CHAPITRE III

VARIABLES ALEATOIRES

1. Introduction

Dans la plupart des **phénomènes aléatoires**, le résultat d'une **épreuve** peut se traduire par une « grandeur » mathématique, très souvent représentée par un nombre entier ou un nombre réel. La notion mathématique qui représente efficacement ce genre de situation concrète est celle de **variable aléatoire** (notée également v.a.). Ainsi le temps de désintégration d'un atome radioactif, le pourcentage de réponses « oui » à une question posée dans un sondage ou le nombre d'enfants d'un couple sont des exemples de variables aléatoires.

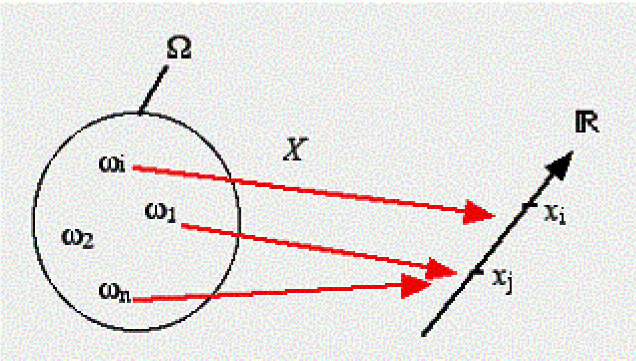
Remarque :

On se limitera ici au cas des variables aléatoires **réelles** (les entiers faisant bien sûr partie des réels).

Etant donné un **espace probabilisé** d'espace fondamental Ω et de probabilité P,

On appelle **variable aléatoire** sur cet espace, toute **application** X de Ω dans $\mathbb{R}$ telle que :

$$X : \mathrm{P}(\Omega) \to \mathbb{R}$$
$$\omega \to X(\omega)$$



A chaque évènement élémentaire ω de Ω correspond un nombre réel x associé à la variable aléatoire X. Comme l'indique le graphe, il n'y a pas obligatoirement autant de valeurs possibles prises par la variable aléatoire X que d'évènements élémentaires. La valeur x correspond à la **réalisation** de la variable X pour l'évènement élémentaire ω.

Exemple :

Si l'on considère la constitution d'une fratrie de deux enfants, l'espace fondamental est constitué des évènements élémentaires suivant : $\Omega = \{GG, GF, FG, FF\}$

Les valeurs possibles prises par la variable aléatoire X, « nombres de fille dans la famille » sont : $X = \{0,1,2\}$

Variables aléatoires discrètes

2.1 Définition

Une variable aléatoire est dite **discrète** si elle ne prend que des **valeurs discontinues** dans un intervalle donné (borné ou non borné). L'ensemble des nombres entiers est discret. En règle générale, toutes les variables qui résultent d'un **dénombrement** ou d'une **numération** sont de types discrètes.

Exemples :

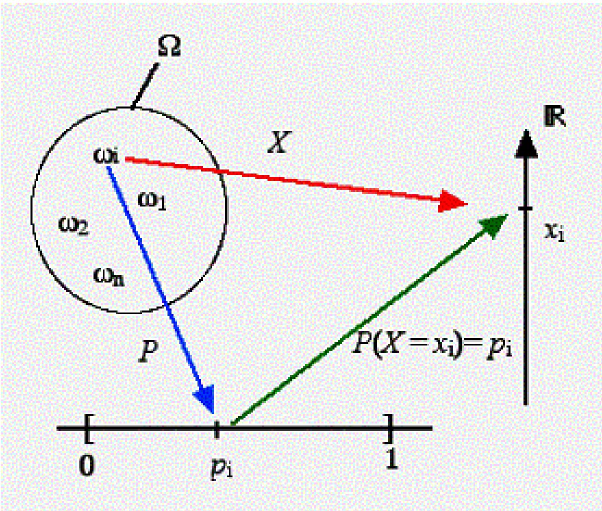
Les variables aléatoires,

- le nombre de petits par porté pour une espèce animale donnée (chat, marmotte, etc.),
- le nombre de bactéries dans 100 ml de préparation,
- le nombre de mutations dans une séquence d'ADN de 10 kb,

Sont des **variables aléatoires discrètes**.

