Université de Nantes Année 2013-2014
L3 Maths-Eco

Feuille 3 : Méthodes d'estimation

1 Méthode des moments

Exercice 1 Soit X une variable aléatoire réelle de loi

$$\mathbb{P}(X=0)=\frac{a}{a+1}, \qquad \mathbb{P}(X=1)=\frac{1}{a+1},$$

$n \in \mathbb{N}^*$, et $(X_1,\dots,X_n)$ un n-échantillon de X. Ici, $a>0$ est un réel inconnu que l'on souhaite estimer à l'aide de $(X_1,\dots,X_n)$. Déterminer un estimateur de a par la méthode des moments.

Exercice 2 Soient X une variable aléatoire réelle de densité

$$f(x)=\begin{cases} \dfrac{1}{2a} & \text{si } x\in[0,a], \\ \dfrac{1}{2(1-a)} & \text{si } x\in]a,1], \\ 0 & \text{sinon,} \end{cases}$$

$n \in \mathbb{N}^*$, et $(X_1,\dots,X_n)$ un n-échantillon de X. Ici, $a\in]0,1[$ est un réel inconnu que l'on souhaite estimer à l'aide de $(X_1,\dots,X_n)$. Déterminer un estimateur de a par la méthode des moments.

Exercice 3 Soient X une variable aléatoire réelle de densité

$$f(x)=\begin{cases} (\theta+1)(\theta+2)(1-x)x^\theta & \text{si } x\in[0,1], \\ 0 & \text{sinon,} \end{cases}$$

$n \in \mathbb{N}^*$, et $(X_1,\dots,X_n)$ un n-échantillon de X. Ici, $\theta>-1$ est un réel inconnu que l'on souhaite estimer à l'aide de $(X_1,\dots,X_n)$. Déterminer un estimateur de θ par la méthode des moments.

Exercice 4 Soient X une variable aléatoire réelle de densité

$$f(x)=p\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}+(1-p)\frac{1}{\sqrt{4\pi}}e^{-\frac{x^2}{4}}, \qquad x\in\mathbb{R}.$$

$n \in \mathbb{N}^*$, et $(X_1,\dots,X_n)$ un n-échantillon de X. Ici, $p\in]0,1[$ est un réel inconnu que l'on souhaite estimer à l'aide de $(X_1,\dots,X_n)$. Déterminer un estimateur de p par la méthode des moments.

2 Méthode du maximum de vraisemblance

Exercice 5 Soit X une variable aléatoire réelle de loi

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \qquad k \in \mathbb{N},$$

$n \in \mathbb{N}^*$, et $(X_1, \ldots, X_n)$ un n-échantillon de X. Ici, $\lambda > 0$ est un réel inconnu que l'on souhaite estimer à l'aide de $(X_1, \ldots, X_n)$. Déterminer l'estimateur du maximum de vraisemblance de λ.

Exercice 6 La hauteur maximale en mètres de la crue annuelle d'un fleuve est une variable aléatoire réelle X de densité

$$f(x) = \begin{cases} \frac{x}{a} e^{-\frac{x^2}{2a}} & \text{si } x \geqslant 0, \\ 0 & \text{sinon.} \end{cases}$$

Soient $n \in \mathbb{N}^*$ et $(X_1, \ldots, X_n)$ un n-échantillon de X. Ici, $a > 0$ est un réel inconnu que l'on souhaite estimer à l'aide de $(X_1, \ldots, X_n)$.

1. Déterminer l'estimateur du maximum de vraisemblance $\widehat{a}_n$ de a.
2. *Application.* Une crue supérieure à 6 mètres serait catastrophique. Pendant 8 ans, on a observé les hauteurs de crue du fleuve en mètres. Les résultats sont : $2,5$; $2,9$; $1,8$; $0,9$; $1,7$; $2,1$; $2,2$; $2,8$. À partir de ces mesures, donner une estimation ponctuelle de a et une estimation de la probabilité d'avoir une catastrophe une année donnée.

Exercice 7 Soit $(Y_1, \ldots, Y_n)$ un vecteur de n variables aléatoires réelles telles que, pour tout $i \in \{1, \ldots, n\}$,

$$Y_i = \frac{\theta}{i^\alpha} + \sigma X_i,$$

où $\alpha \geqslant 0$ et $\sigma > 0$ sont des réels connus, $(X_1, \ldots, X_n)$ est un vecteur de n variables aléatoires réelles indépendantes et identiquement distribuées suivant la loi Normale centrée réduite, et $\theta > 0$ est un réel inconnu que l'on souhaite estimer à l'aide des $(Y_1, \ldots, Y_n)$.

1. Déterminer la vraisemblance et la log-vraisemblance de $(Y_1, \ldots, Y_n)$.
2. Déterminer l'estimateur du maximum de vraisemblance $\widehat{\theta}_n$ de θ.
3. Est-il sans biais ? Déterminer les valeurs de α pour lesquelles l'estimateur $\widehat{\theta}_n$ converge.

