# Résumé de sup : probabilités

## I. Espaces probabilités finis

### 1) Univers, événements

L'ensemble des résultats possibles d'une expérience aléatoire est un ensemble $\Omega$ appelé **univers**. $\Omega$ est l'ensemble des cas possibles ou des éventualités ou des issues. En sup, $\Omega$ est fini.

Si $\Omega$ est un univers fini. Une partie de $\Omega$ est un **événement**. L'ensemble des événements est donc $\mathscr{P}(\Omega)$.
$\Omega$ est l'événement certain, $\varnothing$ est l'événement impossible, un singleton $\{\omega\}$ (où $\omega \in \Omega$) est un événement élémentaire.

### 2) Opérations sur les événements

Si $A$ et $B$ sont deux événements, $C_\Omega A$ est l'événement contraire de $A$, $A \cup B$ est la réunion de $A$ et $B$, $A \cap B$ est l'intersection de $A$ et $B$.

$A$ et $B$ sont incompatibles ssi $A \cap B = \varnothing$. Si $A \subset B$, on dit que $A$ implique $B$.

Un système complet d'événements est une famille $(A_i)_{1 \leqslant i \leqslant n}$ telle que $\forall i \neq j$, $A_i \cap A_j = \varnothing$ et $\displaystyle\bigcup_{1 \leqslant i \leqslant n} A_i = \Omega$.

### 3) Probabilité

Soit $\Omega$ un univers fini. Une **probabilité** sur $\Omega$ est une application $P$ de $\mathscr{P}(\Omega)$ dans $[0, 1]$ telle que
1) $P(\Omega) = 1$
2) pour tous événements $A$ et $B$ tels que $A \cap B = \varnothing$, $P(A \cup B) = P(A) + P(B)$.

Dans ce cas, $(\Omega, P)$ est un espace probabilisé.

### 4) Calculs de probabilités

**Théorème.**
- $P(\varnothing) = 0$.
- $P\left(\overline{A}\right) = 1 - P(A)$.
- Si $A \subset B$, $P(A) \leqslant P(B)$ (croissance d'une probabilité). Dans ce cas, $P(B \setminus A) = P(B) - P(A)$.
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- Si $A_1, \ldots A_n$ sont deux à deux incompatibles, $P(A_1 \cup \ldots \cup A_n) = P(A_1) + \ldots + P(A_n)$

  Si de plus $(A_i)_{1 \leqslant i \leqslant n}$ est un système complet d'événements, alors $P(A_1) + \ldots + P(A_n) = 1$ et pour tout événement $B$, $P(B) = \displaystyle\sum_{i=1}^{n} P(B \cap A_i)$.

**Théorème.** Pour tout $\omega$ de $\Omega$, on pose $p_\omega = P(\{\omega\})$.
- $\displaystyle\sum_{\omega \in \Omega} p_\omega = 1$
- $\forall A \in \mathscr{P}(\Omega)$, $P(A) = \displaystyle\sum_{\omega \in A} p_\omega$.

**Théorème (cas de l'équiprobabilité des cas possibles).**
Si $\forall \omega \in \Omega,\ p_\omega = \dfrac{1}{\text{card}(\Omega)}$, alors $\forall A \in \mathscr{P}(\Omega)$, $P(A) = \dfrac{\text{card}(A)}{\text{card}(\Omega)} = \dfrac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$.

## II. Probabilités conditionnelles

Soit $A$ un événement tel que $P(A) \neq 0$. La probabilité de $B$ sachant $A$ est $P_A(B) = \dfrac{P(B \cap A)}{P(A)}$.

**Théorème.** L'application $\begin{array}{rccc} P_A : & \mathscr{P}(\Omega) & \rightarrow & \mathbb{R} \\ & B & \mapsto & P_A(B) \end{array}$ est une probabilité sur $\Omega$.

**Théorème.** Pour tous $A$ et $B$, $\begin{array}{rl} P(A \cap B) & = P_A(B) \text{ si } P(A) \neq 0 \\ & = P_B(A) \text{ si } P(B) \neq 0 \end{array}$.

