



Une information sur les matrices de covariance : la liaison-information

Christian Michaux

► **To cite this version:**

Christian Michaux. Une information sur les matrices de covariance : la liaison-information. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 1973. Français. <tel-00284034>

HAL Id: tel-00284034

<https://tel.archives-ouvertes.fr/tel-00284034>

Submitted on 2 Jun 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N° d'ordre

THESE

présentée à

L'UNIVERSITE SCIENTIFIQUE ET MEDICALE DE GRENOBLE

pour obtenir

LE GRADE DE DOCTEUR DE TROISIEME CYCLE

"Informatique"

par

Christian MICHAUX

— o —

UNE INFORMATION SUR LES MATRICES DE COVARIANCE

LA LIAISON - INFORMATION

— o —

Thèse soutenue le 30 Juin 1973 devant la commission d'examen

PRESIDENT : MONSIEUR J. KUNTZMANN

EXAMINATEURS : MESSIEURS G. ROMIER

M. TERRENOIRE

B. VAN CUTSEM

F. PECCOUD

Président : Monsieur Michel SOUTIF

Vice-Président : Monsieur Gabriel CAU

PROFESSEURS TITULAIRES

MM. ANGLES D'AURIAC Paul
ARNAUD Georges
ARNAUD Paul
AUBERT Guy
AYANT Yves
Mme BARBIER Marie-Jeanne
MM. BARBIER Jean-Claude
BARBIER Reynold
BARJON Robert
BARNOUD Fernand
BARRA Jean-René
BARRIE Joseph
BENOIT Jean
BERNARD Alain
BESSON Jean
BEZES Henri
BLAMBERT Maurice
BOLLIET Louis
BONNET Georges
BONNET Jean-Louis
BONNET-EYMARD Joseph
BONNIER Etienne
BOUCHERLE André
BOUCHEZ Robert
BOUSSARD Jean-Claude
BRAVARD Yves
BRISSONNEAU Pierre
BUYLE-BODIN Maurice
CABANAC Jean
CABANEL Jean
CALAS François
CARRAZ Gilbert
CAU Gabriel
CAUQUIS Georges
CHABAUTY Claude
CHARACHON Robert
CHATEAU Robert
CHENE Marcel
COEUR André
CONTAMIN Robert
COUDERC Pierre
CRAYA Antoine

Mécanique des fluides
Clinique des maladies infectieuses
Chimie
Physique
Physique approfondie
Electrochimie
Physique expérimentale
Géologie appliquée
Physique nucléaire
Biosynthèse de la cellulose
Statistiques
Clinique chirurgicale
Radioélectricité
Mathématiques Pures
Electrochimie
Chirurgie générale
Mathématiques Pures
Informatique (IUT B)
Electrotechnique
Clinique ophtalmologique
Pathologie médicale
Electrochimie Electrométallurgie
Chimie et Toxicologie
Physique nucléaire
Mathématiques Appliquées
Géographie
Physique du solide
Electronique
Pathologie chirurgicale
Clinique rhumatologique et hydrologie
Anatomie
Biologie animale et pharmacodynamie
Médecine légale et Toxicologie
Chimie organique
Mathématiques Pures
Oto-Rhino-Laryngologie
Thérapeutique
Chimie papetière
Pharmacie chimique
Clinique gynécologique
Anatomie Pathologique
Mécanique

Mme	DEBELMAS Anne-Marie	Matière médicale
MM.	DEBELMAS Jacques	Géologie générale
	DEGRANGE Charles	Zoologie
	DESRE Pierre	Métallurgie
	DESSAUX Georges	Physiologie animale
	DODU Jacques	Mécanique appliquée
	DOLIQUE Jean-Michel	Physique des plasmas
	DREYFUS Bernard	Thermodynamique
	DUCROS Pierre	Cristallographie
	DUGOIS Pierre	Clinique de Dermatologie et Syphiligraphie
	FAU René	Clinique neuro-psychiatrique
	FELICI Noël	Electrostatique
	GAGNAIRE Didier	Chimie physique
	GALLISSOT François	Mathématiques Pures
	GALVANI Octave	Mathématiques Pures
	GASTINEL Noël	Analyse numérique
	GEINDRE Michel	Electroradiologie
	GERBER Robert	Mathématiques Pures
	GIRAUD Pierre	Géologie
	KLEIN Joseph	Mathématiques Pures
Mme	KOFLER Lucie	Botanique et Physiologie végétale
MM.	KOSZUL Jean-Louis	Mathématiques Pures
	KRAVTCHENKO Julien	Mécanique
	KUNTZMANN Jean	Mathématiques appliquées
	LACAZE Albert	Thermodynamique
	LACHARME Jean	Biologie végétale
	LAJZEROWICZ Joseph	Physique
	LATREILLE René	Chirurgie générale
	LATURAZE Jean	Biochimie pharmaceutique
	LAURENT Pierre-Jean	Mathématiques appliquées
	LEDRU Jean	Clinique médicale B
	LLIBOUTRY Louis	Géophysique
	LOUP Jean	Géographie
Mlle	LUTZ Elisabeth	Mathématiques Pures
MM.	MALGRANGE Bernard	Mathématiques Pures
	MALINAS Yves	Clinique obstétricale
	MARTIN-NOEL Pierre	Seméiologie médicale
	MASSEPORT Jean	Géographie
	MAZARE Yves	Clinique médicale A
	MICHEL Robert	Minéralogie et Pétrographie
	MOURIQUAND Claude	Histologie
	MOUSSA André	Chimie nucléaire
	NEEL Louis	Physique du solide
	OZENDA Paul	Botanique
	FAUTHENET René	Electrotechnique
	PAYAN Jean-Jacques	Mathématiques Pures
	PEBAY-PEYROULA Jean-Claude	Physique
	PERRET René	Servomécanismes
	PILLET Emile	Physique industrielle
	RASSAT André	Chimie systématique
	RENARD Michel	Thermodynamique
	REULOS René	Physique industrielle
	RINALDI Renaud	Physique
	ROGET Jean	Clinique de pédiatrie et de puériculture
	SANTON Lucien	Mécanique
	SEIGNEURIN Raymond	Microbiologie et Hygiène
	SENGEL Philippe	Zoologie
	SILBERT Robert	Mécanique des fluides
	SOUTIF Michel	Physique générale

MM.	TANCHE Maurice	Physiologie
	TRAYNARD Philippe	Chimie générale
	VAILLAND François	Zoologie
	VALENTIN Jacques	Physique nucléaire
	VAUQUOIS Bernard	Calcul électronique
Mme	VERAIN Alice	Pharmacie galénique
M.	VERAIN André	Physique
Mme	VEYRET Germaine	Géographie
MM.	VEYRET Paul	Géographie
	VIGNAIS Pierre	Biochimie médicale
	YOCOZ Jean	Physique nucléaire théorique

PROFESSEURS ASSOCIES

MM.	BULLEMER Bernhard	Physique
	HANO JUN-ICHI	Mathématiques Pures
	STEPHENS Michaël	Mathématiques appliquées

PROFESSEURS SANS CHAIRE

MM.	BEAUDOING André	Pédiatrie
Mme	BERTRANDIAS Françoise	Mathématiques Pures
MM.	BERTRANDIAS Jean-Paul	Mathématiques appliquées
	BIAREZ Jean-Pierre	Mécanique
	BONNETAIN Lucien	Chimie minérale
Mme	BONNIER Jane	Chimie générale
MM.	CARLIER Georges	Biologie végétale
	COHEN Joseph	Electrotechnique
	COUMES André	Radioélectricité
	DEPASSEL Roger	Mécanique des fluides
	DEPORTES Charles	Chimie minérale
	GAUTHIER Yves	Sciences biologiques
	GAVEND Michel	Pharmacologie
	GERMAIN Jean-Pierre	Mécanique
	GIDON Paul	Géologie et Minéralogie
	GLENAT René	Chimie organique
	HACQUES Gérard	Calcul numérique
	JANIN Bernard	Géographie
Mme	KAHANE Josette	Physique
MM.	MULLER Jean-Michel	Thérapeutique
	PERRIAUX Jean-Jacques	Géologie et Minéralogie
	POULOUJADOFF Michel	Electrotechnique
	REBECQ Jacques	Biologie (CUS)
	REVOL Michel	Urologie
	REYMOND Jean-Charles	Chirurgie générale
	ROBERT André	Chimie papetière
	DE ROUGEMONT Jacques	Neurochirurgie
	SARRAZIN Roger	Anatomie et chirurgie
	SARROT-REYNAULD Jean	Géologie
	SIBILLE Robert	Construction mécanique
	SIROT Louis	Chirurgie générale
Mme	SOUTIF Jeanne	Physique générale

MAITRES DE CONFERENCES ET MAITRES DE CONFERENCES AGREGES

Mle	AGNIUS-DELORD Claudine	Physique pharmaceutique
	ALARY Josette	Chimie analytique
MM.	AMBLARD Pierre	Dermatologie
	AMBROISE-THOMAS Pierre	Parasitologie
	ARMAND Yves	Chimie
	BEGUIN Claude	Chimie organique
	BELORIZKY Elie	Physique
	BENZAKEN Claude	Mathématiques appliquées
	BILLET Jean	Géographie
	BLIMAN Samuel	Electronique (EIE)
	BLOCH Daniel	Electrotechnique
Mme	BOUCHE Liane	Mathématiques (CUS)
MM.	BOUCHET Yves	Anatomie
	BOUVARD Maurice	Mécanique des fluides
	BRODEAU François	Mathématiques (IUT B)
	BRUGEL Lucien	Energétique
	BUISSON Roger	Physique
	BUTEL Jean	Orthopédie
	CHAMBAZ Edmond	Biochimie médicale
	CHAMPETIER Jean	Anatomie et organogénèse
	CHIAVERINA Jean	Biologie appliquée (EFP)
	CHIBON Pierre	Biologie animale
	COHEN-ADDAD Jean-Pierre	Spectrométrie physique
	COLOMB Maurice	Biochimie médicale
	CONTE René	Physique
	COULOMB Max	Radiologie
	CROUZET Guy	Radiologie
	DURAND Francis	Métallurgie
	DUSSAUD René	Mathématiques (CUS)
Mme	ETERRADOSSI Jacqueline	Physiologie
MM.	FAURE Jacques	Médecine légale
	GENSAC Pierre	Botanique
	GIDON Maurice	Géologie
	GRIFFITHS Michaël	Mathématiques appliquées
	GROULADE Joseph	Biochimie médicale
	HOLLARD Daniel	Hématologie
	HUGONOT Robert	Hygiène et Médecine préventive
	IDELMAN Simon	Physiologie animale
	IVANES Marcel	Electricité
	JALBERT Pierre	Histologie
	JOLY Jean-René	Mathématiques Pures
	JOUBERT Jean-Claude	Physique du solide
	JULLIEN Pierre	Mathématiques Pures
	KAHANE André	Physique générale
	KUHN Gérard	Physique
	LACOUME Jean-Louis	Physique
Mme	LAJZEROWICZ Jeannine	Physique
MM.	LANCIA Roland	Physique atomique
	LE JUNIER Noël	Electronique
	LEROY Philippe	Mathématiques
	LOISEAUX Jean-Marie	Physique nucléaire
	LONGEQUEUE Jean-Pierre	Physique nucléaire
	LUU DUC Cuong	Chimie organique
	MACHE Régis	Physiologie végétale
	MAGNIN Robert	Hygiène et Médecine préventive
	MARECHAL Jean	Mécanique
	MARTIN-BOUYER Michel	Chimie (CUS)

MM.	MAYNARD Roger	Physique du solide
	MICHOULIER Jean	Physique (IUT A)
	MICLOUD Max	Maladies infectieuses
	MOREAU René	Hydraulique (INP)
	NEGRE Robert	Mécanique
	PARAMELLE Bernard	Pneumologie
	PECCOUD François	Analyse (IUT B)
	PEFFEN René	Métallurgie
	PELMONT Jean	Physiologie animale
	PERRET Jean	Neurologie
	PERRIN Louis	Pathologie expérimentale
	PFISTER Jean-Claude	Physique du solide
	PHELIP Xavier	Rhumatologie
Mlle	RIERY Yvette	Biologie animale
MM.	RACHAIL Michel	Médecine interne
	RACINET Claude	Gynécologie et obstétrique
	RENAUD Maurice	Chimie
	RICHARD Lucien	Botanique
Mme	RINAUDO Marquerite	Chimie macromoléculaire
MM.	ROMIER Guy	Mathématiques (IUT B)
	SHOM Jean-Claude	Chimie générale
	STIEGLITZ Paul	Anesthésiologie
	STOEBNER Pierre	Anatomie pathologique
	VAN CUTSEM Bernard	Mathématiques appliquées
	VEILLON Gérard	Mathématiques appliquées (INP)
	VIALON Pierre	Géologie
	VOOG Robert	Médecine interne
	VROUSSOS Constantin	Radiologie
	ZADWORNY François	Electronique

MAITRES DE CONFERENCES ASSOCIES

MM.	BOUDOURIS Georges	Radioélectricité
	CHEEKE John	Thermodynamique
	GOLDSCHMIDT Hubert	Mathématiques
	SIDNEY STUARD	Mathématiques Pures
	YACOUD Mahmoud	Médecine légale

CHARGES DE FONCTIONS DE MAITRES DE CONFERENCES

Mme	BERIEL Hélène	Physiologie
Mme	RENAUDET Jacqueline	Microbiologie

Fait le 30 mai 1972.

Je tiens à exprimer mes plus vifs remerciements à

*Monsieur le Professeur J. KUNTZMANN,
qui a bien voulu me faire l'honneur de présider
le jury de cette thèse.*

*Monsieur G. ROMIER,
Professeur à l'Université II de Grenoble,
qui a suivi avec intérêt le développement de cette étude.*

*Messieurs B. VAN CUTSEM et F. PECCOUD,
qui ont bien voulu examiner ce travail
et faire partie du jury.*

*Je suis particulièrement reconnaissant à Monsieur M. TERRENOIRE
qui a inspiré et dirigé cette étude et m'a aidé constamment par ses
nombreux conseils et suggestions.*

*Je remercie vivement Monsieur le Professeur C.F. PICARD
pour l'aide qu'il m'a apportée et les encouragements qu'il m'a prodigués
tout au long de ce travail.*

*Je remercie enfin tous ceux qui ont permis la réalisation matérielle
de cette thèse.*

TERMINOLOGIE

* Si X et Y sont deux variables aléatoires

E(X) et V(X) représentent respectivement l'espérance et la variance de X.

