MATHÉMATIQUES ET SCIENCES HUMAINES

A. ASTIÉ

Comparaisons par paires et problèmes de classement. Estimation et tests statistiques

Mathématiques et sciences humaines, tome 32 (1970), p. 17-44

<http://www.numdam.org/item?id=MSH_1970__32__17_0>

M.S.H. (8e année, no 32, 1970, p. 17-44)

COMPARAISONS PAR PAIRES ET PROBLÈMES DE CLASSEMENT

ESTIMATION ET TESTS STATISTIQUES

par

A. ASTIÉ *

Cet article traite, du point de vue des tests statistiques, de l'un des modèles classiques de la comparaison par paires. Pour un point de vue très différent, mais complémentaire, on pourra lire J. Marschak, « Binary choice constraints », in Arrow, Karlin et Suppes (eds.), Mathematical methods in social sciences, *Stanford University Press, 1960.*

G. Th. G.

I. INTRODUCTION

Dans la méthode des comparaisons par paires, des objets sont présentés par paire à un ou plusieurs juges.

Quand une « paire » d'objets est présentée à un juge il doit choisir celui des deux qu'il préfère; nous supposerons ici qu'il ne lui est pas permis de les déclarer égaux (mais cette éventualité peut être envisagée, cf. par exemple, Singh et Thompson (1968) et Rao et Kupper (1967)).

A partir d'un ensemble de n objets, on peut comparer certaines ou toutes les $n\,\dfrac{(n-1)}{2}$ paires possibles (ceci sera l'expérience de base dont on fera éventuellement des répétitions). L'ensemble de comparaisons ainsi obtenu nous donne une relation binaire sur l'ensemble des objets: nous l'appellerons relation de comparaison.

Une relation de comparaison est plus générale qu'un classement (car l'objet a peut être préféré à b, b préféré à c et c préféré à a). L'existence de ces incohérences peut être due à différentes raisons, par exemple au fait qu'il existe plusieurs critères de classement.

Le problème sera donc: à partir d'un ensemble de comparaison par paires, peut-on en déduire que la relation de comparaison sous-jacente est une relation d'ordre ? (y a-t-il « assez peu » d'incohérences pour cela ?).

En dehors de son intérêt pratique (il est parfois impossible de comparer autrement les objets), l'intérêt théorique de la méthode des comparaisons par paires par rapport à une méthode de classement

* Laboratoire de Statistique, Faculté des Sciences, Toulouse.

direct des objets par le juge sera donc de faire apparaître éventuellement des incohérences permettant ainsi de contester l'existence d'un classement (ou plus exactement de contester que le juge a fait son choix en accord avec un classement).

Nous ne présenterons pas ici un exposé exhaustif sur les méthodes de comparaisons par paires. Les idées fondamentales proviennent de Kendall et Smith (1940), Slater (1961), Thompson et Remage (1964), Bradley et Terry (1952); on trouvera ici une synthèse de ces divers points de vue présentée de manière à former un ensemble cohérent.

Nous essaierons essentiellement de faire apparaître les rapports qui relient entre elles ces méthodes, ainsi que leur différences et éventuellement leurs faiblesses.

Notons que ce problème peut être rapproché des problèmes d'agrégation des préférences individuelles (effet Condorcet) essentiellement étudiés par Guilbaud, et dont Monjardet (1968) a fait une bibliographie récente. Si ces deux sortes de problèmes diffèrent dans leurs fondements (il ne s'agit pas du même type de données), il est toutefois intéressant de les rapprocher au niveau des techniques de résolution (nous en verrons un exemple avec l'algorithme de E. Jacquet-Lagrèze).

II. GÉNÉRALITÉS

1) *Définitions et notations*

Soit: $X = x_1, \ldots, x_n$ un ensemble de n objets.

Nous définissons une relation de comparaison comme étant une relation binaire R possédant les propriétés:

(i) non-réflexivité: $(x, x) \notin R$;

(ii) anti-symétrie: $(x, y) \in R \Rightarrow (y, x) \notin R$.

Cette relation est complète si elle possède en outre la propriété:

$$\forall x \in X, \forall y \in X, x \neq y, (x, y) \notin R \Rightarrow (y, x) \in R.$$

Une relation de comparaison est une relation d'ordre strict R_0 si elle possède en plus la transitivité:

$$(x, y) \in R_0, (y, z) \in R_0 \Rightarrow (x, z) \in R_0.$$

Si la relation de comparaison est complète, l'ordre strict est total.

Nous représenterons la relation binaire R sur X par le graphe simple $G = (X, R)$ dont R est l'ensemble des arcs.

Si R est une relation de comparaison, $G = (X, R)$ est un graphe simple orienté (anti-symétrie) et sans boucle (non-réflexivité); si de plus il est complet, il sera fréquemment appelé graphe de tournoi.

Nous noterons: $\Omega = \{ = (X, R) \mid R$ est une relation de comparaison complète sur $X \}$, c'est-à-dire Ω est l'ensemble des graphes de tournoi à n sommets.

Le cardinal de Ω est $|\Omega| = 2^{\binom{n}{2}}$.

Nous noterons: $\Omega_0 = \{ G = (X, R_0) \mid R_0$ est une relation d'ordre strict total sur $X \}$.

Nous avons:

$$\Omega_0 \subset \Omega \qquad \text{et} \qquad |\,\Omega_0\,| = n\,!$$

Ω_0 est l'ensemble des graphes sans circuit de Ω.

Nous dirons que $||\,a_{ij}\,||$ $i = 1, \ldots, n$; $j = 1, \ldots, n$ est la matrice associée à $G = (X, R)$ si:

$$a_{ij} = 1 \Leftrightarrow (x_i, x_j) \in R$$
$$a_{ij} = 0 \Leftrightarrow (x_i, x_j) \notin R,$$

on a donc: $a_{ii} = 0$ quel que soit $i = 1, \ldots, n$.

Les définitions et propriétés de théorie des graphes sont empruntées à Berge (1958).

2) *Modèle probabiliste*

Soit: $n\,(n-1)$ variables aléatoires A_{ij} $(i, j = 1, \ldots, n;\ i \neq j)$ a deux valeurs possibles 0 et 1:

$$\left(A_{ij} = \begin{cases} 1 \text{ si } x_i \text{ est préféré à } x_j \text{ dans une comparaison} \\ 0 \text{ si } x_j \text{ est préféré à } x_i \text{ dans une comparaison} \end{cases}\right)$$

Les v.a. A_{ij} $(i < j)$ sont supposées indépendantes et $A_{ij} + A_{ji} = 1$.

On pose:
$$\begin{cases} P_r\,(A_{ij} = 1) = \pi_{ij} \\ P_r\,(A_{ij} = 0) = 1 - \pi_{ij} \\ \pi_{ij} + \pi_{ji} = 1. \end{cases}$$

Soit $G = (X, R)$ le graphe aléatoire à valeurs dans Ω de matrice associée $A = ||\,A_{ij}\,||$.

La loi de probabilité de G (c.a.d. la loi conjointe des v.a. A_{ij}) dépend des paramètres π_{ij}:

$$\text{quel que soit } G \in \Omega,\ P_r\,(G) = \prod_{i<j} \pi_{ij}^{a_{ij}}\, \pi_{ji}^{a_{ji}}$$

en notant $||\,a_{ij}\,||$ la matrice associée à G.

Remarquons que $G^* \in \Omega$ (de matrice associée a_{ij}^*) sera tel que, pour tout $G \in \Omega$, $P_r\,(G^*) \geqslant P_r\,(G)$ si, et seulement si: $a_{ij}^* = 1$ dès que $\pi_{ij} > \pi_{ji}$.

Soit:

$$\Omega^* = \{\, G^* \in \Omega \mid \quad P_r\,(G^*) \geqslant P_r\,(G) \text{ quel que soit } G \in \Omega \,\}$$

$$= \{\, G^* \in \Omega \mid \quad \pi_{ij} \geqslant \frac{1}{2} \Rightarrow a_{ij}^* = 1 \,\},$$

nous noterons aussi $\Omega^*\,(\pi)$,

— si $\pi_{ij} \neq \frac{1}{2}$ pour tout (i, j), G^* est déterminé de manière unique c.a.d. $|\,\Omega^*\,| = 1$;

— s'il existe (i, j) tel que $\pi_{ij} = \pi_{ji} = \frac{1}{2}$, on peut prendre indifféremment $a_{ij}^* = 0$ ou $a_{ij}^* = 1$;

c.a.d., s'il y a k couples non ordonnés (i, j), tels que $\pi_{ij} = \pi_{ji} = \frac{1}{2}$, $| \Omega^* | = 2^k$.

Nous nous intéresserons aux graphes G* comme *représentant la relation de comparaison sous-jacente aux comparaisons observées.*

En particulier si: $\Omega^* \cap \Omega_0 \neq \emptyset$ (c.a.d. $\exists\ G_0^* \in \Omega^* \cap \Omega_0$), la relation de comparaison sous-jacente est une relation d'ordre strict total R_0^* donnée par $G_0^* = (X, R_0^*)$.

Le problème sera donc de déterminer G_0^* et de savoir quelle confiance on pourra accorder au classement ainsi obtenu.

Dans certains cas, on sera amené à introduire des hypothèses supplémentaires, c'est le cas des modèles à valeurs intrinsèques: on suppose que chaque objet x_i a une valeur intrinsèque π_i, π_{ij} étant une fonction de π_i et π_j.

