MAT7381 Chapitre 2
Loi normale multidimensionnelle

2.1 Moments d'un vecteur aléatoire

Soit $\mathbf{Y} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1p} \\ y_{21} & y_{22} & \cdots & y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{np} \end{bmatrix}$ une matrice aléatoire.

Définition. *L'espérance* $\mathbf{M} = \mathrm{E}(\mathbf{Y})$ d'une matrice aléatoire $\mathbf{Y}$ est la matrice des espérances :

$$\mathbf{M} = \mathrm{E}(\mathbf{Y}) = \begin{bmatrix} \mathrm{E}(y_{11}) & \mathrm{E}(y_{12}) & \cdots & \mathrm{E}(y_{1q}) \\ \mathrm{E}(y_{21}) & \mathrm{E}(y_{22}) & \cdots & \mathrm{E}(y_{2q}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{E}(y_{n1}) & \mathrm{E}(y_{n2}) & \cdots & \mathrm{E}(y_{nq}) \end{bmatrix} = \begin{bmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1q} \\ \mu_{21} & \mu_{22} & \cdots & \mu_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n1} & \mu_{n2} & \cdots & \mu_{nq} \end{bmatrix}$$

Les propriétés suivantes sont aisément démontrées :

Théorème 2.1.1 *Si* $\mathbf{Y}$ *est une matrice aléatoire carrée et* $\mathbf{A}$ *une matrice constante, alors*

$$\mathrm{E}(\mathbf{AY}) = \mathbf{A}\mathrm{E}(\mathbf{Y}).$$

et si $\mathbf{Y}$ *est carrée,*

$$\mathrm{E}[tr(\mathbf{Y})] = tr[\mathrm{E}(\mathbf{Y})]$$

Démonstration Exercice ■

Définition

La *matrice de covariance* $\mathbf{\Sigma} = \mathrm{Var}(\boldsymbol{y})$ d'un *vecteur* aléatoire $\boldsymbol{y} \in \mathbb{R}^n$ d'espérance $\boldsymbol{\mu}$ est une matrice $\mathbf{\Sigma}$ définie par

$$\mathrm{Var}(\boldsymbol{y}) = \mathbf{\Sigma} = \mathrm{E}[(\boldsymbol{y}-\boldsymbol{\mu})(\boldsymbol{y}-\boldsymbol{\mu})'] = \begin{bmatrix} \mathrm{E}(y_1-\mu_1)^2 & \mathrm{E}(y_1-\mu_1)(y_2-\mu_2) & \cdots & \mathrm{E}(y_1-\mu_1)(y_q-\mu_q) \\ \mathrm{E}(y_2-\mu_2)(y_1-\mu_1) & \mathrm{E}(y_2-\mu_2)^2 & \cdots & \mathrm{E}(y_2-\mu_2)(y_q-\mu_q) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{E}(y_q-\mu_q)(y_1-\mu_1) & \mathrm{E}(y_q-\mu_q)(y_2-\mu_2) & \cdots & \mathrm{E}(y_q-\mu_q)^2 \end{bmatrix}$$

Si $\sigma_{ij} = \mathrm{Cov}(y_i\,;\,y_j)$ et $\sigma_{ii} = \sigma_i^2 = \mathrm{Var}(y_i)$, les composantes de la matrice $\mathbf{\Sigma}$ s'écrivent

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix}$$

Remarque. *La matrice* $\mathbf{\Sigma}$ *est nécessairement symétrique.* ■

La covariance entre deux vecteurs aléatoires $\boldsymbol{y}_1$ et $\boldsymbol{y}_2$ de moyennes $\boldsymbol{\mu}_1$ et $\boldsymbol{\mu}_2$ est une matrice définie par

$$\mathrm{Cov}(\boldsymbol{y}_1\,;\,\boldsymbol{y}_2) = \mathrm{E}[(\boldsymbol{y}_1-\boldsymbol{\mu}_1)(\boldsymbol{y}_2-\boldsymbol{\mu}_2)']$$

Si $\boldsymbol{y}' = [\boldsymbol{y}_1' \mid \boldsymbol{y}_2']$ et si $\mathbf{\Sigma} = \mathrm{Var}(\boldsymbol{y})$ est partitionnée de façon conforme

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

alors $\boldsymbol{\Sigma}_{11} = \text{Var}(\boldsymbol{y}_1)$, $\boldsymbol{\Sigma}_{22} = \text{Var}(\boldsymbol{y}_2)$, $\boldsymbol{\Sigma}_{12} = \text{Cov}(\boldsymbol{y}_1 ; \boldsymbol{y}_2)$, $\boldsymbol{\Sigma}_{21} = (\boldsymbol{\Sigma}_{12})' = \text{Cov}(\boldsymbol{y}_2 ; \boldsymbol{y}_1)$.

Théorème 2.1.2. *Soit* $\boldsymbol{y}$ *un vecteur aléatoire* $n \times 1$ *de moyenne* $\boldsymbol{\mu}$ *et de matrice de covariance* $\boldsymbol{\Sigma}$, $\mathbf{L}$ *une matrice* $n \times p$, $\mathbf{M}$ *une matrice* $n \times q$ *et* $\boldsymbol{a}$ *un vecteur* $p \times 1$. *Alors*

$$\text{E}(\mathbf{L}'\boldsymbol{y} + \boldsymbol{a}) = \mathbf{L}'\boldsymbol{\mu} + \boldsymbol{a}, \quad \text{Var}(\mathbf{L}'\boldsymbol{y} + \boldsymbol{a}) = \mathbf{L}'\boldsymbol{\Sigma}\mathbf{L}, \tag{2.1.1}$$

et

$$\text{Cov}(\mathbf{L}'\boldsymbol{y}, \mathbf{M}'\boldsymbol{y}) = \mathbf{L}'\boldsymbol{\Sigma}\mathbf{M} \tag{2.1.2}$$

Il est utile de noter la généralisation suivante de la propriété bien connue $\text{Var}(X) = \text{E}(X^2)-[\text{E}(X)]^2$.

$$\text{Var}(\boldsymbol{y}) = \text{E}(\boldsymbol{y}\boldsymbol{y}') - \boldsymbol{\mu}\boldsymbol{\mu}' \tag{2.1.3}$$

Remarque. *Puisque pour tout* ℓ, $\ell'\boldsymbol{\Sigma}\ell$ *est la variance de* $\ell'\boldsymbol{y}$, *on a que* $\ell'\boldsymbol{\Sigma}\ell \geq 0$ *pour tout* ℓ. *Donc* $\boldsymbol{\Sigma}$ *est semi définie positive. Elle est en fait définie positive, à moins qu'il existe un vecteur* ℓ *tel que* $\ell'\boldsymbol{\Sigma}\ell = \mathbf{0}$, *donc à moins qu'il existe un vecteur* ℓ *tel que* $\ell'\boldsymbol{y}$ *est constante.* ■

Théorème 2.1.3. *Si* $\mathbf{A}$ *est une matrice réelle symétrique,* $\boldsymbol{y}$ *un vecteur aléatoire de moyenne* $\boldsymbol{\mu}$ *et de matrice de covariance* $\boldsymbol{\Sigma}$, *alors*

$$\text{E}(\boldsymbol{y}'\mathbf{A}\boldsymbol{y}) = tr(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} \tag{2.1.4}$$

Démonstration. $\text{E}[\boldsymbol{y}'\mathbf{A}\boldsymbol{y}] = \text{E}[tr(\boldsymbol{y}'\mathbf{A}\boldsymbol{y})] = \text{E}[tr(\mathbf{A}\boldsymbol{y}\boldsymbol{y}')] = tr[\text{E}(\mathbf{A}\boldsymbol{y}\boldsymbol{y}')] = tr[\mathbf{A}\text{E}(\boldsymbol{y}\boldsymbol{y}')] = tr[\mathbf{A}\{\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}'\}] = tr(\mathbf{A}\boldsymbol{\Sigma}) + tr(\mathbf{A}\boldsymbol{\mu}\boldsymbol{\mu}') = tr(\mathbf{A}\boldsymbol{\Sigma}) + tr(\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}) = tr(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$ ■

2.2 La fonction de densité normale multidimensionnelle

Un vecteur aléatoire $\boldsymbol{y}' = (y_1 ; ... ; y_n)$ suit une *loi normale de dimension n* si sa fonction de densité est donnée par :

$$f(y_1, \ldots, y_n) = \frac{1}{(2\pi)^{n/2} |\boldsymbol{\Sigma}|^{1/2}} e^{-(1/2)(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})}$$

Le vecteur aléatoire $\boldsymbol{y}$ est alors de moyenne $\boldsymbol{\mu}$ et de matrice de covariance $\boldsymbol{\Sigma}$, et on écrit $\boldsymbol{y} \sim \mathcal{N}_n(\boldsymbol{\mu} ; \boldsymbol{\Sigma})$.

Lorsque $n = 1$, la fonction f est la fonction de densité normale usuelle,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty$$

de moyenne $\text{E}(y) = \int xf(x)dx = \mu$ et de variance $\text{Var}(y) = \int(x-\mu)^2 f(x)dx = \text{E}(y - \mu)^2 = \sigma^2$.

Remarque. *de la fonction de densité normale multidimensionnelles suppose que la matrice de covariance* $\boldsymbol{\Sigma}$ *est non-singulière (auquel cas on dit que la* distribution *est non-singulière). La normalité multidimensionnelle* peut *se définir sans cette restriction, et les conclusions des théorèmes qui suivent s'appliquent même lorsque* $|\boldsymbol{\Sigma}| = \mathbf{0}$. *Dans ce cours, cependant, nous nous restreindrons aux lois normales non singulières.* ■

Lois marginales et conditionnelles

Théorème 2.2.1. *Soit* $\boldsymbol{y} \sim \mathcal{N}_n(\boldsymbol{\mu}\,;\boldsymbol{\Sigma})$ e*t supposons que* $\boldsymbol{y}$, $\boldsymbol{\mu}$ et $\boldsymbol{\Sigma}$ *sont partitionnés de la façon suivante, où* $\boldsymbol{y}_1$ *est de dimension* n_1 et $\boldsymbol{y}_2$ de dimension $n_2 = n - n_1$.

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \tag{2.2.1}$$

Alors la loi marginale de $\boldsymbol{y}_1$ *est*

$$\boldsymbol{y}_1 \sim \mathcal{N}_{n_1}(\boldsymbol{\mu}_1;\boldsymbol{\Sigma}_{11}) \tag{2.2.2}$$

et la loi conditionnelle de $\boldsymbol{y}_2$ *est normale de moyenne et de matrice de covariances*

$$\mathrm{E}(\boldsymbol{y}_2 \mid \boldsymbol{y}_1) = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{y}_1 - \boldsymbol{\mu}_1), \quad \mathrm{Var}(\boldsymbol{y}_2 \mid \boldsymbol{y}_1) = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{22.1} \tag{2.2.3}$$

Démonstration Considérons la transformation $\boldsymbol{z} = \begin{pmatrix} \boldsymbol{z}_1 \\ \boldsymbol{z}_2 \end{pmatrix} = \mathbf{A}(\boldsymbol{y}-\boldsymbol{\mu})$ où $\mathbf{A} = \begin{bmatrix} \mathbf{I}_{n_1} & 0 \\ -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \mathbf{I}_{n_2} \end{bmatrix}$. La transformation inverse est $(\boldsymbol{y}-\boldsymbol{\mu}) = \mathbf{A}^{-1}\boldsymbol{z} = \begin{bmatrix} \mathbf{I}_{n_1} & 0 \\ \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \mathbf{I}_{n_2} \end{bmatrix}\begin{pmatrix} \boldsymbol{z}_1 \\ \boldsymbol{z}_2 \end{pmatrix}$. Le jacobien est égal à 1 et la fonction de densité de $\boldsymbol{z}$ est

$$g(z_1, \ldots, z_n) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} e^{-(1/2)\left(\mathrm{A}^{-1}\boldsymbol{z}\right)'\Sigma^{-1}\left(\mathrm{A}^{-1}\boldsymbol{z}\right)} = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} e^{-(1/2)\boldsymbol{z}'[\mathrm{A}\Sigma\mathrm{A}']^{-1}\boldsymbol{z}}$$

