

Avant – Propos

Ce cours de **Probabilités** est destiné surtout aux étudiants de L2 Mathématiques (dans le cadre du système L.M.D.), il couvre le programme officiel. Le contenu de ce cours est consacré au programme du module Probabilités, à savoir :

- 1) Rappels sur les probabilités.
- 2) Variables aléatoires à une dimension.
- 3) Lois de probabilités absolument continues usuelles.

J'espère que ce support aidera les étudiants de deuxième année Mathématiques à assimiler les Probabilités et plus particulièrement les variables aléatoires. Enfin, des erreurs peuvent être relevées, prière de les signaler à l'auteur.

Table des matières

1	Variables aléatoires à une dimension	1
1.1	Loi de probabilité d'une variable aléatoire	2
1.2	Fonction de répartition	3
1.3	Variables aléatoires discrètes	5
1.4	Loi de Probabilité	5
1.5	Espérance	6
1.6	Variance et écart-type	8
1.7	Lois de probabilités usuelles (cas discèt)	11
1.7.1	Loi uniforme sur $\{1, \dots, n\}$	11
1.7.2	Loi de Bernoulli	12
1.7.3	Loi binomiale	12
1.7.4	Loi hypergéométrique	13
1.7.5	Loi géométrique	14
1.7.6	Loi de Poisson	15
1.8	Travaux dirigès	17

1

Variables aléatoires à une dimension

Dans tout ce chapitre, Ω est un ensemble au plus dénombrable, et on note ω l'élément générique de Ω et $\mathcal{P}(\Omega)$ l'ensemble des parties de Ω .

Définition 1.0.1. (*Variable aléatoire*)

Soit Ω l'univers d'une expérience aléatoire muni d'une tribu \mathcal{A} (ou σ -algèbre \mathcal{A}) et \mathbb{R} est muni de la tribu $\mathcal{B}_{\mathbb{R}}$. Une application

$$X : \begin{cases} \Omega \rightarrow \mathbb{R} \\ \omega \mapsto X(\omega) \end{cases}$$

est une variable aléatoire, si pour tout événement $B \in \mathcal{B}_{\mathbb{R}}$, $X^{-1}(B) \in \mathcal{A}$.

Exemple 1.0.1. On jette deux dés, et on pose $\Omega = \{(i, j), (i, j) \in \{1, 2, 3, 4, 5, 6\}^2\}$.

$$X : \omega = (a, b) \mapsto a + b \text{ (somme des deux dés)}$$

Exemple 1.0.2. On lance une fléchette contre une cible, et on pose $\Omega = \mathbb{R}^2$.

$$D : \omega = (x, y) \mapsto \sqrt{x^2 + y^2}$$

(distance euclidienne du point d'impact au centre de la cible).

Il est naturel d'exiger que l'on puisse calculer la probabilité que " X tombe dans l'intervalle $[a, b]$ ". Il faut donc que :

$X^{-1}([a, b]) \in \mathcal{A}$. On rappelle que pour toute partie B de \mathbb{R} , on pose :

$$X^{-1}(B) = \{\omega \in \Omega / X(\omega) \in B\}.$$

C'est l'image réciproque de la partie B , il n'est pas question de bijection réciproque de X dans la notation $X^{-1}(B)$.

Notation. On notera v.a.r. pour variable aléatoire réelle.

Propriété 1.0.1. *L'ensemble des variables aléatoires réelles définies sur un même espace (Ω, \mathcal{A}) , muni des lois usuelles, a une structure d'algèbre. En particulier, la somme et le produit de deux variables aléatoires réelles sont des variables aléatoires réelles.*

Définition 1.0.2. *Soit X une variable aléatoire réelle. Les événements du type $X^{-1}([a, b])$, $X^{-1}(]a, b])$, $X^{-1}([a, b])$ et $X^{-1}(a)$ sont appelés événements liés à X .*

La tribu engendrée par ces événements s'appelle tribu engendrée par X ; c'est par définition une tribu incluse dans \mathcal{A} , notée \mathcal{A}_X .

1.1 Loi de probabilité d'une variable aléatoire

Définition 1.1.1. *On appelle loi de probabilité (ou distribution de probabilité) d'une v.a.r. X définie sur un espace $(\Omega; \mathcal{A}; \mathbf{P})$ l'application*

$$\mu : \begin{cases} \mathcal{B}_{\mathbb{R}} \rightarrow [0, 1] \\ B \mapsto \mathbf{P}(X^{-1}(B)) \end{cases}$$

Remarque 1.1.1. Par définition $X^{-1}(B) = \{\omega \in \Omega / X(\omega) \in B\}$, et cet événement est souvent noté $\langle\langle X \in B \rangle\rangle$.

Propriété 1.1.1. *L'application μ ainsi définie est une probabilité sur l'espace probabilisable $(\mathbb{R}; \mathcal{B}_{\mathbb{R}})$. Elle est appelée probabilité image de \mathbf{P} par X .*

Preuve.

Pour montrer que μ est une probabilité sur $(\mathbb{R}; \mathcal{B}_{\mathbb{R}})$, on vérifie les trois axiomes.

Positivité : $\forall B \in \mathcal{B}_{\mathbb{R}}, \mu(B) = \mathbf{P}(X^{-1}(B)) \geq 0$.

Totalité : $\mu(\mathbb{R}) = \mathbf{P}(X \in \mathbb{R}) = \mathbf{P}(\Omega) = 1$.

Additivité : soit B_n une suite d'éléments de $\mathcal{B}_{\mathbb{R}}$ deux à deux disjoints.

$$\begin{aligned} \mu\left(\bigcup_n B_n\right) &= \mathbf{P}\left(X^{-1}\left(\bigcup_n B_n\right)\right) \\ &= \mathbf{P}\left(\bigcup_n X^{-1}(B_n)\right) \\ &= \sum_n \mathbf{P}(X^{-1}(B_n)) \text{ (additivité de } \mathbf{P}) \\ &= \sum_n \mu(B_n) \end{aligned}$$

Notation. L'application μ , est parfois notée \mathbf{P}_X .

