CONTRÔLE DE CONNAISSANCES
ALÉATOIRE, MAP 311
3 juillet 2017

Durée : 2 heures.
Sujet proposé par B. Jourdain.
Polycopié, copies des transparents et notes de PC autorisées.

Ce sujet est constitué d'un exercice et un problème indépendants. Il est possible d'admettre le résultat d'une question pour l'utiliser dans la suite.

Exercice : file de voitures (9 points/30)

Sur une route à une seule voie et un seul sens de circulation, on considère une file infinie de véhicules numérotés $0, 1, 2, 3, \dots$. On suppose qu'ils circulent tous à la même vitesse prise égale à 1. Ils sont séparés par des distances $(L_j)_{j\geq 1}$ comptées de l'arrière du véhicule $j-1$ à l'avant du véhicule j qui suit.
On suppose que pour que le véhicule j s'arrête, il lui faut, à partir du moment où le véhicule $j-1$ qui est devant lui commence à freiner, une distance $R_j + D_j$ où R_j est la distance parcourue pendant le délai de réaction du conducteur du véhicule j et D_j est la distance d'immobilisation de ce véhicule.
On suppose que les triplets $(L_j, R_j, D_j)_{j\geq 1}$ sont i.i.d. avec L_1, D_1 et R_1 indépendantes et positives.

On va étudier les risques de collision lorsqu'un objet encombrant tombe du véhicule 0 en tête de file et empêche la circulation. Pour cela on supposera pour simplifier que :

— lorsqu'un automobiliste freine, il tente d'arrêter son véhicule sur la distance la plus courte possible, peu importe s'il laisse de la place devant lui,

— les vitesses de décélération sont telles que les collisions n'ont lieu qu'entre un véhicule en mouvement et un véhicule arrêté.

On pose $D_0 = 0$ et pour $n \geq 1$, on note $B_n = \{R_n + D_n < L_n + D_{n-1}\}$.

1. Montrer que $\mathbb{P}(B_1) \leq \mathbb{P}(B_2)$. Comparer $\mathbb{P}(B_2)$ et $\mathbb{P}(B_3)$.
2. Dans cette question, on suppose que R_1, L_1 et D_1 possèdent des densités respectivement notées p_R, p_L et p_D.
   (a) Exprimer en fonction de p_R, p_D et $\overline{F}_L(\ell) = \mathbb{P}(\ell < L_1)$ la probabilité pour que le véhicule 1 ne percute pas l'obstacle.
   (b) Calculer cette probabilité lorsque p_R, p_L et p_D sont les densités exponentielles de paramètres respectifs θ_R, θ_L et θ_D.

3. Expliquer pourquoi l'événement "les véhicules de 1 à n s'arrêtent sans collision" s'écrit $A_n = \bigcap_{k=1}^n B_k$.
4. On suppose que $\mathbb{P}(R_1 \geq L_1) > 0$ et on note p cette probabilité.
   (a) Montrer que $\mathbb{P}(D_2 \geq D_1) = \frac{1+\mathbb{P}(D_2=D_1)}{2}$ et que $\mathbb{P}(B_2^c) \geq \frac{p}{2}$.
   (b) Que peut-on dire des événements $(B_{2k})_{k\geq 1}$? En déduire que $\lim_{n\to+\infty} \mathbb{P}(A_{2n}) = 0$.
   (c) Conclure qu'il y a presque sûrement au moins un accident.
   (d) On suppose que D_1 est déterministe égale à d_F. Déterminer la loi de l'indice N du premier véhicule qui n'arrive pas à s'arrêter à temps.
   Lorsque $\mathbb{P}(R_1 + d_F < L_1) > 0$, quelle est la loi conditionnelle de $N-1$ sachant $N \geq 2$?

♦ ♦
♦

Problème : loi de Pareto

Pour $a \in \mathbb{R}$ et $b > 0$, on appelle loi de Pareto de paramètre (a, b) et on note $\mathcal{P}(a, b)$ la loi de densité $p_1(x, a, b) = \frac{b}{(x-a)^{b+1}} \mathbf{1}_{[a+1,+\infty[}(x)$. On note $X \sim \mathcal{P}(a, b)$ si la variable aléatoire X est distribuée suivant la loi de Pareto de paramètre (a, b). Cette loi est par exemple utilisée pour modéliser la distribution des revenus dans une population.

I) Analyse probabiliste (7 points/30)

Soit $X \sim \mathcal{P}(a, b)$.

