

Université Saad Dahlab Blida1

Faculté de Médecine

Première Année de Médecine et Médecine Dentaire

Responsable de l'enseignement : Dr BENILHA.S



MODULE DE BIOMATHEMATIQUES / STATISTIQUES

Statistiques descriptives : Définition préliminaires (effectif, pourcentage, ratio, taux/Ensemble/Population/Echantillon/Elément/Individu/Echantillon

Dr BENILHA.S

Statistiques descriptives : Définition préliminaires (effectif, pourcentage, ratio, taux/Ensemble/Population/Echantillon/Elément/Individu/Echantillon

I-Définition de la statistique descriptive :

Ensemble des méthodes qui permettent de décrire les unités statistiques qui composent une population.

Buts de la statistique descriptive :

Toute série d'observations comporte un certain nombre de données relatives à un ou plusieurs caractères ou encore **variables**.

Le but des statistiques descriptives est de décrire un ensemble d'observations à l'aide de quelques éléments caractéristiques (tableaux, graphiques, numériques).

II-Concepts de base en statistique :

Population :

Désigne l'ensemble des entités ou des individus qui font l'objet de l'étude.

Echantillon :

Une partie (sous -ensemble) de la population dont l'effectif (la taille) est très réduit et qui permet d'effectuer l'étude expérimentale.

Individu ou unité statistique :

Unité de base sur laquelle l'observation réalise un certain nombre de mesures (une personne, un ménage, une ville,).

III- Le caractère en statistique :

Contrairement à une constante -caractéristique ayant la même valeur pour tous les individus- **une variable** comporte nécessairement plus d'une modalité, les modalités sont les différentes catégories que peut présenter une variable.

Exemple : le « **groupe sanguin dans le système ABO** » est **une variable** car il comporte les modalités : groupe O, groupe A, groupe B, groupe AB.

Donc, chaque individu d'un échantillon ou d'une population sera décrit relativement à un ou plusieurs **caractères** (on dit aussi **variables**).

Types des caractères :

1-Variables qualitatives :

Une variable qualitative est un caractère dont les modalités s'expriment par des qualités et non pas par des valeurs numériques.

On dit d'un caractère qu'il est qualitatif, si ses diverses modalités ne sont pas mesurables.

Le « **groupe sanguin** dans le système ABO », le « **sexe** » ou « genre » avec ses deux modalités : masculin et féminin sont deux variables qualitatives,

Les variables qualitatives à **deux modalités** sont souvent désignées sous le vocable de « variables **dichotomiques** » parce que de type « l'un ou l'autre ».

Tandis que les variables comportant plus de deux modalités, comme le « groupe sanguin » sont appelées **multichotomiques**.

Classification des caractères qualitatifs :

***Nominale** : Ordre n'existe pas (ex: couleur des yeux,...).

Les variables multichotomiques dont les modalités ne sont pas soumises à un tel ordre sont appelées variables **nominales**.

Exemples :

-**groupe sanguin** dans le système ABO (O, A, B, AB).

-**statut matrimonial** (célibataire, marié, divorcé, veuf).

***Ordinale** : Peuvent- être **ordonnées**

Exemples :

- Stade de la maladie au moment du diagnostic : stades I à IV pour la maladie de Hodgkin.

- Gravité des lésions engendrées par les accidents de la route : mineure, modérée, sérieuse, sévère, critique.

- Intensité de la douleur : faible, modérée, importante.

- Niveau d'instruction : analphabète, primaire, secondaire, supérieur.

2-Caractère quantitatif :

Un caractère est dit quantitatif, si ses diverses modalités s'expriment par des valeurs numériques (**mesurables**).

A chacune de ses modalités, on peut associer un nombre, ce nombre est appelé **variable statistique**.

Exemples : l'âge, le poids, la taille, la glycémie, le nombre de personnes par ménage, le temps, etc.

Classification des variables quantitatives :

***Variables discrètes** : quand la variable ne peut prendre que des valeurs isolées (nombre entier).

Exemple: nombre d'enfants par famille, nombre d'accidents du travail par ateliers, nombre de globules par unité de volume sanguin »,

***Variables continues :** Quand la variable est susceptible de prendre toute valeur appartenant à son intervalle de variation.

Une variable quantitative continue a des modalités en nombre infini qui se situent à un point quelconque d'une échelle numérique,

Exemples :

La glycémie, l'heure du jour, l'âge, la taille », la tension artérielle, le taux d'urée sanguine, le taux de cholestérol, le temps depuis l'instauration d'un traitement » ...

IV-Effectif :

L'effectif ou fréquence **absolue** est le nombre d'individus par classe.

Ce dénombrement donne lieu à une représentation des données sous forme de tableau.

