Variables aléatoires discrètes

Dans tout ce chapitre, même si ce n'est pas rappelé à chaque fois, on suppose que $(\Omega, \mathscr{T}, \mathbb{P})$ est un espace probabilisé.

Dans ce chapitre, si E est un ensemble fini ou dénombrable et $(u_x)_{x\in E}$ une famille de réels indexée par E, la notation $\sum_{x\in E} u_x$ désigne la somme au sens des familles sommables ou des familles positives selon les cas.

1 Variables aléatoires discrètes

1.1 Définition

Définition. On appelle variable aléatoire discrète (en abrégé v.a.d.) toute application X de Ω dans un ensemble quelconque telle que son ensemble-image $X(\Omega)$ soit **fini ou dénombrable** et telle que pour tout $x \in X(\Omega)$, $X^{-1}(\{x\}) \in \mathscr{T}$.

On appelle variable aléatoire discrète réelle toute variable aléatoire discrète à valeurs dans $\mathbb{R}$, v.a.d. complexe quand elle est à valeurs dans $\mathbb{C}$.

En général, on note avec des lettres majuscules droites les variables aléatoires : X, Y, S, etc.

Si $U \subset X(\Omega)$, on note plutôt $X^{-1}(U)$ sous la forme préférée des probabilistes $\{X \in U\}$ ou même $(X \in U)$.

Remarque.

— Comme son nom ne l'indique pas, une variable aléatoire n'est pas une variable, mais une fonction ! La terminologie a été fixée à une époque ancienne où la notion n'était pas encore parfaitement claire.

— Toute fonction constante est une variable aléatoire, appelé variable aléatoire certaine.

— Si A est un événement, on note 1_A l'application de Ω dans $\mathbb{R}$ définie par :
$$1_A(\omega) = 1 \text{ si } \omega \in A \text{ et } 1_A(\omega) = 0 \text{ sinon.}$$
La v.a.r. 1_A est appelé fonction indicatrice de A.

1.2 Probabilité-image d'une v.a.d.

Proposition 1. *Soit X une v.a.d. à valeurs dans E.*

À toute partie U de $X(\Omega)$, on associe $\mathbb{P}_X(U) = \mathbb{P}(\{X \in U\}) = \mathbb{P}(X \in U)$.

Alors $\mathbb{P}_X$ est une probabilité sur $(X(\Omega), \mathscr{P}(X(\Omega)))$, appelé probabilité-image de X ou loi de X.

Remarque. L'intérêt de la notion de variable aléatoire est de déplacer les calculs de probabilité dans l'univers Ω souvent inconnu, muni d'une tribu inconnue dans un ensemble fini ou dénombrable $X(\Omega)$ bien plus agréable, muni de la tribu simple $\mathscr{P}(X(\Omega))$!

Dans notre pratique des probabilités, nous supposerons toujours l'existence des variables aléatoires qu'on considère ! Voyons néanmoins un résultat de réalisation effective.

1.3 Loi d'une v.a.r.

On rappelle qu'une distribution discrète (de probabilités) est une famille de réels positifs $(p_x)_{x\in L}$ où L est un ensemble fini ou dénombrable (non vide) telle que $\sum_{x\in L} p_x = 1$.

Les exemples suivants sont fondamentaux et sont à connaître.

Exemples. Avec L un ensemble fini :

— distribution uniforme sur L : $\mathscr{U}(L) = (p_x)_{x\in L}$ où $p_x = \dfrac{1}{\operatorname{card} L}$

— distribution de Bernoulli de paramètre $p \in [0,1]$: $\mathscr{B}(p) = (p_0, p_1)$ où $p_1 = p$, $p_0 = 1 - p$

— distribution binomiale de paramètre $(n,p) \in \mathbb{N} \times [0,1]$: $\mathscr{B}(n,p) = (p_k)_{k\in[\![0,n]\!]}$ où $p_k = \dbinom{n}{k} p^k (1-p)^{n-k}$

Exemples. Avec L dénombrable :

— distribution géométrique de paramètre $p \in]0,1[$: $\mathscr{G}(p) = (p_x)_{x\in\mathbb{N}^*}$ où $p_x = (1-p)^{x-1}p$

— distribution de Poisson de paramètre $\lambda > 0$: $\mathscr{P}(\lambda) = (p_k)_{k\in\mathbb{N}}$ où $p_k = e^{-\lambda}\dfrac{\lambda^k}{k!}$

Définition. On dit qu'une v.a.d. X sur Ω suit la loi discrète $\mathscr{L}$ quand

$$X(\Omega) \subset L \text{ et } \forall x \in L \quad \mathbb{P}(X = x) = p_x$$

où $(p_x)_{x\in L}$ est la distribution de probabilités nommée par le même nom que $\mathscr{L}$.

Quand on demande la loi d'une variable aléatoire X, on demande donc de déterminer un ensemble L tel que $X(\Omega) \subset L$ et pour chaque $a \in L$, la valeur de $\mathbb{P}(X = a)$. En général, on a souvent $X(\Omega) = L$, mais parfois ce n'est pas ce qui est le plus simple.

Par abus de langage, on confond la loi et la distribution de probabilités associée.

