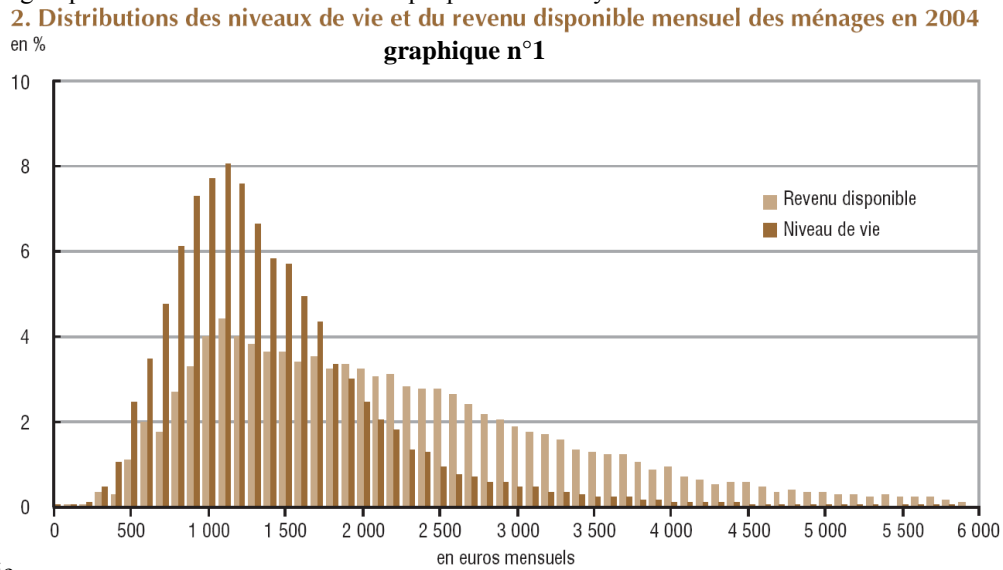
Les données collectées lors d'une enquête sont regroupées en vue d'une étude statistique pour être analysées et commentées.

Les données de la série sont regroupées dans un tableau des effectifs ou des fréquences puis représentées sur un graphique en vue d'une interprétation.

questions de lecture de graphiques:

Quel est le pourcentage des ménages qui ont un revenu disponible mensuel entre 3000 et 3100 € ?

Que peut-on dire de la répartition des revenus disponibles des ménages en 2004 ?



Surligner l'histogramme correspondant à la série des revenus disponibles des ménages

Note : au-delà de 6000 euros, les distributions ne sont pas représentées.
Lecture : le pas de l'histogramme est de 100 euros ; la hauteur de la barre de coordonnées n en abscisse est donc égale à la proportion de revenus ou niveaux de vie compris entre n et $n+100$ euros : ainsi 5,7 % des individus ont un niveau de vie mensuel compris entre 1 500 euros et 1 600 euros et 3,6 % des ménages ont un revenu disponible compris entre 1 500 euros et 1 600 euros.
Source : Insee-DGI, enquête Revenus fiscaux 2004.

II- caractéristiques de dispersion d'une série statistique : médianes , quartiles , déciles .

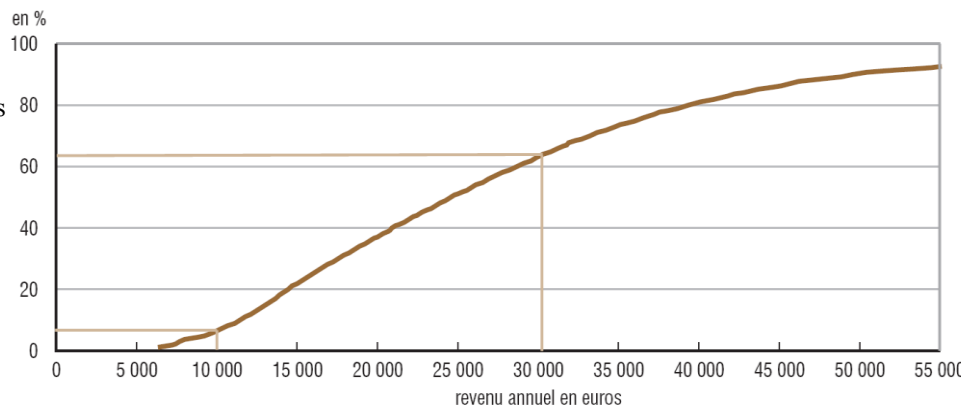
Pour mettre en évidence les caractéristiques de dispersion de la série statistique, on détermine la médiane, les quartiles (et les déciles).

Pour cela, on établit le tableau des fréquences cumulées croissantes ou des effectifs cumulés croissants (obtenu directement à l'aide de logiciels lorsqu'il y a de nombreuses données).

Puis on construit un graphique pour faciliter l'interprétation

A l'aide de la lecture du graphique ci-contre, compléter le tableau suivant :

1. Répartition des revenus disponibles des ménages en 2004
graphique n°2



Champ : ménages dont le revenu déclaré au fisc est positif ou nul et dont la personne de référence n'est pas étudiante.
Lecture : 57 % des ménages ont en 2004 un revenu disponible annuel compris entre 10 000 et 30 000 euros.
Source : enquête Revenus fiscaux 2004, Insee-DGI.

Premier décile D1	Premier quartile Q1	médiane	Troisième quartile Q3	Dernier décile D9

Construire le diagramme en boîte appelé aussi « boîte à moustache » associé à cette série statistique

Quel est le pourcentage des ménages qui ont un revenu disponible supérieur à 40 000 € ?

Représenter D1, D9, Q1 , la médiane et Q3 sur le graphique n°1 (on divisera par 12 les valeurs obtenues dans le tableau ci dessus). Comment les interpréter sur le graphique n°1 ?

