CNAM MULHOUSE . Ramm Algebra Center

Formation Ingénieur Informatique

Mathématiques : PROBABILITES

Cours Michel GOZE

Chapitre 5

Variables aléatoires finies ou continues

Soit $(\Omega, \mathfrak{F}, P)$ un espace probabilisé. En général l'ensemble Ω est un ensemble relativement compliqué sur lequel il est difficile de faire des opérations mathématiques. Imaginons par exemple, une étude sur la population des anchois en Méditérannée. Cette population est difficilement discernable. D'où l'idée naturelle de remplacer cet ensemble Ω par un ensemble mathématique plus accessible aux opérations, $\mathbb{R}$, par exemple, en privilégiant un caractère particulier de Ω. Dans tout ce chapitre, on s'intéressera essentiellement au cas où l'ensemble Ω est fini. Dans ce cas, on prendra comme tribu $F = \mathcal{P}(\Omega)$.

1. Définition et Exemples

1.1. **Définition d'une variable aléatoire finie.** Dans le cas où $\Omega = \{\omega_1, \cdots, \omega_N\}$ est un espace probabilisé fini, la notion de variable aléatoire (parfois appelée aussi caractère) est très simple.

Définition 1. *Soit* $(\Omega, \mathfrak{F} = \mathcal{P}(\Omega), P)$ *un espace probabilisé fini. On appelle variable aléatoire sur cet espace toute fonction*

$$X : \Omega \to \mathbb{R}.$$

Une telle variable aléatoire est dite discrète.

Une variable aléatoire prend, dans ce cas fini, qu'un nombre fini de valeurs $x_1, \cdots, x_s$ que nous convenons de classer dans le sens croissant :

$$x_1 < x_2 < \cdots < x_s.$$

L'ensemble fini $\Omega_X = \{x_1, \cdots, x_s\}$ est un sous-ensemble de $\mathbb{R}$ appelé l'ensemble fondamental de la variable aléatoire X. Nous voyons déjà l'intérêt de cette notion de variable aléatoire, nous transposons l'étude des probabilités dans des sous-ensembles (ici finis) de $\mathbb{R}$, alors que l'ensemble de base Ω était plus ou moins bien défini mathématiquement.

Pour chacune de ces valeurs x_i, nous pouvons considérer le sous-ensemble

$$X^{-1}(x_i) = \{\omega \in \Omega,\ X(\omega) = x_i\}.$$

Attention à la notation X^{-1} qui ne signifie pas prendre l'inverse ou la fonction réciproque de X qui n'existe pas en général. Mais cet ensemble $X^{-1}(x_i)$ est l'ensemble de tous les évènements élémentaires ω qui ont pour image le réel fixé x_i, et par définition de x_i, cet ensemble est non vide.

On s'intéressera également au sous-ensemble de Ω défini par

$$X^{-1}(]-\infty, x]) = \{\omega \in \Omega, \quad X(\omega) \leq x\}$$

pour tout $x \in \mathbb{R}$. Si $x < x_1$, il est clair que cet ensemble est vide. Si $x_i \leq x < x_{i+1}$, alors

$$X^{-1}(]-\infty, x]) = X^{-1}(x_1) \cup \cdots \cup X^{-1}(x_i).$$

De même, on définit pour tout segment $[a, b]$ de $\mathbb{R}$, le sous-enesemble de Ω :

$$X^{-1}([a, b]) = \{\omega \in \Omega, \quad a \leq X(\omega) \leq b\}.$$

1.2. **Exemples.**

(1) Soit Ω l'ensemble des épreuves consistant à peser tous les habitants de Mulhouse. La fonction : " à un individu on fait correspondre sa taille" est une variable aléatoire.

(2) La fonction qui à un dé fait correspondre sa marque est une variable aléatoire. Dans ce cas, on a

$$\Omega = \Omega_X = \{1, 2, 3, 4, 5, 6\}.$$

(3) Considérons toujours le lancer d'un dé cubique. La fonction Y qui aux marques paires fait correspondre 1 et aux impiares 0 est une variable aléatoire. Dans ce cas, on a

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

et

$$\Omega_Y = \{0, 1\}.$$

Notons que

$$Y^{-1}(0) = \{1, 3, 5\}$$

et

$$Y^{-1}(1) = \{2, 4, 6\}.$$

(4) Soit Ω l'ensemble des élèves de la classe, $\mathfrak{F} = \mathcal{P}(\Omega)$. L'application $X : \Omega \to \mathbb{R}$ qui à un élève fait correspondre son âge est une variable aléatoire et

$$X^{-1}(]-\infty, x]) = \{\omega \in \Omega, X(\omega) \leq x\} \in \mathfrak{F}$$

correspondra à l'ensemble des élèves qui ont moins (éventuellement égal) de $[x]$ ans, où $[x]$ est la partie entière du nombre réel x.

(5) Considérons le lancer d'un dé et notons par f_i la face du dé correspondant au chiffre i. Alors $\Omega = \{f_1, \cdots, f_6\}$ et X donnée par $X(f_i) = i$ est une variable aléatoire.

(6) On note par h et t les côtés pile et face d'une pièce de monnaie. L'expérience assoiée au lancer de cette pièce correspond à $\Omega = \{h, t\}$. L'application $X : \Omega \to \mathbb{R}$ donnée par

$$X(h) = 1, \ X(t) = 0$$

est une variable aléatoire sur Ω. Si nous considérons maintenant l'expérience associée à n lancers successifs, alors $\Omega_1 = \Omega^n$. Un évènements élémentaire est un mot du type $\omega = hhtth...t$ de longueur n écrit avec les seules lettres h et t. L'application

$$X : \Omega \to \mathbb{R}$$

définie par

$$X(\omega) = \text{nombre de lettres } h \text{ dans le mot } \omega$$

est une variable aléatoire. Elle décrit le nombre de fois où pile est obtenu au cours de n lancers successifs.

1.3. **Variables aléatoires continues.** On suppose ici que l'espace probabilisé est donné par le triplet $(\Omega, \mathfrak{F}, P)$, Ω n'étant pas nécessairement fini et $\mathfrak{F}$ une tribu de Ω.

