

# Chapitre 13

## Variables aléatoires discrètes

### 13.1 Espaces probabilisés

#### 13.1.1 Tribus

**Définition 13.1** Une tribu  $\mathcal{A}$  sur l'ensemble  $\Omega$  est une partie de  $\mathcal{P}(\Omega)$ , c'est-à-dire un ensemble de parties de  $\Omega$  telle que :

- $\Omega \in \mathcal{A}$ ,
- pour tout  $A$  de  $\mathcal{A}$ ,  $\Omega - A$  est dans  $\mathcal{A}$  (stabilité par passage au complémentaire),
- pour tout suite  $(A_n)_{n \in \mathbb{N}}$  d'éléments de  $\mathcal{A}$   $\bigcup_{n \in \mathbb{N}} A_n$  est dans  $\mathcal{A}$  (stabilité par union d'énumérable).

**Vocabulaire 13.1** On parle parfois de  $\sigma$ -algèbre au lieu de tribu, ce qui explique la notation  $\mathcal{A}$ .

**Proposition 13.1** Une tribu est une partie de  $\mathcal{P}(\Omega)$  stable pour les opérations ensemblistes (union, intersection et passage au complémentaire) finies ou dénombrables.

#### 13.1.2 Evènements

**Vocabulaire 13.2** Les éléments d'une tribu s'appellent des évènements.

**Vocabulaire 13.3** Un couple  $(\Omega, \mathcal{A})$  formé d'un ensemble et d'une tribu sur cet ensemble s'appelle un espace probabilisable.

### 13.2 Propriétés élémentaires des probabilités

**Définition 13.2** Une probabilité sur l'espace probabilisable  $(\Omega, \mathcal{A})$  est une application  $\mathbb{P}$  définie sur  $\mathcal{A}$  à valeurs dans  $[0, 1]$ , telle que :

- $\mathbb{P}(\Omega) = 1$ .
- Si  $(A_n)_{n \geq 0}$  est une suite d'évènements disjoints alors  $\mathbb{P}(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n)$ , ce qui implique en particulier que la série est convergente.

**Vocabulaire 13.4** Un triplet  $(\Omega, \mathcal{A}, P)$  formé d'un ensemble, d'une tribu sur cet ensemble et d'une probabilité sur cette tribu s'appelle un espace probabilisé.

**Proposition 13.2** Si  $\mathbb{P}$  est une probabilité sur  $(\Omega, \mathcal{A})$  alors

- $\mathbb{P}(\emptyset) = 0$ ,
- Si  $A$  et  $B$  sont deux évènements disjoints  $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ .

On retrouve donc les propriétés des probabilités définies sur les ensembles finis.

**Proposition 13.3** Si  $\Omega$  est dénombrable et si  $\mathcal{A} = \mathcal{P}(\Omega)$ , la donnée d'une probabilité sur  $(\Omega, \mathcal{A})$  est équivalente à la donnée d'une famille de réels positifs  $(p_\omega)_{\omega \in \Omega}$  de somme 1. L'identification se fait via la relation  $\mathbb{P}(\omega) = p_\omega$ . On a alors pour toute partie  $A$  de  $\Omega$  :  $\mathbb{P}(A) = \sum_{\omega \in A} p_\omega$ .

On peut alors étendre toute la théorie des probabilités sur les ensembles finis en remplaçant les sommes finies par des familles dont il faudra néanmoins vérifier qu'elles sont sommables.

**Définition 13.3** Une famille finie ou dénombrable  $(A_i)_{i \in I}$  d'évènements est un système complet d'évènements si elle constitue une partition de  $\Omega$ .

**Vocabulaire 13.5** On peut étendre cette définition et parler de système quasi-complet, lorsque  $\sum_{i \in I} \mathbb{P}(A_i) = 1$  (les évènements restent disjoints). Si on ajoute à cette famille l'évènement  $\Omega - \bigcup_{i \in I} A_i$  on obtient un système complet d'évènement. Même si le terme quasi-complet n'est pas au programme et est rarement employé on verra régulièrement apparaître de tels systèmes.

**Proposition 13.4 (Formule des probabilités totales)** Si  $B$  est un évènement et  $(A_i)_{i \in I}$  est un système complet d'évènement :

$$\mathbb{P}(B) = \sum_{i \in I} \mathbb{P}(B \cap A_i).$$

**Vocabulaire 13.6** Un évènement est presque sûr si sa probabilité est égale à 1. Une propriété définie sur  $\Omega$  est presque sûre si l'ensemble des éléments de  $\Omega$  qui la vérifie est un évènement presque sûr. Un évènement est négligeable si sa probabilité est nulle.

