

REVUE DE STATISTIQUE APPLIQUÉE

S. S. ZHANG

C. GENEST

Étude d'un test de confirmation des priorités dans le cadre du procédé d'analyse hiérarchique

Revue de statistique appliquée, tome 44, n° 2 (1996), p. 81-103

http://www.numdam.org/item?id=RSA_1996__44_2_81_0

© Société française de statistique, 1996, tous droits réservés.

L'accès aux archives de la revue « Revue de statistique appliquée » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

ÉTUDE D'UN TEST DE CONFIRMATION DES PRIORITÉS DANS LE CADRE DU PROCÉDÉ D'ANALYSE HIÉRARCHIQUE

S.S. Zhang, C. Genest

*Département de mathématiques et de statistique
Université Laval, Québec G1K 7P4*

RÉSUMÉ

Le procédé d'analyse hiérarchique de Saaty (1977) permet d'estimer l'importance relative qu'un individu accorde à des objets à partir de leur comparaison deux à deux sur une échelle numérique. Genest & Zhang (1995) ont proposé une façon de bâtir des régions de confiance pour le vecteur de priorités sous-jacent au moyen d'un pivot s'appuyant sur la métrique de Hilbert. Cet article examine le biais et la puissance d'un test statistique déduit de cette procédure. Des comparaisons sont ensuite faites avec le test du rapport de vraisemblance découlant de la formulation du problème en terme d'un modèle linéaire. Elles permettent de conclure à la similarité et à la complémentarité des deux approches.

Mots-clés : *analyse de la variance, comparaisons par paire, étude de Monte-Carlo, loi de l'étendue studentisée, métrique de Hilbert, procédé d'analyse hiérarchique, test de Tukey.*

SUMMARY

The Analytic Hierarchy Process was developed by Saaty (1977) as a means of eliciting individual priorities for a set of items, based on their pairwise comparison on a ratio scale. Genest & Zhang (1995) have shown how to exploit the pivotal properties of a statistic derived from Hilbert's metric to construct confidence regions for the underlying priority vector. This paper examines the bias and the power of a test deduced from this procedure. Numerical comparisons are also made with the likelihood ratio test suggested by an equivalent formulation of this problem in terms of a linear model. These results establish the similarity and complementarity of the two approaches.

Keywords : *analysis of variance, analytic hierarchy process, Hilbert's metric, Monte Carlo simulations, paired comparisons, Studentized range distribution, Tukey's test.*

1. Introduction

Imaginons que l'on veuille juger de la capacité de l'œil humain à percevoir la distance relative des choses en fonction de leur intensité lumineuse. Pour ce faire, on pourrait par exemple disposer $n \geq 3$ objets identiques à différents endroits dans une pièce obscure et demander à un sujet, assis dans un coin et coiffé d'un casque de

mineur, d'évaluer leur éloignement relatif en les éclairant de sa torche. Pour faciliter la tâche du répondant, il serait alors naturel d'exiger de lui qu'il procède au moyen de comparaisons par paire et que dans chaque cas, il indique lequel des deux objets est le plus éloigné, et par quel facteur.

En supposant que les objets soient numérotés de 1 à n et que chaque paire ait donné lieu à une seule comparaison, le jeu de données serait alors constitué de $n(n-1)/2$ observations $r_{ij} > 0$ reflétant, pour chaque $i \neq j$, le degré d'éloignement relatif de l'objet i par rapport à l'objet j . Si le répondant estimait par exemple que $r_{ij} = 2$ ou que $r_{ji} = 1/2$, cela signifierait qu'à son avis, l'objet i est deux fois plus éloigné de lui que l'objet j .

Réduit à sa plus simple expression, le procédé d'analyse hiérarchique de Saaty (1977, 1994) vise à analyser ce type de données en modélisant les réponses du sujet comme suit

$$r_{ij} = \frac{w_i}{w_j} \epsilon_{ij} \quad (1)$$

en terme de paramètres $w_1 > 0, \dots, w_n > 0$ (représentant ici le véritable degré relatif d'éloignement des n objets) et de perturbations aléatoires $\epsilon_{ij} > 0$ permettant de tenir compte d'éventuelles incohérences dans la formulation des réponses.

De façon générale, les paramètres w_i sont appelés des priorités, puisqu'ils fournissent un classement des objets sur une échelle numérique, soit en terme d'éloignement relatif comme ci-haut, soit en terme de toute autre caractéristique objective ou subjective, tel leur taille, leur poids, leur volume, leur valeur marchande, sentimentale, hédonique, et ainsi de suite. En principe, on devrait bien sûr avoir $r_{ij} = w_i/w_j$ pour tous $1 \leq i, j \leq n$. En pratique, toutefois, les difficultés inhérentes à la comparaison mentale de plusieurs objets mènent la plupart du temps à des incohérences dans les données; ainsi pourra-t-on avoir $r_{ij} \times r_{jk} \neq r_{ik}$ ou, pis encore, $r_{ij} > 1$ et $r_{jk} > 1$ mais $r_{ik} < 1$ pour certains $1 \leq i, j, k \leq n$. Ce sont ces aberrations que les termes ϵ_{ij} permettent de prendre en considération dans le modèle (1).

D'un point de vue statistique, l'analyse de ce type de données est facilitée par la reformulation de l'équation (1) en terme du modèle linéaire

$$y_{ij} = p_i - p_j + e_{ij} \quad (2)$$

dans lequel $y_{ij} = \log(r_{ij})$, $p_i = \log(w_i)$ et $e_{ij} = \log(\epsilon_{ij})$. Par souci de commodité, on notera dans la suite $r_{ji} = 1/r_{ij}$, $\epsilon_{ji} = 1/\epsilon_{ij}$ et $r_{ii} = \epsilon_{ii} = 1$ pour tous $1 \leq i, j \leq n$. Par voie de conséquence, on aura dès lors $y_{ji} = -y_{ij}$, $e_{ji} = -e_{ij}$ et $y_{ii} = e_{ii} = 0$ pour tous $1 \leq i, j \leq n$. de Jong (1984) signale que les estimations des moindres carrés des p_i sont de la forme

$$\hat{p}_i = \sum_{j=1}^n y_{ij}/n$$

à une constante additive près. La convention $\sum_{i=1}^n w_i = 1$, à laquelle on adhèrera ici, fait en sorte que ces estimations soient uniques. Si de plus les perturbations aléatoires

e_{ij} sont supposées d'espérance nulle et de même variance σ^2 , un estimateur sans biais de ce paramètre est alors donné par

$$\hat{\sigma}^2 = \frac{1}{(n-1)(n-2)} \sum_{i=1}^n \sum_{j=1}^n (y_{ij} - \hat{p}_i + \hat{p}_j)^2.$$

Sous les postulats de normalité et d'indépendance des e_{ij} pour $1 \leq i < j \leq n$, toutes ces estimations sont à vraisemblance maximale.

Dans des travaux récents motivés par une remarque de Saaty (1993), Genest & Zhang (1995) ont proposé de s'appuyer sur la métrique de Hilbert pour construire des régions de confiance pour le vecteur-colonne $w = (w_1, \dots, w_n)'$. Si $a = (a_1, \dots, a_n)'$ et $b = (b_1, \dots, b_n)'$ sont deux tels vecteurs à coordonnées strictement positives, la distance de Hilbert entre eux est définie par

$$D(a, b) = \log \left\{ \frac{\max_i (a_i/b_i)}{\min_i (a_i/b_i)} \right\}.$$

Partant de l'observation que la loi du rapport $\sqrt{n}D(\hat{w}, w)/\hat{\sigma}$ est la même quelles que soient les valeurs hypothétiques de w et de σ^2 , Genest & Zhang (1995) ont montré que l'ensemble $C_{1-\alpha}$ des vecteurs w pour lesquels

$$\frac{\max_i (\hat{w}_i/w_i)}{\min_i (\hat{w}_i/w_i)} \leq \exp(p_\alpha \hat{\sigma}/\sqrt{n})$$

constitue une région de confiance à $100 \times (1 - \alpha)\%$ pour w , dans la mesure où la valeur critique p_α représente le quantile d'ordre $1 - \alpha$ de la loi de l'étendue studentisée d'un échantillon aléatoire de taille n d'une population normale centrée réduite.

Le but de cet article est d'étudier les propriétés du test de l'hypothèse H_0 : $w = w^*$ déduit de cette procédure et d'en comparer le comportement à celui du test du rapport de vraisemblance découlant de la représentation du problème sous la forme d'un modèle linéaire. Les deux critères de test sont précisés au §2, ainsi que leur puissance sous les postulats de normalité et d'indépendance des perturbations aléatoires e_{ij} pour tous $1 \leq i < j \leq n$. Après avoir constaté l'absence de biais des deux tests au §3, on compare leur puissance numériquement, dans le cas spécial où $n = 3$ et $w^* = (1/3, 1/3, 1/3)'$ d'abord, puis pour de plus grandes valeurs de n par un choix judicieux de contre-hypothèses contiguës. Les résultats, illustrés par des graphiques au §4, font ressortir la similarité et la complémentarité des deux approches. Les simulations dont on fait état au §5 permettent également de voir qu'on peut s'affranchir du postulat de normalité sans forcément mettre en péril la qualité de l'inférence fondée sur ces tests concurrents.