Définition 2. *Soit $(\Omega, \mathfrak{F} = \mathcal{P}(\Omega), P)$ un espace probabilisé. On appelle variable aléatoire sur cet espace toute fonction*

$$X : \Omega \to \mathbb{R}.$$

Cette variable est dite continue si l'ensemble $X(\Omega)$ est un intervalle ou une réunion d'intervalles de $\mathbb{R}$.

La description d'une loi continue diffère de celles des lois discrètes puisque pour une variable aléatoire continue X, la probabilité que X prenne une valeur bien précise x est nulle :

$$P(X = x) = 0.$$

Nous verrons ceci un peu plus loin. Prenons par exemple comme variable aléatoire X taille d'un individu. On aura $P(X = 1,654321) = 0$.

1.4. **Somme, produit de variables aléatoires.** Dans ce qui suit, sauf précision contraire, on ne distinguera pas les cas continus et discrets.

Soient X et Y deux variables aléatoires sur l'espace probabilisé $(\Omega, \mathfrak{F}, P)$. Comme X et Y sont à valeurs dans $\mathbb{R}$ on peut définir la somme et le produit de ces deux applications :

$$X + Y : \Omega \to \mathbb{R}$$

telle que

$$(X + Y)(\omega) = X(\omega) + Y(\omega).$$

Par exemple, si $X(\Omega) = \{x_1, \cdots, x_p\}$ et $Y(\Omega) = \{y_1, \cdots, y_l\}$, alors

$$(X + Y)(\Omega) = \{x_i + y_j,\ i = 1, \cdots, p,\ j = 1, \cdots, l\}.$$

On définit de même le produit

$$XY : \Omega \to \mathbb{R}$$

par

$$(XY)(\omega) = X(\omega)Y(\omega).$$

Par exemple, si $X(\Omega) = \{x_1, \cdots, x_p\}$ et $Y(\Omega) = \{y_1, \cdots, y_l\}$, alors

$$(XY)(\Omega) = \{x_i y_j,\ i = 1, \cdots, p,\ j = 1, \cdots, l\}.$$

1.5. **Image d'une variable aléatoire.** Soit X une variable aléatoire sur $(\Omega, \mathfrak{F}, P)$ et soit

$$\varphi : \mathbb{R} \to \mathbb{R}$$

une fonction (continue, dérivable, indéfiniment de fois dérivable) d'une variable réelle.

Définition 3. *On appelle variable aléatoire image de X par φ, l'application*

$$Y = \varphi(X) : \Omega \to \mathbb{R}$$

définie par

$$\varphi(X)(\omega) = \varphi(X(\omega)).$$

2. Loi de probabilité d'une variable aléatoire

2.1. **Définition.**

Définition 4. *Soit $(\Omega, \mathfrak{F}, P)$ un espace probabilisé et soit X une variable aléatoire sur cet espace. La loi de probabilité de X est la probabilité sur $\mathbb{R}$ définie par*

$$P_X(]-\infty, x]) = P(\{\omega \in \Omega, \ X(\omega) \leq x\}).$$

En particulier on a :

$$P_X(]a, b]) = P(\{\omega \in \Omega, \ a < X(\omega) \leq b\})$$

ou bien

$$P_X(\{x\}) = P(\{\omega \in \Omega, \ X(\omega) = x\}).$$

Remarquons que P_X est bien une loi de probabilité sur $\mathbb{R}$.

2.1.1. **Cas où Ω est fini.** Posons $\Omega_X = \{x_1, \cdots, x_N\}$. Cette loi de probabilité est entièrement déterminée par les couples (x_i, p_i) avec

$$p_i = P(X^{-1}(x_i)).$$

Notons que

$$p_1 + \cdots + p_N = 1.$$

Exemples.

(1) Une urne contient 3 boules, deux rouges et une blanche. On effectue trois tirages successifs et on note le rang d'apparition de la boule blanche. Ici $\Omega = \{(r, r, b), (r, b, r), (b, r, r)\}$ correspondant au tirage des trois boules et la variable aléatoire

$$X : \Omega \to \{1, 2, 3\}$$

correspond au rang de la boule blanche. La loi de X est donc

$$P_X(]-\infty, x]) = \begin{cases} 0 & \text{si } x < 1 \\ 1/3 & \text{si } 1 \leq x < 2 \\ 1/3 & \text{si } 2 \leq x < 3 \\ 1/3 & \text{si } 3 \leq x \end{cases}$$

On a aussi
$$p_1 = P(X^{-1}(1)) = p_2 = P(X^{-1}(2)) = p_3 = P(X^{-1}(3)) = 1/3.$$

(2) On lance un dé deux fois successivement. Soit X la variable aléatoire correspondant à la somme des chiffres sortis. Ici
$$\Omega = \{(1,1),(1,2),\cdots,(4,6),(5,6),(6,6)\}.$$
C'est un ensemble fini contenant 36 éléments. La variable aléatoire est la fonction
$$X : \Omega \to \mathbb{R}$$
définie par
$$X((i,j)) = i + j$$
avec $(i,j) \in \Omega$. Ainsi X est à valeurs dans $\{2,3,\cdots,12\}$. On a, par exemple,
$$P_X(2) = \frac{1}{36}.$$
Plus généralement la loi de probabilité de X vérifie
$$\begin{cases} P_X(k) = \dfrac{k-1}{36} & \text{si } 2 \leq k \leq 6, \\ P_X(k) = \dfrac{13-k}{36} & \text{si } 7 \leq k \leq 12. \end{cases}$$

Remarque. Rappelons que $P(X = x)$ signifie $P(\{\omega \in \Omega, X(\omega) = x\})$. Soit $A \in \mathcal{P}(\Omega)$. C'est un ensemble fini. Posons
$$A = \{\omega_1, \cdots, \omega_p\}.$$
Alors
$$P(A) = \sum_{k=1}^{p} P(\{\omega_k\}).$$
Soit $x \in \mathbb{R}$ et soit $A = \{\omega \in \Omega,\ X(\omega) \leq x\}$.. Alors
$$P_X(]-\infty, x]) = P(A).$$
On en déduit bien, que la donnée des valeurs $P(X = x)$ détermine complètement la loi de X.

