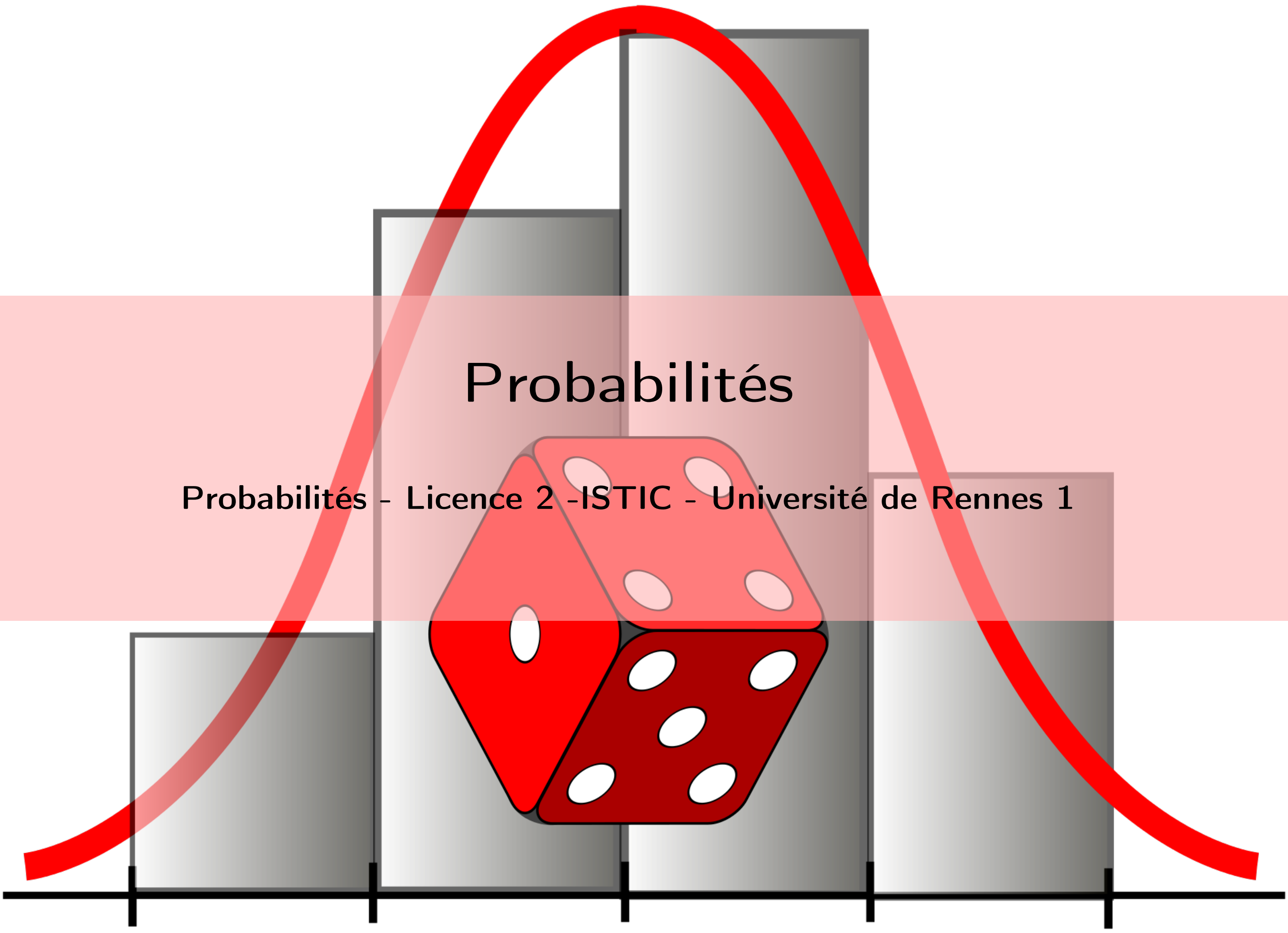
Probabilités

Probabilités - Licence 2 -ISTIC - Université de Rennes 1



Ce polycopié est basé sur un polycopié fait par Bert Wiest :
`https://perso.univ-rennes1.fr/stephane.leborgne/PS1Cours.pdf`

Le template utilisé pour le rendu du polycopié est basé sur le Legrand Orange Book disponible sur le web à l'adresse suivante :
`https://www.latextemplates.com/template/the-legrand-orange-book`

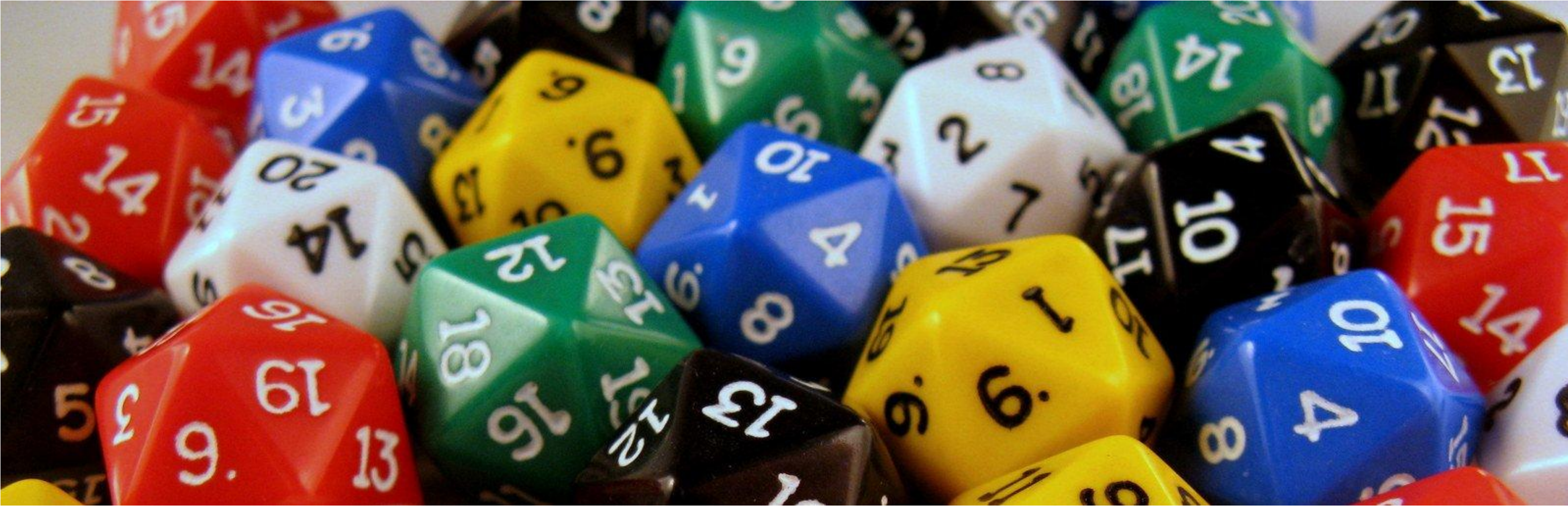
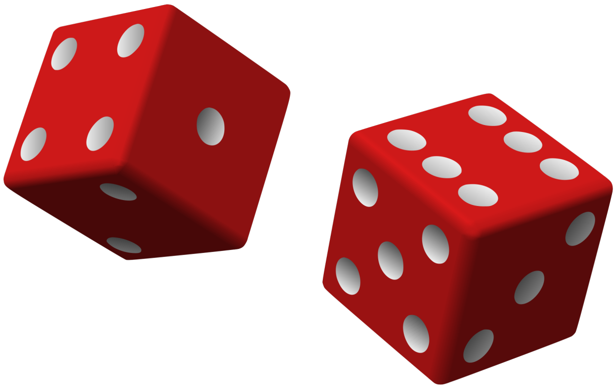


Table des matières



1. Ensembles et dénombrements

1 Univers, événement, manipulation ensembliste

Notation. *(Modélisation d'une expérience aléatoire)*
Pour une expérience (ou "épreuve") aléatoire donnée, nous noterons Ω l'ensemble (nous dirons **l'"Univers"***) de tous les résultats possibles de l'expérience. Un "***événement***" est par définition un sous-ensemble de Ω.*

Exemple 1.1 Je tire une carte au hasard d'un jeu de 32 cartes. Alors :

$$\Omega = \{\text{7 de pique, 8 de pique, ... , as trèfle}\} \qquad \text{(32 éléments)}.$$

L'événement "10 d'une couleur quelconque" a quatre éléments :

$$A = \{\text{10 de pique, 10 de coeur, 10 de carreau, 10 de trèfle}\}.$$

On a bien : $A \subset \Omega$.

Exemple 1.2 Je jette deux fois de suite un dé. Alors

$$\begin{aligned}\Omega &= \{(1,1),(1,2),\ldots,(1,6),(2,1),(2,2),\ldots\ldots,(6,5),(6,6)\} \\ &= \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\} \qquad (6 \cdot 6 = 36 \text{ éléments}) \\ &= \{1,2,3,4,5,6\}^2\end{aligned}$$

L'événement "deux fois le même nombre"$= \{(1,1),\ldots,(6,6)\}$ a 6 éléments.

Notation. $\forall$ = *pour tout (quelque soit),* $\exists$ = *il existe.*

Notation. *Pour un ensemble Ω, on note :*

— $\mathscr{P}(\Omega)$ *l'ensemble des parties* (l'ensemble des sous-ensembles) *de Ω. C.à.d., chaque élément de $\mathscr{P}(\Omega)$ est un sous ensemble de Ω. (Par exemple, si $\Omega = \{A,B,C\}$, alors*

$$\mathscr{P}(\Omega) = \{\emptyset, \{A\}, \{B\}, \{C\}, \{A,B\}, \{A,C\}, \{B,C\}, \{A,B,C\}\})$$

Remarque : si Ω a n éléments, alors $\mathscr{P}(\Omega)$ a 2^n éléments.

— *Complémentaire : si $A \subset \Omega$, on note $A^c = \{x \in \Omega \mid x \notin A\}$ le* complémentaire.

— *Différence : si $A \subset \Omega$ et $B \subset \Omega$, alors $A \setminus B = \{x \in A \mid x \notin B\}$.*

— *La différence symétrique : si $A \subset \Omega$ et $B \subset \Omega$, alors on note*

$$A \Delta B = (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$$

la différence symétrique.

Proposition 1.3 (Propriétés de ces opérations)

1. $A \cup B = B \cup A$.
2. $A \cap B = B \cap A$.
3. $(A \cup B) \cup C = A \cup (B \cup C)$ donc on peut écrire $A \cup B \cup C$.
4. même chose pour $\cap$.
5. $(A^c)^c = A$.
6. $(A \cup B)^c = A^c \cap B^c$.
7. $(A \cap B)^c = A^c \cup B^c$.

Démonstration. Admis (mais plus ou moins évident...). ■

Exercice Donner un exemple qui montre qu'en général, $(A \setminus B) \setminus C \neq A \setminus (B \setminus C)$ ■

Définition 1.4 (Produit cartésien) Si A et B sont deux ensembles, alors $A \times B$ est l'ensemble des couples ordonnés :

$$A \times B = \{(x, y) \mid x \in A, y \in B\}$$

Exemple 1.5 Si je jette un dé *et* joue à pile ou face, alors j'aurai l'univers Ω

$$\Omega = \{1, \ldots, 6\} \times \{P, F\}$$

avec $6 \cdot 2 = 12$ éléments : $\Omega = \{(1,P), (2,P), \ldots, (6,P), (1,F), \ldots, (6,F)\}$.

Définition 1.6 Soit Ω un ensemble. Si Ω n'a qu'un nombre fini d'éléments, alors ce nombre est appelé le *cardinal* de Ω, et noté $\mathrm{Card}(\Omega) = \#\Omega = |\Omega|$.

Proposition 1.7 (Propriétés du cardinal)

1. $\mathrm{Card}(\emptyset) = 0$
2. Si $A \subset B$ alors $\mathrm{Card}(A) \leqslant \mathrm{Card}(B)$.
3. Si $A \subset \Omega$, alors $\mathrm{Card}(A^c) = \mathrm{Card}(\Omega) - \mathrm{Card}(A)$.
4. Si $A, B \subset \Omega$ sont disjoints ($A \cap B = \emptyset$), alors $\mathrm{Card}(A \cup B) = \mathrm{Card}(A) + \mathrm{Card}(B)$.
5. En général, $\mathrm{Card}(A \cup B) = \mathrm{Card}(A) + \mathrm{Card}(B) - \mathrm{Card}(A \cap B)$.
6. $\mathrm{Card}(A \setminus B) = \mathrm{Card}(A) - \mathrm{Card}(A \cap B)$.
7. Si A et B sont deux ensembles finis, alors $\mathrm{Card}(A \times B) = \mathrm{Card}(A) \cdot \mathrm{Card}(B)$.

Démonstration. Admis. ■

Exemple 1.8 Parmi 30 étudiants (E), 20 parlent Anglais (A), 15 parlent Chinois (C), et 10 les Deux (D). Combien d'étudiants ne parlent ni anglais ni chinois ?

On cherche :

$$\begin{aligned}
\mathrm{Card}\big((A \cup C)^c\big) &= \mathrm{Card}(E) - \mathrm{Card}(A \cup C) \\
&= 30 - (\mathrm{Card}(A) + \mathrm{Card}(C) - \mathrm{Card}(D)) \\
&= 30 - (20 + 15 - 10) \\
&= 5
\end{aligned}$$

2 Quelques exemples fondamentaux de cardinaux

2.A. Factoriel

La notation suivante est bien utile :

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1 \qquad \text{prononcé "factoriel } n \text{ ou } n \text{ factoriel"}$$

Par convention, on pose $0! = 1$. Les premières valeurs de factoriel sont :

n	$n!$
0	1
1	1
2	2
3	6
4	24
5	120
6	720

2.B. Arrangements avec répétition

Exemple 1.9 (Nombre de combinaisons d'un antivol) Le nombre de combinaisons possibles pour un antivol à code à 4 chiffres est $10 \times 10 \times 10 \times 10 = 10^4$. En effet, il y a 10 choix pour le premier chiffre (0, 1, 2, ..., 9), 10 pour le second, encore 10 pour le troisième et enfin 10 pour le quatrième.

En terme plus théorique :

Proposition 1.10 Pour tous entiers $n,k \geqslant 1$. Le nombre de manières de choisir k objets parmi n objets, chaque objet pouvant être choisi plusieurs fois et l'ordre du choix étant pris en compte est n^k.

On parle du nombre d'*arrangements avec répétition* (ou avec remise).

2.C. Arrangement sans répétition

Exemple 1.11 (Nombre de quintés) Pendant une course hippique où 10 chevaux sont en compétitions. Combien y a-t-il de quintés possibles ? Réponse : $\underbrace{10 \times 9 \times 8 \times 7 \times 6}_{5 \text{ termes}} = 30\,240$.

En terme plus théorique :

Proposition 1.12 Pour tous entiers $n,k \geqslant 1$ avec $k \leqslant n$. Le nombre de manières de choisir k objets parmi n objets, chaque objet ne pouvant être choisi qu'au plus une fois et l'ordre du choix étant pris en compte est $A_n^k = \underbrace{n \times (n-1) \times (n-2) \times \cdots \times (n-k+1)}_{k \text{ termes}} = \frac{n!}{(n-k)!}$.

On parle du nombre d'*arrangements sans répétition* (ou sans remise).

Démonstration. On a :

— n choix pour le 1^{er} objet.

— $n-1$ choix pour le 2^{eme} objet.

— $n-2$ choix pour le 3^{eme} objet.

— $\vdots$

— $n-j$ choix pour le $(j+1)^{\text{eme}}$ objet.

— $\vdots$

— $n-(k-1)$ choix pour le k^{eme} objet.

On obtient ainsi :

$$A_n^k = n \times (n-1) \times (n-2) \times \cdots \times (n-k+1)$$

Enfin :

$$\frac{n!}{(n-k)!} = \frac{n \times (n-1) \times \cdots \times (n-k+1) \times (n-k) \times (n-k-1) \times \cdots \times 2 \times 1}{(n-k) \times (n-k-1) \times \cdots \times 2 \times 1} = A_n^k$$

■

R Soit Ω un ensemble avec n éléments. Combien de permutations des éléments de Ω y a-t-il, c.à.d. combien y a-t-il de listes ordonnées contenant tous les éléments de Ω (sans répétition) ? Ou encore combien y a-t-il d'arrangements sans répétitions de n objets parmi n objets ?
Réponse : $A_n^n = n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 2 \cdot 1 = \frac{n!}{0!}$.

2.D. Combinaisons sans répétition

Exemple 1.13 Combien y a-t-il de grilles de loto ? Au loto, on choisit 6 nombres parmi 49, l'ordre n'a pas d'importance et on ne peut pas choisir deux fois le même nombre.

Réponse : $\frac{49\times48\times47\times46\times45\times44}{6\times5\times4\times3\times2\times1} = 13\,983\,816$.

Explication : On va noter C_{49}^{6} le nombre de grilles de loto. On va commencer par compter le nombre d'arrangements sans répétition mais d'une autre façon que celle de la section précédente.

On compte le nombre de façons de choisir 6 objets parmi 49 en tenant compte de l'ordre. On choisit tout d'abord 6 objets parmi 49 sans tenir compte de l'ordre, on a C_{49}^{6} choix, puis on ordonne les 6 objets, il y a 6! façons de faire, on a donc :

$$A_{49}^{6} = C_{49}^{6} \times 6!$$

Proposition 1.14 Pour tous entiers $n, k \geqslant 0$ avec $k \leqslant n$. Le nombre de manières de choisir k objets parmi n objets, chaque objet ne pouvant être choisi qu'au plus une fois et l'ordre du choix n'étant pris en compte est $C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!}$.

On parle du nombre de *combinaisons sans répétition* (ou sans remise).

R Le nombre de combinaisons sans répétitions de k objets parmi n objets compte aussi le nombre de sous-ensembles à k éléments d'un ensemble à n éléments.

2.E. Résumé

On résume avec des urnes et des boules.

Théorème 1.15 Étant donnée une urne avec n boules, étiquetées $1, \ldots, n$. On tire k boules. Il y a 4 façons de procéder.

— Deux possibilités : sans ou avec remise des boules entre deux tirages.

— Deux sous-possibilités : quand on note les résultats, on peut tenir compte de l'ordre d'apparition des nombres (regarder les "arrangements") ou l'ignorer ("combinaisons").

Le nombre de résultats possibles est dans le tableau suivant :

	Sans remise	Avec remise
Attention à l'ordre (arrangements)	$A_n^k = \dfrac{n!}{(n-k)!} = \underbrace{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)}_{k \text{ termes}}$	$n^k = \underbrace{n\cdot n\cdot\ldots\cdot n}_{k \text{ termes}}$
En ignorant l'ordre (combinaisons)	$C_n^k = \binom{n}{k} = \dfrac{n!}{(n-k)!k!}$	$\Gamma_n^k := C_{n+k-1}^{n-1} = C_{n+k-1}^{k} = \dfrac{(n+k-1)!}{k!(n-1)!}$

Exemple 1.16 (Des exemples de ses 4 modes de tirage dans la vie courante)

1. Nombre de manières de choisir ses Top 10 ($p=10$) parmi n chansons $= A_n^{10}$.
2. Nombre de textes possibles de longueur k (alphabet avec n lettres) $= n^k$.
3. Nombre de résultats du loto 6 sur 49 (ou p parmi n) $= C_{49}^6$.
4. Nombre de distributions possibles des voix quand k électeurs votent pour n candidats $= \Gamma_n^k$.

Il ne reste que la formule pour Γ_n^k à démontrer.

Démo et résultat hors-programme. Combinaisons avec répétition : On veut démontrer que le nombre de combinaisons avec répétition est de C_{n+k-1}^k. Dessinons $n+k-1$ cercles. Par exemple, si $n=4$ et $k=9$, on dessine 12 cercles. Supposons que le tirage a donné j_1 fois la boule 1, ... , j_n fois la boule n (avec $j_1+\ldots+j_n=k$). Cela me donne une façon de choisir $n-1$ cercles parmi les $n+k-1$ de la façon suivante : j'ignore les j_1 premières boules, je sélectionne la j_1+1ème, ignore encore j_2 boules, sélectionne la suivante, etc. Par exemple, si $j_1=2$, $j_2=4$, $j_3=0$, $j_4=3$, je dessine

$$\bigcirc\bigcirc\otimes\bigcirc\bigcirc\bigcirc\bigcirc\otimes\otimes\bigcirc\bigcirc\bigcirc$$

Le résultat du tirage est uniquement déterminé (en ignorant l'ordre) par le choix des $n-1$ cercles, et chaque choix de $n-1$ cercles peut s'obtenir de cette façon. Il y a C_{n+k-1}^{n-1} différents choix de $n-1$ cercles, et donc C_{n+k-1}^{n-1} combinaisons. ■

3 Propriétés des coefficients binomiaux

Définition 1.17 Les nombres C_n^k (= # façons de choisir k éléments parmi n) s'appellent les *coefficients binomiaux*. Autre notation : $\binom{n}{k}$ prononcé "k parmi n".

$$\begin{aligned}\binom{n}{k} &= \text{Nb. de résultats possibles lors d'un tirage de } k \text{ boules parmi } n \text{ boules distinctes, sans remise.}\\ &= \text{Nb. de sous-ensembles à } k \text{ éléments d'un ensemble à } n \text{ éléments.}\\ &= \frac{n!}{k!(n-k)!}\end{aligned}$$

Exemple 1.18 — Nombres de full au poker. Si on distribue 5 cartes au hasard d'un jeu de 52 cartes à un joueur, il y a :

— $\binom{52}{5} = 2\ 598\ 960$ donnes possibles.

