PROBABILITE SUR UN ENSEMBLE FINI

Ω est un ensemble fini non vide. On note $\mathcal{P}(\Omega)$ l'ensemble des parties de Ω.

• **Vocabulaire**

1. Ω est l'univers ou univers des possibles.
2. Toute partie A de Ω est appelée événement.
3. Pour tout élément ω de Ω, $\{\omega\}$ est appelé événement élémentaire.
4. L'évènement $A \cup B$ se lit "A ou B", l'évènement $A \cap B$ se lit "A et B".
5. $\emptyset$ est l'événement impossible, et Ω est l'événement certain.
6. Pour A et B de $\mathcal{P}(\Omega)$, si $A \cap B = \emptyset$, ils sont dits événements incompatibles.
7. Pour $A \in \mathcal{P}(\Omega)$, son complémentaire $\overline{A}$ est l'événement contraire.

• **Probabilité**

Une probabilité P est une application $P : \mathcal{P}(\Omega) \to [0,1]$ satisfaisant : $P(\Omega) = 1$ (*normalisation*) et si A et B sont deux événements incompatibles, $P(A \cup B) = P(A) + P(B)$ (*additivité*).

$\to$ Le triplet $(\Omega, \mathcal{P}(\Omega), P)$ est ce qu'on appelle un espace probabilisé.

• **Propriétés**

1. $P(\emptyset) = 0$.
2. $P(\overline{A}) + P(A) = 1$.
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
4. Si $\Omega = \{\omega_1, ..., \omega_n\}$ est de cardinal $n \in \mathbb{N}^*$, alors l'additivité donne :
   - si P est une probabilité, elle est déterminée par la donnée des nombres $p_k := P(\{\omega_k\})$.
   - réciproquement, avec n nombres réels positifs $p_1, ..., p_n$ de somme 1, on définit une probabilité P en posant $P(\{\omega_k\}) = p_k$.

• **Equiprobabilité**

Il y a équiprobabilité quand la probabilité est uniforme, c'est à dire quand $p_1 = p_2 = ... = p_n = \frac{1}{n}$.

Dans ce cas on a $P(A) = \dfrac{\text{Card}(A)}{\text{Card}(\Omega)} = \dfrac{\text{"}\textit{nombre de cas favorables}\text{"}}{\text{"}\textit{nombre de cas possibles}\text{"}}$.

• **Indépendance**

Deux événements A et B sont dits indépendants si $P(A \cap B) = P(A).P(B)$.

• **Probabilité conditionnelle**

Soient A et B deux événements tel que $P(A) \neq 0$. Alors la probabilité de B sachant A est $P(B|A) = \frac{P(A \cap B)}{P(A)}$.

$\to$ P_A définie par $P_A(B) = P(B|A)$ est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$.

$\to$ Si A et B sont deux événements de probabilité non nulle, alors :

$$A \text{ et } B \text{ indépendants} \iff P(B|A) = P(B) \iff P(A|B) = P(A)$$

• **Formules**

Soit $(A_1, ..., A_m)$ une partition de Ω telle que pour tout i, $P(A_i) \neq 0$ et $B \in \mathcal{P}(\Omega)$.

$\to$ Formule des probabilités totales : $P(B) = \displaystyle\sum_{i=1}^{n} P(B|A_i).P(A_i)$.

$\to$ Formule de Bayes[1] : Si $P(B) \neq 0$, alors $P(A_j|B) = \dfrac{P(B|A_j).P(A_j)}{\sum_{i=1}^{n} P(B|A_i).P(A_i)}$.

[1] Il s'agit juste d'écrire $P(A \cap B)$ de deux manières différentes.

VARIABLE ALEATOIRE SUR UN ENSEMBLE FINI

Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé fini, $\mathrm{Card}(\Omega) = n$ et $\Omega = \{\omega_1, ..., \omega_n\}$.