Covariance (X,Y) sera la covariance de X et de Y.

* $\mathbb{R}^+ = \{x \in \mathbb{R} / x > 0\}$

* $\mathbb{R}_0^+ = \mathbb{R}^+ \cup \{0\}$

* Etant donné n variables aléatoires Y_1, Y_2, \dots, Y_n , on notera

- $\text{COV}(Y_1, Y_2, \dots, Y_n)$ leur matrice de covariance

- $\text{COR}(Y_1, Y_2, \dots, Y_n)$ leur matrice de corrélation

- $\text{DIAG}(Y_1, Y_2, \dots, Y_n)$ leur matrice de covariance si elles sont non corrélées.

* Une matrice (n x n) aura n lignes et n colonnes.

- $C = ((c_{ij}))$ sera la matrice de terme général c_{ij} , i indice de ligne et j indice de colonne.

- Soit $C = ((c_{ij}))$ une matrice (n x n) ; on notera C_i^j la matrice suivante :

$$C_i^j = \begin{pmatrix} c_{ii} & c_{ij} \\ c_{ij} & c_{jj} \end{pmatrix}$$

* C_n, R_n, D_n représentent respectivement l'ensemble des matrices de covariance (n x n), l'ensemble des matrices de corrélation (n x n), et l'ensemble des matrices de covariance diagonales (n x n).

$$- C = \bigcup_{n=2}^{\infty} C_n$$

$$- R = \bigcup_{n=2}^{\infty} R_n$$

$$- D = \bigcup_{n=2}^{\infty} D_n$$

* Soit $C \in C_n$ d'éléments diagonaux $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$ non nuls. On dira que $R \in R_n$ est déduite de $C \in C_n$ si

$$R = \begin{pmatrix} \frac{1}{\sigma_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sigma_n} \end{pmatrix} \subset \begin{pmatrix} \frac{1}{\sigma_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sigma_n} \end{pmatrix}$$

* Etant donné une arborescence $A = (X, \Gamma)$, on notera :

- x_0 sa racine
- (x, y) l'arc joignant le sommet x au sommet y
- $A(x)$ l'ensemble des ascendants de x dans A .

INTRODUCTION

Etant donné un vecteur aléatoire Y dont les composantes sont les variables aléatoires Y_1, Y_2, \dots, Y_n définies sur une population Ω , et sa matrice de covariance associée, nous allons définir dans les deux premiers chapitres une information qui permette de "mesurer" la liaison entre ces n variables.

Si, en plus du vecteur aléatoire Y , nous disposons d'une variable aléatoire X , définie également sur Ω , nous nous proposons dans le chapitre III de définir une fonction qui puisse "mesurer" la liaison entre X et Y .

Les applications pratiques de ces "liaisons-informations" à l'analyse factorielle et à des problèmes de segmentation conduisent à des résultats satisfaisants comme le montreront les résultats numériques exposés dans le dernier chapitre.

CHAPITRE I

LIAISON - INFORMATION

INTRODUCTION

Nous allons nous intéresser aux applications de $C \rightarrow \mathbb{R}_0^+$ soit $\mathcal{F}(C \rightarrow \mathbb{R}_0^+)$ qui devront vérifier certaines propriétés qui ont été introduites en accord avec les résultats suivants :

- Le coefficient de corrélation linéaire entre deux variables aléatoires X et Y est un nombre compris entre -1 et $+1$, égal à ± 1 si la liaison entre X et Y est linéaire. De plus, si l'on modifie linéairement le codage de ces variables aléatoires, sa valeur est inchangée.

- Soit $R \in R_n$ et soit $D \in D_n$ la matrice diagonale semblable à R (R est diagonalisable, puisque R est symétrique réelle).

Si V_i est le vecteur propre associé à la valeur propre λ_i alors λ_i représente la variance de V_i [8]. Donc la matrice D est la matrice de covariance des vecteurs propres de R .

I - PROPRIETES DE LA LIAISON-INFORMATION

1. Remarques préliminaires

- Les covariances se calculent à partir des variables prises deux à deux. Nous respecterons cette notion de couple en introduisant une propriété d'additivité des informations des variables prises deux à deux (propriété I-1).

- L'information sera un indicateur de désordre pour les variables aléatoires, elle croîtra donc en même temps que les variances. De plus, si nous remarquons que toute matrice diagonale (2 x 2) peut se décomposer en un produit de deux matrices diagonales de la façon suivante :

$$\begin{pmatrix} u & 0 \\ 0 & v \end{pmatrix} = \begin{pmatrix} u & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & v \end{pmatrix}$$

nous pouvons introduire une propriété multiplicative faisant intervenir cette décomposition (propriété I-2).

- Dans l'introduction du Chapitre I, nous avons vu que la matrice diagonale d'une matrice de corrélation R est la matrice de covariance des vecteurs propres de R. Or, la diagonalisation étant uniquement un changement des vecteurs de base, nous introduirons une propriété de conservation d'information lors d'une diagonalisation (propriété I-3).

2. Propriétés

Les applications $I \in \mathcal{F} (C \rightarrow \mathbb{R}_0^+)$ qui vérifieront les quatre propriétés suivantes, seront appelées liaison-information.

PROPRIETE I-1

Soit $C \in C_n$ alors

$$I(C) = h(n) \sum_{i=1}^{n-1} \sum_{j=i+1}^n I(C_i^j) \quad \forall n \geq 2, n \in \mathbb{N}$$

h étant une application de $\mathbb{N} \rightarrow \mathbb{R}^+$

PROPRIETE I-2

Soit $C \in C_2$ une matrice dont les éléments diagonaux sont σ_1^2 et σ_2^2 et soit R la matrice déduite de C alors

$$I(C) = \psi(\sigma_1^2, \sigma_2^2) I(R)$$

où ψ est une application de $\mathbb{R}_0^+ \times \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$ vérifiant les propriétés suivantes :

- ψ admet des dérivées partielles continues par rapport à chacune de ses variables en tous points de l'ensemble de définition.

- $\psi(u,v)$ est non décroissante en u et v si $(u,v) \in \mathbb{R}_0^+ \times \mathbb{R}_0^+$ et est strictement croissante en u et v si $(u,v) \in \mathbb{R}^+ \times \mathbb{R}^+$.

$$- \psi(u,v) = \psi(u,1) \psi(1,v) \quad \forall (u,v) \in \mathbb{R}_0^+ \times \mathbb{R}_0^+$$

PROPRIETE I-3

Soient $R \in R$ et $D \in D$ semblable à R alors

$$I(R) = I(D)$$

PROPRIETE I-4

Si $C \in C$ est une matrice unitaire alors

$$I(C) = 1$$

II - CONSEQUENCES DES QUATRE PROPRIETES PRECEDENTES

1. Conséquence I-1

L'application h de la propriété I-1 a pour expression :

$$h(n) = \frac{2}{n(n-1)} \quad \forall n \in \mathbb{N}, n \geq 2.$$

DEMONSTRATION

Soit C la matrice de covariance unitaire de dimension $(n \times n)$ alors

$$I(C) = 1 \quad (\text{propriété I-4})$$

De plus $C_i^j = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ donc $I(C_i^j) = 1$.

En vertu de la propriété I-1, on obtient :

$$1 = h(n) \sum_{i=1}^{n-1} \sum_{j=i+1}^n 1 \quad \forall n \geq 2, n \in \mathbb{N}.$$

$$\text{D'où } h(n) = \frac{2}{n(n-1)} \quad \forall n \geq 2, n \in \mathbb{N}.$$

2. Conséquence I-2

$$\psi(u, v) = \psi(v, u) \quad \forall (u, v) \in [0, 2] \times [0, 2]$$

DEMONSTRATION

Soit $R = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix} \in R$, $r \in [-1, 1]$

alors $D_1 = \begin{pmatrix} 1-r & 0 \\ 0 & 1+r \end{pmatrix}$ et $D_2 = \begin{pmatrix} 1+r & 0 \\ 0 & 1-r \end{pmatrix}$

sont les deux matrices diagonales semblables à R .

Comme $I(R) \doteq I(D_1) = I(D_2)$ (propriété I-3)

et que $I(R) = \psi(1-r, 1+r) I \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \psi(1+r, 1-r) I \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ (propriété I-2)

alors $\psi(1-r, 1+r) = \psi(1+r, 1-r) \quad \forall r \in [-1, 1]$

et donc que $\psi(u, v) = \psi(v, u) \quad \forall (u, v) \in [0, 2] \times [0, 2]$.

3. Conséquence I-3

Soit $R \in R_n$ et $D \in D_n$ semblable à R .

Si $D = ((\lambda_i))$ $i=1,2,\dots,n$ et $R = ((r_{ij}))$ $i,j=1,2,\dots,n$

$$\text{alors } I(R) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \psi(1-r_{ij}, 1+r_{ij}) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \psi(\lambda_i, \lambda_j)$$

DEMONSTRATION

$$\text{Soit } R_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ alors } I(R_2) = 1 \quad (\text{propriété I-4})$$

$$- I(R) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n I(R_i^j) \quad \forall n \geq 2, n \in \mathbb{N} \quad (\text{propriété I-1})$$

Une des matrices diagonales $D(i,j)$ semblable à R_i^j est telle que

$$D(i,j) = \begin{pmatrix} 1-r_{ij} & 0 \\ 0 & 1+r_{ij} \end{pmatrix}$$

$$\text{comme } I(R_i^j) = I(D(i,j)) \quad (\text{propriété I-3})$$

$$\text{et comme } I(D(i,j)) = \psi(1-r_{ij}, 1+r_{ij}) I(R_2)$$

$$\text{alors } I(R) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \psi(1-r_{ij}, 1+r_{ij}) \quad \forall n \in \mathbb{N} \\ n \geq 2.$$

$$- \text{On a } I(R) = I(D) \quad (\text{propriété I-3})$$

$$\text{et } I(D) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n I(D_i^j) \quad (\text{propriété I-1})$$

$$\text{Comme } I(D_i^j) = \psi(\lambda_i, \lambda_j) I(R_2) \quad (\text{propriété I-2})$$

$$\text{alors } I(R) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \psi(\lambda_i, \lambda_j)$$

REMARQUE

$$I(R_2) = \psi(1,1) I(R_2) \quad (\text{propriété I-2})$$

$$\text{donc } \psi(1,1) = 1.$$

DEMONSTRATION

- Soit $\Psi(r)$ l'application de $[-1,1] \rightarrow \mathbb{R}_0^+ \times \mathbb{R}_0^+$ définie de la façon suivante :

$$\Psi(r) = (1+r, 1-r) \quad \forall r \in [-1,1]$$

Ψ est continuellement dérivable $\forall r \in [-1,1]$

Or, la fonction ψ admet des dérivées partielles continues par rapport à chacune de ses variables, donc la fonction $g = \psi \circ \Psi$ est continuellement dérivable $\forall r \in [-1,1]$.

- $\psi(1+r, 1-r) = \psi(1-r, 1+r)$ (conséquence I-2)

Or, $\psi(1-r, 1+r) = g(-r)$ par définition de g .

Donc $g(r) = g(-r) \quad \forall r \in [-1,1]$.

6. Conséquence I-6

Soit $f(x) = \psi(1+x, 1) \quad \forall x \in [-1, \infty]$

- i) $f(x)$ est continuellement dérivable $\forall x \in [-1, \infty]$
- ii) $f'(x) > 0 \quad \forall x \in]-1, \infty]$
 $f'(x) \geq 0$ si $x = -1$.

DEMONSTRATION

- De la même façon que l'on a montré que $g(r) = \psi(1+r, 1-r)$ est continuellement dérivable $\forall r \in [-1,1]$, on montrerait que $f(x)$ est continuellement dérivable $\forall x \in [-1, \infty]$.

- $\psi(1+x, 1)$ est une fonction strictement croissante si $x > -1$ et non décroissante si $x \geq -1$. (propriété I-2).

donc $f'(x) > 0 \quad \forall x \in]-1, \infty]$
 $f'(x) \geq 0$ si $x = -1$.

7. Conséquence I-7

La fonction ψ vérifie les deux relations suivantes :

$$\text{i) } \psi(1+r,1) + \psi(1-r,1) = 2 \quad \forall r \in [-1,1]$$

$$\text{ii) } 3\psi(1+r,1-r) = \psi(1-r,1-r) + 2\psi(1-r,1+2r) \quad \forall r \in [0,1]$$

DEMONSTRATION

- Soit $R \in R_3$ définie de la façon suivante :

$$R = \begin{pmatrix} 1 & r & 0 \\ r & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{où } r \in [-1,1]$$

Le polynôme caractéristique est

$$P(\lambda) = (1-\lambda)^3 - r^2(1-\lambda) = (1-\lambda)(1-\lambda+r)(1-\lambda-r)$$

Les valeurs de λ qui annulent $P(\lambda)$ sont donc

$$\lambda_1 = 1+r, \lambda_2 = 1-r, \lambda_3 = 1.$$

En conséquence, la matrice $D \in D$ semblable à R peut s'écrire

$$D = \begin{pmatrix} 1+r & 0 & 0 \\ 0 & 1-r & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Calculons $I(R)$ en utilisant la conséquence I-2 :

$$I(R) = \frac{1}{3} [\psi(1-r,1+r) + 2 \psi(1,1)]$$

$$I(R) = \frac{1}{3} [\psi(1+r,1-r) + \psi(1+r,1) + \psi(1-r,1)]$$

donc comme $\psi(1,1) = 1$

$$\psi(1+r,1) + \psi(1-r,1) = 2 \quad \forall r \in [-1,1]$$

- Soit $R \in R_3$ définie de la façon suivante :

$$R = \begin{pmatrix} 1 & r & r \\ r & 1 & r \\ r & r & 1 \end{pmatrix} \quad \text{où } r \in [0,1]$$

Le polynôme caractéristique de la matrice R est :

$$P(\lambda) = (1-\lambda)^3 + 2r^3 - 3r^2(1-\lambda) = -(\lambda-1+r)^2 (\lambda-1-2r)$$

Les valeurs propres de R sont donc :

$$\lambda_1 = \lambda_2 = 1-r, \quad \lambda_3 = 1+2r$$

En conséquence la matrice $D \in D$ semblable à R peut s'écrire :

$$D = \begin{pmatrix} 1-r & 0 & 0 \\ 0 & 1-r & 0 \\ 0 & 0 & 1+2r \end{pmatrix}$$

Calculons $I(R)$ en utilisant la conséquence I-3 :

$$I(R) = \frac{1}{3} [3\psi(1-r, 1+r)]$$

$$I(R) = \frac{1}{3} [\psi(1-r, 1-r) + 2\psi(1-r, 1+2r)]$$

donc $3\psi(1-r, 1+r) = \psi(1-r, 1-r) + 2\psi(1-r, 1+2r) \quad \forall r \in [0,1].$

III - EXPRESSION DE LA LIAISON-INFORMATION

Dans ce paragraphe, nous allons déterminer la fonction $I \in \mathcal{F}(C \rightarrow \mathbb{R}_0^+)$ qui vérifiera les quatre propriétés et leurs conséquences.

