



« Mon intention, disait Malthus, est seulement de montrer que le principe des lois en faveur des pauvres repose sur une erreur » (MALTHUS «Essai sur le principe de population »)
« Elles ne tendent qu'à appauvrir les riches » (RICARDO : « Principes »)

Les données

L'Office Européen EUROSTAT adopte depuis le 01 janvier 2021, une classification territoriale appelée NUTS : **Nomenclature of territorial units for statistics**

Trois niveaux d'agrégation sont distingués et dénommés :

NUTS1 : 92 régions

NUTS2 : 242 régions

NUTS3 : 1166 régions.

On dispose ci-dessous au niveau NUTS3, d'un échantillon de 300 régions pour lesquelles a été dénombré le nombre d'enfants de trois ans scolarisés (Variable x_i = nombre d'enfants) en 2020.

Le tableau ci-dessous (voir plus bas) résulte d'un tri croissant des 300 régions (col 1) selon ce nombre d'enfants, de $x_{MIN} = 20$ à $x_{MAX} = 115305$ (col 2). Ce tri donne le rang de 1 à 300 (col 3). (REMARQUE : Une version EXCEL de ce tableau est disponible)

On souhaite réaliser un traitement de cette distribution statistique, en répondant aux questions suivantes EN TROIS PARTIES :

PARTIE 1°) TRAITEMENT STATISTIQUE CLASSIQUE

- 1- La variable x_i est *discrète*. Est-il souhaitable de la traiter comme une *variable continue* ?

Le nombre d'enfants est un objet non divisible. Donc la variable est discrète. Elle prend 300 modalités (300 nombres d'enfants différents). Il est souhaitable de la traiter comme une variable continue, dont les modalités sont regroupées en classes. Car, dans le cas contraire, la colonne n_i serait pour chaque modalité égale à 1 (chaque région ayant son nombre d'enfants), le total étant $N = 300$, l'effectif global.

- 2- Quel serait le nombre de classes optimal (N) si on se base sur les règles de Brooks-Carruthers, et sur celle de Huntsberger.

Les deux règles donnent *le nombre de classes optimal (k)*, connaissant l'effectif global (N):

règle de Brooks-Carruthers : $k < 5 + \log N$

règle de Huntsberger : $k \approx 1 + (3,3 \times \log N)$

Pour $N = 300$ Brooks-Carruthers $k < 5 + \log(300) \rightarrow k < 5 + 2,477 = 7,5$ classes, soit 7 classes

Huntsberger $k \approx 1 + 3,3 \log(300) \rightarrow k \approx 1 + 3,3 \times 2,477 = 9$ classes.

Il est possible d'opter pour l'une ou l'autre 7 ou 9, mais il est préférable d'opter pour la valeur intermédiaire (moyenne), soit $k = 8$ classes.

- 3- Appliquer ce nombre optimal de classes au tableau ci-dessous, pour CONSTITUER DES CLASSES, de x_{Min} à x_{Max} , ayant le même **EFFECTIF MOYEN** (une classe pouvant diverger). Réaliser pour cela le tableau de distribution en N classes.

L'effectif moyen (n_i) que devrait contenir chacune des 8 classes est donné par :

$n_i = N / k = 300 / 8 = 37,5$ soit 37 pays par classe, le reliquat (x) pouvant être versé dans la dernière classe (« fourre tout »). Soit $300 = (7 \times 37) + x \rightarrow x = 300 - 259 = 41$ qui sera l'effectif de la dernière classe.

L'ébauche du tableau de distribution est alors, x_{MIN} et x_{MAX} étant connus :

Ebauche du tableau de distribution

| classes i | x_{i-} | x_{i+} | n_i |
|-------------|----------|----------|-------|
| 1 | 20 | ? | 37 |
| 2 | ? | ? | 37 |
| 3 | ? | ? | 37 |
| 4 | ? | ? | 37 |
| 5 | ? | ? | 37 |
| 6 | ? | ? | 37 |
| 7 | ? | ? | 37 |
| 8 | ? | 115305 | 41 |
| Ensemble | | | 300 |

rang 1 = xMIN ou première région en haut de l'énoncé

? bornes de classes à combler

rang 300 = xMAX ou dernière région en bas de l'énoncé

On détermine les bornes manquantes, en incrémentant dans le tableau de l'énoncé de 37 régions en 37 régions.

Soit pour la première classe :

| RÉGIONS | Variable | RANG |
|---------------------------------------|----------|------|
| 2020 | x_i | |
| Kentrikschweiz | 229 | 1 |
| Åland | 281 | 2 |
| Région lémanique | 331 | 3 |
| Ionia Nisia | 675 | 4 |
| Dorsio Aigalo | 695 | 5 |
| Valle d'Aosta/Valleé d'Aoste | 912 | 6 |
| Dytiki Makedonia | 929 | 7 |
| Ciudad de Ceuta | 996 | 8 |
| Kastamonu, Çankiri, Sinop | 1178 | 9 |
| Sierca Ellada | 1182 | 10 |
| Ciudad de Melilla | 1205 | 11 |
| Notio Aigalo | 1244 | 12 |
| Ipeiros | 1307 | 13 |
| Çonguldak, Karabük, Bartın | 1319 | 14 |
| Agri, Kars, Iğdir, Ardahan | 1565 | 15 |
| Peloponnisos | 1577 | 16 |
| Erzurum, Erzurum, Bayburt | 1666 | 17 |
| Ticino | 1726 | 18 |
| Região Autónoma dos Açores (PT) | 1769 | 19 |
| Dytiki Ellada | 1796 | 20 |
| Região Autónoma da Madeira (PT) | 1832 | 21 |
| Kirillikale, Aksaray, Niğde, Nevşehir | 1835 | 22 |
| Molise | 1920 | 23 |
| Malatya, Elazığ, Bingöl, Tunceli | 2029 | 24 |
| Thessalia | 2264 | 25 |
| Anatoliki Makedonia, Thraki | 2314 | 26 |
| Van, Mus, Bitlis, Hakkari | 2350 | 27 |
| Burgentland (AT) | 2369 | 28 |
| Mardin, Batman, Sırnak, Şirnak | 2462 | 29 |
| Kriti | 2478 | 30 |
| Kayseri, Sivas, Yozgat | 2495 | 31 |
| Konya, Karaman | 2579 | 32 |
| La Rioja | 2720 | 33 |
| Gaziantep, Adiyaman, Kilis | 2812 | 34 |
| Corse | 2907 | 35 |
| Prov. Luxembourg (BE) | 3166 | 36 |
| Kuzeydogu Anadolu | 3231 | 37 |

On lit : $n_1 = 37$ régions $\rightarrow x_{i+} = 3231$

donc la première classe s'écrit $[20 ; 3231[$ et comporte 37 régions.

Ce qui permet de compléter le tableau de distribution :

| classes i | x_{i-} | x_{i+} | n_i |
|-------------|----------|----------|-------|
| 1 | 20 | 3231 | 37 |
| 2 | 3231 | ? | 37 |
| 3 | ? | ? | 37 |
| 4 | ? | ? | 37 |
| 5 | ? | ? | 37 |
| 6 | ? | ? | 37 |
| 7 | ? | ? | 37 |
| 8 | ? | 115305 | 41 |
| Ensemble | | | 300 |

Pour la seconde : dans la colonne « rang » : $37 + 37 = 74$, donc :

| | | |
|--------------------------------------|------|----|
| Kuzeydogu Anadolu | 3231 | 37 |
| Kärnten | 3318 | 38 |
| Algarve | 3349 | 39 |
| Vorarlberg | 3446 | 40 |
| Balikesir, Çanakkale | 3561 | 41 |
| Innlandet | 3571 | 42 |
| Martinique | 3632 | 43 |
| Hatay, Kahramanmaraş, Osmaniye | 3644 | 44 |
| Tekirdağ, Edirne, Kırklareli | 3748 | 45 |
| Basilicata | 3896 | 46 |
| Manisa, Afyonkarahisar, Kütahya | 3943 | 47 |
| Mellersta Norrland | 4008 | 48 |
| Salzburg | 4049 | 49 |
| Prov. Brabant wallon | 4133 | 50 |
| Samsun, Tokat, Çorum, Amasya | 4133 | 51 |
| Doğu Karadeniz | 4197 | 52 |
| Trabzon, Ordu, Giresun, Rize, Artvin | 4197 | 53 |
| Orta Anadolu | 4330 | 54 |
| Cantabria | 4366 | 55 |
| Ortadoğu Anadolu | 4379 | 56 |
| Malta | 4391 | 57 |
| Nisia Aigalou, Kriti | 4414 | 58 |
| Provincia Autonoma di Trento | 4485 | 59 |
| Severozapaden | 4671 | 60 |
| Suyane | 4794 | 61 |
| Nord-Norge | 4819 | 62 |
| Provincia Autonoma di Bolzano/Bd | 4856 | 63 |
| Mayotte | 4878 | 64 |
| Alentejo | 4903 | 65 |
| Severen tsenträlen | 4936 | 66 |
| Guaadeloupe | 5038 | 67 |
| Prov. Namur | 5076 | 68 |
| Trøndelag | 5152 | 69 |
| Bremen | 5331 | 70 |
| Kentriki Makedonia | 5491 | 71 |
| Övre Norrland | 5502 | 72 |
| Region Juzne i Istocne Srbije | 5640 | 73 |
| Sanliurfa, Diyarbakir | 5655 | 74 |

On lit : $n_2 = 37$ régions $\rightarrow x_{i+} = 5665$

donc la seconde classe s'écrit $[3231 ; 5665[$ et comporte 37 régions.

