**Chapitre 2**

**Étude d’une variable statistique discrète**

Le caractère statistique (ou variable statistique, dénotée V.S) 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* : Ω→ {*x*1*, x*2*, ..., xn*}*,*

avec Card(Ω) := *N* est le nombre d’individus dans notre étude.

Nous allons utiliser souvent l’exemple ci-dessous pour illustrer les énoncés de ce cha-pitre.

**Exemple 8**

*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 :*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *xi* | *0* | *1* | *2* | *3* | *4* | *5* | *6* |
|  |  |  |  |  |  |  |  |
| *ni (Eﬀectif)* | *18* | *32* | *66* | *41* | *32* | *9* | *2* |
|  |  |  |  |  |  |  |  |

*Nous avons*

*–* Ω *ensemble des familles.*

*– ω une famille.*

*– X nombre d’enfants par famille*

*X*:*ω*→*X*(*ω*)*.*

12 2.1. EFFECTIF PARTIEL - EFFECTIF CUMULÉ

*On lit, à la famille ω, on associe X*(*ω*) = *le nombre d’enfants de cette famille.*

**2.1 Eﬀectif partiel - eﬀectif cumulé**

On étudie ici un caractère statistique numérique représenté par une suite *xi* décrivant la valeur du caractère avec *i* varie de 1 à *k*.

**2.1.1 Eﬀectif partiel (fréquence absolue)**

**Définition 7**

*Pour chaque valeur xi, on pose par définition*

*ni* = *Card*{*ω* ∈Ω : *X*(*ω*) = *xi*}*.*

*ni : le nombre d’individus qui ont le même xi, ça s’appelle eﬀectif partiel de xi.*

**

|  |  |  |
| --- | --- | --- |
| *ni* | *w* | *xi* |
|  |  |

Figure 2.1: Le nombre d’individus qui prennent la valeur *xi*.

**Exemple 9**

*Dans l’exemple* [*8,*](#page23) *on a* 66 *est le nombre de familles qui ont* 2 *enfants.*

|  |  |  |  |
| --- | --- | --- | --- |
| *xi* | · · · | *2* | · · · |
| *ni (Eﬀectif)* | · · · | *66* | · · · |

|  |  |
| --- | --- |
| 2.2. FRÉQUENCE PARTIELLE - FRÉQUENCE CUMULÉE | 13 |

**2.1.2 Eﬀectif cumulé**

**Définition 8**

*Pour chaque valeur xi, on pose par définition*

*Ni* = *n*1+ *n*2+ *...* + *ni.*

*L’eﬀectif cumulé Ni d’une valeur est la somme de l’eﬀectif de cette valeur et de tous les eﬀectifs des valeurs qui précèdent.*

**Exemple 10**

*Dans l’exemple* [*8,*](#page23) *on a* 50 *est le nombre de familles qui ont un nombre d’enfants inférieur à* 1*. Nous le regardons dans le tableau suivant :*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *xi* | *0* | *1* | *2* | *3* | *4* | *5* | *6* |
|  |  |  |  |  |  |  |  |
| *Ni* | *18* | *50* | *116* | *157* | *189* | *198* | *200* |

**Interprétation** :*Ni*est le nombre d’individus dont la valeur du caractère est inférieurou égale à *xi*. De ce fait, l’eﬀectif total est donné par

*n*

X

*N* =card{Ω) = *ni.*

*i*=1

Dans notre exemple précédent, nous avons *N* = 200.

**2.2 Fréquence partielle - Fréquence cumulée**

Typiquement les eﬀectifs *ni* sont grands et il est intéressant de calculer des grandeurs permettant de résumer la série.

**2.2.1 Fréquence partielle (fréquence relative)**

**Définition 9**

*Pour chaque valeur xi, on pose par définition*

*fi* := *Nni .*

14 2.2. FRÉQUENCE PARTIELLE - FRÉQUENCE CUMULÉE

*fi s’appelle la fréquence partielle de xi. La fréquence d’une valeur est le rapport de l’eﬀectif de cette valeur par l’eﬀectif total.*

**Remarque 3**

*On peut remplacer fi par fi* ×100 *qui représente alors un pourcentage.*

**Interprétation :** *fi*=est le pourcentage des*ω*tel que*X*(*ω*) =*xi*.

**Exemple 11**

*Dans l’exemple précédent,* 0*,* 33 := *il y a* 33*% de familles dont le nombre d’enfants égale à* 2*. Ce pourcentage est calculé de la façon suivante (N* = 200*) :*

|  |  |  |  |
| --- | --- | --- | --- |
| *xi* | · · · | *2* | · · · |
| *ni (Eﬀectif)* | · · · | *66* | · · · |
| *fi (fréquence)* | · · · |  | 66 | = 0*.*33 | · · · |
| 200 |

Nous pouvons conclure la propriété suivante.

