Probabilités et statistique

Les variables aléatoires à une dimension

Module 2

Dr Diagne

Plan

- **Variable aléatoire**

  Les variables aléatoires discrètes

  Les variables aléatoires continues

- **Espérance mathématique**

  Fonction de variable aléatoire

  Moments

- **Variance**

  Transformation linéaire

1. Variable aléatoire

- ***Définition***

Il s'agit de variable dont la valeur dépend du hasard.

Exemple

a) **X** : « le nombre d'ordinateurs défectueux »

b) **Y** : « la durée de vie d'une ampoule électrique en heures ».

On note R_X l'ensemble des valeurs possibles de la variable aléatoire X.

Il existe deux types de variables aléatoires: *discrètes* et *continues*.

A. Les variables aléatoires *discrètes*

La variable aléatoire X est discrète si l'ensemble R_X est fini ou infini dénombrable.

Exemple

a) X : " la valeur obtenue en lançant un dé ".

$$R_X = \{1,2,3,4,5,6\} \text{ est un ensemble fini.}$$

La variable X est donc discrète.

b) Y : « nombre d’étudiants connectés à Internet en une journée »

$R_Y = \{0, 1, 2, 3, 4, ...\}$ est infini et dénombrable

Y est une variable discrète

c) Z : « la qualité d’une ligne de connexion téléphonique »

$\boldsymbol{R_Z = \{1, 2, 3, 4, 5\}}$ ***est fini.***

1 : Passabe, 2 : Assez bien, 3 : bien, 4 : Très bien, 5 : Excellent.

Z est une variable discrète.

d) $\dot{T} : R_T = \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, ...\right\}$ est infini et dénombrabl e

T est une variable discrète

e) ***U :*** $.\boldsymbol{R_U} = [0\ ,\ 1]$ ***est un ensemble infini non dénombrabl e.***

U n'est pas une variable discrète.

Fonction de masse

On considère la variable aléatoire X avec $\mathrm{R}_{\mathrm{X}} = \{ x_1, x_2, x_3, \ldots, x_n \}$

X prend une valeur x_i avec une probabilité $p(x_i)$.

On note $$p : R_x \longrightarrow [0, 1]$$

$$x_i \longrightarrow p(x_i) = P(X = x_i)$$

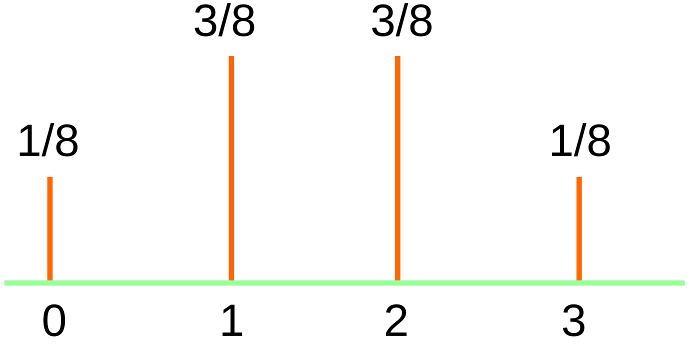
Avec $0 \leq p(x_i) \leq 1$ ***et*** $\sum_{i=1}^{n} p(x_i) = 1$

$p(x)$ désigne la fonction de masse de la variable X .

Exemple

X : « le nombre de piles dans 3 lancers d'une pièce de monnaie.»

X	0	1	2	3
p(x)	1/8	3/8	3/8	1/8



Exemples

- X : << le nombre de succés en n épreuves >>

$$p(x)=\binom{n}{x}p^{x}(1-p)^{n-x} \quad x=0,1,2,\ldots,n\,; \quad 0<p<1.$$

- Y : << le nombre d'épreuves pour obtenir 1 succés >>

$$p(y)=(1-p)^{y-1}p \quad , \quad y=1,\ 2,\ 3,\ \ldots.. \text{ et } \quad 0<p<1.$$

- Z :<< le nombre de succés en temps discret >>

$$p(z)=\frac{e^{-\lambda}\lambda^{z}}{z!} \qquad z=0,1,\ 2,\ldots. \text{ et } \ \lambda>0.$$

Remarques

- X : $R_X=\{0,1,2,3\}$

$$P(X\le 2)=P(X=0)+P(X=1)+P(X=2)=p(0)+p(1)+p(2)$$

$$P(X<2)=P(X=0)+P(X=1)=p(0)+p(1)$$

- Y : $R_Y=\{0,1,2,3,\ldots..,100\}$

$$P(X\le 99)=P(X=0)+P(X=1)+P(X=2)+\ldots..+P(X=99)=p(0)+p(1)+p(2)+\ldots..+p(99)$$

$$=1-P(X=100)=1-p(100)$$

B. Les variables aléatoires *continues*

- X ne peut pas prendre de valeur réelle fixe.
- R_X est toujours un intervalle réel.
- $P(X = x) = 0$ pour tout réel x.