III- Comparaison moyenne et médiane

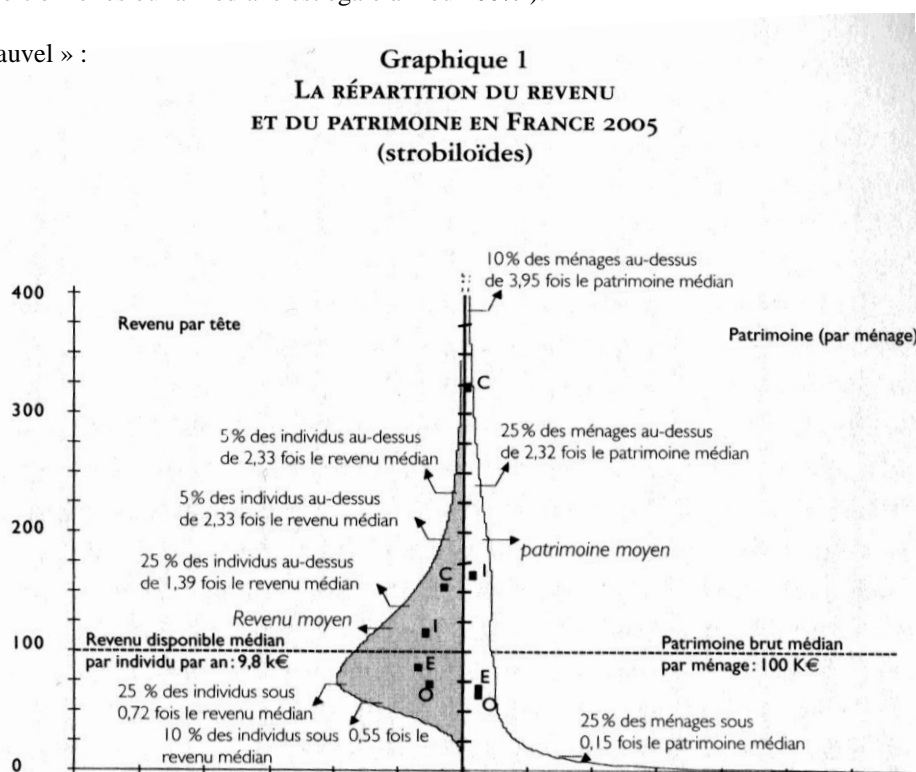
	2000	2001	2002	2002 rétropolée ¹	2003	2004
Revenu disponible médian (euros constants)	23 450	24 000	24 390	24 930	24 750	24 600
Revenu disponible moyen (euros constants)	27 600	28 200	28 480	29 070	29 010	28 340
Rapport interdécile (D9/D1)	4,5	4,4	4,4	4,4	4,4	4,3

Indiquer la position du revenu disponible moyen sur chacun des graphique : n°1 , n°2 et diagramme en boîte
Que signifie le constat d'un revenu moyen supérieur au revenu médian ?

IV- Différents indicateurs utilisés en SES:

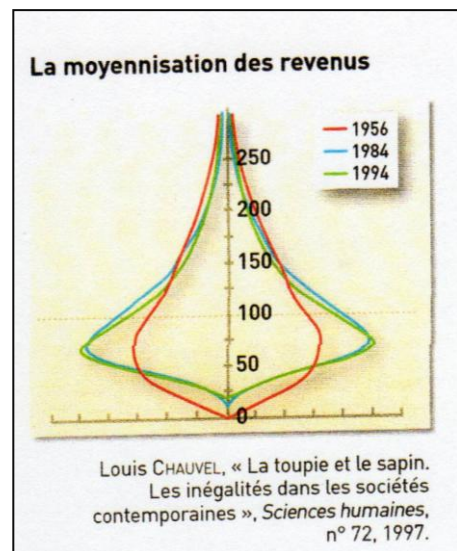
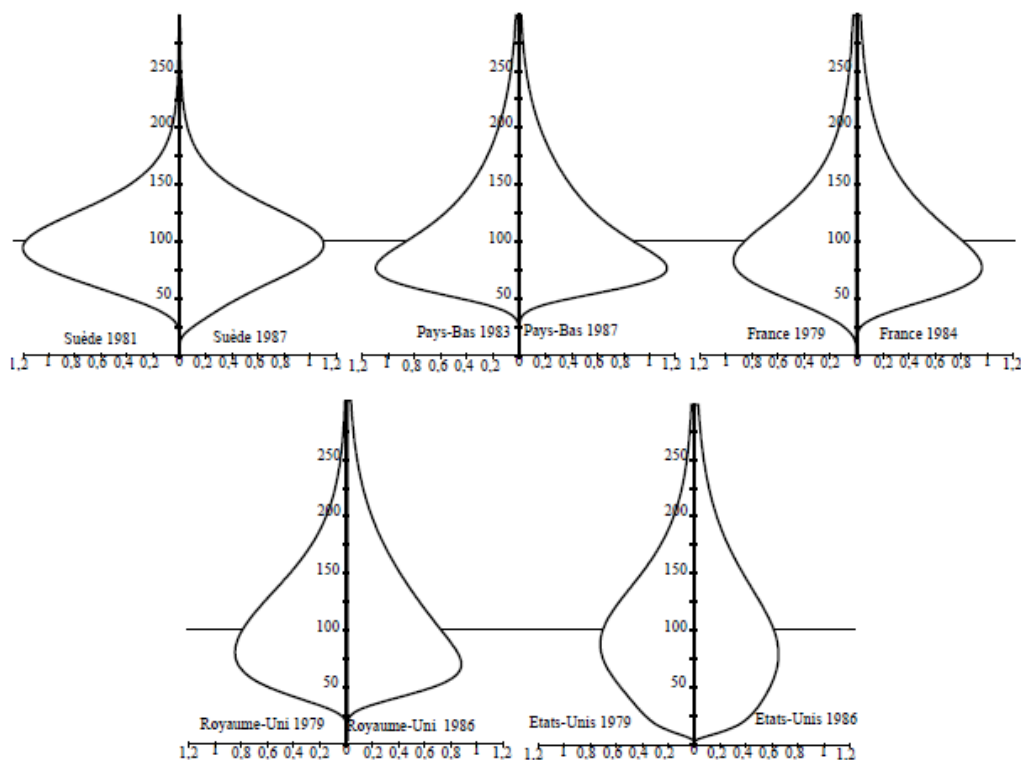
Pour observer la dispersion de cette série statistique et pouvoir comparer avec d'autres (soit avec des séries de revenus disponibles d'autres pays ou avec des séries d'autres dates ou de comparer deux types de grandeurs à la même date et dans le même pays) , on peut utiliser plusieurs indicateurs : Q3-Q1, D9/D1 , construction des courbes de Lorenz associées et le coefficient de Gini (voir TD sur la construction de telles courbes) , mais aussi l'analyse de « strobiloïdes » construits avec des séries médianisées (c'est-à-dire avec des valeurs proportionnelles où la médiane est égale à 1 ou 100%).

