MATHÉMATIQUES ET SCIENCES HUMAINES

V. GIAKOUMAKIS

B. MONJARDET

Coefficients d'accord entre deux préordres totaux. Comparaison ordinale des coefficients

Mathématiques et sciences humaines, tome 98 (1987), p. 69-87

<http://www.numdam.org/item?id=MSH_1987__98__69_0>

Math. Sci. hum.(25^{e}année, n°98, *1987*, p.69-87)

COEFFICIENTS D'ACCORD ENTRE DEUX PREORDRES TOTAUX COMPARAISON ORDINALE DES COEFFICIENTS

V. GIAKOUMAKIS*, B. MONJARDET**

1. INTRODUCTION

Dans un texte précédent (Giakoumakis et Monjardet, 1986) nous avons étudié quinze coefficients d'accord entre deux préordres totaux dont douze généralisaient le classique tau de Kendall entre deux ordres totaux, tandis que trois généralisaient le non moins classique rho de Spearman (en particulier tous ces coefficients varient entre -1 et $+1$). Dans ce texte nous complétons ce travail en étudiant d'une part les inégalités entre ces coefficients, d'autre part les préordonnances associées. Dans le premier cas, nous considérons d'abord les inégalités "élémentaires" entre les douze coefficients généralisant tau , puis celles entre les trois coefficients généralisant rho . Il n'a pas été jugé utile de chercher d'éventuelles inégalités entre ces deux types de coefficients, puisque l'on sait que dans le cas des ordres totaux, il n'existe pas d'inégalité simple entre tau et rho (cf., par exemple, Kendall 1970 et Monjardet et Leconte de Poly Barbut 1986). Dans le second cas nous montrons que les préordonnances (totales) associées à sept coefficients d'accord généralisant tau sont toutes différentes, ce qui signifie qu'étant donné deux tels coefficients σ et σ' on peut toujours trouver deux couples (P_1,P_2) et (P_3,P_4) de préordres totaux pour lesquels $\sigma(P_1,P_2)\leq\sigma(P_3,P_4)$ tandis que $\sigma'(P_1,P_2) > \sigma'(P_3,P_4)$.

Nous avons parlé plus haut d'inégalité élémentaire entre deux coefficients σ et σ' ; nous entendons par là que pour tout couple de préordres totaux l'inégalité $\sigma(P_1,P_2) \leq \sigma'(P_1,P_2)$(ou l'inégalité inverse) est vérifiée.

Plus généralement, nous considérons un domaine $\mathcal{C}$ formé par un ensemble de couples de préordres totaux, et un ensemble Σ de coefficients d'accords

* Université Paris 5
** Université Paris 5 et Centre d'Analyse et de Mathématique Sociale

définis sur Θ . Nous définissons une relation d'ordre $\underset{\Theta}{\leq}$ sur Σ en posant pour $\sigma,\sigma' \in \Sigma$:

$$\sigma \underset{\Theta}{\leq} \sigma' \Leftrightarrow \text{pour tout } (P_1,P_2) \in \Theta \text{ , } \sigma(P_1,P_2) \leq \sigma'(P_1,P_2)$$

L'ordre strict correspondant est défini par $\sigma \underset{\Theta}{<} \sigma' \Leftrightarrow \sigma \underset{\Theta}{\leq} \sigma'$ et il existe (P_1,P_2) tel que $\sigma(P_1,P_2) < \sigma'(P_1,P_2)$. Nous disons que σ et σ' sont incomparables pour l'ordre $\underset{\Theta}{\leq}$ si $\sigma \underset{\Theta}{\nleq} \sigma'$ et $\sigma' \underset{\Theta}{\nleq} \sigma$, i.e. s'il existe un couple (P_1,P_2) de préordres totaux avec $\sigma(P_1,P_2) < \sigma'(P_1,P_2)$ et un autre couple avec l'inégalité inverse; dans un tel cas nous notons $\sigma \underset{\Theta}{\|} \sigma'$. Lorsque l'ensemble Θ aura été clairement précisé, nous allègerons les notations en remplaçant par exemple $\underset{\Theta}{\leq}$ par $\leq$.

Nous représentons l'ordre ainsi obtenu sur Σ par son classique diagramme (cf. par exemple, Barbut et Monjardet 1970). Dans un tel diagramme les coefficients sont représentés par des points du plan. Le point représentant un coefficient σ est relié par un segment de droite "montant" à tout point représentant un coefficient σ' qui "couvre" σ (i.e. $\sigma < \sigma'$ et il n'existe pas σ'' avec $\sigma < \sigma'' < \sigma'$). Dans un tel diagramme le coefficient σ est inférieur au coefficient σ' si et seulement si, il existe un "trajet" ascendant de σ jusqu'à σ' ; dans le cas contraire σ et σ' sont comparables.

Donnons maintenant un résumé du contenu de ce papier. A la section 2 nous rappelons les notions indispensables portant sur les préordres totaux en précisant les notations utilisées dans ce texte. La section 3 est consacrée à l'ordre entre les douze coefficients généralisant tau , ordre étudié pour divers domaines de couples de préordres totaux. Nous considérons d'abord l'ensemble noté $\mathcal{P}^{2*}$ des couples (P_1,P_2) pour lesquels P_1 et P_2 sont différents du préordre "universel" X^2 (où tous les éléments de X sont équivalents); cette légère restriction provient de ce que si P_1 ou P_2 égalent X^2 , les valeurs de certains de ces douze coefficients peuvent être indéterminées; nous éliminons donc ce cas qu'il serait fastidieux d'examiner en détail. Les conclusions de cette étude peuvent se lire sur la figure représentant le diagramme de l'ordre partiel entre les douze coefficients, ce diagramme est "complet" à une exception près : nous conjecturons sans voir pu le démontrer en général, qu'on a toujours $\tau_b \leq \tau_4$; cette conjecture est représentée sur la figure par un trait pointillé entre τ_b et τ_4 . Dans tous les autres cas l'ordre (qui peut être l'incomparabilité) entre deux coefficients a été déterminé; ainsi sur les 66 relations possibles entre les douze coefficients il y a 25 (ou 26) inégalités strictes et 41 (ou 40) incomparabilités. Dans le reste de cette section on détermine de même l'ordre partiel entre les douze coefficients, mais pour les domaines restreints suivants de couples de préordres totaux : ensemble des couples pour lesquels le nombre d'accords stricts moins le nombre de désaccords stricts entre les deux préordres totaux est soit positif ou nul, soit négatif; ensemble des couples (P_1,P_2) de préordres totaux pour lesquels soit $P_1 \subset P_2$, soit $P_1 \subset P_2^r$ (préordre total réciproque de P_2) ; évidemment dans ces deux derniers cas beaucoup de coefficients prennent la même valeur, et par exemple, dans le dernier cas, on obtient un ordre total entre les coefficients dont les valeurs diffèrent.

