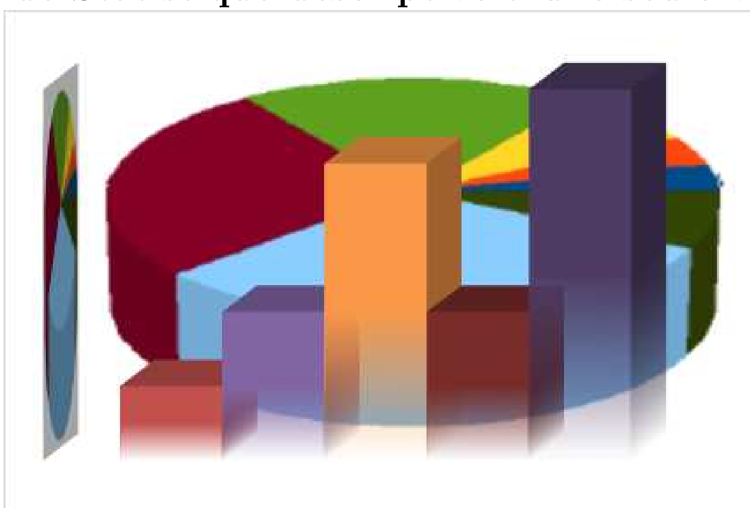
UNIVERSITE DE BLIDA 1

FACULTÉ DES SCIENCES

DEPARTEMENT DU TRONC COMMUN MATHEMATIQUES ET INFORMATIQUE(MI)

Cours de Statistique descriptive à une seule variable



Réalisé par : Mme Messaoudi Nadia

Chargée de cours : **1 er année MI Section A**

Courriel : messlina2012@gmail.com

Année universitaire 2019- 2020

Chapitre 2

Représentation graphique d'une variable statistique

Nous présentons dans ce chapitre quelques méthodes graphiques d'une variable statistique : qualitative, quantitative discrète et continue.
Le graphique est un outil de communication visuelle et rapide qui permet de :
Visualiser d'un seul coup d'oeil les principales caractéristiques.
Mettre en évidence les tendances (soit à droite ou à gauche)
Voir s'il y 'a répétition du phénomène.

2.1 Représentation graphique : caractère qualitatif

Il existe plusieurs tyes de représentations graphiques pour un caractère qualitatif. Les plus courants sont les suivants : le diagramme en tuyaux d'orgue et le diagramme à secteurs circulaires.

2.1.1 Diagramme en tuyaux d'orgue

C'est un graphique qui à chaque modalités x_i de la variable qualitative x associe un rectangle de base constante dont la hauteur est proportionnelle à l'effectif n_i (ou la fréquence f_i). Ces rectangles sont séparés les uns des autres par des distances égales.(voir figure 2.1)

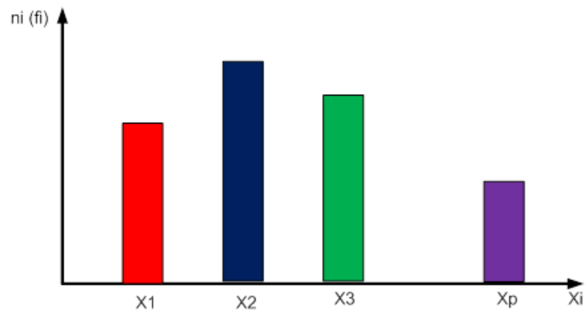


FIGURE 2.1 – Diagramme en tuyaux d'orgue

D'après l'exemple 1 des couleurs des voitures dans un garage parking, le diagramme correspondant est le suivant :

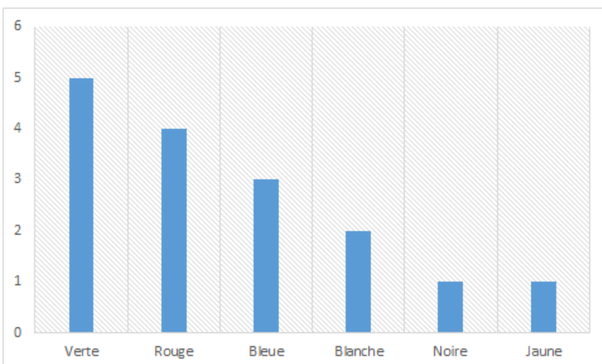


FIGURE 2.2 – Diagramme en tuyaux d'orgue des Couleurs des véhicules dans un parking

