I. ESPACES PROBABILISÉS

1. Probabilité sur un univers fini

On note $\Omega = \{\omega_1, \dots, \omega_n\}$ l'univers de notre expérience.
On souhaite définir la probabilité des événements de manière à modéliser au mieux la situation étudiée.
En général, on cherche à mettre en évidence une situation d'équiprobabilité, ou on définit la probabilité des événements à partir de fréquences d'observation sur une grand nombre de réalisation de l'expérience.
Ici, nous allons définir la notion théorique de probabilité que notre modèle devra respecter.

Définition 1 : On appelle (mesure de) **probabilité** sur Ω toute application $\mathbb{P} : \mathscr{P}(\Omega) \to [0,1]$ telle que

a. $\mathbb{P}(\Omega) = 1$

b. Pour tous événements incompatibles A et B, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

2. Propriété d'une probabilité

Proposition 1

Soit $(\Omega, \mathscr{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.

a. $\mathbb{P}(\emptyset) = 0$

b. $\forall A \in \mathscr{P}(\Omega)$, $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$ et $\forall A,\ B \in \mathscr{P}(\Omega)$, $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$

c. $\forall A,\ B \in \mathscr{P}(\Omega)$, $A \subset B \implies \mathbb{P}(A) \leqslant \mathbb{P}(B)$

d. $\forall A,\ B \in \mathscr{P}(\Omega)$, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

3. Probabilité conditionnelle

Définition 2 : Soit B un événement de probabilité non nulle. Étant donné un événement A, la probabilité de A sachant B est $\mathbb{P}(A|B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$.

Remarque 1 :

- Explication : A et B sont deux événements liés à l'épreuve $\mathscr{E}$. On répète n fois cette épreuve et on observe n_B fois l'événement B et $n_{A\cap B}$ l'événement $A \cap B$. La fréquence de A quand B est réalisé est $\dfrac{n_{A\cap B}}{n_B}$, ce que l'on pose égal à $\mathbb{P}(A|B)$.
  On a alors $\mathbb{P}(A|B) = \dfrac{n_{A\cap B}}{n} \dfrac{n}{n_B}$, et en considérant que $\mathbb{P}(A \cap B) \simeq \frac{n_{A\cap B}}{n}$ et $\mathbb{P}(B) \simeq \frac{n_B}{n}$ on retrouve notre définition.
- On note aussi $\mathbb{P}_B(A)$ la probabilité de A sachant B.

Proposition 2

Soit B un événement de probabilité non nulle.
L'application $\mathbb{P}_B : \begin{array}{ccc} \mathscr{P}(\Omega) & \to & [0,1] \\ A & \mapsto & \mathbb{P}(A|B) \end{array}$ est une (mesure de) probabilité sur Ω.

Théorème 1 Formule des probabilités composées

Soit $(A_i)_{1 \leqslant i \leqslant m}$ des événements tels que $\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_m) \neq 0$. On a alors :

$$\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_m) = \mathbb{P}(A_m | A_1 \cap \dots \cap A_{m-1})\ \mathbb{P}(A_{m-1} | A_1 \cap \dots \cap A_{m-2}) \dots\ \mathbb{P}(A_2 | A_1)\ \mathbb{P}(A_1)\ .$$

Théorème 2 Formule de Bayes

Soit A et B deux événements de probabilité non nulle. On a $\mathbb{P}(B|A) = \dfrac{\mathbb{P}(B)}{\mathbb{P}(A)} \mathbb{P}(A|B)$.

4. Formule des probabilités totales

Soit $(\Omega, \mathscr{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.

Définition 3 : Soit $n \in \mathbb{N}^*$ et $(A_i)_{i\in[\![1,n]\!]}$ une famille finie d'événements.
On dit que $(A_i)_{i\in[\![1,n]\!]}$ est un **système complet d'événements** si
- les événements A_i sont 2 à 2 incompatibles.
- la réunion $\bigcup\limits_{i\in[\![1,n]\!]} A_i$ est égale à Ω.

Exemple 1
- Soit A un événement, $(A, \overline{A})$ forme un système complet d'événements.
- On lance un dé équilibré et note le résultat. On note A_i : obtenir i, pour $i \in [\![1,6]\!]$.
  $(A_i)_{1\leqslant i\leqslant 6}$ forme un système complet d'événements.
- On lance une pièce équilibrée, on la lance n fois et on note le résultat. On note P_i (resp. F_i) "obtenir PILE (resp. FACE) au i-ème lancer". (P_1, F_1) est un système complet d'événements mais pas (P_1, F_2) ...

Théorème 3 Formule des probabilités totales

Soit $(A_i)_{i\in[\![1,n]\!]}$ un système complet fini d'événements.
Pour tout événement B, $\mathbb{P}(B) = \sum\limits_{i=1}^{n} \mathbb{P}(B\cap A_i) = \sum\limits_{i=1}^{n} \mathbb{P}(B|A_i)\mathbb{P}(A_i)$.

5. Indépendance

Soit $(\Omega, \mathscr{P}(\Omega), \mathbb{P})$ un espace probabilisé.

Définition 4 : Soient A et B deux événements. A et B sont dits **indépendants** si $\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Remarque 2 :
- A et B sont indépendants si et seulement si $\mathbb{P}(A|B) = \mathbb{P}(A)$. Cela signifie que la probabilité que A soit réalisé ne dépend pas du fait que B soit réalisé ou non (on peut échanger les rôles de A et B dans cette remarque).
- Un événement de probabilité nulle ou égale à 1 est indépendant de tout autre événement.
- Deux événements incompatibles et de probabilités non nulles **ne sont pas** indépendants.

