Chapitre 13 : Variables aléatoires finies

ECE3 Lycée Carnot

9 février 2011

Introduction

Pour introduire cette nouvelle notion, absolument fondamentale en probabilités (tellement d'ailleurs que vous n'entendrez plus parler que de ça jusqu'à la fin de l'année, à peu de choses près), prenons un exemple très classique : on lance simultanément (ou successivement, ça ne change pas grand chose) quatre pièces équilibrées. L'univers Ω des résultats de l'expérience est un ensemble à $2^4 = 16$ éléments constitué des suites de quatre Pile ou Face. On peut naturellement déjà se poser plein de questions concernant cet univers, mais il arrive qu'on ait envie de considérer des résultats qui ne sont pas directement ceux de l'expérience. Par exemple, on veut étudier plus particulièrement le nombre de Piles obtenus lors de ces quatre lancers de pièces. Ce nombre de Piles est un entier directement associé au résultat de l'expérience (si vous connaissez le résultat, vous connaissez le nombre de Piles). Eh bien, une variable aléatoire, c'est exactement ça : une application qui, à chaque élément de Ω, associe un nombre réel.

1 Variables aléatoires finies

1.1 Définition, notations

Définition 1. Soit $(\Omega, \mathcal{T})$ un espace probabilisable. Une **variable aléatoire** (réelle) X sur Ω est une application $X : \Omega \to \mathbb{R}$ telle que $\forall x \in \mathbb{R}$, $\{\omega \in \Omega | X(\omega) \leqslant x\} \in \mathcal{T}$.

Remarque 1. On note $X(\Omega)$ l'ensemble des valeurs prises par la variable aléatoire X (qui est bien l'image de l'ensemble Ω par l'application X).

Exemple : Dans l'exemple explicité en introduction (lancers de quatre pièces équilibrées), en notant X le nombre de Piles obtenus, X est une variable aléatoire, et $X(\Omega) = \{0; 1; 2; 3; 4\}$.
Exemple : L'application qui à chaque français associe sa taille est une variable aléatoire sur l'ensemble de la population française. On a ici $X(\Omega) \subset [0; 3]$ (j'ai pris large).

La définition que je viens de donner étant très générale, nous allons très rapidement nous restreindre à un cas particulier : **pour la suite du chapitre, on supposera que l'univers Ω est fini**. Dans ce cas, une variable aléatoire sera simplement une application de Ω dans $\mathbb{R}$ (et même le plus souvent dans $\mathbb{N}$), qui prendra donc nécessairement un nombre fini de valeurs (et on peut oublier la condition technique de la définition générale).

Définition 2. Soit X une variable aléatoire sur un univers Ω. On note habituellement $X = x$, l'événement $\{\omega \in \Omega | X(\omega) = x\}$. On utilisera de même la notation $X \leqslant x$ pour l'événement $\{\omega \in \Omega | X(\omega) \leqslant x\}$ (et $X \geqslant x$; $X < x$ et $X > x$ pour des évèvements similaires).

Exemple : Ainsi, si on reprend l'exemple du lancer de quatre pièces (et toujours avec X le nombre de Piles), on pourra écrire $P(X = 1) = \dfrac{4}{16} = \dfrac{1}{4}$ (il y a quatre cas sur les 16 possibles pour lesquels on obtient un seul Pile), ou encore $P(X \geqslant 3) = \dfrac{5}{16}$ (cinq cas valables sur 16).

Proposition 1. Soient X et Y deux variables aléatoires sur un même univers Ω, alors $X+Y$, XY, λX (où λ est un réel quelconque), $\max(X,Y)$ et $\min(X,Y)$ sont également des variables aléatoires.

Pas de démonstration, c'est évident, ce sont aussi des applications de Ω dans $\mathbb{R}$.

Proposition 2. Soit X une variable aléatoire sur Ω et $g : \mathbb{R} \to \mathbb{R}$ une fonction, alors $g(X) : \omega \mapsto g(X(\omega))$ est aussi une variable aléatoire (notée $g(X)$).

