Université Paris I, Panthéon - Sorbonne

Première Année Master M.A.E.F. 2007 – 2008

Statistiques I

Contrôle continu n°1, décembre 2007

Examen de 1h30. Tout document ou calculatrice est interdit.

1. On considère une suite $(\varepsilon_i)_{i\in\mathbb{N}}$ de variables aléatoires indépendantes et identiquement distribuées suivant une loi $\mathcal{N}(0,1)$. On pose $X_0=0$ et on définit:
$$X_n=\alpha\,X_{n-1}+\varepsilon_n, \quad \text{pour tout } n\in\mathbb{N}^*$$
où $\alpha\in]-1,1[$ est inconnu.
   (a) Montrer que pour tout $k\in\mathbb{N}^*$, $X_k=\displaystyle\sum_{i=1}^k\alpha^{k-i}\,\varepsilon_i$.
   (b) Pour tout $k\in\mathbb{N}^*$, déterminer $\mathbb{E}(X_k)$, $\mathrm{var}(X_k)$ et la loi de X_k. En déduire que si $\alpha\neq 0$, $(X_n)_{n\in\mathbf{N}^*}$ n'est pas une suite de variables identiquement distribuées.
   (c) Pour $1\leq i\leq j$, montrer que $\mathrm{cov}(X_i,X_j)=\alpha^{j-i}\dfrac{1-\alpha^{2i}}{1-\alpha^2}$. Montrer que si $\alpha\neq 0$, $(X_n)_{n\in\mathbf{N}^*}$ n'est pas une suite de variables indépendantes.
   (d) On suppose que $(X_1,\dots,X_n)$ avec $n\in\mathbb{N}^*$ est connu Montrer que
$$\overline{X_n}=\frac{1}{n}\sum_{i=1}^n X_i=\frac{1}{n}\sum_{k=1}^n\Big(\frac{1-\alpha^{n-k+1}}{1-\alpha}\Big)\varepsilon_k.$$
En déduire la loi de $\overline{X_n}$, puis que $\sqrt{n}\,(1-\alpha)\,\overline{X_n}\ \xrightarrow[n\to\infty]{\mathcal{L}}\ \mathcal{N}(0,1)$. $\overline{X_n}$ est-il un estimateur convergent de α?
   (e) Soit $\widehat{\sigma}_n^2=\dfrac{1}{n}\displaystyle\sum_{i=1}^n X_i^2$. Montrer que $\mathrm{var}(\widehat{\sigma}_n^2)=\dfrac{2}{n^2}\displaystyle\sum_{1\leq i,j\leq n}\mathrm{cov}^2(X_i,X_j)$ et $\displaystyle\lim_{n\to\infty} n\,\mathrm{var}(\widehat{\sigma}_n^2)=\dfrac{2(1+\alpha^2)}{(1-\alpha^2)^3}$.
En déduire que $\widehat{\sigma}_n^2\ \xrightarrow[n\to+\infty]{\mathcal{P}}\ (1-\alpha^2)^{-1}$ ainsi que l'expression d'un estimateur convergent de $|\alpha|$.

2. Soit X une variable aléatoire dont la loi est absolument continue par rapport à la mesure de Lebesgue sur $[0,1]$ et telle que sa densité par rapport à cette mesure soit de la forme:
$$f_{a,b}(x)=(a\,x+b)\,\mathbb{I}_{0\leq x\leq 1},$$
avec $(a,b)\in\mathbb{R}$. Soit $(X_1,\dots,X_n)$ un échantillon de n v.a.i.i.d. de même loi que X.
   (a) Donner des conditions sur (a,b) pour que $f_{a,b}$ soit bien une densité. Montrer que l'on peut écrire $f_{a,b}$ sous la forme $g_\theta(x)=\dfrac{2}{1-2\theta}\,(x-\theta)\,\mathbb{I}_{0\leq x\leq 1}$ en précisant l'ensemble Θ des valeurs possibles pour θ.
   (b) Préciser le modèle statistique associé à $(X_1,\dots,X_n)$. On peut montrer (ne pas le faire! sauf s'il vous reste du temps) que ce modèle n'appartient pas à la famille exponentielle.
   (c) Montrer que $\widehat{S}_n=(X_1,\dots,X_n)$ est une statistique exhaustive pour ce modèle (redémontrer cette propriété du cours...). En considérant la fonction $h(x_1,\dots,x_n)=x_1-x_1^2-1/6$, montrer que cette statistique n'est pas complète.
   (d) On pose $Y=X-\theta$. Déterminer la loi de Y.
   (e) Soit $(Y_1,\dots,Y_n)$ un échantillon de n v.a.i.i.d. de même loi que Y. Le modèle statistique associé à $(Y_1,\dots,Y_n)$ appartient-il à la famille exponentielle?
   (f) Déterminer une statistique exhaustive $\widehat{T}_n=(\widehat{T}_1,\widehat{T}_2,\widehat{T}_3)$ pour ce modèle. Montrer qu'elle est également minimale, dès que $n\geq 3$.