Les modèles usuels sont de la forme:

$$\pi_{ij} = H(\pi_i - \pi_j)$$

où H est une fonction non décroissante variant de 0 à 1 et telle que: $H(-x) + H(x) = 1$ pour tout x,

$$\pi_{ij} \geqslant \frac{1}{2} \text{ est équivalent à } \pi_i \geqslant \pi_j.$$

Donc la relation d'ordre R_0^* définie précédemment, correspond bien au classement des objets par ordre décroissant de leurs valeurs π_i.

Il est clair que $\pi_{ij} = H(\pi_i - \pi_j)$ est une hypothèse plus forte que celle selon laquelle la relation de comparaison est une relation d'ordre.

3) *Tests de signification*

On peut considérer le problème de test comme se décomposant en deux étapes:

Test I

Éprouver l'existence d'une relation d'ordre sur X, c'est-à-dire tester H_0^1 contre H_1^1 avec:

$$H_0^1: \ \Omega^* \cap \Omega_0 \neq \emptyset \Leftrightarrow \forall\ (i, j, k);\ \pi_{ij} > \frac{1}{2} \text{ et } \pi_{jk} > \frac{1}{2} \Rightarrow \pi_{ki} \leqq \frac{1}{2}$$

$$H_1^1: \ \Omega^* \cap \Omega_0 = \emptyset \Leftrightarrow \exists\ (i, j, k),\ \pi_{ij} > \frac{1}{2},\quad \pi_{jk} > \frac{1}{2},\quad \pi_{ki} > \frac{1}{2}$$

Test II

Si ce premier test n'est pas significatif, on accepte $\Omega^* \cap \Omega_0 \neq \emptyset$, c.a.d. l'existence d'une relation d'ordre sur X. Il reste alors à éprouver l'hypothèse que les comparaisons ont été effectuées « au hasard »,

c.a.d. tester H_0^2 contre H_1^2 (nous sommes dans le cas: $\Omega^* \cap \Omega_0 \neq \emptyset$):

$$H_0^2: \ \Omega^* \cap \Omega_0 = \Omega_0 \Leftrightarrow \forall\ (i, j),\ \pi_{ij} = \frac{1}{2}$$

$$H_1^2: \ \Omega^* \cap \Omega_0 \neq \Omega_0 \Leftrightarrow \exists\ (i, j),\ \pi_{ij} \neq \frac{1}{2}.$$

Montrons que les deux écritures de H_0^2 sont équivalentes:

$$\Omega^* \cap \Omega_0 = \Omega_0 \Leftrightarrow \forall\ G_0 \in \Omega_0, \quad \forall\ G_0' \in \Omega_0: P_r(G_0') = P_r(G_0).$$

Prenons pour G_0' l'ordre inverse de G_0 (c.a.d. $R_0 \cap R_0' = \emptyset$); $G_0 \in \Omega^*$ donc, :

$$\text{si } (x_i, x_j) \in R_0, \quad \pi_{ij} \geqq \frac{1}{2}, \ (x_i, x_j) \in R_0 \Rightarrow (x_j, x_i) \in R_0', \quad \text{et} \quad G_0' \in \Omega^* \Rightarrow \pi_{ji} \geqq \frac{1}{2}.$$

donc:

$$\pi_{ij} = \pi_{ji} = \frac{1}{2}$$

et ceci quel que soit i et j.

La réciproque est évidente.

Mais en pratique, on ne fera que le test II car l'hypothèse H_0^1 est une hypothèse composite peu maniable, alors que H_0^2 est une hypothèse simple.

Cela signifie qu'on admet *a priori*, qu'il n'y a que deux éventualités: les comparaisons sont effectuées « au hasard », ou en accord avec un classement.

A titre d'exemple, précisons la structure des hypothèses pour $n = 3$:

l'espace paramétrique est le cube de côté 1: $\Pi\ \pi = (\pi_{12}, \pi_{13}, \pi_{23})$.

Les plans $\pi_{12} = \frac{1}{2}, \pi_{13} = \frac{1}{2}, \pi_{23} = \frac{1}{2}$ divisent le cube C en 8 cubes de côté $\frac{1}{2}$.

Notons:

$$K = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$$

$$C_1 = \left\{ \pi \in \Pi: \pi_{12} \leqq \frac{1}{2}, \pi_{13} \leqq \frac{1}{2}, \pi_{23} \leqq \frac{1}{2} \right\}$$

$$C_2 = \left\{ \pi \in \Pi: \pi_{12} \geqq \frac{1}{2}, \pi_{13} \leqq \frac{1}{2}, \pi_{23} \leqq \frac{1}{2} \right\}$$

$$C_3 = \left\{ \pi \in \Pi : \pi_{12} \leqq \frac{1}{2}, \pi_{13} \geqq \frac{1}{2}, \pi_{23} \leqq \frac{1}{2} \right\}$$

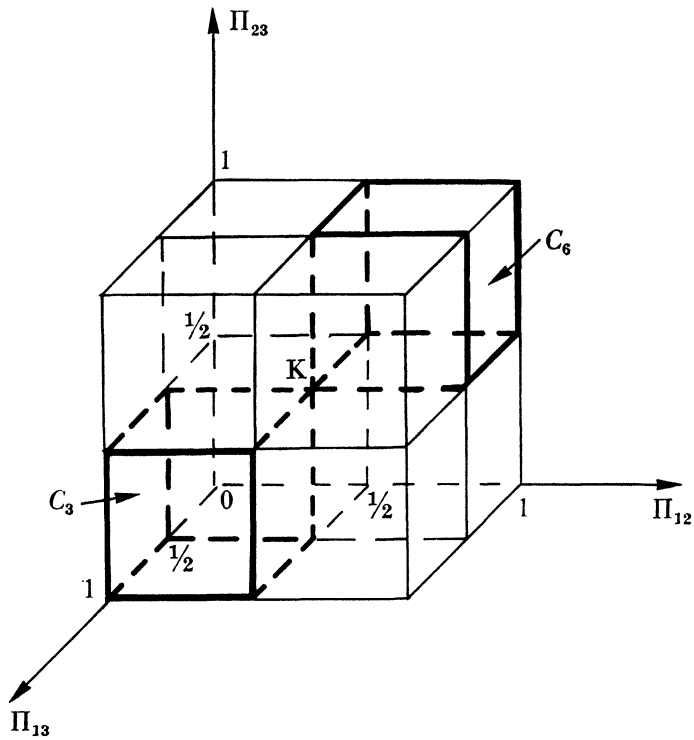
$$C_4 = \left\{ \pi \in \Pi : \pi_{12} \leqq \frac{1}{2}, \pi_{13} \leqq \frac{1}{2}, \pi_{23} \geqq \frac{1}{2} \right\}$$

$$C_5 = \left\{ \pi \in \Pi : \pi_{12} \geqq \frac{1}{2}, \pi_{13} \geqq \frac{1}{2}, \pi_{23} \leqq \frac{1}{2} \right\}$$

$$C_6 = \left\{ \pi \in \Pi : \pi_{12} \geqq \frac{1}{2}, \pi_{13} \leqq \frac{1}{2}, \pi_{23} \geqq \frac{1}{2} \right\}$$

$$C_7 = \left\{ \pi \in \Pi : \pi_{12} \leqq \frac{1}{2}, \pi_{13} \geqq \frac{1}{2}, \pi_{23} \geqq \frac{1}{2} \right\}$$

$$C_8 = \left\{ \pi \in \Pi : \pi_{12} \geqq \frac{1}{2}, \pi_{13} \geqq \frac{1}{2}, \pi_{23} \geqq \frac{1}{2} \right\}$$



Nous noterons:

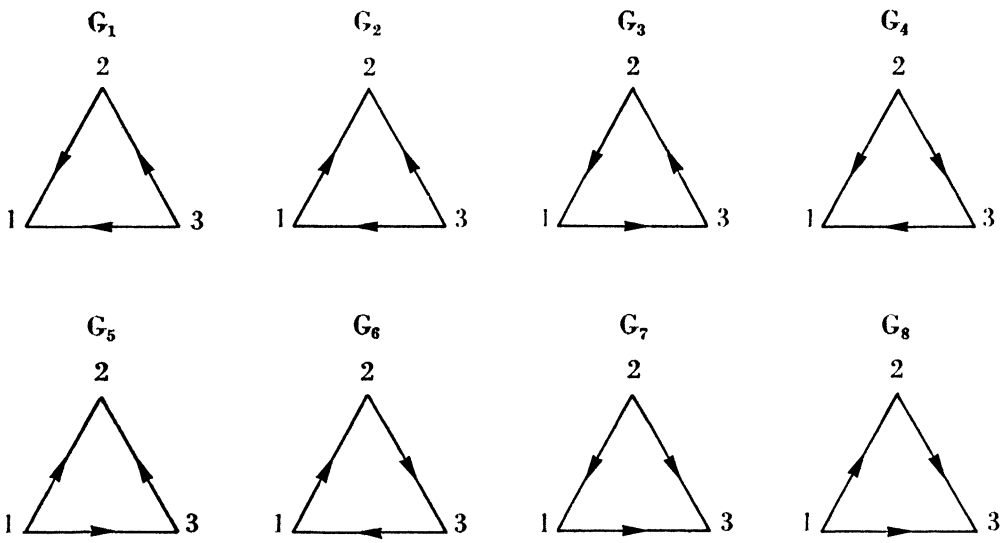
$$\overset{0}{C}_1 = \left\{ \pi \in \Pi : \pi_{12} < \frac{1}{2}, \pi_{13} < \frac{1}{2}, \pi_{23} < \frac{1}{2} \right\}$$

Il y a correspondance biunivoque entre C_k et $G \in \Omega$ définie par:

$$C_k \Leftrightarrow G_k$$

où:

$$\Omega^* (\pi) = \{G_k\} \quad \forall \pi \in \overset{0}{C}_k .$$



$G_3 \notin \Omega_0$, $G_6 \notin \Omega_0$

H_0^1: $\pi \in C_1 \cup C_2 \cup C_4 \cup C_5 \cup C_7 \cup C_8$

H_1^1: $\pi \in \overset{0}{C_3} \cup \overset{0}{C_6}$

si H_1^0 acceptée:

H_0^2: $\pi = K$

H_1^2: $\pi \in C_1 \cup C_2 \cup C_4 \cup C_5 \cup C_7 \cup C_8 - \{K\}$.

