

Exercices : Statistiques

Exercice 1 :

Au poste de péage, on compte le nombre de voitures se présentant sur une période de 5mn. Sur 100 observations de 5mn, on obtient les résultats suivants :

Nombre de voitures	1	2	3	4	5	6	7	8	9	10	11	12
Nombre d'observations	2	8	14	20	19	15	9	6	2	3	1	1

1) Construire la table des fréquences et le diagramme en bâtons en fréquences de la série du nombre de voitures. 2) Calculer la moyenne et l'écart-type de cette série. 3) Déterminer la médiane, les quartiles et tracer le box-plot. 4) Étudier la symétrie de la série.

Exercice 2 :

Entre 1996 et 2000, les salaires ont évolués en France : le salaire moyen a augmenté, le salaire médian est resté stable, l'écart inter-quartile ainsi que l'étendue ont augmenté. Interpréter ces données.

Exercice 2 : Paradoxe de Simpson

On dispose de deux traitements contre les calculs rénaux. Le traitement A et le traitement B.

Sur une étude, on a montré que :

- Pour les malades souffrant de petits calculs, le traitement A a soigné 81 malades sur 87 et le traitement B, 234 sur 270. - Pour les malades souffrant de gros calculs, le traitement A a soigné 192 malade sur 263 et le traitement B, 55 sur 80.

Quel traitement est le plus efficace ?

Exercice 3 :

Une maladie mortelle touche 100 personnes sur 1000 dans la population. Un test est mis au point pour détecter la présence de la maladie. On évalue qu'il est positif dans 99% des cas lorsque le patient est atteint (vrai positif) et qu'il est négatif dans deux tiers des cas quand le patient n'est pas malade.

Le traitement de cette maladie est lui-même très dangereux et est fatal pour une personne sur deux.

Un patient obtient un résultat positif. Le médecin doit-il décider d'opérer ?

Exercice 4 :

Dans une expérience, on compare la croissance de haricots avec et sans engrais. On obtient les données suivantes :

engrais	oui	oui	oui	oui	oui	non	non	non	non	non
Masse au bout de 20 jours (g)	35.2	32.2	31.4	30.0	33.7	21.0	28.0	25.0	27.6	25.4

Puis on recommence l'expérience, et on obtient cette fois-ci les données suivantes :

engrais	oui	oui	oui	oui	oui	non	non	non	non	non
Masse au bout de 20 jours (g)	47.2	31.3	24.1	17.8	42.1	40.4	10.4	25.4	35.8	15.0

1. Pour chaque expérience, déterminer la moyenne, la médiane, l'étendue, l'écart-type et tracer les boîtes à moustache. Conclusion ?

Exercice 5 :

Dans un groupe de garçons et de filles de 19 ans, on a mesuré la taille et la masse de chaque personne, on obtient les résultats suivants :

Sexe	G	G	G	G	G	F	F	F	F	F	F	F	F	F	F
Taille	180	185	178	186	180	160	165	180	175	172	178	163	173	173	169
Masse	72	83	65	81	80	58	64	71	60	68	60	73	64	61	65

1. Pour le groupe des garçons, des filles et le groupe complet : calculer moyenne, médiane, écart-type, quartiles. Représenter la boîte à moustache. Conclusions.
2. Représenter le nuage des points. Calculer, pour le groupe des garçons, des filles et le groupe complet, le point moyen et le coefficient de corrélation.
3. Déterminer la droite des moindres carrés dans chaque cas. Conclusion.

Exercice 7 : Autres moyennes

Pour deux nombres réels strictement positifs, a et b , on définit la moyenne géométrique par $g(a, b) = \sqrt{ab}$ et la moyenne harmonique par $h(a, b) = \frac{2}{\frac{1}{a} + \frac{1}{b}}$. On notera $m(a, b)$ la moyenne arithmétique de a et b .

1. Un train parcourt l'aller d'un trajet à la vitesse v_1 et le retour à la vitesse v_2 . Quelle est sa vitesse moyenne sur le trajet ?
2. Un salaire augmente de 5% puis de 10%. Quelle est l'augmentation finale ? Quelle est l'augmentation moyenne ?
3. Montrer que $h(a, b) \leq g(a, b) \leq m(a, b)$
4. Déterminer une relation entre $g(a, b)$, $h(a, b)$ et $m(a, b)$.

Exercice 8 : Ajustement affine

Lors d'une épidémie, on a relevé, à intervalle régulier le nombre de cas déclarés. On a obtenu les résultats suivants :

Rang du relevé : x_i	0	1	2	3	4	5
Nombre de cas y_i :	600	690	794	913	1045	1205

1. Représenter le nuage de point. Calculer le coefficient de corrélation. Un ajustement affine semble-t-il pertinent ?
2. Déterminer la droite d'ajustement affine à la méthode des moindres carrés. Combien peut-on prévoir de cas pour le sixième relevé ?
3. Une autre méthode, appelée méthode de Mayer, consiste à déterminer la droite passant par les points moyens de deux sous groupes des données de départ. Appliquer cette méthode. Conclusion ?