Variables aléatoires discrètes

Ce chapitre fait appel aux notions de probabilités élémentaires vues au lycée.

1 Exemple

Dans de nombreux jeux, on fait intervenir le hasard en observant la somme des points marqués par deux dés. Considérons le jet d'un dé blanc et d'un dé rouge et notons S la somme des points obtenus. On modélise cette expérience en prenant l'équiprobabilité sur l'univers

$$\Omega = \{1, 2, 3, 4, 5, 6\}^2.$$

Un événement élémentaire ω est un couple (i, j) où i est le résultat du dé blanc et j, le résultat du dé rouge.
Le tableau suivant résume tous les cas possibles. Pour tout $\omega = (i, j)$, on a $S(\omega) = i + j$.

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

On a défini une application S de Ω dans l'ensemble des points possibles $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.
On dit que S est une **variable aléatoire** sur Ω. Ce qui nous intéresse ici n'est pas ω mais $S(\omega)$.
On cherche donc à déterminer la probabilité que la somme des points soit égale à k pour tout k entier de 2 à 12. En prenant l'équiprobabilité pour $\mathcal{P}(\Omega)$, on obtient :

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

On considère donc un nouvel ensemble d'événements élémentaires noté

$$\Omega' = \{2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9,\ 10,\ 11,\ 12\}\ ,$$

muni de la probabilté $\mathbf{P}_S$ définie dans le tableau ci-dessus. Cette nouvelle probabilité s'appelle **loi de la variable** S.

2 Définition d'une variable aléatoire

Soient Ω et Ω' deux ensembles finis ou dénombrables et X une application de Ω vers Ω'. Lorsqu'on choisit une structure d'espace probabilisé discret $(\Omega, \mathcal{P}(\Omega), \mathbf{P})$, on dit que X est une variable aléatoire sur $(\Omega, \mathcal{P}(\Omega), \mathbf{P})$ et à valeurs dans Ω'. C'est alors la probabilité $\mathbf{P}$ qui est utilisée pour donner un sens aux probabilités d'événements de la forme $\ll X \in A \gg$, où A est inclus dans Ω', ce qui permet de définir une nouvelle probabilité sur Ω' que l'on note $\mathbf{P}_X$.

$$\forall\, A \in \mathcal{P}(\Omega'), \qquad \mathbf{P}_X(A) = \mathbf{P}(X \in A) = \mathbf{P}\{\omega : \omega \in \Omega \quad \text{et} \quad X(\omega) \in A\}\ .$$

Une variable aléatoire peut donc être vue comme un couple formé d'une application X et d'une probabilité sur son ensemble de départ. $(\Omega', \mathcal{P}(\Omega'), \mathbf{P}_X)$ est un nouvel espace probabilisé, construit à partir de $(\Omega, \mathcal{P}(\Omega), \mathbf{P})$ et de X. On dit que la loi P_X est la loi de probabilité de X.

3 Fonction de répartition d'une variable réelle

Pour une variable aléatoire à valeurs dans $\mathbb{R}$, on définit la fonction de répartition de manière analogue à celle des variables statistiques. Les fréquences sont remplacées par des probabilités. Soit X une variable aléatoire réelle. La fonction de répartition de X, notée F, donne la probabilité $F(x)$ pour que X tombe à gauche de x.

$$\begin{aligned} F : \mathbb{R} &\longrightarrow [0,1] \\ x &\longmapsto F(x) = \mathbf{P}(X \leq x) = \mathbf{P}_X(]-\infty, x]) \ . \end{aligned}$$

Par complémentaire, on a bien sûr $\mathbf{P}(X > x) = 1 - F(x)$.

F est une fonction en escalier, croissante, continue à droite et vérifiant

$$\lim_{x \to -\infty} F(x) = 0 \qquad \text{et} \qquad \lim_{x \to +\infty} F(x) = 1 \ .$$

4 Moyenne et variance d'une variable réelle

Elles sont définies de manière similaire à celles des variables statistiques, mais les variables aléatoires pouvant prendre une infinité de valeurs, elles n'existent pas toujours. À la place du mot «moyenne», on utilise souvent le mot «espérance» en particulier quand on parle de gain ou de durée de vie.

Espérance Soit X une variable aléatoire réelle dont la distribution de probabilité est donnée par le tableau

Valeurs de X	x_1	x_2	...	x_n	...
Probabilités	p_1	p_2	...	p_n	...

1er cas : le nombre de valeurs possibles pour X est fini et égal à n. Alors l'espérance de X existe et vaut

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

2ème cas : le nombre de valeurs de X ayant une probabilité non nulle est infini dénombrable. Si la série basée sur la suite $(p_i \, x_i)_{i \geq 1}$ est convergente alors sa somme est l'espérance de X

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

Sinon X n'a pas d'espérance.

