
TP 4 – Statistiques bivariées

Exercice 1

Données : tabac.xls

Question 1 :

- 1) partir du caractère « nombre de cigarettes », créer une nouvelle colonne prenant la valeur « fumeur » si le nombre de cigarettes est strictement supérieur à 0 et « non fumeur » sinon.
- 2) Etablir le tableau de contingence des effectifs entre « problèmes pulmonaires » et « fumeurs ».
- 3) Quel pourcentage de l'échantillon est fumeur ? Quel pourcentage de l'échantillon présente des problèmes pulmonaires ? Quel pourcentage de l'échantillon est non fumeur et avec des troubles pulmonaires ?

Question 2 :

- 1) Etablir le profil ligne.
- 2) Parmi les personnes présentant un trouble pulmonaire, quel pourcentage est fumeur ?
- 3) Comparer avec les fréquences marginales. Calculer un indice vous permettant de confirmer vos impressions.

Exercice 2

Données : herbicide.xls

Question 1 :

- 1) Représenter la moyenne du taux de plants survivants en fonction de l'herbicide et de la plante.
- 2) Quel herbicide vous semble le plus efficace et quel est celui qui est le plus néfaste ?
- 3) Pouvez-vous vous faire une idée a priori des rapports de corrélation entre le taux et la plante puis entre le taux et l'herbicide. Calculer ces rapports et vérifier.

Question 2 :

- 1) Dans le cas où les plantes n'ont subi aucun traitement, établir le lien entre le nombre de plants (x) et le nombre de plants survivants (y). En déduire le nombre de plants survivants dans le cas où le nombre de plants est 100.

- 2) Quel pourcentage de la variance de y est expliquée par la droite de régression ? Conclusion.

Question 3 :

Dans le cas de l'herbicide 1, peut-on établir un lien entre nombre de plants et nombre de plants survivants ? Expliquer pourquoi. Que faudrait-il faire pour déterminer ce lien ?