2.1.2 Diagramme à secteurs circulaires

Le diagramme à secteurs est un graphique qui divise un disque (ou demi secteur) en secteurs angulaires dont les angles au centre sont proportionnelles aux effectifs (ou fréquences) de chaque modalité.
Pour une modalité i, d'effectif n_i , l'angle au centre α_i correspondant est donné (en degré) à l'aide de la règle de trois de la manière suivante :

$$N \longrightarrow 360°$$
$$n_i \longrightarrow \alpha_i^\circ$$

D'où :

$$\alpha_i^\circ = \frac{n_i}{N} \times 360° = f_i \times 360°$$

On peut adopter pour la représentation graphique le diagramme semi-circulaire, l'angle au centre est égal à :

$$\alpha_i^\circ = \frac{n_i}{N} \times 180^\circ = f_i \times 180^\circ$$

Exemple : D'après l'exemple 1, on peut calculer l'angle correspondant à chacune des couleurs, d'où : $\alpha_{verte} = \frac{5}{16} \times 360 = 112.5^\circ$, $\alpha_{rouge} = 90^\circ$, $\alpha_{bleue} = 67.5^\circ$, $\alpha_{blanche} = 45^\circ$, $\alpha_{noire} = 22.5^\circ$, $\alpha_{jaune} = 22.5^\circ$. Le diagramme correspondant est le suivant :

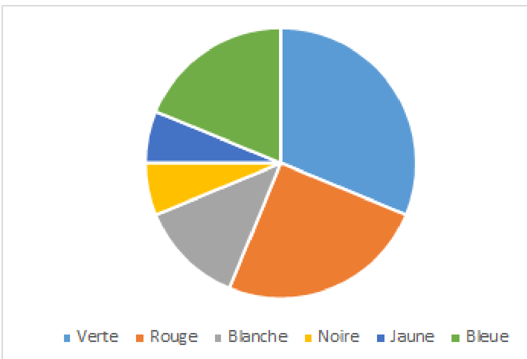


FIGURE 2.3 – Diagramme circulaire des couleurs des véhicules dans un parking

2.2 Représentation : caractère quantitatif discret

2.2.1 Diagramme en bâtons

On appelle diagramme en bâtons, un graphique qui associe à chaque valeur x_i de la variable discrète x sur l'axe des abscisses, un segment (bâton) vertical dont la hauteur est proportionnelle à l'effectif (ou la fréquence) sur l'axe des ordonnées. On suppose les valeurs observées de la variable quantitative discrète sont ordonnées par un ordre croissant (ou décroissant)(voir figure 2.4).

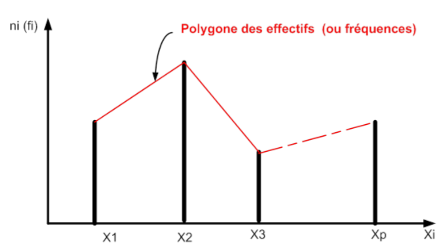


FIGURE 2.4 – Diagramme en bâtons

Définition : Polygone des effectifs (ou des fréquences) d'une variable discrète est la courbe qui joint les sommets des batons du diagramme en bâtons.

Exemple :
On donne la distribution statistique du nombre de pièces par logement (après dénombrement des résultats de l'enquête)

Nombre de pièces	1	2	3	4	5	6	Total
Nombre de logements n_i	5	10	20	30	25	10	100
Fréquence f_i	0.05	0.1	0.2	0.3	0.25	0.1	1
$F_i cum \nearrow$	0.05	0.15	0.35	0.65	0.9	1	
$F_i cum \searrow$	1	0.95	0.85	0.65	0.35	0.1	

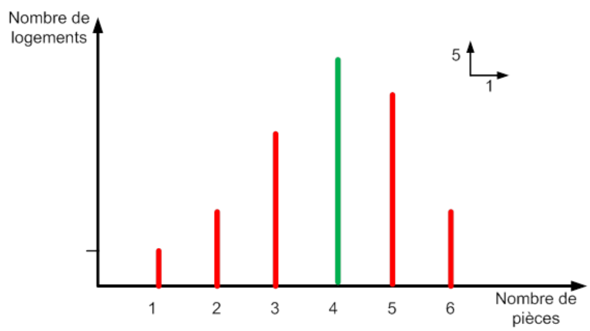


FIGURE 2.5 – Diagramme en bâtons de nombre de pièces par logements

Remarque : On peut représenter le diagramme en bâtons des fréquences.

