Présentation des chapitres de mathématiques de MP
Probabilités discrètes

Lorsque l'énoncé d'un exercice est accompagné du symbole « ⚝ » dans la marge, cela signifie qu'il ne peut pas être traité avec toute la rigueur mathématique attendue tant qu'on est cantonné aux outils de MPSI (j'indique alors clairement ce qu'il est nécessaire d'admettre). Vous pourrez essayer de les traiter à nouveau, rigoureusement, au moment adéquat l'année prochaine.

Les probabilités de 2[e] année consistent surtout en l'extension des résultats de 1[re] année, dans un cadre où les univers peuvent éventuellement être de cardinal infini. C'est souvent nécessaire lorsqu'on veut modéliser une expérience faisant apparaître (au moins en esprit) une infinité d'étapes, ou bien un nombre fini d'étapes qui peut néanmoins être arbitrairement élevé. Un exemple standard est celui d'une expérience ayant probabilité p d'aboutir à un succès, et qu'on répète jusqu'à enfin observer un succès (ou, cela revient au même si l'on y réfléchit bien : on répète l'expérience une infinité de fois et on observe le rang d'apparition du premier succès).

Exercice 1. Soient $n \in \mathbb{N} \setminus \{0\}$ et $p \in]0{,}1[$. On lance n fois une pièce non équilibrée, ayant probabilité p de faire face. Mathématiquement, cela peut se modéliser par la donnée de n variables aléatoires indépendantes $X_1, \dots, X_n$ suivant la loi de Bernoulli de paramètre p. Soit X la variable aléatoire valant $+\infty$ si l'on n'obtient jamais face, et valant $k \in [\![1, n]\!]$ si le premier face apparaît au k^{e} lancer. Autrement dit : $X = \inf \{i \in [\![1, n]\!] \mid X_i = 1\}$ (avec la convention $\inf \emptyset = +\infty$).

1. Donner la loi de X.
2. ⚝ On reprend la même expérience , mais en lançant une infinité de fois la pièce. Que s'attend-on à avoir pour $\mathrm{P}(X = k)$, pour tout $k \in \mathbb{N} \setminus \{0\}$, dans ce cas ? (Nous manquons à ce stade de théorèmes pour justifier que le raisonnement très tentant à effectuer est effectivement correct.)

On définira la loi géométrique pour modéliser ce genre d'expérience.

L'extension aux univers infinis ne pose presque pas de difficulté, tant qu'on se limite aux univers *en bijection avec* $\mathbb{N}$ (c'est-à-dire aux ensembles dits *dénombrables*). Pour comprendre le type de problème posé hors de ce cadre, on pourra méditer l'exercice suivant :

Exercice 2. On lance une pièce une infinité de fois, ce qu'on modélise par l'univers : $\Omega = \{0{,}1\}^{\mathbb{N}}$. Il est tentant d'affirmer que la probabilité uniforme $\{0{,}1\}$ en induit une sur Ω, comme c'est le cas pour les produits cartésiens finis. Notons P une telle probabilité sur Ω, si elle existe.

1. Soit $\omega \in \Omega$. Pour tout $n \in \mathbb{N}$, notons A_n l'évènement : « les n premiers lancers coïncident avec les n premiers termes de ω. » Calculer $\mathrm{P}(A_n)$ pour tout $n \in \mathbb{N}$, et en déduire : $\mathrm{P}(\{\omega\}) = 0$.
2. Comparer $\mathrm{P}(\Omega)$ et $\sum_{\omega \in \Omega} \mathrm{P}(\{\omega\})$. Est-ce que le résultat vous paraît normal ?