1. Théorème I-1

Etant donné une matrice $R \in R_2$, où $R = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}$

la seule fonction $I(R) = \psi(1+r, 1-r)$ $r \in [-1, 1]$ vérifiant les quatre propriétés précédentes est la fonction

$$g(r) = 1 - \alpha^2 r^2 \text{ où } \alpha \text{ quelconque } \in]0, 1]$$

NOTATION

Soit la fonction $f(r) = \psi(1+r, 1)$ $\forall r \in [-1, 1]$.

- f est continuellement dérivable sur tout l'ensemble de définition (conséquence I-6)
- f vérifie la relation :

$$f(r) + f(-r) = 2 \quad \forall r \in [-1, 1] \quad (\text{conséquence I-7}).$$

Lemme I-1

La dérivée de la fonction f vérifie :

$$f'(r) = f' \left(\frac{r}{2} \right) \quad \forall r \in [-1, 1]$$

DEMONSTRATION

Dérivons la relation

$$f(r) + f(-r) = 2 \quad \forall r \in [-1, 1].$$

On obtient

$$f'(r) = f'(-r) \quad \forall r \in [-1, 1].$$

Dans la conséquence I-7, nous avons démontré la relation :

$$3 \psi(1+r, 1-r) = \psi(1-r, 1-r) + 2 \psi(1-r, 1+2r) \quad \forall r \in [0, 1]$$

Or, $\psi(u, v) = \psi(u, 1) \psi(1, v)$ (propriété I-2)

donc $3 \psi(1+r, 1) \psi(1, 1-r) = \psi(1-r, 1) \psi(1, 1-r) + 2 \psi(1-r, 1) \psi(1, 1+2r) \quad \forall r \in [0, 1]$

ou encore

$$3 f(r) f(-r) = f^2(-r) + 2 f(-r) f(2r) \quad \forall r \in [0, 1]$$

Puisque f est une fonction croissante (conséquence I-6) positive ou nulle, $f(-r)$ ne peut s'annuler que pour $r = 1$.

Donc $3 f(r) = f(-r) + 2 f(2r) \quad \forall r \in [0, 1[$

Or, $f(r) + f(-r) = 2 \quad \forall r \in [-1, 1]$

donc $2 f(r) = 1 + f(2r) \quad \forall r \in [0, 1[$

Dérivons cette dernière relation :

$$f'(r) = f'(2r) \quad \forall r \in [0, 1[$$

Or $f'(r) = f'(-r) \quad \forall r \in [-1, 1]$

et $f'(2r) = f'(-2r) \quad \forall r \in [-\frac{1}{2}, \frac{1}{2}]$

Donc $f'(-r) = f'(-2r) \quad \forall r \in [0, \frac{1}{2}]$

et donc $f'(r) = f'(2r) \quad \forall r \in [-\frac{1}{2}, 1[$

ou encore $f'(r') = f'(\frac{r'}{2}) \quad \forall r' \in [-1, 1]$

DEMONSTRATION DU THEOREME I-1

Appliquons le théorème des accroissements finis à la fonction f :

Soit $r \in [-1,1]$, $\forall \theta$ et $\theta' \in]0,1[$ tel que

$$\textcircled{1} \quad \begin{cases} f(r) = f(0) + r f'(\theta r) \\ f(-r) = f(0) - r f'(-\theta' r) \end{cases}$$

Additionnons membre à membre ces deux égalités en remarquant que $f(0) = \psi(1,1) = 1$

$$f(r) + f(-r) = 2 + r [f'(\theta r) - f'(-\theta' r)]$$

Comme $f(r) + f(-r) = 2 \quad \forall r \in [-1,1]$

Alors $r [f'(\theta r) - f'(-\theta' r)] = 0$

- $r \neq 0 \Rightarrow f'(\theta r) = f'(-\theta' r)$

- $r \rightarrow 0 \Rightarrow f'(\theta r) \rightarrow f'(0)$

et $f'(-\theta' r) \rightarrow f'(0)$ puisque f' est continue (conséquence I-6).

Donc $f'(\theta r) = f'(-\theta' r) \quad \forall r \in [-1,1]$

Or, dans le lemme I-1, nous avons démontré que

$$f'(r) = f'\left(\frac{r}{2}\right) \quad \forall r \in [-1,1]$$

Donc $f'(\theta r) = f'\left(\frac{\theta r}{2}\right) = \dots = f'\left(\frac{\theta r}{2^n}\right) = \dots \quad \forall n \in \mathbb{N}$

Or, f' est continue

Donc $\lim_{n \rightarrow \infty} f'\left(\frac{\theta r}{2^n}\right) = f'(0)$

et donc $f'(\theta r) = f'(-\theta' r) = f'(0)$

Reportons ce résultat dans les relations $\textcircled{1}$

$$\begin{cases} f(r) = 1 + r f'(0) \\ f(-r) = 1 - r f'(0) \end{cases}$$

Or $g(r) = \psi(1+r, 1-r) = f(r) f(-r) \quad \forall r \in [-1,1]$ (propriété I-2)

$$\text{Donc } g(r) = 1 - r^2 f'^2(0)$$

$$f'(0) > 0 \quad (\text{conséquence I-6})$$

$$\text{et } g(1) \geq 0 \quad (\text{propriété I-3})$$

$$\text{donc } f'^2(0) \in]0,1]$$

$$\text{Posons } f'(0) = \alpha$$

$$\text{alors } g(r) = 1 - \alpha^2 r^2 \quad r \in [-1,1], \alpha \in]0,1]$$

$$\text{Si } R = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix} \quad r \in [-1,1] \quad \text{la liaison-information est}$$

$$I(R) = \psi(1+r, 1-r) = g(r).$$

Il est aisé de vérifier que $I(R)$ vérifie les quatre propriétés I-1, I-2, I-3 et I-4.

De plus, cette fonction g est unique puisque

$$f'(\theta r) = f'(-\theta' r) = f'(0) \quad \forall r \in [-1,1] \quad \theta \text{ et } \theta' \in]0,1]$$

$$\text{et que } g(r) = f(r) f(-r) = 1 - r^2 f'(\theta r) f'(-\theta' r).$$

REMARQUE

$$\text{Soit } R_1 \in R_2, \quad R_1 = \begin{pmatrix} 1 & \alpha \\ \alpha & 1 \end{pmatrix}$$

$$I(R_1) = g(1) = 1 - \alpha^2$$

donc α^2 représente le complémentaire à 1 de l'information contenue dans la matricé R_1 .

2. Corollaire I-1 du théorème I-1

La liaison-information d'une matrice $R = ((r_{ij})) \in R_n$ a pour expression :

$$I(R) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n (1 - \alpha^2 r_{ij}^2) \quad \begin{array}{l} \alpha \in]0,1[\\ r_{ij} \in [-1,1] \\ \forall i,j = 1,2,\dots,n \end{array}$$

DEMONSTRATION

La conséquence I-2 :

$$I(R) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \psi(1-r_{ij}, 1+r_{ij}) \quad \forall R \in R_n$$

et le théorème I-1 :

$$\psi(1+r, 1-r) = 1 - \alpha^2 r^2 \quad \forall \alpha \in]0,1], r \in [-1,1]$$

nous permettent d'écrire :

$$I(R) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n (1 - \alpha^2 r_{ij}^2) \quad \alpha \in]0,1]$$

REMARQUE

Le théorème I-1 nous a permis de préciser la fonction $\psi(u,v)$ si $(u,v) \in [0,2] \times [0,2]$ où $u = 1+r$ et $v = 1-r$.

$$\psi(u,v) = 1 - \alpha^2(u-1)(1-v) = 1 - \alpha^2 + \alpha^2 uv.$$

Il nous reste à la déterminer si $(u,v) \in [2,\infty] \times [2,\infty]$.

3. Théorème I-2

Soit $R \in R_n$ et soit $D \in D_n$ la matrice diagonale semblable à R .

$$R = ((r_{ij})) \quad D = ((\lambda_i)) \quad i, j = 1, 2, \dots, n.$$

$$\text{alors} \quad \sum_{i=1}^{n-1} \sum_{j=i+1}^n \det(R_{ij}^j) = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \lambda_i \lambda_j.$$

DEMONSTRATION

Nous allons démontrer ce théorème par récurrence.

1) Le théorème est vrai si $n = 2$, $n = 3$; en effet

- $n = 2$

$$R = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix} \quad D = \begin{pmatrix} 1-r & 0 \\ 0 & 1+r \end{pmatrix}$$

$$\det(R) = 1 - r^2 = \det(D).$$

- $n = 3$

$$\text{Soient } R = \begin{pmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{pmatrix} \quad \text{et } D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

$$\det(R - \lambda I) = -\lambda^3 + 3\lambda^2 + \lambda(a^2 + b^2 + c^2 - 3) + 2abc - b^2 - c^2 - a^2 + 1$$

$$\det(R - \lambda I) = 0 = -(\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3)$$

et par identification, on obtient :

$$\begin{cases} \lambda_1 \lambda_2 \lambda_3 = 1 + 2abc - b^2 - c^2 - a^2 \\ \lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3 = 3 - (a^2 + b^2 + c^2) \\ \lambda_1 + \lambda_2 + \lambda_3 = 3 \end{cases}$$

donc dans le cas $n = 3$

$$\lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3 = \det(R_1^2) + \det(R_1^3) + \det(R_2^3)$$

2) On suppose le théorème vrai si R est de dimension $(n-1 \times n-1)$,
démontrons-le pour R de dimension $(n \times n)$.

Soient $R \in R_{n-1}$ $R = ((r_{ij}))$ $i, j = 1, 2, \dots, n-1$

et D semblable à R $D = ((\lambda_i))$ $i = 1, 2, \dots, n-1$

$$\text{alors } \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \det(R_{ij}^j) = \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \lambda_i \lambda_j.$$

Considérons maintenant la matrice suivante :

$$R' = \begin{pmatrix} 1 & a_{12} & \dots & a_{1n} \\ a_{21} & & & \\ \vdots & & & \\ \vdots & & & \\ a_{n,1} & & & \end{pmatrix} \begin{matrix} \left(\begin{matrix} R \end{matrix} \right) \end{matrix}$$

où R est la matrice définie précédemment

et soit $D' = ((\lambda'_i))$ $i = 1, 2, \dots, n$ la matrice semblable à R' .

Nous allons chercher dans le polynôme caractéristique de R' le coefficient de λ^{n-2} .

Développons $\det(R' - \lambda I)$ par rapport à la première colonne.

$$\det(R' - \lambda I) = (1 - \lambda) \det(R - \lambda I) + \sum_{j=2}^n (-1)^{1+j} a_{j,1} A_{j,1}$$

où $A_{j,1}$ est le mineur de $a_{j,1}$ soit

$$A_{j,1} = \begin{vmatrix} a_{12} & a_{13} & \dots & a_{1,j-1} & a_{1j} & \dots & a_{1n} \\ 1-\lambda & r_{12} & \dots & \vdots & \vdots & \dots & r_{1,n-1} \\ r_{21} & 1-\lambda & \dots & \vdots & \vdots & \dots & \\ \dots & \dots & \dots & \vdots & \vdots & \dots & \\ r_{j-2,1} & \dots & \dots & 1-\lambda & r_{j-2,j-1} & \dots & \\ r_{j,1} & \dots & \dots & r_{j,j-1} & 1-\lambda & \dots & \\ \dots & \dots & \dots & \dots & \dots & \dots & \\ r_{n-1,1} & \dots & \dots & \dots & \dots & \dots & 1-\lambda \end{vmatrix}$$

Le coefficient de λ^{n-3} dans $\det(R - \lambda I)$ est

$$(-1)^{n-3} \left[\frac{(n-1)(n-2)}{2} - \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} r_{ij}^2 \right]$$

et celui de λ^{n-2} est $(-1)^{n-2} \times (n-1)$.

