# Feuille de TD n° 12

*Résultats limites, Chaînes de Markov*

2 juin 2022

**Exercice 1.**

1. Rappeler l'inégalité de Bienaymé Tchebychev.
2. Soit $(Y_n)$ une suite de variables aléatoires mutuellement indépendantes, de même loi de probas, et de carré intégrable.

   On pose $S_n = \displaystyle\sum_{k=1}^{n} Y_k$.

   Montrer que $\forall\, a \in ]0, +\infty[$, $P\left(\left|\dfrac{S_n}{n} - E(Y_1)\right| \geq a\right) \leq \dfrac{V(Y_1)}{na^2}$.
3. **Application**
   (a) On prend une urne avec 2 boules rouges et 3 boules noires. On tire, sans remise, une boule dans cette urne, de façon aléatoire uniforme et indépendante.
   Pour $n \geq 1$, on pose $Y_n$ la v.a. qui donne la couleur de la boule obtenue au tirage numéro $n$. (rouge=1, noire=0)
   Donner la loi de la v.a. $Y_n$.
   Calculer $\mathbb{E}(Y_n)$.
   (b) Á partir de quel nombre de tirages peut-on garantir à plus de 95% que la proportion de boules rouges obtenues restera comprise entre $0,35$ et $0,45$ ?

1. Soit $a \in ]0, +\infty[$. Pour toute variable aléatoire $X$ admettant un moment d'ordre 2, on a :
$P(|X - E(X)| \geq a) \leq \dfrac{V(X)}{a^2}$.

2. On pose $X = \dfrac{S_n}{n}$.
Par linéarité de l'espérance et comme toutes les variables $Y_i$ ont la même espérance, on a $E(X) = E(Y_1)$.
De plus, comme les variables sont mutuellement indépendantes, on a $V(X) = \dfrac{1}{n^2}V(S_n) = \dfrac{1}{n}V(Y_1)$.
Alors, en appliquant 1. à X, on obtient le résultat souhaité.

3. $\forall i \in \mathbb{N}^*$, on considère la variable aléatoire $Y_i$ valant 1 si la $i^{\text{ième}}$ boule tirée est rouge et 0 sinon.
$Y_i$ suit une loi de Bernoulli de paramètre $p$ avec $p = \dfrac{2}{5} = 0,4$ .
Les variables $Y_i$ suivent la même loi, sont mutuellement indépendantes et admettent des moments d'ordre 2.
On a d'après le cours, $\forall i \in \mathbb{N}$, $E(Y_i) = 0,4$ et $V(Y_i) = 0,4(1 - 0,4) = 0,24$.
On pose $S_n = \displaystyle\sum_{i=1}^{n} Y_i$. $S_n$ représente le nombre de boules rouges obtenues au cours de $n$ tirages.
Alors $T_n = \dfrac{\sum_{i=1}^{n} Y_i}{n}$ représente la proportion de boules rouges obtenues au cours de $n$ tirages.
On cherche à partir de combien de tirages on a $P(0,35 \leq T_n \leq 0,45) > 0,95$.
Or $P(0,35 \leq T_n \leq 0,45) = P\left(0,35 \leq \dfrac{S_n}{n} \leq 0,45\right) = P\left(-0,05 \leq \dfrac{S_n}{n} - E(Y_1) \leq 0,05\right)$
$= P\left(\left|\dfrac{S_n}{n} - E(Y_1)\right| \leq 0,05\right) = 1 - P\left(\left|\dfrac{S_n}{n} - E(Y_1)\right| > 0,05\right)$.
On a donc $P(0,35 \leq T_n \leq 0,45) = 1 - P\left(\left|\dfrac{S_n}{n} - E(Y_1)\right| > 0,05\right)$.
Or, d'après la question précédente, $P\left(\left|\dfrac{S_n}{n} - E(Y_1)\right| \geq 0,05\right) \leq \dfrac{0,24}{n(0,05)^2}$.
Donc $P(0,35 \leq T_n \leq 0,45) \geq 1 - \dfrac{0,24}{n(0,05)^2}$.
Il suffit alors pour répondre au problème de chercher à partir de quel rang $n$, on a $1 - \dfrac{0,24}{n(0,05)^2} \geq 0,95$.
La résolution de cette inéquation donne $n \geq \dfrac{0,24}{0,05^3}$ c'est-à-dire $n \geq 1920$.