Propriété 1.1.2. *Si X est une v.a.r. sur (Ω, \mathcal{A}) et si g est bien définie et continue par morceaux sur $X(\Omega)$, alors goX est une v.a.r. sur (Ω, \mathcal{A}) .*

Preuve. Soit $B \in \mathcal{B}_{\mathbb{R}}$, Comme g est continue par morceaux, $g^{-1}(B) \in \mathcal{B}_{\mathbb{R}}$ et comme X est une v.a.r. sur (Ω, \mathcal{A}) , $X^{-1}(g^{-1}(B)) \in \mathcal{A}$, c'est-à-dire $(goX)^{-1}(B) \in \mathcal{A}$.

Exemple 1.1.1. On lance un dé équilibré, on pose $\Omega = \{1, 2, 3, 4, 5, 6\}$, $\mathcal{A} = \mathcal{P}(\Omega)$ et on prend pour \mathbf{P} la probabilité uniforme. On réalise un gain nul si l'on obtient 1, un gain de 1\$ si l'on obtient 2, 3 ou 4 et de 2\$ si le résultat est 5 ou 6. On note X la v.a.r. égale au gain obtenu.

On a donc $X(\Omega) = \{0, 1, 2\}$. La loi de X , c'est-à-dire l'application μ définie auparavant, est entièrement déterminée par $\mathbf{P}(X = 0)$, $\mathbf{P}(X = 1)$ et $\mathbf{P}(X = 2)$, c'est-à-dire par les poids de probabilité affectés aux éléments de son univers image $X(\Omega)$. Par exemple pour $B = \{1, 2\}$

$$\begin{aligned} \mathbf{P}(X \in B) &= \mathbf{P}(X^{-1}(B)) \\ &= \mathbf{P}(X^{-1}(\{1\}) \cup X^{-1}(\{2\})) \\ &= \mathbf{P}(X^{-1}(\{1\})) + \mathbf{P}(X^{-1}(\{2\})) \\ &= \mathbf{P}(\{2, 3, 4\}) + \mathbf{P}(\{5, 6\}) \\ &= \mathbf{P}(\{2\}) + \dots + \mathbf{P}(\{6\}) \\ &= \frac{5}{6} \end{aligned}$$

1.2 Fonction de répartition

Définition 1.2.1. Soit X une variable aléatoire réelle définie sur un espace probabilisable $(\Omega; \mathcal{A}; \mathbf{P})$. On appelle fonction de répartition de X l'application :

$$\begin{aligned} F : \mathbb{R} &\rightarrow [0, 1] \\ x &\mapsto \mathbf{P}(X \leq x) \end{aligned}$$

Dans l'exemple précédent, on obtient

$$F(x) = \begin{cases} 0 & \text{si } 0 < x \\ \frac{1}{6} & \text{si } 0 \leq x < 1 \\ \frac{2}{3} & \text{si } 1 \leq x < 2 \\ 1 & \text{si } x \geq 2 \end{cases}$$

Propriété 1.2.1. Toute fonction de répartition F possède les propriétés suivantes :

1. F est croissante.
2. $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$.
3. F est continue à droite en tout point x de \mathbb{R} .
4. F a une limite à gauche en tout point x de \mathbb{R} et :

$$\forall x \in \mathbb{R}, \quad F(x) - \lim_{t \rightarrow x^-} F(t) = \mathbf{P}(X = x)$$

Preuve.

1. Soit $y \geq x$: $F(y) - F(x) = \mathbf{P}(X \leq y) - \mathbf{P}(X \leq x) = \mathbf{P}(x < X \leq y) \geq 0$, en effet F est croissante.
2. Comme F est monotone (en l'occurrence croissante), on a l'équivalence suivante :

$$\lim_{x \rightarrow -\infty} F(x) = 0 \Leftrightarrow \lim_{n \rightarrow +\infty} F(-n) = 0.$$

Or $F(-n) = \mathbf{P}(X \in]-\infty; -n])$. Notons pour tout $n \in \mathbb{N}$, $A_n = X^{-1}(]-\infty; -n])$. La suite d'événements (A_n) est décroissante au sens de l'inclusion et on a donc, grâce aux propriétés de la probabilité \mathbf{P} . Donc la suite de réels $(\mathbf{P}(A_n))$ est convergente et :

$$\begin{aligned} \lim_{n \rightarrow +\infty} \mathbf{P}(A_n) &= \mathbf{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) \\ \implies \lim_{n \rightarrow +\infty} F(-n) &= \mathbf{P}(\emptyset) \end{aligned}$$

Même démarche pour la limite en $+\infty$.

3. Soit $x \in \mathbb{R}$ fixé. Comme F est croissante, il suffit d'étudier la convergence de la suite de réels $(F(x + \frac{1}{n}))$ pour étudier l'existence et la valeur de la limite à droite de F en x . Par définition :

$$F\left(x + \frac{1}{n}\right) = \mathbf{P}\left(X \in \left] -\infty; x + \frac{1}{n} \right] \right)$$

En notant $A_n = X^{-1}\left(\left] -\infty; x + \frac{1}{n} \right] \right)$, on obtient une suite d'événements décroissante pour l'inclusion, donc la suite de réels $(\mathbf{P}(A_n))$ converge et

$$\begin{aligned} \lim_{n \rightarrow +\infty} \mathbf{P}(A_n) &= \mathbf{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) \\ \implies \lim_{n \rightarrow +\infty} F\left(x + \frac{1}{n}\right) &= \mathbf{P}\left(X^{-1}\left(\left] -\infty; x \right] \right)\right) = F(x) \end{aligned}$$

4. Là encore, il suffit d'étudier la suite $(F(x - \frac{1}{n}))$

$$F\left(x - \frac{1}{n}\right) = \mathbf{P}\left(X \in \left] -\infty; x - \frac{1}{n} \right] \right)$$

