

Objectif du TD : réinvestissement des connaissances de statistiques, médiane, diagrammes en boîte, quartiles, déciles, rapport interdécile.

Les données, tableaux et graphiques proviennent de l'étude de l'INSEE : « les revenus et patrimoines des ménages, éditions 2006 et 2012 ».

Exemple : le niveau de vie des individus en France en 2009.

I- caractéristiques de dispersion d'une série statistique :

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

Un graphique de base : le regroupement par classe

remarque : sur ce graphique de l'INSEE, les effectifs sont ramenés en pourcentage de l'effectif total (c'est-à-dire du nombre total des individus), en ordonnée sont donc représentées les fréquences de la série statistique. (Les données sont regroupées par ordre croissant du niveau de vie)

l'avantage de ce type de graphique est de voir la répartition des revenus par classe d'amplitude 100 € mais ne permet pas une interprétation en terme de dispersion

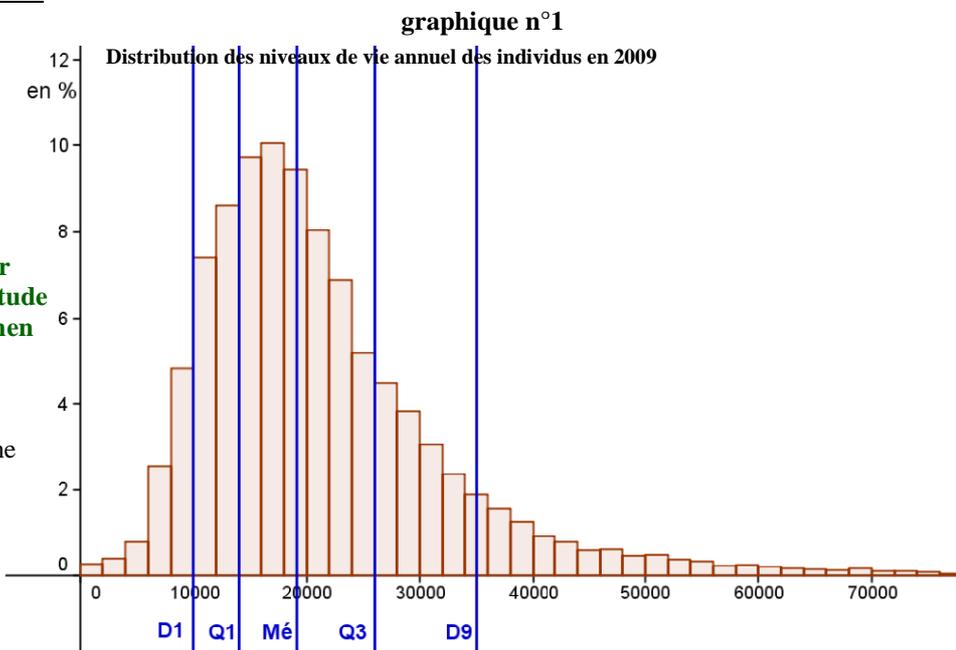
1) médianes, quartiles, déciles

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

Pour cela, on utilise les effectifs cumulés et la lecture de graphiques ou tableaux.

Compléter le tableau suivant en mettant en Evidence sur les deux graphiques, les valeurs trouvées.

Premier décile D1	Premier quartile Q1	médiane	Troisième quartile Q3	Dernier décile D9
10 000	14 000	19 000	26 000	35 000

