

Exercices de Statistiques

Université de Lorraine

Estimation et théorie des tests

Clément Dell'Aiera

1 Principe de Neyman : décision à 2 points

1. Soit f la densité d'une loi de probabilité sur \mathbb{R} , et \mathcal{E} l'expérience statistique engendré par un n -échantillon de loi $p_\theta(x) = f(x - \theta)$. On suppose que $\Theta = \{0, \theta_0\}$ avec $\theta_0 \neq 0$. On veut tester $H_0|\theta = 0$ contre $H_1|\theta = \theta_0$.
 - (a) Décrire l'expérience statistique et donner la vraisemblance du modèle.
 - (b) Donner la zone de rejet du test de Neyman-Pearson de niveau α associé à H_0 et H_1 .
2. L'expérimentateur observe une seule réalisation d'une v.a. X de loi de Poisson de paramètre $\theta > 0$. On veut tester $H_0|\theta = \theta_0$ contre $H_1|\theta = \theta_1$, où $\theta_0 \neq \theta_1$.
 - (a) Donner la zone de rejet du test de Neyman-Pearson de niveau α associé.
 - (b) Sachant que $\mathbb{P}_{\theta_0}(X > 9) = 0.032$ et $\mathbb{P}_{\theta_1}(X > 8) = 0.068$, donner une zone de rejet explicite pour $\alpha = 0.05 = 5\%$. Le test est-il optimal ?

2 Neyman-Pearson : familles à rapport de vraisemblance monotone

1. Soit \mathcal{E} l'expérience statistique engendrée par un n -échantillon de loi normale $\mathcal{N}(\theta, \sigma^2)$, où σ^2 est connu, et $\theta \in \Theta = \mathbb{R}$. On souhaite tester $H_0|\theta = \theta_0$ contre $H_1|\theta = \theta_1$, où $\theta_0 < \theta_1$.
 - (a) Décrire le modèle ainsi que la vraisemblance. On choisira la mesure de Lebesgue comme mesure dominante.
 - (b) Calculer le rapport de vraisemblance

$$\frac{f(\theta_1, Z)}{f(\theta_0, Z)}.$$

- (c) Donner la zone de rejet pour le test de Neyman-Pearson associé.

2. Pour la même expérience statistique, on a un test optimal (uniformément plus puissant) de H_0 contre H_1 donné par la région de rejet

$$\mathcal{R} = \{\bar{X}_n > c\}$$

où c est solution de $\mathbb{P}_{\theta_0}(\bar{X}_n > c) = \alpha$.

- (a) Calculer explicitement la valeur de la constante $c = c(\theta_0, \alpha)$.
 (b) Calculer la puissance de ce test.

3 Exercice

L'expérimentateur observe 2 échantillons indépendants X_1, \dots, X_n et Y_1, \dots, Y_m de tailles distinctes $n \neq m$, de lois respectives $\mathcal{N}(\mu_1, \sigma_1^2)$ et $\mathcal{N}(\mu_2, \sigma_2^2)$. Il souhaite tester

$$H_0 : \mu_1 = \mu_2 \quad \text{contre} \quad H_1 : \mu_1 \neq \mu_2.$$

Si $s_{n,1}^2 = \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X}_n)^2$ et $s_{m,2}^2 = \frac{1}{m} \sum_{j=1}^m (Y_j - \bar{Y}_m)^2$, construire un test basé sur la statistique

$$T_{n,m} = \frac{\bar{X}_n - \bar{Y}_m}{\sqrt{s_{n,1}^2 + s_{m,2}^2}}$$

et étudier sa consistance.

4 Neyman-Pearson : loi exponentielle

On observe un n -échantillon $\underline{x} = (X_1, \dots, X_n)$ de variables iid de loi exponentielle de paramètre $\lambda > 0$, de densité

$$x \mapsto \lambda \exp(-\lambda x) 1_{x \geq 0}.$$

1. Rappeler l'espérance et la variance (les calculer si besoin) d'une loi exponentielle de paramètre λ . On rappelle que $2\lambda \sum_{j=1}^n X_j$ suit alors une loi du χ^2 à $2n$ degrés de liberté.
2. Ecrire le modèle statistique engendré par l'observation \underline{x} .
3. Calculer l'estimateur du maximum de vraisemblance $\hat{\lambda}^{MV}$ de λ .
4. Montrer que $\hat{\lambda}^{MV}$ est asymptotiquement normal, et calculer sa variance limite.
5. Soient $0 < \lambda_0 < \lambda_1$. Construire un test d'hypothèse de

$$H_0 : \lambda = \lambda_0 \quad \text{contre} \quad H_1 : \lambda = \lambda_1$$

de niveau α et uniformément plus puissant. Expliciter le choix du seuil définissant la région critique. Montrer que le test est consistant, i.e. que l'erreur de seconde espèce du test tend vers 0 lorsque $n \rightarrow \infty$.

5 Intervalles de confiance

Soient $\mu \in \mathbb{R}$, $\sigma > 0$ et $\alpha \in (0, 1)$. On observe un n -échantillon $\underline{x} = (X_1, \dots, X_n)$ de variables iid de loi normale $\mathcal{N}(\mu, \sigma^2)$.

1. Donner un intervalle de confiance de niveau α pour μ , si σ^2 est connu, puis si σ^2 est inconnu.
2. Donner un intervalle de confiance de niveau α pour σ^2 , si μ est connu, puis si μ est inconnu.