— Le nombres de donne qui sont un full (Brelan + paire) est donné par :

$$\binom{13}{1}\cdot\binom{4}{3}\cdot\binom{12}{1}\cdot\binom{4}{2} = 13\cdot 4\cdot 12\cdot 6 = 3744$$

choix de la valeur du brelan, puis les 3 cartes du brelan, puis le choix de la valeur de la paire, puis les deux cartes de la paire. (Proba = 0.0014)

Les coefficients binomiaux reviennent régulièrement, il est bon de connaître leurs propriétés de bases.

Théorème 1.19 (Propriétés des coefficients binomiaux)

a. $\forall n,\ C_n^0 = C_n^n = 1$.

b. $\forall n,\ C_n^1 = C_n^{n-1} = n$.

c. $\forall (k,n)$ avec $0 \leqslant k \leqslant n$, $C_n^k = C_n^{n-k}$.

d. Si $1 \leqslant k \leqslant n$, alors $k \cdot C_n^k = n \cdot C_{n-1}^{k-1}$.

e. *Formule du triangle de Pascal* : si $1 \leqslant k \leqslant n$, alors

$$C_n^k = C_{n-1}^{k-1} + C_{n-1}^k$$

f. *Formule du binôme de Newton* : si $a, b \in \mathbb{R}$, $n \in \mathbb{N}$, alors

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

g. $\sum_k C_n^k = 2^n$.

h. $\sum_k (-1)^k C_n^k = 0$.

R La propriété (e) permet de calculer les C_n^k par le “triangle de Pascal”.

Démonstration.

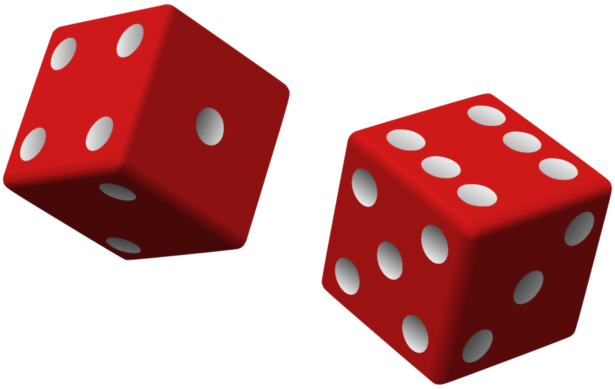
(a), (b) Évident.

(e)

$$\begin{aligned} C_{n-1}^{k-1} + C_{n-1}^k &= \frac{(n-1)!}{(k-1)!(n-k)!} + \frac{(n-1)!}{k!(n-k-1)!} \\ &= \frac{k \cdot (n-1)! + (n-k) \cdot (n-1)!}{k!(n-k)!} \\ &= \frac{n!}{k!(n-k)!} = C_n^k \end{aligned}$$

(c), (d), (f), (g), (h) laissé en Exo. ∎

Exemple 1.20 Un exemple de (f) : $(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$



2. Espaces probabilisés

1 C'est quoi ?

Définition 2.1 — Fausse définition. Soit Ω un ensemble. Une *probabilité* sur Ω est une fonction

$$\mathbb{P}: \mathscr{P}(\Omega) \to [0,1]$$

qui à chaque événement $A \subset \Omega$ associe un nombre, qu'on appelle la *probabilité de* A. Cette fonction doit satisfaire :

1. $\mathbb{P}(\Omega) = 1$.
2. Si $A_1, A_2, A_3, \ldots$ est une collection (finie ou infinie) de sous-ensembles *disjoints* de Ω, alors
$$\mathbb{P}(A_1 \cup A_2 \cup A_3 \cup \ldots) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) + \ldots$$

R [Interprétation] Si $\Omega = \{$résultats d'une expérience aléatoire$\}$, alors Ω est de "volume" 1, et $\mathbb{P}$ mesure le volume de l'événement A.

Définition 2.2 — Correction de la définition.

— Si Ω est fini ou si $\Omega = \mathbb{N}$ ou si $\Omega = \mathbb{Z}$, alors la définition est déjà correcte !

— En revanche, pour $\Omega = \mathbb{R}$, il n'existe pas beaucoup de telle fonction $\mathbb{P}$ (Thm difficile). En général, on ne peut pas demander que $\mathbb{P}(A)$ soit défini pour **tous** les sous-ensembles $A \subset \Omega$, mais que pour ceux appartenant à une certaine famille $\mathscr{A} \subset \mathscr{P}(\Omega)$ qui s'appelle une *tribu* ou *σ-algèbre* de Ω.
Dans le cadre de ce cours, on fera comme si tout événement possède une proba. C'est certes faux, mais cela n'induit pas de graves erreurs dans l'immédiat.

Un *espace probabilisé* est un couple $(\Omega, \mathbb{P})$ où $\mathbb{P}$ est une probabilité sur Ω.

Exemple 2.3 (Ω est fini.) **Expérience aléatoire :** Tirer une fois dans une urne contenant 2 boules vertes, une rouge, une bleue.
$\Omega = \{V, R, B\}$, $\mathbb{P}(V) = 0,5(= 50\%)$, $\mathbb{P}(R) = 0,25(= 25\%)$, $\mathbb{P}(B) = 0,25$,
$\mathbb{P}(\text{“pas rouge”}) = \mathbb{P}(V \cup B) = 0,75$.

Exemple 2.4 (Ω est infini.) **Expérience aléatoire :** choisir un nombre réel x aléatoirement uniformément dans l'intervalle $[0,1]$.
Beaucoup de logiciels sur ordinateurs ont cette fonction.
$\Omega = [0,1]$. Pour tout intervalle $[a,b] \subset [0,1]$, quelle est la probabilité que x soit dans $[a,b]$?
Rép : $\mathbb{P}([a,b])$=longueur de l'intervalle= $b-a$.
Remarquez : pour tout nombre $c \in [0,1]$, on a $\mathbb{P}(\{c\}) = 0$!

2 Premières propriétés

Proposition 2.5 Si $\mathbb{P}$ est une probabilité sur Ω, et si $A, B \subset \Omega$, alors :

a. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

b. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

c. Si $A \subset B$ (l'événement A implique B), alors $\mathbb{P}(A) \leqslant \mathbb{P}(B)$.

Démonstration.

a.

$$\begin{aligned}
\mathbb{P}(A \cup B) &= \mathbb{P}((A \setminus B) \cup (A \cap B) \cup (B \setminus A)) \qquad \text{(réunion disjointe)} \\
&= \mathbb{P}(A \setminus B) + \mathbb{P}(A \cap B) + \mathbb{P}(B \setminus A) \\
&= \mathbb{P}(A \setminus B) + \mathbb{P}(A \cap B) + \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B) - \mathbb{P}(A \cap B) \\
&= \Big(\mathbb{P}(A \setminus B) + \mathbb{P}(A \cap B)\Big) + \Big(\mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B)\Big) - \mathbb{P}(A \cap B) \\
&= \mathbb{P}((A \setminus B) \cup (A \cap B)) + \mathbb{P}((B \setminus A) \cup (A \cap B)) - \mathbb{P}(A \cap B) \\
&= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)
\end{aligned}$$

b. $\Omega = A \cup A^c$ (réunion disjointe), donc

$$\begin{aligned}
\mathbb{P}(A) + \mathbb{P}(A^c) &= \mathbb{P}(A \cup A^c) \\
&= \mathbb{P}(\Omega) \\
&= 1
\end{aligned}$$

c. $B = A \cup (B \setminus A)$ (réunion disjointe), donc $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geqslant \mathbb{P}(A)$ ∎

Notation.

- $\emptyset$ *est l'événement impossible.*
- *Si* $A \subset \Omega$ *avec* $\mathbb{P}(A) = 0$, *alors* A *est* presque impossible.

- Ω *est l'événement certain.*
- *Si* $A \subset \Omega$ *avec* $\mathbb{P}(A) = 1$*, alors* A *est* presque certain.
- *Si* A *et* B *sont disjoints on dit aussi qu'ils sont des* événements incompatibles.

Définition 2.6 Les événements $\{\omega\}$ pour $\omega \in \Omega$, autrement dit les singletons, s'appellent les *événements élémentaires*.

Notation. *Lorsque* A, B *sont deux événements* ***disjoints*** *d'un univers* Ω*, on notera* $A \sqcup B$ *leur union. Autrement dit, lorsqu'on écrit* $A \sqcup B$ *cela sous-entend que* $A \cap B = \emptyset$*.*

3 Le cas où Ω est fini

3.A. Les probabilités des événements élémentaires suffisent

Proposition 2.7 Supposons Ω est fini : $\Omega = \{\omega_1, \ldots, \omega_n\}$. Alors pour spécifier une probabilité $\mathbb{P}$ sur Ω, il faut et il suffit de se donner les nombres $p_1 = \mathbb{P}(\omega_1), \ldots, p_n = \mathbb{P}(\omega_n)$ tels que :

$$\begin{cases} \textbf{1.} \quad \forall i, \quad p_i \geqslant 0; \\ \textbf{2.} \quad \sum_{i=1}^{n} p_i = 1. \end{cases}$$

En effet, étant donné une telle famille de nombres $(p_i)_i$, on peut bien calculer la probabilité de tous événements A via :

$$\mathbb{P}(A) = \sum_{i, \omega_i \in A} p_i \qquad \text{p.ex} \quad \mathbb{P}(\{\omega_2, \omega_4, \omega_7\}) = p_2 + p_4 + p_7.$$

Démonstration. Il faut juste vérifier la formule pour $\mathbb{P}(A)$, on a :

$$\mathbb{P}(A) = \mathbb{P}(\bigsqcup_{\omega \in A} \mathbb{P}(\{\omega\}) = \mathbb{P}(\bigsqcup_{i, \omega_i \in A} \mathbb{P}(\{\omega_i\}) = \sum_{i, \omega_i \in A} \mathbb{P}(\{\omega_i\}) = \sum_{i, \omega_i \in A} p_i$$

En particulier, en prenant $A = \Omega$, on trouve :

$$1 = \mathbb{P}(\Omega) = \sum_{i=1}^{n} p_i.$$

■

Exemple 2.8 Dans l'Exemple 2.3, $\mathbb{P}(V \sqcup B) = \mathbb{P}(V) + \mathbb{P}(B) = 0{,}75$.

3.B. Le cas équiprobable

Définition 2.9 Un espace probabilisé $(\Omega, \mathbb{P})$ est *équiprobable* lorsque les événéments élémentaires ont tous la même probabilité $p > 0$, c-à-d :

$$\forall \omega, \omega' \in \Omega, \qquad \mathbb{P}(\{\omega\}) = \mathbb{P}(\{\omega'\}) = p > 0.$$

Proposition 2.10 — Cas de l'équiprobabilité. Soit $(\Omega, \mathbb{P})$ un espace probabilisé équiprobable. Alors :

1. Ω est fini.
2. $\forall \omega \in \Omega$, $\mathbb{P}(\{\omega\}) = \dfrac{1}{\mathrm{Card}(\Omega)}$.
3. $\forall A \in \mathscr{P}(\Omega)$, $\mathbb{P}(A) = \dfrac{\mathrm{Card}(A)}{\mathrm{Card}(\Omega)}$.

Démonstration. On note p la probabilité des événements élémentaires. Supposons Ω infini, il existe alors une suite infinie d'éléments distincts $(\omega_i)_{i\geqslant 1}$ de Ω et on a :

$$1 = \mathbb{P}(\Omega) \leqslant \mathbb{P}\Big(\bigsqcup_{i\geqslant 1} \{\omega_i\}\Big) = \sum_{i\geqslant 1} p_i = \infty \times p$$

Absurde donc Ω est fini. On écrit $\Omega = \{\omega_1, ..., \omega_n\}$, on fait le calcul analogue du précédent :

$$1 = \mathbb{P}(\Omega) = \mathbb{P}\Big(\bigsqcup_{i=1}^{n} \{\omega_i\}\Big) = \sum_{i=1}^{n} p_i = n \times p$$

D'où :

$$p = \frac{1}{n}.$$

Enfin, si $A \subset \Omega$, on a :

$$\mathbb{P}(A) = \mathbb{P}\Big(\bigsqcup_{i,\omega_i\in A} \mathbb{P}(\{\omega_i\}\Big) = \sum_{i,\omega_i\in A} p_i = \sum_{i,\omega_i\in A} p = p \times \sum_{i,\omega_i\in A} 1 = p \times \mathrm{Card}(A) = \frac{\mathrm{Card}(A)}{\mathrm{Card}(\Omega)}.$$

■

Exemple 2.11 Dans l'Exemple 1.2[a], $\mathbb{P}$("deux fois le même nombre") $= 6/36 = 1/6$

a. "Je jette deux fois de suite un dé".

Exemple 2.12 Un exemple célèbre : Quelle est la proba P_n que parmi n personnes, au moins deux ont le même anniversaire ?[a] **Réponse :**

$$P_n = 1 - \mathbb{P}(n \text{ personnes ont toutes des anniversaires différents})$$

Pour calculer ça, soit

$$\Omega = \text{ Arrangements avec répétition de } n \text{ anniversaires } = \{1 \text{ janv}, \ldots, 31 \text{ déc}\}^n$$

donc $\mathrm{Card}(\Omega) = 365^n$. Soit

$$A = \{ \text{Arrangements sans répétition de } n \text{ anniversaires}\}$$

donc $\mathrm{Card}(A) = A_{365}^n = 365 \cdot 364 \cdot ... \cdot (365 - n + 1)$. Enfin résultat :

$$P_n = 1 - \frac{A_{365}^n}{365^n}$$

Calcul : si $n = 23$, alors P_n est un peu plus grand que $0,5$! !
si $n = 47$, alors $0,95$.

a. Pas forcément avec le même âge, juste le même jour anniversaire. On oublie les complications du type : 29 février etc.

4 Le cas où Ω est infini mais dénombrable

Définition 2.13 L'ensemble Ω est *dénombrable* s'il existe une liste infinie mentionnant tous ses éléments, c.à.d.,

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \ldots\}$$

R Les ensembles $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ sont dénombrables. Si E est dénombrable alors pour tout entier $r \geqslant 1$, l'ensemble E^r est dénombrable.
Les ensembles $\mathbb{R}$ et $\mathscr{P}(\mathbb{N})$ ne sont pas dénombrables.

Définition 2.14 L'ensemble Ω est *au plus dénombrable* s'il est fini ou dénombrable.

Proposition 2.15 — Les probabilités des événements élémentaires suffisent. Supposons Ω est dénombrable : $\Omega = \{\omega_1, ..., \omega_n, ...\}$. Alors pour spécifier une probabilité $\mathbb{P}$ sur Ω, il faut et il suffit de se donner les nombres $(p_i)_{i\geqslant 1}$ tels que :

$$\begin{cases} \textbf{1.} \quad \forall i \geqslant 1, \quad p_i \geqslant 0; \\ \textbf{2.} \quad \displaystyle\sum_{i=1}^{\infty} p_i = 1. \end{cases}$$

En effet, étant donné une telle famille de nombres $(p_i)_i$, on peut bien calculer la probabilité de tous événements A via :

$$\mathbb{P}(A) = \sum_{i, \omega_i \in A} p_i.$$

Démonstration. Admis. ■

Exemple 2.16 On définit la probabilité suivante sur $\Omega = \mathbb{N}^*$ en posant :

$$\text{Pour tout } k \in \Omega, \qquad \mathbb{P}(\{k\}) = p_k = \frac{1}{2^k}.$$

1. Vérifier que $\mathbb{P}$ définit bien une probabilité.
2. Calculer la probabilité de l'événement $A = \{2, 4\}$.
3. Calculer la probabilité des événements $P = \{k \in \Omega \,|\, k \text{ est pair}\}$ et $I = \{k \in \Omega \,|\, k \text{ est impair}\}$.

Solution :

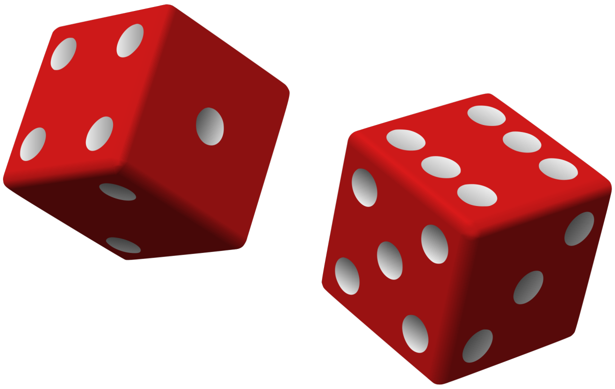
1. *Clairement les p_k sont positifs. Il reste à vérifier que leur somme fait* 1*. On a :*

$$\sum_{k\geqslant 1} p_k = \sum_{k\geqslant 1} \frac{1}{2^k} = \frac{1/2}{1-1/2} = 1.$$

2. $$\mathbb{P}(\{2,4\}) = p_2 + p_4 = \frac{1}{2^2} + \frac{1}{2^4} = \frac{4}{16} + \frac{1}{16} = \frac{5}{16}.$$

3. $$\mathbb{P}(P) = \sum_{k\geqslant 1, k\ pair} p_k = \sum_{n\geqslant 1} p_{2n} = \sum_{n\geqslant 1} \frac{1}{2^{2n}} = \sum_{n\geqslant 1} \frac{1}{4^n} = \frac{1/4}{1-1/4} = \frac{1}{3}.$$

   Par passage au complémentaire, on obtient $\mathbb{P}(I) = 2/3$.



3. Probabilité conditionnelle et indépendance

1 C'est quoi ?

Exemple 3.1 Soit $\Omega = \{\text{hommes français de 30-60 ans}\}$. Trois sous-ensembles :
$A = \{$ ceux dont la pointure des chaussures est plus élevée que la moyenne $\}$
$B = \{$ ceux avec un nombre d'années d'études plus élevé que la moyenne $\}$
$C = \{$ ceux avec revenus plus élevés que la moyenne $\}$

A et B sont (à peu près) indépendants : si je sais que Monsieur X appartient à B, ça ne donne pas d'information sur sa probabilité d'appartenir aussi à A. Donc

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

Autrement dit :

$$\underbrace{\frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}}_{\text{Proba. d'appartenir à } A \text{ parmi ceux qui appartiennent à } B} = \mathbb{P}(A)$$

En revanche, nombre d'années d'étude et revenu sont positivement corrélés : si je sais que Monsieur X appartient à C, ses chances d'appartenir aussi à B sont plus grandes que $\mathbb{P}(B)$. Donc

$$\mathbb{P}(B \cap C) > \mathbb{P}(B) \cdot \mathbb{P}(C)$$

Autrement dit :

$$\underbrace{\frac{\mathbb{P}(B \cap C)}{\mathbb{P}(C)}}_{\text{Proba. d'appartenir à } B \text{ parmi ceux qui appartiennent à } C} > \mathbb{P}(B)$$

(Il y a aussi la possibilité d'une corrélation négative, où $\mathbb{P}(B \cap C) < \mathbb{P}(B) \cdot \mathbb{P}(C)$)

Définition 3.2 (Indépendance de deux événements)
Soit Ω un univers muni d'une probabilité $\mathbb{P}$. Soient $A, B \subset \Omega$ deux événements. Les deux événements A et B sont *indépendants* lorsque

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Définition 3.3 (Probabilité conditionnelle)
Soit Ω un univers muni d'une probabilité $\mathbb{P}$. Soient $A, B \subset \Omega$ deux événements. Supposons que $\mathbb{P}(B) > 0$. La *probabilité conditionnelle* de A sachant B, est définie comme suit :

$$\mathbb{P}(A|B) = \mathbb{P}_B(A) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Exercice Soit Ω un univers muni d'une proba $\mathbb{P}$ et $B \subset \Omega$ avec $\mathbb{P}(B) > 0$. Alors la probabilité conditionnelle $\mathbb{P}(\cdot|B)$ est bien une proba sur Ω. ■

Lemme 3.4 Deux événements A et B avec $\mathbb{P}(A) > 0$ et $\mathbb{P}(B) > 0$ sont indépendants
si et seulement si $\mathbb{P}(A|B) = \mathbb{P}(A)$,
si et seulement si $\mathbb{P}(B|A) = \mathbb{P}(B)$.

Démonstration. Immédiate à partir des définitions. ■

Exemple 3.5

- Deux lancers indépendants d'une pièce potentiellement truquée :
  Pile avec proba p, Face avec proba $1 - p$.

  Soit $\Omega = \{(\text{Pile}, \text{Face}), (\text{Pile}, \text{Pile}), (\text{Face}, \text{Pile}), (\text{Face}, \text{Face})\}$, $A = \{\text{premier lancer Pile}\} = \{(\text{Pile}, \text{Pile}), (\text{Pile}, \text{Face})\}$ et $B = \{\text{deuxième lancer pile}\} = \{(\text{Face}, \text{Pile}), (\text{Pile}, \text{Pile})\}$.

  L'indépendance des deux lancers nous donnent le tableau suivant pour les probas des éléments de Ω.

1^{er}-lancer \ 2^{eme}-lancer	Pile	Face
Pile	p^2	$p(1-p)$
Face	$(1-p)p$	$(1-p)^2$

  Alors $\mathbb{P}(A) = p^2 + p(1-p) = p$, $\quad \mathbb{P}(B) = (1-p)p + p^2 = p$, $\quad \mathbb{P}(A \cap B) = p^2 = \mathbb{P}(A) \cdot \mathbb{P}(B)$
- On joue à pile ou face jusqu'à la première apparition de Face. Alors :

$$\mathbb{P}(\text{ Première apparition de Face au } n\text{-ème lancer }) = p^{n-1}(1-p)$$

  Interprétation : Temps d'attente jusqu'au premier succès. On va utiliser cet exemple plus tard (Exemple 4.22).

