Université Hassiba Benbouali Chlef
Faculté des Sciences Exactes & Informatique
Département de Mathématiques

Année universitaire 2021/2022
Master II : Math.Appli & Stat
Module : Stat Non Parametrique

Devoir n=02 : estimation NP de la densité de probabilité

Exercice 01 :

Parmi les fonctions suivantes définies sur **R**, déterminer les quelles sont la densité d'une variable aléatoire à densité. Calculer le cas échéant leur fonction de répartition et préciser si elles a dmettent une espérance.

1. $f_1(x) = \begin{cases} \cos x & \text{si } x \in [0, \pi/2] \\ 0 & \text{sinon.} \end{cases}$

2. $f_2(x) = \dfrac{1}{1+x^2},\ x \in \mathbb{R}$

3. $f_3(x) = \dfrac{e^x}{(e^x+1)^2},\ x \in \mathbb{R}$

4. $f_4(x) = \begin{cases} 1+x & \text{si } x \in [-1, 0] \\ 1-x & \text{si } x \in [0, 1] \\ 0 & \text{sinon} \end{cases}$

5. $f_5(x) = \begin{cases} \frac{1}{|x|^3} & \text{si } |x| > 1 \\ 0 & \text{sinon} \end{cases}$

6. $f_6(x) = \sin x + 1,\ x \in \mathbb{R}.$

Exercice 02 : Soit X une variable aléatoire réelle admettant comme densité $f(x) = c\exp(- \mid x \mid);\quad x \in \mathbb{R}$: **1.** Montrer que c = 1/2. **2.** Calculer la fonction de répartition de X. **3.** Calculer $\mathbb{E}[X]$.

4. Pour tout $n \in \mathbb{N}$, on appelle I_n l'intégrale définie par

$$I_n = \int_0^{+\infty} x^n \exp(-x)dx$$

(a) Combien vaut I_0 ?

(b) Montrer que pour tout $n \in N^*$ $I_n = nI_{n-1}$. En déduire que pour tout $n \in \mathbb{N}, I_n = n!$.

5. Pout tout $n \in \mathbb{N}$, calculer $\mathbb{E}[X^{2n}]$. En déduire $V[X]$.

6. Pour tout $n \in \mathbb{N}$, que vaut $\mathbb{E}[X^{2n+1}]$?

Exercice 03 : Soit X une variable aléatoire de loi de Bernoulli de paramètre p $(0 < p < 1)$. On note $Y = 1 - X$.

(a) Quelle est la loi de Y ? (b) Quelle est la loi du vecteur $(X, Y)'$? (c)Déduire $Cov(X, Y)$.

Exercice 04 : Les variables aléatoires X et Y ont la densité conjointe

$$f(x, y) = \begin{cases} e^{-x^2 y}, & x \geq 1\ et\ y \geq 0; \\ 0, & \text{sinon.} \end{cases}$$

1. Calculer $\mathbf{P}(X^2Y > 1)$. 2) Calculer les densités marginales $f_X(x)$ et $f_Y(y)$.
   X et Y sont-elles indépendantes ?

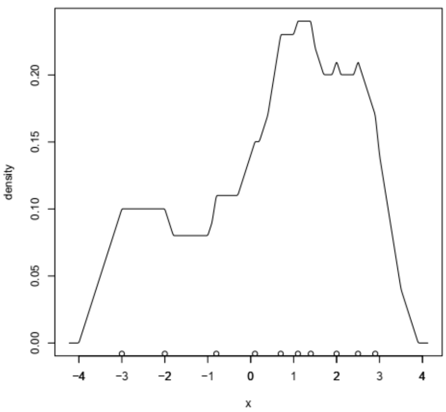
Exercice 05 : Les variables aléatoires X et Y ont la densité conjointe

$$f(x,y) = \begin{cases} 2xy + \frac{3}{2}y^2, & 0<x<1 \text{ et } 0<y<1\,; \\ 0, & \text{sinon.} \end{cases}$$

1. Vérifier que $f(x,y)$ est une densité ?.Trouver les densités marginales $f_X(x)$ et $f_Y(y)$.
2. Trouver les densités conditionnelles $f_{X|Y=y}(x)$ et $f_{Y|X=x}(y)$.
3. Calculer $\mathbf{P}((X,Y) \in [0,1/2] \times [0,1/2])$. Trouver $\mathbf{P}(X < Y)$. et $\mathbb{E}(Y|X=x)$.
4. Soit la variable aléatoire $Z = \mathbb{E}(Y|X)$. Quelle est la distribution de Z ? Trouver $\mathbb{E}(Z)$.

Exercice 05 : This figure is a kernel density estimate for n = 10. Circles represent the observations. One of the following five kernels is used to draw : - Epanechnikov - Biweight - Gaussian triangular -uniform.

(1-) Which kernel is used ? (2-) Determine h from the figure.



Exercice 06 : La fonction hist trace l'histogramme de l'échantillon x. La syntaxe est

- hist(x, proba = TRUE), le choix du nombre de classes est optimisé pour minimiser l'erreur quadratique moyenne.
- On peut fixer les classes en ajoutant l'argument breaks. hist(x, proba = TRUE , breaks = p) avec p un entier, le nombre de classes est approximativement p et les classes sont de même longueur.
- hist(x, proba = TRUE , breaks = a) avec a un vecteur, les coordonnées de a définissent les classes de l'histogramme. Il y a donc length(a) - 1 classes.

On cherche a estimer la densité par un histogramme en utilisant le nombre de classes optimal de la fonction hist. Pour n = 50, 500, 1000, 10000 :

a) Tracer l'histogramme calculé sur les n premières observations $X_1, ..., X_n$. **b)** Superposer la densité de la loi théorique c'est à dire la densité de la loi N(0,1). **c)** Commenter les résultats.

Tracer les 4 graphiques sur une même fenêtre en utilisant par(mfrow=c(2,2)).