

Variables aléatoires à densité

I. Fonction de répartition d'une variable aléatoire réelle

On fixe un espace probabilisé $(\Omega, \mathcal{A}, P)$. Une variable aléatoire réelle est une application $X : \Omega \rightarrow \mathbb{R}$ telle que $\{X \leq x\}$ soit un événement pour tout réel x . Dans le cas discret, la loi de X se lit sur les nombres $P(X = k)$. En revanche, pour une durée de vie ou une erreur de mesure, chaque valeur isolée a une probabilité nulle. C'est pourquoi on décrit la loi par une fonction qui cumule les probabilités.

Définition :

La **fonction de répartition** de X est la fonction $F_X : \mathbb{R} \rightarrow [0,1]$ définie par $F_X(x) = P(X \leq x)$.

Théorème :

La fonction F_X est croissante, continue à droite en tout point, et vérifie $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$. De plus, pour $a < b$, on a $P(a < X \leq b) = F_X(b) - F_X(a)$ et $P(X = a) = F_X(a) - \lim_{x \rightarrow a^-} F_X(x)$.

Démonstration :

D'abord, si $x \leq y$, alors $\{X \leq x\} \subset \{X \leq y\}$, donc $F_X(x) \leq F_X(y)$. Ensuite, la suite d'événements $\{X \leq x + 1/n\}$ décroît vers $\{X \leq x\}$. Par continuité décroissante de P , on

obtient $F_X(x + 1/n) \rightarrow F_X(x)$; la monotonie donne alors la continuité à droite. De même, $\{X \leq -n\}$ décroît vers l'ensemble vide et $\{X \leq n\}$ croît vers Ω , d'où les deux limites. Enfin, $\{X < a\}$ est la réunion croissante des $\{X \leq a - 1/n\}$, ce qui donne la formule du saut.

Ainsi, la fonction de répartition présente un saut en chaque valeur de probabilité non nulle. Autrement dit, F_X est continue en a si et seulement si $P(X = a) = 0$. On admet aussi que F_X caractérise la loi : deux variables de même fonction de répartition ont la même loi.

Exemple :

Si X suit la loi de Bernoulli de paramètre p , alors $F_X(x) = 0$ pour $x < 0$, $F_X(x) = 1 - p$ pour $0 \leq x < 1$ et $F_X(x) = 1$ pour $x \geq 1$. La fonction est en escalier, avec deux sauts de hauteurs $1 - p$ et p .

II. Variables aléatoires à densité

1. Définition et premières propriétés

Le cas continu remplace la somme des $P(X = k)$ par une intégrale généralisée. Les intégrales de ce chapitre sont donc des intégrales impropres, dont la convergence se justifie toujours.

Définition :

Une **densité de probabilité** est une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ positive, continue sauf en un nombre fini de points, telle que $\int_{-\infty}^{+\infty} f(t) dt$ converge et vaille 1. On dit que X est une **variable à densité** de densité f si, pour tout réel x ,

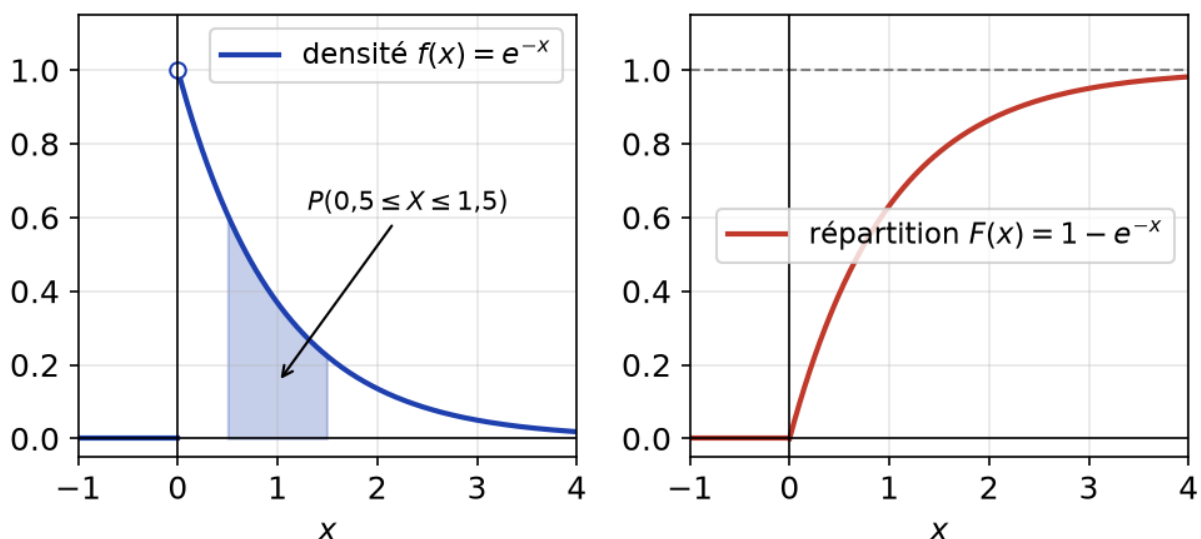
$$F_X(x) = \int_{-\infty}^x f(t) dt.$$

Propriété :