En notant $A_n = X^{-1}\left(\left] -\infty; x - \frac{1}{n} \right] \right)$, on obtient une suite d'événements croissante pour l'inclusion, donc la suite de réels $(\mathbf{P}(A_n))$ converge et :

$$\begin{aligned} \lim_{n \rightarrow +\infty} \mathbf{P}(A_n) &= \mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) \\ \implies \lim_{n \rightarrow +\infty} F\left(x - \frac{1}{n}\right) &= \mathbf{P}\left(X^{-1}\left(\left] -\infty; x \right] \right)\right) = \mathbf{P}(X < x) \end{aligned}$$

F a donc une limite à gauche en x , et :

$$F(x) - \lim_{t \rightarrow x^-} F(t) = \mathbf{P}(X \leq x) - \mathbf{P}(X < x) = \mathbf{P}(X = x)$$

Théorème 1.2.1. Toute fonction $F : \mathbb{R} \rightarrow [0; 1]$ telle que :

1. F est croissante,
2. $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$,
3. F est continue à droite,
est la fonction de répartition d'une v.a.r. X .

Preuve.

1.3 Variables aléatoires discrètes

Définition 1.3.1. Une v.a.r. X est dite discrète si $X(\Omega)$ est fini ou dénombrable.

Un ensemble dénombrable est un ensemble en bijection avec \mathbb{N} . X est donc discrète si $X(\Omega)$ est de la forme $X(\Omega) = \bigcup_{i \in I} \{x_i\}$, $I \subseteq \mathbb{N}$

Exemple 1.3.1. Une urne contient 3 boules rouges et 4 boules noires. On extrait successivement avec remise 2 boules de l'urne. L'univers est $\Omega = \{R_1, R_2, R_3, N_1, N_2, N_3, N_4\}$ et on considère comme tribu $\mathcal{A} = \mathcal{P}(\Omega)$. On mise au départ 10 \$, et on gagne 8 \$ par boule rouge obtenue. Soit X la v.a.r. prenant pour valeur le gain final. On a : $X(\Omega) = \{-10, -2, +6\}$.

Exemple 1.3.2. Dans le cas du jet de deux dés, la somme S des deux résultats obtenus est une variable aléatoire discrète à valeurs dans $F = \{2, 3, \dots, 12\}$ dont nous avons calculé la loi :

$$\begin{aligned} \mathbf{P}(S = 2) = \mathbf{P}(S = 12) &= 1/36 & \mathbf{P}(S = 3) = \mathbf{P}(S = 11) &= 1/18 \\ \mathbf{P}(S = 4) = \mathbf{P}(S = 10) &= 1/12 & \mathbf{P}(S = 5) = \mathbf{P}(S = 9) &= 1/9 \\ \mathbf{P}(S = 6) = \mathbf{P}(S = 8) &= 5/36 & \mathbf{P}(S = 7) &= 1/6. \end{aligned}$$

Il faut noter que la loi de S est la probabilité uniforme dont on doit munir l'ensemble $\{2, 3, \dots, 12\}$ lorsque l'on s'intéresse à la somme de deux résultats obtenus.

1.4 Loi de Probabilité

Soit X une v.a.r. discrète, d'univers image $X(\Omega) = \{x_i, i \in I\}$ (où I est une partie de \mathbb{N}). D'après la définition vue au chapitre précédent, il s'agit de déterminer l'application :

$$\mu : \begin{cases} \mathcal{B}_{\mathbb{R}} \rightarrow [0; 1] \\ B \mapsto \mathbf{P}(X^{-1}(B)) \end{cases}$$

Or cette application est entièrement déterminée dès que l'on connaît $(\mathbf{P}(X = x_i))_{i \in I}$.

En effet, on a, par l'additivité de \mathbf{P} :

$$\forall B \in \mathcal{B}_{\mathbb{R}}, \mu(B) = \mathbf{P}(X \in B) = \left(\bigcup_{i/x_i \in B} (X = x_i) \right) = \sum_{i/x_i \in B} \mathbf{P}(X = x_i).$$

Définition 1.4.1. Donner la loi d'une v.a.r. discrète X , c'est donner $X(\Omega)$, et pour tout $x_i \in X(\Omega)$, la valeur de $\mathbf{P}(X = x_i)$.

Exemple 1.4.1. On reprend l'exemple précédent du jet de deux dés, la somme S des deux résultats obtenus est une variable aléatoire discrète avec la lois de probabilité suivante :

$X(\Omega)$	2	3	4	5	6	7	8	9	10	11	12
$\mathbf{P}(S = s_i)$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

Propriété 1.4.1. *Puisque P vérifie l'axiome de totalité, on a évidemment :*

$$\sum_{i \in I} \mathbf{P}(X = x_i) = 1$$

Preuve.

Il suffit d'écrire que les événements $X^{-1}(\{x_i\}), i \in I$, sont disjoints et que leur réunion est égale à Ω .

Propriété 1.4.2. *Soit I une partie non vide de \mathbb{N} .*

La donnée de deux familles de réels $(x_i)_{i \in I}$ et $(p_i)_{i \in I}$ tels que :

$$1. \forall i \in I, p_i \geq 0;$$

$$2. \sum_{i \in I} p_i = 1,$$

permet de définir la loi de probabilité d'une v.a.r. discrète X en posant :

$$X(\Omega) = \{x_i, i \in I\} \quad \text{et} \quad \forall i \in I, \mathbf{P}(X = x_i) = p_i$$

Exemple 1.4.2. *Soit X d'univers image $X(\Omega) = \mathbb{N}$ telle que*

$$\forall n \in \mathbb{N}, \quad \mathbf{P}(X = n) = e^{-2} \frac{2^n}{n!}$$

On a bien la positivité des $\mathbf{P}(X = n)$ et $\sum_{n=0}^{+\infty} \mathbf{P}(X = n) = 1$.

