Probabilités

Table des matières

1 Combinatoire

1.1 Choix

Soit $E = \{e_1, e_2, \cdots, e_n\}$ un ensemble fini de cardinal n. Alors il existe n façons distinctes de choisir un élément e de E. Le comptage du nombre de choix d'un élément dans un ensemble revient donc à la détermination du cardinal de cet ensemble. A cette fin, il est bon de se souvenir que l'ensemble $A \times B = \{(a, b) | a \in A,\ b \in B\}$ est de cardinal $|A| \times |B|$. En particulier, si r est un entier, le cardinal de E^r est $|E|^r$. Il existe donc n^p façons différentes de choisir p fois de suite un élément dans un ensemble ayant n éléments.

Exemples : Il existe $7 \times 12 = 84$ choix d'un jour de la semaine et d'un mois de l'année. Il existe $6^3 = 216$ résultats possibles lorsque l'on tire 3 fois de suite un dé usuel (à six faces). A l'heure actuelle, il existe 14 candidats déclarés à l'élection présidentielle de 2012. Sauf en cas de victoire au premier tour, deux de ces candidats s'affronteront au second tour. Il existe donc $14 \times 2 = 28$ possibilités distinctes (théoriques) de vote en cas d'une élection à deux tours.

1.2 Les fonctions cruciales du dénombrement

1.2.1 Ordonner un ensemble à n éléments

Soit $E = \{e_1, e_2, \cdots, e_n\}$ un ensemble fini de cardinal n. Un ordre sur E est le choix d'une énumération des éléments de E. Par définition, il existe factoriel n ordre sur E, ce que l'on note $n!$. L'entier $n!$ est égal à $n \times (n-1) \times (n-2) \times \cdots \times 3 \times 2 \times 1$. Ceci se démontre par récurrence de la manière suivante. Pour ordonner un E, il suffit de choisir un élément e que l'on place en première place. Ceci peut se faire de n façons distinctes. Ensuite, il reste à ordonner l'ensemble $E - \{e\}$, qui est de cardinal $n-1$. Il existe donc $(n-1)!$ façons de procéder. Il y a donc $n \times (n-1)!$ façons d'ordonner E. Donc $n! = n(n-1)\cdots 2 \cdot 1$.

1.2.2 Choisir p éléments parmi n

Soit $E = \{e_1, e_2, \cdots, e_n\}$ un ensemble fini de cardinal n et soit p un entier. Un p-arrangement $(e_{i_j})_{1 \leq j \leq p}$ de E est la donnée de p éléments distincts de E dans un ordre donné (donc (e_1, e_2) et (e_2, e_1) ne sont pas le même 2-arrangement). On désigne le nombre de p-arrangements distincts d'éléments de E par A_n^p.

Pour construire un p-arrangement, il faut choisir un élément e de E, que l'on désigne comme le premier élément de l'arrangement, puis choisir un $(p-1)$-arrangement de $E - \{e\}$. Réciproquement, tous les p-arrangements sont construits de cette manière. Les p-arrangements vérifient donc la relation $A_n^p = nA_{n-1}^{p-1}$. En répétant cette relation, on obtient :

$$A_n^p = n(n-1)A_{n-2}^{p-2} = n(n-1)\cdots(n-p+2)A_{n-p+1}^{p-p+1} = n(n-1)\cdots(n-p+2)(n-p+1)$$

Notons que, un peu par convention et aussi un peu pour des raisons de cohérence interne, on considère qu'il existe un unique 0-arrangement (et non pas 0) dans un ensemble à n éléments lorsque $n > 0$. Donc $A_n^0 = 1$.

De manière alternative, et quasiment équivalente, on peut remarquer qu'un p-arrangement se prolonge en un ordre sur les n éléments de E de $(n-p)!$ façons distinctes et que tous les ordres sur E sont construits de cette manière. Donc :

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

1.2.3 Choisir p éléments parmi n, sans se préoccuper de l'ordre

Soit $E = \{e_1, e_2, \cdots, e_n\}$ un ensemble fini de cardinal n et soit p un entier. Un p-uplet $(e_{i_j})_{1 \leq j \leq p}$ de E est la donnée de p éléments distincts de E sans ordre prescrit (donc (e_1, e_2) et (e_2, e_1) sont le même 2-uplet). On désigne par $\binom{n}{p}$ le nombre de façons distinctes de choisir des p-uplet dans E. Il est bon de remarquer qu'il existe une façon unique de choisir un 0-uplet dans E, car la seule possibilité est de prendre le p-uplet vide, tandis qu'il n'existe aucune façon de choisir un p-uplet si $p < 0$. Ainsi, $\binom{n}{0} = 1$ et $\binom{n}{-p} = 0$ si $p > 0$.

Pour construire un p-uplet, il suffit de construire un p-arrangement. Par définition de $p!$, il existe $p!$ p-arrangements distincts qui produisent par cette construction le même p-uplet. Donc :

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

Remarquons que choisir un p-uplet b dans E est la même chose que de choisir le $(n-p)$-uplet des éléments que l'on n'a pas choisi dans b. Donc $\binom{n}{p} = \binom{n}{n-p}$. Cela est également manifeste à partir de la dernière expression de l'équation précédente.

1.3 Quelques propriétés avancées de combinatoires

1.3.1 Interprétation polynomiale

Considérons le polynôme $(x+y)^n = (x+y)(x+y)\cdots(x+y)$. Lorsque que l'on développe ce polynôme, on doit pour chacune des n parenthèses choisir x ou y. Le terme de degré x^k correspond donc au nombre de façons distinctes de choisir k fois x parmi n possibilités, et donc par définition à $\binom{n}{k}$. Donc :

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Cette interprétation permet de démontrer certaines propriétés des $\binom{n}{k}$. Par exemple, remplacer x et y par 1 montre que :

$$(1+1)^n = 2^n = \sum_{k=0}^{n} \binom{n}{k}$$

Soit s un entier. L'égalité $(x+y)^n = (x+y)^{n-s}(x+y)^s$ montre que :

$$\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = \sum_{i=0}^{s} \binom{s}{i} x^i y^{s-i} \sum_{j=0}^{n-s} \binom{n-s}{j} x^j y^{n-s-j}$$

En identifiant le terme en $x^k y^{n-k}$ de chaque côté de l'équation précédente, on obtient :

$$\binom{n}{k} = \sum_{i=0}^{k} \binom{s}{i} \binom{n-s}{k-i}$$

Ou encore :

$$\binom{n}{k} = \sum_{i+j=k} \binom{s}{i} \binom{n-s}{j}$$

1.3.2 Coefficients multinomiaux

On peut généraliser la notion de p-uplet en la notion de $(p_1, \cdots, p_s)$-multiplet. Étant donné des entiers $(p_i)_{1\leq i\leq s}$ vérifiant $p_1 + p_2 + \cdots + p_s = n$, un $(p_1, \cdots, p_s)$-multiplet est la donnée de p_1 éléments distincts, puis de p_2 éléments distincts parmi ceux qui restent, et ainsi de suite. Le nombre de $(p_1, \cdots, p_s)$-multiplets est par définition le nombre multinomial $\begin{pmatrix} n \\ p_1, \cdots, p_s \end{pmatrix}$. Les nombres multinomiaux vérifient la relation :

$$\begin{pmatrix} n \\ p_1, \cdots, p_s \end{pmatrix} = \frac{n!}{p_1!p_2!\cdots p_s!}$$

1.3.3 Triangle de Pascal

Pour construire un p-uplet d'éléments de E de cardinal n, il suffit de choisir tout d'abord si ce p-uplet contiendra ou non le dernier élément de E. Si oui, alors il reste à choisir un $(p-1)$-uplet de $E-\{e_n\}$ de cardinal $n-1$. Sinon, alors notre p-uplet est un p-uplet de $E-\{e_n\}$. Réciproquement, tout p-uplet d'éléments de E est construit de la manière précédente. Donc $\binom{n}{p} = \binom{n-1}{p-1} + \binom{n}{p-1}$. Si l'on dispose les coefficients binomiaux en un triangle dont les lignes correspondent à n constant et les colonnes à p constant, alors la somme de deux termes consécutifs d'un ligne est égale au terme de la ligne suivante en dessous du deuxième terme. Le dessin que l'on obtient est appelé le triangle de Pascal (même s'il est apparaît déjà dans les travaux de mathématiciens indiens du X° siècle).

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1
1 7 21 35 35 21 7 1
1 8 28 56 70 56 28 8 1
1 9 36 84 126 126 84 36 9 1
1 10 45 120 210 252 210 120 45 10 1

1.3.4 Rangement de p éléments indistincts dans n cellules

Supposons que l'on dispose de n espaces et que l'on souhaite y ranger p éléments identiques. Voici une façon de procéder pour obtenir le nombre possible de tels rangements. On peut imaginer que le rangement se fasse de la manière suivante : on compte à partir de 1 le nombre d'éléments que l'on place dans la première cellule ; à tout moment, on peut crier stop et on s'arrête puis on recommence le comptage des éléments dans la deuxième cellule, et l'on continue jusqu'à atteindre p. Afin que toutes les cellules soient remplies (même de zéro élément), il faut crier $n-1$ fois stop. Pour que tous les éléments soient placés, il faut compter de 1 à p. La personne qui décrit le rangement de cette manière prononce donc $n-1+p$ mots. Le rangement est défini par la position des $n-1$ stops. Donc le nombre de tels rangements est $\binom{n-1+p}{n-1} = \binom{n+p-1}{p}$.

1.4 Exercices

Un dé usuel comporte 6 faces numérotées de 1 à 6. Un jeu de cartes usuel contient 52 cartes divisées en 13 cartes numérotées de 1 à 13 de 4 couleurs.

Exercice 1 De combien de façon différente peut-on obtenir 7 en lançant deux dés ? De combien de façons différentes peut-on tirer une carte dans un jeu de cartes ? Deux cartes ? Cinq cartes ? Combien existe-t-il de tirages de cinq cartes où toutes les cartes sont de la même couleur ? Combien existe-t-il de tirages de cinq cartes où deux cartes (au moins) parmi les cinq ont la même valeur ?

