Bert Wiest
Université de Rennes 1
UFR Mathématiques et IRMAR

Cours PS1

Organisation Il y a 15h de cours, dont 2h pour les contrôles continus. Détails sur ma page `http://perso.univ-rennes1.fr/bertold.wiest`

Prérequis Je vais utiliser sans explication

- théorie des ensembles très élémentaire : réunion $A \cup B$, intersection, A et B sont "disjoints" si $A \cap B = \emptyset$, $A \subset B$ sous-ensemble
- un peu de différentiation et intégration de fonctions $f : \mathbb{R} \to \mathbb{R}$.
  - dérivée de polynômes, fonction exponentielle, dérivée de composée de fns.
  - primitive de polynômes, fonction exp, intégration par parties
- suites
  - définition de "convergence"
  - $1 + x + x^2 + x^3 + \ldots = \frac{1}{(1-x)}$ pour $x \in]-1, 1[$
- fonction exponentielle $e^x = \exp(x)$:
  - $\exp(0) = 1$, $\exp'(x) = \exp(x)$, et caractérisé par ça.
  - $\lim_{n\to\infty}(1 + \frac{x}{n})^n = e^x$
  - $\exp(x) = 1 + x + x^2/2! + x^3/3! + x^4/4! + \ldots$ (fait : ça converge)

Littérature

- Alain Combrouze, Probabilités 1, Presses universitaires de France
- Sylvie Méléard, Aléatoire, Les éditions de l'école polytechnique

Ce qu'il faut faire pour ce cours :

- Après chaque cours magistral, relire le cours pour le comprendre, et apprendre les définitions et résultats principaux.
- Participer activement aux TDs. Quand c'est demandé, aussi essayer de faire des exercices à la maison.

[Exemples d'utilisation des probas :

- Statistiques par exemple pour math-éco (intervalle de confiance etc). Exemple typique d'une "variable aléatoire": durée de vie d'une personne. C'est une quantité inconnue pour chaque individu, mais ça suit une distribution de probabilités bien comprise et étudié par les statisticiens. (Ou : prix d'une action en bourse etc.)
- Physique erreurs de mesure (qui ont tendance à être distribués selon une Gaussienne = distribution "normale" en forme de cloche)
- info et télécom : Sur internet ou transmission sur tél portable, probabilité de perte de bits ou de paquets, ou la distribution de la durée de voyage sur internet d'un paquet. Comment adapter ses algorithmes de transmissions à l'évolution de minute en minute de cette probabilité ?
- info : files d'attente (par ex pour accès aux serveurs)]

1. Ensembles et dénombrements

Notation 1.1. (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 "événément" est par définition un sous-ensemble de Ω.

Exemple 1.2. Je tire une carte (choisie au hasard) d'un jeu avec 32 cartes. Alors $\Omega = \{$7 de pique, 8 de pique, ... , as trèfle$\}$ (32 éléments). L'événement "10 d'une quelconque couleur" a quatre éléments : $A = \{$10 pique, 10 coeur, 10 carreau, 10 trèfle$\}$. Remarquez : $A \subset \Omega$.

Exemple 1.3. Je jette deux fois de suite un dé. Alors

$$\begin{aligned}\Omega &= \{(1,1),(1,2),...,(1,6),(2,1),(2,2),......,(6,5),(6,6)\} \\ &= \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}\end{aligned}$$

a $6 \cdot 6 = 36$ éléments. L'événement "deux fois le même nombre"$= \{(1,1),...,(6,6)\}$ a 6 éléments.

Remarque 1.4. C'étaient des exemples particulièrement agréables car tous les éléments de Ω arrivent avec la même probabilité. [Donc, pour calculer la probabilité de "deux fois le même nombre", il suffit de calculer $\frac{6}{36} = \frac{1}{6}$. On va y revenir.]

Notation 1.5. $\forall$ = pour tout (quelque soit), $\exists$ = il existe

Notation 1.6. Pour un ensemble Ω, on note

- $\mathcal{P}(\Omega)$ l'ensemble de parties de Omega . C.à.d., chaque élément de $\mathbb{P}(\Omega)$ est un sous ensemble de Ω. (Par exemple, si $\Omega = \{A, B, C\}$, alors
  $$\mathcal{P}(\Omega) = \{\emptyset, \{A\}, \{B\}, \{C\}, \{A,B\}, \{A,C\}, \{B,C\}, \{A,B,C\}\})$$
  Remarque : si Ω a n éléments, alors $\mathcal{P}(\Omega)$ a 2^n éléments.

- Complémentaire : si $A \subset \Omega$, on note $A^c = \{x \in \Omega \mid x \notin A\}$ le complémentaire. [Dessin]
- Différence : si $A \subset \Omega$ et $B \subset \Omega$, alors $A \setminus B = \{x \in A \mid x \notin B\}$. [Dessin]
- 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. [Dessin]

Proposition 1.7. *(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) $(A^c)^c = A$ *[Dessin]*
(6) $(A \cup B)^c = A^c \cap B^c$ *[Dessin]*
(7) $(A \cap B)^c = A^c \cup B^c$

Démonstration admis (mais plus ou moins évident...)

Exercice 1.8. Donner un exemple qui montre qu'en général, $(A \setminus B) \setminus C \neq A \setminus (B \setminus C)$

Définition 1.9. (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 1.10. si je jette un dé *et* joue au pile ou face, alors j'aurai l'univers
$$\Omega = \{1, ..., 6\} \times \{P, F\}$$
avec $6 \cdot 2 = 12$ éléments : $(1, P)$, $(2, P)$,..., $(6, P)$, $(1, F)$,...,$(6, F)$.

Définition 1.11. 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 $card(\Omega)$.

Proposition 1.12. (1) $card(\emptyset) = 0$
(2) *Si* $A \subset B$ *alors* $card(A) \leqslant card(B)$
(3) *si* $A \subset \Omega$, *alors* $card(A^c) = card(\Omega) - card(A)$
(4) *si* $A, B \subset \Omega$ *sont disjoints* ($A \cap B = \emptyset$), *alors* $card(A \cup B) = card(A) + card(B)$
(5) *en général,* $card(A \cup B) = card(A) + card(B) - card(A \cap B)$ *- DESSIN*

(6) $card(A \setminus B) = card(A) - card(A \cap B)$
(7) *Si A et B sont deux ensembles finis, alors* $card(A \times B) = card(A) \cdot card(B)$

Démonstration admis [juste dire quelque mots sur (5)]

Exemple 1.13. parmi 30 étudiants (E), 20 parlent Anglais (A), 15 parlent Chinois (C), et 10 les Deux (D). Combien parlent ni anglais ni chinois ?

Rép : DESSIN. On cherche

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

Quatre exemples fondamentaux de cardinaux

Rappel 1.14. Soit Ω un ensemble avec n éléments. Combien de permutations des éléments de Ω y a-t-il, c.à.d. combien de listes (ordonnées) qui contiennent tous les éléments de Ω (sans répétition) ? Réponse : $n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 2 \cdot 1 = n!$ ("n factoriel").

Théorème 1.15. *Étant donnée une urne avec n boules, étiquetés $1, \ldots, n$. On fait p tirages.*
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 répétition	Avec répétition
attention à l'ordre (arrangements)	$A_n^p := \frac{n!}{(n-p)!}$ $= n \cdot (n-1) \cdot \ldots \cdot (n-p+1)$ **(1)**	$n^p = n \cdot n \cdot \ldots \cdot n$ **(2)**
en ignorant l'ordre (combinaisons)	$C_n^p := \frac{n!}{(n-p)!p!}$ **(3)**	$\Gamma_n^p := C_{n+p-1}^{n-1}$ $= \frac{(n+p-1)!}{p!(n-1)!}$ **(4)**

Exemple 1.16. (1) Nombre de manières de choisir ses Top 10 ($p = 10$) parmi n chansons
(2) Nombre de textes possibles de longueur p (alphabet avec n lettres)
(3) Nombre de résultats du loto 6 sur 49 (ou p sur n)
(4) Nombre de distributions possibles des voix quand p électeurs votent pour n candidats

[Question pour moi-même : où est-ce qu'on se servira du Γ_n^p ?]

Démonstration (1) arrangements sans répétition :
pour le premier tirage il y a n possibilités
pour le deuxième tirage il y a $n-1$ possibilités
...
pour le $p$ème tirage il y a $n-p+1$ possibilités
donc en total $n \cdot (n-1) \cdot \ldots \cdot (n-p+1)$ possibilités

(3) combinaisons sans répétition : les p boules tirés ont pu être tiré dans $p!$ ordres différentes. Chaque combinaison correspond donc à $p!$ arrangements différents. Le nombre de combinaisons est $\frac{A_n^p}{p!}$

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

(2) facile

(4) On veut démontrer que le nombre de combinaisons possibles est de C_{n+p-1}^{n-1}. Dessinons $n+p-1$ cercles. Par exemple, si $n=4$ et $p=9$, on dessine 12 cercles

$$\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc\bigcirc$$

Supposons le tirage a donné k_1 fois la boule $1, \ldots, k_n$ fois la boule n (avec $k_1 + \ldots + k_n = p$). Cela me donne une façon de choisir $n-1$ cercles parmi les $n+p-1$ de la façon suivante : j'ignore les k_1 premières boules, je sélectionne la k_1+1ème, ignore encore k_2 boules, sélectionne la suivante, etc. Par exemple, si $k_1 = 2$, $k_2 = 4$, $k_3 = 0$, $k_4 = 3$, 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+p-1}^{n-1} différents choix de p cercles, et donc C_{n+p-1}^{n-1} combinaisons. $\square$

Théorème 1.18. *(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) $\forall (k,n)$, $C_n^k = C_n^{n-k}$
(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

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

NE PAS DIRE en cours mais mettre sur la feuille d'exos :
* $\sum_k C_n^k = 2^n$
* $\sum_k (-1)^k C_n^k = 0$

Remarque 1.19. par (e) on peut calculer les C_n^k par le "triangle de Pascal". [Expliquer, montrer le calcul au moins jusqu'à la ligne 4.]

Démonstration (a), (b) faciles.
(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) 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

Définition fausse 2.1. Soit Ω un ensemble. Une *probabilité* sur Ω est une fonction

$$\mathbb{P} \colon \mathcal{P}(\Omega) \to [0,1]$$

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

$$\mathbb{P}(A_1 \cup A_2 \cup A_3 \cup \ldots) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) + \ldots$$

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

Correction de la définition 2.2. Si Ω est fini ou si $\Omega = \mathbb{N}$ ou si $\Omega = \mathbb{Z}$, alors la définition est déjà correcte ! En revanche, pour $\Omega = \mathbb{R}$, par exemple, aucune telle fonction $\mathbb{P}$ existe. On ne devrait 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 $\mathcal{A} \subset \mathbb{P}(\Omega)$ qui s'appelle une *tribu* ou *σ-algèbre* de Ω.

Dans le cadre de ce cours, penser que tout événement possède une proba est certes faux, mais n'induit pas de graves erreurs dans l'immédiat.

