Travaux dirigés, feuille 1 : Probabilités discrètes

Exercice 1
Une urne contient b boules bleues et r boules rouges. On les tire une à une sans remise.

1) A l'aide du dénombrement, calculer la probabilité que la première boule rouge soit obtenue au k-ième tirage. Peut-on faire ce calcul indifféremment dans
   a) l'ensemble des permutations de $b+r$ objets
   b) l'ensemble des suites de $b+r$ éléments à valeurs dans {bleu, rouge} dont b "bleu" et r "rouge"
   c) l'ensemble des choix ordonnés sans remise de k objets parmi $b+r$
   d) l'ensemble des suites de k éléments à valeurs dans {bleu, rouge} dont au plus b "bleu" et r "rouge"
muni de la probabilité uniforme ? Pourquoi ?

2) Retrouver ce résultat à l'aide du conditionnement.

Exercice 2
On dispose de deux urnes A et B contenant des proportions respectives de boules blanches et noires de $(\frac{1}{3}, \frac{2}{3})$ et de $(\frac{2}{3}, \frac{1}{3})$. On tire une des urnes au hasard avec équiprobabilité puis on procède, avec cette urne, à une série de tirages avec remise.

1) Quelle est la probabilité que la n-ième boule tirée soit blanche ?

2) Quelle est la probabilité que la 3-ième boule tirée soit noire sachant que les deux premières l'ont été ?

3) Quelle est la probabilité que l'on ait choisi l'urne A sachant que les n premières boules tirées sont noires ?

Exercice 3
Soient X et Y deux variables aléatoires indépendantes de loi géométrique sur $\mathbb{N}$ de paramètre $0 < 1-a < 1$ (i.e. $\forall k \in \mathbb{N}$, $\mathbb{P}[X=k] = \mathbb{P}[Y=k] = a^k(1-a)$). On pose $M = \min(X,Y)$ et $Z = Y - X$.

1) Quelle est la loi de M ? Indication : on pourra calculer $\mathbb{P}(M \geq k)$.

2) Quelle est la loi du couple (M, Z) ?

3) En déduire la loi de Z. Que peut-on dire de M et Z ?

Problème 1 : la marche aléatoire sur $\mathbb{Z}$

Ce problème répond à la question : le promeneur au hasard sur $\mathbb{Z}$ revient-il à son point de départ ? On considère une suite de variables aléatoires i.i.d. $(X_n)_{n\geq 1}$ sur $(\Omega, \mathcal{A}, \mathbb{P})$ de loi de Bernoulli à valeurs dans $\{+1, -1\}$ de paramètre $p \in]0,1[$, i.e. $\mathbb{P}_{X_1} = p\delta_1 + (1-p)\delta_{-1}$. On étudie leurs sommes partielles $S_n = X_1 + \ldots + X_n$: S_n représente la position du promeneur à l'instant n, X_n correspond au pas qu'il fait entre l'instant $n-1$ et l'instant n. On note N ($N \geq 0$) le nombre de retours du promeneur en 0, i.e., $N = \sum_{n\geq 1} \mathbb{1}_{S_n=0}$.

1. (a) Montrer que $\mathbb{P}(N \geq m+1) = \mathbb{P}(N \geq 1)\mathbb{P}(N \geq m)$, puis que $\mathbb{P}(N \geq m) = \mathbb{P}(N \geq 1)^m$ pour $m \geq 1$.
   Indication : on pourra définir $T = \inf\{n \geq 1 : S_n = 0\}$ et remarquer que $\{N \geq m+1\} = \bigcup_{k\geq 1} (\{T = k\} \cap \{N \geq m+1\})$.

   (b) Déterminer la loi de N.

   (c) En déduire que,
   — soit $\mathbb{E}[N] < \infty$ (auquel cas $N < \infty$ p.s.),
   — soit $\mathbb{E}[N] = \infty$ et alors $N = \infty$ p.s..
   Dans le premier cas, on dit que 0 est transient (ou transitoire), dans le second cas, on dit que 0 est recurrent.
2. (a) Exprimer l'espérance $\mathbb{E}X_n$ en fonction de $\mu = 2p-1$. Montrer que pour tout $n \geq 1$
   $$\mathbb{E}\big[S_n\big] = n\mu\,, \qquad \mathbb{E}\big[(S_n - n\mu)^2\big] = 4np(1-p)$$
   et pour tout $n \geq 2$
   $$\mathbb{E}\big[(S_n - n\mu)^4\big] \leq 6n^2$$

   (b) En notant
   $$Y = \sum_{n\geq 1} \left(\frac{S_n - n\mu}{n}\right)^4 \,,$$
   en déduire que $\mathbb{E}[Y] < \infty$, puis que $\lim_{n\to\infty} S_n/n = \mu$ p.s. Conclure alors que
   $$p > 1/2 \Rightarrow \lim_{n\to\infty} S_n = +\infty \quad \text{p.s.}$$
   $$p < 1/2 \Rightarrow \lim_{n\to\infty} S_n = -\infty \quad \text{p.s.}$$

   (c) Avec la question 1), conclure que si $p \neq 1/2$, 0 est transient.
3. (a) On suppose maintenant que $p = 1/2$. Calculer $\mathbb{P}(S_{2n} = k)$ pour $k \in \mathbb{Z}$.

