

J. P. BENZÉCRI

**Sur le calcul des taux d'inertie dans
l'analyse d'un questionnaire, addendum
et erratum à [BIN. MULT.]**

Les cahiers de l'analyse des données, tome 4, n° 3 (1979),
p. 377-378

http://www.numdam.org/item?id=CAD_1979__4_3_377_0

© Les cahiers de l'analyse des données, Dunod, 1979, tous droits réservés.

L'accès aux archives de la revue « Les cahiers de l'analyse des données » 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/>

SUR LE CALCUL DES TAUX D'INERTIE DANS L'ANALYSE D'UN QUESTIONNAIRE, ADDENDUM ET ERRATUM À [BIN. MULT.] [TAUX QUEST.]

par J. P. Benzécri ⁽¹⁾

Partons de la référence aux correspondances continues : Supposons que, comme il est souvent le cas même pour des questions proprement dites, une variable continue x soit sous-jacente à l'ensemble des modalités J_q . Quand on subdivise indéfiniment l'intervalle de variation de x , les facteurs tendent à définir une fonction $\varphi_\alpha(x)$ sur cet intervalle ; $\varphi_\alpha(x)$ étant la limite de $\varphi_\alpha(j)$ pour la modalité j qui comprend x , limite qui existe dans le cadre de modèles satisfaisant à des hypothèses convenables.

Cependant si l'on calcule les pourcentages d'inertie sur le tableau k_{IJ} , ces pourcentages tendent nécessairement vers zéro, même pour le 1^{er} facteur. Ceci est absurde ; car les facteurs introduits par la subdivision des modalités (bien que causant la croissance indéfinie de la trace) n'ont pas d'intérêt ; (cf *Cahiers* Vol. II n° 1 § 2. C. 1). D'autre part contrairement à ce qui a été publié les taux τ'_α (calculés pour le tableau de Burt k'_{JJ}) tendent eux aussi vers zéro bien qu'ils soient pour les premiers facteurs, supérieurs aux τ_α . Précisons donc ce qu'il convient de substituer aux calculs usuels de taux.

Le tableau k_{IJ} , donc aussi le tableau de Burt k'_{JJ} , produit en général Card J - Card Q facteurs non triviaux ; (lesquels, engendrent le sous-espace de R^J formé des fonctions qui ont moyenne nulle sur chacun des J_q ; tandis que les fonctions constantes sur chacun des J_q donnent d'une part le facteur trivial constant sur J relatif à la v. p. 1 ; et d'autre part Card Q - 1 facteurs triviaux relatifs à la v. p. 0). La trace étant (Card J - Card Q)/Card Q, la valeur moyenne d'une valeur propre λ_α non triviale est (1/Card Q) ; quant aux λ'_α , qui sont les

(1) Professeur de statistique à l'université P. & M. Curie

carrés des λ_α , leur valeur moyenne est supérieure à $(1/\text{Card } Q)^2$ (en effet la moyenne du carré est égale au carré de la moyenne, augmenté de la variance). La trace tr' est donc supérieure à $(\text{Card } J - \text{Card } Q) / (\text{Card } Q)^2$; et tend vers l'infini avec $\text{Card } J$ si $\text{Card } Q$ reste fixé.

Cependant la valeur typique $\lambda_\alpha = 1/\text{Card } Q$ apparaît dans deux cas.

D'une part dans l'analyse du tableau λ_{SJ} qui correspond à un ensemble de sujets donnant une fois et une seule chaque système de réponses : dans ce cas trivial, tous les facteurs sont relatifs à la v. p. $\lambda = 1/\text{Card } Q$. D'autre part dans le cas où $\text{Card } Q = 2$: l'analyse de k_{IJ} fournit sur $J = J_{q1} \cup J_{q2}$ les mêmes facteurs que l'analyse du tableau rectangulaire $J_{q1} \times J_{q2}$ (tableau de correspondance binaire usuel) mais avec des v. p. qui sont (si on note ρ la v. p. relative à $J_{q1} \times J_{q2}$) :

$$\lambda = (1 \pm \rho^{1/2})/2 \quad ; \quad \text{d'où} : \lambda' = (1 \pm \rho^{1/2})^2/4 \quad ;$$

$$\rho = 4(\lambda - 0,5)^2 = 4(\lambda'^{1/2} - 0,5)^2 \quad ;$$

Dans ce cas, seuls comptent les facteurs relatifs à $\lambda = (1 + \rho^{1/2})/2$; (les autres étant des artéfacts) et il faut les rapporter à la v. p. ρ , dans les calculs de pourcentage.

Dans le cas général d'un questionnaire à $\text{Card } Q$ questions, on se restreindra donc également dans les calculs de pourcentage aux facteurs relatifs à une v. p. supérieure à $1/\text{Card } Q$; de plus, on attribuera à chacun de ces facteurs l'inertie ρ :

$$\rho(\lambda) = (\text{Card } Q / (\text{Card } Q - 1))^2 (\lambda - (1/\text{Card } Q))^2$$

cette quantité varie de 1 à 0 quand λ varie de 1 à $(1/\text{Card } Q)$. Dans le cas où $\text{Card } Q$ est très élevé la valeur ρ ainsi calculée diffère peu de $\lambda' = \lambda^2$; mais en général elle donnera une idée plus juste de l'importance relative des facteurs. Les taux modifiés seront donc :

$$\tau m_\alpha = \rho(\lambda_\alpha) / \sum \{ \rho(\lambda_{\alpha'}) \mid \alpha' \in Z \quad ; \quad \lambda_{\alpha'} > (1/\text{Card } Q) \}$$

la somme au dénominateur étant étendue à l'ensemble Z des valeurs propres supérieures à $1/\text{Card } Q$.

Remarque : Si on analyse le tableau obtenu en mettant à zéro les blocs diagonaux du tableau de Burt, on a les mêmes facteurs avec pour valeurs propres justement les $(\lambda - (1/\text{Card } Q))^2$; de plus seuls sont directs les facteurs pour lesquels λ dépasse $(1/\text{Card } Q)$. ceci permet d'interpréter les calculs ci-dessus en terme d'inertie relative aux facteurs directs d'un tableau de Burt modifié.