24 Probabilités

I Espace probabilisé

1 Tribu

Comme vu en première année dans le cadre des probabilités finies, on appelle **univers**, noté en général Ω, l'ensemble des **issues** ou **résultats** ou **réalisations** d'une expérience aléatoire.

Commençons par rappeler quelques situations modèles dans le cadre des univers finis : tirage de p boules dans une urne en contenant n, numérotées de 1 à n.

Exemple

E 1 – **Tirages successifs, avec remise** : Dans ce cas, l'ordre est important, et il peut y avoir répétition. On peut choisir comme modèle de résultat un p-uplet (ou p-liste) d'éléments de $[\![1,n]\!]$.
On pose $\Omega = [\![1,n]\!]^p$, et alors $|\Omega| = n^p$.
Il s'agit aussi du nombre d'applications d'un ensemble à p éléments vers un ensemble à n éléments (à chaque tirage, on associe sa boule).

E 2 – **Tirages successifs, sans remise** : Dans ce cas, l'ordre est important, et il ne peut pas y avoir répétition. On peut choisir comme modèle de résultat un p-uplet d'éléments deux à deux distincts (ou p-arrangements) de $[\![1,n]\!]$.
On pose $\Omega = \mathscr{A}_p([\![1,n]\!])$, et alors

$$|\Omega| = A_n^p = n(n-1)\cdots(n-p+1) = \frac{n!}{(n-p)!}$$

(notation hors-programme).
Il s'agit aussi du nombre d'injections d'un ensemble à p éléments vers un ensemble à n éléments (à chaque tirage, on associe sa boule).

E 3 – **Tirage simultané** : Dans ce cas, l'ordre n'est pas important, et il ne peut pas y avoir répétition. On peut choisir comme modèle de résultat une partie à p éléments de $[\![1,n]\!]$ (ou p-combinaison).
On pose $\Omega = \mathscr{P}_p([\![1,n]\!])$, et alors

$$|\Omega| = \binom{n}{p} = \frac{n(n-1)\cdots(n-p+1)}{p!} = \frac{n!}{p!(n-p)!}.$$

Rappelons également qu'une même expérience peut donner lieu à différents univers possible selon ce que l'on souhaite observer (par exemple : carte d'une main vs couleur seulement de la carte, ou bien résultat d'un dé vs parité de ce résultat, etc.)

C'est encore plus vrai pour des univers infinis : le cadre formel que l'on va se donner prévoit que certaines parties de l'univers Ω seulement soient « observables » (les événements), afin de définir une probabilité dans ce cadre plus général.

Remarque

R 1 – Conformément aux habitudes probabilistes, on note, pour $A \subset \Omega$, $\overline{A} = \Omega \setminus A$ le complémentaire dans Ω de A. Cela n'a absolument rien à voir avec l'adhérence topologique, donc.

Définition 1 : Tribu

Soit Ω un ensemble. On appelle **tribu** sur Ω toute partie $\mathscr{A}$ de $\mathscr{P}(\Omega)$ telle que

(i)

(ii) **Stabilité par passage au complémentaire** :

(iii) **Stabilité par réunion dénombrable** :

Le couple $(\Omega, \mathscr{A})$ est appelé **espace probabilisable**, et les éléments de $\mathscr{A}$ ses **événements**.

Exemple

E 4 – $\mathscr{P}(\Omega)$ est une tribu sur Ω (dite **discrète**)

E 5 – $\{\varnothing, \Omega\}$ est une tribu sur Ω (dite **grossière**)

E 6 – Si A est une partie non vide de Ω, distincte de Ω, la plus petite tribu sur Ω contenant A est

Le vocabulaire vu en première année reste valable :

- Si $A \in \mathscr{A}$, $\overline{A}$ est l'**événement contraire** (qui est bien un événement).
- L'événement $\varnothing$ est appelé **événement impossible**.
- On dit que deux événements A et B sont **incompatibles** lorsque $A \cap B = \varnothing$.

Ne pas confondre issue = résultat = réalisation avec événement !

D'après le programme officiel, *la manipulation de tribu n'est pas un objectif du programme* : elles servent de cadre théorique mais, dans la pratique, on n'attend pas nécessairement de les préciser.

Propriété 1 : des tribus

Une tribu est stable par réunion finie, par intersection dénombrable, par intersection finie.

Ainsi dit, si $\mathscr{A}$ est une tribu sur Ω,

(i) $\Omega \in \mathscr{A}$

(ii) Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathscr{A}$,

(iii) Si $(A_n)_{0\leqslant n\leqslant N}$ est une famille finie d'éléments de $\mathscr{A}$,

2 Probabilité

Définition 2 : Probabilité

Soit $(\Omega, \mathscr{A})$ un espace probabilisable. Une **probabilité** sur $(\Omega, \mathscr{A})$ est une application $\mathbb{P}$ définie sur $\mathscr{A}$ telle que

(i) $\forall A \in \mathscr{A} \quad \mathbb{P}(A) \in$

(ii) $\mathbb{P}(\Omega) =$

(iii) **σ-additivité** : Si $(A_n)_{n\in\mathbf{N}}$ est une suite d'événements deux à deux disjoints (incompatibles),

On dit alors que le triplet $(\Omega, \mathscr{A}, \mathbb{P})$ est un **espace probabilisé**.

Propriété 2 : d'une probabilité

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé, A et B des événements.

(i) $\mathbb{P}(\varnothing) =$

(ii) *Si A et B sont deux événements incompatibles,* $\mathbb{P}(A \sqcup B) =$

Plus généralement, $\mathbb{P}\left(\bigsqcup_{n=0}^{N} A_n\right) =$

(iii) *Si $A \subset B$,* $\mathbb{P}(B \setminus A) =$ *Si A et B sont quelconques,* $\mathbb{P}(B \setminus A) =$

(iv) $\mathbb{P}(A \cup B) =$

(v) ***Croissance*** *: si $A \subset B$,*

Remarque

R 2 – Pour trois événements,

$\mathbb{P}(A \cup B \cup C) =$

Et plus généralement, **Formule de Poincaré** (HP) : pour toute famille $(A_1, \dots, A_n) \in \mathscr{P}(\Omega)^n$, qui

peut se montrer par récurrence par exemple, ou, plus simplement, en remarquant que $\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = 1 - \mathbb{P}\left(\bigcap_{i=1}^{n} \overline{A_i}\right) = 1 - \mathbb{E}\left(\mathbb{1}_{\bigcap_{i=1}^{n} \overline{A_i}}\right) = \dots$

Propriété 3 : Probabilité d'une réunion au plus dénombrable

Si $(A_i)_{i\in I}$ est une famille au plus dénombrable d'événements deux à deux incompatibles, alors $(\mathbb{P}(A_i))_{i\in I}$ est sommable et $\mathbb{P}\left(\bigsqcup_{i\in I} A_i\right) = \sum_{i\in I} \mathbb{P}(A_i)$.

Définition 3 : Distribution de probabilités

Soit Ω un ensemble. On appelle **distribution de probabilités** sur Ω toute famille d'éléments de $\mathbb{R}^+$ indexée par Ω et somme (finie) égale à 1.

On appelle **support** d'une telle distribution $(p_\omega)_{\omega\in\Omega}$ l'ensemble $\{\omega \in \Omega,\ p_\omega \neq 0\}$.

Propriété 4 : Support au plus dénombrable

Le support d'une distribution de probabilités est toujours au plus dénombrable.

3 Cas très simple : univers fini

Si Ω est fini, on prend généralement $\mathscr{A} = \mathscr{P}(\Omega)$, et la propriété de σ-additivité est équivalente à la propriété

Si A et B sont deux événements disjoints, alors
$\mathbb{P}(A \sqcup B) = \mathbb{P}(A) + \mathbb{P}(B)$.

Propriété 5 : Probabilité finie associée à une distribution

Si $\Omega = \{\omega_1, \dots, \omega_m\}$, $\mathbb{P}$ est entièrement définie par la donnée d'une distribution de probabilités $(p_{\omega_i})_{1\leqslant i\leqslant m}$ telle que pour tout $i \in [\![1, m]\!]$, $\mathbb{P}(\{\omega_i\}) = p_{\omega_i}$. Et, pour toute partie A de Ω,

$\mathbb{P}(A) =$

Les probabilités des événements élémentaires déterminent donc $\mathbb{P}$.

4 Cas simple : univers dénombrable

Ici, on garde la propriété de σ-additivité, que l'on ne peut plus remplacer par la simple additivité.

Ici encore, il n'y a pas d'obstacle à prendre la tribu « discrète », c'est-à-dire $\mathscr{P}(\Omega)$. On obtient :

Propriété 6 : Probabilité discrète associée à une distribution

Soit Ω un ensemble dénombrable. Pour toute distribution de probabilités $(p_\omega)_{\omega\in\Omega}$, il existe une unique probabilité sur $\mathscr{P}(\Omega)$ telle que

$$\forall \omega \in \Omega, \qquad \mathbb{P}(\{\omega\}) = p_\omega$$

Cette probabilité vérifie

$\forall A \in \mathscr{P}(\Omega), \qquad \mathbb{P}(A) =$

Donc, encore une fois, $\mathbb{P}$ est définie de manière unique par les probabilités des singletons. Pour un univers fini ou dénombrable, les tribus d'évènements n'ont donc pas grand intérêt.

