# Vecteurs gaussiens

On considère $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

# 1 Introduction

## 1.1 Définitions

Rappelons la définition des variables aléatoires gaussiennes réelles.

**Définition 1**

- *Une variable aléatoire réelle $Z$ est dite* **gaussienne centrée réduite** *si elle admet pour densité par rapport à la mesure de Lebesgue sur $\mathbb{R}$ la fonction :*

$$\boxed{f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)}$$

*On note $Z \rightsquigarrow \mathcal{N}(0,1)$.*

- *Une variable aléatoire réelle $X$ est dite* **gaussienne** *s'il existe $(\mu, \sigma) \in \mathbb{R} \times \mathbb{R}^+$ et $Z \rightsquigarrow \mathcal{N}(0,1)$ tels que $X = \mu + \sigma Z$. La densité de $X$ est alors*

$$\boxed{f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}$$

*On note $X \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$. Quand $\sigma = 0$, on dit que $X$ est une variable gaussienne dégénérée.*

Une variable gaussienne est caractérisée par sa fonction caractéristique, donnée par la proposition suivante :

**Théorème 1**
*La fonction caractéristique de $X \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$ est donnée par*

$$\boxed{\forall t \in \mathbb{R}, \quad \varphi_X(t) = \exp\left(it\mu - \frac{\sigma^2 t^2}{2}\right)}$$

*Preuve :* $\varphi_X$ se calculer à l'aide de $\varphi_Z$ où $Z \rightsquigarrow \mathcal{N}(0,1)$ et on montre que

$$\forall t \in \mathbb{R}, \quad \varphi_Z(t) = -t\varphi'_Z(t)$$

□

**Définition 2**
*Un vecteur aléatoire $X$ à valeurs dans $\mathbb{R}^d$ est dit* **gaussien** *si toute combinaison linéaire de ses composantes est une variable aléatoire gaussienne.*
*Si $X = {}^t(X_1, \ldots, X_d)$ est un vecteur gaussien, on définit son* **vecteur moyenne** $\mathbb{E}(X)$ *par*

$$\mathbb{E}(X) = {}^t\left(\mathbb{E}(X_1), \ldots, \mathbb{E}(X_d)\right)$$

*et sa* **matrice de variance-covariance** $Var(X)$ *par*

$$Var(X) = \mathbb{E}\left( (X - \mathbb{E}(X)) \times {}^t(X - \mathbb{E}(X)) \right)$$

Notons que $Var(X)$ est symétrique et

$$\forall i, j = 1 \dots d, \quad Var(X)_{i,j} = cov(X_i, X_j)$$

**Remarque :** Si $(X_1, \dots, X_n)$ est un $n$-échantillon de loin gaussienne, alors on a évidemment que $X = {}^t(X_1, \dots, X_n)$ est un vecteur gaussien dont la matrice de variance-covariance est proportionnelle à $I_d$.

## 1.2 Propriétés des vecteurs gaussiens

Donnons la fonction caractéristique d'un vecteur gaussien et les conséquences importantes qui en découlent.

**Théorème 2**
*Soit $X = {}^t(X_1, \dots, X_d)$ un vecteur gaussien. On note $m = \mathbb{E}(X)$ et $\Sigma = Var(X)$. On a que $X$ admet pour fonction caractéristique la fonction*

$$\boxed{\forall u \in \mathbb{R}^d, \quad \varphi_X(u) = \mathbb{E}\left[\exp(i\,{}^t uX)\right] = \exp(i\,{}^t um - {}^t u\Sigma u)}$$

*La loi de $X$ est donc entièrement déterminée par $m$ et $\Sigma$. On note $X \rightsquigarrow \mathcal{N}(m, \Sigma)$.*

*Preuve :* Il suffit de remarquer que $\forall u \in \mathbb{R}^d, \quad {}^t uX \rightsquigarrow \mathcal{N}({}^t um, {}^t u\Sigma u)$.

