Table des matières

I La base théorique

I.1 Fonction probabilité

I.1.1 Définition (Événement)

On se donne un ensemble Ω appelé univers. Ω est l'ensemble des réalisations envisageables pour une expérience aléatoire.

Un événement est un sous-ensemble de Ω qui regroupe donc une ou plusieurs des réalisations possibles.

On définit souvent un événement par une phrase. Par exemple, pour un lancé de dé à 6 faces, l'ensemble Ω peut-être $[\![1,6]\!]$. On peut alors définir les événements :

— Obtenir un 2. (c'est l'ensemble $\{2\}$ de cardinal 1)

— Obtenir un résultat impair (c'est l'ensemble $\{1,3,5\}$)

— ...

I.1.2 Définition

Pour définir une probabilité, il faut d'abord donner un univers Ω et un ensemble T contenant des événements.

Une probabilité est alors une fonction $\mathbb{P}$ qui associe à chaque événement un réel dans $[0,1]$ et qui vérifie en plus une condition : si les $(A_n)_{n\in\mathbb{N}}$ sont des événements **deux à deux disjoints** alors

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

c'est-à-dire que cette série converge, et que sa somme est un réel de $[0,1]$.

I.2 Variables aléatoires

I.2.1 Définition

— Une variable aléatoire réelle est une fonction $X : \Omega \to \mathbb{R}$. À chaque réalisation possible, on associe un nombre réel.

X doit de plus vérifier la condition (voir la remarque suivante) que l'ensemble $X^{-1}(\{x\})$ est un événement pour tout $x \in X(\Omega)$.

— $X(\Omega)$ est l'ensemble des valeurs que X peut prendre.

— On dit que X est une variable aléatoire discrète ssi $X(\Omega)$ est fini ou dénombrable.

I.2.2 Variable aléatoire et événements

— Il est rarement demandé de prouver qu'une fonction est une variable aléatoire.

— Il faut toujours bien noter quel est l'ensemble $X(\Omega)$.

— Par définition, si x est l'une des valeurs possibles pour X, on peut considérer l'événement $(X = x)$ qui se lit "X prend la valeur x" et pas "X égal x". Noter d'ailleurs les parenthèses.

Par abus de notation, on note $\mathbb{P}(X = x)$ à la place de $\mathbb{P}((X = x))$ exactement de la même manière que l'on note $f(x,y)$ au lieu de $f((x,y))$ pour des fonctions de deux variables.

Il faut tout de même garder en tête que $X = x$ (sans ses parenthèses) n'a pas de sens.

— On obtient également des événements en considérant, pour $a \in \mathbb{R}$,

$$(X \leqslant a), (X > a), \dots.$$

Les probabilités de ces événements peuvent toujours se mettre sous la forme d'une somme de probabilités de la forme $\mathbb{P}(X = x_n)$.

— Si X est une variable discrète (ce qui sera toujours le cas pour nous), on peut numéroter ses valeurs :

$$X(\Omega) = \{x_0, x_1, x_2, \dots\} = \{x_n;\ n \in \mathbb{N}\}$$

Très souvent, la numérotation est naturelle et on peut s'en passer ($X(\Omega) = \mathbb{N}$ par exemple).

— La loi d'une variable aléatoire est la fonction

$$\left\{ \begin{array}{rcl} X(\Omega) & \to & [0,1] \\ x & \mapsto & \mathbb{P}(X = x) \end{array} \right.$$

Pour donner la loi de X il faut donc :

1. Exhiber clairement $X(\Omega)$.
2. Pour chaque valeur possible x, donner $\mathbb{P}(X = x)$.

— Avec la numérotation précédente pour $X(\Omega)$, on a

$$\sum_{n=0}^{+\infty} \mathbb{P}(X = x_n) = 1,$$

car les événements $((X = X_n))_{n \in \mathbb{N}}$ forment un système complet d'événements (voir plus loin).

— En pratique, la vérification que l'égalité précédente a bien lieu suffit pour vérifier que X est une variable aléatoire.

II Calculs sur les événements

II.1 Union, intersection

II.1.1 Cas favorable : l'union disjointe

Si des événements $(A_n)_n$ sont deux à deux **disjoints** ou **incompatibles** (ils représentent par exemple une disjonction de cas), alors

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

La probabilité d'être dans l'un de ces cas se calcule par somme.

Évidemment, cette formule est valable même pour seulement deux ou 3 événements (et la somme a un nombre fini de termes).