5 Cas moins simple : univers non dénombrable

Dans le cas où l'univers est infini indénombrable c'est plus compliqué : on peut montrer que pour un tirage à pile ou face infini non dénombrable, modélisé par $\{0,1\}^{\mathbb{N}}$ (non dénombrable par argument diagonal de Cantor), la seule valeur possible pour le probabilité d'un événement élémentaire est... 0.

Pourquoi ? Intuitivement, si la probabilité d'obtenir un pile est $p \in]0,1[$, alors la probabilité d'obtenir n piles de suite de va être $p^n \xrightarrow[n\to+\infty]{} 0$... Donc, il est légitime de penser que l'événement « obtenir que des piles » a une probabilité nulle, par exemple.

C'est donc moins simple, on en peut pas se contenter des événements élémentaires, mais complètement hors-programme.

6 Continuités croissante et décroissante

Propriété 7 : Continuité croissante

Soit $(A_n)_{n\in\mathbb{N}}$ une suite croissante (pour l'inclusion) d'événements :

$$\forall\, n \in \mathbb{N},\quad A_n \subset A_{n+1}$$

Alors

Remarque

R 3 – Continuité car $\bigcup_{n=0}^{+\infty} A_n$ est en quelque sorte la « limite » de la suite croissante $(A_n)_{n\in\mathbb{N}}$.

Remarquons aussi que $A_k = \bigcup_{n=0}^{k} A_n$.

Corollaire 1 : Limite d'une probabilité d'une réunion

Soit $(A_n)_{n\in\mathbb{N}}$ une suite quelconque d'événements, alors

Propriété 8 : Continuité décroissante

Soit $(A_n)_{n\in\mathbb{N}}$ une suite décroissante (pour l'inclusion) d'événements :

$$\forall\, n \in \mathbb{N},\quad A_{n+1} \subset A_n$$

Alors

Corollaire 2 : Limite d'une probabilité d'une intersection

Soit $(A_n)_{n\in\mathbb{N}}$ une suite quelconque d'événements, alors

7 Inégalité de Boole

Propriété 9 : Inégalité de Boole

Soit $(A_n)_{n\in\mathbb{N}}$ une suite quelconque d'événements. Alors, dans $[0,+\infty]$,

Remarque

R 4 – Où, si la série à termes positifs $\sum \mathbb{P}(A_n)$ diverge, on lira la formule $\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) \leqslant +\infty$, ce qui ne dit rien.

Si la série converge et a une somme $\geqslant 1$, le résultat ne dit rien non plus.

8 Négligeabilité

a Événements négligeables

Définition 4 : Événement négligeable

On dit qu'un événement A est **négligeable** lorsque $\mathbb{P}(A) = 0$.

Remarque

R 5 – L'événement impossible est négligeable.

Un événement négligeable n'est pas en général impossible.

Exemple

E 7 – Dans le jeu de Pile ou Face infini, l'événement « la pièce donne Pile un nombre fini de fois » est négligeable.

Propriété 10 : Réunion finie ou dénombrable

Une réunion (respectivement intersection) finie ou dénombrable d'événements négligeables est négligeable.

Propriété 11 : Partie d'un événement négligeable

Si A et B sont deux événements tel que $A \subset B$, si B est négligeable, A l'est.

b Événements, propriétés presque sûr·e·s

Définition 5 : Événement presque sûr

Un événement A est **presque sûr**, ou **presque certain**, lorsque $\mathbb{P}(A) = 1$, ce qui équivaut à dire que $\overline{A}$ est négligeable.

Une propriété est dite **presque sûre** lorsque l'ensemble des éléments de Ω qui ont cette propriété est un événement presque sûr.

Propriété 12 : Réunion, intersection au plus dénombrable

Toute réunion (respectivement intersection) finie ou dénombrable d'événements presque sûrs l'est encore.

II Conditionnement et indépendance

Les notions vues en première année se généralise sans problème particulier.

1 Conditionnement

Définition 6

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé, B un événement tel que $\mathbb{P}(B) > 0$. Pour tout événement $A \in \mathscr{A}$, on définit la **probabilité conditionnelle de A sachant B** par

(Se lit en général « probabilité de A sachant B »)

Remarque

R 6 – ⚠ Il n'y a toujours pas d'« événement conditionnel $A|B$ » (élément de $\mathscr{A}$) : ce n'est qu'une notation signifiant qu'on se place en observateur de l'événement A sachant que l'événement B est déjà réalisé.
Mais la notation $\mathbb{P}_B$ peut aussi être trompeuse, car c'est la même que celle de la loi d'une variable aléatoire.

Propriété 13 : Probabilité... conditionnée

$\mathbb{P}_B : A \in \mathscr{A} \mapsto \mathbb{P}(A|B)$ est une probabilité sur $(\Omega, \mathscr{A})$.

... et donc toutes les propriétés des probabilités, toutes les formules qui vont suivre peuvent être appliquées à des probabilités conditionnelles.

Lorsque plusieurs conditions s'enchaînent, il suffit de les intersecter : « $\mathbb{P}(A|B|C)$ » $= \mathbb{P}_C(A|B) = \mathbb{P}(A|B \cap C)$.

2 Probabilités composées

Propriété 14 : Formule des probabilités composées

Soit $n \geqslant 2$, $A_1, \dots, A_n$ des événements de l'espace probabilisé fini $(\Omega, \mathscr{A}, \mathbb{P})$ tels que $\mathbb{P}(A_1 \cap \dots \cap A_{n-1}) > 0$.

Remarque

R 7 – À nouveau, cela correspond à notre intuition : on réalise A_1, puis A_2 sachant que A_1 l'est, puis A_3 sachant que A_1 et A_2 le sont, etc. On se sert donc en général de cette formule lorsque l'on a des événements successifs, **chronologiques**.

3 Probabilités totales

Définition 7 : Système complet et quasi-complet d'événements

Soit (Ω, P) un espace probabilisé, I un ensemble fini ou dénombrable. On dit que la famille $(A_i)_{i \in I}$ d'événements est un **système complet d'évènements** lorsque

On dit que la famille $(A_i)_{i \in I}$ d'événements est un **système quasi-complet d'évènements** lorsque

Remarque

R 8 – Si on impose de plus les A_i non vides, ce qui se fait parfois, $\{A_i\}_{i \in I}$ est une partition de Ω.

R 9 – Un système complet d'événement est quasi-complet, mais la réciproque est fausse.

Propriété 15 : Formule des probabilités totales

Si $(A_i)_{i \in I}$ où I est fini ou dénombrable est un système complet d'événements, alors pour tout événement B,

Si, de plus, pour tout i, $\mathbb{P}(A_i) > 0$,

Si certains événements sont négligeables, alors les $B \cap A_i$ le seront aussi et il suffit de remplacer la somme pour $i \in I$ par la somme pour $i \in J = \{i \in I,\ \mathbb{P}(A_i) > 0\}$.

Remarque

R 10 – La formule des probabilités totales est utile lorsque l'on fait une expérience aléatoire en plusieurs étapes. Elle permet de raisonner par disjonction de cas, suivant le résultat de la première étape.
⚠ ne pas confondre $\mathbb{P}(B \cap A_i)$ et $\mathbb{P}(B \mid A_i)$!

Exercice 1 : CCINP 101

Exercice 2 : CCINP 107

4 Formule de Bayes

Propriété 16 : Formule de Bayes

Si A, B sont des événements non négligeables, alors

Si, de plus, $\overline{A}$ n'est pas négligeable,

Plus généralement, si $(A_i)_{i \in I}$ (I fini ou dénombrable) est un système complet d'événements non négligeables, on a

Remarque

R 11 – Formule permettant de remonter le temps, appelée aussi formule de probabilité des causes.

Exercice 3 : CCINP 105

III Événements indépendants

1 Couple d'événements indépendants

Définition 8 : Indépendance de deux événements

Deux événements A et B d'un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ sont dits **indépendants** lorsque

On note $A \perp\!\!\!\perp B$ lorsque A et B sont indépendants.

Propriété 17 : Caractérisation par probabilités conditionnelles

Deux événements A et B d'un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ tels que $\mathbb{P}(B) > 0$ sont ***indépendants*** *si et seulement si*

Remarque

R 12 – Cela traduit bien notre intuition : que B soit réalisé ou non, la probabilité de A ne change pas.

R 13 – Bien sûr, si $\mathbb{P}(A) > 0$, cela s'écrit aussi $\mathbb{P}(B \mid A) = \mathbb{P}(B)$.

