

Statistique L01 Psychologie

Enseignant: PUTOIS Benjamin
Laboratoire E.M.C. (salle K18)
Benjamin.Putois@univ-lyon2.fr

GÉNÉRALITÉS

13 semaines de cours plus examen sur table en amphi (1h30) la 14^e semaine avec sujet commun aux différents groupes.

Examen avec utilisation des tables, d'un formulaire distribué par l'enseignant et d'une calculatrice.

3 parties :

1. Statistiques descriptives
2. Loi normale
3. Corrélation

1. Demandez aux étudiants ce qu'ils ont fait un bac S, L et ES...

Statistique L01 Psychologie

Definition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne
Mode
Médiane

Paramètre de Dispersion

Ecart-type
Quartiles

Représentation Graphique

Origine:

L'origine du mot « statistique » remonte au latin classique *status* (état) qui, par une série d'évolutions successives, aboutit au français *statistique*, attesté pour la première fois en 1771.

Les statistiques et la statistique:

Les statistiques peuvent être définie comme l'étude méthodique des faits sociaux qui définissent un Etat, par des procédés numériques (dénombrements, inventaires, recensements,...)

La statistique est l'ensemble de techniques d'interprétation mathématique appliquées à des phénomènes (ex : faits sociaux) pour lesquels une étude exhaustive de tous les facteurs est impossible à cause de leur grand nombre ou de leur complexité.

Histoire des statistiques:

De tous temps, les chefs d'Etat ont souhaité déterminer la puissance des nations qu'ils dirigeaient à l'aide de recensements partiels ou complets (population, territoire, production,...)

Au début du XXe Siècle, un débat oppose les partisans des recensements (réalisés sur l'ensemble de la population) et des sondages (réalisés sur un échantillon représentatif de la population).

Statistique L01 Psychologie

Définition

Population et échantillon

Variable aléatoire

Variable continue et discrète

Série statistique

Variable continue

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Population et échantillon

Les recensements ne sont pas toujours possibles, ni souhaitables. Dans certains cas, ils peuvent être trop chers (comme, par exemple, des enquêtes sur toute la population d'un pays). Ils peuvent aussi contenir des erreurs. Parfois, ils sont carrément aberrants (mesurer la solidité moyenne d'un type de voiture en lançant toutes les voitures de ce type contre un mur serait commercialement inacceptable).

Pour pallier ces inconvénients, on a recours au sondage statistique, qui consiste à déduire les propriétés de toute une population à partir de l'analyse d'un échantillon est capital que l'échantillon soit choisi et analysé de manière adéquate. En particulier, il faut que l'échantillon soit *représentatif* de la population. Un échantillon non représentatif est dit *biaisé*.

STATISTIQUES DESCRIPTIVES

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Une étude statistique a pour objet de formuler des lois valables pour tout un ensemble d'êtres ou d'éléments qui constituent une population (ex : êtres humains, animaux, plantes...). Pour des raisons pratiques et économiques, l'étude porte en fait sur un groupe restreint des membres d'une population : un échantillon.

Ex : si on s'intéresse à la longueur des tiges d'une certaine variété de blé, on sélectionne seulement un échantillon soit de manière aléatoire soit par une méthode dite des quotas, soit le plus souvent grâce à des méthodes mixtes.

L'échantillon est donc toujours plus petit que la population.

STATISTIQUES DESCRIPTIVES

Définition

Population et échantillon

Variable aléatoire

Variable continue et discrète

Série statistique

Variable continue

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

La notion de variable aléatoire

Dans une population, on s'intéresse à un ou plusieurs traits (ou caractéristiques, ou caractères) des individus.

Ex : sexe, CSP, score obtenu à un test...

Ce trait doit être présent chez chacun des membres de la population.

Cette caractéristique est appelée en statistique une *variable aléatoire (V.A.)*

Variable car elle peut prendre plusieurs valeurs (au moins 2) que l'on appelle des *modalités*

STATISTIQUES DESCRIPTIVES

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

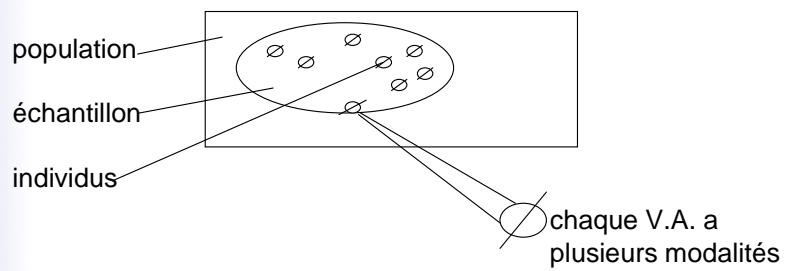
Ecart-type

Quartiles

Représentation Graphique

Ex : la variable « sexe » a deux modalités : homme ou femme.

Si on étudie les âges des personnes d'un groupe de 20 à 30 ans, cette variable a 11 modalités si on considère l'âge en années mais beaucoup plus (132) si c'est en mois ou encore en jour.



Aléatoire car les valeurs observées de la variable obéissent (en grande partie) aux lois du hasard.

STATISTIQUES DESCRIPTIVES

Définition

Population et échantillon

Variable aléatoire

Variable continue et discrète

Série statistique

Variable continue

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Ex :

population = ensemble des copies de la promotion des étudiants de 10 année de 2002-2003.

Variable (notée X) = note à l'examen de statistique prenant les modalités 0,1,...,20 (dans N)

Si je tire au hasard 10 copies dans la population de copies, j'obtiens théoriquement 10 valeurs de cette variable réparties aléatoirement à l'intérieur de toutes les notes possibles.

STATISTIQUES DESCRIPTIVES

Définition

Population et échantillon

Variable aléatoire

Variables continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Variable continue ou discontinue/discrète

Une variable est dite continue lorsqu'elle est susceptible de prendre toutes les valeurs (numériques) possibles (donc souvent un nombre infini) à l'intérieur d'un intervalle.

Dans ce cas, il est parfois nécessaire de faire des regroupements des modalités en intervalles. Pour d'éventuels calculs, nous verrons ensuite que l'on remplace souvent l'intervalle par son milieu.

Une variable est dite discrète ou discontinue lorsqu'elle ne peut prendre que des valeurs isolées.

STATISTIQUES DESCRIPTIVES

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne
Mode
Médiane

Paramètre de Dispersion

Ecart-type
Quartiles

Représentation Graphique

2. SERIES STATISTIQUES :

C'est l'ensemble des couples (modalités x_i - effectif n_i) pour une variable donnée.