Proposition 2. *Soit X une variable aléatoire qui suit la loi $\mathscr{L} = (p_x)_{x\in L}$.*

La famille d'événements $(\{X = x\})_{x\in L}$ est un système quasi-complet d'événements, appelé système quasi-complet d'événements associé à X.

En particulier, $\sum_{x\in L} \mathbb{P}(X = x) = 1$.

Exercices **:**

1) Une urne contient $n+1$ boules numérotées de 0 à n. On procède à deux tirages avec remise et on note S la somme des numéros des deux boules. Déterminez la loi de S.
2) On lance une infinité de fois une pièce dont la probabilité de tomber sur *pile* vaut $p \in]0,1[$. On note X la variable aléatoire qui compte le nombre de lancers nécessaires pour obtenir une première fois *pile*. Déterminez la loi de X en faisant des hypothèses raisonnables sur les lancers.

Théorème de réalisation.

Théorème 1. *Soit $\mathscr{L}$ une distribution discrète de probabilités.*

Alors il existe un espace probabilisé $(\Omega, \mathscr{T}, \mathbb{P})$ et une variable aléatoire X sur Ω tels que X suive la loi $\mathscr{L}$.

On note $X \sim \mathscr{L}$ pour signifier que X suit la loi $\mathscr{L}$, ou que $X \sim Y$ pour signifier que X et Y suivent la même loi (ce qui ne présuppose rien sur les univers de départ de X et Y : on a seulement $X(\Omega) = Y(\Omega')$ et les égalités numériques $\forall x \in X(\Omega) \quad \mathbb{P}(X = x) = \mathbb{P}(Y = x)$).

1.4 Cas des v.a.d. à valeurs réelles ou complexes

Proposition 3. *L'ensemble des v.a.d. réelles sur l'espace probabilisé $(\Omega, \mathscr{T}, \mathbb{P})$ est un $\mathbb{R}$-e.v. : on peut donc additionner, multiplier par un scalaire des v.a.d. réelles. On peut aussi les multiplier entre elles.*

Si X est une v.a.d. réelle sur Ω et si f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$ définie sur $X(\Omega)$, alors la composée $f \circ X$ est une v.a.d. réelle définie sur Ω, qu'on note plutôt $f(X)$.

On a bien sûr le même résultat avec les variables complexes.

La loi de $f(X)$ est donc théoriquement donnée par :

$$\text{si } y \in f(X)(\Omega), \text{ alors } \mathbb{P}(f(X) = y) = \sum_{x\in f^{-1}(\{y\})\cap X(\Omega)} \mathbb{P}(X = x).$$

En pratique, cette loi est souvent difficilement calculable.

Exercices **:**

3) Soit X une v.a.r. suivant une loi de Poisson $\mathscr{P}(\lambda)$ et $p \in \mathbb{N}^*$. On pose $Y = X^2 - (2p+1)X$. Déterminez la loi de Y.

2 Espérance

2.1 Définitions

Définition. Soit X une v.a.d. à valeurs dans $[0,+\infty]$.
On appelle alors espérance de X le nombre $\mathbb{E}(X) = \displaystyle\sum_{x\in X(\Omega)} x\,\mathbb{P}(X=x)$ (au sens des familles positives) : c'est un réel si la famille $(x\,\mathbb{P}(X=x))_{x\in X(\Omega)}$ est sommable, $+\infty$ sinon.

Une variable réelle positive possède donc toujours une espérance, éventuellement infinie. Dans le cas général, on ne peut pas toujours définir l'espérance.

Définition. Soit X une v.a.d. complexe.
On dit que X possède une espérance finie quand la famille $(x\,\mathbb{P}(X=x))_{x\in X(\Omega)}$ est sommable. Dans ce cas, on appelle espérance de X le nombre $\mathbb{E}(X) = \displaystyle\sum_{x\in X(\Omega)} x\,\mathbb{P}(X=x)$ (au sens des familles sommables).

On note L^1 l'ensemble des variables à espérance finie. On voit souvent écrit $X \in L^1$ pour signifier en abrégé que la variable X possède une espérance finie.

L'espérance de X est donc la moyenne de ses valeurs possibles, pondérées par leurs probabilités respectives, quand cela a un sens.

Exemples.
- L'espérance d'une v.a.r. certaine égale à a est a.
- Si A est un événement, l'espérance de son indicatrice est sa probabilité : $\mathbb{E}(1_A) = \mathbb{P}(A)$.

***Exercices* :**

4) Soit $\lambda \in \mathbb{R}$ et $\alpha > 0$. Déterminez une condition sur λ et α pour qu'existe une v.a.r. X qui suit une loi $\left(\dfrac{\lambda}{n^\alpha}\right)_{n\in\mathbb{N}^*}$.
À quelle condition X possède-t-elle une espérance finie ?
5) Calculez l'espérance d'une variable X qui suit une loi géométrique de paramètre $p \in\,]0,1[$.

Proposition 4. *Toute v.a. finie possède une espérance finie.*
Toute v.a.d. complexe bornée possède une espérance finie.

2.2 Propriétés

Une propriété qui découle des th. de comparaison classiques sur les familles sommables.

Proposition 5. *Soit X, Y deux v.a.d. complexes.*
Si $|X| \leqslant |Y|$ et $Y \in L^1$, alors $X \in L^1$.
Si $X \in L^1$, alors $|X| \in L^1$ et $|\mathbb{E}(X)| \leqslant \mathbb{E}(|X|)$ (inégalité triangulaire).