Ce qui permet de compléter le tableau de distribution :

| classes i | x_{i-} | x_{i+} | n_i |
|-------------|----------|----------|-------|
| 1 | 20 | 3231 | 37 |
| 2 | 3231 | 5665 | 37 |
| 3 | 5665 | ? | 37 |
| 4 | ? | ? | 37 |
| 5 | ? | ? | 37 |
| 6 | ? | ? | 37 |
| 7 | ? | ? | 37 |
| 8 | ? | 115305 | 41 |
| Ensemble | | | 300 |

L'opération est identique (ajouter 37) jusqu'à la 7eme classe. Le reliquat de 4 régions étant versé dans la dernière classe pour laquelle $x_{i+} = x_{MAX} = 115305$.

Soit alors l'ébauche complète du tableau de distribution

| classes i | x_{i-} | x_{i+} | n_i |
|-------------|----------|----------|-------|
| 1 | 20 | 3231 | 37 |
| 2 | 3231 | 5665 | 37 |
| 3 | 5665 | 9157 | 37 |
| 4 | 9157 | 11955 | 37 |
| 5 | 11955 | 16155 | 37 |
| 6 | 16155 | 23529 | 37 |
| 7 | 23529 | 36275 | 37 |
| 8 | 36275 | 115305 | 41 |
| Ensemble | | | 300 |

- 4- PASSER A L'ETAPE de la mesure de la **TENDANCE CENTRALE**, en vous limitant à LA MOYENNE, au MODE, et à la MEDIANE.

On calcule dans le tableau TOUTES les valeurs requises pour déterminer les indicateurs de tendance centrale :

$$\bar{x} = \sum_{i=1}^n f_i \cdot c_{xi}$$

a) La MOYENNE théorique

b) La MEDIANE $x_{Mé} = x_{i-} + \left(a_i \frac{F(x_{Mé}) - F(x_{i-})}{F(x_{i+}) - F(x_{i-})} \right)$

c) Le MODE $x_{Mo} = x_i / (f_i/a_i) \% MAX$

Le tableau de distribution devient :

| classe i | x_{i-} | x_{i+} | n_i | f_i | $f_i\%$ | $F(x_{i+})$ | $F(x_{i-})$ | a_i | f_i/a_i | $f_i/a_i * 100000$ |
|------------|----------|----------|-------|-------|---------|-------------|-------------|---------|------------|--------------------|
| 1 | 20 | 3231 | 37 | 0,123 | 12,33% | 12,3% | 0,0% | 3211 | 3,841E-05 | 3,84 |
| 2 | 3231 | 5665 | 37 | 0,123 | 12,33% | 24,7% | 12,3% | 2434 | 5,0671E-05 | 5,07 |
| 3 | 5665 | 9157 | 37 | 0,123 | 12,33% | 37,0% | 24,7% | 3492 | 3,5319E-05 | 3,53 |
| 4 | 9157 | 11955 | 37 | 0,123 | 12,33% | 49,3% | 37,0% | 2798 | 4,4079E-05 | 4,41 |
| 5 | 11955 | 16155 | 37 | 0,123 | 12,33% | 61,7% | 49,3% | 4200 | 2,9365E-05 | 2,94 |
| 6 | 16155 | 23529 | 37 | 0,123 | 12,33% | 74,0% | 61,7% | 7374 | 1,6725E-05 | 1,67 |
| 7 | 23529 | 36275 | 37 | 0,123 | 12,33% | 86,3% | 74,0% | 12746 | 9,6762E-06 | 0,97 |
| 8 | 36275 | 115305 | 41 | 0,137 | 13,67% | 100,0% | 86,3% | 79030 | 1,7293E-06 | 0,17 |
| Ens | | | 300 | 1 | 100% | | 100,0% | 115 285 | | |

a) La MOYENNE

La formule de la moyenne consiste en une estimation. Or, il est possible dans certains cas de déterminer non la moyenne théorique, mais la MOYENNE REELLE. Ce cas est celui où on connaît la masse à répartir (MS), et l'effectif global (N).

C'est le cas dans cet exemple où :

MS = Nombre total d'enfants = 5832142 (total de la colonne « variable x_i ») et N = nombre total de régions = 300

Donc \bar{x} REELLE = MS/N = 5832142/300 = **19440**

b) La MEDIANE

Le tableau de distribution doit être complété de la manière suivante :

| classe i | x_{i-} | x_{i+} | n_i | f_i | $f_i\%$ | $F(x_{i+})$ | $F(x_{i-})$ |
|------------|----------|----------|-------|-------|---------|-------------|-------------|
| 1 | 20 | 3231 | 37 | 0,123 | 12,33% | 12,3% | 0,0% |
| 2 | 3231 | 5665 | 37 | 0,123 | 12,33% | 24,7% | 12,3% |
| 3 | 5665 | 9157 | 37 | 0,123 | 12,33% | 37,0% | 24,7% |
| 4 | 9157 | 11955 | 37 | 0,123 | 12,33% | 49,3% | 37,0% |
| 5 | 11955 | 16155 | 37 | 0,123 | 12,33% | 61,7% | 49,3% |
| 6 | 16155 | 23529 | 37 | 0,123 | 12,33% | 74,0% | 61,7% |
| 7 | 23529 | 36275 | 37 | 0,123 | 12,33% | 86,3% | 74,0% |
| 8 | 36275 | 115305 | 41 | 0,137 | 13,67% | 100,0% | 86,3% |
| Ens | | | 300 | 1 | 100% | | 100,0% |

Puisque $x_{M\acute{e}} = x_{i-} + \left(a_i \frac{F(x_{M\acute{e}}) - F(x_{i-})}{F(x_{i+}) - F(x_{i-})} \right)$, les valeurs encadrantes sont :

$F(x_{i-}) = 49,3\%$; $F(x_{i+}) = 61,7\%$; $x_{i-} = 11955$; $x_{i+} = 16155$

Donc $x_{M\acute{e}} = 11955 + (16155 - 11955) (61,7 - 50 / 61,5 - 49,3) = 15918$

c) Le mode $x_{M0} = [3231 ; 5665[/ (f_i/a_i) \times 100.000 = 5,07 \text{ MAX}$, soit le centre de classe $Cx2 = 4448$

5- Illustrer votre résultat en réalisant **L'HISTOGRAMME** de la distribution et le **POLYGONE** des fréquences.

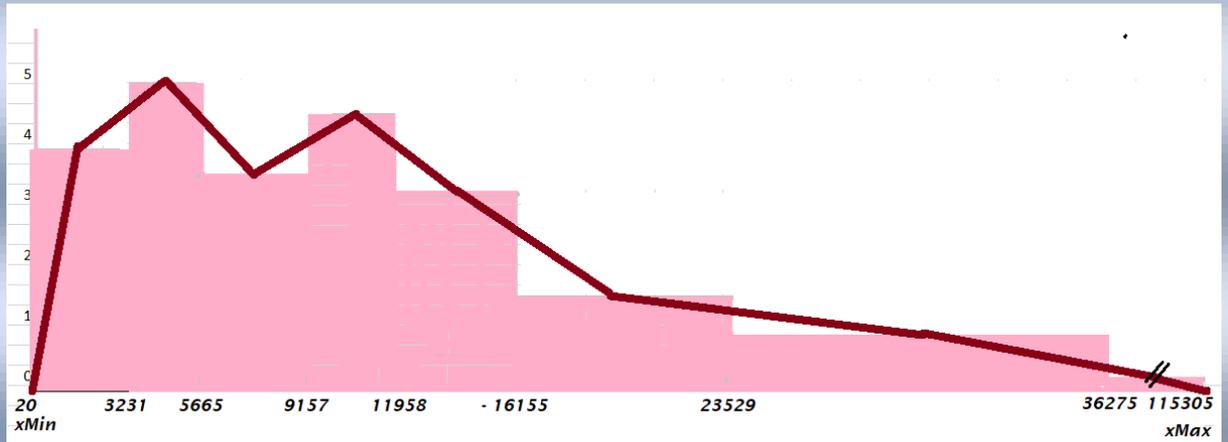
En ordonnée : calcul des fréquences par unité d'amplitude (f_i/a_i) et (f_i/a_i)% sur lesquelles on opère toutes les transformations souhaitables pour rendre l'ordonnée lisible, et simple à construire . On choisit ici de calculer $(f_i/a_i) \times 100000$ (voir tableau ci-dessous).

