Feuille de TD 2

Exercice 1 : (loi de Poisson) Soit $X_1, \ldots, X_n$ des variables aléatoires i.i.d suivant une loi de Poisson de paramètre λ et

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{et} \quad S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2$$

1. Montrer que $\overline{X}_n$ est un estimateur sans biais et asymptotiqyuement normal de λ.
2. Montrer que S_n^2 est un estimateur sans biais de λ et montrer sa normalité asymptotique. On rappelle que $\mathbb{E}\left[(X_1-\lambda)^4\right] = \lambda + 3\lambda^2$.
3. Quel estimateur privilégier ?
4. Montrer que
   (a) $\sqrt{n}\frac{\overline{X}_n-\lambda}{\sqrt{\overline{X}_n}} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)$
   (b) $\sqrt{n}\frac{\overline{X}_n-\lambda}{S_n} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)$
   (c) $\sqrt{n}\left(g\left(\overline{X}_n\right) - g(\lambda)\right) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)$ pour un bon choix de g.
5. Déterminer les intervalles de confiance correspondant, lequel est le meilleur ?

Exercice 2 : (loi uniforme) Soit $X_1, \ldots, X_n$ un échantillon de loi uniforme sur $[0,\theta]$.

1. Le modèle est-il régulier ?
2. Par la méthode des moments, proposer un estimateur de θ. Montrer sa consistance et sa normalité asymptotique. En déduire un intervalle de confiance asymptotique de niveau $1-\alpha$.
3. Soit $\theta_0 > 0$, proposer un test de niveau asymptotique α pour tester

$$H0 : \theta = \theta_0 \qquad \text{contre} \qquad H_1 : \theta \neq \theta_0.$$

4. Calculer l'estimateur du maximum de vraisemblance et calculer son risque quadratique
5. Calculer la loi limite de $n\left(\theta - \theta_n^{MV}\right)$.
6. Déterminer $c_{\alpha,n}$ tel que $\left[X_{(n)}, c_{\alpha,n}X_{(n)}\right]$ soit un intervalle de confiance de niveau $1-\alpha$.
7. Calculer la médiane de X_1 et en déduire un nouvel estimateur de θ. Donner sa normalité asymptotique ainsi qu'un nouvel intervalle de confiance asymptotique.

8. Quel intervalle choisiriez vous ?
9. Soit $\lambda_0 > 0$. En déduire un test de niveau α pour tester

$$H0 : "\lambda = \lambda_0" \qquad \text{contre} \qquad H1 : "\lambda \neq \lambda_0".$$

Exercice 3 : (Loi de Pareto) Soit $X_1, \ldots, X_n$ des variables aléatoires i.i.d de densité de Pareto

$$f_\theta(x) = \frac{\theta}{x^{\theta+1}} \mathbf{1}_{x \geq 1}.$$

avec $\theta > 0$.

1. On suppose $\theta > 1$, estimer θ par la méthode des moments. A-t-on consistance ? normalité asymptotique ?
2. Que se passe-t-il si $\theta \in (0, 1)$?
3. Calculer la médiane de X et en déduire un nouvel estimateur de θ, pour tout $\theta > 0$. Donner sa normalité asymptotique.
4. Donner l'estimateur du maximum de vraisemblance pour $\theta > 0$. Est-il asymptotiquement normal ?
5. Le modèle est-il régulier ? L'estimateur du maximum de vraisemblance est-il asymptotiquement efficace ?

Exercice 4 : lois translatées

1. Soit $\theta \in \mathbb{R}$ et une variable aléatoire X de densité $f_\theta(x) = f(x - \theta)$ par rapport à la mesure de Lebesgue, pour une certaine densité f sur $\mathbb{R}$ connue et un paramètre inconnu $\theta \in \mathbb{R}$. Montrer que si f est de classe C^1 et si

$$I = \int_{\mathbb{R}} \frac{(f'(x))^2}{f(x)} \mathbf{1}_{f(x)>0} dx < +\infty,$$

alors le modèle $(f_\theta)_{\theta \in \mathbb{R}}$ est régulier d'information de Fisher $I(\theta) = I$.

2. Pour chacun des modèles suivants, dire si le modèle est régulier et si oui calculer l'information de Fisher :
   (a) $f_\theta(x) = \exp(-(x-\theta)^2/2)/(2\pi)^{1/2}, \theta \in \mathbb{R}$;
   (b) $f_\theta(x) = \theta^{-1} \mathbf{1}_{[0,\theta]}(x), \theta > 0$;
   (c) $f_\theta(x) = c_1(1 - (x-\theta)^2) \mathbf{1}_{[\theta-1,\theta+1]}(x), \theta \in \mathbb{R}$, pour une constante $c_1 > 0$ à déterminer ;
   (d) $f_\theta(x) = c_2(1 - (x-\theta)^2)^2 \mathbf{1}_{[\theta-1,\theta+1]}(x), \theta \in \mathbb{R}$, pour une constante $c_2 > 0$ à déterminer.