**Proposition 13.5 (Continuité croissante)** Si  $(A_n)_{n \in \mathbb{N}}$  est une suite croissante d'évènements alors

$$\lim_{n \rightarrow +\infty} \mathbb{P}(A_n) = P \left( \bigcup_{n=0}^{+\infty} A_n \right).$$

**Proposition 13.6 (Continuité décroissante)** Si  $(A_n)_{n \in \mathbb{N}}$  est une suite décroissante d'évènements alors

$$\lim_{n \rightarrow +\infty} \mathbb{P}(A_n) = P \left( \bigcap_{n=0}^{+\infty} A_n \right).$$

**Proposition 13.7** Si  $(A_n)_{n \in \mathbb{N}}$  est une suite d'évènements alors

$$P \left( \bigcup_{n=0}^{+\infty} A_n \right) \leq \sum_{n=0}^{+\infty} \mathbb{P}(A_n)$$

**Remarque 13.1** On remarquera que le second membre peut être égal à  $+\infty$ . C'est une technique qui sera beaucoup utilisée en théorie des probabilités. Si  $(x_i)_{i \in I}$  est une famille de réels positifs on peut toujours parler de sa somme. Elle est finie si et seulement la famille est sommable. On peut utiliser directement les techniques de conservation des inégalités et de sommation par paquets à condition de se limiter aux familles de réels positifs.

**Proposition 13.8** Toute réunion finie ou dénombrable d'évènements négligeables est un évènement négligeable.

## 13.3 Probabilités conditionnelles et indépendance

On se donne dans ce qui va suivre un espace probabilisé  $(\Omega, \mathcal{A}, P)$ , auquel il sera fait implicitement référence.

### 13.3.1 Probabilité conditionnelle

**Définition 13.4** Si  $A$  et  $B$  sont deux évènements, avec  $\mathbb{P}(B) > 0$ , on appelle probabilité conditionnelle de  $A$  sachant  $B$  (souvent simplifiée en « probabilité de  $A$  si  $B$  ») le nombre

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

**Proposition 13.9 (Formule des probabilités composées)**

— Si  $A$  et  $B$  sont deux évènements avec  $\mathbb{P}(B) \neq 0$  :

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

— Si  $(A_1, \dots, A_p)$  sont des évènements avec  $\mathbb{P}(A_1 \cap \dots \cap A_{p-1}) \neq 0$  :

$$\mathbb{P}(A_1 \cap \dots \cap A_p) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2) \cdots \mathbb{P}(A_p|A_1 \cap \dots \cap A_{p-1}).$$

**Remarque 13.2**  $\mathbb{P}(B|A_1 \cap \dots \cap A_p)$  se note aussi  $\mathbb{P}(B|A_1, \dots, A_p)$

**Proposition 13.10 (Formule des probabilités totales)** Si  $(A_i)_{i \in I}$  est un système complet d'évènements, avec  $\mathbb{P}(A_i) > 0$  pour tout  $i$ , alors, pour tout évènement  $B$  :

$$\mathbb{P}(B) = \sum_{i \in I} \mathbb{P}(B|A_i)\mathbb{P}(A_i).$$

**Proposition 13.11 (Formule de Bayes)** Avec les mêmes hypothèses que dans la proposition précédente on aura, si de plus  $\mathbb{P}(B) > 0$

$$\forall j \in I \quad \mathbb{P}(A_j|B) = \frac{\mathbb{P}(A_j)\mathbb{P}(B|A_j)}{\sum_{i \in I} \mathbb{P}(A_i)\mathbb{P}(B|A_i)},$$

en particulier si  $A$  et  $B$  sont deux évènements de probabilité non nulle :

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

### 13.3.2 Evènement indépendants

**Définition 13.5**

— Deux évènements  $A$  et  $B$  sont indépendants si et seulement si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

— Des évènements  $A_i$ ,  $i \in I$  sont mutuellement indépendants si et seulement si pour toute suite finie d'éléments distincts de  $I$

$$\mathbb{P}(A_{i_1} \cap \dots \cap A_{i_p}) = \mathbb{P}(A_{i_1}) \dots \mathbb{P}(A_{i_p}).$$

## 13.4 Variables aléatoires discrètes

### 13.4.1 Définition

**Définition 13.6** Etant donné un ensemble  $E$  et un espace probabilisé  $(\Omega, \mathcal{A}, P)$ , une variable aléatoire discrète est une application  $X$  de  $\Omega$  dans  $E$  telle que  $X(\Omega)$  soit fini ou dénombrable et que, pour tout  $x$  de  $X^{-1}(\{x\})$  appartient à  $\mathcal{A}$ .