En abscisse : les bornes x_i , à l'échelle. Pour déterminer l'échelle, on effectue les transformations ci-dessous, qui montre que les classes 1 à 7 exigent 15,7 cms d'abscisse totale, seule la dernière possède une amplitude démesurée ($a_8/2 = 34,3$ cms !). Pour celle-ci on adoptera la convention de la discontinuité (voir tableau ci-dessous et l'histogramme).

Soit les calculs :

| abscisse | | | | | | Ordonnée | |
|----------|--------|-----------|---------------------|------------|------------------|---------------|------------------------|
| f_i | a_i | a_i/e | $(col2) \times 100$ | $(col3)/2$ | $(col4)$ cumulée | (f_i/a_i) | $(col6) \times 100000$ |
| (Col0) | col1 | col2 | col3 | col4 (cms) | col5 (cms) | col6 | col7 |
| 0,123 | 3211 | 0,0278527 | 2,8 | 1,4 | 1,4 | 0,00003840963 | 3,8 |
| 0,123 | 2434 | 0,0211129 | 2,1 | 1,1 | 2,4 | 0,00005067105 | 5,1 |
| 0,123 | 3492 | 0,0302902 | 3,0 | 1,5 | 4,0 | 0,00003531882 | 3,5 |
| 0,123 | 2798 | 0,0242703 | 2,4 | 1,2 | 5,2 | 0,00004407910 | 4,4 |
| 0,123 | 4200 | 0,0364315 | 3,6 | 1,8 | 7,0 | 0,00002936508 | 2,9 |
| 0,123 | 7374 | 0,0639632 | 6,4 | 3,2 | 10,2 | 0,00001672543 | 1,7 |
| 0,123 | 12746 | 0,1105608 | 11,1 | 5,5 | 15,7 | 0,00000967624 | 1,0 |
| 0,137 | 79030 | 0,6855185 | 68,6 | 34,3 | 50,0 | 0,00000172930 | 0,2 |
| 1 | 115285 | 1 | 100,0 | 50,0 | | | |

L'histogramme correspondant, et le *polygone des fréquences (courbe)* :



6- PASSER A L'ETAPE de la mesure de la DISPERSION, en allant jusqu'au CV = **Coefficient de variation (en %)**

On applique successivement les formules suivantes :

le coefficient de variation (en %) $CV = \frac{\sigma_{(x)}}{\bar{x}} \times 100\%$

L'écart type $\sigma = (\sigma^2)^{1/2}$ = racine carrée de la variance

La variance $\sigma^2 = \sum_{i=1}^n (f_i \cdot x_i^2) - (\bar{x})^2$ Signifie que La variance est égale à la moyenne des carrés des valeurs moins le carré de la moyenne

= Moyenne des carrés des valeurs – carré de la moyenne

Le tableau de distribution doit être complété

| classe i | xi- | xi+ | ni | fi | fi% | F(xi+) | F(xi-) | ai | fi/ai | fi/ai * 100000 | Moyenne des carrés des valeurs | | |
|----------|-------|--------|-----|-------|--------|--------|--------|---------|------------|----------------|--------------------------------|------------------|--------------------|
| | | | | | | | | | | | Cxi | Cxi ² | fiCxi ² |
| 1 | 20 | 3231 | 37 | 0,123 | 12,33% | 12,3% | 0,0% | 3211 | 3,841E-05 | 3,84 | 1625,5 | 2642250,25 | 325877,53 |
| 2 | 3231 | 5665 | 37 | 0,123 | 12,33% | 24,7% | 12,3% | 2434 | 5,0671E-05 | 5,07 | 4448 | 19784704 | 2440113,5 |
| 3 | 5665 | 9157 | 37 | 0,123 | 12,33% | 37,0% | 24,7% | 3492 | 3,5319E-05 | 3,53 | 7411 | 54922921 | 6773826,9 |
| 4 | 9157 | 11955 | 37 | 0,123 | 12,33% | 49,3% | 37,0% | 2798 | 4,4079E-05 | 4,41 | 10556 | 111429136 | 13742927 |
| 5 | 11955 | 16155 | 37 | 0,123 | 12,33% | 61,7% | 49,3% | 4200 | 2,9365E-05 | 2,94 | 14055 | 197543025 | 24363640 |
| 6 | 16155 | 23529 | 37 | 0,123 | 12,33% | 74,0% | 61,7% | 7374 | 1,6725E-05 | 1,67 | 19842 | 393704964 | 48556946 |
| 7 | 23529 | 36275 | 37 | 0,123 | 12,33% | 86,3% | 74,0% | 12746 | 9,6762E-06 | 0,97 | 29902 | 894129604 | 110275984 |
| 8 | 36275 | 115305 | 41 | 0,137 | 13,67% | 100,0% | 86,3% | 79030 | 1,7293E-06 | 0,17 | 75790 | 5744124100 | 785030294 |
| Ens | | | 300 | 1 | 100% | | 100,0% | 115 285 | | | | | 991509608 |

$$\sigma^2 = \sum_{i=1}^n (f_i \cdot x_i^2) - (\bar{x})^2 \quad (\text{avec } x_i = Cx_i) = 991509608 - 19440^2 = 991509608 - 377913600 = 613596008$$

$$\sigma = (\sigma^2)^{1/2} = (613596008)^{1/2} = 24770,87$$

$$CV = \frac{\sigma_{(x)}}{\bar{x}} \times 100\% = (24770,87/19440) \times 100\% = 127\%$$

PARTIE 2°) LA METHODE DE L'AMPLITUDE MOYENNE UNIFORME

- 7- On veut s'affranchir des règles de Brooks-Carruthers, et de Huntsberger. On choisit de constituer les CLASSES en leur donnant la **MEME AMPLITUDE MOYENNE** (que l'on appellera (\bar{a}_i)). **Calculer \bar{a}_i , l'amplitude moyenne, à l'aide du tableau de distribution (de l'étape 1).**

Calcul de (\bar{a}_i) l'amplitude moyenne. On dispose dans le tableau distribution des colonnes :

| classe i | x_{i-} | x_{i+} | a_i |
|------------|----------|----------|---------|
| 1 | 20 | 3231 | 3211 |
| 2 | 3231 | 5665 | 2434 |
| 3 | 5665 | 9157 | 3492 |
| 4 | 9157 | 11955 | 2798 |
| 5 | 11955 | 16155 | 4200 |
| 6 | 16155 | 23529 | 7374 |
| 7 | 23529 | 36275 | 12746 |
| 8 | 36275 | 115305 | 79030 |
| Ens | | | 115 285 |

L'amplitude moyenne serait égal à :

$$\bar{a}_i = \text{étendue} / \text{nb classes} = 115285/8 = 14410,625$$

Pour éviter de travailler avec cette étendue décimales, la seule exacte (puisque $14410,625 \times 8 = 115285$, l'étendue), on accepte d'arrondir, soit à 14410, soit à 14411. On retient ici $\bar{a}_i = 14411$ ce qui éloigne le moins de l'étendue exacte (puisque $14411 \times 8 = 115288 \cong 115285$).

On parvient ainsi à un tableau dont les seuls éléments connus sont :

| classe i | x_{i-} | x_{i+} | a_i |
|------------|----------|----------|--------|
| 1 | 20 | | 14411 |
| 2 | | | 14411 |
| 3 | | | 14411 |
| 4 | | | 14411 |
| 5 | | | 14411 |
| 6 | | | 14411 |
| 7 | | | 14411 |
| 8 | | 115308 | 14411 |
| Ens | | | 115288 |

On connaît

L'amplitude moyenne = $\bar{a}_i = 14411$

$x_{\text{Min}} = 20$

$x_{\text{Max}} = 115308$ (au lieu de 115305) légèrement modifié par nous du fait du choix d'arrondi, de sorte que l'étendue

$$x_{\text{Max}} - x_{\text{Min}} = e = 115308 - 20 = 115288$$

$$= \sum_{i=1}^8 \bar{a}_i = 8 \times 14411 = 115288$$

- 8- Ayant l'amplitude de chaque classe, retraiter le tableau des données ci-dessous, pour créer un nouveau tableau de distribution, comportant les NOUVELLES CLASSES (de x_{MIN} à x_{MAX}).

Le choix d'une amplitude unique entraîne un retraitement du tableau de l'énoncé. Ce retraitement consiste à dénombrer les régions, selon les 8 classes, toutes d'amplitude ($x_{i+} - x_{i-} = \bar{a}_i = 14411$).

