TD n°5 – Tests fondés sur la vraisemblance

stoehr@ceremade.dauphine.fr

Exercice 1. Soit $(X_1, \ldots, X_n)$ un n-échantillon de la loi de Weibull $W(\theta, a, k)$, de paramètres réels a, θ, k strictement positifs, dont la densité est définie pour $x \in \mathbb{R}$ par :

$$f(x \mid \theta, a, k) = \frac{k}{\theta}(x-a)^{k-1} \exp\left(-\frac{(x-a)^k}{\theta}\right) \mathbb{1}_{\{x>a\}}.$$

Dans la suite, on suppose que les paramètres a et k sont connus. Le seul paramètre inconnu est alors θ. On considère le problème de test suivant

$$H_0 : \theta = \theta_0 \quad \text{contre} \quad H_1 : \theta = \theta_1, \quad \text{avec} \quad 0 < \theta_0 < \theta_1.$$

1. Montrer que la variable aléatoire $Y = 2(X-a)^k/\theta$ suit la loi $\chi^2(2)$.

2. **(a)** Montrer que l'estimateur du maximum de vraisemblance de θ est

$$\widehat{\theta} = \frac{1}{n}\sum_{i=1}^{n}(X_i - a)^k.$$

(b) En déduire un intervalle de confiance bilatère pour θ de niveau exactement $1-\alpha$, $\alpha \in]0,1[$.

3. Montrer que l'on a une région critique de la forme

$$\mathscr{R} = \left\{(X_1, \ldots, X_n) : \sum_{i=1}^{n}(X_i - a)^k \geq c\right\}, \quad \text{avec } c \text{ une constante.}$$

4. Déterminer un test de taille $\alpha \in]0,1[$.

5. Exprimer, en fonction de θ_0, θ_1 et d'un quantile de la loi $\chi^2(2n)$, le risque de seconde espèce et la puissance de ce test.

6. Préciser comment varient la région critique, le risque de seconde espèce et la puissance de ce test en fonction de α puis de θ_1.

Exercice 2. On teste sur une population de $n = 75$ individus la capacité à réussir un exercice de proprioception. L'expérience repose sur le modèle statistique $\mathscr{P} = \{\mathscr{B}(p) \mid p \in]0,1[\}$ et on note $X = (X_1, \ldots, X_n)$ un n-échantillon suivant ce modèle. On a obtenu les résultats suivant $\overline{X}_n = 0.476$ et on souhaite réaliser le test statistique suivant

$$H_0 : p = 0.45 \quad \text{et} \quad H_1 : p = 0.5.$$

1. Justifier que la zone de rejet pour un test de taille $\alpha \in]0,1[$ est de la forme

$$\mathscr{R}_\alpha = \left\{ X_1, \dots, X_n : S_n = \sum_{i=1}^{n} X_i > k_\alpha \right\}.$$

2. Calculer l'erreur de première espèce, l'erreur de seconde espèce et la puissance du test en fonction de k_α.

3. Est-il possible de construire un test de taille α pour tout $\alpha \in]0,1[$?

4. Donner la conclusion, l'erreur de seconde espèce et la fonction puissance du test pour les niveaux de test suivants : 1%, 5%, 10%.

5. Calculer la p-valeur. Que pouvez-vous conclure ?

Exercice 3. Parmi les familles suivantes, lesquelles sont à rapport de vraisemblance monotone ?

1. $(X_1, \cdots, X_n)$ est un n-échantillon de la loi $\mathcal{N}(\mu, \sigma^2)$, avec $\mu \in \mathbb{R}$ inconnu et $\sigma^2 \in \mathbb{R}_+^*$ connu.

2. $(X_1, \cdots, X_n)$ est un n-échantillon de la loi $\mathcal{N}(\mu, \sigma^2)$, avec $\mu \in \mathbb{R}$ connu et $\sigma^2 \in \mathbb{R}_+^*$ inconnu.

3. $(X_1, \cdots, X_n)$ un n-échantillon de la loi $\mathscr{E}(\theta)$, $\theta \in \mathbb{R}_+^*$.