**Vocabulaire 13.7** Si  $E \subset \mathbb{R}$ , on parle de variable aléatoire réelle.

### 13.4.2 Loi d'une variable aléatoire

**Définition 13.7** On appelle loi d'une variable aléatoire discrète  $X$  à valeurs dans  $E$  la fonction, notée  $\mathbb{P}_X$ , définie sur  $\mathcal{P}(E)$  par

$$\mathbb{P}_X(A) = \mathbb{P}(X \in A).$$

**Remarque 13.3**  $X \in A$  est une notation simplifiée pour l'évènement  $\{\omega \in \Omega; X(\omega) \in A\}$ . Cet ensemble est bien un évènement car il est égal à

$$\bigcup_{a \in X(\Omega) \cap A} X^{-1}(\{a\})$$

et puisque  $X(\Omega)$  est fini ou dénombrable, cet ensemble est une réunion finie ou dénombrable d'évènements, c'est donc un évènement.

Si  $X$  est une variable aléatoire réelle on utilisera de la même façon les raccourcis  $X > a$ ,  $X \geq a$ ,  $X < b$ ,  $X \leq b$  et  $a \leq X \leq b$  et toute notation similaire.

**Remarque 13.4** La notation  $\mathbb{P}_X$  justifiée car on pourrait vérifier que  $\mathbb{P}_X$  est une probabilité sur  $(E, \mathcal{P}(E))$ . On parle de probabilité image de  $\mathbb{P}$  par  $X$  (notion hors-programme). On écrit  $X \sim \mathcal{L}$  pour exprimer que la loi de  $X$  est  $\mathcal{L}$ . Si deux variables aléatoires discrètes suivent la même loi on écrit  $X \sim Y$ . On remarquera que ces variables aléatoires peuvent très bien ne pas être définies sur un même espace probabilisable.

## 13.5 Couples de variables aléatoires, variables aléatoires indépendantes

### 13.5.1 $n$ -uplets de variables aléatoires. Loi conjointe, lois marginales

Toutes les variables aléatoires concernées par la suite de cet exposé sont des variables aléatoires discrètes définies sur le même espace probabilisé. Nous ne le rappellerons pas.

**Proposition 13.12** *Si  $X$  et  $Y$  sont deux variables aléatoires à valeurs dans  $E$  et  $F$ , alors la fonction  $(X, Y)$  est une variable aléatoire discrète à valeurs dans  $E \times F$ . De même si  $X_1, \dots, X_n$  sont  $n$  variables aléatoires, à valeurs dans  $E_1, \dots, E_n$ , la fonction  $(X_1, \dots, X_n)$  est une variable aléatoire discrète à valeurs dans  $E_1 \times \dots \times E_n$ .*

**Vocabulaire 13.8** *On parle de couple de variables aléatoires, de  $n$ -uplet de variables aléatoires. Si  $X_1, \dots, X_n$  sont toutes des variables aléatoires discrètes réelles  $(X_1, \dots, X_n)$  s'appelle un vecteur aléatoire (discret).*

**Définition 13.8** *Si  $X$  et  $Y$  sont deux variables aléatoires discrètes la loi de  $(X, Y)$  s'appelle la loi conjointe de  $X$ .*

Soit  $X$  et  $Y$  deux variables aléatoires discrètes. Notons  $X(\Omega) = \{x_k; k \in K\}$ ,  $Y(\omega) = \{y_l; l \in L\}$  leurs images, qui sont finis ou dénombrables. Alors  $(X, Y)(\Omega) \subset \{(x_k, y_l); (k, l) \in K \times L\} = X(\Omega) \times Y(\Omega)$ . La loi de conjointe de  $(X, Y)$  est déterminée à l'aide  $\{p_{k,l} = \mathbb{P}((X, Y) = (x_k, y_l)); (k, l) \in K \times L\}$  par

$$\mathbb{P}_{X,Y}(A) = \sum_{(x_k, y_l) \in A} p_{k,l}.$$

C'est une généralisation du résultat vu pour une variable aléatoire. L'image de  $(X, Y)$  n'étant pas toujours facile à déterminer on somme sur un ensemble plus grand, généralement  $X(\Omega) \times Y(\Omega)$ , voire  $E \times F$ , en ajoutant des probabilités nulles pour les singletons qui ne sont pas dans l'image.

**Définition 13.9** *La loi marginale de  $X$  est définie par  $\mathbb{P}_X(A) = \mathbb{P}_{X,Y}(A \times F)$ , la loi marginale de  $Y$  est définie par  $\mathbb{P}_Y(B) = \mathbb{P}_{X,Y}(E \times B)$*

**Remarque 13.5** *Les lois marginales sont les lois des coordonnées, dont on pourrait vérifier qu'il s'agit de variables aléatoires.*

Avec les notations précédentes on aura :  $\mathbb{P}_X(\{x_k\}) = \sum_{l \in L} p_{k,l}$ , et une formule similaire pour  $\mathbb{P}_Y$ .

