Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2015

Épreuve de mathématiques

Durée : 4h

L'objet du problème est l'étude du nombre de records d'une permutation via la suite de variables aléatoires $(R_n)_{n\in\mathbb{N}^*}$ introduite dans la partie II. La partie III utilise les notations et des résultats de la partie II.

Partie I Une distance entre lois de variables aléatoires

Toutes les variables aléatoires de cette partie sont définies sur le même espace probabilisé $(\Omega, \mathscr{T}, \mathbf{P})$.

1. Soit X, Y et Z trois variables aléatoires à valeurs dans $\mathbb{N}$.

a) Justifier la convergence de la série de terme général $\Big|\mathbf{P}([X=n]) - \mathbf{P}([Y=n])\Big|$.

On note $d(X,Y) = \dfrac{1}{2}\displaystyle\sum_{n=0}^{+\infty}\Big(\Big|\mathbf{P}([X=n]) - \mathbf{P}([Y=n])\Big|\Big)$.

b) Que dire des variables X et Y quand $d(X,Y)=0$?

Donner un exemple simple de deux variables X et Y distinctes et telles que $d(X,Y)=0$.

c) Prouver l'inégalité : $d(X,Z) \leq d(X,Y) + d(Y,Z)$.

2. a) On pose $A = \Big\{k \in \mathbb{N} ; \mathbf{P}([X=k]) \geq \mathbf{P}([Y=k])\Big\}$. Vérifier l'égalité :

$$d(X,Y) = \Big|\mathbf{P}([X \in A]) - \mathbf{P}([Y \in A])\Big|.$$

b) Soit B une partie de $\mathbb{N}$.

Prouver l'inégalité : $\Big|\mathbf{P}([X \in B]) - \mathbf{P}([Y \in B])\Big| \leq d(X,Y)$.

On pourra faire intervenir les parties $B \cap A$ et $B \cap A^c$.

3. Soit $p \in [0,1]$, X une variable suivant la loi de Bernoulli de paramètre p et Y une variable suivant la loi de Poisson de paramètre p.

a) Justifier, pour tout réel x, l'inégalité : $\mathrm{e}^x \geq 1 + x$.

b) Établir l'égalité $d(X,Y) = p(1-\mathrm{e}^{-p})$ et en déduire la majoration $d(X,Y) \leq p^2$.

Dans la fin de cette partie, on considère un entier n au moins égal à 3, un entier $N > n$, et on note I la matrice identité de $\mathrm{M}_N(\mathbb{R})$ et R la matrice de $\mathrm{M}_N(\mathbb{R})$ dont tous les coefficients sont nuls sauf ceux de la sur-diagonale qui valent 1, c'est-à-dire $R = \begin{pmatrix} 0 & 1 & & (0) \\ & \ddots & \ddots & \\ & (0) & \ddots & 1 \\ & & & 0 \end{pmatrix}$.

On considère une n-liste $(p_1, p_2, \ldots, p_n)$ de réels de $[0,1]$, n variables aléatoires X_1, X_2,..., X_n (resp. Y_1, Y_2,..., Y_n), indépendantes suivant des lois de Bernoulli (resp. Poisson) de paramètres respectifs p_1, p_2,..., p_n.

On note, pour tout entier $i \in [\![1,n]\!]$, $P_i = (1-p_i)I + p_i R$ et $Q_i = p_i(R-I)$.

4. a) Montrer que la première ligne du produit P_1P_2 est constituée des $n+1$ réels $\mathbf{P}([X_1+X_2=0])$, $\mathbf{P}([X_1+X_2=1])$, ..., $\mathbf{P}([X_1+X_2=n])$ suivis de termes nuls.

b) On note $U_n = \displaystyle\sum_{k=1}^{n} X_k$. Montrer que la première ligne du produit $P_1P_2\ldots P_n$ est constituée des $n+1$ réels $\mathbf{P}([U_n=0])$, $\mathbf{P}([U_n=1])$, ..., $\mathbf{P}([U_n=n])$ suivis de termes nuls.

5. On rappelle (ou on admet) que si $A \in \mathrm{M}_N(\mathbb{R})$, la suite de terme général $\displaystyle\sum_{k=0}^{r}\frac{A^k}{k!}$ converge, quand r tend vers $+\infty$, vers une matrice notée $\exp A$. *La convergence d'une suite de matrices s'entendant comme la convergence coefficient par coefficient.* On **admet** que si deux matrices A et B de $\mathrm{M}_N(\mathbb{R})$ commutent alors on a $\exp(A+B) = \exp A \times \exp B$.

a) Soit i un entier de $[\![1,n]\!]$ et r un entier naturel. Prouver l'égalité :

$$\sum_{k=0}^{r} \frac{Q_i^k}{k!} = \sum_{j=0}^{r} \frac{p_i^j}{j!}\Big(\sum_{k=0}^{r-j} \frac{(-1)^k\, p_i^k}{k!}\Big) R^j.$$

b) En déduire que $\exp Q_i = \displaystyle\sum_{j=0}^{N-1} \frac{p_i^j\, \mathrm{e}^{-p_i}}{j!} R^j$.

c) On note $V_n = \displaystyle\sum_{k=1}^{n} Y_k$. Montrer que la première ligne du produit $\exp Q_1 \times \exp Q_2 \times \ldots \times \exp Q_n$ est constituée par les N réels $\mathbf{P}([V_n = 0])$, $\mathbf{P}([V_n = 1])$, …, $\mathbf{P}([V_n = N-1])$.