   (b) En déduire que
   $$\forall k \in \mathbb{Z}\,, \qquad \mathbb{P}(S_{2n} = k) \leq \mathbb{P}(S_{2n} = 0)$$
   puis que $\mathbb{P}(S_{2n} = 0) \geq 1/(4n+1)$ (on pourra remarquer que $|S_{2n}| \leq 2n$).

   (c) Conclure que 0 est recurrent si $p = 1/2$.

Problème 2 : le collectionneur de coupons

Soit $n \geq 1$. Soit $(X_k^n)_{k\geq 1}$ une suite de v.a. i.i.d. de loi uniforme sur $\{1, 2, \ldots, n\}$, i.e., pour tout $i \in \{1, ..., n\}$ on a $\mathbb{P}(X_k^n = i) = 1/n$. Soit

$$T_n = \inf\{m \geq 1 \,|\, \{X_1^n, \ldots, X_m^n\} = \{1, \ldots, n\}\}$$

le premier instant où toutes les valeurs possibles ont été observées.

1) Montrer que $T_n < \infty$ p.s.

2) Soit $\tau_k^n = \inf\{m \geq 1 \,|\, \mathrm{Card}\{X_1^n, \ldots, X_m^n\} = k\}$, pour $k \in \{1 \ldots n\}$, et $\sigma^n = (\sigma_1^n, \ldots, \sigma_n^n)$ où $\sigma_k^n = X_{\tau_k^n}^n$. Montrer que les variables $(\sigma^n, \tau_2^n - \tau_1^n, \ldots, \tau_n^n - \tau_{n-1}^n)$ sont indépendantes, et déterminer leur loi respective.

3) Exprimer T_n en fonction de τ_k^n pour $k \in \{1 \ldots n\}$. Exprimer $\mathbb{E}(T_n)$ sous forme de somme, puis en donner un équivalent quand n tend vers l'infini.

4) Montrer que $\mathrm{Var}(T_n) \leq Cn^2$ où C est une constante. En déduire que $\frac{T_n}{n \ln n}$ converge vers 1 en probabilité, c'est-à-dire que

$$\forall \varepsilon > 0\,, \quad \lim_{n\to\infty} \mathbb{P}\left[\left|\frac{T_n}{n \ln n} - 1\right| > \varepsilon\right] = 0\,.$$

Problème 3 : le graphe aléatoire d'Erdös-Rényi

Un graphe est un ensemble de points, dont certaines paires sont directement reliées par un lien. On considère, pour $n \geq 1$ entier et $p \in (0,1)$, un ensemble de n points (appelés sommets) noté $V = \{1, \ldots, n\}$. Pour chacune des $n(n-1)/2$ paires de sommets $\{i, j\}$ avec $i \neq j$, on note $T_{i,j} = T_{j,i} = 1$ (respectivement, $T_{i,j} = T_{j,i} = 0$) selon que i et j sont *directement reliés* ou non ; on dit dans le premier cas qu'il y a une arête qui les joint. Notez qu'un sommet n'est pas directement relié à lui-même. Le graphe est défini par la donnée des sommets et des arêtes présentes.

On suppose que les variables $(T_{i,j})_{1 \leq i < j \leq n}$ sont indépendantes et de même loi de Bernoulli de paramètre p, i.e., $\mathbb{P}(T_{i,j} = 1) = p, \mathbb{P}(T_{i,j} = 0) = 1 - p$.

1) Dans cette question, on ne fait aucune hypothèse supplémentaire sur p.

   a) Quelle est la loi du nombre $N_n = \sum_{1 \leq i < j \leq n} \mathbb{1}_{\{T_{i,j}=1\}}$ d'arêtes du graphe ? Quelle est celle du nombre D_i de sommets directement reliés à un sommet i fixé ?

   b) Un sommet i est dit isolé s'il n'est relié à aucun sommet dans le graphe, i.e. si $D_i = 0$. Calculer la probabilité que le sommet 1 soit isolé. Vérifier que le nombre Z_n de sommets isolés est une somme de n variables aléatoires de Bernoulli *non indépendantes*. Calculer $\mathbb{E}Z_n$.

2) On suppose que $p = p_n = \frac{\lambda \ln n}{n}$, avec $\lambda > 1$. Montrer que $\lim_{n\to\infty} \mathbb{E}Z_n = 0$, puis que $\lim_{n\to\infty} \mathbb{P}(Z_n \geq 1) = 0$.

3) On suppose que $p = p_n = \frac{\theta}{n}$, avec $\theta > 0$.

   a) Soit k un entier. Quelle est la limite de $\mathbb{P}(D_i = k)$ quand $n \to \infty$?

   b) On dit que le triangle ijk est dans le graphe si $T_{i,j} = T_{j,k} = T_{i,k} = 1$; on identifie les triangles ijk, jki, et tous ceux obtenus par permutation. Calculer l'espérance du nombre T_n de triangles dans le graphe, puis sa limite quand $n \to \infty$.

4) On suppose maintenant que $p = p_n = \frac{\lambda \ln n}{n}$, avec $0 < \lambda < 1$.

   a) Pour $i \neq j$, calculer $\mathbb{P}(D_i = D_j = 0)$ en fonction de p, puis la valeur de $\mathbb{E}(Z_n^2)$.

   b) Pour toute variable aléatoire positive Z, montrer que $\mathbb{P}(Z = 0) \leq \mathrm{Var}(Z)/(\mathbb{E}Z)^2$.

   c) En conclure enfin que $\lim_{n\to\infty} \mathbb{P}(Z_n \geq 1) = 1$