Table des matières

I Cadre théorique

I.1 Ensembles dénombrables

Définition 1
Soit E. On dit que E est dénombrable ssi il existe une application $\varphi : \mathbb{N} \to E$ bijective. En d'autre termes, on peut écrire $E = \{x_0, x_1, \dots\}$ sans oublier un seul élément.

Théorème 1

1. $\mathbb{N}\backslash\{0\}$ est dénombrable.
2. $\mathbb{Z}$ est dénombrable.
3. $\mathbb{N} \times \mathbb{N}$ et $\mathbb{Z} \times \mathbb{Z}$ sont dénombrables.
4. Si E et F sont dénombrables alors $E \times F$ est dénombrable.

I.2 Espaces probabilisés

Définition 2
Soit Ω un ensemble que l'on appellera univers. Une **tribu** sur Ω est un sous ensemble T de $\mathcal{P}(\Omega)$ (les éléments de T sont des sous ensembles de Ω) qui vérifie les 3 conditions :

1. $\Omega \in T$
2. $\forall A \in T \ A^C = \overline{A} = \Omega \backslash A \in T$.
3. Si $(A_n) \in T^{\mathbb{N}}$ alors $\bigcup\limits_{n\in\mathbb{N}} A_n \in T$.

Les éléments de T (qui sont des ensembles, rappelons le) sont des **événements**. Le couple (Ω, T) est un **espace probabilisable**.

Proposition 1
Soit (Ω, T) un espace probabilisable.

1. $\emptyset \in T$.
2. Si $(A_n)_{n\in\mathbb{N}} \in T^{\mathbb{N}}$ $\bigcap\limits_{n\in\mathbb{N}} A_n \in T$. De plus, $\bigcap\limits_{n\in\mathbb{N}} A_n = \overline{\bigcup\limits_{n\in\mathbb{N}} A_n^c}$

Définition 3
Soit Ω un ensemble et T une tribu sur Ω. Une **probabilité** sur Ω est une fonction $\mathbb{P}$ qui associe à chaque événement A une probabilité $\mathbb{P}(A) \in [0,1]$ avec les contraintes suivantes :

1. $\mathbb{P}(\Omega) = 1$
2. Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements incompatibles deux à deux (ie disjoints deux à deux), alors

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n) : \text{propriété de } \sigma\text{-additivité}$$

En particulier, toute série de la forme précédente doit converger vers un nombre dans $[0,1]$.

Le triplet $(\Omega, T, \mathbb{P})$ est appelé un **espace probabilisé**. Dans la suite du cours, nous utiliserons ces notations.

Définition 4
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 $\bigcup\limits_{n\in\mathbb{N}} A_n = \Omega$.

Définition 5
Soit A un événement.

1. Si $A \neq \emptyset$ et $\mathbb{P}(A) = 0$ on dit que A est **négligeable**.
2. Si $A \neq \Omega$ et $\mathbb{P}(A) = 1$ on dit que A est presque sûr.

I.3 Propriétés des probabilités

Proposition 2 (Définition d'une probabilité)
Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements incompatibles deux à deux, alors

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n) \in [0,1]$$

et toute série de la forme précédente (série de probabilités d'événements deux à deux incompatibles) est une série à termes positifs convergente, de limite (somme de la série) dans $[0,1]$

Proposition 3 (Adaptation de la 1ère année)
Soit $(\Omega, T, \mathbb{P})$ un espace probabilisé. Soient A, B deux événements et $(A_n)_{n\in\mathbb{N}}$ une suite d'événements.

1. $\mathbb{P}(\emptyset) = 0$
2. $\mathbb{P}(A^c) = \mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.
3. Si $A \subset B$ alors $\mathbb{P}(A) \leqslant \mathbb{P}(B)$.
4. $\mathbb{P}(A\cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A\cap B)$ et $\mathbb{P}(A\cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A\cup B)$.
5. $\mathbb{P}\left(\bigcup\limits_{k=0}^{N} A_k\right) \leqslant \sum\limits_{k=0}^{N} \mathbb{P}(A_k)$ pour tout $N \in \mathbb{N}$.

