

Chapitre 17 : Variables aléatoires

PTSI B Lycée Eiffel

28 avril 2021

Un médecin annonce à un de ses patients :

« J'ai une bonne et une mauvaise nouvelle, je commence par la mauvaise.

Vous avez une maladie grave dont on ne guérit qu'avec une probabilité $\frac{1}{10}$. »

« Et la bonne nouvelle docteur ? »

« Mes neuf derniers patients sont morts... »

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 pourra ainsi noter PFPF le résultat de l'expérience consistant à tirer Pile au premier lancer, à nouveau Pile au deuxième, Face au troisième et Pile au quatrième). 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. Ces variables aléatoires donnent fréquemment lieu à des développements mathématiques faisant intervenir des techniques de calcul qui pour nous se limiteront cette année aux calculs de sommes (et l'an prochain aux calculs de séries, quand vous aurez étendu la notion de variable aléatoire à des univers infinis). Par ailleurs, certaines situations classiques en probabilités donnent lieu à l'apparition de variables aléatoires usuelles (schémas de Bernoulli notamment) pour lesquelles nous démontrerons quelques résultats dans ce chapitre qui nous éviteront de refaire en permanence les mêmes calculs (d'espérance ou de variance notamment).

Objectifs du chapitre :

- être capable d'appliquer ses connaissances en dénombrement pour déterminer correctement la loi d'une variable aléatoire.
- maîtriser les techniques de calcul de l'espérance et de la variance.
- savoir repérer sans hésitation les variables aléatoires suivant une loi binômiale, et celles qui n'en suivent pas une.

1 Variables aléatoires finies.

Comme dans le chapitre de probabilités précédent, on se contentera de travailler sur des univers finis dans tout ce chapitre.

1.1 Définition, notations.

Définition 1. Soit Ω l'univers associé à une expérience aléatoire. Une **variable aléatoire** (réelle) X sur Ω est une application $X : \Omega \rightarrow \mathbb{R}$.

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 : 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). Sur le même univers, on pourrait définir une autre variable aléatoire associant par exemple à chaque français son groupe de musique préféré. Mais ce ne serait bien sûr pas une variable aléatoire réelle (les résultats ne seraient même pas numériques) et n'aurait donc aucun intérêt mathématique (calculer la moyenne du groupe de musique préféré des français n'aurait par exemple absolument aucun sens, alors que calculer une taille moyenne est nettement plus envisageable).

Les trois exemples numérotés qui suivent seront repris dans l'ensemble du chapitre pour illustrer les différentes notions abordées et donner des exemples de calculs (d'espérance et de variance notamment).

Exemple 1 : On lance deux dés équilibrés à six faces, et on note X la somme des résultats obtenus sur les deux dés. On aura $X(\Omega) = \{2, 3, \dots, 12\}$.

Exemple 2 : Dans l'exemple explicité en introduction (lancer 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 3 : 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 simultanément, le plus petit ne peut pas être plus grand que 3, ce serait bien sûr différent si on effectuait des tirages successifs avec remise).

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

Exemple : Ainsi, si on reprend l'exemple du lancer de quatre pièces (et toujours avec X le nombre de Piles), on pourra écrire $P(X = 1)$ pour désigner la probabilité de l'événement « On a obtenu exactement une fois Pile lors de nos quatre lancers de pièce », raccourci extrêmement pratique. Ici, en l'occurrence, $P(X = 1) = \frac{4}{16} = \frac{1}{4}$ (il y a quatre cas sur les 16 possibles pour lesquels on obtient un seul Pile, il faut choisir quelle pièce est tombée sur Pile). De façon similaire, $P(X \geq 3) = \frac{5}{16}$ (cinq cas valables sur 16, les quatre où on a obtenu trois Piles et un Face et l'unique tirage ayant donné quatre Piles).

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. Plus généralement, X^2 , X^3 ou n'importe quelle expression de la forme $f(X)$, où f est une fonction définie sur $X(\Omega)$, est une variable aléatoire sur Ω .

Pas de démonstration, c'est évident, ce sont aussi des applications de Ω dans $\mathbb{R}$. Il faut simplement avoir à l'esprit qu'on peut « effectuer des calculs » avec des variables aléatoires, ce qui interviendra très régulièrement dans les calculs de variance.

