Chapitre 19

# Probabilités

## Table des matières


**1 Espaces probabilisés** **2**
- 1.1 Univers, évènements . . . 2
- 1.2 Probabilité . . . 2
- 1.3 Probabilité conditionnelle . . . 3
- 1.4 Indépendance de deux évènements . . . 4
- 1.5 Indépendance de $n$ évènements . . . 4

**2 Loi d'une variable aléatoire** **4**
- 2.1 Définition et propriétés . . . 4
- 2.2 Lois usuelles . . . 6
- 2.3 Couple de variables aléatoires . . . 7
- 2.4 Généralisation à un $n$-uplet de variables aléatoires . . . 8
- 2.5 Indépendance de deux variables aléatoires . . . 8
- 2.6 Indépendance de $n$ variables aléatoires . . . 9

**3 Espérance et Variance** **10**
- 3.1 Espérance . . . 10
- 3.2 Variance . . . 13

**4 Tableau récapitulatif des lois usuelles** **16**

# 1 Espaces probabilisés

## 1.1 Univers, évènements

**Définition : vocabulaire probabiliste**

Soit $\Omega$ un ensemble fini que l'on appelle **univers**.

1. Un sous-ensemble de $\Omega$ est appelé **évènement**.
2. L'ensemble de tous les évènements est donc $\mathscr{P}(\Omega)$.
3. Si $\omega \in \Omega$, $\{\omega\}$ est appelé **évènement élémentaire**.
4. L'évènement $\Omega$ est appelé **évènement certain**.
5. L'évènement $\varnothing$ est appelé **évènement impossible**.
6. Si $A$ et $B$ sont deux évènements, $A \cup B$ (resp. $A \cap B$) est appelé évènement «$A$ ou $B$», (resp. «$A$ et $B$»).
7. Si $A \cap B = \varnothing$, les évènements $A$ et $B$ sont dits **incompatibles**.
8. L'évènement $\overline{A}$ est appelé **évènement contraire** de $A$».

**Exemples 1.**
- On lance un dé à six faces, qui est $\Omega$ ? Quelles sont les évènements élémentaires ? Quel est l'évènement «le résultat est pair» ?
- On lance $n$ fois un dé et on note $S_k$ l'évènement «le dé a fait un six lors du $k$-ième lancer». Donner $\Omega$, écrire explicitement $S_k$. Écrire, en fonction de $S_k$ les évènements «on a obtenu un six lors des deux premiers lancers», «on a obtenu un six à chaque lancer», «on n'a jamais obtenu un six», «on a obtenu au moins une fois un six», «on a obtenu au moins deux six d'affilée».

**Définition d'un système complet d'événements**

Un **système complet d'évènements** est formé d'évènements $(A_i)_{1 \leqslant i \leqslant p}$ deux à deux incompatibles tels que $\Omega = \bigcup_{i=1}^{p} A_i$.

**Exemples 2.**
- Si $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, alors $\{\omega_1\}$, $\{\omega_2\}$, ..., $\{\omega_n\}$ forment un système complet d'évènements.
- Les évènements «la face du dé est pair» et «la face du dé est impair» forment un système complet d'évènements.

## 1.2 Probabilité

**Définition d'une probabilité sur un univers fini**

Une application $\mathbb{P}\colon \mathscr{P}(\Omega) \to [\,0\,;1\,]$, telle que $\mathbb{P}(\Omega) = 1$ et pour tout $(A, B) \in \mathscr{P}(\Omega)^2$, si $A \cap B = \varnothing$, alors $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ est appelée **probabilité** sur $\Omega$. Le couple $(\Omega, \mathbb{P})$ est appelé **espace probabilisé**.

**Exemple 3.** L'application $\mathbb{P}\colon A \mapsto |A|/|\Omega|$ est une probabilité sur $\Omega$, appelée **probabilité uniforme sur** $\Omega$.

**Remarque 1.** Il faut bien comprendre que ce qui compte, ce n'est pas tant l'ensemble $\Omega$, mais bien la probabilité qu'on attribue aux évènements. Par exemple, si on lance un dé, $\Omega = [\![\,1\,;6\,]\!]$, si le dé est équilibré alors $\mathbb{P}$ sera la probabilité uniforme, mais sinon ce sera une autre probabilité.

À partir de maintenant, $(\Omega, \mathbb{P})$ désigne un espace probabilisé.

**Proposition n° 1 : propriétés des probabilités**

Soit $A$, $B$, $A_1$, …, $A_n$ des évènements.

1. $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$
2. $\mathbb{P}(\varnothing) = 0$
3. $\mathbb{P}(A \backslash B) = \mathbb{P}(A) - P(A \cap B)$
4. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
5. Si $B \subset A$ alors $\mathbb{P}(A \backslash B) = \mathbb{P}(A) - \mathbb{P}(B)$ et $\mathbb{P}(B) \leqslant \mathbb{P}(A)$
6. $\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \leqslant \sum_{i=1}^{n} \mathbb{P}(A_i)$
7. Si $A_1, \ldots, A_n$ sont 2 à 2 incompatibles, alors $\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i)$

**Définition d'une distribution de probabilités**

Une **distribution de probabilité** sur $\Omega = \{\omega_1, \dots, \omega_n\}$ est une famille de réels positifs $(p_1, \dots, p_n)$ de somme 1.

**Proposition n° 2 : une probabilité est entièrement caractérisée par sa distribution de probabilité**

Soit une distribution de probabilité $(p_1, \dots, p_n)$ sur $\Omega$. Il existe une unique probabilité $\mathbb{P}$ telle que $\mathbb{P}(\{\omega_i\}) = p_i$ pour tout $i \in [\![ 1\,;n ]\!]$.

**Exemple 4.** Considérons un dé et donc $\Omega = [\![ 1\,;6 ]\!]$, $p_1 = p_2 = p_3 = p_4 = p_5 = 1/12$ et $p_6 = 7/12$, alors il existe une probabilité tel que la probabilité de tirer six soit $7/12$ contre $1/12$ pour les autres faces.

## 1.3 Probabilité conditionnelle

**Définition d'une probabilité conditionnelle**

Soit $B$ un évènement tel que $\mathbb{P}(B) > 0$, on définit la **probabilité conditionnelle** de $A$ sachant $B$ par $\mathbb{P}(A|B) = \mathbb{P}_B(A) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$.

**Exemple 5.** On lance un dé à six faces équilibré, quelle est la probabilité de tirer un nombre pair sachant que l'on a tiré un nombre premier ?

**Proposition n° 3 : la probabilité conditionnelle est une probabilité**

Si $\mathbb{P}(B) > 0$, alors $\mathbb{P}_B$ est une probabilité.

**Remarque 2.** Si $\mathbb{P}(B) > 0$, alors on a toujours $\mathbb{P}(B)\mathbb{P}(A|B) = \mathbb{P}(A \cap B)$.
Si $\mathbb{P}(B) = 0$, on pose, par convention, $\mathbb{P}(A|B) = 0$, ainsi la relation $\mathbb{P}(B)\mathbb{P}(A|B) = \mathbb{P}(A \cap B)$ reste vraie.

**Théorème n° 1 : formule des probabilités totales**

Soit $(\Omega, \mathbb{P})$ un espace probabilisé, $(A_1, A_2, \dots, A_n)$ un système complet d'évènements. Pour tout $B \in \mathscr{P}(\Omega)$ :

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B \cap A_i) = \sum_{i=1}^{n} \mathbb{P}(B|A_i)\mathbb{P}(A_i)$$

**Exemple 6.** On tire deux boules successivement et sans remise d'une urne contenant $x$ boules rouges et $y$ boules bleues avec $x \geqslant 1$ et $y \geqslant 1$. Quelle est la probabilité que la deuxième boule soit bleue ?

**Théorème n° 2 : formules des probabilités composées**

Soient $(\Omega, \mathbb{P})$ un espace probabilisé et une famille d'évènements $(A_1, A_2, \dots, A_n)$ alors :

$$\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}_{A_1}(A_2)\mathbb{P}_{A_1 \cap A_2}(A_3) \dots \mathbb{P}_{A_1 \cap A_2 \cap \dots A_{n-1}}(A_n) = \mathbb{P}(A_1) \prod_{k=1}^{n-1} \mathbb{P}_{\bigcap_{i=1}^{k} A_i}(A_{k+1})$$

**Exemple 7.** On tire successivement et sans remise trois billes d'une urne contenant $x \geqslant 2$ billes rouges et $y \geqslant 1$ bleues. Quelle est la probabilité que les deux premières soient rouges mais pas la dernière ?