   Si le sens à donner à cette somme vous embête, remplacez Ω dans l'indexation par des parties FINIES $\Omega_n \subseteq \Omega$ *telles que : $\forall n \in \mathbb{N}$, $\Omega_n \subseteq \Omega_{n+1}$, et : $\Omega = \bigcup_{n \in \mathbb{N}} \Omega_n$, puis prenez la limite quand $n \to +\infty$.*
3. Vérifier que Ω n'est pas en bijection avec $\mathbb{N}$. Pour cela, raisonner par l'absurde en supposant l'existence d'une bijection $f : \mathbb{N} \to \Omega$, et montrer que l'élément de Ω suivant fournit une contradiction : $\omega = (1 - f(n)_n)_{n \in \mathbb{N}}$. *On a noté $f(n)_n$ le n^{e} terme de la suite $f(n)$.*

À partir du moment où le cadre théorique est correctement défini pour faire des probabilités sur des univers infinis (modulo une certaine contrainte formulée ci-dessus), de nouvelles questions peuvent se poser :

— sur la réalisation asymptotique d'un évènement lorsqu'on itère une expérience un nombre arbitrairement grand de fois ;

— sur la « convergence » d'une suite de variables aléatoires, ou du moins leur approximation (puisque le terme de convergence n'est pas défini en MP pour des variables aléatoires).

1 Étude asymptotique d'un évènement

À partir du moment où il est permis de formuler des expériences ayant une infinité d'étapes (ou du moins un nombre fini d'étapes pouvant être arbitrairement élevé, comme : « on répète le lancer d'une pièce tant qu'on n'a pas obtenu "face" »), il est possible de s'interroger sur la probabilité d'observer tel phénomène *au moins une fois* dans cette série infinie d'expériences (par exemple : avoir au moins une fois « face » dans une série infinie de lancers d'une pièce), ou d'observer tel phénomène *systématiquement*. Formuler de tels évènements mathématiquement nécessite d'écrire une réunion ou intersection infinie d'évènements :

Exercice 3. On lance une pièce une infinité de fois, et on le modélise avec l'univers $\Omega = \{0,1\}^{(\mathbb{N}\setminus\{0\})}$ (l'exercice 2 montre que cette modélisation peut être problématique, mais pour ce que je demande dans cet exercice ce ne sera pas le cas puisque je ne vais pas parler de probabilités : seulement d'évènements). Dans cette modélisation, on peut imaginer que le nombre 1 représente « face ».

Pour tout $i \in \mathbb{N} \setminus \{0\}$, soit A_i l'évènement : « le i^{e} lancer donne "face" ».

1. Écrire en termes ensemblistes l'évènement : « on obtient "face" à chaque lancer de cette série infinie de lancers ».
2. Écrire en termes ensemblistes l'évènement : « on obtient "face" au moins une fois dans cette série infinie de lancers ».
3. Écrire en termes ensemblistes l'évènement : « on obtient "face" une infinité de fois dans cette série infinie de lancers ».
4. Écrire en termes ensemblistes l'évènement : « on finit par obtenir exclusivement "face" au-delà d'un certain nombre de lancers ».

Avec les outils de 1^{re} année, calculer la probabilité d'une réunion ou intersection infinie est inaccessible. En seconde année, nous donnerons un certain nombre de résultats le permettant, dépendant des hypothèses sur les évènements de la réunion ou intersection. L'un d'eux est le *théorème de continuité monotone* qui nous dit essentiellement que le raisonnement naïf suivant : « lorsqu'on veut calculer la probabilité d'un évènement se produisant en une infinité d'étapes, il suffit de faire le calcul pour n étapes, puis de prendre la limite quand $n \to +\infty$ », est un raisonnement correct.

Un autre cadre faisant intervenir une étude asymptotique de la probabilité d'un évènement, est donné par les chaînes de Markov : on répète une certaine expérience une infinité de fois, dont l'état à la $(n+1)^{\text{e}}$ itération dépend uniquement de l'issue de la n^{e} itération (absence de mémoire). On se demande quelles issues de l'expérience sont les plus susceptibles d'advenir après un très grand nombre d'étapes. Nous donnons de tels exemples dans les exercices ci-dessous :

Exercice 4. Une abeille va chaque jour sur l'une des deux fleurs A et B. Au jour 0, elle va à la fleur A. À chaque nouvelle journée, il y a une probabilité $p \in]0,1[$ qu'elle aille sur la même fleur que la veille. Pour tout entier n, on note A_n l'évènement : « l'abeille est sur la fleur A au jour n » et B_n l'évènement : « l'abeille est sur la fleur B au jour n ». On pose $a_n = \mathrm{P}(A_n)$ et $b_n = \mathrm{P}(B_n)$.