**Théorème (formule des probabilités totales).** Soit $(A_i)_{1 \leqslant i \leqslant n}$ un système complet d'événements tels que $\forall i \in [\![1, n]\!]$, $P(A_i) \neq 0$, alors

$$\forall B \in \mathscr{P}(\Omega),\ P(B) = \sum_{i=1}^{n} P(A_i) \times P_{A_i}(B).$$

En particulier, si $P(A) \neq 0$ et $P\left(\overline{A}\right) \neq 0$, $P(B) = P(A) \times P_A(B) + P\left(\overline{A}\right) \times P_{\overline{A}}(B)$.

**Théorème (formule de Bayes (inversion d'une probabilité conditionnelle)).** Soit $(A_i)_{1\leqslant i\leqslant n}$ un système complet d'événements tels que $\forall i\in[\![1,n]\!]$, $P(A_i)\neq 0$, alors pour tout $B$ tel que $P(B)\neq 0$,

$$\forall i\in[\![1,n]\!],\ P_B(A_i)=\frac{P(A_i)\times P_{A_i}(B)}{\displaystyle\sum_{i=1}^{n}P(A_i)\times P_{A_i}(B)}.$$

En particulier, $P_B(A)=\dfrac{P(A)\times P_A(B)}{P(A)\times P_A(B)+P\left(\overline{A}\right)\times P_{\overline{A}}(B)}$.

# III. Indépendance

$A$ et $B$ sont indépendants si et seulement si $P(A\cap B)=P(A)\times P(B)$. Si $P(A)\neq 0$, il revient au même de dire $P_A(B)=P(B)$.

**Théorème.** Si $A$ et $B$ sont indépendants, alors $A$ et $\overline{B}$, $\overline{A}$ et $B$, $\overline{A}$ et $\overline{B}$ sont indépendants.

Soient $A_1$, ..., $A_n$, $n$ événements.

$A_1$, ..., $A_n$ sont **deux à deux indépendants** $\Leftrightarrow \forall i\neq j,\ P(A_i\cap A_j)=P(A_i)\times P(A_j)$.

$A_1$, ..., $A_n$ sont **indépendants** $\Leftrightarrow \forall I\subset[\![1,n]\!],\ P\left(\bigcap_{i\in I}A_i\right)=\prod_{i\in I}P(A_i)$.

**Théorème.** indépendants $\underset{\not\Leftarrow}{\Rightarrow}$ deux à deux indépendants.

# IV. Variables aléatoires sur un univers fini

## 1) Variables aléatoires. Loi d'une variable aléatoire

Soit $\Omega$ un univers fini. Une variable aléatoire associée à cet univers est une application $X$ de $\Omega$ dans un certain ensemble $E$. Si $E=\mathbb{R}$, $X$ est une variable aléatoire réelle.

**Variable indicatrice.** Soit $A$ un événement. La variable $X:\begin{array}{ccc}\Omega&\to&\mathbb{R}\\ \omega&\mapsto&\left\{\begin{array}{l}1\text{ si }\omega\in A\\ 0\text{ si }\omega\notin A\end{array}\right.\end{array}$ est la variable indicatrice de l'événement $A$. On peut la noter $1_A$. Elle est utilisée dans une démonstration de l'inégalité de Bienaymé-Tchebychev.

**Quelques notations.**

• Si $X$ est une variable aléatoire sur $\Omega$ et $f$ une application définie sur $X(\Omega)$, on peut définir $f\circ X$ (souvent notée $f(X)$). Par exemple, $X^2$, $\sqrt{X}$, $e^X$ ...

• Si $A$ est une partie de $E$ ($E=\mathbb{R}$ en général), l'événement $\{X\in A\}$ est l'événement $X^{-1}(A)=\{\omega\in\Omega/\ X(\omega)\in A\}$.
Si $x$ est un élément de $E$, l'événement $\{X=x\}$ est l'événement $X^{-1}(\{x\})=\{\omega\in\Omega/\ X(\omega)=x\}$.
Si $X$ est une variable réelle, $\{X\leqslant x\}=X^{-1}(]-\infty,x])=\{\omega\in\Omega/\ X(\omega)\leqslant x\}$.