Soit X de densité f . Alors F_X est continue sur $\mathbb{R}$, donc $P(X = a) = 0$ pour tout réel a . De plus, F_X est dérivable en tout point x où f est continue, avec $F'_X(x) = f(x)$. Enfin, pour $a \leq b$,

$$P(a \leq X \leq b) = P(a < X < b) = \int_a^b f(t) dt.$$

Par conséquent, les inégalités larges ou strictes donnent la même probabilité. En revanche, la valeur $f(x)$ n'est pas une probabilité : elle peut dépasser 1. En effet, seule l'aire sous la courbe a un sens probabiliste. La figure ci-dessous montre la densité et la fonction de répartition de la loi exponentielle de paramètre 1 ; l'aire colorée vaut $P(0,5 \leq X \leq 1,5)$.



Méthode :

Pour **vérifier qu'une fonction f est une densité** : 1) montrer que $f \geq 0$; 2) vérifier la continuité sauf en un nombre fini de points ; 3) justifier la convergence de $\int_{-\infty}^{+\infty} f$, par exemple par comparaison ; 4) calculer cette intégrale et vérifier qu'elle vaut 1. Si f dépend d'une constante c , la condition 4 détermine c .

Exemple :

Soit $f(x) = \frac{c}{1+x^2}$. Cette fonction est continue et positive si $c \geq 0$. De plus,
 $\int_{-A}^B \frac{dt}{1+t^2} = \arctan B - \arctan(-A)$ tend vers π . Donc f est une densité si et seulement si $c = 1/\pi$. C'est la **loi de Cauchy**.

2. Reconnaître une variable à densité

Souvent, on connaît d'abord F_X . Le critère suivant permet alors de conclure.

Théorème :

Si F_X est continue sur $\mathbb{R}$ et de classe C^1 sauf en un nombre fini de points, alors X est à densité. Une densité est obtenue en posant $f = F'_X$ là où F_X est dérivable, et une valeur positive arbitraire ailleurs.

Remarque :

Une densité n'est pas unique : modifier f en un nombre fini de points ne change aucune intégrale. On parle donc d'une densité de X , et non de la densité.

III. Les lois usuelles à densité

1. Loi uniforme

Définition :

Soit $a < b$. La variable X suit la **loi uniforme** sur $[a, b]$, notée $\mathcal{U}([a, b])$, si elle admet pour densité $f(x) = \frac{1}{b-a}$ sur $[a, b]$ et $f(x) = 0$ ailleurs. Sa fonction de répartition vaut $\frac{x-a}{b-a}$ sur $[a, b]$, 0 avant et 1 après.

Ainsi, la probabilité d'un sous-intervalle de $[a, b]$ est proportionnelle à sa longueur. C'est l'analogue continu de l'équiprobabilité.

2. Loi exponentielle

Définition :

Soit $\lambda > 0$. La variable X suit la **loi exponentielle** de paramètre λ , notée $\mathcal{E}(\lambda)$, si elle a pour densité $f(x) = \lambda e^{-\lambda x}$ pour $x \geq 0$ et $f(x) = 0$ pour $x < 0$. Alors $F_X(x) = 1 - e^{-\lambda x}$ pour $x \geq 0$.

Propriété :

Si $X \sim \mathcal{E}(\lambda)$, alors X est **sans mémoire** : pour tous $s, t \geq 0$,
 $P(X > s + t \mid X > s) = P(X > t)$.

Démonstration :

En effet, $P(X > t) = e^{-\lambda t}$. Comme $\{X > s+t\} \subset \{X > s\}$, le quotient vaut
 $e^{-\lambda(s+t)} / e^{-\lambda s} = e^{-\lambda t}$.

Cette propriété est l'analogue de celle de la loi géométrique dans le cas discret. Elle modélise des durées de vie sans vieillissement, comme la désintégration d'un noyau radioactif.

3. Loi normale

On admet ici le résultat de Gauss $\int_{-\infty}^{+\infty} e^{-t^2/2} dt = \sqrt{2\pi}$, qui se démontre par une intégrale double en coordonnées polaires.

