

Examen: Session Normale

Nom et Prénom* : N° d'examen :

REMARQUES IMPORTANTES

- Les téléphones portables doivent être éteints.
 - Aucun document n'est autorisé.
 - Seules les calculatrices non programmables sont autorisées.
 - Les exercices sont indépendants. Ils ne sont pas classés par ordre de difficulté.
-

Questions de cours: (7pts)

Cocher la (les) bonne(s) réponse(s):

1. Une série statistique peut avoir plusieurs modes.
 Oui Non
2. La médiane d'une série statistique est son premier quartile.
 Oui Non Une autre proposition:
3. Les valeurs inférieures ou égales au premier quartile représente 50% ou plus de la population.
 Oui Non Une autre proposition:
4. On peut tirer d'une boîte à moustaches le mode les quartiles.
 Oui Non Une autre proposition:
5. Soient $(x_j)_{1 \leq j \leq n}$ et $(y_j)_{1 \leq j \leq n}$ deux séries statistiques telles que $y_j = 2x_j - 7$. Soient \bar{x} , \bar{y} et s_x^2 , s_y^2 leurs moyennes et variances respectives. Alors
 $\bar{y} = \bar{x} - 7$ $s_x^2 = 2s_y^2$ $\bar{y} = 4\bar{x}$ $s_y^2 = 4s_x^2$
6. On représente les séries qualitatives par
 Diagramme en secteurs Histogramme Diagramme en bâtons
7. Pour comparer la dispersion de deux séries exprimées dans les mêmes unités, on utilise
 la moyenne le mode la médiane l'écart interquartile
8. Les indices servent à comparer les états d'une même variable dans deux situations différentes, une situation prise comme référence et une autre situation que l'on compare à la première.
 Oui Non

*Feuille à rendre avec la copie.

Exercice 1: (9pts)

Le tableau suivant donne la moyenne Y des maxima de tension artérielle en fonction de l'âge X d'une population donnée :

Âge : X	36	42	48	54	60	66
Tension : Y	12	13.5	12.5	14.25	15.5	15

1. Préciser la nature des variables étudiées.
2. a) Représenter graphiquement la série d'observations (x_i, y_i) de la variable couple (X, Y) , en précisant la droite D des points extrêmes $E_1(36; 12)$ et $E_2(66; 15)$.
b) On propose l'équation $y = \alpha x + \beta$ de la droite D comme un modèle de prévision des valeurs de Y en fonction de X . Déterminer les réels α et β .
c) En déduire une prévision \hat{y}_7 de tension artérielle pour une personne âgée de 78 ans.
3. Calculer les moyennes \bar{x} et \bar{y} des séries d'observations x_i, y_i en question.
4. Calculer le coefficient de corrélation linéaire entre X et Y , et interpréter le résultat.
5. a) Déterminer l'équation de la droite Δ de régression linéaire de Y en fonction de X .
b) En déduire une nouvelle prévision \hat{y}'_7 de tension artérielle pour une personne de 78 ans.
6. On définit l'erreur quadratique moyenne d'un modèle linéaire de prévision $D : \alpha x + \beta$ par

$$EQM(D) = \frac{\sum_{i=1}^n (y_i - \alpha x_i - \beta)^2}{n},$$

où (x_i, y_i) désigne la série d'observations de la variable (X, Y) . Les valeurs d'un modèle de prévision sont plus vraisemblables, si son erreur quadratique moyenne est minimale.

- a) Calculer $EQM(D)$ et $EQM(\Delta)$.
- b) Lequel des deux résultats \hat{y}_7 et \hat{y}'_7 paraît le plus vraisemblable ? (Justifier votre réponse).

Exercice 2: (4pts)

Le tableau ci-après donne la répartition de 18 483 personnes d'une ville **A** suivant:

- Qu'elles ont été vaccinées ou non contre la fièvre typhoïde;
- Qu'elles ont, ou non, été atteintes par la maladie.

Manifestation de la maladie	Oui	Non	Effectif Marginal
Vaccination			
Oui	56	6759	
Non	272	11396	
Effectif Marginal			

1. Remplir le tableau ci-dessus. En déduire le pourcentage des personnes vaccinées.
2. Dresser le tableau des profils lignes.
3. Les deux variables en question sont-elles indépendantes? (Justifier votre réponse).