Concrètement, on cherche les bornes de classe manquantes ($x_{i-}; x_{i+}$), partant de $x_{\text{Min}} = 20$, de sorte que chaque classe soit conforme à une amplitude $\bar{a}_i = 14411 = x_{i+} - x_{i-}$

Soit alors les bornes de classes suivantes obtenues par incrémentations successives égales à 14411 :

| classe i | x_{i-} | x_{i+} | a_i |
|------------|----------|----------|--------|
| 1 | 20 | 14431 | 14411 |
| 2 | 14431 | 28841 | 14411 |
| 3 | 28841 | 43252 | 14411 |
| 4 | 43252 | 57663 | 14411 |
| 5 | 57663 | 72073 | 14411 |
| 6 | 72073 | 86484 | 14411 |
| 7 | 86484 | 100894 | 14411 |
| 8 | 100894 | 115305 | 14411 |
| Ens | | | 115288 |

Les bornes de classes étant trouvées, il reste à dénombrer le nombre de régions appartenant à chaque classe, c'est-à-dire les *effectifs catégoriels* n_i . Soit :

| classe i | x_{i-} | x_{i+} | a_i | n_i |
|------------|----------|----------|--------|-------|
| 1 | 20 | 14431 | 14411 | 167 |
| 2 | 14431 | 28841 | 14411 | 67 |
| 3 | 28841 | 43252 | 14411 | 32 |
| 4 | 43252 | 57663 | 14411 | 12 |
| 5 | 57663 | 72073 | 14411 | 10 |
| 6 | 72073 | 86484 | 14411 | 6 |
| 7 | 86484 | 100894 | 14411 | 2 |
| 8 | 100894 | 115305 | 14411 | 4 |
| Ens | | | 115288 | 300 |

Tableau qui est celui de la distribution des 300 régions selon le nombre d'enfants scolarisés, en 8 classes d'amplitude égale.

- 9- PASSER A L'ETAPE de la mesure de la TENDANCE CENTRALE, en vous limitant à LA MOYENNE, au MODE, et à la MEDIANE.

La Tendance centrale est calculée de la même manière que précédemment. Les colonnes nécessaires dans le nouveau tableau étant :

| x_{i-} | x_{i+} | a_i | n_i | f_i | $f_i\%$ | $F(x_{i+})$ | $F(x_{i-})$ | $(f_i/a_i)\%$ | (f_i/a_i) pour 10000 |
|----------|----------|--------|-------|-------|---------|-------------|-------------|---------------|------------------------|
| 20 | 14431 | 14411 | 167 | 0,557 | 55,7% | 55,7% | 0,0% | 0,0039% | 3,8628% |
| 14431 | 28841 | 14411 | 67 | 0,223 | 22,3% | 78,0% | 55,7% | 0,0015% | 1,5497% |
| 28841 | 43252 | 14411 | 32 | 0,107 | 10,7% | 88,7% | 78,0% | 0,0007% | 0,7402% |
| 43252 | 57663 | 14411 | 12 | 0,040 | 4,0% | 92,7% | 88,7% | 0,0003% | 0,2776% |
| 57663 | 72073 | 14411 | 10 | 0,033 | 3,3% | 96,0% | 92,7% | 0,0002% | 0,2313% |
| 72073 | 86484 | 14411 | 6 | 0,020 | 2,0% | 98,0% | 96,0% | 0,0001% | 0,1388% |
| 86484 | 100894 | 14411 | 2 | 0,007 | 0,7% | 98,7% | 98,0% | 0,0000% | 0,0463% |
| 100894 | 115305 | 14411 | 4 | 0,013 | 1,3% | 100,0% | 98,7% | 0,0001% | 0,0925% |
| | | 115285 | 300 | 1 | 100,0% | | 1 | | |

a) La MOYENNE est comme précédemment la MOYENNE REELLE et ne change donc pas :

$$\bar{x} \text{ REELLE} = MS/N = 5832142/300 = \mathbf{19440}$$

d) La MEDIANE $\bar{x}_{Mé} = x_i^- + \left(a_i \frac{F(x_{Mé}) - F(x_i^-)}{F(x_i^+) - F(x_i^-)} \right)$

Les valeurs encadrantes lisibles dans le tableau sont :

$x_{i-} = 20$ et $x_{i+} = 14431$, $F(x_{i-}) = 0\%$ et $F(x_{i+}) = 55,7\%$

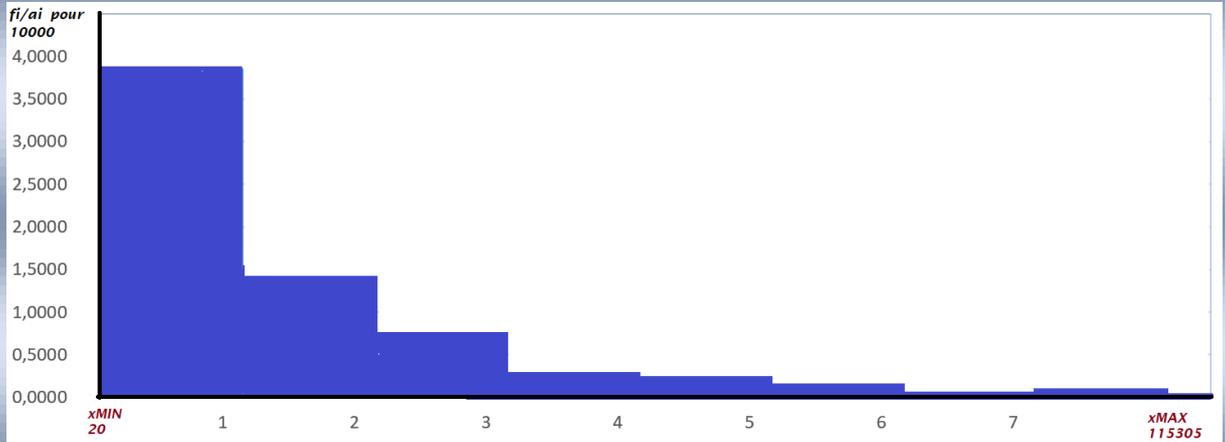
donc $\bar{x}_{Mé} = 20 + (14411(50/55,7)) = 20 + (14411(0,8976)) = 20 + 12936 = 12956$

c) Le MODE $\bar{x}_{Mo} = x_i / (f_i/a_i)\% \text{ MAX}$

La classe modale est la classe 1 soit (« 20 ; 14431 ») puisque $a_1 = 14431 - 20 = 14411$) dont la fréquence par unité d'amplitude est

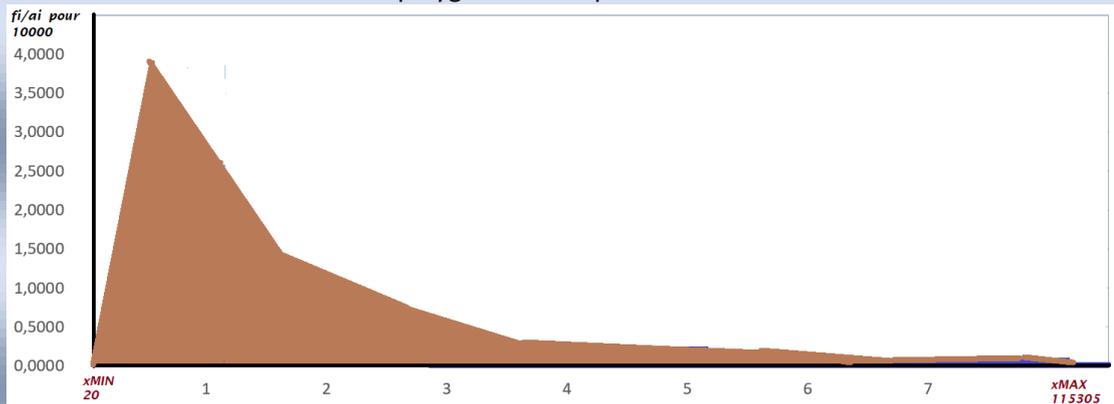
$((f_i/a_i) \text{ pour } 10000) = 3,86 \text{ MAX}$; Le Centre de classe modal étant $Cx_1 = (20+14431)/2 = 14451/2 = 7225$

10- Illustrer votre résultat en réalisant L'HISTOGRAMME de la distribution et le POLYGONE des fréquences.



(Les nombres 1 à 7 indiquent les bornes de classes, entre lesquelles la différence est constante = 14411)- voir tableau-

On en déduit immédiatement le polygone de fréquences



11- PASSER A L'ETAPE de la mesure de la DISPERSION, en allant jusqu'au CV = Coefficient de variation (en %)