Exemple 3.6 On jette deux dés.
$\Omega = \{1,2,3,4,5,6\}^2$.
A = la somme des deux dés est $4 = \{(1,3),(2,2),(3,1)\}$.
B = le premier dé donne $1 = \{(1,1),\ldots,(1,6)\}$.
$A \cap B = \{(1,3)\}$. Alors

$$\mathbb{P}(A) = 3/36 = 1/12 \text{ et } \mathbb{P}(B) = 6/36 = 1/6, \text{ donc } \mathbb{P}(A) \cdot \mathbb{P}(B) = 1/72$$

mais $\mathbb{P}(A \cap B) = 1/36$. Donc $\mathbb{P}(A \cap B) > \mathbb{P}(A) \cdot \mathbb{P}(B)$ - les deux événements ont une corrélation positive. En fait $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{1}{6} > \frac{1}{12} = \mathbb{P}(A)$

Proposition 3.7 Si A et B sont indépendants, alors A^c et B sont aussi indépendants.

Démonstration. On a

$$B = (A \cap B) \sqcup (A^c \cap B)$$

On en déduit que :

$$\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B) = \mathbb{P}(A)\mathbb{P}(B) + \mathbb{P}(A^c \cap B)$$

Ainsi,

$$\mathbb{P}(A^c \cap B) = \big(1 - \mathbb{P}(A)\big)\mathbb{P}(B).$$

■

2 Trois formules importantes

Proposition 3.8 (Formule des probas composées)
Si $A_1, \ldots, A_n$ sont des événements de Ω tels que $\mathbb{P}(A_1 \cap \ldots \cap A_n) > 0$, alors

$$\mathbb{P}(A_1 \cap \ldots \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|A_1 \cap A_2) \cdot \ldots \cdot \mathbb{P}(A_n|A_1 \cap \ldots \cap A_{n-1})$$

Démonstration. Par récurrence. Pour $n = 2$, c'est la définition de la probabilité conditionnelle.
$n - 1 \Longrightarrow n$:

$$\begin{aligned}\mathbb{P}((A_1 \cap \cdots \cap A_{n-1}) \cap A_n) &\overset{\text{définition}}{=} \mathbb{P}(A_1 \cap \ldots \cap A_{n-1}) \cdot \mathbb{P}(A_n|A_1 \cap \ldots \cap A_{n-1}) \\ &\overset{\text{récurrence}}{=} \mathbb{P}(A_1) \cdot \ldots \cdot \mathbb{P}(A_{n-1}|A_1 \cap \ldots \cap A_{n-2}) \cdot \mathbb{P}(A_n|A_1 \cap \cdots \cap A_{n-1}))\end{aligned}$$

■

Exemple 3.9 Urne avec 3 boules rouges, 3 blanches. On fait trois tirages sans remise. $\mathbb{P}(BBB) = ?$

Réponse : événements B_i = blanc au $i$ème tirage.

$$\begin{aligned}\mathbb{P}(BBB) &= \mathbb{P}(B_1 \cap B_2 \cap B_3)\\ &= \mathbb{P}(B_1)\cdot\mathbb{P}(B_2|B_1)\cdot\mathbb{P}(B_3|B_1\cap B_2)\\ &= 3/6\cdot 2/5\cdot 1/4\\ &= 1/20\end{aligned}$$

Définition 3.10 Une collection (finie ou infinie) $B_1, B_2, B_3, \ldots \subset \Omega$ de sous-ensembles de Ω forme une partition de Ω si

- ils sont deux-à-deux disjoints et
- $B_1 \sqcup B_2 \sqcup B_3 \sqcup \ldots = \Omega$.

R [Rappel] Si $\Omega = B_1 \sqcup B_2 \sqcup B_3 \sqcup \ldots$ est une partition, alors

$$\mathbb{P}(\Omega) = \mathbb{P}(B_1) + \mathbb{P}(B_2) + \mathbb{P}(B_3) + \ldots$$

Exemple 3.11 Pour tout $B \subset \Omega$, on a une partition $\Omega = B \cup B^c$.

Lemme 3.12 (Formule des probabilités totales) Soit $B_1, B_2, \ldots \subset \Omega$ une collection finie ou infinie d'événements qui forme une partition de Ω. Alors

$$\mathbb{P}(A) = \sum_i \mathbb{P}(A \cap B_i) = \sum_i \mathbb{P}(B_i)\cdot\mathbb{P}(A|B_i)$$

Démonstration. Exercice ■

Proposition 3.13 (Formule de Bayes) Sous la même hypothèse que 3.12, et si $\mathbb{P}(A) > 0$, alors pour $i = 1, 2, \ldots,$

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(B_i)\cdot\mathbb{P}(A|B_i)}{\mathbb{P}(B_1)\cdot\mathbb{P}(A|B_1) + \mathbb{P}(B_2)\cdot\mathbb{P}(A|B_2) + \ldots}$$

Démonstration. Numérateur $= \mathbb{P}(A \cap B_i)$, dénominateur $\stackrel{3.12}{=} \mathbb{P}(A)$. □ ■

Application typique de la formule de Bayes : un événement observé peut avoir plusieurs causes, et on veut calculer leur probabilité.

Astuce pratique : ne pas mémoriser la formule, mais dessiner un arbre de probabilités. Comme dans l'exemple qui suit :

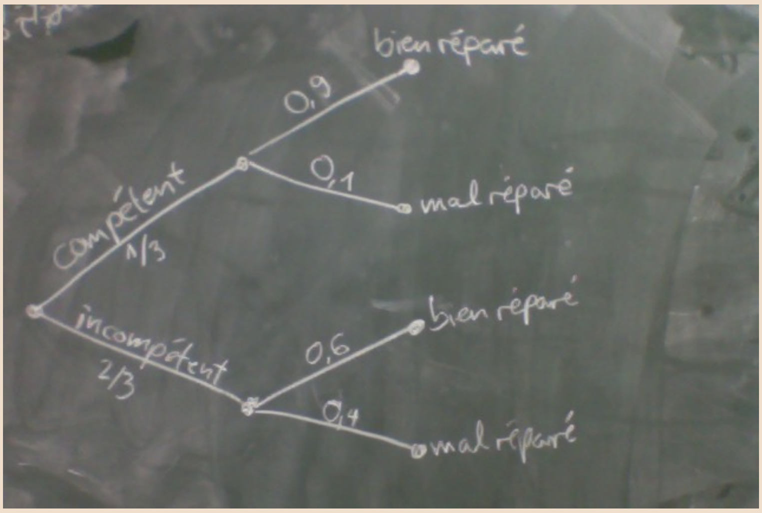
Exemple 3.14 Il y a deux types de plombiers : compétents (qui réparent mes WC correctement avec une proba de 0,9) et incompétents (proba 0,6). La proba qu'un plombier choisi au hasard dans l'annuaire soit compétent est ⅓. J'appelle un plombier au hasard, et il répare mes WC correctement. Quelle est la probabilité qu'il soit en fait compétent ?

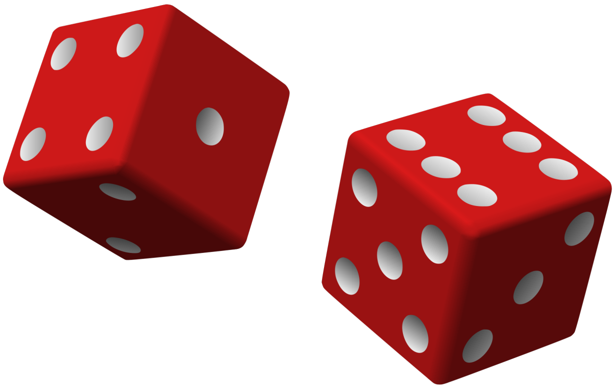
Réponse : épreuve aléatoire : appeler au hasard, attendre le résultat de la réparation. Dans le dessin on voit :

$$\mathbb{P}(\text{comp.}|\text{bien réparé}) = \frac{\mathbb{P}(\text{comp. et bien rép.})}{\mathbb{P}(\text{bien réparé})} = \frac{\mathbb{P}(\text{branche 1})}{\mathbb{P}(\text{branche 1 ou 3})} = \frac{\frac{1}{3}\cdot 0,9}{\frac{1}{3}\cdot 0,9+\frac{2}{3}\cdot 0,6} = \frac{0,3}{0,7} = \frac{3}{7}$$

Donc une seule réparation réussie n'est pas un signal fort qu'il est compétent ! Formellement, événements C =compétent, R =bien réparé.

$$\mathbb{P}(C|R) \overset{Bayes}{=} \frac{\mathbb{P}(C)\cdot\mathbb{P}(R|C)}{\mathbb{P}(C)\cdot\mathbb{P}(R|C)+\mathbb{P}(C^c)\cdot\mathbb{P}(R|C^c)} = \frac{\frac{1}{3}\cdot 0,9}{\frac{1}{3}\cdot 0,9+\frac{2}{3}\cdot 0,6} = \frac{3}{7}$$





4. Variables aléatoires discrètes

1 Variables aléatoires

1.A. C'est quoi ?

Définition 4.1 — Fausse définition. Considérons un ensemble Ω muni d'une probabilité $\mathbb{P}$. Une *variable aléatoire* (abréviation v.a., en anglais "random variable") sur Ω est une fonction $X : \Omega \to \mathbb{R}$.

Définition 4.2 — Correction de la définition. (hors programme) Ω est muni d'une *tribu* $\mathscr{A} \subset \mathscr{P}(\Omega)$, et la fonction X ne doit pas être trop sauvage : pour tout $t \in \mathbb{R}$, l'ensemble $X^{-1}(]-\infty, t]) \subset \Omega$ doit appartenir à la tribu $\mathscr{A}$.

Exemple 4.3

- Je jette deux dés. Alors $\Omega = \{(1,1),(1,2),...(6,5),(6,6)\}$. On peut définir $X(i,j) = i+j$ (la somme des deux dés).
- Je joue à pile ou face n fois. Alors $\Omega = \{P,F\}^n$, et on peut définir deux variables aléatoires X et Y : pour tout $\omega \in \Omega$, soit
  * $X(\omega) =$ nombre d'apparitions de "P"
  * $Y(\omega) = \begin{cases} 0 & \text{si } \omega \text{ ne contient que des "F"} \\ k & \text{si la première apparition de "P" est lors du } k\text{-ème lancer.} \end{cases}$
- $\Omega = \{$êtres humains vivants$\}$, pour tout $\omega \in \Omega$, $\mathbb{P}(\omega) = 1/7$milliards, $X(\omega) =$âge que ω aura au moment de sa mort.
- Un exemple d'une v.a. à valeurs dans tout $\mathbb{R}_+$. Épreuve aléatoire : je sélectionne une

particule d'une substance radioactive, et j'attends sa décomposition. $\Omega = \{$atomes de la substance$\}$, $X : \Omega \to \mathbb{R}$, $\omega \mapsto$ temps jusqu'à la décomposition.

Proposition 4.4 Si X et Y sont deux v.a. et $\lambda \in \mathbb{R}$, alors la somme $X+Y$, le produit $X \cdot Y$, et le multiple scalaire $\lambda \cdot X$ sont aussi des v.a..

Démonstration. Admis. ∎

1.B. Fonction de répartition

Pour tout nombre $x \in \mathbb{R}$, on écrit "$\{X = x\}$" pour l'événement $\{\omega \in \Omega \mid X(\omega) = x\}$, idem avec $<, >, \leqslant, \geqslant$. Pour tout événement $A \subset \mathbb{R}$, on écrit "$\{X \in A\}$" pour l'événement $\{\omega \in \Omega \mid X(\omega) \in A\}$.

Définition 4.5 La *fonction de répartition* F_X d'une variable aléatoire X est la fonction réelle définie par :

$$F_X(t) = \mathbb{P}(X \leqslant t)$$

Définition 4.6 On dit que deux variables aléatoires X et Y suivent la même loi lorsque pour tout $A \subset \mathbb{R}$, on a :

$$\mathbb{P}(X \in A) = \mathbb{P}(Y \in A)$$

Théorème 4.7 Deux variables aléatoires suivent la même loi si et seulement si leurs fonctions de répartitions sont égales.

2 Variables aléatoires discrètes

Définition 4.8 Soit $X : \Omega \to \mathbb{R}$ une variable aléatoire. Un réel $x \in \mathbb{R}$ est un *atome* de X lorsque $\mathbb{P}(X = x) > 0$. On note $\mathrm{Atom}(X)$ l'ensemble des atomes de X.

R L'ensemble $\mathrm{Atom}(X)$ est au plus dénombrable.

Définition 4.9 Une variable aléatoire est discrète lorsque $\mathbb{P}(X \in \mathrm{Atom}(X) = 1)$.

Proposition 4.10 Soit X variable aléatoire X si son image $X(\Omega) = \{X(\omega) \mid \omega \in \Omega\}$ est au plus dénombrable alors X est *discrète*.

Démonstration. Par définition $\mathbb{P}(X \in X(\Omega)) = 1$ et $\mathrm{Atom}(X) \subset X(\Omega)$. On a :

$$1 = \mathbb{P}(X \in X(\Omega)) = P(X \in \mathrm{Atom}(X)) + P(X \in X(\Omega) \setminus \mathrm{Atom}(X))$$

Comme $X(\Omega)$ est au plus dénombrable, on a :

$$P(X \in X(\Omega) \setminus \mathrm{Atom}(X)) = \sum_{a \in X(\Omega) \setminus \mathrm{Atom}(X)} \underbrace{\mathbb{P}(X = a)}_{=0} = 0$$

∎

R Dans ce cours, toutes les v.a. discrètes X que nous étudierons en détail seront très sages : elles satisferont :

- Ω est fini ou dénombrable, et
- $X(\Omega) = \text{Atom}(X) \subset \mathbb{Z}$.

Notation. *Soit $X \colon \Omega \to \mathbb{R}$ la v.a. discrète de l'Exemple 4.3(1er) (X =somme de deux dés) alors :*

$$\mathbb{P}(X = 3) = \mathbb{P}(\{(1,2),(2,1)\}) = \frac{2}{36} = 0,05555\ldots$$

Définition 4.11 La *loi* de probabilité (en anglais : probability distribution) d'une variable aléatoire discrète X est la probabilité $\mathbb{P}_X$ sur l'ensemble $X(\Omega) \subset \mathbb{R}$ définie par

$$\mathbb{P}_X(\{x_i\}) \;=\; \mathbb{P}(X = x_i).$$

R On la représente souvent par un tableau avec en première ligne la liste des $(x_i)_i$ et en seconde ligne les valeurs de $\mathbb{P}(X = x_i)$.

Commentaire : C'est une définition qui a l'air abstraite, mais c'est un objet intuitif et très important. Quand on étudie la *loi* d'une variable aléatoire, on oublie tout le modèle du monde réel sous-jacent, toute l'expérience aléatoire etc : pour spécifier la loi on doit juste dire quel nombre est tiré avec quelle probabilité. On peut penser que la loi d'une variable aléatoire est une boîte noire avec un bouton et un écran qui peut afficher juste un nombre. Chaque fois qu'on appuie sur le bouton, la boite affiche un nombre au hasard – mais les fréquences relatives des nombres sont distribuées selon la loi en question.

Exemple 4.12 (voir 4.3(2ème)) Je joue à pile ou face 2 fois. $\Omega = \{(P,P),(P,F),(F,P),(F,F)\}$, $X \colon \Omega \to \{0,1,2\}$, $\omega \mapsto$ nombre de "pile". Loi de la v.a. X :

$X(\Omega) = \{0,1,2\}$

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

[Donc ici, la boite noire dit "0" dans 25% des cas, "1" dans 50% des cas, et "2" dans 25% des cas.]

R [Première visualisation] La loi d'une variable aléatoire discrète peut être représentée visuellement par son *diagramme en bâtons*. Par exemple pour la loi de l'Exemple 4.12, voir la figure 4.1.

R [Deuxième visualisation] Le graphe de la fonction de répartition. (Remarquer que $\lim_{t\to-\infty} F = 0$, que F est croissante, et que $\lim_{t\to+\infty} F = 1$. Dans le cas de l'exemple 4.12 :

Attention : Ne pas confondre *diagramme en bâtons* et *fonction de répartition*.

Proposition 4.13 Si X et Y sont discrètes alors $X+Y$, $X \cdot Y$ et λX le sont aussi.

Démonstration. Admis. ∎

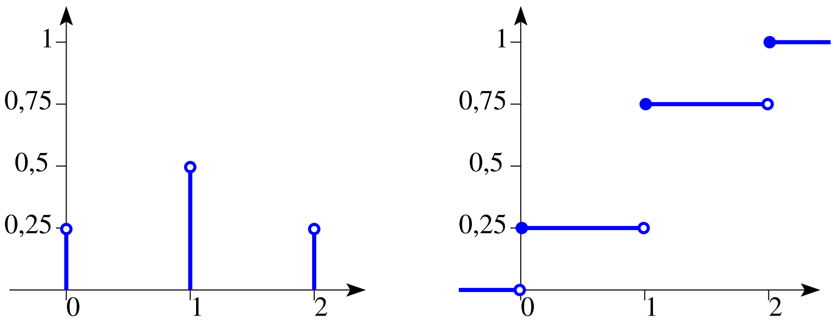


FIGURE 4.1 – Diagramme en bâtons et fonction de répartition de l'Exemple 4.12

Proposition 4.14 Deux variables aléatoires discrètes X et Y ont la même loi si et seulement si :

$$\mathrm{Atom}(X) = \mathrm{Atom}(Y) \quad \text{et} \quad \forall a \in \mathrm{Atom}(X), \text{ on a : } \mathbb{P}(X = a) = \mathbb{P}(Y = a).$$

R Quand on demande la loi d'une variable aléatoire discrète, on demande l'ensemble $\mathrm{Atom}(X)$ ou l'ensemble $X(\Omega)$ (on passe sur la subtilité entre les deux qui n'apparait pas en pratique) ET le tableau 1ère ligne les éléments de $X(\Omega)$, seconde ligne les valeurs $\mathbb{P}(X = x)$ pour $x \in X(\Omega)$. Cf Exemple 4.12.

R [Rappel sur les suites et séries.] Considérons la série $\sum a_k$, où $(a_k)_{k\in\mathbb{N}}$ est une suite de nombre réels.

- On dit que la série $\sum a_k$ est *convergente* si la suite des sommes partielles $S_k = \sum_{k=1}^{k} a_k$ est convergente. Exemple : série convergente $\sum_{k=1}^{\infty} \frac{1}{2^i} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots = 1$.
- Une condition nécessaire pour que la série $\sum a_k$ converge est que $(a_k)_k \to 0$. Cette condition n'est pas suffisante - par exemple $\sum_{k=1}^{\infty} \frac{1}{i} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots = \infty$ (série divergente)
- Soit $q \in [0,1[$. Alors $(1-q)\sum_{k=0}^{\infty} q^i = 1$. Par changement de variables $p = 1-q$ on obtient :

$$p \cdot \sum_{k=0}^{\infty} (1-p)^i = 1$$

- Soit $\alpha > 0$. La série de terme général $a_k = \frac{1}{k^\alpha}$ est convergente si et seulement si $\alpha > 1$.
- Si $0 \leqslant a_k \leqslant b_k$ et la série Σb_k est convergente alors la série Σa_k est aussi convergente.

R [Construire une variable aléatoire discrète à l'aide d'une série positive et convergente] Soit $\sum a_k$ une série convergente tel que $a_k > 0$ pour tout k. On note $L = \sum_{k=0}^{\infty} a_k$. On note $p_k = a_k/L$. Ainsi, $\sum_{k=0}^{\infty} p_k = 1$ et on peut définir une variable aléatoire X dont la loi est donnée par :

$$X(\Omega) = \mathbb{N} \qquad \text{et} \qquad \mathbb{P}(X = k) = p_k.$$

3 Des lois discrètes classiques

3.A. La loi uniforme discrète

Notation. *Soient* $p, q \in \mathbb{Z}$ *deux entiers tels que* $p \leqslant q$, *on note :* $[\![p, q]\!] = \{p, \ldots, q\}$.

Définition 4.15 (**Loi uniforme** - en anglais “uniform distribution”)
On dit qu'une v.a. $X : \Omega \to \mathbb{R}$ suit une *loi uniforme sur* $[\![1, n]\!]$ lorsque :

- $X(\Omega) = [\![1, n]\!]$, et
- pour tout $k \in [\![1, n]\!]$ on a $\mathbb{P}(X = k) = 1/n$.

Notation : $X \sim \mathscr{U}([\![1, n]\!])$

Exemple 4.16

- Épreuve aléatoire : Lancer d'un dé.

$$\Omega = \{1, 2, \ldots, 6\}, \; X : \Omega \to \{1, \ldots, 6\}, \omega \mapsto \text{le numéro de la face obtenue.}$$

  Alors $X \sim \mathscr{U}([\![1, 6]\!])$.
- Épreuve aléatoire : une urne contient n boules, numérotés $1, \ldots, n$. On en prend une au hasard.

$$\Omega = \{\text{les boules}\}, \; X : \Omega \to \{1, \ldots, n\}, \omega \mapsto \text{le numéro de la boule } \omega.$$

  Alors $X \sim \mathscr{U}([\![1, n]\!])$.

3.B. La loi de Bernoulli

Notation (Convention). *Dans un jeu de pile ou face, on va interpréter Pile comme un succès et Face comme un échec.*

Définition 4.17 (**Loi de Bernoulli**)
On dit qu'une v.a. $X : \Omega \to \mathbb{R}$ suit une *loi de Bernoulli de paramètre* p (avec $0 < p < 1$) lorsque :

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

Notation $X \sim \mathscr{B}(1, p)$.