2.2.2 Diagramme cumulatif de la variable discrète

On appelle diagramme cumulatif, la courbe cumulative représentative de la fonction de répartition d'un caractère x l'application $F : \Re \to [0, 1]$, $F(x)$ est la proportion des individus dont le caractère est inférieur ou égale à x ($x \in \Re$). Cette fonction est constante dans l'intervalle séparant deux valeurs possibles consécutives, qui vaut :

$$F(x) = \begin{cases} 0 & , x < x_1 \\ F_1 & , x_1 \leq x < x_2 \\ F_j & , x_j \leq x < x_{j+1} \\ 1 & , x \geq x_p \end{cases}$$

où :

$$F_j = f_j cum \nearrow = \sum_{k=1}^{j} f_k, j = 1, ..., p$$

La représentation graphique de la fonction de répartition est la courbe cumulative en escalier, donnée par la figure 2.6 (a).

(a) courbe cumulative (b) Courbe en escalier inverse

FIGURE 2.6 – Courbe cumulative en escalier

• On peut tracer de manière analogue la courbe cumulative des fréquences cumulatives décroissantes, ce qu'on appellera courbe cumulative inversée (voir figure 2.6 (b)).

Exemple : Consédirons l'exemple 3 du nombre de pièces par logement donné précédemment , les courbes en escalier correspondantes sont :

(a) Courbe cumulative des *Ficum* ↗ (b) Courbe cumulative des *Ficum* ↘

FIGURE 2.7 – Courbes cumulatives du nombre de pièces par logement

2.3 Représentation graphique : caractère quantitatif continu

2.3.1 Tableau statistique d'une variable quantitative continue

Une variable continue est souvent regroupée en classe, son tableau statistique est le suivant :

Classes	$[e_1, e_2[$	$e_2, e_3[$	....	$[e_k, e_{k+1}[$
Centre x_i	x_1	x_2	...	x_k
n_i	n_1	n_2	...	n_k
f_i	f_1	f_2	...	f_k
$Ficum \nearrow$	F_1	F_2	...	F_k

Une classe est un intervalle fermé à gauche et ouvert à droite du type $[e_i,e_{i+1}[$, où :
e_i : la borne inférieure de la classe numéro i.
e_{i+1} : la borne supérieure de la classe numéro i.
L'amplitude d'une classe i est $a_i = e_{i+1} - e_i$.
Le centre de la classe i est x_i : $x_i = \frac{e_i+e_{i+1}}{2}$.
n_i : effectif de la classe i .
f_i : fréquence de la classe i.

2.3.2 Histogramme

Une variable continue repartie en classes. Sa représentation se fait à l'aide d'un histogramme. On construit des rectangles juxtaposés dont la base (sur l'axe des abscisses) est l'amplitude de la classe et la hauteur est proportionnelle aux effectifs (ou fréquences). La surface du rectangle doit être égale (ou proportionnelle) à l'effectif (ou la fréquence) de cette classe.
- Deux cas peuvent se présenter lors de la construction de l'histogramme :
a/ les classes ont toutes la même amplitude.
b/ les classes ont des amplitudes inégales.
• Cas a/ Amplitudes toutes égales :
Si les amplitudes de classes sont égales, la hauteur des rectangles correspond aux effectifs (ou aux fréquences) des classes.

Remarquons que les classes ont les mêmes amplitudes ($a_1 = a_2 = ... = a_k = a$) l'aire de chaque rectangle est proportionnel à l'effectif (ou la fréquence) de la classe correspondante. L'aire de l'histogramme est égale :

$$n_1a_1 + n_2a_2 + ... + n_ka_k = \sum_{i=1}^{k} n_i = a \times N$$

(ou bien $a \times 1$).

