1 Probabilités

Objectifs :

- Variable aléatoire discrète : loi de probabilité, espérance, variance et l'écart-type
- Á l'aide de simulations, faire le lien avec la moyenne et la variance d'une série de données
- Connaitre les propriétés de l'espérance et de la variance
- Être capable de modéliser la répétition d'expériences identiques et indépendantes à deux ou trois issues à l'aide d'un arbre pondéré ; savoir exploiter cette modélisation pour calculer des probabilités
- Reconnaitre une épreuve de Bernoulli, un schéma de Bernoulli ; déterminer la loi de probabilités associée
- Calculer l'espérance et la variance d'une variable aléatoire suivant une loi binomiale
- Représenter graphiquement une loi binomiale $\mathscr{B}(n;p)$
- Connaitre la définition et les propriétés des coefficients binomiaux, être capable d'utiliser le triangle de Pascal
- Être capable d'utiliser la loi binomiale pour prendre une décision à partir d'une fréquence d'échantillonnage

Aperçu historique : Le mot probabilité apparaît tout d'abord dans les traductions d'Aristote, où il signifie qu'une idée est communément admise par tous. Pendant le Moyen-âge et la Renaissance, des erreurs de traduction amèneront à utiliser ce terme pour désigner la vraisemblance d'une idée.
L'étude des probabilités telles que nous les connaissons est apparue avec la notion de risque au XII^e^ siècle, pour l'établissement de contrats commerciaux, puis au XVI^e^ siècle avec la généralisation des contrats d'assurance maritime.
Mais le véritable début de la théorie des probabilités date des travaux de Pierre de Fermat, Blaise Pascal et Christian Huygens dans la deuxième moitié du XVII^e^ siècle : ils commençaient à s'intéresser à la théorie de jeux. Le terme probabilités a alors pris son sens actuel, et Jakob Bernoulli a traité ce sujet en tant que théorie mathématique. Au XVIII^e^ siècle, Gabriel Cramer donne un cours sur la logique probabiliste, puis au XIX^e^ siècle la théorie moderne des probabilités en mathématiques apparaît. Au début du XX^e^ siècle, l'axiomatique de Kolmogorov propose un cadre formel rigoureux pour la théorie des probabilités, permettant enfin que celle-ci devienne une branche des mathématiques à part entière.
La famille Bernoulli vivait à Bâle (Suisse). Entre le XVII^e^ et le XIX^e^ siècle, cette famille, qui était à l'origine une famille de marchands, a fourni à son temps une véritable dynastie de grands mathématiciens. Parmi eux, les frères Johann et Jakob ne s'entendaient pas très bien mais étaient tous les deux de grands admirateurs de Leibniz (ils se sont rangés à ses côtés lors de la querelle qui l'a opposé à Newton pour la paternité du calcul différentiel et intégral). C'est Jakob (1654-1705) qui a donné son nom à « l'épreuve de Bernoulli » : professeur à l'université de Bâle à partir de 1682, il fut nommé associé de l'Académie des sciences de Paris en 1699 et de celle de Berlin en 1701. C'est dans *Ars Conjectandi*, « l'art de conjecturer » (ouvrage publié à titre posthume par son neveu Nicolas Bernoulli en 1713), que Jakob Bernoulli consolide et développe la théorie des probabilités. Il y expose sa théorie des permutations et des combinaisons, qui constitue aujourd'hui les fondements de la combinatoire.

1. Variables aléatoires

a. Définitions et rappels

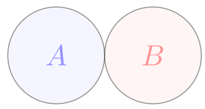
On considère une expérience aléatoire conduisant à n issues, et on note $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ l'ensemble des issues. Ω est « l'univers » des possibles. On définit la probabilité P en donnant la probabilité de chaque issue, c'est-à-dire les nombres $P(\omega_1), P(\omega_2), \ldots, P(\omega_n)$ vérifiant :

- $\forall i \in \{0, 1, \ldots, n\}, 0 \leqslant P(\omega_i) \leqslant 1$
- $\sum_{i=1}^{i=n} P(\omega_i) = P(\omega_1) + P(\omega_2) + \cdots + P(\omega_n) = 1$

Définition 1.1 Un événement A est une partie de Ω.
La probabilité de A, notée $P(A)$, est la somme des probabilités des issues appartenant à A.

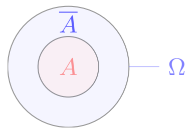
Remarques :

- A est un événement impossible si et seulement si $A = \emptyset$.
  Dans ce cas $P(A) = 0$
- A est un événement certain si et seulement si $A = \Omega$.
  Dans ce cas $P(A) = 1$



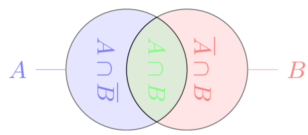
A et B incompatibles

- A et B sont incompatibles si $A \cap B$ est impossible, c'est-à-dire si $P(A \cap B) = 0$. Dans ce cas $P(A \cup B) = P(A) + P(B)$.
  L'événement $A \cup B$ est réalisé quand A **ou** B est réalisé. Ce « ou » est, en principe, inclusif mais pour des événements incompatibles, il prend le sens d'un ou exclusif.
- A et B sont des événements contraires si $A \cap B = \emptyset$ et $A \cup B = \Omega$.
  Dans ce cas on note $B = \overline{A}$ et on a $P(\overline{A}) = 1 - P(A)$.



A et $\overline{A}$ contraires

- pour des événements quelconques A et B, on a
  $$P(A \cup B) + P(A \cap B) = P(A) + P(B)$$
  Analysons la situation en termes d'ensembles incompatibles :
  $A = (A \cap B) \cup (A \cap \overline{B})$ et $B = (A \cap B) \cup (\overline{A} \cap B)$,
  d'où $A \cup B = [(A \cap \overline{B}) \cup (A \cap B)] \cup (\overline{A} \cap B)$
  et donc $P(A \cup B) = P(A) + P(\overline{A} \cap B)$,
  mais, comme $P(\overline{A} \cap B) = P(B) - P(A \cap B)$,
  on a finalement $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.



A et B compatibles

Exemple 1 – Une urne contient 100 boules numérotées de 0 à 99. On tire dans cette urne une boule et on note son numéro.
Appelons A l'événement « être divisible par 5 », B : « se terminer par 5 », C : « être divisible par 2 » et D : « être divisible par 3 ».
Il est clair que Ω est l'ensemble des 100 numéros, et que chaque numéro est équiprobable.
A peut être réalisé de 20 façons différentes car $A = \{0, 5 = 1 \times 5, \ldots, 95 = 19 \times 5\}$ et donc $P(A) = \frac{20}{100}$.
De même, on a $P(B) = \frac{10}{100}$, $P(C) = \frac{50}{100}$ et $P(D) = \frac{34}{100}$ (car $D = \{0, 3 = 1 \times 3, \ldots, 99 = 33 \times 3\}$).
$P(A \cap B) = P(B) = \frac{10}{100}$ et $P(A \cup B) = P(A) = \frac{20}{100}$ car $B \subset A$.
$P(A \cap C) = \frac{10}{100}$ car $A \cap C$: « être divisible par 10 », et
$P(A \cup C) = P(A) + P(C) - P(A \cap C) = \frac{20+50-10}{100} = \frac{60}{100}$.
$P(A \cap D) = \frac{4}{100}$ car $A \cap C = \{0, 30, 60, 90\}$, et
$P(A \cup D) = P(A) + P(D) - P(A \cap D) = \frac{20+34-4}{100} = \frac{50}{100}$.
$P(B \cap C) = 0$ car $A \cap C = \emptyset$, et donc $P(B \cup C) = P(B) + P(C) = \frac{10+50}{100} = \frac{60}{100}$.
$P(B \cap D) = \frac{3}{100}$ car $B \cap D = \{15, 45, 75\}$, et donc
$P(B \cup D) = P(B) + P(D) - P(B \cap D) = \frac{10+34-3}{100} = \frac{41}{100}$.
$P(C \cap D) = \frac{17}{100}$ car $C \cap D =$: « être divisible par 6 » $(C \cap D = \{0, 6 = 1 \times 6, \ldots, 96 = 16 \times 6\})$ et donc $P(C \cup D) = P(C) + P(D) - P(C \cap D) = \frac{50+34-17}{100} = \frac{67}{100}$.

Définition 1.2 Lorsque toutes les issues ont la même probabilité – on dit qu'il y a équiprobabilité des issues – s'il y a n issues possibles, la probabilité de chacune est $\frac{1}{n}$.

Remarque :
On reconnait une situation d'équiprobabilité au fait que les issues sont interchangeables. Par exemple, un dé cubique bien équilibré a six faces interchangeables. Que l'un ou l'autre d'entre elles apparaisse lors d'un tirage du dé, cela arrive avec la même probabilité de $\frac{1}{6}$. Quand on lance une pièce équilibrée, on obtient une face ou l'autre avec la même probabilité de $\frac{1}{2}$ car les deux faces sont interchangeables.

Propriété 1.1 Dans le cas d'une situation d'équiprobabilité à n issues possibles, un événement A qui se réalise pour k issues a une probabilité de $P(A) = \frac{k}{n}$. On résume cela en écrivant

$$P(A) = \frac{\text{nombre de cas favorables à A}}{\text{nombre de cas possibles}}$$

Exemple 2 – Dans un jeu de 52 cartes, il y a 13 cartes de niveaux différents $(\text{as}, 2, 3, \ldots, \text{valet}, \text{dame}, \text{roi})$, dans 4 couleurs différentes (cœur, carreau, trèfle, pique).
➲ Supposons que l'on tire une carte dans le jeu.
Quelles sont les probabilités de A : « tirer un cœur » et de B : « tirer une tête (valet, dame ou roi) » ?
Il s'agit d'une situation d'équiprobabilité à 52 issues possibles.
A est réalisé pour 13 d'entre elles, donc $P(A) = \frac{13}{52} = \frac{1}{4}$. On aurait pu raisonner en termes de couleurs : puisqu'il y en a 4 et qu'elles sont équiprobables, la probabilité de tirer l'une d'elle est $\frac{1}{4}$.
B est réalisable de $3 \times 4 = 12$ façons, donc $P(B) = \frac{12}{52} = \frac{3}{13}$. On aurait pu raisonner en termes de niveaux : puisqu'il y a 13 niveaux équiprobables par couleur dont 3 sont des têtes, la probabilité de tirer l'une de ces têtes est $\frac{3}{13}$.
➲ Supposons que l'on tire deux cartes du jeu.
Quelle est la probabilité de A : « tirer deux cœurs », de B : « tirer deux têtes » et de $A \cup B$?
Le nombre de mains de deux cartes est $52 \times 51 \div 2 = 1326$ (on choisit la première carte, puis la seconde et on divise par deux car chaque main possible a été comptée deux fois). Toutes ces mains sont équiprobables.
A est réalisé pour $13 \times 12 \div 2 = 78$ d'entre elles, donc $P(A) = \frac{78}{1326} = \frac{1}{17}$.
B est réalisé pour $12 \times 2 \div 2 = 12$ d'entre elles, donc $P(B) = \frac{12}{1326} = \frac{2}{221}$.
Pour déterminer $P(A \cup B)$, commençons par déterminer $P(A \cap B)$: il n'y a que 3 mains favorables à $A \cap B = \{(\text{V}\heartsuit,\text{R}\heartsuit),(\text{V}\heartsuit,\text{D}\heartsuit),(\text{D}\heartsuit,\text{R}\heartsuit)\}$, donc $P(A \cap B) = \frac{3}{1326} = \frac{1}{442}$.
On en déduit que $P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{78}{1326} + \frac{12}{1326} - \frac{3}{1326} = \frac{87}{1326}$.