Exercice 5 : (Loi exponentielle translatée) On observe un échantillon $X_1, \ldots, X_n$ dont la loi admet la densité

$$f_\theta(x) = \exp(-(x-\theta)) \mathbf{1}_{[\theta,+\infty[}(x),$$

où θ est un paramètre réel inconnu.

1. Rappeler la loi de $n(X_{(1)} - \theta)$. En déduire un intervalle de confiance pour θ de niveau $1 - \alpha$.
2. Soit $\alpha \in]0,1[$. On souhaite tester au niveau α

$$H_0 : \theta \geq 0 \text{ contre } H_1 : \theta < 0.$$

   (a) Construire un test à partir de l'intervalle de confiance de la question 2, calculer sa puissance et donner son allure (pour n et α fixés). Quelle est sa taille $\alpha^\star$?

   (b) Proposer un autre test qui soit, lui, de taille α.

   (c) Calculer la fonction puissance du test. La représenter en fonction de θ pour n et α fixés.

   (d) Comment varie la puissance en fonction de α ? en fonction de n ?

Exercice 6 : Soit $\theta > 0$ un paramètre inconnu. On considère la densité f_θ définie pour tout $x \in \mathbb{R}$ par

$$f_\theta(x) = \frac{\theta}{x^2}\mathbf{1}_{x \geq \theta}.$$

Dans ce qui suit, on considère des variables aléatoires $X_1, \ldots, X_n$ indépendantes et de densité f_θ.

1. Calculer la fonction de répartition de X_1.
2. Calculer la médiane de X_1 et en déduire un estimateur de θ. Est-il consistant ? Asymptotiquement normal ?
3. Par la méthode des moments, proposer un estimateur. Est-il consistant ? Asymptotiquement normal ?
4. Déterminer l'estimateur du maximum de vraisemblance $\hat{\theta}_n$. Déterminer la loi limie de $n\left(\hat{\theta}_n - \theta\right)$.
5. Le modèle $(f_\theta)_{\theta > 0}$ est-il régulier ?
6. Pour tout $\epsilon > 0$, calculer

$$\mathbb{P}\left[\hat{\theta}_n \geq (1+\epsilon)\theta\right].$$

   En déduire un intervalle de confiance non asymptotique de niveau $1 - \alpha$.

Exercice 7 : Estimation de deux paramètres Soit $Y \sim \mathcal{E}(\lambda)$ avec $\lambda > 0$ inconnu. Soit $X = Y + \theta$ avec $\theta > 0$ inconnu. On admettra que X a pour densité $f_{\theta,\lambda}$ définie pour tout x par

$$f_{\theta,\lambda}(x) = \lambda \exp\left(-\lambda(x-\theta)\right)\mathbf{1}_{[\theta,+\infty[}(x).$$

Soient $X_1, \ldots, X_n$ des variables aléatoires i.i.d de même loi que X.

1. Calculer $\mathbb{E}[X]$ et $\mathbb{V}[X]$.
2. On suppose θ "fixé", le modèle (f_λ) avec $f_\lambda = f_{\theta,\lambda}$ est il régulier ?
3. On suppose λ "fixé", le modèle (f_θ) avec $f_\theta = f_{\theta,\lambda}$ est-il régulier ?
4. Déterminer l'estimateur du maximum de vraisemblance de θ.

5. Calculer la fonction de répartition de $X_{(1)}$.
6. En déduire l'erreur quadratique moyenne de $X_{(1)}$.
7. Déduire de la question 5

$$\sqrt{n}\left(X_{(1)}-\theta\right) \xrightarrow[n\to+\infty]{\mathbb{P}} 0.$$

Que pouvez-vous en déduire ?

8. En déduire un estimateur de λ.
9. Montrer que l'estimateur de λ est consistant. Pour s'aider, on admettra que si (A_n) et (B_n) sont deux suites de variables aléatoires convergeant en probabilités vers a et b, et $g : I \times J \longrightarrow \mathbb{R}$ est une fonction continue en (a, b), alors

$$g\left(A_n, B_n\right) \xrightarrow[n\to+\infty]{\mathbb{P}} g(a,b).$$

avec I, J des intervalles ouvert de $\mathbb{R}$.

10. Montrer que

$$\sqrt{n}\left(\overline{X}_n-\hat{\theta}_n-\frac{1}{\lambda}\right) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}\left(0,\frac{1}{\lambda^2}\right).$$

11. En déduire la normalité asymptotique de l'estimateur de λ.
12. Soit $\alpha \in (0,1)$, en déduire un intervalle de confiance asymptotique de λ.
13. Montrer que

$$n\left(X_{(1)}-\theta\right) \sim \mathcal{E}(\lambda).$$

14. Pour tout $\alpha \in (0,1)$, donner le quantile $q_{\lambda,1-\alpha}$ d'ordre $1-\alpha$ de la loi exponentielle de paramètre λ.
15. Donner la convergence de

$$n\left(X_{(1)}-\theta\right)-\left(q_{\hat{\lambda}_n,1-\alpha}-q_{\lambda,1-\alpha}\right).$$

avec $q_{\hat{\lambda}_n} = \frac{-\ln(\alpha)}{\hat{\lambda}_n}$.