Donc le coefficient de λ^{n-2} dans $\det(R' - \lambda I)$ est

$$(-1)^{n-2} \left[(n-1) + \frac{(n-1)(n-2)}{2} - \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} r_{ij}^2 + \sum_{j=2}^n (-1)^{1+j} (-1)^{I(i_1, i_2, \dots, i_{n-1})} a_{1j}^2 \right]$$

où $(-1)^{I(i_1, i_2, \dots, i_{n-1})}$ est la signature de la permutation

$$i_1 = 2 ; i_2 = 3 ; \dots ; i_{j-2} = j-1 ; i_{j-1} = 1 ; i_j = j ; \dots ; i_{n-1} = n-1 ;$$

donc $(-1)^{I(i_1, i_2, \dots, i_{n-1})} = (-1)^{j-2}$

d'où le coefficient de λ^{n-2} est :

$$(-1)^{n-2} \left[\frac{n(n-1)}{2} - \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} r_{ij}^2 - \sum_{j=2}^n a_{1j}^2 \right] \quad (1)$$

Comme $\lambda_1^i, \lambda_2^i, \dots, \lambda_n^i$ sont les valeurs propres de R^i

$$\det(R^i - \lambda I) = (-1)^n (\lambda - \lambda_1^i) (\lambda - \lambda_2^i) \dots (\lambda - \lambda_n^i)$$

donc le coefficient de λ^{n-2} est aussi

$$(-1)^n \sum_{i=1}^{n-1} \sum_{j=i+1}^n \lambda_i^i \lambda_j^i \quad (2)$$

Et en égalant (1) et (2) nous obtenons le résultat cherché

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^n \det(R_i^{ij}) = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \lambda_i^i \lambda_j^i$$

4. Corollaire I-2 du théorème I-2

La fonction $\psi(u, v)$ pourra prendre comme valeur :

$$1 - \alpha^2 + \alpha^2 uv \quad \alpha \in]0, 1] \quad (u, v) \in \mathbb{R}_0^+ \times \mathbb{R}_0^+$$

DEMONSTRATION

Soit $R = ((r_{ij})) \in R_n$ et $D = ((\lambda_i)) \in D_n$ semblable à R

$$\lambda_i \in [0, n] \quad i = 1, 2, \dots, n$$

Dans le corollaire I-1, nous avons donné la valeur de

$$I(R) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n (1 - \alpha^2 r_{ij}^2) \quad \alpha \in]0, 1]$$

ou encore

$$I(R) = 1 - \alpha^2 + \frac{2\alpha^2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n (1 - r_{ij}^2)$$

$$\text{Comme } \sum_{i=1}^{n-1} \sum_{j=i+1}^n (1 - r_{ij}^2) = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \lambda_i \lambda_j \quad (\text{théorème I-2})$$

on obtient

$$I(R) = 1 - \alpha^2 + \frac{2\alpha^2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \lambda_i \lambda_j = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n [1 - \alpha^2 + \alpha^2 \lambda_i \lambda_j]$$

$$\text{Or } I(R) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \psi(\lambda_i, \lambda_j) \quad (\text{conséquence I-2})$$

donc on pourra prendre comme fonction ψ

$$\psi(\lambda_i, \lambda_j) = 1 - \alpha^2 + \alpha^2 \lambda_i \lambda_j \quad \lambda_i, \lambda_j \in [0, n]$$

REMARQUE

Nous avons déjà vu que α^2 représente le complémentaire à 1 de l'information contenue dans la matrice :

$$R_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$\alpha^2 = 1 - I(R_1).$$

De plus, $I(R_1)$ représente la valeur minimale de l'information de toute matrice de corrélation d'ordre 2.

$$I(R_1) = \min_{R \in R_2} I(R)$$

On pourra donc prendre comme minimum sans restriction de généralité

$$I(R_1) = 0 \quad \text{ou encore} \quad \alpha^2 = 1$$

$$\text{donc } \psi(u, v) = uv \quad \forall (u, v) \in \mathbb{R}_0^+ \times \mathbb{R}_0^+$$

5. Conséquence I-8

La liaison-information d'une matrice $C \in C_n$ dont les éléments diagonaux sont

$$\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2 \quad \text{est :}$$

$$I(C) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \sigma_i^2 \sigma_j^2 (1-r_{ij}^2)$$

DEMONSTRATION

Soit $C \in C_n$ et $R \in R_n$ déduite de C .

$$I(C) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \psi(\sigma_i^2, \sigma_j^2) I(R_i^j) \quad (\text{propriétés I-1, I-2})$$

Or $\psi(u, v) = uv$ d'après la remarque précédente et

$$I(R_i^j) = \psi(1-r_{ij}, 1+r_{ij}) = 1-r_{ij}^2$$

$$\text{donc } I(C) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \sigma_i^2 \sigma_j^2 (1-r_{ij}^2) \quad \forall C \in C_n$$

REMARQUE

$$r_{ij} = \text{covariance}(Y_i, Y_j) / \sigma_i \sigma_j$$

$$\text{et } \sigma_i^2 \sigma_j^2 (1-r_{ij}^2) = \sigma_i^2 \sigma_j^2 - \text{covariance}(Y_i, Y_j)^2$$

donc

$$I(C) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \det C_i^j \quad \forall C \in C_n$$

CHAPITRE II

INDEPENDANCE DES PROPRIETES ENONCEES
AU CHAPITRE I

INTRODUCTION

Dans le chapitre précédent, nous avons imposé à la fonction $I \in \mathcal{F}(C \rightarrow \mathbb{R}_0^+)$ de vérifier quatre propriétés. Il nous reste à nous assurer qu'aucune d'entre elles n'est déductible des trois autres. Afin d'y parvenir nous montrerons que tous les systèmes de trois propriétés, déduits du système de départ par la suppression d'une quelconque des propriétés, nous conduisent à au moins deux fonctions de

$$\mathcal{F}(C \rightarrow \mathbb{R}_0^+).$$

REMARQUE

La suppression de la propriété, I-4 nous conduit à une infinité de fonctions se déduisant toutes de celle définie au chapitre I par une transformation affine.

I - SUPPRESSION DE LA PROPRIETE I-1

Soit $C \in \mathcal{C}$ et prenons comme fonction

$$I \in \mathcal{F}(C \rightarrow \mathbb{R}_0^+) \text{ la fonction } I(C) = \det(C).$$

Cette fonction vérifie les propriétés I-2, I-3, I-4.

1) Si $C \in \mathcal{C}$ est unitaire alors $I(C) = \det(C) = 1$.

2) Soit $C \in C_2$ d'éléments diagonaux σ_1^2, σ_2^2 et soit R la matrice déduite de C .

$$\text{Alors } \det C = \sigma_1^2 \sigma_2^2 \det R.$$

$$\text{Donc } \psi(u,v) = uv.$$

Cette fonction ψ vérifie évidemment toutes les conditions analytiques imposées dans la propriété I-2.

3) Soit $R \in R$ et $D \in D$ semblable à R alors

$$\det R = \det D$$

$$\text{donc } I(R) = I(D).$$

4) La propriété I-1 par contre n'est pas vérifiée en général.

Soit $D = ((\lambda_i)) \in D_n$

$$I(D) = h(n) \sum_{i=1}^{n-1} \sum_{j=i+1}^n \psi(\lambda_i, \lambda_j) I\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) = h(n) \sum_{i=1}^{n-1} \sum_{j=i+1}^n \lambda_i \lambda_j \quad (1)$$

d'autre part

$$I(D) = \det D = \lambda_1 \lambda_2 \lambda_3 \dots \lambda_n \quad (2)$$

Comme en général l'égalité entre (1) et (2) n'est pas réalisée, alors la propriété I-1 n'est pas vérifiée.

II - SUPPRESSION DE LA PROPRIÉTÉ I-2

Pour cette propriété, nous procéderons en deux étapes. D'abord suppression de

$$\psi(u,v) = \psi(u,1) \psi(1,v) \quad \forall (u,v) \in \mathbb{R}_0^+ \times \mathbb{R}_0^+$$

puis suppression de la propriété I-2 dans son ensemble.

1. Suppression de $\psi(u,v) = \psi(u,1) \psi(1,v) \quad \forall (u,v) \in \mathbb{R}_0^+ \times \mathbb{R}_0^+$

Soit $C \in C_n$ d'éléments diagonaux $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$

Prenons comme $I \in \mathcal{F}(C \rightarrow \mathbb{R}_0^+)$ la fonction suivante :

$$I(C) = \frac{1}{n} \sum_{i=1}^n \sigma_i^2$$

Dans ce cas $\psi(u,v) = \frac{u+v}{2} \quad \forall (u,v) \in \mathbb{R}_0^+ \times \mathbb{R}_0^+$

- Si $C \in C$ est unitaire alors $I(C) = 1$

$$- I(C) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n I(C_{ij}^j)$$

En effet

$$\text{si } C \in R \quad I(C) = I(C_{ij}^j) = 1$$

$$\text{si } C \in C-R \quad I(C_{ij}^j) = \psi(\sigma_i^2, \sigma_j^2) I(R_{ij}^j) = \frac{\sigma_i^2 + \sigma_j^2}{2}$$

donc

$$I(C) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{\sigma_i^2 + \sigma_j^2}{2} = \frac{1}{n} \sum_{i=1}^n \sigma_i^2$$

- $\psi(u,v) = \frac{u+v}{2} \quad \forall (u,v) \in \mathbb{R}_0^+ \times \mathbb{R}_0^+$ vérifie les conditions analytiques imposées à ψ sauf

$$\psi(u,v) = \psi(u,1) \psi(1,v) \text{ qui n'est vérifiée que pour } u=1 \text{ ou } v=1.$$

- Soit $R \in R_n$ et $D \in D_n$ semblable à R .

$$D = ((\lambda_i)) \text{ et } \sum_{i=1}^n \lambda_i = n$$

donc $I(R) = I(D) = 1$.

2. Suppression de la propriété I-2

Soit $C \in \mathcal{C}$ alors $I(C) = 1$.

$$I \in \mathcal{F}(C \rightarrow \mathbb{R}_0^+)$$

Les propriétés I-1, I-3, I-4 sont évidemment vérifiées et la propriété I-2 est devenue inutile.

On peut cependant définir la fonction ψ telle que

$$I(C) = \psi(\sigma_1^2, \sigma_2^2) I(R) \quad C \in \mathcal{C}_2 \text{ et } R \text{ déduite de } C.$$

ψ sera la fonction constante

$$\psi(u, v) = 1 \quad \forall (u, v) \in \mathbb{R}_0^+ \times \mathbb{R}_0^+.$$

$\psi(u, v) = 1$ vérifie la condition

$$\psi(u, v) = \psi(u, 1) \psi(1, v) \quad \forall (u, v) \in \mathbb{R}_0^+ \times \mathbb{R}_0^+$$

mais ne vérifie pas les conditions de croissance imposées à ψ dans la propriété I-2.

REMARQUE

Nous venons de montrer que deux des conditions imposées à ψ ne sont pas déductibles des autres propriétés. Il en est de même évidemment pour la condition de dérivabilité. La propriété I-2 ne peut donc être déduite des autres propriétés.

III - SUPPRESSION DE LA PROPRIETE I-3

Soit $C \in \mathcal{C}$ et prenons comme fonction

$$I \in \mathcal{F}(C \rightarrow \mathbb{R}_0^+) \text{ la fonction}$$

$$I(C) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \sigma_i^4 \sigma_j^4 (1-r_{ij}^2)^2$$

Dans ce cas $\psi(u,v) = u^2 v^2 \quad \forall (u,v) \in \mathbb{R}_0^+ \times \mathbb{R}_0^+$

Cette fonction vérifie les propriétés I-1, I-2, I-4.

1 - La propriété I-1 est vérifiée par définition de I.

2 - Toutes les conditions imposées à ψ sont vérifiées en effet $\psi(u,v) = u^2 v^2$ est continuellement dérivable, croissante en u et v , et décomposable en $\psi(u,1) \psi(1,v) = (u^2 \times 1)(1 \times v^2) = u^2 v^2$.

3 - Si C est unitaire $I(C) = 1$.

4 - La propriété I-3 n'est pas satisfaite puisque

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^n (1-r_{ij}^2)^2 \quad \text{est en général différent de}$$

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^n (\lambda_i \lambda_j)^2$$

Prenons par exemple les matrices semblables suivantes :

$$R = \begin{pmatrix} 1 & r & 0 \\ r & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et } D = \begin{pmatrix} 1-r & 0 & 0 \\ 0 & 1+r & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$I(R) = \frac{2}{6} [(1-r^2)^2 + 2] = \frac{1}{3} (3 - 2r^2 + r^4)$$

$$I(D) = \frac{1}{3} [(1-r^2)^2 + (1-r)^2 + (1+r)^2] = \frac{1}{3} (3 + r^4)$$

donc si $r \neq 0$ $I(R) \neq I(D)$.

IV - CONCLUSION

Nous pouvons donc conclure qu'aucune des quatre propriétés du Chapitre I n'est déductible des trois autres puisqu'il nous est toujours possible de trouver au moins une fonction, différente de celle définie au Chapitre I, vérifiant trois des quatre propriétés I-1, I-2, I-3, I-4.

CHAPITRE III

LIAISON-INFORMATION CONDITIONNELLE

INTRODUCTION

Dans ce chapitre, nous allons nous intéresser aux applications $\mathcal{F}(C \rightarrow \mathbb{R}_0^+)$ qui permettront de mesurer la liaison existant entre des variables aléatoires Y_1, Y_2, \dots, Y_n et une variable aléatoire X . Nous appellerons ces applications liaisons-informations conditionnelles, et nous les noterons $I_X(C)$ si $C = \text{COV}(X, Y_1, Y_2, \dots, Y_n)$. Dans tout ce chapitre, X sera une variable aléatoire centrée et réduite.

I - DEFINITIONS

Soient $n+1$ variables aléatoires X, Y_1, Y_2, \dots, Y_n , et soit $C = \text{COV}(X, Y_1, Y_2, \dots, Y_n)$.

Les deux définitions que nous allons proposer se déduisent immédiatement de la liaison-information définie dans le chapitre I.

1. DEFINITION III-1

$$I_X^{(1)}(C) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \det C_{oij}$$

où $C_{oij} = \text{COV}(X, Y_i, Y_j)$

2. DEFINITION III-2

$$I_X^{(2)}(C) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \sigma_i^2 \sigma_j^2 (1 - \rho_{o/ij}^2)$$

où $\rho_{o/ij}$ représente le coefficient de corrélation multiple de X par rapport à Y_i et Y_j . (L'indice zéro correspond à la variable X).

II - PROPRIETES

1. Propriété III-1

Si $C_{oij} = \text{COV}(X, Y_i, Y_j)$ et si $\rho_{o/ij}$ est le coefficient de corrélation multiple de X par rapport à Y_i et Y_j alors

$$\det C_{oij} = (\det C_i^j) (1 - \rho_{o/ij}^2)$$

DEMONSTRATION

Cette propriété est évidente puisqu'elle représente une des définitions du coefficient de corrélation multiple.

REMARQUE

$\det C_{oij} \geq 0$ puisque $\det C_{ij}$ et $1 - \rho_{o/ij}^2$ sont deux quantités positives ou nulles.

2. Propriété III-2

Si $C \in \mathcal{C}$ alors $I_X^{(1)}(C) \leq I(C)$

DEMONSTRATION

$$I_X^{(1)}(C) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \det C_{oij}$$

$$I(C) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \det C_i^j$$

Soit $\Delta I^{(1)} = I(C) - I_X^{(1)}(C)$

$$\Delta I^{(1)} = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n (\det C_i^j - \det C_{oij})$$

or $\det C_{oij} = \det C_i^j (1 - \rho_{o/ij}^2)$ (propriété III-1)

Donc $\det C_i^j - \det C_{oij} = \det C_i^j \times \rho_{o/ij}^2$

et $\Delta I^{(1)} \geq 0$ car $\det C_i^j \times \rho_{o/ij}^2 \geq 0$.