**Proposition 1**

*Soit fi défini comme précédemment. Alors,*

*n*

X

*fi* = 1*.*

*i*=1

*Démonstration.* Rappelons que

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | *n* |  |  |  |  |  |  |
|  |  | X |  |  |  |  |  |
|  |  |  | *ni* = *N.* |  |
|  |  | *i*=1 |  |  |  |  |  |
| Ce qui implique que |  |  |  |  |  |  |  |  |
| *n* | *fi* = | *n* | *ni* | = | 1 | *n* | *ni* = 1*.* |
| X |  | X |  |  |  |  | X |  |
| *i*=1 |  | *i*=1 | *N* |  | *N* | *i*=1 |  |
|  |  |  |  |  |  |

|  |  |
| --- | --- |
| 2.2. FRÉQUENCE PARTIELLE - FRÉQUENCE CUMULÉE | 15 |

**2.2.2** **Fréquence cumulée**

**Définition 10**

*Pour chaque valeur xi, on pose par définition*

*Fi* = *f*1+ *f*2+ *...* + *fi.*

*La quantité Fi s’appelle la fréquence cumulée de xi.*

**Interprétation** :*Fi*=est le pourcentage des*ω*tel que la valeur*X*(*ω*)est inférieureou égale à *xi*.

**Exemple 12**

*- Dans l’exemple précédent,* 0*.*785 *représente* 78*.*5% *de familles dont le nombre d’en-fants est inférieur ou égale à* 3*.*

*- Dans un deuxième exemple, nous nous intéressons aux nombres d’erreurs d’assem-blage sur un ensemble d’appareils,*

|  |  |  |
| --- | --- | --- |
| *Nombre d’erreurs* | *Nombre d’appareils* | *Fréquences cumulées* |
|  |  |  |
| 0 | *101* | *0.26* |
|  |  |  |
| 1 | *140* | *0.61* |
|  |  |  |
| 2 | *92* | *0.84* |
|  |  |  |
| 3 | *42* | *0.94* |
|  |  |  |
| 4 | *18* | *0.99* |
|  |  |  |
| 5 | *3* | *1* |
|  |  |  |

*Nous avons* 94% *des appareils qui ont un nombre d’erreurs d’assemblage inférieur ou égale à* 3*.*

Nous avons vu que les tableaux sont un moyen souvent indispensable, en tous cas très utile, de classification et de présentation des unités d’une population statistique. Dans le pa-ragraphe suivant, nous allons voir comment on traduit ses tableaux en graphique permettant aussi de résumer d’une manière visuelle les données.

16 2.3. REPRÉSENTATION GRAPHIQUE DES SÉRIES STATISTIQUES

**2.3** **Représentation graphique des séries statistiques**

On distingue les méthodes de représentation d’une variable statistique en fonction de la nature de cette variable (qualitative ou quantitative). Les représentations recommandées et les plus fréquentes sont les tableaux et les diagrammes (graphe).

Le graphique est un support visuel qui permet :

**La synthèse** : visualiser d’un seul coup d’œil les principales caractéristiques (mais onperd une quantité d’informations), voir Figure [2.2.](#page28)



Figure 2.2: Quelques caractéristiques du graphique

**La découverte** : met en évidence les tendances.

**Le contrôle** : on aperçoit mieux les anomalies sur un graphique que dans un tableau. **La recherche des régularités** : régularité dans le mouvement, répétition du phéno-mène.

**2.3.1 Distribution à caractère qualitatif**

A partir de l’observation d’une variable qualitative, deux diagrammes permettent de représenter cette variable : le diagramme en bandes (dit tuyaux d’orgue) et le diagramme à secteurs angulaires (dit camembert).

|  |  |
| --- | --- |
| 2.3. REPRÉSENTATION GRAPHIQUE DES SÉRIES STATISTIQUES | 17 |

**Tuyaux d’orgues**

Nous portons en abscisses les modalités, de façon arbitraire. Nous portons en ordonnées des rectangles dont la longueur est proportionnelle aux eﬀectifs, ou aux fréquences, de chaque modalité (voir Figure [2.3)](#page29).



Figure 2.3: Tuyaux d’orgues

**Diagramme par secteur (diagramme circulaire)**

Les diagrammes circulaires, ou semi-circulaires, consistent à partager un disque ou un demi-disque, en tranches, ou secteurs, correspondant aux modalités observées et dont la surface est proportionnelle à l’eﬀectif, ou à la fréquence, de la modalité (voir Figure [2.4)](#page29).