III. COMPARAISONS COMPLÈTES (sans répétition)

Le résultat d'une expérience est un graphe observé G, appartenant à Ω, de matrice associée:

$$a = || a_{ij} ||.$$

On se propose de tester (test II de l'introduction) H_0^2 contre H_1^2.

Nous pouvons interpréter ces hypothèses de la manière suivante :

H_0^2: le graphe aléatoire G est uniformément réparti dans Ω :

$$\forall G \in \Omega, \quad P_r(G) = \frac{1}{\binom{n}{2}}$$

H_1^2: G n'est pas uniformément réparti dans Ω et une de ses valeurs modales appartient à Ω_0:

$$\exists G_0^* \in \Omega_0 : \begin{cases} \forall G \in \Omega, P_r(G_0^*) \geqq P_r(G) \quad \text{et} \\ \exists G_0 \in \Omega_0, P_r(G_0^*) > P_r(G_0). \end{cases}$$

On doit choisir une région critique R_c, contenue dans Ω. Si le graphe observé G appartient à R_c, on rejettera H_0^2, c'est-à-dire qu'on acceptera l'existence d'une relation d'ordre sur X.

Le problème est de choisir R_c.

Il semble naturel de prendre Ω_0 contenu dans R_c: en effet, il n'y a pas de raison de favoriser *a priori* certains classements.

Ensuite on complétera R_c par les éléments de Ω « les plus proches » de Ω_0 (en un sens qui sera précisé ultérieurement) de manière à avoir $P_r(R_c \mid H_0^2) = \alpha$ (α étant le seuil de signification choisi pour le test).

On voit apparaître ici une difficulté du problème.

Alors que, pour le choix de la région critique, tous les éléments de Ω_0 jouent le même rôle, lorsque nous accepterons H_1^2 nous conclurons à l'existence d'une relation d'ordre sur X: R_0^*, non nécessairement unique, correspondant à $G_0^* \in \Omega^* \cap \Omega_0$.

R_0^* étant une relation d'ordre sur X, notons $(x_{i_1}, \ldots, x_{i_n})$ le classement qui lui correspond.

Puisque G_0^* appartient à Ω^*, $\pi_{i_k\, i_h} \geqq \frac{1}{2}$ pour tout $k < h$ et $P_r(G_0^*) = \sup_{G \in \Omega} P_r(G)$, soit $\overline{G_0} = (X, \overline{R_0})$, le graphe de Ω_0 tel que $R_0^* \cap \overline{R_0} = \emptyset$, le classement qui lui correspond est: $(x_{i_n}, \ldots, x_{i_1})$.

$$P_r(\overline{G_0}) = \prod_{h < k} \pi_{i_h\, i_k}$$

donc:

$$P_z(\overline{G_0}) = \inf_{G \in \Omega} P_z(G),$$

or, G_0^* et $\overline{G_0}$ jouent le même rôle au niveau du choix de R_c.

La puissance du test étant:

$$P_r(R_c \mid H_1^2) = \sum_{G \in R_c} P_r(G \mid H_1^2),$$

et le seuil étant:

$$\alpha = P_r(R_c \mid H_0^2) = \sum_{G \in R_c} P_r(G \mid H_0^2),$$

nous voyons que si:

$$P_r(G_0^* \mid H_1^2) > P_r(G_0^* \mid H_0^2),$$

par contre:

$$P_r(\overline{G_0} \mid H_1^2) < P_r(\overline{G_0} \mid H_0^2)$$

et pour certaines valeurs de π, il sera possible d'avoir:

$$P_r(R_c \mid H_1^2) < P_r(R_c \mid H_0^2),$$

c'est-à-dire que le test sera *biaisé.*

On obtient des régions critiques différentes suivant la façon dont on définira la « distance » d'un graphe G à Ω_0.

1) *Test de Kendall et Smith*

On définira l'écart d'incohérence de $G = (X, R)$ et $G' = (X, R')$ appartenant à Ω par:

$$e(G, G') = \sum_{i} \sum_{j} \sum_{k} \left| t_{ijk} - t'_{ijk} \right|$$

où:

$$t_{ijk} = \begin{cases} 1 & \text{si } (x_i, x_j) \in \mathrm{R} \text{ et } (x_j, x_k) \in \mathrm{R} \text{ et } (x_k, x_i) \in \mathrm{R}, \text{ ou} \\ & (x_k, x_j) \in \mathrm{R} \text{ et } (x_j, x_i) \in \mathrm{R} \text{ et } (x_i, x_k) \in \mathrm{R} \\ 0 & \text{dans le cas contraire.} \end{cases}$$

t'_{ijk} est défini de manière analogue pour R′.

L'écart de G à une partie ω de Ω sera:

$$e(\mathrm{G}, \omega) = \min_{\mathrm{G}' \in \omega} e(\mathrm{G}, \mathrm{G}'),$$

en particulier:

$$e(\mathrm{G}, \Omega_0) = \sum_{i} \sum_{j} \sum_{k} t_{ijk}$$

est le nombre de circuits de longueurs 3 de G (en effet, si $\mathrm{G}' \in \Omega_0$, $t'_{ijk} = 0$ pour tout (i, j, k)). Remarquons que pour cet écart, un graphe G appartenant à Ω est « équidistant » de tout G_0 appartenant à Ω_0.

On vérifie que: $\Omega_0 = \{ \mathrm{G} \in \Omega : e(\mathrm{G}, \Omega_0) = 0 \}$.

Posons:

$$\Omega_k = \{ \mathrm{G} \in \Omega : e(\mathrm{G}, \Omega_0) = k \} \qquad k = 0, 1, 2, \ldots$$

Nous prendrons donc comme région critique Ω_0 et les graphes les plus proches de Ω_0 (au sens de l'écart ci-dessus),

$$\mathrm{R}_k = \bigcup_{k=0}^{u} \Omega_k$$

u étant le plus grand entier positif ou nul vérifiant:

$$\mathrm{P_r}\left(\bigcup_{k=0}^{u} \Omega_k \mid H_0^2 \right) \leqq \alpha.$$

Si E_n est la variable aléatoire égale au nombre de circuits de longueur 3 du graphe aléatoire G sur n sommets, ceci s'écrit:

$$\mathrm{P_r}(\mathrm{R}_k \mid H_0^2) = \mathrm{P_r}(\mathrm{E}_n \leqq u \mid H_0^2) = \sum_{k=0}^{u} \frac{|\Omega_k|}{|\Omega|}.$$

Des tables donnant la distribution de E_n sous H_0^2 ont été établies par Kendall et Smith (1940) pour $3 \leqq n \leqq 7$ et par Alway (1962) pour $8 \leqq n \leqq 10$.

On montre de plus que sous H_0^2:

$$\mu = \mathrm{E}(\mathrm{E}_n) = \frac{1}{4}\binom{n}{3} \qquad \sigma^2 = \mathrm{Var}(\mathrm{E}_n) = \frac{3}{16}\binom{n}{3}$$

et $\frac{E_n - \mu}{\sigma}$ converge en loi vers une variable normale réduite lorsque n tend vers l'infini.

Soit G le graphe observé, de matrice associée $||\, a_{ij} \,||$.

Posons $s_i = \sum_{j(\neq i)} a_{ij}$, s_i sera le score observé de x_i.

$$e\,(G, \Omega_0) = \binom{n}{3} - \sum_{i=1}^{n} \binom{s_i}{2} = \frac{n\,(n-1)\,(2\,n-1)}{12} - \frac{1}{2} \sum_{i=1}^{n} s_i^2$$

$$e\,(G, \Omega_0) \leqslant \begin{cases} \dfrac{n^3 - n}{24} & \text{si } n \text{ est impair} \\ \dfrac{n^3 - 4\,n}{2} & \text{si } n \text{ est pair} \end{cases}$$

$e\,(G, \Omega_0)$ est donc aisément calculable, puisque c'est une fonction des s_i, donc si, pour le graphe observé G, on a $e\,(G, \Omega_0) \leqq u$, on rejettera H_0^2 et on admettra qu'il existe une relation d'ordre sur X.

