Examen du 19 Octobre 2018

Exercise 1. On considère un n-échantillon $(X_1, \cdots, X_n)$ de variables aléatoires de loi X. On note F la fonction de répartition de X et F_n la fonction de répartition empirique associée à cet échantillon.

1. Montrer que si Z est une variable aléatoire de loi exponentielle de paramètre λ alors λZ est une variable aléatoire de loi exponentielle de paramètre 1.

2. Construire un test d'appartenance à la famille des lois exponentielles, c'est-à-dire un test de
$$H_0 : F \in \mathcal{F} \text{ contre } \; H_1 : F \notin \mathcal{F}$$
où
$$\mathcal{F} = \{\text{fonction de répartition } G : \exists\lambda > 0, \text{ tel que } G(x) = (1 - \exp(-\lambda x))I_{x\geq 0}\}.$$
(on demande de préciser la statistique du test, de montrer que sa loi sous H_0 est libre et d'écrire l'expression de la p-valeur de ce test.)

3. La loi de la statistique du test de la question 2 est tabulée. On fournit ci-dessous quelques quantiles intéressants:

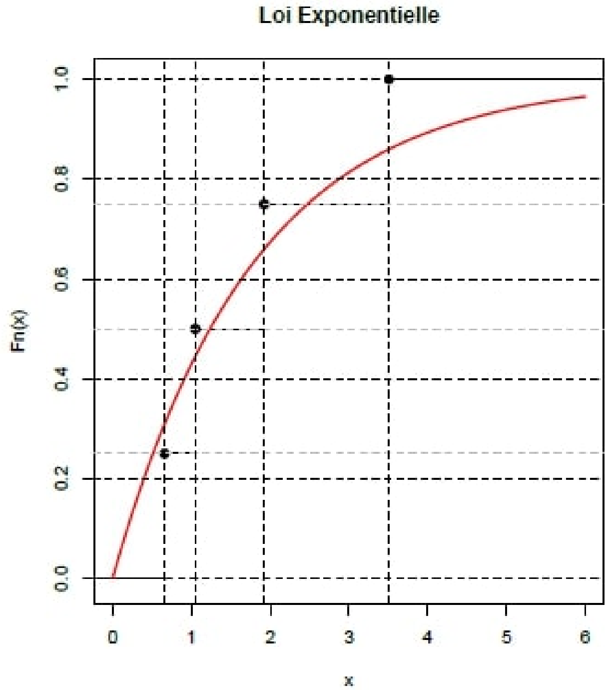
Quantiles	$q_{5\%}$	$q_{10\%}$	$q_{90\%}$	$q_{95\%}$
Stat du test d'appartenance à la loi expo	0.21	0.23	0.44	0.48

Considérons la réalisation d'un échantillon de taille $n = 4$:
$$0.66, \; 3.51, \; 1.92, \; 1.05.$$
Nous cherchons à tester si cet échantillon est distribué selon une loi exponentielle. Pour cela nous proposons d'appliquer le test précédemment construit. Sur la figure, ci-dessous, nous avons tracé la fonction de répartition empirique correspondant à l'échantillon donné. Nous avons tracé aussi la fonction de répartition de la loi exponentielle de paramàtre $\hat{\lambda}$ ($\hat{\lambda} = 0.56$).

4. Par une lecture graphique sur la figure, donner la valeur de la statistique du test.

5. En utilisant les quantiles donnés ci-dessus, effectuer le test pour un niveau 5%.

6. Conclure.



Exercise 2. Avec le logiciel R, on exécute la commande suivante

```
> binom.test(4,23,1/2, "l")
on obtient :
Exact binomial test
data: 4 and 23
number of successes = 4, number of trials = 23, p-value = 0.0013
alternative hypothesis: true probability of success is less than 0.5
```

De quel test d'hypothèse s'agit-il ? préciser les hypothèses H_0 et H_1 ? Donner l'expression mathématique qui permet d'avoir *p-value* $= 0.0013$. Que serait la décision pour une erreur de première espèce 0.05.

Exercice 3. On a mesuré le rythme des battements cardiaques de 9 sujets pris au hasard parmi les étudiants se présentant spontanément au camion de centre de transfusion sanguine pour faire un don du sang. Après le don du sang, les étudiants prennent une collation sur place. L'infirmière en profite pour leur mesurer leur rythme cardiaque. Les résultats sont les suivants exprimés en nombre de battements cardiaques par minute :

Sujet	1	2	3	4	5	6	7	8	9
Mesure avant le don du sang	78	101	77	75	92	72	64	87	65
Mesure après le don du sang	84	80	67	72	79	76	66	79	60

```
> MesureAV<-c(78,101,77,75,92,72,64,87,65)
> MesureAP<-c(84,80,67,72,79,76,66,79,60)
> MesureAP-MesureAV
[1]   6 -21 -10  -3 -13   4   2  -8  -5
```

Déterminer à l'aide du test de Wilcoxon et de la table ci-dessous l'influence du don du sang sur le battement cardiaque des étudiants. Au préalable, on formulera clairement les hypothèses nulle et alternative testées.

k	0	1	2	3	4	5	6	7	8	9
$P[W_9^+ \leq k]$	0.0019	0.0039	0.0058	0.0097	0.0136	0.0195	0.0273	0.0371	0.0488	0.0644

k	10	11	12	13	14	15	16	17	18	19
$P[W_9^+ \leq k]$	0.0820	0.1015	0.125	0.1503	0.1796	0.2128	0.2480	0.2851	0.3261	0.3671