2.2 Loi de probabilité

Une variable aléatoire est caractérisée par l'ensemble des valeurs qu'elle peut prendre et par l'expression mathématique de la probabilité de ces valeurs. Cette expression s'appelle **la loi de probabilité** (ou **distribution de probabilité**) de la variable aléatoire.



La loi de probabilité d'une **variable aléatoire discrète** est entièrement déterminée par les probabilités p_i des évènements $\{X = x_i\}$, x_i parcourant l'univers image $X(\Omega)$. La **loi de probabilité** est donnée par les $(x_i, p_i)_i$

Remarque :

Afin de simplifier l'écriture, nous noterons pour la suite du cours : $P(\{X = x_i\})$ équivalent à $P(X = x_i)$ ou p_i

Exemple :

Dans le cas de la constitution d'une fratrie de deux enfants, si l'on fait l'hypothèse que la probabilité d'avoir un garçon est égale à celle d'avoir une fille ($\frac{1}{2}$), alors la distribution de probabilité ou **loi de probabilité du nombre de filles** dans une fratrie de deux enfants est :

Ensemble des évènements possibles ω	Valeurs de la variable aléatoire X	Probabilités associées à la variable X $P(X = x_i)$ ou p_i
G et G	**0**	**1/4**
F et G ou G et F	**1**	**1/2**
F et F	**2**	**1/4**

Si $P(F) = P(G) = \frac{1}{2}$, alors

1. $P((F \cap G) \cup (G \cap F)) = P(F \cap G) + P(G \cap F)$ Propriétés d'additivité

Avec $(F \cap G) \cap (G \cap F) = \Phi$ évènements incompatibles

2. $P(F \cap G) = P(F)P(G)$ Propriété d'indépendance

D'où $P(F \cap G) \cup (G \cap F) = P(X = 1) = \left(\frac{1}{2} \times \frac{1}{2}\right) + \left(\frac{1}{2} \times \frac{1}{2}\right) = \frac{1}{2}$

Remarque :

Une loi de probabilité n'est établie que si $\sum_{i=1}^{i=n} p_i = 1$, la somme étant étendue à tous les indices i

2.3 Fonction de répartition

On appelle **fonction de répartition** d'une variable aléatoire X, la fonction F_X telle que :

$$\begin{aligned} F_X &: \mathbb{R} \to \mathbb{R} \\ t &\to F_X(x) = P(X < x) \end{aligned}$$

Concrètement la fonction de répartition correspond à la **distribution des probabilités cumulées**. Le plateau atteint par la fonction de répartition correspond à la valeur de probabilité 1 car $\sum_{i=1}^{i=n} p_i = 1$.

L'importance pratique de la fonction de répartition est qu'elle permet de calculer la probabilité de tout intervalle dans $\mathbb{R}$.

Propriétés : les propriétés associées à la fonction de répartition sont les suivantes :

Soit F_X la fonction de répartition d'une **variable aléatoire discrète** X alors :

(P1) $\forall t \in \mathbb{R} : 0 \leq F_x(x) \leq 1$

(P2) F_X est croissante sur $\mathbb{R}$

(P3) $\lim_{x \to -\infty} F_X(x) = 0$ et $\lim_{x \to +\infty} F_X(x) = 1$

(P4) si $a \leq b$: $P(a \leq X \leq b) = F_X(b) - F_X(a)$

Exemple :

On considère l'évènement ω « lancer de 3 pièces ». On introduit une variable aléatoire X définie par $X(\omega)$ « nombre de piles de l'évènement ω ». La loi de probabilité de X est :