Définition 5 : Soit $(A_i)_{i\in[\![1,m]\!]}$ une famille finie d'événements.
- On dit que les A_i sont **mutuellement indépendants** si pour tout $k \in [\![2,m]\!]$,
  pour tous $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant m$, $\mathbb{P}(A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_k}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2})\ldots\mathbb{P}(A_{i_k})$.
- On dit que les A_i sont **deux à deux indépendants** si
  pour tout $(i,j) \in [\![1,m]\!]^2$ tels que $i \neq j$, $\mathbb{P}(A_i\cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$.

Remarque 3 : L'indépendance mutuelle implique l'indépendance deux à deux mais la réciproque n'est pas vraie.

Exemple 2 : On lance deux dés non truqués, un bleu et un rouge. On définit les événements suivants
A : "le dé rouge amène un numéro pair", B : "le dé bleu amène un numéro pair", C : "la somme des numéros obtenus est paire".
Montrer que (A, B, C) sont deux à deux indépendants mais pas mutuellement indépendants.

Remarque 4 : Pour calculer la probabilité d'une intersection d'événements qui ne sont pas mutuellement indépendants, on utilisera la formule des probabilités composées.

II. VARIABLES ALÉATOIRES DISCRÈTES

1. Définition

Soit $(\Omega, \mathscr{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.

Définition 6 : *Variable aléatoire discrète*
Une variable aléatoire réelle X est une application définie sur Ω et à valeurs dans $\mathbb{R}$.
On dit que X est une **variable aléatoire réelle finie** si $X(\Omega)$ est une partie finie de $\mathbb{R}$.

Remarque 5 :

- Une variable aléatoire (VA) est **une application** $X : \Omega \to E$ où E est un ensemble quelconque.
  On désigne habituellement les variables aléatoires par des lettres majuscules X, Y, Z.
  Dans ce cours, on travaille avec des VA réelles donc $E = \mathbb{R}$. Les valeurs prises par l'application X sont donc réelles.
- **Notation des événements** : Soit X une variable aléatoire réelle finie sur Ω.
  - Pour tout $x \in X(\Omega)$, l'ensemble $X^{-1}(\{x\})$ est un événement (car c'est une partie de Ω), qui sera noté $(X = x)$.
    On a donc $(X = x) = X^{-1}(\{x\}) = \{\omega \in \Omega \ / \ X(\omega) = x\}$.
  - De même, pour $U \subset \mathbb{R}$, on note $(X \in U)$ l'événement $X^{-1}(U) = \{\omega \in \Omega \ / \ X(\omega) \in U\}$.
  - Si on décrit les valeurs de $X(\Omega)$: $X(\Omega) = \{x_k,\ k \in [\![1, n]\!]\}$, on peut alors écrire $(X \in U) = \displaystyle\bigcup_{x_k \in U} (X = x_k)$.
  - Pour une variable aléatoire réelle X et $a \in \mathbb{R}$, on note $(X \leqslant a)$ l'événement $(X \in]-\infty, a])$.
    On a donc $(X \leqslant a) = \{\omega \in \Omega \ / \ X(\omega) \leqslant a\}$.
    De même, on définit $(X < a)$, $(X \geqslant a)$, $(X > a)$.

Exemple 3
On lance un dé non truqué n fois (avec $n \in \mathbb{N}^*$ fixé) et on note le nombre de 6 obtenus, ce qui définit une variable aléatoire notée X. Ici, $X(\Omega) = [\![0, n]\!]$, dans ce cas, $X(\Omega)$ est fini.
La situation n'est pas la même si l'énoncé ne précise pas le nombre de lancer. Par exemple, si on lance jusqu'à obtenir un 6, et X désigne le rang du 6 obtenu. On prendra (plus tard) $X(\Omega) = \mathbb{N}^*$.

Exemple 4
On lance un dé non truqué quatre fois et on note X la variable aléatoire égale au nombre de 6 obtenus. Décrire $X(\Omega)$ l'ensemble des valeurs possibles prises par X, puis l'événement $(X = 2)$.

2. Loi de probabillité

Soit X une variable aléatoire réelle finie sur $(\Omega, \mathscr{P}(\Omega), \mathbb{P})$. On note $X(\Omega) = \{x_k,\ k \in [\![1, n]\!]\}$ avec $n \in \mathbb{N}^*$.

Définition 7 :
- La famille $((X = x_k))_{k \in [\![1,n]\!]}$ est un système complet d'événements associés à X.
- Donner la loi de probabilité de X, c'est donner $X(\Omega)$ et $\mathbb{P}(X = x)$ pour tout $x \in X(\Omega)$.

Remarque 6 :

- L'application $\mathbb{P}_X : \begin{array}{ccc} \mathscr{P}(X(\Omega)) & \to & [0,1] \\ A & \mapsto & \mathbb{P}(X \in A) \end{array}$ définit une probabilité sur $(X(\Omega), \mathscr{P}(X(\Omega)))$.
  On dit que $\mathbb{P}_X$ est la loi de X ou que X suit la loi $\mathbb{P}_X$.

  On remarquera que $\forall A \in \mathscr{P}(X(\Omega))$, $\mathbb{P}_X(A) = \displaystyle\sum_{x \in A} \mathbb{P}(X = x) = \sum_{\omega \in \Omega \ / \ X(\omega) = x} \mathbb{P}(\omega)$.
- Puisque $((X = x_k))_{k \in [\![1,n]\!]}$ est un système complet d'événements, on vérifiera toujours que $\displaystyle\sum_{k=1}^{n} \mathbb{P}(X = x_k) = 1$.