2. Présentation des tests et calcul de leur puissance

Dans le cadre du modèle linéaire (2), la statistique canonique permettant de tester l'hypothèse H_0 : $w = w^*$ s'exprime en terme des $p_i^* = \log(w_i^*)$ comme le

rapport pondéré $F = n_2 S_1^2 / n_1 S_2^2$ de deux sommes de carrés résiduelles, à savoir :

$$S_1^2 = \frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^n (\hat{p}_i - p_i^* - \hat{p}_j + p_j^*)^2 \text{ et } S_2^2 = \frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^n (y_{ij} - \hat{p}_i + \hat{p}_j)^2,$$

lesquelles sont indépendantes et respectivement de loi du khi-deux à $n_1 = n - 1$ et $n_2 = (n - 1)(n - 2)/2$ degrés de liberté sous les postulats d'indépendance et de normalité des perturbations aléatoires mentionnés au §1. Ce test du rapport de vraisemblance rejette H_0 au seuil de $100 \times \alpha\%$ dès que F est supérieur au quantile $F_{n_1, n_2}(\alpha)$ d'ordre $1 - \alpha$ de la loi de Fisher-Snedecor. Quant au test déduit de la procédure de Genest & Zhang (1995), il rejettera H_0 lorsque le vecteur w^* n'appartient pas à la région de confiance $C_{1-\alpha}$ définie plus haut, c'est-à-dire si la statistique $Q = \sqrt{n}D(\hat{w}, w^*)/\hat{\sigma}$ est supérieure au quantile p_α d'ordre $1 - \alpha$ de la loi de l'étendue studentisée.

Les propositions suivantes permettent de déterminer la puissance et l'absence de biais des deux tests pour une contre-hypothèse générale $w \neq w^*$. Dans la suite, I_n dénote la matrice identité de taille n et e_n représente le vecteur-colonne $(1, \dots, 1)'$ de longueur n .

Proposition 2.1.

Supposons que pour $1 \leq i < j \leq n$, les perturbations aléatoires e_{ij} du modèle (2) soient mutuellement indépendantes, de loi normale d'espérance nulle et de variance σ^2 . Sous l'hypothèse nulle $H_0 : w = w^*$, la statistique F obéit à une loi de Fisher-Snedecor à $n_1 = n - 1$ et $n_2 = (n - 1)(n - 2)/2$ degrés de liberté. Lorsque $w \neq w^*$, elle est décentrée et a pour paramètre d'excentricité $\lambda = n(p - p^*)'B(p - p^*)/\sigma^2$, où $B = I - ee'/n$ est la matrice dite de centrage. Dans ce dernier cas, la probabilité que la statistique F prenne une valeur supérieure à $c > 0$ est donnée par :

$$1 - \sum_{k=0}^{\infty} \left(\frac{\lambda^k}{2^k k!} e^{-\lambda/2} \right) \frac{\Gamma\left(\frac{n_1 + n_2}{2} + k\right)}{\Gamma\left(\frac{n_1}{2} + k\right) \Gamma\left(\frac{n_2}{2}\right)} \int_0^{\frac{cn_1}{cn_1 + n_2}} t^{\frac{n_1}{2} + k - 1} (1 - t)^{\frac{n_2}{2} - 1} dt. \quad (3)$$

Démonstration.

C'est une application immédiate du deuxième théorème fondamental de la théorie des moindres carrés, tel qu'énoncé en page 191 de l'ouvrage de Rao (1973). Pour le voir, dénotons par Y et E les vecteurs-colonnes de longueur $n(n - 1)/2$ ayant respectivement pour composantes les réalisations de y_{ij} et de e_{ij} , pour $1 \leq i < j \leq n$. On peut alors écrire le modèle (2) sous la forme matricielle classique $Y = Xp + E$, où X est une matrice d'incidence de dimension $n(n - 1)/2 \times n$ comportant exactement deux entrées non-nulles par ligne, à savoir un 1 et un -1. L'hypothèse nulle $H_0 : w = w^*$ est équivalente à la condition que p et p^* ne diffèrent que par une constante, soit encore $H'(p - p^*) = 0$, où H est une matrice de dimension $n \times (n - 1)$

définie comme suit :

$$H = \begin{pmatrix} e'_{n-1} \\ -I_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 \end{pmatrix}.$$

D'après le résultat énoncé par Rao (1973), les sommes de carrés S_1^2 et S_2^2 définies plus haut sont indépendantes et suivent respectivement des lois du khi-deux à $n-1$ et à $(n-1)(n-2)/2$ degrés de liberté sous l'hypothèse nulle, puisque $\text{rang}(X) = \text{rang}(H) = n-1$. Rao mentionne de plus que S_1^2 est une variable aléatoire du khi-deux décentrée lorsque H_0 n'est pas vérifiée. Un calcul simple montre que son paramètre d'excentricité est alors égal à $\lambda = n(p-p')'B(p-p')/\sigma^2$, tel qu'annoncé.

La formule permettant de calculer la probabilité que la statistique F prenne une valeur supérieure à $c > 0$ se déduit immédiatement des résultats exposés dans le chapitre 30 du livre de Johnson & Kotz (1970).□

La définition suivante facilitera l'énoncé de la deuxième proposition.

Définition 2.2.

Soient $Z_1, \dots, Z_n$ un échantillon aléatoire de loi normale centrée réduite et T une variable aléatoire du khi-deux à m degrés de liberté indépendante de tous les Z_i . On appelle loi de l'étendue studentisée de paramètres n et m la loi du rapport $\{\max_i(Z_i) - \min_i(Z_i)\}/\sqrt{T/m}$. De plus, on nomme loi de l'étendue studentisée décentrée de mêmes paramètres la loi du rapport

$$\{\max_i(Z_i + \theta_i) - \min_i(Z_i + \theta_i)\}/\sqrt{T/m}. \quad (4)$$

L'excentricité de la distribution est alors caractérisée, à un scalaire près, par le vecteur-colonne $\Theta = (\theta_1, \dots, \theta_n)'$.

Proposition 2.3.

Supposons que pour $1 \leq i < j \leq n$, les perturbations aléatoires e_{ij} du modèle (2) soient mutuellement indépendantes, de loi normale d'espérance nulle et de variance σ^2 . Sous l'hypothèse nulle $H_0 : w = w^*$, la statistique $Q = \sqrt{n}D(\hat{w}, w^*)/\hat{\sigma}$ obéit à la loi de l'étendue studentisée de paramètres n et $n_2 = (n-1)(n-2)/2$. Lorsque $w \neq w^*$, cette loi est décentrée et a pour paramètre d'excentricité $\Theta = \sqrt{n}(p-p')/\sigma$. Dans ce dernier cas, la probabilité que Q prenne une valeur supérieure à $c > 0$ est donnée par la formule

$$1 - \sum_{i=1}^n \int_0^\infty \int_{-\infty}^\infty \varphi(x - \theta_i) \prod_{i \neq j=1}^n \{\Phi(x - \theta_j + cs/\sqrt{n_2}) - \Phi(x - \theta_j)\} \psi(s) dx ds, \quad (5)$$

exprimée en terme de la fonction de densité ψ d'une variable khi à n_2 degrés de liberté¹, de la fonction de densité normale centrée réduite φ et de la fonction de répartition correspondante Φ .

Démonstration.

Il n'est pas nécessaire de vérifier la première partie de la proposition, puisqu'elle se déduit de la deuxième en posant $w^* = w$. Pour déterminer la loi de la statistique Q de façon générale, on part du fait que $n_2 \hat{\sigma}^2 / \sigma^2 = S_2^2$ est une variable du khi-deux à $n_2 = (n-1)(n-2)/2$ degrés de liberté quelle que soit la valeur de w . Tel qu'établi dans la démonstration de la Proposition 6.1 de l'article de Genest & Zhang (1995), cette somme de carrés est indépendante du vecteur $\bar{e} = (\bar{e}_1, \dots, \bar{e}_n)'$ dont

les éléments sont définis par $\bar{e}_i = \sum_{j=1}^n e_{ij} / n$. Or

$$\frac{\sqrt{n}D(\hat{w}, w^*)}{\sigma} = \frac{\sqrt{n}}{\sigma} \{ \max_i (\bar{e}_i + p_i - p_i^*) - \min_i (\bar{e}_i + p_i - p_i^*) \}.$$

L'argument sera donc complet si on peut réexprimer cette variable, appelons-la $\tilde{Q}$, sous la même forme que le numérateur de (4) en terme de variables aléatoires indépendantes $Z_1, \dots, Z_n$ de loi normale centrée réduite.