Le strobiloïde dit de « Chauvel » :



Note: le strobiloïde représente la forme de la pyramide sociale correspondant à la répartition du revenu ou du patrimoine. La courbe (zones grisées) est d'autant plus large qu'une part importante de la population est située précisément à ce niveau. Si 100 représente la médiane des revenus, le large renflement au centre du strobiloïde permet de révéler une forte classe moyenne située à égale distance des extrêmes. Du côté des patrimoines, au contraire, il n'existe pas de classe moyenne dans la mesure où la population est largement étirée entre l'extrême opulence et l'extrême dénuement en termes d'accumulation. Les revenus sont exprimés en francs par tête dans le ménage: l'individu médian se trouve dans un ménage où le revenu annuel est de 9,8 K€ par individu. Le patrimoine est exprimé par ménage. Pour le revenu, C, I, E et O représentent la médiane du revenu respectivement des cadres, des professions intermédiaires, des employés et des ouvriers. Pour le patrimoine, ce sont les moyennes, les médianes n'ayant pas été calculées dans les publications de l'INSEE.

Pour autant, entre la moitié et le double du revenu médian, on compte 80% de la population. Entre les quatre cinquièmes du revenu médian (médiane -20%) et les cinq quarts (médiane +25%), 34%. Alors que les titulaires des plus bas revenus se raréfient rapidement à mesure que l'on descend en deçà de 75% du revenu médian, tout comme les plus riches lorsque l'on s'élève dans l'échelle des revenus, nous avons à ce niveau intermédiaire de la répartition du revenu une certaine accumulation de la population: la « classe de revenu médian ».



O.F.C.E.

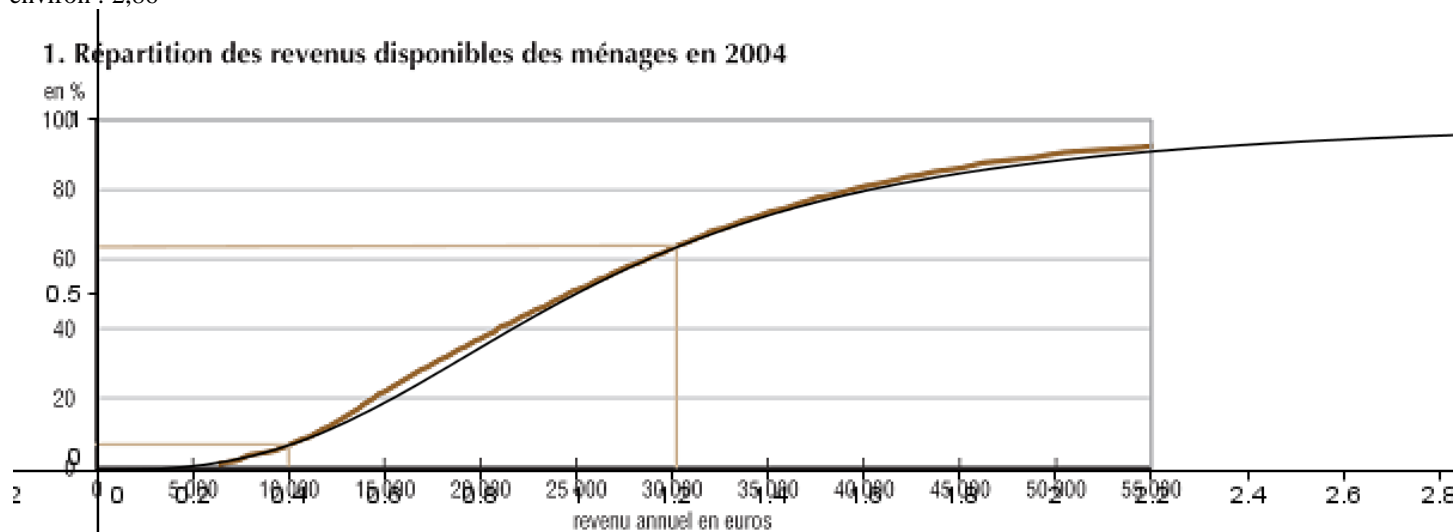
69, quai d'Orsay, 75007 PARIS - Tel: 44. 18. 54. 31.

V- Modélisation mathématique du strobiloïde des revenus

L'allure des séries des revenus a été l'objet d'étude de mathématiciens afin d'en dégager des indicateurs. Par exemple, Pareto au XIXème mais n'avait que des données tronquées (difficulté d'obtenir des données), puis Champernown etc Ces études débouchèrent sur une étude mathématique générale des séries statistiques.

1 - Prenons par exemple une courbe de Champernown du type $F(R) = 1 - \frac{1}{1+R^\alpha}$ pour modéliser le graphique n°2, avec R le revenu médianisé, c'est-à-dire que la valeur de la médiane est ramenée à 1 (on divise toutes les données par la valeur médiane). L'avantage est de pouvoir comparer la forme de différentes séries mais en perdant les informations dues aux inégalités des différentes valeurs médianes observées.