1.2 Loi d'une variable aléatoire.

Pour pouvoir dire des choses intéressantes sur une variable aléatoire, la première chose à étudier est 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 de calculer la probabilité de chaque résultat. On présente souvent les résultats sous forme d'un tableau, comme nous allons le faire en reprenant nos trois exemples cités plus haut. Dans certains cas, on peut également donner une formule exprimant $P(X = k)$ en fonction de k , mais une telle formule valable pour toutes les valeurs possibles de k n'est pas toujours évidente à déterminer (quand il y en a une !). Ce sera le cas des variables aléatoires usuelles que nous étudierons à la fin de ce chapitre.

Exemple 1 : Pour le lancer de deux dés, la somme des résultats obtenus a pour loi :

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

Démonstration. C'est un calcul de probabilités tout à fait classique. Lorsqu'on lance deux dés à six faces, on a $6^2 = 36$ résultats équiprobables possibles (en considérant bien sûr ici comme résultats les paires d'entiers compris entre 1 et 6, par exemple (2, 4) ou (5, 5)). Sur ces 36 résultats, un seul donne une valeur de X égale à 2 (on a tiré un double 1), deux donnent une valeur de X égale à 3 (on a tiré un 1 sur un des deux dés et un 2 sur l'autre, ce qui correspond bien à deux cas différents), etc. $\square$

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

Démonstration. On a déjà fait une bonne partie du travail plus haut : il y a 16 résultats différents possibles, un seul (FFFF) qui correspond à l'événement $X = 0$, quatre qui donnent $X = 1$ (on doit choisir la pièce qui tombe sur Pile) etc. Les plus observateurs d'entre vous remarqueront qu'on peut ici trouver une formule générale : $P(X = k) = \frac{\binom{4}{k}}{16}$. On retrouvera ce type de formule plus loin dans le chapitre quand nous définirons les lois binomiales. $\square$

Exemple 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 exemple, 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 $\binom{5}{3}$, donc 6 tirages sur 10). On obtient en tout cas :

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

Proposition 2. Les événements $(X = k)_{k \in X(\Omega)}$ forment un système complet d'événements. On a donc $\sum_{k \in X(\Omega)} P(X = k) = 1$.

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 . $\square$

2 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 (imaginez par exemple qu'on reprenne l'étude de la somme des résultats d'un lancer de dés de notre exemple 1, mais avec dix dés au lieu de deux, bon courage pour trouver la loi de X), 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, et se rangent dans deux catégories donnant des informations complémentaires :

- les paramètres de **position** donnent une idée globale des valeurs prises par la variable aléatoire. Dans cette catégorie se situent la moyenne (qu'on privilégiera en pratique dans nos calculs) mais aussi par exemple la médiane (valeur qui sépare en deux parties égales l'ensemble des valeurs prises par la variable aléatoire, en tenant compte des probabilités d'apparition de ces valeurs). Dans certains cas, connaître la médiane est plus pertinent que la moyenne (par exemple, si on étudie les salaires des français, un calcul de moyenne est peu représentatif car les plus gros salaires sont tellement énormes qu'ils contribuent « trop » à augmenter la moyenne, quand la médiane sera plus représentative du salaire d'un français « moyen »).
- les paramètres de **dispersion** servent à mesurer la répartition des valeurs prises par la variable aléatoire autour du paramètre de position (moyenne ou médiane). C'est le cas de l'écart-type (associé en général au calcul de moyenne) ou de l'écart absolu moyen (associé au calcul de médiane, que nous n'étudierons pas). À moyenne fixée, plus l'écart-type est gros, plus les valeurs prises par la variable sont (en moyenne !) « loin » de la moyenne.

2.1 Espérance.

Définition 4. L'**espérance** d'une variable aléatoire X est définie par $E(X) = \sum_{k \in X(\Omega)} kP(X = k)$.

Remarque 3. Il s'agit bel et bien d'un calcul de moyenne pondérée, les coefficients étant égaux aux probabilités $P(X = k)$. La somme des coefficients vaut donc 1, ce qui explique qu'il n'y ait pas de division à faire pour calculer cette moyenne. Si on dispose de la loi de X sous forme de tableau, le calcul de l'espérance est donc très simple, on multiplie chaque valeur par la probabilité qui se trouve juste en-dessous et on fait la somme.