2) *Test de Slater*

On définira la distance [1] entre deux graphes G et G' $\in \Omega$ par:

$$d\,(G, G') = \sum_i \sum_j |\, a_{ij} - a'_{ij} \,|$$

où

$||\, a_{ij} \,||$ est la matrice associée à G et
$||\, a'_{ij} \,||$ est la matrice associée à G'.

et la distance d'un graphe G $\in \Omega$ à une partie $\omega \subset \Omega$.

$$d\,(G, \omega) = \min_{G' \in \omega} d\,(G, G')$$

en particulier nous aurons :

$$d\,(G, \Omega_0) = \min_{G_0 \in \Omega_0} d\,(G, G_0)$$

On a:

$$\Omega_0 = \{\, G \in \Omega : d\,(G, \Omega_0) = 0 \,\}.$$

Posons:

$$\omega_k = \{\, G \in \Omega : d\,(G, \Omega_0) = k \,\} \qquad k = 0, 1, \ldots \quad (\omega_0 = \Omega_0)$$

1. Cette distance a été envisagée par M. Barbut.

Nous prendrons la région critique:

$$R_S = \bigcup_{k=0}^{v} \omega_k$$

où v est le plus grand entier positif ou nul vérifiant:

$$P_r\left[\bigcup_{k=0}^{v} \omega_k \mid H_0^2\right] \leqq \alpha.$$

Au graphe aléatoire (à n sommets) G, nous associerons la variable aléatoire $D_n = d(G, \Omega_0)$ et nous écrirons:

$$P_r\left[\bigcup_{k=0}^{v} \omega_k \mid H_0^2\right] = P_r\left[D_n \leqq v \mid H_0^2\right] = \frac{\sum_{k=0}^{v} |\omega_k|}{|\Omega|}$$

Avec les résultats de Reid (1969), on voit que, pour tout $n \geqq 2$ on a:

$$D_n \leqslant \begin{cases} \dfrac{n(n-4)}{4} + 1 & \text{si } n \text{ est pair} \\ \dfrac{(n-1)(n-3)}{4} + 1 & \text{si } n \text{ est impair} \end{cases}$$

Si $n \geqslant 8$, on a même:

$$D_n \leqq \begin{cases} \dfrac{n(n-4)}{4} & \text{si } n \text{ est pair} \\ \dfrac{(n-1)(n-3)}{4} & \text{si } n \text{ est impair} \end{cases}$$

Slater donne des tables de la distribution de D_n pour $2 \leqq n \leqq 8$ et $0 \leqq d_n \leqq 8$.

Remage et Thompson (1966) donnent un algorithme de programmation dynamique qui permet de déterminer l'ensemble des graphes $G_0 \in \Omega_0$ tels que $d(G, G_0) = d(G, \Omega_0)$, on en déduit immédiatement la valeur observée de D_n, c'est-à-dire $d(G, \Omega_0)$.

Si $d(G, \Omega_0) \leqslant v$, on rejettera H_0^2 et on admettra l'existence d'une relation d'ordre sur X, donnée par:

$$G_0 \in \Omega_0 \quad \text{tel que} \quad d(G, G_0) = d(G, \Omega_0),$$

elle n'est pas unique puisqu'il peut y avoir plusieurs graphes G_0 qui vérifient cela.

3) *Conclusion sur les méthodes de Kendall et de Slater*

— Lorsque nous rejettons H_0^2, nous admettons l'existence d'une relation d'ordre sur X.

— Dans le cas du test de Slater, nous obtenons directement la relation d'ordre (G_0 tel que: $d(G, G_0) = d(G, \Omega_0)$, qui n'est pas nécessairement unique.

— Par contre, pour le test de Kendall, le test ne particularise aucun élément de Ω_0 ($e(G, G_0) = e(G, \Omega_0)$, $\forall\ G_0 \in \Omega_0$).

Il semble cependant logique d'adopter un classement par ordre décroissant des scores S_i :

$$\text{soit} \quad S_i = \sum_{j(\neq i)} A_{ij},$$

sous H_0^2 on a $E(S_i) = \dfrac{n-1}{2}$.

Posons:

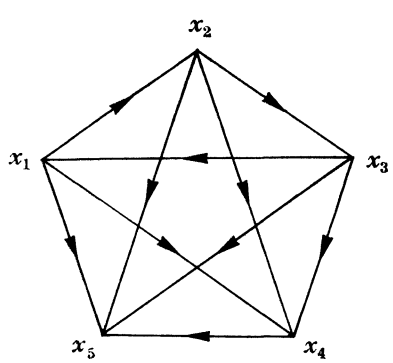
$$T = \sum_{i=1}^{n} \left(S_i - \frac{n-1}{2} \right)^2$$

T est d'autant plus grand que les scores S_i sont plus éloignés de cette valeur moyenne $\dfrac{n-1}{2}$.

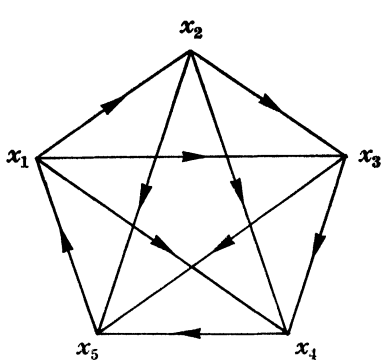
Or il est facile de montrer que $T = -2\,E_n + \dfrac{n^3 - n}{12}$.

La région critique du test de Kendall correspond aux petites valeurs de E_n, donc aux grandes valeurs de T. Cela montre que le test de Kendall conduit à rejeter H_0^2 si les scores sont suffisamment dispersés autour de la moyenne $\dfrac{n-1}{2}$.

Exemples



G_1



G_2

Pour G_1: $t_{123} = 1$ et $t_{ijk} = 0$ pour tout $(i, j, k) \neq (1, 2, 3)$, donc:

$e(G_1, \Omega_0) = 1$
$d(G_1, \Omega_0) = 1$

(il suffit de renverser un quelconque des 3 arcs (x_1, x_2); (x_2, x_3); (x_3, x_1) pour obtenir un graphe sans circuit).

Pour G_2: $t_{125} = 1 \quad t_{135} = 1 \quad t_{145} = 1$

$$e\,(G_2, \Omega_0) = 3$$
$$d\,(G_2, \Omega_0) = 1$$

(il suffit de renverser l'arc (x_5, x_1) pour avoir un graphe sans circuit).

On a $d\,(G_1, \Omega_0) = d\,(G_2, \Omega_0)$, donc G_1 et G_2 seront traités de manière analogue par le test de Slater.

Par contre, $e\,(G_2, \Omega_0) > e\,(G_1, \Omega_0)$ donc, suivant le seuil choisi, nous pouvons être amené à rejeter H_0^2 lorsque G_1 est observé alors qu'on accepte H_0^2 avec G_2.

Remarquons que dans les deux cas, le classement $(x_1, x_2, x_3, x_4, x_5)$ est un classement par ordre décroissant des scores qui est en désaccord avec:

— l'arc (x_3, x_1) de G_1 tel que $s_1^{(1)} = s_3^{(1)} = 3$,

— l'arc (x_5, x_1) de G_2 tel que $s_1^{(2)} = 3$, $s_5^{(2)} = 1$.

En notant $s_i^{(j)}$, $i = 1, \ldots, n$; $j = 1, 2$; le score de x_i correspondant à G_j.

Alors que la méthode de Slater ne considère que le nombre d'arcs à intervertir pour arriver à un graphe transitif, la méthode de Kendall pondère les arcs à intervertir d'autant plus lourdement que les extrémités de ces arcs ont des scores plus différents.

Remarquons que, d'une manière générale, si le graphe observé G est tel que: $(x_i, x_j) \in R$ et $s_j > s_i$, le fait de remplacer l'arc (x_i, x_j) par (x_j, x_i) diminue de $s_j - s_i + 1$ le nombre de circuits de longueur 3 du graphe.

4) *Une généralisation de la méthode de Slater au cas de comparaisons complètes avec répétitions*

Soient $G_1, G_2, \ldots, G_m$ les graphes fournis par m juges, ce sont les valeurs observées de m graphes aléatoires indépendants et distribués comme G.

Soit $\| a_{ijk} \|$ la matrice associée à G_k $(k = 1, 2, \ldots, m)$.

Dans le cas d'une observation $(m = 1)$, la méthode de Slater conduisait à déterminer $G_0 \in \Omega_0$ tel que $d\,(G, G_0) = \min_{G_0' \in \Omega_0} d\,(G, G_0')$.

Ici il faudra déterminer $G_0 \in \Omega_0$ qui soit le plus proche simultanément de $G_1, G_2, \ldots, G_m$.

Par exemple, nous pouvons prendre le graphe $G_0 \in \Omega_0$ qui minimise:

$$\sum_{k=1}^{m} d\,(G_k, G_0) = \sum_{k=1}^{m} \sum_{i<j} | a_{ijk} - a_{ijo} |$$

en notant $\| a_{ijo} \|$ la matrice associée à $G_0 = (X, R_0)$.

$$\sum_{k=1}^{m} \sum_{i<j} | a_{ijk} - a_{ijo} | = \sum_{i<j} \sum_{k=1}^{m} | a_{ijk} - a_{ijo} | = \sum_{\substack{(i,j) \\ (x_i, x_j) \in R_0}} s_{ji}$$

donc :

$$\min_{G_0 \in \Omega_0} \sum_{k=1}^{m} d(G_k, G_0) = \min_{G_0 \in \Omega_0} \sum_{\substack{(i,j) \\ (x_i, x_j) \in R_0}} s_{ji}$$

donc $G_0 = (X, R_0)$ défini de cette manière minimise le nombre de préférences observées qui sont en désaccord avec l'ordre R_0.

Remarque.

Pour la détermination pratique de R_0, on pourra utiliser l'algorithme proposé par E. Jacquet-Lagrèze, algorithme permettant d'échanger les objets i et j dans une matrice carrée en suivant le critère $\sum_{i<j} s_{ij}$ maximum (puisque $s_{ij} + s_{ji} = m$, $\sum_{i<j} s_{ij}$ maximum est équivalent à $\sum_{i<j} s_{ij}$ minimum).

IV. CAS GÉNÉRAL

m_{ij} répétitions de la comparaison de la paire (x_i, x_j) $(m_{ij} > 0)$.