Deux façons de présenter les données:

Données brutes:

Note	2	5	12	7	10	5	12	7	4	12	3	17	3	6	5
------	---	---	----	---	----	---	----	---	---	----	---	----	---	---	---

Données Ordonnées:

Note	2	3	4	5	6	7	10	12	17
effectif	1	2	1	3	1	2	1	3	1

SERIES STATISTIQUES

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Exemple sur une variable discontinue : nombre de mots rappelés (d'une liste de 5 mots) après 5 minutes de délai

Nombre de mots (6 Modalités) (x_i)	Nombre de sujets concernés Effectif partiel (n_i)	Fréquence relative	Effectif cumulé Nombre de sujets
0	1		
1	2		
2	3		
3	5		
4	5		
5	10		
	Effectif total N=		

Effectif total : Notation : N.

Fréquence relative : C'est le rapport de l'effectif partiel à l'effectif total

$$\frac{\text{Effectif de la modalité}}{N}$$

N

Permet d'obtenir des pourcentages

EXO 01:

Trouver le nombre moyen de mot rappelés par sujets:

Il faut donc trouver le nombre de mots rappelés au total : $2+6+15+20+50 = 93$ mots pour 26 sujets.

* 4% des personnes n'ont rappelé aucun mot

** 6 sujets ont rappelé 2 mots ou moins

SERIES STATISTIQUES

Définition

Population et échantillon

Variable aléatoire

Variable continue et discrète

Série statistique

Variable continue

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Exemple sur une variable continue : poids de nouveau-nés dans une maternité.

Les pesées ont été faites à 10 grammes près et ont conduit à des résultats s'échelonnant entre 2,24 kg et 4,49 kg.

Ces données ont été regroupées par intervalles (8 intervalles).

Par exemple: [2,5 ; 2,8]

Valeur centrale des classes :

(Borne supérieure de l'intervalle – borne inférieure de l'intervalle)/2 + plus petite valeur de l'intervalle.

Ex : milieu de [2.5 ; 2.8] = $(2.8 - 2.5) / 2 = 0.15 \rightarrow 0.15 + 2.5 = 2.65$

Ou plus simplement (borne supérieure de l'intervalle + borne inférieure de l'intervalle)/2

ATTENTION aux différences entre intervalles ouverts ou fermés. Si la variable est discrète il vaut mieux travailler sur des intervalles fermés.

SERIES STATISTIQUES

Définition

Population et échantillon
Variable aléatoire
Variable continue et discrète

Série statistique

Variable continue

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Trouver le centre de classe, fréquence relative et effectif cumulé

Modalités Regroupées en intervalles x_i	Effectifs partiels n_i	Centre de classe (x_i)	Fréquence relative	Effectif cumulé
[2,2 ; 2,5[5			
[2,5 ; 2,8[11			
[2,8 ; 3,1[24			
[3,1 ; 3,4[40			
[3,4 ; 3,7[42			
[3,7 ; 4,0[20			
[4,0 ; 4,3[13			
[4,3 ; 4,6[6			
	N =		Total =	

3. LES ECHELLES DE MESURE :

ECHELLES DE MESURE

Définition

Population et échantillon

Variable aléatoire

Variable continue et discrète

Série statistique

Variable continue

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

3. LES ECHELLES DE MESURE :

Lorsqu'on fait une mesure, que ce soit une mesure d'une grandeur physique ou d'une grandeur psychologique (sensation, réponse comportementale, etc...), on utilise toujours une échelle de mesure pour reporter la « valeur » de ce que l'on observe. Mais mesurer n'implique pas obligatoirement l'utilisation du système numérique.

Il existe en fait différents niveaux de mesures correspondant à différentes échelles de mesure. C'est ainsi que Stevens a proposé une distinction entre 4 types d'échelles vérifiant d'une manière plus ou moins contraignante les propriétés arithmétiques des nombres. Ce sont les échelles nominales, ordinales, d'intervalle et de rapport. Selon le type d'échelle utilisée, des opérations plus ou moins complexes pourront être effectuées sur les nombres appliqués aux objets.

ECHELLES DE MESURE

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Les échelles nominales

C'est le niveau le plus bas des échelles, l'échelle la plus simple qui apporte le moins d'information. Elle permet seulement une classification des données (échelle qualitative). Pour désigner les classes, on peut utiliser des nombres mais aussi n'importe quel autre symbole.

Exemples:

Sexe avec 2 classes (ou modalités) : femmes et hommes.

Situation familiale (célibataire, marié, concubinage, veuf, séparé, etc...)

etc...

Une échelle nominale n'a ni origine ni unité. Les nombres utilisés sur ces échelles possèdent très peu de propriétés. C'est l'échelle de mesure la moins puissante. Les nombres sont ici seulement des symboles permettant de distinguer les objets entre eux sur l'échelle. Avec de telles échelles, on ne peut donc que dénombrer les observations de chaque classe.

ECHELLES DE MESURE

Définition

Population et échantillon

Variable aléatoire

Variable continue et discrète

Série statistique

Variable continue

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Les échelles ordinales.

On atteint un niveau plus élevé de mesure si on ajoute une relation d'ordre à la propriété d'identité des échelles nominales. Les échelles ordinales possèdent donc toutes les propriétés des échelles nominales (classification des données) et permettent en plus une ordination des données. Elles possèdent une structure d'ordre.

Exemples:

Tranches d'âges (moins de 20 ans, 20-25, 26-35, 36-45, 46-55, plus de 55)

Échelle de certitude (incertain, plutôt incertain, plutôt certain, certain)

Estimation de la gêne due au bruit de la circulation sur une échelle à 5 catégories (1, 2, 3, 4, 5), 1 correspondant à pas du tout gêné, et 5 à très gêné.

ECHELLES DE MESURE

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Les échelles d'intervalle

Elles possèdent les propriétés des échelles précédentes (classification et ordre) et en plus se caractérisent par des intervalles égaux entre les différentes classes. L'échelle a donc ici une unité. Quant à l'origine, elle est définie, mais de façon arbitraire.

Le fait de pouvoir maintenant définir la distance entre 2 objets permet de calculer des différences entre des mesures. On peut donc utiliser tous les traitements numériques basés sur des différences. On peut ainsi calculer une moyenne arithmétique: c'est la valeur pour laquelle la somme des différences entre toutes les valeurs observées et cette valeur est nulle. De la même manière on a le droit de calculer l'étalement des valeurs autour de cette moyenne: c'est l'écart-type (variance).