Exemple 2.3. où Ω est fini. Expérience aléatoire : tirer une fois d'une urne avec 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. où Ω est infini. Expérience : 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 $\mathbb{P}(\{c\}) = 0$!

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)

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

(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 $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geqslant \mathbb{P}(A)$ $\square$

Notation 2.6. $\emptyset$ 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.*

Le cas particulier où Ω est fini

Observation 2.7. Supposons Ω est fini : $\Omega = \{\omega_1, ..., \omega_n\}$. Alors pour spécifier une probabilité $\mathbb{P}$ sur Ω, il suffit de donner les nombres $p_1 = \mathbb{P}(\omega_1), \dots, p_n = \mathbb{P}(\omega_n)$.

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

Proposition 2.9. *(cas de l'équiprobabilité) Supposons que Ω est fini et tous les éléments de Ω ont la même probabilité $p_i = \frac{1}{card(\Omega)}$. Alors pour tout événement $A \subset \Omega$,*

$$\mathbb{P}(A) = \frac{card(A)}{card(\Omega)}$$

Dém C'est une conséquence immédiate de l'Observation 2.7.

Exemple 2.10. Dans l'Exemple 1.3, $\mathbb{P}$(“deux fois le même nombre”) $= 6/36 = 1/6$

Exemple 2.11. Un exemple célèbre : Quelle est la proba P_n que parmi n personnes il y a au moins deux avec le même anniversaire ? [Pas forcément avec le même âge, juste le même anniversaire. On oublie complications : 29 février etc.] Réponse :

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

Pour calculer ça, soit

$$\Omega = \text{ arrangements de } n \text{ anniversaires } = \{1 \text{ janv}, \dots, 31 \text{ déc}\}^n$$

– donc $card(\Omega) = 365^n$. Soit

$$A = \{ \text{ arrangements de } n \text{ anniversaires distincts}\}$$

donc $card(A) = A_{365}^n = 365 \cdot 364 \cdot ... \cdot (365 - n + 1)$. 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 $n = 47$, $0,95$

Le cas où Ω est infini mais dénombrable

Définition 2.12. L'ensemble Ω est *dénombrable* s'il existe une liste (éventuellement infinie) mentionnant tous ses éléments, c.à.d.,

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

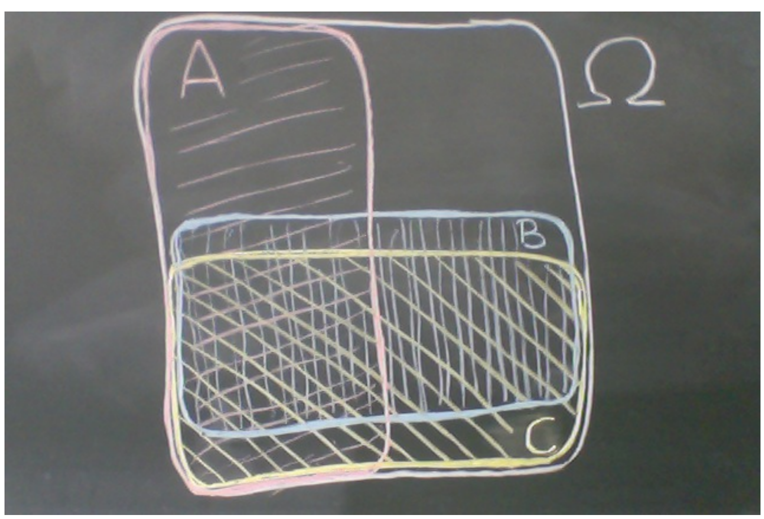
Proposition 2.13. *(analogue de 2.7) Pour spécifier une proba sur Ω dénombrable, il suffit de spécifier $p_i = \mathbb{P}(\omega_i)$ pour $i \in \mathbb{N}$, avec $p_1 + p_2 + p_3 + ... = 1$.*

Dém admis

Exemple 2.14. Soit $\Omega = \mathbb{N}$, $\mathbb{P}(1) = \frac{1}{2}$, $\mathbb{P}(2) = \frac{1}{4}$, $\mathbb{P}(3) = \frac{1}{8}$, $\mathbb{P}(4) = \frac{1}{16}$, etc. (Remarquez que en effet $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots = 1$). Cette information suffit pour calculer la probabilit de tout sous-ensemble de $\mathbb{N}$, par exemple $\mathbb{P}(\{2, 4\}) = \frac{1}{2} + \frac{1}{16}$

3. Probabilité conditionnelle et indépendance

Exemple 3.1. Soit $\Omega = \{$hommes français de 30-60 ans$\}$. Trois sous-ensembles :
$A = \{$ ceux de taille de chaussures plus élevé que la médiane $\}$
$B = \{$ ceux avec nombre d'années d'études plus élevé que la médiane $\}$
$C = \{$ ceux avec revenus plus élevés que la médiane $\}$



A et B sont (à peu près) indépendant : 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 qu'une Monsieur X appartient à C, ses chances chances d'appartenir aussi à B sont plus grands que $\mathbb{P}(C)$. 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. Soit Ω un univers muni d'une probabilité $\mathbb{P}$. Soient $A, B \subset \Omega$ des événements.

(a) Supposons que $\mathbb{P}(B) > 0$. Alors $\mathbb{P}(A|B)$, la *probabilité conditionnelle* de A sachant que B est réalisée (ou proba conditionnelle en B), est définie comme suit :

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

(b) Supposons que $\mathbb{P}(A) > 0$ et $\mathbb{P}(B) > 0$. Alors A et B sont *indépendants* si $\mathbb{P}(A\cap B) = \mathbb{P}(A)\cdot\mathbb{P}(B)$

Exercice 3.3. Soit Ω un univers muni d'une proba $\mathbb{P}$, et soit $B \subset \Omega$ avec $\mathbb{P}(B) > 0$. Alors la probabilité conditionnelle $\mathbb{P}(.|B)$ 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. (a) Deux lancés successives et indépendantes d'une pièce truquée : Pile avec proba p, Face avec proba $1-p$. Soit
$A = \{\text{premier lancer pile}\}$, $B = \{\text{deuxième lancer pile}\}$.
Alors $\mathbb{P}(A) = p$, $\mathbb{P}(B) = p$, $\mathbb{P}(A \cap B) = \mathbb{P}((\text{pile,pile})) = p^2 = \mathbb{P}(A) \cdot \mathbb{P}(B)$

(b) On joue au pile ou face successivement, 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 4.18)]

Exemple 3.6. On jette deux dés.
A = la somme des deux dés est $4 = \{(1,3),(2,2),(3,1)\}$
B = le premier dé donne $1 = \{(1,1),...,(1,6)\}$.
$A \cap B = (1,3)$. 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 Par le Lemme 3.4 il suffit de montrer que $\mathbb{P}(A^c|B) = \mathbb{P}(A^c)$. Or

$$\mathbb{P}(A^c|B) \overset{3.3}{=} 1 - \mathbb{P}(A|B) = 1 - \mathbb{P}(A) = \mathbb{P}(A^c)$$

où la deuxième égalité vient du Lemme 3.4 et l'hypothèse que A, B indépendants. □

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 proba conditionnelle.
$n-1 \Longrightarrow n$:

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

Exemple 3.9. Urne avec 3 boules rouges, 3 blanches. Je tire trois fois 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
- $B_1 \cup B_2 \cup B_3 \cup \ldots = \Omega$

Rappel 3.11. Si $\Omega = B_1 \cup B_2 \cup B_3 \cup \ldots$ est une partition, alors

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

Exemple 3.12. Pour tout $B \subset \Omega$, on a une partition $\Omega = B \cup B^c$

Lemme 3.13. *(Formule des probabilités totales) Soit $B_1, B_2, ... \subset \Omega$ une collection finie ou infinie d'événements qui forme une partition de Ω. Alors*

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

Démonstration exercice $\square$

Proposition 3.14. *(Formule de Bayes) Sous la même hypothèse que 3.13, et si $\mathbb{P}(A) > 0$, alors pour $i = 1, 2, \ldots,$*

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

Démonstration Numérateur $= \mathbb{P}(A \cap B_i)$, dénominateur $\overset{3.13}{=} \mathbb{P}(A)$. $\square$

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

Exemple 3.15. 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 $\frac{1}{3}$. 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. Événements C =compétent, R =bien réparé

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

Donc une seule réparation réussie n'est pas un signal fort qu'il est compétent !

4. Variables aléatoires discrètes

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

Correction de la définition 4.2. Ω est muni d'une *tribu* $\mathcal{A} \subset \mathcal{P}(\Omega)$, 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 $\mathcal{A}$.

Exemple 4.3. (a) 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).

(b) Je joue au 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"
* $Y(\omega) = \begin{cases} 0 & \text{si } \omega \text{ ne contient que des “F”} \\ k & \text{si la première apparition de “P” est lors du } k\text{-ème lancer.} \end{cases}$

(c) $\Omega = \{\text{êtres humains vivants}\}$, pour tout $\omega \in \Omega$, $\mathbb{P}(\omega) = 1/7\text{millard}$, $X(\omega)$ =âge que ω aura au moment de sa mort. [C'est une v.a. très étudiée par les statisticiens, en particuliers ceux qui travaillent pour des assurances.]

(d) un exemple d'une v.a. à valeurs dans $\mathbb{R}$. Épreuve aléatoire : je sélectionne une particule d'une substance radioactive, et j'attends sa décomposition. $\Omega = \{\text{atomes de la substance}\}$, $X\colon \Omega \to \mathbb{R}$, $\omega \mapsto$ temps jusqu'à la décomposition.

Variables aléatoires discrètes

Définition 4.4. Une variable aléatoire X est dite *discrète* si son image $X(\Omega) = \{X(\omega) \mid \omega \in \Omega\}$ est dénombrable.

Remarque 4.5. Dans ce cours, toutes les v.a. discrètes X que nous étudierons en détail seront très sages : elles satisfont
* $X(\Omega) \subset \mathbb{N}$, et
* Ω est fini ou dénombrable.

Notation 4.6. Soit $X\colon\ \Omega \to \mathbb{R}$ une v.a. discrète. Alors pour tout nombre $x_i \in X(\Omega)$, on écrit "$X = x_i$" pour l'événement $\{\omega \in \Omega \mid X(\omega) = x_i\}$. Par exemple,

$$\mathbb{P}(X = x_i) \stackrel{Notation}{=} \mathbb{P}(\{\omega \in \Omega \mid X(\omega) = x_i\})$$

Définition 4.7. La *loi* de probabilité (en anglais: probability distribution) d'une variable aléatoire discrète X est la probabilité $\mathbb{P}_X$ sur l'ensemble $\Omega' = X(\Omega) \subset \mathbb{R}$ définie par

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

[C'est une définition très importante – quand on étudie la loi d'une variable aléatoire, on oublie tout le modèle sous-jacent, toute l'exprience aléatoire etc, on fait juste attention quelle valeur numérique est tirée avec quelle probabilité. On peut penser que la loi d'une variable aléatoire est une boîte noire avec un bouton et un haut-parleur. Chaque fois qu'on appuie sur le bouton, la boite dit un nombre au hasard – mais les fréquences rélatives des nombres sont distribués selon la loi en question.]