**Théorème n° 3 : formule de Bayes**

Soit $B$ un évènement tel que $\mathbb{P}(B) > 0$, alors $\mathbb{P}_B(A) = \dfrac{\mathbb{P}(A)\mathbb{P}_A(B)}{\mathbb{P}(B)}$

**Exemple 8.** Considérons une maladie telle que la probabilité qu'une personne soit atteinte soit de 1/1000 et que l'on ait un test pour savoir si un patient est infecté :

- Si le patient est malade, le résultat sera positif avec une probabilité de 99/100.
- Si un patient est sain, le résultat sera négatif avec une probabilité de 95/100 .

Supposons que le test d'un patient soit positif, quelle est la probabilité qu'il soit vraiment malade ?

### 1.4 Indépendance de deux évènements

**Définition de l'indépendance de deux évènements**

On dit que deux évènements $A$ et $B$ sont **indépendants** si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

**Remarque 3.** Lorsque $\mathbb{P}(B) > 0$, alors $A$ et $B$ sont indépendants ssi $\mathbb{P}(A|B) = \mathbb{P}(A)$. Ainsi, la probabilité d'obtenir $A$ est la même si on sait que $B$ est réalisé.

**Proposition n° 4 : indépendance des évènements complémentaires**

Si $A$ et $B$ sont indépendants, alors $A$ et $\overline{B}$ le sont aussi, de même pour $\overline{A}$ et $B$, de même $\overline{A}$ et $\overline{B}$.

### 1.5 Indépendance de $n$ évènements

**Définition de l'indépendance de $n$ évènements**

On dit que les évènements $A_1, A_2, \dots, A_n$ sont **indépendants** si $\forall J \subset [\![ 1\,;n ]\!] \quad \mathbb{P}\left(\bigcap_{j\in J} A_j\right) = \prod_{j\in J} \mathbb{P}(A_j)$

**Exemple 9.** Si $n = 2$ ou $n = 3$ quelle(s) vérification(s) faut-il faire pour démontrer que les évènements sont indépendants ?

**Remarque 4.** Si les événements $A_1, A_2, \dots, A_n$ sont indépendants, alors toute sous-famille l'est aussi.

**Péril imminent l'indépendance de $n$ évènements n'est pas l'indépendance deux à deux**

Il est possible que $A_1$ et $A_2$ soient indépendants, de même entre $A_2$ et $A_3$ de même qu'entre $A_1$ et $A_3$ sans que $A_1$, $A_2$ et $A_3$ soient indépendants.

**Exemple 10.** Soit $\Omega = [\![ 0\,;3 ]\!]$ muni de sa probabilité uniforme, on note $A_1 = \{0, 1\}$, $A_2 = \{1, 2\}$ et $A_3 = \{0, 2\}$, montrer que $A_1$ et $A_2$ sont indépendants, de même $A_1$ et $A_3$ puis $A_2$ et $A_3$ mais que $A_1$, $A_2$ et $A_3$ ne sont pas indépendants.

**Proposition n° 5 : indépendance des évènements complémentaires**

Si $A_1, A_2, \dots, A_n$ sont des évènements indépendants et $B_i \in \{A_i, \overline{A_i}\}$, alors $B_1, B_2, \dots, B_n$ sont indépendants.

**Exemple 11.** Soit un entier $n \geqslant 7$. On lance $n$ fois une pièce dont la probabilité de faire pile est $p \in ]\,0\,;1\,[$. On note $A_k$ l'évènement «on a obtenu un pile lors du $k$-ième lancer» et on suppose que les $A_1, \dots, A_n$ sont indépendants. Quelle est la probabilité d'obtenir un pile lors des trois premiers lancers puis quatre faces lors des quatre lancers suivants ?

## 2 Loi d'une variable aléatoire

### 2.1 Définition et propriétés

**Définition d'une variable aléatoire**

Une application $X : \begin{cases} \Omega \longrightarrow E \\ \omega \longmapsto X(\omega) \end{cases}$ où $E$ est un ensemble est appelée **variable aléatoire** sur $\Omega$ à valeurs dans $E$

.

**Remarque 5.** On remarque *malicieusement* qu'une variable aléatoire n'a rien d'aléatoire et n'a rien d'une variable...

**Exemples 12.**
1. On lance deux dés et on note $X_1$ : somme des résultats obtenus.
2. On lance $n$ dés et on note $X_2$ : nombre de 6 obtenus.
3. On tire au hasard simultanément 3 boules d'une urne qui contient 3 boules blanches et 3 rouges et on note $X_3$ : nombre de boules rouges.

À partir de maintenant $X : \Omega \to E$ est une variable aléatoire.

**Définition vocabulaire**

1. $X(\Omega) = \{X(\omega) \mid \omega \in \Omega\}$ est appelé **univers image**.
2. Soit $B \subset E$, on note l'évènement $(X \in B) = \{X \in B\} = X^{-1}(B) = \{\omega \in \Omega,\ X(\omega) \in B\}$.
3. Pour $x \in E$, on note l'évènement $(X = x) = X^{-1}(\{x\}) = \{\omega \in \Omega,\ X(\omega) = x\}$.
4. Si $E = \mathbb{R}$, on note l'évènement $(X < x) = X^{-1}(]-\infty\,;x[) = \{\omega \in \Omega,\ X(\omega) < x\}$, idem pour $(X \geqslant x)$ etc.

**Exemple 13.** Déterminer l'univers image des VA de l'exemple 12 puis déterminer les évènements $(X_1 = 7)$ et $(X_2 = 0)$.

**Proposition n° 6 : système complet d'événements associé à une variable aléatoire**

Soit une VA $X : \Omega \to E = \{e_1, \dots, e_n\}$, alors $(X = e_1), \dots, (X = e_n)$ forment un SCE appelé **SCE des valeurs possibles** de $X$. En particulier, $\sum_{i=1}^{n} \mathbb{P}(X = e_i) = 1$ pour $A \subset E$, $\mathbb{P}(X \in A) = \sum_{a \in A} \mathbb{P}(X = a)$

$\mathbb{P}_X : A \mapsto \mathbb{P}(X \in A)$ est une probabilité sur $E$ et est déterminée par $(\mathbb{P}(X = e))_{e \in E}$.

**Démonstration de la proposition n° 6 :** Notons $E = \{x_1, \dots, x_p\}$ et posons $A_i = (X = x_i)$, pour $i \in [\![1\,;p]\!]$. Fixons $(i,j) \in [\![1\,;n]\!]^2$ tel que $i \neq j$. Soit $\omega \in A_i \cap A_j$, alors $\omega \in A_i$ et $\omega \in A_j$. Donc, $X(\omega) = x_i$ et $X(\omega) = x_j$, par conséquent $x_i = x_j$. Or, comme $i \neq j$, $x_i \neq x_j$. Ainsi, $A_i \cap A_j = \varnothing$. De plus, soit $\omega \in \Omega$. Il existe alors existe $j \in [\![1\,;n]\!]$, tel que $x_j = X(\omega)$, donc $\omega \in (X = x_j)$, ainsi, $\omega \in A_j \subset \bigcup_{i=1}^{p} A_i$, d'où $\omega \in \bigcup_{i=1}^{p} A_i$. On a donc démontré que $\Omega \subset \bigcup_{i=1}^{p} A_i$. Comme l'inclusion réciproque est vraie, on a bien $\Omega = \bigcup_{i=1}^{p} A_i$. Comme les intersections deux à deux sont disjointes, on a bien un SCE.

