3. Théorèmes asymptotiques. Tests, Théorème de Neyman-Pearson

Objectifs : Savoir appliquer la loi des grands nombres, le théorème central limite, la méthode delta et le Lemme de Slutsky. Savoir appliquer le théorème de Neyman-Pearson pour des test d'hypothèses simples. Les exercices 3.1 à 3.2 sont à faire pendant le TD, les 3.3 à 3.6 sont à chercher de votre côté.

Exercice 3.1 (Théorèmes asymptotiques)**.** Supposons que $X_1, \ldots, X_n$ sont i.i.d., centrées, de variance finie $\sigma^2 > 0$ inconnue. On note classiquement $\bar{X}$ la moyenne empirique et $\hat{\sigma}^2$ l'estimateur non biaisé de la variance.

1. Écrire $\hat{\sigma}^2$ en fonction des X_i.
2. Étudier la convergence presque sûre de $\bar{X}$ et de $\hat{\sigma}^2$.
3. Montrer que
$$\frac{\sqrt{n}\bar{X}}{\sqrt{\hat{\sigma}^2}} \xrightarrow[n\to\infty]{(d)} \mathcal{N}(0,1)\,.$$
.
4. On suppose ici que $\sigma^2 = 1$ et qu'il est connu. Donner la limite en loi de
$$\sqrt{n}(e^{\bar{X}/2} - 1)\,.$$

.

Exercice 3.2 (Test optimal pour la loi exponentielle)**.** Soient $X_1, \ldots, X_n$ des variables aléatoires i.i.d. de loi exponentielle de paramètre[2] $\lambda > 0$. On souhaite tester les hypothèses simples $\mathcal{H}_0 : \lambda = \lambda_0$ contre $\mathcal{H}_1 : \lambda = \lambda_1$ avec $\lambda_1 > \lambda_0$.

On rappelle la propriété suivante : la somme de n variables aléatoires indépendantes de loi exponentielle de paramètre $\lambda = 1$ suit une loi Gamma $\Gamma(n,1)$, dont on notera $q_\beta^{\Gamma,n}$ le quantile d'ordre β.

1. Déterminer un test uniformément plus puissant (UPP) de niveau $\alpha \in]0,1[$ pour ce problème. Montrer que c'est un test dont la région de rejet peut s'écrire
$$R = \left\{ \sum_{i=1}^n X_i < \frac{u(\alpha)}{\lambda_0} \right\},$$
où $u(\alpha)$ est une constante dépendant de α dont on ne demande pas l'expression dans cette question.
2. Montrer que sous $\mathcal{H}_0$, $\lambda_0 X_1 \sim \mathrm{Exp}(1)$. En déduire la valeur de $u(\alpha)$.
3. On notera F_n la fonction de répartition de la loi $\Gamma(n,1)$. Exprimer la puissance $\pi(\lambda_1)$ du test précédent pour toute valeur de $\lambda_1 > \lambda_0$, en fonction de F_n. Commenter sa monotonie en λ_1.
4. Montrer par ailleurs que le test précédent est uniformément plus puissant (UPP) pour tester $\mathcal{H}_0 : \lambda = \lambda_0$ contre $\mathcal{H}_1 : \lambda > \lambda_0$.

Exercice 3.3 (Asymptotique du maximum de vraisemblance)**.** Reprenons le modèle où les $X_1, \ldots, X_n$ sont i.i.d. et dont la loi est à densité sur $\mathbb{R}$ donnée par
$$x \mapsto \mathbb{1}_{x\geq 1}\frac{a-1}{x^a}$$

[2] on rappelle que c'est loi à densité $x \mapsto \mathbb{1}_{x\geq 0}\lambda e^{-\lambda x}$.

avec $a > 1$ un paramètre. On a établi que l'estimateur du maximum de vraisemblance de a est donné par

$$\hat{a}_{MV} = 1 + \left(\frac{1}{n}\sum_{i=1}^{n} \log X_i\right)^{-1}$$

dans l'exercice 2.3.

1. Etudier la distribution de $\log X_1$ en calculant sa fonction de répartition, et en déduire que $\hat{a}_{MV}$ est également consistant.
2. Calculer l'information de Fisher $I_1(a)$ de X_1. En admettant que le modèle est régulier, montrer que $\hat{a}_{MV}$ est asymptotiquement normal et donner ses caractéristiques.

Exercice 3.4 (Test de variance symétrique à deux échantillons)**.** *On admettra dans cet exercice que si Z_1, Z_2 sont deux variables i.i.d. gaussiennes standard, alors Z_1/Z_2 suit une loi de Cauchy de densité sur $\mathbb{R}$ donnée par $x \mapsto \frac{1}{\pi(1+x^2)}$.*

On travaille avec deux échantillons indépendants, X_1 et X_2. On suppose que $0 < \sigma_1^2 < \sigma_2^2$, et on souhaite tester

$$\mathcal{H}_0 : X_1 \sim \mathcal{N}(0, \sigma_1^2),\ X_2 \sim \mathcal{N}(0, \sigma_2^2)$$

contre

$$\mathcal{H}_1 : X_1 \sim \mathcal{N}(0, \sigma_2^2),\ X_2 \sim \mathcal{N}(0, \sigma_1^2)\,.$$

On souhaite travailler avec des tests symétriques ϕ, tels que $\phi(x_1, x_2) = 1 - \phi(x_2, x_1)$ presque partout.

1. Montrer que si ϕ est un test symétrique de niveau α et de puissance β, alors $\alpha = 1 - \beta$.
2. Écrire le rapport de vraisemblance L, et construire un estimateur symétrique ϕ_0 basé sur L pour ce problème.
3. Montrer que le niveau $\alpha(\phi_0)$ et la puissance $\beta(\phi_0)$ de ϕ_0 vérifient :

$$\alpha(\phi_0) = 1 - \beta(\phi_0) = \frac{2}{\pi}\tan^{-1}\left(\frac{\sigma_1}{\sigma_2}\right)\,.$$

Exercice 3.5 (Plus vite que $\sqrt{n}$?)**.** Reprenons l'exercice 3.1, en supposant que $\sigma^2 = 1$ et qu'il est connu.

1. Quelle est la limite en loi de $\sqrt{n}(\cos(\bar{X}) - 1)$?
2. Trouver une suite $(a_n)_{n\geq 1}$ telle que $a_n(\cos(\bar{X}) - 1)$ converge en loi vers une v.a. Z qui n'est pas presque sûrement constante, dont on précisera la loi.

Exercice 3.6. Reprenons l'exercice 3.2.

1. Sous quelle condition sur a ces variables sont-elles d'espérance finie ?
2. Sous la condition de la question 1, donner un estimateur $\hat{a}_{MM}$ de a via la méthode des moments. Prouver qu'il est consistant.
3. On suppose dans toute la suite que la condition de la question 1 est vérifiée. A l'aide du TCL et de la méthode Delta, montrer que $\hat{\alpha}_{MM}$ est asymptotiquement normal et donner ses caractéristiques.
4. Pour n grand, quel est l'estimateur à privilégier selon vous, entre $\hat{a}_{MM}$ et $\hat{a}_{MV}$?