Mais $[\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}'] = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & 0 \\ 0 & \boldsymbol{\Sigma}_{22.1} \end{bmatrix}$ et $\boldsymbol{z}'[\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}']^{-1}\boldsymbol{z} = \boldsymbol{z}_1'\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{z}_1 + \boldsymbol{z}_2'\boldsymbol{\Sigma}_{22.1}^{-1}\boldsymbol{z}_2$ et $|\boldsymbol{\Sigma}| = |\boldsymbol{\Sigma}_{11}||\boldsymbol{\Sigma}_{22.1}|$. Donc

$$g(z_1, ;\ldots, ; z_n) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}_{11}|^{1/2}|\boldsymbol{\Sigma}_{22.1}|^{1/2}} e^{-(1/2)\boldsymbol{z}_1'\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{z}_1} e^{-(1/2)\boldsymbol{z}_2'\boldsymbol{\Sigma}_{22.1}^{-1}\boldsymbol{z}_2}$$

$$= \frac{1}{(2\pi)^{n_1/2}|\boldsymbol{\Sigma}_{11}|^{1/2}} e^{-(1/2)\boldsymbol{z}_1'\Sigma_{11}^{-1}\boldsymbol{z}_1} \times \frac{1}{(2\pi)^{n_2/2}|\boldsymbol{\Sigma}_{22.1}|^{1/2}} e^{-(1/2)\boldsymbol{z}_2'\Sigma_{22.1}^{-1}\boldsymbol{z}_2}$$

La loi marginale de $\boldsymbol{z}_1$ est $\displaystyle\int \frac{1}{(2\pi)^{n_1/2}|\boldsymbol{\Sigma}_{11}|^{1/2}} e^{-(1/2)\boldsymbol{z}_1'\Sigma_{11}^{-1}\boldsymbol{z}_1} \times \frac{1}{(2\pi)^{n_2/2}|\boldsymbol{\Sigma}_{22.1}|^{1/2}} e^{-(1/2)\boldsymbol{z}_2'\Sigma_{22.1}^{-1}\boldsymbol{z}_2} d\boldsymbol{z}_2$

$$= \frac{1}{(2\pi)^{n_1/2}|\boldsymbol{\Sigma}_{11}|^{1/2}} e^{-(1/2)\boldsymbol{z}_1'\Sigma_{11}^{-1}\boldsymbol{z}_1} \times \int \frac{1}{(2\pi)^{n_2/2}|\boldsymbol{\Sigma}_{22.1}|^{1/2}} e^{-(1/2)\boldsymbol{z}_2'\Sigma_{22.1}^{-1}\boldsymbol{z}_2} d\boldsymbol{z}_2 \ .$$

L'intégrale est celle d'une densité normale multidimensionnelle, et vaut donc 1. La loi marginale de $\boldsymbol{z}_1$ est donc $\dfrac{1}{(2\pi)^{n_1/2}|\boldsymbol{\Sigma}_{11}|^{1/2}} e^{-(1/2)\boldsymbol{z}_1'\Sigma_{11}^{-1}\boldsymbol{z}_1}$. C'est la densité de $\boldsymbol{y}_1 - \boldsymbol{\mu}_1$, une normale de moyenne $\mathbf{0}$ et de matrice de covariance $\boldsymbol{\Sigma}_{11}$. On trouve aisément que la densité de $\boldsymbol{y}_1$ est une normale de moyenne $\boldsymbol{\mu}_1$ et de matrice de covariance $\boldsymbol{\Sigma}_{11}$. On obtient immédiatement, en divisant la densité conjointe par la marginale de $\boldsymbol{z}_1$, que la densité conditionnelle de $\boldsymbol{z}_2$ étant donné $\boldsymbol{z}_1 = y_1 - \mu_1$ est normale de matrice de covariance $\boldsymbol{\Sigma}_{22.1}$. Puisque

$$\boldsymbol{z}_2 = y_2 - \left(\boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(y_1 - \mu_1)\right),$$

nous concluons que l'espérance conditionnelle est $\boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{y}_1 - \boldsymbol{\mu}_1)$ ■

Fonctions linéaires d'un vecteur normal

Théorème 2.2.2. *Soit* $y \sim \mathcal{N}_n(\boldsymbol{\mu}\,;\boldsymbol{\Sigma})$ *et* $\mathbf{A}$ *une matrice* $n \times n$ *non singulière. Soit* $\boldsymbol{z} = \mathbf{A}\boldsymbol{y}$. *Alors*

$$\boldsymbol{z} \sim \mathcal{N}_n(\mathbf{A}\boldsymbol{\mu}\,;\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}').$$

Démonstration. La transformation inverse est $\boldsymbol{y} = \mathbf{A}^{-1}\mathbf{z}$. Le jacobien de la transformation est $\mathbf{A}^{-1}$.

La fonction de densité de z est donc $\dfrac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} e^{-(1/2)(\mathbf{A}^{-1}\mathbf{z}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{A}^{-1}\mathbf{z}-\boldsymbol{\mu})}|\mathbf{A}^{-1}| =$

$$\frac{1}{(2\pi)^{n/2}|\mathbf{A\Sigma A}|^{1/2}} e^{-(1/2)(\mathbf{A}^{-1}\mathbf{z}-\mathbf{A}^{-1}\mathbf{A}\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{A}^{-1}\mathbf{z}-\mathbf{A}^{-1}\mathbf{A}\boldsymbol{\mu})} = \frac{1}{(2\pi)^{n/2}|\mathbf{A\Sigma A}'|^{1/2}} e^{-(1/2)(\mathbf{z}-\mathbf{A}\boldsymbol{\mu})'(\mathbf{A}')^{-1}\boldsymbol{\Sigma}^{-1}\mathbf{A}^{-1}(\mathbf{z}-\mathbf{A}\boldsymbol{\mu})}$$

$= \dfrac{1}{(2\pi)^{n/2}|\mathbf{A\Sigma A}'|^{1/2}} e^{-(1/2)(\mathbf{z}-\mathbf{A}\boldsymbol{\mu})'(\mathbf{A\Sigma A}')^{-1}(\mathbf{z}-\mathbf{A}\boldsymbol{\mu})}$, ce qui est la densité d'une normale $\mathcal{N}_n(\mathbf{A}\boldsymbol{\mu}$; $\mathbf{A\Sigma A}')$ ■

Corollaire *Soit* $\boldsymbol{y} \sim \mathcal{N}_n(\boldsymbol{\mu}\,;\boldsymbol{\Sigma})$ *et* $\mathbf{L}$ *une matrice* $n \times q$ *de plein rang* q . *Soit* $\boldsymbol{z} = \mathbf{L}'\boldsymbol{y}$. *Alors*

$$\boldsymbol{z} \sim \mathcal{N}_q(\mathbf{L}'\boldsymbol{\mu}\,;\mathbf{L}'\boldsymbol{\Sigma}\mathbf{L}).$$

Démonstration Soit $\mathbf{M}$ une matrice $n\times(n\text{-}q)$ telle que $\mathbf{A} = \begin{bmatrix}\mathbf{L}'\\ \mathbf{M}'\end{bmatrix}$ est non singulière. Alors $\mathbf{A}\boldsymbol{y}$ $\sim \mathcal{N}_n(\mathbf{A}\boldsymbol{\mu}\,;\mathbf{A\Sigma A}')$. Mais $\mathbf{A}\boldsymbol{y} = \begin{bmatrix}\mathbf{L}'\boldsymbol{y}\\ \mathbf{M}'\boldsymbol{y}\end{bmatrix} = \begin{bmatrix}\boldsymbol{z}\\ \boldsymbol{w}\end{bmatrix}$, et $\mathbf{A\Sigma A}' = \begin{bmatrix}\mathbf{L}'\boldsymbol{\Sigma}\mathbf{L} & \mathbf{L}'\boldsymbol{\Sigma}\mathbf{M}\\ \mathbf{M}'\boldsymbol{\Sigma} & \mathbf{M}'\boldsymbol{\Sigma}\mathbf{M}\end{bmatrix}$. Le résultat découle du théorème 2.2.1. ■

Remarque. *Si* $\boldsymbol{\Sigma}$ *est non singulière et* $\mathbf{L}_{n\times q}$ *est de rang q, alors* $\boldsymbol{z}$ *a une distribution non-singulière.* ■

Indépendance de variables normales

La covariance entre deux variables aléatoires indépendantes est nulle, mais deux variables dont la covariance est nulle ne sont pas nécessairement indépendantes. Cependant, si X et Y suivent conjointement une loi normale, alors l'implication inverse (covariance nulle $\Rightarrow$ indépendance) *est* vraie. Voici l'énoncé général.

Théorème 2.2.3 *Soit* $\boldsymbol{y} \sim \mathcal{N}_n(\boldsymbol{\mu}\,;\boldsymbol{\Sigma})$ *et supposons que* $\boldsymbol{y}$, $\boldsymbol{\mu}$ *et* $\boldsymbol{\Sigma}$ *sont partitionnés* :

$$\boldsymbol{y} = \begin{pmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\end{pmatrix},\ \boldsymbol{\mu} = \begin{pmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{pmatrix}\ \ \boldsymbol{\Sigma} = \begin{pmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{pmatrix} \qquad (1.2.1)$$

où $\boldsymbol{y}_1$ *et* $\boldsymbol{y}_2$ *sont de dimensions* n_1 *et* $n_2 = n\text{-}n_1$, *respectivement. Alors les vecteurs* $\boldsymbol{y}_1$ *et* $\boldsymbol{y}_2$ *sont stochastiquement indépendants si et seulement si* $\boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21} = \mathbf{0}$.

Démonstration On sait déjà que la condition est nécessaire. Pour montrer qu'elle est suffisante il suffit de constater que la loi conditionnelle de $\boldsymbol{y}_2$ est normale de moyenne et de matrice de covariances $\mathrm{E}(\boldsymbol{y}_2\,|\,\boldsymbol{y}_1) = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{y}_1 - \boldsymbol{\mu}_1) = \boldsymbol{\mu}_2$ et de variance $\mathrm{Var}(\boldsymbol{y}_2\,|\,\boldsymbol{y}_1) =$ $\boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{22}$, donc indépendante *de* $\boldsymbol{y}_1$. ■

Théorème 2.2.4 *Deux fonctions linéaires* $\mathbf{L}_1'\boldsymbol{y}$ *et* $\mathbf{L}_2'\boldsymbol{y}$ *d'un vecteur* $\boldsymbol{y} \sim \mathcal{N}_n(\boldsymbol{\mu}\,;\,\boldsymbol{\Sigma})$ *sont indépendantes si et seulement si* $\mathbf{L}_1'\boldsymbol{\Sigma}\mathbf{L}_2 = \mathbf{0}$.

Démonstration Soit $\mathbf{L} = (\mathbf{L}_1\,;\,\mathbf{L}_2)$ une matrice partitionnée $q\times p$. Nous supposerons que $\mathbf{L}$ est de rang plein. Selon le corollaire du théorème 1.2.2, le vecteur $\mathbf{L}'\boldsymbol{y} \sim \mathcal{N}(\mathbf{L}'\boldsymbol{\mu}\,;\,\mathbf{L}'\boldsymbol{\Sigma}\mathbf{L})$. Mais

$$\mathbf{L}'\boldsymbol{y} = \begin{pmatrix} \mathbf{L}_1'\boldsymbol{y} \\ \mathbf{L}_2'\boldsymbol{y} \end{pmatrix}, \text{ et } \boldsymbol{\Sigma} = \begin{pmatrix} \mathbf{L}_1'\boldsymbol{\Sigma}\mathbf{L}_1 & \mathbf{L}_1'\boldsymbol{\Sigma}\mathbf{L}_2 \\ \mathbf{L}_2'\boldsymbol{\Sigma}\mathbf{L}_1 & \mathbf{L}_2'\boldsymbol{\Sigma}\mathbf{L}_2 \end{pmatrix} \qquad (2.2.2)$$

Le résultat découle du théorème 1.2.3 ■

Transformations orthogonales

Si les composantes $y_1, \ldots, y_n$ d'un vecteur $\boldsymbol{y} = (y_1, \ldots, y_n)'$ sont indépendantes, les composantes $z_1, \ldots, z_m$ du vecteur $\boldsymbol{z} = \mathbf{L}'\boldsymbol{y}$, fonction linéaire des y_i, ne le sont généralement plus. Quelles sont les transformations linéaires qui préservent l'indépendance ? Ce sont les transformations *orthogonales*.