Figure 2.4: Diagramme par secteur

Le degré d’un secteur est déterminé à l’aide de la règle de trois de la manière suivante :

* −→ 360◦

*ni* −→ *di* (*degr*é *de la modalit*é *i*)*.*

18 2.3. REPRÉSENTATION GRAPHIQUE DES SÉRIES STATISTIQUES

Donc,

*di* = *ni* ×360*.*

*N*

**2.3.2 Distribution à caractère quantitatif discret**

A partir de l’observation d’une variable quantitative discrète, deux diagrammes per-mettent de représenter cette variable : le diagramme en bâtons et le diagramme cumulatif (voir ci-dessous).

Pour l’illustration, nous prenons l’exemple précédent de départ (nombre d’enfants par famille). Nous rappelons le tableau statistique associe.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *xi* | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|  |  |  |  |  |  |  |  |
| *ni* | 18 | 32 | 66 | 41 | 32 | 9 | 2 |
|  |  |  |  |  |  |  |  |

**Diagramme à bâtons**

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 (voir Figure [2.5)](#page30).



Figure 2.5: Diagramme à bâtons

**2.3.3 Représentation sous forme de courbe et fonction de répartition**

Nous avons déjà abordé les distributions cumulées d’une variable statistique. Nous allons dans cette partie exploiter ses valeurs cumulées pour introduire la notion de la fonction de répartition. Cette notion ne concerne que les variables quantitatives.

|  |  |
| --- | --- |
| 2.3. REPRÉSENTATION GRAPHIQUE DES SÉRIES STATISTIQUES | 19 |

Soit la fonction *Fx* : R → [0*,* 1] définie par

*Fx*(*x*) :=pourcentage des individus dont la valeur du caractère est≤ *x.*

Cette fonction s’appelle la fonction de répartition du caractère *X*.

**Remarque 4**

*Pour tout i* ∈{1*, . . . , n*}*, on a*

*Fx*(*xi*) = *Fi.*

*La courbe de Fx passe par les points* (*x*1*, F*1)*,* (*x*2*, F*2)*, ... et* (*xn, Fn*)*.*

En se basant sur notre exemple, la courbe de *Fx* est représentée ci-dessous (Figure [2.6)](#page31)

sur

R =] − ∞*,* 0[ ∪ [0*,* 1[ ∪ *....* ∪ [6*,* +∞[*.*

Dans ce cas, nous avons

– Si *x <* 0, alors *Fx*(*x*) = 0.

– Si *x* ∈ [0*,* 1[, alors *Fx*(*x*) = 0*.*09.

...

– Si *x* ≥ 6, alors *Fx*(*x*) = 1.

Cette courbe s’appelle "la courbe cumulative des fréquences". La courbe cumulative est une courbe en escalier représentant les fréquences cumulées relatives.



Figure 2.6: Représentation d’une variable quantitative discrète par la courbe cumulative.

20 2.4. PARAMÈTRES DE POSITION

**Proposition 2**

*La fonction de répartition satisfait, pour i* ∈{1*, . . . , n*}*,*

*– l’égalité, Fx*(*xi*) = *Fi,*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | 0*,* |  |  | *si* | *x < x*1*,* |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  | *F* |  | *,* | *si* | *x* | 1 |  |  | *x < x ,* |  |
|  |  |  |  |  | 1 |  |  |  |  |  |  |  | 2 |  |
|  | *F* | *x* |  |  |  |  |  |  |  |  | ≤ |  |  |  |  |
| *– l’expression,* |  | *x*( ) = |  |  |  |  |  |  |  |  |  |  |  |  |  | *.* |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  | *F* | *,* | *si* | *x* | *i* | ≤ | *x < x* | *i*+1 | *,* |
|  |  |  |  |  | *i* |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  | 1 | *,* |  |  | *si* | *x* |  |  | *x* | *.* |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  | *n* |  |  |  |
|  |  |  |  |  |  |  |  |  |  | ≥ |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**2.4 Paramètres de position (caractéristique de tendance cen-**

**trale)**

Les indicateurs statistiques de tendance centrale (dits aussi de position) considérés fréquemment sont la moyenne, la médiane et le mode.

**Le mode**

Le mode d’une V.S est la valeur qui a le plus grand eﬀectif partiel (ou la plus grande fréquence partielle) et il est dénoté par *Mo*.