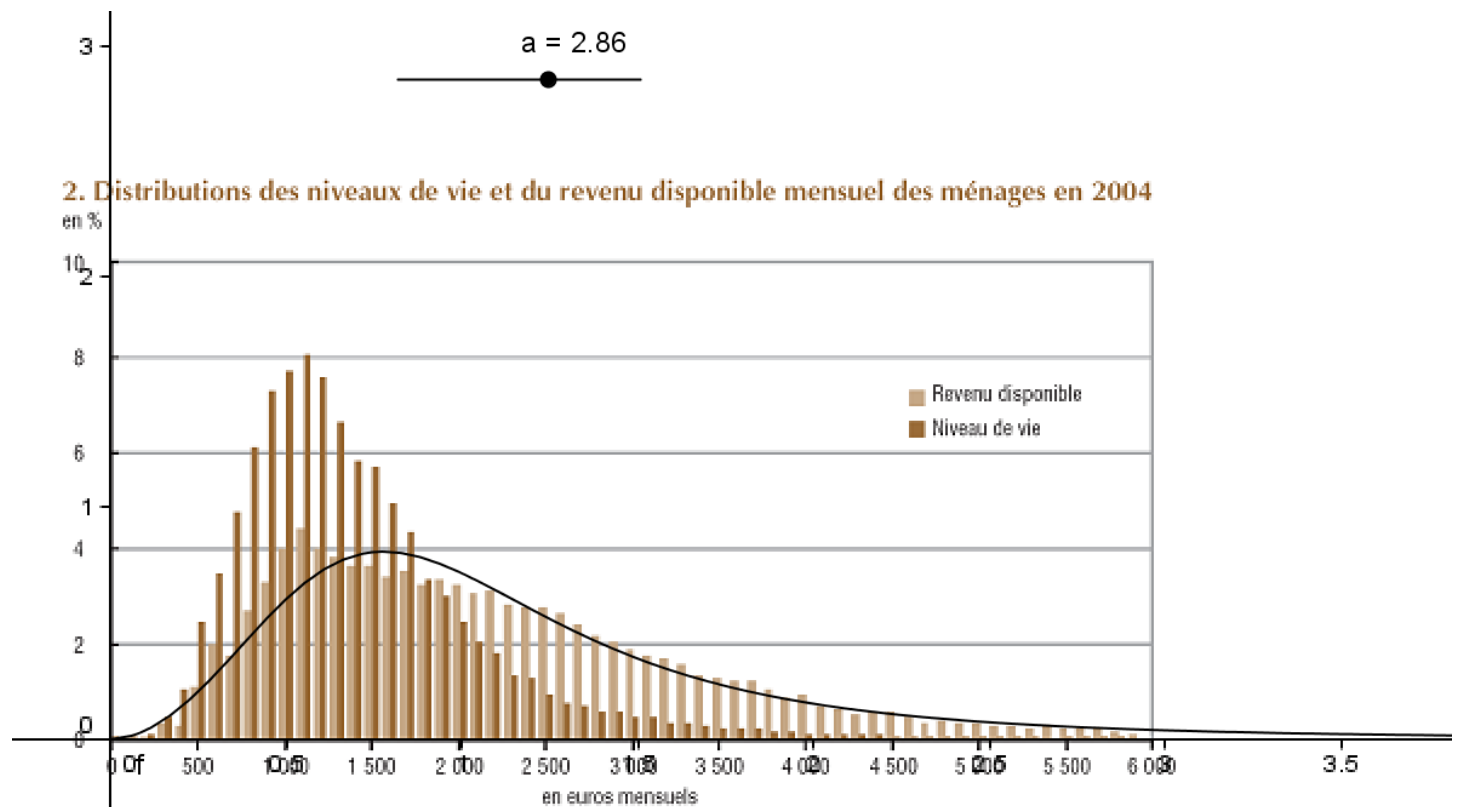
A l'aide du logiciel geogebra on observe que pour la série des revenus étudiée précédemment, une valeur de α possible est environ : 2,86



Cette courbe vérifie $F(0)=0$ et $\lim_{R \rightarrow +\infty} F(R) = 1$, elle correspond à la fonction de répartition des revenus, celle des revenus cumulés.

2- La courbe du graphique n°1 correspond à la fonction densité des revenus et peut être modélisée par la dérivée de la fonction F :