1. Montrer que $(X-a)^{-b}$ suit la loi uniforme sur $[0, 1]$ et en déduire sans calcul que $\ln(X-a)$ suit la loi exponentielle de paramètre b.
2. À l'aide du changement de variables $u = \frac{1}{x-a}$ vérifier que

$$\forall c \in]-1, b[,\ \mathbb{E}[(X-(a+1))^c] = b\int_0^1 (1-u)^c u^{b-c-1} du.$$

   Reconnaître l'intégrale d'une densité usuelle à la constante de normalisation près et en déduire que

$$\forall c \in]-1, b[,\ \mathbb{E}[(X-(1+a))^c] = \frac{\Gamma(b-c)\Gamma(c+1)}{\Gamma(b)}. \tag{1}$$

   où Γ est la fonction gamma d'Euler définie par $\forall a > 0$, $\Gamma(a) = \int_0^\infty x^{a-1}e^{-x}dx$ qui vérifie $\Gamma(a+1) = a\Gamma(a)$ et $\forall n \in \mathbb{N}^*$, $\Gamma(n) = (n-1)!$.

Les deux questions qui suivent sont indépendantes du reste du problème.
Soient $b, c > 0$, $Y \sim \mathcal{P}(0, b)$, $Z \sim \mathcal{P}(0, c)$ et ε une variable aléatoire de Bernoulli de paramètre $\frac{c}{b+c}$ indépendantes.

3. Déterminer la densité de $S = \varepsilon Y + \frac{1-\varepsilon}{Z}$.
4. Déterminer la loi de (R, Z) où $R = Y/Z$. Les variables aléatoires R et Z sont-elles indépendantes ? Vérifier que R et S ont même loi.

II) Estimation statistique (14 points/30)

On observe une réalisation de $X = (X_1, \ldots, X_n)$ où les X_i sont des variables aléatoires indépendantes et identiquement distribuées de loi commune dans $\{\mathcal{P}(a,b) : a \in \mathbb{R}, b > 0\}$.

5. Soient $a \in \mathbb{R}, b > 0$. Déterminer la vraisemblance $p_n(x, a, b)$ pour $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Exprimer à l'aide de $\min_{1 \leq i \leq n} x_i$ la valeur $a_n(x) \in \mathbb{R}$ qui maximise $a \mapsto p_n(x, a, b)$ pour tout $(x, b) \in \mathbb{R}^n \times]0, +\infty[$.
6. Calculer la log-vraisemblance $l_n(x, a, b)$. Quelle valeur $b_n(x, a) \in]0, +\infty[$ maximise-t-elle $b \mapsto l_n(x, a, b)$ pour $a \in \mathbb{R}$ et $x \in]a+1, +\infty[^n$ fixés ? En déduire l'estimateur du maximum de vraisemblance (EMV) $(\hat{a}_n, \hat{b}_n)$ du couple (a, b).
7. (a) Calculer $\mathbb{P}_{(a,b)}(X_1 > x)$ pour $x \geq a+1$. En déduire que sous $\mathbb{P}_{(a,b)}$, $M_n = \min_{1 \leq i \leq n} X_i \sim \mathcal{P}(a, nb)$. Calculer $\mathbb{E}_{(a,b)}(\hat{a}_n - a)$ pour $n > 1/b$. L'EMV est-il sans biais ?

   (b) À l'aide de l'équation (1), vérifier que $\lim_{n\to\infty} n^3 \mathbb{E}_{(a,b)}[|\hat{a}_n - a|^3] = \frac{6}{b^3}$ et en déduire que $\mathbb{P}_{(a,b)}\left(\lim_{n\to\infty} \sqrt{n}(\hat{a}_n - a) = 0\right) = 1$.
8. On pose $R_n = \frac{1}{n}\sum_{i=1}^n \ln(X_i - a) - \frac{1}{n}\sum_{i=1}^n \ln(X_i - \hat{a}_n)$.

   (a) À l'aide de la question 1, donner les comportements asymptotiques lorsque $n \to \infty$ sous $\mathbb{P}_{(a,b)}$ de $\frac{1}{n}\sum_{i=1}^n \ln(X_i - a)$ et $\sqrt{n}\left(\frac{1}{b} - \frac{1}{n}\sum_{i=1}^n \ln(X_i - a)\right)$.

   (b) Avec l'inégalité $\forall x > 0$, $\ln(x) \leq x - 1$, vérifier que $R_n \leq \frac{1}{n}\sum_{i=1}^n \frac{\hat{a}_n - a}{X_i - \hat{a}_n}$. En déduire que $0 \leq R_n \leq \hat{a}_n - a$.

   (c) Montrer que $(\hat{a}_n, \hat{b}_n)$ converge $\mathbb{P}_{(a,b)}$ p.s. vers (a, b) lorsque $n \to \infty$.

   (d) En écrivant que $\sqrt{n}(\hat{b}_n - b) = b\hat{b}_n \times \left(\sqrt{n}\left(\frac{1}{b} - \frac{1}{n}\sum_{i=1}^n \ln(X_i - a)\right) + \sqrt{n}R_n\right)$, vérifier que $\sqrt{n}(\hat{b}_n - b)$ converge en loi vers $\mathcal{N}(0, b^2)$ lorsque $n \to \infty$.

♦ ♦
♦