Exemples :

L'effectif ou fréquence absolue est le nombre d'individus correspondant à une modalité donnée d'une variable.

Exemple 1 :

La distribution de 50 malades selon le sexe, parmi ces 50 malades, 15 sont de sexe masculin et 35 de sexe féminin, les effectifs correspondant à chacune des deux modalités sont 15 et 35.

Exemple 2 :

On a dénombré sur un ensemble de 180 sujets, les individus qui appartenaient aux différents groupes sanguins (tableau).

Tableau : Description de l'échantillon des groupes sanguins.

<i>A+</i>	<i>A-</i>	<i>B+</i>	<i>B-</i>	<i>AB+</i>	<i>AB-</i>	<i>O+</i>	<i>O-</i>
<i>80</i>	<i>10</i>	<i>20</i>	<i>5</i>	<i>5</i>	<i>2</i>	<i>50</i>	<i>8</i>

V-Fréquence relative :

On peut définir les fréquences relatives (pourcentage ou proportion) qui sont, pour chaque classe, le rapport de son effectif au nombre total d'individus de la série des mesures.

Le rapport entre l'effectif d'une modalité de variable et l'effectif total de la série sur laquelle cette variable est mesurée. Le numérateur fait obligatoirement partie du dénominateur.

La somme des fréquences relatives est égale à 1.

Parfois, les résultats sont exprimés en pourcentage, chacune des fréquences relatives étant multipliée par 100. La fréquence relative s'exprime généralement en pourcentage parce que celui-ci est plus expressif.

Tableau : fréquences relatives d'un échantillon selon le groupe sanguin

A	B	AB	O
50	14	4	32

VI-Fréquences cumulées (relatives et absolues)

Les fréquences cumulées sont utilisées pour les données ordinales qui présentent des classes ordonnées.

Exemple :

Sur un échantillon de 500 malades cancéreux, on a noté le stade de la maladie.

On peut résumer ou présenter ces données par des fréquences relatives.

Les résultats obtenus sont présentés par le tableau.

Tableau : Répartition selon le stade de la maladie.

<i>Stade</i>	<i>Nombre de malades</i>	<i>Fréquence relative (%)</i>	<i>Fréquence relative cumulée (%)</i>
1	350	70	70
2	110	22	92
3	30	6	98
4	10	2	100

Interprétation : cette présentation permet de dire, par exemple, que **92%** des sujets examinés ont un stade inférieur ou égal à 2.

VII- Ratio :

Le Ratio : est le rapport entre les effectifs de deux modalités (classes) d'une même variable. Donc le numérateur et le dénominateur sont de même nature, mais exclusifs l'un de l'autre. C'est un nombre sans unité.

Le sex-ratio : est le rapport numérique des sexes à la naissance, le sex-ratio = (effectif des garçons) / (effectif des filles).

Si le sex-ratio est égal à 1.06 (ou 106%), cela signifie que pour 100 naissances féminines, il y a 106 naissances masculines.

Lorsqu'on n'a plus affaire à des naissances, on peut parler de ratio hommes/femmes.

Exemple :

Dans une population de 100 individus, on observe 49 hommes et 51 femmes.

Quel est le ratio H/F (Sexe ratio) ? $49/51 = 0,96$

VIII -Taux :

Le Taux : est un rapport qui prend en compte la notion de temps.

C'est La probabilité de survenue d'un évènement au cours du temps.

Le numérateur(N) : les individus ayant subis un évènement pendant une période de temps déterminé

Dénominateur (D): ensemble des individus susceptibles de connaître cet évènement pendant cette période.

Au (N) figurent des individus ayant subi un événement pendant une période de temps déterminé et au (D) figure l'ensemble des individus susceptibles de connaître l'événement pendant cette période (la population à risque).

Exemple :

Dans une population de 500 personnes, on a relevé au cours de l'année 2002,74 cas d'infarctus aigue du myocarde, le taux d'incidence est donc de 148/1000.

-la cantine d'une école recevant 250 enfants a été le siège d'une toxi-infection alimentaire collective.52 enfants ont présenté l'affection.

Le taux d'attaque est de $52/250 = 208$ pour 1000.

IV-Conclusion :

La Statistiques descriptives est la science des données, c'est un ensemble de méthodes et techniques qui impliquent la collection, la classification, le résumé, l'organisation, l'analyse et l'interprétation d'une information numérique.

V - Références bibliographiques :

- Ancelle T. Statistique Épidémiologie. Édition 2002.
- Statistique appliquée aux sciences de la vie ; Valentin Rousson, édition 2017.
- Méthodes statistiques : medecine-biologie.Jean BOUYER.édition 2017
- Précis d'épidémiologie BEZZAOUCHA Abdeljalil.