

STATISTIQUE DESCRIPTIVE BIVARIÉE - TD

EXERCICE 1 (*Data CO2 du logiciel R*)

On cherche à étudier la quantité de CO₂ absorbée par la plante *Echinochloa crus-galli* dans diverses conditions. Une expérience est réalisée conjointement au Québec et dans l'état du Mississippi aux Etats-Unis sur différents spécimens de la plante. Pour une partie des spécimens, un traitement est appliqué toutes les nuits. On mesure la concentration de CO₂ dans l'air ainsi que le taux de CO₂ absorbé par la plante. Les résultats de l'expérience sont résumées dans le tableau ci dessous.

Specimen	1	2	3	4	5	6	7	8
Lieu	Quebec							
Traitement	Non	Non	Non	Non	Non	Non	Oui	Oui
Concentration	95	175	250	350	500	675	95	175
Absorption	16.0	30.4	34.8	37.2	35.3	39.2	14.2	24.1
Specimen	9	10	11	12	13	14	15	16
Lieu	Quebec	Missi						
Traitement	Oui	Non						
Concentration	250	95	175	250	350	500	675	1000
Absorption	30.3	10.6	19.2	26.2	30.0	30.9	32.4	35.5
Specimen	17	18	19	20	21	22	23	24
Lieu	Missi							
Traitement	Non	Oui						
Concentration	95	95	175	250	350	500	675	1000
Absorption	12.0	10.5	14.9	18.1	18.9	19.5	22.2	21.9

- Déterminer la population étudiée, l'unité statistique, les variables et leurs types.

Variables Lieu et Traitement

- On travaille d'abord sur la distribution jointe des deux variables.
 - Donner le tableau de contingence en effectifs de la distribution jointe de ces deux variables.
 - Combien y a-t-il de plantes traitées au Québec ?
 - Combien y a-t-il de plantes dans le Mississippi ?
 - Combien y a-t-il de plantes non traitées ?
 - Donner le tableau de la distribution jointe en fréquence de ces deux variables.
 - Quelle est la proportion de plantes non traitées et dans le Mississippi ?
- Quelle est la distribution marginale de la variable Lieu ? Quelle est la proportion de plantes qui sont au Québec ?
- On s'intéresse maintenant aux distributions conditionnelles
 - Donner les tableaux des distributions conditionnelles en fréquences de la variable Lieu, conditionnellement à chaque modalité de la variable Traitement.

- (b) Donner les tableaux des distributions conditionnelles en fréquences de la variable Traitement, conditionnellement à chaque modalité de la variable Lieu.
- (c) Quelle est la proportion de plantes traitées sachant que les plantes sont au Quebec ?
- (d) Quelle est la proportion de plantes du Mississippi sachant qu'elles sont traitées ?

Variables Lieu et Concentration

5. On travaille d'abord sur la distribution jointe des deux variables.
 - (a) Donner le tableau de contingence en fréquences de la distribution jointe de ces deux variables. On considerera la variable Concentration comme une variable quantitative discrète.
 - (b) Quelle est la proportion de plantes au Quebec et dans une concentration ambiante de 500 mL/L ?
 - (c) Donner les distributions marginales de la variable Lieu et de la variable Concentration.
 - (d) Quelle est la proportion de plantes dans un environnement de 500 mL/L ?
6. On s'intéresse maintenant aux distributions conditionnelles.
 - (a) Donner les tableaux des distributions conditionnelles en fréquences de la variable Lieu, conditionnellement à chaque modalité de la variable Concentration.
 - (b) Donner les tableaux des distributions conditionnelles en fréquences de la variable Concentration, conditionnellement à chaque modalité de la variable Lieu.
 - (c) Quelle est la proportion de plantes venant du Quebec sachant qu'elles ont bénéficié d'une concentration de 250 mL/L ?
 - (d) Quelle est la proportion de plantes ayant bénéficié d'une concentration de 350 mL/L sachant qu'elles sont au Mississippi ?

