Durée de l'épreuve : 2h00 - Calculatrice autorisée, aucun document n'est permis

Exercice 1

Vous menez une petite étude sur les loisirs des français. Vous interrogez 415 individus croisés lors du *salon du livre*. Vous leur demandez, entre autre, combien de temps par jour, en moyenne, ils regardent la télévision et combien de romans ils ont lu l'année dernière. Le tableau suivant présente les résultats croisés des réponses à ces deux questions.

$X_i \backslash Y_j$	[0 ; 12]	[13 ; 25]	[26 ; 38]	[39 ; 51]	Total
[0 ; 3[heures	66	54	52	49	221
[3 ; 6[heures	58	34	22	6	120
[6 ; 9[heures	59	9	4	2	74
Total	183	97	78	57	415

X_i = temps moyen par jour, passé devant la télévision (en heures).
Y_j = nombre de romans lus l'année passée (de 0 à 51).

1) Déterminez la valeur de n_{21}, f_{32}, n_3 et f_2 *(1 point)*

2) Calculez la moyenne, la variance et l'écart type des Xi et des Yj *(1,5 point)*

3) Calculez la médiane et le mode des Xi *(1 points)*

Rappel : $Mo = L + a \times \left[\frac{e_s - e_i}{(e_s - e_i) + (e_s - e_i)} \right]$

4) Comparez la moyenne, le mode et la médiane des Xi. Quelle information peut-on en tirer sur la symétrie de cette distribution ? Vérifiez cette interprétation en calculant les deux coefficients d'asymétrie de Pearson s et $\beta 1$. Enfin, calculez et interprétez le coefficient d'aplatissement β_2 de Pearson des Xi. *(3 points)*

5) Déterminez le coefficient de corrélation et le coefficient de détermination entre Xi et Yj. Le coefficient de corrélation est-il statistiquement significatif ? Expliquez ces résultats et tirez une conclusion de cette étude. *(3 points)*

6) Déterminez l'équation de régression $Y=aX+b$. D'après cette équation, en moyenne, combien de romans par an devrait lire un individu qui passe une heure par jour à regarder la télévision ? Ce résultat vous paraît-il généralisable à la population française ou présente-t-il un biais ? *(2 points)*

7) Vous choisissez un individu au hasard parmi ces 415 interrogés. Quelle est la probabilité :
- qu'il regarde la télé plus de 3h par jour ?
- qu'il ait lu plus de 25 romans dans l'année, sachant qu'il regarde la télé moins de 3h par jour ?
- qu'il regarde la télé au moins 6h par jour, sachant qu'il a lu moins de 13 romans l'année dernière ? *(1,5 point)*

8) Vous décidez de changer d'approche statistique en agrégeant les résultats. Vous croisez les deux variables binaires suivantes : le fait d'être ou non un gros consommateur de télé (plus de 3 heures par jour *versus* moins de 3 heures par jour) et le fait d'être ou non un gros lecteur (plus de 12 romans par an *versus* 12 romans ou moins par an). Vérifier l'existence d'un lien statistiquement significatif entre ces deux nouvelles variables. *(2 points)*

Exercice 2

18 individus ont été sélectionnés pour un entretien de recrutement. Après une présentation de groupe, ils vont tous être auditionnés pendant 10 minutes exactement, chacun leur tour. Parmi eux, il y a 10 gestionnaires et 8 ingénieurs. L'ordre de passage est aléatoire.

9) Quelle est la probabilité que le premier ingénieur ne soit auditionné qu'après le passage de 5 gestionnaires ? *(1 point)*

10) Quelle est la probabilité que le second ingénieur ne soit auditionné qu'après le passage de 5 gestionnaires ? *(1,5 point)*

11) Quelle est la probabilité que les trois premiers candidats auditionnés soient des gestionnaires puis qu'une alternance ingénieur-gestionnaire-ingénieur-gestionnaire-ingénieur (...) ait lieu jusqu'au dernier candidat ? *(1,5 point)*

12) Soit X la variable aléatoire égale au nombre de minutes attendues par un candidat donné. Déterminez l'espérance E(X), la variance V(X) et l'écart-type $\sigma(X)$. *(1 point)*