**Définition 13.10** *Si  $\mathbb{P}_X(\{x\})$  (qui est égal à  $\mathbb{P}(X = x)$ ) est non nul, on peut définir la probabilité conditionnelle de  $Y$  sachant  $X = x$  par  $\mathbb{P}(Y = y | X = x) = \mathbb{P}((X, Y) = (x, y)) / \mathbb{P}(X = x)$ . On définirait de même  $\mathbb{P}(Y = y | X > x)$  ou toute autre probabilité conditionnelle similaire.*

**Remarque 13.6** *Les définitions précédentes s'étendent facilement aux  $n$ -uplets de variables aléatoires.*

### 13.5.2 Variables aléatoires indépendantes

**Définition 13.11** *Deux variables aléatoires  $X$  et  $Y$  sont indépendantes si et seulement si pour tout couple d'ensembles :*

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B).$$

**Remarque 13.7** *On utilise ici deux abréviations :  $X \in A$  au lieu de  $\{\omega; X(\omega) \in A\}$  et la virgule à la place de la conjonction (ou de l'intersection si on effectue d'abord l'interprétation ensembliste de  $X \in A$  et  $Y \in B$ ).*

**Proposition 13.13** *Deux variables aléatoires discrètes,  $X$  à valeurs dans  $E$  et  $Y$  à valeurs dans  $F$  sont indépendantes si et seulement si*

$$\forall (x, y) \in E \times F \quad \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y).$$

On peut étendre la définition de l'indépendance à un  $n$ -uplet de variables aléatoires.

**Définition 13.12** *Les variables aléatoires  $(X_i)_{i \in I}$  sont mutuellement indépendantes si et seulement si pour tout  $n$ -uplet d'ensembles  $(A_1, \dots, A_n)$  et tout  $n$ -uplet d'indices  $(i_1, \dots, i_n)$  :*

$$\mathbb{P}(X_{i_1} \in A_1, \dots, X_{i_n} \in A_n) = \mathbb{P}(X_{i_1} \in A_1)\mathbb{P}(X_{i_2} \in A_2) \cdots \mathbb{P}(X_{i_n} \in A_n).$$

**Remarque 13.8** Des variables mutuellement indépendantes sont deux à deux indépendantes, mais la réciproque est fautive.

Le théorème qui suit, très utile, sera admis.

**Théorème 13.1** Si  $X_1, \dots, X_n$  sont des variables aléatoires mutuellement indépendantes, si  $p$  est un entier compris entre 1 et  $n - 1$ , si  $f$  et  $g$  sont deux fonctions alors les variables aléatoires  $f(X_1, \dots, X_p)$  et  $g(X_{p+1}, \dots, X_n)$  sont indépendantes.

**Remarque 13.9**  $f(X_1, \dots, X_p)$  désigne la fonction  $\omega \mapsto f(X_1(\omega), \dots, X_p(\omega))$ . On pourrait vérifier qu'il s'agit bien d'une variable aléatoire.

Le théorème suivant va nous permettre d'étudier d'un point de vue probabiliste le comportement de suites de variables aléatoires. Sa démonstration est hors-programme, et ce n'est pas regrettable.

**Théorème 13.2** Soit  $(\mathcal{L}_n)_{n \in \mathbb{N}}$  une suite de lois. Alors il existe un espace probabilisé  $(\Omega, \mathcal{A}, P)$  et une suite  $(X_n)_{n \in \mathbb{N}}$  de variables aléatoires mutuellement indépendantes de telle manière que  $X_n$  suive la loi  $\mathcal{L}_n$ .

**Exemple 13.1** Modélisation du jeu de pile ou face infini.

## 13.6 Lois usuelles

### 13.6.1 Loi géométrique

**Définition 13.13** Une variable aléatoire à valeur dans  $\mathbb{N}^*$  suit une loi géométrique de paramètre  $p$  appartenant à  $]0, 1[$  si et seulement si  $\mathbb{P}(X = k) = (1 - p)^{k-1}p$ . On écrit  $X \sim \mathcal{G}(p)$ .

**Proposition 13.14** La loi géométrique est la loi du rang du premier succès dans une suite d'épreuves de Bernoulli mutuellement indépendantes de paramètres  $p$ .

**Proposition 13.15** Les lois géométriques sont les lois sans mémoire sur  $\mathbb{N}^*$ . Elles sont caractérisées par :

$$\mathbb{P}(X = n + k | X > k) = \mathbb{P}(X = n).$$

### 13.6.2 Loi de Poisson

**Définition 13.14** Une variable aléatoire à valeurs dans  $\mathbb{N}$  suit une loi de Poisson de paramètre  $\lambda$  appartenant à  $\mathbb{R}^*$  si et seulement si  $\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$ . On écrit  $X \sim \mathcal{P}(\lambda)$ .