Reprenons nos notations : $X(\Omega) = \{x_k,\ k \in [\![1, n]\!]\}$ et $p_k = \mathbb{P}(X = x_k)$ pour $k \in [\![1, n]\!]$.
On peut alors décrire la probabilité de tout événement $X \in U$, à l'aide des p_k : $\mathbb{P}(X \in U) = \displaystyle\sum_{k \in [\![1,n]\!] \ / \ x_k \in U} p_k$.

3. Les lois usuelles

a. *Loi uniforme*

Définition 8 : Soit X une variable aléatoire réelle finie (i.e. $X(\Omega)$ est une partie finie de $\mathbb{R}$).
On dit que la variable aléatoire X suit **la loi uniforme sur** $X(\Omega)$ (on note $X \hookrightarrow \mathcal{U}(X(\Omega))$) si

$$\mathrm{Card}(X(\Omega)) = N \text{ et pour tout } x \in X(\Omega),\ \mathbb{P}(X = x) = \frac{1}{N}.$$

Situations de référence

➢ On lance un dé équilibré à 6 faces, X est la VA égale au résultat du dé. X suit la loi uniforme sur $[\![1,6]\!]$.
➢ On tire une boule dans une urne contenant 10 boules numérotées de 1 à 10. X, VA égale au numéro de la boule obtenue, suit une loi uniforme sur $[\![1,10]\!]$.

b. *Loi de Bernoulli*

Définition 9 : Soit $p \in]0,1[$, on note $q = 1 - p$.
On dit que la variable aléatoire X suit **la loi de Bernoulli de paramètre** p (on note $X \hookrightarrow \mathcal{B}(p)$) si

$$X(\Omega) = \{0,1\} \text{ et } \mathbb{P}(X = 1) = p\ ,\ \mathbb{P}(X = 0) = 1 - p = q\ .$$

Situations de référence

➢ On réalise une expérience avec deux issues possibles, ou pour laquelle un événement est considéré comme un succès.
➢ Lancer de pièces : deux issues possibles Pile ou Face, $X = 0$ si on obtient $\mathbb{P}$, $X = 1$ si on obtient F. Si la pièce est équilibrée, on a $X \hookrightarrow \mathcal{B}(\frac{1}{2})$.
➢ Lancer de dé : on désigne par succès le fait d'avoir un 6. On peut alors définir une VA X égale à 1 en cas de succès et à 0 sinon. Alors $X \hookrightarrow \mathcal{B}(\frac{1}{6})$.
Dans la même expérience, on considère comme succès le fait d'avoir un résultat supérieur strictement à 4, Y, la VA associée à ce succès suit une loi de Bernoulli de paramètre $p = \dfrac{2}{3}$.

c. *Loi binomiale*

Définition 10 : Soit $n \in \mathbb{N}^*$ et $p \in [0,1]$, on note $q = 1 - p$.
On dit que la variable aléatoire X suit **la loi binomiale de paramètres** n **et** p (on note $X \hookrightarrow \mathcal{B}(n,p)$) si

$$X(\Omega) = [\![0,n]\!] \text{ et pour tout } k \in [\![0,n]\!],\ \mathbb{P}(X = k) = \binom{n}{k} p^k q^{n-k}\ .$$

Situations de référence

➢ La situation générale à reconnaître est la suivante : on réalise n épreuves de Bernoulli **indépendantes**, X est la VA qui compte le nombre de succès sur les n épreuves.
➢ On effectue n lancers successifs d'une pièce équilibrée, la VA X compte le nombre de "Face" obtenu au cours des n lancers. X suit une loi binomiale de paramètres n et $p = \dfrac{1}{2}$.
➢ Pour n lancers successifs d'un dé équilibré, X compte le nombe de 6 obtenus : $X \hookrightarrow \mathcal{B}(n, \dfrac{1}{6})$.

Dans l'exemple 3, X compte le nombre de succès dans une répétition de n épreuves de Bernoulli indépendantes. X suit donc une loi binomiale de paramètres n et $\dfrac{1}{6}$ (probabilité d'obtenir un 6, le succès).
On conclut que, pour tout $k \in [\![0,n]\!]$, $\mathbb{P}(X = k) = \binom{n}{k} \dfrac{1}{6^k} (\dfrac{5}{6})^{n-k}$.

4. Fonction de répartition

On considère ici $(\Omega, \mathbb{P})$ et X, une variable aléatoire réelle finie sur Ω.
On note $X(\Omega) = \{x_k,\ k \in [\![1,n]\!]\}$, et on rapelle que c'est une partie finie de $\mathbb{R}$.

Soit $x \in \mathbb{R}$. On note $(X \leqslant x)$ l'événement $X \in]-\infty, x]$, (c'est aussi $\bigcup\limits_{\substack{k \in [\![1,n]\!] \\ x_k \leqslant x}} (X = x_k)$, la réunion disjointe des événements $(X = x_k)$ pour $x_k \leqslant x$).

Définition 11 : La fonction de répartition de X est la fonction $F_X : \mathbb{R} \to \mathbb{R}$ définie par

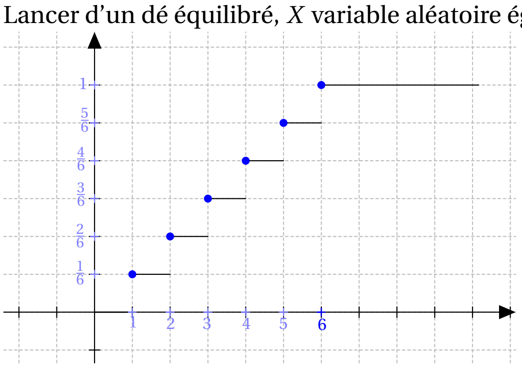
$$\forall x \in \mathbb{R},\ F_X(x) = \mathbb{P}(X \leqslant x)\ .$$

Remarque 7 :

- F est une fonction en escalier, à valeurs dans $[0,1]$, ne prenant qu'un nombre fini de valeurs distinctes.
- On a, pour $x \in \mathbb{R}$, $F_X(x) = \sum\limits_{\substack{k \in [\![1,n]\!] \\ x_k \leqslant x}} \mathbb{P}(X = x_k)$.

