TD 02 – Variables Aléatoires (corrigé)

Exercice 1. *Indépendance*

Soient X,Y et Z des variables aléatoires indépendantes et identiquement distribuées, de loi uniforme sur $\{-1, 1\}$. Les affirmations suivantes sont-elles vraies ou fausses ?

1. XY et Z sont indépendantes.
   ☞ Oui (lemme de groupement par paquets).
2. XY et YZ sont indépendantes.
   ☞ Oui, on peut le calculer. Pour tous $a, b \in \{-1, 1\}$, on a
   $$\mathbf{P}\{XY = a, YZ = b\} = \frac{1}{2}\mathbf{P}\{X = a, Z = b \mid Y = 1\} + \frac{1}{2}\mathbf{P}\{X = -a, Z = -a \mid Y = -1\} = \frac{1}{4}.$$
3. X, Y et XY sont indépendantes.
   ☞ Non, $\mathbf{P}\{(X = 1) \cap (Y = 1) \cap (XY = -1)\} = 0$ par exemple.

Exercice 2. *Rouge et Vert*

1. Supposons que l'on commence avec une urne contenant 2 boules, une rouge et une verte. On répète la procédure suivante jusqu'à ce que l'urne contienne n boules : à chaque étape, on tire une balle uniformément de l'urne, et on la remet ainsi qu'une autre boule de même couleur dans l'urne. Montrer que le nombre de boules rouges a la même probabilité d'être n'importe quel nombre entre 1 et $n - 1$.

   ☞ Soit X_n la variable aléatoire qui compte le nombre de balles rouges au début de la $(n - 1)$ième étape (c'est-à-dire lorsqu'il y a n boules dans la corbeille). Ainsi $\mathbf{P}\{X_2 = 1\} = 1$ puisqu'au début de la première étape, il y a exactement une balle rouge dans la corbeille.

   Soit Y_n la variable aléatoire qui vaut 1 si la balle tiré à la $(n - 1)$ième étape est rouge, et 0 sinon.

   Montrons par récurrence sur n que $\forall i \in \{1, \dots, n - 1\}, \mathbf{P}\{X_n = i\} = \frac{1}{n-1}$.

   Pour $i \in \{2, \dots n - 2\}$, on a $X_n = i$ si et seulement l'un des deux cas suivants se présente :

— Il y avait précédemment i balles rouges, et l'on a tiré une boule verte ;
— Il y avait précédemment $i-1$ balles rouges, et l'on a tiré une boule rouge.

Comme ces deux évenements sont disjoints, on a :

$$\begin{aligned}
\mathbf{P}\{X_n = i\} &= \mathbf{P}\{X_{n-1} = i \text{ et } Y_{n-1} = 0\} + \mathbf{P}\{X_{n-1} = i-1 \text{ et } Y_{n-1} = 1\} \\
&= \mathbf{P}\{Y_{n-1} = 0 \mid X_{n-1} = i\} \cdot \mathbf{P}\{X_{n-1} = i\} \\
&\quad + \mathbf{P}\{Y_{n-1} = 1 \mid X_{n-1} = i-1\} \cdot \mathbf{P}\{X_{n-1} = i-1\} \\
&= \tfrac{(n-1-i)}{(n-1)} \cdot \tfrac{1}{n-2} + \tfrac{i-1}{n-1} \cdot \tfrac{1}{n-2} \text{ par hypothèse de récurrence} \\
&= \tfrac{n-2}{n-1} \cdot \tfrac{1}{n-2} = \tfrac{1}{n-1}
\end{aligned}$$

De plus, par le même genre d'arguments,

$$\mathbf{P}\{X_n = 1\} = \mathbf{P}\{X_{n-1} = 1 \text{ et } Y_{n-1} = 0\} = \frac{1}{n-2} \cdot \frac{n-2}{n-1} = \frac{1}{n-1}$$

et

$$\mathbf{P}\{X_n = n-1\} = \mathbf{P}\{X_{n-1} = n-2 \text{ et } Y_{n-1} = 1\} = \frac{1}{n-2} \cdot \frac{n-2}{n-1} = \frac{1}{n-1}$$

