MESURES ET ANALYSES STATISTIQUES DE DONNÉES

Probabilités

Master Génie des Systèmes Industriels, mentions ACCIE et RIM

Université du Littoral - Côte d'Opale, La Citadelle

Laurent SMOCH

(smoch@lmpa.univ-littoral.fr)

Septembre 2012

Laboratoire de Mathématiques Pures et Appliquées Joseph Liouville
Université du Littoral, zone universitaire de la Mi-Voix, bâtiment H. Poincarré
50, rue F. Buisson, BP 699, F-62228 Calais cedex

Table des matières

Chapitre 3

Les variables aléatoires réelles

3.1 Introduction - Exemple

Une épreuve consiste à jeter deux dés discernables A et B. On considère la somme des numéros apparus sur la face supérieure des dés. On peut symboliser les résultats à l'aide du tableau ci-dessous :

B A	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

L'univers Ω correspondant à cette expérience aléatoire est défini par

$$\Omega = \{(i, j), i = 1, 2, 3, 4, 5, 6 \text{ et } j = 1, 2, 3, 4, 5, 6\}, \text{ avec } \mathrm{Card}(\Omega) = 6^2 = 36$$

où chaque événement élémentaire est équiprobable. On dispose donc d'un espace probabilisé $(\Omega, \mathcal{P}(\Omega), p)$.
Considérons l'application

$$\begin{array}{rcl} X : \Omega & \to & \mathbb{R} \\ (i, j) & \mapsto & i + j \end{array}$$

L'ensemble des valeurs prises par X est appelé univers image et vaut

$$X(\Omega) = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}.$$

On définit ainsi une fonction de probabilité p_X associée à la variable X, déduite de p par

$$p_X(x) = p(\{X = x\}) = p(\{X^{-1}(x)\})$$

Par exemple, $p_X(3) = p(\{X = 3\}) = p(\{(1, 2), (2, 1)\}) = \dfrac{2}{36} = \dfrac{1}{18}$.

X est une variable aléatoire réelle car associée à une épreuve aléatoire réelle (à valeurs dans $\mathbb{R}$) et sa fonction de probabilité p_X est déduite de p.

x_i	2	3	4	5	6	7	8	9	10	11	12	Total
$p(\{X = x_i\})$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	1

En notant $p_i = p(\{X = x_i\})$ on a $\sum_i p_i = 1$.

Remarque 3.1.1 À la même épreuve aléatoire on peut associer plusieurs variables aléatoires par exemple $Y(i,j) = i.j$, $Z(i,j) = \sup(i,j)$, ...

3.2 Définitions et propriétés

1. Une **variable aléatoire réelle** définie sur un espace probabilisé $(\Omega, p(\Omega), p)$ est une application de Ω dans $\mathbb{R}$ définie par

$$\boxed{\begin{array}{rcl} X : \Omega & \to & \mathbb{R} \\ \omega_i & \mapsto & X(\omega_i) = x_i \end{array}}$$

L'ensemble des valeurs prises par la variable aléatoire X s'appelle l'**univers image** et est défini par $X(\Omega) = \{x_1, x_2, \ldots, x_n\}$. ω_i étant un des résultats possibles de l'expérience, $X(\omega_i)$ est en quelque sorte la caractéristique de ω_i qui nous intéresse.
Comme on ne peut pas prévoir avec certitude quel ω_i on obtiendra à l'issue de l'expérience aléatoire, on ne peut pas non plus prévoir avec certitude quelle valeur prendra $X(\omega_i)$. C'est pourquoi la fonction X est souvent interprétée comme une "grandeur aléatoire". On note

$$\boxed{\{X = x_i\} = \{\omega_i \in \Omega / X(\omega_i) = x_i\}}$$

Chaque événement $\{X = x_i\}$ a une probabilité p_i de se produire

$$\boxed{p(\{X = x_i\}) = p_X(x_i) = p_i}$$

2. On appelle **loi de probabilité** de la variable aléatoire X l'application

$$\boxed{\begin{array}{rcl} f : X(\Omega) & \to & [0;1] \\ x_i & \mapsto & f(x_i) = p(\{X = x_i\}) \end{array}}$$

Définir la loi de probabilité d'une variable aléatoire revient à déterminer la probabilité de chacun des événements $\{X = x_i\}$ c'est-à-dire $p_i = p(\{X = x_i\})$. On vérifiera alors que

$$\boxed{\sum_i p_i = 1}$$

3. On peut citer les propriétés essentielles suivantes :

$$\boxed{\begin{array}{l} \text{Si } X \text{ et } Y \text{ sont deux variables aléatoires définies sur un même espace probabilisé } (\Omega, \mathcal{P}(\Omega), p) \text{ alors} \\ \qquad X.Y, \quad X+Y, \quad aX+bY, \quad X+a, \quad aX, \quad X^2 \\ \text{sont des variables aléatoires } (a \text{ et } b \text{ étant deux réels}). \end{array}}$$

Exemple 3.2.1 Soit l'épreuve consistant à jeter un dé équilibré, $\Omega = \{1, 2, 3, 4, 5, 6\}$ est l'univers associé à cette expérience. La distribution de probabilité est équiprobable (puisque le dé est équilibré). Soient les variables aléatoires définies comme suit :

- $$\begin{array}{rcl} X : \Omega & \to & \mathbb{R} \\ \omega_i & \mapsto & X(\omega_i) = \begin{cases} 1 & \text{si } \omega_i \in \{1,2,3,5\} \\ 0 & \text{si } \omega_i \in \{4,6\} \end{cases} \end{array}.$$

  L'univers image de la variable X est $X(\Omega) = \{0, 1\}$. Sa loi de probabilité est définie par

  $p_X(1) = p(\{X = 1\}) = p(\{1, 2, 3, 5\}) = \dfrac{4}{6} = \dfrac{2}{3}$,

  $p_X(0) = p(\{X = 0\}) = p(\{4, 6\}) = \dfrac{2}{6} = \dfrac{1}{3}$.