Dans la section 4, après avoir rappelé l'expression des trois coefficients généralisant rho , nous étudions les inégalités entre eux.

Dans la dernière section, on montre que certains des coefficients étudiés sont "ordinalement" différents en ce sens que les préordonnances qu'on peut leur associer sont différentes. On s'est limité ici à montrer ce résultat pour sept des coefficients généralisant tau , mais il pourrait être étendu à tous les coefficients considérés dans cet article.

Les références finales sont limitées, puisque les problèmes étudiés ici l'ont été très peu auparavant, et que toutes les références concernant les coefficients d'accord considérés peuvent être trouvées dans Giakoumakis et Monjardet (1968).

2. RAPPELS ET NOTATIONS

Dans tout ce texte, X désigne un ensemble fini, à n éléments.

Un *préordre total* P sur X est une relation binaire sur X (c'est-à-dire une partie de X^2), réflexive ($(x,x) \in P$ pour tout $x \in X$) , transitive ($(x,y) \in P$ et $(y,z) \in P$ impliquent $(x,z) \in P$) et totale ($(x,y) \notin P$ implique $(y,x) \in P$).

Si de plus P est antisymétrique ($(x,y) \in P$ et $(y,x) \in P$ impliquent $x = y$) P est un *ordre total*.

Les propriétés classiques des préordres totaux sont bien connues (cf. par exemple, Barbut et Monjardet 1970). Nous introduisons les notations utilisées dans ce texte en rappelant quelques faits fondamentaux :

. Tout préordre total P est égal à $I + 0$ où

$I = \{(x,y) \in X^2$ tel que $(x,y) \in P$ et $(y,x) \in P\}$

est la partie symétrique de P ,

$0 = \{(x,y) \in X^2$ tel que $(x,y) \in P$ et $(y,x) \notin P\}$

est la partie asymétrique de P , et où

$+$ est la notation pour la réunion de deux ensembles disjoints.

. La partie I de P est une relation d'équivalence (i.e. une relation réflexive, symétrique et transitive) à laquelle est associée une partition π de X . Par définition, les classes du préordre P , sont les classes de cette partition π . On utilisera parfois la notation $\mathring{I} : \mathring{I} = I - \{(x,x), x \in X\}$.

. Le préordre total P induit un ordre total sur les classes de π donc une *partition ordonnée* de X (i.e. une partition π de X munie d'un ordre total sur les classes de π , noté $\underset{\pi}{\leq}$). Inversement, à toute partition ordonnée de X , notée $(\pi, \underset{\pi}{\leq})$ correspond un préordre total sur X (cette correspondance étant bijective).

Ce fait justifie qu'un préordre total P soit souvent noté par sa partition ordonnée associée i.e. par $C_1 < C_2 < \ldots < C_t$ (où les C_i sont les classes de P , ordonnée suivant l'ordre $\underset{\pi}{\leq}$). En particulier si toutes les classes C_i ont un seul élément, i.e. si P est un ordre total, on le notera $x_1 < x_2 < \ldots < x_n$ ou plus simplement $x_1 x_2 \ldots x_n$, $x_i \in X$.

Inversement, si P n'a qu'une seule classe X , cela signifie qu'il est composé de tous les couples de X , et on le notera donc X^2 .

<u>Partition et paramètres fondamentaux associés à une paire de préordres totaux</u>

Soient $P = O + I$, $P' = O' + I'$ deux préordres totaux sur X . On a :

$X^2 = O + I + O^r = O' + I' + O'^r$ où O^r (respectivement O'^r) désigne la relation réciproque de O (respectivement de O') :

$O^r = \{(x,y) \in X^2 \,/\, (y,x) \in O\}$.

On appelle partition (fondamentale) associée à la paire P,P' , la partition de X^2 définie dans le tableau 1 ci-dessous; compte tenu des égalités entre les effectifs de certaines classes de cette partition, ces effectifs définissent cinq nombres différents (au plus), dénotés a, b_1, b_2, c et d :

	O'	I'	O'^r
O	$a = \lvert O \cap O' \rvert$	$b_2 = \lvert O \cap I' \rvert$	$d = \lvert O \cap O'^r \rvert$
I	$b_1 = \lvert I \cap O' \rvert$	$c = \lvert I \cap I' \rvert$	$b_1 = \lvert I \cap O'^r \rvert$
O^r	$d = \lvert O^r \cap O' \rvert$	$b_2 = \lvert O^r \cap I \rvert$	$a = \lvert O^r \cap O'^r \rvert$

P P^r

<u>Tableau 1.</u>

En posant $b = b_1 + b_2$ et $c^* = \frac{c-n}{2}$ on a :

$$n^2 = 2(a+b+d) + c \quad \text{et} \quad n(n-1) = 2(a+b+c^*+d) .$$

Les effectifs des classes de cette partition jouent le rôle des paramètres fondamentaux pour calculer un grand nombre de coefficients d'accord entre P et P' .

On trouvera l'expression des douze coefficients généralisant tau étudiés ici au début du paragraphe suivant et celle des trois coefficients généralisant rho au paragraphe 4.1.

3. COMPARAISONS ORDINALES DE DOUZE COEFFICIENTS GENERALISANT LE TAU DE KENDALL

3.1. Expression des douze coefficients

Tous ces coefficients s'expriment en fonction des paramètres associés à un couple de préordres totaux, définis dans la section précédente. Nous donnons ci-dessous les formules correspondantes; une expression telle que $\tau_a = \frac{2(a-d)}{n(n-1)}$ signifie que $\tau_a(P,P') = \frac{2[a(P,P') - d(P,P')]}{n(n-1)}$.

$$\tau_1 = \frac{2(a-d) + 2c^* - 2b}{n(n-1)} \qquad \tau_2 = \frac{2(a-d) + 2c^* - b}{n(n-1)}$$

$$\tau_3 = \frac{2(a-d) + 2c^*}{n(n-1)} \qquad \tau_4 = \frac{2(a-d) + 2c^* + b}{n(n-1)}$$

$$\tau_5 = \frac{2(a-d) + 2c^* + 2b}{n(n-1)}$$

$$\tau_F = \frac{2(a-d) - 2c^*}{n(n-1)} \qquad \tau_a = \frac{2(a-d)}{n(n-1)}$$

$$e = \frac{a-d}{a+b+d} \qquad d_a = \frac{a-d}{a+b_2+d} \qquad d_b = \frac{a-d}{a+b_1+d}$$

$$\tau_b = \frac{a-d}{[(a+b_2+d)(a+b_1+d]^{\frac{1}{2}}}$$

$$\gamma = \frac{a-d}{a+d}$$

3.2. L'ordre partiel entre les coefficients : cas général

Par cas général, rappelons que nous considérons ici l'ensemble $\Theta = \mathcal{P}^{2*}$ des couples de préordres totaux sur X , dont aucun n'est égal au préordre X^2 .