Exemple

$$\boldsymbol{R_x} = \left]-\infty\,,\ +\infty\right[\ \ ;\ \ \boldsymbol{R_x} = \left[0\,,\ +\infty\right[\ \ ;\ \ \boldsymbol{R_x} = \left[0\,,\ \boldsymbol{1}\right] \cup \left]3\,,\ \boldsymbol{5}\right]$$

➢ ***Fonction de densité***

La fonction f est une densité de probabilité associée à la variable aléatoire continue X si elle satisfait les conditions suivantes:

1. $f(x) \geq 0$ pour tout réel x.
2. $\int_{R_X} f(x)dx = 1$

l'intégrale est prise sur R_X.

Exemple

X : « le temps d'attente en minutes à un arrêt bus »

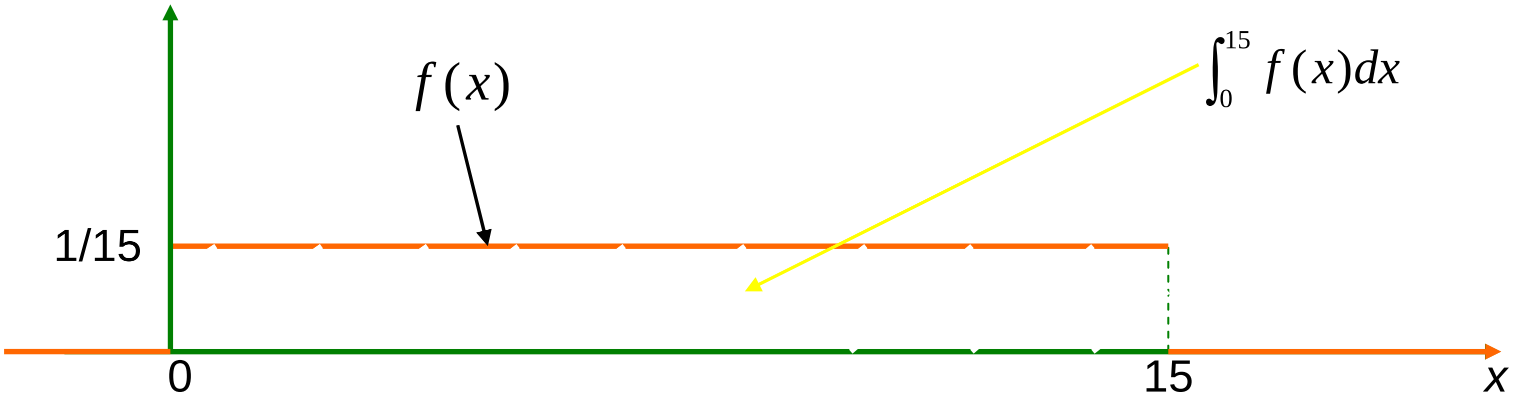
$$\boldsymbol{f(x)} = \begin{cases} \boldsymbol{\frac{1}{15}} & \boldsymbol{si} \ \ 0 \le \boldsymbol{x} \le 15 \\ 0 & \boldsymbol{sinon} \end{cases}$$

$$\boldsymbol{R_x} = [0\ ,\ \boldsymbol{15}]$$

$f(x)$ est elle une densité de probabilités ?