Exercice 8 Soient X une variable aléatoire réelle de densité

$$f(x) = \begin{cases} (\alpha - 1)x^{-\alpha} & \text{si } x \geqslant 1, \\ 0 & \text{sinon,} \end{cases}$$

$n \in \mathbb{N}^*$, et $(X_1, \ldots, X_n)$ un n-échantillon de X. Le paramètre $\alpha > 2$ est un réel inconnu.

1. Déterminer l'estimateur du maximum de vraisemblance $\widehat{\alpha}_n$ de α.

2. On pose $\beta = \dfrac{\alpha - 2}{\alpha - 1}$. Déterminer l'estimateur du maximum de vraisemblance $\widehat{\beta}_n$ de β.

Exercice 9 Soient X la variable aléatoire réelle de densité

$$f(x) = \begin{cases} \dfrac{3}{(x-\theta)^4} & \text{si } x \geqslant 1+\theta, \\ 0 & \text{sinon.} \end{cases}$$

$n \in \mathbb{N}^*$, et $(X_1, \ldots, X_n)$ un n-échantillon de X. Ici, $\theta > 0$ est un réel inconnu que l'on souhaite estimer à l'aide de $(X_1, \ldots, X_n)$.

1. Calculer $\mathbb{E}(X - \theta)$ et $\mathbb{E}\left((X-\theta)^2\right)$. En déduire $\mathbb{E}(X)$ et $\mathbb{V}(X)$.
2. Déduire du résultat de la question 1- un estimateur $\widetilde{\theta}_n$ de θ en utilisant la méthode des moments. Calculer son risque quadratique.
3. Calculer l'estimateur du maximum de vraisemblance $\widehat{\theta}_n$ de θ.
4. Calculer la fonction de répartition de $\widehat{\theta}_n - \theta$. En déduire sa densité.
5. Calculer le risque quadratique de $\widehat{\theta}_n$. Entre $\widehat{\theta}_n$ et $\widetilde{\theta}_n$, quel est le meilleur estimateur de θ lorsque n est supposé être grand ?
6. Calculer le biais de $\widehat{\theta}_n$, et en déduire un estimateur sans biais $\widehat{\theta}_n^*$ de θ. Montrer, sans calcul intégral, que, lorsque n est supposé être grand, $\widehat{\theta}_n^*$ est un meilleur estimateur de θ que $\widehat{\theta}_n$.

Exercice supplémentaire. Soit $X_1, \ldots, X_n$ des va i.i.d. ayant des moments d'ordre 4. On pose

$$S_n^2 = \sum_{i=1}^{n} (X_i - \bar{X}_i)^2/n$$

Le but de cet exercice est de calculer le risque quadratique $\mathbb{E}((S_n^2 - \sigma^2)^2)$ de l'estimateur S_n^2 de $\sigma^2 = \mathbb{V}(X_1)$. On rappelle que $\sum_{i=1}^{n-1} i = n(n-1)/2$.

1. Montrer que l'on peut supposer sans perte de généralité que les X_i sont centrées. On fera cette hypothèse dans la suite.
2. Montrer que
$$\mathbb{E}((S_n^2 - \sigma^2)^2) = \mathbb{V}(S_n^2) + b^2(S_n^2)$$
où $b(S_n^2) = \mathbb{E}(S_n^2) - \sigma^2$.
3. Démontrer que :
$$S_n^2 = \frac{n-1}{n^2} \sum_{i=1}^{n} X_i^2 - \frac{2}{n^2} \sum_{k<l} X_k X_l$$
En déduire que $b(S_n^2) = -\sigma^2/n$.

4. Montrer que

$$\mathbb{C}(\sum_{i=1}^{n} X_i^2, \sum_{k<l} X_k X_l) = 0, \quad \mathbb{V}(\sum_{k<l} X_k X_l) = n(n-1)\sigma^4/2$$

En déduire que :

$$\mathbb{V}(S_n^2) = \frac{n-1}{n^3}\left((n-1)\mathbb{E}(X_1^4) - (n-3)(\mathbb{E}(X_1^2))^2\right)$$

et calculer la valeur du risque quadratique de S_n^2.

Exercice à rendre. Soit X une variable aléatoire réelle de loi de Bernoulli de paramètre θ, avec $0 < \theta < 1$. Et $(X_1, \ldots, X_n)$ un n-échantillon de X.

1. Déterminer un estimateur de θ par la méthode des moments.
2. Déterminer l'estimateur du maximum de vraisemblance de θ.
3. Les estimateurs précédents sont-il sans biais ?
4. On cherche à estimer la variance $\theta(1-\theta)$. On propose l'estimateur $\hat{V} = \bar{X}(1-\bar{X})$. Etudier le biais de cet estimateur, sa convergence presque sure et la convergence en loi de $\sqrt{n}(\hat{V} - \theta(1-\theta))$.
5. Corriger l'estimateur $\hat{V}$ pour proposer un nouvel estimateur $\hat{V}'$ de $\theta(1-\theta)$ qui est sans biais. Etudier sa convergence presque sure et la convergence en loi de $\sqrt{n}(\hat{V}' - \theta(1-\theta))$.