

# Chapitre 1

## Variables aléatoires

### 1.1 Bases des probabilités

Dans ce module, nous rappelons très rapidement les fondements de la théorie des probabilités.

#### 1.1.1 Axiomes de probabilité

La théorie des probabilité est l'ensemble des techniques mathématiques décrivant les résultats d'expériences dont l'issue n'est pas connue à l'avance. Soient  $\zeta_i$  le résultat de la  $i$ ème expérience,  $A, B, \dots = \{ \text{quelques } \zeta_i \}$  un évènement,  $\Omega = \{ \text{tous les } \zeta_i \}$  l'évènement certain et  $\emptyset$  l'évènement impossible. La probabilité de l'évènement  $A$  est le réel  $P(A)$  satisfaisant aux trois axiomes

A1.  $P(A) \geq 0$

A2.  $P(\Omega) = 1$

A3. Dans le cas où  $\Omega$  est un ensemble fini, si  $A \cap B = \emptyset$  alors  $P(A \cup B) = P(A) + P(B)$ .

A3'. Dans le cas où  $\Omega$  est un ensemble infini, si  $A_1, A_2, \dots$  est une suite d'évènements telle que  $A_i \cap A_j = \emptyset$  pour  $i \neq j$  alors  $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ .

Ces axiomes entraînent les corollaires suivants :

P1.  $P(\bar{A}) = 1 - P(A)$  où  $\bar{A} = \Omega \setminus A$ .

P2.  $P(A) \leq 1$ .

P3.  $P(\emptyset) = 0$ .

P4.  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ .

P5. Si  $A \subseteq B$ ,  $P(A) \leq P(B)$ .

### 1.1.2 Probabilité conditionnelle

La probabilité de l'évènement  $A$  étant donné que l'évènement  $B$  s'est déroulé est définie par

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

si  $P(B) \neq 0$  et par  $P(A|B) = 0$  si  $P(B) = 0$ . Elle possède deux propriétés importantes. Soit  $[A_1, A_2, \dots, A_n]$  une partition de  $\Omega$ , c'est-à-dire un ensemble d'évènements tels que

$$A_1 \cup \dots \cup A_n = \Omega \quad \text{et} \quad A_i \cap A_j = \emptyset \quad \text{pour} \quad i \neq j.$$

Alors

P6. (Théorème des probabilités totales)

$$P(B) = \sum_{i=1}^n P(B|A_i)P(A_i).$$

P7. (Règle de Bayes)

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^n P(B|A_j)P(A_j)}.$$

### 1.1.3 Indépendance

Deux évènements  $A$  et  $B$  sont *indépendants* lorsque

$$P(A \cap B) = P(A)P(B).$$

Trois évènements  $A$ ,  $B$  et  $C$  sont indépendants si et seulement si

$$P(A \cap B) = P(A)P(B)$$

$$P(A \cap C) = P(A)P(C)$$

$$P(B \cap C) = P(B)P(C)$$

et

$$P(A \cap B \cap C) = P(A)P(B)P(C).$$

On généralise aisément cette définition à un nombre quelconque d'évènements indépendants.

## 1.2 Variable aléatoire

Une *variable aléatoire* (v.a.) est une fonction qui assigne à chaque résultat  $\zeta$  d'une expérience un réel  $X(\zeta)$  tel que

C1. l'ensemble  $\{\zeta | X(\zeta) \leq x\}$  est un évènement pour tout  $x \in \mathbb{R}$  ;

C2. les probabilités des évènements  $\{\zeta | X(\zeta) = \infty\}$  et  $\{\zeta | X(\zeta) = -\infty\}$  est nulle, i.e.  $P(X = \infty) = P(X = -\infty) = 0$ .

### 1.2.1 Fonction de répartition

Une v.a. est caractérisée par sa *fonction de répartition*  $F_X(x)$  qui est la probabilité que  $X$  prenne une valeur inférieure ou égale au réel  $x$  :

$$F_X(x) = P(X \leq x) = P(A) \quad (1.1)$$

où  $A$  est l'événement  $A = \{\zeta \in \Omega | X(\zeta) \leq x\}$ . Elle possède les propriétés suivantes :

P8.  $0 \leq F_X(x) \leq 1$ .

P9.  $\lim_{x \rightarrow -\infty} F_X(x) = 0$  et  $\lim_{x \rightarrow \infty} F_X(x) = 1$ .

P10. Si  $a < b$ ,  $F_X(a) \leq F_X(b)$ .

P11.  $P(a < X \leq b) = F_X(b) - F_X(a)$

P12.  $F_X(x)$  est continue à droite, i.e.  $F_X(x) = \lim_{\varepsilon \rightarrow 0, \varepsilon > 0} F_X(x + \varepsilon) = F_X(x^+)$ .