**Loi de probabilité d'une variable aléatoire.** Soit $X$ une variable aléatoire sur un espace probabilisé fini $(\Omega,P)$.
L'application $\begin{array}{ccc}X(\Omega)&\to&[0,1]\\ x&\mapsto&P(X=x)\end{array}$ est une probabilité sur $X(\Omega)$ appelée loi de $X$. La loi de $X$ peut aussi être l'application plus générale $\begin{array}{ccc}\mathscr{P}(X(\Omega))&\to&[0,1]\\ A&\mapsto&P(X\in A)\end{array}$. On note $P_X$ la loi de $X$.

**Théorème.** $\displaystyle\sum_{x\in X(\Omega)}P(X=x)=1$. Pour toute partie $A$ de $X(\Omega)$, $P(X\in A)=\displaystyle\sum_{x\in A}P(X=x)$.

**Théorème (loi de $f(X)$).** La loi de $f(X)$ est :

$$\forall y\in f(X(\Omega)),\ P(f(X)=y)=\sum_{x\in f^{-1}(\{y\})}P(X=x).$$

Par exemple, si $Y=X^2$, $P(Y=1)=P\left(X^2=1\right)=P(X=1)+P(X=-1)$.

## 2) Espérance, variance, écart-type

**a) Espérance** Si $X$ prend les valeurs $x_1$, ..., $x_n$, l'espérance de $X$ est

$$E(X)=\sum_{k=1}^{n}x_kP(X=x_k)=\sum_{x\in X(\Omega)}xP(X=x).$$

L'espérance de la variable indicatrice $1_A$ d'un événement $A$ est $P(A)$.

**Théorème (linéarité).** L'espérance est une forme linéaire c'est-à-dire $E(X+Y)=E(X)+E(Y)$ et $E(\lambda X)=\lambda E(X)$.
En particulier, $E(aX+b)=aE(X)+b$.

Si $X$ est d'espérance nulle, $X$ est centrée. Si $X$ est une variable réelle quelconque, $X-E(X)$ est la variable centrée associé à $X$.

**Théorème (positivité, croissance).** Si $X$ est une variable aléatoire réelle positive, alors $E(X)\geqslant 0$.
Si $X$ est $Y$ sont des variables aléatoires telles que $X\leqslant Y$, alors $E(X)\leqslant E(Y)$.

**Théorème (inégalité de Markov).** Si $X$ est une variable réelle positive,

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

**Démonstration.** Soit $a>0$. Soit $A=\{X\geqslant a\}$. Soit $\omega\in\Omega$.

- Si $\omega\in A$, $1_A(\omega)=1$ et $\dfrac{X}{a}(\omega)=\dfrac{X(\omega)}{a}\geqslant\dfrac{a}{a}=1=1_A(\omega)$.
- Si $\omega\notin A$, $1_A(\omega)=0\leqslant\dfrac{X(\omega)}{a}$.

Donc, $\forall\omega\in\Omega$, $1_A(\omega)\leqslant\dfrac{X}{a}(\omega)$ ou encore $1_A\leqslant\dfrac{X}{a}$. Par croissance de l'espérance, $E\left(1_A\right)\leqslant E\left(\dfrac{X}{a}\right)=\dfrac{E(X)}{a}$ avec $E\left(1_A\right)=P(A)=E(X\geqslant a)$.

**Théorème de transfert.** L'espérance de $f(X)$ est $E(f(X))=\displaystyle\sum_{x\in X(\Omega)}f(x)P(X=x)$.

### b) Variance, écart-type.

**Définition.** Le moment d'ordre $k$ de $X$ est $E\left(X^k\right)=\displaystyle\sum_{x\in X(\Omega)}x^kP(X=x)$.

**Définition.** La variance de $X$ est $V(X)=E\left((X-E(X))^2\right)=\displaystyle\sum_{x\in X(\Omega)}P(X=x)\times(x-E(X))^2$.

**Théorème (formule de Koenig-Huygens).** $V(X)=E\left((X-E(X))^2\right)=E\left(X^2\right)-(E(X))^2$.

**Théorème.** $V(aX+b)=a^2V(X)$.

**Définition.** L'écart-type de $X$ est $\sigma(X)=\sqrt{V(X)}$. Une variable $X$ telle que $E(X)=0$ et $\sigma(X)=1$ est dite centrée réduite.
Si $X$ est une variable d'écart-type non nul, $\dfrac{X-E(X)}{\sigma(X)}$ est centrée réduite et est la variable centré réduite associée à $X$.