- $\begin{array}{rccl} Y & : & \Omega & \to & \mathbb{R} \\ & & \omega_i & \mapsto & Y(\omega_i) = \omega_i^2 \end{array}$.

  L'univers image de la variable Y est $Y(\Omega) = \{1, 4, 9, 16, 25, 36\}$. Sa loi de probabilité est définie par

Valeurs de Y : y_i	1	4	9	16	25	36	Total
$p(\{Y = y_i\}) = p_i$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

Considérons la variable $Z = X + 2$. Cette variable a pour distribution de probabilité

Valeurs de Z : z_i	2	3	Total
$p(\{Z = z_i\}) = p_i$	$\frac{1}{3}$	$\frac{2}{3}$	1

Considérons la variable $T = 2X$. Cette variable a pour distribution de probabilité

Valeurs de T : t_i	0	2	Total
$p\{T = t_i\} = p_i$	$\frac{1}{3}$	$\frac{2}{3}$	1

Considérons les variables $U = X + Y$ et $V = X \times Y$. On a les résultats suivants :

ω_i	1	2	3	4	5	6
$p(\omega_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$X(\omega_i)$	1	1	1	0	1	0
$Y(\omega_i)$	1	4	9	16	25	36
$(X + Y)(\omega_i)$	2	5	10	16	26	36
$(X \times Y)(\omega_i)$	1	4	9	0	25	0

L'univers image de la variable $U = X + Y$ est $U(\Omega) = \{2, 5, 10, 16, 26, 36\}$, celui de la variable V est $V(\Omega) = \{0, 1, 4, 9, 25\}$. Les distributions des variables aléatoires U et V sont :

$U = X + Y$	2	5	10	16	26	36	Total
p_U	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

$V = X \times Y$	0	1	4	9	25	Total
p_V	$\frac{2}{6} = \frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

3.3 Types de variables aléatoires réelles

Soit X une variable aléatoire réelle, on envisage trois types de variables aléatoires réelles.

- **discrète finie**, si l'ensemble $X(\Omega)$ est fini. Les variables précédentes en sont des exemples.
- **discrète infinie**, si l'ensemble $X(\Omega)$ est infini dénombrable.

  Exemple 3.3.1 Soit l'expérience aléatoire qui consiste à lancer une pièce de monnaie, on considère la variable aléatoire X qui désigne le rang d'apparition du premier "face" obtenu. L'univers image est $\mathbb{N}^\star$. L'événement $\{X = k\}$ est obtenu en lançant k fois la pièce, les $k - 1$ premiers résultats étant "pile", le k-ième face. D'où

$$p(\{X=k\}) = \underbrace{\frac{1}{2}\times\ldots\times\frac{1}{2}}_{k-1 \text{ fois}}\times\frac{1}{2} = \left(\frac{1}{2}\right)^{k-1}\times\frac{1}{2} = \left(\frac{1}{2}\right)^{k}$$

- **continue**, si $X(\Omega)$ est un intervalle de $\mathbb{R}$ ou une réunion d'intervalles.

Cette distinction est importante car les techniques développées pour étudier les variables aléatoires sont très différentes selon que les variables sont discrètes ou continues.

3.4 Fonction de répartition d'une variable aléatoire discrète

3.4.1 Définition

Soit une variable aléatoire X définie sur un espace probabilisé $(\Omega, \mathcal{P}(\Omega), p)$. La **fonction de répartition** de X est définie par

$$\boxed{\begin{array}{rcl} F:\mathbb{R} & \to & [0;1] \\ a & \mapsto & F(a)=p(\{X\leq a\}) \end{array}}$$

3.4.2 Représentation graphique - Exemple

Une boîte contient 4 boules numérotées de 1 à 4. Une deuxième boîte contient 3 boules numérotées de 1 à 3. On prélève au hasard une boule de la première boîte puis une boule de la seconde boîte. Soit X la variable aléatoire représentant la valeur absolue de la différence des nombres indiqués sur les boules tirées. On se propose de

- déterminer la loi de probabilité de la variable X,
- représenter graphiquement sa fonction de répartition.

1. Si B_1 et B_2 sont respectivement les boules tirées dans la première et dans la deuxième boîte, on a le tableau ci-dessous :

B_1 \ B_2	1	2	3
1	0	1	2
2	1	0	1
3	2	1	0
4	3	2	1

L'univers associé à l'expérience décrite précédemment est défini par $\Omega = \{(i,j), i = 1,2,3,4 \text{ et } j = 1,2,3\}$. La variable aléatoire X est définie par

$$\begin{array}{rcl} X:\Omega & \to & \mathbb{R} \\ (i,j) & \mapsto & |i-j| \end{array}$$

où i et j sont respectivement les numéros des boules des urnes B_1 et B_2. L'univers image de la variable X est $X(\Omega) = \{0,1,2,3\}$.