Exemple 4.18 Épreuve aléatoire : on joue à pile ou face avec une pièce qui tombe sur pile avec une probabilité p. Donc

$$\Omega = \{\text{pile, face}\}, \; \mathbb{P}(\text{pile}) = p, \; \mathbb{P}(\text{face}) = 1 - p$$

$$X : \Omega \to \{0, 1\}, \; \text{pile} \mapsto 1, \text{face} \mapsto 0.$$

Alors $X \sim \mathscr{B}(1,p)$.

3.C. La loi binomiale

Définition 4.19 **(La loi binomiale)**
On dit qu'une v.a. $X : \Omega \to \mathbb{R}$ suit une *loi binomiale de paramètre* (n,p) lorsque :

- $X(\Omega) = [\![0,n]\!]$ et
- pour $k \in [\![0,n]\!]$ on a $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.

Notation : $X \sim \mathscr{B}(n,p)$.

R Pour $n = 1$ on obtient la loi binomiale.

Exemple 4.20 (Loi binomiale - cette loi compte le nombre de succès parmi n essais indépendants.)
Épreuve aléatoire : avec la même pièce, on joue à pile ou face n fois, et on compte le nombre de "pile".

$$\Omega = \{\text{pile, face}\}^n \text{ et } X : \Omega \to [\![0,n]\!],\ \omega \mapsto \text{ nombre de "pile"}.$$

Si ω contient k fois "Pile" et $n-k$ fois "Face", alors $\mathbb{P}(\omega) = p^k(1-p)^{n-k}$. Or, il y a C_n^k façons de distribuer les k "Pile" sur les n lancers, donc

$$\mathbb{P}(X = k) = C_n^k \cdot p^k (1-p)^{n-k}.$$

Autrement dit, $X \sim \mathscr{B}(n,p)$.

3.D. La loi géométrique

Définition 4.21 **(Loi géométrique)**
On dit qu'une v.a. $X : \Omega \to \mathbb{R}$ suit une *loi géométrique de paramètre* p lorsque :

- $X(\Omega) = \{1,2,3,4,...\} = \mathbb{N}^*$ et
- $\mathbb{P}(X = k) = p \cdot (1-p)^{k-1}$.

Notation : $X \sim \mathscr{G}(p)$.

R Ceci est bien une loi car $\sum_k \mathbb{P}(X = k) = 1$ par le rappel 2(c).

Exemple 4.22 (Loi géométrique - cette loi mesure le temps d'attente jusqu'au premier succès d'une suite d'essais indépendants.) Épreuve aléatoire : toujours avec la même pièce, on joue à pile ou face jusqu'au premier pile. Alors

$$\Omega = \{P, FP, FFP, FFFP, FFFFP, ...\}$$

et

$$X : \Omega \to \mathbb{N},\quad \omega \mapsto \text{longueur du mot } \omega$$

La probabilité de $(F...FP)$ ($k-1$ fois la lettre F, suivie par P) est $(1-p)^{k-1} \cdot p$

Exercice Démontrer que c'est la "loi d'une v.a. sans mémoire" : si $T>0$ et $\tau>0$ alors

$$\mathbb{P}(X>\tau)=\mathbb{P}(X>T+\tau \mid X>T).$$

- Indication : montrer d'abord que $\mathbb{P}(X>k)=(1-p)^k$.
- Interprétation : Pas de mémoire, car si j'ai déjà obtenu 10 fois face, alors ça ne m'apporte aucune nouvelle information sur la probabilité future de Pile. (Ceci est à contraster avec la variable aléatoire X suivante : $\Omega=\{\text{êtres humains}\}$, $X(\omega)=$durée de vie de ω (en années). La probabilité que la durée de vie d'une personne prise au hasard est supérieure à $k=$30 ans est bien différente de la probabilité que la durée de vie d'une personne prise parmi celles de durée de vie au moins $K=70$ ans soit supérieure à $K+k=70+30=100$ ans.)

3.E. La loi de Poisson

Définition 4.23 **(Loi de Poisson) [**d'après Siméon Denis Poisson, 1781 - 1840**]**
On dit qu'une v.a. $X:\Omega\to\mathbb{R}$ suit une *loi de Poisson de paramètre* λ (où $\lambda\in\mathbb{R},\lambda>0$) lorsque :
- $X(\Omega)=\{0,1,2,3,\ldots\}$ et
- $\mathbb{P}(X=k)=e^{-\lambda}\cdot\frac{\lambda^k}{k!}$.

Notation $X\sim\mathscr{P}(\lambda)$.

R Ceci est bien une loi, car $\sum_{k=0}^{\infty}\mathbb{P}(X=k)=e^{-\lambda}\cdot\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=1$.

R Occurrences dans la vraie vie :
- Nombre de décompositions de particules par seconde dans une substance radioactive.
- Nombre de requêtes par microseconde à google.fr autour de 9h.
- Nombre de fautes de frappe par page dans un livre.

Il n'y a pas de modèle simple, c'est une loi limite : soit $\lambda>0$ fixé. Soit $n\in\mathbb{N}$ très grand, et soit $p=\lambda/n$ (donc p est très petit). Je fais n fois un essai dont la probabilité de succès est de p (beaucoup d'essais avec individuellement peu de chances). La variable aléatoire X compte le nombre de succès (en moyenne il y en a toujours $n\cdot p=\lambda$). Alors $X\sim\mathscr{P}(\lambda)$ (quand $n\to\infty$). Plus formellement

Proposition 4.24 Soit $\lambda>0$. Soit $X_n\sim\mathscr{B}(n,\lambda/n)$ et $Y\sim\mathscr{P}(\lambda)$. Alors :

$$\text{pour tout } k\in\mathbb{N},\qquad \mathbb{P}(X_n=k)\xrightarrow[n\to\infty]{}\mathbb{P}(Y=k).$$

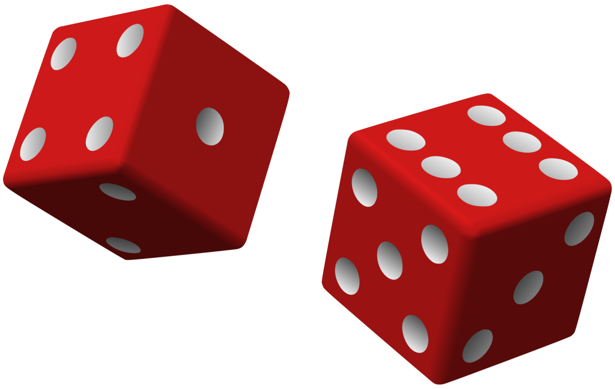
Démonstration. On fixe un $k \in \mathbb{N}$. Soit $n \geqslant k$,

$$\begin{aligned}
\mathbb{P}(X_n = k) &= C_n^k p^k (1-p)^{n-k} \\
&= \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&= \frac{\lambda^k}{k!} \frac{n(n-1)\cdots(n-k+1)}{n^k} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}
\end{aligned}$$

Pour λ et k fixés, on regarde la limite quand $n \to +\infty$:

$$\frac{n(n-1)\cdots(n-k+1)}{n^k} \longrightarrow 1, \quad \left(1 - \frac{\lambda}{n}\right)^n \longrightarrow e^{-\lambda} \quad \left(1 - \frac{\lambda}{n}\right)^{-k} \longrightarrow 1,$$

et donc $\mathbb{P}(X_n = k) \longrightarrow e^{-\lambda} \frac{\lambda^k}{k!} = \mathbb{P}(Y = k)$. ∎



5. Espérance et Variance

1 L'espérance d'une v.a.

1.A. C'est quoi ?

Définition 5.1 *L'espérance* (anglais : expected value) d'une v.a. discrète X est le nombre

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x \cdot \mathbb{P}(X = x).$$

R Idée : c'est la moyenne ! ! Quand vous entendez "espérance", pensez "moyenne" ! On verra plus tard une explication possible du fait qu'on dit espérance plutôt que la moyenne.

R (Interprétation) Une v.a. est "une variable qui ne prend pas une valeur numérique fixe, mais différentes valeurs avec des probas différentes". L'espérance est la moyenne de toutes ces valeurs, pondérées selon leur probabilité.

R Interprétation en termes d'argent : on me propose le jeu suivant. D'abord je paye un prix de participation de A Euros. Ensuite on tire un nombre réel selon une certaine loi décidée à l'avance. Si le nombre tiré est x, alors je reçois x Euros. Question : jusqu'à quel prix A est-ce que je devrais accepter ? Réponse : à long terme, si la participation coûte moins que (l'espérance de la loi) Euros, alors je fais un profit.

Par définition, l'espérance d'une v.a. ne dépend que de sa loi. L'interprétation est peut-être plus claire avec une autre formule :

Proposition 5.2 Si Ω est fini ou dénombrable, alors

$$\mathbb{E}(X) = \sum_{\omega\in\Omega} \mathbb{P}(\omega)\cdot X(\omega).$$

Exemple 5.3 Pour un groupe Ω de n adultes on veut calculer la moyenne d'âge.
Soit $X : \Omega \to \mathbb{N}$, $\omega \mapsto \text{age}(\omega)$.

$$\begin{aligned}
\text{Moyenne d'age} &= \frac{\sum_{\omega\in\Omega} X(\omega)}{n} \\
&= \sum_{\omega\in\Omega} \frac{1}{n}\cdot X(\omega) \qquad \text{(ecriture Prop 5.2)} \\
&= \sum_{\alpha=18,19,\ldots,150} \alpha \cdot \underbrace{\mathbb{P}(X=\alpha)}_{\text{(Proportion de personnes ayant exactement } \alpha \text{ annees)}} .
\end{aligned}$$

(ecriture Def 5.1)

Démonstration.

$$\sum_{\omega\in\Omega} \mathbb{P}(\omega)\cdot X(\omega) = \sum_{x\in X(\Omega)} \sum_{\substack{\omega\in\Omega\\ X(\omega)=x}} \mathbb{P}(\omega)\cdot x = \sum_{x\in X(\Omega)} x \sum_{\substack{\omega\in\Omega\\ X(\omega)=x}} \mathbb{P}(\omega) = \sum_{x\in X(\Omega)} x\,\mathbb{P}(X=x)$$

■

Exemple 5.4 (Espérance infinie)
Sur $\Omega = \{1,2,3,4,\ldots\}$ soit $\mathbb{P}(k) = \frac{6}{\pi^2}\cdot\frac{1}{k^2}$.

- C'est une proba - en effet, on peut montrer que $1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\ldots = \frac{\pi^2}{6}$.
- Soit $X(k) = k$. Cette v.a. n'a pas d'espérance, car :

$$\sum_{k=1}^{\infty} k\cdot\frac{6}{\pi^2}\cdot\frac{1}{k^2} = \sum_{k=1}^{\infty}\frac{6}{\pi^2}\cdot\frac{1}{k} = \frac{6}{\pi^2}\cdot\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\ldots\right)$$

est une série divergente.

Définition 5.5 On dit qu'une v.a. discrète *admet une espérance* (est *intégrable*) lorsque : la série $\sum_{x\in X(\Omega)} |x|\cdot\mathbb{P}(X=x)$ converge, on note alors :

$$\mathbb{E}(|X|) = \sum_{x\in X(\Omega)} |x|\cdot\mathbb{P}(X=x) < \infty.$$

1.B. Propriétés

Proposition 5.6 (Propriétés de l'espérance) Soit Ω muni d'une probabilité $\mathbb{P}$. Soient X, Y deux variables aléatoires définies sur Ω . Alors :

1. Addition : $\mathbb{E}(X+Y) = \mathbb{E}(X)+\mathbb{E}(Y)$.
2. Multiplication par un réel : $\forall b \in \mathbb{R}, \qquad \mathbb{E}(b\cdot X) = b\cdot\mathbb{E}(X)$.

3. Addition d'un réel : $\forall a \in \mathbb{R}, \qquad \mathbb{E}(X+a) = \mathbb{E}(X) + a.$

Démonstration. On suppose Ω fini ou dénombrable pour simplifier.

1. *Addition.*

$$\begin{aligned}\mathbb{E}(X+Y) &= \sum_{\omega\in\Omega} \Big(X(\omega)+Y(\omega)\Big)\mathbb{P}(\omega)\\ &= \sum_{\omega\in\Omega} X(\omega)\mathbb{P}(\omega) + \sum_{\omega\in\Omega} Y(\omega)\mathbb{P}(\omega)\\ &= \mathbb{E}(X)+\mathbb{E}(Y)\end{aligned}$$

3. *Multiplication par un réel.*

$$\mathbb{E}(bX) = \sum_{\omega\in\Omega} bX(\omega)\mathbb{P}(\omega) = b\sum_{\omega\in\Omega} X(\omega)\mathbb{P}(\omega) = b\cdot\mathbb{E}(X)$$

2. *Addition d'un réel.*

$$\mathbb{E}(X+a) = \mathbb{E}(X)+\mathbb{E}(a) = \mathbb{E}(X)+a\mathbb{E}(1) = \mathbb{E}(X)+a$$

■

Théorème 5.7 (Théorème de transfert - cas discret) Soit Ω muni d'une probabilité $\mathbb{P}$, X une v.a. discrète et $h:\mathbb{R}\to\mathbb{R}$ une fonction réelle . Alors :

$$\mathbb{E}\big(h(X)\big) \;=\; \sum_{x\in X(\Omega)} h(x)\cdot\mathbb{P}(X=x).$$

Démonstration.

$$\begin{aligned}\sum_{x\in X(\Omega)} h(x)\cdot\mathbb{P}(X=x) \;&=\; \sum_{y\in h\circ X(\Omega)} \quad \sum_{\substack{x\in X(\Omega);\\ h(x)=y}} y\cdot\mathbb{P}(X=x)\\ &=\; \sum_{y\in h\circ X(\Omega)} y \sum_{\substack{x\in X(\Omega);\\ h(x)=y}} \mathbb{P}(X=x)\\ &=\; \sum_{y\in h\circ X(\Omega)} y\,\mathbb{P}(h(X)=y)\\ &=\; \mathbb{E}(h\circ X)\end{aligned}$$

■

Exemple 5.8 Le théorème de transfert appliqué à :

1. $h(x)=x^2$ donne :

$$\mathbb{E}\big(X^2\big) = \sum_{x\in X(\Omega)} x^2\cdot\mathbb{P}(X=x).$$

2. $h(x)=x^3$ donne :

$$\mathbb{E}\big(X^3\big) = \sum_{x\in X(\Omega)} x^3\cdot\mathbb{P}(X=x).$$

3. $h(x) = e^x$ donne :
$$\mathbb{E}\left(e^X\right) = \sum_{x \in X(\Omega)} e^x \cdot \mathbb{P}(X = x).$$

2 La variance d'une v.a.

2.A. C'est quoi ?

Définition 5.9 Soit X une variable aléatoire sur Ω et supposons que $\mathbb{E}(X)$ existe. La *variance* de X est le nombre
$$\mathbb{V}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right)$$
L'écart-type (anglais : standard deviation) de X est
$$\sigma(X) = \sqrt{\mathbb{V}(X)}$$

R Interprétation : $\mathbb{V}(X)$ petit $\Longrightarrow$ valeurs de X concentrées autour de $\mathbb{E}(X)$. $\mathbb{V}(X)$ grand $\Longrightarrow$ valeurs de X éparpillées.

R Si X est discrète, alors
$$\mathbb{V}(X) = \sum_{x \in X(\Omega)} \left(x - \mathbb{E}(X)\right)^2 \mathbb{P}(X = x) = \sum_{\omega \in \Omega} \left(X(\omega) - \mathbb{E}(X)\right)^2 \mathbb{P}(\omega)$$

2.B. Propriétés

Proposition 5.10 Soit $X : \Omega \to \mathbb{R}$ une v.a. Alors
1. $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$
2. $\mathbb{V}(X) \geqslant 0$
3. $\mathbb{V}(X + a) = \mathbb{V}(X)$ pour $a \in \mathbb{R}$
4. $\mathbb{V}(\lambda X) = \lambda^2 \mathbb{V}(X)$ pour $\lambda \in \mathbb{R}$
5. $\sigma(\lambda X) = |\lambda| \cdot \sigma(X)$

Démonstration. Il suffit d'utiliser la Proposition 5.6, pour (1) :
1.
$$\begin{aligned}
\mathbb{V}(X) &= \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \mathbb{E}\left(X^2 - 2\mathbb{E}(X) \cdot X + \mathbb{E}(X)^2\right) \\
&= \mathbb{E}(X^2) - \mathbb{E}(2\mathbb{E}(X) \cdot X) + \mathbb{E}(\mathbb{E}(X)^2) \\
&= \mathbb{E}(X^2) - 2\mathbb{E}(X) \cdot \mathbb{E}(X) + (\mathbb{E}(X))^2 \\
&= \mathbb{E}(X^2) - (\mathbb{E}(X))^2
\end{aligned}$$
2. La v.a. $X - \mathbb{E}(X)$ est positive donc $\mathbb{V}(X) \geqslant 0$.
3. De l'égalité :
$$(X + a) - \mathbb{E}(X + a) = X + a - \mathbb{E}(X) - a = X - \mathbb{E}(X)$$
On obtient :
$$\mathbb{V}(X + a) = \mathbb{V}(X).$$

4. De l'égalité :
$$(\lambda X) - \mathbb{E}(\lambda X) = \lambda(X - \mathbb{E}(X))$$
On obtient :
$$\mathbb{V}(\lambda X) = \lambda^2 \mathbb{V}(X).$$

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

5. Conséquence directe de (4).

■

3 Espérance et variance des variables aléatoires discrètes usuelles

Théorème 5.11 L'espérance et la variance des v.a. discrètes précédemment décrites sont dans le Tableau 5.1.

Démonstration.

- Soit X une v.a. de loi uniforme $X \sim \mathscr{U}([\![1,n]\!])$. On met dans un coin de sa tête, les deux formules suivantes :

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2} \qquad \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$

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

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

Ainsi :

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

- Soit X une v.a. de loi de Bernoulli $X \sim \mathscr{B}(1,p)$. Alors

$$\mathbb{E}(X) = 1 \cdot \mathbb{P}(X=1) + 0 \cdot \mathbb{P}(X=0) = 1 \cdot p + 0 \cdot (1-p) = p,$$

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = 1^2 \cdot \mathbb{P}(X=1) + 0^2 \cdot \mathbb{P}(X=0) - p^2 = p - p^2 = p \cdot (1-p).$$

- Soit X une v.a. de loi binomiale $X \sim \mathscr{B}(n,p)$. Notons $q = 1-p$. Alors

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k=0}^{n} k \cdot \mathbb{P}(X=k) \\
&= \sum_{k=0}^{n} k \cdot \binom{k}{n} p^k q^{n-k} \\
&= \sum_{k=0}^{n} k \cdot \frac{n!}{k!(n-k)!} p^k q^{n-k} \\
&= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} p^{k-1} q^{(n-1)-(k-1)} \\
&= np \sum_{i=0}^{n-1} \frac{(n-1)!}{i!((n-1)-i)!} p^i q^{(n-1)-i} \\
&= np \sum_{i=0}^{n-1} \binom{i}{n-1} p^i q^{(n-1)-i} \\
&= np(p+q)^{n-1} \\
&= np
\end{aligned}$$

- Soit X une v.a. de poisson $X \sim \mathscr{P}(\lambda)$. Alors

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k=0}^{\infty} k \cdot \mathbb{P}(X=k) \\
&= \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} \\
&= \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} \\
&= \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} \\
&= \lambda \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda} \\
&= \lambda.
\end{aligned}$$

- Soit X une v.a. de loi géométrique $X \sim \mathscr{G}(p)$. Notons $q = 1-p$. On pose :

$$f(q) = \sum_{k=0}^{\infty} q^k = 1 + q + q^2 + q^3 + \cdots + q^k + \cdots = \frac{1}{1-q}$$

On a (il faudrait le justifier) :

$$f'(q) = \sum_{k=1}^{\infty} k q^{k-1} = 0 + 1 + 2q + 3q^2 + \cdots + k q^{k-1} + \cdots = \left(\frac{1}{1-q}\right)' = \frac{1}{(1-q)^2}$$

$$\begin{aligned}\mathbb{E}(X) &= \sum_{k=1}^{\infty} kq^{k-1}p = p \cdot \sum_{k=1}^{\infty} kq^{k-1} = pf'(q)\\ &= p \cdot \frac{1}{(1-q)^2} = p \cdot \frac{1}{p^2}\\ &= \frac{1}{p}\end{aligned}$$

- Le reste du tableau sera admis.

■

TABLE 5.1 – **Lois discrètes classiques**

Dénomination	**Loi**	**Espérance**	**Variance**
Loi Uniforme $X \sim \mathcal{U}([\![1,n]\!])$	$X(\Omega) = [\![1,n]\!]$ $\mathbb{P}(X=k) = \frac{1}{n}$	$\mathbb{E}(X) = \frac{n+1}{2}$	$\mathbb{V}(X) = \frac{n^2-1}{12}$
Loi de Bernoulli $X \sim \mathcal{B}(1,p)$	$X(\Omega) = \{0,1\}$ $\mathbb{P}(X=0) = 1-p$ $\mathbb{P}(X=1) = p$	$\mathbb{E}(X) = p$	$\mathbb{V}(X) = p(1-p)$
Loi Binomiale $X \sim \mathcal{B}(n,p)$	$X(\Omega) = [\![0,n]\!]$ $\mathbb{P}(X=k) = C_n^k p^k (1-p)^{n-k}$	$\mathbb{E}(X) = np$	$\mathbb{V}(X) = np(1-p)$
Loi Géométrique $X \sim \mathcal{G}(p)$	$X(\Omega) = [\![1,+\infty[\![$ $\mathbb{P}(X=k) = p(1-p)^{k-1}$	$\mathbb{E}(X) = \frac{1}{p}$	$\mathbb{V}(X) = \frac{1-p}{p^2}$
Loi de Poisson $X \sim \mathcal{P}(\lambda)$	$X(\Omega) = [\![0,+\infty[\![$ $\mathbb{P}(X=k) = e^{-\lambda}\frac{\lambda^k}{k!}$	$\mathbb{E}(X) = \lambda$	$\mathbb{V}(X) = \lambda$

Les cases en rouge seront rappelées pendant les examens, les cases non-coloriés sont à savoir.

