Chapitre 13 : probabilités sur un univers infini

Tables des matières

I. Introduction

Ce chapitre complète et généralise les résultats établis au chapitre 8, dans lequel on se cantonnait aux univers finis. Les principales notions sont reprises dans ce nouveau cadre. Certaines changent peu, voire pas du tout, mais dans certains cas, les énoncés doivent être adaptés et des précautions prises. Par exemple, les sommes qui intervenaient dans le chapitre 8 (somme des probabilités ou espérance) sont alors remplacées par des séries dont la convergence n'est a priori pas assurée.

II. Cadre généralisé des probabilités

1) Exemples fondamentaux

Exemple 1 (Univers infinis).
1. On jette un pièce et on note le premier rang pour lequel on tombe sur Pile. Alors $\Omega = \mathbb{N}^*$ est infini et dénombrable.
2. On jette indéfiniment un dé à six faces et on note la suite des résultats obtenus. $\Omega = [\![1\,;6]\!]^{\mathbb{N}}$ est l'ensemble des suites d'éléments de $[\![1\,;6]\!]$, donc infini et dénombrable.
3. On note le temps d'attente d'un bus, qui est un nombre aléatoire dans l'intervalle $[0\,;10]$. Cette univers est infini et non dénombrable (c'est à dire beaucoup plus grand que $\mathbb{N}$).

.

Exemple 2 (Intersections/réunions infinies d'événements). ♡ ✎ On considère la seconde expérience modélisée dans l'exemple précédent et on note, pour tout entier naturel non nul i, A_i : « le résultat du ième lancer est 1 ».

Exprimer les événements suivants à l'aide des événements A_i :
1. B_i : « Le premier 1 apparaît au i-ème lancer »
2. C : « Au moins un des résultats est un 1 »
3. D : « Aucun lancer ne donne un 1 »

2) Espace probabilisable

Définition. Soit Ω un ensemble et $A \subset \mathcal{P}(\Omega)$. $\mathcal{A}$ est une tribu de parties de Ω si :
1. $\Omega \in \mathcal{A}$
2. $\forall A \in \mathcal{A}$, on a $\bar{A} \in \mathcal{A}$
3. Si pour tout $n \in \mathbb{N},\ A_n \in \mathcal{A}$, alors $\displaystyle\bigcup_{n=0}^{+\infty} A_n \in \mathcal{A}$. Le couple $(\Omega,\ \mathcal{A})$ est appelé espace probabilisable et les éléments de $\mathcal{A}$ sont appelés les événements.

Remarque . $\mathcal{P}(\Omega)$ est évidemment une tribu. C'est la seule qu'on utilise lorsque Ω est un ensemble fini. Lorsque Ω est infini, $\mathcal{P}$ devient énorme et il est alors souvent commode de considérer des tribus plus petites qui suffisent pour appliquer notre modèle. De plus, on peut montrer qu'il est impossible de construire une probabilité sur $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$, ce qui nous oblige à considérer des tribus plus petites qui permettent les calculs de probabilités.

Exemple 3. ✎ Soit $\Omega = \{1;\ 2;\ 3;\ 4\}$. Donner une tribu à deux éléments et une tribu à quatre éléments.

Proposition 1.

Avec les notations de la définition précédente, on a :

1. $\emptyset \in \mathcal{A}$.
2. Les intersections et réunions d'éléments de $\mathcal{A}$ sont des éléments de $\mathcal{A}$.
3. Si pour tout $n \in N,\ A_n \in \mathcal{A}$, alors $\displaystyle\bigcap_{n=0}^{+\infty} A_n \in \mathcal{A}$
4. $\displaystyle\overline{\bigcup_{n=0}^{+\infty} A_n} = \bigcap_{n=0}^{+\infty} \overline{A_n}$ et $\displaystyle\overline{\bigcap_{n=0}^{+\infty} A_n} = \bigcup_{n=0}^{+\infty} \overline{A_n}$

La définition suivante a été vue au chapitre 8. A présent, l'union peut concerner une infinité d'ensembles.

