

# Chapitre 2

## Vecteurs aléatoires gaussiens

### 1 Vecteurs aléatoires

Certaines notions que l'on a définies dans le chapitre précédent que pour les variables aléatoires réelles se transposent pour les variables aléatoires vectorielles.

On dispose dans  $\mathbb{R}^d$  du produit scalaire usuel :

$$\langle x | y \rangle = x_1 y_1 + \cdots + x_d y_d$$

si  $x = (x_1, \dots, x_d)$  et  $y = (y_1, \dots, y_d) \in \mathbb{R}^d$ .

#### 1.1 Fonction caractéristique

Si  $X : \Omega \rightarrow \mathbb{R}^d$  est une variable aléatoire, on définit sa **fonction caractéristique**  $\Phi_X : \mathbb{R}^d \rightarrow \mathbb{C}$  par :

$$\boxed{\Phi_X(u) = \mathbb{E}(e^{i\langle u | X \rangle})}, \quad \forall u \in \mathbb{R}^d,$$

où  $(\langle u | X \rangle)(\omega) = \langle u | X(\omega) \rangle, \forall \omega \in \Omega$ . Autrement dit, si  $X = (X_1, \dots, X_d)$  et  $u = (u_1, \dots, u_d)$  :

$$\boxed{\Phi_X(u) = \mathbb{E}(e^{i(u_1 X_1 + \cdots + u_d X_d)})}$$

**Convention.** Il est d'usage d'identifier les vecteurs de  $\mathbb{R}^d$  à des **matrices-colonnes** ; ainsi on écrira le vecteur aléatoire  $X = (X_1, \dots, X_d)$  sous la forme :

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix}$$

et alors :

$${}^t X = (X_1 \cdots X_d)$$

est une matrice-ligne (noter l'absence maintenant des virgules).

Si  $u = \begin{pmatrix} u_1 \\ \vdots \\ u_d \end{pmatrix}$ , le produit scalaire s'écrira, en identifiant les éléments de  $\mathbb{R}$

à des matrices à une ligne et une colonne, comme le produit des matrices :

$$\langle u | X \rangle = {}^t X \cdot u = {}^t u \cdot X.$$

Comme lorsque  $d = 1$ , on a :

**Théorème 1.1** *Deux v.a. possédant la même fonction caractéristique ont la même loi : si  $X$  et  $Y$  sont deux v.a. à valeurs dans  $\mathbb{R}^d$  telles que  $\Phi_X = \Phi_Y$ , alors  $\mathbb{P}_X = \mathbb{P}_Y$ .*

**Preuve.** Elle est la même que celle faite dans le cas  $d = 1$ , en utilisant l'unité approchée définie par  $p_a(y) = \frac{1}{a^d} p(\frac{y}{a})$ ,  $a > 0$ , avec  $p(y) = p(y_1, \dots, y_d) = \frac{1}{\pi^d(1+y_1^2) \cdots (1+y_d^2)}$ .  $\square$

**Corollaire 1.2** *Soit  $X_k: \Omega \rightarrow \mathbb{R}^{d_k}$ ,  $1 \leq k \leq n$ , des v.a.; elles sont indépendantes si et seulement si :*

$$\Phi_{[X_1, \dots, X_n]}(t_1, \dots, t_n) = \Phi_{X_1}(t_1) \cdots \Phi_{X_n}(t_n), \quad \forall t_k \in \mathbb{R}^{d_k}, 1 \leq k \leq n.$$

**Preuve.** Le Théorème de Fubini donne :

$$\Phi_{X_1}(t_1) \cdots \Phi_{X_n}(t_n) = \int_{\mathbb{R}^{d_1 + \dots + d_n}} e^{i(t_1 x_1 + \dots + t_n x_n)} d(\mathbb{P}_{X_1} \otimes \dots \otimes \mathbb{P}_{X_n})(x_1, \dots, x_n);$$

d'où le résultat, puisque le membre de droite est la f.c d'une v.a. suivant la loi  $\mathbb{P}_{X_1} \otimes \dots \otimes \mathbb{P}_{X_n}$ .  $\square$

## 1.2 Espérance

On dira que  $X = (X_1, \dots, X_d): \Omega \rightarrow \mathbb{R}^d$  est *intégrable* (resp., de carré *intégrable*, resp. de puissance  $r^{\text{ème}}$  *intégrable* :  $X \in L^r(\Omega; \mathbb{R}^d)$ ) si chacune de ses composantes  $X_1, \dots, X_d: \Omega \rightarrow \mathbb{R}$  l'est. On notera, pour  $X$  intégrable :

$$\mathbb{E}(X) = (\mathbb{E}(X_1), \dots, \mathbb{E}(X_d)),$$

ou, matriciellement :

$$\mathbb{E}(X) = \begin{pmatrix} \mathbb{E}(X_1) \\ \vdots \\ \mathbb{E}(X_d) \end{pmatrix}.$$

Il est facile de voir que la v.a.  $X \in L^r(\Omega; \mathbb{R}^d)$  si et seulement si :

$$\mathbb{E}(\|X\|^r) < +\infty.$$

**Proposition 1.3** Soit  $A: \mathbb{R}^d \rightarrow \mathbb{R}^{d'}$  une application **linéaire**, et  $X: \Omega \rightarrow \mathbb{R}^d$  un vecteur aléatoire intégrable. Alors :

$$\boxed{\mathbb{E}(AX) = A\mathbb{E}(X)}.$$

Dans cette proposition, de la même manière que l'on a identifié les vecteurs à des matrices-colonnes, on **identifie** l'application linéaire  $A$  à sa matrice; donc :

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,d} \\ \vdots & & \vdots \\ \alpha_{d',1} & \cdots & \alpha_{d',d} \end{pmatrix}.$$

**Preuve.** Il suffit de l'écrire; on a :

$$AX = \left( \sum_{l=1}^d \alpha_{k,l} X_l \right)_{1 \leq k \leq d'};$$

chacune des composantes de  $AX$  est intégrable; donc  $AX$  est intégrable et :

$$\begin{aligned} \mathbb{E}(AX) &= \left( \mathbb{E} \left( \sum_{l=1}^d \alpha_{k,l} X_l \right) \right)_{1 \leq k \leq d'} \\ &= \left( \sum_{l=1}^d \alpha_{k,l} X_l \mathbb{E}(X_l) \right)_{1 \leq k \leq d'} = A\mathbb{E}(X). \quad \square \end{aligned}$$