Pour chaque couple (i, j) on a un échantillon de m_{ij} v.a. indépendantes et distribuées comme A_{ij}.

On les notera: A_{ijk}, $k = 1, 2, \ldots, m_{ij}$.

La vraisemblance de l'échantillon global sera:

$$L(a, \pi) = \prod_{i<j} \prod_{k=1}^{m_{ij}} \pi_{ij}^{a_{ijk}} \pi_{ji}^{a_{jik}}$$

posons:

$$S_{ij} = \sum_{k=1}^{m_{ij}} A_{ijk} \qquad \text{et} \qquad s_{ij} = \sum_{k=1}^{m_{ij}} a_{ijk}\,.$$

Remarquons que:

$$A_{ijk} + A_{jik} = 1 \qquad S_{ij} + S_{ji} = m_{ij}$$

$$L(a, \pi) = \prod_{i<j} \pi_{ij}^{s_{ij}} \pi_{ji}^{s_{ji}}\,.$$

S_{ij} constitue une statistique exhaustive pour π_{ij}.

1) *Estimation d'un classement*

Si nous estimons les paramètres π_{ij} par la méthode du maximum de vraisemblance, sans aucune hypothèse sur leurs valeurs, nous trouvons:

$$\hat{\pi}_{ij} = \frac{s_{ij}}{m_{ij}}$$

Ceci nous définit $\Omega^* (\hat{\pi}) = \{ G \in \Omega \mid \hat{\pi}_{ij} > \frac{1}{2} \Rightarrow a_{ij} = 1 \}$ et tout graphe de $\Omega^* (\hat{\pi})$ est une estimation du maximum de vraisemblance de la relation de comparaison représentée par $G^* \in \Omega^*$.

Mais il est évident que les graphes de $\Omega^* (\hat{\pi})$ ainsi obtenus n'appartiennent pas nécessairement à Ω_0 (ils peuvent comporter des circuits).

Nous allons donc admettre, *a priori*, l'hypothèse que nous avons noté H_0^1: $\Omega^* \cap \Omega_0 \neq \emptyset$ (en fait on a déjà vu qu'on admet toujours cette hypothèse *a priori* car elle n'est pas facile à tester).

Et sous l'hypothèse H_0^1 nous allons estimer les paramètres π_{ij} par la méthode du maximum de vraisemblance, nous noterons $\tilde{\pi}_{ij}$ leurs estimations.

Parmi les graphes de $\Omega^* (\tilde{\pi})$ il y en aura au moins un appartenant à Ω_0 (puisque $\Omega^* \cap \Omega_0 \neq \emptyset$)

Notons Π l'espace paramétrique de dimension $\binom{n}{2}$:

$$\text{si } \pi \in \Pi : \pi = (\pi_{12}, \pi_{13}, \ldots, \pi_{ij}, \ldots, \pi_{n-1\, n}).$$

L'estimation $\tilde{\pi}$ de π est définie par:

$$L (a, \tilde{\pi}) = \sup_{\pi \in \Pi_0} L (a, \pi)$$

où:

$$\Pi_0 = \{ \pi \in \Pi : \Omega^* (\pi) \cap \Omega_0 \neq \emptyset \}.$$

$\Omega^* (\tilde{\pi}) \cap \Omega_0$ est alors l'ensemble des classements estimés par la méthode du maximum de vraisemblance.

$$\text{Log } L (a, \pi) = \sum_{i<j} (s_{ij} \text{ Log } \pi_{ij} + s_{ji} \text{ Log } \pi_{ji}) = \sum_{i<j} u_{ij} (\pi_{ij})$$

où:

$$u_{ij} (\pi_{ij}) = s_{ij} \text{ Log } \pi_{ij} + s_{ji} \text{ Log } \pi_{ji}$$

est maximum pour $\pi_{ij} = \hat{\pi}_{ij}$.

Montrons que:

$$\tilde{\pi}_{ij} = \begin{cases} \hat{\pi}_{ij} \\ \text{ou} \\ \frac{1}{2} \end{cases}$$

— si $\hat{\pi} \in \Pi_0$, il est évident que $\tilde{\pi}_{ij} = \hat{\pi}_{ij}$ pour tout (i, j);

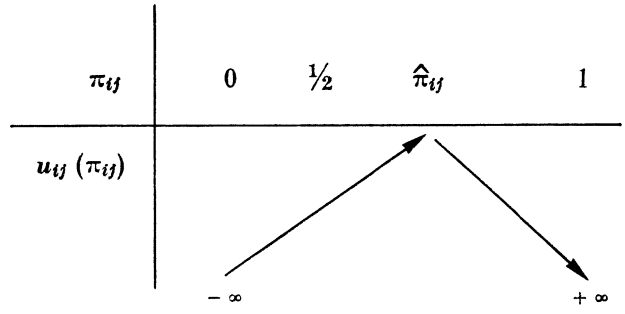
— si $\hat{\pi} \notin \Pi_0$,

$$\Omega^*(\hat{\pi}) \cap \Omega_0 = \emptyset \Rightarrow \forall\, \hat{G} = (X, \hat{R}) \in \Omega^*(\hat{\pi}), \qquad \hat{G} \notin \Omega_0$$
$$\Omega^*(\tilde{\pi}) \cap \Omega_0 \neq \emptyset \Rightarrow \exists\, \tilde{G} = (X, \tilde{R}) \in \Omega^*(\tilde{\pi}), \qquad \tilde{G} \in \Omega_0$$

puisque $\hat{G} \notin \Omega_0$ et $\tilde{G} \in \Omega_0$, il existe (i, j) tel que $(x_i, x_j) \in \hat{R}$ et $(x_i, x_j) \notin \tilde{R}$, c'est-à-dire il existe (i, j):

$$\hat{\pi}_{ij} \geqq \frac{1}{2} \qquad \text{et} \qquad \tilde{\pi}_{ij} \leqq \frac{1}{2}.$$

Nous pouvons représenter sur un tableau, les variations de la fonction $u_{ij}(\pi_{ij})$.



Il est clair que si $\hat{\pi}_{ij} \geqq \frac{1}{2}$, on a:

$$u_{ij}(\pi_{ij}) \leqq u_{ij}\left(\frac{1}{2}\right) \qquad \forall\, \pi_{ij} \leqq \frac{1}{2}$$

donc:

$$\tilde{\pi}_{ij} = \frac{1}{2}.$$

$\tilde{\pi}$ est défini par:

$$L(a, \tilde{\pi}) = \sup_{\pi \in \Pi_0} L(a, \pi)$$

ceci équivaut à:

$$\frac{L(a, \hat{\pi})}{L(a, \tilde{\pi})} = \inf_{\pi \in \Pi_0} \frac{L(a, \hat{\pi})}{L(a, \pi)}.$$

Or:

$$\operatorname{Log} \frac{L(a, \hat{\pi})}{L(a, \tilde{\pi})} = \sum_{i<j} \left| u_{ij}(\hat{\pi}_{ij}) - u_{ij}(\tilde{\pi}_{ij}) \right|_j ;$$

les termes de la somme tels que $\tilde{\pi}_{ij} = \hat{\pi}_{ij}$ vont disparaître, il va rester:

$$\text{Log}\frac{L(a,\hat{\pi})}{L(a,\tilde{\pi})} = \sum_{\substack{(i,j)\\ \hat{\pi}_{ij} \neq \tilde{\pi}_{ij}}} \left| u_{ij}(\hat{\pi}_{ij}) - u_{ij}\left(\frac{1}{2}\right)\right|.$$

Or:

$$u_{ij}(\hat{\pi}_{ij}) = m_{ij}(\hat{\pi}_{ij}\,\text{Log}\,\hat{\pi}_{ij} + \hat{\pi}_{ji}\,\text{Log}\,\hat{\pi}_{ji})$$

en remplaçant s_{ij} par $m_{ij}\,\hat{\pi}_{ij}$ et $u_{ij}\left(\frac{1}{2}\right) = m_{ij}\,\text{Log}\,\frac{1}{2} = -\,m_{ij}\,\text{Log}\,2$.

Donc:

$$\text{Log}\frac{L(a,\hat{\pi})}{L(a,\tilde{\pi})} = \sum_{\substack{(i,j)\\ \hat{\pi}_{ij} \neq \tilde{\pi}_{ij}}} m_{ij}(\hat{\pi}_{ij}\,\text{Log}\,\hat{\pi}_{ij} + \hat{\pi}_{ji}\,\text{Log}\,\hat{\pi}_{ji} + \text{Log}\,2)$$

or: $\hat{\pi}_{ij}\,\text{Log}\,\hat{\pi}_{ij} + \hat{\pi}_{ji}\,\text{Log}\,\hat{\pi}_{ji}$ est minimum pour $\hat{\pi}_{ij} = \hat{\pi}_{ji} = \frac{1}{2}$, donc la quantité $u_{ij}(\hat{\pi}_{ij}) - u_{ij}\left(\frac{1}{2}\right)$ est d'autant plus petite que m_{ij} est petit et $\hat{\pi}_{ij}$ voisin de $\frac{1}{2}$.

Cette méthode nous conduira donc à inverser en priorité, des arcs de $\hat{G} \in \Omega^*(\hat{\pi})$ pour lesquels un petit nombre de comparaisons avaient été faites ou pour lesquels:

$$\hat{\pi}_{ij} = \frac{s_{ij}}{m_{ij}} \neq \frac{1}{2} \text{ c'est-à-dire } s_{ij} \neq s_{ji}.$$

Cas particulier: $m_{ij} = 1$ pour tout (i, j).