2. On se donne maintenant une urne avec n boules rouges et $100-n$ boules vertes, où n est choisi uniformément entre 0 et 100. On tire aléatoirement une boule de l'urne, elle est rouge, et on la retire. La prochaine boule tirée aléatoirement a-t-elle plus de chances d'être rouge ou verte ? ☞ On va remplacer 100 par N et traiter le cas général. On écrit r_k l'événement "la k-ième boule tirée est rouge" et b_n l'événement "n boules sont rouges". On souhaite calculer $P(r_2|r_1)$. Tout d'abord, $P(b_n) = \frac{1}{N+1}$, $P(r_1|b_n) = \frac{n}{N}$ et $P(r_1 \cap r_2|b_n) = \frac{n(n-1)}{N(N-1)}$. Ensuite, on va calculer $P(r_1)$ et $P(r_1 \cap r_2)$:

$$P(r_1) = \sum_{n=0}^{N} P(r_1|b_n)P(b_n) = \frac{1}{N+1}\sum_{n=0}^{N}\frac{n}{N} = \frac{1}{N(N+1)}\frac{(N+1)N}{2} = \frac{1}{2} \tag{1}$$

$$P(r_1 \cap r_2) = \sum_{n=0}^{N} P(r_1 \cap r_2|b_n)P(b_n) = \frac{1}{N+1}\sum_{n=0}^{N}\frac{n(n-1)}{N(N-1)} = \frac{1}{3} \tag{2}$$

Ainsi, $P(r_2|r_1) = \frac{P(r_1 \cap r_2)}{P(r_1)} = \frac{2}{3}$. Il est donc plus probable d'obtenir une autre boule de la même couleur que la première.

Exercice 3. *Le problème des rencontres*

On se donne une urne contenant n boules numérotées de 1 à n, et on procède à une succession de tirages sans remise jusqu'à vider l'urne. On s'intéresse aux évènements $E_i =$ « la $i$ème boule tirée porte le numéro i».

1. Proposer un espace de probabilité pour modéliser cette expérience.

☞ On prend :
- $\Omega =$ l'ensemble des permutations de $\{1, 2, \ldots, n\}$,
- $\mathcal{A} = \mathcal{P}(\Omega)$,
- $\mathbf{P}\{\{\sigma\}\} = 1/n!$ (équipartition).

2. Calculer la probabilité des évènements suivants :

(a) E_i pour $1 \leq i \leq n$,

(b) $E_i \cap E_j$ pour $1 \leq i < j \leq n$,

(c) $\bigcap_{j=1}^{r} E_{i_j}$ pour $1 \leq i_1 < \cdots < i_r \leq n$.

☞ E_i arrive lorsqu'on tire la boule i à l'étape i. Or, compter le nombre de permutations de n éléments où $\sigma(i) = i$ revient à compter les permutations de $n-1$ éléments (les $j \neq i$ - on pourrait s'amuser à construire la bijection entre les deux ensembles pour justifier l'égalité).

Ainsi $\mathbf{P}\{E_i\} = \dfrac{(n-1)!}{n!} = \dfrac{1}{n}$.

Pour $E_i \cap E_j$, on est ramené aux permutations de $n-2$ éléments.

Donc $\mathbf{P}\{E_i \cap E_j\} = \dfrac{(n-2)!}{n!} = \dfrac{1}{n(n-1)}$.

De même, on arrive à $\mathbf{P}\left\{\bigcap_{i_1 < \cdots < i_r} E_{i_j}\right\} = \dfrac{1}{n(n-1)\ldots(n-r+1)}$.

3. Calculer la probabilité que l'évènement E_i se produise pour au moins un i.

Quelle est la limite de cette probabilité lorsque n tend vers l'infini ?

☞ La quantité cherchée est donc $P_n = \mathbf{P}\left\{\bigcup_{1 \leq i \leq n} E_i\right\}$.

D'après la formule de Poincaré, on a :

$$\begin{aligned} P_n &= \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} \mathbf{P}\left\{\bigcap_{1 \leq j \leq k} A_{i_j}\right\} \\ &= \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k!}. \end{aligned}$$

On reconnaît presque le développement en série $e^{-1} = \sum_{k=0}^{n} \frac{(-1)^k}{k!}$. En fait, on a $P_n - 1 \to -e^{-1}$ donc la limite vaut en fait $1 - e^{-1}$.

