

Promotion : Master 2 "Biotechnologie Microbienne"

Durée: 1h30

Corrigé type d'examen de Bioinformatique 2023/2024

Exercice 1: Pour mesurer la dépendance entre l'âge et le risque cardio-vasculaire, on a observé 12 patients, pour lesquels on dispose de l'âge en années (variable X), et du logarithme du dosage en d-dimères (variable Y). On donne les quantités suivantes:

$$\sum x_i = 596, \quad \sum x_i^2 = 32435, \quad \sum y_i = -5.2, \quad \sum y_i^2 = 4.3, \quad \sum x_i y_i = -188.58.$$

1. Calculer le coefficient de corrélation linéaire de X et Y. Que peut-on en déduire ?

(1pt) Les moyennes: $\bar{x} = \frac{\sum x_i}{N} = \frac{596}{12} = 49.66, \quad \bar{y} = \frac{\sum y_i}{N} = \frac{-5.2}{12} = -0.43$

(1pt) La covariance entre x et y: $\sigma_{xy} = \frac{\sum x_i y_i}{N} - \bar{x} * \bar{y} = \frac{-188.58}{12} - (49.66)(-0.43) = 5.80$

(2pt) Les variances: $\sigma_x^2 = \frac{\sum x_i^2}{N} - (\bar{x})^2 = \frac{32435}{12} - (49.66)^2 = 236.139, \quad \sigma_y^2 = \frac{\sum y_i^2}{N} - (\bar{y})^2 = \frac{4.3}{12} - (-0.43)^2 = 0.17$

(1pt) Le coefficient de corrélation: $r = \frac{\sigma_{xy}}{\sigma_x * \sigma_y} = \frac{5.80}{\sqrt{236.139 * 0.17}} = 0.91$

(1pt) Comme le coefficient de corrélation linéaire r est proche de 1 alors la relation linéaire positive entre les variables est forte.

2. Donner l'équation de la droite de régression de Y en X.

la droite de régression de Y en X à pour équation $y = ax + b$ tel que:

$$a = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{5.80}{236.139} = 0.024 \text{ et } b = \bar{y} - a\bar{x} = -1.654, \text{ donc l'équation de la droite est: } y=5.8 x-1.654 \text{ (1pt)}$$

3. Quelle valeur de Y prévoyez-vous pour un individu de 60 ans? $y = 5.8 * 60 - 1.654 = -0.179$ (1pt)

Exercice 2 - QCM - Entourer la bonne réponse

Le taux normal de glycémie est 1 g/l de sang, on dose la glycémie chez 120 sujets diabétique depuis 4 heures. On observe somme $\sum x_i = 144$ et $\sum x_i^2 = 198$ de l'échantillon.

1) le taux de glycémie moyen de la population est estimé à: a) 1.8 b) 1.2 c) 2.5

2) la variance du taux de glycémie de la population est estimé à: a) 0.18 b) 0.21 c) 0.31

Maintenant, on veut tester si ces sujets sont hyperglycémiques au risque de 5%?

3) Lors de la réalisation de ce test, la première étape correspond à:

a) Calculer la probabilité « p » b) Calculer la statistique du test c) Formuler les hypothèses.

4) Dans ce cas, on réalisera un test: a) unilatéral à droite b) unilatéral à gauche c) bilatéral.

5) La statistique du test suit approximativement une loi de: a) Student b) $\mathcal{N}(0, 1)$ c) khi-deux

6) La statistique du test est égale à: a) $T = 10.90$ b) $T = 0.087$ c) $T = 4.78$

7) Au risque de 5%: a) on rejette H_0 b) on accepte H_0 c) on rejette H_1

Exercice 3 - QCM - Entourer la bonne réponse

On cherche à estimer la proportion de lapins touchés par la myxomatose dans un département. 150 lapins sont prélevés: 33 s'avèrent infectés et subiront un traitement, les autres seront vaccinés puis remis dans leur milieu.

1) La taille de l'échantillon étudié est: a) 12 b) 138 c) 150

2) La fréquence de lapins infectés par la myxomatose dans cet échantillon est égale à:

a) 8 % b) 8,7 % c) 22 %

3) L'intervalle de confiance au niveau de confiance de 95 % est :

a) [0,037;0,12] b) [0,138;0,302] c) [0,154;0,286]
