Chapitre 3. Probabilités discrètes

Table des matières

1 Rappels en théorie des ensembles et dénombrements

1.1 Ensemble des parties d'un ensemble

Soit E un ensemble. L'ensemble de ses parties est noté $\mathcal{P}(E)$.

1.1.1 Opérations sur $\mathcal{P}(E)$

On définit les deux lois de composition interne **réunion** et **intersection** par :

$$A \cup B = \{x \in E \,/\, x \in A \quad \textbf{ou} \quad x \in B\} \qquad et \qquad A \cap B = \{x \in E \,/\, x \in A \quad \textbf{et} \quad x \in B\}$$

Ces deux lois sont commutatives, associatives, et distributives l'une par rapport à l'autre :

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

1.1.2 Complémentaire

A tout élément A de $\mathcal{P}(E)$ on peut associer son complémentaire $\bar{A}$ défini par :

$$\bar{A} = \{x \in E \,/\, x \notin A\}$$

Propriétés :

1. $\forall A \in \mathcal{P}(E) \qquad \bar{\bar{A}} = A$
2. $\forall A \in \mathcal{P}(E) \qquad A \cup \bar{A} = E \quad et \quad A \cap \bar{A} = \emptyset$
3. $\forall A \in \mathcal{P}(E), \quad \forall B \in \mathcal{P}(E) \qquad \overline{A \cup B} = \bar{A} \cap \bar{B} \qquad et \qquad \overline{A \cap B} = \bar{A} \cup \bar{B}$

Cette dernière propriété est connue sous le nom de **loi de Morgan**.

1.1.3 Partition

On appelle partition de E toute famille $\{E_i \,/\, i \in I\}$ de $\mathcal{P}(E)$ telle que :

$$\bigcup_{i \in I} E_i = E \qquad et \qquad \forall (i,j) \in I^2 \quad E_i \cap E_j = \emptyset$$

1.1.4 Produit cartésien

Soient p ensembles $E_1, E_2, \ldots, E_p$. Leur produit cartésien, noté $E_1 \times E_2 \times \ldots \times E_p$, est l'ensemble des p-uples $(x_1, x_2, \ldots, x_p)$ tels que, pour $1 \leq i \leq p \qquad x_i \in E_i$.
Exemple : Dans les chapitres précédents, nous avons utilisé l'ensemble $\mathbb{R}^n$, produit cartésien de n fois l'ensemble $\mathbb{R}$:

$$\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R}$$

1.2 Dénombrement

Dénombrer un ensemble, c'est déterminer le nombre de ses éléments, c'est-à-dire son cardinal, noté $card(E)$. Dans la suite de ce paragraphe, on utilisera uniquement des ensembles de cardinal fini.
Les paragraphes suivants concernent des méthodes de dénombrement, soit de façon générale, soit d'ensembles particuliers.

1.2.1 Cardinal d'une réunion d'éléments de $\mathcal{P}(E)$

Théorème 1 : $\boxed{card(A \cup B) = card(A) + card(B) - card(A \cap B)}$

Généralisation : formule du crible de Poincaré

Théorème 1bis :

$$card(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} card(A_i) - \sum_{1\leq i<j\leq n} card(A_i \cap A_j) + \sum_{1\leq i<j<k\leq n} card(A_i \cap A_j \cap A_k)$$
$$+ \ldots + (-1)^{n-1} card(A_1 \cap A_2 \cap \ldots \cap A_n)$$

Cas d'une partition :

$$card(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} card(A_i)$$

En effet dans ce cas, les intersections deux à deux, trois à trois, etc, sont toutes vides.

1.2.2 Cardinal d'un produit cartésien

Théorème 2 :

$$card(E_1 \times E_2 \times \ldots \times E_p) = \prod_{i=1}^{p} card(E_i)$$

1.2.3 p-listes

Soit E un ensemble de cardinal n (entier naturel).

Définition 1 : Une p-liste d'éléments de E est un élément de $E^p = E \times E \times \ldots \times E$.
C'est une **suite ordonnée** de p éléments de E.

D'après le paragraphe précédent, le nombre de p-listes de E est donc n^p.
Exemple : Le nombre de "mots" de 4 lettres écrits avec un alphabet de 26 lettres est 26^4.

1.2.4 p-listes d'éléments distincts, ou arrangements

Définition 2 : Soit $p \leq n$. Une p-liste d'éléments distincts de E, ou arrangement de p éléments de E, est une **suite ordonnée** de p éléments distincts de E.

Soit $(x_1, x_2, \ldots, x_p)$ une p-liste d'éléments distincts de E. Il y a n choix possibles pour x_1, puis, pour chacun de ces choix, $n-1$ choix possibles pour x_2, etc. Le nombre de p-listes d'éléments distincts de E est donc :

$$A_n^p = n(n-1)(n-2)\ldots(n-p+1) = \frac{n!}{(n-p)!}$$

Définition 3 : Une permutation de E est un arrangement des n éléments de E.