- L'événement $\{X=0\}$ est obtenu de 3 façons :
  - en tirant les boules n°1 des boîtes 1 et 2 dont la probabilité est $\frac{1}{4}\times\frac{1}{3}=\frac{1}{12}$, les événements liés aux deux boîtes étant indépendants,
  - en tirant les boules n°2 des boîtes 1 et 2 dont la probabilité est aussi $\frac{1}{12}$,
  - en tirant les boules n°3 des boîtes 1 et 2 de probabilité $\frac{1}{12}$.

  Par conséquent, $p(\{X=0\}) = \frac{1}{12}+\frac{1}{12}+\frac{1}{12}=\frac{3}{12}=\frac{1}{4}$.
- L'événement $\{X=1\}$ est obtenu avec les couples $(1,2)$, $(2,1)$, $(3,2)$, $(2,3)$ et $(4,3)$. Donc

$$p(\{X = 1\}) = 5 \times \frac{1}{4} \times \frac{1}{3} = \frac{5}{12}.$$

– L'événement $\{X = 2\}$ est obtenu avec les couples $(3, 1)$, $(1, 3)$ et $(4, 2)$. Donc
$$p(\{X = 2\}) = 3 \times \frac{1}{4} \times \frac{1}{3} = \frac{3}{12} = \frac{1}{4}.$$

– L'événement $\{X = 3\}$ est obtenu avec le couple $(4, 1)$. Donc $p(\{X = 3\}) = \frac{1}{4} \times \frac{1}{3} = \frac{1}{12}$.

On en déduit la distribution de probabilité de la variable X :

x_i	0	1	2	3	Total
p_i	$\frac{3}{12}$	$\frac{5}{12}$	$\frac{3}{12}$	$\frac{1}{12}$	1

2. Représentation graphique de la fonction de répartition. On peut distinguer plusieurs cas selon la valeur prise par a :
   - $a < 0$,
     $F(a) = p(\{X \leq a\}) = p(\{X < 0\}) = p(\emptyset) = 0$
   - $0 \leq a < 1$,
     $F(a) = p(\{X \leq a\}) = p(\{0 \leq X < 1\}) = p(\{X = 0\}) = \frac{3}{12} = \frac{1}{4}$
   - $1 \leq a < 2$,
     $F(a) = p(\{X \leq a\}) = p(\{0 \leq X < 2\}) = p(\{X = 0\}) + p(\{X = 1\}) = \frac{3}{12} + \frac{5}{12} = \frac{2}{3}$
   - $2 \leq a < 3$,
     $F(a) = p(\{X \leq a\}) = p(\{0 \leq X < 3\}) = p(\{X = 0\}) + p(\{X = 1\}) + p(\{X = 2\}) = \frac{3}{12} + \frac{5}{12} + \frac{3}{12} = \frac{11}{12}$
   - $a \geq 3$, $F(a) = p(\{X \leq a\}) = p(\{0 \leq X \leq 3\}) = p(\Omega) = 1$

   d'où la représentation graphique

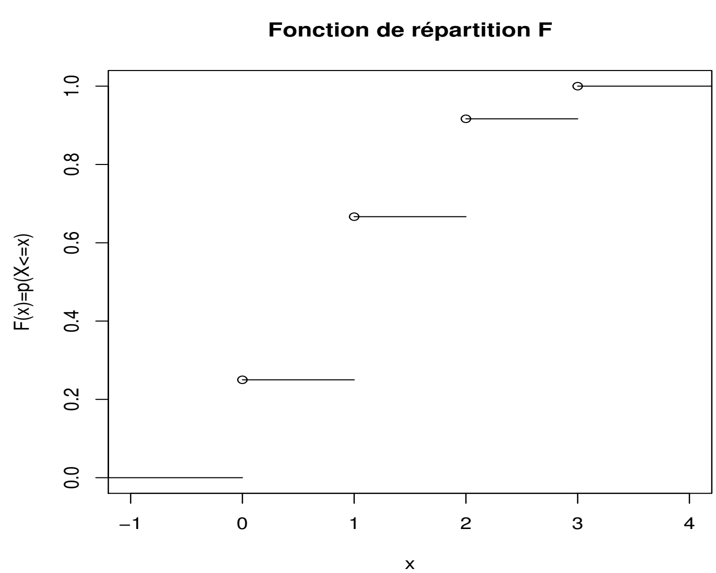

FIGURE 3.1

où le symbole "∘" indique que la valeur en abscisse est exclue.

On donne ci-dessous le code sous R permettant de générer ce diagramme en escalier :

```
> x<-c(0,1,2,3)
> probx<-c(3/12,5/12,3/12,1/12)
> frepx<-cumsum(probx)
> diagesc<-stepfun(x,c(0,frepx))
> plot(diagesc,vertical=FALSE,main=''Fonction de répartition F'',ylab=''F(x)=p(X<=x)'')
```

La fonction `stepfun` retourne un objet de type `function` qui définit une fonction constante par morceaux et continue à droite. La syntaxe est `F=stepfun(x, z)` où

– le vecteur x contient les points de discontinuité de la fonction,
– le vecteur z est de la forme $c(a, y)$ où a est la valeur prise par F avant le premier point de discontinuité et y les valeurs de la fonction aux points de discontinuité x.

Pour représenter graphiquement la fonction F, on utilise `plot` ou `lines` (pour superposer), avec l'argument `vertical=FALSE`.