R 14 – ⚠ Ne pas confondre l'indépendance de deux événements et le fait qu'ils soient incompatibles. Ces notions s'excluent en général. En effet, si A et B sont incompatibles, $A \cap B = \varnothing$, donc $\mathbb{P}(A \cap B) = 0$. Si A et B sont de probabilité non nulle, ils ne sont pas indépendants. (Ce qui se comprend car $A \subset \overline{B}$ par exemple).

R 15 – Contrairement à l'incompatibilité qui est une notion ensembliste, l'indépendance est une notion probabiliste : elle dépend de la probabilité dont est muni Ω.

R 16 – Il n'est pas toujours facile de prédire si deux événements sont indépendants.

Naturellement, si deux événements sont indépendants, leurs complémentaires le sont. Plus précisément :

Propriété 18 : Indépendance et complémentaire

Si deux événements A et B d'un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ sont indépendants, alors

- *A et $\overline{B}$ sont indépendants,*
- *$\overline{A}$ et B sont indépendants,*
- *$\overline{A}$ et $\overline{B}$ sont indépendants.*

Remarque

R 17 – Si A, B sont indépendants et A, C aussi, on ne peut rien dire en général de A et $B \cap C$ et de A et $B \cup C$.

Exercice 4

On lance deux dés équilibrés et l'on considère les événements A « le premier dé amène un nombre pair », B « le second dé amène un nombre pair » et C « les deux dés amènent des nombres de même parité ».

Montrer que A, B, C sont deux à deux indépendants mais que A n'est indépendant ni de $B \cap C$, ni de $B \cup C$.

2 Famille d'événements indépendants

Définition 9 : Événements indépendants vs 2 à 2 indépendants

Soit $(A_i)_{i \in I}$ avec I fini ou dénombrable une famille d'événements.

- Les A_i sont dit **deux à deux indépendants** lorsque pour tout $i \neq j$, A_i et A_j sont indépendants, c'est-à-dire que $\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$.
- Les A_i sont dit **indépendants**, lorsque

Remarque

R 18 – C'est une propriété très forte : elle demande de vérifier énormément conditions ! En général, les espaces probabilisés sont construits pour avoir des événements indépendants et on n'a donc pas à le vérifier à la main.

R 19 – L'indépendance est stable par extraction de sous-familles.

Propriété 19 : Indépendants ⇒ 2 à 2 $\perp\!\!\!\perp$

Si les A_i sont indépendants, alors ils sont deux à deux indépendants.
La réciproque est fausse si $n \geqslant 3$.

Remarque

R 20 – Attention c'est l'inverse des nombres premiers / polynômes entre eux :

indépendants globalement ⇒ deux à deux.

R 21 – Si les événements A_i sont deux à deux indépendants et si pour tout i on pose $B_i = A_i$ ou $\overline{A_i}$, alors les B_i sont deux à deux indépendants d'après la propriété vue précédemment. Cela se généralise aux événements indépendants :

Propriété 20 : Passages au complémentaire dans l'indépendance

Si les événements A_i pour $i \in I$ sont indépendants et si pour tout $i \in I$ on pose $B_i = A_i$ ou $\overline{A_i}$, alors les B_i sont indépendants.

Exercice 5 : Indicatrice d'Euler

Soit $\Omega = [\![1, n]\!]$ où n est un entier non premier supérieur ou égal à 2, muni de la probabilité uniforme. Si $d|n$, on note $A_d = \{kd \mid k \in \Omega \text{ et } kd \in \Omega\}$.

1. **Quelle est la probabilité de A_d ?**
2. **Soit P l'ensemble des diviseurs premiers de n.**
   (a) **Démontrer que $(A_p)_{p\in P}$ est une famille d'événements indépendants.**
   (b) **En déduire le cardinal $\varphi(n)$ de l'ensemble A des nombres inférieurs ou égaux à n et premiers avec n (indicatrice d'Euler).**

Exercice 6

Si $A_1, \dots, A_n$ sont indépendants, $1 \leqslant p \leqslant n-1$, Montrer que les événements suivants sont indépendants :

- $\bigcap_{i=1}^{p} A_i$ **et** $\bigcap_{i=p+1}^{n} A_i$,
- $\bigcup_{i=1}^{p} A_i$ **et** $\bigcap_{i=p+1}^{n} A_i$.
- $\bigcup_{i=1}^{p} A_i$ **et** $\bigcup_{i=p+1}^{n} A_i$,

IV VARIABLES ALÉATOIRES DISCRÈTES

On se donne une espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$.

1 Définition

Définition 10 : Variable aléatoire discrète

Soit E un ensemble quelconque. Une application $X : \Omega \to E$ est appelée **variable aléatoire discrète** sur $(\Omega, \mathscr{A}, \mathbb{P})$ lorsqu'elle vérifie

(i)

(ii)

Elle est dite **réelle** lorsque $E \subset \mathbb{R}$.

Remarque

R 22 – La notation $(X = x)$ est un peu déroutante, cela revient par exemple à noter $\pi\mathbb{Z} = \sin^{-1}(\{0\}) = (\sin = 0)$.
Si A est une partie de E, on note $(X \in A)$ l'événement $X^{-1}(A) = \{\omega \in \Omega, X(\omega) \in A\}$.

R 23 – On note aussi, pour une variable aléatoire réelle,

$$(X \leqslant x) = X^{-1}(]-\infty, x]) = \{\omega \in \Omega,\ X(\omega) \leqslant x\}$$

et on introduit de la même façon, $(X < x)$, $(X \geqslant x)$, $(X > x)$.

R 24 – Enfin, si f est une fonction définie sur $X(\Omega)$, on note $f(X)$ la fonction $f \circ X$. Est-ce une variable aléatoire ? Oui. Voir ci-après.

R 25 – La deuxième condition est là pour qu'on puisse calculer la probabilité $\mathbb{P}(X = x)$ pour tout $x \in X(\Omega)$.
Si $x \in E \setminus X(\Omega)$, $(X = x) = \varnothing \in \mathscr{A}$ également.

R 26 – On ne demande pas que E soit fini ou dénombrable, seulement que $X(\Omega)$ le soit : si des valeurs de E ne sont pas atteintes, on peut s'en débarrasser.
On ne demande pas non plus que l'univers Ω soit fini ou dénombrable.

R 27 – Lorsque l'univers est fini ou dénombrable, on choisit $\mathscr{A} = \mathscr{P}(\Omega)$ et toute fonction de Ω dans E est une variable aléatoire discrète.

Exemple : fondamental

E 8 – Si F événement de l'univers Ω, alors

$$\mathbb{1}_F : \left| \begin{array}{ccl} \Omega & \longrightarrow & \{0,1\} \\ \omega & \longmapsto & \begin{cases} 1 & \text{si } \omega \in F \\ 0 & \text{si } \omega \notin F \end{cases} \end{array} \right.$$

est une variable aléatoire.
$\mathbb{P}(\mathbb{1}_F = 1) = \mathbb{P}(F)$ et $\mathbb{P}(\mathbb{1}_F = 0) = \mathbb{P}\left(\overline{F}\right)$.

Propriété 21 : SCE associé à une variable aléatoire

$\left((X = x)\right)_{x\in X(\Omega)}$ *est un système complet d'événements appelé* ***système complet d'événements associé à*** X.

Remarque

R 28 – On peut remplacer $X(\Omega)$ par E, ce qui revient à ajouter des ensembles vides.

Propriété 22 : Les parties de $X(\Omega)$ sont des événements

Soit X une variable aléatoire discrète sur $(\Omega, \mathscr{A}, \mathbb{P})$. Alors pour toute partie A de $X(\Omega)$, $(X \in A) \in \mathscr{A}$.

Propriété 23 : Une fonction d'une v.a.d. est une v.a.d.

Si $X : \Omega \to E$ est une variable aléatoire discrète, si $f : E \to F$ est une fonction (ou application) quelconque, alors $f \circ X$, notée $f(X)$ est une variable aléatoire discrète.

2 Loi

On fixe X une variable aléatoire discrète sur $(\Omega, \mathscr{A}, \mathbb{P})$.

Définition 11 : Loi d'une v.a.d.

L'application

$$\mathbb{P}_X : \left| \begin{array}{ccc} \mathscr{P}(X(\Omega)) & \longrightarrow & \mathbb{R} \\ A & \longmapsto & \mathbb{P}(X \in A) \end{array} \right.$$

est appelée **loi de X**.

Propriété 24 : La loi est une probabilité

$\mathbb{P}_X$ est une probabilité sur l'espace probabilisable $\big(X(\Omega), \mathscr{P}(X(\Omega))\big)$.

Remarque

R 29 – Comme $X(\Omega)$ est au plus dénombrable, il n'est pas choquant de choisir $\mathscr{P}(X(\Omega))$ comme tribu.

Propriété 25 : Expression de la loi de X

Si $A \in \mathscr{P}(X(\Omega))$, $\mathbb{P}_X(A) =$

Corollaire 3

La loi de X est uniquement déterminée par la donnée de $(\mathbb{P}(X = x))_{x \in X(\Omega)}$, famille sommable de réels dans $[0,1]$ de somme 1.

