Séance I : Espace de probabilité

Exercice I.1.
On suppose que l'on a autant de chance d'avoir une fille ou un garçon à la naissance. Votre voisin de palier vous dit qu'il a deux enfants.

1. Quelle est la probabilité qu'il ait au moins un garçon ?
2. Quelle est la probabilité qu'il ait un garçon, sachant que l'aînée est une fille ?
3. Quelle est la probabilité qu'il ait un garçon, sachant qu'il a au moins une fille ?
4. Vous téléphonez à votre voisin. Une fille décroche le téléphone. Vous savez que dans les familles avec un garçon et une fille, la fille décroche le téléphone avec probabilité p, quelle est la probabilité que votre voisin ait un garçon ?
5. Vous sonnez à la porte de votre voisin. Une fille ouvre la porte. Sachant que l'aîné(e) ouvre la porte avec probabilité p, et ce indépendamment de la répartition de la famille, quelle est la probabilité que votre voisin ait un garçon ?

△

Exercice I.2.
Les laboratoires pharmaceutiques indiquent pour chaque test sa sensibilité α, qui est la probabilité que le test soit positif si le sujet est malade, et sa spécificité β, qui est la probabilité que le test soit négatif si le sujet est sain. Sachant qu'en moyenne il y a un malade sur 1000 personnes, calculer la probabilité pour que vous soyez un sujet sain alors que votre test est positif, avec $\alpha = 98\%$ et $\beta = 97\%$. Calculer la probabilité d'être malade alors que le test est négatif. Commentaire.

△

Exercice I.3.
Le joueur A possède deux dés à six faces, et le joueur B posséde un dé à douze faces. Le joueur qui fait le plus grand score remporte la mise (match nul si égalité). Le jeu est-il équilibré ? On calculera la probabilité que A gagne et la probabilité d'avoir un match nul.

△

Exercice I.4.
On considère trois cartes : une avec les deux faces rouges, une avec les deux faces blanches, et une avec une face rouge et une face blanche. On tire une carte au hasard. On expose une face au hasard. Elle est rouge. Parieriez-vous que la face cachée est blanche ? Pour vous aider dans votre choix :

1. Déterminer l'espace de probabilité.
2. Calculer la probabilité que la face cachée est blanche sachant que la face visible est rouge.

△

Exercice I.5.
La **formule du crible**. Soit $A_1, \cdots, A_n$ des évènements.

1. Montrer que $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$.
2. Montrer la formule du crible par récurrence.
$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{p=1}^{n}(-1)^{p+1} \sum_{1\leq i_1<\cdots<i_p\leq n} \mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_p}). \tag{I.1}$$
3. Montrer, par récurrence sur n, que pour $1 \leq m \leq n$,
$$\sum_{p=1}^{m}(-1)^{p+1} \sum_{1\leq i_1<\cdots<i_p\leq n} \mathbb{P}(A_{i_1} \cap \cdots \cap A_{j_p})$$
est une majoration (resp. minoration) de $\mathbb{P}(\bigcup_{i=1}^{n} A_i)$ lorsque m est impair (resp. pair).

$\triangle$

Exercice I.6.
On utilise dans cet exercice la formule du crible (I.1) (cf exercice I.5).

1. Pour fêter leur réussite à un concours, n étudiants se donnent rendez-vous dans un châlet. En entrant chaque personne dépose sa veste dans un vestiaire. Au petit matin, quand les esprits ne sont plus clairs, chacun prend au hasard une veste. Quelle est la probabilité pour qu'une personne au moins ait sa propre veste ?
2. En déduire le nombre de permutations de $\{1, \ldots, n\}$ sans point fixe (problème formulé par P. R. de Montmort en 1708)[1]
3. En s'inspirant de la question 1, calculer la probabilité $\pi_n(k)$ pour que k personnes exactement aient leur propre veste.
4. Calculer la limite $\pi(k)$ de $\pi_n(k)$ quand n tend vers l'infini. Vérifier que la famille $(\pi(k), k \in \mathbb{N})$ détermine une probabilité sur $\mathbb{N}$. Il s'agit en fait de la loi de Poisson.

$\triangle$

Exercice I.7.
On utilise dans cet exercice la formule du crible (I.1) (cf exercice I.5). Soit $1 \leq k \leq n$.

1. Calculer à l'aide de la formule du crible le nombre de surjections de $\{1, \cdots, n\}$ dans $\{1, \cdots, k\}$.
2. En déduire $\left\{ n \atop k \right\}$, le nombre de partitions d'un ensemble à n éléments en k sous-ensembles non vides. Les nombres $\left\{ n \atop k \right\}$ sont appelés les nombres de Stirling de deuxième espèce.
3. Montrer que
$$\left\{ n \atop k \right\} = \left\{ n-1 \atop k-1 \right\} + k \left\{ n-1 \atop k \right\}, \quad \left\{ n \atop 1 \right\} = 1, \quad \left\{ n \atop k \right\} = 0 \text{ si } k > n.$$

$\triangle$

[1] Voir L. Takacs (The problem of coincidences, *Arch. Hist. Exact Sci.* 21 :3 (1980), 229-244) pour une étude historique du problème des coïncidences vu par les probabilistes.

Séance II : V. a. discrètes

Exercice II.1.
Soit X une variable aléatoire de loi de Poisson de paramètre $\theta > 0$ (i.e. $\mathbb{P}(X = k) = \mathrm{e}^{-\theta} \dfrac{\theta^k}{k!}$, $k \in \mathbb{N}$).

1. Vérifier que $\dfrac{1}{1+X}$ est une variable aléatoire intégrable. Calculer $\mathbb{E}\left[\dfrac{1}{1+X}\right]$.
2. Calculer $\mathbb{E}\left[\dfrac{1}{(1+X)(2+X)}\right]$ et en déduire $\mathbb{E}\left[\dfrac{1}{2+X}\right]$.

$\triangle$

Exercice II.2.
Soit X_1, X_2 des variables aléatoires indépendantes de loi de Poisson de paramètres respectifs $\theta_1 > 0$ et $\theta_2 > 0$.