Définition. Soit $(\Omega,\ \mathcal{A})$ un espace probabilisable, et soit $(A_i)_{i\in I}$ une famille d'événements de $\mathcal{A}$. Cette famille est un système complet d'événements (SCE) si :

1. Les A_i sont 2 à 2 incompatibles :
$$\forall i,\ j \in I, A_i \cap A_j = \emptyset$$
2. Leur réunion est Ω : $\displaystyle\bigcup_{i\in I} A_i = \Omega$

Exemple 4. ✎ Les familles $(A_i)_{i\in\mathbb{N}^*}$ et $(B_i)_{i\in\mathbb{N}^*}$ de l'exemple 2 sont-elles des systèmes complets d'événements ?

3) Espace probabilisé

Une fois les espaces probabilisables définis, on a un cadre théorique qui permet potentiellement de mesurer les chances que les différents événements se produisent. Mais il y a de nombreuses manières de le faire. Par exemple, si on considère une pièce équilibrée ou non dans l'exemple 1.1, les événements resteront les mêmes mais le calcul des probabilités sera différent. La définition suivante généralise ce qu'est une probabilité au cas des univers infini.

Définition. Soit $(\Omega,\ \mathcal{A})$ un espace probabilisable. On appelle probabilité sur $(\Omega,\ \mathcal{A})$ toute application P de $\mathcal{A}$ dans $[0\,;1]$ vérifiant :

1. $P(\Omega) = 1$.
2. Si $(A_n)_{n\in n\in\mathbb{N}}$ est une famille d'événements 2 à 2 incompatibles, alors $\sum P(A_n)$ converge et : $\displaystyle P(\bigcup_{n=0}^{+\infty} A_n) = \sum_{n=0}^{+\infty} P(A_n)$. On appelle alors $(\Omega,\ \mathcal{A},\ P)$ un espace probabilisé.

Proposition 2.

$(\Omega,\ \mathcal{A},\ P)$ un espace probabilisé, $(A_i)_{i\in I}$ un SCE. Alors la série $\displaystyle\sum_{i\in I} P(A_i)$ converge et sa somme vaut 1.

Définition. Soit $(\Omega,\ \mathcal{A},\ P)$ un espace probabilisé.

E est dit négligeable si $P(E) = 0$ et presque sûr (ou vrai presque sûrement) si $P(E) = 1$.

Remarque . ⚠Dans le cas fini, les notions d'événements impossibles et d'événements négligeables se confondent. De même pour les notions d'événements certains et presque sûrs. Mais dans ce nouveau cadre, ce n'est plus le cas. Il faut être particulièrement attentif à cette différence lorsqu'on considère des SCE et la formule des probabilités totales.

Exemple 5. ✎ Dans le cadre de l'exemple 2 avec un dé équilibré, que dire intuitivement des événements C et D ?

Les théorèmes suivants, valables dans tout espace probabilisé $(\Omega,\ \mathcal{A},\ P)$, permettent de démontrer ce genre de résultats.

Théorème 3.« Théorème de la limite monotone »

1. Si $(A_n)_{n\in\mathbb{N}}$ est une suite croissante d'événements (ie $\forall n \in \mathbb{N}, A_n \subset A_{n+1}$), alors $P(\bigcup_{n=0}^{+\infty} A_n) = \lim_{n\to+\infty} P(A_n)$.
2. Si $(A_n)_{n\in\mathbb{N}}$ est une suite décroissante d'événements (ie $\forall n \in N, A_{n+1} \subset A_n$), alors $P(\bigcap_{n=0}^{+\infty} A_n) = \lim_{n\to+\infty} P(A_n)$.

Théorème 4.« Corollaire du théorème de la limite monotone »

1. Si $(A_n)_{n\in\mathbb{N}}$ est une suite quelconque d'événements, alors $P(\bigcup_{n=0}^{+\infty} A_n) = \lim_{n\to+\infty} P(\bigcup_{i=0}^{n} A_i)$.
2. Si $(A_n)_{n\in\mathbb{N}}$ est une suite quelconque d'événements, alors $P(\bigcap_{n=0}^{+\infty} A_n) = \lim_{n\to+\infty} P(\bigcap_{i=0}^{n} A_i)$.

