Chapitre 4 : Les paramètres de concentration

Objectif général du chapitre :

Comprendre les paramètres de concentration.

Introduction :

Définition :

Un indicateur de concentration est un nombre qui mesure la relation entre fraction d'effectifs et leur part de masse totale de la variable ($\sum n_i c_i$). La concentration peut être représentée graphiquement par la courbe de Lorenz et mesurée soit à l'aide de coefficient de Gini, soit à l'aide de la médiale.

Seules les distributions positives et continues peuvent être soumises à une étude de concentration.

I. La courbe de concentration de Lorenz

Objectifs spécifiques :

- Saisir la manière de déterminer la courbe de concentration de Lorenz.

Durée : 0,45 H.

Contenu :

1. Définition :

Soit une variable statistique X positive et continue telle que sa masse totale ($\sum n_j c_j$) est partagée par les individus de la population. Désignant par :

$q_i = \sum_{j=1}^{i} n_j c_j / \sum_{j=1}^{k} n_j c_j \iff q_i = \sum_{j=1}^{i} f_i c_i / \sum_{j=1}^{k} f_j c_j$: la part de cet individu dans la masse totale.

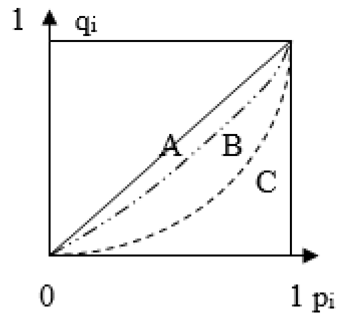
Et $p_i = F(e_i)$: proportion des individus.

La courbe de concentration est la représentation graphique des points (p_i , q_i).

2. Particularité de la courbe de concentration :

- Elle est représentée dans un carré de côté 1 ou 100 %.
- Elle passe par les points (0,0) et (1,1).

- Elle se trouve toujours en dessous de sa première bissectrice (qi = pi qui traduit une concentration nulle indiquant une répartition parfaitement égalitaire de la masse totale de la variable).
- Elle est croissante et elle a une forme exponentielle.



Plus la courbe de concentration est éloignée de la première bissectrice, plus la concentration est forte donc, l'inégalité est forte.

A : faible concentration : répartition faiblement inégalitaire.

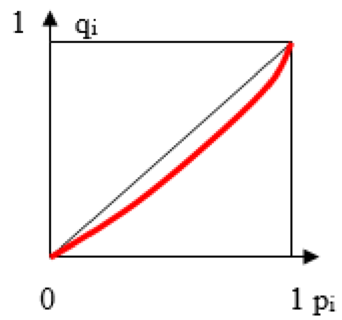
B : forte concentration : répartition fortement inégalitaire.

C : totale concentration : répartition parfaitement inégalitaire.

4) Exemple : La répartition d'ouvrier selon le salaire mensuel.

Salaire	n_i	f_i (%)	$p_i=F(e_i)$ (%)	c_i	$n_i c_i$	$\sum_{j=1}^{i} n_j c_j$	$q_i = \sum_{j=1}^{i} n_j c_j / \sum_{j=1}^{k} n_j c_j$ (%)
[50-100[	20	8	8	75	1500	1500	3
[50-100[	40	16	24	125	5000	6500	13,1
[50-100[	60	24	48	175	10500	1700	34,2
[50-100[	75	30	78	225	16875	33875	68,1
[50-100[	40	16	94	275	11000	44875	90,2
[50-100[	15	6	100	325	4875	49750	1
Total	250	100	-------------	------	49750	------------	--------------

8% des individus reçoivent 3% des salaires, 48 % des individus reçoivent moins de 200 D.



La courbe de concentration est proche de la première bissectrice d'où la concentration est faible : il s'agit d'une répartition faiblement inégalitaire de la masse des salaires.

II. L'indice de concentration

Objectifs spécifiques :

- Maîtriser le calcul de l'indice de concentration.

Durée : 0,45 H.

Contenu :

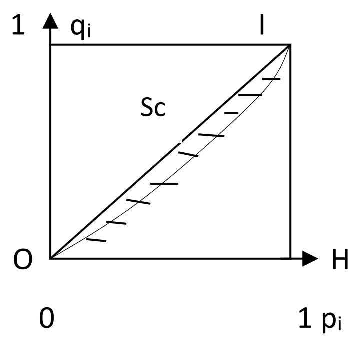
1. *Définition :*

On désigne par :

Sc : surface de concentration délimitée par la courbe de concentration et la diagonale OI.

