Master Ingénierie Statistique .

Statistique Non paramétrique .
Anne Philippe
Estimation fonctionnelle

Exercice 1.

Soit X une variable aléatoire continue de densité

$$f(x) = \frac{e^x}{(1+5e^x)^{1.2}} \qquad \forall x \in \mathbb{R}.$$

(1) Calculer la fonction de répartition de la loi de X.
(2) Calculer la fonction de quantile de la loi de X.

Exercice 2.

Soit X une variable aléatoire continue de densité définie par

$$f(x) = \begin{cases} 3(x-1)(2-x) & \text{si } x \in]1,2[\\ 3(x-3)(4-x) & \text{si } x \in]3,4[\\ 0 & \text{sinon} \end{cases}$$

Quelle est la médiane de la loi de X ?

Exercice 3.

Soit $X_1, ..., X_n$ n variables aléatoires iid suivant la loi qui admet pour fonction de répartition F. Soit $a < b$ des nombres fixés tels que $F(a) > 0$ et $f(b) < 1$. On pose $\theta = T(F) = F(b) - F(a)$.

1) Donner l'estimateur par injection de θ. On le note $\hat{\theta}_n$.
2) Calculer l'erreur quadratique (erreur au sens L^2) de cet estimateur.
3) Montrer que $\sqrt{n}(\hat{\theta}_n - \theta)$ converge en loi vers une variable aléatoire gaussienne.
4) Déterminer un intervalle de confiance de niveau asymptotique $1-\alpha$ pour le paramètre θ.

Exercice 4.

Soit X une variable aléatoire appartenant à L^3. On note μ et σ^2 la moyenne et la variance de X Le coefficient d'asymétrie - qui mesure le manque de symétrie d'une distribution - est défini par

$$\kappa = \frac{E((X-\mu)^3)}{\sigma^3}.$$

1) Quelle est la valeur de κ pour la loi normale $\mathcal{N}(0,1)$?
2) Soit $X_1, ..., X_n$ n variables aléatoires iid suivant la même loi que $X \in L^3$. Quel est l'estimateur par injection du paramètre κ.
3) Montrer que cet estimateur converge presque surement vers θ.

Exercice 5.

Soit $X_1, ..., X_n$ n variables aléatoires iid suivant la loi de fonction de répartition F. On suppose que $F^- = F^{-1}$.

1) Quelle est la loi de $F(X_1)$.
2) Montrer que la loi de la variable aléatoire $\sup_{x\in\mathbb{R}} |F_n(x) - F(x)|$ ne depend pas de la loi de X_1.

Exercice 6.

Soit $X_1, ..., X_n$ n variables aléatoires iid suivant la loi de fonction de répartition F. On fixe $x \in \mathbb{R}$). Montrer que pour tout $b > 0$

$$P(|\hat{F}_n(x) - F(x)| > b) \leq \frac{1}{4nb^2}$$

et en déduire un intervale de confiance de niveau $1 - \alpha$ pour $F(x)$

Exercice 7.

Soit $(X_1, \ldots, X_n)$ n variables aléatoires indépendantes et de même loi. On suppose que la loi de X_1 admet une densité f. On observe $(x_1, ..., x_n)$ une réalisation de $(X_1, \ldots, X_n)$. On dispose des informations suivantes à propos de l'histogramme de ces données

Classe I_i	$]0,2]$	$]2,4]$	$]4,7]$	$]7,11]$	$]11,15]$
Hauteur h_i	0.245	0.130	0.050	0.020	h_5

On rappelle que l'histogramme est une fonction constante par morceaux qui définit une densité.

1) Calculer la valeur de h_5.
2) A partir des informations disponibles, peut-on calculer la valeur du processus empirique $\hat{F}_n$ au point $t = 7$? si oui donner la valeur sinon donner un encadrement de cette valeur.
3) Même question pour $t = 10$.
4) On dispose de l'information supplémentaire $\hat{F}_n(3) = 0.6$. Construire l'histogramme des observations $(x_1, ..., x_n)$ associé aux classes suivantes :

$$]0,2]\ ,\quad]2,3],\quad]3,4],\quad]4,7],\quad]7,11],\quad]11,15].$$