16. Déterminer un intervalle de confiance asymptotique de niveau $1-\alpha$ pour θ.

Exercice 8 (Loi Bêta dilatée et translatée) : Soit $X_1,\ldots,X_n$ i.i.d. de densité :

$$x \in \mathbb{R} \mapsto f_k(x,\theta) = c_k\left[1-(x-\theta)^2\right]^{k-1}\mathbf{1}_{[\theta-1,\theta+1]}(x),$$

où $k \geq 1$ est connu et $c_k = \frac{\Gamma(2k)}{\Gamma(k)^2 2^{2k-1}}$. Cette densité est celle d'une variable aléatoire qui s'obtient par homothétie de rapport 2 et translation de $\theta-1$ d'une variable de loi $\mathrm{Beta}(k,k)$.
Dans toute la suite, on appellera $\hat{\theta}_n$ un estimateur du maximum de vraisemblance et on étudiera ses propriétés.

Partie 1 On se place tout d'abord dans la situation où $k > 2$.

1. Existence de $\hat{\theta}_n$.
   (a) Pour tout $\theta \in \mathbb{R}$, montrer que l'intervalle $[X_{(n)} - 1, X_{(1)} + 1]$ est d'intérieur non-vide $\mathbb{P}_\theta$-p.s. et que si $\hat{\theta}_n$ existe, alors $\hat{\theta}_n \in [X_{(n)} - 1, X_{(1)} + 1]$.
   (b) Montrer que pour $\theta \in]X_{(n)} - 1, X_{(1)} + 1[$, la dérivée $\ell_n'(\theta)$ de la log-vraisemblance s'écrit :
   $$\ell_n'(\theta) = (k-1)\left(\sum_{i=1}^n \frac{1}{1-(X_i-\theta)} - \sum_{i=1}^n \frac{1}{1+(X_i-\theta)}\right).$$
   Prouver que
   i. $\ell_n'(\theta) \xrightarrow[\theta \to X_{(n)}-1]{} \infty$;
   ii. $\ell_n'(\theta) \xrightarrow[\theta \to X_{(1)}+1]{} -\infty$;
   iii. $\theta \in]X_{(n)} - 1, X_{(1)} + 1[\mapsto \ell_n'(\theta)$ est strictement décroissante.
   (c) En déduire que l'estimateur du maximum de vraisemblance est l'unique solution de $\ell_n'(\theta) = 0$. Montrer que la résolution de $\ell_n'(\theta) = 0$ nécessite de rechercher les racines d'un polynôme de degré élevé.
2. Consistance de $\hat{\theta}_n$.
   (a) Montrer que l'on peut trouver une constante C (ne dépendant pas de θ) telle que, pour $\epsilon \in [0, 1]$:
   $$\mathbb{P}_\theta(X_1 \leq \theta - 1 + \epsilon) \geq C\epsilon^k.$$
   En déduire que :
   $$\mathbb{P}_\theta\left((X_{(1)} + 1) - \theta \geq \epsilon\right) = \mathbb{P}_\theta\left(X_{(1)} - (\theta - 1) \geq \epsilon\right) \leq (1 - C\epsilon^k)^n.$$
   (b) Par symétrie, montrer que pour tout $\epsilon \in [0, 1]$:
   $$\mathbb{P}_\theta(\theta + 1 - x \leq X_1) \geq C\epsilon^k$$
   et
   $$\mathbb{P}_\theta\left(\theta - (X_{(n)} - 1) \geq \epsilon\right) = \mathbb{P}_\theta\left((\theta + 1) - X_{(n)} \geq \epsilon\right) \leq (1 - C\epsilon^k)^n.$$
   (c) En déduire que $\hat{\theta}_n$ est consistant. Est-il fortement consistant ?
3. Le modèle statistique $\{f_k(\cdot, \theta), \theta \in \mathbb{R}\}$ est-il régulier ? Si oui, calculer l'information de Fisher et conclure quant à la normalité asymptotique de $\hat{\theta}_n$.
   Indication : Si $\alpha > 0$ et $\beta > 0$, alors $\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$.

Partie 2 On suppose maintenant que $k \leq 2$.

1. Que dire des propriétés de $\hat{\theta}_n$?

2. Soit $\alpha > 0$. À l'aide de la question 2, montrer que pour tout $\epsilon > 0$:

$$\mathbb{P}_\theta\left(n^\alpha|\hat{\theta}_n - \theta| > \epsilon\right) \leq 2\left(1 - C\frac{\epsilon^k}{n^{\alpha k}}\right)^n.$$

3. Conclure sur la vitesse de convergence de $\hat{\theta}_n$. Quand $k < 2$, cette convergence est-elle plus rapide que lorsque le modèle est régulier (i.e. $k > 2$) ? Qu'en est-il lorsque $k = 2$?