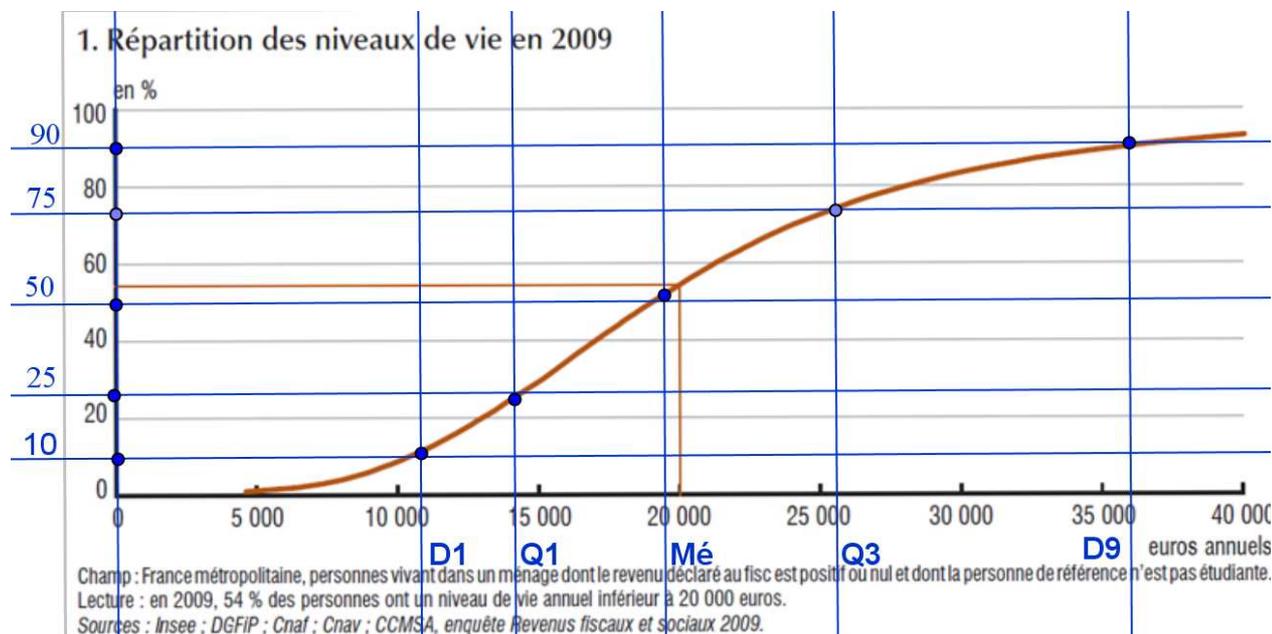


Source : d'après le graphique de la distribution des revenus et des niveaux de vie relatif à 2003 et le graphique des répartitions pour 2009, publiés par l'Insee.