Le complément (orange) au tableau permet le calcul de la dispersion

| xi- | xi+ | ai | ni | fi | fi% | F(xi+) | F(xi-) | (fi/ai)% | (fi/ai) pour 10000 | Cxi | cxi ² | fiCxi ² |
|--------|--------|--------|-----|-------|--------|--------|--------|----------|--------------------|----------|------------------|--------------------|
| 20 | 14431 | 14411 | 167 | 0,557 | 55,7% | 55,7% | 0,0% | 0,0039% | 3,8628% | 7225,313 | 52205140,72 | 29060861,67 |
| 14431 | 28841 | 14411 | 67 | 0,223 | 22,3% | 78,0% | 55,7% | 0,0015% | 1,5497% | 21635,94 | 468113899,7 | 104545437,6 |
| 28841 | 43252 | 14411 | 32 | 0,107 | 10,7% | 88,7% | 78,0% | 0,0007% | 0,7402% | 36046,57 | 1299355209 | 138597888,9 |
| 43252 | 57663 | 14411 | 12 | 0,040 | 4,0% | 92,7% | 88,7% | 0,0003% | 0,2776% | 50457,2 | 2545929032 | 101837161,3 |
| 57663 | 72073 | 14411 | 10 | 0,033 | 3,3% | 96,0% | 92,7% | 0,0002% | 0,2313% | 64867,83 | 4207835369 | 140261179 |
| 72073 | 86484 | 14411 | 6 | 0,020 | 2,0% | 98,0% | 96,0% | 0,0001% | 0,1388% | 79278,46 | 6285074220 | 125701484,4 |
| 86484 | 100894 | 14411 | 2 | 0,007 | 0,7% | 98,7% | 98,0% | 0,0000% | 0,0463% | 93689,09 | 8777645585 | 58517637,23 |
| 100894 | 115305 | 14411 | 4 | 0,013 | 1,3% | 100,0% | 98,7% | 0,0001% | 0,0925% | 108099,7 | 11685549464 | 155807326,2 |
| | | 115285 | 300 | 1 | 100,0% | | 1 | | | | | 854328976,3 |

$$\sigma^2 = \sum_{i=1}^n (f_i \cdot x_i^2) - (\bar{x})^2 \quad (\text{avec } x_i = Cx_i) = 854328976,3 - 19440^2 = 854328976,3 - 377913600 = 476415376,3$$

$$\sigma = (\sigma^2)^{1/2} = (476415376,3)^{1/2} = 21827$$

$$CV = \frac{\sigma_{(x)}}{\bar{x}} \times 100\% = (21827/19440) \times 100\% = 112\%$$

PARTIE 3°) L'ANALYSE DES CONSEQUENCES DU CHANGEMENT DE METHODE

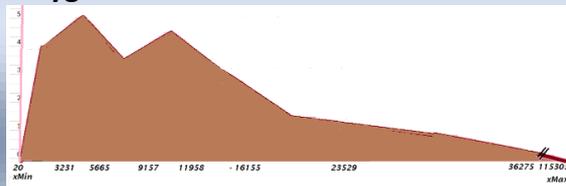
Dans cette partie, vous cherchez :

- 1) D'abord, à tirer tous les enseignements que vous pouvez, du passage de la partie 1 à la partie 2 (tendance centrale, dispersion, histogramme...) .

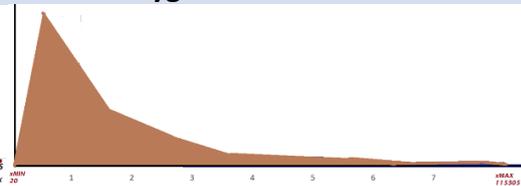
Suivant les résultats la modification de la méthode n'a qu'une faible incidence sur la **dispersion**. Les deux coefficients de variation étant très proches : 127% (méthode 1) et 112% (méthode 2). Dispersion qui dans les deux cas ressort comme élevée.

D'autres conséquences sont révélées par les deux polygones des fréquences, rappelés ci-dessous :

Polygone méthode 1



Polygone méthode 2



La relation de Pearson

La relation de Pearson étudie les relations entre indicateurs de position (ou de tendance centrale : CTC). Soit la comparaison

| CTC | Methode 1 | Methode 2 |
|---------|-----------|-----------|
| MOYENNE | 19440 | 19440 |
| MEDIANE | 15918 | 12956 |
| MODE | 4448 | 7225 |

Dans les deux cas, on observe que : $\bar{x} > xM_{\acute{e}} > xM_0$, inégalité qui permet de conclure que les deux distributions sont étalées à droite. Ce que prouve le polygone.

Toutefois si la distribution 1 paraît légèrement centrée, la distribution 2 est fortement tirée à gauche, vers les classes les plus petites.

- 2) Puis à comprendre l'effet sur la distribution du passage à une **amplitude moyenne uniforme**, en étudiant :

21) LA FORME : polygone des fréquences, et coefficients de forme ;
L'enseignement de la *fonction de répartition* (à réaliser pour chaque méthode, dans le même graphique) peut ici être ajoutée.

On calcule les indicateurs de forme : Kurtosis (ou aplatissement ou Gamma 2) et Skewness (ou asymétrie ou Gamma1). Les résultats donnés immédiatement sous EXCEL sont :

| | Methode 1 | Methode 2 |
|------------------|-----------|------------|
| KURTOSIS | 8 | 4,7 |
| ASYMETRIE | 2 | 1,7 |

Le signe positif du Kurtosis indique que les deux distributions connaissent bien un *pic*, et ne sont pas aplaties. La distribution 1 étant moins aplatie que la distribution 2 (Kurto 8 > 4).

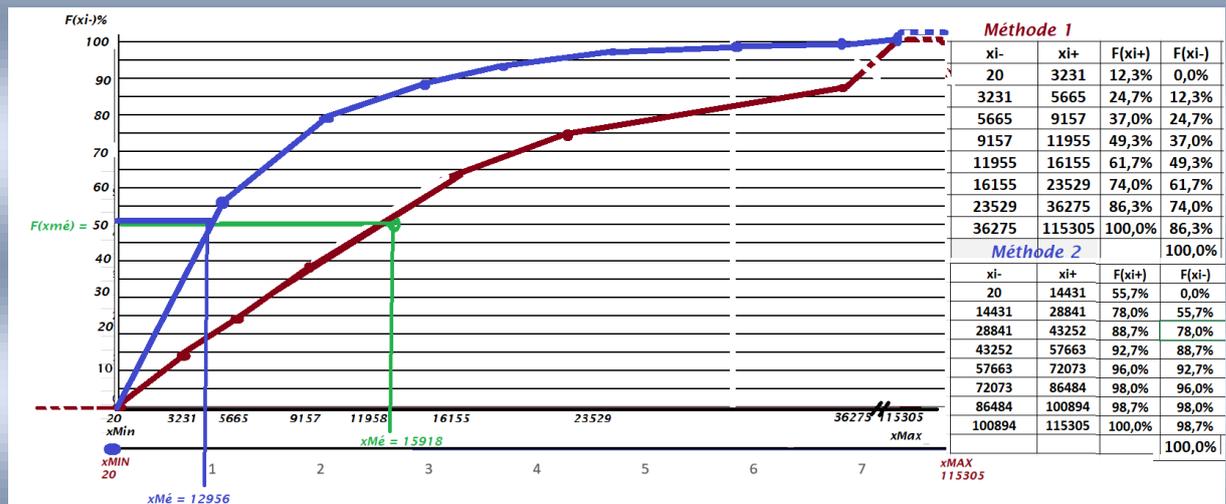
Le signe positif du coefficient d'asymétrie confirme la relation de Pearson. Les deux distributions sont bien étalées à droite, de manière grossièrement identique.

Conclusion :

le phénomène le plus marquant révélé par le Kurtosis et par le polygone des fréquences a trait aux conséquences du choix d'une amplitude uniforme, au lieu d'un équilibre des effectifs, comme le veut la règle. L'amplitude uniforme a *déséquilibré les effectifs catégoriels*. Elle a « gonflé » les effectifs des classes les plus petites, et minimisé ceux des classes les plus élevées. Ce phénomène semble même obéir à une proportionnalité, telle que le polygone 2 se traduit par l'apparition d'une courbe *strictement convexe* :



En représentant les deux fonctions de répartition, on peut observer ce *déversement des effectifs*, par le déplacement de la courbe vers le haut et la gauche (vers la courbe en bleu). La Médiane est en conséquence tirée à gauche.



(NB : ce graph comporte évidemment deux échelles d'abscisse différentes)

C'est en comparant à nouveau les deux méthodes que l'on peut trouver une explication mathématique à ce résultat.

| | Methode 1 | Methode 2 |
|-------------------|------------------------------|-------------------------|
| nombre de classes | 8 | 8 |
| Effectifs n_i | quasi constants ou donnés | variables et incertains |
| amplitude a_i | variable ou incertaine | constante ou donnée |
| bornes de classes | dépendantes des n_i | dépendantes de a_i |

On lit que les relations de dépendance diffèrent :

Méthode 1 : nb classes et N \rightarrow n_i \rightarrow bornes \rightarrow a_i amplitude \Leftrightarrow *effectifs présumés et amplitude résultante*

Méthode 2 : nb classes et N \rightarrow a_i amplitude \rightarrow bornes \rightarrow n_i \Leftrightarrow *amplitude présumée et effectifs résultants*

Mais la méthode 2 n'est aléatoire qu'en apparence. Les effectifs sont en fait dépendants de l'amplitude uniforme, car ils suivent une règle de distribution. En comprenant cette règle, on comprend la courbe convexe.

Le bon sens veut en effet qu'à priori le nombre d'enfants scolarisés est normalement élevé dans quelques régions très étendues (ou à forte densité), et faible dans nombre de régions dont la densité est restreinte. Donc à une forte amplitude est associé un faible effectif de régions, et à une faible amplitude est associé un effectif croissant de régions. C'est pourquoi, en adoptant une amplitude uniforme (égale à une moyenne), on réduit les amplitudes élevées, et on accroît les amplitudes faibles. Le mouvement naturel vers une amplitude moyenne est en effet :



Ce mouvement comporte un phénomène lié (et non aléatoire) de *déversement automatique des effectifs n_i , des classes aux amplitudes élevées, vers les classes aux amplitudes les plus faibles*, et ceci de manière progressive. Ceci jusqu'à concentrer les effectifs dans la classe modale 1, dont la fréquence $f_1\% = 55,7\%$ (ou $f_i/a_i = 3,9 \text{ p.}10000$).