Variables Traitement et Absorption

7. On travaille d'abord sur la distribution jointe des deux variables.
 - (a) Donner le tableau de contingence en fréquences de la distribution jointe de ces deux variables. On considerera des classes pour la variable Absorption : [10; 20[, [20; 30[et [30; 40[.
 - (b) Quelle est la proportion de plantes ayant eu un traitement et une absorption comprise entre 20 et 30 μL ?
 - (c) Donner les distributions marginales de chaque variable.
8. On s'intéresse maintenant aux distributions conditionnelles.
 - (a) Donner les tableaux des distributions conditionnelles en fréquences de la variable Traitement, conditionnellement à chaque modalité de la variable Absorption.
 - (b) Donner les tableaux des distributions conditionnelles en fréquences de la variable Absorption, conditionnellement à chaque modalité de la variable Traitement.
 - (c) Quelle est la proportion de plantes traitées sachant qu'elles ont absorbé entre 10 et 20 μL ?
 - (d) Quelle est la proportion de plantes ayant absorbé entre 30 et 40 μL sachant qu'elles n'ont pas été traitées ?
9. Calculer les quartiles des distributions conditionnelles de la variable Absorption conditionnellement à la variable Traitement.

10. Représenter graphiquement les distributions conditionnelles à l'aide de boîtes à moustaches.
11. Calculer les moyennes et variances conditionnelles de cette distribution. Commenter.
12. Calculer la moyenne et la variance marginale de la variable Absorption. Commenter.
13. Faire le lien entre les moyennes marginale et conditionnelles et entre les variances marginale et conditionnelles.
14. Calculer la variance inter-groupes et la variance intra-groupes de la variable Absorption selon les groupes de Traitement. Commenter.
15. Retrouver la formule de décomposition de la variance.
16. Calculer le rapport de corrélation de la variable Absorption par rapport à la variable Traitement. Interpréter.
17. Indicateur de Fisher
 - (a) Calculer l'indicateur de Fisher de Y par rapport à X , Interpréter
 - (b) On donne le seuil de significativité $s_{5\%}(24, 2) = 4.35$, conclure.
 - (c) Le logiciel R, donne une $p - \text{valeur}$ strictement inférieure à 5%, conclure.

EXERCICE 2

Dans un hôpital parisien, on a prélevé la pression artérielle sur 30 patients qui rendent visite à un généraliste le lundi matin. La pression est divisée en deux temps : la pression artérielle systolique qui est la valeur de la pression artérielle au moment de la contraction ventriculaire et la pression artérielle diastolique qui est la valeur de la pression artérielle au moment de la relaxation ventriculaire. Les observations sont contenues dans le tableau suivant (X représente la pression systolique et Y la pression diastolique (en mmHg)).

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	12	11	12	13	14	15	11	12	11	11	15	11	11	13	11
Y	7	8	8	6	7	7	8	6	8	8	7	7	8	7	8
<hr/>															
i	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
X	12	15	11	12	11	12	13	12	11	11	12	11	11	11	14
Y	6	7	6	7	8	7	7	8	7	9	7	7	8	7	9

1. Déterminer la population étudiée, l'unité statistique, les variables et leurs types.
2. Distributions jointe et marginale
 - (a) Donner les tableaux de contingence en effectifs et en fréquences de la distribution jointe de ces deux variables.
 - (b) Déterminer les distributions marginales de X et Y en effectifs et en fréquences.
 - (c) Calculer les moyennes et variances marginales de X et Y .
 - (d) Quel est le pourcentage de personnes ayant la mesure de pression 11-8 ?
 - (e) Quel est le pourcentage de personnes ayant une pression systolique égale à 12 ?
3. Distributions conditionnelles
 - (a) Déterminer les distributions conditionnelles en effectifs et en fréquences des variables Y sachant X et Y sachant X .
 - (b) Quel est le pourcentage de personnes ayant une pression diastolique égale à 6 sachant qu'elles ont une pression systolique de 11 ?

- (c) Quel est le pourcentage de personnes ayant une pression diastolique égale à 12 sachant qu'elles ont une pression systolique de 7 ?
 - (d) Calculer les moyennes et les variances conditionnelles de Y sachant X et de X sachant Y.
4. Retrouver le lien entre la moyenne marginale et les moyennes conditionnelles.
 5. Retrouver le lien entre la variance marginale et les variances conditionnelles.