Conséquence : Le nombre de permutations de E est $n!$.

1.2.5 Combinaisons

Définition 4 : Soit $p \leq n$. Une combinaison de p éléments de E (combinaison de p éléments parmi n) est un sous-ensemble (ou partie) de E contenant p éléments.

La différence entre une p-liste d'éléments distincts de E et une combinaison de p éléments de E, c'est qu'il n'y a pas de notion d'ordre dans une combinaison.
Le nombre de p-listes d'éléments distincts parmi n est donc le nombre de combinaisons de p parmi n, multiplié par le nombre d'"ordres" possibles sur ces p éléments, autrement dit par le nombre de permutations sur ces p éléments. On a donc :

$$p!\binom{n}{p} = \frac{n!}{(n-p)!} \quad \text{et donc} \quad \binom{n}{p} = \frac{n!}{p!(n-p)!}$$

Les nombres $\binom{n}{p}$ sont appelés **coefficients du binôme**.

1.2.6 Nombre d'applications d'un ensemble E vers un ensemble F.

Soit E un ensemble de cardinal n et F un ensemble de cardinal p.

Le nombre d'applications de E vers F est le nombre de n-listes d'éléments de F. En effet chaque élément de E admet une image unique par chaque application. Une application f est donc caractérisée par la n-liste des n éléments $f(x_1), f(x_2), \ldots, f(x_n)$, où $E = \{x_1, x_2, \ldots, x_n\}$.
Le nombre d'applications de E vers F est donc p^n.

Le nombre d'applications injectives de E vers F est le nombre de n-listes d'éléments distincts de F. En effet une application injective est une application telle que 2 éléments distincts de E ont forcément 2 images distinctes.
Le nombre d'injections de E vers F est donc $\frac{p!}{(p-n)!}$.
On remarque que ce nombre est non nul seulement si $p \geq n$.

Si $card(E) = card(F) = n$, le nombre de bijections de E vers F est $n!$.

1.2.7 Utilisation d'une bijection pour dénombrer un ensemble

Théorème 3 : Soient E et F deux ensembles de cardinaux finis.
S'il existe une **bijection** de E vers F, $card(E) = card(F)$.

Exemple 1. Soit $E = [\![1, n]\!]$, avec $n \geq 4$, et soit $S_4(E) = \{(x, y, z, t) \,/\, x < y < z < t\}$.
A tout quadruplets (x, y, z, t) de $S_4(E)$ on associe le sous-ensemble de E à 4 éléments $\{x, y, z, t\}$. L'application ainsi définie est une bijection. Or le nombre de sous-ensembles de E à 4 éléments est $\binom{n}{4}$.
Donc $card(S_4(E)) = \binom{n}{4}$.

Exemple 2. (Cas particulier du problème des "trous de Kaplansky")
Soit n un entier tel que $n \geq 5$ et $A = \{(a, b, c) \in [\![1, n]\!] \,/\, a < b < c \text{ et } a, b, c \text{ ne sont pas consécutifs}\}$
Soit f l'application définie sur A par : $f((a, b, c)) = (a, b-1, c-2)$.
L'ensemble image de cette application est l'ensemble B des triplets (x, y, z) tels que $1 \leq x < y < z \leq n-2$.
Chaque triplet (x, y, z) de B a pour unique antécédent dans A : $(x, y+1, z+2)$.
Donc l'application f est une bijection de A vers B.
Par conséquent, par analogie avec le résultat de l'exemple 1,

$$card(A) = \binom{n-2}{3}$$

1.2.8 Utilisation d'une partition

Pour dénombrer un ensemble, on peut le décomposer en parties disjointes plus faciles à dénombrer.

Exemple 3 : dénombrement de $\mathcal{P}(E)$
Soit E un ensemble de cardinal n.
Soit E_k le sous-ensemble de $\mathcal{P}(E)$ constitué des parties de E de cardinal k.
Alors $card(E_k) = \binom{n}{k}$.
Comme les ensembles E_k sont tous disjoints, et que $\bigcup_{0 \leq k \leq n} E_k = \mathcal{P}(E)$, $\{E_k \,/\, 0 \leq k \leq n\}$ est une partition de $\mathcal{P}(E)$.
On a donc : $card(\mathcal{P}(E)) = \sum_{k=0}^{n} E_k = \sum_{k=0}^{n} \binom{n}{k} = 2^n$
(D'après la formule du binôme de Newton.)

Remarque : On peut également démontrer ce résultat par récurrence.