Exemple 1 : On calcule un peu laborieusement l'espérance de la somme de nos deux chiffres lors du lancer de dés : $E(X) = \frac{2 + 6 + 12 + 20 + 30 + 42 + 40 + 36 + 30 + 22 + 12}{36} = \frac{252}{36} = 7$. Ce résultat tout simple vous semble logique ? On reviendra dessus un peu plus loin (il y a une méthode plus simple pour retrouver le résultat). On ne pourrait d'ailleurs sûrement pas calculer par cette méthode l'espérance de la somme des chiffres obtenus lorsqu'on lance dix dés (on serait incapables de donner la loi) alors qu'elle devrait logiquement valoir 35.

Exemple 2 : Reprenons l'exemple de quatre lancers de pièce, où X était le nombre de Piles 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 3 : Pour nos cinq jetons dans l'urne, l'espérance vaut $E(X) = \frac{6 + 6 + 3}{10} = \frac{3}{2}$. Tout ce qu'on peut constater c'est que c'est cohérent puisque l'espérance se situe entre les valeurs extrêmes prises par la variable X (et plus proche de 1 qui est la valeur qu'on obtiendra la majorité du temps).

Proposition 3. La variable aléatoire constante $X : \omega \mapsto a, a \in \mathbb{R}$, a une espérance $E(X) = a$.

Démonstration. C'est évidemment complètement trivial puisque dans ce cas la variable (qui ne l'est pas vraiment) prend comme unique valeur a , avec probabilité 1. En fait, cette propriété sert simplement à supprimer les espérances quand elles sont autour d'une constante dans certains calculs formels. $\square$

Définition 5. Soit A un évènement inclus dans notre univers Ω . La **variable indicatrice de l'évènement** A , notée $\mathbf{1}_A$, est la variable aléatoire définie par $\mathbf{1}_A(\omega) = 1$ si $\omega \in A$, et $\mathbf{1}_A(\omega) = 0$ sinon.

Proposition 4. L'espérance de la variable aléatoire indicatrice $\mathbf{1}_A$ vaut $P(A)$.

Démonstration. C'est bien parce que je suis consciencieux que je fais une preuve. La loi de $\mathbf{1}_A$ est extrêmement simple, elle prend la valeur 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 donc bien $P(A)$. En fait, ces variables indicatrices jouent un rôle essentiel dans le principe des schémas de Bernoulli que nous évoquerons dans la dernière partie du cours, et permettent aussi de calculer plus simplement certaines espérances, comme nous allons le voir ci-dessous. $\square$

Théorème 1. Linéarité de l'espérance.

Si X et Y sont deux variables aléatoires définies sur le même univers Ω , et a, b deux réels, on a $E(aX + bY) = aE(X) + bE(Y)$. En particulier, on aura toujours $E(X + Y) = E(X) + E(Y)$, et $E(aX) = aE(X)$, 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 donc esquivée. Cette linéarité est toutefois « évidente » d'un point de vue intuitif. Prenons un exemple classique et toujours réjouissant, celui des notes qu'on attribue à nos chers élèves. Lors d'une interro de maths particulièrement peu révisée, j'obtiens à la fin de ma correction une moyenne brute égale à 5 sur 20. Je décide dans un élan de générosité d'ajouter trois points à tout le monde, quel effet cela aura-t-il sur la moyenne ? Ça va bien sûr l'augmenter de trois points. C'est exactement ce que dit la propriété $E(X + b) = E(X) + b$. Si au lieu d'augmenter tout le monde de trois points, je décide carrément de doubler toutes les notes, la moyenne sera multipliée par deux. C'est ce qu'affirme la formule $E(aX) = aE(X)$. Enfin, si je décide de faire une deuxième interro pour rattraper la première, deuxième interro où on atteint une moyenne de 7, et que j'additionne les deux notes pour en former une nouvelle, la moyenne de la classe sera de 12. C'est ce qui découle de la dernière formule $E(X + Y) = E(X) + E(Y)$. $\square$

Exemple : Cette propriété très simple est mine de rien extrêmement utile (c'est même la propriété fondamentale à maîtriser sur l'espérance). On lance par exemple successivement 100 dés équilibrés à six faces (tant qu'à faire, maintenant qu'on a une méthode efficace, on peut se le permettre). On note X la somme des résultats obtenus. Calculer l'espérance directement demande un certain courage (la loi de X est une horreur absolue), mais on peut en fait s'en sortir beaucoup plus simplement ! Notons X_i la variable aléatoire donnant le résultat du lancer du dé numéro i . On peut constater

que $X = X_1 + X_2 + \cdots + X_{100}$. Or, toutes les variables X_i ont la même espérance, celle de la variable donnant le résultat d'un lancer de dé à six faces, qui vaut $\frac{7}{2}$ (cette variable prend chacune des valeurs 1, 2, 3, 4, 5 et 6 avec probabilité $\frac{1}{6}$, je vous laisse détailler le calcul si nécessaire). On a donc $E(X) = E(X_1 + X_2 + \cdots + X_{100}) = E(X_1) + E(X_2) + \cdots + E(X_{100}) = \frac{7}{2} + \cdots + \frac{7}{2} = 350$ (résultat intuitivement évident, soit dit en passant).