**Exemple 13**

*Dans l’exemple* [*8,*](#page23) *le mode est égal à* 2 *qui correspondant au plus grand eﬀectif.*

|  |  |
| --- | --- |
| 2.4. PARAMÈTRES DE POSITION | 21 |

**Remarque 5**

*On peut avoir plus d’un mode ou rien.*

**La médiane**

On appelle médiane la valeur Me de la V.S X qui vérifie la relation suivante :

*Fx*(*M e*−) *<* 0*.*5≤ *Fx*(*M e*+) = *Fx*(*M e*)*.*

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

**Exemple 14**

*Dans l’exemple* [*8,*](#page23) *la relation*

*Fx*(0) = 0 *<* 0*.*5≤ *Fx*(0+) = 0*.*09

*n’est pas satisfaite. Donc, la médiane est diﬀérente de* 0*. Par contre, nous avons*

*Fx*(2−) = 0*.*25 *<* 0*.*5≤ *Fx*(2+) = *F* (2) = 0*.*58*.*

*Donc, M e* = 2*.*

**La moyenne**

On appelle moyenne de *X*, la quantité

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *x* =1 | *n* | *nixi* = | *n* | *fixi,* |
|  |  |  | X |  | X |  |
|  |  |  |  |  |  |  |
|  |  | *N* | *i*=1 |  | *i*=1 |  |
|  |  |  |  |  |

avec *N* = Card(Ω). On peut donc exprimer et calculer la moyenne dite "arithmétique" avec des eﬀectifs ou avec des fréquences.

**Exemple 15**

*Si x* = 2*.*46*, alors nous avons au moyenne une famille de quartier a* 2*.*46 *d’enfants.*

La valeur de la moyenne est abstraite. Comme dans l’exemple précédent, *x* = 2*.*46 est un chiﬀre qui ne correspond pas à un fait concret.

22 2.5. PARAMÈTRES DE DISPERSION (VARIABILITÉ)

La moyenne arithmétique dont on vient d’indiquer la formule est dite moyenne pon-dérée ; cela signifie que chaque valeur de la variable est multipliée (pondérée) par un coef-ficient, ici par l’eﬀectif *ni* qui lui correspond. Dans ce cas, chaque valeur *xi* de la variable intervient dans le calcul de la moyenne autant de fois qu’elle a été observée. On parle de moyenne arithmétique simple quand on n’eﬀectue pas de pondération. Par exemple, si 5 étudiants ont pour âge respectif 18, 19, 20, 21 et 22 ans, leur âge moyen est donné par (18 + 19 + 20 + 21 + 22)*/*5 = 20 ans.

**Remarque 6**

*Nous mentionnons qu’il existe d’autres moyennes que la moyenne arithmétique*

**2.5 Paramètres de dispersion (variabilité)**

Les indicateurs statistiques de dispersion usuels sont l’étendue, la variance et l’écart-

type.

**L’étendue**

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

* = *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).

**La variance**

On appelle variance de cette série statistique *X*, le nombre

*n*

X

*V ar*(*X*) = *fi*(*x* − *xi*)2

*i*=1

On dit que la variance est la moyenne des carrés des écarts à la moyenne *x*. Les « écarts à la moyenne » sont les (*x* −*xi*), les « carrés des écarts à la moyenne » sont donc les (*x* −*xi*)2. En faisant la moyenne de ces écarts, on trouve la variance.

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

|  |  |
| --- | --- |
| 2.5. PARAMÈTRES DE DISPERSION (VARIABILITÉ) | 23 |

**Théorème 1**

*Soit* (*xi, ni*) *une série statistique de moyenne x et de variance V ar*(*X*)*. Alors,*