Théorème 2.2.5 Soit $\boldsymbol{y} \sim \mathcal{N}_n(\boldsymbol{\mu}\,;\,\sigma^2\mathbf{I}_n)$, $\mathbf{P}$ *une matrice* $n \times m$ $(m \le n)$ *dont les colonnes sont orthonormales* : $\mathbf{P}'\mathbf{P} = \mathbf{I}_m$. Soit $\boldsymbol{z} = \mathbf{P}'\boldsymbol{y}$. Alors $\boldsymbol{z} \sim \mathcal{N}_m(\mathbf{P}'\boldsymbol{\mu}\,;\,\sigma^2\mathbf{I}_m)$.

Démonstration Découle directement du théorème 2.1.3. ■

Quelles sont les transformations qui transforment des variables dépendantes en variables indépendantes ? Considérons un vecteur $\boldsymbol{y} \sim \mathcal{N}_n(\boldsymbol{\mu}\,;\,\boldsymbol{\Sigma})$. La matrice $\boldsymbol{\Sigma}$, étant symétrique réelle peut être diagonalisée par une matrice orthogonale $\mathbf{P}$: $\mathbf{P}'\boldsymbol{\Sigma}\mathbf{P} = \mathbf{D}$ (ou $\boldsymbol{\Sigma} = \mathbf{PDP}'$) pour une certaine matrice diagonale $\mathbf{D}$ dont les composantes de la diagonale sont toutes non nulles. $\mathbf{D}$ a donc une « racine carrée » $\mathbf{D}^{1/2}$. On a alors une transformation qui réduit le vecteur $\boldsymbol{y}$ en un vecteur $\boldsymbol{z}$ dont les composante sont indépendantes et de variance 1.

Théorème 2.2.6 Si $\boldsymbol{z} = \mathbf{P}'\boldsymbol{y}$, alors $\boldsymbol{z} \sim \mathcal{N}_n(\mathbf{P}'\boldsymbol{\mu}\,;\,\mathbf{D})$, et donc les composantes de $\boldsymbol{z}$ sont indépendantes. Et si $\boldsymbol{u} = \mathbf{D}^{-1/2}\mathbf{P}'\boldsymbol{y}$, alors $\boldsymbol{u} \sim \mathcal{N}_n(\mathbf{D}^{-1/2}\mathbf{P}'\boldsymbol{\mu}\,;\,\mathbf{I}_n)$.

Démonstration Exercice. ■

Remarque *La norme d'un vecteur normal*

Plusieurs décisions en statistique sont basées sur l'importance d'un écart entre deux quantités, en particulier l'écart entre une observation y et sa moyenne μ. *Lorsque ces variables sont scalaires, l'écart est naturellement défini par le carré de la différence entre y et* μ, $(y-\mu)^2$. *Lorsqu'il s'agit de* vecteurs *aléatoire, la distance entre* $\boldsymbol{y}$ *sa moyenne* $\boldsymbol{\mu}$ *est le carré* $\|\boldsymbol{y} - \boldsymbol{\mu}\|^2$ *de la norme* $\|\boldsymbol{y} - \boldsymbol{\mu}\|$ *du vecteur* $\boldsymbol{y} - \boldsymbol{\mu}$. *Une norme n'étant pas unique, il est nécessaire de s'interroger sur ce qui constitue un choix raisonnable. Supposons pour commencer que les composantes du vecteur* $\begin{bmatrix} y_1-\mu_1 \\ y_2-\mu_2 \\ \vdots \\ y_p-\mu_p \end{bmatrix}$ *sont indépendantes. La norme classique* $\|\boldsymbol{y} - \boldsymbol{\mu}\| = \sqrt{\sum_{i=1}^p (y_i-\mu_i)^2}$ *n'est généralement pas satisfaisante car elle accorde le même poids à chaque écart* $y_i - \mu_i$, *ce qui n'est presque jamais raisonnable. Supposons, par exemple, que les* μ_i *sont les moyennes de certaines mesures crâniennes et les* $w_i = y_i-\mu_i$ *sont les écarts entre les mesures d'un individu et la moyenne* μ_i. *Si* μ_1 *est la largeur moyenne des orbites des yeux (de l'ordre de* 2,4 cm, *environ) et* μ_2 *est le diamètre frontal maximum (environ* 12 cm), *il est clair que les* *écarts $w_1$et w_2, exprimés en centimètres, ne sont pas comparables. Il convient de les exprimer en* nombre d'écarts-types*, c'est-à-dire, de mesurer les écarts par leur cote Z,* $z_i = \dfrac{y_i - \mu_i}{\sigma_i}$ *. La norme de* $\boldsymbol{y}$-$\boldsymbol{\mu}$ *serait alors définie par* $\|\boldsymbol{y} - \boldsymbol{\mu}\| = \sqrt{\sum_{i=1}^{p}\left(\dfrac{y_i - \mu_i}{\sigma_i}\right)^2}$. Le carré de la norme, en langage matriciel, s'exprime comme $(\boldsymbol{y}-\boldsymbol{\mu})'\mathbf{D}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})$, *où* $\mathbf{D} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_p^2 \end{bmatrix}$ *est la matrice de covariance de* $\boldsymbol{y}$.

Cette norme est satisfaisante lorsque les variables sont indépendantes, mais elle fait défaut lorsqu'elles ne le sont pas. Par exemple, supposons que $\boldsymbol{z}_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ *et* $\boldsymbol{z}_2 = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$ *sont les cotes Z correspondant à la taille et le poids de deux individus. La somme des carrés est* 8 *dans les deux cas. Ces deux individus sont-ils vraiment aussi éloignés de la moyenne l'un que l'autre ? Pas vraiment : le deuxième l'est beaucoup plus, pour la raisons suivante : une taille supérieure à la moyenne, devrait normalement être accompagnée d'un poids supérieur à la moyenne. C'est ce qui fait que le premier individu est plus normal, moins excentrique, contrairement au deuxième, qui est beaucoup moins corpulent qu'il ne devrait l'être, compte tenu de sa taille.*

La mesure $(\boldsymbol{y}-\boldsymbol{\mu})'\mathbf{D}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})$ *proposée plus haut suggère que dans le cas d'un vecteur de loi* $\mathcal{N}_p(\boldsymbol{\mu}\,;\,\boldsymbol{\Sigma})$*, la norme appropriée est* $\sqrt{(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})}$ *. Cette norme est telle que la fonction de densité* $\mathcal{N}_p(\boldsymbol{\mu}\,;\,\boldsymbol{\Sigma})$ *est constante sur les ellipsoïdes* $(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}) = k$*, ce qui veut dire que deux vecteurs* $\boldsymbol{y}_1$-$\boldsymbol{\mu}$ et $\boldsymbol{y}_2$-$\boldsymbol{\mu}$ *sont considérés comme de même longueur si la densité aux points* $\boldsymbol{y}_1$ *et* $\boldsymbol{y}_2$ *est la même. La mesure* $(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})$ *est appelée* distance de Mahalanobis. ■

2.3 Loi khi-deux (centrale) (χ^2)

Définition La loi du χ^2 est la distribution d'une somme de carrés de variables aléatoires normales centrées réduites indépendantes. Soit $z_1, z_2, \ldots, z_\nu$ des variables indépendantes et identiquement distribuées (*i.d.d.*) selon la loi $\mathcal{N}(0\,;\,1)$ et

$$X = z_1^2 + z_2^2 + \ldots + z_\nu^2 .$$

Alors $X \sim \chi^2_\nu$, c'est-à-dire, X est distribuée selon la loi khi-deux avec ν degrés de liberté .

Fonction de densité La fonction de densité d'une variable de loi χ^2_ν est

$$\frac{1}{\Gamma(\nu/2)\,2^{\nu/2}}\, x^{\nu/2-1} e^{-x/2}, \; x > 0 \tag{2.3.1}$$

Lorsque $y \sim \chi^2_\nu$, on a $E(y) = \nu$ et $\mathrm{Var}(y) = 2\nu$.

Remarque *La loi khi-deux est un cas particulier de la loi gamma dont la fonction de densité est* $\dfrac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x^2/2}$*,* $x > 0$*. La loi khi-deux est donc une loi gamma de paramètres* $\alpha = \nu/2$ et $\beta = 2$.

La propriété $\int_0^\infty x^{\alpha-1} e^{-x^2/2} dx = \Gamma(\alpha)\beta^\alpha$ *qui en découle se révèle souvent utile pour éviter de calculer des intégrales.* ■

Fonction génératrice des moments La fonction génératrice des moments d'une variable de loi χ^2_ν est

$$\mathrm{M}(t) = (1 - 2t)^{-\nu/2} \quad (t < 1/2). \tag{2.3.2}$$

Démonstration

$\mathrm{M}(t) = \mathrm{E}(e^{tX}) = \dfrac{1}{\Gamma(\nu/2)2^{\nu/2}}\displaystyle\int_0^\infty e^{tx}x^{\nu/2}e^{-x/2}dx = \dfrac{1}{\Gamma(\nu/2)2^{\nu/2}}\displaystyle\int_0^\infty x^{\nu/2-1}e^{-(1/2)(1-2t)x}dx =$

$\dfrac{\Gamma(\nu/2)[2/(1-2t)]^{\nu/2}}{\Gamma(\nu/2)2^{\nu/2}} = (1\text{-}2t)^{-\nu/2}$, le résultat voulu.

Si $\boldsymbol{y}$ est un vecteur normal, il suffit que les covariances entre les composantes de $\boldsymbol{y}$ soient nulles pour que celles-ci soient indépendantes. Nous avons donc,

$$\text{Si } \boldsymbol{y} \sim \mathcal{N}_n(\mathbf{0}\,;\,\mathbf{I}_n), \text{ alors } \boldsymbol{y}'\boldsymbol{y} \sim \chi^2_n \tag{2.3.3}$$

Additivité et soustractivité de variables de loi khi-deux

Les théorèmes suivants caractérisent les situations où une somme de variables de loi χ^2 est de loi χ^2.

Théorème 2.3.1 (*Additivité*)

Si $X \sim \chi^2_r$, $Y \sim \chi^2_s$ et si X et Y sont indépendantes, alors $Z = X + Y \sim \chi^2_{r+s}$.

Démonstration. Si M_X et M_Y sont les fonctions génératrices des moments (*fgm*) de X et de Y, respectivement, alors la *fgm* M_Z de Z est, grâce à l'indépendance,

$$M_Z(t) = M_X(t)M_Y(t) = (1\text{-}2t)^{-r/2}(1\text{-}2t)^{-s/2} = (1\text{-}2t)^{-(r+s)/2}$$

ce qui est la *fgm* d'une loi χ^2_{r+s} ■

Théorème 2.3.2. (*Soustractivité*) *Soit $Z = X + Y$* où *$Z \sim \chi^2_n$, $X \sim \chi^2_r$, $n > r$, et X, Y sont indépendantes*.

Alors $Y \sim \chi^2_{n-r}$.