Ainsi, $\sum_{i=1}^{p} \mathbb{P}((X = x_i)) = \sum_{i=1}^{p} \mathbb{P}(A_i) = 1$. Et d'après la formule des probabilités totales,

$$\mathbb{P}(X \in B) = \sum_{i=1}^{p} \mathbb{P}((X \in B) \cap (X = x_i))$$

Calculons $(X \in B) \cap (X = x_i)$.
- Si $x_i \notin B$, alors prenons $\omega \in (X \in B) \cap (X = x_i)$, ainsi $X(\omega) \in B$ et $X(\omega) = x_i \notin B$, ce qui est impossible, donc $(X \in B) \cap (X = x_i) = \varnothing$, ainsi $\mathbb{P}((X \in B) \cap (X = x_i)) = 0$.
- Si $x_i \in B$, alors $(X = x_i) \subset (X \in B)$, dès lors, $(X \in B) \cap (X = x_i) = (X = x_i)$. Ainsi, $\mathbb{P}((X \in B) \cap (X = x_i)) = \mathbb{P}((X = x_i))$.

Ainsi $\mathbb{P}(X \in B) = \sum_{i=1}^{p} \mathbb{P}((X \in B) \cap (X = x_i)) = \sum_{x_i \in B} \mathbb{P}(X = x_i)$. ■

**Définition de la loi de probabilité**

Soit $X$ une variable aléatoire de $(\Omega, \mathbb{P})$ à valeurs dans $E$. La **loi** de $X$ est la famille $(\mathbb{P}_X(\{x\}))_{x \in E}$.

**Remarque 6.** Trouver la loi de probabilité de $X$ revient à déterminer la valeur de $\mathbb{P}(X = x)$ pour tout $x \in X(\Omega)$.

**Exemple 14.** Déterminer la loi de probabilité de la variable aléatoire $X_1$ de l'exemple 12.

**Remarque 7.** Soit $(p_1, \dots, p_n)$ une distribution de probabilités sur $E = \{e_1, \dots, e_n\}$, il existe une variable aléatoire $X$ définie sur un certain espace probabilisé $(\Omega, \mathbb{P})$ et à valeurs dans $E$ telle que, pour tout $i \in [\![1\,;n]\!]$, $\mathbb{P}(X = e_i) = p_i$ .

**Exemple 15.** Ainsi, il existe une variable aléatoire réelle $X$ sur $\Omega$ tel que $\mathbb{P}(X = 0) = 1/2$, $\mathbb{P}(X = 1) = 1/4$ et $\mathbb{P}(X = 2) = 1/4$ sans avoir à définir $X$ ou $\Omega$. D'ailleurs, souvent, cela ne nous importera peu.

**Définition de deux variables aléatoires de même loi**

Si $\mathbb{P}_X = \mathbb{P}_Y$, alors on note $X \sim Y$

**Exemple 16.** Soit $X$ une variable aléatoire telle que $\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = 1/2$ (on dit que $X$ est une VA de Rademacher), quelle est la loi de $Y = -X$ ?

**Attention : avoir la même loi ne veut pas dire être égales**

Si $X \sim Y$, cela ne signifie pas que $X = Y$

**Définition de l'image d'une variable aléatoire**

Soit $f : E \to F$. Alors $f \circ X$, est une nouvelle VA, notée $f(X)$ et appelée **variable aléatoire image**.

**Proposition n° 7 : probabilité de l'image d'une variable aléatoire**

Soient $X$ une variable sur $(\Omega, \mathbb{P})$ à valeurs dans $E$ et $f : E \to F$ et $Y = f \circ X = f(X)$ alors : $Y(\Omega) = f(X(\Omega))$ et

$$\forall y \in Y(\Omega) \qquad \mathbb{P}(Y = y) = \sum_{x \in f^{-1}(\{y\})} \mathbb{P}(X = x)$$

**Démonstration de la proposition n° 7 :** Notons $X(\Omega) = \{x_1, \dots, x_p\}$. Soit $y \in \mathbb{R}$. Alors,

$$\begin{aligned} y \in Y(\Omega) &\iff \exists \omega \in \Omega \quad y = Y(\omega) = f(X(\omega)) \\ &\iff \exists \omega \in \Omega \quad \exists i \in [\![ 1 ; p ]\!] \quad X(\omega) = x_i \quad y = f(x_i) \\ &\iff \exists i \in [\![ 1 ; p ]\!] \quad y = f(x_i) \iff y \in \{f(x_1), \dots, f(x_p)\} \end{aligned}$$

Dès lors, $Y(\Omega) = \{f(x_1), \dots, f(x_p)\}$. Soit $y \in \Omega$, d'après la formule des probabilités totales :

$$\mathbb{P}(Y = y) = \sum_{i=1}^{p} \mathbb{P}((Y = y) \cap (X = x_i)) = \sum_{i=1}^{p} \mathbb{P}((f(X) = y) \cap (X = x_i))$$

Calculons $\mathbb{P}(f(X) = y) \cap (X = x_i))$

- Si $f(x_i) \neq y$. Soit $\omega \in (f(X) = y) \cap (X = x_i)$, alors $X(\omega) = x_i$, et $y = f(X(\omega)) = f(x_i) \neq y$ ce qui est impossible, donc $(f(X) = y) \cap (X = x_i) = \varnothing$. Par conséquent, $\mathbb{P}((f(X) = y) \cap (X = x_i)) = 0$.
- Si $f(x_i) = y$. Remarquons que $(f(X) = y) \cap (X = x_i)) \subset (X = x_i)$. Réciproquement, soit $\omega \in (X = x_i)$, alors $X(\omega) = x_i$, et $f(X(\omega)) = f(x_i) = y$, donc $\omega \in (f(X) = y)$. Ainsi, $\omega \in (X = x_i) \cap (f(X) = y)$, on a ainsi montré l'inclusion $(X = x_i) \subset (f(X) = y) \cap (X = x_i)$. Par double inclusion, il vient $(f(X) = y) \cap (X = x_i) = (X = x_i)$, ainsi $\mathbb{P}((f(X) = y) \cap (X = x_i)) = \mathbb{P}(X = x_i)$.

Ainsi, $\mathbb{P}(Y = y) = \sum\limits_{\substack{i=1 \\ \text{si } f(x_i)=y}}^{n} \mathbb{P}(X = x_i)$. ■

**Exemple 17.** On reprend l'exemple de $X_3$. On associe un gain algébrique à ce tirage nommé $G$ : on gagne 2 euros par boule rouge obtenue et on perd 1 euro par boule non rouge. Définir $G$ en fonction de $X_3$ et déterminer $G(\Omega)$ ainsi que sa loi. Idem si on considère que le gain estle carré de la différence entre un et nombre de boules rouges.

## 2.2 Lois usuelles

**Définition de la loi uniforme (modélise le tirage au hasard de façon équitable)**

Soit $X$ une VA à valeurs dans un ensemble fini $E$. On dit que $X$ suit une loi **uniforme** sur $E$ si pour tout $e \in E$, $\mathbb{P}(X = e) = \dfrac{1}{|E|}$. On note $X \sim \mathscr{U}(E)$.

**Exemple 18.** Si $X \sim \mathscr{U}(\llbracket 1\,;n \rrbracket)$, alors $k \in \llbracket 1\,;n \rrbracket$, $\mathbb{P}(X=k) = \dfrac{1}{n}$

**Définition de la loi de Bernoulli (modélise une expérience à 2 issues)**

On dit que $X$ suit la loi de **Bernoulli** de paramètre $p \in [\,0\,;1\,]$ si $\begin{cases} \mathbb{P}(X=1) &= & p \\ \mathbb{P}(X=0) &= & 1-p \end{cases}$. On note $X \sim \mathscr{B}(p)$.

**Définition d'une loi binomiale (compte les succès dans $n$ VA de Bernoulli indépendantes)**

Si pour tout $k \in \llbracket 0\,;n \rrbracket$, $\mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$, on dit que $X$ suit une loi **binomiale** de paramètres $n \in \mathbb{N}^*$ et $p \in [\,0\,;1\,]$. On note $X \sim \mathscr{B}(n,p)$.

**Remarque 8.** Si une variable $X$ compte le nombre de succès de $n$ VA de Bernoulli indépendantes et de paramètre $p$, alors $X \sim \mathscr{B}(n,p)$ où $p$ est la probabilité de succès à chaque expérience (sera démontré plus tard).

