

REVUE DE STATISTIQUE APPLIQUÉE

H. EL MAÂCHE

Y. LEPAGE

Mesures d'association vectorielle basées sur une matrice de corrélation

Revue de statistique appliquée, tome 46, n° 4 (1998), p. 27-43

http://www.numdam.org/item?id=RSA_1998__46_4_27_0

© Société française de statistique, 1998, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

MESURES D'ASSOCIATION VECTORIELLE BASÉES SUR UNE MATRICE DE CORRÉLATION

H. El Maâche, Y Lepage

Département de mathématiques et de statistique, Université de Montréal

RÉSUMÉ

Trois mesures d'association vectorielle sont construites à partir de la matrice de corrélation échantillonnale. Ces mesures servent à tester l'absence de liaison entre deux vecteurs de lois elliptiques. À partir de la loi asymptotique de la matrice de corrélation échantillonnale, la loi asymptotique des mesures d'association vectorielle est obtenue sous l'hypothèse d'absence de liaison et sous une suite d'hypothèses alternatives. Un exemple est donné pour illustrer les différents tests et une étude expérimentale permet de comparer ces tests à ceux basés sur la matrice de covariance.

Mots-clés : Mesures d'association vectorielle, matrice de corrélation, lois elliptiques, lois asymptotiques, tests d'hypothèses.

ABSTRACT

Three measures of multivariate association are constructed using the correlation matrix. These measures are utilised to test the lack of multivariate association of two vectors of elliptical distributions. The asymptotic distributions of the measures are obtained under the hypothesis of lack of multivariate association and under a sequence of alternatives from the asymptotic distribution of the correlation matrix. The tests are illustrated by an example and an empirical study is performed to compare the new tests with those based on the covariance matrix.

Keywords : Multivariate association, correlation matrix, asymptotic distributions, elliptical distributions, hypothesis testing.

1. Introduction

On veut éprouver l'hypothèse nulle d'absence de liaison entre deux vecteurs aléatoires dont la loi appartient à la famille des lois elliptiques. En utilisant la matrice de covariance échantillonnale, Muirhead (1982) a proposé un test asymptotique pour tester l'absence de liaison entre plusieurs vecteurs aléatoires. Dans le même contexte, en utilisant la mesure d'Escoufier (1973) basée sur la matrice de covariance, Cléroux & Ducharme (1989) ont proposé un autre test pour l'absence de liaison entre deux vecteurs aléatoires. À partir des mesures d'Escoufier (1973), de Stewart & Love (1968) et Cramer & Nicewander (1979), Allaire & Lepage (1990) ont généralisé le travail

de Cléroux & Ducharme (1989) en proposant trois autres tests pour tester l'absence de liaison entre plusieurs vecteurs aléatoires.

En pratique, il arrive souvent que l'échelle des variables n'est pas la même et que les variances respectives sont très différentes. Pour pallier à ce problème, on centre et réduit chaque variable afin de donner un poids égal à chacune d'elles dans l'analyse. Dans de telles situations, la mesure d'Escoufier (1973) et la mesure de Stewart & Love (1968) ne sont pas appropriées puisqu'elles ne sont pas invariantes sous des transformations linéaires des vecteurs d'observations. Ce problème a déjà été soulevé par Allaire & Lepage (1990) et aussi par Roy & Cléroux (1993). Pour y remédier, on applique les mesures d'Escoufier (1973), de Stewart & Love (1968) et Cramer & Nicewander (1979) à la matrice de corrélation plutôt qu'à la matrice de covariance. On en déduit trois nouveaux tests asymptotiques pour éprouver l'absence de liaison entre deux vecteurs aléatoires provenant d'une loi elliptique. La loi asymptotique des mesures est obtenue sous l'hypothèse nulle d'absence de liaison (Théorème 3) et sous une suite d'alternatives convergeant vers l'hypothèse nulle (Théorème 4). Ces résultats originaux reposent sur la connaissance explicite de la matrice de covariance de la loi asymptotique de la matrice de corrélation d'un échantillon provenant d'une loi elliptique (Théorèmes 1 et 2). Un exemple illustre les nouveaux tests et une étude expérimentale permet de comparer les niveaux et les puissances des trois tests basés sur la matrice de corrélation à ceux construits à partir de la matrice de covariance.

2. Notation

Soit $X = (X^{(1)}, \dots, X^{(p)}, X^{(p+1)}, \dots, X^{(m)})'$ un vecteur aléatoire de dimension m , de matrice de covariance $\Sigma = (\sigma^{(k,l)})$ et de matrice de corrélation $\mathcal{P} = (\rho^{(k,l)})$ où $\rho^{(k,l)} = \sigma^{(k,l)} / \sqrt{\sigma^{(k,k)}} \sqrt{\sigma^{(l,l)}}$, k et $l = 1, \dots, m$. À partir d'un échantillon X_1, \dots, X_n , on construit les matrices de covariance empirique $S = (s_{kl})$ et de corrélation empirique $R = (r_{kl})$ où $r_{kl} = s_{kl} / \sqrt{s_{kk}} \sqrt{s_{ll}}$, k et $l = 1, \dots, m$. On partitionne le vecteur X en deux sous-vecteurs $X^{[1]}$ et $X^{[2]}$ de dimension respective p et q ($p + q = m$) et les matrices \mathcal{P} et R :

$$\mathcal{P} = \begin{pmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{21} & \mathcal{P}_{22} \end{pmatrix} \quad \text{et} \quad R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}$$

où $\mathcal{P}_{11} : p \times p$ et $R_{11} : p \times p$.

Les mesures d'Escoufier (1973), de Stewart & Love (1968) et de Cramer & Nicewander (1979) appliquées à la matrice de corrélation échantillonnale sont

$$RV^{(R)} = \frac{\text{tr}(R_{12}R'_{12})}{\sqrt{\text{tr}(R_{11}^2) \text{tr}(R_{22}^2)}} \quad \text{pour la mesure d'Escoufier (1973)}$$

$$SL^{(R)} = \frac{\text{tr}(R_{12}R_{22}^{-1}R'_{12})}{p} \quad \text{pour la mesure de Stewart & Love (1968)}$$

$$\text{et } CN^{(R)} = \frac{\text{tr}(R_{11}^{-1}R_{12}R_{22}^{-1}R'_{12})}{p} \quad \text{pour la mesure de Cramer \& Nicewander (1979), où on a tenu compte que } \text{tr}(R_{11}) = \text{tr}(\mathcal{P}_{11}) = p.$$

Notons que ce dernier

coefficient est identique à celui que l'on obtiendrait si l'on remplaçait les R_{ij} par les S_{ij} .

Au niveau de la population, si on applique les mêmes mesures à la matrice de corrélation, on obtient

$$\rho_{RV}^{(R)} = \frac{\text{tr}(\mathcal{P}_{12}\mathcal{P}'_{12})}{\sqrt{\text{tr}(\mathcal{P}_{11}^2)\text{tr}(\mathcal{P}_{22}^2)}} \text{ pour la mesure d'Escoufier (1973),}$$

$$\rho_{SL}^{(R)} = \frac{\text{tr}(\mathcal{P}_{12}\mathcal{P}_{22}^{-1}\mathcal{P}'_{12})}{p} \text{ pour la mesure de Stewart \& Love (1968)}$$

et $\rho_{CN}^{(R)} = \frac{\text{tr}(\mathcal{P}_{11}^{-1}\mathcal{P}_{12}\mathcal{P}_{22}^{-1}\mathcal{P}'_{12})}{p}$ pour la mesure de Cramer & Nicewander (1979).

Les trois nouvelles mesures construites sont invariantes au changement de position et d'échelle et possèdent les propriétés suivantes :

(i) $\rho M^{(R)} = 0$ si et seulement si $\mathcal{P}_{21} = 0$, pour $M = RV, SL$ et CN .