Correction La valeur du premier dé peut être choisi parmi 6. Une fois connue la valeur du premier dé, la valeur du second dé est imposée. Il y a donc 6 tirages possibles dont la valeur est 7. Il existe 52 tirages d'une carte dans un jeu. Il existe $\binom{52}{2}$ tirages de deux cartes et $\binom{52}{5}$ tirages de cinq cartes. Un tirage de cinq cartes de la même couleur correspond à un choix d'une des cartes couleurs puis au choix de 5 cartes parmi les 13 cartes de cette couleur. Il existe donc $4\binom{13}{5}$ tirages possibles de cinq cartes d'une même couleur. Pour dénombrer les tirages contenant au moins une paire, il est plus aisé de compter les tirages ne comportant pas de paires. Dans ce cas, il faut choisir 5 valeurs parmi les 13 possibles. Ensuite, il faut choisir l'une des 4 couleurs pour chacune des 5 cartes. Donc le nombre de tirages sans paire est $4^5\binom{13}{5}$. Donc le nombre de tirage dont deux cartes au moins ont la même valeur est $\binom{52}{5} - 4^5\binom{13}{5}$.

Exercice 2 Un wagon de RER comporte 30 places assises. 18 voyageurs montent et s'asseyent tous. Combien de répartition différentes des places occupées existe-t-il ? Si l'on suppose que les voyageurs sont tous distincts, de combien de façons différentes peuvent-ils s'asseoir ?

Correction Une répartition des places occupées est un choix de 18 places parmi les 30. Il en existe donc $\binom{30}{18}$. Si de plus on supposer que les voyageurs sont tous distincts, une répartition des voyageurs est obtenue en choisissant une répartition des places assises et un ordre sur les voyageurs. Il y a donc $18!\binom{30}{18}$ différentes répartitions.

Exercice 3 Combien d'anagrammes du mot PROBABILITÉS peut on former ? Combien d'anagrammes peut-on former avec 11 des 12 lettres du mot PROBABILITÉS ?

Correction Si l'on se souvient des coefficients multinomiaux, on remarque qu'une[1] anagramme de PROBABILITÉS correspond au choix de la position des deux i parmi les 12 possibles, puis des deux b parmi les 10 restantes, puis du p, du r, du o, du a, du l, du t, du e et du s. Il y a donc $\binom{12}{2,2,1,1,1,1,1,1,1,1}$ anagrammes possibles. Alternativement, on peut remarquer qu'il existe 12! ordres possibles sur les lettres de PROBABILITÉS, mais que la permutation des deux b et des deux i produit la même anagramme. Parmi ces 119750400 anagrammes, mon ordinateur m'assure qu'il

[1] Ce n'est pas une faute de frappe

n'existe que deux mots appartenant à la langue française : probabilités et probabiliste. Pour former une anagramme avec 11 des douze lettres, il faut commencer par choisir la lettre parmi les 10 que l'on va supprimer. Il y a donc 10 choix possibles. Si l'on choisit l'une des 8 lettres uniques, il reste alors $\frac{11!}{4}$ anagrammes. Sinon, on a choisi une des deux lettres doubles. Il reste alors $\frac{11!}{2}$ anagrammes. Il existe donc en tout $3 \times 11!$ anagrammes. Parmi ces 119750400 anagrammes (autant donc que d'anagrammes complètes, mais c'est une coïncidence), mon ordinateur m'assure qu'il n'existe que deux mots appartenant à la langue française : probabilité et bipolarités.

Exercice 4 Parmi n personnes, on décide de former des groupes de travail. Un tel groupe est formé d'un certain nombre de personnes et d'un représentant du groupe (un groupe n'est donc pas vide). Combien de groupes différents peut-on former. En déduire la formule suivante :

$$\sum_{s=0}^{n} s \binom{n}{s} = n2^{n-1}$$

Correction Pour former un groupe, on peut commencer par choisir le représentant. Il y a alors n choix. Ensuite, pour les $n-1$ personnes restantes, on a le choix entre les mettre dans le groupe ou non. Il y a donc un choix entre deux possibilités pour chaque personne, donc 2^{n-1} possibilités. Il y a donc $n2^{n-1}$ groupes. Alternativement, pour construire un groupe, on peut commencer par choisir un nombre s de membres et les s membres du groupes. Il y a $\binom{n}{s}$ groupes à s membres. Il faut ensuite encore choisir un représentant du groupe. Il y a donc $s\binom{n}{s}$ groupes de s membres. Il y a donc

$$\sum_{s=0}^{n} s \binom{n}{s} = n2^{n-1}$$

groupes.

Exercice 5 Montrer :

$$\sum_{s=0}^{n} \binom{n}{s} = 2^n, \ \sum_{s=0}^{n} (-1)^s \binom{n}{s} = 0$$

Correction La première somme est le nombre total d'ensemble d'éléments que l'on peut former avec n éléments. Pour former un tel ensemble, il faut choisir pour chacun des n éléments s'il appartient ou non à l'ensemble. Il y a donc un choix parmi deux possibilités pour chacun des éléments, soit 2^n choix. La seconde somme compare le nombre de s-uplets de cardinal pair avec le nombre de p-uplets de cardinal impair. Si n est impair, le choix d'un p-uplet pair revient au choix d'un $(n-p)$-uplet impair. Il y a donc autant de p-uplet pair que de p-uplet impair. Si n est pair, il y a d'après ce que l'on vient de voir autant de p-uplet pairs ne contenant pas n que de p-uplets impairs ne contenant pas n. Un p-uplet pair contenant n est la même chose qu'un p-uplet impair ne contenant pas n. Il y en a donc autant que de p-uplets pairs ne contenant pas n ; ce qui est aussi le nombre de p-uplets impairs contenant n.

Alternativement, on peut remarquer que la deuxième somme est le développement de $(1-1)^n$.

2 Fondations axiomatiques des probabilités

2.1 Langage de la théorie des ensembles

On rappelle que si X est un ensemble, on peut effectuer les opérations d'intersection, union et passage au complémentaire sur les parties de X. On dit qu'un ensemble est dénombrable s'il est fini ou s'il peut s'écrire sous la forme $X = \{x_1, x_2, x_3 \cdots\}$.

2.2 Espace des événements et probabilités

On considère un ensemble Ω, que l'on appelle l'univers ainsi qu'un sous-ensemble T des parties de Ω vérifiant les propriétés suivantes :

1. T est non-vide.
2. Si $A \in T$ (donc $A \subset \Omega$) alors le complémentaire $\bar{A}$ de A dans Ω est aussi dans T.
3. Si $\mathcal{A} = \{A_1, A_2 \cdots\}$ est un ensemble dénombrable d'éléments de T, alors l'union de tous les A_i est également dans T.

Dans le cas, fréquent dans ce cours, où l'ensemble Ω est un ensemble fini, on peut remplacer la troisième condition par la condition suivante : si A et B sont dans T, alors $A \cup B$ est dans T. Un sous-ensemble T vérifiant les propriétés ci-dessus est appelé une tribu sur les parties de Ω.

Les éléments de T sont appelés les éléments de l'espace des événements. Notons que d'après la première propriété de T, il existe un élément $A \in T$. D'après la deuxième, le complémentaire $\bar{A}$ de A est dans T. D'après la troisième, l'union $A \cup \bar{A} = \Omega$ est dans T. L'ensemble Ω tout entier est donc toujours dans T. Son complémentaire, l'ensemble vide $\varnothing$, est donc également toujours dans T.

Une probabilité sur (Ω, T) est une fonction qui associe à un événement (donc à un élément $A \in T$) un nombre réel et qui vérifie les propriétés suivantes :

1. Pour tout $A \in T$, $0 \leq p(A) \leq 1$.
2. $p(\Omega) = 1$.
3. Si $\mathcal{A} = \{A_1, A_2 \cdots\}$ est un ensemble dénombrable d'éléments de T deux à deux disjoints, alors :
$$p(A_1 \cup A_2 \cup \cdots) = \sum_{i=1}^{\infty} p(A_i)$$

Dans le cas, fréquent dans ce cours, où l'ensemble Ω est un ensemble fini, on peut remplacer la troisième condition par la condition suivante : si A et B sont dans T et $A \cap B = \varnothing$, alors $p(A \cup B) = p(A) + p(B)$.

On appelle espace probabilisé la donnée de (Ω, T, p).

Ces propriétés impliquent que $p(\varnothing) = 0$, que $p(\bar{A}) = 1 - p(A)$ et que $p(A \cup B) = p(A) + p(B) - p(A \cap B)$.

2.3 Trois exemples fondamentaux

2.3.1 La situation finie équiprobable

Considérons la situation où Ω est l'ensemble fini des éventualités possibles d'une expérience combinatoire. Prenons pour T l'ensemble des parties de Ω. Définissons p par la formule $p(A) = |A|/|\Omega|$. Alors (Ω, T, p) vérifient toutes les propriétés des fondations axiomatiques des probabilités.

Exemples :

1. Pour modéliser le lancer d'un dé équilibré, on peut prendre $\Omega = \{1, 2, 3, 4, 5, 6\}$. La probabilité de tirer un 6 est alors $p(\{6\}) = 1/6$. La probabilité de tirer un 3 ou un 4 est $p(\{3, 4\}) = 2/6 = 1/3$. La probabilité de tirer un nombre pair est $p(\{2, 4, 6\}) = 3/6 = 1/2$.

2. Pour modéliser le lancer de deux pièces équilibrées, on peut prendre $\Omega = \{PP, PF, FP, FF\}$.

3. Pour modéliser le choix d'un groupe de trois personnes (sans ordre) dans un groupe de 20 (dont 14 hommes et 6 femmes), on peut prendre $\Omega = \{\text{groupes de 3 personnes parmi 20}\}$. Alors la probabilité que ce groupe contienne une personne x donnée est $p(A) = \binom{19}{2} / \binom{20}{3}$. La probabilité qu'un groupe ne contienne que des hommes est $p(B) = \binom{14}{3} / \binom{20}{3}$. La probabilité qu'un groupe ne contienne que des hommes et une personne x donnée est $p(A \cap B) = 0$ si x n'est pas un homme et $p(A \cap B) = \binom{13}{2} / \binom{20}{3}$ si x est un homme.