Nombre de piles	0	1	2	3
$P(X = x_i)$	$1/8$	$3/8$	$3/8$	$1/8$
F_X	$1/8$	$4/8$	$7/8$	1

Dans le cas d'une variable aléatoire discrète, on utilise un **diagramme en bâtons** pour visualiser la **distribution de probabilités** et une fonction en escalier pour **la fonction de répartition.**

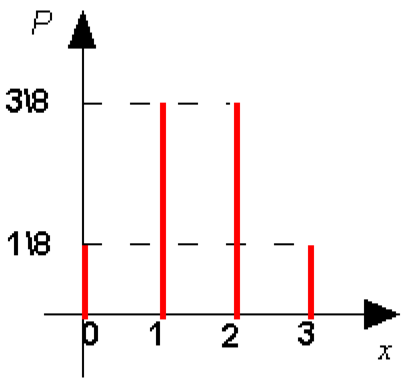
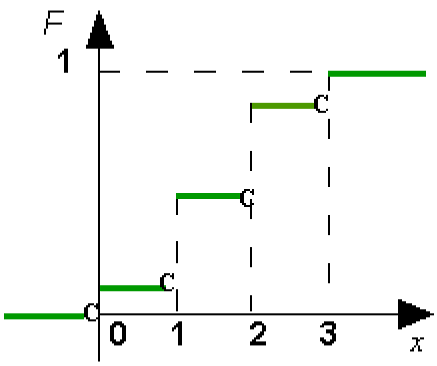


Diagramme en bâtons



Fonction de répartition

3. Variables aléatoires continues

3.1 Définition

Une variable aléatoire est dite **continue** si elle peut prendre **toutes les valeurs** dans un intervalle donné (borné ou non borné). En règle générale, toutes les variables qui résultent d'une **mesure** sont de type continu.

Exemples : Les variables aléatoires,

- le masse corporelle des individus pour une espèce animale donnée,

- taux de glucose dans le sang,

Sont des **variables aléatoires continues**.

3.2 Fonction densité de probabilité

Dans le cas d'une **variable aléatoire continue**, la loi de probabilité associe une probabilité à chaque ensemble de valeurs définies **dans un intervalle donné**. En effet, pour une variable aléatoire continue, la probabilité associée à l'évènement $\{X = a\}$ est nulle, car il est impossible d'observer exactement cette valeur.

On considère alors la probabilité que la variable aléatoire X prenne des valeurs comprises dans un intervalle $[a,b]$ tel que $P(a \le X \le b)$

Lorsque cet intervalle tend vers 0, la valeur prise par X tend alors vers une fonction que l'on appelle **fonction densité de probabilité** ou **densité de probabilité**.

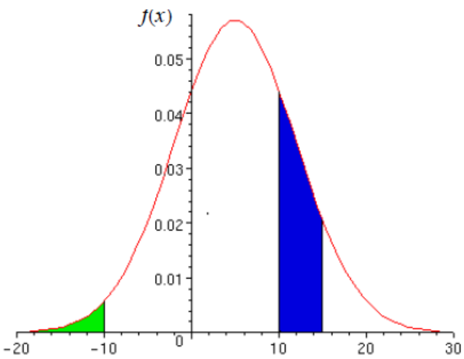
On appelle **densité de probabilité** toute application continue: $\begin{array}{l} f:\mathbb{R} \to \mathbb{R} \\ x \to f(x) \end{array}$

Telle que : (P1) $x \in \mathbb{R} : f(x) \ge 0$. (P2) $\int_{-\infty}^{+\infty} f(x)dx = 1$

Soit une f**onction densité de probabilité** $f(x)$:

(1) l'aire hachurée **en vert** correspond à la probabilité $P(X < -10)$

(2) l'aire hachurée **en bleu** correspond à la probabilité $P(+10 < X < +15)$



Remarque :

Cette fonction densité de probabilité est une loi de probabilité car l'aire sous la courbe est égale à 1 pour toutes les valeurs de x définies

Réciproquement :

Une variable aléatoire X définie sur un univers Ω est dite absolument continue, s'il existe une **fonction densité de probabilité** f telle que : $\forall t \in \mathbb{R} : P(X < x) = \int_{-\infty}^{x} f(t)dt$ (Voir graphe ci-dessus).