(ii) Lorsque $p = q = 1$, les trois mesures correspondent au carré du coefficient de corrélation simple entre les variables $X^{(1)}$ et $X^{(2)}$.

(iii) $0 \leq \rho M^{(R)} \leq 1$, pour $M = RV, SL$ et CN .

(iv) $M^{(R)}$ est un estimateur convergent de $\rho M^{(R)}$ pour $M = RV, SL$ et CN .

3. Distribution asymptotique de $n^{1/2} \text{vec } R$.

La distribution asymptotique de $\sqrt{n} \text{vec } R$, où $\text{vec } R$ est le vecteur obtenu en empilant les colonnes de la matrice R , qui est présentée dans le cas général par Steiger & Hakstian (1982) ne permet pas d'obtenir les éléments de la matrice de covariance asymptotique de ce vecteur comme fonction des éléments de la matrice de corrélation de la population \mathcal{P} , autrement dit comme fonction seulement des moments d'ordre deux. En faisant l'hypothèse d'ellipticité, on peut donner une expression matricielle qui dépend de \mathcal{P} pour cette matrice de covariance asymptotique.

THÉORÈME 1. *Soit R la matrice de corrélation d'un échantillon de taille n issu d'un vecteur aléatoire $X = (X_1, \dots, X_m)'$ dont la loi est elliptique de paramètres μ et V , de matrice de corrélation \mathcal{P} et de coefficient d'aplatissement κ . Alors la loi asymptotique, lorsque $n \rightarrow \infty$, de $\sqrt{n}(\text{vec } R - \text{vec } \mathcal{P})$ est une loi multinormale $\mathcal{N}_{m^2}(O, (1 + \kappa)\Lambda)$ avec*

$$\begin{aligned} \Lambda = & \frac{1}{2} \left[(I \otimes \mathcal{P})E(\mathcal{P} \otimes \mathcal{P})E(I \otimes \mathcal{P}) + (\mathcal{P} \otimes I)E(\mathcal{P} \otimes \mathcal{P})E(\mathcal{P} \otimes I) \right. \\ & + (I \otimes \mathcal{P})E(\mathcal{P} \otimes \mathcal{P})E(\mathcal{P} \otimes I) + (\mathcal{P} \otimes I)E(\mathcal{P} \otimes \mathcal{P})E(I \otimes \mathcal{P}) \left. \right] \\ & + (\mathcal{P} \otimes \mathcal{P}) + (\mathcal{P} \otimes \mathcal{P})K - (I \otimes \mathcal{P})E(\mathcal{P} \otimes \mathcal{P}) \\ & - (\mathcal{P} \otimes \mathcal{P})E(I \otimes \mathcal{P}) - (\mathcal{P} \otimes \mathcal{P})E(\mathcal{P} \otimes I) - (\mathcal{P} \otimes I)E(\mathcal{P} \otimes \mathcal{P}) \quad (3.1) \end{aligned}$$

où la matrice I désigne la matrice identité $m \times m$, $E = \sum_{i=1}^m H_{ii} \otimes H_{ii}$ avec H_{ij} , $i, j = 1, \dots, m$, étant la matrice $m \times m$ ayant 1 à la position (i, j) et 0 ailleurs et $K = K_{mm} = \sum_{i=1}^m \sum_{j=1}^m H_{ij} \otimes H'_{ij}$ désigne la matrice de commutation $m^2 \times m^2$.

Démonstration. D'après le théorème de Steiger & Hakstian (1982), on a

$$\sqrt{n} \operatorname{vec}(R - \mathcal{P}) \xrightarrow{\mathcal{L}} \mathcal{N}_{m^2}(0, \Omega_R)$$

où les éléments de Ω_R sont donnés par

$$\begin{aligned} \omega^{(ij,kl)} = & \rho^{(ij,kl)} + \frac{1}{4} \rho^{(i,j)} \rho^{(k,l)} (\rho^{(ii,kk)} + \rho^{(ii,ll)} + \rho^{(jj,kk)} + \rho^{(jj,ll)}) \\ & - \frac{1}{2} \rho^{(i,j)} (\rho^{(ii,kl)} + \rho^{(jj,kl)}) - \frac{1}{2} \rho^{(k,l)} (\rho^{(ij,kk)} + \rho^{(ij,ll)}), \end{aligned}$$

avec $\rho^{(ij,kl)} = \sigma^{(ij,kl)} (\sigma^{(i,i)} \sigma^{(j,j)} \sigma^{(k,k)} \sigma^{(l,l)})^{-1/2}$ et $\sigma^{(ij,kl)} = E[(X^{(i)} - \mu^{(i)})(X^{(j)} - \mu^{(j)})(X^{(k)} - \mu^{(k)})(X^{(l)} - \mu^{(l)})]$, $i, j, k, l = 1, \dots, m$. Puisque X est elliptique, on a (voir Muirhead (1982))

$$\rho^{(ij,kl)} = (1 + \kappa) (\rho^{(i,j)} \rho^{(k,l)} + \rho^{(i,k)} \rho^{(j,l)} + \rho^{(i,l)} \rho^{(j,k)}) \quad (3.2)$$

et ainsi, on obtient

$$\begin{aligned} \omega^{(ij,kl)} = & (1 + \kappa) \left[\frac{1}{2} \rho^{(i,j)} \rho^{(k,l)} [(\rho^{(i,k)})^2 + (\rho^{(i,l)})^2 + (\rho^{(j,k)})^2 + (\rho^{(j,l)})^2] \right. \\ & - \rho^{(i,j)} [\rho^{(i,k)} \rho^{(i,l)} + \rho^{(j,k)} \rho^{(j,l)}] - \rho^{(k,l)} [\rho^{(i,l)} \rho^{(j,l)} + \rho^{(i,k)} \rho^{(j,k)}] \\ & \left. + [\rho^{(i,k)} \rho^{(j,l)} + \rho^{(i,l)} \rho^{(j,k)}] \right]. \quad (3.3) \end{aligned}$$

On identifie la matrice à laquelle correspond chaque élément du membre de droite de la somme $\omega^{(ij,kl)}$. On a par exemple

- $\rho^{(i,k)} \rho^{(j,l)}$ est évidemment l'élément générique de la matrice $\mathcal{P} \otimes \mathcal{P}$.
- $\rho^{(i,l)} \rho^{(j,k)}$ est l'élément générique de la matrice $(\mathcal{P} \otimes \mathcal{P})K$. En effet, si on pose $e'_i = (0, \dots, 0, 1, 0, \dots, 0)$, le vecteur de dimension m dont toutes les composantes

sont nulles sauf la $i^{\text{ième}}$ qui prend la valeur 1, on a

$$\begin{aligned}
 & \sum_{ijkl}^m (e_i \otimes e_j) [\rho^{(i,l)} \rho^{(j,k)}] (e_k \otimes e_l)' = \sum_{ijkl}^m (e_i \otimes e_j) [(e_i' \mathcal{P} e_l) \otimes (e_j' \mathcal{P} e_k)] (e_k' \otimes e_l') \\
 &= \sum_{ijkl}^m (e_i \otimes e_j) (e_i' \otimes e_j') (\mathcal{P} \otimes \mathcal{P}) (e_l \otimes e_k) (e_k' \otimes e_l') \\
 &= \sum_{ijkl}^m (H_{ii} \otimes H_{jj}) (\mathcal{P} \otimes \mathcal{P}) (H_{lk} \otimes H_{lk}') = (I \otimes I) (\mathcal{P} \otimes \mathcal{P}) K = (\mathcal{P} \otimes \mathcal{P}) K.
 \end{aligned}$$