1.5 Espérance

Il s'agit de la moyenne des valeurs de X pondérées par leur probabilité d'apparition :

Définition 1.5.1. *Soit X une v.a.r. discrète.*

On appelle espérance de X si elle existe le réel, noté $\mathbb{E}(X)$:

$$\mathbb{E}(X) = \sum_{i=1}^n x_i \mathbf{P}(X = x_i),$$

lorsque $X(\Omega) = \{x_1, \dots, x_n\}$, ou :

$$\mathbb{E}(X) = \sum_{i=1}^{+\infty} x_i \mathbf{P}(X = x_i),$$

lorsque $X(\Omega) = \mathbb{N}^$ et que cette série est convergente.*

Exemple 1.5.1. *En reprenant le jet des deux dès et de considérer la somme des dés. En effet,*

$$\mathbb{E}(X) = 2 \times \frac{1}{36} + 3 \times \frac{1}{18} + 4 \times \frac{1}{12} + 5 \times \frac{1}{9} + 6 \times \frac{5}{36} + 7 \times \frac{1}{9} + 8 \times \frac{5}{36} + 9 \times \frac{1}{9} + 10 \times \frac{1}{12} + 11 \times \frac{1}{18} + 12 \times \frac{1}{36} = 6.6111$$

Exemple 1.5.2. Soit X de loi : $X(\Omega) = \mathbb{N}$ et $\forall n \in \mathbb{N}$, $\mathbf{P}(X = n) = e^{-2} \frac{2^n}{n!}$. La série $\sum_{n \geq 0} n \mathbf{P}(X = n)$ est convergente (utiliser par exemple la règle de D'Alembert), donc $\mathbb{E}(X)$ existe, et :

$$\sum_{n=0}^{+\infty} n \mathbf{P}(X = n) = \sum_{n=0}^{+\infty} n e^{-2} \frac{2^n}{n!} = e^{-2} \sum_{n=1}^{+\infty} \frac{2^n}{(n-1)!} = e^{-2} \times 2e^2 = 2$$

Propriété 1.5.1. Si $X(\Omega)$ possède un maximum et un minimum, alors $\mathbb{E}(X)$ existe et :

$$x_{\min} \leq \mathbb{E}(X) \leq x_{\max}$$

Preuve. Pour le cas général, Soit X une variable aléatoire telle que $X(\Omega) = \mathbb{N}$

$$\forall i \in \mathbb{N}, x_{\min} \leq x_i \leq x_{\max}$$

$$\implies \sum_{i=0}^{+\infty} p_i x_{\min} \leq \sum_{i=0}^{+\infty} p_i x_i \leq \sum_{i=0}^{+\infty} p_i x_{\max}$$

D'où l'encadrement souhaité (puisque $\sum_{i=0}^{+\infty} p_i = 1$).

Même démarche dans le cas fini.

Théorème 1.5.1. Si Ω est fini ou dénombrable, alors :

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega) \mathbf{P}(\{\omega\})$$

sous réserve de convergence absolue.

Preuve. On note $A_i = \{\omega \in \Omega / X(\omega) = x_i\}$, alors :

$$\mathbf{P}(X = x_i) = \mathbf{P}(X^{-1}(\{x_i\})) = \mathbf{P}(A_i) = \sum_{\omega \in A_i} \mathbf{P}(\{\omega\}).$$

On obtient donc :

$$\begin{aligned} \sum_i x_i \mathbf{P}(X = x_i) &= \sum_i x_i \sum_{\omega \in A_i} \mathbf{P}(\{\omega\}) \\ &= \sum_i \sum_{\omega \in A_i} x_i \mathbf{P}(\{\omega\}) \\ &= \sum_i \sum_{\omega \in A_i} X(\omega) \mathbf{P}(\{\omega\}) \quad \text{car } \omega \in A_i \implies X(\omega) = x_i \\ &= \sum_{\omega \in \Omega} X(\omega) \mathbf{P}(\{\omega\}) \quad \text{car } \bigcup_i A_i = \Omega. \end{aligned}$$

Ce théorème est essentiel car il permet de montrer facilement la linéarité de l'opérateur E , ce qui n'a rien d'évident avec les formules de la définition.

Propriété 1.5.2. Soient X et Y deux v.a.r. discrètes sur Ω qui possèdent une espérance, et soit $\lambda \in \mathbb{R}$.

Alors $X + \lambda Y$ est une v.a.r. discrète sur Ω qui possède une espérance, et :

$$\mathbb{E}(X + \lambda Y) = \mathbb{E}(X) + \lambda \mathbb{E}(Y)$$

Autrement dit, l'opérateur E est une forme linéaire sur l'espace vectoriel des v.a.r. discrètes sur Ω d'espérance finie.

Preuve. La série $\sum_{\omega \in \Omega} (X(\omega) + \lambda Y(\omega)) \mathbf{P}(\{\omega\})$ est absolument convergente comme somme de deux séries absolument convergentes, et on peut séparer la somme en deux (puisque chacune a bien un sens) :

$$\begin{aligned} \sum_{\omega \in \Omega} (X(\omega) + \lambda Y(\omega)) \mathbf{P}(\{\omega\}) &= \sum_{\omega \in \Omega} X(\omega) \mathbf{P}(\{\omega\}) + \lambda \sum_{\omega \in \Omega} Y(\omega) \mathbf{P}(\{\omega\}) \\ &= \mathbb{E}(X) + \lambda \mathbb{E}(Y) \end{aligned}$$

Propriété 1.5.3. Soient X et Y deux v.a.r. discrètes, $M \in \mathbb{R}_+^*$ et $a \in \mathbb{R}$, alors :

- 1) Si $X \geq 0$, $\mathbb{E}(X) \geq 0$.
- 2) Si $X \leq Y$, $\mathbb{E}(X) \leq \mathbb{E}(Y)$.
- 3) Si $|X| \leq M$, $\mathbb{E}(X) \leq \infty$.
- 4) Si $|X| = a$, $\mathbb{E}(X) = a$.