Exemple 4.8. (voir 4.3(b)) Je joue au pile et 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	0	1	2
$\mathbb{P}(X = x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

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

Première visualisation 4.9. La loi d'une variable aléatoire discrète peut être représentée visuellement par son *diagramme en bâtons*. Par exemple pour la loi de l'Exemple 4.8: [Dessiner le diagramme en bâtons de 4.8]

Définition 4.10. La *fonction de répartition* d'une loi est la fonction $F : \mathbb{R} \to [0, 1]$, $t \mapsto \mathbb{P}(X \leqslant t)$.

Deuxième visualisation 4.11. Le graphe de la fonction de répartition. (Remarquer que $lim_{t\to-\infty}F=0$, qu F est croissante, et que $lim_{t\to\infty}F=1$. Dans le cas de l'exemple 4.8 :

[Dessiner la fonction de répartition]

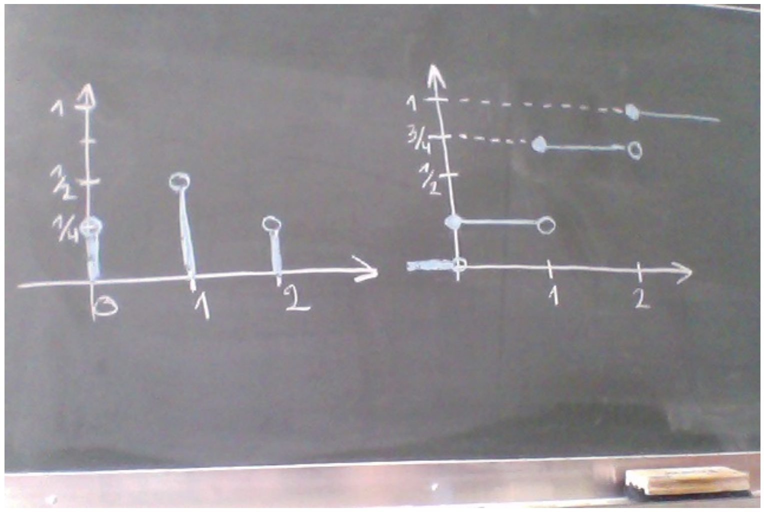


Diagramme en bâtons et fonction de répartition de l'Exemple 4.8

Proposition 4.12. *Si X et Y sont deux v.a. et $\lambda \in \mathbb{R}$, alors la somme $X+Y$, le produit $X \cdot Y$, multiple scalaire $\lambda \cdot X$ est aussi une loi. Si X et Y sont discrets alors $X+Y$, $X \cdot Y$ et λX le sont aussi.*

Démonstration admis

Rappel 4.13. sur les suites et séries. Considérons la série $\sum_{i=1}^{+\infty} a_i$, où $(a_i)_{i\in\mathbb{N}}$ est une suite de nombre réels.

(a) On dit la série est *convergente* si la suite des sommes partielles $S_k = \sum_{i=1}^{k} a_i$ est convergente.

(b) Une condition nécessaire pour que $\sum_{i=1}^{\infty} a_i$ converge est que $(a_i) \to 0$. Cette condition n'est pas suffisante - par exemple $\sum_{i=1}^{\infty} \frac{1}{i} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + ... = \infty$ (série divergente)

(c) Soit $q \in [0, 1[$. Alors $(1-q)\sum_{i=0}^{\infty} q^i = 1$. Par changement de variables $p = 1-q$ on obtient :

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

Lois discrètes classiques

Exemple 4.14. (Loi uniforme - en anglais "uniform distribution") Épreuve aléatoire : une urne contient n boules, numérotés $1, ..., n$. On en prend une au hasard.

$$\Omega = \{\text{les boules}\},\ X\colon\ \Omega \to \{1, ..., n\}, \omega \mapsto \text{le numero de la boule } \omega$$

[Pour $n = 6$ on obtient un modèle du dé non-pipé.]

Notation $[\![1, n]\!] = \{1, ..., n\}$

Définition On dit une v.a. $X\colon\ \Omega \to \mathbb{R}$ suit une loi uniforme sur $[\![1, n]\!]$ si
* $X(\Omega) = [\![1, n]\!]$, et
* pour tout $k \in [\![1, n]\!]$ on a $\mathbb{P}(X = k) = \frac{1}{n}$

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

DESSINER diagramme en batons

Convention 4.15. Dans un jeu de pile ou face, on va interpréter Pile comme un succès et Face comme un échec.

Exemple 4.16. (Loi de Bernoulli) Épreuve aléatoire : on joue au pile ou face avec une pièce qui tombe sur pile avec une probabilité p. Donc

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

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

Définition On dit une v.a. $X\colon\ \Omega \to \mathbb{R}$ suit une loi de Bernoulli de paramètre p (avec $0 < p < 1$) si
* $X(\Omega) = \{0, 1\}$ et
* $\mathbb{P}(X = 0) = 1 - p$, $\mathbb{P}(X = 1) = p$.

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

DESSINER diagramme en batons

Une généralisation :

Exemple 4.17. (Loi binomiale - cette loi compte le nombre de succès parmi n essais indépendants.) Épreuve aléatoire : avec la même pièce, on joue au pile ou face n fois, et on compte le nombre de "pile".

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

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

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

Définition On dit une v.a. $X\colon\ \Omega \to \mathbb{R}$ suit une loi binomiale de paramètres (n,p), si
* $X(\Omega) = [\![0,n]\!]$
* pour $k \in [\![0,n]\!]$ on a $\mathbb{P}(X=k) = C_n^k p^k(1-p)^{n-k}$.

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

[Pour $n=1$ on obtient la loi binomiale.]
DESSINER diagramme en batons.)

Exemple 4.18. (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 au pile ou face jusqu'au premier P. Alors

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

et

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

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

Définition On dit une v.a. $X : \Omega \to \mathbb{R}$ suit une loi géométrique de paramètre p si
* $X(\Omega) = \{1, 2, 3, 4, ...\}$
* $\mathbb{P}(X=k) = p \cdot (1-p)^{k-1}$ [Remarquez que $\sum \mathbb{P}(X=k) = 1$ par 4.13(c)]

Notation $X \sim \mathcal{G}(p)$

Exercice Démontrer que c'est la "loi d'une v.a. sans mémoire" : si $K > 0$ et $k > 0$ alors

$$\mathbb{P}(X > k) = \mathbb{P}(X > K + k \mid X > K)$$

(Indication : montrer d'abord que $\mathbb{P}(X > k) = (1-p)^k$)

[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.)]

Exemple 4.19. (la loi hypergéométrique) Pas le temps.

Exemple 4.20. (Loi de Poisson) [d'après Siméon Denis Poisson, 1781 - 1840]

Définition On dit une v.a. $X\colon \Omega \to \mathbb{R}$ suit une loi de Poisson de paramètre λ (où $\lambda \in \mathbb{R}, \lambda > 0$) si
* $X(\Omega) = \{0, 1, 2, 3, ...\}$
* $\mathbb{P}(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$

Notation $X \sim \mathcal{P}(\lambda)$

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

Remarque 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 $\lambda \in \mathbb{R}$ 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 $X \sim \mathcal{P}(\lambda)$ (quand $n \to \infty$). Plus formellement

Proposition 4.21. *Soit $\lambda \in \mathbb{R}$ fixé. Soit $X_n \sim \mathcal{B}(n, \frac{\lambda}{n})$, une variable binomiale avec paramètres n et $p = \frac{\lambda}{n}$. Soit $Y \sim \mathcal{P}(\lambda)$ (Poisson). Alors pour tout $k \in \mathbb{N}$,*

$$\mathbb{P}(X_n = k) \overset{n\to\infty}{\longrightarrow} \mathbb{P}(Y = k)$$

Démonstration

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

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

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

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

[Idée : c'est la moyenne !! Quand vous entendez "espérance", pensez "moyenne" !]

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

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.

[Interprétation en termes d'argent : on me propose le jeu suivant. J'achte un billet pour 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 du billet est-ce que je devrais accepter ? Réponse : à long terme, si le billet coûte moins de (l'espérance de la loi) Euros, alors je fais un profit.]

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

$$\begin{aligned}\text{moyenne d'age} &= \frac{\sum_{\omega\in\Omega} X(\omega)}{n} \\ &= \sum_{\omega\in\Omega} \frac{1}{n} \cdot X(\omega) \text{ (ecriture Prop 5.2)} \\ &= \sum_{\alpha=18,19,\dots,150} \alpha \cdot \text{(Proportion de personnes ayant exactement } \alpha \text{ annees).(ecriture Def 5.1)}\end{aligned}$$

Exemple 5.4. Sur $\Omega = \{1, 2, 3, 4, ...\}$ 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.

Proposition 5.5. *(Propriétés de l'espérance) Soit Ω muni d'une probabilité $\mathbb{P}$. Soit X une variable aléatoire discrète définie sur Ω. Soit $a, b \in \mathbb{R}$, et $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction réelle. Alors:*

(1) *Addition:* $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$
(2) *Addition d'un réel:* $\mathbb{E}(X+a) = \mathbb{E}(X) + a$
(3) *Multiplication par un réel:* $\mathbb{E}(b \cdot X) = b \cdot \mathbb{E}(X)$
(4) *Composition: Si X est une V.A. discrète pouvant prendre ses valeurs parmi les valeurs x_i, alors*
$$\mathbb{E}(\varphi(X)) = \textstyle\sum_{x_i \in X(\Omega)} \varphi(x_i) \cdot \mathbb{P}(X = x_i)$$

Dém (1) *Addition.* A partir de la definition
$$\begin{aligned}\mathbb{E}(X+Y) &= \sum_{\omega \in \Omega} (X(\omega) + Y(\omega))\, \mathbb{P}(\{\omega\}) \\ &= \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\}) + \sum_{\omega \in \Omega} Y(\omega)\mathbb{P}(\{\omega\}) \\ &= \mathbb{E}(X) + \mathbb{E}(Y)\end{aligned}$$
(3) *Multiplication par un réel.*
$$\mathbb{E}(bX) = \sum_{x_i \in X(\Omega)} b x_i \mathbb{P}(X = x_i) = b \sum_{x_i \in X(\Omega)} x_i \mathbb{P}(X = x_i) = b \cdot \mathbb{E}(X)$$

(2) & (4) : exercice. □

Définition 5.6. Soit X une variable aléatoire sur Ω. 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
$$\sigma(X) = \sqrt{\mathbb{V}(X)}$$

Remarque 5.7. Interprétation : $\mathbb{V}(X)$ petit $\Longrightarrow$ valeurs de X concentrées autour de $\mathbb{E}(X)$. $\mathbb{V}(X)$ grand $\Longrightarrow$ valeurs de X éparpillées.

