Estimation paramétrique

Exercice 1. Soit $\mathcal{P} = \{\mathbb{P}_\theta, \theta \in \Theta\}$ un modèle paramétrique telle que la loi de probabilité $\mathbb{P}_\theta$ admette une variance σ^2 inconnue (elle dépend évidemment de θ).

1. Montrer que la variance empirique $V_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2$ est un estimateur biaisé mais asymptotiquement sans biais de σ^2.
2. En déduire que la variance empirique corrigée $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2$ est un E.S.B. de σ^2.

Exercice 2. On rappelle que, par définition, si $Y_1, ..., Y_n$ sont i.i.d. de loi $\mathcal{N}(0,1)$, alors $\sum_{i=1}^n |Y_i|^2$ suit la loi du χ^2 à n degrés de libertés, notée $\chi^2(n)$.

1. Montrer que si $Y \sim \mathcal{N}(0,1)$, alors $\mathbb{E}[Z^4] = 3$. En déduire que si $Z \sim \chi^2(d)$ alors $\mathbb{E}[Z] = d$ et $\mathrm{Var}(Z) = 2d$.
2. On considère le modèle gaussien bi-dimensionnel $\mathcal{P} = \{\mathbb{P}_\theta = \mathcal{N}(\theta), \theta = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}^{+*}\}$. On note comme dans l'exercice précédent V_n^2 et S_n^2 les variance empirique et variance empirique corrigée de l'échantillon $(X_1, ..., X_n)$. On rappelle que $(n-1)S_n^2/\sigma^2 \sim \chi^2(n-1)$. Calculer le risque quadratique des deux estimateurs S_n^2 et V_n^2 de σ^2. En déduire que V_n^2 est préférable à S_n^2.

Exercice 3. On considère le modèle uniforme continu $\mathcal{P} = \{\mathbb{P}_\theta = \mathcal{U}(0,\theta) : \theta > 0\}$ et un échantillon associé $(X_1, ..., X_n)$.

1. On pose $\hat{\theta}_n = 2\bar{X}_n$. Montrer que $\hat{\theta}$ est un E.S.B. de θ et calculer son risque quadratique.
2. On rappel que $\hat{\theta}_n^{MV} = \max_{1\leqslant i \leqslant n} X_i$ est l'E.M.V. de θ.
   (a) Montrer que $\hat{\theta}_n^{MV}$ est une variable à densité, de densité $g(x) = \frac{n}{\theta^n} x^{n-1} \mathbb{1}_{[0,\theta]}(x)$.
   (b) Montrer que $\hat{\theta}_n^{MV}$ est un estimateur biaisé de θ. Est-il asymptotiquement sans biais ?
   (c) Montrer que $T_n = \frac{n+1}{n}\hat{\theta}_n^{MV}$ est un E.S.B. de θ.
3. Comparer les estimateurs $\hat{\theta}_n$, $\hat{\theta}_n^{MV}$ et T_n.

Exercice 4. On considère le modèle paramétrique $\mathbb{P}_\theta = \mathcal{E}(\frac{1}{\theta})$ avec $\theta > 0$ et $(X_1, ..., X_n)$ un échantillon associé.

— Soit $\theta > 0$ et $X \sim \mathcal{E}(\frac{1}{\theta})$. Déterminer la valeur $g(\theta) = \mathbb{P}_\theta(X > \theta^2)$.

— On note N_n le nombre de variables aléatoires X_i de l'échantillon $(X_1, ..., X_n)$ qui dépassent θ^2. Donner la loi de N_n. En déduire un estimateur T_n de $g(\theta)$.

— On pose $T_n' = e^{-\bar{X}_n}$. Montrer que T_n' est un estimateur biaisé de $g(\theta)$. Est-il asymptotiquement sans biais ?

Exercice 5. Soit $(X_1, ..., X_n)$ un n-échantillon d'une loi d'espérance μ et de variance σ^2. Quelle(s) condition(s) doivent vérifier les réels a_i pour que $\sum_{i=1}^n a_i X_i$ soit un estimateur sans biais de μ ? Déterminer parmi ces estimateurs sans biais celui qui est de variance minimale et donner sa variance.

Exercice 6. Soient Y et Z deux variables aléatoires indépendantes suivant des lois exponentielles de paramètres inconnus λ et μ. On considère un échantillon $((Y_1, Z_1), ..., (Y_n, Z_n))$ de même loi que (Y, Z) et on suppose qu'on observe uniquement les variables aléatoires $X_i = \min(Y_i, Z_i)$.

1. Quelle est la loi de X_i ?
2. Le modèle statistique est il identifiable ? Quelle fonction de (λ, μ) est identifiable ?
3. Quel est l'estimateur du maximum de vraisemblance de $\gamma = \lambda + \mu$ fondé sur l'observation des (X_i) ? Quel est l'estimateur du maximum de vraisemblance de $\gamma = \lambda + \mu$ fondé sur l'observation des (Y_i, Z_i) ?
4. Comparer ces deux estimateurs.