• Variable aléatoire

Une variable aléatoire sur Ω est une application $X : \Omega \to \mathbb{R}$.
→ L'ensemble image, dit aussi univers image, $X(\Omega)$ est un ensemble fini $\{x_1, ..., x_m\}$ avec $m \leq n$.
→ Pour tout $E \subset \mathbb{R}$, $X^{-1}(E)$ est un événement noté $X \in E$.
→ $(X = a)$, $(X < a)$ et $(a \leq X \leq b)$ sont les événements $X^{-1}(\{a\})$, $X^{-1}(]-\infty, a[)$ et $X^{-1}([a, b])$.
→ L'ensemble des v.a. forme une algèbre commutative sur $\mathbb{R}$.

• Fonction de répartition

Soit X une v.a., la fonction de répartition de X est l'application $F : \mathbb{R} \to [0, 1]$ définie[2] pour tout $x \in \mathbb{R}$ par $F(x) = P(X \leq x)$.
→ C'est une fonction continue à droite et croissante sur $\mathbb{R}$.
→ Dans le cas présent, c'est une fonction en escalier telle que pour tout $x < \inf X(\Omega)$, $F(x) = 0$, et pour tout $x \geq \sup X(\Omega)$, $F(x) = 1$.

• Loi de probabilité

La loi de X est l'application $\begin{array}{ccc} \mathcal{J} & \to & [0,1] \\ I & \mapsto & P(X^{-1}(I)) \end{array}$, où $\mathcal{J}$ est l'ensemble des intervalles de $\mathbb{R}$.
→ Dans le cas présent, il suffit juste de connaître les $P(X = x_i)$.
→ La loi de probabilité est une probabilité sur $X(\Omega)$. (En fait c'est une probabilité sur $\mathbb{R}$.)

• Espérance mathématique *(moyenne)*

L'espérance d'une v.a. X est le nombre réel $\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega).P(\{\omega\}) = \sum_{i=1}^{n} X(\omega_i).P(\{\omega_i\})$.
→ L'espérance mathématique est une forme linéaire sur le $\mathbb{R}$-espace vectoriel des v.a.
→ On a également la formule[3] suivante $\mathbb{E}(X) = \sum_{x \in X(\Omega)} x.P(X = x) = \sum_{j=1}^{m} x_j.P(X = x_j)$.
→ Si X est une v.a. constante, i.e. $X(\Omega) = \{a\}$, alors $\mathbb{E}(X) = a$.

• Variance

La Variance d'une v.a. X, est le nombre réel positif $V(X) = \mathbb{E}\Big([X - \mathbb{E}(X)]^2\Big)$.
→ La variance est une forme quadratique positive sur le $\mathbb{R}$-espace vectoriel des v.a., de cône isotrope le sous-espace des v.a. constantes, et satisfaisant : $\forall a \in \mathbb{R}, \forall b \in \mathbb{R}, V(aX+b) = a^2 V(X)$.
→ $\sigma(X) := \sqrt{V(X)}$ est l'écart-type de X.
→ On calcule souvent $V(X)$ par le Théorème de Koenig : $V(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$.

• Moments

Soit X une variable aléatoire, et $k \in \mathbb{N}^*$.
Le moment d'ordre k de X est $\mathbb{E}(X^k)$.
Le moment centré d'ordre k de X est $\mathbb{E}\Big([X - \mathbb{E}(X)]^k\Big)$.
→ L'espérance est le moment d'ordre 1 et la variance est le moment centré d'ordre 2.

---

[2] Il sagit de la définition au programme, mais on rencontre aussi la définition $F(x) = P(X < x)$.
[3] Double avantage : il y a moins de termes à sommer, et il n'y a pas besoin d'expliciter Ω.

PROBABILITES DISCRETES

On veut définir la notion de probabilité sur un univers discret Ω, ainsi que celle de variable aléatoire discrète, i.e. $X(\Omega)$ est discret. Les cas d'un univers fini et des v.a. ayant un nombre fini de valeurs ayant été traités, cette page sera exclusivement consacrée au cas dénombrable.

• **Probabilité sur un univers dénombrable**

Soit $\Omega = \{\omega_n ; n \in \mathbb{N}\}$ un univers dénombrable.

Tout marche de la même façon que dans le cas fini en prenant $(\Omega, \mathcal{P}(\Omega), P)$ comme espace probabilisé. La probabilité P est définie sur les événements élémentaires par $P(\{\omega_n\}) = p_n$, où la suite $(p_n)_{n\in\mathbb{N}}$ est une suite quelconque satisfaisant $p_n \geq 0$ et $\displaystyle\sum_{n=0}^{\infty} p_n = 1$.