1. Les numéros, ou nombres utilisés sur ces échelles correspondent à des rangs. Mais la distance entre les rangs n'est pas obligatoirement la même. La forme de l'histogramme des données n'a donc à priori aucune importance. On peut en effet l'aplatir ou le resserrer en modifiant les distances entre les valeurs de l'échelle. Les échelles ordinales n'ont donc ni une origine ni une unité fixe.

ECHELLES DE MESURE

Définition

Population et échantillon

Variable aléatoire

Variable continue et discrète

Série statistique

Variable continue

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Un exemple d'échelle non psychologique : l'échelle centigrade (ou Celsius) et l'échelle Fahrenheit des températures sont des échelles d'intervalle: l'écart entre 10 et 11° Celsius est le même qu'entre 15 et 16° Fahrenheit. Mais l'origine et l'unité sont arbitraires. En effet, on passe de l'une à l'autre par une simple transformation linéaire: $F = 32 + 1.8 C$.

Exemple : $0^{\circ}C = 32^{\circ}F$, et $100^{\circ}C = 212^{\circ}F$.

Ces échelles comme toutes les échelles d'intervalle n'ont donc pas d'origine fixe (il n'existe pas de zéro absolu avec Celsius et Fahrenheit). Les échelles d'intervalle sont en fait suffisantes pour avoir le droit d'effectuer tous les calculs statistiques que l'on verra par la suite.

ECHELLES DE MESURE

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Echelle de rapport:

Avec les échelles de rapport, on rajoute une propriété supplémentaire. L'origine n'est plus ici arbitraire. Cela permet de décider si 2 rapports entre des valeurs sont identiques ou non.

Exemples:

Températures en degrés Kelvin: le 0°K est le zéro absolu = -273.15°C .

Longueur, poids, surface, temps, etc ...

Lorsqu'on transforme l'unité de telles échelles, le rapport entre les différents niveaux est constant : rapport entre 2m et 1m et le même qu'entre 200 cm et 100 cm (unité arbitraire). De même, on a le droit de dire que 40 cm est 2 fois plus grand que 20 cm. En revanche, on ne peut pas dire que 20°C est 2 fois plus chaud que 10°C , car si on transforme les $^{\circ}\text{C}$ en $^{\circ}\text{F}$, on obtient 68°F et 40°F . Leur rapport n'est plus égal à 2.

Avec ce genre d'échelles, toutes les opérations arithmétiques sont possibles.

POSITION ET DISPERSION

Définition

Population et échantillon
Variable aléatoire
Variable continue et discrète

Série statistique

Variable continue
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne
Mode
Médiane

Paramètre de Dispersion

Ecart-type
Quartiles

Représentation Graphique

L'ensemble des données brutes d'une série statistique est très difficile à interpréter. Il est donc nécessaire de définir et de calculer un ensemble de valeurs caractéristiques (ou paramètres). Ces valeurs permettent de rendre compte de l'information de façon condensée et précise.

Il existe deux sortes de paramètres :

► Les paramètres de position ou de premier ordre comme la moyenne, le mode ou encore la médiane. Ces paramètres permettent de calculer les valeurs centrales et ainsi de rendre compte de l'ordre de grandeur de la série étudiée.

► Les paramètres de dispersion ou de second ordre comme la variance, l'écart type ou encore les quantiles. Ces paramètres permettent de préciser le degré de dispersion des différentes valeurs de la série étudiée autour d'une valeur centrale.

Évoquer la *distinction fondamentale* entre les valeurs théoriques des paramètres (de la population) et les valeurs empiriques (de l'échantillon)

POSITION ET DISPERSION

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Exemple de l'importance des deux paramètres:

Prenons l'exemple de 2 classes d'élèves, de même effectif. Nous voulons savoir quelle est la classe la meilleure à une épreuve de maths. Quels paramètres vont nous permettre une comparaison la plus objective?

Il faut donc d'un côté concevoir le niveau globale de chacune des classes, c'est-à-dire le niveau moyen de la classe et savoir si il y a beaucoup d'élèves qui sont proche ou éloigné de ce niveau global.

	Note par élève						Moyenne
	a	b	c	d	e	f	
CLASSE A	0	2	8	12	18	20	10
CLASSE B	10	10	10	10	10	10	10
CLASSE C	0	0	0	20	20	20	10

La moyenne est un paramètre de position, mais nous voyons bien qu'en ayant la même moyen les classes sont vraiment très différentes...

POSITION ET DISPERSION

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

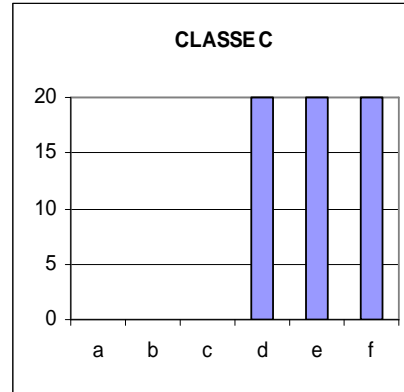
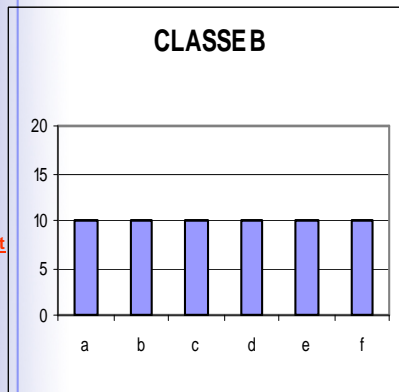
Moyenne
Mode
Médiane

Paramètre de Dispersion

Ecart-type
Quartiles

Représentation Graphique

C'est pour cela que le paramètre de dispersion est fondamental



PARAMETRE DE POSITION

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

1. LES PARAMÈTRES DE POSITION

MOYENNE :

Ce calcul est possible uniquement sur des échelles au minimum d'intervalles. On la note \bar{X}

2 formules possibles :

► Sur les données brutes : $\bar{X} = \frac{1}{N} \sum_{i=1}^N x_i$

Ex : Note en anglais sur l'année : 12,15,6, 12. =
(12+14+6+12)/4 = 11

Remarque : différencier \bar{X} = la variable et \bar{x} = une des valeurs prise par la variable

► Sur les données regroupées par modalité: $\bar{X} = \frac{1}{N} \sum_{i=1}^m n_i x_i$

Même exemple : moyenne = (1x6 + 2x12 + 1x14)/4 = 11

PARAMETRE DE POSITION

Définition

Population et échantillon

Variable aléatoire

Variable continue et discrète

Série statistique

Variable continue

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

EXO_03 :