La propriété 5.1 suppose que l'on puisse dénombrer les cas favorables et les cas possibles. Le dénombrement n'est pas toujours aisé et manipule parfois de très grands nombres. L'objet de ce chapitre n'étant pas de faire une étude systématique des méthodes de dénombrement, nous nous limiterons à deux exemples.

Exemple 3 – ➲ J'écoute le dernier CD de ma chanteuse préférée en mode « shuffle » : les 12 chansons passent toutes les 12, une fois et une seule, dans un ordre aléatoire. J'aimerai calculer la probabilité que les chansons 3 et 7 passent l'une après l'autre.
Le nombre d'ordres différents des 12 chansons (les cas possibles) est $12 \times 11 \times 10 \times \cdots \times 2 \times 1$, un nombre très grand que l'on note 12! (factorielle 12) et qui vaut 479 001 600. Pour déterminer le nombre de cas favorables, il faut procéder méthodiquement :

- On choisit la place des deux chansons parmi les douze places. Il y en a 11, la liste est $\{(1, 2), (2, 3), \ldots, (11, 12)\}$.
- Pour chacune de ces places, il faut choisir un des deux ordres possibles pour les deux chansons : $(3, 7)$ ou $(7, 3)$.
- On doit aussi choisir l'ordre des 10 autres chansons : il y a $10! = 3\,628\,800$ ordres possibles.

Il y a donc, en tout $11 \times 2 \times 10!$ cas favorables. La probabilité cherchée est finalement $\frac{11\times2\times10!}{12!} = \frac{2}{12} = \frac{1}{6}$. Deux chansons données passent l'une après l'autre une fois sur six, en moyenne.

➲ J'ai invité des amis à une soirée. Nous sommes douze et j'aimerai déterminer la probabilité que deux d'entre nous au moins aient la même date d'anniversaire.

Quand on voit au moins, on doit penser à l'événement contraire. Ici, le contraire de E :« deux d'entre nous au moins aient la même date d'anniversaire » est $\overline{E}$:« chacun a une date d'anniversaire différente ». Cherchons la probabilité de cet événement. On va estimer qu'il n'y a que 365 dates d'anniversaire possibles (pas d'année bissextile) et que chacune est équiprobable pour les 12 personnes.

Le nombre de cas possibles est 365^{12} un nombre colossal approximativement égal à $5,6 \times 10^{30}$. Ce résultat distingue des listes de dates d'anniversaire égales, mais on doit procéder ainsi pour conserver l'équiprobabilité des listes. On dénombre des listes ordonnées, où les personnes sont identifiées (par exemple par l'ordre alphabétique de leur nom).

Le nombre de cas favorables est le nombre de listes de dates sans répétition d'une certaine date. On peut choisir la 1^{re} date parmi 365, mais la 2^{e} est choisie parmi 364, la 3^{e} parmi 363, etc. Il y a donc $365 \times 364 \times 363 \times \cdots \times 354$ cas favorables. Un nombre très grand que l'on peut noter $\frac{365!}{353!}$ (après simplification, les deux nombres sont en effet égaux).

Finalement, $P(\overline{E}) = \frac{365!}{353!\times365^{12}}$. Ce nombre est plus facile à calculer en l'écrivant $\frac{365}{365} \times \frac{364}{365} \times \frac{363}{365} \times \cdots \times \frac{354}{365}$ et en utilisant un tableur, ou bien en écrivant un petit programme.

Ce programme est très simple, nous en donnons une version en Python ci-contre. Il nous fournit alors le résultat $P(\overline{E}) \approx 0,832975$, d'où $P(E) \approx 0,167025$.
Pour information, c'est à partir de $n = 23$ personnes assemblées que la probabilité de E devient supérieure à $0,5$.

```
N=12
P=1
for I in range(N) :
    P*=(365-I)/365
print("P(non E)={}".format(P))
print("P(E)={}".format(1-P))
```

Définition 1.3 (variable aléatoire) Si, à chaque issue ω_i de Ω, on peut associer un réel $X(\omega_i)$, on dit que l'on a défini une variable aléatoire X sur Ω. Les probabilités associées aux différentes valeurs de X est appelée « loi » de la variable X.

Remarques :

- Une variable aléatoire est une fonction définie sur Ω ; ce n'est donc pas une variable (ce sont les issues qui sont variables ici) et ce n'est pas non plus aléatoire puisque chaque issue ω_i est associée à la valeur $X(\omega_i)$ de façon bien déterminée (le hasard n'a pas sa place dans cette association).
- Une variable aléatoire permet de définir des événements. Si l'expérience aléatoire est un jeu dans lequel un joueur gagne ou perd de l'argent, on écrira $X = 5$ l'événement réalisé lorsque le gain du joueur est 5 (l'unité monétaire est généralement précisée par l'énoncé). L'évènement $X \leqslant 0$ correspond à une perte, tandis que $10 \leqslant X < 20$ correspond à un gain appartenant à l'intervalle $[10; 20[$. etc.
- La notation ensembliste est modifiée puisqu'en notant $X = x$, en réalité, on désigne l'ensemble des issues ω_i telles que $X(\omega_i) = x$, c'est-à-dire $\{\omega_i \in \Omega, X(\omega_i) = x\}$.
- La loi de probabilité de X est une fonction $f : k \mapsto P(X = k)$. Cette fonction agit sur l'ensemble, noté $X(\Omega)$, des valeurs possibles de X. On désigne parfois cette fonction par le terme de « loi de distribution » de X ou encore de « densité » de X.
- La somme de toutes les probabilités $P(X = k)$ quand k décrit l'ensemble $X(\Omega)$ est égal à 1, puisque les événements $X = k$ sont deux à deux disjoints, leur réunion formant Ω.

Exemple 4 – On lance une pièce jusqu'à obtenir pile et on note le rang k de ce premier pile. Pour chaque réalisation de cette expérience, on note $X = k$ l'événement donnant la longueur (le temps d'attente) du tirage. X est une variable aléatoire qui peut prendre n'importe quelle valeur entière non nulle.

- $X = 1$ correspond à l'obtention de pile au 1$^{\text{er}}$ lancé ; $P(X=1) = \frac{1}{2}$.
- $X = 2$ correspond à l'obtention de face au 1$^{\text{er}}$ lancé et de pile au 2$^{\text{e}}$; $P(X=2) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.
- $X = 3$ correspond à l'obtention de face aux 2 premiers lancés et de pile au 3$^{\text{e}}$; $P(X=3) = \frac{1}{2^3} = \frac{1}{8}$.
- Il en va de même pour toutes les valeurs de k, $X = k$ correspond à l'obtention de face aux $k-1$ premiers lancés et de pile au k^{e} et on a $P(X=k) = \frac{1}{2^k}$.

La loi de la variable aléatoire X est ainsi connue avec précision pour n'importe quelle valeur de k et on peut vérifier que la somme infinie de toutes ces probabilités est bien égale à 1. Ici il s'agit de la somme des n termes d'une suite géométrique de premier terme $\frac{1}{2}$ et de raison $\frac{1}{2}$ lorsque n tend vers $+\infty$ (voir le chapitre sur les suites). On peut déjà déterminer $P(X \leqslant k)$ qui correspond à cette somme jusqu'à $n = k$:

$$P(X \leqslant k) = \frac{1}{2} \times \frac{1 - \frac{1}{2^k}}{1 - \frac{1}{2}} = 1 - \frac{1}{2^k}$$

Pour la limite, on voit bien que $\frac{1}{2^k}$ tend vers 0, d'où le résultat.

b. Paramètres d'une variable aléatoire

Nous allons définir les différents paramètres (espérance, variance, écart-type) pour le cas d'un ensemble $X(\Omega)$ fini. Bien entendu, ces définitions s'étendront également aux ensembles infinis (à condition qu'ils soient dénombrables), comme on en rencontre fréquemment (voir l'exercice 86).

Définition 1.4 Soient X une variable aléatoire telle que $X(\Omega) = \{x_1, x_2, \ldots, x_n\}$.
En notant $p_i = P(X = x_i)$ les différentes valeurs de la loi de probabilité de X :

- L'espérance de X est le réel, noté $E(X)$, défini par $E(X) = \sum\limits_{i=1}^{n} p_i x_i$
- La variance de X est le réel positif, noté $V(X)$, défini par $V(X) = \sum\limits_{i=1}^{n} p_i \Big(x_i - E(X)\Big)^2$
- L'écart-type de X est le réel positif, noté $\sigma(X)$, défini par $\sigma(X) = \sqrt{V(X)}$

Remarques :

- L'espérance mathématique d'une variable aléatoire est un paramètre central, équivalent probabiliste de la notion statistique de moyenne. De même, la variance et l'écart-type sont des mesures de la dispersion que l'on calcule aussi en statistiques. La variance est donc une sorte de moyenne des carrés des écarts à l'espérance. On pourrait mesurer la dispersion en utilisant une autre formule, par exemple la moyenne des valeurs absolues des écarts à l'espérance. Si on réalise l'expérience aléatoire un très grand nombre de fois, les paramètres statistiques (fréquences, moyenne, écart-type) se rapprochent des résultats théoriques calculés en probabilités.
- Dans un jeu où X est associé aux gains possibles, $E(X)$ représente le gain moyen que l'on peut espérer. Le jeu est dit équitable si $E(X) = 0$, favorable au joueur si $E(X) > 0$ et défavorable au joueur si $E(X) < 0$.

Exemple 5 – Lorsqu'on lance deux dés cubiques, chacune des faces étant numérotée entre 1 et 6, il y a 36 résultats équiprobables. La variable X donnant la somme des dés prend les valeurs entières comprises entre 2 et 12, selon la loi de probabilité suivante :

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

$E(X) = \sum\limits_{i=1}^{12} p_i x_i = 2 \times \frac{1}{36} + 3 \times \frac{2}{36} + 4 \times \frac{3}{36} + 5 \times \frac{4}{36} + 6 \times \frac{5}{36} + 7 \times \frac{6}{36} + 8 \times \frac{5}{36} + \cdots + 12 \times \frac{1}{36} = \frac{252}{36} = 7$

$V(X) = \sum\limits_{i=1}^{12} p_i \Big(x_i - 7\Big)^2 = (2-7)^2 \times \frac{1}{36} + (3-7)^2 \times \frac{2}{36} + (4-7)^2 \times \frac{3}{36} + \cdots + (12-7)^2 \times \frac{1}{36} = \frac{210}{36} = \frac{35}{6} \approx 5,83$

$\sigma(X) = \sqrt{\frac{35}{6}} \approx 2,42$

Exemple 6 – Pour donner un exemple de calcul de ces paramètres, dans le cas d'un ensemble $X(\Omega)$ infini dénombrable, on peut reprendre l'exemple 86 où la loi de probabilité se résume à l'égalité $p_i = P(X = x_i) = \frac{1}{2^i}$, valable pour tout $i \geqslant 1$.