**Exemple 19.** Soit une urne qui contient 10 billes blanches, 3 rouges et 12 noires. On tire au hasard, successivement et avec remise, 7 boules. On note $X$ la VA qui compte le nombre de boules rouges obtenues. Quelle est la loi de $X$ ?

## 2.3 Couple de variables aléatoires

Soient $X$ (resp. $Y$) une VA définie sur $(\Omega, \mathbb{P})$ à valeurs dans $E$ (resp. $F$).

**Définition de la loi conjointe**

Le couple $(X,Y)$ est une variable à aléatoire sur $(\Omega, \mathbb{P})$ à valeurs dans $E \times F$. La **loi conjointe** de $X$ et $Y$ est la loi du couple $(X,Y)$ : $\begin{cases} X(\Omega) \times Y(\Omega) & \longrightarrow [\,0\,;1\,] \\ (x,y) & \longmapsto \mathbb{P}((X=x) \cap (Y=y)) \end{cases}$.

**Remarques 9.**
- On note $\mathbb{P}(X=x, Y=y) = \mathbb{P}((X=x) \cap (Y=y))$.
- Si $p = |E|$ et $q = |F|$ sont petits, alors on consigne la loi conjointe dans un tableau à $p$ lignes et $q$ colonnes.

**Exemple 20.** On lance deux dés équilibrés et de façon indépendante et on note $X$ le plus grand nombre obtenu et $Y$ le plus petit, donner la loi conjointe de $(X,Y)$.

**Attention à l'univers image**

Ce n'est pas parce que $x$ est une issue de $X$ et que $y$ est une issue de $Y$ que $(x,y)$ est une issue de $(X,Y)$ : $(X,Y)(\Omega) \subset X(\Omega) \times Y(\Omega)$ mais il n'y a pas forcément égalité.

**Définition de la loi marginale**

On appelle **première loi marginale** de $(X,Y)$ la loi de $X$, **seconde loi marginale** de $(X,Y)$ la loi de $Y$.

**Proposition n° 8 : calcul des lois marginales en fonction de la loi conjointe**

Soit $(X,Y)$ un couple de VA, alors la loi marginale de $X$ est donnée par (idem pour $Y$) :

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

**Exemple 21.** Calculer la loi marginale de $X$ dans l'exemple 20.

**Remarque 10.** Étant donnée la loi conjointe, on peut calculer les lois marginales, par contre, les lois marginales seules ne suffisent pas à retrouver la loi conjointe.

**Exemple 22.** Donner deux exemples de variables aléatoires $(X, Y)$ tels que les lois marginales soient des lois uniformes sur $\{0; 1\}$.

**Définition loi conditionnelle de $X$ sachant un évènement $A$**

Soit $X$ une variable aléatoire sur $(\Omega, \mathbb{P})$ à valeurs dans $E$ et $A$ un évènement de $\Omega$ tel que $\mathbb{P}(A) > 0$. On appelle **loi conditionnelle** de $X$ sachant que $A$ l'application $x \mapsto \mathbb{P}(X = x|A) = \dfrac{\mathbb{P}((X = x) \cap A)}{\mathbb{P}(A)}$ définie sur $E$ et à valeurs dans $[\,0\,;1\,]$.

**Exemple 23.** En conservant l'exemple 20, calculer la loi de $X$ sachant l'évènement $(Y = 3)$.

## 2.4 Généralisation à un $n$-uplet de variables aléatoires

**Définition de la loi d'un $n$-uplet**

Soit $X = (X_1, X_2, \ldots, X_n)$ un $n$-uplet de VA, on appelle **loi conjointe** de $(X_1, X_2, \ldots, X_n)$ la loi de $X$ :

$$(x_1, x_2, \ldots, x_n) \mapsto \mathbb{P}(X = (x_1, x_2, \ldots, x_n)) = \mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

La loi de $X_i$ est appelé $i$-ième loi marginale.

**Remarque 11.** Si $X = (X_1, X_2, \ldots, X_n)$ est un $n$-uplet de VA, alors la $i$-ième marginale de $X$ est :

$$\mathbb{P}(X_i = x_i) = \sum_{(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)} \mathbb{P}((X_1, \ldots, X_n) = (x_1, \ldots, x_n))$$

## 2.5 Indépendance de deux variables aléatoires

**Définition de l'indépendance de 2 variables aléatoires**

On dit que deux VA, définies sur $\Omega$, $X$ et $Y$ sont **indépendantes** (noté $X \perp\!\!\!\perp Y$) si :

$$\forall A \subset X(\Omega) \quad \forall B \subset Y(\Omega) \qquad \mathbb{P}((X \in A) \cap (Y \in B)) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

**Proposition n° 9 : équivalence de l'indépendance**

$X \perp\!\!\!\perp Y \iff \forall (x, y) \in X(\Omega) \times Y(\Omega) \quad \mathbb{P}((X = x) \cap (Y = y)) = \mathbb{P}(X = x) \times \mathbb{P}(Y = y)$

**Démonstration de la proposition n° 9 :** Supposons que pour tout $(A, B) \in X(\Omega) \times Y(\Omega)$, $(X \in A)$ et $(Y \in B)$ sont indépendants. Alors, pour $(x, y) \in X(\omega) \times Y(\Omega)$, on pose $A = \{x\}$ et $B = \{y\}$, on a alors $(X \in A)$ et $(Y \in B)$ sont indépendants, soit $\mathbb{P}((X \in A) \cap (Y \in B)) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$. D'où

$$\mathbb{P}((X = x) \cap (Y = y)) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

Réciproquement, que pour tout $x \in X(\Omega)$ et $y \in Y(\Omega)$, $\mathbb{P}((X \in A) \cap (Y \in B)) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$. Soit $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, alors

$$\begin{aligned}
\mathbb{P}((X \in A) \cap (Y \in B)) &= \sum_{(x,y) \in A \times B} \mathbb{P}((X = x) \cap (Y = y)) \\
&= \sum_{x \in A} \sum_{y \in B} \mathbb{P}(X = x) \times \mathbb{P}(Y = y) = \sum_{x \in A} \mathbb{P}(X = x) \sum_{y \in A} \mathbb{P}(Y = y) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)
\end{aligned}$$

■

**Proposition n° 10 : indépendance de variables aléatoires images**
Si $f\colon X(\Omega) \to E$ et $g\colon Y(\Omega) \to E$, et $X \perp\!\!\!\perp Y$ alors $f(X) \perp\!\!\!\perp g(Y)$

**Démonstration de la proposition n° 10 :** Soit $x \in E$ et $y \in Y$. Alors

$$\mathbb{P}((f(X)=x)\cap(g(Y)=y)) = \mathbb{P}((X \in f^{-1}(\{x\}))\cap(Y \in g^{-1}(\{y\}))) = \mathbb{P}(X \in f^{-1}(\{x\}))\times\mathbb{P}(Y \in g^{-1}(\{y\})) = \mathbb{P}(f(X)=x)\times\mathbb{P}(g(Y)=y)$$

Ainsi, $f(X)$ et $g(Y)$ sont indépendantes. ■

**Exemple 24.** Ainsi, si $X$ et $Y$ sont des VA réelles indépendantes, alors $X^2 \perp\!\!\!\perp Y \times \sin(Y^3)$.

## 2.6 Indépendance de $n$ variables aléatoires

**Définition de l'indépendance de $n$ variables aléatoires**

On dit que $n$ variables aléatoires $X_1, \dots, X_n$ sont **indépendantes** si :

$$\forall\, (A_1, A_2, \dots, A_n) \in \mathscr{P}(X_1(\Omega)) \times \dots \times \mathscr{P}(X_n(\Omega)) \qquad \mathbb{P}\left(\bigcap_{i=1}^{n}(X_i \in A_i)\right) = \prod_{i=1}^{n}\mathbb{P}(X_i \in A_i)$$

**Remarque 12.** L'indépendance de $n$ variables aléatoires sert à modéliser la répétition d'expériences où les résultats précédents n'ont pas de conséquences sur les expériences à venir. Par exemple, si $X_k$ représente la valeur du dé lors du $k$-ième lancer, alors il paraît raisonnable de supposer que $X_1$, $X_2, \dots, X_n$ sont indépendantes.