2.3.2 La situation finie pondérée

C'est la même situation que dans la situation finie équiprobable, mais on ne suppose plus que p est définie par $p(A) = |A|/|\Omega|$.

Exemples :

1. Deux joueurs de ping-pong ont remarqué que lorsqu'ils jouent l'un contre l'autre, A gagne 2 fois sur 3 contre B. Ils décident de jouer trois parties consécutives et de déclarer vainqueur celui qui aura remporté deux parties. On peut prendre pour Ω l'ensemble suivant :

$$\{\text{Mots de trois lettres formés des lettres } A \text{ et } B\}$$

   Mais alors, la probabilité de $p(AAA)$ est $(2/3)^3$ tandis que la probabilité de BBB est $(1/3)^3$. La probabilité que A l'emporte est $p(AAA) + 3p(AAB)$.

2. Certaines particules élémentaires s'appellent des bosons (par exemple les photons, le noyau de l'atome de Carbone, le noyau de l'atome d'Hélium) tandis que d'autres s'appellent des fermions (par exemple les élections et les protons). Si l'on considère p bosons pouvant occuper n états physiques discernables, la probabilité que le système physique entier soit dans un état donnée est $1/\binom{n+p-1}{p}$. Si l'on considère p fermions pouvant occuper n états physiques discernables, la probabilité que le système physique entier soit dans un état donnée est $1/\binom{n}{p}$. Si l'on compare à la situation combinatoire, ceci montre d'une part que les bosons et les

fermions sont par nature indiscernables et d'autre part que deux bosons ou plus peuvent occuper le même état alors que deux fermions sont toujours dans des états distincts.

2.3.3 La situation infinie dénombrable

Dans cette situation Ω est infinie mais dénombrable. Les éléments de Ω peuvent donc être identifiés si on le souhaite avec $\mathbb{N}$. On prend pour T l'ensemble des parties de A. La probabilité p doit vérifier l'équation :

$$\sum_{n=0}^{\infty} p(n) = p(0) + p(1) + p(2) + p(3) + \cdots = 1$$

En particulier, pour tout $\varepsilon > 0$, l'ensemble des ω tel que $p(\omega) \geq \varepsilon$ est fini. En effet, dans le cas contraire, l'inégalité

$$\sum_{n=0}^{N} p(n) \geq \varepsilon |\{0 \leq \omega \leq N | p(\omega) \geq \varepsilon\}|$$

montre que la somme des $p(\omega)$ de 0 à N est plus grande que 1 pour N assez grand.

Exemples :

1. $p(n) = e^{-1} \frac{1}{n!}$
2. Considérons la situation où un joueur lance une pièce jusqu'à tomber sur pile et compte le nombre de lancers qu'il doit faire. *A priori*, on ne peut exclure le fait que la pile ne tomber jamais sur pile, et donc que le nombre de lancers sera infini. On peut modéliser cette situation en prenant $p(n) = 1/2^n$ et vérifier que la somme des probabilités est bien 1.
3. Plus généralement, on peut associer une probabilité aux éléments de $\mathbb{N}$ à partir de toutes séries (absolument) convergentes en normalisant la somme de la série. Par exemple, la fonction p définie par

$$\begin{aligned} p : \quad \mathbb{N} &\longrightarrow [0,1] \\ n &\longmapsto \begin{cases} 0 \text{ si } n = 0 \\ \frac{6}{\pi^2 n^2} \text{ si } n > 0 \end{cases} \end{aligned}$$

   est une probabilité.

3 Probabilités conditionnelles

3.1 Fondations axiomatiques des probabilités conditionnelles

3.1.1 Définition

On considère Ω un ensemble, T une tribu sur les parties de Ω et p une probabilité sur T. Soit B un événement. L'ensemble T_B des parties de Ω défini par $\{A \cap B | A \in T\}$ est alors une tribu sur les parties de B. Si de plus $p(B) \neq 0$, alors $p(\cdot|B)$ définie par

$$\begin{aligned} p(\cdot|B) : \quad T_B &\longrightarrow \mathbb{R} \\ A &\longmapsto \frac{p(A \cap B)}{p(B)} \end{aligned}$$

est une probabilité sur T_B. On appelle la probabilité $p(\cdot|B)$la probabilité conditionnelle à B. Intuitivement, elle mesure la probabilité qu'un événement A se produise sachant que l'événement B s'est produit.

Exemples :

1. Quatre objets discernables sont placés dans quatre urnes discernables. Sachant que B les deux premiers objets sont dans des urnes distinctes, calculons la probabilité que A trois objets soient dans la même urne (on suppose que tous les arrangements sont équiprobables). Trois objets sont dans la même urne si et seulement si les deux objets restants sont placés dans l'urne du premier objet ou dans l'urne du deuxième objet. Deux rangements conviennent donc sur les $4^2 = 16$ rangements possibles. La probabilité $p(A|B)$ que trois objets soient dans la même urne est donc $1/8$. Par ailleurs, $p(B)$ est bien $3/4$ et la probabilité de $A \cap B$ est bien $3/32$.

2. Soit une famille ayant deux enfants. Sachant que B cette famille a un garçon, quelle est la probabilité que A les deux enfants soient des garçons (on suppose que tous les arrangements familiaux possibles sont équiprobables) ? Intuitivement, on a envie de répondre que cette probabilité est $1/2$, car le sexe de l'un des enfants n'a pas *a priori* d'incidence sur le sexe de l'autre enfant. Nous allons voir que notre intuition n'est pas bonne. En effet, $p(B) = 3/4$ et $p(A \cap B) = p(A) = 1/4$. Donc $p(A|B) = 1/3$. Notre erreur provient du fait que les trois répartitions possibles des sexes de deux enfants (à savoir fille/fille, fille/garçon et garçon/garçon) ne sont pas équiprobables.

La formule $p(A \cap B) = p(A|B)p(B)$ permet de déduire la probabilité d'une intersection de la probabilité conditionnelle. Il est instructif de la généraliser au cas d'une intersection multiple.

$$\begin{aligned} p(A \cap B \cap C) &= p(A \cap (B \cap C)) = p(A|B \cap C)p(B \cap C) \\ &= p(A|B \cap C)p(B|C)p(C) \end{aligned}$$

Imaginons que $C = C \cap D$ dans les équations précédentes. On obtient alors :

$$\begin{aligned} p(A \cap B \cap C \cap D) &= p(A \cap (B \cap C \cap D)) = p(A|B \cap C \cap D)p(B \cap C \cap D) \\ &= p(A|B \cap C \cap D)p(B|C \cap D)p(C \cap D) \\ &= p(A|B \cap C \cap D)p(B|C \cap D)p(C|D)p(D) \end{aligned}$$

Et plus généralement :

$$p\left(\bigcap_{i=1}^{n} A_i\right) = p(A_1)\prod_{i=2}^{n} p\left(A_i|\bigcap_{j=1}^{i-1} A_j\right) \tag{1}$$

3.1.2 Indépendance

Lorsque $p(A|B) = p(A)$, on dit que les événements A et B sont indépendants. Si $p(A) \neq 0$ et si A et B sont indépendants, alors $p(A|B) = \frac{p(A \cap B)}{p(B)} = p(A)$ donc $p(B) = \frac{p(A \cap B)}{p(A)} = p(B|A)$ donc B et A sont indépendants. Lorsque deux événements sont indépendants :

$$p(A \cap B) = p(A)p(B)$$

Plus généralement, si $A_1, \cdots, A_n$ sont des événements mutuellement indépendants, la formule (1) donne :

$$p\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} A_i$$

Ces relations ne demeurent bien évidemment pas vraies si A et B ne sont pas indépendants.

3.2 Deux théorèmes fondamentaux

3.2.1 Formule des probabilités totales

Théorème 1. *Soit A un événement et $(B_i)_{1\leq i\leq n}$ une famille d'événements disjoints telle que $p(\bigcup_{i=1}^{n} B_i) = 1$. Alors :*

$$p(A) = \sum_{i=1}^{n} p(A|B_i)p(B_i)$$

Démonstration. Si $i \neq j$, alors $B_i \cap B_j = \varnothing$. Donc, $(A \cap B_i) \cap (A \cap B_j) = \varnothing$ Donc :

$$p((A \cap B_i) \cup (A \cap B_j)) = p(A \cap B_i) + p(A \cap B_j) = p(A|B_i)p(B_i) + p(A|B_j)p(B_j)$$

Soit C le complémentaire de l'union des B_i. Alors $p(C) = 0$. Donc :

$$\begin{aligned} p(A) &= p((A \cap C) \cup (A \cap \bigcup_{i=1}^{n} B_i)) = p(A \cap C) + p(A \cap \bigcup_{i=1}^{n} B_i) \\ &= \sum_{i=1}^{n} p(A \cap B_i) = \sum_{i=1}^{n} p(A|B_i)p(B_i) \end{aligned}$$

□

Un cas fréquent d'application de ce théorème est celui où l'on prend $(B_i) = (B, \bar{B})$.

3.2.2 Formule de Bayes

Théorème 2. *Soit A et B deux événements tels que $p(A)p(B) \neq 0$. Alors :*

$$p(A|B)p(B) = p(B|A)p(A)$$

Démonstration. En effet, les deux membres de l'équation sont égaux à $p(A \cap B)$. □

Exemple : Chaque jour, la probabilité que le RER arrive en retard est de $1/4$. Lorsque le RER arrive en retard, un étudiant arrive en retard à son cours avec une probabilité de $9/10$. Lorsque le RER arrive à l'heure, cet étudiant arrive à en retard à son cours avec une probabilité de $1/5$. Aujourd'hui, l'étudiant est arrivé à l'heure. Quelle est la probabilité que le RER soit arrivé en retard ? Soit A l'événement "l'étudiant arrive à l'heure" et B l'événement "le RER arrive en retard". D'après l'énoncé, $p(A|B) = 1/10$, $p(A|\bar{B}) = 4/5$ et $p(B) = 1/4$. Donc :

$$p(A) = p(A|B)p(B) + p(A|\bar{B})p(\bar{B}) = \frac{1}{10}\frac{1}{4} + \frac{4}{5}\frac{3}{4} = \frac{5}{8}$$

D'après la formule de Bayes :

$$p(B|A) = p(A|B)\frac{p(B)}{p(A)} = \frac{1}{10}\frac{1}{4}\frac{8}{5} = \frac{1}{25}$$

3.3 Exemples

Prévalence d'une allèle récessive dans une population hétérozygote Dans une population, environ 5% des hommes sont daltoniens. Supposons que le daltonisme soit lié à la présence d'une certaine allèle récessive sur le chromosome X. Nous allons déterminer la prévalence de cette allèle parmi les chromosomes X de la population, en déduire la proportion des femmes touchées par le daltonisme, puis la proportion des personnes touchées, et enfin la probabilité qu'une personne daltonienne soit un homme. Notons H l'événement "être un homme", F l'événement "être une femme", D l'événement "être daltonien", X_1 l'événement "porter l'allèle sur le premier chromosome X" et X_2 l'événement "porter l'allèle sur le second chromosome X". Alors $p(D|H) = p(X_1|H) = 1/20$. Il est raisonnable de considérer que les événements X_1 et H sont indépendants, que les événements X_1 et F sont indépendants, que les événements $X_1 \cap F$ et $X_2 \cap F$ sont indépendants et enfin que la probabilité $p(X_2 \cap F)$ est égale à $p(X_1)$.