Nous désignerons par  $S_X$  l'ensemble des valeurs que peut prendre la variable aléatoire  $X$  :  $S_X = \{X(\zeta) | \zeta \in \Omega\}$ . Si  $X$  prend un ensemble dénombrable de valeurs,  $S_X = \{x_1, x_2, x_3, \dots\}$ , alors  $X$  est appelée *variable aléatoire discrète*. Sa fonction de répartition  $F(x)$  est une fonction en escalier, présentant une discontinuité à gauche en chacune des valeurs  $x_i$  :

$$p_i = P(X = x_i) = F_X(x_i) - \lim_{\varepsilon \rightarrow 0, \varepsilon > 0} F_X(x_i - \varepsilon) = F_X(x_i) - F_X(x_i^-).$$

Si  $F_X(x)$  est continue, on parlera d'une *variable aléatoire continue*. Dans ce cas, la probabilité que  $X$  prenne exactement une valeur  $x$  particulière est nulle :  $P(X = x) = 0$ . Enfin, dans les autres cas,  $F(x)$  est discontinue bien que n'étant pas une fonction en escalier. On a alors une *variable aléatoire mixte*. Ces définitions peuvent être généralisées pour des v.a. complexes.

### 1.2.2 Densité de probabilité

La fonction de *densité de probabilité* est définie par

$$f_X(x) = \frac{dF_X(x)}{dx}. \quad (1.2)$$

Elle possède les propriétés suivantes :

P13.  $f_X(x) \geq 0$ .

P14.  $\int_{-\infty}^{\infty} f_X(x) dx = 1$ .

P15.  $P(X \leq a) = F_X(a) = \int_{-\infty}^a f_X(x) dx$ .

P16.  $P(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(x) dx$ . Cette propriété entraîne que  $P(x < X \leq x + \Delta x) = f_X(x) \Delta x$  lorsque  $\Delta x \rightarrow 0$ .

Dans le cas d'une v.a. discrète, l'utilisation des impulsions de Dirac  $\delta(\cdot)$  permet d'écrire (1.2) sous la forme

$$f_X(x) = \sum_i p_i \delta(x - x_i)$$

où  $p_i = P(X = x_i)$ . Les propriétés P13 à P16 deviennent alors

P13'.  $p_i \geq 0$ .

P14'.  $\sum_i p_i = 1$ .

P15'.  $P(X \leq a) = F_X(a) = \sum_{x_i \leq a} p_i$ .

P16'.  $P(X = x_i) = F_X(x_i) - F_X(x_i^-) = p_i$

### 1.2.3 Fonction d'une variable aléatoire

Soit  $g(\cdot)$  une fonction continue, comme par exemple celle représentée à la figure 1.1. Si  $X$  est une v.a. continue, l'expression

$$Y = g(X)$$

désigne également une v.a. continue (Pour être tout à fait rigoureux sur le plan mathématique, il faut vérifier que la fonction  $g(x)$  soit telle que  $Y$  satisfasse aux deux conditions C1 et C2. Ce sera toujours le cas dans la suite). Soit  $y$  une valeur particulière que peut prendre la v.a.  $Y$ , et soient  $x_1, x_2, \dots, x_m$  et les  $m$  racines respectives de l'équation  $y = g(x)$ , c'est-à-dire  $y = g(x_i)$  pour  $1 \leq i \leq m$ . Nous allons traiter l'exemple de la fonction  $g(x)$  tel qu'il est représenté à la figure 1.1, où  $m = 3$ .

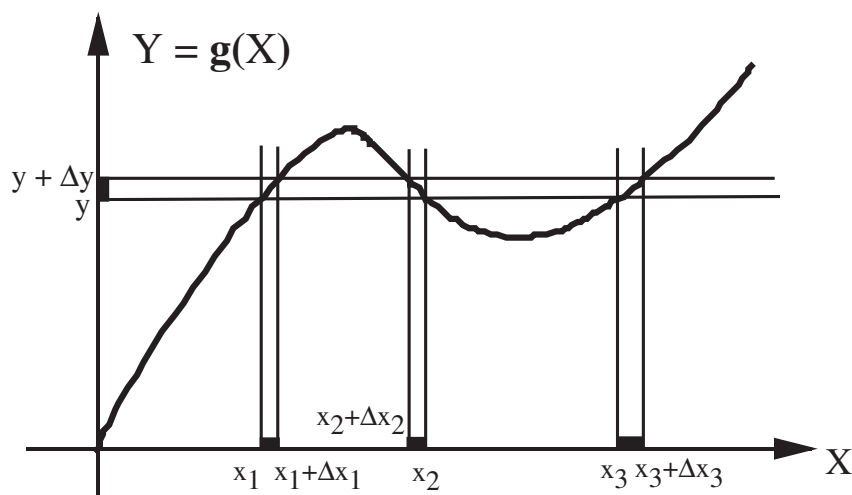


FIGURE 1.1 – Fonction  $g(X)$  d'une variable aléatoire  $X$ .