4. X de loi $\mathscr{B}(n, p)$ avec $n \in \mathbb{N}$ connu et $p \in]0,1[$ inconnu.

Exercice 4. Soit $(X_1, \dots, X_n)$ un n-échantillon de la loi de Weibull $W(\theta, a, k)$, de paramètres réels a, θ, k strictement positifs, où a et k sont supposés connus. On considère le problème de test suivant

$$H_0 : \theta \le 1 \quad \text{contre} \quad H_1 : \theta > 1.$$

1. Proposer un test uniformément plus puissant (UPP) de taille $\alpha \in]0,1[$.

2. Donner la p-valeur de ce test en fonction de la fonction de répartition F d'une loi du χ^2.

3. Quelle commande utilise-t-on pour évaluer F avec R ?

4. **Application numérique.** Soient $a = 0$, $k = 0.5$, $n = 15$ et $\sum_{i=1}^{n} \sqrt{x_i} = 20.23$. On donne également les valeurs suivantes : $F(20.23) = 0.0895853$ et $F(40.46) = 0.9037421$. Calculer la p-valeur pour ce test et en déduire la décision au niveau $\alpha = 5\%$.

Exercice 5 ***(À connaître – Test gaussien).*** On souhaite mettre sur le marché un médicament B pour remplacer le médicament A, dont le but est de soulager la douleur après une opération chirurgicale. Ce médicament B ne sera mis sur le marché que s'il est significativement plus efficace que le médicament A. On a collecté les données (fictives) résumées ci-dessous.

	Médicament A	**Médicament** B
Taille de l'échantillon	$n = 55$	
Moyenne empirique (en heures)	$\bar{X}_A = 4.64$	$\bar{X}_B = 4.03$
Écart-type empirique (en heures)	$S_A = 1.25$	$S_B = 1.82$

On fait l'hypothèse des modèles statistiques suivant : $\{\mathcal{N}(\mu_A, \sigma^2) \mid \mu_A \in \mathbb{R}, \sigma \in \mathbb{R}_+^*\}$ pour le médicament A et $\{\mathcal{N}(\mu_B, \sigma^2) \mid \mu_B \in \mathbb{R}, \sigma \in \mathbb{R}_+^*\}$ pour le médicament B. On note $X = (X_1, \ldots, X_n)$ le n-échantillon observé pour le médicament A et $Y = (Y_1, \ldots, Y_n)$ le n-échantillon observé pour le médicament B.

Partie I – On suppose que la variance est connue : $\sigma^2 = 2$.

1. Donner la loi de $\overline{X}_n - \overline{Y}_n$.
2. Écrire les hypothèses H_0 et H_1 et construire un test uniformément plus puissant de taille $\alpha \in]0,1[$.
3. Déterminer l'erreur de seconde espèce et la fonction puissance de ce test de taille α.
4. Donner la conclusion, l'erreur de seconde espèce et la fonction puissance du test pour les tailles de test suivantes : 1%, 5%, 10%.
5. Calculer la p-valeur. Que pouvez-vous conclure ?

Partie II – On suppose à présent que la variance est inconnue.

1. Construire un test de taille $\alpha \in]0,1[$.
2. Déterminer l'erreur de seconde espèce et la fonction puissance de ce test de taille α.
3. Donner la conclusion, l'erreur de seconde espèce et la fonction puissance du test pour les tailles de test suivantes : 1%, 5%, 10%.
4. Calculer la p-valeur. Que pouvez-vous conclure ?

Exercice 6. Une usine fabrique des machines. On suppose que le temps aléatoire X au bout duquel une machine tombe en panne suit une loi exponentielle $\varepsilon(\lambda)$, $\lambda \in \mathbb{R}_+^*$ inconnu. Les machines sont vendues à une grande surface avec laquelle elle passe un contrat sur la qualité des machines fournies. Sur ce contrat on veut que la probabilité pour qu'une machine tombe en panne avant un instant donnée T soit plus petite qu'une quantité donnée. On dit que la convention est respectée si $\lambda \leq \lambda_0$, où $\lambda_0 = 1/2$. Pour s'en assurer, le fournisseur lance régulièrement des études dans lesquelles il observe au hasard n machines. On note $X_1, \ldots, X_n$ le n-échantillon suivant la loi $\varepsilon(\lambda)$ correspondant aux temps de panne observés. Lors de la dernière étude, pour $n = 12$ machines, il obtient $\overline{X}_n = 2.61$.

1. Écrire les hypothèses H_0 et H_1 et construire un test uniformément plus puissant de taille $\alpha \in]0,1[$.
2. Déterminer l'erreur de seconde espèce et la fonction puissance de ce test de taille α.
3. Donner la conclusion, l'erreur de seconde espèce et la fonction puissance du test pour les tailles de test suivantes : 1%, 5%, 10%.
4. Calculer la p-valeur. Que pouvez-vous conclure ?
5. Le point de vue de la grande surface serait-il le même que celui du fournisseur ?