II.1.2 Cas général

Considérons deux événements A, B pas forcément disjoints. En notant $\bar{A}, \bar{B}$ les événements contraires, on a

$$\mathbb{P}(A \cup B) = 1 - \mathbb{P}(\bar{A} \cap \bar{B})$$

En général, une probabilité d'intersection est plus facilement calculable.

II.1.3 Inégalités et variables aléatoires

On considère X une variable aléatoire réelle et $x \in \mathbb{R}$.

$$\mathbb{P}(X \geqslant x) = 1 - \mathbb{P}(X < x)$$

Si on sait que X est à valeurs entières ($X(\Omega) \subset \mathbb{N}$) et si $n \in \mathbb{N}$

$$\mathbb{P}(X = n) = \mathbb{P}(X \geqslant n) - \mathbb{P}(X \geqslant n+1) = \mathbb{P}(X \leqslant n) - \mathbb{P}(X \leqslant n-1)$$

II.1.4 Intersections

Pour A, B deux événements

$$\mathbb{P}(A \cap B) = \begin{cases} \mathbb{P}(A)\mathbb{P}(B) \text{ si } A, B \text{ sont indépendants} \\ \mathbb{P}(A|B) \times \mathbb{P}(B) \text{ sinon} \end{cases}$$

Le cas général pour des événements non indépendants se traite en utilisant la formule des probabilités composées qui peut se reconstruire pas à pas en considérants un par un les événements en jeu.

II.2 Systèmes complets

II.2.1 Définition

Notons, pour $n \in \mathbb{N}$, A_n un événement (donc A_n est un sous-ensemble de Ω).

On dit que $(A_n)_{n\in\mathbb{N}}$ est un système complet d'événements ssi $\forall (i,j) \in \mathbb{N}^2\ i \neq j \Rightarrow A_i \cap A_j = \emptyset$ (disjoints 2 à 2) et $\bigcup\limits_{n\in\mathbb{N}} A_n = \Omega$.

II.2.2 Définition

Notons, pour $n \in \mathbb{N}$, A_n un événement (donc A_n est un sous-ensemble de Ω).

On dit que $(A_n)_{n\in\mathbb{N}}$ est un système quasi-complet d'événements ssi

$$\forall (i,j) \in \mathbb{N}^2\ i \neq j \Rightarrow A_i \cap A_j = \emptyset$$

(les A_n sont disjoints 2 à 2) et

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = 1.$$

En d'autres termes, la probabilité qu'aucun des A_n ne soit réalisé est nulle, et l'un exactement des A_n est réalisé presque sûrement (avec une probabilité 1, un et un seul des A_n est réalisé).

II.2.3 Exemples fondamentaux

- Si A est un événement, alors $(A, \bar{A})$ est un système complet d'événements.
- Si X est une variable aléatoire discrète, la famille $((X = x))_{x\in X(\Omega)}$ est un système complet d'événements. On considère tous les événements possibles "X prend la valeur...".

II.2.4 L'idée fondamentale

Si B est un événement (lié à une variable aléatoire ou non) et $(A_n)_{n\in\mathbb{N}}$ est un système quasi-complet (ou complet)

$$\mathbb{P}(B) = \sum_{n=0}^{+\infty} \mathbb{P}(B \cap A_n)$$

où la somme peut avoir un nombre fini de termes.

II.2.5 Applications

- Probabilités totales : $\mathbb{P}(B) = \sum\limits_{n=0}^{+\infty} \mathbb{P}(B|A_n) \times \mathbb{P}(A_n)$.
- Utilisation d'une variable aléatoire : notons $X(\Omega) = \{x_n ; n \in \mathbb{N}\}$

  $$\mathbb{P}(B) = \sum_{n=0}^{+\infty} \mathbb{P}(B \text{ et } X = x_n)$$

  Ceci est très souvent utilisé dans le cas où B est défini en utilisant une autre variable aléatoire.
- Formule de Bayes : lorsque cela a du sens, on a évidemment (revenir à la définition d'une probabilité conditionnelle)

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

  Quand on applique cette formule, $\mathbb{P}(B)$ est très souvent calculée par la formule des probabilités totales (car on connaît des probabilités de la forme $\mathbb{P}(B|...)$).

III Calculs sur les variables aléatoires

III.1 Lois usuelles

III.1.1 Loi uniforme

— Notation : $X \hookrightarrow \mathcal{U}(E)$ où $E = \{x_1, \dots, x_n\}$ est un ensemble fini.

