Méthodes – Probabilités

1 ✓ Dénombrement : révisions de 1^re^ année

1.1 Formulaire de dénombrement

Nous rappelons quelques formules de dénombrements que l'on utilise souvent, *en particulier lorsque nous avons une probabilité uniforme à étudier*. En effet, si P est une probabilité **uniforme** sur un espace probabilisable $(\Omega, \mathcal{P}(\Omega))$, où Ω est un ensemble FINI, alors :

$$\forall A \in \mathcal{P}(\Omega), \quad \mathrm{P}(A) = \frac{\mathrm{card}(A)}{\mathrm{card}(\Omega)}.$$

Savoir calculer $\mathrm{P}(A)$ revient donc à savoir calculer le nombre d'éléments dans A et Ω. Or les évènements et l'univers sont souvent construits comme des listes vérifiant certaines contraintes, des partitions de Ω, etc. Il est donc essentiel **de savoir les reconnaître** d'une part, et d'en déduire leur cardinal d'autre part, grâce aux rappels ci-dessous.

Dans ce qui suit, n et p sont des entiers naturels, et pour le nombre d'arrangements ou de combinaisons on a de plus l'hypothèse : $n \geqslant p$.

1. Si E a n éléments, alors E^p a n^p éléments (en particulier, $[\![1, n]\!]^p$ a n^p éléments), et plus généralement :

   $$\boxed{\mathrm{card}(E_1 \times \cdots \times E_p) = \mathrm{card}(E_1) \times \mathrm{card}(E_2) \times \cdots \times \mathrm{card}(E_p).}$$

   Quand reconnaître un ensemble de la forme $E_1 \times \cdots \times E_p$. Cela apparaît lorsque l'univers d'une expérience aléatoire est l'ensemble des résultats donnés par p lancers de dés, de pièces, etc. : on consigne tous ces résultats dans un p-uplet.

   Par exemple, si Christian Renauld et Lionel Messire s'affrontent en lançant des dés, respectivement à 6 et 12 faces, alors les résultats de l'expérience sont des couples (a, b) avec $a \in [\![1,6]\!]$ et $b \in [\![1,12]\!]$. Donc l'univers Ω est $[\![1,6]\!] \times [\![1,12]\!]$. D'après cette formule : $\mathrm{card}(\Omega) = 6 \times 12$.

   S'ils les lancent dix fois, alors les résultats de l'expérience sont des couples $((a_1, b_1), (a_2, b_2), \ldots, (a_{10}, b_{10}))$ avec les a_i dans $[\![1,6]\!]$ et les b_i dans $[\![1,12]\!]$. Donc l'univers Ω est dans ce cas $([\![1,6]\!] \times [\![1,12]\!])^{10}$.

   Si toutes les composantes du p-uplet sont prises dans le même ensemble E, alors l'ensemble des p-uplets est E^p. Par exemple, si nous lançons dix dés à 6 faces, l'ensemble des résultats est $[\![1,6]\!]^{10}$.

2. Le nombre de permutations d'un ensemble à n éléments (bijections de cet ensemble dans lui-même) est $n!$:

   $$\boxed{\text{Nombre de permutations d'un ensemble à } n \text{ éléments} = n!.}$$

   Quand reconnaître un nombre de permutations. On les trouve dans un contexte analogue au cas des arrangements : voir le point suivant.

3. Le nombre d'arrangements, ou d'injections d'un ensemble à p éléments dans un ensemble à n éléments, ou de listes possibles de p éléments choisis parmi n éléments possibles, quand $n \geqslant p$, est :

   $$A_n^p = \frac{n!}{(n-p)!}$$

   $$\boxed{\begin{array}{c}\text{Nombre de listes de } p \text{ éléments choisis parmi } n \text{ possibles} \\ = A_n^p = \dfrac{n!}{(n-p)!} = n \times (n-1) \times \cdots \times (n-(p-1)).\end{array}}$$

   Quand reconnaître un nombre d'arrangements, c'est-à-dire l'ensemble des listes possibles à p éléments choisis par n éléments possibles. Ils apparaissent dès qu'il s'agit de choisir p éléments DIFFÉRENTS au sein d'un ensemble de n possibilités. Cela peut donc apparaître en bien des circonstances, souvent quand on veut dénombrer le nombre de cas favorables d'un évènement :

— quand on *classe* des objets numérotés (par exemple pour le podium d'une course de chevaux) ;
— quand on répète une expérience p fois, et que les issues sont différentes à chaque fois, le nombre total d'issues possibles étant n (par exemple pour p lancers de dés, où l'on se demande le nombre de façons d'avoir des numéros différents à chaque fois).

Ces deux situations ne sont pas exhaustives.

4. Le nombre de combinaisons, ou parties à p éléments d'un ensemble à n éléments, est :

$$\binom{n}{p} = \frac{n!}{(n-p)!p!}.$$

$$\boxed{\text{Nombre de combinaisons de } p \text{ éléments choisis parmi } n \text{ possibles} = \binom{n}{p} = \frac{n!}{(n-p)!p!} = \frac{n \times (n-1) \times \cdots \times (n-(p-1))}{p \times (p-1) \times \cdots \times 2}.}$$

Quand reconnaître un nombre de combinaisons, c'est-à-dire l'ensemble des parties à p éléments dans un ensemble à n éléments. Ce nombre de combinaisons apparaît dès qu'il s'agit de choisir p éléments DIFFÉRENTS parmi n, *mais qu'il serait indifférent de les choisir suivant un ordre précis* (c'est là la différence avec le nombre d'arrangements). On a coutume de simplement dire qu'on choisit p éléments INDISCERNABLES, ou SANS ORDRE. Cela s'utilise donc pour dénombrer :
— le nombre de choix de p objets parmi n, où tout ce qui nous intéresse est ce qui constitue l'ensemble de ces p objets, sans comparer les uns aux autres (exemples : le nombre de façons de tirer p boules dans une urne, le nombre de doubles paires possibles au poker – on se moque éperdument de distinguer les deux cartes au sein de chaque paire, etc. –)
— le nombre de RANGS, dans une expérience décomposée en n étapes, où l'on observe un certain phénomène ou prend une certaine décision (exemples : comptage des positions de k succès en n répétitions d'une expérience, ou comptage des moments où l'on « monte » si l'on veut compter le nombre de chemins de longueur n partant de (0,0), où à chaque abscisse entière on passe à la suivante en augmentant ou diminuant d'une unité l'ordonnée) ; à noter c'est souvent dans ce cas-là que vous y pensez le moins.

La distinction avec le nombre d'arrangements est parfois subtile. Nous en reparlons autour des exemples 1 à 3. → page 3

1.2 Comment retrouver les formules rapidement

Nous proposons des arguments informels, qui suffisent pour une recherche rapide au brouillon :
— **retrouver** card(E^p) **:** l'ensemble E^p est constitué des p-uplets $(x_1, \ldots, x_p)$ avec les x_i dans E ; il y a card(E) possibilités pour le choix de la première coordonnée x_1, et pour chacun de ces choix il y a encore card(E) possibilités pour le choix de la deuxième coordonnée x_2 ; par conséquent, si nous cherchons le nombre de possibilités pour (x_1, x_2), il y en a :

$$\sum_{x_1 \in E} \text{card}(E) = \text{card}(E) \times \text{card}(E) = (\text{card}(E))^2.$$

En continuant le raisonnement avec la troisième coordonnée, etc., jusqu'à la p^{e} (il y a chaque fois card(E) choix possibles), on obtient en tout :

$$\underbrace{\text{card}(E) \times \cdots \times \text{card}(E)}_{p \text{ fois}} = (\text{card}(E))^p$$

p-uplets possibles d'éléments de E, donc : $\text{card}(E^p) = (\text{card}(E))^p$.
— **retrouver le nombre de permutations d'un ensemble à n éléments :** nous avons n éléments distincts à permuter ; disons pour simplifier qu'il s'agit des entiers de $[\![1, n]\!]$; pour définir une permutation de ces entiers, on choisit :
— le premier entier de la liste : il y a n possibilités ;

  - le deuxième entier de la liste : ce peut être n'importe lequel, tant qu'il est différent du premier entier choisi, ce qui fait $n-1$ possibilités ;
  - le troisième entier de la liste : ce peut être n'importe lequel, tant qu'il est différent des deux premiers entiers choisis, ce qui fait $n-2$ possibilités ;
  - et ainsi de suite, jusqu'à avoir épuisé la liste des entiers.

  Il y a donc en tout, par principe multiplicatif :

$$n \times (n-1) \times (n-2) \times \cdots \times 1 = n!$$

  façons de permuter les entiers de $[\![1, n]\!]$.
- **retrouver le nombre de listes de p éléments choisis parmi n possibles :** le principe est proche du précédent : nous avons p éléments distincts à ordonner, choisis dans une liste de n éléments ; disons pour simplifier qu'il s'agit des entiers de $[\![1, n]\!]$; pour définir une liste de p éléments de $[\![1, n]\!]$, on choisit :
  - le premier entier de la liste : il y a n possibilités ;
  - le deuxième entier de la liste : ce peut être n'importe lequel, tant qu'il est différent du premier entier choisi, ce qui fait $n-1$ possibilités ;
  - le troisième entier de la liste : ce peut être n'importe lequel, tant qu'il est différent des deux premiers entiers choisis, ce qui fait $n-2$ possibilités ;
  - et ainsi de suite, jusqu'au choix du dernier entier de la liste : à ce stade, il reste $n-(p-1)$ choix possibles, puisque le p^{e} entier de la liste doit être distinct des $p-1$ précédents.

  Il y a donc en tout, par principe multiplicatif :

$$A_n^p = n \times (n-1) \times (n-2) \times \cdots \times (n-(p-1)) = \frac{n!}{(n-p)!}$$

  façons de faire une liste de p éléments distincts de $[\![1, n]\!]$.
- **retrouver le nombre de combinaisons de p éléments choisis parmi n possibles :** pour retrouver la formule pour $\binom{n}{p}$, il est nécessaire de savoir la retrouver (ou de s'en souvenir) pour A_n^p, parce qu'on utilise une relation très simple entre A_n^p et $\binom{n}{p}$: on note que pour avoir le nombre de listes de p éléments choisis parmi n possibles, il suffit de compter le nombre de façons de choisir p éléments parmi n, et de permuter toutes ces possibilités pour avoir tous les ordres possibles.
  Un exemple pour vous en convaincre : si l'on choisit deux entiers dans $[\![1,4]\!]$, les choix possibles sont $\{1,2\}$, $\{1,3\}$, $\{1,4\}$, $\{2,3\}$, $\{2,4\}$ et $\{3,4\}$ (cela fait six possibilités). Pour avoir le nombre de listes de deux éléments choisis dans $[\![1,4]\!]$, je permute tous ces couples et on obtient :

$$\begin{array}{cccccc} (1,2) & (2,1) & (1,3) & (3,1) & (1,4) & (4,1) \\ (2,3) & (3,2) & (2,4) & (4,2) & (3,4) & (4,3) \end{array}$$

  Il apparaît effectivement toutes les possibilités. On obtient $12 = A_4^2$, comme prévu.
  Revenons au cas général. Cela se traduit mathématiquement par :

$$A_n^p = \binom{n}{p} \times (\text{nombre de permutations de } p \text{ éléments}) = \binom{n}{p} \times p!.$$

  On en déduit : $\displaystyle\binom{n}{p} = \frac{A_n^p}{p!} = \frac{n!}{(n-p)!p!}$.

Notons que la distinction entre A_n^p et $\binom{n}{p}$ (qui revient à la question : faut-il compter l'ordre ou non dans une liste d'éléments ? faut-il les distinguer ?) **est une source fréquente d'erreurs**, qui tient à la modélisation d'une expérience (ladite modélisation n'étant pas unique ; ainsi, le recours aux deux quantités peut parfois faire sens). Illustrons la difficulté sur un exemple, où l'on tiendrait compte d'un ordre à tort.

Exemple 1. Soit $r \in \mathbb{N} \setminus \{0\}$. On lance une pièce équilibrée jusqu'à obtenir face pour la r^{e} fois. Pour tout $n \in \mathbb{N} \setminus \{0\}$, on se demande quelle est la probabilité que ce r^{e} face soit observé après n lancers.

Soit X la variable aléatoire donnant le nombre de lancers nécessaires pour obtenir le r^e face. Un raisonnement informel (mais facile à formaliser) serait le suivant : si $n \in \mathbb{N} \setminus \{0\}$, alors avoir $(X = n)$ signifie qu'en n lancers, nous avons vu r fois « face » (avec probabilité $\frac{1}{2}$) et $n - r$ fois « pile » (avec probabilité $\frac{1}{2}$). Les lancers étant indépendants (hypothèse implicite évidente), on en déduit :

$$\mathrm{P}(X = n) = (\text{nombre de façons d'avoir } r \text{ fois « face » en } n \text{ lancers}) \times \left(\frac{1}{2}\right)^r \times \left(\frac{1}{2}\right)^{n-r}.$$

Tout d'abord, on sait que le r^e lancer est un face, il n'y a pas le choix. Il reste donc à compter le nombre de façons d'avoir $r - 1$ fois « face » dans les $n - 1$ premiers lancers. Et là, on peut être perturbé : faut-il utiliser A_{n-1}^{r-1} (arranger $r - 1$ éléments parmi $n - 1$, en tenant compte de l'ordre), ou $\binom{n-1}{r-1}$ (choisir $r - 1$ éléments parmi $n - 1$, sans tenir compte de l'ordre) ?

Il serait tentant de penser qu'ici, comme il ne revient pas au même d'avoir pile, puis face, et d'avoir face, puis pile (ce sont deux résultats d'expérience différents), l'ordre compte ; et que par conséquent c'est A_{n-1}^{r-1} qui est pertinent pour dénombrer ces possibilités.

Ce raisonnement est FAUX, parce qu'on se trompe sur la nature des objets à dénombrer : ce qu'on dénombre est le choix des RANGS des lancers qui donnent « face ». Par exemple, si on obtient la succession de lancers : $(\mathrm{F}, \mathrm{F}, \mathrm{P}, \mathrm{F})$ (ici $r = 3$ et $n = 4$), alors les positions choisies sont l'ensemble $\{1,2\}$ (n'oublions pas qu'on a établi ci-dessus que le r^e face n'intervient pas dans notre dénombrement : il est nécessairement à la fin). Qu'on choisisse 2 et 1, ou 1 et 2, cela revient au même, on obtient toujours la succession $(\mathrm{F}, \mathrm{F}, \mathrm{P}, \mathrm{F})$. Donc les $r - 1$ RANGS retenus parmi les $n - 1$ possibles ne dépendent pas de l'ordre : on les dénombre à l'aide des coefficients binomiaux, et non des arrangements.

En conclusion :

$$\forall n \in \mathbb{N} \setminus [\![0, r-1]\!], \quad \mathrm{P}(X = n) = \binom{n-1}{r-1} \times \frac{1}{2^n}.$$

Voir aussi la section suivante pour plus de commentaires sur les coefficients binomiaux.

Tout est récapitulé dans la figure 1.

FIGURE 1 – Dénombrement : rappels.

Ensemble à dénombrer	Cardinal	Exemples de situation pratique
$E_1 \times E_2 \times \cdots \times E_p$	$\mathrm{card}(E_1) \times \mathrm{card}(E_2) \times \cdots \times \mathrm{card}(E_p)$	Dénombrement des résultats de p expériences. Le résultat de la i^e expérience est un élément de E_i. Cas d'un lancer de deux dés à six faces : $[\![1,6]\!] \times [\![1,6]\!]$.
Permutations de n éléments	$n!$	Comme pour les arrangements, avec $p = n$: voir ci-dessous.
Listes de p éléments choisis parmi n	$A_n^p = \dfrac{n!}{(n-p)!} = n \times (n-1) \times \cdots \times (n-(p-1))$	Dénombrement des résultats de p expériences, mais la différence avec ci-dessus est qu'on veut des résultats DIFFÉRENTS à chacune des expériences. Ou encore : classement de p éléments choisis parmi n.
Combinaisons de p éléments parmi n	$\dbinom{n}{p} = \dfrac{n!}{(n-p)!p!}$	On choisit p objets au sein de n indiscernables (au sens qu'ils jouent le même rôle vis-à-vis du problème), et on dénombre les possibilités.