Exemple 4 : Soit $A = \{(X,Y) \in (\mathcal{P}(E))^2 \,/\, X \subset Y\}$.
On pose $A_k = \{(X,Y) \in (\mathcal{P}(E))^2 \,/\, X \subset Y \quad et \quad card(Y) = k\}$.
On définit ainsi une partition de A :

$$card(A) = \sum_{k=0}^{n} card(A_k) = \sum_{k=0}^{n} \binom{n}{k} 2^k = 3^n$$

1.3 Propriétés des coefficients du binôme

1.3.1 Conventions

Si $k > n$,ou $k < 0$, on convient que : $\binom{n}{k} = 0$

1.3.2 Symétrie :

$$\binom{n}{n-p} = \binom{n}{p}$$

1.3.3 Formule du triangle de Pascal :

$$\forall (n,p) \;\; / \;\; 1 \leq p \leq n \qquad \binom{n}{p} = \binom{n-1}{p-1} + \binom{n-1}{p}$$

1.3.4

$$\forall (n,k) \;\; / \;\; 1 \leq k \leq n \qquad k\binom{n}{k} = n\binom{n-1}{k-1}$$

1.3.5 Formule du binôme de Newton

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

1.3.6 Somme sur l'indice "d'en haut"

$$\sum_{k=n}^{m} \binom{k}{n} = \binom{m+1}{n+1}$$

1.3.7 Formule de Vandermonde

$$\sum_{i=0}^{n} \binom{n}{i} \binom{m}{k-i} = \binom{n+m}{k}$$

2 Généralités sur les probabilités

2.1 Définitions, vocabulaire

Définition 5 :

Soit un ensemble Ω, appelé **univers**.
On appelle **tribu**, ou σ**-algèbre**, de parties de Ω, tout sous-ensemble $\mathcal{A}$ de $\mathcal{P}(\Omega)$ tel que :
• Ω est élément de $\mathcal{A}$
• $\forall A \in \mathcal{A}, \qquad \overline{A} \in \mathcal{A}$
• Pour toute famille finie ou dénombrable (A_i) d'éléments de $\mathcal{A}$,
$\bigcup A_i \in \mathcal{A}$. ($\mathcal{A}$ est stable pour la réunion).

Le couple $(\Omega, \mathcal{A})$ est alors appelé **univers probabilisable**.
Les éléments de Ω sont des **éventualités**, les éléments de $\mathcal{A}$ sont des **événements**.

Application : Si on effectue une expérience aléatoire (c'est-à-dire dont le résultat dépend du hasard), Ω est l'ensemble des résultats possibles de cette expérience.
Par exemple si on lance un dé à 6 faces numérotées de 1 à 6, on pourra prendre $\Omega = \{1, 2, 3, 4, 5, 6\}$ et $\mathcal{A} = \mathcal{P}(\Omega)$.

Il y a deux façons de considérer les événements : soit comme un élément de $\mathcal{A}$, c'est-à-dire une partie de Ω, soit comme "quelque chose" qui peut être vrai ou faux, arriver ou ne pas arriver, autrement dit une **proposition logique**.
En reprenant l'exemple du lancer de dé, l'événement $\{2, 4, 6\}$ correspond à : "Le résultat du lancer est un nombre pair".

On dit que l'éventualité ω **réalise** l'événement A si $\omega \in A$. Cela revient à dire que le résultat ω de l'expérience aléatoire vérifie la proposition qui définit A.
Par exemple, le résultat "2" réalise l'événement "Le résultat du lancer est un nombre pair" : $2 \in \{2, 4, 6\}$.

Le tableau suivant établit la correspondance entre le vocabulaire ensembliste et le vocabulaire logique.

Ω	Evénement certain
$\emptyset$	Evénement impossible
$A \cap B = \emptyset$	A et B sont incompatibles
$\overline{A}$	Evénement contraire de A
$A \cap B$	A et B
$A \cup B$	A ou B
$A \subset B$	$A \Rightarrow B$
$A = B$	$A \Leftrightarrow B$
$\omega \in A$	ω réalise A

Définition 6 :

Un **système complet d'événements** est une famille finie ou dénombrable d'éléments de $\mathcal{A}$ incompatibles 2 à 2 et dont la réunion est Ω. C'est donc une partition de Ω, formée d'événements.

Propriété : Soit $(A_i)_{i \in I}$ un système complet d'événements de l'univers Ω. Alors tout événement B Se décompose sur ce S.C.E. de la façon suivante :

$$B = \bigcup_{i \in I} (B \cap A_i)$$

Cas particulier : Pour tout événement A, $\{A, \overline{A}\}$ est un système complet d'événements.

Définition 7 :

Une **probabilité** sur un univers probabilisable $(\Omega, \mathcal{A})$ est une application P de $\mathcal{A}$ vers $[0,1]$ vérifiant :
• $P(\Omega) = 1$
• Si A et B sont deux événements incompatibles, $P(A \cup B) = P(A) + P(B)$
On dit alors que $(\Omega, \mathcal{A}, P)$ est un univers probabilisé.