**Théorème (inégalité de Bienaymé-Tchebychev).**

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

**Démonstration.** On applique l'inégalité de Markov à la variable $\left(\dfrac{X-E(X)}{\varepsilon}\right)^2$. L'événement $\{|X-E(X)|\geqslant\varepsilon\}$ est l'événement $\left\{\left(\dfrac{X-E(X)}{\varepsilon}\right)^2\geqslant 1\right\}$. Puisque la variable $\left(\dfrac{X-E(X)}{\varepsilon}\right)^2$ est positive et que $1>0$,

$$P\{|X-E(X)|\geqslant\varepsilon\}\leqslant\frac{1}{1}E\left(\left(\frac{X-E(X)}{\varepsilon}\right)^2\right)=\frac{1}{\varepsilon^2}E\left((X-E(X))^2\right)=\frac{V(X)}{\varepsilon^2}.$$

# V. Couples de variables aléatoires, $n$-uplets de variables aléatoires

## 1) Couples, $n$-uplets

**Définition.** Soient $\Omega$ un univers fini et $X$ et $Y$ deux variables aléatoires sur $\Omega$ à valeurs dans $E$ et $E'$ respectivement.
L'application $\begin{array}{rccc}(X,Y): & \Omega & \to & E\times E'\\ & \omega & \mapsto & (X(\omega),Y(\omega))\end{array}$ est un **couple** de variables aléatoires sur $\Omega$. Si $E=E'=\mathbb{R}$, $(X,Y)$ est une couple de variables aléatoires réelles sur $\Omega$.
Plus généralement, un $n$-uplet de variables aléatoires réelles sur $\Omega$ est $\begin{array}{rccc}(X_1,\dots,X_n): & \Omega & \to & \mathbb{R}^n\\ & \omega & \mapsto & (X_1(\omega),\dots,X_n(\omega))\end{array}$

**Définition.** Si $X$ et $Y$ sont deux variables aléatoires sur l'espace probabilisé fini $(\Omega, P)$, alors la **loi conjointe** de $X$ et $Y$ est la loi du couple $(X, Y)$. Donner la loi conjointe du couple $(X, Y)$, c'est donner les $P((X, Y) = (x, y)) = P(\{X = x\} \cap \{Y = y\})$, $x \in X(\Omega)$, $y \in Y(\Omega)$. Les **lois marginales** (car on les retrouve en marge) du couple $(X, Y)$ sont les lois de $X$ et de $Y$.

**Théorème.** La loi conjointe détermine les lois marginales :

$$\forall x \in X(\Omega),\ P(X = x) = \sum_{y \in Y(\Omega)} P(\{X = x\} \cap \{Y = y\}).$$

$$\forall y \in Y(\Omega),\ P(Y = y) = \sum_{x \in X(\Omega)} P(\{X = x\} \cap \{Y = y\}).$$

Par exemple, si la loi du couple $(X, Y)$ est

| X \ Y | c | d |
|---|---|---|
| a | $\frac{1}{12}$ | $\frac{1}{6}$ |
| b | $\frac{1}{8}$ | $\frac{5}{8}$ |

la première loi marginale du couple $(X, Y)$ est $P(X = a) = P((X = A) \cap (Y = c)) + P(X = a) \cap (Y = d)) = \frac{1}{12} + \frac{1}{6} = \frac{1}{4}$, $P(X = b) = \frac{1}{8} + \frac{5}{8} = \frac{3}{4}$ ...

| X \ Y | c | d | loi de X |
|---|---|---|---|
| a | $\frac{1}{12}$ | $\frac{1}{6}$ | $\frac{1}{4}$ |
| b | $\frac{1}{8}$ | $\frac{5}{8}$ | $\frac{3}{4}$ |
| loi de Y | $\frac{5}{24}$ | $\frac{19}{24}$ | 1 |