Proposition 4 (Système complet d'événements)
Si $(A_n)_{n\in\mathbb{N}}$ est un système complet d'événements et B est un événement, alors

1. $\mathbb{P}\left(\bigcup\limits_{n\in\mathbb{N}} A_n\right) = 1$
2. $\mathbb{P}(B) = \sum\limits_{n=0}^{+\infty} \mathbb{P}(B\cap A_n)$

On retrouve le cours de première année en prenant un système complet fini (tous les A_n sont vides sauf les quelques premiers).

Théorème 2
Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements.

1. Si $(A_n)_{n\in\mathbb{N}}$ est croissante au sens de l'inclusion ($\forall n \in \mathbb{N}\ A_n \subset A_{n+1}$) alors

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mathbb{P}(A_n)$$

2. Si $(A_n)_{n\in\mathbb{N}}$ est décroissante au sens de l'inclusion ($\forall n \in \mathbb{N}\ A_{n+1} \subset A_n$ alors

$$\mathbb{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mathbb{P}(A_n)$$

Le résultat important est l'existence de ces limites.

Proposition 5 (Sous-additivité)
Soit $(A_n)_{n\in\mathbb{N}} \in T^{\mathbb{N}}$ telle que $\sum\limits_{n\geqslant 0} \mathbb{P}(A_n)$ converge. Alors $\mathbb{P}\left(\bigcup\limits_{n\in\mathbb{N}} A_n\right) \leqslant \sum\limits_{n=0}^{+\infty} \mathbb{P}(A_n)$.

II Calcul de probabilités

II.1 Probabilités conditionnelles

Définition-Proposition 1
Soit B un événement tel que $\mathbb{P}(B) > 0$.

1. Pour un événement A, la probabilité de A sachant B est $\mathbb{P}(A|B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$
2. L'application $\mathbb{P}_B : \left\{\begin{array}{lcl} T & \to & [0,1] \\ A & \mapsto & \mathbb{P}_B(A) = \mathbb{P}(A|B) \end{array}\right.$ est une probabilité. C'est la probabilité conditionnelle sachant B.

Proposition 6 (Formule des probabilités composées)

1. Pour A, B des événements, si $\mathbb{P}(B) > 0$ alors $\mathbb{P}(A\cap B) = \mathbb{P}(A|B)\mathbb{P}(B)$. Rappelons en plus que $\mathbb{P}(A\cup B) = 1 - \mathbb{P}(\overline{A}\cap\overline{B})$
2. Si $A_1, \dots, A_n$ sont des événements tels que $\mathbb{P}\left(\bigcap\limits_{k=1}^{n-1} A_k\right) \neq 0$ alors

$$\mathbb{P}\left(\bigcap_{k=1}^{n} A_k\right) = \mathbb{P}(A_1)\times\mathbb{P}(A_2|A_1)\times\mathbb{P}(A_3|A_1\cap A_2)\times\cdots\times\mathbb{P}\left(A_n|\bigcap_{k=0}^{n-1} A_k\right)$$

Proposition 7 (Probabilité totales)
Il s'agit de traduire les propriétés des probabilités vis-à-vis de l'intersection en termes de probabilités conditionnelles.

1. Soit A un événement ni négligeable ni presque sûr ($\mathbb{P}(A) \in]0,1[$). Alors $A, \overline{A}$ forment un système complet d'événements et pour tout événement B,

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

2. Pour $(A_n)_{n\in\mathbb{N}}$ un système complet d'événements (y compris fini) et B un événement

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

où l'on convient que $\mathbb{P}(B|A_n)\mathbb{P}(A_n)=0$ si $\mathbb{P}(A_n)=0$.

Proposition 8 (Formule de Bayes)
Soient A, B deux événements non négligeables ($\mathbb{P}(A)>0$ et $\mathbb{P}(B)>0$. Alors $\mathbb{P}(B|A)=\frac{\mathbb{P}(B)}{\mathbb{P}(A)}\mathbb{P}(A|B)$.

II.2 Evénements indépendants

Définition 6
Soient A, B deux événements
On dit que A et B sont indépendants ssi $\mathbb{P}(A\cap B)=\mathbb{P}(A)\mathbb{P}(B)$.