La section suivante vous sera plus utile si vous voulez être en mesure de dénombrer sans vous sentir capables de manipuler ce vocabulaire ensembliste (évidemment, si nous formalisons des dénombrements, ce n'est pas sans raison : une approche intuitive ne sera d'aucun recours dans les cas subtils). Nous y donnons aussi plus d'exemples en guise d'illustrations.

1.3 Méthodes de dénombrement : mise en pratique

1.3.1 Passage au complémentaire

Pour dénombrer une partie d'un ensemble, il est parfois plus simple **de dénombrer son complémentaire**, par exemple pour se ramener à un ensemble plus petit. Le cardinal de la partie considérée est alors la différence entre le cardinal de l'ensemble global et celui du complémentaire de cette partie. Cela permet aussi de passer d'un « au moins » ou « au plus » à un « exactement », *parfois* plus facile à dénombrer.

Ce principe vaut pour les probabilités, avec un intérêt supplémentaire si l'on considère le complémentaire d'une union ou d'une intersection. En effet, le complémentaire « transforme » une union en intersection, et inversement, en vertu des formules : $\overline{A \cap B} = \bar{A} \cup \bar{B}$ et $\overline{A \cup B} = \bar{A} \cap \bar{B}$. Il peut être intéressant de transformer une union en intersection, de sorte à pouvoir ensuite utiliser l'indépendance des évènements ($\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$). À l'inverse, il peut être intéressant de transformer une intersection en union, *si elle est disjointe*, de sorte à pouvoir utiliser la σ-additivité d'une probabilité ($\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)$). Voir les sections 2.2 et 2.3 pour plus de détails. → page 12

→ page 14

Exemple 2. On choisit au hasard trois élèves parmi les quarante-deux de la classe de PSI 2022/2023, et on se demande quelle est la probabilité qu'il y ait au moins une fille parmi ces trois élèves. Puisqu'il y a neuf filles dans cette classe, on pourrait déterminer la probabilité qu'il y en ait respectivement une, deux, ou trois parmi les trois élèves tirés au hasard, mais il est plus direct de déterminer la probabilité qu'il n'y en ait aucune (cela se ramène à compter le nombre de façons de lister trois garçons sur les trente-trois possibles : cela donne A_{33}^3 possibilités), puis de passer au complémentaire. Cette probabilité égale donc :

$$1 - \frac{A_{33}^3}{A_{42}^3} = 1 - \frac{33!/30!}{42!/39!} = 1 - \frac{33 \times 32 \times 31}{42 \times 41 \times 40} = \frac{753}{1435} \approx 0{,}52.$$

J'ai choisi de compter des arrangements, mais puisque les choix d'élèves sont *a priori* simultanés, on pourrait très bien utiliser les coefficients binomiaux pour notre calcul. Le résultat final est le même. Voir ci-dessous.

1.3.2 Indiscernable implique choix simultanés

Lorsqu'on doit choisir p éléments INDISCERNABLES **parmi** n éléments donnés (sans considération d'ordre et sans répétition), on effectue un choix SIMULTANÉ de ces éléments à l'aide des combinaisons.

Dans la rédaction, on indique alors sobrement que l'on choisit simultanément p éléments **parmi** n et que l'on a $\binom{n}{p}$ possibilités de le faire.

L'adjectif « indiscernable » ne concerne pas la nature des objets choisis mais seulement leur fonction dans le dénombrement considéré.

Ainsi, lorsque votre serviteur choisit deux élèves au hasard pour lui chercher un café, les deux élèves choisis, bien que parfaitement distincts au sens de l'état civil, doivent être considérés comme indiscernables puisque chacun des deux a la même fonction : aller chercher un café.

Exemple 3. Dans l'ancienne formule du Loto, un tirage était constitué de 6 numéros dans $[\![1,49]\!]$. Le nombre de tirages possibles était donc le nombre de choix simultanés de 6 numéros **parmi** 49, ce qui donnait $\binom{49}{6} = 13983816$ tirages possibles.

Ci-dessus, je vous invite à choisir des objets indiscernables de façon simultanée. Rien n'interdit *a priori* de les choisir successivement (il faudrait alors compter le nombre d'arrangements de 6 parmi 49), mais cette démarche successive est très sujette aux erreurs car elle induit, sur les éléments choisis, un ordre qui n'a pas lieu d'être.

En procédant ainsi, on compterait plusieurs fois chaque choix, ce qui nécessiterait de corriger l'effectif trouvé. Par exemple, choisir les numéros 1, 4, 5, 7, 10, 22, ou 4, 5, 22, 10, 7, 1, donne la même grille du Loto, et plus généralement toute permutation de ces six numéros donnent la même grille. Si l'on tient compte de l'ordre et qu'on veut éviter les doublons dans le décompte, il faudrait donc diviser le nombre total de choix comptés par le nombre de permutations possibles de chaque choix : il y en a 6! par choix. On retrouverait alors $\dfrac{A_{49}^6}{6!} = 13983816$ tirages possibles.

1.3.3 Principe du « et »

Si le dénombrement d'un ensemble se décompose en une succession de p étapes offrant respectivement n_1, n_2, ..., n_p possibilités, où chacun des nombres n_i ne dépend que de l'étape i, le nombre total d'issues est égal à $n_1 \times n_2 \times \times \cdots \times n_p$ (le **produit**), parce que chaque choix d'une étape doit être associé à chaque choix de tout autre étape.

Dans la rédaction d'une solution de ce type, on dénombre chacune des étapes en les séparant par des « **et** ».

Exemple 4. Imaginons qu'il n'existe pas de principe de tête de série ni de chapeau pour le tirage au sort de la Coupe du Monde de football. On se demande alors combien de combinaisons possibles peut-on avoir d'un groupe de quatre équipes avec exactement une équipe de chaque continent représenté : Afrique, Amérique (Nord et Sud regroupés), Asie, Europe, sachant qu'il y a en lice :
— cinq équipes africaines ;
— huit équipes américaines ;
— cinq équipes asiatiques ;
— quatorze équipes européennes.
Nous voulons une équipe africaine (cinq possibilités), **et** une équipe américaine (huit possibilités), **et** une équipe asiatique (cinq possibilités), **et** une équipe européenne (quatorze possibilités). Cela donne $5 \times 8 \times 5 \times 14 = 2800$ possibilités.

Exemple 5. La classe de PSI 2022/2023, formée de trente-trois garçons et neuf filles, décide de monter une pièce de théâtre dont les personnages sont des professeurs. Il y a cinq hommes (MM. Bricard, Girot, Hauroigné, Louis, Winckler) et deux femmes (Mmes Nicolas-Pierre, Louis).

En plus de ces rôles, il y a un chœur de six élèves (composé indifféremment de garçons ou de filles) pour chanter les tourments de cette équipe pédagogique sur un ton psalmodique. On veut alors savoir combien il y a de distributions possibles des rôles : il s'agit de choisir cinq garçons parmi les trente-trois de la classe (A_{33}^5 possibilités – les rôles sont discernables), **et** deux filles parmi les dix de la classe (A_9^2 possibilités – les rôles sont discernables), **et** six élèves parmi les trente-cinq restants ($\binom{35}{6}$ possibilités – les six membres du chœur ne sont pas distingués). Il y a donc :

$$A_{33}^5 \times A_9^2 \times \binom{35}{6} = 3\,328\,424\,367\,206\,400$$

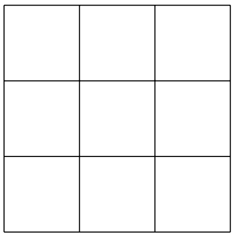
distributions possibles.

1.3.4 Principe du « ou »

Dans de nombreuses situations où l'on veut dénombrer un ensemble, on est amené à découper celui-ci en plusieurs sous-ensembles complémentaires et DISJOINTS qui sont faciles à dénombrer individuellement. Le cardinal de l'ensemble global est alors la **somme** des cardinaux des sous-ensembles de la partition considérée.

Dans la rédaction d'une solution de ce type, on distingue chacun des cas en les séparant par des « **ou** » (exclusifs) ou, ce qui revient au même, des « soit » (mais ce n'est pas un terme de la logique).

Exemple 6. Pour dénombrer le nombre de carrés (de toutes tailles) dans le quadrillage :



on constate qu'on a des carrés de taille 1×1 (il y en a neuf), **ou** des carrés de taille 2×2 (il y en a quatre), **ou** des carrés de taille 3×3 (il y en a un). Il y a donc, en tout, $9 + 4 + 1 = 14$ carrés dans ce quadrillage.

En résumé, « **et** » = **produit**, et « **ou** » = **somme** (en dénombrement).

2 Comment calculer une probabilité

On rappelle les différentes formules du cours servant à calculer des probabilités dans la figure 2.

FIGURE 2 – Formules de calculs de probabilités (rappels).

DÉFINITIONS, PROPRIÉTÉS DES PROBABILITÉS

$$\mathrm{P}(\emptyset) = 0, \quad \mathrm{P}(\Omega) = 1, \quad \mathrm{P}\left(\bar{A}\right) = 1 - \mathrm{P}(A), \quad \mathrm{P}(A \setminus B) = \mathrm{P}(A) - \mathrm{P}(A \cap B),$$

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B), \quad A \subseteq B \Rightarrow \mathrm{P}(A) \leqslant \mathrm{P}(B).$$

Définition d'une probabilité conditionnelle : $\mathrm{P}_B(A) = \dfrac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$ (si $\mathrm{P}(B) > 0$).
Définition d'évènements indépendants : $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$.

FORMULES UTILES

— Probabilités composées : $\mathrm{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathrm{P}(A_1)\mathrm{P}_{A_1}(A_2)\cdots\mathrm{P}_{A_1 \cap \cdots \cap A_{n-1}}(A_n)$.
— Probabilités totales : $(A_n)_{n \in \mathbb{N}}$ système complet d'évènements $\Rightarrow \mathrm{P}(B) = \sum\limits_{n=0}^{+\infty} \mathrm{P}(A_n)\mathrm{P}_{A_n}(B)$.
— Continuité croissante : $\forall n \in \mathbb{N},\ A_n \subseteq A_{n+1} \Rightarrow \mathrm{P}\left(\bigcup\limits_{n=0}^{+\infty} A_n\right) = \lim\limits_{n \to +\infty} \mathrm{P}(A_n)$.
— Continuité décroissante : $\forall n \in \mathbb{N},\ A_{n+1} \subseteq A_n \Rightarrow \mathrm{P}\left(\bigcap\limits_{n=0}^{+\infty} A_n\right) = \lim\limits_{n \to +\infty} \mathrm{P}(A_n)$.
— Formule de Bayes : $\mathrm{P}_B(A) = \mathrm{P}_A(B)\frac{\mathrm{P}(A)}{\mathrm{P}(B)}$.

Toute formule n'apparaissant pas dans ce formulaire est en général fautive et pure invention de votre part. Les enjeux de ce chapitre sont d'apprendre à les utiliser pour calculer toutes les probabilités rigoureusement. Si on ne le fait pas, en se contentant uniquement de l'intuition, alors ce chapitre ne vous apprendra rien. Vous savez d'expérience que les probabilités sont pleines de vérités contre-intuitives, et ces formules sont justement là pour fournir le cadre rigoureux évitant les écueils de l'intuition.

Il y a donc deux difficultés à surmonter dans ce chapitre :
— comment écrire correctement, d'un point de vue ensembliste, les évènements dont nous cherchons à calculer les probabilités (s'agit-il d'intersections ? de réunions ? si oui, de quels évènements ? s'il y a plusieurs façons d'écrire un évènement comme intersections et réunions d'autres évènements, comment les choisir pour que ce soit pertinent ?)
— une fois que nous avons reconnu une intersection ou une réunion d'évènements, il y a plusieurs formules semblant s'appliquer au calcul de sa probabilité : comment choisir la plus pertinente ?

Le deuxième point me paraît être le plus problématique pour vous. Vous êtes nombreux, une fois que vous avez écrit un évènement correctement comme réunions et intersections d'évènements plus simples, à ne pas utiliser cette écriture, et à calculer « à l'intuition » la probabilité demandée. C'est là que vous vous trompez la plupart du temps. Vous passez à côté de l'intérêt de l'écriture ensembliste comme réunion et intersection d'évènements plus simples : elle est là pour vous aider à ne pas vous tromper, justement !

J'ai souvent vu le raisonnement suivant (dont la conclusion est certes correcte) :

> Deux joueurs A et B lancent une pièce truquée à tour de rôle (en commençant par A), et la partie s'arrête dès qu'un joueur fait « pile » : le joueur ayant fait pile a alors gagné. La probabilité de faire « pile » est égale à p. Pour tout $n \in \mathbb{N} \setminus \{0\}$, on cherche la probabilité que A gagne à son n^{e} lancer. Si, pour tout $k \in \mathbb{N}$, on note $\mathrm{F}_{\mathrm{A},k}$ l'évènement : « le joueur A fait face au k^{e} lancer » (notation analogue pour l'évènement « le joueur B fait face au k^{e} lancer »), alors la probabilité recherchée est :
>
> $$\mathrm{P}(\mathrm{F}_{\mathrm{A},1} \cap \mathrm{F}_{\mathrm{B},1} \cap \mathrm{F}_{\mathrm{A},2} \cap \mathrm{F}_{\mathrm{B},2} \cap \cdots \cap \mathrm{F}_{\mathrm{A},n-1} \cap \mathrm{F}_{\mathrm{B},n-1} \cap \overline{\mathrm{F}_{\mathrm{A},n}}).$$
>
> Comme A et B ont une probabilité $1-p$ de faire « face » à chaque lancer, et qu'ils le font $n-1$ si l'on veut que l'évènement ci-dessus se réalise, et comme A a une probabilité p de faire « pile » au n^{e} lancer, on a :
>
> $$\mathrm{P}(\mathrm{F}_{\mathrm{A},1} \cap \mathrm{F}_{\mathrm{B},1} \cap \mathrm{F}_{\mathrm{A},2} \cap \mathrm{F}_{\mathrm{B},2} \cap \cdots \cap \mathrm{F}_{\mathrm{A},n-1} \cap \mathrm{F}_{\mathrm{B},n-1} \cap \overline{\mathrm{F}_{\mathrm{A},n}}) = (1-p)^{n-1} \times (1-p)^{n-1} \times p = (1-p)^{2(n-1)}p.$$

Bien que le résultat final soit juste (et puisse, d'ailleurs, se justifier rapidement avec l'usage de variables aléatoires de lois usuelles : comment ?), comment justifiez-vous MATHÉMATIQUEMENT ce raisonnement intuitif ? Et OÙ utilisez-vous le fait que ce soit une intersection d'évènements ?

Ce qui vous effraie est que vous trouvez les formules (rappelées dans la figure 2) compliquées à interpréter, ← page 7 en particulier la formule des probabilités composées à cause de ses nombreux conditionnements. L'objectif des prochaines sections est de vous convaincre qu'utiliser ces formules n'est pas si compliqué qu'il n'y paraît, et de montrer comment identifier les situations où elles sont pertinentes.

2.1 ✓ Formaliser un évènement : intersections, réunions, etc.

Le moyen fondamental de traduire un énoncé concret (écrit en français et non en langage mathématique) en une réunion ou une intersection d'évènements, est entièrement contenu dans les deux principes suivants :

$$
\begin{aligned}
(A \cup B) &= A \text{ OU } B \text{ se réalise} = \begin{cases} \text{AU MOINS UN des évènements} \\ A \text{ ou } B \text{ se réalise} \end{cases} \\
(A \cap B) &= A \text{ ET } B \text{ se réalisent tous les deux}
\end{aligned}
$$

En résumé, « ou » = réunion, et « et » = intersection. Cela entraîne les cas particuliers suivants, que vous croiserez souvent.

