

Résumé de sup : probabilités

I. Espaces probabilités finis

1) Univers, événements

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

Si Ω est un univers fini. Une partie de Ω est un **événement**. L'ensemble des événements est donc $\mathcal{P}(\Omega)$.

Ω est l'événement certain, \emptyset 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 $A \cap B = \emptyset$. Si $A \subset B$, on dit que A implique B .

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

3) Probabilité

Soit Ω un univers fini. Une **probabilité** sur Ω est une application P de $\mathcal{P}(\Omega)$ dans $[0, 1]$ telle que

1) $P(\Omega) = 1$

2) pour tous événements A et B tels que $A \cap B = \emptyset$, $P(A \cup B) = P(A) + P(B)$.

Dans ce cas, (Ω, P) est un espace probabilisé.

4) Calculs de probabilités

Théorème.

- $P(\emptyset) = 0$.
- $P(\overline{A}) = 1 - P(A)$.
- Si $A \subset B$, $P(A) \leq P(B)$ (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, \dots, A_n sont deux à deux incompatibles, $P(A_1 \cup \dots \cup A_n) = P(A_1) + \dots + P(A_n)$

Si de plus $(A_i)_{1 \leq i \leq n}$ est un système complet d'événements, alors $P(A_1) + \dots + P(A_n) = 1$ et pour tout événement B ,

$$P(B) = \sum_{i=1}^n P(B \cap A_i).$$

Théorème. Pour tout ω de Ω , on pose $p_{\omega} = P(\{\omega\})$.

- $\sum_{\omega \in \Omega} p_{\omega} = 1$
- $\forall A \in \mathcal{P}(\Omega), P(A) = \sum_{\omega \in A} p_{\omega}$.

Théorème (cas de l'équiprobabilité des cas possibles).

Si $\forall \omega \in \Omega, p_{\omega} = \frac{1}{\text{card}(\Omega)}$, alors $\forall A \in \mathcal{P}(\Omega), P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)} = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$.

II. Probabilités conditionnelles

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

Théorème. L'application $P_A : \mathcal{P}(\Omega) \rightarrow \mathbb{R}$ est une probabilité sur Ω .

$$B \mapsto P_A(B)$$

Théorème. Pour tous A et B , $P(A \cap B) = P_A(B) \times P(A)$ si $P(A) \neq 0$.
 $= P_B(A) \times P(B)$ si $P(B) \neq 0$

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

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

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

Théorème (formule de BAYES (inversion d'une probabilité conditionnelle)). Soit $(A_i)_{1 \leq i \leq n}$ un système complet d'événements tels que $\forall i \in \llbracket 1, n \rrbracket, P(A_i) \neq 0$, alors pour tout B tel que $P(B) \neq 0$,

$$\forall i \in \llbracket 1, n \rrbracket, P_B(A_i) = \frac{P(A_i) \times P_{A_i}(B)}{\sum_{i=1}^n P(A_i) \times P_{A_i}(B)}.$$

En particulier, $P_B(A) = \frac{P(A) \times P_A(B)}{P(A) \times P_A(B) + P(\bar{A}) \times P_{\bar{A}}(B)}$.

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 \bar{B} , \bar{A} et B , \bar{A} et \bar{B} sont indépendants.

Soient A_1, \dots, A_n , n événements.

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

Théorème. indépendants $\begin{matrix} \Rightarrow \\ \neq \end{matrix}$ deux à deux indépendants.

IV. Variables aléatoires sur un univers fini

1) Variables aléatoires. Loi d'une variable aléatoire

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

Variable indicatrice. Soit A un événement. La variable $X : \Omega \rightarrow \mathbb{R}$ est la variable indicatrice de $\omega \mapsto \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$

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

- 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, $\{X \leq x\} = X^{-1}([-\infty, x]) = \{\omega \in \Omega / X(\omega) \leq x\}$.

Loi de probabilité d'une variable aléatoire. Soit X une variable aléatoire sur un espace probabilisé fini (Ω, P) .