→ La probabilité d'un événement $A \in \mathcal{P}(\Omega)$ est le nombre[4] $P(A) = \displaystyle\sum_{a\in A} P(\{a\})$.

• **Remarques**

1- Pour garder la cohérence de la théorie, il faut ajouter une nouvelle condition à la définition d'une probabilité. C'est ce qu'on appelle la σ-additivité :

si $(A_n)_{n\in\mathbb{N}}$ sont des événements deux à deux incompatibles alors $P\Big(\displaystyle\bigcup_{n\in\mathbb{N}} A_n\Big) = \sum_{n=0}^{\infty} P(A_n)$.

2- L'équiprobabilité n'a plus aucun sens.

3- En revanche, tout ce qui concerne l'indépendance, la notion de probabilité conditionnelle, ainsi que les formules des probabilités totales et de Bayes reste valable.

• **Variable aléatoire réelle discrète**

Une variable aléatoire réelle discrète est une application $X : \Omega \to \mathbb{R}$ telle que :

1- Ω est un univers muni d'une probabilité.

2- L'ensemble image $X(\Omega)$ est fini ou dénombrable[5].

3- Pour tout $x \in X(\Omega)$, $X^{-1}(\{x\})$ est un événement noté $(X = x)$.

• **Moments**

Le moment d'ordre k est, s'il y a convergence absolue, $\mathbb{E}(X^k) = \displaystyle\sum_{x\in X(\Omega)} x^k P(X = x)$.

→ Espérance (si la série converge absolument) : $\mathbb{E}(X) = \displaystyle\sum_{x\in X(\Omega)} x\,P(X = x)$.

→ Variance (si la série converge) : $V(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \displaystyle\sum_{x\in X(\Omega)} \Big(x - \mathbb{E}(X)\Big)^2 P(X = x)$.

• **Fonction génératrice**

On appelle fonction génératrice d'une v.a. X à valeurs dans $\mathbb{N}$ la série entière définie au moins pour $s \in [-1, 1]$ par $\Phi_X(s) := \mathbb{E}(s^X) = \displaystyle\sum_{n=0}^{\infty} P(X = n)s^n$.

→ Si cela a un sens, on a les moments en dérivant: $\mathbb{E}(X) = \Phi'_X(1)$, $\mathbb{E}(X^2) = \Phi''_X(1) + \Phi'_X(1)$.

→ Si X et Y sont à valeurs dans $\mathbb{N}$ et indépendantes : $\Phi_{X+Y} = \Phi_X.\Phi_Y$.

[4]La notation $P(A)$ a un sens car il s'agit soit d'une somme finie soit d'une série convergente à termes positifs.

[5]Ce qui est le cas si Ω est fini ou dénombrable.

PROBABILITES : LE CAS GENERAL

• **Tribu ou σ-algèbre**

On appelle tribu $\mathcal{T}$ sur Ω un sous ensemble de $\mathcal{P}(\Omega)$ satisfaisant :

1. Ω et $\emptyset$ sont dans $\mathcal{T}$.
2. Si $A \in \mathcal{T}$, alors son complémentaire $\overline{A} \in \mathcal{T}$.
3. Si $A \in \mathcal{T}$ et $B \in \mathcal{T}$, alors $A \cup B \in \mathcal{T}$.
4. Si $(A_n)_{n\in\mathbb{N}}$ est une collection dénombrable d'éléments de $\mathcal{T}$, alors $\displaystyle\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{T}$.

→ Les éléments de $\mathcal{T}$ sont appelés **événements**.
→ Le couple $(\Omega, \mathcal{T})$ est un espace probabilisable.
→ Par complémentarité, toute intersection finie ou dénombrable d'éléments de $\mathcal{T}$ est dans $\mathcal{T}$.