En particulier, pour tout  $a \in \mathbb{R}^d$ , on a :

$$\boxed{\mathbb{E}(\langle a | X \rangle) = \langle a | \mathbb{E}(X) \rangle},$$

ou, autrement écrit :

$$\boxed{\mathbb{E}({}^t a \cdot X) = {}^t a \cdot \mathbb{E}(X)}.$$

**Définition 1.4** Si  $M = (X_{k,l})_{\substack{1 \leq k \leq n \\ 1 \leq l \leq s}}$  est une matrice formée de variables aléatoires réelles  $X_{k,l}: \Omega \rightarrow \mathbb{R}$ , on dit que c'est une **matrice aléatoire**. On dit qu'elle est intégrable si chacune des  $X_{k,l}$  l'est, et l'on pose alors :

$$\mathbb{E}(M) = (\mathbb{E}(X_{k,l}))_{k,l}.$$

**Proposition 1.5** Soit  $M$  une matrice aléatoire et  $A$  et  $B$  deux matrices (**non** aléatoires; on dit qu'elles sont **déterministes**) telles que les produits  $AM$  et  $MB$  existent et soient intégrables. On a :

$$\boxed{\mathbb{E}(AM) = A\mathbb{E}(M) \quad \text{et} \quad \mathbb{E}(MB) = \mathbb{E}(M)B}.$$

**Preuve.** Il suffit de regarder la définition du produit des matrices. □

**Définition 1.6** Si  $X: \Omega \rightarrow \mathbb{R}^d$  est un vecteur aléatoire de carré intégrable, on définit sa **matrice de covariance**, avec les notations matricielles, par :

$$K_X = \mathbb{E} \left[ (X - \mathbb{E}(X)) \cdot {}^t(X - \mathbb{E}(X)) \right].$$

On notera que :

$$(X - \mathbb{E}(X)) \cdot {}^t(X - \mathbb{E}(X)) = \begin{pmatrix} X_1 - \mathbb{E}(X_1) \\ \vdots \\ X_d - \mathbb{E}(X_d) \end{pmatrix} \cdot (X_1 - \mathbb{E}(X_1) \cdots X_d - \mathbb{E}(X_d))$$

est une matrice carrée d'ordre  $d$ , dont les termes sont :

$$(X_k - \mathbb{E}(X_k))(X_l - \mathbb{E}(X_l)), \quad 1 \leq k, l \leq d.$$

Lorsque  $X, Y \in L^2(\Omega)$  sont deux *v.a. réelles* de carré intégrable, on définit leur **covariance** par :

$$\text{cov}(X, Y) = \mathbb{E} [(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

Lorsque  $Y = X$ , on obtient la *variance* de  $X$ .

Revenant au cas où  $X = (X_1, \dots, X_d)$  est un vecteur aléatoire, on voit que sa matrice de covariance est :

$$K_X = \left( \text{cov}(X_k, X_l) \right)_{1 \leq k, l \leq d}.$$

**Proposition 1.7** La matrice de covariance est symétrique et **positive**. La forme quadratique associée est non dégénérée si et seulement si, en tant qu'éléments de  $L^2(\Omega, \mathbb{P})$ , les vecteurs  $X_1 - \mathbb{E}(X_1), \dots, X_d - \mathbb{E}(X_d)$  sont **linéairement indépendants**.

**Preuve.** La symétrie est évidente.

Rappelons qu'une matrice symétrique est positive si la forme bilinéaire (ou la forme quadratique) associée est une forme positive. Il faut donc voir que, pour

tout  $u = \begin{pmatrix} u_1 \\ \vdots \\ u_d \end{pmatrix}$ , on a :

$${}^t u K_X u \geq 0.$$

Mais cela résulte du lemme suivant.

**Lemme 1.8** Pour tout  $u \in \mathbb{R}^d$ , on a :

$${}^t u K_X u = \mathbb{E} \left( [\langle u | X - \mathbb{E}(X) \rangle]^2 \right).$$

**Preuve.** C'est un simple calcul ; grâce à la Proposition 1.3, on a :

$${}^t u K_X u = {}^t u \mathbb{E} ([X - \mathbb{E}(X)]. {}^t [X - \mathbb{E}(X)]) u = \mathbb{E} ({}^t u. [X - \mathbb{E}(X)]. {}^t [X - \mathbb{E}(X)]. u) ;$$

d'où le résultat, puisque  ${}^t u. [X - \mathbb{E}(X)] = \langle u | X - \mathbb{E}(X) \rangle$  et que  ${}^t [X - \mathbb{E}(X)]. u = \langle X - \mathbb{E}(X) | u \rangle = \langle u | X - \mathbb{E}(X) \rangle$ .  $\square$

Cette formule montre de plus que la forme quadratique  $u \mapsto {}^t u K_X u$  est dégénérée si, et seulement si, il existe  $u \neq 0$  tel que  $\langle u | X - \mathbb{E}(X) \rangle = 0$ , ce qui signifie que l'on a, dans  $L^2(\mathbb{P})$ ,  $u_1(X_1 - \mathbb{E}(X_1)) + \dots + u_d(X_d - \mathbb{E}(X_d)) = 0$  pour des  $u_1, \dots, u_d$  non tous nuls ; autrement dit, que  $X_1 - \mathbb{E}(X_1), \dots, X_d - \mathbb{E}(X_d)$  sont linéairement indépendants dans  $L^2(\mathbb{P})$ .  $\square$

**Remarque.** Il faut bien faire attention que si des *v.a.r.*, non constantes, sont *indépendantes*, elles sont *linéairement indépendantes* (sinon, l'une peut s'exprimer comme combinaison linéaire des autres ; elle ne peut donc être indépendante de celles-ci), l'inverse est bien sûr loin d'être vrai.

**Exemple.** Si  $X$  suit la loi uniforme sur  $[0, 1]$ ,  $X$  et  $X^2$ , qui ne sont, bien sûr, pas indépendantes, sont linéairement indépendantes, car si  $aX + bX^2 = 0$ , alors, d'une part, en prenant l'espérance, on a  $\frac{a}{2} + \frac{b}{3} = 0$ , puisque :

$$\mathbb{E}(X^k) = \int_0^1 x^k dx = \frac{1}{k+1} ;$$

mais on a aussi, d'autre part,  $aX^2 + bX^3 = 0$ , d'où  $\frac{a}{3} + \frac{b}{4} = 0$  ; donc  $a = b = 0$ .

Notons que, par définition, la covariance de deux *v.a.r.*  $X$  et  $Y$  est le produit scalaire, dans  $L^2(\mathbb{P})$ , des *v.a.r.* centrées  $X - \mathbb{E}(X)$  et  $Y - \mathbb{E}(Y)$  :

$$\text{cov}(X, Y) = \langle X - \mathbb{E}(X) | Y - \mathbb{E}(Y) \rangle .$$

La covariance est donc nulle (on dit alors que  $X$  et  $Y$  sont *non corrélées*, ou *décorrélées*, si et seulement si les *v.a.r.* centrées  $X - \mathbb{E}(X)$  et  $Y - \mathbb{E}(Y)$  sont orthogonales. Lorsque  $X$  et  $Y$  sont indépendantes, on a  $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$  ; donc  $\text{cov}(X, Y) = 0$ . Ainsi :

**Proposition 1.9** *Si  $X$  et  $Y$  sont deux v.a.r. indépendantes, alors  $\text{cov}(X, Y) = 0$ , et  $X - \mathbb{E}(X)$  et  $Y - \mathbb{E}(Y)$  sont orthogonales dans  $L^2(\mathbb{P})$ .*

On avait déjà vu cela dans la Remarque suivant le Corollaire 14 du Chapitre 1.

