

Université de Montréal

Une classe de tests basés sur des fonctions des
rangs pour l'hypothèse d'absence de liaison entre
plusieurs vecteurs aléatoires

par

Sylvie Gervais

Département de mathématiques et de statistique
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de
Philosophiæ Doctor (Ph.D.)
en Statistique

février 2001

© Sylvie Gervais, 2000



QA

3

US4

2001

v. 003

Université de Montréal

Faculté des études supérieures


Cette thèse intitulée

Une classe de tests basés sur des fonctions des
rangs pour l'hypothèse d'absence de liaison entre
plusieurs vecteurs aléatoires

présentée par

Sylvie Gervais

a été évaluée par un jury composé des personnes suivantes :


(président-rapporteur)

Yves Lepage

(directeur de recherche)

Aziz Lazraq

(co-directeur)

Jean-François Angers

(membre du jury)

Yves Escoufier

(examineur externe)


(représentant du doyen)

Thèse acceptée le :

7 février 2001

SOMMAIRE

Les objectifs de ce travail sont tout d'abord, la construction d'une vaste classe de tests pour l'hypothèse d'absence de liaison entre plusieurs vecteurs aléatoires où chaque membre de la classe est une forme quadratique de la matrice d'association. Après avoir obtenu les distributions asymptotiques conjointes de ces statistiques de tests sous H_0 et sous une suite de contre-hypothèses convergeant vers H_0 , et avoir examiné quelques cas particuliers faisant partie de la classe, on procède à une étude de la performance des tests de la classe.

Pour étudier le comportement asymptotique des tests, on utilise la notion d'efficacité asymptotique relative. On obtient des expressions analytiques de l'efficacité asymptotique relative pour certains cas particuliers de la classe. On propose ensuite une méthode pour obtenir des efficacités asymptotiques relatives de façon numérique qui s'inspire de l'approche de Nyblom et Mäkeläinen (1983).

Finalement, on conclut l'étude de la performance des tests en analysant les niveaux et puissances expérimentaux obtenus à la suite d'une étude de Monte Carlo.

REMERCIEMENTS

Cette thèse est l'aboutissement de plusieurs années de travail parsemées de divers obstacles et défis à relever. Je tiens donc à remercier toutes les personnes qui m'ont aidée directement ou indirectement à accomplir ce travail. Plus particulièrement je tiens à remercier les personnes suivantes.

D'abord mon directeur Yves Lepage qui a su me guider et me soutenir à travers toutes les étapes qui ont mené à l'accomplissement de ce travail. Mon co-directeur, Aziz Lazraq pour ses judicieux conseils et son implication dans ce projet.

J'aimerais remercier aussi les professeurs et étudiants du département de mathématiques et de statistique qui ont contribué de près ou de loin à faire progresser mon travail et tout particulièrement mon ami Denis Larocque.

Je voudrais bien entendu souligner le très important support moral de ma mère Irène, mon père Jean-Claude, mon frère Eddy et ma défunte grand-maman Jeanne-D'arc qui aurait été bien fière de sa petite-fille.

Merci aussi au CRSNG et FCAR pour le support financier.

Je dédis ce travail à mon fils Jérémy ainsi qu'à mon époux Sylvain dont la présence et le support ont été déterminants dans la réussite de ce projet.

Maman a enfin réussi mon trésor!

Table des matières

Sommaire	iv
Remerciements	v
Table des figures	ix
Liste des tableaux	x
Introduction	1
Chapitre 1. Une classe de tests basés sur des fonctions des rangs	8
1.1. La matrice d'association	8
1.2. L'hypothèse d'indépendance	12
1.3. Définition de la classe de statistiques	16
Chapitre 2. Distributions asymptotiques des statistiques de la classe \mathcal{C}	21
2.1. Distribution asymptotique de T	21
2.2. Distribution asymptotique conjointe des R_{kl} sous $H_0^{(K)}$	23
2.3. Tests d'indépendance entre plusieurs vecteurs aléatoires	34
2.4. Distribution asymptotique des statistiques sous une suite de contre- hypothèses	38

Chapitre 3. Étude de cas particuliers	46
3.1. Quelques cas particuliers pour la mesure d'association	46
3.1.1. La mesure de Stewart et Love (1968)	46
3.1.2. La mesure d'Escoufier (1973)	50
3.1.3. La mesure de Cramer et Nicewander (1979)	54
3.2. Quelques cas particuliers pour la fonction des rangs	57
3.2.1. Les cotes linéaires	57
3.2.2. Les cotes normales	59
Chapitre 4. Comparaison de la performance des tests	62
4.1. L'efficacité asymptotique relative	62
4.2. La fonction de puissance des tests de la classe \mathcal{C}	66
4.3. Comparaison asymptotique dans certains cas particuliers	70
4.3.1. Classe de tests basés sur la mesure de Cramer et Nicewander (1979)	70
4.3.2. Approche de Nyblom et Mäkeläinen (1983)	91
4.3.3. Quelques comparaisons numériques	95
Chapitre 5. Puissances et niveaux expérimentaux des tests	109
5.1. Méthodologie	110
5.2. Résultats expérimentaux	114
5.3. Interprétation des résultats	115
Conclusion	143
Annexe A. Généralisation du théorème 2.2	145

Annexe B. Contiguïté sous l'hypothèse de multinormalité 149
Bibliographie 160

Table des figures

5.1	Comparaison des tests pour la loi multinormale avec C00	122
5.2	Comparaison des tests pour la loi multinormale avec C15	123
5.3	Comparaison des tests pour la loi multinormale avec C20	124
5.4	Comparaison des tests pour la loi multinormale avec C25	125
5.5	Comparaison des tests pour la loi logistique avec C00	126
5.6	Comparaison des tests pour la loi logistique avec C15	127
5.7	Comparaison des tests pour la loi logistique avec C20	128
5.8	Comparaison des tests pour la loi logistique avec C25	129
5.9	Comparaison des tests pour la loi générale avec C00	130
5.10	Comparaison des tests pour la loi générale avec C15	131
5.11	Comparaison des tests pour la loi générale avec C20	132
5.12	Comparaison des tests pour la loi générale avec C25	133

Liste des tableaux

4.1	Valeurs propres et vecteurs propres des matrices Θ_{11} , Θ_{22} , $\Theta_{11} \otimes \Theta_{22}$ et $\Theta_{11} \otimes I_2$	97
4.2	Valeurs des paramètres θ_{ij} et a_{ij} pour les cotes normales et linéaires lorsque X est de loi multinormale	101
4.3	Efficacités asymptotiques relatives pour X de loi multinormale et $K = 2$	102
4.4	Valeurs des paramètres θ_{ij} et a_{ij} pour les cotes normales et linéaires lorsque X est de loi logistique	105
4.5	Efficacités asymptotiques relatives pour X de loi logistique et $K = 2$.	106
4.6	Efficacités asymptotiques relatives pour X de loi multinormale, $K = 3$	107
4.7	Efficacités asymptotiques relatives pour X de loi logistique, $K = 3$...	107
5.1	Niveaux et puissances expérimentaux pour X de loi multinormale avec $K = 2$ et $n = 50$	114
5.2	Niveaux et puissances expérimentaux pour X de loi multinormale avec $K = 3$ et $n = 50$	115
5.3	Niveaux et puissances expérimentaux pour X de loi multinormale avec $K = 2$ et $n = 100$	116
5.4	Niveaux et puissances expérimentaux pour X de loi multinormale avec $K = 3$ et $n = 100$	117

5.5	Niveaux et puissances expérimentaux pour X de loi logistique multivariée avec $K = 2$ et $n = 50$	117
5.6	Niveaux et puissances expérimentaux pour X de loi logistique multivariée avec $K = 3$ et $n = 50$	118
5.7	Niveaux et puissances expérimentaux pour X de loi logistique multivariée avec $K = 2$ et $n = 100$	118
5.8	Niveaux et puissances expérimentaux pour X de loi logistique multivariée avec $K = 3$ et $n = 100$	119
5.9	Niveaux et puissances expérimentaux pour X de loi générale multivariée avec $K = 2$ et $n = 50$	119
5.10	Niveaux et puissances expérimentaux pour X de loi générale multivariée avec $K = 3$ et $n = 50$	120
5.11	Niveaux et puissances expérimentaux pour X de loi générale multivariée avec $K = 2$ et $n = 100$	120
5.12	Niveaux et puissances expérimentaux pour X de loi générale multivariée avec $K = 3$ et $n = 100$	121
5.13	Tableau d'analyse de variance à un facteur ($n=216$).....	135
5.14	Valeurs-p pour les analyses de différences significatives entre les niveaux obtenus pour les 9 différents tests	136
5.15	Tableau d'analyse de variance à un facteur (structure des sous-vecteurs avec $n=50$, $C15$ et la loi multinormale)	138

INTRODUCTION

Le problème de l'indépendance des différentes composantes d'un vecteur constitue une partie importante de l'analyse multidimensionnelle. Outre l'intérêt purement théorique, le recours à cette théorie est très courant dans plusieurs disciplines scientifiques, entre autre, dans le domaine des sciences sociales, économiques ou encore biologiques. Dès qu'on s'intéresse à l'étude du lien pouvant exister entre plusieurs groupes de variables, les mesures et les tests de liaison sont des outils de choix pour répondre à nos interrogations.

Dans un premier temps, définissons le problème de façon plus formelle. On considère un vecteur aléatoire de dimension p composé de K sous-vecteurs $X = (X^{(1)'} \dots X^{(K)'})'$ où le sous-vecteur $X^{(i)}$ est de dimension p_i pour $i = 1, \dots, K$ et $\sum_{i=1}^K p_i = p$, avec $\mu^{(i)} = E(X^{(i)})$ et $\sum_{i,j} = E(X^{(i)} - \mu^{(i)})(X^{(j)} - \mu^{(j)})'$ pour $i, j = 1, \dots, K$ où A' représente la transposée de la matrice A .

On se questionne à savoir s'il existe un lien entre les K sous-vecteurs de X . On peut chercher à mesurer ce lien à l'aide de mesures d'association vectorielle. De telles mesures existent en grand nombre dans la littérature, chacune ayant ses caractéristiques propres. Comme par exemple, certaines d'entre elles sont fonction des corrélations canoniques; c'est le cas notamment des mesures d'Hotelling (1936), de Cramer (1974), Roseboom (1965), Stewart et Love (1968), Coxhead (1974), Shaffer et Gillo (1974), Cramer et Nicewander (1979) ou encore Ramsey, ten Berge et Styan (1984).

D'autres mesures sont plutôt basées sur le fait que deux vecteurs sont parfaitement corrélés si une transformation orthogonale permet de les faire coïncider, c'est le cas des mesures d'Escoufier (1973) et de Stephens (1979).

Les mesures d'association vectorielle permettent de quantifier le degré de liaison entre les K sous-vecteurs, mais c'est un test d'indépendance ou de non association qui permettra de vérifier cette hypothèse. Plusieurs auteurs se sont penchés sur cette question et un bon nombre d'approches ont été proposées. Soulignons en passant que certains de ces tests sont des tests d'indépendance alors que d'autres sont plutôt des tests d'absence de liaison ou de non association. Faisons un bref survol de la littérature à ce sujet.

Mentionnons d'abord quelques tests paramétriques. Si X suit une loi multinormale, on peut utiliser le test d'indépendance bien connu, le test du rapport de vraisemblance maximale qu'on peut retrouver, par exemple, dans le livre d'Anderson (1984). Toujours dans le cas de la multinormalité, en 1988, Lazraq et Cléroux (1988) présentent des tests d'indépendance basés sur les corrélations canoniques, sur la mesure de Stewart et Love (1968) et sur la mesure d'Escoufier (1973). Dans un cadre plus général, si la distribution du vecteur X appartient à la famille de distributions elliptiques, des tests basés sur les mesures de Stewart et Love (1968), Escoufier (1973) et Cramer et Nicewander (1979) ont été présentés par Allaire et Lepage (1990).

Ces tests d'absence de liaison sont essentiellement basés sur la matrice de covariance Σ . Ainsi, lorsque X est de loi autre que la loi multinormale, la performance des tests est affectée puisque comme on le sait, la matrice de covariance n'est alors plus un outil aussi performant pour mesurer la liaison entre vecteurs.

Une alternative intéressante à l'utilisation de la matrice de covariance classique est d'utiliser la matrice de corrélation. C'est d'ailleurs ce qu'on fait El Maâche

et Lepage (1998) qui ont dérivé des tests d'absence de liaison dans le cas où la distribution du vecteur X appartient à la famille de distributions elliptiques pour $K = 2$ sous-vecteurs.

Une alternative intéressante à la matrice de covariance classique est d'utiliser la matrice d'association qui est composée des paramètres d'association, que l'on définira de façon précise au chapitre 1. Notons simplement pour le moment que ce sont des fonctionnelles régulières de la distribution de X , $F(x)$ et sont définis pour une classe plus vaste de distributions. Il s'agit d'une approche non paramétrique où l'on utilise certaines fonctions des rangs des observations. D'autres avantages sont liés à l'utilisation de tests basés sur de telles fonctionnelles régulières, entre autre, l'élimination des limites quant à la distribution des observations, le fait que les données peuvent être de type ordinal et être observées selon des échelles de mesure différentes, et enfin, rappelons que l'approche basée sur les rangs est plus robuste vis-à-vis des observations aberrantes (Cléroux, Lazraq et Lepage (1994)).

Ainsi, plusieurs tests dans la littérature sont basés sur la matrice d'association. Mentionnons d'abord, Puri, Sen et Gokhale (1970) qui ont développé des tests d'absence d'association entre plusieurs vecteurs aléatoires en remplaçant la matrice de covariance par la matrice d'association dans la définition de la statistique du test du rapport de vraisemblance maximale (Anderson (1984)). Dans leur article, les auteurs ont comparé les tests obtenus au test classique pour deux fonctions des rangs. On relate ces résultats au chapitre 3. Pour le moment, notons simplement que les tests obtenus sont convergents et relativement efficaces.

Mentionnons aussi les travaux de Cléroux *et al* (1994) et Cléroux, Lazraq et Lepage (1995) dans lesquels les auteurs ont introduit des tests basés sur les mesures de Stewart et Love (1968) et d'Escoufier (1973) selon l'approche de Allaire et Lepage (1990) (inspirée des travaux de Lazraq et Cléroux (1988)), en utilisant

la matrice d'association dans le cas spécifique où les fonctions des rangs sont les cotes linéaires. Les auteurs ont étudié le cas où $K = 2$ et $K = 3$ et ont obtenus des niveaux et puissances expérimentaux.

Avant de présenter la classe de tests dont il est question dans le présent travail, soulignons aussi les récents travaux de Dauxois et Nkiet (1997) et Dauxois et Nkiet (1998) qui présentent une classe de tests basés sur des corrélations canoniques linéaires qui mène à des tests d'absence d'association (1997) et une classe de tests basés sur des corrélations canoniques non linéaires qui mène elle, à des tests d'indépendance (1998). Le cas $K = 2$ sous-vecteurs est considéré dans les deux cas et aucune étude de la performance des tests n'a été présentée.

L'objectif de ce travail est de construire une vaste classe de tests définis à partir de la matrice d'association, qui englobe plusieurs tests existants. Dans un deuxième temps, d'étudier la performance de ces tests en comparant les tests tant à l'intérieur de la classe qu'avec des tests existants ne faisant pas partie de la classe.

Dans le premier chapitre, nous définissons la classe de tests. Brièvement, pour qu'une statistique fasse partie de la classe, elle doit pouvoir s'exprimer comme une forme quadratique de la matrice d'association et satisfaire une série de conditions dont nous discuterons amplement au chapitre 1. Mais pour le moment, soulignons simplement que l'approche utilisée pour construire la classe de tests est la même que celle utilisée par Cléroux *et al* (1994) et Cléroux *et al* (1995), mais en généralisant le choix de la mesure d'association considérée et les fonctions de cotes utilisées pour définir la matrice d'association. Par exemple, en considérant les cotes linéaires et la mesure de Stewart et Love (1968), on retrouve le test proposé par les auteurs dans l'article de Cléroux *et al* (1994).

Au chapitre 2, on détermine la distribution asymptotique des tests de la classe sous l'hypothèse d'absence de liaison et sous une suite de contre-hypothèses qui convergent vers H_0 ¹. On y discute aussi des conditions très particulières sous lesquelles les tests peuvent être considérés comme des tests d'indépendance. Les techniques de démonstrations sont analogues à celles de Cléroux *et al* (1994) et Cléroux *et al* (1995) mais en utilisant un résultat plus général d'un théorème de Puri *et al* (1970).

Le chapitre 3 est entièrement consacré à l'étude de cas particuliers. Nous établissons les résultats pour des mesures d'association particulières, soit, la mesure de Stewart et Love (1968), la mesure d'Escoufier (1973) et la mesure de Cramer et Nicewander (1979) et pour deux différentes fonctions de cotes, les cotes normales et les cotes linéaires. Ces choix ont été motivés par la volonté de comparer les performances et résultats expérimentaux à ce qui est fait dans la littérature, mais bien entendu, il serait possible d'étudier plusieurs cas particuliers et c'est d'ailleurs une des première étape qui suivra ce travail.

La deuxième partie du travail se concentre sur l'étude de la performance des tests. Jusqu'à maintenant dans la littérature, la performance des tests dont il est question ici a été évaluée essentiellement de façon expérimentale, c'est-à-dire par des études de simulation pour illustrer le comportement des tests dans différentes situations données, notamment, pour de petites tailles échantillonales.

Nous avons donc attaqué, au chapitre 4, le calcul de l'efficacité asymptotique relative des tests de la classe non seulement entre eux, mais aussi par rapport à des tests n'appartenant pas à la classe. Dans un premier temps, on étudie la sous-classe de tests obtenue en utilisant la mesure de Cramer et Nicewander (1979).

1. En annexe B, nous présentons une démonstration à l'effet que la suite de contre-hypothèses considérée est contiguë à l'hypothèse nulle lorsque les observations proviennent d'une loi multinormale.

On constate qu'il est possible d'obtenir des résultats analytiques très intéressants étant donné la forme particulière des distributions des statistiques appartenant à cette sous-classe.

On commence par déduire une expression générale pour l'efficacité asymptotique relative entre deux tests appartenant à cette sous-classe, quelque soient la distribution des observations et les fonctions de cotes considérées. On obtient ensuite une expression explicite pour l'efficacité asymptotique relative du test basé sur les cotes normales par rapport à celui basé sur les cotes linéaires, dans le cas où les observations proviennent d'une loi multinormale avec des structures de covariance données (théorème 4.2 et 4.3). On y établit alors la supériorité du test basé sur les cotes normales dans ces situations précises.

Par la suite, on évalue l'efficacité asymptotique relative du test basé sur la mesure de Cramer et Nicewander (1979) et sur les cotes normales par rapport au test de Puri *et al* (1970) et au test paramétrique du rapport de vraisemblance maximale.

Finalement, on présente une façon d'obtenir une valeur numérique pour calculer l'efficacité asymptotique relative de n'importe quel couple de tests dont la distribution est de la forme de celle des tests appartenant à la classe; c'est-à-dire une combinaison linéaire de variables aléatoires de loi khi-deux décentrées. Cette méthode repose essentiellement sur l'approche de Nyblom et Mäkeläinen (1983) et l'utilisation de l'algorithme d'Imhof (1961) dont nous discutons aussi au chapitre 4.

Pour clore le chapitre, on illustre cette méthode en présentant des calculs d'efficacité asymptotique relative pour chacun des cas particuliers considérés au chapitre 3, dans le cas où les observations proviennent d'une loi logistique ou d'une loi normale.

Notons que pour tous les cas considérés, les tests basés sur la mesure de Cramer et Nicewander (1979) donnent les efficacités asymptotiques relatives les plus élevées.

Finalement, au chapitre 5, on présente une étude expérimentale de la performance de six tests faisant partie de la classe, les tests de Puri *et al* (1970) et le test du rapport de maximum de vraisemblance pour un ensemble de situations données qui couvrent tout ce qui a été fait dans la littérature concernant ces tests. Encore une fois, le test basé sur la mesure de Cramer et Nicewander (1979) ressort parmi les tests considérés dans cette étude expérimentale.

Chapitre 1

UNE CLASSE DE TESTS BASÉS SUR DES FONCTIONS DES RANGS

1.1. LA MATRICE D'ASSOCIATION

Considérons un vecteur aléatoire de dimension p composé de K sous-vecteurs

$$X = \begin{pmatrix} X^{(1)} \\ \vdots \\ X^{(K)} \end{pmatrix} = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix}$$

où le sous-vecteur $X^{(k)}$ est de dimension p_k pour $k = 1, \dots, K$ et $\sum_{k=1}^K p_k = p$.

Dans le présent travail, le vecteur formé d'une part de K sous-vecteurs $X = (X^{(1)'}, \dots, X^{(K)'})'$ est aussi formé de p variables $X = (X_1 \cdots X_p)'$. On utilisera les indices i et j lorsqu'on fera référence au couple de variables X_i et X_j et les indices k et l serviront à identifier le couple de sous-vecteurs $X^{(k)}$ et $X^{(l)}$. Ainsi, on a i et $j = 1, \dots, p$ et, k et $l = 1, \dots, K$.

Définissons les quantités suivantes :

- $F(\mathbf{x})$: distribution conjointe des composantes du vecteur aléatoire X
où $\mathbf{x} = (x_1, \dots, x_p)'$,
- $F^{(k)}(x^{(k)})$: distribution conjointe du sous-vecteur $X^{(k)}$ pour $k = 1, \dots, K$,
- $F_{[i]}(x)$: distribution marginale de X_i pour $i = 1, \dots, p$,

- $F_{[i,j]}(x,y)$: distribution conjointe de (X_i, X_j) pour i et $j = 1, \dots, p$ ($i \neq j$).

Pour chaque sous-vecteur, notons l'espérance et la matrice de covariance par

$$\mu^{(k)} = E(X^{(k)}) \text{ et } \Sigma_{kl} = COV(X^{(k)}, X^{(l)}) = E(X^{(k)} - \mu^{(k)})(X^{(l)} - \mu^{(l)})'$$

pour k et $l = 1, \dots, K$ où A' représente la transposée de la matrice A . Enfin, posons

$$\mu = \begin{pmatrix} \mu^{(1)} \\ \vdots \\ \mu^{(K)} \end{pmatrix} \text{ et } \Sigma = \begin{pmatrix} \Sigma_{11} & \cdots & \Sigma_{1K} \\ \vdots & \ddots & \vdots \\ \Sigma_{K1} & \cdots & \Sigma_{KK} \end{pmatrix}$$

qui représentent respectivement le vecteur des moyennes et la matrice de covariance du vecteur aléatoire X .

Pour $i = 1, \dots, p$, soit $J_{(i)}(u)$, $0 < u < 1$, une fonction absolument continue sur $(0,1)$ et satisfaisant les conditions suivantes :

Condition 1.1. (i) $\int_0^1 J_{(i)}(u) du = 0$ et (ii) $\int_0^1 J_{(i)}^2(u) du = 1$.

Posons $Y_i = J_{(i)}(F_{[i]}(X_i))$ pour $i = 1, \dots, p$. Cette quantité est appelée la fonction grade de X_i , $i = 1, \dots, p$ (voir Hoeffding (1948)). On définit la matrice d'association Θ , de dimensions $p \times p$, à partir des éléments θ_{ij} où

$$\begin{aligned} \theta_{ij} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J_{(i)}(F_{[i]}(x)) J_{(j)}(F_{[j]}(y)) dF_{[i,j]}(x,y) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Y_i Y_j dF_{[i,j]}(x,y) = E_{F_{[i,j]}}(Y_i Y_j) \end{aligned} \quad (1.1)$$

pour i et $j = 1, \dots, p$ avec $E_{F_{[i,j]}}$ représentant l'espérance par rapport à la distribution $F_{[i,j]}$. Les éléments θ_{ij} , i et $j = 1, \dots, p$, de la matrice Θ sont appelés paramètres d'association. Notons qu'avec la condition 1.1, on a $\theta_{ii} = 1$ pour $i = 1, \dots, p$ et $|\theta_{ij}| \leq 1$ pour $i \neq j$, i et $j = 1, \dots, p$. Enfin, remarquons qu'étudier

l'indépendance entre X_i et X_j est équivalent à étudier l'indépendance entre Y_i et Y_j par la proposition suivante tirée de Puri *et al* (1970).

Proposition 1.1. *Soient $J_{(r)}(u)$ et $J_{(s)}(u)$ deux fonctions absolument continues et satisfaisant à la condition 1.1. Si $J_{(r)}(u)$ et $J_{(s)}(u)$ ne sont pas constantes pour $0 < u < 1$, alors X_r et X_s sont indépendantes si et seulement si Y_r et Y_s sont indépendantes pour r et $s = 1, \dots, p$ ($r \neq s$).*

On partitionne la matrice d'association de la façon suivante :

$$\Theta = \begin{pmatrix} \Theta_{11} & \cdots & \Theta_{1K} \\ \vdots & \ddots & \vdots \\ \Theta_{K1} & \cdots & \Theta_{KK} \end{pmatrix}$$

où Θ_{kl} est une matrice de dimensions $p_k \times p_l$ représentant l'association entre les sous-vecteurs $X^{(k)}$ et $X^{(l)}$ du vecteur X .

Considérons maintenant un échantillon de vecteurs aléatoires indépendants

$$X^{[1]} = \begin{pmatrix} X_1^{(1)} \\ \vdots \\ X_1^{(K)} \end{pmatrix}, \dots, X^{[n]} = \begin{pmatrix} X_n^{(1)} \\ \vdots \\ X_n^{(K)} \end{pmatrix}$$

de même loi que X avec

$$\bar{X} = \begin{pmatrix} \bar{X}^{(1)} \\ \vdots \\ \bar{X}^{(K)} \end{pmatrix} \text{ et } S = \begin{pmatrix} S_{11} & \cdots & S_{1K} \\ \vdots & \ddots & \vdots \\ S_{K1} & \cdots & S_{KK} \end{pmatrix}$$

où

$$\bar{X}^{(k)} = \frac{1}{n} \sum_{\alpha=1}^n X_{\alpha}^{(k)} \text{ et } S_{kl} = \frac{1}{n-1} \sum_{\alpha=1}^n (X_{\alpha}^{(k)} - \bar{X}^{(k)})(X_{\alpha}^{(l)} - \bar{X}^{(l)})'$$

pour k et $l = 1, \dots, K$.

Puisque les paramètres d'association présentés dans (1.1) sont définis de façon analogue aux fonctionnelles étudiées par von Mises (1947), Puri et Sen (1971) proposent d'utiliser la même approche pour construire des estimateurs pour θ_{ij} .

Condition 1.2. On considère pour $i = 1, \dots, p$, une suite de fonctions $\{J_{n(i)}(u)\}_{n \geq 1}$ telles que $J_{n(i)}(u)$ est absolument continue sur $(0,1)$ et converge uniformément pour $0 < u < 1$ vers $J_{(i)}(u)$ lorsque $n \rightarrow \infty$.

Par souci de cohérence avec la condition 1.1 imposée aux fonctions $J_{(i)}(\cdot)$, on choisit les fonctions $J_{n(i)}(\cdot)$ telles que pour $i = 1, \dots, p$,

Condition 1.3. (i) $\sum_{j=1}^n J_{n(i)}\left(\frac{j}{n}\right) = 0$ et (ii) $\frac{1}{n} \sum_{j=1}^n J_{n(i)}^2\left(\frac{j}{n}\right) = 1$.

On construit ensuite pour $i = 1, \dots, p$ les cotes par

$$J_{n(i)}\left(F_{n[i]}^*(x_i)\right)$$

où $F_{n[i]}^*(\cdot) = \frac{n}{n+1}F_{n[i]}(\cdot)$ et $F_{n[i]}(\cdot)$ est la distribution marginale expérimentale de X_i c'est-à-dire

$$F_{n[i]}(x_i) = [\text{nombre de } X_{i\alpha} \text{ parmi } X_{i1}, \dots, X_{in} \leq x_i] / n,$$

où $X_{i\alpha}$ représente la $i^{\text{ième}}$ composante du $\alpha^{\text{ième}}$ vecteur de l'échantillon aléatoire, $\alpha = 1, \dots, n$.

On obtient ainsi un estimateur pour θ_{ij} , i et $j = 1, \dots, p$, en posant

$$\hat{\theta}_{ij} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J_{n(i)}\left(F_{n[i]}^*(x_i)\right) J_{n(j)}\left(F_{n[j]}^*(x_j)\right) dF_{n[i,j]}(x_i, x_j) \quad (1.2)$$

où $F_{n[i,j]}(x,y)$ est la distribution conjointe expérimentale du couple (X_i, X_j) , c'est-à-dire

$$F_{n[i,j]}(x,y) = [\text{nombre de couples } (X_{i\alpha}, X_{j\alpha}) \text{ tels que } X_{i\alpha} \leq x \text{ et } X_{j\alpha} \leq y] / n,$$

$\alpha = 1, \dots, n$.

On peut aussi écrire pour i et $j = 1, \dots, p$,

$$\hat{\theta}_{ij} = \frac{1}{n} \sum_{\alpha=1}^n J_{n(i)}(F_{n[i]}^*(x_{i\alpha})) J_{n(j)}(F_{n[j]}^*(x_{j\alpha})) = t_{ij}. \quad (1.3)$$

On définit l'estimateur de la matrice d'association par la matrice de dimensions $p \times p$ formée des t_{ij} :

$$\hat{\Theta} = T = \begin{pmatrix} t_{11} & \cdots & t_{1p} \\ \vdots & \ddots & \vdots \\ t_{p1} & \cdots & t_{pp} \end{pmatrix} = \begin{pmatrix} T_{11} & \cdots & T_{1K} \\ \vdots & \ddots & \vdots \\ T_{K1} & \cdots & T_{KK} \end{pmatrix} \quad (1.4)$$

où les sous-matrices T_{kl} sont de dimensions $p_k \times p_l$ conformément aux sous-matrices Θ_{kl} de la matrice d'association Θ .

De plus, si on représente le déterminant d'une matrice A par $|A|$, on a la proposition suivante tirée de Puri et Sen (1966).

Proposition 1.2. *Si la matrice Θ est définie positive, alors $|T|$ converge en probabilité vers $|\Theta|$ lorsque $n \rightarrow \infty$, si pour un certain $\delta > 0$, $E|t_{ii}|^{\frac{(2+\delta)}{2}} < \infty$.*

Notons enfin que comme chacune des composantes de la matrice T converge en probabilité vers les composantes correspondantes de la matrice Θ (voir Puri et Sen (1971), lemme 8.5.1, p.359), la matrice T est un estimateur convergent de la matrice Θ .

1.2. L'HYPOTHÈSE D'INDÉPENDANCE

Le problème qui nous intéresse concerne l'hypothèse d'indépendance entre K sous-vecteurs de X . Dans cette section, nous présentons comment s'expriment les hypothèses qui seront confrontées ainsi que les conditions sous lesquelles on devra parler de tests d'absence de liaison plutôt que de tests d'indépendance.

La classe de tests présentée dans le présent travail étant basée sur la matrice d'association, étudions, dans un premier temps, de quelle façon l'indépendance et les paramètres d'association sont reliés.

Définissons les quantités suivantes introduites par Puri et Sen (1971) :

- La fonction de dépendance entre deux variables aléatoires X_i et X_j :

$$\Omega_{ij}(x_i, x_j) = \frac{F_{[i,j]}(x_i, x_j)}{F_{[i]}(x_i) \cdot F_{[j]}(x_j)}. \quad (1.5)$$

- La fonction de dépendance des K sous-vecteurs du vecteur X :

$$\Omega(x) = \frac{F(\mathbf{x})}{\prod_{k=1}^K F^{(k)}(x^{(k)})}. \quad (1.6)$$

- La fonction de dépendance totale entre les p variables aléatoires qui composent le vecteur X :

$$\Omega_p(x) = \frac{F(\mathbf{x})}{\prod_{i=1}^p F_{[i]}(x_i)}.$$

Avec ces définitions, on obtient les trois équivalences suivantes :

- $\Omega_{ij}(x_i, x_j) = 1, \forall (x_i, x_j)$ si et seulement si X_i et X_j sont indépendantes,
- $\Omega(x) = 1, \forall x$ si et seulement si $X^{(1)}, \dots, X^{(K)}$ sont indépendants,
- $\Omega_p(x) = 1, \forall x$ si et seulement si X_1, \dots, X_p sont indépendantes.

Par (1.5) et (1.6), on constate aussi que $\Omega(x) = 1$ est équivalent à $\Omega_{ij}(x_i, x_j) = 1$ pour toutes les variables X_i et X_j qui n'appartiennent pas au même sous-vecteur.

Nous sommes maintenant en mesure de présenter les liens entre la notion d'indépendance et les paramètres d'association qui sont à la base des tests étudiés. En fait, en utilisant (1.5), on constate que les paramètres d'association peuvent

s'exprimer en terme de la fonction de dépendance de la façon suivante (voir Puri *et al* (1970)) :

Proposition 1.3. *Si les fonctions $J_{(i)}$ satisfont la condition 1.1, $|J_{(i)}(1)| < \infty$ et $|J_{(i)}(0)| < \infty$ et si les dérivées $J'_{(i)}(u)$ existent pour $i = 1, \dots, p$, on peut alors écrire*

$$\begin{aligned} \theta_{ij} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [F_{[i,j]}(x_i, x_j) - F_{[i]}(x_i)F_{[j]}(x_j)] \\ &\quad \cdot J'_{(i)}(F_{[i]}(x_i)) J'_{(j)}(F_{[j]}(x_j)) dF_{[i]}(x_i) dF_{[j]}(x_j) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\Omega_{ij}(x_i, x_j) - 1] F_{[i]}(x_i) F_{[j]}(x_j) J'_{(i)}(F_{[i]}(x_i)) \\ &\quad \cdot J'_{(j)}(F_{[j]}(x_j)) dF_{[i]}(x_i) dF_{[j]}(x_j) \end{aligned} \tag{1.7}$$

En exprimant θ_{ij} sous la forme (1.7), on obtient directement la proposition suivante.

Proposition 1.4. *Soit X_i et X_j deux variables aléatoires. Si X_i et X_j sont indépendantes alors $\theta_{ij} = 0$.*

Nous avons donc que l'indépendance entraîne que les paramètres d'association sont nuls, mais qu'en est-il de la réciproque?

Dans le cadre paramétrique, lorsqu'on utilise la matrice de covariance classique Σ pour étudier l'indépendance, il est bien connu que $\sigma_{ij} = Cov(X_i, X_j) = 0$ n'entraîne pas nécessairement l'indépendance entre X_i et X_j pour i et $j = 1, \dots, p$ ($i \neq j$), sauf dans le cas de la multinormalité. De la même façon, dans le cas qui nous intéresse, la réciproque sera vraie seulement sous certaines conditions. En fait, on veut montrer que pour i et $j = 1, \dots, p$ ($i \neq j$),

si $\theta_{ij} = 0$ alors les variables aléatoires X_i et X_j sont indépendantes. (1.8)

Donc en supposant $|J_{(i)}(1)| < \infty$ et $|J_{(i)}(0)| < \infty$ et si les dérivées $J'_{(i)}(u)$ existent pour $i = 1, \dots, p$, on peut exprimer θ_{ij} sous la forme (1.7). On constate donc que (1.8) est vraie si les fonctions $J_{(i)}(u)$ et $J_{(j)}(u)$ sont monotones (croissantes ou décroissantes) en u , $0 < u < 1$, et si $\Omega_{ij}(x_i, x_j) \geq 1$ (ou si $\Omega_{ij}(x_i, x_j) \leq 1$) pour tout $(x_i, x_j) \in \mathfrak{R}^2$.

La condition de monotonie des fonctions J n'est pas vraiment restrictive puisque la plupart des fonctions J qu'on considère en pratique sont monotones en u , mais la condition sur la fonction de dépendance l'est beaucoup plus. Puri *et al* (1970) font mention de quelques cas particuliers qui respectent cette condition, mais de façon générale, (1.8) n'est pas nécessairement vérifiée, comme c'est le cas, entre autre, dans les travaux de Puri *et al* (1970) et dans ceux de Cléroux *et al* (1994) et Cléroux *et al* (1995).

Terminons donc cette section en notant que l'hypothèse nulle $H_0^{(K)}$ représentant l'indépendance entre les K sous-vecteurs de X peut ainsi s'exprimer en terme de la fonction de dépendance (1.5) :

$$H_0^{(K)} : \Omega(x) = 1, \quad (1.9)$$

et l'hypothèse nulle représentant l'absence de liaison entre les K sous-vecteurs de X peut s'exprimer en terme des matrices d'association :

$$H_0^{*(K)} : \Theta_{kl} = 0 \text{ pour } 1 \leq k < l \leq K. \quad (1.10)$$

Dans la prochaine section, nous définissons donc une classe de statistiques de tests pour étudier cette hypothèse.

1.3. DÉFINITION DE LA CLASSE DE STATISTIQUES

Considérons pour k et $l = 1, \dots, p$, des mesures d'association vectorielles théoriques de la forme

$$\varrho_{kl} = tr(\Theta_{lk}\Delta_{kl}\Theta_{lk}) \quad (1.11)$$

où l'opérateur $tr(A)$ représente la trace de la matrice A , $\Delta_{kl} = \Delta(\Theta_{kk}, \Theta_{ll})$ est une matrice de dimensions $p_k \times p_l$ et où les sous-matrices Θ_{kk} , Θ_{ll} et Θ_{kl} , k et $l = 1, \dots, K$, de la matrice Θ qui sont formées des éléments θ_{ij} définis par (1.1).

En utilisant le résultat suivant que l'on retrouve, entre autre, dans Muirhead (1982),

$$tr(BCD) = [vec(B)']'(I \otimes C)[vec(D)], \quad (1.12)$$

où I est la matrice identité, \otimes désigne le produit de Kronecker et $vec(A)$ est le vecteur formé en empilant les colonnes de la matrice A les unes sous les autres, on peut, de façon équivalente, exprimer ϱ_{kl} (1.11) comme une forme quadratique en Θ_{kl}

$$\varrho_{kl} = [vec(\Theta_{lk})]' \Lambda_{kl} [vec(\Theta_{lk})] \quad (1.13)$$

où $\Lambda_{kl} = I_{p_k} \otimes \Delta_{kl}$, k et $l = 1, \dots, K$ est une matrice de dimensions $p_k p_l \times p_k p_l$.

Puisque pour k et $l = 1, \dots, K$, $\varrho_{kl} \geq 0$, on en déduit que sous $H_0^{(K)}$, $\sum \sum_{1 \leq k < l \leq K} \varrho_{kl} = 0$. Il est donc naturel de penser à des tests basés sur des statistiques de la forme

$$\sum_{1 \leq k < l \leq K} R_{kl} \quad (1.14)$$

où R_{kl} est un estimateur de ϱ_{kl} ayant des propriétés intéressantes dans le cadre du problème qui nous intéresse. On voudra notamment connaître la distribution asymptotique conjointe de ces estimateurs sous $H_0^{(K)}$ ainsi que sous une suite de

contre-hypothèses convergeant vers $H_0^{(K)}$ et on voudra aussi que cet estimateur converge en probabilité vers ϱ_{kl} .

On rejettera donc l'hypothèse d'absence d'association entre les K sous-vecteurs de X pour des valeurs élevées de la statistique (1.14).

On a vu précédemment que chaque paramètre d'association échantillonnal converge en probabilité vers le paramètre échantillonnal correspondant, c'est-à-dire $t_{ij} \xrightarrow{\mathcal{P}} \theta_{ij}$. On obtient donc un estimateur naturel en remplaçant les paramètres d'associations θ_{ij} par leurs estimateurs t_{ij} définis par (1.3) dans l'expression de ϱ_{kl} .

Ainsi, en posant, $\hat{\Delta}_{kl} = \Delta(T_{kk}, T_{ll})$ et $\hat{\Lambda}_{kl} = I_{p_k p_l} \otimes \hat{\Delta}_{kl}$ pour k et $l = 1, \dots, K$, les quantités

$$R_{kl} = \text{tr} \left(T_{lk} \hat{\Delta}_{kl} T_{lk} \right), \quad (1.15)$$

ou de façon équivalente

$$R_{kl} = [\text{vec}(T_{lk})]' \hat{\Lambda}_{kl} [\text{vec}(T_{lk})], \quad (1.16)$$

sont des estimateurs convergents pour ϱ_{kl} . Nous appellerons ces quantités des mesures d'association vectorielle échantillonnales.

Notons que puisque les mesures d'association vectorielle (1.13) sont fonctions des paramètres d'association θ_{ij} définis par (1.1), elles sont donc fonctions des $J_{(i)}(\cdot)$, $i = 1, \dots, p$, ainsi que des matrices Λ_{kl} qui définissent les formes quadratiques en Θ_{kl} , $1 \leq k < l \leq K$. Il serait ainsi plus précis d'exprimer les quantités (1.13) de la façon suivante :

$$\varrho_{kl}^{\Lambda_{kl}, J} = [\text{vec}(\Theta_{lk})]' \Lambda_{kl} [\text{vec}(\Theta_{lk})].$$

D'une façon équivalente, il serait plus précis d'exprimer les mesures d'association vectorielle échantillonnales (1.16) de la façon suivante :

$$R_{kl}^{\hat{\Lambda}_{kl}, J_n} = [\text{vec}(T_{lk})]' \hat{\Lambda}_{kl} [\text{vec}(T_{lk})], \quad (1.17)$$

où $\hat{\Lambda}_{kl} = \Delta(T_{kk}, T_{ll})$ et $J_{n(i)}(\cdot)$ respectent les conditions 1.2 et 1.3.

La classe que nous considérons est ainsi constituée des statistiques de la forme (1.17).

Dans le prochain chapitre, on obtient la distribution asymptotique conjointe des statistiques (1.17) sous $H_0^{(K)}$ ainsi que sous une suite de contre-hypothèses qui convergent vers $H_0^{(K)}$. Nous verrons donc à ce moment, en quoi les conditions qui suivent sont essentielles pour obtenir les résultats asymptotiques, mais pour le moment, commençons par présenter ces conditions afin d'être en mesure de définir formellement la classe de tests.

Condition 1.4. *Les matrices Λ_{kl} et $\hat{\Lambda}_{kl}$ sont symétriques.*

Condition 1.5. *Les fonctions $J_{n(i)}(u)$ et $J_{(i)}(u)$ sont telles que*

(i) $\lim_{n \rightarrow \infty} J_{n(i)}(u) = J_{(i)}(u)$ existe pour $0 < u < 1$ et est non constante pour $i = 1, \dots, p$. De plus, cette limite vérifie la condition 1.1.

(ii) $J_{n(i)}(1) = o(n^{1/4})$, $i = 1, \dots, p$.

(iii) Pour $i, j = 1, \dots, p$,

$$\begin{aligned} & \int_{I_{n(i)}} \int_{I_{n(j)}} \{ J_{n(i)}(F_{n[i]}^*(x_i)) J_{n(j)}(F_{n[j]}^*(x_j)) \\ & \quad - J_{(i)}(F_{[i]}(x_i)) J_{(j)}(F_{[j]}(x_j)) \} dF_{n[i,j]}(x_i, x_j) \\ & = o_p(n^{-1/2}) \end{aligned}$$

où $I_{n(k)} = \{x_k : 0 < F_{n(k)}(x_k) < 1\}$ pour $k = 1, \dots, p$ et $X_n = o_p(b_n)$ si $\frac{X_n}{b_n} \xrightarrow{\mathcal{P}} 0$ où $\xrightarrow{\mathcal{P}}$ désigne la convergence en probabilité lorsque $n \rightarrow \infty$.

(iv) Il existe une constante C telle que pour $i = 1, \dots, p$,

$$1) |J_{(i)}(u)| \leq C [u(1-u)]^{-\alpha} \text{ pour } 0 < \alpha < \frac{1}{8},$$

$$2) |J'_{(i)}(u)| \leq C [u(1-u)]^{-1},$$

$$3) |J''_{(i)}(u)| \leq C [u(1-u)]^{-2}.$$

où $J'_{(i)}(u)$ et $J''_{(i)}(u)$ désignent respectivement les dérivées première et seconde de $J_{(i)}(u)$.

Remarquons que contrairement aux apparences, la condition 1.5 n'est pas restrictive. Par exemple, les fonctions des rangs usuelles telles que les cotes normales ou linéaires respectent ces conditions. Mentionnons aussi un théorème dû à Bhuchongkul (1964) qui permet d'obtenir des fonctions satisfaisant la condition 1.5.

Théorème 1.1. *Si pour $i = 1, \dots, p$, $J_{n(i)}\left(\frac{\alpha}{n+1}\right)$ est l'espérance de la α -ième statistique d'ordre d'un échantillon de n variables dont la distribution est la fonction inverse de $J_{(i)}(\cdot)$ et si la condition 1.5 (iv) est vérifiée, alors les conditions 1.5 (i), (ii) et (iii) sont aussi vérifiées.*

Nous sommes maintenant en mesure de présenter la classe de statistiques avec laquelle nous étudierons le problème d'absence de liaison entre K sous-vecteurs de X .

Définition 1.1. *Une statistique $R_{kl}^{\hat{\Lambda}_{kl}, J_n}$ de la forme (1.17) fait partie de la classe \mathcal{C} si les conditions 1.1, 1.2, 1.3, 1.4 et 1.5 sont satisfaites.*

Pour clore ce chapitre, rappelons que la classe \mathcal{C} comprend les statistiques de tests présentées par Cléroux *et al* (1994) et Cléroux *et al* (1995). Nous verrons d'ailleurs ces cas particuliers au chapitre 3.