3.3 Fonction de répartition

Si comme pour les variables aléatoires discrètes, on définit la **fonction de répartition** de X par :

$$\boxed{\begin{array}{l} F_X : \mathbb{R} \to \mathbb{R} \\ \quad t \ \to F_X(x) = P(X < x) \end{array}}$$

Alors la relation entre la **fonction de répartition** F_X et la fonction **densité de probabilité** $f(x)$ est la suivante :

$$\boxed{\forall t \in \mathbb{R} : F_X(x) = P(X < x) = \int_{-\infty}^{x} f(t)dt}$$

La fonction de répartition F_X est **la primitive** de la fonction densité de probabilité f , et permet d'obtenir les probabilités associées à la variable aléatoire X , en effet :

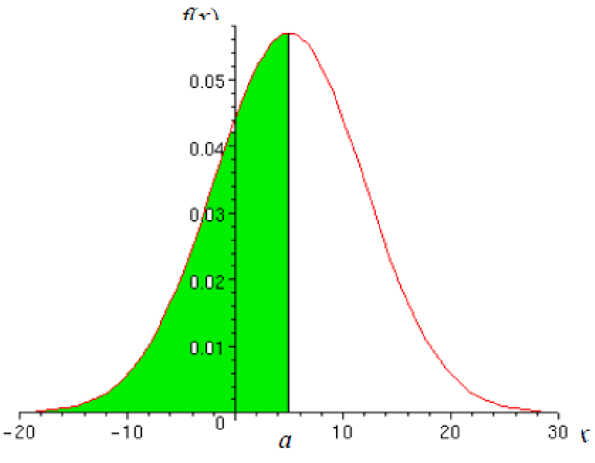
Soit X une **variable aléatoire absolument continue** de densité f et de fonction de répartition F_X , alors :

(P$_1$) $P(a \le X \le b) = F_X(b) - F_X(a) = \int_a^b f(x)dx, (a < b)$

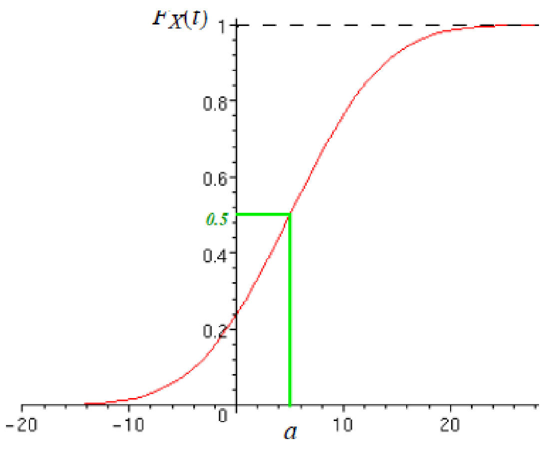
(P$_2$) . $\forall a \in \mathbb{R} : P(X = a) = 0$ si f est continue à droite du point a.

Remarque : La propriété P$_2$ implique que $P(X \le x) = P(X < x)$.

La **fonction de répartition** correspond aux ***probabilités cumulées*** associées à la variable aléatoire continue sur l'intervalle d'étude (graphe ci-dessous).