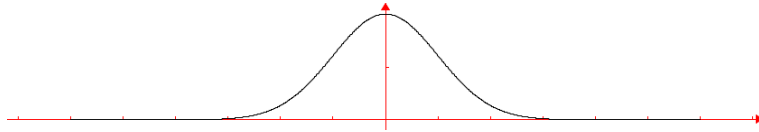
**Corollaire 1.10** *Si  $X_1, \dots, X_d : \Omega \rightarrow \mathbb{R}$  sont indépendantes, la matrice de covariance du vecteur aléatoire  $X = (X_1, \dots, X_d)$  est diagonale.*

La réciproque est évidemment en général fautive, puisque le fait que la matrice de covariance soit diagonale signifie seulement que les composantes du vecteur aléatoire sont deux-à-deux non corrélées.

## 2 Vecteurs aléatoires gaussiens

Rappelons que la **loi gaussienne** (ou **normale**) **centrée réduite** est la loi  $\mathcal{N}(0, 1)$  sur  $\mathbb{R}$  dont la densité est :

$$f_G(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$



Si  $G \sim \mathcal{N}(0, 1)$ , sa *f.c.* est  $\Phi_G(t) = e^{-t^2/2}$ ; son espérance est nulle :  $\mathbb{E}(G) = 0$ , et sa variance est  $\text{Var}(G) = 1$ .

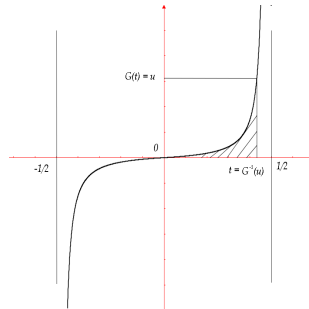
Si  $G \sim \mathcal{N}(0, 1)$ , alors  $\sigma G + m \sim \mathcal{N}(m, \sigma^2)$  et, inversement, si  $X \sim \mathcal{N}(m, \sigma^2)$ , alors  $G = \frac{X-m}{\sigma} \sim \mathcal{N}(0, 1)$ .

On **conviendra** de dire que les **constantes** sont des **gaussiennes dégénérées** (correspondant à  $\sigma = 0$ ).

On peut “ visualiser une *v.a.r.* gaussienne centrée réduite  $G$  ainsi : prenons  $\Omega = ] -\frac{1}{2}, \frac{1}{2} [$ , muni de sa tribu borélienne, avec la probabilité  $\mathbb{P}$  qui est la mesure de Lebesgue  $\lambda$ .

Soit  $G$  la fonction impaire sur  $\Omega$  qui, sur  $[0, \frac{1}{2} [$ , est la fonction réciproque de :

$$u \mapsto \int_0^u e^{-x^2/2} \frac{dx}{\sqrt{2\pi}}$$



On a bien, puisque  $G^{-1}$  est strictement croissante :

$$\mathbb{P}(0 \leq G \leq u) = \lambda([0, G^{-1}(u)]) = G^{-1}(u) = \int_0^u e^{-x^2/2} \frac{dx}{\sqrt{2\pi}}.$$

**Définition 2.1** On dit qu'un vecteur aléatoire  $X: \Omega \rightarrow \mathbb{R}^d$  est un **vecteur aléatoire gaussien** si la v.a.r.  $\varphi(X): \Omega \rightarrow \mathbb{R}$  est gaussienne, pour toute forme linéaire  $\varphi: \mathbb{R}^d \rightarrow \mathbb{R}$ .

En d'autres termes, si  $X = (X_1, \dots, X_d)$ , les v.a.r.  $a_1X_1 + \dots + a_dX_d$  doivent être gaussiennes, pour **tout choix** des nombres réels  $a_1, \dots, a_d$ .

On dira souvent **vecteur gaussien** au lieu de "vecteur aléatoire gaussien".

Il est clair que :

**Proposition 2.2** Si  $X: \Omega \rightarrow \mathbb{R}^d$  est un vecteur gaussien et si  $A: \mathbb{R}^d \rightarrow \mathbb{R}^{d'}$  est une application **linéaire**, alors  $AX: \Omega \rightarrow \mathbb{R}^{d'}$  est encore un vecteur gaussien.

puisque  $\psi \circ A$  est une forme linéaire sur  $\mathbb{R}^d$ , pour toute forme linéaire  $\psi$  sur  $\mathbb{R}^{d'}$ .

Il résulte aussi immédiatement de la définition que l'on a :

**Proposition 2.3** Si  $X = (X_1, \dots, X_d): \Omega \rightarrow \mathbb{R}^d$  est un vecteur gaussien, alors chaque v.a.r.  $X_k: \Omega \rightarrow \mathbb{R}$ , est gaussienne.

puisque les applications  $(x_1, \dots, x_d) \mapsto x_k$  sont des formes linéaires sur  $\mathbb{R}^d$ .

L'inverse est faux, comme on le verra. Néanmoins, on a :

**Théorème 2.4** Soit  $X_1, \dots, X_d: \Omega \rightarrow \mathbb{R}$  des v.a.r. gaussiennes ; si elles sont **indépendantes**, alors le vecteur  $X = (X_1, \dots, X_d)$  est gaussien.

**Preuve.** Pour tous réels  $a_1, \dots, a_d \in \mathbb{R}$ , les v.a.r.  $a_1X_1, \dots, a_dX_d$  sont encore gaussiennes et indépendantes ; alors (voir Chapitre 1, Corollaire 5.10), la v.a.r.  $a_1X_1 + \dots + a_dX_d$  est encore gaussienne.  $\square$

**Théorème 2.5** Si  $X$  est un vecteur gaussien, sa fonction caractéristique vaut :

$$\Phi_X(u) = \exp\left(i {}^t u \cdot \mathbb{E}(X) - \frac{1}{2} {}^t u K_X u\right), \quad \forall u \in \mathbb{R}^d,$$

où  $K_X$  est la matrice de covariance de  $X$ .

On notera que lorsque  $X$  est une v.a.r. gaussienne ( $d = 1$ ), alors  $u \in \mathbb{R}$  et  ${}^t u K_X u = \text{Var}(X)u^2 = \sigma^2 u^2$ ,  ${}^t u \cdot \mathbb{E}(X) = mu$  ; cette formule étend bien le cas scalaire.