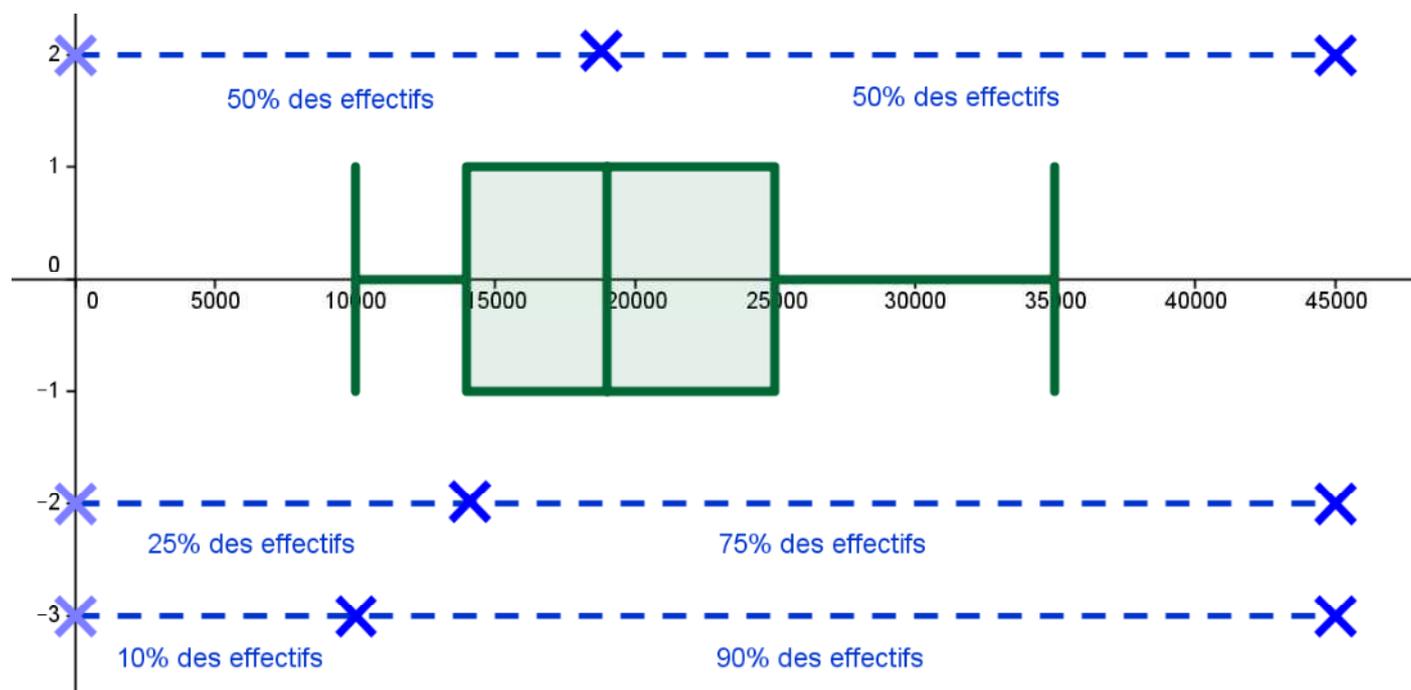
Lecture : le pas de l'histogramme est de 2000 € ; la hauteur de la barre de coordonnées n en abscisse est donc égale à la proportion de revenus ou niveaux de vie compris entre n et n+2000 ; ainsi 9,4% des individus ont un niveau de vie annuel en 2009 compris entre 18000€ et 20000€ et 5,4% des ménages ont un revenu disponible compris entre 18000€ et 20000€

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

graphique n°2



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



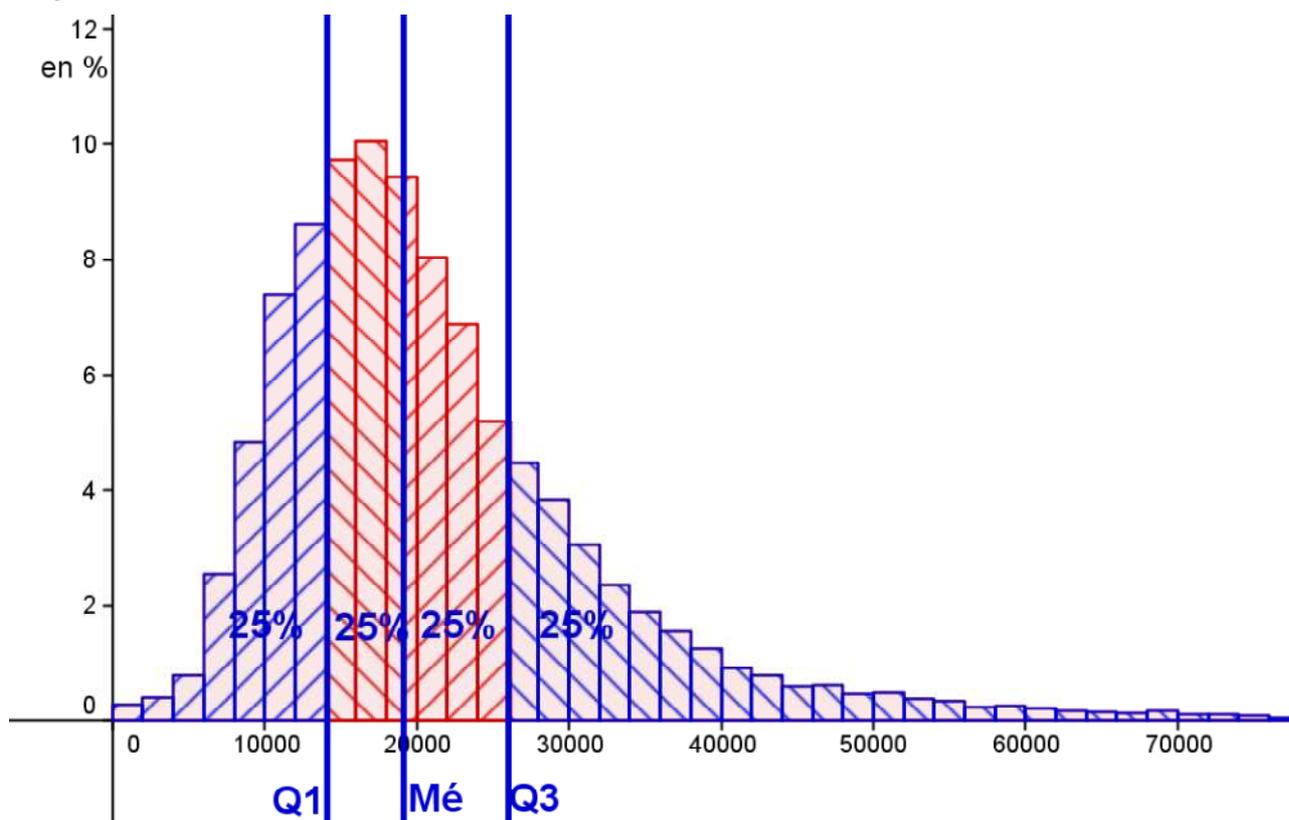
En math, on demandera souvent la valeur de l'écart interquartile : $Q3 - Q1 = 25\ 000 - 14\ 000 = 11\ 000\ €$