**Exercice 2.** Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires, mutuellement indépendantes, avec $X_n$ suivant une loi de Bernoulli de paramètre $p_n \in ]0, 1[$.

1. On pose $X = \frac{X_1 + \ldots + X_n}{n}$.
   Montrer que $Var(X) \leq \frac{1}{4n}$.

2. Soit $\varepsilon > 0$.
   En utilisant l'inégalité de Bienaymé-Tchebychev, montrer que l'on a
$$\mathbb{P}\left(\left|\frac{X_1+\cdots+X_n}{n} - \frac{1}{n}\sum_{i=1}^n p_i\right| < \varepsilon\right) \to_{n\to+\infty} 1$$

---

Soit $X = \dfrac{X_1+\cdots+X_n}{n}$. On a
$$\mathbb{E}(X) = \frac{1}{n}\sum_{k=1}^n p_k \text{ et } \operatorname{Var} X = \frac{1}{n^2}\sum_{k=1}^n \operatorname{Var}(X_i) = \frac{1}{n^2}\sum_{k=0}^n p_k(1-p_k).$$
Et comme $x(1-x) \leq \dfrac{1}{4}$ sur $[0,1]$, on obtient $\operatorname{Var}(X) \leq \dfrac{1}{4n}$.
L'inégalité de Bienaymé-Tchabychev montre que
$$\mathbb{P}(|X-\mathbb{E}(X)| \geq \varepsilon) \leq \frac{\operatorname{var} X}{\varepsilon^2}$$
et donc
$$\mathbb{P}\left(\left|\frac{X_1+\cdots+X_n}{n} - \sum_{k=1}^n p_k\right| < \varepsilon\right) \geq 1 - \frac{1}{4n\varepsilon^2}$$
et quand $n \to +\infty$ la quantité de droite tend bien vers 1.

**Exercice 3** (Une chaîne de Markov homogène infinie, irréductible)**.**
Une mouche se déplace dans $\mathbb{Z}^3$.
Pour tout instant $n \geq 0$, on note $S_n$ la v.a. discrète qui donne sa position.
Soient $(X_n)_{n\geq 1}$, $(Y_n)_{n\geq 0}$ et $(Z_n)_{n\geq 0}$ des suites de variables aléatoires, mutuellement indépendantes, telles que :
$\mathbb{P}(X_n = 1) = \mathbb{P}(X_n = -1) = \mathbb{P}(Y_n = 1) = \mathbb{P}(Y_n = -1) = 1/2, \mathbb{P}(Z_n = 1) = \mathbb{P}(Z_n = -1) = 1/2$, $\forall n \geq 0$.
On pose $R_n = (X_n, Y_n, Z_n)$.
On suppose que $S_{n+1} = S_n + (X_{n+1}, Y_{n+1}, Z_{n+1}) = S_n + R_{n+1}$ et que $S_0 = (0,0,0)$.
On remarquera que $\tilde{X}_n = \dfrac{X_n+1}{2}$ suit une loi de Bernoulli $\mathcal{B}(\frac{1}{2})$. De même pour $\tilde{Y}_n$ et $\tilde{Z}_n$.

1. Montrer que $(S_n)_{n\geq 0}$ est une chaîne de Markov, et qu'elle est homogène. On pourra exprimer $S_n$ en fonction des $R_i$.
2. Soit $n \geq 0$.
   Que vaut $\mathbb{P}(S_n = (0,0,0))$ quand $n$ est impair ?
3. Soit $p \geq 0$.
   Montrer que $\mathbb{P}(S_{2p} = (0,0,0)) = \left(\binom{2p}{p}\left(\dfrac{1}{2}\right)^{2p}\right)^3$.
4. En utilisant la formule de Stirling, donner un équivalent, quand $p \to +\infty$, de $\mathbb{P}(S_{2p} = (0,0,0))$.
5. On note $B$ l'événement "la mouche passe une infinité de fois par $(0,0,0)$". Montrer, en utilisant les propriétés des mesures de probabilité, que $\mathbb{P}(B) = 0$.