1. Calculer la loi de $X_1 + X_2$.
2. Calculer la loi de X_1 sachant $X_1 + X_2$. Reconnaître cette loi.
3. Calculer $\mathbb{E}[X_1|X_1 + X_2]$.

$\triangle$

Exercice II.3.
Un gardien de nuit doit ouvrir une porte dans le noir, avec n clefs dont une seule est la bonne.

1. Donner la loi de probabilité du nombre X d'essais nécessaires s'il essaie les clefs une à une sans utiliser deux fois la même. Calculer l'espérance et la variance de X.
2. Lorsque le gardien est ivre, il mélange toutes les clefs à chaque tentative. Identifier la loi de X. Rappeler l'espérance et la variance de X.
3. Le gardien est ivre un jour sur trois. Sachant qu'un jour n tentatives ont été nécessaires pour ouvrir la porte, quelle est la probabilité que le gardien ait été ivre ce jour là ? Calculer la limite quand n tend vers l'infini.

$\triangle$

Exercice II.4.
On désire modéliser la loi du temps d'attente d'une panne de machine à l'aide d'une lois sans mémoire : la probabilité pour que la machine tombe en panne après la date $k + n$ sachant qu'elle fonctionne à l'instant n est indépendante de n. Plus précisément, on dit qu'une variable aléatoire X à valeurs dans $\mathbb{N}$ est de loi sans mémoire si $\mathbb{P}(X > k + n|X > n)$ est indépendant de $n \in \mathbb{N}$ pour tout $k \in \mathbb{N}$.

1. Montrer que la loi géométrique de paramètre p est sans mémoire.
2. Caractériser toutes les lois sans mémoire des variables aléatoires X à valeurs dans $\mathbb{N}^*$. On pourra calculer $\mathbb{P}(X > 1 + n)$ en fonction de $\mathbb{P}(X > 1)$.

3. Caractériser toutes les lois sans mémoire des variables aléatoires X à valeurs dans $\mathbb{N}$.

$\triangle$

Exercice II.5.
Soit T_1 et T_2 deux variables aléatoires indépendantes de loi géométrique de paramètre p_1 et p_2.

1. Calculer et reconnaître la loi de $\min(T_1, T_2)$.
2. Calculer la loi jointe de $\min(T_1, T_2)$ et $T_1 - T_2$.
3. En déduire que $\min(T_1, T_2)$ est indépendant de $\mathbf{1}_{\{T_1 \leq T_2\}}$. Quelle est la loi de $\mathbf{1}_{\{T_1 \leq T_2\}}$?
4. Déduire également de la question 2. que $R = \max(T_1, T_2) - \min(T_1, T_2)$ est indépendant de $\min(T_1, T_2)$.
5. Calculer la loi de R conditionnellement à $\{R \neq 0\}$. Reconnaître cette loi quand $p_1 = p_2$.

$\triangle$

Exercice II.6.
On considère une urne contenant $r \geq 1$ boules rouges et $b \geq 1$ boules bleues. On tire au hasard les boules **sans remise**.

1. Combien existe-t-il de tirages complets possibles ?
2. On note X_n le nombre de boules rouges obtenues alors que l'on a tiré n boules. Calculer la loi de X_n.
3. Reconnaître la loi limite de X_n quand $r \to \infty$ et $r/(r+b) \to p \in]0, 1[$.
4. On note $Y_k = 1$ si la $\text{k}^{\text{ième}}$ boule est rouge et $Y_k = 0$ sinon. Quelle est la loi de $(Y_1, \ldots, Y_{r+b})$?
5. En déduire que $Y_1, \ldots, Y_{r+b}$ ont même loi. Calculer la loi de Y_1.
6. Exprimer X_n en fonction de $Y_1, \ldots, Y_n$. Calculer $\mathbb{E}[X_n]$ et $\text{Var}(X_n)$.
7. On note S_n le nombre de boules rouges lors d'un tirage aléatoire de n boules de l'urne **avec remise**. Quelle est la loi de S_n ? Comparer avec la question 3.
8. Calculer $\mathbb{E}[S_n]$ et $\text{Var}(S_n)$. Comparer avec la question 6.

$\triangle$

Exercice II.7.
On considère une urne contenant $r \geq 1$ boules rouges et $b \geq 1$ boules bleues. On tire au hasard les boules **sans remise**.

1. On note T_1 le nombre de boules qu'il faut tirer pour obtenir une boule rouge. Calculer la loi de T_1.
2. Quelle est la loi limite quand $r \to \infty$ et $r/(r+b) \to p \in]0, 1[$. Comparer la loi limite avec le premier temps d'obtention d'une boule rouge dans un tirage **avec remise**.
3. On note $Z_1 = T_1 - 1$ et pour $i \in \{2, \ldots, r\}$, Z_k le nombre de boules bleues obtenues entre la $\text{(k-1)}^{\text{ième}}$ boule rouge et la $\text{k}^{\text{ième}}$ boule rouge. Enfin on note Z_{r+1} le nombre de boules bleues obtenues après la dernière boule rouge. Calculer la loi de $(Z_1, \ldots, Z_{r+1})$.
4. En déduire que $Z_1, \ldots, Z_{r+1}$ ont même loi. Calculer $Z_1 + \cdots + Z_{r+1}$, puis calculer $\mathbb{E}[T_1]$. Cette méthode ne permet pas néanmoins de calculer $\text{Var}(T_1)$. On admet que $\text{Var}(T_1) = \dfrac{rb(b+r+1)}{(r+1)^2(r+2)}$.
5. Que se passe-t-il pour $r \to \infty$ et $r/(r+b) \to p \in]0, 1[$?

$\triangle$

Séance III : V.a. réelles

Exercice III.1.
Soit X une variable aléatoire réelle de densité $f(x) = x\,\mathrm{e}^{-x^2/2}\,\mathbf{1}_{\{x>0\}}$.

