

STATISTIQUE I

Despina Baskiotis

EISTI

12 décembre 2009

Généralités

Les STATISTIQUES : Ensemble des résultats obtenus à la suite de nombreuses observations d'un même phénomène, présentées le plus souvent sous forme d'un tableau ou d'un graphique.

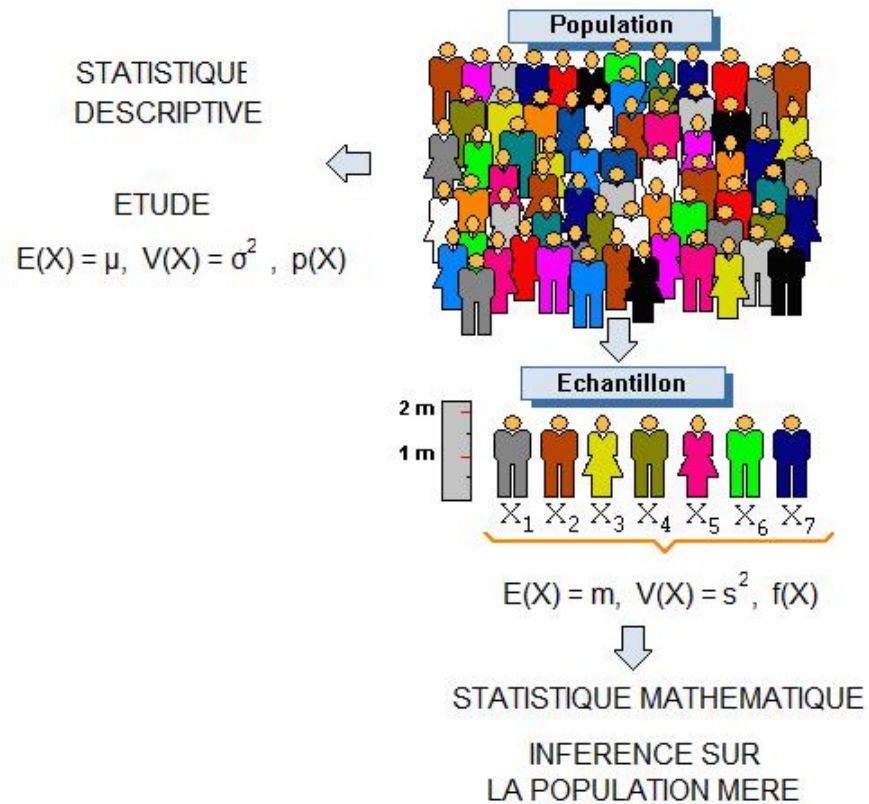
La STATISTIQUE : Ensemble de moyens et méthodes scientifiques, qui permettent d'analyser quantitativement l'information c.a.d. les "STATISTIQUES".

Remarques : Les erreurs faites lors d'une expérience sur un petit nombre d'essais ont beaucoup d'incidences sur les conclusions.

La STATISTIQUE :

- **Statistique descriptive** : Son objectif est de décrire et de résumer de façon formalisée des informations collectées de façon exhaustive.
- **La Statistique mathématique** : Elle permet, grâce aux méthodes mathématiques, l'étude d'une population en se basant sur une partie de cette population.
- **Le Calcul des Probabilités** : C'est l'outil mathématique.

ÉCHANTILLONNAGE



Les frontières entre ces trois parties restent très floues : l'étude des lois de probabilités déborde sur la statistique mathématique et cette dernière prend ses bases dans certains domaines de la statistique descriptive.

Résumons

- **PROBABILITÉ** : On connaît le modèle probabiliste, (théorie)
- **STATISTIQUE** : On veut déterminer le modèle à partir des observations

exemple :

On lance $n = 9$ fois une pièce.

