

Fiche de TD n°5 : Tests

1] Test de Student et de Fisher : exercice à savoir faire et à penser à utiliser dans les autres exercices.

Soit (X_1, \dots, X_n) un n échantillon de loi $\mathcal{N}(\theta, \sigma^2)$. Soit (x_1, \dots, x_n) une observation de ce n -échantillon.

Pour $u = (u_1, \dots, u_n) \in \mathbb{R}^n$, on note $\bar{u}_n = \frac{1}{n} \sum_{i=1}^n u_i$ et $s_n^2(u) = \frac{1}{n-1} \sum_{i=1}^n (u_i - \bar{u}_n)^2$.

1) Le premier but de l'exercice est de construire un test pur de niveau $\alpha \in (0, 1)$ dans le problème de test de $\mathcal{H}_0 : \theta \geq \theta_0$ contre l'alternative $\mathcal{H}_1 : \theta < \theta_0$.

Soit t le quantile d'ordre α de la loi $\mathcal{T}(n-1)$.

Montrer que le test pur de région de rejet

$$R_{Student} = \{u \in \mathbb{R}^n : \bar{u}_n \leq \theta_0 + t \frac{s_n(u)}{\sqrt{n}}\},$$

appelé test de Student, est de niveau α . La procédure de décision consiste donc à rejeter \mathcal{H}_0 au niveau α lorsque $(x_1, \dots, x_n) \in R_{Student}$.

2) Le deuxième but de l'exercice est de construire un test pur de niveau $\alpha \in (0, 1)$ dans le problème de test de $\mathcal{H}_0 : \sigma \geq \sigma_0$ contre l'alternative $\mathcal{H}_1 : \sigma < \sigma_0$.

Soit t le quantile d'ordre α de la loi χ_{n-1}^2 .

Montrer que le test pur de région de rejet

$$R_{Fisher} = \{u \in \mathbb{R}^n : s_n^2(u) \leq \sigma_0^2 \frac{t}{n-1}\},$$

appelé test de Fisher, est de niveau α . La procédure de décision consiste donc à rejeter \mathcal{H}_0 au niveau α lorsque $(x_1, \dots, x_n) \in R_{Fisher}$.

2] L'objectif de l'étude est de préciser si la teneur en sucre d'un soda respecte les recommandations d'un organisme de santé indépendant, à savoir une valeur inférieure à 0.100 g/ml. Dans des conditions expérimentales assurant l'indépendance des observations, 10 bouteilles de la production sont analysées. Sur cette observation, on observe un taux moyen de 0.103 g/ml de sucre et un écart-type de 0.010 g/ml. On suppose que les 10 observations sont issues d'une même loi gaussienne.

1) Préciser le modèle statistique.

2) Construire le test en se plaçant du point de vue du directeur de l'usine de fabrication du soda. Quelle serait sa conclusion au niveau 5% ?

3) Construire le test en se plaçant du point de vue de l'organisme de santé. Quelle serait sa conclusion au niveau 5% ?

3] Soit X_1, X_2, \dots, X_n issus d'une loi $\mathcal{E}(\lambda)$. On souhaite tester $\mathcal{H}_0 : \lambda = 1/2$ vs. $\mathcal{H}_1 : \lambda = 1$. Quelle est la région de rejet au niveau 0,05 pour le test de Neyman Pearson simple ?

4] Soit à tester $\mathcal{H}_0 : \lambda = \lambda_0$ vs. $\mathcal{H}_1 : \lambda \neq \lambda_0$ pour le paramètre λ de la loi $\mathcal{E}(\lambda)$ à partir d'un échantillon de taille n . Établir formellement le test du test des vraisemblances maximales. Application : établir le test pour $\lambda_0 = 1/4$ et $n = 30$ en utilisant la loi asymptotique du test des vraisemblances maximales.