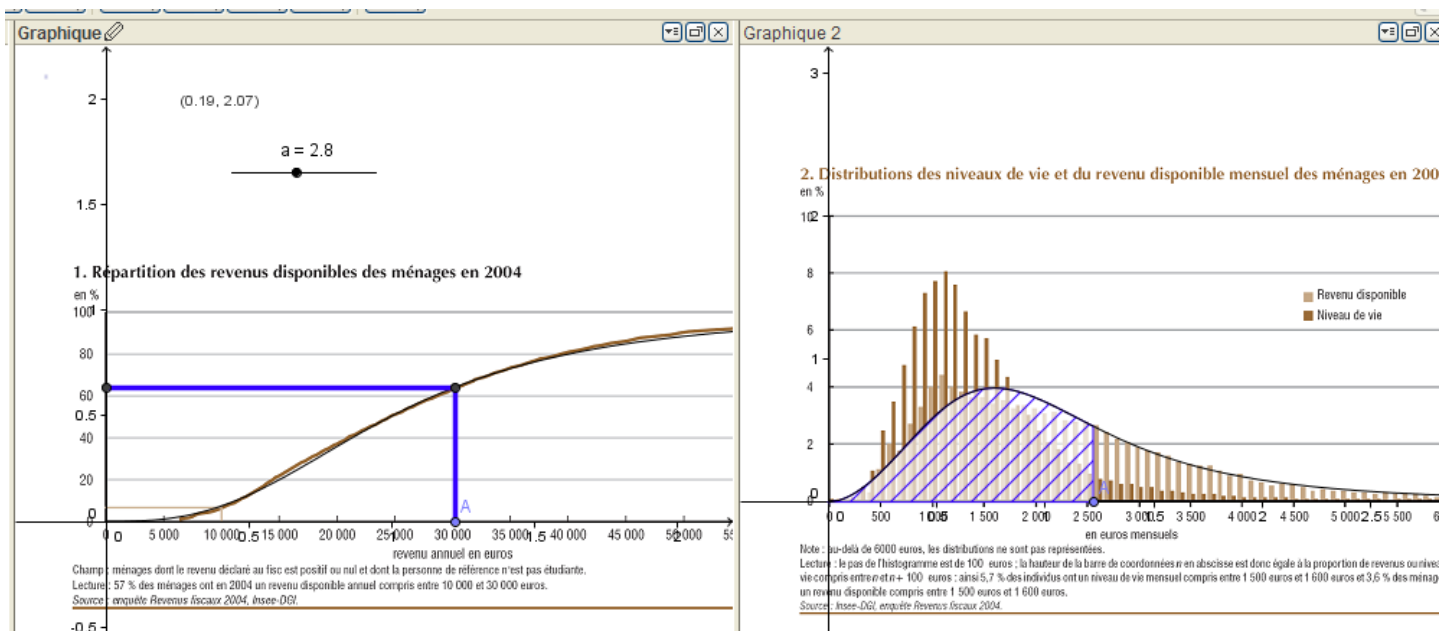
$$F'(R) =$$



L'aire située sous la courbe est bien égale à 1 c'est-à-dire à 100% des revenus (voir interprétations graphiques précédentes ou les effectifs , ie les fréquences sont interprétées en termes d'aire) et c'est pour cette raison que les unités ne sont pas respectées sur l'axe des ordonnées (aire de tranches de revenus de largeur infiniment petite).

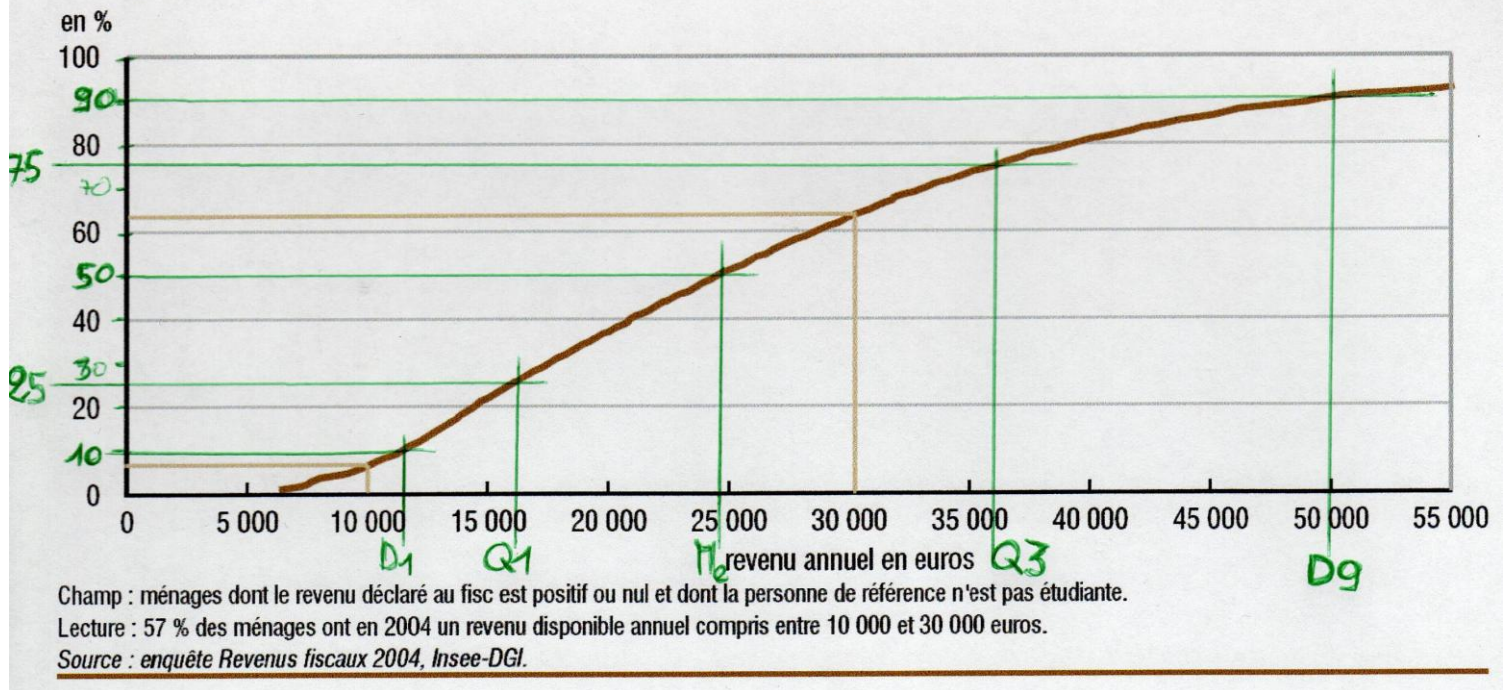
On retrouve ce résultat par le calcul intégral en interprétant les deux courbes avec la notion d'aire : $F(r) = \int_0^r F'(x)dx = p(x < r)$ c'est-à-dire la proportion des revenus disponibles inférieur à r.

$$\lim_{r \rightarrow +\infty} \int_0^r F'(x)dx = \lim_{r \rightarrow +\infty} F(r) - F(0) = 1 - 0 = 1 \quad \text{noté} \quad \int_0^{+\infty} F'(x)dx = 1$$