Fonction densité de probabilité $f(x)$



Fonction de répartition F_X

L'aire **hachurée en vert** sous la courbe de la fonction densité de probabilité correspond à la probabilité $P(X < a)$ et vaut **0,5** car ceci correspond exactement à la moitié de l'aire totale sous la courbe. Cette probabilité correspond à la valeur de la fonction de répartition au **point** d'**inflexion de la courbe.**

Propriété : Les propriétés associées à la fonction de répartition sont les suivantes :

Soit F_X la fonction de répartition d 'une variable aléatoire **absolument continue** X alors :

(P$_1$) F_X est continue sur $\mathbb{R}$, dérivable en tout point où f est continue et alors $F_X{}' = f$

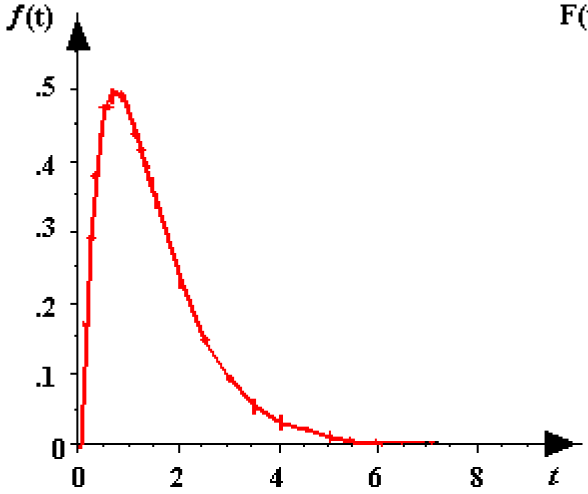
(P$_2$) F_X est croissante sur $\mathbb{R}$

(P$_3$) F_X est à valeurs dans $[0,1]$

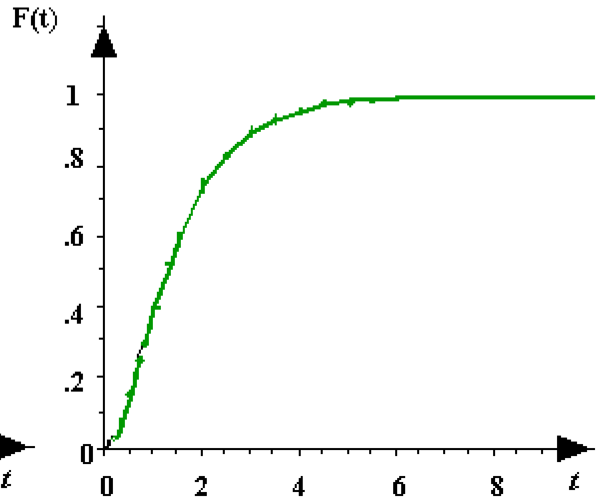
(P$_4$) $\lim_{x \to -\infty} F_X(x) = 0$ et $\lim_{x \to +\infty} F_X(x) = 1$

Exemple : Dans une population de **canards colverts**, lors d'une leur lieu de repos. Ainsi à $t = 0$, la surface de l'étang regagne l'étang entre les temps t_1 et t_2 (en minutes) alerte, l'ensemble des individus quittent est déserte et la probabilité qu'un canard est donnée par : $\int_{t_1}^{t_2} f(t)dt$ avec $f(t) = 2e^{-t} - 2e^{-2t}$ qui représente la fonction **densité de probabilité**.

La primitive de $f(t)$, $F_T(t)$, **fonction de répartition** est de la forme : $F_T(t) = 1 + e^{-2t} - 2e^{-t}$



Fonction de densité de probabilité



Fonction de répartition

L'évolution de la recolonisation de l'étang par les canards colverts en fonction du temps est donnée par la **courbe rouge**. On observe ainsi que plus de 50 % des canards se posent sur l'étang au cours des 2 premières minutes qui suivent l'alerte. Au bout de 7 minutes, tous les canards ont regagné l'étang. La distribution des probabilités cumulées est donnée sur la **courbe verte.**

4. Espérance et Variance

Une loi de probabilité peut être caractérisée par certaines valeurs typiques correspondant aux notions de valeur centrale, de dispersion et de forme de distribution.