2.1.2. **Cas où X est continue.** Il n'est ainsi pas possible de définir la loi de X par la donnée des probabilités des événements élémentaires. En effet nous avons signalé que dans ce cas $P(X = x) = 0$. Dans le cas fini, il est clair que deux variables aléatoires sont égales si elles prennent les mêmes valeurs. Nous avons totefois dans ce cas la notion plus générale que celle de l'égalité.

Définition 5. *Soient X et Y deux variables aléatoires sur l'espace probabilisé $(\Omega, \mathfrak{F}, P)$. Elles sont égales presque partout si l'ensemble*
$$A = \{\omega \in \Omega,\ X(\omega) \neq Y(\omega)\}$$
qui est dans $\mathfrak{F}$ vérifie
$$P(A) = 0.$$

Nous savons que cela ne signifie pas que A soit vide

Théorème 1. *Soient X et Y deux variables aléatoires sur l'espace probabilisé $(\Omega, \mathfrak{F}, P)$. Si elles sont égales presque partout, alors $P_X = P_Y$.*

La réciproque est fausse. Le fait que X et Y aient la même loi de probabilité ne signifie pas qu'elles soient égales ni même égale presque partout.

Exemple. On lance une pièce équilibrée et on considère le tirage pile ou face dont le résultat est 0 pour pile et 1 pour face.. On effectue trois lancers. Soit X la variable aléatoire correspondant au nombre de fois ou pile apparaît et soit Y la variable aléatoire correspondant au nombre de fois où face apparaît. On a $\Omega = \{0,1\}^3$ et $\mathfrak{F} = \mathcal{P}(\Omega)$. Un résultat s'écrit $\omega = (\alpha_1, \alpha_2, \alpha_3)$ avec $\alpha_i = 0$ ou 1. Alors

$$P(\{\omega\}) = (\frac{1}{2})^3.$$

Soit k un entier inférieur ou égal à 3. Notons par A_k le sous ensemble de Ω défini par

$$A_k = \{\omega \in \Omega, X(\omega) = k\}.$$

Alors

$$P_X(\{k\}) = P(\{\omega, X(\omega) = k) = P(A_k) = C_3^k(\frac{1}{2})^3.$$

De même

$$P_Y(\{k\}) = C_3^k(\frac{1}{2})^3.$$

Les variables X et Y ont la même loi de probabilité.

2.2. **La loi de Bernoulli, loi binomiale.**

Définition 6. *Soit X une variable aléatoire sur un espace probabilisé fini $(\Omega, \mathfrak{F}, P)$ à valeurs dans $\{0, 1, \cdots, n\}$ et soit $p \in [0,1]$.*

(1) On dit X suit la loi de Bernoulli de paramètre p si X est à valeurs dans $\{0,1\}$ et

$$P_X(\{0\}) = p, \; P_X(\{1\}) = 1 - p.$$

On note cette loi $\mathcal{B}(1,p)$.

(2) On dit X suit la loi binomiale de paramètre p si X est à valeurs dans $\{0,1\}$ et

$$P_X(\{k\}) = C_n^k p^k (1-p)^{n-k}$$

pour tout $k = 0, 1, \cdots, n$. On note cette loi $\mathcal{B}(n,p)$.

Comme X est finie, sa loi est bien déterminée dès que l'on connait les probabilités des singletons $\{k\}$. On peut vérifier à titre d'exercice que si X suit la loi de binomilale $\mathcal{B}(n, 1/2)$, alors

$$P_X(\mathbb{R}) = P_X(\{0, 1, \cdots, n\}) = \frac{1}{2}^n \sum_{k=0}^{n} C_n^k = \frac{1}{2}^n (1+1)^n = 1.$$

Nous avons vu, dans les chapitres précédents, que la loi de Bernoulli correspondait à une expérience à deux issues, succés et echec et p est la probabilité pour avoir un succés. La loi binomiale correspond au résultat de la somme de n expériences indépendantes, chacune d'elles suivant la loi de Bernoulli.

3. Fonction de répartition d'une variable aléatoire, densité

3.1. Fonction de répartition. Rappelons la convention :

$$(X \leq x) = \{\omega \in \Omega,\ X(\{\omega) \leq x\}.$$

Définition 7. *Soit X une variable aléatoire sur un espace probabilisé $(\Omega, \mathfrak{F}, P)$. On appelle fonction de répartition de X la fonction numérique*

$$F_X : \mathbb{R} \to \mathbb{R}$$

définie par

$$F_X(x) = P(X \leq x)$$

pour tout $x \in \mathbb{R}$.

On peut voir facilement que cette fonction vérifie les propriétés suivantes :

(1) $F_X(b) - F_X(a) = P(a < X \leq b)$

(2) Pour tout x, $0 \leq F(x) \leq 1$

(3) F est une fonction croissante.

Supposons que X ne prenne que les valeurs $x_1, x_2, \cdots, x_s$ avec $x_1 < x_2 < \cdots < x_s$. Posons

$$P_X(\{x_i\} = p_i.$$

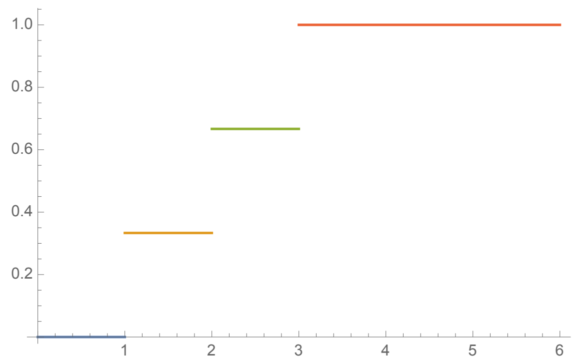
Alors la fonction de répartition F_X est une fonction en escalier, c'est-à-dire constante par morceaux, croissante vérifiant

(1) $F_X(x) = 0$ pour tout $x < x_1$,

(2) $F_X(x) = p_1 + \cdots + p_i$ pour tout $x \in [x_i, x_{i+1}[$

(3) $F_X(1)$ pour tout $x <\geq x_r$

Elle est discontinue pour $x = x_1, \cdots, x_s$, le saut de discontinuité est égal à p_i en chacun de ces points x_i. Le graphe d'une fonction de répartition ressemble à :



Exemples.