Exemple : Si X est une variable aléatoire, X^2 en est également une.

1.2 Loi d'une variable aléatoire

L'intérêt des variables aléatoires est bien entendu d'étudier la probabilité d'apparition de chacun des résultats possibles :

Définition 3. Soit X une variable aléatoire, la **loi de probabilité** de X est la donnée des probabilités $P(X=k)$, pour toutes les valeurs k prises par X (c'est-à-dire pour $k \in X(\Omega)$).

Remarque 2. Pour calculer la loi d'une variable aléatoire, il suffit donc de déterminer toutes les valeurs qu'elle peut prendre, puis calculer la probabilité de chaque résultat.

Exemple : Reprenons notre exemple du nombre de Piles sur quatre lancers de pièces. On peut présenter la loi de X sous la forme d'un tableau :

k	0	1	2	3	4
$P(X=k)$	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$

Proposition 3. Les événements $(X=k)_{k \in X(\Omega)}$ forment un système complet d'événements. On a donc $\sum\limits_{k \in X(\Omega)} P(X=k) = 1$.

Démonstration. Ces événements sont incompatibles (on ne peut pas avoir à la fois $X(\omega) = k$ et $X(\omega) = k'$ pour des valeurs différentes de k et k'). Leur réunion est bien Ω tout entier puisque chaque élément ω de Ω a une image par X. □

Exemple : Dans une urne se trouvent cinq jetons numérotés de 1 à 5. On en tire 3 simultanément et on note X le plus petit des trois numéros tirés. On a ici $X(\Omega) = \{1; 2; 3\}$ (si on tire trois jetons, le plus petit ne peut pas être plus grand que 3). Pour déterminer la loi, le plus simple est de dénombrer tous les cas possibles (il n'y en a que 10), même si on peut exprimer les probabilités à l'aide de coefficients binomiaux (par exmple, pour avoir $X=1$, il faut tirer le jeton 1 puis deux autres parmi les 4 restants, soit $\binom{4}{2}$ tirages favorables sur les $\binom{5}{3}$). on obtient en tout cas :

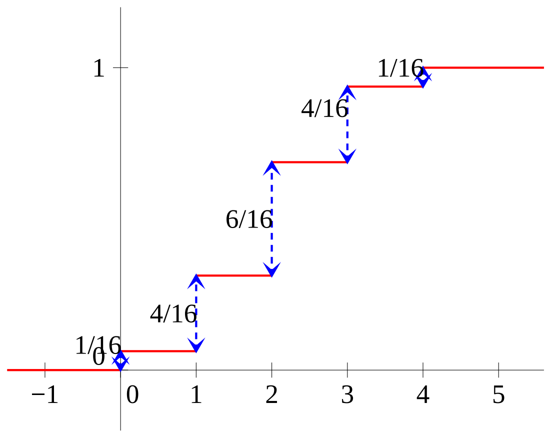
k	1	2	3
$P(X=k)$	$\frac{6}{10}$	$\frac{3}{10}$	$\frac{1}{10}$

1.3 Fonction de répartition

La fonction de répartition est simplement une autre façon de représenter la loi d'une variable aléatoire. Dans le cas des lois finies, elle n'apporte absolument aucune information supplémentaire, et son utilité est donc limitée. Mais vous verrez (surtout l'an prochain) que c'est une notion essentielle dans le cadre des variables aléatoires continues, où la représentation de la loi sous forme de tableau n'a plus de sens.

Définition 4. La **fonction de répartition** d'une variable aléatoire X est la fonction $F_X : \mathbb{R} \to \mathbb{R}$ définie par $F_X(x) = P(X \leqslant x)$.