Soit  $A$  l'événement

$$A = \{\zeta \mid y < Y(\zeta) \leq y + \Delta y\}.$$

D'après la propriété P16, la probabilité que cet évènement  $A$  se produise, c'est-à-dire que  $Y$  soit comprise entre  $y$  et  $y + \Delta y$  est donnée par

$$P(A) = P(y < Y \leq y + \Delta y) = f_Y(y)|\Delta y| \quad (1.3)$$

si  $\Delta y \rightarrow 0$ . D'autre part, l'évènement  $A$  peut s'exprimer comme

$$A = \{\zeta \mid y < g(X(\zeta)) \leq y + \Delta y\}$$

ou encore, puisque  $g(x)$  est continue, par

$$A = \{\zeta \mid x_1 < X(\zeta) \leq x_1 + \Delta x_1\} \cup \{\zeta \mid x_2 + \Delta x_2 < X(\zeta) \leq x_2\} \cup \{\zeta \mid x_3 < X(\zeta) \leq x_3 + \Delta x_3\}$$

d'où,  $\Delta x_i$  tendant vers 0 quand  $\Delta y \rightarrow 0$ ,

$$P(A) = f_X(x_1)|\Delta x_1| + f_X(x_2)|\Delta x_2| + f_X(x_3)|\Delta x_3|. \quad (1.4)$$

Comme

$$g'(x_i) = \lim_{\Delta x_i \rightarrow 0} \frac{\Delta y}{\Delta x_i},$$

en égalant les valeurs (1.3) et (1.4) de  $P(A)$ , on trouve

$$f_Y(y) = \sum_i \frac{f_X(x_i)}{|g'(x_i)|}. \quad (1.5)$$

Cette formule est tout à fait générale, et s'applique à tous les cas où  $X$  est continue.

#### 1.2.4 Espérance d'une fonction d'une variable aléatoire

L'*espérance* de la fonction  $g(X)$ , où  $X$  est une v.a. de densité de probabilité  $f_X(x)$ , est définie par

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x)dx. \quad (1.6)$$

pour autant que cette intégrale converge absolument. Plusieurs définitions importantes se ramènent à un cas particulier de la définition donnée ci-dessus. Par exemple, si  $g(X) = X^n$ , l'espérance porte alors le nom de *moment d'ordre  $n$*  :

$$E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x)dx. \quad (1.7)$$

Le moment d'ordre 1 est encore appelé *moyenne* de la v.a.  $X$  et est noté

$$\mu_X = E[X] = \int_{-\infty}^{\infty} x f_X(x)dx. \quad (1.8)$$

La v.a.  $Y = X - \mu_X$  est une *variable aléatoire centrée* vu que  $E[Y] = 0$ . La *variance* de la v.a.  $X$  est le moment d'ordre 2 de la v.a  $Y = X - \mu_X$ , et l'*écart-type*  $\sigma_X$  est la racine carrée de la variance :

$$\begin{aligned} VAR[X] = \sigma_X^2 &= E[(X - \mu_X)^2] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx \\ &= E[X^2] - \mu_X^2 = \int_{-\infty}^{\infty} x^2 f_X(x) dx - \mu_X^2. \end{aligned} \quad (1.9)$$

Dans le cas discret, ces deux grandeurs s'écrivent respectivement, avec  $P(X = x_i) = p_i$ ,

$$\mu_X = E[X] = \sum_i p_i x_i$$

et

$$VAR[X] = \sigma_X^2 = E[(X - \mu_X)^2] = \sum_i p_i (x_i - \mu_X)^2.$$

Les moments peuvent ne pas exister pour certains v.a. Par exemple, la densité de probabilité de la v.a. de Cauchy est

$$f_X(x) = \frac{1}{\pi(1+x^2)}.$$

On peut vérifier qu'il s'agit effectivement d'une densité de probabilité car

$$\int_{-\infty}^{\infty} f_X(x) dx = \int_{-\infty}^{\infty} \frac{dx}{\pi(1+x^2)} = \frac{1}{\pi} [\text{Arctg}(x)]_{-\infty}^{\infty} = \frac{1}{\pi} \left[ \frac{\pi}{2} - \frac{-\pi}{2} \right] = 1.$$

Par contre, comme

$$\int_{-\infty}^{\infty} |x f_X(x)| dx = \int_{-\infty}^{\infty} \left| \frac{x}{\pi(1+x^2)} \right| dx = 2 \int_0^{\infty} \frac{x}{\pi(1+x^2)} dx = \frac{1}{\pi} \left[ \ln(1+x^2) \right]_0^{\infty} = \infty$$

la fonction  $x f_X(x)$  n'est pas absolument intégrable, et  $E[X]$  n'est pas défini.

### 1.2.5 Transformées

L'opérateur d'espérance mathématique d'une fonction de v.a., donné par (1.6), appliqué à la v.a. complexe  $\exp(j\omega X)$ , où  $j$  est la racine carrée de -1, définit la *fonction caractéristique*  $\Phi_X(\omega)$  de la v.a.  $X$  :

$$\Phi_X(\omega) = E[e^{j\omega X}] = \int_{-\infty}^{\infty} e^{j\omega x} f_X(x) dx. \quad (1.10)$$

On remarque que cette fonction est le complexe conjugué de la transformée de Fourier de la fonction  $f_X(x)$ . Elle existe toujours car

$$\Phi_X(\omega) \leq |\Phi_X(\omega)| = \left| \int_{-\infty}^{\infty} e^{j\omega x} f_X(x) dx \right| \leq \int_{-\infty}^{\infty} |e^{j\omega x} f_X(x)| dx = \int_{-\infty}^{\infty} f_X(x) dx = 1$$

et est maximale en  $\omega = 0$  car  $\Phi_X(0) = 1$ . La transformée inverse permet d'obtenir la densité de probabilité à partir de la fonction caractéristique par le fait que