Remarque 5.8. Si X est discrète, alors
$$\mathbb{V}(X) = \sum_{x_k \in X(\Omega)} (x_k - \mathbb{E}(X))^2\, \mathbb{P}(X = x_k)$$

Proposition 5.9. *Soit Ω l'univers muni d'une probabilité P, soit $X\colon \Omega \to \mathbb{R}$ 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) $\mathbb{V}(\lambda X) = \lambda^2 X$ *pour* $\lambda \in \mathbb{R}$
(5) $\sigma(\lambda X) = |\lambda| \cdot \sigma(X)$

Dém On va démontrer (1), en utilisant les propriétés de Proposition 5.5

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

Exercice : Montrer les propriétés (2),(3) et (4). □

Théorème 5.10. *L'espérance et la variance des v.a. discrètes précédemment décrites sont dans le Tableau 1.*

Démonstration partielle • Soit X une v.a. de loi uniforme $X \sim \mathcal{U}([1,n])$. Alors

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

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

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

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

Exercice : calculer l'espérance de la loi binomiale $\mathcal{B}(n,p)$.

• Soit X une v.a. de loi géométrique $X \sim \mathcal{G}(p)$. Notons $q = 1-p$. Alors

$$\begin{aligned}\mathbb{E}(X) = &\sum_{k=1}^{\infty} kq^{k-1}p = p\cdot\sum_{k=1}^{\infty} kq^{k-1} \overset{(*)}{=} p\cdot\frac{d}{dq}\left(\sum_{k=0}^{\infty} q^k\right) =\\ = &\ p\cdot\frac{d}{dq}\left(\frac{1}{1-q}\right) = p\cdot\frac{1}{(1-q)^2} = p\cdot\frac{1}{p^2} =\\ = &\ \frac{1}{p}\end{aligned}$$

où à $(*)$ je triche un peu – ça demanderait plus de justification.

• Exercice : calculer l'espérance de la loi de Poisson.

• Le reste du tableau sera admis. $\square$

Définition 5.11. Une *médiane* d'une variable aléatoire X est un nombre réel m satisfaisant:

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

Remarque 5.12. 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.)

6. Variables aléatoires continues

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 extrème : tous les nombres appartenant 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 6.1. (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 6.2. Une v.a. $X\colon \Omega \to \mathbb{R}$ est *continue* s'il existe une fonction $f\colon \mathbb{R} \to \mathbb{R}_{\geqslant 0}$ 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 de la loi de X).

Définition 6.3. La *loi* d'une v.a. continue $X\colon \Omega \to \mathbb{R}$ est la probabilité $\mathbb{P}_X$ sur $\mathbb{R}$ déterminée par

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

Remarque 6.4. Une fonction $f\colon \mathbb{R} \to \mathbb{R}$ qui est la densité d'une loi satisfait deux propriétés :

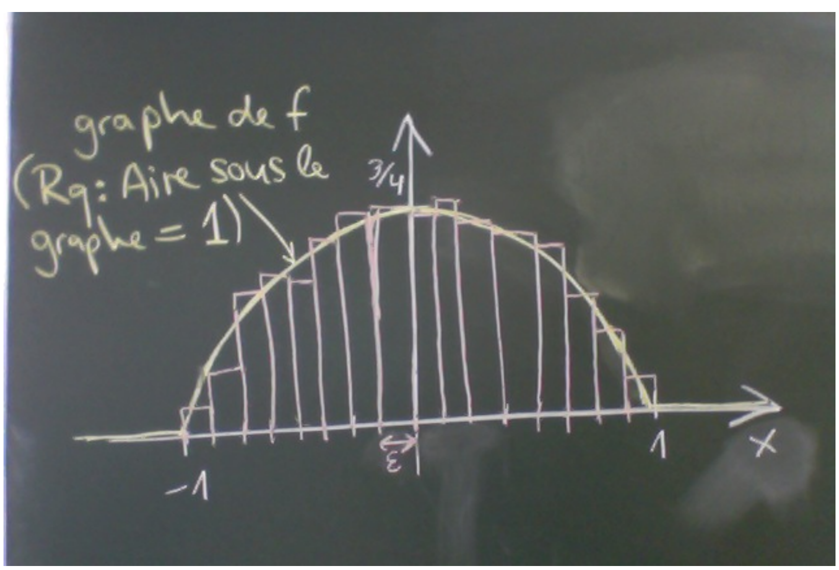
(1) $f(x) \geqslant 0$ pour tout $x \in \mathbb{R}$,
(2) f est intégrable et $\int_{-\infty}^{+\infty} f(x) = \mathbb{P}(X \in]-\infty, +\infty[) = \mathbb{P}(X \in \mathbb{R}) = 1$

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

Exemple 6.5.

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

S'il pleut N gouttes de pluie sur $\mathbb{R}$ dont les positions sont tirés selon la loi de densité f, alors il pleut le plus fort en 0 et pas du tout en dehors de $[-1,1]$ (nuage au-dessus de $[-1,1]$). Dessinons un histogramme : on partitionne $\mathbb{R}$ en boîtes de largeur ϵ, et on compte quelle boîte a reçu combien de gouttes. Si dans dans une certaine boîte il y a k gouttes, alors on dessine une barre de hauteur $\frac{k}{\epsilon \cdot N}$. Alors l'histogramme va ressembler au graphe de f.



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

Exemple 6.6. La durée de vie (en minutes) d'une particule élémentaire d'une certaine substance radioactive peut être modélisé par une variable aléatoire de densité

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

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

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

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

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

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

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

alors

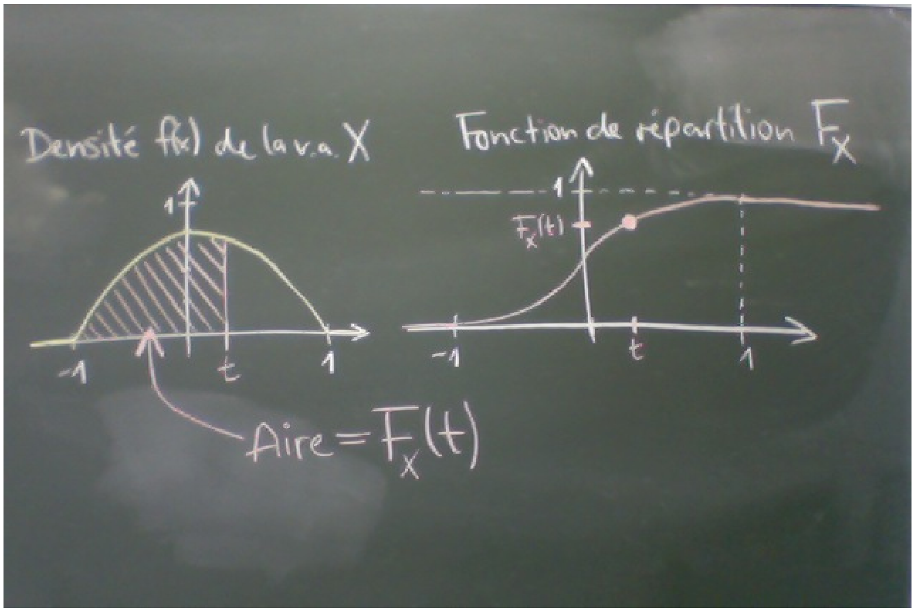
$$\mathbb{P}(X \leqslant t) = 1 - e^{-\lambda t}$$

Définition 6.7. La *fonction de répartition* F_X d'une variable aléatoire continue X est définie comme dans le cas discret :

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

Donc si X est de densité f, alors

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



La densité de l'exemple 6.5 et sa fonction de répartition F_X.

Remarque 6.8. 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.

Remarque 6.9. Si f est la densité de la v.a. X, alors $F_X(t) = \int_{-\infty}^{t} f(x)\, dx$. Donc F_X est une primitive de f, et $F_X' = f$.

Définition 6.10. 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 6.11. 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

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

Remarque 6.12. Comme dans le cas discret on démontre

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

Aussi, les analogues des Propositions 5.5 et 5.9 sont valables pour les v.a. continues.

Lois continues classiques

Voir le Tableau 2.

Exemple 6.13. (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 $[0, 1]$. Formellement :

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

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

Notation $X \sim \mathcal{U}(a, b)$.

DESSINS de la densité et de la fonction de répartition $F_X(t)$. Bien indiquer que l'aire sous la densité est 1, et rappeler que l'aire à gauche de t est $F_X(t)$.

Calculs 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 $[c, d]$.

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

(donc la valeur moyenne est le milieu de l'intervalle $[a, b]$), et

$$\begin{aligned} \mathbb{V}(X) &= \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \int_a^b \frac{x^2}{b-a}\, dx - \left(\frac{a+b}{2}\right)^2 \\ &= \left[\frac{x^3}{3(b-a)}\right]_{x=a}^{x=b} - \left(\frac{a+b}{2}\right)^2 = \frac{b^3 - a^3}{3(b-a)} - \frac{a^2 + 2ab + b^2}{4} \\ &= \ldots \left(= \tfrac{4b^3 - 4a^3}{12(b-a)} - \tfrac{3(b-a)(a^2+2ab+b^2)}{12(b-a)} = \tfrac{-a^3+3a^2b-3ab^2+b^3}{12(b-a)} = \tfrac{(b-a)^3}{12(b-a)}\right) \\ &= \frac{(b-a)^2}{12} \end{aligned}$$

Exemple 6.14. La loi exponentielle

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

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

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

La loi exponentielle est l'analogue continu de la loi géométrique. Par exemple, dans un standard téléphonique, je note les moments exacts d'appels. Si je commence à regarder à 9h17 pile, alors le temps d'attente jusqu'au premier appel va suivre une loi géométrique.

Plus formellement, en TD on démontrera : si un certain événement arrive en moyenne λ fois par minute ($\lambda \in \mathbb{R}$), mais les différentes occurrences sont indépendantes, alors le temps d'attente (en minutes) jusqu'à la première occurrence de l'événement suit une loi exponentielle $\mathcal{E}(\lambda)$.

Calculs • Fonction de répartition : $F_X(t) = \mathbb{P}(X \leqslant t) = 1 - e^{-\lambda t}$ (Dém.: Exemple 6.6).

DESSIN densité (bien indiquer aire sous f est 1) et fonction de répartition

• Espérance : $\mathbb{E}(X) = \frac{1}{\lambda}$. Démonstration:

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

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 $u' = 1$, $v = \frac{1}{-\lambda} e^{-\lambda x}$).

• Variance : par un calcul semblable (exercice) on montre :

$$\mathbb{V}(X) = \frac{1}{\lambda^2}$$

Exercice Démontrer que c'est la loi d'une v.a. sans mémoire (voir Exercice 4.15). [Pas de mémoire, c.à.d., si je sais qu'il n'y a pas eu d'appel téléphonique depuis 10min, ça ne m'apporte aucune information sur la probabilité d'appels futurs.]

Exemple 6.15. 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 PS1).

Définition 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).

On va admettre • $\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{(x-\mu)^2}{2\sigma^2}}\ dx = 1$ donc c'est bien une densité de probabilité. [Les étudiants de maths, math-éco et prépa ingénieurs verront la démonstration de cette égalité en deuxième année d'études, en VAR.]
• $\mathbb{E}(X) = \mu$
• $\mathbb{V}(X) = \sigma^2$

La loi normale est l'analogue continu de la loi binomiale. Rappellons qu'une v.a. X suit une loi binomiale $\mathcal{B}(n, \frac{1}{2})$ si

$$\mathbb{P}(X = k) = \frac{1}{2^n} C_n^k$$

Ses valeurs sont donc proportionnelles aux valeurs de la $n$ème ligne du triangle de Pascal (en forme de "cloche"). Entre autres, le TCL donnera un sens exact à la phrase : "Pour des grandes valeurs de n, la loi Binomiale $\mathcal{B}(n,p)$ est presque égale à une loi normale". Voir aussi le Tableau 2.