• **Exemples :**
1. $\mathcal{P}(\Omega)$ et $\{\emptyset, \Omega\}$ sont respectivement la plus grande et la plus petite tribu sur Ω.
2. La tribu Borélienne de $\mathbb{R}$: c'est la plus petite tribu de $\mathbb{R}$ contenant l'ensemble, noté $\mathcal{J}$, des intervalles de $\mathbb{R}$. On la notera $\mathcal{B}$. **(H.P.)**

• **Probabilité**

Une probabilité P est une application $P : \mathcal{T} \to [0, 1]$ satisfaisant

1. $P(\Omega) = 1$ (*normalisation*).
2. Soit $(A_n)_{n\in\mathbb{N}}$ des événements 2 à 2 disjoints : $\displaystyle P\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \sum_{n=0}^{\infty} P(A_n)$ (σ-*additivité*).

→ Le triplet $(\Omega, \mathcal{T}, P)$ est un espace probabilisé.

• **Propriétés :**
1. $P(\emptyset) = 0$.
2. $P(\overline{A}) + P(A) = 1$.
3. Si A et B sont incompatibles alors $P(A \cup B) = P(A) + P(B)$. (*additivité*)
4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. (*formule du crible*)
5. Soit n événements A_1, A_2, ... A_n, la généralisation de la formule du crible – dite également de Poincaré – est

$$P(A_1 \cup A_2 \cup ... \cup A_n) = \sum_{k=1}^{n-1} (-1)^{k+1} \sum_{\substack{\{i_1, ..., i_k\} \subset [1, n] \\ \mathrm{Card}(\{i_1, ..., i_k\}) = k}} P(A_{i_1} \cap ... \cap A_{i_k})$$

6. $\displaystyle P\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) \leq \sum_{n=0}^{\infty} P(A_n)$. (*Inégalité de Boole*)
7. Si $(A_n)_{n\in\mathbb{N}}$ est une suite croissante d'événements $\displaystyle P\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \lim_{n\to\infty} P(A_n)$.

• **Remarques**

1. Si Ω est discret, on peut prendre $\mathcal{T} = \mathcal{P}(\Omega)$ comme tribu.
2. Si Ω n'est pas discret, on ne peut plus caractériser une probabilité quelconque par les événements élémentaires ; on ne prend plus la tribu $\mathcal{P}(\Omega)$.
3. Rien ne change en ce qui concerne l'indépendance, les probabilités conditionnelles ainsi que les formules des probabilités totales et de Bayes.
4. Si $A \in \mathcal{T}$ tel que $A \neq \Omega$ et $P(A) = 1$ (resp. $A \neq \emptyset$ et $P(A) = 0$), on dit que A est un événement presque sûr ou presque certain (resp. A est un événement presque impossible ou négligeable).

VARIABLE ALEATOIRE REELLE A DENSITE

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

• **Variable aléatoire réelle** (v.a.r.)
Une v.a.r. est une application $X : \Omega \to \mathbb{R}$ telle que pour tout intervalle I de $\mathbb{R}$, $X^{-1}(I) \in \mathcal{T}$.
→ La loi de X est la probabilité P_X sur $(\mathbb{R}, \mathcal{B})$ caractérisée par : $P_X(I) = P(X \in I)$ pour tout intervalle I de $\mathbb{R}$.
→ Si X et Y sont deux v.a.r. sur $(\Omega, \mathcal{T}, P)$, alors $X + Y$ et $X.Y$ sont des v.a.r. **(Admis)**

• **Fonction de répartition**
La fonction de répartition d'une v.a.r. X est l'application $F : \mathbb{R} \to \mathbb{R}$ définie pour tout x réel par $F(x) = P(X \leq x)$.
→ La fonction de répartition est une application croissante, continue à droite et ayant une limite à gauche en tout point de $\mathbb{R}$ (c'est une fonction cadlag).
→ La fonction de répartition de la v.a.r. X caractérise complètement la loi de X. Exemples :

$$P(a < X \leq b) = F(b) - F(a) \quad ; \quad P(X < a) = \lim_{x \to a \,;\, x < a} F(x)$$

→ La fonction de répartition de la v.a.r. X vérifie $\lim\limits_{x \to -\infty} F(x) = 0$ et $\lim\limits_{x \to +\infty} F(x) = 1$.
→ x_0 est un point de discontinuité de F si et seulement si $P(X = x_0) \neq 0$.
→ Si F est continue, on dit que X est une v.a.r. continue.