**Preuve.** L'application  $x \mapsto {}^t u \cdot x = \langle u | x \rangle$  est une forme linéaire. La v.a.r.  ${}^t u \cdot X$  est donc gaussienne. Son espérance est  $m = \mathbb{E}({}^t u \cdot X) = {}^t u \cdot \mathbb{E}(X)$ , et sa variance est  $\sigma^2 = \text{Var}({}^t u \cdot X) = {}^t u K_X u$  ; en effet :

**Lemme 2.6** Pour tout vecteur aléatoire  $X$ , on a :  ${}^t u K_X u = \text{Var}({}^t u \cdot X)$ .

**Preuve du lemme.** On a vu (Lemme 1.8) que  ${}^t u K_X u = \mathbb{E}([\langle u | X - \mathbb{E}(X) \rangle]^2)$ ; mais :

$$\langle u | X - \mathbb{E}(X) \rangle = {}^t u \cdot (X - \mathbb{E}(X)) = {}^t u \cdot X - {}^t u \cdot \mathbb{E}(X) = {}^t u \cdot X - \mathbb{E}({}^t u \cdot X);$$

donc :

$${}^t u K_X u = \mathbb{E}([{}^t u \cdot X - \mathbb{E}({}^t u \cdot X)]^2) = \text{Var}({}^t u \cdot X). \quad \square$$

**Suite de la preuve du Théorème 2.5.** Donc, pour tout  $v \in \mathbb{R}$ , on a :

$$\Phi_{{}^t u \cdot X}(v) = \exp\left(ivm - \frac{1}{2}\sigma^2 v^2\right).$$

Comme :

$$\Phi_{{}^t u \cdot X}(v) = \mathbb{E}(e^{iv\langle u | X \rangle}) = \mathbb{E}(e^{iv\langle u | X \rangle}),$$

on obtient, en prenant  $v = 1$  :

$$\Phi_X(u) = \mathbb{E}(e^{i\langle u | X \rangle}) = \Phi_{{}^t u \cdot X}(1) = \exp\left(i{}^t u \cdot \mathbb{E}(X) - \frac{1}{2}{}^t u K_X u\right). \quad \square$$

Comme corollaire, on obtient le **très important** résultat suivant.

**Théorème 2.7** Si  $X = (X_1, \dots, X_d)$  est un **vecteur gaussien**, alors ses composantes sont indépendantes si et seulement si elles sont non corrélées :  $\text{cov}(X_j, X_k) = 0, \forall j \neq k$ , c'est-à-dire si et seulement si sa matrice de covariance est diagonale.

Autrement dit, si et seulement si  $X_1 - \mathbb{E}(X_1), \dots, X_d - \mathbb{E}(X_d)$  sont orthogonales dans  $L^2(\mathbb{P})$ .

**Preuve.** Nous savons déjà que des *v.a.r.* indépendantes sont non corrélées. Inversement, si elles ne sont pas corrélées, la *matrice de covariance* de  $X$  est **diagonale** :

$$K_X = \begin{pmatrix} \sigma_1^2 & \dots & 0 \\ & \ddots & \\ 0 & \dots & \sigma_d^2 \end{pmatrix},$$

avec  $\sigma_k^2 = \text{Var}(X_k)$ . Si l'on note  $m_k = \mathbb{E}(X_k)$ , on a  $\mathbb{E}(X) = (m_1, \dots, m_d)$ .

Comme  $X$  est un *vecteur gaussien*, on a, pour  $u = (u_1, \dots, u_d) \in \mathbb{R}^d$  :

$$\begin{aligned} \Phi_X(u) &= \exp\left(i\langle u | \mathbb{E}(X) \rangle - \frac{1}{2}{}^t u K_X u\right) \\ &= \exp\left(i(u_1 m_1 + \dots + u_d m_d) - \frac{1}{2}(\sigma_1^2 u_1^2 + \dots + \sigma_d^2 u_d^2)\right) \\ &= \exp\left(iu_1 m_1 - \frac{1}{2}\sigma_1^2 u_1^2\right) \times \dots \times \exp\left(iu_d m_d - \frac{1}{2}\sigma_d^2 u_d^2\right) \\ &= \Phi_{X_1}(u_1) \cdots \Phi_{X_d}(u_d), \end{aligned}$$



ce qui prouve l'indépendance de  $X_1, \dots, X_d$ .  $\square$

**Remarque.** Il est indispensable de savoir au préalable que le vecteur  $X$  est gaussien.

**Exemple.** Soit  $Y$  une *v.a.r.* de loi  $\mathcal{N}(0, 1)$  et  $\varepsilon$  une *v.a.r.* de Rademacher :  $\mathbb{P}(\varepsilon = -1) = \mathbb{P}(\varepsilon = 1) = \frac{1}{2}$ , indépendante de  $Y$ , et posons  $Z = \varepsilon Y$ . Alors :

a)  $Z \sim \mathcal{N}(0, 1)$ .

En effet, pour tout borélien  $A$  de  $\mathbb{R}$ , on a :

$$\mathbb{P}(Z \in A) = \mathbb{P}(\varepsilon Y \in A) = \mathbb{P}(\varepsilon = 1, Y \in A) + \mathbb{P}(\varepsilon = -1, -Y \in A);$$

mais l'indépendance donne d'une part :

$$\mathbb{P}(\varepsilon = 1, Y \in A) = \mathbb{P}(\varepsilon = 1) \mathbb{P}(Y \in A) = \frac{1}{2} \mathbb{P}(Y \in A),$$

et, d'autre part :

$$\mathbb{P}(\varepsilon = -1, -Y \in A) = \mathbb{P}(\varepsilon = -1) \mathbb{P}(-Y \in A) = \frac{1}{2} \mathbb{P}(-Y \in A) = \frac{1}{2} \mathbb{P}(Y \in A),$$

puisque  $(-Y)$  suit la même loi que  $Y$ . On a donc  $\mathbb{P}(Z \in A) = \mathbb{P}(Y \in A)$ , de sorte que  $Z$  suit la même loi que  $Y$ .

b)  $\text{cov}(Y, Z) = \mathbb{E}(YZ) = \mathbb{E}(Y \cdot \varepsilon Y) = \mathbb{E}(\varepsilon Y^2) \stackrel{\text{II}}{=} \mathbb{E}(\varepsilon) \mathbb{E}(Y^2) = 0 \times 1 = 0$ .

Pourtant, les *v.a.r.*  $Y$  et  $Z$  ne sont pas indépendantes (car sinon  $Y$  serait indépendante de  $Z^2 = (\varepsilon Y)^2 = Y^2$ , ce qui n'est pas). Il en résulte que le vecteur  $X = (Y, Z)$  n'est **pas gaussien**, bien que ses composantes soient des *v.a.r.* gaussiennes.  $\square$

Nous avons vu que toute matrice de covariance est symétrique positive. Inversement, nous allons voir que toute telle matrice est la matrice de covariance d'un vecteur gaussien. C'est la généralisation de ce qui se passe pour la dimension 1, où la moyenne et la variance d'une *v.a.r.* gaussienne peuvent être données arbitrairement. Cela permet aussi d'assurer l'existence de vecteurs gaussiens en dehors du cas où les composantes sont indépendantes.