**Proposition n° 11 : indépendance de $n$ variables aléatoires.**
$X_1, \dots, X_n$ sont des variables aléatoires indépendantes si et seulement si :

$$\forall\, (x_1, x_2, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega) \qquad \mathbb{P}\left(\bigcap_{i=1}^{n}(X_i = x_i)\right) = \prod_{i=1}^{n}\mathbb{P}(X_i = x_i)$$

**Remarque 13.** Si $X_1, \dots, X_n$ sont indépendantes et $i \neq j$ alors $X_i$ et $X_j$ sont indépendantes (indépendance deux à deux). Plus généralement si $X_1, \dots, X_n$ sont indépendantes, alors toute sous-famille est indépendante.

**Péril imminent la réciproque est fausse**
Si $X$, $Y$ et $Z$ sont telles que $X$ et $Y$ sont indépendantes, $X$ et $Z$ aussi et $Y$ et $Z$ également, cela n'implique pas que $X$, $Y$ et $Z$ sont indépendantes. De même, avec plus de trois variables aléatoires.

**Exemple 25.** Soit $X$ tel que $\mathbb{P}(X=1) = \mathbb{P}(X=-1) = \dfrac{1}{2}$, $Y$ indépendante de $X$ de même loi, on pose $Z = XY$, alors $X \perp\!\!\!\perp Y$, $X \perp\!\!\!\perp Z$ et $Y \perp\!\!\!\perp Z$. Mais, $X$, $Y$ et $Z$ ne sont pas indépendantes.

**Solution de l'exemple 25 :** En effet, $Z(\Omega) = \{-1, 1\}$,

$$\begin{aligned}\mathbb{P}(Z=1) &= \mathbb{P}((X=Y=1)\cup(X=Y=-1)) = \mathbb{P}(X=Y=1)+\mathbb{P}(X=Y=-1)\\ &= \mathbb{P}(X=1)\mathbb{P}(Y=1)+\mathbb{P}(X=-1)\mathbb{P}(Y=-1) = \frac{1}{2}\end{aligned}$$

On a également $\mathbb{P}(Z=-1) = 1-\mathbb{P}(Z=1) = \dfrac{1}{2}$. De plus, pour tout $(a,b) \in \{-1,1\}^2$,

$$\mathbb{P}((Z=a)\cap(X=b)) = \mathbb{P}(Y=a/b\cap X=b) = \frac{1}{4} = \mathbb{P}(Z=a)\times\mathbb{P}(X=b)$$

Ainsi, $X$ et $Z$ sont indépendants. De même, $Y$ et $Z$ sont indépendants. Cependant

$$\mathbb{P}((X=1)\cap(Y=1)\cap(Z=-1)) = 0 \neq \mathbb{P}(X=1)\mathbb{P}(Y=1)\mathbb{P}(Z=1) = \frac{1}{8}$$

**Exemple fondamental : somme de $n$ variables aléatoires de Bernoulli**
Soit $X_1, \dots, X_n$ $n$ VA indépendantes, de même loi : $X_i \sim \mathscr{B}(p)$, alors $X = X_1 + \dots + X_n \sim \mathscr{B}(n,p)$.

**Démonstration de l'exemple fondamental :** Tout d'abord, comme $X_i(\Omega) = \{0,1\}$, on a $X(\Omega) \subset [\![ 0\,;n ]\!]$. Soit $k \in [\![ 0\,;n ]\!]$. Notons $A_k = \{(x_1, x_2, \dots, x_n) \in \{0,1\}^n, \quad \sum_{i=1}^{n} x_i = k\}$. Ainsi par union disjointe puis par indépendance puis en utilisant le fait que $\text{Card}(A_k) = \binom{n}{k}$, on a

$$\begin{aligned}
\mathbb{P}(X=k) &= \mathbb{P}\left( \bigcup_{(x_1,x_2,\dots,x_n)\in A_k} (X_1 = x_1) \cap (X_2 = x_2) \cap \dots \cap (X_n = x_n) \right) \\
&= \sum_{(x_1,x_2,\dots,x_n)\in A_k} \mathbb{P}((X_1 = x_1) \cap (X_2 = x_2) \cap \dots \cap (X_n = x_n)) \\
&= \sum_{(x_1,x_2,\dots,x_n)\in A_k} \mathbb{P}(X_1 = x_1) \times \mathbb{P}(X_2 = x_2) \times \dots \times \mathbb{P}(X_n = x_n) \\
&= \sum_{(x_1,x_2,\dots,x_n)\in A_k} p^k (1-p)^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}
\end{aligned}$$

**Théorème n° 4 : lemme des coalitions** *(admis)*

Soit $(X_1, X_2, \dots, X_n)$ une famille de variables aléatoires indépendantes à valeurs dans $E$, soit $f\colon E^m \to F$ et $g\colon E^{n-m} \to F$, alors les variables aléatoires $f(X_1, X_2, \dots, X_m)$ et $g(X_{m+1}, X_{m+2}, \dots, X_n)$ sont indépendantes.

**Remarque 14.** On peut faire plus de deux coalitions : les variables aléatoires $f(X_1, \dots, X_4)$, $g(X_5, \dots, X_8)$, $h(X_9, X_{10})$, ..., $m(X_{20}, \dots, X_{25})$ sont indépendantes.

**Exemple 26.** Si $X_1$, $X_2, \dots, X_7$ sont indépendantes, alors $X_1 X_2$, $\exp(X_3) + X_4 \sin(X_5)$, $X_6$, ${X_7}^8$ sont indépendantes.

# 3 Espérance et Variance

On cherche à calculer des indicateurs permettant de décrire $X$ une variable aléatoire réelle. On se ne travaillera donc dans cette partie qu'avec des variables aléatoires réelles.

## 3.1 Espérance

**Définition de l'espérance**

Pour $X(\Omega) = \{x_1, \dots, x_p\}$, on appelle **espérance** de $X$ : $\mathbb{E}(X) = \sum_{k=1}^{p} x_k \mathbb{P}(X = x_k) = \sum_{x \in X(\Omega)} x \mathbb{P}(X = x)$

On dit que $X$ est **centrée** si $\mathbb{E}(X) = 0$.

**Remarque 15.** $\mathbb{E}(X)$ est une moyenne pondérée (par les probabilités) des valeurs prises par $X$.

**Exemples : espérances usuelles à connaître**

1. Si $X \sim \mathscr{U}([\![ 1\,;n ]\!])$ alors $\mathbb{E}(X) = \dfrac{n+1}{2}$
2. Si $X \sim \mathscr{B}(p)$ alors $\mathbb{E}(X) = p$
3. Si $X \sim \mathscr{B}(n,p)$ alors $\mathbb{E}(X) = np$
4. Si $A$ est un évènement de $\Omega$, alors $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$

**Démonstration des espérances usuelles :** Montrons ces résultats :

1. Si $X \sim \mathscr{U}([\![ 1\,;n ]\!])$, alors $X(\omega) = [\![ 1\,;n ]\!]$, ainsi
$$\mathbb{E}(X) = \sum_{k=1}^{n} k\mathbb{P}(X=k) = \sum_{k=1}^{n} \frac{k}{n} = \frac{1}{n} \sum_{k=1}^{n} = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}$$
2. Si $X \sim \mathscr{B}(p)$, alors $X(\Omega) = \{0,1\}$, donc $\mathbb{E}(X) = 0\mathbb{P}(X=0) + 1\mathbb{P}(X=1) = p$.

3. Si $X \sim \mathscr{B}(n,p)$, alors $X(\Omega) = [\![ 0\,;n ]\!]$, donc

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

Or[1], pour $k \in [\![ 1\,;n ]\!]$, $k\binom{n}{k} = n\binom{n-1}{k-1}$ On obtient donc

$$\mathbb{E}(X) = \sum_{k=1}^{n} n\binom{n-1}{k-1}p^k(1-p)^{n-k} \underset{j=k-1}{=} n\sum_{j=0}^{n-1}\binom{n-1}{j}p^{j+1}(1-p)^{n-(j+1)} = np\sum_{j=0}^{n-1}\binom{n-1}{j}p^j(1-p)^{(n-1)-j}$$

En reconnaissant la formule du binôme de Newton, on trouve $\mathbb{E}(X) = np(p+(1-p))^{n-1} = np$.