2.1.1 Cas particulier fréquent : évènement décrit en une succession d'étapes

Supposons que pour observer un certain évènement A, on passe d'abord par la réalisation d'étapes A_1, A_2, …, A_n (par exemple, si A est obtenu après n lancers d'une pièce ou d'un dé, ou après n tirages dans une urne, etc., et que chaque lancer ou tirage nécessite la vérification d'une condition pour que A se réalise). Alors informellement :

$$A = A_1 \text{ se réalise ET } A_2 \text{ se réalise ET } \ldots \text{ ET } A_n \text{ se réalise.}$$

Formellement, cela se traduit par : $A = A_1 \cap A_2 \cap \cdots \cap A_n = \bigcap_{i=1}^{n} A_i$, d'après le principe donné ci-dessus.

Remarque. Bien sûr, dans un exercice, c'est souvent à VOUS que revient l'initiative de définir les A_i. L'examinateur ne lit pas dans vos pensées et n'est pas tenu de deviner les évènements de votre intersection, si vous ne les définissez pas.

Exemple 7. On lance quatre fois un dé équilibré à six faces (ce qu'on peut modéliser avec l'univers $\Omega = [\![1{,}6]\!]^4$ muni de la probabilité uniforme), et on se demande la probabilité d'avoir alternativement un nombre pair et un nombre impair. Contentons-nous de la description ensembliste de l'évènement voulu. Pour cela, si l'on note pour tout $i \in [\![1{,}4]\!]$ l'évènement P_i : « avoir un nombre pair au i^e lancer », et I_i l'évènement : « avoir un nombre impair au i^e lancer » (noter que I_i est le contraire de P_i), on veut l'évènement :

P_1 se réalise ET I_2 se réalise ET P_3 se réalise ET I_4 se réalise,

OU

I_1 se réalise ET P_2 se réalise ET I_3 se réalise ET P_4 se réalise.

La première ligne ci-dessus s'écrit ensemblistement : $P_1 \cap I_2 \cap P_3 \cap I_4$. Le second s'écrit : $I_1 \cap P_2 \cap I_3 \cap P_4$. On veut que l'un OU l'autre se réalise. Donc l'évènement dont on cherche la probabilité est :

$$(P_1 \cap I_2 \cap P_3 \cap I_4) \cup (I_1 \cap P_2 \cap I_3 \cap P_4) \quad \text{(le parenthésage est important).}$$

Exercice 1. En déduire la probabilité demandée.

2.1.2 Cas particulier fréquent : évènement qu'on veut observer au moins une fois

On répète une certaine expérience n fois, et on se demande la probabilité qu'un certain évènement A soit observé AU MOINS UNE FOIS en ces n expériences. Pour cela, si pour tout $i \in [\![1, n]\!]$ l'on note A_i l'évènement : « observer A à la i^e expérience », alors on note qu'observer au moins une fois A équivaut à :

A_1 se réalise OU A_2 se réalise OU ... OU A_n se réalise.

Donc l'évènement étudié est : $A_1 \cup A_2 \cup \cdots \cup A_n = \bigcup\limits_{i=1}^{n} A_i$.

Noter que l'union n'est pas disjointe *a priori* : on pourrait observer plusieurs fois l'évènement A.

Dit plus intelligemment : A est observé au moins une fois si et seulement s'il existe (au moins un) $i \in [\![1, n]\!]$ tel que A_i se réalise. Interprété ainsi, on arrive plus facilement à généraliser la réflexion au cas suivant : il arrive qu'on répète une expérience un nombre infini de fois (au moins en esprit). Dans ce cas, dire que A est observé au moins une fois signifie qu'il existe (au moins un) $n \in \mathbb{N} \setminus \{0\}$ tel que A_n se réalise. Donc l'évènement « observer au moins une fois A » est :

$$\bigcup_{n=1}^{+\infty} A_n.$$

Cette fois-ci l'entier n peut être arbitrairement grand, qui sait ? vu qu'il y a une infinité d'itérations de l'expérience, on ne peut pas se contenter de considérer tous les cas jusqu'à un entier fixé. C'est pourquoi la réunion va jusqu'à l'infini.

Remarque. Bien sûr, dans un exercice, c'est souvent à VOUS que revient l'initiative de définir les A_i. L'examinateur ne lit pas dans vos pensées et n'est pas tenu de deviner les évènements de votre réunion, si vous ne les définissez pas.

2.1.3 Cas particulier fréquent : évènement qu'on veut observer plusieurs fois (éventuellement une infinité)

On répète une certaine expérience n fois, et on se demande la probabilité qu'un certain évènement A soit observé À CHACUNE de ces n expériences. Pour cela, si pour tout $i \in [\![1, n]\!]$ l'on note A_i l'évènement : « observer A à la i^e expérience », alors on note qu'observer systématiquement A équivaut à :

A_1 se réalise ET A_2 se réalise ET ... ET A_n se réalise.

Donc l'évènement étudié est : $A_1 \cap A_2 \cap \cdots \cap A_n = \bigcap_{i=1}^{n} A_i$.

Dit plus intelligemment : A est observé à chaque expérience si et seulement si, pour TOUT $i \in [\![1, n]\!]$, l'évènement A_i se réalise. Interprété ainsi, on arrive plus facilement à généraliser la réflexion au cas suivant : il arrive qu'on répète une expérience un nombre infini de fois (au moins en esprit). Dans ce cas, dire que A est observé à chaque fois signifie que pour TOUT $n \in \mathbb{N} \setminus \{0\}$ l'évènement A_n se réalise. Donc l'évènement « observer systématiquement A en une infinité d'expériences » est :

$$\bigcap_{n=1}^{+\infty} A_n.$$

Cette fois-ci l'entier n peut être arbitrairement grand, qui sait ? vu qu'il y a une infinité d'itérations de l'expérience, on ne peut pas se contenter de considérer tous les cas jusqu'à un entier fixé. C'est pourquoi l'intersection va jusqu'à l'infini.

Remarque. Bien sûr, dans un exercice, c'est souvent à VOUS que revient l'initiative de définir les A_i. L'examinateur ne lit pas dans vos pensées et n'est pas tenu de deviner les évènements de votre intersection, si vous ne les définissez pas.

Mise en garde 1. À manipuler les variables n comme des quantités abstraites ou symboliques, sous prétexte qu'elles sont notées avec des lettres et non des entiers explicites, vous faites très souvent les bêtises suivantes et je vous somme d'y prendre garde :

> Si l'on répète une infinité de fois une expérience, que A est l'évènement : « observer au moins une fois tel phénomène », et que pour tout $n \in \mathbb{N} \setminus \{0\}$ on définit A_n comme l'évènement : « on observe tel phénomène à la n^{e} expérience », alors A N'EST PAS LA RÉUNION SUIVANTE : $A_1 \cup \cdots \cup A_n$.
> De même, si A est l'évènement : « observer systématiquement la réalisation de tel phénomène », alors IL EST TOTALEMENT FAUX D'ÉCRIRE : $A = A_1 \cap \cdots \cap A_n$.
> En effet, dans ces deux cas, vous ne décrivez pas A, mais l'évènement : « observer au moins une fois tel phénomène entre la première et la n^{e} expérience » (dans le cas de la réunion), et l'évènement : « observer tel phénomène à chacune des n premières expériences » dans le cas de l'intersection, où n est un entier donné (qui, d'ailleurs, N'EST MÊME PAS DANS LA DÉFINITION DE A, donc mal défini !).
> Les descriptions correctes sont respectivement, pour la réunion : $A = \bigcup_{n=1}^{+\infty} A_n = A_1 \cup A_2 \cup A_3 \cup \cdots$ (ne s'arrête jamais), et pour l'intersection : $A = \bigcap_{n=1}^{+\infty} A_n = A_1 \cap A_2 \cap A_3 \cap \cdots$ (ne s'arrête jamais).
> Par contre, si vous ne répétez pas une infinité de fois l'expérience, mais un nombre n de fois (fini, donc), alors les descriptions ci-dessus sont correctes.
>
> Il vous arrive aussi de faire l'erreur exactement inverse : écrire $A = \bigcup_{n=1}^{+\infty} A_n$ alors qu'il n'intervient qu'un nombre fini possible d'évènements. Faites la différence entre une réunion finie, et une réunion infinie ! De même pour les intersections. Le seul moyen de ne pas se tromper est de prendre le temps de se demander : combien peut-il y avoir d'expériences en tout ? à laquelle veut-on observer tel phénomène ? Ce sont les entiers qui répondent à ces questions qui indexent les réunions ou intersections étudiées. Si les entiers à convenir peuvent être arbitrairement grands, alors la réunion ou intersection étudiée est indexée par un ensemble infini.

2.1.4 Cas extrêmement fréquent : évènement décrit avec des variables aléatoires

Dans tous les domaines de ce chapitre, retenez ce conseil important :

> **Dans la mesure du possible, définissez les évènements étudiés à l'aide de variables aléatoires réelles discrètes.**

C'est très souvent faisable : une variable aléatoire n'est rien d'autre qu'une grandeur (un nombre) mesurée dans une expérience aléatoire. Or, en sciences, à peu près tout peut être exprimé avec des grandeurs !

C'est important non seulement parce que vous pouvez gagner un temps significatif s'il y a des variables aléatoires de lois usuelles (et c'est souvent le cas), mais aussi parce qu'elles facilitent la description des évènements. Rappelons que si X est une variable aléatoire réelle discrète, alors :

$$(X \in U) = \bigcup_{x \in U} (X = x),$$

où l'union fait intervenir des évènements incompatibles : cela fournit « gratuitement » une écriture de l'évènement $(X \in U)$ à l'aide d'évènements plus simples ! (dont on a les probabilités si la loi de X est connue)

Exemple 8. Si X est une variable aléatoire réelle discrète, disons à valeurs dans $\mathbb{N}$, et qu'on se demande quelle est la probabilité de l'évènement $(X \geqslant k)$ (où k est un entier naturel fixé), alors on note qu'écrit INFORMELLEMENT au brouillon, on a :

$$(X \geqslant k) = (X = k) \text{ OU } (X = k+1) \text{ OU } (X = k+2) \text{ OU } (X = k+3) \text{ OU } (X = k+4) \text{ OU } \dots$$

et donc, FORMELLEMENT :

$$(X \geqslant k) = (X = k) \cup (X = k+1) \cup (X = k+2) \cup (X = k+3) \cup (X = k+4) \cup \cdots = \bigcup_{n=k}^{+\infty} (X = n).$$

la réunion va alors de $n = k$ à l'infini, parce que $(X \geqslant k)$ est vérifié dès que $(X = n)$ *pour n'importe quel entier n supérieur ou égal à k*. Dit moins formellement : dans notre réunion écrite avec des points de suspension, $(X = n)$ apparaît pour n'importe quelle valeur de n entre k et l'infini. Alors, par σ-additivité (les évènements $(X = n)$, pour $n \geqslant k$, sont toujours incompatibles pour une variable aléatoire X) :

$$\boxed{\mathrm{P}(X \geqslant k) = \mathrm{P}\left(\bigcup_{n=k}^{+\infty} (X = n)\right) = \sum_{n=k}^{+\infty} \mathrm{P}(X = n).}$$

Si, par contre, on voulait décrire l'évènement : $(a \leqslant X \leqslant b)$, où a et b sont des entiers naturels, alors on écrirait d'abord INFORMELLEMENT, au brouillon :

$$(a \leqslant X \leqslant b) = (X = a) \text{ OU } (X = a+1) \text{ OU} \dots \text{ OU } (X = b-1) \text{ OU } (X = b),$$

et donc, FORMELLEMENT : $(a \leqslant X \leqslant b) = (X = a) \cup (X = a+1) \cup \cdots \cup (X = b-1) \cup (X = b) = \bigcup\limits_{n=a}^{b} (X = n)$. Là encore, la σ-additivité permettrait d'en déduire la probabilité de $(a \leqslant X \leqslant b)$.

Exemple 9. On lance un dé à 20 faces – c'est-à-dire un icosaèdre – (disons non équilibré pour que la réponse ne soit pas triviale ; je ne donne pas les probabilités de tomber sur chaque face car là n'est pas le sujet). On se demande la probabilité que le résultat soit un nombre pair. Notons A cet évènement.

Pour correctement formaliser le problème, et suivant le conseil de définir des variables aléatoires dès que possible, soit X la variable aléatoire donnant le résultat du dé. On a, d'abord informellement :

$$A = (X \text{ pair}) = (X = 2) \text{ OU } (X = 4) \text{ OU } (X = 6) \text{ OU} \cdots \text{ OU } (X = 20)$$

et donc, formellement : $A = (X \text{ pair}) = (X = 2) \cup (X = 4) \cup (X = 6) \cup \cdots \cup (X = 20) = \bigcup\limits_{k=1}^{10} (X = 2k)$, parce qu'un entier pair entre 1 et 20 s'écrit $2k$ avec k entre 1 et 10. Alors, par σ-additivité :

$$\mathrm{P}(A) = \mathrm{P}\left(\bigcup_{k=1}^{10} (X = 2k)\right) = \sum_{k=1}^{10} \mathrm{P}(X = 2k).$$

Exercice 2. Des joueurs A et B s'affrontent dans un jeu de pile ou face. Chaque joueur lance une pièce, et gagne s'il fait pile (si les deux joueurs font pile, il y a match nul et on rejoue). Pour tout $n \in \mathbb{N} \setminus \{0\}$, on note A_n l'évènement : « le joueur A fait pile à la n^{e} manche » et B_n l'évènement : « le joueur B fait pile à la n^{e} manche ».

Écrire correctement les évènements :

1. Le joueur A remporte la partie avant la cinquième manche.
2. Les joueurs A et B font match nul jusqu'à la huitième manche.
3. Les joueurs A et B font match nul jusqu'à la n^{e} manche.
4. Les joueurs A et B font match nul jusqu'à la fin des temps.
5. Le joueur A remporte la partie.
6. Le joueur A remporte la partie après un nombre pair de manches.

Bien sûr, on peut définir les évènements en jeu plus synthétiquement en recourant aux variables aléatoires donnant le rang du premier pile pour A et B respectivement : l'intérêt est qu'elles suivent une loi géométrique. Refaire l'exercice en les utilisant.

Exercice 3. Deux joueurs A et B lancent à tour de rôle une pièce. Le joueur A commence, et la partie s'achève dès qu'un joueur obtient « face ».

Écrire correctement, à l'aide d'intersections d'évènements simples :

1. Le joueur A gagne lors de son n^{e} lancer.
2. Le joueur B gagne lors de son n^{e} lancer.
3. Le joueur A gagne.
4. Le joueur B gagne.
5. Le jeu ne s'arrête jamais.

Exercice 4. On lance une pièce une infinité de fois. Pour tout $n \in \mathbb{N} \setminus \{0\}$, on note A_n l'évènement : « la pièce tombe sur pile au n^{e} lancer ». Écrire, à l'aide des évènements de la famille $(A_n)_{n\in\mathbb{N}}$:

1. On obtient au moins une fois pile dans cette série infinie de lancers.
2. On n'obtient jamais pile dans cette série infinie de lancers.
3. On obtient systématiquement pile dans cette série infinie de lancers.
4. On obtient alternativement pile et face dans cette série infinie de lancers.
5. Au moins une fois, dans cette série infinie de lancers, on obtient deux piles consécutifs.

2.2 ✓ Calculer la probabilité d'une réunion d'évènements

Vous avez affaire à la probabilité d'une **réunion** d'évènements lorsque vous devez considérer la probabilité qu'il se produit AU MOINS UNE des issues parmi une liste d'issues possibles.

Pour calculer la probabilité d'une réunion d'évènements, demandez-vous si :

— ces évènements sont **incompatibles** (c'est-à-dire : peuvent-ils ou non se produire simultanément ?) ;
— ces évènements forment une **suite croissante** (si l'un se réalise, est-ce que les suivants aussi ?) ;
— ces évènements sont **indépendants** (est-ce que la réalisation de l'un influe sur la probabilité de réalisation des autres ?) ;
— ces évènements ne sont que deux, et la réalisation de l'un conditionne la réalisation de l'autre.

