Chapitre 31

Variable

Hugo Salou MP2I

Dernière mise à jour le 15 juin 2022

TABLE DES MATIÈRES

I Définitions **2**

II Exemples de lois **5**

III Couples de variables aléatoires **7**

IV Familles de variables aléatoires **10**

V Espérance d'une variable aléatoire **12**

VI Variance **17**

VII Covariance (Hors-Programme) **22**

VIII Loi des grands nombres **25**

Première partie

Définitions

Définition: Une variable aléatoire est une application sur un espace probabilisé (Ω, P) dans E où E est un ensemble quelconque :

$$X : \Omega \longrightarrow E.$$

Si $E \subset \mathbb{R}$, on dit que X est réelle.

Si E est un espace vectoriel, on dit que X est un vecteur aléatoire.

Si $E \subset \mathscr{M}_{n,p}(\mathbb{K})$, on dit que X est une matrice aléatoire.

Exemple:
On lance 2 dés bien équilibrés et on note S la somme des points sur ces deux dés.

On pose $\Omega = [\![1, 6]\!]^2$, P l'équiprobabilité et

$$\begin{aligned} S : \Omega &\longrightarrow \mathbb{R} \\ (\omega_1, \omega_2) &\longmapsto \omega_1 + \omega_2. \end{aligned}$$

$S(\Omega) = [\![2, 12]\!]$.

s	2	3	4	5	6	7	8	9	10	11	12
$P(S = s)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Dans la suite du chapitre, (Ω, P) est un espace probabilisé.

Proposition: Soit $X : \Omega \to E$ une variable aléatoire avec $E = X(\Omega)$. L'application

$$\begin{aligned} P_X : \mathscr{P}(E) &\longrightarrow [0, 1] \\ A &\longmapsto P\left(X^{-1}(A)\right) \end{aligned}$$

est une probabilité sur E.

Preuve:
- $\forall A \in \mathscr{P}(E)$, $X^{-1}(A) \in \mathscr{P}(\Omega)$ donc $P\big(X^{-1}(A)\big) \in [0, 1]$.
- $P_X(E) = P\big(X^{-1}(E)\big) = P(\Omega) = 1$.
- Soient $A, B \in \mathscr{P}(E)$ avec $A \cap B = \varnothing$.

$$\begin{aligned} P_X(A \cup B) &= P\big(X^{-1}(A \cup B)\big) \\ &= P\big(X^{-1}(A) \cup X^{-1}(B)\big) \\ &= P\big(X^{-1}(A)\big) + P\big(X^{-1}(B)\big) \\ &= P_X(A) + P_X(B). \end{aligned}$$

□

Définition: Soit $X : \Omega \to E$ une variable aléatoire. P_X est la loi de X.

Remarque (Notations):
Soit $X : \Omega \to E$ une variable aléatoire.
- Soit $A \in \mathscr{P}(E)$. $X^{-1}(A)$ est noté $(X \in A)$. D'où $P_X(A) = P(X \in A)$.

— Soit $x \in E$. $X^{-1}\big(\{x\}\big)$ est noté $(X = x)$; $P_X\big(\{x\}\big) = P(X = x)$.
— Si $E \subset \mathbb{R}$, $X^{-1}\big(]-\infty, x]\big)$ est noté $(X \leqslant x)$ et donc $P\big(X^{-1}\big(]-\infty, x]\big)\big) = P(X \leqslant x)$. De même avec les autres inégalités strictes et larges.

Remarque (Notation):
On note $X \sim Y$ si X et Y suivent la même loi, i.e. $P_X = P_Y$.

Deuxième partie

Exemples de lois

Définition: Soit $X : \Omega \to E$ une variable aléatoire. On dit que X suit une loi uniforme si P_X est l'équiprobabilité sur $X(\Omega)$. On note alors $X \sim \mathscr{U}\big(X(\Omega)\big)$.

Définition: Soit $X : \Omega \to \{0, 1\}$. On note $p = P(X = 1) \in [0, 1]$. On dit que X suit la loi de Bernoulli de paramètre p. On note alors $X \sim \mathscr{B}(p)$.