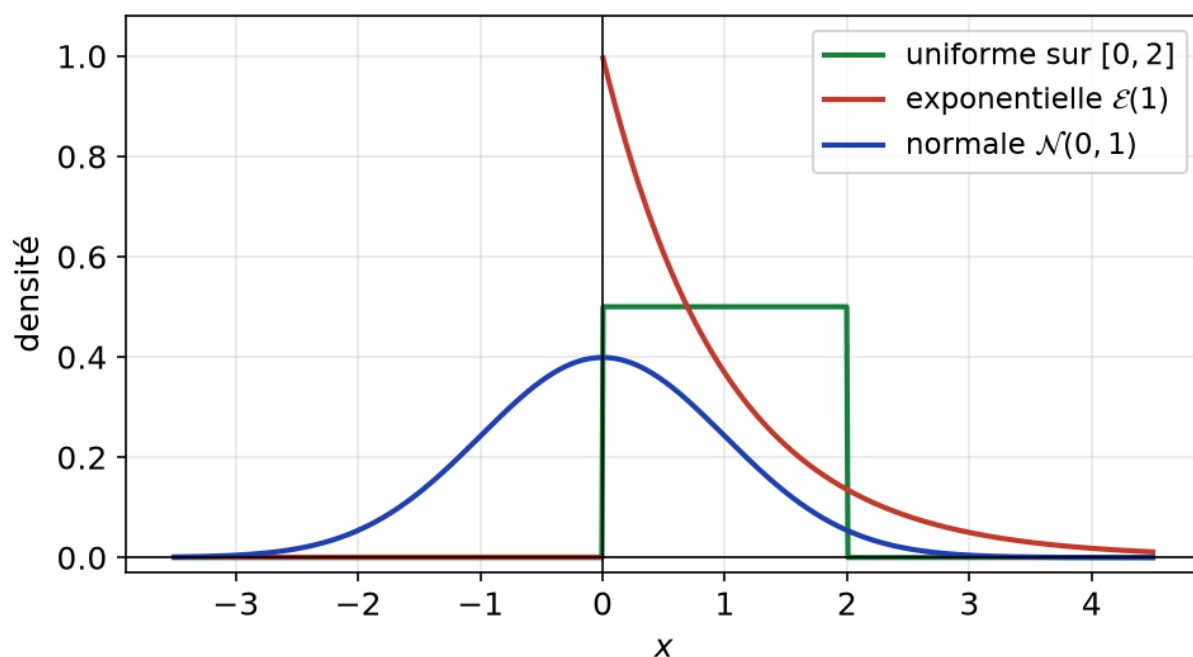
Définition :

La **loi normale centrée réduite** $\mathcal{N}(0,1)$ a pour densité $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. Sa fonction de répartition se note Φ . Plus généralement, pour $m \in \mathbb{R}$ et $\sigma > 0$, la loi $\mathcal{N}(m, \sigma^2)$ a pour densité

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right).$$

La figure ci-dessous compare les densités des trois lois usuelles. On y voit le plateau de la loi uniforme, la

décroissance de la loi exponentielle et la cloche symétrique de la loi normale.



Propriété :

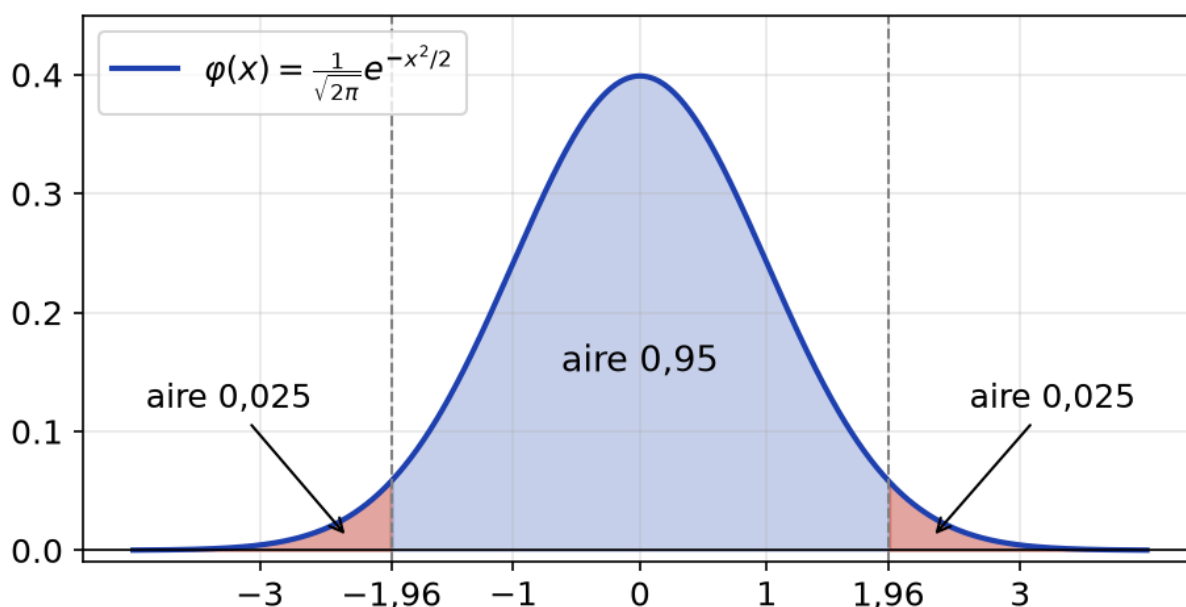
Par parité de φ , on a $\Phi(-x) = 1 - \Phi(x)$ pour tout réel x . Par conséquent, pour $a > 0$,
 $P(|Z| \leq a) = 2\Phi(a) - 1$. De plus, $X \sim \mathcal{N}(m, \sigma^2)$ si et seulement si
 $Z = \frac{X - m}{\sigma} \sim \mathcal{N}(0, 1)$.