3.4.3 Représentation graphique - Cas général

Soit X une variable aléatoire définie sur $(\Omega, \mathcal{P}(\Omega), p)$ avec $p_i = p(\{X = x_i\})$ et $X(\Omega) = \{x_1, x_2, \ldots, x_n\}$ l'univers image de X, les x_i étant rangés par ordre croissant. Comme $F(a) = p(\{X \leq a\})$,

- $a < x_1$, $F(a) = 0$
- $x_1 \leq a < x_2$, $F(a) = p(\{X = x_1\}) = p_1$
- $x_2 \leq a < x_3$, $F(a) = p(\{X = x_1\}) + p(\{X = x_2\}) = p_1 + p_2$

  $\vdots$
- $x_i \leq a < x_{i+1}$, $F(a) = p_1 + p_2 + \ldots + p_i$

  $\vdots$
- $a \geq x_n$, $F(a) = 1$.

On en déduit la représentation graphique ci-dessous : On peut faire les remarques suivantes :

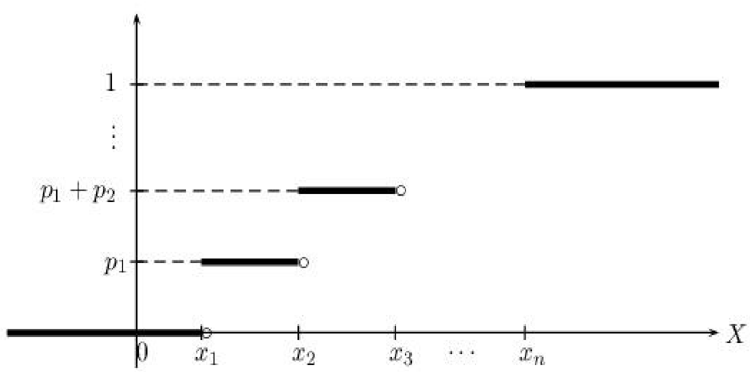

Figure 3.2

- la fonction de répartition est une fonction en escaliers (fonction constante par intervalles),
- la fonction F est croissante au sens large,
- la fonction F est continue sauf aux points x_i où elle est continue à droite : $\lim\limits_{x \overset{>}{\mapsto} x_i} F(x) = 0$,
- $\lim\limits_{x \mapsto +\infty} F(x) = 1$, $\lim\limits_{x \mapsto -\infty} F(x) = 0$,
- $p(\{a < X \leq b\}) = F(b) - F(a)$,
- $F(x_{i+1}) - F(x_i) = p_{i+1} = p(\{X = x_{i+1}\})$.

3.5 Fonction de répartition d'une variable aléatoire continue - Densité de probabilité

3.5.1 Définition

On dit que la variable aléatoire X, de fonction de répartition F, est absolument continue s'il existe une fonction f définie sur $\mathbb{R}$ et possédant les propriétés suivantes :

- f est positive ou nulle : $\forall x \in \mathbb{R},\ f(x) \geq 0$,
- f est continue sauf peut-être en un nombre fini de points,
- $\displaystyle\int_{-\infty}^{+\infty} f(t)dt = 1$,
- F est définie par $F(x) = \displaystyle\int_{-\infty}^{x} f(t)dt$ ou alors $F'(x) = f(x)$ en tout point où f est continue.

La fonction f, appelée **densité de probabilité**, définit la loi de probabilité de la variable X.

3.5.2 Interprétation graphique

Voici une représentation graphique possible de la fonction de densité f

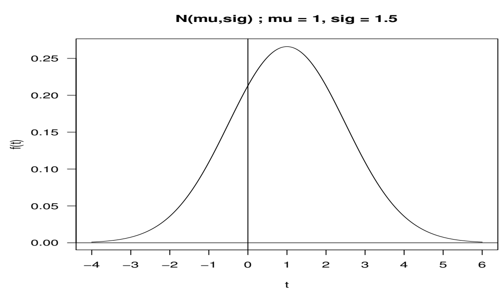

FIGURE 3.3

1. $F(a) = \displaystyle\int_{-\infty}^{a} f(t)dt$ est la surface du domaine plan limité par la courbe, la droite des abscisses et la droite d'équation $x = a$

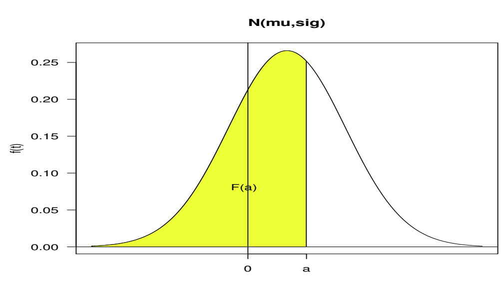

FIGURE 3.4

2. $F(b) - F(a) = \displaystyle\int_{-\infty}^{b} f(t)dt - \int_{-\infty}^{a} f(t)dt = \int_{a}^{b} f(t)dt = p(\{a \leq X \leq b\})$

   On donne le code sous R pour générer le graphique :

```
> x.c <- seq(from = -4, to = 6, length = 1000)
> plot(x.c, dnorm(x.c, 1, 1.5), main =''N(mu,sig) ; mu = 1, sig = 1.5'',
```