Lemme 1

(1) $\tau_1 \leq \tau_2 \leq \tau_3 \leq \tau_4 \leq \tau_5$

(2) $\tau_F \leq \tau_a \leq \tau_3$

Ces inégalités sont évidentes d'après l'expression des coefficients.

Lemme 2

(3) $\operatorname{Max}(\gamma, d_a, d_b) \leq \tau_5$

En remplaçant $n(n-1)$ par $2(a+b+c^*+d)$, on montre que l'inégalité γ inférieure ou égal à τ_5 est équivalente à l'inégalité triviale $0 \leq 4d(b+c^*)$. On vérifie de même les deux autres inégalités.

Lemme 3

(4) $\tau_F \leq e \leq \tau_3$

On vérifie que la première inégalité est équivalente à l'inégalité triviale $0 \leq 2c^*(b+2a)$ et la seconde à l'inégalité triviale $0 \leq 2c^*(b+d)$.

Lemme 4

(5) $\tau_b \leq \tau_5$

Posons $a+b_2+d = |O| = o$, $a+b_2+d = |O'| = o'$, $2(b_1+c^*) = |\mathring{I}| = i$, $2(b_2+c^*) = |\mathring{I}'| = i'$.

On a :

$$\tau_b = \frac{a-d}{\sqrt{oo'}}$$

$$\tau_5 = \frac{2(a-d)+2c^*+2(b_1+b_2)}{n(n-1)} = \frac{2(a-d)}{n(n-1)} + \frac{i+i'-2c^*}{n(n-1)}$$

$$= \frac{2\sqrt{oo'}}{n(n-1)}\tau_b + \frac{i+i'-2c^*}{n(n-1)}$$

On a donc :

$$\tau_5 \geq \tau_b \Leftrightarrow \frac{2\sqrt{oo'}}{n(n-1)}\tau_b + \frac{i+i'-2c^*}{n(n-1)} \geq \tau_b \Leftrightarrow \frac{i+i'-2c^*}{n(n-1)-2\sqrt{oo'}} \geq \tau_b$$

(Si $n(n-1) = 2\sqrt{oo'}$, $\tau_5 = \tau_b$) .

Cette dernière inégalité est vraie, puisqu'on a toujours $i+i'-2c^* \geq n(n-1) - 2\sqrt{oo'}$.

En effet soit $o \geq o'$; puisque $o' = \frac{n(n-1) - i'}{2}$ on a $n(n-1) - 2\sqrt{oo'} \leq n(n-1) - 2o' = i'$; mais $i + i' - 2c^* \geq i'$ puisque $i - 2c^* = i - |\overset{\circ}{I} \cap \overset{\circ}{I}'| \geq o$.

Donc $\frac{i+i'-2c^*}{n(n-1)-2\sqrt{oo'}} \geq 1 \geq \tau_b$ □

Lemme 5

(6) Si $P \subseteq P'$, $\tau_b \leq \tau_4$.

Si $P = P'$ on a $\tau_b = \tau_4 = +1$. Soit $P \subset P'$, on a $d = b_1 = o$,

$|O| = a+b_2 > |O'| = a$, $\tau_b = (\frac{a}{a+b_2})^{\frac{1}{2}}$, $\tau_4 = \frac{2(a+c^*)+b_2}{n(n-1)} = \frac{n(n-1 - b_2}{n(n-1)}$

L'inégalité $\tau_b \leq \tau_4$ est alors équivalente à l'inégalité

$2(a+b_2)b_2\ n(n-1) \leq b_2\ n^2(n-1)^2 + (a+b_2)b_2^2$. Mais celle-ci est vraie puisque $a + b_2 = |O| \leq \frac{n(n-1)}{2}$ implique $2(a+b_2)\ b_2\ n(n-1) \leq b_2\ n^2(n-1)^2$ (on a d'ailleurs τ_b strictement inférieur à τ_4).

Remarque :

Nous conjecturons qu'on a toujours (7) $\tau_b \leq \tau_4$. Cette inégalité est trivialement vraie dans le cas $a - d$ négatif (car on a alors $\tau_b \leq \tau_a \leq \tau_4$), et nous avons montré qu'elle était vraie dans d'autres cas particuliers. Il reste alors à trouver une preuve générale (ou un contre exemple ...!)

A part les inégalités (1) à (6) établies dans les lemmes précédents et - sans doute - l'inégalité (7), il n'existe pas d'autres inégalités entre les douze coefficients d'accords, dans le cas général (i.e. quand le domaine de définition de ces coefficients est P^{2*}). Ce résultat peut s'exprimer de la manière suivante :

Proposition

L'ordre partiel entre les douze coefficients d'accord définis sur P^{2*} est représenté par le diagramme de la Figure 1.

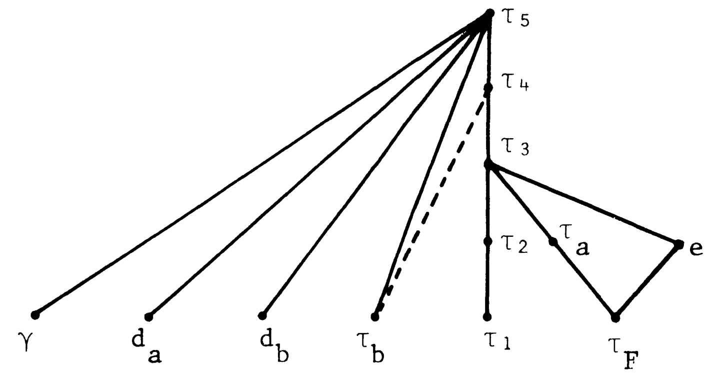

Figure 1.

Sur ce diagramme le trait pointillé entre τ_b et τ_4 signifie que l'inégalité entre ces deux coefficients est seulement conjecturée. Si cette conjecture est vraie, l'ordre partiel figuré ci-dessus correspond à 26 inégalités et 40 incomparabilités entre les coefficients.