Dans ces conditions, $p(X_1) = 1/20$ et

$$\begin{aligned} p(D|F) = p(X_1 \cap X_2|F) &= \frac{p(X_1 \cap F \cap X_2 \cap F)}{p(F)} \\ &= \frac{p(X_1 \cap F)p(X_2 \cap F)}{p(F)} = \frac{p(X_1 \cap F)p(X_1)}{p(F)} \\ &= \frac{p(X_1)^2 p(F)}{p(F)} = p(X_1)^2 = \frac{1}{400} \end{aligned}$$

De plus :

$$p(D) = p(D|M)p(M) + p(D|F)p(F)$$

donc $p(D) = \frac{20}{800} + \frac{1}{800} = 21/800$ et $p(M|D) = p(D|M)\frac{p(M)}{p(D)}$ donc $p(M|D) = \frac{5\cdot 800}{200\cdot 21} = 20/21$.

Contrôle de qualité dans une chaîne de montage Une chaîne de montage est formée de trois machines A, B et C. Elles produisent respectivement 25%, 35% et 40% de la production totale de l'usine. De plus, 5%, 4% et 2% de leur productions respectives présentent un défaut. Un objet est choisi au hasard et l'on constate qu'il est défectueux. Quelle est la probabilité qu'il provienne de la machine A ?

D'après la formule des probabilités totales :

$$p(D) = p(D|A)p(A) + p(D|B)p(B) + p(D|C)p(C) = \frac{1}{20}\frac{1}{4} + \frac{1}{25}\frac{7}{20} + \frac{1}{50}\frac{2}{5} = 69/2000$$

Donc :

$$p(A|D) = p(D|A)\frac{p(A)}{p(D)} = \frac{1}{20}\frac{2000}{69 \cdot 4} = 25/69 \simeq 36\%$$

Boite de chocolats Une boite de chocolat divisée en 6 compartiments contient un nombre n inconnu, mais compris entre 0 et 5, de chocolats. On suppose que toutes les valeurs possibles du nombre de chocolat sont équiprobables et que toutes les répartitions sont équiprobables. Vous vous approchez du paquet pour prendre un chocolat dans le compartiment 1. Quelle est votre probabilité d'attraper effectivement un chocolat ? Juste avant de tendre la main, quelqu'un vous passe devant et prend un chocolat dans le compartiment 1. Quelle est maintenant votre probabilité d'attraper un chocolat dans le compartiment 2 ?

On note i l'événement "Il y a un chocolat dans le compartiment i" et n l'événement "Il y a n chocolats dans la boite". Notons que par symétrie (ou par le calcul), la probabilité $p(i)$ ne dépend pas de i. Dans une répartition contenant n chocolat, la probabilité que le compartiment a soit plein $\binom{5}{n-1} / \binom{6}{n}$[2]. Donc la probabilité $p(1)$ de trouver un chocolat dans le compartiment 1 est :

$$\begin{aligned} p(1) &= \sum_{n=0}^{5} p(1|n)p(n) \\ &= \sum_{n=0}^{5} \frac{\binom{5}{n-1}}{6\binom{6}{n}} \\ &= 5/12 \simeq 0,41 \end{aligned}$$

Calculons maintenant $p(2|1)$. Dans une répartition contenant n chocolats, la probabilité que les deux compartiments 1 et 2 soient pleins est $p(1 \cap 2|n) = \binom{4}{n-2} / \binom{6}{n}$. Donc :

$$\begin{aligned} p(2|1) &= \frac{p(1 \cap 2)}{p(1)} \\ &= \sum_{n=0}^{5} \frac{12\binom{4}{n-2}}{30\binom{6}{n}} \\ &= 8/15 \simeq 0,53 \end{aligned}$$

De façon peut être contre-intuitive, voir quelqu'un manger un chocolat de la boite juste avant nous augmente donc notre probabilité de manger un chocolat.

4 Variable aléatoire

4.1 Définition

Une variable aléatoire est une fonction définie sur un espace probabilisé. Si X est une variable aléatoire à valeurs dans un ensemble E dénombrable, alors la fonction

$$\begin{aligned} p(X = \cdot): \quad E &\longrightarrow [0,1] \\ x \in E &\longmapsto p(X = x) \end{aligned}$$

est une fonction sur E à valeurs dans $[0,1]$ que l'on appelle la distribution de X. Si f est la distribution d'une variable aléatoire X à valeurs dans un ensemble E dénombrable, alors :

$$\sum_{x \in E} f(x) = 1$$

Si X est une variable aléatoire sur (Ω, T, p) et si $B \in T$ est de probabilité non-nulle alors X définit par restriction à $(B, T_B, p(\cdot|B))$ une variable aléatoire $X|B$ sur $(B, T_B, p(\cdot|B))$. La distribution de $X|B$ est $p(X = \cdot|B)$.

[2]On rappelle que le coefficient binomial $\binom{n}{p}$ est nul si p est strictement négatif.

Soit X et Y deux variables aléatoires sur (Ω, T, p) à valeurs dans E et F dénombrables. La distribution jointe de X et de Y est la fonction

$$p(X=\cdot, Y=\cdot): \quad \begin{array}{rl} E \times F & \longrightarrow [0,1] \\ (x,y) \in E \times F & \longmapsto p(X=x \cap Y=y) \end{array}$$

La formule

$$p(X=x) = \sum_{y \in F} p(X=x \cap Y=y)$$

montre que la distribution jointe de X et Y détermine la distribution de X, et par symétrie celle de Y. En revanche, la connaissance de la distribution de X et de celle de Y ne donne en général aucune information sur la distribution jointe de X et Y. Soit F' le sous-ensemble $\{f \in F | p(Y=f) > 0\}$. La distribution conditionnelle de X connaissant Y est la fonction

$$p(X=\cdot|Y=\cdot): \quad \begin{array}{rl} E \times F' & \longrightarrow [0,1] \\ (x,y) \in E \times F' & \longmapsto p(X=x|Y=y) \end{array}$$

Lorsque $p(X=\cdot|Y=\cdot)$ est égal à la fonction $p(X=\cdot)p(Y=\cdot)$, on dit que les variables aléatoires X et Y sont indépendantes.

Exemples :

1. On lance 3 dés à six faces. Soit $\Omega = \{1, \cdots, 6\}^3$. Les valeurs D_1, D_2 et D_3 obtenues sur chaque dé sont des variables aléatoires sur l'espace probabilisé $(\Omega, \mathcal{P}(\Omega), p)$ à valeurs dans $\{1, \cdots, 6\}$. Les trois distributions de ces variables vérifient $f(i) = 1/6$ et sont donc égales. La variable aléatoire $S = D_1 + D_2 + D_3$ est à valeur dans $\{3, \cdots, 18\}$. La variable aléatoire $Q = D_1/D_2$ est à valeur dans $[0,6]$ (par exemple) mais sa distribution g vérifie $g(x) = 0$ si $x \notin \{a/b | (a,b) \in \{1, \cdots, 6\}\}$. La variable aléatoire N égale au nombre de 6 obtenus est à valeurs dans $\{0,1,2,3\}$. Les variables aléatoires D_1 et D_2 sont indépendantes tandis que les variables aléatoires S et N ne le sont pas : $p(S=18|N=3) = 1$ alors que $p(S=18)p(N=3) = 1/6^9$.

2. Un étudiant estime avoir neuf chances sur dix d'obtenir la moyenne à chacun de ces 5 examens. Le nombre d'examen où il aura effectivement obtenu la moyenne est une variable aléatoire X sur $(\Omega = \{0,1\}^5, \mathcal{P}(\Omega), p)$ à valeurs dans $\{0, \cdots, 5\}$. Elle vérifie $p(X=0) = 1/10^5$ et $p(X \geq 3) \simeq 99\%$.

3. Chaque année, un conducteur à une probabilité x d'avoir un accident. La fonction X donnant le nombre d'années avant son prochain accident est une variable aléatoire sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}), p)$. Ici, on ne précisera pas ce qu'est la probabilité p ; ce serait possible mais assez délicat. L'événement $X = k$ signifie que le conducteur n'a pas eu d'accidents pendant $k-1$ année puis qu'il a eu un accident l'année k. Donc $p(X=k) = (1-x)^{k-1}x$.

4.2 Lois fondamentales

4.2.1 Loi équirépartie

On dit qu'une variable X suit la loi équirépartie si X est à valeur dans un ensemble fini de cardinal n et si $p(X=x) = 1/n$.

4.2.2 Tirage de Bernoulli

On dit qu'une variable X suit la loi $B(p)$ de tirage de Bernoulli de paramètre p si X est à valeur dans $\{0,1\}$ et si $p(X=1)=p$.

4.2.3 Loi binomiale

On dit qu'une variable X suit la loi binomiale $B(n,p)$ de paramètres n et p si X est à valeur dans $\{0,n\}$ et si :

$$p(X=s)=\binom{n}{s}p^s(1-p)^{n-s}$$

On remarque que la loi $B(p)$ est la loi $B(1,p)$ et que X suit la loi $B(n,p)$ lorsque X compte le nombre de succès (identifié avec la valeur 1) dans une répétition de n tirages de Bernoulli.