La linéarité découle encore de la théorie des familles sommables.

Proposition 6. *Soit X, Y deux v.a.d. complexes admettant chacune une espérance finie, λ un complexe. Alors $X+Y$ et λX admettent une espérance finie et*

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

En particulier, pour tout $(a,b)\in\mathbb{C}^2$, $\mathbb{E}(aX+b) = a\mathbb{E}(X) + b$.

Si X est une v.a.r. positive, alors $\mathbb{E}(X) \geqslant 0$.
Si X et Y sont deux v.a.r. telles que $X \leqslant Y$, alors $\mathbb{E}(X) \leqslant \mathbb{E}(Y)$.

Proposition 7. *Soit X une v.a.d. réelle positive.*
Alors X a une espérance nulle si et s.si X est presque sûrement nulle, i.e. $\mathbb{P}(X=0)=1$.

Définition. On dit qu'une v.a.d. complexe X est centrée si et seulement si $\mathbb{E}(X)=0$.

À toute v.a.d. complexe X, on associe une v.a.d. complexe centrée : $X-\mathbb{E}(X)$.

2.3 Théorème de transfert

Théorème 2. *Soit X une v.a.d. complexe. Soit f une fonction de $\mathbb{C}$ dans $\mathbb{C}$ définie sur $X(\Omega)$.*
Alors la v.a.d. complexe $f(X)$ possède une espérance finie si et s.si la famille $(f(x)\,\mathbb{P}(X=x))_{x\in X(\Omega)}$ est sommable et dans ce cas, on a l'égalité

$$\mathbb{E}(f(X))=\sum_{x\in X(\Omega)} f(x)\,\mathbb{P}(X=x)$$

Ce théorème permet de calculer directement l'espérance de $f(X)$ sans devoir calculer la loi de $f(X)$: il suffit de connaître celle de X.

Exercices :

6) Soit $p\in\,]0,1[$, $X\sim\mathscr{G}(p)$. Calculez $\mathbb{E}(e^{-X})$.

Enfin, un petit résultat parfois utile.

Proposition 8. *Soit X une v.a.d. à valeurs dans $\mathbb{N}$.*
Alors $\mathbb{E}(X)=\sum_{n=1}^{+\infty}\mathbb{P}(X\geqslant n)$.

2.4 Inégalité de Markov

Proposition 9. *Soit X une v.a.d. réelle positive admettant une espérance finie. Alors*

$$\forall a>0\quad \mathbb{P}(X\geqslant a)\leqslant\frac{\mathbb{E}(X)}{a}$$

Cette inégalité n'a d'intérêt que pour $a>\mathbb{E}(X)$, sinon on majore une probabilité par un nombre plus grand que 1.

Il arrive souvent qu'on ne sache pas calculer explicitement la loi d'une variable aléatoire, mais qu'on arrive à calculer son espérance (par exemple comme somme de variables aléatoires simples). Les inégalités de ce type permettent quand même parfois d'obtenir des résultats à propos des probabilités de certains événements.

Corollaire 1. *Soit X une v.a.d. complexe admettant une espérance finie. Alors*

$$\forall a>0\quad \mathbb{P}(|X|\geqslant a)\leqslant\frac{\mathbb{E}(|X|)}{a}$$

3 Variance d'une variable réelle

3.1 Moments d'ordre 2

Proposition 10. *Soit X une v.a.d. réelle.*
Si X^2 possède une espérance finie (on dit aussi que X admet un moment d'ordre 2), alors X possède aussi une espérance finie. Dans ce cas, pour tout $a\in\mathbb{R}$, $(X+a)^2$ possède une espérance finie.

On note L^2 l'ensemble des v.a.d. réelle X telles que X^2 est d'espérance finie, *i.e.* $\mathbb{E}(X^2)<+\infty$.

Proposition 11. *Soit X, Y deux v.a.d. réelles dans L^2.*
Alors XY est dans L^1 et $\mathbb{E}(XY)^2 \leqslant \mathbb{E}(X^2)\,\mathbb{E}(Y^2)$

Il s'agit bien sûr de l'inégalité de Cauchy-Schwarz.

3.2 Variance et écart-type

Définition. Soit X une v.a.d. réelle telle que $X \in L^2$.
On appelle variance de X le nombre $\mathbb{V}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right)$.
On appelle écart-type de X le réel $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

La variance (ou l'écart-type) mesure la dispersion de X autour de sa moyenne. Une variable de variance 1 est dite réduite.

Proposition 12. *Soit X une v.a.d. réelle telle que $X \in L^2$.*
Alors $\mathbb{V}(X) = 0$ si et s.si X est presque sûrement constante.

En général, on calcule la variance par la formule suivante.

Proposition 13. *(Formule de Huyghens)*
Soit X une v.a.d. réelle possédant une variance finie. Alors

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

Exercices :

7) Calculez la variance d'une variable suivant une loi géométrique de paramètre $p \in]0,1[$.