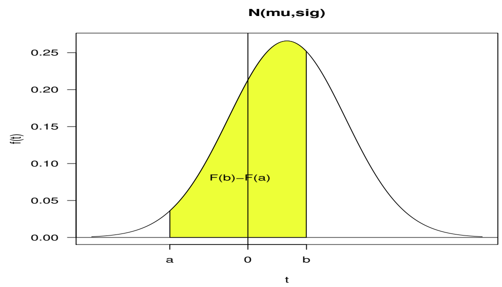

FIGURE 3.5

```
+ xlab =''t'', ylab =''f(t)'', las = 1, xaxt =''n'', type =''l'')
> polycurve <- function(x, y, base.y = min(y), ...) {
+ polygon(x = c(min(x), x, max(x)), y = c(base.y, y, base.y),
+ ...)
+ }
> bons <- which(x.c >= -2 & x.c <= 1.5)
> polycurve(x.c[bons], dnorm(x.c, mu, sd)[bons], base.y = 0, col =''yellow'')
> abline(v=0)
> abline(h=0)
> axis(1, x, x)
> text(-1.8,0.01,''a'')
> text(1.2,0.01,''b'')
> text(-0.2,0.08,''F(b)-F(a)'')
```

3. $\displaystyle\int_{-\infty}^{+\infty} f(t)dt = 1$, l'aire du domaine plan limité par la courbe et la droite des abscisses vaut 1.
4. $p(\{X = a\}) = \displaystyle\int_a^a f(t)dt = 0$. Il est très important de comprendre que cette condition n'implique pas que les nombres $p(\{a \leq X \leq b\})$ sont nuls mais seulement les formules du type $p(\{X \leq a\}) = p(\{X < a\})$ ou $p(\{X \geq a\}) = p(\{X > a\})$ ou encore $p(\{a \leq X \leq b\}) = p(\{a < X < b\})\ldots$

3.5.3 Propriétés de la fonction de répartition

- F est croissante au sens large
- $\lim\limits_{x\mapsto -\infty} F(x) = 0, \ \lim\limits_{x\mapsto +\infty} F(x) = 1$
- F est dérivable sauf aux points où f n'est pas continue et $F' = f$.

3.5.4 Exemple

On considère la densité de probabilité définie par : $f(x) = \begin{cases} \dfrac{3}{x^4} & \text{si } x \geq 1 \\ 0 & \text{si } x < 1 \end{cases}$.

- f est continue sur $\mathbb{R}/\{1\}$.
- $\forall x \in \mathbb{R}, \ f(x) \geq 0$
- $\displaystyle\int_{-\infty}^{+\infty} f(t)dt = \int_1^{+\infty} \frac{3}{t^4}dt = \left[-\frac{1}{t^3}\right]_1^{+\infty} = 1$ car $\lim\limits_{t\mapsto +\infty} \dfrac{1}{t^3} = 0$.

- Lorsque $x < 1$, $f(x) = 0 \Rightarrow F(x) = c$ mais $\lim_{x \mapsto -\infty} F(x) = 0$ donc $F(x) = 0$.

  Lorsque $x \geq 1$, $f(x) = \frac{3}{x^4} \Rightarrow F(x) = -\frac{1}{x^3} + c$ mais $\lim_{x \mapsto +\infty} F(x) = 1$ donc $F(x) = 1 - \frac{1}{x^3}$.
- $p(\{X \leq 2\}) = F(2) = 1 - \frac{1}{2^3} = 1 - \frac{1}{8} = \frac{7}{8}$,

  $p(\{2 < X \leq 3\}) = F(3) - F(2) = (1 - \frac{1}{3^3}) - (1 - \frac{1}{2^3}) = \frac{1}{2^3} - \frac{1}{3^3} = \frac{19}{216}$.

3.6 Moments d'une variable aléatoire

3.6.1 Espérance mathématique

L'**espérance mathématique** (ou **moment d'ordre** 1, noté m_1) d'une variable aléatoire réelle X est notée $E(X)$ avec

- si X est discrète, $E(X) = \sum_i p_i x_i$. L'espérance est dans ce cas la moyenne pondérée par les probabilités des différentes valeurs possibles que peut prendre la variable X, elle représente la valeur moyenne de la variable aléatoire X.
- si X est absolument continue, $E(X) = \int_{-\infty}^{+\infty} t f(t) dt$.

Exemple 3.6.1 1. Si X est discrète, on peut reprendre l'exemple de la section 3.4.2. avec les deux boîtes contenant 4 et 3 boules. On a les résultats suivants :

x_i	0	1	2	3	Total
p_i	$\frac{3}{12}$	$\frac{5}{12}$	$\frac{3}{12}$	$\frac{1}{12}$	1
$p_i x_i$	0	$\frac{5}{12}$	$\frac{6}{12}$	$\frac{3}{12}$	$\frac{14}{12}$

On en déduit que $E(X) = \sum_i p_i x_i = \frac{14}{12} = \frac{7}{6}$

2. Si X est absolument continue, on peut considérer l'exemple de la section 3.5.4. :

$$f(x) = \begin{cases} \frac{3}{x^4} & \text{si } x \geq 1 \\ 0 & \text{si } x < 1 \end{cases}$$

alors,

$$E(X) = \int_{-\infty}^{+\infty} t f(t) dt = \int_1^{+\infty} \frac{3}{t^3} dt = \left[-\frac{3}{2t^2} \right]_1^{+\infty}.$$

Or $\lim_{t \mapsto +\infty} \frac{1}{t^2} = 0$ donc $E(X) = 0 - \left(-\frac{3}{2} \right) = \frac{3}{2}$.

