Université A- mira de Bejaia
Faculté des sciences économiques, commerciales et des sciences de gestion
Département SEGC(LMD)
Module stat I
Enseignante : Dr. BERRAH

Chapitre 6 : Les mesures de concentration

La concentration consiste à mesurer les inégalités dans la répartition des valeurs d'une variable entre les individus de la population étudiée. Autrement dit, On cherche à savoir, si la variable statistique étudiée, est concentrée sur quelques individus ou au contraire elle est bien répartie entre les individus.

Par exemple, la concentration des revenus mesure la manière dont les revenus totaux sont répartis entre les individus d'une population. Une forte concentration des revenus signifie qu'une part importante des revenus totaux est détenue par un petit nombre d'individus, tandis qu'une faible concentration indique une répartition plus équitable.

Cette notion s'applique aux variables quantitatives prenant des valeurs positives et ayant une signification concrète lorsqu'elles sont agrégées. On l'utilise par exemple pour analyser la concentration des salaires, des revenus, des richesses, des parts de marché, etc.

1-Définitions de base : Afin de mesurer la concentration, il convient de définir les notions suivantes :

1-1-Valeurs globales (la masse totale : $\acute{N}$**):** xi représentent les valeurs du caractère étudié (ou les centres des classes), ni les effectifs correspondants. Les valeurs globales de la série est la somme de toutes les unités de la variable prises par l'ensemble des individus de la population, autrement dit c'est la somme des produits : $\acute{N} = \sum$ni.xi.

1-2-Médiale (Notée **Ml**) : La Médiale est la valeur du caractère qui partage en deux parties égales la masse totale ($\acute{N} = \sum$ni.xi), elle s'exprime dans la même unité que la variable étudiée. La médiale est une médiane que l'on calcule non plus sur les effectifs (ni) de la série mais sur la somme du produit ni.xi (ou ci étant le centre de classe). Elle est déterminée selon les étapes suivantes :

- Calculer les valeurs de la masse: $n_i\grave{} = n_i \times x_i$.
- Calculer les effectifs des masses cumulés croissant $\grave{N}^{\uparrow}$
- Déterminer la classe médiale [x_i ; x_{i+1} [: c'est la classe qui correspond à la valeur : $\frac{\acute{N}}{2}$ sur la colonne croissante $\grave{N}^{\uparrow}$. Sachant : $\acute{N} = \sum_{i=1}^{K} \acute{n}_i$
- Appliquer la formule suivante :

$$ML = x_i + \frac{\frac{\sum nixi}{2} - \text{N}^{\uparrow\text{précedent}}_{(\text{nixi})}}{\text{nixi (de classe médiale)}}(a_i) = x_i + \frac{\frac{\acute{N}}{2} - \acute{\text{N}}^{\uparrow}_{\text{précedent}}}{\acute{\text{n}}_{(\text{de la classe médiale})}}(a_i)$$

Sachant :

x_i : Borne inférieure de la classe médiale.

a_i : Amplitude de la classe médiale.

$\acute{\text{N}}^{\uparrow}_{\text{précedent}}$: Effectif des masses cumulé croissant de la classe située avant la classe médiale.

$\acute{\text{n}}_{\text{de la classe médiale}}$: Effectif de la masse (ni*xi) de la classe médiale.

Remarque 1 : on peut calculer la médiale avec les fréquences relatives de la masse (f_i'), sachant : $f_i' = \frac{n_i \times x_i}{\sum n_i \times x_i} = \frac{n_i'}{\sum_{i=1}^{k} n_i'} = \frac{n_i'}{N'}$

Alors la médiale est égale à :

$$ML = x_i + \frac{0.5 - \acute{\text{F}}^{\uparrow}_{\text{précedent}}}{\acute{\text{f}}_{ML}}(a_i)$$

Remarque 2 : La médiale est toujours supérieure ou égale à la médiane.

2- comment mesurer la concentration ?

On mesure la concentration par deux méthodes : la méthode algébrique et la méthode graphique.

2-1- La méthode algébrique : La démarche est la suivante :

1) Calculer la médiane (Mé) de la série

2) Calculer la médiale (Ml) de la série.

3) Calculer l'écart (Δ M) entre la médiale et la médiane : Δ M = ML - Mé

4) Calculer l'étendue de la série (E) : $E = x_{max} - x_{min}$

5) Calculer le rapport : $S = \frac{\Delta M}{E}$

- si $S \geq 0.5$ La concentration est forte

- si $S \prec 0.5$ La concentration est faible

- si $S = 0$ La concentration est nulle (situation égalité parfaite)

Exemple : une entreprise présente la répartition des revenus de manière suivante

Salaire (xi)	[500-1500 [	[1500-2500 [	[2500-5500[	total
ni	50	125	25	200

1-Calculer la médiane.

