IUT HSE — Probabilités et statistiques

Université de Bordeaux — 2014/2015

Feuille 4.

Exercice 1. Dans le cadre de travaux de recherche sur la durée de la saison de végétation en montagne, des stations météorologiques sont installées à différentes altitudes. La température moyenne (variable Y en degrés Celsius) ainsi que l'altitude (variable X en mètres) de chaque station données dans le tableau ci-dessous :

altitude	1040	1230	1500	1600	1740	1950	2200	2530	2800	3100
température	7.4	6	4.5	3.8	2.9	1.9	1	1.2	1.5	4.5

(1) Calculer le coefficient de corrélation linéaire.
(2) Calculer les estimations des paramètres a, b pour la régression linéaire de Y sur X.
(3) Quelle température moyenne prévoyez-vous à $1100\,m$?

Exercice 2. On veut prédire la hauteur H d'un arbre en fonction de son diamètre D. Pour faire une régression linéaire, on effectue un changement de variable en posant $Y = \ln(H)$ et $X = \ln(D)$. Voici les mesures faites sur 5 arbres :

X	1.61	1.20	0.97	0.51	0.42
Y	2.22	2.27	2.38	2.60	2.65

(1) Représentez ce nuage de points
(2) Donner le coefficient de corrélation linéaire entre X et Y.
(3) Donner l'équation de la droite de régression de Y par rapport à X. représentez cette droite sur le même graphique que le nuage de points.
(4) Donner la hauteur prévue d'un arbre de diamètre $0.7\,m$.

Exercice 3. 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 :

$$\sum x_i = 596; \sum x_i^2 = 32435; \sum y_i = 5.2; \sum y_i^2 = 4.3; \sum x_i y_i = 188.58.$$

(1) Calculer le coefficient de corrélation linéaire de X et Y .
(2) Calculer l'équation de la droite de régression linéaire de Y sur X.

Exercice 4. On souhaite vérifier si la prise de poids d'un jeune mouton en un an (variable Y en kilogrammes) dépend de son poids initial (variable X également en kilogrammes). Sur 100 moutons, on donne les résultats suivants :

$$\sum x_i = 3460; \sum y_i = 4130; \sum x_i^2 = 127000; \sum y_i^2 = 180000; \sum x_i y_i = 151000.$$

(1) Calculez la moyenne et l'écart type empiriques de X et Y.
(2) Calculer le coefficient de corrélation linéaire.
(3) Estimer les paramètres a, b pour la régression linéaire de Y sur X.
(4) Selon ce modèle combien un mouton de poids initial 36 kg devrait-il prendre de poids ?

Exercice 5. La dure de vie moyenne d'un échantillon non-exhaustif de 100 tubes fluorescents a été établie par le calcul à 1570 heures avec un écart-type de 120 heures. Soit μ la durée de vie moyenne en heure d'un tube fluorescent, tester les hypothèses suivantes:

(1) l'hypothèse $\mu = 1600$ heures relativement à l'hypothèse $\mu \neq 1600$ heures, avec un niveau de signification 0.05, puis de 0.01.
(2) l'hypothèse $\mu \leq 1600$ heures relativement à l'hypoth'ese $\mu > 1600$ heures, avec un niveau de signification de 0.05, puis de 0.01.

Exercice 6. Chez un individu adulte, le logarithme du dosage en d-dimères, variable que nous noterons X, est modélisé par une loi normale d'espérance μ et de variance σ^2. La variable X est un indicateur de risque cardio-vasculaire: on considère que chez les individus sains, $\mu = 1$, alors que chez les individus à risque, $\mu = 0$. Dans les deux cas, la valeur de σ^2 est la même : $\sigma^2 = 0.09$.

Le Dr. House ne souhaite pas alarmer inutilement ses patients. Quelles hypothèses H_0 et H_1 choisira-t-il de tester ? Donner la règle de décision pour son test, au seuil de 1%, et au seuil de 5%.

Exercice 7. Une machine à emballer est censée produire des paquets de 1 kg. Le poids réel des paquets est modélisé par une variable aléatoire suivant une loi normale dont l'écart-type vaut 20 g. Par contre, il est possible de régler le poids moyen des paquets.

(1) Le responsable de la production décide de ne pas mettre à la vente les paquets dont le poids s'écarterait trop de la valeur nominale de 1 kg. Quelles hypothèses H_0 et H_1 doit-il tester ? Établir la règle de décision de ce test aux seuils de 5% et 1%.
(2) Le patron de l'usine prétend que les paquets mis à la vente sont souvent trop lourds, ce qui fait perdre de l'argent l'usine. Quelles hypothses H_0 et H_1 le responsable de production doit-il tester ? Établir la règle de décision de ce test aux seuils de 5% et 1%.