L'espérance mathématique a les propriétés suivantes :

- Si X est une variable aléatoire constante ($X = k$) de probabilité 1 alors $E(X) = k$.
- Si X et Y sont deux variables aléatoires et a et b deux réels

$$\boxed{E(aX + bY) = aE(X) + bE(Y)}$$

En particulier,

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

$$E(aX) = aE(X),$$
$$E(X+Y) = E(X) + E(Y).$$

On peut généraliser ces résultats à n variables aléatoires $X_1, \ldots, X_n$. Soient $a_1 \ldots, a_n$ n réels alors

$$\boxed{E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i E(X_i)}$$

<u>*Preuve*</u> : Dans le cas de deux variables aléatoires absolument continues dont les densités de probabilité sont f_1 et f_2 :

$$\begin{aligned} E(a_1X_1 + a_2X_2) &= \int_{-\infty}^{+\infty} t(a_1f_1 + a_2f_2)(t)dt = a_1 \int_{-\infty}^{+\infty} tf_1(t)dt + a_2 \int_{-\infty}^{+\infty} tf_2(t)dt \\ &= a_1E(X_1) + a_2E(X_2). \end{aligned}$$

La démonstration est similaire lorsque les variables X_1 et X_2 sont discrètes.
Le cas à n variables se traite de la même façon. ■

3.6.2 Variable centrée

Soit X une variable aléatoire réelle d'espérance mathématique $E(X)$. La variable

$$Z = X - E(X)$$

est la **variable aléatoire centrée** associée à X. On peut alors vérifier que $\boxed{E(Z) = 0}$.

<u>*Preuve*</u> : $E(Z) = E(X - E(X)) = E(X) - E(X) = 0$ car $E(X)$ est une constante et si $a \in \mathbb{R}$, $E(X + a) = E(X) + a$. ■

3.6.3 Moments d'ordre k

Soit X une variable aléatoire réelle. On nomme **moment d'ordre** k ($k \in \mathbb{N}^\star$) l'espérance mathématique de la variable X^k. On le note

$$\boxed{m_k = E(X^k)}$$

- Dans le cas d'une variable discrète, $m_k = \displaystyle\sum_i p_i x_i^k$.
- Dans le cas d'une variable aléatoire absolument continue de densité de probabilité f, $m_k = \displaystyle\int_{-\infty}^{+\infty} t^k f(t)dt$.

Remarque 3.6.1 On a en particulier $m_1 = E(X)$.

3.6.4 Moments centrés d'ordre k

On appelle **moments centrés d'ordre** k de la variable aléatoire réelle X, que l'on note δ_k, les moments d'ordre k de la variable centrée $Z = X - E(X)$.

- Dans le cas d'une variable discrète, $\delta_k = E(X - E(X))^k = \displaystyle\sum_i p_i (x_i - E(X))^k$.
- Dans le cas d'une variable absolument continue, $\delta_k = \displaystyle\int_{-\infty}^{+\infty} [t - E(X)]^k f(t)dt$.

Remarque 3.6.2 On a en particulier $\delta_1 = 0$.

3.6.5 Variance et écart-type

1. On appelle **variance** de la variable aléatoire X le moment centré d'ordre 2 de cette variable.

$$\boxed{V(X) = \delta_2 = E(X - E(X))^2}$$

On a la définition équivalente suivante :

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

Preuve : $E(X - E(X))^2 = E(X^2 - 2XE(X) + (E(X))^2)$ or $E(X)$ est une constante donc $V(X) = E(X^2) - 2E(X)E(X) + (E(X))^2 = E(X^2) - E(X)^2$. ■

– Si X est une variable discrète (admettant une espérance et telle que la série $\sum_i p_i x_i^2$ est convergente),

$$V(X) = \sum_i p_i(x_i - E(X))^2 = \left(\sum_i p_i x_i^2\right) - (E(X))^2.$$

– Si X est absolument continue, $V(X) = \int_{-\infty}^{+\infty} (t - E(X))^2 f(t)dt = \left(\int_{-\infty}^{+\infty} t^2 f(t)dt\right) - (E(X))^2$.

Exemple 3.6.2 (a) Si X est discrète, on peut reprendre l'exemple de la section 3.4.2. avec les deux boîtes contenant 4 et 3 boules. On a les résultats suivants :

x_i	0	1	2	3	Total
p_i	$\frac{3}{12}$	$\frac{5}{12}$	$\frac{3}{12}$	$\frac{1}{12}$	1
$p_i x_i$	0	$\frac{5}{12}$	$\frac{6}{12}$	$\frac{3}{12}$	$E(X) = \frac{14}{12} = \frac{7}{6}$
$p_i x_i^2$	0	$\frac{5}{12}$	$\frac{12}{12}$	$\frac{9}{12}$	$E(X^2) = \frac{26}{12} = \frac{13}{6}$

On trouve ainsi $V(X) = \frac{13}{6} - \left(\frac{7}{6}\right)^2 = \frac{29}{36}$.

