Chapitre 2

Variables Aléatoires Discrètes

2.1 Notion de variable aléatoire

Définition 2.1. *On appelle* variable aléatoire *(abrégée* v.a.*) le résultat d'une épreuve aléatoire lorsque l'issue de celle-ci peut être représentée par un nombre. Elle est dite* discrète *lorsqu'elle ne prend que des valeurs en* nombre fini *ou* infini dénombrable *et* continue *lorsque l'ensemble des valeurs prises couvre* $\mathbb{R}$ *tout entier ou une partie de* $\mathbb{R}$ *(donc en* nombre infini *et* non dénombrable*).*

Une variable aléatoire est généralement désignée par une lettre majuscule X, Y, Z, etc. et peut également être définie en tant qu'application depuis l'univers Ω dans l'ensemble des valeurs qu'elle peut prendre, noté $X(\Omega)$:

$$X \;:\; \begin{array}{ccc} \Omega & \longrightarrow & X(\Omega) \\ \omega & \longmapsto & x \end{array},$$

où $\omega \in \Omega$ représente une réalisation particulière de l'épreuve en question. Pour une v.a. discrète, l'ensemble des valeurs prises $X(\Omega)$ est *fini* ou *dénombrable*. Dans la plupart des applications et des exemples, $X(\Omega)$ sera une partie de $\mathbb{N}$ (ensemble des entiers naturels), voire $\mathbb{N}$ tout entier (dans le cas d'une variable aléatoire prenant un ensemble infini dénombrable de valeurs).

Exemple 2.1.

1. *Résultat d'un jet de dé à 6 faces.*
   Si l'on note X le résultat du lancer, il s'agit alors d'une variable aléatoire discrète à valeurs dans $X(\Omega) = \{1, 2, 3, 4, 5, 6\}$. Dans ce cas particulier, $X(\Omega)$ est confondu avec l'univers Ω, et l'application définissant la v.a. est l'identité. C'est un cas particulier, puisque l'expérience aléatoire sous-jacente fourni déjà un nombre.
2. *Lancer de 2 pièces de monnaie.*
   Pour le jet de 2 pièces identiques, dont les issues sont notées P (pour Pile) et F (pour Face), on considère l'univers

   $$\Omega = \{PP, PF, FP, FF\},$$

   où la première (resp. deuxième) lettre représente l'issue du premier (resp. deuxième) jet, obtenant ainsi tous les événements élémentaires de l'expérience. Cet univers n'est pas composé de nombres mais par exemple le nombre de fois où Pile (P) est apparu constitue une variable aléatoire :

$$X : \Omega \longrightarrow X(\Omega) = \{0, 1, 2\} \subset \mathbb{N}.$$

Il s'agit d'une variable discrète ne prenant qu'un nombre fini de valeurs (0,1 ou 2 : pour $\omega = FF$, on a $X(\omega) = X(FF) = 0$, et pour les autres réalisations, on a $X(PF) = X(FP) = 1$, $X(PP) = 2$).

3. *Tirage du loto.*
   Le résultat d'un tirage classique du loto est une variable aléatoire discrète X à valeurs dans $\{0, \dots, 49\}$ pour le premier numéro. À chaque réalisation ω de l'expérience, on obtient un entier naturel $X(\omega)$ compris entre 0 et 49.

On note génériquement x_i les valeurs prises par X. Ainsi l'ensemble des valeurs possibles pour X s'écrit $X(\Omega) = \{x_1, \dots, x_n\}$ ou $X(\Omega) = \{x_i, i = 1, \dots, n\}$ pour une variable aléatoire prenant n différentes valeurs, ou encore $X(\Omega) = \{x_i, i \geq 1\}$ pour une infinité dénombrable de valeurs possibles.

Les événements $\{X = x_i\}$, engendrés par ces différentes valeurs prises, constituent les *événements élémentaires de la variable aléatoire* X. Dans le cas ci-dessus du lancer de deux pièces, les événements élémentaires seront ainsi notés $\{X = 0\}$ (« Aucun *Pile* n'a été tiré »), $\{X = 1\}$ (« Un *Pile* – et un seul – a été tiré au cours des deux lancers ») et $\{X = 2\}$ (« Deux *Piles* ont été tirés »).

Remarque 2.1. Les événements élémentaires d'une variable aléatoire ne sont pas nécessairement équiprobables, même si les événements élémentaires de Ω sont équiprobables.

2.2 Loi d'une variable aléatoire discrète

Pour caractériser le comportement aléatoire d'une v.a. discrète X, on introduira sa *loi* essentiellement de deux manières : soit directement en fournissant les probabilités des événements élémentaires, soit indirectement en s'intéressant à des données cumulées par l'intermédiaire de sa *fonction de répartition.*

Définition 2.2. *La* loi d'une variable discrète X *est une probabilité* $\mathbb{P}_X$ *définie par*

$$\begin{array}{rcl} \mathbb{P}_X : X(\Omega) & \longrightarrow & [0\,;1] \\ x & \longmapsto & \mathbb{P}_X(x) = \mathbb{P}(X = x). \end{array}$$

Selon le contexte, on note la probabilité que X prenne la valeur x de différentes manières : $\mathbb{P}(\{X = x\})$, $\mathbb{P}(X = x)$, $p(x)$ ou p_x. Pour une valeur possible x_i de X, on note aussi souvent $p_i = \mathbb{P}(X = x_i)$.

On vérifie aisément que cette application est bien une probabilité (qui vérifie donc les axiomes fondamentaux) définie sur l'univers $X(\Omega)$ constitué par l'ensemble des valeurs prises par X.

Exemple 2.2. Soit le jet d'un dé (équilibré) à 6 faces, et X le résultat d'un jet. On a $X(\Omega) = \{1,2,3,4,5,6\}$. On vérifie directement l'axiome de l'événement certain et l'additivité pour des événements disjoints :

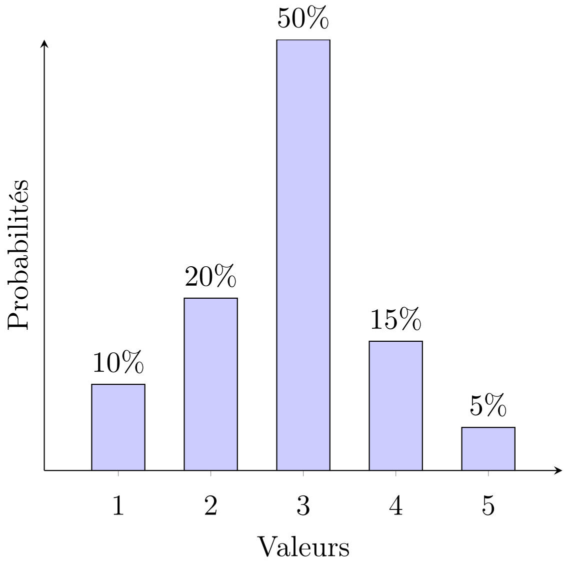
$$\begin{aligned}
\mathbb{P}_X\left(X(\Omega)\right) &= \mathbb{P}_X\left(\{1,2,3,4,5,6\}\right) = \mathbb{P}\left(X \in \{1,2,3,4,5,6\}\right) \\
&= \mathbb{P}\left(\{X=1\} \cup \cdots \cup \{X=6\}\right) \\
&= \mathbb{P}(X=1) + \cdots + \mathbb{P}(X=6) \\
&= \frac{1}{6} + \cdots + \frac{1}{6} \\
&= 1.
\end{aligned}$$