- $f(x) = \frac{1}{15}$ ou 0 donc $f(x) \ge 0$
- $\int_{R_X} f(x)dx = \int_0^{30} \frac{1}{15} dx = \frac{1}{15}\int_0^{30} 1\, dx = \frac{1}{15}[x]_0^{30} = \frac{1}{30}(30-0) = \frac{30}{30} = 1$

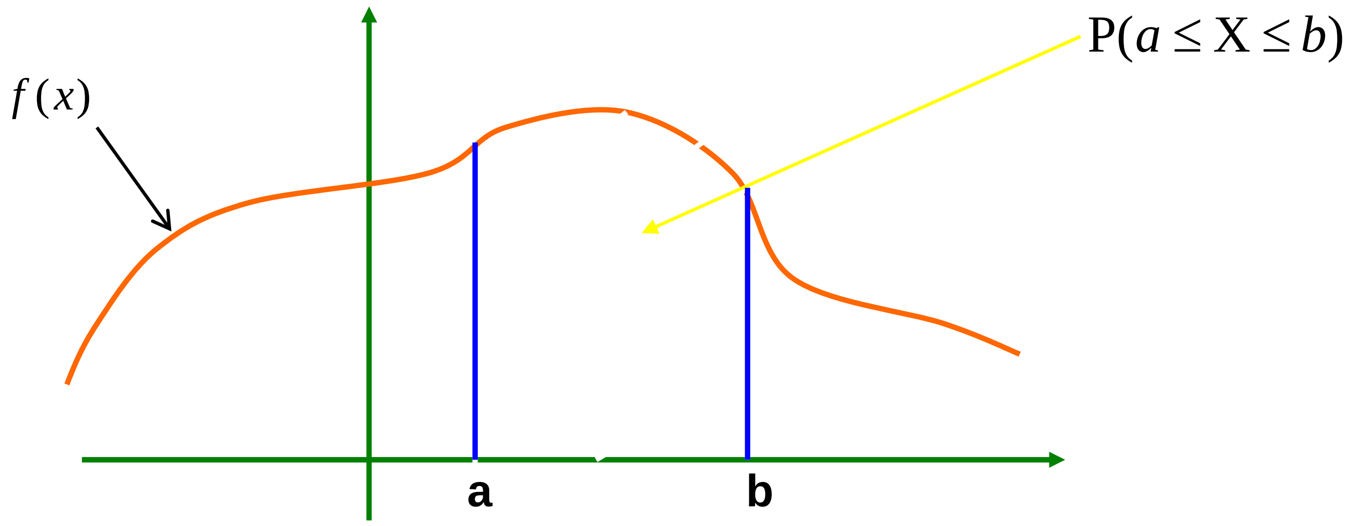
Donc $f(x)$ est bien une densité de probabilité.



Interprétation

- $P(a \leq X \leq b) = \int_a^b f(x)dx$

C'est l'aire sous la courbe $f(x)$ entre a et b.



Fonction de répartition

La fonction de répartition $F(x)$ d'une variable aléatoire X est définie par:

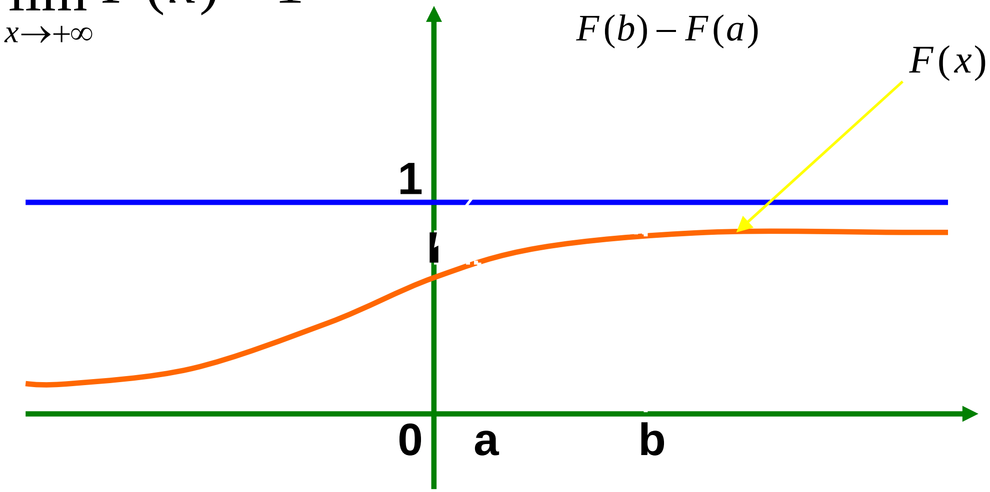
$$F : \mathbf{R} \to [0 \;\; 1]$$
$$x \to F(x) = P(X \leq x)$$

