

TP2: Statistiques descriptives bivariées

Le Chenadec Gilles

September 7, 2017

1 Identité (1 point :-)

NOM :

PRENOM :

Si binôme, il y a (ni trinôme ni plus-nôme)

NOM :

PRENOM :

2 Informations préliminaires (mais néanmoins importantes)

- Le TP est noté.
- Il vous est demandé de commencer le compte-rendu de ce TP en séance et de le déposer dans Moodle, 1 semaine jour pour jour après la séance (si la séance a lieu, le lundi 7 de 8h10 à 12h15, la date limite est le lundi 14 à 12h15, un rendu après est à ranger dans la catégorie "risque").
- Vous devez déposer dans Moodle ce fichier Word dans lequel vous reporterez tous vos résultats:
 - réponses textuelles aux questions,
 - si vous avez besoin de mettre des équations :
 - * word est équipé d'un éditeur d'équations
 - * utilisez l'appareil photo de votre smartphone si vous préférez le papier crayon
 - codes matlab (ctrl-C dans l'éditeur matlab, ctrl-V dans Word),
 - figures (menu Edit d'une figure Matlab puis copy figure puis ctrl-V dans Word) avec indication des abscisses et ordonnées,
 - valeurs numériques des résultats des calculs, etc.
- Il est très fortement conseillé :
 - de prendre le clavier à tour de rôle pour progresser
 - de justifier la réponse aux questions posées
 - de ne pas plagier les phrases ni du cours, ni des camarades (l'école est dotée du logiciel anti-plagiat compilatio qui compare les documents rendus par les étudiants et les documents disponibles sur internet).
- Pour vous aider :
 - Les enseignants sont là à votre disposition
 - Cours Matlab sur Moodle :
 - * **cours**
 - * **fichier**

3 Statistiques Descriptives bvariées sur des données d'Iris

Une fois réalisés les graphiques pour chaque variable prise séparément, l'étude peut porter sur la relation entre deux variables. On parle de croisement de deux variables ou d'étude bvariée. L'analyse bvariée a pour but de comprendre la relation entre deux variables. Les moyens d'analyse diffèrent selon la nature des variables considérées.

3.1 Etude de la longueur du pétale en fonction de la largeur du pétale

Représentation graphique La représentation graphique liant deux variables quantitatives est le nuage de points ou l'histogramme 2D.

- 1) Tracer le nuage de points de la longueur du pétale en fonction de la largeur du pétale pour les 150 iris contenus dans les données (plot, scatter ou scatterhist). Ne pas oublier de mettre des titres sur les axes. Décrire le nuage de points.
- 2) Tracer l'histogramme bidimensionnel associé (hist3). A quoi correspond chaque baton ?

Coefficient de corrélation et droite de régression

- 3) Tracer le nuage de points de la longueur du pétale en fonction de la largeur du pétale pour les 150 iris contenus dans les données
- 4) Rappeler la définition de coefficient de corrélation et calculer le par la fonction (corrcoeff).
- 5) Donner l'équation de la droite de régression linéaire, créer une fonction permettant de la calculer à partir de deux variables X et Y et tracer la sur le même graphique
- 6) Analyser le lien entre les deux variables.

3.2 Etude de la longueur de pétale selon les différentes espèces

Nous abordons ici l'analyse des liens entre une variable qualitative et une variable quantitative. Les représentations graphiques permettant de lier une variable qualitative et une variable quantitative sont l'histogramme et la boîte à moustaches.

Représentations graphiques

- 7) Représenter sur une même figure l'histogramme par espèce (trois histogrammes avec une couleur pour chaque espèce)
- 8) Représenter une boîte à moustache par espèce

Mesures de corrélation

- 9) Calculer le rapport de corrélation lié à la décomposition de la variance en variance intra-classe et interclasse. Qu'en concluez-vous?