

# PARTIE 2

## COMPLEMENTS ET EXTENSIONS DE L'ANALYSE STATISTIQUE IMPLICATIVE

### Chapitre 1 : Statistique de rangs et Analyse Statistique Implicative

Jean-Claude Régnier \* et Régis Gras \*\*

\* Université de Lyon - UMR 5191 ICAR  
ENS-LSH 15, Parvis René Descartes BP 7000 69342 LYON cedex 07  
[jean-claude.regnier@univ-lyon2.fr](mailto:jean-claude.regnier@univ-lyon2.fr)

\*Equipe Connaissances et Décision, Laboratoire d'Informatique de Nantes Atlantique  
Ecole Polytechnique de l'Université de Nantes, UMR 6241  
La Chantrerie BP 60601 44306 Nantes cedex  
[regisgra@club-internet.fr](mailto:regisgra@club-internet.fr)

**Résumé.** Nous discutons de l'apport de la méthode d'analyse statistique implicative au sens de R. Gras, à l'étude de la concordance/discordance des rangs accordés par des juges à des objets. Cette dernière est à comprendre au sens de Friedman ou de Kendall. Ici nous comparons une analyse de préférences exprimées par les rangs, avec l'analyse de la propension entre variables modales de J. B. Lagrange. Nous nous affranchissons de l'hypothèse d'absence de lien *a priori* entre les variables. Nous affectons d'une mesure de qualité des énoncés de la forme : « si l'objet a est rangé par les juges alors, généralement, l'objet b est rangé à un rang meilleur par les mêmes juges », et représentons par un graphe les relations de préférences de l'ensemble des objets rangés. Nous nous limitons aux deux cas des rangements complets et incomplets mais sans *ex aequo* de q objets par k juges. Le texte présenté ici reprend en partie (Régnier et Gras, 2005)

## 1 Introduction.

Rappelons la problématique de la concordance des rangements de q objets par k juges. Considérons q objets soumis au jugement individuel et indépendant de k ( $k > 2$ ) individus d'un ensemble E qui doivent fournir un rangement complet et sans *ex aequo* de ces objets. Il s'agit de savoir si on peut dégager de l'ensemble des k jugements, une concordance

relativement à un rangement de ces  $q$  objets. C'est une des formes de **l'analyse des préférences** où, ici, nous n'envisageons pas de comparaison par paires.

La réalisation d'une telle opération peut être décrite par le tableau suivant :

Objets Individus	$O_1$	$O_2$	...	$O_j$	...	$O_q$
<b>1</b>	$R_{11}^{(q)}$	$R_{12}^{(q)}$		$R_{1j}^{(q)}$		$R_{1q}^{(q)}$
...						
<b>i</b>	$R_{i1}^{(q)}$	$R_{i2}^{(q)}$		$R_{ij}^{(q)}$		$R_{iq}^{(q)}$
...						

TAB. 1 - *Tableau de rangs*

où chaque ligne  $i$  ( $i = 1, \dots, k$ ) est une permutation de  $(1, 2, \dots, q)$ , réalisations du vecteur aléatoire  $(R_{ij}^{(q)})_{j=1, \dots, q}$ . Il y a donc  $q!$  réalisations possibles par ligne. Les contraintes de rangement complet et sans *ex aequo* conduisent à ce que chaque ligne  $i$  correspondant à l'individu  $i$  vérifie la relation  $\sum_{j=1}^q R_{ij}^{(q)} = \frac{q(q+1)}{2}$ . Sous l'hypothèse  $H_0$ , généralement légitime, de l'équiprobabilité des  $q!$  rangements des  $q$  objets par l'individu  $i$ , la probabilité pour l'objet  $j$ ,  $j \in \{0, 1, \dots, q\}$ , d'être placé au rang  $r$ ,  $r \in \{0, 1, \dots, q\}$  est  $\frac{1}{q}$ . La probabilité pour que l'objet  $j_1$ ,  $j_1 \in \{0, 1, \dots, q\}$ , soit placé au rang  $r_1$ ,  $r_1 \in \{0, 1, \dots, q\}$ , et l'objet  $j_2$ ,  $j_2 \in \{0, 1, \dots, q\} - \{j_1\}$ , soit placé au rang  $r_2$ ,  $r_2 \in \{0, 1, \dots, q\}$  et  $r_2 \neq r_1$  est de  $\frac{1}{q(q-1)}$ . Les caractéristiques des variables aléatoires  $R_{ij}^{(q)}$  sont :

Les caractéristiques des variables aléatoires  $R_{ij}^{(q)}$  sont :

Espérance	Variance	Covariance	Corrélation
$E(R_{ij}^{(q)}) = \frac{q+1}{2}$	$V(R_{ij}^{(q)}) = \frac{q^2-1}{12}$	$Cov(R_{ij_1}^{(q)}; R_{ij_2}^{(q)}) = -\frac{q+1}{12}$	$Corr(R_{ij_1}^{(q)}; R_{ij_2}^{(q)}) = -\frac{1}{q-1}$

TAB. 2 - *Tableau des caractéristiques des variables rangs aléatoires*

## 2 Approche par les statistiques de Friedman et de Kendall

Considérons le tableau TAB.1. Les sommes marginales des lignes sont constantes et égales à la somme des rangs attribués par un même individu. Il comporte  $(q!)^k$  réalisations équiprobables sous l'hypothèse  $H_0$  d'équiprobabilité des permutations caractérisant le rangement des  $q$  objets effectué indépendamment par chaque individu  $i$ . Désignons par  $S(O_j)$  la somme des rangs attribués respectivement par les  $k$  individus à l'objet  $O_j$ . Pour avoir la distribution de probabilité exacte de  $S(O_j)$ , il suffirait de construire les  $(q!)^k$  tableaux, opération, concrètement impossible dans les cas intéressants. Sous  $H_0$ , nous pouvons cependant connaître les caractéristiques des variables  $S(O_j)$ ,  $j \in \{0, 1, \dots, q\}$ .

Espérance	Variance
$E(S(O_j)) = k \frac{q+1}{2}$	$V(S(O_j)) = k \frac{q^2 - 1}{12}$

TAB. 3 - Tableau des caractéristiques des variables aléatoires  $S(O_j)$ , somme des rangs

## 2.1 Interprétation des $S(O_j)$ pour expliciter un rangement concordant

Nous pouvons utiliser l'ordre des sommes des rangs attribués à chaque objet pour déterminer un rangement que nous interprétons comme étant celui qui « s'approche au mieux » de chaque point de vue des juges. Deux cas extrêmes peuvent apparaître.

### 2.1.1 Concordance « parfaite »

Cette situation correspond au cas où les juges attribueraient respectivement à chaque objet le même rang.

Il y a  $q!$  tableaux de ce type sur les  $(q!)^k$  possibles. Sous  $H_0$ , cet événement a donc une probabilité  $\left(\frac{1}{q!}\right)^{k-1}$  de survenir.