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

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

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 Wikipdia sur la Loi Normale.]

[DESSIN densité et Φ. Rappeller que Φ = aire sous f à gauche de t.]

Exemple 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 :

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

Donc

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

Exemple 6.16. On dit une variable aléatoire X est *log-normale* si $Y = \ln(X)$ est normale.

Exemple 6.17. Si $X_1, \ldots, X_k$ sont des v.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 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 le factoriel : $\Gamma(n+1) = n!$.

Exemple 6.18. Une autre loi utilisée en statistiques est la *loi logistique* (voir feuille distribuée), et il y a plein d'autres.

7. Indépendance de variables aléatoires

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

Définition 7.1. Soient $X, Y\colon \Omega \to \mathbb{R}$ 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

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

sont indépendantes; autrement dit, si

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

Remarque 7.2. Si X et Y sont discrètes il suffit de vérifier que pour tout $x \in X(\Omega)$ et $y \in X_2(\Omega)$,

$$\mathbb{P}(X = x \text{ et } Y = y) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y)$$

Proposition 7.3. *Soient X et Y deux variables indépendantes. Alors*

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

Démonstration (que dans le cas où X, Y sont discrètes).

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

Définition 7.4. La *covariance* de deux v.a. X et Y est le nombre

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

Proposition 7.5. $Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$

Démonstration

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

Remarque 7.6. Si X et Y sont deux v.a. *indépendantes*, alors d'après Proposition 7.3, $\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$, et donc

$$Cov(X, Y) = 0$$

Exemple 7.7. ATTENTION, la réciproque est fausse

$$X, Y \text{indep} \overset{\Longrightarrow}{\not\Longleftarrow} Cov(X, Y) = 0$$

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

$$Cov(X, Y) = 0$$

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 $\mathbb{P}(X = 0, Y = 0) \neq \mathbb{P}(X = 0) \cdot \mathbb{P}(Y = 0)$ donc

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

Interprétation 7.8. Si $Cov(X, Y) > 0$, 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 $Cov(X, Y) < 0$, alors X et Y sont "négativement correlés".

Théorème 7.9. *(Propriétés de la covariance) Soient X et Y deux variables aléatoires. Alors*

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

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

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

8. Loi des grands nombres et théorème central limite

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.

Exemples : (a) J'ai une pièce et je veux savoir quelle est la probabilité p de "Pile" (juste ou truquée ?). Soit $\Omega = \{\text{Pile, Face}\}$, $\mathbb{P}(\text{Pile}) = p$, $\mathbb{P}(\text{Face}) = 1 - p$, $X\colon\ \Omega \to \{0,1\}$, Pile $\mapsto 1$, Face $\mapsto 0$, on veut estimer $\mathbb{E}(X) = p$.

(b) Soit $\Omega = \{\text{adultes français}\}$ $X(\omega) = 1$ si ω va voter Hollande, $X(\omega) = 0$ sinon. Je veux estimer $\mathbb{E}(X) =$ pourcentage du vote pour Hollande. Je peux téléphoner 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 $s_n := \frac{x_1+\ldots+x_n}{n}$. Alors, pour n très grand $s_n \overset{approx}{=} \mu$.

Questions : (a) pourquoi ? (b) Comment choisir n pour avoir 95% confiance que

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

Théorème 8.1. *(Inégalité de Markov) Soit $X : \Omega \to \mathbb{R}_+$ une variable aléatoire à valeurs non–négatives. Alors pour tout nombre réel $\lambda > 0$*

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

Remarque : 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 lois !]

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 8.2. *(Inégalité de Tchebychev, ou de Bienaymé–Tchebychev) Soit $X\colon \Omega \to \mathbb{R}$ une v.a. sur un éspace probabilisé (Ω, 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 les événements loin de la moyenne – il dit que les évenéments extrèmes sont rares.

Remarque : 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 tous les lois.]

Démonstration

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

Exemple 8.3. Soit X une v.a., et 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 :

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

Donc la probabilité est inférieure à 1%

Exemple 8.4. 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

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

Définition 8.5. Deux v.a. $X, Y\colon\ \Omega \to \mathbb{R}$ sont indépendantes et identiquement distribuées (i.i.d.), si X et Y sont indépendantes et si elles suivent la même loi de probabilité: $\mathbb{P}(X \leqslant a) = \mathbb{P}(Y \leqslant a)$.

[La *loi* des grands nombres - ici, le mot "loi" veut dire "règle" ou "théorème".]

Théorème 8.6. *(La loi faible des grands nombres) Soit $X_1, X_2, \ldots, X_n, \ldots$ une suite de variables indépendantes et identiquement distribuées. On suppose que toutes admettent l'espérance finie $\mathbb{E}(X_i) = \mu$. Alors pour tout $\epsilon > 0$,*

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

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

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

Démonstration du théorème pour le cas special quand les variables considérées ont variance $\mathbb{V}(X_i) = \sigma^2$ *finie.* Dans ce cas, l'indépendance implique, d'après le Théorème 7.9(4),

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

D'après la linéarité de l'espérance,

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

et d'après Tchebychev

$$\mathbb{P}\left(|S_n - \mu| > \epsilon\right) \;\leqslant\; \frac{\sigma^2}{n\epsilon^2} \;\overset{n\to\infty}{\longrightarrow}\; 0, \qquad \text{c.q.f.d.}$$

[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,

$$s_n = \frac{x_1 + \ldots + x_n}{n} \overset{n\to\infty}{\longrightarrow} \mathbb{E}(X)]$$

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*

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

Alors la loi de la v.a.

$$\sqrt{n}\left(\frac{S_n - \mu}{\sigma}\right)$$

tend vers la loi normale $\mathcal{N}(0,1)$ lorsque $n \to \infty$. Ceci veut dire que la fonction de répartition de $\sqrt{n}\left(\frac{S_n-\mu}{\sigma}\right)$ tend, en tout point, vers la fonction de répartition Φ de la loi normale :

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

Démonstration admise.

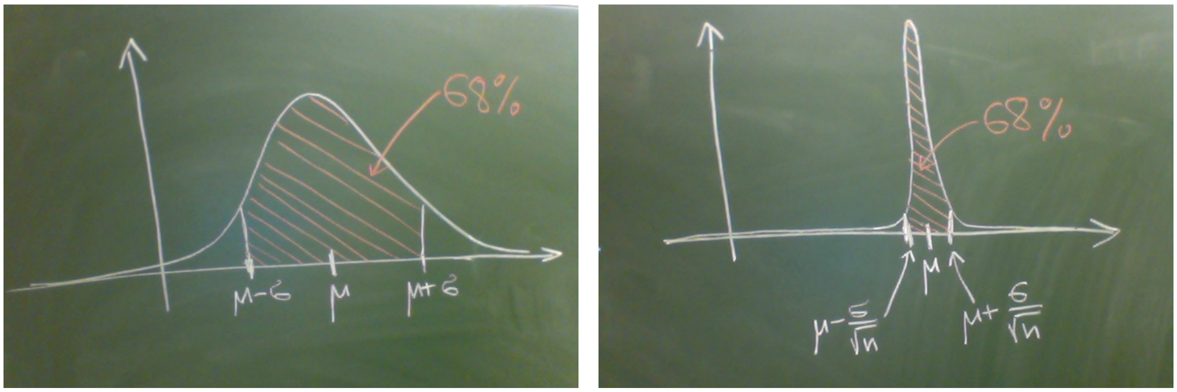
[C'est un théorème concernant une limite, et le théorème est central à la théorie de la probabilité - d'où le nom bizarre.]

Exemple 8.9. Soient $X_1, \ldots, X_{10}$ des variables aléatoires indépendantes uniformes sur intervalle $(0,1)$. On cherche à évaluer $\mathbb{P}(\frac{X_1+\ldots+X_{10}}{10} > 0{,}6)$.

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

$$\begin{aligned}
\mathbb{P}\left(\frac{X_1 + \ldots + X_{10}}{10} > 0,6\right) &= 1 - \mathbb{P}\left(S_{10} \leqslant 0,6\right) \\
&= 1 - \mathbb{P}\left(\sqrt{10}\cdot\frac{S_{10} - \frac{1}{2}}{\sqrt{\frac{1}{12}}} \leqslant \sqrt{10}\cdot\frac{0,6 - \frac{1}{2}}{\sqrt{\frac{1}{12}}}\right) \\
&= 1 - \mathbb{P}\left(\qquad\qquad\qquad \leqslant \sqrt{1,2}\right) \\
&\overset{TCL}{\approx} 1 - \Phi(\sqrt{1,2}) \approx 0,1367.
\end{aligned}$$

Interprétation du TCL : On a vu en TD que dans la loi $\mathcal{N}(\mu, \sigma^2)$, à peu près 68% des nombres tirés sont entre $-\sigma$ et σ (et 95, 4% sont entre -2σ et 2σ). Si X est une v.a. arbitraire d'espérance μ et d'écart-type σ (variance σ^2), alors pour grand n, la v.a. S_n ressemble à une normale $\mathcal{N}(\mu, \frac{\sigma^2}{n})$, d'espérance μ et d'écart-type $\frac{\sigma}{\sqrt{n}}$.