**Corollaire 1 (Propriété de linéarité)**
*Soit $X = {}^t(X_1, \dots, X_d)$ un vecteur gaussien. On note $m = \mathbb{E}(X)$ et $\Sigma = Var(X)$. On a pour toute matrice $A$ possédant $d$ colonnes et pour tout vecteur $b \in \mathbb{R}^d$,*

$$AX + b \ \rightsquigarrow\ \mathcal{N}(Am + b\ ,\ A\Sigma\,{}^t A)$$

**Corollaire 2 (Propriété pour l'indépendance)**
*Soit $X = {}^t(X_1, \dots, X_d)$ un vecteur gaussien. Pour tout $(i, j) \in \{1, \dots, d\}^2$ tel que $i \neq j$, $X_i$ et $X_j$ sont indépendantes si et seulement si $cov(X_i, X_j) = 0$.*

**Remarque :** Les composantes d'un vecteur gaussien sont des variables aléatoires gaussiennes mais la réciproque est fausse. En effet, on considère $X \rightsquigarrow \mathcal{N}(0, 1)$ et $\varepsilon \rightsquigarrow B(0.5)$ indépendante de $X$. Alors $X_1 = X$ et $X_2 = (2\varepsilon - 1)X$ sont des variables gaussiennes mais ${}^t(X_1, X_2)$ n'est pas un vecteur gaussien. Notons que dans cet exemple, $cov(X_1, X_2) = 0$ mais que $X_1$ et $X_2$ ne sont pas indépendantes.

**Proposition 3 (Propriété pour l'espérance conditionnelle)**
*Soit $(Y, X_1, \dots, X_d)$ un vecteur gaussien. Alors $\mathbb{E}(Y \mid X_1, \dots, X_d)$ est une fonction affine de $(X_1, \dots, X_d)$.*

*Preuve :* : Soit $p_F(Y)$ la projection de $Y$ sur $F = Vect(1, X_1, \dots, X_d)$ pour le produit scalaire associé à l'espérance. Donc $\mathbb{E}\left[(Y - p_F(Y))Z\right] = 0$ pour toute variable $Z \in F$. Avec $Z = 1$, on déduit que $\mathbb{E}[Y - p_F(Y)] = 0$. Puis, pour toute variable $Z \in \{X_1, \dots, X_d\}$, le vecteur $(Y - p_F(Y), X_1, \dots, X_d)$ étant gaussien, $0 = \mathbb{E}\left[(Y - p_F(Y))Z\right] = cov(Y - p_F(Y), Z)$ montre que $Y - p_F(Y)$ et $Z$ sont indépendantes. Donc $Y - p_F(Y)$ est indépendante de toute fonction de $(X_1, \dots, X_d)$ et $p_F(Y) = \mathbb{E}(Y \mid X_1, \dots, X_d)$. □

A l'aide de la fonction caractéristique, on démontre le Théorème Centrale Limite Vectoriel.

**Théorème 4 (Théorème Central Limite Vectoriel)**
*Soient $X_1, \ldots, X_n$ des vecteurs aléatoires de $\mathbb{R}^d$ i.i.d. admettant un moment d'ordre 2. On note $m$ leur espérance et $\Gamma$ leur matrice de variance-covariance. Alors,*

$$\boxed{\sqrt{n}\left(\overline{X}_n - m\right) \xrightarrow[n\to+\infty]{loi} \mathcal{N}(0,\Gamma)}$$

*Preuve :* On calcule pour tout $n$ la fonction caractéristique de $Z_n = \sqrt{n}(\overline{X}_n - m)$ :

$$\forall u \in \mathbb{R}^d, \quad \varphi_{Z_n}(u) = \mathbb{E}\left[\exp(i\,{}^t u Z_n)\right]$$

On a par le Théorème Central Limite que ${}^t u Z_n \xrightarrow[n\to+\infty]{\text{loi}} \mathcal{N}(0, {}^t u \Gamma u)$. Donc

$$\forall u \in \mathbb{R}^d, \quad \varphi_{Z_n}(u) \xrightarrow[n\to+\infty]{} \exp\left(-\frac{1}{2}\,{}^t u \Gamma u\right)$$

□

**Théorème 5**
*Soit $X = {}^t(X_1, \ldots, X_d)$ un vecteur gaussien. On note $m = \mathbb{E}(X)$ et $\Sigma = Var(X)$. $X$ admet une densité $f$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$ si et seulement si* $\det(\Sigma) \neq 0$.