REMARQUE

Le signe de $\Delta I^{(2)} = I(C) - I_X^{(2)}(C)$ est imprévisible puisque

$$\Delta I^{(2)} = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n (\rho_{o/ij}^2 - r_{ij}^2)$$

III - CONCLUSION

Deux définitions ont été proposées dans ce chapitre pour la liaison-information conditionnelle. Il semblerait que la première soit la meilleure en vertu de la propriété II-2. Cependant, nous les conserverons toutes les deux pour les applications pratiques et nous déciderons ensuite celle qui paraîtra le mieux mesurer la liaison entre X et $Y_1 Y_2 \dots Y_n$.

CHAPITRE IV

APPLICATIONS PRATIQUES

INTRODUCTION

La première application consiste à "classer" des variables aléatoires corrélées entre elles à l'aide de l'analyse factorielle et de calculer la liaison-information apportée par chacune d'elles.

La seconde concerne une application de la liaison-information conditionnelle à un problème de segmentation.

I - APPLICATION 1

1. Analyse factorielle

L'analyse factorielle a pour but de figurer géométriquement dans un espace euclidien de faible dimension un ensemble d'informations qui se présentent sous la forme de variables aléatoires corrélées entre elles [2] ; [8].

Soient n variables aléatoires

$$Y_1, Y_2, \dots, Y_n$$

et soit $R \in R_n$ leur matrice de corrélation

$$R = ((r_{kl}))$$

Si nous diagonalisons la matrice R , nous obtenons n valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n$ et leurs vecteurs propres correspondants V_1, V_2, \dots, V_n .

Géométriquement, R est la matrice d'inertie du nuage de points constitué dans \mathbb{R}^n à partir des n variables aléatoires, les vecteurs propres sont les axes principaux d'inertie et les valeurs propres les moments principaux d'inertie.

Statistiquement λ_i est la variance de la variable aléatoire V_i définie comme une combinaison linéaire des variables aléatoires Y_1, Y_2, \dots, Y_n .

Soit $D = ((\lambda_i))$ la matrice diagonale semblable à R et soit

$$V_i = a_{i1} Y_1 + a_{i2} Y_2 + \dots + a_{in} Y_n \quad \forall i = 1, 2, \dots, n$$

le vecteur propre associé à λ_i .

Le scalaire a_{ik} représente "l'influence" de Y_k sur l'axe factoriel V_i . Il nous est donc possible de classer les Y_k par ordre d'influence croissante pour la variable V_i ayant la plus grande variance λ_{Max} .

2. Liaison-information apportée par une variable

Etant donné n variables aléatoires Y_1, Y_2, \dots, Y_n et leur matrice de corrélation R , nous appellerons liaison-information apportée par la variable Y_i (dans R) la quantité

$$I [Y_i(R)] = \frac{1}{(n-1)} \sum_{\substack{j=1 \\ j \neq i}}^n \det R_i^j$$

REMARQUES

- Si $R = ((r_{ij}))$ alors

$$I [Y_i(R)] = 0 \quad \text{si} \quad r_{ij}^2 = 1 \quad \forall j = 1, 2, \dots, n \quad j \neq i$$

$$I [Y_i(R)] = 1 \quad \text{si} \quad r_{ij}^2 = 0 \quad \forall j \neq i = 1, 2, \dots, n.$$

- La liaison-information de la matrice R est la valeur moyenne des liaisons-informations apportées par chacune des variables :

$$I(R) = \frac{1}{n} \sum_{i=1}^n I[(Y_i(R))]$$

3. Applications pratiques

Dans le paragraphe 1, nous avons vu comment il était possible de classer les Y_k par ordre d'influence croissante. Nous pouvons maintenant comparer cet ordre avec celui que l'on obtiendrait en classant par ordre décroissant les liaisons-informations apportées par toutes les variables Y_i .

Intuitivement, ces deux classements devraient avoir des similitudes. La propriété suivante va d'ailleurs établir que pour $n = 3$ ces deux classements sont identiques.

Propriété IV-1

Soit $R = ((r_{ij}))$ la matrice de corrélation des variables aléatoires Y_1, Y_2, Y_3 et soit $D = ((\lambda_i)) \quad i=1,2,3$ la matrice diagonale semblable à R.

A la valeur propre maximale $\lambda_M = \text{Max}(\lambda_1, \lambda_2, \lambda_3)$ correspond

le vecteur propre $V_M \begin{vmatrix} x \\ y \\ z \end{vmatrix}$

Alors la condition nécessaire et suffisante pour que $x^2 > y^2 > z^2$ est que

$$I [Y_1(R)] < I [Y_2(R)] < I [Y_3(R)]$$

Avant d'établir la propriété IV-1, nous allons démontrer la propriété suivante :

Propriété IV-2

Un minorant de $(1-\lambda_M)^2$ est le complémentaire à 1 de la liaison-information de la matrice R :

$$(1-\lambda_M)^2 > 1-I(R) = \frac{1}{3} [r_{12}^2 + r_{13}^2 + r_{23}^2]$$

DEMONSTRATION DE LA PROPRIETE IV-2

Soient $R = \begin{pmatrix} 1 & r_{12} & r_{13} \\ r_{12} & 1 & r_{23} \\ r_{13} & r_{23} & 1 \end{pmatrix}$ et $D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$

$$P(\lambda) = \det(R-\lambda I) = (1-\lambda)^3 - (r_{12}^2 + r_{13}^2 + r_{23}^2)(1-\lambda) + 2r_{12} r_{13} r_{23}$$

On a de plus $\lambda_M = \text{Max}(\lambda_1, \lambda_2, \lambda_3)$

$$\text{et } P(\lambda_1) = P(\lambda_2) = P(\lambda_3) = 0$$

Pour plus de commodité, nous choisirons

$$\lambda_1 > \lambda_2 > \lambda_3$$

$$\text{Soit } Y = P(1-X) = X^3 - (r_{12}^2 + r_{13}^2 + r_{23}^2) X + 2r_{12} r_{13} r_{23}$$

$$\text{et soit } \alpha_1 = 1-\lambda_1, \alpha_2 = 1-\lambda_2, \alpha_3 = 1-\lambda_3$$

$$P(1-\alpha_1) = P(1-\alpha_2) = P(1-\alpha_3) = 0$$

$$\text{et } \alpha_1 < \alpha_2 < \alpha_3$$

$$Y' = 3X^2 - (r_{12}^2 + r_{13}^2 + r_{23}^2)$$

$$\text{et donc } Y' = 0 \text{ si } X = \pm \sqrt{\frac{r_{12}^2 + r_{13}^2 + r_{23}^2}{3}} = \pm b$$

X	$-\infty$	α_1	$-b$	α_2	b	α_3	$+\infty$
Y'		+	0	-	0	+	
Y	$-\infty$						$+\infty$

Comme Y possède 3 racines

α_1 se place entre $-\infty$ et $-b$

α_2 se place entre $-b$ et $+b$

α_3 se place entre b et $+\infty$

donc $\alpha_1 < -b$

$$\text{donc } (1-\lambda_1) < -\sqrt{\frac{r_{12}^2+r_{13}^2+r_{23}^2}{3}}$$

$$\text{ou encore } (1-\lambda_M)^2 > \frac{r_{12}^2+r_{13}^2+r_{23}^2}{3} = 1-I(R)$$

DEMONSTRATION DE LA PROPRIETE IV-1

Exprimons x, y, z en fonction de λ_1 (supposé maximal), r_{12}, r_{13}, r_{23}

$$\begin{cases} (1-\lambda_1) x + r_{12} y + r_{13} z = 0 \\ r_{12} x + (1-\lambda_1) y + r_{23} z = 0 \\ r_{13} x + r_{23} y + (1-\lambda_1) z = 0 \end{cases}$$

$$x = - \frac{\begin{vmatrix} r_{13} & r_{12} \\ r_{23} & 1-\lambda_1 \end{vmatrix}}{\begin{vmatrix} 1-\lambda_1 & r_{12} \\ r_{12} & 1-\lambda_1 \end{vmatrix}} z$$

$$y = - \frac{\begin{vmatrix} 1-\lambda_1 & r_{13} \\ r_{12} & r_{23} \end{vmatrix}}{\begin{vmatrix} 1-\lambda_1 & r_{12} \\ r_{12} & 1-\lambda_1 \end{vmatrix}} z$$

$$x = - \frac{\begin{vmatrix} r_{12} & r_{13} \\ r_{23} & 1-\lambda_1 \end{vmatrix}}{\begin{vmatrix} 1-\lambda_1 & r_{13} \\ r_{13} & 1-\lambda_1 \end{vmatrix}} y$$

$$\text{Soit encore} \quad x = - \frac{(1-\lambda_1)r_{13} - r_{12}r_{23}}{(1-\lambda_1)^2 - r_{12}^2} z \quad \textcircled{a}$$

$$y = - \frac{(1-\lambda_1)r_{23} - r_{12}r_{13}}{(1-\lambda_1)^2 - r_{12}^2} z \quad \textcircled{b}$$

$$x = - \frac{r_{12}(1-\lambda_1) - r_{13}r_{23}}{(1-\lambda_1)^2 - r_{13}^2} y \quad \textcircled{c}$$

DEMONSTRATION 1

Si $x^2 > y^2 > z^2$ alors $I[Y_1(R)] < I[Y_2(R)] < I[Y_3(R)]$?

$$\textcircled{a} \quad \Leftrightarrow (1-\lambda_1)^2 r_{13}^2 + r_{12}^2 r_{23}^2 - 2r_{12}r_{13}r_{23}(1-\lambda_1) - (1-\lambda_1)^4 - r_{12}^4 + 2r_{12}^2(1-\lambda_1)^2 > 0$$

$$\textcircled{b} \quad \Leftrightarrow (1-\lambda_1)^2 r_{23}^2 + r_{12}^2 r_{13}^2 - 2r_{12}r_{13}r_{23}(1-\lambda_1) - (1-\lambda_1)^4 - r_{12}^4 + 2r_{12}^2(1-\lambda_1)^2 > 0$$

$$\textcircled{c} \quad \Leftrightarrow (1-\lambda_1)^2 r_{12}^2 + r_{13}^2 r_{23}^2 - 2r_{12}r_{13}r_{23}(1-\lambda_1) - (1-\lambda_1)^4 - r_{13}^4 + 2r_{13}^2(1-\lambda_1)^2 > 0$$

$$\text{Or } P(\lambda_1) = 0 \Rightarrow -2r_{12}r_{13}r_{23} - (1-\lambda_1)^3 = -(r_{12}^2 + r_{13}^2 + r_{23}^2)(1-\lambda_1)$$

donc

$$\textcircled{a} \quad \Leftrightarrow (r_{12}^2 - r_{23}^2)(1-\lambda_1)^2 + r_{12}^2(r_{23}^2 - r_{12}^2) > 0$$

$$\textcircled{b} \Leftrightarrow (r_{12}^2 - r_{13}^2) (1 - \lambda_1)^2 + r_{12}^2 (r_{13}^2 - r_{12}^2) > 0$$

$$\textcircled{c} \Leftrightarrow (r_{13}^2 - r_{23}^2) (1 - \lambda_1)^2 + r_{13}^2 (r_{23}^2 - r_{13}^2) > 0$$

et donc

$$\textcircled{a} \Leftrightarrow \textcircled{1} \begin{cases} (1 - \lambda_1)^2 > r_{12}^2 \\ r_{12}^2 > r_{23}^2 \end{cases} \quad \text{ou} \quad \textcircled{1'} \begin{cases} (1 - \lambda_1)^2 < r_{12}^2 \\ r_{12}^2 < r_{23}^2 \end{cases}$$

$$\textcircled{b} \Leftrightarrow \textcircled{2} \begin{cases} (1 - \lambda_1)^2 > r_{12}^2 \\ r_{12}^2 > r_{13}^2 \end{cases} \quad \text{ou} \quad \textcircled{2'} \begin{cases} (1 - \lambda_1)^2 < r_{12}^2 \\ r_{12}^2 < r_{13}^2 \end{cases}$$

$$\textcircled{c} \Leftrightarrow \textcircled{3} \begin{cases} (1 - \lambda_1)^2 > r_{13}^2 \\ r_{13}^2 > r_{23}^2 \end{cases} \quad \text{ou} \quad \textcircled{3'} \begin{cases} (1 - \lambda_1)^2 < r_{13}^2 \\ r_{13}^2 < r_{23}^2 \end{cases}$$

$$\textcircled{1} + \textcircled{2} + \textcircled{3} \Rightarrow (1 - \lambda_1)^2 > r_{12}^2 > r_{13}^2 > r_{23}^2$$

$$\textcircled{1'} + \textcircled{2} + \textcircled{3} \Rightarrow \text{incompatible entre eux}$$

$$\textcircled{1'} + \textcircled{2'} + \textcircled{3} \Rightarrow \text{incompatible entre eux}$$

$$\textcircled{1'} + \textcircled{2'} + \textcircled{3'} \Rightarrow \text{incompatible avec } (1 - \lambda_1)^2 > \frac{r_{12}^2 + r_{13}^2 + r_{23}^2}{3}$$

$$\textcircled{1} + \textcircled{2'} + \textcircled{3} \Rightarrow \text{incompatible entre eux}$$

$$\textcircled{1} + \textcircled{2'} + \textcircled{3'} \Rightarrow \text{incompatible entre eux}$$

$$\textcircled{1} + \textcircled{2} + \textcircled{3'} \Rightarrow \text{incompatible entre eux}$$

$$\textcircled{1'} + \textcircled{2} + \textcircled{3'} \Rightarrow \text{incompatible entre eux}$$

$$\text{donc } x^2 > y^2 > z^2 \Rightarrow r_{12}^2 > r_{13}^2 > r_{23}^2$$

$$\text{et donc } r_{12}^2 + r_{13}^2 > r_{12}^2 + r_{23}^2$$

$$r_{12}^2 + r_{13}^2 > r_{13}^2 + r_{23}^2$$

$$r_{12}^2 + r_{23}^2 > r_{13}^2 + r_{23}^2$$

$$\text{soit } I[Y_1(R)] < I[Y_2(R)] < I[Y_3(R)]$$

DEMONSTRATION 2

Si $I[Y_1(R)] < I[Y_2(R)] < I[Y_3(R)]$ alors $x^2 > y^2 > z^2$?