*n*

X

*V ar*(*X*) = *fix*2*i* − *x*2*.*

*i*=1

*Démonstration.* Par définition, nous avons

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *V ar*(*X*) = *fi*(*x* − *xi*)2 | = *N* |  |  |  |  |  |  |  |  |  |  |  | *n* | *n* | *.* |
|  |  | *ni*(*x* − *xi*)2 = X |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | *ni*( |  | − *xi*)2 |  |
| *n* |  |  |  |  |  |  |  |  | 1 |  | *n* |  |  |  |  |  | *x* |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | *i*=1 |  |  |  |  |  |  |
| X |  |  |  |  |  |  |  |  |  | X |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | *ni* |  |
| *i*=1 |  |  |  |  |  |  |  |  |  | *i*=1 |  |  |  |  |  |  |  |  |  | X |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | *i*=1 |  |
| Donc, |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | *n* |  |  |  |  |  |  |  |  |  |  |  |  | *n* |  |  |  |  | *n* |  |  |  |  | *.* |  |
| *V ar*(*X*) =X *n* |  |  |  |  |  |  |  | = X |  |  |  |  |  |  |  |  |  |
|  |  |  |  | *ni*( |  |  | − *xi*)2 |  |  |  | *ni*( |  | 2 | + *xi*2 | − 2 | *xxi*) |  |
|  |  |  |  | *x* |  |  |  | *x* |  |
|  |  | *i*=1 |  |  |  |  |  |  |  |  |  |  |  |  | *i*=1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | X |  |  |  |  |  |  |  |  |  |  |  |  |  |  | X |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  | *ni* |  |  |  |  |  |  |  |  |  |  |  | *ni* |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | *i*=1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  | *i*=1 |  |  |  |  |  |  |  |  |  |  |
| Par égalité, nous avons |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | *n* |  |  |  |  |  |  |  | *n* |  |  |  |  | *n* |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | X |  |  |  |  |  |  |  | X |  |  |  | X |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  | *ni* |  | 2 |  |  |  |  | *nix*2 |  |  |  |  | 2*ni* |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  | *x* |  |  |  |  |  |  |  |  | *xxi* |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | *i* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *V ar*(*X*) = |  | *i*=1 |  |  |  |  | + |  | *i*=1 | − |  | *i*=1 |  |  |  |  | *.* |  |  |  |  |  |
|  | *n* |  |  |  |  |  | *n* |  |  | *n* |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | X |  |  |  |  |  |  |  | X |  |  |  |  | X |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  | *ni* |  |  |  |  | *ni* |  |  |  |  |  | *ni* |  |  |  |  |  |  |
|  |  |  |  |  | *i*=1 |  |  |  |  |  |  |  | *i*=1 |  |  |  |  | *i*=1 |  |  |  |  |  |  |  |  |  |  |

Ce qui implique que

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 2 | *n* | 2 |  |  | 2 |
| X |  |  |
|  |  |  | *nixi*2 |  |  |  |  |  | 1 | *n* |  |
|  |  |  | *i*=1 |  |  |  |  |  |  | X |  |
| *V ar*(*X*) = *x* + |  |  |  |  |  |  |  |
| *ni* |  |  |  |  |  |  |  |
| *n* | −2*xx* =−*x* + *N i*=1 *nixi .* |
|  |  |  | X |  |  |  |  |  |  |  |  |

*i*=1

**Remarque 7**

*Dans l’utilisation de la formule du théorème précédent, il faut veiller à remplacer x par sa valeur approchée la plus précise possible.*

24 2.6. EXERCICES CORRIGÉS

**L’écart type**

La quantité

q

*σX* = *V ar*(*x*)

s’appelle l’écart type de la V.S *X*.

**Remarque 8**

*Le paramètre σx mesure la distance moyenne entre x et les valeurs de X (voir Figure* [*2.*](#page36)*7). Il sert à mesurer la dispersion d’une série statistique autour de sa moyenne.*

*– Plus il est petit, plus les caractères sont concentrés autour de la moyenne (on dit que la série est homogène).*

*– Plus il est grand, plus les caractères sont dispersés autour de la moyenne (on dit que la série est hétérogène).*

**

Figure 2.7: La dispersion d’une série statistique autour de sa moyenne

**2.6** **Exercices corrigés**

**Exercice 7**

*- Le tableau suivant donne la répartition selon le groupe sanguin de 40 individus pris au hasard dans une population,*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | *Groupes sanguins* | *A* | *B* | *AB* | *O* |  |
|  |  |  |  |  |  |  |  |
|  |  | *L’eﬀectif* | *20* | *10* | *n*3 | *5* |  |
|  |  |  |  |  |  |  |  |
| *1.* | *Déterminer la variable statistique et son type.* |  |  |  |
| *2.* | *Déterminer l’eﬀectif des personnes ayant un groupe sanguin AB.* |
| *3.* | *Donner toutes les représentations graphiques possibles de cette distribution.* |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

|  |  |
| --- | --- |
| 2.6. EXERCICES CORRIGÉS | 25 |

**Solution** 1 - La population dans cette étude est les40personnes. Donc*N*= 40. La variablestatistique est le groupe sanguin des individus et elle est qualitative.

2 - L’eﬀectif total est égal à 40. Par conséquent,

4

X

*N* = 40 = *ni.*

*i*=1

Alors,

20 + 10 + *n*3 + 5 = 40*.*

Ce qui implique que *n*3 = 5.

3- Nous avons deux représentations possibles "Tuyaux d’orgue" et "Diagramme en sec-teur".

*Effectif*

**

|  |  |  |
| --- | --- | --- |
| *20* | B | A |
| *15* | O |
|  |
| *10* |  |
| *5* | AB |  |
|  |  |
| *o ABBA* | *Groupes sanguins* |  |

Figure 2.8: A gauche "Tuyaux d’orgue" et à droite "Diagramme en secteur"

Les angles dans la figure [2.8](#page37) se calcule en utilisant la règle de trois.