Interprétation : l'écart des niveaux de vie entre les 25 % des individus « les plus riches » et les 25% des individus « les plus pauvres » est supérieur à 11 000 €.

En SES, on demandera plutôt les rapports interdéciles : exemple $\frac{D9}{D1} = \frac{35\ 000}{10\ 000} \approx 3,5$

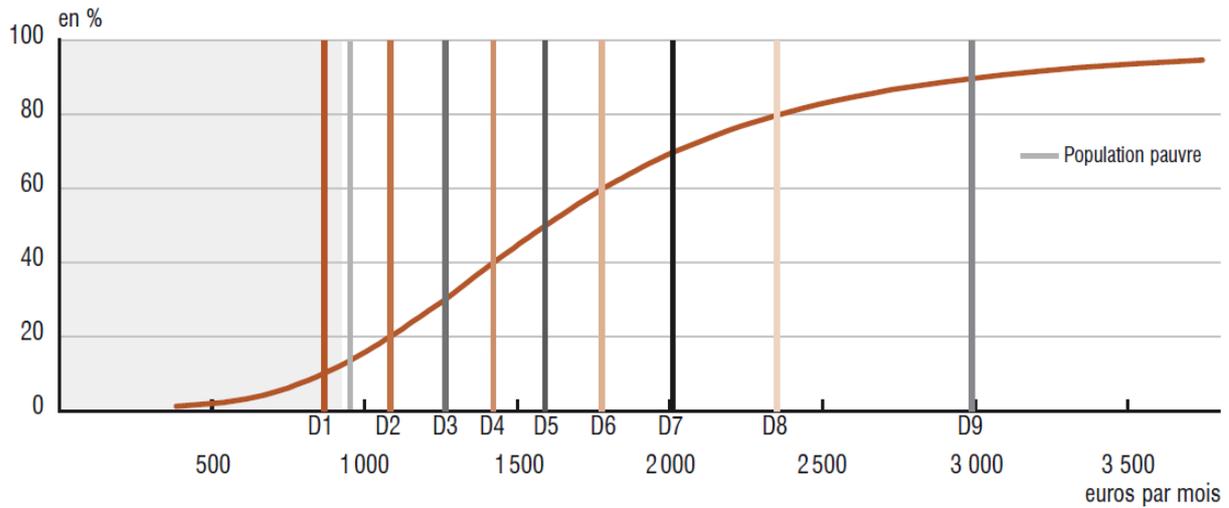
Interprétation : le niveau de vie des 10% des individus les plus riches est au moins 3,5 fois supérieur au revenu des 10 % des individus les plus pauvres.