Chapitre 2

DISTRIBUTIONS ASYMPTOTIQUES DES STATISTIQUES DE LA CLASSE \mathcal{C}

Dans le chapitre précédent, on a introduit la classe \mathcal{C} de statistiques pour étudier le problème d'absence de liaison entre K sous-vecteurs de X . Voyons maintenant quelle est la distribution asymptotique des statistiques faisant partie de la classe, sous $H_0^{(K)}$ (donnée par (1.9)) ainsi que sous une suite de contre-hypothèses qui converge vers $H_0^{(K)}$.

2.1. DISTRIBUTION ASYMPTOTIQUE DE T

Puisque les statistiques de la classe \mathcal{C} s'expriment comme des formes quadratiques du vecteur $vec(T)$, c'est-à-dire

$$R_{kl}^{\hat{\Lambda}_{kl}, J_n} = [vec(T_{lk})]' \hat{\Lambda}_{kl} [vec(T_{lk})],$$

k et $l = 1, \dots, K$, on a besoin, dans un premier temps, d'établir la distribution asymptotique conjointe des éléments de T .

Rappelons que la matrice T est formé des estimateurs des paramètres d'association, soit

$$t_{ij} = \hat{\theta}_{ij} = \frac{1}{n} \sum_{\alpha=1}^n J_{n(i)}(F_{n[i]}^*(x_{i\alpha})) J_{n(j)}(F_{n[j]}^*(x_{j\alpha})),$$

i et $j = 1, \dots, p$.

Puri *et al* (1970) ont démontré le résultat suivant quelque soit la distribution des observations F .

Théorème 2.1. *Soient les matrices Θ et T dont les éléments sont respectivement définis par les équations (1.1) et (1.3) et supposons que les fonctions $J_{(i)}(\cdot)$ et $J_{n(i)}(\cdot)$, $i = 1, \dots, p$ satisfont la condition 1.5. Alors, lorsque $n \rightarrow \infty$, on a*

$$\sqrt{n} (\text{uvec}(T) - \text{uvec}(\Theta)) \xrightarrow{\mathcal{L}} N_f(0, \Phi)$$

où $f = \frac{p(p-1)}{2}$ et $\Phi = (\phi_{(i,j),(r,s)})$ où pour i, j, r et $s = 1, \dots, p$,

$$\phi_{(i,j),(r,s)} = \begin{cases} \text{Var} \left[\sum_{l=1}^3 U_{(i,j);l}^{(\alpha)} \right] & \text{si } (i,j) = (r,s), \\ \sum_{k=1}^3 \sum_{l=1}^3 \text{Cov} \left(U_{(i,j);k}^{(\alpha)}, U_{(r,s);l}^{(\alpha)} \right) & \text{si } (i,j) \neq (r,s), \end{cases}$$

avec

$$\begin{aligned} U_{(r,s);1}^{(\alpha)} &= J_{(r)}(F_{[r]}(X_{r\alpha})) J_{(s)}(F_{[s]}(X_{s\alpha})), \\ U_{(r,s);2}^{(\alpha)} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (I(X_{r\alpha} \leq x_r) - F_{[r]}(x_r)) J_{(s)}(F_{[s]}(x_s)) J'_{(r)}(F_{[r]}(x_r)) \\ &\quad \cdot dF_{[r,s]}(x_r, x_s), \\ U_{(r,s);3}^{(\alpha)} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (I(X_{s\alpha} \leq x_s) - F_{[s]}(x_s)) J_{(r)}(F_{[r]}(x_r)) J'_{(s)}(F_{[s]}(x_s)) \\ &\quad \cdot dF_{[r,s]}(x_r, x_s) \end{aligned}$$

$$\text{où } (X_{r\alpha} \leq u) = \begin{cases} 1 & \text{si } X_{r\alpha} \leq u, \\ 0 & \text{sinon} \end{cases} \quad \text{et } \alpha = 1, \dots, n.$$

La quantité $\text{uvec}(T)$ désigne le vecteur de dimension $\frac{p(p-1)}{2}$ formé en empilant les colonnes de T mais en omettant les éléments sur et sous la diagonale principale. Enfin, la notation $\xrightarrow{\mathcal{L}}$ représente la convergence en loi.

La démonstration détaillée présentée dans l'article de Puri *et al* (1970) traite du cas où les observations $X_{i\alpha}$ sont indépendantes et identiquement distribuées selon $F_{[i]}(x)$ et où les couples $(X_{i\alpha}, X_{j\alpha})$ indépendants et identiquement distribués selon $F_{[i,j]}(x, y)$, pour $\alpha = 1, \dots, n$ et i et $j = 1, \dots, p$, ($i \neq j$). Ce résultat

demeure pourtant valide dans le cas où les observations $X_{i\alpha}$ sont indépendantes et de loi $F_{[i]}^{(\alpha)}(x)$ et les couples $(X_{i\alpha}, X_{j\alpha})$ sont indépendants et de loi $F_{[i,j]}^{(\alpha)}(x, y)$, pour $\alpha = 1, \dots, n$ et i et $j = 1, \dots, p$, ($i \neq j$).

Or, comme nous le verrons au théorème 2.7, ce résultat plus général est essentiel pour obtenir la distribution asymptotique des statistiques sous une suite de contre-hypothèses convergeant vers $H_0^{(K)}$. Mentionnons d'ailleurs que ce résultat plus général est tout aussi essentiel dans le cadre des travaux de Puri *et al* (1970) tout comme dans ceux de Cléroux *et al* (1994) et Cléroux *et al* (1995). Comme on retrouve très peu de détails sur le passage du cas où les observations sont indépendantes et identiquement distribuées au cas plus général dans l'article de Puri *et al* (1970) et puisque ce passage n'est pas trivial pour qui n'est pas familier avec cette théorie, nous présentons dans l'annexe 1 les principales étapes du raisonnement permettant la généralisation du théorème 2.1.

2.2. DISTRIBUTION ASYMPTOTIQUE CONJOINTE DES R_{kl} SOUS $H_0^{(K)}$

On veut obtenir la distribution asymptotique conjointe des statistiques appartenant à la classe \mathcal{C} sous l'hypothèse nulle $H_0^{(K)} : \Omega(x) = 1$. Par le théorème 2.1, puisque les fonctions $J_{(i)}(\cdot)$ et $J_{n(i)}(\cdot)$, $i = 1, \dots, p$ satisfont la condition 1.5, on a alors sous $H_0^{(K)}$

$$\sqrt{n}(\text{uvec}(T)) \xrightarrow{\mathcal{L}} N_f(0, \Phi_0)$$

puisque l'on a alors $\Theta_{kl} = 0$ pour $1 \leq k < l \leq K$.

Déterminons maintenant la nature des éléments de la matrice de covariance Φ_0 .

Théorème 2.2. *Sous la condition 1.1 et si X_r et X_s sont deux variables aléatoires telles que $F_{[r,s]}(x_r, x_s) = F_{[r]}(x_r) \cdot F_{[s]}(x_s)$, alors*

- (i) $U_{(r,s);2}^{(\alpha)}$ et $U_{(r,s);3}^{(\alpha)}$ sont des variables aléatoires dégénérées en 0;
- (ii) $E\left(U_{(r,s);l}^{(\alpha)}\right) = 0$ pour $l = 1, 2, 3$.

Démonstration:

- (i) Puisque $F_{[r,s]}(x_r, x_s) = F_{[r]}(x_r) \cdot F_{[s]}(x_s)$, on peut écrire

$$U_{(r,s);2}^{(\alpha)} = \int_{-\infty}^{\infty} ((F_{1X_{r\alpha}}(x_r) - F_{[r]}(x_r)) J'_{(r)}(F_{[r]}(x_r)) dF_{[r]}(x_r) \\ \cdot \int_{-\infty}^{\infty} J_{(s)}(F_{[s]}(x_s)) dF_{[s]}(x_s)$$

et

$$U_{(r,s);3}^{(\alpha)} = \int_{-\infty}^{\infty} ((F_{1X_{s\alpha}}(x_s) - F_{[s]}(x_s)) J'_{(s)}(F_{[s]}(x_s)) \int_{-\infty}^{\infty} J_{(r)}(F_{[r]}(x_r)) dF_{[r]}(x_r).$$

Mais comme par la condition 1.1, $\int_{-\infty}^{\infty} J_{(r)}(F_{[r]}(x_r)) dF_{[r]}(x_r) = 0$, le résultat suit.

- (ii) Par (i), on a $E\left(U_{(r,s);2}^{(\alpha)}\right) = E\left(U_{(r,s);3}^{(\alpha)}\right) = 0$. D'autre part, comme

$$E\left(U_{(r,s);1}^{(\alpha)}\right) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J_{(r)}(F_{[r]}(X_{r\alpha})) J_{(s)}(F_{[s]}(X_{s\alpha})) dF_{[r,s]}(x_r, x_s) = \theta_{rs},$$

on obtient le résultat par la proposition 1.4.

C.Q.F.D.

On obtient ainsi, quelque soit la distribution de X , le résultat suivant :

Théorème 2.3. *Sous la condition 1.1, la matrice de covariance de la distribution asymptotique de $uvec(T)$ sous $H_0^{(K)}$ est $\Phi_0 = (\phi_{0(i,j),(r,s)})$ où pour i, j, r et $s =$*

$1, \dots, p,$

$$\phi_{0(i,j),(r,s)} = \begin{cases} 1 & \text{si } (i,j) = (r,s), \\ \theta_{ir}\theta_{js} & \text{si } ((i,j),(r,s)) \in C_1, \\ 0 & \text{si } ((i,j),(r,s)) \in C_2, \\ \nu_{ijrs} & \text{sinon} \end{cases}$$

avec

$$C_1 = \{(i,j),(r,s) : (i,j) \neq (r,s) \text{ et } t_{ij} \text{ et } t_{rs} \text{ appartiennent à une même sous-matrice } T_{kl}, 1 \leq k < l \leq K\},$$

et

$$C_2 = \{(i,j),(r,s) : (i,j) \neq (r,s) \text{ et } t_{ij} \text{ et } t_{rs} \text{ appartiennent respectivement aux matrices } T_{k'l'} \text{ et } T_{kl} \text{ où } (k',l') \neq (k,l) \text{ avec } 1 \leq k' < l' \leq K \text{ et } 1 \leq k < l \leq K\}.$$

Seulement certaines sous-matrices de Φ sont importantes pour obtenir nos résultats. Elles sont représentées par les cas où $(i,j) = (r,s)$, $((i,j),(r,s)) \in C_1$ et $(i,j) = (r,s)$, $((i,j),(r,s)) \in C_2$. C'est pourquoi il n'est pas nécessaire de connaître de façon explicite les quantités ν_{ijrs} .

Avant de procéder à la démonstration, illustrons la différence entre les ensembles C_1 et C_2 définis dans le théorème 2.3, à l'aide d'un exemple.

Exemple 2.1 Considérons un vecteur X pour lequel $K = 3$, $p = 7$, $p_1 = 2$,

$p_2 = 3$ et $p_3 = 2$, c'est-à-dire

$$X = \begin{pmatrix} X^{(1)} \\ X^{(2)} \\ X^{(3)} \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \\ \begin{pmatrix} X_3 \\ X_4 \\ X_5 \end{pmatrix} \\ \begin{pmatrix} X_6 \\ X_7 \end{pmatrix} \end{pmatrix}.$$

Dans ce cas particulier, la matrice d'association échantillonnale est de la forme

$$T = \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{pmatrix} = \begin{pmatrix} \begin{array}{cc|ccc|cc} t_{11} & t_{12} & t_{13} & t_{14} & t_{15} & t_{16} & t_{17} \\ t_{21} & t_{22} & t_{23} & t_{24} & t_{25} & t_{26} & t_{27} \\ \hline t_{31} & t_{32} & t_{33} & t_{34} & t_{35} & t_{36} & t_{37} \\ t_{41} & t_{42} & t_{43} & t_{44} & t_{45} & t_{46} & t_{47} \\ t_{51} & t_{52} & t_{53} & t_{54} & t_{55} & t_{56} & t_{57} \\ \hline t_{61} & t_{62} & t_{63} & t_{64} & t_{65} & t_{66} & t_{67} \\ t_{71} & t_{72} & t_{73} & t_{74} & t_{75} & t_{76} & t_{77} \end{array} \end{pmatrix}.$$

Ainsi, sous $C_1, (i,j) \neq (r,s)$ et les variables X_i et X_r appartiennent au même sous-vecteur $X^{(k)}$ et les variables X_j et X_s appartiennent au même sous-vecteur $X^{(l)}$. Par exemple, pour les composantes t_{13} et t_{24} qui appartiennent à la même sous-matrice T_{12} et $(1,3) \neq (2,4)$, les variables X_1 et X_2 appartiennent à $X^{(1)}$ et les variables X_3 et X_4 appartiennent à $X^{(2)}$.

Sous $C_2, (i,j) \neq (r,s)$ et t_{ij} et t_{rs} appartiennent respectivement aux matrices $T_{k'l'}$ et T_{kl} où $(k',l') \neq (k,l)$ avec $1 \leq k' < l' \leq K$ et $1 \leq k < l \leq K$. Dans le cas présent, il y a nécessairement un couple (ou les deux) (X_i, X_r) et/ou (X_j, X_s) dont les composantes n'appartiennent pas au même sous-vecteur de X . Par exemple,

pour t_{15} et t_{36} qui appartiennent respectivement aux sous-matrices T_{12} et T_{23} , on constate que les variables X_1 et X_3 n'appartiennent pas au même sous-vecteur et il en est de même pour les variables X_5 et X_6 .

Présentons maintenant la démonstration du théorème 2.3.

Démonstration:

i. Si $(i,j) = (r,s)$, on a alors

$$\begin{aligned}\phi_{(i,j),(r,s)} &= \phi_{(r,s),(r,s)} = Var \left[\sum_{l=1}^3 U_{(r,s);l}^{(\alpha)} \right] = Var \left[U_{(r,s);1}^{(\alpha)} \right] \\ &= E \left[\left(U_{(r,s);1}^{(\alpha)} \right)^2 \right] - E^2 \left[U_{(r,s);1}^{(\alpha)} \right] = E \left[\left(U_{(r,s);1}^{(\alpha)} \right)^2 \right]\end{aligned}$$

par le théorème 2.2. Ainsi, on obtient

$$\begin{aligned}\phi_{(r,s),(r,s)} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J_{(r)}^2 (F_{[r]}(X_{r\alpha})) J_{(s)}^2 (F_{[s]}(X_{s\alpha})) dF_{[r,s]}(x_r, x_s) \\ &= \int_{-\infty}^{\infty} J_{(r)}^2 (F_{[r]}(X_{r\alpha})) dF_{[r]}(x_r) \cdot \int_{-\infty}^{\infty} J_{(s)}^2 (F_{[s]}(X_{s\alpha})) dF_{[s]}(x_s).\end{aligned}$$

Mais comme $\int_{-\infty}^{\infty} J_{(s)}^2 (F_{[s]}(X_{s\alpha})) dF_{[s]}(x_s) = 1$ par la condition 2.1, on obtient le résultat.

ii. Si $(i,j) = (r,s)$, $((i,j),(r,s)) \in C_1$. On a alors t_{ij} et t_{rs} appartiennent à une même sous-matrice T_{kl} où $1 \leq k < l \leq K$. Cela équivaut à dire que les variables X_i et X_r appartiennent au même sous-vecteur $X^{(k)}$ tandis que les variables X_j et X_s appartiennent au même sous-vecteur $X^{(l)}$. Supposons d'abord que $i \neq r$ et $j \neq s$. On a alors

$$\begin{aligned}\phi_{(i,j),(r,s)} &= \sum_{k=1}^3 \sum_{l=1}^3 Cov \left(U_{(i,j);k}^{(\alpha)}, U_{(r,s);l}^{(\alpha)} \right) = Cov \left(U_{(i,j);1}^{(\alpha)}, U_{(r,s);1}^{(\alpha)} \right) \\ &= E \left[U_{(i,j);1}^{(\alpha)} \cdot U_{(r,s);1}^{(\alpha)} \right]\end{aligned}$$

par le théorème 2.2. Par conséquent, on obtient

$$\begin{aligned}\phi_{(i,j),(r,s)} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J_{(i)} (F_{[i]}(x_i)) J_{(j)} (F_{[j]}(x_j)) J_{(r)} (F_{[r]}(x_r)) \\ &\quad \cdot J_{(s)} (F_{[s]}(x_s)) dF_{[i,j,r,s]}(x_i, x_j, x_r, x_s).\end{aligned}$$

Comme sous $H_0^{(K)}$, les sous-vecteurs $X^{(k)}$ et $X^{(l)}$ sont indépendants, tout sous-vecteur de $X^{(k)}$ est indépendant de tout sous-vecteur de $X^{(l)}$. Par conséquent, on obtient

$$\begin{aligned}\phi_{(i,j),(r,s)} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J_{(i)}(F_{[i]}(x_i)) J_{(r)}(F_{[r]}(x_r)) dF_{[i,r]}(x_i, x_r) \\ &\quad \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J_{(j)}(F_{[j]}(x_j)) J_{(s)}(F_{[s]}(x_s)) dF_{[j,s]}(x_j, x_s) \\ &= \theta_{ir} \cdot \theta_{js}.\end{aligned}$$

Supposons maintenant le cas où $i = r$ et $j \neq s$. On a maintenant $X_i \in X^{(k)}$

et $\begin{pmatrix} X_j \\ X_s \end{pmatrix} \in X^{(l)}$. Ainsi,

$$\begin{aligned}\phi_{(i,j),(i,s)} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J_{(i)}^2(F_{[i]}(x_i)) J_{(j)}(F_{[j]}(x_j)) J_{(s)}(F_{[s]}(x_s)) \\ &\quad \cdot dF_{[i,j,s]}(x_i, x_j, x_s) \\ &= \int_{-\infty}^{\infty} J_{(i)}^2(F_{[i]}(x_i)) dF_{[i]}(x_i) \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J_{(j)}(F_{[j]}(x_j)) J_{(s)}(F_{[s]}(x_s)) \\ &\quad \cdot dF_{[j,s]}(x_j, x_s) \\ &= \theta_{ii} \cdot \theta_{js} = \theta_{js}\end{aligned}$$

De la même façon, si $i \neq r$ et $j = s$, on a $\begin{pmatrix} X_i \\ X_r \end{pmatrix} \in X^{(k)}$ et $X_j \in X^{(l)}$ et

donc que

$$\begin{aligned}\phi_{(i,j),(r,j)} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J_{(i)}^2(F_{[i]}(x_i)) J_{(j)}(F_{[j]}(x_j)) J_{(r)}(F_{[r]}(x_r)) \\ &\quad \cdot dF_{[i,j,r]}(x_i, x_j, x_r) \\ &= \int_{-\infty}^{\infty} J_{(j)}^2(F_{[j]}(x_j)) dF_{[j]}(x_j) \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J_{(i)}(F_{[i]}(x_i)) J_{(r)}(F_{[r]}(x_r)) \\ &\quad \cdot dF_{[i,r]}(x_i, x_r) \\ &= \theta_{jj} \cdot \theta_{ir} = \theta_{ir}.\end{aligned}$$

iii. Si $(i,j) = (r,s)$, $((i,j),(r,s)) \in C_2$, on a alors t_{ij} et t_{rs} qui appartiennent respectivement aux matrices $T_{k'l'}$ et T_{kl} où $(k',l') \neq (k,l)$ avec $1 \leq k' < l' \leq$

K et $1 \leq k < l \leq K$. De la même façon que pour l'ensemble C_1 , on trouve

$$\phi_{(i,j),(r,s)} = \theta_{ir}\theta_{js}.$$

Mais puisque dans le cas présent il y a nécessairement au moins un des deux couples (X_i, X_r) et (X_j, X_s) dont les composantes n'appartiennent pas au même sous-vecteur de X , l'hypothèse d'indépendance des K sous-vecteurs permet de déduire que

$$\phi_{(i,j),(r,s)} = 0$$

puisque au moins un des paramètres d'association θ_{ir} et θ_{js} est nul.

C.Q.F.D.

Par les théorèmes 2.1, 2.2 et 2.3, on déduit le corollaire suivant quelque soit la distribution de X .

Corollaire 2.1. *Sous la condition 1.5 et $H_0^{(K)}$, on a lorsque $n \rightarrow \infty$*

$$\sqrt{n} \begin{pmatrix} \text{vec}(T_{12}) \\ \text{vec}(T_{13}) \\ \vdots \\ \text{vec}(T_{(K-1)K}) \end{pmatrix} \xrightarrow{\mathcal{L}} N_g(0, \Psi)$$

où

$$g = \sum_{1 \leq k} \sum_{k < l \leq K} p_k p_l = \frac{1}{2} \left(p^2 - \sum_{k=1}^K p_k^2 \right)$$

et

$$\Psi = \begin{bmatrix} \Theta_{11} \otimes \Theta_{22} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Theta_{(K-1)(K-1)} \otimes \Theta_{KK} \end{bmatrix}.$$

La différence entre les théorèmes 2.2 et 2.3 et le corollaire 2.1 réside dans le fait que le vecteur pour lequel on obtient la normalité asymptotique n'est plus constitué de toutes les composantes de T . Dans les théorèmes 2.2 et 2.3, on énonce le résultat pour le vecteur $uvec(T)$ de dimension f alors que dans le corollaire 2.1, on s'intéresse au vecteur $(vec(T_{12})', vec(T_{13})', \dots, vec(T_{(K-1)K})')'$ de dimension g . La différence entre ces deux vecteurs est d'abord qu'à l'intérieur du vecteur $uvec(T)$, on retrouve des composantes t_{ij} qui appartiennent aux sous-matrices T_{kk} pour $k = 1, \dots, K$ contrairement au vecteur $(vec(T_{12})', vec(T_{13})', \dots, vec(T_{(K-1)K})')'$. De plus, les composantes qui se retrouvent dans les deux vecteurs ne se présentent pas dans le même ordre.

Pour mieux visualiser ces différences, revenons à l'exemple 2.1. Dans ce cas particulier, $uvec(T) = (\boxed{t_{12}}, t_{13}, t_{23}, t_{14}, t_{24}, \boxed{t_{34}}, t_{15}, t_{25}, \boxed{t_{35}}, \boxed{t_{45}}, t_{16}, t_{26}, t_{36}, t_{46}, t_{56}, t_{17}, t_{27}, t_{37}, t_{47}, t_{57}, \boxed{t_{67}})'$ où les quantités encadrées représentent celles qui appartiennent à des sous-matrices T_{kk} pour $k = 1, \dots, K$, et

$$\begin{pmatrix} vec(T_{12}) \\ vec(T_{13}) \\ vec(T_{23}) \end{pmatrix} = (t_{13}, t_{23}, t_{14}, t_{24}, t_{15}, t_{25}, t_{16}, t_{26}, t_{17}, t_{27}, t_{36}, t_{46}, t_{56}, t_{37}, t_{47}, t_{57})'$$

Afin de déterminer la distribution asymptotique conjointe des statistiques appartenant à la classe \mathcal{C} , on a besoin du résultat suivant concernant la distribution des formes quadratiques que l'on retrouve dans Johnson et Kotz (1972), p.150.

Théorème 2.4. *Soient X un vecteur aléatoire de loi $N_p(0, V)$ et A une matrice symétrique de dimensions $p \times p$. Alors, la forme quadratique $X'AX$ a la même distribution que*

$$\sum_{i=1}^p \lambda_i W_i^2$$

où pour $i = 1, \dots, p$, λ_i sont les valeurs propres de VA et W_i sont des variables aléatoires indépendantes et de même loi $N(0,1)$.

On est maintenant en mesure d'obtenir la distribution asymptotique conjointe des statistiques $R_{kl}^{\hat{\Lambda}_{kl}, J_n} \in \mathcal{C}$ pour $1 \leq k < l \leq K$, qui s'expriment, rappelons-le, sous la forme

$$R_{kl}^{\hat{\Lambda}_{kl}, J_n} = [\text{vec}(T_{lk})]' \hat{\Lambda}_{kl} [\text{vec}(T_{lk})],$$

où $\hat{\Lambda}_{kl} = \Lambda(T_{kk}, T_{ll})$ est une matrice symétrique de dimensions $p_k p_l \times p_k p_l$ fonction des sous-matrices T_{kk} et T_{ll} , k et $l = 1, \dots, K$.

Théorème 2.5. Soit $R_{kl}^{\hat{\Lambda}_{kl}, J_n} \in \mathcal{C}$. Lorsque $n \rightarrow \infty$ et sous $H_0^{(K)}$, on a

$$n \begin{pmatrix} R_{12}^{\hat{\Lambda}_{12}, J_n} \\ R_{13}^{\hat{\Lambda}_{13}, J_n} \\ \vdots \\ R_{(K-1)K}^{\hat{\Lambda}_{(K-1)K}, J_n} \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} Y_{12} \\ Y_{13} \\ \vdots \\ Y_{(K-1)K} \end{pmatrix}$$

où les $\frac{K(K-1)}{2}$ variables aléatoires Y_{kl} sont indépendantes et distribuées comme

$$\sum_{m=1}^{p_k p_l} \lambda_{klm} W_{klm}^2$$

avec pour $1 \leq k < l \leq K$ et $m = 1, \dots, p_k p_l$, les variables aléatoires W_{klm} indépendantes et de même loi $N(0,1)$ et λ_{klm} les valeurs propres de $\Lambda_{kl}(\Theta_{kk} \otimes \Theta_{ll})$ où $\Lambda_{kl} = \Lambda(\Theta_{kk} \otimes \Theta_{ll}) = I \otimes \Delta(\Theta_{kk}, \Theta_{ll})$.

Pour démontrer ce résultat, on procède de manière analogue à la démonstration proposée par Cléroux *et al* (1994) et Cléroux *et al* (1995).

Démonstration: On a pour $1 \leq k < l \leq K$, $R_{kl}^{\hat{\Lambda}_{kl}, J_n} = [\text{vec}(T_{lk})]' \hat{\Lambda}_{kl} [\text{vec}(T_{lk})]$ avec les quantités $\hat{\Lambda}_{kl}$ et $J_{n(i)}(\cdot)$ fixées et conformes aux conditions de la classe \mathcal{C} .

On sait, par le corollaire 2.1 que lorsque $n \rightarrow \infty$,

$$\sqrt{n} \begin{pmatrix} \text{vec}(T_{12}) \\ \text{vec}(T_{13}) \\ \vdots \\ \text{vec}(T_{(K-1)K}) \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} Z_{12} \\ Z_{13} \\ \vdots \\ Z_{(K-1)K} \end{pmatrix}$$

où les vecteurs aléatoires Z_{kl} sont indépendants et de loi $N_{p_{kl}}(0, \Theta_{kk} \otimes \Theta_{ll})$ pour $1 \leq k < l \leq K$.

On sait également que T est un estimateur convergent pour Θ , ainsi, lorsque $n \rightarrow \infty$, on a

$$\begin{pmatrix} \text{vec}(T_{11}) \\ \text{vec}(T_{22}) \\ \vdots \\ \text{vec}(T_{KK}) \end{pmatrix} \xrightarrow{\mathcal{P}} \begin{pmatrix} \text{vec}(\Theta_{11}) \\ \text{vec}(\Theta_{22}) \\ \vdots \\ \text{vec}(\Theta_{KK}) \end{pmatrix}. \quad (2.1)$$

Par le théorème 4.4 de Billingsley (1968), on obtient lorsque $n \rightarrow \infty$

$$\begin{pmatrix} \sqrt{n} \text{vec}(T_{12}) \\ \vdots \\ \sqrt{n} \text{vec}(T_{(K-1)K}) \\ \text{vec}(T_{11}) \\ \vdots \\ \text{vec}(T_{KK}) \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} Z_{12} \\ \vdots \\ Z_{(K-1)K} \\ \text{vec}(\Theta_{11}) \\ \vdots \\ \text{vec}(\Theta_{KK}) \end{pmatrix}. \quad (2.2)$$

D'autre part, les formes quadratiques $R_{kl}^{\hat{\Lambda}_{kl}, J^n} = [\text{vec}(T_{lk})]' \hat{\Lambda}_{kl} [\text{vec}(T_{lk})]$ sont des fonctions continues en tous leurs arguments pour $1 \leq k < l \leq K$. Donc, en utilisant le théorème de Slutsky (voir Billingsley (1968)) et les relations de

convergence (2.1) et (2.2), puisque

$$h(\text{vec}(T_{11}), \dots, \text{vec}(T_{KK}), \sqrt{n}\text{vec}(T_{12}), \dots, \text{vec}(t_{(K-1)K}))$$

est une fonction continue en tous ses arguments, alors

$$h(\text{vec}(T_{11}), \dots, \text{vec}(T_{KK}), \sqrt{n}\text{vec}(T_{12}), \dots, \sqrt{n}\text{vec}(T_{(K-1)K}))$$

$$\xrightarrow{\mathcal{L}}$$

$$h(\text{vec}(\Theta_{11}), \dots, \text{vec}(\Theta_{KK}), Z_{12}, \dots, Z_{(K-1)K}),$$

et on déduit ainsi que pour $1 \leq k < l \leq K$, $R_{kl}^{\hat{\Lambda}_{kl}, J_n} \xrightarrow{\mathcal{L}} \varrho_{kl}$ où $\varrho_{kl} = Z'_{kl} \Lambda_{kl} Z_{kl}$ et où $\Lambda_{kl} = \Lambda(\Theta_{kk}, \Theta_{ll})$ est une matrice symétrique par la condition 1.4. On obtient donc lorsque $n \rightarrow \infty$,

$$n \begin{pmatrix} R_{12}^{\hat{\Lambda}_{12}, J_n} \\ R_{13}^{\hat{\Lambda}_{13}, J_n} \\ \vdots \\ R_{(K-1)K}^{\hat{\Lambda}_{(K-1)K}, J_n} \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} Z'_{12} \Lambda_{12} Z_{12} \\ Z'_{13} \Lambda_{13} Z_{13} \\ \vdots \\ Z'_{(K-1)K} \Lambda_{(K-1)K} Z_{(K-1)K} \end{pmatrix}.$$

Puisque pour $1 \leq k < l \leq K$ et $m = 1, \dots, p_k p_l$, les vecteurs aléatoires Z_{kl} sont indépendants, les formes quadratiques

$$Y_{kl} = Z'_{kl} \Lambda_{kl} Z_{kl}$$

sont indépendantes et par le théorème 2.4, distribuées comme

$$\sum_{m=1}^{p_k p_l} \lambda_{klm} W_{klm}^2$$

où les variables aléatoires W_{klm} sont indépendantes et de même loi $N(0,1)$ et λ_{klm} sont les valeurs propres de $\Lambda_{kl}(\Theta_{kk} \otimes \Theta_{ll})$.

C.Q.F.D.

2.3. TESTS D'INDÉPENDANCE ENTRE PLUSIEURS VECTEURS ALÉATOIRES

Dans la section précédente, on a obtenu la distribution asymptotique conjointe de mesures d'association échantillonnales $R_{kl}^{\hat{\Lambda}_{kl}, J_n} \in \mathcal{C}$ pour $1 \leq k < l \leq K$ sous $H_0^{(K)}$. En se servant de ces résultats, on peut maintenant présenter formellement la classe de tests basés sur les statistiques appartenant à la classe \mathcal{C} pour étudier $H_0^{(K)}$.

Différents raisonnements peuvent permettre de construire un test pour l'hypothèse d'indépendance de K sous-vecteurs. Voyons quelques-unes de ces approches.

Si X suit une loi multinormale, il est alors possible d'exprimer l'hypothèse d'indépendance en terme des matrices de covariance :

$$H_0^{(K)} : \Sigma_{kl} = 0 \text{ pour } 1 \leq k < l \leq K.$$

On peut alors, par exemple, utiliser le test du rapport de vraisemblance maximale (voir Anderson (1984)). Ce test est basé sur la statistique

$$\lambda = \frac{|S|^{\frac{n}{2}}}{\prod_{k=1}^K |S_{kk}|^{\frac{n}{2}}} \quad (2.3)$$

où l'opérateur $|A|$ représente le déterminant de la matrice A et les matrices S et S_{kk} sont les matrices de covariance échantillonnales telles que définies au chapitre 1.

Sous $H_0^{(K)}$, $\lambda^{\frac{2}{n}}$ possède la même distribution que

$$\prod_{i=2}^K \prod_{j=1}^{p_i} V_{ij}$$

où les variables aléatoires V_{ij} sont indépendantes et de loi bêta $(\frac{1}{2}[n+1-p_i^*-j], \frac{1}{2}p_i^*)$ où $p_i^* = \sum_{l=1}^{i-1} p_l$. Plus facile à utiliser, la distribution asymptotique de λ sous $H_0^{(K)}$

lorsque $n \rightarrow \infty$ est donnée (voir Anderson (1984)) par

$$T^{(A)} = -2\rho \log \lambda = -n\rho \log \lambda^{\frac{2}{n}} \xrightarrow{\mathcal{L}} \chi_g^2$$

où $\rho = 1 - \frac{3}{2n} - \frac{p^3 - \sum_{k=1}^K p_k^3}{3n(p^2 - \sum_{k=1}^K p_k^2)}$ et

$$g = \sum_{1 \leq k < l \leq K} p_k p_l = \frac{1}{2} \left(p^2 - \sum_{k=1}^K p_k^2 \right).$$

Toujours sous l'hypothèse de la multinormalité, on peut aussi utiliser le test proposé par Nagao (1973) basé sur la statistique

$$T_N = \frac{1}{2} \text{tr} (SS_D^{-1} - I)^2$$

où S est la matrice de covariance échantillonnale et

$$S_D = \begin{pmatrix} S_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & S_{KK} \end{pmatrix}.$$

La distribution asymptotique de T_N sous $H_0^{(K)}$ lorsque $n \rightarrow \infty$ est

$$T_N \xrightarrow{\mathcal{L}} \chi_g^2.$$

Si la distribution du vecteur X appartient à la famille des distributions elliptiques, Muirhead et Waternaux (1980) suggèrent un test basé sur la statistique (2.3) pour laquelle on a

$$-2 \log \lambda \xrightarrow{\mathcal{L}} (1 + \kappa) \chi_g^2$$

où κ est le coefficient d'aplatissement de la famille de distributions elliptiques.

Enfin, lorsque la distribution de X est quelconque, on ne peut plus utiliser la matrice de covariance pour exprimer l'hypothèse d'indépendance. Puri *et al* (1970) proposent un test analogue au test de rapport de vraisemblance maximale

classique (2.3) en remplaçant les matrices de covariance échantillonnales classiques par les matrices d'association échantillonnales (1.4). Le test est basé sur la statistique

$$\lambda^* = \frac{|T|}{\prod_{k=1}^K |T_{kk}|} \quad (2.4)$$

où les paramètres d'association échantillonnaires (1.3) sont définis pour toutes fonctions $J_{n(i)}(\cdot)$ satisfaisant à la condition 1.5. Les auteurs montrent que sous $H_0^{(K)}$,

$$T^{(PSG)} = -n \log \lambda^* \xrightarrow{\mathcal{L}} \chi_g^2.$$

La classe de tests qui nous intéresse est construite à partir des statistiques $R_{kl}^{\hat{\Lambda}_{kl}, J_n}$ appartenant à la classe \mathcal{C} selon l'approche décrite à la section 1.3. Rappelons que cette approche est inspirée de celle proposée par Lazraq (1989), Allaire et Lepage (1990) et par la suite adaptée par Cléroux *et al* (1994) et Cléroux *et al* (1995) tel que décrit dans l'introduction.

On a donc des statistiques $R_{kl}^{\hat{\Lambda}_{kl}, J_n} \in \mathcal{C}$. Or, par le théorème 2.5, on a sous $H_0^{(K)}$

$$nR_{kl} \xrightarrow{\mathcal{L}} Y_{kl}$$

où pour $1 \leq k < l \leq K$, les Y_{kl} sont indépendantes et distribuées comme

$$\sum_{m=1}^{p_k p_l} \lambda_{klm} W_{klm}^2.$$

Selon l'approche présentée à la section 1.3, on obtient des tests de la forme

$$\text{rejeter } H_0^{(K)} \text{ si } T_{J_n}^{\hat{\Lambda}_{kl}} > c_\alpha \quad (2.5)$$

où

$$T_{J_n}^{\hat{\Lambda}_{kl}} = n \sum_{1 \leq k < l \leq K} \sum R_{kl}^{\hat{\Lambda}_{kl}, J_n}$$

avec $\hat{\Lambda}_{kl} = \Lambda(T_{kk}, T_{ll})$, une matrice symétrique fonction des sous-matrices T_{kk} et T_{ll} et c_α est le quantile d'ordre $1 - \alpha$ de la distribution de

$$\sum_{1 \leq k < l \leq K} \sum_{m=1}^{p_k p_l} \lambda_{klm} W_{klm}^2$$

où les quantités λ_{klm} et W_{klm} sont définies dans le théorème 2.5. Notons qu'en pratique, il faudra bien sûr estimer les quantités λ_{klm} en remplaçant les paramètres d'association θ_{ij} par les paramètres d'association échantillonnaires t_{ij} dans l'expression des quantités λ_{klm} , $1 \leq k < l \leq K$ et $m = 1, \dots, p_k p_l$.

Les tests de la forme (2.5) forment ainsi une classe de tests basés sur des fonctions des rangs pour vérifier l'hypothèse nulle $H_0^{(K)} : \Omega(x) = 1$. Au chapitre 3, on examinera en détails plusieurs cas particuliers de tests qui appartiennent à cette classe, entre autre, les tests obtenus par Cléroux *et al* (1994), Cléroux *et al* (1995) et Lazraq, Lepage et Cléroux (1995). De plus, au chapitre 4, on verra comment calculer les valeurs des points critiques c_α en utilisant l'algorithme d'Imhof (1961).

Pour clore cette section, soulignons que les tests de la classe sont convergents. En effet, on a pour $1 \leq k < l \leq K$

$$R_{kl} \xrightarrow{\mathcal{P}} Q_{kl},$$

où $\xrightarrow{\mathcal{P}}$ désigne la convergence en probabilité lorsque $n \rightarrow \infty$.

Lorsque l'hypothèse $H_0^{(K)}$ est fautive, on a $\varrho_{kl} > 0$ pour au moins un couple (k, l) , $1 \leq k < l \leq K$, d'où

$$\lim_{n \rightarrow \infty} P \left(n \sum_{1 \leq k < l \leq K} R_{kl} > c_\alpha \right) = 1.$$

2.4. DISTRIBUTION ASYMPTOTIQUE DES STATISTIQUES SOUS UNE SUITE DE CONTRE-HYPOTHÈSES

Pour obtenir la puissance asymptotique des tests appartenant à la classe de tests définis par (2.5) à la section 2.3, et pour éventuellement, au chapitre 4, comparer la performance des tests, on a besoin de la distribution asymptotique conjointe des $R_{kl}^{\hat{\Lambda}_{kl}, J_n} \in \mathcal{C}$ sous la contre-hypothèse $H_1 : \Omega(x) \neq 1$.

Tout comme lorsqu'on établit la distribution conjointe des $R_{kl}^{\hat{\Lambda}_{kl}, J_n} \in \mathcal{C}$ sous $H_0^{(K)}$, on recourt à des méthodes asymptotiques pour obtenir des expressions pour la fonction de puissance.

Si $T_{J_n}^{\hat{\Lambda}_{kl}}$ est la statistique du test, la puissance limite du test pour $H_0^{(K)}$ est

$$\lim_{n \rightarrow \infty} P \left(T_{J_n}^{\hat{\Lambda}_{kl}} > c_\alpha | H_1 \right)$$

où c_α est tel que le test est de niveau asymptotique α , c'est-à-dire que

$$\lim_{n \rightarrow \infty} P \left(T_{J_n}^{\hat{\Lambda}_{kl}} > c_\alpha | H_0^{(K)} \right) = \alpha.$$

Toutefois, les tests de la forme (2.5) étant convergents, on a

$$\lim_{n \rightarrow \infty} P \left(T_{J_n}^{\hat{\Lambda}_{kl}} > c_\alpha | H_1 \right) = 1.$$

Alors, comment comparer la performance asymptotique de tests dont les puissances asymptotiques convergent toutes vers 1? Pour contourner ce problème, on utilise une approche qui consiste à considérer une suite de contre-hypothèses

$\{H_{1n}\}_{n \geq 1}$ qui converge vers l'hypothèse nulle, $H_0^{(K)}$, avec une vitesse de convergence adéquate pour s'assurer que la puissance du test proposé est comprise dans l'intervalle ouvert $(0,1)$. Pour déterminer une vitesse de convergence adéquate, on se sert de la notion de contiguïté due à Le Cam (1960).

De façon générale, soit un échantillon aléatoire $X = (X_1, \dots, X_n)$ dont la distribution est donnée par

$$F(x_1, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n), \quad -\infty < x_1, \dots, x_n < \infty. \quad (2.6)$$

Soit une hypothèse nulle simple H_0 et considérons une suite de contre-hypothèses $\{H_{1n}\}_{n \geq 1}$. Considérons pour $n = 1, 2, \dots$, un espace de probabilité $(\mathcal{X}_n, \mathcal{A}_n, \mu_n)$ où \mathcal{X}_n est l'ensemble de toutes les valeurs possibles $x = (x_1, \dots, x_n)$ de X , \mathcal{A}_n est la tribu borélienne des sous-ensembles de \mathcal{X}_n et μ_n est une mesure de probabilité définie par l'équation (2.6). De plus, soit $p_n^{(0)}$, la mesure de probabilité de X sous H_0 et $p_n^{(n)}$, la mesure de probabilité de X sous H_{1n} .

Rappelons que si pour tout $A_n \in \mathcal{A}_n$ tel que $P_n^{(0)}(A_n) \rightarrow 0$, alors $P_n^{(n)}(A_n) \rightarrow 0$ lorsque $n \rightarrow \infty$, on dit que les densités $p_n^{(n)}$ sont contiguës à la densité $p_n^{(0)}$ où $dP_n^{(0)} = p_n^{(0)} d\mu_n$ et $dP_n^{(n)} = p_n^{(n)} d\mu_n$. Par abus de langage, on dit souvent que la suite de contre-hypothèses $\{H_{1n}\}_{n \geq 1}$ est contiguë à l'hypothèse nulle H_0 .

Choisir la suite de contre-hypothèses $\{H_{1n}\}_{n \geq 1}$ de façon à ce que les densités $p_n^{(n)}$ soient contiguës à $p_n^{(0)}$, la densité sous $H_0^{(K)}$, nous assure qu'on obtient la bonne vitesse de convergence pour le calcul de la puissance asymptotique.

Considérons la suite de contre-hypothèses suivantes :

$$H_{1n} : \Omega_{ij}(x, y) = 1 + \frac{\omega_{ij}(x, y)}{\sqrt{n}} \quad (2.7)$$

pour toutes les variables X_i et X_j qui n'appartiennent pas au même sous-vecteur de X , où $\Omega_{ij}(x, y) = \frac{F_{[i,j]}(x, y)}{F_{[i]}(x) \cdot F_{[j]}(y)}$ est la fonction de dépendance introduite au chapitre 1 et $\omega_{ij} \neq 0$.

Si les fonctions $J_{(i)}(\cdot), i = 1, \dots, p$, satisfont la condition 1.1, on peut exprimer θ_{ij} par (1.7) et ainsi, sous H_{1n} définie par (2.7), on a

$$\Theta_{kl} = \frac{A_{kl}}{\sqrt{n}} \text{ pour } 1 \leq k < l \leq K$$

où la matrice A_{kl} est symétrique et est formée des éléments a_{ij} , i et $j = 1, \dots, p$, où

$$a_{ij} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \omega_{ij} F_{[i]}(x_i) F_{[j]}(x_j) J'_i(F_{[i]}(x_i)) J'_j(F_{[j]}(x_j)) dF_{[i]}(x_i) dF_{[j]}(x_j). \quad (2.8)$$

En annexe B, on montre que la suite de contre-hypothèses $\{H_{1n}\}_{n \geq 1}$ définie par (2.12) est contiguë à l'hypothèse nulle H_0 dans le cas où X suit une loi multinormale.

Le théorème suivant donne la loi asymptotique de

$$(vec(T_{12})', vec(T_{13})', \dots, vec(T_{(K-1)K})')'$$

quelque soit la distribution des observations sous l'hypothèse (2.7).