6. On note, pour toute matrice $A = (a_{i,j})_{1\le i,j\le N}$ de $\mathrm{M}_N(\mathbb{R})$, $\|A\| = \displaystyle\max_{1\le i\le N} \sum_{j=1}^{N} |a_{i,j}|$.

a) Prouver, pour tout couple (A,B) d'éléments de $\mathrm{M}_N(\mathbb{R})$ les inégalités

$$\|A+B\| \le \|A\| + \|B\| \quad \text{et} \quad \|AB\| \le \|A\|\,\|B\|.$$

b) Prouver, pour tout entier $i \in [\![1,n]\!]$, l'inégalité $\left\|\exp Q_i\right\| \le 1$.

c) En remarquant que

$$\prod_{i=1}^{n} P_i - \prod_{i=1}^{n} \exp Q_i = (P_1 - \exp Q_1)\Big(\prod_{i=2}^{n} P_i\Big) + (\exp Q_1)\Big(\prod_{i=2}^{n} P_i - \prod_{i=2}^{n} \exp Q_i\Big),$$

prouver l'inégalité

$$\left\|\prod_{i=1}^{n} P_i - \prod_{i=1}^{n} \exp Q_i\right\| \le \sum_{i=1}^{n} \left\|P_i - \exp Q_i\right\|.$$

d) En exprimant, pour tout entier i de $[\![1,n]\!]$, la matrice $P_i - \exp Q_i$ à l'aide des puissances de la matrice R, prouver l'inégalité $\left\|P_i - \exp Q_i\right\| \le 2p_i^2$.

7. a) Déduire des questions précédentes l'inégalité $d(U_n, V_n) \le \displaystyle\sum_{i=1}^{n} p_i^2$.

b) Soit $\lambda > 0$. En appliquant le résultat précédent, retrouver la propriété d'approximation usuelle de la loi binomiale de paramètres n et $\dfrac{\lambda}{n}$ par une loi de Poisson.

Partie II Records d'une permutation

On rappelle qu'une permutation d'un ensemble non vide E est une bijection de E sur E. On note, pour tout entier naturel n non nul, $\mathscr{S}_n$ l'ensemble des permutations de l'ensemble $[\![1,n]\!]$ (des entiers k tels que $1 \leq k \leq n$) et $\sigma := (\sigma_1, \sigma_2, \ldots, \sigma_n)$ la permutation σ qui envoie l'entier 1 sur σ_1, 2 sur σ_2, ..., n sur σ_n.
On note $\mathscr{P}(\mathscr{S}_n)$ l'ensemble des parties de $\mathscr{S}_n$ et on munit l'espace probabilisable $(\mathscr{S}_n, \mathscr{P}(\mathscr{S}_n))$ de la probabilité uniforme notée $\mathbf{P}$. Ainsi, pour tout élément σ de $\mathscr{S}_n$, on a $\mathbf{P}(\{\sigma\}) = \dfrac{1}{n!}\cdot$

Soit $n \in \mathbb{N}^*$ et $k \in [\![1,n]\!]$. On dit qu'une permutation $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ présente un record au rang k si, pour tout entier i vérifiant $1 \leq i \leq k$, on a $\sigma_i \leq \sigma_k$. Ainsi, en particulier, toute permutation présente un record au rang 1. On note $R_n(\sigma)$ le nombre de records que compte la permutation σ. On définit ainsi une variable aléatoire R_n sur $\mathscr{S}_n$. Bien sûr, la variable R_1 est certaine égale à 1. On note, pour tout entier naturel n non nul, $H_n = \displaystyle\sum_{k=1}^{n} \frac{1}{k}\cdot$

1. Montrer que la suite $(H_n - \ln n)_{n \in \mathbb{N}^*}$ est monotone convergente. On note γ la limite de cette suite.

Dans toute la suite de cette partie **II** on considère un entier n au moins égal à 3.

2. Déterminer la loi de R_3, son espérance et sa variance.

3. Déterminer les probabilités $\mathbf{P}([R_n = 1])$ et $\mathbf{P}([R_n = n])$.

4. **a)** Soit p un entier de $[\![2,n]\!]$. Déterminer le nombre de permutations de $[\![1,n]\!]$ ayant exactement 2 records lesquels sont atteints aux rangs 1 et p.

b) Prouver l'égalité : $\mathbf{P}([R_n = 2]) = \dfrac{1}{n} \displaystyle\sum_{k=1}^{n-1} \frac{1}{k}\cdot$

c) Donner un équivalent de $\mathbf{P}([R_n = 2])$ quand n tend vers l'infini.