Méthode :

Pour **calculer une probabilité normale** : centrer et réduire, puis lire la table de Φ . Par exemple, si $X \sim \mathcal{N}(m, \sigma^2)$, alors $P(X \leq x) = \Phi\left(\frac{x - m}{\sigma}\right)$. Pour une valeur négative, on utilise $\Phi(-u) = 1 - \Phi(u)$.

Exemple :

On lit dans la table $\Phi(1,96) \approx 0,975$. Donc $P(|Z| \leq 1,96) \approx 0,95$: la loi normale centrée réduite prend ses valeurs dans $[-1,96; 1,96]$ avec une probabilité de 95 %. La figure ci-dessous représente cette aire centrale et les deux queues de 2,5 %.



IV. Espérance, variance et théorème de transfert

1. Espérance d'une variable à densité

Dans le cas discret, $E(X) = \sum_k k P(X = k)$ sous réserve de convergence absolue. Dans le cas continu, la somme devient une intégrale.

Définition :

Soit X de densité f . On dit que X admet une **espérance** si $\int_{-\infty}^{+\infty} |t| f(t) dt$ converge. On

pose alors $E(X) = \int_{-\infty}^{+\infty} t f(t) dt$.

Attention :

La convergence absolue est exigée. Par exemple, pour la loi de Cauchy, $\frac{|t|}{\pi(1+t^2)} \sim \frac{1}{\pi|t|}$ en $\pm\infty$. Cette intégrale diverge, donc la loi de Cauchy n'a pas d'espérance, bien que sa densité soit paire.

2. Théorème de transfert

Théorème :

Théorème de transfert. Soit X de densité f , nulle hors d'un intervalle I , et g continue sur I sauf en un nombre fini de points. Alors $g(X)$ admet une espérance si et seulement si

$\int_I |g(t)| f(t) dt$ converge, et dans ce cas

$$E(g(X)) = \int_I g(t) f(t) dt.$$

Ce théorème est admis en L2. Son intérêt est considérable : il calcule $E(g(X))$ sans déterminer la loi de $g(X)$. Il s'agit du parallèle exact de la formule discrète $E(g(X)) = \sum_k g(k) P(X = k)$.

3. Variance

Définition :

Si X^2 admet une espérance, on dit que X admet un moment d'ordre 2. Alors X admet une espérance et la **variance** $V(X) = E((X - E(X))^2)$ vérifie la formule de König-Huygens $V(X) = E(X^2) - E(X)^2$. L'écart type est $\sigma(X) = \sqrt{V(X)}$.

Propriété :

L'espérance est linéaire et croissante. Pour tous réels a, b , on a $E(aX + b) = aE(X) + b$ et $V(aX + b) = a^2 V(X)$.

Proposition :

Voici les moments des lois usuelles.

- Si $X \sim \mathcal{U}([a, b])$, alors $E(X) = \frac{a+b}{2}$ et $V(X) = \frac{(b-a)^2}{12}$.
- Si $X \sim \mathcal{E}(\lambda)$, alors $E(X) = \frac{1}{\lambda}$ et $V(X) = \frac{1}{\lambda^2}$.
- Si $X \sim \mathcal{N}(m, \sigma^2)$, alors $E(X) = m$ et $V(X) = \sigma^2$.

Démonstration :

Traisons la loi exponentielle. Pour $A > 0$, une intégration par parties donne

$$\int_0^A \lambda t e^{-\lambda t} dt = [-t e^{-\lambda t}]_0^A + \int_0^A e^{-\lambda t} dt = -A e^{-\lambda A} + \frac{1 - e^{-\lambda A}}{\lambda}.$$

Quand $A \rightarrow +\infty$, on obtient $E(X) = 1/\lambda$. De même,

$\int_0^{+\infty} \lambda t^2 e^{-\lambda t} dt = \frac{2}{\lambda} E(X) = \frac{2}{\lambda^2}$, donc $V(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$. Pour la loi normale, la fonction $t \varphi(t)$ est impaire et intégrable, donc $E(Z) = 0$. Ensuite, une intégration par parties avec $\varphi'(t) = -t\varphi(t)$ donne $E(Z^2) = \int t \cdot t\varphi(t) dt = \int \varphi(t) dt = 1$. Enfin, $X = m + \sigma Z$ donne $E(X) = m$ et $V(X) = \sigma^2$.

Exemple :

Si $X \sim \mathcal{E}(1)$, le théorème de transfert donne $E(e^{X/2}) = \int_0^{+\infty} e^{t/2} e^{-t} dt = 2$. En revanche, $E(e^X)$ n'existe pas, car $\int_0^{+\infty} e^t e^{-t} dt$ diverge.

V. Loi de l'image $g(X)$ d'une variable à densité

On cherche maintenant la loi de $Y = g(X)$, et pas seulement son espérance. Deux méthodes coexistent.

Méthode :

Par la fonction de répartition. Pour tout y , on écrit $F_Y(y) = P(g(X) \leq y)$. On résout l'inéquation $g(X) \leq y$ en X , puis on exprime le résultat avec F_X . Enfin, on dérive pour obtenir une densité, après avoir vérifié le critère du II.2.