1. Pour tout entier n, exprimer a_{n+1} et b_{n+1} en fonction de a_n et b_n.
2. En remarquant que $a_n + b_n = 1$, déterminer les expressions de a_n et b_n en fonction de n pour tout $n \in \mathbb{N}$.
3. Vers quoi tendent les suites $(a_n)_{n\geqslant 0}$ et $(b_n)_{n\geqslant 0}$? Interpréter.

Exercice 5. On considère un carré ABCD et son centre, noté O. Lorsqu'un jeton est placé sur l'un de ces cinq points, il doit, à chaque tour, se déplacer vers l'un des points voisins de manière équiprobable. Au départ, le jeton est en A. Pour tout $n \in \mathbb{N} \setminus \{0\}$, on note p_n la probabilité que le jeton soit en O après n tours. Démontrer que : $\forall n \in \mathbb{N} \setminus \{0\}$, $p_{n+1} = \frac{1}{3}(1 - p_n)$, et en déduire l'expression de p_n en fonction de $n \in \mathbb{N} \setminus \{0\}$. Quelle est la limite de $(p_n)_{n\geqslant 1}$ quand n tend vers $+\infty$?

Dans ces deux exercices, on s'en sort parce qu'il n'y a qu'un nombre raisonnable d'issues à l'expérience à chaque étape. Si l'expérience est plus complexe, ce n'est plus possible... À moins de tirer profit des enseignements du chapitre de réduction des matrices. Plus précisément, on y illustre comment nous pouvons expliciter des suites réelles ou complexes, même lorsqu'elles vérifient des relations de récurrence sophistiquées, en passant par l'étude d'une suite de vecteurs colonnes $(X_n)_{n\geqslant 0}$ vérifiant une relation de récurrence d'ordre 1 : $\forall n \in \mathbb{N}$, $X_{n+1} = AX_n$, avec A une matrice carrée. Nous pourrons en faire autant avec les chaînes de Markov, grâce à un bon usage de la formule des probabilités totales :

Exercice 6. Pour étudier les capacités d'adaptation des souris, des biologistes ont mis au point le protocole suivant. Ils placent une souris dans une boîte comportant trois issues A, B et C. Les deux premières conduisent à un cul-de-sac alors que la troisième conduit à un morceau de gruyère. Après que la souris a choisi son issue, les scientifiques la remettent au centre de la boîte pour répéter l'expérience. Ils observent les résultats suivants :

— si la souris choisit la sortie A, elle sort la fois suivante en B ou C de façon équiprobable ;

— si elle choisit la sortie B, elle sort la fois suivante en A ou C avec la même probabilité ;

— si elle choisit la sortie C, elle la choisit de nouveau systématiquement la fois suivante.

Pour tout $n \in \mathbb{N} \setminus \{0\}$, on note A_n (respectivement B_n, C_n) l'évènement : « la souris a choisi la sortie A (respectivement B, C) à la n^{e} expérience », et on pose : $a_n = \mathrm{P}(A_n)$, $b_n = \mathrm{P}(B_n)$ et $c_n = \mathrm{P}(C_n)$, et enfin : $X_n = \begin{pmatrix} a_n \\ b_n \\ c_n \end{pmatrix}$. Déterminer une matrice $M \in \mathrm{M}_3(\mathbb{R})$ telle que : $\forall n \in \mathbb{N} \setminus \{0\}$, $X_{n+1} = MX_n$.

On utilisera la formule des probabilités totales.

Exercice 7. Christian Renauld, Lionel Messire et Célian Baptiste font un « jeu » de survie. On suppose que chacun d'entre eux choisit au hasard (uniformément) une cible parmi les autres joueurs survivants, et lui tire dessus. Pour simplifier, les tirs des différents participants ont la même probabilité de succès $\frac{2}{3}$, et les succès des tirs sont indépendants les uns des autres. Tout le monde tire en même temps. On répète ce « jeu » jusqu'à ce qu'il n'y ait plus qu'un seul survivant, ou aucun.