Exemple 5

Lancer d'un dé équilibré, X variable aléatoire égale au résultat obtenu.



Proposition 3

- F_X est croissante et à valeurs dans $[0,1]$.
- $\lim\limits_{x \to -\infty} F_X(x) = 0$ et $\lim\limits_{x \to +\infty} F_X(x) = 1$.

Remarque 8 :

a. Soit $x \in \mathbb{R}$, on a $\mathbb{P}(X > x) = 1 - \mathbb{P}(X \leqslant x)$ car $\overline{(X > x)} = (X \leqslant x)$. Donc $\mathbb{P}(X < x) = 1 - F_X(x)$.

b. Pour X une variable aléatoire telle que $X(\Omega) \subset \mathbb{N}$, on a $\forall n \in \mathbb{N}$, $\mathbb{P}(X = n) = F_X(n) - F_X(n-1)$.
En effet, $(X = n) \cup (X \leqslant n-1) = (X \leqslant n)$, et l'union est disjointe.
On peut donc déterminer la loi de probabilité de X à partir de sa fonction de répartition.

5. Image d'une variable aléatoire par une fonction

Définition 12 : Soit X une variable aléatoire sur $(\Omega, \mathbb{P})$ et $f : \mathbb{R} \to \mathbb{R}$ définie sur $X(\Omega)$.
$f \circ X$ est une variable aléatoire sur $(\Omega, \mathbb{P})$, que l'on note $f(X)$.

➢ Pour déterminer la loi de probabilité de $Y = f(X)$, on commence par identifier $Y(\Omega) = f(X(\Omega))$ (image directe de $X(\Omega)$ par f).

➢ Pour $y \in Y(\Omega)$, on a $(Y = y) = \left\{\omega \in \Omega\ :\ f \circ X(\omega) = y\right\} = \left\{\omega \in \Omega\ :\ X(\omega) \in f^{-1}(\{y\})\right\} = \bigcup\limits_{x \in f^{-1}(\{y\})} \{\omega \in \Omega\ :\ X(\omega) = x\}$.

Donc $(Y = y) = \bigcup\limits_{x \in f^{-1}(\{y\})} (X = x)$, et comme l'union est disjointe : $\mathbb{P}(Y = y) = \sum\limits_{x \in f^{-1}(\{y\})} \mathbb{P}(X = x)$.

III. ESPÉRANCE ET VARIANCE

1. Espérance

Définition 13 : Soit X une variable aléatoire réelle finie à valeurs dans $X(\Omega) = \{x_k,\ k \in [\![1,n]\!]\}$.
L'espérance de X est définie par $\mathbb{E}(X) = \sum\limits_{x \in X(\Omega)} x\, \mathbb{P}(X = x) = \sum\limits_{k=1}^{n} x_k\, P(X = x_k)$.

Remarque 9 : On vérifiera que $\mathbb{E}(X) = \sum_{\omega\in\Omega} X(\omega)\mathbb{P}(\omega)$.

Exemple 6 ♥

➢ <u>Loi uniforme</u> : Soit $N \in \mathbb{N}^*$, si $X \hookrightarrow \mathscr{U}([\![1,N]\!])$, alors $\mathbb{E}(X) = \dfrac{N+1}{2}$.
➢ <u>Loi de Bernoulli</u> : Soit $p \in]0,1[$, si $X \hookrightarrow \mathscr{B}(p)$, alors $\mathbb{E}(X) = p$.
➢ <u>Loi binomiale</u> : Soit $n \in \mathbb{N}^*$ et $p \in [0,1]$, si $X \hookrightarrow \mathscr{B}(n,p)$, alors $\mathbb{E}(X) = np$.

<u>Proposition</u> 4

a. *Linéarité*
Soient X et Y deux variables aléatoires finies sur Ω. Soient a, b des réels. On a alors $\mathbb{E}(aX+bY) = a\mathbb{E}(X)+b\mathbb{E}(Y)$.

b. *Positivité*
Soit X un variable aléatoire finie sur Ω ne prenant que des valeurs réelles positives (i.e. $\forall\omega\in\Omega,\ X(\omega)\geqslant 0$). On a alors $\mathbb{E}(X)\geqslant 0$.

c. *Croissance*
Soient X et Y deux variables aléatoires finies sur Ω.
Si $\forall\omega\in\Omega, X(\omega)\leqslant Y(\omega)$ alors $\mathbb{E}(X)\leqslant\mathbb{E}(Y)$.

Remarque 10 : Un cas particulier de la linéarité est le suivant :
pour X une variable aléatoire finie, pour tout $(a,b)\in\mathbb{R}^2$, $\mathbb{E}(aX+b) = a\mathbb{E}(X)+b$.

<u>Proposition</u> 5 *Théorème du transfert*

Soit X une variable aléatoire réelle finie, soit f une application à valeurs réelles définie sur $X(\Omega)$.
Alors $\mathbb{E}(f(X)) = \sum_{x\in X(\Omega)} f(x)\mathbb{P}(X=x)$.