- *Si* $\det(\Sigma) = 0$*, la loi de $X - m$ est presque sûrement portée par un espace vectoriel engendré par les vecteurs propres associés aux valeurs propres non nulles de $\Sigma$.*
- *Si* $\det(\Sigma) \neq 0$,

$$\boxed{\forall x \in \mathbb{R}^d, \quad f(x) = \left(\frac{1}{\sqrt{2\pi}}\right)^d \frac{1}{\sqrt{\det(\Sigma)}} \exp\left(-\frac{{}^t(x-m)\Sigma^{-1}(x-m)}{2}\right)}$$

*Preuve :* La matrice $\Sigma$ est symétrique. Donc il existe $U$ une matrice orthogonale (composée des vecteurs propres de $\Sigma$ notés $u_1, u_2, \ldots, u_d$) et il existe $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > 0$ (si $r = rg(\Sigma) \leq d$) tels que

$$\Sigma = U \Gamma\, {}^t U \quad \text{avec} \quad \Gamma = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

Si $\det(\Sigma) = 0$, on a $r < d$. Pour $i \in \{r+1, \ldots, d\}$, $\mathbb{E}\left[({}^t u_i(X-m))^2\right] = {}^t u_i \Sigma u_i = 0$. Donc ${}^t u_i(X-m) = 0$ *p.s.* et $X - m$ prend ses valeurs dans $Vect(u_1, \ldots, u_r)$ qui est de mesure de Lebesgue nulle dans $\mathbb{R}^d$.

Si $\det(\Sigma) \neq 0$, $U\sqrt{\Gamma}$ est inversible. On pose $Y \rightsquigarrow \mathcal{N}(0, I_d)$. Alors $U\sqrt{\Gamma}Y + m \rightsquigarrow X$. Pour toute fonction $g$ continue bornée,

$$\begin{aligned} \mathbb{E}(g(X)) &= \mathbb{E}(g(U\sqrt{\Gamma}Y + m)) \\ &= \int_{\mathbb{R}^d} g(U\sqrt{\Gamma}y + m) \left(\frac{1}{\sqrt{2\pi}}\right)^d \exp\left(-\frac{\|y\|^2}{2}\right) dy \\ &= \int_{\mathbb{R}^d} g(x) \left(\frac{1}{\sqrt{2\pi}}\right)^d \frac{1}{\sqrt{\det(\Sigma)}} \exp\left(-\frac{{}^t(x-m)\Sigma^{-1}(x-m)}{2}\right) dx \end{aligned}$$

$\square$

## 2 Théorème de Cochran, lois du $\chi^2$ et de Student

Dans tout ce paragraphe, nous nous placerons dans $\mathbb{R}^d$ muni du produit scalaire euclidien et on notera $\|.\|$ la norme euclidienne dans $\mathbb{R}^d$.

**Proposition-Définition 6**
*Soit $X$ un vecteur gaussien de $\mathbb{R}^d$ tel que $\mathbb{E}(X) = m$ et $Var(X) = I_d$. La loi de $\|X\|^2$ ne dépend que de $d$ et $\|m\|$. On note*

$$\|X\|^2 \rightsquigarrow \chi^2(d, \|m\|^2)$$

*et on dit que $\|X\|^2$ suit une* **loi du $\chi^2$** *(qui est* **décentrée** *si $\|m\| \neq 0$). L'entier $d$ est le* **nombre de degrés de liberté**, *$\|m\|^2$ est le* **paramètre de décentrage**.
*Lorsque $\|m\| = 0$, on note plus simplement $\|X\|^2 \rightsquigarrow \chi^2(d)$.*