4.1 Espérance mathématique

L'espérance d'une variable aléatoire $E(X)$ correspond à la moyenne des valeurs possibles de X pondérées par les probabilités associées à ces valeurs. C'est un paramètre de position qui correspond au moment d'ordre 1 de la variable aléatoire X. C'est l'équivalent de la moyenne arithmétique $\bar{X}$. En effet lorsque le nombre d'épreuves n est grand, $\bar{X}$ tend vers $E(X)$.

1. Variables aléatoires discrètes

Si X est une **variable aléatoire discrète** définie sur un univers probabilisé Ω, on appelle espérance de X, le réel défini par : $E(X)=\sum_{\omega\in\Omega} X(\omega)P(\omega)$

Remarque :

Si $X(\Omega)$ est infini, on n'est pas sûr que l'espérance existe. L'espérance mathématique est également notée $\mu(X)$, μ_X ou encore μ si aucune confusion n'est à craindre.

Nous pouvons donner une autre définition de l'espérance d'une **variable aléatoire discrète** X si à $\omega\in\Omega$, on associe l'image x telle que $X(\omega)=x$.

Théorème :

Si X est une **variable aléatoire discrète** de loi de probabilité $(x_i, p_i)_i$ définit sur un nombre fini n d'évènements élémentaires alors : $E(X)=\sum_{i=1}^{i=n} x_i p_i$

Exemples :

(1) Si l'on reprend l'exemple d'une fratrie de deux enfants, l'espérance du variable aléatoire « nombre de filles « est : $E(X)=\sum_{i=1}^{i=3} x_i p_i = 0\times\frac{1}{4}+1\times\frac{1}{2}+2\times\frac{1}{4}=1$

Si l'on observe un nombre suffisant de fratries de 2 enfants, on attend **en moyenne une fille** par fratrie.

(2) Quand est-il de l'espérance de la variable aléatoire X de valeurs 0, 1, 2 et 3 avec respectivement les probabilités 0,1 ; 0,2 ; 0, 3 et 0,4 ?

Réponse :

2. Variables aléatoires continues

Si X est une **variable aléatoire absolument continue** de densité f, on appelle **espérance** de X, le réel $E(X)$, défini par : $E(X)=\int_{-\infty}^{+\infty} x f(x)dx$ si cette intégrale est convergente.

Remarque :

Dans cet exemple, la variable étudiée t ne peut prendre que des valeurs dans $[0,+\infty[$

4.1.3 Propriétés de l'espérance

Les propriétés de **l'espérance** valent aussi bien pour une variable aléatoire discrète ou une variable aléatoire absolument continue.

Si X et Y sont deux variables aléatoires définies sur un même univers Ω, admettant une espérance, alors :

(P_1) $E(X+Y)=E(X)+E(Y)$

(P_2) $E(aX)=aE(X) \quad / a\in\mathbb{R}$

(P_3) Si $X\geq 0$ alors $E(X)\geq 0$

(P_4) Si $E(X+Y)=E(X)+E(Y)$ est un caractère constant tel que : $\forall\omega\in\Omega : X(\omega)=k ; Alors : E(X)=k$

Remarque :

Dans le cas continu $E(X+Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)f(x,y)dxdy$. La propriété P_1 est vérifiée quelques soient les relations de dépendance ou d'indépendance statistique entre les deux variables.