4. $X(\Omega) = \{0,1\}$, $\mathbb{E}(X) = 0\mathbb{P}(X=0) + 1\mathbb{P}(X=1) = \mathbb{P}(\{\omega \in \Omega,\ \mathbb{1}(\omega)=1\}) = \mathbb{P}(A)$. ■

**Proposition n° 12 : écriture théorique de $\mathbb{E}(X)$**

Si $\Omega = \{\omega_1, \omega_2, \dots, \omega_m\}$ $\quad \mathbb{E}(X) = \sum_{i=1}^{m} X(\omega_i)\mathbb{P}(\{\omega_i\}) = \sum_{\omega\in\Omega} X(\omega)\mathbb{P}(\{\omega\})$

**Démonstration de la proposition n° 12 :** Soit $x \in X(\Omega)$, alors $(X=x) = \bigcup_{\substack{\omega\in\Omega \\ \text{si } X(\omega)=x}} \{\omega\}$ (union d'ensembles deux à deux disjoints) Ainsi, $\mathbb{P}(X=x) = \sum_{\substack{\omega\in\Omega \\ \text{si } X(\omega)=x}} \mathbb{P}(\{\omega\})$. Dès lors,

$$\mathbb{E}(X) = \sum_{x\in X(\Omega)} x\mathbb{P}(X=x) = \sum_{x\in X(\Omega)} x \sum_{\substack{\omega\in\Omega \\ \text{si } X(\omega)=x}} \mathbb{P}(\{\omega\}) = \sum_{\omega\in\Omega} X(\omega)\mathbb{P}(\{\omega\})$$ ■

**Proposition n° 13 : propriétés de l'espérance**

Soient $X$ et $Y$ deux variables aléatoires et $\lambda \in \mathbb{R}$ alors :

1. Linéarité de l'espérance : $\mathbb{E}(\lambda X + Y) = \lambda\mathbb{E}(X) + \mathbb{E}(Y)$
2. Espérance d'une variable aléatoire constante : $\mathbb{E}(\lambda) = \lambda$
3. Positivité de l'espérance : si $X \geqslant 0$ (pour tout $\omega \in \Omega, X(\omega) \geqslant 0$), alors $\mathbb{E}(X) \geqslant 0$
4. Croissance de l'espérance : si $X \leqslant Y$ $(\forall \omega \in \Omega, X(\omega) \leqslant Y(\omega))$ alors, $\mathbb{E}(X) \leqslant \mathbb{E}(Y)$
5. Inégalité triangulaire : $|\mathbb{E}(X)| \leqslant \mathbb{E}(|X|)$

**Exemple 27.** Si $X \sim \mathscr{B}(n,p)$, alors $\mathbb{E}(X) = np$

**Théorème n° 5 formule de transfert**

Soient $X$ une VA à valeurs dans $E$ et $f : E \to \mathbb{R}$ et $Y = f(X)$ alors : $\mathbb{E}(Y) = \sum_{x\in X(\Omega)} f(x)\mathbb{P}(X=x)$

1. Formule du maire : dans une ville de $n$ citoyens, on a un conseil municipal de $k$ personnes dont un maire (avec $k \geqslant 1$ forcément) Combien y a-t-il de possibilités ? Pour choisir le conseil municipal de $k$ personnes parmi les $n$ citoyens, il y a $\binom{n}{k}$ possibilités. Une fois ce conseil municipal choisi, il y a $k$ choix possibles pour le maire. Cela fait donc $k\binom{n}{k}$. Sinon, on choisit d'abord le maire parmi les $n$ citoyens de la ville, il reste $k-1$ personnes à choisir pour le conseil municipal parmi les $n-1$ citoyens restant. Cela fait $\binom{n-1}{k-1}$ possibilités, en multipliant par $n$, cela fait $n\binom{n-1}{k-1}$, on a donc compter de deux façons différente la même chose, donc $k\binom{n}{k} = n\binom{n-1}{k-1}$.

**Démonstration du théorème n° 5 :** En utilisant l'expression de la loi de $Y$ en fonction de $X$, on obtient :

$$\mathbb{E}(Y) = \sum_{y\in Y(\Omega)} y\mathbb{P}(Y=y) = \sum_{y\in Y(\Omega)} y \sum_{\substack{x\in X(\Omega)\\ \text{si } f(x)=y}} \mathbb{P}(X=x) = \sum_{y\in Y(\Omega)} \sum_{\substack{x\in X(\Omega)\\ \text{si } f(x)=y}} f(x)\mathbb{P}(X=x) = \sum_{x\in X(\Omega)} f(x)\mathbb{P}(X=x)$$

■

**Remarque 16.** Ce théorème permet de calculer $\mathbb{E}(Y)$ sans connaître la loi de $Y$ mais à partir de la loi de $X$. On remarque que ce théorème s'applique si $X$ est un couple ou un $n$-uplet de variables aléatoires.

**Exemples 28.**

1. Si $X \sim \mathscr{U}(\llbracket 1\,;n\,\rrbracket)$ calculer $\mathbb{E}(X^2)$
2. Si $X \sim \mathscr{U}(\llbracket 1\,;n\,\rrbracket)$ calculer $\mathbb{E}(3^X)$
3. Si $X \sim \mathscr{B}(n,p)$, calculer $\mathbb{E}(X^2)$

**Proposition n° 14 : inégalité de Markov**

Soient $a>0$ et $X$ une variable aléatoire **positive**. Alors $\mathbb{P}(X \geqslant a) \leqslant \dfrac{\mathbb{E}(X)}{a}$

**Démonstration de la proposition n° 14 :** Comme, $X$ est positive $x\mathbb{P}(X=x) \geqslant 0$ pour $x \in X(\omega)$ et $x<a$, ainsi :

$$\mathbb{E}(X) = \sum_{x\in X(\Omega)} x\mathbb{P}(X=x) \geqslant \sum_{\substack{x\in X(\Omega)\\ x\geqslant a}} x\mathbb{P}(X=x) \geqslant \sum_{\substack{x\in X(\omega)\\ x\geqslant a}} a\mathbb{P}(X=x) = a\mathbb{P}(X\geqslant a)$$

■

**Exemple 29.** Si dans une classe, la moyenne au DS est 6, alors que peut-on dire de la probabilité d'avoir une note supérieure ou égale à 18 ?

**Proposition n° 15 : espérance du produit de variables aléatoires indépendantes**

Soient $X$ et $Y$ deux VA indépendantes alors $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$

Si $X_1, X_2 \ldots, X_n$ sont indépendantes alors $\mathbb{E}\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} \mathbb{E}(X_i)$

**Démonstration de la proposition n° 15 :** En utilisant l'écriture théorique de l'espérance :

$$\begin{aligned}
\mathbb{E}(XY) &= \sum_{\omega\in\Omega} (XY)(\omega)\mathbb{P}(\{\omega\}) = \sum_{x\in X(\Omega)} \sum_{y\in Y(\Omega)} \sum_{\omega\in(X=x)\cap(Y=y)} X(\omega)Y(\omega)\mathbb{P}(\{\omega\}) \\
&= \sum_{x\in X(\Omega)} \sum_{y\in Y(\Omega)} xy \sum_{\omega\in(X=x)\cap(Y=y)} \mathbb{P}(\{\omega\}) \\
&= \sum_{x\in X(\Omega)} \sum_{y\in Y(\Omega)} xy\mathbb{P}((X=x)\cap(Y=y)) \\
&= \sum_{x\in X(\Omega)} \sum_{y\in Y(\Omega)} xy\mathbb{P}(X=x)\times\mathbb{P}(Y=y) \\
&= \sum_{x\in X(\Omega)} x\mathbb{P}(X=x) \sum_{y\in Y(\Omega)} y\mathbb{P}(Y=y) = \mathbb{E}(X)\mathbb{E}(Y)
\end{aligned}$$

■

**Péril imminent : la réciproque est fausse**

Si $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$, ça ne prouve pas forcément que $X$ et $Y$ sont indépendantes.