Exemple : Reprenons notre exemple standard de lancer de quatre pièces, dont la loi a été donnée plus haut. La courbe de F_X ressemble à ceci (à chaque fois qu'on atteint une des valeurs appartenant à $X(\Omega)$, on fait un bond dont la hauteur est la probabilité correspondante) :



Proposition 4. Si X est une variable aléatoire finie, la fonction F_X est une fonction en escalier (c'est-à-dire qu'on peut découper $\mathbb{R}$ en un nombre fini d'intervalles sur lesquels la fonction est constante), dont les sauts se produisent pour les valeurs k appartenant à $X(\Omega)$ et ont pour hauteur $P(X = k)$. Dans le cas général, une fonction de répartition vérifie toujours les propriétés suivantes :

- La fonction F_x est croissante.
- $\lim_{x \to -\infty} F_X(x) = 0$ et $\lim_{x \to +\infty} F_X(x) = 1$.
- La fonction F_x est continue à droite en tout réel.

Proposition 5. Lien entre loi et fonction de répartition.
Soit X une variable aléatoire de fonction de répartition F_X. Alors

- $\forall x \in \mathbb{R}$, $F_X(x) = \sum_{k \in X(\Omega) | k \leqslant x} P(X = k)$
- dans l'autre sens, $\forall k \in X(\Omega)$, $P(X = k) = F_X(k) - \lim_{x \to k^-} F_X(x)$

On a plus généralement, pour tous réels x et y, $P(x < X \leqslant y) = F_X(y) - F_X(x)$.

Exemple : Pour mieux comprendre l'intérêt de cette notion, prenons un exemple continu (sans l'étudier en détail). On note X la variable aléatoire donnant le temps d'attente (en heures) d'un client aléatoire à un guichet de la Poste. On supposer pour fixer les idées que $X(\Omega) = [0; 4]$ (au bout de 4 heures, le client en aura vraiment assez d'attendre). Pour ce genre de variable aléatoire, la fonction F_X ne sera plus une fonction en escalier mais simplement une fonction croissante « ordinaire » (elle a par exemple toutes les chances de ne pas comporter de sauts comme dans le cas d'une variable finie, mais plutôt d'être continue). Déterminer la probabilité d'attendre entre 1 et 2 heures au guichet revient d'après la dernière formule donnée dans la propriété précédente à calculer $F_X(2) - F_X(1)$.

1.4 Moments d'une variable aléatoire

Lorsqu'on s'intéresse à une variable aléatoire pouvant prendre un grand nombre de valeurs (et même dans les autres cas !), il peut être intéressant de donner, en plus de la loi de la variable qui ne sera pas toujours une donnée très lisible, des caractéristiques d'ensemble de cette loi, comme la moyenne des valeurs prises (pondérées par leur probabilité d'apparition). Ces paramètres sont les mêmes que ceux qu'on étudie en statistiques, nous allons plus particulièrement nous intéresser à l'espérance (qui n'est autre que la moyenne évoquée plus haut, c'est un paramètre de position) et à l'écart-type (paramètre de dispersion, qui mesure la répartition des valeurs autour de l'espérance).

1.4.1 Espérance

Exemple : À un devoir, un élève de prépa ayant décidé de ne réviser qu'un sujet bien précis estime avoir :

- une chance sur 10 de tomber sur le bon sujet et d'avoir 18.
- trois chances sur 10 de ne pas tomber sur le bon sujet mais de sauver un 8 en pipotant.
- six chances sur 10 de sécher lamentablement et d'obtenir un 2 bien mérité.

La note moyenne que peut espérer avoir cet élève à son devoir est de $\frac{1}{10} \times 18 + \frac{3}{0} \times 8 + \frac{6}{10} \times 2 = 5,4$. Ce calcul est un calcul d'espérance mathématique, celle de la variable aléatoire donnant la note de l'élève à son devoir.

Définition 5. L'**espérance** d'une variable aléatoire X est définie par la formule

$$E(X) = \sum_{k \in X(\Omega)} kP(X = k)$$

Remarque 3. Il s'agit bel et bien d'un calcul de moyenne avec coefficients égaux à $P(X = k)$, la somme des coefficients valant ici 1.