Proposition 14. *Si a, b sont deux réels et X est une v.a.d. réelle dans L^2, alors $aX+b \in L^2$ et $\mathbb{V}(aX+b) = a^2\mathbb{V}(X)$*

Remarque. Si X est une v.a.d. réelle d'espérance m et de variance v, on pose $X' = \dfrac{1}{\sqrt{v}}(X - m)$.
X' est alors une v.a.d. réelle centrée dont la variance vaut 1 : on l'appelle la v.a.d. réelle centrée réduite associée à X.

3.3 Inégalité de Bienaymé-Tchebychev

Proposition 15. *Soit X une v.a.d. réelle possédant un moment d'ordre 2. Alors*

$$\forall \varepsilon > 0 \quad \mathbb{P}\big(|X - \mathbb{E}(X)| \geqslant \varepsilon\big) \leqslant \frac{\mathbb{V}(X)}{\varepsilon^2}$$

La même remarque que pour l'inégalité de Markov s'applique : cette inégalité n'a d'intérêt que pour des valeurs assez grandes de ε, sinon on majore une probabilité par 1.

Exercices :

8) On lance n fois un dé à 6 faces. Comment choisir n pour que la probabilité d'obtenir moins d'une fois sur deux le nombre 6 soit au moins égale à $\dfrac{3}{4}$?

3.4 Généralisation

Définition. Soit X une v.a.d. réelle, $k \in \mathbb{N}^*$.
On dit que X admet un moment d'ordre k quand X^k a une espérance finie.

Proposition 16. *Soit X une v.a.d. réelle, $k \in \mathbb{N}^*$.*
Si X possède un moment d'ordre k, alors pour tout $\ell \in [\![1, k]\!]$, X possède un moment d'ordre ℓ.

4 Lois classiques

4.1 Loi uniforme

Définition. Soit $(\Omega, \mathbb{P})$ un espace probabilisé et L un ensemble fini non vide.
On dit qu'une variable aléatoire suit la loi uniforme sur L quand $X(\Omega) = L$ et $\mathbb{P}_X$ est la probabilité uniforme sur L, autrement dit si pour tout $x \in L$, $\mathbb{P}(X = x) = \dfrac{1}{\text{card}(L)}$.
On note alors $X \sim \mathscr{U}(L)$.

Le cas le plus courant est celui des variables à valeurs entières.

Définition. Soit $(\Omega, \mathbb{P})$ un espace probabilisé et a, b deux entiers tels que $a \leqslant b$.
On dit qu'une v.a.r. suit la loi uniforme sur $[\![a, b]\!]$ quand $X(\Omega) = [\![a, b]\!]$ et $\mathbb{P}_X$ est la probabilité uniforme sur $[\![a, b]\!]$, autrement dit si pour tout $k \in [\![a, b]\!]$, $\mathbb{P}(X = k) = \dfrac{1}{b - a + 1}$.
On note alors $X \sim \mathscr{U}([\![a, b]\!])$.

Dans ce cas, on a $\mathbb{E}(X) = \dfrac{a+b}{2}$ et $\mathbb{V}(X) = \dfrac{(b-a)(b-a+2)}{12}$.
Souvent, on a $a = 1$ et $b = n$, donc dans ce cas, $\mathbb{E}(X) = \dfrac{n+1}{2}$ et $\mathbb{V}(X) = \dfrac{n^2 - 1}{12}$.
Exemple. Si on note X le nombre obtenu après un lancer de dé non truqué, X suit la loi uniforme sur $[\![1, 6]\!]$.

4.2 Loi de Bernoulli

Définition. Soit $p \in [0, 1]$. On dit qu'une v.a.r. X suit la loi de Bernoulli de paramètre p quand $X(\Omega) \subset \{0, 1\}$ et $\mathbb{P}(X = 1) = p$.
On note alors $X \sim \mathscr{B}(p)$.

Dans ce cas, on note souvent $q = 1 - p$: $\mathbb{E}(X) = p$ et $\mathbb{V}(X) = pq$.
Exemples.
— Toute expérience aléatoire à deux issues peut être représentée par une variable de Bernoulli en notant 0 et 1 les deux issues. Le cas typique est le lancer d'une pièce (équilibrée si $p = 1/2$, non équilibrée sinon).
— En particulier toute expérience dont seul la réussite ou l'échec importe peut être représentée par une variable de Bernoulli : 1 représente la réussite, 0 l'échec.

4.3 Loi binomiale

Définition. Soit $p \in [0, 1]$. On dit qu'une v.a.r. X suit la loi binomiale de paramètre (n, p) quand $X(\Omega) \subset [\![0, n]\!]$ et pour tout $k \in [\![0, n]\!]$, $\mathbb{P}(X = k) = \dbinom{n}{k} p^k (1-p)^{n-k}$.
On note alors $X \sim \mathscr{B}(n, p)$.

En notant $q = 1 - p$: $\mathbb{E}(X) = np$ et $\mathbb{V}(X) = npq$.

Proposition 17. *On considère une suite de n expériences aléatoires indépendantes qui suivent une loi de Bernoulli de même paramètre p. On note X le nombre de réussites dans cette répétition d'expériences, appelée schéma de Bernoulli. Alors la loi de X est la loi binomiale de paramètre (n, p).*

Exemples.
— On lance n fois une pièce dont la probabilité de tomber sur *pile* vaut p. Alors si X est le nombre de fois où on tombe sur *pile*, X suit une loi binomiale de paramètre (n, p).
— On fait n tirages successifs avec remise dans une urne contenant une proportion p de boules blanches. Le nombre de boules blanches tirées suit la loi binomiale de paramètre (n, p).