Pour ce faire, on doit d'abord remarquer que le vecteur-colonne $V = (V_1, \dots, V_n)'$ de composantes

$$V_i = \frac{n}{\sigma \sqrt{n-1}} (\bar{e}_i + p_i - p_i^*)$$

obéit à une loi normale multidimensionnelle d'espérance $\mu = n(p - p^*) / \sigma \sqrt{n-1}$ et de covariance $\Sigma = (1 - \rho)I_n + \rho e_n e_n'$ où $\rho = -1/(n-1)$. Ceci découle directement du fait, déjà souligné dans la démonstration de la Proposition 3.1 de l'article de Genest & Zhang (1995), que le vecteur $\bar{e}$ est lui-même de loi normale multivariée d'espérance nulle et de structure d'équivariance $B\sigma^2/n$. D'après un résultat de Gupta, Pillai & Steck (1964), il est alors possible d'exprimer les variables V_i sous la forme

$$V_i - \mu_i = \sqrt{-\rho} Z_0 + \sqrt{1 - \rho} Z_i$$

en fonction de variables aléatoires normales centrées réduites indépendantes $Z_1, \dots, Z_n$ et d'un aléa supplémentaire Z_0 de même loi tel que $\text{cov}(Z_i, Z_0) = -\sqrt{-\rho}/\sqrt{1 - \rho}$

¹ Par définition, $\psi(s) = 2^{1-n_2/2} s^{n_2-1} \exp(-s^2/2) / \Gamma(n_2/2)$ pour $s > 0$ et vaut zéro partout ailleurs.

pour tout $1 \leq i \leq n$. Dès lors, on a

$$\begin{aligned}\tilde{Q} &= \sqrt{\frac{n-1}{n}} \left\{ \max_i(V_i) - \min_i(V_i) \right\} \\ &= \max_i \left(Z_i + \frac{\mu_i}{\sqrt{1-\rho}} \right) - \min_i \left(Z_i + \frac{\mu_i}{\sqrt{1-\rho}} \right) \\ &= \max_i(Z_i + \theta_i) - \min_i(Z_i + \theta_i)\end{aligned}$$

puisque $\mu_i = \sqrt{1-\rho} \theta_i$ pour tout $1 \leq i \leq n$. Ceci achève la démonstration de la deuxième partie de la proposition.

Enfin, pour déduire la formule (5), remarquons que

$$\begin{aligned}P(Q \leq c) &= \int_0^\infty P\left(\tilde{Q} \leq \frac{cs}{\sqrt{n_2}} \mid \frac{\sqrt{n_2}\hat{\sigma}}{\sigma} = s\right) \psi(s) ds \\ &= \int_0^\infty P\left(\tilde{Q} \leq \frac{cs}{\sqrt{n_2}}\right) \psi(s) ds,\end{aligned}$$

puisque les variables $\tilde{Q}$ et $\hat{\sigma}$ sont indépendantes et que le rapport $\sqrt{n_2}\hat{\sigma}/\sigma$ obéit à une loi khi à n_2 degrés de liberté. Le résultat est alors une conséquence immédiate de l'exercice 2.3.2 de l'ouvrage de David (1981), lequel stipule que

$$P(\tilde{Q} \leq q) = \sum_{i=1}^n \int_{-\infty}^{\infty} \varphi(x - \theta_i) \prod_{i \neq j=1}^n \{\Phi(x - \theta_j + q) - \Phi(x - \theta_j)\} dx$$

quel que soit $q > 0$. $\square$

3. Absence de biais des tests

Il est bien connu que sous les postulats d'indépendance et de normalité, le test du rapport de vraisemblance F est sans biais. Pour voir que le test fondé sur la statistique $Q = \sqrt{n}D(\hat{w}, w^*)/\hat{\sigma}$ l'est également, on s'inspire de l'approche utilisée par Ramachandran (1956) pour démontrer l'absence de biais d'un test d'égalité des moyennes proposé par Tukey dans un contexte d'analyse de la variance.

À cette fin, on opère d'abord le changement de variable $t = x - \theta_i$ dans le $i^{\text{ème}}$ terme de la somme (5), ce qui permet de réécrire cette formule sous la forme $1 - \beta$ avec

$$\beta = \sum_{i=1}^n \int_0^\infty \int_{-\infty}^{\infty} \varphi(t) \prod_{i \neq j=1}^n \{\Phi(t + \theta_i - \theta_j + cs/\sqrt{n_2}) - \Phi(t + \theta_i - \theta_j)\} \psi(s) dt ds.$$

On procède ensuite à une reparamétrisation du vecteur Θ en posant $\eta_{i-1} = \theta_1 - \theta_i$ pour $1 \leq i \leq n$. Il vient alors

$$\beta = \sum_{i=1}^n \int_0^\infty \int_{-\infty}^\infty \varphi(t) \prod_{i \neq j=1}^n \{ \Phi(t + \eta_{j-1} - \eta_{i-1} + cs/\sqrt{n_2}) - \Phi(t + \eta_{j-1} - \eta_{i-1}) \} \psi(s) dt ds,$$

ce qui s'écrit encore

$$\begin{aligned} \beta = \sum_{i=1}^{n-1} \int_0^\infty \psi(s) \int_{-\infty}^\infty \varphi(z) \int_{z-\eta_i}^{z-\eta_i+cs/\sqrt{n_2}} \varphi(t) dt \prod_{i \neq j=1}^{n-1} \int_{z-\eta_i+\eta_j}^{z-\eta_i+\eta_j+cs/\sqrt{n_2}} \varphi(t) dt dz ds \\ + \int_0^\infty \psi(s) \int_{-\infty}^\infty \varphi(z) \prod_{j=1}^{n-1} \int_{z+\eta_j}^{z+\eta_j+cs/\sqrt{n_2}} \varphi(t) dt dz ds \quad (6) \end{aligned}$$

après avoir isolé le terme $i = 1$ et remplacé $i - 1$ par i dans la somme restante. L'expression (6) correspondant en tout point à l'équation (3.2) de l'article de Ramachandran, la démarche exposée au §4 de son travail permet de conclure. Vu ce lien entre le test de Tukey et celui déduit de la procédure de Genest & Zhang (1995), on dira par la suite que Q est la statistique de Tukey.

4. Comparaisons numériques

Pour se faire une idée comparative de la puissance du test classique et de celui de Tukey, des graphiques illustrant la fluctuation de la probabilité de rejeter l'hypothèse nulle $H_0 : w = e/n$ en fonction de la vraie valeur de w ont été produits à partir de simulations de type Monte-Carlo dans le cas où $n = 3$ pour des tests de seuil nominal $\alpha = 5\%$. Cette approche a été préférée à l'évaluation numérique des formules (3) et (5) pour des raisons de coût et de temps de calcul. Chaque graphique est basé sur 10 000 observations de la valeur de la statistique calculées à partir de matrices de réponses d'éléments r_{ij} de la forme (1) où $\log(\epsilon_{ij})$ obéit à une loi normale centrée de variance $\sigma^2 = 1, 1/4$ ou $1/100$ selon le cas. Les figures 1-7 sont typiques des résultats obtenus. Les courbes d'isoprobabilité des deux fonctions de puissance y sont représentées sur un système d'axes ayant w_1 pour abscisse et w_2 pour ordonnée. Ces graphiques font clairement ressortir la très grande similarité de comportement des deux tests; ils montrent également que ni l'un ni l'autre de ces tests n'est uniformément plus puissant que son compétiteur.

Afin de pousser la comparaison graphique au delà du cas $n = 3$, il est nécessaire de spécifier une série de contre-hypothèses contiguës dépendant d'un paramètre unidimensionnel, disons $\gamma \geq 0$. Un choix commode consiste à prendre

$$w_i(\gamma) = \frac{\exp \{ \gamma(n - 2i + 1)/n \}}{\sum_{j=1}^n \exp \{ \gamma(n - 2j + 1)/n \}} \quad (7)$$