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.



2) les déciles en SES

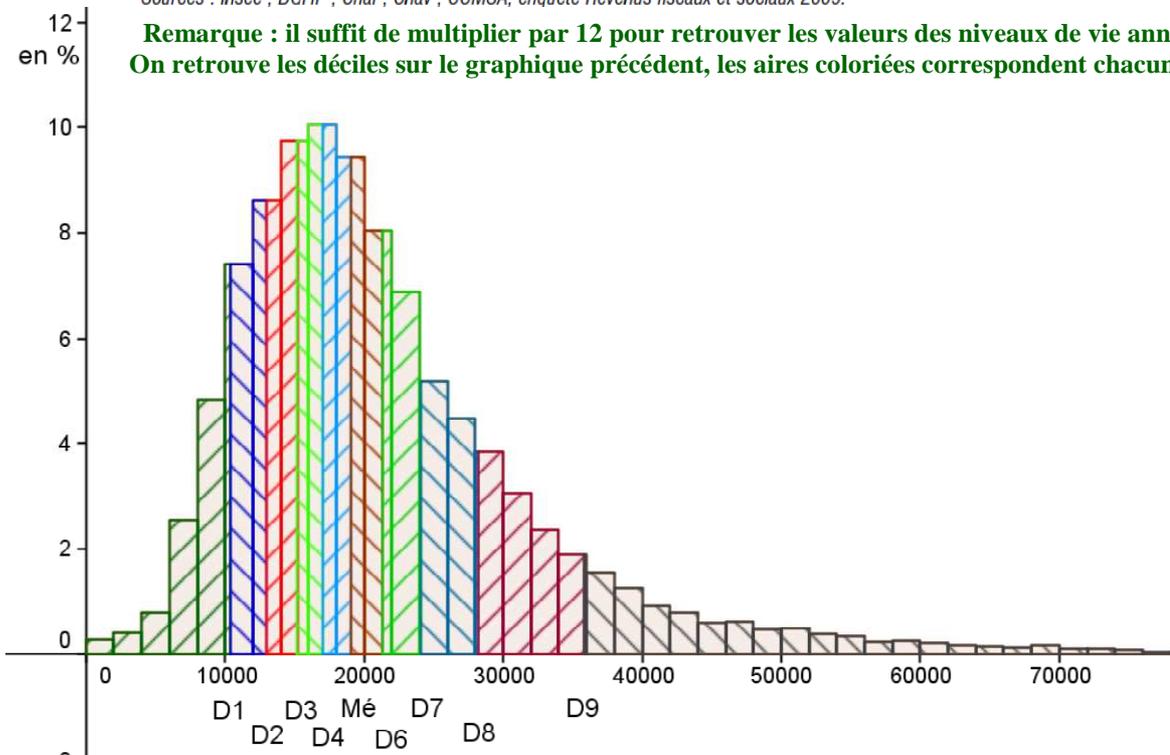
1. Répartition de la population selon le niveau de vie en 2009



Champ : France métropolitaine, personnes vivant dans un ménage dont le revenu déclaré au fisc est positif ou nul et dont la personne de référence n'est pas étudiante.
Lecture : D1 à D9 désignent les 9 déciles de niveaux de vie, seuils qui partagent la population en 10 sous-populations d'effectifs égaux. 70 % des personnes vivent avec moins de 2010 euros par mois.

Sources : Insee ; DGFIP ; Cnaf ; Cnav ; CCMSA, enquête Revenus fiscaux et sociaux 2009.