Exemple : Reprenons l'exemple de quatre lancers de pièce, où X était le nombre de Pile obtenu. On aura $E(X) = 0 \times \frac{1}{16} + 1 \times \frac{4}{16} + 2 \times \frac{6}{16} + 3 \times \frac{4}{16} + 4 \times \frac{1}{16} = 2$. Le résultat est bien conforme à l'intuition qu'on a de la moyenne de la variable aléatoire X.

Exemple : Lors d'une tombola, 1000 personnes ont misé 2 euros. Il y a 100 personnes qui gagnent un lot d'une valeur de 5 euro, 10 gagnent un lot d'une valeur de 10 euros, 3 personnes gagnent un lot d'une valeur de 100 euros et enfin une personne gagne le gros lot, d'une valeur de 600 euros. Naturellement, les 886 personnes restantes ne gagnent rien. On note X la variable aléatoire correspondant au gain. On a $E(X) = 0 \times \frac{886}{1000} + 5 \times \frac{100}{1000} + 10 \times \frac{10}{1000} + 100 \times \frac{3}{1000} + 600\frac{1}{1000} = 1.5$. Autrement dit, chaque participant gagnera en moyenne 1.5 euro, ou plutôt en perdra 0.5 sur les deux qu'il avait misés. On comprend mieux sur cet exemple l'origine du terme espérance, et accessoirement la façon dont la Française des Jeux se remplit les poches.

Définition 6. Soit A un évènement inclus dans notre univers Ω. On appelle **variable indicatrice de l'évènement** A, et on note $\mathbf{1}_A$, la variable aléatoire définie par $\mathbf{1}_A(\omega) = 1$ si $\omega \in A$, et $\mathbf{1}_A(\omega) = 0$ sinon.

Proposition 6. La variable aléatoire constante $X : \omega \mapsto a$, $a \in \mathbb{R}$, a une espérance $E(X) = a$. L'espérance d'une variable aléatoire indicatrice $\mathbf{1}_A$ vaut $P(A)$.

Démonstration. C'est bien parce que je suis consciencieux que je fais une preuve. Dans le premier cas, la loi de X est simple : a avec probabilité 1. On a donc $E(X) = 1 \times a = a$ en appliquant la définition de l'espérance. Dans le second, la loi de $\mathbf{1}_A$ est à peine plus compliquée, 1 si $\omega \in A$ c'est-à-dire avec probabilité $P(A)$ et 0 sinon, donc avec probabilité $1 - P(A)$. L'espérance vaut bien $P(A)$. □

Proposition 7. Linéarité de l'espérance.
Si X et Y sont deux variables aléatoires définies sur le même univers Ω, et a, b deux réls, on a $E(aX + bY) = aE(X) + bE(Y)$. En particulier, on aura toujours $E(X + Y) = E(X) + E(Y)$; $E(aX) = aE(X)$, ou encore en utilisant l'espérance d'une variable constante calculée plus haut, $E(X + b) = E(X) + b$.

Démonstration. La preuve est un peu formelle et sera esquivée cette année. □

Exemple : Cette propriété très simple est mine de rien bien utile (c'est même la propriété fondamentale à maitriser sur l'espérance). On lance par exemple successivement 90 dés. On note X le nombre de 6 obtenus. Calculer l'espérance directement demande un certain courage (la loi de X est une horreur absolue), mais on peut ruser ! Notons A_i l'évènement « On tire un 6 au lancer numéro i » et $\mathbf{1}_{A_i}$ la variable indicatrice correspondante. On peut constater que $X = \mathbf{1}_{A_1} + \mathbf{1}_{A_2} + \dots + \mathbf{1}_{A_{90}}$ (en effet, additionner les variables indicatrices revient à ajouter 1 à chaque fois qu'un 6 sort, et 0 sinon, ce qui revient exactement à compter le nombre de 6). O a donc $E(X) = E(\mathbf{1}_{A_1}) + E(\mathbf{1}_{A_2}) + \dots + E(\mathbf{1}_{A_{90}}) = P(A_1) + P(A_2) + \dots + P(A_{90}) = \frac{1}{6} + \dots + \frac{1}{6} = \frac{90}{6} = 15$ (résultat intuitivement évident, soit dit en passant).