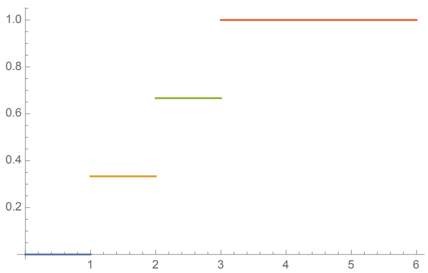
(1) **La loi de Dirac.** Une variable aléatoire X sur l'espace $\Omega = \{\omega_1, \cdots, \omega_N\}$ suit la loi de Dirac d'indice $a \in \mathbb{R}$ si

$$X(\omega_i) = a$$

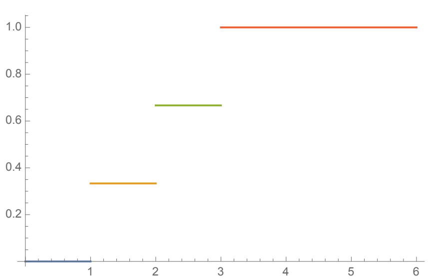
pour tout $i = 1, \cdots, n$. La Fonction de répartition vérifie donc

$$\begin{cases} F_X(x) = 0, \text{ si } x < a, \\ F_X(x) = 1, \text{ si } x \geq a. \end{cases}$$

Le graphe de $F_X(x)$ est (ici $a = 2, 5$) :



(2) Soit X une variable aléatoire sur Ω dont la loi est celle de Bernoulli. On dira dans ce cas que X est une variable aléatoire de Bernoulli de paramètre p si elle prend les valeurs 1 et 0 avec les probabilités respectives p et $q = 1 - p$. Sa fonction de répartition est (ici $p = 1/3$) :



Proposition 1.
$$P(X > x) = 1 - F_X(x).$$

Démonstration. En effet, les évènements $\{X > x\}$ et $\{X \leq x\}$ sont disjoints et leur réunion est Ω. On en déduit

$$P(X > x) + P(X \leq x) = 1$$

d'où la proposition.

3.2. **Fonctions de répartitions des variables continues. Densité.** Dans les exemples précédents, qui ne concernent que des variables finies, les fonctions de répartitions sont des fonctions en escalier. Il n'en est plus de même dans le cas continu.

3.2.1. **Probabilité d'un évènement élémentaire.** Revenons à la remarque précédente dans laquelle on mentionnait que la probabilité d'un événement élémentaire pour une variable continue était nul. Ceci nécessite un peu d'explication. On dit qu'une fonction d'une variable réelle $F : \mathbb{R} \to \mathbb{R}$ est continue en un point x_0 si

$$\lim_{x \to x_0} F(x) = F(x_0).$$

Ceci n'est pas toujours vérifiée, du moins dans un cadre général, pour les fonctions de répartition. Elles vérifient néanmoins la propriété

$$\lim_{x \to x_0, x \geq x_0} F(x) = F(x_0).$$

On dit qu'elle sont continues à droite. Si tel est le cas, on a alors, en général

$$\lim_{x \to x_0, x \leq x_0} F(x) \neq F(x_0)$$

et on note

$$\lim_{x \to x_0, x \leq x_0} F(x) = F(x_0^-).$$

On a alors

$$P(X = x_0) = F_X(x_0) - F_X(x_0^-).$$

Par contre,

Proposition 2. *Supposons que la fonction de répartition $F_X(x)$ de X soit continue en tout point. Alors*

$$P(X = x) = 0.$$

3.2.2. **Fonction de densité.** On suppose maintenant que la fonction $F_X(x)$ est continue et dérivable.

Proposition 3. *Supposons que la fonction de répartition $F_X(x)$ de X soit dérivable en tout point. On appelle fonction de densité de X la fonction dérivée de F_X :*

$$f_X(x) = \frac{dF_X(x)}{dx}.$$

Comme $F_X(x)$ est une fonction croissante, sa dérivée $f_X(x)$ est positive. Rappelons que si $f(x)$ est la dérivée de la fonction $F(x)$, alors $F(x)$ est une primitive de $f(x)$ et on a

$$\int_a^b f(x)dx = F(b) - F(a).$$

Cette intégrale représente l'aire de la partie du plan situé entre le segment $[a, b]$ de l'axe des abscisses et la courbe $f_X(x)$. On en déduit

Proposition 4.
$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

Pour calculer en fonction de la densité les probabilités $P(X \leq x)$ nous avons besoin de la notion d'intégrale généralisée :

Définition 8. *Soit $f(x)$ une fonction continue. Alors*

(1) $\displaystyle\int_{-\infty}^{b} f(t)dt = \lim_{a \to -\infty} \int_a^b f(t)dt.$

(2) $\displaystyle\int_{a}^{+\infty} f(t)dt = \lim_{b \to +\infty} \int_a^b f(t)dt.$

(3) $\displaystyle\int_{-\infty}^{+\infty} f(t)dt = \lim_{b \to +\infty} \int_c^b f(t)dt + \lim_{a \to -\infty} \int_a^c f(t)dt.$

Bien entendu, certaines de ces limites peuvent ne pas exister. Mais ceci est un autre problème que nous n'aborderons pas ici.

