

UNIVERSITE ABDELMALEK ESSAADI

F.S.J.E.S. TANGER

DEPARTEMENT ECONOMIE ET GESTION

SUPPORT DE COURS STATISTIQUE DESCRIPTIVE

Année académique 2020-2021

Semestre I

Groupes : E & F

SEANCE I

Prof. BAKKALI Imad

primadbakkali@gmail.com
imbakkali@gmail.com

PLAN DE COURS :

A. GENERALITES

I. NOTIONS STATISTIQUES.

II. TERMINOLOGIES STATISTIQUES.

III. REPRESENTATIONS GRAPHIQUES.

B CARACTERISTIQUES DE TENDANCE CENTRALE

IV. PARAMETRES DE POSTIONS

MODE.

MOYENNES.

MEDIANE.

QUARTILES, DECILES, CENTILES.

C. CARACTERISTIQUES DE DISPERSION

V. PARAMETRE DE DISPERSION

ETENDUE.

ECART ABSOLU MOYEN.

INTERVALLES ENTRE QUARTILES, DECILES, CENTILES.

MEDIALE.

VARIANCE, ECART-TYPE et COEFICIENT DE VARIATION.

A. GENERALITES

I. NOTIONS STATISTIQUES

Dans un sens général, la « statistique » est l'ensemble des méthodes scientifiques à partir desquelles sont recueillies, présentées, résumées et analysées des données. Dans un sens plus étroit, le terme de « statistique » est employé pour désigner des données ou des résultats obtenus à partir de ces données, on parle ainsi de statistiques démographiques, de statistiques sur les revenus, le chômage, etc.

La statistique est un ensemble de principes et de méthodes scientifiques pour recueillir, classer, résumer et communiquer des données numériques en vue de leur utilisation pour en tirer des conclusions et prendre des décisions.

La collecte, l'ordre et les analyses des données sont appelées statistiques. La science de l'interprétation et de la présentation des données analysées est l'objectif premier des statistiques.

Les statistiques comprennent tous les aspects de la collecte d'informations. Cela peut se faire au moyen d'enquêtes, de comptes de la population et de bien des façons.

L'objet de la statistique descriptive est de décrire, c'est-à-dire de résumer ou représenter, par des statistiques, les données disponibles quand elles sont nombreuses.

Les statistiques affectent tous les aspects de la vie moderne. Elles sous-tendent de nombreuses décisions des pouvoirs publics, des entreprises...La statistique est indispensable au développement économique.

Les statistiques descriptives sont largement utilisées dans le univers des affaires pour se procurer des données utiles. L'ensemble de données à partir duquel

l'analyse statistique est effectuée récupère beaucoup d'informations qui sont déjà connues de tout le monde, mais elles sont présentées dans un impact significatif qu'elles ont créé sur une certaine situation.

Il existe deux types importants de méthodes statistiques largement utilisées. L'une concerne les statistiques descriptives et l'autre les statistiques inférentielles.

Les statistiques descriptives et inférentielles sont les deux divisions principales de la science des statistiques. Ces deux combinaisons offrent de nombreux avantages pour identifier les besoins futurs. Les statistiques sont la base de nombreuses décisions instruites prises par un gouvernement ou une organisation.

Les statistiques descriptives et les statistiques inférentielles aident à résoudre de nombreux problèmes à résoudre. La meilleure partie des statistiques est qu'elles répondent aux besoins futurs. La description passe par certaines propriétés telles que les mesures de positions et également les mesures de dispersion. Les statistiques descriptives cèdent les informations de manière significative en utilisant des graphiques et des tableaux.

La statistique descriptive contient un ensemble de méthodes, pour décrire avec des outils appropriés des ensembles nombreux et dégager l'essentiel de l'information qui en résulte. Elle utilise des modes de représentations graphiques comme des courbes de fréquences et des histogrammes. Elle utilise également des caractéristiques obtenues par un calcul algébrique :

- Indicateurs de valeur centrale : la moyenne, la médiane, le mode.
- Indicateurs de dispersion autour d'une valeur centrale : variance, écart-type.

