Université Dr.Moulay Tahar de Saida.
Faculté des Sciences
Département de mathématiques

12 novembre 2013
Durée: 1h30

Concours d'accès à l'Ecole doctorale
Modèles stochastiques, Statistique et Applications
1^{ere} Epreuve: Statistique Para.-NonPara.
Sujet 1

Exercice 1. *Soient $f \in L^1$ (une fonction intégrable) et K un noyau borné, intégrable et vérifiant $\int_{\mathbb{R}} K(x)dx = 1$ et $|xK(x)| \longrightarrow 0$ quand $x \longrightarrow \infty$. Montrer que f est continue en tout point de x et*

$$\lim_{h_n \longrightarrow 0} (f * K_h)(x) = f(x)$$

où $K_h(.) = \frac{1}{h} K\left(\frac{.}{h}\right)$ et $f * g(x) = \int g(x-y)f(y)dy$.

Exercice 2.

(1) *Soient deux populations gaussiennes P_1 et P_2. On suppose que $m_1 = m_2 = m$ et $\sigma_1 = \sigma_2 = \sigma$. On extrait de chaque population un échantillon de taille n_1 et n_2 respectivement. Montrer que*

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{(n_1-1)S_1'^2(n_2-1)S_2'^2}{n_1+n_2-2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_1}}}$$

suit une loi de Student à $n_1 + n_2 - 2$ degrés de liberté.

(2) *Soit P une population réparti suivant une loi normale $\mathcal{N}(m, \sigma)$. Deux échantillons de tailles respectives $n_1 = 12$ et $n_2 = 18$ ont pour moyennes $\overline{x}_1 = 22.23$ et $\overline{x}_2 = 21.98$ et écart-type $S_1' = 0.18$ et $S_2' = 0.23$ respectivement. Peut-on conclure que ces deux échantillons proviennent d'une même population pour un seuil de signification de 5%.*

Comme K est d'intégrale qui vaut 1, on a:

Exercice 3. *Soit $X_1, X_2, \dots X_n$ un n-échantillon de X de fonction de répartition F et de densité f et soit $x \in \mathbb{R}$ tel que $F(x) < 1$. Admettant que que la fonction de hasard $\lambda(x) = \frac{f(x)}{1-F(x)}$ est un paramètre fonctionnel.*

(1) *Déduire un estimateur $\lambda_n(x)$ pour $\lambda(x)$ par la méthode du noyau.*
(2) *Montrer que*

$$\lambda_n(x) - \lambda(x) = \frac{f_n(x) - f(x)}{1 - f_n(x)} + (F_n(x) - F(x)) \frac{\lambda(x)}{1 - F_n(x)}$$

(3) *En utilisant la convergence presque complète de $F_n(x)$ vers $F(x)$ montrer qu'il existe $\delta > 0$ tel que:*

$$\sum_n \mathbb{P}\left[(1 - F_n(x)) < \delta\right] < \infty$$

(4) *Etudier la convergence presque complète de l'estimateur $\lambda_n(x)$.*

$$\begin{aligned}
(f * K_{h_n})(x) - f(x) &= \int f(x-y)K_{h_n}(y)dy - f(x)\int K_{h_n}(y)dy \\
&= \int (f(x-y) - f(x))\, K_{h_n}(y)dy \\
&= \int\limits_{|y|\leq\delta} (f(x-y) - f(x))\, K_{h_n}(y)dy \\
&- \int\limits_{|y|>\delta} (f(x-y) - f(x))\, K_{h_n}(y)dy
\end{aligned}$$

Ainsi, pour tout $\delta > 0$:

$$\begin{aligned}
\left|(f * K_{h_n})(x) - f(x)\right| &\leq \sup_{|y|\leq\delta} |f(x-y) - f(x)| \int\limits_{|y|\leq\delta} K_{h_n}(y)dy + \\
&\int\limits_{|y|>\delta} (f(x-y) - f(x))\, K_{h_n}(y)dy \\
&\leq \sup_{|y|\leq\delta} |f(x-y) - f(x)| \int_{|t|\leq\delta|h_n|^{-1}} K(t)dt + \\
&\int_{|y|>\delta} \frac{|f(x-y)|}{|y|} \left|\frac{y}{h_n}\right| K\left(\frac{y}{h_n}\right) dy + |f(x)| \int_{|t|>\delta|h_n|^{-1}} K(t)dt
\end{aligned}$$

On obtient donc pour h_n fixé et $\delta \longrightarrow 0$ le premier terme du membre droite de cete inégalité tend vers 0 Ã cause de la continuité de f, le second terme est majoré par $\delta^{-1} \int_{\mathbb{R}} |f| \sup_{|t|>\delta|h_n|^{-1}} |t||K(t)|dt$ qui tend vers 0 quand $h_n \longrightarrow 0$ et le troisième terme $|f(x)| \int_{|t|>\delta|h_n|^{-1}} K(t)dt \longrightarrow 0$ quand $h_n \longrightarrow 0$.

Ainsi on vérifie que si f est continue en x

$$\left|(f * K_{h_n})(x) - f(x)\right| \underset{h_n\longrightarrow 0}{\longrightarrow} 0 \implies \lim_{h_n\longrightarrow 0} (f * K_{h_n})(x) = f(x)$$