Nous n'étudions pas dans ce cours, d'exemple qui illustre ce cas car il nécessite la notion d'analyse sur les séries. Cependant, nous retrouverons à la fin de ce chapitre, deux lois usuelles de variables pouvant prendre un nombre infini de valeurs.

Exemples de l'introduction : espérance de la variable S somme des deux dés

$$E(S) = \sum_{i=1}^{11} p_i \, x_i = \frac{252}{36} = 7 \ .$$

Variance Si la variable aléatoire réelle X admet une espérance $E(X)$ et si la variable $(X - E(X))^2$ admet aussi une espérance, on l'appelle variance de X et on la note $\mathbf{V}(X)$. On a alors

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

Lorsque $\mathbf{V}(X)$ existe, on note $\sigma(X)$ sa racine carrée et on l'appelle écart–type de X.

5 Couple aléatoire

Définition 5.1
Soient X et Y deux variables aléatoires discrètes définies sur le même espace probabilisé $(\Omega, \mathcal{P}(\Omega), \mathbf{P})$, *l'application :*

$$\begin{aligned} \Omega &\longrightarrow \mathbb{R}^2 \\ \omega &\longmapsto (X(\omega), Y(\omega)) \end{aligned}$$

est appelée couple aléatoire discret des marginales X et Y et notée (X, Y).

Définition 5.2
La loi $\mathbf{P}_{X,Y}$ du couple (X, Y) est la probabilité définie sur l'ensemble des parties de $\mathbb{R}^2$ par :

$$\forall\, B \subset \mathbb{R}^2, \quad \mathbf{P}_{X,Y} = \mathbf{P}\left(\{\omega \in \Omega,\ (X(\omega), Y(\omega)) \in B\}\right) \ .$$

Les lois $\mathbf{P}_X$ et $\mathbf{P}_Y$ des v.a. X et Y sont appelées lois marginales du couple.

Dans la suite les ensembles de valeurs possibles pour les v.a. marginales X et Y seront notés :

$$\begin{aligned} X(\Omega) &= \{x_0, x_1, ..., x_i, ...\} \quad \text{et} \\ Y(\Omega) &= \{y_0, y_1, ..., y_j, ...\} \ . \end{aligned}$$

La loi du couple (X, Y) est caractérisée par les probabilités

$$\mathbf{P}_{X,Y}(\{(x_i, x_j)\}) = \mathbf{P}((X = x_i) \cap (Y = y_j)), \qquad \text{pour } x_i \in X(\Omega),\ y_j \in Y(\Omega) \ .$$

Proposition 5.1
Si (X, Y) est un couple aléatoire, ses lois marginales peuvent se calculer par :

$$\begin{aligned} \forall\, x_i \in X(\Omega), &\qquad \mathbf{P}_X(x_i) = \sum_{y_j \in Y(\Omega)} \mathbf{P}((X = x_i) \cap (Y = y_j)), \\ \forall\, y_j \in Y(\Omega), &\qquad \mathbf{P}_Y(y_j) = \sum_{x_i \in X(\Omega)} \mathbf{P}((X = x_i) \cap (Y = y_j)) \ . \end{aligned}$$

6 Variables aléatoires indépendantes

Définition 6.1 (Indépendance de deux variables aléatoires)
Soient X et Y deux variables aléatoires discrètes définies sur le même espace probabilisé $(\Omega, \mathcal{P}(\Omega), \mathbf{P})$. *On dit que les variables aléatoires X et Y sont indépendantes si et seulement si*

$$(\forall\, x_i \in X(\Omega)),\ (\forall y_j \in Y(\Omega)), \quad \mathbf{P}((X = x_i) \cap (Y = y_j)) = \mathbf{P}(X = x_i)\mathbf{P}(Y = y_j) \ .$$

Proposition 6.1
Soient X et Y deux variables aléatoires discrètes indépendantes, f et g deux fonctions dont les domaines de définition contiennent respectivement $X(\Omega)$ et $Y(\Omega)$. Alors les variables $f(X)$ et $f(Y)$ sont indépendantes.

7 Propriétés de l'espérance et de la variance

7.1 Transformation affine

Soit X une variable aléatoire réelle qui admet une espérance $\mathbf{E}(X)$ et une variance $\mathbf{V}(X)$. Si $(a,\, b) \in \mathbb{R}^2$ alors

$$\begin{aligned} \mathbf{E}(a\,X+b) &= a\,\mathbf{E}(X)+b \\ \mathbf{V}(a\,X+b) &= a^2\mathbf{V}(X)\ . \end{aligned}$$