Définition 7. Une variable aléatoire X est dite **centrée** si $E(X) = 0$.

Proposition 8. Si X est une variable aléatoire d'espérance m, la variable aléatoire $X - m$ est centrée. On l'appelle **variable aléatoire centrée associée à** X.

Démonstration. Par linéarité, $E(X - m) = E(X) - E(m) = m - m = 0$. □

Proposition 9. Si X est une variable aléatoire positive (c'est-à-dire que $X(\Omega) \subset \mathbb{R}^+$), on a $E(X) \geqslant 0$. Si X, Y sont deux variables aléatoires telles que $X \leqslant Y$ (c'est-à-dire que $\forall \omega \in \Omega$, $X(\omega) \leqslant Y(\omega)$), alors $E(X) \leqslant E(Y)$.

Démonstration. C'est une fois de plus évident. Tous les termes intervenant dans le calcul de l'espérance de X étant positifs, la somme sera nécessairement positive. Pour la deuxième propriété, on peut utiliser une ruse classique : si $X \leqslant Y$, la variable aléatoire $Y - X$ est positive, donc $E(Y - X) \geqslant 0$. Or, $E(Y - X) = E(Y) - E(X)$, ce qui nous donne l'inégalité voulue. □

Théorème 1. (théorème du transfert) Soit X une variable aléatoire et $g : \mathbb{R} \to \mathbb{R}$ une fonction, alors on a $E(g(X)) = \sum_{k \in X(\Omega)} g(k)P(X = k)$.

Démonstration. On admettra ce résultat qui est un peu technique à prouver. C'est évident dans le cas où les images par g des valeurs k sont distinctes, mais un peu plus pénible à rédiger dans le cas général. □

Exemple : Si on cherche à calculer $E(X^2)$, il suffit de faire le calcul de somme suivant : $\sum_{k \in X(\Omega)} k^2 P(X = k)$ (autrement dit, on élève les valeurs au carré et on ne touche pas aux probabilités).

1.4.2 Moments d'ordre supérieur

Définition 8. Soit X une variable aléatoire et r un entier strictement positif, le **moment d'ordre** r de X, noté $m_r(X)$, est l'espérance de la variable aléatoire X^r. Autrement dit (en utilisant le théorème du transfert) $m_r(X) = \sum_{k \in X(\Omega)} k^r P(X = k)$.

Remarque 4. Le moment d'ordre 1 de X n'est autre que l'espérance de X.

Définition 9. La **variance** $V(X)$ d'une variable aléatoire X est le moment d'ordre 2 de la variable aléatoire centrée associée à X. Autrement dit, $V(X) = E((X - E(X))^2)$. L'écart-type σ de la variable aléatoire X est défini par $\sigma(X) = \sqrt{V(X)}$.

Que représente cette variance ? Il s'agit, techniquement, d'une moyenne de carrés d'écarts à la moyenne. Pourquoi prendre le carré ? Tout simplement car la moyenne des écarts à la moyenne est nulle. Pour réellement mesurer ces écarts, il faut « les rendre positifs », ce qui se fait bien en les élevant au carré. On pourrait également penser à prendre leur valeur absolue, mais cela aurait moins de propriétés intéressantes pour le calcul. Par contre, pour « effacer »la mise au carré, on reprend ensuite la racine carrée du résultat obtenu pour définir l'écart-type. L'écart-type représente donc (comme son nom l'indique) un écart moyen entre les valeurs prises par X et la moyenne de X (plus il est grand, plus les valeurs prises par X sont étalées).

Proposition 10. La variance d'une variable aléatoire est toujours positive.
On a la formule $V(aX+b) = a^2V(X)$.