Propriétés :

1. $P(\emptyset) = 0$
2. Pour tout couple (A, B) d'événements, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
3. Généralisation : Formule du crible de Poincaré :

Théorème 4 :

$$P(\bigcup_{i=1}^{n} A_i = \sum_{i=1}^{n} P(A_i) + \ldots + (-1)^{p-1} \sum_{1 \leq i_1 < i_2 < \ldots < i_p \leq n} P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_p}) + \ldots$$
$$+(-1)^{n-1} P(A_1 \cap A_2 \cap \ldots \cap A_n)$$

4. Condition pour qu'une famille d'événements soit un système complet d'événements :

Théorème 5 :

$(A_i)_{i \in I}$ est un système complet d'événements si, et seulement si :
• Chaque éventualité réalise un des A_i et un seul, ou :
• $\displaystyle\sum_{i \in I} P(A_i) = 1$ et les (A_i) sont incompatibles 2 à 2.

2.2 Probabilité sur un univers fini

Soit $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ un univers de cardinal n.
Soit $p_1, p_2, \ldots, p_n$ n nombres réels positifs ou nuls tels que : $\displaystyle\sum_{i=1}^{n} p_i = 1$.
Alors ces nombres définissent une probabilité sur $(\Omega, \mathcal{P}(\Omega))$, de la façon suivante :

$$\forall A \in \mathcal{P}(\Omega), \qquad P(A) = \sum_{i/\omega_i \in A} p_i$$

En particulier : $P(\{\omega_i\}) = p_i$.

Le cas le plus fréquent est l'**équiprobabilité**, ou **probabilité uniforme** :

Définition 8 :

P est la probabilité uniforme sur Ω lorsque tous les p_i sont égaux entre eux,
et on a donc : $\quad P(\{\omega_i\}) = \dfrac{1}{card(\Omega)} = \dfrac{1}{n}$

Dans ce cas, $\qquad \forall A \in \mathcal{P}(\Omega), \qquad P(A) = \dfrac{card(A)}{card(\Omega)} = \dfrac{nombre\,de\,cas\,favorables}{nombre\,de\,cas\,possibles}$

Les problèmes de calcul de probabilités sont ramenés à des problèmes de dénombrements.

2.3 Probabilités conditionnelles

Définition 9 :

Soit A un événement d'un univers probabilisé $(\Omega, \mathcal{A}, P)$, tel que $P(A) \neq 0$.
La probabilité conditionnelle relativement à A est l'application P_A de $\mathcal{A}$ vers $[0,1]$ définie par :

$$\forall B \in \mathcal{A} \qquad P_A(B) = \frac{P(B \cap A)}{P(A)}$$

L'application P_A est bien une probabilité. En effet :
• $P_A(\Omega) = \dfrac{P(A \cap \Omega)}{P(A)} = \dfrac{P(A)}{P(A)} = 1$
• Si B et C sont incompatibles, $B \cap A$ et $C \cap A$ sont incompatibles et

$$P_A(B \cup C) = \frac{P((B \cup C) \cap A)}{P(A)} = \frac{P((B \cap A) \cup (C \cap A))}{P(A)} = \frac{P(B \cap A) + P(C \cap A)}{P(A)} = P_A(B) + P_A(C)$$

Propriétés :

1. **Formule des probabilités composées**

   Théorème 6 :

   Soit $(A_i)_{1 \leq i \leq n}$ une famille d'événements telle que $P(A_1 \cap A_2 \cap \ldots \cap A_n) \neq 0$. Alors :
   $P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P_{A_1}(A_2)P_{A_1 \cap A_2}(A_3) \ldots P_{A_1 \cap A_2 \cap \ldots \cap A_{n-1}}(A_n)$

   Cette formule est utile en particulier pour les problèmes de tirages au sort successifs non indépendants entre eux.

2. **Formule des probabilités totales.**

   Théorème 7 :

   Soit $(A_i)_{i \in I}$ un système complet d'événements. Pour tout événement B :
   $$P(B) = \sum_{i \in I} P(B \cap A_i) = \sum_{i \in I} P(A_i) P_{A_i}(B)$$

   Cette formule sert très fréquemment, dès que l'on cherche à calculer la probabilité d'un événement "compliqué" : on décompose alors cet événement à l'aide d'un système complet d'événements judicieusement choisi.

3. **Formule de Bayes**

   Théorème 8 :

   Soit $(A_i)_{i \in I}$ un système complet d'événements. Pour tout événement B tel que $P(B) \neq 0$:
   $$P_B(A_j) = \frac{P(A_j) P_{A_j}(B)}{\displaystyle\sum_{i \in I} P(A_i) P_{A_i}(B)}$$

2.4 Indépendance

Définition 10 : Deux événements A et B sont **indépendants** si, et seulement si, $P(A \cap B) = P(A)P(B)$

Remarques

1. L'indépendance de deux événements dépend de la probabilité P considérée.
2. Elle se généralise à une famille finie ou dénombrable d'événements : on a alors deux notions d'indépendance.
   • Les événements $(A_i)_{i\in I}$ sont mutuellement (ou globalement) indépendants si, et seulement si,
   $P\left(\bigcap_{i\in I} A_i\right) = \prod_{i\in I} P(A_i)$
   • Les événements $(A_i)_{i\in I}$ sont deux à deux indépendants si, et seulement si,
   $\forall (i,j)\in I^2,\quad i\neq j,\quad P(A_i\cap A_j) = P(A_i)P(A_j)$.
3. Si A et B sont indépendants, et si $P(A)P(B)\neq 0$, $P_A(B) = P(B)$ et $P_B(A) = P(A)$.
   Le conditionnement par un événement indépendant ne change rien à la probabilité.