Donner la loi d'une variable aléatoire correspond à fournir les probabilités des ses événements élémentaires. Ces données peuvent être présentées sous la forme d'un *tableau* à 2 lignes avec autant de colonne que de valeurs possibles pour la v.a. : une première ligne contient les valeurs possibles $x_1, \ldots, x_n$ prises par X et une seconde ligne contient leurs probabilités respectives. Ce tableau peut aussi être représenter graphiquement sous forme d'un *diagramme en bâtons.*

Exemple 2.3. 1. Soit X une variable aléatoire à valeurs dans $\{1,2,3,4,5\}$ avec $p_1 = 0,1$, $p_2 = 0,2$, $p_3 = 0,5$, $p_4 = 0,15$ et $p_5 = 0,05$. Alors le tableau s'écrit

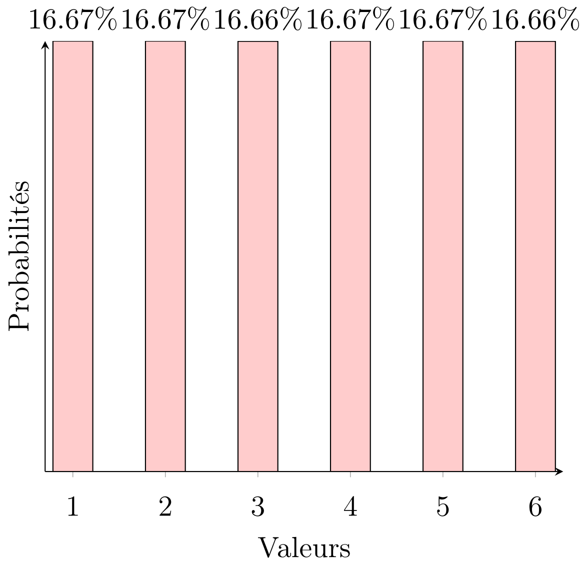
x	1	2	3	4	5
$\mathbb{P}(X=x)$	$0,1$	$0,2$	$0,5$	$0,15$	$0,05$

et son diagramme en bâtons est



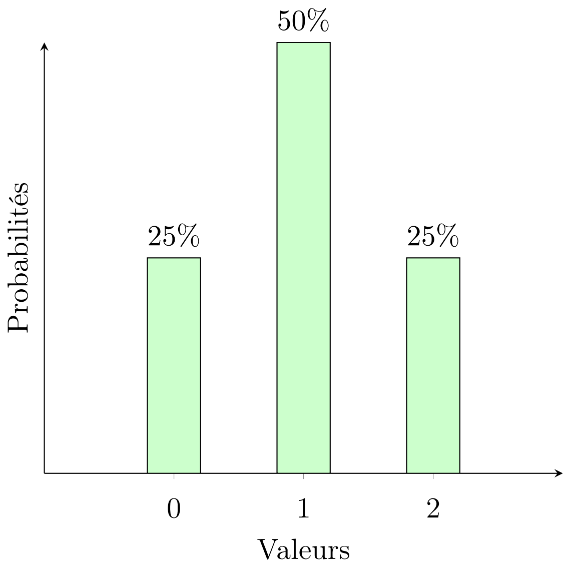
2. Pour l'exemple d'un dé à six faces équilibrées, ces probabilités sont identiques, notées $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = p$ de sorte que l'on doit avoir $6p = 1$ soit $p = \dfrac{1}{6} \approx 16,67\%$. Il s'agit en fait d'une loi *uniforme.*

x	1	2	3	4	5	6
$\mathbb{P}(X=x)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$



3. Pour les exemples des deux lancers d'une même pièce *équilibrée*, avec X comptant le nombre de *Piles*, on obtient $p_0 = \dfrac{1}{4}$, $p_1 = \dfrac{1}{2}$, $p_2 = \dfrac{1}{4}$.

x	0	1	2
$\mathbb{P}(X = x)$	1/4	1/2	1/4



2.2.1 Fonction de répartition

La loi d'une variable aléatoire X peut aussi être caractérisée à partir de sa *fonction de répartition* F_X. Il s'agit d'une fonction qui à chaque réel x associe *la probabilité que la variable aléatoire* X *soit inférieure ou égale à* x, c'est-à-dire :

$$\begin{array}{rccl} F_X & : & \mathbb{R} & \longrightarrow & [0\,;1] \\ & & x & \longmapsto & F_X(x) = \mathbb{P}(X \leq x). \end{array}$$

À chaque valeur x, $F_X(x)$ correspond au *cumul* des probabilités des événements élémentaires qui sont inférieurs ou égaux à x.

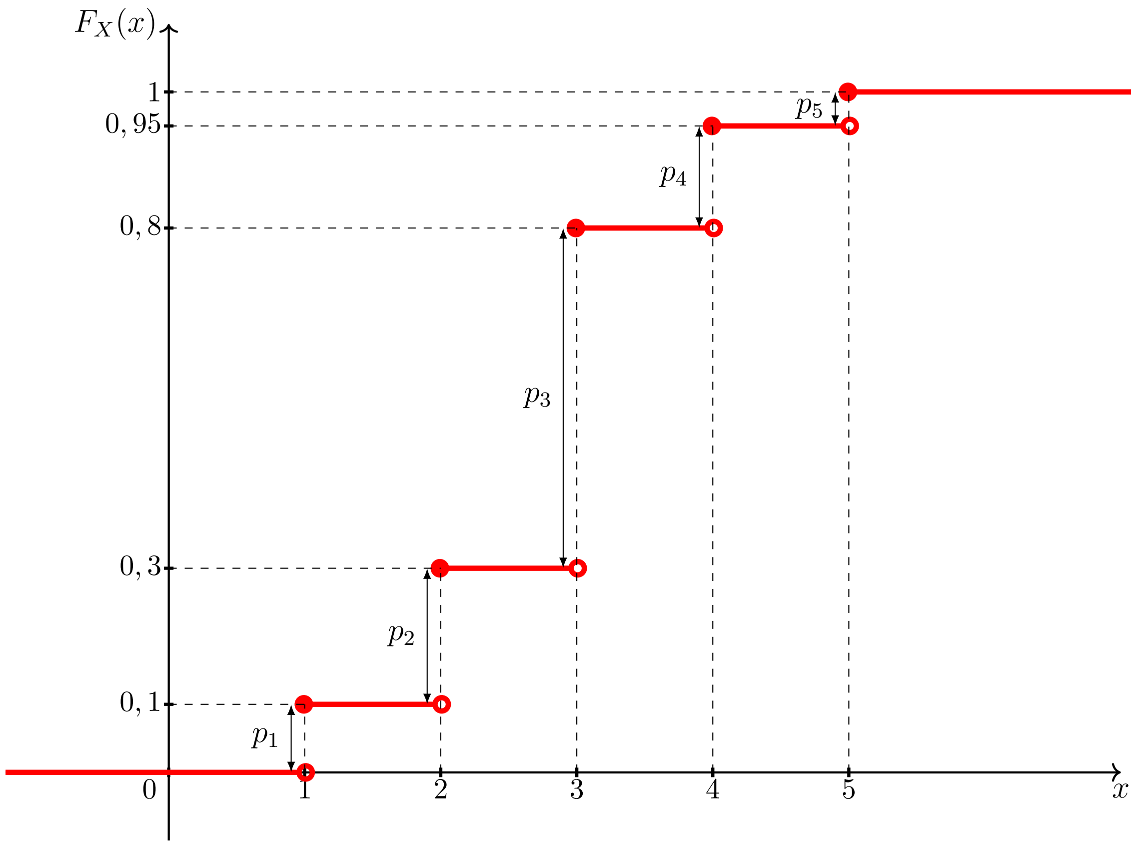
Dans le cas discret, F_X ne change de valeurs qu'aux valeurs ponctuelles $x_1, \ldots, x_n$ prises par X. Si on ordonne ces valeurs $x_1 < \cdots < x_n$, il suffit alors d'additionner les probabilités élémentaires :

$$\forall 1 \leq i \leq n, \quad F_X(x_i) = \mathbb{P}(X \leq x_i) = \mathbb{P}(X = x_1) + \cdots + \mathbb{P}(X = x_i) = p_1 + \cdots + p_i,$$

qui est inférieure à 1 si $i < n$ et égale à 1 si $i = n$.