1. Vérifier que f est une densité de probabilité.
2. Montrer que $Y = X^2$ est une variable aléatoire continue, dont on précisera la densité. Reconnaître la loi de Y.
3. Calculer l'espérance et la variance de Y

$\triangle$

Exercice III.2.
Soit Z une variable aléatoire de Cauchy (de densité $\dfrac{1}{\pi}\dfrac{1}{1+x^2}$).

1. Calculer et reconnaître la loi de $1/Z$.
2. Pour quelles valeurs de $\alpha \in \mathbb{R}$ est-ce que $|Z|^\alpha$ est intégrable ?

$\triangle$

Exercice III.3.
Soit X une variable aléatoire positive de densité f. Montrer soigneusement (en distinguant $\mathbb{E}[X^r] = \infty$ et $\mathbb{E}[X^r] < \infty$) que

$$\mathbb{E}[X^r] = \int_0^\infty r x^{r-1} \mathbb{P}(X > x) dx \quad \text{si} \quad r > 0.$$

Donner une formule analogue pour $r < 0$. $\triangle$

Exercice III.4.
Votre ami choisit deux nombres positifs sans vous faire part de la manière dont il les choisit. Après avoir lancé une pièce équilibrée, il vous donne le plus petit s'il a obtenu face et le plus grand sinon. Vous devez parier s'il vous a donné le plus petit ou le plus grand. Votre objectif est de maximiser votre probabilité de gagner votre pari.

1. Vous lancez une pièce équilibrée ou non. Si vous obtenez face, vous pariez qu'il vous a donné le plus petit, sinon vous pariez qu'il vous a donné le plus grand. Quelle est la probabilité de gagner votre pari ?
2. Vous simulez une variable aléatoire positive continue Z ayant pour support $\mathbb{R}^+$. Si le nombre donné par votre ami est plus petit que Z, alors vous pariez qu'il vous a donné le plus petit, sinon vous pariez qu'il vous a donné le plus grand. Quelle est la probabilité de gagner votre pari ?
3. On suppose que les deux nombres de votre ami, ont été obtenus par simulation suivant une loi (continue de densité strictement positive sur $]0, \infty[$) donnée et connue de vous. Déterminer votre stratégie optimale (i.e. la loi de Z que l'on ne suppose plus continue). Quelle est alors la probabilité de gagner votre pari ?

$\triangle$

Exercice III.5.
Soient X et Y deux variables aléatoires dont les fonctions de répartitions, notées respectivement F et G, sont continues et strictement croissantes.

1. Calculer la loi de $G(Y)$.
2. Calculer et reconnaître la loi de $F^{-1}(G(Y))$.
3. En déduire une méthode pour simuler la variable aléatoire X (il s'agit de la méthode d'inversion de la fonction de répartition).

$\triangle$

Exercice III.6.
On modélise le temps d'attente de pannes de machines par des variables aléatoires sans mémoire i.e.

$$\mathbb{P}(X > t + s | X > t) = \mathbb{P}(X > s) \quad t \geq 0, s \geq 0.$$

1. Montrer que les lois exponentielles sont sans mémoire.
2. Déterminer les variables aléatoires positives dont la loi admet une densité, qui sont sans mémoire. On montrera que la fonction $\bar{F}(t) = \mathbb{P}(X > t)$ satisfait une équation différentielle que l'on justifiera et que l'on résoudra.

$\triangle$

Exercice III.7.
Les lois exponentielles apparaissent comme des lois limites pour des lois géométriques changées d'échelle. On peut mettre en évidence d'autres propriétés.

1. Montrer que si T est une variable aléatoire exponentielle de paramètre λ, alors $\lfloor Tm \rfloor + 1$, où $\lfloor x \rfloor$ représente la partie entière de x, et $m > 0$, est une variable aléatoire géométrique dont on déterminera le coefficient.
2. Soit T une variable aléatoire positive telle que $\lfloor T2^n \rfloor + 1$ est une variable aléatoire géométrique pour tout $n \in \mathbb{N}^*$. On note p_n son paramètre. Établir une relation de récurrence entre $q_n = 1 - p_n$ et $q_{n+1} = 1 - p_{n+1}$. Montrer que

   $$\{\lfloor T2^n \rfloor + 1 \geq \lfloor 2^n x \rfloor + 2\} \subset \{T \geq x\} \subset \{\lfloor T2^n \rfloor + 1 \geq \lfloor 2^n x \rfloor\}.$$

   En conclure que T suit une loi exponentielle. On déterminera son paramètre en fonction de q_0.
3. Soit T une variable aléatoire positive telle que : il existe une suite $(m_n, n \geq 1)$ croissante avec $m_0 > 0$ et $\lim_{n\to\infty} m_n = \infty$, et pour tout $n \geq 1$, $\lfloor Tm_n \rfloor + 1$ est une variable aléatoire géométrique. Montrer que T suit une loi exponentielle. On exprimera son paramètre à l'aide de celui de $\lfloor Tm_0 \rfloor$. (On montrera d'abord que $\mathbb{P}(T > x) = \lim_{n\to\infty} \mathbb{P}(\lfloor Tm_n \rfloor > \lfloor m_n x \rfloor)$ pour $x \geq 0$.)

$\triangle$

Séance IV : V.a. vectorielles

Exercice IV.1.
Soit Y une variable aléatoire de loi exponentielle $\lambda > 0$ et ε une variable aléatoire discrète indépendante de Y et telle que $\mathbb{P}(\varepsilon = 1) = \mathbb{P}(\varepsilon = -1) = 1/2$. Montrer que la variable aléatoire $Z = \varepsilon Y$ est à densité et la calculer. Cette loi est appelée loi exponentielle symétrique. △

Exercice IV.2.
Soit $(U_n, n \in \mathbb{N}^*)$ une suite de variables aléatoires indépendantes de loi uniforme sur $[0, 1]$. Soit $\theta > 0$.