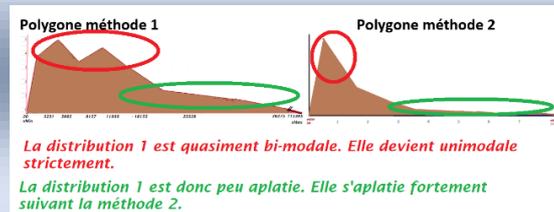
Mathématiquement ce phénomène peut être apparenté à celui des poids et mesures.

Il faut en effet peu de tonnes pour une énorme quantité de grammes. De même faut il peu de régions très denses, contre une grande quantité de petites régions. Ce qui est vrai, si et seulement si *existe une unité de mesure stable du poids* : le gramme, ou ici l'amplitude a_i . De sorte qu'il devient indifférent de raisonner sur $f_i\%$, ou bien $f_i/a_i\%$ (puisque l'on peut simplifier par le dénominateur), toute comme il est mathématiquement indifférent de raisonner en tonnes ou en grammes.

Telle est la conséquence principale du choix d'une amplitude moyenne et constante : *avoir déversé l'effectif des régions de manière continue et proportionnelle des plus denses régions, vers les moins denses*.

Cette méthode présente donc l'inconvénient de supprimer la part de l'aléa ou du hasard dans la distribution des effectifs, et donc celui de prédéterminer, et d'accentuer, la *forme de*

la distribution. Une partie de l'information résultant de la méthode 1 disparaît au centre. Ce que l'on peut illustrer par la comparaison ci-dessous :



22) LA CONCENTRATION (IG, xML, iV)

a) Le calcul de l'Indice de Gini (IG) nécessite les colonnes suivantes dans chaque tableau :

| METHODE CLASSIQUE | | | | | | | | |
|-------------------|--------|---------|---------|-----------|-------|-------|-------|-------------------|
| xi- | xi+ | Si | gi réel | gi réel % | Gi+ | Gi- | fi | fi[(Gi- + Gi+)/2] |
| 20 | 3231 | 64144 | 0,0110 | 1,100% | 0,011 | 0 | 0,123 | 0,0007 |
| 3231 | 5665 | 163942 | 0,0281 | 2,811% | 0,039 | 0,011 | 0,123 | 0,0031 |
| 5665 | 9157 | 274474 | 0,0471 | 4,706% | 0,086 | 0,039 | 0,123 | 0,0077 |
| 9157 | 11955 | 384091 | 0,0659 | 6,586% | 0,152 | 0,086 | 0,123 | 0,0147 |
| 11955 | 16155 | 524648 | 0,0900 | 8,996% | 0,242 | 0,152 | 0,123 | 0,0243 |
| 16155 | 23529 | 719785 | 0,1234 | 12,342% | 0,365 | 0,242 | 0,123 | 0,0375 |
| 23529 | 36275 | 1112633 | 0,1908 | 19,078% | 0,556 | 0,365 | 0,123 | 0,0568 |
| 36275 | 115305 | 2588425 | 0,4438 | 44,382% | 1,000 | 0,556 | 0,137 | 0,1063 |
| | | 5832142 | 1,0000 | 100,000% | | 1,000 | 1,000 | U = 0,2511 |

| METHODE DE L'AMPLITUDE MOYENNE UNIFORME | | | | | | | | |
|---|--------|---------|---------|-----------|-------|-------|-------|-------------------|
| xi- | xi+ | Si | gi réel | gi réel % | Gi+ | Gi- | fi | fi[(Gi- + Gi+)/2] |
| 20 | 14431 | 1150915 | 0,1973 | 19,73% | 0,197 | 0 | 0,557 | 0,055 |
| 14431 | 28841 | 1311337 | 0,2248 | 22,48% | 0,422 | 0,197 | 0,223 | 0,069 |
| 28841 | 43252 | 1098227 | 0,1883 | 18,83% | 0,610 | 0,422 | 0,107 | 0,055 |
| 43252 | 57663 | 593631 | 0,1018 | 10,18% | 0,712 | 0,610 | 0,040 | 0,026 |
| 57663 | 72073 | 660848 | 0,1133 | 11,33% | 0,826 | 0,712 | 0,033 | 0,026 |
| 72073 | 86484 | 388578 | 0,0666 | 6,66% | 0,892 | 0,826 | 0,020 | 0,017 |
| 86484 | 100894 | 190281 | 0,0326 | 3,26% | 0,925 | 0,892 | 0,007 | 0,006 |
| 100894 | 115305 | 438325 | 0,0752 | 7,52% | 1,000 | 0,925 | 0,013 | 0,013 |
| | | 5832142 | 1 | 100,00% | | 1,000 | 1,000 | U = 0,267 |

Explication :

Les classes étant connues, ainsi que la *MASSE TOTALE A REPARTIR* MS = 5832142 enfants dans les 300 régions, il est possible dans le tableau de l'énoncé de dénombrer le nombre d'enfants au sein de chacune des classes = Si.

On en déduit alors la part REELLE (et non estimée) de chaque classe dans la masse, c'est-à-dire : $gi = Si/MS$.

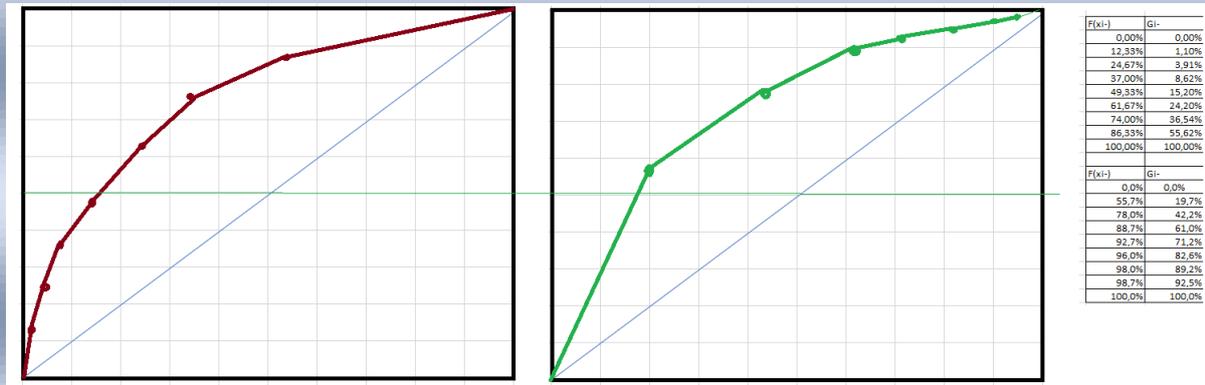
L'Indice de GINI est : $IG = 2(0,5 - U)$ avec $U = \sum(fi \cdot (Gi- + Gi+)/2)$ (voir tableau)

On obtient ainsi :

$$IG1 = 2(0,5 - 0,2511) = 0,498 \text{ soit } 50\%$$

$$IG2 = 2(0,5 - 0,267) = 0,466 \text{ soit un peu moins de } 50\%$$

b) La courbe de LORENZ-GINI



c) Les MEDIALES (xML) et LE COEFFICIENT DE VARIATION (iV)

On extrait les valeurs encadrantes nécessaires à l'interpolation de xML dans chaque tableau

Tableau de la courbe 1

| | | | |
|-------|--------|--------|--------|
| 23529 | 36275 | 74,00% | 36,54% |
| 36275 | 115305 | 86,33% | 55,62% |

Les valeurs encadrantes sont :

$$x_{i-} = 23529 \quad G(x_{i-}) = 36,54$$

$$x_{i+} = 36275 \quad G(x_{i+}) = 55,62$$

$$x_{ML1} = 23529 + (12746 \cdot (50 - 36,54)) / (55,62 - 36,54) = 32520,7 \text{ pour } x_{M\acute{e}} = 15918$$

on en d\eduit iV, le coefficient de variation

$$(x_{ML} - x_{M\acute{e}}) / e = (32521 - 15918) / 115285 = 0,1440 \text{ soit } 14,4\%$$

Tableau de la courbe 2

| | | | |
|-------|-------|-------|-------|
| 28841 | 43252 | 78,0% | 42,2% |
| 43252 | 57663 | 88,7% | 61,0% |

$$x_{i-} = 28841 \quad G(x_{i-}) = 42,2$$

$$x_{i+} = 43252 \quad G(x_{i+}) = 61$$

$$x_{ML2} = 28841 + (14411 \cdot (50 - 42,2)) / (61 - 42,2) = 34820 \text{ pour } x_{M\acute{e}} = 12956$$

$$(x_{ML} - x_{M\acute{e}}) / e = (34820 - 12956) / 115285 = 0,19 \text{ soit } 19\%$$

Leçon DE L'ANALYSE DE LA CONCENTRATION

On peut dresser les constats suivants :

Le premier constat est celui d'une concentration \elev\ee (IG \cong 50%) dans les deux cas. Elle est aussi traduite, par la forte proximit\ee des deux M\ediales), c'est-\a-dire de la modalit\ee xi qui partage la masse en deux parts \eegales, et celle des deux iV. De m\eme en est il des formes similaires des deux courbes de Gini.