Exemple 2.4. Soit X une variable aléatoire à valeurs dans $\{1, 2, 3, 4, 5\}$ avec $p_1 = 0,1$, $p_2 = 0,2$, $p_3 = 0,5$, $p_4 = 0,15$ et $p_5 = 0,05$. Alors sa fonction de répartition est définie par

$$F_X(x) = \mathbb{P}(X \leq x) = \begin{cases} 0 & \text{si} \quad x \in]-\infty; 1[\\ 0,1 & \text{si} \quad x \in [1; 2[\\ 0,3 & \text{si} \quad x \in [2; 3[\\ 0,8 & \text{si} \quad x \in [3; 4[\\ 0,95 & \text{si} \quad x \in [4; 5[\\ 1 & \text{si} \quad x \in [5; +\infty[\end{cases}$$



Fonction de répartition de X.

Propriété 2.1. *Si X est une v.a. discrète de fonction de répartition F_X, alors*

1. *F_X est* croissante en escalier *(càdlàg, continue à droite limite à gauche), avec*

$$\lim_{x\to-\infty} F_X(x) = 0 \quad et \quad \lim_{x\to+\infty} F_X(x) = 1.$$

2. *Pour tous $a, b \in \mathbb{R}$, $a < b$, $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$.*

2.2.2 Espérance mathématique et variance

L'*espérance mathématique* d'une variable aléatoire X est sa moyenne théorique : les valeurs prises par la v.a. sont pondérées par leurs probabilités élémentaires.

Pour une v.a. discrète X prenant n valeurs $x_1, \ldots, x_n$, on définit l'espérance $\mathbb{E}[X]$ par

$$\mathbb{E}[X] = \sum_{i=1}^{n} x_i \cdot \mathbb{P}(X = x_i) = \sum_{i=1}^{n} x_i \cdot p_i = x_1 p_1 + x_2 p_2 + \cdots + x_n p_n.$$

Exemple 2.5. Lors du lancer de 2 pièces identiques, le nombre moyen de *Piles* espéré correspond à l'espérance mathématique de la variable aléatoire X comptant le nombre de *Piles*, et vaut

$$\mathbb{E}[X] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1.$$

Exemple 2.6. Un autre exemple est celui du calcul de la moyenne d'un étudiant sur un semestre : on multiplie la note obtenue dans chaque matière par le nombre de crédits associé, on les additionne et on divise le tout par le nombre total de crédits du semestre. On fait donc la moyenne pondérée de notes de chaque matière (la pondération des notes étant égale à la proportion de crédits que chaque matière représente sur le semestre).

Propriété 2.2.

1. **Variable aléatoire *p.s.* constante** *: si Z est une variable prenant la valeur c avec probabilité 1, alors $\mathbb{E}[Z] = c$.*

2. **Linéarité de l'espérance** *: si X est une variable aléatoire définie sur un univers Ω et a,b deux réels, alors*

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b.$$

En particulier $\mathbb{E}[2X] = 2\mathbb{E}[X]$, $\mathbb{E}\left[\frac{1}{2}X\right] = \frac{1}{2} \cdot \mathbb{E}[X]$, $\mathbb{E}[aX] = a\mathbb{E}[X]$, *etc.*

3. **L'espérance est croissante** *: si X et Y sont deux variables aléatoires définies sur le même univers Ω et que $X \leq Y$ avec probabilité 1, alors $\mathbb{E}[X] \leq \mathbb{E}[Y]$.*

4. **Inégalité de Markov** *: Soit X une v.a. d'espérance finie. Alors*

$$\forall a > 0, \quad \mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}.$$

5. **Formule de transfert discrète** *(ou « de changement de variable ») : Si X est une variable aléatoire discrète et g une fonction à valeurs réelles définie sur $X(\Omega)$, alors $Y = g(X)$ est aussi une variable aléatoire définie sur le même univers Ω et on peut déterminer la loi de Y si on connaît celle de X. L'espérance de Y peut être calculée à partir de sa loi, mais également directement à partir de celle de X grâce à la formule de transfert :*
$$\mathbb{E}[Y] = \mathbb{E}[g(X)] = \sum_{x \in X(\Omega)} g(x)\mathbb{P}_X(x).$$

Exemple 2.7. Soient X une v.a. et $Y = X^2 = g(X)$, où $g(x) = x^2$. Alors
$$\mathbb{E}[Y] = \sum_{y \in Y(\Omega)} y \cdot \mathbb{P}_Y(y) = \sum_{x \in X(\Omega)} x^2 \cdot \mathbb{P}_X(x).$$

Pour déterminer la loi de Y, on le fait à partir de celle de X comme suit : pour $y < 0$, on a $\mathbb{P}_Y(y) = \mathbb{P}(Y = y) = 0$, tandis que pour $y \geq 0$, on a
$$\begin{aligned} \mathbb{P}_Y(y) &= \mathbb{P}(Y = y) = \mathbb{P}(|X| = \sqrt{y}) \\ &= \mathbb{P}(\{X = \sqrt{y}\} \cup \{X = -\sqrt{y}\}) \\ &= \mathbb{P}(X = \sqrt{y}) + \mathbb{P}(X = -\sqrt{y}) \\ &= \mathbb{P}_X(\sqrt{y}) + \mathbb{P}_X(-\sqrt{y}). \end{aligned}$$

Exemple 2.8. Soient X prenant n valeurs, notées x_i, et $Y = X^k$. Alors, la formule de transfert nous indique que l'espérance de X^k est
$$\mathbb{E}[Y] = \mathbb{E}[X^k] = \sum_{i=1}^{n} x_i^k \cdot \mathbb{P}(X = x_i),$$
que l'on appelle le moment (simple) d'ordre k de la v.a. X, noté communément m_k.

Remarque 2.2. L'espérance $\mathbb{E}[X]$ n'est qu'un indicateur moyen et n'est pas suffisant pour caractériser la loi d'une variable aléatoire. Pour décrire plus précisément le comportement de X, sans caractériser complètement la loi de X, on peut s'intéresser à la répartition attendue des écarts à cette moyenne. Cependant, si on considère simplement la différence $X - \mathbb{E}[X]$, on obtient une variable aléatoire d'espérance nulle, c'est-à-dire un écart moyen à l'espérance nul : par linéarité, $\mathbb{E}[X - \mathbb{E}[X]] = 0$ (on dit que l'on *centre* la variable). On pourrait considérer l'espérance de $|X - \mathbb{E}[X]|$ (*écart absolu moyen*), mais on considère plus souvent l'espérance de la variable aléatoire $(X - \mathbb{E}[X])^2$. Il s'agit du *moment centré d'ordre 2* de X, ou *variance* de X.