Remarque

R 30 – Ainsi, pour décrire la loi d'une variable aléatoire, on se contente de préciser $X(\Omega)$ et les $\mathbb{P}(X = x)$ pour $x \in X(\Omega)$.
On verra plus loin les lois usuelles à connaître parfaitement.

Notation 1 : ~

Si X et Y suivent la même loi, on note $X \sim Y$.
Si X suit une loi $\mathscr{L}$, on note $X \sim \mathscr{L}$.

Exercice 7 : CCINP 109

Exercice 8 : CCINP 104

Propriété 26 : Loi de $f(X)$

La loi de $Y = f(X)$ est donnée par $\forall\, y \in f(X(\Omega))$,

De la même manière, on obtient par exemple :

Propriété 27 : Loi d'une somme, d'un produit

Si X et Y sont des variables aléatoires,

V Familles de variables aléatoires

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ espace probabilisé.

1 Définition et lois

a Couple de variables aléatoires discrètes

Les notions vues en première année se généralise sans problème particulier.

Définition – Propriété 1

Soit X, Y variables aléatoires discrètes sur Ω à valeurs dans E, E'. L'application

$$Z : \left| \begin{array}{ccc} \Omega & \longrightarrow & E \times E' \\ \omega & \longmapsto & (X(\omega), Y(\omega)) \end{array} \right.$$

est une variable aléatoire discrète appelée **couple** $Z = (X, Y)$.

Remarque

R 31 – $(X,Y)(\Omega) \subset X(\Omega) \times Y(\Omega)$ et il n'y a pas égalité en général.

R 32 – On note indifféremment $((X,Y) = (x,y))$ ou $(X = x) \cap (Y = y)$ ou $(X = x \text{ et } X = y)$ ou $(X = x, Y = y)$ ces événements.

Propriété 28 : SCE associé à un couple

Soit (X,Y) un couple de variables aléatoires discrètes. Alors la famille d'événements $\big((X,Y)=(x,y)\big)_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est un système complet d'événements appelé ***système complet d'événements associé au couple*** *(X,Y).*

Remarque

R 33 – On sait donc que

- Si X et Y sont deux variables aléatoires discrètes définies sur un espace probabilisé $(\Omega,\mathscr{A},P)$, le « couple »

$$(X,Y)\ :\ \omega\longmapsto (X(\omega),Y(\omega))$$

est une variable aléatoire discrète.

- Si $Z\ :\ \Omega\to E$ est une variable aléatoire discrète, si $f\ :\ E\to F$ est une fonction quelconque, alors $f\circ Z$, notée $f(Z)$ est une variable aléatoire discrète.

Et constatons que donc, si X et Y sont des variables aléatoires discrètes réelles définies sur un même univers probabilisé, alors $X+Y$, XY, $\min(X,Y)$, $\max(X,Y)$ sont des variables aléatoires réelles discrètes.
Bien sûr, il y aussi $\Gamma\left(\text{Arctan}\left(1+X^2+Y^2\right)\right)$, mais on n'a cité que quelques exemples fréquemment utiles.

Pour calculer les lois :

$$\mathbb{P}(X+Y=z)=\sum_{x,y\ \mid\ x+y=z}\mathbb{P}\big(X=x,Y=y\big)$$

$$\text{et}\quad \mathbb{P}(XY=z)=\sum_{x,y\ \mid\ xy=z}\mathbb{P}\big(X=x,Y=y\big)$$

b Loi conjointe

Définition 12 : Loi conjointe

Soit (X,Y) un couple de variable aléatoires discrètes. On appelle **loi conjointe** de (X,Y) la loi $\mathbb{P}_{(X,Y)}$ de la variable aléatoire (X,Y).

Remarque

R 34 – Vu la propriété précédente, cette loi est déterminée par $\mathbb{P}(X=x,Y=y)$ pour $(x,y)\in X(\Omega)\times Y(\Omega)$. Lorsque les variables aléatoires sont finies, cette loi peut être représentée dans un tableau à double entrée.

Exemple

E 9 – On lance deux dés, X est la v.a. égale au plus grand des nombres, Y celle du plus petit. On pose $\Omega=[\![1,6]\!]^2$ muni de la probabilité uniforme. Calculer la loi conjointe de (X,Y).

c Lois marginales

Définition 13 : Lois marginales

Si (X,Y) est un couple de variables aléatoires discrètes, les lois de X et de Y sont appelées **première et seconde lois marginales du couple.**

Propriété 29 : Loi conjointe détermine lois marginales

La loi conjointe de (X,Y) détermine les lois marginales de (X,Y) mais la réciproque est fausse.

d Lois conditionnelles

Définition 14 : Loi conditionnelle

Soit (X,Y) un couple de variables aléatoires discrètes. Pour tout $x\in X(\Omega)$ tel que $\mathbb{P}(X=x)\neq 0$, la **loi conditionnelle de Y sachant** $(X=x)$ est la loi de Y pour la probabilité conditionnelle $\mathbb{P}_{(X=x)}$.
Elle est donc déterminée par, pour tout $y\in Y(\Omega)$,

$$\mathbb{P}(Y=y\mid X=x)=\frac{\mathbb{P}(X=x,Y=y)}{\mathbb{P}(X=x)}.$$

Remarque

R 35 – Les lois conditionnelles de Y sachant $(X=x)$ et la loi de X permettent de déterminer la loi conjointe de (X,Y) :

- Soit $\mathbb{P}(X=x)=0$ et alors $\mathbb{P}(X=x,Y=y)\leqslant\mathbb{P}(X=x)=0$ donc $\mathbb{P}(X=x,Y=y)=0$,
- soit $\mathbb{P}(X=x)\neq 0$ et

$$\mathbb{P}(X=x,Y=y)=\mathbb{P}(Y=y\mid X=x)\mathbb{P}(X=x).$$

2 Extension aux n-uplets

Définition 15 : n-uplets de variables aléatoires

Soit $(X_1,\dots,X_n)$ un n-uplet de variables aléatoires discrètes. C'est encore une variable aléatoire discrète appelé **vecteur aléatoire discret** de dimension n.
La **loi conjointe** de $(X_1,\dots,X_n)$ est déterminée par les $\mathbb{P}(X_1=x_1,\dots,X_n=x_n)$ où pour tout i, $x_i\in X_i(\Omega)$.
Les lois de $X_1,\dots,X_n$ sont les **lois marginales** de $(X_1,\dots,X_n)$.

Définition 16 : Loi conditionnelle pour n variables

Si $x_1,\dots,x_{n-1}$ sont fixés, tel que $\mathbb{P}(X_1=x_1,\dots,X_{n-1}=x_{n-1})>0$, la **loi conditionnelle** de X_n sachant $(X_1=x_1,\dots,X_{n-1}=x_{n-1})$ est déterminée par

$$\mathbb{P}(X_n=x_n\mid X_1=x_1,\dots,X_{n-1}=x_{n-1})=\frac{\mathbb{P}(X_1=x_1,\dots,X_n=x_n)}{\mathbb{P}(X_1=x_1,\dots,X_{n-1}=x_{n-1})}$$

pour tout x_n.

Remarque

R 36 – Lorsque l'on a la propriété

$$\mathbb{P}\left(X_{i+1}=x_{i+1} \mid X_1=x_1,\ldots,X_i=x_i\right)=\mathbb{P}(X_{i+1}=x_{i+1} \mid X_i=x_i)$$

(phénomène sans mémoire), on dit que la famille $(X_1,\ldots X_n)$ de variables aléatoires est **markovienne**.

3 Indépendance

a Cas d'un couple de variable

Définition 17 : Indépendance

Soient X, Y deux variables aléatoires discrètes sur l'espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$.

X et Y sont dites **indépendantes** si pour tout $(A,B) \in \mathscr{P}(X(\Omega)) \times \mathscr{P}(Y(\Omega))$, les événements $(X \in A)$ et $(Y \in B)$ sont indépendants, c'est-à-dire

On note parfois $X \perp\!\!\!\perp Y$.

Propriété 30 : Caractérisation par des événements élémentaires

X et Y sont indépendantes si et seulement si pour tout $(x,y) \in X(\Omega) \times Y(\Omega)$, $(X=x)$ et $(Y=y)$ sont indépendants, c'est-à-dire

Remarque

R 37 – Si X et Y sont indépendantes, la donnée des lois marginales de (X,Y) détermine sa loi conjointe.

Propriété 31 : Caractérisation par les lois conditionnelles

Soit (X,Y) couple de variables aléatoires. Il y a équivalence entre

(i) Les variables aléatoires X et Y sont indépendantes.

(ii) Pour tout $y \in Y(\Omega)$ tel que $\mathbb{P}(Y=y)>0$, la loi de X sachant $(Y=y)$ est la même que la loi de X.

(iii) Pour tout $x \in X(\Omega)$ tel que $\mathbb{P}(X=x)>0$, la loi de Y sachant $(X=x)$ est la même que la loi de Y.

