

Exercice 3 : Diversité des exploitations agricoles en France métropolitaine en 2007

CORRIGÉ

On assimilera dans cet exercice la taille des exploitations agricoles à la surface réellement utilisée pour la production agricole (surface agricole utilisée, SAU). La précision des résultats des calculs sera la même que celle des données du tableau 1.

1) Question préliminaire

Indiquez le sens statistique des données du tableau 1 :

- Quelle est la population étudiée ?
- Quelle est la variable étudiée ? De quel type de variable s'agit-il ?
- Quelle est la masse statistique ? A partir de quelle(s) données pourrait-elle être calculée ? Avec quelle(s) hypothèse(s) ?

Corrigé

- La population statistique est ici l'ensemble des exploitations agricoles de France métropolitaine en 2007.
- La variable de classement de cette population est ici la taille (ou leur superficie). Il s'agit d'une variable quantitative continue. Les superficies ont été agrégées en groupes d'amplitudes inégales.
- La masse statistique est ici la superficie totale correspondant aux différents types d'exploitation. Par exemple, la surface agricole totale utilisée par l'ensemble des exploitations de moins de 10 hectares est de 142 000 ha. Cette dernière peut être obtenue sans hypothèse par le produit de la taille moyenne des exploitations de moins de 10 ha par le nombre d'exploitations de moins de 10 ha. On peut également approcher ce résultat par le produit du centre de la tranche de surfaces (par exemple 5 ha pour les exploitations de moins de 10 ha) par le nombre d'exploitations agricoles appartenant à cette tranche, en faisant l'hypothèse d'une répartition linéaire des exploitations sur l'intervalle de superficies considéré.

2) Calculez la moyenne, la variance, l'écart-type et le coefficient de variation. Exprimez le plus clairement possible (et de manière concise) la moyenne et l'écart-type.

Corrigé

La moyenne correspond ici à la taille moyenne des exploitations. Il s'agit du rapport entre, d'une part, la surface totale de l'ensemble des exploitations et, d'autre part, le nombre total d'exploitations agricoles. Comme on assimile dans cet exercice la superficie de chaque exploitation à la surface agricole utilisée (ce qui a d'ailleurs plus de sens), on connaît ces deux grandeurs : la surface agricole

utilisée totale est de 25 210 300 ha et le nombre total d'exploitations est de 326 300. La taille moyenne des exploitations est donc égale à :

$$\bar{t} = \frac{25\,210,3}{326,3} = 77,3 \text{ ha}$$

La variance correspond à la moyenne des carrés des écarts à la moyenne. Il faut donc ici :

- a) déterminer pour chaque classe d'exploitations l'écart entre la taille moyenne de ces exploitations (t_i) et la taille moyenne de l'ensemble des exploitations (77,3 ha) ;
- b) élever au carré cet écart ;
- c) pondérer le carré de l'écart à la moyenne par le poids démographique (n_i / N) de la classe d'exploitations considérée ;
- d) on somme ensuite ce type de résultat obtenu pour les différentes classes.

On peut ici déterminer précisément la taille moyenne de chaque groupe d'exploitations. En effet, pour chacun d'eux, on connaît leur masse statistique ($SAU_i = m_i$), c'est-à-dire la somme de leurs superficies. Par exemple, on peut lire dans le tableau 1 que les 34 900 exploitations agricoles de moins de 10 ha correspondent à une superficie agricole totale de 142 000 ha, soit en moyenne 4,1 ha par exploitation agricole de moins de 10 ha.

$$t_i = \frac{m_i}{n_i} = \frac{SAU_i}{n_i}$$

$$Var(t) = \sum_{i=1}^6 \frac{n_i}{N} \times (t_i - \bar{t})^2$$

Calcul des t_i :

i	Exploitations professionnelles	SAU _i (m _i)	Nombre (n _i)	Taille moyenne (t _i)
	Total	25 210,3	326,3	77,3
1	Moins de 10 ha (1)	142,0	34,9	4,1
2	10 à 25 ha	664,2	38,2	17,4
3	25 à 50 ha	2 419,8	64,8	37,3
4	50 à 100 ha	7 243,4	100,2	72,3
5	100 à 200 ha	9 539,2	69,5	137,3
6	200 ha ou plus	5 201,7	18,7	278,2