**Exercice 8**

* *Le gérant d’un magasin vendant des articles de consommation courante a relevé pour un article particulier qui semble connaître une très forte popularité, le nombre d’articles vendus par jour. Son relevé a porté sur les ventes des mois de Mars et Avril, ce qui correspond à*

*52 jours de vente. Le relevé des observations se présente comme suit :*

*7138109121089106147159111211125141181014128*

*5713121611911111212151451499141311101112915.*

1. *Quel type est la variable statistique étudiée.*
2. *Déterminer le tableau statistique en fonction des eﬀectifs, des fréquences, des eﬀectifs cumulés et des fréquences cumulés.*
3. *Tracer le diagramme des bâtonnés associé à la variable X.*
4. *Soit Fx la fonction de répartition. Déterminer Fx.*
5. *Calculer le mode Mo et la moyenne arithmétique x.*

26 2.6. EXERCICES CORRIGÉS

1. *Déterminer à partir du tableau puis à partir du graphe, la valeur de la médiane Me.*
2. *Calculer la variance et l’écart-type.*

**Solution** 1 - La population est les52jours et la variable statistique étudiée est le nombred’articles vendus par jour. Son type est bien évidement quantitatif discret (nombre).

2 - Le tableau statistique est donné par

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *xi* | 5 | 6 | 7 |  | 8 |  | 9 |  | 10 | 11 | 12 | 13 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *ni* | 3 | 1 | 3 |  | 4 |  | 7 |  | 5 | 8 | 8 | 3 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *fi* | 3/52 | 1/52 | 3/52 |  | 4/52 |  | 7/52 |  | 5 /52 | 8 /52 | 8/52 | 3 /52 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Ni* | 3 | 4 | 7 |  | 11 |  | 18 |  | 23 | 31 | 39 | 42 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| *Fi* | 3/52 | 4/52 | 7/52 | 11/52 | 18/52 |  | 23/52 | 31/52 | 39/52 | 42/52 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  | 14 |  | 15 |  | 16 |  |  | P |  |  |  |  |
|  |  |  |  | 6 |  |  | 3 |  |  | 1 |  |  | *N*=52 |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  | 6 /52 |  | 3/52 |  | 1/52 |  | 1 |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  | 48 |  | 51 |  | 52 |  |  | ? |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  | 48/52 |  | 51/52 |  | 1 |  |  | ? |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

3 - L’élaboration du diagramme des bâtonnets de *X*,

*ni ou fi*

**

|  |  |
| --- | --- |
| *8* |  |
| *7* |  |
| *6* |  |
| *5* |  |
| *4* |  |
| *3* |  |
| *2* |  |
| *1* |  |
| *56* | *78910111213141516 xi* |
|  | Figure 2.9: Diagramme à bâtons |
|  |  |

|  |  |
| --- | --- |
| 2.6. EXERCICES CORRIGÉS | 27 |
| 4 - La fonction de répartition est donnée par |
|  |  |  | 0*,* |  | *si x <* 5*,* |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  | 3*/*52*,* | *si* 5 *x <* 6*,* |
|  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  | ≤ |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  | 4*/*52*,* | *si* 6 *x <* 7*,* |
|  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| *F* | *x* |  |  |  |  |  |  | ≤ |  |
|  | *x*( ) = |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  | 7*/*52*,* | *si* 7 |  | *x <* 8*,* |
|  |  |  | . | . |  | . | . | ≤ |  |
|  |  |  |  | . |  | . |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  | 1*,* |  | *si x* |  | 16*.* |
|  |  |  |  |  |
|  |  |  |  |  |  |  |  | ≥ |  |
|  |  |  |  |  |  |  |  |  |

5 - Le mode est la valeur de la variable qui a le plus grand eﬀectif, c’est à dire, *ni* = 8. Donc,

*Mo* = 11 et *Mo* = 12*.*

La moyenne arithmétique est donnée par ;

* + 1. 12

1 XX

*x* = *N i*=1 *nixi* = *i*=1 *fixi.*

Par conséquent,

*x* =521(3×5 + 1×6 + 3×7 + *...* + 1×16) =55552= 10*.*67*.*

6 - La médiane est la valeur de la variable qui divise la population de la série statistique en deux parties égales. Nous avons,

*Fx*(11−) =2352 *<* 0*.*5≤ *Fx*(11+) = *F* (*M e*) =3152*.*

Donc, *M e* = 11.