Autre exemple : Un professeur assure l'enseignement des mathématiques dans 2 classes de 6ème différentes A et B. Lors d'un contrôle des connaissances effectué en fin de trimestre, les notes obtenues par chaque élève sont les suivantes:

Classe A : 7 - 9 - 9 - 9 - 13 - 10 - 10 - 10 - 8 - 10 - 10 - 10
- 11 - 10 - 11 - 11 - 11 - 7 - 12 - 9 - 9 - 10 - 10 - 10 -
11 - 8 - 12 - 8 - 9 - 11 - 9

Classe B : 5 - 17 - 3 - 14 - 6 - 16 - 16 - 2 - 16 - 15 - 6 - 14
- 4 - 4 - 4 - 18 - 1 - 17 - 15 - 3 - 13 - 2 - 19 - 17 - 3 - 16
- 15 - 5 - 5 - 4 - 16 - 16 - 5 - 4

Calculer les moyennes avec les 2 formules

PARAMETRE DE POSITION

Définition

Population et échantillon

Variable aléatoire

Variable continue et discrète

Série statistique

Variable continue

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

La moyenne graphique:

Si vous avez deux points dans l'espace? La moyenne de leur position est le milieu du segment qui les rejoint.

A
+

B
+

PARAMETRE DE POSITION

Définition

Population et échantillon

Variable aléatoire

Variable continue et discrète

Série statistique

Variable continue

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

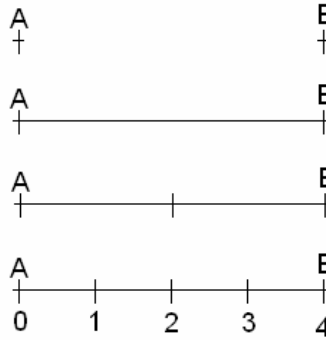
Ecart-type

Quartiles

Représentation Graphique

La moyenne graphique:

Si vous avez deux points dans l'espace? La moyenne de leur position est le milieu du segment qui les rejoint.



$$\text{Moyenne sur l'échelle: } m = (0+4)/2 = 2$$

PARAMETRE DE POSITION

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

La moyenne graphique:

Dans un espace à plusieurs dimensions et d'effectif sup.
à 2, la médiane s'exprime par le :

Exemple: espace à 2D:

PARAMETRE DE POSITION

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

B) MODE – CLASSE MODALE :

Le *mode* correspond à la valeur de la variable qui contient l'effectif le plus élevé. Il est possible de le calculer \forall le type d'échelle.

La *classe modale* correspond à la classe de la variable qui contient l'effectif le plus élevé.

Ex avec une variable nominale : Enquête sur la représentation que se fait le public de la psychologie et des psychologues.

À votre avis, pourquoi les gens vont voir un psychologue ?

	Nombre de réponses
Parce qu'ils sont dans un état anormal	8
Pour trouver de l'aide	9
Pour se connaître	8
Pour des problèmes d'orientation	10
Autres réponses	5

* Calculer N :

* Trouver le mode :

PARAMETRE DE POSITION

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Classe A : 7 - 9 - 9 - 9 - 13 - 10 - 10 - 10 - 8 - 10 - 10 - 10 -
11 - 10 - 11 - 11 - 11 - 7 - 12 - 9 - 9 - 10 - 10 - 10 - 11 - 8 -
12 - 8 - 9 - 11 - 9

Classe B : 5 - 17 - 3 - 14 - 6 - 16 - 16 - 2 - 16 - 15 - 6 - 14 -
4 - 4 - 4 - 18 - 1 - 17 - 15 - 3 - 13 - 2 - 19 - 17 - 3 - 16 - 15 -
5 - 5 - 4 - 16 - 16 - 5 - 4

Calculer le mode de chaque classe. Faites un graphique.

PARAMETRE DE POSITION

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

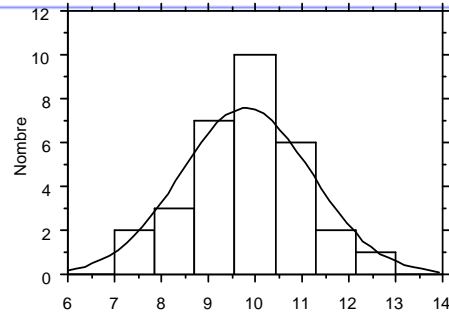
Médiane

Paramètre de Dispersion

Ecart-type

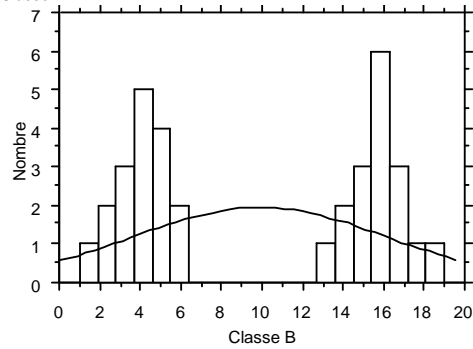
Quartiles

Représentation Graphique



Classe A

Mode Classe A :



Mode Classe B :

PARAMETRE DE POSITION

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Exercice avec des intervalles : score obtenu a un test :

Score	[2 ; 3]	[4 ; 5]	[6 ; 7]	[8 ; 9]	[10 ; 11]
Effectif	12	30	12	90	25
Σ effectifs					

Classe modale=

Mode = centre de la classe modale=

PARAMETRE DE POSITION

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

C) MÉDIANE – CLASSE MÉDIANE :

La médiane correspond à la valeur de la variable qui partage l'effectif total en deux parties égales.

Pour la trouver : on calcule $N/2$ puis on regarde dans l'effectif cumulé à partir de quelle valeur de la variable cet effectif est atteint.

Exemple :

- Soit un échantillon de 9 personnes dont le poids est :

45 – 68 – 89 – 74 – 62 – 56 – 49 – 52 – 63 kg

classés par ordre croissant :

45 – 49 – 52 – 56 – 62 – 63 – 68 – 74 – 89 kg

$\underbrace{45 - 49 - 52 - 56}_{4}$
 \uparrow
 $\underbrace{62 - 63 - 68 - 74 - 89}_{4}$

 médiane

- Si le nombre d'individus est pair, on prend la moyenne entre les deux valeurs centrales :

$\underbrace{45 - 49 - 52 - 55 - 56}_{5}$
 $\underbrace{62 - 63 - 68 - 74 - 89}_{5}$

$$\text{médiane} = \frac{56 + 62}{2} = 59 \text{ kg}$$

PARAMETRE DE POSITION

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

C) MÉDIANE – CLASSE MÉDIANE :

La médiane correspond à la valeur de la variable qui partage l'effectif total en deux parties égales.

