

# Analyse bivariée (2)

STA101 TD2

2023-2024

## Exercice 1

On étudie le mécanisme de reconnaissance des mots chez les enfants d'école primaire. On présente à l'écran une suite de lettres qui forment soit un mot soit un non-mot. On demande à l'élève de décider le plus rapidement possible si ce qui est à l'écran est un mot ou un non-mot. A l'issue de l'expérience on mesure les temps de réaction et le pourcentage d'erreur. Dans cet exercice, on se contentera d'étudier les temps de réaction. Les temps de réaction mesurés (en ms) sont reportés dans le tableau ci dessous

CP	3139	2972	3055	3098
CE1	1564	1499	1518	1429
CE2	1135	1029	1032	1021
CM1	748	884	769	794
CM2	663	743	775	683

On donne  $s_{CP} = 61.9$ ,  $s_{CE1} = 48.6$ ,  $s_{CE2} = 46.8$ ,  $s_{CM1} = 51.8$ ,  $s_{CM2} = 45.0$

1. Représenter les données graphiquement
2. Mesurer le lien entre le temps de réaction et le niveau de l'élève
3. Comment pourrait-on tester si le lien apparent est significatif?
4. Conclure quant au lien entre les deux variables

On donne le quantile de la loi de Fisher d'ordre 0.95 à 4 et 15 ddl : 3.06

## Exercice 2

On a relevé la couleur des yeux  $X$  et des cheveux  $Y$  de  $n = 592$  femmes. Les résultats sont donnés dans le tableau de contingence suivant ventilant la population des 592 femmes suivant leur couleurs des yeux et des cheveux

	brun	chatain	roux	blond	total
marron	68	119	26	7	220
noisette	15	54	14	10	93
vert	5	29	14	16	64
bleu	20	84	17	94	215
total	108	286	71	127	592

1. Etablir le tableau des profils-lignes et des profils-colonnes
2. Déterminer les effectifs attendus sous l'hypothèse d'indépendance
3. Calculer le critère du  $\chi^2$  entre  $X$  et  $Y$ . Comment peut-on interpréter cette valeur?
4. Calculer le  $C$  de Cramer
5. Conclure quant au lien entre les deux variables

On donne le quantile de la loi du chi-deux d'ordre 0.95 à 9 ddl : 16.9

**Exercice 3.** Une entreprise  $P$  est composée de deux établissements  $P_1$  et  $P_2$ . Le tableau ci-dessous donne la répartition des salaires par catégorie socio-professionnelle. Les salaires sont exprimés en milliers d'euros :

Catégories	$P_1$		$P_2$	
	Classe ( $x_1$ )	Effectif ( $n_1$ )	Classe ( $x_2$ )	Effectif ( $n_2$ )
Ouvriers	[20,25[	60	[15,20[	5
Employés	[30,35[	95	[25,60[	15
Cadres	[80,100[	5	[70,80[	30

Pour répondre aux différentes questions, on considèrera que tous les individus d'une même classe ont un salaire identique, égal au centre de la classe.

1. Calculer le salaire moyen dans l'entreprise
2. Calculer la variance du salaire au sein de l'entreprise.
3. Calculer les moyennes et les variances pour chacun des établissements
4. Calculer la variance inter-établissement
5. Calculer la variance intra-établissement
6. Vérifier que la somme des variance intra et inter est égale à la variance totale
7. Calculer le rapport de corrélation entre les variables établissement et salaire

NB : on rappelle les formules de calcul de moyenne et variance empirique d'une série  $(x_1, \dots, x_n)$  dans le cas où chaque observation est affectée d'un poids  $p_i$  ( $1 \leq i \leq n$ ) (la somme des poids étant égale à 1):

$$\bar{x} = \sum_{i=1}^n p_i x_i \text{ (moyenne), } s_x^2 = \sum_{i=1}^n p_i (x_i - \bar{x})^2 \text{ (variance)}$$