3 Variables aléatoires : généralités

3.1 Définitions, exemple

Définition 11 : Soit $(\Omega, \mathcal{A}, P)$ un univers probabilisé. Une variable aléatoire réelle X est une application de Ω vers $\mathbb{R}$ telle que, pour tout intervalle I de $\mathbb{R}$, $X^{-1}(I)\in\mathcal{A}$.

On peut le comprendre comme suit : X est une quantité mesurable qui dépend du hasard (et qui est liée à l'expérience aléatoire effectué, à partir de laquelle on définit l'univers probabilisé), et la condition signifie que, pour tout intervalle I de $\mathbb{R}$, la proposition "X est un nombre réel appartenant à I" est un **événement**.

Définition 12 : L'ensemble de valeurs d'une variable aléatoire X définie sur $(\Omega, \mathcal{A}, P)$, noté $X(\Omega)$, est l'ensemble image de Ω par l'application X.
C'est donc l'ensemble des valeurs **possibles** de X à l'issue de l'expérience aléatoire.

Cas discret : Supposons que Ω est un ensemble fini : $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$.
Alors $\mathcal{A} = \mathcal{P}(\Omega)$, et l'ensemble $X(\Omega) = \{X(\omega_1), X(\omega_2), \dots, X(\omega_n)\}$ est également fini.
Comme X n'est pas forcément injective, l'ensemble de valeurs de X peut comporter moins de n éléments distincts : posons $X(\Omega) = \{x_1, x_2, \dots, x_p\}$.
Dans ce cas X est une variable aléatoire si, et seulement si,

$$\forall i\,/\,1\leq i\leq p,\qquad X^{-1}(x_i)\in\mathcal{A}$$

Autrement dit, la proposition $(X = x_i)$ est un **événement**.
Cet événement est l'ensemble des éventualités (= éléments de Ω) qui **réalisent** $X = x_i$.
On peut donc calculer sa **probabilité**. On peut raisonner de même lorsque Ω est dénombrable, et donc $X(\Omega)$ aussi.

Définition 13 : Soit X une variable aléatoire telle que $X(\Omega) = \{x_i\,/\,i\in I\}$.
La loi de probabilité de X est la suite des couples (x_i, p_i), $i\in I$, tel que
$p_i = P(X = x_i)$.

Exemple : Soit un dé équilibré dont les faces sont : AS, ROI, DAME, VALET, 10, 9.
On lance le dé ; les faces 9 et 10 rapportent 1 point, VALET : 4 points, ROI ou DAME : 6 points, AS : 10 points.

Soit X le nombre de points obtenus après un lancer.
Alors par exemple : $X^{-1}(6) = \{\text{ROI, DAME}\}$.
Autrement dit les résultats possibles qui réalisent l'événement $(X = 6)$ sont ROI et DAME.
La probabilité de cet événement est : $P(X = 6) = \frac{2}{6} = \frac{1}{3}$.
$X(\Omega) = \{1, 4, 6, 10\}$ et la loi de probabilité de X est donnée par le tableau suivant :

x_i	1	4	6	10
p_i	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{6}$

3.2 Fonction de répartition

Définition 14 : Soit X une variable aléatoire. La fonction de répartition F de la variable X est la fonction définie sur $\mathbb{R}$ par : $\forall x \in \mathbb{R} \qquad F(x) = P(X \leq x)$

Si X a pour ensemble de valeurs $X(\Omega) = \{x_i \,/\, i \in I\}$, pour tout réel x on a :

$$F(x) = \sum_{i/x_i \leq x} P(X = x_i)$$

Par exemple si $X(\Omega) = \{1, 2, \ldots, n\}$,

$$F(k) = \sum_{i=1}^{k} P(X = i)$$

Cette formule donne la fonction de répartition à partir de la loi de probabilité. Dans certains problèmes on peut être amené à trouver la loi de probabilité, connaissant la fonction de répartition, lorsque X est à valeurs entières.
En remarquant que : $(X \leq k) \Leftrightarrow (X = k) \cup (X < k) \Leftrightarrow (X = k) \cup (X \leq k - 1)$, on obtient :

Théorème 9 : Si X est une variable aléatoire à valeurs dans $\mathbb{N}$,
$\forall k \in X(\Omega) \qquad P(X = k) = F(k) - F(k - 1)$