Exercice 1 ✍

➢ Soit X la VA égale au résultat du lancer d'un dé équilibré. Calculer $\mathbb{E}(\dfrac{1}{X})$.
➢ Soit $X \hookrightarrow \mathscr{B}(n,p)$, avec $p\in]0,1[$. Déterminer $\mathbb{E}(X(X-1))$. En déduire $\mathbb{E}(X^2)$

2. <u>Variance</u>

Définition 14 : Soit X une variable aléatoire réelle finie sur Ω.
On appelle **variance** de X le réel $\mathbb{V}(X) = \mathbb{E}((X-\mathbb{E}(X))^2)$.
On appelle **écart-type** de X le réel $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

Remarque 11 : On sait aussi que $(X-\mathbb{E}(X))^2$ est une variable aléatoire à valeurs positives, donc son espérance est positive. Cela justifie la définition de $\sigma(X)$.

<u>Proposition</u> 6 *Formule de Koenig*

Soit X une variable aléatoire réelle finie.
On a $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$.

démo : On calcule en utilisant la linéarité de l'espérance
$\mathbb{E}((X-\mathbb{E}(X))^2) = \mathbb{E}(X^2 - 2\mathbb{E}(X)X + (\mathbb{E}(X))^2) = \mathbb{E}(X^2) - 2\mathbb{E}(X)\,\mathbb{E}(X) + (\mathbb{E}(X))^2 = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$.

Exemple 7 ♥

➢ <u>Loi uniforme</u> : Soit $N\in\mathbb{N}^*$, si $X \hookrightarrow \mathscr{U}([\![1,N]\!])$ alors $\mathbb{V}(X) = \dfrac{N^2-1}{12}$.
➢ <u>Loi de Bernoulli</u> : Soit $p\in]0,1[$, si $X \hookrightarrow \mathscr{B}(p)$, alors $\mathbb{V}(X) = pq$.
➢ <u>Loi binomiale</u> : Soit $n\in\mathbb{N}^*$ et $p\in[0,1]$, $q = 1-p$, si $X \hookrightarrow \mathscr{B}(n,p)$ alors $\mathbb{V}(X) = npq$.

dem pour la loi binomiale.
➢ On écrit $\mathbb{V}(X) = \mathbb{E}(X(X-1)) + \mathbb{E}(X) - (\mathbb{E}(X))^2$ (formule déduite de la formule de Koenig car $\mathbb{E}(X^2) = \mathbb{E}(X(X-1)) + \mathbb{E}(X)$).

➢ On connaît $\mathbb{E}(X) = np$. Donc $\mathbb{E}(X) - (\mathbb{E}(X))^2 = np(1-np)$.
➢ On calcule $\mathbb{E}(X(X-1))$, avec le théorème du transfert (on note $q = 1-p$) :

$$\begin{aligned}
\mathbb{E}(X(X-1)) &= \sum_{k=0}^{n} k(k-1)\mathbb{P}(X=k) = \sum_{k=0}^{n} k(k-1)\binom{n}{k} p^k q^{n-k},\\
&= \sum_{k=2}^{n} k(k-1)\binom{n}{k} p^k q^{n-k}, \text{ car les termes sont nuls pour } k=0 \text{ et } k=1,\\
&= \sum_{k=2}^{n} n(k-1)\binom{n-1}{k-1} p^k q^{n-k}, \text{ grâce à LA formule, voir annexe,}\\
&= \sum_{k=2}^{n} n(n-1)\binom{n-2}{k-2} p^k q^{n-k}, \text{ enocre grâce à LA formule,}\\
&= n(n-1)\sum_{l=0}^{n-2} \binom{n-2}{l} p^{l+2} q^{n-l-2}, \text{ par changement d'indice } l = k-2,\\
&= n(n-1)p^2, \text{ avec la formule du binôme de Newton et } (p+q)^{n-2} = 1
\end{aligned}$$

➢ On conclut pour la variance : $\mathbb{V}(X) = n(n-1)p^2 + np(1-np) = np((n-1)p + 1 - np) = np(1-p) = npq$.

Proposition 7

> Soit X une variable aléatoire finie.
> Pour tout $(a,b) \in \mathbb{R}^2$, on a $\mathbb{V}(aX+b) = a^2\mathbb{V}(X)$.

Proposition 8

> Soit X une variable aléatoire réelle finie sur Ω.
> - Si $\mathbb{E}(X^2) = 0$ alors $\mathbb{P}(X=0) = 1$.
> - Si $\mathbb{V}(X) = 0$ alors $\mathbb{P}(X = \mathbb{E}(X)) = 1$.

3. Inégalité de Markov et de Bienaymé-Tchebychev

Proposition 9 *Inégalité de Markov*

> Soit X une variable aléatoire réelle finie à valeurs positives. Soit $a > 0$, on a alors $\mathbb{P}(X \geqslant a) \leqslant \dfrac{\mathbb{E}(X)}{a}$.

Remarque 12 :

- L'inégalité de Markov a pour intérêt de donner une borne à une probabilité lorsque l'on connaît l'espérance. Cette borne n'est cependant en général pas optimale.
- Si X n'est pas positive, on peut utiliser $|X|$: $\mathbb{P}(|X| \geqslant a) \leqslant \dfrac{\mathbb{E}(|X|)}{a}$.
- On a $\mathbb{P}(X^2 \geqslant a^2) \leqslant \dfrac{\mathbb{E}(X^2)}{a^2}$ ou encore $\mathbb{P}(|X| \geqslant a) \leqslant \dfrac{\mathbb{E}(X^2)}{a^2}$.