### 2.1.2 Hétérogénéité « parfaite » des rangements

Cette situation correspond au cas où l'attribution des rangs aux  $q$  objets produirait par des effets de compensation des sommes de rangs égales ou presque égales à une unité près. Le nombre de tableaux vérifiant ces conditions est à l'évidence très supérieur aux  $q!$  tableaux de la concordance « parfaite ». Intuitivement, sous  $H_0$ , nous avons beaucoup plus de chance d'obtenir un tableau du type « hétérogénéité parfaite » que du type « concordance parfaite ».

### 2.1.3 Entre « concordance parfaite » et « hétérogénéité parfaite »

La question demeure quant à l'interprétation des cas intermédiaires. Quelle décision prendre si nous obtenons des résultats compris entre ces deux extrêmes ? La **statistique  $F_{q,k}$  de Friedman** ou celle  **$W$  de Kendall** constituent des outils d'aide à la prise de cette décision.

$$F_{q,k} = \frac{12k}{q(q+1)} \sum_{j=1}^q \left( \frac{S(O_j)}{k} - \frac{q+1}{2} \right)^2$$

$$W = \frac{\sum_{j=1}^q \left( S(O_j) - k \frac{(q+1)}{2} \right)^2}{\frac{k^2(q^3 - q)}{12}} = \frac{S_K}{S_{K \max}}$$

Ces deux statistiques sont reliées par :  $F_{q,k} = k(q-1)W$ . Elles constituent des mesures d'écart entre la situation observée et celle de l'hétérogénéité « parfaite ».

## 2.2 Test de concordance des rangements de q objets par k juges, W de Kendall

La loi de W sous  $H_0$  ne dépend que des  $(q!)^k$  résultats possibles. Elle est théoriquement connaissable mais pratiquement impossible en raison de la taille des calculs en jeu. On obtient les valeurs critiques à partir d'une table qui a été établie lorsque  $q < 7$ . Sinon pour  $q \geq 7$ , on utilise la loi de la statistique  $k(q-1)W$  qui est approximativement la loi du  $\chi^2_{(q-1)}$  à  $q-1$  ddl.

## 3 Analyse statistique implicative, outil d'exploration des rangements complets et sans *ex aequo*.

À l'origine (Gras et al, 2001), comme nous avons pu le voir dans la première partie de cet ouvrage, l'analyse statistique implicative s'applique aux variables binaires en produisant des règles, puis aux classes de variables binaires en produisant des méta-règles (R-règles). Cette méthode d'analyse de données constitue un outil pour expliciter statistiquement des structures quasi-implicatives. Nous rappelons le tableau de contingence qui constitue la référence sur laquelle se fonde le raisonnement quasi-implicatif.

		Variable <b>b</b>		
		1	0	
Variable <b>a</b>	1	$n(a \wedge b)$	$n(a \wedge \neg b)$	$n(a)$
	0	$n(\neg a \wedge b)$	$n(\neg a \wedge \neg b)$	$n(\neg a)$
		$n(b)$	$n(\neg b)$	$n$

TAB. 4 - Tableau de contingence de référence du raisonnement

Nous supposons que  $n(a) < n(b)$  et nous nous intéressons aux individus qui contredisent l'implication mathématique dont l'effectif est  $n(a \wedge \neg b)$ .

$0 \leq n(a \wedge \neg b) < \frac{n(a)n(\neg b)}{n}$	L'effectif observé étant inférieur à l'effectif théorique signifie que la dépendance entre la présence de (a) et l'absence de (b) est répulsive. En terme d'implication statistique, l'observation qui va dans le sens $(a) \Rightarrow (b)$
$0 < \frac{n(a)n(\neg b)}{n} < n(a \wedge \neg b)$	L'effectif observé étant supérieur à l'effectif théorique signifie que la dépendance entre la présence du caractère (a) et l'absence du caractère (b) est attractive. En terme d'implication statistique, nous interprétons une observation qui va dans le sens de la négation de $(a) \Rightarrow (b)$

TAB. 5 - Tableau des interprétations qui fonde le raisonnement par quasi-implication.

Cela conduit à définir une implication statistique admissible à un niveau de confiance fixé. Par la suite le rapprochement entre variables de rangs et analyse statistique implicative apparaît avec l'extension de cette théorie aux variables et classes de variables modales à valeurs sur  $[0; 1]$  de IR. Dans ce passage des variables binaires vers les variables modales,

J.B. Lagrange (Lagrange, 1998) avait introduit la notion de « **relation de propension** entre deux variables à valeurs dans l'intervalle  $[0 ; 1]$  ». Comme nous l'avons exposé en Partie 1, il a été défini un *indice (coefficient) de propension*  $\tilde{q}(a, \neg b)$  d'une variable modale a vers une variable modale b ainsi qu'une *intensité de propension*  $\Phi(-\tilde{q}(a, \neg b))$ . Le modèle adopté est une stricte extension de l'implication statistique entre variables binaires. Dans le cadre de l'approche par l'analyse implicative, le recours aux variables de rangs est apparu dans les travaux de Marc Bailleul (1994). Pour explorer les représentations de l'enseignement des mathématiques, il a été proposé un protocole qui amène les sujets interrogés à fournir un rangement. Dans le choix de l'outil mathématique d'analyse, M.Bailleul inscrit la contrainte de la prise en compte d'un indice non-symétrique permettant de répondre à la question : *dans quelle mesure peut-on affirmer, de façon relativement sûre, que le choix par un individu de telle variable va entraîner le choix de telle autre variable ?* Et concernant le cas des variables de rangs qui nous intéresse : *dans quelle mesure le choix de telle variable avec tel rang entraîne-t-il le choix de telle autre variable avec un rang meilleur ou égal à celui affecté à la première ?*

### 3.1 Transformation d'une variable de rang en une variable modale

En premier lieu, nous nous plaçons dans le cas d'un rangement complet et sans *ex æquo*. En général il est demandé aux k individus de ranger les q objets par ordre de préférence décroissante. Ainsi le rang 1 est le plus important tandis que le rang q correspond à l'objet le moins important. Or dans le modèle proposé par J.B Lagrange la valeur 1 représente le plus haut degré d'adhésion possible tandis que le degré minimum d'adhésion est coté par 0. Dans ces conditions, la transformation suivante est une solution à ce problème :

$$X_{ij}^{(q)} = \frac{q - R_{ij}^{(q)}}{q - 1}$$

qui respecte la contrainte sémantique.