Calcul de la variance, de l'écart-type et du coefficient de variation :

$$Var(t) = \left[\frac{34,9}{326,3} \times (4,1 - 77,3)^2 \right] + \dots + \left[\frac{18,7}{326,3} \times (278,2 - 77,3)^2 \right] = 4\,387 \text{ ha}^2$$

$$\sigma_t = \sqrt{Var(t)} = \sqrt{4\,387} = 66,2 \text{ ha}$$

$$CV = \frac{\sigma_t}{\bar{t}} = \frac{66,2}{77,4} = 0,86 = 86 \%$$

La taille moyenne des exploitations agricoles est de 77,4 ha. Mais il y a une grande hétérogénéité des superficies puisqu'une grande partie des exploitations sont comprises entre 11,2 et 143,6 ha (77,4 ha \pm 66,2 ha), soit une plage de variation dont les bornes se situent dans un rapport de 1 à 13 ! On a donc une grande diversité des exploitations, ce qui se concrétise par un coefficient de variation proche de 1 (86 %).

3) Calculez la médiane (Me) et la médiale (MI). Comparez là encore le plus clairement possible (et toujours de manière concise) ces deux indicateurs.

Tracez sur un carré de Gini la courbe de Lorenz et calculez l'indice de Gini.

Corrigé

a) Rappels : la médiane est la valeur de la variable (ici la superficie) qui partage l'effectif total de la population étudiée (ici les exploitations agricoles) en deux groupes de même effectif. La médiale est la valeur de la variable (toujours la superficie) qui partage la masse statistique totale (ici la superficie totale) en deux masses égales.

Il faut donc au préalable calculer :

- les effectifs (d'exploitations) cumulés (nc_i) puis les fréquences relatives cumulées (fc_i) (ou directement ces dernières) ;
- le cumul des masses statistiques (mc_i) puis le cumul des masses statistiques relatives (mcr_i).

Le tableau 1 présente déjà la répartition en % des exploitations agricoles selon la superficie. Il faut donc juste cumuler ces pourcentages pour obtenir la série des fréquences relatives cumulées.

Le tableau 1 indique aussi déjà les masses statistiques (SAU), ainsi que la part de chacune de ces masses statistiques dans la surface agricole utilisée totale de France métropolitaine. Il faut donc juste cumuler ces fréquences pour déterminer le cumul des masses statistiques relatives.

Tableau 2 : Calcul du cumul des fréquences relatives (fc_i) et du cumul des masses statistiques relatives (mcr_i)

Exploitations professionnelles	Part (en %) f_i	Cumul des parts fc_i	SAU (en %) mr_i	Cumul des SAU (en %) mcr_i
Moins de 10 ha	10,7	10,7	0,6	0,6
10 à 25 ha	11,7 ⁽¹⁾	22,4 ⁽²⁾	2,6 ⁽³⁾	3,2 ⁽⁴⁾
25 à 50 ha	19,9	42,3	9,6	12,8
50 à 100 ha	30,7	73	28,7	41,5
100 à 200 ha	21,3	94,3	37,9	79,4
200 ha ou plus	5,7	100,0	20,6	100,0
Total	100,0		100,0	

Lecture du tableau :

(1) 11,7 % des exploitations ont une superficie comprise entre 10 et 25 ha.

(2) 22,4 % des exploitations ont une superficie inférieure à 25 ha.

(3) Les exploitations agricoles dont la taille est comprise entre 10 et 25 ha concentrent 2,6 % de la surface agricole totale utilisée de France métropolitaine.

(4) Les exploitations agricoles dont la taille est inférieure à 25 ha concentrent 3,2 % de la surface agricole totale utilisée de France métropolitaine.

Calcul de la médiane :

$$Me = 50 + \left[\left[\frac{(50 \% - 42,3 \%)}{(73,0 \% - 42,3 \%)} \right] \times (100 - 50) \right] = 62,5 \text{ ha}$$

Calcul de la médiale :