Pour montrer la proposition précédente, on considère d'abord les inégalités entre les valeurs absolues de certains de ces coefficients, inégalités dérivant directement de leurs expressions :

$$|\tau_a| \leq |e| \leq \min(|d_a|,|d_b|) \leq |\tau_b| \leq \max(|d_a|,|d_b|) \leq \gamma$$

Deux quelconques des coefficients vérifiant une telle inégalité sont évidemment incomparables; on en déduit 17 incomparabilités. Les autres incomparabilités du diagramme seront montrées ultérieurement pour un domaine restreint de couples de préordres (cf. 3.3) et seront donc à fortiori valables dans le cas général.

Remarque

A part les inégalités évidentes résultant des formules, il ne semble pas y avoir eu d'études générales sur la comparaison des douze coefficients ci-dessus. En particulier, les lemmes 2 à 5 ci-dessus sont (apparemment) nouveaux. Lingoes (1979) "soupçonne" que l'ordre entre quatre coefficients qu'il considère (τ_1 , τ_5 , τ_a et γ) est partiel, mais il écrit faussement qu'on a toujours $\tau_5 \geq \tau_1 \geq \tau_a$ (puisqu'en fait τ_1 et τ_a sont incomparables). Il remarque aussi que τ_5 peut s'écrire $\frac{a-d+b+c^*}{a+d+b+c^*}$, donc que si $b+c^* = |\mathring{I} \cup \mathring{I}'|$ tend vers zéro, τ_5 tend (supérieurement) vers γ .

3.3. L'ordre partiel entre les douze coefficients : cas $a-d \geq 0$.

On se restreint ici à l'ensemble des couples (P,P') de préordres totaux pour lesquels $a(P,P') - d(P,P') \geq 0$ (et P et P' différents de X^2) . Toutes les inégalités du paragraphe précédent restent évidemment valables dans ce cas particulier, mais il s'en ajoute d'autres. Puisque dans ce cas les coefficients τ_a , e , d_a , d_b , τ_b et γ sont toujours positifs, on a

(8) $\tau_F \leq \tau_a \leq e \leq \min(d_a,d_b) \leq \tau_b \leq \max(d_a,d_b) \leq \gamma$.

On va montrer qu'en fait il n'existe pas dans ce cas d'autres inégalités supplémentaires que celles données par (8), ce qui peut s'exprimer de la manière suivante ;

Proposition

L'ordre partiel entre les douze coefficients définis sur l'ensemble des couples de préordres totaux pour lesquels $a-d \geq 0$ (avec a ou $d \neq 0$) est représenté par le diagramme de la Figure 2.

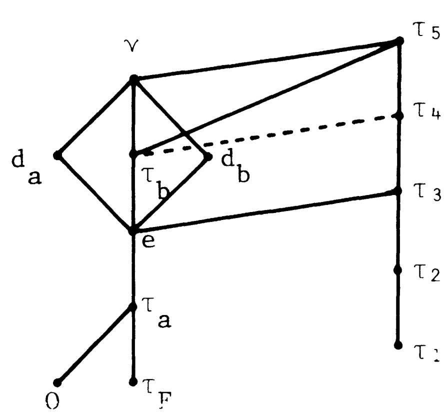

Figure 2.

Là encore dans le diagramme ci-dessus, le pointillé entre τ_b et τ_4, signifie que cette inégalité est seulement conjecturée. Si on la suppose vraie, l'ordre partiel ci-dessus correspond à 42 inégalités et 24 incomparabilités entre les coefficients. Remarquons que sur ce diagramme nous avons ajouté un treizième point $\underline{o}$, qui correspondrait à la valeur d'un coefficient toujours nul. Ceci permet de visualiser sur le diagramme les coefficients σ pour lesquels on a toujours $\sigma(P,P') \geq 0$; c'est en fait le cas de tous à l'exception de τ_F, τ_1 et τ_2.

Pour démontrer la proposition précédente, il faut démontrer les 27 incomparabilités représentées sur le diagramme (24 plus 3 pour le point $\underline{o}$). Une incomparabilité entre deux coefficients σ_1 et σ_2 s'établit en montrant qu'on peut avoir pour certains couples $\sigma_1(P,P') < \sigma_2(P,P')$ et pour d'autres $\sigma_1(P,P') > \sigma_2(P,P')$; ceci peut se faire directement à l'aide d'exemples, ou indirectement en utilisant la transitivité de la relation d'ordre. Il nous suffira en fait d'utiliser quatre exemples que nous donnons ci-dessous.

Exemple 1.	$P = 1 < 2 < \ldots < 7$	$P' = 1 < 2 \ldots 7$
Exemple 2.	$P = 1 < 2 < 3\ 4\ 5\ 6$	$P' = 1 < 2\ 3\ 4\ 5\ 6$
Exemple 3.	$P = 1 < 2\ 3\ 4 < 5$	$P' = 1 < 5 < 2\ 3\ 4$
Exemple 4.	$P = 1\ 2 < 3\ 4 < 5 < 6$	$P' = 1 < 2\ 4 < 3\ 5\ 6$

Démontrons ces incomparabilités en commençant par les trois concernant $\underline{o}$. Il est clair qu'il existe P,P' pour lequel $\tau_F > 0$, $\tau_2 \geq \tau_1 > 0$ (prendre par exemple $P = P'$ ordre total); dans l'exemple 1, on a $\tau_F = \frac{2}{7} > 0 > \tau_2 = -.07 > \tau_1 = -.42$; dans l'exemple 2, on a $\tau_2 = \frac{3}{5} > 0 > \tau_F = -.07$. Donc $\underline{o} \parallel \tau_1$, $\underline{o} \parallel \tau_2$, $\underline{o} \parallel \tau_F$.

Si $P \subset P' \neq X^2$, on a $\gamma = +1 \geq \tau_4 \geq \tau_1$. Dans l'exemple 3, on a $\gamma = .14 < \tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5 = .4$. Donc $\gamma \parallel \tau_1$ et $\gamma \parallel \tau_4$. On en déduit $\gamma \parallel \tau_2$ et $\gamma \parallel \tau_3$, car si, par exemple, $\gamma \leq \tau_2$ on aurait $\gamma \leq \tau_4$, et si $\tau_2 \leq \gamma$, on aurait $\tau_1 \leq \gamma$.

On a $\tau_F \parallel \tau_1$ et $\tau_F \parallel \tau_2$; en effet $\tau_1 < \tau_2 \leq \tau_F$ impliquerait $\tau_2 < \gamma$ ce qui est faux; et dans l'exemple 1, on a $\tau_F = .28 > 0 > \tau_2 > \tau_1$.

