Université de Marseille
Licence de Mathématiques, 3eme année, probabilités-Statistique
Examen du 30 juin 2016

Le polycopié du cours, les notes de cours et les notes de TD sont autorisés.
L'examen contient 3 exercices. Le barème est sur 28 points.

Exercice 1 (Indépendance et inégalité. Barème : 10 points) Soit (Ω, T, P) un espace probabilisé et X, Y deux v.a.r..

1. On suppose dans cette question que X et Y sont indépendantes et que $Y > X$ p.s..
   (a) Montrer que pour tout $t \in \mathbb{R}$, on a $P(\{Y < t\})(1 - P(\{X < t\})) = 0$.
   (b) Montrer qu'il existe $t \in \mathbb{R}$ tel que $P(\{Y < t\}) = 0$.
   (c) On pose $a = \sup\{t \in \mathbb{R};\ P(\{Y < t\}) = 0\}$. Montrer que $a \in \mathbb{R}$, $P(\{Y < a\}) = 0$ et $P(\{X > a\}) = 0$.
2. On suppose que X modélise la durée de vie d'un moustique et Y la durée de vie d'un éléphant.
   (a) Le web suggère de prendre comme loi pour X une loi exponentielle de paramètre α et comme loi pour Y une loi exponentielle de paramètre β, avec $0 < \beta < \alpha$. (On rappelle que la loi exponentielle de paramètre λ est la loi de densité f_λ avec $f_\lambda(x) = \lambda e^{-\lambda x}$ pour $x \geq 0$ et $f_\lambda(x) = 0$ pour $x < 0$.)
   Peut-on considèrer que X et Y sont indépendantes et que $Y > X$ p.s. (c'est-à-dire que l'éléphant vit presque sûrement plus longtemps que le moustique. . .) ?
   (b) On choisit maintenant pour X une loi de densité dont la densité a pour support l'intervalle $]0, a]$ et pour Y une loi de densité dont la densité a pour support l'intervalle $]b, c]$, avec $0 < a < b < c$ (on ne s'intéresse pas à la mortalité infantile des éléphants). A t-on $Y > X$ p.s. ? Peut-on supposer que X et Y sont indépendantes ?

Exercice 2 (Exponentielle d'une v.a. gaussienne, barème 3 points) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X une v.a.r. t.q. $X \sim \mathcal{N}(0, 1)$. Soit $Y = \exp(X)$. Calculer l'espérance et la variance de Y.

Exercice 3 (Statistique. Barème : 15 points) Soit X une variable aléatoire dont la loi a pour densité

$$f_{\mu,\lambda}(x) := \frac{\lambda}{2} e^{-\lambda|x-\mu|}$$

où $\lambda > 0$ et $\mu \in \mathbb{R}$ sont deux paramètres. Soit $X_1, \ldots, X_n$ un échantillon de taille n de X.

1. Vérifier que $f_{\mu,\lambda}$ est une fonction de densité, et la représenter graphiquement (on pourra éventuellement prendre $\mu = 3$ et $\lambda = 1$, par exemple, pour effectuer le dessin).
2. Calculer $E(X)$ et expliquer le résultat à l'aide de la question précédente. En déduire un estimateur $\hat{\mu}$ de μ par la méthode des moments.
3. On pose $Y = |X - \mu|$.
   (a) Quelles sont la valeurs possibles de Y ? Calculer la fonction de répartition de Y.
   (b) En déduire la loi de Y, et donner sans calcul $E(Y)$ et $\mathrm{V}(Y)$, puis $E(Y^2)$.
   (c) En déduire que $\mathrm{V}(X) = \frac{2}{\lambda^2}$. Donner alors un estimateur $\hat{\lambda}$ de λ par la méthode des moments (i.e. en supposant que l'estimateur $\hat{\lambda}$ est la valeur de λ qui permet d'avoir l'égalité $V(X) = S_n^2(X_1, \ldots, X_n)$ où $S_n^2(X_1, \ldots, X_1)$ est la variance empirique de l'échantillon).
4. On cherche à estimer λ et μ par la méthode du maximum de vraisemblance.
   (a) Donner l'expression de la vraisemblance $L(x_1, \ldots, x_n, \mu, \lambda)$ et de son logarithme $\ln(L(x_1, \ldots, x_n, \mu, \lambda))$, où $x_1, \ldots, x_n$ sont des réalisations de $X_1, \ldots, X_n$
   (b) On suppose que la valeur de μ est connue, et on veut estimer λ. Calculer l'estimateur $\hat{\lambda}'$ de λ par la méthode du maximum de vraisemblance.
   (c) On suppose que la valeur de λ est connue. Montrer qu'une valeur $\hat{\mu}'$ de μ maximisant la vraisemblance est une valeur minimisant la fonction
   $$\varphi : y \mapsto |x_1 - y| + \cdots + |x_n - y|.$$

(d) Déterminer les points où φ est continue et dérivable. Quitte à effectuer une permutation des x_i on pourra supposer $x_1 \leq x_2 \leq \ldots \leq x_n$. Calculer alors la dérivée de φ aux points où elle est définie, et étudier les variations de φ sur $\mathbb{R}$. En déduire qu'une valeur $\hat{\mu}'$ de μ maximisant la vraisemblance est une médiane des x_i, i.e. une valeur $\hat{\mu}'$ telle que

$$Card\{i \mid x_i \leq \hat{\mu}'\} = Card\{i \mid x_i \geq \hat{\mu}'\}.$$

5. On dispose des observations $x_1, \ldots, x_{15}$ suivantes :

$$7,7, \ ; \ 6,7 \ ; \ 7,0 \ ; \ 8,1 \ ; \ 7,0 \ ; \ 6,9 \ ; \ 6,9 \ ; \ 7,0 \ ; \ 5,7 \ ; \ 9,0 \ ; \ 7,2 \ ; \ 6,7 \ ; \ 7,0 \ ; \ 6,7 \ ; \ 6,9.$$

(a) Calculer la moyenne, la médiane et la variance de ces données.

(b) On suppose que ces données proviennent d'une expérience qui a une densité $f_{\mu,\lambda}$ avec μ et λ inconnus. Calculer les estimations de μ et λ issues des estimateurs donnés aux questions 2, 3 et 4.