Pour tout $n \in \mathbb{N} \setminus \{0\}$, on note les évènements Z_n : « il n'y a plus de survivant après n tirs (ou un nombre de tirs inférieur à n) », U_n : « il reste un survivant après n tirs (ou un nombre de tirs inférieur à n) », D_n : « il reste deux survivants après n tirs », et enfin T_n : « il reste trois survivants après n tirs ». Pour tout $n \in \mathbb{N} \setminus \{0\}$, soit X_n le vecteur colonne de $\mathrm{M}_{4,1}(\mathbb{R})$ dont les coordonnées sont respectivement $\mathrm{P}(\mathrm{Z}_n)$, $\mathrm{P}(\mathrm{U}_n)$, $\mathrm{P}(\mathrm{D}_n)$ et $\mathrm{P}(\mathrm{T}_n)$. Montrer :

$$\forall n \in \mathbb{N} \setminus \{0\}, \quad X_{n+1} = \begin{pmatrix} 1 & 0 & 4/9 & 2/27 \\ 0 & 1 & 4/9 & 5/9 \\ 0 & 0 & 1/9 & 1/3 \\ 0 & 0 & 0 & 1/27 \end{pmatrix} X_n.$$

La détermination de la dernière colonne nécessite de la patience, parce qu'il faut considérer plusieurs cas, pour passer de trois survivants à deux ou un, selon que chaque joueur ait une cible différente ou non. *On utilisera la formule des probabilités totales.*

Sachant étudier asymptotiquement des suites « géométriques » à valeurs matricielles, nous saurons en déduire des probabilités asymptotiques. Ce n'est pas le seul chapitre où l'algèbre linéaire viendra au secours d'autres domaines des mathématiques (voir le chapitre sur les équations différentielles linéaires).

2 Approximation de variables aléatoires

L'approximation de variables aléatoires est en fait une idée parmi les plus naturelles des probabilités, et nous allons l'approfondir (en restant dans le cadre du programme, qui n'introduit pas de variables aléatoires dont l'image est non dénombrable : il s'avère que c'est un facteur très limitant).

Un exemple où, je crois, l'idée est naturelle : si l'on veut modéliser (informatiquement ou en esprit) ce qu'est un tirage uniformément au hasard d'un nombre RÉEL de [0,1], on peut remplacer l'intervalle [0,1] par la donnée de N points équirépartis dans cet intervalle, et tirer au hasard un nombre parmi ces N points : si N est arbitrairement grand, l'intuition laisse entendre que « cela revient presque au même ». Formalisons :

Exercice 8. Pour tout $n \in \mathbb{N} \setminus \{0\}$, on note E_n l'ensemble : $E_n = \left\{ \dfrac{k}{n} \mid k \in [\![0, n]\!] \right\}$, et X_n est une variable aléatoire réelle discrète suivant la loi uniforme sur E_n.

1. Soit $x \in [0,1]$. Montrer : $\forall n \in \mathbb{N} \setminus \{0\}$, $\mathrm{P}(X_n \leqslant x) = \dfrac{\lfloor nx \rfloor + 1}{n+1}$, et en déduire :
$$\lim_{n \to +\infty} \mathrm{P}(X_n \leqslant x) = x.$$
2. Pour tout $(a, b) \in [0,1]^2$ tel que : $a \leqslant b$, montrer : $\lim_{n \to +\infty} \mathrm{P}(X_n \in [a, b]) = b - a$.
3. Soit $(c, d) \in \mathbb{R}^2$ tel que : $c < d$. On pose à présent : $\forall n \in \mathbb{N}\setminus\{0\}$, $E_n = \left\{ a + k\dfrac{b-a}{n} \mid k \in [\![0, n]\!] \right\}$, et X_n est une variable aléatoire réelle discrète suivant la loi uniforme sur E_n. Pour tout $(a, b) \in [c, d]^2$ tel que : $a \leqslant b$, montrer : $\lim_{n \to +\infty} \mathrm{P}(X_n \in [a, b]) = \dfrac{b-a}{d-c}$.

C'est le même principe qui est utilisé dans la méthode de Monte-Carlo.