Démonstration.

Poser $B_n = \bigcup_{i=0}^{n} A_i$, et appliquer le théorème de la limite monotone. On procède de même pour l'intersection.

Exemple 6. ♡ ✎ On reprend le cadre de l'exemple 2 avec un dé équilibré et les lancers mutuellement indépendants. On note de plus C_n : « au moins un des résultats est un 1 lors des n premiers lancers ».

1. (a) Exprimer C_n à l'aide des A_i.
   (b) Déduire que $P(C_n) = 1 - P(\overline{A_1} \cap \cdots \cap \overline{A_n})$
   (c) Calculer alors $P(C_n)$.
   (d) Justifier que $(C_n)_{n\in\mathbb{N}^*}$ est une suite croissante d'événements.
   (e) Montrer que C est presque sûr et D négligeable.
   (f) Exprimer C à l'aide des A_i. La suite $(A_n)_{n\in\mathbb{N}^*}$ est-elle une suite monotone d'événements ?
   (g) Appliquer la variante du théorème de la limite monotone pour retrouver que C est presque sûr.
2. (a) Exprimer D en fonction des A_i, puis en fonction des B_i.
   (b) Redémontrer que D est négligeable de deux manières.

4) Généralisation des propriétés des probabilités

Presque tout ce qui a été vu dans le cas fini fonctionne à l'identique dans le cas infini : définition de la probabilité conditionnelle, indépendance d'événements, formule de Bayes, formule des probabilités composées.

Il faut simplement généraliser l'indépendance mutuelle d'événements et la formule des probabilités totales.

Définition. Les événements d'une famille infinie $(A_n)_{n\in\mathbb{N}}$ sont mutuellement indépendants si, pour toute partie finie I de N, on a : $P(\bigcap_{n\in I} A_n) = \prod_{n\in I} P(A_n)$

Théorème 5.

Soit $(A_n)_{n\in\mathbb{N}}$ un SCE alors pour tout événement B, la série $\sum P(B\cap A_n)$ converge et $P(B) = \sum_{n=0}^{+\infty} P(B\cap A_n)$

Si de plus, pour tout $n \in N$, on a $P(A_n) \neq 0$, on a $P(B) = \sum_{n=0}^{+\infty} P_{A_n}(B)P(A_n)$

Remarque . Si seulement un nombre fini d'événements sont non négligeables, on peut quand même écrire $P(B) = \sum_{n \in I} P_{A_n}(B)P(A_n)$, pour I l'ensemble des indices d'événements non négligeables. De sorte que les événements négligeables participent à la construction du SCE mais n'interviennent plus dans la formule finale (mais il ne faut pas les oublier !).

Exemple 7. ✎ On se place dans le cadre de l'exemple 2.

1. Exprimer B_n en fonction des A_i. Déduire $P(B_n)$. Vérifier que l'on a bien $\sum_{k=1}^{+\infty} P(B_k) = 1$
2. Calculer la probabilité de l'événement E : « La première fois qu'un 1 est sorti, il est suivi d'un autre 1 ».
3. On joue au jeu : jeter un dé jusqu'à obtenir un 1. On appelle alors n le nombre de lancers nécessaires pour obtenir ce 1. On jette alors une pièce équilibrée n fois. On gagne si lors de ces n lancers de pièce, au moins un Pile est sorti. Soit F : « gagner à ce jeu ». Déterminer $P(F)$.

III. Généralités sur les variables aléatoires

1) Définition généralisée. Système complet d'événements associé à une variable aléatoire

Dans le chapitre 8, on avait défini une variable aléatoire comme une application qui, à un événement élémentaire, associe un nombre réel. Dans le cas infini, l'idée est la même mais il faut modifier la définition pour des raisons théoriques.

Définition.