Si la variable aléatoire X est continue de densité $f(x)$ alors

$$F(a) = P(X \le a) = \int_{-\infty}^{a} f(x)dx$$

Propriétés

1. $F^{'}(x) = f(x) \ge 0 \Rightarrow F(x)$ est non décroissante.
2. $F(b) - F(a) = P(a \le X < b)$
3. $\lim\limits_{x \to -\infty} F(x) = 0$
4. $\lim\limits_{x \to +\infty} F(x) = 1$

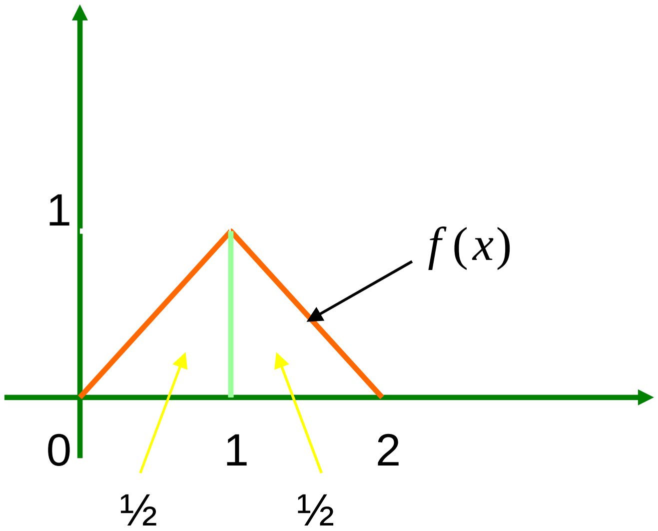


Exemple

X : variable aléatoire continue

$$f(x) = \begin{cases} x & \text{si } 0 \le x \le 1 \\ 2 - x & \text{si } 1 < x \le 2 \end{cases}$$

$$R_X = [0 \quad 1] \cup]1 \quad 2]$$



a) $f(x)$ est-elle une densité de probabilité ?

- $0 \le x < 1 \Rightarrow x \ge 0 \;$ et $\; 1 < x \le 2 \Rightarrow 2 - x \ge 0$

  $\Rightarrow f(x) \ge 0$

- $\int_{R_X} f(x)dx = \int_0^1 x\,dx + \int_1^2 (2-x)dx$

$$= \left[\tfrac{x^2}{2}\right]_0^1 + \left[2x - \tfrac{x^2}{2}\right]_1^2$$

$$= \left(\tfrac{1}{2} - 0\right) + \left[\left(4 - \tfrac{4}{2}\right) - \left(2 - \tfrac{1}{2}\right)\right] = 1$$

Donc $f(x)$ est une densité de probabilité.

b) Déterminer la fonction de répartition de X.

$$F(x)=P(X\le x)=\int_{-\infty}^{x} f(u)du$$

$$\bullet\, x\le 0 \Rightarrow F(x)=\int_{-\infty}^{x} 0\,du=0$$

$$\bullet\, 0<x\le 1 \Rightarrow F(x)=\int_{-\infty}^{x} f(u)du=\int_{-\infty}^{0} 0\,du+\int_{0}^{x} u\,du$$

$$=0+\left[\tfrac{u^2}{2}\right]_0^x=\tfrac{x^2}{2}$$

$$\bullet\, 2<x \Rightarrow F(x)=\int_{-\infty}^{x} f(u)du=\int_{-\infty}^{0} 0\,du+\int_{0}^{1} u\,du+\int_{1}^{2}(2-u)\,du+\int_{2}^{x} 0\,du$$

$$=0+\left[\tfrac{u^2}{2}\right]_0^1+\left[2u-\tfrac{u^2}{2}\right]_1^2+0=\left(\tfrac{1}{2}-0\right)+\left[\left(4-\tfrac{4}{2}\right)-\left(2-\tfrac{1}{2}\right)\right]=1$$

$$F(x)=\begin{cases} 0 & \text{si} \quad x\le 0 \\ \frac{x^2}{2} & \text{si} \ 0<x\le 1 \\ -\frac{x^2}{2}+2x-1 & \text{si} \ 1<x\le 2 \\ 1 & \text{si} \ 2<x \end{cases}$$

c) Calculer la probabilité $\mathrm{P}\left(\frac{1}{2}\le \mathrm{X}\le \frac{3}{2}\right)$.