7.2 Variance d'une somme

Si X et Y sont deux variables aléatoires définies sur le même espace probabilisé et qui ont toutes deux une espérance et une variance alors $(X+Y)$ a une espérance et une variance et on a

$$\begin{aligned} \mathbf{E}(X+Y) &= \mathbf{E}(X)+\mathbf{E}(Y) \\ \mathbf{V}(X+Y) &= \mathbf{V}(X)+\mathbf{V}(Y)+2\,\mathbf{Cov}(X,Y)\ , \end{aligned}$$

où $\mathbf{Cov}(X,Y)$, la *covariance* de $(X,\, Y)$ est définie par

$$\begin{aligned} \mathbf{Cov}(X,\, Y) &= \mathbf{E}((X-\mathbf{E}(X))\,(Y-\mathbf{E}(Y))) \\ &= \mathbf{E}(X\,Y)-\mathbf{E}(X)\,\mathbf{E}(Y)\ . \end{aligned}$$

Le coefficient de corrélation linéaire est défini, lorsque $\sigma(X)\,\sigma(Y)\neq 0$ par

$$r(X,Y)=\frac{\mathbf{Cov}(X,\, Y)}{\sigma(X)\,\sigma(Y)}\ .$$

Attention, comme l'espérance et la variance, la covariance n'existe pas toujours.

Proposition 7.1
*Soient X et Y deux variables aléatoires discrètes **indépendantes**. Alors* $\mathbf{Cov}(X,\, Y)=0$.

Attention ! La réciproque est fausse.
Mais la contraposée de la propriété nous permet de déduire :
Si $\mathbf{Cov}(X,\, Y)\neq 0$, alors les variables X et Y ne sont pas indépendantes.

8 Lois discrètes usuelles

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega,\, \mathcal{P}(\Omega),\, \mathbf{P})$.

8.1 Loi de Bernoulli

Définition 8.1
La v.a. X suit une loi de Bernoulli de paramètre p ($p\in[0;\,1]$) si elle ne prend que des valeurs 0 et 1 avec :

$$\begin{aligned} \mathbf{P}(X=1) &= p \\ \mathbf{P}(X=0) &= 1-p=q\ . \end{aligned}$$

Notation : $X\sim B(p)$.

Elle permet la modélisation de toute expérience où l'on a 2 possibilités (pile ou face, réussite ou échec ...). On peut remarquer que $X^2 = X$, donc X^2 suit également la loi $\sim B(p)$ (de même pour X^n, $n \geq 1$). On a donc

$$E(X) = p$$

$E(X^2) = p$ et

$$V(X) = p - p^2 = p(1-p).$$

8.2 Loi uniforme sur un ensemble fini de réels

Définition 8.2
La v.a. X suit la loi uniforme sur l'ensemble des réels $\{x_1, x_2, ..., x_n\}$ si :

$$\forall\, k \in \mathbb{N}, \quad \mathbf{P}(X = x_k) = \frac{1}{n}\ .$$

Par exemple, le nombre de points indiqué par un dé suit la loi uniforme sur $\{1, 2, 3, 4, 5, 6\}$. L'espérance d'une loi uniforme sur $\{x_1, x_2, ..., x_n\}$ est égale à

$$E(X) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

et sa variance

$$V(X) = \frac{1}{n}\sum_{i=1}^{n} x_i^2 \ - \ (E(X))^2.$$

8.3 Loi binomiale

Définition 8.3
La v.a. X suit une loi binomiale de paramètres n et p ($n \in \mathbb{N}^*$ et $p \in [0;\, 1]$) si l'ensemble des valeurs est $X(\Omega) = \{0,\, 1, ..., n\}$ et :

$$\forall\, k \in \mathbb{N}, \quad \mathbf{P}(X = k) = C_n^k p^k\,(1-p)^{n-k}\ .$$

Notation : $X \sim B(n,\, p)$.

La formule ci–dessus définit bien une loi de probabilité puisque les $C_n^k p^k\,(1-p)^{n-k}$ sont positifs et

$$\sum_{k=0}^{n} C_n^k p^k\,(1-p)^{n-k} = (p + (1-p))^n = 1^n = 1$$

en appliquant la formule du binôme de Newton.

Espérance et variance de la loi binomiale $B(n,\, p)$.
Les espérances de X et X^2 sont données par

$$E(X) = \sum_{k=1}^{n} k\mathbf{P}(X = k) = \sum_{k=1}^{n} kC_n^k p^k\,(1-p)^{n-k}$$

$$E(X^2) = \sum_{k=1}^{n} k^2\mathbf{P}(X = k) = \sum_{k=1}^{n} k^2 C_n^k p^k\,(1-p)^{n-k}$$

qui ne se prêtent pas très bien au calcul. Pour évaluer $E(X)$ et $V(X)$ il vaut mieux utiliser un raisonnement probabiliste fondé sur des propriétés de variables aléatoires. Pour cela on utilise le fait que la loi binomiale est celle du nombre X de succès dans la répétition de n expériences de Bernoulli ayant la même probabilité de succès p. On note X_i le résultat de l'expérience numéro i. La loi de chaque variable X_i est une loi de Bernoulli.