4. Combien y a-t-il de façons de placer huit tours sur un échiquier de telle sorte qu'aucune d'entre elles en attaque une autre ? Qu'en est-il si on impose en plus que la diagonale principale soit vide ?
☞ Pour placer 8 tours avec aucune en prise, il faut mettre 1 tour par ligne et par colonne. Il suffit donc par exemple de choisir successivement la ligne pour la tour sur la colonne i pour $1 \leq i \leq 8$. C'est un tirage sans remise dans $\{1, 2, \ldots, 8\}$, donc il y a $8! = 40320$ possibilités.

Pour le deuxième cas, on est ramené au problème des rencontres pour $n = 8$ (on impose en effet que pour tout i, le $i$ème tirage ne donne pas la ligne i). Donc on a $8!/P_8 = 14833$ possibilités.

Exercice 4. *La martingale classique*
On considère une version simplifiée de la roulette où on obtient la couleur noire avec la probabilité $\frac{1}{2}$, la couleur rouge sinon. Une joueuse gagne le double de sa mise si la balle tombe sur la couleur qu'elle a choisie, elle perd sa mise sinon. Une stratégie de jeu populaire est la suivante : au premier tour, la joueuse mise 1 euro. Tant qu'elle perd, elle double sa mise (elle parie 2^{k-1} euros au k-ième tour).

1. Montrer qu'en suivant cette stratégie, la joueuse finit par gagner 1 euro.
☞ Soit T_n l'événement "le jeu dure au moins n parties". $\mathbf{P}\{T_n\} = \frac{1}{2^{n-1}}$ On peut d'abord montrer que le jeu finit presque surement, i.e. : $\mathbf{P}\{\bigcap T_n\} = \lim_n \mathbf{P}\{T_n\} = 0$ (continuité décroissante). Puis, si le joueur gagne au rang k, son gain est alors de :

$$-\sum_{i=1}^{k} 2^{i-1} + 2^k = 1 \tag{3}$$

2. Soit X la variable aléatoire qui mesure la perte maximale avant de gagner. Montrer que $\mathbf{E}[X]$ est non bornée.
☞ Soit W la v.a. qui note le tour gagnant.

$$\mathbf{E}[X] = \sum_{k=1}^{+\infty} \mathbf{P}\{W = k\} \times (2^{k-1} - 1) = \sum_{k=1}^{+\infty} \frac{2^{k-1} - 1}{2^k} \tag{4}$$

3. Soit X_j le montant gagné ou perdu lors du tour j (X_j vaut 0 si la joueuse a gagné lors d'un tour précédent.) Calculer $\mathbf{E}\left[X_j\right]$, et montrer que, en utilisant la linéarité de l'espérance, l'espérance du gain vaut 0. Est ce que la linéarité de l'espérance tient dans ce cas ?

☞ On a :

$$\mathbf{E}\left[X_j\right] = \underbrace{0}_{\text{j'ai déjà gagné}} + \underbrace{\frac{1}{2^j}}_{\text{je gagne au tour } j} \times \underbrace{2^{j-1}}_{\text{gain au tour } j} + \underbrace{\frac{1}{2^j}}_{\text{je continue après le tour } j} \times \underbrace{-2^{j-1}}_{\text{mise perdue au tour } j} = 0 \tag{5}$$

Soit G le gain.

$$\mathbf{E}\left[G\right] = \mathbf{E}\left[\sum_1^\infty X_j\right] \underbrace{=}_{\text{si linéarité}} 0 \tag{6}$$

Or, l'espérance du gain vaut 1 comme vu à la question 1. Pas de linéarité de l'espérance, car $\mathbf{E}\left[\sum^\infty |X_j|\right]$ ne converge pas.

Exercice 5. *Débiaiser des bits*

Supposons que vous ayez une machine qui produit des bits aléatoires i.i.d. de loi de Bernoulli de paramètre p, mais que vous ne connaissiez pas la valeur de $p \in]0,1[$.

1. Proposez un algorithme qui utilise la machine pour produire un bit de loi uniforme sur $\{0,1\}$.

   ☞ L'algorithme naturel est de produire des paires de bits jusqu'à obtenir une paire de bits distincts, et de retourner 0 ou 1 selon que la paire ainsi produite est 01 ou 10. Le nombre de paires de bits générées est une variable aléatoire de loi géométrique de paramètre $q := 2p(1-p)$, en particulier il est fini (et l'algorithme termine) presque sûrement.

