Université de Biskra
Facultés SESNV
Département de Mathématique

2020/2019
Master 2 (Stat)

****

Examen – Module Statistique Non Paramétrique

****

Exercice 1 (4 points) **Soit f_n un estimateur à noyau de la densité f, et soient les deux noyaux d'Epanechinkov et Quartique (Biweight) donnés par**

$$\mathbf{K}_e(t) = \frac{3}{4}\left(1-t^2\right).\mathbf{1}_{(|t|\leq 1)} \quad \text{et} \quad \mathbf{K}_q(t) = \frac{15}{16}\left(1-t^2\right)^2.\mathbf{1}_{(|t|\leq 1)}$$

On donne les quantités suivantes:

$$\boldsymbol{\mu}_{Ke} = \int \mathbf{K}_e^2(t)\,\mathbf{dt} = \mathbf{0,6} \qquad \text{et} \qquad \boldsymbol{\mu}_{Kq} = \int \mathbf{K}_q^2(t)\,\mathbf{dt} = \mathbf{0,714}.$$

(a) Rappeler la formule de l'effecacité relative au noyau optimal d'Epanechinkov, notée Eff (K).

(b) Donner la valeur de Eff (K_q) du noyau Quartique (Biweight).

Exercice 2 (8 points) **Soit F une fonction de répartition sur $\mathbb{R}$ et soit $\theta \in \mathbb{R}_+$ un paramètre inconnu. On dispose d'un échantillon iid $(X_1, ..., X_n)$ de fonction de répartition :**

$$P(X \leq x) = F_\theta(x) = F(x-\theta).$$

On considère aussi la variable aléatoire $V_n = \sum_{i=1}^{n} 1_{(X_i > 0)}$.

(a) Pour $n \in \mathbb{N}^*$ fixé, montrer que V_n suit une loi Binomiale de paramètres n et p à préciser.

(b) Montrer que la loi limite de $(V_n - np)/\sqrt{n}$ est gaussienne. Préciser la moyenne et la variance de cette loi limite.

(c) Déterminer la loi limite des deux statistiques S_n et de W_n suivantes:

$$S_n = \left(\frac{V_n}{n}\right)^2 \qquad \text{et} \qquad W_n = \frac{n}{V_n}.$$

Exercice 3 (8 points) **On étudie l'arrivée des voitures à un poste de péage sur une autoroute, pendant la durée de temps unitaire d'une munite. Soit X la variable aléatoire à valeurs dans $\mathbb{N}$ comptant le nombre des voitures arrivées à un poste de péage sur une autoroute, par munite. On souhaite tester si le processus peut-être assimilé à une loi de Poisson de paramètre λ à l'aide d'un test d'adéquation de Khi-2. Soient les résultats obtenus à partir de 170 mesures:**

Débit par munite O_k :	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Fréquence N_k :	**15**	**33**	**12**	**37**	**30**	**10**	**9**	**5**	**12**	**7**

(a) Justifier le choix $\hat{\boldsymbol{\lambda}} = 4,45$ du paramètre $\boldsymbol{\lambda}$.

(b) Donner la statistique de test d'ajustement de Khi-deux pour le problème de test considéré.

(c) Calculer la valeur de cette statistique, sachant que la probabilité p_k que C soit égal à k lorsque C est une variable suivant la loi de Poisson de paramètre $\hat{\boldsymbol{\lambda}} = 4,45$ est donnée par :

k :	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
p_k:	**0.05**	**0.12**	**0.17**	**0.19**	**0.17**	**0.13**	**0.08**	**0.04**	**0.02**	**0.01**

(d) Quelle est la conclusion de ce test pour un niveau $\boldsymbol{\alpha} = 5\%$?