Définition 7
Soient $A_1,\dots,A_n$ des événements. On dit qu'ils sont mutuellement indépendants ssi

$$\forall I\subset [\![1,n]\!]\mathbb{P}(\bigcap_{i\in I}A_i)=\prod_{i\in I}\mathbb{P}(A_i)$$

Proposition 9
Si (A, B) sont indépendants, il en est de même de $(A^c, B), (A, B^c), (A^c, B^c)$. On peut généraliser ce résultat à des événements mutuellement indépendants (et mettre des complémentaires ou non où bon nous semble).

III Variables aléatoires

III.1 Lois

Définition 8

1. Une variable aléatoire **discrète** est une fonction $X:\Omega\to\mathbb{R}$ où $X(\Omega)$ (l'ensemble des valeurs de X) est dénombrable ou fini.
2. Si H est un ensemble de valeurs de X (on a donc $H\subset X(\omega)$), on note $(X\in H)$ l'événement "X prend l'une des valeurs dans H", c'est à dire l'ensemble $X^{-1}(H)$ (image réciproque par la fonction X de l'ensemble H).
3. Si x est l'une des valeurs que peut prendre X (ie. $x\in X(\Omega)$), on note $(X=x)$ l'événement $X^{-1}(\{x\})$, c'est à dire l'événement "X prend la valeur x"

Théorème 3
Soit X une variable aléatoire discrète sur Ω. Notons $\{x_n|\ n\in\mathbb{N}\}$ l'ensemble de ses valeurs. Alors $((X=x_n))_{n\in\mathbb{N}}$ est un système complet d'événements.

Définition 9
Soit X une variable aléatoire discrète. La loi de X est l'application

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

Avec les notations du théorème précédent, il s'agit de donner, $\mathbb{P}(X=x_n)$, pour tout $n\in\mathbb{N}$.

III.2 Loi usuelles

Définition 10
On dit qu'une variable aléatoire X suit la loi géométrique de paramètre $p\in]0,1[$ (on note $X\hookrightarrow\mathcal{G}(p)$) ssi $\forall k\in\mathbb{N}\backslash\{0\}$ $\mathbb{P}(X=k)=p(1-p)^{k-1}$.
En particulier, l'ensemble des valeurs de X est $X(\Omega)=\mathbb{N}\backslash\{0\}$

Proposition 10
Soit X une variable aléatoire, $X\hookrightarrow\mathcal{G}(p)$ pour un $p\in]0,1[$. Soient $n,k\in\mathbb{N}\backslash\{0\}$. $\mathbb{P}(X>n+k|X>n)=\mathbb{P}(X>k)$.
On dit que la loi géométrique est sans mémoire.

Définition 11
Soit $\lambda\in]0,+\infty[$. On dit qu'une variable aléatoire X à valeurs dans $\mathbb{N}$ suit la loi de Poisson de paramètre λ (noté $X\hookrightarrow\mathcal{P}(\lambda)$) ssi $\forall k\in\mathbb{N}$ $\mathbb{P}(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}$.

III.3 Loi conjointe, indépendance

Définition 12
Soient X, Y deux variables aléatoires discrètes sur Ω. On note $\{x_n|\ n\in\mathbb{N}\}$ et $\{y_m|\ m\in\mathbb{N}\}$ les valeurs possibles de X et Y respectivement.

1. La **loi conjointe** du couple (X, Y) est la loi décrite par la donnée de $\mathbb{P}(X=x_n, Y=y_m)$ pour toutes les valeurs de n et m.
   Il s'agit de la loi de la variable discrète $Z=(X,Y)$.
2. Les lois marginales de la loi conjointe de (X, Y) sont les lois de X et Y.
3. Pour $n_0\in\mathbb{N}$ fixé tel que $\mathbb{P}(X=x_{n_0})\neq 0$, la loi conditionnelle de Y sachant $(X=x_{n_0})$ est la loi donnée par $\mathbb{P}(Y=y_m|X=x_{n_0})$

Définition 13
Soient X, Y deux variables aléatoires discrètes sur Ω. on dit qu'elles sont indépendantes ssi $\forall x, y \in X(\Omega) \times Y(\Omega) \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$ ie ssi les événements $(X = x)$ et $(Y = y)$ sont deux à deux indépendants pour toutes les valeurs possibles de x et y.