1. Donner la loi de $X_k = -\log(U_k)/\theta$.
2. Donner la loi de $\sum_{k=1}^{n} X_k$.
3. Calculer la loi de N défini par $N = \inf \left\{n; \prod_{k=1}^{n+1} U_k < \mathrm{e}^{-\theta}\right\}$.
4. En déduire une méthode pour simuler des variables aléatoires de Poisson.

△

Exercice IV.3.
On considère un gâteau circulaire avec une cerise sur le bord. On découpe le gâteau en deux parts en coupant suivant deux rayons choisis au hasard.

1. Avec quelle probabilité la part contenant la cerise est-elle plus petite que la part ne contenant pas la cerise ?
2. Quelle est la longueur angulaire moyenne de la part contenant la cerise ?

△

Exercice IV.4.
Soit X et Y deux variables aléatoires indépendantes de loi respective $\Gamma(\lambda, a)$ et $\Gamma(\lambda, b)$ avec $a, b, \lambda \in]0, \infty[$.

1. Calculer la loi du couple $(X + Y, \dfrac{X}{X + Y})$.
2. Montrer que $X + Y$ et $\dfrac{X}{X + Y}$ sont indépendantes et identifier leur loi.

△

Exercice IV.5.
On désire déterminer la distribution des vitesses des molécules d'un gaz monoatomique parfait à l'équilibre (loi de Maxwell (1859)).

1. Soit (X, Y, Z), un vecteur aléatoire continu à valeurs dans $\mathbb{R}^3$ dont la loi est invariante par rotation autour de l'origine et dont les composantes X, Y, Z sont indépendantes.

Caractériser[2] les lois marginales de X, Y et Z dans le cas où les densités des lois marginales sont des fonctions $C^1(\mathbb{R}^3, \mathbb{R})$.

2. On représente la vitesse d'une molécule d'un gaz monoatomique parfait à l'équilibre dans un repère orthonormal par un vecteur aléatoire $V = (V_1, V_2, V_3)$. Le choix du repère étant arbitraire, il est naturel de supposer que la loi de V est invariante par rotation. Il est de plus naturel de supposer que les coordonnées de V sont indépendantes. Si on suppose de plus que la loi de V possède une densité dérivable, on en déduit que le vecteur V vérifie les propriétés de la question 1). Déterminer la densité de probabilité de la vitesse d'une molécule, sachant que l'énergie cinétique moyenne d'un atome du gaz de masse m est $\frac{3}{2}kT$ où k est la constante de Boltzmann et T la température du gaz. (Pour des molécules à plusieurs atomes, l'énergie cinétique moyenne tient compte d'effets complexes comme la rotation, les oscillations... La loi de Maxwell n'est plus vérifiée dans ces cas.)
3. Montrer que si X et Y sont deux variables aléatoires indépendantes de loi respective $\Gamma(\lambda, a)$ et $\Gamma(\lambda, b)$, alors la loi de $X + Y$ est une loi gamma dont on précisera les paramètres.
4. Calculer la loi de V_1^2. En déduire la loi de $|V|^2$ et la loi de $|V| = \sqrt{V_1^2 + V_2^2 + V_3^2}$ dite loi de Maxwell.

△

Exercice IV.6.
Soit $(X_n, n \geq 1)$ une suite de variables aléatoires indépendantes de loi exponentielle de paramètre rescpectif $\lambda_n > 0$. Montrer que les trois conditions suivantes sont équivalentes.

1. $\sum_{n\geq 1} \lambda_n^{-1} < \infty$.
2. $\mathbb{E}\left[\sum_{n\geq 1} X_n\right] < \infty$.
3. $\mathbb{P}\left(\sum_{n\geq 1} X_n < \infty\right) > 0$.

Pour $3 \Rightarrow 1$, on pourra considérer $\mathbb{E}[\mathrm{e}^{-\sum_{n\geq 1} X_n}]$.

△

Exercice IV.7.
Soit $n \geq 2$ et $X_1, \ldots, X_n$ une suite de n variables aléatoires indépendantes de loi uniforme sur $[0, 1]$. On note $Y = \min_{1\leq i\leq n} X_i$ et $Z = \max_{1\leq i\leq n} X_i$.

1. Calculer la loi de (Y, Z).
2. En déduire la loi de Y et la loi de Z. Reconnaître ces deux lois.
3. Calculer $\mathbb{E}[Y|Z]$.
4. Calculer $\mathbb{E}[g(Y/Z)|Z]$, pour une fonction g mesurable bornée. En déduire puis reconnaître la loi de Y/Z conditionnellement à Z. Retrouver le résultat de la question 3.
5. Montrer que $(1 - Z, 1 - Y)$ a même loi que (Y, Z).
6. En déduire que $(1 - Z)/(1 - Y)$ est indépendant de Y.

△

[2] En fait, on peut, en utilisant les fonctions caractéristiques, caractériser toutes les lois des vecteurs aléatoires qui sont invariantes par rotation autour de l'origine et dont les coordonées sont indépendantes, voir l'exercice VI.8. Hormis la variable nulle, on ne trouve pas d'autres lois que celles obtenues sous les hypothèses de cet exercice.

Séance V : Convergence

Exercice V.1.
Théorème de Weierstrass (1885) : "Toute fonction continue sur un intervalle fermé borné est limite uniforme d'une suite de polynômes".

Cet exercice s'inspire de la démonstration de Bernstein du théorème de Weierstrass. Soit $(X_n, n \geq 1)$ une suite de variables aléatoires indépendantes de loi de Bernoulli de paramètre $x \in [0,1]$. Pour $n \geq 1$, on considère la moyenne empirique $\bar{X}_n = \frac{1}{n}\sum_{k=1}^n X_k$. Soit $h : [0,1] \to \mathbb{R}$ une fonction continue. Soit $\delta > 0$. On pose $\Delta_n = \{\left|\bar{X}_n - x\right| > \delta\}$.