2- Calculer la médiale.

3- Calculer l'écart Δ M = médiale – médiane.

4- Comment elle est la mesure de concentration des salaires ?

Solution :

xi	**ni**	f_i	$F\uparrow$	ci	n_i' (Masses salariales =($n_i \times c_i$)	f_i' $(=\frac{n_i \times c_i}{\sum n_i \times c_i})$	$F'\uparrow$
[500-1500 [	50	0.25	0.25	1000	50000	0.125	0.125
[1500-2500 [	125	0.625	0.875	2000	250000	0.625	0.75
[2500-5500[	25	0.125	1	4000	100000	0.25	1
total	200	1	-	-	400000	1	-

1-La médiane : Classe médiane : on compare 0.5 par apport aux valeurs de la colonne $F\uparrow$, donc la classe médiane est [1500 ; 2500 [

$$\text{Donc } M_{é} = x_i + \frac{0.5 - \mathrm{F}^{+}_{\text{Précedent}}}{\mathrm{f}_{\text{i mediane}}}(a_i) = 1500 + \frac{0.5 - 0.25}{0.625}(1000) = 1900$$

2-La médiale : Classe médiale : on compare 0.5 par apport aux valeurs de la colonne $F'\uparrow$

donc la classe médiale est [1500 ; 2500 [

Donc : $ML = x_i + \frac{0.5 - \mathrm{F}^{\uparrow}_{\text{précedent}}}{\mathrm{f}_{ML}}(a_i)$

$$ML = 1500 + \frac{0.5 - 0.125}{0.625}1000 = 2100$$

3-l'écart Δ M = médiale – médiane = 2100-1900=200

4-Mesure de la concentration :

l'étendue = X_{max} - X_{min} = 5500-500 = 5000

On calcule le rapport : $S = \frac{\Delta M}{E} = \frac{200}{5000}$ = 0.04

La concentration des salaires est : $0.04 \prec 0.5$ et donc la concentration est faible.

2-2- La méthode Graphique : La démarche est la suivante :

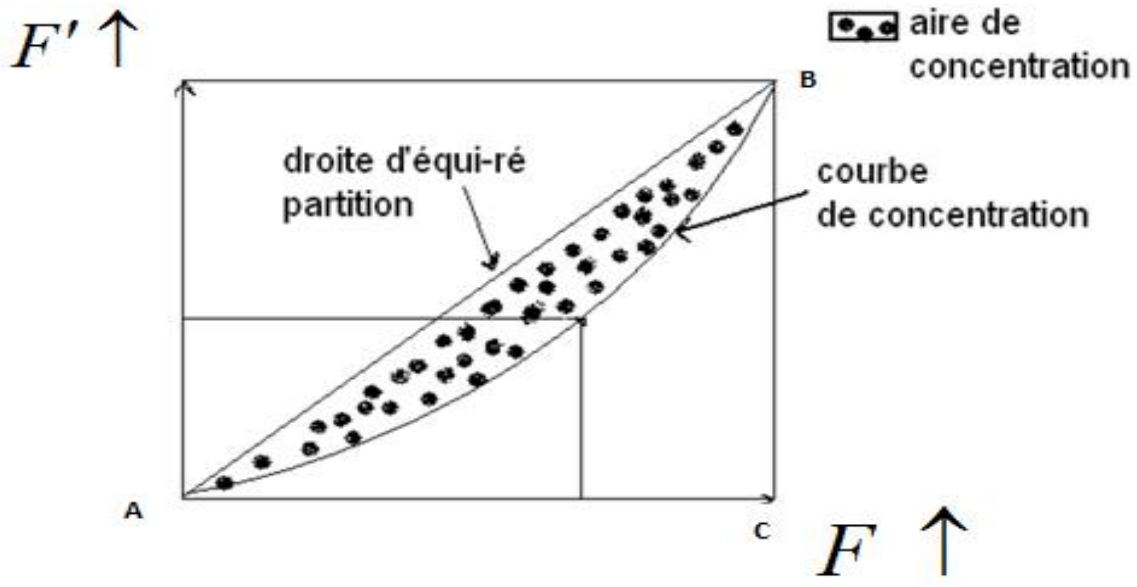
1) Tracer la courbe de concentration (s'appelle aussi courbe de LORENZ), Elle se construit sur un repère orthonormé à partir des fréquences cumulées relatives (voir le schéma ci-après) :

- Sur l'axe des abscisses : il porte les fréquences relatives cumulées croissante ($F_i \uparrow$)
- Sur l'axe des ordonnées : il porte les fréquences relatives des masses cumulées croissantes ($F(ni * xi) \uparrow$).