### 3.2 Le modèle de la quantification de la propension de J.B. Lagrange

Étant données deux variables modales a et b,  $(a_i)_{i=1, \dots, n}$  et  $(b_i)_{i=1, \dots, n}$  deux vecteurs de  $[0 ; 1]^n$  représentant une réalisation de ces variables auprès des n individus d'un ensemble P. « Il y a propension de a vers b si on rencontre en moyenne peu d'individus pour lesquels  $(a_i)$  est fort et  $(b_i)$  l'est moins . La moyenne sur P des  $a_i(1-b_i)$  est pris comme indice de non-propension de a vers b. Cette valeur moyenne est comparée à celle qui résulterait d'une hypothèse d'absence de lien entre a et b ». J.B. Lagrange propose alors la statistique  $\tilde{Q}(a, \neg b)$  approximativement distribuée comme la variable de Laplace Gauss centrée réduite, dont la réalisation donne le *coefficient de propension de a vers b* le nombre à partir de

$$\tilde{q}(a, \neg b) = \frac{\sum_{i \in P} a_i(1-b_i) - m_a(1-m_b)}{n} - \frac{m_a(1-m_b)}{\sqrt{\frac{(v_a + m_a^2)(v_b + (1-m_b)^2)}{n}}}$$

à partir

duquel il calcule l'*intensité de la propension*. I.C. Lerman (1992) obtenait déjà et

indépendamment une formule comparable par extension du cas booléen au cas numérique de différentes formes de l'hypothèse d'absence de liaison, dans le contexte de la similarité.

### 3.3 Application du modèle de la propension aux variables de rangs

L'exploration de la structure du tableau TAB.1 de rangs par l'analyse statistique de la propension requiert un réexamen des conditions d'application du modèle de Lagrange. Afin de pouvoir comparer les approches, nous supposons que les valeurs modales sont réparties de façon régulière (équidistante) sur  $[0 ; 1]$ . Les  $k$  événements  $\{i \in P\}$  qui définissent les lignes du tableau, sont évidemment supposés a priori indépendants. Là où un problème nous paraît surgir, dans le cadre du traitement modal en terme de rangs, c'est dans le choix *a priori* d'une hypothèse d'indépendance entre les deux variables  $A_i$  et  $B_i$ . Dans la procédure de rangement comme nous l'avons précisé plus haut, chaque composante du vecteur  $i$ , c'est-à-dire chaque variable aléatoire rattachée au critère  $i$  et à l'objet  $j$ , dépend des  $q-1$  autres objets. Il n'est donc pas possible ici de considérer l'indépendance de deux variables  $A_i$  et  $B_i$  — dont la réalisation donne respectivement le rang de l'objet A et celui de l'objet B — comme acceptable. Elle est en effet incompatible avec la procédure même qui conduit l'individu à se confronter à la totalité des  $q$  objets pour produire finalement une permutation de  $\{1, \dots, q\}$ . Ceci nous montre qu'il n'est pas possible de transporter le modèle modal de Lagrange dans notre modèle de rangements complets sans *ex æquo*. Comment peut-on exprimer la structure issue du rangement complet et sans *ex æquo* des  $q$  objets par  $k$  sujets en terme de relation de propension ?

Réinterprétons ce que propose Lagrange, c'est à dire la quantité  $s = \frac{\sum_{i \in P} a_i(1 - b_i)}{n}$ , dans

le contexte des variables de rangs. Dire que  $a_i$  est fort, traduit le fait que l'objet A considéré, est placé à un haut degré de préférence. Dire que  $b_i$  est faible, traduit le fait que l'objet B considéré, est placé à un bas degré de préférence. La relation de propension de a vers b devrait traduire le fait que « la tendance à placer l'objet A à un rang  $r_A$  entraîne celle à placer l'objet B à un rang  $r_B$  correspondant à une préférence de B sur A. » Dans notre modélisation du choix préférentiel, B est préféré à A si  $\text{rang}(B) < \text{rang}(A)$  et donc si  $0 \leq \frac{q - \text{rang}(A)}{q - 1} < \frac{q - \text{rang}(B)}{q - 1} \leq 1$  Dans ces conditions, en conservant la statistique

$$\tilde{Q}(a, -b) = \frac{Z - E[A_i(1 - B_i)]}{\sqrt{\frac{E([A_i(1 - B_i)]^2)}{n}}}$$

$B_i$  sont dépendantes pour le calcul de l'espérance et de la variance de Z. Nous conservons l'approximation gaussienne de la variable aléatoire indice Z dans la mesure où nous restons dans les conditions posées par Lagrange : d'une part, pour chaque  $i$ , l'évènement  $\{i \in P\}$  et les évènements définis par la variable  $A_i(1 - B_i)$  sont indépendants, d'autre part, les variables  $T_i$  sont indépendantes et identiquement distribuées.

### 3.4 Expression du coefficient de propension pour des variables rangs

En renonçant à l'indépendance, nous avons recalculé l'espérance et la variable de Z, ce qui nous a conduits au résultat suivant par lequel le coefficient de propension s'exprime ainsi

$$\tilde{q}_{(q)}(a, -b) = \frac{\frac{\sum_{i \in P} a_i(1-b_i)}{n} - \frac{1}{2} + \frac{(q-2)(3q-1)}{12(q-1)^2}}{\sqrt{\frac{20q^4 - 46q^3 + 49q^2 - 41q + 24}{180n(q-1)^4}}} = \frac{\frac{\sum_{i \in P} a_i(1-b_i)}{n} - \left[ \frac{1}{4} + \frac{(q+1)}{12(q-1)^2} \right]}{\sqrt{\frac{20q^4 - 46q^3 + 49q^2 - 41q + 24}{180n(q-1)^4}}}$$

### 3.5 Compatibilité de la relation de propension permutationnelle avec l'ordre des moyennes

Considérons deux variables a et b issues de la transformation des rangs attribués aux objets  $O_1$  à  $O_q$ . Notons  $m_a$  et  $m_b$  leurs moyennes respectives et  $\Phi$  la fonction de répartition de la variable de Laplace Gauss centrée réduite. Nous démontrons aisément que :

**Proposition 1 :**  $\Phi(-\tilde{q}_{(q)}(a, -b)) \geq \Phi(-\tilde{q}_{(q)}(b, -a))$  si et seulement si  $m_a \leq m_b$

**Proposition 2 :**  $\Phi(-\tilde{q}_{(q)}(a, -b)) \geq \Phi(-\tilde{q}_{(q)}(b, -a))$  si et seulement si Rang(Objet B)  $\leq$  Rang(Objet A) dans le rangement synthétique.

Ceci signifie que le rangement de l'objet A à un degré de préférence donné « entraîne » le rangement de l'objet B à un degré de préférence plus élevé. Ces deux propositions permettent l'accord sémantique de l'indice  $\tilde{q}$  avec l'intensité implicative où la variable a serait moins bien classée que la variable b. Il advient alors que l'ordre des objets issu de l'ordre des intensités de **propension permutationnelle** est compatible avec celui issu de l'ordre des sommes des rangs attribués par les juges aux q objets dans le modèle du W de Kendall.