*Preuve :* Soit $Y \in \mathbb{R}^d$ tel que $Y \rightsquigarrow \mathcal{N}(m', I_d)$ avec $\|m\| = \|m'\|$. Il existe $U$ matrice orthogonale telle que $m = Um'$. Donc $UY \rightsquigarrow \mathcal{N}(m, I_d) \rightsquigarrow X$ et

$$\|Y\|^2 = \|UY\|^2 \rightsquigarrow \|X\|^2$$

$\square$

**Proposition 7**
*Si $Z_d \rightsquigarrow \chi^2(d)$, on montre que la densité de $Z_d$ est la fonction $f$ telle que*

$$\boxed{\forall x \in \mathbb{R}, \quad f(x) = \frac{\exp\left(-\dfrac{x}{2}\right) x^{\frac{d}{2}-1}}{2^{\frac{d}{2}}\Gamma\left(\dfrac{d}{2}\right)} \mathbf{1}_{\mathbb{R}^+}(x)}$$

*avec*

$$\forall a > 0, \quad \Gamma(a) = \int_0^{+\infty} e^{-x} x^{a-1} dx$$

*On a :*

$$\boxed{\mathbb{E}(Z_d) = d}, \quad \boxed{Var(Z_d) = 2d}$$

Voici le résultat principal :

**Théorème 8 (de Cochran)**
*Soit $E_1 \oplus \ldots \oplus E_r$ une décomposition de $\mathbb{R}^d$ en sous-espaces deux à deux orthogonaux de dimensions respectives $d_1, \ldots, d_r$. Si $X \rightsquigarrow \mathcal{N}(m, I_d)$, les vecteurs aléatoires $X_{E_1}, \ldots, X_{E_r}$, projections orthogonales de $X$ sur $E_1, \ldots, E_r$ sont indépendants, les variables aléatoires $\|X_{E_1}\|^2, \ldots, \|X_{E_r}\|^2$ sont indépendantes et*

$${}^t\left(\|X_{E_1}\|^2, \ldots, \|X_{E_r}\|^2\right) \quad \rightsquigarrow \quad {}^t\left(\chi^2(d_1, \|m_{E_1}\|^2), \ldots, \chi^2(d_r, \|m_{E_r}\|^2)\right)$$

*où $m_{E_1}, \ldots, m_{E_r}$ sont les projections de $m$ sur $E_1, \ldots, E_r$.*

*Preuve :* Soit $(e_{j1},\dots,e_{jd_j})$ une base orthonormée de $E_j$. On a

$$\forall j=1\dots d,\quad X_{E_j}=\sum_{k=1}^{d_j} e_{jk}\,{}^t e_{jk}X$$

Les variables ${}^t e_{jk}X$ sont indépendantes de la loi $\mathcal{N}({}^t e_{jk}m,1)$ donc les vecteurs aléatoires $X_{E_1},\dots,X_{E,d}$ sont indépendants. Pour achever la preuve, il suffit alors de remarquer que

$$\forall j=1\dots d,\quad \|X_{E_j}\|^2=\sum_{k=1}^{d_j}\left({}^t e_{jk}X\right)^2$$

$\square$

Voici une application importante du théorème de Cochran :

**Proposition 9**
*Soit $X=(X_1,\dots,X_n)$ un $n$-échantillon de loi $\mathcal{N}(\mu,\sigma^2)$. Prenons les estimateurs suivants pour l'estimation de $\mu$ et $\sigma^2$ :*

$$\overline{X}_n=\frac{1}{n}\sum_{i=1}^n X_i,\qquad S_n^2=\frac{1}{n}\sum_{i=1}^n (X_i-\overline{X}_n)^2$$