Question : Quelle est la probabilité d'obtenir "PPPPFFFFPF" (I)

Réponse : :

- Si p est connu, on a $p^4(1 - p)^5$ et donc on peut la calculer.
- Si p est inconnu, on va essayer de l'estimer.

c.a.d : On fera une estimation à partir de (I), et on trouve $\hat{p} = \frac{4}{9}$.

TERMINOLOGIE

- **Population mère** : La population étudiée.
- **Échantillon** : Les individus extraits de la population mère.
Les calculs utilisant l'échantillon sont des calculs **certain**s.
- **Inference Statistique** : C'est l'extrapolation des résultats à la population mère.
C'est de nature probabiliste.

- **Variable ou Caractère** : la caractéristique observée.
Aspect particulier commun à tous les individus d'une population.
 - **Variables quantitatives** : les caractères mesurables.
 - **Qualitatives** : les variables non mesurables.
- **Modalités d'un caractère** : les différentes valeurs prises par la variable.
Les modalités permettent le classement des individus de la population.

- **Distribution statistique ou série statistique** : les modalités possibles d'une variable, complétée, pour chacune, par l'effectif correspondant.
- **Fréquence** : proportion obtenue en divisant l'effectif relatif à une modalité par l'effectif total.
- **Effectif cumulé** : (correspondant à la modalité X_p) : C'est la somme des effectifs de toutes les modalités précédentes (y compris l'effectif de la modalité X_p).

- **Expérience** : Une opération faite sous des conditions contrôlées ayant comme but de découvrir ou vérifier des événements.
- **Recensement** : Examen de tous les individus.
- **Sondage** : prélèvement de n individus \Rightarrow Échantillon de taille n

Objectif de la Statistique Mathématique :

Recueillir et Analyser des données

Le travail du statisticien :

1. Conception : Définir l'objet étudié, le questionnaire, les variables.
2. Collecte : Recueil des données.
3. Apurement : Vérification et élimination des observations aberrantes.
4. Modélisation : Construction du modèle probabiliste à aide des outils mathématiques.

La Statistique Mathématique :

Les méthodes dont l'objectif est de faire de l' **Inference** sur un modèle statistique à partir d'un échantillon.

Ces méthodes se situent en aval dans le traitement de données et reposent sur des hypothèses probabilistes.

Procédure suivie :

1. Spécification du modèle probabiliste
 - Détermination du modèle
 - Estimation des ses paramètres
 - Validation du modèle
2. Mise en œuvre des techniques adaptées :
 - Économétrie
 - Analyse de la variance
 - Analyse des données
 - ...

Plan du cours :

1. Rappel des résultats essentiels des probabilités
2. Statistique paramétrique
 - Échantillonnage
 - Estimation statistique : ponctuelle et par I.D.C.
 - Théorie de décision
 - Analyse de la variance
3. Statistique non paramétrique
 - Régression simple et multiple
 - Traitement des données
 - Analyse des données
 - Structuration des données

But du cours

Permettre à l'élève de voir les méthodes et les bases mathématiques de la statistique.

Ce cours est une simple initiation.

Les activités du statisticien

Il trouve sa place dans :

- la Production industrielle
- la Conception des systèmes
- l'Économie, la Finance
- la Gestion et Administration, les Assurances
- les Sciences (en général)
- le Contrôle de qualité et Ingénierie
- la Démographie
- l'Environnement et Météorologie
- la Médecine sociale et préventive
- l'Épidémiologie
- l'Industrie pharmaceutique et Agro-alimentaire

Connaissances requises :

Cours de probabilité 1ère année, c.a.d :

- Espace probabilisé
- Variables aléatoires.
- Distribution d'une v.a.
- Fonctions de densité
- Fonctions de répartition
- Variance, covariance et espérance d'une v.a.
- les lois de probabilité

BIBLIOGRAPHIE

- Polycopié de Statistique
- Probabilités -exercices corrigés : D.Ghorbanzadeh
- Probability and Random Processes : G. Grimmett
- One Thousand Exercises in probability : G. Grimmett
- Initiation aux probabilités : Ross