### 3.6 Apport de l'exploration des rangements des q objets à l'aide de la relation de propension

Les propriétés précédentes laissent penser que finalement la relation de propension ne fait qu'apporter une information redondante à celle fournie par l'application de la procédure W de Kendall. Toutefois elle nous paraît fournir une information plus analytique sur la structure de quasi-ordre affectée d'un niveau de confiance  $1-\alpha$ , qui organise l'ensemble des  $q(q-1)$  paires d'objets, que ne le permet l'approche par la procédure du W de Kendall qui fournit un rangement global synthétique. À partir des définitions posées par J.B. Lagrange (Lagrange 1998 p.82) concernant les notions de la relation de propension, de chemins de propension, de filiation et de variables jumelles, nous cherchons à étudier la structure qui organise les q objets en fonction des rangements. En particulier, nous interprétons la propriété de

covariance positive comme signifiant la tendance pour le rangement des deux objets A et B de recevoir des valeurs attribuées par les juges qui vont dans le même sens.

Dans la procédure de Kendall, le q-uplet des moyennes  $(m_j)_{j=1, \dots, q}$  est l'information à partir de laquelle nous déterminons le rangement synthétique des q objets par les juges. Il est clair que les valeurs de ces q moyennes dont la somme est constante, contiennent le niveau de préférence auquel chaque juge place chaque objet en relation avec les q-1 autres. Or dans la règle de rangement adoptée dans la procédure de Kendall, ce niveau n'est pas pris en compte. C'est alors que dans l'approche par la relation de propension (permutationnelle), nous pouvons voir qu'il est rendu compte de ce phénomène. En effet si nous évaluons

$$\Delta P_{(q)}^{(a,b)} = \frac{\frac{\sum_{i \in P} a_i(1-b_i)}{n} - \frac{\sum_{i \in P} b_i(1-a_i)}{n}}{\sqrt{\frac{20q^4 - 46q^3 + 49q^2 - 41q + 24}{180n(q-1)^4}}} = \frac{\frac{\sum_{i \in P} a_i}{n} - \frac{\sum_{i \in P} b_i}{n}}{\sqrt{\frac{20q^4 - 46q^3 + 49q^2 - 41q + 24}{180n(q-1)^4}}}$$

Cet écart entre les deux coefficients de propension est proportionnel à l'écart entre les deux moyennes. Plus l'écart entre les moyennes est fort, plus les coefficients de propension de a vers b et de b vers a sont différenciés. De là alors que  $m_a \leq m_b$  permet de ranger l'objet B avant l'objet A, la quantité  $\Delta P_{(q)}^{(a,b)} \leq 0$  complète la conclusion en fournissant un autre critère d'aide à l'interprétation basé sur une mesure, à savoir celle de l'intensité de la propension, et le niveau de confiance  $1-\alpha$  dont elle rend compte.

### 3.7 Étude des situations extrêmes

#### 3.7.1 Concordance « parfaite »

La structure à explorer est celle d'un tableau dont les composantes de chaque colonne sont constantes. Cela signifie que tous les juges émettent le même jugement respectivement à chaque objet. Il en résulte que le vecteur  $(m_j)_{j=1, \dots, q}$  des moyennes des q colonnes est égal, à une permutation près, au vecteur  $(0, \frac{1}{q-1}; \frac{2}{q-1}; \dots; \frac{q-2}{q-1}; 1)$  de  $[0; 1]^q$ , tandis que le vecteur  $(v_j)_{j=1, \dots, q}$  des variances des q colonnes est égal au q-vecteur nul. Le calcul des valeurs

possibles de  $s = s(A \Rightarrow B) = \frac{\sum_{i \in P} a_i(1-b_i)}{n}$  met en évidence que l'indice de propension dans le cadre de l'analyse implicative devient :

$$\tilde{q}(a, -b) = \frac{s(A \Rightarrow B) - m_a(1-m_b)}{\frac{m_a(1-m_b)}{\sqrt{n}}} = 0 \text{ si } m_a(1-m_b) \neq 0.$$

Dans la sémantique de cette méthode d'analyse, cette situation où l'on ne peut se prononcer ni sur la propension de a vers b, ni sur sa réciproque correspond à une structure triviale de l'ensemble des rangs et, en fait, à un vide informationnel : si tous les juges sont d'accord, les préférences vont de soi, ce sont des données *a priori*.

Si  $m_a(1-m_b) = 0$ , le rapport a une forme indéterminée  $\frac{0}{0}$ . Cette situation se produit quand l'objet A est au niveau de préférence 0 — le plus bas — pour tous les juges ou quand l'objet B est situé au niveau de préférence 1 — le plus haut — pour tous les juges. Replacé dans le contexte des rangements des  $q$  objets, cela signifie que l'objet B est l'objet préféré — donc  $\text{rang}(B) = 1$  — alors que l'objet A est celui qui est placé en dernier rang de préférence — donc  $\text{rang}(A) = q$  —. Il ressortait de l'approche par la statistique W de Kendall que l'ensemble des objets possédait une structure d'ordre triviale significative. En analyse implicative, l'indétermination est levée par un passage à la limite, dans le cadre de la topologie de Frechet. La valeur de l'indice définie par continuité est nulle et correspond alors à une intensité d'implication statistique égale à 0,50. C'est aussi ce que vaut l'intensité d'implication dans le cas d'indépendance de  $a$  et  $b$ . En ce qui concerne le facteur de variance, il est indéterminé quand  $m_a(1-m_b)=0$  et infini dans les autres cas. Si nous revenons au coefficient de propension intégrant l'approche permutatonnelle, il devient :

$$\tilde{q}_{(q)}(a, -b) = \frac{s(A \Rightarrow B) - \frac{1}{2} + \frac{(q-2)(3q-1)}{12(q-1)^2}}{\sqrt{\frac{20q^4 - 46q^3 + 49q^2 - 41q + 24}{180n(q-1)^4}}} = \frac{m_a(1-m_b) - \frac{1}{2} + \frac{(q-2)(3q-1)}{12(q-1)^2}}{\sqrt{\frac{20q^4 - 46q^3 + 49q^2 - 41q + 24}{180n(q-1)^4}}}$$

Or, la propension de  $a$  vers  $b$  n'est sémantiquement envisageable que si  $0 \leq m_a \leq m_b \leq 1$ . De là  $0 \leq m_a(1-m_b) \leq m_b(1-m_b)$ . Comme  $m_b(1-m_b)$  est toujours au plus égal à  $\frac{1}{4}$ , alors le numérateur de  $\tilde{q}_{(q)}(a, -b)$  est toujours négatif. En effet,  $-\frac{1}{2} + \frac{(q-2)(3q-1)}{12(q-1)^2}$  est supérieur à  $\frac{1}{4}$ . Ce résultat montre que, sur la base de la propension permutatonnelle, il est possible de calculer la qualité numérique de cette propension qui reste significative d'une tendance de  $b$  à dominer généralement  $a$ . Il n'y a donc pas contradiction entre les deux approches, mais l'approche permutatonnelle est plus précise.