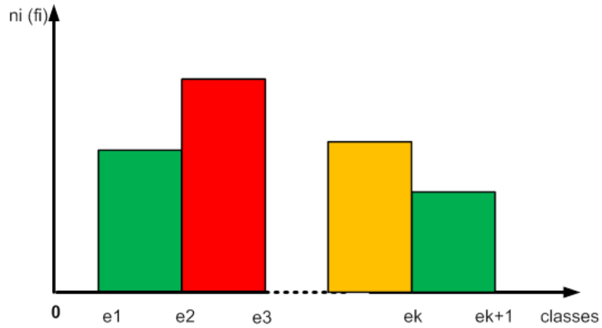


FIGURE 2.8 – Histogramme (classes d'amplitudes égales)

Exemple : Soit la distribution statistique correspondante à l'ancienneté du personnel cadre dans une entreprise de 50 employés :

Durée en années	[0,4[	[4,8[	[8,12[	[12,16[	[16,20[
Effectif	8	22	8	7	5
f_i	0.16	0.44	0.16	0.14	0.1
$f_i cum \nearrow$	0.16	0.6	0.76	0.9	1

Les classes ont toutes des amplitudes égales :

$$a_1 = a_2 = ... = a_5 = a = 4$$

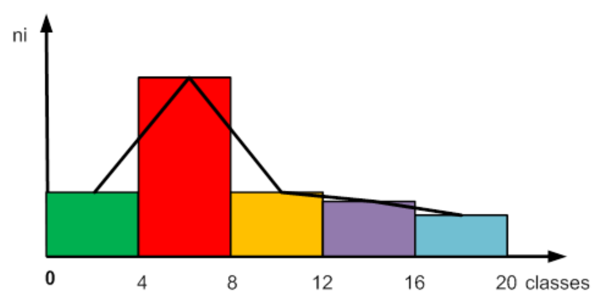


FIGURE 2.9 – Histogramme de l'ancienneté du personnel cadre dans une entreprise

Polygone des effectifs (ou des fréquences)

Le polygone des effectifs (ou des fréquences) est obtenu en joignant les milieux des sommets des rectangles des classes d'amplitudes égales de l'histogramme (voir figure 2.9).

Remarque :

Le polygone des fréquences donne une vision plus réaliste de la distribution en éliminant les ruptures entre les classes. Il permet également de percevoir la dissymétrie de la distribution.

Cas b/ Classes d'amplitudes inégales : correction de l'histogramme

La correction de l'histogramme signifie que la représentation graphique des effectifs (ou fréquences) est ramené à la même échelle exprimée en unité d'amplitude notée "au" (sorte de même dénominateur).

L'unité d'amplitude au (ou a) est déterminée par le PGCD des amplitudes a_i. L'effectif (ou la fréquence) corrigé est égale à :

$$n_i^c = \frac{n_i}{a_i} \times au$$

(pour la fréquence corrigée : $f_i^c = \dfrac{f_i}{a_i} \times au$). L'histogramme correspondant est le suivant :

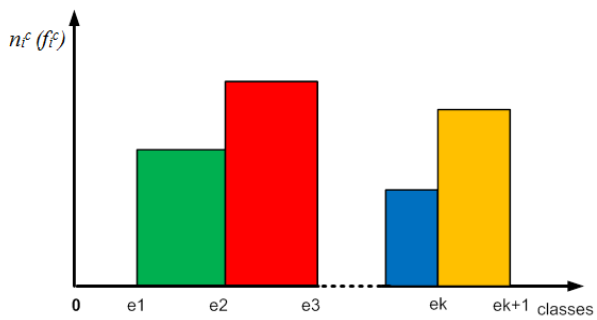


FIGURE 2.10 – Histogramme (classes d'amplitudes inégales)

L'aire de l'histogramme des effectifs est proportionnelle soit à l'effectif total N ou la fréquence totale 1 :

$$n_1^c a_1 + n_2^c a_2 + ... + n_k^c a_k = \frac{n_1}{a_1} a_1 \times a + \frac{n_2}{a_2} a_2 \times a + ... + \frac{n_k}{a_k} a_k \times a = a \times \sum_{i=1}^{k} n_i = a \times N$$

Exemple : Reprenons l'exemple précédent, en regroupant les effectifs des deux dernières classes de valeurs en une classe unique. Que deviendrait l'histogramme ?