**Proposition 13.16** Si  $X_n \sim \mathcal{B}(n, p_n)$  avec  $\lim_{n \rightarrow +\infty} np_n = \lambda$  alors

$$\forall k \in \mathbb{N} \quad \lim_{n \rightarrow +\infty} \mathbb{P}(X_n = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

On dit que la loi de Poisson est la loi des événements rares.

Dans la pratique, on sera en présence d'une variable aléatoire suivant une loi  $\mathcal{B}(N, p)$ ,  $N$  grand,  $p$  petit, de telle sorte que le produit  $\lambda = Np$  reste petit devant  $N$ . Ce que dit la proposition précédente, sans qu'elle ne donne aucun résultat quantitatif sur la zone de validité du résultat et la qualité de l'approximation, c'est que pour  $k$  petit devant  $n$ , en général dans les environs de  $\lambda$ , la probabilité  $\mathbb{P}(X = k)$  est à-peu-près égale à  $e^{-\lambda} \frac{\lambda^k}{k!}$ .

Par exemple, soit un lot de 5000 vis, choisies au hasard. La probabilité pour que l'une d'elle soit défectueuse est  $p = 0,001$ , notons  $X$  la variable aléatoire renvoyant le nombre de vis défectueuses, et  $Y$  une variable aléatoire suivant une loi de Poisson de paramètre  $\lambda = Np$ . L'espérance du nombre de vis défectueuses est  $Np = \lambda = 5$ .

Voici les lois de  $X$  et  $Y$  restreintes aux petite valeurs de  $k$ . On voit que la loi de Poisson réalise une excellente approximation.

INSERER UN TABLEAU (mieux encore un graphique)

Si on choisit maintenant  $p = 0.1$  la valeur attendue est 500. Le tableau suivant montre que dans le voisinage de la valeur attendue l'approximation n'est pas bonne.

INSERER UN TABLEAU

## 13.7 Espérance

### 13.7.1 Définition

**Définition 13.15** Soit  $X$  une variable aléatoire à valeurs positives. Son espérance, notée  $E(X)$ , à valeur dans  $\mathbb{R} \cup \{+\infty\}$  est  $E(X) = \sum_{x \in X(\Omega)} \mathbb{P}(X = x)x$ .

Au lieu de sommer sur  $X(\Omega)$ , on pourrait sommer sur tout ensemble fini ou dénombrable contenant  $X(\Omega)$ . Cela rend de grands services dans la rédaction des démonstrations.

**Définition 13.16** Si  $X$  est une variable aléatoire réelle, elle est dite d'espérance finie si la famille  $(\mathbb{P}(X = x))_{x \in X(\Omega)}$  est sommable. Son espérance est alors  $E(X) = \sum_{x \in X(\Omega)} \mathbb{P}(X = x)x$ .

On définirait de même l'espérance d'une variable aléatoire à valeurs complexes.

**Vocabulaire 13.9** Une variable aléatoire est centrée si elle est d'espérance finie et cette espérance est nulle.

### 13.7.2 Espérance des lois usuelles

**Proposition 13.17** Les espérances des lois usuelles sont :

- $E(\mathcal{B}(p)) = p$ ,
- $E(\mathcal{B}(n, p)) = np$ ,
- $E(\mathcal{G}(p)) = \frac{1}{p}$ ,
- $E(\mathcal{B}(\lambda)) = \lambda$ .

### 13.7.3 Propriétés de l'espérance

**Proposition 13.18** Soit  $X$  une variable aléatoire réelle,  $(A_i)_{i \in I}$  un système complet d'évènements tel que pour tout  $i$  de  $I$   $X$  soit constante sur  $A_i$  de valeur  $x_i$ . Alors  $X$  est d'espérance finie si et seulement si la famille  $(x_i \mathbb{P}(A_i))_{i \in I}$  est sommable et dans ce cas

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

On remarquera le lien avec l'intégrale des fonctions en escalier.

**Proposition 13.19** Soit  $X$  et  $Y$  deux variables aléatoires réelles. Si  $Y$  est d'espérance finie et si  $|X| \leq Y$  alors  $X$  est d'espérance finie et  $|E(X)| \leq E(|X|) \leq E(Y)$ .

