TESTS STATISTIQUES

SIMON COSTE ET PIERRE MONTAGNON

Exercice 1. La législation en vigueur impose aux aéroports des normes relatives au bruit émis par les avions au décollage et à l'atterrissage. Une limite de 80 dB est fixée pour les zones habitées proches d'un aéroport. Au-delà de cette limite, l'aéroport doit indemniser les riverains.

Les habitants d'un village proche de l'aéroport assurent que la limite de 80 dB est dépassée en moyenne par un certain type d'avions et portent plainte. Des experts sont convoqués ; ils admettent que l'intensité du bruit émis par ce type d'avions peut être modélisée par une variable aléatoire suivant une loi normale de moyenne m et de variance 49.

Ils décident alors de réaliser le test suivant :

$$\mathrm{H}_0 : m = 80 \text{ contre } \mathrm{H}_1 : m < 80$$

et fixent le risque d'erreur admissible à 1%.

(1) Montrer que le test préserve les intérêts des riverains, le risque d'autoriser des avions trop bruyants à voler étant contrôlé de façon privilégiée. Quelle est la signification du risque de deuxième espèce ?

(2) L'enregistrement de l'intensité du bruit dû au passage d'un échantillon de $n = 40$ avions du type concerné a donné une intensité moyenne de 79 dB. Pour faciliter la modélisation, on considère que l'intensité du bruit émis par un avion est indépendante de l'intensité du bruit émis par un autre avion. Soit $\overline{\mathrm{X}_{40}}$ la variable aléatoire qui, à chaque échantillon aléatoire de 40 avions du type concerné, associe la moyenne des intensités de bruit émis par ces avions.

   (a) Sous l'hypothèse H_0, quelle est la loi de $\overline{\mathrm{X}_{40}}$?

   (b) Déterminer la région critique du test.

   (c) Énoncer la règle de décision du test.

   (d) Dans le cas de l'échantillon observé, H_0 est-elle rejetée ?

(3) La compagnie qui commercialise ce type d'avions affirme que l'intensité moyenne du bruit émis par ces avions est de 78 dB avec un écart type égal à 7 dB. En tenant ces affirmations pour vraies, calculer la probabilité que l'aéroport verse à tort des indemnités aux riverains à la suite du test conçu précédemment.

Date: 6 octobre 2020.

Exercice 2. Dans l'*Amour de l'Art*, Pierre Bourdieu et Alain Darbel réalisent des tableaux de contingence des réponses à différentes questions en fonction de l'appartenance à une classe sociale. On trouve par exemple page 210 les réponses par classe sociale à la question suivante :

Avez-vous des livres sur l'art ? Si oui, lesquels ?

Les réponses ont été codées en quatre catégories :
— Sans réponse
— Non
— Oui sans préciser lesquels
— Oui en précisant lesquels

Les auteurs obtiennent alors le tableau suivant :

	SR	Non	Oui (non précis)	Oui (précis)	Total	Effectifs
Classes populaires	0%	69,5%	13%	17,5%	100%	53
Classes moyennes	0%	40%	15%	45%	100%	98
Classes supérieures	2%	9%	15%	74%	100%	99

On souhaite tester l'indépendance entre la variable « classe sociale » et la variable « possession d'un livre d'art » à partir de ce tableau.

(1) Reconstruire les effectifs observés dans chaque cellule n_{ij} à partir des fréquences en ligne.

(2) Calculer les effectifs théoriques $\tilde{n}_{ij}$ sous l'hypothèse d'indépendance des deux distributions.

(3) Calculer pour chaque cellule du tableau la contribution à la statistique du χ^2.

(4) Réaliser le test d'indépendance du χ^2 à 95% : on précisera soigneusement le choix des hypothèses H_0 et H_1 et la signification du risque de première espèce.

(5) Comment calculer la p-valeur du test, et quel est l'intérêt d'une telle valeur ?

Exercice 3. On étudie l'arrivée des appels dans un central téléphonique en comptabilisant le nombre d'appels reçus par seconde pendant 200 secondes consécutives. On suppose les observations indépendantes. Les résultats obtenus sont répertoriés ci-dessous.

Nombre d'appels par seconde	0	1	2	3	4	5	6	7	8	9	10	11
Effectifs observés	6	15	40	42	37	30	10	9	5	3	2	1

On donne à présent les tables suivantes :

	0	1	2	3	4
$\mathbf{P}_{3,7}(\mathrm{X} = k)$	0,0247	0,0915	0,1692	0,2087	0,1931

	5	6	7	8	$\geqslant 9$
$\mathbf{P}_{3,7}(\mathrm{X} = k)$	0,1429	0,0881	0,0466	0,0215	<0,01

où X est une variable aléatoire de loi $\mathcal{P}(3{,}7)$ sous $\mathbf{P}_{3,7}$.

(1) Tester au niveau 5% l'adéquation de la loi empirique du nombre d'appels reçus par seconde à une loi de Poisson de paramètre 3,7.

(2) Tester au niveau 5% l'adéquation de la loi empirique du nombre d'appels reçus par seconde à la famille des lois de Poisson.

Exercice 4 (égalité des moyennes)**.** Désireux de trouver à se loger pour l'année scolaire à venir, un étudiant collecte des informations sur les loyers des studios mis en location à Cachan. Il obtient 6 observations à partir d'annonces dans une agence immobilière et 300 observations sur un site spécialisé. On suppose que les observations réalisées sont toutes indépendantes et que celles qui constituent le sous-échantillon 1 (resp. 2) suivent une même loi gaussienne de moyenne inconnue μ_1 (resp. μ_2) et d'écarts-types respectifs 187 et 300.

(1) Proposer un test d'égalité des moyennes pour ces deux sous-échantillons.

(2) La moyenne des loyers observés à l'agence est égale à 493, celle des loyers observés sur le site est égale à 550. Évaluer la p-value du test compte tenu de ces observations.