**Remarque :** La suite $(S_n)_n$ est une marche aléatoire dans $\mathbb{R}^3$.
Dans $\mathbb{R}^2$ (voir TD 11, Ex 1), on obtient que la probabilité de revenir une infinité de fois en $(0,0)$ est 1. Dans $\mathbb{R}^3$, revenir une infinité de fois en $(0,0,0)$ est de probabilité 0. Ce résultat est très important chez les marches aléatoires.

---

1. Soit $(s_i) \in (\mathbb{Z}^3)^n$ tel que $(S_{n-1} = s_{n-1}, \ldots, S_1 = s_1)$ est un événement de probabilité non nulle. On note $s_0 = (0,0,0)$. On note, de plus, pour tout $i \in [\![1,n]\!]$, $r_i = s_i - s_{i-1}$ et $R_i = S_i - S_{i-1}$. On remarque que la famille $(R_i)$ est une famille de variables aléatoires mutuellement indédendantes. On a :
$$\begin{aligned}\mathbb{P}(S_n = s_n | S_{n-1} = s_{n-1}, \ldots, S_1 = s_1) &= \frac{\mathbb{P}(S_n = s_n, S_{n-1} = s_{n-1}, \ldots, S_1 = s_1)}{\mathbb{P}(S_{n-1} = s_{n-1}, \ldots, S_1 = s_1)} \\ &= \frac{\mathbb{P}(R_n = r_n, R_{n-1} = r_{n-1}, \ldots, R_1 = r_1)}{\mathbb{P}(R_{n-1} = r_{n-1}, \ldots, R_1 = r_1)} \\ &= \mathbb{P}(R_n = r_n)\end{aligned}$$
   Les variables aléatoires $R_n$ et $S_{n-1} = R_{n-1} + \ldots + R_1$ sont indépendantes. On a donc :
$$\begin{aligned}\mathbb{P}(S_n = s_n | S_{n-1} = s_{n-1}) &= \frac{\mathbb{P}(R_n = r_n, S_{n-1} = s_{n-1})}{\mathbb{P}(S_{n-1} = s_{n-1})} \\ &= \mathbb{P}(R_n = r_n)\end{aligned}$$
   Donc la suite $(S_n)$ est une chaîne de Markov.
   De plus, la probabilité $\mathbb{P}(S_n = s_n | S_{n-1} = s_{n-1}) = \mathbb{P}(R_n = r_n)$ ne dépend pas de $n$ car, par hypothèse, les $R_n$ suivent toutes la même loi.
2. On remarque que $\displaystyle\sum_{k=1}^n X_k = 2\sum_{k=1}^n \tilde{X}_k - n$ où $\sum_{k=1}^n \tilde{X}_n$ suit une loi binomiale $\mathcal{B}(n, \frac{1}{2})$.
   Donc, en particulier, si $n$ est impair alors $\displaystyle\sum_{k=1}^n X_k$ est impair presque sûrement. D'où $\left(\displaystyle\sum_{k=1}^n X_k = 0\right)$ est presque impossible.

Donc

$$0 \leq \mathbb{P}(S_n = (0,0,0)) \leq P\left(\sum_{k=1}^{n} X_k = 0\right) = 0.$$

3. De même que dans la question précédente, on note $\sum_{k=1}^{n} X_k = 2\sum_{k=1}^{n} \tilde{X}_k - n$, $\sum_{k=1}^{n} Y_k = 2\sum_{k=1}^{n} \tilde{Y}_k - n$ et $\sum_{k=1}^{n} Z_k = 2\sum_{k=1}^{n} \tilde{Z}_k - n$ où $\sum_{k=1}^n \tilde{X}_k$, $\sum_{k=1}^n \tilde{Y}_k$ et $\sum_{k=1}^n \tilde{Z}_n$ suivent une loi binomiale $\mathcal{B}(n, \frac{1}{2})$. Les variables aléatoires $(X_k)$, $(Y_k)$ et $(Z_k)$ sont mutuellement indépendantes donc $\sum_{k=1}^n \tilde{X}_k$, $\sum_{k=1}^n \tilde{Y}_k$ et $\sum_{k=1}^n \tilde{Z}_k$ sont mutuellement indépendantes. D'où