4.2.4 Loi géométrique

On dit qu'une variable X suit la loi géométrique $G(p)$ de paramètre p si X est à valeur dans $\mathbb{N}^*$ et si $p(X=n)=(1-p)^{n-1}p$. On remarque que X suit la loi géométrique lorsque X compte le nombre d'essais avant le premier succès dans une répétition de tirages de Bernoulli.

4.2.5 Loi de Poisson

On dit qu'une variable X suit la loi de Poisson $P(\lambda)$ de paramètre λ si X est à valeur dans $\mathbb{N}$ et si :

$$p(X=n)=e^{-\lambda}\frac{\lambda^n}{n!}$$

4.2.6 Loi hypergéométrique

On dit qu'une variable X suit la loi hypergéométrique $H(N,n,p)$ de paramètres N, n et p si X est à valeur dans $\{0,n\}$ et si :

$$p(X=s)=\frac{\binom{n}{s}\binom{N-n}{Np-s}}{\binom{N}{Np}}$$

Considérons un tirage (sans remise) de n boules d'une urne contenant N boules, $N_r=Np$ étant rouges et $N_n=N(1-p)$ étant noires. Le nombre de tels tirages est $\binom{N}{n}$. Soit P le nombre de tels tirages contenant exactement s boules rouges. Un tel tirage contient exactement $n-s$ boules noires choisies parmi $N(1-p)$ et s boules rouges choisies parmi Np. Donc

$$P=\binom{Np}{s}\binom{N(1-p)}{n-s}$$

et la probabilité qu'un tirage comporte exactement s boules rouges est :

$$\begin{aligned}
p = \frac{\binom{Np}{s}\binom{N(1-p)}{n-s}}{\binom{N}{n}} &= \frac{N_r!N_n!n!(N-n)!}{s!(N_r-s)!(n-s)!(N_n-n+s)!N!} \\
&= \frac{n!}{s!(n-s)!} \cdot \frac{(N-n)!}{(N_r-s)!(N-n-N_r+s)!} \cdot \frac{N_r!}{N!(N_r-s)!} \\
&= \frac{\binom{n}{s}\binom{N-n}{Np-s}}{\binom{N}{Np}}
\end{aligned}$$

On remarque donc que X suit la loi $H(N, n, p)$ lorsque X compte le nombre de boules rouges tirées dans un tirage (sans remise) de n boules d'une urne contenant N boules, Np étant rouges.

4.3 Espérance, variance

4.3.1 Espérance

Soit X une variable aléatoire à valeurs dans $\mathbb{R}$. L'espérance de X est définie par

$$E(X) = \sum_{x \in \mathbb{R}} xp(X = x) \tag{2}$$

lorsque la somme précédente est convergente. L'espérance est en particulier bien définie lorsque la variable aléatoire X est à valeurs dans un ensemble fini avec probabilité 1. En effet, la somme (2) devient dans ce cas une somme finie. La définition montre également qu'une variable aléatoire constante est égale à son espérance et réciproquement qu'une variable aléatoire est égale à son espérance si elle est constante avec probabilité 1. Il est usuel d'interpréter l'espérance de X comme la valeur moyenne de X mais il n'est pas si facile de donner un sens précis à cette assertion.

Théorème 3. *Soit X, Y deux variables dont les espérances sont définies. Alors $E(X+Y)$ est définie et $E(X+Y) = E(X) + E(Y)$.*

Démonstration.

$$\begin{aligned}
E(X+Y) &= \sum_{(x,y) \in \mathbb{R}^2} (x+y)P(X = x \cap Y = y) \\
&= \sum_{(x,y) \in \mathbb{R}^2} xP(X = x \cap Y = y) + \sum_{(x,y) \in \mathbb{R}^2} yP(X = x \cap Y = y) \\
&= \sum_{x \in \mathbb{R}} x \sum_{y \in \mathbb{R}} P(X = x \cap Y = y) + \sum_{y \in \mathbb{R}} y \sum_{x \in \mathbb{R}} P(X = x \cap Y = y)
\end{aligned}$$

Or :

$$\sum_{y \in \mathbb{R}} P(X = x \cap Y = y) = P(X = x), \quad \sum_{x \in \mathbb{R}} P(X = x \cap Y = y) = P(Y = y)$$

Donc :

$$E(X+Y) = E(X) + E(Y)$$

□

L'espérance est donc linéaire, ce qui signifie que $E(aX + bY) = aE(X) + bE(Y)$ pour tous réels a, b et toutes variables aléatoires X, Y.

4.3.2 Covariance, variance

Soit X et Y deux variables aléatoires telles que X^2 et Y^2 aient une espérance. La covariance de X et de Y est alors l'espérance de $(X - E(X))(Y - E(Y))$. Soit :

$$\operatorname{Cov}(X, Y) = E\left((X - E(X)(Y - E(Y)))\right)$$

Par additivité de l'espérance, la covariance de X et Y est aussi $E(XY) - E(X)E(Y)$. En particulier, si X et Y sont indépendantes alors $\operatorname{Cov}(X, Y) = 0$. La réciproque n'est pas vraie en général : il existe des variables aléatoires de covariance nulle et qui ne sont pas indépendantes. La covariance de X avec elle-même s'appelle la variance et est notée $\operatorname{Var}(X)$. Par additivité de l'espérance, $\operatorname{Var}(X)$ est égale à $E(X^2) - E(X)^2$.

Par définition, la variance est l'espérance d'une variable aléatoire positive ou nulle donc est positive ou nulle. Ceci montre que $E(X^2) \geq E(X)^2$. En cas d'égalité, la variable aléatoire $(X - E(X))^2$ est nulle donc X est égale à son espérance avec probabilité 1, ou de manière équivalente, X est constante avec probabilité 1.

Théorème 4. *Soit $X_1, \cdots, X_n$ des variables aléatoires dont les variances sont définies. Soit S_n la somme des X_s. Alors :*

$$\operatorname{Var}(S_n) = \sum_{s=1}^{n} \operatorname{Var}(X_s) + 2\sum_{s<t} \operatorname{Cov}(X_s, X_t)$$

En particulier, si les X_s sont mutuellement indépendantes, alors :

$$\operatorname{Var}(S_n) = \sum_{s=1}^{n} E(X_s^2)$$

Démonstration. En effet :

$$E(S_n) = \sum_{s=1}^{n} E(X_s),\ (S_n - E(S_n))^2 = \left(\sum_{s=1}^{n} (X_s - E(X_s))\right)^2$$

Donc :

$$(S_n - E(S_n))^2 = \sum_{s=1}^{n} (X_s - E(X_s))^2 + 2\sum_{s<t} (X_s - E(X_s))(X_t - E(X_t))$$

En prenant les espérances des deux côtés de l'équation précédente, on obtient la première assertion. La seconde s'en déduit car la covariance de deux variables aléatoires indépendantes est nulle. □

La variance n'est pas linéaire. Le théorème précédent montre que $\operatorname{Var}(aX + b) = a^2 \operatorname{Var}(X)$.

4.3.3 Coefficient de corrélation

La formule $\operatorname{Cov}(aX, aY) = a^2 \operatorname{Cov}(X, Y)$ montre que la variance d'une variable aléatoire correspondant à une grandeur physique ayant une unité dépend du choix de l'unité. Pour éviter, cet effet fâcheux, il est commode d'introduire le coefficient de corrélation $\operatorname{Cor}(X, Y)$ défini par :

$$\operatorname{Cor}(X, Y) = \frac{\operatorname{Cov}(X, Y)}{\sqrt{\operatorname{Var}(X) \operatorname{Var}(Y)}}$$

Alors $\text{Cor}(aX+b, cY+d) = \text{Cor}(X,Y)$. Si X et Y sont des variables aléatoires indépendantes, alors $\text{Cor}(X,Y) = 0$. Notons que la réciproque n'est pas vraie : si $\text{Cor}(X,Y) = 0$, tout ce que l'on peut conclure est que l'une des variables n'est pas une fonction linéaire de l'autre. D'un point de vue scientifique, il est également important de distinguer corrélation et explication : par exemple, la consommation mensuelle de glaces est fortement corrélée à la fréquence des accidents d'avions liés aux oiseaux dans les aéroports français sans pour autant que l'on puisse supposer que ce sont les glaces qui provoquent des accidents d'avions ou réciproquement que les accidents d'avions font acheter des glaces. La raison est tout simplement que les gens mangent plus de glaces l'été et qu'il y a plus d'oiseaux en France l'été que l'hiver.

4.4 Espérance et variance des lois fondamentales

4.4.1 Loi équirépartie

Soit X suivant la loi équirépartie à valeurs dans $\{1, \cdots, n\}$. Alors :

$$E(X) = \sum_{s=1}^{n} s \cdot p(X=s) = \frac{1}{n}\sum_{s=1}^{n} s = \frac{n+1}{2}$$

Ceci montre que si les correcteurs notaient au hasard entre 0 et 20, alors les candidats pourraient espérer obtenir la moyenne à leurs examens (l'espérance de la variable aléatoire N qui représente la note étant alors en effet égale à 10).

La variance de X est définie par :

$$\text{Var}(X) = E(X^2) - E(X^2) = \frac{1}{n}\sum_{s=1}^{n} s^2 - \left(\frac{n+1}{2}\right)^2 = \frac{(n+1)(n-1)}{12}$$

4.4.2 Tirage de Bernoulli

La définition montre directement que l'espérance d'un tirage de Bernoulli est p et que la variance d'un tirage de Bernoulli est $p(1-p)$.