Selon les cas de figure, voici comment en déduire la probabilité d'une réunion :

— si les évènements de la famille $(A_n)_{n\in\mathbb{N}}$ sont **incompatibles** deux à deux, alors par σ-additivité :

$$\mathrm{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathrm{P}(A_n) \ ;$$

— si $(A_n)_{n\geqslant 0}$ est une suite d'évènements **croissante**, alors d'après le théorème de continuité croissante :

$$\mathrm{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} \mathrm{P}(A_n) \ ;$$

— si $(A_1, \dots, A_n)$ est une famille d'évènements **indépendants**, alors par passage au complémentaire et définition de l'indépendance :

$$\mathrm{P}\left(\bigcup_{i=1}^{n} A_i\right) = 1 - \mathrm{P}\left(\overline{\bigcup_{i=1}^{n} A_i}\right) = 1 - \mathrm{P}\left(\bigcap_{i=1}^{n} \overline{A_i}\right) = 1 - \prod_{i=1}^{n} \mathrm{P}(\overline{A_i}) = 1 - \prod_{i=1}^{n} (1 - \mathrm{P}(A_i)) \ ;$$

et on peut généraliser cette formule à une famille infinie quitte à passer à la limite ensuite (grâce au théorème de continuité croissante) ;

— si A et B sont deux évènements ne vérifiant rien de tout cela, alors d'après la formule de Moivre :

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B) \ .$$

Cette formule est utile si vous savez calculer $\mathrm{P}(A \cap B)$ grâce à un conditionnement : $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}_A(B)$, ou si $A \cap B$ admet une description plus simple que $A \cup B$.
Par exemple, si on tire un nombre au hasard entre 1 et 60, qu'on prend A l'évènement « être un multiple de 3 » et B l'évènement « être un multiple de 2 », alors $A \cup B$: « être un multiple de 3 ou 2 » est plus délicat à décrire rapidement (attention à ne pas compter les doublons, lorsqu'on énumère les multiples de 3, puis les multiples de 2 !) que l'évènement $A \cap B$: « être un multiple de 3 ET 2 », c'est-à-dire : « être un multiple de 6 » (on voit alors immédiatement qu'il y en a dix entre 1 et 60).
Une raison pour laquelle $A \cap B$ est souvent plus simple à déterminer est que c'est un ensemble plus petit que $A \cup B$, A et B.

Il est très important de souligner qu'en général, **vous savez dès le début dans quel cas de figure vous êtes**, et je dirais même que vous avez modélisé le problème et choisi les évènements de la réunion de sorte à vous placer sciemment dans ces contextes favorables :

— on a naturellement des évènements **incompatibles** lorsque :

— X est une variable aléatoire et qu'on prend $A_n = (X = x_n)$; plus généralement, on rappelle que :

$$(X \in U) = \bigcup_{x \in U} (X = x),$$

et tous les $(X = x)$ sont incompatibles (on illustre ce type d'évènements dans la section 2.1) ; ← page 8

— les A_n forment l'ensemble des issues possibles d'une expérience, par exemple :

A_n : « avoir le (premier) succès à la n^{e} itération de l'expérience ».

— on a naturellement une suite d'évènements **croissante** lorsqu'on répète une certaine expérience un nombre éventuellement infini de fois, et qu'on cherche à étudier les évènements :
« l'expérience finit par se terminer (sous condition) », ou :
« on observe au moins une fois tel phénomène dans cette infinité d'expériences », etc. ;
il suffit en effet de prendre pour A_n :

A_n : « la partie s'est terminée en moins de n itérations de l'expérience »,

A_n : « on observe au moins une fois tel phénomène dans les n premières itérations de l'expérience »,

où nous vous laissons vérifier la monotonie ;

— on a des évènements **indépendants** lorsque qu'on répète une expérience plusieurs fois (lancer un dé, par exemple) sans que le résultat d'une expérience n'influe sur les suivantes ; si A_n désigne un évènement concernant la n^{e} itération de l'expérience (et uniquement celle-ci), alors en général les A_n forment une famille d'évènements indépendants (noter que c'est souvent une hypothèse donnée dans l'énoncé).

On n'avance donc pas à l'aveugle et nous sommes méthodiques.

Ces conseils sont récapitulés dans la figure 3.

FIGURE 3 – Cas de la probabilité d'une réunion.

σ-ADDITIVITÉ (évènements incompatibles)

Si les A_n sont INCOMPATIBLES (synonyme : DISJOINTS), alors : $\mathrm{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathrm{P}(A_n)$.
Cadre d'application. Il faut que les évènements A_i ne puissent pas se produire simultanément.
Exemple fréquent. Les A_n sont toutes les issues possibles d'une expérience. Plus généralement : A_n est de la forme $(X = x_n)$ avec X une variable aléatoire réelle discrète.

THÉORÈME DE CONTINUITÉ CROISSANTE (suite croissante d'évènements)

Continuité croissante : $\forall n \in \mathbb{N},\ A_n \subseteq A_{n+1} \Rightarrow \mathrm{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} \mathrm{P}(A_n)$.
Sans hypothèse : $\mathrm{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{N\to+\infty} \mathrm{P}\left(\bigcup_{n=0}^{N} A_n\right)$.
Cadre d'application. C'est le contraire ici : si A_n se réalise, alors tous les suivants aussi.
Exemple fréquent. L'évènement A_n est l'observation d'AU MOINS un succès en n itérations d'une expérience.

PASSAGE AU COMPLÉMENTAIRE (évènements indépendants)

$\mathrm{P}\left(\bigcup_{i=1}^{n} A_i\right) = 1 - \mathrm{P}\left(\overline{\bigcup_{i=1}^{n} A_i}\right) = 1 - \mathrm{P}\left(\bigcap_{i=1}^{n} \overline{A_i}\right)$.
Cadre d'application. En général, on gagne à passer aux intersections si l'on sait que les évènements sont INDÉPENDANTS. Dans ce cas, on peut écrire : $\mathrm{P}\left(\bigcup_{i=1}^{n} A_i\right) = 1 - \prod_{i=1}^{n} \mathrm{P}\left(\overline{A_i}\right) = 1 - \prod_{i=1}^{n} (1 - \mathrm{P}(A_i))$. Si la réunion est infinie : d'abord écrire la réunion finie, puis utiliser le théorème de continuité croissante.
Exemple fréquent. Les évènements A_i sont de la forme $(X_i \in U_i)$ où les X_i sont des variables aléatoires indépendantes par hypothèse (résultats de différents dés ou pièces, par exemple).

LA FORMULE DE MOIVRE

$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$.
Cadre d'application. Vous ne la connaissez que dans le cas d'une réunion de deux évènements, ce qui limite déjà son cadre d'utilisation. Il n'est intéressant d'utiliser cette formule ici que si vous connaissez $\mathrm{P}(A \cap B)$, sinon vous êtes dans une impasse. C'est possible si vous connaissez la probabilité conditionnelle de B sachant A (ou de A sachant B), ou si A et B sont indépendants.
Exemple fréquent. Les évènements A et B sont de la forme $A = (X \in U)$ et $B = (Y \in V)$, où X et Y sont deux variables aléatoires dont on connaît soit l'indépendance, soit la loi conjointe. On rencontre aussi le cas où $A = (X \in U)$ et $B = (X \in V)$, et cela ramène le calcul de $\mathrm{P}(X \in U \cup V)$ à celui de $\mathrm{P}(X \in U \cap V)$: intéressant si $U \cap V$ est plus simple à décrire.

2.3 ✓ Calculer la probabilité d'une intersection d'évènements

Vous avez affaire à la probabilité d'une **intersection** d'évènements lorsque vous devez considérer la probabilité que se réalisent TOUS les évènements d'une liste. C'est en particulier le cas lorsque, pour qu'il se réalise un certain évènement après n étapes, il faut d'abord que se réalisent des évènements $A_1, \ldots, A_n$ à chaque étape.

Pour calculer la probabilité d'une intersection d'évènements, demandez-vous si :
— ces évènements sont **indépendants** (c'est-à-dire : la réalisation de l'un influe-t-elle sur les autres ?) ;
— ces évènements forment une **suite décroissante** (si l'un ne se réalise pas, est-ce que les suivants non plus ? si l'un se réalise, est-ce que les précédents sont aussi réalisés nécessairement ?) ;
— ces évènements **suivent un ordre logique ou chronologique naturel** ;
— ces évènements ne sont que deux, et la **non réalisation** simultanée des deux évènements est plus facile à étudier que leur réalisation simultanée.

Dans le cas d'évènements incompatibles, l'intersection est un évènement impossible et sa probabilité est trivialement nulle.

Selon les cas de figure, voici comment en déduire la probabilité d'une intersection :
— si $(A_1, \dots, A_n)$ est une famille d'évènements **indépendants**, alors par définition de l'indépendance :

$$\mathrm{P}\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} \mathrm{P}(A_i)\,;$$

et on peut généraliser cette formule à une famille infinie quitte à passer à la limite ensuite (grâce au théorème de continuité décroissante) ;
— si $(A_n)_{n\geqslant 0}$ est une suite d'évènements **décroissante**, alors d'après le théorème de continuité décroissante :

$$\mathrm{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} \mathrm{P}(A_n)\ ;$$

— si les évènements de la famille $(\overline{A_n})_{n\in\mathbb{N}}$ sont **incompatibles** deux à deux (attention, la propriété n'est pas sur la famille $(A_n)_{n\in\mathbb{N}}$, mais sur les complémentaires !), alors par passage au complémentaire et σ-additivité on a :

$$\mathrm{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = 1 - \mathrm{P}\left(\bigcup_{n=0}^{+\infty} \overline{A_n}\right) = 1 - \sum_{n=0}^{+\infty} \mathrm{P}(\overline{A_n}) = 1 - \sum_{n=0}^{+\infty}(1 - \mathrm{P}(A_n))\ ;$$

— si $(A_1, \dots, A_n)$ correspond naturellement à une suite d'étapes, dans la réalisation logique ou chronologique d'un évènement, alors d'après la formule des probabilités composées :

$$\mathrm{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathrm{P}(A_1)\mathrm{P}_{A_1}(A_2)\cdots \mathrm{P}_{A_1\cap\cdots\cap A_{n-1}}(A_n).$$

Il est très important de souligner qu'en général, **vous savez dès le début dans quel cas de figure vous êtes**, et je dirais même que vous avez modélisé le problème et choisi les évènements d'intersection de sorte à vous placer sciemment dans ces contextes favorables. Nous avons déjà commenté le cas des évènements incompatibles et indépendants dans les sections 2.1 et 2.2. Nous avons aussi mentionné, dans la section 2.1, le cas d'un évènement obtenu par succession d'étapes. De plus :

← page 8
← page 12

— on a naturellement une suite d'évènements **décroissante** lorsqu'on répète une certaine expérience un nombre éventuellement infini de fois, et qu'on cherche à étudier les évènements :

« l'expérience ne se termine jamais (sous condition) », ou :

« on observe systématiquement tel phénomène dans cette infinité d'expériences », etc. ;

il suffit en effet de prendre pour A_n :

A_n : « la condition requise pour achever l'expérience ne s'est jamais vérifiée en les n premières itérations »,

A_n : « on observe systématiquement tel phénomène dans les n premières itérations de l'expérience »,

où nous vous laissons vérifier la monotonie ;

— $\boxed{\text{on a naturellement une famille d'évènements } (\overline{A_n})_{n\in\mathbb{N}} \textbf{ incompatibles}}$ lorsqu'à chaque fois qu'on fait l'expérience, dans un couple d'évènements (A_{n_1}, A_{n_2}), il y a toujours au moins un des deux qui se réalise, par exemple :

> On répète une infinité de fois une expérience, et A_n est l'évènement :
> « il existe un résultat de cette suite d'expériences qui est supérieur ou égal au résultat de la n^{e} expérience » ;

On n'avance donc pas à l'aveugle et nous sommes méthodiques.

Ces conseils sont récapitulés dans la figure 4.

Figure 4 – Cas de la probabilité d'une intersection.

Indépendance

Si les A_i sont indépendants, alors : $\mathrm{P}\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} \mathrm{P}(A_i)$.
Cadre d'application. Comme l'indique l'hypothèse : il faut l'indépendance. Veuillez ne pas l'inventer si elle n'est pas vérifiée. Si l'intersection est infinie : d'abord écrire l'intersection finie, puis utiliser le théorème de continuité décroissante.
Exemple fréquent. Les évènements A_i sont de la forme $(X_i \in U_i)$ où les X_i sont des variables aléatoires indépendantes par hypothèse (résultats de différents dés ou pièces, par exemple).

Théorème de continuité décroissante (suite décroissante d'évènements)

Continuité décroissante : $\forall n \in \mathbb{N},\ A_{n+1} \subseteq A_n \Rightarrow \mathrm{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} \mathrm{P}(A_n)$.
Sans hypothèse : $\mathrm{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{N\to+\infty} \mathrm{P}\left(\bigcap_{n=0}^{N} A_n\right)$.
Cadre d'application. Si A_n se réalise, alors tous les précédents se sont aussi réalisés (ce sont des évènements de moins en moins probables).
Exemple fréquent. L'évènement A_n est l'observation systématique d'un succès en n itérations d'une expérience.

Passage au complémentaire (complémentaires incompatibles)

$\mathrm{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = 1 - \mathrm{P}\left(\overline{\bigcap_{n=0}^{+\infty} A_n}\right) = 1 - \mathrm{P}\left(\bigcup_{n=0}^{+\infty} \overline{A_n}\right)$.
Cadre d'application. On n'y gagne à passer aux intersections que si l'on sait que les évènements complémentaires $\overline{A_n}$ sont disjoints. Dans ce cas, on peut écrire : $\mathrm{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = 1 - \sum_{n=0}^{+\infty}(1 - \mathrm{P}(A_n))$.
Exemple fréquent. Quand on considère deux évènements de la famille $(A_n)_{n\in\mathbb{N}}$ il y en a toujours l'un des deux qui se réalise.

Formule des probabilités composées (succession logique ou chronologique d'étapes)

Probabilités composées : $\mathrm{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathrm{P}(A_1)\mathrm{P}_{A_1}(A_2)\cdots\mathrm{P}_{A_1\cap\cdots\cap A_{n-1}}(A_n)$.
Cadre d'application. Elle remplace la formule en cas de non indépendance. Elle n'est pas pertinente s'il n'y a pas une chronologie claire : les évènements doivent suivre un certain ordre de réalisation (logique ou chronologique).
Exemple fréquent. Si un évènement se décompose en n étapes (n tirages de boules sans remise, etc.).

Remarque. On pourrait se demander pourquoi je n'inclus pas le cas d'une famille (finie) et monotone d'évènements. En effet, si $A_n \subseteq A_{n-1} \subseteq \cdots \subseteq A_1$, alors $\bigcap\limits_{k=1}^{n} A_k = A_n$ (pour vous en convaincre, faites un dessin !), et donc : $\mathrm{P}\left(\bigcap\limits_{k=1}^{n} A_k\right) = \mathrm{P}(A_n)$. Résultat analogue si les inclusions sont dans l'autre ordre, en remplaçant $\mathrm{P}(A_n)$ par $\mathrm{P}(A_1)$. Cela semble bien plus simple que de calculer la probabilité d'une inclusion de n évènements !

La raison pour laquelle je ne mentionne pas le cas d'une famille finie et monotone, est que cette simplicité est trompeuse : il n'en est rien en général. En effet, lorsqu'on est dans ce type de configuration, l'égalité $\bigcap\limits_{k=1}^{n} A_k = A_n$ est souvent la traduction du fait que pour accomplir A_n, il faut passer par des étapes A_k. En regardant comment, un à un, on obtient A_1, puis A_2, etc. (en regardant la probabilité de réalisation à chaque étape), puis A_n, on se ramène à n calculs de probabilités plus simples (implicitement : on utilise la formule des probabilités composées). En regardant A_n directement au lieu de l'intersection, on perd cette fragmentation en étapes simples.

2.4 ✓ Synthèse : conseils pour entamer un exercice, et passer de l'intuition à la démonstration

En présence d'un exercice contenant un énoncé en français, et qu'on doit traduire mathématiquement pour en déduire des probabilités, je vous recommande de suivre les conseils suivants, qui sont éclairés par les conseils des sections précédentes.