*Alors, on a :*

- *$\overline{X}_n$ et $S_n^2$ sont des variables aléatoires indépendantes.*
- *Les lois de ces variables sont explicites :*

$$\overline{X}_n \rightsquigarrow \mathcal{N}\left(\mu,\frac{\sigma^2}{n}\right),\qquad \frac{nS_n}{\sigma^2}\rightsquigarrow \chi^2(n-1)$$

*Preuve :* On pose pour tout $i=1\dots n$, $Y_i=\dfrac{X_i-m}{\sigma}$. On a alors que $(Y_1,\dots,Y_n)$ est un $n$-échantillon de loi $\mathcal{N}(0,1)$. On pose ensuite $e={}^t(1,\dots,1)$ et $E=Vect(e)$. On a alors

$$\mathbb{R}^n=E\oplus E^\perp$$

Les projections de $Y={}^t(Y_1,\dots,Y_n)$ sur $E$ et $E^\perp$, $Y_E$ et $Y_{E^\perp}$ sont indépendantes et valent

$$Y_E=\frac{1}{n}\sum_{i=1}^n Y_i\times e,\qquad Y_{E^\perp}=\begin{pmatrix} Y_1-\frac{1}{n}\sum_{i=1}^n Y_i\\ \vdots \\ Y_n-\frac{1}{n}\sum_{i=1}^n Y_i\end{pmatrix}$$

On a

$$\frac{1}{\sigma}(\overline{X}_n-\mu)\times e=Y_E,\qquad \frac{nS_n}{\sigma^2}=\|Y_{E^\perp}\|^2$$

$\square$

Ce résultat nous permet de construire des intervalles de confiance pour l'estimation de $\mu$ et $\sigma^2$ à l'aide de la définition suivante.

**Définition 3**

*Si $X$ et $Y$ sont deux variables aléatoires indépendantes telles que*

- $X \rightsquigarrow \mathcal{N}(\mu, 1)$,
- $Y \rightsquigarrow \chi^2(d)$,

*Alors, la loi de la variable*

$$\boxed{Z = \frac{X}{\sqrt{\frac{Y}{d}}}}$$

*est appelée* **loi de Student** *(***décentrée** *si $\mu \neq 0$)* **à** $d$ **degrés de liberté**. *On note*

$$Z \rightsquigarrow t(d, \mu)$$

*Si le paramètre de décentrage $\mu$ est nul, on note plus simplement*

$$Z \rightsquigarrow t(d)$$

**Proposition 10**

*Si $Z_d \rightsquigarrow t(d)$, on montre que la densité de $Z_d$ est la fonction $f$ telle que*

$$\boxed{\forall x \in \mathbb{R}, \quad f(x) = \frac{\Gamma\left(\frac{d+1}{2}\right)}{\sqrt{d\pi}\Gamma\left(\frac{d}{2}\right)} \left(1 + \frac{x^2}{d}\right)^{-\frac{d+1}{2}}}$$

*avec*

$$\forall a > 0, \quad \Gamma(a) = \int_0^{+\infty} e^{-x} x^{a-1} dx$$

*Pour $d > 1$, on a*

$$\boxed{\mathbb{E}(Z_d) = 0}$$

*Pour $d > 2$, on a*

$$\boxed{Var(Z_d) = \frac{d}{d-2}}$$

*On a également*

$$\boxed{Z_d \xrightarrow[n \to +\infty]{loi} Z \qquad \text{avec } Z \rightsquigarrow \mathcal{N}(0,1)}$$

Comme la loi de Normale, la loi de Student est symétrique mais ses queues sont plus épaisses que celles de la loi normale. On déduit de la définition précédente que

$$\frac{\sigma^{-1}\sqrt{n}(\overline{X}_n - \mu)}{\sigma^{-1}\sqrt{\frac{nS_n}{n-1}}} = \frac{\overline{X}_n - \mu}{\sqrt{\frac{S_n}{n-1}}} \rightsquigarrow t(n-1)$$