Remarque : il suffit de multiplier par 12 pour retrouver les valeurs des niveaux de vie annuels. On retrouve les déciles sur le graphique précédent, les aires coloriées correspondent chacune à 10 % des effectifs.



Compléter les phrases suivantes :

- D6 = 21 310 €, ce qui signifie que 60 % des individus ont un niveau de vie inférieur à 21 310 € en France , en 2009, par an.
- 10 % des individus ont un niveau de vie compris entre D5 et D6 .
- 60 % des individus ont un niveau de vie compris entre D2 et D8 .
- 30 % des individus ont un niveau de vie supérieur à 24 100 €
- 60 % des individus ont un niveau de vie supérieur à 17 050 €

Distribution des niveaux de vie

en euros
2009 par an

	2009
1er décile (D1)	10 410
2ème décile (D2)	13 010
3ème décile (D3)	15 160
4ème décile (D4)	17 050
Médiane (D5)	19 080
6ème décile (D6)	21 310
7ème décile (D7)	24 100
8ème décile (D8)	28 170
9ème décile (D9)	35 840
Rapport interdécile (D9/D1)	3,44

Lecture : en 2009, 10 % des individus ont un niveau de vie inférieur à 10 410 euros.

Attention : $D9 - D8 = 7\,670 \text{ €}$ et $D2 - D1 = 3\,400 \text{ €}$, ce qui signifie que l'écart de niveau de vie entre les 10% des individus situés dans la tranche du 9^{ème} décile est bien plus grand que l'écart des niveaux de vie entre les 10% des individus dans la tranche du deuxième décile.

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 :] 19 080 ; 21 310]. Les 10 % des individus ayant un niveau de vie supérieur à celui des 50 % les moins riches et inférieur à celui des 40 % les plus riches, ont un niveau de vie compris entre 19 080 € et 21 310 € .

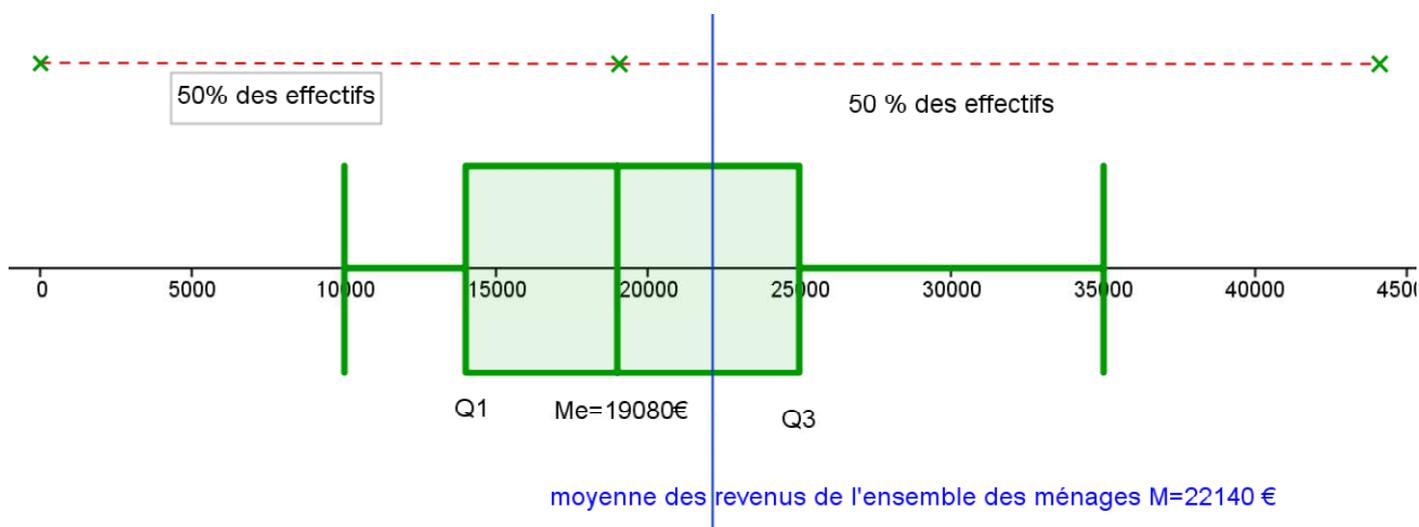
Remarque n° 2 : Ces caractéristiques de dispersion permettent d'essayer de définir le « seuil de pauvreté » : selon l'insee 60% du niveau de vie médian.