La *variance* mesure la déviation moyenne de X autour de la moyenne espérée $\mathbb{E}[X]$ et est définie par
$$\mathbb{V}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right], \tag{2.1}$$
qui pour une variable aléatoire discrète prenant un nombre fini n de valeurs $x_1, \ldots, x_n$ se calcule via la somme pondérée suivante
$$\mathbb{V}(X) = \sum_{i=1}^{n} p_i \cdot (x_i - \mathbb{E}[X])^2.$$

Remarque 2.3. $\mathbb{V}(X)$ est toujours positive puisqu'il s'agit de l'espérance du carré d'une variable aléatoire (qui prend toujours des valeurs positives, donc d'une moyenne espérée au minimum supérieure ou égale à 0).

La variance peut également se calculer à l'aide de la *formule de Koenig* :

$$\mathbb{V}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \sum_{i=1}^{n} p_i x_i^2 - \left(\sum_{i=1}^{n} p_i x_i \right)^2.$$

En effet, en développant l'équation (2.1), on a

$$\begin{aligned} \mathbb{V}(X) &= \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[X^2 - 2X\mathbb{E}[X] + \mathbb{E}[X]^2\right] \\ &= \mathbb{E}[X^2] - 2\mathbb{E}[X]\mathbb{E}[X] + \mathbb{E}[\mathbb{E}[X]^2] = E[X^2] - \mathbb{E}[X]^2 \end{aligned}$$

car $\mathbb{E}[X] \in \mathbb{R}$.

Exemple 2.9. Lorsque X correspond au nombre de *Piles* obtenu lors du lancer de deux pièces équilibrées, dont l'espérance a été déterminée égale à 1, la variance est

$$\mathbb{V}(X) = \frac{1}{4} \cdot (0-1)^2 + \frac{1}{2} \cdot (1-1)^2 + \frac{1}{4} \cdot (2-1)^2 = \frac{1}{2}.$$

Propriété 2.3.
1. **Positivité** : *Pour toute variable aléatoire X, $\mathbb{V}(X) \geq 0$, et $\mathbb{V}(X) = 0$ si et seulement si X est p.s. constante.*
2. **Non-linéarité de la variance** : *Pour toute variable aléatoire X et pour tous $a, b \in \mathbb{R}$,*
$$\mathbb{V}(aX + b) = a^2\, \mathbb{V}(X).$$
3. **Inégalité de Bienaymé-Tchebyshev** : *Soit X une v.a. de variance finie, alors pour tout $a > 0$,*
$$\mathbb{P}\left(|X - \mathbb{E}[X]| \geq a\right) \leq \frac{\mathbb{V}(X)}{a^2}.$$

Pour mesurer la dispersion d'une variable aléatoire X, on considère souvent en statistique *l'écart-type*, lié à la variance par :

$$\sigma_X = \sigma(X) = \sqrt{\mathbb{V}(X)}.$$

2.3 Couples aléatoires discrets

Définition 2.3. *Un* couple aléatoire discret *est un couple (X, Y) de variables aléatoires définies sur le même univers Ω et à valeurs dans*

$$X(\Omega) \times Y(\Omega) = \big\{(x, y) : x \in X(\Omega), y \in Y(\Omega)\big\}.$$

On notera $\{X = x, Y = y\}$ l'événement $\{X = x\} \cap \{Y = y\}$.

2.3.1 Loi du couple

Définition 2.4. *On appelle* loi de probabilité *ou* loi jointe *du couple* (X, Y) *l'application*

$$\begin{array}{rcccl} \mathbb{P}_{XY} & : & X(\Omega) \times Y(\Omega) & \longrightarrow & [0\,;1] \\ & & (x, y) & \longmapsto & \mathbb{P}_{XY}(x, y) = \mathbb{P}(X = x, Y = y). \end{array}$$

En pratique, ces probabilités jointes sont données à l'aide d'un tableau à double entrée dont les lignes correspondent aux valeurs possibles $x_i \in X(\Omega)$ prises par X, les colonnes à celles $y_j \in Y(\Omega)$ et l'élément de la ligne i et de la colonne j à la probabilité jointe $\mathbb{P}(X = x_i, Y = y_j)$. Le calcul de ces valeurs est réalisé selon les hypothèses sur les événements élémentaires de Ω.

Remarque 2.4. La somme de toutes les probabilités du tableau (donc de la loi jointe) doit être égal à 1 puisque c'est une probabilité,

$$\sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} \mathbb{P}_{XY}(x, y) = 1.$$

Exemple 2.10. Soit une urne contenant 3 boules indiscernables numérotées $\{1, 2, 3\}$. On tire successivement, sans remise, deux boules de l'urne. Soit X et Y les numéros obtenus aux 1$^{\text{er}}$ et 2$^{\text{nd}}$ tirages respectivement. Les résultats du second tirage dépendent de ceux du premier : si une boule est tirée au premier tirage, elle ne peut l'être au second ; on a donc $\mathbb{P}(X = 1, Y = 1) = \mathbb{P}(X = 2, Y = 2) = \mathbb{P}(X = 3, Y = 3) = 0$. À part ces cas, on trouve une probabilité positive et identique à toute autre combinaison possible de sorte que la loi du couple est donnée par le tableau suivant :

$x \backslash y$	1	2	3
1	0	1/6	1/6
2	1/6	0	1/6
3	1/6	1/6	0

La somme des probabilités des cas où les numéros des 2 tirages sont différents est égale à 1. Comme les probabilités sont les mêmes, chacune de ces combinaisons a une probabilité de 1/6.

2.3.2 Lois marginales

Il se peut que, connaissant la loi du couple, on ne veuille s'intéresser qu'à une seule de ses coordonnées/variables : on parlera alors de *loi marginale*.

Définition 2.5. *Soit* (X, Y) *un couple aléatoire discret. On appelle* loi marginale de X *l'application* $\mathbb{P}_X$ *de* $X(\Omega)$ *dans* $[0\,;1]$ *définie pour tout* $x \in X(\Omega)$ *par*

$$\mathbb{P}_X(x) = \mathbb{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbb{P}_{XY}(x, y).$$

On définit de manière analogue la loi marginale de Y, *soit*

$$\forall y \in Y(\Omega), \quad \mathbb{P}_Y(y) = \sum_{x \in X(\Omega)} \mathbb{P}_{XY}(x, y).$$

Les lois marginales sont des probabilités qui reviennent à ne se concentrer que sur une seule des v.a. du couple.

Exemple 2.11. Dans l'exemple précédent, la loi marginale de X est ainsi obtenue en sommant les lignes du tableau de la loi jointe, et est donnée par le tableau suivant :

x	1	2	3
$\mathbb{P}_X(x)$	$1/3$	$1/3$	$1/3$

,

tandis que l'on obtient la loi marginale de Y en sommant les colonnes :

y	1	2	3
$\mathbb{P}_Y(y)$	$1/3$	$1/3$	$1/3$

.

2.3.3 Formule de transfert

D'une manière générale, on peut calculer l'espérance d'une fonction g des deux variables X et Y, définie sur $X(\Omega) \times Y(\Omega)$, grâce à la loi du couple en écrivant

$$\mathbb{E}[g(X,Y)] = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} g(x,y) \cdot \mathbb{P}_{XY}(x,y) = \sum_{x\in X(\Omega)} \sum_{y\in Y(\Omega)} g(x,y) \cdot \mathbb{P}_{XY}(x,y).$$

Exemple 2.12. Si $g(X,Y) = aX + bY$, avec a, b deux réels, alors :

$$\mathbb{E}[aX + bY] = \sum_{x\in X(\Omega)} \sum_{y\in Y(\Omega)} (ax + by) \cdot \mathbb{P}_{XY}(x,y).$$

Remarque 2.5. La propriété de linéarité de l'espérance est bien vérifiée avec deux v.a., car on trouve à partir de l'exemple précédent :

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y] = a \sum_{x\in X(\Omega)} x \cdot \mathbb{P}_X(x) + b \sum_{y\in Y(\Omega)} y \cdot \mathbb{P}_Y(y),$$

étant donné la définition de la loi marginale.