Exemple :

Soit $Z \sim \mathcal{N}(0,1)$ et $Y = Z^2$. Pour $y \leq 0$, $F_Y(y) = 0$. Pour $y > 0$, on a $F_Y(y) = P(-\sqrt{y} \leq Z \leq \sqrt{y}) = 2\Phi(\sqrt{y}) - 1$. En dérivant, $f_Y(y) = \frac{1}{\sqrt{y}} \varphi(\sqrt{y}) = \frac{1}{\sqrt{2\pi y}} e^{-y/2}$ pour $y > 0$. C'est la loi du khi-deux à un degré de liberté.

Théorème :

Méthode de la fonction muette. Si, pour toute fonction h continue et bornée, on a

$$E(h(Y)) = \int_{-\infty}^{+\infty} h(y) q(y) dy \text{ avec } q \text{ une densité, alors } Y \text{ admet } q \text{ pour densité.}$$

En pratique, on écrit $E(h(g(X)))$ par le théorème de transfert, puis on effectue le changement de variable $y = g(x)$. Ainsi, lorsque g est une bijection $\mathcal{C}^1$ strictement monotone d'un intervalle I sur J , on obtient la formule

$$f_Y(y) = f_X(g^{-1}(y)) \left| (g^{-1})'(y) \right| \text{ pour } y \in J.$$

Exemple :

Soit $U \sim \mathcal{U}([0,1])$ et $Y = -\ln(U)$. Pour h continue bornée,

$E(h(Y)) = \int_0^1 h(-\ln u) du$. Le changement de variable $u = e^{-y}$, $du = -e^{-y} dy$, donne $\int_0^{+\infty} h(y) e^{-y} dy$. Donc $Y \sim \mathcal{E}(1)$. Ce résultat sert à **simuler** une loi exponentielle à partir d'un générateur uniforme.

Attention :

Si g n'est pas injective, comme $x \rightarrow x^2$ sur $\mathbb{R}$, il faut découper le domaine. Dans ce cas, la fonction de répartition reste souvent la voie la plus sûre.

VI. Couples à densité, indépendance et somme

1. Densité d'un couple et densités marginales

Définition :

Un couple (X, Y) admet la **densité** $f: \mathbb{R}^2 \rightarrow \mathbb{R}_+$, intégrable d'intégrale 1, si pour tout domaine D raisonnable (par exemple un rectangle ou un domaine élémentaire),

$$P((X, Y) \in D) = \iint_D f(x, y) \, dx \, dy.$$

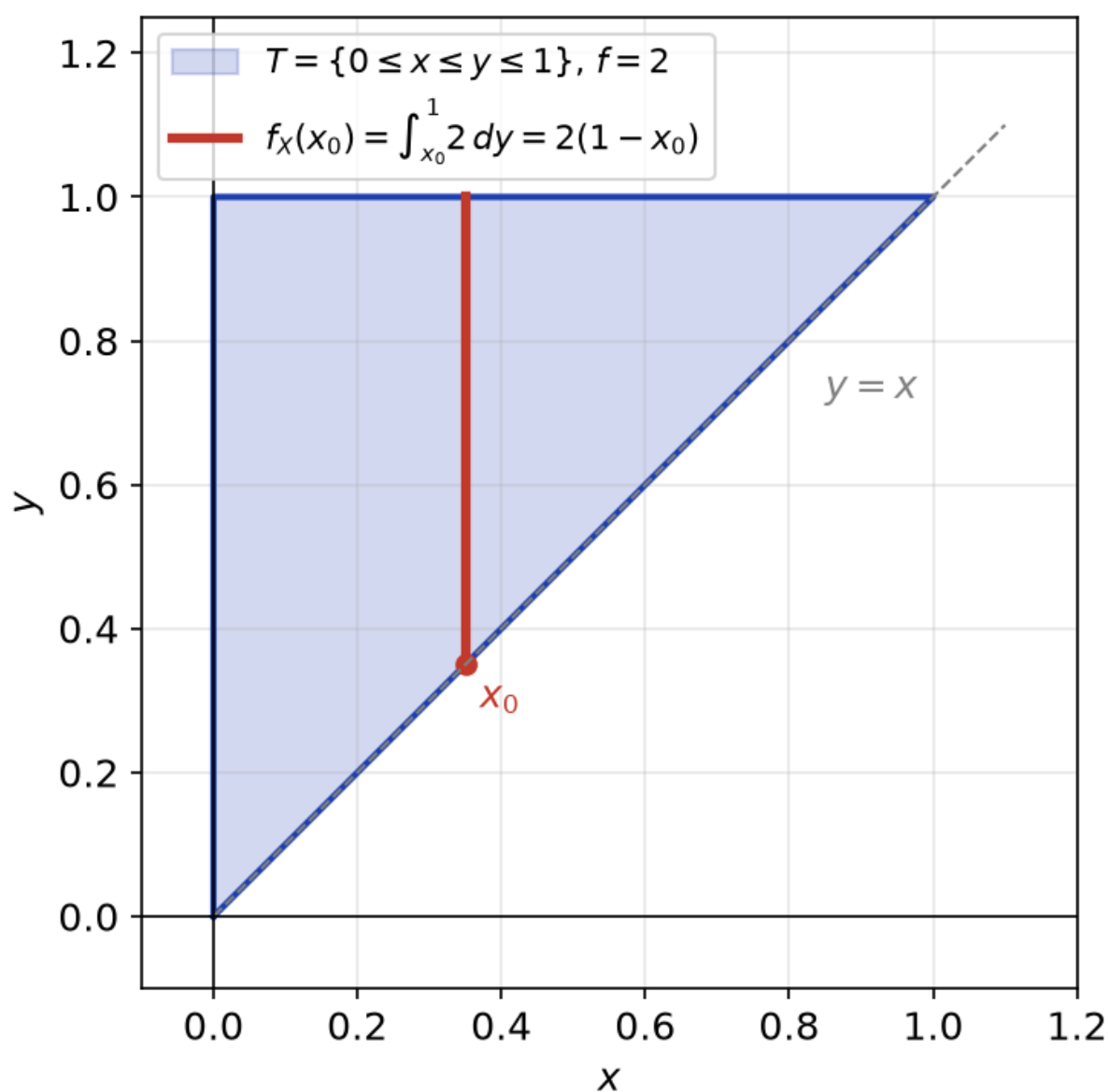
Propriété :