De cette façon, on peut trouver la matrice à laquelle appartient l'élément en question. Ainsi, on obtient que

$$\begin{aligned}
 & \rho^{(i,j)} \rho^{(k,l)} (\rho^{(j,k)})^2 \text{ est l'élément générique de } (\mathcal{P} \otimes I) E(\mathcal{P} \otimes \mathcal{P}) E(I \otimes \mathcal{P}), \\
 & \rho^{(i,j)} \rho^{(k,l)} (\rho^{(i,l)})^2 \text{ est l'élément générique de } (I \otimes \mathcal{P}) E(\mathcal{P} \otimes \mathcal{P}) E(\mathcal{P} \otimes I), \\
 & \rho^{(i,j)} \rho^{(k,l)} (\rho^{(i,k)})^2 \text{ est l'élément générique de } (I \otimes \mathcal{P}) E(\mathcal{P} \otimes \mathcal{P}) E(I \otimes \mathcal{P}), \\
 & \rho^{(i,j)} \rho^{(k,l)} (\rho^{(j,l)})^2 \text{ est l'élément générique de } (\mathcal{P} \otimes I) E(\mathcal{P} \otimes \mathcal{P}) E(\mathcal{P} \otimes I), \\
 & \rho^{(i,j)} \rho^{(i,k)} \rho^{(i,l)} \text{ est l'élément générique de } (I \otimes \mathcal{P}) E(\mathcal{P} \otimes \mathcal{P}), \\
 & \rho^{(i,j)} \rho^{(j,k)} \rho^{(j,l)} \text{ est l'élément générique de } (\mathcal{P} \otimes I) E(\mathcal{P} \otimes \mathcal{P}), \\
 & \rho^{(k,l)} \rho^{(i,l)} \rho^{(j,l)} \text{ est l'élément générique de } (\mathcal{P} \otimes \mathcal{P}) E(\mathcal{P} \otimes I), \\
 & \rho^{(k,l)} \rho^{(i,k)} \rho^{(j,k)} \text{ est l'élément générique de } (\mathcal{P} \otimes \mathcal{P}) E(I \otimes \mathcal{P}).
 \end{aligned}$$

On peut, par exemple, vérifier que $\rho^{(i,j)} \rho^{(k,l)} (\rho^{(i,l)})^2$ est l'élément générique de la matrice $(I \otimes \mathcal{P}) E(\mathcal{P} \otimes \mathcal{P}) E(\mathcal{P} \otimes I)$. Il suffit d'écrire

$$\begin{aligned}
 & (e_i \otimes e_j)' (I \otimes \mathcal{P}) E(\mathcal{P} \otimes \mathcal{P}) E(\mathcal{P} \otimes I) (e_k \otimes e_l) \\
 &= \sum_{st}^m (e_i' \otimes e_j') [(I \otimes \mathcal{P}) (H_{tt} \otimes H_{tt}) (\mathcal{P} \otimes \mathcal{P}) (H_{ss} \otimes H_{ss}) (\mathcal{P} \otimes I)] (e_k \otimes e_l) \\
 &= \sum_{st}^m (e_i' H_{tt} \mathcal{P} H_{ss} \mathcal{P} e_k) \otimes (e_j' \mathcal{P} H_{tt} \mathcal{P} H_{ss} e_l) = (e_i' \mathcal{P} e_l) (e_l' \mathcal{P} e_k) (e_j' \mathcal{P} e_i) (e_i' \mathcal{P} e_l) \\
 &= \rho^{(i,l)} \rho^{(l,k)} \rho^{(j,i)} \rho^{(i,l)} = \rho^{(i,j)} \rho^{(k,l)} (\rho^{(i,l)})^2.
 \end{aligned}$$

c.q.f.d.

4. Distributions asymptotiques de $n^{1/2}$ vec R_{21} .

On veut éprouver l'hypothèse nulle $H_0 : \mathcal{P}_{21} = 0$. Cette hypothèse n'implique évidemment pas l'indépendance entre $X^{[1]}$ et $X^{[2]}$ mais seulement la non corrélation

entre $X^{[1]}$ et $X^{[2]}$. Considérons aussi la suite d'hypothèses alternatives

$$H_{1:n} : \mathcal{P}_{21} = \frac{1}{\sqrt{n}}A = \left(\frac{a_{ij}}{\sqrt{n}} \right) \quad \text{avec} \quad |a_{ij}| \leq \sqrt{n},$$

où A est une matrice donnée de dimension $q \times p$, $i = p+1, \dots, m$ et $j = 1, \dots, p$.

La loi asymptotique de $\sqrt{n} \text{vec } R_{21}$ sous H_0 et sous la suite $H_{1:n}$ convergeant vers H_0 se déduit comme cas particulier de la première partie du théorème suivant qui donne, à partir du théorème 1, la loi asymptotique de $\sqrt{n} \text{vec } R_{21}$ avec une expression matricielle pour la matrice de covariance asymptotique.

THÉORÈME 2. *Sous les conditions du théorème 1, on a*

$$(i) \sqrt{n}(\text{vec } R_{21} - \text{vec } \mathcal{P}_{21}) \xrightarrow{\mathcal{L}} \mathcal{N}_{pq}(O, (1 + \kappa)D), \text{ avec}$$

$$\begin{aligned} D = & \frac{1}{2} \left[(I_p \otimes \mathcal{P}_{21})E_p(\mathcal{P}_{11} \otimes \mathcal{P}_{11})E_p(I_p \otimes \mathcal{P}_{12}) \right. \\ & + (\mathcal{P}_{12} \otimes I_q)E_q(\mathcal{P}_{22} \otimes \mathcal{P}_{22})E_q(\mathcal{P}_{21} \otimes I_q) \\ & + (I_p \otimes \mathcal{P}_{21})E_p(\mathcal{P}_{12} \otimes \mathcal{P}_{12})E_q(\mathcal{P}_{21} \otimes I_q) \\ & \left. + (\mathcal{P}_{12} \otimes I_q)E_q(\mathcal{P}_{21} \otimes \mathcal{P}_{21})E_p(I_p \otimes \mathcal{P}_{12}) \right] \\ & + (\mathcal{P}_{11} \otimes \mathcal{P}_{22}) + (\mathcal{P}_{12} \otimes \mathcal{P}_{21})K_{qp} - (I_p \otimes \mathcal{P}_{21})E_p(\mathcal{P}_{11} \otimes \mathcal{P}_{12}) \\ & - (\mathcal{P}_{11} \otimes \mathcal{P}_{21})E_p(I_p \otimes \mathcal{P}_{12}) - (\mathcal{P}_{12} \otimes I_q)E_q(\mathcal{P}_{21} \otimes \mathcal{P}_{22}) \\ & - (\mathcal{P}_{12} \otimes \mathcal{P}_{22})E_q(\mathcal{P}_{21} \otimes I_q), \end{aligned}$$

où pour

$$H_{ij} = \begin{pmatrix} H_{ij}^{11} & H_{ij}^{12} \\ H_{ij}^{21} & H_{ij}^{22} \end{pmatrix},$$

$$E_p = \sum_{t=1}^p H_{tt}^{11} \otimes H_{tt}^{11}, \quad E_q = \sum_{t=p+1}^m H_{tt}^{22} \otimes H_{tt}^{22}, \quad p + q = m$$

$$\text{et } K_{qp} = \sum_{i=p+1}^m \sum_{j=1}^p H_{ij}^{21} \otimes H_{ij}^{21'} \text{ désigne la matrice de commutation } qp \times qp.$$

(ii) Sous H_0 , $\sqrt{n} \text{vec } R_{21}$ converge en loi vers une loi multinormale

$$\mathcal{N}_{pq}(O, (1 + \kappa)(\mathcal{P}_{11} \otimes \mathcal{P}_{22})).$$

(iii) Sous $H_{1:n}$, $\sqrt{n} \text{vec } R_{21}$ converge en loi vers une loi multinormale

$$\mathcal{N}_{pq}(\text{vec } A, (1 + \kappa)(\mathcal{P}_{11} \otimes \mathcal{P}_{22})).$$

Démonstration. (i) Soit la matrice $B = B_1 \otimes B_2$ avec $B_1 = [I_p \ 0] : p \times m$ et $B_2 = [0 \ I_q] : q \times m$. Alors, B vec $R = (B_1 \otimes B_2)$ vec $R = \text{vec}(B_2 R B_1') = \text{vec} R_{21}$ et d'après le théorème 1, on obtient que $\sqrt{n}(\text{vec} R_{21} - \text{vec} \mathcal{P}_{21})$ converge en loi vers $\mathcal{N}_{pq}(0, (1 + \kappa) B \Lambda B')$.

