

Exercice 1.

On veut tester l'efficacité d'un nouveau traitement contre les migraines. On dispose d'un échantillon de 18 personnes sujettes aux migraines à qui on fournit une quantité égale de pilules correspondant au nouveau traitement (A) et de pilules d'aspirine standard (B). Lorsqu'ils ont utilisé l'intégralité des deux jeux de pilules on demande à chaque patient de juger quel type de pilule (A ou B) a été le plus efficace. Sur les 18 patients, 12 déclarent que le nouveau traitement (A) est plus efficace que l'ancien (B). Comment tester l'efficacité du nouveau traitement ? (on développera ce test et on donnera l'expression de sa p-valeur)

Exercice 2.

On considère un échantillon X_1, \dots, X_n de variables iid réelles de loi diffuse, et $m \in \mathbb{R}$. On souhaite tester l'hypothèse H_0 : "la médiane de la loi des X_i est égale à m ", contre l'hypothèse alternative H_1 : "la médiane de la loi des X_i est strictement inférieure à m ". On considère la statistique de test

$$S = \sum_{i=1}^n \mathbb{I}_{X_i > m}$$

1. Montrer que sous H_0 , la statistique S est libre en loi de X , et qu'on a alors $S \sim \text{Bin}(n, 1/2)$.
2. Le 11 janvier 2014, la vitesse maximale sur le périphérique parisien est passée de 80 km/h à 70 km/h. Un conducteur s'interroge sur l'impact de cette mesure sur son temps de trajet total en voiture depuis une ville de la banlieue Sud. Une longue série de données a permis de montrer qu'avec un périphérique limité à 80 km/h, le temps de trajet médian est de 38 min. Suite au passage à une vitesse limitée à 70 km/h, il recueille un échantillon de $n = 23$ temps de trajet en minutes :

26 29 31 26 32 29 37 37 29 37 30 47 41 37 45 35 35 36 32 30 43 30 35.

Ayant le sentiment que son temps de trajet a diminué, il souhaite tester l'hypothèse nulle que ces observations sont des réalisations d'une loi dont la médiane est toujours de 38 minutes, contre l'hypothèse alternative que la médiane est moindre. Calculer la p-valeur du test à l'aide de la table ci-dessous.

k	3	4	5	6	7	8	9	10	11	12
$P[S \leq k]$	0.0002	0.0013	0.0053	0.0173	0.0466	0.105	0.2024	0.3388	0.5	0.6612
k	13	14	15	16	17	18	19	20	21	22
$P[S \leq k]$	0.7976	0.895	0.9534	0.9827	0.9947	0.9987	0.9998	1	1	1

Exercise 3.

On dit qu'une variable aléatoire V suit la loi de Pareto de paramètres $(m, \alpha) \in R^{*+} \times R^{*+}$ si la densité de la loi de V , notée $f_{m,\alpha}$ satisfait pour tout $x \in R$,

$$f_{m,\alpha}(x) = \alpha m^\alpha x^{-\alpha-1} \mathbb{1}_{x>m}$$

On observe $Z = (Z_1, \dots, Z_n)$ un n-échantillon de loi de Pareto de paramètres $(m, 1)$.

1. Déterminer \hat{m} l'estimateur du maximum de vraisemblance de $m \in R^+$.
2. Quelle est la loi de l'échantillon $m^{-1}Z = (m^{-1}Z_1, \dots, m^{-1}Z_n)$?
3. On considère la statistique de test suivante :

$$h_n(Z) = \sup_{t \geq \hat{m}} \left| \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{Z_i > t} - \left(\frac{t}{\hat{m}}\right)^{-1} \right|$$

Montrer que si la loi des Z_i appartient à \mathcal{F}_1 , la famille des distributions de Pareto de paramètres m et 1, alors la loi de la statistique $h_n(Z)$ ne dépend pas de m .

4. Proposer un test de niveau α pour tester si la distribution d'une suite i.i.d. d'observations appartient à \mathcal{F}_1 (donner la définition de ce test et justifier votre réponse).