$E(X) = \sum\limits_{i=1}^{+\infty} p_i x_i = \frac{1}{2} + 2 \times \frac{1}{4} + 3 \times \frac{1}{8} + 4 \times \frac{1}{16} + \ldots$

Pour déterminer la valeur exacte de cette somme infinie, on se trouve encore un peu démuni. Mais une approche algorithmique nous en donnera une très bonne valeur approchée.
Programmons le calcul approché de $E(X)$ et donnons différentes valeurs de sortie.

n	5	10	20	40	80
$\sum\limits_{i=1}^{n} \frac{i}{2^i}$	$1,7825$	$1,98828$	$1,99997901$	$1,99999999996$	1,999999999999

Il semblerait que $E(X) = 2$; les valeurs ne changent pas entre 40 et 80, signe que la capacité du calcul informatique est dépassée pour $n = 80$. En prenant $E(X) = 2$, on peut compléter la caractérisation de cette variable aléatoire et calculer une valeur approchée de $V(X)$ et de $\sigma(X)$. On trouve alors, au vu des valeurs suivantes, que $V(X) = 2$ et donc $\sigma(X) = \sqrt{2} \approx 1,414213562$:

n	5	10	20	40	80
$e_n = \sum\limits_{i=1}^{n} \frac{(i-2)^2}{2^i}$	$1,15625$	$1,900390$	$1,99961$	$1,999999998$	1,999999999
$\sigma_n = \sqrt{e_n}$	$1,07529$	$1,378546$	$1,41407$	$1,414213561$	1,414213562

Soit X une variable aléatoire. Cette variable X prenant des valeurs réelles, on peut être amené à effectuer des opérations sur elle. Par exemple, on peut doubler ses valeurs et s'intéresser à la variable aléatoire $Y = 2X$. On peut ajouter une quantité fixe et s'intéresser à la variable aléatoire $Z = X + k$.

Propriété 1.2 Soient X une variable aléatoire et a et b deux réels. Les paramètres de la variable aléatoire $X' = aX + b$ sont $E(X') = E(aX + b) = aE(X) + b$ et $V(X') = V(aX + b) = a^2V(X)$.

Démonstration Ces propriétés découlent des propriétés de la multiplication et de l'addition, ainsi que des définitions, notamment $\sum p_i = 1$, $\sum p_i x_i = E(X)$ et $\sum p_i(x_i - E(X))^2 = V(X)$:

$$\begin{aligned}
E(aX+b) &= (ax_1+b)\times p_1 + (ax_2+b)\times p_2 + \ldots + (ax_n+b)\times p_n \\
&= (ax_1p_1 + ax_2p_2 + \ldots + ax_np_n) + (bp_1 + bp_2 + \ldots + bp_n) \\
&= a(x_1p_1 + x_2p_2 + \ldots + x_np_n) + b(p_1 + p_2 + \ldots + p_n) \\
&= aE(X) + b \times 1 = aE(X) + b
\end{aligned}$$

$$\begin{aligned}
V(aX+b) &= \sum_{i=1}^{n} p_i(ax_i + b - E(aX+b))^2 = \sum_{i=1}^{n} p_i(ax_i + b - aE(X) - b)^2 \\
&= \sum_{i=1}^{n} p_i(ax_i - aE(X))^2 = \sum_{i=1}^{n} p_i \times a^2 \times (x_i - E(X))^2 \\
&= a^2 \times \sum_{i=1}^{n} p_i \times (x_i - E(X))^2 = a^2V(X)
\end{aligned}$$

Conséquence : On a donc aussi $\sigma(aX+b) = \sqrt{V(aX+b)} = \sqrt{a^2V(X)} = |a|\sqrt{V(X)} = |a|\sigma(X)$

Exemple 7 – Le Duc de Toscane a remarqué que, lorsqu'on tire 3 dés, la somme 10 tombe un peu plus souvent que la somme 9, alors qu'on peut obtenir chacune de ces deux sommes de six façons :

- $9 = 6+2+1 = 5+3+1 = 5+2+2 = 4+4+1 = 4+3+2 = 3+3+3$
- $10 = 6+3+1 = 6+2+2 = 5+4+1 = 5+3+2 = 4+4+2 = 4+3+3$

Il s'ouvre alors de ce paradoxe à Galilée (1554-1642) qui rédige un cours traité sur la question, jetant ainsi les bases de la théorie des probabilités. Celui-ci dénombre les $6^3 = 216$ issues équiprobables pour le tirage de 3 dés ainsi que les cas favorables :

- pour 9, il y a $6+6+6+3+3+1=25$ possibilités équiprobables (on distingue 6 cas quand les chiffres sont différents, par exemple $621, 612, 126, 162, 216, 261$; 3 cas quand il y a répétition d'un chiffre, par exemple $522, 252, 225$ et un seul cas pour 333).
- pour 10, il y a $6+6+6+3+3+3=27$ possibilités équiprobables.

En appelant X la variable aléatoire « somme des dés », on a $P(X=9)=\frac{25}{216}$ et $P(X=10)=\frac{27}{216}$.
L'observation du duc était bien fondée.
Imaginons qu'un joueur mise 5 € sur la somme $X=10$ avec la promesse de recevoir 8 fois sa mise s'il gagne. En appelant Y la variable aléatoire donnant son gain algébrique (-5 € s'il perd et 40 € s'il gagne), quelle est l'espérance de gain du joueur ?
$E(Y)=40\times\frac{27}{216}-5\times\frac{216-27}{216}=\frac{1080-945}{216}=\frac{-135}{216}=0,625$ €.
Ce jeu avantage donc le joueur. Si on souhaite le rendre équitable, il suffit d'abaisser de $0,625$ € chacun des montants : le joueur « gagne » $-5-0,625=-5,625$ € quand il perd (sa mise passe donc à $5,625$ €), et quand il gagne, on lui verse $40-0,625=39,375$ €, soit $39,375\div 5,625=7$ fois sa mise. La nouvelle espérance est ainsi abaissée de $0,625$ € et le jeu devient équitable.
On a utilisé la propriété 5.2 : $E(Y+b)=E(Y)+b$, soit ici $E(Y-0,625)=E(Y)-0,625$.
Maintenant, on peut conserver la même espérance nulle en multipliant par un coefficient quelconque a, car la même propriété 5.2 affirme que $E(aY)=aE(Y)$. Pour éviter d'avoir à miser $5,625$ €, on multiplie cette mise par $\frac{5}{5{,}625}=\frac{8}{9}$. Ainsi, la nouvelle mise revient à 5 € et le gain en cas de sortie du 10 devient $39,375\times\frac{8}{9}=35$ €.
Globalement, on a effectué le changement $Y'=\frac{5}{5{,}625}(Y-0,625)=\frac{8}{9}Y-\frac{5}{9}$ et la propriété 5.2 qui affirme que $E(Y')=E(\frac{8}{9}Y-\frac{5}{9})=\frac{8}{9}E(Y)-\frac{5}{9}$. Comme $E(Y)=0,625$ ici, on obtient $E(Y')=0$.

Propriété 1.3 Autre formule pour la variance Soit X une variable aléatoire d'espérance $E(X)$ et de variance $V(X)$. Alors on a $V(X)=E(X^2)-\Big(E(X)\Big)^2$

Démonstration

$$\begin{aligned}
V(X) &= \sum_{i=1}^{n} p_i\Big(x_i-E(X)\Big)^2=\sum_{i=1}^{n}p_ix_i^2-2\sum_{i=1}^{n}p_ix_iE(X)+\sum_{i=1}^{n}p_i\Big(E(X)\Big)^2\\
&= E(X^2)-2\sum_{i=1}^{n}p_ix_iE(X)+\sum_{i=1}^{n}p_iE(X)\times E(X)=E(X^2)-\sum_{i=1}^{n}\Big(p_iE(X)\,(2x_i-E(X))\Big)\\
&= E(X^2)-E(X)\left(\sum_{i=1}^{n}2p_ix_i-\sum_{i=1}^{n}p_iE(X)\right)=E(X^2)-E(X)\left(2E(X)-E(X)\sum_{i=1}^{n}p_i\right)\\
&= E(X^2)-E(X)\,(2E(X)-E(X)\times 1)=E(X^2)-E(X)\times E(X)=E(X^2)-\Big(E(X)\Big)^2
\end{aligned}$$

Remarques : Cette propriété est un cas particulier d'un théorème connu sous le nom de « Kœnig-Huyghens » qui dit que $E((X-a)^2)=V(X)+(E(X)-a)^2$.
En effet, en prenant $a=0$, on obtient $E(X^2)=V(X)+\Big(E(X)\Big)^2$.
Pour démontrer ce théorème, on peut partir de l'égalité $x_i-a=x_i-E(X)+E(X)-a$ qui, élevée au carré s'écrit $(x_i-a)^2=(x_i-E(X))^2+(E(X)-a)^2+2\times(x_i-E(X))(E(X)-a)$.
Multiplions ensuite cette nouvelle égalité par p_i et sommons les quantités obtenues pour toutes les valeurs de i entre 1 et n. On obtient :

$$E((X-a)^2)=V(X)+(E(X)-a)^2+2\times(E(X)-a)\times\sum_{i=1}^{n}p_i(x_i-E(X))$$

Le dernier terme de cette somme est nul car
$\sum\limits_{i=1}^{n}p_i(x_i-E(X))=\sum\limits_{i=1}^{n}p_ix_i-E(X)\sum\limits_{i=1}^{n}p_i=E(X)-E(X)=0$, d'où le résultat.

Cette propriété facilite le calcul algorithmique de $V(X)$ puisque l'on n'a pas besoin de connaître « à l'avance » la valeur de $E(X)$: on calcule en parallèle $\sum p_i x_i$ (pour l'obtention de $E(X)$) et $\sum p_i x_i^2$ (pour l'obtention de $E(X^2)$). Supposons que l'on cherche la variance de l'exemple 87 où on s'intéresse à X la somme de deux dés. Au lieu de faire comme dans l'exemple, on peut calculer simultanément :

$E(X) = \sum_{i=1}^{12} p_i x_i = 2 \times \frac{1}{36} + 3 \times \frac{2}{36} + 4 \times \frac{3}{36} + \cdots + 12 \times \frac{1}{36} = \frac{252}{36} = 7$

$E(X^2) = \sum_{i=1}^{12} p_i x_i^2 = 2^2 \times \frac{1}{36} + 3^2 \times \frac{2}{36} + 4^2 \times \frac{3}{36} + \cdots + 12^2 \times \frac{1}{36} = \frac{1974}{36}$

Ensuite, on effectue $E(X^2) - \left(E(X)\right)^2 = \frac{1974}{36} - \left(\frac{252}{36}\right)^2 = \frac{7560}{36^2} = \frac{35}{6}$.