Propriété 32 : Fonctions de variables aléatoires indépendantes

Si X, Y sont des variables aléatoires indépendantes, f, g définies sur $X(\Omega)$ et $Y(\Omega)$ respectivement, alors

Exemple

E 10 – Si X et Y sont indépendantes, pour tous $m, n \in \mathbb{N}$, X^m et Y^n le sont.

Remarque

R 38 – En reprenant un calcul précédent, on obtient, si X, Y indépendantes,

$$\mathbb{P}(X+Y=z)=$$

où l'on peut remplacer $X+Y$ (et $x+y$) par n'importe quelle fonction de X et Y.

b Variables aléatoires indépendantes

Définition 18 : Variables aléatoires indépendantes

Des variables aléatoires discrètes $X_1,\ldots,X_n$ sont dites **indépendantes** lorsque pour toutes parties A_1 de $X_1(\Omega)$, ..., A_n de $X_n(\Omega)$, les événements $(X_1 \in A_1)$, ..., $(X_n \in A_n)$ le sont.

Une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires discrètes est une suite de variables aléatoire indépendantes lorsque pour tout $n \in \mathbb{N}$, $X_1,\ldots,X_n$ le sont.

Si, de plus, elles ont même loi, on dit que ce sont des **variables aléatoires indépendantes identiquement distribuées** (vaiid).

Propriété 33 : Caractérisation par des événements élémentaires

$X_1,\ldots,X_n$ sont indépendantes si et seulement si pour tout $(x_1,\ldots,x_n) \in X_1(\Omega)\times\cdots\times X_n(\Omega)$, les événements $(X_1=x_1)$, ..., $(X_n=x_n)$ le sont.

Remarque

R 39 – n expériences aléatoires indépendantes peuvent être modélisées par n variables aléatoires indépendantes. Le résultat de la i^e expérience est noté X_i et

$$\mathbb{P}(X_1=x_1,\ldots,X_n=x_n)=\mathbb{P}(X_1=x_1)\cdots\mathbb{P}(X_n=x_n).$$

R 40 – Comme pour les événements, indépendants $\Rightarrow$ indépendants deux à deux, mais la réciproque est fausse si $n>2$.

Exemple

E 11 – Si X_1, X_2 vaiid finies de loi uniforme $\mathscr{U}(2)$ sur $\{-1,1\}$. $X_3 = X_1 \times X_2$. Montrer que X_1, X_2, X_3 sont deux à deux indépendantes mais pas indépendantes globalement.

Propriété 34 : Fonctions de variables aléatoires indépendantes

Si $(X_n)_{n\in\mathbb{N}}$ est une famille de variables aléatoires indépendantes, pour tout $n \in \mathbb{N}$, f_n définie $X_n(\Omega)$, alors

Propriété 35 : Lemme des coalitions

Soit $n, m \in \mathbb{N}$ tels que $0 < m < n$, $X_1, \dots, X_m, \dots, X_n$ des variables aléatoires discrètes indépendantes sur $(\Omega, \mathscr{A}, \mathbb{P})$, f définie sur $X_1(\Omega) \times \cdots \times X_m(\Omega)$ et g définie sur $X_{m+1}(\Omega) \times \cdots \times X_n(\Omega)$.

Le résultat s'étant à plus de deux coalitions.

Théorème 1

Soit $(\mathscr{L}_n)_{n\in\mathbb{N}}$ une suite de lois de probabilités discrètes.

Il existe un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ et une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires discrètes indépendantes sur $(\Omega, \mathscr{A}, \mathbb{P})$ tels que pour tout $n \in \mathbb{N}$, $X_n \sim \mathscr{L}_n$.

Exemple

E 12 – Un jeu de pile ou face infini se modélise (naturellement) par une suite d'épreuves de Bernoulli indépendantes.

Démonstration : Admis

VI Lois usuelles

1 Loi Uniforme

Définition 19 : Loi uniforme

On dit que qu'une variable aléatoire **finie** X suit une **loi uniforme** lorsque pour tout $x \in X(\Omega)$,

$$\mathbb{P}_X(\{x\}) = \mathbb{P}(X = x) = \frac{1}{n}$$

où $n = |X(\Omega)|$, c'est-à-dire que pour tout $A \subset X(\Omega)$, $\mathbb{P}_X(A) = \frac{|A|}{n}$.

On note alors $X \sim \mathscr{U}(n)$.

Exemple

E 13 – Si on tire un dé à n faces ou si on tire une boule dans une urne qui en contient n (numérotée), alors la variable aléatoire du résultat suit $\mathscr{U}(n)$.

Remarque

R 41 –

R 42 – ⚠ cela ne concerne pas de la probabilité $\mathbb{P}$ initiale : $\mathbb{P}_X$ peut être uniforme sans que $\mathbb{P}$ le soit.
Si, par exemple, on lance un dé à 6 faces truqué tel que l'on obtient 1 ou 6 avec une probabilité $1/4$ et $2, 3, 4$ ou 5 avec probabilité $1/8$, X variable aléatoire $\mathbb{1}_{2\mathbb{N}}$, alors $\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = 1/2$ donc $X \sim \mathscr{U}(2)$ alors que $\mathbb{P}$ n'est pas la probabilité uniforme.

R 43 – On a donc $\mathbb{P}_X = \sum_{x\in X(\Omega)} \frac{\delta_x}{n}$.

2 Loi de Bernoulli

Définition 20 : Loi de Bernoulli

On dit que X suit une **loi de Bernoulli de paramètre** $p \in [0, 1]$ lorsque X est à valeurs dans $E = \{0, 1\}$, $\mathbb{P}(X = 1) = p$ et $\mathbb{P}(X = 0) = q = 1 - p$.

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

Exemple : Situation type

E 14 – Variable aléatoire étudiant le succès (1) d'un événement donné ou son échec (0).

Propriété 36 : Ce sont les fonctions indicatrices

Les variables aléatoires qui suivent une loi de Bernoulli de paramètre p sont exactement les fonctions indicatrices des parties F de Ω telles que $\mathbb{P}(F) = p$.

Remarque

R 44 – Deux variables aléatoires de Bernoulli sont indépendantes si et seulement si $(X = 1)$ et $(Y = 1)$ le sont.

R 45 – Si X prend deux valeurs a et b distinctes, alors $Y = \frac{X - a}{b - a}$ suit une loi de Bernoulli de paramètre $p = \mathbb{P}(X = b)$.
Autrement dit, $X = a + (b - a)Y$ où Y suit une loi de Bernoulli.

3 Loi binomiale

Lors de la répétition de n expériences de Bernoulli indépendantes, la probabilité d'avoir $k \leqslant n$ succès s'écrit $\binom{n}{k} p^k (1-p)^{n-k}$ où p est la probabilité d'un succès.

Si on appelle X la variable aléatoire du nombre de succès, à valeurs dans $[\![0, n]\!]$, alors elle suit la loi donnée par

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Remarquons que l'on peut écrire $X = X_1 + \cdots + X_n$ où X_i est la variable aléatoire de Bernoulli succès à la i^{e} répétition.

Définition 21 : Loi binomiale

On dit que X suit une **loi binomiale de paramètre** (n,p) **où** $p \in [0,1]$ lorsque X est à valeurs dans $[\![0,n]\!]$ et pour tout $k \in [\![0,n]\!]$,

$$\mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} p^k q^{n-k}$$

avec $q = 1-p$. On note alors $X \sim \mathscr{B}(n,p)$.

Exemple : Situation type

E 15 – Nombre de succès dans la répétition de n expériences de Bernoulli indépendantes.

Remarque

R 46 – $\mathbb{P}_X = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \delta_k$.

R 47 – $\mathscr{B}(1,p) = \mathscr{B}(p)$.

R 48 – La formule du binôme redonne (ou se retrouve par) $\sum_{k=0}^{n} \mathbb{P}(X=k) = 1$.

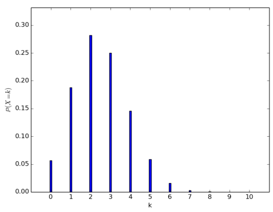

FIGURE 1 – Loi $\mathscr{B}(10, 1/4)$

Exercice 9

Si $X \sim \mathscr{B}(n,p)$ alors $Y = n - X \sim \mathscr{B}(n,q)$.

4 Loi géométrique

Soit $p \in]0,1[$.

On lance une infinité de fois une pièce donnant pile avec probabilité p. Les lancers sont indépendants.

Soit X_n la variable aléatoire du succès au n^{e} lancer : elle vaut 1 si c'est pile, et 0 sinon.

$(X_n)_{n\in\mathbb{N}^*}$ est une suite de vadiid, toutes de loi $\mathscr{B}(p)$.

Soit X la variable aléatoire du rang du premier succès : pour tout $\omega \in \Omega$, $X(\omega) =$

Soit $n \in \mathbb{N}^*$.