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-j\omega x} \Phi_X(\omega) d\omega. \quad (1.11)$$

Les deux fonctions  $f_X(x)$  et  $\Phi_X(\omega)$  peuvent donc être utilisées de manière équivalente. Si deux v.a. continues ont la même fonction caractéristique, elles auront la même densité de probabilité, et vice-versa. L'intérêt de la fonction caractéristique apparaît dans certains calculs, notamment celui des moments de  $X$ . En effet, en dérivant  $\Phi_X(\omega)$  par rapport à  $\omega$  et en évaluant cette dérivée en  $\omega = 0$ , (1.10) entraîne que

$$\left. \frac{d\Phi_X(\omega)}{d\omega} \right|_{\omega=0} = \int_{-\infty}^{\infty} jxe^{j\omega x} f_X(x) dx \Big|_{\omega=0} = j \int_{-\infty}^{\infty} x f_X(x) dx = jE[X], \quad (1.12)$$

d'où

$$E[X] = \frac{1}{j} \left. \frac{d\Phi_X(\omega)}{d\omega} \right|_{\omega=0}.$$

De même, en dérivant (1.12) par rapport à  $\omega$  et en évaluant cette deuxième dérivée en  $\omega = 0$ , on trouve

$$E[X^k] = \frac{1}{j^k} \left. \frac{d^k \Phi_X(\omega)}{d\omega^k} \right|_{\omega=0}. \quad (1.13)$$

De la même manière qu'on peut étendre la transformée de Fourier à la transformée de Laplace en prenant un complexe  $s$  au lieu du nombre imaginaire  $j\omega$ , on étend la définition de fonction caractéristique à celle de *fonction génératrice des moments*

$$\hat{\Phi}_X(s) = E[e^{sX}] = \int_{-\infty}^{\infty} e^{sx} f_X(x) dx. \quad (1.14)$$

On utilise aussi parfois dans certains développements la *fonction génératrice de cumulants*, encore appelée *seconde fonction caractéristique*,

$$\Psi_X(\omega) = \ln \Phi_X(\omega).$$

Enfin, dans le cas d'une v.a. discrète  $X$  qui prend des valeurs uniformément espacées  $S_X = \{0, 1, 2, 3, \dots\} = \mathbb{N}$ , on peut définir une *fonction génératrice de probabilité*, qui est l'espérance de la fonction  $z^X$ , i.e.

$$G_X(z) = E[z^X] = \sum_{k=0}^{\infty} z^k P(X = k) = \sum_{k=0}^{\infty} z^k p_k \quad (1.15)$$

et est, à nouveau au signe de l'exposant près, la transformée en  $z$  de la fonction (discrète)  $p_k$ . On retrouve les probabilités  $p_k$  en dérivant  $k$  fois  $G_X(z)$  par rapport à  $z$ , et en évaluant cette dérivée en  $z = 0$  :

$$p_k = P(X = k) = \frac{1}{k!} \left. \frac{d^k G_X(z)}{dz^k} \right|_{z=0}. \quad (1.16)$$

Par contre, en dérivant  $k$  fois  $G_X(z)$  par rapport à  $z$  mais en évaluant cette dérivée à présent en  $z = 1$ , on trouve les moments factoriaux de  $X$ , à partir desquels on peut retrouver tous les moments de  $X$  :

$$E[X(X-1)(X-2)\dots(X-k+1)] = \left. \frac{d^k G_X(z)}{dz^k} \right|_{z=1}. \quad (1.17)$$

Remarquons finalement que

$$G_X(1) = \sum_{k=0}^{\infty} p_k = 1.$$

### 1.2.6 Inégalités

Très souvent en pratique, on ne connaît pas complètement en pratique la densité de probabilité  $f_X(x)$  d'une v.a. aléatoire  $X$ . Il est donc très utile de pouvoir borner une probabilité ou l'espérance d'une fonction de cette v.a.

#### Inégalité de Markov

Une première borne est obtenue dans le cas d'une v.a.  $X$  qui ne prend pas de valeurs négatives ( $X \geq 0$ ) en constatant que, si  $a > 0$ ,

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^a x f_X(x) dx + \int_a^{\infty} x f_X(x) dx \\ &\geq \int_a^{\infty} x f_X(x) dx \geq \int_a^{\infty} a f_X(x) dx = a \int_a^{\infty} f_X(x) dx = a P(X \geq a), \end{aligned}$$

d'où

$$P(X \geq a) \leq E[X]/a \tag{1.18}$$

avec  $X \geq 0$  et  $a > 0$ . Cette inégalité porte le nom d'*inégalité de Markov*.

#### Inégalité de Tchebytcheff

A présent, considérons la v.a.  $D = (X - \mu_X)^2$ , dont la moyenne est  $E[D] = E[(X - \mu_X)^2] = \sigma_X^2$ , c'est-à-dire la variance de la v.a.  $X$ . Posons  $b = \sqrt{a}$  dans l'inégalité de Markov, qui devient donc

$$P(D \geq b^2) \leq \sigma_X^2/b^2.$$

Comme  $P(D \geq b^2) = P((X - \mu_X)^2 \geq b^2) = P(|X - \mu_X| \geq b)$ , cette inégalité devient

$$P(|X - \mu_X| \geq b) \leq \sigma_X^2/b^2 \tag{1.19}$$

et est connue sous le nom d'*inégalité de Tchebytcheff*.