5] Soit une loi mère $\mathcal{P}(\lambda)$ où λ est inconnu. On veut tester $\mathcal{H}_0 : \lambda = \lambda_0$ vs. $\mathcal{H}_1 : \lambda \neq \lambda_0$. Montrer que la région de rejet pour le test des vraisemblances maximales est de la forme $x \in [c_1, c_2]^c$ avec une certaine contrainte liant c_1 et c_2 . Application : pour $\lambda_0 = 5$ et $n = 10$ résoudre au plus proche de la solution du test des vraisemblances maximales en utilisant une région de rejet conservatrice par rapport au niveau 0,05. Calculer le niveau exact de cette règle.

6] Afin de tester la satisfaction des clients à un service donné, on effectue un sondage et on définit une variable aléatoire Y_i de la façon suivante :

$Y_i = 1$ si le client i est satisfait

$Y_i = 0$ si le client i n'est pas satisfait

A l'aide d'un échantillon (Y_1, \dots, Y_n) de même loi de Bernoulli de paramètre λ on désire tester les hypothèses $\mathcal{H}_0 : \lambda = 0.58$ vs. $\mathcal{H}_1 : \lambda = 0.48$.

1) Construire la vraisemblance des observations y_1, \dots, y_n et expliciter la région de rejet de \mathcal{H}_0 du test de Neyman-Pearson (pour l'application numérique, on choisira un risque de première espèce $\lambda = 0.1$).

2) Déterminer la puissance de ce test.

7 Un producteur de pneus envisage de changer la méthode de fabrication. La distribution de la durée de vie de ses pneus traditionnels est connue : moyenne 64 000 km, écart-type 8 000 km ; elle est pratiquement gaussienne. Dix pneus sont fabriqués avec la nouvelle méthode et une moyenne de 67 300 km est constatée. En supposant que la nouvelle fabrication donnerait une distribution à peu près gaussienne et de même variance, testez l'efficacité de la nouvelle méthode au niveau $\alpha = 0,05$. Tracez la fonction puissance de ce test. (aide : test de $\mathcal{H}_0 : \lambda < \lambda_0$ contre $\mathcal{H}_1 : \lambda \geq \lambda_0$)

8 Une enquête sur la gêne causée par la proximité d'un aéroport a donné, par sexe, les résultats suivants :

Gêne \ Sexe	Sexe		
	Femmes	Hommes	tous
Aucune	75	35	110
Faible	25	27	52
Moyenne	17	8	25
Forte	3	12	15

Identifier la situation d'échantillonnage et poser l'hypothèse nulle correspondant à la question informelle : la gêne est-elle identique pour les deux sexes ? Tester cette hypothèse nulle.

9 On donne, pour une agglomération, la répartition du nombre de jours sans accident, avec un accident etc., parmi 50 jours d'observation tirés au hasard dans une année :

nombre d'accidents	nombre de jour
0	21
1	18
2	7
3	3
4	1

Tester que la répartition du nombre quotidien d'accidents suit une loi de Poisson. Aide : on effectuera le test du khi-deux en regroupant les catégories de façon à ne pas avoir de fréquences inférieures à 5. Pour simplifier on estimera λ par l'estimation usuelle du maximum de vraisemblance.

10 On lance 450 fois un dé à 6 faces. On obtient les résultats suivants

Numéro de la face observé	1	2	3	4	5	6
Effectif observé	62	50	76	68	111	83

Est ce que le dé est bien équilibré ? (On pourra faire un test du chi 2 d'adéquation).

11 Dans un hôpital psychiatrique, on a recensé les résultats de 490 patients face à un certain médicament, on a noté la saison pendant laquelle le médicament a été pris ainsi que le résultat.

Saison	Réaction au médicament	non réaction au médicament
Printemps	55	64
Eté	59	60
Automne	52	63
Hiver	60	77

Y-a-t-il un lien entre la saison et la réaction au médicament au seuil 5% ? (On pourra faire un test du chi 2 d'indépendance).