Proposition 11
Soient X, Y deux variables indépendantes et $G \subset X(\Omega), H \subset Y(\Omega)$. Alors $\mathbb{P}((X, Y) \in G \times H) = \mathbb{P}(X \in G) \times \mathbb{P}(Y \in H)$.

Proposition 12
Si X, Y sont des variables aléatoires discrètes indépendantes et si on peut calculer $f(X)$ et $g(Y)$ pour des fonctions f et g alors $f(X)$ et $g(Y)$ sont indépendantes.

Définition 14
Soit $(X_n)_{n\in\mathbb{N}}$ n variables aléatoires discrètes sur Ω. On dit qu'elles sont mutuellement indépendantes ssi pour $k \in \mathbb{N}\backslash\{0\}$ et $i_1, \ldots, i_k \in \mathbb{N}$, $\forall x_1 \in X_{i_1}(\Omega), \ldots, x_k \in X_{i_k}(\Omega) \mathbb{P}(X_{i_1} = x_1$ et $\ldots$ et $X_{i_k} = x_k) = \prod_{j=1}^{k} \mathbb{P}(X_{i_j} = x_i)$.

Autrement dit, on peut calculer toute probabilité d'intersection finie par produit.

IV Fonctions et probabilités

IV.1 Fonction de répartition

Définition 15
Soit X une variable aléatoire discrète. On appelle fonction de répartition de X et on note F_X la fonction

$$F_X : \left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \mathbb{P}(X \leqslant x) \end{array} \right.$$

Proposition 13
Avec les notations de la définition, on a :

1. F_X est croissante sur $\mathbb{R}$.
2. $\lim\limits_{x\to-\infty} F_X(x) = 0$.
3. $\lim\limits_{x\to+\infty} F_X(x) = 1$.

IV.2 Fonction génératrice

Définition 16
Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. La fonction génératrice (ou série génératrice) de X est la fonction

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

G_X est définie au moins sur le segment $[-1, 1]$, $\mathcal{C}^\infty$ sur $]-1, 1[$ et $G_X(1) = 1$.

Théorème 4
Si X et Y sont deux variables aléatoires à valeurs dans $\mathbb{N}$ et **indépendantes**, notons R_X et R_Y les rayons de convergence de G_X et G_Y respectivement. Posons également $r = \min(R_X, R_Y)$

Alors G_{X+Y} est de rayon $R \geqslant r$ et

$$\forall t \in]-r, r[\ G_{X+Y}(t) = G_X(t)G_Y(t)$$

IV.3 Espérance, variance

Définition 17
Soit X une variable aléatoire discrète à valeurs dans $X(\Omega) = \{x_n|\ n \in \mathbb{N}\}$. On dit que X est d'espérance finie ssi $\sum\limits_{n\in\mathbb{N}} x_n\mathbb{P}(X = x_n)$ converge **absolument**.

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

Théorème 5 (Théorème de transfert)
Soit X une variable aléatoire discrète et f une fonction définie sur $X(\Omega) = \{x_n|\ n \in \mathbb{N}\}$ à valeurs réelles. $f(X)$ est d'espérance finie ssi $\sum\limits_{n\geqslant 0} f(x_n)\mathbb{P}(X = x_n)$ est absolument convergente.

Alors $E(f(X)) = \sum\limits_{n=0}^{+\infty} f(x_n)\mathbb{P}(X = x_n)$. Ainsi l'espérance de $f(X)$ est déterminée par la loi de X.

Définition-Proposition 2
Soit X une variable aléatoire discrète. Si X^2 est d'espérance finie alors X aussi. Dans ce cas :

1. on appelle **variance** de X le nombre réel positif $V(X) = E\left((X - E(X))^2\right) = E(X^2) - E(X)^2$.
2. on appelle **écart-type** de X le nombre réel positif $\sigma(X) = \sqrt{V(X)}$.

Si $\sigma(X)=1$, on dit que X est réduite.

Théorème 6
Soit X une variable aléatoire à valeurs dans $\mathbb{N}$ et G_X sa fonction génératrice.