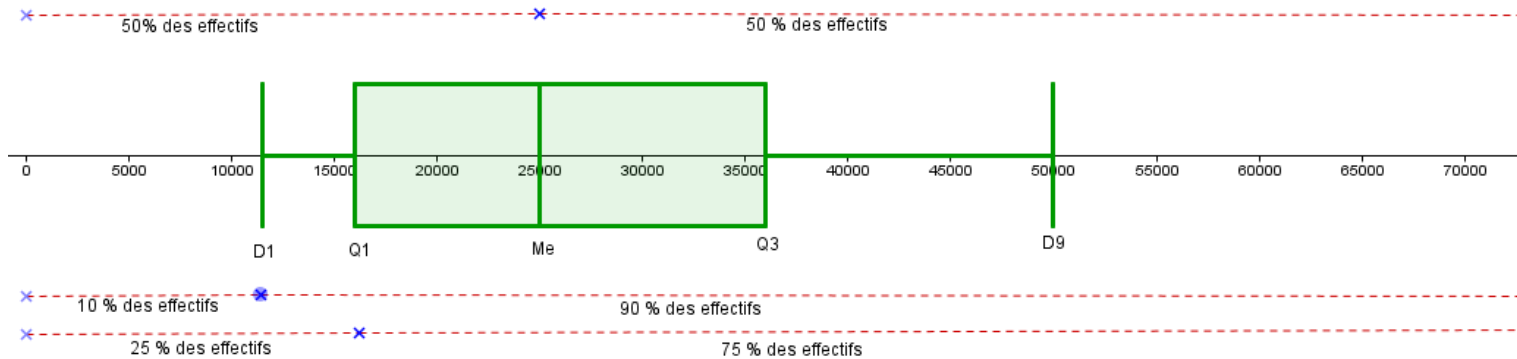
3- La modélisation des courbes de revenus a intéressé le sociologue Louis Chauvel pour en dégager des paramètres qui pourraient permettre des comparaisons fines et exploitables en termes d'aide à la décision sur les politiques de réduction des inégalités. Il utilise des fonctions plus compliquées que celles de Champernown avec trois paramètres permettant de prendre en compte les bas de la courbe du strobiloïde (les plus bas revenus) et les hauts (les hauts revenus) . Voir animation avec geogebra.

1. Répartition des revenus disponibles des ménages en 2004



Premier décile D1	Premier quartile Q1	médiane	Troisième quartile Q3	Dernier décile D9
11 500	16 000	25 000	36 000	50 000

Attention : D1, Q1, Me, Q3 et D9 sont du type des valeurs du caractère de la série statistique.



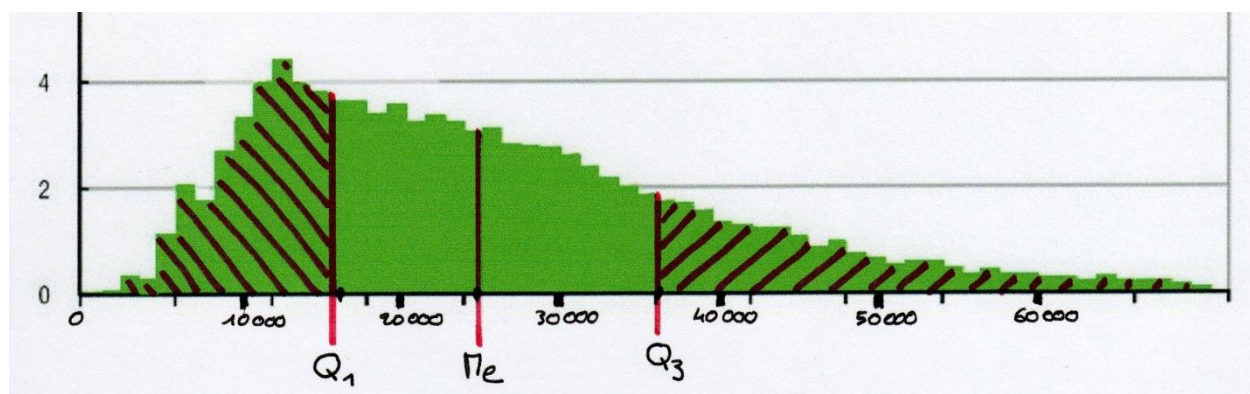
En math, on demandera souvent la valeur de l'écart interquartile : $Q3 - Q1 = 36000 - 16000 = 20000 \text{ €}$

Interprétation : l'écart des revenus annuels disponibles des ménages entre les 25 % les plus riches et les 25% les plus pauvres est supérieur à 20 000 €.

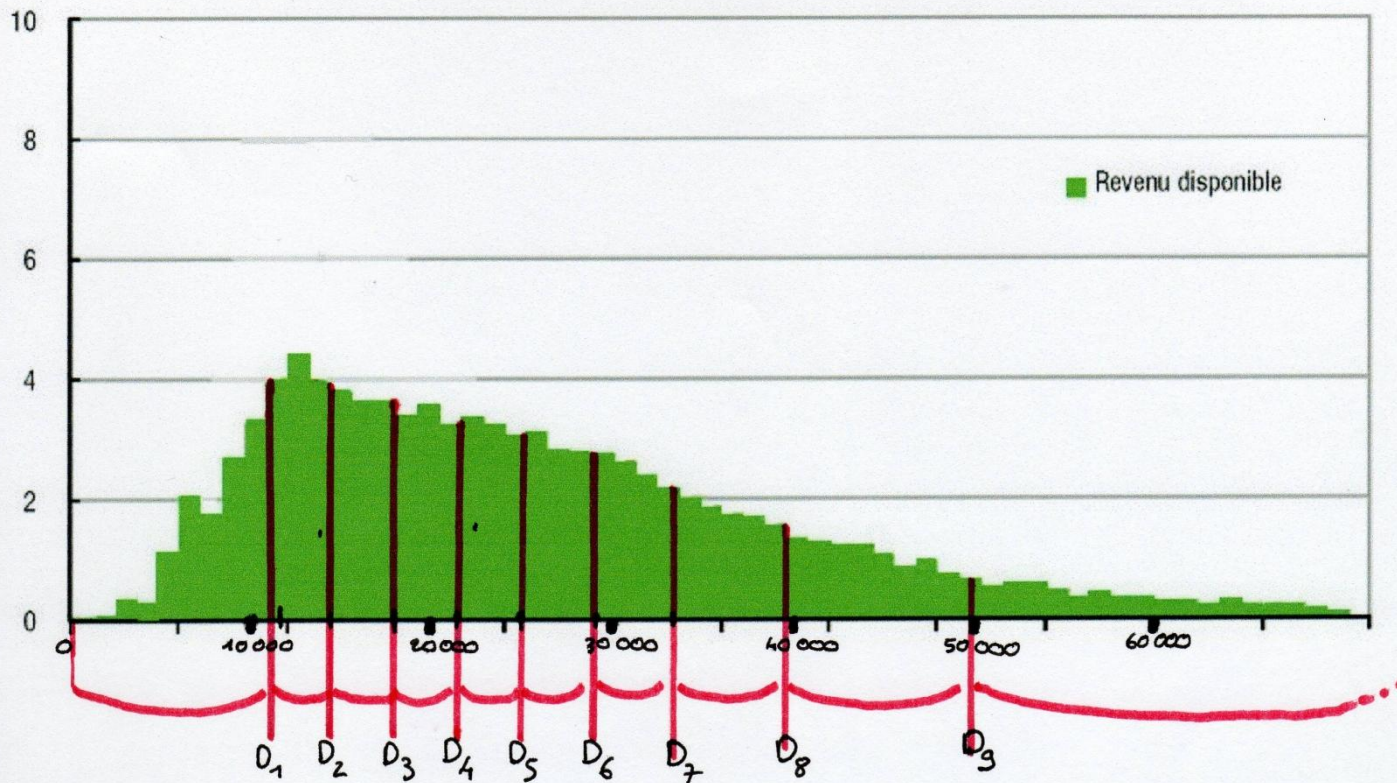
En SES, on demandera plutôt les rapports interdéciles : exemple $\frac{D9}{D1} = \frac{50000}{11500} \approx 4,3$