4.4 Loi géométrique

Définition. Soit $p \in]0,1[$. On dit qu'une v.a.d. réelle X suit la loi géométrique de paramètre p quand $X(\Omega) \subset \mathbb{N}^* \cup \{+\infty\}$ et pour tout $k \in \mathbb{N}^*$, $\mathbb{P}(X=k) = (1-p)^{k-1}p$.

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

En notant $q = 1-p$: $\mathbb{P}(X=+\infty) = 0$, $\mathbb{E}(X) = \dfrac{1}{p}$, $\mathbb{V}(X) = \dfrac{q}{p^2}$.

Proposition 18. *La loi géométrique de paramètre p est la loi du rang du premier succès dans un schéma de Bernoulli infini de paramètre p.*

Exemples.

— On lance une pièce dont la probabilité de tomber sur *pile* vaut p. Alors si X est le premier rang pour lequel le lancer donne *pile*, X suit une loi géométrique de paramètre p.

— On fait des tirages successifs avec remise dans une urne contenant une proportion p de boules blanches. Le premier rang où on tire une boule blanche suit la loi géométrique de paramètre p.

4.5 Loi de Poisson

Définition. Soit $\lambda > 0$. On dit qu'une v.a.d. réelle X suit la loi de Poisson de paramètre λ quand $X(\Omega) = \mathbb{N}$ et pour tout $k \in \mathbb{N}$, $\mathbb{P}(X=k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$.

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

Dans ce cas, $\mathbb{E}(X) = \lambda$, $\mathbb{V}(X) = \lambda$.

La loi de Poisson est la loi qui compte le nombre d'événements dans un intervalle de temps fixé, quand ces événements se produisent à une certaine fréquence connue et indépendamment du temps écoulé depuis le précédent. Le paramètre λ est alors le nombre moyen d'événements se produisant durant la durée fixée et X compte le nombre d'événements qui se sont effectivement produits durant la durée fixée.

Par exemple, si des statistiques montrent que le passage à une borne de péage se fait à une fréquence de f véhicules par heure, alors on choisit souvent la loi de Poisson $\mathscr{P}(f \times h)$ pour compter le nombre de véhicules qui passe dans une durée de h heures : $\mathbb{P}(X=k) = e^{-f.h}\dfrac{(f.h)^k}{k!}$ est la probabilité pour que sur un intervalle de temps de h heures, k voitures soient effectivement passées à la borne.

Souvent la fréquence f est faible, on utilise cette loi pour l'étude des événements dits « rares ».

5 Couples de variables aléatoires

5.1 Généralités

Définition. Soit X, Y deux v.a.d. La fonction $\omega \mapsto (X(\omega), Y(\omega))$ est appelée couple de variables aléatoires. Si X et Y sont deux v.a.r. le couple est dit couple de v.a.d. réelle

On peut reprendre le même schéma de présentation que pour une seule v.a.r.

Proposition 19. *Soit (X, Y) un couple de v.a.d.*
Alors la famille d'événements $(\{X = x\} \cap \{Y = y\})_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est un système complet d'événements appelé s.c.e. associé au couple (X, Y).

La probabilité $\mathbb{P}(\{X = x\} \cap \{Y = y\})$ est souvent notée $\mathbb{P}(X = x, Y = y)$.

Définition. Soit X, Y deux v.a.d. La loi conjointe de X et Y est la loi du couple (X, Y).

Dans le cas d'univers fini, on la représente souvent par un tableau à double entrée : si $X(\Omega) = \{x_1, \dots, x_m\}$ et $Y(\Omega) = \{y_1, \dots, y_n\}$, alors on place $\mathbb{P}(X = x_i, Y = y_j)$ sur la i-ème ligne et la j-ème colonne.

5.2 Lois marginales

Définition. Soit X, Y deux v.a.d. Les lois marginales du couple (X, Y) sont les lois de X et de Y.

Proposition 20. *Soit X, Y deux v.a.d. Alors*
- ▷ *pour tout $x \in X(\Omega)$, $\mathbb{P}(X = x) = \displaystyle\sum_{y\in Y(\Omega)} \mathbb{P}(X = x, Y = y)$,*
- ▷ *pour tout $y \in Y(\Omega)$, $\mathbb{P}(Y = y) = \displaystyle\sum_{x\in X(\Omega)} \mathbb{P}(X = x, Y = y)$.*

Si on a représenté la loi conjointe de (X, Y) sous forme d'un tableau, on obtient les lois marginales en additionnant les probabilités sur chaque ligne ou chaque colonne (d'où le nom « lois marginales » : celle qu'on note en marge du tableau).

***Exercices* :**

9) Une urne contient n boules numérotées de 1 à n. On tire une boule avec remise : on note X le numéro de la boule. Puis on tire X boules et on note Y le maximum des numéros des boules tirées. Déterminez la loi de X, la loi conjointe du couple (X, Y) et déduisez-en la loi de Y.

Remarque. La loi d'une variable aléatoire ne peut pas dépendre d'une autre variable aléatoire !