1. X possède une espérance finie ssi G_X est dérivable en 1 et alors $E(X)=G'_X(1)$.
2. X possède une variance finie ssi G_X est deux fois dérivable en 1 et alors

$$V(X)=G''_X(1)+G'_X(1)-\left(G'_X(1)\right)^2.$$

Proposition 14 (Propriétés de l'espérance)
Soient X,Y deux variables aléatoires discrètes sur Ω.

1. Linéarité. Soient $\lambda,\mu\in\mathbb{R}$. $E(\lambda X+\mu Y)=\lambda E(X)+\mu E(Y)$.
2. Positivité : si $X\geqslant 0$ alors $E(X)\geqslant 0$.
3. Croissance. Si $\forall\omega\in\Omega\ X(\omega)\leqslant Y(\omega)$ (que l'on note $X\leqslant Y$) alors $E(X)\leqslant E(Y)$.
4. Si X et Y sont **indépendantes** alors $E(XY)=E(X)E(Y)$.

Proposition 15
Soit X une variable aléatoire réelle et $a,b\in\mathbb{R}$. $V(aX+b)=a^2V(X)$.

IV.4 Covariance

Définition-Proposition 3
Soient X,Y deux variables aléatoires discrètes. Si X et Y admettent un moment d'ordre 2 alors la variable $(X-E(X))(Y-E(Y))$ est d'espérance finie.

Dans ce cas on appelle covariance de X et Y le réel

$$\operatorname{cov}(X,Y)=E((X-E(X))(Y-E(Y)))$$

.

Proposition 16
Dans les conditions de la définition précédente :

1. $\operatorname{cov}(X,Y)=E(XY)-E(X)E(Y)$.
2. **Si** X et Y sont indépendantes **alors** $\operatorname{cov}(X,Y)=0$.
3. la covariance est bilinéaire et symétrique.

Proposition 17
Soit X,Y deux variables aléatoires admettant une variance finie. Alors $X+Y$ est de variance finie et

$$V(X+Y)=V(X)+V(Y)+2\operatorname{cov}(X,Y)$$

Théorème 7 (Cauchy-Schwartz)
On a $|\operatorname{cov}(X,Y)|\leqslant\sqrt{V(X)V(Y)}$

Définition 18
Soient X,Y deux variables aléatoires de variance finie et non nulle. Le coefficient de corrélation de X et Y est

$$\operatorname{cor}(X,Y)=\frac{\operatorname{cov}(X,Y)}{\sqrt{V(X)V(Y)}}=\frac{\operatorname{cov}(X,Y)}{\sigma(X)\sigma(Y)}\in[-1,1]$$

V Etude asymptotique

V.1 Interprétation de la loi de Poisson

Proposition 18
Soit $\lambda>0$.

On considère une suite $(X_n)_{n\geqslant 1}$ de variables aléatoires telles que $X_n\hookrightarrow\mathcal{B}(n,p_n)$ où $p_n\underset{+\infty}{\sim}\frac{\lambda}{n}$.

Pour $k\in\mathbb{N}$ fixé, on a $\lim\limits_{n\to+\infty}\mathbb{P}(X_n=k)=e^{-\lambda}\frac{\lambda^k}{k!}$.

V.2 Loi des grands nombres

Théorème 8 (Inégalité de Markov)
Soit X une variable aléatoire réelle à valeurs positives, d'espérance finie.

$$\forall a>0\ \mathbb{P}(X\geqslant a)\leqslant\frac{E(X)}{a}.$$

Théorème 9 (Inégalité de Bienaymé-Tchebychev)
Soit X une variable aléatoire de variance finie.

$$\forall\varepsilon>0\ \mathbb{P}(|X-E(X)|>\varepsilon)\leqslant\frac{V(X)}{\varepsilon^2}.$$

Théorème 10 (Loi faible des grands nombres)
Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires deux à deux indépendantes et de même loi, admettant un moment d'ordre 2.

On pose, pour $n\geqslant 1$, $S_n=\sum\limits_{k=1}^{n}X_k$ et on note $m=E(X_1)$ l'espérance commune aux X_k.

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

. Pour un $\eta>0$ fixé, la limite est nulle.