Il suffit alors d'expliciter la matrice $B \Lambda B'$. Pour cela, on reprend terme à terme la somme des matrices constituant Λ .

Considérons par exemple la matrice $(I \otimes \mathcal{P})E(\mathcal{P} \otimes \mathcal{P})E(I \otimes \mathcal{P})$ et calculons la première matrice :

$$\begin{aligned} \overset{\circ}{A} &= B(I \otimes \mathcal{P})E(\mathcal{P} \otimes \mathcal{P})E(I \otimes \mathcal{P})B' \\ &= \sum_{ts}^m \sum_{ts}^m (B_1 \otimes B_2)(I \otimes \mathcal{P})(H_{tt} \otimes H_{tt}) \\ &\quad (\mathcal{P} \otimes \mathcal{P})(H_{ss} \otimes H_{ss})(I \otimes \mathcal{P})(B_1' \otimes B_2') \\ &= \sum_{ts}^m (B_1 H_{tt}) \mathcal{P} (B_1 H_{ss})' \otimes (B_2 \mathcal{P} H_{tt}) \mathcal{P} (B_2 \mathcal{P} H_{ss})' \end{aligned}$$

puisque

$$B_1 H_{tt} = [D_{tt} \ 0] : p \times m = \begin{cases} 0 : p \times m & \text{si } t > p, \\ [D_{tt} \ 0] : p \times m & \text{sinon,} \end{cases}$$

avec $D_{tt} : p \times p$ ayant 1 à la position (t, t) et 0 ailleurs.

De même, on a

$$H_{ss} B_1' = (B_1 H_{ss})' = \begin{cases} 0 : m \times p & \text{si } s > p, \\ [D_{ss} \ 0]' : m \times p & \text{sinon.} \end{cases}$$

Donc si $t > p$ ou $s > p$, on obtient $\overset{\circ}{A} = 0$.

Pour $1 \leq t \leq p$ et $1 \leq s \leq p$, on a

$$B_2 \mathcal{P} H_{tt} = [0 \ I_q] \begin{pmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{21} & \mathcal{P}_{22} \end{pmatrix} \begin{pmatrix} D_{tt} & 0 \\ 0 & 0 \end{pmatrix} = [\mathcal{P}_{21} D_{tt} \ 0] : q \times m,$$

d'où

$$\begin{aligned} (B_2 \mathcal{P} H_{tt}) \mathcal{P} (B_2 \mathcal{P} H_{ss})' &= [\mathcal{P}_{21} D_{tt} \ 0] \begin{pmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{21} & \mathcal{P}_{22} \end{pmatrix} \begin{pmatrix} D_{ss} \mathcal{P}_{12} \\ 0 \end{pmatrix} \\ &= \mathcal{P}_{21} D_{tt} \mathcal{P}_{11} D_{ss} \mathcal{P}_{12} \end{aligned}$$

et

$$(B_1 H_{tt}) \mathcal{P} (B_1 H_{ss})' = [D_{tt} \ 0] \begin{pmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{21} & \mathcal{P}_{22} \end{pmatrix} \begin{pmatrix} D_{ss} \\ 0 \end{pmatrix} = D_{tt} \mathcal{P}_{11} D_{ss}.$$

Ainsi, on a

$$\begin{aligned} \overset{\circ}{A} &= \sum_{t=1}^p \sum_{s=1}^p (D_{tt} \mathcal{P}_{11} D_{ss}) \otimes (\mathcal{P}_{21} D_{tt} \mathcal{P}_{11} D_{ss} \mathcal{P}_{12}) \\ &= (I_p \otimes \mathcal{P}_{21}) E_p (\mathcal{P}_{11} \otimes \mathcal{P}_{11}) E_p (I_p \otimes \mathcal{P}_{12}). \end{aligned}$$

En remarquant que certaines matrices de Λ sont des transposées d'autres et que $K_{mm}(B'_1 \otimes B'_2) = (B'_2 \otimes B'_1) K_{qp}$, on obtient par le même calcul le résultat cherché.

(ii) Si H_0 est vraie, c'est-à-dire $\mathcal{P}_{21} = 0$, on a alors

$$\sqrt{n} \text{vec } R_{21} \xrightarrow{\mathcal{L}} \mathcal{N}_{pq}(O, (1 + \kappa)(\mathcal{P}_{11} \otimes \mathcal{P}_{22})).$$

(iii) D'après la partie (i), $\sqrt{n} \text{vec } R_{21}$ converge en loi vers une loi multinormale dont la matrice de covariance est $(1 + \kappa) \mathcal{P}_{11} \otimes \mathcal{P}_{22}$, puisque tous les termes de la matrice D , autre que $\mathcal{P}_{11} \otimes \mathcal{P}_{22}$, contiennent \mathcal{P}_{12} ou sa transposée et que les éléments de cette matrice sont par hypothèse de la forme a_{ij}/\sqrt{n} , pour $i = p + 1, \dots, m$ et $j = 1, \dots, p$.

D'autre part, en substituant la valeur de \mathcal{P}_{21} , on déduit que $\sqrt{n} \text{vec } R_{21}$ converge en loi vers une loi multinormale de moyenne $\text{vec } A$. Par conséquent, par Serfling (1980) (voir lemme A, p.20), $\sqrt{n} \text{vec } R_{21}$ converge en loi vers une loi multinormale $\mathcal{N}_{pq}(\text{vec } A, (1 + \kappa) \mathcal{P}_{11} \otimes \mathcal{P}_{22})$.

c.q.f.d.

On remarque que pour $m = 2$ et pour ρ et r désignant respectivement le coefficient de corrélation de la population et de l'échantillon, on a

$$D = \frac{1}{2} [\rho^2 + \rho^2 + \rho^4 + \rho^4] - \rho^2 - \rho^2 - \rho^2 - \rho^2 + 1 + \rho^2 = (1 - \rho^2)^2$$

et on retrouve le résultat $\sqrt{n}(r - \rho) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (1 + \kappa)(1 - \rho^2)^2)$ (voir Muirhead (1982)).

5. Distributions asymptotiques des mesures

Avec les résultats de la section précédente, on peut établir les lois asymptotiques des $nRV^{(R)}$, $nSL^{(R)}$ et $nCN^{(R)}$ sous H_0 et sous $H_{1,n}$ convergeant vers H_0 .

THÉORÈME 3. Soit R la matrice de corrélation d'un échantillon de taille n issu d'un vecteur aléatoire $X = (X_1, \dots, X_m)'$ dont la loi est elliptique de paramètres μ et V , de matrice de corrélation \mathcal{P} et de coefficient d'aplatissement κ . Alors, sous l'hypothèse H_0 , on a

1. $nRV^{(R)} \xrightarrow{\mathcal{L}} \frac{1 + \kappa}{\sqrt{\text{tr}(\mathcal{P}_{11}^2) \text{tr}(\mathcal{P}_{22}^2)}} \sum_{k=1}^p \sum_{j=1}^q \lambda_k \mu_j U_{kj}^2$, où les variables aléatoires U_{kj} sont indépendantes et de même loi $\mathcal{N}(0, 1)$, $k = 1, \dots, p$ et $j = 1, \dots, q$,
2. $nSL^{(R)} \xrightarrow{\mathcal{L}} \frac{1 + \kappa}{p} \sum_{k=1}^p \lambda_k Z_k^2$, où les variables aléatoires Z_k^2 , $k = 1, \dots, p$, sont indépendantes et de même loi χ^2 avec q degrés de liberté,
3. $nCN^{(R)} \xrightarrow{\mathcal{L}} \frac{1 + \kappa}{p} \chi_{pq}^2$,

avec λ_k , $k = 1, \dots, p$, et μ_j , $j = 1, \dots, q$, étant les valeurs propres de \mathcal{P}_{11} et \mathcal{P}_{22} respectivement.