Proposition 5.
$$P(X \leq x) = \int_{-\infty}^{x} f(t)dt.$$

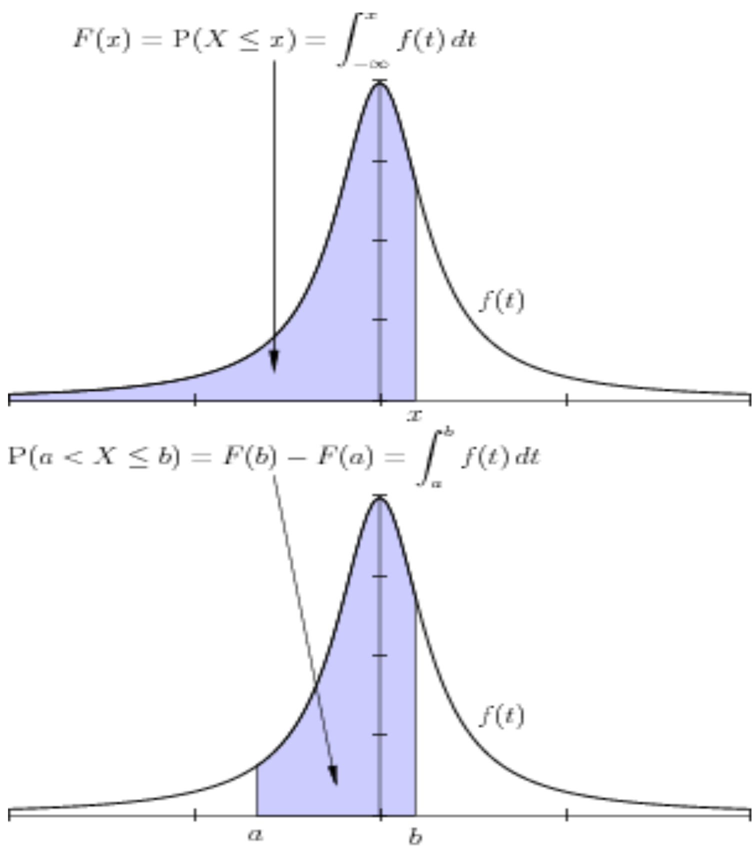
On en déduit

Proposition 6. *Une fonction f est une fonction de densité sur un intervalle I si*

(1) f est continue et positive sur I,

(2) L'aire sous la courbe et délimitée par l'axe des abscisses est égale à 1.

Voici à quoi ressemble un graphe de fonction de densité :



4. Espérance,écart-type d'une variable aléatoire discrète

4.1. **Espérance mathématique.** Soit X une variable aléatoire prenant un nombre fini de valeurs $x_1 < \cdots < x_s$.

Définition 9. *On appelle espérance mathématique de X (ou parfois valeur moyenne de X), le nombre réel, noté $E(X)$ défini par*

$$E(X) = \sum_{i=1}^{s} x_i p_i$$

où $p_i = P(X - x_i)$.

Cette définition d'espérance se généralise en considérant les moments d'ordre k :

Définition 10. *On appelle moment d'ordre k ($k \in \mathbb{N},\ k \neq 0$), de X , le nombre réel, noté m_k, défini par*

$$E(X) = \sum_{i=1}^{s} x_i^k p_i$$

où $p_i = P(X - x_i)$.

On a donc

$$E(X) = m_1(X).$$

L'espérance s'interprète donc comme l'abscisse du centre de gravité des masses $p_1, \cdots, p_s$ et le moment d'ordre 2 comme un moment d'inertie de l'ensemble de ces masses.

Théorème 2. *On a*

(1) $m_k(X) = E(X^k)$,

(2) Si X et Y sont deux variables aléatoires sur Ω de même loi, alors

$$E(X+Y) = E(X) + E(Y),$$

(3) si a, b sont deux réels, alors

$$E(aX + b) = aE(X) + b,$$

(4) Si X et Y sont de plus **indépendantes***, alors*

$$E(XY) = E(X)E(Y).$$

Démonstration.

(1) La première propriété découle directement de la définition du moment d'ordre k.

(2) Les variables X et Y ont la même loi de probabilité. Ceci signifie que X prend les valeurs $x_1 < \cdots < x_s$ et Y prend les valeurs $y_1 < \cdots < y_s$ avec

$$P(X = x_i) = P(Y = y_i) = p_i.$$

Ainsi

$$E(X+Y) = \sum_{i=1}^{s}(x_i + y_i)p_i = \sum_{i=1}^{s} x_i p_i + \sum_{i=1}^{s} y_i p_i = E(X) + E(Y).$$

(3) On suppose que le scalaire b est la variable aléatoire qui fait correspondre à tout élément de ω la valeur b. Alors

$$E(aX + b) = \sum_{i=1}^{s}(ax_i + b)p_i = a\sum_{i=1}^{s} x_i p_i + b\sum_{i=1}^{s} p_i = aE(X) + b.$$

(4) Si maintenant les variables X et Y sont de plus supposées indépendantes, alors les valeurs de XY sont les scalaires $x_i y_j$ avec $1 \leq i, j \leq s$. Comme

$$P(XY = x_i y_j) = p_i p_j,$$

alors

$$E(XY) = \sum_{i,j=1}^{s} x_i y_y p_i p_j = \sum_{i=1}^{s} x_i p_i \sum_{j=1}^{s} y_y p_j = E(X)E(Y).$$

Exemples.

(1) L'espérance d'une variable de Dirac de paramètre a est
$$E(X) = a.$$
(2) L'espérance d'une variable de Bernoulli de papmètre p est
$$E(X) = p.$$

4.2. **Variance, écart-type.** Soit X une variable aléatoire et $E(X)$ son espérance. Considérons la nouvelle variable
$$Z = X - E(X)$$
où $E(X)$ est considérée comme la variable aléatoire
$$E(X)(\omega) = E(X)$$
pour tout $\omega \in \Omega$. Cette variable Z est appelée l'écart de X.

Définition 11. *On a*

(1) On appelle variance de X l'espérance de la variable écart $Z^2 = (X - E(X))^2$. On la note $V(X)$ et donc
$$V(X) = E((X - E(X))^2).$$
(2) On appelle écart-type de X le scalaire $\sigma(X)$ défini par
$$\sigma(X) = \sqrt{V(X)}.$$

On montre facilement

Théorème 3. *On a*
$$V(X) = E((X - m_1)^2) = e(X^2) - (E(X))^2$$
et donc
$$\sigma^2 = m_2 - m_1^2.$$

Exemples.

(1) La variance et l'écart-type d'une variable de Dirac de paramètre a est nul :
$$V(X) = 0.$$
(2) La variance et l'écart-type d'une variable de Bernoulli de papmètre p est
$$V(X) = p(1-p), \quad \sigma(X) = \sqrt{p(1-p)}.$$
(3) Une variable aléatoire suit la loi binomiale $B(p, q, n)$ si Les valeurs de X sont $\{0, 1, \cdots, n\}$ et si la probabilité pour que $X = k$ est
$$P(X = k) = C_n^k p^k q^{n-k}.$$
Dans ce cas
$$E(X) = np, \; V(X) = npq, \; \sigma(X) = \sqrt{npq}.$$

5. Espérance, Variance, Ecart type d'une variable aléatoire continue

5.1. **Espérance.** Nous n'abordons pas ici les problèmes d'existence des intégrales généralisées qui suivent.