En notant $t_{n-1,1-\frac{\alpha}{2}}$ le quantile d'ordre $1 - \frac{\alpha}{2}$ pour la loi $t(n-1)$ et $c_{n-1,1-\alpha}$ le quantile d'ordre $1-\alpha$ pour la loi $\chi^2(n-1)$, un intervalle de confiance de niveau de confiance exactement égal à $1-\alpha$ pour $\mu$ est

$$I_{n,\alpha} = \left[\overline{X}_n - t_{n-1,1-\frac{\alpha}{2}}\sqrt{\frac{S_n}{n-1}}\,,\, \overline{X}_n + t_{n-1,1-\frac{\alpha}{2}}\sqrt{\frac{S_n}{n-1}}\right]$$

et un intervalle de confiance de niveau de confiance exactement égal à $1-\alpha$ pour $\sigma^2$ est :

$$J_{n,\sigma} = \left[\frac{nS_n}{n_{n-1,1-\alpha}}\,,\, +\infty\right[$$

On déduit de ces intervalles de confiance les tests de taille $\alpha$ de $\mu = \mu_0$ conte $\mu \neq \mu_0$ et de $\sigma^2 = \sigma_0^2$ contre $\sigma^2 < \sigma_0^2$.

Remarquons que l'on obtient une région de confiance de niveau de confiance $1-2\alpha$ pour l'estimation de $\theta = (\mu, \sigma^2)$ en considérant $I_{n,\sigma} \times J_{n,\sigma}$.

# 3 Test d'ajustement du $\chi^2$

Dans cette partie, on considère une variable aléatoire discrète $X$ à valeurs dans $\{a_1, \ldots, a_d\}$. On se donne $d$ réels strictement positifs $p_1, \ldots, p_d$ tels que $p_1 + \ldots + p_d = 1$ et on désire tester

$$(H_0): \quad \forall i \in \{1, \ldots, d\}, \quad \mathbb{P}(X = a_i) = p_i$$

contre

$$(H_1): \quad \exists i \in \{1, \ldots, d\}, \quad \mathbb{P}(X = a_i) \neq p_i$$

Pour cela, on dispose d'un $n$-échantillon $(X_1, \ldots, X_n)$ de même loi que $X$. On utilise la méthode des moments pour estimer $p_i$ et on note

$$\forall i \in \{1, \ldots, d\}, \quad N_{ni} = \sum_{j=1}^{n} \mathbf{1}_{X_j = a_i}, \quad \widehat{p_i} = \frac{N_{ni}}{n}$$

Sous $(H_0)$, pour tout $i \in \{1, \ldots d\}$, $\widehat{p_i}$ est un estimateur fortement consistant et sans biais de $p_i$. Donc si $(H_0)$ est vraie, il y a tout lieu de penser que $\widehat{p} = {}^t(\widehat{p_1}, \ldots, \widehat{p_d})$ sera "proche" de $p = {}^t(p_1, \ldots, p_d)$. Comment mesurer la distance entre $\widehat{p}$ et $p$ ? On introduit la ***pseudo-distance du*** $\chi^2$ entre $\widehat{p}$ et $p$ :

$$D_n^2(\widehat{p}, p) = n \sum_{i=1}^{d} \frac{(\widehat{p_i} - p_i)^2}{p_i}$$

Lorsque $n$ est grand, sa limite est connue et surtout indépendante de $p$, ce qui va nous permettre de résoudre notre problème de test. On a en effet le théorème suivant :

**Théorème 11**

- *Sous* $(H_0)$,

$$D_n^2(\widehat{p}, p) \xrightarrow[n \to +\infty]{loi} \chi^2(d-1)$$

- *Sous* $(H_1)$,

$$D_n^2(\widehat{p}, p) \xrightarrow[n \to +\infty]{p.s.} +\infty$$

*Preuve :* On pose $\forall j \in \{1, \ldots n\}, \quad Z_j = {}^t\left(\frac{1}{\sqrt{p_1}}(\mathbf{1}_{X_j = a_i} - p_1), \ \ldots\ , \frac{1}{\sqrt{p_d}}(\mathbf{1}_{X_j = a_i} - p_d)\right)$.