Par hypothèse $r_{12}^2 > r_{13}^2 > r_{23}^2$

$$\frac{x^2}{z^2} = \frac{(1-\lambda_1)^2 r_{13}^2 + r_{12}^2 r_{23}^2 - 2r_{12} r_{13} r_{23} (1-\lambda_1)}{(1-\lambda_1)^4 + r_{12}^4 - 2r_{12}^2 (1-\lambda_1)^2} = \frac{N_1}{D_1}$$

Calculons $N_1 - D_1$

$$N_1 - D_1 = (1-\lambda_1)^2 r_{13}^2 + r_{12}^2 r_{23}^2 - 2r_{12} r_{13} r_{23} (1-\lambda_1) - (1-\lambda_1)^4 - r_{12}^4 + 2r_{12}^2 (1-\lambda_1)^2$$

$$\text{Or } P(\lambda_1) = 0 \Rightarrow (1-\lambda_1)^3 + 2r_{12} r_{13} r_{23} = (r_{13}^2 + r_{12}^2 + r_{23}^2) (1-\lambda_1)$$

$$\text{Donc } N_1 - D_1 = [(1-\lambda_1)^2 - r_{12}^2] (r_{12}^2 - r_{23}^2)$$

Etudions le signe de $(1-\lambda_1)^2 - r_{12}^2$

$$P(1-r_{12}) = r_{12}^3 - (r_{12}^2 + r_{13}^2 + r_{23}^2) r_{12} + 2r_{12} r_{13} r_{23} = -r_{12} (r_{13} - r_{23})^2$$

$$P(1+\alpha_1) = -\alpha_1^3 + (r_{12}^2 + r_{13}^2 + r_{23}^2) \alpha_1 + 2r_{12} r_{13} r_{23} = 4r_{12} r_{13} r_{23}$$

$$P'(1-r_{12}) = 2r_{12}^2 - r_{13}^2 - r_{23}^2 > 0 \text{ par hypothèse}$$

$$P'(1+\alpha_1) = 3\alpha_1^2 - (r_{12}^2 + r_{13}^2 + r_{23}^2) > 0 \text{ (propriété IV-2)}$$

$$- r_{12} < 0$$

$$P'(1-r_{12}) > 0 \text{ et } P(1-r_{12}) > 0$$

donc

$$\alpha_1 < r_{12} < -b \quad (\text{propriété IV-2})$$

$$\text{et donc } \alpha_1^2 > r_{12}^2$$

$$- r_{12} > 0$$

$$P'(1-r_{12}) > 0 \text{ et } P(1-r_{12}) < 0$$

donc cf propriété IV-2.

$$b < r_{12} < \alpha_3 \quad (\text{propriété IV-2})$$

$$\text{de plus } b < -\alpha_1 < +\infty$$

$$\text{comme } P(1-r_{12}) < P(1+\alpha_1)$$

$$\text{car } -r_{12}(r_{13}-r_{23})^2 < 4r_{12}r_{13}r_{23}$$

$$\text{puisque } r_{12}[r_{13}+r_{23}]^2 > 0$$

$$\text{donc } 0 < b < r_{12} < -\alpha_1$$

$$\text{et donc } \alpha_1^2 > r_{12}^2$$

$$\text{On a donc démontré que } (1-\lambda_1)^2 - r_{12}^2 > 0$$

Comme $r_{12}^2 > r_{23}^2$ par hypothèse

$$N_1 - D_1 = [(1-\lambda_1)^2 - r_{12}^2] (r_{12}^2 - r_{23}^2) > 0$$

$$\text{et donc } x^2 > z^2$$

On montrerait de même que $y^2 > z^2$ et $x^2 > y^2$

En effet ,

$$\frac{y^2}{z^2} = \frac{N_2}{D_2} \Rightarrow N_2 - D_2 = [(1-\lambda_1)^2 - r_{12}^2] (r_{12}^2 - r_{13}^2) > 0$$

$$\frac{x^2}{y^2} = \frac{N_3}{D_3} \Rightarrow N_3 - D_3 = [(1-\lambda_1)^2 - r_{13}^2] (r_{13}^2 - r_{23}^2) > 0$$

$$\text{donc } x^2 > y^2 > z^2.$$

REMARQUE

Cette propriété que nous venons d'établir pour $n=3$ n'est pas vraie dans le cas général comme nous pourrions le constater dans les applications numériques étudiées en Annexe 2.

Cependant, on peut penser que si en plus des relations entre coefficients de corrélation on fait intervenir une relation entre la valeur propre maximale et ces mêmes coefficients, on pourrait établir la propriété IV-1 dans le cas général.

II - APPLICATION 21. Introduction

Soit un ensemble Ω , une tribu α d'évènements dans Ω et une probabilité P définie sur α .

Nous supposons donné un système complet d'évènements E sur (Ω, α) fini soit

$$E = (E_1, E_2, \dots, E_k)$$

Nous disposons par ailleurs de n systèmes complets d'évènements Q_1, Q_2, \dots, Q_n . Si l'on note $Q_j^1, Q_j^2, \dots, Q_j^{a(Q_j)}$ les éléments du système Q_j , nous dirons que Q_j est une question de base $a(Q_j)$ et d'issues $Q_j^1, Q_j^2, \dots, Q_j^{a(Q_j)}$.

Nous cherchons à opérer une partition $\Omega_1, \Omega_2, \dots, \Omega_k$ de Ω à l'aide des questions Q_j $j=1,2,\dots,n$ telle que si $\omega \in \Omega$, $\exists j$ tel que $P_{ij} = P(\omega \in E_i / \omega \in \Omega_j)$ soit "grande" et donc que les P_{is} $\forall s \neq j$ soient "petites".

Nous dirons que les Q_j sont les questions "explicatives" et E la variable "à expliquer".

Les résultats obtenus ne seront vraiment intéressants que si le nombre de questions "explicatives" intervenant dans la constitution d'un Ω_i $i = 1, 2, \dots, k$ sera réduit, c'est-à-dire que pour chaque E_i on désire connaître les questions explicatives les plus discriminantes.

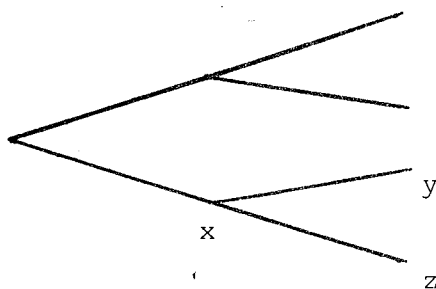
Nous allons présenter deux méthodes utilisant la notion de pseudoquestionnaire.

2. Rappel sur les pseudoquestionnaires [17]

On appelle pseudoquestionnaire construit sur Q_1, Q_2, \dots, Q_n le triplet (A, u, B) où :

- A est un graphe arborescent (X, Γ)
- u est une application de l'ensemble Y des sommets non terminaux de A dans (Q_1, Q_2, \dots, Q_n) vérifiant
 - i) $|\Gamma_y| = a(u(y)) \quad \forall y \in Y$
 - ii) la restriction de u à l'ensemble des sommets de tout chemin de A est injective
- $B = \{b_y \mid y \in Y\}$ où b_y est une bijection entre l'ensemble des arcs sortant de y et celui des issues de la question $u(y)$.

Exemple



$u(x) = q$ question d'issues q_1 et q_2

$b_x((x,y)) = q_1$

$b_x((x,z)) = q_2$

Soit $K = (A, u, B)$ un pseudoquestionnaire :

- On dit par extension que x est un sommet de K si x est un sommet de A
- On appelle rang d'un sommet de K le nombre de ses ascendants dans A.
- On appelle hauteur le nombre H tel qu'il n'existe pas de sommets de rang supérieur à H et tel qu'il existe un sommet de rang H.

- Soit $z \in K$ et soit z_1, z_2, \dots, z_k ses ascendants.

$$b_{z_1}((z_1, z_2)) = Q_{i_1}^{\alpha_1}$$

$$b_{z_2}((z_2, z_3)) = Q_{i_1}^{\alpha_2}$$

.

.

$$b_{z_k}((z_k, z)) = Q_{i_1}^{\alpha_k}$$

On appelle probabilité de z la probabilité de l'évènement associé

$$\mathcal{E}(z) = \prod_{j=1}^k Q_{i_1}^{\alpha_j}$$

et vecteur de probabilité associé au sommet z l'élément $P(E/z)$ ayant pour composantes :

$$P(T_j/E(z)) \quad j = 1, 2, \dots, \ell$$

que l'on note par abus de langage $P(T_j/z)$.

- Etant donné un nombre $\varepsilon \in]0, 1[$, un sommet z est interprétable à ε près si dans K il existe $j_0 \in [1, \dots, \ell]$ tel que

$$P(E_{j_0}/z) \geq 1 - \varepsilon.$$

3. Nouvelle formulation du problème

Etant donné ε , il s'agit de construire un pseudoquestionnaire $K^* = (A, u, B)$ sur Q_1, Q_2, \dots, Q_n de hauteur maximale R_0 tel que la probabilité de l'ensemble des sommets terminaux de K^* non interprétables à ε près soit "minimale".

Nous construirons ce pseudoquestionnaire par récurrence, c'est-à-dire si $K^*(R) = (A_r, u_r, B_r)$ est la restriction de hauteur R de K^* et si

x est un sommet de rang R dans $K^*(R)$, il suffit soit de décider que x est terminal dans $K^*(R+1)$, soit de définir :

- $u_{R+1}(x)$. L'arborescence A_{R+1} est alors déterminée en raison de la condition $a(u_{R+1}(x)) = |\Gamma_x|$
- une bijection entre les arcs sortant de x ainsi construits et les issues de la question $u_{R+1}(x)$.

Afin de déterminer cette application u , nous allons proposer deux méthodes : l'une faisant appel à l'information de Shannon, l'autre à la liaison-information.

REMARQUE

Il est inutile de soulever la question d'existence d'un pseudoquestionnaire dont la probabilité des sommets non interprétables à ϵ près soit inférieur à un seuil fixé à l'avance.

En effet, la construction de K^* de hauteur R_0 (R_0 est en général inférieur à 5 ou 6 dans les applications pratiques) est relativement de faible coût. On se contentera de vérifier à posteriori l'intérêt du résultat.

4. Méthode 1

a) Rappel sur l'information de Shannon

On appelle entropie du système E pour la mesure P la quantité $H_P(E)$ définie par

$$H_P(E) = \sum_{j=1}^{\ell} P(E_j) \text{Log} \frac{1}{P(E_j)}$$

Etant donné un élément A de α , on notera P_A la mesure définie sur (Ω, α) par :

$$P_A(B) = P(B/A) = \frac{P(A \cap B)}{P(A)} \quad \forall B \in \alpha$$

Si $S = (S_1, S_2, \dots, S_n)$ est un système complet d'évènements, on appelle entropie de E conditionnée par S la quantité $H_P(E/S)$ définie par

$$H_P(E/S) = \sum_{i=1}^n P(S_i) H_{P_{S_i}}(E)$$

On appelle information apportée par S sur E pour la mesure P , la quantité $I_P(E,S)$ définie par :

$$I_P(E,S) = H_P(E) - H_P(E/S)$$

b) Exposé de la méthode 1 [17]

Considérons le pseudoquestionnaire $K_0^* = (A,u,B)$ défini de la façon suivante :

- un sommet est terminal dans K_0^* si et seulement si il est interprétable à ϵ près ou de rang égal à R_0 .
- pour un sommet y non terminal, on choisit $u(y)$ tel que :

$$I_{P_y}(u(y),E) = \max_j [I_{P_y}(Q_j,E) / Q_j \in D(y)]$$

où $D(y) = \{Q_j \quad j=1,2,\dots,n : u(x) \neq Q_j \quad \forall x \in A(y)\}$

P_y est la mesure engendrée par $P(E/y)$

Cette heuristique consiste donc à poser en chaque sommet la question qui apporte le plus d'information sur l'état du système E .

Le pseudoquestionnaire K_0^* n'est sans doute pas le pseudoquestionnaire optimum, car l'information transmise par K_0^* n'est pas nécessairement maximale bien que l'information transmise par chaque question de K_0^* le soit, mais on peut toutefois penser que K_0^* approche de l'optimum si l'on se réfère aux résultats obtenus par l'algorithme de Fano [5] dans le cas des questionnaires et à de récents développements sur les pseudoquestionnaires [18] [19].

c) Inconvénient de la méthode 1

L'information de Shannon nécessite des questions de type qualitatif.

Dans la méthode 1, on est donc obligé, pour des questions de type quantitatif, de constituer des classes avant de pouvoir décider quelle sera la question à poser à un rang donné.

Dans la méthode 2 par contre, les classes ne seront constituées dans une question de type quantitatif que lorsqu'elle aura été choisie comme "optimale" à un certain rang.

5. Méthode 2

a) Introduction

Considérons d'une part n applications Y_1, Y_2, \dots, Y_n de Ω dans la droite réelle \mathbb{R} de la forme

$$Y_j(\omega) = y_j^i \quad \text{si } \omega \in Q_j^i \quad j = 1, 2, \dots, n$$

et d'autre part une application X de Ω dans \mathbb{R} de la forme :

$$X(\omega) = x_i \quad \text{si } \omega \in E_i$$

Si R est la matrice de corrélation de ces $n+1$ variables aléatoires, nous pouvons calculer la liaison-information conditionnelle de type 1 et de type 2 soit :

$$I_X^{(1)}(R) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \det R_{0ij}$$

et

$$I_X^{(2)}(R) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n (1 - \rho_{0/ij}^2)$$

l'indice zéro correspondant à la variable X .

Nous allons proposer deux critères de choix des questions en chaque sommet.

b) 1er critère de choix

Considérons le pseudoquestionnaire $K_1^* = (A, u, B)$ défini de la façon suivante :

- un sommet est terminal dans K_1^* si et seulement si il est interprétable à ϵ près ou s'il est de rang R_0 .