Dans le cas général on remarque que si $x \in [x_i, x_{i+1}[$, $F(x) = F(x_i)$: la fonction F est constante sur tout intervalle $[x_i, x_{i+1}[$: la fonction de répartition est donc une **fonction en escalier**.
On peut donc résumer cette fonction de répartition au tableau de valeurs concernant $X(\Omega)$. Pour l'exemple précédent :

x_i	1	4	6	10
$F(x_i)$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{5}{6}$	1

3.3 Composée par une fonction g

Soit X une variable aléatoire d'ensemble de valeurs $X(\Omega) = \{x_i \,/\, i \in I\}$, et de loi de probabilité $\{(x_i, p_i)\}$ avec :

$$\forall i \in I \quad P(X = x_i) = p_i$$

Soit g une fonction réelle de variable réelle, définie sur un intervalle contenant $X(\Omega)$.
On pose $Y = g(X)$. Alors Y est une variable aléatoire, dont l'ensemble de valeurs est $Y(\Omega) = \{g(x_i) \,/\, i \in I\}$ et dont la loi de probabilité est donnée par :

$$\forall y \in Y(\Omega) \qquad P(Y = y) = \sum_{i/g(x_i)=y} P(X = x_i)$$

Exemple : Soit X la variable aléatoire dont la loi de probabilité est donnée par le tableau suivant :

x_i	-2	-1	0	1	2
p_i	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{1}{10}$	$\frac{4}{10}$	$\frac{1}{10}$

et la fonction g définie par : $g(x) = x^2$, d'où $Y = X^2$.
Alors $Y(\Omega) = \{0, 1, 4\}$ et par exemple :

$$P(Y = 4) = P(X = -2) + P(X = 2) = \frac{2}{10}$$

et la loi de probabilité de Y est donnée par le tableau :

y_i	0	1	4
$P(Y = y_i)$	$\frac{1}{10}$	$\frac{7}{10}$	$\frac{2}{10}$

Cas particulier : Si g est une application **injective** sur $X(\Omega)$, alors les événements $(X = x_i)$ et $(g(X) = g(x_i))$ sont équivalents, d'où :

$$\forall i \in I \qquad P(g(X) = g(x_i)) = P(X = x_i)$$

3.4 Espérance et variance d'une variable aléatoire

Définition 15 : Soit X une variable aléatoire d'ensemble de valeurs $X(\Omega) = \{x_i / i \in I\}$.
L'**espérance** de X est le nombre réel : $\quad E(X) = \sum_{i \in I} x_i P(X = x_i)$

Pour l'exemple précédent : $\quad E(X) = 1 \times \frac{1}{6} + 4 \times \frac{1}{3} + 6 \times \frac{1}{3} + 10 \times \frac{1}{6} = \frac{14}{3}$

Cas particulier : Très souvent la variable est à valeurs dans $\mathbb{N}$. On a donc l'écriture suivante, par exemple si $X(\Omega) = [\![1, n]\!]$:

$$E(X) = \sum_{k=1}^{n} kP(X = k)$$

Généralisation à un ensemble de valeurs dénombrable :
Si $X(\Omega) = \mathbb{N}$, on considère la série de terme général $nP(X = n)$. Si cette série est convergente, X admet une espérance, et :

$$E(X) = \sum_{n=0}^{+\infty} nP(X = n)$$

Si la série est divergente, la variable X n'admet pas d'espérance.

Théorème de transfert

Théorème 10 : Si X est une variable aléatoire d'ensemble de valeurs $X(\Omega) = \{x_i \,/\, i \in I\}$, et g une fonction définie sur un intervalle contenant $X(\Omega)$, alors, sous réserve de convergence :

$$E(g(X)) = \sum_{i \in I} g(x_i) P(X = x_i)$$

Par exemple dans le cas où $X(\Omega) = [\![1, n]\!]$, on a :

$$E\left(g(X)\right) = \sum_{k=1}^{n} g(k)P(X = k)$$

Prenons le cas de la fonction “carré” :

$$E\left(X^2\right) = \sum_{k=1}^{n} k^2 P(X = k)$$

Définition 16 : Soit X une variable aléatoire d'ensemble de valeurs fini ou dénombrable.
La **variance** de X est le nombre réel, s'il existe : $V(X) = E\left[(X - E(X))^2\right]$
L'**écart-type** de X est le nombre réel : $\sigma(X) = \sqrt{V(X)}$

Théorème 11 : Sous réserve d'existence de la variance, $V(X) = E(X^2) - (E(X))^2$

Cette formule est connue sous le nom de ***formule de Koenig-Huyghens***.

Propriétés :

Théorème 12 : $E(aX + b) = aE(X) + b$ $\qquad V(aX + b) = a^2V(X)$ $\qquad \sigma(aX + b) = |a|\sigma(X)$

Ces relations se démontrent à l'aide du théorème de transfert.