7 - Nous commençons par la variance,

*n*

*V ar*(*X*) = *N*1X *nix*2*i* − *x*2*.*

*i*=1

Après calcule, on trouve

* *ar*(*X*) = 7*.*64*.*

Par conséquent, l’écart type est calculé à partir de

|  |  |  |  |
| --- | --- | --- | --- |
|  | *σX* =q | *V ar*(*x*) | = 2*.*76*.* |
|  |  |  |  |

28 2.6. EXERCICES CORRIGÉS

**Exercice 9**

* *On considère deux groupes d’étudiants. Nous relevons leurs notes d’examens dans les deux tableaux suivants :*

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Note (groupe A) |  | 8 |  |  | 9 |  | 10 |  | 11 |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Eﬀectif |  | 2 |  |  | 2 |  | 1 |  | 1 |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| Note (groupe B) |  | 6 |  | 8 |  | 9 |  | 13 |  | 14 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Eﬀectif |  | 2 |  | 2 |  | 2 |  | 1 |  |  | 1 |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

*Calculer la moyenne et l’écart type de chaque groupe. Comparer les deux groupes.*

**Solution** Dans un premier temps, nous remarquons que l’eﬀectif total du groupe A est égalà 6 et celui du groupe B est égal à 8.

En utilisant la formule de la moyenne, nous obtenons

*xA* = 9*.*2 et *xB* = 9*.*1*.*

On remarque que les moyennes sont très proches. Peut-on pour autant conclure que ces deux groupes ont des niveaux identiques ?

Nous répondons à cette question après le calcule des écarts type. Ils sont donnés

par

*σXA* = 1*.*11 et *σXB* = 2*.*8*.*

Nous remarquons que même si les deux groupes ont des moyennes quasiment identiques, le groupe B est beaucoup plus dispersé que le groupe A car *σXB* *> σXA* . Les étudiants de ce groupe ont des notes plus irréguliers. On peut dire donc que le groupe B est moins homogènes

que le groupe A. En observant les valeurs du tableau, on voit que c’est cohérent.

**Exercice 10**

* *Un quartier résidentiel comprend* 99 *unités d’habitation ayant une valeur locative moyenne de* 10000 *Da. Deux nouvelles unités d’habitation sont construites dans le quartier : l’une a une valeur locative de* 7000 *Da et l’autre, une villa luxueuse, a une valeur locative de* 114000 *Da.*

*– Quelle est la nouvelle moyenne de valeur locative pour le quartier ?*

*– Pouvait-on s’attendre à de tel résultat ?*

**Solution** - Le nouveau total des mesures de valeur locative est

(99 × 10000) + 7000 + 114000 = 1111000*.*

|  |  |
| --- | --- |
| 2.7. EXERCICES SUPPLÉMENTAIRES | 29 |

Le nouveau total d’individus statistiques est 99 + 2 = 101. La nouvelle moyenne est donc

1111000 = 110000*.*

101

* On pouvait s’attendre à une augmentation de la moyenne car l’une des deux nouvelles valeurs est très nettement au dessus de la moyenne initiale.

**2.7** **Exercices supplémentaires**

**Exercice 11**

* *Pour déterminer le type de logement* (*F* 2*, F* 3*, ...*) *à construire, on étudie* 20 *familles selon leur nombre d’enfants. Durant l’expérience, on note les résultats suivants :*

1*,*3*,*5*,*5*,*3*,*2*,*4*,*4*,*7*,*0*,*2*,*4*,*3*,*7*,*0*,*5*,*4*,*2*,*3*,*2

*– Déterminer, la population, l’unité (individu), la variable statistique et les modalités.*

*– Déterminer le tableau statistique avec xi, ni, fi et Fi.*

X

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Modalité | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |

*ni*

*Ni*

*fi*

*Fi*

**Exercice 12**

*- Voici le tableau des pourcentages obtenu pour la variable " Mode de logement" :*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *xi* | "Cité U" | "Studio" | "Résidence" | "Maison" | "Autre" | TOTAL |
|  |  |  |  |  |  |  |
| % | 4.8 | 16.5 | 38.6 | 28.6 | 11.6 | 100 |
|  |  |  |  |  |  |  |

*Sachant que la taille de l’échantillon N* = 189*, retrouver les eﬀectifs pour chaque modalité.*

**Exercice 13**

* *Au poste de péage, on compte le nombre de voitures se présentant sur une période de* 5 *min. Sur* 100 *observations de* 5 *min, on obtient les résultats suivants :*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 30 |  |  |  |  |  | 2.7. EXERCICES SUPPLÉMENTAIRES |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | *Nombre de voitures* | *1* | *2* | *3* | *4* |  | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | *Nombre d’observations* | *2* | *8* | *14* | *20* |  | *19* | *15* | *9* | *6* | *2* | *3* | *1* | *1* |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

1. *Construire la table des fréquences et le diagramme en bâtons en fréquences de la série du nombre de voitures.*
2. *Calculer la moyenne et l’écart-type de cette série.*
3. *Déterminer la médiane.*