Théorème 2.6. *Sous la condition 1.5 et l'hypothèse (2.7), on a lorsque $n \rightarrow \infty$*

$$\sqrt{n} \begin{pmatrix} vec(T_{12}) \\ vec(T_{13}) \\ \vdots \\ vec(T_{(K-1)K}) \end{pmatrix} \xrightarrow{\mathcal{L}} N_g(\mu, \Psi)$$

$$\text{où } g = \frac{1}{2} \left(p^2 - \sum_{k=1}^K p_k^2 \right),$$

$$\mu = \begin{pmatrix} vec(A_{12}) \\ vec(A_{13}) \\ \vdots \\ vec(A_{(K-1)K}) \end{pmatrix} \text{ et } \Psi = \begin{bmatrix} \Theta_{11} \otimes \Theta_{22} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Theta_{(K-1)(K-1)} \otimes \Theta_{KK} \end{bmatrix}.$$

Démonstration: Posons $X_{i\alpha}$ *i.i.d* $F_{[i]}^{(0)}(x_i)$ et $X_{j\alpha}$ *i.i.d* $F_{[j]}^{(0)}(x_j)$ sous $H_0^{(K)}$. Sous H_{1n} , on a $\Omega_{ij}(x,y) = 1 + \frac{\omega_{ij}(x,y)}{\sqrt{n}}$ et donc

$$F_{[i,j]}^{(n)}(x,y) = F_{[i]}^{(0)}(x)F_{[j]}^{(0)}(y) + \frac{\omega_{ij}(x,y)}{\sqrt{n}}F_{[i]}^{(0)}(x)F_{[j]}^{(0)}(y).$$

Ainsi, les marginales sous H_{1n} , sont données par

$$\begin{aligned} F_{[i]}^{(n)}(x) &= \lim_{y \rightarrow \infty} F_{[i,j]}^{(n)}(x,y) \\ &= F_{[i]}^{(0)}(x) + \frac{F_{[i]}^{(0)}(x)}{\sqrt{n}} \lim_{y \rightarrow \infty} \omega_{ij}(x,y) \end{aligned}$$

et

$$F_{[j]}^{(n)}(y) = F_{[j]}^{(0)}(y) + \frac{F_{[j]}^{(0)}(y)}{\sqrt{n}} \lim_{x \rightarrow \infty} \omega_{ij}(x,y).$$

Ce qui signifie que les observations ne sont plus *i.i.d* comme c'était le cas sous $H_0^{(K)}$: $F_{[i,j]}^{(0)}(x,y) = F_{[i]}^{(0)}(x)F_{[j]}^{(0)}(y)$. Cependant, comme nous l'avons mentionné à la section 2.1, la démonstration du théorème 2.1 ne nécessite pas des observations *i.i.d*. Rappelons que les principales étapes du raisonnement pour s'en convaincre se retrouve à l'Annexe A.

Ainsi, puisque les fonctions $J_{n(i)}(\cdot)$ et $J_{(i)}(\cdot)$, $i = 1, \dots, p$ satisfont la condition 1.5, par le théorème 2.1, on a

$$\sqrt{n}(\text{uvec}(T) - \text{uvec}(\Theta)) \xrightarrow{\mathcal{L}} N_f(0, \Phi).$$

Or, si X suit une loi multinormale $N_p(\mu, \Sigma)$ alors le sous-vecteur BX suit une loi multinormale $N_q(B\mu, B\Sigma B')$ où B est une matrice de dimensions $q \times p$. Ici, le sous-vecteur de $\sqrt{n}(\text{uvec}(T))$ qu'on considère est

$$B\sqrt{n}(\text{uvec}(T)) = \sqrt{n} \begin{pmatrix} \text{vec}(T_{12}) \\ \text{vec}(T_{13}) \\ \vdots \\ \text{vec}(T_{(K-1)K}) \end{pmatrix},$$

et ainsi,

$$B = \begin{pmatrix} 0_p & 1_p & 0_p & 0_p & \cdots & 0_p \\ 0_p & 0_p & 1_p & 0_p & \cdots & 0_p \\ 0_p & 0_p & 0_p & 1_p & \cdots & 0_p \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0_p & 0_p & 0_p & 0_p & \cdots & 1_p \\ 0_p & 0_p & 2_p & 0_p & \cdots & 0_p \\ 0_p & 0_p & 0_p & 2_p & \cdots & 0_p \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0_p & 0_p & 0_p & 0_p & \cdots & 2_p \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0_p & 0_p & 0_p & 0_p & \cdots & K_p \end{pmatrix},$$

où i_p est un vecteur $1 \times p$ de 0 avec un 1 à la i -ième position et $0_p = (0, 0, \dots, 0)$.

Par exemple, $1_p = (1, 0, \dots, 0)$.

En remplaçant les quantités Θ_{kl} par leur valeur sous H_{1n} , soit $\frac{A_{kl}}{\sqrt{n}}$, on obtient donc

$$B\sqrt{n}(\text{uvec}(T)) \xrightarrow{\mathcal{L}} N_g(\mu, \Psi)$$

où

$$\mu = \begin{pmatrix} \text{vec}(A_{12}) \\ \text{vec}(A_{13}) \\ \vdots \\ \text{vec}(A_{(K-1)K}) \end{pmatrix} \text{ et } \Psi = \begin{bmatrix} \Theta_{11} \otimes \Theta_{22} & \cdots & \Theta_{1(K-1)} \otimes \Theta_{2K} \\ \vdots & \ddots & \vdots \\ \Theta_{(K-1)1} \otimes \Theta_{K2} & \cdots & \Theta_{(K-1)(K-1)} \otimes \Theta_{KK} \end{bmatrix}.$$

Enfin, puisque les blocs non diagonaux de Ψ sont de la forme

$$\Theta_{kl} \otimes \Theta_{k'l'} = \frac{A_{kl}}{\sqrt{n}} \otimes \frac{A_{k'l'}}{\sqrt{n}},$$

et qu'ils convergent vers 0 lorsque $n \rightarrow \infty$, on obtient le résultat.

C.Q.F.D.

Dans Johnson et Kotz (1972), p.151, on trouve le résultat suivant.

Théorème 2.7. *Soient X un vecteur aléatoire de loi $N_p(\mu, V)$ et A une matrice symétrique de dimensions $p \times p$. Alors la forme quadratique $X'AX$ a la même distribution que*

$$\sum_{i=1}^p \lambda_i Z_i^2$$

où pour $i = 1, \dots, p$, λ_i sont les valeurs propres de VA , Z_i sont des variables aléatoires indépendantes de loi $N(\delta_i, 1)$, $\delta_i = c_i'(VA)^{-\frac{1}{2}}\mu$, et c_i sont les vecteurs propres orthonormés associés aux valeurs propres λ_i .

Le prochain théorème nous permet d'obtenir la distribution asymptotique conjointe des $R_{kl}^{\hat{\Lambda}_{kl}, J_n} \in \mathcal{C}$ sous la suite de contre-hypothèses $\{H_{1n}\}_{n \geq 1}$ définie par (2.7).

Théorème 2.8. *Si $R_{kl}^{\hat{\Lambda}_{kl}, J_n} \in \mathcal{C}$ et sous H_{1n} , donné par l'équation (2.7), on a lorsque $n \rightarrow \infty$*

$$n \begin{pmatrix} R_{12}^{\hat{\Lambda}_{12}, J_n} \\ R_{13}^{\hat{\Lambda}_{13}, J_n} \\ \vdots \\ R_{(K-1)K}^{\hat{\Lambda}_{(K-1)K}, J_n} \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} V_{12} \\ V_{13} \\ \vdots \\ V_{(K-1)K} \end{pmatrix}$$

où les $\frac{K(K-1)}{2}$ variables V_{kl} pour $1 \leq k < l \leq K$ sont indépendantes et distribuées comme

$$\sum_{m=1}^{p_k p_l} \lambda_{klm} W_{klm}^2$$

avec pour $m = 1, \dots, p_k p_l$, les variables W_{klm} indépendantes de loi $N(\delta_{klm}, 1)$, λ_{klm} les valeurs propres de $\Lambda_{kl}(\Theta_{kk} \otimes \Theta_{ll})$, $\delta_{klm} = c'_{klm} \left(\Theta_{kk}^{-\frac{1}{2}} \otimes \Theta_{ll}^{-\frac{1}{2}} \right) \text{vec}(A_{lk})$ et c_{klm} le vecteur propre orthonormé correspondant à la valeur propre λ_{klm} .

Démonstration: On a pour $1 \leq k < l \leq K$, $R_{kl}^{\hat{\Lambda}_{kl}, J_n} = [\text{vec}(T_{lk})]' \hat{\Lambda}_{kl} [\text{vec}(T_{lk})]$ avec les quantités $\hat{\Lambda}_{kl}$ et $J_{n(i)}(\cdot)$ fixées et conformes aux conditions de la classe \mathcal{C} . Par le théorème 2.6, on a

$$\sqrt{n} \begin{pmatrix} \text{vec}(T_{12}) \\ \text{vec}(T_{13}) \\ \vdots \\ \text{vec}(T_{(K-1)K}) \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} Z_{12} \\ Z_{13} \\ \vdots \\ Z_{(K-1)K} \end{pmatrix}$$

où les vecteurs aléatoires Z_{kl} sont indépendants de loi $N(\text{vec}(A_{kl}), \Theta_{kk} \otimes \Theta_{ll})$ pour $1 \leq k < l \leq K$. On procède ensuite de façon similaire à la démonstration du théorème 2.5.

En utilisant le fait que T est un estimateur convergent pour Θ et le théorème (4.4) de Billingsley (1968), on a la relation de convergence (2.1) lorsque $n \rightarrow \infty$.

Puisque les formes quadratiques $R_{kl}^{\hat{\Lambda}_{kl}, J_n} = [\text{vec}(T_{lk})]' \hat{\Lambda}_{kl} [\text{vec}(T_{lk})]$ sont des fonctions continues en tous leurs arguments pour $1 \leq k < l \leq K$, par les relations de convergence (2.1) et (2.2) et le théorème de Slutsky, on a donc lorsque $n \rightarrow \infty$

$$n \begin{pmatrix} R_{12}^{\hat{\Lambda}_{12}, J_n} \\ R_{13}^{\hat{\Lambda}_{13}, J_n} \\ \vdots \\ R_{(K-1)K}^{\hat{\Lambda}_{(K-1)K}, J_n} \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} Z'_{12} \Lambda_{12} Z_{12} \\ Z'_{13} \Lambda_{13} Z_{13} \\ \vdots \\ Z'_{(K-1)K} \Lambda_{(K-1)K} Z_{(K-1)K} \end{pmatrix}.$$

Puisque les vecteurs aléatoires Z_{kl} sont indépendants et que les matrices $\hat{\Lambda}_{kl}$ sont symétriques par la condition (1.4), les formes quadratiques

$$Y_{kl} = Z'_{kl} \Lambda_{kl} Z_{kl}$$

sont indépendantes et par le théorème 2.7, distribuées comme

$$\sum_{m=1}^{p_k p_l} \lambda_{klm} W_{klm}^2$$

où pour $m = 1, \dots, p_k p_l$, les variables W_{klm} sont indépendantes de loi $N(\delta_{klm}, 1)$, λ_{klm} sont les valeurs propres de $\Lambda_{kl}(\Theta_{kk} \otimes \Theta_{ll})$ où $\Lambda_{kl} = \Lambda(\Theta_{kk} \otimes \Theta_{ll}) = I \otimes \Delta(\Theta_{kk}, \Theta_{ll})$, $\delta_{klm} = c'_{klm} \left(\Theta_{kk}^{-\frac{1}{2}} \otimes \Theta_{ll}^{-\frac{1}{2}} \right) \text{vec}(A_{lk})$ et c_{klm} est le vecteur propre orthonormé correspondant à la valeur propre λ_{klm} .

C.Q.F.D.

Pour clore ce chapitre, soulignons simplement le fait qu'en pratique, les paramètres de la distribution asymptotique, soit les coefficients et les paramètres de décentralité des combinaisons linéaires de khi-deux, sont estimés en remplaçant les θ_{ij} par les paramètres d'association échantillonnaux t_{ij} dans l'expression de ces différents paramètres.

Chapitre 3

ÉTUDE DE CAS PARTICULIERS

3.1. QUELQUES CAS PARTICULIERS POUR LA MESURE D'ASSOCIATION

Dans les deux premiers chapitres, on a présenté nos résultats pour le cas le plus général possible. On veut maintenant voir ce qu'il advient de ces résultats pour des cas particuliers. La première étape sera d'examiner nos résultats pour trois différentes mesures d'association : Stewart et Love (1968), Escoufier (1973) et Cramer et Nicewander (1979).

3.1.1. La mesure de Stewart et Love (1968)

À partir d'un modèle de régression linéaire multivarié pour exprimer $X^{(k)}$ en fonction de $X^{(l)}$, Stewart et Love (1968) ont introduit une mesure d'association entre $X^{(k)}$ et $X^{(l)}$, qui a ensuite été généralisée par Gleason (1976). Cette mesure, qu'on appelle aussi un indice de redondance, peut s'exprimer comme une moyenne des carrés des coefficients de corrélation multiple entre les composantes du vecteur $X^{(k)}$ et le vecteur $X^{(l)}$, pondérée par les variances des composantes de $X^{(k)}$. L'indice de redondance, qu'on notera RI_{kl} représente en somme la proportion de la variance de $X^{(k)}$ expliquée par $X^{(l)}$. On peut aussi exprimer RI_{kl} en terme des

éléments de la matrice de covariance par

$$RI_{kl} = \frac{tr(S_{kl}S_{ll}^{-1}S_{lk})}{tr(S_{kk})}$$

pour $k \neq l$, k et $l = 1, \dots, K$. Par le résultat (1.12), on peut écrire

$$RI_{kl} = [vec(S_{lk})]' \frac{(I_{p_k} \otimes S_{ll}^{-1})}{tr(S_{kk})} [vec(S_{lk})].$$

Dans la population, l'indice de redondance est

$$\rho I_{kl} = \frac{tr(\Sigma_{kl}\Sigma_{ll}^{-1}\Sigma_{lk})}{tr(\Sigma_{kk})} = [vec(\Sigma_{lk})]' \frac{(I_{p_k} \otimes \Sigma_{ll}^{-1})}{tr(\Sigma_{kk})} [vec(\Sigma_{lk})].$$

Notons enfin pour $k \neq l$, k et $l = 1, \dots, K$, que $0 \leq RI_{kl} \leq 1$, RI_{kl} est égal au carré du coefficient de corrélation simple quand $p_k = p_l = 1$, RI_{kl} se réduit au carré du coefficient de corrélation multiple si $p_k = 1$ et $p_l > 1$ et cette mesure n'est pas symétrique et n'est invariante que pour des transformations linéaires de plein rang de $X^{(l)}$ ou des transformations orthogonales de $X^{(k)}$ (voir Cléroux *et al* (1994)).

En se basant sur la mesure de Stewart et Love (1968), on obtient une mesure d'association basée sur des fonctions des rangs en remplaçant la matrice de covariance par la matrice d'association, soit

$$R_{kl}^{(SL)} = \frac{tr(T_{kl}T_{ll}^{-1}T_{lk})}{tr(T_{kk})} = [vec(T_{lk})]' \frac{(I_{p_k} \otimes T_{ll}^{-1})}{tr(T_{kk})} [vec(T_{lk})].$$

Cette mesure fait effectivement partie de la classe de mesures qui nous intéressent puisqu'elle s'exprime sous la forme (1.17) avec

$$\hat{\Lambda}_{kl} = \frac{(I_{p_k} \otimes T_{ll}^{-1})}{tr(T_{kk})} \text{ pour } 1 \leq k < l \leq K.$$

De plus, on a $\hat{\Lambda}_{kl} \xrightarrow{\mathcal{P}} \Lambda_{kl}$ où

$$\Lambda_{kl} = \frac{(I_{p_k} \otimes \Theta_{ll}^{-1})}{tr(\Theta_{kk})}$$

est continue en tous ses arguments pour $1 \leq k < l \leq K$.

On peut ainsi utiliser cette mesure pour vérifier l'hypothèse d'indépendance et en se servant des résultats du chapitre 2, on obtient les deux résultats suivants qui découlent directement des théorèmes 2.5 et 2.8, puisque

$$\Lambda_{kl}(\Theta_{kk} \otimes \Theta_{ll}) = (I_{p_k} \otimes \Theta_{ll}^{-1}) \frac{(\Theta_{kk} \otimes \Theta_{ll})}{\text{tr}(\Theta_{kk})} = \frac{(\Theta_{kk} \otimes I_{p_l})}{\text{tr}(\Theta_{kk})}$$

et que les valeurs propres de $(\Theta_{kk} \otimes I_{p_l})$ sont $\lambda_1^{(k)}, \dots, \lambda_{p_k}^{(k)}$, les valeurs propres de Θ_{kk} chacune avec multiplicité p_l .

Théorème 3.1. *Si les fonctions $J_{(i)}(\cdot)$ et $J_{n(i)}(\cdot)$, $i = 1, \dots, p$ satisfont la condition (1.5) et sous $H_0^{(K)}$, on a lorsque $n \rightarrow \infty$*

$$n \begin{pmatrix} R_{12}^{(SL)} \\ R_{13}^{(SL)} \\ \vdots \\ R_{(K-1)K}^{(SL)} \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} Y_{12} \\ Y_{13} \\ \vdots \\ Y_{(K-1)K} \end{pmatrix}$$

où les $\frac{K(K-1)}{2}$ variables aléatoires Y_{kl} pour $1 \leq k < l \leq K$ sont indépendantes et distribuées comme

$$\frac{1}{\text{tr}(\Theta_{kk})} \sum_{r=1}^{p_k} \sum_{s=1}^{p_l} \lambda_r^{(k)} W_{klrs}^2$$

où les variables aléatoires W_{klrs} sont indépendantes et identiquement distribuées de loi $N(0,1)$ et $\lambda_1^{(k)}, \dots, \lambda_{p_k}^{(k)}$ sont les valeurs propres de Θ_{kk} pour $k = 1, \dots, K$.

Théorème 3.2. *Si les fonctions $J_{(i)}(\cdot)$ et $J_{n(i)}(\cdot)$, $i = 1, \dots, p$ satisfont la condition (1.5) et sous l'hypothèse (2.7), on a lorsque $n \rightarrow \infty$*

$$n \begin{pmatrix} R_{12}^{(SL)} \\ R_{13}^{(SL)} \\ \vdots \\ R_{(K-1)K}^{(SL)} \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} V_{12} \\ V_{13} \\ \vdots \\ V_{(K-1)K} \end{pmatrix}$$

où les $\frac{K(K-1)}{2}$ variables aléatoires V_{kl} pour $1 \leq k < l \leq K$ sont indépendantes et distribuées comme

$$\frac{1}{\text{tr}(\Theta_{kk})} \sum_{r=1}^{p_k} \sum_{s=1}^{p_l} \lambda_r^{(k)} W_{klrs}^2$$

où $\lambda_1^{(k)}, \dots, \lambda_{p_k}^{(k)}$ sont les valeurs propres de Θ_{kk} pour $k = 1, \dots, K$, les variables aléatoires W_{klrs} sont indépendantes de loi $N(\delta_{klrs}, 1)$ avec

$$\delta_{klrs} = c'_{klrs} \left(\Theta_{kk}^{-\frac{1}{2}} \otimes \Theta_{ll}^{-\frac{1}{2}} \right) \text{vec}(A_{lk}),$$

c_{klrs} étant le vecteur propre orthonormé correspondant à la valeur propre $\lambda_r^{(k)}$ avec multiplicité p_l de la matrice $(\Theta_{kk} \otimes I_{p_l})$.

La classe de tests résultant de la mesure de Stewart et Love (1968) est donc de la forme :

$$\text{rejeter } H_0^{(K)} \text{ si } n \sum_{1 \leq k < l \leq K} R_{kl}^{(SL)} > c_\alpha^{(SL)} \quad (3.1)$$

où $c_\alpha^{(SL)}$ est le quantile d'ordre $1 - \alpha$ de la distribution de

$$\sum_{1 \leq k < l \leq K} \sum_{r=1}^{p_k} \sum_{s=1}^{p_l} \frac{1}{\text{tr}(\Theta_{kk})} \lambda_r^{(k)} W_{klrs}^2.$$

On verra au chapitre 4 comment calculer les quantités $c_\alpha^{(SL)}$ en se servant de l'algorithme d'Imhof (1961). Remarquons enfin que si on utilise les cotes linéaires,

c'est-à-dire

$$J_{n(i)}\left(\frac{\alpha}{n}\right) = \sqrt{\frac{12}{n^2 - 1}} \left(\alpha - \frac{n+1}{2}\right)$$

et

$$J_{(i)}(u) = \sqrt{12} \left(u - \frac{1}{2}\right),$$

pour $i = 1, \dots, p$, $\alpha = 1, \dots, n$ et pour $0 < u < 1$, on retrouve les résultats de Cléroux *et al* (1994).

3.1.2. La mesure d'Escoufier (1973)

Alors que la mesure de Stewart et Love(1968) est basée sur la prévision de $X^{(k)}$ par $X^{(l)}$, la mesure d'Escoufier (1973) est définie dans un contexte de distance entre deux matrices de données.

Considérons les matrices suivantes pour $k = 1, \dots, K$,

$$Y_k = \left((X_1^{(k)} - \bar{X}^{(k)}) \quad \dots \quad (X_n^{(k)} - \bar{X}^{(k)}) \right).$$

On définit la distance induite¹ entre les matrices Y_k et Y_l pour k et $l = 1, \dots, K$, $k \neq l$ par

$$\text{dist}(Y_k, Y_l) = \left\| \frac{Y_k' Y_k}{\sqrt{\text{tr}(Y_k' Y_k)^2}} - \frac{Y_l' Y_l}{\sqrt{\text{tr}(Y_l' Y_l)^2}} \right\|$$

où pour toute matrice carrée A , la fonction $\|A\| = \sqrt{\text{tr}(A'A)}$ est une norme.

Soit la mesure

$$RV_{kl} = \frac{\text{tr}(S_{kl} S_{lk})}{\sqrt{\text{tr}(S_{kk}^2) \text{tr}(S_{ll}^2)}}.$$

1. $\text{dist}(Y_k, Y_l)$ n'est pas vraiment une distance à proprement parler puisque $\text{dist}(Y_k, Y_l) = 0$ n'entraîne pas $Y_k = Y_l$ mais plutôt, $\frac{Y_k' Y_k}{\sqrt{\text{tr}(Y_k' Y_k)^2}} = \frac{Y_l' Y_l}{\sqrt{\text{tr}(Y_l' Y_l)^2}}$. C'est pourquoi on parle de "distance induite".

Escoufier (1973) vérifie que $dist(Y_k, Y_l) = \sqrt{2}\sqrt{1 - RV_{kl}}$. On a donc, $dist(Y_k, Y_l) = 0$ si et seulement si $RV_{kl} = 1$.

Remarquons que $0 \leq RV_{kl} \leq 1$. Lorsque $p_k = p_l = 1$, RV_{kl} devient le carré du coefficient de corrélation simple entre les variables X_k et X_l . Enfin, la mesure d'Escoufier (1973) est symétrique, elle est invariante sous des transformations orthogonales de $X^{(k)}$ et $X^{(l)}$ mais elle ne l'est pas sous des transformations linéaires de $X^{(k)}$ et $X^{(l)}$.

Au niveau de la population, on pose pour $k, l = 1, \dots, K$,

$$\rho V_{kl} = \frac{tr(\Sigma_{kl}\Sigma_{lk})}{\sqrt{tr(\Sigma_{kk}^2)tr(\Sigma_{ll}^2)}}.$$

On constate donc que $\rho V_{kl} = 0$ si et seulement si $\Sigma_{kl} = 0$.

En se basant sur la mesure d'Escoufier (1973), on obtient une mesure d'association basée sur des fonctions des rangs en remplaçant la matrice de covariance par la matrice d'association, soit

$$R_{kl}^{(E)} = \frac{tr(T_{kl}T_{lk})}{\sqrt{tr(T_{kk}^2)tr(T_{ll}^2)}},$$

qui peut s'écrire de façon équivalente sous la forme

$$R_{kl}^{(E)} = \frac{[vec(T_{lk})]' [vec(T_{lk})]}{\sqrt{tr(T_{kk}^2)tr(T_{ll}^2)}}$$

en utilisant le résultat (1.12).

Comme on l'a vu pour la mesure $R_{kl}^{(SL)}$, on constate que la mesure $R_{kl}^{(E)}$ fait aussi partie de la classe de mesures qui nous intéressent puisse qu'elle s'exprime sous la forme (1.17) avec

$$\hat{\Lambda}_{kl} = \frac{1}{\sqrt{tr(T_{kk}^2)tr(T_{ll}^2)}} \text{ pour } 1 \leq k < l \leq K.$$

De plus, on a $\hat{\Lambda}_{kl} \xrightarrow{\mathcal{P}} \Lambda_{kl}$ où

$$\Lambda_{kl} = \frac{1}{\sqrt{\text{tr}(\Theta_{kk}^2)\text{tr}(\Theta_{ll}^2)}}$$

est continue en tous ses arguments pour $1 \leq k < l \leq K$.

Théorème 3.3. *Si les fonctions $J_{(i)}(\cdot)$ et $J_{n(i)}(\cdot)$, $i = 1, \dots, p$ satisfont la condition (1.5) et sous $H_0^{(K)}$, on a lorsque $n \rightarrow \infty$*

$$n \begin{pmatrix} R_{12}^{(E)} \\ R_{13}^{(E)} \\ \vdots \\ R_{(K-1)K}^{(E)} \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} Y_{12} \\ Y_{13} \\ \vdots \\ Y_{(K-1)K} \end{pmatrix}$$

où les $\frac{K(K-1)}{2}$ variables aléatoires Y_{kl} pour $1 \leq k < l \leq K$ sont indépendantes et distribuées comme

$$\frac{1}{\sqrt{\text{tr}(\Theta_{kk}^2)\text{tr}(\Theta_{ll}^2)}} \sum_{r=1}^{p_k} \sum_{s=1}^{p_l} \lambda_r^{(k)} \lambda_s^{(l)} W_{klrs}^2$$

où les variables aléatoires W_{klrs} sont indépendantes et identiquement distribuées de loi $N(0,1)$ et $\lambda_1^{(k)}, \dots, \lambda_{p_k}^{(k)}$ sont les valeurs propres de Θ_{kk} pour $k = 1, \dots, K$.

Théorème 3.4. *Si les fonctions $J_{(i)}(\cdot)$ et $J_{n(i)}(\cdot)$, $i = 1, \dots, p$ satisfont la condition (1.5) et sous l'hypothèse (2.7), on a lorsque $n \rightarrow \infty$*

$$n \begin{pmatrix} R_{12}^{(E)} \\ R_{13}^{(E)} \\ \vdots \\ R_{(K-1)K}^{(E)} \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} V_{12} \\ V_{13} \\ \vdots \\ V_{(K-1)K} \end{pmatrix}$$

où les $\frac{K(K-1)}{2}$ variables aléatoires V_{kl} pour $1 \leq k < l \leq K$ sont indépendantes et distribuées comme

$$\frac{1}{\sqrt{\text{tr}(\Theta_{kk}^2)\text{tr}(\Theta_{ll}^2)}} \sum_{r=1}^{p_k} \sum_{s=1}^{p_l} \lambda_r^{(k)} \lambda_s^{(l)} W_{klrs}^2$$

où $\lambda_1^{(k)}, \dots, \lambda_{p_k}^{(k)}$ sont les valeurs propres de Θ_{kk} pour $k = 1, \dots, K$, les variables aléatoires W_{klrs} sont indépendantes de loi $N(\delta_{klrs}, 1)$ avec

$$\delta_{klrs} = c'_{klrs} \left(\Theta_{kk}^{-\frac{1}{2}} \otimes \Theta_{ll}^{-\frac{1}{2}} \right) \text{vec}(A_{lk}),$$

c_{klrs} étant le vecteur propre orthonormé correspondant à la valeur propre $\lambda_r^{(k)} \lambda_s^{(l)}$ de la matrice $(\Theta_{kk} \otimes \Theta_{ll})$.

La classe de tests résultant de la mesure d'Escoufier (1973) est donc de la forme :

$$\text{rejeter } H_0^{(K)} \text{ si } n \sum_{1 \leq k < l \leq K} R_{kl}^{(E)} > c_\alpha^{(E)} \quad (3.2)$$

où $c_\alpha^{(E)}$ est le quantile d'ordre $1 - \alpha$ de la distribution de

$$\sum_{1 \leq k < l \leq K} \sum_{r=1}^{p_k} \sum_{s=1}^{p_l} \frac{1}{\sqrt{\text{tr}(\Theta_{kk}^2)\text{tr}(\Theta_{ll}^2)}} \lambda_r^{(k)} \lambda_s^{(l)} W_{klrs}^2.$$

On remarque encore une fois que si on utilise les cotes linéaires, c'est-à-dire

$$J_{n(i)} \left(\frac{\alpha}{n} \right) = \sqrt{\frac{12}{n^2 - 1}} \left(\alpha - \frac{n+1}{2} \right)$$

et

$$J_{(i)}(u) = \sqrt{12} \left(u - \frac{1}{2} \right),$$

pour $i = 1, \dots, p$, $\alpha = 1, \dots, n$ et pour $0 < u < 1$, on retrouve les résultats de Cléroux *et al* (1995).

3.1.3. La mesure de Cramer et Nicewander (1979)

Tout comme la mesure de Stewart et Love (1968), la mesure de Cramer et Nicewander (1979) est une mesure de redondance en ce sens qu'elle est basée sur la prévision d'un ensemble de variables par un autre.

En fait, la mesure de Cramer et Nicewander (1979), notons-la RI_{kl}^* , est la quantité représentée par la mesure de Stewart et Love (1968) entre $X^{(k)*}$ et $X^{(l)}$ où $X^{(k)*} = OS_{kk}^{-\frac{1}{2}}X^{(k)}$ et O est une matrice orthogonale. De façon plus précise, on transforme $X^{(k)}$ de façon à ce que ses composantes ne soient plus corrélées et aient une variance égale à 1.

Ainsi (voir Allaire et Lepage (1990)), on a

$$RI_{kl}^* = \frac{tr(S_{kk}^{-1}S_{kl}S_{ll}^{-1}S_{lk})}{p_k}.$$

Tout comme RI_{kl} , cette mesure peut s'exprimer en termes des corrélations canoniques entre $X^{(k)}$ et $X^{(l)}$; lorsque $p_k = p_l = 1$, RI_{kl}^* est le carré du coefficient de corrélation simple entre les variables X_k et X_l et lorsque $p_k = 1$, RI_{kl}^* est le carré du coefficient de corrélation multiple entre la variable X_k et le vecteur $X^{(l)}$. De plus, $0 \leq RI_{kl}^* \leq 1$ et cette mesure est symétrique et invariante sous des transformations linéaires de plein rang si on suppose $p_k \leq p_l$ (voir Allaire et Lepage (1990)).

Au niveau de la population, on pose pour k et $l = 1, \dots, K$,

$$\rho I_{kl}^* = \frac{tr(\Sigma_{kk}^{-1}\Sigma_{kl}\Sigma_{ll}^{-1}\Sigma_{lk})}{p_k}.$$

Encore une fois, en utilisant le résultat (1.12), on peut écrire

$$\rho I_{kl}^* = [vec(\Sigma_{lk})]' \left(\frac{\Sigma_{kk}^{-1} \otimes \Sigma_{ll}^{-1}}{p_k} \right) [vec(\Sigma_{lk})],$$

et puisque $(\Sigma_{kk}^{-1} \otimes \Sigma_{ll}^{-1}) = (\Sigma_{kk} \otimes \Sigma_{ll})^{-1}$ est définie positive, $\rho I_{kl}^* = 0$ si et seulement si $\Sigma_{kl} = 0$.

En procédant de la même façon que pour les mesures RI_{kl} et RV_{kl} , on obtient une mesure basée sur RI_{kl}^* en remplaçant S_{kl} par T_{kl} dans la définition de RI_{kl}^* pour k et $l = 1, \dots, K$, soit

$$R_{kl}^{(CN)} = \frac{\text{tr}(T_{kk}^{-1} T_{kl} T_{ll}^{-1} T_{lk})}{p_k} = [\text{vec}(T_{lk})]' \left(\frac{T_{kk}^{-1} \otimes T_{ll}^{-1}}{p_k} \right) [\text{vec}(T_{lk})].$$

La mesure $R_{kl}^{(CN)}$ fait aussi partie de la classe de mesures (1.17) avec

$$\hat{\Lambda}_{kl} = \frac{T_{kk}^{-1} \otimes T_{ll}^{-1}}{p_k} \text{ pour } 1 \leq k < l \leq K.$$

De plus, on a $\hat{\Lambda}_{kl} \xrightarrow{\mathcal{P}} \Lambda_{kl}$ où

$$\Lambda_{kl} = \frac{\Theta_{kk}^{-1} \otimes \Theta_{ll}^{-1}}{p_k} \quad (3.3)$$

est continue en tous ses arguments pour $1 \leq k < l \leq K$.

Puisque $\Lambda_{kl}(\Theta_{kk} \otimes \Theta_{ll}) = \frac{I_{p_k p_l}}{p_k}$ et que les valeurs propres de la matrice identité $I_{p_k p_l}$ sont $p_k p_l$ fois la valeur 1, en utilisant directement les théorèmes 2.5 et 2.8, on obtient les deux résultats suivants.

Théorème 3.5. *Si les fonctions $J_{(i)}(\cdot)$ et $J_{n(i)}(\cdot)$, $i = 1, \dots, p$ satisfont la condition (1.5) et sous $H_0^{(K)}$, on a lorsque $n \rightarrow \infty$*

$$n \begin{pmatrix} R_{12}^{(CN)} \\ R_{13}^{(CN)} \\ \vdots \\ R_{(K-1)K}^{(CN)} \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} Y_{12} \\ Y_{13} \\ \vdots \\ Y_{(K-1)K} \end{pmatrix}$$

où les $\frac{K(K-1)}{2}$ variables aléatoires Y_{kl} pour $1 \leq k < l \leq K$ sont indépendantes et distribuées comme

$$\frac{1}{p_k} \chi_{p_k p_l}^2.$$

Théorème 3.6. Si les fonctions $J_{(i)}(\cdot)$ et $J_{n(i)}(\cdot)$, $i = 1, \dots, p$ satisfont la condition (1.5) et sous l'hypothèse (2.7), on a lorsque $n \rightarrow \infty$

$$n \begin{pmatrix} R_{12}^{(CN)} \\ R_{13}^{(CN)} \\ \vdots \\ R_{(K-1)K}^{(CN)} \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} V_{12} \\ V_{13} \\ \vdots \\ V_{(K-1)K} \end{pmatrix}$$

où les $\frac{K(K-1)}{2}$ variables aléatoires V_{kl} pour $1 \leq k < l \leq K$ sont indépendantes et distribuées comme

$$\frac{1}{p_k} \chi_{p_k p_l}^2 (\delta_{kl}^2)$$

où

$$\delta_{kl}^2 = [\text{vec}(A_{lk})]' (\Theta_{kk} \otimes \Theta_{ll})^{-1} [\text{vec}(A_{lk})].$$

La classe de tests résultant de la mesure de Cramer et Nicewander (1979) est donc de la forme :

$$\text{rejeter } H_0^{(K)} \text{ si } n \sum_{1 \leq k} \sum_{< l \leq K} R_{kl}^{(CN)} > c_\alpha^{(CN)} \quad (3.4)$$

où $c_\alpha^{(CN)}$ est le quantile d'ordre $1 - \alpha$ de la distribution de

$$\sum_{1 \leq k} \sum_{< l \leq K} \frac{V_{kl}}{p_k}$$

où les variables aléatoires V_{kl} sont indépendantes et de loi $\chi_{p_k p_l}^2$.

Cette mesure n'a pas été étudiée par Cléroux *et al* (1994) et Cléroux *et al* (1995) dans un cadre non paramétrique, mais Allaire et Lepage (1990) ont considéré la mesure de Cramer et Nicewander (1979) dans le cas paramétrique, c'est-à-dire en utilisant R_{kl}^* .

3.2. QUELQUES CAS PARTICULIERS POUR LA FONCTION DES RANGS

Le choix des fonctions J est déterminant pour la qualité du test résultant. On verra dans la seconde partie de ce travail, que dépendamment du contexte, le choix d'une fonction des rangs J plutôt qu'une autre peut mener à de meilleurs résultats. Par exemple, lorsque les vecteurs sont de loi multinormale, on verra que l'efficacité asymptotique relative d'un test de la classe basé sur les cotes normales est supérieur à celle d'un test basé sur les cotes linéaires. Mais pour le moment, voyons simplement ce qu'il advient de nos résultats dans les cas où on utilise les cotes linéaires, les cotes normales et les cotes définies comme l'espérance de la α -ième statistique d'ordre.

3.2.1. Les cotes linéaires

Si on choisit d'utiliser les cotes linéaires, on a pour $i = 1, \dots, p$ et $\alpha = 1, \dots, n$,

$$J_{n(i)}\left(\frac{\alpha}{n}\right) = \sqrt{\frac{12}{n^2 - 1}} \left(\alpha - \frac{n + 1}{2}\right) \quad (3.5)$$

qui converge uniformément lorsque $n \rightarrow \infty$ vers

$$J_{(i)}(u) = \sqrt{12} \left(u - \frac{1}{2}\right), \quad 0 < u < 1. \quad (3.6)$$

Il est facile de vérifier que les fonctions (3.5) et (3.6) vérifient les conditions 1.1, 1.2, 1.3 et 1.5.

Les paramètres d'association deviennent alors pour i et $j = 1, \dots, p$,

$$\theta_{ij}^{(L)} = 12 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(F_{[i]}(x_i) - \frac{1}{2} \right) \left(F_{[j]}(x_j) - \frac{1}{2} \right) dF_{[i,j]}(x_i, x_j) \quad (3.7)$$

qui correspond au coefficient de corrélation grade entre X_i et X_j (voir Puri et Sen (1971)). De plus, puisque $F_{n[i]}(X_{i\alpha}) = \frac{R_{i\alpha}}{n}$ où $R_{i\alpha}$ représente le rang de $X_{i\alpha}$ parmi X_{i1}, \dots, X_{in} pour $i = 1, \dots, p$ et $\alpha = 1, \dots, n$, les estimateurs des θ_{ij} sont

$$t_{ij}^{(L)} = \frac{12}{n(n^2 - 1)} \sum_{\alpha=1}^n \left(R_{i\alpha} - \frac{n+1}{2} \right) \left(R_{j\alpha} - \frac{n+1}{2} \right) \quad (3.8)$$

qui correspond au coefficient de corrélation des rangs de Spearman entre X_i et X_j pour $i, j = 1, \dots, p$.

On peut donc utiliser les classes de tests de la section 3.1 en remplaçant les éléments de la matrice d'association et la matrice d'association échantillonnale par les quantités (3.7) et (3.8) respectivement.

Nous verrons au prochain chapitre, au lemme 4.1, que lorsque les observations sont de loi multinormale et lorsqu'on utilise les cotes linéaires, les paramètres d'association peuvent s'exprimer en terme des coefficients de corrélation ρ_{ij} de la façon suivante :

$$\theta_{ij} = \frac{6}{\pi} \arcsin \left(\frac{\rho_{ij}}{2} \right).$$

Notons finalement, comme on l'a souligné dans la section précédente, que dans les cas encore plus particuliers où on a recours à la mesure de Stewart et Love (1968) ou la mesure d'Escoufier (1973), on retrouve les tests obtenus par Cléroux *et al* (1994) et Cléroux *et al* (1995).

3.2.2. Les cotes normales

Considérons maintenant pour $i = 1, \dots, p$, les cotes définies à partir de l'inverse d'une distribution normale centrée et réduite, soit $J_{(i)}(u) = \Phi^{-1}(u)$ avec

$$J_{n(i)}\left(\frac{\alpha}{n+1}\right) = \Phi^{-1}\left(\frac{\alpha}{n+1}\right) \quad (3.9)$$

qui converge lorsque $n \rightarrow \infty$ vers $J_{(i)}(u)$ pour $0 < u < 1$.

Notons que les fonctions (3.9) satisfont 1.1, 1.2, 1.3 et 1.5 (voir Chernoff et Savage (1958), Govindarajulu, Le Cam et Raghavachari (1966), Puri (1964) ou encore, Puri et Sen (1966)).

Les mesures d'association correspondantes sont

$$t_{ij}^{(\Phi)} = \frac{1}{n} \sum_{\alpha=1}^n \Phi^{-1}\left(\frac{R_{i\alpha}}{n+1}\right) \Phi^{-1}\left(\frac{R_{j\alpha}}{n+1}\right), \quad (3.10)$$

et la quantité

$$\frac{nt_{ij}^{(\Phi)}}{\sum_{\alpha=1}^n \left[\Phi^{-1}\left(\frac{i}{n+1}\right)\right]^2}$$

correspond au coefficient de corrélation de type van der Waerden entre les variables X_i et X_j (Hájek et Šidák (1967)).

Les paramètres d'association deviennent

$$\theta_{ij}^{(\Phi)} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi^{-1}(F_{[i]}(x)) \Phi^{-1}(F_{[j]}(y)) dF_{[i,j]}(x,y). \quad (3.11)$$

Différentes classes de tests correspondantes sont obtenues en utilisant les résultats de la section 3.1 pour des matrices d'association et des matrices d'association échantillonnables dont les éléments sont respectivement de la forme (3.11) et (3.10) pour i et $j = 1, \dots, p$.

Certains auteurs dont Chernoff et Savage (1958), Govindarajulu, Le Cam et Raghavachari (1965), Puri (1964) et Puri et Sen (1966), ont aussi construit des

cotes à partir de la distribution inverse de certaines lois continues, comme par exemple, la loi logistique, double-exponentielle et plusieurs autres, mais nous ne nous y attarderons pas dans le présent travail.

Notons qu'Hoeffding (1953) a montré que les cotes normales étaient asymptotiquement équivalentes aux cotes basées sur l'espérance de la α -ième statistique d'ordre d'un échantillon de n variables de loi normale, soient

$$J_{n(i)} \left(\frac{\alpha}{n+1} \right) = E(V^{(\alpha)}).$$

où $V^{(1)} < \dots < V^{(n)}$ sont les statistiques d'ordre d'un échantillon de n variables aléatoires V_1, \dots, V_n , de loi normale centrée et réduite. Pour ces cotes, la condition 2.1 sur les fonctions J est vérifiée (voir entre autre, Puri et Sen (1966)). Dans ce cas, on obtient

$$t_{ij}^{(E_\Phi)} = \frac{1}{n} \sum_{\alpha=1}^n E(V^{(R_{i\alpha})}) E(V^{(R_{j\alpha})}),$$

et la quantité

$$\frac{nt_{ij}^{(E_\Phi)}}{\sum_{\alpha=1}^n [E(V^{(\alpha)})]^2}$$

est appelée le coefficient de corrélation de Fisher-Yates entre les variables X_i et X_j (voir Hájek et Šidák (1967)).

Étant donné l'équivalence asymptotique avec les cotes normales et la difficulté à calculer ces quantités, ces cotes présentent peu d'intérêt en pratique.

Notons enfin qu'au prochain chapitre, plus précisément au lemme 4.1, on montre que dans le cas particulier où les observations proviennent d'une distribution multinormale et lorsqu'on utilise les cotes normales, les paramètres d'association $\theta_{ij} = \rho_{ij}$ où ρ_{ij} est le coefficient de corrélation entre X_i et X_j pour i et $j = 1, \dots, p$ ($i \neq j$).

Passons donc maintenant à la deuxième partie de ce travail qui consiste à étudier et comparer la performance des tests de la classe \mathcal{C} .

Chapitre 4

COMPARAISON DE LA PERFORMANCE DES TESTS

4.1. L'EFFICACITÉ ASYMPTOTIQUE RELATIVE

Dans le chapitre 1, on a construit une classe de tests pour vérifier l'hypothèse de l'indépendance entre plusieurs sous-vecteurs. On a aussi présenté plusieurs tests existants pour le même problème. Une question naturelle se pose : dans la classe de tests, y a-t-il un test plus performant que les autres pour un problème particulier et comment se comparent les tests de la classe par rapport aux tests déjà existants.

Le choix d'un test performant est une partie très importante de la statistique et la façon la plus naturelle de comparer des tests est d'examiner leur fonction de puissance. On peut aussi procéder à une comparaison numérique par simulation informatique. Cependant, lorsque c'est possible, on préfère évidemment comparer les tests de façon analytique et un outil important pour le faire est l'efficacité asymptotique relative.

La notion d'efficacité asymptotique des tests, introduite à la fin des années quarante, diffère de la notion, mieux connue, d'efficacité asymptotique des estimateurs dont on parlait déjà dans les années vingt. Précisons la définition de cette notion.

Soient deux suites de statistiques basées sur n observations, $\{T_{1n}\}_{n \geq 1}$ et $\{T_{2n}\}_{n \geq 1}$, et supposons qu'on veuille confronter des hypothèses de la forme :

$$H_0 : \theta = \theta_0 \text{ versus } H_1 : \theta > \theta_0.$$

Si, à un niveau fixé α , et pour une alternative $\theta > \theta_0$, le test basé sur la suite de statistiques $\{T_{1n}\}_{n \geq 1}$ doit compter sur n_1 observations pour atteindre une certaine puissance alors que le test basé sur la suite de statistiques $\{T_{2n}\}_{n \geq 1}$ n'a besoin que de n_2 observations ($n_1 > n_2$), on aura tendance à croire que le test basé sur la suite de statistiques $\{T_{2n}\}_{n \geq 1}$ est plus performant que celui basé sur la suite de statistiques $\{T_{1n}\}_{n \geq 1}$ pour cette alternative. Donc, en notant par $n_1(\alpha, \beta, \theta)$, la taille échantillonnale nécessaire à la suite $\{T_{1n}\}$ pour avoir un test de puissance β au niveau α pour la valeur θ de l'alternative, et par $n_2(\alpha, \beta, \theta)$, la taille échantillonnale nécessaire à la suite $\{T_{2n}\}$ pour avoir un test de puissance β au niveau α pour la valeur θ de l'alternative, on définit l'efficacité relative de $\{T_{1n}\}$ par rapport à $\{T_{2n}\}$ par

$$e_{1,2}(\alpha, \beta, \theta) = \frac{n_2(\alpha, \beta, \theta)}{n_1(\alpha, \beta, \theta)}.$$

Bien que l'idée soit très simple, le calcul de cette quantité s'avère souvent extrêmement compliqué. Même pour les statistiques les plus simples, il peut être très difficile, voire même impossible de calculer $n_1(\alpha, \beta, \theta)$ et $n_2(\alpha, \beta, \theta)$.