### 3.7.2 Hétérogénéité « parfaite »

La structure à explorer est celle d'un tableau dont les composantes de chaque colonne sont telles que leurs sommes sont constantes ou presque à une unité près. Dans cette situation, le vecteur  $(m_j)_{j=1,\dots,q}$  des moyennes des  $q$  colonnes est égal au  $q$ -vecteur constant  $(\frac{1}{2}; \dots; \frac{1}{2})$  ou encore à une permutation près, au  $q$ -vecteur  $(\frac{1}{2 + \frac{1}{2n(q-1)}}; \frac{1}{2 - \frac{1}{2n(q-1)}}; \dots; \frac{1}{2 + \frac{1}{2n(q-1)}}; \frac{1}{2 - \frac{1}{2n(q-1)}})$  de  $[0; 1]^q$

De là nous en déduisons que nécessairement

$$m_a(1-m_b) \in \left\{ \frac{1}{4}; \frac{1}{4} \left( \frac{1}{2n(q-1)} \right)^2; \left( \frac{1}{2 - \frac{1}{2n(q-1)}} \right)^2; \left( \frac{1}{2 + \frac{1}{2n(q-1)}} \right)^2 \right\}$$

Si nous considérons que le vecteur  $(m_j)_{j=1,\dots,q}$  des moyennes des  $q$  colonnes est approximativement égal au  $q$ -vecteur constant  $(\frac{1}{2}; \dots; \frac{1}{2})$ , le coefficient de propension de J.B. Lagrange s'écrit maintenant :

$$\tilde{q}(a, -b) = \frac{\sum_{i \in P} a_i(1-b_i)}{n} - \frac{1}{4} \sqrt{\frac{(v_a + \frac{1}{4})(v_b + \frac{1}{4})}{n}}$$

### 3.7.3 Information apportée par la covariance : $\text{cov}(a; b)$

Dans l'approche de J.B. Lagrange, le coefficient de propension peut s'exprimer en fonction de la covariance par la formule :  $\tilde{q}(a, -b) = \frac{-\text{cov}(a;b)}{\sqrt{\frac{(v_a + m_a^2)(v_b + (1-m_b)^2)}{n}}}$  d'où

**Proposition 3 :**  $\Phi(-\tilde{q}(a, -b)) \geq 0,5 \Leftrightarrow \text{cov}(a;b) \geq 0$ .

La covariance de  $a$  et  $b$  contient une part importante de l'information relative à la relation de propension de  $a$  vers  $b$ . Dans notre approche permutationnelle, la relation n'est plus aussi directe. En effet nous avons :

$$\tilde{q}_{(q)}(a, -b) = \frac{m_a(1-m_b) - \text{cov}(a;b) - \left[ \frac{1}{4} + \frac{(q+1)}{12(q-1)^2} \right]}{\sqrt{\frac{20q^4 - 46q^3 + 49q^2 - 41q + 24}{180n(q-1)^4}}}$$

Si nous tenons compte du fait que les valeurs prises par  $a$  et  $b$  sont dans  $[0; 1]$ , nous pouvons déduire quelques informations relatives aux encadrements de diverses expressions :

**Proposition 4 :**  $0 \leq m_a(1-m_b) \leq 1$  et  $-1 \leq \text{cov}(a; b) \leq 1$

Par ailleurs sachant que  $\frac{1}{4} < \frac{1}{4} + \frac{(q+1)}{12(q-1)^2} \leq \frac{1}{2}$  quand le nombre d'objets à ranger est supérieur à 2, alors le coefficient de propension peut être ainsi encadré :

**Proposition 5 :** si  $q > 2$ ,

$$-\frac{3}{2} \sqrt{\frac{20q^4 - 46q^3 + 49q^2 - 41q + 24}{180n(q-1)^4}} \leq \tilde{q}_{(q)}(a, -b) < \frac{7}{4} \sqrt{\frac{20q^4 - 46q^3 + 49q^2 - 41q + 24}{180n(q-1)^4}}$$

Il advient aussi que la condition  $\text{cov}(a;b) \geq 0$  n'est plus une condition nécessaire et suffisante pour que le coefficient de propension soit négatif. En effet elle n'est plus

nécessaire car, dans l'exemple que nous aborderons ci-dessous, nous rencontrons des cas (Annexe) où  $cov(a ; b) < 0$  et  $\tilde{q}_{(q)}(a, -b) < 0$  — ainsi dans notre étude  $cov(V45E ; V45C) < 0$  et  $\tilde{q}_{(q)}(a, -b) = -3,3491 < 0$  —. Elle n'est pas non plus suffisante puisque nous avons aussi  $cov(a ; b) > 0$  et  $\tilde{q}_{(q)}(a, -b) > 0$  — ainsi dans notre étude  $cov(V45A ; V45J) > 0$  et  $\tilde{q}_{(q)}(a, -b) = 1.4981 > 0$  — Nous considérons alors la condition  $cov(a ; b) \geq 0$  comme une condition qui réduit la relation de propension aux couples d'objets tels que les variables a et b soient corrélées positivement, prenant en compte ainsi le fait que les valeurs attribuées aux deux objets vont dans le même sens. Par là même, elle se rapporte à la même classe de variables modales mise en œuvre dans la situation que celle déterminée dans l'approche de J.B. Lagrange.

#### 4 Analyse statistique implicative comme outil d'exploration de la structure des rangements (cas du rangement incomplet et sans *ex æquo*)

Nous allons maintenant considérer le cas où il est demandé aux individus de procéder à un choix de t objets parmi les q proposés, puis de les trier par ordre décroissant de préférence. Nous pouvons considérer alors que le résultat est une réalisation du type du tableau TAB. 1, mais avec seulement t (t < q) cases sur q comportant une des permutations (1, 2, ..., t). Nous proposons alors la modélisation suivante qui nous ramène à une forme traitable selon l'approche de W de Kendall. Les individus choisissent t objets et les rangent de 1 à t. Après quoi, nous postulons que cela reviendrait à dire que les (q-t) objets non choisis sont placés en *ex æquo* aux rangs t+1 à t-(q-t) = q. Nous leur attribuons alors le rang

moyen  $\frac{1}{q-t} \sum_{j=1}^{j=q-t} (t+j) = \frac{t+1+q}{2}$ . Nous appliquons alors la procédure du W de Kendall en

prenant en compte l'existence de ces *ex æquo*. Cette prise en compte des *ex æquo* modifie la valeur  $S_{K_{max}}$  obtenue dans le cas de la concordance parfaite entre les k juges qui fournissent tous le même rangement, c'est à dire à une permutation près (1, 2, 3, ..., t,  $\frac{t+1+q}{2}$ , ...,  $\frac{t+1+q}{2}$ ). Il vient donc :  $S_{K_{max}} = \frac{k^2}{12} [q^3 - (q-t)^3 - t]$  De là nous obtenons la

statistique  $W_{(q,t)}^* = \frac{\sum_{j=1}^q \left( S(O_j) - k \frac{(q+1)}{2} \right)^2}{\frac{k^2}{12} [q^3 - (q-t)^3 - t]}$  que nous utiliserons de la même façon que la

statistique W de Kendall. Notons que si t=q-1 nous retrouvons le cas du rangement complet.