Définition: Soit $X : \Omega \to [\![0, n]\!]$. On dit que X suit la loi binomiale de paramètres n et p si

$$\forall k \in [\![0, n]\!]\,,\ P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

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

Théorème: On considère une expérience aléatoire qui n'a que deux issues possibles : succès ou échec. On répète n fois à l'identique cette expérience, et on note X le nombre de succès. Alors $X \sim \mathscr{B}(n, p)$ où P est la probabilité de succès. □

Troisième partie

Couples de variables aléatoires

Exemple:
On lance 2 dés bien équilibrés ; on note S la somme des points sur ces deux dés et M le maximum.

S \ M	1	2	3	4	5	6	loi de S
2	$1/36$	0	0	0	0	0	$1/36$
3	0	$2/36$	0	0	0	0	$2/36$
4	0	$1/36$	$2/36$	0	0	0	$3/36$
5	0	0	$2/36$	$2/36$	0	0	$4/36$
6	0	0	$1/36$	$2/36$	$2/36$	0	$5/36$
7	0	0	0	$0/36$	$2/36$	$2/36$	$6/36$
8	0	0	0	$1/36$	$2/36$	$2/36$	$5/36$
9	0	0	0	0	$2/36$	$2/36$	$4/36$
10	0	0	0	0	$1/36$	$2/36$	$3/36$
11	0	0	0	0	0	$2/36$	$2/36$
12	0	0	0	0	0	$1/36$	$1/36$
loi de M	$1/36$	$3/36$	$5/36$	$7/36$	$9/36$	$11/36$	

Définition: Soient $X : \Omega \to E$ et $Y : \Omega \to F$ deux variables aléatoires.

La loi du couple (X, Y) est P_Z où

$$\begin{aligned} Z : \Omega &\longrightarrow E \times F \\ \omega &\longmapsto \big(X(\omega), Y(\omega)\big) \end{aligned}$$

i.e. c'est la donnée de

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

Définition: Soient $X : \Omega \to E$ et $Y : \Omega \to F$. Les lois marginales du couple (X, Y) sont P_X et P_Y.

Proposition: Avec les notations précédentes,

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

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

Preuve:
probabilités totales □

Remarque:
Si

$$\forall y \in Y(\Omega),\ P(Y = y) \neq 0$$

alors

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

Définition: Soient $X : \Omega \to E$ et $Y : \Omega \to F$ deux variables aléatoires. On dit que X

et Y sont indépendantes si

$$\forall A \subset X(\Omega),\ \forall B \subset Y(\Omega),\ P(X \in A, Y \in B) = P(X \in A)\ P(Y \in B).$$

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

Proposition: Avec les notations précédentes,

$$X \perp\!\!\!\perp Y \iff \forall x \in X(\Omega), \forall y \in Y(\Omega),\ P(X = x, Y = y) = P(X = x)P(Y = y).$$

Preuve: " $\Longrightarrow$ " immédiat
" $\Longleftarrow$ " Soient $A = \{x_1, \dots, x_n\} \subset X(\Omega)$ et $B = \{y_1, \dots, y_p\} \subset Y(\Omega)$.

$$\begin{aligned}
P(X \in A, Y \in B) &= P\left(\left(\dot{\bigcup_{i=1}^{n}}(X = x_i)\right) \cap (Y \in B)\right) \\
&= P\left(\dot{\bigcup_{i=1}^{n}}\left((X = x_i) \cap (Y \in B)\right)\right) \\
&= \sum_{i=1}^{n} P\big((X = x_i) \cap (Y \in B)\big) \\
&= \sum_{i=1}^{n} P\left((X = x_i) \cap \left(\dot{\bigcup_{j=1}^{p}}(Y = y_j)\right)\right) \\
&= \sum_{i=1}^{n} P\left(\dot{\bigcup_{j=1}^{p}}\left((X = x_i) \cap (Y = y_i)\right)\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{p} P(X = x_i, Y = y_j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{p} P(X = x_i)\, P(Y = y_j) \\
&= \sum_{i=1}^{n} P(X = x_i \sum_{j=1}^{p} P(Y = y_j) \\
&= P(X \in A)\ P(Y \in B).
\end{aligned}$$