Les incomparabilités $\gamma \,||\, \tau_i$ et $\tau_F \,||\, \tau_i$ (pour $i = 1,2$) impliquent immédiatement $\tau_a \,||\, \tau_i$, $e \,||\, \tau_i$, $d_a \,||\, \tau_i$, $d_b \,||\, \tau_i$, $\tau_b \,||\, \tau_i$. On en déduit immédiatement des expressions de τ_b , d_a et d_b que ces trois coefficients sont incomparables.

On a $\tau_3 \,||\, d_b$ et $\tau_4 \,||\, d_b$; en effet $\tau_4 < d_b$ impliquerait $\tau_4 < \gamma$ ce qui est faux; et dans l'exemple 1, on a $\tau_3 = .28 < \tau_4 = .64 < d_b = 1$. On montre de même $\tau_3 \,||\, d_a$ et $\tau_4 \,||\, d_a$.

Enfin on a $\tau_3 \,||\, \tau_b$; en effet $\tau_3 < \tau_b$ est impossible (car impliquant $\tau_3 < \gamma$) , et dans l'exemple 4, on a $\tau_3 = .6 < \tau_b = .75$.

On vérifie qu'on a bien obtenu les 27 incomparabilités annoncées, donc que la proposition est démontrée.

N.B. : Dans les quatre exemples utilisés, on avait $a - d > 0$; l'ordre partiel de la Figure 2 est donc aussi valable si tous les couples d'ordres totaux vérifient cette inégalité stricte. Dans le cas des couples vérifiant l'égalité $a = d$, l'ordre partiel se réduit à celui représenté Figure 3.

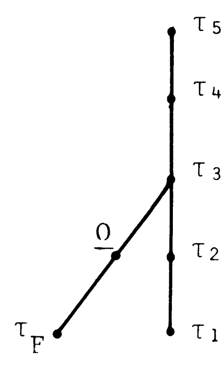

Figure 3

En effet, on a dans ce cas

$$\tau_F = -\frac{2c^*}{n(n-1)} \leq \underline{o} = \tau_a = e = d_a = d_b = \tau_b = \gamma \leq \tau_3 = \frac{2c^*}{n(n-1)}$$

Pour montrer les quatre incomparabilités restantes, il suffit d'exhiber un exemple avec $\tau_1 > \underline{o}$ et un autre avec $\tau_F > \tau_2$; or si $P = 1 < 234 < 567 < 89\times$ et $P' = 234 < 89\times < 567 < 1$, on a $a = d = 18$ et $\tau_1 = .2 > 0$; et pour $P = 12 < 34$ et $P' = 13 < 24$, on a $\tau_F = 0 > \tau_2 = -.17$.

3.4. L'ordre partiel entre les douze coefficients; cas $a - d < 0$

On se restreint ici à l'ensemble des couples (P,P') de préordres totaux pour lesquels $a(P,P') - d(P,P') < 0$ (et P et P' différents de X^2). On a maintenant les inégalités évidentes :

(9) $\gamma \leq \min(d_a,d_b) \leq \tau_b \leq \max(d_a,d_b) \leq e \leq \tau_a \leq 0$.

Là aussi il n'existe pas d'autres inégalités que celles du cas général et celles de (9), ce qui s'exprime par le résultat suivant.

Proposition

L'ordre partiel entre les douze coefficients définis sur l'ensemble des couples de préordre totaux pour lesquels $a - d < 0$ est représenté par le diagramme de la Figure 4.

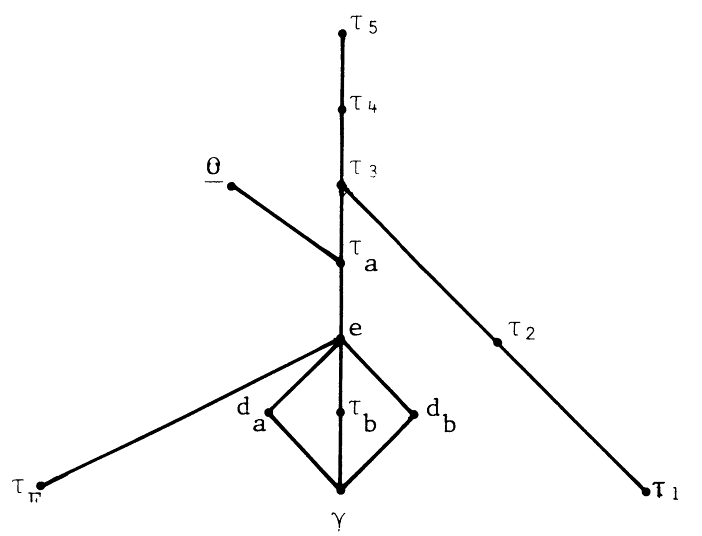

Figure 4

Comme dans les deux diagrammes précédents, nous avons fait figurer dans celui-ci, un point $\underline{o}$ représentant un coefficient identiquement nul. Ce diagramme montre qu'il y a 45 inégalités et 21 incomparabilités entre les douze coefficients dans ce cas.

Pour démontrer la proposition, il faut donc montrer ces 21 incomparabilités plus les cinq concernant le $\underline{o}$. La démonstration est analogue à celle du paragraphe précédent et nous la présentons rapidement; elle utilise les exemples suivants :

Exemple 2' : $P = 1 < 2 < 3456$; $P' = 23456 < 1$

Exemple 3' : $P = 1 < 234 < 5$; $P' = 234 < 5 < 1$

Exemple 5 : $P = 1 < 2 < 3 < 4 < 5$; $P' = 45 < 1 < 2 < 3$

Exemple 6 : $P = 1 < 23 < 45 < 6$; $P' = 5 < 46 < 2 < 1 < 3$

. On a $\underline{\rho} \| \tau_i$, $i = 1$ à 5 , car pour l'exemple 3' , $0 < \tau_1 = \tau_2 = \frac{1}{5} \leq \tau_3 \leq \tau_4 \leq \tau_5$, et pour l'exemple 6 , $\tau_1 < \tau_2 < \tau_3 < \tau_4 < \tau_5 = .33 < \rho$.

. On a $\gamma \| \tau_i$, $i = 1$ et 2 , car pour l'exemple 3', $= -\frac{1}{7} < \tau_1 = \tau_2 = \tau_3 = \frac{1}{5}$, et pour l'exemple 5 , $\tau_1 < \tau_2 = -.35 < \gamma = -.33$.

. On déduit de $\underline{\rho}$ et $\gamma \| \tau_i$, $i = 1$ et 2 , que d_a , τ_b , d_b , e et τ_a sont incomparables à τ_i , $i = 1$ et 2 .