1. Identifier, à l'aide de l'énoncé, les différents évènements et probabilités en jeu (ne pas oublier les conditionnements éventuels).
2. Écrire en terme d'intersections et d'unions les relations entre les évènements, si possible. Si l'énoncé s'y prête, identifier à l'aide de mots la variable aléatoire à considérer, et dans ce cas :
   (a) Identifier correctement la loi de probabilité qu'elle suit en essayant de reconnaître dans le problème une situation type (Bernoulli, binomiale, géométrique... d'abord préciser les valeurs possibles que peut prendre cette variable, pour exclure des possibilités), et déterminer les paramètres de la loi. Attention au fait que certaines variables aléatoires ne suivent des lois usuelles que si l'on rajoute un conditionnement (pour fixer un paramètre) : voir plus bas.
   (b) Utiliser les formules théoriques pour déterminer les probabilités demandées, en suivant les conseils des sections précédentes.

   Plus généralement :
3. À l'aide des méthodes données dans les sections 2.2 et 2.3, ramener les probabilités à calculer à celles connues, en repérant entre autres les schémas classiques suivants : ← page 12 ← page 14
   — une succession logique ou chronologique d'étapes simples ; (penser à la formule des probabilités composées)
   — des évènements de la forme $(X \in U)$ où X est une variable aléatoire (penser à la σ-additivité) ;
   — des évènements de la forme « avoir au moins un succès en une infinité d'expériences » ou « avoir toujours un succès en une infinité d'expériences » (penser au théorème de continuité monotone) ;
   — une variable aléatoire qui serait usuelle si on FIXAIT la valeur d'une autre variable aléatoire (penser à la formule des probabilités totales : voir section 3 et plus spécifiquement la section 3.1). → page 18

   Ce n'est pas exhaustif : référez-vous à toute la discussion qui précède.
4. Penser au besoin aux formules de dénombrement données dans la section 1, surtout en cas de loi uniforme. ← page 1

Notez bien que la formule des probabilités totales et la formule de Bayes n'ont pas été mentionnées dans les sections précédentes. Et pour cause : elles ne s'utilisent pas lorsque nous avons une réunion ou intersection d'évènements. Leur cadre d'application n'est donc pas le même :

— la formule des probabilités totales s'obtient pour passer de probabilités conditionnelles (« avec contraintes supplémentaires ») à une probabilité « tout court », et formalise le principe d'arbre de probabilités : les applications les plus fréquentes de cette formule sont dans la section 3, mais plus

généralement elle permet d'exprimer une probabilité $\mathrm{P}(B)$ **quand on sait quelle est la probabilité qu'elle découle de chaque cause possible** (l'ensemble des causes forme alors un système complet d'évènements) : on exprime $\mathrm{P}(B)$ en fonction des probabilités conditionnelles mieux connues $\mathrm{P}_{A_i}(B)$ où les A_i sont les différentes « causes » possibles ;
— la formule de Bayes s'applique plus particulièrement aux probabilités conditionnelles, et sert à inverser le conditionnement ; pour voir en quoi cela peut être pertinent, voir la section 4. → page 26

3 ✓ Deux applications fréquentes de la formule des probabilités totales

3.1 ✓ Lois marginales et lois conjointes

Les exercices qui suivent nous montrent comment gérer le cas où deux variables aléatoires X et Y sont en jeu, avec X et Y qui semblent suivre des lois usuelles, *mais avec les paramètres de la seconde qui dépendent de la première* (c'est en général le cas lorsque deux expériences se suivent, et que le résultat de la première influe sur le déroulement de la seconde). Dans ce cas, on ne peut pas dire que Y suit une loi usuelle : pour cela, il faudrait que ses paramètres soient des entiers FIXES.

On y remédie *via* conditionnement, afin de FIXER la valeur des paramètres ; c'est-à-dire, on calcule la loi de Y en supposant $(X=\ell)$ réalisé pour $\ell \in X(\Omega)$: on reconnaît alors en $\mathrm{P}_{(X=\ell)}(Y=k)$ une probabilité de variable aléatoire de loi usuelle, et on en déduit $\mathrm{P}(Y=k)$ grâce à la formule des probabilités totales :

$$\boxed{\mathrm{P}(Y=k) = \sum_{\ell \in X(\Omega)} \mathrm{P}(X=\ell)\mathrm{P}_{(X=\ell)}(Y=k).}$$

Il reste à calculer cette somme pour en déduire la loi de probabilité de Y.

Remarque. N'ENLEVEZ PAS LES CONDITIONNEMENTS SANS RAISON !

Nous illustrons ici l'approche sur les deux exemples les plus classiques rencontrés en exercices :
— une loi géométrique suivie d'une loi binomiale ;
— une loi de Poisson suivie d'une loi binomiale.

Exemple 10. (une loi géométrique suivie d'une loi binomiale) Un individu joue avec une pièce, qui a une probabilité $p \in]0,1[$ de tomber sur pile. Dans un premier temps, il lance la pièce jusqu'à obtenir pour la première fois pile. On note N le nombre de lancers nécessaires. Dans un deuxième temps, il lance N fois cette même pièce et on note X le nombre de piles obtenus au cours de cette seconde série de lancers.

La variable aléatoire N compte le nombre de répétitions d'une expérience (lancer une pièce) avant d'observer le premier succès (tomber sur pile). Comme la probabilité de tomber sur pile égale p à chaque lancer, on en déduit que N suit une loi géométrique de paramètre p. On a $N(\Omega) = \mathbb{N} \setminus \{0\}$, et :

$$\forall n \in \mathbb{N} \setminus \{0\}, \quad \mathrm{P}(N=n) = (1-p)^{n-1}p.$$

La variable aléatoire X, elle, ressemble à une loi binomiale de paramètres N et p, mais il n'en est rien : il faudrait pour cela que le nombre N de lancers soit FIXE, or il peut éventuellement varier à chaque fois qu'on reproduit l'expérience. On est donc dans le cas de figure ci-dessus : on y remédie en calculant les probabilités conditionnelles où l'on FIXE le paramètre N.

Soit $n \in \mathbb{N} \setminus \{0\}$. La variable aléatoire X, *sachant* $(N=n)$, compte le nombre de succès (tomber sur pile) dans une répétition de n épreuves de Bernoulli indépendantes (lancer une pièce). Par conséquent, la loi conditionnelle de X sachant $(N=n)$ est une loi binomiale, de paramètres n et p. On en déduit :

$$\forall k \in [\![0,n]\!], \quad \mathrm{P}_{(N=n)}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

N'ENLEVEZ PAS LE CONDITIONNEMENT !

À présent, en utilisant la formule des probabilités totales avec le système complet d'évènements $((N=n))_{n\in\mathbb{N}\setminus\{0\}}$, on obtient pour tout $k \in \mathbb{N}\setminus\{0\}$:

$$\mathrm{P}(X=k) = \sum_{n=1}^{+\infty} \underbrace{\mathrm{P}\left((N=n)\cap(X=k)\right)}_{=0 \text{ si } n<k} = \sum_{n=k}^{+\infty} \mathrm{P}(N=n)\mathrm{P}_{(N=n)}(X=k) = p^{k+1}(1-p)^{k-1}\sum_{n=k}^{+\infty}\binom{n}{k}(1-p)^{2(n-k)}.$$

On doit distinguer le cas $k=0$ du cas $k\neq 0$, parce que dans le cas où $k=0$, il n'y a aucun entier $n\in\mathbb{N}\setminus\{0\}$ tel que $n<k$, et par conséquent la somme reste $\sum_{n=1}^{+\infty}\mathrm{P}(N=n)\mathrm{P}_{(N=n)}(X=0)$, au lieu de devenir $\sum_{n=0}^{+\infty}\mathrm{P}(N=n)\mathrm{P}_{(N=n)}(X=0)$ pour $k=0$ (comme ce serait le cas ci-dessus, si l'on ne faisait pas attention).

Nous avons une première version (non simplifiée) de la loi de probabilité de X. Le calcul de cette somme est l'objet de l'exercice 5 de ce document. → page 20

Exemple 11. (une loi de Poisson suivie d'une loi binomiale) Un insecte pond des œufs, dont le nombre N suit une loi de Poisson de paramètre λ. Chaque œuf a une probabilité p d'éclore (et l'éclosion d'un œuf est indépendante de celle des autres). On note X la variable aléatoire égale au nombre d'insectes nés : on cherche la loi et l'espérance de X.

La variable aléatoire X ressemble à une loi binomiale de paramètres N et p, mais il n'en est rien : il faudrait pour cela que le nombre N d'œufs soit FIXE, or il peut éventuellement varier à chaque fois qu'on reproduit l'expérience. On est donc dans le cas de figure de la section : on y remédie en calculant les probabilités conditionnelles où l'on FIXE le paramètre N.

Soit $k\in\mathbb{N}$. La variable aléatoire X, *sachant* $(N=k)$, compte le nombre de succès (l'œuf éclôt) dans une répétition de k épreuves de Bernoulli indépendantes (lancer une pièce). Par conséquent, la loi conditionnelle de X sachant $(N=k)$ est une loi binomiale, de paramètres k et p. On en déduit :

$$\forall \ell\in[\![0,k]\!],\quad \mathrm{P}_{(N=k)}(X=\ell) = \binom{k}{\ell}p^{\ell}(1-p)^{k-\ell}.$$

N'ENLEVEZ PAS LE CONDITIONNEMENT !

À présent, en utilisant la formule des probabilités totales avec le système complet d'évènements $((N=k))_{k\in\mathbb{N}}$, on obtient :

$$\forall \ell\in\mathbb{N},\quad \mathrm{P}(X=\ell) = \sum_{k=0}^{+\infty}\underbrace{\mathrm{P}\left((N=k)\cap(X=\ell)\right)}_{=0 \text{ si } k<\ell} = \sum_{k=\ell}^{+\infty}\mathrm{P}(N=k)\mathrm{P}_{(N=k)}(X=\ell) = e^{-\lambda}p^{\ell}\sum_{k=\ell}^{+\infty}\frac{\lambda^k}{k!}\binom{k}{\ell}(1-p)^{k-\ell}.$$

Nous avons une première version (non simplifiée) de la loi de probabilité de X. Le calcul de cette somme est l'objet de l'exercice 6. → page 20

Cela nécessite de savoir manipuler certaines sommes, en général infinies. Les principales sommes utilisées sont ci-dessous.

3.1.1 Identités utiles

Vous aurez souvent besoin :

— en présence **d'une variable aléatoire suivant une loi uniforme**, des identités :

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2},\qquad \text{et : } \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6};$$

— en présence **d'une variable aléatoire suivant une loi géométrique**, des identités :

$$\begin{array}{rlcl} \forall x \in]-1,1[, & \displaystyle\sum_{n=0}^{+\infty} x^n & = & \dfrac{1}{1-x}, \qquad \forall x \in]-1,1[, \quad \displaystyle\sum_{n=1}^{+\infty} n x^{n-1} = \dfrac{1}{(1-x)^2}, \\ \forall x \in]-1,1[, & \displaystyle\sum_{n=2}^{+\infty} n(n-1)x^{n-2} & = & \dfrac{2}{(1-x)^3}, \qquad \dots \qquad \dots \\ \forall k \in \mathbb{N},\ \forall x \in]-1,1[, & \displaystyle\sum_{n=k}^{+\infty} \frac{n!}{(n-k)!} x^{n-k} & = & \dfrac{k!}{(1-x)^{k+1}}, \end{array}$$

obtenues par dérivation terme à terme du développement en série entière de $x \mapsto \dfrac{1}{1-x}$; on déduit de la dernière égalité la *formule du binôme négatif* (et c'est plutôt cette identité qui apparaît en probabilités) :

$$\forall k \in \mathbb{N},\ \forall x \in]-1,1[, \quad \sum_{n=k}^{+\infty} \binom{n}{k} x^n = \frac{x^k}{(1-x)^{k+1}};$$

— en présence **d'une variable aléatoire suivant une loi de Poisson**, de l'identité :

$$\forall x \in \mathbb{R}, \quad \sum_{n=0}^{+\infty} \frac{x^n}{n!} = e^x;$$

— en présence **d'une variable aléatoire suivant une loi binomiale**, de l'identité :

$$\forall n \in \mathbb{N},\ \forall (x,y) \in \mathbb{R}^2, \quad \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = (x+y)^n \tag{1}$$

et de celles obtenues par dérivation :

$$\begin{array}{lll} \forall n \in \mathbb{N} \setminus \{0\}, & \forall (x,y) \in \mathbb{R}^2, & \displaystyle\sum_{k=1}^{n} k \binom{n}{k} x^{k-1} y^{n-k} = n(x+y)^{n-1}, \\ \forall n \in \mathbb{N} \setminus \{0,1\}, & \forall (x,y) \in \mathbb{R}^2, & \displaystyle\sum_{k=2}^{n} k(k-1) \binom{n}{k} x^{k-2} y^{n-k} = n(n-1)(x+y)^{n-2}, \end{array}$$

etc.,

dont on déduit par exemple : $\displaystyle\sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = np$ (comment ? on l'utilise notamment pour le calcul de l'espérance d'une variable aléatoire suivant une loi binomiale de paramètres n et p).

Remarque. Le calcul de $\displaystyle\sum_{k=1}^{n} k \binom{n}{k} x^{k-1} y^{n-k}$, $\displaystyle\sum_{k=2}^{n} k(k-1) \binom{n}{k} x^{k-2} y^{n-k}$, etc., ne s'effectue pas nécessairement par dérivation. Il est en effet possible « d'éliminer » les facteurs en k à l'aide de la relation : $k\dbinom{n}{k} = n\dbinom{n-1}{k-1}$, valable pour tous entiers $n \geqslant 1$ et $k \geqslant 1$ (pourquoi ? savoir la démontrer, car ce n'est pas une formule du cours). Essayez de retrouver les identités ci-dessus grâce à cette égalité.

Pour les identités obtenues par dérivations et autres méthodes de calculs de sommes, revoir à profit le document *Méthodes* du chapitre sur les séries entières, section 2.2, *Calculs de sommes*.

Exercice 5. Reprendre l'exemple 10, et utiliser une des identités rappelées ci-dessus pour en déduire : ← page 18

$$\forall k \in \mathbb{N} \setminus \{0\}, \quad \mathrm{P}(X = k) = \frac{(1-p)^{k-1}}{(2-p)^{k+1}}, \text{ et } \mathrm{P}(X=0) = \frac{1-p}{2-p}.$$

Exercice 6. Reprendre l'exemple 11, et en déduire que X suit une loi de Poisson de paramètre λp. On pourra ← page 19

songer à un changement d'indice de sommation, qui fait apparaître une somme de série exponentielle.

Bien sûr, rien n'interdit d'intégrer au lieu de dériver, pour obtenir de nouvelles identités. Nous pouvons en avoir besoin, si l'on est conduit à calculer, par exemple : $\displaystyle\sum_{k=0}^{n} \frac{1}{k+1}\binom{n}{k} p^k(1-p)^{n-k}$ (où p serait une probabilité). La présence de $\dfrac{1}{k+1}$ incite à intégrer l'identité (1).

Exemple 12. (une loi binomiale suivie d'une loi de Bernoulli) Un tireur touche sa cible avec une probabilité $p \in]0,1[$ à chaque tir. Il tire n fois d'affilée, et on note N le nombre de tirs réussis (les tirs sont indépendants). Une fois cette série de n tirs achevée, il recule de N mètres (de sorte que sa probabilité de succès soit désormais de $\dfrac{p}{N+1}$), et tire une nouvelle fois. On note X la variable aléatoire valant 1 s'il réussit ce dernier tir, et 0 sinon. L'objectif est de déterminer la loi de probabilité de X.

Mettons en application la stratégie de cette section. La variable aléatoire N compte le nombre de succès (toucher sa cible) dans une série de n épreuves de Bernoulli de paramètre p indépendantes (tirer sur la cible), donc N suit une loi binomiale de paramètres n et p. On a donc $N(\Omega) = [\![0,n]\!]$, et : $\forall k \in [\![0,n]\!]$, $\mathrm{P}(N=k) = \dbinom{n}{k} p^k(1-p)^{n-k}$. De plus, si $(N=k)$ est réalisé, pour $k \in [\![0,n]\!]$, alors X sachant $(N=k)$ est une loi de Bernoulli de paramètre $\dfrac{p}{k+1}$. C'est-à-dire : $\mathrm{P}_{(N=k)}(X=1) = \dfrac{p}{k+1}$, et : $\mathrm{P}_{(N=k)}(X=0) = 1 - \dfrac{p}{k+1}$.