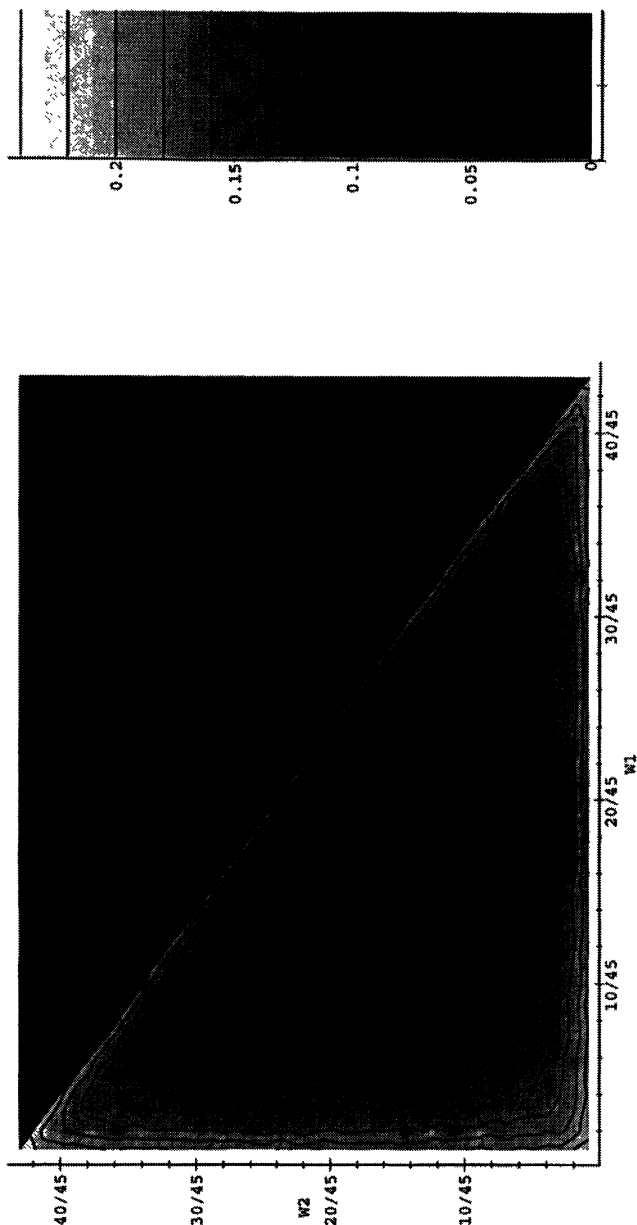


FIGURE 1 :

Puissance observée de la procédure de Tukey lors de 10 000 tests de l'hypothèse $H_0 : w = (1/3, 1/3, 1/3)'$ au seuil de 5%, lorsque les perturbations pseudo-aléatoires affectant les comparaisons par paire suivent une loi normale d'espérance nulle et de variance $\sigma^2 = 1$ sur l'échelle logarithmique.

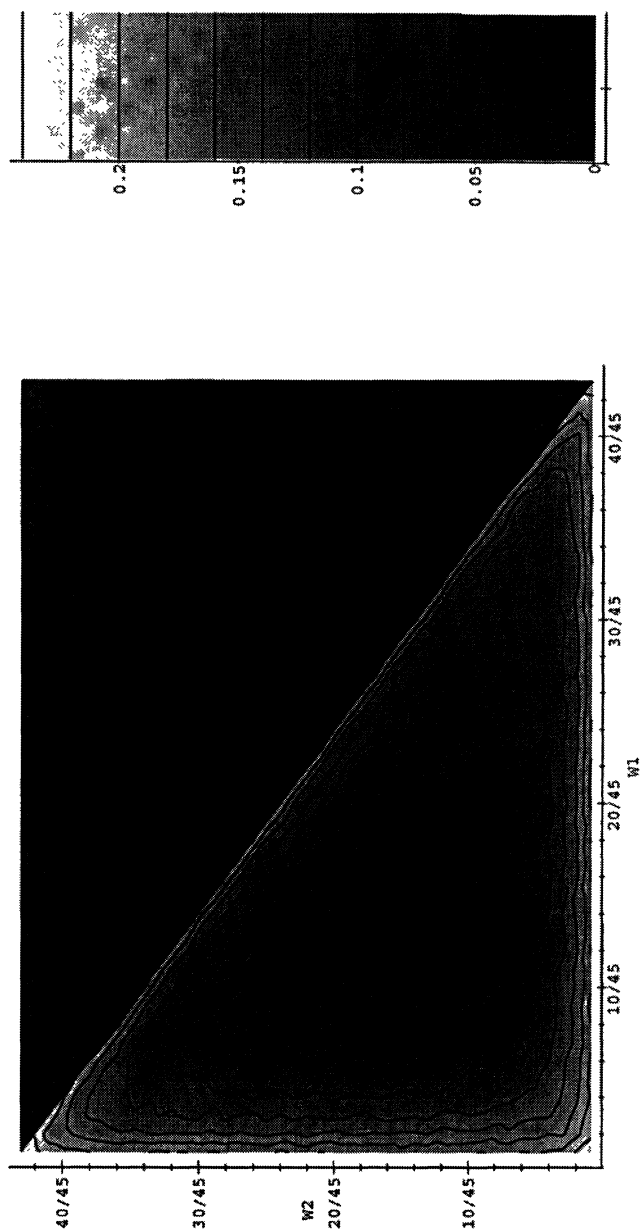


FIGURE 2 :

Puissance observée de la procédure classique lors de 10 000 tests de l'hypothèse $H_0 : w = (1/3, 1/3, 1/3)'$ au seuil de 5%, lorsque les perturbations pseudo-aléatoires affectant les comparaisons par paire suivent une loi normale d'espérance nulle et de variance $\sigma^2 = 1$ sur l'échelle logarithmique.

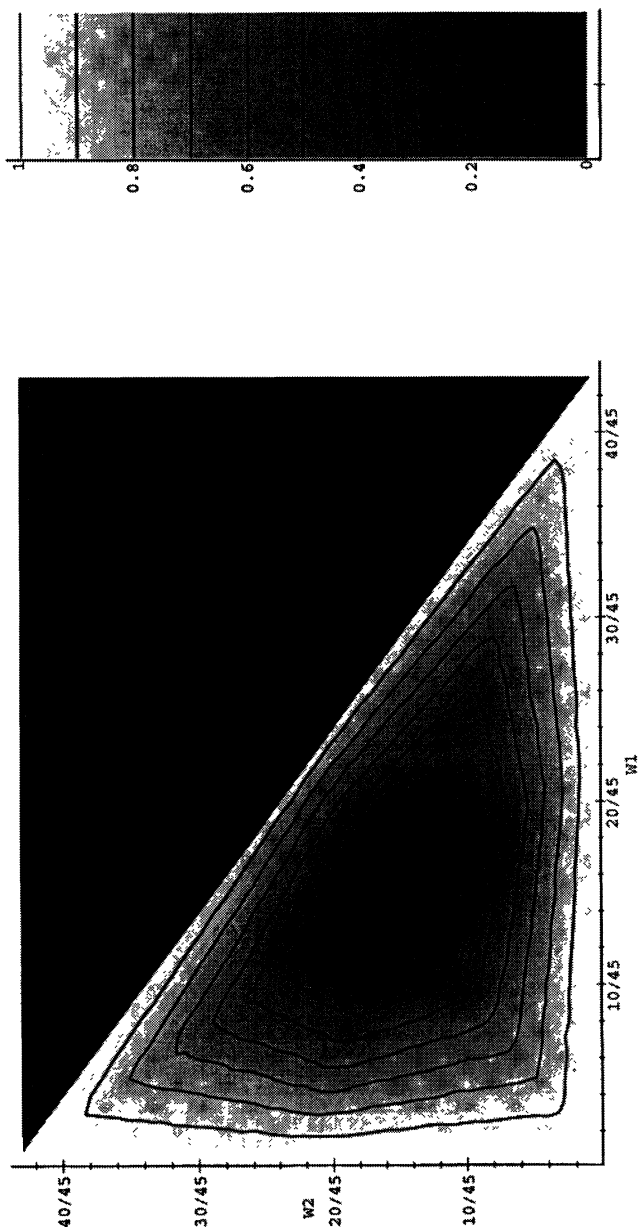


FIGURE 3 :

Puissance observée de la procédure de Tukey lors de 10 000 tests de l'hypothèse $H_0 : w = (1/3, 1/3, 1/3)'$ au seuil de 5%, lorsque les perturbations pseudo-aléatoires affectant les comparaisons par paire suivent une loi normale d'espérance nulle et de variance $\sigma^2 = 1/100$ sur l'échelle logarithmique.