2.3.4 Lois conditionnelles

Définition 2.6. *Soit (X,Y) un couple aléatoire discret. On appelle* loi conditionnelle de X sachant Y*, l'application $\mathbb{P}_{X|Y}$ de $X(\Omega)$ dans $[0\,;1]$ définie pour tout couple $(x,y) \in X(\Omega)\times Y(\Omega)$ par*

$$\mathbb{P}_{X|Y}(x|y) = \mathbb{P}(X = x|Y = y) = \frac{\mathbb{P}_{XY}(x,y)}{\mathbb{P}_Y(y)}.$$

On définit de manière analogue la loi conditionnelle de Y sachant X*, notée $\mathbb{P}_{Y|X}$, par*

$$\forall (x,y) \in X(\Omega) \times Y(\Omega), \quad \mathbb{P}_{Y|X}(y|x) = \mathbb{P}(Y = y|X = x) = \frac{\mathbb{P}_{XY}(x,y)}{\mathbb{P}_X(x)}.$$

Remarque 2.6. Il y a autant de lois conditionnelles de X (sachant Y) qu'il y a de valeurs de Y.

Exemple 2.13. Dans l'exemple précédent, la loi conditionnelle de Y sachant que le chiffre 1 a été tiré au premier tirage (soit sachant que $\{X = 1\}$) est donnée par le tableau :

y	1	2	3
$\mathbb{P}_{Y\|X}(y\|1)$	0	1/2	1/2

On peut également calculer la loi du couple (soit la loi jointe) à partir des lois conditionnelles et des lois marginales *en toutes circonstances* grâce au théorème suivant :

Théorème 2.1. *Soit (X, Y) un couple aléatoire discret. La* formule des probabilités composées *permet d'écrire la loi jointe de la façon suivante*

$$\mathbb{P}_{XY}(x,y) = \begin{cases} \mathbb{P}_X(x) \cdot \mathbb{P}_{Y|X}(y|x) & si \quad \mathbb{P}_X(x) \neq 0 \\ \mathbb{P}_Y(y) \cdot \mathbb{P}_{X|Y}(x|y) & si \quad \mathbb{P}_Y(y) \neq 0 \\ 0 & sinon \end{cases}.$$

2.3.5 Variables aléatoires indépendantes, covariance et corrélation

2.3.5.a Indépendance de variables aléatoires discrètes

Si X et Y sont deux variables aléatoires discrètes à valeurs respectivement dans $X(\Omega)$ et $Y(\Omega)$ (où leurs cardinaux peuvent éventuellement être infinis), on peut montrer que X et Y sont indépendantes si et seulement si pour tout $x \in X(\Omega)$ et tout $y \in Y(\Omega)$,

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y).$$

Remarque 2.7. Cette formule se généralise aux cas de n variables aléatoires discrètes $X_1, \ldots,$ X_n à valeurs dans $X_1(\Omega), \ldots, X_n(\Omega)$ respectivement, de la façon suivante

$$\forall x_1 \in X_1(\Omega), \ldots, x_n \in X_n(\Omega), \quad \mathbb{P}\left(\bigcap_{i=1}^{n} \{X_i = x_i\}\right) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i).$$

Relativement aux fonctions de répartitions de X et Y, lorsque X et Y sont indépendantes, on a :

$$\mathbb{P}(X \leq x, Y \leq y) := \mathbb{P}(\{X \leq x\} \cap \{Y \leq y\}) = \mathbb{P}(X \leq x) \cdot \mathbb{P}(Y \leq y) = F_X(x) \cdot F_Y(y).$$

Intuitivement, cela signifie que la connaissance de l'une de ces variables n'influence pas la loi de l'autre.

Remarque 2.8. Plus généralement, n v.a. $X_1, X_2, \ldots, X_n$ sont indépendantes si pour tous réels $x_1, \ldots, x_n$,

$$\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \leq x_i) = \prod_{i=1}^{n} F_{X_i}(x_i).$$

Propriété 2.4. *Si X et Y sont deux variables aléatoires* indépendantes *définies sur le même univers Ω, alors*

1. $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$.
2. $\mathbb{V}(X + Y) = \mathbb{V}(X - Y) = \mathbb{V}(X) + \mathbb{V}(Y)$.

2.3.5.b Covariance et corrélation

La variance d'une somme de v.a. n'est en général pas égale à la somme des variances. Pour calculer la variance d'une somme de deux v.a., on introduit la *covariance de deux variables aléatoires* :

$$\mathrm{Cov}(X,Y) = \mathbb{E}\big[(X-\mathbb{E}(X))(Y-\mathbb{E}[Y])\big] = \mathbb{E}[X\cdot Y] - \mathbb{E}[X]\cdot\mathbb{E}[Y].$$

Propriété 2.5. *Soient X, Y et Z trois v.a. définies sur un même univers Ω.*

1. $\mathrm{Cov}(X,X) = \mathbb{V}(X)$.
2. **Symétrie** : $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$.
3. **Bilinéarité** : *Soient $a,b,c,d \in \mathbb{R}$. Alors*

$$\mathrm{Cov}(aX+bY,Z) = a\,\mathrm{Cov}(X,Z) + b\,\mathrm{Cov}(Y,Z),$$

$$\mathrm{Cov}(X,cY+dZ) = c\,\mathrm{Cov}(X,Y) + d\,\mathrm{Cov}(X,Z).$$

4. **Indépendance** : *Si $X \perp\!\!\!\perp Y$, alors* $\mathrm{Cov}(X,Y) = 0$.
   La **réciproque est fausse** : *deux variables non-corrélées, c'est-à-dire pour lesquelles la relation précédente est vraie, peuvent ne pas être indépendantes.*

Exemple 2.14. On a deux boîtes A et B et deux billes numérotées 1 et 2. Chaque bille est placée au hasard dans l'une des deux boîtes. On a alors 4 cas possibles équiprobables :

1. les deux billes sont dans la boîte A ;
2. les deux billes sont dans la boîte B ;
3. la bille n° 1 est dans la boîte A et la bille n° 2 est dans la boîte B ;
4. la bille n° 1 est dans la boîte B et la bille n° 2 est dans la boîte A.

Soient X la v.a. représentant le nombre de billes dans la boîte A et Y la v.a. représentant le nombre de boîtes vides. Alors $X(\Omega) = \{0,1,2\}$ et $Y(\Omega) = \{0,1\}$. De plus,

$$\mathbb{E}[X] = 2\cdot\frac{1}{4} + 0\cdot\frac{1}{4} + 1\cdot\frac{1}{4} + 1\cdot\frac{1}{4} = 1,$$

$$\mathbb{E}[Y] = 1\cdot\frac{1}{4} + 1\cdot\frac{1}{4} + 0\cdot\frac{1}{4} + 0\cdot\frac{1}{4} = \frac{1}{2}.$$

Pour le produit, on a

1. $XY = 2$ si les deux billes sont dans la boîte A ;
2. $XY = 0$ si les deux billes sont dans la boîte B ;
3. $XY = 0$ si la bille n° 1 est dans la boîte A et la bille n° 2 est dans la boîte B ;
4. $XY = 0$ si la bille n° 1 est dans la boîte B et la bille n° 2 est dans la boîte A.

