

Statistiques



Dr ZAIZ Khaoula

Université d' El Oued

Faculté de sciences exactes.

Département de mathématiques

Email : khaoula.za@gmail.com

Table des matières



Objectifs	3
I - Chapitre 2 : Série statistiques à une variable	4
1. Étude d'une variable statistique discrète	4
2. Effectif, fréquence	5
2.1. Effectif	5
2.2. Fréquence	5
3. Exercice : Connaissance	6
4. Exercice : Application	6
5. Représentation graphique des séries statistiques	7
5.1. Généralités	7
6. Distribution à caractère qualitatif	8
6.1. Tuyaux d'orgues	8
6.2. Diagramme par secteur (diagramme circulaire)	8
6.3. Distribution à caractère quantitatif discret	9
7. Exercice : Analyse	10
8. Paramètres de position (caractéristique de tendance centrale)	11
8.1. Le mode	11
8.2. La médiane	11
8.3. La moyenne	12
9. Paramètres de dispersion (variabilité)	12
9.1. L'étendue	12
9.2. La variance	12
9.3. L'écart type	13
10. Série de TD	13
Bibliographie	14
Webographie	15

Objectifs

Le cours vise principalement à introduire et faire méditer les concepts fondamentaux et méthodes élémentaires de la statistique pour permettre un apprentissage autonome ultérieur de méthodes complémentaires.

- **Connaissance**
Connaître les définitions de base de statistique, aussi les différents types de variables statistiques et ses caractères.
- **Compréhension**
Préciser les types de variables statistiques, si elle est qualitative, quantitative, discrète ou continue.
- **Application**
Calculer les effectifs, fréquences, les paramètres de position et les paramètres de dispersion.
Représenter graphiquement de variables statistiques.
- **Analyse**
Classer les variables statistiques d'après ses caractères et ses graphes.
- **Jugement évaluation**
Analyser les propriétés d'une série statiques d'après ces caractéristique de tendance centrale et variabilité.
Évaluation finale de contenu étudiant.

Chapitre 2 : Série statistiques à une variable



Une série statistique est simplement une liste de mesures (qui provient généralement d'une étude). Chaque mesure s'appelle une valeur. Une valeur peut être un nombre, mais ce n'est pas obligatoirement le cas.

1. Étude d'une variable statistique discrète

◆ Rappel

Le **caractère statistique*** peut prendre un nombre fini raisonnable de valeurs (note, nombre d'enfants, nombre de pièces, ...). Dans ce cas, le caractère statistique étudié est alors appelé un caractère discret.


Dans toute la suite du chapitre, nous considérons la situation suivante :

$$X: \Omega \rightarrow \{x_1, \dots, x_n\}$$

avec $\text{Card}(\Omega) := N$ est le nombre d'individus dans notre étude.


2. Effectif, fréquence

2.1. Effectif

 *Définition : Effectif partiel*

L'**effectif partiel** : est le nombre de fois où cette valeur apparaît.

L'**effectif total** N : est le nombre total d'individus de la population étudiée.

 *Exemple*

Marie a lancé 20 fois un dé à six faces. Voici les nombres qui sont sortis :


2 - 6 - 5 - 2 - 4 - 1 - 6 - 3 - 1 - 4 - 2 - 1 - 6 - 5 - 6 - 3 - 2 - 1 - 2 - 4.

1. Ce relevé s'appelle une série statistique.
2. On réunit ces résultats dans un tableau d'effectifs en regroupant les valeurs de dés :

nombre x_i	1	2	3	4	5	6	total
effectifs n_i	4	5	2	3	2	4	20

Par exemple, l'**effectif** de la valeur « 4 » est de 3. Ceci veut dire que le nombre 4 est sorti 3 fois.

L'**effectif total** est de 20 lancers.

 *Définition : Effectif cumulé*

Pour chaque valeur x_i , l'**effectif cumulé** N_i d'une valeur est la somme de l'**effectif** de cette valeur et de tous les effectifs des valeurs qui précèdent.

Autrement dit

$$N_i = n_1 + n_2 + \dots + n_i$$


 *Exemple*

Une enquête réalisée dans un village porte sur le nombre d'enfants à charge par famille.