**Théorème 2.8 (Théorème d'existence)** *Pour tout  $m \in \mathbb{R}^d$  et toute matrice réelle  $K$  carrée d'ordre  $d$ , symétrique positive, il existe un vecteur gaussien de moyenne  $m$  et de matrice de covariance  $K$ .*

**Preuve.** Il suffit de le voir pour  $m = 0$ .

Nous pouvons choisir à notre gré l'espace de probabilité sur lequel la variable aléatoire sera définie. Choisissons  $(\Omega, \mathcal{A}) = (\mathbb{R}^r, \mathcal{Bor}(\mathbb{R}^r))$ , où  $r$  est le rang de la matrice  $K$  (on suppose  $K$  non nulle, car si  $K = 0$ , la variable aléatoire nulle convient), et munissons-le de la probabilité gaussienne  $\mathbb{P}$ , de densité :

$$\gamma(t) = \frac{1}{(\sqrt{2\pi})^r} \exp\left(-\frac{1}{2} \sum_{j=1}^r t_j^2\right) = \frac{1}{(\sqrt{2\pi})^r} \exp\left(-\frac{1}{2} \|t\|_2^2\right).$$

Considérons la variable aléatoire :

$$\begin{aligned} Y_0: \quad \Omega &\longrightarrow \mathbb{R}^r \\ t &\longmapsto t. \end{aligned}$$

C'est évidemment une variable aléatoire gaussienne : sa loi est  $\mathbb{P}_{Y_0} = \mathbb{P} = \gamma \cdot \lambda_d$ , centrée, de matrice de covariance égale à  $I_r$ , la matrice unité d'ordre  $r$ .

Comme  $K$  est symétrique positive, il existe une matrice  $(d \times r)$   $A$  telle que  $K = A \cdot {}^t A$ . Alors  $Y = AY_0 : \Omega \rightarrow \mathbb{R}^d$  est un vecteur gaussien (Proposition 2.2), centré, et sa matrice de covariance est :

$$\begin{aligned} \mathbb{E}(Y \cdot {}^t Y) &= \mathbb{E}((AY_0) \cdot {}^t (AY_0)) = \mathbb{E}(A \cdot Y_0 \cdot {}^t Y_0 \cdot {}^t A) = A \cdot \mathbb{E}(Y_0 \cdot {}^t Y_0) \cdot {}^t A \\ &= A \cdot I_r \cdot {}^t A = A \cdot {}^t A = K. \end{aligned} \quad \square$$

La preuve que l'on vient de faire permet d'obtenir facilement le résultat suivant.

**Corollaire 2.9** *Un vecteur gaussien  $X : \Omega \rightarrow \mathbb{R}^d$  possède une densité, par rapport à la mesure de Lebesgue sur  $\mathbb{R}^d$ , si et seulement si sa matrice de covariance  $K_X$  est inversible. Cette densité est alors :*

$$f_X(x) = \frac{1}{(\sqrt{2\pi})^d} \frac{1}{\sqrt{|\det K_X|}} \exp\left(-\frac{1}{2} {}^t [x - \mathbb{E}(X)] \cdot K_X^{-1} \cdot [x - \mathbb{E}(X)]\right).$$

**Preuve.** On peut d'abord supposer que  $\mathbb{E}(X) = 0$ ; ensuite, comme cet énoncé ne concerne que la loi de la v.a.  $X$ , on peut supposer que  $X$  est donné par l'égalité  $X = AY_0$ , comme dans la preuve du Théorème 2.8, avec  $A \cdot {}^t A = K_X$ .

Lorsque  $K_X$  n'est pas inversible, son rang  $r$  est  $< d$ ; alors  $A$  est de taille  $(d \times r)$ , et  $X = AY_0$  prend ses valeurs, presque sûrement, dans un sous-espace de  $\mathbb{R}^d$  de dimension  $r < d$ ;  $X$  ne peut donc avoir de densité (ce sous-espace est de mesure de Lebesgue nulle, alors que pour  $\mathbb{P}_X$ , il est de mesure 1).

Si  $K_X$  est inversible, son rang est  $r = d$ , et  $A$  est inversible. Comme  $X = AY_0$ , on peut utiliser la formule de changement de variable :

**Théorème 2.10 (formule de changement de variable)** *Soit  $O$  un ouvert de  $\mathbb{R}^d$  et  $O'$  un ouvert de  $\mathbb{R}^{d'}$ , tels qu'il existe un difféomorphisme  $\varphi$  de classe  $\mathcal{C}^1$  de  $O$  sur  $O'$ . Alors, pour toute fonction  $f : O \rightarrow \mathbb{R}$  mesurable positive ou intégrable, on a :*

$$\int_O f(u) du = \int_{O'} f(\varphi^{-1}(v)) |\text{Jac } \varphi^{-1}(v)| dv.$$

via son corollaire :

**Corollaire 2.11** *Sous les hypothèses du Théorème 2.10, soit  $U : \Omega \rightarrow \mathbb{R}^d$  une variable aléatoire prenant presque sûrement ses valeurs dans  $O$ . Alors, si  $U$  a une densité de probabilité  $f_U$ ,  $V = \varphi(U)$  possède aussi une densité de probabilité, donnée par :*

$$f_V(v) = f_U(\varphi^{-1}(v)) |\text{Jac } \varphi^{-1}(v)|.$$

Or ici  $\varphi = A$  est linéaire ; donc :

$$|\text{Jac } A^{-1}(y)| = |\det(A^{-1})| = \frac{1}{|\det A|} = \frac{1}{\sqrt{|\det K_X|}};$$

et par conséquent, puisque  $Y_0$  possède la densité  $f_{Y_0} = \gamma$ ,  $X$  possède une densité, donnée par :

$$f_X(x) = \frac{1}{\sqrt{|\det K_X|}} \gamma(A^{-1}x) = \frac{1}{(\sqrt{2\pi})^d} \frac{1}{\sqrt{|\det K_X|}} \exp\left(-\frac{1}{2} \|A^{-1}x\|_2^2\right),$$

d'où le résultat, puisque :

$$\begin{aligned} \|A^{-1}x\|_2^2 &= {}^t(A^{-1}x).(A^{-1}x) = {}^tx.{}^tA^{-1}.A^{-1}.x = {}^tx.(A.{}^tA)^{-1}.x \\ &= {}^tx.K_X^{-1}.x. \end{aligned} \quad \square$$

**Preuve du Corollaire 2.11.** Pour tout borélien  $B$ , on a :

$$\begin{aligned} \mathbb{P}_V(B) &= \int_{\mathbb{R}^{d'}} \mathbb{1}_B(v) d\mathbb{P}_V(v) = \mathbb{E}(\mathbb{1}_B \circ V) = \mathbb{E}(\mathbb{1}_B \circ \varphi(U)) \\ &= \int_{\mathbb{R}^d} (\mathbb{1}_B \circ \varphi)(u) d\mathbb{P}_U(u) = \int_{\mathbb{R}^d} (\mathbb{1}_B \circ \varphi)(u) f_U(u) du \\ &= \int_{\mathbb{R}^{d'}} \mathbb{1}_B(v) f_U(\varphi^{-1}(v)) |\text{Jac } \varphi^{-1}(v)| dv. \end{aligned} \quad \square$$

## 3 Statistique gaussienne

### 3.1 Résultats théoriques

**Définition 3.1** On dit qu'une suite  $(X_1, \dots, X_n)$  de variables aléatoires **indépendantes** et suivant toutes la **même loi**  $\mathbb{P}_X$ , est un  **$n$ -échantillon** de la loi  $\mathbb{P}_X$ , ou  **$n$ -échantillon** de  $X$ .