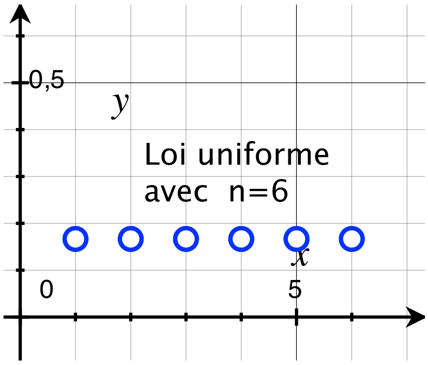
4 La médiane

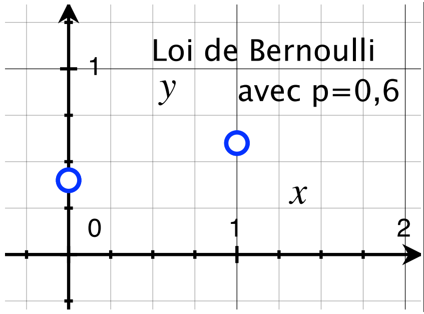
Définition 5.12 Une *médiane* d'une variable aléatoire X est un nombre réel m satisfaisant :

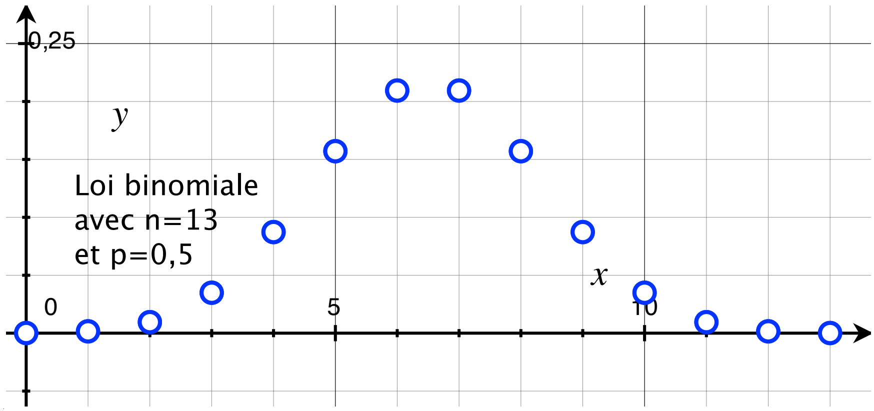
$$\mathbb{P}(X \leqslant m) \geqslant \frac{1}{2} \quad \text{et} \quad \mathbb{P}(X \geqslant m) \geqslant \frac{1}{2}$$

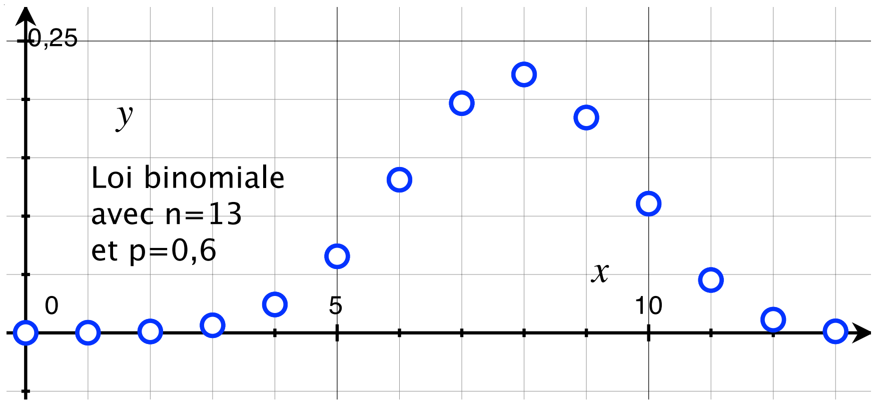
R La médiane ne dépend que de la loi de X. Il peut y avoir plusieurs valeurs possibles pour m.

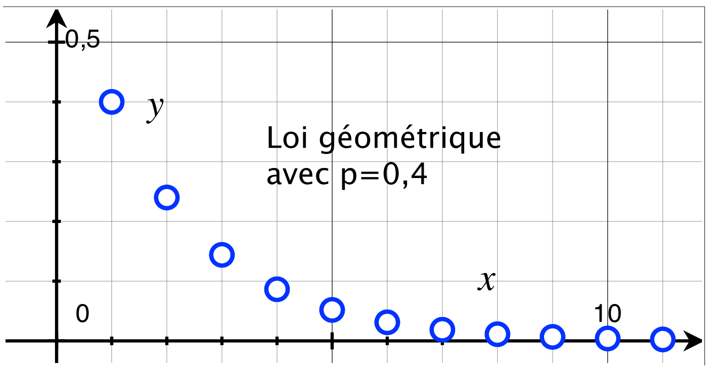
Exemple 5.13 Pour un dé équilibré, tout nombre entre 3 et 4 est une médiane. (En pratique, prendre 3,5, le milieu de l'intervalle.)

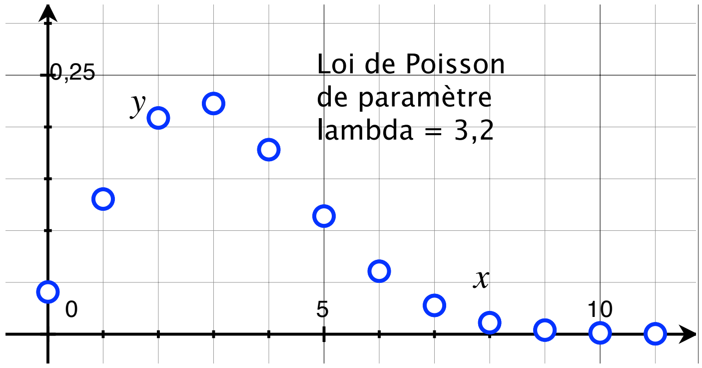

0,5
y
Loi uniforme
avec n=6
x
0
5

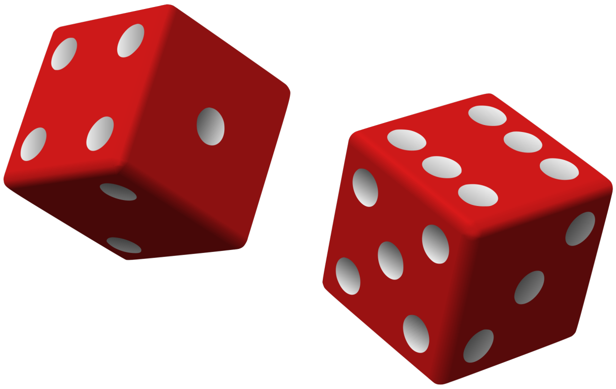

Loi de Bernoulli
1
y
avec p=0,6
x
0
1
2


0,25
y
Loi binomiale
avec n=13
et p=0,5
x
0
5
10


0,25
y
Loi binomiale
avec n=13
et p=0,6
x
0
5
10


0,5
y
Loi géométrique
avec p=0,4
x
0
10


0,25
y
Loi de Poisson
de paramètre
lambda = 3,2
x
0
5
10



6. Variables aléatoires continues

1 C'est quoi ?

Si vous me dites : "Choisissez un nombre réel aléatoire", je ne sais pas quoi faire. Vous devez spécifier selon quelle *loi* je dois tirer le nombre.

Jusqu'ici nous avons vu des lois *discrètes*, c.à.d. les nombres étaient tirés parmi un sous-ensemble discret de $\mathbb{R}$ (typiquement $\mathbb{Z}$ ou $\mathbb{N}$).

Regardons maintenant l'autre extrême : tous les nombres appartenant à un certain intervalle (par exemple à $[0,1]$ ou à $]-\infty,+\infty[=\mathbb{R}$) peuvent être tirés, mais chaque nombre individuellement apparait avec probabilité 0.

Notation. *(Rappel) Si $I \subset \mathbb{R}$ est un sous-ensemble, et X une v.a., alors*

$$\mathbb{P}(X \in I) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in I\})$$

Définition 6.1 Une v.a. $X : \Omega \to \mathbb{R}$ est *continue* s'il existe une fonction $f : \mathbb{R} \to \mathbb{R}_+$ telle que pour tout intervalle $I \subset \mathbb{R}$,

$$\mathbb{P}(X \in I) = \int_I f(x)\, \mathrm{d}x$$

La fonction f s'appelle alors la *densité* de X (ou la densité de la loi de X).

Définition 6.2 La *loi* d'une v.a. continue $X : \Omega \to \mathbb{R}$ est la probabilité $\mathbb{P}_X$ sur $\mathbb{R}$ déterminée par

$$\mathbb{P}_X(I) = \mathbb{P}(X \in I) \quad \text{pour tout intervalle} \quad I \subset \mathbb{R}$$

R Une fonction $f : \mathbb{R} \to \mathbb{R}$ qui est la densité d'une loi satisfait deux propriétés :

1. $f(x) \geqslant 0$ pour tout $x \in \mathbb{R}$,
2. $\int_{-\infty}^{+\infty} f(x) = \mathbb{P}(X \in]-\infty, +\infty[) = \mathbb{P}(X \in \mathbb{R}) = 1$ (en particulier, f est intégrable)

Réciproquement, on peut montrer que chaque fonction satisfaisant (1) & (2) est la densité d'une loi de probabilité.

R Si X est une variable aléatoire continue alors :

$$\forall a \in \mathbb{R}, \qquad \mathbb{P}(X = a) = 0.$$

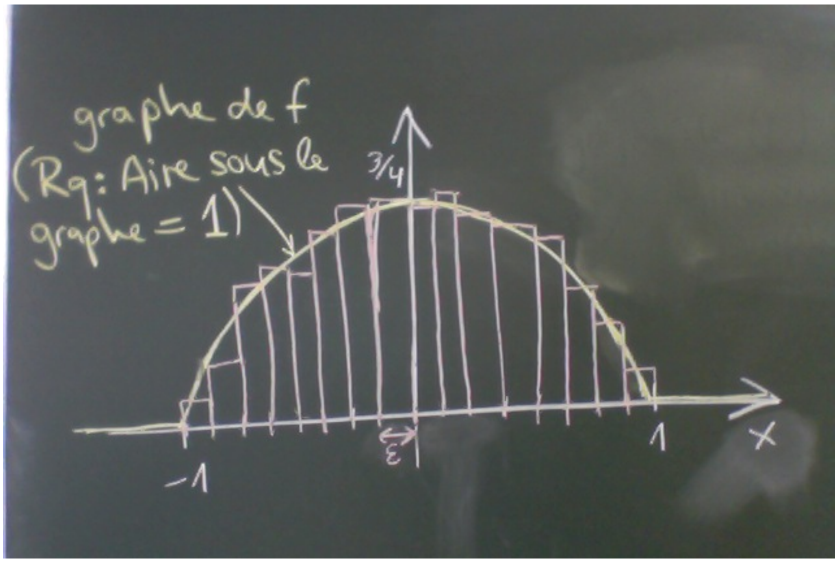
En effet,

$$\mathbb{P}(X = a) = \int_a^a f(x)dx = 0.$$

Exemple 6.3

$$f(x) = \begin{cases} \frac{3}{4}(1-x^2) & \text{si } x \in [-1,1] \\ 0 & \text{sinon} \end{cases} = \frac{3}{4}(1-x^2)\mathbb{1}_{[0,1]}(x)$$

Si l'on sélectionne N points de $\mathbb{R}$ au hasard, tirés selon la loi de densité f, alors le nuage de points est le plus dense en 0, et aucun point en dehors de $[-1,1]$. Dessinons un histogramme : on partitionne $\mathbb{R}$ en boîtes de largeur ϵ, et on compte quelle boîte a reçu combien de points. Si dans une certaine boîte il y a k points, on dessine une barre de hauteur $\frac{k}{\epsilon \cdot N}$. Alors l'histogramme va ressembler au graphe de f.



Densité d'une variable aléatoire X (jaune) et un histogramme (rouge)

Définition 6.4 Soit E un ensemble et A un sous-ensemble $A \subset E$ alors on note : $\mathbb{1}_A : E \to \mathbb{R}$ définie par :

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{cases}.$$

Exemple 6.5 La durée de vie (en minutes) d'une particule élémentaire d'une certaine substance radioactive peut être modélisé par une variable aléatoire de densité

$$f(x) = \begin{cases} \frac{1}{100} e^{-\frac{x}{100}} & \text{si } x \geqslant 0 \\ 0 & \text{si } x < 0 \end{cases} = \frac{1}{100} e^{-\frac{x}{100}} \mathbb{1}_{\mathbb{R}_+}(x)$$

Question : quelle est la probabilité que cette durée soit comprise entre 50 et 150 minutes ? Réponse :

$$\mathbb{P}(50 \leqslant X \leqslant 150) = \int_{50}^{150} f(x)\,dx = \int_{50}^{150} \frac{1}{100} e^{-x/100}\,dx =$$
$$= \left[-e^{-x/100}\right]_{x=50}^{x=150} = e^{-1/2} - e^{-3/2} \approx 0,384$$

Question : quelle est la probabilité que la durée soit inférieure ou égale à 300 minutes ? Réponse :

$$\mathbb{P}(X \leqslant 300) = \int_{-\infty}^{300} f(x)\,dx = \int_{0}^{300} \frac{1}{100} e^{-x/100}\,dx =$$
$$= \left[-e^{-x/100}\right]_{x=0}^{x=300} = e^{0} - e^{-\frac{300}{100}} = 1 - e^{-3} \approx 0,95$$

Plus généralement, le calcul précédent démontre : si $\lambda \in \mathbb{R}_+$ est un nombre positif et X est une v.a. de densité

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geqslant 0 \\ 0 & \text{si } x < 0 \end{cases} = \lambda e^{-\lambda x} \mathbb{1}_{\mathbb{R}_+}(x)$$

alors

$$\mathbb{P}(X \leqslant t) = \begin{cases} 1 - e^{-\lambda t} & \text{si } t \geqslant 0 \\ 0 & \text{si } t < 0 \end{cases} = \left(1 - e^{-\lambda t}\right) \mathbb{1}_{\mathbb{R}_+}(x)$$

2 Fonction de répartition

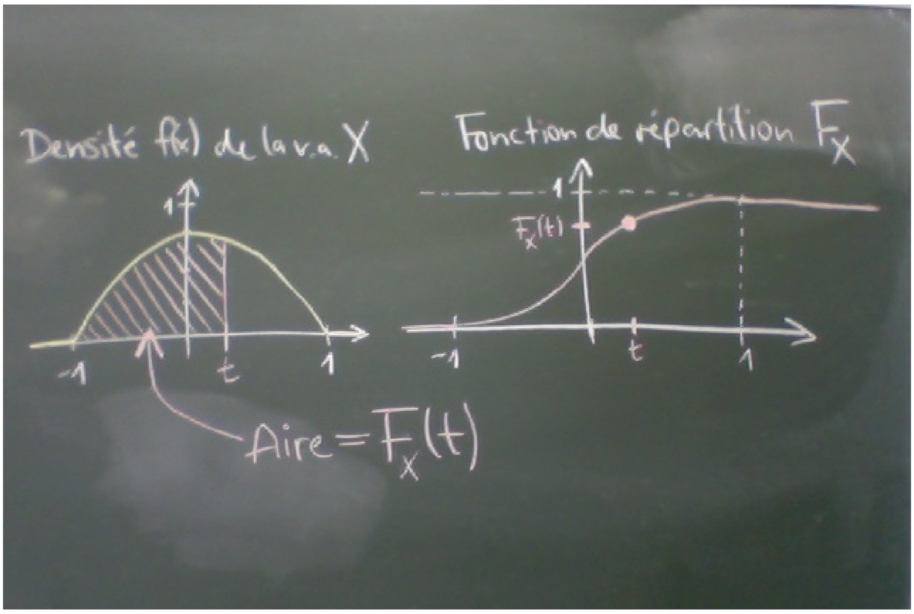
Définition 6.6 (Rappel) La *fonction de répartition* F_X d'une variable aléatoire X est la fonction réelle définie par :

$$F_X(t) = \mathbb{P}(X \leqslant t)$$

Donc si X est de densité f, alors

$$F_X(t) = \mathbb{P}(X \leqslant t) = \int_{-\infty}^{t} f(x)\,dx = \text{Aire sous le graphe de } f \text{ à gauche de la droite } x = t$$

R Si f est la densité de la v.a. X, alors $F_X(t) = \int_{-\infty}^{t} f(x)\,dx$, donc F_X est dérivable par morceaux et $F'_X = f$ et donc F_X est une primitive de f.



La densité de l'Exemple 6.3 et sa fonction de répartition F_X.

3 Médiane, espérance et variance

Définition 6.7 Soit X une variable aléatoire continue X de densité f. Alors une *médiane* de X est un nombre m tel que $\mathbb{P}(X \leqslant m) = \frac{1}{2}$ (et donc $\mathbb{P}(X \geqslant m) = \frac{1}{2}$), c.à.d., $F_X(m) = \frac{1}{2}$. (Géométriquement, l'aire sous f à gauche de m doit être égal à l'aire sous f à droite de m.)

Définition 6.8 L'*espérance* $\mathbb{E}(X)$ d'une variable aléatoire continue X de densité f est le nombre

$$\mathbb{E}(X) \stackrel{\text{def}}{=} \int_{-\infty}^{+\infty} x \cdot f(x)\, dx$$

Définition 6.9 Soit X une variable aléatoire. Notons μ son espérance : $\mu = \mathbb{E}(X)$. Alors la *variance* $\mathbb{V}(X)$ de X est le nombre

$$\mathbb{V}(X) \stackrel{\text{def}}{=} \mathbb{E}((X-\mu)^2) = \int_{-\infty}^{+\infty} (x-\mu)^2 \cdot f(x)\, dx$$

L'*écart-type* de X est

$$\sigma(X) \stackrel{\text{def}}{=} \sqrt{\mathbb{V}(X)}$$

Les analogues des Propositions 5.6 et 5.10 sont valables pour les v.a. continues (avec des démonstrations semblables).

Proposition 6.10 (Propriétés de l'espérance, de la variance et de l'écart-type) Soit Ω muni d'une probabilité $\mathbb{P}$. Soient X, Y deux variables aléatoires définies sur Ω . Alors :

1. Linéarité : $\forall \lambda, \mu \in \mathbb{R}, \qquad \mathbb{E}(\lambda X + \mu Y) = \lambda \mathbb{E}(X) + \mu \mathbb{E}(Y)$.
2. Addition d'un réel : $\forall a \in \mathbb{R}, \qquad \mathbb{E}(X + a) = \mathbb{E}(X) + a$.

3. $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.
4. $\mathbb{V}(X) \geqslant 0$ et si $\mathbb{V}(X) = 0$ alors $X = \mathbb{E}(X)$.
5. $\forall a \in \mathbb{R}, \qquad \mathbb{V}(X+a) = \mathbb{V}(X)$.
6. $\forall \lambda \in \mathbb{R}, \qquad \mathbb{V}(\lambda X) = \lambda^2 \mathbb{V}(X)$.
7. $\forall \lambda \in \mathbb{R}, \qquad \sigma(\lambda X) = |\lambda| \cdot \sigma(X)$

Théorème 6.11 (Théorème de transfert)
Si X est un variable aléatoire continue de densité f alors :

$$\text{Pour toute fonction } h : \mathbb{R} \to \mathbb{R}, \qquad \mathbb{E}(h(X)) = \int_{\mathbb{R}} h(x) f(x) dx.$$

4 Lois continues classiques

Voir le Tableau 6.2.

4.A. La loi uniforme (continue)

Définition 6.12 (**La loi uniforme**)
Une variable aléatoire X suit la *loi uniforme* sur l'intervalle $[a,b]$, si la densité est donnée par

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leqslant x \leqslant b \\ 0 & \text{sinon} \end{cases} = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x)$$

Notation $X \sim \mathcal{U}(a,b)$.
Dessin de la densité et de la fonction de répartition $F_X(t)$.

Proposition 6.13 Soit X une v.a. continue qui suit la loi uniforme sur $[a,b]$.

- Si $[c,d] \subset [a,b]$, alors
$$\mathbb{P}(X \in [c,d]) = \int_c^d \frac{1}{b-a}\, dx = \frac{d-c}{b-a}$$
= la proportion de l'intervalle $[a,b]$ recouvert par $[c,d]$.
- $\mathbb{E}(X) = \frac{b+a}{2}$, donc la valeur moyenne est le milieu de l'intervalle $[a,b]$.
- $\mathbb{V}(X) = \frac{(b-a)^2}{12}$.

Démonstration.

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{1}{b-a} \int_a^b x dx = \frac{1}{b-a} \left[\frac{x^2}{2} \right]_{x=a}^{x=b} = \frac{b+a}{2},$$

$$\begin{aligned}
\mathbb{E}(X^2) &= \int_{-\infty}^{+\infty} x^2 \cdot f(x)dx \\
&= \int_a^b \frac{x^2}{b-a}\,dx \\
&= \left[\frac{x^3}{3(b-a)}\right]_{x=a}^{x=b} \\
&= \frac{b^3-a^3}{3(b-a)} \\
&= \frac{(b-a)(a^2+ab+b^2)}{3(b-a)} \\
&= \frac{a^2+ab+b^2}{3}
\end{aligned}$$

$$\begin{aligned}
\mathbb{V}(X) &= \mathbb{E}(X^2)-\mathbb{E}(X)^2 \\
&= \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2 \\
&= \frac{4(a^2+ab+b^2)-3(a^2+2ab+b^2)}{12} \\
&= \frac{a^2+b^2-2ab}{12} \\
&= \frac{(b-a)^2}{12}
\end{aligned}$$

■

Exemple 6.14 (La loi uniforme) Idée : il pleut sur la droite réelle $\mathbb{R}$. Je note les positions exactes des gouttes qui tombent sur l'intervalle $[0,1]$. Elles seront distribués selon une loi uniforme (même densité partout) sur $[0,1]$.