Soit $(\Omega,\ \mathcal{A})$ un espace probabilisable. X est une variable aléatoire réelle définie sur $(\Omega,\ \mathcal{A})$ si X est une application de Ω dans $\mathbb{R}$ telle que, pour tout élément x de $\mathbb{R}$, l'ensemble $[X \leqslant x] = \{\omega \in \Omega, X(\omega) \leqslant x\}$ est dans $\mathcal{A}$. On note encore $X(\Omega)$ l'univers image. X est une variable aléatoire discrète si $X(\Omega) = \{x_i,\ i \in I\}$, où I est un sous-ensemble de $\mathbb{N}$ ou de $\mathbb{Z}$, éventuellement infini. X est une variable aléatoire finie si $X(\Omega)$ est un ensemble fini.

Remarque .

1. Le fait de ne plus centrer la définition sur les événements élémentaires vient du fait que dans le cas infini et notamment non dénombrable, tous ces événements peuvent être de probabilité nulle, ce qui ne donnerait finalement aucune information sur la loi de X.
2. La condition théorique qui est ajoutée dans cette définition sert à garantir qu'on puisse bien calculer la probabilité que $X \leqslant x$ puisque la probabilité P est définie sur $\mathcal{A}$. Tout ceci était automatiquement vérifié dans le cas dans le cas fini car la tribu choisi est $P(\Omega)$.
3. Le fait de choisir $[X \leqslant x]$ dans la définition plutôt que $[X < x]$, $[X \geqslant x]$, $[X > x]$ n'est pas important. Le même choix sera fait plus tard pour la fonction de répartition de X.
4. On ne vous demandera pas : « démontrer que X est une variable aléatoire ».

Définition. Le système complet d'événements associé à une variable aléatoire X est la famille d'événement $([X = x])_{x \in X(\Omega)}$ (qui peut être infinie et même non dénombrable).

2) Fonction de répartition d'une variable aléatoire

On se place dans un espace probabilisé $(\Omega,\ \mathcal{A},\ P)$ quelconque et X une variable aléatoire réelle sur cet espace.

Définition. On appelle fonction de répartition de X la fonction définie sur $\mathbb{R}$ par : $\forall x \in \mathbb{R}, F_X(x) = P(X \leqslant x)$

Théorème 6.« Propriétés de F_X »

1. $\forall x \in \mathbb{R}, 0 \leqslant F_X(x) \leqslant 1$.
2. F_X est croissante sur $\mathbb{R}$.
3. $\lim\limits_{x \to -\infty} F_X(x) = 0$ et $\lim\limits_{x \to +\infty} F_X(t) = 1$.
4. F_X est continue à droite en tout point x de $\mathbb{R}$ c'est à dire $\lim\limits_{t \to x^+} F_X(t) = F_X(x)$.
5. F_X admet un limite à gauche en tout point x de $\mathbb{R}$ et $\lim\limits_{t \to x^-} F_X(t) = F_X(x) - P([X = x])$.
6. Soient $a \leqslant b$ deux réels. $P([a < X \leqslant b]) = F_X(b) - F_X(a)$.

Exemple 8. ✎ On lance une pièce de monnaie bien équilibrée trois fois de suite. Soit X la variable aléatoire donnant le nombre de Face obtenues et soit $Y = \begin{cases} 1 & \text{si deux côtés identiques apparaissent successivement} \\ 0 & \textbf{sinon} \end{cases}$
Dessiner F_X et F_Y.

Remarque . Contrairement à X, qui ne dépend que de Ω et de $\mathcal{A}$, F_X dépend aussi de la probabilité P.

Proposition 7.« continuité de F_X »

F_X est continue sur $\mathbb{R}$ si et seulement si, pour tout réel x, on a $P([X = x]) = 0$.

Remarque . Ce point établit la différence entre une variable aléatoire discrète, dont la fonction de répartition possède toujours des discontinuités (éventuellement une infinité dénombrable) et une variable aléatoire continue, dont la fonction de répartition est continue sur $\mathbb{R}$.

3) Loi d'une variable aléatoire

On se place dans un espace probabilisé $(\Omega, \ \mathcal{A}, \ P)$ quelconque et X et Y deux variables aléatoires réelles sur cet espace.

