***Étude d’une variable statistique Discrète***

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.

***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*.

**Exemple**

*Dans l’exemple précédent,* 66 *est le nombre de familles qui ont* 2 *enfants.*



***Effectif cumulé***

**Définition**

*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**

*Dans l’exemple précédent :* 50 *est le nombre de familles qui ont un nombre d’enfant inférieur à* 1*. Nous le regardons dans le tableau suivant :*

 

***Fréquence partielle - Fréquence cumulée***

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

***Fréquence partielle (fréquence relative)***

**Définition**

*Pour chaque valeur xi, on pose par définition*

 

*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**

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

*Soit fi défini comme précédemment. Alors,*



***Fréquence cumulée***

**Définition** *Pour chaque valeur xi, on pose par définition*

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

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

*.*

***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 cumulatif (voir ci-dessous).

Pour l’illustration, nous prenons l’exemple précédent de départ (nombre d’enfants par famille)

***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



***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.

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*.

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



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



***Paramètres de position (caractéristique de tendance centrale)***

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 *M*0.



**Remarque**

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

***La médiane***

La série ordonnée par ordre croissant. On appelle cette série ordonnée la statistique d’ordre. Cette notation, très usuelle en statistique, permet de définir la médiane de manière très synthétique.

**-Si n est impair**



**– Si n est pair**



***La moyenne***

On appelle moyenne de *X*, la quantité



***Étude d’une variable statistique continue***

Nous rappelons qu’une variable statistique (V.S) quantitative concerne une grandeur mesurable. Ses valeurs sont des nombres exprimant une quantité et sur lesquelles les opérations arithmétiques (addition, multiplication, etc,...) ont un sens.

**Définition**

*On appelle V.S continue (ou caractère continu) toute application de Ω* *et à valeurs réelles et qui prend un nombre "important" de valeurs (Les caractères continus sont ceux qui ont une infinité de modalités).*

**Remarque 9**

*Comment étudier ce caractère ?*

***Réponse*** *: Partager les valeurs prises par X en classes de valeurs.*

***Classe de valeurs***

**Définition**



Dès qu’un caractère est identifié en tant que continu, ces modalités *Ck* = [*Lk,Lk*+1[

sont des intervalles avec

– *Lk* : borne inférieure.

– *Lk*+1 : borne supérieure.

– *ak* = *Lk*+1 − *Lk* : son amplitude, son pas ou sa longueur.

– *Ck* = *xk* = (*Lk*+1 + *Lk*)*/*2 : son centre.



***Effectif et fréquence d’une classe***



**Définition**

*Le nombre*

*fi* := *ni/N*

*est appelé la fréquence partielle de Ci.*

*On appelle l’effectif cumulé de Ni la quantité*



*On appelle la fréquence cumulée de Ci la quantité*



***Fonction de répartition***

**Notation** : Nous allons noter par

*Ci* = [*x*min = *a*0*, x*min+1 = *a*1[*.*



***Paramètres de tendance centrale***

On note par *Ci* le centre de la classe *Ci* et nous considérons *fi* la fréquence partielle de *Ci*.

***La moyenne***

**Définition**

*La quantité*



***Le mode***

La définition suivante permet de comprendre la démarche à suivre pour calculer le mode d’une manière exacte et qui se trouve dans une des classes appelée "classe modale".

Pour avoir une valeur exacte, le mode se calcule de la manière suivante



Avec

Xm : limite inférieure de la classe modale ;

i : amplitude de la classe modale ;

Δi : écart d’effectif entre la classe modale et la classe inférieure la plus proche

Δs : écart d’effectif entre la classe modale et la classe supérieure la plus proche

***LA MEDIANE***

On cherche la classe contenant le 𝒏/𝟐 individu de l’échantillon. Cette classe est appelée la classe médiane. En supposant que tous les individus de cette classe sont uniformément répartis à l’intérieur, la médiane se calcule de la façon suivante par interpolation linéaire :



xm : limite inférieure de la classe médiane ;

a : amplitude de la classe médiane ;

ni : effectif de la classe médiane

Ni. : Effectif cumulé inférieur à xm

n : taille de l’échantillon

***Exercice 1***

*1/Soit la répartition des salaires journaliers des employés d’une usine « A » :*

**

*1. Déterminer la population statistique, le caractère étudié et sa nature.*

*2. Tracer l’histogramme des effectifs et le polygone des effectifs.*

*2/Soit la répartition des salaires journaliers des 620 employés d’une usine « B » :*

**

*1. Compléter le tableau.*

***Exercice 2***

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



*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.*

***Exercice 3***

*- 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 :*

*7 13 8 10 9 12 10 8 9 10 6 14 7 15 9 11 12 11 12 5 14 11 8 10 14 12 8*

*5 7 13 12 16 11 9 11 11 12 12 15 14 5 14 9 9 14 13 11 10 11 12 9 15.*

*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.*

*6. Déterminer à partir du tableau puis à partir du graphe, la valeur de la médiane Me.*

*7. Calculer la variance et l’écart-type.*