Proposition 5. Une variable aléatoire positive (c'est-à-dire que $X(\Omega) \subset \mathbb{R}^+$) a nécessairement une espérance positive.

Si X, Y sont deux variables aléatoires telles que $X \leq Y$ (c'est-à-dire que $\forall \omega \in \Omega, X(\omega) \leq Y(\omega)$), alors $E(X) \leq E(Y)$.

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 $X \leq Y$, la variable aléatoire $Y - X$ est positive, donc $E(Y - X) \geq 0$. Or, par linéarité de l'espérance, $E(Y - X) = E(Y) - E(X)$, ce qui nous donne l'inégalité voulue. On remarquera en passant que les propriétés fondamentales de l'espérance (linéarité, positivité) rappellent fortement celles de l'intégrale. Ce n'est bien entendu pas du tout un hasard. $\square$

Définition 6. Une variable aléatoire X est **centrée** si $E(X) = 0$.

Proposition 6. 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, $E(X - m) = E(X) - E(m) = m - m = 0$. $\square$

Théorème 2. (théorème du transfert) Soit X une variable aléatoire et $g : \mathbb{R} \rightarrow \mathbb{R}$ 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. $\square$

Remarque 4. Ce théorème signifie qu'on peut calculer très facilement des espérances de variables aléatoires obtenues en appliquant une fonction à la variable X . Ainsi, par exemple, $E(X^2) = \sum_{k \in X(\Omega)} k^2 P(X = k)$ (autrement dit, on élève les valeurs au carré et on garde les probabilités intactes par rapport au calcul de l'espérance de X). Un principe simple à retenir quand on fait ce genre de calcul : on ne modifie jamais les probabilités, mais uniquement les valeurs, dans le calcul d'espérance.

2.2 Variance et moments d'ordre supérieur.

Définition 7. 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 5. Le moment d'ordre 1 de X n'est autre que l'espérance de X .

Définition 8. 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'écart 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 7. La variance d'une variable aléatoire est toujours positive.
On a la formule $V(aX + b) = a^2 V(X)$.

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 $a^2 E((X - E(X))^2) = a^2 V(X)$. $\square$

Remarque 6. Une variable aléatoire a une variance (et un écart-type) nulle si et seulement si elle est constante.

Théorème 3. 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, $V(X) = E((X - E(X))^2) = E(X^2) - 2E(X)E(X) + E(E(X))^2 = E(X^2) - 2E(X)^2 + (E(X))^2 = E(X^2) - (E(X))^2$. $\square$

Remarque 7. En pratique, c'est à peu près systématiquement via la formule de König-Huygens que nous effectuerons nos calculs de variance, car il est plus simple de calculer $E(X^2)$ que de calculer directement $E((X - E(X))^2)$.

Définition 9. Une variable aléatoire est dite **réduite** si son écart-type (et donc sa variance) vaut 1.

Proposition 8. Si X est une variable aléatoire, la **variable aléatoire centrée réduite associée à X** est $X^* = \frac{X - E(X)}{\sigma(X)}$ (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, $V\left(\frac{X - E(X)}{\sigma(X)}\right) = \left(\frac{1}{\sigma(X)}\right)^2 V(X) = 1$. $\square$

Exemple 1 : 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. Un calcul pénible donne $E(X^2) = \frac{2^2 \times 1 + 3^2 \times 2 + \dots + 12^2 \times 1}{36} = \frac{1974}{36}$ donc (formule de König-Huygens) $V(X) = E(X^2) - E(X)^2 = \frac{35}{6}$. L'écart-type vaut donc $\sigma(X) \simeq \sqrt{\frac{35}{6}} \simeq 2.415$.