On donne pour différentes valeurs de m, le quantile d'ordre $0,95$ d'une χ^2 à m ddl $\chi^2_{(m)}$:

$$\boldsymbol{\chi}^2_{(7)} = \mathbf{14,07} \qquad \boldsymbol{\chi}^2_{(8)} = \mathbf{15,51} \qquad \boldsymbol{\chi}^2_{(9)} = \mathbf{16,92}$$

Corrigé-type de l'Epreuve du Module : Statistique Non Paramétrique

Exercice 1 (4 points) Soit f_n un estimateur à noyau de la densité f.
(a) L'effecacité relative au noyau optimal d'Epanechinkov est

$$Eff(K) = \left(\frac{\mu_{opt}\sigma_{opt}}{\mu_c\sigma_c}\right)^{4/5}, \quad \text{avec } \mu = \int K^2(t)\,dt \quad \text{et} \quad \sigma^2 = \int t^2 K(t)\,dt.$$

(b) $Eff(K_q)$ du noyau Quartique (Briweight)**:** On a

$$\mu_{opt} = \int K^2_{opt}(t)\,dt = \int_{-1}^{1}\left(\frac{3}{4}(1-t^2)\right)^2 dt = 0,6 \qquad \mu_q = \int K^2_c(t)\,dt = \int_{-1}^{1}\left(\frac{15}{16}(1-t^2)^2\right)^2 dt = 0,714$$

$$\sigma^2_{opt} = \int t^2 K(t)\,dt = \int_{-1}^{1}\frac{3}{4}(1-t^2)\,t^2 dt = 0,2 \qquad \sigma^2_q = \int t^2 K(t)\,dt = \int_{-1}^{1}\frac{15}{16}(1-t^2)^2\,t^2 dt = 0,143$$

Donc : $Eff(K_q) = \left(\frac{\mu_{opt}\sigma_{opt}}{\mu_c\sigma_c}\right)^{4/5} = \left(\frac{0.6\sqrt{0.2}}{0.714\sqrt{0.143}}\right)^{4/5} = 0.995.$

Exercice 2 (8 points) $P(X \le x) = F_\theta(x) = F(x-\theta)$ et $V_n = \sum_{i=1}^{n} 1_{(X_i>0)}$.
(a) Il est clair que $1_{(X_i>0)}$ est une v.a de Bernoulli de paramètre

$$p = P\left(1_{(X_i>0)} = 1\right) = P(X>0) = 1 - P(X<0) = 1 - F_\theta(0) = 1 - F(-\theta)$$

Donc, pour $n \in \mathbb{N}^*$ fixé, V_n est la somme de n v.a de Bernoulli, elle suit donc, une loi Binomiale de paramèters n et p.
(b) La loi limite de $(V_n - np)/\sqrt{n}$: d'après le TCL :

$$\sqrt{n}\left(\frac{V_n}{n} - p\right) = \frac{1}{\sqrt{n}}(V_n - np) \xrightarrow{L} N(0, p(1-p)) \qquad \text{quand } n \to \infty$$

(c) En applicant la méthode delta à $\frac{V_n}{n}$, nous obtenons:

$$S_n = g\left(\frac{V_n}{n}\right) \text{ tq } g(t) = t^2 \qquad \text{donc} \qquad \sqrt{n}\left(S_n - p^2\right) \xrightarrow{L} N\left(0, 4p^3(1-p)\right)$$

$$W_n = g\left(\frac{V_n}{n}\right) \text{ tq } g(t) = t^{-1} \qquad \text{donc} \qquad \sqrt{n}\left(W_n - p^{-1}\right) \xrightarrow{L} N\left(0, (1-p)/p^3\right)$$

Exercice 3 (8 points) On souhaite tester si le processus peut-être assimilé à une loi de Poisson de paramètre λ : $H_0 : X \sim P^0 = P(\lambda)$
(a) Le choix $\hat{\lambda} = 4,45$ du paramètre λ est dù à l'estimation par moments de λ : Comme $X \sim P(\lambda)$ donc $E(X) = \lambda$, donc $\hat{\lambda} = \bar{X} = 4,45$.
(b) Statistique de test d'ajustement de Khi-deux :

$$D\left(P_n, P^0\right) = \sum_{k=1}^{m=10} \frac{(N_k - np_k)^2}{np_k} \sim \chi^2_{(m-1-1)}$$

(c) La valeur de cette statistique : $D\left(P_n, P^0\right) = 0.57$.
(d) Au niveau $\alpha = 5\%$, la loi limite de la stat de Khi-2 est une $\chi^2_{(m-1-1)} = \chi^2_{(8)}$ car la loi contient un paramètre inconu λ. Puisque

$$D\left(P_n, P^0\right) = 0,57 << \chi^2_{(8)} = 15,51$$

En conclus que, H_0 est acceptée à 95%, le processus est un Poisson.