Démonstration. La première propriété découle immédiatement de la définition du moment d'ordre 2, qui est une somme de termes positifs. Pour la deuxième, c'est du calcul un peu formel. Il faut calculer l'espérance de $(aX+b-E(aX+b))^2$. Or, par linéarité de l'espérance, $E(aX+b) = aE(X)+b$ dont l'expression précédente vaut $(aX+b-aE(X)-b)^2 = a^2(X-E(X))^2$, dont l'espérance vaut $a^2E((X-E(X))^2) = a^2V(X)$. □

Remarque 5. Une variable aléatoire a une variance (et un écart-type) nulle si et seulement si elle est constante.

Théorème 2. Théorème de König-Huygens : $V(X) = E(X^2) - (E(X))^2$.

Démonstration. C'est à nouveau un calcul très formel : $(X-E(X))^2 = X^2 - 2E(X)X + (E(X))^2$, donc, par linéarité de l'espérance, $V(X) = E((X-E(X))^2) = E(X^2) - 2E(X)E(X) + E(E(X))^2 = E(X^2) - 2E(X)^2 + (E(X))^2 = E(X^2) - (E(X))^2$. □

Remarque 6. En pratique, c'est à peu près systématiquement via la formule de König-Huygens que nous effectuerons nos calculs de variance.

Définition 10. Une variable aléatoire est dite **réduite** si son écart-type (et donc sa variance) vaut 1.

Proposition 11. Si X est une variable aléatoire, la **variable aléatoire centrée réduite associée à** X est $X^* = \dfrac{X-E(X)}{\sigma(X)}$ (qui est, vous vous en seriez doutés, centrée et réduite).

Démonstration. On a déjà vu plus haut que $X - E(X)$ était centrée, la diviser par l'écart-type ne va pas changer cela. De plus, $V\left(\dfrac{X-E(X)}{\sigma(X)}\right) = \left(\dfrac{1}{\sigma(X)}\right)^2 V(X) = 1$ □

Exemple : Pour vous montrer qu'un calcul d'écart-type à la main est en général très fastidieux, prenons l'exemple classique du lancer de deux dés, où l'on note X la somme des deux chiffres obtenus. La loi de X est donnée par le tableau suivante :

k	2	3	4	5	6	7	8	9	10	11	12
$P(X=k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Grâce à des calculs élémentaires mais pénibles, on obtient $E(X) = 7$ (logique), puis $E(X^2) = \dfrac{2^2\times 1 + 3^2 \times 2 + \cdots + 12^2 \times 1}{36} = \dfrac{1974}{36}$ donc $V(X) = E(X^2) - E(X)^2 = \dfrac{35}{6}$. L'écart-type vaut donc $\sigma(X) \simeq \sqrt{\dfrac{35}{6}} \simeq 2.415$.

2 Lois usuelles finies

Certaines lois de probabilité interviennent suffisamment régulièrement lorsqu'on étudie des variables aléatoires dans des cas classiques (lancers de dés ou de pièces, tirages de boules dans des urnes, bref toutes les bêtises qu'on aime bien vous infliger dans les exercices de probas) pour qu'il soit intéressant de les étudier une bonne fois pour toutes (et accessoirement de leur donner un nom) et d'en retenir les caractéristiques (espérance et variance notamment). Nous en étudierons quatre dans ce chapitre, et deux autres quand nous aurons étudié de façon plus approfondie les variables aléatoires infinies.

2.1 Loi uniforme

Exemple fondamental : Dans une urne se trouvent n boules numérotées de 1 à n. On en tire une au hasard et on note X le numéro obtenu.

Définition 11. On dit que la variable aléatoire X suit une **loi uniforme sur** $\{1;\dots;n\}$, et on note $X \sim \mathcal{U}(\{1;\dots;n\})$, si $X(\Omega) = \{1;\dots;n\}$ et $\forall k \in \{1;\dots;n\}$, $P(X=k) = \dfrac{1}{n}$.

Proposition 12. Si $X \sim \mathcal{U}(\{1;\dots;n\})$, on a $E(X) = \dfrac{n+1}{2}$ et $V(X) = \dfrac{n^2-1}{12}$.

Démonstration. Pour l'espérance, on a $E(X) = \displaystyle\sum_{k=1}^{k=n} k \times \frac{1}{n} = \frac{1}{n}\sum_{k=1}^{k=n} k = \frac{1}{n}\frac{n(n+1)}{2} = \frac{n+1}{2}$.