Démonstration. Soit $\mathrm{M}_Z(t)$, $\mathrm{M}_X(t)$ et $\mathrm{M}_Y(t)$ les *fgm* de Z, X et Y, respectivement. Puisque Z est une somme de variables aléatoires indépendantes, sa *fgm* est le produit des *fgm* de ses facteurs : $M_Z(t) = \mathrm{M}_X(t)\mathrm{M}_Y(t)$. Puisque $\mathrm{M}_Z(t) = (1\text{-}2t)^{-n/2}$, $\mathrm{M}_X(t) = (1\text{-}2t)^{-r/2}$, nous avons $\mathrm{M}_Y(t) = (1\text{-}2t)^{-(n-r)/2}$ ce qui est la *fgm* d'une χ^2_{n-r}. ■

***Formes quadratiques générales* ($\boldsymbol{y}'\mathbf{A}\boldsymbol{y}$)**

Théorème 2.3.3 *Soit $\boldsymbol{y} \sim \mathcal{N}_n(\mathbf{0}\,;\,\mathbf{I}_n)$, $\mathbf{A}$ une matrice réelle symétrique de rang r. Alors*

(*i*) Si $\mathbf{A}^2 = \mathbf{A}$, alors $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} \sim \chi^2_r$ où $r = r(\mathbf{A})$;

(*ii*) Si $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} \sim \chi^2_s$, alors $\mathbf{A}^2 = \mathbf{A}$, $s = r(\mathbf{A})$.

Démonstration.

(*i*) *Démonstration* de i) Supposons que $\mathbf{A}^2 = \mathbf{A}$. Alors il existe une matrice $\mathbf{P}_{n\times r}$ telle que $\mathbf{P}'\mathbf{P} = \mathbf{I}_r$ et $\mathbf{A} = \mathbf{P}\mathbf{P}'$. Alors $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} = \boldsymbol{y}'\mathbf{P}\mathbf{P}'\boldsymbol{y} = (\mathbf{P}'\boldsymbol{y})'(\mathbf{P}'\boldsymbol{y}) = \boldsymbol{z}'\boldsymbol{z}$, où $\boldsymbol{z} = \mathbf{P}'\boldsymbol{y} \sim \mathcal{N}_r(\mathbf{0}\,;\,\mathbf{I}_r)$.

Donc $\boldsymbol{z}'\boldsymbol{z}$ suit une loi χ^2_r.

(ii) *Démonstration* de ii) Supposons donc que $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} \sim \chi^2_s$. Puisque **A** est réelle et symétrique, il existe une matrice $n \times r$ **P** telle que $\mathbf{A} = \mathbf{PDP}'$, où **D** est une matrice $r \times r$ diagonale dont les éléments de la diagonale sont les valeurs propres $\lambda_1, \ldots, \lambda_r$ non nulles de **A** et $\mathbf{P}'\mathbf{P} = \mathbf{I}_r$.

Donc $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} = \boldsymbol{y}'\mathbf{PDP}'\boldsymbol{y} = \boldsymbol{z}'\mathbf{D}\boldsymbol{z} = \Sigma\lambda_i z_i^2$ où $\mathbf{z} = \mathbf{P}'\boldsymbol{y} \sim \mathcal{N}_r(\mathbf{0} ; \mathbf{I}_r)$. Soit $M(t)$ la *fgm* de $\boldsymbol{y}'\mathbf{A}\boldsymbol{y}$.

D'une part, puisque $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} \sim \chi^2_s$, on a

$$\mathrm{M}(t) = (1\text{-}2t)^{-s/2}.$$

D'autre part, $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} = \Sigma\lambda_i z_i^2$ où $z_i^2 \sim \chi^2_1$ de *fgm* $M_i(t) = (1\text{-}2t)^{-½}$ et les z_i^2 sont indépendantes. Puisque la *fgm* de $\lambda_i z_i^2$ est $M_i(\lambda_i t)$, on a

$$M(t) = \prod_{i=1}^{r} M_i(\lambda_i t) = \prod_{i=1}^{r}(1-2\lambda_i t)^{-1/2}$$

Il suit

$$(1\text{-}2t)^{-s/2} = \prod_{i=1}^{r}(1-2\lambda_i t)^{-1/2} \Rightarrow (1\text{-}2t)^{s} = \prod_{i=1}^{r}(1-2\lambda_i t)$$

d'où $\lambda_i = 1$, $i = 1, \ldots, r$ et $r = s$. Il suit $\mathbf{D} = \mathbf{I}_r$, $\mathbf{A} = \mathbf{PP}'$ et alors $\mathbf{A}^2 = \mathbf{PP}'\mathbf{PP}' = \mathbf{A}$. ■

2.4 Loi khi-deux non centrale

Si $\boldsymbol{y} \sim \mathcal{N}_n(\boldsymbol{\mu} ; \mathbf{I}_n)$, alors $X = \boldsymbol{y}'\boldsymbol{y}$ suit une loi appelée loi khi-deux à n degrés de liberté et paramètre de non centralité $\lambda = \boldsymbol{\mu}'\boldsymbol{\mu}$: $X = \boldsymbol{y}'\boldsymbol{y} \sim \chi^2_n(\lambda)$. Lorsque le paramètre de non centralité est nul, la loi khi-deux $\chi^2_n(0)$ est en fait une khi-deux centrale ordinaire.

Théorème 2.4.1

La fonction génératrice des moments d'une variable $X = \sum_{i=1}^n X_i = \sum_{i=1}^n y_i^2$, *où les* y_i *sont indépendantes,* $y_i \sim \mathcal{N}(\mu_i ; 1)$ *est*

$$M_n(t) = e^{t\lambda/(1-2t)}(1-2t)^{-n/2}$$

Démonstration

Considérons d'abord la fonction génératrice des moments $M_W(t)$ d'une variable $W = y^2$, où $y \sim \mathcal{N}(\mu ; 1)$:

$$\mathrm{M}_W(t) = \mathrm{E}(e^{tW}) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{ty^2} e^{-(1/2)(y-\mu)^2} dy = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-(1/2)[(1-2t)y^2-2\mu y+\mu^2]} dy$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-(1/2)[(1-2t)y^2-2\mu y]-\mu^2/2} dy$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\mu^2/2+\mu^2/[2(1-2t)]}\int_{-\infty}^{\infty} e^{-(1/2)(1-2t)[y^2-2\mu y/(1-2t)+\mu^2/(1-2t)^2]} dy =$$

$$e^{-(\mu^2/2)[1-(1-2t)^{-1}]}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-(1/2)(1-2t)[y-\mu/(1-2t)]^2} dy = e^{t\mu^2/(1-2t)}(1-2t)^{-1/2}.$$

La fonction génératrice des moments de X est donc $\prod_{i=1}^n M_{X_i}(t) =$

$$\prod_{i=1}^{n} e^{t\mu_i^2/(1-2t)}(1-2t)^{-1/2} = (1-2t)^{-n/2} e^{t\sum_{i=1}^{n}\mu_i^2/(1-2t)} = e^{t\lambda/(1-2t)}(1-2t)^{-n/2}.$$ ■

Théorème 2.4.2

La fonction de densité d'une variable X de loi $\chi_n^2(\lambda)$ *est donnée par*

$$f_n(x) = \sum_{j=0}^{\infty} \frac{e^{-\lambda/2}(\lambda/2)^j}{j!} f_{n+2j}(x)$$

où f_{n+2j} *désigne la fonction de densité d'une variable de loi* $\chi_{n+2j}^2(0)$ *.*

Démonstration

La fonction génératrice des moments de X peut s'écrire comme

$M_X(t) = e^{t\lambda/(1-2t)}(1-2t)^{-n/2} =$

$e^{-\lambda/2}\left(e^{(\lambda/2)/(1-2t)}\right)(1-2t)^{-n/2} = e^{-\lambda/2}\left[\sum_{j=0}^{\infty}\frac{1}{j!}\left(\frac{\lambda/2}{1-2t}\right)^j\right](1-2t)^{-n/2} =$

$\sum_{j=0}^{\infty}\frac{e^{-\lambda/2}(\lambda/2)^j}{j!}(1-2t)^{-(n/2+j)}$. Or chaque terme de cette dernière somme est bien celle de la fonction génératrice des moments d'une variable de densité $f_{n+2j}(x)$. ■

Remarque *La densité d'une loi khi-deux non-centrale est donc une série de densités khi-deux centrales dont les coefficients* $\frac{e^{-\lambda/2}(\lambda/2)^j}{j!}$ *sont curieusement les valeurs de la fonction de probabilité d'une variable de loi de Poisson de paramètre* $\lambda/2$ ■.

Formes quadratiques générales **(*y*'A*y*)**

Théorème 2.4.3 *Soit* $\boldsymbol{y} \sim \mathcal{N}_n(\boldsymbol{\mu}\,;\,\mathbf{I}_n)$, $\mathbf{A}$ *une matrice réelle symétrique de rang r. Alors*

(*i*) *Si* $\mathbf{A}^2 = \mathbf{A}$, *alors* $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} \sim \chi_r^2(\lambda)$ *où* $r = r(\mathbf{A})$ *et* $\lambda = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$;

(*ii*) *Si* $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} \sim \chi_s^2(\lambda)$, *alors* $\mathbf{A}^2 = \mathbf{A}$, $s = r(\mathbf{A})$ *et* $\lambda = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$.

Démonstration.

(*i*) *Démonstration* de (*i*) Supposons que $\mathbf{A}^2 = \mathbf{A}$. Alors il existe une matrice $\mathbf{P}_{n\times r}$ telle que $\mathbf{P}'\mathbf{P} = \mathbf{I}_r$ et $\mathbf{A} = \mathbf{P}\mathbf{P}'$. Alors $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} = \boldsymbol{y}'\mathbf{P}\mathbf{P}'\boldsymbol{y} = (\mathbf{P}'\boldsymbol{y})'(\mathbf{P}'\boldsymbol{y}) = \boldsymbol{z}'\boldsymbol{z}$, où $\boldsymbol{z} = \mathbf{P}'\boldsymbol{y} \sim \mathcal{N}_r(\mathbf{P}'\boldsymbol{\mu}\,;\,\mathbf{I}_r)$. Donc $\boldsymbol{z}'\boldsymbol{z}$ suit une loi $\chi_r^2(\lambda)$ avec $\lambda = (\mathbf{P}'\boldsymbol{\mu})'(\mathbf{P}'\boldsymbol{\mu}) = \boldsymbol{\mu}'\mathbf{P}\mathbf{P}'\boldsymbol{\mu} = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$.

(*ii*) *Démonstration* de (*ii*) Omise ■

Remarque *Certains auteurs définissent le paramètre de non centralité par* $\lambda = \boldsymbol{\mu}'\boldsymbol{\mu}/2$, *ce qui simplifie légèrement l'expression des fonctions de densité et génératrice des moments.* ■

2.5 Indépendance de formes quadratiques

Quand est-ce que deux formes quadratiques $\boldsymbol{y}'\mathbf{A}\boldsymbol{y}$ et $\boldsymbol{y}'\mathbf{B}\boldsymbol{y}$ de loi khi-deux sont indépendantes ? Il est nécessaire de le savoir afin d'en déduire la distribution de leur somme ou de leur quotient. Le théorème suivant donne une réponse pour des variables de matrice de covariance $\mathbf{I}$. Nous démontrerons le cas général plus loin.

Théorème 2.5.1 *Soit* $\boldsymbol{y} \sim \mathcal{N}_n(\boldsymbol{\mu}\ ;\ \mathbf{I}_n)$, **A** *et* **B** *deux matrices symétriques. Soit* $\boldsymbol{y}'\mathbf{A}\boldsymbol{y}$ et $\boldsymbol{y}'\mathbf{B}\boldsymbol{y}$ *deux formes quadratiques de loi khi-deux* (*centrale ou non*). Alors $\boldsymbol{y}'\mathbf{A}\boldsymbol{y}$ et $\boldsymbol{y}'\mathbf{B}\boldsymbol{y}$ *sont indépendantes si et seulement si* $\mathbf{AB} = \mathbf{0}$.