**Proposition 13.20** L'ensemble des variables aléatoires possédant une espérance finie est un sous-espace vectoriel de l'espace des variables aléatoires réelles et l'application qui à  $X$  associe  $E(X)$  est linéaire, positive et croissante, ce qui se traduit par :

- $E(X + Y) = E(X) + E(Y)$ ,
- $E(\lambda X) = \lambda E(X)$ ,
- $X \geq 0 \Rightarrow E(X) \geq 0$ ,
- $X \geq Y \Rightarrow E(X) \geq E(Y)$ ,

où  $X$  et  $Y$  sont des variables aléatoires réelles et  $\lambda$  un réel.

**Proposition 13.21 (Formule de transfert)** Soit  $X$  une variable aléatoire discrète,  $f$  une fonction définie sur  $X(\Omega)$  à valeurs dans  $\mathbb{R}$ , alors  $f(X)$  est d'espérance finie si et seulement si la famille  $(f(x)\mathbb{P}(X = x))_{x \in X(\Omega)}$  est sommable et dans ce cas :

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

**Proposition 13.22 (Inégalité de Markov)** Soit  $X$  une variable aléatoire positive d'espérance finie, et  $t$  un réel strictement positif. Alors :

$$\mathbb{P}(X \geq t) \leq \frac{E(X)}{t}.$$

**Proposition 13.23** Si  $X$  et  $Y$  sont deux variables aléatoires réelles indépendantes d'espérance finie alors  $XY$  est d'espérance finie et de plus :

$$E(XY) = E(X)E(Y).$$

## 13.8 Variance, covariance

Dans cette section une variable aléatoire est une variable aléatoire discrète à valeurs réelles.

**Définition 13.17** Si  $p$  est un entier, la variable aléatoire  $X$  admet un moment d'ordre  $p$  si et seulement la variable aléatoire  $X^p$  admet une espérance. Le moment d'ordre  $p$  de  $X$  est alors  $E(X^p)$

**Proposition 13.24** Si une variable aléatoire admet un moment d'ordre 2 alors elle est d'espérance finie.

Cela résulte de  $|X| \leq \frac{1+X^2}{2}$ .

**Proposition 13.25 (Inégalité de Cauchy-Schwarz)** Si les variables aléatoires  $X$  et  $Y$  admettent un moment d'ordre 2, alors  $XY$  est d'espérance finie et

$$(E(XY))^2 \leq E(X^2)E(Y^2).$$

En particulier  $E(X)^2 \leq E(X^2)$

On écrit  $E((X + tY)^2) \geq 0$  pour tout  $t$ . Le trinôme en  $t$  est positif, donc son discriminant est négatif ou nul.

**Proposition 13.26** L'ensemble des variables aléatoires admettant un moment d'ordre 2 est un sous-espace vectoriel (de l'espace des variables aléatoires).

Résulte de  $(X + Y)^2 \leq 2(X^2 + Y^2)$ .

**Définition 13.18** Si  $X$  admet un moment d'ordre 2 sa variance est  $V(X) = E(X^2) - (E(X))^2 = E((X - E(X))^2)$ . Son écart type est  $\sigma(X) = \sqrt{V(X)}$ .

Cette variance est bien définie et positive car si  $X$  admet un moment d'ordre 2,  $X$  admet aussi un moment d'ordre 1 et d'après la proposition précédente  $E(X^2) - (E(X))^2 \geq 0$ .

**Vocabulaire 13.10** Une variable est réduite si sa variance (ou son écart type) est égal à 1.

**Proposition 13.27** Si  $X$  est une variable aléatoire admettant un moment d'ordre 2 et si  $a$  et  $b$  sont deux réels alors  $V(aX + b) = a^2V(X)$ .

**Exemple 13.2** Si  $\sigma(X) > 0$  la variable  $\frac{X - E(X)}{\sigma(X)}$  est réduite et centrée.

**Proposition 13.28** Les lois usuelles du programme admettent toute un moment d'ordre 2 et en particulier une variance. Plus précisément :

- $V(\mathcal{B}(p)) = p - p^2 = p(1 - p) = pq$  (traditionnellement  $1 - p$  est noté  $q$ ),
- $V(\mathcal{B}(n, p)) = np(1 - p) = npq$ ,
- $V(\mathcal{G}(p)) = \frac{1-p}{p^2} = \frac{q}{p^2}$ ,
- $V(\mathcal{P}(\lambda)) = \lambda$ .

**Théorème 13.3 (Inégalité de Bienaymé-Tchebychev)** Soit  $X$  une variable aléatoire admettant un moment d'ordre 2 alors, pour tout  $\epsilon > 0$  :

$$\mathbb{P}(|X - E(X)| \geq \epsilon) \leq \frac{V(X)}{\epsilon^2}.$$

Cette inégalité est l'inégalité de Markov appliquée à  $Y = (X - E(X))^2$ .