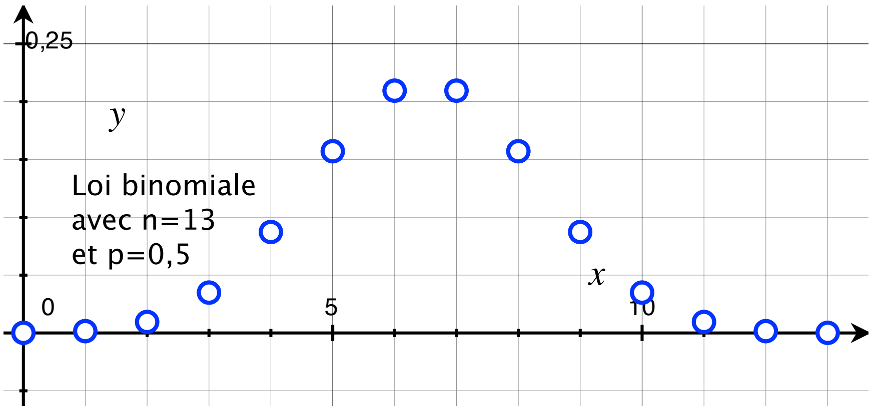


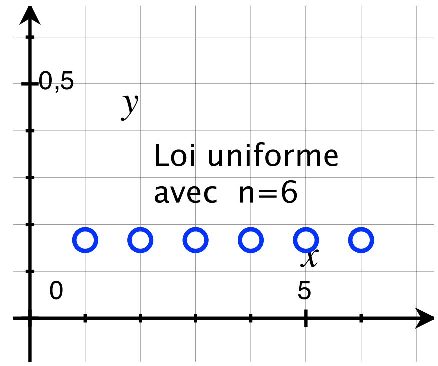
Dessin : la densité de $\mathcal{N}(\mu, \sigma^2)$ et $\mathcal{N}(\mu, \frac{\sigma^2}{n})$

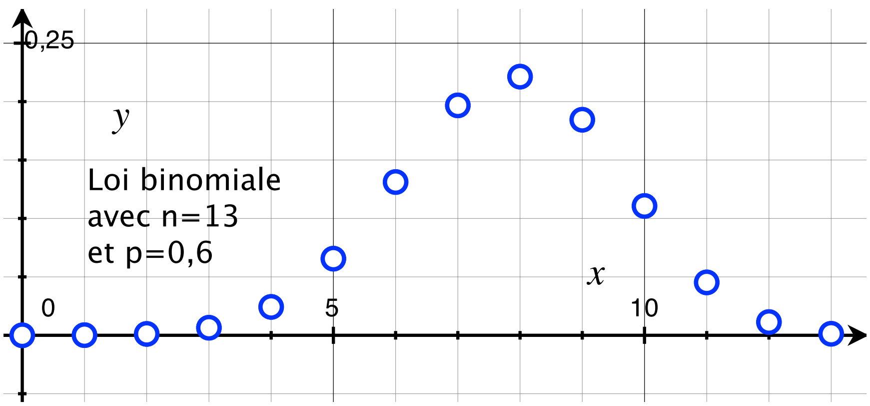
Si je tire n nombres et je calcule leur moyenne s_n, j'ai $\sim$ 68% confiance que le résultat est à distance au plus $\frac{\sigma}{\sqrt{n}}$ de la vraie espérance μ. Si je veux diviser par 3 l'erreur d'estimation attendu, je dois multiplier n, le nombre de tirages, par 9.

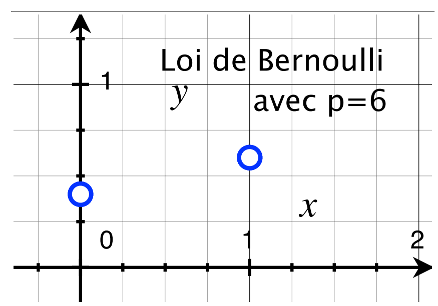
TABLE 1. **Lois discrètes classiques**

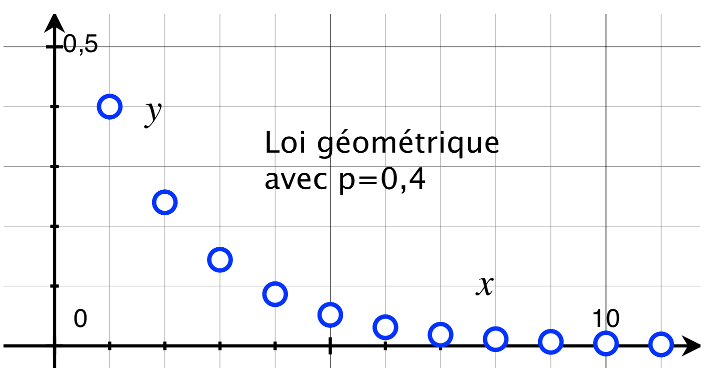
Dénomination	**Loi**	**Espérance**	**Variance**
Loi Uniforme $X \sim \mathcal{U}(\llbracket 1, n \rrbracket)$	$X(\Omega) = \llbracket 1, n \rrbracket$ $\mathbb{P}(X = k) = \frac{1}{n}$	$\mathbb{E}(X) = \frac{n+1}{2}$	$\mathbb{V}(X) = \frac{n^2-1}{12}$
Loi de Bernoulli $X \sim \mathcal{B}(1, p)$	$X(\Omega)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


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


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


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


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


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

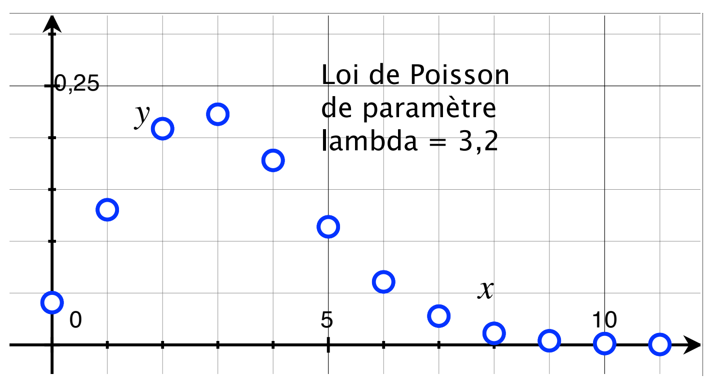
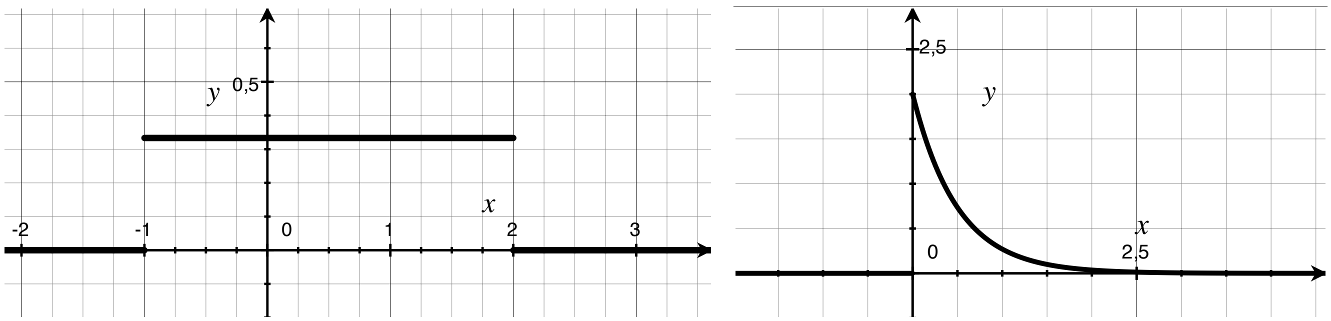
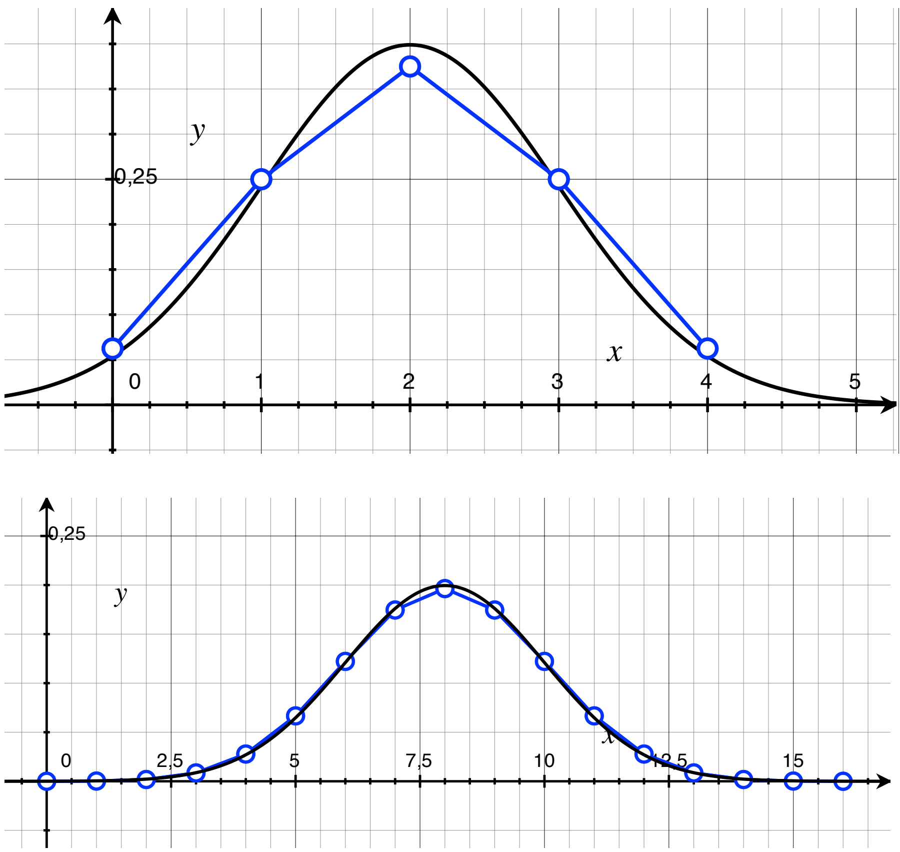

0,25
Loi de Poisson
de paramètre
lambda = 3,2
y
x
0
5
10

Table 2. **Lois discrètes classiques**

Dénomination	**Densité**	**Espérance**	**Variance**
Loi Uniforme $X \sim \mathcal{U}([a,b])$	$f(x) = \frac{1}{b-a}, \quad x \in [a,b]$ $f(x) = 0, \quad x \notin [a,b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Loi Exponentielle $X \sim \mathcal{E}(\lambda)$	$f(x) = \lambda e^{-\lambda x}, \quad x \geqslant 0$ $f(x) = 0, \quad x < 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Loi Normale $X \sim \mathcal{N}(\mu, \sigma^2)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \ \forall x \in \mathbb{R}$	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 $k \in \mathbb{N}$	$f(x) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}, \ x \geqslant 0$ 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 $\mathcal{U}(-1,2)$. Sa densité est de $\frac{1}{3}$ sur l'intervalle $[-1,2]$ et de 0 en-dehors de cet intervalle. (b) Loi exponentielle $\mathcal{E}(2)$



Comparaison de la loi Binomiale $\mathcal{B}(n,\frac{1}{2})$ avec la loi normale de la même espérance et la même variance, pour $n=4$ et $n=16$

Exercices pour le cours PS1

1. Ensembles et dénombrements

Exercice 1.1. Dans l'ensemble $\Omega = \{1, 2, 3, 4, 5, 6\}$, on considère les trois sous-ensembles

$$A = \{1, 2, 3, 4\}, \quad B = \{1, 3, 5\}, \quad C = \{3, 4, 5, 6\}$$

Déterminer les sous-ensembles suivants

- $A \cup (B \cap C)$ et $(A \cup B) \cap C$
- $(A \cup B)^c$
- $A \backslash B$
- $A \cap B \cap C$
- $A \Delta B$

Exercice 1.2. Trouver un exemple d'ensembles A, B, C tels que $(A\backslash B)\backslash C \neq A\backslash(B\backslash C)$

Exercice 1.3. Soit Ω un ensemble, et $A \subset \Omega$ un sous-ensemble. Supposons $card(A) = p$, et $card(\Omega) = n$.

(a) Quel est le nombre de sous-ensembles de Ω ?

(b) Quel est le nombre de sous-ensembles de Ω contenant A ?

(c) Quel est le nombre de sous-ensembles de Ω disjoints de A ?