Démonstration. Puisque les matrices R_{11} et R_{22} convergent en probabilité respectivement vers \mathcal{P}_{11} et \mathcal{P}_{22} , $\sqrt{n} \text{vec } R_{21}$ converge en loi vers le vecteur aléatoire Z de loi $\mathcal{N}_{pq}(O, (1+\kappa)(\mathcal{P}_{11} \otimes \mathcal{P}_{22}))$ et comme on peut écrire $n \text{tr}(R_{12}R_{21}) = (\sqrt{n} \text{vec } R_{21})'(\sqrt{n} \text{vec } R_{21})$,

$$nRV^{(R)} \xrightarrow{\mathcal{L}} \frac{Z'Z}{\sqrt{\text{tr}(\mathcal{P}_{11}^2) \text{tr}(\mathcal{P}_{22}^2)}} = \frac{(1 + \kappa)}{\sqrt{\text{tr}(\mathcal{P}_{11}^2) \text{tr}(\mathcal{P}_{22}^2)}} \sum_{k=1}^p \sum_{j=1}^q \lambda_k \mu_j U_{kj}^2$$

avec λ_k , $k = 1, \dots, p$, et μ_j , $j = 1, \dots, q$, étant les valeurs propres de \mathcal{P}_{11} et \mathcal{P}_{22} respectivement et les variables aléatoires U_{kj} sont indépendantes et de même loi $\mathcal{N}(0, 1)$, $k = 1, \dots, p$ et $j = 1, \dots, q$. La démonstration s'achève en remarquant que

$$n \text{tr}(R_{12}R_{22}^{-1}R_{21}) = (\sqrt{n} \text{vec } R_{21})'(I_p \otimes R_{22}^{-1})(\sqrt{n} \text{vec } R_{21}) \xrightarrow{\mathcal{L}} Z'(I_p \otimes \mathcal{P}_{22}^{-1})Z,$$

$$n \text{tr}(R_{11}^{-1}R_{12}R_{22}^{-1}R_{21}) = (\sqrt{n} \text{vec } R_{21})'(R_{11} \otimes R_{22})^{-1}(\sqrt{n} \text{vec } R_{21}) \text{ et}$$

$$(\sqrt{n} \text{vec } R_{21})'(R_{11} \otimes R_{22})^{-1}(\sqrt{n} \text{vec } R_{21}) \xrightarrow{\mathcal{L}} Z'(\mathcal{P}_{11} \otimes \mathcal{P}_{22})^{-1}Z.$$

c.q.f.d.

THÉORÈME 4. Si les conditions du théorème 3 sont satisfaites, alors sous $H_{1:n}$, on a

1. $nRV^{(R)} \xrightarrow{\mathcal{L}} \frac{1 + \kappa}{\sqrt{\text{tr}(\mathcal{P}_{11}^2) \text{tr}(\mathcal{P}_{22}^2)}} \sum_{k=1}^p \sum_{j=1}^q \lambda_k \mu_j U_{kj}^2$, où les variables aléatoires U_{kj} sont indépendantes et de loi $\mathcal{N}(\delta_{kj}, 1)$, $k = 1, \dots, p$ et $j = 1, \dots, q$,

2. $nSL^{(R)} \xrightarrow{\mathcal{L}} \frac{1+\kappa}{p} \sum_{k=1}^p \lambda_k \mathcal{X}_{k(q)}^2(\delta_k^2)$, où les variables aléatoires $\mathcal{X}_{k(q)}^2(\delta_k^2)$, $k = 1, \dots, p$, sont indépendantes et suivent des lois \mathcal{X}^2 décentrées avec q degrés de liberté et paramètre de décentralité δ_k^2 ,
3. $nCN^{(R)} \xrightarrow{\mathcal{L}} \frac{1+\kappa}{p} \mathcal{X}_{pq}^2(\delta^2)$, où la variable aléatoire $\mathcal{X}_{pq}^2(\delta^2)$ suit une loi \mathcal{X}^2 avec pq degrés de liberté et paramètre de décentralité δ^2 ,
- les λ_k , $k = 1, \dots, p$, et μ_j , $j = 1, \dots, q$, étant les valeurs propres de \mathcal{P}_{11} et \mathcal{P}_{22} respectivement, a_k et b_j leurs vecteurs propres orthonormés associés et

$$\delta_{kj}^2 = \frac{\text{tr}(A'b_j b_j' \mathcal{P}_{22}^{-1} A \mathcal{P}_{11}^{-1} a_k a_k')}{1+\kappa}, \quad k = 1, \dots, p \text{ et } j = 1, \dots, q,$$

$$\delta_k^2 = \frac{\text{tr}(A' \mathcal{P}_{22}^{-1} A \mathcal{P}_{11}^{-1} a_k a_k')}{1+\kappa}, \quad k = 1, \dots, p,$$

$$\delta^2 = \frac{\text{tr}(A' \mathcal{P}_{22}^{-1} A \mathcal{P}_{11}^{-1})}{1+\kappa}.$$

Démonstration. La démonstration est analogue à celle du théorème 3 avec un paramètre de décentralité introduit puisque le vecteur de la moyenne de la loi asymptotique de $\sqrt{n} \text{vec } R_{21}$ est non nul.

En effet, désignons par a_k et b_j les vecteurs propres orthonormés associés respectivement à λ_k et μ_j . Le vecteur $v_{kj} = a_k \otimes b_j$ correspond alors au vecteur propre orthonormé associé à la valeur propre $\lambda_k \mu_j$. On a alors (voir Baldessari (1967)) le paramètre de décentralité

$$\begin{aligned} \delta_{kj}^2 &= \frac{(\text{vec } A)' [v_{kj} v_{kj}' (\mathcal{P}_{11}^{-1} \otimes \mathcal{P}_{22}^{-1})] (\text{vec } A)}{1+\kappa} \\ &= \frac{(\text{vec } A)' [(a_k \otimes b_j) (a_k' \otimes b_j') (\mathcal{P}_{11}^{-1} \otimes \mathcal{P}_{22}^{-1})] (\text{vec } A)}{1+\kappa} \\ &= \frac{(\text{vec } A)' (a_k a_k' \mathcal{P}_{11}^{-1} \otimes b_j b_j' \mathcal{P}_{22}^{-1}) (\text{vec } A)}{1+\kappa} \\ &= \frac{\text{tr}(A' b_j b_j' \mathcal{P}_{22}^{-1} A \mathcal{P}_{11}^{-1} a_k a_k')}{1+\kappa} \end{aligned}$$

avec $k = 1, \dots, p$ et $j = 1, \dots, q$.

Pour SL , $b_j = e_j = (0, \dots, 0, 1, 0, \dots, 0)' \in \mathbb{R}^q$ est un vecteur propre associé à la valeur propre 1 de multiplicité q ; on a le résultat en sommant sur j les δ_{kj}^2 car U_{kj} , $k = 1, \dots, p$, sont indépendantes. Pour CN , $a_k = e_k = (0, \dots, 0, 1, 0, \dots, 0)' \in \mathbb{R}^p$ est un vecteur propre associé à la valeur propre 1 de multiplicité p ; le résultat s'obtient en effectuant la double sommation.

c.q.f.d.