- $(X > n) =$

$\mathbb{P}(X > n) =$

- $(X = n) =$

$\mathbb{P}(X = n) =$

- En passant au contraire,

$\mathbb{P}(X \leqslant n) =$

- Soit A l'événement « N'obtenir que des faces », et $A_n = (X > n)$.

Alors (A_n) est décroissante et $A = \bigcap_{n\in\mathbb{N}^*} A_n$.

Par continuité décroissante,

$\mathbb{P}(A_n) =$

Définition 22 : Loi géométrique

Soit $p \in]0,1[$ et X une variable aléatoire discrète. On dit que X suit une **loi géométrique** de paramètre p si X est à valeurs dans et

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

Remarque

R 49 – Première loi dénombrable du programme. On vérifie bien $\sum_{n=1}^{+\infty} \mathbb{P}(X=n) = 1$.

R 50 – De nouveau, on calcule (à savoir faire !)

$\mathbb{P}(X > n) =$

Exemple : Situation type

E 16 – Le rang du premier succès dans une répétition infinie d'épreuves de Bernoulli indépendantes de paramètre p suit $\mathscr{G}(p)$.

5 Loi de Poisson

Définition 23 : Loi de Poisson

Soit $\lambda \in \mathbb{R}_+^*$ et X une variable aléatoire discrète. On dit que X suit une **loi de Poisson de paramètre** λ si X est à valeurs dans et

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

Remarque

R 51 – On vérifie bien $\sum_{n=0}^{+\infty} \mathbb{P}(X=n) = 1$.

6 Propriétés des lois usuelles

a Somme de n vaiid de Bernoulli

Propriété 37 : Importante !

Si $X_1, \dots, X_n$ vaiid de loi $\mathcal{B}(p)$, alors

$$X_1 + \dots + X_n \sim$$

Remarque

R 52 – Plus généralement, si les X_i indépendantes suivent $\mathcal{B}(n_i, p)$, alors $X_1 + \dots + X_n \sim \mathcal{B}(\sum n_i, p)$.

b Approximation d'une loi de Poisson par des lois binomiales

Propriété 38 : Approximation d'une loi de Poisson par des lois binomiales

Soit $\lambda > 0$, $(p_n)_n \in]0,1[^{\mathbb{N}}$ tel que $np_n \to \lambda$, $(X_n)_n$ une suite de variables aléatoires discrètes réelles.

On suppose que pour tout $n \in \mathbb{N}$, $X_n \sim \mathcal{B}(n, p_n)$.

Alors, pour tout $k \in \mathbb{N}$, $\mathbb{P}(X_n = k) \xrightarrow[n \to +\infty]{} \dfrac{\lambda^k}{k!} e^{-\lambda}$.

Remarque

R 53 – Une loi binomiale $\mathcal{B}(n, p)$ (qui peut être vue comme nombre de succès dans la répétition de n épreuve de Bernoulli avec probabilité p de succès) peut donc être approchée par une loi de Poisson $\mathcal{P}(\lambda)$ où $\lambda = np$ à condition que n soit grand et $p = \dfrac{\lambda}{n}$ soit petit.

La loi de Poisson est qualifiée de **loi des événements rares**.

7 Exercices CCINP

Exercice 10 : CCINP 98

Exercice 11 : CCINP 95

VII Espérance

On fixe ici un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

1 Définition

Définition 24 : Espérance d'une variable aléatoire discrète réelle positive

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

L'**espérance de** X est, par définition, dans $[0, +\infty]$,

On a donc $\mathbb{E}(X) \in \mathbb{R} \cup \{+\infty\}$ suivant la sommabilité ou non de la famille réelle positive $(\mathbb{P}(X = x)x)_{x \in X(\Omega)}$.

Lorsque cette famille est sommable, X est dite d'**espérance finie** $\mathbb{E}(X) \in \mathbb{R}$ et on note $X \in L^1$.

Propriété 39 : Cas d'une variable aléatoire entière

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

Définition 25 : Espérance d'une variable aléatoire discrète réelle ou complexe

Soit X une variable aléatoire réelle ou complexe discrète.

Lorsque la famille $(\mathbb{P}(X = x)\ x)_{x \in X(\Omega)}$ est sommable, on dit que X **est d'espérance finie**, on note $X \in L^1$ et on définit son **espérance**

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

Dans le cas contraire, X n'a pas d'espérance (pas plus infinie que finie)

Lorsque X est d'espérance finie et $\mathbb{E}(X) = 0$, X est dite **centrée**.

Remarque

R 54 – Une expression équivalente à « X est d'espérance finie » est « X a un moment d'ordre 1 ».

R 55 – Une variable aléatoire réelle finie est toujours d'espérance finie (programme de MP2I).

R 56 – Une variable aléatoire à valeurs dénombrables $(x_n)_{n \in \mathbb{N}}$ est d'espérance finie si et seulement si la série $\sum_{n \in \mathbb{N}} \mathbb{P}(X = x_n)x_n$ converge **absolument**, et alors $\mathbb{E}(X) = \sum_{n=0}^{+\infty} \mathbb{P}(X = x_n)x_n$.

R 57 – Ne pas confondre « centrée » et « symétrique » : si $P(X = 1) = P(X = 2) = P(X = 3) = P(X = -6) = 1/4$, la variable aléatoire X est bien centrée. Pour autant, X et $-X$ n'ont pas même loi.

2 Théorème de transfert

Théorème 2 : de transfert

Soit X un variable aléatoire discrète, f une fonction définie sur $X(\Omega)$, à valeurs réelles.

$f(X)$ est d'espérance finie si et seulement si la famille $\left(\mathbb{P}(X = x)\ f(x)\right)_{x \in X(\Omega)}$ est sommable, et dans ce cas

Remarque

R 58 – Pas besoin de connaître la loi de $f(X)$!

Corollaire 4 : Espérance du module

X a une espérance finie si et seulement si $|X|$ a une espérance finie.
Le cas échéant,

Corollaire 5 : Sur un univers fini ou dénombrable

Uniquement dans le cas où Ω est fini ou dénombrable, X est d'espérance finie si et seulement si $\left(\mathbb{P}(\{\omega\})X(\omega)\right)_{\omega\in\Omega}$ est sommable et dans ce cas

Remarque

R 59 – Vu en MP2I dans le cas fini.

Exercice 12 : CCINP 97

3 Propriétés de l'espérance

Une espérance peut être vue comme une intégrale, ce qui rend toutes ces propriétés naturelles.

Propriété 40 : de l'espérance

X et Y désignent deux variables aléatoires réelles ou complexes discrètes.

(i) *Si X est constante presque sûrement, c'est-à-dire qu'on a $a\in\mathbb{K}$ tel que $\mathbb{P}(X=a)=1$, alors elle est d'espérance finie $\mathbb{E}(X)=$*

(ii) ***Linéarité** : si $X, Y\in L^1$ et $\lambda\in\mathbb{K}$, $X+\lambda Y\in L^1$ et*

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

(iii) ***Positivité** : si $X\in L^1$ est à valeurs réelles, positive presque sûrement ie $\mathbb{P}(X\geqslant 0)=1$, alors*

***Positivité améliorée** : si X est à valeurs réelles, positive presque sûrement et si*

(iv) ***Croissance** : si $X, Y\in L^1$ sont à valeurs réelles et si $X\leqslant Y$ presque sûrement, alors*

(v) *Si $X\in L^1$, est centrée et appelée **variable aléatoire centrée associée à** X.*

(vi) ***Inégalité triangulaire** : Si $X\in L^1$, $|X|\in L^1$ et*

(vii) *Si $Y\in L^1$ et $|X|\leqslant Y$, alors $X\in L^1$ et*

En particulier, si X est bornée, elle est d'espérance finie.

Corollaire 6 : Espace vectoriel L^1

L'ensemble L^1 des variables aléatoires discrètes sur $(\Omega,\mathscr{A},\mathbb{P})$ admettant une espérance finie est donc un $\mathbb{R}$-espace vectoriel et $X\mapsto\mathbb{E}(X)$ est une forme linéaire sur L^1.

4 Espérances des lois usuelles

Propriété 41 : Espérance des lois usuelles

(i) *Si $X\sim\mathscr{B}(p)$, alors $\mathbb{E}(X)=$*

(ii) *Si $X\sim\mathscr{B}(n,p)$, alors $\mathbb{E}(X)=$*

(iii) *Si $X\sim\mathscr{G}(p)$, alors $\mathbb{E}(X)=$*

(iv) *Si $X\sim\mathscr{P}(\lambda)$, alors $\mathbb{E}(X)=$*

Remarque

R 60 – L'espérance d'une loi uniforme est la moyenne arithmétiques des valeurs (en nombre fini) prises par la variable aléatoire.