5. On introduit, pour tout entier i de $[\![1,n]\!]$, la variable aléatoire T_i, définie sur $(\mathscr{S}_n, \mathscr{P}(\mathscr{S}_n), \mathbf{P})$, qui, à chaque élément σ de $\mathscr{S}_n$, associe la valeur 1 si σ présente un record au rang i et égale à 0 sinon.

a) Montrer que, pour tout entier i de $[\![1,n]\!]$, T_i suit la loi de Bernoulli de paramètre $\dfrac{1}{i}\cdot$

b) Calculer l'espérance de la variable R_n et en donner un équivalent simple quand n tend vers l'infini.

6. **a)** Soit (i,j) un couple d'entiers vérifiant $2 \leq i < j \leq n$.
En calculant la probabilité $\mathbf{P}([T_i = 1] \cap [T_j = 1])$, justifier l'indépendance des variables T_i et T_j.

b) Calculer la variance de la variable R_n et en donner un équivalent simple quand n tend vers l'infini.

7. On **admet** l'indépendance mutuelle des variables T_1, T_2,..., T_n.
Établir, pour tout entier k de $[\![2,n]\!]$, l'égalité :

$$\mathbf{P}([R_n = k]) = \sum_{2 \leq i_2 < i_3 < \cdots < i_k \leq n} \frac{1}{i_2} \frac{1}{i_3} \cdots \frac{1}{i_k} \prod_{\substack{2 \leq j \leq n \\ j \notin \{i_2, \ldots, i_k\}}} (1 - \frac{1}{j}).$$

a) En déduire l'égalité : $\mathbf{P}([R_n = 3]) = \dfrac{1}{n} \displaystyle\sum_{1 \le i < j \le n-1} \frac{1}{i\,j}$.

b) Prouver l'équivalence : $\mathbf{P}([R_n = 3]) \sim \dfrac{1}{2}\dfrac{\ln^2 n}{n}$.

Partie III Deux résultats asymptotiques

1. `Un premier résultat`

a) Soit $\varepsilon > 0$. Prouver, pour tout entier n assez grand, l'inclusion entre événements :

$$\left[\,\left|\frac{R_n}{\ln n} - 1\right| \ge \varepsilon\,\right] \subset \left[\,\left|\frac{R_n}{\ln n} - \frac{H_n}{\ln n}\right| \ge \frac{\varepsilon}{2}\,\right].$$

b) Soit $\varepsilon > 0$.

(i) Établir l'égalité : $\displaystyle\lim_{n \to +\infty} \mathbf{P}\left(\left[\,\left|\frac{R_n}{\ln n} - \frac{H_n}{\ln n}\right| \ge \varepsilon\,\right]\right) = 0$.

(ii) En déduire l'égalité : $\displaystyle\lim_{n \to +\infty} \mathbf{P}\left(\left[\,\left|\frac{R_n}{\ln n} - 1\right| \ge \varepsilon\,\right]\right) = 0$.

2. `Un second résultat`

On rappelle que si X est une variable aléatoire à valeurs dans $\mathbb{N}$ on note G_X sa fonction génératrice c'est-à-dire la fonction qui, à chaque réel t de $[0,1]$, associe

$$G_X(t) = \mathbf{E}(t^X) = \sum_{k=0}^{+\infty} \mathbf{P}([X = k])\,t^k.$$

a) Déterminer la fonction génératrice d'une variable de Bernoulli de paramètre $p \in [0,1]$.

b) Déterminer la fonction génératrice d'une variable de Poisson de paramètre $\lambda > 0$.

c) Soit $n \in \mathbb{N}^*$, et $X_1, X_2, \ldots, X_n$ des variables aléatoires, à valeurs dans $\mathbb{N}$, mutuellement indépendantes et $S_n = \displaystyle\sum_{k=1}^{n} X_k$. Justifier l'égalité : $G_{S_n} = \displaystyle\prod_{k=1}^{n} G_{X_k}$.

d) Soit $Y, X_1, X_2, \ldots$ une suite de variables aléatoires à valeurs dans $\mathbb{N}$. On **admet** le résultat suivant : si, pour tout réel t de $[0,1]$, la suite $(G_{X_n}(t))_{n \in \mathbb{N}^*}$ converge vers $G_Y(t)$, alors la suite $(X_n)_{n \in \mathbb{N}}$ converge en loi vers Y, c'est-à-dire que, pour tout entier naturel k,

$$\lim_{n \to +\infty} \mathbf{P}([X_n = k]) = \mathbf{P}([Y = k]).$$

Soit m un entier naturel au moins égal à 2 et $n = 2m$. On conserve les notations de la partie II et on pose $W_n = \displaystyle\sum_{k=m+1}^{2m} T_k$ (qui compte le nombre aléatoire de records que présente une permutation de $[\![1, 2m]\!]$ entre les rangs $m+1$ et $2m$).

Prouver, pour tout réel t de $[0,1]$, l'égalité : $G_{W_n}(t) = \displaystyle\prod_{i=m+1}^{2m} \left(1 + \frac{t-1}{i}\right)$.

e) En déduire que, lorsque l'entier m tend vers $+\infty$, la suite de terme général W_n converge en loi vers une variable aléatoire qu'on identifiera.