• **Densité**
Une application $f : \mathbb{R} \to \mathbb{R}^+$ est appelée densité si :
1- f est positive 2- f est continue par morceaux sur $\mathbb{R}$ 3- $\displaystyle\int_{-\infty}^{+\infty} f(t)dt = 1$

• **Variable aléatoire à densité**
Une v.a.r. est dite à densité, s'il existe une application de densité $f : \mathbb{R} \to \mathbb{R}^+$ telle que la fonction de répartition F de X est donnée pour tout $x \in \mathbb{R}$ par $F(x) = P(X \leq x) = \displaystyle\int_{-\infty}^{x} f(t)dt$.
→ la v.a.r. X possède une loi de densité f si et seulement si $P(X \in I) = \displaystyle\int_I f(t)dt$ pour tout intervalle I.
→ X est une v.a.r. à densité si et seulement si F est **continue** et C^1 par morceaux sur $\mathbb{R}$. Dans ce cas, la densité f est alors déterminée par $f(x) = F'(x)$ pour tout x où F est dérivable.

• **Moments des v.a.r. à densité**
Le moment d'ordre k de X est, s'il y a convergence absolue, $\mathbb{E}(X^k) = \displaystyle\int_{-\infty}^{+\infty} t^k f(t)dt$.
→ Espérance (si l'intégrale converge absolument) : $\mathbb{E}(X) = \displaystyle\int_{-\infty}^{+\infty} t\, f(t)dt$.
→ Variance (si l'intégrale converge) : $V(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \displaystyle\int_{-\infty}^{+\infty} \Big(t - \mathbb{E}(X)\Big)^2 f(t)dt$.
→ Soit $g : \mathbb{R} \to \mathbb{R}$ continue par morceaux. Si l'intégrale converge absolument, $g(X)$ est une v.a.r. admettant pour espérance

$$\mathbb{E}(g(X)) = \int_{-\infty}^{+\infty} g(t)f(t)dt$$

INDEPENDANCE - VECTEURS ALEATOIRES

Toutes les v.a.r. seront supposées définies sur un même espace probabilisé $(\Omega, \mathcal{T}, P)$.

• **Variables aléatoires indépendantes**

Deux v.a.r. X et Y sont dites indépendantes, si pour tout intervalle I et J de $\mathbb{R}$, les événements $(X \in I)$ et $(Y \in J)$ sont indépendants.

$\rightarrow$ Dans le cas où $X(\Omega)$ et $Y(\Omega)$ sont discrets, il suffit de regarder l'indépendance des événements $(X = x)$ et $(Y = y)$ pour $x \in X(\Omega)$ et $y \in Y(\Omega)$.

• **Indépendance mutuelle**

n v.a.r. X_1, ..., X_n sont dites indépendantes si pour tout I_1, ..., I_n, des intervalles de $\mathbb{R}$, les n événements $(X_1 \in I_1)$, ..., $(X_n \in I_n)$ sont mutuellement indépendants.

$\rightarrow$ Dans le cas où les X_i sont des v.a.r. discrètes, il suffit de regarder l'indépendance mutuelle des événements $(X_i = x_i)$ où les x_i décrivent $X_i(\Omega)$ pour tout i.

$\rightarrow$ Soit X_1, ..., X_n des v.a.r. indépendantes et deux applications continues $f : \mathbb{R}^p \rightarrow \mathbb{R}$ et $g : \mathbb{R}^{n-p} \rightarrow \mathbb{R}$, alors les v.a.r. $f(X_1, ..., X_p)$ et $g(X_{p+1}, ..., X_n)$ sont indépendantes. **(Admis)**

• **Vecteurs aléatoires**

$V = (X_1, ..., X_n)$ est un vecteur aléatoire de $\mathbb{R}^n$ si X_1, ..., X_n, ses n composantes, sont des variables aléatoires réelles.

• **Vecteurs aléatoires discrets**

$V = (X_1, ..., X_n)$ est un vecteur aléatoire discret de $\mathbb{R}^n$ si ses composantes sont des v.a.r. discrètes. La loi de V est la donnée des valeurs $P(V = (x_1, \ldots, x_n))$ où $(x_1, \ldots, x_n)$ décrit les valeurs de V.