Selon le problème rencontré, différentes approches ont été proposées pour obtenir une expression pour l'efficacité asymptotique relative des tests. L'approche de Pitman (1949) est, sans contredit, la méthode la plus populaire et la plus utilisée. Partant du principe que ce sont les alternatives près de l'hypothèse nulle qu'un bon test doit être capable de déceler, Pitman (1949) propose de calculer la valeur limite de $e_{1,2}(\alpha, \beta, \theta)$ lorsque $\theta \rightarrow \theta_0$.

Selon l'approche de Pitman (1949), on confronte des hypothèses de la forme

$$H_0 : \theta = \theta_0 \text{ versus } H_{1n} : \theta_n = \theta_0 + \frac{\delta_n}{n^r}$$

où $r > 0$ et $\delta_n > 0$ est tel que $\delta_n \rightarrow \delta > 0$ lorsque $n \rightarrow \infty$. Notons aussi par $\gamma_{1n}(\theta_n)$ et $\gamma_{2n}(\theta_n)$ les fonctions de puissance des tests de niveau α basés respectivement sur les suites de statistiques $\{T_{1n}\}_{n \geq 1}$ et $\{T_{2n}\}_{n \geq 1}$. Soit une suite $\{n^* = h(n)\}_{n \geq 1}$ telle que

$$\lim_{n \rightarrow \infty} \gamma_{1n}(\theta_n) = \lim_{n \rightarrow \infty} \gamma_{2n^*}(\theta_n). \quad (4.1)$$

Alors l'efficacité asymptotique relative de $\{T_{2n}\}$ par rapport à $\{T_{1n}\}$ dans le sens de Pitman est donnée par

$$e_{2,1}^P = \lim_{n \rightarrow \infty} \frac{n}{n^*} \quad (4.2)$$

si cette limite existe et est indépendante du choix de $\{\theta_n\}$ et $\{n^*\}$ satisfaisant l'équation (4.1).

Cette approche est principalement utilisée dans le cas particulier où les fonctions de puissance sont telles que

$$\frac{T_{in} - \mu_{in}(\theta)}{\sigma_{in}(\theta)} \xrightarrow{\mathcal{L}} N(0,1)$$

où $\mu_{in}(\theta)$ et $\sigma_{in}(\theta)$ désignent respectivement l'espérance et la variance de T_{in} , avec

- i. $\mu'_{in}(\theta_0) > 0$,
- ii. $\lim_{n \rightarrow \infty} \frac{\mu'_{in}(\theta_n)}{\mu'_{in}(\theta_0)} = 1$, $\lim_{n \rightarrow \infty} \frac{\sigma_{in}(\theta_n)}{\sigma_{in}(\theta_0)} = 1$,
- iii. $\lim_{n \rightarrow \infty} \frac{1}{n^r} \frac{\mu'_{in}(\theta_n)}{\sigma_{in}(\theta_0)} = c^1 > 0$,

1. La quantité c représente l'efficacité du test basé sur $\{T_{ni}\}$.

pour $i = 1, 2$ et où μ'_{in} désigne la dérivée de μ_{in} . Alors (4.2) devient

$$e_{2,1}^P = \lim_{n \rightarrow \infty} \left(\frac{\mu'_{2n}(\theta_0) \sigma_{1n}(\theta_0)}{\mu'_{1n}(\theta_0) \sigma_{2n}(\theta_0)} \right)^{\frac{1}{r}}. \quad (4.3)$$

En observant le développement qui permet d'obtenir ce résultat, on constate que la normalité asymptotique n'est pas essentielle. En fait, il suffit de pouvoir exprimer les puissances des deux tests sous la même forme. Par exemple, si les fonctions de puissance sont des khi-deux décentrés, l'efficacité asymptotique relative dans le sens de Pitman (équation (4.3)) se résume au rapport des paramètres de décentralité. On retrouve dans Kendall et Stuart (1977) une excellente présentation de la démarche à suivre dans le cas non normal.

Il y a cependant des situations où les statistiques des tests n'ont pas la même forme de distribution sous H_0 et sous la contre-hypothèse. Dans ces cas, le calcul de l'efficacité asymptotique relative de Pitman s'avère ou bien impossible ou alors très complexe et le résultat peut dépendre du niveau α et de la puissance β .

Une alternative à l'approche de Pitman a été proposée par Bahadur (1960), Bahadur (1967) et Bahadur (1971). Son approche consiste à évaluer la valeur limite de $e_{1,2}(\alpha, \beta, \theta)$ lorsque $\alpha \rightarrow 0$. Le principe se défend bien puisqu'un niveau petit est évidemment souhaitable pour un test statistique.

Bahadur (1967) montre que sous des conditions simples, l'efficacité asymptotique relative de Pitman et celle de Bahadur donnent le même résultat lorsque les deux statistiques sont asymptotiquement de loi normale. Wieand (1976) apporte un résultat encore plus intéressant en obtenant des conditions pour lesquelles l'efficacité de Pitman existe dans les cas non normal et peut être obtenue en utilisant l'approche de Bahadur.

D'autres approches moins utilisées ont aussi été présentées. Par exemple, Hodges et Lehmann (1956) examinent la valeur limite de $e_{1,2}(\alpha, \beta, \theta)$ lorsque

$\beta \rightarrow 1$. Des approches intermédiaires pour lesquelles on étudie la valeur limite de $e_{1,2}(\alpha, \beta, \theta)$ en fixant un des trois paramètres (α , β ou θ) et en considérant les valeurs limites des deux autres paramètres ($\alpha \rightarrow 0$, $\beta \rightarrow 1$ ou $\theta \rightarrow \theta_0$) ont été proposées par Chernoff (1952) et Kallenberg (1983).

En résumé, il existe deux catégories de mesures de la performance des tests : il y a les mesures d'efficacité locales et non locales. Les mesures d'efficacité locales nécessitent le recours à une suite de contre-hypothèses contiguës à l'hypothèse nulle; on mesure la performance des tests pour des valeurs aux alentours des valeurs sous H_0 . L'efficacité de Pitman est sans contredit la plus populaire représentante des mesures d'efficacité de cette catégorie. D'un autre côté, les mesures d'efficacité non locales, comme l'efficacité de Bahadur par exemple, mesurent la performance des tests pour une alternative fixée lorsque $n \rightarrow \infty$ et leur utilisation nécessite donc la connaissance de la distribution asymptotique sous une alternative fixée.

Dans notre problème, comme nous ne connaissons la distribution asymptotique des statistiques que sous une suite de contre-hypothèses qui convergent vers l'hypothèse nulle, nous utiliserons donc une approche basée sur celle de Pitman pour comparer la performance de nos tests.

4.2. LA FONCTION DE PUISSANCE DES TESTS DE LA CLASSE \mathcal{C}

Pour utiliser l'approche de Pitman, la première étape consiste à obtenir la fonction de puissance des tests à comparer.

On se rappelle que les statistiques des tests appartenant à la classe sont de la forme

$$T_{J_n}^{\hat{\Lambda}_{kl}} = n \sum_{1 \leq k < l \leq K} R_{kl}^{\hat{\Lambda}_{kl}, J_n},$$

et que sous H_{1n} : $\Omega_{ij}(x,y) = 1 + \frac{\omega_{ij}(x,y)}{\sqrt{n}}$ pour toutes les variables X_i et X_j qui n'appartiennent pas au même sous-vecteur de X (équation (2.7)), elles sont distribuées comme une combinaison linéaire de variables aléatoires de loi khi-deux décentrées. Autrement dit, les statistiques $T_{J_n}^{\hat{\Lambda}_{kl}}$ ont la même distribution que

$$Q = \sum_{r=1}^m \lambda_r W_r^2$$

où pour $r = 1, \dots, m$, les variables aléatoires W_r^2 sont indépendantes, de loi khi-deux avec 1 degré de liberté et de paramètre de décentralité δ_r^2 où les quantités δ_r, λ_r et m sont définies par le théorème 2.8. Ainsi, on a besoin de connaître la fonction de puissance d'une statistique de la forme de Q .

Il est bien connu (voir entre autres, Johnson et Kotz (1972)) que la fonction caractéristique d'une variable aléatoire X de loi khi-deux avec ν degrés de liberté et paramètre de décentralité δ^2 est de la forme

$$\varphi_X(t) = (1 - 2it)^{-\frac{\nu}{2}} \exp \left\{ \frac{\delta^2 it}{1 - 2it} \right\} \quad \text{pour } t \in \mathfrak{R}.$$

Posons pour $r = 1, \dots, m$, $\varphi_r(t)$ la fonction caractéristique de W_r^2 et $\varphi_Q(t)$ la fonction caractéristique de Q . On a donc

$$\begin{aligned} \varphi_Q(t) &= \prod_{r=1}^m \varphi_r(\lambda_r t) \\ &= \prod_{r=1}^m (1 - 2i\lambda_r t)^{-\frac{1}{2}} \exp \left\{ \sum_{r=1}^m \frac{\delta_r^2 i\lambda_r t}{1 - 2i\lambda_r t} \right\}. \end{aligned}$$

En utilisant la formule d'inversion qui permet d'obtenir la distribution d'une variable aléatoire à partir de sa fonction caractéristique, Imhof (1961) montre que

$$P(Q > x) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\sin \theta(u)}{u\rho(u)} du \quad (4.4)$$

où pour $u \in \mathfrak{R}$,

$$\theta(u) = \frac{1}{2} \sum_{r=1}^m \left[h_r \tan^{-1}(\lambda_r u) + \frac{\delta_r^2 \lambda_r u}{1 + \lambda_r^2 u^2} \right] - \frac{1}{2} x u$$

et

$$\rho(u) = \prod_{r=1}^m (1 + \lambda_r^2 u^2)^{\frac{1}{4} h_r} \exp \left\{ \frac{1}{2} \sum_{r=1}^m \frac{(\delta_r \lambda_r u)^2}{1 + \lambda_r^2 u^2} \right\}.$$

Les paramètres h_r , pour $r = 1, \dots, m$, sont les degrés de liberté des variables aléatoires W_r^2 dans la forme quadratique $Q = \sum_{r=1}^m \lambda_r W_r^2$. Dans notre cas, $h_r = 1$ pour $r = 1, \dots, m$.

Pour résoudre ce genre d'intégrale, on doit avoir recours à des méthodes numériques d'intégration. La méthode qu'on utilisera est celle proposée par Imhof (1961) qu'on présentera à la section 4.3.2. Mentionnons simplement pour le moment que l'algorithme d'Imhof (1961) permet d'intégrer de façon numérique l'expression (4.4) en contrôlant les erreurs d'intégration associées. On peut ainsi obtenir une estimation aussi précise que l'on veut pour les points critiques de la distribution de Q .

On arrive cependant à la conclusion que l'équation (4.4) ne nous permet donc pas d'obtenir une expression analytique pour l'efficacité asymptotique relative de deux tests avec ce type de distribution.

D'ailleurs, le type de problème qui nous concerne fait survenir plusieurs obstacles au calcul de l'efficacité asymptotique relative. Notons d'abord que, comme on l'a mentionné à la fin de la section précédente, étant donné qu'on ne connaît pas les distributions asymptotiques sous une contre-hypothèse fixe, l'utilisation des méthodes de mesure d'efficacité non locales, comme l'efficacité de Bahadur, par exemple, est impossible. Pour la même raison, les résultats de Wieand (1976) ne sont pas applicables.

D'autre part, même l'utilisation directe de l'approche de Pitman est en général impossible. Comme on le sait, la distribution des statistiques de la classe est une combinaison linéaire de variables aléatoires de loi khi-deux décentrées. Or, les coefficients des khi-deux sont fonction des éléments des matrices d'association Θ_{kk} pour $k = 1, \dots, K$ et cette fonction dépend des valeurs propres des matrices Δ_{kl} qui définissent la statistique du test, soit

$$T_{J_n}^{\hat{\Delta}_{kl}} = n \sum_{1 \leq k} \sum_{l \leq K} tr(T_{kl} \hat{\Delta}_{kl} T_{lk}).$$

Ainsi, les distributions de deux tests de la classe sont généralement de formes différentes, ce qui rend impossible l'utilisation directe de l'approche de Pitman.

En fait, la seule façon d'obtenir des distributions asymptotiques de même forme, est d'avoir tous les coefficients des khi-deux égaux et ceci n'est possible que si les valeurs propres des matrices Δ_{kl} sont toutes les mêmes. Or, c'est précisément ce qui se passe dans le cas où on utilise la mesure de Cramer et Nicewander (1979). En effet, dans ce cas particulier, lorsque les tailles des sous-vecteurs sont toutes égales, les valeurs propres des matrices Δ_{kl} définies dans (3.3) sont toutes les mêmes et donc la distribution asymptotique de la statistique correspondante suit une loi khi-deux décentrée sous H_{1n} .

On verra donc dans la prochaine section que si on considère certains cas particuliers des tests de la classe, il est possible, en adaptant l'approche de Pitman de façon adéquate, de contourner ces obstacles.

4.3. COMPARAISON ASYMPTOTIQUE DANS CERTAINS CAS PARTICULIERS

4.3.1. Classe de tests basés sur la mesure de Cramer et Nicewander (1979)

Le premier cas intéressant est celui de la classe de tests basés sur la mesure de Cramer et Nicewander (1979). Dans ce cas spécifique, la distribution asymptotique de la statistique sous la suite de contre-hypothèses (2.7) convergeant vers $H_0^{(K)}$,

$$H_{1n} : \Omega_{ij}(x_i, x_j) = 1 + \frac{\omega_{ij}(x_i, x_j)}{\sqrt{n}}$$

pour toutes les variables X_i et X_j qui n'appartiennent pas au même sous-vecteur de X , où la fonction de dépendance (1.5) est donnée par

$$\Omega_{ij}(x_i, x_j) = \frac{F_{[i,j]}(x_i, x_j)}{F_{[i]}(x_j) \cdot F_{[j]}(x_i)},$$

est

$$T^{(CN)} \xrightarrow{\mathcal{L}} \sum_{1 \leq k} \sum_{<l \leq K} \frac{1}{p_k} \chi_{p_k p_l}^2(\delta_{kl}^2)$$

où les $\chi_{p_k p_l}^2(\delta_{kl}^2)$ sont des variables aléatoires indépendantes de loi khi-deux décentrée avec $p_k p_l$ degrés de liberté et les paramètres de décentralité δ_{kl}^2 tels que définis dans la théorème 3.6, sont donnés par

$$\delta_{kl}^2 = [\text{vec}(A_{lk})]' (\Theta_{kk} \otimes \Theta_{ll})^{-1} [\text{vec}(A_{lk})]. \quad (4.5)$$

Si on suppose que les sous-vecteurs sont tous de même dimension, c'est-à-dire que $p_1 = p_2 = \dots = p_K = p_0$, on obtient

$$T^{(CN)} \xrightarrow{\mathcal{L}} \frac{1}{p_0} \chi_d^2 \left(\sum_{1 \leq k} \sum_{<l \leq K} \delta_{kl}^2 \right),$$

où $d = \sum \sum_{1 \leq k < l \leq K} p_0^2 = \frac{K(K-1)}{2} p_0^2$. Ainsi, les statistiques $p_0 T^{(CN)}$ sont toutes asymptotiquement de loi khi-deux de degrés de liberté d et de paramètres de décentralité $\sum \sum_{1 \leq k < l \leq K} \delta_{kl}^2$. On élimine ainsi le problème décrit précédemment qui survient lorsque la distribution asymptotique des statistiques est une combinaison linéaire de variables aléatoires de loi khi-deux. La seule quantité qui dépend de la fonction $J(\cdot)$ qu'on utilise est le paramètre de décentralité.

Comme on l'a souligné à la section précédente, il est bien connu que l'efficacité asymptotique dans le sens de Pitman entre deux tests dont les distributions asymptotiques sous la contre-hypothèse sont deux lois khi-deux qui ne diffèrent que par leur paramètre de décentralité, est le rapport des paramètres de décentralité (voir par exemple Puri et Sen (1966)). Ainsi, en utilisant ce résultat, on trouve une expression pour l'efficacité asymptotique relative de deux tests basés sur la mesure de Cramer et Nicewander (1979) pour deux fonctions de cotes différentes.

Théorème 4.1. *Posons $T_1^{(CN)}$ la statistique basée sur la fonction de cotes $J^{(1)}(\cdot)$ et $T_2^{(CN)}$ la statistique basée sur la fonction de cotes $J^{(2)}(\cdot)$. De plus, notons le paramètre de décentralité correspondant à la statistique $T_s^{(CN)}$ par*

$$\delta_{kl}^{2(s)} = \left[\text{vec}(A_{lk}^{(s)}) \right]' \left(\Theta_{kk}^{(s)} \otimes \Theta_{ll}^{(s)} \right)^{-1} \left[\text{vec}(A_{lk}^{(s)}) \right], \quad (4.6)$$

pour $s = 1, 2$.

Alors, l'efficacité asymptotique relative du test basé sur $T_1^{(CN)}$ par rapport à celui basé sur $T_2^{(CN)}$ s'exprime sous la forme

$$e_{1,2} = \frac{\sum \sum_{1 \leq k < l \leq K} \delta_{kl}^{2(1)}}{\sum \sum_{1 \leq k < l \leq K} \delta_{kl}^{2(2)}}. \quad (4.7)$$

On a vu au chapitre 2 (équation (2.7)) qu'en termes de la fonction de dépendance (1.5), la contre-hypothèse s'écrit

$$H_{1n} : \Omega_{ij}(x_i, x_j) = 1 + \frac{\omega_{ij}(x_i, x_j)}{\sqrt{n}}$$

pour toutes les variables X_i et X_j qui n'appartiennent pas au même sous-vecteur de X , où

$$\Omega_{ij}(x_i, x_j) = \frac{F_{[i,j]}(x_i, x_j)}{F_{[i]}(x_i) \cdot F_{[j]}(x_j)}.$$

Ainsi, on obtient

$$\omega_{ij}(x_i, x_j) = \sqrt{n} \left(\frac{F_{[i,j]}(x_i, x_j)}{F_{[i]}(x_i) F_{[j]}(x_j)} - 1 \right). \quad (4.8)$$

Considérons le cas où X suit une loi multinormale de dimension p_0^2 , de moyenne 0 et matrice de covariance Σ^* où $\Sigma^* = ((\rho_{ij}^*))$ avec

$$\rho_{ij}^* = \begin{cases} 1 & \text{si } i = j, \\ \rho_{ij} & \text{si } X_i \text{ et } X_j \text{ appartiennent au même sous-vecteur,} \\ \frac{\lambda_{ij}}{\sqrt{n}} & \text{si } X_i \text{ et } X_j \text{ n'appartiennent pas au même sous-vecteur.} \end{cases} \quad (4.9)$$

On a alors pour i et $j = 1, \dots, p$,

$$F_{[i,j]}(x_i, x_j, \rho_{ij}^*) = \int_{-\infty}^{x_i} \int_{-\infty}^{x_j} \frac{1}{2\pi \sqrt{1 - \rho_{ij}^{*2}}} \exp \left\{ -\frac{1}{2(1 - \rho_{ij}^{*2})} (u^2 - 2\rho_{ij}^* uv + v^2) \right\} \cdot du dv,$$

$$F_{[i]}(x_i) = \Phi(x_i) = \int_{-\infty}^{x_i} \frac{1}{2\sqrt{\pi}} e^{-\frac{t^2}{2}} dt,$$

et la fonction de densité de (X_i, X_j) est donnée par

$$\psi(x_i, x_j, \rho_{ij}^*) = \frac{1}{2\pi \sqrt{1 - \rho_{ij}^{*2}}} \exp \left\{ -\frac{1}{2(1 - \rho_{ij}^{*2})} (x_i^2 - 2\rho_{ij}^* x_i x_j + x_j^2) \right\}.$$

Pour le calcul de (4.7), on utilise le fait qu'on peut exprimer la suite de contre-hypothèses (2.7) en termes des sous-matrices d'association

$$H_{1n} : \Theta_{kl} = \frac{A_{kl}}{\sqrt{n}} \text{ pour } 1 \leq k < l \leq K,$$

où les éléments de la matrice A_{kl} sont donnés par (équation (2.8))

$$a_{ij} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \omega_{ij}(x_i, x_j) F_{[i]}(x_i) F_{[j]}(x_j) J'_{(i)}(F_{[i]}(x_i)) J'_{(j)}(F_{[j]}(x_j)) \cdot dF_{[i]}(x_i) dF_{[j]}(x_j).$$

On doit donc obtenir les valeurs des paramètres d'association θ_{ij} et des éléments des matrices A_{kl} , soient a_{ij} , pour différentes fonctions J , i et $j = 1, \dots, p$ et $1 \leq k < l \leq K$. On constate donc qu'on doit calculer la valeur des paramètres d'association pour les variables appartenant au même sous-vecteur tandis que le calcul des quantités a_{ij} se limite au cas où les variables X_i et X_j n'appartiennent pas au même sous-vecteur, i et $j = 1, \dots, p$. Ainsi, pour le calcul des θ_{ij} , on utilise $\rho_{ij}^* = \rho_{ij}$ (X_i et X_j appartiennent au même sous-vecteur) et pour le calcul des a_{ij} , on utilise $\rho_{ij}^* = \frac{\lambda_{ij}}{\sqrt{n}}$ (X_i et X_j n'appartiennent pas au même sous-vecteur).

Obtenons d'abord une expression pour les paramètres ω_{ij} définis par l'équation (4.8) en utilisant le fait que le vecteur suit une loi multinormale. Développons la fonction $F_{[i,j]}(x_i, x_j) = F_{[i,j]}(x_i, x_j, \rho_{ij}^*)$, pour X_i et X_j n'appartenant pas au même sous-vecteur, autour de $\rho_{ij}^* = 0$. On obtient (voir Puri *et al* (1970))

$$\begin{aligned} F_{[i,j]}(x_i, x_j, \rho_{ij}^*) &= F_{[i,j]}(x_i, x_j, 0) + F'_{[i,j]}(x_i, x_j, 0) \rho_{ij}^* + F''_{[i,j]}(x_i, x_j, 0) \frac{\rho_{ij}^{*2}}{2!} + \dots \\ &= \Phi(x_i) \Phi(x_j) + \frac{\lambda_{ij}}{\sqrt{n}} \phi(x_i) \phi(x_j) + \dots \end{aligned}$$

où $\phi(x)$ est la densité d'une $N(0,1)$ et $F'_{[i,j]}$ et $F''_{[i,j]}$ représentent respectivement les dérivées première et seconde de la fonction $F_{[i,j]}$. On obtient donc, lorsque

$n \rightarrow \infty$,

$$\omega_{ij} = \lambda_{ij} \frac{\phi(x_i)\phi(x_j)}{\Phi(x_i)\Phi(x_j)}.$$

Comparons maintenant le test obtenu en utilisant les cotes normales à celui obtenu en utilisant les cotes linéaires.

Lemme 4.1. *Si X suit une loi multinormale de dimension p_0^2 , de moyenne 0 et de matrice de covariance Σ^* donnée par (4.9), et si $J_{(i)}^{(1)}(u) = \Phi^{-1}(u)$ pour $i = 1, \dots, p$ alors*

$$\theta_{ij}^{(1)} = \rho_{ij} \quad \text{et} \quad a_{ij}^{(1)} = \lambda_{ij}.$$

Démonstration: Pour $i = 1, \dots, p$, les paramètres d'association deviennent

$$\theta_{ij}^{(1)} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_i x_j dF_{[i,j]}(x_i, x_j) = E(X_i X_j) = \rho_{ij}.$$

D'autre part, en utilisant le fait que

$$\frac{df^{-1}(b)}{db} = f^{-1'}(b) = \frac{1}{f'(f^{-1}(b))},$$

(voir Spivak (1980), p.222), les paramètres $a_{ij}^{(1)}$ deviennent

$$\begin{aligned} a_{ij}^{(1)} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \lambda_{ij} \frac{\phi(x_i)\phi(x_j)}{\Phi(x_i)\Phi(x_j)} \cdot \Phi(x_i)\Phi(x_j) \cdot \frac{1}{\phi[\Phi^{-1}(\Phi(x_i))]} \frac{1}{\phi[\Phi^{-1}(\Phi(x_j))]} \\ &\quad \cdot d\Phi_{[i]}(x_i) d\Phi_{[j]}(x_j) \\ &= \lambda_{ij}. \end{aligned}$$

C.Q.F.D.

Lemme 4.2. *Si X suit une loi multinormale de dimension p_0^2 , de moyenne 0 et de matrice de covariance Σ^* donnée par (4.9), et si $J_{(i)}^{(2)}(u) = \sqrt{12}(u - \frac{1}{2})$ pour*

$i = 1, \dots, p$ alors

$$\theta_{ij}^{(2)} = \frac{6}{\pi} \arcsin\left(\frac{\rho_{ij}}{2}\right) \quad \text{et} \quad a_{ij}^{(2)} = \frac{3}{\pi} \lambda_{ij}.$$

Démonstration: Pour i et $j = 1, \dots, p$, on obtient

$$\begin{aligned} a_{ij}^{(2)} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \lambda_{ij} \frac{\phi(x_i)\phi(x_j)}{\Phi(x_i)\Phi(x_j)} \cdot \Phi(x_i)\Phi(x_j) \cdot \sqrt{12}\phi(x_i)\sqrt{12}\phi(x_j) dx_i dx_j \\ &= 12 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi^2(x_i)\phi^2(x_j) dx_i dx_j \\ &= \frac{3}{\pi} \lambda_{ij}, \end{aligned}$$

puisque $\int_{-\infty}^{\infty} \phi^2(x) dx = \frac{2}{\sqrt{\pi}}$.

Pour les paramètres d'association, on a

$$\begin{aligned} \theta_{ij}^{(2)} &= 12 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (F_{[i]}(x_i) - \frac{1}{2})(F_{[j]}(x_j) - \frac{1}{2}) dF_{[i,j]}(x_i, x_j) \\ &= 12 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi(x_i)\Phi(x_j)\psi(x_i, x_j, \rho_{ij}^*) dx_i dx_j - 3 \end{aligned}$$

en utilisant le fait que $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi(x_i)\psi(x_i, x_j, \rho_{ij}^*) dx_i dx_j = \frac{1}{2}$.

L'évaluation de cette intégrale très complexe est possible grâce à une idée ingénieuse de Hotelling et Pabst (1936) qui consiste à remarquer que

$$\frac{\partial}{\partial \rho_{ij}^*} \psi(x_i, x_j, \rho_{ij}^*) = \frac{\partial^2}{\partial x_i \partial x_j} \psi(x_i, x_j, \rho_{ij}^*).$$

Comme

$$\begin{aligned} \frac{\partial}{\partial \rho_{ij}^*} \theta_{ij}^{(2)} &= 12 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi(x_i)\Phi(x_j) \frac{\partial}{\partial \rho_{ij}^*} \psi(x_i, x_j, \rho_{ij}^*) dx_i dx_j \\ &= 12 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi(x_i)\Phi(x_j) \frac{\partial^2}{\partial x_i \partial x_j} \psi(x_i, x_j, \rho_{ij}^*) dx_i dx_j, \end{aligned}$$

en intégrant par parties, on obtient

$$\begin{aligned} \frac{\partial}{\partial \rho_{ij}^*} \theta_{ij}^{(2)} &= 12 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\partial}{\partial x_i} \Phi(x_i) \frac{\partial}{\partial x_j} \Phi(x_j) \psi(x_i, x_j, \rho_{ij}^*) dx_i dx_j \\ &= \frac{6}{\pi \sqrt{4 - \rho_{ij}^{*2}}}. \end{aligned}$$

Par conséquent, on trouve

$$\theta_{ij}^{(2)} = \frac{6}{\pi} \arcsin \left(\frac{\rho_{ij}}{2} \right).$$

C.Q.F.D.

Nous voilà maintenant en mesure de calculer $e_{1,2}$. La proposition suivante établit la supériorité du test basé sur les cotes normales par rapport à celui basé sur les cotes linéaires si on suppose $\rho_{ij} = \rho$ et $\lambda_{ij} = \lambda$ pour i et $j = 1, \dots, p$, pour la classe de tests basée sur la mesure de Cramer et Nicewander (1979).

Théorème 4.2. *Soit $T_1^{(CN)}$, la statistique basée sur les cotes normales et $T_2^{(CN)}$, la statistique basée sur les cotes linéaires. Si X suit une loi multinormale $N_{p_0^2}(0, \Sigma^{**})$ où pour i et $j = 1, \dots, p$,*

$$\rho_{ij}^{**} = \begin{cases} 1 & \text{si } i = j, \\ \rho & \text{si } X_i \text{ et } X_j \text{ appartiennent au même sous-vecteur,} \\ \frac{\lambda}{\sqrt{n}} & \text{si } X_i \text{ et } X_j \text{ n'appartiennent pas au même sous-vecteur,} \end{cases}$$

avec $-1 \leq \lambda \leq 1$, et si $p_1 = p_2 = \dots = p_K = p_0$, $p_0 \geq 2$, alors pour tout $\rho > \frac{-1}{(p_0-1)}$, on a

$$e_{1,2} = \frac{\pi^2 \left(1 + (p_0 - 1) \frac{6}{\pi} \arcsin \left(\frac{\rho}{2}\right)\right)^2}{9 (1 + (p_0 - 1)\rho)^2} > 1.$$

Démonstration: On sait que

$$e_{1,2} = \frac{\sum \sum_{1 \leq k < l \leq K} \delta_{kl}^{(1)2}}{\sum \sum_{1 \leq k < l \leq K} \delta_{kl}^{(2)2}},$$

où pour $1 \leq k < l \leq K$ et $s = 1, 2$, on a

$$\delta_{kl}^{(s)2} = \left[\text{vec}(A_{kl}^{(s)}) \right]' \left(\Theta_{kk}^{(s)-1} \otimes \Theta_{ll}^{(s)-1} \right) \left[\text{vec}(A_{kl}^{(s)}) \right].$$

De plus, par les lemmes 4.1 et 4.2, on a $\theta_{ij}^{(1)} = \rho$, $\theta_{ij}^{(2)} = \frac{6}{\pi} \arcsin\left(\frac{\rho}{2}\right)$, $a_{ij}^{(1)} = \lambda$ et $a_{ij}^{(2)} = \frac{3}{\pi}\lambda$. Pour simplifier la présentation des résultats, posons $\theta_{ij}^{(s)} = r_s$, c'est-à-dire que $r_1 = \rho$ et $r_2 = \frac{6}{\pi} \arcsin\left(\frac{\rho}{2}\right)$. On peut alors écrire

$$\Theta_{kk}^{(s)} = \begin{pmatrix} 1 & r_s & \cdots & r_s \\ r_s & 1 & \cdots & r_s \\ \vdots & \vdots & \ddots & \vdots \\ r_s & r_s & \cdots & 1 \end{pmatrix}_{p_0 \times p_0}.$$

En utilisant la propriété suivante sur les matrices (voir Muirhead (1982), p.580)

$$(A + BCD)^{-1} = A^{-1} - A^{-1}BC(CDA^{-1}BC + C)^{-1}CDA^{-1},$$

où les matrices A, B, C et D sont respectivement de dimensions $n \times n$, $n \times m$, $m \times m$ et $m \times n$ et où A , $(A + BCD)$ et $(CDA^{-1}BC + C)$ sont des matrices régulières¹, on obtient

$$\Theta_{kk}^{(s)-1} = c \begin{pmatrix} 1 + (p_0 - 2)r_s & -r_s & \cdots & -r_s \\ -r_s & 1 + (p_0 - 2)r_s & \cdots & -r_s \\ \vdots & \vdots & \ddots & \vdots \\ -r_s & -r_s & \cdots & 1 + (p_0 - 2)r_s \end{pmatrix}_{p_0 \times p_0},$$

1. Une matrice A est régulière s'il existe une matrice B telle que $AB = I$ où I est la matrice identité.

où $c = \frac{1}{(1-r_s)(1+(p_0-1)r_s)}$. Ainsi, $r_s \neq 1$ et $r_s \neq \frac{-1}{p_0-1}$. Posons

$$B = \begin{pmatrix} 1 + (p_0 - 2)r_s & -r_s & \cdots & -r_s \\ -r_s & 1 + (p_0 - 2)r_s & \cdots & -r_s \\ \vdots & \vdots & \ddots & \vdots \\ -r_s & -r_s & \cdots & 1 + (p_0 - 2)r_s \end{pmatrix}_{p_0 \times p_0}$$

On peut ainsi écrire

$$\Theta_{kk}^{(s)-1} \otimes \Theta_{ll}^{(s)-1} = \frac{1}{(1-r_s)^2(1+(p_0-1)r_s)^2} (B \otimes B).$$

D'autre part, on vérifie facilement que

$$\begin{pmatrix} 1 & \cdots & 1 \end{pmatrix}_{1 \times p_0^2} (B \otimes B)_{p_0^2 \times p_0^2} = \begin{pmatrix} (1-r_s)^2 & \cdots & (1-r_s)^2 \end{pmatrix}_{1 \times p_0^2}$$

et on obtient

$$\begin{aligned} & \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix} (B \otimes B) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} (1-r_s)^2 & \cdots & (1-r_s)^2 \end{pmatrix}_{1 \times p_0^2} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{p_0^2 \times 1} = p_0^2 (1-r_s)^2. \end{aligned}$$

Notons $a_{ij}^{(s)} = v_s$, c'est-à-dire que $v_1 = \lambda$ et $v_2 = \frac{3}{\pi}\lambda$. Puisque pour $s = 1, 2$, $\text{vec}(A_{kl}^{(s)}) = v_s \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$, où $A_{kl}^{(s)}$ est la matrice formée des paramètres $a_{ij}^{(s)}$

correspondant aux sous-vecteurs $X^{(k)}$ et $X^{(l)}$, on obtient

$$\delta_{kl}^{(s)2} = v_s^2 \frac{p_0^2 (1-r_s)^2}{(1-r_s)^2 (1+(p_0-1)r_s)^2} = \frac{v_s^2 p_0^2}{(1+(p_0-1)r_s)^2}.$$

D'où, on trouve

$$e_{1,2} = \frac{\sum \sum \delta_{kl}^{(1)2}}{\sum \sum \delta_{kl}^{(2)2}} = \frac{\frac{K(K-1)}{2} \frac{\lambda^2 p_0^2}{(1+(p_0-1)\rho)^2}}{\frac{K(K-1)}{2} \left(\frac{3}{\pi}\right)^2 \frac{\lambda^2 p_0^2}{(1+(p_0-1)\frac{6}{\pi} \arcsin(\frac{\rho}{2}))^2}}$$

$$= \frac{\pi^2 (1 + (p_0 - 1)\frac{6}{\pi} \arcsin(\frac{\rho}{2}))^2}{9 (1 + (p_0 - 1)\rho)^2}.$$

Il faut d'abord remarquer que $e_{1,2}$ est définie partout sauf pour $\rho = \frac{-1}{(p_0-1)}$ et $p_0 \neq 1$. Lorsque $p_0 = 1$, le problème de l'indépendance de K sous-vecteurs aléatoires se ramène à l'indépendance de K variables aléatoires. Les paramètres de décentralité $\delta_{ij}^{(s)2}$ deviennent

$$\delta_{ij}^{(s)2} = a_{ij}^{(s)2} = \begin{cases} \lambda^2 & \text{si } s = 1, \\ \frac{9}{\pi^2} \lambda^2 & \text{si } s = 2. \end{cases}$$

Ainsi, on trouve

$$e_{1,2} = \frac{\sum \sum \delta_{ij}^{(1)2}}{\sum \sum \delta_{ij}^{(2)2}} = \frac{\frac{K(K-1)}{2} \lambda^2}{\frac{K(K-1)}{2} \frac{9}{\pi^2} \lambda^2} = \frac{\pi^2}{9} > 1,$$

et le test basé sur les cotes normales est asymptotiquement plus efficace au test basé sur les cotes linéaires pour $p_0 = 1$. Lorsque $\rho < \frac{-1}{(p_0-1)}$, $e_{1,2}$ n'est pas toujours supérieure ou égale à 1, mais on peut montrer que $e_{1,2} > 1$ pour $\rho > \frac{-1}{(p_0-1)}$ et $p_0 \geq 2$. Lorsque $\rho > \frac{-1}{(p_0-1)}$, on a $(1 + (p_0 - 1)\frac{6}{\pi} \arcsin(\frac{\rho}{2})) > 0$ et $(1 + (p_0 - 1)\rho) > 0$. Or, $e_{1,2} > 1$ si et seulement si $g(\rho) > 1$ où

$$g(\rho) = \frac{\pi}{3} \left(\frac{1 + (p_0 - 1)\frac{6}{\pi} \arcsin(\frac{\rho}{2})}{1 + (p_0 - 1)\rho} \right).$$

Ceci équivaut à montrer que

$$1 + (p_0 - 1)\frac{6}{\pi} \arcsin\left(\frac{\rho}{2}\right) > \frac{3}{\pi} (1 + (p_0 - 1)\rho)$$

et cette dernière inéquation est équivalente à $h(\rho) > 0$ où

$$h(\rho) = \left(1 - \frac{3}{\pi}\right) + (p_0 - 1) \frac{6}{\pi} \arcsin\left(\frac{\rho}{2}\right) - \frac{3}{\pi}(p_0 - 1)\rho.$$

Puisque

$$\frac{d}{d\rho}h(\rho) = -\frac{3(p_0 - 1)}{\pi} + \frac{3(p_0 - 1)}{\pi\sqrt{1 - \frac{\rho^2}{4}}},$$

on a $\frac{d}{d\rho}h(\rho) = 0$ si et seulement si $\rho = 0$. Le seul point critique de $h(\rho)$ est $\rho = 0$ et $h(0) = 1 - \frac{3}{\pi} = 0,0450703$. Une façon de déterminer si cette valeur est un minimum serait d'examiner la dérivée seconde de $h(\rho)$, mais dans notre cas, $h''(0) = 0$ et on ne peut donc pas s'en servir pour déterminer si $\rho = 0$ est un minimum. D'autre part, on sait que $-\frac{1}{(p_0-1)} < \rho \leq 1$ et notre objectif est de montrer que $h(\rho) \geq 0$. Or,

$$\lim_{\rho \rightarrow \left(\frac{-1}{p_0-1}\right)^+} h(\rho) = 1 - \frac{6}{\pi}(p_0 - 1) \arcsin\left(\frac{1}{2(p_0 - 1)}\right)$$

et

$$h(1) = p_0 \left(1 - \frac{3}{\pi}\right).$$

Il suffit donc de vérifier que $\lim_{\rho \rightarrow \left(\frac{-1}{p_0-1}\right)^+} h(\rho) > 0$ et $h(1) > 0$. En effet, puisque la fonction $h(\rho)$ est continue sur $\left(-\frac{1}{p_0-1}, 1\right]$, que l'unique point critique se trouve en $\rho = 0$ et qu'en ce point $h(0) > 0$, alors si $h(\rho) > 0$ pour les valeurs limites $-\frac{1}{p_0-1}$ et 1, on a $h(\rho) > 0$ pour tout $\rho \in \left(-\frac{1}{p_0-1}, 1\right]$.

On voit directement que $h(1) > 0$ puisque $p_0 \geq 2$. D'un autre côté,

$$\lim_{\rho \rightarrow \left(\frac{-1}{p_0-1}\right)^+} h(\rho) > 0$$

si et seulement si

$$\frac{6}{\pi}(p_0 - 1) \arcsin\left(\frac{1}{2(p_0 - 1)}\right) < 1,$$

ce qui est équivalent à dire que $r(p_0) < 0$ où

$$r(p_0) = \arcsin\left(\frac{1}{2(p_0 - 1)}\right) - \frac{\pi}{6(p_0 - 1)}.$$

On a

$$r'(p_0) = \frac{\pi}{6(p_0 - 1)^2} - \frac{1}{2(p_0 - 1)^2 \sqrt{1 - \frac{1}{4(p_0 - 1)^2}}}.$$

La fonction $r'(p_0) = 0$ si et seulement si $p_0 = 1 + \frac{1}{2\sqrt{1 - \frac{9}{\pi^2}}} = 2,68445$ ou $p_0 = 1 - \frac{1}{2\sqrt{1 - \frac{9}{\pi^2}}} = -0,684453$. Puisque $p_0 \geq 2$, on ne considère que la première solution, soit $p_0 = 2,68445$. Ainsi, pour $p_0 \geq 2$, la fonction $r(p_0)$ a un seul point critique.

D'autre part, la dérivée seconde de $r(p_0)$ est donnée par

$$r''(p_0) = \frac{1}{8 \left(1 - \frac{1}{4(p_0 - 1)^2}\right)^{\frac{3}{2}}} + \frac{1}{(p_0 - 1)^3} \sqrt{1 - \frac{1}{4(p_0 - 1)^2}} - \frac{\pi}{3(p_0 - 1)^3}.$$

Ainsi, puisque $r''(2,68445) = 0,0105852 > 0$, on sait que $p_0 = 2,68445$ est un minimum sur $[2, \infty)$. Or, $r(2) = 0$, $r(2,68445) = -0,00946833$ et $r(p_0) \leq 0$ pour tout $p_0 > 2,68445$. Puisque $p_0 = 2,68445$ est le seul point critique de $r(p_0)$ sur $[2, \infty)$, on a donc $r(p_0) < 0$ pour tout $p_0 \geq 2$. Par conséquent, on a

$$\lim_{\rho \rightarrow \left(\frac{-1}{p_0 - 1}\right)^+} h(\rho) > 0,$$

d'où le résultat escompté.

C.Q.F.D.

Avec le dernier théorème, on a montré que pour la classe de tests obtenus à partir de la mesure de Cramer et Nicewander (1979) et lorsque X suit une loi multinormale avec une structure de covariance particulière, le test basé sur les cotes normales est toujours asymptotiquement plus efficace que le test basé sur les cotes linéaires. De plus, on a obtenu une expression de l'efficacité asymptotique relative de deux tests construits à partir de la mesure de Cramer et Nicewander

(1979). Notons pour $k = 1, 2$, $T_k^{(CN)}$, la statistique basée sur les cotes $J_{(i)}^{(k)}(\cdot)$, $i = 1, \dots, p$. Si X suit une loi multinormale $N_{p_0^2}(0, \Sigma^{**})$ où pour i et $j = 1, \dots, p$,

$$\rho_{ij}^{**} = \begin{cases} 1 & \text{si } i = j, \\ \rho & \text{si } X_i \text{ et } X_j \text{ appartiennent au même sous-vecteur,} \\ \frac{\lambda}{\sqrt{n}} & \text{si } X_i \text{ et } X_j \text{ n'appartiennent pas au même sous-vecteur,} \end{cases}$$

et si $p_1 = p_2 = \dots = p_K = p_0$, $p_0 \geq 2$, alors pour tout $\rho > \frac{-1}{(p_0-1)}$, on peut écrire

$$e_{1,2} = \frac{s_1^2 (1 + (p_0 - 1)r_2)^2}{s_2^2 (1 + (p_0 - 1)r_1)^2}, \quad (4.10)$$

où s_k est la valeur constante qu'on obtient en calculant les $a_{ij}^{(k)}$ et r_k , la valeur constante obtenue en calculant les $\theta_{ij}^{(k)}$ pour $k = 1, 2$. Rappelons que dans le cas particulier où on compare les cotes normales aux cotes linéaires, l'expression (4.10) devient

$$e_{1,2} = \frac{\pi^2 (1 + (p_0 - 1) \frac{\rho}{\pi} \arcsin(\frac{\rho}{2}))^2}{9 (1 + (p_0 - 1)\rho)^2}. \quad (4.11)$$

Examinons maintenant le cas où on étudie l'indépendance de deux sous-vecteurs, c'est-à-dire $K = 2$, et supposons de plus que $p_1 = p_2 = 2$. On a alors

$$X = \left(\left(\begin{array}{c} X_1 \\ X_2 \end{array} \right) \right. \\ \left. \left(\begin{array}{c} X_3 \\ X_4 \end{array} \right) \right).$$

Comme pour le théorème 4.2, le théorème suivant établit encore une fois la supériorité du test obtenu en utilisant les cotes normales par rapport à celui obtenu en utilisant les cotes linéaires, mais cette fois dans le cas où $K = 2$ et X suit une loi multinormale de dimension p_0^2 , de moyenne 0 et matrice de covariance

Σ^* où $\Sigma^* = ((\rho_{ij}^*))$ avec ρ_{ij}^* définis par (4.9), c'est-à-dire

$$\rho_{ij}^* = \begin{cases} 1 & \text{si } i = j, \\ \rho_{ij} & \text{si } X_i \text{ et } X_j \text{ appartiennent au même sous-vecteur,} \\ \frac{\lambda_{ij}}{\sqrt{n}} & \text{si } X_i \text{ et } X_j \text{ n'appartiennent pas au même sous-vecteur.} \end{cases}$$

On remarque que cette matrice est plus générale que celle utilisée dans le théorème 4.2, puisqu'alors $\rho_{ij} = \rho$ et $\lambda_{ij} = \lambda$ pour i et $j = 1, \dots, p$. Toutefois, en considérant cette structure de covariance plus générale, on doit se contenter d'une démonstration numérique.