Alors, d'après la formule des probabilités totales :

$$\mathrm{P}(X=1) = \sum_{k=0}^{n} \mathrm{P}(N=k)\mathrm{P}_{(N=k)}(X=1) = \sum_{k=0}^{n} \binom{n}{k} p^{k+1}(1-p)^{n-k} \times \frac{1}{k+1}.$$

Cette somme ressemble à la formule du binôme de Newton rappelée en (1), avec $x = p$ et $y = 1-p$, mais il y a ce facteur $\dfrac{1}{k+1}$ en trop. On l'obtient en intégrant : $\forall (x,y) \in \mathbb{R}^2$, $(x+y)^n = \displaystyle\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$, entre 0 et $x \in \mathbb{R}$, et on obtient alors : ← page 20

$$\forall x \in \mathbb{R}, \quad \sum_{k=0}^{n} \binom{n}{k} y^{n-k} \int_0^x t^k \mathrm{d}t = \int_0^x (t+y)^n \mathrm{d}t,$$

et donc : $\forall x \in \mathbb{R}$, $\displaystyle\sum_{k=0}^{n} \binom{n}{k} \frac{x^{k+1}}{k+1} y^{n-k} = \frac{(x+y)^{n+1}}{n+1} - \frac{y^{n+1}}{n+1}$. En posant $x = p$ et $y = 1-p$, on conclut :

$$\sum_{k=0}^{n} \frac{1}{k+1}\binom{n}{k} p^{k+1}(1-p)^{n-k} = \frac{1-(1-p)^{n+1}}{n+1}.$$

Autrement dit : $\mathrm{P}(X=1) = \dfrac{1-(1-p)^{n+1}}{n+1}$, et : $\mathrm{P}(X=0) = 1 - \dfrac{1-(1-p)^{n+1}}{n+1}$. En conclusion :

$$X \sim \mathcal{B}\left(\frac{1-(1-p)^{n+1}}{n+1}\right).$$

On peut « s'amuser » en tentant les ving-cinq combinaisons possibles entre lois usuelles. Je marque d'un symbole « ☹ » les combinaisons peu intéressantes à mon avis. J'ai retiré des combinaisons qui conduisent à des sommes impossibles à calculer ou estimer à notre niveau.

Exercice 7. (une loi uniforme suivie d'une loi uniforme) ☹

1. Soient $n \in \mathbb{N} \setminus \{0\}$ et $k \in [\![1, n-1]\!]$. On choisit le tirage au hasard suivant d'un nombre entre 1 et n : d'abord, on décide si l'on choisira un nombre dans l'ensemble $A = [\![1,k]\!]$, ou dans l'ensemble $B = [\![k+1,n]\!]$; l'ensemble A a une probabilité k d'être choisi, et l'ensemble B a une probabilité $n-k$ d'être choisi ; on choisit ensuite au hasard (uniformément) un nombre dans l'ensemble retenu ; soit N le résultat de ce tirage. On note aussi X la variable aléatoire valant 0 si l'ensemble A est retenu, et valant 1 sinon.

(a) Donner la loi de X et, pour tout $k \in [\![0,1]\!]$, la loi conditionnelle de N sachant $(X = k)$.

(b) Donner la loi de probabilité de N. Interprétation ?

2. Soit $n \in \mathbb{N} \setminus \{0\}$. On tire au hasard un nombre entre 1 et n, et on note X le résultat obtenu. Ensuite, on tire au hasard (uniformément) un nombre entre 1 et X, et on note Y le résultat.

(a) Expliciter la loi de X et, pour tout $k \in [\![1, n]\!]$, la loi conditionnelle de Y sachant $(X = k)$.

(b) Donner la loi de probabilité de Y (il apparaît une somme qu'on ne peut pas simplifier).

(c) Prenons $n = 6$ (lancer d'un dé équilibré à six faces). Expliciter la loi de probabilité de Y dans ce cas.

(d) Soit $k \in [\![1, n]\!]$. Donner un équivalent asymptotique, quand $n \to +\infty$, de la probabilité de $(Y = k)$.

Exercice 8. (une loi uniforme suivie d'une loi de Bernoulli) Un joueur lance un dé équilibré à six faces deux fois de suite. On note X le résultat du premier lancer, et Y est la variable aléatoire valant 1 si le résultat du second lancer est supérieur ou égal à X, valant 0 sinon.

1. Donner la loi de X et, pour tout $k \in [\![1,6]\!]$, la loi conditionnelle de Y sachant $(X = k)$.
2. Montrer : $\mathrm{P}(Y = 1) = \dfrac{7}{12}$, et : $\mathrm{P}(Y = 0) = \dfrac{5}{12}$.
3. Changement subtil de règle, et gain à la clé : le joueur est susceptible de gagner X euros s'il s'arrête au terme du premier lancer. S'il décide de faire le second lancer, il gagne $2X$ euros en cas de résultat supérieur ou égal à celui du premier lancer, et 0 euros sinon. Calculer l'espérance du gain dans les deux cas : le joueur a-t-il intérêt à rejouer ? *(on aura besoin de calculer la probabilité de l'évènement $(XY = n)$; à quelle condition sur X et Y cette égalité est-elle possible ? penser à isoler le cas où $n = 0$)*

Exercice 9. (une loi uniforme suivie d'une loi binomiale) ☺

1. On lance un dé équilibré à six faces, et on note N le résultat obtenu. Ensuite, on lance N fois une pièce équilibrée. Soit X le nombre de « piles » obtenus à la fin de cette expérience.

(a) Expliquer pourquoi X ne suit pas une loi usuelle.

(b) Donner la loi de N et, pour tout $k \in [\![1,6]\!]$, la loi conditionnelle de X sachant $(N = k)$.

(c) Montrer :

$$\mathrm{P}(X = 0) = \frac{21}{128}, \quad \mathrm{P}(X = 1) = \frac{5}{16}, \quad \mathrm{P}(X = 2) = \frac{33}{128}, \quad \mathrm{P}(X = 3) = \frac{1}{6},$$

$$\mathrm{P}(X = 4) = \frac{29}{384}, \quad \mathrm{P}(X = 5) = \frac{1}{48}, \quad \mathrm{P}(X = 6) = \frac{1}{384}.$$

2. On lance un dé équilibré à six faces, et on note N le résultat obtenu. Puis on lance six fois ce même dé, et on note X le nombre de fois où, dans cette série de lancers, le résultat obtenu est supérieur ou égal à N.

(a) Pour tout $k \in [\![1,6]\!]$, donner la loi conditionnelle de X sachant $(N = k)$.

(b) Montrer :

$$\mathrm{P}(X = 0) = \frac{20515}{6^7}, \quad \mathrm{P}(X = 1) = \frac{6035}{6^6}, \quad \mathrm{P}(X = 2) = \frac{39885}{6^7}, \quad \mathrm{P}(X = 3) = \frac{40060}{6^7},$$

$$\mathrm{P}(X = 4) = \frac{39885}{6^7}, \quad \mathrm{P}(X = 5) = \frac{6035}{6^6}, \quad \mathrm{P}(X = 6) = \frac{67171}{6^7}.$$

Exercice 10. (une loi uniforme suivie d'une loi géométrique) Christian Renauld lance un dé équilibré à six faces et obtient un résultat N. Lionel Messire, dans son éternelle jalousie, cherche à l'égaler et lance ce même dé jusqu'à obtenir un résultat supérieur ou égal à N. On note X le nombre de lancers qui furent nécessaires pour que Lionel Messire trouve satisfaction. ☺

1. Pour tout $k \in [\![1,6]\!]$, donner la loi conditionnelle de X sachant $(N = k)$.
2. Montrer :
$$\forall n \in \mathbb{N} \setminus \{0\}, \quad \mathrm{P}(X = n) = \frac{1}{6}\sum_{k=1}^{6}\left(\frac{k-1}{6}\right)^{n-1}\left(1 - \frac{k-1}{6}\right),$$
et donner un équivalent asymptotique de $\mathrm{P}(X = n)$ quand $n \to +\infty$.

Exercice 11. (une loi uniforme suivie d'une loi de Poisson) Un client à la poste choisit au hasard un des guichets n° 1, n° 2 ou n° 3, avec équiprobabilité. Soit N la variable aléatoire donnant le numéro du guichet qu'il a choisi. ☹
Il s'avère que dans cette poste, la plupart des clients choisissent les guichets les plus proches : ainsi le nombre de clients au guichet n° k, pour $k \in \{1,2,3\}$, suit une loi de Poisson de paramètre $4 - k$.

On note X la variable aléatoire donnant le nombre de personnes patientant devant notre client.

1. Interpréter l'énoncé en termes de loi conditionnelle.
2. Montrer :
$$\forall n \in \mathbb{N}, \quad \mathrm{P}(X = n) = \frac{3^n e^{-3} + 2^n e^{-2} + e^{-1}}{3n!}.$$

Exercice 12. (une loi de Bernoulli suivie d'une loi uniforme) Un professeur veut envoyer un élève de sa classe de PSI (p garçons, $n-p$ filles, donc n élèves au total) au tableau. Tourmenté par l'idée d'envoyer aussi souvent des garçons et des filles, il choisit d'abord s'il enverra un garçon ou une fille avec équiprobabilité : on note X la variable aléatoire valant 0 s'il choisit d'envoyer un garçon, et 1 sinon. Il choisit ensuite au hasard (uniformément) un élève dans l'effectif retenu. On note Y l'élève envoyé (les garçons sont numérotés de 1 à p et les filles de $p + 1$ à n). ☹

1. Donner la loi de X et, pour tout $m \in [\![0,1]\!]$, la loi conditionnelle de Y sachant $(X = m)$.
2. Montrer :
$$\forall k \in [\![1, p]\!], \quad \mathrm{P}(Y = k) = \frac{1}{2p}, \quad \forall k \in [\![p + 1, n]\!], \quad \mathrm{P}(Y = k) = \frac{1}{2(n-p)}.$$

Exercice 13. (une loi de Bernoulli suivie d'une loi de Bernoulli) Un basketteur vise le panier à une distance raisonnable. Il a une probabilité p de réussir son lancer en faisant entrer la balle dans le panier (empochant les deux points). Soit X la variable aléatoire valant 0 s'il échoue dans ce lancer, et valant 1 sinon.

S'il réussit son premier lancer, il s'enhardit et recule : de la sorte, sa probabilité de succès est désormais de $p' < p$, et s'il réussit son lancer alors il empoche trois points au lieu de deux. S'il ne réussit pas son premier lancer, il reste à même distance et gagne toujours deux points en cas de succès. Soit Y la variable aléatoire valant 0 s'il échoue dans ce lancer, et valant 1 sinon.

1. Donner la loi de X et, pour tout $k \in [\![0,1]\!]$, la loi conditionnelle de Y sachant $(X = k)$.
2. Montrer : $\mathrm{P}(Y = 1) = p(p'+(1-p))$, et : $\mathrm{P}(Y = 0) = (1-p)^2+p(1-p')$. Ainsi : $Y \sim \mathcal{B}\left(p(p' + (1 - p))\right)$.
3. Soit G le nombre de points gagnés après ses deux lancers. Calculer l'espérance de G.

Exercice 14. (une loi de Bernoulli suivie d'une loi binomiale)

1. Lors d'une séance de tirs au but au football (il y a cinq tirs), l'Association Athlétique de Madré envoie en premier son joueur phare, Antoine Homgris. Il a une probabilité $\frac{9}{10}$ de réussir son tir. Les quatre tireurs suivants ont une probabilité $\frac{3}{4}$, chacun, de réussir leur tir, *sauf si Homgris rate le sien* : dans ce cas, leur confiance s'étiole et leur probabilité de succès descend à $\frac{1}{4}$. Il y a par contre indépendance entre ces quatre derniers tirs.

   On note X la variable aléatoire valant 1 si Homgris réussit son tir, 0 sinon, et on note N le nombre de tirs réussis par ses quatre coéquipiers. On remarque que $X + N$ donne le nombre de succès au total.

   (a) Donner la loi de X et, pour tout $k \in [\![0,1]\!]$, la loi conditionnelle de N sachant $(X = k)$.

(b) Montrer :

$$\forall n \in [\![0,4]\!], \quad \mathrm{P}(N=n) = \frac{27}{320}\frac{3^n + 3^{2-n}}{(4-n)!n!}.$$

(c) Donner la loi de probabilité de $X+N$, puis son espérance.

2. Dans une fête foraine, un stand appâte les riverains en proposant une partie gratuite d'un jeu de hasard. Seulement, pour les inciter à multiplier les parties, cette partie gratuite donne une probabilité $\frac{3}{4}$ de gagner. Les parties régulières (payantes) ont une probabilité $\frac{1}{4}$ de victoire.
Un joueur appâté décide, en cas de victoire à la partie offerte, de refaire une partie (payante) quatre fois de suite. En cas de défaite, il n'est guère motivé et ne rejoue que deux fois de suite.
On note X la variable aléatoire valant 1 si la partie gratuite est gagnée, 0 sinon, et N est la variable aléatoire donnant le nombre de parties gagnées ensuite. On note que $X+N$ donne le nombre de succès au total.

(a) Donner la loi de X et, pour tout $k \in [\![0,1]\!]$, la loi conditionnelle de N sachant $(X=k)$.

(b) Montrer :

$$\forall n \in [\![0,4]\!], \quad \mathrm{P}(N=n) = \frac{1}{4\cdot 3^n}\left(\frac{243}{256}\binom{4}{n} + \frac{9}{16}\binom{2}{n}\right).$$

(c) Donner la loi de probabilité de $X+N$.

(d) On suppose que chaque partie gagnée, y compris la partie offerte, a un gain de douze euros en cas de victoire. Chaque partie payante coûte cinq euros. On note G le gain : calculer son espérance.

Exercice 15. (une loi de Bernoulli suivie d'une loi de Poisson) Un client à la poste choisit au hasard un des guichets n° 1 ou n° 2. Puisque le guichet n° 1 est le plus proche de l'entrée, il le choisit avec probabilité $\frac{4}{5}$. ☹

Mais il n'est pas le seul à avoir ce réflexe : ainsi le nombre de clients au guichet n° 1 suit une loi de Poisson de paramètre 5 et celui au guichet n° 2 suit une loi de Poisson de paramètre 1.

On note X la variable aléatoire donnant le nombre de personnes patientant devant notre client.

1. Interpréter l'énoncé en termes de loi conditionnelle.
2. Montrer :

$$\forall n \in \mathbb{N}, \quad \mathrm{P}(X=n) = \frac{4\cdot 5^n e^{-5} + e^{-1}}{5n!}.$$

Exercice 16. (une loi de Poisson suivie d'une loi de Bernoulli) Chaque jour, le nombre N de courriers électroniques reçus par un certain professeur de mathématiques, de la part de ses élèves, suit une loi de Poisson de paramètre $\lambda > 0$.

Chaque fin de journée, il essaie de répondre tant bien que mal à tous les messages reçus, mais plus il en reçoit et moins il a de chance d'arriver à bout de sa besogne : il a une probabilité égale à $\dfrac{p}{N+1}$ de répondre à tous les messages, où $p \in]0,1]$. On note X la variable aléatoire valant 1 s'il arrive au bout de tous les messages, 0 sinon.

1. Pour tout $k \in \mathbb{N}$, donner la loi conditionnelle de X sachant $(N=k)$.
2. Montrer : $\mathrm{P}(X=1) = \dfrac{p(1-e^{-\lambda})}{\lambda}$, et : $\mathrm{P}(X=0) = 1 - \dfrac{p(1-e^{-\lambda})}{\lambda}$. Ainsi : $X \sim \mathcal{B}\left(\dfrac{p(1-e^{-\lambda})}{\lambda}\right)$.