**Définition 13.19** La covariance de deux variables aléatoires  $X$  et  $Y$  admettant chacune un moment d'ordre 2 est

$$\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y).$$

On retrouve  $\text{Cov}(X, X) = V(X)$

**Exercice 13.1** Montrer de même que  $4\text{Cov}(X, Y) = V(X + Y) - V(X - Y)$ , ou  $2\text{Cov}(X, Y) = V(X + Y) - V(X) - V(Y)$ .

**Proposition 13.29** Si  $X$  et  $Y$  sont deux variables indépendantes, admettant des moments d'ordre 2 alors :

$$V(X + Y) = V(X) + V(Y).$$

**Proposition 13.30** Si  $X_1, \dots, X_n$  sont des variables aléatoires admettant des moments d'ordre 2 :

$$V(X_1 + \dots + X_n) = \sum_{i=1}^n V(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j).$$

Il en résulte :

**Proposition 13.31** Si  $X_1, \dots, X_n$  sont des variables aléatoires deux à deux indépendantes admettant des moments d'ordre 2 :

$$V(X_1 + \dots + X_n) = \sum_{i=1}^n V(X_i) = V(X_1) + \dots + V(X_n).$$

## 13.9 Loi faible des grands nombres

**Théorème 13.4 (Loi faible des grands nombres)** Si  $(X_n)_{n \geq 0}$  est une suite de variables aléatoires deux à deux indépendantes, de même loi, admettant un moment d'ordre 2 alors, si on note  $S_n = X_1 + \dots + X_n$  et  $m = E(X_1) (= E(X_2) = \dots = E(X_m))$  :

$$\lim_{n \rightarrow +\infty} P\left(\left|\frac{S_n}{n} - m\right| \geq \epsilon\right) = 0.$$

## 13.10 Fonctions génératrices

**Définition 13.20** Soit  $X$  une variable aléatoire à valeurs dans  $\mathbb{N}$ . Sa fonction génératrice est

$$G_X(t) = E(t^X) = \sum_{n=0}^{+\infty} \mathbb{P}(X = n)t^n.$$

Le programme ne précise pas son domaine de définition. Au sens étendu de l'espérance des variables aléatoires positives cette fonction est au moins définie sur  $\mathbb{R}^+$ , à valeurs dans  $\mathbb{R}^+ \cup \{+\infty\}$ .

**Proposition 13.32** La série entière définissant la fonction génératrice a un rayon de convergence au moins égal à 1 et converge normalement donc uniformément sur le disque fermé de centre 0 et de rayon 1, en particulier sur  $[-1, 1]$ . Il en résulte que  $G_X$  est continue sur  $[-1, 1]$  au moins et de classe  $C^\infty$  sur  $] -1, 1[$  au moins. De plus puisque  $\mathbb{P}(X = n) = G_x^{(n)}(0)$  la fonction génératrice détermine la loi de  $X$ . Deux variables aléatoires ayant même fonction génératrice sont de même loi.

**Proposition 13.33 (Fonctions génératrice des lois usuelles)** Les fonctions génératrices des lois usuelles sont données par :

- Si  $X \sim \mathcal{B}(p)$  alors  $G_X(t) = (1 - p) + pt = q + pt$ ,
- si  $X \sim \mathcal{B}(n, p)$  alors  $G_X(t) = ((1 - p) + pt)^n = (q + pt)^n$ ,
- si  $X \sim \mathcal{G}(p)$  alors  $G_X(t) = \frac{pt}{1 - qt}$ ,
- si  $X \sim \mathcal{P}(\lambda)$  alors  $G_X(t) = e^{(t-1)\lambda}$ .

**Proposition 13.34**  $X$  possède une espérance si et seulement si  $G_X$  est dérivable à gauche en 1. On alors

$$E(X) = G'_X(1).$$

**Lemme 13.1** Soit  $f : [0, 1] \rightarrow \mathbb{R}$  une fonction somme sur cet intervalle d'une série entière  $\sum_{n \geq 0} a_n x^n$  à coefficients positifs. Alors  $f$  est dérivable en 1 si et seulement si  $\sum_{n \geq 1} n a_n$  converge. Dans ce cas on a  $f'(1) = \sum_{n=1}^{+\infty} n a_n$ .