• **Vecteurs aléatoires à densité**

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$ une densité, c'est-à-dire le produit d'une fonction continue positive sur $\mathbb{R}^n$ par la fonction indicatrice d'un ensemble "géométriquement simple" (on peut prendre tout $\mathbb{R}^n$) et telle que $\displaystyle\int_{\mathbb{R}^n} f(x)dx_1 \ldots dx_n = 1$.

Un vecteur aléatoire $V = (X_1, \ldots, X_n)$ de $\mathbb{R}^n$ est de densité f si pour tous intervalles I_i de $\mathbb{R}$ on a

$$P(V \in I_1 \times \ldots \times I_n) = P((X_1 \in I_1) \cap \ldots \cap (X_n \in I_n)) = \int_{I_1} \ldots \int_{I_n} f(x)dx.$$

$\rightarrow$ dans ce cas, on a par exemple que X_1 est de densité $x_1 \mapsto \displaystyle\int_{\mathbb{R}^{n-1}} f(x_1, y_2, \ldots, y_n)dy_2 \ldots dy_n$.

Si $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ est le produit d'une fonction continue positive sur $\mathbb{R}^n$ par la fonction indicatrice d'un ensemble "géométriquement simple" et si φf est intégrable sur $\mathbb{R}^n$ alors $\varphi \circ V$ est une v.a.r. dont l'espérance est $\mathbb{E}(\varphi \circ V) = \displaystyle\int_{\mathbb{R}^n} \varphi(x)f(x)dx$.

• **Indépendance de v.a.r. à densité**

les n v.a.r. qui forment les composantes du vecteur aléatoire $V = (X_1, \ldots, X_n)$ de densité f sont indépendantes si et seulement si il existe n densités $f_1, \ldots, f_n$ telles que

$$\forall x_1, \ldots, x_n, \qquad f(x_1, \ldots, x_n) = f_1(x_1) \ldots f_n(x_n).$$

• **Covariance de deux variables aléatoires**

Soit X et Y deux v.a.r., la covariance de X et Y est le réel

$$\mathrm{Cov}(X,Y) := \mathbb{E}\Big([X-\mathbb{E}(X)].[Y-\mathbb{E}(Y)]\Big)$$

On la calcule souvent par la formule de Koenig : $\mathrm{Cov}(X,Y) = \mathbb{E}(X.Y) - \mathbb{E}(X).\mathbb{E}(Y)$.
Dans le cas où X et Y sont des v.a.r. discrètes, la covariance existe toujours. Dans le cas où ce sont des v.a.r. à densité, la covariance existe sous l'hypothèse que X et Y ont une variance finie (et donc une espérance finie). Si le vecteur (X,Y) a pour densité f, on calcule $\mathbb{E}(X.Y)$ en écrivant: $\mathbb{E}(X.Y) = \displaystyle\int_{\mathbb{R}^2} xyf(x,y)dxdy$.
$\rightarrow V(X+Y) = V(X) + V(Y) + 2\,\mathrm{Cov}(X,Y)$.
$\rightarrow |\mathrm{Cov}(X,Y)| \leq \sqrt{V(X)}\sqrt{V(Y)}$ *(Inégalité de Cauchy-Schwartz)*

X et Y indép. $\boxed{\Longrightarrow}$ $\mathrm{Cov}(X,Y) = 0 \iff \mathbb{E}(X.Y) = \mathbb{E}(X).\mathbb{E}(Y) \iff V(X+Y) = V(X)+V(Y)$

• **Espérance et variance d'une somme de v.a.r.**

On a toujours $\mathbb{E}(X_1 + \ldots + X_n) = \mathbb{E}(X_1) + \ldots + \mathbb{E}(X_n)$.
Si les v.a.r. $X_1,\ldots,X_n$ sont mutuellement indépendantes et admettent une variance on a $V(X_1 + \ldots + X_n) = V(X_1) + \ldots + V(X_n)$.

APPROXIMATIONS ET CONVERGENCES

• **Inégalité de Bienaymé-Tchébitchef**

Soit X une v.a.r. d'espérance μ et de variance σ^2 : pour tout $t > 0$, $P(|X - \mu| \geq t) \leq \dfrac{\sigma^2}{t^2}$.