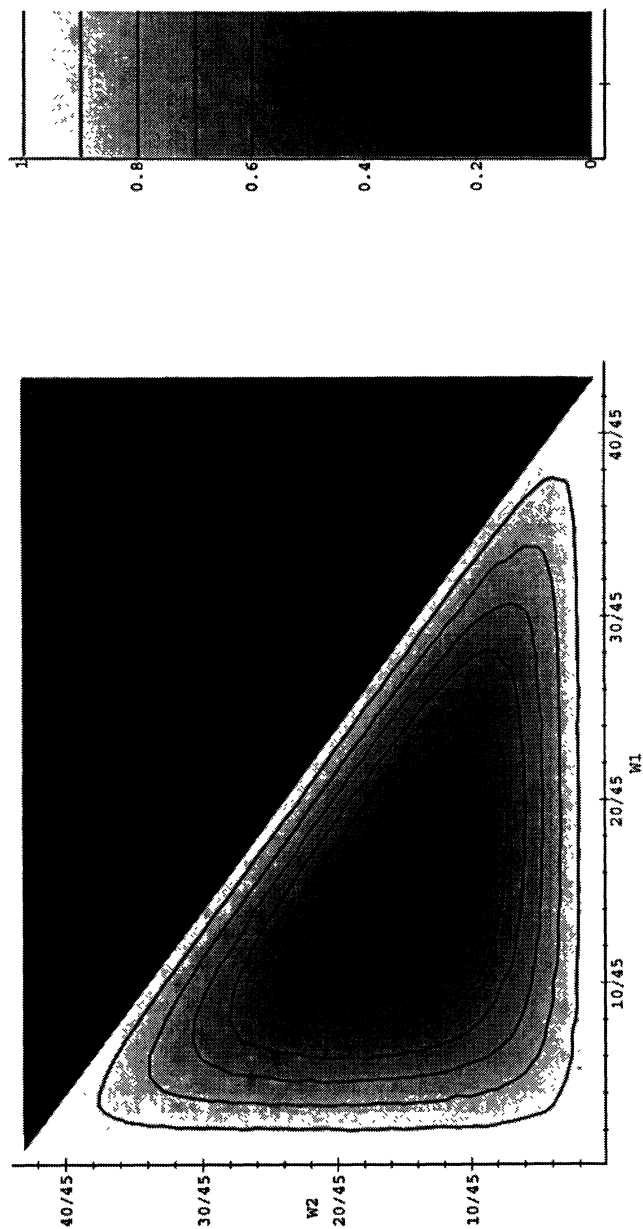


FIGURE 4 :

Puissance observée de la procédure classique lors de 10 000 tests de l'hypothèse $H_0 : w = (1/3, 1/3, 1/3)'$ au seuil de 5%, lorsque les perturbations pseudo-aléatoires affectant les comparaisons par paire suivent une loi normale d'espérance nulle et de variance $\sigma^2 = 1/100$ sur l'échelle logarithmique.

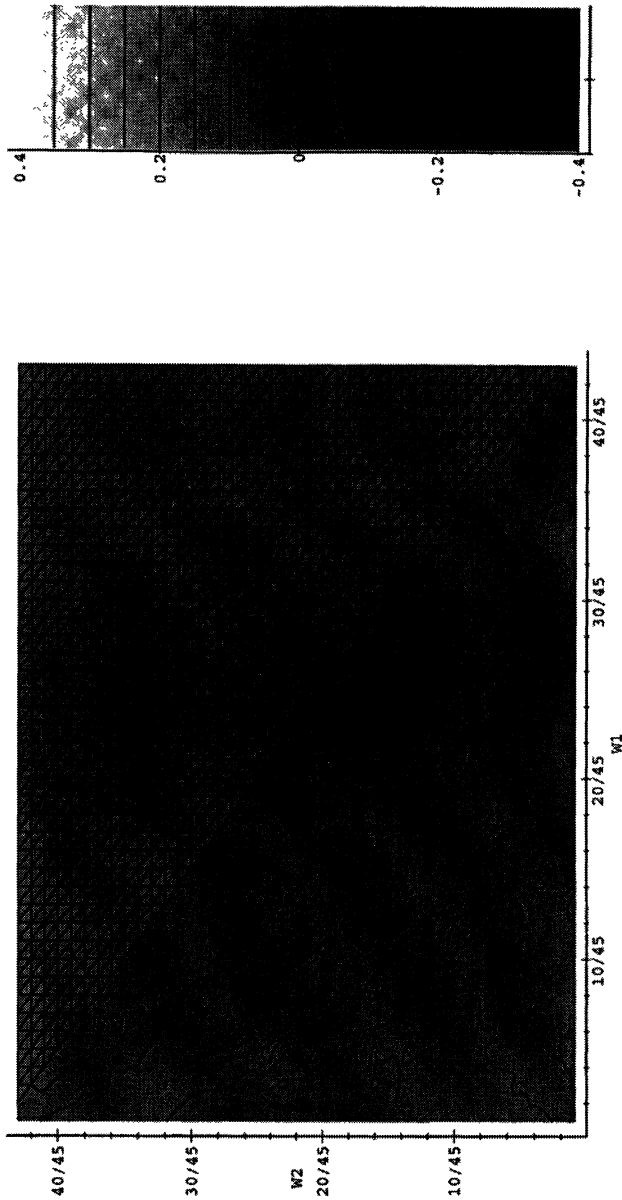


FIGURE 5 :

Différence observée entre la puissance de la procédure de Tukey et celle de la procédure classique lors de 10 000 tests de l'hypothèse $H_0 : w = (1/3, 1/3, 1/3)'$ au seuil de 5%, lorsque les perturbations pseudo-aléatoires affectant les comparaisons par paire suivent une loi normale d'espérance nulle et de variance $\sigma^2 = 1$ sur l'échelle logarithmique.

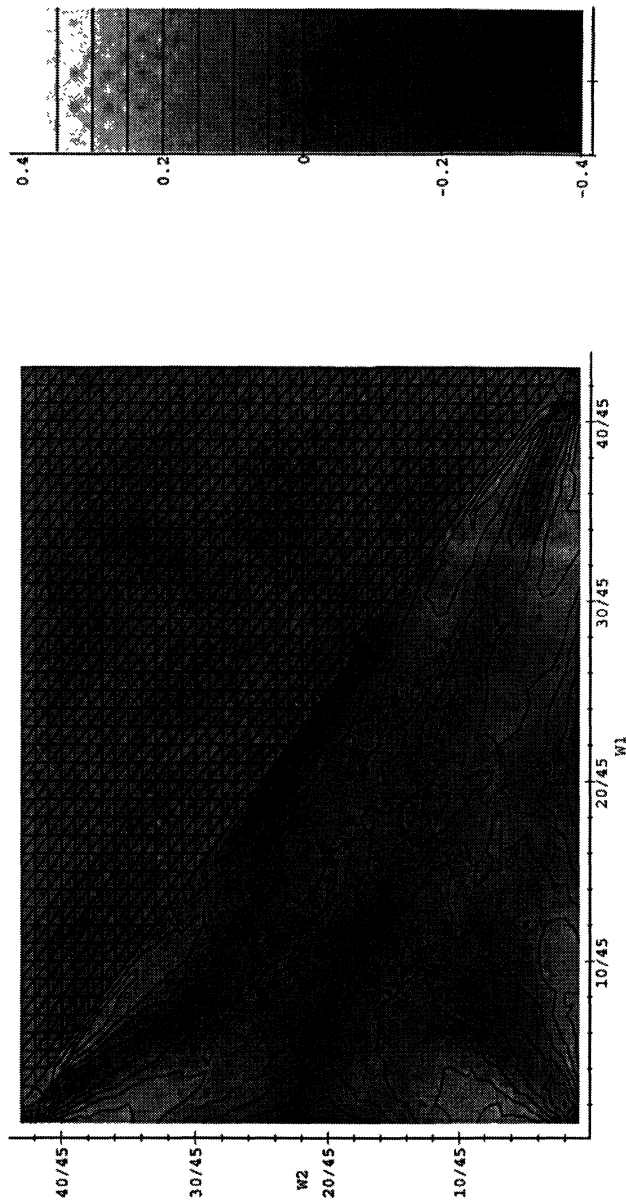


FIGURE 6 :

Différence observée entre la puissance de la procédure de Tukey et celle de la procédure classique lors de 10 000 tests de l'hypothèse $H_0 : w = (1/3, 1/3, 1/3)'$ au seuil de 5%, lorsque les perturbations pseudo-aléatoires affectant les comparaisons par paire suivent une loi normale d'espérance nulle et de variance $\sigma^2 = 1/4$ sur l'échelle logarithmique.