L'inférence statistique est l'ensemble de techniques permettant d'induire les caractéristiques d'un groupe général (la population) à partir de celles d'un groupe particulier (l'échantillon), en fournissant une mesure de la certitude de la prédiction : la probabilité d'erreur.

Le tableau ci-dessous présente une comparaison entre statistique descriptives et statistiques inférentielles :

Paramètres de comparaison	Statistique Descriptive	Statistique Inférentielle
Utilités	<ul style="list-style-type: none"> - Organise, analyse et présente les données d'une manière parfaitement significative pour plus d'investigation. -Recueillir des observations portant sur des sujets présentant une certaine propriété. 	<ul style="list-style-type: none"> -Aide à connaître la population en examinant un échantillon de celle-ci. -Explique la probabilité d'un événement futur. - Les statistiques inférentielles sont pertinentes lorsqu'il est difficile ou impossible d'examiner chaque membre d'une population entière.
Méthode	Fournit les données de manière significative grâce à des représentations graphiques, des graphiques et des tableaux ou une représentation graphique	<ul style="list-style-type: none"> -Aide à comparer et à prévoir à partir des données. -utilisent un échantillon aléatoire de données d'une population afin de décrire cette dernière et de faire des déductions à son sujet
Fonctionnalité principale	Donne normalement des données statistiques que tout le monde connaît.	Va au-delà des données. Il donne des conclusions sur la population étudiée et aide à apprendre le comportement de la population en pourcentage.
Résultats	Propose des graphiques et également des tableaux.	Offre la probabilité d'une occurrence particulière.

Source :Elaboration propre .

II. TERMINOLOGIE STATISTIQUE :

Le statisticien doit définir avec précision les termes qu'il utilise pour décrire et interpréter les données. L'emploi rigoureux et la compréhension des termes utilisés permettent d'obtenir une description fiable du phénomène étudié.

2.1 POPULATION

Une population (ou population mère) est un ensemble dont chaque élément est un individu ou une unité statistique. Les termes de population et d'individus sont employés aussi bien lorsqu'il s'agit d'un ensemble d'êtres humains : « la population résidente au Maroc », « les salariés d'une entreprise », etc., que d'un ensemble d'objets inanimés :

« La production automobile pour une année », « le stock des machines à une date donnée », et même d'ensembles abstraits ou des événements : « ensemble des jours d'une année », « la série du revenu national depuis vingt ans ». Chaque observation porte sur une unité statistique.

La population soumise à l'analyse statistique doit être définie avec précision afin que l'ensemble considéré soit déterminé sans ambiguïté de sorte qu'un individu quelconque puisse y être affecté sans incertitude.

La population est donc l'ensemble de référence sur lequel porte les observations, par exemple l'ensemble des personnes qui ont réussi le baccalauréat l'an passé, l'ensemble des automobiles en état de circuler au 1er janvier de cette année, etc.

On peut dire que La population est l'ensemble des unités statistiques définissant le champ de l'étude.

2.2 INDIVIDU OU UNITE STATISTIQUE

Une unité statistique ou « Individu » est une unité d'observation ou de mesure pour laquelle des données sont recueillies ou dérivées, c'est l'élément de base pour tabuler les données statistiques.

L'Unité statistique est donc un élément de la population, l'ensemble des individus constitue la population.

Exemples :

Citoyens d'un pays (unité statistique = citoyen)

Ensemble de la production d'une usine (unité statistique = produit).

Les entreprises de la zone franche Tanger Free Zone (unité statistique = une entreprise).

2.3 ECHANTILLON

L'échantillon est un ensemble d'individus représentatifs d'une population. L'échantillonnage vise à obtenir une meilleure connaissance d'une ou plusieurs population(s) ou sous-populations(s) par l'étude d'un nombre d'échantillons jugé statistiquement représentatif. Le recours à un plan d'échantillonnage répond en général à une contrainte pratique (manque de temps, de place, évaluation destructive d'une production, coût financier...) interdisant l'étude exhaustive de la population.

Un des objets de la statistique est d'étudier des caractères attachés à certains ensembles, qui constituent selon un terme emprunté à la démographie, une population.