Durée en années	[0,4[	[4,8[	[8,12[	[12,20[
Effectif n_i	8	22	8	12
a_i	4	4	4	8
$n_i^c = \dfrac{n_i \times au}{a_i}$	8	22	8	6

$$au = a = PGCD(a_1, a_2, a_3, a_4) = PGCD(4, 8) = 4$$

Le polygone des effectifs (ou des fréquences) est obtenu en joignant les milieux de l'unité d'amplitude des sommets des rectangles de l'histogramme.

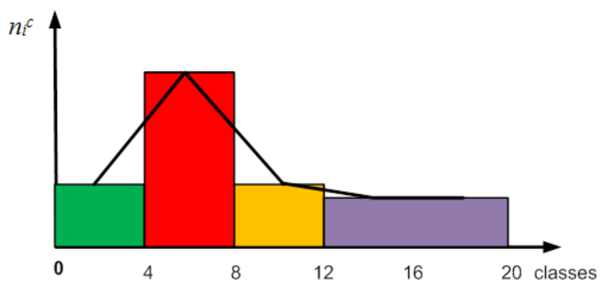


FIGURE 2.11 – Histogramme (classes d'amplitudes inégales)

Remarque On ajoute toujours deux classes fictives aux extrémités pour tracer le polygones des effectifs (ou fréquences). La surface sous le polygone des effectifs (ou fréquences) est égale à l'effectif total N (ou 1).

2.3.3 Courbe cumulative

La courbe cumulative est obtenue en joignant dans un système d'axes orthogonaux , les points d'abscisses (borne supérieure de la classe i et d'ordonnée $F_i cum \nearrow$ (fréquence cumulée croissante) correspondante par des segments de droite.
De la même manière, on peut tracer la courbe cumulative des fréquences cumulées décroissantes $F_i cum \searrow$ en portant les points dont les abscisses représentent la borne inférieure de chaque classe et les ordonnées la fréquences cumulées correspondantes puis en reliant ces points par des segments de droite.

- On peut aussi tracer les courbes des effectifs cumulés croissants ou décroissants.
La figure suivante représente les polygones des effectifs (ou fréquences) cumulés.

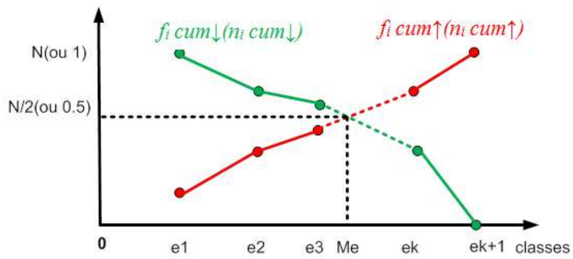


FIGURE 2.12 – Courbes des fréquences (ou effectifs) cumulées croissantes et décroissantes

Remarque : Le point d'intersection des deux courbes est un paramètre de position appellé *médiane* qui possède l'ordonnée $\frac{N}{2}$ (ou $\frac{1}{2}$). Ce paramètre est décrit en détaille dans le chapitre suivant.

Exemple : Reprenons l'exemple d'ancienneté du personnel cadre d'une entreprise.

Durée en années	[0,4[	[4,8[	[8,12[	[12,16[	[16,20[
Effectif	8	22	8	7	5
f_i	0.16	0.44	0.16	0.14	0.1
$f_i cum \nearrow$	0.16	0.6	0.76	0.9	1
$f_i cum \searrow$	1	0.84	0.4	0.24	0.1

Les courbes des frq́uences cumulées croissantes et décroissantes sont illustrées par la figure suivante :

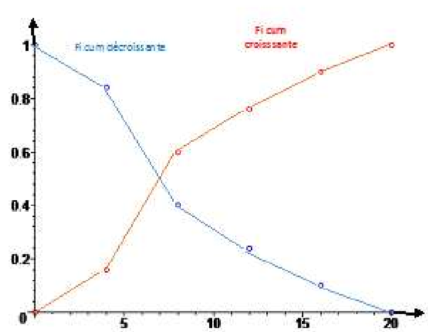


FIGURE 2.13 – Courbes des fréquences cumulées croissantes et décroissantes