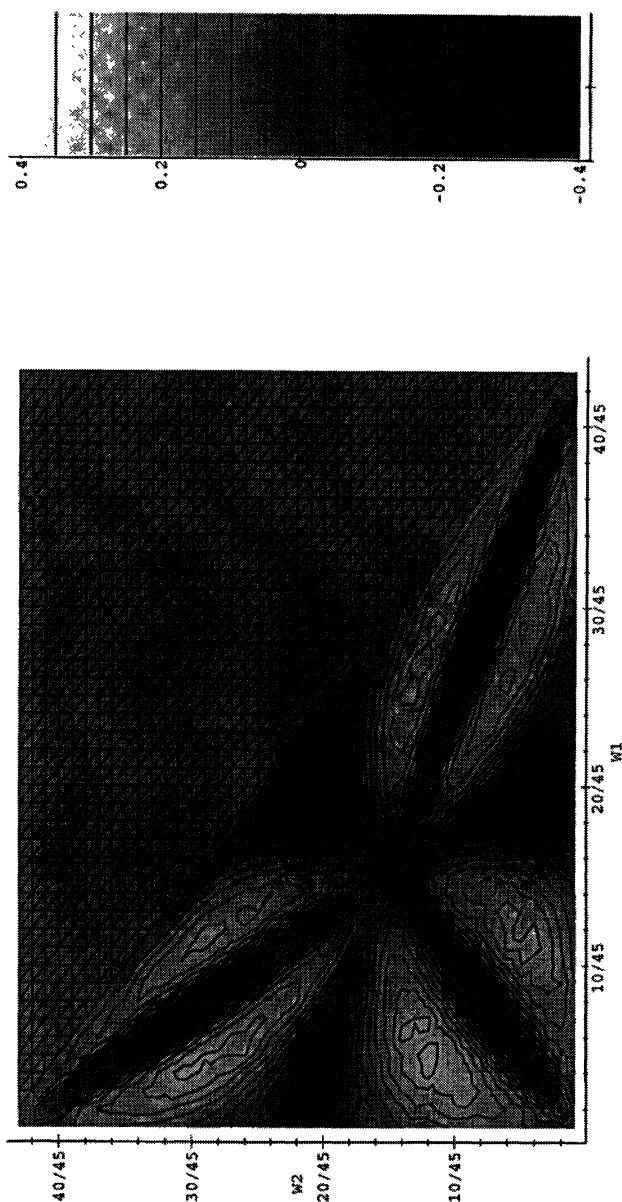


FIGURE 7 :

Différence observée entre la puissance de la procédure de Tukey et celle de la procédure classique lors de 10 000 tests de l'hypothèse $H_0 : w = (1/3, 1/3, 1/3)'$ au seuil de 5%, lorsque les perturbations pseudo-aléatoires affectant les comparaisons par paire suivent une loi normale d'espérance nulle et de variance $\sigma^2 = 1/100$ sur l'échelle logarithmique.

pour $1 \leq i \leq n$, de sorte que $\gamma = 0$ corresponde à l'hypothèse nulle $H_0 : w = e/n$. La figure 8 permet de visualiser la courbe que décrit le vecteur $w(\gamma)$ dans le simplexe lorsque $n = 3$. Ce choix de contre-hypothèses se justifie du fait que le vecteur dont les composantes sont définies en (7) est à la fois la solution de Saaty et celle des moindres carrés logarithmiques au problème de l'estimation des priorités d'une matrice de réponses cohérente au sens ordinal dont tous les éléments à l'extérieur de la diagonale principale seraient égaux à $a = e^\gamma > 1$ ou à son inverse multiplicatif. Pour de plus amples détails sur le rôle central que jouent les matrices de ce type dans l'analyse de comparaisons par paire exprimées sur une échelle ordinale, on peut se référer à l'article de Genest, Lapointe & Drury (1993).

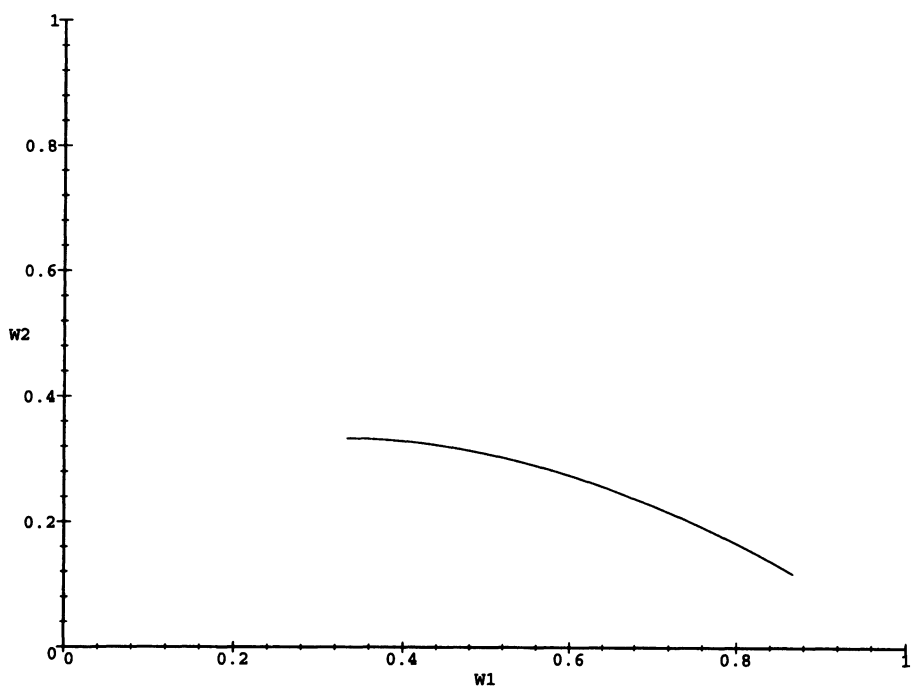


FIGURE 8 :

Lieu géométrique des points $w = (w_1, w_2)'$ de composantes $w_i(\gamma)$ définies par l'équation (7) pour $i = 1, 2$

Les figures 9 et 10 illustrent l'évolution de la puissance du test classique et du test de Tukey à mesure que la vraie valeur de $w(\gamma)$ s'éloigne du point e/n . Ces résultats ont été obtenus à partir de tests effectués aux seuils de 5% et de 1% sur 10 000 matrices de réponses de taille $n = 3, 5, 7$ et 9. Les simulations ont été réalisées en supposant une loi normale d'espérance nulle et de variance $\sigma^2 = 1, 1/4$ et $1/100$ pour le logarithme des perturbations aléatoires affectant les comparaisons par paire. La figure 9 montre que pour les contre-hypothèses choisies, les deux tests ont un comportement très semblable lorsque $n = 5$ et ce, sans égard à la valeur de σ^2 ou au seuil nominal des tests. Les résultats numériques confirment également l'absence de biais des deux tests rapportée au §3. Quant à la figure 10, elle met en évidence la détérioration graduelle du test de Tukey à mesure que n croît. Cette observation est toutefois limitée à la série de contre-hypothèses choisie, puisque ni l'un ni l'autre des deux tests n'est uniformément plus puissant que son compétiteur. Dans le cas d'espèce considéré, la supériorité du test classique reste d'ailleurs toute relative, puisque quand $n = 7$, la différence de puissance entre les deux tests est inférieure à 6% lorsque $\alpha = 0,05$ et à 8% quand $\alpha = 0,01$. Ces différences s'élèvent respectivement à 8,6% et à 12,8% lorsque $n = 9$.

Si on ajoute à ces différentes observations le fait que la méthode des comparaisons par paire devient rapidement impraticable pour des valeurs de n supérieures à 10, il apparaît que le choix du test à utiliser pour confirmer des priorités dans le contexte du procédé d'analyse hiérarchique se réduit essentiellement à une affaire de goût. Cette conclusion est d'ailleurs en tout point semblable à celle que David (1953) formulait à propos des mérites relatifs du test de Tukey et du test du rapport de vraisemblance dans le cadre de l'analyse de la variance.

5. Étude de robustesse

Les résultats théoriques et expérimentaux exposés dans les paragraphes précédents supposent tous la normalité du logarithme des perturbations aléatoires ϵ_{ij} du modèle (1). Afin d'évaluer dans quelle mesure le test de Tukey et son compétiteur classique dépendent de ce postulat, leur puissance a été étudiée sous d'autres conditions à l'aide de simulations de type Monte-Carlo. Trois genres de lois pour $e_{ij} = \log(\epsilon_{ij})$ ont été utilisés à cette fin, à savoir :

- (1) la loi uniforme sur l'intervalle $(-\sqrt{3}, \sqrt{3})$;
- (2) la loi de Laplace (ou double exponentielle) de densité $\exp(-\sqrt{2}|x|)/\sqrt{2}$ sur la droite réelle;
- (3) un mélange de deux lois normales $N(0, \tau^2)$ et $N\left(0, \frac{1 - \pi\tau^2}{1 - \pi}\right)$ en proportions $\pi > 0$ et $1 - \pi$ respectivement.

Toutes ces lois étant de variance un, les résultats de simulations sont directement comparables à ceux obtenus lorsque e_{ij} suit une loi normale centrée réduite. Ce choix de lois a été guidé par le fait que les ailes de la distribution normale sont plus déployées que celles de la loi uniforme, mais moins que celles des deux autres lois.