1. Montrer que $\mathbb{P}(\Delta_n) \leq \delta^{-2}\mathbb{E}[(\bar{X}_n - x)^2]$. Majorer $\mathbb{P}(\Delta_n)$ indépendamment de $x \in [0,1]$.
2. Déterminer $\lim_{n\to\infty} \sup_{x\in[0,1]} \left|h(x) - \mathbb{E}[h(\bar{X}_n)]\right|$, en écrivant

$$\left|h(x) - h(\bar{X}_n)\right| = \left|h(x) - h(\bar{X}_n)\right| \mathbf{1}_{\Delta_n} + \left|h(x) - h(\bar{X}_n)\right| \mathbf{1}_{\Delta_n^c}.$$

3. Quelle est la loi de $n\bar{X}_n$?
4. En déduire que

$$\lim_{n\to\infty} \sup_{x\in[0,1]} \left|h(x) - \sum_{k=0}^{n} \binom{n}{k} h(k/n) x^k (1-x)^{n-k}\right| = 0.$$

5. Soit $f : \mathbb{R}^+ \to \mathbb{R}$ continue bornée. Montrer, en s'inspirant des questions précédentes, que pour tout $x \in \mathbb{R}^+$,

$$\lim_{n\to\infty} \left|f(x) - \sum_{k=0}^{\infty} \mathrm{e}^{-nx} \frac{(nx)^k}{k!} f(k/n)\right| = 0.$$

   Si l'on suppose f uniformément continue, la convergence ci-dessus est-elle uniforme en x ? (Prendre par exemple $f(x) = \cos(x)$ pour $x_n = \pi n$.)

△

Exercice V.2.
En considérant une suite de variables aléatoires indépendantes de loi uniforme sur $[0,1]$, calculer à l'aide de la loi des grands nombres

$$\lim_{n\to\infty} \int_{[0,1]^n} f(\frac{x_1 + \cdots + x_n}{n}) \, dx_1 \cdots dx_n,$$

où f est une application continue bornée de $\mathbb{R}$ dans $\mathbb{R}$. △

Exercice V.3.
Soit X une variable aléatoire à valeurs dans $\mathbb{N}$.

1. Montrer que $\sum_{k=1}^{+\infty} \mathbb{P}(X \geq k) = \mathbb{E}[X]$.

On suppose que X est intégrable. Soit $(X_n, n \geq 1)$ une suite de variables aléatoires de même loi que X.

2. Soit $m \in \mathbb{N}^*$. Montrer que la variable aléatoire $Y_m = \sum_{n=1}^{+\infty} \mathbf{1}_{\{\frac{X_n}{n} \geq \frac{1}{m}\}}$ est p.s. finie.
3. En déduire que $\frac{X_n}{n}$ tend p.s. vers 0 quand n tend vers l'infini.
4. Montrer que $\frac{1}{n} \max(X_1, \ldots, X_n)$ tend p.s. vers 0 quand n tend vers l'infini.
5. On suppose que X est une variable aléatoire intégrable à valeurs dans $\mathbb{R}$ et non plus dans $\mathbb{N}$. Montrer que $\frac{1}{n} \max(X_1, \ldots, X_n)$ tend p.s. vers 0 quand n tend vers l'infini.

△

Exercice V.4.
Majoration du théorème des grandes déviations[3].

Soit $(X_n, n \geq 1)$ une suite de variables aléatoires indépendantes et de même loi que X. On suppose que X est une variable aléatoire bornée non constante de moyenne $\mu = \mathbb{E}[X]$. Le but de cet exercice est de trouver une majoration exponentielle de l'évènement rare $\{|\bar{X}_n - \mu| \geq \varepsilon)\}$ avec $\varepsilon > 0$, où $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$. Cette majoration est un exemple particulier des résultats de la théorie des grandes déviations.

On note Φ la transformée de Laplace de X définie par $\Phi(\theta) = \mathbb{E}[\mathrm{e}^{\theta X}]$ pour $\theta \in \mathbb{R}$.

1. Montrer que Φ est de classe C^∞. Calculer ses dérivées.
2. Montrer que $(\Phi')^2 \leq \Phi\Phi''$. En déduire que la fonction $\lambda = \log(\Phi)$ est convexe. Vérifier que $\lambda(0) = 0$.

On considère la transformée de Legendre de λ, I, définie pour $x \in \mathbb{R}$ par

$$I(x) = \sup_{\theta \in \mathbb{R}} \{\theta x - \lambda(\theta)\}.$$

3. Montrer que I est convexe, positive et nulle en μ.
4. Montrer que $\mathbb{P}(\bar{X}_n \geq \mu + \varepsilon) \leq \mathrm{e}^{-\theta(\mu+\varepsilon)+n\lambda(\theta/n)}$ pour $\theta \geq 0$. En déduire que
$$\mathbb{P}(\bar{X}_n \geq \mu + \varepsilon) \leq \mathrm{e}^{-nI(\mu+\varepsilon)} = \mathrm{e}^{-n \inf\{I(x); x \geq \mu+\varepsilon\}}.$$
5. Montrer que
$$\mathbb{P}(|\bar{X}_n - \mu| \geq \varepsilon) \leq 2\,\mathrm{e}^{-n \inf\{I(x); |x-\mu| \geq \varepsilon\}}. \tag{V.2}$$
6. Calculer explicitement la majoration exponentielle quand X suit la loi de Bernoulli de paramètre $p \in]0, 1[$.

La majoration exponentielle (V.2) est en fait vraie dans un cadre plus général (i.e. pour des variables aléatoires non bornées).

△

[3] Les théorèmes des grandes déviations étudient les équivalents logarithmiques des probabilités d'évènements rares. L'exemple typique d'évènement considéré est $\{|\bar{X}_n - \mathbb{E}[X_1]| > \varepsilon\}$, dont l'étude est due à Cramèr (1938). D'autre part, la théorie des grandes déviations est un sujet d'étude qui connaît un essor important depuis les années 1980.

Séance VI : Convergence en loi

Exercice VI.1.
Soit $(T_n, n \geq n_0)$ une suite de variables aléatoires de loi géométrique de paramètre $p_n = \dfrac{\theta}{n}$ avec $n_0 > \theta > 0$. Montrer que la suite $\left(\dfrac{T_n}{n}, n \geq n_0\right)$ converge en loi et déterminer sa limite.