L'application $X(\Omega) \rightarrow [0, 1]$ est une probabilité sur $X(\Omega)$ appelée loi de X. La loi de X peut aussi être l'application

$$x \mapsto P(X = x)$$

plus générale $\mathcal{P}(X(\Omega)) \rightarrow [0, 1]$. On note P_X la loi de X.

$$A \mapsto P(X \in A)$$

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

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 $Y = X^2$, $P(Y = 1) = P(X^2 = 1) = P(X = 1) + P(X = -1)$.

2) Espérance, variance, écart-type

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

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

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 associée à X .

Théorème (positivité, croissance). Si X est une variable aléatoire réelle positive, alors $E(X) \geq 0$.

Si X et Y sont des variables aléatoires telles que $X \leq Y$, alors $E(X) \leq E(Y)$.

Théorème (inégalité de MARKOV). Si X est une variable réelle positive,

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

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

- Si $\omega \in A$, $1_A(\omega) = 1$ et $\frac{X}{a}(\omega) = \frac{X(\omega)}{a} \geq \frac{a}{a} = 1 = 1_A(\omega)$.
- Si $\omega \notin A$, $1_A(\omega) = 0 \leq \frac{X(\omega)}{a}$.

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

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

b) **Variance, écart-type.**

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

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

Théorème (formule de KOENIG-HUYGENS). $V(X) = E((X - E(X))^2) = E(X^2) - (E(X))^2$.

Théorème. $V(aX + b) = a^2 V(X)$.

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, $\frac{X - E(X)}{\sigma(X)}$ est centrée réduite et est la variable centrée réduite associée à X .

Théorème (inégalité de BIENAYMÉ-TCHEBYCHEV).

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

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

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

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

1) **Couples, n-uplets**

Définition. Soient Ω un univers fini et X et Y deux variables aléatoires sur Ω à valeurs dans E et E' respectivement. L'application $(X, Y) : \Omega \rightarrow E \times E'$ est un **couple** de variables aléatoires sur Ω . Si $E = E' = \mathbb{R}$, (X, Y) est une couple de variables aléatoires réelles sur Ω .

Plus généralement, un n -uplet de variables aléatoires réelles sur Ω est

$$(X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n \\ \omega \mapsto (X_1(\omega), \dots, X_n(\omega))$$

Définition. Si X et Y sont deux variables aléatoires sur l'espace probabilisé fini (Ω, \mathcal{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

$X \backslash Y$	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 $P(X = a) = P((X = a) \cap (Y = c)) + P(X = a) \cap (Y = d)) = \frac{1}{12} + \frac{1}{6} = \frac{1}{4}$,
 $P(X = b) = \frac{1}{8} + \frac{5}{8} = \frac{3}{4} \dots$

$X \backslash Y$	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 $Y = y$:

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

Si pour tout $x \in X(\Omega)$, $P(X = x) \neq 0$, on peut définir la loi de Y sachant que $X = x$: $\forall y \in Y(\Omega)$, $P_{X=x}(Y = y) = \frac{P((X = x) \cap (Y = y))}{p(X = x)}$.

Les lois conditionnelles sont déterminées par la loi conjointe et les lois marginales 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, \dots, 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, \dots, X_n sont indépendantes si et seulement si $\forall (x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$, les événements $\{X_1 = x_1\}, \dots, \{X_n = x_n\}$ sont indépendants.

Théorème. Si les variables X_1, \dots, X_n sont indépendantes, alors les variables X_1, \dots, X_n sont deux à deux indépendantes. Réciproque fautive.

Théorème. Si les variables X_1, \dots, X_n sont indépendantes, alors pour toutes fonctions f_1, \dots, f_n , les variables $f_1(X_1), \dots, f_n(X_n)$ 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 $E(XY) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} xy P((X = x) \cap (Y = y))$.

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, $V(X_1 + \dots + X_n) = \sum_{i=1}^n V(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{cov}(X_i, X_j)$.

Théorème. $|\text{cov}(X, Y)| \leq \sigma(X)\sigma(Y)$.

b) Cas de variables indépendantes

Théorème. Si X et Y sont indépendantes,

- $E(XY) = E(X)E(Y)$;
- $\text{cov}(X, Y) = 0$;
- $V(X + Y) = V(X) + V(Y)$.

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

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

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

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