Dans l'application suivante où on calcule la variance de l'exemple 86, on ne connait $E(X)$ avec certitude qu'en passant à la limite. La propriété énoncée apparait donc encore plus utile puisqu'elle permet d'estimer $E(X)$ et $V(X)$ à chaque étape du calcul :

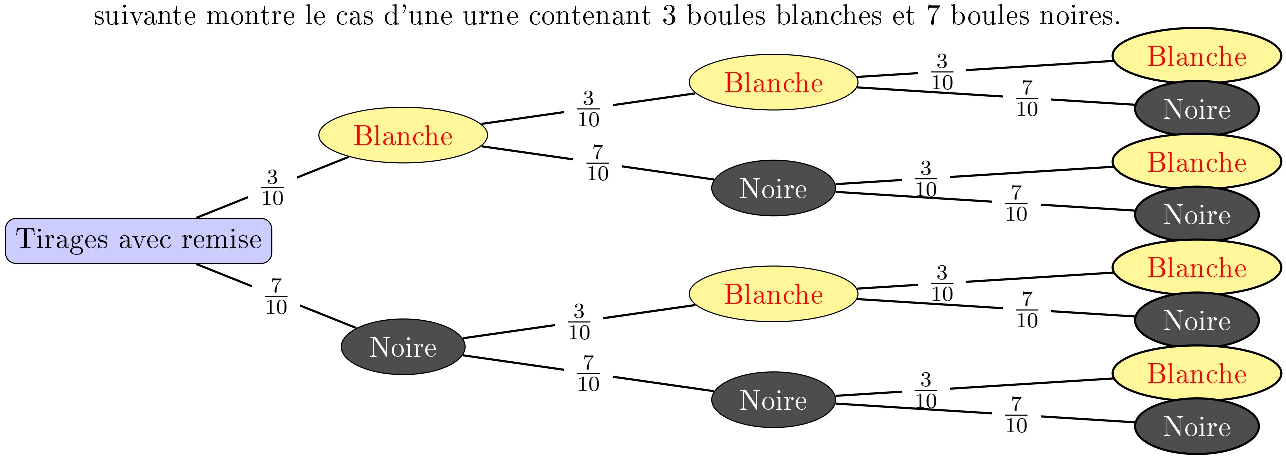
	n	5	10	20	40	80
Estimation de E(X)	$\sum p_i x_i = \sum_{i=1}^{n} \frac{i}{2^i}$	$1,625$	$1,9785$	$1,99995$	$1,999999999$	1,9999999999
	$\sum p_i x_i^2 = \sum_{i=1}^{n} \frac{i^2}{2^i}$	$4,406$	$5,8574$	$5,99953$	$5,999999998$	5,9999999999
Estimation de V(X)	$\sum p_i x_i^2 - (\sum p_i x_i)^2$	$1,233$	$1,9041$	$1,99962$	$1,999999998$	2,0

c. Répétition d'expériences

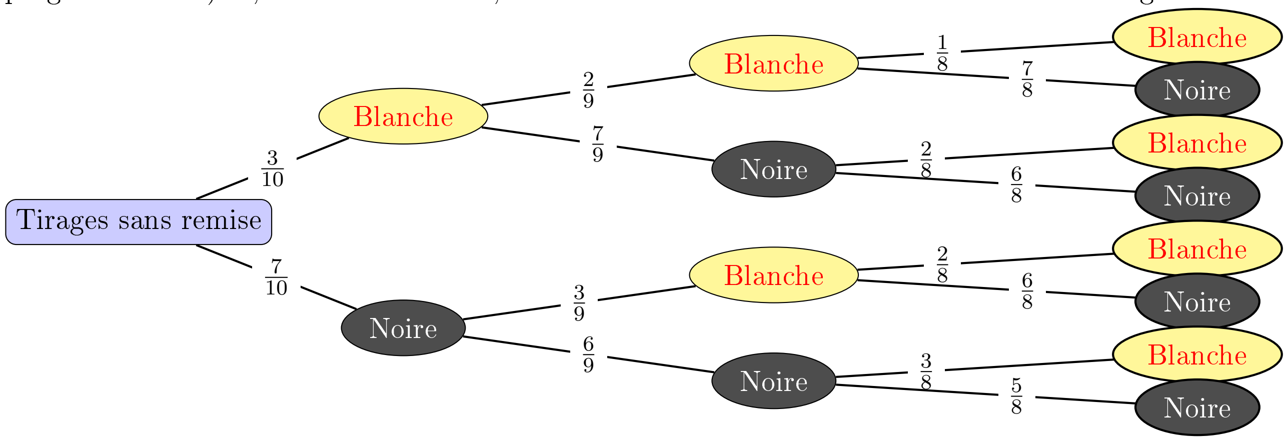
Définition 1.5 Deux expériences sont indépendantes si le résultat de la première n'influence pas le résultat de la seconde, c'est-à-dire qu'il ne modifie pas la loi de probabilités de celle-ci.

Exemples :

- Je lance un dé à six faces (dé cubique) en même temps qu'un dé à quatre faces (dé tétraédrique). Ces deux expériences simultanées sont indépendantes car le résultat du dé cubique n'influence pas le résultat dé tétraédrique. Même chose si je lance deux dés cubiques : chacun des dés indique un résultat indépendant de l'autre. Le processus expérimental importe peu : je peux lancer les dés ensemble ou séparément, sans changer la loi de probabilité des résultats.
- Si je lance une pièce de monnaie jusqu'à obtenir un « pile », les tirages successifs sont indépendants, la pièce n'ayant pas de mémoire : ce n'est pas parce que j'ai déjà obtenu 5 fois « face » que j'ai plus de chance de faire « pile » au sixième lancé. Les lancés sont indépendants. L'existence du sixième lancé, par contre, dépend du résultat du cinquième puisqu'on ne l'effectue qu'à condition d'avoir obtenu un « face. »
- Dans un vidage d'urne (expérience souvent choisie car elle permet d'illustrer des situations très différentes), si on remet la boule tirée après avoir noté sa couleur, son numéro, les tirages successifs sont indépendants. Si on ne remet pas la boule, alors chacun des tirages influe sur la loi de probabilités de la boule suivante. Les tirages ne sont pas indépendants. L'illustration suivante montre le cas d'une urne contenant 3 boules blanches et 7 boules noires.



Ci-dessus : les tirages se font « avec remise ». Il y a indépendance entre les différents tirages. Les probabilités sont inchangées d'une étape à l'autre.
Ci-dessous : les tirages se font « sans remise ». Il n'y a pas indépendance entre les différents tirages. Les probabilités changent d'une étape à l'autre : le nombre de boules diminue (on vide l'urne progressivement) et, selon les résultats, le nombre de boules d'une couleur diminue ou stagne.



Propriété 1.4 Un arbre de probabilités représente la répétition d'expériences aléatoires : les nœuds (sommets) sont les issues de ces expériences et les nombres inscrits sur les branches (arêtes, arcs) sont les probabilités de ces événements. Que les expériences soient dépendantes ou indépendantes les unes des autres, ces arbres vérifient toujours les propriétés suivantes :

- les probabilités des branches partant d'un même nœud ont une somme égale à 1
- la probabilité d'un « chemin » est le produit des probabilités des branches de ce chemin
- la probabilité d'un événement réalisé par plusieurs chemins est la somme des probabilités de ces chemins

Démonstration En partant d'un même nœud, on représente les différentes issues d'une expérience aléatoire. Ces issues sont incompatibles et forment l'univers Ω des possibles de cette expérience. La somme des probabilités vaut donc 1.
Le produit des probabilités sur un chemin se théorise en termes de probabilités conditionnelles : on note $P(A/B)$ la probabilité de réaliser A sachant que B a été réalisé. Cette probabilité est égale à $\frac{P(A\cap B)}{P(B)}$, ce qu'on peut écrire aussi $P(A\cap B) = P(B) \times P(A/B)$. L'événement $A \cap B$ est le chemin « nœud B puis nœud A » dont la probabilité est obtenue par le produit des probabilités inscrites sur les branches de ce chemin. D'une façon intuitive, si B se produit 3 fois sur 10 (comme dans l'urne ci-dessus) et que A se produit 7 fois sur 9 après que A se soit produit, alors la succession B puis A se produit les $\frac{7}{9}$ de $\frac{3}{10}$ du temps, soit les $\frac{7}{9} \times \frac{3}{10}$ du temps.
La définition 5.1 de la probabilité d'un événement implique que si plusieurs chemins réalisent un événement particulier, alors la probabilité de cet événement est la somme des probabilités de ces chemins.

Exemple 8 – En reprenant l'exemple de notre urne contenant 3 boules blanches et 7 noires. On tire successivement 3 boules et on envisage les deux scénarios (avec ou sans remise). Notons X la variable aléatoire comptant le nombre de boules blanches obtenues et déterminons la loi de X.
➟ $X = 0$ est l'événement « obtenir 3 boules noires » qui n'est réalisé que sur un chemin.
Sans remise, $P(X = 0) = \frac{7}{10} \times \frac{6}{9} \times \frac{5}{8} = \frac{210}{720} = \frac{7}{24} \approx 0,292$.
Avec remise, $P(X = 0) = \left(\frac{7}{10}\right)^3 = \frac{343}{1000} = 0,343$.
➟ $X = 1$ est l'événement « obtenir 2 boules noires et 1 blanche ».
Cet événement est réalisé par 3 chemins, selon le rang d'apparition de la boule blanche.
Sans remise, $P(X = 1) = \frac{3}{10} \times \frac{7}{9} \times \frac{6}{8} + \frac{7}{10} \times \frac{3}{9} \times \frac{6}{8} + \frac{7}{10} \times \frac{6}{9} \times \frac{3}{8} = 3 \times \frac{126}{720} = \frac{21}{40} = 0,525$.
Avec remise, $P(X = 1) = 3 \times \frac{3}{10} \times \left(\frac{7}{10}\right)^2 = \frac{441}{1000} = 0,441$.

➟ De même, on montre que, sans remise $P(X=2)=\frac{3\times 42}{720}=\frac{7}{40}=0,175$ et, avec remise $P(X=2)=\frac{189}{1000}=0,189$.
➟ Sans remise $P(X=3)=\frac{6}{720}=\frac{1}{120}\approx 0,008$ et, avec remise $P(X=3)=\frac{27}{1000}=0,027$.
Vérifions ces résultats et calculons les paramètres de X pour ces deux scénarios :

Sans remise	Vérification : $\sum_{k=0}^{3} P(X=k)=\frac{7}{24}+\frac{21}{40}+\frac{7}{40}+\frac{1}{120}=\frac{35+63+21+1}{120}=1$. $E(X)=0\times\frac{7}{24}+1\times\frac{21}{40}+2\times\frac{7}{40}+3\times\frac{1}{120}=\frac{63+42+3}{120}=\frac{108}{120}=\frac{9}{10}=0,9$. $E(X^2)=0\times\frac{7}{24}+1\times\frac{21}{40}+4\times\frac{7}{40}+9\times\frac{1}{120}=\frac{63+84+9}{120}=\frac{156}{120}=\frac{13}{10}=1,3$. D'où $V(X)=E(X^2)-E(X^2)=1,3-0,9^2=0,49$ et $\sigma(X)=\sqrt{0,49}=0,7$.
Avec remise	Vérification : $\sum_{k=0}^{3} P(X=k)=\frac{343+441+189+27}{1000}=1$. $E(X)=\frac{0\times 373+1\times 441+2\times 189+3\times 27}{1000}=\frac{900}{1000}=0,9$. $E(X^2)=\frac{0\times 373+1\times 441+4\times 189+9\times 27}{1000}=\frac{1440}{1000}=1,44$. D'où $V(X)=E(X^2)-E(X^2)=1,44-0,9^2=0,63$ et $\sigma(X)=\sqrt{0,63}\approx 0,79$.

2. Loi binomiale

a. Épreuve de Bernoulli

Définition 1.6 Une expérience aléatoire admettant deux issues est appelée épreuve de Bernoulli.
- l'une des issues est appelée « succès », notée S ; sa probabilité est p.
- l'autre issue est appelée « échec », notée $\overline{S}$; sa probabilité est $1-p$

On dit que la variable aléatoire X prenant la valeur 1 en cas de succès et la valeur 0 en cas d'échec suit une loi de Bernouilli de paramètre p et on appelle X une variable de Bernoulli.