Proposition 10 *Inégalité de Bienaymé-Tchebychev*

> Soit X une variable aléatoire réelle finie. On a alors, en notant $\mu = \mathbb{E}(X)$ et $\sigma = \sqrt{\mathbb{V}(X)}$.
>
> $$\forall \varepsilon > 0,\ \mathbb{P}(|X-\mu| \geqslant \varepsilon) \leqslant \frac{\sigma^2}{\varepsilon^2}\ .$$

démo : On applique l'inégalité de Markov à $Y = (X-\mu)^2$.
Par définition de $\mathbb{V}(X)$, $\mathbb{E}(Y) = \mathbb{V}(X) = \sigma^2$. On conclut en utilisant le fait que $\mathbb{P}(Y \geqslant \varepsilon^2) = \mathbb{P}(|X-\mu| \geqslant \varepsilon)$.

Remarque 13 : L'inégalité de Bienaymé-Tchebychev a pour intérêt de donner une majoration de la probabilité de l'événement $(|X - \mathbb{E}(X)| \geqslant \varepsilon)$, qui donne la probabilité de l'écart de X à son espérance, à ε près. Pour $\varepsilon = 1$, on voit que la variance donne une majoration de cette probabilité. Pour ε devenant grand, cette probabilité tend vers 0.

IV. COUPLES DE VARIABLES ALÉATOIRES

1. Lois conjointes, marginales et conditionnelles

On considère deux variables aléatoires réelles finies X et Y sur un même espace probabilisé $(\Omega, \mathscr{P}(\Omega), \mathbb{P})$.

Définition 15 : On parle du couple (X, Y) de variables aléatoires réelles finies sur Ω.
➢ La loi de probabilité de ce couple est appelé **loi conjointe** de (X, Y). C'est la donnée de

- l'ensemble $X(\Omega) \times Y(\Omega)$ des valeurs prises par le couple (X, Y),
- $\forall (x, y) \in X(\Omega) \times Y(\Omega)$, $\mathbb{P}(X = x, Y = y)$, qui désigne la probabilité de l'événement $(X = x) \cap (Y = y)$.

➢ Les lois de X et de Y sont appelées **lois marginales** du couple (X, Y).

Dans le cas d'un couple de variables aléatoires réelles finies, on donne parfois la loi conjointe de (X, Y) sous forme d'un tableau. Les lois marginales s'obtiennent en marge du tableau (pour des VA finies). En effet, $(Y = y)_{y \in Y(\Omega)}$ est un système complet d'événements donc

$\boxed{\text{pour tout } x \in X(\Omega), \mathbb{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbb{P}(X = x, Y = y)}$.

Exemple 8

On tire deux boules dans une urne contenant 4 boules blanches et 6 boules rouges.
On travaille avec remise, et on note X la VA égale à 1 si la première boule tirée est blanche et 0 sinon, Y la VA égale à 1 si la seconde boule est blanche et 0 sinon.

Voici la loi conjointe du couple (X, Y) :

X/Y	0	1	
0	$\frac{9}{5^2}$	$\frac{6}{5^2}$	
1	$\frac{6}{5^2}$	$\frac{4}{5^2}$	

La loi marginale de X est $\mathbb{P}(X = 0) = \frac{3}{5}$ et $\mathbb{P}(X = 1) = \frac{2}{5}$. La loi marginale de Y est la même.

On reprend l'expérience, mais **sans remise** : établir la loi conjointe du couple, et donner les lois marginales.

Exercice 2

On lance deux dés équilibrés, on note U le maximum des résultats obtenus et V le minimum de ces résultats.
Établir la loi conjointe du couple (U, V) puis les lois marginales.

Définition 16 : Soit (X, Y) un couple de variables aléatoires réelles finies.
Soit $x \in X(\Omega)$ tel que $\mathbb{P}(X = x) > 0$, on appelle **loi conditionnelle de Y sachant** $(X = x)$, la loi de probabilité $\mathbb{P}_{(X=x)}$ sur $Y(\Omega)$. Celle-ci est définie par $\forall y \in Y(\Omega)$, $\mathbb{P}_{(X=x)}(Y = y) = \dfrac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(X = x)}$.

On définit de même, pour $y \in Y(\Omega)$ tel que $\mathbb{P}(Y = y) > 0$, la loi conditionnelle de X sachant $(Y = y)$.
On parlera aussi de la loi de X conditionnée à $(Y = y)$.

Remarque 14 : On a donc pour $(x, y) \in X(\Omega) \times Y(\Omega)$ tels que $\mathbb{P}(X = x) > 0$ et $\mathbb{P}(Y = y) > 0$,

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

2. Variables aléatoires indépendantes

Définition 17 : Soit (X, Y) un couple de variables aléatoires réelles finies sur Ω.
On dit que (X, Y) est un couple de variables aléatoires indépendantes si

$$\forall (x, y) \in X(\Omega) \times Y(\Omega),\ \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y) .$$

Proposition 11

Soit (X,Y) un couple de variables aléatoires indépendantes.
➢ Pour toute partie $A \subset X(\Omega)$ et toute partie $B \subset Y(\Omega)$, on a $\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$.
➢ Pour toutes fonctions f, définie sur $X(\Omega)$, et g, définie sur $Y(\Omega)$, les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

Définition 18 : Si $(X_1, \dots, X_n)$ est une famille finie de variables finies sur Ω.
• On dit que ces variables aléatoires sont **mutuellement indépendantes** si
$\forall (x_1, \dots, x_n) \in X_1(\Omega), \times \cdots \times X_n(\Omega)$, les événements $(X = x_i)$ sont mutuellement indépendants.
• On dit que ces variables aléatoires sont 2 **à** 2 **indépendantes** si
$\forall (i,j) \in [\![1,n]\!]^2$ tel que $i \neq j$, les variables aléatoires X_i et X_j sont indépendantes.

