Fiche de TD 1 : Méthodes d'estimation

Exercice 1 (Loi Gamma et Méthode des moments)
Soit X une variable aléatoire de Gamma $G(\alpha, \beta)$ dont la densité :

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \mathbb{I}_{\mathbb{R}_+}(x), \qquad \text{avec} \qquad \alpha,\ \beta > 0 \qquad \text{et} \qquad \Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt.$$

1. Calculer $\mathbb{E}(X)$ et $\mathrm{Var}(X)$
2. Supposons que nous disposons un échantillon aléatoire $X_1, \ldots, X_n$ de X, estimer les paramètres α et β par la méthode des moments.

Exercice 2 (Loi exponentielle et Méthode des moments)
Soit X une variable aléatoire de loi exponentielle de paramètre $\lambda > 0$ dont la densité :

$$f(x) = \lambda e^{-\lambda x} \mathbb{I}_{\mathbb{R}_+}(x)$$

1. Calculer $\mathbb{E}(X)$ et proposer un EMM de λ.
2. Calculer $\mathbb{E}(X^2)$ et proposer un nouvel EMM de λ.
3. Soit $t_0 > 0$, on définit $Y_i = \mathbb{I}_{\{X_i > t_0\}}, i = 1, \ldots, n$, étudier la distribution des Y_i et proposer un nouvel EMM de λ basé sur les Y_i.

Exercice 3
Soit $X_1, \ldots, X_n$ un échantillon aléatoire d'une variable aléatoire X dont la distribution dépend d'un paramètre inconnu a. Déterminer un estimateur de a par la méthode des moments dans les cas suivants :

1. La distribution X est donnée par

$$\mathbb{P}(X = 0) = \frac{a}{a+1} \qquad \text{et} \qquad \mathbb{P}(X = 1) = \frac{1}{a+1}, \qquad a \in \mathbb{R}_+^*$$

2. La distribution de X est donnée par la densité suivante :

$$f(x) = \frac{1}{2a} \mathbb{I}_{[0,a]}(x) + \frac{1}{2(1-a)} \mathbb{I}_{]a,1]}(x), \qquad \text{où} \qquad a \in]0, 1[.$$

3. La distribution de X est donnée par la densité suivante :

$$f(x) = \frac{a}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} + \frac{1-a}{\sqrt{4\pi}} e^{-\frac{x^2}{4}}, \qquad x \in \mathbb{R} \qquad \text{et} \qquad a \in]0, 1[.$$

Exercice 4
Soit $X_1, \ldots, X_n$ un échantillon aléatoire d'une variable aléatoire $X \sim \mathcal{U}([0, \theta])$ où $\theta > 0$ un paramètre inconnu.

1. Montrer qu'on peut écrire la fonction de la vraisemblance de cet échantillon sous cette forme :

$$L(x_1, \ldots, x_n; \theta) = \frac{1}{\theta^n} \mathbb{I}_{x_{(1)} \geq 0} \mathbb{I}_{x_{(n)} \leq \theta},$$

   où $x_{(1)} = \min_{i=1\ldots,n} x_i$ et $x_{(n)} = \max_{i=1\ldots,n} x_i$.
2. Montrer que $X_{(n)} = \max_{i=1\ldots,n} x_i$ est l'estimateur du maximum de vraisemblance de θ
3. Déterminer la loi de $X_{(n)}$.
4. Calculer $\mathbb{E}(X_{(n)})$ et en déduire que $X_{(n)}$ est un estimateur asymptotiquement sans biais pour θ.
5. Calculer $\mathrm{Var}(X_{(n)})$ et en déduire que $X_{(n)}$ est un estimateur consistant pour θ.

Exercice 5

Nous nous intéressons à la modélisation de la durée de vie X d'un équipement électronique. Il est pertinent de considérer que cette durée de vie suit une distribution aléatoire de type exponentiel. Cependant, la valeur du paramètre λ associé à cette distribution reste inconnue.

<u>Partie 1</u>

Soit $X_1, \ldots, X_n$ un échantillon aléatoire de X.

1. Donner un estimateur du maximum de vraisemblance de λ.
2. Donner un estimateur du maximum de vraisemblance de la probabilité que la durée de vie d'un matériel soit supérieur à un temps $t_0 > 0$ fixé :

$$\alpha = \mathbb{P}(X > t_0).$$

3. Quels estimateurs de λ et de α obtient-on si on utilise la méthode des moments ?

<u>Partie 2</u>

Imaginons maintenant que les données sur ces durées de vie proviennent de l'expérience suivante : au temps t = 0, un équipement est installé sur un banc d'essai. Lorsqu'il tombe en panne, il est immédiatement remplacé par un équipement identique et neuf, sans prendre en compte le temps nécessaire au remplacement. Ce processus se poursuit jusqu'au temps t_0 fixé. On désigne par K le nombre de pannes observées pendant l'intervalle $[0, t_0]$.

1. Calculer la probabilité que $K = 0$.
2. On note T_k le temps écoulé jusqu'à la k-ème panne observée. C'est à dire que $T_k = X_1 + \ldots + X_k$. Montrer que la loi de la v.a.r. T_k est une Gamma $G(k, \lambda)$ (Ind. On pourra utiliser la transformée de Laplace ou la fonction caractéristique).
3. Exprimer l'événement $K = k$ en fonction d'événements liant les variables aléatoire T_k et X_{k+1}. En déduire que la loi de K est une loi de Poisson, dont on déterminera la valeur du paramètre.

Exercice 6

Un agriculteur possède un champ carré dont il veut estimer la superficie. Quand il mesure un côté de son champ, il sait (un statisticien de passage lui a confirmé), ou il suppose, 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 une estimation combinant les deux mesures, telle que :

$$s_3 = s_1 s_2 = 24.99, \qquad s_4 = \frac{s_1 + s_2}{2} = 25.01 \qquad \text{et} \qquad s_5 = \left(\frac{s_1 + s_2}{2}\right)^2 = 25 \quad ?$$

Faut-il recommencer ses mesures jusqu'à ce qu'il trouve deux résultats identiques, ou bien combiner intelligemment n mesures pour construire des estimations du type s_4 ou s_5 (généralisées à ces n mesures) ?

Soit μ la longueur réelle (inconnue) du côté du champ de l'agriculteur, et notons X la variable aléatoire qui représente la mesure qu'obtient l'agriculteur lorsqu'il mesure le coté de son champ.

1. Préciser la distribution de X et la fonction $g(\mu)$ que l'agriculteur souhaite à estimer.
2. Étudier les cinq estimateurs proposés. On calculera notamment leurs biais, variances et risques quadratiques moyens.
   Indication : Si $X \sim \mathcal{N}(m, \sigma^2)$ *alors* $\mathrm{Var}(X^2) = 2(\sigma^2 + 2m^2\sigma^2)$
3. Quel estimateur proposez-vous à l'agriculteur parmi les cinq estimateurs que vous venez de comparer ?
4. Donner les estimateurs qui généralisent s_4 et s_5 au cas où l'agriculteur a pu faire n mesures du coté de son champ. Effectuer la même étude qu'à la question 2 pour ces estimateurs. Étudier également leurs consistance.
5. Donner l'estimateur du maximum de vraisemblance de $g(\mu)$ et l'étudier s'il est différent de ceux considérés précédemment.