Supposons que  $\sum_{n \geq 1} n a_n$  converge. Puisque  $\sum_{n \geq 0} a_n$  converge on sait déjà que  $f$  est continue sur  $[0, 1]$ , de classe  $C^\infty$  sur  $[0, 1[$  et en particulier de classe  $C^1$  sur cet intervalle. De plus

$$\forall x \in [0, 1[ \quad f'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1} = \sum_{n=1}^{+\infty} u_n(x).$$

On a

$$\forall x \in [0, 1[ \quad \forall n \in \mathbb{N}^* \quad |u_n(x)| \leq n a_n.$$

Or  $\sum_{n \geq 1} na_n$  converge donc  $\sum_{n \geq 1} u_n$  converge normalement, donc uniformément sur  $[0, 1[$ . Puisque  $\lim_{x \rightarrow 1^-} u_n(x) = na_n$ , pour tout  $n$ , on peut appliquer le théorème de permutation des limites. On en déduit  $\lim_{x \rightarrow 1^-} f'(x) = \sum_{n=1}^{+\infty} na_n$ . Ce résultat, joint aux hypothèses déjà vérifiées sur  $f$ , permet d'affirmer que  $f$  est de classe  $\mathcal{C}^1$  sur  $[0, 1]$ , avec  $f'(1) = \sum_{n=1}^{+\infty} na_n$ .

Réciproquement. On suppose  $f$  dérivable en 1. Donc  $f'(1) = \lim_{x \rightarrow 1^-} \frac{f(1) - f(x)}{1 - x}$  existe. Or :

$$\forall x \in [0, 1[ \quad g(x) = \frac{f(1) - f(x)}{1 - x} = \sum_{n=1}^{+\infty} a_n(1 + x + \dots + x^{n-1}) = \sum_{n=1}^{+\infty} v_n(x).$$

Chaque  $v_n$  est une fonction croissante, les  $a_n$  sont positifs donc  $g$  est une fonction croissante. En particulier on aura

$$\forall N \in \mathbb{N}^* \quad \forall x \in [0, 1[ \quad \sum_{n=1}^N v_n(x) \leq g(x) \leq \lim_{x \rightarrow 1^-} g(x) = f'(1).$$

On peut faire tendre  $x$  vers  $1^-$  dans l'expression la plus à gauche. On obtient ainsi :

$$\forall N \in \mathbb{N}^* \quad \sum_{n=1}^N na_n \leq f'(1).$$

Ceci permet d'affirmer que la série à termes positifs  $\sum_{n \geq 1} na_n$  est convergente. Le sens direct montrer alors que sa somme est  $f'(1)$ .

**Proposition 13.35**  *$X$  possède un moment d'ordre 2 si et seulement si  $G_X$  est deux fois dérivable à gauche en 1. On alors*

$$V(X) = G_X''(1) + G_X'(1) - (G_X'(1))^2.$$

Si  $G_X$  est deux fois dérivable à gauche en 1, alors en appliquant deux fois le lemme on en déduit que  $\sum_{n \geq 1} nP(X = n)$  converge ainsi que  $\sum_{n \geq 2} n(n-1)a_n$ , avec  $G_X'(1) = \sum_{n=0}^{+\infty} na_n$  et  $G_X''(1) = \sum_{n=0}^{+\infty} n(n-1)P(X = n)$ , et réciproquement.

On en déduit donc que  $G_X$  est deux fois dérivable en 1 si et seulement si  $E(X)$  et  $E(X(X-1))$  sont finies. Or  $X^2 = X(X-1) + X$  et  $X(X-1) = X^2 - X$ . Donc  $G_X$  est deux fois dérivable en 1 si et seulement si  $E(X)$  et  $E(X^2)$  sont finies. Or  $E(X)$  est finie dès que  $E(X^2)$  est finie.

Finalement  $G_X$  est deux fois dérivable en 1 si et seulement si  $E(X^2)$  est finie. C'est-à-dire si et seulement si  $X$  admet un moment d'ordre 2.

De plus dans ce cas  $G_X'(1) = E(X)$  et  $G_X''(1) = E(X(X-1)) = E(X^2) - E(X)$ , soit  $E(X) = G_X'(1)$  et  $E(X^2) = G_X''(1) + G_X'(1)$ . L'expression de la variance s'en déduit.

**Remarque 13.10** *La même technique pourrait être utilisée pour justifier l'existence et déterminer les valeurs de moments d'ordres supérieurs.*

**Remarque 13.11** *Le résultat de la proposition précédente permet de retrouver facilement les espérances et variances des lois usuelles*

**Proposition 13.36** *La fonction génératrice d'une somme de variables aléatoires indépendantes est le produit de leurs fonctions génératrices.*

En effet :

$$G_{X_1 + \dots + X_n}(t) = E(t^{X_1 + \dots + X_n}) = E(t^{X_1} t^{X_2} \dots t^{X_n}).$$

Mais puisque les  $X_i$  sont mutuellement indépendantes, les  $t^{X_i}$  aussi, donc

$$G_{X_1 + \dots + X_n}(t) = E(t^{X_1} t^{X_2} \dots t^{X_n}) = E(t^{X_1}) E(t^{X_2}) \dots E(t^{X_n}) = G_{X_1}(t) G_{X_2}(t) \dots G_{X_n}(t).$$