4.2 Variance

La variance d'une variable aléatoire $V(X)$ est l'espérance mathématique du carré de l'écart à l'espérance mathématique. C'est un paramètre de dispersion qui correspond au moment centré d'ordre 2 de la variable aléatoire X. C'est l'équivalent de la variance observée S^2. En effet lorsque le nombre d'épreuves n est grand, S^2 tend vers $V(X)$

Si X est une variable aléatoire ayant une espérance $E(X)$, on appelle **variance** du X le réel :

$$V(X)=E\left(\left[X-E(X)\right]^2\right)$$

$$V(X) = E\left(\left[X - E(X)\right]^2\right)$$

$$= E\left(X^2 - 2XE(X) + \left[E(X)\right]^2\right)$$

Autre Notation : $= E\left(X^2\right) - 2E\left[XE(X)\right] + E\left[E(X)^2\right]$

$$= E\left(X^2\right) - 2E(X)^2 + E(X)^2$$

$$= E\left(X^2\right) - E(X)^2$$

Alors : $V(X) = E\left(X^2\right) - E(X)^2$

Remarque :

Si $X(\Omega)$ est infini, il n'est nullement évident que $V(X)$ existe. De plus comme $\left[X - E(X)\right]^2 \geq 0$ nécessairement $V(X) \geq 0$. Par définition, une **variance est toujours positive**

La variance est également notée σ^2 si aucune confusion n'est à craindre.

Si X est une variable aléatoire ayant une variance $V(X)$, on appelle **écart-type** de X, le réel : $\sigma(X) = \sqrt{V(X)}$

Remarque :

L'écart-type permet de disposer d'un paramètre de dispersion qui s'exprime dans les **mêmes unités** que la variable aléatoire elle-même.

4.2.1 Variables aléatoires discrètes

Si X est une variable aléatoire **discrète** de loi de probabilité $\left(x_i, p_i\right)_{i=1,n}$ définie sur un nombre fini (n) d'évènements élémentaires alors la variance est égale à : $V(X) = \sum_{i=1}^{n}\left(x_i - E(X)\right)^2 p_i = \sum_{i=1}^{n} x_i^2 p_i - E(X)^2$

<u>**Exemple :**</u>

(1) Si l'on reprend l'exemple d'une <u>fratrie de deux enfants</u>, la variance de la variable aléatoire « nombre de filles « est :

$$V(X) = \frac{1}{4}(0-1)^2 + \frac{1}{2}(1-1)^2 + \frac{1}{4}(2-1)^2 = \frac{1}{2}$$

$$V(X) = \frac{1}{2} \quad donc : \sigma(X) = \sqrt{\frac{1}{2}} = 0.7$$

(2) Dans le cas de la loi de probabilité du <u>nombre de piles</u> lors du lancer de 3 pièces, quelles sont les valeurs de l'espérance et de la variance de cette loi ?

<u>Réponse</u>.

4.2.2 Variables aléatoires continues

Si X est une variable aléatoire **continue** donnée par sa densité de probabilité alors la variance de X est le nombre réel positif tel que :

$$V(X) = \int_{-\infty}^{+\infty} x^2 f(x)dx - E(X)^2$$

$$V(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f(x)dx$$

Exemple :

Dans le cadre de la recolonisation de l'étang par la population de canard colvert, la variance de la loi de probabilité est : $V(T) = \int_{-\infty}^{+\infty} (t - E(T))^2 f(t)dt = \frac{5}{4}$

$Alors : \sigma(T) = \sqrt{5/4} = 1.12$

4.2.3 Propriétés de la variance

Si X est une variable aléatoire admettant une variance alors :

(P_1) $\forall a \in \mathbb{R} : V(aX) = a^2 V(X)$

(P_2) $\forall (a,b) \in \mathbb{R}^2 : V(aX + b) = a^2 V(X)$

(P_3) $V(X) = 0 \Leftrightarrow X = E(X)$

Exercice 1 : Le nombre X de kilogrammes de tomates récoltés dans un jardin en une semaine, est une variable aléatoire dont la distribution de probabilité est la suivante :

X	0	1	2	3
$P(X = x)$	0,1	0,5	0,3	0,1

1. Quelle est l'espérance mathématique de X et quelle est sa variance ?
2. Pendant les six semaines de la saison de récolte, la distribution de probabilité reste la même ; rechercher l'espérance mathématique et la variance de la va. Y : récolte totale en six semaines.