• **Convergence en probabilité**

Une suite $(X_n)_{n\in\mathbb{N}}$ de v.a.r. converge en probabilité vers une v.a.r. X si pour tout $\eta > 0$,

$$\lim_{n\to+\infty} P(|X_n - X| < \eta) = 1$$

$\rightarrow$ Il faut bien sûr que toutes ces v.a.r. soient définies sur un même espace probabilisé $(\Omega, \mathcal{T}, P)$.

• **Loi faible des grands nombres**

Soit $(X_n)_{n\geq 1}$ une suite de v.a.r. définies sur un même espace probabilisé, indépendantes et de même loi, d'espérance μ et de variance σ^2. Alors, la moyenne $\overline{X}_n := \frac{1}{n}(X_1 + ... + X_n)$ converge en probabilité vers μ. (Remarque : $\mathbb{E}(\overline{X}_n) = \mu$ et $V(\overline{X}_n) = \frac{\sigma^2}{n} \to 0$.)

• **Convergence presque sûre**

Une suite $(X_n)_{n\in\mathbb{N}}$ de v.a.r. converge presque sûrement vers une v.a.r. X s'il existe un ensemble N de probabilité nulle tel que $X_n(\omega) \to X(\omega)$ pour tout $\omega \in \Omega \setminus N$.

• **Loi forte des grands nombres**

Soit $(X_n)_{n\geq 1}$ une suite de v.a.r. définies sur un même espace probabilisé, indépendantes et de même loi d'espérance μ. Alors, la moyenne $\overline{X}_n := \frac{1}{n}(X_1 + ... + X_n)$ converge presque sûrement vers μ.

• **Convergence en loi**

Une suite $(X_n)_{n\in\mathbb{N}}$ de v.a.r. converge en loi vers une v.a.r. X si $P(X_n \in I) \to P(X \in I)$ pour tout intervalle I.

• **Théorème de la limite centrale** **(Admis)**

Soit $(X_n)_{n\geq 1}$ une suite de v.a.r. définies sur un même espace probabilisé, indépendantes et de même loi, d'espérance μ et de variance σ^2. Alors, la suite de v.a.r. $Z_n := \dfrac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} = \dfrac{1}{\sqrt{n}} \displaystyle\sum_{i=1}^{n} \dfrac{X_i - \mu}{\sigma}$ converge en loi vers une v.a.r. de loi $\mathcal{N}(0, 1)$. (Remarque : $\mathbb{E}(Z_n) = 0$ et $V(Z_n) = 1$.)

LES LOIS AU PROGRAMME

Il est a noté qu'il n'est jamais nécessaire d'expliciter $(\Omega, \mathcal{T}, P)$, seule son existence est requise.

• **Deux exemples de loi uniforme** $P(A) = \dfrac{\text{"}mesure\ de\ A\text{"}}{\text{"}mesure\ totale\text{"}}$

1. $X(\Omega) = \{x_1, ..., x_n\}$ est fini de cardinal $n \in \mathbb{N}^*$. La loi est uniforme si $P(X = x_i)$ ne dépend pas de i. On a donc $P(X = x_i) = \frac{1}{n}$. $\mathbb{E}(X) = \frac{x_1 + ... + x_n}{n}$ et $V(X) = ...$
2. $X(\Omega) = [a, b]$, $a < b$. La loi est uniforme si $P(\alpha \leq X \leq \beta)$ ne dépend que de la longueur du segment $[a, b] \cap [\alpha, \beta]$. X est une v.a.r. de densité f définie par $f(x) = \frac{1}{b-a}$ si $x \in [a, b]$ et $f(x) = 0$ sinon. $\mathbb{E}(X) = \frac{b-a}{2}$ et $V(X) = \frac{1}{12}(b-a)^2$.

• **La loi de Bernoulli** $\mathcal{B}(p)$ ("Pile ou Face")
$X(\Omega) = \{0, 1\}$ symbolisant $\{$*échec,succès*$\}$. On pose $P(X = 1) = p$ et $P(X = 0) = q = 1 - p$.
$\mathbb{E}(X) = p$ et $V(X) = pq = p(1-p)$.