Écrire par exemple que $Y(\Omega) = [\![X, n]\!]$ dans l'exercice précédent est un non-sens, car X n'est pas un nombre, mais une variable aléatoire (*i.e.* une fonction !).

Écrire $Y(\Omega) = [\![i, n]\!]$ où i est la valeur de X après le premier tirage est tout aussi absurde !

Il faut imaginer que l'expérience a lieu dans un lieu qui vous est caché et que l'opérateur ne vous donne que la valeur de Y. On laisse l'opérateur répéter son expérience un grand nombre de fois et on attend qu'il vous donne les valeurs de Y pour chacune des expériences. Qu'allez-vous obtenir ? Toutes les valeurs possibles de Y, qui peut varier de 1 à n. Certainement pas quelque chose qui dépend de X, puisque vous ne savez même pas quelle expérience réelle a lieu dans le lieu caché : pour vous, observateur extérieur, X n'existe pas.

5.3 Lois conditionnelles

Définition. Soit X, Y deux v.a.d.
Pour tout $y \in Y(\Omega)$ tel que $\mathbb{P}(Y = y) \neq 0$, la loi de X sachant $(Y = y)$ est la loi de X dans l'espace probabilisé $(\Omega, \mathbb{P}_{(Y=y)})$:

pour tout $x \in X(\Omega)$, $\mathbb{P}_{(Y=y)}(X = x) = \dfrac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)}$.

On déduit de toutes les définitions précédentes les relations suivantes.

Proposition 21. *Soit X, Y deux v.a.d.*
On suppose que pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, $\mathbb{P}(X = x) \neq 0$ et $\mathbb{P}(Y = y) \neq 0$. Alors pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$,

- ▷ $\mathbb{P}(X = x, Y = y) = \mathbb{P}(Y = y) \times \mathbb{P}_{(Y=y)}(X = x) = \mathbb{P}(X = x) \times \mathbb{P}_{(X=x)}(Y = y)$
- ▷ $\mathbb{P}(X = x) = \displaystyle\sum_{y \in Y(\Omega)} \mathbb{P}(Y = y) \times \mathbb{P}_{(Y=y)}(X = x)$
- ▷ $\mathbb{P}(Y = y) = \displaystyle\sum_{x \in X(\Omega)} \mathbb{P}(X = x) \times \mathbb{P}_{(X=x)}(Y = y)$

5.4 Covariance

Proposition 22. *Soit X, Y deux v.a.d. réelles ayant une variance finie.*
Alors XY possède une espérance finie.

Définition. Dans les conditions précédentes, on appelle covariance du couple (X, Y) le nombre

$$\mathrm{Cov}(X, Y) = \mathbb{E}\Big((X - \mathbb{E}(X)).(Y - \mathbb{E}(Y))\Big)$$

Si de plus, $\mathbb{V}(X)$ et $\mathbb{V}(X)$ sont non nulles, on appelle coefficient de corrélation du couple (X, Y) le nombre

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sigma(X)\sigma(Y)}$$

Comme pour la variance, on a une expression plus simple.

Proposition 23. *Avec les mêmes hypothèses,* $\mathrm{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$.

L'application Cov a des propriétés classiques de bilinéarité.

Proposition 24.
Cov *est une application bilinéaire.*
De plus, pour tout couple de v.a.d. réelles (X, Y) ayant une variance finie,

$$\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\,\mathrm{Cov}(X, Y)$$

Proposition 25 (Inégalité de Cauchy-Schwarz)**.** *Pour tout couple de v.a.d. réelles (X, Y) ayant une variance finie, $|\mathrm{Cov}(X, Y)| \leqslant \sigma(X)\sigma(Y)$.*
Si de plus $\sigma(X)$ et $\sigma(X)$ sont non nuls, $\rho(X, Y) \in [-1, +1]$ et $|\rho(X, Y)| = 1$ si et s.si il existe $(a, b) \in \mathbb{R}^2$ tel que $Y = aX + b$ presque sûrement.

6 Indépendance de variables aléatoires

6.1 Généralités

Définition. Soit X, Y deux v.a.d.
On dit que X et Y sont indépendantes quand pour toute partie A de $X(\Omega)$ et toute partie B de $Y(\Omega)$, les événements $\{X \in A\}$ et $\{Y \in B\}$ sont indépendants,
c'est-à-dire quand $\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A) \times \mathbb{P}(Y \in B)$.

Dans ce cas, on note $X \perp\!\!\!\perp Y$ pour signaler que X et Y sont indépendantes.

On peut se restreindre aux événements des s.c.e. associés à X et Y (résultat admis).

Proposition 26. *Soit X, Y deux v.a.d.*
X et Y sont indépendantes si et seulement si pour toute partie $(x, y) \in X(\Omega) \times Y(\Omega)$, les événements $\{X = x\}$ et $\{Y = y\}$ sont indépendants, c'est-à-dire si $\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \times \mathbb{P}(Y = y)$.

Remarque. En général, la connaissance des lois marginales de (X, Y) ne permet pas de retrouver la loi conjointe de (X, Y).

Dans le cas où les variables aléatoires X et Y sont indépendantes, c'est possible.

Donc pour montrer que deux variables X et Y ne sont pas indépendantes, il suffit de trouver deux valeurs x, y de X, Y resp. telles que $\mathbb{P}(X = x, Y = y) \neq \mathbb{P}(X = x) \times \mathbb{P}(Y = y)$.

