CPES 2 – Probabilités approfondies 2015-2016
Devoir Maison à rendre le lundi 14 mars

Igor Kortchemski – igor.kortchemski@cmap.polytechnique.fr

Il est demandé de soigneusement numéroter les questions. Lors de la correction, il sera fait grand cas de la clarté, de la concision et de la précision de la rédaction.

Exercice 1. On considère deux variables aléatoires X et Y définies sur le même espace de probabilité à valeurs dans $\{0,1\}$, dont la loi jointe est donnée par le tableau suivant :

(x,y)	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$\mathbb{P}(X=x,Y=y)$	$\frac{1}{2}-a$	$a+\frac{1}{3}$	b	$\frac{1}{6}-2a$

(1) Quelles valeurs sont autorisées pour a et b ?

(2) Calculer en fonction de a et b les lois marginales de X et de Y.

(3) Quelles valeurs de a et b correspondent à un couple (X,Y) de variables indépendantes ?

Exercice 2. Soit X une variable aléatoire réelle dont une densité est $f(t)=ce^{-|t|}$ pour $t\in\mathbb{R}$, où $c>0$ est une constante.

(1) Déterminer c.

(2) Soit $Y=X^2$. Déterminer la loi de Y.

(3) Soit Z la variable aléatoire définie par $Z=1$ si $X>2$ et $Z=0$ si $X\leq 2$. Déterminer la loi de Z.

Exercice 3. Soit X une variable aléatoire qui suit une loi de Poisson de paramètre $\lambda>0$.

(i) Calculer $\mathbb{E}\left[\frac{1}{X+1}\right]$ et $\mathbb{E}\left[\frac{1}{(X+1)(X+2)}\right]$. En déduire la valeur de $\mathbb{E}\left[\frac{1}{X+2}\right]$.

(ii) Soit Y une variable aléatoire telle que pour tout $k\geq 0$, la loi conditionnelle de Y sachant que $X=k$ est une loi de Poisson de paramètre k (informellement, Y est une loi de Poisson de paramètre X !). Calculer $\mathbb{E}[Y^2]$.

Exercice 4. Soient $n\geq 1$ et $\theta>0$. On considère des variables aléatoires $U_1,\dots,U_n$ indépendantes qui sont distribuées uniformément sur le segment $[0,\theta]$. On pose

$$S=\frac{n+1}{n}\cdot\max(U_1,\dots,U_n).$$

Montrer que S est une variable aléatoire réelle à densité, et donner une expression de sa densité.

Problème 1. On considère une assemblée qui comporte 1600 députés, qui ont formé 16000 commissions différentes, chacune comportant 80 députés. Le but de l'exercice est de prouver qu'il existe forcément au moins deux commissions qui ont au moins quatre députés en commun. Pour cela, nous allons utiliser des probabilités !

Il est possible d'admettre le résultat d'une question à condition de le marquer explicitement, et de l'utiliser ultérieurement.

(1) Soient $a, b \geq 0$ deux nombres réels positifs. Montrer que $a^2 + b^2 \geq \frac{1}{2}(a+b)^2$.

(2) Soit $n \geq 2$ un entier. On considère maintenant $a_1, \ldots, a_n \geq 0$ des nombres réels positifs.

   a) On note $P(x)$ le polynôme de degré deux défini par $P(x) = \sum_{i=1}^{n} (a_i x - 1)^2$. Quel est le discriminant de P ? En déduire que
$$\sum_{i=1}^{n} a_i^2 \geq \frac{1}{n} \left(\sum_{i=1}^{n} a_i \right)^2 .$$

   b) On pose $b = (a_1 + \cdots + a_n)/n$. Montrer que
$$\frac{1}{n} \sum_{i=1}^{n} \binom{a_i}{2} \geq \frac{b(b-1)}{2}.$$

(3) Soit X une variable aléatoire à valeurs dans $\mathbb{N} = \{0, 1, 2, \ldots\}$, définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, et qui admet une espérance. Montrer qu'il existe $\omega \in \Omega$ tel que $X(\omega) \geq \mathbb{E}[X]$.

(4) Pour $1 \leq i \leq 1600$, notons a_i le nombre de commissions auquel appartient le i-ième député.

   a) Que vaut $\sum_{i=1}^{1600} a_i$?

   b) Soit C le nombre de manières différentes qu'on choisir deux commissions différentes parmi les 16000 commissions. Que vaut C ? Pour $1 \leq i \leq 1600$, de combien de manières différentes peut-on choisir deux commissions différentes contenant toutes les deux le i-ième député (on exprimera la réponse en fonction de a_i) ?

   c) Choisissons deux commissions différentes au hasard, uniformément parmi les C manières de choisir deux commissions différentes. Ensuite, pour $1 \leq i \leq 1600$, notons X_i la variable aléatoire définie par
$$X_i = \begin{cases} 1 & \text{si le } i\text{-ième député appartient aux deux commissions choisies} \\ 0 & \text{sinon.} \end{cases}$$

   d) Soit $X = X_1 + X_2 + \cdots + X_{1600}$. Que représente la quantité X ?

   e) Pour $1 \leq i \leq 1600$, montrer que $\mathbb{P}(X_i = 1) = \binom{a_i}{2} / \binom{16000}{2}$.

   f) Montrer que $\mathbb{E}[X] \geq 3.995$.

   g) En déduire qu'il existe deux commissions qui ont au moins quatre députés en commun.