Théorème 4.3. *Considérons la mesure de Cramer et Nicewander (1979) avec $K = 2$ et $p_1 = p_2 = 2$. Soit $T_1^{(CN)}$, la statistique basée sur les cotes normales et $T_2^{(CN)}$, la statistique basée sur les cotes linéaires. Si X suit une loi multinormale de moyenne 0 et de matrice de covariance Σ^* définie par (4.9), alors pour tout $\rho_{ij}^* \in [-1, 1]$, on a*

$$e_{1,2} \geq 1.$$

avec inégalité stricte pour $\rho_{ij} \in (-1, 1)$.

Notons que l'expression analytique pour $e_{1,2}$ se retrouve à l'équation (4.16) qui suit la démonstration.

Démonstration: Posons

$$Y = \begin{pmatrix} \lambda_{13} \\ \lambda_{14} \\ \lambda_{23} \\ \lambda_{24} \end{pmatrix}, \quad A = \left(\Theta_{11}^{(1)} \otimes \Theta_{22}^{(1)} \right)^{-1} \quad \text{et} \quad B = \left(\Theta_{11}^{(2)} \otimes \Theta_{22}^{(2)} \right)^{-1}.$$

On peut alors écrire

$$e_{1,2} = \frac{Y'AY}{\left(\frac{3}{\pi}\right)Y'B\left(\frac{3}{\pi}\right)Y}$$

puisque $a_{ij}^{(1)} = \lambda_{ij}$ et $a_{ij}^{(2)} = \frac{3}{\pi}\lambda_{ij}$ pour $i = 1,2$ et $j = 3,4$. Ainsi, $e_{1,2} \geq 1$ si et seulement si $Y'(A - \frac{9}{\pi^2}B)Y \geq 0$. Il suffit donc de montrer que la matrice $A - \frac{9}{\pi^2}B$ est définie positive, ce qui est vrai si et seulement si toutes les valeurs propres de $A - \frac{9}{\pi^2}B$ sont supérieures ou égales à 0.

Or, les valeurs propres de $A - \frac{9}{\pi^2}B$ sont

$$\begin{aligned} v_1 = & -9 + \pi^2 + 9\rho_{12} + 9\rho_{34} - 9\rho_{12}\rho_{34} - 6\pi \arcsin\left(\frac{\rho_{12}}{2}\right) \\ & - 6\arcsin\left(\frac{\rho_{34}}{2}\right) + 36\arcsin\left(\frac{\rho_{12}}{2}\right)\arcsin\left(\frac{\rho_{34}}{2}\right) / \\ & [(\rho_{12} - 1)(\rho_{34} - 1)(\pi - 6\arcsin\left(\frac{\rho_{12}}{2}\right))(\pi - 6\arcsin\left(\frac{\rho_{34}}{2}\right))], \end{aligned}$$

$$\begin{aligned} v_2 = & 9 - \pi^2 - 9\rho_{12} + 9\rho_{34} - 9\rho_{12}\rho_{34} + 6\pi \arcsin\left(\frac{\rho_{12}}{2}\right) \\ & - 6\arcsin\left(\frac{\rho_{34}}{2}\right) + 36\arcsin\left(\frac{\rho_{12}}{2}\right)\arcsin\left(\frac{\rho_{34}}{2}\right) / \\ & [(\rho_{12} - 1)(\rho_{34} + 1)(\pi - 6\arcsin\left(\frac{\rho_{12}}{2}\right))(\pi + 6\arcsin\left(\frac{\rho_{34}}{2}\right))], \end{aligned}$$

$$\begin{aligned} v_3 = & 9 - \pi^2 + 9\rho_{12} - 9\rho_{34} - 9\rho_{12}\rho_{34} - 6\pi \arcsin\left(\frac{\rho_{12}}{2}\right) \\ & 6\arcsin\left(\frac{\rho_{34}}{2}\right) + 36\arcsin\left(\frac{\rho_{12}}{2}\right)\arcsin\left(\frac{\rho_{34}}{2}\right) / \\ & [(\rho_{12} + 1)(\rho_{34} - 1)(\pi + 6\arcsin\left(\frac{\rho_{12}}{2}\right))(\pi - 6\arcsin\left(\frac{\rho_{34}}{2}\right))], \end{aligned}$$

$$\begin{aligned} v_4 = & -9 + \pi^2 - 9\rho_{12} - 9\rho_{34} - 9\rho_{12}\rho_{34} + 6\pi \arcsin\left(\frac{\rho_{12}}{2}\right) \\ & 6\arcsin\left(\frac{\rho_{34}}{2}\right) + 36\arcsin\left(\frac{\rho_{12}}{2}\right)\arcsin\left(\frac{\rho_{34}}{2}\right) / \\ & [(\rho_{12} + 1)(\rho_{34} + 1)(\pi + 6\arcsin\left(\frac{\rho_{12}}{2}\right))(\pi + 6\arcsin\left(\frac{\rho_{34}}{2}\right))], \end{aligned}$$

Vérifions donc que les quatre valeurs propres v_1 , v_2 , v_3 et v_4 sont toujours supérieures ou égales à 0.

Remarquons d'abord que puisque pour $(i,j) = (1,2)$ ou $(3,4)$, $-1 \leq \rho_{ij} \leq 1$.
on a

$$(\rho_{ij} - 1) \leq 0, \quad (4.12)$$

$$(\rho_{ij} + 1) \geq 0, \quad (4.13)$$

et enfin,

$$-\frac{\pi}{6} \leq \arcsin\left(\frac{\rho_{ij}}{2}\right) \leq \frac{\pi}{6}$$

ce qui implique que

$$\left(\pi - 6 \arcsin\left(\frac{\rho_{ij}}{2}\right)\right) \geq 0. \quad (4.14)$$

Pour les équations (4.12), (4.13) et (4.14), on obtient une inégalité stricte dès que $\rho_{ij} \in (-1,1)$.

- (i) Par (4.12), (4.13) et (4.14), on sait que le dénominateur de v_1 est ≥ 0 , c'est-à-dire

$$(\rho_{12} - 1)(\rho_{34} - 1) \left(\pi - 6 \arcsin\left(\frac{\rho_{12}}{2}\right)\right) \left(\pi - 6 \arcsin\left(\frac{\rho_{34}}{2}\right)\right) \geq 0,$$

avec inégalité stricte pour $\rho_{ij} \in (-1,1)$. Il suffit donc de montrer que le numérateur de v_1 est toujours supérieur ou égal à 0 pour conclure que $v_1 \geq 0$ pour tout $-1 \leq \rho_{ij} \leq 1$, $(i,j) = (1,2)$ ou $(3,4)$. Or le numérateur de v_1 est une fonction de deux variables, (ρ_{12}, ρ_{34}) et cette fonction implique des fonctions transcendentes, ce qui rend le calcul des extrémums et des racines beaucoup plus complexes. Cependant, en utilisant des algorithmes numériques, on peut trouver le minimum d'une fonction de plusieurs variables, même lorsque celle-ci contient des fonctions transcendentes. On utilise donc un programme de la librairie NAG associée à FORTRAN pour trouver la valeur minimale du numérateur de v_1 . Ce programme est basé sur

l'algorithme modifié de Newton (voir Gill, Murray et Wright (1981)) pour trouver le minimum d'une fonction. On trouve que la valeur minimale pour le numérateur de v_1 est 0 et cette valeur est obtenue si $\rho_{12} = 1$ ou $\rho_{34} = 1$.

- (ii) Par (4.12), (4.13) et (4.14), on sait que le dénominateur de v_2 est ≤ 0 , c'est-à-dire

$$(\rho_{12} - 1)(\rho_{34} + 1) \left(\pi - 6 \arcsin \left(\frac{\rho_{12}}{2} \right) \right) \left(\pi + 6 \arcsin \left(\frac{\rho_{34}}{2} \right) \right) \leq 0,$$

avec inégalité stricte pour $\rho_{ij} \in (-1,1)$. On veut donc montrer que le numérateur de v_2 est toujours inférieur ou égal à 0 pour conclure que $v_2 \geq 0$ pour tout $-1 \leq \rho_{ij} \leq 1$, $(i,j) = (1,2)$ ou $(3,4)$. Encore une fois, en utilisant l'algorithme modifié de Newton, on obtient que la valeur maximale pour le numérateur de v_2 est 0 et cette valeur est obtenue pour $\rho_{12} = 1$ et pour toute valeur de ρ_{34} .

- (iii) On démontre que $v_3 \geq 0$ exactement de la même façon qu'on l'a fait pour v_2 puisque la seule différence entre ces deux valeurs propres est qu'on remplace ρ_{12} par ρ_{34} et vice-versa dans v_3 .
- (iv) Finalement, par (4.12), (4.13) et (4.14), on sait que le dénominateur de v_4 est ≥ 0 , c'est-à-dire

$$(\rho_{12} + 1)(\rho_{34} + 1) \left(\pi + 6 \arcsin \left(\frac{\rho_{12}}{2} \right) \right) \left(\pi + 6 \arcsin \left(\frac{\rho_{34}}{2} \right) \right) \geq 0,$$

avec inégalité stricte pour $\rho_{ij} \in (-1,1)$. On montre donc, de la même façon que dans les cas précédents, c'est-à-dire en utilisant l'algorithme modifié de Newton, que la valeur minimale du numérateur de v_4 est 0 et cette valeur est obtenue lorsque $\rho_{12} = -1$ ou $\rho_{34} = -1$. On peut donc en conclure que $v_4 \geq 0$ pour tout $-1 \leq \rho_{ij} \leq 1$, $(i,j) = (1,2)$ ou $(3,4)$, avec inégalité stricte pour $\rho_{ij} \in (-1,1)$.

C.Q.F.D.

On sait maintenant que le test basé sur les cotes normales est toujours asymptotiquement plus efficace (pour $\rho_{ij} \in (-1,1)$) ou équivalent (pour $\rho_{ij} = 1$ ou -1) au test basé sur les cotes linéaires lorsqu'on se retrouve dans ce cas particulier mais on est aussi en mesure de calculer la valeur de l'efficacité asymptotique relative $e_{1,2}$ quelque soit la distribution des observations ou les fonctions de cotes considérées, en calculant

$$e_{1,2} = \frac{\delta_{12}^{2(1)}}{\delta_{12}^{2(2)}}$$

où pour $s = 1,2$,

$$\delta_{kl}^{2(s)} = \left[\text{vec}(A_{kl}^{(s)}) \right]' \left(\Theta_{kk}^{(s)} \otimes \Theta_{ll}^{(s)} \right)^{-1} \left[\text{vec}(A_{lk}^{(s)}) \right]$$

Puisqu'ici, $K = 2$, $p_1 = p_2 = 2$, on a

$$X = \left(\begin{array}{c} \left(\begin{array}{c} X_1 \\ X_2 \end{array} \right) \\ \left(\begin{array}{c} X_3 \\ X_4 \end{array} \right) \end{array} \right).$$

Ainsi,

$$\Theta = \begin{pmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{pmatrix}, \text{ où } \Theta_{11} = \begin{pmatrix} 1 & \theta_{12} \\ \theta_{21} & 1 \end{pmatrix} \text{ et } \Theta_{22} = \begin{pmatrix} 1 & \theta_{34} \\ \theta_{43} & 1 \end{pmatrix},$$

$$A_{12} = \begin{pmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{pmatrix} \text{ et } \text{vec}(A_{12}) = (a_{13} \ a_{23} \ a_{14} \ a_{24}).$$

De plus, on a

$$\Theta_{11}^{-1} = \frac{1}{1 - \theta_{12}^2} \begin{pmatrix} 1 & -\theta_{12} \\ -\theta_{12} & 1 \end{pmatrix} \quad \text{et} \quad \Theta_{22}^{-1} = \frac{1}{1 - \theta_{34}^2} \begin{pmatrix} 1 & -\theta_{34} \\ -\theta_{34} & 1 \end{pmatrix}.$$

Ainsi,

$$\begin{aligned} \delta_{12}^{2(s)} &= \left[a_{13}^{2(s)} + a_{14}^{2(s)} + a_{23}^{2(s)} + a_{24}^{2(s)} - 2a_{23}^{(s)} a_{24}^{(s)} \theta_{34}^{(s)} \right. \\ &\quad \left. - 2a_{14}^{(s)} \theta_{12}^{(s)} \left(a_{24}^{(s)} - a_{23}^{(s)} \theta_{34}^{(s)} \right) \right. \\ &\quad \left. - 2a_{13}^{(s)} \left(a_{23}^{(s)} \theta_{12}^{(s)} + \left(a_{14}^{(s)} - a_{24}^{(s)} \theta_{12}^{(s)} \right) \theta_{34}^{(s)} \right) \right] / \\ &\quad \left[\left(\theta_{12}^{2(s)} - 1 \right) \left(\theta_{34}^{2(s)} - 1 \right) \right]. \end{aligned} \tag{4.15}$$

Dans le cas particulier où X est de loi multinormale et qu'on compare les cotes normales aux cotes linéaires, puisque

cotes normales	cotes linéaires
$\theta_{12}^{(1)} = \rho_{12}$	$\theta_{12}^{(2)} = \frac{6}{\pi} \arcsin \left(\frac{\rho_{12}}{2} \right)$
$\theta_{34}^{(1)} = \rho_{34}$	$\theta_{34}^{(2)} = \frac{6}{\pi} \arcsin \left(\frac{\rho_{34}}{2} \right)$
$a_{13}^{(1)} = \lambda_{13}$	$a_{13}^{(2)} = \frac{3}{\pi} \lambda_{13}$
$a_{14}^{(1)} = \lambda_{14}$	$a_{14}^{(2)} = \frac{3}{\pi} \lambda_{14}$
$a_{23}^{(1)} = \lambda_{23}$	$a_{23}^{(2)} = \frac{3}{\pi} \lambda_{23}$
$a_{24}^{(1)} = \lambda_{24}$	$a_{24}^{(2)} = \frac{3}{\pi} \lambda_{24}$

on obtient

$$\begin{aligned}
e_{1,2} = & [\lambda_{13}^2 + \lambda_{14}^2 + \lambda_{23}^2 + \lambda_{24}^2 - 2\lambda_{23}\lambda_{24}\rho_{34} \\
& - 2\lambda_{14}\rho_{12}(\lambda_{24} - \lambda_{23}\rho_{34}) - 2\lambda_{13}(\lambda_{23}\rho_{12} + (\lambda_{14} - \lambda_{24}\rho_{12})\rho_{34}) \cdot \\
& (\pi^2 - 36 \arcsin^2(\frac{\rho_{12}}{2})) (\pi^2 - 36 \arcsin^2(\frac{\rho_{34}}{2}))] / \\
& [9(\rho_{12}^2 - 1)(\rho_{34}^2 - 1) \{-12 \arcsin^2(\frac{\rho_{12}}{2}) \cdot \\
& ((\lambda_{13}\lambda_{23} + \lambda_{14}\lambda_{24})\pi - 6(\lambda_{14}\lambda_{23} + \lambda_{13}\lambda_{24}) \arcsin(\frac{\rho_{34}}{2})) + \\
& \pi [(\lambda_{13}^2 + \lambda_{14}^2 + \lambda_{23}^2 + \lambda_{24}^2)\pi - 12(\lambda_{13}\lambda_{14} + \lambda_{23}\lambda_{24}) \arcsin(\frac{\rho_{34}}{2})\}}] \} \\
\end{aligned} \tag{4.16}$$

Notons enfin que si on suppose que

$$\begin{cases} \rho_{12} = \rho_{34} = \rho, \\ \lambda_{13} = \lambda_{14} = \lambda_{23} = \lambda_{24} = \lambda, \end{cases}$$

alors

$$e_{1,2} = \frac{\pi^2 \left(1 + \frac{6}{\pi} \arcsin\left(\frac{\rho}{2}\right)\right)^2}{9(1 + \rho)^2},$$

qui correspond bien à l'expression (4.11) pour $p_0 = 2$.

En résumé, lorsqu'on utilise la mesure de Cramer et Nicewander (1979), dès qu'on considère K vecteurs de même taille, c'est-à-dire $p_1 = \dots = p_K = p_0$, le calcul de l'efficacité asymptotique relative est relativement simple; il suffit d'évaluer les paramètres de décentralité (4.6). Cependant, lorsqu'on veut une expression analytique de l'efficacité asymptotique relative, il devient rapidement compliqué de le faire pour des cas plus généraux. Cependant, lorsqu'on veut comparer les cotes normales aux cotes linéaires, on réussit à obtenir une expression pour $e_{1,2}$ en supposant la multinormalité et en donnant une structure particulière à la matrice de covariance dans les propositions 4.2 et 4.3.

Avant de clore cette section, comparons la performance du test $T_{\Phi}^{(CN)}$, le test obtenu en utilisant les cotes normales avec deux tests existants dans la littérature

qu'on a présenté au chapitre 2 : le test du rapport de vraisemblance maximale ($T^{(A)}$) et le test proposé par Puri *et al* (1970) ($T^{(PSG)}$) dans le cas où $K = 2$ et où X suit une loi multinormale de dimension $p = p_1 + p_2$ de moyenne 0 et de matrice de covariance $\Sigma^* = ((\rho_{ij}^*))$ où pour $i = 1, \dots, p_1$ et $j = 1, \dots, p_2$,

$$\rho_{ij}^* = \begin{cases} 1 & \text{si } i = j, \\ \rho_{ij} & \text{si } X_i \text{ et } X_j \text{ appartiennent au même sous-vecteur,} \\ \frac{\lambda_{ij}}{\sqrt{n}} & \text{si } X_i \text{ et } X_j \text{ n'appartiennent pas au même sous-vecteur.} \end{cases}$$

Sous ces conditions et sous une suite de contre-hypothèses $H_{1n} : \Theta_{12} = \frac{A_{12}}{\sqrt{n}}$, Puri *et al* (1970) établissent que

$$T^{(PSG)} \xrightarrow{\mathcal{L}} \chi_{p_1 p_2}^2 \left([vec(A_{12})]' (\Theta_{11}^{-1} \otimes \Theta_{22}^{-1}) [vec(A_{21})] \right)$$

et

$$T^{(A)} \xrightarrow{\mathcal{L}} \chi_{p_1 p_2}^2 (\delta_{12}^{2*}),$$

où δ_{12}^{2*} est obtenu en remplaçant θ_{ij} par ρ_{ij} et a_{ij} par λ_{ij} dans (4.15).

On constate immédiatement que $e_{T_{\Phi}^{(CN)}, T^{(A)}} = 1$, autrement dit, le test du rapport de vraisemblance maximale et le $T_{\Phi}^{(CN)}$ sont asymptotiquement équivalents. On constate qu'ici, les tailles des sous-vecteurs p_1 et p_2 sont arbitraires.

D'autre part, si $p_1 = p_2 = 2$, en utilisant le théorème 4.3, on peut comparer la performance de $T_{\Phi}^{(CN)}$ à celle de $T_{\Phi}^{(PSG)}$ et $T_L^{(PSG)}$ où $T_{\Phi}^{(PSG)}$ et $T_L^{(PSG)}$ représentent respectivement les tests de Puri *et al* (1970) obtenus en utilisant les cotes normales (Φ) et les cotes linéaires (L), c'est-à-dire

$$T_J^{(PSG)} = -n \log \lambda^* \quad \text{où} \quad \lambda^* = \frac{|T|}{\prod_{k=1}^K |T_{kk}|},$$

où les matrices et sous-matrices d'association T et T_{kk} , $k = 1, \dots, K$, sont obtenues en utilisant les cotes J . Ainsi, par la proposition 4.3, on peut montrer que

dans le cas où X suit une loi multivariée avec la structure de covariance donnée par (4.9), on a

$$e_{T_{\Phi}^{(CN)}, T_L^{(PSG)}} \geq 1,$$

et

$$e_{T_{\Phi}^{(CN)}, T_{\Phi}^{(PSG)}} = 1.$$

On sait qu'en utilisant la mesure de Cramer et Nicewander (1979), on est en mesure d'exprimer la distribution asymptotique des statistiques correspondantes par des khi-deux décentrés sous H_{1n} , ce qui a permis de faire toutes les comparaisons analytiques présentées jusqu'ici. Malheureusement, il n'en est pas de même pour les autres mesures de la classe : on se retrouve avec des statistiques dont les distributions sont des combinaisons linéaires de variables aléatoires de loi khi-deux et comme les coefficients des combinaisons linéaires de khi-deux sont définis à partir des valeurs propres des matrices Θ_{ii} , $i = 1, \dots, K$, ils sont différents d'un test à l'autre, on perd donc l'avantage d'avoir le rapport de deux khi-deux décentrés, ce qui implique qu'on ne peut utiliser directement l'approche de Pitman pour le calcul des efficacités asymptotiques relatives puisque les distributions sont alors de formes différentes.

On peut cependant contourner cet obstacle en utilisant une astuce proposée par Nyblom et Mäkeläinen (1983) et l'algorithme d'Imhof (1961) dont on a brièvement parlé dans la section précédente, qui permettront d'obtenir une approximation pour l'efficacité asymptotique relative de deux tests de la classe.

4.3.2. Approche de Nyblom et Mäkeläinen (1983)

Voyons d'abord en quoi consiste l'utilisation de l'algorithme d'Imhof (1961) dans le cadre de notre problème.

Pour réussir à comparer des tests de formes différentes, nous allons utiliser le théorème suivant tiré de Nyblom et Mäkeläinen (1983). Mais rappelons d'abord quelques notations introduites au début du chapitre 4. On considère deux suites de statistiques $\{T_{1n}\}_{n \geq 1}$ et $\{T_{2n}\}_{n \geq 1}$ pour confronter les hypothèses

$$H_0 : \theta = \theta_0 \text{ versus } H_{1n} : \theta_n = \theta_0 + \frac{\delta_n}{n^r}$$

et qui, sous H_{1n} , ont pour fonctions de puissance $\gamma_{1n}(\theta_n)$ et $\gamma_{2n}(\theta_n)$ respectivement.

Supposons que les limites

$$\lim_{n \rightarrow \infty} \gamma_{in}(\theta_n) = \gamma_i(\delta), \quad i = 1, 2$$

existent, on a alors le théorème suivant.

Théorème 4.4. *Si δ_1 et δ_2 sont tels que $\gamma_1(\delta_1) = \gamma_2(\delta_2) = \gamma$ et si γ_1 et γ_2 sont strictement croissantes, alors*

$$e_{1,2} = \left(\frac{\delta_2}{\delta_1} \right)^{\frac{1}{r}}.$$

Ainsi, supposons deux statistiques telles que

$$T_{1n} \xrightarrow{\mathcal{L}} \begin{cases} \sum a_{1k} \chi_1^2(0) & \text{sous } H_0^{(K)}, \\ \sum a_{1k} \chi_1^2(b_{1k}) & \text{sous } H_{1n}, \end{cases}$$

et

$$T_{2n} \xrightarrow{\mathcal{L}} \begin{cases} \sum a_{2k} \chi_1^2(0) & \text{sous } H_0^{(K)}, \\ \sum a_{2k} \chi_1^2(b_{2k}) & \text{sous } H_{1n}, \end{cases}$$

et définissons $c_{1\alpha}$ et $c_{2\alpha}$ tels que

$$P(T_{1n} > c_{1\alpha} | H_0^{(K)}) = \alpha \text{ et } P(T_{2n} > c_{2\alpha} | H_0^{(K)}) = \alpha.$$

L'idée, pour comparer la performance de ces deux tests, est de trouver les quantités $\delta_1^2(\alpha, \gamma)$ et $\delta_2^2(\alpha, \gamma)$ telles que

$$P\left(\sum a_{1k} \chi_1^2(\delta_1^2(\alpha, \gamma) b_{1k}) > c_{1\alpha} | H_{1n}\right) = \gamma,$$

et

$$P\left(\sum a_{2k} \chi_1^2(\delta_2^2(\alpha, \gamma) b_{2k}) > c_{2\alpha} | H_{1n}\right) = \gamma,$$

et l'efficacité asymptotique relative de T_{1n} par rapport à T_{2n} , qui dépend alors de α et γ , est donnée par

$$e_{1,2}(\alpha, \gamma) = \frac{\delta_1^2(\alpha, \gamma)}{\delta_2^2(\alpha, \gamma)}. \quad (4.17)$$

Donc, pour calculer $e_{1,2}(\alpha, \gamma)$, on doit obtenir les valeurs $\delta_1^2(\alpha, \gamma)$ et $\delta_2^2(\alpha, \gamma)$. Pour ce faire, on doit être en mesure de calculer les probabilités associées à la distribution de $\sum a_k \chi_{r_k}^2(b_k)$ et c'est précisément ce que l'algorithme d'Imhof (1961) permet d'obtenir.

Présentons donc brièvement en quoi consiste cette méthode numérique, mais soulignons d'abord qu'on retrouve dans la thèse de Lazraq (1989), une annexe consacrée à la distribution des formes quadratiques ainsi qu'à l'algorithme d'Imhof (1961). Nous référons donc le lecteur aux travaux de Lazraq (1989) pour plus de détails.

Considérons la forme quadratique $Q = X'AX$ où le vecteur aléatoire X suit une loi $N_p(0, \Sigma)$. On a vu dans la section 4.2 que la fonction de répartition de Q s'exprime sous la forme

$$P(Q > x) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\sin \theta(u)}{u \rho(u)} du \quad (4.18)$$

où pour $u \in \mathfrak{R}$,

$$\theta(u) = \frac{1}{2} \sum_{r=1}^m \left[h_r \tan^{-1}(\lambda_r u) + \frac{\delta_r^2 \lambda_r u}{1 + \lambda_r^2 u^2} \right] - \frac{1}{2} x u$$

et

$$\rho(u) = \prod_{r=1}^m (1 + \lambda_r^2 u^2)^{\frac{1}{4} h_r} \exp \left\{ \frac{1}{2} \sum_{r=1}^m \frac{(\delta_r \lambda_r u)^2}{1 + \lambda_r^2 u^2} \right\}$$

où λ_r sont les m valeurs propres distinctes de $A\Sigma$, chacune de multiplicité h_r .

Imhof (1961) suggère de procéder par intégration numérique pour calculer $P(Q > x)$. En utilisant le fait que

$$\lim_{u \rightarrow 0} \frac{\sin \theta(u)}{u \rho(u)} = \frac{1}{2} \sum_{r=1}^m \lambda_r (h_r + \delta_r^2) - \frac{1}{2} x$$

et

$$\lim_{u \rightarrow \infty} \theta(u) = \begin{cases} -\infty & \text{si } x > 0, \\ \infty & \text{si } x < 0, \\ \frac{\pi}{4} \sum_{r=1}^m h_r \lambda_r |\lambda_r|^{-1} & \text{si } x = 0. \end{cases}$$

Puisque la fonction $u\rho(u)$ croît de façon monotone vers ∞ , il en déduit qu'on peut procéder à l'intégration numérique de (4.18) sur l'intervalle fini $0 \leq u \leq U$. La précision de l'approximation obtenue dépendra ainsi de 2 sources d'erreur: l'erreur de troncature, c'est-à-dire

$$\frac{1}{\pi} \int_U^\infty \frac{\sin \theta(u)}{u \rho(u)} du$$

et l'erreur d'intégration due à la méthode numérique utilisée pour intégrer

$$\frac{1}{\pi} \int_0^U \frac{\sin \theta(u)}{u \rho(u)} du.$$

De plus, il montre que l'erreur de troncature est bornée supérieurement par T_U où

$$T_U^{-1} = \pi k U^k \prod_{r=1}^m |\lambda_r|^{\frac{1}{2} h_r} \exp \left\{ \frac{1}{2} \sum_{r=1}^m \delta_r^2 \lambda_r^2 U^2 (1 + \lambda_r^2 U^2)^{-1} \right\}$$

avec $k = \frac{1}{2} \sum_{r=1}^m h_r$.

Lorsqu'on veut utiliser une procédure d'intégration numérique, on détermine d'abord l'erreur de troncature qu'on est prêt à tolérer et ensuite, on choisit la valeur U correspondante. On retrouve dans Lazraq (1989) ou encore Abrahamse et Koerts (1969) un algorithme simple pour obtenir en quelques étapes une valeur de U offrant une erreur de troncature acceptable.

En ce qui concerne l'erreur d'intégration, elle résulte de la méthode d'intégration numérique utilisée. Dans le cadre du présent travail, nous utiliserons la règle de Simpson dont le principe est d'intégrer une fonction $f(x)$ sur un intervalle $[a, b]$ en divisant l'intervalle en parties égales jusqu'à ce que le degré de précision voulu soit atteint. On peut trouver le programme Fortran FQUAD qu'on a utilisé pour obtenir une approximation pour $F(x) = P(Q > x)$ dans le livre de Abrahamse et Koerts (1969) ou dans la thèse de Lazraq (1989).

4.3.3. Quelques comparaisons numériques

En utilisant les résultats de la section précédente, procédons maintenant à la comparaison de la performance des tests de la classe pour quelques cas particuliers; la loi normale bivariée et la loi logistique bivariée.

Supposons d'abord $X = \left(\begin{array}{c} \left(X_1 \right) \\ \left(X_2 \right) \\ \left(X_3 \right) \\ \left(X_4 \right) \end{array} \right)$, c'est-à-dire $K = 2$ et $p_1 = p_2 = 2$.

Définissons quelques quantités dont on aura besoin.

On a d'abord besoin des matrices

$$\Theta_{11} = \begin{pmatrix} 1 & \theta_{12} \\ \theta_{12} & 1 \end{pmatrix}, \quad \Theta_{22} = \begin{pmatrix} 1 & \theta_{34} \\ \theta_{34} & 1 \end{pmatrix}$$

et

$$A_{12} = \begin{pmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{pmatrix}.$$

Soit le théorème suivant que l'on retrouve, entre autre, dans Muirhead (1982), à la page 588.

Théorème 4.5. *Soit A une matrice $m \times m$ symétrique, non définie négative. On a alors*

$$A^{\frac{1}{2}} = HD^{\frac{1}{2}}H',$$

où $D = \text{diag}(\lambda_1, \dots, \lambda_m)$, c'est-à-dire une matrice diagonale formée de $\lambda_1, \dots, \lambda_m$, les valeurs propres de la matrice A , la matrice H est telle que $H'AH = D$, donc $H = (h_1 \cdots h_m)$ où pour $i = 1, \dots, m$, les h_i sont les vecteurs propres associés aux valeurs propres de A , et enfin, $D^{\frac{1}{2}} = \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_m})$.

En utilisant le théorème 4.5, on obtient

$$\Theta_{11}^{-\frac{1}{2}} = \frac{1}{4} \left(\frac{1}{\sqrt{1+\theta_{12}}} + \frac{1}{\sqrt{1-\theta_{12}}} \right) \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \quad (4.19)$$

et

$$\Theta_{22}^{-\frac{1}{2}} = \frac{1}{4} \left(\frac{1}{\sqrt{1+\theta_{34}}} + \frac{1}{\sqrt{1-\theta_{34}}} \right) \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \quad (4.20)$$

De plus,

$$\Theta_{11}^{-1} = \frac{1}{1-\theta_{12}^2} \begin{pmatrix} 1 & -\theta_{12} \\ -\theta_{12} & 1 \end{pmatrix} \quad \text{et} \quad \Theta_{22}^{-1} = \frac{1}{1-\theta_{34}^2} \begin{pmatrix} 1 & -\theta_{34} \\ -\theta_{34} & 1 \end{pmatrix}. \quad (4.21)$$

Et enfin, on obtient aussi le tableau 4.1 qui présente les valeurs propres ainsi que les vecteurs propres associés pour les matrices Θ_{11} , Θ_{22} , $\Theta_{11} \otimes \Theta_{22}$ et $\Theta_{11} \otimes I_2$.

TABLEAU 4.1. Valeurs propres et vecteurs propres des matrices Θ_{11} , Θ_{22} , $\Theta_{11} \otimes \Theta_{22}$ et $\Theta_{11} \otimes I_2$

matrice	valeurs propres	vecteurs propres
Θ_{11}	$1 - \theta_{12}$	$\begin{pmatrix} -1 & 1 \end{pmatrix}'$
	$1 + \theta_{12}$	$\begin{pmatrix} 1 & 1 \end{pmatrix}'$
Θ_{22}	$1 - \theta_{34}$	$\begin{pmatrix} -1 & 1 \end{pmatrix}'$
	$1 + \theta_{34}$	$\begin{pmatrix} 1 & 1 \end{pmatrix}'$
$\Theta_{11} \otimes \Theta_{22}$	$1 + \theta_{12} - \theta_{34} - \theta_{12}\theta_{34}$	$\begin{pmatrix} -1 & 1 & -1 & 1 \end{pmatrix}'$
	$1 - \theta_{12} + \theta_{34} - \theta_{12}\theta_{34}$	$\begin{pmatrix} -1 & -1 & 1 & 1 \end{pmatrix}'$
	$1 - \theta_{12} - \theta_{34} + \theta_{12}\theta_{34}$	$\begin{pmatrix} 1 & -1 & -1 & 1 \end{pmatrix}'$
	$1 + \theta_{12} + \theta_{34} + \theta_{12}\theta_{34}$	$\begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}'$
$\Theta_{11} \otimes I_2$	$1 - \theta_{12}$	$\begin{pmatrix} 0 & -1 & 0 & 1 \end{pmatrix}'$
	$1 - \theta_{12}$	$\begin{pmatrix} -1 & 0 & 1 & 0 \end{pmatrix}'$
	$1 + \theta_{12}$	$\begin{pmatrix} 0 & 1 & 0 & 1 \end{pmatrix}'$
	$1 + \theta_{12}$	$\begin{pmatrix} 1 & 0 & 1 & 0 \end{pmatrix}'$

Si on utilise la mesure de Stewart et Love (1968), par le théorème 3.2, on a

$$T^{(SL)} \xrightarrow{\mathcal{L}} \sum_{1 \leq i < j \leq K} \sum_{k=1}^{p_i} \sum_{l=1}^{p_j} \frac{1}{\text{tr}(\Theta_{ii})} \lambda_k^{(i)} \lambda_l^{(j)} (\delta_{ijkl}^2)$$

où $\delta_{ijkl} = c'_{ijkl} \left(\Theta_{ii}^{-\frac{1}{2}} \otimes \Theta_{jj}^{-\frac{1}{2}} \right) \text{vec}(A_{ji})$, $\lambda_1^{(i)} \geq \dots \geq \lambda_{p_i}^{(i)}$ sont les valeurs propres de Θ_{ii} et c_{ijkl} est le vecteur propre correspondant à la valeur propre $\lambda_k^{(i)}$ de multiplicité p_j de la matrice $(\Theta_{ii} \otimes I_{p_j})$.

Ainsi, dans le cas $K = 2$ et $p_1 = p_2 = 2$, et en utilisant (4.19), (4.20), (4.21) et le tableau 4.1, on obtient

$$T^{(SL)} \xrightarrow{\mathcal{L}} \sum_{k=1}^4 b_k^{(SL)} \chi_1^2 \left(d_k^{(SL)} \right) \quad (4.22)$$

où

$$b_1^{(SL)} = b_2^{(SL)} = \frac{1}{2}(1 - \theta_{12}) \quad , \quad b_3^{(SL)} = b_4^{(SL)} = \frac{1}{2}(1 + \theta_{12}),$$

$$d_1^{(SL)} = (0 \quad -1 \quad 0 \quad 1) \left(\Theta_{11}^{-\frac{1}{2}} \otimes \Theta_{22}^{-\frac{1}{2}} \right) \begin{pmatrix} a_{13} \\ a_{23} \\ a_{14} \\ a_{24} \end{pmatrix},$$

$$d_2^{(SL)} = (-1 \quad 0 \quad 1 \quad 0) \left(\Theta_{11}^{-\frac{1}{2}} \otimes \Theta_{22}^{-\frac{1}{2}} \right) \begin{pmatrix} a_{13} \\ a_{23} \\ a_{14} \\ a_{24} \end{pmatrix},$$

$$d_3^{(SL)} = (0 \quad 1 \quad 0 \quad 1) \left(\Theta_{11}^{-\frac{1}{2}} \otimes \Theta_{22}^{-\frac{1}{2}} \right) \begin{pmatrix} a_{13} \\ a_{23} \\ a_{14} \\ a_{24} \end{pmatrix},$$

$$d_4^{(SL)} = (1 \quad 0 \quad 1 \quad 0) \left(\Theta_{11}^{-\frac{1}{2}} \otimes \Theta_{22}^{-\frac{1}{2}} \right) \begin{pmatrix} a_{13} \\ a_{23} \\ a_{14} \\ a_{24} \end{pmatrix}.$$

Pour la mesure d'Escoufier (1973), par le théorème 3.4, on a

$$T^{(E)} \xrightarrow{\mathcal{L}} \sum_{1 \leq i} \sum_{< j \leq K} \frac{1}{\sqrt{\text{tr}(\Theta_{ii}^2) \text{tr}(\Theta_{jj}^2)}} \sum_{k=1}^{p_i} \sum_{l=1}^{p_j} \lambda_k^{(i)} \lambda_l^{(j)} \chi_1^2(\delta_{ijkl}^2)$$

où $\delta_{ijkl} = c'_{ijkl} \left(\Theta_{ii}^{-\frac{1}{2}} \otimes \Theta_{jj}^{-\frac{1}{2}} \right) \text{vec}(A_{ji})$ et c_{ijkl} est le vecteur propre correspondant à la valeur propre $\lambda_k^{(i)} \lambda_l^{(j)}$ de la matrice $(\Theta_{ii} \otimes \Theta_{jj})$.

Encore une fois, en se servant des résultats (4.19), (4.20), (4.21) et du tableau 4.1, on obtient, dans le cas $K = 2$ et $p_1 = p_2 = 2$,

$$T^{(E)} \xrightarrow{\mathcal{L}} \sum_{k=1}^4 b_k^{(E)} \chi_1^2(d_k^{(E)}) \quad (4.23)$$

où

$$b_1^{(E)} = \frac{(1 - \theta_{12})(1 - \theta_{34})}{\sqrt{4(1 + \theta_{12}^2)(1 + \theta_{34}^2)}}, \quad b_2^{(E)} = \frac{(1 - \theta_{12})(1 + \theta_{34})}{\sqrt{4(1 + \theta_{12}^2)(1 + \theta_{34}^2)}},$$

$$b_3^{(E)} = \frac{(1 + \theta_{12})(1 - \theta_{34})}{\sqrt{4(1 + \theta_{12}^2)(1 + \theta_{34}^2)}}, \quad b_4^{(E)} = \frac{(1 + \theta_{12})(1 + \theta_{34})}{\sqrt{4(1 + \theta_{12}^2)(1 + \theta_{34}^2)}},$$

$$d_1^{(E)} = \begin{pmatrix} -1 & 1 & -1 & 1 \end{pmatrix} \left(\Theta_{11}^{-\frac{1}{2}} \otimes \Theta_{22}^{-\frac{1}{2}} \right) \begin{pmatrix} a_{13} \\ a_{23} \\ a_{14} \\ a_{24} \end{pmatrix},$$

$$d_2^{(E)} = \begin{pmatrix} -1 & -1 & 1 & 1 \end{pmatrix} \left(\Theta_{11}^{-\frac{1}{2}} \otimes \Theta_{22}^{-\frac{1}{2}} \right) \begin{pmatrix} a_{13} \\ a_{23} \\ a_{14} \\ a_{24} \end{pmatrix},$$

$$d_3^{(E)} = (1 \quad -1 \quad -1 \quad 1) \left(\Theta_{11}^{-\frac{1}{2}} \otimes \Theta_{22}^{-\frac{1}{2}} \right) \begin{pmatrix} a_{13} \\ a_{23} \\ a_{14} \\ a_{24} \end{pmatrix},$$

$$d_4^{(E)} = (1 \quad 1 \quad 1 \quad 1) \left(\Theta_{11}^{-\frac{1}{2}} \otimes \Theta_{22}^{-\frac{1}{2}} \right) \begin{pmatrix} a_{13} \\ a_{23} \\ a_{14} \\ a_{24} \end{pmatrix}.$$

Enfin, pour la mesure de Cramer et Nicewander (1979), par le théorème 3.6, on a

$$T^{(CN)} \xrightarrow{\mathcal{L}} \sum_{1 \leq i < j \leq K} \sum_{p_i} \frac{1}{p_i} \chi_{p_i p_j}^2(\delta_{ij}^2)$$

où $\delta_{ij} = [\text{vec}(A_{ji})]' (\Theta_{ii}^{-1} \otimes \Theta_{jj}^{-1}) [\text{vec}(A_{ji})]$. Ainsi, dans le cas $K = 2$ et $p_1 = p_2 = 2$, et en utilisant (4.19), (4.20), (4.21) et le tableau 4.1, on obtient

$$T^{(CN)} \xrightarrow{\mathcal{L}} \frac{1}{2} \chi_4^2(d^{(CN)}) \quad (4.24)$$

où

$$d^{(CN)} = (a_{13} \quad a_{23} \quad a_{14} \quad a_{24}) \left(\Theta_{11}^{-1} \otimes \Theta_{22}^{-1} \right) \begin{pmatrix} a_{13} \\ a_{23} \\ a_{14} \\ a_{24} \end{pmatrix}.$$

Donc, on connaît les distributions asymptotiques de $T^{(SL)}$, $T^{(E)}$ et $T^{(CN)}$ qui sont de la forme

$$\sum_{k=1}^4 b_k^{(\cdot)} \chi_1^2(d_k^{(\cdot)}),$$

où les coefficients $b_k^{(\cdot)}$ et les paramètres de décentralité $d_k^{(\cdot)}$ sont fonction des quantités θ_{34} , a_{13} , a_{14} , a_{23} , a_{24} et ont été obtenus précédemment.

Ainsi, pour des valeurs fixées de θ_{34} , a_{13} , a_{14} , a_{23} , a_{24} , on connaît explicitement les paramètres de la distribution et donc, en utilisant l'algorithme d'Imhof (1961) de la façon décrite à la section 4.3.2, on est en mesure de calculer les efficacités asymptotiques voulues.

Rappelons que dans un premier temps, on calcule les valeurs critiques $c_\alpha^{(\cdot)}$ telles que $P\left(T^{(\cdot)} > c_\alpha^{(\cdot)} | H_0\right) = \alpha$ pour des valeurs de α fixées (dans ce travail, on considère les valeurs $\alpha = 0,05; 0,10$). Ensuite, toujours en utilisant l'algorithme d'Imhof (1961) de manière itérative, on obtient les quantités $\delta^{(\cdot)2}(\alpha, \beta)$ pour des valeurs γ fixées (les valeurs $\gamma = 0,75; 0,80$ sont considérées ici). Autrement dit, on cherche $\delta^{(\cdot)2}(\alpha, \beta)$ tel que

$$P\left(\sum b_k^{(\cdot)} \chi_1^2\left(\delta_1^2(\alpha, \gamma) d_k^{(\cdot)}\right) > c_\alpha^{(\cdot)}\right) = \gamma.$$

Enfin, l'efficacité asymptotique relative d'un test par rapport à un autre est tout simplement le rapport des quantités $\delta^{(\cdot)2}(\alpha, \beta)$ correspondant à chacun de ces tests.

Supposons d'abord que X suit une loi multinormale $N_4(0, \Sigma^*)$ avec $\Sigma^* = ((\rho_{ij}^*))$ donnée par (4.9). On a vu à la section précédente, aux lemmes 4.1 et 4.2, qu'on a alors

TABLEAU 4.2. Valeurs des paramètres θ_{ij} et a_{ij} pour les cotes normales et linéaires lorsque X est de loi multinormale

cotes	θ_{ij}	a_{ij}
normales	ρ_{ij}	λ_{ij}
linéaires	$\frac{6}{\pi} \arcsin\left(\frac{\rho_{ij}}{2}\right)$	$\frac{3}{\pi} \lambda_{ij}$

En posant $\rho_{12} = \rho_{34} = 0,5$, $\lambda_{13} = 0,1$, $\lambda_{14} = 0,2$, $\lambda_{23} = 0,4$ et $\lambda_{24} = 0,6$, par le théorème 4.4, l'équation (4.17) et l'utilisation de l'algorithme d'Imhof (1961), on

obtient le tableau 4.3 où $e_{\Phi,L}^{(SL)}$, $e_{\Phi,L}^{(E)}$, et $e_{\Phi,L}^{(CN)}$ représentent respectivement l'efficacité asymptotique relative des cotes normales par rapport aux cotes linéaires pour les mesures de Stewart et Love (1968), Escoufier (1973) et Cramer et Nicewander (1979). Les quantités $e_{T_{\Phi}^{(u)}, T_{\Phi}^{(v)}}$ représentent l'efficacité asymptotique relative du test basé sur la mesure u par rapport à celui basé sur la mesure v en utilisant les cotes normales où u et v représentent les mesures de Stewart et Love (1968), Escoufier (1973) et Cramer et Nicewander (1979).