6. Tests et applications

Le théorème 3 permet de construire trois tests asymptotiques pour tester l'hypothèse nulle H_0 . À l'aide des mesures d'Escoufier (1973), de Stewart & Love (1968) et Cramer & Nicewander (1979), les régions critiques des tests de niveau asymptotique α sont donnés respectivement par

$$\begin{aligned} &\text{rejeter } H_0 \text{ si } nRV^{(R)} > C_\alpha^{(1)}, \text{ rejeter } H_0 \text{ si } nSL^{(R)} > C_\alpha^{(2)} \\ &\text{ou rejeter } H_0 \text{ si } nCN^{(R)} > C_\alpha^{(3)} \end{aligned}$$

où $C_\alpha^{(h)}$, pour $h = 1, 2, 3$, est le quantile d'ordre $1 - \alpha$ de la loi correspondante donnée par le théorème 3.

On a $M^{(R)} \xrightarrow{\mathcal{P}} \rho M^{(R)}$ lorsque $n \rightarrow \infty$, pour $M = RV, SL$ ou CN . Lorsque H_0 est fautive, $\rho M^{(R)} > 0$ et alors pour $M = RV, SL$ ou CN et $h = 1, 2, 3$, on conclut que $\lim_{n \rightarrow \infty} P(nM^{(R)} > C_\alpha^{(h)}) = 1$. Les trois tests proposés sont donc convergents.

Cependant, la fonction de répartition d'une somme pondérée de lois χ^2 indépendantes ne possède pas une forme explicite. On évalue numériquement les quantiles donnés par le théorème 3 à l'aide de l'algorithme d'Imhof (1961) dont un programme en Fortran est donné dans Koerts & Abrahams (1969).

Puisqu'en général certains paramètres sont inconnus, on les remplace par des estimateurs convergents, c'est-à-dire \mathcal{P}_{11} , \mathcal{P}_{22} , μ_l et λ_k sont remplacés par R_{11} , R_{22} et leur valeurs propres respectives $\hat{\mu}_l$ et $\hat{\lambda}_k$. Dans une étude faite, dans un cadre général, par Roy & Cléroux (1993) la précision des valeurs critiques obtenues par l'algorithme d'Imhof (1961) en remplaçant les paramètres inconnus par des estimateurs convergents est analysée. Les auteurs montrent que le niveau obtenu en utilisant l'algorithme d'Imhof (1961) converge en probabilité, lorsque $n \rightarrow \infty$, vers le niveau nominal.

Un exemple. L'exemple qu'on considère ici a été traité par Cléroux & Ducharme (1989) en utilisant la mesure d'Escoufier (1973) basée sur la matrice de covariance. Toutefois, les données indiquent que les variances des variables sont très différentes et n'ont pas la même échelle de mesure. Même si on centre et réduit chaque variable, on ne peut appliquer la mesure d'Escoufier (1973) ou celle de Stewart & Love (1968), basées sur la matrice de covariance échantillonnale parce que ces mesures ne sont pas invariantes par rapport aux transformations linéaires. Mais, basées sur la matrice de corrélation échantillonnale, les trois mesures sont invariantes aux changements d'échelle et de position et par conséquent, elles s'appliquent parfaitement à cet exemple. Cette propriété constitue un avantage important par rapport aux tests paramétriques d'absence de liaison proposés par Cléroux & Ducharme (1989) ou par Allaire & Lepage (1990). D'ailleurs, les derniers auteurs soulèvent ce problème de la nature de la matrice utilisée par les trois mesures.

On considère donc l'ensemble de données de Werner, Tolls, Hultin & Mellecker (1970). Les données sont constituées des résultats d'analyses sanguines de 188 patients. Pour chaque patient, on mesure 8 variables : âge, taille, poids, utilisation de pilule anticonceptionnelle, cholestérol, albumine, calcium et acide urique. La

variable pilule anticonceptionnelle, qui est une variable binaire, n'est pas considérée ici. De plus, 7 vecteurs sont omis à cause d'observations manquantes. Ainsi, la taille de l'échantillon est réduite à $n = 181$. Les 7 variables considérées se regroupent naturellement en deux sous-ensembles de variables. Le premier est désigné par le sous-vecteur $X^{[1]}$ constitué des trois premières variables (biologiques). Le deuxième est désigné par le sous-vecteur $X^{[2]}$ constitué des quatre dernières variables (biochimiques). On a donc deux sous-vecteurs de dimension respective $p = 3$ et $q = 4$ (avec $p + q = m = 7$).

À l'aide de chacun des trois tests proposés, on rejette très fortement la non-association entre les deux sous-vecteurs. Le tableau 1 donne les différentes valeurs des statistiques correspondantes, leurs points critiques au niveau nominal 5 % et les niveaux critiques.

TABLEAU 1
*Tests d'absence de liaison basés sur la matrice de corrélation,
points critiques et niveaux critiques*

	$nRV^{(R)}$	$nSL^{(R)}$	$nCN^{(R)}$
Valeur des statistiques	16,63	22,48	20,12
Point critique $C_{0,05}^{(h)}$	6,83	9,04	8,62
Niveau critique	$4,06 \times 10^{-5}$	$3,57 \times 10^{-6}$	$7,51 \times 10^{-6}$

7. Étude expérimentale

À l'aide d'une simulation, on va évaluer le comportement des trois tests développés précédemment et les comparer aux trois tests compétiteurs, basés sur la matrice de covariance $nM^{(C)}$, $M = RV, SL$ et CN , étudiés par Allaire & Lepage (1990). Rappelons que $CN^{(R)} = CN^{(C)}$.

On utilise la méthode de Monte Carlo pour comparer le niveau et la puissance expérimentalement. Cette étude est réalisée à titre indicatif puisqu'on se limite seulement au cas $p = 2$, $q = 2$ ($p + q = m$) et au niveau 5 %. Le nombre d'échantillons générés est 10000 pour trois types de lois multivariées elliptiques : une loi multinormale, deux lois multinormales ϵ -contaminées ($\epsilon = 0,05$ et $0,1$) et deux lois t multivariées de degrés de liberté 25 et 5 respectivement. Elles sont utilisées avec la même matrice de covariance Σ telle que

$$\Sigma_{11} = I_p, \Sigma_{22} = I_q \text{ et } \Sigma_{12} = \Sigma'_{21} = C_{00}, C_{10}, C_{15} \text{ et } C_{20}$$

où les matrices C_{xy} représentent les matrices $p \times q$ dont tous les éléments sont égaux à 0, xy . Par exemple, tous les éléments de C_{15} sont égaux à 0,15. Ce type de matrices de covariance a été utilisé par Cléroux, Lazraq & Lepage (1995).

Ainsi, les lois utilisées sont les suivantes. La loi multinormale $\mathcal{N}_4(0, \Sigma)$ avec un coefficient d'aplatissement $\kappa = 0$. La loi multinormale ϵ -contaminée peut s'écrire comme (voir Johnson (1987), p.55)

$$(1 - \epsilon)\mathcal{N}_4(O, V) + \epsilon\mathcal{N}_4(O, \sigma^2 V) \text{ avec } \Sigma = (1 - \epsilon + \epsilon\sigma^2)V.$$

En fixant $\sigma^2 = 0,25$ et en considérant $\epsilon = 0,05$ et $\epsilon = 0,1$, on obtient respectivement $\kappa = 0,028$ et $\kappa = 0,059$ car

$$\kappa = \frac{1 + \epsilon(\sigma^4 - 1)}{(1 + \epsilon(\sigma^2 - 1))^2} - 1$$

(voir Muirhead (1982), p.49). La loi t multivariée avec ν degrés de liberté et paramètre V s'obtient (voir Johnson (1987), p. 118) par

$$Y = \left(\sqrt{\frac{S}{\nu}} \right)^{-1} Z \tag{7.4}$$

où Z suit une loi $\mathcal{N}_4(0, V)$, S suit une loi χ^2_ν , Z et S sont indépendantes, $\Sigma = \nu V/(\nu - 2)$ et $\kappa = 2/(\nu - 4)$. Par conséquent, pour $\nu = 25$ et $\nu = 5$ on trouve respectivement $\kappa = 0,095$ et $\kappa = 2$. On a choisi $\nu = 5$ afin de considérer une loi dont les ailes sont plus relevées que celles d'une loi multinormale.