On a:

$$\hat{\pi}_{ij} = s_{ij} = 0 \text{ ou } 1 \qquad u_{ij}(\hat{\pi}_{ij}) = 0\,\text{Log}\,0 + 1\,\text{Log}\,1 = 0$$

et $\Omega^*(\hat{\pi}) = \{ G \}$ graphe observé.

On a alors:

$$\text{Log}\frac{L(a,\hat{\pi})}{L(a,\tilde{\pi})} = \sum_{\substack{(i,j)\\ \hat{\pi}_{ij} \neq \tilde{\pi}_{ij}}} \text{Log}\,2.$$

Or $\hat{\pi}$ est défini par:

$$\text{Log}\frac{L(a,\hat{\pi})}{L(a,\tilde{\pi})} = \inf_{\pi \in \pi_s} \text{Log}\frac{L(a,\hat{\pi})}{L(a,\tilde{\pi})}$$

donc:

$$\text{Log}\,\frac{L\,(a,\,\hat{\pi})}{L\,(a,\,\tilde{\pi})} = \text{Log}\,2 \times \text{d}\,(\text{G},\,\Omega_0)\,.$$

Donc ici l'estimation du classement, qui est donnée par l'ensemble des graphes de $\Omega^*\,(\tilde{\pi}) \cap \Omega_0$ est:

$$\Omega^*\,(\tilde{\pi}) \cap \Omega_0 = \{\,G_0 \in \Omega_0 \mid d\,(G,\,G_0) = d\,(G,\,\Omega_0)\,\}\,.$$

On retrouve donc les classements obtenus par la méthode de Slater.

Remarque

Dans le cas où $m_{ij} = m$ pour tout (i, j), De Cani (1969) utilise des méthodes de programmation linéaire pour déterminer l'estimation du maximum de vraisemblance des classements.

2) *Test basé sur le rapport de vraisemblance*

Comme nous l'avons vu dans l'introduction, nous admettons *a priori* l'hypothèse H_0^1: $\Omega^* \cap \Omega_0 \neq \emptyset$ et nous effectuons le test de H_0^2 contre H_1^2:

$$H_0^2\colon\ \Omega^* \cap \Omega_0 = \Omega_0 \Leftrightarrow \pi_{ij} = \frac{1}{2}\quad \forall\,(i, j)$$

$$H_1^2\colon\ \Omega^* \cap \Omega_0 \neq \Omega_0\ .$$

Le rapport de vraisemblance est:

$$\lambda = \frac{L\,(a \mid H_0^2)}{\sup\limits_{\pi} L\,(a \mid H_0^2 \cup H_1^2)}$$

$$\sup_{\pi} \text{L}\,(\text{a} \mid \text{H}_0^2 \cup \text{H}_1^2) = \sup_{\pi} L\,(a \mid \Omega^* \cap \Omega_0 \neq \emptyset) = \text{L}\,(a, \tilde{\pi})$$

donc:

$$\lambda = \frac{L\left(a, \frac{1}{2}\right)}{L\,(a,\,\tilde{\pi})}$$

notons:

$$\text{J} = \{\,(i, j),\quad i < j,\quad \hat{\pi}_{ij} = \tilde{\pi}_{ij}\,\}$$

$$\text{J'} = \{\,(i, j),\quad i < j,\quad \hat{\pi}_{ij} \neq \tilde{\pi}_{ij}\,\}$$

c'est-à-dire pour tout: $(i, j) \in \text{J'}$, $\tilde{\pi}_{ij} = \frac{1}{2}$.

Cela permet de simplifier λ:

$$\lambda = \frac{\prod\limits_{\text{J}} \left(\frac{1}{2}\right)^{m_{ij}}}{\prod\limits_{\text{J}} \hat{\pi}_{ij}^{\,s_{ij}}\ \hat{\pi}_{ji}^{\,s_{ji}}}$$

$$\text{Log } \lambda = - \sum_J m_{ij} \left[\hat{\pi}_{ij} \text{ Log } \hat{\pi}_{ij} + \hat{\pi}_{ji} \text{ Log } \hat{\pi}_{ji} + \text{Log } 2 \right].$$

La région critique R_c correspond aux petites valeurs de λ:

$$\lambda < K \Leftrightarrow \sum_J m_{ij} \left[\hat{\pi}_{ij} \text{ Log } \hat{\pi}_{ij} + \hat{\pi}_{ji} \text{ Log } \hat{\pi}_{ji} + \text{Log } 2 \right] > C \quad \left(\hat{\pi}_{ij} = \frac{s_{ij}}{m_{ij}} \right).$$

Si $m_{ij} = m$: la région critique correspond à:

$$\sum_J \left[\hat{\pi}_{ij} \text{ Log } \hat{\pi}_{ij} + \hat{\pi}_{ji} \text{ Log } \hat{\pi}_{ji} \right] + \text{Log } 2 \times | J | > C.$$

La méthode de De Cani (1969) permettant de déterminer $\tilde{\pi}$, nous donnera J, donc cette quantité est calculable.

C devra être déterminé en fonction du seuil α de manière à avoir:

$$P_r \left[\sum_J \left[\hat{\pi}_{ij} \text{ Log } \hat{\pi}_{ij} + \hat{\pi}_{ji} \text{ Log } \hat{\pi}_{ji} \right] + | J | \text{ Log } 2 > C \mid H_0^2 \right] = \alpha.$$

Ceci n'est pas facile à manier, cependant il sera toujours possible d'approcher C par des méthodes de simulation.

Si $m_{ij} = 1$:

$$| J | = C_n^2 - d (G, \Omega_0)$$

$$\lambda < K \Leftrightarrow C_n^2 - d (G, \Omega_0) > C \Leftrightarrow d (G, \Omega_0) < C'$$

on retrouve le test de Slater.

V. MODÈLES A VALEURS INTRINSÈQUES

Nous supposerons *a priori* que chaque objet x_i a une valeur intrinsèque π_i ($1 \leqq i \leqq n$).

Les π_{ij} seront alors définis en fonction des π_i suivant les modèles choisis.

Il est difficile de justifier *a priori* une forme particulière de π_{ij} mais le minimum que doit vérifier un modèle est que $\pi_{ij} = f(\pi_i, \pi_j)$ soit une fonction non décroissante en π_i et non croissante en π_j.

Les modèles qui ont été considérés sont en général de la forme: $\pi_{ij} = H (\pi_i - \pi_j)$ où H est une fonction non décroissante variant de 0 à 1 et telle que $H (- x) + H (x) = 1$, quel que soit x.

Ce modèle vérifie les conditions minima ci-dessus.

a) Modèle de Scheffé:

$$H (x) = \frac{1}{2} + x \qquad - \frac{1}{2} < x < \frac{1}{2}$$

qui nous donne:

$$\pi_{ij} = \frac{1}{2} + (\pi_i - \pi_j).$$

b) Modèle de Thurstone-Mosteller (ou distribution normale):

$$H(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt.$$

c) Modèle de Bradley-Terry. Il est basé sur l'idée, intuitivement séduisante, que le rapport:

$$\frac{\pi_{ij}}{\pi_{ji}} = \frac{\pi_i}{\pi_j}$$

comme $\pi_{ij} + \pi_{ji} = 1$, cela revient à poser:

$$\pi_{ij} = \frac{\pi_i}{\pi_i + \pi_j}.$$

Remarquons que ceci se ramène à la forme générale ci-dessus en posant:

$$H(x) = \frac{1}{1 + e^{-x}},$$

on a:

$$H(\pi_i - \pi_j) = \frac{e^{\pi_i}}{e^{\pi_i} + e^{\pi_j}}$$

en posant:

$$\pi_i' = e^{\pi_i}, \quad H(\pi_i - \pi_j) = \frac{\pi_i'}{\pi_i' + \pi_j'},$$

Nous allons étudier ce dernier modèle.

1) *Le modèle de Bradley-Terry*

On a m_{ij} répétitions de la comparaison (x_i, x_j).

On impose les conditions:

$$\begin{cases} \pi_i \geqq 0 \\ \sum_{i=1}^{n} \pi_i = 1 \end{cases}$$

(ces conditions ne sont pas restrictives, il est toujours possible de s'y ramener).

L'hypothèse de base du modèle est:

$$\pi_{ij} = \frac{\pi_i}{\pi_i + \pi_j}.$$

La vraisemblance s'écrit à présent:

$$L(a, \pi) = \prod_{i<j} \pi_{ij}^{s_{ij}} \pi_{ji}^{s_{ji}}$$

où:

$$s_{ij} = \sum_{k=1}^{m_{ij}} a_{ijk}$$

$$L(a, \pi) = \prod_{i<j} \frac{\pi_i^{s_{ij}} \pi_j^{s_{ji}}}{(\pi_i + \pi_j)^{m_{ij}}} = \frac{\prod_{i=1}^{n} \pi_i^{s_i}}{\prod_{i<j} (\pi_i + \pi_j)^{m_{ij}}}$$

où:

$$s_i = \sum_{j(\neq i)} s_{ij}.$$

Soit $S_i = \sum_{j(\neq i)} S_{ij}$, $(S_1, \ldots, S_n)$ constitue une statistique exhaustive pour $(\pi_1, \ldots, \pi_n)$.

Les équations du maximum de vraisemblance sont:

$$\frac{s_i}{\hat{\pi}_i} = \sum_{j(\neq i)} \frac{m_{ij}}{\hat{\pi}_i + \hat{\pi}_j} \qquad i = 1, 2, \ldots, n.$$

Ces équations ne se résolvent pas simplement, on trouvera dans Ford (1957) un processus itératif pour le calcul des $\hat{\pi}_i$.