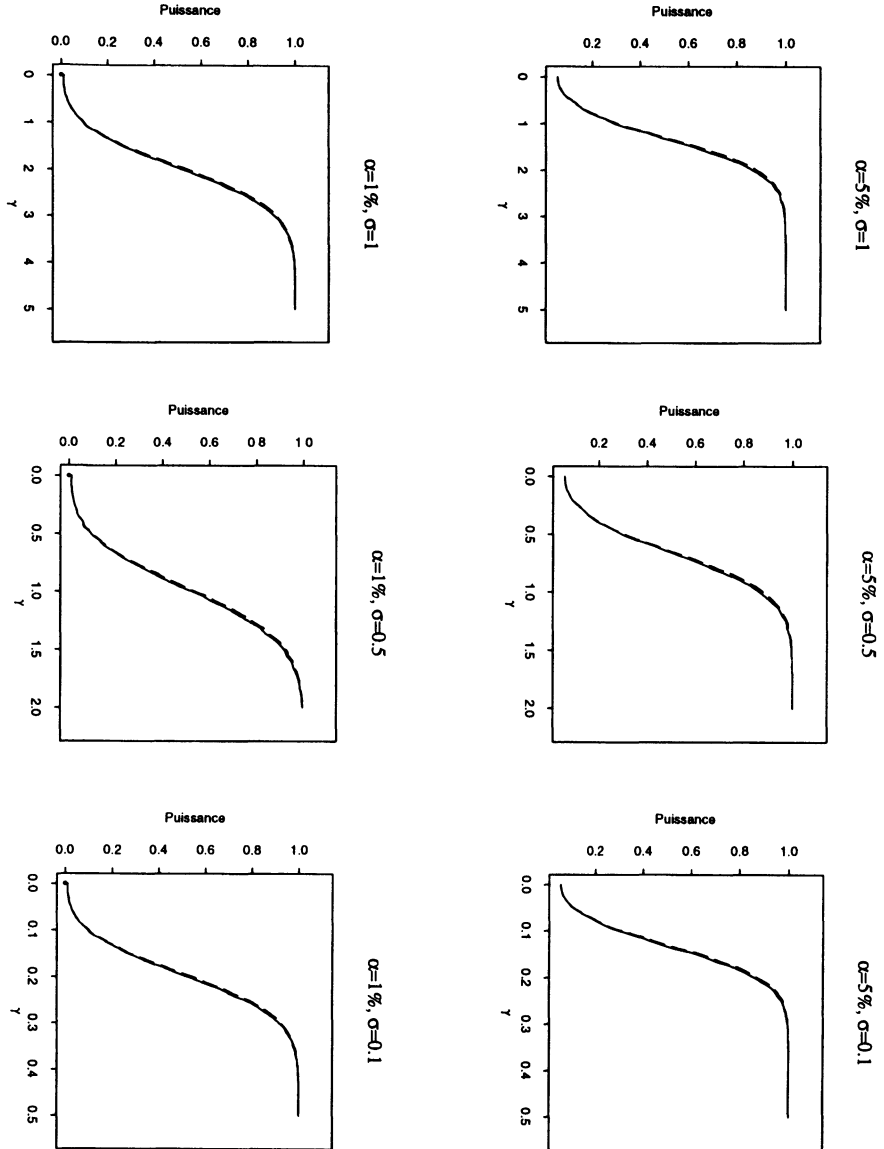


FIGURE 9 :

Puissance observée des procédures classique (ligne pointillée) et de Tukey (ligne continue) lors de 10 000 tests confrontant les hypothèses $H_0 : w = (1/5, 1/5, 1/5, 1/5, 1/5)'$ et $H_1 : w = w(\gamma)$ aux seuils de 5% et 1%, lorsque les perturbations pseudo-aléatoires affectant les comparaisons par paire suivent une loi normale d'espérance nulle et de variance $\sigma^2 = 1, 1/4$ et $1/100$.

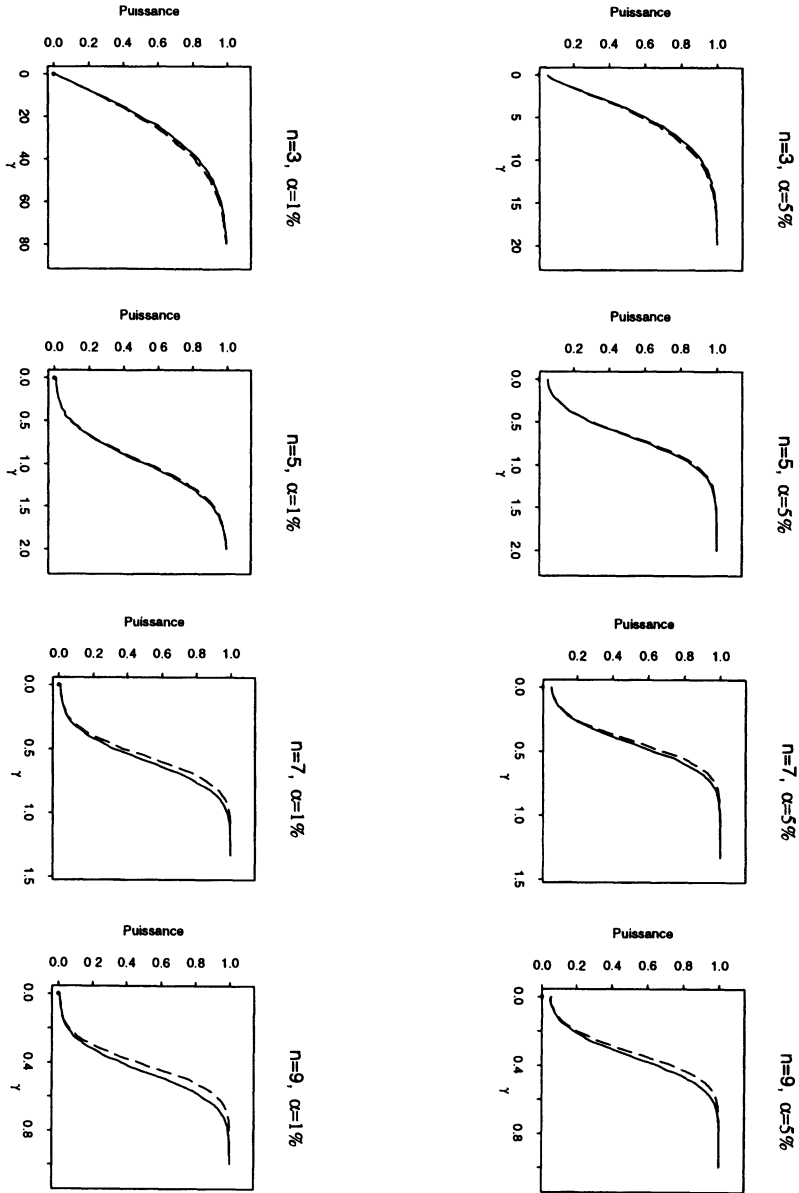


FIGURE 10 :

Puissance observée des procédures classique (ligne pointillée) et de Tukey (ligne continue) lors de 10 000 tests confrontant les hypothèses $H_0 : w = e/n$ et $H_1 : w = w(\gamma)$ aux seuils de 5% et 1%, lorsque $n = 3, 5, 7$ et 9 et que les perturbations pseudo-aléatoires affectant les comparaisons par paire suivent une loi normale d'espérance nulle et de variance $\sigma^2 = 1/4$.

Les figures 11 et 12 sont le résultat de 10 000 tests de l'hypothèse $H_0 : w = e/n$ effectués au seuil de 5% et de 1% quand $n = 5$. Chaque graphique illustre l'évolution de la puissance des deux tests concurrents en fonction de $\gamma \geq 0$, le paramètre définissant les contre-hypothèses contiguës (7).

La figure 11 met en parallèle les situations où la loi de e_{ij} est normale, uniforme ou double exponentielle. La figure 12 permet de faire des comparaisons avec différents mélanges de lois normales. Comme on peut le constater, le comportement des deux tests est le même dans tous les cas. Il semble donc n'y avoir que peu de conséquence à supposer la normalité lorsque la loi des perturbations est en réalité uniforme, double exponentielle ou un mélange de lois normales. La généralité de cette conclusion mériterait toutefois d'être testée à l'aide d'autres séries de contre-hypothèses. Comme on le signale dans le §2.10 de l'ouvrage de Miller (1981), bien peu de choses sont connues à ce jour sur la robustesse des procédures statistiques définies à partir de l'étendue studentisée.

6. Conclusion

Tous les résultats théoriques et expérimentaux rapportés dans cet article portent à croire qu'il n'existe pas de différence majeure entre la statistique de Tukey et celle du rapport de vraisemblance, vus comme critères d'un test de confirmation des priorités dans le cadre du procédé d'analyse hiérarchique de Saaty. L'une et l'autre procédure sont sans biais et de puissance comparable pour de petites valeurs de n sous les postulats d'indépendance mutuelle et de normalité des perturbations aléatoires affectant les comparaisons par paire effectuées par le répondant. Par ailleurs, les deux statistiques semblent se comporter adéquatement même lorsque ledit postulat de normalité s'avère non fondé.

Il serait intéressant, dans des travaux ultérieurs, d'examiner la sensibilité relative de ces deux tests en cas de dépendance ou d'hétérogénéité de la variance des termes d'erreur e_{ij} du modèle linéaire (2). La structure de covariance proposée par de Jong (1984) serait un point de départ naturel pour une telle étude. Un projet plus ambitieux consisterait à identifier le type de contre-hypothèses pour lequel le test de Tukey s'avère moins puissant que son compétiteur. Bien peu de progrès semblent avoir été réalisés dans ce domaine depuis les travaux de David (1953).