$$\begin{aligned}
\mathbb{P}(S_{2p} = (0,0,0)) &= \mathbb{P}(2\textstyle\sum_{k=1}^{2p} \tilde{X}_k - 2p = 0, 2\sum_{k=1}^{2p} \tilde{Y}_k - 2p = 0, 2\sum_{k=1}^{2p} \tilde{Z}_k - 2p = 0)\\
&= \mathbb{P}(2\textstyle\sum_{k=1}^{2p} \tilde{X}_k - 2p = 0)\mathbb{P}(2\sum_{k=1}^{2p} \tilde{Y}_k - 2p = 0)\mathbb{P}(2\sum_{k=1}^{2p} \tilde{Z}_k - 2p = 0)\\
&= \mathbb{P}(\textstyle\sum_{k=1}^{n} \tilde{X}_k = p)\mathbb{P}(\sum_{k=1}^{n} \tilde{Y}_k = p)\mathbb{P}(\sum_{k=1}^{n} \tilde{Z}_n = p)\\
&= \left(\binom{2p}{p}\frac{1}{2^{2p}}\right)^3
\end{aligned}$$

4. D'après la formule de Stirling on a :

$$\begin{aligned}
\binom{2p}{p} &= \frac{(2p)!}{p!p!}\\
&\underset{p\to+\infty}{\sim} \frac{\sqrt{4\pi p}\left(\frac{2p}{e}\right)^{2p}}{((\sqrt{2\pi p}\left(\frac{p}{e}\right)^p)^2}\\
&= \frac{2^{2p}}{\sqrt{\pi p}}
\end{aligned}$$

On a donc :

$$\mathbb{P}(S_{2p} = (0,0,0)) \underset{p\to+\infty}{\sim} \frac{1}{(\pi p)^{3/2}}.$$

5. L'événement $B$ : « la mouche passe une infinité de fois par $(0,0,0)$ » est égal à $\limsup_k (S_k = (0,0,0)) = \bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} (S_k = (0,0,0))$.

Or d'après la question précédente $\mathbb{P}(S_{2p} = (0,0,0)) \underset{n\to+\infty}{\sim} \frac{1}{(\pi p)^{3/2}}$. Donc la série est convergente.

Ainsi, avec le lemme de Borel-Cantelli, on obtient $(\bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} (S_k = (0,0,0))) = 0$.

Donc presque sûrement la mouche ne revient pas à sa position initiale.

## *Pour aller plus loin . . .*

**Exercice 4.** Soit $n \geq 2$. $n$ personnes jouent au jeu suivant :
Chacun mise 1 euro, et écrit Pile ou Face sur du papier sans le montrer aux autres. Puis, un responsable lance une pièce équilibrée.
Si c'est Pile, tous ceux qui ont écrit Pile gagnent. Si c'est Face, tous ceux qui ont écrit Face gagnent.
Les gagnants se partagent les $n$ euros (sous forme de fraction).
Si personne ne gagne, le responsable prend les $n$ euros.

1. Soit $k \in \{1, ..., n\}$.
   On note $X_k$ la v.a. qui indique la somme que reçoit le joueur numéro $k$ à la fin du jeu.
   Calculer l'espérance de $X_k$.
2. Dans cette question, on suppose qu'une $n+1$-ème personne arrive avant qu'on ne lance la pièce.
   On demande au joueur numéro $n$ s'il accepte que cette personne participe aussi.
   Le joueur $n$ veut maximiser l'espérance de ses gains. Doit-il accepter le nouveau joueur ?
   Le responsable veut lui aussi maximiser l'espérance de ses gains. Doit-il accepter le nouveau joueur ?
3. Soit $f$ une fonction convexe sur un intervalle $I$ de $\mathbb{R}$.
   (a) Soit $I$ un intervalle, et $f : I \to \mathbb{R}$ une fonction convexe. On rappelle qu'elle vérifie
   $$\forall x, y \in I,\ \forall t \in [0,1],\ f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$
   Quel raisonnement faut-il utiliser pour montrer le résultat suivant ?
   $\forall x_1, ..., x_n \in I$, $\forall t_1, \ldots, t_n \in [0,1]$ tels que $\sum_{i=1}^n t_i = 1$, on a :
   $$f(t_1x_1 + ... + t_nx_n) \leq t_1f(x_1) + ... + t_nf(x_n).$$
   (b) Soit $X$ une v.a. discrète telle que $X(\Omega)$ est fini et inclus dans $I$.
   Montrer que $f(\mathbb{E}(X)) \leq \mathbb{E}(f(X))$.
4. On se place à nouveau dans un jeu à $n$ joueurs.
   (a) Vérifier que $\mathbb{E}(X_k^2) = \dfrac{n^2}{2}\sum_{r=0}^{n-1} \dfrac{1}{(1+r)^2}\binom{n-1}{r}\left(\dfrac{1}{2}\right)^{n-1}$.
   (b) En déduire que
   $$\mathbb{E}(X_k^2) \geq \frac{2n^2}{(n+1)^2}.$$