**Exercice 14**

*- Dans une petite localité, on a relevé le nombre de pièces par appartement :*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *Nombre de pièces* | *1* | *2* | *3* | *4* | *5* | *6* | *7* |
|  |  |  |  |  |  |  |  |
| *Nombre d’appartements* | *48* | *72* | *96* | *64* | *39* | *25* | *3* |
|  |  |  |  |  |  |  |  |

*Le « nombre de pièces par appartement » est à considérer comme une variable aléatoire discrète à valeurs entières. (A l’interprétation, il faudra préciser que les « demi pièces » ne sont pas comptabilisées).*

*– Déterminer le tableau statistique.*

*– Tracer le diagramme des bâtonnés et la courbe des fréquences cumulées associés à la variable statistique.*

*– Calculer la moyenne et l’écart-type de cette série.*

*– Déterminer la médiane.*

**Exercice 15**

*- Une machine coupe des barres de* 12 *cm. Mais malheureusement, elle n’est pas bien réglée et les longueurs varient autour de la valeur attendue. Une étude sur* 185 *barres donnent les résultats suivants :*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | *Longueurs en cm* | *11.5* | *11.6* | *11.7* | *11.8* | *11.9* | *12* | *12.1* | *12.2* | *12.3* |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | *Eﬀectif* | *3* | *15* | *16* | *16* | *18* | *20* | *25* | *25* | *28* |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| *1.* | *Quel type est la variable statistique étudiée.* |  |  |  |  |  |  |
| *2.* | *Déterminer le tableau statistique.* |  |  |  |  |  |  |  |  |
| *3.* | *Tracer le diagramme des bâtonnés et la courbe des fréquences cumulées associés à la* |
|  | *variable statistique.* |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |

|  |  |
| --- | --- |
| 2.7. EXERCICES SUPPLÉMENTAIRES | 31 |

*4. Calculer la moyenne et l’écart type.*

**Exercice 16**

*- La répartition en 2016 du nombre de pièces des résidences principales en Algérie est*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | *Nombre de pièces* | *1* | *2* | *3* | *4* |  | *5* | *6 et plus* |
|  |  |  |  |  |  |  |  |  |  |
|  |  | *Eﬀectif (*×103*)* | *1200* | *2700* | *4700* | *5500* |  | *3500* | *2500* |
| *–* | *Donner l’individu de la population statistique étudiée.* |  |  |
| *–* | *Quelle est la nature de la variable étudiée ?* |  |  |  |  |
| *–* | *Quel outil graphique peut-on utiliser pour représenter cette variable ?* |

**Exercice 17**

*- On observe le nombre d’arrivées des clients à un bureau de poste pendant un intervalle de temps donné (disant* 10 *minutes). En répétant* 100 *fois cette observation, on obtient les résultats suivants :*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *Nombre d’arrivées* | *1* | *2* | *3* | *4* | *5* | *6* | *Total* |
|  |  |  |  |  |  |  |  |
| *Nombre d’observations* | *15* | *25* | *26* | *20* | *7* | *7* | *100* |
|  |  |  |  |  |  |  |  |

*– Représenter graphiquement ces résultats (pour les eﬀectifs et pour les fréquences cumulées).*

*– Calculer la valeur de la moyenne arithmétique, de la médiane, de la variance et de l’écart type des résultats.*

**Exercice 18**

* *On mesure les diamètres de troncs d’arbres d’une même espèce. On étudie* 400 *échantillons. On obtient les résultats suivants :*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Diamètre en cm | 25 | 26 | 27 | 28 | 29 | 30 |
|  |  |  |  |  |  |  |
| Pourcentage | 10% | 15% | 30% | 35% | 5% | 5% |
|  |  |  |  |  |  |  |

*On donne :*

(25 × 0*.*1) + (26 × 0*.*15) + (27 × 0*.*3) + (28 × 0*.*35) + (29 × 0*.*05) + (30 × 0*.*05) = 27*.*25*.*

(252 × 0*.*1) + (262 × 0*.*15) + (272 × 0*.*3) + (282 × 0*.*35) + (292 × 0*.*05) + (302 × 0*.*05) = 744*.*05*.*

1. *Établir le tableau statistique en fonction des eﬀectifs et des fréquences relatives.*
2. *Quel est le diamètre moyen de ces troncs d’arbres ?*

32 2.7. EXERCICES SUPPLÉMENTAIRES

1. *Déterminer la variance puis l’écart-type de la série statistique résumée dans le tableau ci-dessus.*
2. *Représenter graphiquement ces résultats (juste pour les eﬀectifs).*
3. *Déterminer le mode et donner son interprétation.*