On note X le nombre d'enfants, les résultats sont données par ce tableau:

x_i	0	1	2	3	4	5	6
n_i (Effectif)	18	32	66	41	32	9	2
N_i (Effectif cumulé)	18	50	116	157	189	198	200

2.2. Fréquence

 *Définition*

On appelle **fréquence**, souvent notée f_i , d'une valeur du caractère le quotient de l'**effectif** de la valeur sur l'**effectif total**.

$$f_i = \frac{\text{effectif de la valeur}}{\text{effectif total}} = \frac{n_i}{N}$$

 **Exemple**

Voici les notes relevées lors d'une interrogation dans une classe.

Note	8	10	12	15	16	20
Effectif n_i	4	8	10	5	2	1

1. L'**effectif total** N est: $4+8+10+5+2+1=30$.
2. La **fréquence** de la note 8 est $\frac{4}{30} = \frac{2}{15}$


Donc, On obtient ainsi le tableau suivant:

Note	8	10	12	15	16	20
Effectif n_i	4	8	10	5	2	1
Fréquence f_i	$\frac{2}{15}$	$\frac{4}{15}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{15}$	$\frac{1}{30}$

 **Remarque**

Soit défini comme précédemment. Alors,

$$\sum_{i=1}^p f_i = 1.$$

 **Définition**

Pour chaque valeur x_i , on pose par définition

$$F_i = f_1 + f_2 + \dots + f_i$$

Le **fréquence cumulé** F_i d'une valeur est la somme de fréquence de cette valeur et de tous les **fréquences** des valeurs qui précèdent.

3. Exercice : Connaissance

l'effectif cumulé de x_n est

Le fréquence cumulé est la division de ; et

4. Exercice : Application

Nous rappelons le tableau statistique associe

x_i	0	1	2	3	4	5	6
n_i	18	32	66	41	32	9	2

Calculer le fréquence de nombre 4

5. Représentation graphique des séries statistiques

5.1. Généralités

◆ Rappel

1. Le graphique est un mode d'expression qui permet "visuellement" de saisir et de mémoriser un certain nombre d'informations. C'est pourquoi, lors de la présentation des résultats statistiques et complémentairement aux tableaux, on utilise souvent une représentation graphique.

Cette **représentation graphique*** peut répondre à deux types d'objectifs:

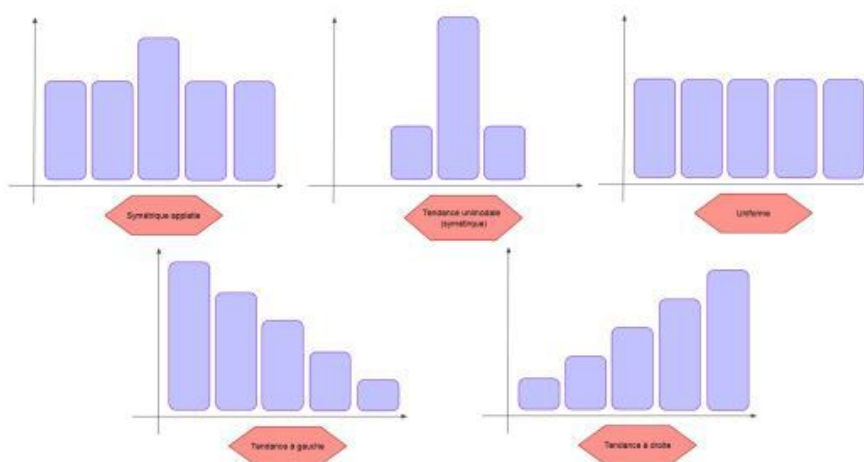
- Être un moyen de communication et permettre de véhiculer une information. C'est ainsi que certains graphiques figurent dans des articles de magazines et journaux, dans des brochures de présentation de résultats commerciaux ou (et) comptables et même dans certaines publicités . Ceci prouve qu'un bon graphique est souvent plus explicite qu'un long discours.
- Être un instrument de travail et permettre une vue d'ensemble synthétique du phénomène étudié, ce qui facilite l'analyse

1. Typologie des graphiques.

En matière statistique, tout graphique est caractérisé par :

- La nature de la série qu'il représente:
 - graphiques chronologiques.
 - graphiques spatiaux, quantitatifs ou qualitatifs
 - Le système de coordonnées qu'il utilise :
 - coordonnées cartésiennes.
 - coordonnées polaires.
 - coordonnées pluri linéaires.