TABLEAU 4.3. *Efficacités asymptotiques relatives pour X de loi multinormale et $K = 2$*

α	γ	$e_{\Phi,L}^{(SL)}$	$e_{\Phi,L}^{(E)}$	$e_{\Phi,L}^{(CN)}$	$e_{T_{\Phi}^{(E)}, T_{\Phi}^{(SL)}}$	$e_{T_{\Phi}^{(CN)}, T_{\Phi}^{(E)}}$	$e_{T_{\Phi}^{(CN)}, T_{\Phi}^{(SL)}}$
0,05	0,75	1,038928	1,065714	1,047908	2,258830	5,027868	11,357097
	0,80	1,037098	1,066155	1,047923	2,323873	5,026810	11,681668
0,10	0,75	1,039856	1,068115	1,047909	2,212435	5,030527	11,129715
	0,80	1,038025	1,066025	1,047944	2,282093	5,030527	11,471291

On constate que les tests basés sur les cotes normales sont toujours asymptotiquement plus efficaces que les tests basés sur les cotes linéaires. Dans le cas particulier où on utilise la mesure de Cramer et Nicewander (1979), on connaît la véritable valeur de l'efficacité asymptotique relative du test basé sur les cotes normales par rapport à celui basé sur les cotes linéaires en posant $\lambda_{13} = 0,1$, $\lambda_{14} = 0,2$, $\lambda_{23} = 0,4$, $\lambda_{24} = 0,6$, $\theta_{12} = 0,5$ et $\theta_{34} = 0,5$ dans l'équation (4.16). On obtient $e_{\Phi,L}^{(CN)} = 1,04792$, ce qui nous donne une bonne idée de la précision de l'approximation obtenue par l'algorithme d'Imhof (1961).

Remarquons enfin que c'est le test basé sur la mesure de Cramer et Nicewander (1979) qui offre la meilleure performance.

Supposons maintenant que X suit une loi logistique multivariée. On considère toujours le cas où

$$X = \begin{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix} \end{pmatrix}.$$

On suppose que pour $i = 1, 2, 3, 4$, la variable aléatoire X_i suit une loi logistique telle que la fonction de répartition est donnée par $F(x) = (1 + e^{-x})^{-1}$. De plus, si (X_i, X_j) n'appartiennent pas au même sous-vecteur, alors la fonction de répartition conjointe de (X_i, X_j) est²

$$F_1(x_i, x_j) = F(x_i)F(x_j) [1 + \alpha_{ij} \{(1 - F(x_i))(1 - F(x_j))\}] \quad (4.25)$$

où $\alpha_{ij} = \frac{\lambda_{ij}}{\sqrt{n}}$. Finalement, si (X_i, X_j) appartiennent au même sous-vecteur, alors la fonction de répartition conjointe de (X_i, X_j) est³

$$F_2(x_i, x_j) = [1 + e^{-x_i} + e^{-x_j}]^{-1}. \quad (4.26)$$

Il nous suffit donc de calculer la valeur des paramètres θ_{12} , θ_{34} , a_{13} , a_{14} , a_{23} et a_{24} pour les cotes linéaires et les cotes normales pour être ensuite en mesure de calculer les efficacités asymptotiques relatives approximatives voulues.

Rappelons que les paramètres a_{ij} sont donnés par

$$a_{ij} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \omega_{ij}(x_i, x_j) F_{[i]}(x_i) F_{[j]}(x_j) J'_{(i)}(F_{[i]}(x_i)) J'_{(j)}(F_{[j]}(x_j)) dF_{[i]}(x_i) dF_{[j]}(x_j)$$

2. On retrouve cette forme de la distribution logistique bivariée dans Gumbel (1961). Le paramètre α_{ij} doit être compris dans l'intervalle $[-1, 1]$.

3. Cette dernière forme de la distribution est souvent appelée la première fonction de distribution de la logistique bivariée (voir Gumbel (1961) ou Johnson et Kotz (1972)).

où $\omega_{ij}(x_i, x_j) = \sqrt{n} \left(\frac{F_1(x_i, x_j)}{F(x_i)F(x_j)} - 1 \right)$. Puisque $\alpha_{ij} = \frac{\lambda_{ij}}{\sqrt{n}}$, on obtient

$$\omega_{ij}(x_i, x_j) = \lambda_{ij}(1 - F(x_i))(1 - F(x_j)).$$

Ces paramètres sont calculés pour les variables X_i et X_j qui n'appartiennent pas au même sous-vecteur. Rappelons aussi que les paramètres d'association θ_{ij} , donnés par

$$\theta_{ij}^{(2)} = 12 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (F_{[i]}(x_i) - \frac{1}{2})(F_{[j]}(x_j) - \frac{1}{2}) dF_{[i,j]}(x_i, x_j)$$

sont calculés pour les variables (X_1, X_2) et (X_3, X_4) , c'est-à-dire les variables qui appartiennent au même sous-vecteur.

Lorsqu'on considère les cotes linéaires, puisque $J'_{(i)}(u) = \sqrt{12}$, on a

$$\begin{aligned} a_{ij}^{(2)} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \lambda_{ij}(1 - F(x_i))(1 - F(x_j))F(x_i)F(x_j)\sqrt{12}\sqrt{12}dF(x_i)dF(x_j) \\ &= 12\lambda_{ij} \int_0^1 \int_0^1 (1-u)(1-v)u v du dv \\ &= \frac{\lambda_{ij}}{3}. \end{aligned}$$

Et puisque la fonction de densité conjointe de (X_i, X_j) est donnée par

$$f_2(x_i, x_j) = \frac{2e^{-x_i-x_j}}{(1 + e^{-x_i} + e^{-x_j})^3},$$

on obtient

$$\theta_{ij}^{(2)} = 0,478416.$$

En ce qui concerne les cotes normales, on doit recourir à des méthodes d'intégration numérique puisque les deux intégrales à évaluer sont de la forme

$$\theta_{ij}^{(1)} = 2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi^{(-1)} \left[(1 + e^{-x})^{-1} \right] \Phi^{(-1)} \left[(1 + e^{-y})^{-1} \right] \frac{e^{-x-y}}{(1 + e^{-x} + e^{-y})^3} dx dy$$

et

$$a_{ij}^{(1)} = \lambda_{ij} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[1 - (1 + e^{-x})^{-1}\right] \left[1 - (1 + e^{-y})^{-1}\right] (1 + e^{-x})^{-1} (1 + e^{-y})^{-1} \cdot \\ \Phi^{(-1)'} \left[(1 + e^{-x})^{-1}\right] \Phi^{(-1)'} \left[(1 + e^{-y})^{-1}\right] \frac{e^{-x}e^{-y}}{(1+e^{-x})^2(1+e^{-y})^2} dx dy.$$

La méthode d'intégration numérique utilisée est tirée de la librairie IMSL associée à FORTRAN (voir le manuel Math/Librairy IMSL, chapitre 4). L'approximation est obtenue par des applications successives de produits des formules de Gauss. Pour plus de détails, voir Atkinson (1978). On obtient les approximations suivantes

$$\theta_{ij}^{(1)} = 0,497027 \text{ et } a_{ij}^{(1)} = \lambda_{ij}(0,318285)$$

avec des erreurs d'approximation estimées respectives de $2,98 \times 10^{-8}$ et $8,94 \times 10^{-8}$.

On obtient donc, dans le cas de la loi logistique définie par (4.25) et (4.26),

TABLEAU 4.4. Valeurs des paramètres θ_{ij} et a_{ij} pour les cotes normales et linéaires lorsque X est de loi logistique

cotes	θ_{ij}	a_{ij}
normales	0,497027	0,318285 λ_{ij}
linéaires	0,478416	$\frac{\lambda_{ij}}{3}$

En posant $\lambda_{13} = 0,1$, $\lambda_{14} = 0,2$, $\lambda_{23} = 0,4$ et $\lambda_{24} = 0,6$ et en utilisant le théorème 4.4, l'équation (4.17), les résultats du tableau 4.4 et l'algorithme d'Imhof (1961), on calcule les efficacités asymptotiques relatives présentées dans le tableau 4.5 pour $\alpha = 0,05; 0,10$ et $\gamma = 0,75$ et $0,80$.

TABLEAU 4.5. *Efficacités asymptotiques relatives pour X de loi logistique et K = 2*

α	γ	$e_{\Phi,L}^{(SL)}$	$e_{\Phi,L}^{(E)}$	$e_{\Phi,L}^{(CN)}$	$e_{T_L^{(E)}, T_L^{(SL)}}$	$e_{T_L^{(CN)}, T_L^{(E)}}$	$e_{T_L^{(CN)}, T_L^{(SL)}}$
0,05	0,75	0,946731	0,972689	0,955405	2,188165	5,132965	11,231773
	0,80	0,946753	0,972846	0,955278	2,245636	5,133369	11,527677
0,10	0,75	0,947932	0,972380	0,955405	2,148428	5,131808	11,025316
	0,80	0,946504	0,972459	0,955411	2,209828	5,132030	11,340903

Comme c'est habituellement le cas lorsqu'on a la distribution logistique (voir Hájek et Šidák (1967), p.16), on constate que les tests basés sur les cotes linéaires sont asymptotiquement plus efficaces que ceux basés sur les cotes normales. Comme dans le cas de la loi multinormale, on remarque une fois de plus que le test basé sur la mesure de Cramer et Nicewander (1979) est nettement asymptotiquement plus efficace que ceux basés sur les mesures de Stewart et Love (1968) et d'Escoufier (1973).

Considérons maintenant le cas $K = 3$ et $p_1 = p_2 = p_3 = 2$. En utilisant les résultats (4.19), (4.20), (4.21), (4.22), (4.23) et (4.24) ainsi que le théorème 4.4, l'équation (4.17) et l'algorithme d'Imhof (1961), on obtient les tableaux 4.6 et 4.7 pour $\alpha = 0,05; 0,10$ et $\gamma = 0,75$ et $0,80$.

En se basant sur le critère de l'efficacité asymptotique relative, on constate donc, encore une fois, la supériorité des tests basés sur les cotes normales par rapport à ceux basés sur les cotes linéaires lorsqu'on a la loi normale et de façon opposée, la supériorité des tests basés sur les cotes linéaires par rapport à ceux basés sur les cotes normales lorsqu'on a la loi logistique.

Il est aussi particulièrement intéressant de remarquer que, pour une loi donnée, les mesures d'efficacité asymptotique relative entre les tests basés sur les

TABLEAU 4.6. *Efficacités asymptotiques relatives pour X de loi multinormale, K = 3*

α	γ	$e_{\Phi,L}^{(SL)}$	$e_{\Phi,L}^{(E)}$	$e_{\Phi,L}^{(CN)}$	$e_{T_{\Phi}^{(E)}, T_{\Phi}^{(SL)}}$	$e_{T_{\Phi}^{(CN)}, T_{\Phi}^{(E)}}$	$e_{T_{\Phi}^{(CN)}, T_{\Phi}^{(SL)}}$
0,05	0,75	1,023642	1,065569	1,047931	2,292709	3,186513	7,305747
	0,80	1,019681	1,065121	1,047913	2,343982	3,195507	7,490213
0,10	0,75	1,044834	1,067357	1,047916	2,289325	3,208555	7,345426
	0,80	1,046536	1,067791	1,047894	2,358583	3,216338	7,585999

TABLEAU 4.7. *Efficacités asymptotiques relatives pour X de loi logistique, K = 3*

α	γ	$e_{\Phi,L}^{(SL)}$	$e_{\Phi,L}^{(E)}$	$e_{\Phi,L}^{(CN)}$	$e_{T_L^{(E)}, T_L^{(SL)}}$	$e_{T_L^{(CN)}, T_L^{(E)}}$	$e_{T_L^{(CN)}, T_L^{(SL)}}$
0,05	0,75	0,950730	0,972646	0,955467	2,190108	3,253307	7,125094
	0,80	0,951023	0,971544	0,955419	2,233326	3,259686	7,279942
0,10	0,75	0,948273	0,968196	0,955406	2,175221	3,267848	7,108293
	0,80	0,946489	0,964114	0,955411	2,225445	3,269815	7,276795

cotes normales et ceux basés sur les cotes linéaires calculées pour les trois mesures de liaison considérées ne semblent pas varier quelques soient les valeurs de α , γ ou que $K = 2$ ou $K = 3$. De plus, lorsqu'on compare la performance entre deux tests basés sur des mesures de liaison différentes, dans toutes les situations étudiées, ce sont les tests basés sur la mesure de Cramer et Nicewander (1979) qui offrent les meilleurs résultats.

Pour conclure ce chapitre, notons que dès qu'on est en mesure de calculer les paramètres θ_{ij} et a_{ij} , on peut obtenir toutes efficacités asymptotiques relatives entre deux tests en utilisant la méthode de comparaison et les résultats présentés dans cette section.

On aurait pu, entre autre, calculer les efficacités asymptotiques relatives pour d'autres valeurs de α , γ , λ_{ij} , ρ_{ij} , p_i ou encore K , mais étant donné que tous ces résultats sont asymptotiques, étudions plutôt la performance de ces tests en présence de petits échantillons. Dans le prochain chapitre, on examinera donc les niveaux et puissances expérimentaux de nos tests.

Chapitre 5

PUISSANCES ET NIVEAUX EXPÉRIMENTAUX DES TESTS

Pour clore ce travail, procédons à une comparaison du comportement expérimental des tests de la classe entre eux et aussi avec des compétiteurs paramétriques et non paramétriques. Pour ce faire, une étude de Monte Carlo est proposée à l'intérieur de laquelle on étudie les niveaux et puissances expérimentaux des tests en faisant varier les tailles échantillonnales, le nombre de sous-vecteurs de X , la taille des sous-vecteurs de X , la loi des observations ainsi que le degré de violation de l'hypothèse nulle sous H_{1n} .

Pour cette étude expérimentale, considérons d'abord des tests appartenant à la classe \mathcal{C} . Premièrement, deux tests basés sur la mesure de Stewart et Love (1968), définis par (3.1). Le premier, $T_{\Phi}^{(SL)}$, basé sur les cotes normales et le second, $T_L^{(SL)}$, basé sur les cotes linéaires. De façon analogue, on choisit deux tests basés sur la mesure d'Escoufier (1973), définis par (3.2); $T_{\Phi}^{(E)}$ et $T_L^{(E)}$, basés respectivement sur les cotes normales et linéaires. Enfin, on considère deux autres tests, basés cette fois sur la mesure de Cramer et Nicewander (1979), définis par (3.4), soit $T_{\Phi}^{(CN)}$ et $T_L^{(CN)}$, encore une fois, basés respectivement sur les cotes normales et linéaires.

À ces tests appartenant à la classe \mathcal{C} , nous comparons un compétiteur paramétrique, soit le test du rapport de vraisemblance maximale noté $T^{(A)}$ et défini

par (2.3), et deux compétiteurs non paramétriques dérivés de la classe de tests proposée par Puri *et al* (1970), définie par (2.4); le premier, $T_{\Phi}^{(PSG)}$, obtenu en utilisant les cotes normales et le second, $T_L^{(PSG)}$, obtenu en utilisant les cotes linéaires.

5.1. MÉTHODOLOGIE

Le comportement de ces neuf tests est donc étudié dans différents contextes en faisant varier les paramètres suivants :

- (i) les tailles échantillonales, $n = 50$ et 100 ;
- (ii) le nombre de sous-vecteurs K ainsi que la taille de chacun de ces sous-vecteurs :
 - pour $K = 2$, on considère $p_1 = p_2 = 2$ et $p_1 = 2, p_2 = 3$;
 - pour $K = 3$, on considère $p_1 = p_2 = p_3 = 2$ et $p_1 = p_2 = 2, p_3 = 3$;
- (iii) trois types de distributions sont considérées :
 - a) la loi multinormale $N_p(0, \Sigma)$ où $p = \sum_{k=1}^K p_k$ et où

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \cdots & \Sigma_{1K} \\ \vdots & \ddots & \vdots \\ \Sigma_{K1} & \cdots & \Sigma_{KK} \end{pmatrix}$$

avec

$$\begin{cases} \Sigma_{kk} = I_{p_k} & \text{pour } k = 1, \dots, K; \\ \Sigma_{kl} = C_{00}, C_{15}, C_{20}, C_{25} & \text{pour } 1 \leq k < l \leq K; \end{cases} \quad (5.1)$$

où C_{xy} représente la matrice $p_k \times p_l$ dont les éléments sont tous égaux à la valeur réelle $0,xy$;

- b) une loi multivariée générale centrée avec une structure de covariance Σ donnée par (5.1). Pour ce faire, on procède de la même façon que

l'ont fait Cléroux *et al* (1994) et Cléroux *et al* (1995), c'est-à-dire qu'on considère un vecteur Y dont les composantes sont indépendantes et respectivement de loi $N(0,1)$, $Exp(1) - 1$, $\Gamma(1,1) - 1$, $\sqrt{12}(Beta(1,1) - 0.5)$, $\frac{\sqrt{3}}{\pi}Logistique(0,1)$, $\sqrt{12}(\chi_1^2 - 1)$ et $(U(0,1) - 0.5)$ où

- $N(\mu, \sigma^2)$ est la loi normale de paramètres μ et σ^2 ;
- $Exp(\theta)$ est la loi exponentielle de paramètre θ ;
- $\Gamma(a, b)$ est la loi gamma de paramètres a et b ;
- $Beta(a, b)$ est la loi bêta de paramètres a et b ;
- $Logistique(a, b)$ est la loi logistique de paramètres a et b ;
- χ_ν^2 est la loi khi-deux centrée avec ν degrés de liberté;
- et enfin, $U(a, b)$ est la loi uniforme sur l'intervalle $[a, b]$.

Ainsi, le vecteur Y est tel que $E(Y) = 0$ et $Var(Y) = I$. Donc, en considérant le vecteur X défini par la transformation $X = CY$, on a $E(X) = 0$ et $Var(X) = CC' = \Sigma$.

- c) Et enfin, une loi logistique multivariée, encore une fois, centrée avec une structure de covariance Σ donnée par (5.1). On procède de la même façon que pour la loi multivariée générale sauf que cette fois, toutes les composantes du vecteur initial Y sont indépendantes de loi $\frac{\sqrt{3}}{\pi}Logistique(0,1)$.

Dans tous les cas, on considère un niveau nominal de 5%. Les résultats expérimentaux sont obtenus de la façon suivante :

Étape 1 : on génère n vecteurs de loi F , X_1, \dots, X_n ;

Étape 2 : on obtient les vecteurs des rangs correspondant à cet échantillon, R_1, \dots, R_n ;

Étape 3: pour i et $j = 1, \dots, p$, on calcule les paramètres d'association échantillonnaux t_{ij} pour les cotes normales et linéaires ainsi que les covariances échantillonnales s_{ij} (pour le test paramétrique);

Étape 4: on calcule la valeur des neuf statistiques à comparer;

Étape 5: on calcule le point critique $\hat{c}_{0,05}$, le 95e quantile de la distribution de chacune des neuf statistiques en utilisant l'algorithme d'Imhof en remplaçant les paramètres inconnus de chaque distribution par les estimateurs convergents obtenus à partir des observations générées, c'est-à-dire, en remplaçant la matrice Θ , et ses valeurs et vecteurs propres correspondants par la matrice d'association échantillonnale T avec ses valeurs et vecteurs propres correspondants¹;

Étape 6: on rejette $H_0^{(K)}$ si la valeur de la statistique est supérieure au point critique obtenu à l'étape 5 pour chacune des neuf statistiques.

On répète les étapes 1 à 6, 1000 fois pour chacune des situations à l'étude pour finalement obtenir les niveaux et puissances expérimentaux présentés dans les tableaux 5.1 à 5.12. Notons que les résultats que l'on retrouve dans ces tableaux sont multipliés par 1000 par soucis de clarté dans la présentation.

Comme on le sait, chacun des résultats de la simulation peut être considéré comme la moyenne de 1000 observations provenant d'une loi binômiale avec $p = 0,05$. Ainsi, pour déterminer, par exemple, si deux valeurs \hat{p}_1 et \hat{p}_2 sont significativement différentes au niveau 5%, on rejettera l'hypothèse de l'égalité des

1. Notons que Roy et Cléroux (1993) ont étudié la précision des valeurs critiques obtenues par l'algorithme d'Imhof (1961) en remplaçant les paramètres inconnus de la distribution par des estimateurs convergents. Ils montrent que le niveau obtenu converge en probabilité vers le niveau nominal lorsque $n \rightarrow \infty$.

proportions si

$$\frac{|\hat{p}_1 - \hat{p}_2|}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{1000} + \frac{\hat{p}_2(1-\hat{p}_2)}{1000}}} \geq 1,96.$$

On sait aussi que le niveau observé sera significativement différent du niveau nominal (5%) s'il n'appartient pas à l'intervalle $[0,0365; 0,0635]$ (intervalle de confiance de niveau 95% pour une proportion). Avant de présenter les résultats, notons enfin que les puissances expérimentales doivent être considérées comme étant la valeur observée plus ou moins 2 erreurs standards maximales, c'est-à-dire $\pm 0,0316$.

5.2. RÉSULTATS EXPÉRIMENTAUX

TABLEAU 5.1. *Niveaux et puissances expérimentaux pour X de loi multinormale avec $K = 2$ et $n = 50$*

TESTS	$p_1 = p_2 = 2$				$p_1 = 2$ et $p_2 = 3$			
	C_{00}	C_{15}	C_{20}	C_{25}	C_{00}	C_{15}	C_{20}	C_{25}
$T^{(A)}$	051	352	625	880	051	481	797	980
$T_{\Phi}^{(PSG)}$	066	372	662	877	075	517	815	977
$T_L^{(PSG)}$	065	378	623	846	077	482	778	964
$T_{\Phi}^{(SL)}$	046	315	588	820	049	425	733	939
$T_L^{(SL)}$	046	309	558	793	044	388	684	915
$T_{\Phi}^{(ESC)}$	048	313	568	804	049	407	698	917
$T_L^{(ESC)}$	042	297	543	772	043	371	645	888
$T_{\Phi}^{(CN)}$	048	325	597	845	051	430	748	964
$T_L^{(CN)}$	049	322	566	810	046	390	711	942

TABLEAU 5.2. Niveaux et puissances expérimentaux pour X de loi multinormale avec $K = 3$ et $n = 50$

TESTS	$p_1 = p_2 = p_3 = 2$				$p_1 = p_2 = 2$ et $p_3 = 3$			
	C_{00}	C_{15}	C_{20}	C_{25}	C_{00}	C_{15}	C_{20}	C_{25}
$T^{(A)}$	057	570	868	978	045	672	933	996
$T_{\Phi}^{(PSG)}$	079	628	894	983	079	741	957	999
$T_L^{(PSG)}$	081	595	865	979	083	721	932	999
$T_{\Phi}^{(SL)}$	047	596	870	976	038	690	930	995
$T_L^{(SL)}$	054	562	834	974	031	663	915	992
$T_{\Phi}^{(ESC)}$	047	583	844	969	047	679	912	983
$T_L^{(ESC)}$	054	550	816	959	032	637	882	975
$T_{\Phi}^{(CN)}$	052	602	891	983	043	702	951	998
$T_L^{(CN)}$	055	591	858	977	036	666	925	998

5.3. INTERPRÉTATION DES RÉSULTATS

Pour faciliter la discussion des résultats obtenus, nous ferons évidemment référence aux résultats des tableaux 5.1 à 5.12, mais aussi aux graphiques 5.1 à 5.12. À chaque combinaison de distribution et alternative considérée correspond un graphique. Il y a donc un graphique pour chacun des 12 cas suivants : pour la loi multinormale avec les structures de covariance C_{00} , C_{15} , C_{20} et C_{25} , pour la loi logistique multivariée avec les structures de covariance C_{00} , C_{15} , C_{20} et C_{25} et enfin, de la même façon, pour la loi multivariée générale avec les structures de covariance C_{00} , C_{15} , C_{20} et C_{25} .

Dans chaque graphique, on retrouve sur l'axe des y les résultats expérimentaux (niveaux ou puissances multipliés par 1000) et sur l'axe des x , les 9 statistiques étudiées où $T1$ fait référence au test $T^{(A)}$, $T2$ au test $T_{\Phi}^{(PSG)}$, $T3$ au test

TABLEAU 5.3. Niveaux et puissances expérimentaux pour X de loi multinormale avec $K = 2$ et $n = 100$

TESTS	$p_1 = p_2 = 2$				$p_1 = 2$ et $p_2 = 3$			
	C_{00}	C_{15}	C_{20}	C_{25}	C_{00}	C_{15}	C_{20}	C_{25}
$T^{(A)}$	055	703	935	996	055	840	993	1000
$T_{\Phi}^{(PSG)}$	067	702	938	992	060	856	995	1000
$T_L^{(PSG)}$	074	658	924	990	064	814	981	1000
$T_{\Phi}^{(SL)}$	060	675	918	989	049	811	983	1000
$T_L^{(SL)}$	060	628	909	984	048	782	971	1000
$T_{\Phi}^{(ESC)}$	056	667	911	987	047	803	972	999
$T_L^{(ESC)}$	059	627	897	981	043	759	963	999
$T_{\Phi}^{(CN)}$	060	684	928	991	049	823	987	1000
$T_L^{(CN)}$	063	635	910	988	045	787	978	1000

$T_L^{(PSG)}$, T_4 au test $T_{\Phi}^{(SL)}$, T_5 au test $T_L^{(SL)}$, T_6 au test $T_{\Phi}^{(E)}$, T_7 au test $T_L^{(E)}$, T_8 au test $T_{\Phi}^{(CN)}$, et enfin, T_9 au test $T_L^{(CN)}$.

Enfin, les 8 courbes de chaque graphique illustrent les résultats pour les situations suivantes :

- a: $K = 2$, $p_1 = p_2 = 2$ et $n = 50$;
- b: $K = 2$, $p_1 = 2$, $p_2 = 3$ et $n = 50$;
- c: $K = 3$, $p_1 = p_2 = p_3 = 2$ et $n = 50$;
- d: $K = 3$, $p_1 = p_2 = 2$, $p_3 = 3$ et $n = 50$;
- A: $K = 2$, $p_1 = p_2 = 2$ et $n = 100$;
- B: $K = 2$, $p_1 = 2$, $p_2 = 3$ et $n = 100$;
- C: $K = 3$, $p_1 = p_2 = p_3 = 2$ et $n = 100$;
- D: $K = 3$, $p_1 = p_2 = 2$, $p_3 = 3$ et $n = 100$.

TABLEAU 5.4. Niveaux et puissances expérimentaux pour X de loi multinormale avec $K = 3$ et $n = 100$

TESTS	$p_1 = p_2 = p_3 = 2$				$p_1 = p_2 = 2$ et $p_3 = 3$			
	C_{00}	C_{15}	C_{20}	C_{25}	C_{00}	C_{15}	C_{20}	C_{25}
$T^{(A)}$	041	929	998	1000	043	965	1000	1000
$T_{\Phi}^{(PSG)}$	059	939	999	1000	059	970	1000	1000
$T_L^{(PSG)}$	059	910	995	1000	062	960	1000	1000
$T_{\Phi}^{(SL)}$	042	935	998	1000	045	971	1000	1000
$T_L^{(SL)}$	044	908	995	999	051	956	1000	1000
$T_{\Phi}^{(ESC)}$	039	922	997	1000	043	958	1000	1000
$T_L^{(ESC)}$	042	902	994	999	046	946	1000	1000
$T_{\Phi}^{(CN)}$	042	940	999	1000	043	974	1000	1000
$T_L^{(CN)}$	046	917	996	1000	048	962	1000	1000

TABLEAU 5.5. Niveaux et puissances expérimentaux pour X de loi logistique multi-variée avec $K = 2$ et $n = 50$

TESTS	$p_1 = p_2 = 2$				$p_1 = 2$ et $p_2 = 3$			
	C_{00}	C_{15}	C_{20}	C_{25}	C_{00}	C_{15}	C_{20}	C_{25}
$T^{(A)}$	056	379	632	859	044	498	809	976
$T_{\Phi}^{(PSG)}$	077	411	657	879	075	548	825	981
$T_L^{(PSG)}$	067	403	651	861	067	530	818	970
$T_{\Phi}^{(SL)}$	055	351	602	822	042	454	749	942
$T_L^{(SL)}$	048	347	581	805	041	438	729	934
$T_{\Phi}^{(ESC)}$	061	359	592	804	038	439	739	921
$T_L^{(ESC)}$	054	341	571	784	039	432	705	904
$T_{\Phi}^{(CN)}$	058	360	604	846	043	458	778	969
$T_L^{(CN)}$	049	348	594	818	044	435	755	954

TABLEAU 5.6. Niveaux et puissances expérimentaux pour X de loi logistique multivariée avec $K = 3$ et $n = 50$

TESTS	$p_1 = p_2 = p_3 = 2$				$p_1 = p_2 = 2$ et $p_3 = 3$			
	C_{00}	C_{15}	C_{20}	C_{25}	C_{00}	C_{15}	C_{20}	C_{25}
$T^{(A)}$	053	577	861	984	047	656	936	999
$T_{\Phi}^{(PSG)}$	089	651	909	994	089	756	964	1000
$T_L^{(PSG)}$	092	639	901	989	094	741	957	1000
$T_{\Phi}^{(SL)}$	050	623	893	988	045	708	937	999
$T_L^{(SL)}$	050	618	882	983	045	698	926	998
$T_{\Phi}^{(ESC)}$	047	610	867	980	040	688	908	989
$T_L^{(ESC)}$	047	603	858	976	039	670	906	988
$T_{\Phi}^{(CN)}$	056	639	902	992	048	713	961	999
$T_L^{(CN)}$	059	617	901	987	044	700	943	998

TABLEAU 5.7. Niveaux et puissances expérimentaux pour X de loi logistique multivariée avec $K = 2$ et $n = 100$

TESTS	$p_1 = p_2 = 2$				$p_1 = 2$ et $p_2 = 3$			
	C_{00}	C_{15}	C_{20}	C_{25}	C_{00}	C_{15}	C_{20}	C_{25}
$T^{(A)}$	048	691	936	998	056	827	993	1000
$T_{\Phi}^{(PSG)}$	058	725	949	998	066	854	993	1000
$T_L^{(PSG)}$	055	704	949	999	064	835	986	1000
$T_{\Phi}^{(SL)}$	050	697	925	996	053	817	983	1000
$T_L^{(SL)}$	042	670	927	996	056	797	978	1000
$T_{\Phi}^{(ESC)}$	049	692	919	994	054	797	976	1000
$T_L^{(ESC)}$	041	663	919	994	051	780	970	998
$T_{\Phi}^{(CN)}$	050	694	935	998	055	825	989	1000
$T_L^{(CN)}$	044	663	943	999	055	800	984	1000

TABLEAU 5.8. Niveaux et puissances expérimentaux pour X de loi logistique multivariée avec $K = 3$ et $n = 100$

TESTS	$p_1 = p_2 = p_3 = 2$				$p_1 = p_2 = 2$ et $p_3 = 3$			
	C_{00}	C_{15}	C_{20}	C_{25}	C_{00}	C_{15}	C_{20}	C_{25}
$T^{(A)}$	055	913	999	1000	053	967	1000	1000
$T_{\Phi}^{(PSG)}$	066	930	1000	1000	068	972	1000	1000
$T_L^{(PSG)}$	066	927	998	1000	069	971	1000	1000
$T_{\Phi}^{(SL)}$	054	931	1000	1000	047	964	1000	1000
$T_L^{(SL)}$	058	931	998	1000	046	964	1000	1000
$T_{\Phi}^{(ESC)}$	055	926	999	1000	043	955	1000	1000
$T_L^{(ESC)}$	050	923	998	1000	048	952	1000	1000
$T_{\Phi}^{(CN)}$	053	939	1000	1000	049	973	1000	1000
$T_L^{(CN)}$	053	935	998	1000	042	973	1000	1000

TABLEAU 5.9. Niveaux et puissances expérimentaux pour X de loi générale multivariée avec $K = 2$ et $n = 50$

TESTS	$p_1 = p_2 = 2$				$p_1 = 2$ et $p_2 = 3$			
	C_{00}	C_{15}	C_{20}	C_{25}	C_{00}	C_{15}	C_{20}	C_{25}
$T^{(A)}$	053	403	665	876	045	459	799	967
$T_{\Phi}^{(PSG)}$	074	654	882	987	060	731	947	998
$T_L^{(PSG)}$	068	534	804	951	062	650	903	989
$T_{\Phi}^{(SL)}$	052	594	841	953	040	653	898	989
$T_L^{(SL)}$	047	478	751	920	040	554	840	973
$T_{\Phi}^{(ESC)}$	051	581	822	949	041	619	873	971
$T_L^{(ESC)}$	046	466	727	901	042	526	802	949
$T_{\Phi}^{(CN)}$	049	601	852	978	040	668	920	996
$T_L^{(CN)}$	043	480	760	932	041	550	856	982

TABLEAU 5.10. Niveaux et puissances expérimentaux pour X de loi générale multivariée avec $K = 3$ et $n = 50$

TESTS	$p_1 = p_2 = p_3 = 2$				$p_1 = p_2 = 2$ et $p_3 = 3$			
	C_{00}	C_{15}	C_{20}	C_{25}	C_{00}	C_{15}	C_{20}	C_{25}
$T^{(A)}$	047	607	909	992	050	703	946	1000
$T_{\Phi}^{(PSG)}$	067	935	997	1000	088	959	997	1000
$T_L^{(PSG)}$	071	864	985	1000	096	902	993	1000
$T_{\Phi}^{(SL)}$	041	930	993	1000	043	943	996	1000
$T_L^{(SL)}$	038	858	981	1000	049	878	991	999
$T_{\Phi}^{(ESC)}$	041	913	992	1000	037	913	994	999
$T_L^{(ESC)}$	036	841	973	998	037	843	981	997
$T_{\Phi}^{(CN)}$	038	931	997	1000	042	945	997	1000
$T_L^{(CN)}$	035	856	983	1000	046	887	993	1000

TABLEAU 5.11. Niveaux et puissances expérimentaux pour X de loi générale multivariée avec $K = 2$ et $n = 100$

TESTS	$p_1 = p_2 = 2$				$p_1 = 2$ et $p_2 = 3$			
	C_{00}	C_{15}	C_{20}	C_{25}	C_{00}	C_{15}	C_{20}	C_{25}
$T^{(A)}$	048	693	955	997	052	838	994	1000
$T_{\Phi}^{(PSG)}$	055	949	997	999	063	974	1000	1000
$T_L^{(PSG)}$	062	856	984	999	066	938	998	1000
$T_{\Phi}^{(SL)}$	051	928	996	999	047	968	999	1000
$T_L^{(SL)}$	053	825	978	999	051	924	996	1000
$T_{\Phi}^{(ESC)}$	052	925	995	999	048	958	999	1000
$T_L^{(ESC)}$	052	814	973	998	048	900	993	1000
$T_{\Phi}^{(CN)}$	050	939	996	999	050	968	999	1000
$T_L^{(CN)}$	052	836	982	999	056	930	997	1000

TABLEAU 5.12. Niveaux et puissances expérimentaux pour X de loi générale multivariée avec $K = 3$ et $n = 100$

TESTS	$p_1 = p_2 = p_3 = 2$				$p_1 = p_2 = 2$ et $p_3 = 3$			
	C_{00}	C_{15}	C_{20}	C_{25}	C_{00}	C_{15}	C_{20}	C_{25}
$T^{(A)}$	055	960	1000	1000	044	978	1000	1000
$T_{\Phi}^{(PSG)}$	054	1000	1000	1000	071	1000	1000	1000
$T_L^{(PSG)}$	063	995	1000	1000	075	997	1000	1000
$T_{\Phi}^{(SL)}$	045	1000	1000	1000	051	1000	1000	1000
$T_L^{(SL)}$	050	996	1000	1000	052	993	1000	1000
$T_{\Phi}^{(ESC)}$	046	999	1000	1000	047	1000	1000	1000
$T_L^{(ESC)}$	048	992	1000	1000	048	994	1000	1000
$T_{\Phi}^{(CN)}$	046	1000	1000	1000	054	1000	1000	1000
$T_L^{(CN)}$	052	996	1000	1000	047	996	1000	1000

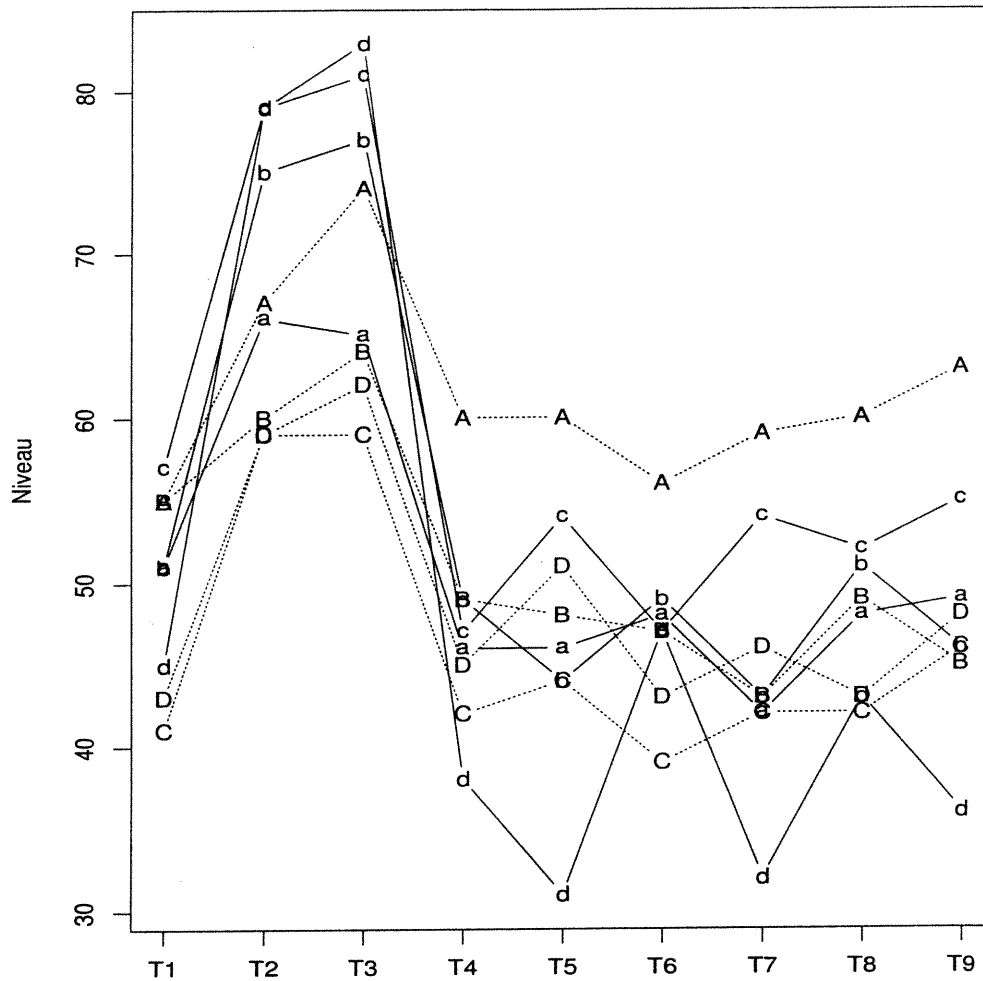


FIGURE 5.1. *Comparison des tests pour la loi multinormale avec C00 où T1 fait référence au test $T^{(A)}$, T2 au test $T_{\Phi}^{(PSG)}$, T3 au test $T_L^{(PSG)}$, T4 au test $T_{\Phi}^{(SL)}$, T5 au test $T_L^{(SL)}$, T6 au test $T_{\Phi}^{(E)}$, T7 au test $T_L^{(E)}$, T8 au test $T_{\Phi}^{(CN)}$, et enfin, T9 au test $T_L^{(CN)}$ et où 8 courbes illustrent les résultats correspondants aux cas a, b, c, d, A, B, C, D décrits précédemment.*

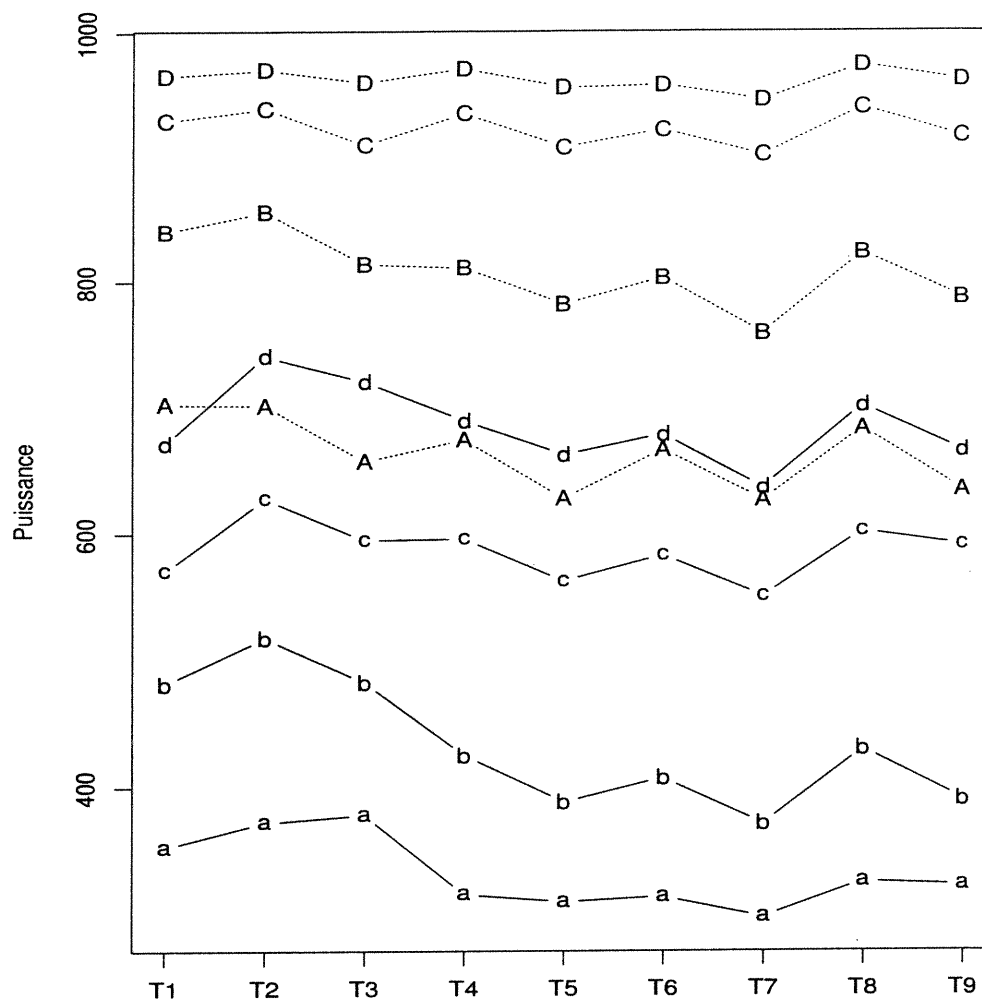


FIGURE 5.2. Comparaison des tests pour la loi multinormale avec C15 où T1 fait référence au test $T^{(A)}$, T2 au test $T_{\Phi}^{(PSG)}$, T3 au test $T_L^{(PSG)}$, T4 au test $T_{\Phi}^{(SL)}$, T5 au test $T_L^{(SL)}$, T6 au test $T_{\Phi}^{(E)}$, T7 au test $T_L^{(E)}$, T8 au test $T_{\Phi}^{(CN)}$, et enfin, T9 au test $T_L^{(CN)}$ et où 8 courbes illustrent les résultats correspondants aux cas a, b, c, d, A, B, C, D décrits précédemment.

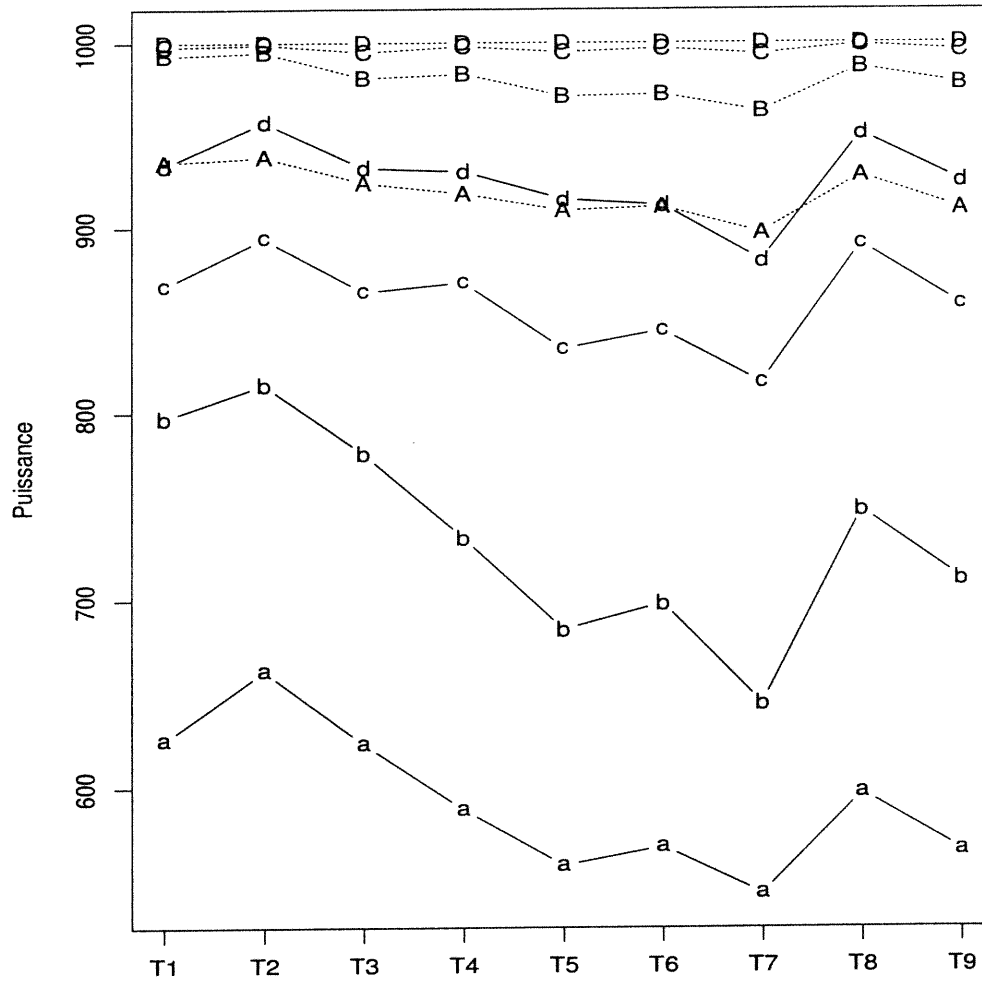


FIGURE 5.3. Comparaison des tests pour la loi multinormale avec C20 où T1 fait référence au test $T^{(A)}$, T2 au test $T_{\Phi}^{(PSG)}$, T3 au test $T_L^{(PSG)}$, T4 au test $T_{\Phi}^{(SL)}$, T5 au test $T_L^{(SL)}$, T6 au test $T_{\Phi}^{(E)}$, T7 au test $T_L^{(E)}$, T8 au test $T_{\Phi}^{(CN)}$, et enfin, T9 au test $T_L^{(CN)}$ et où 8 courbes illustrent les résultats correspondants aux cas a, b, c, d, A, B, C, D décrits précédemment.