□

Quatrième partie

Familles de variables aléatoires

Définition: Soit $(X_i)_{i\in I} \in \prod_{i\in I} E_i^{\Omega}$ une famille de variables aléatoires. On dit que ces variables sont indépendantes si, pour toute partie finie J de I, et pour tout $(A_j)_{j\in J} \in \prod_{j\in J} \mathscr{P}(E_j)$,

$$P\left(\bigcap_{j\in J}(X_j \in A_j)\right) = \prod_{j\in J} P(X_j \in A_j).$$

Proposition: Avec les notations précédentes, $(X_i)_{i\in I}$ est une famille de variables aléatoires indédpendantes si et seulement si

$$\forall J \subset I,\ J \text{ finie},\ \forall (x_j)_{j\in J} \in \prod_{j\in J} X_j(\Omega),\ P\left(\bigcap_{j\in J}(X_j = x_j)\right) = \prod_{j\in J} P(X_j = x_j).$$

□

Proposition: Soit $(X_i)_{i\in\mathbb{N}}$ une suite de variables aléatoires suivant la même loi de Bernoulli $\mathscr{B}(p)$. Alors,

$$\forall n \in \mathbb{N}^*,\ \sum_{i=1}^{n} X_i \sim \mathscr{B}(n,p).$$

Proposition: Soient $X \sim \mathscr{B}(n,p)$ et $Y \sim \mathscr{B}(m,p)$.

Si $X \perp\!\!\!\perp Y$, alors $X + Y \sim \mathscr{B}(n+m,p)$. □

Cinquième partie

Espérance d'une variable aléatoire

Dans ce paragrape, toutes les variables aléatoires sont à valeurs réelles.

Définition: Soit $X : \Omega \to \mathbb{R}$. L'espérance de X est

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

Exemple (À connaître): 1. Avec $X \sim \mathscr{B}(p)$, on a $E(X) = \sum_{x \in \{0,1\}} xP(X = x) = p$.

2. Avec $X \sim \mathscr{B}(n, p)$, on a

$$\begin{aligned} E(X) &= \sum_{k=0}^{n} k\, P(X = k) \\ &= \sum_{k=1}^{n} k \binom{n}{k} p^k (1-p)^{n-k} \\ &= n \sum_{k=1}^{n} \binom{n-1}{k-1} p^k (1-p)^{n-k} \\ &= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-k-1} \\ &= np(p + 1 - p)^{n-1} \\ &= np \end{aligned}$$

Lemme: Soit $X : \Omega \to \mathbb{R}$.

$$E(X) = \sum_{\omega \in \Omega} X(\omega)\, P\big(\{\omega\}\big).$$

Preuve:

$$\begin{aligned} \sum_{w \in \Omega} X(\omega)\, P\big(\{\omega\}\big) &= \sum_{x \in X(\Omega)} \sum_{\substack{\omega \in \Omega \\ X(\omega) = x}} X(\omega)\, P\big(\{\omega\}\big) \\ &= \sum_{x \in X(\Omega)} \sum_{\substack{\omega \in \Omega \\ X(\omega) = x}} xP\big(\{\omega\}\big) \\ &= \sum_{x \in X(\omega)} x \sum_{\substack{\omega \in \Omega \\ X(\omega) = x}} P\big(\{\omega\}\big) \\ &= \sum_{x \in X(\Omega)} xP\left(\bigcup_{\substack{\omega \in \Omega \\ X(\omega) = x}}^{\cdot} \{\omega\} \right) \\ &= \sum_{x \in X(\Omega)} x\, P(X = x) \end{aligned}$$

□

Théorème (linéarité de l'espérance)**:** Soient X et Y deux variables aléatoires réelles, et α, β deux réels. Alors,

$$E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y).$$

Preuve:

$$\begin{aligned}\alpha E(X) + \beta E(Y) &= \alpha \sum_{\omega\in\Omega} X(\omega)P\big(\{\omega\}\big) = \beta \sum_{\omega\in\Omega} Y(\omega)P\big(\{\omega\}\big)\\ &= \sum_{\omega\in\Omega} \big(\alpha X(\omega) + \beta Y(\omega)\big)P\big(\{\omega\}\big)\\ &= E(\alpha X + \beta Y).\end{aligned}$$

□

Corollaire: Soit X une variable aléatoire suivant la loi binomiale $\mathscr{B}(n,p)$. On a $E(X) = np$.

Preuve:
Soient $X_1, \dots, X_n$ des variables indépendantes suivant toutes $\mathscr{B}(p)$.