Exercice 1.4. On veut placer n convives autour d'une table circulaire avec n chaises. Combien y a-t-il de dispositions possibles, sachant que deux dispositions sont identiques si chaque convive a les mêmes voisins.

Exercice 1.5. Combien de séries de résultats possibles (tenant compte de l'ordre) y a-t-il si l'on jette un dé quatre fois ? Et combien de séries contenant au moins un 6 ?

Exercice 1.6. Un jeu de cartes contient 32 cartes (16 noires et 16 rouges). On tire trois fois (sans remise). Combien de séries de résultats y a-t-il ? Et combien d'entre eux contiennent exactement une carte rouge ?

Exercice 1.7. (Coefficients binomiaux)

(a) Si $0 \leqslant k \leqslant n$, alors $C_n^k = C_n^{n-k}$

(b) Si $1 \leqslant k \leqslant n$, alors $k \cdot C_n^k = n \cdot C_{n-1}^{k-1}$

(c) Si $a, b \in \mathbb{R}$ et $n \in \mathbb{N}$, alors $(a+b)^n = \sum_{k=0}^n C_n^k a^k b^{n-k}$ (formule du binôme de Newton)

(d) Si $n \in \mathbb{N}$, alors $\sum_{k=0}^n C_n^k = 2^n$

(e) Si $n \in \mathbb{N}$, alors $\sum_{k=0}^n (-1)^k C_n^k = 0$. (Par exemple : $C_4^0 - C_4^1 + C_4^2 - C_4^3 + C_4^4 = 0$)

Exercice 1.8. Il y a 4 types de gâteaux dans une patisserie. De combien de façons peut-on acheter 7 gâteaux ?

2. Espaces probabilisés

Exercice 2.1. De combien de manières peut-on mettre 8 personnes autour d'une table ronde avec 8 places

(a) si aucune restriction est mise.

(b) s'il y a une famille de 3 qui doit être assis ensemble.

(c) Si on place les 8 invités aléatoirement, quelle est la probabilité que la famille est assise ensemble par pure chance ?

Exercice 2.2. On considère un jeu avec 32 cartes.

(a) Combien y a-t-il de donnes de 4 cartes ?

(b) Combien y a-t-il de donnes de 4 cartes contenant exactement deux rois ?

(c) Quelle est la probabilité qu'une donne de 4 cartes contienne exactement deux rois ?

Exercice 2.3. Une urne contient 2 boules rouges et 5 noires. Les joueurs A et B tirent à tour de rôle une boule, sans remise, jusqu'à ce que une boule rouge sorte (A commence). Quelle est la probabilité que ce soit A qui tire la première boule rouge ?

Exercice 2.4. Dans une lotérie, le joueur doit choisir 8 nombres entre 1 et 40. Le tirage sélectionne 8 nombres parmi les 40. En admettant que le tirage est équiprobable pour les C_{40}^{8} combinaisons, quelle est la probabilité que le joueur ait

(a) les 8 bons nombres ?

(b) 7 parmi les 8 bons nombres ?

(c) aucun bon nombre ?

3. Probabilité conditionnelle et indépendance

Exercice 3.1. (Démonstration que la probabilité conditionnelle est bien une probabilité) Soit $\mathbb{P}$ une probabilité sur un univers Ω, et soit $B \subset \Omega$ un événement avec $\mathbb{P}(B) > 0$. Montrer que la fonction qui à l'événement A associe le nombre $\mathbb{P}(A|B)$ est une probabilité sur Ω.

Exercice 3.2. Pour une famille avec exactement deux enfants, calculer les probabilités suivantes.

(a) Sachant qu'il y a au moins une fille, quelle est la probabilité que l'autre enfant est aussi une fille ?

(b) Sachant que l'ainée est une fille, quelle est la probabilité que l'autre enfant est aussi une fille ?

Exercice 3.3. (Indépendance multiple) Commençons avec une définition : trois événements A, B, C sont dits *deux à deux indépendants* si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) \quad \text{et} \quad \mathbb{P}(A \cap C) = \mathbb{P}(A) \cdot \mathbb{P}(C) \quad \text{et} \quad \mathbb{P}(B \cap C) = \mathbb{P}(B) \cdot \mathbb{P}(C).$$

Ils sont dits *mutuellement indépendants* si

$$\text{ils sont deux à deux indépendants et en plus } \mathbb{P}(A \cap B \cap C) = \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C)$$

Soit, par exemple $\Omega = \{\text{familles avec deux enfants}\}$. Considérons les trois événéments $A =$"la fratrie est mixte", $B =$"l'enfant ainé est une fille", $C =$"le cadet est un garçon".

(a) Montrer que ces trois événements sont deux à deux indépendants.

(b) Montrer que ces trois événements ne sont pas mutuellement indépendants.

Exercice 3.4. Dans une population on trouve une proportion de $\frac{1}{10000}$ individus qui portent un certain virus. Il y a un test pour la présence de ce virus. Ce test n'est pas parfait : si un individu porte le virus, alors le test le détecte avec une probabilité de 0,99. Si un individu ne porte pas le virus, le test donne un résultat positif (érroné) avec probabilité de 0,001.

(a) Si on tire un individu au hasard de la population, on lui fait passer le test, et le résultat est positif, quelle est la probabilité qu'il porte vraiment le virus ?

(b) À première vue, le résultat obtenu en (a) est extêmement surprenant ! Expliquez-le en quelques phrases françaises.

Exercice 3.5. (Le jeu des trois portes) Le jeu des trois portes était un jeu télévisé populaire (Let's make a deal) diffusé dans les années 1970 aux États-Unis. Le joueur est placé devant 3 portes fermées. Derrière l'une d'elles se trouve une voiture que le joueur peut gagner. Derrière les deux autres se trouve une chèvre. Le présentateur connait la position de la voiture. Le joueur doit d'abord désigner une porte. Puis le présentateur ouvre une porte qui est ni celle choisie par le candidat, ni celle qui cache la voiture. Le candidat a alors le droit ou bien d'ouvrir la porte initialement choisie, ou bien de changer son choix vers la troisième porte.

(a) Quelles sont ses chances de gagner la voiture s'il garde son choix initial ?

(b) Quelles sont ses chances de gagner la voiture s'il choisit la troisième porte ?

Exercice 3.6. (Difficile)

Dans une voiture de TGV toutes les places sont vendues. La première personne qui arrive est un peu tête-en-l'air, et se met aléatoirement à une place (mais dans la bonne voiture). Chaque personne suivante qui arrive applique la règle suivante : si sa place est encore libre, elle se met sur sa place prévue, si sa place est déjà prise, elle se met aléatoirement sur une des places encore libres. Quelle est la probabilité que la dernière personne qui arrive pourra se mettre sur sa place prévue ? (Indication : cette question peut se faire sans aucun calcul et sans mentionner les probabilités conditionnelles. Or, pour avoir une idée, vous pouvez calculer les cas où une voiture de TGV n'a que 2, 3, ou 4 sièges.)

Exercice 3.7. A quelle condition deux événements incompatibles sont-ils indépendants?

Exercice 3.8. On considère trois cartes: une avec les deux faces rouges, une avec les deux faces blanches, et une avec une face rouge et une face blanche. On tire une carte au hasard. On expose une face au hasard. Elle est rouge. Parieriez-vous que la face cachée est blanche?

Exercice 3.9. Avant de partir en vacances tu pries ton voisin de bien vouloir arroser une plante. Sans arrosage, elle mourra avec la probabilité $0,8$; avec arrosage, elle mourra avec la probabilité $0,15$. Tu es sûr à 90 % que ton voisin l'arrosera.

(1) Quelle est la probabilité que la plante soit vivante à ton retour?
(2) Si elle est morte, quelle est la probabilité que le voisin ait oublié de l'arroser?

Exercice 3.10. Soit n un entier positif. On considère n individus $I_1, I_2, \ldots, I_n$; ces individus mentent avec probabilité $p\,(0 < p < 1)$, et leurs comportements sont indépendants. Une information (sous forme de oui ou non) est donnée à I_1 qui la transmet à I_2, ... qui la transmet à I_n, qui l'annonce au monde. Quelle est la probabilité p_n pour que l'information soit fidèlement transmise, c.à.d. que l'annonce de I_n coincide avec l'information donnée à I_1 ? Montrer que $\lim_{n\to\infty} = \frac{1}{2}$. (Indication : pour la deuxième question, se souvenir des solutions des question 1(d) et (e) sur la feuille 2.)

4. Variables aléatoires discrètes

Exercice 4.1. La fonction de répartition d'une v.a. Y est la suivante:

$$F_Y(x) = \begin{cases} 0 & \text{si } x < 0 \\ 0.5 & \text{si } 0 \leqslant x < 1 \\ 0.6 & \text{si } 1 \leqslant x < 2 \\ 0.8 & \text{si } 2 \leqslant x < 3 \\ 0.9 & \text{si } 3 \leqslant x < 3.5 \\ 1 & \text{si } x \geqslant 3,5 \end{cases}$$

Calculer la loi de probabilité de Y.

Exercice 4.2. On lance deux dés honnêtes. On note X le plus grand des numéros obtenus. Déterminer la loi de X.

Exercice 4.3. Soit X une variable aléatoire qui suit une loi géométrique de paramètre p – donc $\mathbb{P}(X = k) = p \cdot (1-p)^{k-1}$ pour $k = 1, 2, 3, \ldots$ Démontrer que X est une "variable aléatoire sans mémoire" :

$$\mathbb{P}(X > k) = \mathbb{P}(X > K + k \mid X > K)$$

Indication : montrer d'abord que $\mathbb{P}(X > k) = (1-p)^k$.

Exercice 4.4. Admettons que le nombre d'erreurs par page dans un livre suive la loi de Poisson de paramètre $\lambda = 0{,}5$. Calculer la probabilité que, sur une page donnée, il y a au moins 3 erreurs.

Exercice 4.5. Soit X une variable aléatoire de Poisson avec paramètre λ. Pour une valeur $k \in \mathbb{N}$ donnée, quelle est la valeur de λ qui maximise $\mathbb{P}(X = k)$?

5. Espérance et variance

Exercice 5.1. On lance deux dés honnêtes. On note X le plus grand des numéros obtenus. Déterminer la loi de X. Calculer son espérance et sa variance.

Exercice 5.2. On lance une pièce de monnaie un certain nombre de fois jusqu'à obtenir "pile". On s'arrête la première fois où on obtient "pile". On touche alors une somme d'argent égale à 2 puissance le nombre de fois où on a obtenu "face". On note X cette somme. Quelle est l'espérance de X?

Exercice 5.3. Soit X une variable aléatoire discrète. Que peut-on dire si la variance de X est 0 ?

Exercice 5.4. Soit X une variable aléatoire. On appelle variable aleatoire centrée réduite associée à X, la variable aléatoire X^* définie par:

$$X^* = \frac{X - \mathbb{E}(X)}{\sigma(X)}.$$

Trouver $\mathbb{E}(X^*)$ et $\mathbb{V}(X^*)$.

Exercice 5.5. Soir X une variable aléatoire telle que $\mathbb{E}(X) = 1$ et $\mathbb{V}(X) = 5$. Calculer $\mathbb{E}((2+X)^2)$ et $\mathbb{V}(4+3X)$.

