Université Paris I, Panthéon - Sorbonne

Première Année Master M.A.E.F. 2006 – 2007

Statistiques I

Contrôle continu n°2, janvier 2007

Examen de 2 h 00. Tout document ou calculatrice est interdit.

1. Soit X une variable aléatoire suivant la loi suivante:

$$P(X=1) = P(Y=-1) = p \quad \text{et} \quad P(X=0) = 1-2p,$$

où p est un paramètre réel inconnu.

(a) Déterminer l'ensemble des valeurs possibles pour p. Calculer EX et var$X$.

(b) On suppose que la suite $(X_i)_{i\in\mathbb{N}}$ est constituée de variables aléatoires indépendantes et identiquement distribuées suivant la même loi que X. Soit un échantillon $(X_1,\ldots,X_n)$. Déterminer le modèle statistique associé à cet échantillon et déterminer une mesure dominant ce modèle. Montrer que le modèle appartient à la famille exponentielle. En déduire une statistique exhaustive complète pour ce modèle. Montrer que p peut être estimé efficacement et donner un tel estimateur. Calculer la borne de Cramer-Rao et vérifier qu'elle est bien atteinte par cet estimateur.

(c) On définit la suite $(Y_i)_{i\in\mathbb{N}^*}$ à partir de $(X_i)_{i\in\mathbb{N}}$ de la manière suivante:

$$Y_{i+1} = X_i \cdot X_{i+1} \quad \text{pour} \quad i \in \mathbb{N}.$$

Déterminer la loi de Y_i. Montrer que $\mathrm{cov}(Y_i, Y_{i+1}) = 0$. Les $(Y_i)_i$ sont-elles indépendantes ?

(d) Montrer que $(|Y_1|,\ldots,|Y_n|)$ est une statistique exhaustive pour le modèle statistique induit par $(Y_1,\cdots,Y_n)$.

2. Soit la variable X qui suit une loi dont la densité f_X par rapport à la mesure de Lebesgue sur $\mathbb{R}$ est, avec $\theta > 0$ et $\alpha > 0$:

$$f_X(x) = K \cdot x^{\alpha} \mathbb{I}_{0\le x\le\theta} \quad \text{pour tout } x \in \mathbb{R},$$

(a) Déterminer K en fonction de α et θ.

(b) Montrer que $Y = \log(\theta/X)$ suit une loi exponentielle dont on précisera le paramètre.

(c) On suppose que la suite $(X_i)_{i\in\mathbb{N}}$ est constituée de variables aléatoires indépendantes et identiquement distribuées suivant la même loi que X. Soit un échantillon $(X_1,\ldots,X_n)$. On suppose que (θ,α) est inconnu. Préciser alors le modèle statistique formé par cet échantillon et la mesure dominante. Ce modèle appartient-il à la famille exponentielle ?

(d) Dans cette question, et uniquement dans cette question, on suppose que θ est connu. Préciser alors le modèle statistique. Ce modèle appartient-il à la famille exponentielle ? Montrer que l'estimateur du maximum de vraisemblance $\tilde{\alpha}_n$ de α existe, est unique et s'écrit:

$$\tilde{\alpha}_n = \frac{1}{\frac{1}{n}\sum_{i=1}^n \log(\theta/X_i)} - 1$$

Montrer que $\tilde{\alpha}_n$ converge presque sûrement vers α et qu'il vérifie un théorème de la limite centrale que l'on précisera. En déduire un intervalle de confiance à 95% sur α pour n grand.

(e) Dans cette question, θ et α sont inconnus. Déterminer une statistique exhaustive pour le modèle. En vous aidant de la question précédente, déterminer l'estimateur du maximum de vraisemblance $(\widehat{\theta}_n, \widehat{\alpha}_n)$ de (θ, α). Déterminer la fonction de répartition de $\log(\theta/\widehat{\theta}_n)$ et en déduire que $\widehat{\theta}_n \xrightarrow[n\to+\infty]{\mathcal{P}} \theta$, puis que $\sqrt{n}\log(\theta/\widehat{\theta}_n) \xrightarrow[n\to+\infty]{\mathcal{P}} 0$.

(f) Soit $(U_n)_{n\in\mathbb{N}}$ et $(V_n)_{n\in\mathbb{N}}$ deux suites de variables aléatoires définies sur le même espace de probabilité. Montrer que si $(U_n)_n$ converge vers une loi P_0 et $(V_n)_n$ converge en probabilité vers 0, alors $(U_n + V_n)_n$ converge en loi vers P_0 (on pourra par exemple majorer la différence de fonctions caractéristiques). En déduire que $\widehat{\alpha}_n$ suit le même théorème de la limite centrale que $\tilde{\alpha}_n$.