(b) Si X est absolument continue, on peut considérer l'exemple de la section 5.4. :

$$f(x) = \begin{cases} \frac{3}{x^4} & \text{si } x \geq 1 \\ 0 & \text{si } x < 1 \end{cases}.$$

On a $E(X) = \frac{3}{2}$. Par conséquent, $V(X) = \int_1^{+\infty} \frac{3}{t^4} t^2 dt - \left(\frac{3}{2}\right)^2 = \left[-\frac{3}{t}\right]_1^{+\infty} - \frac{9}{4}$ or $\lim_{t \mapsto +\infty} \frac{3}{t} = 0$ donc $V(X) = 3 - \frac{9}{4} = \frac{3}{4}$.

2. La variance a les propriétés suivantes :
   – Si X est une variable aléatoire constante $(X = k)$ de probabilité 1 alors $V(X) = 0$.
   – $$\boxed{V(aX) = a^2 V(X)}$$
   Preuve : $V(aX) = E(aX - (E(aX)))^2 = E(aX - a(E(X)))^2 = a^2 E(X - E(X))^2 = a^2 V(X)$ ■
   – $$\boxed{V(aX + b) = a^2 V(X)}$$
   Preuve : Démonstration identique ■

La variance n'a pas de propriété de linéarité et en général, $V(X + Y) \neq V(X) + V(Y)$.

3. On a vu que si la variable aléatoire X est multipliée par un réel a, sa variance est multipliée par a^2 alors que son espérance est multipliée par a. Pour cette raison, lorsque l'on veut mesurer de façon homogène la moyenne et la dispersion, on utilise plutôt que la variance la notion d'écart-type.

L'**écart-type** de la variable X, noté σ_X, est la racine carrée de la variance de X :

$$\boxed{\sigma_X = \sqrt{V(X)} = \sigma(X)}$$

Remarque 3.6.3 En utilisant le dernier résultat sur la variance, on a l'égalité $\sigma(aX+b) = |a|\sigma(X)$.

4. Soit X une variable aléatoire, d'espérance mathématique $E(X)$ et d'écart-type $\sigma(X)$. On appelle variable centrée réduite la variable

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

Cette variable vérifie les propriétés $\boxed{E(X^\star) = 0}$ et $\boxed{\sigma(X^\star) = 1}$

<u>*Preuve*</u> :

- $E(X^\star) = E\left(\dfrac{X - E(X)}{\sigma(X)}\right) = \dfrac{1}{\sigma(X)} E(X - E(X)) = \dfrac{1}{\sigma(X)}(E(X) - E(X)) = 0,$
- $V(X^\star) = \dfrac{1}{\sigma^2(X)} V(X - E(X)) = \dfrac{1}{\sigma^2(X)} V(X) = \dfrac{\sigma^2(X)}{\sigma^2(X)} = 1.$ ■

3.6.6 Moments factoriels

On appelle **moment factoriel** d'ordre $k \geq 1$, noté $\Gamma_{[k]}$, l'espérance mathématique de la variable $X(X-1)\dots(X-k+1)$ soit

$$\boxed{\Gamma_{[k]} = E[X(X-1)\dots(X-k+1)]}$$

Remarque 3.6.4 On peut facilement vérifier que $E(X) = m_1 = \Gamma_{[1]}$ et $V(X) = m_2 - (m_1)^2 = \Gamma_{[2]}$.

3.7 Exercices

Exercice 35 Dans une grande surface, on a relevé sur une longue période le nombre d'articles de type A vendus.
L'étude statistique permet d'admettre que la variable aléatoire X qui associe à un jour ouvrable choisi au hasard pendant un mois le nombre d'articles de type A vendus ce jour là a une probabilité définie par le tableau suivant.
On donnera les valeurs approchées arrondies à 10^{-2} près des résultats.

Nombre x_i de pièces vendues	0	1	2	3	4	5	6
$p(\{X = x_i\})$	0,10	0,16	0,25	0,30	0,13	0,05	0,01

1. Représentez graphiquement la fonction de répartition de la variable aléatoire X.
2. Calculez l'espérance mathématique $E(X)$ de la variable aléatoire X. Que représente $E(X)$?
3. Calculez la variance et l'écart-type de la variable aléatoire X

Exercice 36 Une partie de loterie consiste à lâcher une bille dans un appareil qui comporte six portes de sortie, numérotées de 1 à 6.
Soit X la variable aléatoire qui à chaque partie associe le numéro de la porte de sortie franchie. Sa loi de probabilité est définie par le tableau suivant :

i	1	2	3	4	5	6
$p(\{X = x_i\})$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

La règle du jeu est la suivante : un joueur mise 2 euros : il reçoit 12 euros si la bille franchit les portes 1 ou 6, 2 euros si elle franchit les portes 3 ou 4. Les portes 2 et 5 ne rapportent rien.
Le gain d'un joueur est la différence entre ce qu'il reçoit à l'issue de la partie et sa mise. Le gain peut donc être éventuellement un nombre négatif ou nul.
Soit Y la variable aléatoire qui à chaque partie effectuée par un joueur donné associe le gain.

1. Quelles sont les valeurs possibles de Y ?
2. Déterminer la loi de probabilité de Y ?
3. Un jeu est équitable si l'espérance mathématique du gain est nulle. Ce jeu est-il équitable ?

Exercice 37 Dans cet exercice, les tirages sont équiprobables.
Un sac contient quatre jetons noirs et quatre jetons blancs. On tire quatre jetons du sac simultanément.
Soit X la variable aléatoire qui à chaque tirage associe le nombre de jetons noirs tirés.
Déterminez la loi de probabilité de X, son espérance mathématique et son écart-type.