#### Inégalité de Jensen

Une troisième inégalité permet de minorer l'espérance d'une fonction convexe et continue d'une v.a. par la fonction de l'espérance de cette v.a. Rappelons qu'une fonction  $g : I \subseteq \mathbb{R} \rightarrow \mathbb{R} \cup \{\infty\}$  est convexe si pour tout  $x, x' \in I$ , et pour tout  $0 \leq t \leq 1$ ,

$$g(tx + (1-t)x') \leq tg(x) + (1-t)g(x').$$

Une fonction  $g$  convexe et continue a la propriété d'être toujours minorée par une fonction linéaire passant par un point du graphe de la fonction  $g$  : pour tout  $a \in \mathbb{R}$ , il existe  $\theta \in \mathbb{R}$  (pouvant donc dépendre de  $a$ ) tel que pour tout  $x \in \mathbb{R}$

$$g(x) \geq g(a) + \theta(x - a). \tag{1.20}$$

Prenons  $a = E[X]$  et remplaçons  $x$  par la v.a.  $X$  dans (1.21). Cette inégalité devient alors  $g(X) \geq g(E[X]) + \theta(X - E[X])$ . En prenant les espérances, il s'ensuit donc que

$$E[g(X)] \geq g(E[X]), \quad (1.21)$$

Cette inégalité porte le nom d'*inégalité de Jensen*.

### 1.3 Exemples de variables aléatoires

Nous allons rappeler les types de v.a. les plus usuels.

- variables aléatoires discrètes : v.a. de Bernoulli, binomiale, géométrique, de Poisson
- variables aléatoires continues : v.a. uniforme, gaussienne, exponentielle, Gamma, Erlang, Chi-carré

#### 1.3.1 Variable aléatoire de Bernoulli

La *fonction indicatrice*  $I_A$  d'un évènement  $A$  lié à une expérience aléatoire est

$$I_A(\zeta) = \begin{cases} 0 & \text{si } \zeta \notin A \\ 1 & \text{si } \zeta \in A. \end{cases} \quad (1.22)$$

Par exemple, si  $A$  consiste à obtenir un chiffre pair lors d'un lancer de dé, on a

$$I_A(\zeta) = \begin{cases} 0 & \text{si } \zeta = 1, 3, 5 \\ 1 & \text{si } \zeta = 2, 4, 6. \end{cases}$$

Cette fonction est donc une v.a.  $X = I_A$ , qui prend une des valeurs de  $S_X = \{0, 1\}$  avec les probabilités  $P(X = 0) = 1 - p$  et  $P(X = 1) = p$ , où  $p = P(A)$ , et est appelée *variable aléatoire de Bernoulli* de paramètre  $p$ .

On calcule aisément sa moyenne

$$\mu_X = E[X] = p, \quad (1.23)$$

sa variance

$$\sigma_X^2 = E[X^2] - \mu_X^2 = p - p^2 = p(1 - p) \quad (1.24)$$

et sa fonction génératrice de probabilité

$$G_X(z) = 1 - p + pz. \quad (1.25)$$

### 1.3.2 Variable aléatoire binomiale

On répète une expérience  $A$  de Bernoulli, de paramètre  $p$ ,  $n$  fois indépendamment. La v.a.  $X$  comptant le nombre de succès obtenus est la *variable aléatoire binomiale* de paramètres  $(n, p)$ , qui s'exprime à partir de la fonction indicatrice par

$$X = I_{A_1} + I_{A_2} + \dots + I_{A_n}$$

Son domaine est donc  $S_X = \{0, 1, 2, \dots, n\}$ . Sa loi de probabilité est donnée par

$$p_k = P(X = k) = C_n^k p^k (1-p)^{(n-k)} \quad (1.26)$$

où

$$C_n^k = \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

A l'aide de la formule du binôme (ou plus simplement en utilisant les techniques du module suivant, basées sur le fait que  $X$  est une somme de v.a. de Bernoulli indépendantes), on trouve

$$\mu_X = E[X] = np \quad (1.27)$$

$$\sigma_X^2 = np(1-p) \quad (1.28)$$

$$G_X(z) = (1-p + pz)^n. \quad (1.29)$$

### 1.3.3 Variable aléatoire géométrique

A présent, on compte le nombre d'essais de Bernoulli avant d'obtenir un succès, les expériences de Bernoulli, de paramètre  $p$ , étant indépendantes. Dans une première version, la *variable aléatoire géométrique* de paramètre  $p$   $X$  compte le nombre d'*échecs* avant le premier succès. Par conséquent, son domaine est  $S_X = \{0, 1, 2, \dots, \}$  et sa loi de probabilité est donnée par