Définition. On appelle loi de X la donnée de toutes les probabilités $P(X \in A)$, où A est une réunion au plus dénombrable d'intervalles de $\mathbb{R}$.

Remarque . Cette définition, très théorique, ne vous servira pas dans la pratique. On en a déjà vu un cas particulier simple dans le cas fini que nous allons généraliser dans la prochaine partie aux variables aléatoires discrètes.

Proposition 8.« la fonction de répartition caractérise la loi »

$F_X = F_Y \Leftrightarrow X$ et Y ont la même loi.

IV. Variables aléatoires discrètes

On se place dans un espace probabilisé $(\Omega, \ \mathcal{A}, \ P)$ quelconque et X une variable aléatoire discrète sur cette espace, ce qui signifie que $X(\Omega) = \{x_i, i \in I\}$, où I est inclus dans $\mathbb{N}$ ou $\mathbb{Z}$.

1) Lois et fonctions de répartition dans le cas discret

Définition. La loi d'une variable aléatoire discrète X est la donnée de toutes les probabilités $P([X = x_i])$, $i \in I$, ce qui revient à donner les probabilités de chacun des événements du système complet d'événements associé à X.

Remarque . 1. Dans le cas fini, c'est la définition donnée au chapitre 8, qui conduit à une présentation sous forme de tableau, du moins lorsque $X(\Omega)$ est assez petit.

2. Dans le cas infini, la représentation sous forme de tableau est évidemment beaucoup moins pratique ! Mais on peut continuer à penser la loi comme un tableau infini.
3. Cette définition ne s'adapte pas aux variables aléatoires continues. En effet, dans ce cas, toutes les probabilités du SCE associées à X sont nulles (voir la proposition 7).

Comme dans le cas fini, la somme des probabilités doit être égale à 1, mais dans le cas infini cela nécessite l'emploi de séries.

Proposition 9.

$\sum_{i\in I} P([X = x_i])$ converge et sa somme vaut 1. Par exemple, si $X(\Omega) = \mathbb{N}$, on a $\sum_{k=0}^{+\infty} P(X = k) = 1$

Les propriétés suivantes de la fonction de répartition sont valables uniquement dans le cas discret.

Proposition 10.« Propriétés supplémentaires de F_X dans le cas discret

1. Pour tout réel x, on a $F_X(x) = \sum_{x_i\leqslant x,\ i\in I} P([X = x_i])$.
2. Pour tout $i \in I$, on a $P([X = x_i]) = F_X(x_i) - F_X(x_{i-1})$.
3. En particulier, si $X(\Omega) = \mathbb{N}$ ou $\mathbb{Z}$, alors, pour tout entier k, on a $P([X = k]) = F_X(k) - F_X(k - 1)$.

2) Transformation d'une variable aléatoire discrète

Définition. Soit g une fonction définie sur $X(\Omega)$ à valeur dans $\mathbb{R}$. Alors, on définit la variable aléatoire $Y = g(X)$ de la manière suivante : $\forall y \in g(X(\Omega))$, on a $P([Y = y]) = \sum_{x\in X(\Omega), g(x)=y} P([X = x])$. Cette variable aléatoire est également discrète et a pour univers image $Y(\Omega) = g(X(\Omega)) = g(x_i),\ i \in I$.

Exemple 9. ✎

1. Soit X une v.a.r telle que, $P([X = 0]) = P([X = 10]) = \frac{1}{4}$ et $P([X = 20]) = \frac{1}{2}$. Donner les lois de $Y = \frac{X - 10}{10}$ et $Z = Y^2$.
2. Soit $Y = e^X$, où X v.a.r discrète. Soit $y \in \mathbb{R}$. Donner $P([Y = y])$ en fonction d'une probabilité portant sur X.

3) Espérance d'une variable aléatoire discrète

La définition d'espérance doit être adaptée au cas infini à l'aide des séries.

