 lerouvillois@math.univ-lyon1.fr

Fiche de TD N^o 3 : ESTIMATION

Exercice 1.

Un organisme de transport en commun souhaite étudier le retard occasionné par ses bus. On constate que le retard d'un bus (temps entre l'horaire prévu et l'arrivée effective d'un bus) peut être modélisé par une variable aléatoire de loi uniforme sur $[0, \theta]$ et que les retards des différents bus sont indépendants. On souhaite estimer le paramètre $\theta > 0$ à partir de l'observation d'un n-échantillon $X = (X_1, ..., X_n)$ de ces retards.

1. Déterminer un estimateur $\hat{\theta}_n$ de θ par la méthode des moments puis $\hat{\theta}'_n$ par la méthode des moindres carrés. Remarquer que ces deux méthodes fournissent le même estimateur.
2. Déterminer l'estimateur T_n du maximum de vraisemblance de θ. Est-ce une statistique exhaustive ? Calculer son espérance et en déduire un nouvel estimateur non biaisé T'_n fonction d'une statistique exhaustive.
3. Comparer les risques quadratiques de $\hat{\theta}_n$ et de T'_n et choisir le meilleur. (On peut montrer que T'_n est le meilleur estimateur non biaisé de θ).
4. Peut-on parler d'estimateur efficace pour ce modèle ?

Exercice 2.

Un agriculteur possède un champ carré dont il veut estimer la superficie. Quand il mesure un côté de son champ, il sait (un ami statisticien lui a confirmé) que l'erreur expérimentale de la mesure est une variable aléatoire de loi normale centrée et de variance σ^2. Il réalise une première mesure de ce côté et trouve une valeur $X_1 = 510$ mètres. Il en déduit une superficie de $S_1 = 26.01$ hectares. Il réalise une deuxième mesure et trouve alors $X_2 = 490$, d'où une valeur de la superficie $S_2 = 24.01$. Il abandonne ses mesures et réfléchit pour savoir quelle est la bonne façon de procéder. Doit-il adopter comme estimation de la surface S_1, S_2 ou un estimation combinant les deux mesures, telle que : $S_3 = X_1 X_2 = 24.99$, $S_4 = \frac{S_1+S_2}{2} = 25.01$ ou $S_5 = \left(\frac{X_1+X_2}{2}\right)^2 = 25$? On se propose d'aider l'agriculteur à résoudre son problème.

1. Préciser le modèle considéré ainsi que la fonction $g(\mu)$ que l'on cherche à estimer. Étudier les cinq estimateurs proposés (biais, variance) et choisir l'estimateur de risque quadratique le plus petit. (Indication : si $X \sim \mathcal{N}(\mu, \sigma^2)$, alors $\mathrm{Var}(X^2) = 2(\sigma^4 + 2\mu^2\sigma^2)$).
2. Donner les estimateurs qui généralisent S_4 et S_5 au cas où l'agriculteur a pu faire n mesures du côté de son champ. Effectuer la même étude qu'à la question 1) pour ces estimateurs. Étudier aussi leur convergence et leur efficacité asymptotique. Conclure.

3. Quel est l'estimateur du maximum de vraisemblance ?

Exercice 3.

Soit $X = (X_1, \ldots, X_n)$ un n-échantillon de loi exponentielle de paramètre 1 et décentrée sur $[\theta, +\infty[$.

1. Déterminer l'estimateur du maximum de vraisemblance de θ. Est-il biaisé ? Est-ce une statistique exhaustive ?
2. En déduire un estimateur sans biais de θ fonction d'une statistique exhaustive. (On peut montrer qu'il s'agit du meilleur estimateur non biaisé de θ).

Exercice 4.

On considère un échantillon $X = (X_1, \ldots, X_n)$ de taille n issu d'une loi de Poisson de paramètre $\theta > 0$. On souhaite estimer le paramètre $e^{-\theta}$.

1. Montrer que $\mathbb{1}_{X_1=0}$ est un estimateur non biaisé de $e^{-\theta}$.
2. Trouver une statistique exhaustive T_n.
3. Calculer $\mathbb{E}[\mathbb{1}_{X_1=0}|T_n]$ et montrer qu'il s'agit d'un estimateur non biaisé de $e^{-\theta}$. (On peut montrer qu'il s'agit du meilleur estimateur non biaisé de $e^{-\theta}$).

Exercice 5.

On possède une pièce de monnaie déséquilibrée qui, quand on la lance, a une probabilité $p \in]0, 1[$ de tomber sur pile et une probabilité $1 - p$ de tomber sur face.

1. On effectue n tirages indépendants de cette pièce. Estimer p par la méthode des moments puis par la méthode du maximum de vraisemblance. Quelles sont les propriétés de l'estimateur ? (biais, variance, efficacité, convergence)
2. On lance la pièce jusqu'à ce qu'elle tombe sur pile et on note Y le nombre de tirages nécessaires.
   a) Quelle est la loi de Y ?
   b) On dispose de n répétitions indépendantes de cette expérience $(Y_1, ..., Y_n)$. Quel est l'estimateur du maximum de vraisemblance de p ? Est-il convergent ?

Exercice 6.

On s'intéresse à la durée de vie de d'un modèle de téléphones portables que l'on peut modéliser par une variable aléatoire exponentielle de paramètre $\theta > 0$ que l'on souhaite estimer. Soit $X = (X_1, \ldots, X_n)$ un n-échantillon de loi exponentielle de paramètre θ.

1. Calculer le Score $S_n(\theta)$ ainsi que l'information de Fisher $I_n(\theta)$.
2. Déterminer l'estimateur du maximum de vraisemblance T_n de $\frac{1}{\theta}$. Est-ce un estimateur efficace de $\frac{1}{\theta}$?

3. Déterminer l'estimateur du maximum de vraisemblance $\hat{\theta}_n$ de θ. Calculer son espérance et en déduire une version $\hat{\theta}'_n$ non biaisée. Est-ce un estimateur efficace de θ ? Est-ce un estimateur asymptotiquement efficace de θ ?

4. Existe-t-il un estimateur non biaisé et efficace de θ ?