Définition 12. *Soit X une variable aléatoire continue de densité $f_X(x)$. L'espérance de X est*
$$E(X) = \int_{-\infty}^{+\infty} t f_X(t)dt.$$

On dit parfois que l'espérance de X est son moment d'ordre 1 et noté $m_1(X)$. Ceci permet de définir le moment d'ordre k :
$$m_k(X) = \int_{-\infty}^{+\infty} t^k f_X(t)dt.$$
La variable aléatoire
$$Z = X - E(X) = X - m_1(X)$$
est appelée l'écart de X.

5.2. **Variance et Ecart-type.** Ces deux invariants de X sont, lorsqu'ils sont définis, donnés par :

Définition 13. *Soit X une variable aléatoire continue de densité $f_X(x)$. La variance de X est*
$$V(X) = \int_{-\infty}^{+\infty} (t - E(X))^2 f_X(t)dt = E(Z^2).$$
L'écart-type est défini par
$$\sigma(X) = \sqrt{V(X)}.$$

On retrouve dans ce cadre-ci les propriétés établies dans le cadre fini :

Proposition 7. *Soient a, b des constantes.*
1. $E(aX + b) = aE(X) + b$,
2. $\sigma^2(aX + b) = a^2\sigma^2(X)$,
3. $E(X + Y) = E(X) + E(Y)$.
4. $\sigma^2(X) = m_2(X) - m_1(X)$.

Si les variables X et Y sont indépendantes
1. $E(XY) = E(X)E(Y)$,
2. $\sigma^2(X + Y) = \sigma^2(X) + \sigma^2(Y)$.

6. La loi normale

6.1. **La loi normale $\mathcal{N}(\mu, \sigma)$.** La loi normale est une des lois de probabilité qui modélise de nombreux phénomènes naturels. Par exemple, le caractère taille de la population est distribué selon une loi normale. Elle est également utilisée pour donner une approximation fine pour des lois de probabilité de variables aléatoires finies, mais lorsque les paramètres sont des grands nombres. Par exemple, dans les lois de binomiales ou de Bernoulli, il est souvent très difficile de calculer numériquement les coefficients binomiaux lorsque les paramètres sont des grands nombres.

Commençons par définir un cas particulier, celui de la loi normale centrée réduite.

Définition 14. *On appelle loi normale centrée réduite, notée $\mathcal{N}(0,1)$ la loi de probabilité d'une variable aléatoire continue dont la fonction de densité est*

$$\varphi_{0,1}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

On peut vérifier, à titre d'exercice, que la fonction $\varphi_{0,1}$ est bien une fonction de densité. Sinon, on l'admet. Sur une calculette graphique, on peut représenter cette fonction, on obtient la célèbre courbe de Gauss, appelée aussi courbe en cloche.

Plus généralement on définit laloi normale de paramètres μ et σ :

Définition 15. *On appelle loi normale $\mathcal{N}(\mu, \sigma)$ de paramètres $\mu, \mu \in \mathbb{R}$ et $\sigma, \sigma \in \mathbb{R}$ et $\sigma > 0$, la loi de probabilité dont la fonction de densité est*

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

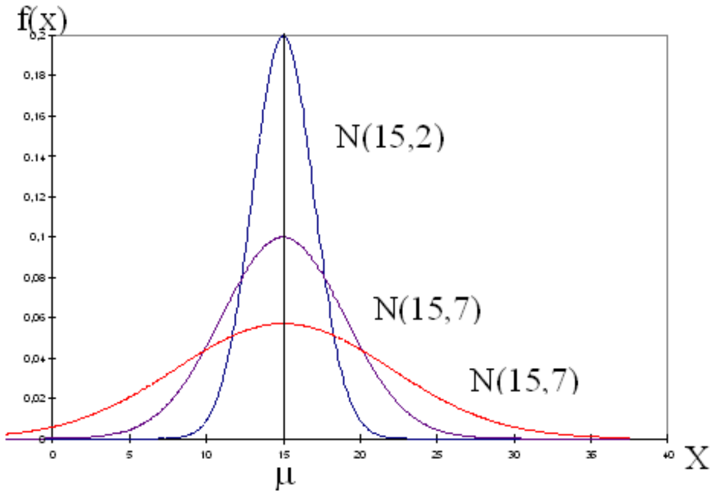
Cette loi est aussi appelée loi de Gauss ou de Laplace-Gauss.

Rappelons que si $f_X(x)$ est la densité de probabilité d'une variable aléatoire continue, alors elle vérifie

(1) $f_X(x) \geq 0$

(2) $\displaystyle\int_{-\infty}^{+\infty} f_X(t)dt = 1.$

On vérifie sans peine, comme l'exponentielle est toujours positive, que la fonction $\varphi_{\mu,\sigma}(x)$ est positive. Nous pouvons tracer son graphe, en utilisant par exemple un logiciel graphique :



Par contre, on admettra, car le calcul nécessite des outils mathématiques sophistiqués :

$$\int_{-\infty}^{+\infty} \varphi_{\mu,\sigma}(x)dx = 1.$$

Proposition 8. *Soit X une variable aléatoire normale, c'est-à-dire dont la fonction de densité suit la loi normale $\mathcal{N}(\mu,\sigma)$. Alors*

(1) Son espérance, ou valeur moyenne, est $E(X) = \mu$,

(2) Son écart-type est σ.

6.2. **Passage de la loi normale $\mathcal{N}(\mu,\sigma)$ à la loi normale centrée réduite $\mathcal{N}(0,1)$.** Rappelons que la loi de densité de la loi normale centrée réduite est

$$\varphi_{0,1}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

et la fonction de répartition

$$\mathfrak{F}_{0,1}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$

Si la variable aléatoire a pour loi de probabilité $\mathcal{N}(0,1)$, alors

$$P(x_1 \leq X \leq x_2) = \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{t^2}{2}} dt.$$

Posons

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-\frac{t^2}{2}} dt.$$

Alors

$$P(x_1 \leq X \leq x_2) = \Phi(x_2) - \Phi(x_1).$$

La fonction Φ est impaire et vérifie

$$\Phi(-\infty) = -\frac{1}{2}, \quad \Phi(+\infty) = \frac{1}{2}.$$

Les valeurs de Φ sont données dans la table en fin de chapitre.