$$p_k = P(X = k) = p(1-p)^k \quad (1.30)$$

On trouve, en recourant aux formules relatives aux progressions géométriques,

$$\mu_X = E[X] = \frac{1-p}{p} \quad (1.31)$$

$$\sigma_X^2 = \frac{1-p}{p^2} \quad (1.32)$$

$$G_X(z) = \frac{p}{1-(1-p)z}. \quad (1.33)$$

Dans la seconde version, la *variable aléatoire géométrique* de paramètre  $p$   $X'$  compte le nombre d'*essais* avant le premier succès. Par conséquent,  $X' = X + 1$ , son domaine est  $S_{X'} = \{1, 2, \dots, \}$  et sa loi de probabilité est donnée par

$$p'_k = P(X' = k) = p(1-p)^{(k-1)} \quad (1.34)$$

On trouve immédiatement, comme  $X' = X + 1$ ,

$$\mu_{X'} = E[X'] = E[X] + 1 = \frac{1}{p} \quad (1.35)$$

$$\sigma_{X'}^2 = \frac{1-p}{p^2} \quad (1.36)$$

$$G_{X'}(z) = \frac{pz}{1 - (1-p)z}. \quad (1.37)$$

### 1.3.4 Variable aléatoire de Poisson

Dans le processus de Poisson, une v.a. aléatoire compte le nombre d'arrivées de clients dans un système pendant une unité de temps, ces arrivées étant soumises à certaines hypothèses. Cette étude sera faite au module 15. Pour le moment, on se contente de définir la *variable aléatoire de Poisson*, de paramètre  $\mu > 0$ , par la loi de probabilité

$$p_k = P(X = k) = \frac{\mu^k}{k!} e^{-\mu}. \quad (1.38)$$

En utilisant le développement de  $\exp(-\mu)$  en série de Taylor, on obtient après quelques manipulations

$$\mu_X = E[X] = \mu \quad (1.39)$$

$$\sigma_X^2 = \mu \quad (1.40)$$

$$G_X(z) = e^{\mu(z-1)}. \quad (1.41)$$

La v.a. de Poisson a la propriété d'être la limite d'une v.a. binomiale  $B(n, p)$ , lorsque  $n \rightarrow \infty$  et  $p \rightarrow 0$ . Plus précisément, soit  $X$  une binomiale  $B(n, p)$  et  $Y$  une v.a. de Poisson de paramètre  $\mu$ . Alors si  $n \rightarrow \infty$  et  $p \rightarrow 0$  tout en gardant  $np$  fini,

$$p_k = P(X = k) = C_n^k p^k (1-p)^{(n-k)} \rightarrow \frac{\mu^k}{k!} e^{-\mu} = P(Y = k)$$

avec  $\mu = np$ . En effet, pour  $k = 0$ , comme  $C_n^0 = 1$ , on a

$$p_0 = (1-p)^n = (1 - \mu/n)^n \rightarrow e^{-\mu}$$

si  $n \rightarrow \infty$ . Pour  $k > 0$ ,

$$\frac{p_{k+1}}{p_k} = \frac{C_n^{k+1} p^{k+1} (1-p)^{(n-k-1)}}{C_n^k p^k (1-p)^{(n-k)}} = \dots = \frac{(1-k/n)\mu}{(k+1)(1-\mu/n)} \rightarrow \frac{\mu}{k+1}$$

si  $n \rightarrow \infty$ , d'où à la limite

$$p_{k+1} \rightarrow \frac{\mu}{k+1} p_k \rightarrow \frac{\mu}{k+1} \frac{\mu}{k} p_{k-1} \rightarrow \dots \rightarrow \frac{\mu^{k+1}}{(k+1)!} p_0 \rightarrow \frac{\mu^{k+1}}{(k+1)!} e^{-\mu}.$$

### 1.3.5 Variable aléatoire uniforme

Une *variable aléatoire uniforme continue* sur un intervalle  $[a, b] = S_X$  est définie par sa densité de probabilité

$$f_X(x) = \begin{cases} 1/(b-a) & \text{si } a \leq x \leq b \\ 0 & \text{sinon.} \end{cases} \quad (1.42)$$

On calcule immédiatement sa moyenne

$$\mu_X = E[X] = \frac{a+b}{2}, \quad (1.43)$$

sa variance

$$\sigma_X^2 = E[X^2] - \mu_X^2 = \int_a^b x^2 \frac{1}{(b-a)} dx - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12} \quad (1.44)$$

et sa fonction caractéristique

$$\Phi_X(\omega) = \int_a^b \frac{1}{(b-a)} e^{j\omega x} dx = \frac{e^{j\omega b} - e^{j\omega a}}{(b-a)j\omega}. \quad (1.45)$$

### 1.3.6 Variable aléatoire gaussienne

La gaussienne est une des variables aléatoires les plus importantes, car c'est une excellente approximation d'une somme d'un grand nombre de v.a. dont la densité de probabilité n'est pas entièrement connue, ce qui est très souvent le cas en pratique.