△

Exercice VI.2.
Soit $(X_n, n \in \mathbb{N}^*)$ une suite de variables aléatoires de loi exponentielle de paramètre λ_n. Étudier la convergence en loi dans les trois cas suivants :

1. $\lim_{n\to\infty} \lambda_n = \lambda \in]0, \infty[$,
2. $\lim_{n\to\infty} \lambda_n = +\infty$,
3. $\lim_{n\to\infty} \lambda_n = 0$.

△

Exercice VI.3.
Soit $(U_n, n \geq 1)$ une suite de variables aléatoires indépendantes de loi uniforme sur l'intervalle $[0, \theta]$, où $\theta > 0$. On pose pour $n \geq 1$, $X_n = \max_{1\leq i\leq n} U_i$.

1. Montrer que $(X_n, n \geq 1)$ converge p.s. et déterminer sa limite. On pourra calculer $\mathbb{P}(|X_n - \theta| > \varepsilon)$ pour $\varepsilon > 0$.
2. Étudier la convergence en loi de la suite $(n(\theta - X_n), n \geq 1)$.

△

Exercice VI.4.
Soit $(U_i, i \in \mathbb{N}^*)$ une suite de variables aléatoires indépendantes de loi uniforme sur $[0, 1]$. Soit $\alpha > 0$.

1. Pour $n \in \mathbb{N}^*$, on pose $X_n = (U_1 \cdots U_n)^{\alpha/n}$. Montrer que la suite $(X_n, n \in \mathbb{N}^*)$ converge presque sûrement et donner sa limite. On pourra considérer dans un premier temps la suite $(\log(X_n), n \in \mathbb{N}^*)$.
2. Montrer que la suite $(Y_n, n \in \mathbb{N}^*)$, définie par $Y_n = (X_n \, \mathrm{e}^{\alpha})^{\sqrt{n}}$, converge en loi et déterminer la loi limite. On calculera la densité de la loi limite s'il s'agit d'une variable aléatoire continue.

△

Exercice VI.5.
Précision des sondages.

1. À quelle précision peut prétendre un sondage sur deux candidats effectué sur un échantillon de 1 000 personnes ? Est-ce que ce résultat dépend de la taille de la population ?

2. En Floride, pour l'élection présidentielle américaine 2000, on compte 6 millions de votants. Sachant qu'il y a eu environ 4 000 voix d'écart, quel est le nombre de personnes qu'il aurait fallu interroger dans un sondage pour savoir avec 95% de chance qui allait être le vainqueur ?

△

Exercice VI.6.
Soit X et X' deux variables aléatoires réelles, indépendantes, de carré intégrable et de même loi $\mathcal{L}$. On désire déterminer les lois $\mathcal{L}$ vérifiant la condition (C) suivante :

$$\frac{X+X'}{\sqrt{2}} \quad \text{et } X \text{ ont même loi.}$$

1. Calculer $\mathbb{E}[X]$.
2. Vérifier que la loi gaussienne centrée réduite vérifie la condition (C).
3. En utilisant le théorème de la limite centrale, déterminer les seules lois vérifiant la condition (C). (Indication : déterminer par récurrence la loi de $2^{n/2} \sum_{k=1}^{2^n} X_k$.)

△

Exercice VI.7.
Soit $(X_n, n \in \mathbb{N}^*)$ une suite de variables aléatoires réelles continues, indépendantes et de même loi. On suppose que la densité, f, de leur loi est bornée, symétrique, continue en 0 et telle que $f(0) > 0$.

1. Montrer que la suite de variable aléatoire $\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}$ converge en loi vers une variable aléatoire de Cauchy dont on déterminera la paramètre en fonction de $f(0)$. On rappelle que
$$\int_0^\infty \frac{1-\cos(u)}{u^2}\, du = \lim_{T\to\infty} \int_0^T \frac{\sin(u)}{u}\, du = \frac{\pi}{2}.$$
2. En déduire que la moyenne harmonique empirique $\tilde{X}_n = \frac{n}{\sum_{i=1}^{n} \frac{1}{X_i}}$ converge en loi vers une loi de Cauchy.

△

Exercice VI.8.
Soit $X = (X_1, \ldots, X_d)$ un vecteur aléatoire à valeurs dans $\mathbb{R}^d$. On suppose que la loi de X est invariante par rotation et que les variables aléatoires $X_1, \ldots, X_d$ sont indépendantes. Le but de cet exercice est de déterminer la loi de X.

1. On pose $g(x) = \psi_{X_1}(\sqrt{x})$ pour $x \geq 0$. Vérifier que g est réelle et solution de
$$\prod_{k=1}^{d} g(v_k) = g(\sum_{k=1}^{d} v_k), \quad \text{pour tout } v_1, \ldots, v_d \in [0, \infty[. \tag{VI.3}$$
2. En déduire que $\psi_{X_1}(u_1) = \mathrm{e}^{-\sigma^2 u_1^2/2}$, pour $\sigma \geq 0$. Montrer que soit $X = 0$ p.s. soit $X_1, \ldots, X_d$ sont de même loi gaussienne centrée.

△

Séance VII : Chaîne de Markov

Exercice VII.1.
En 1873, Galton publie un problème concernant le calcul de la probabilité d'extinction des noms de familles. N'obtenant pas de réponse satisfaisante, il contacte Watson qui fournit une réponse partielle. Ce n'est qu'à partir de 1930 que ce problème attire à nouveau l'attention et obtient alors une réponse détaillée[4].

Le but du problème qui suit est, à partir d'un modèle élémentaire d'évolution de population, appelé modèle de Galton-Watson, de déterminer cette probabilité d'extinction.