Pour la variance, on va utiliser la formule de König-Huygens. On a $E(X^2) = \displaystyle\sum_{k=1}^{k=n} k^2 \times \frac{1}{n} = \frac{1}{n} \times \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$ donc $V(X) = E(X^2) - (E(X))^2 = \dfrac{(n+1)(2n+1)}{6} - \dfrac{(n+1)^2}{4} = \dfrac{2(n+1)(2n+1) - 3(n+1)^2}{12} = \dfrac{4n^2+6n+2-3n^2-6n-3}{12} = \dfrac{n^2-1}{12}$. □

Remarque 7. À partir d'une loi uniforme prenant ses valeurs entre 1 et n, on construit facilement une loi dont la probabilité est uniforme entre deux entiers m et p (il suffit d'ajouter une constante). La loi ainsi construite a une espérance égale à $\dfrac{m+p}{2}$ et une variance égale à $\dfrac{(m-p+1)^2-1}{12}$.

2.2 Loi de Bernoulli

Exemple fondamental : On lance une pièce mal équilibrée pour laquelle la probabilité d'obtenir Pile vaut p et on note X la variable aléatoire valant 1 si on tombe sur Pile et 0 si on tombe sur face.

Définition 12. On dit que la variable aléatoire X suit une **loi de Bernoulli de paramètre** p (avec $p \in [0;1]$) si $X(\Omega) = \{0;1\}$; $P(X=1) = p$ et $P(X=0) = 1-p$. On le note $X \sim \mathcal{B}(1,p)$.

Remarque 8. Cette loi est aussi appelée loi indicatrice de paramètre p, puisqu'elle apparait essentiellement dans le cas où X est la variable aléatoire indicatrice d'un événement.

Proposition 13. Si $X \sim \mathcal{B}(1,p)$, alors $E(X) = p$ et $V(X) = p(1-p)$.

Démonstration. Pour l'espérance, on a déjà fait le calcul un peu plus haut. On a par ailleurs de la même façon $E(X^2) = p$, donc $V(X) = p - p^2 = p(1-p)$. □

Remarque 9. On utilise surtout en pratique des sommes de variables aléatoires suivant des lois de Bernoulli, comme on a déjà pu le faire dans le cas du lancer successif de 90 dés.

2.3 Loi binômiale

Exemple fondamental : Une urne contient des boules blanches et noires, avec une proportion p de boules blanches (et donc une proportion $1-p$ de boules noires). On tire n boules **avec remise** dans l'urne et on note X le nombre de boules blanches obtenues.

Définition 13. On dit qu'une variable aléatoire X suit une **loi binomiale de paramètre** (n,p) (avec $n \in \mathbb{N}^*$ et $p \in [0;1]$) si $X(\Omega) = \{0;\dots;n\}$ et $\forall k \in \{0;\dots;n\}$, $P(X=k) = \dbinom{n}{k} p^k (1-p)^{n-k}$. On le note $X \sim \mathcal{B}(n,p)$.

Remarque 10. Si $n=1$, la loi binomiale de paramètre $(1,p)$ n'est autre que la loi de Bernoulli de paramètre p, ce qui justifie l'emploi de la même notation.

Remarque 11. Une autre façon de voir une loi binômiale est de considérer que la variable aléatoire correspondante compte le nombre de réussites quand on tente n fois de suite (de façon indépendante) un tirage ayant une probabilité p de réussir.

Proposition 14. Si $X \sim \mathcal{B}(n,p)$, alors $E(X) = np$ et $V(X) = np(1-p)$ (on note parfois $q = 1-p$, auquel cas on a $V(X) = npq$).