4.4.3 Loi binomiale

Soit X suivant la loi binomiale. Par définition :

$$E(X) = \sum_{s=0}^{n} s \binom{n}{s} p^s (1-p)^{n-s}$$

On peut calculer l'espérance de X en évaluant cette somme (par exemple en reconnaissant la somme d'une dérivée et d'une série binomiale). Il est néanmoins beaucoup plus aisé de remarquer que X est la somme de n tirages de Bernoulli. Donc $E(X) = np$ d'après le théorème 3. De même :

$$\text{Var}(X) = \sum_{s=0}^{n} s^2 \binom{n}{s} p^s (1-p)^{n-s} - n^2 p^2$$

A nouveau, on peut évaluer cette somme (par exemple en reconnaissant une combinaison linéaire d'une dérivée seconde, d'une dérivée première et d'une série binomiale). Il est néanmoins beaucoup plus aisé d'appliquer le théorème 4. Les tirages de Bernoulli étant indépendants, on obtient ainsi :

$$\mathrm{Var}(X) = \sum_{s=0}^{n} \mathrm{Var}(X_s) = np(1-p)$$

4.4.4 Loi géométrique

Soit X suivant la loi géométrique $G(p)$ de paramètre p. Calculons l'espérance (en la considérant implicitement comme une fonction de p) :

$$\begin{aligned} E(X) &= \sum_{s=1}^{\infty} s(1-p)^{s-1}p = -p\left(\sum_{s=1}^{\infty}(1-p)^s\right)' \\ &= -p\left(\frac{1}{p}\right)' = \frac{1}{p} \end{aligned}$$

Le calcul de la variance est similaire mais un peu plus complexe :

$$\begin{aligned} E(X^2) &= \sum_{s=1}^{\infty} s^2(1-p)^{s-1}p = p\sum_{s=1}^{\infty} s(s-1)(1-p)^{s-1} + p\sum_{s=1}^{\infty}(1-p)^{s-1} \\ &= -p(1-p)\left(\frac{1}{p}\right)'' + \frac{1}{p} = \frac{2-p}{p^2} \end{aligned}$$

Donc :

$$\mathrm{Var}(X) = \frac{1-p}{p^2}$$

4.4.5 Loi de Poisson

Soit X une variable suivant la loi de Poisson $\mathcal{P}(\lambda)$. Alors :

$$\begin{aligned} E(X) &= \sum_{n=0}^{\infty} np(X=n) = \sum_{n=0}^{\infty} ne^{-\lambda}\frac{\lambda^n}{n!} \\ &= \lambda e^{-\lambda}\sum_{n=0}^{\infty} n\frac{\lambda^{n-1}}{n!} = \lambda e^{-\lambda}\frac{d}{d\lambda}\left(\sum_{n=0}^{\infty}\frac{\lambda^n}{n!}\right) \\ &= \lambda e^{-\lambda}\frac{\mathrm{d}}{\mathrm{d}\lambda}(e^{\lambda}) = \lambda \end{aligned}$$

Ce résultat nous permets de décrire beaucoup de variables aléatoires suivant la loi de Poisson. Soit M un phénomène aléatoire qui a les caractéristiques suivantes :

1. Une occurrence de M peut intervenir à tout moment.
2. Toutes les occurrences de M sont indépendantes.
3. L'espérance du nombre d'occurrences de M pendant un intervalle de temps de longueur t est égale à λt.

Soit X la variable aléatoire qui compte le nombre d'occurrences de M pendant un intervalle donné de longueur 1. Nous allons montré que X suit la loi de Poisson $\mathcal{P}(\lambda)$. Considérons en effet un autre variable aléatoire Y comptant le nombre d'occurrences d'un processus N vérifiant les mêmes propriétés que M. Soit un intervalle de temps de longueur 1. Supposons que le nombre totale d'occurrences de M et N soit égal à n pendant cet intervalle. Il existe alors 2^n distributions distinctes des occurrences de M et N et $\binom{n}{s}$ distributions distinctes ayant exactement s occurrences de M. Donc :

$$p(X = s|X + Y = n) = \frac{1}{2^n}\binom{n}{s}$$

Or :

$$p(X = s|X + Y = n) = p(X = s \cap Y = n - s|X + y = n) = \frac{p(X = s \cap Y = n - s)}{p(X + Y = n)}$$

Les variables X et Y sont indépendantes donc :

$$p(X = s|X + Y = n) = \frac{p(X = s)p(Y = n - s)}{p(X + y = n)}$$

Donc :

$$p(X = s)p(Y = n - s) = \frac{p(X + Y = n)}{2^n}\binom{n}{s}$$

Appliquons l'équation précédente à $s = 0$ et $s = 1$. Alors :

$$p(X = 0)p(Y = n) = \frac{p(X + Y = n)}{2^n},\ p(X = 1)p(Y = n - 1) = \frac{np(X + Y = n)}{2^n}$$

Donc :

$$\frac{p(X = 0)p(Y = n)}{p(X = 1)p(Y = n - 1)} = \frac{1}{n}$$

Soit encore :

$$p(Y = n) = \frac{1}{n} \cdot \frac{p(X = 1)}{p(X = 0)} p(Y = n - 1)$$

Notons $\mu = p(X = 1)/p(X = 0)$. Alors :

$$p(Y = n) = \frac{\mu}{n} p(Y = n - 1) = \frac{\mu^n}{n!} p(Y = 0)$$

La somme des probabilités doit être égales à 1 donc :

$$\sum_{n=0}^{\infty} p(Y = n) = p(Y = 0)\sum_{n=0}^{\infty}\frac{\mu^n}{n!} = p(Y = 0)e^{\mu} = 1$$

Soit encore :

$$p(Y = 0) = e^{-\mu},\ p(Y = n) = e^{-\mu}\frac{\mu^n}{n!}$$

La variable aléatoire Y suit donc la loi de Poisson $\mathcal{P}(\mu)$. De plus, elle est d'espérance λ par hypothèse. Donc $\mu = \lambda$. Par symétrie, la variable aléatoire suit également $\mathcal{P}(\lambda)$. Le calcul ci-dessus montre également que si X_1 et X_2 suivent toutes les deux des lois de Poisson de paramètres respectifs λ_1 et λ_2 et si X_1 et X_2 sont indépendantes, alors $X_1 + X_2$ suit la loi de Poisson de paramètre $\mathcal{P}(\lambda_1 + \lambda_2)$.

Calculons maintenant $E(X^2)$.

$$\begin{aligned}
E(X^2) &= e^{-\lambda}\sum_{s=0}^{\infty} s^2 \frac{\lambda^s}{s!} = e^{-\lambda}\sum_{s=1}^{\infty} s \frac{\lambda^s}{(s-1)!} \\
&= e^{-\lambda}\sum_{s=1}^{\infty} (s-1)\frac{\lambda^s}{(s-1)!} + e^{-\lambda}\sum_{s=1}^{\infty}\frac{\lambda^s}{(s-1)!} \\
&= \lambda^2 e^{-\lambda}\sum_{s=2}^{\infty}\frac{\lambda^{s-2}}{(s-2)!} + \lambda e^{-\lambda}\sum_{s=1}^{\infty}\frac{\lambda^{s-1}}{(s-1)!} \\
&= \lambda^2 + \lambda
\end{aligned}$$

Donc $\mathrm{Var}(X) = \lambda$.

4.4.6 Loi hypergéométrique

Soit X une variable suivant la loi hypergéométrique $H(N, n, p)$. On peut interpréter X comme le nombre de boules rouges tirées dans un tirage de n boules rouges ou blanches (sans remplacement). Donc X est la somme des variables aléatoires X_s telles que $X_s = 0$ si la boule s n'est pas rouge et $X_s = 1$ si la boule s est rouge. Les boules jouant toutes le même rôle, les variables X_s suivent toutes la même loi. Fixer $s = 1$ montre que cette loi est un tirages de Bernoulli. Donc :

$$E(X_s) = \frac{r}{r+b} = p,\ \mathrm{Var}(X_s) = \frac{rb}{(r+b)^2} = p(1-p)$$

Donc $E(X) = np$. La variable aléatoire X_sX_t est un tirage de Bernoulli de probabilité de succès $r(r-1)/(r+b)(r+b-1)$. Donc :

$$\mathrm{Cov}(X_s, X_t) = \frac{-rb}{(r+b)^2(r+b-1)}$$

Et :

$$\mathrm{Var}(X) = np(1-p)\left(1 - \frac{n-1}{N-1}\right)$$

On remarque donc qu'un tirage sans remplacement à la même espérance et une variance plus faible qu'un tirage avec remplacement.

4.5 Paradoxe du temps d'attente

4.5.1 Motivation et formalisation

Un RER arrive en moyenne toutes les 6 minutes. Si l'on arrive au hasard sur le quai du RER, combien de temps va-t-on attendre en moyenne ? La réponse à cette question est contre-intuitive et un peu déprimante : au minimum 3 minutes, et parfois bien plus. Nous verrons que dans certain cas, tout se passe comme si arriver au hasard sur le quai du RER revenait à arriver juste après le dernier passage.

Formalisons le problème de la manière suivante. L'arrivée du i-ème RER est un événement aléatoire X_i et on note L_i la variable aléatoire $X_{i+1} - X_i$. On suppose que toutes les variables L_i suivent la même loi L et on suppose que $\mathrm{Var}(L)$ existe. L'assertion comme quoi un RER passe en moyenne toutes les n minutes est l'assertion que $E(L) = n$. On note $A_{s,n}$ l'événement "arriver sur le quai dans la subdivision $[s, s+1]$ d'un intervalle de longueur n sachant que l'on sait que cet intervalle est de longueur n". L'hypothèse d'arrivée au hasard sur le quai signifie que la probabilité de $A_{s,n}$ ne dépend pas de s. Soit T le temps d'attente. On suppose pour simplifier que $T|A_{s,n}$ est $n - s - 1/2$ (alternativement, on pourrait considérer les événements $A_{x,n}$ "arriver en $x \in [0, n]$" et poser $T|A_{x,n} = n - x$; les deux modélisations donnent le même résultat). On cherche à calculer $E(T)$.

4.5.2 Deux exemples

Supposons tout d'abord que la variable aléatoire L soit constante. Sa valeur est alors égale à n. La variable aléatoire T suit alors la loi équirépartie sur $\{1/2, 3/2 \cdots, n-1-1/2, n-1/2\}$ donc son espérance est $n/2$.

Supposons maintenant que la variable aléatoire L suive la loi géométrique $G(p)$; ce qui serait le cas si au lieu d'attendre le RER, on attendait le premier face dans une série de pile ou face en arrivant au hasard au milieu d'une série de lancers. Alors, au facteur $1/2$ de normalisation près, T suit également la loi géométrie $G(p)$ donc $E(T) = 1/p - 1/2 = E(L) - 1/2$. Dans ce cas, on voit qu'en arrivant au hasard, tout se passe comme si on venait de rater le dernier RER.