Problème 2. On fixe un entier $N \geq 1$. Soient $X_1, X_2, \ldots, X_n, \ldots$ des variables aléatoires indépendantes de loi uniforme sur $\{1, 2, \ldots, N\}$. On imagine que les X_i sont des cartes à collectionner vendues chez le marchand de journaux. À chaque achat on a une image dans la collection $\{1, \ldots, N\}$ et on cherche à réunir toute la collection (c'est-à-dire une image de chaque sorte) sachant qu'on peut tomber plusieurs fois sur la même... Plus formellement on note pour $i \in \{0, 1, 2, 3, \ldots, N\}$

$$T_i^{(N)} = \inf\{k \geq 1 : \mathrm{Card}\{X_1, X_2, \ldots, X_k\} = i\},$$

et par convention $T_0^{(N)} = 0$. C'est donc le premier instant où on a i images différentes et $T_N^{(N)}$ est le nombre d'images qu'il a fallu acheter pour avoir la collection complète.

(1) Justifier (en français) pourquoi les variables aléatoires $T_{i+1}^{(N)} - T_i^{(N)}$ sont (mutuellement) indépendantes pour $i \in \{0, \ldots, N-1\}$.

(2) Montrer que $T_{i+1}^{(N)} - T_i^{(N)}$ suit une loi géométrique dont on calculera le paramètre.

(3) En déduire l'espérance et la variance de la variable $T_N^{(N)}$.

(4) On rappelle que $\sum_{k \geq 1} k^{-2} = \frac{\pi^2}{6}$. Montrer que lorsque $N \to \infty$, on a

$$\mathbb{E}\left[T_N^{(N)}\right] \sim N \log(N) \quad \text{et} \quad \mathsf{Var}(T_N^{(N)}) \sim \frac{\pi^2}{6} N^2.$$

(5) Soit X une variable aléatoire réelle discrète qui admet un moment d'ordre 2. Montrer que pour tout $\epsilon > 0$,

$$\mathbb{P}\left(X > \epsilon\right) \leq \frac{\mathbb{E}\left[X^2\right]}{\epsilon^2}.$$

(6) En déduire que pour tout $\epsilon > 0$ on a

$$\mathbb{P}\left(\left|\frac{T_N^{(N)}}{N \log N} - 1\right| > \epsilon\right) \quad \underset{N \to \infty}{\longrightarrow} \quad 0.$$

(7) Pour tout $i \in \{1, \ldots, N\}$ on note $A_{i,m}$ l'évènement $\{i \notin \{X_1, \ldots, X_m\}\}$. Calculer la probabilité de $A_{i,m}$.

(8) En déduire que pour tout $x \geq 0$ on a

$$\mathbb{P}(T_N^{(N)} \geq \log(N)N + xN) \leq e^{-x}.$$

Problème 3. On fixe un entier $n \geq 1$. Soient $X_1, \dots, X_n$ des variables aléatoires réelles discrètes indépendantes qui admettent toutes un moment d'ordre 2. On suppose que $\mathbb{E}[X_i] = 0$ pour tout $1 \leq i \leq n$ (mais on ne suppose pas que ces variables ont même loi). On pose $S_n = X_1 + \cdots + X_n$. Le but de ce problème est de montrer que

$$\mathbb{P}\left(\max_{1\leq k\leq n} |S_k| \geq t\right) \leq \frac{1}{t^2} \sum_{1\leq i\leq n} \mathrm{Var}(X_i).$$

(Cette inégalité est connue sous le nom de « Inégalité maximale de Kolmogorov ».)

Il est possible d'admettre le résultat d'une question à condition de le marquer explicitement, et de l'utiliser ultérieurement.

(1) Pour $1 \leq k \leq n$, calculer $\mathbb{E}[S_n - S_k]$.

(2) Exprimer $\mathrm{Var}(S_n)$ en fonction de $\mathrm{Var}(X_1), \dots, \mathrm{Var}(X_n)$.

(3) Soit A l'événement $A = \{\max_{1\leq k\leq n} |S_k| \geq t\}$. Pour $1 \leq k \leq n$, on introduit l'événement

$$A_k = \bigcap_{1\leq j<k} \{|S_j| < t\} \cap \{|S_k| \geq t\}.$$

Montrer que les événements $A_1, A_2, \dots, A_n$ sont deux à deux disjoints et que $A = A_1 \cup A_2 \cup \cdots \cup A_n$.

(4) Montrer que $\mathbb{E}[S_k^2 \mathbb{1}_{A_k}] \geq t^2 \mathbb{P}(A_k)$, où $\mathbb{1}_{A_k}$ est la variable aléatoire qui vaut 1 si A_k est réalisé et 0 sinon.

(5) Expliquer rapidement pourquoi $\mathbb{1}_A = \sum_{k=1}^n \mathbb{1}_{A_k}$ et en déduire que $\mathbb{E}[S_n^2] \geq \sum_{k=1}^n \mathbb{E}[S_n^2 \mathbb{1}_{A_k}]$.

(6) Montrer que

$$\mathbb{E}\left[S_n^2 \mathbb{1}_{A_k}\right] \geq \mathbb{E}\left[S_k^2 \mathbb{1}_{A_k}\right] + 2\mathbb{E}[S_n - S_k]\,\mathbb{E}[S_k \mathbb{1}_{A_k}].$$

Indication. On pourra écrire $S_n^2 = (S_k + (S_n - S_k))^2$.

(7) En déduire que $\mathbb{E}[S_n^2] \geq t^2 \mathbb{P}(A)$ et conclure.

(8) *Application :* majorer $\mathbb{P}(\max_{1\leq k\leq n} |S_k| \geq t)$ lorsque X_i suit une loi de Poisson de paramètre i.