Corollaire 7 : Cas d'une fonction indicatrice

Soit A un événement de notre tribu $\mathscr{A}$. Alors $\mathbb{1}_A$ a une espérance finie et $\mathbb{E}(\mathbb{1}_A)=$

Exercice 13 : Formule de Poincaré

En remarquant que $\mathbb{P}\left(\bigcup_{i=1}^n A_i\right)=1-\mathbb{P}\left(\bigcap_{i=1}^n \overline{A_i}\right)=1-\mathbb{E}\left(\mathbb{1}_{\bigcap_{i=1}^n \overline{A_i}}\right)$, prouver la formule de Poincaré.

5 Exercices CCINP

Exercice 14 : CCINP 102

Exercice 15 : CCINP 103

Exercice 16 : CCINP 106

Exercice 17 : CCINP 108

Exercice 18 : CCINP 111

6 Espérance et indépendance

Propriété 42 : Espérance et indépendance

Soit X, $Y\in L^1$ indépendantes. Alors $XY\in L^1$, et

Réciproque fausse en général.
*Plus généralement, si $(X_1,\dots,X_n)$ est une famille de variables aléatoires **indépendantes** d'espérance finie, alors $\prod_{i=1}^n X_i$ l'est et*

VIII Variance, écart-type et covariance

On fixe ici un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$. Les variables aléatoires considérées sont à valeurs réelles.

1 Espace L^2

Sous réserve d'existence, les moments (dénomination hors programme) d'une variable aléatoire sont les $\mathbb{E}\left(|X|^k\right)$ pour $k \in \mathbb{N}^*$. Ce sont des paramètres numériques qui donnent des renseignements sur sa loi. En général, on se limite aux moments d'ordre 1 (espérance) et d'ordre 2 (permet d'obtenir la variance).

Notation 2 : L^2

Soit X une variable aléatoire réelle discrète.

On note $X \in L^2$ lorsque X^2 est d'espérance finie (ce qu'on peut noter $\mathbb{E}(X^2) < +\infty$ car X^2 est à valeurs réelles positives).

Propriété 43 : Inégalité de Cauchy-Schwarz

Si deux variables aléatoires réelles discrètes $X, Y \in L^2$, leur produit $XY \in L^1$, et

avec égalité si et seulement si X et Y sont colinéaires presque sûrement, c'est-à-dire

Corollaire 8

L^2 est un $\mathbb{R}$-espace vectoriel.

Propriété 44 : $L^2 \subset L^1$

Si $X \in L^2$, $X \in L^1$.

Remarque

R 61 – Donc L^2 est un sous-espace de L^1.

2 Variance et écart-type

Définition 26 : Variance, écart-type, variable réduite

Soit $X \in L^2$.

On appelle **variance** de X le nombre

On appelle **écart-type** de X le nombre

Lorsque $\mathbb{V}(X) = 1$, X est dite **réduite**.

Remarque

R 62 – $\mathbb{V}(X)$ est le moment d'ordre 2 de la variable aléatoire centrée associée à X : $X - \mathbb{E}(X)$. Par positivité de l'espérance, $\mathbb{V}(x) \geqslant 0$ donc l'écart-type est bien défini.

R 63 – L'écart-type s'interprète comme une distance euclidienne dans $\mathbb{R}^n$ entre le vecteur dont les coordonnées sont les valeurs prises par X et le vecteur dont toutes les coordonnées valent $\mathbb{E}(X)$. C'est donc un indicateur de dispersion de X autour de sa moyenne $\mathbb{E}(X)$.

R 64 – Ne pas hésiter à noter $m = \mathbb{E}(X)$. Il est plus facile à visualiser $\mathbb{E}(X - m) = \mathbb{E}(X) - m = 0$ que $\mathbb{E}(X - \mathbb{E}(X)) = \mathbb{E}(X) - \mathbb{E}(X) = 0$, par exemple.

R 65 – D'après la formule de transfert, si les valeurs prises par X sont les x_i pour $i \in I$ et $m = \mathbb{E}(X)$,

$$\mathbb{V}(X) = \sum_{i \in I} \mathbb{P}(X = x_i)(x_i - m)^2.$$

R 66 – Plus la variance (et donc l'écart-type) est petit, plus X est concentrée autour de sa moyenne $m = \mathbb{E}(X)$.

Le cas extrême est pour une variable aléatoire constante : $\mathbb{V}(X) = 0$.

Réciproquement, si $\mathbb{V}(X) = 0$, alors

$$\forall x \in X(\Omega), \quad \mathbb{P}(X = x) = 0 \text{ ou } x = m = \mathbb{E}(X).$$

Autrement dit, $\mathbb{P}(X \neq m) = 0$ ou encore $\mathbb{P}(X = m) = 1$: X est constante presque sûrement.

Exercice 19 : CCINP 100

Propriété 45 : de la variance

Soit $X \in L^2$.

(i) ***Formule de Kœnig-Huygens*** *:*

(ii) *Si $a, b \in \mathbb{R}$,*

(iii) *Si $\sigma(X) \neq 0$, est centrée réduite, appelée* ***variable aléatoire centrée réduite associée à X****.*

Remarque

R 67 – La deuxième formule est intuitive au sens où une translation des valeurs de X ne perturbe la distance à la moyenne, et comme cette distance est au carré, une homothétie de rapport a la multiplie par a^2.

3 Covariance

Définition 27 : Covariance

Soit $(X, Y) \in \left(L^2\right)^2$ un couple de variables aléatoires réelles discrètes admettant un moment d'ordre 2.

On appelle **covariance** du couple (X, Y) le nombre

Lorsque $\mathrm{Cov}(X, Y) = 0$, X et Y sont dites **non corrélées**.

Remarque

R 68 – La covariance mesure la corrélation entre les variations de X et de Y dans le sens où elle est positive lorsque X et Y s'écartent de leur moyenne dans le même sens, et négative si c'est dans le sens opposé.

R 69 – Cela ressemble à un produit scalaire et ce n'est pas un hasard ! On vérifie facilement qu'il s'agit d'une forme bilinéaire positive. La variance correspond au carré de la « norme » (et donc l'écart-type à la « norme ».)
Cela permet par exemple d'appliquer l'inégalité de Cauchy-Schwarz (mais sans le cas d'égalité) : $\mathrm{Cov}(X,Y)^2 \leqslant \mathbb{V}(X)\mathbb{V}(Y)$ ie $|\mathrm{Cov}(X,Y)| \leqslant \sigma(X)\sigma(Y)$.

Propriété 46 : de la covariance

Soient X, $Y \in L^2$ deux variables aléatoires réelles discrète admettant un moment d'ordre 2.

(i) Cov *est une forme bilinéaire positive.*

(ii) ***Formule de Kœnig-Huygens*** *:*

(iii) $\mathbb{V}(X+Y) =$

(iv) *Si* $X \perp Y$, $\mathrm{Cov}(X,Y) = 0$ *et la réciproque est fausse.*

Remarque

R 70 – $\mathrm{Cov}(X,X) = 0 \Longrightarrow X$ constante presque sûrement.

R 71 – $\dfrac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)} = \mathrm{Cov}\left(\dfrac{X}{\sigma(X)} \times \dfrac{Y}{\sigma(Y)}\right) \in]-1,1[$ est le cœfficient de corrélation de X et Y.

4 Variance d'une somme de variables aléatoires

Propriété 47 : Variance d'une somme

Soient $X_1, \dots, X_n \in L^2$.

(i) $X_1 + \cdots + X_n \in L^2$ *et*

(ii) *Si* $X_1, \dots, X_n$ *sont décorrélées deux à deux* $(i \neq j \Rightarrow \mathrm{Cov}(X_i, X_j) = 0)$,

En particulier, si $X_1, \dots, X_n$ sont des vaiid,

$$\mathbb{V}(X_1 + \cdots + X_n) = n\mathbb{V}(X_1).$$

5 Cas des lois usuelles

Propriété 48 : Espérance et variance des lois usuelles

(i) *Si* $X \sim \mathscr{B}(p)$,

(ii) *Si* $X \sim \mathscr{B}(n,p)$,

(iii) *Si* $X \sim \mathscr{G}(p)$,

(iv) *Si* $X \sim \mathscr{P}(\lambda)$,

IX Inégalités de Markov et de Bienaymé-Tchebychev, Loi faible des grands nombres

Propriété 49 : Inégalité de Markov

Soit $X \in L^1$ une variable aléatoire discrète admettant une espérance finie. Pour tout $a > 0$,

Remarque

R 72 – Si X est à valeurs positives, on a donc aussi $\mathbb{P}(X \geqslant a) \leqslant \dfrac{\mathbb{E}(X)}{a}$.

R 73 – On a aussi $\mathbb{P}(|X| > a) \leqslant \dfrac{\mathbb{E}(|X|)}{a}$.

Propriété 50 : Inégalité de Bienaymé-Tchebychev

Soit $X \in L^2$ une variable aléatoire réelle discrète admettant un moment d'ordre 2, $m = \mathbb{E}(X)$, $a > 0$.

c'est-à-dire, en notant m l'espérance de X et σ son écart-type,

Remarque

R 74 – On a aussi $\mathbb{P}(|X - m| \geqslant a) \leqslant \dfrac{\sigma^2}{a^2}$.