Pour la trouver : on calcule $N/2$ puis on regarde dans l'effectif cumulé à partir de quelle valeur de la variable cet effectif est atteint.

Classe A : 7 - 9 - 9 - 9 - 13 - 10 - 10 - 10 - 8 - 10 - 10 - 10 - 11 - 10 - 11 - 11 - 11 - 7 - 12 - 9 - 9 - 10 - 10 - 10 - 11 - 8 - 12 - 8 - 9 - 11 - 9

X_i	7	8	9	10	11	12	13
n_i	2	3	7	10	6	2	1
Σn_i	2	5	12	22	28	30	31

$$N/2 = 15.5$$

La médiane est la note du 16ème sujet (15 avant et 15 après) donc 10

PARAMETRE DE POSITION

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Classe B : 5 - 17 - 3 - 14 - 6 - 16 - 16 - 2 - 16 - 15 - 6 - 14 -
4 - 4 - 4 - 18 - 1 - 17 - 15 - 3 - 13 - 2 - 19 - 17 - 3 - 16 - 15 -
5 - 5 - 4 - 16 - 16 - 5 - 4

X_i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
n_i	1	2	3	5	4	2	0	0	0	0	0	0	1	2	3	6	3	1	1	0
$\sum n_i$	1	3	6	11	15	17	17	17	17	17	17	17	18	20	23	29	32	33	34	34

$N/2 = 17$

La médiane est la note entre le 17ème et le 18ème sujet donc toutes les valeurs entre 6 et 12 peuvent être considérées comme médianes.

PARAMETRE DE POSITION

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

La classe médiane est la classe qui partage l'effectif total en deux parties égales.

Pour la trouver : $N/2$ puis on regarde dans l'effectif cumulé à partir de quel intervalle ce nombre est atteint.

Par approximation, la médiane correspond alors au centre de la classe médiane

Ex : Soit la série statistique suivante :

	Effectif	Effectif cumulé
[0 ;10[17	17
[10 ;20[15	32
[20 ;30[26	58
[30 ;40[30	88
[40 ;50[17	105
[50 ;60[5	110

* Calculer N :

* Classe modale :

* Mode :

* Classe médiane :

* Médiane par approximation :

PARAMETRE DE POSITION

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne
Mode

Médiane

Paramètre de Dispersion

Ecart-type
Quartiles

Représentation Graphique

Mais on peut aussi la calculer précisément.

Calcul précis de la médiane avec des intervalles :

$$M = a + \frac{b - a}{F(b) - F(a)} \times (i - F(a))$$

i = effectif cumulé correspondant à la médiane

a = valeur de la variable correspondant à la borne inférieure de l'intervalle

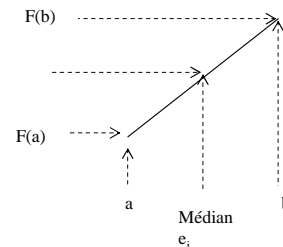
b = valeur de la variable correspondant à la borne supérieure de l'intervalle

$F(a)$ = effectif cumulé avant la borne inférieure

$F(b)$ = effectif cumulé jusqu'à la borne supérieure

Avec l'exemple :

$$M = 20 + \frac{30 - 20}{58 - 32} \times (55 - 32) = \mathbf{28.85}$$



PARAMETRE DE POSITION

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne
Mode

Médiane

Paramètre de Dispersion

Ecart-type
Quartiles

Représentation Graphique

Autre exemple :

Score	Effectif	Effectif cumulé
[15 ; 20[18	18
[20 ; 30[23	41
[30 ; 40[15	56
[40 ; 50[31	87
[50 ; 60[26	113

**Classe modale :*

** Mode :*

** Classe médiane :*

** Médiane par approximation :*

** Calcul précis :*

PARAMETRE DE DISPERSION

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

A) VARIANCE ET ÉCART TYPE

- La variance et l'écart type sont des indices de dispersion qui permettent de voir comment les données sont réparties autour de la moyenne. Ces 2 paramètres sont de même nature puisque la variance est le carré de l'écart type (ou l'écart type est la racine carré de la variance). L'écart type est mesuré dans la même unité que la variable alors que la variance est mesurée dans cette unité au carré.

La variance d'une série statistique $S(x_i, n_i)$ avec $1 \leq i \leq p$ correspond à la moyenne des carrés des écarts à la moyenne de cette série. Elle se note : var ou σ^2 . Elle est toujours ≥ 0 .

L'écart type correspond à l'écart moyen par rapport à la moyenne. Il est également toujours ≥ 0

PARAMETRE DE DISPERSION

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne
Mode
Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Écart type d'une population :
Sur les données brutes

Écart type d'une population :

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^m n_i (x_i - \bar{x})^2} = \sqrt{\frac{1}{N} \sum_{i=1}^m n_i x_i^2 - \bar{x}^2}$$

Sur les données brutes

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2 - \bar{x}^2}$$

La formule complète est rarement utilisée et dans la pratique, on lui préfère la formule de Koenig :

Attention : **seul xi est élevé au carré**

PARAMETRE DE DISPERSION

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne
Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Distinction entre échantillon et population : l'écart type d'un échantillon calculé avec cette formule est sous-estimé.

On préfère généralement utiliser la formule suivante :

Écart type (non-biaisé) d'un échantillon :

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^m n_i (x_i - \bar{x})^2} = \sqrt{\frac{1}{N-1} (\sum_{i=1}^m n_i x_i^2 - N \cdot \bar{x}^2)}$$

Ou sur les données brutes

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} = \sqrt{\frac{1}{N-1} (\sum_{i=1}^N x_i^2 - N \cdot \bar{x}^2)}$$

Mais pour plus de simplicité, cette année seul σ sera calculé

PARAMETRE DE DISPERSION

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne
Mode
Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

- Plus il y a d'individu disperser autour de la moyenne plus l'écart type est grand
- Plus les scores des individus sont éloignés autour de la moyenne plus l'écart-type est grand
- Donc l'écart-type est l'expression de la dispersion des scores et du nombre d'individu autour de la moyenne.