Par le Théorème Central Limite Vectoriel, on a

$$\frac{1}{\sqrt{n}} \left( \sum_{j=1}^{n} Z_j \right) \xrightarrow[n \to +\infty]{\text{loi}} \mathcal{N}(0, I_d - \sqrt{p}\, {}^t\sqrt{p})$$

avec $\sqrt{p} = {}^t(\sqrt{p_1}, \ldots, \sqrt{p_d})$. Donc

$$\sqrt{n}\, {}^t\left( \frac{1}{\sqrt{p_1}}(\widehat{p_1} - p_1), \ \ldots\ , \frac{1}{\sqrt{p_d}}(\widehat{p_d} - p_d) \right) \xrightarrow[n \to +\infty]{\text{loi}} \mathcal{N}(0, I_d - \sqrt{p}\, {}^t\sqrt{p})$$

En utilisant la fonction continue $f$ définie par $\forall x = {}^t(x_1, \ldots, x_d), \quad f(x) = \|x\|^2 = \sum_{j=1}^{d} x_j^2$, on obtient

$$D_n^2(\widehat{p}, p) \xrightarrow[n \to +\infty]{\text{loi}} f(V)$$

où $V$ est une variable aléatoire telle que $V \rightsquigarrow \mathcal{N}(0, I_d - \sqrt{p}\,{}^t\sqrt{p})$ et qui a donc même loi que la projection de $W \rightsquigarrow \mathcal{N}(0, I_d)$ sur $\left(Vect(\sqrt{p})\right)^{\perp}$. Donc

$$f(V) \rightsquigarrow \chi^2(d-1)$$

$\square$

Pour tester $(H_0)$ contre $(H_1)$, on considère donc le test asymptotique de taille $1-\alpha$

$$\varphi(X_1, \dots, X_n) = \mathbf{1}_{D_n^2(\widehat{p}, p) > c_{d-1,1-\alpha}}$$

où $c_{d-1,1-\alpha}$ est le quantile d'ordre $1-\alpha$ de la loi $\chi^2(d-1)$. Remarquons que la puissance du test tend vers 1 quand $n \to +\infty$.

**Remarque 1** : L'approximation par la loi limite est correcte si pour tout $i \in \{1, \dots, d\}$, $np_i \geq 5$. Si ce n'est pas le cas, il faut effectuer un regroupement par classes.

**Remarque 2** : On peut utiliser ce test lorsque la loi de $X$ est continue. SI $X$ est à valeurs dans $\Omega$, on construit une partition finie de $\Omega$ et on applique ce qui précède. Tout le problème porte sur le choix de cette partition.

**Un exemple :** Pour tester sa théorie génétique, Mendel croisa des pois tous jaunes et lisses et obtint à la première génération des pois jaunes ou verts et lisses ou ridés. Plus précisément, il obtint 315 pois jaunes et lisses, 108 pois verts et lisses, 101 pois jaunes et ridés et 32 pois verts et ridés. Est-ce que ces observations confirment ou infirment la théorie mendélienne ?

Sous cette approche, la proportion $p$ de chacune des 4 classes précédentes est $p = {}^t\left(\frac{9}{16}, \frac{3}{16}, \frac{3}{16}, \frac{1}{16}\right)$. On teste donc

$$(H_0): \quad p = {}^t\left(\frac{9}{16}, \frac{3}{16}, \frac{3}{16}, \frac{1}{16}\right)$$

contre

$$(H_1): \quad p \neq {}^t\left(\frac{9}{16}, \frac{3}{16}, \frac{3}{16}, \frac{1}{16}\right)$$

On a $c_{3,0.95} = 7.815$. Comme sous $(H_0)$, $D_{556}^2(\widehat{p}, p) = 0.47$, on accepte $(H_0)$.