Devoir de Statistiques

Exercice 1

Un statisticien observe un n-échantillon $X_1, \ldots, X_n$ de la loi $\mathbb{P}_{\theta,\lambda}$ telle que

$$\frac{\mathrm{d}\mathbb{P}_{\theta,\lambda}}{\mathrm{d}x} = \lambda \mathbb{1}_{[0,1]}(x) + \frac{1-\lambda}{\theta}\mathbb{1}_{[0,\theta]}(x),$$

où $\lambda \in [0,1[$ et $\theta \geq 1$ sont inconnus.

1) Montrer que si l'on exclut une valeur de θ à préciser, le modèle est identifiable.
2) Montrer que $\widehat{\theta} = \max_i X_i$ est un estimateur (faiblement) consistant de θ.
3) Calculer $\mathbb{E}_{\theta,\lambda}(X_1)$ et proposer un estimateur $\widehat{\lambda}$ de λ.
4) Montrer que si $\theta > 1$, alors $\widehat{\lambda}$ est un estimateur consistant. Que devient la convergence de $\widehat{\lambda}$ si $\theta = 1$?
5) Expliquer comment simuler une variable aléatoire X de loi $\mathbb{P}_{\theta,\lambda}$.
6) Illustrer cet exercice sous Scilab. On pourra simuler des échantillons, construire les estimateurs et observer leur consistance.
7) *Facultatif : trouver une procédure de test pour tester $\theta = 1$ contre $\theta > 1$.*

Exercice 2

On a testé le taux de cholestérol sanguin de 25 hommes et on a obtenu les valeurs suivantes

164 272 261 248 235 192 203 278 268 230 242 305 286
310 345 289 326 335 297 328 400 292 194 338 252

On suppose qu'on est dans un modèle gaussien. Donner un intervalle de confiance à 90% pour la moyenne et un autre pour la variance. On pourra chercher les quantiles nécessaires dans Scilab.

Exercice 3

On observe $X_1, \ldots, X_N$ indépendantes avec $X_1, \ldots, X_n$ i.i.d. de loi $\mathcal{P}(\lambda^a)$ et $X_{n+1}, \ldots, X_N$ i.i.d. de loi $\mathcal{P}(\lambda^b)$ où λ est inconnu et $a \geq 1$ et $b \geq 1$ sont connus. On pose $m = N - n$.

1) Ecrire l'équation définissant l'estimateur du maximum de vraisemblance $\widehat{\lambda}_N$ de λ.
2) Calculer l'information de Fisher $I_N(\lambda)$.
3) On suppose à présent que $a = 1$ et $b = 2$, ainsi que $N = 2m = 2n$. Calculer explicitement $\widehat{\lambda}_N$, puis calculer $\lim_{N\to\infty} I_N(\lambda)/N = I(\lambda)$.
4) Prouver que $T_n = \sum_{i=1}^n X_i + 2\sum_{i=n+1}^N X_i$ obéit à un théorème de la limite centrale.
5) En déduire que $\sqrt{N}(\widehat{\lambda}_N - \lambda)$ converge en loi vers une normale centrée de variance $\sigma^2(\lambda)$. Comparer $\sigma^2(\lambda)$ à $I^{-1}(\lambda)$.
6) Illustrer la convergence en loi ci-dessus sous Scilab.