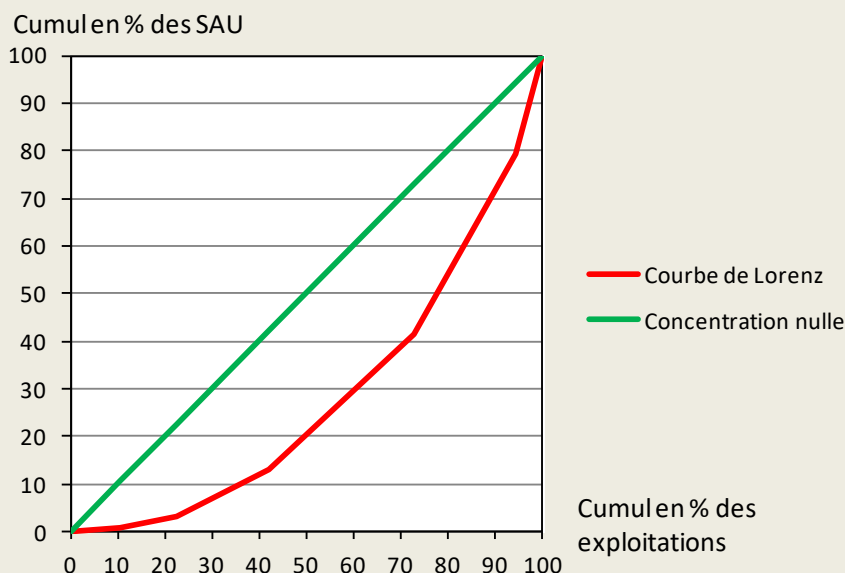
$$Ml = 100 + \left[\left[\frac{(50,0 \% - 41,5 \%)}{(79,4 \% - 41,5 \%)} \right] \times (200 - 100) \right] = 122,4 \text{ ha}$$

Une exploitation agricole sur deux a une superficie inférieure à 62,5 ha. Mais il faut agréger toutes les exploitations agricoles de moins de 122,4 ha pour atteindre la moitié de la surface agricole utilisée totale.

Si l'on compare les fréquences cumulées relatives et les masses statistiques relatives cumulées, on peut dire que : près des trois quarts (73 %) des exploitations agricoles comptent moins de 100 ha. Elles correspondent seulement à 41,5 % de la surface agricole totale utilisée en France métropolitaine. Une petite partie (27 %) des exploitations agricoles (les plus grosses, celles comptant au moins 100 ha) concentre donc une large majorité (58,5 %) des terres agricoles utilisées. Mieux : les exploitations de plus de 200 ha, qui représentent 5,7 % des exploitations agricoles, concentrent plus de 20 % des terres utilisées pour l'agriculture. Il y a donc une concentration assez forte. L'indice de Gini devrait le confirmer.

b) La courbe de Lorenz (figure 1) est tracée à partir de la série des fréquences relatives cumulées (en abscisses) et des masses statistiques relatives cumulées (en ordonnées).

Figure 1 : La concentration d'une grande partie des terres agricoles utilisées dans un nombre plus réduit d'exploitations agricoles

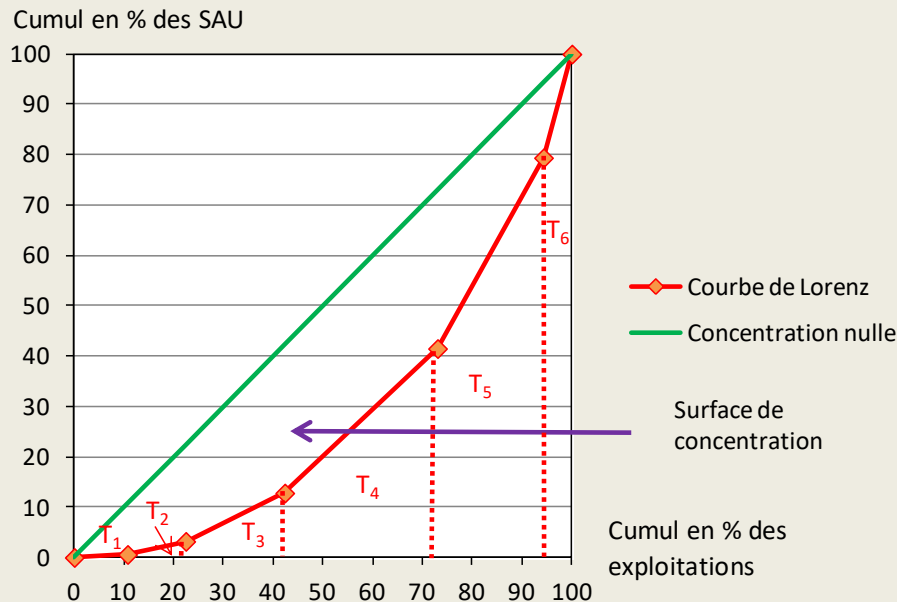


L'indice de Gini est le rapport entre la surface de concentration observée et la surface de concentration maximale (figure 2). La surface de concentration maximale est, dans un carré de Gini de surface 100×100 , de 5 000. La surface de concentration observée est obtenue en effectuant la

différence entre la surface de concentration maximale et la surface située sous la courbe de Lorenz. Cette dernière peut être décomposée en une série de trapèzes de surface S_{T_i} :

$$S_{T_i} = b \times \frac{(H + h)}{2}$$

Figure 2 : Calcul de l'indice de Gini à partir de la somme des surfaces des trapèzes situés sous la courbe de Lorenz