On dit aussi que les variables aléatoires  $X_1, \dots, X_n$  sont **indépendantes et identiquement distribuées**, en abrégé : *i.i.d.*.

**Définition 3.2** Soit  $(X_1, \dots, X_n)$  un  $n$ -échantillon d'une v.a.r.  $X \in L^2(\mathbb{P})$ .

1) La v.a.r.

$$\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$$

est appelée la **moyenne empirique** de l'échantillon.

2) La v.a.r.

$$\bar{V}_n = \frac{1}{n} [(X_1 - \bar{X}_n)^2 + \dots + (X_n - \bar{X}_n)^2]$$

est appelée la **variance empirique** de l'échantillon.

Notons que les Statisticiens préfèrent prendre pour variance empirique  $\frac{n}{n-1}\bar{V}_n$ , car son espérance est égale à  $\text{Var}(X)$ ; on dit que c'est un *estimateur sans biais*.

### Définition 3.3

1) On appelle **loi du chi-deux à  $d$  degrés de liberté** la loi de la somme de  $d$  carrés de gaussiennes centrées réduites indépendantes :  $Y_1^2 + \dots + Y_d^2$ . On la note  $\chi_d^2$ .

2) On appelle **loi de Student à  $d$  degrés de liberté** la loi de  $X/\sqrt{K/d}$ , où  $X \sim \mathcal{N}(0, 1)$ ,  $K \sim \chi_d^2$  et  $X \perp\!\!\!\perp K$  ( $X$  indépendante de  $K$ ). On la note  $T_d$ .

Student est le pseudonyme de William Sealey Gosset (1876–1937).

L'intérêt de ces lois vient du résultat suivant.

### Théorème 3.4 (Théorème de Cochran)

Soit  $(X_1, \dots, X_n)$  un  $n$ -échantillon gaussien, de loi  $\mathcal{N}(m, \sigma^2)$ ,  $\sigma > 0$ .

Alors :

1)  $\bar{X}_n \sim \mathcal{N}(m, \frac{\sigma^2}{n})$ ;

2)  $(\frac{n}{\sigma^2})\bar{V}_n \sim \chi_{n-1}^2$ ;

3)  $\bar{X}_n$  et  $\bar{V}_n$  sont indépendantes;

4)  $\frac{\bar{X}_n - m}{\sqrt{\bar{V}_n/(n-1)}} \sim T_{n-1}$ .

**Remarque.** Les points essentiels sont 2) et 3).

On notera que dans le 2), on **perd un degré de liberté**.

Dans le 2), il est sans doute plus naturel d'écrire :

$$\left(\frac{n}{\sigma^2}\right)\bar{V}_n = \frac{1}{\sigma^2} [(X_1 - \bar{X}_n)^2 + \dots + (X_n - \bar{X}_n)^2].$$

**Preuve.** 1) Les *v.a.r.*  $X_1, \dots, X_n$  étant gaussiennes et indépendantes, leur somme est encore gaussienne. Donc  $\bar{X}_n$  est gaussienne; de plus :  $\mathbb{E}(\bar{X}_n) = \frac{1}{n}(\mathbb{E}(X_1) + \dots + \mathbb{E}(X_n)) = \frac{1}{n} \times nm = m$ , et  $\text{Var}(\bar{X}_n) = \frac{1}{n^2}(\text{Var}(X_1) + \dots + \text{Var}(X_n)) = \frac{1}{n^2} \times n\sigma^2 = \frac{\sigma^2}{n}$ .

2) Commençons par nous ramener au cas centré réduit en posant  $Y_k = \frac{X_k - m}{\sigma}$ ;  $Y_k \sim \mathcal{N}(0, 1)$  et :

$$\bar{Y}_n = \frac{1}{n}(Y_1 + \dots + Y_n) = \frac{\bar{X}_n - m}{\sigma};$$

$$\bar{W}_n = \frac{1}{n}[(Y_1 - \bar{Y}_n)^2 + \dots + (Y_n - \bar{Y}_n)^2] = \frac{1}{\sigma^2} \bar{V}_n.$$

Il s'agit donc de montrer que :

$$n\bar{W}_n \sim \chi_{n-1}^2.$$

Notons pour cela  $(e_1, \dots, e_n)$  la base canonique de  $\mathbb{R}^n$ , et choisissons une autre base orthonormée  $(f_1, \dots, f_n)$  de  $\mathbb{R}^n$  en imposant que :

$$f_n = \frac{e_1 + \dots + e_n}{\sqrt{n}}.$$

La matrice de passage a donc la forme suivante :

$$U = \begin{pmatrix} \cdot & \dots & \cdot & 1/\sqrt{n} \\ \vdots & & \vdots & 1/\sqrt{n} \\ \cdot & \dots & \cdot & 1/\sqrt{n} \end{pmatrix}$$

et elle est orthogonale, puisque les bases sont orthonormées.

Les *v.a.r.*  $Y_1, \dots, Y_n$  étant gaussiennes et indépendantes, le vecteur  $Y = (Y_1, \dots, Y_n)$  est gaussien. Il en résulte que le vecteur  ${}^tUY = Z = (Z_1, \dots, Z_n)$  est aussi gaussien. Comme sa matrice de covariance :

$$\begin{aligned} K_Z &= \mathbb{E}(Z \cdot {}^tZ) = \mathbb{E}((UY) \cdot {}^t(UY)) = \mathbb{E}({}^tU \cdot Y \cdot {}^tY \cdot U) \\ &= {}^tU \mathbb{E}(Y \cdot {}^tY) U = {}^tU \cdot K_Y \cdot U = {}^tU \cdot I_n \cdot U = {}^tU \cdot U = I_n \end{aligned}$$

(puisque  $U$  est orthogonale) est l'identité, en particulier diagonale, *les composantes de  $Z$  sont indépendantes* (Théorème 2.7), *et de loi  $\mathcal{N}(0, 1)$ .*

