

Seuls les documents distribués avec le sujet sont acceptés. Les calculatrices sont acceptées.
Les exercices sont indépendants et peuvent être traités dans l'ordre que vous souhaitez.

Exercice 1

Un responsable de ventes de magasin analyse l'évolution de son chiffre d'affaires sur la dernière période. Il relève également le montant des frais de publicité engagés sur la même période. Il dresse le tableau suivant (les montants sont exprimés en centaines d'euros).

| | | | | | | | | | | |
|---------------------------|-----|-----|-----|------|-----|-----|-----|-----|-----|-----|
| <i>Frais de publicité</i> | 10 | 6 | 6,5 | 11,5 | 11 | 8 | 7 | 6,5 | 11 | 9 |
| <i>Chiffre d'affaires</i> | 250 | 220 | 228 | 262 | 268 | 244 | 240 | 222 | 259 | 246 |

- 1) Calculer les paramètres de position (moyenne, médiane) des 2 séries.
- 2) Calculer les paramètres de dispersion (écart-type, écart inter-quartile) des 2 séries.
- 3) Calculer le coefficient de corrélation entre les frais de publicité et le chiffre d'affaires.
Peut-on conclure à une dépendance linéaire entre les frais de publicité et le chiffre d'affaires?
- 4) Tracer le graphique (frais de publicité/ chiffre d'affaires). Peut-on modéliser le problème par une régression linéaire? Si oui, calculer les coefficients de la droite de régression.

Exercice 2

Avant le Tour de France, le médecin d'une équipe cycliste désire tester l'efficacité d'une nouvelle recette indécélable aux contrôles usuels. Après avoir ingurgité la mixture, le coureur cycliste effectue un test d'effort sur un vélo de laboratoire. Deux relevés sont effectués.

- 1) Le premier relevé est le temps, exprimé en secondes, écoulé avant que le rythme cardiaque atteigne son maximum. Les données sont les suivantes:

| | | | | | | | | | | |
|----------|-----|------|-----|-----|-----|-----|------|-----|-----|-----|
| Cycliste | A | B | C | D | E | F | G | H | I | J |
| Temps | 215 | 238* | 300 | 370 | 420 | 460 | 530* | 540 | 620 | 700 |

Les nombres suivis d'une étoile correspondent à des coureurs qui ont du interrompre l'expérience.

- a) En utilisant la méthode de Kaplan-Meier, estimer la fiabilité $R(t)$.
 - b) Calculer le coefficient de corrélation de la série $(t ; \ln(R(t)))$.
 - c) Peut-on justifier l'utilisation d'une loi exponentielle pour modéliser ce temps ? Si oui, estimer le paramètre de cette loi.
- 2) Après la première période où le rythme cardiaque augmente, le coureur entre dans une phase où le rythme est à peu près stable jusqu'à l'apparition de signes d'épuisement se traduisant par une baisse du rythme cardiaque. Les temps, exprimés en minutes, durant lequel le rythme est stable sont donnés ci-dessous:

| | | | | | | | | |
|----------|----|----|----|----|----|----|---|----|
| Cycliste | A | B | C | D | E | F | G | H |
| Temps | 36 | 70 | 49 | 16 | 21 | 12 | 6 | 26 |

- a) Déterminer une estimation de $R(t)$, par la méthode des rangs médians.
- b) Justifier graphiquement l'utilisation d'une loi exponentielle pour modéliser ce temps, et estimer le paramètre de cette loi.

Exercice 3

Une laiterie produit des camemberts au lait cru. La masse X , exprimée en g, d'un camembert tiré au hasard dans la production, est distribuée selon une loi normale.

On tire un échantillon de 17 camemberts que l'on pèse et dont le tableau suivant fournit les masses :

| | | | | | | | | | | | | | | | | | |
|--------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Masses | 250 | 254 | 254 | 253 | 256 | 250 | 257 | 251 | 253 | 255 | 250 | 255 | 252 | 261 | 252 | 251 | 255 |
|--------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

- 1) Donner la moyenne et la variance de cet échantillon.
- 2) Donner une estimation ponctuelle de la variance s^2 de la production.
- 3) Déterminer une estimation par intervalle de confiance à 95% de la masse moyenne m de la production
- 4) Quelle taille minimale donner à un échantillon pour obtenir un intervalle de confiance pour m , au niveau 95%, d'amplitude inférieure à 1, avec une variance $s^2 = 8,51$.
- 5) Le responsable de fabrication estime que plus de 15 % des camemberts de la production ont une masse supérieure à 257 g. On tire un échantillon aléatoire simple de 200 camemberts. On constate que 40 d'entre eux pèsent plus de 257 g. Au vu de cet échantillon, peut-on conclure, au seuil de signification 5 %, que le responsable de fabrication a raison ?

Exercice 4

Dans le but de tester l'influence de l'heure de prise sur l'efficacité d'un médicament, on a effectué une enquête sur un échantillon de 600 malades d'un hôpital.

La variable « moment de prise du médicament » a trois modalités : matin, après-midi, soir.

Sur 200 malades ayant pris le médicament le matin, on a observé 141 guérisons.

Sur 200 malades ayant pris le médicament l'après-midi, on a observé 125 guérisons.

Sachant qu'on a observé en tout 420 cas de guérison dans cet échantillon :

- 1) Etablir un tableau synthétisant les données.
- 2) Peut-on dire au seuil de 5% qu'il existe un lien entre le moment de la journée auquel ce médicament est administré et son efficacité ?

Exercice 5

On constitue un échantillon de nourrissons français en prenant 40 filles et 35 garçons.

Au seuil de 5%, cet échantillon est-il représentatif de la population des bébés en France vis-à-vis du sexe, sachant que sur le plan national il y a 48,78% de bébés filles ?