Interprétation : le revenu des 10% les plus riches est au moins 4,3 fois supérieur au revenu des 10 % les plus pauvres.

Vision des quartiles et l'écart interquartile sur l'histogramme des fréquences (graphique n°1). On a hachuré les 25% des effectifs des valeurs en dessous de Q1 et les 25% des effectifs des valeurs au dessus de Q3 (les aires des parties hachurées sont donc égales), 50 % des effectifs sont donc entre Q1 et Q3.

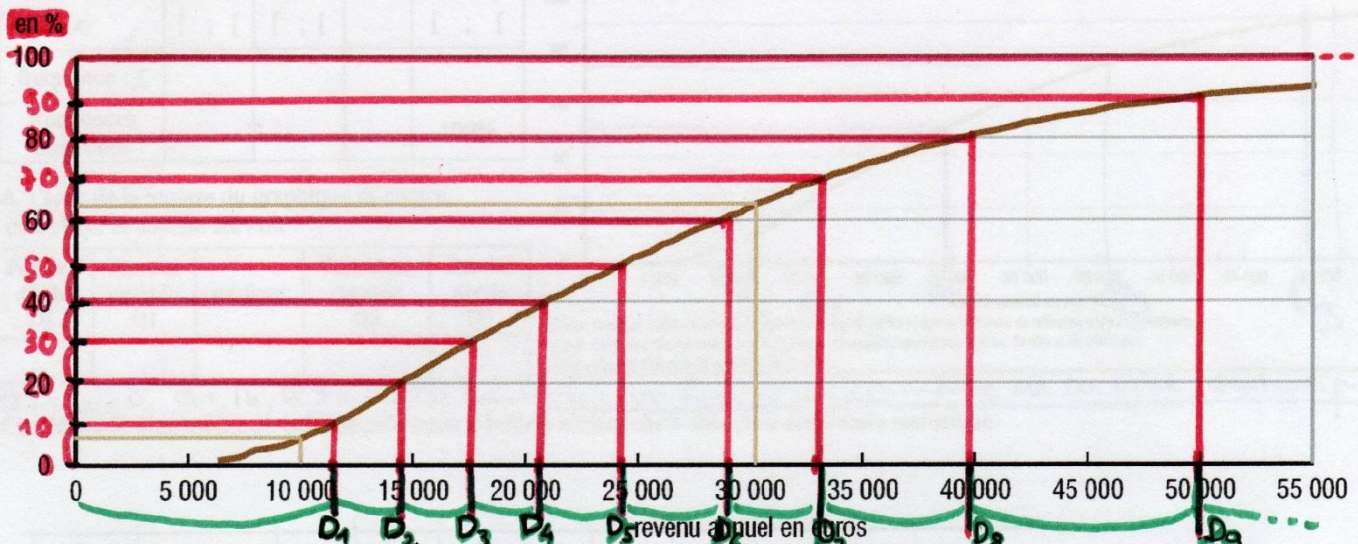


Remarque : en SES. On emploie aussi le terme de *décile* pour désigner les ensembles (au nombre de 10) des données situées entre deux déciles (déciles pointés) par exemple pour D6 : ensemble des valeurs de la série obtenues entre 50 % et 60 % des effectifs) . On obtiendrait pour D6 :] 25 000 ; 29 000]. Les 10 % des ménages ayant un revenu supérieur à celui des 50 % les moins riches et inférieur à celui des 40 % les plus riches, ont un revenu compris entre 25 000 € et 29 000 € .



Les aires vertes entre chaque décile sont égales.

1. Répartition des revenus disponibles des ménages en 2004



2. Revenu disponible en 2004 par décile

Tranche de revenu disponible	Limite supérieure de tranche (décile)	Précision du décile	Revenu moyen
Inférieur à D1	11 480	+/- 130	8 790
D1 à D2	14 410	+/- 120	13 010
D2 à D3	17 580	+/- 140	15 940
D3 à D4	20 940	+/- 150	19 250
D4 à D5	24 600	+/- 150	22 780
D5 à D6	28 620	+/- 170	26 570
D6 à D7	33 170	+/- 190	30 840
D7 à D8	39 360	+/- 250	36 080
D8 à D9	49 550	+/- 370	43 930
Supérieur à D9			72 170

Champ : ménages dont le revenu déclaré au fisc est positif ou nul et dont la personne de référence n'est pas étudiante.