2. On souhaite maintenant produire n bits indépendants de loi uniforme sur $\{0,1\}$. Proposez un algorithme, et déterminez une valeur de t telle que la probabilité d'utiliser la machine plus de tn fois soit inférieure à $\frac{1}{100}$ pour n assez grand. *Évidemment, plus t est petit, mieux c'est.*

   ☞ Si on répète l'algorithme précédent, le nombre d'utilisations N de la machine suit la loi de

   $$2(X_1 + \cdots + X_n)$$

   où les variables aléatoires (X_i) sont i.i.d. de loi géométrique de paramètre q. Vous êtes nombreux à avoir majoré $\mathrm{P}(N \geq t)$ par l'inégalité de Markov.

C'est correct, mais donne un résultat mauvais : on a $\mathsf{P}(N \geq tn) \leq \frac{1}{tp(1-p)}$ ce qui nécessite $t = \frac{100}{p(1-p)}$. Cela peut être amélioré en utilisant l'inégalité de Tchebychev : en effet $\mathbf{E}N = \frac{n}{p(1-p)}$ et $\mathbf{Var}\left[N\right] = O(n)$ (peu importe la constante exacte, il suffit de dire qu'une v.a. géométrique a une variance finie), donc pour tout $\varepsilon > 0$,

$$\mathsf{P}(N \geq (1+\varepsilon)\mathbf{E}N) \leq \frac{\mathbf{Var}\left[N\right]}{\varepsilon^2(\mathbf{E}N)^2} = O\left(\frac{1}{\varepsilon^2 n}\right).$$

En particulier, tout réel $t > \frac{1}{p(1-p)}$ convient, une amélioration d'un facteur 100...

Le rendement peut être amélioré car l'algorithme de la question 1 gaspille beaucoup de bits sans en extraire d'information. Par exemple, on peut regarder si les paires de (paires de bits identiques) que l'on a jetées sont identiques ou pas, et déclarer que l'on arrête l'algorithme lorsqu'elles sont différentes, en retournant 0 ou 1 selon la situation 00 11 ou 11 00. Cela se prête naturellement à une formulation récursive, que vous trouverez expliquée ici `https://en.wikipedia.org/wiki/Bernoulli_process#Randomness_extraction`. On peut montrer que tout $t > 1/h(p)$ convient, où $h(p) = -p\log_2 p - (1-p)\log_2(1-p)$ est la fonction d'entropie binaire, et que cette borne est optimale. Il est remarquable que pour p proche de $1/2$, le rendement obtenu est proche de 1 !

Exercice 6. *Running Time*

Soit $\mathcal{A}$ un algorithme déterministe qui prend en entrée une chaîne de n bits et dont l'espérance du temps d'exécution est $\mathcal{O}(n^2)$ si l'entrée est choisie aléatoirement de manière uniforme.

1. Soit $f(n)$ une fonction tendant vers $+\infty$ avec n. Montrer que la probabilité que le temps d'exécution soit supérieur à $n^2 f(n)$ tend vers zéro quand n tend vers l'infini.

☞ Le but ici est d'utiliser l'inégalité de Markov. Soit X le temps d'exécution de l'algorithme.

$$\mathbf{P}\left\{X \geq n^2 \cdot f(n)\right\} \leq \frac{\mathbf{E}\left[X\right]}{n^2 \cdot f(n)} \leq \frac{c \cdot n^2}{n^2 \cdot f(n)} \leq \frac{c}{f(n)} \longrightarrow_{n\to\infty} 0 \ .$$

2. Que pouvons nous en déduire sur le temps d'exécution dans le pire cas ?

☞ Pour avoir une borne supérieure sur le temps dans le pire cas, on utilise le fait que les entrées sont distribuées uniformément. Comme chaque entrée est choisie avec probabilité $1/2^n$, on a que si $\mathbf{P}\{X \geq t\}$ est non nulle, elle doit être au moins égale à $1/2^n$ (car au moins une entrée donnera un temps de calcul supérieur à t). On a vu à la question précédente que

$$\mathbf{P}\left\{X \geq n^2 \cdot f(n)\right\} \leq \frac{c}{f(n)}.$$

Pour que cette quantité soit inférieure à $1/2^n$, il faut que $f(n) \geq c2^n$. On en déduit que le temps d'exécution dans le pire cas est borné par $cn^2 2^n = O(n^2 2^n)$.