3.2 ✓ Chaînes de Markov

Une chaîne de Markov apparaît dans un exercice lorsque nous reproduisons une *même* expérience une infinité de fois (ou un nombre n de fois, mais il faut alors quelques précautions), dans les mêmes conditions, où :

— lors de chaque expérience, les issues possibles sont toujours les mêmes : pour tout $n \in \mathbb{N}$, on note $A_{1,n}, \dots, A_{k,n}$ les k résultats possibles de la n^{e} expérience ; puisque nécessairement un de ces évènements se réalise, il s'agit d'un système complet d'évènements (en général $k \in \{2,3,4\}$) ;

— pour tout $n \in \mathbb{N}$, on connaît les probabilités des différentes issues de la $(n+1)^{\text{e}}$ expérience, *à la condition de connaître l'état à la n^{e} expérience* : cela s'interprète en termes de conditionnements par la connaissance explicite de $\mathrm{P}_{A_{j,n}}(A_{i,n+1})$ pour tous $i, j \in [\![1, k]\!]$;
— ces probabilités conditionnelles $\mathrm{P}_{A_{j,n}}(A_{i,n+1})$ ne dépendent pas de n : ce qu'il se passe est semblable à chaque réitération de l'expérience et les probabilités ne changent donc pas.

Si ces trois conditions sont réunies, alors on pose :

$$\forall n \in \mathbb{N}, \quad X_n = \begin{pmatrix} \mathrm{P}\left(A_{1,n}\right) \\ \vdots \\ \mathrm{P}\left(A_{k,n}\right) \end{pmatrix} \in \mathrm{M}_{k,1}(\mathbb{R}).$$

Pour tout $n \in \mathbb{N}$, la formule des probabilités totales, appliquée au système complet d'évènements $(A_{1,n}, \ldots, A_{k,n})$, permet d'écrire :

$$\forall i \in [\![1, k]\!], \quad \mathrm{P}\left(A_{i,n+1}\right) = \sum_{j=1}^{k} \mathrm{P}(A_{j,n}) \mathrm{P}_{A_{j,n}}\left(A_{i,n+1}\right),$$

ce qui s'écrit matriciellement : $\forall n \in \mathbb{N}$, $X_{n+1} = MX_n$, où la matrice $M = \Big(\big(\mathrm{P}_{A_{j,n}}(A_{i,n+1})\big)\Big)_{1 \leqslant i,j \leqslant k}$ est connue (et ne dépend pas de n : important, sinon ce qui suit est faux). Cela nous permet de déduire par récurrence immédiate que :

$$\forall n \in \mathbb{N}, \quad X_n = M^n X_0,$$

et l'exercice de probabilités se réduit à un calcul de puissances de matrice : on y parvient grâce à la méthode de division euclidienne de X^n par un polynôme annulateur de M, ou *via* réduction (ce qui est souvent préférable, si M est diagonalisable en particulier et c'est presque toujours le cas). Les méthodes ont été vues dans le chapitre de réduction des endomorphismes : revoir éventuellement son document *Méthodes*, section 4.

Le vecteur colonne X_0 est connu grâce à l'état initial de l'expérience : si on sait qu'on est à l'état $A_{i_0,0}$ en début d'expérience, alors X_0 est un vecteur n'ayant que des 0, sauf un 1 à la coordonnée numéro i_0. En multipliant M^n par X_0, on en déduit X_n, et donc $\mathrm{P}(A_{i,n})$ pour tout $i \in [\![1, k]\!]$ et tout $n \in \mathbb{N}$: on connaît les probabilités de chaque issue à chaque itération de l'expérience, ce qui permet en général d'en déduire tout ce qu'on veut dans le problème.

Astuce. La matrice M admet *toujours*, dans ce cas de figure, la valeur propre 1 (la somme des coefficients de chaque colonne égale 1). cela facilite le calcul du polynôme caractéristique : sommez toutes les lignes, vous obtiendrez déjà le facteur $(\lambda - 1)$.

Exercice 17. Démontrer cette affirmation.

Parfois, on vous demandera plutôt d'écrire la relation de récurrence sous la forme :

$$\begin{pmatrix} a_{n+1} & b_{n+1} & c_{n+1} \end{pmatrix} = \begin{pmatrix} a_n & b_n & c_n \end{pmatrix} M^\top.$$

Notez qu'il s'agit simplement de prendre la transposée de l'égalité ci-dessus. Le calcul des sous-espaces propres de $M^\top$ peut avoir l'avantage supplémentaire qu'on sait d'office que $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ est un vecteur propre de $M^\top$ (mais pas de M), du fait que la somme des coefficients de chaque ligne égale 1.

3.2.1 Cas de l'ordre 2

Si nous sommes dans le cas où le système complet d'évènements de l'énoncé est de la forme $(A_n, \overline{A_n})$ (pour tout $n \in \mathbb{N}$), et si on pose $b_n = \mathrm{P}(\overline{A_n})$, alors la même méthode que ci-dessus permet d'écrire une relation de la forme :

$$\begin{pmatrix} a_{n+1} \\ b_{n+1} \end{pmatrix} = M \begin{pmatrix} a_n \\ b_n \end{pmatrix}.$$

Donc, de même, calculer M^n permet d'en déduire a_n et b_n. Mais ici, en vérité, on peut faire BEAUCOUP PLUS SIMPLE ! En effet, si nous avons une relation de la forme : $a_{n+1} = \alpha a_n + \beta b_n$ (grâce à la formule des probabilités totales) alors, en utilisant le fait que $a_n + b_n = 1$, on en déduit :

$$\forall n \in \mathbb{N}, \quad a_{n+1} = (\alpha - \beta)a_n + \beta.$$

Ainsi $(a_n)_{n\geqslant 0}$ est une suite *arithmético-géométrique* ! Vous DEVEZ savoir déterminer l'expression générale d'une suite arithmético-géométrique, et en déduire ainsi a_n pour tout $n \in \mathbb{N}$: nul besoin de réduire une matrice !

Rappels : comment expliciter une suite arithmético-géométrique Pour déterminer l'ensemble des suites $(u_n)_{n\geqslant 0}$ à valeurs complexes et vérifiant :

$$\forall n \in \mathbb{N}, \quad u_{n+1} = au_n + b,$$

où $a \in \mathbb{C} \setminus \{1\}$ (si $a = 1$ alors c'est une suite arithmétique et on sait en déterminer la forme) et $b \in \mathbb{C}^*$, on procède ainsi :

1. On détermine le nombre $\lambda \in \mathbb{C}$ vérifiant : $\lambda = a\lambda + b$, vous trouvez : $\lambda = \dfrac{b}{1-a}$ (ne vous fatiguez pas à retenir son expression : résolvez l'équation à chaque fois).
2. On introduit la suite $(v_n)_{n\geqslant 0}$ définie par : $\forall n \in \mathbb{N}$, $v_n = u_n - \lambda$. On calcule v_{n+1} pour tout $n \in \mathbb{N}$; vous trouvez : $v_{n+1} = av_n$, et on en déduit que $(v_n)_{n\geqslant 0}$ est une suite géométrique de raison a, c'est-à-dire : $\forall n \in \mathbb{N}$, $v_n = a^n v_0$.
3. On conclut en écrivant :

$$\forall n \in \mathbb{N}, \quad u_n = v_n + \lambda = a^n v_0 + \frac{b}{1-a} \quad \left(= a^n u_0 + b \times \frac{1-a^n}{1-a}\right).$$

Pour retrouver non rigoureusement la forme d'une suite arithmético-géométrique Vous pouvez faire AU BROUILLON le raisonnement par réitération suivant :

$$\begin{aligned}
&u_n = au_{n-1} + b = a\left(au_{n-2} + b\right) + b = a^2 u_{n-2} + ab + b,\\
&u_n = a^2 u_{n-2} + b(a+1) = a^2\left(au_{n-3} + b\right) + b(a+1) = a^3 u_{n-3} + a^2 b + b(a+1) = a^3 u_{n-3} + b(a^2 + a + 1),\\
&\quad\vdots
\end{aligned}$$

jusqu'à obtenir : $u_n = a^n u_0 + b \times \sum\limits_{k=0}^{n-1} a^k$. Cette dernière somme est la somme des termes consécutifs d'une suite géométrique de raison a, et on sait donc que : $\sum\limits_{k=0}^{n-1} a^k = \frac{1-a^n}{1-a}$. On retrouve l'expression ci-dessus.

Exercice 18. Retrouver la forme ci-dessus des suites vérifiant $u_{n+1} = au_n + b$ pour tout $n \in \mathbb{N}$, grâce au principe :

$$\text{solutions d'une équation linéaire} = \text{solutions homogènes} + \text{solution particulière},$$

dont nous avons établi toute la généralité en algèbre linéaire.

Exercice 19. Retrouver la forme ci-dessus des suites vérifiant $u_{n+1} = au_n + b$ pour tout $n \in \mathbb{N}$, grâce à la technique du télescopage que votre serviteur illustre dans le document *Méthodes* sur les suites récurrentes, section 2.4.

4 ✓ Cadre d'application de la formule de Bayes

On rappelle la formule de Bayes :

$$\mathrm{P}_B(A) = \mathrm{P}_A(B) \times \frac{\mathrm{P}(A)}{\mathrm{P}(B)}.$$

Elle est extrêmement facile à démontrer : sachez le faire. Cela vous évitera de permuter accidentellement $\mathrm{P}(A)$ et $\mathrm{P}(B)$ dans la fraction (en particulier si les notations A et B sont échangées par rapport à la formule que vous avez apprise).

Contrairement à ce que les étudiants pensent souvent, il est insensé de penser que la formule de Bayes nous est utile dès que nous avons une probabilité conditionnelle à calculer. Après tout, aussi bien $\mathrm{P}_B(A)$ que $\mathrm{P}_A(B)$ sont des probabilités conditionnelles dans la formule ci-dessus : est-ce qu'il faudrait toutes les deux les calculer avec la formule de Bayes, vu qu'elles sont conditionnelles ? Mais on tournerait alors en rond !

N'oublions pas que la science cherche à RÉSOUDRE DES PROBLÈMES : les formules sont des moyens de les résoudre. Savoir quand utiliser la formule de Bayes revient donc à savoir quel problème elle tente de résoudre. Pour cela, on note qu'en général, lorsqu'on étudie une expérience aléatoire, les probabilités conditionnelles $\mathrm{P}_A(B)$ sont connues **lorsque B est une conséquence de A**. Je parle de conséquence CHRONOLOGIQUE ou LOGIQUE : par exemple, dans une expérience de dynamique de populations, si A est « avoir un enfant » et B est « avoir un garçon », alors B n'est bien sûr pas conséquence de A sur le plan chronologique, puisque A et B se produisent simultanément le cas échéant. Par contre c'est une conséquence sur le plan logique : pour qu'une famille ait des garçons, encore faut-il D'ABORD qu'elle ait des enfants. Ceci étant dit, $\mathrm{P}_A(B)$ est souvent une donnée de départ quand B est conséquence de A (obtenue par exemple par une analyse fréquentielle dans la vraie vie : on se place dans le cas de figure A, et on regarde alors combien de fois on observe B), mais $\mathrm{P}_B(A)$ est en général inconnu parce qu'on ne se pose pas la question dans cet ordre : il est rare qu'on observe une conséquence B sans connaître sa cause A.

En résumé : si B est une conséquence (logique ou chronologique) de A, alors $\mathrm{P}_A(B)$ est en général connu ou facile à obtenir, mais pas $\mathrm{P}_B(A)$. **C'est là que réside l'intérêt la formule de Bayes** : lorsque, étant donnée une conséquence B observée, nous voulons trouver la probabilité $\mathrm{P}_B(A)$ que ce soit dû à la cause A, cette formule permet d'inverser le conditionnement pour se ramener à un lien de causalité plus logique et mieux connu, en se ramenant à la probabilité $\mathrm{P}_A(B)$.

Pour songer à l'utiliser, en fin de compte : lorsqu'on vous demande de calculer une probabilité conditionnelle $\mathrm{P}_B(A)$ (guetter les « sachant que », « on admet que », « on sait que », pour reconnaître une probabilité conditionnelle) :

— interprétez toutes les données de l'énoncé à l'aide de probabilités conditionnelles, et voyez si $\mathrm{P}_A(B)$ y figure (ou $\mathrm{P}_A(\bar{B})$) ;

— voyez si $\mathrm{P}_A(B)$ peut s'interpréter comme la loi conditionnelle d'une variable aléatoire de loi usuelle (en général binomiale ou géométrique), ou du moins se calculer facilement.

Si aucun de ces deux points n'est vérifié, alors la formule de Bayes n'est d'aucun recours et il vous faut une autre approche. Mais le cas échéant, utilisez-la.

Calculer $\mathrm{P}(B)$ en pratique : formule des probabilités totales. Comme nous l'avons dit plus haut, en général dans une expérience nous connaissons non pas $\mathrm{P}(B)$, mais les probabilités d'avoir B sachant qu'une certaine cause A est observée. Autrement dit, on connaît plutôt $\mathrm{P}_A(B)$. On s'y ramène à l'aide de la formule des probabilités totales :

$$\mathrm{P}_B(A) = \mathrm{P}_A(B)\frac{\mathrm{P}(A)}{\mathrm{P}(A)\mathrm{P}_A(B) + \mathrm{P}(\bar{A})\mathrm{P}_{\bar{A}}(B)}$$

Ce recours est extrêmement fréquent. Attention néanmoins : il peut très bien y avoir plus que deux causes éventuelles ($\bar{A}$ et A) de B. C'est notamment le cas si la probabilité de B dépend d'un certain paramètre n qui peut varier si on refait l'expérience (autrement dit : B dépend d'une variable aléatoire). On verra souvent cet usage-là : si X est une variable aléatoire telle qu'on connaisse $\mathrm{P}_{(X=x)}(B)$ pour tout $x \in X(\Omega)$, alors d'après la formule des probabilités totales :

$$\mathrm{P}_B(A) = \mathrm{P}_A(B)\frac{\mathrm{P}(A)}{\sum\limits_{x\in X(\Omega)} \mathrm{P}(X = x)\mathrm{P}_{(X=x)}(B)}$$

(dans presque tous les cas rencontrés, A est aussi un évènement de la forme $(X = n)$).

Pour vous mettre la puce à l'oreille (« faut-il utiliser la formule des probabilités totales comme dans le premier cas, ou le second ? »), notez que si A est un évènement de la forme $(X = n)$ où X est une variable aléatoire, alors $\bar{A} = (X \neq n) = \bigcup_{y<n} (X = y) \cup \bigcup_{y>n} (X = y)$: c'est un ensemble abominable si X peut prendre beaucoup de valeurs différentes, et déterminer $\mathrm{P}_{\bar{A}}(B)$ est alors infaisable. Évitez donc la première formule si A est défini par une condition de la forme « telle grandeur égale n » (sauf si cette grandeur ne peut prendre que deux valeurs différentes : là, le complémentaire reste simple à décrire).

5 ✓ Penser à l'inégalité de Bienaymé-Tchebychev, et correctement l'utiliser

Le cadre d'application de l'inégalité de Bienaymé-Tchebychev : $\forall \varepsilon > 0$, $\mathrm{P}\left(|X - \mathrm{E}(X)| > \varepsilon\right) \leqslant \frac{\mathrm{V}(X)}{\varepsilon^2}$ qui donne, par passage au complémentaire :

$$\mathrm{P}\left(|X - \mathrm{E}(X)| \leqslant \varepsilon\right) \geqslant 1 - \frac{\mathrm{V}(X)}{\varepsilon^2} \tag{2}$$

est facile, puisqu'elle ne s'utilise que dans une configuration : lorsqu'on veut déterminer la probabilité qu'une variable aléatoire prenne des valeurs proches ou lointaines de sa valeur moyenne théorique (l'espérance). Plus précisément, l'inégalité de Bienaymé-Tchebychev sert en général à répondre à un énoncé du type :

« déterminer la probabilité que telle grandeur soit comprise entre m_1 et m_2 »,

ou :

« déterminer pour quelles valeurs (d'un certain paramètre) la probabilité, que telle grandeur soit comprise entre m_1 et m_2, est supérieure ou égale à p »,

Vous devez y penser dès que vous voyez un tel énoncé.