**Définition (lois conditionnelles).** Si pour tout $y \in Y(\Omega)$, $P(Y = y) \neq 0$, on peut définir la loi de $X$ sachant que $Y = y$ : $\forall x \in X(\Omega)$, $P_{Y=y}(X = x) = \dfrac{P((X = x) \cap (Y = y))}{p(Y = y)} = \dfrac{P((X, Y) = (x, y))}{p(Y = y)}$.
Si pour tout $x \in X(\Omega)$, $P(X = x) \neq 0$, on peut définir la loi de $Y$ sachant que $X = x$ : $\forall y \in Y(\Omega)$, $P_{X=x}(Y = y) = \dfrac{P((X = x) \cap (Y = y))}{p(X = x)}$.
Les lois conditionnelles sont déterminées par la loi conjointe et les lois marginale et donc par la loi conjointe uniquement.

### 2) Indépendance

#### a) de deux variables

**Définition.** $X$ et $Y$ sont indépendantes si et seulement si $\forall (x, y) \in X(\Omega) \times Y(\Omega)$, $P((X = x) \cap (Y = y)) = P(X = x) \times P(Y = y)$.

#### b) d'un $n$-uplet de variables

**Définition.** $X_1, \ldots, X_n$ sont deux à deux indépendantes si et seulement si $\forall i \neq j$, $X_i$ et $X_j$ sont indépendantes. Ceci équivaut à $\forall i \neq j$, $\forall (x_i, x_j) \in X_i(\Omega) \times X_j(\Omega)$, $P((X_i = x_i) \cap (X_j = x_j)) = P(X_i = x_i) \times P(X_j = x_j)$.

**Définition.** $X_1, \ldots, X_n$ sont indépendantes si et seulement si $\forall (x_1, \ldots, x_n) \in X_1(\Omega) \times X_n(\Omega)$, les événements $\{X_1 = x_1\}$, $\ldots \{X_n = x_n\}$ sont indépendants.

**Théorème.** Si les variables $X_1, \ldots, X_n$ sont indépendantes, alors les variables $X_1, \ldots, X_n$ sont deux à deux indépendantes. Réciproque fausse.

**Théorème.** Si les variables $X_1, \ldots, X_n$ sont indépendantes, alors pour toutes fonctions $f_1, \ldots, f_n$, les variables $f_1(X_1)$, $\ldots$, $f_n(X_n)$ sont indépendantes.

### 3) Covariance

#### a) Cas général

**Définition.** La covariance du couple $(X, Y)$ est $\text{cov}(X, Y) = E((X - E(X))(Y - E(Y)))$.

**Théorème.** $\text{cov}(X, Y) = E(XY) - E(X)E(Y)$ avec $E(XY) = \displaystyle\sum_{(x,y) \in X(\Omega) \times Y(\Omega)} xy\, P((X = x) \cap (Y = y))$.

**Théorème.** $V(X + Y) = V(X) + V(Y) + 2\text{cov}(X, Y)$ et donc aussi $\text{cov}(X, Y) = \frac{1}{2}(V(X + Y) - V(X) - V(Y))$.

Plus généralement, $V(X_1+\ldots+X_n)=\displaystyle\sum_{i=1}^{n}V(X_i)+2\sum_{1\leqslant i<j\leqslant n}\operatorname{cov}(X_i,X_j)$.

**Théorème.** $|\operatorname{cov}(X,Y)|\leqslant\sigma(X)\sigma(Y)$.

**b) Cas de variables indépendantes**

**Théorème.** Si $X$ et $Y$ sont indépendantes,
- $E(XY)=E(X)E(Y)$ ;
- $\operatorname{cov}(X,Y)=0$ ;
- $V(X+Y)=V(X)+V(Y)$.

**Théorème.** Si $X_1, \ldots, X_n$ sont **deux à deux indépendantes** (et en particulier si $X_1, \ldots, X_n$ sont indépendantes),

$$V(X_1+\ldots+X_n)=V(X_1)+\ldots+V(X_n).$$

Si $X_1, \ldots, X_n$ sont **indépendantes** (et pas seulement deux à deux indépendantes),

$$E(X_1\ldots X_n)=E(X_1)\times\ldots\times E(X_n).$$