Démonstration

i) Supposons que $\boldsymbol{y}'\mathbf{A}\boldsymbol{y}$ et $\boldsymbol{y}'\mathbf{B}\boldsymbol{y}$ sont des khi-deux indépendantes. Alors **A** et **B** sont idempotentes et $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} + \boldsymbol{y}'\mathbf{B}\boldsymbol{y} = \boldsymbol{y}'(\mathbf{A}+\mathbf{B})\boldsymbol{y}$ est de loi khi-deux. Donc **A**+**B** est idempotente : $(\mathbf{A}+\mathbf{B})^2 = \mathbf{A}^2 + \mathbf{B}^2 + \mathbf{AB} + \mathbf{BA} = \mathbf{A} + \mathbf{B} \Rightarrow \mathbf{A} + \mathbf{B} + \mathbf{AB} + \mathbf{BA} = \mathbf{A} + \mathbf{B} \Rightarrow \mathbf{AB} + \mathbf{BA} = \mathbf{0}$ $\Rightarrow \mathbf{AB} = -\mathbf{BA} \Rightarrow \mathbf{ABB} = -\mathbf{BAB} \Rightarrow \mathbf{AB} = -\mathbf{BAB}$, ce qui entraîne que **AB** est symétrique. Donc $\mathbf{AB} = (\mathbf{AB})' = \mathbf{B}'\mathbf{A}' = \mathbf{BA}$. Alors $\mathbf{AB} + \mathbf{BA} = 0 \Rightarrow \mathbf{AB} + \mathbf{AB} = 0 \Rightarrow \mathbf{AB} = \mathbf{0}$. C'est ce qu'il fallait démontrer.

ii) Supposons que $\mathbf{AB} = \mathbf{0}$. Puisque $\boldsymbol{y}'\mathbf{A}\boldsymbol{y}$ et $\boldsymbol{y}'\mathbf{B}\boldsymbol{y}$ sont de loi khi-deux, A et B sont idempotentes. Donc il existe des matrices **P** et **Q** telles que $\mathbf{A} = \mathbf{PP}'$, $\mathbf{B} = \mathbf{QQ}'$, et $\mathbf{P}'\mathbf{P}$ et $\mathbf{Q}'\mathbf{Q}$ sont des matrices identité. Si $\mathbf{AB} = \mathbf{0}$, alors $\mathbf{PP}'\mathbf{QQ}' = \mathbf{0} \Rightarrow \mathbf{P}'\mathbf{PP}'\mathbf{QQ}'\mathbf{Q} = \mathbf{0} \Rightarrow$ $\mathbf{IP}'\mathbf{QI} = \mathbf{0} \Rightarrow \mathbf{P}'\mathbf{Q} = \mathbf{0}$, ce qui permet de conclure que $\mathbf{P}'\boldsymbol{y}$ et $\mathbf{Q}'\boldsymbol{y}$ sont indépendantes. Pour montrer que $\boldsymbol{y}'\mathbf{A}\boldsymbol{y}$ et $\boldsymbol{y}'\mathbf{B}\boldsymbol{y}$ le sont,il suffit de montrer que $\boldsymbol{y}'\mathbf{A}\boldsymbol{y}$ est fonction de $\mathbf{P}'\boldsymbol{y}$ et que $\boldsymbol{y}'\mathbf{B}\boldsymbol{y}$ est fonction de $\mathbf{Q}'\boldsymbol{y}$. On a
$\boldsymbol{y}'\mathbf{A}\boldsymbol{y} = \boldsymbol{y}'\mathbf{PP}'\boldsymbol{y} = (\mathbf{P}'\boldsymbol{y})'(\mathbf{P}'\boldsymbol{y})$; et $\boldsymbol{y}'\mathbf{B}\boldsymbol{y} = \boldsymbol{y}'\mathbf{QQ}'\boldsymbol{y} = (\mathbf{Q}'\boldsymbol{y})'(\mathbf{Q}'\boldsymbol{y})$ est fonction de $\mathbf{Q}'\boldsymbol{y}$, ce qui conclut la démonstration. ■

Nous allons maintenant généraliser les deux théorèmes précédents : nous considérerons le cas où la matrice de covariance est une matrice définie positive quelconque **Σ** .

Théorème 2.5.2 *Soit* $\boldsymbol{y} \sim \mathcal{N}_n(\boldsymbol{\mu}\ ;\ \boldsymbol{\Sigma})$, **A** *une matrice réelle symétrique de rang r. Alors*

i) *Si* $\mathbf{A\Sigma A} = \mathbf{A}$, *alors* $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} \sim \chi^2_r(\lambda)$ *où* $r = r(\mathbf{A})$ *et* $\lambda = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$;

ii) *Si* $\boldsymbol{y}'\boldsymbol{\Sigma}\boldsymbol{y} \sim \chi^2_s(\lambda)$, *alors* $\mathbf{A\Sigma A} = \mathbf{A}$, $s = r(\mathbf{A})$ *et* $\lambda = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$.

Démonstration Puisque **Σ** est définie positive, il existe une matrice symétrique non singulière **E** telle que $\boldsymbol{\Sigma} = \mathbf{E}^2$. Alors $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} = \boldsymbol{y}'\mathbf{E}^{-1}\mathbf{EAEE}^{-1}\boldsymbol{y} = \boldsymbol{z}'\mathbf{EAE}\boldsymbol{z}$, où $\boldsymbol{z} = \mathbf{E}^{-1}\boldsymbol{y} \sim \mathcal{N}(\mathbf{E}^{-1}\boldsymbol{\mu}\ ;\ \mathbf{I}\ )$. Nous pouvons alors appliquer le théorème 3 à la forme quadratique $\boldsymbol{z}'\mathbf{EAE}\boldsymbol{z}$. Dans la partie (*i*) du théorème, la condition d'idempotence devient $\mathbf{EAEEAE} = \mathbf{EAE}$, ce qui, étant donné que **E** est non singulière est équivalent à $\mathbf{AEEA} = \mathbf{A}$, c'est à dire $\mathbf{A\Sigma A} = \mathbf{A}$. Si cette condition est vérifiée, alors d'après le théorème 3, $\boldsymbol{z}'\mathbf{EAE}\boldsymbol{z} \sim \chi^2_n(\lambda)$ où $r = r(\mathbf{EAE}) = r(\mathbf{A})$ grâce à la non singularité de **E** ; et $\lambda = (\mathbf{E}^{-1}\boldsymbol{\mu})'\mathbf{EAE}(\mathbf{E}^{-1}\boldsymbol{\mu}) = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$.

Réciproquement, si $\boldsymbol{z}'\mathbf{EAE}\boldsymbol{z} \sim \chi^2_s(\lambda)$, alors, par le théorème 3, $\mathbf{EAEEAE} = \mathbf{EAE} \Leftrightarrow \mathbf{AEEA} = \mathbf{A}$ (par la non singularité de **E**), ce qui est équivalent à $\mathbf{A\Sigma A} = \mathbf{A}$; $s = r(\mathbf{EAE}) = r(\mathbf{A})$; et $\lambda = (\mathbf{E}^{-1}\boldsymbol{\mu})'\mathbf{EAE}(\mathbf{E}^{-1}\boldsymbol{\mu}) = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$. ■

Et voici la généralisation de la deuxième partie du théorème 2.5.1 concernant l'indépendance de formes quadratiques.

Théorème 2.5.3 *Soit* $\boldsymbol{y} \sim \mathcal{N}_n(\boldsymbol{\mu}\ ;\ \boldsymbol{\Sigma})$, **A** *et* **B** *deux matrices symétriques. Soit* $\boldsymbol{y}'\mathbf{A}\boldsymbol{y}$ *et* $\boldsymbol{y}'\mathbf{B}\boldsymbol{y}$ *deux formes quadratiques de loi khi-deux* (*centrale ou non*). *Alors* $\boldsymbol{y}'\mathbf{A}\boldsymbol{y}$ *et* $\boldsymbol{y}'\mathbf{B}\boldsymbol{y}$ *sont indépendantes si et seulement si* $\mathbf{A\Sigma B} = \mathbf{0}$.

Démonstration Puisque $\mathbf{\Sigma}$ est définie positive, il existe une matrice symétrique non singulière $\mathbf{E}$ telle que $\mathbf{\Sigma = E}^2$. Alors $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} = \boldsymbol{y}'\mathbf{E}^{-1}\mathbf{EAEE}^{-1}\boldsymbol{y} = \boldsymbol{z}'\mathbf{EAE}\boldsymbol{z}$, et $\boldsymbol{y}'\mathbf{B}\boldsymbol{y} = \boldsymbol{y}'\mathbf{E}^{-1}\mathbf{EBEE}^{-1}\boldsymbol{y} = \boldsymbol{z}'\mathbf{EBE}\boldsymbol{z}$ où $\boldsymbol{z} = \mathbf{E}^{-1}\boldsymbol{y} \sim \mathcal{N}(\mathbf{E}^{-1}\boldsymbol{\mu} ; \mathbf{I})$. D'après le théorème 3.3.1, $\boldsymbol{z}'\mathbf{EAE}\boldsymbol{z}$, et $\boldsymbol{z}'\mathbf{EBE}\boldsymbol{z}$ sont indépendantes si et seulement si $\mathbf{EAEEBE = 0}$. Mais étant donnée la non singularité de $\mathbf{E}$, ceci entraîne que $\mathbf{AEEB = 0}$, c'est-à-dire, que $\mathbf{A\Sigma B = 0}$. ■

Et voici un théorème général concernant la décomposition d'une forme quadratique en une somme de formes quadratiques indépendantes. Le théorème est connu sous le nom de *théorème de Cochran.*

Théorème 2.5.4 *Soit* $\boldsymbol{y} \sim \mathcal{N}_n(\mathbf{0} ; \mathbf{I}_n)$. *Soit* $\mathbf{A}$ *une matrice idempotente et symétrique de rang r, et* $\mathbf{A} = \mathbf{A}_1 + ... + \mathbf{A}_k$, $r(\mathbf{A}_i) = r_i$, $Q = \boldsymbol{y}'\mathbf{A}\boldsymbol{y}$, $Q_i = \boldsymbol{y}'\mathbf{A}_i\boldsymbol{y}$, $i = 1, ... , k$. *Alors les trois théorèmes suivants sont équivalents :*

i) $r(\mathbf{A}) = \sum_{i=1}^{k} r(\mathbf{A}_i)$; ii) $\mathbf{A}_i^2 = \mathbf{A}_i$ et $\mathbf{A}_i\mathbf{A}_j = \mathbf{0}$; iii) *Les* Q_i *sont indépendantes et* $Q_i \sim \chi^2_{r_i}$.

Démonstration. Omise. ■

Si $\mathbf{\Sigma}$ n'est que semi définie positive et non définie positive, nous avons la généralisation suivante :

Théorème 2.5.5. *Soit* $\boldsymbol{y} \sim \mathcal{N}_n(\mathbf{0} ; \mathbf{\Sigma})$ *où* $\mathbf{\Sigma}$ *est semi définie positive. La forme quadratique* $\boldsymbol{y}'\mathbf{A}\boldsymbol{y}$ *suit une loi* χ^2 *si et seulement si* $\mathbf{\Sigma A \Sigma A \Sigma = \Sigma A \Sigma}$, *auquel cas, le nombre de degrés de liberté est* $r = tr(\mathbf{A\Sigma})$.