Afin de générer des nombres aléatoires, on a utilisé le programme GGNSM provenant de la programmation IMSL. On a ainsi généré 10000 échantillons indépendants selon la loi multinormale, loi multinormale ϵ -contaminée ($\epsilon = 0,05$ et $\epsilon = 0,1$) et les lois χ^2_{25} et χ^2_5 pour obtenir les observations désirées. Des tailles échantillonnales de 50, 100, 150, 200, 250, 300 et 1000 sont considérées.

Les tableaux 2 et 3 présentent les résultats de la simulation pour les cinq lois. Les lois y sont disposées selon l'ordre croissant de leur coefficient d'aplatissement. De la loi multinormale à la loi t_5 , on a $\kappa = 0,028, 0,059, 0,095$ et 2 . Afin de pouvoir juger du niveau expérimental des cinq tests asymptotiques et de leur puissance expérimentale, un niveau expérimental sera jugé acceptable si le niveau nominal 5 % appartient à l'intervalle de confiance au niveau 95 %. Ceci sera réalisé lorsque le niveau expérimental, correspondant à la colonne C_{00} , varie entre $456 \cdot 10^{-4}$ et $544 \cdot 10^{-4}$.

Les résultats de cette étude font apparaître que pour des lois présentant des ailes plus relevées que celles d'une loi multinormale, les niveaux expérimentaux ne sont acceptables que pour de très grandes tailles échantillonnales. Pour ce type de lois, les cinq tests asymptotiques convergent lentement, en fonction de la taille n , vers un niveau acceptable. Cependant, les tests basés sur la matrice de corrélation échantillonnale convergent plus rapidement que ceux basés sur la matrice de covariance échantillonnale. Ceci est mis en évidence par l'étude expérimentale avec la loi t_5 dont le coefficient d'aplatissement est 2. On peut le constater en examinant les résultats expérimentaux pour la loi t_5 donnés dans le tableau 2 et notamment dans le tableau 3 pour la taille 1000. Il ressort donc de cette étude expérimentale qu'on devrait utiliser ces tests avec prudence lorsque la population sous-jacente suit une loi elliptique avec des ailes plus relevées que celles d'une loi multinormale.

TABLEAU 2
*Puissance expérimentale ($\times 10000$) des tests au niveau 5%
pour des lois multivariées de structure de covariance donnée avec $p = 2$ et $q = 2$.*

n	Tests	Multinormale						0,05-contaminée						0,1-contaminée						t avec $\nu=25$						t avec $\nu=5$					
		C_{00}	C_{10}	C_{15}	C_{20}	C_{00}	C_{10}	C_{15}	C_{20}	C_{00}	C_{10}	C_{15}	C_{20}	C_{00}	C_{10}	C_{15}	C_{20}	C_{00}	C_{10}	C_{15}	C_{20}	C_{00}	C_{10}	C_{15}	C_{20}	C_{00}	C_{10}	C_{15}	C_{20}		
50	nRV ^(C)	587	1854	3749	6207	596	1839	3707	6135	634	1884	3692	6148	625	1828	3644	5935	892	1705	2925	4584	892	1705	2925	4584	892	1705	2925	4584		
	nSL ^(C)	601	1853	3773	6301	608	1826	3689	6179	666	1860	3652	6209	619	1777	3580	6028	691	1479	2648	4504	691	1479	2648	4504	691	1479	2648	4504		
	nCN ^(C)	579	1818	3759	6412	625	1800	3713	6340	672	1823	3677	6275	603	1781	3557	6126	562	1345	2486	4371	562	1345	2486	4371	562	1345	2486	4371		
100	nRV ^(R)	581	1837	3701	6178	622	1814	3684	6088	644	1836	3614	6067	619	1791	3605	5881	697	1512	2698	4391	697	1512	2698	4391	697	1512	2698	4391		
	nSL ^(R)	578	1838	3769	6296	625	1805	3705	6164	648	1846	3641	6181	621	1794	3594	6005	638	1427	2562	4380	638	1427	2562	4380	638	1427	2562	4380		
	nCN ^(R)	579	1818	3759	6412	625	1800	3713	6340	672	1823	3677	6275	603	1781	3557	6126	562	1345	2486	4371	562	1345	2486	4371	562	1345	2486	4371		
150	nRV ^(C)	539	3324	6752	9219	532	3171	6541	9109	611	3207	6429	9030	595	3158	6358	8942	794	2361	4503	6941	794	2361	4503	6941	794	2361	4503	6941		
	nSL ^(C)	540	3331	6767	9301	534	3165	6631	9200	599	3216	6483	9131	568	3145	6418	9026	667	2219	4437	6981	667	2219	4437	6981	667	2219	4437	6981		
	nCN ^(C)	544	3311	6832	9376	532	3188	6683	9299	593	3197	6577	9228	572	3087	6476	9152	556	2072	4380	7078	556	2072	4380	7078	556	2072	4380	7078		
150	nRV ^(R)	551	3299	6689	9205	558	3174	6543	9299	609	3212	6420	9025	575	3092	6333	8951	682	2202	4385	6879	682	2202	4385	6879	682	2202	4385	6879		
	nSL ^(R)	547	3296	6747	9302	532	3172	6635	9203	596	3207	6486	9130	557	3115	6413	9018	623	2128	4377	6943	623	2128	4377	6943	623	2128	4377	6943		
	nCN ^(R)	544	3311	6832	9376	532	3188	6683	9299	593	3197	6577	9228	572	3087	6476	9152	556	2072	4380	7078	556	2072	4380	7078	556	2072	4380	7078		
150	nRV ^(C)	532	4756	8583	9881	535	4667	8448	9868	552	4493	8327	9842	557	4426	8263	9812	770	3047	5866	8371	770	3047	5866	8371	770	3047	5866	8371		
	nSL ^(C)	538	4786	8637	9898	541	4654	8502	9894	569	4541	8363	9875	543	4440	8327	9836	637	2948	5861	8429	637	2948	5861	8429	637	2948	5861	8429		
	nCN ^(C)	535	4814	8722	9907	542	4658	8583	9912	558	4524	8425	9897	535	4465	8390	9868	578	2832	5851	8515	578	2832	5851	8515	578	2832	5851	8515		
150	nRV ^(R)	527	4767	8595	9878	532	4599	8433	9866	566	4486	8301	9848	549	4387	8274	9810	652	2942	5773	8371	652	2942	5773	8371	652	2942	5773	8371		
	nSL ^(R)	538	4776	8632	9893	543	4634	8503	9893	579	4503	8375	9872	541	4422	8345	9838	609	2895	5821	8426	609	2895	5821	8426	609	2895	5821	8426		
	nCN ^(R)	535	4814	8722	9907	542	4658	8583	9912	558	4524	8425	9897	535	4465	8390	9868	578	2832	5851	8515	578	2832	5851	8515	578	2832	5851	8515		