Calcul de la surface de concentration :

$$S_{concentration} = 5\,000 - \sum_{i=1}^6 S_{T_i}$$

$$S_{T_1} = (10,7 - 0,0) \times \frac{(0,6 + 0,0)}{2} = 3,2$$

$$S_{T_2} = (22,4 - 10,7) \times \frac{(3,2 + 0,6)}{2} = 22,2$$

$$S_{T_3} = (42,3 - 22,4) \times \frac{(12,8 + 3,2)}{2} = 159,2$$

$$S_{T_4} = (73,0 - 42,3) \times \frac{(41,5 + 12,8)}{2} = 833,5$$

$$S_{T_5} = (94,3 - 73,0) \times \frac{(79,4 + 41,5)}{2} = 1\,287,6$$

$$S_{T_6} = (100,0 - 94,3) \times \frac{(100,0 + 79,4)}{2} = 511,3$$

$$\sum_{i=1}^6 S_{T_i} = 3,2 + 22,2 + 159,2 + 833,5 + 1\,287,6 + 511,3 = 2\,817,0$$

$$S_{Concentration} = 5\,000 - 2\,817 = 2\,183$$

$$I_{Gini} = \frac{2\,183}{5\,000} = 0,44 = 44\%$$

4) Commentez les données du tableau 1 à l'aide des caractéristiques de dispersion et de concentration de votre choix.

Proposition de corrigé

En 2007, la surface agricole totale utilisée en France métropolitaine couvre 25,2 millions d'hectares ; elle est répartie dans 326,3 mille exploitations agricoles.

Les superficies des exploitations agricoles françaises sont de tailles très inégales. Certes, 3 exploitations sur 10 comptent de 50 à 100 ha. Mais on compte quasiment autant de petites exploitations (22,4 % ont moins de 25 ha) que de très grandes (27,0 % comptent au moins 100 ha). De ce fait, la moyenne reflète très imparfaitement le profil des exploitations françaises : la taille moyenne est de 77,3 ha ... mais $\pm 66,2$ ha (qui correspond à l'écart moyen autour de la taille moyenne) !

Ces exploitations pèsent d'un poids inégal sur le potentiel agricole français : en effet, les exploitations comptant au moins 100 ha, qui représentent à peine plus du quart (27 %) des exploitations métropolitaines, regroupent près de 60 % (plus précisément 58,5 %) de l'ensemble des terres agricoles utilisées en France. À elles seules, les exploitations comptant au moins 200 ha (5,7 % des exploitations agricoles françaises) concentrent plus d'1/5^e des terres agricoles utilisées en métropole.

La mise en relation de deux chiffres résume bien cette assez forte concentration des terres cultivées par un nombre réduit d'exploitations :

- une exploitation agricole sur deux a une superficie inférieure à 62,5 ha (taille médiane) ;
- mais il faut agréger toutes les exploitations agricoles de moins de 122,4 ha (taille médiale) pour atteindre la moitié de la surface agricole utilisée totale.

Annexes : Données

Tableau 1 : Les exploitations agricoles selon la taille (en milliers) et surface agricole utilisée (SAU, en milliers d'hectares)

Exploitations professionnelles	Nombre (en milliers)	Part (%)	SAU (en milliers)	Part (%)
Total	326,3	100,0	25 210,3	100,0
Moins de 10 ha (1)	34,9	10,7	142,0	0,6
10 à 25 ha	38,2	11,7	664,2	2,6
25 à 50 ha	64,8	19,9	2 419,8	9,6
50 à 100 ha	100,2	30,7	7 243,4	28,7
100 à 200 ha	69,5	21,3	9 539,2	37,9
200 ha ou plus	18,7	5,7	5 201,7	20,6

(1) : y compris les exploitations sans superficie agricole utilisée.

Champ : France métropolitaine.

Source : service de la statistique et de la prospective (SSP), enquête Structure.