$$\mathrm{P}\left(\tfrac{1}{2}\le \mathrm{X}\le \tfrac{3}{2}\right)=F\left(\tfrac{3}{2}\right)-F\left(\tfrac{1}{2}\right)=\left(-\tfrac{(3/2)^2}{2}+2\times\tfrac{3}{2}-1\right)-\tfrac{(1/2)^2}{2}=\frac{3}{4}$$

d) Calculer la probabilité $\mathrm{P}(\mathrm{X}\ge \frac{5}{3})$

$$\mathrm{P}\left(\mathrm{X}\ge \tfrac{5}{3}\right)=1-\mathrm{P}\left(\mathrm{X}<\tfrac{5}{3}\right)=1-\mathrm{P}\left(\mathrm{X}\le \tfrac{5}{3}\right) \qquad \text{X est une variable continue.}$$

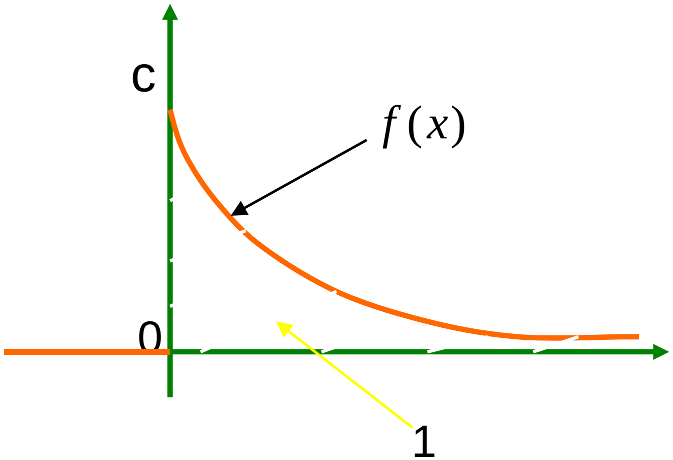
$$=1-F\left(\tfrac{5}{3}\right)=1-\left(-\tfrac{(5/3)^2}{2}+2\times\tfrac{5}{3}-1\right)=\tfrac{1}{18}$$

Application

X : « durée de vie en années d'une catégorie d'ordinateurs »

La densité de probabilité est définie par:

$$f(x) = \begin{cases} 0 & si \quad x < 0 \\ c\, e^{-0.05\,x} & si \quad x \geq 0 \end{cases}$$



c est une constante positive.

$$R_X = [0\,,\ +\infty[$$

a) Quelle est la valeur de c ?

$$\int_0^{+\infty} ce^{-0.05x}dx = 1 \Leftrightarrow c\int_0^{+\infty} e^{-0.05x}dx = 1 \Leftrightarrow c\left[-\tfrac{1}{0.05}e^{-0.05x}\right]_0^{+\infty} = 1 \Leftrightarrow c\left[-\tfrac{1}{0.05}(e^{-\infty} - e^{0})\right] = 1$$

$$\Leftrightarrow c\left[-\tfrac{1}{0.05}(e^{-\infty} - e^{0})\right] = 1 \Leftrightarrow c\left[-\tfrac{1}{0.05}(0-1)\right] = 1$$

$$\Leftrightarrow \frac{c}{0.05} = 1 \Leftrightarrow c = 0.05$$

b) Quelle est la probabilité qu'un ordinateur de cette catégorie ait une durée de vie entre 15 et 25 ans ?

$$P(15 \leq X \leq 25) = \int_{15}^{25} 0.05e^{-0.05x}dx = \left[-e^{-0.05x}\right]_{15}^{25} = -e^{-0.05\times 25} + e^{-0.05\times 15} = 0.18$$

c) Déterminer la fonction de répartition de X .

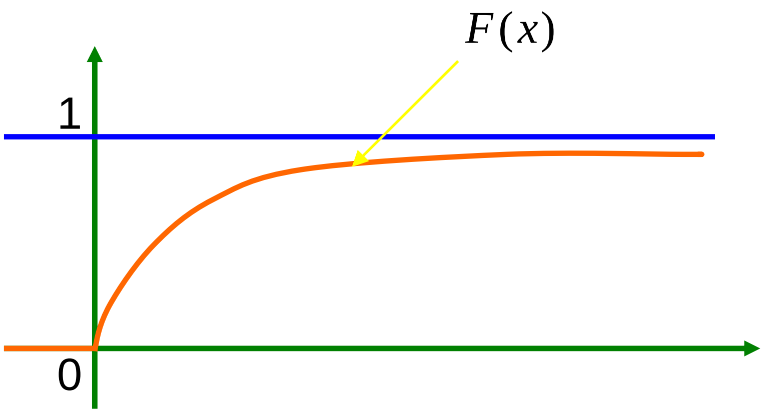
- $x < 0 \Rightarrow F(x) = \int_{-\infty}^{x} f(u)du = \int_{-\infty}^{x} 0\, du = 0$