##### 4.1 Variable modale associée au rangement de t objets parmi q réalisé par k individus

Nous adoptons la transformation qui définit la variable modale par la formule suivante :

$$X_{i,j}^{(q,t)} = \frac{(t+1) - R_{i,j}^q}{t} I_{\{1,2,\dots,t\}}(R_{i,j}^q)$$

## 4.2 Expression du coefficient de propension pour des variables de rangs (cas du rangement incomplet sans *ex aequo*)

Dans ce contexte particulier, le coefficient de propension s'exprime alors de la façon suivante où  $t$  est le nombre d'objets à ranger parmi les  $q$  objets soumis au jugement de  $n$  individus :

$$\tilde{q}_{(q,t)}(a, -b) = \frac{m_a(1 - m_b) - \text{cov}(a; b) - \frac{(t+1)(2t+1)}{6tq}}{\sqrt{\frac{(t+1)[(t-1)(-40t^3 - 36t^2 - 5t - 6) + 30t^2(2t+1)(q-1)]}{180nq(q-1)t^3}}}$$

Nous constatons que pour  $t=q-1$ ,  $\tilde{q}_{(q,q-1)}(a, -b) = \tilde{q}_{(q)}(a, -b)$

## 5 Application à l'exploration des représentations des étudiants à l'égard des objectifs d'enseignement de la statistique

Dans une recherche que J.C Régnier (Régnier, 2002) a conduite sur la thématique de l'enseignement-apprentissage de la statistique dans un enseignement à distance, les représentations des étudiants relativement à la statistique ont été explorées, en particulier, à partir des rangements des objectifs déclarés de son enseignement. Durant l'année 2001-2002, un échantillon de 125 étudiants concernés a été soumis à un questionnaire dont une question consistait à ranger les 10 objectifs de l'enseignement selon un ordre décroissant d'importance. De cet ensemble, nous avons obtenu un sous-échantillon d'observation de 73 sujets dont nous traitons les réponses fournies relativement au rangement des 10 objectifs suivants :

<i>Énoncés des objectifs</i>	<i>Rang</i>
V45A) d'expliciter les questions d'une problématique dont les réponses relèvent d'une approche statistique,	
V45B) de décrire, traiter, analyser des données de manière pertinente dans le cadre d'une étude en particulier dans le domaine éducatif,	
V45C) de faire le lien entre la réflexion analytique sur des questions relevant du champ de l'éducation, leur formalisation et leur traitement quantitatif,	
V45D) de lire avec un regard critique et distancié, les conclusions de diverses études statistiques apparaissant dans des rapports de recherche en Sciences de l'Éducation,	
V45E) de poursuivre de façon autonome et personnalisée un apprentissage en statistique afin d'enrichir ses acquis,	
V45F) de poser un regard plus positif à l'égard d'un domaine largement exploité dans les <i>media</i> , dans le sens de ne pas considérer les résultats dans l'ordre du tout ou rien mais en les replaçant judicieusement dans leur domaine de validité,	
V45G) d'exploiter des notions et des démarches mathématiques à des fins d'outils, et de ce fait de modifier dans un sens positif le rapport souvent négatif que nombre entretient avec cette science,	
V45H) de s'exercer à un raisonnement intégrant l'idée de "risque d'erreur" dans l'énoncé de ses conclusions.	
V45I) de s'exercer à l'interprétation de phénomènes éducatifs sur la base de données statistiques sur des "faits éducatifs" et sur des relations entre ces "faits"	
V45J) de s'exercer à la communication des résultats des analyses des données en distinguant clairement le modèle utilisé, de la réalité qu'il est supposé représenter, en séparant bien les traitements menés à l'intérieur du modèle, des interprétations reformulées dans le contexte du problème.	

TAB. 6 - *Tableau des q=10 items à ranger par ordre d'importance décroissante.*

Le calcul des sommes  $S(O_j)_{j=1,\dots,10}$  donne les résultats suivant :

Objectifs	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10
	V45A	V45B	V45C	V45D	V45E	V45F	V45G	V45H	V45I	V45J
Sommes	322	276	338	349	552	408	459	500	392	419
Rangs issus de l'ordre des sommes	2	1	3	4	10	6	8	9	5	7

TAB. 7 - *Tableau des q=10 items rangés par ordre d'importance décroissante selon l'approche W de Kendall.*

La mise en œuvre de la procédure du test W de Kendall fournit les informations suivantes : la réalisation empirique de la statistique W vaut  $w \approx 0,14775$  tandis que la valeur critique au seuil de  $\alpha = 0.01$  vaut  $w_c \approx 0,03298$ .

Elles conduisent à rejeter l'hypothèse  $H_0$  d'indépendance ou d'hétérogénéité des rangements fournis par les étudiants et à admettre à un niveau  $\alpha=0.01$  de risque de 1<sup>ère</sup> espèce qu'il existe une tendance à s'accorder sur l'ordre d'importance des objectifs :

1	2	3	4	5	6	7	8	9	10
(O2	O1	O3	O4	O9	O6	O10	O7	O8	O5)
(V45B	V45A	V45C	V45D	V45I	V45F	V45J	V45G	V45H	V45E)

### 5.1 Exploration des 73 rangements des 10 objectifs par l'analyse de la propension statistique

Nous avons transformé le tableau des rangs par la conversion  $X_{ij}^{(q)} = \frac{q - R_{ij}^{(q)}}{q - 1}$  afin d'obtenir un tableau dont les valeurs sont les réalisations de variables modales dans  $[0 ; 1]$ . Nous avons ensuite soumis ce tableau de données à la procédure de l'analyse de la propension avec le logiciel CHIC 2.3 (Couturier 2001).

Items	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10
Objectifs	V45A	V45B	V45C	V45D	V45E	V45F	V45G	V45H	V45I	V45J
Occurrence	45,333	50,444	43,556	42,333	19,778	35,778	30,111	25,556	37,556	34,556
Moyenne	0,621	0,691	0,597	0,580	0,271	0,490	0,412	0,350	0,514	0,473
Ecart Type	0,264	0,256	0,254	0,308	0,298	0,333	0,342	0,271	0,269	0,334
Variance	0,070	0,065	0,064	0,095	0,089	0,111	0,117	0,073	0,073	0,111
Facteur de variance	3,3642	3,2665	3,7348	2,5700	2,2291	2,2537	2,0739	3,1046	3,4428	2,2400

TAB. 8 - Tableau des valeurs caractéristiques des variables modales Objectifs.

Nous obtenons le graphe « implicatif » suivant :

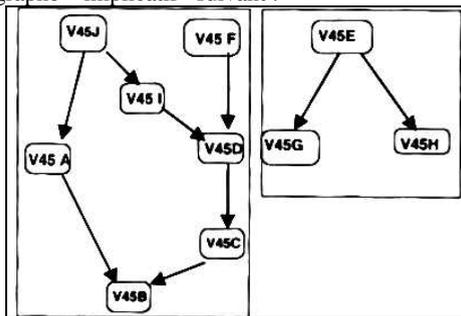


FIG. 1 – Graphe implicatif à partir de l'approche JB Lagrange de la relation de propension.