Donc $\mathbb{P}(XY = 2) = 1/4$ et $\mathbb{P}(XY = 0) = 3/4$, d'où

$$\mathbb{E}[XY] = 2 \cdot \frac{1}{4} + 0 \cdot \frac{3}{4} = \frac{1}{2},$$

et $\mathrm{Cov}(X, Y) = \frac{1}{2} - 1 \cdot \frac{1}{2} = 0$. Or $\mathbb{P}(X = 1, Y = 1) = 0$ car cet événement est impossible et $\mathbb{P}(X = 1)\mathbb{P}(Y = 1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$, donc les v.a. X et Y ne sont pas indépendantes mais leur covariance est nulle.

On peut calculer la variance de la somme de deux v.a. X et Y de la manière suivante :

$$\begin{aligned} \mathbb{V}(X+Y) &= \mathbb{E}\left[(X+Y-\mathbb{E}[X+Y])^2\right] = \mathbb{E}\left[(X-\mathbb{E}[X]+Y-\mathbb{E}[Y])^2\right] \\ &= \mathbb{E}\left[(X-\mathbb{E}[X])^2\right] + \mathbb{E}\left[(Y-\mathbb{E}[Y])^2\right] + 2\mathbb{E}\left[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])\right] \\ &= \mathbb{V}(X) + \mathbb{V}(Y) + 2\,\mathrm{Cov}(X,Y). \end{aligned}$$

En particulier,

$$\mathbb{V}(X-Y) = \mathbb{V}(X) + \mathbb{V}(Y) - 2\,\mathrm{Cov}(X,Y).$$

Cette formule se généralise à la somme de n v.a. $X_1, \ldots, X_n$ de la façon suivante :

$$\mathbb{V}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{V}(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j).$$

La covariance est à valeurs réelles et en la normalisant par le produit des écart-types des v.a., on obtient la *corrélation* qui est à valeurs dans $[-1\,;1]$ et est définie par

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}.$$

Elle mesure la dépendance linéaire entre les v.a. X et Y. Si $X \perp\!\!\!\perp Y$, alors $\rho(X,Y) = 0$ (mais la **réciproque est fausse** en général, comme pour la covariance).

2.4 Lois usuelles discrètes

On considère dans toute cette section une variable aléatoire discrète sur un univers quelconque Ω. Lorsque X prend n valeurs, l'ensemble $X(\Omega)$ des valeurs prises par X est désigné par $\{x_1, \ldots, x_n\} = \{x_i\}_{i=1,\ldots,n}$, et $\{x_i\}_{i\in\mathbb{N}}$ lorsque X prend une infinité dénombrable de valeurs. Le comportement aléatoire d'une telle variable X peut prendre de multiples formes selon les divers phénomènes étudiés et la loi d'une variable aléatoire peut prendre toute forme envisageable. Cependant, certains paramètres objectifs de caractérisation (espérance, variance, etc.) permettent de dégager des comportements récurrents et des familles de lois typiques permettent une modélisation approchée raisonnable de la plupart des phénomènes aléatoires courants. Nous décrivons ici les lois discrètes les plus importantes – **la loi uniforme, la loi de Bernoulli, la loi binomiale, le loi géométrique et la loi de Poisson** – à travers certains exemples de modélisations que l'on traitera également en TD.

2.4.1 Loi uniforme sur $\{1, \ldots, n\}$

Elle modélise des situations d'*équiprobabilité*. On dit qu'une variable aléatoire X suit une loi uniforme sur $\{1, \ldots, n\}$ (noté $X \sim \mathcal{U}_{\{1,\ldots,n\}}$) lorsqu'elle prend ses valeurs dans $\{1, \ldots, n\}$ avec d'identiques probabilités élémentaires. Puisque leur somme doit valoir 1, on en déduit qu'elles doivent toutes être égales à $1/n$:

$$\forall\, k = 1, \ldots, n, \quad \mathbb{P}(X = k) = \frac{1}{n}.$$

On note également ces probabilités p_k, $p(k)$ ou $\mathbb{P}_X(k)$. Elles ont la particularité d'être indépendantes de k.

Espérance et variance. En utilisant le fait que la somme des entiers de 1 à n est égale à $\dfrac{n(n+1)}{2}$, on obtient

$$\mathbb{E}[X] = \sum_{k=1}^{n} k \cdot \frac{1}{n} = \frac{1}{n} \sum_{k=1}^{n} k = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}.$$

En utilisant le fait que la somme des carrés des entiers de 1 à n est égale à $\dfrac{n(n+1)(2n+1)}{6}$, on obtient

$$\mathbb{E}[X^2] = \sum_{k=1}^{n} k^2 \cdot \frac{1}{n} = \frac{1}{n} \sum_{k=1}^{n} k^2 = \frac{1}{n} \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6},$$

d'où

$$\begin{aligned}
\mathbb{V}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} \\
&= \frac{2(2n^2+3n+1)}{12} - \frac{3(n^2+2n+1)}{12} = \frac{n^2-1}{12}.
\end{aligned}$$

Exemple 2.15. Soit X la v.a. représentant le résultat d'un jet de dé équilibré à six faces. Les $n = 6$ modalités possibles, $x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4, x_5 = 5, x_6 = 6$, ont toutes pour probabilité élémentaire $1/6$:

$$\forall\, k = 1, \ldots, 6, \quad \mathbb{P}_X(k) = \mathbb{P}(X = k) = \frac{1}{6},$$

et on peut calculer $\mathbb{E}[X] = \dfrac{7}{2}$ et $\mathbb{V}(X) = \dfrac{35}{12}$.

2.4.2 Loi de Bernoulli

Cette loi est celle de toute variable aléatoire X modélisant une expérience dont l'issue ne possède qu'une alternative de type « succès ou échec », « vrai ou faux », « marche ou arrêt », « pile ou face », etc. Un *succès* est représenté par l'événement $\{X = 1\}$ tandis que $\{X = 0\}$ correspond à un *échec*. Puisque l'on a $\mathbb{P}(X = 0) = 1 - \mathbb{P}(X = 1)$, la loi de X ne dépend que d'un paramètre (la probabilité de succès notée p) ; on parle alors de *la loi de Bernoulli de paramètre* p (on le note $X \sim \mathcal{B}(p)$) caractérisée par

$$\mathbb{P}(X = 1) = p, \quad \mathbb{P}(X = 0) = 1 - p.$$

Espérance et variance. On a facilement

$$\mathbb{E}[X] = 0 \cdot \mathbb{P}(X = 0) + 1 \cdot \mathbb{P}(X = 1) = p, \quad \mathbb{E}[X^2] = 0^2 \cdot \mathbb{P}(X = 0) + 1^2 \cdot \mathbb{P}(X = 1) = p,$$

$$\text{et} \quad \mathbb{V}(X) = p(1 - p).$$

2.4.3 Loi binomiale $\mathcal{B}(n, p)$

La loi binomiale est la loi de probabilité de toute variable aléatoire représentant une suite d'épreuves de Bernoulli possédant les propriétés suivantes :

- chaque épreuve donne lieu à deux éventualités exclusives de probabilités respectives p et $q = 1 - p$ (identiques pour toutes les épreuves) ;
- les épreuves répétées sont *indépendantes* les unes des autres ;
- la variable aléatoire correspondante prend pour valeur le nombre de succès dans une suite de n épreuves.

Deux paramètres (le nombre n d'épreuves (identiques et indépendantes) répétées et la probabilité p de succès dans l'épreuve de Bernoulli) caractérisent la loi. On dit alors que X suit une loi binomiale de paramètres n et p (noté $X \sim \mathcal{B}(n, p)$), à valeurs dans $X(\Omega) = \{0, 1, 2, \ldots, n\}$.