De plus, comme la matrice  $U$  est orthogonale, elle définit une transformation *isométrique* de  $\mathbb{R}^n$  ; donc  $\|Y\|^2 = \|Z\|^2$  (c'est-à-dire que pour tout  $\omega \in \Omega$ , on a  $\|Y(\omega)\|_{\mathbb{R}^n}^2 = \|Z(\omega)\|_{\mathbb{R}^n}^2$ ), soit :

$$Y_1^2 + \dots + Y_n^2 = Z_1^2 + \dots + Z_n^2;$$

d'où, puisque  $Z_n = \frac{Y_1 + \dots + Y_n}{\sqrt{n}} = \sqrt{n} \bar{Y}_n$  :

$$\begin{aligned} n\bar{W}_n &= (Y_1 - \bar{Y}_n)^2 + \dots + (Y_n - \bar{Y}_n)^2 \\ &= Y_1^2 + \dots + Y_n^2 - 2(Y_1 + \dots + Y_n) \bar{Y}_n + n\bar{Y}_n^2 \\ &= Y_1^2 + \dots + Y_n^2 - n\bar{Y}_n^2 = Z_1^2 + \dots + Z_n^2 - Z_n^2 \\ &= Z_1^2 + \dots + Z_{n-1}^2 \sim \chi_{n-1}^2. \end{aligned}$$

3) Comme  $Z_1, \dots, Z_n$  sont indépendantes, les deux *v.a.r.* :

$$\bar{Y}_n = \frac{1}{\sqrt{n}} Z_n \quad \text{et} \quad \bar{W}_n = \frac{1}{n} (Z_1^2 + \dots + Z_{n-1}^2)$$

le sont aussi. Donc  $\bar{X}_n$  et  $\bar{V}_n$  sont aussi indépendantes.

4) Résulte maintenant de ce qui précède et de la définition de la loi de Student.  $\square$

## 3.2 Application pratique

Considérons la situation suivante.

Un produit commercialisé est vendu dans des boîtes portant la mention “contenance 500 grammes”.

En fait, la quantité de produit dans chaque boîte n’est pas rigoureusement constante et dépend des aléas de la fabrication (fiabilité de la machine, entre-autres). La fonction qui à chaque boîte lui associe sa contenance peut donc être considérée comme une *variable aléatoire réelle*  $X$ . La mention de contenance de 500 grammes annoncée par le fabricant est la *moyenne*  $\mathbb{E}(X)$  de cette *v.a.r.* .

Pour contrôler la contenance, on ne pèse pas toutes les boîtes, mais on effectue des **tests**. On choisit donc, *au hasard*, un certain nombre de boîtes dont on pèse le contenu. Si le nombre total de boîtes est très grand par rapport au nombre de boîtes choisies, on peut considérer que la situation est la même à chaque prélèvement, et donc que l’on renouvelle à chaque fois la *même opération* ; autrement dit, au lieu de considérer que l’on a  $n$  valeurs (*occurrences*)  $X(\omega_1), \dots, X(\omega_n)$  de la *v.a.r.*  $X$ , on considère que l’on fait  $n$  fois la même expérience, et donc que l’on a  $n$  *v.a.r.*  $X_1, \dots, X_n$ , ayant toutes la même loi que  $X$ , et dont on mesure une valeur :  $X_1^{obs} = X_1(\omega), \dots, X_n^{obs} = X_n(\omega)$ . Ces *v.a.r.* peuvent de plus être considérées comme *indépendantes*, puisque l’on choisit les boîtes *au hasard*. On a donc un *n-échantillon* de  $X$ .

Nous allons faire (parce c’est plus simple !) l’hypothèse que  $X$  suit une *loi normale*  $\mathcal{N}(m, \sigma^2)$ . Il n’y a évidemment aucune raison que ce soit réellement le cas. Néanmoins, on peut le faire, en théorie, de façon approximative, puisque l’on verra (Théorème Limite Central), qu’en un certain sens, toutes les lois peuvent être approximées par des gaussiennes (on peut, par exemple, prendre pour  $X_1, \dots, X_n$ , non pas directement les valeurs observées, mais des moyennes de valeurs de lots observés). C’est le travail du Statisticien de déterminer la loi suivie, ou, du moins, avec quelle approximation on peut dire que telle loi est suivie.

L’intérêt de la loi normale est qu’elle est très peu dispersée autour de sa moyenne (il y a une très petite probabilité que l’on soit loin de sa moyenne).

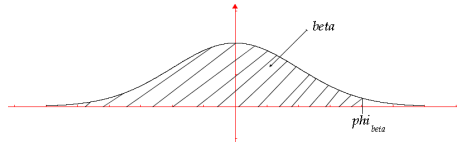
**Définition 3.5** Soit  $Z$  une *v.a.r.*, à valeurs dans  $\mathbb{R}$ . Supposons, pour simplifier, que sa fonction de répartition  $F_Z$  soit continue et strictement croissante. Pour  $0 \leq \beta \leq 1$ , on note  $\varphi_\beta$  l’unique nombre réel vérifiant :

$$F_Z(\varphi_\beta) = \mathbb{P}(Z \leq \varphi_\beta) = \beta$$

(c’est-à-dire que  $\varphi_\beta = F_Z^{-1}(\beta)$ ). On dit que c’est le **quantile d’ordre**  $\beta$ .

Pour  $\beta = 1/2$ ,  $\varphi_{1/2}$  est la **médiane**.  $\varphi_{1/4}$  est le premier quartile et  $\varphi_{3/4}$  le troisième quartile.

On obtient ces valeurs en utilisant des calculatrices, ou des logiciels ; autrefois, mais encore aujourd’hui, on utilisait des tables de valeurs.



Pour  $Z \sim \mathcal{N}(0, 1)$ , les valeurs les plus couramment utilisées sont :

$$\varphi_{0,975} \approx 1,96 \quad \text{et} \quad \varphi_{0,90} \approx 1,282,$$

et surtout :

$$\varphi_{0,95} \approx 1,645.$$

On a  $\varphi_0 = -\infty$  et  $\varphi_1 = +\infty$ .

Comme la densité de  $\mathbb{P}_Z$  est paire ( $Z$  est symétrique), on a :

$$\varphi_{1-\alpha} = -\varphi_\alpha.$$



On a donc :

$$\mathbb{P}(-\varphi_{1-\alpha} \leq Z \leq \varphi_{1-\beta}) = 1 - (\alpha + \beta).$$

On peut bien sûr prendre  $\alpha = \beta$ , mais il prendre  $\alpha \neq \beta$  sera aussi utile.

Soit maintenant un  $n$ -échantillon  $(X_1, \dots, X_n)$  de la loi  $\mathcal{N}(m, \sigma^2)$ .