On considère un individu masculin à l'instant 0, et on note Z_n le nombre de descendants masculin de cet individu à la n-ième génération ($Z_0 = 1$ par convention). On suppose que les nombres de garçons de chaque individu sont indépendants et de même loi qu'une variable aléatoire, ξ, à valeurs entières. Plus précisément, soit $(\xi_{i,n}, i \geq 1, n \geq 0)$ une suite doublement indicée de variables aléatoires indépendantes de même loi que ξ. Le nombre d'individus de la $n+1$-ième génération est la somme des garçons des individus de la n-ième génération : pour $n \geq 0$,

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_{i,n},$$

avec la convention que $Z_{n+1} = 0$ si $Z_n = 0$. On note

$$\eta = \mathbb{P}(\text{il existe } n \geq 0 \text{ tel que } Z_n = 0)$$

la probabilité d'extinction de la population. Pour $k \in \mathbb{N}$, on note $p_k = \mathbb{P}(\xi = k)$, et l'on suppose que $\boxed{p_0 > 0}$ (sinon presque sûrement la population ne s'éteint pas).

I Calcul de la probabilité d'extinction

1. Montrer que η est la limite croissante de la suite $(\mathbb{P}(Z_n = 0), n \geq 0)$.

On suppose que ξ est intégrable, et on pose $m = \mathbb{E}[\xi]$.

2. Calculer $\mathbb{E}[Z_{n+1}|Z_n]$. En déduire que $\mathbb{E}[Z_n] = m^n$.
3. Montrer que si $m < 1$, alors $\eta = 1$, i.e. la population s'éteint presque sûrement.

On note ϕ la fonction génératrice de ξ, et ϕ_n la fonction génératrice de Z_n (et $\phi_0(z) = z$ pour $z \in [0,1]$).

4. Calculer $\mathbb{E}[z^{Z_{n+1}}|Z_n]$ pour $z \in [-1,1]$. En déduire que $\phi_{n+1} = \phi_n \circ \phi$, puis que $\phi_{n+1} = \phi \circ \phi_n$.
5. Montrer que $\mathbb{P}(Z_{n+1} = 0) = \phi(\mathbb{P}(Z_n = 0))$. En déduire que η est solution de l'équation

$$\phi(x) = x. \tag{VII.4}$$

[4]Voir l'article de D. Kendall. Branching processes since 1873, *J. London Math. Soc.* (1966), **41** pp. 385-406.

6. Calculer $\phi'(1)$. Vérifier que si $m \geq 1$, alors ϕ est strictement convexe sur $[0,1]$. Tracer le graphe $z \mapsto \phi(z)$ pour $z \in [0,1]$.
7. En déduire que si $m = 1$, alors $\eta = 1$.

On suppose dorénavant que $m > 1$.

8. Montrer que (VII.4) possède une unique solution $x_0 \in]0,1[$.
9. Montrer que $\mathbb{P}(Z_n = 0) \leq x_0$ pour tout $n \geq 0$. En déduire que $\eta = x_0$.

II Comportement asymptotique sur un exemple

Les données concernant les U.S.A. en 1920 pour la population masculine (cf la référence (4) en bas de page 13) sont telles que l'on peut modéliser la loi de ξ sous la forme

$$p_0 = \alpha, \quad \text{et pour } k \geq 1, \quad p_k = (1-\alpha)(1-\beta)\beta^{k-1}, \tag{VII.5}$$

avec $0 < \alpha < \beta < 1$. On suppose dorénavant que la loi de ξ est donnée par (VII.5).

1. Calculer $m = \mathbb{E}[\xi]$, vérifier que $m > 1$ et calculer η, l'unique solution de (VII.4) dans $]0,1[$, où ϕ est la fonction génératrice de ξ. Application numérique (cf la note (4) en bas de page 13) : $\alpha = 0.4813$ et $\beta = 0.5586$.
2. Vérifier que $\dfrac{\phi(z)-1}{\phi(z)-\eta} = m\dfrac{z-1}{z-\eta}$. En déduire $\phi_n(z)$, où $\phi_1 = \phi$ et pour $n \geq 2$, $\phi_n = \phi \circ \phi_{n-1}$.
3. Calculer la fonction caractéristique de XY, où X et Y sont indépendants, X est une variable aléatoire de Bernoulli de paramètre $p \in [0,1]$, et Y une variable aléatoire exponentielle de paramètre $\lambda > 0$.
4. Montrer que la suite $(m^{-n}Z_n, n \geq 1)$, où Z_n est une variable aléatoire de fonction génératrice ϕ_n, converge en loi vers une variable aléatoire Z dont on reconnaîtra la loi.

△

Exercice VII.2 (Chaîne de Markov à deux états)**.**
On considère l'espace à deux états $E = \{a, b\}$. La matrice stochastique la plus générale s'écrit

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}.$$

1. Calculer la ou les probabilités invariantes. À quelle condition existe-t-il une seule probabilité invariante ?
2. On suppose que la chaîne de Markov associée à P est irréductible (i.e. $P(a,b) > 0$ et $P(b,a) > 0$). Montrer que les puissances de P s'écrivent

$$P^n = \begin{pmatrix} 1-p & p \\ 1-p & p \end{pmatrix} + \gamma^n \begin{pmatrix} p & -p \\ -1+p & 1-p \end{pmatrix}.$$

   où l'on déterminera p et γ.
3. Montrer que si la chaîne de Markov est irréductible avec $P(a,a) > 0$ ou $P(b,b) > 0$, alors $\lim\limits_{n\to\infty} \nu_0 P^n$ est indépendant de la loi initiale ν_0.

△

Exercice VII.3.
Soit $(X_n, n \geq 1)$ une suite de variables aléatoires indépendantes de loi de Bernoulli de paramètre 1/2. On note $T_m = \inf\{k \geq m; X_k = 1, \ldots, X_{k-m+1} = 1\}$. Proposer une méthode pour calculer $\mathbb{P}(T_m \leq n)$ à l'aide d'une chaîne de Markov. △

Séance VII : Révision

Exercice VII.1.
Soit $(X_n, n \geq 1)$ une suite de variables aléatoires indépendantes de loi de Poisson de paramètre 1. On note $S_n = \sum_{k=1}^n X_k$.