4.5.3 Cas général

Voici comment procéder en général. Soit I la variable aléatoire définie par la longueur de l'intervalle entre la dernière et la prochaine occurrence de X observée par l'observateur. La clé du problème est de comprendre que I ne suit pas la même loi que L. En effet, si $\mathbf{p}(L = \ell_1) = \mathbf{p}(L = \ell_2)$ avec $\ell_1 < \ell_2$, alors $\mathbf{p}(I = \ell_1) \leq \mathbf{p}(I = \ell_2)$ car un observateur arrivant au hasard a plus de chance d'arriver dans un intervalle plus long que dans un intervalle plus court. Formellement, I est une variable aléatoire donc est une fonction :

$$\begin{aligned} I : \quad (\Omega, \tau, \mathbf{p}) &\longrightarrow (\mathbb{N}, \mathcal{P}(\mathbb{N}), \mathbf{p}') \\ e &\longmapsto I(e) = n \end{aligned}$$

Donc :

$$\mathbf{p}(I = n) = \mathbf{p}'(\{n\}) = \frac{1}{E(L)} \sum_{x \in I^{-1}(n)} n\mathbf{p}(x) = \frac{n\mathbf{p}(L = n)}{E(L)}$$

L'observateur arrivant au hasard donc le temps d'attente sachant que l'intervalle observée est de longueur n est équirépartie. Donc, pour $0 \leq s \leq n-1$ et $t = s + 1/2$

$$\mathbf{p}(T = t | I = n) = \frac{1}{n}$$

Donc :

$$\mathbf{p}(T = t) = \sum_{n \geq t} \frac{\mathbf{p}(L = n)}{E(L)}$$

Et donc :

$$\begin{aligned} E(T) &= \sum_{t=0}^{\infty} \sum_{n \geq t} t \frac{\mathbf{p}(L = n)}{E(L)} = \frac{1}{E(L)} \sum_{n=1}^{\infty} \sum_{s=0}^{n-1} (s + \frac{1}{2}) \mathbf{p}(L = n) \\ &= \frac{1}{E(L)} \sum_{n=1}^{\infty} \frac{n^2}{2} \mathbf{p}(L = n) = \frac{E(L^2)}{2E(L)} \end{aligned}$$

Comme $\mathrm{Var}(L) = E(L^2) - E(L)^2$:

$$E(T) = \frac{E(L)}{2}\left(1 + \frac{\mathrm{Var}(L)}{E(L)^2}\right)$$

4.5.4 Conclusion et applications

Le temps d'attente observée est toujours supérieur à $E(L)/2$ et l'égalité est réalisée si et seulement si L est de variance nulle, donc constante. A espérance égale, mieux vaut toujours avoir une variance faible si l'on veut éviter d'attendre. Lorsque L suit la loi géométrique, on retrouve bien que l'espérance du temps d'attente est égale à l'espérance de la longueur de l'intervalle entre deux RER moins $1/2$. Lorsque L suit la loi de Poisson, l'espérance du temps d'attente est même égal à l'espérance de la longueur de l'intervalle entre deux RER. Dans ce cas, arriver au hasard revient donc systématiquement à voir les portes se fermer juste devant soi. Le fait que le temps d'attente soit toujours supérieur à l'espérance du temps entre deux occurrences a des conséquences importantes en pratique. Tout d'abord, cela explique pourquoi nous avons du mal à croire la RATP quand elle affirme qu'il y a un RER toutes les 10 minutes aux heures de pointe. En effet, nous avons tendance à croire que cela signifie que nous attendrons alors en moyenne cinq minutes et nous venons de voir qu'il n'en est rien. Ensuite, si au lieu d'imaginer quelqu'un en train d'attendre son RER, on imagine quelqu'un qui mesure la fiabilité d'une installation. Attendre le prochain incident et multiplier par deux surestimera la fiabilité du réseau ; ce qui est bien évidemment dommageable pour la sécurité de l'installation. Par exemple, si au moment de souscrire votre police d'assurance, vous attendez votre prochain accident puis vous multiplier le temps d'attente par deux pour estimer combien d'accidents vous aurez en moyenne par an, vous souscrirez une assurance qui ne vous protégera pas assez. Dans la mesure où beaucoup d'accidents suivent la loi de Poisson, vous risquez même de prendre une police d'assurance 2 fois trop laxiste.

4.6 Exercices

4.6.1 Lois binomiales itérées

Énoncé Supposons donnée une expérience dont le nombre de succès suit la loi binomiale $B(n, p)$. L'objectif de cet exercice est de déterminer la loi suivie par le nombre de succès si l'on répète l'expérience pour les instances d'échecs de la première expérience.

Un groupe de n candidats passent un examen. Pour chacun d'entre eux, la probabilité de réussir est $p = 1/2$ et celle d'échouer est $q = 1 - p$. Soit X_1 la variable aléatoire qui compte le nombre de candidats reçus.

1. Quelle est la loi de X_1 et quelle est son espérance ?

Les candidats ayant échoués peuvent passer une session de rattrapage. On note X_2 le nombre de succès dans la session de rattrapage.

2. (a) Montrer que si $0 \le r \le n - s$, on a :

$$\mathbf{p}(X_2 = r | X_1 = s) = \binom{n-s}{r} p^r q^{n-s-r}$$

(b) Montrer l'égalité :

$$\binom{n-s}{r}\binom{n}{s} = \binom{n-r}{s}\binom{n}{r} \tag{3}$$

(c) Montrer que :

$$\mathbf{p}(X_2 = r) = \sum_{s=0}^{n-r} \mathbf{p}(X_2 = r | X_1 = s)\mathbf{p}(X_1 = s)$$

(d) Déduire de la question précédente et de l'équation (3) que X_2 suit la loi binomiale $B(n, pq)$.

3. En déduire la loi suivie par le nombre de succès dans 3 itérations de ce procédé, puis par s itérations.

Correction

1. La variable aléatoire X_1 suit la loi binomiale $B(n, p)$. Donc $E(X_1) = np$.

2. (a) Le nombre de succès obtenus lorsque l'on réalise $n - s$ tirages de Bernoulli indépendants suit la loi binomiale de paramètres $B(n-s, p)$. La probabilité d'obtenir r succès est donc $\binom{n-s}{r} p^r q^{n-s-r}$.

   (b) Le terme de gauche de l'équation est le nombre de couples (a, b) où a est un choix de s éléments parmi n et b un choix de r éléments parmi les $n-s$ restants. Le terme de droite de l'équation est le nombre de couples (b, a) où b est un choix de r éléments parmi n et a un choix de s éléments parmi les $n - r$ restants. L'application $(a, b) \mapsto (b, a)$ est une bijection entre ces deux ensembles donc ces deux nombres sont égaux. Alternativement, on peut utiliser la définition :

$$\begin{aligned}\binom{n-s}{r}\binom{n}{s} &= \frac{(n-s)!}{r!(n-s-r)!}\frac{n!}{s!(n-s)!} = \frac{n!}{s!r!(n-r-s)!} \\ &= \frac{(n-r)!}{s!(n-r-s)!}\frac{n!}{r!(n-r)!} = \binom{n-r}{s}\binom{n}{r}\end{aligned}$$

   (c) Si $X_2 = r$, alors $\{X_1 = s\}_{s=0,1,\ldots,n-r}$ est un système complet d'événements. On applique alors le théorème des probabilités totales pour obtenir l'égalité demandée.

   (d) On a :

$$\begin{aligned}\mathbf{p}(X_2 = r) &= \sum_{s=0}^{n-r} \mathbf{p}(X_2 = r | X_1 = s)\mathbf{p}(X_1 = s) \\ &= \sum_{s=0}^{n-r} \binom{n-s}{r} p^r q^{n-s-r} \binom{n}{s} p^s q^{n-s} \\ &= \sum_{s=0}^{n-r} \binom{n-r}{s}\binom{n}{r} p^{r+s} q^{2n-2s-r} \\ &= \binom{n}{r}(pq)^r \sum_{s=0}^{n-r} \binom{n-r}{s} p^s (q^2)^{(n-r)-s} \\ &= \binom{n}{r}(pq)^r (p+q^2)^{n-r}\end{aligned}$$

Or $pq + p + q^2 = (p+q)q + p = p + q = 1$. Si $p_2 = pq$, on a donc bien :

$$\mathbf{p}(X_2 = r) = \binom{n}{r} p_2^r (1-p_2)^{n-r}$$

Ainsi, la variable aléatoire X_2 suit la loi binomiale $B(n, pq)$. Alternativement, on peut remarquer que X_2 compte le nombre de succès dans n tirages de Bernoulli indépendants pour lesquels le succès est défini par la combinaison échec/réussite dans deux passages d'examens successifs, et donc dont la probabilité de succès est pq.

3. La variable aléatoire X_3 compte le nombre de succès dans n tirages de Bernoulli indépendants pour lesquels le succès est défini par la combinaison échec/échec/réussite et sui donc la loi binomiale $B(n, p(1-p)^2)$. Par le même raisonnement, la variable aléatoire X_s suit la loi binomiale $B(n, p(1-)^{s-1})$.

4.6.2 Variance des lois binomiales pondérées

Énoncé L'objectif de cet exercice est de démontrer que parmi les variables aléatoires comptant le nombre de succès dans n tirages de Bernoulli indépendants et d'espérance np, c'est la loi binomiale (correspondant donc au cas où les tirages ont tous la même probabilité de succès) qui a la plus grande variance.

Soit $X_1, X_2, \cdots, X_n$ des tirages de Bernoulli indépendants tels que la probabilité de succès de X_i soit p_i. Soit S la variable aléatoire égale à la somme des X_i.

1. Calculer l'espérance de S.
2. Calculer la variance de S.
3. Soit $0 \leq p \leq 1$. Parmi toutes les distributions de p_i telles que

$$\sum_{i=1}^{n} p_i = np$$

quelle est celle telle que la variance de S soit maximale ?