Comme  $\bar{X}_n \sim \mathcal{N}(m, \frac{\sigma^2}{n})$ , on a :

$$Z = \frac{\bar{X}_n - m}{\sqrt{\frac{\sigma^2}{n}}} \sim \mathcal{N}(0, 1);$$

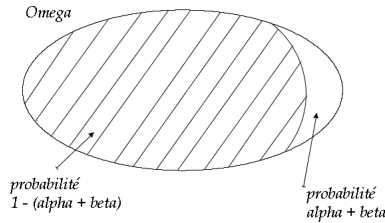
donc :

$$\mathbb{P}\left(-\sqrt{\frac{\sigma^2}{n}} \varphi_{1-\alpha} \leq \bar{X}_n - m \leq \sqrt{\frac{\sigma^2}{n}} \varphi_{1-\beta}\right) = 1 - (\alpha + \beta).$$

Pour chaque valeur de  $\omega \in \Omega$ , l'intervalle :

$$\left[\bar{X}_n(\omega) - \sqrt{\frac{\sigma^2}{n}} \varphi_{1-\alpha}; \bar{X}_n(\omega) + \sqrt{\frac{\sigma^2}{n}} \varphi_{1-\beta}\right]$$

est appelé **intervalle de confiance** pour la moyenne  $m$  **au niveau**  $1 - (\alpha + \beta)$  ( $1 - (\alpha + \beta)$  est le **niveau de confiance**) : il y a une probabilité  $1 - (\alpha + \beta)$  que la moyenne  $m$  soit dans cet intervalle.



Revenons au problème des boîtes de conserve.

Supposons que l'on ait fait  $n = 10$  mesures indépendantes, et que l'on ait observé les valeurs suivantes :

495 ; 497 ; 498 ; 498 ; 498 ; 501 ; 501 ; 501 ; 502 ; 504.

Supposons aussi que l'on sache (d'une façon ou d'une autre, par exemple parce que c'est l'imprécision due à la machine remplissant les boîtes) que l'écart-type est  $\sigma = 1$  gramme (en moyenne les contenances s'écartent de la valeur 500 grammes que de 1 gramme).

On a alors :

$$\bar{X}_n^{obs} = \bar{X}_10(\omega) = 499,5.$$

**Choisissons** (c'est le testeur qui décide de cela) de prendre un **niveau de confiance** de 90% (soit 0,9). On dit aussi que l'on prend un **coefficient de sécurité** de 90%, ou que l'on se laisse une **marge d'erreur** de 10%.

Il y a plusieurs façons de procéder.

1) *Supposons que l'on veuille être impartial.* On prendra alors  $\alpha = \beta = 0,05$ . Comme  $\varphi_{0,95} \approx 1,645$ , on aura, aux approximations près :

$$-\sqrt{\frac{1}{10}} \times 1,645 \leq 499,5 - m \leq \sqrt{\frac{1}{10}} \times 1,645,$$

soit, puisque  $\frac{1,645}{\sqrt{10}} \approx 0,52$  :

$$498,98 \leq m \leq 500,02.$$

Cet intervalle de confiance ne permet pas de conclure si la contenance est respectée ou non.

2) *Supposons que le test soit fait par le fabricant.* Il choisira  $\alpha = 0$  et  $\beta = 0,10$ . Alors, comme  $\varphi_{0,90} \approx 1,282$  et  $\frac{1,282}{\sqrt{10}} \approx 0,41$ , on trouve :

$$m \geq 499,5 - 0,41 = 499,09.$$

Avec cette façon de procéder, on ne peut rien conclure, mais un fabricant peu scrupuleux pourra prétendre qu'il n'y a rien à lui reprocher !

3) *Si le test est fait par une association de consommateurs,* celle-ci prendra  $\beta = 0$  et  $\alpha = 0,10$ . On trouve :

$$m \leq 499,5 + 0,41 = 499,91.$$



On peut donc dire, avec une confiance de 90%, que le fabricant ne respecte pas la contenance.

Nous venons de supposer que l'on connaissait l'écart-type. Ce n'est en général pas le cas. On doit alors utiliser la *variance empirique*. Rappelons que :

$$\frac{\bar{X}_n - m}{\sqrt{\bar{V}_n/(n-1)}} \sim T_{n-1}.$$

En notant  $t_{n-1,\beta}$  les quantiles de cette loi, on a donc :

$$\mathbb{P}\left(-\sqrt{\frac{\bar{V}_n}{n-1}}t_{n-1,\alpha} \leq \bar{X}_n - m \leq \sqrt{\frac{\bar{V}_n}{n-1}}t_{n-1,\beta}\right) = 1 - (\alpha + \beta).$$

Dans notre exemple, on a :

$$\bar{V}_n^{obs} = \bar{V}_n(\omega) = \frac{1}{10} \sum_{k=1}^{10} (X_k(\omega) - \bar{X}_n(\omega))^2 \approx 6,65;$$

l'écart-type observé est donc  $\sigma^{obs} \approx 2,58$ . Comme  $t_{9;0,95} \approx 1,833$ , on a, avec un niveau de confiance de 90% ( $\alpha = \beta = 0,05$ ) :  $497,92 \leq m \leq 501,08$ .

On peut vouloir aussi avoir une *estimation de l'écart-type*  $\sigma$ . Pour cela, on utilise que  $\frac{n}{\sigma^2}\bar{V}_n \sim \chi_{n-1}^2$ . On a donc :

$$\mathbb{P}\left(\frac{n}{\sigma^2}\bar{V}_n \leq \chi_{n-1,\alpha}^2\right) = \alpha.$$

Si l'on veut une estimation supérieure, on écrira :

$$\mathbb{P}\left(\sigma^2 \leq \frac{n\bar{V}_n}{\chi_{n-1,\alpha}^2}\right) = 1 - \alpha.$$

Pour  $\alpha = 0,05$ , on aura, avec une sécurité de 95%, en utilisant la valeur  $\chi_{9;0,05}^2 \approx 3,325$  :

$$\sigma^2 \leq \frac{10 \times 6,65}{3,325} = 20,$$

soit  $\sigma \leq 4,5$ .

Si l'on veut une estimation inférieure, on écrira :

$$\mathbb{P}\left(\sigma^2 \geq \frac{n\bar{V}_n}{\chi_{n-1,\alpha}^2}\right) = \alpha;$$

alors, en prenant  $\alpha = 0,95$ , on a  $\chi_{9;0,95}^2 \approx 16,919$  et, avec une sécurité de 95%, on a :

$$\sigma^2 \geq \frac{10 \times 6,65}{16,919} \approx 3,93,$$

soit  $\sigma \geq 1,98$ .

La différence est importante car le nombre de mesures  $n = 10$  est petit :  $\frac{10}{3,325} \approx 3$  et  $\frac{10}{16,919} \approx 0,6$ . Pour  $n = 30$  :  $\frac{30}{\chi_{29;0,05}^2} \approx \frac{30}{17,708} \approx 1,7$  et  $\frac{30}{\chi_{29;0,95}^2} \approx \frac{30}{42,557} \approx 0,7$ .