**Exemple 30.** $X \sim \mathscr{U}(\llbracket -1\,;1 \rrbracket)$ et $Y = 1 - X^2$

**Solution de l'exemple 30 :** En effet, $XY = X(1-X^2) = 0$. Or, $\mathbb{E}(X) = 0$ et $\mathbb{E}(XY) = 0$. De plus,

$$\mathbb{P}((X=0) \cap (Y=0)) = \mathbb{P}((X=0) \cap (X^2=1)) = 0$$

Tandis que $\mathbb{P}(X=0) \times \mathbb{P}(Y=0) = \frac{1}{3} \times \frac{2}{3}$. Ainsi, $X$ et $Y$ ne sont indépendantes mais $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

## 3.2 Variance

**Définition de la variance et de l'écart-type**

Soit $X$ une variable aléatoire, on appelle **variance** de $X$ le réel positif $\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$
On appelle **écart-type** de $X$ le réel positif $\sigma(X) = \sqrt{\mathbb{V}(X)}$

**Remarque 17.** La variance mesure la moyenne les carrés des écarts de $X$ par rapport à $\mathbb{E}(X)$.
Par la formule de transfert, $\mathbb{V}(X) = \sum_{x \in X(\Omega)} (x - \mathbb{E}(X))^2 \mathbb{P}(X = x)$.

**Proposition n° 16 : propriétés de la variance**

Soit $X$ une variable aléatoire alors :

1. $\mathbb{V}(X) \geqslant 0$
2. $\forall (a,b) \in \mathbb{R}^2 \quad \mathbb{V}(aX+b) = a^2\mathbb{V}(X)$ (la variance est quadratique)
3. $\mathbb{V}(X) = 0 \iff \mathbb{P}(X = \mathbb{E}(X)) = 1$ ($X$ est presque sûrement constante)
4. $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$ (formule de König-Huygens)
5. Si $\mathbb{V}(X) > 0$, alors $Y = \dfrac{X - \mathbb{E}(X)}{\sigma(X)}$ est une variable centrée réduite ($\mathbb{E}(Y) = 0$ et $\mathbb{V}(Y) = 1$)

**Démonstration de la proposition n° 16 :**

1. Comme $(X - \mathbb{E}(X))^2 \geqslant 0$ par positivité de l'espérance, $\mathbb{V}(X) \geqslant 0$.
2. $\mathbb{V}(aX+b) = \mathbb{E}((aX+b-\mathbb{E}(aX+b))^2) = \mathbb{E}((aX+b-a\mathbb{E}(X)-b)^2) = \mathbb{E}(a^2(X-\mathbb{E}(X))^2) = a^2\mathbb{V}(X)$
3. Pour cette démonstration on va séparer les valeurs que peut prendre $X$ entre celles qui ont une probabilité non nulle et celles qui ont une probabilité nulle : notons $B = \{x \in X(\Omega),\ \mathbb{P}(X=x) = 0\}$ et $C = X(\Omega) \backslash B$, alors $X(\Omega) = B \cup C$ et $B \cap C = \varnothing$, ainsi
$$\mathbb{V}(X) = \sum_{x \in B} (x - \mathbb{E}(X))^2 \underbrace{\mathbb{P}(X=x)}_{=0} + \sum_{x \in C} (x - \mathbb{E}(X))^2 \underbrace{\mathbb{P}(X=x)}_{\neq 0} = \sum_{x \in C} (x - \mathbb{E}(X))^2 \mathbb{P}(X=x)$$
Ainsi, supposons que $\mathbb{V}(X) = \sum_{x \in C} (x - \mathbb{E}(X))^2 \mathbb{P}(X=x) = 0$ alors $x \in C$, $(x - \mathbb{E}(X))^2 \mathbb{P}(X=x) = 0$ (somme nulle de termes positifs) alors pour tout $x \in C$, $x = \mathbb{E}(X)$. Ainsi, $C = \{\mathbb{E}(X)\}$ Or,
$$\mathbb{P}(X \in B) = \sum_{x \in B} \mathbb{P}(X = x) = 0$$
Donc $\mathbb{P}(X \in C) = 1 - \mathbb{P}(X \in B) = 1$, ainsi $\mathbb{P}(X = \mathbb{E}(X)) = 1$. Ainsi, $X$ est quasi-certaine égale à $\mathbb{E}(X)$. Réciproquement, supposons que $X$ est quasi-certaine égale à $\mathbb{E}(X)$. Alors $X(\Omega) = \{\mathbb{E}(X), x_2, \ldots, x_p\}$, avec $\mathbb{P}(X = \mathbb{E}(X)) = 1$ et donc $i \in \llbracket 2\,;p \rrbracket$, $\mathbb{P}(X = x_i) = 0$. En réutilisant la formule de transfert,
$$\mathbb{V}(X) = \sum_{x \in X(\Omega)} (x - \mathbb{E}(X))^2 \mathbb{P}(X=x) = (\mathbb{E}(X) - \mathbb{E}(X))^2 \mathbb{P}(X = \mathbb{E}(X)) + 0 = 0$$
4. En utilisant une identité remarquable et la linéarité de l'espérance :
$$\begin{aligned} \mathbb{V}(X) &= \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \mathbb{E}\left(X^2 - 2X\mathbb{E}(X) + \mathbb{E}(X)^2\right) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \end{aligned}$$ ■

**Exemples de variances des lois usuelles à connaître :**

1. Si $X \sim \mathscr{U}(\llbracket 1\,;n \rrbracket)$, alors $\mathbb{V}(X) = \dfrac{n^2-1}{12}$
2. Si $X \sim \mathscr{B}(p)$, alors $\mathbb{V}(X) = p(1-p)$
3. Si $X \sim \mathscr{B}(n,p)$, alors $\mathbb{V}(X) = np(1-p)$

**Démonstration des variances usuelles à connaître :** En utilisant la formule de transfert pour calculer $\mathbb{E}(X^2)$, on obtient :

1. $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \sum_{k=1}^{n} k^2 \frac{1}{n} - \frac{(n+1)^2}{4} = \frac{n(n+1)(2n+1)}{6n} - \frac{(n+1)^2}{4} = \frac{n+1}{12}(4n+2-(3n+3)) = \frac{n^2-1}{12}$.
2. Comme $X^2 = X$, on a $\mathbb{V}(X) = \mathbb{E}(X) - \mathbb{E}(X)^2 = p - p^2 = p(1-p)$.
3. En utilisant [2] $k(k-1)\binom{n}{p} = n(n-1)\binom{n-2}{k-2}$ valable pour $k \in [\![ 2\,;n ]\!]$, on obtient :

$$\begin{aligned}
\mathbb{V}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{E}(X(X-1)) + \mathbb{E}(X) - \mathbb{E}(X)^2 = \sum_{k=0}^{n} \left( k(k-1)\binom{n}{k} p^k (1-p)^{n-k} \right) + np - (np)^2 \\
&= \sum_{k=2}^{n} \left( k(k-1)\binom{n}{k} p^k (1-p)^{n-k} \right) + np - (np)^2 \\
&= \sum_{k=2}^{n} \left( n(n-1)\binom{n-2}{k-2} p^k (1-p)^{n-k} \right) + np - (np)^2 \\
&= n(n-1)p^2 \sum_{j=0}^{n-2} \left( \binom{n-2}{j} p^j (1-p)^{n-2-j} \right) + np - (np)^2 \\
&= n(n-1)p^2 (p + (1-p))^{n-2} + np - n^2p^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p)
\end{aligned}$$

### Définition de la covariance de deux variables aléatoires

Soit $X$ et $Y$ deux VA réelles, on définit la **covariance** de $X$ et $Y$ par $\quad \operatorname{cov}(X,Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y)))$
On dit que $X$ et $Y$ sont **non corrélées** si $\operatorname{cov}(X,Y) = 0$.

### Proposition n° 17 : formule de la covariance

Si $X$ et $Y$ sont deux VA réelles, alors $\quad \operatorname{cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$

**Remarque 18.**
- $\mathbb{V}(X) = \operatorname{cov}(X,X)$
- Si $X$ et $Y$ sont indépendantes, alors $X$ et $Y$ sont non corrélées.

### Attention la réciproque est fausse

Il est possible d'être non corrélées sans être indépendantes.