Le second constat est que toutefois la concentration n'est pas de m\eme nature. Ceci est la principale diff\erence. Ce dont il faut rendre compte.

La courbe de Gini permet de lire cette diff\erence, que l'on retrouve dans le tableau de distribution.

C'est au extrêmes de chaque courbe de Gini, que se lit la modification de la concentration.

Deux segments de droite peuvent être comparés :

Le segment supérieur de la courbe 1 : il fait place dans la courbe 2 à une prolifération de points, lesquels sont maintenant situés sous le segment ;

Le segment inférieur de la courbe 2 il était antérieurement composé d'une prolifération de points. Le segment est maintenant situé sous les points.

En résumé, la modification dans la courbe 2 peut être lue ainsi :

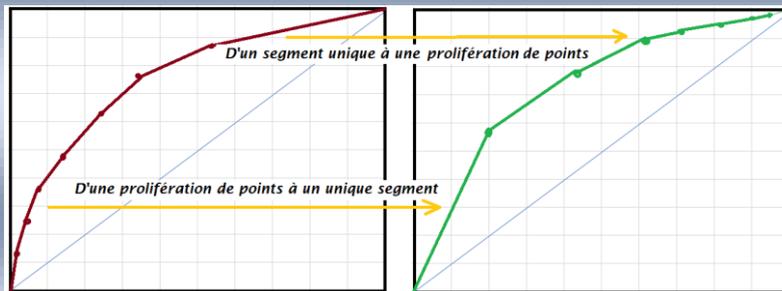
A l'extrémité supérieur : un segment est devenu une prolifération de points, en leur étant supérieur.

Le phénomène nouveau est ici la prolifération des points.

A l'extrémité inférieur : une prolifération de points, est devenue un segment, qui leur est inférieur.

Le phénomène nouveau est ici le segment.

Ce que l'on peut lire dans les deux courbes de la manière suivante :



Le troisième constat est déduit du précédent. On peut dire que la même concentration ($IG \cong 50\%$) a basculé. La méthode de l'amplitude moyenne a contribué à *redistribuer la masse*, des classes les plus denses vers les classes les moins denses.

Ce phénomène est transparent dans le tableau des $F(x_i-)$ et $G_i(x_i-)$, où on observe la différence aux extrémités :

| COURBE 1 | | 86,33% de la population se partage 55,62% de la masse. Donc $(100\% - 86,33\% = 13,67\%)$ de la population bénéficie de $(100\% - 55,62\% = 44,38\%)$ de la masse | COURBE 2 | | Au lieu que 12,33% de la population se partage 1,1% de la masse, c'est maintenant 55,7% qui bénéficie de 19,7% de la masse |
|-----------|---------|--|-----------|--------|--|
| $F(x_i-)$ | G_i- | | $F(x_i-)$ | G_i- | |
| 0,00% | 0,00% | | 0,0% | 0,0% | |
| 12,33% | 1,10% | | 55,7% | 19,7% | |
| 24,67% | 3,91% | | 78,0% | 42,2% | |
| 37,00% | 8,62% | | 88,7% | 61,0% | |
| 49,33% | 15,20% | | 92,7% | 71,2% | |
| 61,67% | 24,20% | | 96,0% | 82,6% | |
| 74,00% | 36,54% | | 98,0% | 89,2% | |
| 86,33% | 55,62% | | 98,7% | 92,5% | |
| 100,00% | 100,00% | | 100,0% | 100,0% | |

3) Conclusion : résumer simplement la conséquence majeure du passage de la méthode classique à celle de l'amplitude moyenne uniforme.

En augmentant la part des classes à faible densité, et en diminuant celle des classes à forte densité, la méthode de l'amplitude moyenne uniforme, réalise un nivellement des régions. Phénomène général et toujours associé au calcul mathématique, et à l'adoption, d'une moyenne.

Ainsi, chez Malthus et Ricardo, les Lois sur les pauvres sont elles dites accroître les revenus des pauvres, en diminuant ceux des riches, comme si un revenu (+ou-) moyen était édicté.