. On a $\tau_F \| \tau_i$, $i = 1$ et 2 , car, $\tau_i \| \underline{\rho}$ implique $\tau_i \nleq \tau_F$, et dans l'exemple 6 , $\tau_1 < \tau_2 = -.63 < \tau_F = -.53$.

. On a $\tau_F \| \tau_b$, γ , d_a, d_b , car, dans l'exemple 3', on a $\tau_F = -.4 < \gamma = \tau_b = d_a = d_b = -.14$, et dans l'exemple 6 , on a $\gamma = -.67 < d_b = -.61 < \tau_b = -.59 < d_a = -.57 < \tau_F = -.53$.

. On a $d_a \| d_b$ et τ_b , $d_b \| \tau_b$.

On vérifie qu'on a bien obtenu ainsi les 26 incomparabilités du diagramme représenté Figure 4.

3.5. L'ordre partiel entre les douze coefficients; cas P comparable avec P' ou P'^r .

3.5.1. Cas $P \subseteq P' \neq X^2$

Les paramètres d'un tel couple vérifient $d = b_1 = 0 < \min(a,b_2)$, (en particulier on est dans le cas $a - d > 0$). On obtient alors les expression suivantes des douze coefficients :

$$\tau_1 = \frac{2(a+c^*-b_2)}{n(n-1)} \qquad \tau_2 = \frac{2(a+c^*)-b_2}{n(n-1)} \qquad \tau_3 = \frac{2(a+c^*)}{n(n-1)}$$

$$\tau_4 = \frac{2(a+c^*)+b_2}{n(n-1)} \qquad \tau_5 = \frac{2(a+c^*+b_2)}{n(n-1)} = \gamma = d_b = +1$$

$$\tau_F = \frac{2(a-c^*)}{n(n-1)} \qquad \tau_a = \frac{2a}{n(n-1)}$$

$$e = d_a = \frac{2a}{n(n-1)-2c^*} = \frac{a}{a+b_2} = \frac{|O'|}{|O|}$$

$$\tau_b = \frac{a}{(a(a+b_2))^{\frac{1}{2}}} = \left(\frac{|O'|}{|O|}\right)^{\frac{1}{2}}$$

On voit donc que certains coefficients deviennent égaux. On va montrer qu'en dehors des modifications apportées par ces égalités l'ordre partiel entre les coefficients est exactement le même que celui obtenu en 3.3 (dans le cas $a - d \geq 0$). Ceci se traduit par la proposition suivante :

Proposition

L'ordre partiel entre les douze coefficients définis sur l'ensemble des couples (P,P') de préordres totaux vérifiant $P \subseteq P' \neq X^2$ est représenté par le diagramme de la Figure 5.

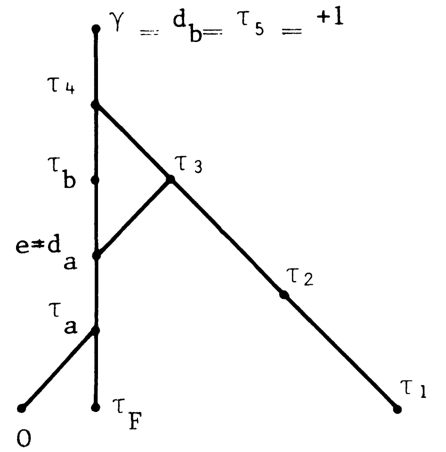

Figure 5

Pour démontrer cette proposition on utilise l'exemple supplémentaire suivant :

Exemple 7 : P = 1 < 2 < 3...14 , P' = 1 < 2...14 .

. On a $\underline{o} \parallel \tau_F$ et τ_2 : on utilise les exemples 1 et 2 (où l'on a $P \subset P'$), pour lesquels on a respectivement $\tau_F = \frac{2}{7} > 0 > \tau_2 = -.07$ et $\tau_F = -.07 < 0 < \tau_2 = \frac{3}{5}$.

. On a $\tau_1 \parallel \tau_b$: si $\tau_b < \tau_1$ on aurait $\underline{o} < \tau_b < \tau_1 < \tau_2$ ce qui est impossible; dans l'exemple 7, on a $\tau_b = .72 < \tau_1 = .736$.

. On a $\tau_3 \parallel \tau_b$: $\tau_3 < \tau_b$ est impossible (puisque $\tau_1 \parallel \tau_b$) ; et dans l'exemple 1, on a $\tau_3 = \frac{2}{7} < \tau_b = (\frac{2}{7})^{\frac{1}{2}}$.

. On a $\underline{o}$, τ_a , e , $\tau_b \parallel \tau_1$ et τ_2 , en utilisant les résultats précédents.

. On a τ_1 , $\tau_2 \| \tau_F$: τ_1 ou $\tau_2 < \tau_F$ est impossible (puisque $\tau_1, \tau_2 \| \tau_a$) ; et dans l'exemple 1, on a $\tau_1 = -.42 < \tau_3 = -.07 < \tau_F = \frac{2}{7}$.

On vérifie que les douze incomparabilités requises ont été établies.

N.B. : Le cas $P \subseteq P' \subset X^2$ (i.e. $d = b_1 = 0 < a$) comporte deux cas particuliers. On peut d'abord avoir aussi $b_2 = 0$, auquel cas on $P = P'$; tous les coefficients prennent alors la valeur $+1$ sauf $\tau_F = \frac{a-c^*}{a+c^*} < \tau_a = \frac{a}{a+c^*}$. On peut aussi avoir $c^* = 0$, auquel cas P est un ordre total, on a alors $\tau_1 < \tau_2 < \tau_F = \tau_a = e = d_a = \tau_3 < \tau_b < \tau_4 < \gamma = d_b = \tau_5 \equiv +1$ avec $\underline{o}$ incomparable à τ_1 et τ_2 , et inférieur à e .

5.3.2. Cas $P \subseteq P'^r \neq X^2$

Les résultats sont similaires au cas précédent : à part les égalités entre coefficients induites par le fait qu'on a dans ce cas $a = b_1 = 0 < d$, on a les mêmes inégalités que dans le cas traité au paragraphe 3.4 (cas $a - d < 0$) . Nous laissons au lecteur le soin de vérifier cette assertion, i.e. la

Proposition

L'ordre partiel entre les douze coefficients définis sur l'ensemble des couples (P,P') de préordres totaux vérifiant $P \subseteq P'^r \neq X^2$ est représenté par le diagramme de la Figure 6.