Si (X, Y) a pour densité f , alors X et Y sont à densité, de **densités marginales**

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) \, dy, \quad f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) \, dx.$$

C'est encore l'analogie du cas discret, où $P(X = i) = \sum_j P(X = i, Y = j)$. Par exemple, si $f = 2$ sur le triangle $T = \{0 \leq x \leq y \leq 1\}$ et 0 ailleurs, alors

$f_X(x) = \int_x^1 2 \, dy = 2(1-x)$ sur $[0,1]$. La figure ci-dessous montre le segment vertical sur lequel on intègre.



Le théorème de transfert s'étend aux couples : $E(g(X,Y)) = \iint g(x,y) f(x,y) \, dx \, dy$ sous réserve de convergence absolue. Par exemple, $E(XY)$ permet de calculer la covariance $\text{Cov}(X,Y) = E(XY) - E(X)E(Y)$.

2. Indépendance

Théorème :

Soient X et Y à densités f_X et f_Y . Elles sont **indépendantes** si et seulement si le couple admet pour densité $(x,y) \rightarrow f_X(x) f_Y(y)$.

Remarque :

Si la densité du couple est non nulle exactement sur un domaine qui n'est pas un produit d'intervalles, comme un triangle, les variables ne sont pas indépendantes. Dans ce cas, le produit $f_X f_Y$ serait non nul sur un rectangle plus grand.

3. Loi d'une somme : produit de convolution

Théorème :

Si X et Y sont indépendantes, de densités f_X et f_Y , et si l'une des deux est bornée, alors $S = X + Y$ est à densité, et

$$f_S(s) = (f_X * f_Y)(s) = \int_{-\infty}^{+\infty} f_X(x) f_Y(s - x) dx.$$

Démonstration :

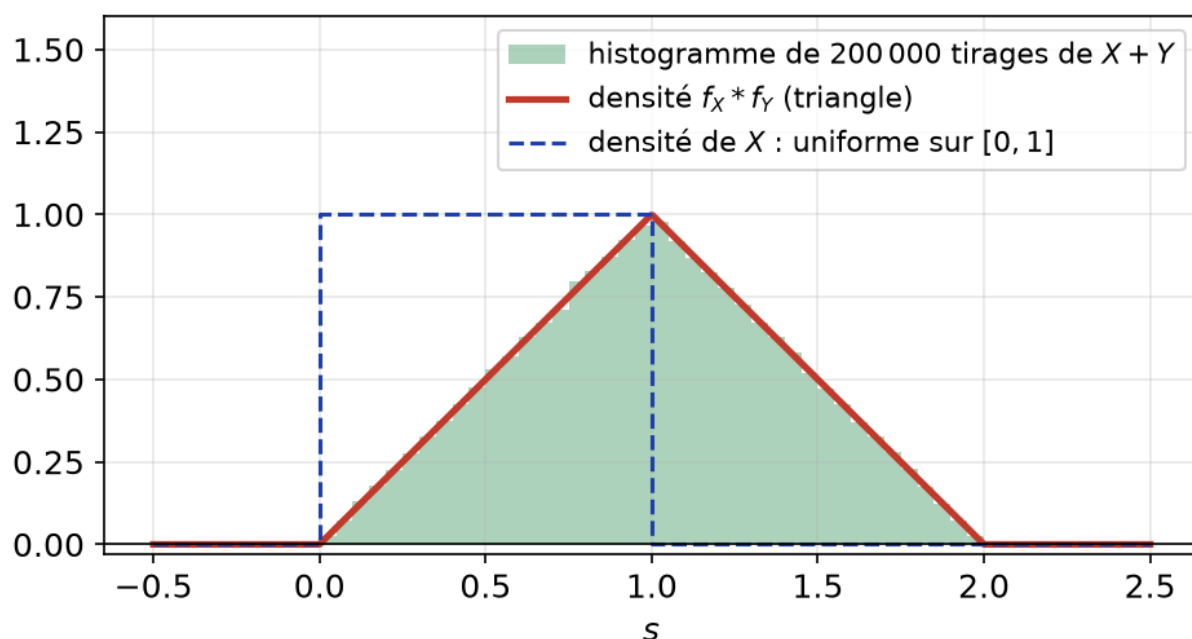
Soit h continue bornée. Par indépendance et transfert,