On a

$$X \sim \sum_{i=1}^{n} X_i$$

donc

$$\begin{aligned}E(X) &= E\left(\sum_{i=1}^{n} X_i\right)\\ &= \sum_{i=1}^{n} E(X_i)\\ &= np\end{aligned}$$

□

Exemple:
On pose $\Omega = S_n$ le groupe symétrique ; P l'équiprobabilité sur Ω et

$$\begin{aligned}X : \Omega &\longrightarrow [\![0, n]\!]\\ \omega &\longmapsto \text{le nombre de points fixes de } \omega\end{aligned}$$

Que vaut $E(X)$?

On pose, pour tout $\omega \in \Omega$, pour tout $i \in [\![0, n]\!]$, $X_i(\omega) = \begin{cases}1 & \text{si } \omega(i) = i\\ 0 & \text{sinon.}\end{cases}$

Ainsi,

$$\forall \omega \in \Omega,\ \sum_{i=1}^{n} X_i(\omega) = X(\omega).$$

D'où

$$\begin{aligned} E(X) &= E\left(\sum_{i=1}^{n} X_i\right) \\ &= \sum_{i=1}^{n} E(X_i) \\ &= \sum_{i=1}^{n} P(X_i = 1) \\ &= \sum_{i=1}^{n} \frac{(n-1)!}{n!} \\ &= \sum_{i=1}^{n} \frac{1}{n} \\ &= \frac{n}{n} \\ &= 1 \end{aligned}$$

On n'a pas eu besoin de déterminer la loi de X pour déterminer l'espérance de X.

Définition: Soient $X : \Omega \to \mathbb{R}$ une variable aléatoire et $f : \mathbb{R} \to \mathbb{R}$.

$$\begin{array}{ccc} \Omega & \xrightarrow{X} & \mathbb{R} \\ & \searrow & \downarrow f \\ & & \mathbb{R} \end{array}$$

La composée $f \circ X$ est notée $f(X)$.

Théorème (formule de transfert)**:** Avec les notations précédentes,

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

Preuve:

$$\begin{aligned} E\big(f(X)\big) &= \sum_{\omega \in \Omega} f\big(X(\omega)\big)\, P\big(\{\omega\}\big) \\ &= \sum_{x \in X(\Omega)} \sum_{\substack{\omega \in \Omega \\ X(\omega) = x}} f\big(X(\omega)\big)\, P\big(\{\omega\}\big) \\ &= \sum_{x \in X(\Omega)} f(x) \sum_{\substack{\omega \in \Omega \\ X(\omega) = x}} P\big(\{\omega\}\big) \\ &= \sum_{x \in X(\Omega)} f(x)\, P(X = x). \end{aligned}$$

□

Exemple:

Avec $X(\Omega) = \{-1, 0, 1\}$, $P(X = -1) = \dfrac{1}{2}$, $P(X = 0) = \dfrac{1}{4}$, $P(X = 1) = \dfrac{1}{4}$, *que vaut* $E(X^2)$ *?*

— Loi de X^2 : $X^2(\Omega) = \{0, 1\}$ avec $P(X^2 = 0) = \dfrac{1}{4}$ et $P(X^2 = 1) = \dfrac{3}{4}$; d'où $E(X^2) = \dfrac{3}{4}$.

— Formule de transfert : $E(X^2) = \displaystyle\sum_{x=-1}^{1} x^2 P(X = x) = \frac{3}{4}$.

Sixième partie

Variance

Dans ce paragrape, toutes les variables aléatoires sont à valeurs réelles.

Définition: Soit $X : \Omega \to \mathbb{R}$ une variable aléatoire. La variance de X est

$$V(X) = E\left(\left(X - E(X)\right)^2\right).$$

Proposition: Avec les notations précédentes,

$$V(X) \geqslant 0.$$

Preuve:
On pose $\mu = E(X)$.

$$V(X) = \sum_{x \in X(\Omega)} \underbrace{(x-\mu)^2 \, P(X=x)}_{\geqslant 0} \geqslant 0.$$

□

Définition: L'écart-type de X est

$$\sigma_X = \sqrt{V(X)}.$$

Définition: On dit que X vérifie presque sûrement une propriété si la probabilité que X vérifie cette propriété vaut 1.