Lecture : pour les 10 % des ménages les plus modestes, le revenu disponible en 2004 est inférieur à 11 480 euros (D1, 1^{er} décile), avec une précision de plus ou moins 130 euros. Il y a donc une probabilité de 95 % pour que la vraie valeur du 1^{er} décile soit comprise dans l'intervalle [11 480 - 130 ; 11 480 + 130]. Le revenu moyen des 10 % des ménages les plus modestes est de 8 790 euros.

III-Moyenne :

	2000	2001	2002	2002 rétropolée ¹	2003	2004
Revenu disponible médian (euros constants)	23 450	24 000	24 390	24 930	24 750	24 600
Revenu disponible moyen (euros constants)	27 600	28 200	28 480	29 070	29 010	28 340
Rapport interdécile (D9/D1)	4,5	4,4	4,4	4,4	4,4	4,3

Interprétation : la proportion des ménages ayant un revenu inférieur au salaire moyen est supérieure à la proportion des ménages ayant un salaire supérieur au revenu moyen . On en conclut que ceux qui ont un salaire supérieur au revenu moyen perçoivent un revenu beaucoup plus élevé que les autres.

La moyenne est sensible aux valeurs extrêmes alors que la médiane ne l'est pas .

Extrait du manuel de SES , terminale ES

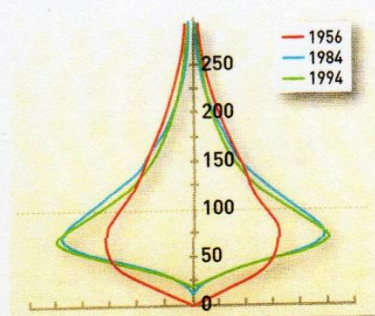
6 La réduction des inégalités de revenus

orsque l'on aborde la dynamique historique, trois résultats sont importants. D'abord, les Trente Glorieuses (1945-1975) apparaissent comme une période extraordinaire d'enrichissement du salariat : alors qu'en 1956, le pouvoir d'achat du salaire réel augmentait d'environ 0,5 % par an dans les années 1990, une année moyenne de la période antérieure à 1975 s'accompagnait d'une hausse de 3,5 %. Ensuite, l'écart entre cadres et ouvriers (nous prenons les employés, le résultat serait presque identique) a diminué après 1968 (le cadre gagnait en moyenne 4 fois plus que l'ouvrier en 1968, et 2,7 en 1984) ; un mouvement considérable de rapprochement

des salaires entre qualifiés et routiniers a donc eu lieu après 1968 ; ce mouvement résulte notamment d'un renouvellement fort de la population des cadres dans les années 1970 avec l'arrivée précoce et massive des jeunes générations diplômées du baby-boom. Enfin, depuis cette date, l'écart est à peu près stable (2,5 en 1998). La fin des Trente Glorieuses a donc été marquée par une dynamique doublement favorable à la classe ouvrière (et aux employés) : les écarts se réduisent dans un contexte d'enrichissement rapide et partagé.

Louis CHAUVEL,
« Le retour des classes sociales ? »
Revue de l'OFCE, n° 79, octobre 2001.

La moyennisation des revenus



Louis CHAUVEL, « La toupie et le sapin. Les inégalités dans les sociétés contemporaines », *Sciences humaines*, n° 72, 1997.

QUESTIONS

- 1 Recenser** : quels sont les éléments qui attestent d'une réduction des inégalités ?
- 2 Déduire** : comme le pouvoir d'achat a crû de 3,5 % par an en moyenne entre 1945 et 1975, par combien a-t-il été multiplié sur la période ?
- 3 Décrire** : d'après le strobiloïde, quels sont les signes d'une moyennisation de la société et d'une réduction des inégalités entre 1956 et 1984 ?

strobiloïde

est un graphique de répartition en forme de toupie (en grec, *strobilos* signifie toupie). Il se construit en mettant en ordonnées le revenu exprimé sous forme d'indice, le revenu médian représentant la base 100, et en abscisses la population concernée par ce revenu. La partie gauche et la partie droite du graphique sont symétriques de part et d'autre de la base 100 à former une toupie.

Entre 0 et 50 % du revenu médian se situe le groupe des plus pauvres. Entre 50 et 200 % du revenu médian se situent les catégories moyennes. Au-delà de 200 % se trouve le groupe des plus riches.

Pour la partie modélisation

Pour les animations sous geogebra: utiliser la version **Geogebra 4** (disponible gratuitement sur le site geogebra).

(remarque : la position du graphique en tant qu'image pose problème par rapport au repère ... c'est du bidouillage donc peut poser une difficulté de lecture)

- 1) **strobiloineceterepartition1.jpg.ggb** : faire bouger le curseur « a » et voir l'évolution des deux courbes sur les deux fenêtres (F et sa dérivée).
- 2) **strobiloetvertical1.ggb** : pour redresser le strobilo verticalement, puis faire bouger « a »
- 3) **strobiloetverticalcomparaison1.ggb** : le même mais comparaison dans une fenêtre avec les strobilo d'autres pays : bouger « a », d'où l'importance d'améliorer l'ajustement des courbes.
- 4) **strobiloineceterepartitionhachure1.jpg.ggb** : met en évidence la notion de fonction de répartition et fonction densité : Faire bouger le point A bleu pour mettre en correspondance l'aire sous le strobilo et la fonction de répartition (on observe l'égalité de l'aire et de la valeur de F dans la fenêtre algèbre) ... avec ou sans images de fond des graphique ?
- 5) **strobiloetverticalcomparaisonchauvel1.ggb** : amélioration avec les fonctions de modélisation de Chauvel, faire bouger les a , b et c . Les valeurs remarquables des différents pays sont données dans le tableau qui doit apparaître quand on demande de faire apparaître l'image 2.