Exemple 2.16. Le nombre de *Piles* obtenus au cours de n lancers indépendants d'une pièce équilibrée est une variable aléatoire discrète, suivant une loi binomiale $\mathcal{B}(n, p)$ avec $p = \frac{1}{2}$, puisque la probabilité de succès est celle d'obtenir un pile pour une pièce équilibrée. On a par ailleurs

$$X = X_1 + \cdots + X_k + \cdots + X_n,$$

où les X_k sont des variables aléatoires de Bernoulli indépendantes de paramètre $p = \frac{1}{2}$, correspondant à une seule épreuve de pile ou face.

Exemple 2.17. Le nombre de boules rouges extraites au cours de n tirages successifs *avec remise* (pour assurer l'indépendance des tirages) d'une boule dans une urne contenant des boules rouges et blanches dans des proportions p et $q = 1 - p$ est une variable aléatoire de loi $\mathcal{B}(n, p)$.

Le nombre de possibilités de k succès en n essais est le nombre de combinaisons C_n^k. Il suffit ensuite de le multiplier par les probabilités de succès et d'échecs pour obtenir les **probabilités élémentaires de la loi binomiale :**

$$\forall k \in \{0, \ldots, n\}, \quad \mathbb{P}(X = k) = C_n^k \cdot p^k \cdot (1 - p)^{n-k}. \tag{2.2}$$

Espérance et variance. La loi binomiale $\mathcal{B}(n, p)$ étant la somme de n lois de Bernoulli $\mathcal{B}(p)$ indépendantes, on a, en posant $X = \sum_{k=1}^{n} B_k$, $B_k \sim \mathcal{B}(p)$ indépendantes,

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{k=1}^{n} B_k\right] = \sum_{k=1}^{n} \mathbb{E}[B_k] = n \cdot p \quad \text{et} \quad \mathbb{V}(X) = \mathbb{V}\left(\sum_{k=1}^{n} B_k\right) = \sum_{k=1}^{n} \mathbb{V}(B_k) = n \cdot p \cdot (1 - p).$$

Exemple 2.18. 1. Un atelier comporte 100 machines identiques. Chaque machine a une probabilité $p = 0,01$ de tomber en panne dans une journée. Lorsque l'on suppose que les machines tombent en panne de manière indépendante, la variable aléatoire X désignant le nombre de machines tombant en panne dans une journée suit une loi $\mathcal{B}(100; 0,01)$ et les probabilités que ce nombre soit égal à k est donc évalué grâce à (2.2). En particulier, on a directement un nombre moyen de pannes égal à $\mathbb{E}[X] = 100 \times 0,01 = 1$, la variance étant $\mathbb{V}(X) = 100 \times 0,01 \times 0,99 = 0,99$.

2. Une machine qui a une probabilité $p = 0,01$ de tomber en panne dans une journée est amenée à fonctionner pendant 20 jours consécutifs. Sous quelles conditions la variable aléatoire X représentant le nombre de jours avec panne pendant ces 20 journées suit-elle une loi binomiale ? Dans ce cas, laquelle ?

   Réponse : En supposant l'indépendance des pannes, i.e. en considérant une restauration à l'identique de chaque machine après une panne, X suit une loi $\mathcal{B}(20; 0,01)$.

3. Au cours de 1 000 lancers d'une pièce équilibrée, le nombre de *Piles* obtenu suit une loi binomiale $\mathcal{B}(1\,000; 0,5)$ et le nombre moyen de *Piles* obtenu sera donc égal à $1\,000 \times 0,5 = 500$.

2.4.4 Loi géométrique

Si, au lieu de s'intéresser au nombre de succès rencontrés au cours d'un certain nombre de tirages, on se focalise sur le nombre d'essais nécessaires pour voir apparaître un succès, on obtient une loi géométrique au lieu d'une loi binomiale.

La *loi géométrique* modélise la loi d'une variable aléatoire X pouvant prendre toutes les valeurs entières non nulles ($X(\Omega) = \mathbb{N}^*$) et qui comptabilise le nombre d'essais nécessaires pour l'obtention d'un succès au cours d'une suite d'épreuves de Bernoulli indépendantes et de même paramètre p (on note alors $X \sim \mathcal{G}(p)$). Grâce à l'indépendance, les probabilités sont alors simplement données par l'expression

$$\forall k \in \mathbb{N}^*, \quad \mathbb{P}(X = k) = p \cdot (1-p)^{k-1}.$$

Le facteur $(1-p)^{k-1}$ correspond à la probabilité d'avoir $k-1$ échecs au $k-1$ premiers lancers et p correspond à la probabilité du succès au $k^{\text{ème}}$ lancer.

Exemple 2.19.
1. Le nombre aléatoire X de *Piles* obtenus avant l'apparition du premier *Face* au cours de lancers indépendants d'une pièce équilibrée suit une loi géométrique de paramètre $\frac{1}{2}$.
2. Le nombre de boules blanches extraites avant l'apparition de la $1^{\text{ère}}$ boule rouge au cours d'une suite de tirages *avec remise* dans une urne contenant des boules rouges et blanches dans des proportions p et $1-p$ suit une loi géométrique de paramètre p.

Espérance et variance. Rappelons que, pour tout $x \in [0,1[$,

$$\begin{aligned}
S(x) &= \sum_{k\geq 0} x^k = \lim_{n\to+\infty} \sum_{k=0}^{n} x^k = \lim_{n\to+\infty} \frac{1-x^n}{1-x} = \frac{1}{1-x}, \\
S'(x) &= \sum_{k\geq 1} k x^{k-1} = \left(\frac{1}{1-x}\right)' = \frac{1}{(1-x)^2}, \\
S''(x) &= \sum_{k\geq 2} k(k-1) x^{k-2} = \left(\frac{1}{(1-x)^2}\right)' = \frac{2}{(1-x)^3}.
\end{aligned}$$

Ainsi, pour $p > 0$,

$$\mathbb{E}[X] = \sum_{k \geq 1} k \cdot p \cdot (1-p)^{k-1} = p \cdot S'(1-p) = \frac{p}{(1-(1-p))^2} = \frac{1}{p}.$$

De même,

$$\begin{aligned} \mathbb{E}[X^2] &= \sum_{k \geq 1} k^2 \cdot p \cdot (1-p)^{k-1} = p(1-p) \sum_{k \geq 1} k(k-1)(1-p)^{k-2} + p \sum_{k \geq 1} k(1-p)^{k-1} \\ &= p(1-p) \cdot S''(1-p) + p \cdot S'(1-p) = \frac{2p(1-p)}{(1-(1-p))^3} + \frac{1}{p} = \frac{2(1-p)}{p^2} + \frac{1}{p} = \frac{2-p}{p^2}. \end{aligned}$$

D'où

$$\mathbb{V}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

2.4.5 Loi de Poisson et approximation

Lorsque le nombre d'épreuves n devient très important, la manipulation de la loi binomiale devient très fastidieuse et une grande part du travail statistique consiste à l'approcher par des lois plus maniables. Lorsque la probabilité du succès est petite, c'est-à-dire lorsque l'on s'intéresse à des événements rares, on peut la remplacer en première approximation par une **loi de Poisson** $\mathcal{P}(\lambda)$. Cette loi a initialement été introduite pour évaluer le nombre aléatoire d'événements de même probabilité réalisés pendant une durée donnée. Elle peut modéliser par exemple le nombre d'appels reçus par un standard téléphonique dans une journée, le nombre de clients se présentant à un guichet en une heure, etc.