3) Rappel : ne pas confondre moyenne et médiane !

2. Quelques indicateurs sur le niveau de vie de 2001 à 2009

	montants annuels en euros constants 2009									
	2001	2002	2003	2004	2005	2006	2007	2008	2009	
Niveau de vie médian	17 410	17 880	17 830	17 740	18 030	18 300	18 690	19 000	19 080	
Niveau de vie moyen	20 230	20 730	20 610	20 590	20 820	21 340	21 690	22 130	22 140	

Que signifie le constat d'un revenu moyen supérieur au revenu médian ?



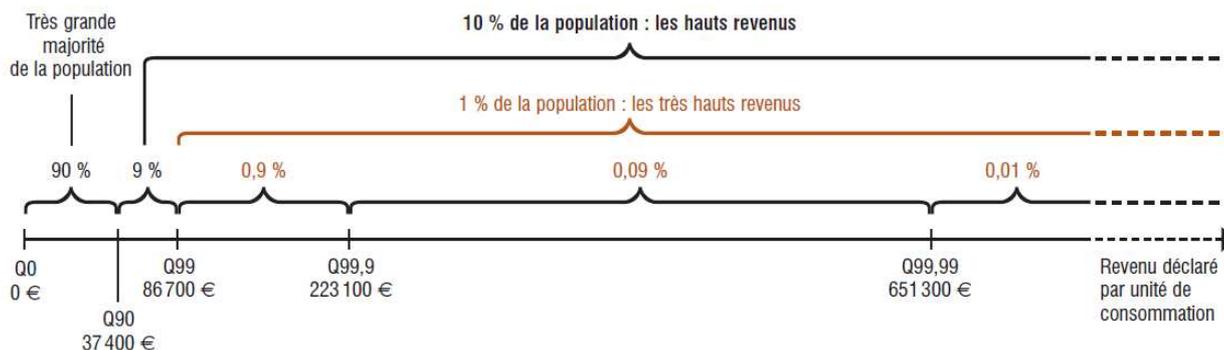
la proportion des individus ayant un niveau de vie inférieur au niveau de vie moyen est supérieure à la proportion des individus ayant un niveau de vie supérieur au niveau de vie moyen . Il y a bien plus de la moitié des individus qui ont un niveau de vie inférieur au niveau de vie moyen.

Comment expliquer cet écart ? (page 310 du manuel de SES)

Les individus qui ont un niveau de vie supérieur au niveau de vie moyen ont des niveaux de vie beaucoup plus élevés que les autres.

La moyenne est sensible aux valeurs extrêmes alors que la médiane ne l'est pas , d'où le choix d'indicateurs calculés à l'aide des déciles et de la médiane plutôt que celui de la moyenne.

1. Échelle des revenus déclarés par unité de consommation en 2009



Champ : France métropolitaine, personnes appartenant à des ménages fiscaux dont les revenus déclarés sont strictement positifs.
Sources : Insee-DGFIP; Revenus fiscaux localisés (RFL) 2009, calculs Insee.

Remarque : la **moyenne** est souvent associée à l'**écart type** de la série.

Pour l'écart type de la série du niveau de vie des individus : 17000€ environ, pour l'écart type de la série des revenus des ménages : 31000€ environ . Comment peut-on interpréter cette différence sur les deux séries ?

La série des niveaux de vie a des valeurs bien plus resserrées autour de la moyenne du niveau de vie que la série des revenus disponibles autour de la moyenne des revenus disponibles.