Propriété 1.5 Soit X une variable aléatoire de Bernoulli de paramètre p.
Les paramètres de X sont : $E(X)=p$; $V(X)=p\times(1-p)$; $\sigma(X)=\sqrt{V(X)}=\sqrt{p\times(1-p)}$

Démonstration On a : $E(X)=0\times(1-p)+1\times p=p$.
De même : $V(X)=E(X^2)-(E(X))^2=0^2\times(1-p)+1^2\times p-p^2=p(1-p)$

Exemple : Tirer une boule de l'urne précédente qui contient 3 boules blanches et 7 noires constitue une épreuve de Bernoulli. Si, on convient d'appeler succès le tirage d'une boule blanche, il s'agit d'une épreuve de Bernoulli de paramètre $p=\frac{3}{10}$. La variable aléatoire de Bernoulli X de paramètre p qui prend la valeur 1 en cas de succès et 0 en cas d'échec a une espérance $E(X)=p=\frac{3}{10}$ et une variance $V(X)=p(1-p)=\frac{21}{100}$.
Si, au contraire, on convient d'appeler succès le tirage d'une boule noire, il s'agit d'une épreuve de Bernoulli de paramètre $q=\frac{7}{10}$. La variable aléatoire de Bernoulli X' de paramètre q qui prend la valeur 1 en cas de succès et 0 en cas d'échec a une espérance $E(X')=q=\frac{7}{10}$ et une variance $V(X')=q(1-q)=\frac{21}{100}$.

b. schéma de Bernoulli

Définition 1.7 Soit une épreuve de Bernoulli de paramètre p. La répétition de n fois ($n\geqslant 1$) cette même épreuve de Bernoulli, de façon indépendante, constitue une expérience aléatoire appelée « schéma de Bernoulli ».

Remarques : L'exemple précédent, où on tire successivement et avec remise, trois boules d'une urne qui en contient 3 blanches et 7 noires est un schéma de Bernoulli, du fait que les tirages sont indépendants les uns des autres. Avec des tirages sans remise, ce ne serait pas un schéma de Bernouilli puisqu'un tirage modifie les probabilités. Si on considère comme « succés » le tirage d'une boule blanche, il s'agit d'un schéma de Bernoulli de paramètres $n = 3$ et $p = \frac{3}{10}$.
Dans un schéma de Bernoulli, si on s'intéresse à l'événement E_k : « obtenir k succés », celui-ci est réalisé par différents chemins qui ont tous la même probabilité : $p^k \times (1-p)^{n-k}$. Il suffit de dénombrer ces chemins pour déterminer la probabilité de l'événement E_k.

Définition 1.8 Dans le cas d'une répétition de n épreuves de Bernoulli identiques et indépendantes, on note $\binom{n}{k}$ ou C_n^k le nombre de chemins aboutissant à k succès exactement ($0 \leqslant k \leqslant n$). L'entier $\binom{n}{k}$ est appelé « coefficient binomial » et lu « k parmi n ».

Notre étude du tirage avec remise de trois boules dans l'urne nous a montré qu'avec 3 épreuves de Bernouilli successives, les coefficients sont : $\binom{3}{0} = 1$, $\binom{3}{1} = 3$, $\binom{3}{2} = 3$ et $\binom{3}{3} = 1$.
On peut remarquer que ces coefficients sont ceux du développement du « binôme » $(a+b)^n$. Par exemple, lorsque $n = 3$, on a $(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$ et on voit bien que les coefficients obtenus sont les $\binom{3}{k}$ déterminés précédemment, avec $0 \leqslant k \leqslant 3$. De même, les coefficients 1, 2 et 1 du développement de $(a+b)^2 = a^2 + 2ab + b^2$ sont les nombres $\binom{2}{k}$, avec $0 \leqslant k \leqslant 2$. Cela explique l'expression « coefficient binomial » pour les désigner. Les propriétés de ces nombres sont nombreuses, nous en verrons quelques-unes plus loin.

Propriété 1.6 Lors d'un schéma de Bernoulli de paramètres n et p, une variable aléatoire comptant le nombre de succès suit une loi de probabilités appelée loi binomiale de paramètres n et p, notée $\mathscr{B}(n,p)$. Soit X une variable aléatoire qui suit une loi binomiale de paramètres n et p. Alors on a :

$$\forall k \in \{0,1,2,...,n\}, \qquad P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Démonstration Cette propriété découle de la définition des coefficients binomiaux qui dénombrent les chemins qui contiennent k succès et qui, du fait de la propriété 5.4, ont la probabilité $p^k(1-p)^{n-k}$ d'advenir.

Exemple 9 – On lance quatre fois de suite un dé équilibré à six faces. On appelle succès l'obtention d'un 6 et on considère la variable aléatoire qui comptabilise le nombre de succès. Cette loi suit une loi binomiale $\mathscr{B}(4, \frac{1}{6})$ car le dé est équilibré et n'a pas de mémoire (les épreuves de Bernoulli sont bien identiques et indépendantes).
On peut construire un arbre qui contient $2^4 = 16$ chemins. L'arbre complet est dessiné plus bas. Il n'est pas d'un grand intérêt excepté, peut-être, pour illustrer le dénombrement que nous allons faire : 1 seul chemin mène à 0 succès et 1 seul mène à 4 succès ; 4 mènent à 1 succès et 4 mènent à 3 succès. Cela fait un total de $1 + 1 + 4 + 4 = 10$ chemins. Il en reste donc $16 - 10 = 6$ pour ceux qui réalisent 2 succès. Nous avons marqué ces six chemins sur l'illustration ci-dessous.
Les coefficients binomiaux pour un schéma de Bernoulli de paramètre $n = 4$ sont donc :

$$\binom{4}{0} = 1,\ \binom{4}{1} = 4,\ \binom{4}{2} = 6,\ \binom{4}{3} = 4 \text{ et } \binom{4}{4} = 1$$

Vous pouvez vérifier que ces nombres sont bien les coefficients du développement de :

$$(a+b)^4 = (a^2 + 2ab + b^2)^2 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$$

La loi de probabilités de la variable X comptabilisant les succès est donnée par la formule suivante :

$$P(X=k) = \binom{4}{k} \times \left(\frac{1}{6}\right)^k \times \left(\frac{5}{6}\right)^{4-k}$$

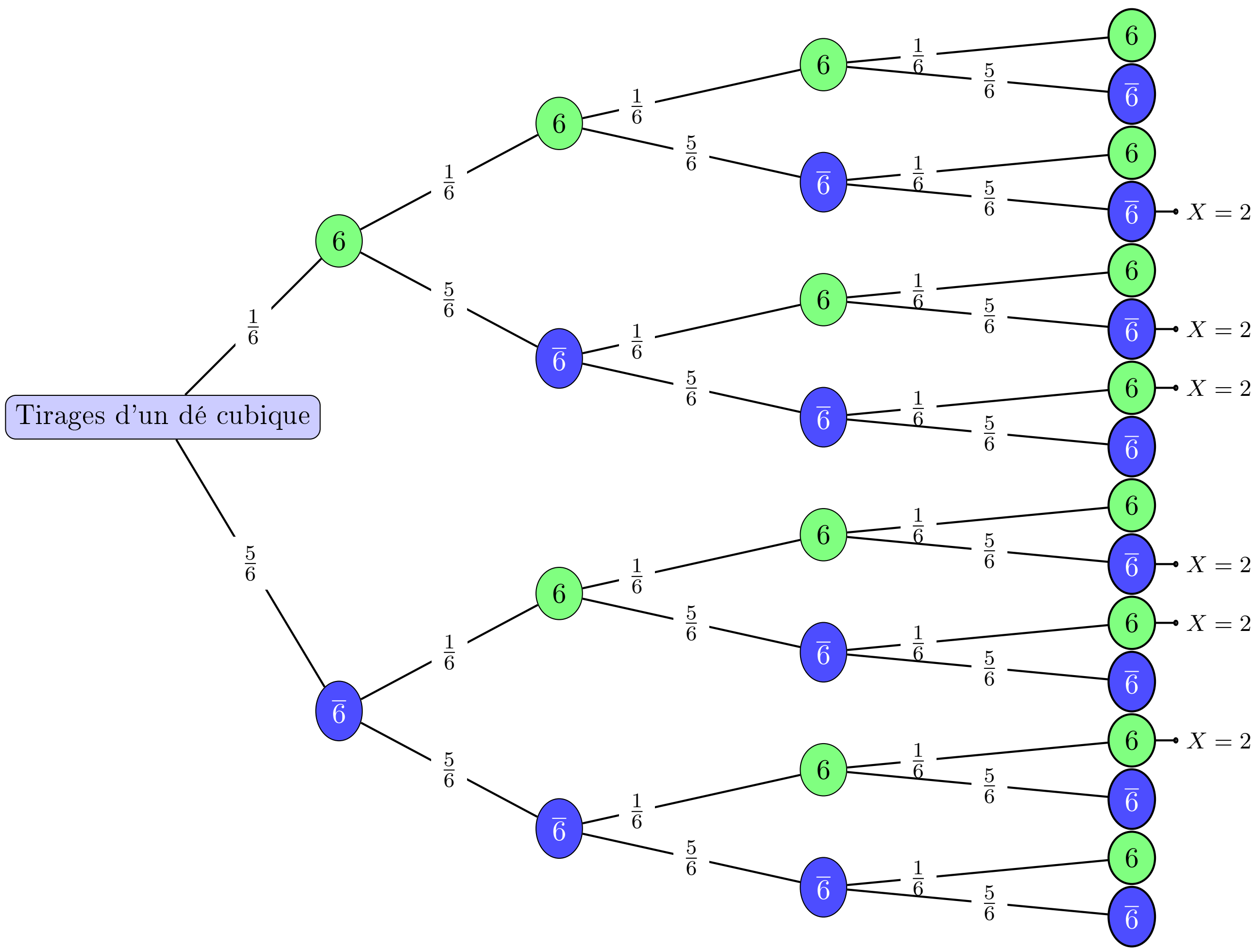
Donnons les résultats de façon plus explicite :

k	0	1	2	3	4
$P(X=k)$	$\left(\frac{5}{6}\right)^4$ $\approx 0,4823$	$4\times\left(\frac{1}{6}\right)\times\left(\frac{5}{6}\right)^3$ $\approx 0,3858$	$6\times\left(\frac{1}{6}\right)^2\times\left(\frac{5}{6}\right)^2$ $\approx 0,1157$	$4\times\left(\frac{1}{6}\right)^3\times\left(\frac{5}{6}\right)$ $\approx 0,0154$	$\left(\frac{1}{6}\right)^4$ $\approx 0,0008$

Calculons l'espérance et la variance de X :

$E(X)=0\times\left(\frac{5}{6}\right)^4+1\times 4\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^3+2\times 6\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2+3\times 4\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)+4\times\left(\frac{1}{6}\right)^4=\frac{2}{3}$

$V(X)=\left[0\times\left(\frac{5}{6}\right)^4+1\times 4\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^3+4\times 6\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2+9\times 4\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)+16\times\left(\frac{1}{6}\right)^4\right]-\left[\frac{2}{3}\right]^2=\frac{5}{9}$



Propriété 1.7 (Paramètres de la loi binomiale)
Soit X une variable aléatoire $\mathscr{B}(n,p)$ qui suit une loi binomiale de paramètres n et p. Alors :

$$E(X)=np \quad ; \quad V(X)=np(1-p) \quad ; \quad \sigma(X)=\sqrt{V(X)}=\sqrt{np(1-p)}$$

Démonstration (Pistes pour une démonstration)
Cette propriété découle d'une propriété de l'espérance, la linéarité :

> si Y_1 et Y_2 sont deux variables de Bernoulli définies sur la même expérience, alors $E(Y_1+Y_2)=E(Y_1)+E(Y_2)$.