C'est une chose d'y penser, mais encore faut-il bien traduire l'énoncé pour reconnaître une probabilité de la forme : $\mathrm{P}\left(|X - \mathrm{E}(X)| \geqslant \varepsilon\right)$ (ou $\mathrm{P}\left(|X - \mathrm{E}(X)| < \varepsilon\right)$). Pour cela :

1. Vous appelez X la fameuse grandeur, dont on cherche la probabilité qu'elle soit comprise entre m_1 et m_2. Autrement dit, avec cette notation, on cherche : $\mathrm{P}(m_1 \leqslant X \leqslant m_2)$.
2. C'est une variable aléatoire (souvent de loi usuelle). Calculez son espérance $\mathrm{E}(X)$.
3. Ensuite, transformez l'encadrement : $(m_1 \leqslant X \leqslant m_2)$, en soustrayant $\mathrm{E}(X)$ à chaque membre de l'encadrement, pour en déduire : $\mathrm{P}(m_1 \leqslant X \leqslant m_2) = \mathrm{P}(m_1 - \mathrm{E}(X) \leqslant X - \mathrm{E}(X) \leqslant m_2 - \mathrm{E}(X))$. À ce stade-là, il y a deux cas possibles, plus ou moins favorables :
   (a) Il est quasiment systématique qu'en fait, m_1 et m_2 soient symétriques par rapport à l'espérance de X, c'est-à-dire : m_1 et m_2 sont de la forme $m_1 = \mathrm{E}(X) - \varepsilon$ et $m_2 = \mathrm{E}(X) + \varepsilon$. Dans ce cas on a directement :
   $$(m_1 \leqslant X \leqslant m_2) = (\mathrm{E}(X) - \varepsilon \leqslant X \leqslant \mathrm{E}(X) + \varepsilon) = (-\varepsilon \leqslant X - \mathrm{E}(X) \leqslant \varepsilon) = \left(|X - \mathrm{E}(X)| \leqslant \varepsilon\right),$$
   comme nous vous laissons le vérifier. Il reste alors à utiliser l'inégalité de Bienaymé-Tchebychev, ou plus précisément sa variante rappelée en (2), pour répondre à la question de l'exercice.
   (Notons que réciproquement, cette discussion vous donne un indice supplémentaire pour reconnaître le cadre d'application de l'inégalité de Bienaymé-Tchebychev : c'est lorsqu'on veut déterminer la probabilité que X soit « proche » (ou non) de sa valeur moyenne.)
   (b) Si on n'est pas dans le cas de figure ci-dessus, ce n'est pas grave : si l'on note m la plus petite valeur entre $|m_1 - \mathrm{E}(X)|$ et $|m_2 - \mathrm{E}(X)|$, alors l'encadrement $m_1 - \mathrm{E}(X) \leqslant X - \mathrm{E}(X) \leqslant m_2 - \mathrm{E}(X)$ est vérifié en particulier si : $-m \leqslant X - \mathrm{E}(X) \leqslant m$, donc :
   $$\begin{aligned}\left(|X - \mathrm{E}(X)| \leqslant m\right) = (-m \leqslant X - \mathrm{E}(X) \leqslant m) &\subseteq (m_1 - \mathrm{E}(X) \leqslant X - \mathrm{E}(X) \leqslant m_2 - \mathrm{E}(X)) \\ &= (m_1 \leqslant X \leqslant m_2),\end{aligned}$$
   et donc, par croissance d'une probabilité :
   $$\mathrm{P}\left(|X - \mathrm{E}(X)| \leqslant m\right) \leqslant \mathrm{P}(m_1 \leqslant X \leqslant m_2),$$
   ce qu'on sait minorer grâce à la variante de l'inégalité de Bienaymé-Tchebychev (2).

6 ✓ Contredire l'indépendance de deux variables aléatoires

Je ne parlerai pas du cas où l'on veut *montrer* l'indépendance de variables aléatoires : il n'y a là rien de spécialement ingénieux à faire, sinon utiliser la définition de l'indépendance... Je m'attarde seulement sur la marche à suivre lorsqu'il faut *infirmer* l'indépendance, parce que dans ce cas vous tombez régulièrement dans un écueil logique (contre lequel on vous met pourtant en garde dès la première semaine de cours de 1re année). À savoir : pour contredire un énoncé devant être universellement vrai (présence d'un « $\forall$ »), il suffit de trouver UN SEUL contre-exemple.

Dans le cas de l'indépendance de variables aléatoires X et Y, donc, pour contredire l'affirmation : « $\forall (x,y) \in X(\Omega) \times Y(\Omega)$, $\mathrm{P}(X=x, Y=y) = \mathrm{P}(X=x)\mathrm{P}(Y=y)$ », c'est une TRÈS MAUVAISE approche de tenter de montrer : « $\forall (x,y) \in X(\Omega) \times Y(\Omega)$, $\mathrm{P}(X=x, Y=y) \neq \mathrm{P}(X=x)\mathrm{P}(Y=y)$ ». Vous n'y parviendrez pas, ou seulement avec beaucoup de peine, si les lois de X et Y font intervenir des quantités compliquées (coefficients binomiaux en général). Pire : ce peut être faux.

Ce que vous devez faire. Vous devez trouver UN choix PARTICULIER de x et y tels que l'on ait : $\mathrm{P}(X=x, Y=y) \neq \mathrm{P}(X=x)\mathrm{P}(Y=y)$. Puisque vous avez le choix, vous pouvez prendre des valeurs de x et y qui facilitent vos calculs (par exemple : x ou y égal à 0, ou 1, etc.).

Mieux encore : quand c'est possible, prenez des valeurs de x et y tels que $(X=x, Y=y)$ soit CLAIREMENT un évènement impossible, tandis que $(X=x)$ et $(Y=y)$ ne le sont pas. Dans ce cas, l'égalité $\mathrm{P}(X=x, Y=y) = \mathrm{P}(X=x)\mathrm{P}(Y=y)$ est nécessairement impossible, puisque le membre de gauche est nul et pas celui de membre de droite : vous n'avez même pas besoin de vous fatiguer à simplifier le produit pour obtenir le résultat souhaité !

Exemple 13. On lance une pièce équilibrée une infinité de fois, et on appelle X le rang du premier pile, Y le rang du second pile. Montrons que X et Y ne sont pas indépendantes (chose intuitive) : pour cela, on note qu'on a nécessairement $X < Y$, donc $(X=x, Y=y)$ est impossible dès qu'on prend $x \geqslant y$. Prenons par exemple $x = y = 2$ pour faciliter les calculs. On a $\mathrm{P}(X=2, Y=2) = 0$ pour la raison qu'on vient de donner. Or $X \sim \mathcal{G}\left(\frac{1}{2}\right)$, donc $\mathrm{P}(X=2) = \frac{1}{2^2} \neq 0$, et un calcul facile (que nous vous laissons effectuer...) montre que : $\mathrm{P}(Y=2) = \frac{1}{2^2} \neq 0$, donc $\mathrm{P}(X=2)\mathrm{P}(Y=2) \neq 0 = \mathrm{P}(X=2, Y=2)$ et on en déduit que X et Y ne sont pas indépendantes.

Demandez-vous pourquoi je n'ai pas pris $x = y = 1$ pour encore plus simplifier les calculs.

Si l'on n'est pas dans le cas très facile où l'évènement $(X=x, Y=y)$ est impossible pour de bonnes valeurs de x et y, vous aurez en général à calculer $\mathrm{P}(X=x, Y=y)$ *via* un conditionnement (vous ne pouvez pas utiliser l'indépendance de X et Y, vu que vous voulez justement la contredire). C'est-à-dire, en écrivant :

$$\mathrm{P}(X=x, Y=y) = \mathrm{P}(X=x)\mathrm{P}_{(X=x)}(Y=y), \quad \text{ou :} \quad \mathrm{P}(X=x, Y=y) = \mathrm{P}(Y=y)\mathrm{P}_{(Y=y)}(X=x).$$

S'il y a davantage que deux variables aléatoires, vous utilisez la formule des probabilités composées.

Le choix du conditionnement dépend de la situation : l'ordre logique ou chronologique de l'expérience permet tantôt de calculer plus facilement la loi conditionnelle de Y (sachant que $X = \cdots$), tantôt l'inverse.

Exemple 14. On lance une pièce équilibrée une infinité de fois, et on appelle X le rang du premier pile à être suivi d'un autre pile, tandis que Y est le rang du premier pile à être suivi de *deux* autres piles. Montrons que X et Y ne sont pas indépendantes (pourquoi est-ce intuitif ?). Pour cela, on note qu'on a facilement : $\mathrm{P}(X=1, Y=1) = \mathrm{P}(X=1)\mathrm{P}_{(X=1)}(Y=1) = \frac{1}{2^2} \times \frac{1}{2} = \frac{1}{2^3}$, tandis que : $\mathrm{P}(X=1) = \frac{1}{2^2}$, et : $\mathrm{P}(Y=1) = \frac{1}{2^3}$. Ainsi : $\mathrm{P}(X=1)\mathrm{P}(Y=1) = \frac{1}{2^5} \neq \mathrm{P}(X=1, Y=1)$, donc X et Y ne sont pas indépendantes.

Exercice 20. Vérifier les calculs de probabilité qui ont été omis dans cet exemple. On peut procéder autrement, si l'on remarque que : $(X=1, Y=1) = (Y=1)$.

7 ✓ Fonctions génératrices : cas où elles prévalent

Rappelons ce qu'est la fonction génératrice d'une variable aléatoire X à valeurs dans $\mathbb{N}$:

$$\forall t \in [-1,1], \quad G_X(t) = \sum_{n=0}^{+\infty} \mathrm{P}(X = n)t^n.$$

Si on la connaît explicitement, alors les coefficients de son développement en série entière donnent la loi de X, tandis que ses dérivées successives en 1 permettent de calculer l'espérance et la variance de X.

En soi, dès qu'on connaît la loi d'une variable aléatoires à valeurs dans $\mathbb{N}$, on pourrait en toutes circonstances étudier sa fonction génératrice pour en déduire son espérance et sa variance. Mais ce n'est pas forcément plus intéressant que par application directe des définitions.

Le cas où, par contre, la fonction génératrice apporte une réelle valeur ajoutée, **est quand nous étudions une somme de variables aléatoires INDÉPENDANTES** (attention, cette hypothèse est essentielle). Dans ce cas songez-y, même si l'énoncé ne vous y invite pas. L'intérêt est que dans ce cas, la fonction génératrice est facile à calculer : si X et Y sont INDÉPENDANTES, alors :

$$\forall t \in [-1,1], \quad G_{X+Y}(t) = G_X(t) \cdot G_Y(t).$$

Si l'on connaît les lois de X et Y, on en déduit leurs fonctions génératrices G_X et G_Y (c'est direct si X et Y suivent des lois usuelles), et donc on en déduit leur produit G_{X+Y} (ne pas passer par la définition du produit de Cauchy : ce n'est pas assez explicite pour ce qu'on veut en général).

La connaissance de G_{X+Y} permet d'en déduire ensuite **la loi de probabilité de** $Z = X + Y$, en développant en série entière en 0 l'application $t \mapsto G_Z(t)$ qu'on a explicitée grâce à la formule ci-dessus ; si on parvient à l'écrire sous la forme : $G_Z(t) = \sum\limits_{n=0}^{+\infty} a_n t^n$ (pour tout t dans un voisinage de 0), où les a_n sont CONNUS, alors du fait que par définition on ait aussi : $\forall t \in [-1,1]$, $G_Z(t) = \sum\limits_{n=0}^{+\infty} \mathrm{P}(Z = n)t^n$, on en déduit par unicité des coefficients : $\forall n \in \mathbb{N}$, $a_n = \mathrm{P}(Z = n)$.

> Si vous avez des facteurs devant la somme de série entière, rentrez-les dans la somme : il faut les compter dans l'expression du coefficient devant t^n ! Si $G_Z(t) = \lambda \sum\limits_{n=0}^{+\infty} b_n t^n$, alors il est faux d'en déduire qu'on a $b_n = \mathrm{P}(Z = n)$ pour tout $n \in \mathbb{N}$: on a $\lambda b_n = \mathrm{P}(Z = n)$ pour tout $n \in \mathbb{N}$, car $G_Z(t) = \sum\limits_{n=0}^{+\infty} \lambda b_n t^n = \sum\limits_{n=0}^{+\infty} \mathrm{P}(Z = n)t^n$.

Je ne parlerai pas tant que cela de l'espérance ni la variance, bien qu'on connaisse les relations :

$$\mathrm{E}(Z) = G_Z'(1), \text{ et : } \mathrm{V}(Z) = G_Z''(1) + G_Z'(1) - (G_Z'(1))^2,$$

puisque *dans le cas précis d'une somme de variables aléatoires*, on la trouve plus directement par linéarité pour l'espérance (nul besoin d'indépendance), et dans le cas de la variance on a $\mathrm{V}(X+Y) = \mathrm{V}(X) + \mathrm{V}(Y)$ si X et Y sont indépendantes : nul besoin de fonction génératrice.

Cette méthode d'utilisation de la fonction génératrice se généralise au cas d'une somme de n variables aléatoires indépendantes. Plus il y a de variables aléatoires, et plus elle devient rentable.

Attention, si vous ne mentionnez pas l'indépendance au moment d'utiliser cette formule, vous faites un oubli crucial et le raisonnement est faux : vous perdez des points. Ce n'est pas une bête hypothèse technique.

Pour ensuite développer en série entière en 0 la fonction génératrice G_Z, songez aux méthodes répertoriées dans le document *Méthodes* du chapitre sur les séries entières, section 3 (*Développer en série entière*).

Exemple 15. Soient X et Y deux variables aléatoires INDÉPENDANTES. On suppose : $X \sim \mathcal{G}(p)$, $Y \sim \mathcal{G}(q)$, avec p et q *différents*. On cherche la loi de probabilité de $Z = X + Y$.

Le fait d'avoir une somme, et l'hypothèse d'indépendance, incite immédiatement à utiliser la fonction génératrice, même si l'énoncé ne le suggère pas. On a, sous ces hypothèses :

$$\forall t \in [-1,1], \quad G_Z(t) = G_X(t)G_Y(t) = \frac{pt}{1-t(1-p)} \cdot \frac{qt}{1-t(1-q)}.$$

On développe cette fonction en série entière en 0. Pour cela, on décompose en éléments simples, et on trouve :

$$\begin{aligned}\forall t \in [-1,1], \quad G_Z(t) &= \frac{pqt^2}{q-p}\left(\frac{1-p}{1-t(1-p)} - \frac{1-q}{1-t(1-q)}\right)\\ &= \frac{pqt^2}{q-p}\left((1-p)\sum_{n=0}^{+\infty}(1-p)^n t^n - (1-q)\sum_{n=0}^{+\infty}(1-q)^n t^n\right)\\ &= \sum_{n=0}^{+\infty}\frac{pq}{q-p}\left((1-p)^{n+1} - (1-q)^{n+1}\right)t^{n+2}\\ &= \sum_{n=2}^{+\infty}\frac{pq}{q-p}\left((1-p)^{n-1} - (1-q)^{n-1}\right)t^{n}.\end{aligned}$$

Or : $\forall t \in [-1,1]$, $G_Z(t) = \sum_{n=0}^{+\infty} \mathrm{P}(Z=n)t^n$. L'unicité des coefficients d'une somme de série entière implique donc :

$$\forall n \in \mathbb{N} \setminus \{0,1\}, \quad \mathrm{P}(Z=n) = \frac{pq}{q-p}\left((1-p)^{n-1} - (1-q)^{n-1}\right),$$

tandis que $\mathrm{P}(Z=0) = \mathrm{P}(Z=1) = 0$ (car le coefficient de t^n, pour $n=0$ et $n=1$, est nul dans le développement en série entière explicite de G_Z). On a obtenu la loi de probabilité de Z.

Exercice 21. Trouver la variance de Z à l'aide de la fonction génératrice (nous vous conseillons de décomposer en éléments simples G_Z pour faciliter le calcul des dérivées, en tenant compte du facteur t^2 au numérateur).

Exercice 22. S'inspirer de ce qui précède pour donner la loi de probabilité de la somme de *trois* variables aléatoires indépendantes suivant des lois géométriques de paramètres tous différents.

Exercice 23.

1. Reprendre l'exemple ci-dessus, mais dans le cas où $p = q$.
2. Généraliser à une somme de n variables aléatoires indépendantes, suivant la même loi géométrique de paramètre p.

Table des matières

Table des figures