Définition. Une variable aléatoire X admet une espérance si la série $\sum_{i\in I} x_i P(X = x_i)$ converge absolument et dans ce cas, l'espérance est la somme de la série, notée $E(X)$. Par exemple, si $X(\Omega) = \mathbb{N}$, on a $E(X) = \sum_{k=0}^{+\infty} kP(X = k)$

Exemple 10. ✎

1. Soit X le rang d'obtention du premier 1 avec un dé équilibré. Montrer que X admet une espérance et la calculer.

2. Une urne contient initialement deux boules, l'une noire et l'autre blanche. On tire une à une des boules dans cette urne en suivant le protocole suivant :
   * Si la boule tirée est noire, on la remet simplement dans l'urne avant le tirage suivant.
   * Si la boule tirée est blanche, on la remet dans l'urne et l'on y rajoute une boule blanche avant le tirage suivant.

   On note X le rang du premier tirage d'une boule noire. On pourra utiliser les événements : B_i :« la ième boule tirée est blanche »et N_i : « la ième boule tirée est noire »

   (a) Déterminer $X(\Omega)$. (b) Calculer $\sum\limits_{k\in X(\Omega)} P(X=k)$ et interpréter. (c) Calculer, si possible $E(X)$.

Proposition 11.« Propriétés de l'espérance »

Si X et Y admettent une espérance, et si $a,\ b$ sont des réels :
1. Linéarité :$E(aX+b)=aE(X)+b$ et $E(X+Y)=E(X)+E(Y)$.
2. Localisation : si $X(\Omega)\subset[x\,;y]$ alors $x\leqslant E(X)\leqslant y$.
3. Positivité : si toutes les valeurs de X sont positives ($X(\Omega)\subset\mathbb{R}_+$) alors $E(X)\geqslant 0$

On se place dans le cadre des opérations sur les variables aléatoires.

Théorème 12.« Théorème de transfert. Cas particulier de la fonction carrée »

$E(g(X))$ existe ssi la série $\sum\limits_{i\in I} g(x_i)P(X=x_i)$ converge absolument et est alors égale à la somme de cette série. Par exemple, avec la fonction carré et $X(\Omega)=\mathbb{N}$, en cas de convergence : $E(X^2)=\sum\limits_{k=0}^{+\infty} k^2P(X=k)$

4) Moments d'ordre r d'une variable aléatoire discrète

Définition. Pour un entier r, lorsque X^r admet une espérance, on dit que X admet un moment d'ordre r et on note $m_r(X)=E(X^r)$

On montre l'existence des moments et on calcule leurs valeurs à l'aide de la formule de transfert.

Exemple 11. ✎
1. Justifier que toute variable aléatoire admet un moment d'ordre 0 et le calculer. Donner l'autre nom du moment d'ordre 1.
2. Que dire des moments des variables aléatoires finies ?

Proposition 13.« Lien entre les différents moments d'une variable aléatoire discrète »

Si X admet un moment d'ordre r, alors X admet un moment d'ordre i pour tout i inférieur à r.

5) Variance d'une variable aléatoire discrète

Définition. On dit que X admet une variance si $(X-E(X))^2$ admet une espérance. La variance est alors $V(X)=E(X-(E(X))^2$. On dit que X admet un écart-type si elle admet une variance, et alors : $\sigma(X)=\sqrt{V(X)}$

Voici un théorème donnant une technique de calcul de la variance qui peut s'avérer plus simple que la définition.

Théorème 14.

[« Koenig-Huygens »] X admet une variance si et seulement si elle admet un moment d'ordre 2, et $V(X)=m_2(X)-(m_1(X))^2=E(X^2)-(E(X))^2$

Exemple 12. ✎ Soit X le rang d'obtention du premier 1 avec un dé équilibré. Montrer que $V(X)$ existe et la calculer.

Remarque . Une variable aléatoire peut donc admettre une espérance sans admettre de variance, mais pas le contraire, en vertu de la proposition sur les moments d'ordre r.

Proposition 15.« Propriétés de la variance »

1. Pour des réels $a,\ b$ et X admettant une variance : $V(aX+b)=a^2V(X)$.
2. $V(X)\geqslant 0$.
3. $V(X)=0$ si et seulement si X est constante (certaine).