— Valeurs possibles : $X(\Omega) = E$.

— Loi : $\forall x \in E\ \mathbb{P}(X = x) = \frac{1}{n} = \frac{1}{\text{Card}(E)}$.

— Interprétation : X prend chacune de ses valeurs avec la même probabilité.

— Exemple : X est le résultat d'un lancé d'un dé à 6 faces équilibré.

III.1.2 Loi de Bernoulli

— Notation : $X \hookrightarrow \mathcal{B}(p)$ où $p \in [0, 1]$

— Valeurs possibles : $X(\Omega) = \{0, 1\}$

— Loi : $\mathbb{P}(X = 1) = p$, $\mathbb{P}(X = 0) = 1 - p$

— Interprétation : s'utilise lorsqu'on peut interpréter le résultat de l'expérience comme un échec (X prendra la valeur 0) ou un succès (X prendra la valeur 1).

— Exemple : pour le même lancé de dé, on considère comme un succès le fait d'obtenir 5 ou plus comme résultat. On crée ainsi une variable aléatoire suivant une loi de Bernoulli de paramètre $p = \frac{1}{3}$.

III.1.3 Loi binomiale

— Notation : $X \hookrightarrow \mathcal{B}(n, p)$ où $n \in \mathbb{N}\backslash\{0\}$ et $p \in [0, 1]$.

— Valeurs possibles : $X(\Omega) = [\![1, n]\!]$.

— Loi : $\forall k \in [\![1, n]\!]\ \mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$.

— Interprétation : on considère maintenant la répétition **indépendante** de n expériences de Bernoulli (succès ou échec) de même paramètre p (probabilité de succès). Le nombre total de succès au cours de ces n expériences suit une loi binomiale de paramètres n et p.

— Remarque : Si on note $X_1, \dots, X_n$ des variables aléatoires indépendantes suivant une même loi de Bernoulli $\mathcal{B}(p)$, alors

$$\sum_{k=1}^{n} X_k = S \hookrightarrow \mathcal{B}(n, p),$$

car la somme des valeurs obtenues pour les X_k est exactement le nombre de 1.

III.1.4 Loi géométrique

— Notation : $X \hookrightarrow \mathcal{G}(p)$ où $p \in]0, 1[$ (on exclut 0 et 1 cette fois).

— Valeurs possibles : $X(\Omega) = \mathbb{N}\backslash\{0\}$ qui est un ensemble infini.

— Loi : $\forall n \in \mathbb{N}\backslash\{0\}\ \mathbb{P}(X = n) = p \times (1 - p)^{n-1}$.

— Interprétation : on répète indéfiniment et de manière indépendante une même expérience de Bernoulli de paramètre p et on note X le rang d'apparition du premier succès. Alors X suit une loi géométrique de paramètre p.

Attention à ne pas confondre ces deux lois.
Loi binomiale : nombre de succès.
Loi géométrique : rang du premier succès.
Pour ces deux lois, les expériences doivent être indépendantes et de même probabilité de succès.

III.1.5 Loi de Poisson

— Notation : $X \hookrightarrow \mathcal{P}(\lambda)$ où $\lambda \in]0, +\infty[$.

— Valeurs possibles : $X(\Omega) = \mathbb{N}$.

— Loi : $\forall n \in \mathbb{N}\ \mathbb{P}(X = n) = e^{-\lambda} \frac{\lambda^n}{n!}$.

— Interprétation : λ est l'espérance de X (voir plus loin).

III.2 Lois conjointes

Rappel : la loi conjointe de deux variables aléatoires X et Y est la loi du couple (X, Y).

III.2.1 Notation

Si x_n est une valeur possible de X ($x_n \in X(\Omega)$) et $y_m \in Y(\Omega)$, on note $(X = x_n, Y = y_m)$ l'événement $(X = x_n)$ et $(Y = y_n)$. Il s'agit donc d'une intersection d'événements.

Dans le cas où X, Y ne prennent qu'un nombre fini de valeurs, on peut représenter la loi conjointe dans un tableau à double entrée.