2) *Étude du classement par les scores*

On a vu que, dans le cas du modèle de Bradley-Terry, le vecteur score $(S_1, S_2, \ldots, S_n)$ constitue une statistique exhaustive pour les paramètres du modèle.

Limitons-nous au cas: $m_{ij} = m$ *pour tout* (i, j).

En pratique, on utilise fréquemment dans ce cas un classement par les scores. Classer par les scores, c'est admettre implicitement que les scores contiennent toute l'information, c'est-à-dire qu'ils constituent une statistique exhaustive pour les paramètres π_{ij}.

Montrons que: la condition nécessaire et suffisante pour que $(S_1, \ldots, S_n)$ constitue une statistique exhaustive pour les paramètres π_{ij} est qu'il existe des valeurs $\pi_1, \ldots, \pi_n$ telles que:

$$\pi_{ij} = \frac{\pi_i}{\pi_i + \pi_j} .$$

Seule, la condition nécessaire reste à démontrer.

Démontrons d'abord que si $(S_1, \ldots, S_n)$ est une statistique exhaustive, la vraisemblance ne dépend que des scores:

$$L(a, \pi) = \prod_{i<j} \pi_{ij}^{s_{ij}} \pi_{ji}^{s_{ji}} \; L\left(a, \frac{1}{2}\right) = \left(\frac{1}{2}\right)^{m\binom{n}{2}} \quad \text{pour tout } a.$$

$(S_1, \ldots, S_n)$ étant une statistique exhaustive: $L(a, \pi) = \gamma(s, \pi)\, k(a)$ (où γ est une fonction qui ne dépend de a que par le vecteur score $s = (s_1, \ldots, s_n)$ et k est une fonction de a qui ne dépend pas de π).

$$L\left(a, \frac{1}{2}\right) = \left(\frac{1}{2}\right)^{m\binom{n}{2}} = \gamma\left(s, \frac{1}{2}\right) k(a) \neq 0 \qquad k(a) = \frac{1}{2^{m\binom{n}{2}}\, \gamma\left(s, \frac{1}{2}\right)} = k'(s)$$

$$L(a, \pi) = \gamma(s, \pi) \times k'(s) = \psi(s, \pi)$$

Donc la vraisemblance ne dépend de a que par l'intermédiaire des scores.

Soient $G_1, G_2, \ldots, G_m$ les graphes observés et supposons que (x_l, x_k, x_r) soit un circuit de G_1, c'est-à-dire:

$$a_{lk1} = 1 \qquad a_{kr1} = 1 \qquad a_{rl1} = 1.$$

Soit G_1' le graphe obtenu en changeant l'orientation de ce circuit:

$$a'_{kl1} = 1 \qquad a'_{rk1} = 1 \qquad a'_{lr1} = 1$$
$$a'_{ij1} = a_{ij1} \qquad \forall (i, j) \neq (l, k), (k, r), (r, l) .$$

Nous noterons $L(a, \pi)$ la vraisemblance de $G_1, G_2, \ldots, G_m$ et $L(a', \pi)$, la vraisemblance de $G_1', G_2, \ldots, G_m$.

$$\frac{L(a, \pi)}{L(a', \pi)} = \frac{\pi_{lk}\, \pi_{kr}\, \pi_{rl}}{\pi_{kl}\, \pi_{rk}\, \pi_{lr}} = \frac{\psi(s, \pi)}{\psi(s', \pi)}.$$

Il est évident que les scores n'ont pas été altérés par le changement d'orientation du circuit de longueur 3 (x_l, x_k, x_r).

$$s = s' \Rightarrow \frac{L(a, \pi)}{L(a', \pi)} = 1 = \frac{\pi_{lk}\, \pi_{kr}\, \pi_{rl}}{\pi_{kl}\, \pi_{rk}\, \pi_{lr}} \qquad \forall (l, k, r)$$

posons:

$$\pi_1 = 1 \qquad \pi_l = \frac{\pi_{l1}}{\pi_{1l}} \qquad \forall\, l\ (\neq 1)$$

$$\frac{\pi_{k1}\,\pi_{1l}\,\pi_{lk}}{\pi_{1k}\,\pi_{l1}\,\pi_{kl}} = 1 \qquad \frac{\pi_{kl}}{\pi_{lk}} = \frac{\pi_{k1}\,\pi_{1l}}{\pi_{1k}\,\pi_{l1}} = \frac{\pi_k}{\pi_l}$$

donc:

$$\forall\,(k, l) \qquad \frac{\pi_{kl}}{\pi_{lk}} = \frac{\pi_k}{\pi_l} \Leftrightarrow \forall\,(k, l) \qquad \pi_{kl} = \frac{\pi_k}{\pi_k + \pi_l}$$

On a donc démontré que si les scores constituent une statistique exhaustive, il existe $\pi_1, \pi_2, \ldots, \pi_n$ tels que π_{ij} se mette sous la forme:

$$\pi_{ij} = \frac{\pi_i}{\pi_i + \pi_j} \qquad \forall\,(i, j)\,.$$

Remarque

Il est facile de se ramener à $\Sigma\,\pi_i = 1$ en posant:

$$\pi_i' = \frac{\pi_i}{\Sigma\,\pi_i}$$

on aura:

$$\Sigma\,\pi_i' = 1 \qquad \text{et} \qquad \pi_{ij} = \frac{\pi_i'}{\pi_i' + \pi_j'}\,.$$

Remarquons que le classement obtenu en estimant les π_i par le maximum de vraisemblance, est le même que le classement par les scores.

En effet, les équations du maximum de vraisemblance donnent:

$$\frac{s_i}{\hat{\pi}_i} = m \sum_{k(\neq i)} \frac{1}{\hat{\pi}_i + \hat{\pi}_k}$$

d'où:

$$s_i - s_j = m\,(\hat{\pi}_i - \hat{\pi}_j) \sum_k \frac{\hat{\pi}_k}{\hat{\pi}_i + \hat{\pi}_j}$$

où $\displaystyle\sum_k \frac{\hat{\pi}_k}{\hat{\pi}_i + \hat{\pi}_j}$ est strictement positif.

Cela montre que:

$$s_i > s_j \Leftrightarrow \hat{\pi}_i > \hat{\pi}_j$$

$$s_i = s_j \Leftrightarrow \hat{\pi}_i = \hat{\pi}_j\,.$$

Buhlmann et Huber (1963) démontrent que le modèle de Bradley-Terry est le seul qui soit tel que le procédé de classement par les scores minimise uniformément le risque (parmi les procédés de classement invariants par permutation) pour toute fonction de perte « raisonnable ».

En conclusion, dans le cas où $m_{ij} = m$ pour tout (i, j):

a) le modèle de Bradley-Terry donne un classement identique au classement par les scores;

b) les scores contiennent toute l'information sur les π_{ij} (c'est-à-dire constituent une statistique exhaustive) si et seulement si le modèle de Bradley-Terry est vrai:

$$\pi_{ij} = \frac{\pi_i}{\pi_i + \pi_j}.$$

Il semble donc qu'adopter un classement par les scores comme étant un bon classement, en particulier dans ce sens qu'il tient compte de toute l'information disponible, est en quelque sorte équivalent à admettre que les π_{ij} s'expriment selon les hypothèses de Bradley-Terry.

3) *Tests de signification pour le modèle de Bradley-Terry*

L'hypothèse de base du modèle de Bradley-Terry est que, pour tout (i, j), π_{ij} peut s'écrire sous la forme:

$$\pi_{ij} = \frac{\pi_i}{\pi_i + \pi_j}.$$

C'est donc cette hypothèse que nous éprouverons en premier lieu:

$$H_0^3: \ \forall\ (i, j),\ \pi_{ij} = \frac{\pi_i}{\pi_i + \pi_j}$$

$$H_1^3: \ \exists\ \ (i, j),\ \pi_{ij} \neq \frac{\pi_i}{\pi_i + \pi_j}.$$

Nous pouvons rapprocher ce test du test I. En effet, nous avons vu dans l'introduction que $H_0^3 \subset H_0^1$; donc si ce test nous conduit à accepter H_0^3, nous aurons accepté H_0^1, mais le rejet de H_0^3 n'implique pas de rejet de H_0^1.

Le test sera basé sur le rapport de vraisemblance:

$$\lambda_1 = \frac{L\ (a, \hat{\pi}_i)}{L\ (a, \hat{\pi}_{ij})} = \frac{\prod_{i=1}^{n} \hat{\pi}_i^{s_i} \prod_{i<j} (\hat{\pi}_i + \hat{\pi}_j)^{-m_{ij}}}{\prod_{i<j} \hat{\pi}_{ij}^{s_{ij}}\ \hat{\pi}_{ji}^{s_{ji}}}$$

où:

$$\hat{\pi}_{ij} = \frac{s_{ij}}{m_{ij}}.$$

La région critique, de la forme $\lambda_1 < C_1$ s'écrira donc:

$$\sum_{i=1}^{n} s_i \operatorname{Log} \hat{\pi}_i - \sum_{i} \sum_{j} m_{ij} \operatorname{Log} (\hat{\pi}_i + \hat{\pi}_j) - \sum_{i} \sum_{j} (s_{ij} \operatorname{Log} s_{ij} + s_{ji} \operatorname{Log} s_{ji}) < K_1.$$

On montre que, si pour tout (i, j) $m_{ij} = m$, $-2 \operatorname{Log} \lambda_1$ converge en loi, lorsque m tend vers l'infini, vers un χ^2 à $\binom{n}{2} - n + 1$ degrés de liberté.

