Université Paris I, Panthéon - Sorbonne

Première Année Master M.A.E.F. 2017 – 2018

Statistiques I

Contrôle continu n°2, décembre 2017

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

1. **(Sur 12 points)** Soit la variable X qui suit une loi dont la densité f_X par rapport à la mesure de Lebesgue sur $]0,1]$ est, avec β et $K \in \mathbb{R}$:
$$f_X(x) = K\, x^\beta \quad \text{pour tout } x \in]0,1],$$
   (a) Déterminer K en fonction de β **(0.5pts)**, en précisant quelle condition doit vérifier β **(0.5pts)**. En déduire $\mathbb{E}(X)$ **(0.5pts)** et $\mathrm{var}(X)$ **(1pt)**, en précisant également les conditions portant sur β.

   (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 observé $(X_1,\dots,X_n)$. On désire estimer β à partir de cet échantillon. Quel est le modèle statistique **(0.5pts)**? Montrer que ce modèle appartient à la famille exponentielle **(0.5pts)**.

   (c) En déduire qu'il n'existe pas d'estimateur sans biais efficace de β **(1.5pts)**.

   (d) Montrer que $-\log(X)$ suit une loi connue dont on précisera le paramètre **(1.5pts)**. En déduire que $\widetilde{\beta}_n = -1-(n-1)\Big(\sum_{i=1}^n \log(X_i)\Big)^{-1}$ est un estimateur sans biais de β (utiliser la loi gamma...) pour $n \geq 2$ **(1.5pts)**, puis qu'il est de variance uniformément minimale parmi les estimateurs sans biais **(1pt)**.

   (e) Montrer que l'estimateur $\widehat{\beta}_n$ de β par maximum de vraisemblance diffère de $\widetilde{\beta}_n$ **(1pt)**. Montrer que $\widehat{\beta}_n$ est biaisé mais asymptotiquement non biaisé **(1pt)**. Montrer sans calcul que $\mathrm{var}(\widetilde{\beta}_n) < \mathrm{var}(\widehat{\beta}_n)$ **(1pts)**.

2. **(Sur 20 points)** Soit une suite $(X_i)_{i\in\mathbf{N}}$ de variables aléatoires indépendantes et identiquement distribuées suivant une loi uniforme sur $\{-m, 1-m, \dots, m-1, m\}$ où $m \in \mathbf{N}$ est inconnu. On observe $(X_1,\dots,X_n)$ où $n \in \mathbf{N}^*$.

   (a) Rappeler la loi de X_1 ainsi que son espérance **(1pt)**.

   (b) Déterminer le modèle statistique induit par $(X_1,\dots,X_n)$ **(0.5pts)**. Est-ce un modèle dominé **(0.5pts)**? Exponentiel **(0.5pts)**?

   (c) Montrer que $\widehat{T} = (\min_{1\leq i\leq n} X_i\,,\, \max_{1\leq i\leq n} X_i)$ est une statistique exhaustive pour le modèle **(1pt)**.

   (d) Montrer que $\widehat{T}' = \max_{1\leq i\leq n}|X_i|$ est aussi une statistique exhaustive **(1pt)**. Montrer qu'elle est minimale **(1.5pts)** alors que T ne l'est pas **(1pt)**.

   (e) Déterminer la loi de $|X_1|$ **(0.5pts)**, et en déduire que $\mathbb{P}(\widehat{T}' = j) = \frac{(2j+1)^n-(2j-1)^n}{(2m+1)^n}$ pour $j \in \{1,2,\dots,m\}$ **(2.5pts)**.

   (f) Soit $h :\ \mathbf{N} \to \mathbf{R}$, quelconque. Déterminer explicitement $\mathbb{E}[h(\widehat{T}')]$ **(0.5pts)**. En déduire que si $\mathbb{E}[h(\widehat{T}')] = 0$ pour tout $m \in \mathbf{N}$, alors $h \equiv 0$ (on pourra commencer par le cas $m = 0$,...) **(2pts)**. Conséquence **(0.5pts)**?

   (g) Déterminer $\widehat{m}_n$ l'estimateur du maximum de vraisemblance de m **(1pt)**, puis montrer que $\mathbb{E}(\widehat{m}_n) = m - b_n(m)$ avec $b_n(m) = \frac{1}{(2m+1)^n}\sum_{i=1}^{m-1}(2i+1)^n$ **(2pts)**.

   (h) A l'aide d'une comparaison avec une intégrale, montrer que $b_n(m) \leq (2m+1)/(n+1)$ **(2pts)**. En déduire que $\mathbb{E}[|\widehat{m}_n - m|] \underset{n\to+\infty}{\longrightarrow} 0$ **(1pt)**, puis que $\widehat{m}_n$ est un estimateur convergent de m **(0.5pts)**. A quelle vitesse converge-t-il **(0.5pts)**?