III.2.2 Lorsqu'elle est donnée : lois marginales

Dans le cas où l'énoncé donne toutes les probabilités de la forme $\mathbb{P}(X = x_n, Y = y_m)$ pour toutes les valeurs possibles de X et Y

1. On peut retrouver la loi de X (première loi marginale) en fixant un $n \in \mathbb{N}$ et une valeur $x_n \in X(\Omega)$:
$$\mathbb{P}(X = x_n) = \sum_{m=0}^{+\infty} \mathbb{P}(X = x_n, Y = y_m)$$
2. On peut retrouver la loi de Y en fixant un $m \in \mathbb{N}$ et une valeur $y_m \in Y(\Omega)$:
$$\mathbb{P}(Y = y_m) = \sum_{n=0}^{+\infty} \mathbb{P}(X = x_n, Y = y_m)$$

Bien remarquer que l'indice change : on somme sur les valeurs de Y pour retrouver la loi de X et vice-versa pour la loi de Y.

III.2.3 Lorsqu'il faut la calculer

Dans le cas où l'énoncé ne donne pas explicitement la loi conjointe :

— il donne un moyen de calculer des probabilités d'intersection (indépendance, probabilité conditionnelle, ...)

— on en déduit la loi conjointe puis les lois marginales.

Souvent, la loi conjointe est définie par cas et certaines probabilités sont nulles. Il faut penser à adapter l'ensemble d'indices des sommes à calculer.

Exemple On lance une pièce truquée qui tombe sur pile avec une probabilité $p \in]0, 1[$. On note X le rang d'apparition du premier pile et Y le rang d'apparition du second. Donner la loi conditionnelle de Y sachant $(X = n)$ pour un $n \in \mathbb{N} \backslash \{0\}$, en déduire la loi conjointe puis la loi de Y.

Réponse : $\forall m \geqslant 2$ $\mathbb{P}(Y = m) = mp(1 - p)^{m-2}$.

III.3 Espérance, variance, covariance

III.3.1 Définition

Soit X une variable aléatoire discrète à valeurs dans $X(\Omega) = \{x_n |\ n \in \mathbb{N}\} \subset \mathbb{R}$. On dit que X est d'espérance finie ssi $\sum x_n \mathbb{P}(X = x_n)$ converge **absolument**.

Dans ce cas, on appelle **espérance de** X et on note $E(X)$ le réel $\sum\limits_{n=0}^{+\infty} x_n \mathbb{P}(X = X_n)$.

Il s'agit de la traduction de "la valeur moyenne de X" : on la calcule comme moyenne pondérée des valeurs par leurs probabilités d'apparition.

III.3.2 Propriétés calculatoires

— Linéarité, positivité, croissance.

— Théorème de transfert. Si X est une variable aléatoire et $Y = f(X)$ est une variable calculée en fonction de X. Avec les notations précédentes pour les valeurs de X, Y est d'espérance finie ssi $\sum f(x_n)\mathbb{P}(X = x_n)$ converge absolument et alors
$$E(Y) = E(f(X)) = \sum_{n=0}^{+\infty} f(x_n)\mathbb{P}(X = x_n)$$
Le principal point de ce théorème est que la connaissance de la loi de Y n'est pas nécessaire au calcul de $E(Y)$.

— Cas classique où $X(\Omega) \subset \mathbb{N}$ (les valeurs de X sont des entiers naturels) En cas de convergence absolue, on a

$$E(Y) = E(f(X)) = \sum_{n=0}^{+\infty} f(n)\mathbb{P}(X = n)$$

— Indépendance : SI X, Y sont indépendantes, alors $E(XY) = E(X)E(Y)$.

III.3.3 Fonction génératrice

Pour X une variable aléatoire à valeurs dans $\mathbb{N}$ on peut poser

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

Il s'agit de la somme d'une série entière définie et continue sur $[-1, 1]$ au moins. Par unicité des coefficients, si on connaît G_X alors on connaît la loi de X.

Lorsque G_X est dérivable en 1, on a

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

Cette formule se retrouve par l'application non justifiée du théorème de dérivation terme à terme.

III.3.4 Variance

— X est de variance finie ssi (théorème de transfert) $\sum x_n^2 \mathbb{P}(X = x_n)$ converge absolument. Si X est à valeurs réelles, cette série est en fait à termes positifs.

— $V(X) = E(X^2) - E(X)^2$ est souvent la forme la plus efficace pour le calcul effectif.

— $V(aX) = a^2(X)$ est une formule homogène : $V(X)$ à la même unité que X^2.

— $V(X + b)$ s'interprète bien : $V(X)$ mesure la "dispersion" des valeurs de X, et translater les valeurs de X ne change pas cette dispersion.