Autrement dit, les variables $(X_1, \dots, X_n)$ sont mutuellement indépendantes si
$\forall (x_1, \dots, x_n) \in X_1(\Omega), \times \cdots \times X_n(\Omega)$, $\mathbb{P}(\bigcap_{i\in[\![1,n]\!]}(X_i = x_i)) = \prod\limits_{i\in[\![1,n]\!]} \mathbb{P}(X_i = x_i)$.

Remarque 15 :
➢ Si $(X_1, \dots, X_n)$ sont mutuellement indépendantes alors elles sont indépendantes 2 à 2, mais la réciproque est en général fausse.

➢ On utilise la mutuelle indépendance pour montrer le résultat suivant : si $(X_1, \dots, X_n)$ sont des variables aléatoires sur un même espace probabilisé $(\Omega, \mathbb{P})$, mutuellement indépendantes et suivant la même loi de Bernoulli de paramètre p, alors $X_1 + X_2 + \cdots + X_n \hookrightarrow \mathscr{B}(n,p)$.

3. Fonctions de deux variables aléatoires
On considère un couple de variables aléatoires (X,Y) réelles et discrètes sur Ω. On s'intéresse à de nouvelles variables aléatoires sur Ω formées à partir de X et Y. On a déjà rencontré $X+Y$ ou $aX+bY$ avec a, b réels. On peut travailler de même avec $X-Y$, $\min(X,Y)$ ou XY.

➢ Cas de la somme de deux variables aléatoires X, Y **indépendantes** : on note $Z = X+Y$.
➞ On sait que $\mathbb{E}(Z) = \mathbb{E}(X) + \mathbb{E}(Y)$.
➞ On peut aussi chercher la loi de Z :
soit $z \in Z(\Omega)$, $\mathbb{P}(Z = z) = \sum\limits_{(x,y)\in X(\Omega)\times Y(\Omega)\ /\ x+y=z} \mathbb{P}(X = x, Y = y) = \sum\limits_{x\in X(\Omega)} \mathbb{P}(X = x, Y = z-x)$.
On peut ensuite utiliser l'indépendance de X et Y, pour obtenir

$$\mathbb{P}(Z = z) = \sum_{x\in X(\Omega)} \mathbb{P}(X = x)\mathbb{P}(Y = z-x)\ .$$

➢ Cas du produit XY :
Proposition 12

Soient (X,Y) un couple de variables aléatoires réelles finies.
a. *Inégalité de Cauchy-Schwarz* : on a $(\mathbb{E}(XY))^2 \leqslant \mathbb{E}(X^2)\mathbb{E}(Y^2)$.
b. (*Admis*) Si X et Y sont indépendantes, alors $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

démo : ➢ **Pour le a.**
➞ Soit $\lambda \in \mathbb{R}$, $(\lambda X + Y)^2$ est une variable aléatoire réelle discrète à valeurs réelles positives.
De plus, $(\lambda X + Y)^2 = \lambda^2 X^2 + 2\lambda XY + Y^2$, donc par hypothèse tous les termes possèdent une espérance finie.
Le théorème de linéarité permet de conclure que $(\lambda X + Y)^2$ admet une espérance finie. De plus, cette variable aléatoire prend des valeurs positives ou nulles, donc $\mathbb{E}((\lambda X + Y)^2) \geqslant 0$.
Or on a aussi, toujours par linéarité : $\mathbb{E}((\lambda X + Y)^2) = \lambda^2\mathbb{E}(X^2) + 2\lambda\mathbb{E}(XY) + \mathbb{E}(Y^2)$.
Donc, $\forall \lambda \in \mathbb{R}$, $\lambda^2\mathbb{E}(X^2) + 2\lambda\mathbb{E}(XY) + \mathbb{E}(Y^2) \geqslant 0$.
➞ On envisage deux cas :
• si $\mathbb{E}(X^2) = 0$, on déduit de la positivité de la fonction $\lambda \mapsto 2\lambda\mathbb{E}(XY) + \mathbb{E}(Y^2)$ que $\mathbb{E}(XY) = 0$. On a ainsi l'inégalité attendue.
• si $\mathbb{E}(X^2) \neq 0$, on a une fonction polynomiale de degré 2 qui est toujours positive. Elle ne s'annule donc pas, ce qui implique que le discriminant du polynôme associé est positif ou nul. En écrivant $\Delta \geqslant 0$, on obtient l'inégalité demandée.

4. Covariance et coefficient de corrélation

Définition 19 : Soit (X,Y) un couple de variables aléatoires réelles finies sur un espace probabilisé $(\Omega,\mathscr{P}(\Omega),\mathbb{P})$. On appelle **covariance de** X **et** Y le réel $\mathrm{Cov}(X,Y)$ défini par

$$\mathrm{Cov}(X,Y)=\mathbb{E}((X-\mathbb{E}(X))(Y-\mathbb{E}(Y)))=\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)\,.$$

Proposition 13

Soient X, Y, Z des variable aléatoire réelles finies.