Exemple 2 : Dans le cas du lancer de quatre pièces, c'est moins affreux. On a $E(X^2) = \frac{4 + 24 + 36 + 16}{16} = \frac{80}{16} = 5$, donc $V(X) = 5 - 4 = 1$. Autrement dit, la variable X est une variable réduite (c'est une coïncidence numérique ici, il n'y a pas de raison profonde à cela).

Exemple 3 : Dernier calcul pour nos histoires de jetons dans une urne, $E(X^2) = \frac{6 + 4 \times 3 + 9}{10} = \frac{27}{10}$, donc $V(X) = \frac{27}{10} - \frac{9}{4} = \frac{9}{20}$. Sans surprise, on obtient ici un écart-type faible puisque les valeurs que peut prendre X sont limitées, et que la variable va en pratique prendre presque tout le temps la valeur 1 ou 2 (dont être à une « distance » $\frac{1}{2}$ de sa moyenne).

2.3 Inégalité de Bienaymé-Tchebychev.

Théorème 4. Inégalité de Bienaymé (Irénee-Jules) - Tchebychev (Pafnouti).

Soit X une variable aléatoire d'espérance m et d'écart-type σ , alors, pour tout réel strictement positif a , on a l'inégalité $P(|X - m| \geq a) \leq \frac{\sigma^2}{a^2}$.

Démonstration. La preuve découle d'une autre inégalité connue sous le nom d'inégalité de Markov : si $X(\Omega) \subset \mathbb{R}^+$ et $a > 0$, $P(X \geq a) \leq \frac{E(X)}{a}$. C'est essentiellement évident : si ce n'était pas le cas, X prendrait des valeurs plus grandes que a avec une probabilité plus grande que $\frac{E(X)}{a}$, ce qui contribuerait à l'espérance pour une quantité plus grande que $E(X)$, ce n'est pas possible quand on n'ajoutera ensuite plus que des valeurs positives dans le calcul de l'espérance. Appliquons donc cette inégalité à la variable $(X - m)^2$, qui est positive, et à a^2 . On sait que $E((X - m)^2) = V(X) = \sigma^2$, donc $P((X - m)^2 \geq a^2) \leq \frac{\sigma^2}{a^2}$. Or, la probabilité de gauche est la même que celle de $|X - m| \geq a$, ce qui prouve l'inégalité de Bienaymé-Tchebychev. $\square$

Remarque 8. Cette inégalité permet de majorer la probabilité d'obtenir des valeurs éloignées de la moyenne pour une variable aléatoire (ce qui a notamment un lien avec la notion d'intervalle de confiance en statistiques), mais elle est hélas très imprécise, et ne donne pas de bons résultats en pratique sauf pour des valeurs réellement extrêmes de la variable aléatoire. En fait, son principal intérêt est de permettre de démontrer rapidement la loi faible des grandes nombres, mais cette dernière n'est pas à votre programme cette année !

3 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 trois dans ce chapitre (dont l'une est un cas particulier d'une autre), et vous en verrez quelques autres l'an prochain (sur des univers infinis).

3.1 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 10. 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 $\forall k \in \{1, \dots, n\}$, $P(X = k) = \frac{1}{n}$.

Proposition 9. Si $X \sim \mathcal{U}(\{1, \dots, n\})$, on a $E(X) = \frac{n+1}{2}$ et $V(X) = \frac{n^2-1}{12}$.

Démonstration. Pour l'espérance, on a $E(X) = \sum_{k=1}^{k=n} k \times \frac{1}{n} = \frac{1}{n} \sum_{k=1}^{k=n} k = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}$. Pour la variance, on va comme toujours utiliser la formule de König-Huygens. On a $E(X^2) = \sum_{k=1}^{k=n} k^2 \times \frac{1}{n} = \frac{1}{n} \times \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$ donc $V(X) = E(X^2) - (E(X))^2 = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{2(n+1)(2n+1) - 3(n+1)^2}{12} = \frac{4n^2 + 6n + 2 - 3n^2 - 6n - 3}{12} = \frac{n^2-1}{12}$. $\square$

Remarque 9. 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 $\frac{m+p}{2}$ et une variance égale à $\frac{(m-p+1)^2-1}{12}$.