Un autre exemple est donné dans l'exercice ci-dessous, qui formalise l'idée que quand on fait un tirage au hasard dans un ensemble extrêmement grand, tirer avec ou sans remise « revient presque au même », puisque même dans le cas d'une remise il est très peu probable de tomber deux fois sur le même élément. L'intérêt de l'approximation est qu'avec remise, le dénombrement est plus simple et on se ramène aisément à une loi usuelle.

Exercice 9. L'équipe de football de Lionel Messire, l'En Avant Montvendre (localisée près de Barcelonne dans la Drôme), possède N joueurs sous contrat fédéral, dont n joueurs qui ont régulièrement recours à des méthodes médicales illicites pour améliorer leurs performances.

Un beau jour, un contrôle anti-dopage fait irruption dans leur vestiaire, et choisit uniformément au hasard ℓ joueurs parmi les N de l'équipe, pour procéder à une analyse sanguine qui détecterait quasi-certainement les joueurs en fraude. On note $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé modélisant cette expérience, et X_N la variable aléatoire comptant, parmi les ℓ joueurs choisis par le contrôle anti-dopage, combien sont *effectivement* dopés.

1. Donner $X_N(\Omega)$, et justifier : $\forall k \in X_N(\Omega)$, $\mathrm{P}(X_N = k) = \dfrac{\binom{n}{k}\binom{N-n}{\ell-k}}{\binom{N}{\ell}}$.
2. Montrer : $\forall k \in \mathbb{N}$, $\dbinom{M}{k} \underset{M \to +\infty}{\sim} \dfrac{M^k}{k!}$.
3. On admet dans cette question que $n = pN$ où $p \in]0,1[$ est une constante. Montrer : $\forall k \in X_N(\Omega)$, $\lim_{N \to +\infty} \mathrm{P}(X_N = k) = \dbinom{\ell}{k} p^k (1-p)^{\ell-k}$. Que reconnaît-on ? À la lumière du paragraphe qui précède l'exercice, expliquer pourquoi ce résultat était prévisible.

Nous étudierons le cas plus particulier d'une suite de variables aléatoires suivant une loi binomiale dont le paramètre n (donnant le nombre d'épreuves) devient arbitrairement grand ; pour songer à des situations réelles où cela se produit, il suffit de penser à l'observation du nombre de personnes entrant dans un magasin dans une journée. Il est possible pour cela de regarder, à chaque minute de la journée, si quelqu'un est entré (cela correspond à $n = 24 \times 60 = 1440$). On y associe une probabilité p_n dont la détermination dans la vraie vie s'obtient par observation fréquentielle. Mais on peut affiner cette observation en observant plutôt à chaque *seconde* de la journée si quelqu'un est entré (on a alors $n = 86400$). Intuitivement, plus on segmente la journée en un nombre n d'intervalles de temps, et plus la modélisation est précise ; mais le prix à payer est que les calculs impliquant la loi binomiale de paramètre (n, p) deviennent de plus en plus complexes : songez à la présence de nombres tels que $n! = 86400!$... On va alors chercher une *approximation* de la loi binomiale, et obtenir ce qu'on appelle la *loi de Poisson*. L'exercice qui suit y initie et permet de conjecturer sa définition :

Exercice 10. Pour tout $n \in \mathbb{N} \setminus \{0\}$, soit X_n une variable aléatoire telle que : $X_n \sim \mathcal{B}(n, p_n)$, avec $p_n \in]0,1[$. On suppose : $\lim\limits_{n \to +\infty} np_n = \lambda \neq 0$. Montrer : $\forall k \in \mathbb{N}$, $\lim\limits_{n \to +\infty} \mathrm{P}(X_n = k) = e^{-\lambda} \dfrac{\lambda^k}{k!}$.