• **La loi binomiale** $\mathcal{B}(n, p)$ (Tirages *échec-succès* avec remise)
$X(\Omega) = \{0, 1, ..., n\}$. Un élément ω de Ω consiste en n épreuves de Bernoulli, $X(\omega)$ est le nombre de succès : $X = X_1 + ... + X_n$ où X_1, ..., X_n sont n v.a.r. de même loi de Bernoulli $\mathcal{B}(p)$, et mutuellement indépendantes. Sa loi est $P(X = k) = C_n^k p^k (1-p)^{n-k}$. $\mathbb{E}(X) = np$ et $V(X) = npq = np(1-p)$.
→ Si X suit la loi $\mathcal{B}(n, p)$ et Y la loi $\mathcal{B}(m, p)$ sont indépendantes, alors $X+Y$ suit la loi $\mathcal{B}(n+m, p)$.

• **La loi hypergéométrique** $\mathcal{H}(N, p, n)$ (Tirages *échec-succès* sans remise)
$X(\Omega) = \{0, 1, ..., n\}$. Dans une population d'effectif initial N, on tire un échantillon sans remise de taille $n \leq N$. X est le nombre de succès obtenus. La proportion de succès initiale est p et d'échec $q = 1 - p$. Sa loi est $P(X = k) = \frac{C_{Np}^k C_{N(1-p)}^{n-k}}{C_N^n}$. $\mathbb{E}(X) = np$ et $V(X) = npq\frac{N-n}{N-1}$.

• **La loi géométrique** (Comptage avec arrêt au premier succès ; succès non compté)
$X(\Omega) = \mathbb{N}$. Soit $p \in]0, 1]$, sa loi est $P(X = k) = (1-p)^k p$. $\mathbb{E}(X) = \frac{1-p}{p}$ et $V(X) = \frac{1-p}{p^2}$.
→ Si on compte le succès : $Y = X + 1$, $Y(\Omega) = \mathbb{N}^*$, $P(Y = k) = (1-p)^{k-1}p$.

• **La loi de Poisson** $\mathcal{P}(\lambda)$
$X(\Omega) = \mathbb{N}$. Soit $\lambda > 0$ appelé paramètre de la loi de Poisson. Sa loi est $P(X = k) = \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$.
$\mathbb{E}(X) = \lambda$ et $V(X) = \lambda$.
→ Si X et Y sont deux v.a.r. indépendantes de loi $\mathcal{P}(\lambda)$ et $\mathcal{P}(\mu)$, alors $X+Y$ est de loi $\mathcal{P}(\lambda+\mu)$.

• **La loi Exponentielle** $\mathcal{E}(\lambda)$
$X(\Omega) = \mathbb{R}$. Soit $\lambda > 0$ appelé paramètre de la loi. $\mathcal{E}(\lambda)$ est la loi de densité $f(t) = \lambda\mathrm{e}^{-\lambda t}$ si $t \geq 0$, et $f(t) = 0$ sinon. $\mathbb{E}(X) = \frac{1}{\lambda}$ et $V(X) = \frac{1}{\lambda^2}$; $F(x) = 1 - \mathrm{e}^{-\lambda x}$ si $x \geq 0$ et $F(x) = 0$ si $x < 0$.

• **La loi Normale ou de Gauss** $\mathcal{N}(\mu, \sigma)$
$X(\Omega) = \mathbb{R}$. $\mathcal{N}(\mu, \sigma)$ est la loi de densité $f(t) = \frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}}$. $\mathbb{E}(X) = m$ et $V(X) = \sigma^2$.
→ Pour tout a et b réels, si X est de loi $\mathcal{N}(\mu, \sigma)$, $aX + b$ est $\mathcal{N}(a\mu + b, |a|\sigma)$. En particulier, $\frac{X-\mu}{\sigma}$ est de loi $\mathcal{N}(0, 1)$, dite loi normale centrée réduite, dont on note Φ la fonction de répartition.
→ Si X et Y sont deux v.a.r. indépendantes de loi respective $\mathcal{N}(\mu, \sigma)$ et $\mathcal{N}(\mu', \sigma')$, alors la loi de $X + Y$ est $\mathcal{N}(\mu + \mu', \sqrt{\sigma^2 + \sigma'^2})$.