- $0 \le x \Rightarrow F(x) = \int_{-\infty}^{x} f(u)du = \int_{-\infty}^{0} 0\, du + \int_{0}^{x} 0.05e^{-0.05u} du$

$$= 0.05\left[-\frac{1}{0.05} e^{-0.05u}\right]_0^x = \left[-e^{-0.05u}\right]_0^x$$

$$= -e^{-0.05x} + 1$$

$$F(x) = \begin{cases} 0 & si \quad x < 0 \\ 1 - e^{-0.05x} & si \quad x \ge 0 \end{cases}$$



d) Quelle est la probabilité que la durée de vie d'un ordinateur de cette catégorie dépasse 30 ans ?

$$\mathrm{P}(\mathrm{X} > 30) = 1 - \mathrm{P}(\mathrm{X} \le 30) = 1 - F(30)$$

$$= 1 - \left(1 - e^{-0.05 \times 30}\right) = e^{-0.05 \times 30} = 0.22$$

2. Espérance mathématique

➢ *Définition*

L'espérance mathématique d'une variable aléatoire X est la moyenne limite pondérée par les probabilités élémentaires des valeurs possibles de la variable. On note $E(\mathrm{X}) = \mu$.

$$E(\mathrm{X}) = \mu = \begin{cases} \sum_{R_\mathrm{X}} x_i\, p(x_i) & \text{si X est une variable discrète} \\ \int_{R_\mathrm{X}} x f(x)\, dx & \text{si X est une variable continue} \end{cases}$$

***Exemple* 1**

X : « lancer un dé une fois »

$$R_\mathrm{X} = \{1, 2, 3, 4, 5, 6\}$$

$$p(x_i) = \tfrac{1}{6} \text{ pour tout } x_i \in R_\mathrm{X}.$$

$$E(\mathrm{X}) = \sum_{i=1}^{6} x_i\, p(x_i) = 1 \times \tfrac{1}{6} + 2 \times \tfrac{1}{6} + 3 \times \tfrac{1}{6} + 4 \times \tfrac{1}{6} + 5 \times \tfrac{1}{6} + 6 \times \tfrac{1}{6} = \tfrac{21}{6} = 3.5$$

Exemple 2

X : variable aléatoire de densité

$$f(x)=\begin{cases} x & si\ \ 0\le x\le 1 \\ 2-x & si\ \ 1<x\le 2 \end{cases}$$

$$R_X=[0,\ 1]\cup]1,\ 2]$$

$$\bullet\ E(X)=\int_{R_X} x\ f(x)dx=\int_0^1 x\ xdx+\int_1^2 x(2-x)dx$$

$$=\int_0^1 x^2dx+\int_1^2(2x-x^2)dx=\left[\tfrac{x^3}{3}\right]_0^1+\left[x^2-\tfrac{x^3}{3}\right]_1^2$$

$$=\left(\tfrac{1}{3}-0\right)+\left[\left(4-\tfrac{8}{3}\right)-\left(1-\tfrac{1}{3}\right]=\tfrac{1}{3}+\tfrac{4}{3}-\tfrac{2}{3}=\tfrac{3}{3}=1\right.$$

Exemple 3

X :" durée de vie en années d'une catégorie d'ordinateurs"

$$f(x)=\begin{cases} 0 & si\ \ x<0 \\ 0.05\,e^{-0.05x} & si\ \ x\ge 0 \end{cases}$$

$$R_X=[0\ \ +\infty[$$

$$\bullet\ E(X)=\int_{R_X} xf(x)dx=\int_0^{+\infty} x\times 0.05\,e^{-0.05x}dx=0.05\int_0^{+\infty}\underbrace{xe^{-0.05x}dx}$$

Intégration par parties : $u=x \qquad u'=1$

$$v'=e^{-0.05x} \qquad v=-\tfrac{1}{0.05}\,e^{-0.05x}$$

$$\bullet\ \int_0^{+\infty} xe^{-0.05x}dx=\left[x(-\tfrac{1}{0.05}e^{-0.05x})\right]_0^{+\infty}-\int_0^{+\infty}-\tfrac{1}{0.05}e^{-0.05x}dx=0+\tfrac{1}{0.05}\left[-\tfrac{1}{0.05}e^{-0.05x}\right]_0^{+\infty}=\tfrac{1}{(0.05)^2}(-e^{-\infty}+e^0)=\frac{1}{(0.05)^2}$$

$$E(X)=0.05\times\frac{1}{(0.05)^2}=\frac{1}{0.05}=20$$

A. Espérance mathématique de fonction d'une variable aléatoire

Soit X une variable aléatoire et $g(x)$ une fonction continue.