Preuve. (Analogie)

Théorème 1.5.2. Théorème de transfert

Soient X une v.a.r. discrètes sur Ω d'univers image $X(\Omega) = \{x_i, i \in I\}$, et g une fonction continue, donc $g \circ X$ est une v.a.r. discrète. On a :

$$\mathbb{E}(g(X)) = \sum_{i \in I} g(x_i) \mathbf{P}(X = x_i)$$

1.6 Variance et écart-type

Définition 1.6.1. On appelle moment d'ordre 2 de X l'espérance, si elle existe, de la variable X^2 . D'après le théorème précédent, c'est donc le réel, s'il existe :

$$\mathbb{E}(X^2) = \sum_{i \in I} x_i^2 \mathbf{P}(X = x_i)$$

Définition 1.6.2. On appelle variance de X l'espérance, si elle existe, de la variable $(X - \mathbb{E}(X))^2$:

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

C'est donc le réel, sous réserve d'existence :

$$\mathbb{E}((X - \mathbb{E}(X))^2) = \sum_{i \in I} (x_i - \mathbb{E}(X))^2 \mathbf{P}(X = x_i)$$

La variance est aussi appelée moment centré d'ordre 2.

Remarque 1.6.1. Son existence suppose en particulier celle de $\mathbb{E}(X)$. On mesure la somme des carrés des écarts entre les valeurs de X et la moyenne, pondérés par les poids de probabilité. C'est donc un indicateur de "dispersion" autour de $\mathbb{E}(X)$.

Propriété 1.6.1. *Si X admet une variance, alors :*

1. $\mathbb{V}(X) \geq 0$

2. $\mathbb{V}(X) = 0 \Leftrightarrow X \text{ est constante.}$

En effet : $\mathbb{V}(X) = 0 \Rightarrow \forall i, (x_i - \mathbb{E}(X))^2 \mathbf{P}(X = x_i) = 0$. On a donc que tous les x_i sont égaux (à $\mathbb{E}(X)$). (En toute rigueur, c'est le cas sauf lorsque $\mathbf{P}(X = x_i) = 0$, c'est-à-dire hors d'un ensemble de mesure nulle : X est donc constante "presque partout").

Proposition 1.6.1. Si X une v.a.r. discrète possède un moment d'ordre 2, alors elle possède un moment d'ordre 1.

Preuve.

On suppose que $\sum_{i \in I} x_i^2 p_i$ converge. On a l'inégalité :

$$\forall x \in \mathbb{R}, |x| \leq x^2 + 1.$$

Comme les séries $\sum_{i \in I} x_i^2 p_i$ et $\sum_{i \in I} p_i$ sont convergentes, la série $\sum_{i \in I} |x_i| p_i$ converge, en vertu du théorème de comparaison pour les séries à termes positifs.

Propriété 1.6.2. *X admet une variance si et seulement si X admet un moment d'ordre 2.*

Preuve.

On a :

$$\sum_{i \in I} (x_i - \mathbb{E}(X))^2 p_i = \sum_{i \in I} x_i^2 p_i - 2\mathbb{E}(X) \sum_{i \in I} x_i p_i + (\mathbb{E}(X))^2.$$

On en déduit l'équivalence entre la convergence de la série $\sum_{i \in I} x_i^2 p_i$ et celle de $\sum_{i \in I} (x_i - \mathbb{E}(X))^2 p_i$.

Le calcul pratique se fait souvent avec la formule de Koenig-Huygens :

Théorème 1.6.1. (*Koenig-Huygens*)

Si X possède une variance, alors :

$$\begin{aligned} \mathbb{V}(X) &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \\ &= \sum_{i \in I} x_i^2 \mathbf{P}(X = x_i) - \left(\sum_{i \in I} x_i \mathbf{P}(X = x_i) \right)^2 \end{aligned}$$

Preuve. On utilise la linéarité de l'espérance :

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{E}(X - \mathbb{E}(X))^2 \\ &= \mathbb{E}(X^2 - 2X\mathbb{E}(X) + (\mathbb{E}(X))^2) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + (\mathbb{E}(X))^2 \\ &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2\end{aligned}$$

Exemple 1.6.1. Pour le jeu des 10\$, on obtient :

$$\mathbb{V}(X) = (-10)^2 \times \frac{16}{49} + (-2)^2 \times \frac{24}{49} + 6^2 \times \frac{9}{49} - \left(\frac{-22}{7}\right)^2 \simeq 31.35$$

Propriété 1.6.3. Soit $(a, b) \in \mathbb{R}^2$. Si X possède une variance, alors la v.a.r. $aX + b$ possède une variance et :

$$\mathbb{V}(aX + b) = a^2\mathbb{V}(X)$$

Preuve. Il faut vérifier que la variable $(aX + b - \mathbb{E}(aX + b))^2$ possède une espérance. Or, par linéarité de E :

$$(aX + b - \mathbb{E}(aX + b))^2 = a^2(X - \mathbb{E}(X))^2.$$

Puisque X possède une variance, on a par définition l'existence de l'espérance de la variable $(X - \mathbb{E}(X))^2$, et donc celle de la variable $a^2(X - \mathbb{E}(X))^2$.

On a immédiatement :

$$\begin{aligned}\mathbb{V}(aX + b) &= \mathbb{E}(a^2(X - \mathbb{E}(X))^2) \\ &= a^2\mathbb{E}(X - \mathbb{E}(X))^2 \\ &= a^2\mathbb{V}(X)\end{aligned}$$

Définition 1.6.3. On appelle variable centrée toute variable X telle que : $\mathbb{E}(X) = 0$, et variable réduite toute variable X telle que : $\mathbb{V}(X) = 1$.

Définition 1.6.4. On appelle l'écart-type de X , lorsque X admet une variance, le réel :

$$\sigma(X) = \sqrt{\mathbb{V}(X)}$$

Remarque 1.6.2. La variable $X - \mathbb{E}(X)$ est centrée et la variable $\frac{1}{\sigma(X)}X$ est réduite.