n	Tests	Multinormale						0,05-contaminée						0,1-contaminée						t avec $\nu=25$						t avec $\nu=5$					
		C_{00}	C_{10}	C_{15}	C_{20}	C_{00}	C_{10}	C_{15}	C_{20}	C_{00}	C_{10}	C_{15}	C_{20}	C_{00}	C_{10}	C_{15}	C_{20}	C_{00}	C_{10}	C_{15}	C_{20}	C_{00}	C_{10}	C_{15}	C_{20}						
200	nRV ^(C)	522	6138	9480	9985	517	6043	9423	9979	516	5808	9345	9979	516	5808	9345	9979	537	5685	9264	9974	692	3707	7053	9165						
	nSL ^(C)	512	6167	9515	9988	528	6068	9458	9984	518	5876	9377	9987	518	5876	9377	9987	533	5700	9301	9981	628	3658	7061	9228						
	nCN ^(C)	522	6194	9533	9992	536	6107	9482	9993	523	5852	9414	9989	523	5852	9414	9989	521	5731	9339	9984	544	3602	7060	9299						
	nRV ^(R)	514	6120	9479	9988	535	6048	9421	9979	510	5824	9328	9980	510	5824	9328	9980	523	5661	9261	9975	626	3619	6992	9170						
	nSL ^(R)	512	6173	9515	9988	533	6089	9447	9985	520	5862	9378	9988	520	5862	9378	9988	516	5697	9299	9981	591	3599	7048	9230						
	nCN ^(R)	522	6194	9533	9992	536	6107	9482	9993	523	5852	9414	9989	523	5852	9414	9989	521	5731	9339	9984	544	3602	7060	9299						
250	nRV ^(C)	527	7182	9799	9999	517	7042	9779	9999	533	6913	9758	9999	533	6913	9758	9999	526	6787	9711	9998	656	4346	7827	9541						
	nSL ^(C)	511	7231	9815	9999	512	7068	9793	9998	536	6940	9776	1000	536	6940	9776	1000	512	6800	9730	9997	575	4325	7890	9547						
	nCN ^(C)	506	7255	9829	10000	526	7054	9812	9999	536	6975	9800	1000	536	6975	9800	1000	499	6818	9742	9999	537	4308	7926	9602						
	nRV ^(R)	517	7202	9802	9999	521	7024	9775	9999	537	6913	9757	9999	537	6913	9757	9999	510	6765	9701	9997	584	4310	7842	9533						
	nSL ^(R)	511	7229	9815	9999	527	7050	9794	9998	537	6941	9778	1000	537	6941	9778	1000	500	6795	9724	9997	554	4277	7893	9547						
	nCN ^(R)	506	7255	9829	10000	526	7054	9812	9999	536	6975	9800	1000	536	6975	9800	1000	499	6818	9742	9999	537	4308	7926	9602						
300	nRV ^(C)	517	8093	9945	10000	493	7935	9935	10000	489	7773	9916	10000	489	7773	9916	10000	496	7691	9914	10000	626	4945	8422	9715						
	nSL ^(C)	516	8122	9946	10000	488	7955	9939	10000	481	7818	9929	10000	481	7818	9929	10000	496	7714	9920	10000	566	4904	8488	9738						
	nCN ^(C)	510	8151	9957	10000	508	7999	9948	10000	475	7871	9937	10000	475	7871	9937	10000	508	7744	9928	10000	508	4845	8520	9757						
	nRV ^(R)	517	8095	9946	10000	486	7909	9937	10000	484	7769	9920	10000	484	7769	9920	10000	505	7680	9911	10000	569	4888	8413	9717						
	nSL ^(R)	512	8119	9950	10000	491	7946	9939	10000	481	7825	9931	10000	481	7825	9931	10000	504	7729	9918	10000	541	4857	8483	9729						
	nCN ^(R)	510	8151	9957	10000	508	7999	9948	10000	475	7871	9937	10000	475	7871	9937	10000	508	7744	9928	10000	508	4845	8520	9757						

TABLEAU 3
Niveau expérimental ($\times 10000$) des tests au niveau nominal 5 %
pour $n = 1000$ et différentes lois.

		Multinor- male	0,05-cont	0,1-cont	t_{25}	t_5
n	Tests	C_{00}	C_{00}	C_{00}	C_{00}	C_{00}
1000	$nRV^{(C)}$	480	478	500	511	587
	$nSL^{(C)}$	474	478	490	512	543
	$nCN^{(C)}$	470	480	493	510	516
	$nRV^{(R)}$	479	477	501	520	544
	$nSL^{(R)}$	482	488	498	509	529
	$nCN^{(R)}$	470	480	493	510	516

Quant à la puissance expérimentale, mis à part le cas de la loi t_5 où les comparaisons ne pourraient se faire que pour de très grandes tailles échantillonnales, on peut conclure que globalement les deux groupes de tests $nM^{(R)}$ et $nM^{(C)}$, pour $M = RV, SL$ et CN , sont équivalents. Aussi, on note que dans les cinq cas, la puissance expérimentale de chaque test croît au fur et à mesure qu'on s'éloigne de l'hypothèse nulle C_{00} , c'est-à-dire que la puissance expérimentale croît avec la valeur xy des matrices C_{xy} . De plus, plus on s'éloigne de l'hypothèse de multinormalité plus la puissance expérimentale des cinq tests diminue. Elle apparaît comme une fonction décroissante du coefficient d'aplatissement.

De façon générale, on peut dire que le test construit avec la mesure de Cramer & Nicewander (1979) (c'est-à-dire les tests $nCN^{(R)} = nCN^{(C)}$) se comporte mieux que ceux construits avec les deux autres mesures. Il stabilise le niveau expérimental à un seuil acceptable, et de plus, sa puissance expérimentale augmente plus rapidement que pour ceux avec les deux autres mesures.

Références

- ALLAIRE J., LEPAGE Y., (1990), Tests de l'absence de liaison entre plusieurs vecteurs aléatoires pour les distributions elliptiques. *Statistique et Analyse des Données*, **15**, 3, 21-46.
- BALDESSARI B., (1967), The Distribution of Quadratic Form of Normal Random Variables. *Ann. Math. Statist.*, **38**, 1700-1704.
- CLÉROUX R., DUCHARME G., (1989), Vector Correlation for Elliptical Distributions. *Commun. Statist.-Theory Meth.*, **18**(4), 1441-1454.
- CLÉROUX R., LAZRAQ A., LEPAGE Y., (1995), Vector Correlation Based on Rank and a Nonparametric Test of no Association Between Vectors. *Commun. Statist.-Theory meth.*, **24**(3), 713-733.

- CRAMER E.M., NICEWANDER G.R., (1979), Some Symetric Invariant Measure of Multivariate Association. *Psychometrika*, **49**, 403-423.
- ESCOUFIER Y., (1973), Le Traitement des variables vectorielles. *Biometrics*, **29**, 751-760.
- IMHOF P., (1961), Computing the Distribution of Quadratic Forms in Normal Variates. *Biometrika*, **48**, 419-426.
- JOHNSON M.E., (1987), *Multivariate Statistical Simulation*, John Wiley, New York.
- KOERTS J., ABRAHAMS A.P.J., (1969), *On the Theory and Application of the General Linear Model*, Rotterdam University Press.
- MUIRHEAD R.J., (1982), *Aspects of Multivariate Statistical Theory*, John Wiley, New York.
- ROY R., CLÉROUX R., (1993), Vector Cross-Correlation in Time Series and Application, *International Statistical Review*, **61**, 3, 447-464.
- SERFLING R.J., (1980), *Approximation Theorems of Mathematical Statistics*, John Wiley, New York.
- STEIGER J.H., HAKSTIAN A.R., (1982), The Asymptotic Distribution of Elements of a Correlation Matrix : Theory and application. *British Journal of Mathematical and Statistical Psychology*, **35**, 208-215.
- STEWART D., LOVE W., (1968), A General Canonical Correlation Index. *Psychological Bulletin*, **70**, 160-163.
- WERNER M., TOLLS M.E.M., HULTIN J.V., MELLECKER J., (1970), Sex and Age Dependence of Serum Calcium, Inorganic Phosphorus, Total Protein and Albumin in a Large Ambulatory Population. *Fifth International Congress on Automation, Advances in Automated Analysis*, **2**, 59-65.