On pourra utiliser la fonction $g : x \mapsto \dfrac{1}{(1+x)^2}$.

---

1. Posons $S_n = X_1 + \cdots + X_n$ la variable aléatoire à valeur dans $\{0, n\}$.
   $(S_n = 0)$ ssi tous les joueurs perdent. Les variables $X_i$ sont indépendantes de loi de Bernoulli de paramètre $1/2$. $\mathbb{P}(S_n = 0) = \left(\frac{1}{2}\right)^n$. On en déduit
   $$\mathbb{E}(S_n) = 0 \times \left(\frac{1}{2}\right)^n + n \times \left[1 - \left(\frac{1}{2}\right)^n\right] = n \times \left[1 - \left(\frac{1}{2}\right)^n\right]$$
   On en déduit que $\mathbb{E}(X_k) = 1 - \left(\frac{1}{2}\right)^n$.
2. Le gain moyen augmente avec le nombre de joueurs, donc le joueur à intérêt d'accepter un nouveau joueur.
   Par contre, le gain moyen du responsable est $\frac{n}{2^n}$ et $x \mapsto \frac{x}{2^x}$ est décroissante pour $x \geq 2$. Donc le responsable a intérêt à refuser le nouvel arrivant.
3. Soit $f : I \to \mathbb{R}$ une fonction convexe.
   (a) Une récurrence avec utilisation du barycentre partiel donne le résultat (c'est du cours).
   (b) Une application directe du théorème de transfert permet de conclure.
4. Si $G_k$ est l'évènement le joueur $k$ gagne et $Y$ le nombre de gagnants différents de $k$, on a
   $\mathbb{P}(X_k = n/r) = \mathbb{P}(Y = r-1 \cap G_k) \overset{\textit{indépendants}}{=} \mathbb{P}(Y = r-1)\mathbb{P}(G_k) = \frac{1}{2} \times \binom{n-1}{r-1}\left(\frac{1}{2}\right)^{n-1}$, car $Y$ suit une loi binomiale de paramètre $n-1$ et $1/2$.
   $$\mathbb{E}(X_k^2) = \sum_{r=1}^{n} \frac{n^2}{r^2}\binom{n-1}{r-1}\left(\frac{1}{2}\right)^n = \frac{n^2}{2}\sum_{r=0}^{n-1}\frac{1}{(1+r)^2}\binom{n-1}{r}\left(\frac{1}{2}\right)^{n-1}$$
   La fonction $g : x \mapsto \frac{1}{(1+x)^2}$ est convexe et soit $U$ une loi binomiale de paramètres $n-1$ et $\frac{1}{2}$. La question précédente donne
   $$\mathbb{E}(X_k^2) = \frac{n^2}{2}\sum_{r=0}^{n-1}\frac{1}{(1+r)^2}\binom{n-1}{r}\left(\frac{1}{2}\right)^{n-1} = \frac{n^2}{2} \times \mathbb{E}(g(U)) \geq \frac{n^2}{2} \times g(\mathbb{E}(U)) = \frac{n^2}{2} \times \frac{1}{\left((n-1) \times \frac{1}{2} + 1\right)^2} = \frac{2n^2}{(n+1)^2}.$$