Une population est constituée d'un ensemble d'individus, exemple: le parc automobile marocain. Un individu est une automobile, les caractères ou variables étudiées peuvent être la puissance, la consommation de carburant, l'âge.

Pour cela, on peut recueillir l'information sur chaque individu composant la population, c'est ce que l'on fait pour un recensement.

Cette façon exhaustive de procéder n'est pas toujours possible ou souhaitable à mettre en œuvre et on extrait l'information sur un certain nombre d'individus composant la population, c'est ce qu'on appelle prélever un échantillon ou encore faire un sondage.

A.PRELEVEMENT D'UN ECHANTILLON ET SONDAGES

Un des objets de la statistique est d'étudier des caractères attachés à certains ensembles, qui constituent selon un terme emprunté à la démographie, une population. Une population est constituée d'un ensemble d'individus. Exemple de population : le parc automobile marocain. Un individu est une automobile, les caractères ou variables étudiées peuvent être la puissance, la consommation de carburant, l'âge,...

Pour cela, on peut recueillir l'information sur chaque individu composant la population, c'est ce que l'on fait pour un recensement.

Cette façon exhaustive de procéder n'est pas toujours possible ou souhaitable à mettre en œuvre et on extrait l'information sur un certain nombre d'individus composant la population, c'est ce qu'on appelle prélever un échantillon ou encore faire un sondage.

B.LES RAISONS D'ECHANTILLONNER

Elles sont très variées, en voici quelques-unes :

- le budget est limité et le coût de la collecte élevé ;
- le manque de temps ne permet pas de recueillir l'information sur autant d'individus qu'on le souhaite ;
- les résultats sont recueillis avec plus de précision car ils sont plus facilement contrôlables du fait du nombre peu élevé d'observations ;

– le calcul d’une marge d’erreur est possible, voir l’application sur les intervalles de confiance.

C. L’INFERENCE STATISTIQUE

Les caractéristiques d’une population ne sont souvent connues qu’avec une certaine imprécision lorsque ces caractéristiques sont étudiées sur un échantillon. Les caractéristiques d’un échantillon reflètent en effet avec une certaine marge d’erreur les caractéristiques de la population.

La statistique inductive ou inférentielle consiste à induire des résultats sur une population à partir d’un échantillon en précisant si possible la marge d’erreur. Ceci fait appel au calcul des probabilités, donc à la statistique mathématique, qui n’est pas l’objet de ce support.

2.4 CARACTERE OU VARIABLE STATISTIQUE

Un individu « ou unité statistique » peut être décrit selon une ou plusieurs composantes qu’on appelle caractères ou variables statistiques.

Pour décrire une population, on classe les individus selon certains attributs que l’on appelle des caractères ou des variables. Le caractère est un critère de classement, il peut présenter plusieurs situations différentes, il prend plusieurs modalités. Les deux modalités du caractère « sexe » sont : masculin et féminin. Ce caractère qui peut prendre deux modalités est dit dichotomique, une illustration de ce type de caractère : on peut « être au chômage » ou « ne pas être au chômage ». Le nombre de modalités d’un caractère dépend de l’information disponible et du but de l’étude.

2.4.1 CARACTERE QUANTITATIF

A. LES VARIABLES QUANTITATIVES OU NUMERIQUES

Les variables quantitatives ou variables statistiques ont des attributs dont les modalités sont exprimées sous forme numérique. Une variable quantitative est soit mesurable soit repérable. À chaque unité statistique est associé un nombre : la valeur de la variable. Pour l'analyse statistique, il est habituel de distinguer les variables discrètes et les variables continues.

• VARIABLE QUANTITATIVE DISCRETE

Une variable dont les valeurs sont obtenues par dénombrement est une variable discrète. C'est par exemple le cas du nombre d'enfants. Une variable statistique est discrète ou discontinue lorsqu'elle ne peut prendre que certaines valeurs isolées – valeurs prises dans \mathbb{N} plus rarement dans \mathbb{Z} . C'est le cas du nombre de personnes qui composent un ménage. Un caractère discret peut prendre une infinité de valeurs dénombrables, il peut aussi n'en prendre que quelques-unes : le nombre d'enfants par familles qui est nécessairement un entier fini.