4 Lois de probabilité usuelles

4.1 Loi uniforme

Définition 17 : La variable X suit une loi uniforme sur $\{x_1, x_2, \dots, x_n\}$ si, et seulement si, pour tout i de 1 à n, $\quad P(X = x_i) = \dfrac{1}{n}$

Notation : $X \hookrightarrow \mathcal{U}(\{x_1, x_2, \dots, x_n\})$

Remarque : La probabilité commune est l'inverse du nombre de valeurs de $X(\Omega)$.
Le cas le plus fréquent est celui où l'expérience aléatoire est un tirage (équiprobable) d'un numéro parmi $\{1, 2, \dots, n\}$. On a alors :

Ensemble de valeurs : $X(\Omega) = \{1, 2, \dots, n\} = [\![1, n]\!]$

Loi de probabilité : $P(X = k) = \dfrac{1}{n}$

Espérance : $E(X) = \dfrac{n+1}{2}$

Variance : $V(X) = \dfrac{n^2 - 1}{12}$

4.2 Loi de Bernoulli

Définition 18 : La variable X suit une loi de Bernoulli de paramètre p si, et seulement si, $X(\Omega) = \{0,1\}$ et $P(X=1) = p$.

Expérience aléatoire : N'importe quelle expérience aléatoire ayant deux issues possibles, appelées succès et échec. La probabilité du succès est p.

Notation : $X \hookrightarrow \mathcal{B}(p)$ ou : $X \hookrightarrow \mathcal{B}(1,p)$

Ensemble de valeurs : $X(\Omega) = \{0,1\}$

Loi de probabilité : $P(X=1) = p$ et $P(X=0) = q = 1-p$

Espérance : $E(X) = p$

Variance : $V(X) = pq = p(1-p)$

Remarque : toute variable aléatoire ayant pour ensemble de valeurs $\{0,1\}$ est une variable de Bernoulli.

Carré d'une variable de Bernoulli : Si $X \hookrightarrow \mathcal{B}(p)$, alors $X^2 = X$.

Produit de deux variables de Bernoulli indépendantes : Si $X \hookrightarrow \mathcal{B}(p_1)$ et $Y \hookrightarrow \mathcal{B}(p_2)$, et si X et Y sont indépendantes, leur produit XY suit une loi de Bernoulli de paramètre p_1p_2.
dém : $XY(\Omega) = \{0,1\}$ et $P(XY=1) = P(X=1 \cap Y=1) = P(X=1)P(Y=1) = p_1p_2$.

4.3 Loi binomiale

Définition 19 : La variable X suit une loi binomiale de paramètres n et p si, et seulement si, $\forall k \in [\![0,n]\!] \qquad P(X=k) = \binom{n}{k} p^k(1-p)^{n-k}$.

Expérience aléatoire : On répète n fois une expérience aléatoire ayant deux issues possibles, appelées succès et échec. La probabilité du succès est p. X est le **nombre de succès** sur n tentatives indépendantes entre elles.

Notation : $X \hookrightarrow \mathcal{B}(n,p)$

Ensemble de valeurs : $X(\Omega) = [\![0,n]\!]$

Loi de probabilité : $P(X=k) = \binom{n}{k} p^k(1-p)^{n-k}$

Espérance : $E(X) = np$

Variance : $V(X) = npq = np(1-p)$

Propriétés :

Théorème 13 : Si $X_1, X_2, \ldots, X_n$ sont n variables de Bernoulli indépendantes de même paramètre p, si $X = \sum_{i=1}^{n} X_i$, X est une variable aléatoire de loi binomiale $\mathcal{B}(n,p)$.

Théorème 14 : Si $X_1 \hookrightarrow \mathcal{B}(n_1, p)$ et $X_2 \hookrightarrow \mathcal{B}(n_2, p)$, X_1 et X_2 **indépendantes**, alors $X_1 + X_2 \hookrightarrow \mathcal{B}(n_1 + n_2, p)$

4.4 Loi hypergéométrique

Définition 20 : La variable X suit une loi hypergéométrique de paramètres N, n et p si, et seulement si,

$$\forall k \in [\![0, n]\!] \qquad P(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, \text{ avec } M = Np.$$

Expérience aléatoire : On prélève un **échantillon** (tirage simultané, ou tirage sans remise) de n objets dans un ensemble de N, ayant une propriété A ou non, la proportion d'objets ayant la propriété A étant p. X est le nombre des objets tirés ayant la propriété A.

Notation : $X \hookrightarrow \mathcal{H}(N, n, p)$

Ensemble de valeurs : $X(\Omega) = [\![max(0, n - N(1-p)), min(n, Np)]\!]$.
Dans la pratique on pourra prendre : $X(\Omega) = [\![0, n]\!]$

Loi de probabilité : $P(X = k) = \dfrac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$

Espérance : $E(X) = np$

Variance : $V(X) = npq\dfrac{N-n}{N-1} = np(1-p)\dfrac{N-n}{N-1}$

Remarque : En étudiant ce type de loi on sera amené à utiliser la formule de Vandermonde.