Démonstration. On a $E(X) = \sum_{k=1}^{k=n} k\binom{n}{k}p^k(1-p)^{n-k}$. On aimerait bien appliquer le binome de Newton, mais il faut pour cela faire disparaitre le k, ce qui est par exemple possible grace à la formule $k\binom{n}{k} = n\binom{n-1}{k-1}$ On a donc $E(X) = \sum_{k=1}^{k=n} n\binom{n-1}{k-1}p^k(1-p)^{n-k} = n\sum_{j=0}^{j=n-1} p^{j+1}(1-p)^{n-1-j} = np(p+1-p)^{n-1} = np$.
Pour la variance, on ne va pas calculer $E(X^2)$ directement, mais passer par $E(X(X-1))$, ce qui va permettre d'utiliser la formule $k(k-1)\binom{n}{k} = n(n-1)\binom{n-2}{k-2}$ (obtenue en appliquant deux fois de suite la formule utilisée dans le calcul précédent). Un calcul extrêmement similaire au précédent donne alors $E(X(X-1)) = n(n-1)p^2$, donc $V(X) = E(X^2) - E(X)^2 = E(X(X-1)+X) - E(X)^2 = n(n-1)p^2 + np - n^2p^2 = np - np^2 = np(1-p)$ □

2.4 Loi hypergéométrique

Exemple fondamental : Dans une urne se trouvent N boules blanches et noires, avec une proportion p de boules blanches. On tire n boules dans l'urne **sans remise** (ou simultanément) et on note X le nombre de boules blanches obtenues.

Définition 14. On dit qu'une variable aléatoire X suit une **loi hypergéométrique de paramètre** (N,n,p) (avec $N \in \mathbb{N}^*$, $1 \leqslant n \leqslant N$ et $p \in [0;1]$) si $X(\Omega) = \{max(0, n-Nq); \dots ; min(n, Np)\}$ (où on a noté $q = 1-p$) et $\forall k \in \{max(0, n-Nq); \dots ; min(n, Np)\}$, $P(X = k) = \dfrac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}}$. On le note $X \sim \mathcal{H}(N,n,p)$.

Proposition 15. Si $X \sim \mathcal{H}(N,n,p)$, alors $E(X) = np$ et $V(X) = npq\dfrac{N-n}{N-1}$.

Démonstration. Pour simplifier les notations, tous les coefficients binomiaux faisant intervenir des entiers négatifs seront considérés comme nuls. On utilise le même type d'astuce que pour la loi binomiale :

$$E(X) = \sum_{k=0}^{n} k\frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}} = \frac{Np}{\binom{N}{n}}\sum_{k=1}^{n}\binom{Np-1}{k-1}\binom{Nq}{n-k} = \frac{Np}{\binom{N}{n}}\sum_{j=0}^{n-1}\binom{Np-1}{j}\binom{Nq}{n-1-j}$$

On peut maintenant appliquer la formule de Vandermonde à notre somme et on obtient

$$E(X) = Np\frac{\binom{Np+Nq-1}{n-1}}{\binom{N}{n}} = Np\frac{\binom{N-1}{n-1}}{\binom{N}{n}} = NP \times \frac{(N-1)!}{(n-1)!(N-n)!} \times \frac{(N-n)!}{N!n!} = Np\frac{n}{N} = np$$

Pour la variance, on utilise à nouveau les mêmes astuces. On commence par calculer

$$E(X(X-1)) = Np(Np-1)\frac{\binom{N-2}{n-2}}{\binom{N}{n}} = Np(Np-1)\times\frac{(N-2)!}{(n-2)!(N-n)!}\times\frac{n!(N-n)!}{N!} = \frac{p(Np-1)n(n-1)}{N-1}$$

comme ci-dessus, puis

$$V(X) = E(X(X-1)) + E(X) - E(X)^2 = \frac{p(Np-1)n(n-1)}{N-1} + np - n^2p^2$$

$$= np\frac{(Np-1)(n-1)+N-1-np(N-1)}{N-1} = np\frac{nNp-Np-n+1+N-1-nNp+np}{N-1}$$

$$= np\frac{N-n-Np+np}{N-1} = np\frac{(1-p)(N-n)}{N-1} = npq\frac{N-n}{N-1}$$

$\square$