- pour un sommet y non terminal on choisit $u(y) = Q_i$ tel que

$$\alpha_i = \sum_{j=1}^n \det R_{0ij} = \min_k [\alpha_k / Q_k \in D(y)]$$

$$\text{où } D(y) = \{Q_k \quad k=1,2,\dots,n : u(x) \neq Q_k \quad \forall x \in A(y)\}$$

Si nous disposons de variables quantitatives, nous pouvons, sans constituer des classes, calculer la matrice de corrélation R . Ainsi nous pouvons constituer des classes dans une variable quantitative Q_i après qu'elle ait été choisie comme "meilleure" question.

Une méthode de constitution de classes a été proposée en annexe 1.

c) 2ème critère de choix

Le principe de la méthode reste le même, seulement pour déterminer la question Q_i correspondant au sommet non terminal y du pseudoquestionnaire K_2^* nous minimisons :

$$\beta_i = \sum_{j=1}^n (1 - \rho_{0ij}^2) = \min_k [\beta_k / Q_k \in D(y)]$$

$$\text{où } D(y) = \{Q_k \quad k=1,2,\dots,n : u(x) \neq Q_k \quad \forall x \in A(y)\}.$$

REMARQUE

Si nous désirons comparer les pseudoquestionnaires obtenus par les deux méthodes précédentes, nous calculerons les informations transmises par chacun d'eux soit

$$I(K_0^{**}), I(K_1^{**}), I(K_2^{**})$$

$$\text{où } I(K_i^{**}) = \sum_{z \in Z_i} P(z) I_{P_z}(u(z), E) \quad i=0,1,2.$$

Z_i représentant l'ensemble des sommets non terminaux de K_i .

En effet $I(K_i)$ sera d'autant plus "grande" que la probabilité de l'ensemble des sommets terminaux non interprétables à ϵ près sera "petite".

ANNEXE 1

METHODE POUR CONSTITUER DES CLASSES DANS UNE VARIABLE DE MESURE

Soient y_1, y_2, \dots, y_N les réponses de la population π à la question Y, variable de mesure.

On supposera $y_1 \neq y_2 \neq \dots \neq y_N$

et de plus $y_1 < y_2 < \dots < y_N$

Soient

$$d_1 = y_2 - y_1$$

$$d_2 = y_3 - y_2$$

.....

$$d_{N-1} = y_N - y_{N-1}$$

Classons par ordre croissant les valeurs d_i

soit $d_{i_1} < d_{i_2} < \dots < d_{i_{N-1}}$

où $d_{i_1} = y_{i_1} - y_{i_1-1} \dots d_{i_{N-1}} = y_{i_{N-1}} - y_{i_{N-1}-1}$

Soit $\bar{d}_i = \frac{1}{N-1} \sum_{j=1}^{N-1} d_{i_j}$

Considérons les rapports suivants :

$$\frac{d_{i_{N-1}}}{\bar{d}_i} > \frac{d_{i_{N-2}}}{\bar{d}_i} > \dots > \frac{d_{i_1}}{\bar{d}_i}$$

et fixons-nous une valeur $\gamma \geq 1$ qui nous permettra de décider automatiquement du nombre de classes.

Si k est tel que $\frac{d_{i_k}}{\bar{d}_i} > \gamma > \frac{d_{i_{k-1}}}{\bar{d}_i}$

Considérons les valeurs de Y suivantes :

$$y_{i_{N-1}}, y_{i_{N-1}-1}, \dots, y_{i_k}, y_{i_k-1}.$$

On obtient ainsi $\ell = 2(N-k)$ valeurs différentes si nous supposons qu'aucun point n'est isolé.

Classons ces ℓ valeurs par ordre croissant et avec une numérotation convenable, nous obtenons :

$$y_{\alpha_1} \quad y_{\alpha_2} \quad \dots \quad y_{\alpha_\ell}$$

Si h est un individu de la population π et $Y(h)$ sa réponse à la question Y , considérons les classes suivantes :

$$A_1 = \{h \in \pi / Y(h) \leq y_{\alpha_1}\}$$

$$A_2 = \{h \in \pi / y_{\alpha_2} \leq Y(h) \leq y_{\alpha_3}\}$$

.....

$$A_{N-k} = \{h \in \pi / y_{\alpha_{\ell-2}} \leq Y(h) \leq y_{\alpha_{\ell-1}}\}$$

$$A_{N-k+1} = \{h \in \pi / Y(h) \geq y_{\alpha_\ell}\}$$

Nous obtenons ainsi $N-k+1$ classes. Assurons-nous qu'aucune classe ne contienne un nombre d'individus inférieur à un "seuil de classe" β fixé au départ, sinon il conviendrait de regrouper ces individus à la classe la plus proche.

REMARQUES

1 - Si certains points sont isolés, il se peut que dans la suite des valeurs

$$y_{i_{N-1}}, y_{i_{N-1}-1}, \dots, y_{i_k}, y_{i_k-1},$$

une valeur soit considérée deux fois, auquel cas l'individu qui aura fourni cette valeur appartiendra à la classe la plus proche.

2 - Pour fixer numériquement γ , on peut utiliser la formule de Bienaymé-Tchekyscheff :

$$P\{|Z-E(Z)| \geq \varepsilon\} \leq \frac{\sigma_Z^2}{\varepsilon^2} \text{ où } Z \text{ est une variable aléatoire quelconque.}$$

La variable Z que nous considérerons aura pour valeur l'ensemble des d_{ij} .

$$Z = \{d_{ij}\}$$

$$E(Z) = \frac{1}{N-1} \sum d_{ij} = \bar{d}_i$$

$$\sigma_Z^2 = \frac{\sum d_{ij}^2}{N-1} - \bar{d}_i^2$$

$$\text{Posons } \alpha^2 = \frac{\sigma_Z^2}{\varepsilon^2} \Rightarrow \varepsilon^2 = \frac{\sigma_Z^2}{\alpha^2}$$

On obtient

$$P\{|d_{ij} - \bar{d}_i| > \frac{\sigma_Z}{\alpha}\} \leq \alpha^2$$

ou

$$P\left\{\left|\frac{d_{ij}}{\bar{d}_i} - 1\right| \geq \frac{\sigma_Z}{\alpha \bar{d}_i}\right\} \leq \alpha^2$$

$$P\left\{1 - \frac{\sigma_Z}{\alpha \bar{d}_i} < \frac{d_{ij}}{\bar{d}_i} < 1 + \frac{\sigma_Z}{\alpha \bar{d}_i}\right\} > 1 - \alpha^2$$

$$P\left\{\gamma_1 < \frac{d_{ij}}{\bar{d}_i} < \gamma_2\right\} > 1 - \alpha^2$$

Ce qui nous intéresse c'est la valeur γ_1 soit

$$1 - \frac{\sigma_Z}{\alpha \bar{d}_i}$$

On pourrait donc, au lieu de fixer numériquement γ , donner une valeur de α .

ANNEXE 2

RESULTATS NUMERIQUES

INTRODUCTION

Un programme a été écrit pour chacune des applications définies dans le chapitre IV. Les données numériques utilisées pour tester les programmes ne correspondent pas à des données réelles provenant d'une enquête réalisée sur une population, mais ont été définies d'une façon aléatoire.

I - ANALYSE FACTORIELLE ET LIAISON-INFORMATION

Etant donné les réponses d'une population π aux questions Y_1, Y_2, \dots, Y_N le programme P_1 calcule :

- la matrice de corrélation R des variables Y_1, Y_2, \dots, Y_N .
- la valeur propre maximale λ_M de R .
- le vecteur propre V_M associé à λ_M .
- les liaisons-informations apportées par les variables Y_1, Y_2, \dots, Y_N .

Des résultats numériques du programme P_1 , on peut tirer les constatations suivantes :

dans certains cas, l'ordre d' "influence" décroissante des variables Y_i sur l'axe factoriel V_M est le même que l'ordre croissant des liaisons-informations apportées par les variables Y_i , et dans d'autres ils peuvent différer quelque peu comme dans l'exemple numérique suivant :

$$\lambda_M = 6,866152 \rightarrow V_M$$

$$\begin{array}{l} a_{M1} = - 0.01492729 \\ a_{M2} = 0.3532097 \\ a_{M3} = 0.3475269 \\ a_{M4} = 0.3520029 \\ a_{M5} = 0.273424 \\ a_{M6} = 0.3455701 \\ a_{M7} = 0.3622366 \\ a_{M8} = 0.3181097 \\ a_{M9} = 0.2977010 \\ a_{M10} = 0.3392907 \end{array}$$

$$\begin{array}{l} 9 \times I[Y_1(R)] = 0,3602151 \\ 9 \times I[Y_2(R)] = 4.919536 \\ 9 \times I[Y_3(R)] = 4.737910 \\ 9 \times I[Y_4(R)] = 4.921384 \\ 9 \times I[Y_5(R)] = 3.058548 \\ 9 \times I[Y_6(R)] = 4.753138 \\ 9 \times I[Y_7(R)] = 5.203061 \\ 9 \times I[Y_8(R)] = 3.895291 \\ 9 \times I[Y_9(R)] = 3.391309 \\ 9 \times I[Y_{10}(R)] = 4.481048 \end{array}$$

- L'ordre d'influence décroissante des Y_i sur V_M est donc :

$$Y_7, Y_2, Y_4, Y_3, Y_6, Y_{10}, Y_8, Y_9, Y_5, Y_1$$

- L'ordre croissant des liaisons-informations apportées par les Y_i est :

$$Y_7, Y_4, Y_2, Y_6, Y_3, Y_{10}, Y_8, Y_9, Y_5, Y_1$$

Donc, malgré quelques exceptions, en fait très limitées dans les exemples étudiés, on peut dire que plus l' "influence" de Y_i sur V_M sera "grande", plus la liaison-information apportée par Y_i sera "petite". Ce résultat confirme d'ailleurs la notion intuitive de la liaison-information introduite dans le chapitre I.

II - PSEUDOQUESTIONNAIRES

Etant donné les réponses de la population π à la variable à expliquer X et aux variables "explicatives" Y_1, Y_2, \dots, Y_N , le programme P_2 permet de construire les pseudoquestionnaires $K_0^{**}, K_1^{**}, K_2^{**}$, définis dans le Chapitre IV.

A la fin de ce paragraphe, vous trouverez trois pseudoquestionnaires construits à partir des résultats numériques du programme P_2 . Dans ce cas particulier $N = 10$ et le rang maximal R_0 a été limité à 4.

Si π_0, π_1, π_2 représentent respectivement les partitions de la population π correspondant à $K_0^{**}, K_1^{**}, K_2^{**}$, on a pu calculer : l'entropie de X conditionnée par π_0 , l'entropie de X conditionnée par π_1 et l'entropie de X conditionnée par π_2 soit :

$$H_P(X/\pi_0) = 1,486295$$

$$H_P(X/\pi_1) = 1,491720$$

$$H_P(X/\pi_2) = 1,906527$$

De même, on a calculé les longueurs de cheminement

$$L[K_i^{**}] = \sum_{x \in Z_i} P(x) r(x) \quad i = 0, 1, 2$$

où $r(x)$ est le rang de x dans K_i^{**} et Z_i l'ensemble des sommets terminaux de K_i^{**} :

$$L[K_0^{**}] = 2,9$$

$$L[K_1^{**}] = 3,48$$

$$L[K_2^{**}] = 3,18$$

Ces résultats numériques nous permettent de constater que la méthode 1 donne de meilleurs résultats aussi bien pour la longueur de cheminement que pour l'information de Shannon.

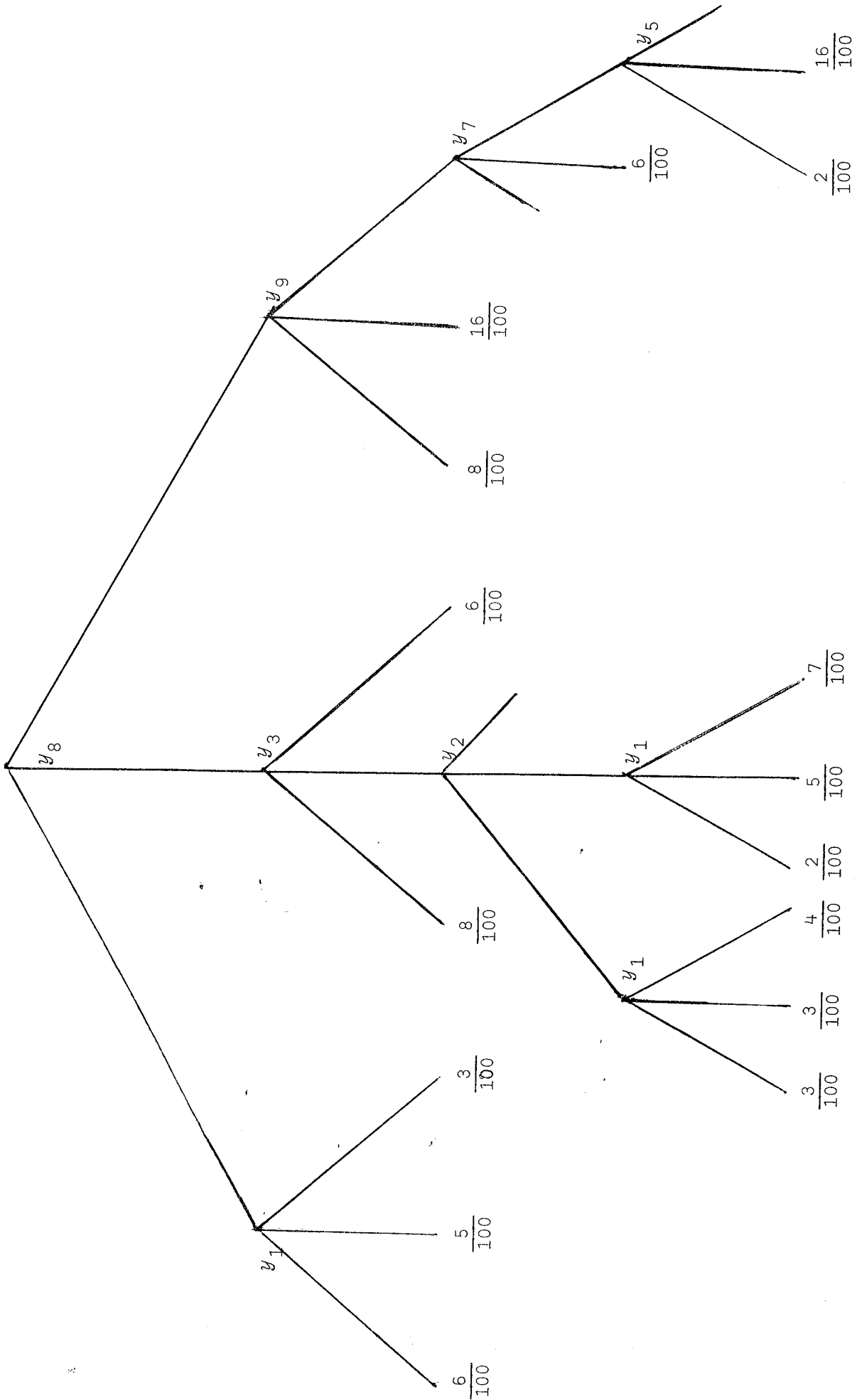
Cependant, les résultats obtenus par la méthode 2 sont comparables à ceux de la méthode 1. De plus, il est impossible, au vu des résultats numériques, de favoriser l'un des critères de choix de la méthode 2 puisque si l'on a