Esquisse de démonstration. $\mathbf{\Sigma}$ peut s'écrire comme $\mathbf{\Sigma = PDP}' = [\mathbf{P}_1 \mid \mathbf{P}_2]\begin{bmatrix}\mathbf{D} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\mathbf{P}_1'\\ \mathbf{P}_2'\end{bmatrix} = \mathbf{P}_1\mathbf{D}\mathbf{P}_1'$, où $\mathbf{P}_1$ est $n\times r$, $r = r(\mathbf{\Sigma})$. $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} = \boldsymbol{y}'\mathbf{PP}'\mathbf{APP}'\boldsymbol{y} = \boldsymbol{z}\mathbf{P}'\mathbf{AP}\boldsymbol{z}$, $\boldsymbol{z} = \mathbf{P}'\boldsymbol{y} = \begin{bmatrix}\mathbf{P}_1'\boldsymbol{y}\\ \mathbf{0}\end{bmatrix} = \begin{bmatrix}\boldsymbol{z}_1\\ \mathbf{0}\end{bmatrix}$, *où* $\boldsymbol{z}_1$ est de moyenne nulle et de matrice de covariance $\mathbf{D}$. Alors $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} = \boldsymbol{z}_1'\mathbf{P}_1'\mathbf{AP}_1\boldsymbol{z}_1$ est de loi χ^2_r si et seulement si
$\mathbf{P}_1'\mathbf{AP}_1\mathbf{DP}_1'\mathbf{AP}_1 = \mathbf{P}_1'\mathbf{AP}_1 \Leftrightarrow \mathbf{P}_1'\mathbf{A\Sigma AP}_1 = \mathbf{P}_1'\mathbf{AP}_1 \Leftrightarrow \mathbf{P}_1\mathbf{DP}_1'\mathbf{A\Sigma AP}_1\mathbf{DP}_1' = \mathbf{P}_1\mathbf{DP}_1'\mathbf{AP}_1\mathbf{DP}_1' \Leftrightarrow \mathbf{\Sigma A \Sigma A \Sigma = \Sigma A \Sigma}$ ■

Loi de Student (loi t)

Définition Soit X et Y deux variables indépendantes, $Y \sim \mathcal{N}(0 ; 1)$ et $X \sim \chi^2_\nu(\lambda)$. Alors la variable

$$t = \frac{Y}{\sqrt{X/\nu}}$$

suit une loi appelée *loi de Student* à ν degrés de liberté et paramètre de non centralité λ, $t \sim t_\nu(\lambda)$. Lorsque $\lambda = 0$, la loi est dite *centrale*. On la désignera par t_ν (0) ou plus simplement par t_ν ■

Loi de Fisher (loi $\mathcal{F}$)

Définition Soit $X_1 \sim \chi^2_{\nu_1}$ et $X_2 \sim \chi^2_{\nu_2}(\lambda)$, X_1 et X_2 indépendantes, alors la variable

$$F = \frac{X_1/\nu_1}{X_2/\nu_2}$$

suit une loi appelée *loi de Fisher*, ou loi $\mathcal{F}$, à ν_1 et ν_2 degrés de liberté et paramètre de non centralité λ : $F \sim \mathcal{F}_{\nu_1;\nu_2}(\lambda)$. Lorsque $\lambda = 0$, la loi est dite *centrale*. On la désignera par $\mathcal{F}_{\nu_1;\nu_2}$ (0)

ou plus simplement par $\mathcal{F}_{\nu_1;\nu_2}$ ■

Remarques

1 Si $F \sim F_{\nu_1;\nu_2}(0)$, alors $\dfrac{1}{F} \sim \mathcal{F}_{\nu_2;\nu_1}(0)$.

2 Si $t \sim t_\nu$, alors $t^2 \sim \mathcal{F}_{1\,;\nu}$.

2.6 Échantillons normaux multidimensionnels

Soit $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n$ n vecteurs aléatoires indépendants chacun de loi $\mathcal{N}_p(\boldsymbol{\mu}\,;\boldsymbol{\Sigma})$, où $\boldsymbol{\mu}$ est un *vecteur* $p\times 1$ et Σ est une matrice définie positive $p\times p$. La densité conjointe de ces vecteurs est

$f(\boldsymbol{y}_1\,;\boldsymbol{y}_2\,;\ldots;\boldsymbol{y}_n\,;\boldsymbol{\mu}\,;\boldsymbol{\Sigma}) =$

$$\prod_{i=1}^n \frac{1}{(2\pi)^{1/2}\,|\boldsymbol{\Sigma}|^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{y}_i-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu})\right\} = \frac{1}{(2\pi)^{n/2}\,|\boldsymbol{\Sigma}|^{n/2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^n(\boldsymbol{y}_i-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu})\right\}$$

Soit $\bar{\boldsymbol{y}} = \frac{1}{n}\sum_{i=1}^n \boldsymbol{y}_i$ le vecteur moyenne des observations et $\mathbf{S} = \frac{1}{n}\sum_{i=1}^n(\boldsymbol{y}_i-\bar{\boldsymbol{y}})(\boldsymbol{y}_i-\bar{\boldsymbol{y}})'$ la matrice de covariance échantillonnale. On peut réécrire la somme dans l'exposant de la fonction de densité comme $\sum_{i=1}^n(\boldsymbol{y}_i-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}) = \sum_{i=1}^n[(\boldsymbol{y}_i-\bar{\boldsymbol{y}})+(\bar{\boldsymbol{y}}-\boldsymbol{\mu})]'\boldsymbol{\Sigma}^{-1}[(\boldsymbol{y}_i-\bar{\boldsymbol{y}})+(\bar{\boldsymbol{y}}-\boldsymbol{\mu})]$

$$= \sum_{i=1}^n(\boldsymbol{y}_i-\bar{\boldsymbol{y}})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_i-\bar{\boldsymbol{y}}) + n(\bar{\boldsymbol{y}}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\boldsymbol{y}}-\boldsymbol{\mu}) + 2\sum_{i=1}^n(\bar{\boldsymbol{y}}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_i-\bar{\boldsymbol{y}})$$

$$= \sum_{i=1}^n(\boldsymbol{y}_i-\bar{\boldsymbol{y}})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_i-\bar{\boldsymbol{y}}) + n(\bar{\boldsymbol{y}}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\boldsymbol{y}}-\boldsymbol{\mu}) + 2(\bar{\boldsymbol{y}}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}\sum_{i=1}^n(\boldsymbol{y}_i-\bar{\boldsymbol{y}})$$

$$= \sum_{i=1}^n(\boldsymbol{y}_i-\bar{\boldsymbol{y}})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_i-\bar{\boldsymbol{y}}) + n(\bar{\boldsymbol{y}}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\boldsymbol{y}}-\boldsymbol{\mu}) = \sum_{i=1}^n tr(\boldsymbol{y}_i-\bar{\boldsymbol{y}})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_i-\bar{\boldsymbol{y}}) + n(\bar{\boldsymbol{y}}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\boldsymbol{y}}-\boldsymbol{\mu})$$

$$= tr\sum_{i=1}^n\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_i-\bar{\boldsymbol{y}})(\boldsymbol{y}_i-\bar{\boldsymbol{y}})' + n(\bar{\boldsymbol{y}}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\boldsymbol{y}}-\boldsymbol{\mu}) = tr\boldsymbol{\Sigma}^{-1}\sum_{i=1}^n(\boldsymbol{y}_i-\bar{\boldsymbol{y}})(\boldsymbol{y}_i-\bar{\boldsymbol{y}})' + n(\bar{\boldsymbol{y}}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\boldsymbol{y}}-\boldsymbol{\mu})$$

$$= n\,tr\boldsymbol{\Sigma}^{-1}\mathbf{S} + n(\bar{\boldsymbol{y}}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\boldsymbol{y}}-\boldsymbol{\mu})$$

Ainsi donc, la fonction de vraisemblance peut s'écrire comme

$$L(\boldsymbol{\mu}\,;\boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{n/2}\,|\boldsymbol{\Sigma}|^{n/2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^n(\boldsymbol{y}_i-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu})\right\}$$
$$= \frac{1}{(2\pi)^{n/2}\,|\boldsymbol{\Sigma}|^{n/2}}\exp\left\{-\frac{n}{2}tr\boldsymbol{\Sigma}^{-1}\mathbf{S} - \frac{1}{2}n(\bar{\boldsymbol{y}}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\boldsymbol{y}}-\boldsymbol{\mu})\right\}.$$

Estimateurs des paramètres

Les valeurs de $\boldsymbol{\mu}$ et $\boldsymbol{\Sigma}$ qui maximisent la fonction de vraisemblance $L(\boldsymbol{\mu}\,;\boldsymbol{\Sigma})$ sont $\boldsymbol{\mu} = \bar{\boldsymbol{y}}$ et $\boldsymbol{\Sigma} = \mathbf{S}$. Pour montrer que $\bar{\boldsymbol{y}}$ maximise L, il suffit de noter que, étant donné que $n(\bar{\boldsymbol{y}}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\boldsymbol{y}}-\boldsymbol{\mu}) \geq 0$, $\exp\left\{-\frac{1}{2}\sum_{i=1}^n(\bar{\boldsymbol{y}}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\boldsymbol{y}}-\boldsymbol{\mu})\right\}$ est inférieur ou égal à 1 pour tout $\boldsymbol{\mu}$, et atteint son maximum de 0 à $\boldsymbol{\mu} = \bar{\boldsymbol{y}}$. Alors $L(\boldsymbol{\mu};\boldsymbol{\Sigma}) \leq L(\bar{\boldsymbol{y}};\boldsymbol{\Sigma})$ pour tout $\boldsymbol{\Sigma}$, et il reste à maximiser ceci par rapport à $\boldsymbol{\Sigma}$.

Nous montrons maintenant que $L(\bar{\boldsymbol{y}};\mathbf{S}) \geq L(\bar{\boldsymbol{y}};\boldsymbol{\Sigma})$ pour tous $\boldsymbol{\Sigma}$ définie positive. Ce qui est équivalent à montrer que $\ln L(\bar{\boldsymbol{y}};\mathbf{S}) - \ln L(\bar{\boldsymbol{y}};\boldsymbol{\Sigma}) \geq 0$.

$$\begin{aligned}\ln L(\bar{y};\mathbf{S}) - \ln L(\bar{y};\boldsymbol{\Sigma}) &= -\frac{n}{2}\ln|\mathbf{S}| - \frac{n}{2}tr\mathbf{S}^{-1}\mathbf{S} + \frac{n}{2}\ln|\boldsymbol{\Sigma}| + \frac{n}{2}tr\boldsymbol{\Sigma}^{-1}\mathbf{S}\\ &= -\frac{n}{2}\ln|\boldsymbol{\Sigma}^{-1}\mathbf{S}| + \frac{n}{2}tr\boldsymbol{\Sigma}^{-1}\mathbf{S} - \frac{np}{2}\\ &= -\frac{n}{2}\ln\prod_{i=1}^{n}\lambda_i + \frac{n}{2}\sum_{i=1}^{p}\lambda_i - \frac{np}{2},\end{aligned}$$

où les λ_i sont les valeurs propres de $\boldsymbol{\Sigma}^{-1}\mathbf{S}$.

Il faut donc montrer que $-\frac{n}{2}\ln\prod_{i=1}^{n}\lambda_i + \frac{n}{2}\sum_{i=1}^{p}\lambda_i - \frac{np}{2} \geq 0 \Leftrightarrow \sum_{i=1}^{p}(\lambda_i - \ln\lambda_i) \geq p$.

Les λ_i sont positifs car **S** est définie positive (avec probabilité 1) ; donc elle possède une « racine carrés » $\mathbf{S}^{1/2}$ et les valeurs propres de $\boldsymbol{\Sigma}^{-1}\mathbf{S}$ sont les mêmes que celles de $\mathbf{S}^{1/2}\boldsymbol{\Sigma}^{-1}\mathbf{S}^{1/2}$, qui est définie positive. Pour montrer que $\sum_{i=1}^{p}(\lambda_i - \ln\lambda_i) \geq p$, il suffit de vérifier que λ_i - ln $\lambda_i \geq 1$, ce qui découle du fait que la dérivée de λ_i - ln λ_i est négative dans (0 ; 1) et positive à droite de 1. Donc son minimum est 1 au point $\lambda_i = 1$, ce qui conclut la démonstration. ■

Distribution des estimateurs

Il est assez facile de démontrer que $\bar{\boldsymbol{y}} \sim \mathcal{N}_p(\boldsymbol{\mu} ; (1/n)\boldsymbol{\Sigma})$, utilisant les fait qu'une fonction linéaire de normales indépendantes est normale. La distribution de la matrice $\mathbf{A} = \sum_{i=1}^{n}(\boldsymbol{y}_i - \bar{\boldsymbol{y}})(\boldsymbol{y}_i - \bar{\boldsymbol{y}})' = n\mathbf{S}$ exige quelques développements.