Appliquée aux n variables de Bernoulli qui constituent le schéma de Bernoulli que l'on cherche à caractériser, cette propriété permet de comprendre que
$E(X)=E(Y_1+Y_2+\cdots+Y_n)=nE(Y_1)=np$, car $E(Y_1)=E(Y_2)=\cdots=E(Y_n)$.
La variance a également cette propriété de linéarité et, comme
$V(Y_1)=V(Y_2)=\cdots=V(Y_n)=p(1-p)$, on peut en déduire que $V(X)=np(1-p)$.

<u>Utilisation de la calculatrice</u> : La calculatrice permet de déterminer les coefficients binomiaux $\binom{n}{k}$, les valeurs de $P(X = k)$ et aussi $P(X \leqslant k)$ pour une variable aléatoire X qui suit une loi binomiale $\mathscr{B}(n, p)$, avec $n \in \mathbb{N}^*$, $p \in [0; 1]$ et k entier tel que $0 \leqslant k \leqslant n$. En se mettant dans le mode « Table », on peut également obtenir un tableau de ces différentes valeurs : on remplace k par X, et on écrit Y1=binomFdp(n,p,X) sur TI ou Bpd(X, n, p) sur Casio. Sur Numworks, le tableau n'est pas (encore) disponible, mais les différentes valeurs sont représentées graphiquement par un diagramme en bâton.

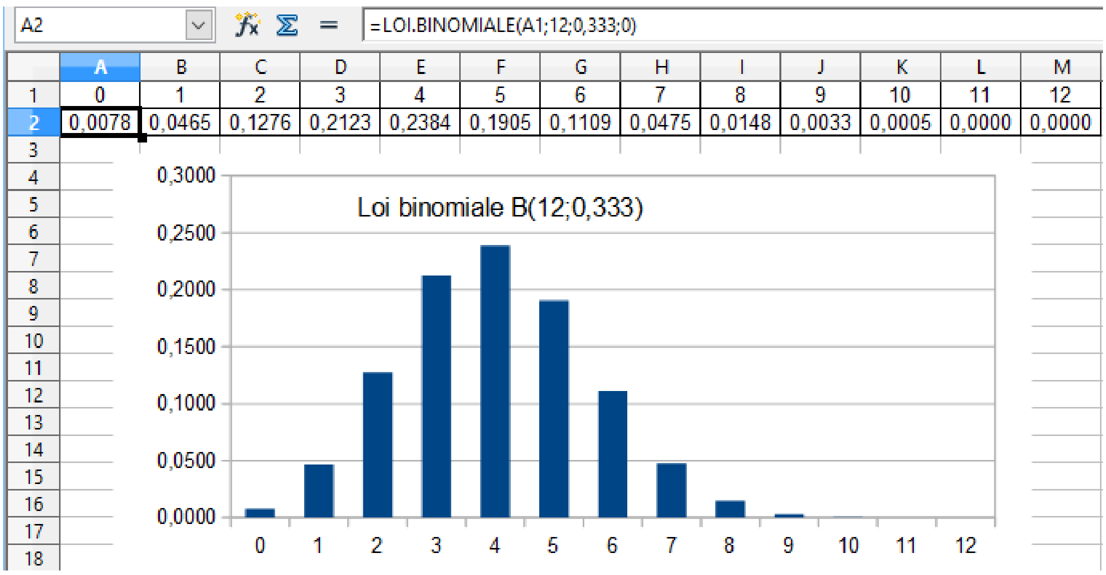
	TI	Casio	Numworks
$\binom{n}{k}$	touche math option PRB choisir 3 (combinaison)	menu RUN option PROB(F3) choisir nCr	menu Calculs, Toolbox (5$^{\text{e}}$ touche), Dénombrement, binomial(n,k), entrer n et k
$P(X = k)$	touche 2nde puis var dans DISTRIB choisir binomFdp(, compléter avec les valeurs de n, p, k et)	menu STAT options DIST(F5) puis BINM(F5). Compléter l'option Bpd(F1) « probability distribution » avec les valeurs de k, n, p	menu Probabilités, choisir Binomiale, entrer n,p, puis Suivant, sélectionner la 1$^{\text{re}}$ icône
$P(X \leqslant k)$	dans l'onglet DISTRIB, choisir binomFRép(et compléter avec les valeurs de n, p, k	Compléter l'option Bcd(F2) « cumulative distribution » avec les valeurs de k, n, p	menu Probabilités, choisir Binomiale, entrer n,p, puis Suivant, sélectionner la 4$^{\text{e}}$ icône

<u>Utilisation du tableur</u> : Avec un tableur (*calc* d'OpenOffice), on a accès aux valeurs utiles :

- les coefficients binomiaux $\binom{n}{k}$: taper la formule « =COMBIN(n ;k) »)
- les valeurs de $P(X = k)$: taper la formule « =LOI.BINOMIALE(k ;n ;p ;0) »)
- les valeurs de $P(X \leqslant k)$: taper la formule « =LOI.BINOMIALE(k ;n ;p ;1) »)

Remarques : avec le tableur *Excel* de Microsoft, il suffit de remplacer les codes 0 et 1 par `False` et `True`. En faisant varier k entre 0 et n dans une feuille de calculs, on obtient simplement le tableau des $P(X = k)$ d'une loi binomiale. Ce tableau fourni à l'assistant graphique conduit à la représentation de cette loi par un diagramme en bâtons.

Pour représenter la loi binomiale $\mathscr{B}(12; 0, 333)$ à l'aide du tableur : commencer par entrer les valeurs de k sur la ligne 1 (de 0 à 12), puis entrer la formule en A2 : « =LOI.BINOMIALE(A1 ;12 ;0,333 ;0) ». Recopier ensuite cette formule en tirant la poignée vers la droite (voir l'illustration).
Pour obtenir le diagramme, sélectionner la plage des données (de la cellule A1 à la cellule M2, plage notée A1 :M2), puis lancer l'assistant graphique (icône Diagramme) où on sélectionne le type de diagramme « Colonne ». Déclarer les données « en ligne » et la « Première ligne comme étiquette ».



c. Coefficients binomiaux

PROPRIÉTÉ 1.8 En notant $n!$ (prononcer « factorielle n ») le nombre $n \times (n-1) \times \ldots \times 2 \times 1$, et en posant $0! = 1$, les nombres $\binom{n}{k}$ se calculent à l'aide de la formule

$$\binom{n}{k} = \frac{n!}{k! \times (n-k)!}$$

DÉMONSTRATION Pour dénombrer les chemins menant à k succès parmi n étapes, il suffit de choisir le premier succès parmi les n étapes, le 2e succès parmi les $n-1$ étapes restantes, etc. jusqu'au ke succès que l'on choisit parmi les $n-k+1$ étapes restantes. Il y a ainsi $n \times (n-1) \times \ldots \times (n-k+1)$ choix possibles. Ce faisant, on a compté plusieurs fois chacune des possibilités. Par exemple, si on doit dénombrer les chemins qui mènent à 3 succès sur 5 étapes, on va compter, selon ce principe, les six triplets (1,2,3), (1,3,2), (2,1,3), (2,3,1), (3,2,1) et (3,1,2) alors qu'ils représentent un seul et même choix. De même pour tous les triplet-solution.
Par conséquent, il faut diviser le nombre calculé par le nombre d'ordres différents des k succès. Ce nombre d'ordres est $k! = k \times (k-1) \times \ldots \times 2 \times 1$ car on peut mettre le premier élément à n'importe laquelle des k places, le 2e élément à n'importe laquelle des $k-1$ places restantes, etc. jusqu'au dernier élément qui va forcément dans la seule place restante. Dans notre exemple, les $k = 3$ éléments peuvent se permuter de $3! = 3 \times 2 \times 1 = 6$ façons différentes, comme on l'a vu.
Finalement, le nombre cherché est égal à $n \times (n-1) \times \ldots \times (n-k+1) \div k!$, quotient qui se simplifie en utilisant le symbole !, car $n \times (n-1) \times \ldots \times (n-k+1) = \frac{n!}{(n-k)!}$. Ainsi on obtient $\binom{n}{k} = \frac{n!}{k! \times (n-k)!}$.

Remarque :
Le nombre $\binom{n}{k}$ que l'on vient de calculer a une utilité dans bien d'autres domaines que la loi binomiale. D'une façon générale, on dit que ce que l'on dénombre ici est le nombre de parties à k éléments distincts (sans répétition) choisis parmi n éléments. On a déjà dit que ce dénombrement s'applique aux développement du binome $(a+b)^n$ car, lorsqu'on doit calculer le coefficient du terme $a^k b^{n-k}$, on dénombre des parties à k éléments parmi n (les k facteurs où on prend le nombre a). C'est exactement la même situation que pour la loi binomiale où les parties à k éléments parmi n sont les places des k succès. Voici d'autres exemples où ces coefficients trouvent une utilité :

- Une assemblée de 20 personnes doit élire un bureau (un président, un secrétaire et un trésorier). Combien de bureaux différents peut-on former ? Il s'agit de former une partie à 3 éléments choisis parmi 20. Leur nombre est donc de $\binom{20}{3} = \frac{20!}{3! \times 17!}$. Remarquons au passage la taille du nombre $20! = 2\,432\,902\,008\,176\,640\,000$. Cela montre que, si la formule obtenue avec les factorielles est plus simple, elle peut poser des problème de calcul (des problèmes d'arrondis ou de dépassement de capacité). Il est préférable d'utiliser la formule initiale : $\binom{20}{3} = \frac{20 \times 19 \times 18}{3 \times 2 \times 1} = 1140$. On peut donc former 1140 bureaux différents.
- Dans l'ancienne formule du Loto, on devait choisir 6 numéros d'une grille en contenant 49. Quelle est la probabilité de gagner le gros lot ? Il y a une seule grille gagnante. Combien y a t-il de grilles différentes de Loto ? Une grille est une partie à 6 éléments choisis parmi 49. Leur nombre est donc de $\binom{49}{6} = \frac{49!}{6! \times 43!}$. On se retrouve confronté au problème mentionné ci-dessus, avec des nombres encore plus grands :

$$49! = 608\,281\,864\,034\,267\,560\,872\,252\,163\,321\,295\,376\,887\,552\,831\,379\,210\,240\,000\,000\,000$$

  Avec la formule initiale, on trouve $\binom{49}{6} = \frac{49 \times 48 \times 47 \times 46 \times 45 \times 44}{6 \times 5 \times 4 \times 3 \times 2 \times 1} = 13\,983\,816$. Comme chaque grille est équiprobable, il y avait une chance sur 13 983 816 de gagner le gros lot en misant sur une grille de l'ancienne formule du Loto.