4.5 Loi géométrique

Définition 21 : La variable X suit une loi géométrique de paramètre p si, et seulement si, $\forall k \in \mathbb{N}^* \qquad P(X = k) = pq^{k-1}$, avec $q = 1 - p$.

Expérience aléatoire : On réalise une suite d'essais indépendants d'une expérience à deux issues, la probabilité du succès étant p. X est le **rang du premier succès**.

Notation : $X \hookrightarrow \mathcal{G}(p)$

Ensemble de valeurs : $X(\Omega) = \mathbb{N}^*$.

Loi de probabilité : $P(X = k) = pq^{k-1}$

Espérance : $E(X) = \dfrac{1}{p}$

Variance : $V(X) = \dfrac{1-p}{p^2}$

Remarque : On rencontre souvent des variables du type : "X est le nombre d'échecs avant le premier

succès". Alors $X = Y - 1$, où $Y \hookrightarrow \mathcal{G}(p)$.
La loi de X est alors : $P(X = k) = pq^k$, $\quad E(X) = \dfrac{1-p}{p}$, $\quad V(X) = \dfrac{1-p}{p^2}$.

4.6 Loi de Poisson

Définition 22 : La variable X suit une loi de Poisson de paramètre λ si, et seulement si, $\forall k \in \mathbb{N} \qquad P(X = k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$.

Expérience aléatoire : Cette loi intervient dans les problèmes d'approximation de loi binomiale, pour des faibles valeurs du paramètre p. Elle est le plus souvent donnée en hypothèse dans un énoncé.

Notation : $X \hookrightarrow \mathcal{P}(\lambda)$

Ensemble de valeurs : $X(\Omega) = \mathbb{N}$.

Loi de probabilité : $P(X = k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$

Espérance : $E(X) = \lambda$

Variance : $V(X) = \lambda$

Propriété :

Théorème 15 : Si $X_1 \hookrightarrow \mathcal{P}(\lambda_1)$ et $X_2 \hookrightarrow \mathcal{P}(\lambda_2)$, X_1 et X_2 **indépendantes**, alors $X_1 + X_2 \hookrightarrow \mathcal{P}(\lambda_1 + \lambda_2)$

Ce théorème se démontre à l'aide de la formule des probabilités totales (voir chapitre 6).

5 Problèmes classiques en probabilités

5.1 Déterminer une loi de probabilité

Soit une variable X, définie à partir d'une expérience aléatoire.

- Soit X suit une loi usuelle, qu'il s'agit de reconnaître. Les critères peuvent être : l'ensemble de valeurs (fini ou infini), s'il s'agit d'un nombre de "succès" sur un nombre de tirages donnés, ces tirages sont-ils indépendants (loi binomiale) ou simultanés (loi hypergéométrique), etc.
- Soit la loi de X est plus compliquée : on décompose alors l'événement $(X = k)$ en réunion disjointe d'intersections d'événements élémentaires, et on applique la formule des probabilités composées ou la formule des probabilités totales.
  On peut également utiliser un dénombrement.
  Suivant l'énoncé, il arrive que la fonction de répartition soit plus facile à déterminer que la loi de probabilité. Si X est une variable à valeurs entières, on applique alors la formule :

$$P(X = k) = P(X \leq k) - P(X \leq k - 1)$$

5.2 Chaîne de Markov

Une expérience aléatoire E à p issues possibles x_i, pour i de 1 à p, est répétée de façon que, pour tout entier naturel n, les hypothèses de E_{n+1} dépendent uniquement des résultats obtenus à l'expérience précédente E_n.
On note X_n la variable aléatoire qui prend pour valeurs une des x_i, à l'issue de l'expérience numéro n.
Il s'agit de déterminer la loi de probabilité de X_n, connaissant :

- la loi de X_0 (état initial).

- les probabilités conditinnelles $P_{X_n=x_j}(X_{n+1}=x_i)$, pour tout i et tout j de 1 à p.
  Ces probabilités sont indépendantes de n. Elles expriment le passage de l'état n à l'état $n+1$.

Méthode :

- Comme les événements $(X_n = x_j)$, pour j de 1 à p, constituent un système complet d'événements, on peut appliquer la formule des probabilités totales pour obtenir, pour tout i de 1 à p, la probabilité $P(X_{n+1}=x_i)$ en fonction des $P(X_n=x_j)$.
- On pose $U_n = \begin{pmatrix} P(X_n=x_1) \\ P(X_n=x_2) \\ \vdots \\ P(X_n=x_p) \end{pmatrix}$, et on déduit du calcul précédent une matrice A telle que :

  $$\forall n \in \mathbb{N} \qquad U_{n+1} = AU_n$$

  Par un raisonnement par récurrence, on en déduit :

  $$\forall n \in \mathbb{N} \qquad U_n = A^n U_0$$

  Il suffit alors de déterminer la puissance n-ième de la matrice A, par exemple en utilisant la diagonalisation de cette matrice (voir chapitre 5).