Feuille d'exercices n°4.

Exercice 1. On désire comparer un nouveau médicament à un médicament existant. On se place du point de vue du patient, qui veut savoir si le nouveau médicament est plus efficace que son médicament habituel. Le taux de guérisons avec ce dernier est de 60%. Par ailleurs, sur 100 individus testés, 72 ont été guéris avec le nouveau médicament.

1. Proposer un modèle statistique adapté aux données et mettre en place un test afin de savoir si le nouveau médicament est plus efficace que le médicament habituel. Conclure lorsque le niveau du test est fixé à 5%.
2. Calculer une approximation de la puissance du test. Tracer un graphe illustrant cette fonction puissance.

Exercice 2. Soit $(\mathbb{R}^n, \{\mathbb{P}_\theta^{\otimes n}\}_{\theta>0})$ où $\mathbb{P}_\theta$ admet comme densité par rapport à la mesure de Lebesgue la fonction

$$f_\theta(x) = \theta^{-1}x^{-1-1/\theta}\mathbf{1}_{[1,\infty[}(x).$$

On note $(X_1, ..., X_n)$ un échantillon associé.

1. Déterminer l'estimateur $\hat{\theta}_n$ du maximum de vraisemblance pour θ. Cet estimateur est-il sans biais, consistant, UMVU ?
2. Pour n grand, donner un intervalle de confiance de niveau $1-\alpha$ ($\alpha \in]0,1[$).
3. On veut tester $H_0 : \theta = 3$ contre $H_1 : \theta = 5/2$. Montrer que la stratégie optimale consiste à décider H_1 si $\hat{\theta}_n < C_n$ et H_0 sinon. On suppose que $n = 50$ et on fixe l'erreur de première espèce à 5%. Calculer :
   - une approximation de C_n
   - une approximation de l'erreur de seconde espèce.
4. On cherche ensuite à tester $H_0 : \theta \geq 3$ contre $H_1 : \theta < 3$. On reprend la procédure de test précédente. Calculer une approximation de la fonction puissance du test associé.
5. On cherche ensuite à tester $H_0 : \theta = 3$ contre $H_1 : \theta \neq 3$. Pourquoi la procédure de test de la question 3 ne convient pas ? Proposer une procédure de test de niveau $\alpha = 5\%$.

Exercice 3. On considère le modèle statistique $(\mathbb{R}^n, \{\mathbb{P}_\theta^{\otimes n}\}_{\theta\in\mathbb{R}})$ où $\mathbb{P}_\theta$ est la loi $\mathcal{N}(\theta, 1)$. On veut tester $H_0 : \theta = 0$ contre $H_1 : \theta = m < 0$.

1. Donner la forme du test de Neyman-Pearson de niveau $\alpha \in]0, 1[$ pour ce problème. Ce test est-il UPP(α) ?
2. (a) Calculer la fonction puissance $m \in \mathbb{R}_-^* \mapsto \pi_n(m)$ de ces tests et tracer son graphe.

   (b) Etudier la convergence simple de π_n lorsque n tends vers l'infini. Peut-on parler de convergence uniforme sur $\mathbb{R}_-^*$?
3. On considère l'alternative dépendant de n, $H_1 : \theta = -Cn^{-\gamma}$, avec $C > 0, \gamma \in \mathbb{R}$. Etudier le comportement de la puissance du test en fonction de γ lorsque n tend vers l'infini.

Exercice 4. On dispose d'un échantillon de n revenus, supposés distribués suivant une loi de Pareto de densité

$$f_{a,k}(x) = \frac{ak^a}{x^{a+1}} \mathbf{1}_{[k,\infty[}(x).$$

Les deux paramètres $a > 0, k > 0$ sont inconnus.

1. Estimer ces paramètres par la méthode du maximum de vraisemblance.
2. On désire tester l'hypothèse $H_0 : a = 1$ contre $H_1 : a \neq 1$. Montrer que tout test du rapport de vraisemblance admet une région de rejet de la forme $\{\phi \leq r_1\} \cup \{\phi \geq r_2\}$, $0 < r_1 < r_2$, pour une fonction $\phi : (\mathbb{R}_+^*)^n \to \mathbb{R}_+$ que l'on précisera.

Exercice 5. Soient $(X_1, ..., X_n)$ et $(Y_1, ..., Y_n)$ deux échantillons de lois exponentielles indépendantes, de paramètres respectifs $\theta > 0$ et $\theta' > 0$ inconnus. On désire tester l'hypothèse $H_0 : \theta = \theta'$ contre $H_1 : \theta \neq \theta'$.

1. Déterminer un test de rapport de vraisemblance.
2. Montrer que ce test peut-être fondé sur la statistique

$$\phi : (x_1, ..., x_n, y_1, ..., y_n) \to \frac{\sum_{i=1}^n x_i}{\sum_{i=1}^n x_i + \sum_{i=1}^n y_i}.$$

3. Déterminer la loi de ϕ sous H_0.

Exercice 6. (suite de l'exercice 5, feuille 3). On reconsidère le cas de la radio généraliste s'intéressant au taux d'audience du segment étudiant. On reprend en particulier la dernière question, où on se demandait si l'audience sur le segment étudiant avait augmentée de manière significative entre deux sondages espacés de 6 mois. Mettre en place un test statistique pour répondre à cette question. Conclure.