

Statistiques et probabilités



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

• Partie statistique •
12 points

Un responsable de l'EMS a réalisé un tableau de synthèse sur le devenir des 460 étudiants diplômés d'un master l'an dernier. Il croise le nombre de mois passés en stage en entreprise pendant l'ensemble de la formation universitaire (minimum 6 mois) avec le salaire mensuel net actuel de ces diplômés (en milliers d'euros). Passer plus de temps en stage pendant ses études permet-il de gagner davantage en début de carrière ?

X_i, Y_j	1,4	2,2	3	3,8	
	[1 ; 1,8]	[1,8 ; 2,6]	[2,6 ; 3,4]	[3,4 ; 4,2]	Total
13	100	105	44	12	261
15	42	46	30	19	137
21	2	4	24	32	62
Total	144	155	98	63	460

X_i = nombre de mois passés en stage pendant l'ensemble des études supérieures

Y_j = salaire net mensuel en milliers d'euros

- 1) Déterminez la valeur de n_{12} , f_{12} , n_j et f_j (1 point)
 - 2) Calculez la moyenne, la variance et l'écart type des X_i et des Y_j (1,5 point)
 - 3) Calculez la médiane (par interpolation linéaire) et le mode des salaires (1 point)
- Rappel : $Mo = L + a \times \left[\frac{e_1 - e}{(e_1 - e) + (e_2 - e_1)} \right]$
- 4) Comparez la moyenne, le mode et la médiane des salaires. Quelle information peut-on en tirer sur la symétrie de leur distribution ? Vérifiez cette interprétation en calculant les deux coefficients d'asymétrie de Pearson s et β_1 . Enfin, calculez et interprétez le coefficient d'aplatissement β_2 de Pearson des Y_j . (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. (2,5 points)

Rappel : Coefficient de corrélation : $r = \frac{COV(x,y)}{\sigma(x)\sigma(y)}$
Coefficient de détermination : r^2

$$T \text{ de Student : } t = r \cdot \sqrt{\frac{n-2}{1-r^2}}$$

6) Déterminez l'équation de régression $Y=aX+b$. D'après cette équation, si l'un des anciens étudiants vous dit qu'il a passé 20 mois en stage pendant ses études, quel devrait être son salaire aujourd'hui ? Expliquez pourquoi ce dernier résultat n'est pas très fiable. Quelles autres variables pourrait-on prendre en compte pour améliorer la fiabilité de cette étude ?
(1,5 point)

7) Vous choisissez un individu au hasard parmi ces 460 anciens étudiants. Quelle est la probabilité :

- qu'il ait réalisé moins de 12 mois de stage ?
- qu'il ait un salaire supérieur à 2600 euros, sachant qu'il a réalisé moins de 18 mois de stage pendant ses études ?
- qu'il ait réalisé plus de 18 mois de stage, sachant qu'il a un salaire actuel supérieur à 1800 euros ?

(1,5 point)

*** Partie probabilités ***

8 points

Exercice 1 : une épidémie touche 3% des habitants de votre quartier. Un test sanguin de dépistage existe mais n'est pas complètement fiable. Ce test vous déclare malade dans 96% des cas si vous êtes effectivement malade, et vous déclare non malade dans 88% des cas si vous n'êtes pas malade. Vous passez le test, et il vous déclare malade. Quelle est la probabilité pour que vous le soyez vraiment ?
(3 points)

Exercice 2 : vous n'êtes pas sûr d'avoir enregistré un fichier avant d'avoir éteint votre ordinateur. Notons p la probabilité que vous l'avez bien enregistré. Le disque dur de votre ordinateur comporte 6 dossiers dans lesquels le fichier est susceptible de se trouver. Quelle est la probabilité que le fichier se trouve dans le sixième dossier sachant que vous ne l'avez pas trouvé dans les 5 premiers dossiers ?
(3 points)

Exercice 3 : vous jetez 3 dés. Quelle est la probabilité d'obtenir 6, 5 et 3 ? Quelle est la probabilité d'obtenir 6, 4 et 4 ?
(2 points)