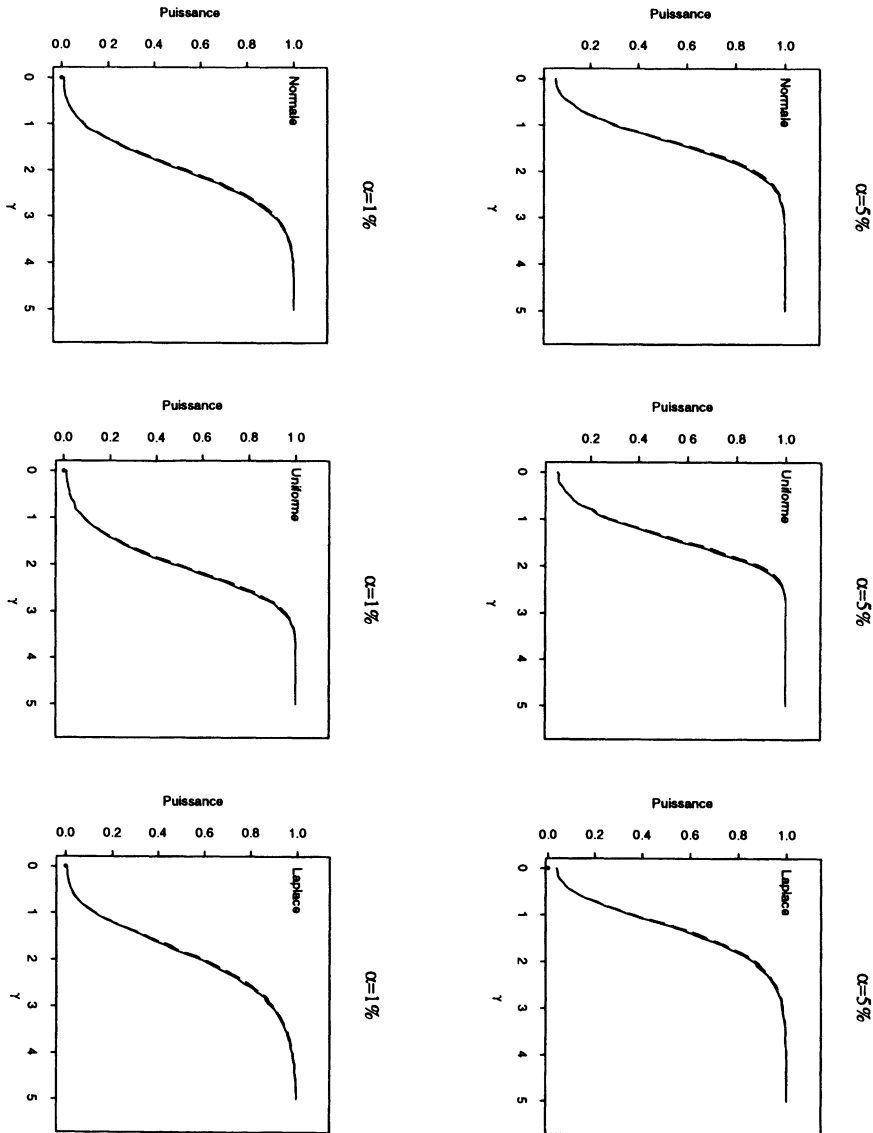


FIGURE 11 :

Puissance observée des procédures classique (ligne pointillée) et de Tukey (ligne continue) lors de 10 000 tests confrontant les hypothèses $H_0 : w = (1/5, 1/5, 1/5, 1/5, 1/5)'$ et $H_1 : w = w(\gamma)$ aux seuils de 5% et 1%, lorsque les perturbations pseudo-aléatoires affectant les comparaisons par paire suivent (a) une loi normale centrée réduite; (b) une loi uniforme sur l'intervalle $(-\sqrt{3}, \sqrt{3})$; (c) une loi double exponentielle de variance 1.

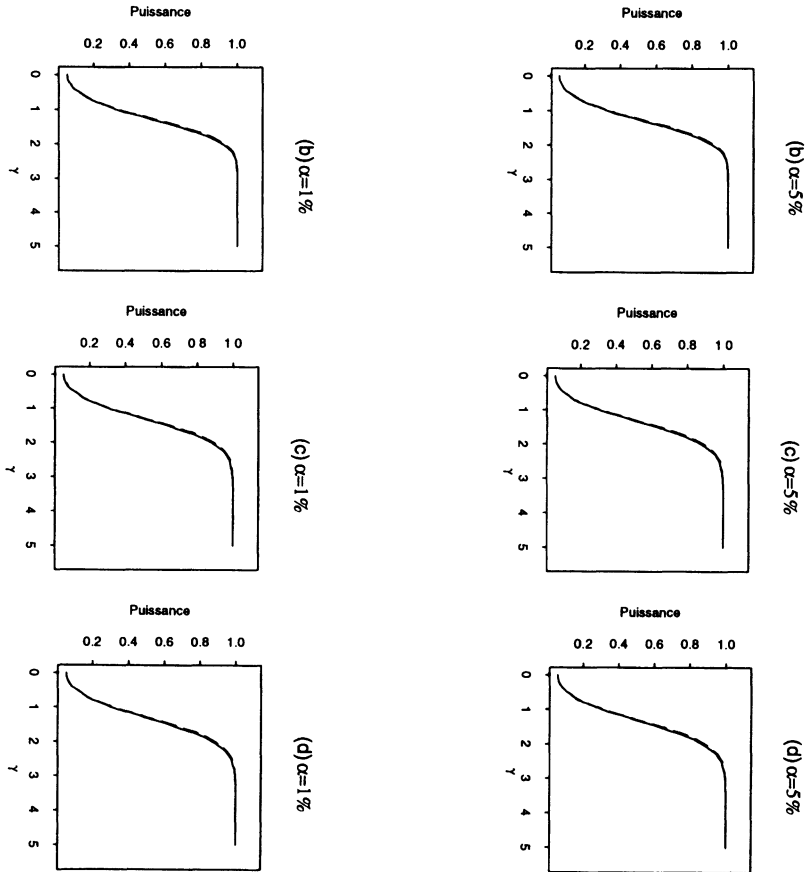


FIGURE 12 :

Puissance observée des procédures classique (ligne pointillée) et de Tukey (ligne continue) lors de 10 000 tests confrontant les hypothèses $H_0 : w = (1/5, 1/5, 1/5, 1/5, 1/5)'$ et $H_1 : w = w(\gamma)$ aux seuils de 5% et 1%, lorsque les perturbations pseudo-aléatoires affectant les comparaisons par paire ont pour distributions un mélange de lois normales de paramètres (a) $\tau^2 = 1/2$ et $\pi = 1/2$; (b) $\tau^2 = 1/2$ et $\pi = 1/10$; (c) $\tau^2 = 9/10$ et $\pi = 1/10$; (d) $\tau^2 = 99/100$ et $\pi = 1/2$.

Bibliographie

- DAVID, H. A. (1953). The power function of some tests based on range. *Biometrika*, **40**, 347-353.
- DAVID, H. A. (1981). *Order Statistics*, deuxième édition. John Wiley & Sons, New York.
- DE JONG, P. (1984). A statistical approach to Saaty's scaling method for priorities. *Journal of Mathematical Psychology*, **28**, 467-478.
- GENEST, C., F. LAPOINTE & S. W. Drury (1993). On a proposal of Jensen for the analysis of ordinal pairwise preferences using Saaty's eigenvector scaling method. *Journal of Mathematical Psychology*, **37**, 575-610.
- GENEST, C. & S.-S. Zhang (1995). Hilbert's metric and the Analytic Hierarchy Process. *Mathematical and Computer Modelling*, sous presse.
- GUPTA, S. S., K. C. S. PILLAI & G. P. Steck (1964). On the distribution of linear functions and ratios of linear functions of ordered correlated normal random variables with emphasis on range. *Biometrika*, **51**, 143-151.
- JOHNSON, N. L. & S. KOTZ (1970). *Continuous Univariate Distributions-2*. John Wiley & Sons, New York.
- MILLER, R. G. Jr (1981). *Simultaneous Statistical Inference*, deuxième édition. Springer-Verlag, New York.
- RAMACHANDRAN, K. V. (1956). On the Tukey test for the equality of means and the Hartley test for the equality of variances. *The Annals of Mathematical Statistics*, **27**, 825-831.
- RAO, C. R. (1973). *Linear Statistical Inference and its Applications*, deuxième édition. John Wiley & Sons, New York.
- SAATY, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, **15**, 234-281.
- SAATY, T. L. (1993). What is relative measurement? The ratio scale phantom. *Mathematical and Computer Modelling*, **17**, 1-12.
- SAATY, T. L. (1994). *Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process*. RWS Publications, Pittsburgh.