Les valeurs prises en compte pour construire ce graphe correspondent aux 9 cas où l'intensité de propension est supérieure à 0,50 — en d'autres termes ceux où le coefficient de propension est positif — et avec lesquelles nous procédons à l'identification des couples

(a ; b) vérifiant la **relation de propension**  $a \Re b \Leftrightarrow \begin{cases} \Phi(-\tilde{q}(a,-b)) \geq 1-\alpha = 0.5 \\ \Phi(-\tilde{q}(a,-b)) \geq \Phi(-\tilde{q}(b,-a)) \end{cases}$  et possédant la propriété de la covariance positive.

Le relation de propension est compatible avec le rangement issu de la procédure W de Kendall. Mais elle s'avère plus informative car, d'une part, l'ordre est partiel et, d'autre part, les relations de préférences sont valuées

Objectifs	O2	O1	O3	O4	O9	O6	O10	O7	O8	O5
	V45B	V45A	V45C	V45D	V45I	V45F	V45J	V45G	V45H	V45E
Rangs	1	2	3	4	5	6	7	8	9	10

Voici les chemins de propension :

*Chemin JAB* : V45J [rang =7]  $\Rightarrow$  V45A [rang =2]  $\Rightarrow$  V45B [rang =1]

*Chemin JIDCB* : V45J [rang =7]  $\Rightarrow$  V45I [rang =5]  $\Rightarrow$  V45D [rang =4]  $\Rightarrow$  V45C [rang =3]  $\Rightarrow$  V45B [rang =1]

*Chemin FDCB* V45F [rang =6]  $\Rightarrow$  V45D [rang =4]  $\Rightarrow$  V45C [rang =3]  $\Rightarrow$  V45B [rang =1]

*Chemin EG* V45E [rang =10]  $\Rightarrow$  V45G [rang =8]

*Chemin EH* V45E [rang =10]  $\Rightarrow$  V45H [rang =9]

Nous constatons le respect de la propriété de la propension à ce que « l'objet  $O_{j_1}$  étant rangé au rang  $r_1$  », « l'objet  $O_{j_2}$  est rangé au rang  $r_2$  qui traduit la préférence de cet objet  $O_{j_2}$  sur l'objet  $O_{j_1}$  ». Nous n'abordons pas ici l'approche sémantique que requiert l'interprétation de ces résultats.

## 5.2 Approche permutatonnelle de la propension de X vers Y

Dans l'approche que nous avons nommée « permutatonnelle » en raison du fait que la base de l'information est une permutation sur  $\{1, \dots, q\}$ , nous suivons le même raisonnement. La seule différence réside dans le fait que le calcul de l'espérance et de la variance de la statistique Z utilisée par J.B. Lagrange tient compte de la particularité de l'information fondée sur la statistique de rangs. Nous obtenons ainsi les résultats relatifs aux 90 couples de variables. 43 couples sont en relation de propension au niveau de confiance  $1-\alpha = 0,50$ . Pour calculer l'intensité de propension, nous utilisons encore la référence à une mesure gaussienne qui demeure monotone par rapport aux coefficients de propension permutatonnelle. Nous ne retenons cependant que les couples en relation au niveau  $1-\alpha \geq 0,90$ . La réalisation du graphe implicatif avec le critère de la covariance positive conduit à la représentation des relations de propension suivante :

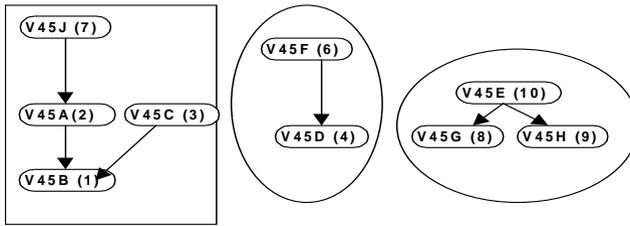


FIG. 2 –: *graphe implicatif dans l'approche permutationnelle de la relation de propension*

Si nous explicitons les chemins de propension nous obtenons alors :

*Chemin JAB* V45J [rang 7]  $\Rightarrow$  V45A [rang 2]  $\Rightarrow$  V45B [rang 1]

*Chemin CB* V45C [rang 3]  $\Rightarrow$  V45B [rang 1] ;

*Chemin FD* V45F[rang 6]  $\Rightarrow$  V45D[rang 4]

*Chemin EG* V45E [rang 10]  $\Rightarrow$  V45G [rang 8] ;

*Chemin EH* V45E [rang 10]  $\Rightarrow$  V45H [rang 9]

Comme nous l'avons démontré, cette structure définie par la relation de propension — *versus* approche permutationnelle — respecte la propriété établissant que « l'objet  $O_{j1}$  étant rangé au rang  $r_1$  », « l'objet  $O_{j2}$  est rangé au rang  $r_2$  qui traduit la préférence de cet objet  $O_{j2}$  sur l'objet  $O_{j1}$  »

## 6 Comparaison des résultats issus des deux approches

Nous avons calculé les écarts des intensités de propension entre les deux modèles (Annexe 4). Ce calcul met en évidence l'incidence sur la valeur de l'intensité de propension quand on ne tient pas compte des contraintes du modèle permutationnel. En moyenne, l'intensité de propension permutationnelle dépasse celle au sens de Lagrange, d'une valeur d'environ 0,13. Toutefois cette intensité de propension permutationnelle reste inférieure à celle au sens de Lagrange, pour 39 couples sur 90, soit dans plus de 43% des cas. Notons enfin que les propriétés topologiques du graphe implicatif, définies par les plus fortes intensités, sont peu affectées par ces différences, les principales liaisons étant donc respectées.

## 7 Conclusion

Nous avons rappelé que l'objectif du test de rang de Kendall consistait à établir un degré de concordance-discordance entre les jugements de  $k$  juges portant sur  $q$  objets, mais que sa référence à une loi de probabilité était délicate et que le résultat se limitait à rejeter l'hypothèse nulle. De même, dans le cas de rangement complet (tous les objets sont rangés) ou incomplet (certains peuvent ne pas l'être), la méthode, dite de propension, appliquée aux variables modales par J.B. Lagrange dans le cadre de la méthode implicative, ne pouvait être compatible avec la notion de rangement qui oblige à relaxer l'indépendance entre les variables. Aussi, privés de cette propriété, nous avons établi les valeurs des paramètres des

nouvelles variables en jeu dans une approche de la même propension exprimée de façon raccourcie par « l'objet b est généralement préféré à l'objet a ». Nous avons vu que la seule covariance, par son signe, n'était plus ni nécessaire ni suffisante pour conjecturer une relation de propension entre objets. Toutefois, une même référence à une mesure gaussienne pour déterminer l'intensité de propension dans les deux approches (permutationnelle et celle de Lagrange) nous a permis alors de comparer, sur une application, les structures d'ordre partiel sur l'ensemble des objets pour les deux relations de propension. On a observé que les différences étaient faibles, particulièrement lorsque les intensités de propension étaient fortes, c'est-à-dire lorsque les relations de préférences étaient nettes.