	échantillons					
	A1	A2	A3	A4	A5	A6
sujet A	10	12	12	14	14	18
sujet B	10	10	12	10	14	10
sujet C	10	10	10	10	10	10
sujet D	10	10	10	10	10	10
sujet E	10	10	8	10	6	10
sujet F	10	8	8	6	6	2

moyenne	10	10	10	10	10	10
sd	0	1,265	1,789	2,53	3,578	5,06

PARAMETRE DE DISPERSION

Définition

Population et échantillon
Variable aléatoire
Variable continue et discrète

Série statistique

Variable continue
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne
Mode
Médiane

Paramètre de Dispersion

Ecart-type
Quartiles

Représentation Graphique

Ex : Voici les notes des élèves de la classe C :

Notes x_i	Effectifs n_i	Centre de classe	$n_i x_i$	$n_i x_i^2$
[0 ; 2[0	1	0	0
[2 ; 4[1	3	3	9
[4 ; 6[2	5	10	50
[6 ; 8[9	7	63	441
[8 ; 10[15	9	135	1215
[10 ; 12[8	11	88	968
[12 ; 14[3	13	39	507
[14 ; 16[0	15	0	0
[16 ; 18[1	17	17	289
[18 ; 20[1	19	19	361

$$\bar{X} = 374/40 = 9,35$$

$$\sigma = \sqrt{\frac{3840}{40} - 9,35^2} = \sqrt{8,58} = 2,93$$

PARAMETRE DE DISPERSION

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne
Mode
Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

autre exemple :

Voici les notes de la classe D :

Notes x_i	Effectifs n_i	centre	$n_i x_i$	$n_i x_i^2$
[0 ;2[10			
[2 ;4[8			
[4 ;6[2			
[6 ;8[1			
[8 ;10[0			
[10 ;12[1			
[12 ;14[2			
[14 ;16[2			
[16 ;18[5			
[18 ;20[3			
Σ	34			

PARAMETRE DE DISPERSION

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne
Mode
Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

B) QUANTILES D'ORDRE α (avec $0 < \alpha < 1$)

On appelle quantile d'ordre α de la série statistique $S(x_i, n_i)$, la modalité de cette série pour laquelle l'effectif cumulé représente le quotient α de l'effectif total.

La médiane est un quantile particulier, elle correspond au quantile d'ordre 1/2 puisque la médiane désigne la valeur de la distribution pour un effectif cumulé de 50 %.

Cas particulier : quartiles, déciles, centiles

$\alpha = 1/4$ de N $\rightarrow X_{\alpha}$ est le premier quartile, noté Q1. Le premier quartile est donc la valeur de la distribution pour un effectif cumulé de 25 % (ou valeur en dessous de laquelle se situent 25% des données).

Idem pour $\alpha = 3/4$ de N \rightarrow troisième quartile (Q3).
Q2 est en fait la médiane.

Si $\alpha = i/10$ de N, avec $i \in [1 ; 9]$ $\rightarrow X_{\alpha}$ est un ième décile.

Si $\alpha = i/100$ de N, avec $i \in [1 ; 99]$ $\rightarrow X_{\alpha}$ est un ième centile.

PARAMETRE DE DISPERSION

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Exemple avec la Classe A :

X_i	7	8	9	10	11	12	13
n_i	2	3	7	10	6	2	1
Σn_i	2	5	12	22	28	30	31

$Q1 = \text{note du } 31/4 = 7.75^\circ \text{ sujet} = 9 = \mathbf{C25}$

$Q2 = \text{médiane} = \text{note du } 31/2 = 15.5^\circ \text{ sujet} = 10$

$Q3 = \text{note du } (31 \cdot 3)/4 = 23.25^\circ \text{ sujet} = 11 = \mathbf{C75}$

$D4 = \text{note du } (31 \cdot 4)/10 = 12.4^\circ \text{ sujet} = 8 = \mathbf{C40}$

$C9 = \text{note du } (31 \cdot 9)/100 = 27.9^\circ \text{ sujet} = 10$

$C25 = \text{note du } (31 \cdot 25)/100 = 7.75^\circ \text{ sujet} = 9$

$C65 = \text{note du } (31 \cdot 65)/100 = 20.15^\circ \text{ sujet} = 10$

Etc...

Remarque :

Il est possible de calculer un *intervalle interquartile*. Il regroupe 50 % de l'effectif total de la distribution et se calcule par la soustraction de $Q1$ à $Q3$.

Ex : Ici, intervalle interquartile : $Q3 - Q1 = 11 - 9 = 2$

PARAMETRE DE DISPERSION

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne
Mode
Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

En cas de distribution des données en intervalles, 2 façons de trouver les quartiles :

a) par approximation :

Score obtenu à un test de mémorisation

Age des candidats	Nombre d'erreurs	Effectif cumulé
[21 ;30]	5	5
[31 ;40]	7	12
[41 ;50]	13	25
[51 ;60]	25	50

Calculer :

Q3 :

D2 :

Q1 :

PARAMETRE DE DISPERSION

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne
Mode
Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

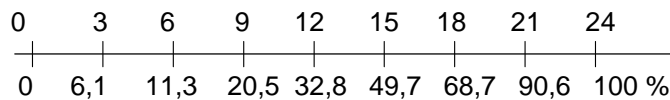
En cas de distribution des données en intervalles, 2 façons de trouver les quartiles :

a) Par calcul précis :

Ex : répartition des accidents domestiques selon les heures de la journée décrite par les données statistiques suivantes :

Horaires	% d'accident domestiques	% cumulé
0h – 3h	6,1	6,1
3h – 6h	5,2	11,3
6h – 9h	9,2	20,5
9h – 12h	12,3	32,8
12h – 15h	15,1	47,9
15h – 18h	20,8	68,7
18h – 21h	21,9	90,6
21h – 24h	9,4	100

ou :



PARAMETRE DE DISPERSION

Définition

Population et échantillon
Variable aléatoire
Variable continue et discrète

Série statistique

Variable continue
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne
Mode
Médiane

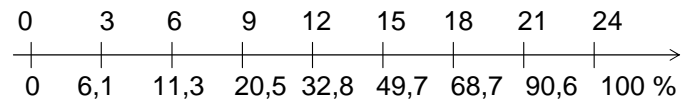
Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

ou :



$$Q_i = a + \frac{b - a}{F(b) - F(a)} \times (i - F(a))$$

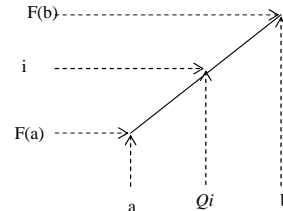
i = effectif cumulé correspondant à la Q_i

a = valeur de la variable correspondant à la borne inférieure de l'intervalle

b = valeur de la variable correspondant à la borne supérieure de l'intervalle

$F(a)$ = effectif cumulé avant la borne inférieure

$F(b)$ = effectif cumulé jusqu'à la borne supérieure



Si l'on cherche Q_1 (horaire avant lequel ont lieu 25 % des accidents domestiques), par approximation, on voit que Q_1 se situe entre 9 et 12 heures. Pour connaître la valeur précise, on applique la formule :

$$Q_{25} = Q_1 = 9 + \frac{12 - 9}{32.8 - 20.5} \times (25 - 20.5) = 10.1$$

REPRESENTATION GRAPHIQUE

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Pour des valeurs discrètes on représente souvent les distributions par des diagrammes en bâtons.