$E(h(X+Y)) = \iint h(x+y) f_X(x) f_Y(y) dx dy$. À x fixé, on pose $s = x + y$ dans l'intégrale en y . Puis le théorème de Fubini pour les fonctions positives permet d'échanger l'ordre d'intégration : $E(h(S)) = \int h(s) \left(\int f_X(x) f_Y(s-x) dx \right) ds$. La méthode de la fonction muette conclut.

Cette formule prolonge la formule discrète $P(X+Y = n) = \sum_k P(X = k) P(Y = n - k)$. En pratique, la difficulté vient des indicatrices : il faut déterminer, pour chaque s , l'ensemble des x où les deux facteurs sont non nuls.

Exemple :

Soient X, Y indépendantes de loi $\mathcal{U}([0,1])$. Alors $f_S(s) = \int_0^1 1_{[0,1]}(s-x) dx$ est la longueur de $[0,1] \cap [s-1, s]$. On trouve $f_S(s) = s$ sur $[0,1]$, $f_S(s) = 2-s$ sur $[1,2]$ et 0 ailleurs. La figure ci-dessous confronte ce triangle à un histogramme de simulations.



Proposition :

Si $X \sim \mathcal{N}(m_1, \sigma_1^2)$ et $Y \sim \mathcal{N}(m_2, \sigma_2^2)$ sont indépendantes, alors
 $X + Y \sim \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$. De même, la somme de n variables indépendantes de loi $\mathcal{E}(\lambda)$ a pour densité $s \mapsto \frac{\lambda^n s^{n-1}}{(n-1)!} e^{-\lambda s}$ sur $\mathbb{R}_+$.

Méthode :

Pour **calculer la loi d'une somme** : 1) justifier l'indépendance ; 2) écrire l'intégrale de convolution avec les indicatrices ; 3) traduire les conditions en un intervalle pour x , en discutant selon s ; 4) calculer, puis vérifier que le résultat est positif et d'intégrale 1.

VII. Approximation de la binomiale par la loi normale

La loi normale apparaît dès qu'on additionne beaucoup de petites contributions indépendantes. Le résultat suivant en est le premier exemple historique.

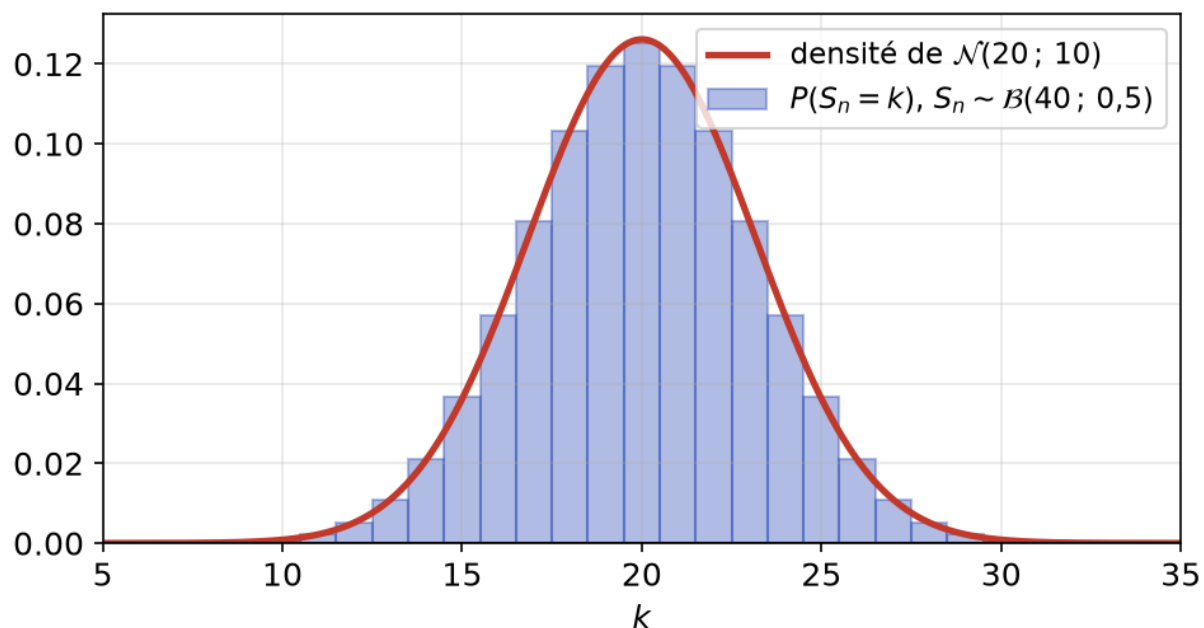
Théorème :