Exercice 7. *Répétitions dans une suite de bits aléatoires*

Soient $(X_1, \ldots, X_n)$ des variables aléatoires i.i.d. de loi uniforme sur $\{0, 1\}$. Une répétition est un sous-mot de $X_1 X_2 \cdots X_n$ du type $00 \cdots 0$ ou $11 \cdots 1$. Par exemple, la suite 00011001 contient 4 répétitions de longueur 2.

1. Pour $p > 1$ fixé, quelle est l'espérance du nombre de répétitions de longueur p ? Montrer que pour $p = \lfloor \log_2 n \rfloor$, cette espérance est de l'ordre de 1.

☞ C'est une application directe du principe de linéarité de l'espérance, où on écrit N comme une somme d'indicatrices. On obtient que l'espérance du nombre N de répétitions de longueur p vaut $\mathbf{E}[N] = \frac{n-p+1}{2^{p-1}}$.
Lorsque $p = \lfloor \log_2 n \rfloor$, on a $2 + o(1) \leq \mathbf{E}[N] \leq 4 + o(1)$.
Attention, on ne peut pas affirmer que $\mathbf{E}[N] \sim 2$: si $\{\cdot\}$ désigne la partie fractionnaire, la quantité $2^{\{\log n\}}$ oscille entre 1 et 2.

2. Montrer que pour $p \leq 0.99 \log_2 n$, la probabilité d'obtenir au moins une répétition de longueur p tend vers 1 quand n tend vers l'infini.

☞ Soit $k = \lfloor n/p \rfloor$. On considère les événements $(A_i)_{0 \leq i \leq k-1}$ définis par $A_i = \{X_{pi+1} = \cdots = X_{pi+p}\}$. Ces événements sont indépendants (groupement par paquets) et de même probabilité $1/2^{p-1}$. On a :

$$\mathbf{P}\{\text{au moins une répétition de longueur } p\} \geq \mathbf{P}\left\{\bigcup A_i\right\} = 1 - \left(1 - \frac{1}{2^{p-1}}\right)^k \geq 1 - \exp\left(\frac{k}{2^{p-1}}\right).$$

On vérifie que cette quantité tend vers 1 lorsque n tend vers l'infini dès lors que $p \leq 0.99 \log_2 n$.

Exercice 8. *Intégration*

Axel souhaite participer à un club de sa nouvelle école (un seul, pour des raisons de temps !). Pendant la semaine d'intégration, les n clubs proposent chacun une activité de découverte, dans un ordre aléatoire. Après chaque activité, Axel peut décider soit de s'inscrire à ce club (et de ne pas aller aux activités de découverte suivantes), soit de ne pas s'y inscrire et de continuer à découvrir des clubs (tout choix est définitif). Bien sûr, Axel aimerait choisir le meilleur club. Iel décide d'utiliser la stratégie suivante : d'abord, participer à m activités, sans inscription ; puis, après la m-ème activité, s'inscrire au premier club qui lui plait strictement plus que tous ceux déjà découverts (on considère qu'il n'y a pas d'ex-aequo).

1. Montrer que la probabilité qu'Axel choisisse le meilleur club est

$$\mathsf{P}_{n,m} = \frac{m}{n} \sum_{j=m+1}^{n} \frac{1}{j-1}.$$

☞ Soit X_j l'évènement "le j-ième club est le meilleur", alors pour tout j on a $\mathbf{P}\left\{X_j\right\} = \frac{1}{n}$.
Soit E_j l'évènement "le j-ième club est le meilleur et est choisi". Si $j \leq m$, $\mathbf{P}\left\{E_j\right\} = 0$. Sinon, si $j > m$, alors $\mathbf{P}\left\{E_j\right\} = \frac{1}{n}\frac{m}{j-1}$. En effet, le meilleur club est en j-ième position avec probabilité $\frac{1}{n}$, et il est choisi si parmi les $j-1$ clubs vus précédemment, le meilleur est parmi les m premiers clubs, ce qui arrive avec probabilité $\frac{m}{j-1}$.

2. En déduire que $\lim_n \max_m \mathsf{P}_{n,m} \geq 1/e$.

☞ Les bornes se trouvent par des calculs d'intégrales :

$$\frac{m}{n}(\ln(n) - \ln(m)) \leq P(n,m) \leq \frac{m}{n}(\ln(n-1) - \ln(m-1))$$

Le maximum de la fonction $\frac{ln(x)}{x}$ est atteint pour $x = e$. Donc on prend $m = \frac{n}{e}$, et on obtient $\lim_n \max_m P(n,m) \geq 1/e$.