La synthèse: visualiser d'un seul coup d'œil les principales caractéristiques (mais on perd une quantité d'informations).



Quelques caractéristiques du graphique

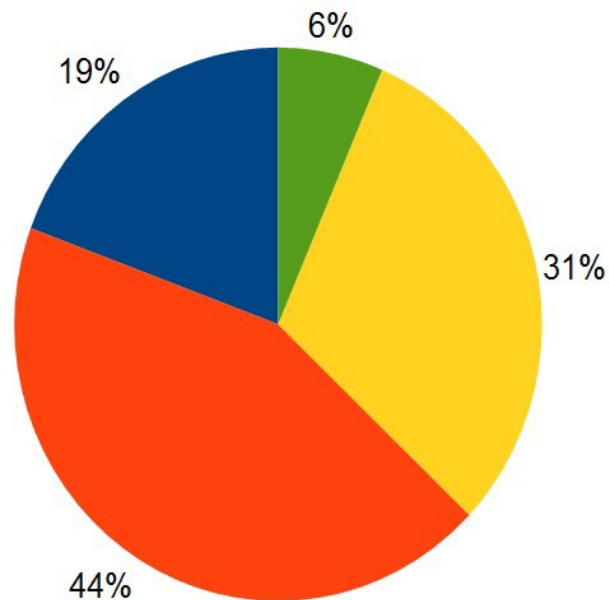


Diagramme par secteur

6.3. Distribution à caractère quantitatif discret

A partir de l'observation d'une variable quantitative discrète, deux diagrammes permettent de représenter cette variable: le diagramme en bâtons et le diagramme cumulé.

6.3.1. Diagramme en bâtons

Définition

On veut représenter cette répartition sous la forme d'un diagramme en bâtons. À chaque marque correspond un bâton. Les hauteurs des bâtons sont proportionnelles aux effectifs représentés

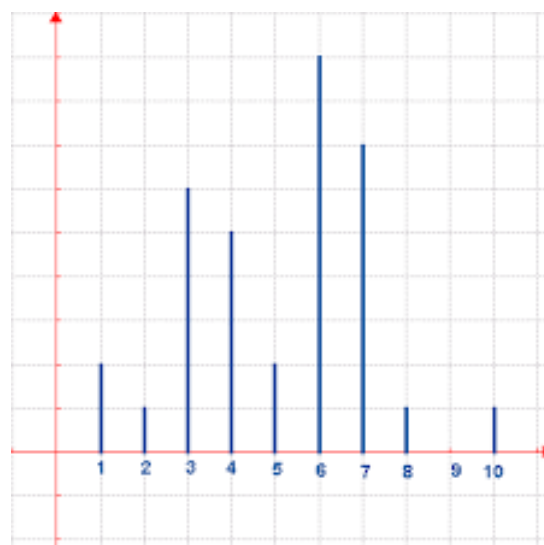
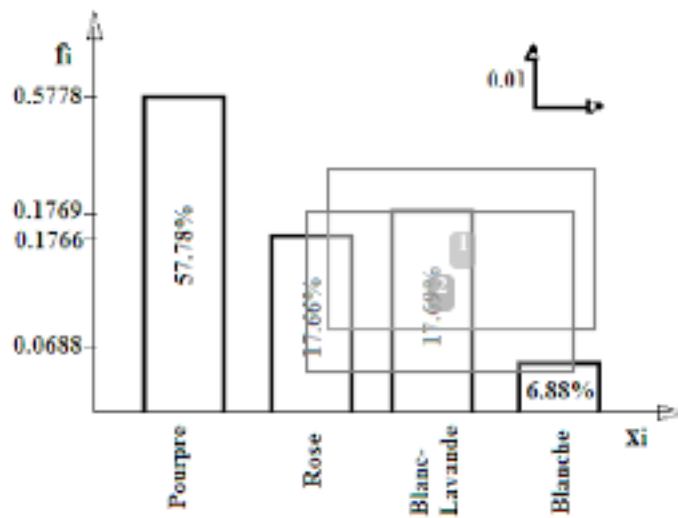


Diagramme à bâtons

7. Exercice : Analyse

Le figure suivant représente une variable statistique qualitative



- 1. vrai
- 2. faux

8. Paramètres de position (caractéristique de tendance centrale)

8.1. Le mode

Définition

Le mode d'une variable statistique est la valeur qui a le plus grand effectif partiel (ou la plus grande fréquence partielle) et il est noté par **Mo**.