On a un résultat sur les composées de variable indépendantes.

Proposition 27. *Soit X, Y deux v.a.d.*
Si X et Y sont indépendantes, alors pour toute fonction f définie sur $X(\Omega)$ et toute fonction g définie sur $Y(\Omega)$, les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

6.2 Espérance et indépendance

On a un résultat remarquable sur les espérances de v.a.r. indépendantes (résultat admis).

Proposition 28. *Soit X, Y deux v.a.d..*
Si X et Y sont indépendantes, alors

- ▷ $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$

  et dans le cas de variables réelles,
- ▷ $\mathrm{Cov}(X, Y) = 0$
- ▷ $\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y)$.

Remarque. La réciproque est fausse.

Deux v.a.d. réelles de covariance nulle sont dites « non corrélées », mais c'est un renseignement très faible sur les variables, à la différence de l'indépendance, qui est une contrainte extrêmement forte.

6.3 Généralisation

Tous les résultats sont admis.

Définition. Soit $(X_i)_{i\in I}$ une famille finie ou dénombrable de v.a.d..
On dit que les variables aléatoires $(X_i)_{i\in I}$ sont mutuellement indépendantes quand pour toute partie finie $J \subset I$, pour toute famille de parties $(A_j)_{j\in J}$ de $\displaystyle\prod_{j\in J} \mathscr{P}(X_j(\Omega))$, les événements $(\{X_j \in A_j\})_{j\in J}$ sont indépendants.

On retrouve la même caractérisation à l'aide des événements des s.c.e. associés au différentes variables aléatoires et on peut même faire plus simple !

Proposition 29. *Soit $(X_i)_{i\in I}$ une famille finie ou dénombrable de v.a.d..*
Les variables aléatoires $(X_i)_{i\in I}$ sont mutuellement indépendantes quand pour toute partie finie $J \subset I$, pour toute famille $(x_j)_{j\in J}$ de $\prod_{j\in J} X_j(\Omega)$, $\mathbb{P}\left(\bigcap_{j\in J}\{X_j = x_j\}\right) = \prod_{j\in J}\mathbb{P}(X_j = x_j)$.

Exercices :

10) Soit (X_1, X_2) et (Y_1, Y_2) deux couples de variables indépendantes telles que pour $i \in \{1, 2\}$, $X_i \sim Y_i$.
Montrez que $X_1 + X_2 \sim Y_1 + Y_2$. Le résultat est-il encore valable en supprimant l'hypothèse d'indépendance ?

Enfin, un petit lemme classique : le lemme des coalitions.

Proposition 30. *Soit $(X_i)_{i\in I}$ une famille finie de v.a.d. réelles indépendantes.*
Soit J, K deux parties de I, non vides et disjointes, f une fonction de $\mathbb{R}^{\operatorname{card} J}$ dans $\mathbb{R}$, g une fonction de $\mathbb{R}^{\operatorname{card} K}$ dans $\mathbb{R}$.
Alors les deux v.a.d. réelles $f(X_j, j \in J)$ et $g(X_k, k \in K)$ sont indépendantes.

En clair, si on partage des v.a.d. réelle indépendantes en deux sous-ensembles n'ayant aucune variable en commun, alors n'importe quoi qui dépend uniquement des variables du premier ensemble est indépendant de n'importe quoi qui dépend uniquement des variables du deuxième ensemble.

Le résultat se généralise à toute famille dénombrable de v.a.d. réelles indépendantes, puisque la définition même d'indépendance impose de ne considérer que des sous-familles finies.

Exemple. Un cas très courant : si $(X_n)_{n\in\mathbb{N}}$ est une famille de v.a.d. réelles indépendantes, alors

— pour tout $n \in \mathbb{N}$, $\sum_{k=0}^{n} X_k$ est indépendante de X_{n+1} ;

— pour tout $(n, p) \in \mathbb{N}^2$ tel que $n < p$, $\sum_{k=0}^{n} X_k$ est indépendante de $\sum_{k=n+1}^{p} X_k$.

On en déduit la généralisation du résultat sur les variances.

Proposition 31. *Soit $X_1, \ldots, X_n$ n v.a.d. réelles indépendantes.*
Alors $\mathbb{V}(X_1 + \ldots + X_n) = \mathbb{V}(X_1) + \ldots + \mathbb{V}(X_n)$.

6.4 Théorème de réalisation

Le théorème de Kolmogorov est un théorème de réalisation sur les familles de variables aléatoires indépendantes (admis et jamais utilisé en pratique).

Théorème 3. *Soit $(\mathscr{L}_i)_{i\in I}$ une famille finie ou dénombrable de lois de probabilités.*
Alors il existe un espace probabilisé $(\Omega, \mathscr{T}, \mathbb{P})$ et des variables aléatoires $(X_i)_{i\in I}$ indépendantes telles que pour tout $i \in I$, $X_i \sim \mathscr{L}_i$.

Un cas particulier très courant.

Définition. On dit qu'une famille finie ou dénombrable de v.a.d. est identiquement distribuée quand toutes les variables aléatoires suivent la même loi.