4.B. La loi exponentielle

Définition 6.15 **(La loi exponentielle)**
Une variable aléatoire X suit la *loi exponentielle de paramètre* λ si la densité est donnée par

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geqslant 0 \\ 0 & \text{sinon} \end{cases} = \lambda e^{-\lambda x}\mathbb{1}_{\mathbb{R}_+}(x)$$

Notation $X \sim \mathscr{E}(\lambda)$.

Proposition 6.16 Soit X une v.a. continue qui suit la loi exponentielle de paramètre λ.

- Fonction de répartition : $F_X(t) = \mathbb{P}(X \leqslant t) = \begin{cases} 1-e^{-\lambda t} & \text{si } t \geqslant 0 \\ 0 & \text{sinon} \end{cases}$.
- Espérance : $\mathbb{E}(X) = \frac{1}{\lambda}$.
- Variance :$\mathbb{V}(X) = \frac{1}{\lambda^2}$.

Démonstration.

- Dém. : Exemple 6.5.

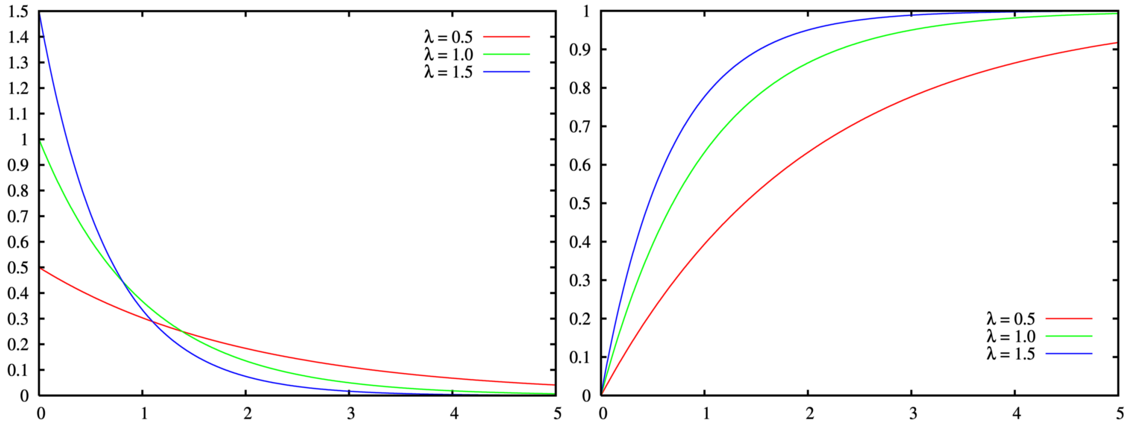


FIGURE 6.1 – Densité et fonction de répartition de la loi exponentielle (Source : wikipedia)

- $$\begin{aligned}\mathbb{E}(X) &= \int_{-\infty}^{0} x \cdot 0 \, dx + \int_{0}^{+\infty} x \cdot \lambda e^{-\lambda x} \\ &= \lambda \cdot \int_{0}^{+\infty} x e^{-\lambda x} \, dx \\ &\overset{(*)}{=} \lambda \cdot \left(\left[\frac{x}{-\lambda} e^{-\lambda x} \right]_{x=0}^{x=+\infty} - \int_{0}^{+\infty} \frac{1}{-\lambda} e^{-\lambda x} \right) \\ &= \lambda \cdot \left(0 - \left[\frac{1}{\lambda^2} e^{-\lambda x} \right]_{x=0}^{x=+\infty} \right) \\ &= \lambda \cdot \frac{1}{\lambda^2} = \frac{1}{\lambda}\end{aligned}$$

  où l'égalité $\overset{(*)}{=}$ découle par intégration par parties : $\int u\, v' = u \cdot v - \int u' v$ avec $u = x$ et $v' = e^{-\lambda x}$ (donc $u' = 1$, $v = \frac{1}{-\lambda} e^{-\lambda x}$).
- Exo.

■

Exemple 6.17 (Interprétation) La loi exponentielle est l'analogue continu de la loi géométrique. Rappel, la loi géométrique modélise le temps d'attente avant une victoire dans un jeu de pile ou face. Ce phénomène est sans mémoire puisque chaque nouveau lancer est indépendant du précédent.

La loi exponentielle modélise le temps de vie d'une particule radioactive, d'une batterie de voiture, d'un composant éléctronique ou le temps de passage d'un individu au comptoir de la poste, le temps d'une conversation téléphonique etc ...

Formellement, les phénomènes sans mémoire (à temps continu) sont modélisés par des lois exponentielles. Une v.a. X est *sans mémoire* lorsque :

$$\forall s, t > 0, \quad \mathbb{P}(X > t + s \,|\, X > t) = \mathbb{P}(X > s)$$

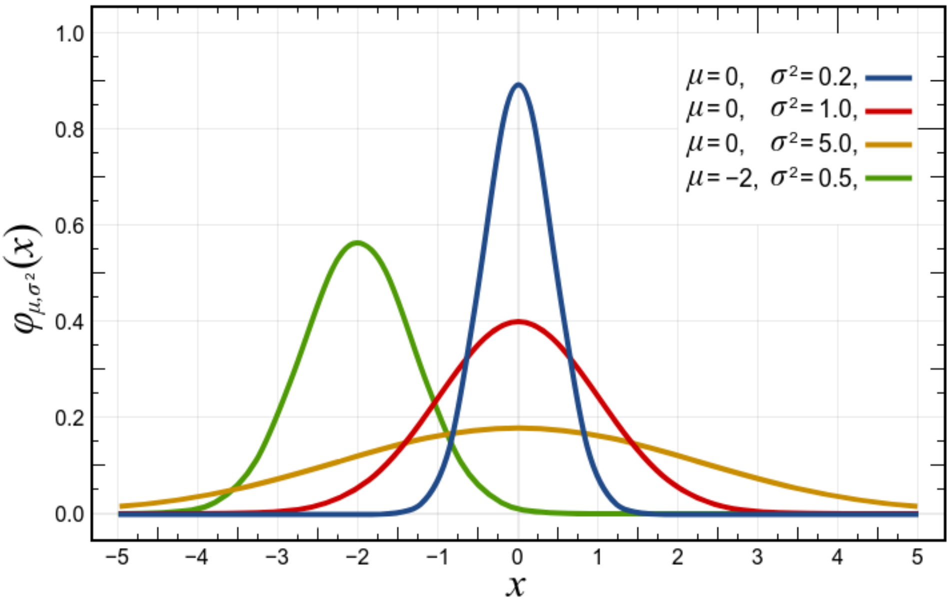


FIGURE 6.2 – Densité de la loi normale pour différents paramètres

qui se récrit :

$$\forall s,t>0,\quad \mathbb{P}(X>t+s)=\mathbb{P}(X>s)\mathbb{P}(X>t))$$

En pratique, cela signifie que si l'on observe qu'un composant électronique à **une durée de vie de 1 an en moyenne** alors la durée de vie de ce composant(en jour) suit une loi exponentielle de densité $f(x)=\frac{1}{365}e^{\frac{-x}{365}}\mathbb{1}_{\mathbb{R}_+}(x)$. On peut alors calculer aisément la probabilité que ce composant fonctionne pendant 2 ans, qui est donné par :

$$\mathbb{P}(\text{vie de comp. machin} > 2 \text{ ans}) = \int_{730}^{\infty}\frac{1}{365}e^{\frac{-x}{365}}dx = \left[-e^{\frac{-x}{365}}\right]_{730}^{\infty} = e^{-2}\approx 0.135$$

Exercice Montrer que la loi exponentielle est bien sans mémoire. ■

4.C. La loi normale

La loi la plus importante est la *loi normale* ou *loi gaussienne*. (Elle est si importante à cause du théorème central limite "TCL", qu'on verra à la fin du cours).

Définition 6.18 (**Loi normale**) Une variable aléatoire X suit la *loi normale d'espérance* $\mu\in\mathbb{R}$ *et de variance* $\sigma^2\in\mathbb{R}_+$ si la densité est donnée par

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Notation $X\sim\mathcal{N}(\mu,\sigma^2)$.

La *loi normale centrée réduite* est la loi $\mathcal{N}(0,1)$ (espérance 0, variance 1).

R On va admettre que :

$$\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx = 1$$

donc c'est bien une densité de probabilité.

Proposition 6.19 Soit X une v.a. continue qui suit la loi normale $\mathcal{N}(\mu,\sigma^2)$.

- $\mathbb{E}(X) = \mu$.
- $\mathbb{V}(X) = \sigma^2$.

Démonstration. On fait le cas $X \sim \mathcal{N}(0,1)$.

$$\begin{aligned}
\mathbb{E}(X) &= \int_{-\infty}^{+\infty} x f(x) dx \\
&= \int_{-\infty}^{+\infty} \underbrace{\frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}_{\text{fct impaire}} dx \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}(X^2) &= \int_{-\infty}^{+\infty} x^2 f(x) dx \\
&= \int_{-\infty}^{+\infty} \frac{x^2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
&= \int_{-\infty}^{+\infty} \underbrace{x}_{v} \underbrace{\frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}_{u'} dx \\
&= \left[x \frac{-1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
&= 0 - 0 + 1
\end{aligned}$$

■

Proposition 6.20 Soient $a \neq 0$ et $b \in \mathbb{R}$,

1. Si X suit une loi normale alors $aX + b$ suit une loi normale.
2. Si $X \sim \mathcal{N}(\mu,\sigma^2)$ alors $aX + b \sim \mathcal{N}(a\mu + b, (a\sigma)^2)$.
3. Si $X \sim \mathcal{N}(\mu,\sigma^2)$ alors $X^* = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$.

Démonstration. On ne démontre pas le premier point mais on remarque que si on pose :

$$f_{(\mu,\sigma^2)}(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad \text{et} \qquad f_{(0,1)}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Alors :

$$f_{(\mu,\sigma^2)}(x) = \frac{1}{\sigma} f_{(0,1)}\left(\frac{x-\mu}{\sigma}\right)$$

Pour le second point, le premier point montre que $aX + B$ suit une loi normale. Les propriétés de l'espérance et de la variance montrent que :

$$\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$$

$$\begin{aligned}
\mathbb{V}(aX + b) &= \mathbb{V}(aX) \\
\mathbb{V}(aX + b) &= \mathbb{V}(aX) \\
&= a^2\mathbb{V}(X)
\end{aligned}$$

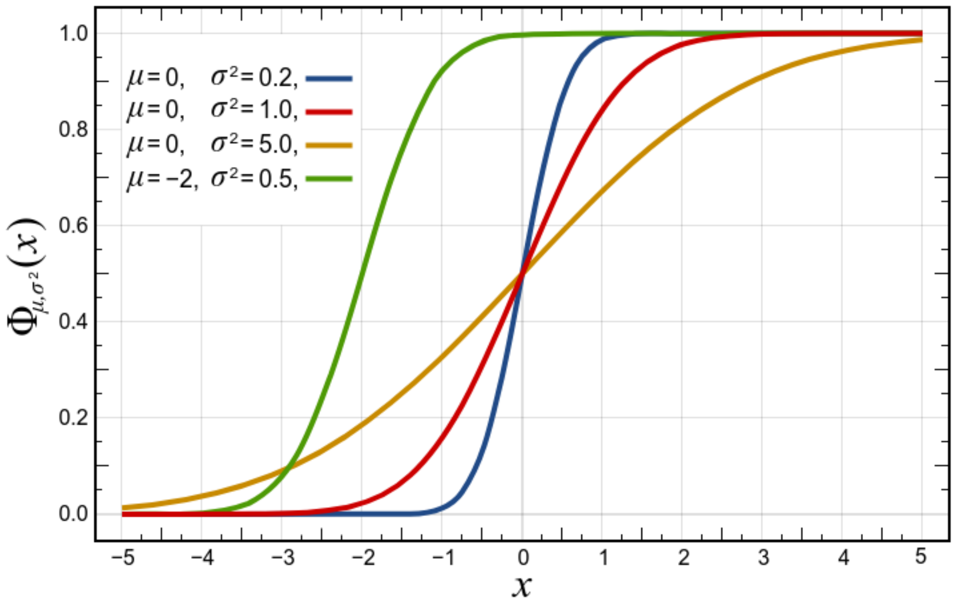


FIGURE 6.3 – Fonction de répartition de la loi normale pour différents paramètres

■

Définition 6.21 Notons $\Phi : \mathbb{R} \to \mathbb{R}_+$ la fonction de répartition d'une loi normale centrée réduite – donc

$$\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

R Il n'y a pas de *formule* pour $\Phi(t)$. Il n'y a que des tables. [des listes avec des valeurs approchées. Par exemple sur la version actuelle de la page Wikipédia sur la Loi Normale.]

Exemple 6.22 On suppose que la taille X, en centimètres, d'un homme âgé de 25 ans est une variable aléatoire normale de paramètres $\mu = 175$ et $\sigma^2 = 36$. Quel est le pourcentage d'hommes de 25 ans ayant une taille supérieure à 185 cm? (Vous avez le droit d'utiliser une table avec les valeurs de Φ.) Réponse : On cherche la probabilité

$$\mathbb{P}(X > 185) = 1 - \mathbb{P}(X \leqslant 185)$$

Nous avons vu en TD que la v.a. $X^* = \frac{X - \mathbb{E}(X)}{\sigma}$ est centrée réduite, càd :

$$\mathbb{E}(X^*) = 0 \qquad \text{et} \qquad \mathbb{V}(X^*) = 1$$

On vient aussi de voir que si X suit une loi normale et $a \neq 0$ et b quelconque alors $aX + b$ suit une loi normale donc :

$$X^* = \frac{X - 175}{6} \sim \mathcal{N}(0,1)$$

t	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	0,5	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,5279	0,53188	0,53586
0,1	0,53983	0,5438	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,6293	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,6591	0,66276	0,6664	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,7054	0,70884	0,71226	0,71566	0,71904	0,7224
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,7549
0,7	0,75804	0,76115	0,76424	0,7673	0,77035	0,77337	0,77637	0,77935	0,7823	0,78524
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,8665	0,86864	0,87076	0,87286	0,87493	0,87698	0,879	0,881	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,9032	0,9049	0,90658	0,90824	0,90988	0,91149	0,91309	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,9222	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,9452	0,9463	0,94738	0,94845	0,9495	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,9608	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,9732	0,97381	0,97441	0,975	0,97558	0,97615	0,9767
2	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,9803	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,983	0,98341	0,98382	0,98422	0,98461	0,985	0,98537	0,98574
2,2	0,9861	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,9884	0,9887	0,98899
2,3	0,98928	0,98956	0,98983	0,9901	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,9918	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,9943	0,99446	0,99461	0,99477	0,99492	0,99506	0,9952
2,6	0,99534	0,99547	0,9956	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,9972	0,99728	0,99736
2,8	0,99744	0,99752	0,9976	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861
3	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99896	0,999
3,1	0,99903	0,99906	0,9991	0,99913	0,99916	0,99918	0,99921	0,99924	0,99926	0,99929
3,2	0,99931	0,99934	0,99936	0,99938	0,9994	0,99942	0,99944	0,99946	0,99948	0,9995
3,3	0,99952	0,99953	0,99955	0,99957	0,99958	0,9996	0,99961	0,99962	0,99964	0,99965
3,4	0,99966	0,99968	0,99969	0,9997	0,99971	0,99972	0,99973	0,99974	0,99975	0,99976
3,5	0,99977	0,99978	0,99978	0,99979	0,9998	0,99981	0,99981	0,99982	0,99983	0,99983
3,6	0,99984	0,99985	0,99985	0,99986	0,99986	0,99987	0,99987	0,99988	0,99988	0,99989
3,7	0,99989	0,9999	0,9999	0,9999	0,99991	0,99991	0,99992	0,99992	0,99992	0,99992
3,8	0,99993	0,99993	0,99993	0,99994	0,99994	0,99994	0,99994	0,99995	0,99995	0,99995
3,9	0,99995	0,99995	0,99996	0,99996	0,99996	0,99996	0,99996	0,99996	0,99997	0,99997
4	0,99997	0,99997	0,99997	0,99997	0,99997	0,99997	0,99998	0,99998	0,99998	0,99998

TABLE 6.1 – Table de Φ

Donc

$$\begin{aligned}\mathbb{P}(X > 185) &= 1 - \mathbb{P}(X \leqslant 185)\\ &= 1 - \mathbb{P}(\tfrac{X-175}{6} \leqslant \tfrac{185-175}{6})\\ &= 1 - \mathbb{P}(X^* \leqslant \tfrac{5}{3})\\ &= 1 - \Phi(\tfrac{5}{3})\\ &\approx 1 - 0,9525 \text{ car } \frac{5}{3} \approx 1,67\\ &= 0,0475 = 4,75\%.\end{aligned}$$

Exemple 6.23 Soit X une variable aléatoire qui suit une loi normale de paramètres (μ,σ). Évaluer les probabilités :

$$\mathbb{P}(|X-\mu| \leqslant \sigma) \qquad \mathbb{P}(|X-\mu| \leqslant 2\sigma)$$

La loi $X^* = \frac{X-\mu}{\sigma}$ est une loi normale centrée réduite et

$$|X-\mu| \leqslant k\sigma \quad \Leftrightarrow \quad |X^*| \leqslant k$$

Donc,

$$\begin{aligned}\mathbb{P}(|X-\mu| \leqslant k\sigma) &= \mathbb{P}(|X^*| \leqslant k)\\ &= 1 - \mathbb{P}(|X^*| > k)\\ &= 1 - \mathbb{P}(|X^*| > k)\\ &= 1 - 2\mathbb{P}(X^* > k)\\ &= 1 - 2(1 - \mathbb{P}(X^* \leqslant k))\\ &= 2\mathbb{P}(X^* \leqslant k) - 1\\ &= 2\Phi(k) - 1\end{aligned}$$

Or, $\Phi(1) = 0,84134$ et $\Phi(2) = 0,97725$ donc :

$$\mathbb{P}(|X-\mu| \leqslant \sigma) \approx 0,68 \qquad \mathbb{P}(|X-\mu| \leqslant 2\sigma) \approx 0,95$$

Pour info, $2\Phi(2,6) - 1 \approx 0,99$ donc

$$\mathbb{P}(|X-\mu| \leqslant 2,6\sigma) \approx 0,99$$

4.D. D'autres lois

Exemple 6.24 On dit une variable aléatoire X est *log-normale* si $Y = \ln(X)$ est normale.

Exemple 6.25 Si $X_1,\ldots,X_k$ sont des v.a. **indépendantes**[a] normales centrées réduites ($X_i \sim \mathcal{N}(0,1)$ pour $i = 1,\ldots,k$), alors $Y = X_1^2 + \ldots + X_k^2$ suit une loi qu'on appelle la *loi du* χ^2 *("chi 2")* *avec k degrés de liberté*. Cette loi est très importante pour des tests statistiques. La densité est

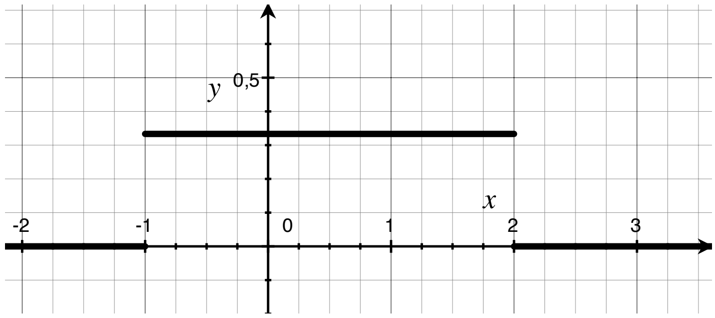
$$f(x) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}$$

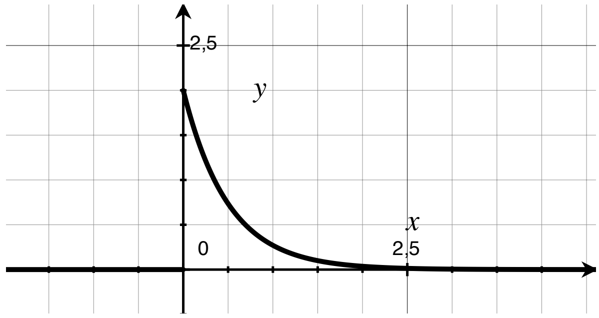
où Γ est la "fonction Γ" – c'est une certaine fonction continue qui interpole la fonction factorielle : $\Gamma(n+1) = n!$.

a. Pas encore défini.

