Statistique inférentielle - TD 5

Exercice 1.

Un statisticien observe $X_1, ..., X_n$ i.i.d. dont la loi admet la densite

$$g_{\theta,\lambda}(x) = \lambda \mathbf{1}_{[0,1]}(x) + \frac{1-\lambda}{\theta}\mathbf{1}_{[0,\theta]}(x),$$

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

1. Montrer que, si on exclut une valeur de θ a preciser, le modèle est identifiable. Sauf indication contraire, on supposera cette condition verifiée dans la suite.
2. Determiner l'EMV de λ lorsque θ est connu, l'EMV de θ lorsque λ est connu, puis l'EMV de (θ, λ) lorsque les 2 paramètres sont inconnus.
3. Etudier la convergence et la loi limite de l'EMV de θ, que λ soit ou non connu.
4. Etudier la consistance de l'EMV de λ, lorsque θ connu, puis lorsque θ est inconnu. Que se passe-t-il si $\theta = 1$?
5. Construire un test de $H_0 : \theta = 1$ contre $H_1 : \theta > 1$ au niveau α. A $\lambda \in [0,1[$ et $\theta > 1$ fixés, étudier la limite de la puissance lorsque n tend vers l'infini.

Exercice 2.

On cherche à analyser la différence de revenu selon le sexe dans une entreprise. Pour ceci, on modélise le revenu R_i du i-ème individu d'un échantillon de n personnes par :

$$R_i = \mu + \beta s_i + \varepsilon_i,\ 1 \leq i \leq n,$$

ou

- on observe les variables aleatoires $R_1, ..., R_n$;
- le vecteur $(s_1, ..., s_n)$ est deterministe avec $s_i = 1$ si le $i-$ème individu est une femme, et $s_i = 0$ sinon ;
- $\varepsilon = (\varepsilon_i)_{1\leq i\leq n}$ est un vecteur gaussien centre de covariance $\sigma^2 I_n$ avec $\sigma^2 > 0$ connue ;
- le parametre $\theta = (\mu, \beta) \in \mathbb{R}^2$ est inconnu.

1. Comment s'interprète le coefficient β ?
2. Calculer la vraisemblance de $R = (R_1, ..., R_n)$. A quelle condition sur $(s_1, ..., s_n)$ l'estimateur du maximum de vraisemblance $\hat{\theta} = (\hat{\mu}, \hat{\beta})$ de θ est-il bien défini ? On supposera dans la suite que cette condition est vérifiée.
3. Déterminer, en fonction de σ^2 et de $(s_1, ..., s_n)$, la loi de $\hat{\beta}$. En déduire un intervalle de confiance pour β.
4. Construire un test de niveau α de
   H_0 : ≪ le salaire ne depend pas du sexe ≫
   contre
   H_1 : ≪ le salaire depend du sexe. ≫
   Calculer sa puissance. Comment varie-t-elle en fonction de α ? en fonc- tion de β ?
5. Proposer un test de niveau α de
   H_0 : ≪ le salaire des hommes est plus eleve que celui des femmes ≫
   contre
   H_1 : ≪ le salaire des hommes est moins eleve que celui des femmes. ≫

Exercice 3.

On observe un échantillon $X_1, ..., X_n$ dont la loi admet la densité $g_\theta(x) = exp(-(x - \theta))\mathbf{1}_{[\theta,+\infty[}(x)$, ou θ est un parametre reel inconnu.

1. Quels estimateurs de θ pouvez-vous proposer en utilisant les méthodes usuelles ?
2. Construire un intervalle de confiance pour θ.
3. Soit $\alpha \in]0, 1[$. On souhaite tester au niveau α :
   $H_0 : \theta \geq 0$ contre $H_1 : \theta < 0$.
   (a) Donner la forme d'une région de rejet.
   (b) Calculer la fonction puissance du test. En déduire la taille du test, puis définir précisément la région de rejet.
   (c) Comment varie la puissance en fonction de α ? en fonction de n ?
4. Proposer un test de niveau α de
   H_0 : « La loi de X_1 est une loi exponentielle » contre H_1 : « La loi de X_1 n'est pas une loi exponentielle (tout en restant dans le modèle paramétrique précédent) »
   Calculer la puissance de ce test.