Supposons maintenant que X soit une variable aléatoire ayant pour loi $\mathcal{N}(\mu,\sigma)$. Associons lui la variable aléatoire

$$Y = \frac{X-\mu}{\sigma}.$$

Il est clair que Y suit encore une loi normale. Son espérance est $E(Y) = 0$ et son écart-type est $\sigma(Y) = \dfrac{\sigma(X)}{\sigma} = 1$. Ainsi Y suit la loi normale centrée.

Proposition 9. *Soit X une variable aléatoire normale, c'est-à-dire dont la fonction de densité suit la loi normale $\mathcal{N}(\mu,\sigma)$. Alors la variable aléatoire*

$$Y = \frac{X-\mu}{\sigma}$$

suit la loi $\mathcal{N}(0,1)$ centrée réduite.

Ce résultat permet de calculer numériquement la probabilité $P(a \leq X \leq b)$ pour une variable suivant une loi normale $\mathcal{N}(\mu,\sigma)$ donnée en transférant le calcul à la variable normale centrée associée et en utilisant la table des valeurs de Φ ou une calculette numérique.

Exemple. Supposons que X suive la loi normale $\mathcal{N}(1,2)$. Alors la variable

$$Y = \frac{X-1}{2}$$

suit la loi normale $\mathcal{N}(0,1)$. Calculons $P(0 \leq X \leq 2)$. On a $X = 2Y+1$. Alors $0 \leq X \leq 2$ implique $0 \leq 2Y+1 \leq 2$ soit

$$-\frac{1}{2} \leq Y \leq \frac{1}{2}.$$

Ainsi

$$P(-\frac{1}{2} \leq Y \leq \frac{1}{2}) = \Phi(\frac{1}{2}) - \Phi(-\frac{1}{2}).$$

Comme ϕ est une fonction impaire $\Phi(\frac{1}{2}) - \Phi(-\frac{1}{2}) = 2\Phi(\frac{1}{2}) = 2 \times 0,19 = 0,38$. Ainsi

$$P(0 \leq X \leq 2) = 0,38.$$

6.3. **Approximation de la loi binomiale par la loi normale.** Rappelons, dans un premier temps, la formule de Stirling, qui est la base de notre théorie d'approximation.

Formule de Stirling. Lorsque n devient grand, alors

$$n! \sim n^n e^{-n}\sqrt{2\pi n}.$$

De cette formule se déduit l'approximation suivante : si $npq \gg 1$, alors

$$\binom{k}{n} p^k q^{n-k} \simeq \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k-np)^2}{2npq}}.$$

La notation $npq \gg 1$ signifie que le nombre npq est grand par rapport à 1, sans trop préciser ce que signifie grand !

Proposition 10. *Soit X une variable aléatoire discrète suivant la loi de Bernoulli. Alors $P(X = k) = \binom{k}{n} p^k q^{n-k}$ et si $npq \gg 1$ alors*

$$P(X = k) \simeq \frac{1}{\sqrt{npq}} \varphi_{0,1}(\frac{k - np}{\sqrt{npq}}).$$

Exemple. On lance 1000 fois une pièce de monnaie. Quelle est la probabilité de faire 500 pile ? Cette probabilité est

$$P(X = 500) = \binom{500}{1000} (0,5)^{500} (0,5)^{500}.$$

D'après le théorème d'approximation ci-dessus, comme $npq = 1000 \times 0,5 \times 0,5 = 250 \gg 1$,

$$P(X = 500) \simeq \frac{1}{\sqrt{250}} \varphi_{0,1}(\frac{500 - 500}{\sqrt{250}}) = \frac{1}{\sqrt{250}} \frac{1}{\sqrt{2\pi}} \simeq 0,0252.$$

7. Fonction caractéristique d'une variable aléatoire continue

Dans le chapitre précédent, nous avons défini la notion de fonction caractéristique d'une variable aléatoire finie ou discrète.

Définition 16. *Soit X une variable aléatoire continue de fonction de densité $f_X(x)$. Sa fonction génératrice est la fonction*

$$\varphi_X(u) = \int_{-\infty}^{+\infty} e^{iux} f_X(x) dx.$$

Ce n'est rien d'autre que la transformée de Fourier de la fonction densité. Elle définit entièrement la loi de probabilité de la variable X. Ceci signifie que deux variables aléatoires ayant même fonction caractéristique ont des fonctions de répartition égales ou égales presque partout. Si m_k sont les moments d'ordre k de X, on a alors,

$$\varphi_X(u) = 1 + ium_1 + \cdots + \frac{(iu)^k}{k!} m_k + \cdots$$

Proposition 11. *Soit X une variable aléatoire suivant la loi normale $\mathcal{N}(\mu, \sigma)$. Sa fonction génératrice est la fonction*

$$\varphi_X(u) = e^{i\mu u - \frac{u^2\sigma^2}{2}}.$$

8. L'inégalité de Bienaymé-Tchebychev

8.1. **L'inégalité.** Etant donnée une variable aléatoire X, son écart-type mesure la tendance qu'à X à s'écarter de sa moyenne. L'inégalité qui suit mesure l'erreur que l'on peut commettre en remplaçant X par sa moyenne.