Exercice 5.6. On tire au hasard 5 cartes d'un jeu de 32 cartes avec remise. Soit X, la variable aléatoire égale au nombre de rois obtenus. Donner la loi de X, son espérance, sa variance, et son écart-type.

Exercice 5.7. Toujours avec un jeu de 32 cartes, on effectue une série infinie de tirages successifs, en remettant chaque fois la carte tirée.

(a) Soit Y, le rang d'apparition du premier roi. Donner la loi de Y, son espérance et sa variance.

(b) Soit Z, le nombre de cartes autres qu'un roi qu'il aura fallu tirer pour obtenir le premier loi. Donner, sans calcul, la loi de Z, son espérance et sa variance.

Exercice 5.8. Pour une variable aléatoire X binomiale d'espérance 6 et de variance 2,4 trouver $\mathbb{P}(X = 5)$.

Exercice 5.9. Une urne contient des jetons numérotés de 1 à n. On les tire un à un sans remise jusqu'à obtenir le plus petit. On note X le nombre de tirages ainsi effectués. Déterminer la loi de X, $\mathbb{E}(X)$, $\mathbb{V}(X)$ et $F_X(x)$.

Exercice 5.10. Une urne contient 2^n papiers sur lesquels sont reproduits les 2^n parties d'un ensemble E à n éléments. On tire un papier au hasard. Soit X, la variable aléatoire égale au cardinal de la partie tirée. Déterminer la loi de X, et donner sans calcul les valeurs de $\mathbb{E}(X)$ et $\mathbb{V}(X)$.

Exercice 5.11. On joue au pile ou face avec une pièce avec $\mathbb{P}(\text{pile}) = p$, et $\mathbb{P}(\text{face}) = 1-p$. On effectue des lancers successifs. Soit X, la variable aléatoire égale au rang de la 2^{eme} apparition de pile. Trouver la loi de X. Trouver l'espérance (difficile) et la variance (encore plus difficile) de X.

6. Variables aléatoires continues

Exercice 6.1. Le nombre de minutes qu'un joueur de base-ball particulier se trouve sur le terrain suit la densité suivante :

$$f(x) = \begin{cases} 0 & \text{si } x < 10 \\ 0,025 & \text{si } 10 \leqslant x < 20 \\ 0,05 & \text{si } 20 \leqslant x < 30 \\ 0,025 & \text{si } 30 \leqslant x < 40 \\ 0 & \text{si } x > 40 \end{cases}$$

Trouver la probabilité que ce joueur soit actif :

(a) plus de 15 minutes.

(b) entre 20 et 35 minutes.

(c) moins de 30 minutes.

Exercice 6.2. Une variable aléatoire X a une densité

$$f(x) = \begin{cases} c \cdot x^4 & \text{si } 0 < x < 1 \\ 0 & \text{sinon} \end{cases}$$

Trouver c, $\mathbb{E}(X)$, $\mathbb{V}(X)$, et la médiane.

Exercice 6.3. Un nombre est choisi au hasard, selon la loi uniforme, sur le segment $[0, 1]$. Trouver la probabilité que le rapport entre le plus petit et le plus grand segment soit inférieur à $\frac{1}{4}$.

Exercice 6.4. Soit N un entier. On choisit indépendamment N nombres réels sur l'intervalle $[0, N]$ selon la loi uniforme $\mathcal{U}(0, N)$. On note X_N le plus petit des nombres obtenus. On va étudier la variable aléatoire X_N :

(a) Pour tout $t \in [0, N]$, calculer $\mathbb{P}(X_N > t)$.

(b) En déduire la fonction de répartition F_{X_N} de X_N.

(c) Pour $t > 0$ fixé, calculer $\lim_{N\to\infty} F_{X_N}(t)$. Indication : vous pouvez admettre la formule (normalement vue en Lycée) : $\lim_{N\to\infty}(1 + \frac{x}{N})^N = e^x$.

(d) Donner un sens exact à la phrase “Pour N très grand, X_N est distribué selon une loi exponentielle $\mathcal{E}(1)$”.

Exercice 6.5. La durée de vie (en minutes) d'une particule élémentaire radioactive peut être modélisée par une variable aléatoire exponentielle, de densité

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

(a) Quelle est la médiane de la durée de vie ?

(b) Quelle est la demi-vie ?

Exercice 6.6. (a) Soit X une variable aléatoire de loi normale centrée réduite $\mathcal{N}(0,1)$. Rappellons que nous notons $\Phi\colon \mathbb{R} \to \mathbb{R}_+$ la fonction de répartition de cette loi. Sachant que $\Phi(-1) = 0,15866\ldots$, $\Phi(0) = 0,5$ et $\Phi(1) = 0,84134$, déterminer la probabilité

$$\mathbb{P}(X \in [\mathbb{E}(X) - \sigma, \mathbb{E}(X) + \sigma])$$

(b) Soit $X \sim \mathcal{N}(\mu, \sigma^2)$, pour $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}_+$ arbitraires. Déterminer la probabilité

$$\mathbb{P}(X \in [\mathbb{E}(X) - \sigma, \mathbb{E}(X) + \sigma])$$

Exercice 6.7. Supposons que vous avez à votre disposition un générateur de nombres aléatoires, qui engendre des nombres réels selon la loi uniforme sur l'intervalle $[0,1]$. Vous souhaitez engendrer des nombres aléatoires, mais pas selon la loi uniforme $\mathcal{U}(0,1)$ mais selon une certaine loi dont vous connaissez la fonction de répartition $F(t)$. Supposons que cette fonction est continue et strictement croissante, de sorte qu'il y a une fonction inverse G – ça veut dire que

$$F(t) = u \quad \Leftrightarrow \quad t = G(u)$$

ou encore $F(G(u)) = u$ pour tout u.

(a) Choisissons Y avec notre générateur, selon la loi uniforme sur $[0,1]$. Montrer que $\mathbb{P}(G(Y) \leqslant t) \;=\; F(t)$.

(b) Donc en pratique, que doit-on faire pour engendrer une v.a. dont la fonction de répartition est F ?

(c) Par exemple, expliquez comment engendrer des nombres aléatoires selon une loi expenentielle $\mathcal{E}(1)$ si l'on n'a qu'un générateur de loi $\mathcal{U}(0,1)$.

Exercice 6.8. (a) Un poste à incendie doit être installé lelong une route forestière de longueur A ($A \in \mathbb{R}$). Si les incendies se déclarent en des points uniformément réparties sur $[0,A]$, où doit-on placer ce poste de façon à minimiser l'espérance entre le poste et l'incendie ?

(b) (difficile) Supposons que la route soit infiniment longue, s'étendant de 0 à l'infini $+\infty$. Si la distance entre un incendie et l'origine est exponentiellement distribué selon une loi $\mathcal{E}(1)$, où doit-on installer le poste à incendie ?

Exercice 6.9. Soit X une variable aléatoire qui est distribué selon la loi exponentielle de paramètre 1, c.à.d., $X \sim \mathcal{E}(1)$. Montrer que X est "sans mémoire", au sens suivant : si $K > 0$ et $k > 0$ alors

$$\mathbb{P}(X > k) = \mathbb{P}(X > K + k \mid X > K)$$

7. Indépendance de variales aléatoires

Exercice 7.1. Soient X et Y, deux variables aléatoires de Bernoulli de même paramètre p, $0 < p < 1$, indépendantes. On définit les variables aléatoires $S = X + Y$ et $D = X - Y$.

(a) Les variables S et D, sont-elles indépendantes?

(b) Calculer $Cov(S, D)$.

Exercice 7.2. Soient X et Y deux variables aléatoires indépendantes. Trouver la loi de $X + Y$ si

(a) $X \sim \mathcal{P}(\lambda_1)$ et $Y \sim \mathcal{P}(\lambda_2)$ (lois de Poisson de paramètres λ_1 et λ_2). Indication : la réponse est que $X + Y$ suit une loi de Poisson de paramètre $\lambda_1 + \lambda_2$.

(b) $X \sim \mathcal{B}(n_1, p)$ et $Y \sim \mathcal{B}(n_2, p)$. Indications : la réponse est que $X + Y$ suit également une loi binomiale : $X + Y \sim \mathcal{B}(n_1 + n_2, p)$. Vous avez le droit d'admettre la formule de Vandermonde :

$$\sum_{k=0}^{n} C_a^k \cdot C_b^{n-k} = C_{a+b}^n$$

Exercice 7.3. On jette deux dés. Soient X et Y respectivement la plus grande et la plus petite des valeurs obtenues.

(a) Calculer la loi de probabilité conditionnelle de Y, sachant que $(X = i)$, pour $i = 1, 2, \ldots, 6$.

(b) X et Y sont-elles indépendantes?

Exercice 7.4. Soient $X \colon \Omega \to \{0, 2, 4\}$ et $Y \colon \Omega \to \{0, 1, 2, 3\}$ deux variables aléatoires. Nous supposons que les probabilités $\mathbb{P}(X = i, Y = j)$ sont comme indiqués dans le tableau suivant.

$X \setminus Y$	0	1	2	3
0	2/48	6/48	3/48	1/48
2	4/48	12/48	6/48	2/48
4	2/48	6/48	3/48	1/48

(a) Déterminer la loi de X. Déterminer la loi de Y.

(b) Les variables aléatoires X et Y, sont-elles indépendantes?

8. Loi des grands nombres et théorème central limite

Exercice 8.1. On suppose que le nombre de pièces sortant d'une usine donnée en l'espace d'une semaine est une variable aléatoire d'espérance 50.

(a) Trouver une borne supérieure sur la probabilité que la production de la semaine prochaine soit d'au moins 75 pièces.

(b)* Y a-t-il une borne sur la probabilité que la production soit inférieure à 30 pièces ?

(c) On sait de plus que la variance de la production hebdomadaire est de 25. Peut-on estimer la probabilité que la production de la semaine prochaine soit strictement comprise entre 40 et 60 pièces?

Exercice 8.2. On considère une variable aléatoire continue X obéissant a la loi uniforme sur l'intervalle $[0, 10]$.

(a) Calculez son espérance et sa variance.

(b) Majorez la quantité $\mathbb{P}(|X-5| \geqslant 5)$ grace à l'inégalité de Tchebychev. Que vaut en fait cette probabilité?

Exercice 8.3. Si $\mathbb{E}(X) = 75$, $\mathbb{E}(Y) = 75$, $\mathbb{V}(X) = 10$, $\mathbb{V}(Y) = 12$ et $Cov(X, Y) = -3$, chercher des bornes supérieures à

(a) $\mathbb{P}(|X - Y| \geqslant 15)$

(b)* $\mathbb{P}(X \geqslant Y + 15)$

(c)* $\mathbb{P}(Y \geqslant X + 15)$

Exercice 8.4. On lance cent fois une pièce de monnaie équilibrée.

(a) Majorez, à l'aide de l'inégalité de Tchebychev, la probabilité d'avoir plus de 70 fois Face ou moins de 30 fois Face à l'issue de ces tirages.

(b) À l'aide du théorème central limite, estimez la même probabilité.