S_{OHI} : Surface du triangle OHI.

L'indice de Gini noté i est défini par : i = Sc / S_{OHI} = ½. D'où, i = 2 Sc.



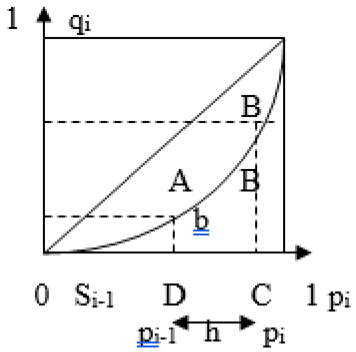
2. *Propriété :*

- 0 ≤ i ≤ 1 : Si i est proche de 0, il y'a une faible concentration et une faible inégalité. Si i est proche de 1 : il y'a une forte concentration et une forte inégalité.
- i est un nombre sans dimension, il permet ainsi de comparer la concentration de deux ou plusieurs distributions différentes.

3. *Calcul pratique par la méthode de Trapèzes :*

Sc : surface de concentration.

Si : surface du trapèze i : ABCD, avec Si = h (b + B) / 2.



D'après le graphique, on a : $S_{OHI} = Sc + \sum S_i$

$$S_{OHI} = Sc + \sum_{i=1}^{k} (½)\,(p_i - p_{i-1})\,(q_{i-1} + q_i)$$

$$S_{OHI} = Sc + (½) \sum_{i=1}^{k} f_i\,(q_{i-1} + q_i)$$

$$Sc = S_{OHI} - (½) \sum_{i=1}^{k} f_i\,(q_{i-1} - q_i) = (1/2) - (½) \sum_{i=1}^{k} f_i\,(q_{i-1} + q_i)$$

$i \approx 2\,Sc = 1 - \sum f_i\,(q_{i-1} + q_i)$, avec $p_0 = q_0 = 0$ pour la première classe.

Application :

Salaire	$q_{i-1} + q_i$	$f_i\,(q_{i-1} + q_i)$
[50-100[	0,03	0,002
[50-100[	0,161	0,026
[50-100[	0,473	0,114
[50-100[	1,023	0,307
[50-100[	1,583	0,253
[50-100[	1,902	0,114
Total	----------	0,816

i = 0,184 : proche de 0, il y'a une faible concentration, la répartition de la masse des salariés entre employés est légèrement inégalitaire, ce qui confirme la conclusion dégagé à partir de la courbe de Lorenz.

III. La médiale

Objectifs spécifiques :

- Identifier la notion de médiale.

Durée : 0,5 H.

Contenu :

1. *Définition :* C'est la valeur de la variable notée M_1 telle que $q(M_1) = 0,5$.

Elle partage la masse totale de la variable ($\sum_{i=1}^{k} n_j c_j$) en deux parties égales.

2. *Détermination :*

Deux cas peuvent se présenter :

- Il existe sur le tableau statistique une valeur e_i telle que $q(e_i) = 0,5$, alors $M_l = e_i$.
- Sinon, en premier lieu, on détermine la classe médiale $[e_{i-1}, e_i[$ telle que :

$q(e_{i-1}) < q(M_l) < q(e_i)$. En deuxième lieu, on détermine M_l par interpolation linéaire :

$M_l = e_{i-1} + [(e_i - e_{i-1}) / (q(e_i) - q(e_{i-1}))]\ [q(M_l) - q(e_{i-1})]$.

Application :

On a : $q(200) = 0,342 < 0,5 < q(250) = 0,681$

$M_l = 200 + [(250 - 200) / (0,681 - 0,342)]\ [0,5 - 0,342] = 223,304$ D.

Interprétation : La moitié de la masse des salaires est distribuée à l'ensemble des ouvriers dont le salaire individuel est inférieur à 223,304 (c.a.d. la médiale). L'autre moitié de la masse des salaires est distribuée à ceux dont le salaire individuel est supérieur à 223,304D.

3. *L'écart entre la médiale et la médiane :*

La médiale sert à déterminer un indicateur de concentration appelé « écart entre M_l et M_e.

$E = M_l - M_e$.

On a toujours, $E \geq 0$, car $M_l \geq M_e$.

Interprétation :

Si $E = 0$: $M_l = M_e$: La concentration est nulle, la répartition est parfaitement égalitaire.

Plus E est grand, plus la concentration est forte, plus la répartition est inégalitaire.

Plus E est petit, plus la concentration est faible, plus la répartition est égalitaire.