Son domaine est  $S_X = \mathbb{R}$ . Une *variable aléatoire gaussienne*  $X$  de paramètres  $(\mu, \sigma)$  est définie par sa densité de probabilité

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (1.46)$$

L'interprétation des paramètres  $\mu$  et  $\sigma$  est fournie par le calcul de sa moyenne et de sa variance. En effet,  $\mu$  n'est autre que la moyenne, car la fonction  $f_X(x)$  est symétrique autour du point  $x = \mu$ ,

$$\mu_X = E[X] = \mu \quad (1.47)$$

tandis que  $\sigma$  est l'écart-type car le calcul de la variance donne

$$\sigma_X^2 = E[X^2] - \mu_X^2 = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx - \mu^2 = \dots = \sigma^2. \quad (1.48)$$

La fonction de répartition doit être calculée numériquement ou à l'aide de tables (fonction  $Q(\cdot)$  ou  $\text{erf}(\cdot)$ )

$$F_X(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\frac{x-\mu}{\sigma}} e^{-\xi^2/2} d\xi = 1 - \text{erf}\left(\frac{x-\mu}{\sigma}\right) \quad (1.49)$$

car cette intégrale n'est pas connue sous forme analytique. Sa fonction caractéristique sera calculée à l'exercice 8 :

$$\Phi_X(\omega) = e^{j\mu\omega - \frac{\sigma^2\omega^2}{2}}. \quad (1.50)$$

Une v.a. gaussienne centrée et normalisée, i.e. dont la moyenne  $\mu = 0$  et l'écart type  $\sigma = 1$ , est encore appelée *variable aléatoire normale*.

### 1.3.7 Variable aléatoire exponentielle

Une *variable aléatoire exponentielle* mesure le temps entre deux arrivées successives dans un processus poissonnien. Elle est à valeurs positives  $S_X = [0, \infty[$  et est paramétrisée par un réel positif  $\lambda > 0$ . Sa densité de probabilité est

$$f_X(x) = \lambda e^{-\lambda x}, \quad (1.51)$$

ce qui permet de calculer sa fonction de répartition, sa moyenne, sa variance et sa fonction caractéristique :

$$F_X(x) = 1 - e^{-\lambda x} \quad (1.52)$$

$$\mu_X = E[X] = \frac{1}{\lambda} \quad (1.53)$$

$$\sigma_X^2 = \frac{1}{\lambda^2} \quad (1.54)$$

$$\Phi_X(\omega) = \int_0^\infty \lambda e^{-\lambda x} e^{j\omega x} dx = \frac{\lambda}{\lambda - j\omega}. \quad (1.55)$$

La v.a. exponentielle est la seule v.a. continue sans mémoire (cfr exercice 4) :

$$P(X > t + T | X > t) = P(X > T).$$

### 1.3.8 Variable aléatoire Gamma

Pour introduire cette v.a., il faut d'abord introduire la fonction  $\Gamma(\cdot)$  définie pour  $u > 0$  par

$$\Gamma(u) = \int_0^\infty x^{u-1} e^{-x} dx.$$

Elle possède les propriétés

$$\begin{aligned} \Gamma(1/2) &= \sqrt{\pi} \\ \Gamma(u+1) &= u\Gamma(u) \\ \Gamma(u) &= (u-1)! \quad \text{si } u \in \mathbb{N}. \end{aligned}$$

Une *variable aléatoire Gamma*  $G(\lambda, \alpha)$  est une v.a. à valeurs strictement positives ( $S_X = ]0, \infty[$ ), paramétrisée par deux réels positifs  $\alpha, \lambda > 0$  et définie par sa densité de probabilité

$$f_X(x) = \frac{\lambda(\lambda x)^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}. \quad (1.56)$$

Sa moyenne, sa variance et sa fonction caractéristique sont données par

$$\mu_X = E[X] = \frac{\alpha}{\lambda} \quad (1.57)$$

$$\sigma_X^2 = \frac{\alpha}{\lambda^2} \quad (1.58)$$

$$\Phi_X(\omega) = \frac{1}{(1 - j\omega/\lambda)^\alpha}. \quad (1.59)$$

Beaucoup de v.a. continues sont des cas particuliers de la v.a. Gamma( $\lambda, \alpha$ ). Par exemple, la v.a.  $G(\lambda, m)$  où  $m \in \mathbb{N}$  est une *variable aléatoire d'Erlang* de paramètres  $(\lambda, m)$ . Comme  $m$  est entier, (1.56) devient

$$f_X(x) = \frac{\lambda(\lambda x)^{m-1}e^{-\lambda x}}{(m-1)!}. \quad (1.60)$$

En particulier, la v.a. d'Erlang lorsque  $m = 1$  est une v.a. exponentielle. Par contre, si  $\lambda = 1/2$  et si  $\alpha = \nu/2$  avec  $\nu \in \mathbb{N}$ , la v.a.  $G(1/2, \nu/2)$  est une *variable aléatoire Chi-carré*  $\chi^2(\nu)$  à  $\nu$  degrés de liberté. Dans ce cas, (1.56) devient