Proposition:

$$V(X) = 0 \iff X \text{ est presque sûrement constante..}$$

□

Proposition (Kœnig-Huygens)**:**

$$V(X) = E(X^2) - E(X)^2.$$

Preuve:
On pose $\mu = E(X)$

$$\begin{aligned} V(X) &= E\big((X-\mu)^2\big) \\ &= E(X^2 - 2\mu X + \mu^2) \\ &= E(X^2) - 2\mu E(X) + \mu^2 E(1) \\ &= E(X^2) - 2\mu^2 + \mu^2 \\ &= E(X^2) - \mu^2 \end{aligned}$$

□

Exemple: 1. Avec $X \sim \mathscr{B}(p)$, on a $X^2 \sim \mathscr{B}(p)$ d'où

$$V(X) = p - p^2 = p(1-p) = pq \leqslant \frac{1}{4} \text{ avec } q = 1-p$$

2. Avec $X \sim \mathscr{B}(n,p)$.

$$\begin{aligned} E(X^2) &= \sum_{k=0}^{n} k^2 \binom{n}{k} p^k q^{n-k} \\ &= \cdots\cdots\cdots \end{aligned}$$

En faisant le calcul, on trouve $V(X) = npq$.

Théorème: Soit $X, Y : \Omega \to \mathbb{R}$.

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

où $\operatorname{Cov}(X,Y) = E(XY) - E(X)\,E(Y)$ est la covariance de X et Y.

Preuve:

$$\begin{aligned} E\big((X+Y)^2\big) &= E(X^2 + Y^2 + 2XY) \\ &= E(X^2) + E(Y^2) + 2E(XY) \end{aligned}$$

et

$$\begin{aligned} \big(E(X+Y)\big)^2 &= \big(E(X) + E(Y)\big)^2 \\ &= E(X)^2 + E(Y)^2 + 2E(X)E(Y). \end{aligned}$$

D'où

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

□

Proposition: Si $X \perp\!\!\!\perp Y$, alors $\operatorname{Cov}(X,Y) = 0$. D'où

$$V(X+Y) = V(X) + V(Y).$$

Preuve:

$$\begin{aligned} E(XY) &= \sum_{\substack{x\in X(\Omega)\\ y\in Y(\Omega)}} xy\, P(X=x, Y=y) \\ &= \sum_{x\in X(\Omega)} \sum_{y\in Y(\Omega)} xy P(X=x)P(Y=y) \\ &= \sum_{x\in X(\Omega)} x\, P(X=x) \sum_{y\in Y(\Omega)} y\, P(Y=y) \\ &= E(X)E(Y) \end{aligned}$$

□

Proposition: Soit $(X_i)_{1\leqslant i\leqslant n}$ des variabls aléatoires définies sur le même espace probabilisé (Ω, P). Alors

$$\begin{aligned} V\left(\sum_{i=1}^{n} X_i\right) &= \sum_{1\leqslant i,j\leqslant n} \mathrm{Cov}(X_i, X_j) \\ &= \sum_{i=1}^{n} V(X_i) + 2 \sum_{1\leqslant i<j\leqslant n} \mathrm{Cov}(X_i, X_j) \end{aligned}$$

□

Corollaire: Avec les notations précédentes, si les X_i sont deux à deux non corrélés (i.e. $\mathrm{Cov}(X_i, X_j) = 0$ pour $i \neq j$), alors

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i).$$

□

Corollaire: Avec les notations précédentes, si les X_i sont indépendantes deux à deux, alors

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i).$$

□

Corollaire: Soit $X \sim \mathscr{B}(n, p)$. Alors,

$$V(X) = np(1-p).$$

Preuve:
On pose $X \sim \sum_{i=1}^{n} X_i$ où les (X_i) sont mutuellement indépendants et $\forall i$, $X_i \sim \mathscr{B}(p)$.

$$\begin{aligned} V(X) &= V\left(\sum_{i=1}^{n} X_i\right) \\ &= \sum_{i=1}^{n} V(X_i) \\ &= \sum_{i=1}^{n} p(1-p) \\ &= np(1-p) \end{aligned}$$