Une nouvelle question se pose alors : les différences entre les deux modèles sont-elles atténuées par l'admission des *ex æquo* où, cette fois, l'hypothèse d'indépendance entre les variables pourrait peut-être y trouver une nouvelle légitimité ?

## Références

- Bailleul, M., (1994) *Analyse statistique implicative : variables modales et contribution des sujets. Application à la modélisation de l'enseignant dans le système didactique*. Thèse Université de Rennes 1
- Couturier, R.,(2001) Traitement de l'analyse statistique implicative dans CHIC , Actes des Journées « Fouille dans les données par la méthode d'analyse implicative », IUFM de Caen, 33-50.
- Friedman M. (1937) The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. Statist. Ass.*, 32, 675-701
- Gras, R., (1979) *Contribution à l'étude expérimentale et à l'analyse de certaines acquisitions cognitives et de certains objectifs didactiques en mathématiques*, Thèse d'Etat, Université de Rennes 1.
- Gras, R., & al., (1996) *L'implication Statistique*. Grenoble: La Pensée Sauvage,
- Gras, R., P. Kuntz et H. Briand, (2001) Les fondements de l'analyse statistique implicative et leurs prolongements pour la fouille de données, *Math. et Sc. Humaines*, n°154-155, pp. 9-29.
- Kendall M.G., et B. Babington Smith (1939)The problem of *m* rankings. *Ann. Math. Statist.*,10, 275-287
- Lagrange J.B. (1998) Analyse implicative d'un ensemble de variables numériques; application au traitement d'un questionnaire aux réponses modales ordonnées, *Revue de Statistique Appliquée* 46 (1) : 71-93.
- Lerman I.C., R. Gras et H. Rostam (1981a) Élaboration et évaluation d'un indice d'implication pour des données binaires, I et II, *Mathématiques et Sciences Humaines*, 74, 5-35 et 75, 5-47.
- Lerman I.C. (1981b) *Classification et analyse ordinale des donnée*. Paris: Dunod,

Régnier J.C. (2006). Étude des difficultés d'apprentissage de la statistique dans le cadre d'un enseignement à distance. *Revue Eduquer : Psychologie et Sciences de l'Éducation*. 14, 15-47

Régnier J.C. et R. Gras (2005). Statistique de rangs et analyse statistique implicative. *Revue de Statistique Appliquée*. 53 (1) : 5-38

## Annexe

Annexe 1		Intensité de propension de X vers Y (JB Lagrange)									
(X⇒Y)	V45A	V45B	V45C	V45D	V45E	V45F	V45G	V45H	V45I	V45J	
V45A		<b>0,7025</b>									
V45B			<b>0,5816</b>								
V45C				<b>0,5825</b>							
V45D											
V45E							<b>0,7953</b>	<b>0,6263</b>			
V45F				<b>0,6708</b>							
V45G											
V45H											
V45I				<b>0,5587</b>							
V45J	<b>0,5190</b>								<b>0,5837</b>		

Annexe 2		Corrélation (X ;Y)									
(X⇒Y)	V45A	V45B	V45C	V45D	V45E	V45F	V45G	V45H	V45I	V45J	
V45A	1	<b>0,249</b>	-0,125	-0,237	-0,190	-0,233	-0,152	-0,090	-0,064	<b>0,016</b>	
V45B		1	<b>0,096</b>	-0,205	-0,020	-0,336	-0,138	-0,179	-0,159	-0,077	
V45C			1	<b>0,097</b>	-0,111	-0,114	-0,134	-0,224	-0,133	-0,183	
V45D				1	-0,389	<b>0,155</b>	-0,182	-0,142	<b>0,063</b>	-0,207	
V45E					1	-0,232	<b>0,259</b>	<b>0,132</b>	-0,253	-0,218	
V45F						1	-0,071	-0,303	-0,005	-0,080	
V45G							1	-0,010	-0,341	-0,402	
V45H								1	-0,120	-0,007	
V45I									1	<b>0,088</b>	
V45J										1	

Annexe 3 Intensité de propension de X vers Y (approche permutatonnelle)  
 [ $1-\alpha = 0.50$ ] et grisé [ $1-\alpha = 0.90$ ]

Les valeurs marquées avec (+) correspondent aux couples de variables dont la covariance est positive

(X⇒Y)	V45A	V45B	V45C	V45D	V45E	V45F	V45G	V45H	V45I	V45J
V45A		0,977+								
V45B										
V45C	0,733	0,973+								
V45D	0,697	0,938	0,793							
V45E	0,999	0,999	0,999	0,995		0,990	0,998+	0,987+	0,994	0,988
V45F	0,900	0,970	0,895	0,952+					0,703	
V45G	0,983	0,997	0,973	0,945					0,755	
V45H	0,997	0,999	0,992	0,991		0,902	0,898		0,972	0,962
V45I	0,925	0,983	0,851	0,880						
V45J	0,974+	0,994	0,899	0,831		0,601			0,821	

Annexe 4 Écart entre l'intensité de propension de X vers Y au sens de l'approche permutatonnelle et celle au sens de Lagrange ( moyenne = 0,13451)

(X⇒Y)	V45A	V45B	V45C	V45D	V45E	V45F	V45G	V45H	V45I	V45J
V45A		0,275	0,111	0,010	-0,404	-0,296	-0,395	-0,453	-0,310	-0,445
V45B	-0,015		-0,163	-0,213	-0,491	-0,289	-0,418	-0,419	-0,385	-0,445
V45C	0,328	0,392		0,088	-0,443	-0,307	-0,400	-0,386	-0,223	-0,315
V45D	0,404	0,638	0,210		-0,277	-0,299	-0,346	-0,411	-0,185	-0,264
V45E	0,754	0,532	0,645	0,922		0,779	0,203	0,361	0,776	0,758
V45F	0,639	0,819	0,510	0,281	-0,333		-0,222	-0,263	0,208	0,028
V45G	0,666	0,682	0,626	0,667	-0,322	0,425		-0,062	0,571	0,362
V45H	0,604	0,724	0,726	0,662	0,048	0,709	0,410		0,592	0,470
V45I	0,483	0,640	0,462	0,321	-0,351	-0,009	-0,202	-0,399		-0,089
V45J	0,455	0,590	0,584	0,558	-0,335	0,187	-0,087	-0,358	0,237	

## Summary

This communication explores the benefit of one data analysis method implicative analysis such as Gras defined it to study ranks concordance-discordance accorded by judges to things. This meaning is the same as in Friedman's or Kendall's works. Then we compare one analysis of judges' preferences through ranks to the propensity analysis between modal variables built by Lagrange. To achieve it, we don't consider the a priori independence hypothesis between variables. Nevertheless, we give a quality measure to "theorem" : "if an A thing is ranked by judges, then generally a B thing is ranked in a better way by the same judges" and we represent by graph relations of preferences for the whole ranked objects. We limit this study to the two cases of complete ranking and incomplete ranking, without ex æquo between q things by k judges.