Définition 1.6.5. On appelle moment d'ordre k l'espérance, si elle existe, de la v.a.r. X^k , c'est-à-dire, d'après le théorème de transfert :

$$\mathbb{E}(X^k) = \sum_{i \in I} x_i^k \mathbf{P}(X = x_i).$$

Remarque 1.6.3. Si $X(\Omega)$ est fini, alors X possède des moments de tout ordre. Dans le cas où $X(\Omega)$ est dénombrable, l'existence du moment d'ordre k suppose par définition la convergence absolue de la série $\sum_{i \in I} x_i^k \mathbf{P}(X = x_i)$, et on a la propriété :

Propriété 1.6.4. *Si X possède un moment d'ordre k , alors tous ses moments d'ordre $k' \leq k$ existent.*

Preuve.

On a l'inégalité suivante : $\forall x \in \mathbb{R}, |x|^{k-1} \leq |x|^k + 1$ (distinguer les cas $|x| \geq 1$ et $|x| \leq 1$). Comme $\sum_{i \in I} |x_i|^k p_i$ et $\sum_{i \in I} p_i$ sont convergentes, le théorème de comparaison pour les séries à termes positifs assure la convergence de $\sum_{i \in I} |x_i|^{k-1} p_i$. Par récurrence finie, on obtient bien que pour tout k' inférieur ou égal k , X possède un moment d'ordre k' .

1.7 Lois de probabilités usuelles (cas discèt)

Il s'agit de passer en revue les lois discrètes usuelles du programme, en donnant pour chacune d'elles l'univers image, les poids de probabilité, l'espérance et la variance, ainsi que le contexte classique où on les rencontre.

1.7.1 Loi uniforme sur $\{1, \dots, n\}$

Définition 1.7.1. *X suit une loi uniforme sur $\{1, \dots, n\}$ si :*

1. $X(\Omega) = \{1, 2, 3, \dots, n\}$,
 2. $\forall i \in \{1, \dots, n\}, \mathbf{P}(X = i) = \frac{1}{n}$.
- On note alors : $X \sim \mathcal{U}(\{1, \dots, n\})$*

Il s'agit simplement, comme son nom l'indique, d'une loi dont tous les poids de probabilité sont identiques (équiprobabilité des événements élémentaires).

Propriété 1.7.1.

$$\mathbb{E}(X) = \frac{n+1}{2} \quad \text{et} \quad \mathbb{V}(X) = \frac{n^2-1}{12}.$$

Preuve. Pour l'espérance

$$\mathbb{E}(X) = \sum_{i=1}^n i \times \frac{1}{n} = \frac{1}{n} \sum_{i=1}^n i = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}.$$

Pour la variance

$$\mathbb{E}(X^2) = \sum_{i=1}^n i^2 \times \frac{1}{n} = \frac{1}{n} \sum_{i=1}^n i^2 = \frac{1}{n} \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6},$$

puis $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$, on obtient : $\mathbb{V}(X) = \frac{n^2-1}{12}$.

1.7.2 Loi de Bernoulli

Définition 1.7.2. Soit $p \in [0, 1]$. X suit la loi de Bernoulli de paramètre p si :

1. $X(\Omega) = \{0, 1\}$,
 2. $\mathbf{P}(X = 0) = 1 - p$ et $\mathbf{P}(X = 1) = p$.
- On note alors : $X \sim \mathcal{B}(p)$

Contexte usuel : on considère un espace probabilisé (Ω, \mathcal{A}, P) et on s'intéresse à un événement S , appelé "succès".

On définit alors une v.a.r.

$$X : \omega \mapsto \begin{cases} 1 & \text{si } \omega \in S \\ 0 & \text{sinon} \end{cases}$$

X est donc simplement l'indicatrice du succès S , et si on note $p = \mathbf{P}(S)$,

X suit une loi de Bernoulli de paramètre p .

Propriété 1.7.2. $\mathbb{E}(X) = p$ et $\mathbb{V}(X) = p(1 - p)$

Remarque 1.7.1. Les calculs des moments de X sont facilités par le fait que :

$$\forall n \geq 1, \quad X^n = X.$$

1.7.3 Loi binomiale

Définition 1.7.3. Soient $n \in \mathbb{N}^*$ et $p \in [0, 1]$. X suit la loi binomiale de paramètres n et p si :

1. $X(\Omega) = \{0, 1, \dots, n\}$,
 2. $\forall k \in \{0, 1, \dots, n\}, \quad \mathbf{P}(X = k) = C_k^n p^k (1 - p)^{n-k}$.
- On note alors : $X \sim \mathcal{B}(n, p)$.

Contexte usuel : comme pour la loi de Bernoulli, on considère une expérience aléatoire modélisée par un espace probabilisé (Ω, \mathcal{A}, P) et on s'intéresse à un événement S .

On note $p = \mathbf{P}(S)$.

On répète n fois cette expérience aléatoire de manière "indépendante", déterminons la loi de X . On a évidemment $X(\Omega) = \{0, 1, \dots, n\}$.

On calcule ensuite les poids de probabilité, en utilisant la définition de la probabilité produit :

$$\begin{aligned} \mathbf{P}(X = 0) &= \mathbf{P}(\bar{S}) \times \mathbf{P}(\bar{S}) \times \dots \times \mathbf{P}(\bar{S}) \text{ (n fois)} \\ &= (1 - p)^n \end{aligned}$$

$$\begin{aligned} \mathbf{P}(X = 1) &= \mathbf{P}(\{S, \bar{S}, \dots, \bar{S}\}) + \mathbf{P}(\{\bar{S}, S, \dots, \bar{S}\}) + \dots + \mathbf{P}(\{\bar{S}, \bar{S}, \dots, S\}) \text{ (n fois)} \\ &= \mathbf{P}(S)\mathbf{P}(\bar{S})\dots\mathbf{P}(\bar{S}) + \mathbf{P}(\bar{S})\mathbf{P}(S)\dots\mathbf{P}(\bar{S}) + \dots + \mathbf{P}(\bar{S})\mathbf{P}(\bar{S})\dots\mathbf{P}(S) \\ &= np(1 - p)^{n-1} \end{aligned}$$

Pour $k \in \{0, 1, \dots, n\}$: l'événement " $X = k$ " est la réunion n-uplets où S apparaît k fois et \bar{S} apparaît $n - k$ fois.