R 75 – le a^2 est logique pour des raisons d'homogénéité (dimension).

R 76 – On retrouve avec l'inégalité de Bienaymé-Thebychev, le fait que si $\mathbb{V}(X)=0$,

$$\forall \varepsilon > 0,\quad \mathbb{P}(|X-\mathbb{E}(X)| \geqslant \varepsilon) = 0.$$

Donc, $\mathbb{P}(|X-\mathbb{E}(X)|>0)=0$, c'est-à-dire $\mathbb{P}(X=\mathbb{E}(X))=1$.

R 77 – En particulier, $\mathbb{P}\left(|X-\mathbb{E}(X)|<a\right) \geqslant 1-\dfrac{\mathbb{V}(X)}{a^2}$.

Intuitivement, en répétant de nombreuses fois un lancer de pièce équilibrée, la fréquence d'apparition de pile doit se rapprocher de $\frac{1}{2}$.

Le théorème suivant permet de donner un cadre théorique à cette intuition.

Théorème 3 : Loi faible des grands nombres

Soit $(X_n)_{n\geqslant 1} \in \left(L^2\right)^{\mathbb{N}^*}$ une suite de variable aléatoires discrètes réelles deux à deux indépendantes identiquement distribuées (de même loi) sur $(\Omega,\mathscr{A},\mathbb{P})$, admettant un moment d'ordre 2. Soit m l'espérance de X_n et σ son écart-type.

On pose enfin $S_n = X_1+\cdots+X_n$.

Pour tout $\epsilon > 0$,

Remarque

R 78 – Parmi tous les échantillons de valeurs possibles $(X_1,\ldots,X_n)$, ceux dont la moyenne (S_n/n) s'éloigne de l'espérance m sont rares, et cette rareté s'accentue avec la taille de l'échantillon ($n\to+\infty$).

Exercice 20 : CCINP 99

X Fonctions génératrices

Dans cette partie, les variables aléatoires sont à valeurs dans $\mathbb{N}$.

1 Définition

Définition 28 : Fonction génératrice

Soit X variable aléatoire discrète sur $(\Omega,\mathscr{A},\mathbb{P})$ **à valeurs dans** $\mathbb{N}$.

On appelle **fonction génératrice associée à** X la fonction

Propriété 51 : des fonctions génératrices

(*i*) *Le rayon de convergence de la série entière $\sum \mathbb{P}(X=n)t^n$ est et elle converge normalement sur*

(*ii*) *Pour tout $t\in[-1,1]$, $G_X(t)=\mathbb{E}($ $)$.*

(*iii*) *G_X est continue sur $[-1,1]$, de classe $\mathscr{C}^\infty$ sur $]-1,1[$ et pour tout $n\in\mathbb{N}$, $\mathbb{P}(X=n)=$*

Propriété 52 : Caractérisation de la loi

Deux variables aléatoires X, Y à valeurs dans $\mathbb{N}$ ont même loi si et seulement si elles ont même fonction génératrice.

Propriété 53 : Lien avec l'espérance et la variance

(*i*) *$X\in L^1$ (est d'espérance finie) si et seulement si G_X est dérivable est 1 et alors*

(*ii*) *$X\in L^2$ si et seulement si G_X est deux fois dérivable est 1 et alors*

On exprime alors $\mathbb{V}(X)$ à l'aide de $G'_X(1)$ et $G''_X(1)$.

2 Cas des lois usuelles

Le programme demande de savoir calculer la fonction génératrice d'une variable aléatoire de Bernoulli, binomiale, géométrique, de Poisson. Allons-y.

3 Somme des variables aléatoires

Propriété 54 : Fonction génératrice d'une somme

Soient $X_1,\ldots,X_n$ des variables aléatoires discrètes indépendantes à valeurs dans $\mathbb{N}$. Alors

Applications

- On retrouve la fonction génératrice d'une loi binomiale à partir de la somme de n vaiid de loi de Bernoulli.
- Une somme de variables aléatoires de loi de Poisson $\mathscr{P}(\lambda_i)$ est encore de loi de Poisson de paramètre la somme des λ_i.
- Une somme de variables aléatoires de loi $\mathscr{B}(n_i,p)$ indépendantes est de loi $\mathscr{B}(\sum n_i,p)$

4 Exercices CCINP

Exercice 21 : CCINP 96

Exercice 22 : CCINP 110

XI FORMULAIRE

Sous réserve d'existence, sommabilité, d'admission de moments, etc. Voici les principales formules du chapitre.

- **Loi de** X : $\boxed{\mathbb{P}_X : A \mapsto \sum_{x\in A} \mathbb{P}(X=x)}$ déterminée par les $\mathbb{P}(X=x)$ pour $x \in X(\Omega)$, positifs de somme 1.
- **Espérance de** X : $\boxed{\mathbb{E}(X) = \sum_{x\in X(\Omega)} \mathbb{P}(X=x)x}$ et si Ω fini ou dénombrable $\boxed{\mathbb{E}(X) = \sum_{\omega\in\Omega} \mathbb{P}(\{\omega\})X(\omega)}$ et si $X(\Omega) \subset \mathbb{N}$, $\boxed{\mathbb{E}(X) = \sum_{n=1}^{+\infty} \mathbb{P}(X \geqslant n).}$
- **Formule de transfert** :

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

- **Variance de** X :

$$\boxed{\mathbb{V}(X) = \mathbb{E}\left((X-\mathbb{E}(X))^2\right) = \mathbb{E}\left(X^2\right) - \mathbb{E}(X)^2.}$$

- **Covariance de** X **et** Y :

$$\boxed{\mathrm{Cov}(X,Y) = \mathbb{E}\left((X-\mathbb{E}(X))(Y-\mathbb{E}(Y))\right) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}$$

nulle si indépendantes.
- **Variance d'une somme** :

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

- **Continuités croissante et décroissante** :
Si $(A_n)_{n\in\mathbb{N}}$ une suite croissante (pour l'inclusion)

$$\boxed{\mathbb{P}(A_k) \xrightarrow[k\to+\infty]{} \mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right)}$$

Si $(A_n)_{n\in\mathbb{N}}$ une suite décroissante (pour l'inclusion)

$$\boxed{\mathbb{P}(A_k) \xrightarrow[k\to+\infty]{} \mathbb{P}\left(\bigcap_{n=0}^{+\infty} A_n\right)}$$

- **Inégalité de Markov** : Si $a>0$, $\boxed{\mathbb{P}(|X| \geqslant a) \leqslant \frac{\mathbb{E}(|X|)}{a}.}$
- **Inégalité de Bienaymé-Tchebychev** : Si $a>0$, $m = \mathbb{E}(X)$ et $\sigma = \sigma(X) = \sqrt{\mathbb{V}(X)}$. $\boxed{\mathbb{P}(|X-m| \geqslant a) \leqslant \frac{\sigma^2}{a^2}}$
- **Loi faible des grands nombres** : Si $\varepsilon>0$, (X_n) une suite de vaiid L^2 d'espérance m et d'écart-type σ,alors

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

- **Inégalité de Cauchy-Schwarz** : Si $X, Y \in L^2$, alors $XY \in L^1$, et

$$\boxed{\mathbb{E}(XY)^2 \leqslant \mathbb{E}\left(X^2\right)\mathbb{E}\left(Y^2\right)}$$

avec égalité si et seulement si X et Y sont colinéaires presque sûrement.

- **Loi de Bernoulli** $\mathscr{B}(p)$:

$$\boxed{\mathbb{P}(X=1)=p \qquad \mathbb{P}(X=0)=1-p=q \qquad \mathbb{E}(X)=p \qquad \mathbb{V}(X)=pq \qquad G_X(t)=q+pt}$$

- **Loi binomiale** $\mathscr{B}(n,p)$:

$$\boxed{\forall\, k \in [\![0,n]\!],\ \ \mathbb{P}(X=k) = \binom{n}{k} p^k q^{n-k} \qquad \mathbb{E}(X)=np \qquad \mathbb{V}(X)=npq \qquad G_X(t)=(q+pt)^n}$$

- **Loi géométrique** $\mathscr{G}(p)$:

$$\boxed{p \in]0,1[\qquad \forall\, n \in \mathbb{N}^*,\ \ \mathbb{P}(X=n) = pq^{n-1} \qquad \mathbb{E}(X) = \frac{1}{p} \qquad \mathbb{V}(X) = \frac{q}{p^2} \qquad G_X(t) = \frac{pt}{1-qt}}$$

- **Loi de Poisson** $\mathscr{P}(\lambda)$:

$$\boxed{\lambda>0 \qquad \forall\, n \in \mathbb{N},\ \ \mathbb{P}(X=n) = \frac{\lambda^n}{n!}\mathrm{e}^{-\lambda} \qquad \mathbb{E}(X)=\lambda \qquad \mathbb{V}(X)=\lambda \qquad G_X(t)=\mathrm{e}^{\lambda(t-1)}}$$