Certaines variables discrètes, comme le nombre de salariés d'une entreprise, pouvant prendre un très grand nombre de valeurs à l'intérieur d'un intervalle de grande amplitude, elles seront traitées comme des variables continues.

• VARIABLE QUANTITATIVE CONTINUE

Lorsque la variable peut prendre toutes les valeurs à l'intérieur d'un intervalle, la variable est dite quantitative continue par exemple la taille d'un individu, le revenu par habitant. Le nombre de modalités possibles est alors infini. La taille d'un individu est une variable continue, les revenus sont considérés comme continus, ce qui n'est pas tout à fait juste puisqu'ils ne peuvent prendre que des valeurs exprimées en centimes Les unités statistiques prenant sur ce type de

variable un nombre très important de valeurs, il est nécessaire que les valeurs de la variable soient regroupées en classes.

Pour obtenir un nombre fini de modalités, les valeurs sont regroupées en classe. Les valeurs d'une variable continue sont mesurables ou repérables, avec un degré de précision déterminé qui n'est pas toujours connu pour les données économiques et sociales.

• DISTINCTION ENTRE VARIABLE QUANTITATIVE DISCRETE ET CONTINUE

La précision des instruments de mesure ou le phénomène étudié ne permettent pas de mesurer une variable avec une extrême précision et une variable continue peut donner l'impression d'être une variable discrète. Par exemple, la taille des individus sera donnée

en centimètres sans décimales.

Dans ce cas, les observations sont des valeurs isolées, il ne s'agit pas pour autant d'un caractère discret car le phénomène étudié varie continûment.

2.4.2 CARACTERE QUALITATIF

Un caractère est qualitatif si chaque modalité ne peut être mesurée ou repérée par un nombre. Toutefois, il est toujours possible de réaliser un codage qui permet de repérer une modalité par un nombre. Sur des variables qualitatives codées, les opérations algébriques n'ont généralement pas de sens.

• CARACTERE QUALITATIF NOMINAL

Les caractères qualitatifs autrement appelés variables nominales ou variables catégorielles ont des attributs dont les différentes modalités ne sont ni mesurables ni repérables. Elles sont constatées. Avec l'usage de l'informatique, on utilise parfois le terme de données qualitatives. Le caractère se subdivise en catégories ou en modalités de la variable auxquelles seront attachés un effectif et une fréquence. C'est le cas pour le sexe, l'état matrimonial, la qualification

professionnelle. Ce sont des noms ou des étiquettes permettant d'identifier une caractéristique de chaque élément. Même s'il n'est pas toujours possible d'y établir un ordre. La présentation des modalités de la variable ne présuppose aucun classement.

•CARACTERE QUALITATIF ORDINAL

Certaines variables appellent naturellement un ordre dans le rangement de leurs catégories comme le niveau de formation. Elles sont repérables selon un type d'échelle plus ou moins légitime. Un caractère ordinal est un caractère qualitatif dans lequel les modalités possibles peuvent être classées dans un ordre spécifique ou dans un ordre naturel quelconque. Les catégories pourront alors donner lieu à un codage par les rangs qui ouvrira une autre gamme de traitements possibles proches de ceux des variables quantitatives.

C'est un caractère qui peut être ordonné, par exemple : le niveau de satisfaction d'un service ferroviaire : pas satisfait, assez satisfait, très satisfait... Dans le cas contraire le caractère qualitatif est nominal, comme par exemple : l'état matrimonial, la couleur des yeux, etc.

Exemple :

Exemples	Type de variable
Rémunérations des salariés dans une entreprise	Variable quantitative continue
Indicateur du moral des ménages	Variable qualitative ordonnée
Écart de rémunération entre hommes et femmes	Variable quantitative continue
Les pays de l'Union Africaine	Caractère qualitatif
Les niveaux de formation des salariés	Variable qualitative ordonnée
Les formes de conventions d'un programme	Caractère qualitatif
Taux de croissance du PIB	Variable quantitative
Prix à la consommation	Variable quantitative
Solde commercial	Variable quantitative