

Répartition des richesses, inégalités

I - A l'aide de statistiques à une variable :

Nous allons rappeler tout le vocabulaire à partir d'un exemple basique : les notes obtenues par des élèves à un devoir de math :

7 ; 15 ; 10 ; 13 ; 17,5 ; 16,5 ; 14,5 ; 7 ; 7 ; 20 ; 20 ; 14 ; 20 ; 8 ; 10,5 ; 12,5 ; 18 ; 8,5 ; 6 ; 10 ; 2 ; 13,5 ; 8,5 ; 14,5

1) La population est l'ensemble des individus ou des unités statistiques étudiées (ici les élèves)
Un échantillon est une partie de cette population.

2) Le caractère étudié peut être qualitatif (classe d'origine, loisir) ou quantitatif (des nombres). Dans ce cas il est discret (age, note) ou continu (taille). Il est alors réparti par classes.
Ici, on pourrait faire des classes pour faire des classes d'élèves par niveau (TB, B, Moyen, Insuffisant)

3) L'effectif est le nombre d'individus d'un échantillon et la fréquence est la proportion d'individus dans cet échantillon.

Ici :

4) On représente les résultats sous forme de graphiques différents suivant si le caractère est qualitatif (bandes, circulaire), discret (points, bâtons) ou continu (histogrammes, polygone des effectifs).

5) Le mode est la valeur ou la classe pour laquelle l'effectif est maximum. Ici :

6) La médiane est la valeur qui sépare la population en deux parties de même effectif.
On peut définir de même les quartiles ou les déciles. On utilise la calculatrice :

7) La moyenne est donnée par la formule : $\bar{x} = \frac{n_1 x_1 + n_2 x_2 + \dots + n_p x_p}{N} = \frac{1}{N} \sum_{i=1}^p n_i x_i$

Ici :

8) L'étendue est la différence entre les valeurs extrêmes. L'intervalle interquartile est la différence entre Q_3 et Q_1 . Il a l'avantage d'être peu sensible aux valeurs extrêmes de la série.

Ici :

9) L'écart moyen est donné par la formule $e_m = \frac{1}{N} \sum_{i=1}^p n_i |x_i - \bar{x}|$ (très peu utilisé à cause des valeurs absolues)

10) La variance est $V = \frac{n_1(x_1 - \bar{x})^2 + n_2(x_2 - \bar{x})^2 + \dots + n_p(x_p - \bar{x})^2}{N} = \frac{1}{N} \sum_{i=1}^p n_i(x_i - \bar{x})^2$

11) L'écart type est $\sigma = \sqrt{V}$. C'est une mesure de dispersion autour de la moyenne. Il a l'avantage d'être calculable et d'utiliser toutes les valeurs de la série.

Ici :

12) Le rapport interdécile D_9/D_1 est le rapport entre le 9ème et le 1^{er} décile. Il met en évidence le rapport entre les premières et les dernières valeurs.

Ici :

II – A l'aide de l'analyse de fonctions :

1) Convexité, point d'inflexion :

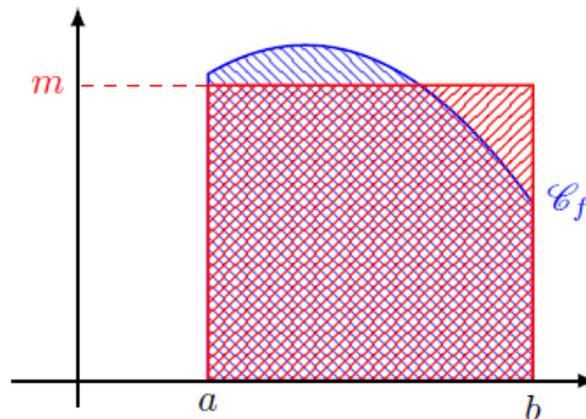
Permet de repérer quand une grandeur change d'inflexion, c'est à dire quand elle croît plus ou moins vite.

2) Valeur moyenne :

On considère une fonction f définie et continue sur un intervalle $[a ; b]$ (avec $a < b$).

On appelle **valeur moyenne de f entre a et b** le nombre m tel que :

$$m = \frac{1}{b-a} \int_a^b f(x) dx$$



Si f est positive sur l'intervalle $[a ; b]$

et si m est la valeur moyenne de la fonction f entre a et b , alors $m(b-a) = \int_a^b f(x) dx$
et l'aire du rectangle de largeur $b-a$ et de longueur m est égale

à l'aire de la partie du plan située sous la courbe C_f et entre les droites d'équations $x=a$ et $x=b$.

Autrement dit, Dans le cas où f est strictement positive sur $[a ; b]$, la valeur moyenne de f correspond à la hauteur du rectangle de largeur $b-a$ ayant la même aire que l'aire sous la courbe.

3) Courbe de Lorenz et indice de Gini :

La **courbe de Lorenz** est la représentation graphique de la fonction qui, à la part x des détenteurs d'une part d'une grandeur, associe la part y de la grandeur détenue.

Par exemple, ci contre, on voit la répartition de richesses dans trois pays différents :

Au Panama, les 60 % les plus pauvres possèdent 20 % des richesses, alors qu'aux Pays Bas, c'est 40 % des richesses.

L'indice de Gini (G) est égal au double de l'aire entre la droite de répartition parfaite et la courbe de Lorenz. C'est un nombre entre 0 et 1. Plus il est proche de 1, plus les inégalités sont importantes.

L'indice de Gini du Panama semble le double de celui des Pays Bas.