□

Proposition (transfert)**:** Soient $X_1, \ldots, X_n$ des variables aléatoires à valeurs réelles toutes définies sur un même espace probabilisé (Ω, P) et $f : \mathbb{R}^n \longrightarrow \mathbb{R}$

$$E\big(f(X_1, \ldots, X_n)\big) = \sum_{x_1 \in X_1(\Omega)} \cdots \sum_{x_n \in X_n(\Omega)} f(x_1, \ldots, x_n) P(X_1 = x_1, \ldots, X_n = x_n).$$

Preuve:

$$\begin{aligned} E\big(f(X_1, \ldots, X_n)\big) &= \sum_{\omega \in \Omega} f\big(X_1(\omega), \ldots, X_n(\omega)\big)\, P\big(\{\omega\}\big) \\ &= \sum_{x_1 \in X_1(\Omega)} \cdots \sum_{x_n \in X_n(\Omega)} \sum_{\substack{\omega \in \Omega \\ \forall i,\, X_i(\omega) = x_i}} f\big(X_1(\omega), \ldots, X_n(\omega)\big)\, P\big(\{\omega\}\big) \\ &= \sum_{x_1 \in X_1(\Omega)} \cdots \sum_{x_n \in X_n(\Omega)} f(x_1, \ldots, x_n) \sum_{\substack{\omega \in \Omega \\ \forall i,\, X_i(\omega) = x_i}} P\big(\{\omega\}\big) \\ &= \sum_{x_1 \in X_1(\Omega)} \cdots \sum_{x_n \in X_n} f(x_1, \ldots, x_n) P(X_1 = x_1, \ldots, X_n = x_n) \end{aligned}$$

□

Septième partie

Covariance (Hors-Programme)

On se place dans une optique de Big Data. On dispose d'un tableau à N lignes : chaque ligne correspond à une observation et chaque colonne à une "mesure" (ou caractéristique).

Ces caractéristiques peuvent être corrélées plus ou moins fortement et contenir plus ou moins d'information.

Plus la variance est grande, plus il y a d'information.

Soient X et Y deux colonnes. D'après l'inégalité de Cauchy-Schwarz :

$$\big|\operatorname{Cov}(X,Y)\big| \leqslant \sigma_X\,\sigma_Y \quad \text{donc} \quad -1 \leqslant \frac{\operatorname{Cov}(X,Y)}{\sigma_X\,\sigma_Y} \leqslant 1.$$

Si $Y = \alpha X + \beta$: $$\left|\frac{\operatorname{Cov}(X,Y)}{\sigma_X,\sigma_Y}\right| = 1.$$

L'objectif est de modifier les colonnes de façon à extraire le plus d'informations possible sur le moins de colonnes possibles.

La matrice de covariance est

$$A = \big(\operatorname{Cov}(X_i,X_j)\big)_{\substack{1\leqslant i\leqslant N\\ 1\leqslant j\leqslant N}}.$$

On aimerait que la matrice A soit diagonale avec de grands coefficients diagonaux.

L'année prochaine, nous verrons le théorème suivant :

> **Théorème** (théorème spectral)**:** Toute matrice symétrique réelle est diagonalisable dans une base orthonormée de vecteurs propres.

Il existe donc des variables aléatoires $Y_1,\dots,Y_k$ combinaisons linéaires des $X_1,\dots,X_k$ telles que

$$\big(\operatorname{Cov}(Y_i,Y_j)\big) = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & (0) \\ & & \ddots & \\ & (0) & & \lambda_k \end{pmatrix}.$$

De plus, le maximum de $V(\alpha_1X_1+\cdots+\alpha_kX_k)$ avec la condition que $\alpha_1^2+\cdots+\alpha_k^2=1$ est $V(Y_1)=\lambda_1$.

Définition (Multiplicateurs de Lagrange)**:** Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ où U est un ouvert de $\mathbb{R}^n$. On cherche $\max\limits_{(x_i)_{i\in[\![1,n]\!]}\in R^n} f(x_1,\dots,x_n)$ avec la contrainte $g(x_1,\dots,x_n)$.