$$f_X(x) = \frac{x^{(\nu-2)/2}e^{-x/2}}{2^{\nu/2}\Gamma(\nu/2)}. \quad (1.61)$$

## 1.4 Exercices

- Démontrer le théorème des probabilités totales P6. Hint : remarquer que l'évènement  $B$  peut encore être écrit comme  $B = B \cap \Omega = B \cap (A_1 \cup A_2 \cup \dots \cup A_n)$ .
- Démontrer, en utilisant le théorème des probabilités totales P6, la règle de Bayes P7.
- Le "run length" coding est utile pour coder simplement de l'information binaire (Bits 0,1) si un symbole (1) a une probabilité d'apparition  $p$  beaucoup plus faible que celle de l'autre symbole (0), comme c'est par exemple le cas pour un fax (0 = blanc, 1 = noir). Il segmente une séquence de symboles binaires, émis indépendamment l'un de l'autre, en des paquets  $f_i$  consistant en soit en une suite de  $i$  symboles 0 terminés par un 1, pour  $0 \leq i \leq n-1$ , soit en une suite de  $n$  0. Par exemple, pour  $n = 3$ , on a la table suivante :

$f_i$	longueur $i$
1	0
01	1
001	2
000	3

Si  $X$  est la variable aléatoire donnant la longueur  $i$  du paquet  $f_i$ , que vaut la probabilité  $P(X = i)$  pour tout  $0 \leq i \leq n$ ? Quelle est la fonction génératrice  $G_X(z)$ ? Que vaut la longueur moyenne  $\mu_X$ ?

- Démontrez qu'une variable aléatoire exponentielle  $X$  est sans mémoire, i.e. que pour tout  $t, T \geq 0$

$$P(X > t + T | X > t) = P(X > T).$$

- Montrer que la v.a. de Poisson  $Y$  de paramètre  $\lambda$  est la forme limite d'une v.a. binomiale  $X$  de paramètre  $(n, p)$  lorsque  $n \rightarrow \infty$  avec  $\lambda = np$  restant fini, en utilisant les fonctions génératrices. Il faut donc montrer que  $G_X(z) \rightarrow G_Y(z)$  lorsque  $n \rightarrow \infty$  avec  $\lambda = np$ .
- Montrer que la moyenne, la variance et la fonction génératrice d'une v.a. géométrique sont respectivement données par (1.31), (1.32) et (1.33).

7. Montrer que la moyenne, la variance et la fonction caractéristique d'une v.a. exponentielle sont respectivement données par (1.53), (1.54) et (1.55).
8. Soit  $X$  une variable aléatoire gaussienne, de moyenne  $\mu$  et de variance  $\sigma^2$ .

(a) Démontrez que sa fonction caractéristique est

$$\Phi_X(\omega) = e^{j\mu\omega - \sigma^2\omega^2/2}.$$

- (b) Si  $\mu = 0$  que valent ses moments d'ordre 3 et 4  $E[X^3]$  et  $E[X^4]$  ?
- (c) Quelle est la densité de probabilité de la variable aléatoire  $Y = aX + b$ , où les constantes  $a, b \in \mathbb{R}$  ? Que valent sa moyenne  $\mu_Y = E[Y]$  et sa variance  $\sigma_Y^2$  ?
9. Soit  $X$  une variable aléatoire uniforme sur l'intervalle  $[-2, 2]$ . Sa densité de probabilité est donc donnée par  $f_X(x) = 1/4$  si  $-2 \leq x \leq 2$  et  $f_X(x) = 0$  sinon. Quelles sont la fonction de répartition  $F_Y(y)$  et la densité de probabilité  $f_Y(y)$  de la variable aléatoire  $Y = g(X)$  où  $g(\cdot)$  est la fonction linéaire par morceaux d'un limiteur, donnée par

$$g(x) = \begin{cases} -1 & \text{si } x < -1 \\ x & \text{si } -1 \leq x \leq 1 \\ 1 & \text{si } 1 < x \end{cases} ?$$

Attention : il s'agira d'une variable aléatoire mixte (une partie continue et l'autre discrète) ! Il est commode de représenter graphiquement  $f_X(x)$ ,  $g(x)$ ,  $F_Y(y)$  et  $f_Y(y)$ .

10. Soit  $X$  une variable aléatoire normale. Quelle est la densité de probabilité de la v.a.  $Y = X^2$  ? Quel nom porte la v.a.  $Y$  ?
11. Soit  $X$  une variable aléatoire uniforme dans l'intervalle  $[0, 2\pi]$ . Quelle est la densité de probabilité de la v.a.  $Y = \sin X$  ? Que valent sa moyenne et sa variance ?
12. Soit  $X$  une v.a. discrète dont le domaine  $S_X = \mathbb{N}$  et telle que

$$P(X = k) \geq P(X = k + 1)$$

pour tout  $k \in \mathbb{N}$ . Démontrez que

$$P(X = k) \leq \frac{2E[X]}{k^2}.$$

13. Soit  $X$  une v.a. ne prenant que des valeurs positives, et dont la fonction génératrice de moment est  $\hat{\Phi}_X(s)$ . Démontrez que pour tout  $x > 0$  et tout  $s > 0$

$$P(X > x) \leq e^{-sx} \hat{\Phi}_X(s).$$

Cette inégalité porte le nom d'inégalité de Chernoff, et est à la base des grandes déviations qu'on verra au module 10.

14. Calculer le moment d'ordre  $n$  d'une variable aléatoire exponentielle de paramètre  $\lambda$ .