$$H_P(X/\pi_1) < H_P(X/\pi_2)$$

on a également

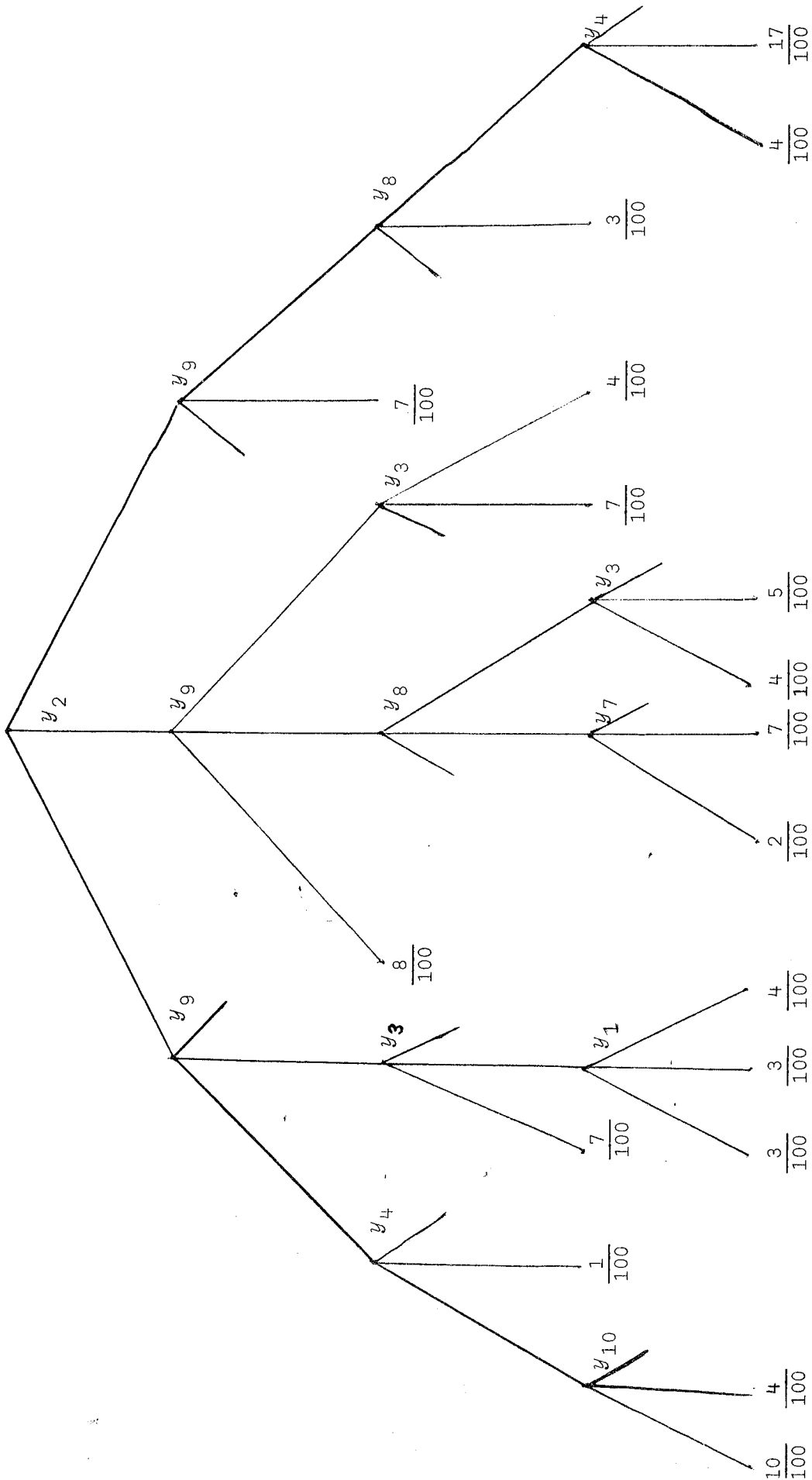
$$L[K_1^{**}] > L[K_2^{**}]$$

Nous laisserons donc à l'utilisateur le soin de décider du meilleur résultat.



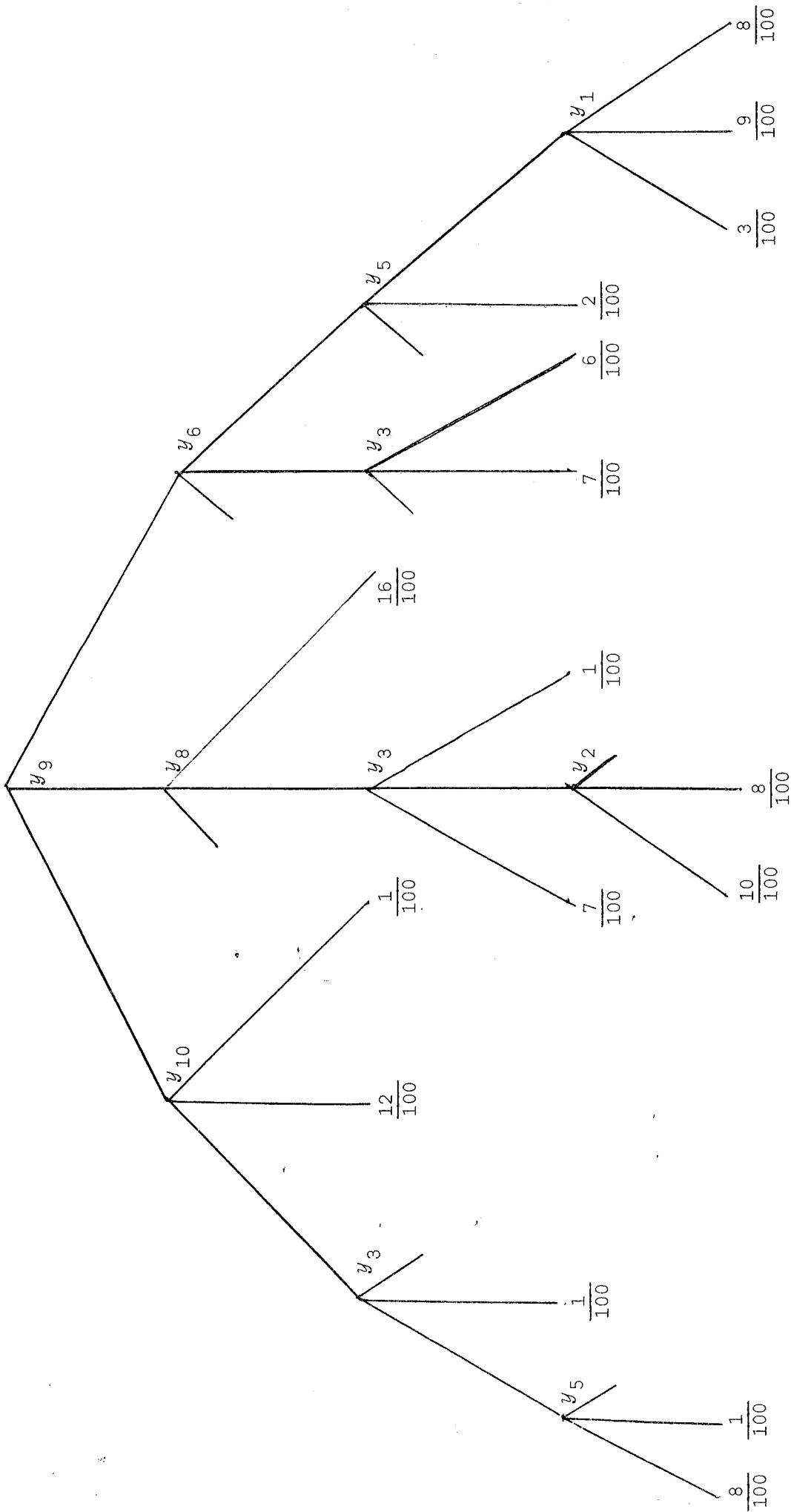
$$u(y_1) = y_1, u(y_2) = Y_2, \dots, u(y_{10}) = Y_{10}$$

$$H_P(X/\pi_0) = 1,486295 \quad L(K_0^*) = 2,9$$



$$u(y_1) = Y_1, u(y_2) = Y_2, \dots, u(y_{10}) = Y_{10}$$

$$H_P(X/\pi_1) = 1,491720 \quad L(K_0^{**}) = 3,48$$



$$u(y_1) = y_1, u(y_2) = y_2, \dots, u(y_{10}) = y_{10}$$

$$H_P(X/\pi_2) = 1,906527 \quad L(K_2^*) = 3,18$$

BIBLIOGRAPHIE

- [1] R.B. ASH "Information Theory"
John Wiley and Sons, Inc 1965
- [2] J.P. BENZECRI "Analyse factorielle"
Cours DEA, Paris
- [3] C. BERGE "Théorie des graphes et ses applications"
Dunod, Paris, 1963
- [4] G. CALOT "Cours de statistiques descriptives"
Dunod, 1964
- [5] R.M. FANO "Transmission of information"
M.I.T. Press Cambridge, Mass
- [6] C. FOURGEAUD "Statistique"
A. FUCHS Dunod, 1967
- [7] S. GUIASU "La théorie mathématique de l'information"
R. THEODORESCU Dunod, Paris, 1968
- [8] M.G. KENDALL "A course in multivariate analysis"
Charles Griffin and Company, 1965
- [9] S. KULLBACK "Information and statistics"
J. Wiley
- [10] J. NEVEU "Bases mathématiques du calcul des probabilités"
Masson, 1964

- [11] Ph. T. NGHIEM
KH VO KHAC "Etude sur les aspects théoriques et pratiques
de la segmentation aux moindres carrés"
RIRO N° 8 - VI - Dunod - 1968
- [12] C.F. PICARD "Théorie des questionnaires"
Gauthier Villars, Paris, 1965
- [13] C.F. PICARD "Introduction à la théorie de l'information"
Cours DEA - Lyon, 1965
- [14] C.F. PICARD "Théorèmes de l'information traitée"
Colloquium on information theory, Debrecen, 1967
- [15] G. ROMIER "Analyse des données et classification"
Groupe de travail du 6-11-1972
- [16] M. TERRENOIRE "Utilisation d'un critère informationnel pour le
dépouillement d'une enquête"
Congrès informatique appliqué à l'économie,
Bucarest, Octobre 1969
- [17] M. TERRENOIRE "Un modèle mathématique de processus d'interroga-
tion : les pseudoquestionnaires"
Thèse d'Etat, Grenoble, 1970
- [18] M. TERRENOIRE "Convergence heuristique pour des processus
d'interrogation"
Convegno sull' informatica teorica
Rome, Février 1973
- [19] M. TERRENOIRE "Convergence de pseudoquestionnaires particuliers"
TOUNISSOUX C.R.A.S. (à paraître).

TABLE DES MATIERES

TERMINOLOGIE	1.
INTRODUCTION	3.
CHAPITRE I - LIAISON-INFORMATION	
Introduction	4.
1) Propriétés de la liaison-information	4.
1) Remarques préliminaires	4.
2) Propriétés	5.
2) Conséquences des propriétés précédentes	6.
1) Conséquence I-1	6.
2) Conséquence I-2	7.
3) Conséquence I-3	7.
4) Conséquence I-4	9.
5) Conséquence I-5	9.
6) Conséquence I-6	10.
7) Conséquence I-7	11.
3) Expression de la liaison-information	13.
1) Théorème I-1 : expression de la fonction ψ dans un cas particulier	13.
2) Corollaire I-1	17.
3) Théorème I-2	17.
4) Corollaire I-2	21.
5) Conséquence I-8	23.
CHAPITRE II - INDEPENDANCE DES PROPRIETES ENONCEES AU CHAPITRE I	
Introduction	24.
1) Suppression de la propriété I-1	24.
2) Suppression de la propriété I-2	25.
1) Suppression de $\psi(u,v)=\psi(u,1)\psi(1,v)$	26.
2) Suppression de la propriété I-2 dans son ensemble ..	27.

3) Suppression de la propriété I-3	27.
4) Conclusion	28.

CHAPITRE III - LIAISON-INFORMATION CONDITIONNELLE

Introduction	29.
1) Définitions	29.
1) Définition III-1	29.
2) Définition III-2	29.
2) Propriétés	30.
1) Propriété III-1	30.
2) Propriété III-2	30.
3) Conclusion	31.

CHAPITRE IV - APPLICATIONS PRATIQUES

Introduction	32.
1) Application 1	32.
1) Rappel sur l'analyse factorielle	32.
2) Liaison-information apportée par une variable	33.
3) Applications pratiques	34.
Propriété IV-1	34.
Propriété IV-2	35.
2) Application 2	41.
1) Introduction	41.
2) Rappel sur les pseudoquestionnaires	42.
3) Nouvelle formulation du problème	43.
4) Méthode 1	44.
5) Méthode 2	46.

Annexe 1 - Méthode pour constituer des classes dans une variable de mesure	49.
---	-----

Annexe 2 - Résultats numériques	53.
---------------------------------------	-----

Bibliographie	60.
---------------------	-----