— $V(X) = 0$ ssi X est presque sûrement constante. C'est-à-dire qu'une seule valeur de X a une image non nulle par la loi de X : $\mathbb{P}(X = x_n) = 0$ sauf pour un x_n.

— Si X est en plus à valeurs dans $\mathbb{N}$, on peut utiliser la fonction génératrice pour calculer $V(X)$. Pour $t \in]-1, 1[$

$$G_X"(t) = \sum_{n=2}^{+\infty} (n^2 - n)\mathbb{P}(X = n)t^{n-2}$$

et on peut en déduire $E(X^2)$ si la formule peut s'appliquer en 1 : $G_X"(1) = E(X^2) - E(X)$.
(on a complété les sommes de séries par des termes nuls.)

III.3.5 Covariance

— Lorsque X, Y sont de variances finies, on peut définir $\text{cov}(X, Y) = E(XY) - E(X)E(Y)$. Cette quantité est nulle lorsque X, Y sont indépendantes.

— La covariance possède les propriétés d'un produit scalaire, sauf le caractère défini. La "norme" associée est la variance. On retrouve ainsi les formules faisant le lien entre les 2.

— L'équivalent du théorème de Pythagore est très utile. Si $X_1, \dots, X_n$ sont indépendantes et de variances finies, alors $\sum\limits_{k=1}^{n} X_k$ est aussi de variance finie et

$$V\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} V(X_k)$$

On retrouve ainsi la variance d'une loi binomiale.

III.4 Inégalités

1. Markov. S'applique à une variable à valeurs positives telle que $E(X)$ existe.

$$\mathbb{P}(X \geqslant a) \leqslant \frac{E(X)}{a} \text{ pour tout } a > 0$$

Cette inégalité est intéressante pour $a > E(X)$

2. Bienaymé-Tchebychev : cette fois la condition est que $V(X)$ existe. Pour $\varepsilon > 0$

$$\mathbb{P}(|X - E(X)| > \varepsilon) \leqslant \frac{V(X)}{\varepsilon^2}$$

   On mesure ici la probabilité que X prenne des valeurs éloignées de son espérance (sa "moyenne").
   Le dénominateur est forcément au carré pour une raison d'homogénéité.

3. Loi faible des grands nombres. On considère maintenant plusieurs variables aléatoires de même loi et de variances finies, indépendantes. Ceci modélise la répétition indépendante d'une même expérience aléatoire.

   Ce théorème s'intéresse à la moyenne arithmétique $M_n = \frac{1}{n} \sum_{k=1}^{n} X_k$ des n premières variables, par rapport à la "moyenne" théorique $m = E(X_1) = E(M_n)$.

$$\forall \eta > 0 \ \mathbb{P}\left(|M_n - m| \geqslant \eta\right) \underset{n \to +\infty}{\to} 0.$$

   On fixe d'abord $\eta > 0$. Alors en faisant tendre le nombre de répétitions vers l'infini (en répétant "suffisamment" de fois), $M_n \in [m - \eta, m + \eta]$ avec une probabilité aussi proche de 1 que voulue.
   Intuitivement, on peut estimer la probabilité théorique de succès en répétant suffisamment de fois une même expérience de Bernoulli.

IV Résumé sur les lois usuelles

Voir le tableau page suivante

Nom	Notation	Valeurs	Loi	Fonctions génératrice	Espérance	Variance
Bernoulli	$\mathcal{B}(p)$	$\{0,1\}$	$\mathbb{P}(X=1)=p,\ \mathbb{P}(X=0)=1-p$	$G_X(t)=1-p+pt,\ t\in\mathbb{R}$	p	$p(1-p)$
Binomiale	$\mathcal{B}(n,p)$	$[\![0,n]\!]$	$\mathbb{P}(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$	$G_X(t)=(1-p+pt)^n,\ t\in\mathbb{R}$	np	$np(1-p)$
Géométrique	$\mathcal{G}(p)$	$\mathbb{N}\backslash\{0\}$	$\mathbb{P}(X=k)=p(1-p)^{k-1}$	$G_X(t)=\frac{pt}{1-(1-p)t},\ t\in]-\frac{1}{1-p},\frac{1}{1-p}[$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Poisson	$\mathcal{P}(\lambda),\lambda>0$	$\mathbb{N}$	$\mathbb{P}(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}$	$G_X(t)=e^{\lambda(t-1)},\ t\in\mathbb{R}$	λ	λ