Théorème de Moivre-Laplace. Soit $p \in]0,1[$ et $S_n \sim \mathcal{B}(n,p)$. Posons $Z_n = \frac{S_n - np}{\sqrt{np(1-p)}}$. Alors, pour tous réels $a < b$,

$$\lim_{n \rightarrow +\infty} P(a \leq Z_n \leq b) = \Phi(b) - \Phi(a).$$

Ce théorème est admis en L2. Il se généralise en théorème central limite. En pratique, on approche $\mathcal{B}(n,p)$ par $\mathcal{N}(np, np(1-p))$ lorsque, par exemple, $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$. La

figure ci-dessous superpose le diagramme de $\mathcal{B}(40; 0,5)$ et la densité de $\mathcal{N}(20; 10)$.



Remarque :

On améliore l'approximation par la **correction de continuité**. Comme S_n est à valeurs entières, on remplace $P(k_1 \leq S_n \leq k_2)$ par $P(k_1 - 0,5 \leq T \leq k_2 + 0,5)$, où T suit la loi normale approchante.

Exemple :

On lance 100 fois une pièce équilibrée. Le nombre de piles S suit $\mathcal{B}(100; 0,5)$, de moyenne 50 et d'écart type 5. Avec correction de continuité,
 $P(45 \leq S \leq 55) \approx 2\Phi(1,1) - 1 \approx 2 \times 0,8643 - 1 \approx 0,729$.

Ce qu'il faut retenir

- La fonction de répartition $F_X(x) = P(X \leq x)$ est croissante, continue à droite, de limites 0 et 1, et caractérise la loi.
- Une densité est positive, continue sauf en un nombre fini de points, d'intégrale généralisée égale à 1 ; alors $F'_X = f$ aux points de continuité de f .
- Pour une variable à densité, $P(X = a) = 0$ et $P(a \leq X \leq b) = \int_a^b f$.
- Lois usuelles : $\mathcal{U}([a,b])$, $\mathcal{E}(\lambda)$ sans mémoire, $\mathcal{N}(m, \sigma^2)$ ramenée à $\mathcal{N}(0,1)$ par $Z = (X - m)/\sigma$.
- L'espérance exige la convergence absolue de $\int t f(t) dt$; le théorème de transfert donne $E(g(X)) = \int g f$.
- La loi de $g(X)$ s'obtient par la fonction de répartition ou par la méthode de la fonction muette.
- Les densités marginales s'obtiennent en intégrant la densité du couple par rapport à l'autre variable ; l'indépendance équivaut à une densité produit.
- La densité d'une somme de variables indépendantes est le produit de convolution $f_X * f_Y$.
- Pour n grand, $\mathcal{B}(n,p)$ s'approche par $\mathcal{N}(np, np(1-p))$, avec correction de continuité.

Questions fréquentes sur variables aléatoires à densité

Pourquoi a-t-on $P(X = a) = 0$ pour une variable à densité ?

Parce que la fonction de répartition $F_X(x) = \int_{-\infty}^x f$ est continue, donc elle n'a aucun saut. Or

$P(X = a)$ est exactement la hauteur du saut de F_X en a . Par conséquent, les inégalités larges et strictes donnent les mêmes probabilités.

Une densité peut-elle prendre des valeurs supérieures à 1 ?

Oui. La valeur $f(x)$ n'est pas une probabilité : seule l'aire sous la courbe en est une. Par exemple, la loi uniforme sur $[0 ; 0,5]$ a pour densité 2 sur cet intervalle.

Quelle différence entre le théorème de transfert et la loi de $g(X)$?

Le théorème de transfert calcule $E(g(X)) = \int g(t) f(t) dt$ sans connaître la loi de $g(X)$. En revanche, déterminer la loi de $g(X)$ demande sa fonction de répartition ou la méthode de la fonction muette. Si seule l'espérance est demandée, le transfert suffit.

Quand peut-on approcher une loi binomiale par une loi normale ?

Lorsque n est grand et que p n'est pas trop proche de 0 ou de 1, par exemple $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$. On remplace alors $\mathcal{B}(n, p)$ par $\mathcal{N}(np, np(1-p))$. La correction de continuité améliore nettement le résultat.

Pour aller plus loin en L2

- Les énoncés : [exercices de maths en L2 sur variables aléatoires à densité](#)
- À maîtriser avant : [Variables aléatoires discrètes](#), [Intégrales généralisées](#), [Intégrales multiples et intégrales curvilignes](#)
- Chapitre précédent : [Intégrales multiples et intégrales curvilignes](#)
- Chapitre suivant : [Géométrie affine](#)
- Tester vos connaissances : [QCM de maths en L2 par chapitre](#)
- Le sommaire : [tous les chapitres de maths de L2](#) et la [licence de maths de L1 à L3](#)