TABLE 6.2 – **Lois continues classiques**

Dénomination	**Densité**	**Espérance**	**Variance**
Loi Uniforme $X \sim \mathscr{U}([a,b])$	$f(x) = \frac{1}{b-a}, \quad x \in [a,b]$ $f(x) = 0, \quad x \notin [a,b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Loi Exponentielle $X \sim \mathscr{E}(\lambda)$	$f(x) = \lambda e^{-\lambda x}, \quad x \geqslant 0$ $f(x) = 0, \quad x < 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Loi Normale $X \sim \mathscr{N}(\mu, \sigma^2)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \ \forall x \in \mathbb{R}$	μ	σ^2
Loi Log-Normale Paramètres μ, σ	$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}}$	$e^{\mu+\frac{\sigma^2}{2}}$	$(e^{\sigma^2}-1)e^{2\mu+\sigma^2}$
Loi χ^2 (Chi-deux) Paramètre $k \in \mathbb{N}$	$f(x) = \frac{1}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}, \ x \geqslant 0$ où Γ est la "fonction Gamma" $f(x) = 0, \quad x < 0$	k	$2k$
Loi Logistique Paramètres μ, s	$f(x) = \frac{e^{-(x-\mu)/s}}{s\left(1+e^{-(x-\mu)/s}\right)^2}$	μ	$\frac{\pi^2}{3}s^2$





(a) Loi uniforme $\mathscr{U}(-1,2)$. Sa densité est de $\frac{1}{3}$ sur l'intervalle $[-1,2]$ et de 0 en-dehors de cet intervalle. (b) Densité de la loi exponentielle $\mathscr{E}(2)$

5 Couple de v.a. conjointement continues

Définition 6.26 Les variables aléatoires X et Y sont dites *conjointement continues* s'il existe une fonction $f:\mathbb{R}^2 \to \mathbb{R}_+$ telle que :

$$\forall [a,b]\times[c,d] \subset \mathbb{R}^2, \quad \mathbb{P}((X,Y)\in[a,b]\times[c,d]) = \int_c^d \int_a^b f(x,y)dxdy$$

La fonction f s'appelle *la densité conjointe (ou simultanée)* de X et Y.

Proposition 6.27 Si les variables aléatoires X et Y sont conjointement continues et le couple (X,Y) est de densité $f:\mathbb{R}^2 \to \mathbb{R}$ alors X et Y sont continues de densité f_X et f_Y et on a :

$$f_X(x) = \int_{\mathbb{R}} f(x,y)dy \qquad \text{et} \qquad f_Y(y) = \int_{\mathbb{R}} f(x,y)dx$$

Exemple 6.28 On prend $f(x,y) = \frac{1}{\pi}\mathbb{1}_{\mathbb{D}}(x,y)$. On calcule, soit $y \in \mathbb{R}$. Si $y \notin [-1,1]$ alors pour tout $x \in \mathbb{R}$, $(x,y) \notin \mathbb{D}$ donc $f_Y(y) = 0$ si $y \notin [-1,1]$.

On suppose à présent $y \in [-1,1]$,

$$\begin{aligned} f_Y(y) &= \int_{\mathbb{R}} \frac{1}{\pi}\mathbb{1}_{\mathbb{D}}(x,y)dx \\ &= \int_{-\sqrt{1-y^2}}^{+\sqrt{1-y^2}} \frac{1}{\pi}dx \\ &= \frac{2\sqrt{1-y^2}}{\pi} \end{aligned}$$

Proposition 6.29 Si $Z = (X,Y) : \Omega \to \mathbb{R}^2$ est une variable aléatoire à densité, c'est à dire qu'il existe $f : \mathbb{R}^2 \to \mathbb{R}_+$ tel que :

$$\forall [a,b]\times[c,d] \subset \mathbb{R}^2, \qquad \mathbb{P}(Z \in [a,b]\times[c,d]) = \int_c^d \int_a^b f(x,y)dxdy$$

Alors X et Y sont à densité et les densités de X et Y sont données par :

$$f_X(x) = \int_{\mathbb{R}} f(x,y)dy \qquad \text{et} \qquad f_Y(y) = \int_{\mathbb{R}} f(x,y)dx$$

On appelle f_X et f_Y *les densités marginales* de Z.

6 Couple de variables aléatoires discrètes

R Soit X, Y deux variables aléatoires discrètes. Le couple (X,Y) est encore une variable aléatoire discrète.

Exemple 6.30 Soient $X, Y : \Omega \to \mathbb{R}$ deux variables aléatoires telles que $X(\Omega) = \{0,1,2,3\}$ et $Y(\Omega) = \{-1,0,1\}$, on suppose de plus que les probabilités que $(X,Y) = (i,j)$ sont données dans le tableau suivant :

Y \ X	0	1	2	3
-1	$1/36$	$2/36$	$4/36$	$1/36$
0	$4/36$	$1/36$	$6/36$	$1/36$
1	$5/36$	$6/36$	$4/36$	$1/36$

1. Ce tableau défini bien une probabilité, car les probas des événements $\mathbb{P}(X = i, Y = j)$ sont bien positive et de somme totale 1.
2. Les lois de X et Y sont données par :

$X(\Omega) = \{0,1,2,3\}$

x	0	1	2	3
$\mathbb{P}(X = x)$	$10/36$	$9/36$	$14/36$	$3/36$

car :

$$\mathbb{P}(X = 0) = \mathbb{P}(X = 0, Y = -1) + \mathbb{P}(X = 0, Y = 0) + \mathbb{P}(X = 0, Y = 1) = 10/36$$

et

$Y(\Omega) = \{-1,0,1\}$

y	-1	0	1
$\mathbb{P}(Y = y)$	$8/36$	$12/36$	$16/36$

3. L' espérance de Y est donnée par :

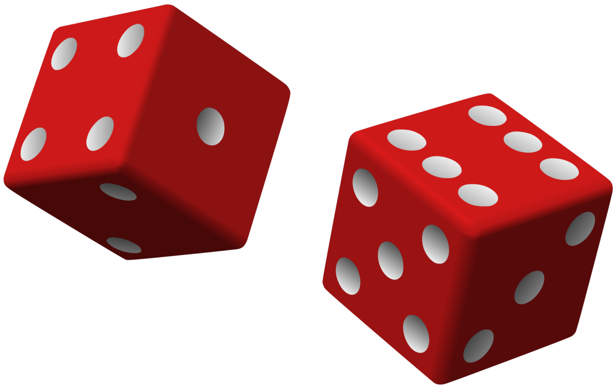
$$\mathbb{E}(Y) = -1 \times 8/36 + 0 \times 12/36 + 1 \times 16/36 = 8/36$$

4. On peut calculer $\mathbb{P}(X = 1|Y = 0)$, autrement dit calculer la probabilité de l'événement $\{X = 1\}$ sachant l'événement $\{Y = 0\}$.

$$\mathbb{P}(X = 1|Y = 0) = \frac{\mathbb{P}(X = 1, Y = 0)}{\mathbb{P}(Y = 0)} = \frac{1/36}{12/36} = 1/12.$$

5. On peut aussi calculer $\mathbb{P}(X+Y=2)$

$$\mathbb{P}(X+Y=2)=\mathbb{P}(X=1,Y=1)+\mathbb{P}(X=2,Y=0)+\mathbb{P}(X=3,Y=-1)={}^{13}\!/_{36}.$$



7. Indépendance de variables aléatoires

1 C'est quoi ?

Idée : Deux v.a. X et $Y : \Omega \to \mathbb{R}$ sont *indépendantes* si la connaissance de la valeur de l'une n'apporte aucune information sur la valeur de l'autre. Exemple : $\Omega = \{\text{hommes adultes en France}\}$, X = revenu, Y = taille de chaussures.

Définition 7.1 Soient $X, Y : \Omega \to \mathbb{R}$ deux v.a. On dit X et Y sont *indépendantes* si pour tous les intervalles $I_1 \subset \mathbb{R}$ et $I_2 \subset \mathbb{R}$, les événements

$$\{\omega \in \Omega \mid X(\omega) \in I_1\} \text{ et } \{\omega \in \Omega \mid Y(\omega) \in I_2\}$$

sont indépendants ; autrement dit, si

$$\mathbb{P}(X \in I_1 \text{ et } Y \in I_2) = \mathbb{P}(X \in I_1) \cdot \mathbb{P}(Y \in I_2)$$

1.A. Cas des variables aléatoires discrètes

R Si X et Y sont discrètes alors X et Y sont indépendantes si et seulement si :

$$\text{pour tout } x \in X(\Omega) \text{ et pour tout } y \in Y(\Omega), \qquad \mathbb{P}(X = x \text{ et } Y = y) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y)$$

1.B. Cas des variables aléatoires à densité

Proposition 7.2 Soit X et Y des variables aléatoires conjointement continues. On note f la densité de la variable aléatoire (X, Y).

S'ils existent $f_X : \mathbb{R} \to \mathbb{R}_+$ et $f_Y : \mathbb{R} \to \mathbb{R}_+$ telles que $f(x, y) = f_X(x) f_Y(y)$ alors :

- X et Y sont indépendantes et à densité.

- Quitte à changer f_X en λf_X et f_Y en $\lambda^{-1} f_Y$, pour un certain $\lambda > 0$, on peut supposer que $\int_{\mathbb{R}} f_X(x)dx = \int_{\mathbb{R}} f_Y(y)dy = 1$, et alors X et Y sont de densité respectives f_X et f_Y.

Démonstration. On commence par vérifier le second point. On sait que :

$$\begin{aligned}1 &= \int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)dxdy \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} f_X(x)f_Y(y)dxdy \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} f_X(x)f_Y(y)dxdy \\ &= \Big(\int_{\mathbb{R}} f_X(x)dx\Big)\Big(\int_{\mathbb{R}} f_Y(y)dy\Big)\end{aligned}$$

On pose $\lambda := \int_{\mathbb{R}} f_Y(y)dy$, $g_X := \lambda f_X$ et $g_Y := \lambda^{-1} f_Y$. On obtient alors :

$$\lambda^{-1} = \int_{\mathbb{R}} f_X(x)dx$$

et

$$\int_{\mathbb{R}} g_X(x)dx = \lambda \int_{\mathbb{R}} f_X(x)dx = \lambda \cdot \lambda^{-1} = 1.$$

De la même façon, on a $\int_{\mathbb{R}} g_Y(y)dy = 1$. On suppose donc à partir de maintenant que :

$$\int_{\mathbb{R}} f_X(x)dx = \int_{\mathbb{R}} f_Y(y)dy = 1$$

On calcule la première marginale. Soit A un intervalle :

$$\begin{aligned}\mathbb{P}(X \in A) &= \mathbb{P}(X \in A, Y \in \mathbb{R}) \\ &= \int_A \int_{\mathbb{R}} f(x,y)dydx \\ &= \int_A \int_{\mathbb{R}} f_X(x)f_Y(y)dydx \\ &= \int_A f_X(x)\Big(\int_{\mathbb{R}} f_Y(y)dy\Big)dx \\ &= \int_A f_X(x)dx \int_{\mathbb{R}} f_Y(y)dy \\ &= \int_A f_X(x)dx \times 1 \\ &= \int_A f_X(x)dx\end{aligned}$$

Donc X est à densité, de densité f_X. De la même façon, on montre que Y est à densité, de densité f_Y.

On montre à présent, l'indépendance de X et Y. Soit A et B deux intervalles :

$$\begin{aligned}
\mathbb{P}(X \in A, Y \in B) &= \int_A \int_B f(x,y)dydx \\
&= \int_A \int_B f_X(x) f_Y(y) dydx \\
&= \int_A f_X(x) \Big(\int_B f_Y(y)dy \Big) dx \\
&= \Big(\int_A f_X(x)dx \Big) \Big(\int_B f_Y(y)dy \Big) \\
&= \mathbb{P}(X \in A)\mathbb{P}(Y \in B)
\end{aligned}$$

■

Proposition 7.3 Si X et Y sont des variables aléatoires indépendantes et continues de densité f_X et f_Y alors elles sont conjointement continues et le couple (X,Y) est de densité $f(x,y) = f_X(x) f_Y(y)$.

Démonstration. Pour tout $A, B \subset \mathbb{R}$ intervalles :

$$\begin{aligned}
\mathbb{P}((X,Y) \in A \times B) &= \mathbb{P}(X \in A \text{ et } Y \in B) \\
&= \mathbb{P}(X \in A)\mathbb{P}(Y \in B) \\
&= \int_A f_X(x)dx \int_B f_Y(y)dy \\
&= \int_{A \times B} f_X(x) f_Y(y) dxdy
\end{aligned}$$

■

Exemple 7.4 Deux personnes se donnent rendez-vous entre 12h et 13h. On note X et Y les variables aléatoires qui modélisent leur temps d'arrivée. On suppose que X et Y sont indépendantes et suivent une loi uniforme continue sur $[0,60]$. Quelle est la probabilité que celui qui arrive en premier attente plus de 10 min ?

On remarque que :

$$\mathbb{P}(X \text{ a plus de 10 min de retard sur } Y) = \mathbb{P}(Y + 10 < X)$$

On cherche donc à calculer $\mathbb{P}(X + 10 < Y) + \mathbb{P}(Y + 10 < X)$.

Comme le problème est symétrique on a $\mathbb{P}(X + 10 < Y) = \mathbb{P}(Y + 10 < X)$.

On note $D = \{(x,y) \in [0,60]^2 \,|\, x+10<y\}$, on a :

$$\begin{aligned}
\mathbb{P}(X+10<Y) &= \int_D \frac{1}{60^2} dxdy \\
&= \int_{y=10}^{y=60} \int_{x=0}^{x=y-10} \frac{1}{60^2} dxdy \\
&= \frac{1}{60^2} \int_{y=10}^{y=60} (y-10)dy \\
&= \frac{1}{60^2} \left[\frac{y^2}{2} - 10y \right]_{y=10}^{y=60} dy \\
&= \frac{1250}{60^2} = \frac{50 \cdot 50}{2 \cdot 60^2}
\end{aligned}$$

D'où : $\mathbb{P}(X+10<Y) + \mathbb{P}(Y+10<X) = 2\frac{1250}{60^2} = \frac{25}{36}$

2 Espérance et indépendance

Définition 7.5 Si X, Y sont des variables aléatoires discrètes alors pour toute fonction $h : \mathbb{R}^2 \to \mathbb{R}$, on a :

$$\mathbb{E}(h(X,Y)) = \sum_{x \in X(\Omega); y \in Y(\Omega)} h(x,y)\mathbb{P}(X=x, Y=y)$$

Définition 7.6 Si X, Y sont des variables aléatoires conjointement continues et le couple (X,Y) est de densité $f : \mathbb{R}^2 \to \mathbb{R}_+$ alors pour toute fonction $h : \mathbb{R}^2 \to \mathbb{R}$, on a :

$$\mathbb{E}(h(X,Y)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x,y) f(x,y) dxdy$$

Proposition 7.7 Soient X et Y deux variables indépendantes. Alors

$$\mathbb{E}(XY) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$$

Démonstration. Dans le cas où X, Y sont discrètes :

$$\begin{aligned}
\mathbb{E}(X) \cdot \mathbb{E}(Y) &= \left(\sum_{x_i \in X(\Omega)} x_i \cdot \mathbb{P}(X=x_i) \right) \cdot \left(\sum_{y_j \in Y(\Omega)} y_j \cdot \mathbb{P}(Y=y_j) \right) \\
&= \sum_{x_i \in X(\Omega)} \sum_{y_j \in Y(\Omega)} x_i y_j \cdot \mathbb{P}(X=x_i) \cdot \mathbb{P}(Y=y_j) \\
&\overset{X,Y \text{ indép}}{=} \sum_{x_i \in X(\Omega)} \sum_{y_j \in Y(\Omega)} x_i y_j \cdot \mathbb{P}(X=x_i, Y=y_j) \\
&= \mathbb{E}(XY)
\end{aligned}$$

Dans le cas où X,Y sont continues :

$$\begin{aligned}\mathbb{E}(XY) &= \int_{\mathbb{R}^2} xyf_X(x)f_Y(y)dxdy \\ &= \int_{\mathbb{R}} xf_X(x)dx \int_{\mathbb{R}} yf_Y(y)dy \\ &= \mathbb{E}(X)\mathbb{E}(Y)\end{aligned}$$

■

3 Indépendance multiple

Définition 7.8 Soient $(X_i)_{i=1,\dots,n} : \Omega \to \mathbb{R}$ n v.a. On dit que les $(X_i)_{i=1,\dots,n}$ sont *indépendantes* si quels que soient les intervalles $(I_i)_{i=1,\dots,n}$, les événements

$$\{\omega \in \Omega \mid X_1(\omega) \in I_1\}, \dots, \{\omega \in \Omega \mid X_n(\omega) \in I_n\}$$

sont indépendants ; autrement dit, si

$$\mathbb{P}(X_1 \in I_1, X_2 \in I_2, \dots, X_n \in I_n) = \mathbb{P}(X_1 \in I_1)\mathbb{P}(X_2 \in I_2)\cdots\mathbb{P}(X_n \in I_n)$$

Exemple 7.9 Soit X_1, ..., X_n n v.a. indépendantes qui suivent toute la loi de Bernoulli de paramètre p. On note $S_n = X_1 + \cdots X_n$. La v.a S_n suit une loi binomiale de paramètre (n,p).

R Cet exemple permet de redémontrer (et de simplifier la preuve) du fait que l'espérance d'une binomiale de paramètre (n,p) est np. Puisque,

$$\begin{aligned}\mathbb{E}(S_n) &= \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n) \\ &= \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_1) \\ &= n\mathbb{E}(X_1) \\ &= np\end{aligned}$$

4 Covariance

Définition 7.10 La *covariance* de deux v.a. X et Y est le nombre :

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

R [Interprétation] Si $\mathrm{Cov}(X,Y) > 0$, alors X et Y sont "positivement correlés" [C.à.d. si X est relativement grand, alors Y a tendance à être relativement grand, aussi], si $\mathrm{Cov}(X,Y) < 0$, alors X et Y sont "négativement correlés".

Proposition 7.11 $\mathrm{Cov}(X,Y) = \mathbb{E}(X\cdot Y) - \mathbb{E}(X)\cdot\mathbb{E}(Y)$

Démonstration.

$$\begin{aligned}\mathrm{Cov}(X,Y) &= \mathbb{E}[XY - Y\mathbb{E}(X) - X\mathbb{E}(Y) + \mathbb{E}(X)\mathbb{E}(Y)]\\ &= \mathbb{E}(XY) - \mathbb{E}(Y)\mathbb{E}(X) - \mathbb{E}(X)\mathbb{E}(Y) + \mathbb{E}(X)\mathbb{E}(Y)\\ &= \mathbb{E}(XY) - \mathbb{E}(Y)\mathbb{E}(X)\end{aligned}$$

■

Proposition 7.12 Si X et Y sont deux v.a. *indépendantes*, alors :

$$\mathrm{Cov}(X,Y) = 0$$

Démonstration. D'après Proposition 7.7, $\mathbb{E}(X\cdot Y) = \mathbb{E}(X)\cdot\mathbb{E}(Y)$. ■

Exemple 7.13 ATTENTION, la réciproque de l'implication X,Y indep $\Longrightarrow \mathrm{Cov}(X,Y) = 0$ est fausse. Par exemple, soient X,Y deux v.a. telles que :

$$\mathbb{P}(X=-1) = \mathbb{P}(X=0) = \mathbb{P}(X=1) = \frac{1}{3}$$

$$Y(\omega) = \begin{cases} 1 & \text{si } X(\omega) = 0 \\ 0 & \text{si } X(\omega) \neq 0 \end{cases}$$

Alors $X(\omega)\cdot Y(\omega) = 0$ pour tout $\omega \in \Omega$, et donc $\mathbb{E}(X\cdot Y) = 0$. Aussi, $\mathbb{E}(X) = (-1)\cdot\frac{1}{3} + 0\cdot\frac{1}{3} + 1\cdot\frac{1}{3} = 0$, donc

$$\mathrm{Cov}(X,Y) = 0$$

En revanche, $\mathbb{P}(X=0, Y=0) = 0$, mais $\mathbb{P}(X=0) = \frac{1}{3}$ et $\mathbb{P}(Y=0) = \frac{2}{3}$, donc $\mathbb{P}(X=0)\cdot\mathbb{P}(Y=0) = \frac{2}{9}$. On conclut que $\mathbb{P}(X=0, Y=0) \neq \mathbb{P}(X=0)\cdot\mathbb{P}(Y=0)$ donc

$$X \text{ et } Y \text{ sont dépendantes.}$$

Exemple 7.14 Les deux côtés d'une pièce sont marqués "0" et "1". On jette cette pièce deux fois – on a donc un univers $\Omega = \{(0,0), (0,1), (1,0), (1,1)\}$, dont chaque élément apparaît avec probabilité $\frac{1}{4}$. Considérons les deux variables aléatoires

$$X((i,j)) = \min(i,j) \quad \text{et} \quad Y((i,j)) = \max(i,j)$$

Quelle est la covariance ?
Réponse : $\mathbb{E}(X) = 0\cdot\frac{3}{4} + 1\cdot\frac{1}{4} = \frac{1}{4}$, $\mathbb{E}(Y) = 0\cdot\frac{1}{4} + 1\cdot\frac{3}{4} = \frac{3}{4}$,
$\mathbb{E}(X\cdot Y) = \frac{1}{4}\cdot\min(0,0)\cdot\max(0,0) + \frac{1}{4}\cdot\min(0,1)\cdot\max(0,1) + \frac{1}{4}\cdot\min(1,0)\cdot\max(1,0) + \frac{1}{4}\cdot\min(1,1)\cdot\max(1,1) = \frac{1}{4}$.
Donc $\mathrm{Cov}(X,Y) = \frac{1}{4} - \frac{3}{4}\cdot\frac{1}{4} = \frac{1}{16}$.
En particulier, les deux variables aléatoires sont dépendantes.

Théorème 7.15 (Propriétés de la covariance) Soient X et Y deux variables aléatoires. Alors :

1. $\mathrm{Cov}(X,X) = \mathbb{V}(X)$,
2. Symétrie : $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$,
3. Bilinéarité :
   $\mathrm{Cov}(a \cdot X, Y) = a \cdot \mathrm{Cov}(X,Y) = \mathrm{Cov}(X, a \cdot Y)$
   $\mathrm{Cov}(X_1 + X_2, Y) = \mathrm{Cov}(X_1,Y) + \mathrm{Cov}(X_2,Y)$
   $\mathrm{Cov}(X, Y_1 + Y_2) = \mathrm{Cov}(X,Y_1) + \mathrm{Cov}(X,Y_2)$
4. $\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2 \cdot \mathrm{Cov}(X,Y)$.
   En particulier, si X et Y sont indépendantes, alors $\mathbb{V}(X+Y) = \mathbb{V}(X) + V(Y)$.