3.2 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 11. 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 10. 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 10. 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 $V(X) = p - p^2 = p(1 - p)$. $\square$

Remarque 11. On utilise surtout en pratique des sommes de variables aléatoires suivant des lois de Bernoulli. On peut voir une variable aléatoire de Bernoulli comme un exemple d'expérience aléatoire où on a seulement deux issues possibles, le « succès » (représenté par la valeur 1 pour la variable aléatoire) et l'« échec » (la variable prend la valeur 0). Ces variables basées sur la dualité succès-échec sont à la base du principe du schéma de Bernoulli : si on répète plusieurs fois de suite une expérience de Bernoulli (la « même » expérience au sens où la probabilité de succès doit rester toujours la même, et où les différentes expériences doivent être indépendantes les unes des autres) et qu'on compte le nombre de succès obtenus lors de ces n essais, on construit une variable aléatoire qui va suivre une loi binômiale, qui nous allons à présent décrire.

3.3 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. Par rapport à la remarque concluant le paragraphe précédent, on répète bien ici n fois une expérience ayant deux issues possibles, le succès étant ici représenté par le tirage d'une boule blanche et l'échec par le tirage d'une boule noire. Il est essentiel que les tirages s'effectuent **avec remise** pour que la probabilité de tirer une boule blanche reste la même à chaque tirage, sinon on obtiendra un tout autre type de loi.

Définition 12. 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 $\forall k \in \{0, \dots, n\}$, $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$. On le note $X \sim \mathcal{B}(n, p)$.

Remarque 12. 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 13. Comme on l'a déjà dit, une loi binômiale est obtenue lorsqu'on compte le nombre de réussites après avoir tenté n fois de suite (de façon indépendante) une expérience aléatoire ayant une probabilité p de réussir. C'est même ainsi qu'on doit repérer les lois binômiales dans un exercice. Attention tout de même, ce n'est pas parce qu'un schéma de Bernoulli est présent dans l'énoncé que toutes les variables aléatoires vont nécessairement suivre des lois binômiales, on peut très bien créer d'autres variables à partir de cette même situation. Par exemple, on lance dix fois de suite une pièce (équilibrée) à Pile ou Face. Si on note X le nombre de Faces obtenues, la variable X suivra une loi binômiale de paramètre $\left(10, \frac{1}{2}\right)$ (on répète dix fois une expérience ayant une probabilité $\frac{1}{2}$ de réussir et on compte simplement le nombre de succès). Mais si, à partir de la même expérience, on note X le numéro du lancer où on a obtenu pour la première fois Pile (en admettant qu'on peut poser $X = 0$ si on n'obtient que des Faces), la variable Y ne suivra pas du tout une loi binômiale.

Proposition 11. 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 $E(X) = \sum_{k=1}^n k \binom{n}{k} p^k (1 - p)^{n-k}$. On aimerait bien appliquer le binôme de Newton, mais il faut pour cela faire disparaître le k , ce qui est par exemple possible grâce à la formule sans

nom $k \binom{n}{k} = n \binom{n-1}{k-1}$. On a donc $E(X) = \sum_{k=1}^{k=n} n \binom{n-1}{k-1} p^k (1-p)^{n-k} = n \sum_{j=0}^{j=n-1} \binom{n-1}{j} p^{j+1} (1-p)^{n-1-j} = np(p+1-p)^{n-1} = np$.

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 sans nom). Un calcul extrêmement similaire au précédent donne alors $E(X(X-1)) = n(n-1)p^2$, donc $V(X) = E(X^2) - E(X)^2 = E(X(X-1) + X) - E(X)^2 = n(n-1)p^2 + np - n^2p^2 = np - np^2 = np(1-p)$. $\square$

Remarque 14. La somme de n variables suivant chacune de façon indépendante une loi de Bernoulli de paramètre p sera toujours une variable binômiale de paramètre (n, p) . C'est même exactement le principe du schéma de Bernoulli.

Remarque 15. Pour les curieux, que se passe-t-il si on reprend le dispositif de notre exemple fondamental (une urne contient des boules blanches en proportion p et noires en proportion $1-p$) mais qu'on effectue n tirages successifs **sans** remise (en notant toujours X le nombre de boules blanches tirées après ces n tirages)? On obtient en fait un autre type de loi classique (mais qui n'est pas à votre programme!), appelée loi hyper-géométrique (quel beau nom). On a aussi des formules pour l'espérance et la variance de telles lois, qui se démontrent en utilisant une belle formule de dénombrement (pas à votre programme non plus!) appelée formule de Vandermonde. Si jamais vous tombez sur une variable de ce type dans un exercice, vous referez bien sûr tous les calculs à la main.