Test II

L'hypothèse H_0^2: $\pi_{ij} = \frac{1}{2}$ $\forall$ (i, j) s'écrit à présent: $\pi_i = \frac{1}{n}$ $\forall$ i.

$$H_0^2: \ \forall\, i = 1, \ldots, \ n \ \ \pi_i = \frac{1}{n}$$

$$H_1^2: \ \exists\, i \in \{ 1, \ldots, n \}: \pi_i \neq \frac{1}{n}.$$

Test basé sur le rapport de vraisemblance:

$$\lambda_2 = \frac{L\left(a, \frac{1}{n}\right)}{L\ (a, \hat{\pi}_i)} = \frac{\left(\frac{1}{2}\right)^{\sum_{i} \sum_{j} m_{ij}}}{\prod_{i=1}^{n} \hat{\pi}_i^{s_i} \prod_{i} \prod_{j} (\hat{\pi}_i + \hat{\pi}_j)^{-m_{ij}}}$$

$$\operatorname{Log} \lambda_2 = - \sum_{i} \sum_{j} m_{ij} \operatorname{Log} 2 - \sum_{i=1}^{n} s_i \operatorname{Log} \hat{\pi}_i + \sum_{i} \sum_{j} m_{ij} \operatorname{Log} (\hat{\pi}_i + \hat{\pi}_j)$$

on rejette H_0^2 si $\lambda_2 < C_2$ c'est-à-dire $B < K_2$ avec:

$$B = \sum_{i} \sum_{j} m_{ij} \operatorname{Log} (\hat{\pi}_i + \hat{\pi}_j) - \sum_{i=1}^{n} s_i \operatorname{Log} \hat{\pi}_i.$$

Dans le cas $m_{ij} = m$ pour tout (i, j), Bradley et Terry (1952) donnent des tables de la distribution de B sous H_0^2 et montrent que lorsque m tend vers l'infini, $-2 \operatorname{Log} \lambda_2$ converge en loi vers un χ^2 à $n - 1$ degrés de liberté.

Remarque

On peut aussi utiliser tout test équivalent du type χ^2 à la place du rapport des vraisemblances.

VI. CONCLUSION

Nous avons laissé de côté le problème posé par le test I de l'hypothèse H_0^1 contre H_1^1.

$$H_0^1\colon \forall\ (i, j, k)\colon \pi_{ij} > \frac{1}{2} \qquad \text{et} \qquad \pi_{jk} > \frac{1}{2} \Rightarrow \pi_{ki} \leqq \frac{1}{2}$$

$$H_1^1\colon \exists\ (i, j, k)\colon \pi_{ij} > \frac{1}{2},\ \pi_{jk} > \frac{1}{2},\ \pi_{ki} > \frac{1}{2}\ .$$

Il est facile de montrer qu'un test basé sur le rapport de vraisemblance conduit, dans le cas particulier de $\binom{n}{2}$ comparaisons sans répétition, à une région critique de la forme: $D > C$ (où D est la statistique utilisée dans le test de Slater: $D = d\ (\mathrm{G}, \Omega_0)$).

Si α est le seuil de signification choisi pour le test, C doit être déterminé par: $\mathrm{P}_r\ [D > C \mid H_0^1] \leqq \alpha$. La difficulté provient du fait que H_0^1 est une hypothèse composite.

Dans les méthodes d'estimation et de test exposées nous avons toujours admis, faute de savoir facilement la tester, l'hypothèse H_0^1.

Que signifie l'hypothèse H_1^1 ?

Cette hypothèse qui, dans le cas de comparaisons complètes sans répétition, serait acceptée pour les grandes valeurs de $D = d\ (\mathrm{G}, \Omega_0)$, signifie que toute relation de comparaison R* définie par $(x_i, x_j) \in \mathrm{R}^*$ dès que $\pi_{ij} > \frac{1}{2}$ n'est pas une relation d'ordre, puisqu'elle n'est pas transitive (tout graphe de Ω^* contient au moins un circuit de longueur 3: (x_i, x_j, x_k).).

Nous pouvons alors essayer d'expliquer l'existence de ces circuits par le fait que la relation de comparaison peut être le résultat de la combinaison de plusieurs relations d'ordre sur X, correspondant aux différents critères de choix employés par le juge.

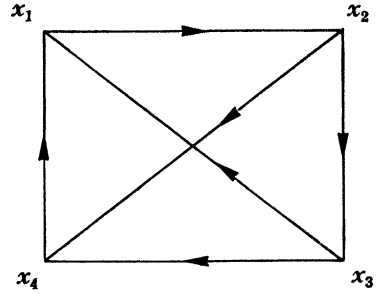
Prenons un exemple: $\mathrm{X} = \{ x_1, x_2, x_3, x_4 \}$.

Supposons qu'il existe trois relations d'ordre sur X, correspondant à trois critères différents:

$$\mathrm{R}_2^1\colon (x_1, x_2, x_3, x_4)$$
$$\mathrm{R}_0^2\colon (x_2, x_3, x_4, x_1)$$
$$\mathrm{R}_0^3\colon (x_3, x_4, x_1, x_2)\ .$$

Si on admet que le juge, comparant la paire (x_i, x_j), fait son choix de manière à être en accord avec la majorité des critères (c'est-à-dire qu'il préférera x_i à x_j si x_i figure avant x_j dans au moins deux des trois classements):

$$(x_1, x_2) \in \mathrm{R}_0^1 \cap \mathrm{R}_0^3 \Rightarrow (x_1, x_2) \in \mathrm{R}.$$



Le graphe observé G = (X, R) ci-dessus contient des circuits bien que les choix du juge soient parfaitement logiques. (Il s'agit de l'effet Condorcet dont on trouvera une étude détaillée dans Guilbaud (1968).)

Les problèmes qui se poseront seront alors: si la relation de comparaison observée contient « trop » de circuits pour être assimilable à une relation d'ordre, peut-on en déduire qu'elle est le résultat de la combinaison de plusieurs relations d'ordre ? Comment les déterminer ?

Notons enfin que Hayashi (1964) a envisagé des problèmes de ce genre: à partir d'un ensemble de comparaisons par paires, il étudie la répartition des objets dans un espace, dont il détermine la dimension, de manière à ce que les coordonnées des points sur chacun des axes représentent le classement des objets selon un des critères.

BIBLIOGRAPHIE

BARBUT M., "Note sur les ordres totaux à distance minimum d'une relation binaire donnée", *Math. Sci. hum.*, n° 17, 1966.

BARBUT M., et MONJARDET B., *Ordre et classification*, tomes 1 et 2, Paris, Hachette, 1970.

BERGE C., *Théorie des graphes et ses applications*, Dunod, 1958.

BRADLEY and TERRY, "The rank analysis of incomplete blocks designs, I, the method of paired comparisons", *Biometrika*, 39, pp. 324-45, 1952.

BRADLEY, "Incomplete block rank analysis : On the appropriateness of the model for a method of paired comparisons", *Biometrics*, 10, n° 3, 1954.

BRUNK, "Mathematical models for ranking from paired comparisons", *J. am. statist. Ass.*, 55, pp. 503 20, 1960.

BUHLMANN and HUBER, "Pairwise comparison and ranking in tournaments", *Ann. math. Statist.*, 34, n° 2, p. 501, 1963.

DAVID, *The method of paired comparisons*, Griffin's statistical monographs, 1963.

DE CANI J. S., "Maximum likelihood paired comparison ranking by linear programming", *Biometrika*, 56, 3, pp. 537-545, 1969.

FORD, "Solution of a ranking problem from binary comparisons", *Amer. math. Month.*, 64, pp. 28-33, 1957.

GUILBAUD G. Th., *Éléments de la théorie mathématique des jeux*, (Monographies de recherche opérationnelle n° 9), Dunod, 1968.

HAYASHI, "Multidimensional quantification of the data obtained by the method of paired comparison", *Ann. Inst. statist. Math.*, 16, pp. 231-245, 1964.

JACQUET-LAGRÈZE E., "L'agrégation des opinions individuelles", *Informatique en Sciences humaines*, n° 4, 1969.

KADANE, "Some equivalence classes in paired comparisons", *Ann. math. Statist.*, 37, n° 2, 1966.

KENDALL M. G. and BABINGTON SMITH, "On the method of paired comparisons", *Biometrika*, 31, pp. 324-345, 1940.

KENDALL M. G., "Further contributions to the theory of paired comparisons", *Biometrics*, 11, n° 1, pp. 43-62, 1955.

MONJARDET B., "Probabilité de l'effet Condorcet", *Math. Sci. hum.*, n° 23, 1968.

MOON, *Tropics on tournaments*, Holt, New York, 1968.

RAO and KUPPER, "Ties in paired comparisons experiments : A generalization of the Bradley-Terry model", *J. amer. statist. Ass.*, 62, n° 317, 1967.

REID, "On sets of arcs containing no cycles in a tournament", *Bull. canad. Math.*, 12, n° 3, pp. 261-264, 1969.

RUSSEL REMAGE and THOMPSON, "Maximum likelihood paired comparisons ranking", *Biometrika*, 53, n°s 1 et 2, p. 143, 1966.

SINGH J. and THOMPSON, "A treatment of ties in paired comparisons", *Ann. math. Statist.*, 39, pp. 2002-15, 1968.

SLATER, "Inconsistencies in a schedule of paired comparisons", *Biometrika*, 48, n°s 3 et 4, p. 303, 1961.

THOMPSON and RUSSEL REMAGE, "Rankings from paired comparisons", *Ann. math. Statist.*, 35, n° 2, p. 739, 1964.