- $\mathrm{Cov}(X,X)=\mathbb{V}(X)$
- $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$
- pour tous $a,\ b\in\mathbb{R}$, $\mathrm{Cov}(aX+bY,Z)=a\mathrm{Cov}(X,Z)+b\mathrm{Cov}(Y,Z)$
- $\mathbb{V}(X+Y)=\mathbb{V}(X)+\mathbb{V}(Y)+2\mathrm{Cov}(X,Y)$

démo

Remarque 16 :

- On a vu que si X et Y sont indépendantes, alors $\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)=0$, donc $\mathrm{Cov}(X,Y)=0$. Attention, la réciproque est fausse.
- On peut donner une généralisation de la propriété sur $X+Y$:

$$\forall(a,b)\in\mathbb{R}^2,\ \mathbb{V}(aX+bY)=a^2\mathbb{V}(X)+b^2\mathbb{V}(Y)+2ab\mathrm{Cov}(X,Y)\,.$$

- On peut aussi généraliser avec plusieurs variables aléatoires $X_1,\ldots,X_n$ dont le carré admet une espérance finie :

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

Définition 20 : *Coefficient de corrélation linéaire*
Soit (X,Y) un couple de variables aléatoires finies, de variances non nulles. Le coefficient de corrélation linéaire de X et Y est le réel

$$\rho(X,Y)=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}}\,.$$

Remarque 17 : Si X et Y sont des variables aléatoires indépendantes alors $\rho(X,Y)=0$.

Proposition 14

Sous les mêmes conditions que dans la définition précédente, on a

- $-1\leqslant\rho(X,Y)\leqslant 1$.
- $|\rho(X,Y)|=1$ si et seulement il existe $a,\ b\in\mathbb{R}$ tels que $Y=aX+b$ est presque sûr.

démo


Jacob Bernoulli (1655-1705)


Andrei Markov (1856-1922)


Andrey Kolmogorov (1903-1987)

Pour en savoir plus :
http://www.bibmath.net/bios/index.php ou
https://www-history.mcs.st-and.ac.uk/BiogIndex.html

Annexe : dénombrement

1. Listes

Soit E un ensemble fini de cardinal $n \in \mathbb{N}^*$.

Définition 21 : On appelle $p-$liste d'éléments de E toute liste $(x_1, x_2, \dots, x_p)$ de p éléments de E.

Proposition 15

Il y a n^p $p-$listes d'éléments de E.

Situation classique : Tirages avec remise dans une urne.

Une urne contient n boules numérotées de 1 à n. On tire successivement avec remise p boules de l'urne et on note $(x_1, \dots, x_p)$ la suite de numéros obtenus. Le nombre de résultats possibles est n^p.

2. Arrangements

Soit E un ensemble fini de cardinal $n \in \mathbb{N}^*$.

Définition 22 : On appelle arrangement de p éléments de E toute $p-$liste $(x_1, x_2, \dots, x_p)$ d'éléments distincts de E.

Proposition 16

Soit $p \in [\![0, n]\!]$. Le nombre d'arrangements de p éléments d'un ensemble E de cardinal n est $A_n^p = n(n-1)\dots(n-p+1)$

Remarque 18 :

- Par convention (et d'après l'expression donnée), $A_n^p = 0$ si $p > n$.
- $A_n^0 = 1$, $A_n^1 = n$, $A_n^2 = n(n-1)$.
- $A_n^n = n!$, et c'est le nombre de permutations possibles d'un ensemble à n éléments.
- $A_n^p = \dfrac{n!}{(n-p)!}$.

Situation classique : Tirages sans remise dans une urne.

Une urne contient n boules numérotées de 1 à n. On tire successivement sans remise p boules de l'urne et on note $(x_1, \dots, x_p)$ la suite de numéros obtenus. Le nombre de résultats possibles est A_n^p.
L'ordre est pris en compte dans la suite des numéros notés.

3. Combinaisons

Soit E un ensemble fini de cardinal $n \in \mathbb{N}^*$.

Définition 23 : Soit $p \in [\![0, n]\!]$. On appelle combinaison de p éléments de E toute partie de E à p éléments.
On note $\begin{pmatrix} n \\ p \end{pmatrix}$ ("combinaison de p parmi n") le nombre de ces parties.

Proposition 17

Pour $n \in \mathbb{N}$ et $p \in [\![0, n]\!]$, $\begin{pmatrix} n \\ p \end{pmatrix} = \dfrac{n!}{p!(n-p)!}$

Remarque 19 : Par convention $\begin{pmatrix} n \\ p \end{pmatrix} = 0$ si $p > n$ ou $p < 0$.

Situation classique : Tirages simultanés dans une urne.

Une urne contient n boules numérotées de 1 à n. On tire simultanément p boules de l'urne et on note $\{x_1, \dots, x_p\}$ l'ensemble des boules obtenues. Le nombre de résultats possibles est $\begin{pmatrix} n \\ p \end{pmatrix}$.
L'ordre n'est pas pris en compte cette situation.

Remarque 20 : Les lancers de dés successifs se dénombrent comme les tirages avec remise, ainsi que les lancers de plusieurs dés.

4. Formules indispensables

- **Symétrie** Pour tout $(n,p) \in \mathbb{N}^2$, avec $0 \leqslant p \leqslant n$, $\begin{pmatrix} n \\ p \end{pmatrix} = \begin{pmatrix} n \\ n-p \end{pmatrix}$.
- **Formule du triangle de Pascal** Pour tout $(n,p) \in \mathbb{N}^2$, $\begin{pmatrix} n \\ p \end{pmatrix} + \begin{pmatrix} n \\ p+1 \end{pmatrix} = \begin{pmatrix} n+1 \\ p+1 \end{pmatrix}$.
- **LA formule** Pour tout $(n,p) \in (\mathbb{N}^*)^2$, $p\begin{pmatrix} n \\ p \end{pmatrix} = n\begin{pmatrix} n-1 \\ p-1 \end{pmatrix}$.
- **Formule de Vandermonde** Pour tout $(a,b) \in \mathbb{N}^3$ et $p \in [\![0, a+b]\!]$, on a $\begin{pmatrix} a+b \\ p \end{pmatrix} = \sum_{k=0}^{p} \begin{pmatrix} a \\ k \end{pmatrix}\begin{pmatrix} b \\ p-k \end{pmatrix}$.