La probabilité de " $X = k$ " est la somme des probabilités de ces événements.

Ils ont tous pour probabilité $p^k(1 - p)^{n-k}$. On a C_n^k n-uplets comportant k fois S et $n - k$ fois \bar{S} , d'où :

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

Propriété 1.7.3. $\mathbb{E}(X) = np$ et $\mathbb{V}(X) = np(1 - p)$.

Preuve. La preuve est purement calculatoire :

Pour l'espérance :

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=0}^n k C_n^k p^k (1 - p)^{n-k} \\ &= \sum_{k=1}^n n C_{n-1}^{k-1} p^k (1 - p)^{n-k} \\ &= np \sum_{k=1}^n C_{n-1}^{k-1} p^{k-1} (1 - p)^{n-1-(k-1)} \\ &= np \end{aligned}$$

Même type de calcul pour la variance, en écrivant $k^2 = k(k - 1) + k$.

Exemple 1.7.1. Soit une urne contenant 6 boules noires et 5 boules rouges indiscernables au toucher. On procède à 10 extractions successives d'une boule, avec remise. Quelle est la loi du nombre de boules rouges obtenues ?

On répète 10 fois la même expérience (puisque les tirages se font avec remise) modélisée par $\Omega = \{N_1, N_2, \dots, N_6, R_1, R_2, \dots, R_5\}$, $A = \mathbf{P}(\Omega)$ et P la probabilité uniforme. Le "succès" est ici : "obtenir une boule rouge" soit $S = \{R_1, \dots, R_5\}$, et donc $\mathbf{P}(S) = \frac{5}{11}$. Par conséquent, en notant X la v.a.r. prenant pour valeur le nombre de boules rouges obtenues, on a :

$$X \sim \mathcal{B}\left(10, \frac{5}{11}\right).$$

1.7.4 Loi hypergéométrique

Définition 1.7.4. Soit $(N, n) \in \mathbb{N}^2$ tel que $n \leq N$ et soit $p \in]0, 1[$ tel que $Np \in \mathbb{N}$.

X suit la loi hypergéométrique de paramètres (N, n, p) si :

1. $X(\Omega) = \{\max(0, n - N(1 - p)), \min(n, Np)\}$.

2. $\forall k \in X(\Omega), \mathbf{P}(X = k) = \frac{C_{Np}^k C_{N(1-p)}^{n-k}}{C_N^n}$.

On note alors : $X \sim \mathcal{H}(N, n, p)$.

Contexte usuel : on considère une urne contenant N boules (indiscernables) : des boules rouges en proportion p (donc $Np \in \mathbb{N}$) et des boules noires en proportion $1 - p$.

On tire simultanément n boules (donc $n \leq N$), et on appelle X la v.a.r. égale au nombre de boules rouges obtenues.

Valeur minimale de X : s'il y a au moins n boules noires, c'est-à-dire si $N(1-p) \geq n$, X peut prendre la valeur 0.

Si le nombre de boules noires est strictement inférieur à n (i.e. $N(1-p) < n$), le nombre de boules rouges minimum que l'on puisse obtenir est $n - N(1-p)$ (on a obtenu toutes les boules noires possibles).

La valeur minimale de X est donc bien : $\max(0, n - N(1-p))$.

Valeur maximale de X : si le nombre de boules rouges est supérieur ou égal à n ($Np \geq n$), on peut obtenir au maximum n boules rouges. Dans le cas contraire ($Np < n$), on ne peut obtenir que Np boules rouges au maximum.

La valeur maximale de X est donc bien : $\min(n, Np)$.

Propriété 1.7.4.

$$\mathbb{E}(X) = np \quad \text{et} \quad \mathbb{V}(X) = \frac{N-n}{N-1} np(1-p).$$

1.7.5 Loi géométrique

Définition 1.7.5. Soit $p \in]0, 1]$. X suit la loi géométrique de paramètre p si :

1. $X(\Omega) = \mathbb{N}^*$
2. $\forall k \in \mathbb{N}^*, \mathbf{P}(X = k) = p(1-p)^{k-1}$

On note alors : $X \sim \mathcal{G}(p)$.

Contexte usuel on considère une expérience aléatoire modélisée par un espace probabilisé (Ω, \mathcal{A}, P) et on s'intéresse à un événement S , de probabilité $\mathbf{P}(S) = p$.

On répète une infinité de fois cette expérience. L'univers est donc $\Omega^{\mathbb{N}}$, c'est-à-dire l'ensemble des suites d'éléments de Ω . On appelle alors X la v.a.r. égale au rang d'apparition du premier succès (loi du "temps d'attente du premier succès"). On a donc bien $X(\Omega) = \mathbb{N}^*$ et :

$$\begin{aligned} \forall k \in \mathbb{N}^*, \mathbf{P}(X = k) &= \mathbf{P}(\underbrace{\bar{S}, \dots, \bar{S}}_{(k-1) \text{ fois}}, S, \dots) \\ &= \mathbf{P}(\bar{S}) \times \dots \times \mathbf{P}(\bar{S}) \times \mathbf{P}(S) \\ &= (1-p)^{k-1} p \end{aligned}$$

Propriété 1.7.5. $\mathbb{E}(X) = \frac{1}{p}$ et $\mathbb{V}(X) = \frac{1-p}{p^2}$

Preuve. Indication :

$$\forall x \in]-1, 1[, \sum_{k=0}^{+\infty} x^k = \frac{1}{1-x}.$$

$$\sum_{k=1}^{+\infty} kx^{k-1} = \frac{1}{(1-x)^2} \text{ et } \sum_{k=2}^{+\infty} k(k-1)x^{k-2} = \frac{2}{(1-x)^3}$$

1.7.6 Loi de Poisson

Définition 1.7.6. Soit $\lambda > 0$. X suit la loi de Poisson de paramètre λ si :

1. $X(\Omega) = \mathbb{N}$

2. $\forall k \in \mathbb{N}, \mathbf{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$.