On pose

$$\begin{aligned} F : \mathbb{R}^{n+1} &\longrightarrow \mathbb{R} \\ (x_1,\dots,x_n,\lambda) &\longmapsto f(x_1,\dots,x_n) + \lambda g(x_1,\dots,x_n). \end{aligned}$$

On cherche les points critiques de F :

$$\begin{aligned} \nabla F(x_1,\dots,x_n,\lambda) = 0 &\iff \begin{cases} \forall i,\ \dfrac{\partial F}{\partial x_i}(x_1,\dots,x_n,\lambda) = 0 \\ \dfrac{\partial F}{\partial \lambda}(x_1,\dots,x_n,\lambda) = 0 \end{cases} \\ &\iff \begin{cases} \forall i,\ \dfrac{\partial f}{\partial x_i}(x_1,\dots,x_n) + \lambda\dfrac{\partial g}{\partial x_i}(x_1,\dots,x_n) = 0 \\ g(x_1,\dots,x_n) = 0 \end{cases} \end{aligned}$$

Si $(x_1,\dots,x_n,\lambda)$ est le maximum de F, alors
$\forall y_1,\dots,y_n,\mu \in \mathbb{R}^{n+1}$ $\qquad F(x_1,\dots,x_n,\lambda) \geqslant F(y_1,\dots,y_n,\mu)$

$$f(x_1,\dots,x_n)+\underbrace{\lambda g(x_1,\dots,x_n)}_{=0}\geqslant f(y_1,\dots,y_n)$$

Huitième partie

Loi des grands nombres

Lemme (inégalité de Марков [a]): Soit $X : \Omega \to \mathbb{R}^+$ une variable aléatoire positive.

$$\forall t \in \mathbb{R}_+^*,\ P(X \geqslant t) \leqslant \frac{E(X)}{t}$$

ou encore,

$$\forall u \in \mathbb{R}_+^*,\ P\big(X \geqslant u\,E(X)\big) \leqslant \frac{1}{u}.$$

a. Марков : Markov

Preuve:
Soit $t \in \mathbb{R}_+^*$.

$$\begin{aligned} E(X) &= \sum_{x \in X(\Omega)} x\,P(X = x) \\ &= \sum_{\substack{x \in X(\Omega) \\ x \geqslant t}} x\,P(X = x) + \sum_{\substack{x \in X(\Omega) \\ x < t}} x\,P(X = x) \\ &\geqslant \sum_{\substack{x \in X(\Omega) \\ x \geqslant t}} t\,P(X = x) + \sum_{\substack{x \in X(\Omega) \\ x < t}} 0 \\ &\geqslant t\,P(X \geqslant t). \end{aligned}$$

□

Lemme (inégalté de Bienaymé-чебышёв [a]): Soit $X : \Omega \to \mathbb{R}$ une variable aléatoire réelle. On pose $\mu = E(X)$ et $\sigma = \sigma_X$. Alors,

$$\forall t \in \mathbb{R}_+^*,\ P\big(|X - \mu| \geqslant t\big) \leqslant \frac{\sigma^2}{t^2}$$

ou encore

$$\forall u \in \mathbb{R}_*^+,\ P\big(|X - \mu| \geqslant u\sigma\big) \leqslant \frac{1}{u^2}.$$

a. чебышёв : Tchebychev

Preuve:
On pose $Y = |X - \mu|^2$, Y est positive.

$$E(Y) = E\left((X - \mu)^2\right) = V(X) = \sigma^2.$$

Donc, d'après l'inégalité de Марков :

$$\forall y \in \mathbb{R}_+^*,\ P(Y \geqslant y) \leqslant \frac{\sigma^2}{y}$$

donc

$$\forall t > 0,\ P\left((X - \mu)^2 \geqslant t^2\right) \leqslant \frac{\sigma^2}{t^2}.$$

Or, $\big(|X-\mu| \geqslant t\big) = \big((X-\mu)^2 \geqslant t^2\big)$, donc

$$P\big(|X-\mu) \geqslant t\big) \leqslant \frac{\sigma^2}{t^2}.$$

□

Remarque (Application):
On considère une expérience de Bernoulli de paramètre $p \in]0,1[$. On répète n fois cette expérience à l'identique et on note Y_n le nombre moyen de succès.