Théorème 4. *Soit X une variable aléatoire sur l'espace probabilité $(\Omega, \mathcal{F}, P)$ telle que $E(X^2)$ soit finie. Si on note par μ son espérance et par σ son écart-type, alors pour tout $\varepsilon \geq 0$,*

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\varepsilon^2}.$$

Par exemple, la probabilité pour que le résultat d'une expérience $X(\omega)$ diffère de μ de plus de 8σ est

$$P(|X - \mu| \geq 8\sigma) \leq \frac{\sigma^2}{(8\sigma^2} = \frac{1}{64}.$$

8.2. **Calcul de π et l'aiguille de Buffon.** On lance une aiguille de longueur a sur un parquet dont les lattes sont de largeur b, $b \geq a$. Buffon calcula la probabilité pour que l'aiguille coupe la rainure entre deux lattes et trouva

$$p = \frac{2a}{\pi b}.$$

Lorsque $a = b$, on trouve

$$p = \frac{2}{\pi}$$

d'où

$$\pi = 2p.$$

On peut espérer calculer, par une bonne estimation de p, une valeur approchée de π. Soit k le nombre de lancers. La loi de probabilité de la variable aléatoire X correspondant à cette expérience est la loi de Bernoulli $\mathcal{B}(k, p)$. Son espérance est donc

$$E(X) = kp$$

et l'écart-type

$$\sigma(X) = \sqrt{kpq}.$$

L'inégalité de Bienaymé-Tchebychev s'écrit

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

ou bien

$$P(|X-\mu| \geq \epsilon\sigma) \leq \frac{1}{\varepsilon^2}$$

soit

$$P(|X-kp| \geq \epsilon\sigma) \leq \frac{1}{\varepsilon^2}.$$

Supposons que $\frac{1}{\varepsilon^2} = \frac{1}{100}$ soit $\varepsilon = 10$. Soit x le nombre d'aiguilles ayant coupée la rainure. Le calcul approché $\widetilde{\pi}$ de π est alors $\widetilde{\pi} = 2k/x$. Supposons que $\widetilde{\pi} = \pi \pm 10^{-3}$. Alors

$$\frac{2k}{\pi + 10^{-3}} \leq x \leq \frac{2k}{\pi - 10^{-3}}.$$

Si $|X-kp| = \varepsilon\sigma$, alors

$$k = \frac{\pi^2}{\varepsilon} 2.10^{-3}\sqrt{kpq}$$

et donc

$$k = \frac{\pi^4}{\varepsilon}^2 4.10^{-6}pq$$

ce qui donne à peu près $k = 560000000$.

EXERCICES

Exercice 1. Une urne contient quatre boules numérotées de 1 à 4. On en tire deux sans remise.

(1) On note par X le plus grand des deux nombres obtenus. Quelle est la loi de probabilité de X ?
(2) On note par Y le plus petit des deux nombres obtenus. Quelle est la loi de probabilité de Y ?

Exercice 2. On lance un dé deux fois successivement. On s'intéresse à la somme des chiffres obtenue après ces deux lancers.

(1) Décrire l'espace probabilisé $(\Omega, \mathfrak{F}, P)$.
(2) Définir la variable aléatoire X associée à cette expérience.
(3) Calculer $X^{-1}(]-\infty, 8]$ et $X^{-1}([3,1,9,2]$.
(4) Déterminer la loi de la variable aléatoire X.

Exercice 3. Une urne contient p boules numérotées de 1 à p. On tire n boules simultanément. Soit X la variable aléatoire correspondant au plus grand numéro tiré et Y au plus petit.

(1) Déterminer les lois de X et Y.
(2) Calculer l'espérance de X

Exercice 4. Calculer l'espérance et la variance d'une variable aléatoire X qui suit la loi uniforme sur un intervalle $[[p, q]]$.

Exercice 5. On considère le lancer d'un dé.

(1) Décrire l'espace probabilisé $(\Omega, \mathfrak{F}, P)$.
(2) On considère l'application $X : \Omega \to \mathbb{R}$ définie par $X(\omega) = 1$ si ω est impair, sinon $X(\omega) = 0$. Montrer que X est une variable aléatoire.
(3) Quelle est sa fonction de répartition.

Exercice 6. On considère un dé cubique truqué dont les faces sont numérotées de 1 à 6. La probabilité d'obtenir une face est proportionnelle au numéro inscrit sur cette face. Soit X la variable aléatoire donnée par le numéro de la face du dessus.

(1) Déterminer la loi de X, calculer son espérance.
(2) On pose $Y = 1/X$. Déterminer la loi de Y, et son espérance.

Exercice 7. Une entreprise souhaite recrute un cadre. n personnes se présentent pour le poste. Chacun d'entre eux passe à tour de rôle un test, et le premier qui réussit le test est engagé. On note p la probabilité de réussir le test. On définit la variable aléatoire X par $X = k$ si le k-ième candidat qui réussit le test est engagé, et $X = n + 1$ si personne n'est engagé.

(1) Déterminer la loi de X.
(2) Calculet $\sum_{k=0}^{n} x^k$ et en déduire $\sum_{k=1}^{n} kx^{k-1}$ pour $x \neq 1$.
(3) En déduire l'espérance de X.
(4) Quelle est la valeur minimale de p pour avoir plus d'une chance sur deux de recruter l'un des candidats ?

Exercice 8. Une grenouille monte les marches d'un escalier (supposé infini) en partant du sol et en sautant ou bien une seule marche, avec probabilité p ou bien deux marches, avec la probabilité $1-p$. On suppose que les sauts sont indépendants les uns des autres.

(1) Dans cette question, on observe n sauts de la grenouille, et on note X_n le nombre de fois où la grenouille a sauté une marche, et Y_n le nombre de marches franchies. Quelle est la loi de X_n ?

(2) Exprimer Y_n en fonction de X_n. En déduire l'espérance et la variance de Y_n.

(3) Pour $k \geq 1$, on note p_k la probabilité que la grenouille passe par la marche k. Que vaut p_1 ? Que vaut p_2 ? Établir une formule de récurrence liant p_k et p_{k-1}. En déduire la valeur de p_k pour $k \geq 1$.

(4) On note désormais Z_n le nombre de sauts nécessaires pour atteindre ou dépasser la n-ième marche. Écrire un algorithme qui simule la variable aléatoire Z_n.

Exercice 9. Dans chaque cas indiquer si f est une fonction de densité :

(1) $$\begin{cases} f(x) = 0, \text{ si } -\infty \leq x \leq 0, \text{ et } 2 \leq x \leq +\infty, \\ f(x) = x \text{ si } 0 \leq x \leq 1, \\ f(x) = 2 - x \text{ si } 1 \leq x \leq 2, \end{cases}$$

(2) $f(x) = \frac{3}{4}(1 - x^2), \quad -1 \leq x \leq 1,$

(3) $f(x) = e^{-x}$ pour $x \geq 0$.