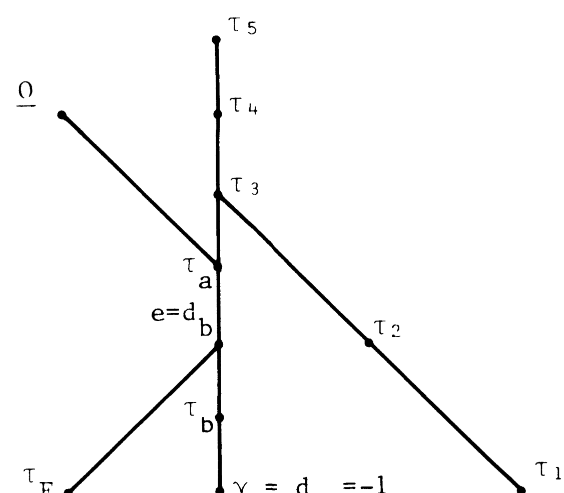

Figure 6.

N.B. : Si de plus on a $P = P'^r$ $(b_2 = 0)$, tous les coefficients prennent la valeur -1 , sauf $\tau_a = -\frac{d}{c^*+d} < \tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5 = \frac{c^*-d}{c^*+d}$. Si d'autre part, on a $a = b_1 = c^* = 0 < d$, P est un ordre total et on obtient : $-1 \equiv \gamma = d_a = \tau_1 < \tau_2 < \tau_b < \tau_3 = \tau_a = \tau_F = e = d_b < \tau_4 < \tau_5$ avec $\underline{o} > e$ et incomparable avec τ_4 et τ_5 .

4. Comparaison ordinale de trois coefficients généralisant le rho de Spearman

Dans ce paragraphe on examine l'ordre entre les coefficients ρ_1 , ρ_a et ρ_b qui généralisent le rho de Spearman dans le cas des ordres totaux. (Pour leurs définitions et étude individuelle cf. Giakoumakis, 1985, Giakoumakis et Monjardet, 1986).

4.1 Expressions de ces coefficients

Ils sont tous les trois basés sur le codage "rang moyen" dont on rappelle brièvement la définition. On fixe une suite $x_1x_2...x_n$ d'éléments de X et on considère un préordre total sur X , $P = C_1 < C_2 < ... < C_t$; on note $n_i = |C_i|$, $i \in [1,t]$. Soit $x_i \in X$ et $x_i \in C_k$.

Par le codage "rang moyen" on fait correspondre à P le vecteur $r = (r_1,...,r_i,...,r_n)$ de $\mathbb{R}^n$ défini de la façon suivante :

$$r_i = \sum_{\ell<k} n_\ell + \frac{n_k+1}{2} .$$

Par exemple, si $P = ab < cde < f$, on aura

$$r = (1{,}5 ,\ 1{,}5 ,\ 4 ,\ 4 ,\ 4 ,\ 6) .$$

Donnons maintenant les formules des trois coefficients .

Posons $P = C_1<...C_i...<C_t$, $n_i = |C_i|$, $i \in [1,t]$

$P' = C'_1<...C'_i...<C'_u$, $n'_i = |C'_i|$, $i \in [1,u]$

$$T = \sum_{i=1}^{t} \frac{(n_i^3 - n_i)}{2} \quad , \quad U = \sum_{i=1}^{u} \frac{(n_i'^3 - n'_i)}{2}$$

On a :

$$\rho_1 = 1 - \frac{6d_E^2(\vec{r},\vec{r}\,')}{n^3-n}$$

$$\rho_a = \frac{n^3-n-6d_E^2(\vec{r},\vec{r}\,')-(T+U)}{n^3-n}$$

$$\rho_b = \frac{n^3-n-6d_E^2(\vec{r},\vec{r}\,')-(T+U)}{(n^3-n-2T)^{1/2}(n^3-n-2U)^{1/2}}$$

où $\vec{r},\vec{r}\,'$ sont les vecteurs "rang moyen" correspondants à P,P' et $d_E(\vec{r},\vec{r}\,')$ la distance euclidienne entre $\vec{r}$ et $\vec{r}\,'$.

4.2. L'ordre partiel entre les trois coefficients généralisant rho .

Il résulte d'abord aisément des formules ci-dessus qu'on a toujours

$$\rho_a \leq \rho_1 \quad |\rho_a| \leq |\rho_b| \text{ et } \rho_a\rho_b \geq 0 .$$

Donc pour ρ_a négatif on a $\rho_b \leq \rho_a \leq \rho_1$.

On dit que deux préordres totaux $P = C_1 < \ldots C_i \ldots < C_t$. $P' = C'_1 < \ldots C_i \ldots < C'_u$ sont de même type si $t = u$, et $|C_i| = |C'_i|$, pour tout $i = 1,\ldots,t$.

Lemme. Soient P et P' deux préordres totaux de même type : alors $\rho_b \leq \rho_1$.

Si P et P' sont de même type, on a $T = U$, et, par conséquent $\rho_b \leq \rho_1$.

$\rho_b = \dfrac{n^3-n-6d_E^2(\vec{r},\vec{r}\,')-2T}{n^3-n-2T}$; puisque $n^3-n-d_E^2(\vec{r},\vec{r}\,') = (n^3-n)\rho_1$, on obtient :

$\rho_b = \dfrac{(n^3-n)\rho_1-2T}{n^3-n-2T}$; alors $\rho_b \leq \rho_1 \Leftrightarrow \dfrac{(n^3-n)\rho_1-2T}{n^3-n-2T} \leq \rho_1 \Leftrightarrow$

$(n^3-n)\rho_1-2T \leq (n^3-n)\rho_1-2T\rho_1$ (puisque $n^3-n-2T > 0$) $\Leftrightarrow 2T \geq 2T\rho_1$, ce qui est vrai puisque $\rho_1 \leq 1$. □

Nous conjecturons qu'on a toujours $\rho_b \leq \rho_1$, sans avoir pu le démontrer. L'ordre partiel entre les trois coefficients est donc représenté par le diagramme suivant de la Figure 7a, où le pointillé entre ρ_b et ρ_1 signifie que l'inégalité $\rho_b \leq \rho_1$ est conjecturée. On a représenté sur la même figure les diagrammes obtenus lorsque ρ_a est positif (Figure 7b) ou négatif (Figure 7c).

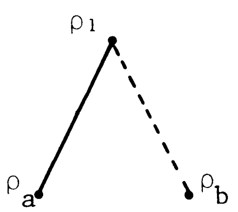

Figure 7a.

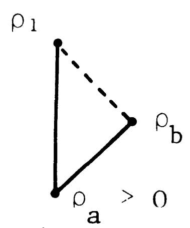

Figure 7b.