Une illustration

Lorsque les valeurs de la variable sont réparties en intervalles, on représente souvent la série statistique par un **histogramme**. L'histogramme est un ensemble de rectangles **juxtaposés** dont la base correspond à l'amplitude de chaque intervalle et dont **l'aire est proportionnelle à l'effectif ni. Donc, ce n'est pas toujours la hauteur qui est proportionnelle à la quantité représentée** lorsque l'on est confronté à une distribution où les intervalles sont inégaux.

:

REPRESENTATION GRAPHIQUE

Définition

Population et échantillon

Variable aléatoire

Variable continue et discrète

Série statistique

Variable continue

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

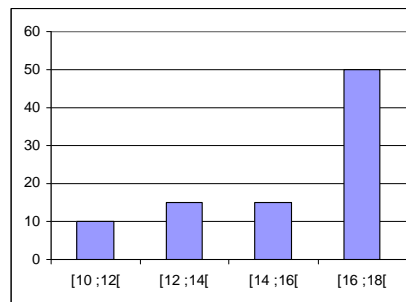
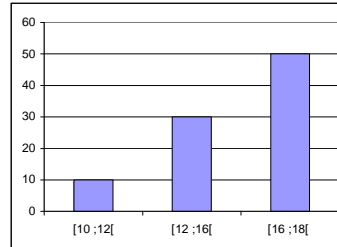
Ecart-type

Quartiles

Représentation Graphique

Ex : Argent de poche par semaine et par âge :

âge	Somme
[10 ;12[10
[12 ;16[30
[16 ;18[50



REPRESENTATION GRAPHIQUE

Définition

Population et échantillon

Variable aléatoire

Variable continue et discrète

Série statistique

Variable continue

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Pour faire un bon graphique:

- Les axes
- L'unité
- Le nom des axes
- La clarté
- Si deux graphique même échelle

REPRESENTATION GRAPHIQUE

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

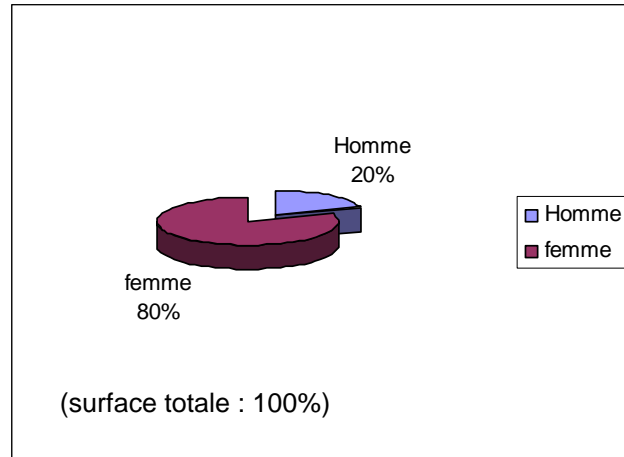
Moyenne
Mode
Médiane

Paramètre de Dispersion

Ecart-type
Quartiles

Représentation Graphique

Pour les échelles nominales, il est préférable d'employer des camemberts:



REPRESENTATION GRAPHIQUE

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

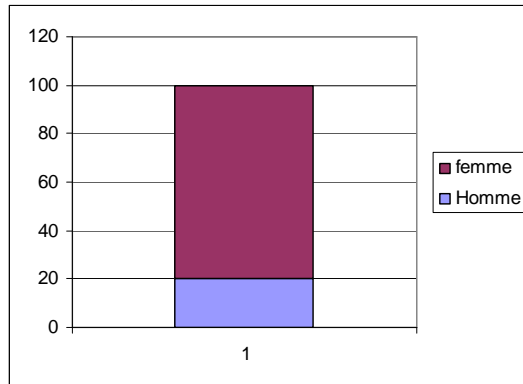
Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Diagrammes en tronçon:



(surface totale : 100%)

REPRESENTATION GRAPHIQUE

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne
Mode
Médiane

Paramètre de Dispersion

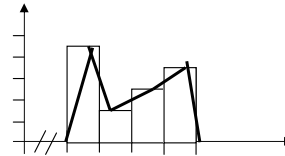
Ecart-type
Quartiles

Représentation Graphique

Dans le cas de répartition par classe d'une série, on représente la distribution par un histogramme et parfois, on relie les valeurs centrales des classes.

Polygone des effectifs

classe	effectif
[10 ;20[22
[20 ;30[8
[30 ;40[12
[40 ;50[17



REPRESENTATION GRAPHIQUE

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne
Mode
Médiane

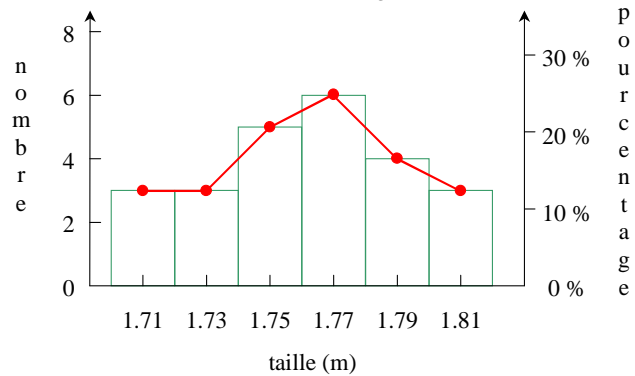
Paramètre de Dispersion

Ecart-type
Quartiles

Représentation Graphique

Polygone des effectifs et polygone des fréquences:

Pour obtenir ce polygone, on raccorde les sommets des barres, au centre de chaque classe, par des segments de droite.



On obtient donc une série de points reliés par des segments de droite. L'abscisse de chaque point correspond au centre de la classe. La hauteur de chaque point (son ordonnée) correspond au nombre d'individus dans la classe (*polygone des effectifs*) ou au pourcentage d'individus dans la classe (*polygone des fréquences*).