Seulement ces approximations de lois de variables aléatoires par passage à la limite (le terme savant est celui de *convergence en loi* de variables aléatoires, mais il n'est pas au programme) ne fournissent pas suffisamment d'information. Par exemple, la sagesse populaire connaît la loi des grands nombres, que nous formulerons en ces termes : si $X_1, \ldots, X_n$ sont n variables aléatoires indépendantes suivant la même loi (imaginez par exemple qu'on lance un dé équilibré à six faces n fois, et que X_i vaut 1 si le i^{e} lancer, et 0 sinon), alors la moyenne empirique $\dfrac{X_1 + \cdots + X_n}{n}$ se rapproche de l'espérance de X_1, d'autant plus que n est très grand. C'est ainsi que vous avez pu observer que lorsqu'on lance un dé un très grand nombre de fois, chaque face apparaît presque avec la proportion $\dfrac{1}{6}$ (qui correspond bien à l'espérance de $X_1 \sim \mathcal{B}\left(\frac{1}{6}\right)$). Mais ce n'est pas très précis : combien de lancers doit-on faire si l'on espère avoir moins de 1% d'écart entre la moyenne empirique et l'espérance ? et si l'on veut 0,1% ? etc. Pour répondre à cette question, on doit plutôt étudier des probabilités du type :

$$\mathrm{P}(|X - Y| \geqslant \varepsilon),$$

où $\varepsilon > 0$ et X et Y sont deux variables aléatoires dont l'une est une approximation de l'autre. En illustration, nous montrerons la *loi faible des grands nombres* selon laquelle, en reprenant les notations ci-dessus :

$$\mathrm{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathrm{E}(X_1)\right| \geqslant \varepsilon\right) = \underset{n \to +\infty}{O}\left(\frac{1}{n}\right)$$

(l'énoncé sera même un peu plus précis que cela : c'est la variance qui permet de mesurer la taille du grand o). Vous pouvez vous convaincre que c'est bien une reformulation savante de la loi des grands nombres telle que nous la formulions ci-dessus.

Les « inégalités de concentration », qui sont d'autres inégalités du même type (en variant les hypothèses sur les variables aléatoires, etc.), sont une source féconde d'exercices de probabilités.

3 Résumé : objectif du chapitre

Après avoir étendu la théorie des probabilités sur un univers fini au cas infini (à condition que l'univers soit en bijection avec $\mathbb{N}$), ce qui consistera notamment en la démonstration de formules permettant de calculer la probabilité d'une réunion ou intersection *infinie* d'évènements, nous verrons comment les outils d'algèbre linéaire de 2^{e} seconde nous permettent d'étudier l'issue la plus probable d'une chaîne de Markov (qui correspond, en gros, à une expérience répétée plusieurs fois, où l'état du système à la $(n+1)^{\text{e}}$ expérience ne dépend que de l'issue de la n^{e}).

Puis une préoccupation majeure du chapitre concernera l'étude asymptotique d'une suite de variables aléatoires, qui se compartimente essentiellement en deux problématiques : 1° approcher des variables aléatoires par des variables aléatoires dont la loi est plus simple (lorsque les paramètres de leurs lois ont un comportement asymptotique appréciable), 2° mesurer la probabilité que l'écart entre une variable aléatoire X et son approximation Y soit supérieur à une quantité ε donnée, grâce à l'inégalité de Markov ou de Bienaymé-Tchebychev. C'est ainsi que nous en déduirons notamment la loi faible des grands nombres.

Enfin, nous introduirons (dans le cas de variables aléatoires à valeurs dans $\mathbb{N}$) la *fonction génératrice* d'une variable aléatoire : c'est une fonction développable en série entière. Nous en parlons dans la présentation du chapitre sur les séries entières, et disions entre autres qu'il était parfois plus intéressant d'étudier une suite $(a_n)_{n \geqslant 0}$ en passant par la fonction $x \mapsto \displaystyle\sum_{n=0}^{+\infty} a_n x^n$. Suivant cette même idée, nous verrons que pour toute variable aléatoire X à valeurs dans $\mathbb{N}$, la fonction $G_X : x \mapsto \sum_{n=0}^{+\infty} \mathrm{P}(X=n)x^n$ l'encode complètement : la connaissance explicite de la fonction G_X permet (au moins en théorie) d'obtenir toutes les informations utiles concernant X : sa loi, son espérance, sa variance, etc. Cela tombe bien, parce que pour certaines variables aléatoires (notamment celles définies comme une somme de variables aléatoires *indépendantes*), il est plus facile d'étudier sa fonction génératrice que la variable aléatoire elle-même, pour des raisons déjà exposées dans le chapitre concerné.