Démonstration. (1), (2), (3) très faciles. Démonstation de (4) :

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

■

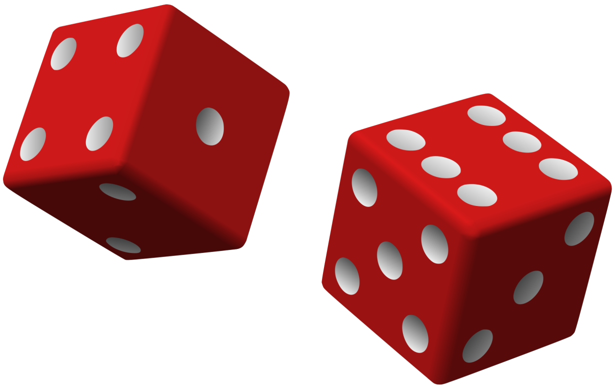
Corollaire 7.16 Si S suit une loi binomiale de paramètre (n,p) alors $\mathbb{V}(S) = np(1-p)$.

Démonstration. On rappelle que si X_1, ..., X_n n sont des v.a. indépendantes qui suivent toute la loi de Bernoulli de paramètre p alors $S = X_1 + \cdots X_n$ suit une loi binomiale de paramètre (n,p).

En particulier,

$$\begin{aligned}
\mathbb{V}(S) &= \mathbb{V}(X_1) + \cdots + \mathbb{V}(X_n) \\
&= \mathbb{V}(X_1) + \cdots + \mathbb{V}(X_1) \\
&= n\mathbb{V}(X_1) \\
&= np(1-p)
\end{aligned}$$

■



8. $\leqslant$ de Bienaymé-Tchebychev, LGN et TCL

Hors-programme pour l'année 2022-23

Motivation Situation typique : il y a une loi de probabilité inconnue dont je veux estimer l'espérance μ. Je ne connais pas la loi, mais je peux tirer des nombres selon elle.

Exemple 8.1

1. J'ai une pièce et je veux savoir quelle est la probabilité p de "Pile" (juste ou truquée ?). Soit $\Omega = \{\text{Pile, Face}\}$, $\mathbb{P}(\text{Pile}) = p$, $\mathbb{P}(\text{Face}) = 1 - p$, $X : \Omega \to \{0,1\}$, Pile $\mapsto 1$, Face $\mapsto 0$, on veut estimer $\mathbb{E}(X) = p$.
2. Soit $\Omega = \{\text{adultes français}\}$, $X(\omega) = 1$ si ω va voter UMP, $X(\omega) = 0$ sinon. Je veux estimer $\mathbb{E}(X) =$ pourcentage du vote pour l'UMP. Je peux téléphoner à des ω au hasard et demander leur intention de vote.

Stratégie évidente pour estimer $\mu = \mathbb{E}(X)$: tirer n nombres $x_1, \ldots, x_n$ pour n assez grand, et calculer leur moyenne $m_n := \frac{x_1 + \ldots + x_n}{n}$. Alors, pour n très grand $m_n \overset{\text{approx}}{=} \mu$.

Questions : **1.** pourquoi ça marche ? **2.** Comment choisir N pour avoir 95% de chance que

$$m_N \in [\mu - 0{,}02,\ \mu + 0{,}02], \quad (\text{autrement dit } |m_N - \mu| \leqslant 0{,}02)\ ?$$

1 Inégalité de Bienaymé–Tchebychev

Théorème 8.2 (Inégalité de Markov) Soit $X : \Omega \to \mathbb{R}_+$ une variable aléatoire à valeurs positives (et supposons que $\mathbb{E}(X)$ existe). Alors pour tout nombre réel $\lambda > 0$

$$\mathbb{P}(X \geqslant \lambda) \leqslant \frac{\mathbb{E}(X)}{\lambda}$$

R Pour $\lambda \leqslant \mathbb{E}(X)$ cet énoncé est vide. [L'inégalité de Markov n'est donc pas très performante, mais a la grande qualité qu'elle marche pour toutes les v.a. positives !]

Démonstration. Pour $\lambda > 0$, soit

$$Y(\omega) = \begin{cases} \lambda & \text{si } X(\omega) \geqslant \lambda \\ 0 & \text{si } X(\omega) \in [0, \lambda[\end{cases}$$

et par hypothèse le cas $X(\omega) < 0$ ne peut pas arriver. On a $X(\omega) \geqslant Y(\omega)$ pour tout $\omega \in \Omega$, et donc

$$\mathbb{E}(X) \geqslant \mathbb{E}(Y).$$

On calcule ensuite

$$\begin{aligned} \mathbb{E}(X) &\geqslant \mathbb{E}(Y) \\ &= 0 \cdot \mathbb{P}(Y = 0) + \lambda \cdot \mathbb{P}(Y = \lambda) \\ &= 0 \cdot \mathbb{P}(X < \lambda) + \lambda \cdot \mathbb{P}(X \geqslant \lambda) \\ &= \lambda \cdot \mathbb{P}(X \geqslant \lambda) \end{aligned}$$

et on déduit $\frac{\mathbb{E}(X)}{\lambda} \geqslant \mathbb{P}(X \geqslant \lambda)$. ■

Théorème 8.3 (Inégalité de Tchebychev, ou de Bienaymé–Tchebychev)
Soit $X : \Omega \to \mathbb{R}$ une v.a. sur un espace probabilisé $(\Omega, \mathbb{P})$. Supposons que $\mu = \mathbb{E}(X)$ et $\sigma^2 = \mathbb{V}(X)$ existent. Pour tout nombre réel $k > 0$ on a

$$\mathbb{P}\left(|X - \mu| \geqslant k\right) \leqslant \frac{\sigma^2}{k^2}$$

Interprétation Tchebychev donne une borne supérieure sur la probabilité des événements loins de la moyenne – il dit que les événements extrèmes sont rares.

R Si $k \leqslant \sigma$, l'écart-type de X, alors l'énoncé est vide. [L'inégalité de Tchebychev n'est donc pas très performante, mais elle marche pour toutes les lois.]

Démonstration.

$$\begin{aligned} \mathbb{P}(|X - \mu| \geqslant k) &= \mathbb{P}\left((X - \mu)^2 \geqslant k^2\right) \\ &\overset{\text{Markov}}{\leqslant} \frac{\mathbb{E}\left((X - \mu)^2\right)}{k^2} \\ &\overset{\text{Def de } \sigma}{=} \frac{\sigma^2}{k^2} \end{aligned}$$

■

Exemple 8.4 Soit X une v.a., supposons que la loi de X n'est pas connue, mais $\mathbb{E}(X) = \mu$ et $\mathbb{V}(X) = \sigma^2$ le sont. Quelle est la probabilité d'obtenir une valeur en dehors de l'intervalle $]\mu - 10\cdot\sigma, \mu + 10\cdot\sigma[$? Autrement dit, que vaut $\mathbb{P}\left(|X-\mu| \geqslant 10\sigma\right)$?

Réponse :

$$\mathbb{P}\left(|X-\mu| \geqslant 10\sigma\right) \overset{\text{Tchebychev}}{\leqslant} \frac{\sigma^2}{10^2\sigma^2} = \frac{1}{100}$$

Donc la probabilité est inférieure à 1%

Exemple 8.5 Le nombre hebdomaire de ventes de voitures pour un certain concessionnaire est une variable aléatoire X d'espérance $\mathbb{E}(X) = 16$ et variance $\mathbb{V}(X) = 9$. Donner une borne inférieure à la probabilité que les ventes de la semaine prochaine se situent entre 10 et 22, bornes incluses.

Réponse : On a $\mathbb{E}(X) = \mu = 16$ et $\mathbb{V}(X) = \sigma^2 = 9$. On cherche une borne inférieure pour $\mathbb{P}(10 \leqslant X \leqslant 22)$.

$$\begin{aligned}
\mathbb{P}(10 \leqslant X \leqslant 22) &= \mathbb{P}(-6 \leqslant X - 16 \leqslant 6) \\
&= \mathbb{P}(|X-16| \leqslant 6) \\
&= 1 - \mathbb{P}(|X-16| > 6) \\
&= 1 - \mathbb{P}(|X-16| \geqslant 7)
\end{aligned}$$

et d'après Tchebychev

$$\mathbb{P}(|X-16| \geqslant 7) \leqslant \frac{9}{7^2} = \frac{9}{49}$$

Donc

$$\mathbb{P}(10 \leqslant X \leqslant 22) \geqslant 1 - \frac{9}{49} = \frac{40}{49} \simeq 0.8$$

2 La loi faible des grands nombres

Définition 8.6 Une famille de v.a. $(X_i)_i : \Omega \to \mathbb{R}$ sont *identiquement distribuées*, si elles suivent la même loi de probabilité, c'est à dire : $\forall a \in \mathbb{R},\ \forall i,j,\ \mathbb{P}(X_i \leqslant a) = \mathbb{P}(X_j \leqslant a)$.

R Très souvent, on considère des familles $(X_i)_i : \Omega \to \mathbb{R}$ v.a. indépendantes et identiquement distribuées (v.a.i.i.d.).

[La *loi* des grands nombres - ici, le mot "loi" veut dire "règle" ou "théorème". Ce n'est pas une loi de probabilité.]

Théorème 8.7 (La loi faible des grands nombres) Soit $X_1, X_2, \dots, X_n, \dots$ une suite de variables indépendantes et identiquement distribuées. On suppose que toutes admettent une espérance finie $\mathbb{E}(X_i) = \mu$. Alors pour tout $\epsilon > 0$,

$$\mathbb{P}\left(\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \mu\right| > \epsilon\right) \longrightarrow 0 \quad \text{lorsque } n \to \infty$$

R (Interprétation) La probabilité que l'estimation $\mathbb{E}(X) \simeq \frac{x_1+\ldots+x_n}{n}$ soit fausse, par une erreur supérieure à ϵ, tend vers 0 quand $n \to \infty$.

Notation. *Si $X_1, X_2, \ldots$ sont des v.a. i.i.d., nous noterons M_n la variable aléatoire "moyenne des n premiers X_i" :*

$$M_n = \frac{X_1+\ldots+X_n}{n}$$

Démonstration. Pour le cas particulier quand les variables considérées ont variance $\mathbb{V}(X_i) = \sigma^2$ *finie*. Dans ce cas, l'indépendance implique, d'après le Théorème 7.15(4),

$$\mathbb{V}(M_n) = \mathbb{V}\left(\frac{X_1+\ldots+X_n}{n}\right) \overset{\text{indep}}{=} \frac{\sum_{i=1}^n \mathbb{V}(X_i)}{n^2} = \frac{n\cdot\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

D'après la linéarité de l'espérance,

$$\mathbb{E}(M_n) = \mathbb{E}\left(\frac{X_1+\ldots+X_n}{n}\right) = \frac{1}{n}\cdot\sum_{i=1}^n \mathbb{E}(X_i) = \frac{1}{n}\cdot n\cdot\mu = \mu$$

et d'après Tchebychev

$$\mathbb{P}\left(\left|M_n-\mu\right| > \epsilon\right) \leqslant \frac{\sigma^2/n}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \overset{n\to\infty}{\longrightarrow} 0.$$

■

R Vous vous imaginez qu'il y a aussi une loi *forte* des grands nombres. Elle dit que pour *presque toute* suite $x_1, x_2, \ldots$ de nombres réels tirés indépendamment selon la loi de X,

$$m_n = \frac{x_1+\ldots+x_n}{n} \underset{n\to\infty}{\longrightarrow} \mathbb{E}(X)$$

3 Théorème central limite

Théorème 8.8 (Théorème central limite "TCL") Soient $X_1, X_2, \ldots : \Omega \to \mathbb{R}$ une suite de variables aléatoires i.i.d d'espérance μ et de variance σ^2. Soit

$$M_n = \frac{X_1+\ldots+X_n}{n}$$

Alors la loi de la v.a. $M_n^* := \dfrac{M_n-\mu}{\frac{\sigma}{\sqrt{n}}}$ tend vers la loi normale $\mathcal{N}(0,1)$ lorsque $n \to \infty$.

Ceci veut dire que la fonction de répartition de $\frac{M_n-\mu}{\frac{\sigma}{\sqrt{n}}}$ tend, en tout point, vers la fonction de répartition Φ de la loi normale :

$$\mathbb{P}\left(\frac{M_n-\mu}{\frac{\sigma}{\sqrt{n}}} \leqslant a\right) \underset{n\to\infty}{\longrightarrow} \Phi(a) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} e^{\frac{-x^2}{2}}\, dx$$

R Il est important de remarquer que M_n^* est centrée réduite, c'est à dire d'espérance 0 et de variance 1.

Démonstration. Admis. ∎

R C'est un théorème concernant une limite, et le théorème est central à la théorie de la probabilité - d'où le nom bizarre.

Exemple 8.9 Soient $X_1,\ldots,X_{100}$ des variables aléatoires indépendantes uniformes sur l'intervalle $[0,1]$. On cherche à évaluer $\mathbb{P}(\frac{X_1+\ldots+X_{100}}{100} > 0{,}53)$.

Réponse : On a $\mathbb{E}(X_i) = \mu = \frac{1}{2}$ et $\mathbb{V}(X_i) = \frac{1}{12}$, donc $\sigma = \sqrt{\frac{1}{12}}$. Alors pour $n = 100$, on a $\frac{\sigma}{\sqrt{n}} = \frac{1}{\sqrt{1200}}$, et

$$\begin{aligned}
\mathbb{P}\left(\frac{X_1+\ldots+X_{100}}{100} > 0{,}53\right) &= 1-\mathbb{P}(M_{100} \leqslant 0{,}53) \\
&= 1-\mathbb{P}\left(\frac{M_{100}-\mu}{\frac{\sigma}{\sqrt{n}}} \leqslant \frac{0{,}53-\mu}{\frac{\sigma}{\sqrt{n}}}\right) \\
&= 1-\mathbb{P}\left(\qquad \leqslant \frac{0{,}03}{\frac{1}{\sqrt{1200}}}\right) \\
&= 1-\mathbb{P}\left(\qquad \leqslant 1{,}039\ldots\right) \\
&\overset{TCL}{\approx} 1-\Phi(1{,}039\ldots) \overset{table}{\approx} 1-0{,}85 = 0{,}15 = \underline{15 \text{ pourcent}}.
\end{aligned}$$

Donc, si je tire 100 nombres selon une loi $\mathscr{U}([0,1])$ et je calcule leur moyenne, alors cette moyenne a des bonnes chances d'être proche de 0,5 : il n'y a que 15% de risque qu'elle soit plus grande que 0,53, et symétriquement 15% de risque qu'elle soit plus petite que 0,47. Il y a donc 70% de probabilité qu'elle soit entre 0,47 et 0,53.

R Interprétation du TCL : On a vu en cours que dans la loi $\mathscr{N}(\mu,\sigma^2)$, à peu près 68% des nombres tirés sont entre $\mu-\sigma$ et $\mu+\sigma$ (et 95,4% sont entre $\mu-2\sigma$ et $\mu+2\sigma$).

Si X est une v.a. arbitraire d'espérance μ et d'écart-type σ (variance σ^2), alors le TCL dit que pour n grand, la v.a. M_n ressemble à une loi normale $\mathscr{N}(\mu, \frac{\sigma^2}{n})$, d'espérance μ et d'écart-type $\frac{\sigma}{\sqrt{n}}$.
Si je tire n nombres et je calcule leur moyenne s_n, j'ai $\sim 68\%$ confiance que le résultat soit à distance au plus $\frac{\sigma}{\sqrt{n}}$ de la vraie espérance μ. Si je veux diviser par 3 l'erreur d'estimation attendu, je dois multiplier n, le nombre de tirages, par 9.

Exemple 8.10 On réalise un sondage pour un référendum sur N personnes. Le résultat du sondage donne 54% pour le oui. Fournir un intervalle de confiance à 95 % pour le résultat final du vote.

Autrement dit, donner un (petit) réel η tel que :

$$\mathbb{P}(\text{proportion de oui} \in [0.54-\eta, 0.54+\eta] \gtrsim 0.95$$

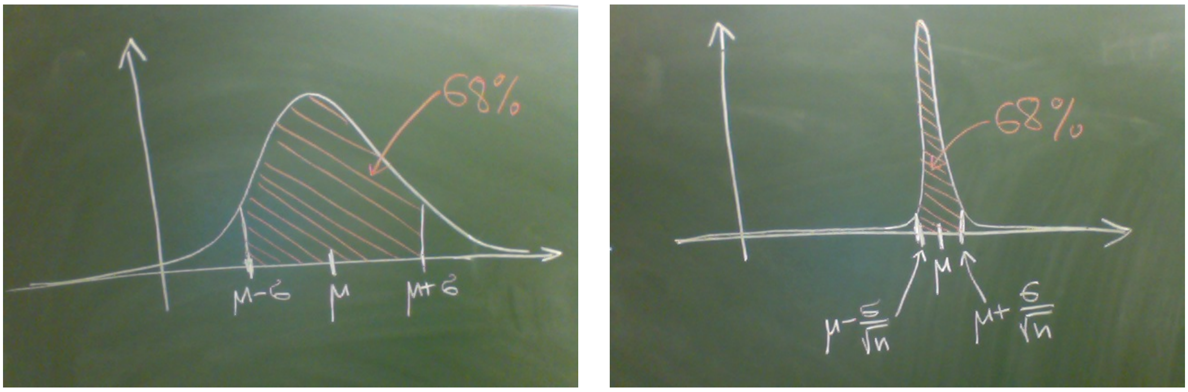


FIGURE 8.1 – La densité de $\mathcal{N}(\mu,\sigma^2)$ et $\mathcal{N}(\mu,\frac{\sigma^2}{n})$

On note X_i la variable aléatoire qui vaut 1 si le sondé i répond "oui", 0 sinon. On note p la proportion de oui, le jour du vote réel. On suppose que les X_i sont indépendantes et identiquement distribuées de loi de Bernoulli de paramètre p. On note

$$M_N = \frac{X_1 + \cdots + X_N}{N}$$

Le nombre M_N est le résultat du sondage. On rappelle que $\mathbb{E}(X_i) = p$ et $\mathbb{V}(X_i) = p(1-p)$. On note G une v.a. qui suit la loi $\mathcal{N}(0,1)$. Le TCL nous dit que :

$$\mathbb{P}\left(\sqrt{N}\left|\frac{M_N - p}{\sqrt{p(1-p)}}\right| \leqslant 2\right) \simeq \mathbb{P}(|G| \leqslant 2) \simeq 0.95$$

Or

$$\mathbb{P}\left(\sqrt{N}\left|\frac{M_N - p}{\sqrt{p(1-p)}}\right| \leqslant 2\right) = \mathbb{P}\left(|M_N - p| \leqslant \frac{2\sqrt{p(1-p)}}{\sqrt{N}}\right)$$

Le problème c'est que l'on ne connait pas p mais on peut utiliser M_N ou remarquer que :

$$p(1-p) \leqslant 1/4$$

Ainsi, on obtient que :

$$\mathbb{P}\left(|M_N - p| \leqslant \frac{1}{\sqrt{N}}\right) \gtrsim 0.95$$

Or,

$$\frac{1}{\sqrt{1000}} \simeq 0.031$$

L'institut de sondage peut donc garantir que le résultat du vote sera dans l'intervalle :

$$[0.509, 0.561]$$

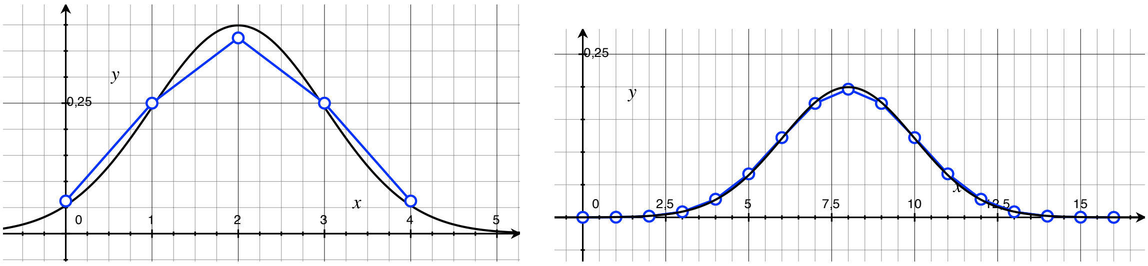


FIGURE 8.2 – Comparaison de la loi Binomiale $\mathscr{B}(n, \frac{1}{2})$ avec la loi normale de la même espérance et la même variance, pour $n = 4$ et $n = 16$. Le diagramme en baton de $\mathscr{B}(n, p)$ converge vers la densité de $\mathcal{N}\Big(np, np(1-p)\Big)$.

Exemple 8.11 Soient $(X_i)_i$ une suite de v.a.i.i.d tel que $X_1 \sim \mathscr{B}(1, p)$, on pose :

$$S_n = X_1 + \cdots + X_n \quad \text{et} \quad M_n = \frac{S_n}{n}$$

Le TCL nous dit que la loi de M_n suit à peu près une loi normale $\mathcal{N}\Big(p, \frac{p(1-p)}{n}\Big)$, donc S_n suit une loi normale $\mathcal{N}\Big(np, np(1-p)\Big)$. Mais S_n suit la loi binomiale $\mathscr{B}(n, p)$. On vient donc de montrer que :

$$\mathscr{B}(n, p) \approx \mathcal{N}\Big(np, np(1-p)\Big)$$