On note alors : $X \sim \mathbf{P}(\lambda)$.

Contexte usuel : c'est la "loi limite" d'une loi binomiale lorsque np reste constant et n tend vers $+\infty$.

Soit $\lambda = np$. On considère une v.a.r. X qui suit une loi binomiale $\mathcal{B}(n, \frac{\lambda}{n})$:

$$\begin{aligned} \forall k \in \mathbb{N}, \mathbf{P}(X = k) &= C_n^k \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k n(n-1)\dots(n-k+1)}{k! n^k} e^{(n-k)\ln(1-\frac{\lambda}{n})} \\ &\xrightarrow{n \rightarrow +\infty} \frac{\lambda^k}{k!} e^{-\lambda} \end{aligned}$$

La loi de Poisson apparaît donc lorsque l'on compte le nombre de succès au cours d'un nombre d'épreuves très élevé, avec une probabilité p du succès très faible : c'est la loi des "événements rares".

Propriété 1.7.6. $\mathbb{E}(X) = \lambda$ et $\mathbb{V}(X) = \lambda$.

Preuve. On a bien la convergence des séries qui suivent (par exemple par la règle de D'Alembert) et :

$$\mathbb{E}(X) = \sum_{k=0}^{+\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^k}{(k-1)!} = e^{-\lambda} \lambda e^{\lambda} = \lambda$$

Pour la variance

$$\begin{aligned} \mathbb{E}(X^2) &= \sum_{k=0}^{+\infty} k^2 e^{-\lambda} \frac{\lambda^k}{k!} \\ &= \sum_{k=0}^{+\infty} (k(k-1) + k) e^{-\lambda} \frac{\lambda^k}{k!} \\ &= \sum_{k=0}^{+\infty} k(k-1) e^{-\lambda} \frac{\lambda^k}{k!} + \sum_{k=0}^{+\infty} k e^{-\lambda} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \sum_{k=2}^{+\infty} \frac{\lambda^k}{(k-2)!} + \lambda \\ &= e^{-\lambda} \lambda^2 e^{\lambda} + \lambda \\ &= \lambda^2 + \lambda \end{aligned}$$

D'où le résultat

$$\mathbb{V}(X) = \lambda$$

1.8 Travaux dirigés

Exercice 1.8.1. Montrer que la loi de probabilité d'une v.a.r. X (l'application μ) est une probabilité sur l'espace probablisable $(\mathbb{R}; \mathcal{B}_{\mathbb{R}})$.

Exercice 1.8.2. Soient X et Y deux v.a.r. de fonctions de répartition respectives F et G définies par :

$$F(x) = \begin{cases} 0 & \text{si } x < -1 \\ \frac{1}{3} & \text{si } -1 \leq x < 2 \\ 1 & \text{si } x \geq 2 \end{cases}$$

$$G(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ x^2 & \text{si } 0 \leq x \leq \frac{1}{2} \\ 1 - \frac{3}{4}e^{\frac{1}{2}-x} & \text{si } x \geq \frac{1}{2} \end{cases}$$

1. Déterminer les réels x pour lesquels $\mathbf{P}(X = x) > 0$ et les réels y pour lesquels $\mathbf{P}(Y = y) > 0$.

2. Soit $T = aX + b$ (a et b réels fixés, $a > 0$). Déterminer la fonction de répartition de T .

3. On suppose X et Y indépendantes, et on pose $Z = \max(X, Y)$, $U = \min(X, F)$. Déterminer la fonction de répartition de Z ainsi que celle de U .

Exercice 1.8.3. Un compteur devrait afficher les valeurs d'une variable aléatoire X suivant une loi $\mathcal{B}(n, p)$. Mais lorsque X est nulle, il affiche un nombre au hasard entre 1 et n . Lorsque X est non nulle, il affiche bien X . Soit Y la variable aléatoire prenant pour valeur le nombre affiché par le compteur.

-) Déterminer la loi de Y , ainsi que son espérance et sa variance.

Exercice 1.8.4.

X est une variable aléatoire qui suit une loi de Poisson de λ .

1) On désigne par Y la variable aléatoire prenant pour valeur :

$$\begin{cases} 0 & \text{lorsque } X = 2p + 1 \ (p \in \mathbb{N}) \\ p & \text{lorsque } X = 2p \ (p \in \mathbb{N}) \end{cases}$$

-) Calculer la probabilité de l'événement " X prend une valeur paire".

-) Déterminer la loi de Y ainsi que son espérance.

2) On note Z la variable : $Z = 4 \left[\frac{X}{2} \right] - 2X + 1$ ($[\]$ désigne la partie entière).

-) Déterminer la loi de Z , ainsi que son espérance et sa variance.

Exercice 1.8.5. Soit X la v.a.r. prenant pour valeurs comme suit :

$$\forall n \in \mathbb{N}^*, \quad \mathbf{P}(X = n) = \frac{4}{n(n+1)(n+2)}.$$

- 1) Vérifier que cette distribution, il s'agit bien de la distribution de probabilité d'une v.a.r.
- 2) Calculer lorsque c'est possible l'espérance et la variance de X .

Indication : Utiliser la décomposition en éléments simples :

$$\frac{4}{n(n+1)(n+2)} = \frac{a}{n} + \frac{b}{n+1} + \frac{c}{n+2}$$

Exercice 1.8.6.

Soit X la v.a.r. prenant pour valeurs les entiers naturels dans $X(\Omega) = \{3k, 3k+1 \text{ tel que } k \in \mathbb{N}\}$, et dont la loi de probabilité est donnée par :

$$\forall n \in X(\Omega), \quad \mathbf{P}(X = n) = \alpha 3^{-n}, \quad \text{où } \alpha \in \mathbb{R}.$$

1. Calculer α .
 2. Calculer, lorsque c'est possible, l'espérance et la variance de X .
- Indication :** Utiliser la série géométrique dérivée une fois :

$$\forall q \in]-1, 1[\left[\sum_{n=1}^{+\infty} nq^{n-1} = \frac{1}{(1-q)^2} \right.$$
