

Démarche Statistique 1

Introduction

Pierre Neuvial, <http://stat.genopole.cnrs.fr/~pneuvial>
Evry, M1 SGO, automne 2014



Remarques préliminaires

Organisation du module

- Cours: Pierre Neuvial
- TD/TP: Morgane Pierre-Jean et Guillem Rigail
 - 2 groupes: GBI et tronc commun
 - utilisation du logiciel R
- Petites interrogations écrites en début de CM
- Devoir surveillé (2h le 21/11); correction le 24/11
- Partiel (2h le 19/12)

Rappel: cours et TD/TP sont obligatoires !

Plan indicatif

- Introduction
- Statistiques descriptives
- Echantillonnage
- Introduction aux tests
- Tests d'égalité
- Tests de comparaison: échantillons appariés
- Tests de comparaison: échantillons non appariés
- Liaison entre deux variables

Références

- R. Clifford Blair & Richard A. Taylor (2008).
Biostatistics for the health sciences.
Pearson Prentice Hall.
- Daudin, Jean-Jacques, Robin, Stéphane, & Vuillet, Colette (1999).
Statistique inférentielle: idées, démarches, exemples.
Presses Universitaires de Rennes.
- Online Statistics Education: An Interactive Multimedia Course of Study.
David Lane, Rice University.
- Prum, Bernard.
La démarche statistique.
Éditions Cépaduès.

Introduction

Définitions

On s'intéresse à des **unités** d'observation (individus ou choses) dont on observe une **caractéristique**

Population

Ensemble des unités

→ *les Français*

Population statistique

Caractéristique de la population

→ *la taille des Français*

Paramètre

Résumé de la caractéristique dans la population

→ *la taille moyenne des Français*

Échantillon

Sous-ensemble d'unités

→ *les étudiants de M1 SGO 2013-2014*

Échantillon statistique

Caractéristique de l'échantillon

→ *la taille des étudiants de M1 SGO 2013-2014*

Estimateur (aussi appelé *statistique*)

Résumé de la caractéristique dans l'échantillon

→ *la taille moyenne des M1 SGO 2013-2014*

Pourquoi des statistiques ?

On ne peut généralement pas mesurer la caractéristique d'intérêt chez tous les individus de la population !

- la population peut être de taille **infinie**
- les mesures ont un **coût**
- la mesure détruit parfois l'objet mesuré

Données et variables

- La mesure d'une caractéristique chez une unité s'appelle une **donnée**
- Une caractéristique pouvant prendre plusieurs valeurs s'appelle une **variable**

Statistique descriptive et inférentielle

Statistique descriptive

Résumer l'information contenue dans un jeu de données:

- par des nombres (moyenne, médiane, écart-type)
- par des graphiques (histogramme, boxplot)

Statistique inférentielle

Utiliser les données relatives à un échantillon pour tirer des conclusions sur l'ensemble de la population:

- estimation de paramètres
- tests d'hypothèse