REPRESENTATION GRAPHIQUE

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne
Mode
Médiane

Paramètre de Dispersion

Ecart-type
Quartiles

Représentation Graphique

Généralisation au polygone des fréquences

Age des candidats	Nbre d'erreurs	fréquence
[10 ;11]	4	$\frac{4}{28} = 0,143$
[12 ;13]	8	0,286
[14 ; 15]	10	0,357
[16 ;17]	6	0,214

Tracer le polygone de fréquence :

REPRESENTATION GRAPHIQUE

Définition

Population et échantillon

Variable aléatoire

Variable continue et discrète

Série statistique

Variable continue

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

Bien interpréter les graphes

Il est courant d'entendre déclarer que l'on fait dire aux statistiques ce que l'on veut. Par exemple, il est possible de présenter les résultats de manière à amener le lecteur peu attentif à accepter une conclusion erronée.

REPRESENTATION GRAPHIQUE

Définition

Population et échantillon

Variable aléatoire

Variable continue et discrète

Série statistique

Variable continue

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

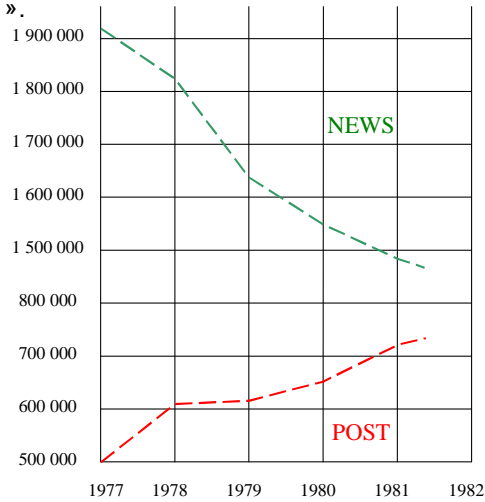
Ecart-type

Quartiles

Représentation Graphique

1. Tirage de journaux concurrents

Le graphique suivant est paru en 1981 dans le New Yorker Post, sous le titre « Ascension du Post, le quotidien préféré des New-Yorkais ».



Le but de ce graphique est de convaincre le lecteur que la croissance du tirage du Post va bientôt l'amener en première position, devant le News qui périclité.

REPRESENTATION GRAPHIQUE

Définition

Population et échantillon

Variable aléatoire

Variable continue et discrète

Série statistique

Variable continue

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique

On remarque deux artifices utilisés pour exagérer la tendance :

1. L'échelle verticale ne démarre pas en zéro.

C'est une présentation acceptable, mais qui renforce les variations apparentes.

2. L'échelle verticale est discontinue.

Alors que deux graduations successives sont séparées de 100.000 unités, on passe brutalement de 800.000 à 1.500.000 dans l'intervalle séparant le Post du News. Les tirages des deux journaux paraissent, de ce fait, beaucoup plus proches que dans la réalité.

Une telle présentation ne serait admissible que si la discontinuité de l'échelle était clairement indiquée, par exemple par des pointillés :

REPRESENTATION GRAPHIQUE

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et

de dispersion

Paramètre de Position

Moyenne

Mode

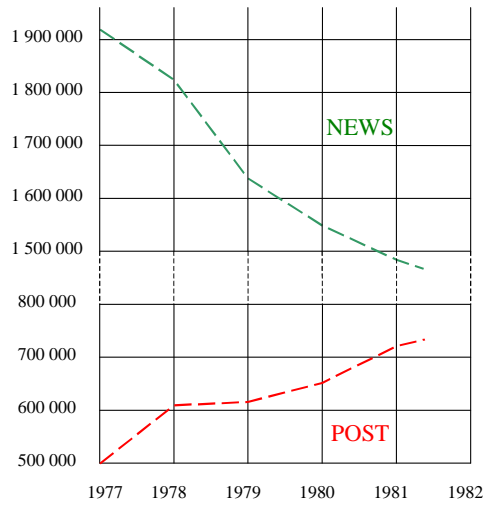
Médiane

Paramètre de Dispersion

Ecart-type

Quartiles

Représentation Graphique



REPRESENTATION GRAPHIQUE

Définition

Population et échantillon
Variable aléatoire
Variable continu et discrète

Série statistique

Variable continu
Variable discrète

Echelles de mesure

Echelles nominales
Echelles ordinales
Echelle d'intervalle
Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

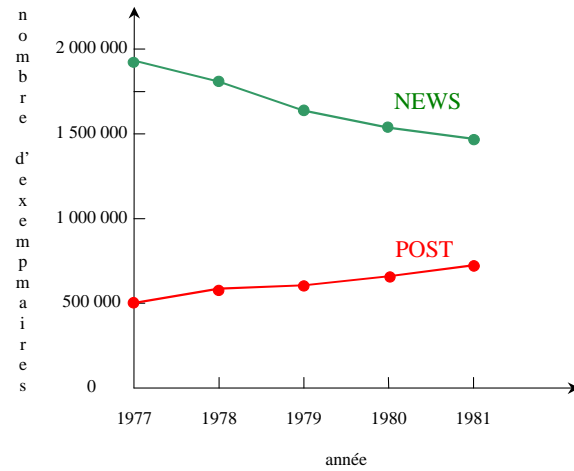
Moyenne
Mode
Médiane

Paramètre de Dispersion

Ecart-type
Quartiles

Représentation Graphique

La version correcte, plus « honnête », du graphique, est la suivante :



On constate immédiatement qu'il reste au Post bien du chemin à parcourir avant d'accéder à la première place.

REPRESENTATION GRAPHIQUE

Définition

Population et échantillon

Variable aléatoire

Variable continu et discrète

Série statistique

Variable continu

Variable discrète

Echelles de mesure

Echelles nominales

Echelles ordinales

Echelle d'intervalle

Echelle de rapport

Paramètres de positions et de dispersion

Paramètre de Position

Moyenne

Mode

Médiane

Paramètre de Dispersion

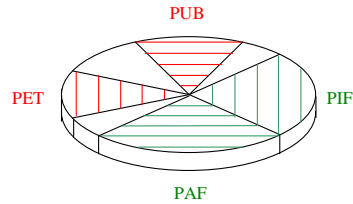
Ecart-type

Quartiles

Représentation Graphique

Le camembert en perspective

Le diagramme sectoriel suivant présente les pourcentages obtenus par 4 partis politiques lors d'une élection.



Une telle présentation en perspective a tendance à faire paraître plus importants les secteurs situés en bas (comme le PAF) ou en haut (comme le PUB) au détriment de ceux de gauche (PET) ou de droite (PIF).

Une présentation "de face" est moins susceptible d'induire le lecteur en erreur.