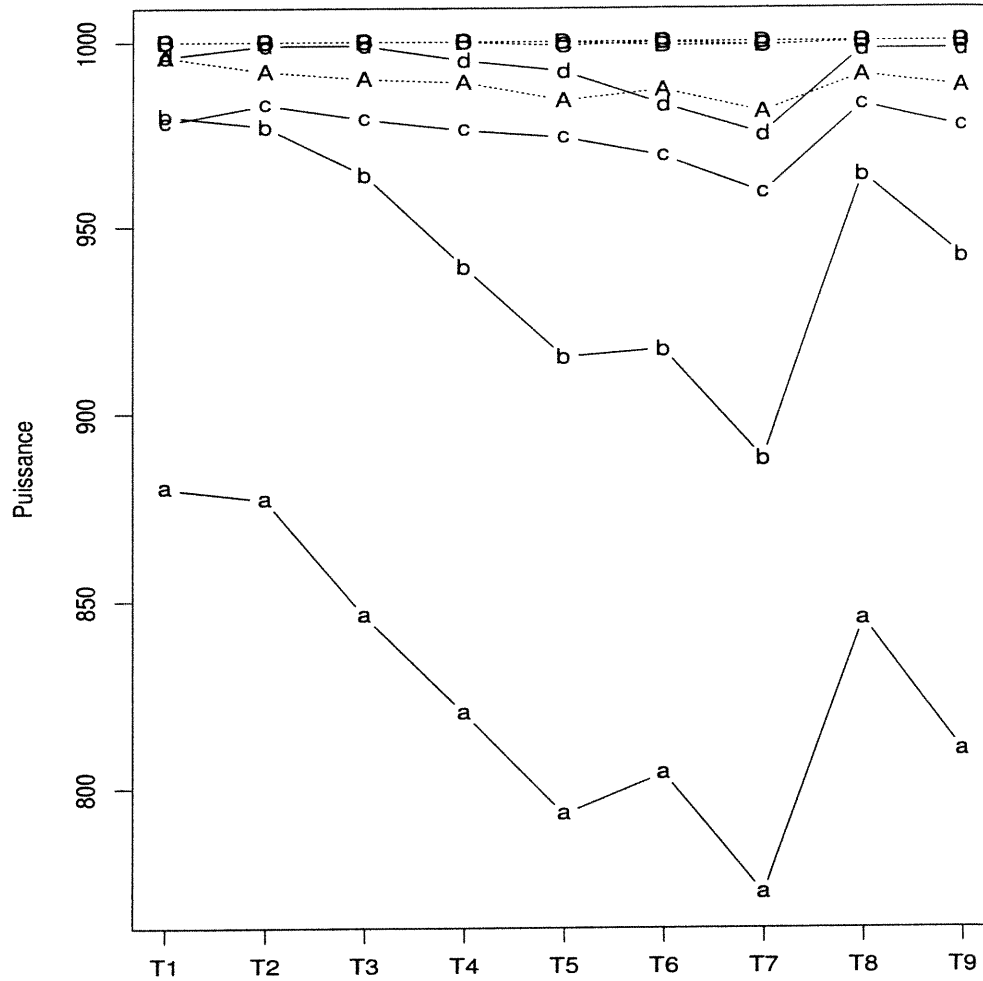


FIGURE 5.4. Comparaison des tests pour la loi multinormale avec C25 où T1 fait référence au test $T^{(A)}$, T2 au test $T_{\Phi}^{(PSG)}$, T3 au test $T_L^{(PSG)}$, T4 au test $T_{\Phi}^{(SL)}$, T5 au test $T_L^{(SL)}$, T6 au test $T_{\Phi}^{(E)}$, T7 au test $T_L^{(E)}$, T8 au test $T_{\Phi}^{(CN)}$, et enfin, T9 au test $T_L^{(CN)}$ et où 8 courbes illustrent les résultats correspondants aux cas a, b, c, d, A, B, C, D décrits précédemment.

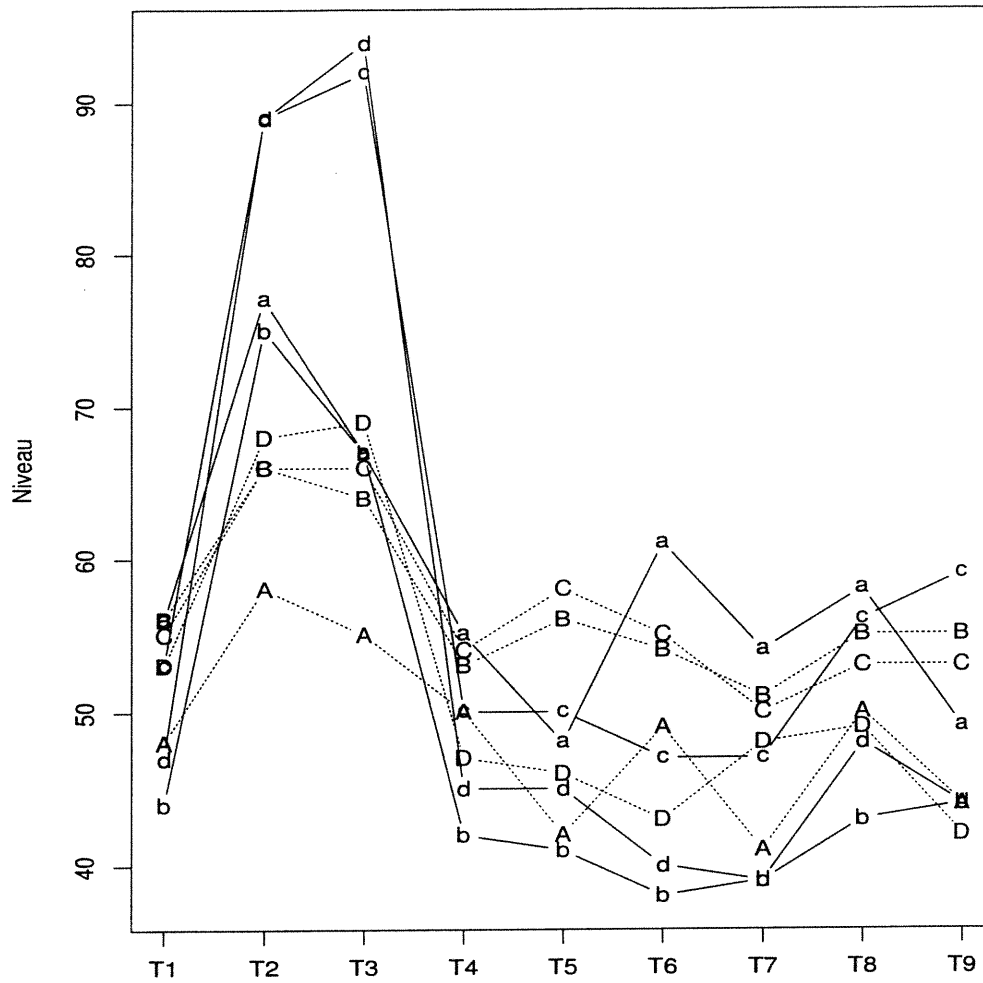


FIGURE 5.5. Comparaison des tests pour la loi logistique avec $C00$ où $T1$ fait référence au test $T^{(A)}$, $T2$ au test $T_{\Phi}^{(PSG)}$, $T3$ au test $T_L^{(PSG)}$, $T4$ au test $T_{\Phi}^{(SL)}$, $T5$ au test $T_L^{(SL)}$, $T6$ au test $T_{\Phi}^{(E)}$, $T7$ au test $T_L^{(E)}$, $T8$ au test $T_{\Phi}^{(CN)}$, et enfin, $T9$ au test $T_L^{(CN)}$ et où 8 courbes illustrent les résultats correspondants aux cas a, b, c, d, A, B, C, D décrits précédemment.

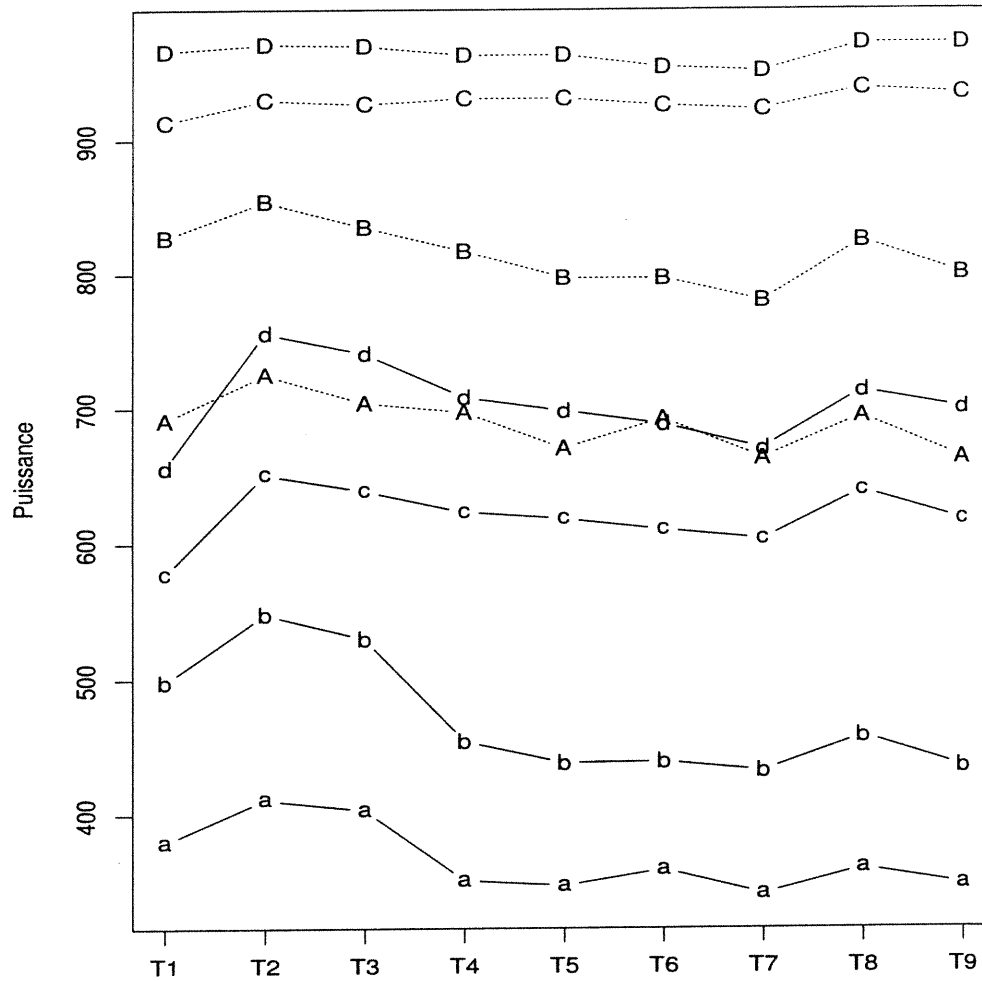


FIGURE 5.6. Comparaison des tests pour la loi logistique avec $C15$ où $T1$ fait référence au test $T^{(A)}$, $T2$ au test $T_{\Phi}^{(PSG)}$, $T3$ au test $T_L^{(PSG)}$, $T4$ au test $T_{\Phi}^{(SL)}$, $T5$ au test $T_L^{(SL)}$, $T6$ au test $T_{\Phi}^{(E)}$, $T7$ au test $T_L^{(E)}$, $T8$ au test $T_{\Phi}^{(CN)}$, et enfin, $T9$ au test $T_L^{(CN)}$ et où 8 courbes illustrent les résultats correspondants aux cas a, b, c, d, A, B, C, D décrits précédemment.

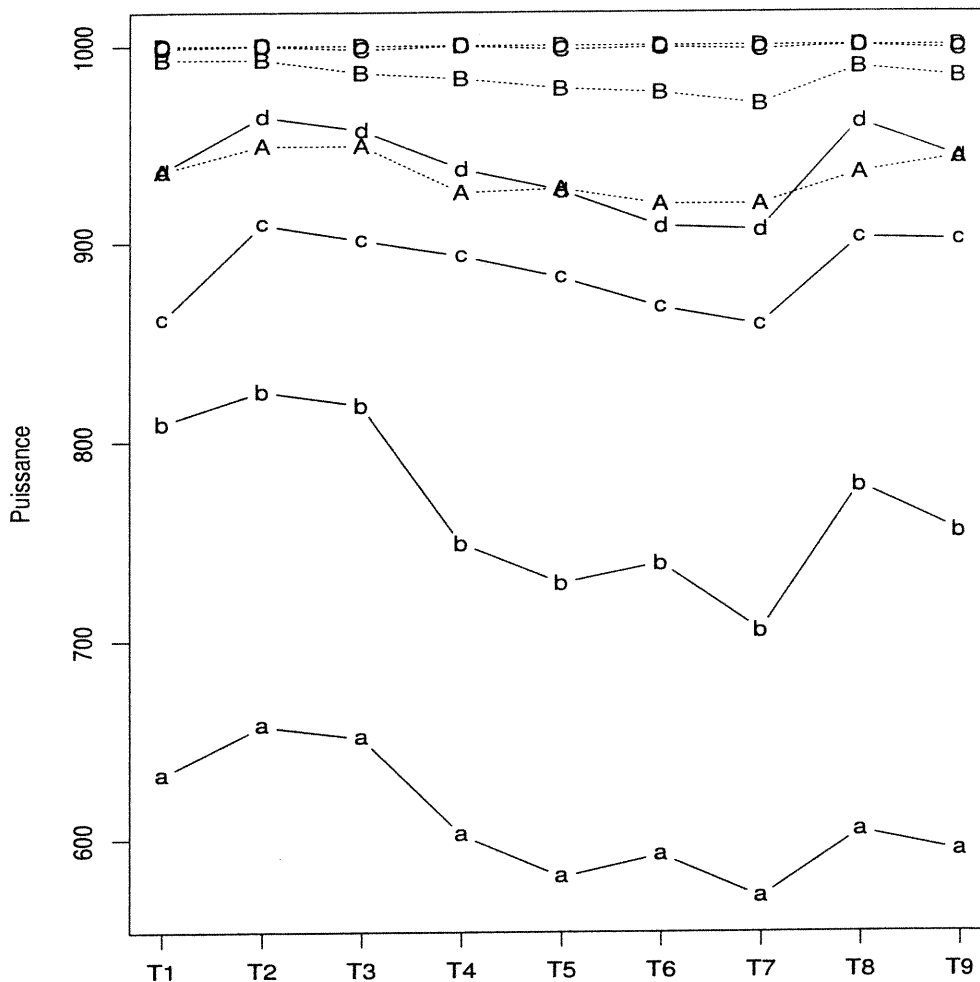


FIGURE 5.7. Comparaison des tests pour la loi logistique avec C20 où T1 fait référence au test $T^{(A)}$, T2 au test $T_{\Phi}^{(PSG)}$, T3 au test $T_L^{(PSG)}$, T4 au test $T_{\Phi}^{(SL)}$, T5 au test $T_L^{(SL)}$, T6 au test $T_{\Phi}^{(E)}$, T7 au test $T_L^{(E)}$, T8 au test $T_{\Phi}^{(CN)}$, et enfin, T9 au test $T_L^{(CN)}$ et où 8 courbes illustrent les résultats correspondants aux a, b, c, d, A, B, C, D décrits précédemment.

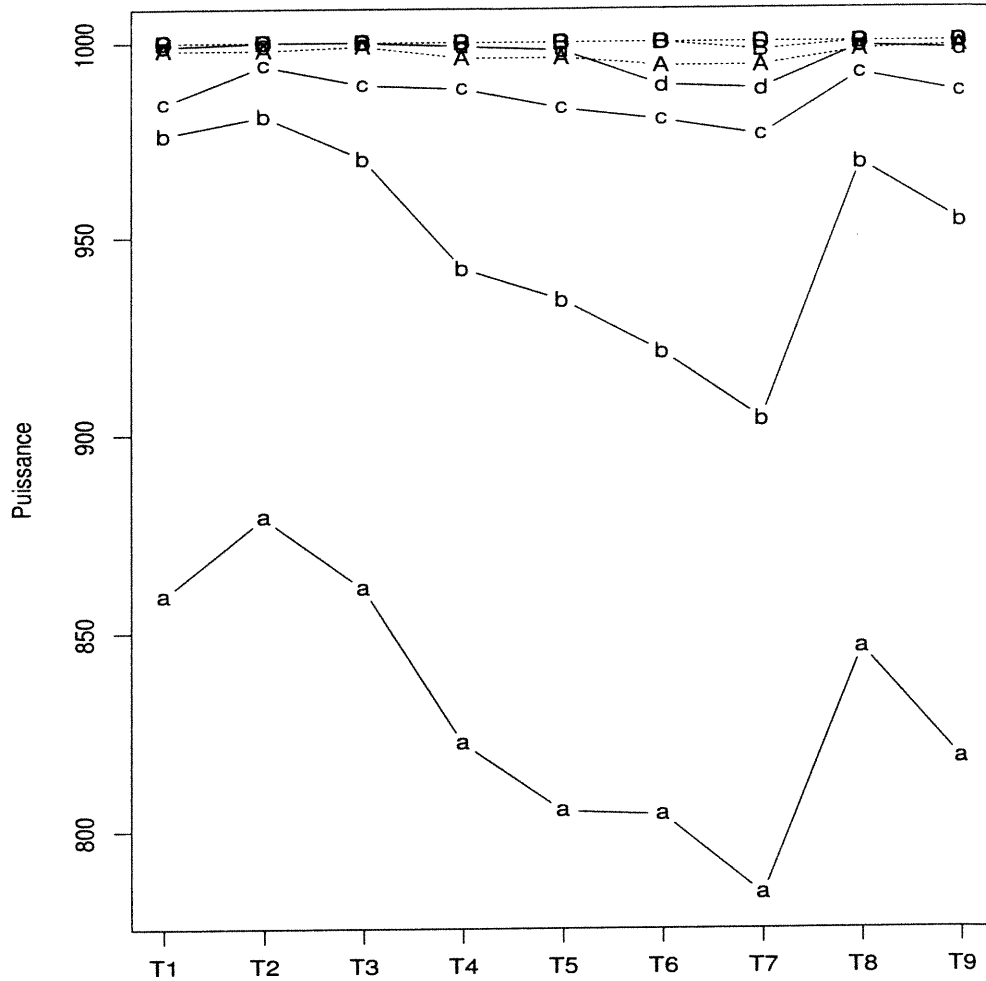


FIGURE 5.8. Comparaison des tests pour la loi logistique avec C25 où T1 fait référence au test $T^{(A)}$, T2 au test $T_{\Phi}^{(PSG)}$, T3 au test $T_L^{(PSG)}$, T4 au test $T_{\Phi}^{(SL)}$, T5 au test $T_L^{(SL)}$, T6 au test $T_{\Phi}^{(E)}$, T7 au test $T_L^{(E)}$, T8 au test $T_{\Phi}^{(CN)}$, et enfin, T9 au test $T_L^{(CN)}$ et où 8 courbes illustrent les résultats correspondants aux cas a, b, c, d, A, B, C, D décrits précédemment.

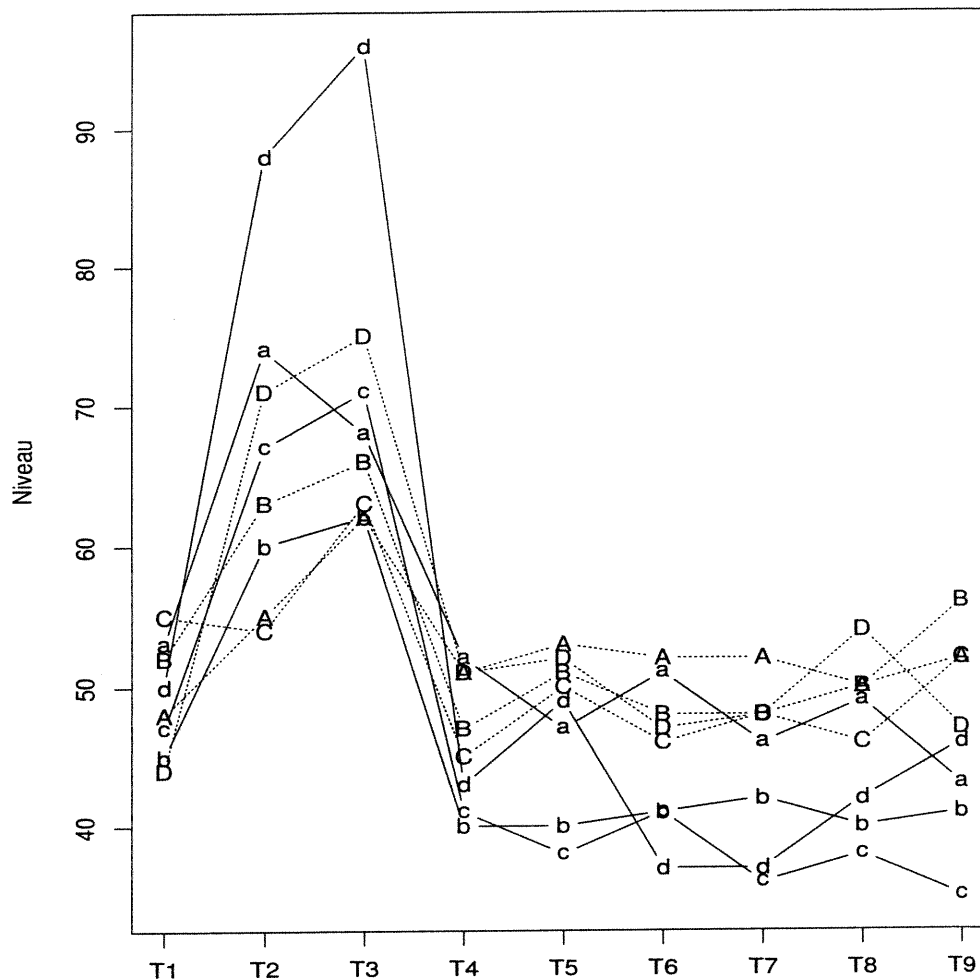


FIGURE 5.9. Comparaison des tests pour la loi générale avec $C00$ où $T1$ fait référence au test $T^{(A)}$, $T2$ au test $T_{\Phi}^{(PSG)}$, $T3$ au test $T_L^{(PSG)}$, $T4$ au test $T_{\Phi}^{(SL)}$, $T5$ au test $T_L^{(SL)}$, $T6$ au test $T_{\Phi}^{(E)}$, $T7$ au test $T_L^{(E)}$, $T8$ au test $T_{\Phi}^{(CN)}$, et enfin, $T9$ au test $T_L^{(CN)}$ et où 8 courbes illustrent les résultats correspondants aux cas a, b, c, d, A, B, C, D décrits précédemment.

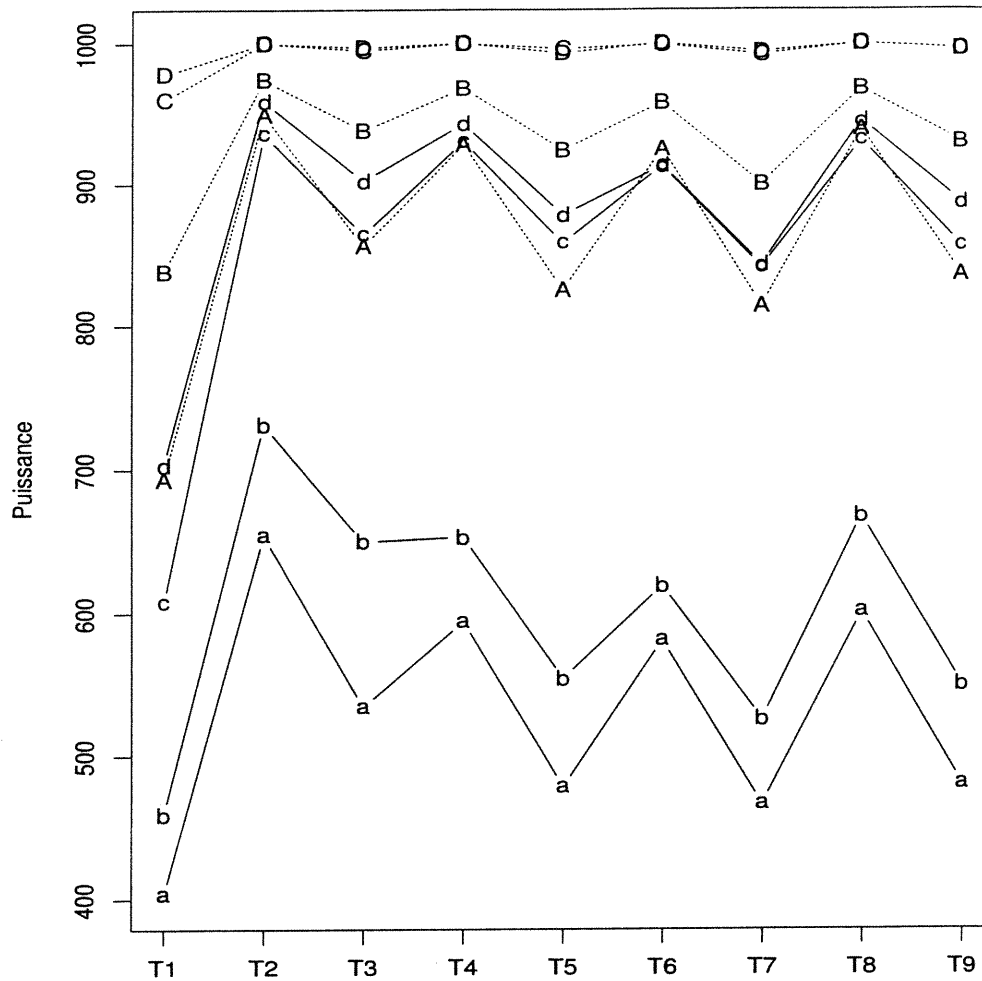


FIGURE 5.10. Comparaison des tests pour la loi générale avec C15 où T1 fait référence au test $T^{(A)}$, T2 au test $T_{\Phi}^{(PSG)}$, T3 au test $T_L^{(PSG)}$, T4 au test $T_{\Phi}^{(SL)}$, T5 au test $T_L^{(SL)}$, T6 au test $T_{\Phi}^{(E)}$, T7 au test $T_L^{(E)}$, T8 au test $T_{\Phi}^{(CN)}$, et enfin, T9 au test $T_L^{(CN)}$ et où 8 courbes illustrent les résultats correspondants aux cas a, b, c, d, A, B, C, D décrits précédemment.

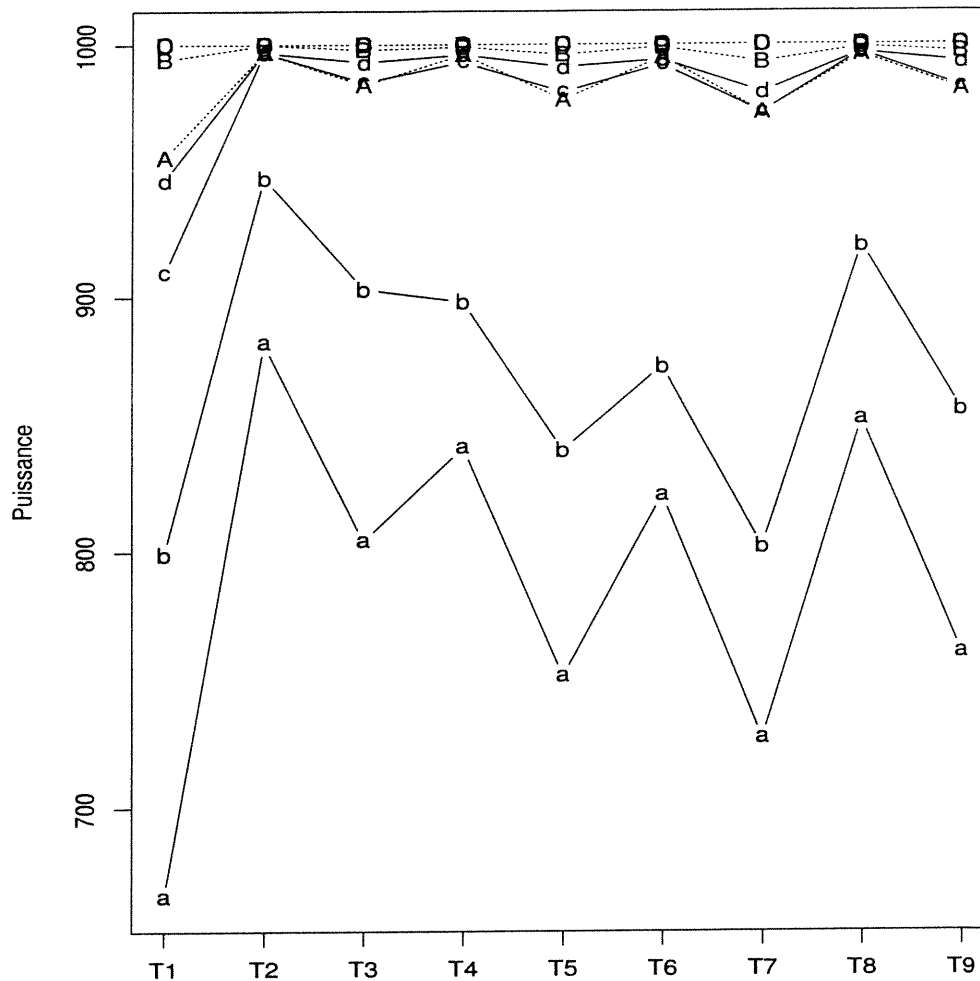


FIGURE 5.11. Comparaison des tests pour la loi générale avec $C20$ où $T1$ fait référence au test $T^{(A)}$, $T2$ au test $T_{\Phi}^{(PSG)}$, $T3$ au test $T_L^{(PSG)}$, $T4$ au test $T_{\Phi}^{(SL)}$, $T5$ au test $T_L^{(SL)}$, $T6$ au test $T_{\Phi}^{(E)}$, $T7$ au test $T_L^{(E)}$, $T8$ au test $T_{\Phi}^{(CN)}$, et enfin, $T9$ au test $T_L^{(CN)}$ et où 8 courbes illustrent les résultats correspondants aux cas a, b, c, d, A, B, C, D décrits précédemment.

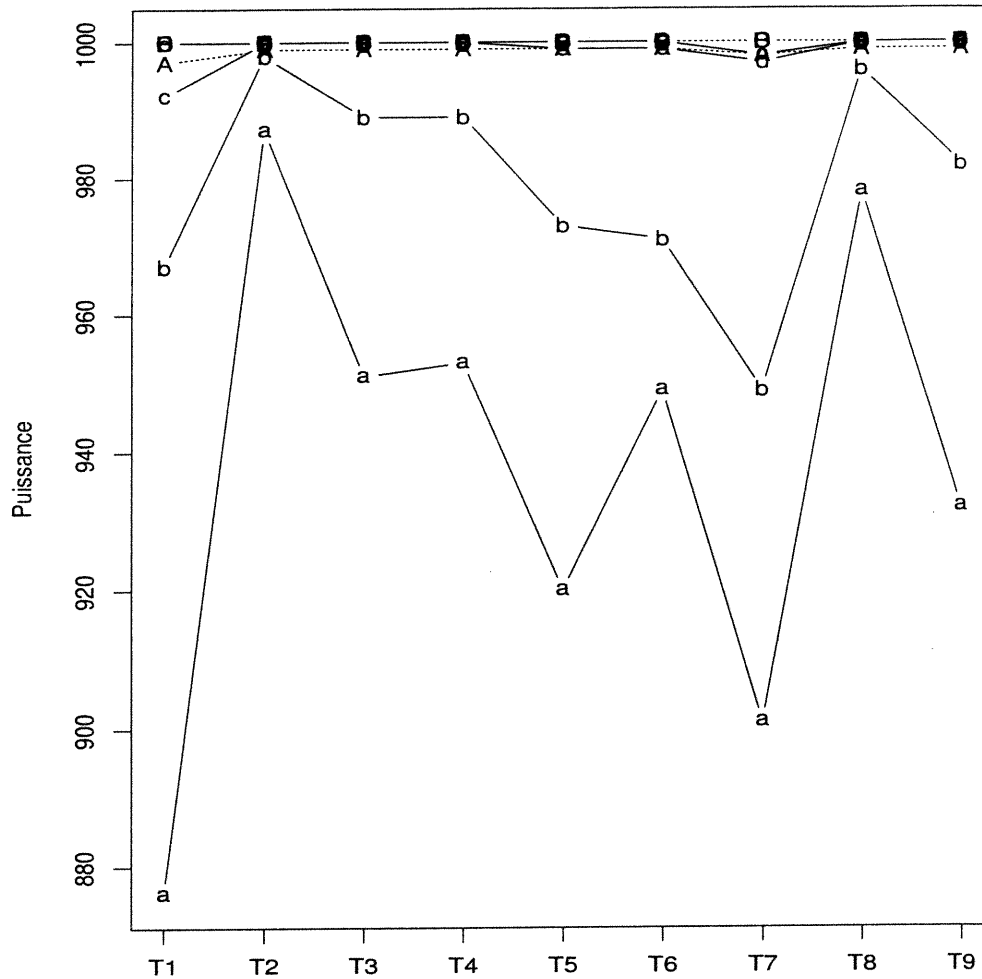


FIGURE 5.12. Comparaison des tests pour la loi générale avec C25 où T1 fait référence au test $T^{(A)}$, T2 au test $T_{\Phi}^{(PSG)}$, T3 au test $T_L^{(PSG)}$, T4 au test $T_{\Phi}^{(SL)}$, T5 au test $T_L^{(SL)}$, T6 au test $T_{\Phi}^{(E)}$, T7 au test $T_L^{(E)}$, T8 au test $T_{\Phi}^{(CN)}$, et enfin, T9 au test $T_L^{(CN)}$ et où 8 courbes illustrent les résultats correspondants aux cas a, b, c, d, A, B, C, D décrits précédemment.

Examinons d'abord les niveaux expérimentaux obtenus. Les graphiques 5.1, 5.5 et 5.9 illustrent de façon très claire que les tests de Puri *et al* (1970) tant

celui basé sur les cotes normales que celui basé sur les cotes linéaires, (T_2 et T_3) surestiment de façon marquée le niveau nominal (0,05). On constate d'ailleurs que ces résultats sont significativement supérieurs au niveau nominal de 0,05 puisqu'ils n'appartiennent pas à l'intervalle $[0,0365; 0,0636]$. Pour étudier si les niveaux expérimentaux de chacun des neuf tests diffèrent de façon significative, procédons à une analyse de variance sur les résultats expérimentaux obtenus. Notons d'abord qu'il est impossible de considérer le modèle incluant les 4 facteurs; taille échantillonnale, structure des sous-vecteurs, le test considéré et la distribution des observations (modèle avec une seule observation par cellule) puisque l'interaction entre les facteurs est significative.

TABLEAU 5.13. *Tableau d'analyse de variance à un facteur (n=216)*

Source	DL	SOMME CARRÉS	CARRÉS MOYENS	VALEUR F	Pr > F
Modèle	73	26460,14815	362,46778	13,94	<,0001
Erreur	142	3693,18519	26,00835		
Total	215	30153,33333			

Source	DL	SOMME CARRÉS	CARRÉS MOYENS	VALEUR F	Pr > F
LOI	2	328,11111	164,05556	6,31	0,0024
TEST	8	19046,08333	2380,76042	91,54	<,0001
TAILLE	1	0,00000	0,00000	0,00	1,0000
SVECT	3	262,85185	87,61728	3,37	0,0204
LOI*TEST	16	156,63889	9,78993	0,38	0,9860
LOI*TAILLE	2	250,11111	125,05556	4,81	0,0095
LOI*SVECT	6	1133,37037	188,89506	7,26	<,0001
TEST*TAILLE	8	2834,58333	354,32292	13,62	<,0001
TEST*SVECT	24	1980,73148	82,53048	3,17	<,0001
TAILLE*SVECT	3	467,66667	155,88889	5,99	0,0007

On doit donc séparer l'analyse en faisant une analyse de variance pour chacune des combinaisons de facteurs possibles. Sauf qu'en pratique, si on voulait faire toutes ces analyses, on se retrouverait à faire 24 analyses de variance à un facteur avec une seule observation par cellule. En tenant compte de ces contraintes, et suite à l'examen des résultats de la simulation, on choisit de faire une analyse de variance pour chacune des distributions considérée et pour les deux cas, $n = 50$ et $n = 100$. Ces analyses permettent de conclure à une différence significative entre les niveaux expérimentaux obtenus pour les différents tests de la simulation. Pour ne pas alourdir la présentation avec les détails de ces analyses de variance, on présente simplement les valeurs-p obtenues dans le tableau 5.3.

TABLEAU 5.14. Valeurs- p pour les analyses de différences significatives entre les niveaux obtenus pour les 9 différents tests

	Distributions		
	normale	logistique	générale
$n = 50$	0,0001	0,0001	0,0001
$n = 100$	0,0110	0,0001	0,0001

De plus, les comparaisons multiples (méthode de Tukey) permettent de constater que les niveaux expérimentaux obtenus pour les deux tests de Puri *et al* (1970) ($T2$ et $T3$) sont significativement différents de ceux obtenus pour les sept autres tests, qui eux, ne diffèrent pas entre eux de façon significative. Cette différence est particulièrement frappante lorsqu'on regarde le graphique 5.5, illustrant les niveaux expérimentaux pour la loi logistique, et plus spécifiquement les courbes c et d ($K = 3$ et $n = 50$). De plus, sachant que le niveau observé est significativement différent du niveau nominal (5%) s'il n'appartient pas à l'intervalle $[0,0365, 0,0635]$, on en déduit que les deux tests de Puri *et al* (1970) diffèrent de façon significative du niveau nominal (5%) alors que ce n'est pas le cas pour les sept autres tests.

Il sera donc important de tenir compte de cette surestimation du niveau des tests lorsque viendra le moment de comparer les puissances expérimentales obtenues pour les tests de Puri *et al* (1970) ($T2$ et $T3$) aux autres tests qui eux, donnent des résultats expérimentaux qui ne diffèrent pas de la valeur nominale (0,05). Notons aussi dans le cas de la loi multinormale que le test basé sur la mesure de Stewart et Love et sur les cotes linéaires ($T5$) et le test basé sur la mesure d'Escoufier (1973) et sur les cotes linéaires aussi ($T7$) sous-estiment de façon importante le niveau nominal dans le cas d , et seulement dans ce cas. Le

cas d représente le cas où $K = 3$, $p_1 = p_2 = 2$, $p_3 = 3$ et $n = 50$. Puisque les résultats ne diffèrent pas du niveau nominal pour la situation D , on peut croire que pour ces deux tests, le fait d'augmenter la taille échantillonnale ($n = 100$) corrige ce problème.

Examinons maintenant les puissances expérimentales obtenues. Il est intéressant de noter que la performance des tests au niveau de la puissance expérimentale s'améliore lorsque l'on passe de la structure de sous-vecteurs a à la structure b , à la structure c , etc... dans l'ordre, jusqu'à la structure D . Et ceci est vrai, quelque soit le test et la distribution considérés. Des analyses de variance pour étudier l'effet des facteurs "structure des sous-vecteurs" et "taille échantillonnale" permettent de vérifier que cette croissance de la puissance d'une structure à l'autre est significative quelque soit le test et la distribution considérés. En raison de l'interaction significative entre les facteurs, on doit encore une fois avoir recours à plusieurs analyses de variance. Afin de ne pas alourdir le texte avec de nombreuses tables d'analyses de variance, on présente les résultats pour un cas particulier dans le tableau suivant et mentionnons simplement que toutes les valeurs- p en rapport avec les résultats mentionnés sont de l'ordre de 0,0001.

TABLEAU 5.15. *Tableau d'analyse de variance à un facteur (structure des sous-vecteurs) avec $n=50$, $C15$ et la loi multinormale)*

Source	DL	SOMME CARRÉS	CARRÉS MOYENS	VALEUR F	Pr > F
Modèle	3	671357,6667	223785,8889	181,99	<,0001
Erreur	32	39348,2222	1229,6319		
Total	35	710705,8889			

Source	DL	SOMME CARRÉS	CARRÉS MOYENS	VALEUR F	Pr > F
SVECT	3	671357,6667	223785,8889	181,99	<,0001

C'est donc dire que premièrement, comme on s'y attendait, les résultats sont supérieurs globalement lorsqu'on considère une taille échantillonnale de 100 comparativement à 50. De plus, le fait de passer de $K = 2$ sous-vecteurs à $K = 3$ sous-vecteurs a un effet profitable, pour chacun des tests considérés au niveau de la puissance. De même, le fait de passer de sous-vecteurs de même taille à des sous-vecteurs de tailles différentes provoque une augmentation significative de la puissance expérimentale obtenue. Le fait d'augmenter le degré de violation de l'hypothèse nulle (c'est-à-dire passer de $C15$ à $C20$ à $C25$), conserve cet état de fait, tout en améliorant les puissances expérimentales obtenues (c'est-à-dire que les valeurs se rapprochent plus et plus rapidement de la valeur 1). Donc, de façon générale, quelque soit la distribution, le test ou le degré de violation de l'hypothèse nulle, les meilleurs résultats au niveau de la puissance expérimentale sont obtenus pour le cas D , c'est-à-dire, pour $K = 3$, $p_1 = p_2 = 2$, $p_3 = 3$ et $n = 100$.

Mais ceci étant vrai pour tous les tests, ça ne nous informe pas sur la supériorité d'un test en particulier. Essayons de voir maintenant si certains tests se démarquent par rapport aux autres.

Au seul examen des graphiques, on voit une certaine supériorité des tests $T2$ - $T3$ (Puri *et al* (1970)) cotes normales et linéaires) et du test $T8$ (Cramer et Nicewander (1973) cotes normales), et peut-être aussi du test $T1$ (test du rapport de vraisemblance maximale) dans le cas où les observations proviennent d'une loi normale. Rappelons d'abord le commentaire fait lors de la discussion concernant les niveaux expérimentaux. On doit être prudent lorsqu'on compare les puissances expérimentales obtenues avec les tests $T2$ et $T3$ puisqu'ils surestiment de façon significative les niveaux expérimentaux; il est donc normal que les puissances expérimentales soient plus élevées sans que ça n'implique que ces tests soient supérieurs. Il est même plus approprié de les exclure des comparaisons.

Si on essaie de dégager une tendance des résultats obtenus, au simple examen des tableaux de résultats et graphiques obtenus, on peut d'ores et déjà voir que le test basé sur la mesure de Cramer et Nicewander et sur les cotes normales ($T8$) semble offrir les résultats les plus intéressants si on regarde la situation globalement. En examinant bien les graphiques 5.4 à 5.12, on constate que la puissance obtenue pour ce test ($T8$) est toujours parmi les plus élevées quelque soit la distribution, la taille échantillonnale, structure de sous-vecteurs et degré de violation de l'hypothèse nulle. Ce qui ne semble pas être le cas pour le test $T1$ (TRVM) par exemple, qui bien que semblant toujours donner des résultats intéressants dans le cas de la distribution multinormale, c'est celui qui donne les moins bons résultats dans le cas de la loi multivariée générale (voir par exemple les graphiques 5.6, 5.9 et 5.12 dans le cas a particulièrement).