Propriété 1.9 Pour tout entier n, on a :

1. $\binom{n}{0} = 1$ et $\binom{n}{1} = n$
2. si $k \leqslant n$ alors $\binom{n}{k} = \binom{n}{n-k}$
3. si $k < n$ alors $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$

Démonstration
1. Il y a un seul chemin qui conduise à 0 succès : celui qui est constitué de n échecs. Si on ne veut qu'un seul succès, il suffit de choisir sa place parmi les n places possibles pour ce succès.
2. Il y a autant de chemins qui mènent à k succès que de chemins qui mènent à k échecs (c'est-à-dire à $n-k$ succès). On l'a déjà dit, les qualificatifs de succès et échec sont relatifs ; en permutant les mots, on ne change pas le nombre de leurs occurrences.
3. Le coefficient $\binom{n+1}{k+1}$ est le nombre de chemins qui mènent à $k+1$ succès dans une répétition de $n+1$ épreuves de Bernoulli. À l'issue des n premières épreuves, il n'y a que trois possibilités :
   - on a déjà eu $k+1$ succès. Il y a $\binom{n}{k+1}$ possibilités auxquelles il suffit d'un échec à la dernière étape
   - on a déjà eu k succès. Il y a $\binom{n}{k}$ possibilités auxquelles il suffit d'un succès à la dernière étape
   - dans tous les autres cas (moins de k succès ou plus de $k+1$ succès), on ne peut aboutir à $k+1$ succès

   Finalement, le nombre de chemins cherché est la somme des dénombrements associés aux deux premières possibilités : $\binom{n}{k} + \binom{n}{k+1}$.

   Remarquons qu'il est très simple de prouver cette relation algébriquement. Il suffit en effet d'écrire les deux expressions avec le même dénominateur, en remarquant que $(n-(k+1)) = (n-k-1)$ est l'entier qui précède $(n-k)$:

$$\binom{n}{k} + \binom{n}{k+1} = \frac{n!}{k!(n-k)!} + \frac{n!}{(k+1)!(n-(k+1))!} = \frac{n!(k+1)}{(k+1)!(n-k)!} + \frac{n!(n-k)}{(k+1)!(n-k)!}$$

$$= \frac{n!(k+1)+n!(n-k)}{(k+1)!(n-k)!} = \frac{n!(k+1+n-k)}{(k+1)!(n-k)!} = \frac{(n+1)!}{(k+1)!(n-k)!} = \binom{n+1}{k+1}$$

Le triangle de Pascal[1] exploite ces propriétés pour le calcul des coefficients $\binom{n}{k}$ de proche en proche. Chaque nombre est obtenu en additionnant deux nombres de la ligne du dessus (propriété 5.9.3) : celui de la même colonne et celui de la colonne précédente comme le montre sur fond jaune le $\binom{5}{2} = 10$ qui est égal à la somme $\binom{4}{1} + \binom{4}{2} = 4 + 6$. Il suffit d'ajouter à cette propriété la 1re colonne de 1 qui vient de la propriété 5.9.1 pour construire le triangle en entier.

n \ k	0	1	2	3	4	5	...
0	1						
1	1	1					
2	1	2	1				
3	1	3	3	1			
4	1	4	6	4	1		
5	1	5	10	10	5	1	
$\vdots$	$\vdots$						$\ddots$

1. Blaise Pascal (1623-1662) : philosophe et mathématicien français. Le triangle qui porte son nom était connu au Moyen-âge, par des mathématiciens persans et chinois et plus tard en Europe, au XVIe siècle, par Tartaglia et Stiefel. Pascal en étudia les propriétés dans son « Traité du triangle arithmétique (1654).

d. Prise de décision

Propriété 1.10 (Intervalle de fluctuation) On considère une population sur laquelle on étudie la proportion p d'individus ayant un caractère C avec $p \in [0,2;0,8]$ et dans laquelle on prélève un échantillon de taille $n \geqslant 25$.
Dans ces conditions, la fréquence f des individus de l'échantillon ayant le caractère C fluctue, dans 95% des cas, à l'intérieur de l'intervalle $\left[p-\frac{1}{\sqrt{n}};p+\frac{1}{\sqrt{n}}\right]$, dit intervalle de fluctuation.

Remarques :
Cela signifie qu'en choisissant un échantillon de taille n au hasard, la probabilité de l'événement « $f \in \left[p-\frac{1}{\sqrt{n}};p+\frac{1}{\sqrt{n}}\right]$ » est supérieure ou égale à $0,95$ sous réserve de satisfaire les conditions sur n et p. Reformulons cette situation dans le contexte de la loi binomiale : le tirage au hasard d'un individu dans une population qui peut présenter un caractère C avec une probabilité p est assimilable à une épreuve de Bernoulli de paramètre p où le succès correspond à « avoir le caractère C ». Le prélèvement d'un échantillon de taille n dans cette population s'assimile alors à un schéma de Bernoulli de paramètres n et p et la variable aléatoire X qui compte le nombre de succès suit la loi binomiale $\mathscr{B}(n,p)$. La variable aléatoire $F=\frac{X}{n}$ correspond à la fréquence du succès.

Définition 1.9 Soit X une variable qui suit une loi binomiale $\mathscr{B}(n,p)$ et $F=\frac{X}{n}$ la variable aléatoire qui représente la fréquence du succès. Un intervalle de fluctuation de F au seuil de 95% est un intervalle de la forme $\left[\frac{a}{n};\frac{b}{n}\right]$ où a et b sont des entiers compris entre 0 et n, tels que

$$P(\frac{a}{n} \leqslant F \leqslant \frac{b}{n}) \geqslant 0,95$$

Dans la pratique :
La condition $P(\frac{a}{n} \leqslant F \leqslant \frac{b}{n}) \geqslant 0,95$ équivaut à $P(a \leqslant X \leqslant b) \geqslant 0,95$. Afin d'obtenir l'intervalle de fluctuation de plus petite amplitude, on cherche les plus petits entiers a et b tels que :

$$P\left(X \leqslant a\right) \geqslant 0,025 \text{ et } P\left(X \leqslant b\right) \geqslant 0,975$$

Ces conditions satisfaites, on aura bien $P(a \leqslant X \leqslant b) \geqslant 0,95$ car
$P\left(X \leqslant b\right) \geqslant 0,975 \Longrightarrow P\left(X > b\right) < 0,025$ et $P\left(X \leqslant a\right) \geqslant 0,025 \Longrightarrow P\left(X < a\right) < 0,025$.
Comme $P\left((X<a)\cup(X>b)\right) = P\left(X<a\right)+P\left(X>b\right) < 0,025+0,025 = 0,05$,
on a $P(a \leqslant X \leqslant b) = P\left(\overline{(X<a)\cup(X>b)}\right) = 1-P\left((X<a)\cup(X>b)\right) \geqslant 0,95$.

Exemple 10 – On constitue des sachets de 100 bonbons `M&M's` à partir d'un mélange contenant des bonbons des 6 couleurs en proportion égale ($f=\frac{1}{6}$ pour chaque couleur). Á l'aide du tableur, on veut déterminer un intervalle de fluctuation au seuil de 95% de la fréquence d'une couleur, disons orange, dans un sachet. Le nombre X de bonbons de couleur orange suit la loi binomiale $\mathscr{B}(100;\frac{1}{6})$. La fréquence des bonbons de couleur orange est donné par la variable aléatoire $F=\frac{X}{100}$. On cherche deux entiers a et b tels que $P(\frac{a}{100} \leqslant F \leqslant \frac{b}{100}) \geqslant 0,95$ autrement dit tels que $P(a \leqslant X \leqslant b) \geqslant 0,95$.
Ci-dessous une copie d'écran du tableur avec les valeurs prise par la variable X, les probabilités $P(X=k)$ et les probabilités cumulées $P(X \leqslant k)$. La partie colorée en orange montre que :

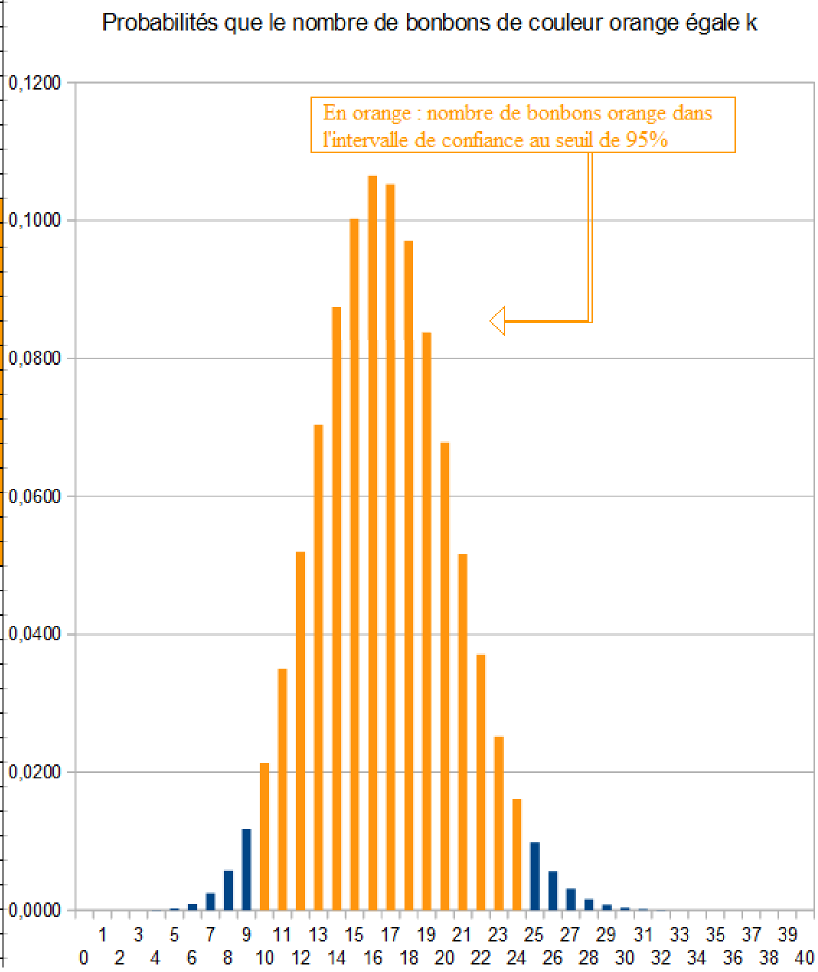
- $P(X \leqslant 10) \geqslant 0,025$ et $a=10$ est la 1re valeur de k où $P(X \leqslant k) \geqslant 0,025$ se produit
- $P(X \leqslant 24) \geqslant 0,975$ et $b=24$ est la 1re valeur de k où $P(X \leqslant k) \geqslant 0,975$ se produit

Donc $P(10 \leqslant X \leqslant 24) \geqslant 0,95$ ou encore $P(\frac{10}{100} \leqslant F \leqslant \frac{24}{100}) \geqslant 0,95$ et donc $\left[\frac{10}{100};\frac{24}{100}\right]$ est un intervalle de fluctuation au seuil de 95% de la fréquence de la couleur orange dans les sachets de 100 `M&M's`. Le risque de voir la fréquence de la couleur orange sortir de cet intervalle est inférieur à 5%. Un échantillon de 100 bonbons contenant moins de 10 bonbons de couleur orange ou plus de 24 bonbons de couleur orange ne serait pas représentatif de cette distribution au seuil de 95%.