Valeurs de X_i	0	1
Probabilités	p	$(1-p)$

On a donc $E(X_i) = p$ et $V(X_i) = p(1-p)$. Comme X est le nombre de total de succès parmi les n expériences, on peut écrire que

$$X = \sum_{i=1}^{n} X_i$$

L'espérance d'une somme est égale à la somme des espérances, on a donc

$$E(X) = \sum_{i=1}^{n} E(X_i) = np.$$

La variance d'une somme de variables aléatoires *indépendantes* est égale à la somme des variances donc

$$V(X) = \sum_{i=1}^{n} V(X_i) = np(1-p).$$

8.4 Loi hypergéométrique

Alors que la loi binomiale intervient dans les tirages avec remise, la loi hypergéométrique correspond aux tirages sans remise.

Exemple 8.1
Dans une production totale de N objets dont M sont défectueux, on prélève au hasard un échantillon de n objets (tirage sans remise). Soit X le nombre aléatoire d'objets défectueux dans l'échantillon. Quelle est sa loi ?

On peut prendre comme espace Ω l'ensemble de tous les échantillons possibles (toutes les parties à n éléments d'un ensemble de cardinal N) muni de l'équiprobabilité. Chaque échantillon a ainsi une probabilité $\frac{1}{C_N^n}$ d'être choisi. Les échantillons réalisant l'événement $\{X = k\}$ sont ceux qui contiennent k objets défectueux et $n-k$ objets non défecteux. Il y a $C_M^k C_{N-M}^{n-k}$ échantillons possibles. Ceci n'est réalisé que pour $0 \leq k \leq M$ et $0 \leq n-k \leq N-M$. On a donc :

$$\begin{aligned} \mathbf{P}(X = k) &= \frac{C_M^k \, C_{N-M}^{n-k}}{C_N^n} \qquad \text{si } 0 \leq k \leq \min(M,\, n) \text{ et} \\ &= 0 \text{ sinon.} \end{aligned}$$

Définition 8.4
La loi définie ci-dessus s'appelle la loi hypergéométrique de paramètres N, M et n. N est l'effectif de la population totale, M celui de la sous-population à laquelle on s'intéresse et n la taille de l'échantillon.
Notation : $X \sim H(N,\, n,\, p)$ avec $p = \frac{M}{N}$.

On peut montrer que

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

Pour une taille d'échantillon n fixée, plus N et M sont grands, moins les tirages sans remise diffèrent des tirages avec remise.

8.5 Loi géométrique

Exemple 8.2
Considérons une suite infinie d'épreuves répétées indépendantes avec même probabilité de succès $p \in]0,1[$. Soit X le numéro de la première épreuve où l'on obtient un succès. Calculer $P(X = k)$ pour tout $k \in \mathbb{N}^*$.
En notant E_i l'événement "succès à la i–ème épreuve",

$$\{X = k\} = \left(\prod_{i=1}^{k-1} \bar{E}_i\right) \cap E_k.$$

D'où par indépendance des épreuves,

$$\mathbf{P}(X = k) = \left(\prod_{i=1}^{k-1} \mathbf{P}(\bar{E}_i)\right) \mathbf{P}(E_k) = (1-p)^{k-1}p\ .$$

Définition 8.5
Une variable aléatoire X suit une loi géométrique de paramètre $p \in]0;1[$ si $X(\Omega) = \mathbb{N}^*$ et

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

Notation : $X \sim G(p)$.

On montre que

$$E(X) = \frac{1}{p} \quad et$$
$$V(X) = \frac{1-p}{p^2}$$

8.6 Loi de Poisson

Définition 8.6
On dit qu'une variable aléatoire X suit la loi de Poisson de paramètre $\lambda > 0$ si $X(\Omega) = \mathbb{N}$ et

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

Notation : $X \sim P(\lambda)$.

Une des raisons de l'importance de cette loi est le théorème de convergence de la loi binomiale vers la loi de Poisson. Lorsqu'on répète un très grand nombre de fois des expériences de Bernoulli indépendantes de même paramètre p très petit, on peut utiliser la loi de probabilité de Poisson. En général, $n > 30$ et $np \in [0.1, 18]$ donnent une bonne approximation.
Elle s'applique par exemple au nombre d'erreurs commises lorsque la probabilité d'erreur est très

petite et le nombre de répétitions indépendantes très grand (fautes de frappe dans l'ensemble d'un livre, pièces défectueuses lorsque les tests sont nombreux et la production de bonne qualité, ...), au nombre d'appels à un serveur ou au nombre de clients qui arrivent dans un magasin pendant un temps donné, etc.
On peut montrer que si X suit la loi $P(\lambda)$ alors

$$E(X) = \lambda \quad \text{et} \quad V(X) = \lambda.$$