Voyons maintenant si ces tendances se révèlent significatives. Il aurait été intéressant de pouvoir procéder à une analyse de variance en considérant tous les facteurs présents, ou enfin avec le plus grand nombre possible de facteurs, et ainsi voir si effectivement certains tests se démarquent selon les différentes

combinaisons de facteurs à l'intérieur d'une seule analyse, mais ceci aurait été possible à condition qu'il n'y ait pas d'interaction entre les différents facteurs ou alors que ces interactions présentent une structure simple qui n'aurait pas gêné à l'établissement de certaines conclusions. Par exemple, si l'interaction entre le facteur test et structure des sous-vecteurs avait été la même quelque soit la loi, la taille échantillonnale et le degré de violation de l'hypothèse nulle. Mais bien évidemment, comme le laissait d'ailleurs entrevoir l'examen des graphiques 5.1 à 5.12, ce pas le cas et la nature des interactions présentes rendent cette analyse problématique. Soulignons aussi que, comme pour toute étude de niveaux et puissances expérimentaux, cela correspond à une situation d'une observation par cellule, ce qui contribue aussi à réduire l'étendue des analyses possibles; par exemple, si on veut comparer les modalités d'un facteur en fixant une valeur précise pour les autres facteurs. On devra alors par exemple se résoudre à procéder à des comparaisons multiples.

Compte tenu de ces contraintes, procédons à diverses comparaisons dans des situations spécifiques.

Pour chacune des 3 distributions considérées, les graphiques permettent clairement de voir que c'est dans les cas *a* et *b*, c'est-à-dire, le cas où $K = 2$, $p_1 = p_2 = 2$ et $n = 50$ et celui où $K = 2$, $p_1 = 2$, $p_2 = 3$ et $n = 50$ que la différence entre la performance des tests est la plus marquée. Ceci est particulièrement frappant lorsqu'on regarde le graphique 5.9, pour le cas de la loi générale avec C_{20} . Les analyses de variance ne permettent pas d'établir la supériorité d'un test en tenant compte de deux facteurs ou plus. Par exemple, pour le modèle qui tient compte des facteurs test et taille échantillonnale, l'effet test n'est pas significatif, quelque soit la loi et le degré de violation de H_0 même en ne considérant que les structures de sous-vecteurs *a* et *b*. Cependant, si on se restreint à la situation

très spécifique où on a la distribution multivariée générale et dans le cas *a*, le test *T8* donne un résultat significativement supérieur au test *T1* et ceci, pour les trois différents degrés de violation de H_0 . La valeur des statistiques pour les tests individuels définis en (??) donnent les résultats suivants qui sont significatifs en s'assurant d'avoir un niveau global de 5% : 9,0338 pour *C15*, 10,011 pour *C20* et 8,9412 pour *C25*.

Remarquons que dans le cas *b*, c'est-à-dire où les sous-vecteurs ne sont pas de même taille, les puissances expérimentales obtenues pour le test *T8* sont encore une fois significativement supérieures au test *T1* mais cette fois, surtout lorsque le degré de violation de H_0 est moindre. En effet, on obtient alors les valeurs suivantes pour les statistiques définies en (??) : 9,6395 pour *C15*, 7,9067 pour *C20* et 4,8404 pour *C25*.

Ainsi, une chose intéressante à souligner ici, est qu'alors qu'aucun test n'est significativement supérieur (même pas le test paramétrique, *T1*, dans le cas de la loi multinormale) au test basé sur la mesure de Cramer et Nicewander (1979) et les cotes normales (*T8*), on a vu qu'il y a des situations particulières où à l'inverse, celui-ci offre un rendement significativement plus élevé. Ce résultat est d'autant plus intéressant qu'il va dans le même sens que les conclusions tirées au chapitre 4 à l'effet que le test basé sur la mesure de Cramer et Nicewander (1979) donne les meilleurs résultats en terme d'efficacité asymptotique relative. Terminons cette section en ajoutant qu'en regroupant les tests basés sur les cotes normales et ceux basés sur les cotes linéaires, il n'est pas surprenant de constater que le groupe de tests basés sur les cotes normales performe mieux que celui basé sur les cotes linéaires lorsqu'on est dans le cas de la distribution multinormale. Cette différence n'est cependant significative que pour les cas *a* et *b*. Notons enfin que la différence entre les résultats obtenus par groupe de tests basés sur les cotes

linéaires comparativement au groupe basé sur les cotes normales, dans le cas de la distribution logistique, n'est pas suffisamment grande pour conclure de façon significative à sa supériorité.

CONCLUSION

Au terme de ce travail, on peut conclure à la réalisation de nos objectifs initiaux. Tout d'abord, la construction d'une vaste classe de tests pour l'hypothèse d'absence de liaison où chaque membre de la classe est une forme quadratique de la matrice d'association. Nous avons ensuite obtenu les distributions asymptotiques conjointes de ces statistiques de tests sous H_0 et sous une suite de contre-hypothèses convergeant vers H_0 . Puisque cette classe compte parmi ses membres plusieurs tests existants, ce travail a permis de généraliser plusieurs résultats que l'on retrouve dans la littérature.

Nous avons ensuite procédé à une étude de la performance des tests de la classe. Dans un premier temps, nous nous sommes intéressés au comportement asymptotique des tests en utilisant la notion d'efficacité asymptotique relative. Dans le cas particulier où les tests sont basés sur la mesure de Cramer et Nicewander (1979), plusieurs résultats analytiques ont pu être obtenus en raison de la forme particulière que prend la distribution des tests. Ces résultats ont montré la supériorité des cotes normales aux cotes linéaires dans le cas où les observations proviennent d'une loi normale, et inversement, la supériorité des cotes linéaires aux cotes normales dans le cas où les observations proviennent d'une loi logistique. De plus, lorsque les observations proviennent d'une loi normale, on a montré l'équivalence asymptotique de trois tests : le test paramétrique du

rapport de vraisemblance maximale, le test non paramétrique de Puri, Sen et Gokhale (1970) et le test de la classe basé sur la mesure de Cramer et Nicewander (1979) avec les cotes normales.

Suite à l'obtention de certains résultats analytiques pour des cas particuliers, des calculs numériques basés sur l'approche de Nyblom et Mäkeläinen (1983) ont révélé la supériorité des tests basés sur la mesure de Cramer et Nicewander (1979) parmi tous les tests de la classe, tant pour la loi normale que pour la loi logistique.

Dans un deuxième temps, on a poursuivi l'étude de la performance des tests en analysant les niveaux et puissances expérimentaux obtenus à la suite d'une étude de Monte carlo. Mentionnons que la supériorité des tests basés sur la mesure de Cramer et Nicewander (1979) constatée au chapitre 4, y est confirmée et ce, non seulement par rapport aux tests de la classe, mais aussi par rapport aux tests du rapport de vraisemblance maximale (TRVM) et à ceux de Puri, Sen et Gokhale (1970) dans l'étude expérimentale.

Plusieurs autres aspects du problème méritent éventuellement d'être approfondis. D'abord sur le plan théorique, mentionnons l'étude des propriétés de la mesure de Cramer et Nicewander (1979) susceptibles d'expliquer sa performance par rapport aux autres mesures étudiées et la recherche d'autres membres de la classe, dans une optique d'optimisation de la performance des tests. Sur le plan numérique, il serait souhaitable de comparer la performance des tests de la classe avec d'autres compétiteurs paramétriques ou non paramétriques de la littérature dont ceux de Dauxois et Kniet (1997, 1998).

Annexe A

GÉNÉRALISATION DU THÉORÈME 2.2

Dans le but d'obtenir la distribution asymptotique conjointe des statistiques de la classe sous une suite de contre-hypothèses qui convergent vers $H_0^{(K)}$, on doit obtenir la distribution asymptotique de $vec(T)$, le vecteur formé éléments de la matrice d'association échantillonnale.

Le théorème 2.2 du chapitre 2, tiré de l'article de Puri, Sen et Gokhale (1970) nous donne la distribution asymptotique de $vec(T)$ lorsque les observations de l'échantillon sont *i.i.d.* de loi F . Lorsque vient le temps d'obtenir la distribution si les observations sont indépendantes, mais de lois différentes, c'est-à-dire,

$$X_\alpha \text{ indépendantes de loi } F_\alpha \text{ pour } \alpha = 1, \dots, n, \quad (\text{A.1})$$

on doit adapter la démonstration de ce théorème à ce cas plus général.

Plusieurs travaux ont d'ailleurs nécessité cette généralisation du théorème 2.2. Mentionnons entre autres les travaux de Cléroux, Lazraq et Lepage (1994, 1995) et bien sûr, ceux de Puri, Sen et Gokhale (1970). On retrouve cependant très peu de détails à l'intérieur de ces articles sur le passage du cas *i.i.d.* au cas A.1. Puisque ce passage n'est pas trivial pour qui n'est pas familier avec cette théorie, nous présentons dans cette annexe les principales étapes du raisonnement permettant la généralisation du théorème 2.1.

On veut montrer que

$$\sqrt{n} (uvec(T) - uvec(\Theta)) \xrightarrow{\mathcal{L}} N(0, \Phi) \quad (\text{A.2})$$

où les quantités $uvec(T)$, $uvec(\Theta)$ et Φ sont définies au chapitre 2 et où les distributions conjointes et marginales sont données par

$$\begin{aligned} F_{[i,j]}(x,y) = F_{[i,j]}^{(n)}(x,y) &= F_{[i]}^{(0)}(x) F_{[j]}^{(0)}(y) \left(1 + \frac{\omega_{ij}(x,y)}{\sqrt{n}} \right) \\ F_{[i]}(x) = F_{[i]}^{(n)}(x) &= F_{[i]}^{(0)}(x) + \frac{F_{[i]}^{(0)}(x)}{\sqrt{n}} \lim_{y \rightarrow \infty} \omega_{ij}(x,y) \\ F_{[j]}(y) = F_{[j]}^{(n)}(y) &= F_{[j]}^{(0)}(y) + \frac{F_{[j]}^{(0)}(y)}{\sqrt{n}} \lim_{x \rightarrow \infty} \omega_{ij}(x,y) \end{aligned}$$

où les fonctions $F_{[i]}^{(0)}(x)$ et $F_{[j]}^{(0)}(y)$ sont les distributions marginales sous l'hypothèse d'indépendance H_0 et où $\omega_{ij}(x,y)$ est défini par (4.8). Soulignons enfin que les distributions empiriques (marginales et conjointes) sont notées $F_{[i]}^*(x)$, $F_{[j]}^*(y)$ et $F_{[i,j]}^*(x,y)$.

L'idée de la démonstration repose sur le fait d'exprimer la matrice d'association de la façon suivante :

$$T = \sum_{i=1}^3 A_{in} + \sum_{i=1}^8 B_{in} \quad (\text{A.3})$$

où

$$\begin{aligned} A_{1n} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J_{(i)}(F_{[i]}(x)) J_{(j)}(F_{[j]}(y)) dF_{[i,j]}^*(x,y) \\ A_{2n} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[F_{[i]}^*(x) - F_{[i]}(x) \right] J_{(j)}(F_{[j]}(y)) J'_{(i)}(F_{[i]}(x)) dF_{[i,j]}(x,y) \\ A_{3n} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[F_{[j]}^*(y) - F_{[j]}(y) \right] J_{(i)}(F_{[i]}(x)) J'_{(j)}(F_{[j]}(y)) dF_{[i,j]}(x,y) \end{aligned}$$

et on retrouve l'expression des termes B dans l'article de Bhuchongkul (1964).

Par la suite, on montre que $\sqrt{n} \sum_{i=1}^3 A_{in}$ est asymptotiquement de loi normale et que B_{in} est $O_p\left(\frac{1}{\sqrt{n}}\right)$. Par (A.3) on a le résultat.

Pour démontrer la normalité asymptotique de $\sqrt{n}\Sigma_{i=1}^3 A_{in}$, Buchongkul (1964) fait remarquer que cette quantité peut s'exprimer comme la moyenne de n variables aléatoires indépendantes, c'est-à-dire

$$\Sigma_{i=1}^3 A_{in} = \frac{1}{n} \Sigma_{r=1}^n \left(U_{(i,j);1}^{(r)} + U_{(i,j);2}^{(r)} + U_{(i,j);3}^{(r)} \right)$$

où

$$\begin{aligned} U_{(i,j);1}^{(r)} &= J_{(i)}(F_{[i]}(X_{ir})) J_{(j)}(F_{[j]}(X_{jr})) \\ U_{(i,j);2}^{(r)} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [I(X_{ir} \leq x_i) - F_{[i]}(x_i)] J'_{(i)}(F_{[i]}(x_i)) J_{(j)}(F_{[j]}(x_j)) \\ &\quad \cdot dF_{[i,j]}(x_i, x_j) \\ U_{(i,j);3}^{(r)} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [I(X_{jr} \leq x_j) - F_{[j]}(x_j)] J_{(i)}(F_{[i]}(x_i)) J'_{(j)}(F_{[j]}(x_j)) \\ &\quad \cdot dF_{[i,j]}(x_i, x_j). \end{aligned}$$

On montre facilement que

$$E_{F_{[i,j]}} \left(U_{(i,j);1}^{(r)} + U_{(i,j);2}^{(r)} + U_{(i,j);3}^{(r)} \right) = \theta_{ij}^{(r)}.$$

On a donc une somme de n variables aléatoires indépendantes de moyenne $\theta_{ij}^{(r)}$. En utilisant le théorème de Bessy-Esséen (voir par exemple Serfling (1980) p.33), il suffit de montrer que le troisième moment de (A.3) est borné pour obtenir la normalité asymptotique.

Afin d'illustrer les techniques qui permettent de démontrer le résultat, examinons un des termes, $|U_{(i,j);1}^{(r)}|^3$ et montrons qu'il est borné. Parce qu'on procède de la même façon en utilisant les mêmes techniques pour tous les autres termes.

Une des astuces consiste à utiliser l'inégalité élémentaire suivante :

$$|ab| \leq \frac{|a|^p}{p} + \frac{|b|^q}{q}, \quad (\text{A.4})$$

pour $p > 1$ et $\frac{1}{p} + \frac{1}{q} = 1$ et la partie (iv) de la condition 1.5.

$$\begin{aligned} E \left| U_{(i,j);1}^{(r)} \right|^3 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| J_{(i)} \left(F_{[i]}^{(r)}(x_{ir}) \right) J_{(j)} \left(F_{[j]}^{(r)}(x_{jr}) \right) \right|^3 dF_{[i,j]}^{(r)}(x_i, x_j) \\ &\leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\left| J_{(i)} \left(F_{[i]}^{(r)}(x_{ir}) \right) \right|^{3(2)}}{2} + \frac{\left| J_{(j)} \left(F_{[j]}^{(r)}(x_{jr}) \right) \right|^{3(2)}}{2} dF_{[i,j]}^{(r)}(x_i, x_j) \end{aligned}$$

en utilisant (A.4) en posant $p = 2$ (et donc $q = 2$). On a donc

$$\begin{aligned} E \left| U_{(i,j);1}^{(r)} \right|^3 &\leq \int_{-\infty}^{\infty} \frac{\left| J_{(i)} \left(F_{[i]}^{(r)}(x_{ir}) \right) \right|^6}{2} dF_{[i]}^{(r)}(x_i) + \int_{-\infty}^{\infty} \frac{\left| J_{(j)} \left(F_{[j]}^{(r)}(x_{jr}) \right) \right|^6}{2} dF_{[j]}^{(r)}(x_j) \\ &= \int_0^1 \frac{\left| J_{(i)}(u) \right|^6}{2} du + \int_0^1 \frac{\left| J_{(j)}(u) \right|^6}{2} du \\ &\leq C \int_0^1 [u(1-u)]^{-6\alpha} du \\ &= C\beta(-6\alpha + 1, -6\alpha + 1) \end{aligned}$$

qui est borné pourvu que $1 - 6\alpha > 0$, donc si $\alpha < 1/6$, mais ceci est vrai puisque $0 < \alpha < 1/8$ (1.5).

Annexe B

CONTIGUITÉ SOUS L'HYPOTHÈSE DE MULTINORMALITÉ

Au chapitre 2, on a choisi la suite de contre-hypothèses

$$H_{1n} : \Theta_{kl} = \frac{A_{kl}}{\sqrt{n}} \text{ pour } 1 \leq k < l \leq K \quad (\text{B.1})$$

où la matrice A_{kl} est symétrique et est formée des éléments a_{ij} où

$$a_{ij} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \omega_{ij} F_{[i]}(x_i) F_{[j]}(x_j) J'_i(F_{[i]}(x_i)) J'_j(F_{[j]}(x_j)) dF_{[i]}(x_i) dF_{[j]}(x_j).$$

On a alors affirmé que dans le cas où X soit une loi multinormale, cette suite de contre-hypothèses est contigüe à l'hypothèse nulle H_0 . Nous allons maintenant le démontrer lorsqu'on considère les cotes linéaires ou les cotes normales.

En utilisant le rapport de vraisemblance

$$L_n(x_n) = \begin{cases} \frac{p_n^{(n)}(x_n)}{p_n^{(0)}(x_n)} & \text{si } p_n^{(0)}(x_n) > 0, \\ 1 & \text{si } p_n^{(0)}(x_n) = p_n^{(n)}(x_n) = 0, \\ \infty & \text{si } p_n^{(0)}(x_n) = 0 < p_n^{(n)}(x_n), \end{cases} \quad (\text{B.2})$$

on obtient le théorème suivant dont on peut retrouver la démonstration, entre autre, dans Hájek et Šidák (1967), p.204.

Théorème B.1. *Si L_n suit asymptotiquement une loi lognormale $(-\frac{1}{2}\sigma^2, \sigma^2)$ où $\sigma^2 > 0$ sous H_0 , alors les densités $p_n^{(n)}$ sont contigües à $p_n^{(0)}$.*

On utilisera donc le théorème B.1 pour démontrer que la suite de contre-hypothèses (B.1) est contigüe à l'hypothèse nulle $H_0^{(K)}$ dans le cas particulier où $K = 2$, le vecteur X suit une loi multinormale $N_p(\mu, \Sigma)$, $p = p_1 + p_2$ et où on a les cotes linéaires ou les cotes normales.

Les lemmes 4.1 et 4.2 du chapitre 4 donnent l'expression des paramètres d'association θ_{ij} par rapport aux coefficients de corrélation entre X_i et X_j , $\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i^2 \sigma_j^2}$, i et $j = 1, \dots, p$ ($i \neq j$), selon qu'on utilise les cotes normales ou linéaires lorsque la distribution des observation est multinormale. On y constate que

$$\theta_{ij} = \begin{cases} \rho_{ij} & \text{pour les cotes normales,} \\ \frac{6}{\pi} \arcsin\left(\frac{\rho_{ij}}{2}\right) & \text{pour les cotes linéaires.} \end{cases}$$

En utilisant ce résultat, l'hypothèse $H_0^{(2)} : \Theta_{12} = 0$ est équivalente à

$$H_0^{(2)} : \Sigma_{12} = 0 \quad (\text{B.3})$$

et la suite de contre-hypothèses $H_{1n} : \Theta_{12} = \frac{A_{12}}{\sqrt{n}}$ peut aussi s'exprimer en terme de la matrice de covariance de la façon suivante:

$$H_{1n} : \Sigma_{12} = \frac{A_{12}^*}{\sqrt{n}}. \quad (\text{B.4})$$

Le lien entre la matrice A_{12}^* et la matrice A_{12} dans (B.1) dépend de la fonction de cotes considérée.

Sous les cotes normales, $\theta_{ij} = \rho_{ij}$ pour i et $j = 1, \dots, p$ ($i \neq j$). Ainsi, puisque $\theta_{ij} = \frac{a_{ij}}{\sqrt{n}}$, on a $\sigma_{ij} = \frac{a_{ij} \sigma_i^2 \sigma_j^2}{\sqrt{n}}$.

Pour les cotes linéaires, on a $\theta_{ij} = \frac{6}{\pi} \arcsin\left(\frac{\rho_{ij}}{2}\right)$. En utilisant le développement en série de Taylor pour $\sin\left(\frac{a_{ij} \pi}{6\sqrt{n}}\right)$, on obtient $\sigma_{ij} = \frac{a_{ij} \pi \sigma_i^2 \sigma_j^2}{3\sqrt{n}} + O(n^{-3/2})$.

Vérifions donc que la suite de contre-hypothèses H_{1n} définie par (B.4) est contigüe à l'hypothèse $H_0^{(2)}$ définie par (B.3).

La fonction de densité de X est donnée par

$$f(x) = (2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (x - \mu)' \Sigma^{-1} (x - \mu) \right\}.$$

Puisque sans perte de généralité, on peut poser $\mu = 0$, les distributions asymptotiques de $x = (X_1, \dots, X_n)$ sous $H_0^{(2)}$ et H_{1n} sont données respectivement par

$$p_n^{(0)}(x) = \prod_{i=1}^n \left[(2\pi)^{-\frac{p}{2}} |\Sigma_0|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} x' \Sigma_0^{-1} x \right\} \right] \quad (\text{B.5})$$

$$p_n^{(n)}(x) = \prod_{i=1}^n \left[(2\pi)^{-\frac{p}{2}} |\Sigma_n|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} x' \Sigma_n^{-1} x \right\} \right],$$

où

$$\Sigma_0 = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix} \quad \text{et} \quad \Sigma_n = \begin{pmatrix} \Sigma_{11} & \frac{A_{12}}{\sqrt{n}} \\ \frac{A_{21}}{\sqrt{n}} & \Sigma_{22} \end{pmatrix}.$$

Théorème B.2. *Si X suit une loi multinormale $N_p(0, \Sigma)$ avec $K = 2$, $p = p_1 + p_2$, alors les densités $p_n^{(n)}$ sous-jacentes à H_{1n} sont contiguës à $p_n^{(0)}$, la densité sous $H_0^{(2)}$.*

Démonstration: En utilisant le théorème (B.1), il suffit de montrer que, sous $H_0^{(2)}$, on a, lorsque $n \rightarrow \infty$

$$\ln L_n \xrightarrow{\mathcal{L}} N\left(-\frac{1}{2}\sigma^2, \sigma^2\right)$$

où L_n est défini par (B.2). Or, ici, on a

$$\begin{aligned} L_n &= \frac{\prod_{i=1}^n \left[(2\pi)^{-\frac{p}{2}} |\Sigma_n|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} X_i' \Sigma_n^{-1} X_i \right\} \right]}{\prod_{i=1}^n \left[(2\pi)^{-\frac{p}{2}} |\Sigma_0|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} X_i' \Sigma_0^{-1} X_i \right\} \right]} \\ &= \frac{|\Sigma_0|^{\frac{n}{2}}}{|\Sigma_n|^{\frac{n}{2}}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n X_i' (\Sigma_n^{-1} - \Sigma_0^{-1}) X_i \right\}, \end{aligned}$$

en utilisant (B.5). Ainsi,

$$\ln L_n = \frac{n}{2} \ln |\Sigma_0| - \frac{n}{2} \ln |\Sigma_n| - \frac{1}{2} \sum_{i=1}^n X_i' (\Sigma_n^{-1} - \Sigma_0^{-1}) X_i.$$

On sait que

$$\Sigma_0^{-1} = \begin{pmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & \Sigma_{22}^{-1} \end{pmatrix} \text{ et } \Sigma_n^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},$$

où

$$\begin{aligned} B_{11} &= \Sigma_{11.2}^{-1}, \\ B_{22} &= \Sigma_{22.1}^{-1}, \\ B_{12} &= -\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22.1}^{-1}, \\ B_{21} &= -\Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11.2}^{-1}, \end{aligned}$$

avec

$$\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \text{ et } \Sigma_{22.1} = \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}.$$

Ainsi, on a

$$\Sigma_{11.2} = \Sigma_{11} - \frac{A_{12} \Sigma_{22}^{-1} A_{21}}{n} \text{ et } \Sigma_{22.1} = \Sigma_{22} - \frac{A_{21} \Sigma_{11}^{-1} A_{12}}{n}$$

d'où

$$\begin{aligned} B_{11} &= \left(\Sigma_{11} - \frac{1}{n} A_{12} \Sigma_{22}^{-1} A_{21} \right)^{-1} \\ &= \left(I - \frac{1}{n} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \right)^{-1} \Sigma_{11}^{-1} \end{aligned}$$

et

$$B_{22} = \left(I - \frac{1}{n} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} A_{12} \right)^{-1} \Sigma_{22}^{-1}.$$

Or, si $P = A + CBD$ on a alors

$$P^{-1} = A^{-1} - A^{-1}CB (B + BDA^{-1}CB)^{-1} BDA^{-1}$$

(voir Muirhead (1982), p.580). Pour obtenir l'expression de $(I-G)^{-1}$, on applique ce dernier résultat en posant $P = I - G$ avec $A = I$, $B = G$, $C = -I$ et $D = I$.

Ainsi, on obtient

$$\begin{aligned}
 P^{-1} &= I + G(G - G \cdot G)^{-1}G \\
 &= I + G \cdot G^{-1}(I - G)^{-1}G \\
 &= I + (I - G)^{-1}G
 \end{aligned}$$

et par récurrence, on déduit

$$\begin{aligned}
 P^{-1} &= I + [I + (I - G)^{-1}G]G \\
 &= I + G + (I - G)^{-1}G^2 \\
 &= I + G + G^2 + (I - G)^{-1}G^3 \\
 &= \dots
 \end{aligned}$$

Alors, on trouve

$$\begin{aligned}
 B_{11} &= \left[I - \frac{1}{n} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \right]^{-1} \Sigma_{11}^{-1} \\
 &= \left\{ I + \frac{1}{n} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} + \left[I - \frac{1}{n} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \right]^{-1} \frac{1}{n^2} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21})^2 \right\} \Sigma_{11}^{-1} \\
 &= \Sigma_{11}^{-1} + \frac{1}{n} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} + H_1 O\left(\frac{1}{n^2}\right),
 \end{aligned}$$

et de la même façon, on obtient

$$B_{22} = \Sigma_{22}^{-1} + \frac{1}{n} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} + H_2 O\left(\frac{1}{n^2}\right),$$

où

$$H_1 = \left[I - \frac{1}{n} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \right]^{-1} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21})^2 \Sigma_{11}^{-1}$$

et

$$H_2 = \left[I - \frac{1}{n} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} A_{12} \right]^{-1} (\Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} A_{12})^2 \Sigma_{22}^{-1}.$$

D'autre part, comme

$$F = \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix} = \Sigma_n^{-1} - \Sigma_0^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} - \begin{pmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & \Sigma_{22}^{-1} \end{pmatrix},$$

on trouve

$$F_{11} = \Sigma_{11}^{-1} + \frac{1}{n} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} + H_1 O\left(\frac{1}{n^2}\right) - \Sigma_{11}^{-1}$$

$$= \frac{1}{n} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} + H_1 O\left(\frac{1}{n^2}\right)$$

$$F_{22} = \frac{1}{n} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} + H_2 O\left(\frac{1}{n^2}\right)$$

$$F_{12} = -\frac{1}{\sqrt{n}} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} = -\frac{1}{\sqrt{n}} \Sigma_{11}^{-1} A_{12} (\Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12})^{-1}$$

$$= -\frac{1}{\sqrt{n}} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} \left(I - \frac{1}{n} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} A_{12}\right)^{-1}$$

$$= -\frac{1}{\sqrt{n}} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} \left(I + \frac{1}{n} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} A_{12} + \left[I - \frac{1}{n} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} A_{12}\right]^{-1}\right.$$

$$\left. \cdot \frac{1}{n^2} (\Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} A_{12})^2\right)$$

$$= -\frac{1}{\sqrt{n}} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} + O\left(\frac{1}{n^{3/2}}\right)$$

$$F_{21} = -\frac{1}{\sqrt{n}} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} + O\left(\frac{1}{n^{3/2}}\right).$$

En posant

$$X_i = \begin{pmatrix} X_i^{(1)} \\ X_i^{(2)} \end{pmatrix}_{p \times 1},$$

où $X_i^{(1)} : p_1 \times 1$ et $X_i^{(2)} : p_2 \times 1$, on peut écrire

$$\begin{aligned}
\ln L_n &= \frac{n}{2} \ln |\Sigma_0| - \frac{n}{2} \ln |\Sigma_n| \\
&\quad - \frac{1}{2} \left\{ \sum_{i=1}^n \begin{pmatrix} X_i^{(1)'} & X_i^{(2)'} \end{pmatrix} \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix} \begin{pmatrix} X_i^{(1)} \\ X_i^{(2)} \end{pmatrix} \right\} \\
&= \frac{n}{2} \ln |\Sigma_0| - \frac{n}{2} \ln |\Sigma_n| - \frac{1}{2} \left\{ \sum_{i=1}^n X_i^{(1)'} F_{11} X_i^{(1)} + \sum_{i=1}^n X_i^{(2)'} F_{22} X_i^{(2)} \right. \\
&\quad \left. + \sum_{i=1}^n X_i^{(2)'} F_{21} X_i^{(1)} + \sum_{i=1}^n X_i^{(1)'} F_{12} X_i^{(2)} \right\} + O_p \left(\frac{1}{\sqrt{n}} \right) \\
&= \frac{n}{2} (\ln |\Sigma_0| - \ln |\Sigma_n|) - \frac{1}{2} \left\{ \frac{1}{n} \sum_{i=1}^n X_i^{(1)'} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1}) X_i^{(1)} \right. \\
&\quad \left. + \frac{1}{n} \sum_{i=1}^n X_i^{(2)'} (\Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1}) X_i^{(2)} \right. \\
&\quad \left. + \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i^{(1)'} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1}) X_i^{(2)} \right. \\
&\quad \left. + \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i^{(2)'} (\Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1}) X_i^{(1)} \right\} + O_p \left(\frac{1}{\sqrt{n}} \right).
\end{aligned}$$

On obtient donc

$$\begin{aligned}
\ln L_n &= \frac{n}{2} \ln |\Sigma_0| - \frac{n}{2} \ln |\Sigma_n| - \frac{1}{2} \left\{ \frac{1}{n} \sum_{i=1}^n X_i^{(1)'} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1}) X_i^{(1)} \right. \\
&\quad \left. + \frac{1}{n} \sum_{i=1}^n X_i^{(2)'} (\Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1}) X_i^{(2)} \right. \\
&\quad \left. + \frac{2}{\sqrt{n}} \sum_{i=1}^n X_i^{(2)'} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1}) X_i^{(1)} \right\} + O_p \left(\frac{1}{\sqrt{n}} \right).
\end{aligned}$$

Nous allons maintenant montrer que la première partie de l'expression de $\ln L_n$, soit

$$\begin{aligned} \frac{n}{2} (\ln |\Sigma_0| - \ln |\Sigma_n|) - \frac{1}{2} \left\{ \frac{1}{n} \sum_{i=1}^n X_i^{(1)'} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1}) X_i^{(1)} \right. \\ \left. + \frac{1}{n} \sum_{i=1}^n X_i^{(2)'} (\Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1}) X_i^{(2)} \right\}, \end{aligned} \quad (\text{B.6})$$

converge en probabilité vers $-\frac{1}{2} \text{tr} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21})$, et que la deuxième partie de l'expression de $\ln L_n$, soit

$$-\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i^{(1)'} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1}) X_i^{(2)}, \quad (\text{B.7})$$

converge en loi vers une loi normale de moyenne 0 et de variance $\text{tr} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21})$. On aura alors le résultat puisque sous $H_0^{(2)}$,

$$\ln L_n \xrightarrow{\mathcal{L}} N \left(-\frac{1}{2} \text{tr} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21}), \text{tr} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21}) \right),$$

qui est bien de la forme $N(-\frac{1}{2}\sigma^2, \sigma^2)$.

Puisque pour un vecteur aléatoire X de moyenne μ et de matrice de covariance V , on a $E(X'AX) = \text{tr}(AV) + \mu' A \mu$ et $\text{Var}(X'AX) = 2\text{tr}(AV)^2 + 4\mu' AV A \mu$ (voir Searle (1971)), on déduit que

$$E \left\{ X_i^{(1)'} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1}) X_i^{(1)} \right\} = \text{tr} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21})$$

et

$$\text{Var} \left\{ X_i^{(1)'} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1}) X_i^{(1)} \right\} = 2\text{tr} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21}).$$

On a donc

$$E \left\{ \frac{1}{n} \sum_{i=1}^n X_i^{(1)'} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1}) X_i^{(1)} \right\} = \text{tr} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21})$$

et

$$\begin{aligned} & \text{Var} \left\{ \frac{1}{n} \sum_{i=1}^n X_i^{(1)'} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1}) X_i^{(1)} \right\} \\ &= \frac{2 \text{tr} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21})}{n} \rightarrow 0 \end{aligned}$$

lorsque $n \rightarrow \infty$ et ainsi,

$$\frac{1}{n} \sum_{i=1}^n X_i^{(1)'} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1}) X_i^{(1)} \xrightarrow{\mathcal{P}} \text{tr} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21}). \quad (\text{B.8})$$

De la même façon, on montre que

$$\frac{1}{n} \sum_{i=1}^n X_i^{(2)'} (\Sigma_{22}^{-1} A_{21} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1}) X_i^{(2)} \xrightarrow{\mathcal{P}} \text{tr} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21}) \quad (\text{B.9})$$

car $\text{tr}(AB) = \text{tr}(BA)$.

Montrons maintenant que

$$\frac{n}{2} (\ln |\Sigma_0| - \ln |\Sigma_n|) \rightarrow \frac{1}{2} \text{tr} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21})$$

lorsque $n \rightarrow \infty$. Comme $|\Sigma_0| = |\Sigma_{11}| |\Sigma_{22}|$ et

$$|\Sigma_n| = \left| \begin{array}{cc} \Sigma_{11} & \frac{1}{\sqrt{n}} A_{12} \\ \frac{1}{\sqrt{n}} A_{21} & \Sigma_{22} \end{array} \right| = \left| \Sigma_{11} - \frac{1}{n} A_{12} \Sigma_{22}^{-1} A_{21} \right| |\Sigma_{22}|,$$

on obtient

$$\begin{aligned} \ln |\Sigma_0| - \ln |\Sigma_n| &= \ln |\Sigma_{11}| - \ln \left| \Sigma_{11} \left(I - \frac{1}{n} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \right) \right| \\ &= - \ln \left| I - \frac{1}{n} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \right|. \end{aligned}$$

Le déterminant que l'on veut calculer est de la forme

$$\begin{vmatrix} 1 - \frac{b_{11}}{n} & -\frac{b_{12}}{n} & \dots & -\frac{b_{1p}}{n} \\ -\frac{b_{12}}{n} & 1 - \frac{b_{22}}{n} & \dots & -\frac{b_{2p}}{n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{b_{1p}}{n} & -\frac{b_{2p}}{n} & \dots & 1 - \frac{b_{pp}}{n} \end{vmatrix},$$

où les éléments de la matrice $\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21}$ sont notés b_{ij} , i et $j = 1, \dots, p$. On obtient d'abord le terme

$$\left(1 - \frac{a_{11}}{n}\right) \left(1 - \frac{a_{22}}{n}\right) \dots \left(1 - \frac{a_{pp}}{n}\right),$$

et tous les autres termes sont au moins d'ordre $\frac{1}{n^2}$ puisque les termes qui forment un déterminant sont toujours composés d'un seul élément de chaque ligne et de chaque colonne. On a donc

$$\left| I - \frac{1}{n} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \right| = 1 - \left(\frac{\sum_{i=1}^p b_{ii}}{n} \right) + O\left(\frac{1}{n^2}\right),$$

et ainsi en développant en série de Taylor,

$$\begin{aligned} n \ln \left| I - \frac{1}{n} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \right| &= -n \left[\left(\frac{\sum_{i=1}^p b_{ii}}{n} \right) + O\left(\frac{1}{n^2}\right) \right] \\ &\rightarrow -\sum_{i=1}^p b_{ii} = -tr(\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21}) \end{aligned}$$

lorsque $n \rightarrow \infty$.

Ainsi, avec (B.8) et (B.9), on obtient donc la convergence en probabilité de (B.6) escomptée.

D'autre part, comme

$$-\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i^{(1)'} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1}) X_i^{(2)}$$

est une forme bilinéaire où $X_i^{(1)}$ suit une loi $N(0, \Sigma_{11})$, $X_i^{(2)}$ suit une loi $N(0, \Sigma_{22})$ et $\Sigma_{12} = Cov(X_i^{(1)}, X_i^{(2)}) = 0$ sous $H_0^{(2)}$, on obtient, en utilisant certains résultats

sur les formes bilinéaires (voir Searle (1971) p.65),

$$\begin{aligned} E \left(X_i^{(1)'} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1}) X_i^{(2)} \right) &= tr \left(\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} Cov(X_i^{(1)}, X_i^{(2)}) \right) \\ &+ E \left(X_i^{(1)'} \right) (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1}) E \left(X_i^{(2)} \right) \\ &= 0 \end{aligned}$$

et, en posant

$$C = \begin{pmatrix} 0 & \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} \\ \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} & 0 \end{pmatrix}$$

et pour $i = 1, \dots, p$,

$$\mu = \begin{pmatrix} E \left(X_i^{(1)} \right) \\ E \left(X_i^{(2)} \right) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$\begin{aligned} Var \left(X_i^{(1)'} (\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1}) X_i^{(2)} \right) &= \frac{1}{2} [tr(C\Sigma)^2 + 2\mu' C \Sigma C \mu] \\ &= tr \left(\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \right), \end{aligned}$$

car

$$C\Sigma \cdot C\Sigma = \begin{pmatrix} \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} & 0 \\ 0 & \Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \end{pmatrix}.$$

Ainsi, par le théorème limite central, on a bien (B.7) qui converge en loi vers une loi normale

$$N \left(0, tr \left(\Sigma_{11}^{-1} A_{12} \Sigma_{22}^{-1} A_{21} \right) \right).$$

C.Q.F.D.

BIBLIOGRAPHIE

- Abrahamse, A. P. J. et Koerts, J. (1969), 'A comparison between the power of the Durbin-Watson test and the power of the BLUS test', *Journal of the American Statistical Association* **64**, 938-948.
- Allaire, J. et 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.
- Anderson, T. W. (1984), *An introduction to multivariate statistical analysis*, New York, Wiley.
- Bahadur, R. R. (1960), 'Stochastic comparison of tests', *The Annals of Mathematical Statistics* **31**, 276-295.
- Bahadur, R. R. (1967), 'Rates of convergence of estimates and test statistics', *The Annals of Mathematical Statistics* **38**, 303-324.
- Bahadur, R. R. (1971), *Some limit theorems in statistics*, Regional conference series in applied mathematics.
- Bhuchongkul, S. (1964), 'A class of nonparametric tests for independent in bivariate populations', *The Annals of Mathematical Statistics* **35**, 138-149.
- Billingsley, P. (1968), *Convergence of probability measures*, Toronto, Wiley.
- Chernoff, H. (1952), 'A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations', *The Annals of Mathematical Statistics* **23**, 493-507.
- Chernoff, H. et Savage, I. R. (1958), 'Asymptotic normality and efficiency of certain nonparametric test statistics', *The Annals of Mathematical Statistics* **29**, 972-994.
- Cléroux, R., Lazraq, A. et Lepage, Y. (1994), 'Indice de redondance basé sur les rangs et inférence non paramétrique', *Revue de la Statistique Appliquée* **42**(2), 79-98.
- Cléroux, R., Lazraq, A. et Lepage, Y. (1995), 'Vector correlation based on ranks and a nonparametric test of no association between vectors', *Communications in Statistics, Part A - Theory and Methods* **24**, 713-733.

- Coxhead, P. (1974), 'Measuring the relationships between two sets of variables', *British journal of mathematical and statistical psychology* **27**, 205–212.
- Cramer, E. M. (1974), 'A generalization of vector correlation and its relation to canonical correlation', *Multivariate Behavioral Research* **9**, 347–352.
- Cramer, E. M. et Nicewander, W. A. (1979), 'Some symmetric, invariant measures of multivariate association', *Psychometrika* **44**, 43–54.
- Dauxois, J. et Nkiet, G. M. (1997), 'Testing for the lack of a linear relationship', *Statistics* **30**, 1–23.
- Dauxois, J. et Nkiet, G. M. (1998), 'Nonlinear canonical analysis and independence tests', *The Annals of Statistics* **26**, 1254–1278.
- El Maâche, H. et Lepage, Y. (1998), 'Mesures d'association vectorielle basées sur une matrice de corrélation', *Revue de la Statistique Appliquée* **46**(4), 27–43.
- Escoufier, Y. (1973), 'Le traitement des variables vectorielles', *Biometrics* **29**, 751–760.
- Gill, P., Murray, W. et Wright, M. H. (1981), *Practical optimization*, Academic Press.
- Gleason, T. C. (1976), 'On redundancy in canonical analysis', *Psychological Bulletin* **83**, 1004–1006.
- Govindarajulu, Z., Le Cam, L. et Raghavachari, M. (1966), 'Generalizations of theorems of Chernoff and Savage on the asymptotic normality of test statistics', *University of California Press* **1**, 609–638.
- Gumbel, E. J. (1961), 'Bivariate logistic distributions', *Journal of the American Statistical Association* **56**, 335–349.
- Hájek, J. et Šidák, Z. (1967), *Theory of rank tests*, New York, Academic Press.
- Hodges, J. et Lehmann, E. L. (1956), 'The efficiency of some nonparametric competitors of the t-test', *The Annals of Mathematical Statistics* **27**, 324–335.
- Hoeffding, W. (1948), 'A class of statistics with asymptotically normal distribution', *The Annals of Mathematical Statistics* **19**, 293–325.
- Hoeffding, W. (1953), 'On the distribution of the expected values of the order statistics', *The Annals of Mathematical Statistics* **24**, 93–100.
- Hotelling, H. (1936), 'Relations between two sets of variables', *Biomtrika* **28**, 321–377.

- Hotelling, H. et Pabst, M. R. (1936), 'Rank correlation and tests of significance involving no assumptions of normality', *The Annals of Mathematical Statistics* **7**, 29–43.
- Imhof, P. (1961), 'Computing the distribution of quadratic forms in normal variates', *Biometrika* **48**, 419–426.
- Johnson, N. L. et Kotz, S. (1972), *Distributions in statistics*, New York, Wiley.
- Kallenberg, O. (1983), *Random Measures*, Academic Press, Brelin.
- Kendall, M. G. et Stuart, A. (1977), *The advanced theory of statistics*, seconde édition, London, Griffin.
- Lazraq, A. (1989), *Inférences sur plusieurs mesures de liaison entre deux vecteurs aléatoires et algorithmes de sélection de variables*, Université de Montréal. Thèse de doctorat.
- Lazraq, A. et Cléroux, R. (1988), 'Étude comparative de différentes mesures de liaison entre deux vecteurs aléatoires et tests d'indépendance', *Statistique et Analyse des Données* **13**(1), 15–38.
- Lazraq, A., Lepage, Y. et Cléroux, R. (1995), 'Tests non paramétriques pour l'indépendance entre plusieurs vecteurs aléatoires', *Publications Institut Statistique Université de Paris* **39**(3), 57–77.
- Le Cam, L. (1960), 'Locally asymptotically normal families of distributions', *University of California Publications in Statistics* **3**, 37–98.
- Muirhead, R. J. (1982), *Aspects of multivariate statistical theory*, New York, Wiley.
- Muirhead, R. J. et Waternaux, C. M. (1980), 'Asymptotic distributions in canonical correlation analysis and other multivariate procedures for nonnormal populations', *Biometrika* **67**, 31–43.
- Nagao, H. (1973), 'On some test criteria for covariance matrix', *The Annals of Statistics* **1**, 700–709.
- Nyblom, J. et Mäkeläinen, T. (1983), 'Comparisons of tests for the presence of random walk coefficients in a simple linear model', *Journal of the American Statistical Association* **78**, 856–864.
- Pitman, E. J. G. (1949), 'Lecture notes on nonparametric statistical inference', *Columbia University*.

- Puri, M. L. (1964), 'Asymptotic efficiency of a class of c -sample tests', *The Annals of Mathematical Statistics* **35**, 102–121.
- Puri, M. L. et Sen, P. K. (1966), 'On a class of multivariate multisample rank-order tests', *Sankhya, Series A, Indian Journal of Statistics* **28**, 353–376.
- Puri, M. L. et Sen, P. K. (1971), *Nonparametric methods in multivariate analysis*, New York, Wiley.
- Puri, M. L., Sen, P. K. et Gokhale, D. V. (1970), 'On a class of rank order tests for independence in multivariate distributions', *Sankhya, Series A, Indian Journal of Statistics* **32**, 271–298.
- Ramsey, J. O., ten Berge et Styau, G. P. H. (1984), 'Matrix correlation', *Psychometrika* **49**, 403–423.
- Roseboom, N. (1965), 'Linear correlation between sets of variables', *Psychometrika* **30**, 57–71.
- Roy, R. et Cléroux, R. (1993), 'Vector cross-correlation in time series and applications', *International Statistical Review* **61**, 447–464.
- Searle, S. R. (1971), *Linear models*, New York, Wiley.
- Serfling, R. J. (1980), *Approximation theorems of mathematical statistics*, New York, Wiley.
- Shaffer, J. P. et Gillo, M. W. (1974), 'A multivariate extension of the correlation ratio', *Educational and Psychological Measurement* **34**, 521–524.
- Spivak, M. (1980), *Calculus*, Publish or Perish, Berkeley.
- Stephens, M. A. (1979), 'Vector correlation', *Biometrika* **66**, 41–48.
- Stewart, D. et Love, W. (1968), 'A general canonical correlation index', *Psychological Bulletin* **70**, 160–163.
- von Mises, R. (1947), 'On the asymptotic distribution of differentiable statistical functions', *The Annals of Mathematical Statistics* **18**, 309–348.
- Wieand, H. S. (1976), 'A condition under which the Pitman and Bahadur approaches to efficiency coincide', *The Annals of Statistics* **4**, 1003–1011.