- $g(X)$ est une variable aléatoire.
- $E[g(X)] = \begin{cases} \sum_{R_X} g(x_i)\,p(x_i) & \text{si X est une variable discrète} \\ \int_{R_X} g(x) f(x)dx & \text{si X est une variable continue} \end{cases}$

Exemple

X : variable aléatoire de densité

$$f(x) = \begin{cases} \frac{x^2}{9} & \text{si } 0 \le x \le 3 \\ 0 & \text{sinon} \end{cases}$$

$$R_X = [0 \quad 3] \qquad\qquad g(X) = X^2 + 1$$

- $E(X) = \int_{R_X} xf(x)dx = \int_0^3 x\frac{x^2}{9}dx = \int_0^3 \frac{x^3}{9}dx = \left[\frac{x^4}{36}\right]_0^3 = \frac{81}{36} = \frac{9}{4}$

- $E[g(X)] = \int_{R_X} g(x)f(x)dx = \int_0^3 (x^2+1)\frac{x^2}{9}dx$

$$= \int_0^3 \left(\frac{x^4}{9} + \frac{x^2}{9}\right)dx = \left[\frac{x^5}{45} + \frac{x^3}{27}\right]_0^3 = \frac{27\times 9}{45} + \frac{27}{27} = \frac{32}{5}$$

- $g[E(X)] = [E(X)]^2 + 1 = \left(\frac{9}{4}\right)^2 + 1 = \frac{97}{16}$
- $E[g(X)] \neq g[E(X)]$

Cas d'une transformation linéaire

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

avec a et b des constantes réelles.

Exemple

$$f(x)=\begin{cases}\frac{x^2}{9} & si\ \ 0\le x\le 3\\ 0 & sinon\end{cases}$$

$R_X=[0,\,3]$ $\qquad g(X)=2X-3$

- $E(X)=\int_{R_X} xf(x)dx=\int_0^3 x\,\frac{x^2}{9}dx=\int_0^3\frac{x^3}{9}dx=\left[\frac{x^4}{36}\right]_0^3=\frac{81}{36}=\frac{9}{4}$

- $g[E(X)]=2\times\frac{9}{4}-3=\frac{3}{2}$

- $E(2X-3)=\int_0^3(2x-3)\frac{x^2}{9}dx$

$$=\int_0^3\left(\frac{2x^3}{9}-\frac{3x^2}{9}\right)dx=\left[\frac{2x^4}{36}-\frac{3x^3}{27}\right]_0^3=\left(\frac{2\times81}{36}-\frac{3\times27}{27}\right)=\frac{3}{2}$$

- $E(2X-3)=2E(X)-3$

B. Moments d'une variable aléatoire

Le k^e moment d'une variable aléatoire X est définie par l'espérance mathématique de X^k. On note $\mu_k = E(X^k)$; $k = 1, 2, 3, \ldots$

$$\mu_k = E(X^k) = \begin{cases} \sum_{x \in R_X} x^k p(x) & \text{si X est une variable discrète} \\ \int_{R_X} x^k f(x) dx & \text{si X est une variable continue} \end{cases}$$

Exemple

$$f(x) = \begin{cases} x & si \ 0 \le x \le 1 \\ 2 - x & si \ 1 < x \le 2 \end{cases}$$

$$R_X = [0 \ \ 1] \cup]1 \ \ 2]$$

$$\bullet \ E(X^k) = \int_{R_X} x^k f(x) dx = \int_0^1 x^k x dx + \int_1^2 x^k (2 - x) dx$$

$$= \int_0^1 x^{k+1} dx + \int_1^2 (2x^k - x^{k+1}) dx = \left[\frac{x^{k+2}}{k+2}\right]_0^1 + \left[\frac{2x^{k+1}}{k+1} - \frac{x^{k+2}}{k+2}\right]_1^2$$

