Contrôle final de Statistiques

Allan Merino

DUT Mesures Physiques, Metz - Mars 2016

Aucun document n'est autorisé. Dans ma bonté légendaire, la calculatrice est quant à elle autorisée. Pensez à bien rédiger vos réponses, et n'hésitez pas à faire apparaître sur la copie des éléments de réponse même si vous n'avez pas entièrement abouti. Voltaire disait : " Il est beau d'écrire ce que l'on pense, c'est le privilège de l'homme. "

Exercice 1

Des mesures du diamètre apparent vertical de la planète Vénus ont donné, en secondes d'arc, les résultats suivants :

$42.7 - 43.01 - 42.76 - 43.63 - 41.6 - 42.95 - 43.18 - 43.1 - 42.56 - 43.48 - 43.06 - 42.87 - 42.78 - 43.2 - 43.39$

Soit X la variable représentant, non pas le diamètre apparent vertical de Vénus, qui lui, est une grandeur parfaitement déterminée, mais les résultats d'une mesure de cette grandeur rendue aléatoire par le fait que de nombreux phénomènes annexes interviennent dans la pratique de cette mesure. Compte tenu du grand nombre de causes diverses d'erreurs de mesure, il est raisonnable de considérer que X suit une répartition normale $\mathcal{N}(\mu, \sigma)$.

1. Donner un intervalle de confiance pour la moyenne μ d'après l'échantillon de mesure obtenu, au seuil 0.05 et en supposant que $\sigma = 0.5$.
2. Combien de mesures faudrait-il avoir au minimum, pour obtenir, à ce même seuil, un intervalle de confiance pour μ de longueur maximale 0.1 ?

Exercice 2

Le lancement d'un nouveau produit nécessitant une réorganisation complète d'un atelier, une entreprise décide de faire une étude de marché sous forme de questionnaire dont on retient en particulier l'information suivante : "la personne interrogée est ou n'est pas intéressée par le nouveau produit". Soit p la proportion (réelle mais inconnue) des personnes intéressées par le nouveau produit. L'entreprise juge qu'il est raisonnable de lancer ce nouveau produit si plus de 50% des personnes sont réellement intéressées. Elle décide donc de faire le test :

$$H_0 : p = 0,5 \qquad \text{contre } H_1 : p > 0,5$$

1. Déterminer la région d'acception de H_0 au seuil de 1% lorsque l'on se base sur les réponses de 100 personnes.

2. Quelle est la conclusion du test et la décision de l'entreprise si 58 personnes sur 100 interrogées déclarent être intéressées par le nouveau produit.

Exercice 3

On cherche à savoir si la fréquence d'une maladie est liée au groupe sanguin. Sur 200 malades observés, on a dénombré 104 personnes du groupe O, 76 personnes du groupe A, 18 personnes du groupe B et 2 personnes du groupe AB. On admettra que dans la population générale la répartition entre les groupes est la suivante :

Groupes	O	A	B	AB
Répartition	47 %	43 %	7 %	3 %

Que concluez-vous ?

Exercice 4

On appelle "fréquence seuil" d'un sportif amateur sa fréquence cardiaque obtenue après trois quarts d'heure d'un effort soutenu de course à pied. Celle-ci est mesurée à l'aide d'un cardio-fréquence-mètre. On cherche à savoir si l'âge d'un sportif a une influence sur sa fréquence seuil. On dipose pour cela de 20 valeurs du couple (x_i, y_i), où x_i est l'âge et y_i la fréquence seuil du sportif. On a obtenu $(\bar{x}, \bar{y}) = (35,6; 170,2)$ et :

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = 1991 \qquad \sum_{i=1}^{n}(y_i - \bar{y})^2 = 189,2 \qquad \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = -195,4$$

1. Calculer la droite des moindres carrés pour le modèle $y = \beta_1 + \beta_2 x + \varepsilon$.
2. Calculer le coefficient de détermination R^2. Commenter la qualité de l'ajustement des données au modèle.
3. Avec ces estimateurs la somme des carrés des résidus vaut $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = 170$. Si on suppose les perturbations ε_i gaussiennes, centrées, indépendantes et de même variance σ^2, en déduire un estimateur $\hat{\sigma}^2$ de σ^2.
4. Donner un estimateur $\hat{\sigma}_2^2$ de la variance de $\hat{\beta}_2$.
5. Tester l'hypothèse $H_0 : \beta_2 = 0$ contre $H_1 : \beta_2 \neq 0$ pour un risque de 5 pourcents. Conclure sur la question de l'influence de l'âge sur la fréquence seuil.

Bonus

1. Notons Φ la fonction de répartition d'une loi normale centrée réduite. Pour tout $z \in \mathbb{R}$, montrer que $\Phi(z) = 1 - \Phi(-z)$.
2. Soit $p \in]0, 1[$ et soit z_p le p-quantile d'une loi normale centrée réduite.
   (a) Montrer que z_p est unique.
   (b) Montrer que $z_p = -z_{1-p}$.

" Ne t'inquiète pas si tu as des difficultés en maths, je peux t'assurer que les miennes sont bien plus importantes ! "
Albert Einstein.