Fiche n°5. Tests.

1. Tests paramétriques

Exercice 1. Un laboratoire de chimie est chargé de conditionner des flacons d'eau de toilette destinés à une parfumerie. On définit une variable aléatoire X associant à chaque flacon le volume de son contenu exprimé en cm^3. On suppose que X suive une loi normale de moyenne μ (inconnue) et d'écart-type (connu) $\sigma = 0.036$. A l'occasion d'une commande, le parfumeur reçoit du laboratoire un lot de flacons. Il envisage d'effectuer un test de conformité de la moyenne μ de la production, avec la valeur $\mu_0 = 43.041$ annoncée par le fournisseur, à partir d'un échantillon de taille $n = 75$. Dans tout ce qui suit, on oppose l'hypothèse nulle $H_0 : \mu = \mu_0$ à l'hypothèse alternative $H_1 : \mu \neq \mu_0$.

(1) Soit $\alpha \in]0,1[$ et $u_{1-\frac{\alpha}{2}}$ le quantile de la loi $\mathcal{N}(0,1)$ d'ordre $1-\frac{\alpha}{2}$. En se plaçant sous l'hypothèse H_0, montrer que
$$\mathbb{P}\left(\overline{X}_n \in \left[\mu_0 - \sqrt{\frac{\sigma^2}{n}}u_{1-\frac{\alpha}{2}},\ \mu_0 + \sqrt{\frac{\sigma^2}{n}}u_{1-\frac{\alpha}{2}}\right]\right) = 1-\alpha$$

(2) En déduire une zone de rejet de l'hypothèse H_0 au seuil de risque de 10%.

(3) Enoncer la règle de décision du test.

(4) Pour réaliser ce test, le parfumeur effectue un prélèvement aléatoire, assimilé à un prélèvement avec remise de 75 flacons pris dans le lot reçu. Les résultats sont donnés dans le tableau suivant :

volume	$[42.930, 42.970[$	$[42.970, 43.010[$	$[43.010, 43.050[$	$[43.050, 43.090[$	$[43.090, 43.130]$
Effectif	2	7	29	19	18

Peut-on affirmer, au niveau de 10%, que la valeur μ_0 annoncée pour μ est correcte ?

Exercice 2. On souhaite contrôler la qualité de fonctionnement d'une usine produisant des ampoules électriques. On note p la probabilité qu'une ampoule produite soit défectueuse. On considère que la fabrication a un régime normale lorsqu'on a $p \leq 10^{-3}$.

(1) On prélève au hasard 500 ampoules indépendantes et on note X le nombre d'ampoules défectueuses de ce lot. Déterminer la loi de la variable aléatoire X. Donner une approximation de cette loi lorsque le régime est normal.

(2) A l'aide de la variable aléatoire X, construire un test permettant de déterminer si la fabrication a un régime normal ou si elle est détérioriée, sachant que l'on souhaite que la probabilité de conclure à tort que le régime est détérioré soit inférieure à 5%. Que peut-on conclure si on observe $X(\omega) = 3$?

(3) Supposons que p soit en réalité $3 \cdot 10^{-3}$. Quelle est la probabilité de déclarer cependant, à l'issue du test de niveau 5%, que le régime de fabrication est normal ?

Exercice 3. Le nombre de personnes dans une file d'attente donnée est une variable aléatoire de loi de Poisson. On a dénombré à dix dates différentes les personnes attendant au guichet de la banque postale d'une certaine ville, de telle sorte que les variables aléatoires $X_1, \ldots, X_{10}$ associées soient indépendantes et de même loi $\mathcal{P}(\theta)$, où θ est inconnu.

(1) Déterminer la loi de $X = X_1 + \cdots + X_{10}$.

(2) Les valeurs observées sont : 3; 0; 1; 4; 2; 3; 1; 4; 0; 2. Peut-on accepter l'hypothèse selon laquelle il y a en moyenne moins d'une personne attendant au guichet de la banque postale au niveau 5% ? *Indication : utiliser le fait que* $\theta \mapsto \mathbb{P}(\mathcal{P}(\theta) \geq a)$ *est croissante.*

2. Tests non paramétriques

Exercice 4. Une enquête sur 160 familles de quatre enfants tirées au hasard dans une population de familles permet d'établir la répartition suivante :

Nombre de filles	0	1	2	3	4
Nombre de familles	16	48	62	30	4

(1) A chaque naissance, la probabilité d'avoir une fille est supposée égale à celle d'avoir un garçon. Quelles hypothèses complémentaires doit-on faire pour ramener à une loi de probabilité simple la variable aléatoire X "nombre de filles dans une famille de 4 enfants" ?

(2) Les effectifs observés sur les 160 familles sont-ils compatibles avec l'hypothèse d'égalité des probabilités des sexes (on choisira un test de niveau 5%) ?

Exercice 5. Une étude est menée dans une université sur l'absentéisme des étudiants. On aimerait savoir si certaines plages horaires sont plus propices à une absence aux cours qu'une autre. Pour cela, au cours d'un mois, on a relevé le nombre d'absences d'étudiants aux cours dans une petite composante à différentes moments de la journée. On souhaite tester l'hypothèse selon laquelle les absences des étudiants se répartissent uniformément tout au long de la journée.

Heures de la journée	Nombre d'étudiants absents
8-10	25
10-12	15
13-15	18
15-17	32

(1) Formuler les hypothèses nulle et alternative et préciser la statistique discriminante.

(2) Au niveau de 5%, peut-on dire que les absences se répartissent uniformément tout au long de la journée ?

Exercice 6. En 1870, le moine Tchèque G. Mendel mène une expérience sur des petits pois en s'intéressant à deux caractères : leurs couleurs, jaunes ou verts, et leurs aspects, lisses ou ridés. Les caractères jaune et lisse sont dominants. [1] En croisant deux variétés pures "Jaune-Lisse" avec "Vert-Ridé", Mendel obtient des petits pois "Jaune-Lisse" qu'il laisse s'autoféconder. Il obtient une nouvelle génération de 556 petits pois dont les phénotypes se répartissent de la manière suivante :

Jaune-Lisse	Jaune-Ridé	Vert-Lisse	Vert-Ridé
315	108	101	32

La deuxième loi de Mendel stipule que les deux caractères étudiés, couleur et aspect, sont indépendants. Peut-on réfuter au risque de première espèce de 5% cette loi à l'aide de l'expérience menée par Mendel (proposer deux tests statistiques).

1. En génétique, un caractère dominant est un caractère qui se manifeste lorsqu'il est présent sur le chromosome.