Comparons ce résultat à l'intervalle de fluctuation prévu par la propriété 5.10 :
$IF_{prop} = \left[\frac{1}{6}-\frac{1}{\sqrt{100}};\frac{1}{6}+\frac{1}{\sqrt{100}}\right] = \left[\frac{1}{15};\frac{4}{15}\right] \approx [0,0667;0,2667]$. L'intervalle de fluctuation donné par la

loi binomiale est $IF_{binom} = [0,1; 0,24]$. On constate que IF_{binom} a une amplitude plus faible que IF_{prop} et que $IF_{binom} \subset IF_{prop}$: les fluctuations prévues par la loi binomiale sont moins dispersées que celles que donne la propriété 5.10.

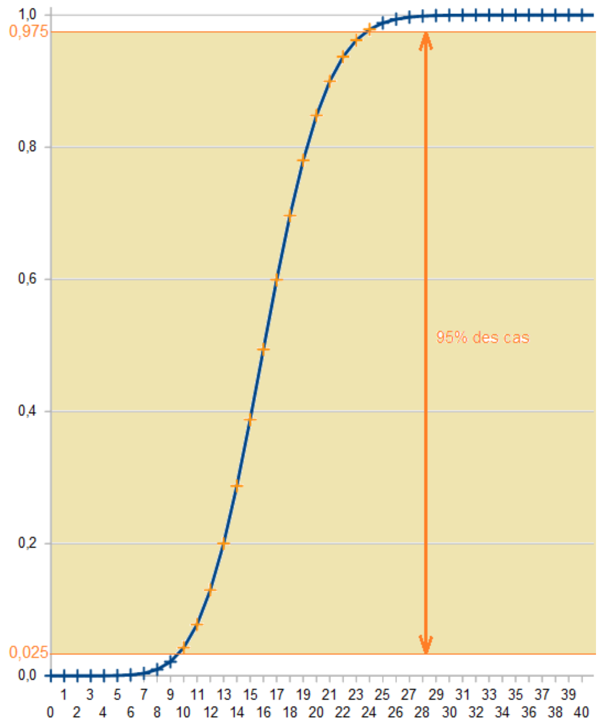
k	P(X=k)	P(X<=k)
0	0,0000	0,0000
1	0,0000	0,0000
2	0,0000	0,0000
3	0,0000	0,0000
4	0,0001	0,0001
5	0,0003	0,0004
6	0,0009	0,0013
7	0,0025	0,0038
8	0,0058	0,0095
9	0,0118	0,0213
10	0,0214	0,0427
11	0,0350	0,0777
12	0,0520	0,1297
13	0,0703	0,2000
14	0,0874	0,2874
15	0,1002	0,3877
16	0,1065	0,4942
17	0,1052	0,5994
18	0,0971	0,6965
19	0,0838	0,7803
20	0,0679	0,8481
21	0,0517	0,8998
22	0,0371	0,9369
23	0,0252	0,9621
24	0,0162	0,9783
25	0,0098	0,9881
26	0,0057	0,9938
27	0,0031	0,9969
28	0,0016	0,9985
29	0,0008	0,9993
30	0,0004	0,9997
31	0,0002	0,9999
32	0,0001	1,0000
33	0,0000	1,0000
34	0,0000	1,0000
35	0,0000	1,0000
36	0,0000	1,0000
37	0,0000	1,0000
38	0,0000	1,0000
39	0,0000	1,0000
40	0,0000	1,0000



La représentation cumulative des probabilités associées à la loi binomiale $\mathscr{B}(100; \frac{1}{6})$ est donnée ci-contre. Il s'agit de la représentation de la fonction dont les valeurs sont fournies par le tableau ci-dessus :

$$k \mapsto P(X \leqslant k)$$

Les droites d'équation $y = 0,025$ et $y = 0,975$ correspondant aux valeurs de coupure $2,5\%$ et $97,5\%$, délimitent une zone (colorée sur le graphique) qui contient exactement 95% de la population. Du fait que les nombres a et b sont des entiers, les valeurs de coupure retenues sont légèrement décalées et le pourcentage de la population contenu entre ces valeurs dépasse toujours légèrement 95%. Ici, la somme des probabilités de $P(X = 10)$ à $P(X = 24)$ est environ égale à $95,7\%$.



Règle de décision :
Dans une population, on ne connaît pas toujours avec exactitude la proportion de la population ayant un caractère C mais on peut faire une hypothèse sur cette proportion, c'est-à-dire se fixer une valeur p de cette proportion, et soumettre cette hypothèse au test d'échantillonnage. Notons f la fréquence observée dans un échantillon de taille n de cette population. Comme on sait que 95% des échantillons doivent avoir une fréquence comprise dans l'intervalle de fluctuation IF déterminé à partir de n et p, on peut établir la règle de décision suivante :

> Si la fréquence f de l'échantillon n'appartient pas à IF, on rejette l'hypothèse faite sur p avec un risque d'erreur de 5%. Dans le cas contraire, on affirme que l'hypothèse est crédible car compatible avec les données d'échantillonnage.

Application : cette règle examine si la variabilité d'un échantillonnage peut s'expliquer par le simple fait du hasard. Lorsque la fréquence f mesurée dans l'échantillon sort de l'intervalle de fluctuation, il faut chercher une autre explication que le hasard pour justifier l'écart entre la fréquence observée et la probabilité hypothétique p attribuée au caractère C.
Si on trouve qu'un sachet de 100 bonbons `M&M's` ne contient que 7 bonbons oranges ou 25 bonbons rouges, d'après ce qui a été vu plus haut, il faut rejeter l'hypothèse que les 6 couleurs sont équiprobables dans ce genre de sachet. La simple fluctuation due au hasard ne permet pas d'expliquer ces valeurs. En affirmant cela, on assume un risque d'erreur de 5%.
Avec les mêmes données numériques, pou déterminer si un dé cubique est bien équilibré on peut faire l'expérience de le tirer 100 fois et de noter le nombre de fois où l'on a obtenu chacune des faces. Selon notre étude, si une quelconque des 6 faces est sortie plus de 24 fois ou moins de 10 fois, alors on trouver ce résultat « anormal » au seuil de 95% et rejeter l'hypothèse d'équiprobabilité des faces.

Le coin du chercheur
❊ Influence du seuil sur la prise de décision

➩ Le seuil de 95% choisi pour accepter ou rejeter une hypothèse n'a pas une valeur objective. On peut modifier ce seuil et examiner une hypothèse avec un seuil plus sévère (par exemple 99%) ou plus souple (par exemple 90%), selon les besoins de l'étude. La méthode étant la même pour les différents seuils envisagés, mettre au point un algorithme qui traite le problème.

- en entrée : la taille de l'échantillon n, la probabilité hypothétique du caractère étudié p , la fréquence f du caractère dans l'échantillon et le seuil s
- en sortie : les valeurs de coupure a et b (et par voie de conséquence les bornes de l'intervalle de fluctuation) ainsi que la réponse décisionnelle (« rejeté » ou « accepté »)

➩ Tester cet algorithme sur une situation concrète, par exemple celle-ci : En France $6,5\%$ des patients sont atteints d'une maladie nosocomiale lors d'une hospitalisation. On réalise une enquête dans les deux hôpitaux d'une ville. Dans l'hôpital A, on a déterminé $8,5\%$ d'infections nosocomiales pour 436 patients et dans l'hôpital B, il y a $8,1\%$ d'infections nosocomiales pour 5612 patients. La situation parait-elle « normale » dans ces deux hôpitaux ?

➩ On peut imaginer le résultat d'une étude décisionnelle : au seuil s_1 l'hypothèse doit être rejetée et au seuil $s_2 > s_1$ l'hypothèse reste compatible avec les données. Dans tous les cas, il existe un seuil critique s^* pour lequel il est difficile de trancher, la fréquence dans l'échantillon étant une des bornes de l'intervalle de fluctuation. Écrire un algorithme qui détermine le seuil critique s^*.

Méthode (Fiche de révision)

Définitions et propriétés à connaître par ♥

➲ Définitions

variable aléatoire	à chaque issue ω_i de Ω est associé un réel $X(\omega_i) = k$
loi de X	Les probabilités associées aux différentes valeurs de X constituent la « loi de probabilités » de la variable X
rang	place dans la suite : rang(u_{n_0})=1 $\Longrightarrow$ rang(u_n)=$n+1-n_0$
déf. par récurrence	u_n est défini par une relation avec u_{n-1} et un terme initial u_0
déf. explicite	u_n est connu directement en fonction de n
constante	à partir d'un certain rang $N \geqslant n_0$, on a $\forall n \geqslant N,\ \ u_{n+1} = u_n$
croissante	à partir d'un certain rang $N \geqslant n_0$, on a $\forall n \geqslant N,\ \ u_{n+1} \geqslant u_n$
majorée	à partir d'un certain rang $N \geqslant n_0$, $\exists M \in \mathbb{R},\ \ \forall n \geqslant N,\ \ u_n \leqslant M$
monotone, bornée	monotone : croissante ou décroissante ; bornée : majorée et minorée
périodique	$\exists p \in \mathbb{N},\ \ \forall n \geqslant N,\ \ u_{n+p} = u_n$; p : période
convergente	admet une limite l quand n tend vers$+\infty$
divergente	n'admet pas de limite finie quand n tend vers$+\infty$
limite	admet une limite finie ou infinie
arithmétique	différence $u_{n+1} - u_n$ constante (raison r) ; suite définie par u_0 et r
géométrique	quotient $\frac{u_{n+1}}{u_n}$ constant (raison q) ; suite définie par u_0 et q

➲ Propriétés

croissante	$u_{n+1} - u_n \geqslant 0$ ou, u_n restant positif, $\frac{u_{n+1}}{u_n} \geqslant 1$
convergente	suite croissante et majorée ou décroissante et minorée
	gendarmes : $v_n \leqslant u_n \leqslant w_n,\ \ \lim v_n = \lim w_n = l \Longrightarrow \lim u_n = l.$
	références : $(\frac{1}{\sqrt{n}})$, $(\frac{1}{n})$, $(\frac{1}{n^2})$, etc. convergent vers 0
divergente vers $+\infty$	suite croissante et non majorée ou décroissante et non minorée
	$u_n \geqslant v_n,\ \ \lim v_n = +\infty \Longrightarrow \lim u_n = +\infty$
	références : $(\sqrt{n})$, (n), (n^2), etc. divergent vers $+\infty$
arithmétique	terme général : $u_n = u_0 + nr$ ou $u_n = u_p + (n-p)r$
	somme : $\sum\limits_{k=0}^{n} u_k = (n+1) \times \frac{2u_0+nr}{2}$ ou $\sum\limits_{k=p}^{p+n-1} u_k = n \times \frac{u_p+u_{p+n-1}}{2}$
	a, b et c consécutifs d'une suite arithm. $\Longleftrightarrow b = \frac{a+c}{2}$
géométrique	terme général : $u_n = u_0 \times q^n$ ou $u_n = u_p \times q^{n-p}$
	somme : $\sum\limits_{k=0}^{n} u_k = u_0 \times \frac{1-q^{n+1}}{1-q}$ ou $\sum\limits_{k=p}^{p+n-1} u_k = n \times \frac{1-q^n}{1-q}$
	a, b et c consécutifs d'une suite géom. $\Longleftrightarrow b^2 = ac$