Exemple

Considérons les couleurs de dix feutres numérotés d'une trousse d'écolier:

noir, jaune, rouge, bleu, vert, bleu, rouge, rouge, noir et rouge.

Le tableau statistique considéré de variable statistique précédent est:

x_i	noir	jaune	bleu	rouge	vert	total
n_i	2	1	2	4	1	10
f_i	0.2	0.1	0.2	0.4	0.1	1

Donc, Le mode est le rouge.

Remarque

On peut avoir plus d'un mode ou rien.

8.2. La médiane

Définition

On appelle médiane la valeur **Me** de la variable statistique qui vérifie la relation suivante :

$$F_x(Me) < 0.5 < F_x'(Me)$$

La médiane partage la série statistique en deux groupes de même effectif.

Remarque

La série doit toujours être ordonnée par ordre croissant ou décroissant.

N : Taille de l'échantillon.

Si, $N = 2p$ (paire) alors :

$$Me = \frac{X_p + X_{p+1}}{2}$$

Si, $N = 2p + 1$ (impaire) alors :

$$Me = X_{p+1}$$

Exemple

- $E = \{1; 5; 5; 6; 89; 22\}$

La moyenne

$$N = 6 = 2 * 3 = 2 * p$$

$$\text{alors, } Me = \frac{X_3 + X_4}{2} = \frac{5 + 6}{2} = 5.5$$

donc, $5 < Me < 6$

- $E = \{1; 5; 6; 7; 8; 10; 24\}$

$$N = 7 = 2 * 3 + 1 = 2 * p + 1$$

$$\text{alors, } Me = X_{p+1} = 7$$

8.3. La moyenne

Définition

On appelle moyenne de \bar{x} , la quantité

$$\bar{x} = \frac{1}{N} \sum_{i=1}^n n_i x_i = \sum_{i=1}^n f_i x_i$$

ou N est l'effectif total de la série statistique.

9. Paramètres de dispersion (variabilité)

Les indicateurs statistiques de dispersion usuels sont **l'étendue**, **la variance** et **l'écart type**.

9.1. L'étendue

Définition

La différence entre la plus grande valeur et la plus petite valeur du caractère, donnée par la quantité

$$e = x_{\max} - x_{\min},$$

s'appelle **l'étendue** de la V.S X.

Le calcul de l'étendue est très simple. Il donne une première idée de la dispersion des observations. C'est un indicateur très rudimentaire et il existe des indicateurs de dispersion plus élaborés (voir ci-dessous)

9.2. La variance

Définition

On appelle variance de cette série statistique X, le nombre

$$Var(X) = \sum_{i=1}^n f_i (\bar{x} - x_i)^2$$

On dit que la variance est la moyenne des carrés des écarts à la moyenne \bar{x} . Les « écarts à la moyenne » sont les $(\bar{x} - x_i)$, les « carrés des écarts à la moyenne » sont donc les $(\bar{x} - x_i)^2$.

En faisant la moyenne de ces écarts, on trouve la variance.

Le théorème suivant (Théorème de König-Huygens) donne une identité remarquable reliant la variance et la moyenne, parfois plus pratique dans le calcul de la variance

Fondamental : (Théorème de König-Huygens)

Soit (x_i, n_i) une série statistique de moyenne \bar{x} et de variance

$\text{Var}(X)$. Alors,

$$\text{Var}(X) = \sum_{i=1}^p f_i x_i^2 - \bar{x}^2$$

9.3. L'écart type

Définition

La quantité

$$\sigma_X = \sqrt{\text{Var}(X)}$$

s'appelle l'écart type de la V.S X.

10. Série de TD

serie2.pdf (cf. serie2)

Bibliographie



D.Dacunha-Castelle and M.Duflo. Probabilités et statistiques : problèmes à temps fixe. Masson, 1982.

Webographie



https://fad.univ-lorraine.fr/pluginfile.php/23861/mod_resource/content/1/co/variable_qualitative.html