Quelques propriétés des échantillons normaux multidimensionnels

- Soit $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ n vecteurs aléatoires indépendants, $\boldsymbol{y}_i \sim \mathcal{N}_p(\boldsymbol{\mu}_i ; \boldsymbol{\Sigma})$. Posons et $\mathbf{Y}' = [\boldsymbol{y}_1 | \boldsymbol{y}_2 | \ldots | \boldsymbol{y}_n]$, et $\mathbf{M}' = [\boldsymbol{\mu}_1 | \boldsymbol{\mu}_2 | \ldots | \boldsymbol{\mu}_n]$.
- Soit $\boldsymbol{a}' = [a_1 ;\ a_2 ;\ \ldots ;\ a_n]$ un vecteur constant et $\boldsymbol{\ell} = \sum_{i=1}^{n} a_i\boldsymbol{y}_i = \mathbf{Y}'\boldsymbol{a}$ une combinaisons linéaire des vecteurs. Alors $\boldsymbol{\ell} \sim \mathcal{N}_p(\mathbf{M}'\boldsymbol{a} ;\ \boldsymbol{a}'\boldsymbol{a}\boldsymbol{\Sigma})$.
- Si $\boldsymbol{a}$ et $\boldsymbol{b}$ sont deux vecteurs fixes tels que $\boldsymbol{a}'\boldsymbol{b} = 0$, $\boldsymbol{a}'\boldsymbol{a} = \boldsymbol{b}'\boldsymbol{b} = 1$, alors $\mathbf{Y}'\boldsymbol{a}$ et $\mathbf{Y}'\boldsymbol{b}$ sont des vecteurs normaux indépendants de matrice de covariance $\boldsymbol{\Sigma}$. Plus généralement, soit **P** est une matrice $n\times m$, $m \leq n$ telle que $\mathbf{P}'\mathbf{P} = \mathbf{I}_m$. Alors les colonnes $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m$ de la matrice $\mathbf{Y}'\mathbf{P}$ sont indépendants, de moyenne $\mathbf{M}'\mathbf{P}$ et de matrice de covariance $\boldsymbol{\Sigma}$. Si $\boldsymbol{\mu}_1 = \ldots = \boldsymbol{\mu}_n = \boldsymbol{\mu}$, $\mathbf{M}' = \boldsymbol{\mu}\boldsymbol{e}'$, alors $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m$ sont de moyenne **0** si les colonnes de **P** sont orthogonales à $\boldsymbol{e} = [1 ; 1 ; \ldots ; 1]'$.

Loi de Wishart

Soit $\mathbf{A} = \sum_{i=1}^{m}\boldsymbol{z}_i\boldsymbol{z}_i'$, où les $\boldsymbol{z}_i$ sont des vecteurs indépendants de loi $\mathcal{N}_p(\mathbf{0} ; \boldsymbol{\Sigma})$. Alors la statistique

$$\mathbf{A} \quad \mathcal{W}_p(\boldsymbol{\Sigma} ; m)$$

suit une loi appelée « loi de Wishart de dimension p, variance $\boldsymbol{\Sigma}$, et m degrés de liberté. On écrit alors $\mathbf{A} \sim \mathcal{W}_p(\boldsymbol{\Sigma} ; m)$.

Quelques propriétés de la loi de Wishart

- Si $\mathbf{A} \sim \mathcal{W}_p(\boldsymbol{\Sigma} ; m)$ et **B** est une matrice $q\times p$, alors $\mathbf{B}'\mathbf{A}\mathbf{B} \sim \mathcal{W}_q(\mathbf{B}'\boldsymbol{\Sigma}\mathbf{B} ; m)$

- Si $\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \sim \mathcal{W}_p(\mathbf{\Sigma} ; m)$, où $\mathbf{A}$ *est* $r\times r$, alors $\mathbf{A}_{11} \sim \mathcal{W}_r(\mathbf{\Sigma}_{11} ; m)$

- Si $\mathbf{A} \sim \mathcal{W}_p(\mathbf{\Sigma} ; m)$, alors $\mathbf{\Sigma}^{-½}\mathbf{A}\mathbf{\Sigma}^{-½} \sim \mathcal{W}_p(\mathbf{I}_p ; m)$.

- Si $\mathbf{A} \sim \mathcal{W}_p(\mathbf{I} ; m)$ et $\mathbf{B}$ est une matrice $p\times q$ $(q \le p)$ telle que $\mathbf{B}'\mathbf{B} = \mathbf{I}_q$, alors $\mathbf{B}'\mathbf{A}\mathbf{B} \sim \mathcal{W}_q(\mathbf{I}_q ; m)$

- Si $\mathbf{A} \sim \mathcal{W}_p(\mathbf{\Sigma} ; m)$ et $\boldsymbol{a}$ est un vecteur fixe tel que $\boldsymbol{a}'\mathbf{\Sigma}\boldsymbol{a} > 0$, alors $\boldsymbol{a}'\mathbf{A}\boldsymbol{a}/\boldsymbol{a}'\mathbf{\Sigma}\boldsymbol{a} \sim \chi^2_m$.
  Cas particulier : $a_{ii}/\sigma_{ii} \sim \chi^2_m$.

- Si $\mathbf{A}_1$ et $\mathbf{A}_2$ sont indépendantes, $\mathbf{A}_1 \sim \mathcal{W}_p(\mathbf{\Sigma} ; m_1)$ et $\mathbf{A}_2 \sim \mathcal{W}_p(\mathbf{\Sigma} ; m_2)$, alors $\mathbf{A}_1 + \mathbf{A}_2 \sim \mathcal{W}_p(\mathbf{\Sigma} ; m_1 + m_2)$.
- Soit $\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}$ une matrice de covariance $(p+q)\times(p+q)$, $\mathbf{\Sigma}_{11}$ étant $p\times p$, et $\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \sim \mathcal{W}_{p+q}(\mathbf{\Sigma} ; m)$, partitionnée de la même façon. Alors $\mathbf{A}_{11.2} = \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} \sim \mathcal{W}_p(\mathbf{A}_{11.2} ; m-q)$ et $\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} \sim \mathcal{W}_p(\mathbf{\Sigma}_{11} ; q)$; $\mathbf{A}_{11.2}$ est indépendante de $\mathbf{A}_{22}$ et de $\mathbf{A}_{12}$. Si $\mathbf{\Sigma}_{12} = \mathbf{0}$, et $\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}$, $\mathbf{A}_{22}$ et $\mathbf{A}_{11.2}$ sont mutuellement indépendants. Notons que si $\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} \\ \mathbf{A}^{21} & \mathbf{A}^{22} \end{bmatrix}$, alors $\mathbf{A}^{11} = \mathbf{A}_{11.2}^{-1}$, d'où on tire que $(\mathbf{A}^{11})^{-1} \sim \mathcal{W}_p(\mathbf{A}_{11.2} ; m-q)$.
- Soit $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ n vecteurs aléatoires indépendants, $\boldsymbol{y}_i \sim \mathcal{N}_p(\boldsymbol{\mu}_i ; \mathbf{\Sigma})$, $\mathbf{Y}' = [\boldsymbol{y}_1 \mid \boldsymbol{y}_2 \mid \ldots \mid \boldsymbol{y}_n]$, et $\mathbf{M} = [\boldsymbol{\mu}_1 \mid \boldsymbol{\mu}_2 \mid \ldots \mid \boldsymbol{\mu}_n]$. Si $\mathbf{P}$ est une matrice $n\times m$, $m \le n$ telle que $\mathbf{P}'\mathbf{P} = \mathbf{I}_m$ et $\mathbf{MP} = \mathbf{0}$, alors $\mathbf{Y}'\mathbf{P}\mathbf{P}'\mathbf{Y} \sim \mathcal{W}_p(\mathbf{\Sigma} ; m)$. En particulier, si $\mathbf{M} = \boldsymbol{\mu}\boldsymbol{e}'$ et $\mathbf{C} = \mathbf{I}_n - \boldsymbol{ee}'/n$, alors $n\mathbf{S} = \sum_{i=1}^n (\boldsymbol{y}_i - \bar{\boldsymbol{y}})(\boldsymbol{y}_i - \bar{\boldsymbol{y}})' = \mathbf{Y}'\mathbf{C}\mathbf{Y} \sim \mathcal{W}_p(\mathbf{\Sigma} ; n-1)$.

- Soit $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ n vecteurs aléatoires indépendants, $\boldsymbol{y}_i \sim \mathcal{N}_p(\boldsymbol{\mu}_i ; \mathbf{\Sigma})$, $\mathbf{Y}' = [\boldsymbol{y}_1 \mid \boldsymbol{y}_2 \mid \ldots \mid \boldsymbol{y}_n]$, et $\mathbf{M}' = [\boldsymbol{\mu}_1 \mid \boldsymbol{\mu}_2 \mid \ldots \mid \boldsymbol{\mu}_n]$. Si $\mathbf{P}$ est une matrice $n\times m$, $m \le n$ telle que $\mathbf{P}'\mathbf{P} = \mathbf{I}_m$ et $\mathbf{M}'\mathbf{P} = \mathbf{0}$, alors $\mathbf{Y}'\mathbf{P}\mathbf{P}'\mathbf{Y} \sim \mathcal{W}_p(\mathbf{\Sigma} ; m)$. En particulier, si $\mathbf{M}' = \boldsymbol{\mu}\boldsymbol{e}'$ et $\mathbf{C} = \mathbf{I}_n - \boldsymbol{ee}'/n$, alors $n\mathbf{S} = \sum_{i=1}^n (\boldsymbol{y}_i - \bar{\boldsymbol{y}})(\boldsymbol{y}_i - \bar{\boldsymbol{y}})' = \mathbf{Y}'\mathbf{C}\mathbf{Y} \sim \mathcal{W}_p(\mathbf{\Sigma} ; n-1)$.

La loi $\mathcal{T}^2$ de Hotelling

La loi $\mathcal{T}^2$ de Hotelling est une généralisation de la loi de Student.

Définition de la loi Soit $\boldsymbol{y} \sim \mathcal{N}_p(\boldsymbol{\mu} ; \mathbf{\Sigma})$ et $\mathbf{A} \sim \mathcal{W}_p(\mathbf{\Sigma} ; m)$, où $\boldsymbol{y}$ et $\mathbf{A}$ sont indépendantes. Alors la statistique

$$m(\boldsymbol{y}-\boldsymbol{\mu})'\mathbf{A}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}) \sim \mathcal{T}^2(p ; m)$$

suit une loi appelée « $\mathcal{T}^2$ de Hotelling » de paramètres p et m (dimension p et m degrés de liberté).

Relation avec la loi $\mathcal{F}$ La loi de Hotelling est liée à la loi $\mathcal{F}$ par la relation

$$\frac{m-p+1}{mp}\mathcal{T}^2(p;m) = \mathcal{F}_{p;m-p+1}$$

Application Une des applications principales de la loi de Hotelling est la suivante :

Soit $\bar{\boldsymbol{x}}$ et **S** la moyenne et la variance d'un échantillon de taille *n* d'une population $\mathcal{N}_p(\boldsymbol{\mu}\ ;\ \boldsymbol{\Sigma})$, alors

$$(n-1)(\bar{\boldsymbol{y}}-\boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\boldsymbol{y}}-\boldsymbol{\mu}) \sim \mathcal{T}^2(p\ ;\ n\text{-}1),\ \text{ou encore,}\ \frac{n-p}{p}(\bar{\boldsymbol{y}}-\boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\boldsymbol{y}}-\boldsymbol{\mu}) \sim \mathcal{F}_{p\ ;n\text{-}p}.$$

La distribution Λ de Wilks

Définition Soit $\mathbf{A} \sim \mathcal{W}_p(\mathbf{I}\ ;\ m)$ *et* $\mathbf{B} \sim \mathcal{W}_p(\mathbf{I}\ ;\ n)$, $m \geq p$, **A** *et* **B** *indépendantes. Alors la variable*

$$\Lambda = \frac{|\mathbf{A}|}{|\mathbf{A}+\mathbf{B}|} = |\mathbf{I}+\mathbf{A}^{-1}\mathbf{B}|^{-1}$$

suit une loi appelée loi de Wilks *de paramètres p, m, et n.* ■

Quelques propriétés de la loi de Wilks

- $\Lambda = \prod_{i=1}^{n} u_i$, où $u_1, \ldots, u_n$ sont indépendantes, u_i de loi bêta de paramètres ½(m+i-p) et p/2.
- Les lois $\Lambda(p\ ;\ m\ ;\ n)$ et $\Lambda(n\ ;\ m+n\text{-}p\ ;\ p)$ sont les mêmes
- $\dfrac{1-\Lambda(p;m;1)}{\Lambda(p;m;1)} \sim \dfrac{p}{m-p+1}\mathcal{F}_{p;m-p+1}$
- $\dfrac{1-\Lambda(1;m;n)}{\Lambda(1;m;n)} \sim \dfrac{n}{m}\mathcal{F}_{n;m}$
- $\dfrac{1-\sqrt{\Lambda(p;m;2)}}{\sqrt{\Lambda(p;m;2)}} \sim \dfrac{p}{m-p+1}\mathcal{F}_{2p;2(m-p+1)}$
- $\dfrac{1-\sqrt{\Lambda(2;m;n)}}{\sqrt{\Lambda(2;m;n)}} \sim \dfrac{n}{m-1}\mathcal{F}_{2n;2(m-1)}$
- Lorsque *m* est grand, on a, approximativement, $-\{m\text{-}½(p\text{-}n+1)\}\ln \Lambda(p\ ;\ m\ ;\ n) \sim \chi^2_{np}$.