1. Montrer que $\dfrac{S_n - n}{\sqrt{n}} \xrightarrow[n\to\infty]{\text{Loi}} \mathcal{N}(0,1)$.
2. Déterminer la loi de S_n, puis montrer que $\lim_{n\to\infty} \mathrm{e}^{-n}(1 + n + \dfrac{n^2}{2!} + \cdots + \dfrac{n^n}{n!}) = \dfrac{1}{2}$.

$\triangle$

Exercice VII.2.
On effectue n séries de 400 tirages de pile ou face avec une pièce équilibrée. On observe les fréquences empiriques de pile $F_1, \ldots, F_n$ dans ces séries.

1. Quelle est (approximativement) la loi de probabilité du nombre N de ces fréquences $(F_i, 1 \leq i \leq n)$ qui ne vérifient pas la condition $0.45 < F_i < 0.55$, lorsque $n = 20$?
2. Est-il plus probable que $N = 0$, que $N = 1$ ou que $N \geq 2$?

$\triangle$

Exercice VII.3.
Soit f une densité sur $\mathbb{R}$ et g une fonction telle que $\int |g(x)|\, f(x)dx < \infty$. On désire calculer $\bar{g} = \int g(x)\, f(x)dx$ par une méthode de Monte-Carlo. Soit $(X_n, n \in \mathbb{N}^*)$ une suite de variables aléatoires indépendantes et de même loi de densité q. On considère $\hat{g}_n = \dfrac{1}{n}\displaystyle\sum_{k=1}^n g(X_k)\dfrac{f(X_k)}{q(X_k)}$.

1. Montrer que $(\hat{g}_n, n \in \mathbb{N}^*)$ converge p.s. vers $\bar{g}$.
2. Donner une condition pour que la suite $(\hat{g}_n, n \in \mathbb{N}^*)$ satisfasse le TCL. Calculer alors la variance asymptotique σ_q^2 et donner un intervalle de confiance pour $\bar{g}$.
3. Montrer que $\sigma_q^2 \geq \left(\int |g(x)|\, f(x)dx\right)^2 - \bar{g}^2$. En déduire le choix optimal de q.

$\triangle$

Exercice VII.4.
Soit $(X_n, n \in \mathbb{N}^*)$ une suite de variables aléatoires réelles indépendantes de carré intégrables. On note $\mu = \mathbb{E}[X_1]$, $\sigma^2 = \mathrm{Var}(X_1)$ et $Z_n = \dfrac{1}{\sqrt{n}}\displaystyle\sum_{k=1}^n (X_k - \mu)$.

1. Rappeler la convergence en loi de la suite $(Z_n, n \in \mathbb{N}^*)$.
2. Établir la convergence de la suite $(Z_{2n} - Z_n, n \in \mathbb{N}^*)$ et donner sa limite.
3. En déduire que la suite $(Z_n, n \in \mathbb{N}^*)$ ne converge pas en probabilité si $\sigma^2 > 0$.

$\triangle$

Exercice VII.5.
Soit $n \geq 2$ et $X_1, \ldots, X_n$ une suite de n variables aléatoires indépendantes de loi uniforme sur $[0,1]$. On note $Y = \min_{1\leq i\leq n} X_i$ et $Z = \max_{1\leq i\leq n} X_i$.

1. Calculer la loi de (Y, Z).
2. En déduire la loi de Y et la loi de Z. Reconnaître ces deux lois.
3. Calculer $\mathbb{E}[Y|Z]$.
4. Calculer $\mathbb{E}[g(Y/Z)|Z]$, pour une fonction g mesurable bornée. En déduire puis reconnaître la loi de Y/Z conditionnellement à Z. Retrouver le résultat de la question 3.
5. Montrer que $(1-Z, 1-Y)$ a même loi que (Y, Z).
6. En déduire que $(1-Z)/(1-Y)$ est indépendant de Y.

△

Exercice VII.6.
Soit (X, Y) un vecteur gaussien de $\mathbb{R}^2$.

1. Trouver $a, b \in \mathbb{R}$ tels que $Y - (aX + b)$ soit indépendant de X.
2. En déduire $\mathbb{E}[Y|X]$ et $\mathrm{Var}(Y - \mathbb{E}[Y|X])$.

△

Exercice VII.7.
Soit T une variable aléatoire de loi de Cauchy. Sa densité est $f(x) = \dfrac{1}{\pi}\dfrac{1}{1+x^2}$, $x \in \mathbb{R}$. Soit θ une variable aléatoire de loi uniforme sur $[0, 2\pi]$.

1. Montrer que $\tan(\theta)$ a même loi que T (on pourra considérer les intervalles $]0, \pi/2[$, $]\pi/2, 3\pi/2[$ et $]3\pi/2, 2\pi[$).

Soit R une variable aléatoire positive indépendante de θ et telle R^2 est de loi exponentielle de paramètre $1/2$. On pose

$$X = R\cos(\theta) \quad \text{et} \quad Y = R\sin(\theta). \tag{VII.6}$$

On rappelle que X et Y sont indépendants de loi $\mathcal{N}(0,1)$ (transformation de Box-Muller).

2. Vérifier que T a même loi que Y/X. En déduire que T a même loi que $1/T$.
3. Vérifier que (X, Y) est un vecteur gaussien. Puis déterminer la loi de $\dfrac{1}{\sqrt{2}}(X-Y, X+Y)$.
   En déduire, à l'aide de la question précédente, que T a même loi que $\dfrac{1+T}{1-T}$.
4. (a) Montrer que $(\cos(2\theta), \sin(2\theta))$ a même loi que $(\cos(\theta), \sin(\theta))$.
   (b) Montrer, à l'aide de (VII.6), que (V, W) a même loi que (X, Y), où
   $$V = \frac{X^2 - Y^2}{\sqrt{X^2 + Y^2}} \quad \text{et} \quad W = \frac{2XY}{\sqrt{X^2+Y^2}}.$$
   (c) En déduire que T a même loi que $\dfrac{1}{2}\left(T - \dfrac{1}{T}\right)$.

△