2-2- La méthode Graphique : La démarche est la suivante :

1) Tracer la courbe de concentration (s'appelle aussi courbe de LORENZ), Elle se construit sur un repère orthonormé à partir des fréquences cumulées relatives (voir le schéma ci-après) :

- Sur l'axe des abscisses : il porte les fréquences relatives cumulées croissante ($F \uparrow$).
- Sur l'axe des ordonnées : il porte les fréquences relatives des masses cumulées croissantes ($F' \uparrow$).



2) Calculer l'indice de Gini (noté IG) : Il est égal au rapport de la surface de concentration sur la surface du triangle ABC.

- Au numérateur, on porte la surface de concentration, comprise entre la bissectrice et la courbe de concentration.
- Au dénominateur on porte la surface du triangle ABC.

$$IG = \frac{surface\ de\ concentration}{surface\ du\ triangle\ ABC} = \frac{surface\ de\ concentration}{(1\times 1)/2} = 2 \times surface\ \ de\ \ concentration$$

L'indice de Gini permet de quantifier l'éloignement de la courbe de concentration à la bissectrice (droite d'égalité parfaite ou droite d'équi-répartition), sa valeur comprise entre 0 et 1 :

- Si IG est proche de 0, la courbe de Lorenz est proche de la diagonale, la concentration est faible.
- Si IG est proche de 1, la concentration est forte.
- Si IG est égal à 0 , la concentration est nulle, elle corresponde à une égalité parfaite.

Si l'on tient à calculer la surface de concentration, on peut se servir de la méthode des trapèzes[1]. Et donc :

$$la\ surface\ de\ concentration = la\ surface\ du\ triangle\ ABC - \sum les\ surfaces\ des\ trapèzes$$

$$\Rightarrow la\ surface\ de\ concentration = 1/2 - \sum surfaces\ des\ trapèzes$$

Rappelant $la\ surface\ de\ trapèz = \dfrac{(la\ base\ \inf érieure + la\ base \sup érieure) \times la\ hauteure}{2}$

$$\Rightarrow \quad la\ surface\ de\ trapèz = \frac{(b_{i-1} + bi) \times f_i}{2}$$

Sachant : b_{i-1} et b_i : indiquent les fréquences relatives des masses cumulées croissantes ($F^{\uparrow}_{(ni\ xi)}$)

fi : fréquence relative.

$$\Rightarrow et\ donc\ la\ \ surface\ de\ concentration = 1/2 - \sum surfaces\ des\ trapèzes = 1/2 - \sum \frac{(b_{i-1} + bi) \times f_i}{2}$$

Comme nous avons :

$$IG = 2 \times surface\ de\ concenation$$

On remplace la valeur du surface de concentration et donc :

$$IG = 2 \times \left(\frac{1}{2} - \sum \frac{(b_{i-1} + bi) \times f_i}{2} \right) = 2 \times \left(\frac{1}{2} - \frac{1}{2} \sum (b_{i-1} + bi) \times f_i \right) = \quad 1 - \sum (b_{i-1} + bi) \times f_i$$

Exemple : le même exemple précédent. Calculer l'indice de Gini et interpréter le résultat.

Solution : on a : $IG = 1 - \sum (b_{i-1} + bi) \times f_i$

Sachant : b_{i-1} et b_i : indiquent les fréquences relatives des masses cumulées croissantes ($F' \uparrow$)

[1] **Rappel** : surface du triangle = la base*la hauteur /2
Surface de trapèze = (la base inférieure + la base supérieure)* la hauteur/2

fi : fréquence relative.

Pour calculer l'indice de Gini, il suffit de faire les calculs suivants :

B_{i-1}	B_i	$B_{i-1} + Bi$	fi	$(B_{i-1} + Bi)*fi$
0	0.125	0.125	0.25	0.03125
0.125	0.75	0.875	0.625	0.546875
0.75	1	1.75	0.125	0.21875
-	-	-	1	0.8

$$IG = 1 - \sum (b_{i-1} + bi) \times f_i = \quad 1 - 0.8 = 0.2$$

Ici, IG = 0.2, il est faible, et donc la concentration des salaires est faible : les salaires sont assez bien répartis sur l'ensemble des salariés.

Série des exercices

Exercice 1 : Le tableau suivant présente la répartition des salaires dans une entreprise :

Classe de salaire (euros)	**Effectif**
[1000 ; 1500[	20
[1500 ; 2000[	30
[2000 ; 2500[	40
[2500 ; 3000[	10

1- Tracez la courbe de Lorenz.

2- Calculez l'indice de Gini. Interprétez les résultats.

Exercice 2 : Le tableau suivant présente la répartition des patrimoines (en milliers d'euros) dans une région :

Classe de patrimoine	Effectif
[0 ; 50[	100
[50 ; 100[	150
[100 ; 200[	200
[200 ; 500[	50

- Calculez l'indice de Gini. Interpréter le résultat.