RÉSUMÉ

1 L'espérance d'une matrice ou d'un vecteur aléatoire **Y** est la matrice (le vecteur) des espérances de ses composantes.

2 La matrice de covariance $\boldsymbol{\Sigma} = \mathrm{E}[(\boldsymbol{y}\text{-}\boldsymbol{\mu})(\boldsymbol{y}\text{-}\boldsymbol{\mu})']$ d'un vecteur aléatoire $\boldsymbol{y}$ de moyenne $\boldsymbol{\mu}$ est une matrice carrée symétrique dont l'élément σ_{ij} est la covariance entre la i^e et la j^e composante de $\boldsymbol{y}$. Les éléments de la diagonale de $\boldsymbol{\Sigma}$ sont donc les variances des composantes de $\boldsymbol{y}$. $\boldsymbol{\Sigma}$ est nécessairement semi définie positive ; elle est définie positive si et seulement si elle est non singulière. $\boldsymbol{\Sigma}$ est singulière s'il existe une fonction linéaire $\boldsymbol{a}'\boldsymbol{y}$ de variance nulle.

3 Soit $\boldsymbol{y}$ un vecteur aléatoire de moyenne $\boldsymbol{\mu}$ et de matrice de covariance $\boldsymbol{\Sigma}$, **L** et **M** des matrices constantes. Alors $\mathrm{E}(\mathbf{L}'\boldsymbol{y} + \boldsymbol{a}) = \mathbf{L}'\boldsymbol{\mu} + \boldsymbol{a}$, $\mathrm{Var}(\mathbf{L}'\boldsymbol{y} + \boldsymbol{a}) = \mathbf{L}'\boldsymbol{\Sigma}\mathbf{L}$, et $\mathrm{Cov}(\mathbf{L}'\boldsymbol{y}, \mathbf{M}'\boldsymbol{y}) = \mathbf{L}'\boldsymbol{\Sigma}\mathbf{M}$. Si y est normale, $\mathbf{L}'\boldsymbol{y} + \boldsymbol{a}$ est normale.

4 Si **A** est une matrice réelle symétrique, $\boldsymbol{y}$ un vecteur aléatoire de moyenne $\boldsymbol{\mu}$ et de matrice de covariance $\boldsymbol{\Sigma}$, alors $\mathrm{E}(\boldsymbol{y}'\mathbf{A}\boldsymbol{y}) = tr(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$

5 Soit $\boldsymbol{y}' = [\boldsymbol{y}_1'\ ;\ \boldsymbol{y}_2']$ un vecteur aléatoire normal partitionné de moyenne $\boldsymbol{\mu} = \begin{pmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{pmatrix}$ et de matrice de

covariance $\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}$. Alors E($\boldsymbol{y}_1 \mid \boldsymbol{y}_2$) = $\boldsymbol{\mu}_1 + \mathbf{\Sigma}_{12}\,\mathbf{\Sigma}_{22}^{-1}(\boldsymbol{y}_2 - \boldsymbol{\mu}_2)$, Var($\boldsymbol{y}_1 \mid \boldsymbol{y}_2$) = $\mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\,\mathbf{\Sigma}_{22}^{-1}\,\mathbf{\Sigma}_{21}$.

$\boldsymbol{y}_1$ et $\boldsymbol{y}_2$ sont indépendants si et seulement si $\mathbf{\Sigma}_{12} = \mathbf{\Sigma}_{21} = \mathbf{0}$.

6 Soit $\boldsymbol{u} = \mathbf{D}^{-1/2}\mathbf{P}'\boldsymbol{y}$, où $\mathbf{P}$ est une matrice orthogonale et $\mathbf{D}$ une matrice diagonale telle que $\mathbf{\Sigma} = \mathbf{PDP}'$. Alors $\boldsymbol{u} \sim \mathcal{N}_n(\mathbf{D}^{-1/2}\mathbf{P}'\boldsymbol{\mu}\ ;\ \mathbf{I}_n)$.

7 On désigne par $\chi_s^2(\lambda)$ la loi khi-deux non centrale à s degrés de liberté et paramètre de non centralité λ. Soit $\boldsymbol{y} \sim \mathcal{N}_n(\boldsymbol{\mu}\ ;\ \mathbf{\Sigma})$, où $\mathbf{\Sigma}$ est définie positive, $\mathbf{A}$ une matrice réelle symétrique de rang r. Alors (i) si $\mathbf{A\Sigma A} = \mathbf{A}$, alors $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} \sim \chi_r^2(\lambda)$ où $r = r(\mathbf{A})$ et $\lambda = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$; (ii) si $\boldsymbol{y}'\mathbf{A}\boldsymbol{y} \sim \chi_s^2(\lambda)$, alors $\mathbf{A\Sigma A} = \mathbf{A}$, $s = r(\mathbf{A})$ et $\lambda = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$. $\lambda = 0$ si et seulement si $\mathbf{A}\boldsymbol{\mu} = \mathbf{0}$.

8 Soit $\boldsymbol{y} \sim \mathcal{N}_n(\boldsymbol{\mu}\ ;\ \mathbf{\Sigma})$, $\mathbf{A}$ et $\mathbf{B}$ deux matrices symétriques. Soit $\boldsymbol{y}'\mathbf{A}\boldsymbol{y}$ et $\boldsymbol{y}'\mathbf{B}\boldsymbol{y}$ deux formes quadratiques de loi khi-deux (centrale ou non). Alors $\boldsymbol{y}'\mathbf{A}\boldsymbol{y}$ et $\boldsymbol{y}'\mathbf{B}\boldsymbol{y}$ sont indépendantes si et seulement si $\mathbf{A\Sigma B} = \mathbf{0}$.

9 Cas particulier : $\mathbf{\Sigma} = \sigma^2\mathbf{I}_n$. Alors $\boldsymbol{y}'\boldsymbol{y}/\sigma^2 = \dfrac{\sum_{i=1}^n y_i^2}{\sigma^2} \sim \chi_n^2(\lambda)$ où $\lambda = \dfrac{\boldsymbol{\mu}'\boldsymbol{\mu}}{\sigma^2}$. $\boldsymbol{y}'\mathbf{A}\boldsymbol{y}/\sigma^2 \sim \chi_r^2(\lambda)$ où $r = r(\mathbf{A})$ et $\lambda = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}/\sigma^2$ si et seulement si $\mathbf{A}$ est idempotente ; si $\mathbf{A}$ et $\mathbf{B}$ sont idempotentes, alors $\boldsymbol{y}'\mathbf{A}\boldsymbol{y}$ et $\boldsymbol{y}'\mathbf{B}\boldsymbol{y}$ sont indépendantes si et seulement si $\mathbf{AB} = \mathbf{0}$.

10 Un quotient $T = \dfrac{Y}{\sqrt{X/\nu}}$ suit une loi appelée *loi de Student* à ν degrés de liberté si X_1 et X_2 sont indépendantes, $X_1 \sim \chi_{\nu_1}^2$ et $X_2 \sim \chi_{\nu_2}^2$.

11 Le quotient $\mathcal{F} = \dfrac{X_1/\nu_1}{X_2/\nu_2}$ suit une loi appelée *loi de Fisher*, ou loi $\mathcal{F}$, à ν_1 et ν_2 degrés de liberté si X_1 et X_2 sont indépendantes, $X_1 \sim \chi_{\nu_1}^2$ et $X_2 \sim \chi_{\nu_2}^2$.

12 Soit $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ n vecteurs aléatoires indépendants, $\boldsymbol{x}_i \sim \mathcal{N}_p(\boldsymbol{\mu}_i\ ;\ \mathbf{\Sigma})$, $\mathbf{Y}' = [\boldsymbol{y}_1 \mid \boldsymbol{y}_2 \mid \ldots \mid \boldsymbol{y}_n]$, $\mathbf{M}' = [\boldsymbol{\mu}_1 \mid \boldsymbol{\mu}_2 \mid \ldots \mid \boldsymbol{\mu}_n]$, $\mathbf{P}$ est une matrice $n\times m$, $m \le n$ telle que $\mathbf{P}'\mathbf{P} = \mathbf{I}_m$. Alors les colonnes $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m$ de la matrice $\mathbf{Y}'\mathbf{P}$ sont indépendantes, de moyenne $\mathbf{M}'\mathbf{P}$ et de matrice de covariance $\mathbf{\Sigma}$. Si $\boldsymbol{\mu}_1 = \ldots = \boldsymbol{\mu}_n = \boldsymbol{\mu}$, $\mathbf{M}' = \boldsymbol{\mu}\boldsymbol{e}'$, alors $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m$ sont de moyenne $\mathbf{0}$ si les colonnes de $\mathbf{P}$ sont orthogonales à $\boldsymbol{e} = [1\ ;\ 1\ ;\ldots\ ;1]'$.

13 Soit $\mathbf{A} = \sum_{i=1}^m \boldsymbol{z}_i\boldsymbol{z}_i'$, où les $\boldsymbol{z}_i$ sont des vecteurs indépendants de loi $\mathcal{N}_p(\mathbf{0}\ ;\ \mathbf{\Sigma})$. Alors la statistique $\mathbf{A} \sim \mathcal{W}_p(\mathbf{\Sigma}\ ;\ m)$.

14 Si $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ sont n vecteurs aléatoires indépendants, $\boldsymbol{y}_i \sim \mathcal{N}_p(\boldsymbol{\mu}\ ;\ \mathbf{\Sigma})$, et $\mathbf{S} = \dfrac{1}{n}\sum_{i=1}^n (\boldsymbol{y}_i - \bar{\boldsymbol{y}})(\boldsymbol{y}_i - \bar{\boldsymbol{y}})'$. Alors $n\mathbf{S} \sim \mathcal{W}_p(\mathbf{\Sigma}\ ;\ n\text{-}1)$, et $\mathbf{S}$ est indépendante de $\bar{\boldsymbol{y}} = \dfrac{1}{n}\sum_{i=1}^n \boldsymbol{y}_i$.

15 Soit $\boldsymbol{y} \sim \mathcal{N}_p(\boldsymbol{\mu}\ ;\ \mathbf{\Sigma})$ et $\mathbf{A} \sim \mathcal{W}_p(\mathbf{\Sigma}\ ;\ m)$, où $\boldsymbol{y}$ et $\mathbf{A}$ sont indépendantes. Alors $m(\boldsymbol{y}-\boldsymbol{\mu})'\mathbf{A}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}) \sim \mathcal{T}^2(p\ ;\ m)$.

16 $\dfrac{m-p+1}{mp}\mathcal{T}^2(p;m) = \mathcal{F}_{p;m-p+1}$.

17 Soit $\bar{\boldsymbol{y}}$ et $\mathbf{S}$ la moyenne et la variance d'un échantillon de taille n d'une population $\mathcal{N}_p(\boldsymbol{\mu}\ ;\ \mathbf{\Sigma})$, alors $(n-1)(\bar{\boldsymbol{y}}-\boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\boldsymbol{y}}-\boldsymbol{\mu}) \sim \mathcal{T}^2(p\ ;\ n\text{-}1)$, ou encore, $\dfrac{n-p}{p}(\bar{\boldsymbol{y}}-\boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\boldsymbol{y}}-\boldsymbol{\mu}) \sim \mathcal{F}_{p\ ;n-p}$.