Correction

1. L'espérance de X_i est égale à p_i donc :

$$E(S) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} p_i$$

2. Les variables aléatoires X_i sont indépendantes donc :

$$\mathrm{Var}(S) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = \sum_{i=1}^{n} p_i(1-p_i) = \sum_{i=1}^{n} p_i - \sum_{i=1}^{n} p_i^2$$

3. D'après la question précédente, la distribution des p_i maximisant la variance de S est celle qui minimise la somme des p_i^2 sous la condition que :

$$\sum_{i=1}^{n} p_i = np$$

Posons $d_i = p_i - p$. Alors :

$$\sum_{i=1}^{n} d_i = 0$$

Minimiser $\sum_{i=1}^{n} p_i^2$ signifie minimiser minimiser $\sum_{i=1}^{n} (d_i + p)^2$ et ceci est équivalent à minimiser $\sum_{i=1}^{n} (d_i + p)^2 - np^2$ car np^2 est constant. Or :

$$\begin{aligned}\sum_{i=1}^{n} (d_i + p)^2 - np^2 &= \sum_{i=1}^{n} \left((d_i + p)^2 - p^2\right) = \sum_{i=1}^{n} \left(d_i^2 + 2pd_i\right) \\ &= \sum_{i=1}^{n} d_i^2\end{aligned}$$

Cette dernière somme étant une somme de termes positifs, elle est minimale si et seulement tous les d_i sont nuls. La variance de S est donc maximale lorsque tous les p_i sont égaux à p et donc quand S suit la loi binomiale de paramètres $B(n,p)$. Combiné avec les résultats du paradoxe du temps d'attente, cet exercice suggère que l'Univers fait tout pour que nous arrivions le plus en retard possible (en effet, le retard d'un RER, et donc la variance de l'intervalle entre deux RER peut être modélisé de manière réaliste par une série de tirages de Bernoulli représentant chacun la possibilité d'un incident entraînant un retard ; une équipe de maintenance va naturellement se concentrer sur les incidents les plus fréquents et entraînant les retards les plus importants, ce qui aura pour tendantiellement l'effet d'égaliser les espérances des tirages de Bernoulli, donc de maximiser la variance et donc de maximiser l'espérance du temps d'attente).

4.6.3 Loi d'arrêt à pile ou face

Énoncé L'objectif de cet exercice est de démontrer que si l'on réalise une série de lancers à pile ou face, le fait de s'arrêter aléatoirement ou de s'arrêter dès que l'on a obtenu un face ne change ni l'espérance du nombre de piles, ni l'espérance du nombre de faces, ni l'espérance du nombre de lancers. En revanche, cela change radicalement l'espérance de la proportion du nombre de piles parmi tous les lancers.

Un joueur lance une pièce équilibrée. Après chaque lancer, il décide aléatoirement de relancer la pièce avec une probabilité $p = 1/2$ ou de s'arrêter avec la même probabilité. On appelle un tirage la donnée d'une suite de piles et de faces obtenue par le lanceur jusqu'à ce qu'il s'arrête. La longueur d'un tirage est le nombre de lancers effectués. On considère l'espace probabilisé $(\Omega, \mathcal{P}(\Omega), \mathbf{p})$ où Ω est l'ensemble des tirages possibles. On admettra que sous ces conditions, la probabilité d'obtenir un certain tirage ne dépend que de la longueur de ce tirage et que la probabilité d'obtenir un tirage de longueur $n \geq 1$ est égale à 2^{-2n}. On note P la variable aléatoire égale au nombre de piles qu'il a obtenus, F la variable aléatoire égale au nombre de faces obtenues, $N = P + F$ la variable aléatoire égale au nombre de lancers effectués.

1. Vérifier que $\mathbf{p}$ est bien une probabilité.

2. Quelle est la loi suivie par la variable aléatoire N ? Quelle est son espérance ?

3. Soit $n \geq 1$ un entier. Quelle est la loi suivie par la variable aléatoire $P|(N = n)$? Que vaut $E(P|N = n)$?

4. Soit $s \geq 0$ un entier. Montrer que :

$$\mathbf{p}(P = s) = \sum_{n \geq 0} \mathbf{p}(P = s | N = n)\mathbf{p}(N = n)$$

5. Calculer $E(P)$ (ou pourra ou bien utiliser la formule des probabilités totales pour relier $\mathbf{p}(P = x)$ à $E(P|N = n)$ ou bien calculer d'abord $E(P + F)$).

6. Montrer que :

$$E\left(\frac{P}{N}\right) = \sum_{n=1}^{\infty}\sum_{s=0}^{n}\frac{s}{n}\mathbf{p}\left(\frac{P}{N} = \frac{s}{n}|N = n\right)\mathbf{p}(N = n) = \sum_{n=1}^{\infty}\mathbf{p}(N = n)\frac{1}{n}E(P|N = n)$$

7. En déduire que :

$$E\left(\frac{P}{N}\right) = \frac{1}{2}$$

On suppose maintenant que le joueur jette le dé jusqu'à obtenir un face, après quoi il s'arrête. On note P', F' et N' le nombre de piles, le nombre de faces et le nombre total de lancers.

8. Quel est la loi suivie par F' ? Quelle est l'espérance de F' ?

9. Quel est la loi suivie par N' ? Quelle est l'espérance de N' ?

10. En déduire l'espérance de P'.

11. Montrer que l'espérance $E\left(\frac{P'}{N'}\right)$ vaut environ $30,6\%$.

12. Deux joueurs lancent leurs pièces côte à côte. Le premier suit la première règle, le second la deuxième.

    a) Si le vainqueur est celui qui obtient le plus de piles, qui a le plus de chance de l'emporter ?

    b) Si le vainqueur est celui qui obtient le plus de faces, qui a le plus de chance de l'emporter ?

    c) Si le vainqueur est celui qui obtient la plus grande proportion de piles, qui a le plus de chance de l'emporter ?

Correction

1. Il suffit de vérifier que :

$$\sum_{T \text{ tirages possibles}} \mathbf{p}(T) = \sum_{n \geq 1} \mathbf{p}(T \text{ de longueur } n) = 1$$

2. La variable aléatoire N compte le nombre d'essais avant d'obtenir un succès dans une répétition de tirages de Bernoulli. Elle suit donc la loi géométrique $G(p)$. Son espérance est donc $1/p$, soit 2.

3. La variable aléatoire $P|(N = n)$ compte le nombre de pile dans n répétition d'un tirage de Bernoulli. Elle suit donc la loi binomiale $B(n, p)$ et son espérance est np soit $n/2$.

4. Soit $s \geq 0$ un entier. La formule des probabilités totales appliquée à la partition de Ω donnée par les sous-ensembles vérifiant $N = n$ donne :

$$\mathbf{p}(P = s) = \sum_{n \geq 0} \mathbf{p}(P = s | N = n) \mathbf{p}(N = n)$$

5. Nous donnons deux solutions. Les variables aléatoires P et F jouent le même rôle dans l'énoncé donc $E(N) = E(P + F) = E(P) + E(F) = 2E(P)$. Donc $E(P) = 1$.

   Alternativement :

$$\begin{aligned}
E(P) &= \sum_{x=0}^{\infty} x\mathbf{p}(P = x) = \sum_{x=0}^{\infty}\sum_{n=1}^{\infty} x\mathbf{p}(P = x|N = n)\mathbf{p}(N = n) \\
&= \sum_{n=1}^{\infty} \mathbf{p}(N = n) \sum_{x=0}^{n} x\mathbf{p}(P = x|N = n) = \sum_{n=1}^{\infty} \mathbf{p}(N = n) E(P|N = n) \\
&= \sum_{n=0}^{\infty} \mathbf{p}(N = n) \frac{n}{2} = \sum_{n=0}^{\infty} \frac{n}{2^{n+1}} \\
&= 1
\end{aligned}$$

6.

$$\begin{aligned}
E\left(\frac{P}{N}\right) &= \sum_{s/n \in \mathbb{R}} \frac{s}{n} \mathbf{p}\left(\frac{P}{N} = \frac{s}{n}\right) = \sum_{n=1}^{\infty}\sum_{s=0}^{n} \frac{s}{n} \mathbf{p}\left(\frac{P}{N} = \frac{s}{n} | N = n\right) \mathbf{p}(N = n) \\
&= \sum_{n=1}^{\infty} \mathbf{p}(N = n) \frac{1}{n} E(P|N = n)
\end{aligned}$$

7. D'après la question précédente :

$$E\left(\frac{P}{N}\right) = \sum_{n=1}^{\infty} \mathbf{p}(N = n) \frac{1}{n} E(P|N = n) = \sum_{n=1}^{\infty} \mathbf{p}(N = n) \frac{1}{n} \frac{n}{2} = \frac{1}{2} \sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{1}{2}$$

8. La variable aléatoire F' est constante égale à 1. Son espérance est donc 1.

9. La variable aléatoire N' compte le nombre d'essais avant un succès dans une répétition de tirages de Bernoulli. Elle suit donc la loi géométrique $G(1/2)$. Son espérance est donc 2.

10. L'espérance est additive donc $E(N') = E(P' + F') = E(P') + E(F') = 1 + E(P')$. Donc $E(P') = 1$.

11. Dans un lancer du joueur 2, N' est égale à $P' + 1$. Donc :

$$\begin{aligned}
E\left(\frac{P'}{N'}\right) &= \sum_{n=0}^{\infty} \frac{n}{n+1} \mathbf{p}\left(\frac{P'}{N'} = \frac{n}{n+1}\right) = \sum_{n=0}^{\infty} \frac{n}{n+1} \mathbf{p}(N' = n + 1) \\
&= \sum_{n=0}^{\infty} \frac{n}{2^{n+1}(n+1)} = 1 - \log(2) \simeq 30,6\%
\end{aligned}$$

12. a) L'espérance de P est égale à l'espérance de P'. Les deux joueurs ont donc autant de chance de l'emporter l'un que l'autre.

b) L'espérance de F est égale à l'espérance de F'. Les deux joueurs ont donc autant de chance de l'emporter l'un que l'autre.

c) En revanche, $E\left(\frac{P}{N}\right) = 1/2$ tandis que $E\left(\frac{P'}{N'}\right) \simeq 0,36$. Il est donc très nettement avantageux de parier sur le premier joueur.