On pose, pour $i \in [\![1,n]\!]$,

$$X_i = \begin{cases} 1 & \text{si la } i\text{-ème expérience est un succès} \\ 0 & \text{sinon.} \end{cases}$$

Les (X_i) sont mutuellement indépendants ; d'où

$$Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Et donc

$$E(Y_n) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}np = p.$$

Comme les (X_i) sont deux à deux indépendantes,

$$V(Y_n) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{1}{n^2}npq = \frac{pq}{n} \text{ où } q = 1-p.$$

D'où, d'après l'inégalité de Bienaymé-чебышёв,

$$\forall \varepsilon > 0,\ P(|Y_n - p| \geqslant \varepsilon) \leqslant \frac{pq}{n\varepsilon^2} \xrightarrow[n\to+\infty]{} 0.$$

On en déduit que

$$\forall \varepsilon > 0,\ P(|Y_n - p| < \varepsilon) \geqslant \underbrace{1 - \frac{pq}{n\varepsilon^2}}_{\text{“seuil”}}.$$

— Par exemple, avec $\varepsilon = 10^{-2}$ et $\text{seuil} = 95\,\%$, on a

$$\begin{aligned} \frac{95}{100} = \text{seuil} = 1 - \frac{pq}{n \times 10^{-4}} &\iff \frac{5}{100} = \frac{pq}{n\,10^{-4}} \\ &\iff 20 = \frac{n\,10^{-4}}{pq} \\ &\iff n = 20\underbrace{pq}_{\leqslant \frac{1}{4}}10^4 \leqslant 5 \times 10^4. \end{aligned}$$

En répétant 50 000 fois l'expérience, on est sûr à 95 % que la valeur observée de $Y_{50\,000}$ est dans l'intervalle $\big[p - 10^{-2}, p + 10^{-2}\big]$.

— Avec $n = 12$ et $\text{seuil} = 95\,\%$, on a

$$\begin{aligned} \frac{95}{100} = \text{seuil} = 1 - \frac{pq}{n\varepsilon^2} &\iff \frac{pq}{n\varepsilon^2} = \frac{5}{100} \\ &\iff \frac{\varepsilon^2 n}{pq} = 20 \\ &\iff \varepsilon^2 = \frac{20}{n}pq \leqslant \frac{5}{12}. \end{aligned}$$

Donc, $\varepsilon = \sqrt{\dfrac{5}{12}} \simeq 0{,}65$.

— Avec $n = 12$ et $\varepsilon = 10^{-2}$,

$$\begin{aligned} P\big(|Y_{12} - p| < \varepsilon\big) &\geqslant 1 - \frac{pq}{12 \cdot 10^{-4}} \\ &\geqslant 1 - \frac{1}{48 \cdot 10^{-4}} \\ &\geqslant \underbrace{1 - \frac{1}{48} \times 10^4}_{<0} \end{aligned}$$

Théorème (loi faible des grands nombres — Hors-programme)**:** Soit $(X_n)_{n\in\mathbb{N}}$ une suite de varialbe aléatoires indépedantes de même loi définies sur le même espace probabilisé (Ω, P). On note μ leur espérance commune.

$$\forall \varepsilon > 0,\ \lim_{n\to+\infty} P\Big(\big|\bar{X}_n - \mu\big| > \varepsilon\Big) = 0$$

où $\forall n,\ \bar{X}_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$.

Preuve:

$$\begin{aligned} \forall n \in \mathbb{N}^*, E\big(\bar{X}_n\big) &= \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \mu \\ V\big(\bar{X}_n\big) &= \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \end{aligned}$$

où σ^2 est la variance commune aux X_i.

Soit $\varepsilon > 0$.

$$0 \leqslant P\Big(\big|\bar{X}_n - \mu\big| > \varepsilon\Big) \leqslant \frac{\sigma^2}{n\varepsilon^2} \xrightarrow[n\to+\infty]{} 0..$$

□

Proposition: Soient $X_1, \ldots, X_n$ des variables aléatoires à valeurs réelles d'un même ensemble probabilisé mutuellement indépendantes. Soient $f : \mathbb{R}^p \to \mathbb{R}$ et $g : \mathbb{R}^{n-p} \to \mathbb{R}$. Alors, $f(X_{i_1}, \ldots, X_{i_p})$ et $g(X_{j_1}, \ldots, X_{j_{n-p}})$ sont indépendantes si

$$\begin{cases} \{i_1, \ldots, i_p\} \cap \{j_1, \ldots, j_{n-p}\} = \varnothing \\ \forall k \neq \ell,\ i_k \neq i_\ell \text{ et } j_k \neq j_\ell. \end{cases}$$