La loi s'exprime à l'aide de la fonction exponentielle et dépend d'un paramètre $\lambda > 0$ qui correspond au nombre moyen d'occurrence du phénomène observé pendant la durée donnée. Plus formellement :

Définition 2.7. *Une variable aléatoire X suit une loi de Poisson de paramètre $\lambda > 0$, notée $\mathcal{P}(\lambda)$, lorsque $X(\Omega) = \mathbb{N}$ et*

$$\forall k \in \mathbb{N}, \quad \mathbb{P}(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}. \tag{2.3}$$

Espérance et variance. En utilisant le développement en série de l'exponentielle, c'est-à-dire

$$\forall x \in \mathbb{R}, \quad e^x = \sum_{k \geq 0} \frac{x^k}{k!},$$

on a

$$\mathbb{E}[X] = \sum_{k \geq 0} k \cdot \mathbb{P}(X = k) = \sum_{k \geq 1} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k \geq 1} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{k \geq 0} \frac{\lambda^k}{k!} = \lambda.$$

De même,

$$\begin{aligned}
\mathbb{E}[X^2] &= \sum_{k\geq 0} k^2 \cdot \mathbb{P}(X=k) = \sum_{k\geq 1} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k\geq 1} k \frac{\lambda^{k-1}}{(k-1)!} \\
&= \lambda e^{-\lambda} \sum_{k\geq 0} (k+1) \frac{\lambda^k}{k!} = \lambda e^{-\lambda} \left(\lambda \sum_{k\geq 1} \frac{\lambda^{k-1}}{(k-1)!} + \sum_{k\geq 0} \frac{\lambda^k}{k!} \right) \\
&= \lambda e^{-\lambda} \left(\lambda \sum_{k\geq 0} \frac{\lambda^k}{k!} + e^{\lambda} \right) = \lambda^2 + \lambda.
\end{aligned}$$

D'où $\mathbb{V}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \lambda$.

Exemple 2.20. Si on sait qu'en général un standard téléphonique reçoit 20 appels dans la journée et que l'on peut modéliser le nombre aléatoire d'appels par une loi de Poisson, on pourra calculer la probabilité d'avoir 5 appels en remplaçant k par 5 et λ par 20 dans la formule (2.3) (soit une loi de Poisson $\mathcal{P}(20)$).

Dans la pratique, des tables donnant les probabilités élémentaires pour différentes valeurs du paramètre sont disponibles et utilisées.

Exemple 2.21 (Propriétés de la loi de Poisson)**.** Un bureau de poste dispose de deux guichets : le 1er est réservé au courrier, le 2nd aux services bancaires. Pour un intervalle de temps $[t_1, t_2]$ donné, on note X_1 et X_2 les v.a. représentant le nombre de clients se présentant indépendamment à l'un et à l'autre des deux guichets. On suppose que X_1 et X_2 suivent respectivement les lois de Poisson $\mathcal{P}(\lambda_1)$ et $\mathcal{P}(\lambda_2)$.

1. Soit Z la v.a. représentant le nombre total de clients durant l'intervalle $[t_1, t_2]$. Alors la loi de Z est une loi de Poisson de paramètre $\lambda_1 + \lambda_2$. En effet, Z est à valeurs dans $\mathbb{N}$ car somme de deux nombres entiers. Ensuite, pour tout $n \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{P}(Z=n) &= \mathbb{P}(X_1 + X_2 = n) = \sum_{k=0}^{n} \mathbb{P}(X_1 + X_2 = n, X_1 = k) \quad \text{probabilités totales} \\
&= \sum_{k=0}^{n} \mathbb{P}(X_2 = n-k, X_1 = k) = \sum_{k=0}^{n} \mathbb{P}(X_2 = n-k)\mathbb{P}(X_1 = k) \text{ car } X_1 \perp\!\!\!\perp X_2 \\
&= \sum_{k=0}^{n} \frac{(\lambda_2)^{n-k}}{(n-k)!} e^{-\lambda_2} \frac{(\lambda_1)^k}{k!} e^{-\lambda_1} \quad \text{car } X_1 \overset{\mathcal{L}}{\sim} \mathcal{P}(\lambda_1) \text{ et } X_2 \overset{\mathcal{L}}{\sim} \mathcal{P}(\lambda_2) \\
&= \frac{e^{\lambda_1+\lambda_2}}{n!} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} (\lambda_1)^k \cdot (\lambda_2)^{n-k} = \frac{e^{\lambda_1+\lambda_2}}{n!} (\lambda_1 + \lambda_2)^n.
\end{aligned}$$

   La dernière égalité est obtenue par la formule du binôme de Newton.

2. La loi de probabilité du nombre de clients se présentant au 1er guichet sachant qu'il y a n clients dans le bureau de poste est la loi conditionnelle de X_1 par rapport à l'événement $\{Z = n\}$, $n \in \mathbb{N}$. On sait qu'alors, conditionnellement à $\{Z = n\}$, X_1 est à valeurs dans

$\{0, 1, \ldots, n\}$ et pour tout $k \in \{0, 1, \ldots, n\}$, on a

$$\begin{aligned}
\mathbb{P}(X_1 = k | Z = n) &= \frac{\mathbb{P}(X_1 = k, Z = n)}{\mathbb{P}(Z = n)} = \frac{\mathbb{P}(X_1 = k, X_2 = n - k)}{\mathbb{P}(Z = n)} \\
&= \frac{\mathbb{P}(X_1 = k)\mathbb{P}(X_2 = n - k)}{\mathbb{P}(Z = n)} \quad \text{car } X_1 \perp\!\!\!\perp X_2 \\
&= \frac{\frac{(\lambda_1)^k}{k!} e^{-\lambda_1} \frac{(\lambda_2)^{n-k}}{(n-k)!} e^{-\lambda_2}}{\frac{(\lambda_1+\lambda_2)^n}{n!} e^{\lambda_1+\lambda_2}} \\
&\qquad \text{car } X_1 \overset{\mathcal{L}}{\sim} \mathcal{P}(\lambda_1),\ X_2 \overset{\mathcal{L}}{\sim} \mathcal{P}(\lambda_2)\ Z \overset{\mathcal{L}}{\sim} \mathcal{P}(\lambda_1 + \lambda_2) \\
&= \frac{n!}{k!(n-k)!} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^k \left(\frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{n-k}.
\end{aligned}$$

Ainsi, la loi conditionnelle de X_1 sachant $\{Z = n\}$ est une loi binomiale de paramètre n et $\frac{\lambda_1}{\lambda_1+\lambda_2}$.

2.4.5.a Approximation de lois binomiales

La loi de Poisson est souvent utilisée comme approximation de certaines lois binomiales pour de grands échantillons, c'est-à-dire des lois binomiales correspondant à des grands nombres n d'épreuves de Bernoulli, lorsque les événements sont rares (p petit) car $\mathbb{V}(X) = np(1 - p)$, où $1 - p$ est proche de 1 donc on la remplace par l'espérance np.

Proposition 2.1. *Soit X une variable aléatoire de loi binomiale $\mathcal{B}(n, p)$ avec n grand et p petit proche de 0. On peut alors approcher la loi de X par une loi de Poisson $\mathcal{P}(\lambda)$ de même espérance, et donc de paramètre*

$$\lambda = n.p$$

Dans la pratique, on considère que l'approximation est bonne lorsque

$$n \geq 30, \quad p \leq 0,1 \quad \text{et} \quad n \cdot p < 15.$$