$$= \left(\frac{1}{k+2} - 0\right) + \left[\left(\frac{2 \times 2^{k+1}}{k+1} - \frac{2^{k+2}}{k+2}\right) - \left(\frac{2}{k+1} - \frac{1}{k+2}\right)\right] = \frac{1 - 2^{k+2} + 1}{k+2} + \frac{2^{k+2} - 2}{k+1}$$

$$\bullet \ E(X^k) = \frac{2^{k+2} - 2}{(k+1)(k+2)}$$

$\bullet$ ***le 1er moment :*** $k = 1 \quad \mu_1 = E(X^1) = \frac{2^3 - 2}{2 \times 3} = \frac{6}{6} = 1 = E(X)$

$\bullet$ ***le 2^e moment :*** $k = 2 \quad \mu_2 = E(X^2) = \frac{2^4 - 2}{3 \times 4} = \frac{14}{12} = \frac{7}{6}$

$\bullet$ ***le 3^e moment :*** $k = 3 \quad \mu_3 = E(X^3) = \frac{2^5 - 2}{4 \times 5} = \frac{30}{20} = \frac{3}{2}$

$\bullet$ ***le 4^e moment :*** $k = 4 \quad \mu_4 = E(X^4) = \frac{2^6 - 2}{7 \times 8} = \frac{62}{56} = \frac{31}{28}$

3. Variance

- ***Définition***

La variance de la variable aléatoire X est définie par le 2e moment centré.

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

- ***Interprétation***

La variance mesure la dispersion des valeurs de X autour de l'espérance mathématique.

Plus la variance est grande plus les valeurs de X sont éloignées de son espérance.

- ***Notation***

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

$$= E(X^2) - [E(X)]^2$$

$$\sigma^2 = \mu_2 - \mu_1^2$$

Exemple

- X : « lancer un dé » ; $R_X = \{1, 2, 3, 4, 5, 6\}$; $p(x) = 1/6$ pour tout $x \in R_X$.

$$\mu_1 = E(X) = \sum_{x \in R_X} x\, p(x) = \tfrac{21}{6}$$

$$\mu_2 = E(X^2) = \sum_{x \in R_X} x^2 p(x) = \tfrac{91}{6}$$

$$\sigma^2 = V(X) = \mu_2 - \mu_1^2 = \tfrac{91}{6} - \left(\tfrac{21}{6}\right)^2 = \tfrac{105}{36}$$

- X : variable aléatoire de densité triangulaire

$$f(x) = \begin{cases} x & si \ \ 0 \le x \le 1 \\ 2 - x & si \ \ 1 < x \le 2 \end{cases}$$

$$R_X = [0 \ \ 1] \cup]1 \ \ 2]$$

$$\mu_1 = 1 \qquad \mu_2 = \tfrac{7}{6}$$

$$\sigma^2 = \tfrac{7}{6} - 1^2 = \frac{1}{6}$$

Cas d'une transformation linéaire

$$V(aX+b)=a^2V(X)$$

a et b sont des constantes réelles.

Remarque

La constante *b* ne joue aucun rôle dans le calcul de la variance.

Exemple

$$f(x)=\begin{cases} x^2/9 & \text{si } 0\le x\le 3 \\ 0 & \text{sinon}\end{cases}$$

$R_x=[0\quad 3]$ $\qquad Y=2X-3$

- $E(X)=\frac{9}{4}$
- $E(X^2)=\int_0^3 x^2\ \frac{x^2}{9}dx\ =\left[\frac{x^5}{45}\right]_0^3=\frac{3^5}{45}=\frac{27}{5}$
- $V(X)=\frac{27}{5}-\left(\frac{9}{4}\right)^2=\frac{27}{80}$

$\Rightarrow V(Y)=\frac{27}{80}=4\times\frac{27}{80}=2^2\times V(X)$

- $E(Y)=E(2X-3)=2E(X)-3=\frac{18}{4}-3=\frac{3}{2}$
- $E(Y^2)=E\left[(2X-3)^2\right]=E(4X^2-12X+9)$
  $=4E(X^2)-12E(X)+9=4\times\frac{27}{5}-12\times\frac{9}{4}+9=\frac{72}{20}=\frac{18}{5}$
- $V(Y)=\frac{18}{5}-\left(\frac{3}{2}\right)^2=\frac{27}{20}$
- $V(2X-3)=2^2\times V(X)$