**Exemple 31.** Si $X \sim \mathscr{U}(\{-1,0,1\})$, alors $\operatorname{cov}(X, X^2) = 0$ alors que $X$ et $X^2$ ne sont pas indépendantes.

### Proposition n° 18 : variance de la somme de variables aléatoires

1. Soient $X$ et $Y$ deux VA alors $\quad \mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\operatorname{cov}(X,Y)$
2. Si $X$ et $Y$ sont non corrélées (ou indépendantes), alors $\quad \mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y)$
3. Soient $X_1, \dots, X_n$, $n$ variables aléatoires, $\quad \mathbb{V}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{V}(X_i) + 2 \sum_{1 \leqslant i < j \leqslant n} \operatorname{cov}(X_i, X_j)$
4. Si $X_1, \dots, X_n$ sont deux à deux non corrélées (ou deux à deux indépendantes), alors $\mathbb{V}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{V}(X_i)$

2. Formule du maire et de son adjoint : dans une ville de $n$ citoyens, on veut un conseil municipal de $k$ personnes dont un maire et un adjoint (avec $k \geqslant 2$ donc). Soit on choisit d'abord le conseil de $k$ personnes, cela fait $\binom{n}{k}$ possibilités, parmi ce groupe, on a $k$ choix pour le maire, à ce maire fixé, il reste $k-1$ pour son adjoint, soit $k(k-1)\binom{n}{k}$ possibilités. Soit on choisit le maire parmi les $n$ personnes, puis son adjoint parmi les $n-1$ personnes restantes, cela fait $n(n-1)$ possibilités, il reste à choisir les $k-2$ personnes pour former le conseil municipal parmi les $n-2$ citoyens restants, soit $n(n-1)\binom{n-2}{k-2}$ possibilités.

**Démonstration de la proposition n° 18 :**

- En utilisant la formule de Koënig-Huygens, les identités remarquables et la linéarité de l'espérance et ce qui précède, on obtient :

$$\begin{aligned}\mathbb{V}(X+Y) &= \mathbb{E}((X+Y)^2) - \mathbb{E}(X+Y)^2 \\ &= \mathbb{E}(X^2+2XY+Y^2) - (\mathbb{E}(X)+\mathbb{E}(Y))^2 \\ &= \mathbb{E}(X^2) + 2\mathbb{E}(XY) + \mathbb{E}(Y^2) - \mathbb{E}(X)^2 - 2\mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(Y)^2 = \mathbb{V}(X) + \mathbb{V}(Y) + 2\text{cov}(X,Y)\end{aligned}$$

- 
- Posons $X = \sum_{i=1}^{n} X_i$, par linéarité de l'espérance, $\mathbb{E}(X) = \sum_{i=1}^{n} \mathbb{E}(X_i)$.

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{E}\left((X-\mathbb{E}(X))^2\right) = \mathbb{E}\left(\left(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n}\mathbb{E}(X_i)\right)^2\right) = \mathbb{E}\left(\left(\sum_{i=1}^{n}(X_i-\mathbb{E}(X_i))\right)^2\right) \\ &= \mathbb{E}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}(X_i-\mathbb{E}(X_i))(X_j-\mathbb{E}(X_j))\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}\left(X_iX_j - \mathbb{E}(X_i)X_j - X_i\mathbb{E}(X_j) + \mathbb{E}(X_i)\mathbb{E}(X_j)\right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}(X_iX_j) - \mathbb{E}(X_i)\mathbb{E}(X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j) + \mathbb{E}(X_i)\mathbb{E}(X_j) = \sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}(X_iX_j) - \mathbb{E}(X_i)\mathbb{E}(X_j)\end{aligned}$$

- 

■

**Remarque 19.** Si $X \sim \mathscr{B}(n,p)$, alors $\mathbb{V}(X) = np(1-p)$

**Théorème n° 6 : inégalité de Bienaymé-Tchebychev**

Soient $X$ une variable aléatoire et $\varepsilon > 0$, alors $\mathbb{P}(|X-\mathbb{E}(X)| \geqslant \varepsilon) \leqslant \dfrac{\mathbb{V}(X)}{\varepsilon^2}$

**Démonstration du théorème n° 6 :** Par croissance des applications $x \mapsto x^2$ et $x \mapsto \sqrt{x}$ sur $\mathbb{R}_+$, $(|X-\mathbb{E}(X)| \geqslant \varepsilon) = (|X-\mathbb{E}(X)|^2 \geqslant \varepsilon^2)$. Ainsi, par application de l'inégalité de Markov, à la variable aléatoire **positive** $|X-\mathbb{E}(X)|^2$ :

$$\mathbb{P}(|X-\mathbb{E}(X)| \geqslant \varepsilon) = \mathbb{P}(|X-\mathbb{E}(X)|^2 \geqslant \varepsilon^2) \leqslant \frac{1}{\varepsilon^2}\mathbb{E}(|X-\mathbb{E}(X)|^2) = \frac{\mathbb{V}(X)}{\varepsilon^2}$$ ■

**Exemple 32.** Supposons qu'on ait dé dont la probabilité d'obtention un six est notée $p$. Pour approximer $p$, on lance ce dé $n$ fois et on note $F$ la fréquence du six. Pour quelle valeur de $n$ la probabilité pour que $F$ soit une approximation de $p$ à 0.01 près est-elle supérieure à 0.9 ?

**Remarque 20.** Sur cet exemple, on a montré que la probabilité d'un évènement est la limite de la fréquence de cet évènement lorsque l'on répète un évènement un «grand» nombre de fois. Cela est conforme à l'intuition, dire qu'un dé est équilibré indique que si on le lance un très grand nombre de fois la fréquence d'une face doit tendre vers 1/6. Cela permet de relier la probabilité à cette notion intuitive ce que l'on avait soigneusement évité jusqu'à présent.

# 4 Tableau récapitulatif des lois usuelles

Les caractéristiques de ce tableau doivent être absolument connu par cœur pour les quatre premières variables aléatoires. Les deux dernières seront vu en PC/PSI. Ainsi l'année prochaine, vous pourrez réviser toutes les lois usuelles sur ce tableau.

| Nom de la loi | Paramètre | Univers image | Loi de probabilité | Espérance | Variance | Interprétation |
|---|---|---|---|---|---|---|
| Constante | $a \in \mathbb{R}$ | $\{a\}$ | $\mathbb{P}(X = a) = 1$ | $a$ | $0$ | Constante |
| Bernoulli | $p \in [\,0\,;1\,]$ | $\{0,1\}$ | $\mathbb{P}(X = 1) = p$<br>$\mathbb{P}(X = 0) = 1 - p$ | $p$ | $p(1-p)$ | Succès vs échec |
| Binomiale | $(p,n) \in [\,0\,;1\,] \times \mathbb{N}^*$ | $[\![\,0\,;n\,]\!]$ | $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$<br>pour $k \in [\![\,0\,;n\,]\!]$ | $np$ | $np(1-p)$ | Nombre de succès dans $n$ Va de Bernoulli de paramètre $p$ indépendantes |
| Uniforme | $n$ | $[\![\,1\,;n\,]\!]$ | $\mathbb{P}(X = k) = \dfrac{1}{n}$<br>pour $k \in [\![\,1\,;n\,]\!]$ | $\dfrac{n+1}{2}$ | $\dfrac{n^2-1}{12}$ | Tirage équitable |
| Géométrique | $p \in ]\,0\,;1\,[$ | $\mathbb{N}^*$ | $\mathbb{P}(X = k) = p(1-p)^{k-1}$<br>pour $k \in \mathbb{N}^*$ | $\dfrac{1}{p}$ | $\dfrac{1-p}{p^2}$ | Donne le premier succès dans une suite de Va indépendantes de Bernoulli de paramètre $p$ |
| Poisson | $\lambda > 0$ | $\mathbb{N}$ | $\mathbb{P}(X = k) = \dfrac{\lambda^k}{k!} \mathrm{e}^{-\lambda}$<br>pour $k \in \mathbb{N}$ | $\lambda$ | $\lambda$ | Désintégration radioactive, arrivé dans une file d'attente, évènements rares etc. |