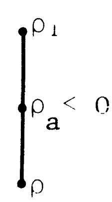

Figure 7c.

5. Comparaison des préordonnances associées à sept coefficients

Soit $\mathcal{P}$ l'ensemble des préordres totaux sur X . Une *préordonnance (totale)* sur $\mathcal{P}$ est un préordre total π défini sur $\mathcal{P}^2$.

Soit σ un coefficient d'accord entre préordres totaux. On lui associe une préordonnance $\pi(\sigma)$ sur $\mathcal{P}$, en posant, pour $p,q \in \mathcal{P}^2$:

$$(p,q) \in \pi(\sigma) \Leftrightarrow \sigma(p) \leq \sigma(q) .$$

Autrement dit $(p,q) \in \pi(\sigma)$ si au sens du coefficient d'accord σ, les deux préordres totaux constituant q sont plus proches que les deux préordres totaux constituant p. Dans la suite, pour alléger les notations, la préordonnance $\pi(\sigma)$ sera notée simplement σ.

Etant donné un couple des préordres totaux (P_i,P_j) et un coefficient σ on note σ^{ij} la valeur de σ qui correspond à ce couple.

Pour montrer que les deux préordonnances associées à deux coefficients σ et σ' sont différentes, il suffit d'exhiber deux couples de préordres totaux (P_1,P_2) et (P_3,P_4) tels que, par exemple :

$$\sigma^{12} \leq \sigma^{34} \quad \text{et} \quad \sigma'^{34} < \sigma'^{12}$$

Proposition

Les préordonnances associées aux sept coefficients τ_1, τ_2, τ_3, τ_4, τ_5, τ_a, τ_F sont toutes différentes.

Preuve

Pour cette preuve on utilise en particulier des couples "extrêmaux" de préordres totaux en ce sens que la valeur du coefficient d'accord est ± 1 (cf. Giakoumakis et Monjardet, 1986).

. Pour $P_1 = P_2$ on a $+1 = \tau_5 = \tau_i$, $i = 1,4$, et pour $P_3 \neq P_4$, $P_3 \cap P_4$ préordre total (par exemple, $P_3 = a < b < cd$, $P_4 = a < bc < d$) on a :

$\tau_i^{34} < \tau_5^{34} = +1$, $i \in [1,4]$ d'où

$\tau_i^{34} < \tau_i^{12} = \tau_5^{12} = \tau_5^{34} = +1 \Rightarrow \tau_5 \neq \tau_i$, $i \in [1,4]$.

. Pour $P_1 = P_2 =$ ordre total on a $\tau_a^{12} = \tau_F^{12} = \tau_5^{12} = +1$ et pour $P_3 \cap P_4$ préordre total non ordre total (par exemple, $P_3 = ab < cd$, $P_4 = ab < cd$) on a : $\tau_F^{34} < \tau_a^{34} < \tau_5^{34} = +1$ d'où

$\tau_F^{34} < \tau_a^{34} < \tau_a^{12} = \tau_F^{12} = \tau_5^{12} = \tau_5^{34} = +1 \quad \tau_5 \neq \tau_a, \tau_F$.

. Pour $P_1 = P_2^r =$ ordre total, on a :

$\tau_1^{12} = \tau_2^{12} = \tau_3^{12} = \tau_4^{12} = \tau_a^{12} = \tau_F^{12} = +1$ et pour $P_3 \cap P_4^r$ ordre total,

$P_3 \neq P_4^r$ (par exemple, $P_3 = ab < c < d$, $P_4 = d < bc < a$) on a

$-1 = \tau_1^{34} < \tau_2^{34}, \tau_3^{34}, \tau_4^{34}, \tau_a^{34}, \tau_F^{34}$ d'où finalement

$\tau_1 \neq \tau_2, \tau_3, \tau_4, \tau_a, \tau_F$.

. Pour $P_1 = ac < bd$, $P_2 = ab < cd$ on a $\tau_2^{12} = -.33$, $\tau_3^{12} = 0$, $\tau_4^{12} = .33$

$\tau_4^{12} = .33$ et pour $P_3 = c < ab < d$, $P_4 = ab < c < d$ on a

$\tau_2^{34} = \tau_3^{34} = \tau_4^{34} = .33$, d'où

$\tau_4 \neq \tau_2, \tau_3$.

. Pour $P_1 = a < b < cd$, $P_2 = abc < d$ on a $\tau_2^{12} = 0$, $\tau_3^{12} = .33$ et pour $P_3 = b < ac < d$, $P_4 = a < bc < d$ on a $\tau_2^{34} = .17$ et $\tau_3^{34} = .33$ d'où

$\tau_2 \neq \tau_3$.

. Pour $P_1 = P_2$ = ordre total on a $\tau_a^{12} = \tau_F^{12} = \tau_2^{12} = \tau_3^{12} = \tau_4^{12} = +1$ et pour $P_3 = P_4 \neq$ ordre total on a : $\tau_a^{34}, \tau_F^{34} < \tau_2^{34} = \tau_3^{34} = \tau_4^{34} = +1$ d'où

$\tau_a, \tau_F \neq \tau_2, \tau_3, \tau_4$.

. Pour $P_1 = P_2^r$ = ordre total on a $\tau_a^{12} = \tau_F^{12} = -1$ et pour $P_3 = P_4^r \neq$ ordre total on a $-1 = \tau_F^{34} < \tau_a^{34}$ d'où $\tau_a \neq \tau_F$. □

BIBLIOGRAPHIE

BARBUT M., MONJARDET B., *Ordre et Classification, Algèbre et Combinatoire*, Paris, Hachette, 1970.

GIAKOUMAKIS V., *Coefficients d'accord entre deux préordres totaux*, Thèse de 3ème cycle, Université de Paris V, 1985.

GIAKOUMAKIS V., MONJARDET B., "Coefficients d'accord entre deux préordres totaux Rapport de recherche, C.M.S., P.027, 1986.

KENDALL M.G., *Rank Correlation Methods*, London, 4th edition, Griffin, 1970.

LINGOES J.C., "Indices of Configural Similarity", in *Geometric Representations of Relational Data*, Lingoes J.C., Roskam E.E., Borg I., edit. Michigan, Mathesis Press, 1979, 675-679.

MONJARDET B., LECONTE DE POLY-BARBUT C., "Valeurs extrémales de la différence de deux coefficients de corrélation de rangs ρ et τ ", Compte-rendu Acad. Sci. Paris, t.303, Série 1, n°10, 1986, (483-486).