

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

Exercice 1

Vous menez une petite étude sur les français et le bio. Vous interrogez 500 individus représentatifs de la population française. Vous leur demandez, entre autre, quel est leur revenu mensuel net, et quel budget ils consacrent en moyenne chaque mois à l'achat de produits alimentaires bio. Le tableau suivant présente les résultats croisés des réponses à ces deux questions.

Remarque : bio = issus de l'agriculture biologique (sans pesticides ni OGM)

$X_i \backslash Y_j$	$[0 ; 100[$	$[100 ; 200[$	$[200 ; 300[$	Total
$[0 ; 1,5[$	121	42	11	174
$[1,5 ; 3[$	96	55	15	166
$[3 ; 4,5[$	55	28	12	95
$[4,5 ; 6[$	35	19	11	65
Total	307	144	49	500

X_i = salaire mensuel net (en milliers d'euros).

Y_j = budget mensuel moyen consacré à l'achat de produits alimentaires bio (en euros).

- 1) Déterminez la valeur de n_{32} , f_{12} , n_2 et $f_{.1}$. (1 point)
- 2) Calculez la moyenne et l'écart type des X_i et des Y_j . (1 point)
- 3) Calculez la médiane et le mode des X_i . (1 point)

$$\text{Rappel : } Mo = L + a \times \left[\frac{e_s - e_i}{(e_s - e_i) + (e_i - e_{i-1})} \right]$$

- 4) Comparez la moyenne, le mode et la médiane des X_i . Quelle information peut-on en tirer sur la symétrie de cette distribution ? Vérifiez cette interprétation en calculant les deux coefficients d'asymétrie de Pearson α et β_1 . Enfin, calculez et interprétez le coefficient d'aplatissement β_2 de Pearson des X_i . (3 points)

- 5) Déterminez le coefficient de corrélation et le coefficient de détermination entre X_i et Y_j . Le coefficient de corrélation est-il statistiquement significatif ? Expliquez ces résultats et tirez une conclusion de cette étude. Peut-on dire que l'importance consacrée aux produits bio est proportionnelle au revenu ? Votre étude comporte-t-elle des biais ? (3 points)

- 6) Déterminez l'équation de régression $Y = aX + b$. D'après cette équation, en moyenne, quel budget mensuel consacre à l'achat de produits alimentaires bio un individu qui a un revenu de 2000 euros ? Ce résultat est-il fiable ? (1,5 point)

- 7) Vous choisissez un individu au hasard parmi ces 500 interrogés. Quelle est la probabilité :
 - qu'il consacre plus de 100 euros de budget par mois à l'achat de produits alimentaires bio ?
 - qu'il consacre plus de 100 euros de budget par mois à l'achat de produits alimentaires bio sachant qu'il gagne plus de 3000 euros par mois ?
 - qu'il gagne moins de 1500 euros par mois, sachant que son budget consacré au bio est supérieur à 200 euros ?
 (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 bio (plus de 200 euros par mois de budget *versus* moins de 200 euros par mois) et le fait d'avoir ou non des revenus supérieurs à 3000 euros nets mensuel. Vérifier l'existence d'un lien statistiquement significatif entre ces deux nouvelles variables. (2 points)

Exercice 2

Un virus touche 2% des exploitations de poulets d'une région donnée. Lorsqu'un des poulets de l'exploitation contracte ce virus, tous les autres poulets de l'exploitation le contractent presque immédiatement. Un test simple et rapide permet de détecter ce virus. Ce test n'est pas complètement fiable et nécessite que le poulet soit abattu et disséqué. Ce test déclare infectés dans 99,5% des cas les poulets effectivement infectés. Il déclare sains dans 88% des cas les poulets effectivement sains.

- 1) Un éleveur fait passer ce test à l'un de ses poulets et il est déclaré infecté. Quelle est la probabilité pour que son poulet soit effectivement infecté ? (1,5 point)
- 2) Si l'on choisit au hasard 10 exploitations différentes de poulets dans la région, quelle est la probabilité que plus d'un tiers d'entre elles soient effectivement infectées ? (1,5 point)
- 3) Soit X la variable aléatoire correspondant au nombre d'exploitations infectées parmi ces 10. Déterminez l'espérance et la variance de X . (1,5 point)
- 4) Imaginons qu'aucune de ces exploitations ne soit infectée. Quelle est la probabilité que le test ne signale aucun cas infecté ? (1,5 point)