FIN

TABLEAU DES 300 REGIONS EUROPEENNES SELON LE NOMBRE D'ENFANTS DE 3 ANS SCOLARISES

| REGIONS 2020 | Variable xi | RANG | | | | | | |
|--------------------------------------|----------------|------|-------|-----|------------------------------------|--|--------|-----|
| Jadranska Hrvatska | | | 7957 | 101 | Pais Vasco | | 18419 | 201 |
| Severozápad | | | 8009 | 102 | Norra Sverige | | 18517 | 202 |
| Sjælland | | | 8079 | 103 | Hamburg | | 18633 | 203 |
| Etelä-Suomi | | | 8308 | 104 | Galicía | | 19308 | 204 |
| Prov. Limburg (BE) | | | 8773 | 105 | Ostra Mellansverige | | 19568 | 205 |
| Region Sumadije i Zapadne Srbije | | | 8804 | 106 | Latvija | | 19615 | 206 |
| Extremadura | | | 8818 | 107 | Centru | | 19697 | 207 |
| Közép-Dunántúl | | | 8922 | 108 | Ege | | 19708 | 208 |
| Norra Mellansverige | | | 9007 | 109 | Pomorskie | | 19833 | 209 |
| Podlaskie | | | 9089 | 110 | Hovedstaden | | 20297 | 210 |
| Zahodna Slovenija | | | 9157 | 111 | Alsace | | 20375 | 211 |
| Vzhodna Slovenija | | | 9266 | 112 | Prov. Antwerpen | | 20435 | 212 |
| Sostines regionas | | | 9292 | 113 | Área Metropolitana de Lisboa | | 20490 | 213 |
| Liguria | | | 9299 | 114 | Haute-Normandie | | 21384 | 214 |
| Yuzhen tsentralen | | | 9392 | 115 | Picardie | | 21970 | 215 |
| Agder og Rogaland (statistical regi | | | 9393 | 116 | Brandenburg | | 22060 | 216 |
| Moravskoslezsko | | | 9420 | 117 | Nord-Vest | | 22124 | 217 |
| Warminko-Mazurskie | | | 9476 | 118 | Dolnoslaskie | | 22600 | 218 |
| Izmir | | | 9587 | 119 | Schleswig-Holstein | | 23126 | 219 |
| Abruzzo | | | 9610 | 120 | Severna i yugoiztochna Bulgaria | | 23203 | 220 |
| Dél-Alföld | | | 9694 | 121 | Västsvrige | | 23369 | 221 |
| Småland med öarna | | | 9756 | 122 | Lorraine | | 23529 | 222 |
| Stredné Slovensko | | | 9827 | 123 | Dunántúl | | 23593 | 223 |
| Észak-Magyarország | | | 9893 | 124 | Srbija - sever | | 24080 | 224 |
| Sardegna | | | 9976 | 125 | Nord-Est | | 24239 | 225 |
| Voreia Ellada | | | 10041 | 126 | Közép-Magyarország | | 24428 | 226 |
| Südosterrreich | | | 10064 | 127 | Yugozapadna i yuzhna tsentralna B | | 24720 | 227 |
| Sar-østlandet (statistical region 20 | | | 10102 | 128 | Westösterreich | | 24802 | 228 |
| Jihozápad | | | 10118 | 129 | Toscana | | 25156 | 229 |
| Vestlandet (statistical region 2016 | | | 10149 | 130 | Norte | | 25373 | 230 |
| Pohjois- ja itä-Suomi | | | 10253 | 131 | Macroregionea patru | | 25689 | 231 |
| Strední Morava | | | 10512 | 132 | Makroregion Centralny | | 25694 | 232 |
| Praha | | | 10527 | 133 | Centre - Val de Loire | | 27417 | 233 |
| Illes Balears | | | 10665 | 134 | Bourgogne - Franche-Comté | | 27973 | 234 |
| Northern and Western | | | 10698 | 135 | Stockholm | | 28004 | 235 |
| Marche | | | 10761 | 136 | Malopolskie | | 29239 | 236 |
| Region Vojvodine | | | 10840 | 137 | Languedoc-Roussillon | | 29646 | 237 |
| Východné Slovensko | | | 10937 | 138 | Makroregion Poludniowo-Zachodni | | 29753 | 238 |
| Güneydogu Anadolu | | | 10939 | 139 | Puglia | | 29779 | 239 |
| Prov. Liège | | | 11072 | 140 | Piemonte | | 29793 | 240 |
| Aragón | | | 11259 | 141 | Noroeste (ES) | | 30071 | 241 |
| Attiki | | | 11423 | 142 | Wielkopolskie | | 30545 | 242 |
| Oberösterreich | | | 11463 | 143 | Istanbul | | 30591 | 243 |
| Länsi-Suomi | | | 11478 | 144 | Midi-Pyrénées | | 30947 | 244 |
| Prov. Vlaams-Brabant | | | 11499 | 145 | Emilia-Romagna | | 31504 | 245 |
| Prov. West-Vlaanderen | | | 11676 | 146 | RUP FR - Régions ultrapériphérique | | 31778 | 246 |
| Ankara | | | 11779 | 147 | Eastern and Midland | | 32053 | 247 |
| Pest | | | 11955 | 148 | Alföld és Észak | | 32852 | 248 |
| Strední Čechy | | | 12003 | 149 | Macroregionea trei | | 33014 | 249 |
| Syddanmark | | | 12059 | 150 | Bretagne | | 33805 | 250 |
| Vest | | | 12143 | 151 | Aquitaine | | 34218 | 251 |
| Zachodniopomorskie | | | 12244 | 152 | Slaskie | | 34470 | 252 |
| Franche-Comté | | | 12437 | 153 | Veneto | | 34670 | 253 |
| Budapest | | | 12473 | 154 | Warszawski stoleczny | | 34708 | 254 |
| Auvergne | | | 12662 | 155 | Ostösterreich | | 34987 | 255 |
| Severovýchod | | | 12836 | 156 | Berlin | | 35157 | 256 |
| Beogradski region | | | 13240 | 157 | Normandie | | 35592 | 257 |
| Észak-Alföld | | | 13265 | 158 | Sachsen | | 36018 | 258 |
| Mecklenburg-Vorpommern | | | 13335 | 159 | Rheinland-Pfalz | | 36275 | 259 |
| La Réunion | | | 13436 | 160 | Sicilia | | 37000 | 260 |
| Sud-Vest Oltenia | | | 13546 | 161 | Région wallonne | | 37416 | 261 |
| Západné Slovensko | | | 13757 | 162 | Noroeste (ES) | | 38606 | 262 |
| Champagne-Ardenne | | | 13864 | 163 | Makroregion Wschodni | | 39050 | 263 |
| Prov. Hainaut | | | 13969 | 164 | Lazio | | 40236 | 264 |
| Kujawsko-Pomorskie | | | 14090 | 165 | Macroregionea doi | | 40572 | 265 |
| Basse-Normandie | | | 14208 | 166 | Macroregionea unu | | 41821 | 266 |
| Midtjylland | | | 14339 | 167 | Pays-de-la-Loire | | 42061 | 267 |
| Bati Anadolu | | | 14358 | 168 | Comunitat Valenciana | | 43115 | 268 |
| Srbija - jug | | | 14444 | 169 | Makroregion Północny | | 43399 | 269 |
| Jihovýchod | | | 14461 | 170 | Centro (ES) | | 44047 | 270 |
| Lubelskie | | | 14582 | 171 | Manner-Suomi | | 45446 | 271 |
| Dogu Marmara | | | 14734 | 172 | Isole | | 46976 | 272 |
| Niederösterreich | | | 15169 | 173 | Östra Sverige | | 47572 | 273 |
| Region Juzne i Istocne Srbije | | | 15197 | 174 | Makroregion Północno-Zachodni | | 50396 | 274 |
| Oslo og Akershus (statistical region | | | 15197 | 174 | Södra Sverige | | 50986 | 275 |
| Calabria | | | 15238 | 175 | Makroregion Województwo Mazow | | 51567 | 276 |
| Yugozapaden | | | 15328 | 176 | Hessen | | 55366 | 277 |
| Podkarpackie | | | 15379 | 177 | Nouvelle-Aquitaine | | 57342 | 278 |
| Helsinki-Uusimaa | | | 15407 | 178 | Provence-Alpes-Côte d'Azur | | 57419 | 279 |
| Bucuresti - Ilfov | | | 15424 | 179 | Grand Est | | 57768 | 280 |
| Bourgogne | | | 15536 | 180 | Occitanie | | 60593 | 281 |
| Centro (PT) | | | 15597 | 181 | Comunidad de Madrid | | 62934 | 282 |
| Prov. Oost-Vlaanderen | | | 15804 | 182 | Makroregion Poludniowy | | 63709 | 283 |
| Región de Murcia | | | 15851 | 183 | Niedersachsen | | 68140 | 284 |
| Canarias | | | 16078 | 184 | Vlaams Gewest | | 68187 | 285 |
| Kontinentalna Hrvatska (NUTS 201 | | | 16155 | 185 | Cataluña | | 69212 | 286 |
| Région de Bruxelles-Capitale/Bruss | | | 16323 | 186 | Continente | | 69712 | 287 |
| Sud-Est | | | 16333 | 187 | Austria | | 69853 | 288 |
| Mazowiecki regionalny | | | 16859 | 188 | Hauts-de-France | | 70740 | 289 |
| Castilla y León | | | 16878 | 189 | Portugal | | 73313 | 290 |
| Poitou-Charentes | | | 16951 | 190 | Lombardia | | 73431 | 291 |
| Vidurio ir vakaru Lietuvos regionas | | | 17194 | 191 | Andalucía | | 79310 | 292 |
| Sachsen-Anhalt | | | 17428 | 192 | Rhône-Alpes | | 80257 | 293 |
| Wien | | | 17449 | 193 | Centro (IT) | | 82267 | 294 |
| Sud - Muntenia | | | 17590 | 194 | Auvergne - Rhône-Alpes | | 92919 | 295 |
| Sydsverige | | | 17862 | 195 | Sur (ES) | | 97362 | 296 |
| Lódzkie | | | 17946 | 196 | Baden-Württemberg | | 101462 | 297 |
| Thüringen | | | 18085 | 197 | Sud | | 108123 | 298 |
| Akdeniz | | | 18108 | 198 | Nord-Ovest | | 113435 | 299 |
| Southern | | | 18168 | 199 | Bayern | | 115305 | 300 |
| Castilla-la Mancha | | | 18351 | 200 | | | | |
| Samsun, Tokat, Çorum, Amasya | | | 4133 | 51 | | | | |
| Dogu Karadeniz | | | 4197 | 52 | | | | |
| Trabzon, Ordu, Giresun, Rize, Artvir | | | 4197 | 53 | | | | |
| Orta Anadolu | | | 4330 | 54 | | | | |
| Cantabria | | | 4366 | 55 | | | | |
| Ortadogu Anadolu | | | 4379 | 56 | | | | |
| Malta | | | 4391 | 57 | | | | |
| Nisia Aigaioi, Kriti | | | 4414 | 58 | | | | |
| Provincia Autonoma di Trento | | | 4485 | 59 | | | | |
| Severozapaden | | | 4671 | 60 | | | | |
| Guyane | | | 4794 | 61 | | | | |
| Nord-Norge | | | 4819 | 62 | | | | |
| Provincia Autonoma di Bolzano/Bo | | | 4856 | 63 | | | | |
| Mayotte | | | 4878 | 64 | | | | |
| Alentejo | | | 4903 | 65 | | | | |
| Severen tsentralen | | | 4936 | 66 | | | | |
| Guadeloupe | | | 5038 | 67 | | | | |
| Prov. Namur | | | 5076 | 68 | | | | |
| Trøndelag | | | 5152 | 69 | | | | |
| Bremen | | | 5331 | 70 | | | | |
| Kentriki Makedonia | | | 5491 | 71 | | | | |
| Övre Norrland | | | 5502 | 72 | | | | |
| Region Juzne i Istocne Srbije | | | 5640 | 73 | | | | |
| Sanliurfa, Diyarbakir | | | 5665 | 74 | | | | |
| Nordjylland | | | 5686 | 75 | | | | |
| Tirol | | | 5844 | 76 | | | | |
| Umbria | | | 6114 | 77 | | | | |
| Limousin | | | 6173 | 78 | | | | |
| Aydin, Denizli, Mugla | | | 6178 | 79 | | | | |
| Bratislavský kraj | | | 6179 | 80 | | | | |
| Comunidad Foral de Navarra | | | 6198 | 81 | | | | |
| Principado de Asturias | | | 6397 | 82 | | | | |
| Bati Karadeniz | | | 6630 | 83 | | | | |
| Severoiztochen | | | 6661 | 84 | | | | |
| Adana, Mersin | | | 6744 | 85 | | | | |
| Steiermark | | | 6746 | 86 | | | | |
| Yugoiztochen | | | 6935 | 87 | | | | |
| Dél-Dunántúl | | | 6948 | 88 | | | | |
| Opolskie | | | 7153 | 89 | | | | |
| Cyprus | | | 7243 | 90 | | | | |
| Bati Marmara | | | 7309 | 91 | | | | |
| Kocaeli, Sakarya, Düzce, Bolu, Yalo | | | 7348 | 92 | | | | |
| Bursa, Eskisehir, Bilecik | | | 7386 | 93 | | | | |
| Saarland | | | 7434 | 94 | | | | |
| Kentriki Ellada | | | 7494 | 95 | | | | |
| Lubuskie | | | 7607 | 96 | | | | |
| Antalya, Isparta, Burdur | | | 7720 | 97 | | | | |
| Nyugat-Dunántúl | | | 7723 | 98 | | | | |
| Swietokrzyskie | | | 7748 | 99 | | | | |
| Friuli-Venezia Giulia | | | 7953 | 100 | | | | |

FIN