Théorème 4. *Soit $\mathscr{L}$ une loi de probabilité.*
Alors il existe un espace probabilisé $(\Omega, \mathscr{T}, \mathbb{P})$ et des variables aléatoires $(X_i)_{i\in I}$ indépendantes et identiquement distribuées qui suivent la loi $\mathscr{L}$.

Dans ce cas, on voit parfois écrit que la famille de v.a.d. est une famille de variables i.i.d.

L'exemple typique est l'étude des suites infinies de lancers d'une pièce : on modélise cette expérience aléatoire par une suite de variables de Bernoulli i.i.d. de même paramètre p.

6.5 Somme de variables indépendantes identiquement distribuées

Plusieurs résultats classiques

Proposition 32. *Soit $X_1, \ldots, X_n$ n v.a.r.*
Si ces n variables aléatoires sont mutuellement indépendantes et suivent la même loi de Bernoulli de paramètre p, alors la somme $X_1 + \ldots + X_n$ suit la loi binomiale de paramètre (n, p).

Proposition 33. *Soit X, Y deux v.a.d. réelles indépendantes suivant une loi de Poisson de paramètres $\lambda > 0$ et $\mu > 0$.*
Alors $X + Y$ suit une loi de Poisson de paramètre $\lambda + \mu$.

7 Loi faible des grands nombres

Proposition 34. *Soit $(X_n)_{n \geqslant 1}$ une suite de v.a.d. réelles i.i.d. admettant une variance finie, alors en posant $S_n = \displaystyle\sum_{k=1}^{n} X_k$, $m = \mathbb{E}(X_i)$, $\sigma^2 = \mathbb{V}(X_i)$, pour tout $\varepsilon > 0$, on a*

$$\mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \xrightarrow[n \to +\infty]{} 0$$

Plus précisément, on a

$$\mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \frac{\sigma^2}{n\varepsilon^2}$$

8 Fonction génératrice d'une variable aléatoire à valeurs entières

8.1 Généralités

Définition. Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N} \cup \{+\infty\}$.

On appelle série génératrice de X la série entière $\sum_{n\geqslant 0} \mathbb{P}(X = n)z^n$.

Quand cette série converge, on appelle $G_X(z)$ sa somme : la fonction $z \mapsto G_X(z)$ est appelée fonction génératrice de X.

On peut l'écrire sa somme comme une espérance : $G_X(z) = \sum_{n=0}^{+\infty} z^n\, \mathbb{P}(X = n) = \mathbb{E}(z^X)$, à condition que la série $\sum_{n\geqslant 0} \mathbb{P}(X = n)z^n$ converge absolument.

Proposition 35. *La série génératrice d'une variable aléatoire X à valeurs dans $\mathbb{N} \cup \{+\infty\}$ est de rayon de convergence au moins 1 et converge normalement sur $[0, 1]$. La fonction G_X est donc continue sur $[0, 1]$. De plus, $G_X(1) = \mathbb{P}(X \neq +\infty) = 1 - \mathbb{P}(X = +\infty)$.*

***Exercices* :**

11) Déterminez la fonction génératrice associée à une variable de Bernoulli.
12) Même question avec une variable qui suit une loi binomiale.
13) Même question avec une variable qui suit une loi uniforme sur $[\![0, n]\!]$.
14) Même question avec une variable qui suit une loi géométrique.
15) Même question avec une variable qui suit une loi de Poisson.

La fonction génératrice caractérise parfaitement la variable aléatoire.

Proposition 36. *Soit X, Y deux variables aléatoires à valeurs dans $\mathbb{N} \cup \{+\infty\}$.
Si $G_X = G_Y$ sur un intervalle $[0, \alpha]$ où $\alpha > 0$, alors $X \sim Y$ (X et Y suivent la même loi).*

8.2 Lien entre espérance et fonction génératrice

Quand $\mathbb{P}(X = +\infty) = 0$, on dit que la variable est presque sûrement à valeurs dans $\mathbb{N}$ (exemple typique : les variables suivant des lois géométriques). C'est évidemment le cas quand X est à valeurs dans $\mathbb{N}$.

Théorème 5. *Soit X une variable aléatoire discrète presque sûrement à valeurs dans $\mathbb{N}$.
X admet une espérance finie si et s.si G_X est dérivable en 1, dans ce cas $\mathbb{E}(X) = G'_X(1)$.
X admet une variance finie si et s.si G_X est deux fois dérivable en 1, dans ce cas $\mathbb{E}(X(X - 1)) = G''_X(1)$.*

Quand $\mathbb{P}(X = +\infty) > 0$, alors X n'est pas d'espérance finie ($\mathbb{E}(X) = +\infty$ dans ce cas).

8.3 Fonction génératrice d'une somme de variables indépendantes

Proposition 37. *Soit X, Y deux variables aléatoires à valeurs dans $\mathbb{N} \cup \{+\infty\}$.
Si X et Y sont indépendantes, alors $G_{X+Y} = G_X.G_Y$.*

Plus généralement, si $X_1, \ldots, X_n$ sont indépendantes, alors $G_{X1+\ldots+X_n} = \prod_{k=1}^{n} G_{X_k}$.

Exemple. Si X et Y sont indépendantes et suivent des lois de Poisson de paramètres $\lambda > 0$ et $\mu > 0$, alors $X + Y$ suit la loi de Poisson de paramètre $\lambda + \mu$.