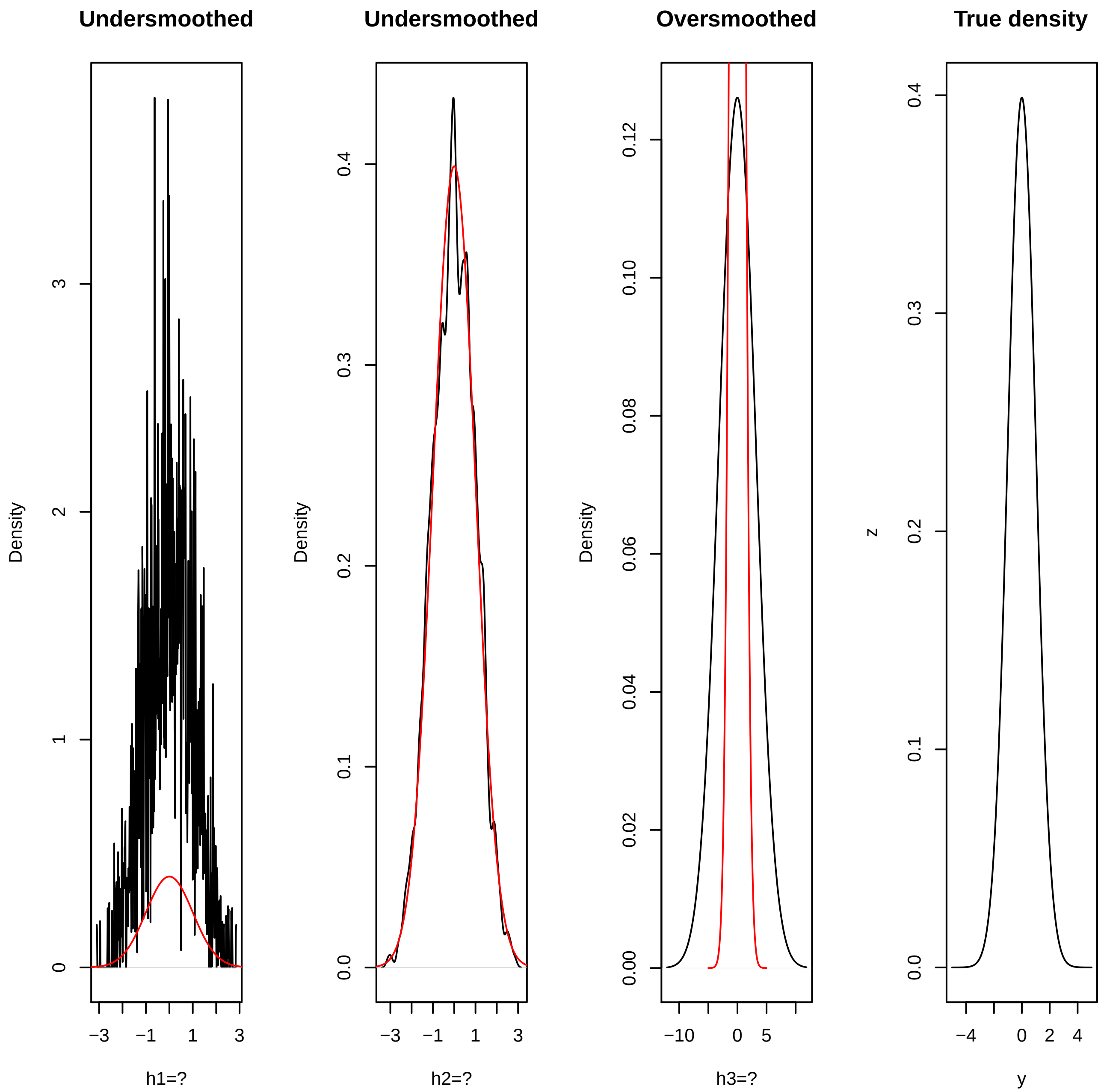
k	20	21	22	23	24	25	26	27	28	29
$P[W_9^+ \leq k]$	0.4101	0.4550	0.5	0.5449	0.5898	0.6328	0.6738	0.7148	0.7519	0.7871

k	30	31	32	33	34	35	36	37	38	39
$P[W_9^+ \leq k]$	0.8203	0.8496	0.875	0.8984	0.9179	0.9355	0.9511	0.9628	0.9726	0.9804

k	40	41	42	43	44	45
$P[W_9^+ \leq k]$	0.9863	0.9902	0.9941	0.9960	0.9980	1

Exercice 4. A partir des données $(x_1, \cdots, x_n)$ issues d'un échantillon d'une loi normale de densité f, on a essayé de représenter graphiquement les estimateurs à noyaux de f : on a pris pour K le noyau gaussien et pour h les trois valeurs suivantes $\{3, 0.001, 0.1\}$. On a obtenu les 3 graphiques suivants : sur chaque graphe il y a une superposition entre l'estimateur de la densité et la densité réelle (le 4ème correspondant à la densité réelle f).

Question: Identifier pour chaque graphe la valeur de h correspondante.


Undersmoothed
Undersmoothed
Oversmoothed
True density
Density
Density
Density
z
h1=?
h2=?
h3=?
y
0
1
2
3
−3
−1
1
3
0.0
0.1
0.2
0.3
0.4
0.00
0.02
0.04
0.06
0.08
0.10
0.12
−10
0
5
−4
0
2
4

Exercice 5. On considère un échantillon $((X_1, Y_1), \cdots, (X_n, Y_n))$ de loi (X, Y) et le modèle de régression non-paramétrique suivant:

$$Y_i = r(X_i) + \epsilon_i, \ \ 1 \leq i \leq n.$$

On suppose que la densité de X, f_X est connue et que f la densité du couple (X, Y) est inconnue. Proposer, en justifiant votre réponse, un estimateur à noyau et linéaire pour r.