

RPM4 2023: TD de statistique

Clément Gauchy

Novembre 2023

Exercice 1:

Soit X_1, \dots, X_n i.i.d. tel que $X_1 \sim \mathcal{U}([-α, α])$ pour $α \geq 0$. On souhaite estimer $α$.

- 1) Quel est le modèle statistique dans ce problème ?
- 2) Ecrivez la vraisemblance de ce modèle statistique pour les données X_1, \dots, X_n
- 3) Etudier le comportement de la vraisemblance en fonction de $α$. En déduire l'estimateur du maximum de vraisemblance $\hat{\alpha}_{MV}$.

Exercice 2: Estimation du temps de détection entre deux particules.

On souhaite estimer le temps d'attente entre la détection de deux photons sur un détecteur. On suppose que le temps d'attente T est une variable aléatoire suivant une loi exponentielle de paramètre λ i.e. sa densité de probabilité est $f_T(t) = \lambda e^{-\lambda t}$. Nous avons en notre possession un jeu de données (T_1, \dots, T_n) i.i.d. de même loi que T .

On définit les lois Gamma $\mathcal{G}(\alpha, \beta)$ par la densité de probabilité $f_{\alpha, \beta}(t) = \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t}$. Sa moyenne est α/β et son mode $(\alpha - 1)/\beta$ pour $\alpha \geq 1$.

- 1) Quel est le modèle statistique ? Selon vous que devons nous estimer dans ce problème ?
- 2) On sait que la loi exponentielle de paramètre λ admet $1/\lambda$ comme moyenne. Proposez un estimateur de λ par méthode des moments.
- 3) Écrivez la vraisemblance de ce modèle. Déterminez l'estimateur du maximum de vraisemblance $\hat{\lambda}_{MV}$ pour λ (utilisez pour cela la log-vraisemblance, on admettra que la log-vraisemblance est concave). Que remarquez vous ?
- 4) On admet que $\sum_{i=1}^n T_i \sim \mathcal{G}(n, \lambda)$. Calculez $\mathbb{E} \left[\frac{1}{\sum_{i=1}^n T_i} \right]$ avec un calcul d'intégrale (on utilisera la propriété $\Gamma(n) = (n - 1)\Gamma(n - 1)$).
- 5) A l'aide de la question 4), calculez le biais de $\hat{\lambda}_{MV}$.

On va désormais utiliser des méthodes bayésiennes. On utilisera comme loi à priori pour λ une loi Gamma de paramètre α et β .

- 6) À l'aide de la règle de Bayes, calculez la loi *a posteriori* de λ sachant (T_1, \dots, T_n) .
- 7) Déterminez l'estimateur de la moyenne et du mode *a posteriori* de λ .

Exercice 3: Estimateur Monte-Carlo pour une étude de radioprotection

On considère un matériau homogène de section efficace macroscopique d'absorption μ pour une particule non spécifiée. On ne considèrera que l'interaction par absorption dans cet exercice. On souhaite protéger un opérateur d'un terme source émetteur de particules énergétiques à l'aide de ce matériau. On suppose que l'opérateur est situé derrière une épaisseur L de ce matériau.

Soit X la variable aléatoire désignant la longueur parcourue dans le matériau par la particule.

- 1) Rappelez sans démonstration la loi de X en utilisant les informations du paragraphe ci dessus
- 2) Pour la radioprotection de l'opérateur, une quantité très importante est $p = \mathbb{P}(X > L)$. Que représente cette quantité ? Pourquoi est elle importante ?
- 3) On suppose que l'on sait générer $(X_i)_{1 \leq i \leq N}$ i.i.d. suivant la loi de X , écrivez \hat{p}_{MC} l'estimateur Monte-Carlo de p .
- 4) Ecrivez la variance de \hat{p}_{MC} . En supposant que $p = 10^{-6}$, combien doit valoir N pour que le coefficient de variation de \hat{p}_{MC} soit de 1% ? (Pour rappel, le coefficient de variation est le ratio écart-type sur moyenne).

On va construire un estimateur par échantillonnage d'importance (*importance sampling*) pour réduire la variance de \hat{p}_{MC} . Pour $0 < \theta < \mu$, on définit la densité d'échantillonnage $h_\theta(x) = (\mu - \theta)e^{-(\mu - \theta)x}$.

- 5) De quelle loi h_θ est elle la densité de probabilité ? Quelle interprétation physique peut on faire si on diminue la section efficace macroscopique μ par θ ?
- 6) On génère $(Y_i)_{1 \leq i \leq N}$ selon la densité de probabilité h_θ . Ecrire l'estimateur par échantillonnage d'importance \hat{p}_{IS} .
- 7) Ecrivez la variance de \hat{p}_{IS} . **Bonus:** Proposez une borne supérieure de cette variance sachant que $e^{-\theta x} < e^{-\theta L}$ pour $x > L$. Trouvez les conditions sur θ, μ et L pour que $\text{Var}(\hat{p}_{IS}) < \text{Var}(\hat{p}_{MC})$