Exercice 38

1. Calculez l'espérance mathématique, la variance et l'écart-type de la variable aléatoire X dont la loi de probabilité est donnée par le tableau :

x_i	2	5	12
$p(\{X = x_i\})$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$

2. Construisez la représentation graphique de la fonction de répartition de la variable aléatoire X.

Exercice 39 Soit X la variable aléatoire qui associe à un mois choisi au hasard le découvert des comptes en banque d'une entreprise. Une étude statistique permet d'admettre que la loi de probabilité de X est donnée par le tableau suivant.

x_i	200000	150000	120000	90000	70000
$p(\{X = x_i\})$	$\frac{1}{10}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{3}$	$\frac{1}{5}$

Calculez l'espérance mathématique $E(X)$ de la variable aléatoire X. Que représente $E(X)$?

Exercice 40 On lance trois pièces de monnaie (non truquées), une de 50 centimes d'euros, une de 1 euro, une de 5 euros.
Un résultat est donné sous la forme d'un triplet, par exemple (P, P, F) où le premier élément est le résultat pour la pièce de 50 centimes d'euros, le deuxième le résultat pour la pièce d'un euro et le troisième pour la pièce de 5 euros.

1. Déterminez l'ensemble des résultats possibles.
2. On gagne 10 euros si on obtient 3 fois face, 2 euros si on obtient 2 fois face et on perd 2 euros si on obtient une seule fois face ou aucune fois face. On note X la variable aléatoire qui à tout lancer des 3 pièces associe le gain obtenu, une perte étant considérée comme un gain négatif.
   (a) Définissez la loi de probabilité de X en présentant les résultats à l'aide d'un tableau.

(b) Calculez l'espérance mathématique $E(X)$ de la variable aléatoire X. Que représente $E(X)$? Le jeu est-il équitable ?

Remarque 3.7.1 Un jeu est équitable lorsque les gains et les pertes s'équilibrent sur un très grand nombre de parties.

Exercice 41 Dans une urne se trouvent 10 boules numérotées de 0 à 9 identiques au toucher. On tire simultanément 2 boules dans l'urne sans remise et on considère l'aléa numérique X qui à chaque tirage associe le plus petit des 2 nombres portés par les 2 boules.

1. Déterminez la distribution de probabilité de X.
2. Construisez la représentation graphique de la fonction de répartition de X dans un repère $(O,\vec{\imath},\vec{\jmath})$.
3. Calculez les probabilités des événements suivants :
   - $A = \{X < 5\}$,
   - $B = \{X \leq 3\}$,
   - $C = \{1 \leq X \leq 6\}$,
   - $D = \{X \geq 7\}$,
   - $E = \{|X - 5| > 2\}$,
   - $G = \{X^2 - 5X + 4 < 0\}$.

Exercice 42 Un dé cubique non truqué admet 3 faces portant un "1", 2 faces portant un "2" et une face portant un "3". On le lance deux fois de suite et on considère la variable aléatoire X associant à chaque tirage la somme des points lus sur la face supérieure du dé obtenue à chacun des deux lancers. Donnez la distribution de probabilité de X.

Exercice 43 Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{P}(\Omega), p)$ avec la distribution de probabilité suivante :

$$\left(\left(-6,\frac{2}{10}\right),\left(-3,\frac{3}{10}\right),\left(0,\frac{2}{10}\right),\left(3,\frac{1}{10}\right),\left(4,\frac{2}{10}\right)\right).$$

Calculez $E(X)$, $V(X)$ et $\sigma(X)$, en déduire la variable centrée réduite X' associée à X.

Exercice 44 Une urne contient 5 boules dont 2 blanches et 3 noires. On tire une boule de l'urne, sans la remettre, 3 fois de suite. Calculez l'espérance mathématique, la variance et l'écart-type de la variable aléatoire X correspondant au nombre de boules blanches tirées.

Exercice 45 On considère une expérience aléatoire consistant à jeter une boule sur la roulette d'un casino. On peut admettre que la boule se déplace sur un cercle de centre O et s'immobilise en un point M. Si A est un point du cercle, la mesure en radians de l'angle $(\widehat{\vec{OA},\vec{OM}})$ est une variable aléatoire prenant ses valeurs dans l'intervalle $[0; 2\pi[$. Si on suppose que la boule a "autant de chances" de s'arrêter sur un point que sur un autre, on montre que X est une variable aléatoire continue à densité et que la densité de X est une fonction f définie sur $\mathbb{R}$ par :

$$f(x) = \begin{cases} \dfrac{1}{2\pi} & \text{si } x \in [0; 2\pi[\\ 0 & \text{sinon} \end{cases}.$$

1. Déterminez la fonction de répartition de X.
2. Calculez les probabilités pour que X soit compris entre 0 et $\dfrac{\pi}{2}$ et entre 0 et π.

Exercice 46 On considère la fonction f définie par : $f(x) = \begin{cases} 0 & \text{si } x < -1 \text{ ou } x > 0 \\ x+1 & \text{si } x \in [-1;0] \\ -x+1 & \text{si } x \in [0;1] \end{cases}$.

1. Montrez que f est une densité de probabilité d'une variable aléatoire X.
2. Déterminez la fonction de répartition de X, on la note F.
3. Écrivez en fonction de F puis calculer $p(\{X < -0,5\}), p(\{-0,5 \leq X \leq 0,5\}), p(\{X > 0,25\})$.
4. Montrez que X admet une espérance et une variance que l'on déterminera.