

J.-P. BENZÉCRI

Pondérations relatives des questions dans un tableau sous forme disjonctive complète

Les cahiers de l'analyse des données, tome 5, n° 2 (1980),
p. 192

http://www.numdam.org/item?id=CAD_1980__5_2_192_0

© Les cahiers de l'analyse des données, Dunod, 1980, 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/>

PONDÉRATIONS RELATIVES DES QUESTIONS DANS UN TABLEAU SOUS FORME DISJONCTIVE COMPLÈTE

par J.-P. Benzécri (1)
 d'après le rapport de stage de D.E.A. de L. Benyamina

Reprenons les notations classiques : I, ensemble des individus ; Q, ensemble des questions ; J_q ensemble des modalités afférentes à la question q de Q ; $J = \cup \{J_q | q \in Q\}$. Alors la somme des contributions absolues à la trace des modalités de la question q est donnée par la formule (cf e. g. ENS 2, la prat. de l'A. des D. ; VI n° 2 pp 322 sqq) :

$$\text{Somme}(q) = \sum \{f_j \| f_I^j - f_I \|^2 | j \in J_q\} = (\text{Card } J_q - 1) / \text{Card } Q.$$

D'autre part, il est commun d'ajuster à des valeurs fixées *a priori*, les parts relatives de la trace afférentes à des blocs homogènes, en multipliant chacun de ces blocs par un coefficient approprié ; la suite des coefficients étant déterminée par un calcul itératif (cf *Cahiers* Vol II n° 3, pp 333 sqq, 1977 ; ou ENS 2, V n° 5, pp 271 sqq). Or dans le cas d'un questionnaire et des blocs J_q , le calcul itératif est inutile, les coefficients peuvent être déterminés directement. Voici comment.

Soit m_q une suite de coefficients positifs, notons $k'(i, j)$ (et de même f' pour les fréquences, fréquences relatives etc.) le tableau obtenu à partir du tableau $k(i, j)$ (en 0-1) en multipliant par m_q le bloc J_q des colonnes afférentes à la question q (i.e. si le sujet i adopte la modalité j de réponse à q, on a $k'(i, j) = m_q$). Il est clair que cette transformation ne modifie pas les profils des colonnes : $f'_I = f_I$; $f'_I^j = f_I^j$; seuls sont modifiés les poids relatifs de celles-ci, i.e. les f_j ; de façon précise, si on note m_{tot} la somme des m_q on a :

$$\forall j \in J_q : f'_j = (m_q / m_{\text{tot}}) \text{Card } Q f_j ; \text{ d'où pour les contributions :}$$

$$\text{Somme}'(q) = \sum \{f'_j \| f_I^j - f_I \|^2 | j \in J_q\} = (\text{Card } J_q - 1) m_q / m_{\text{tot}}.$$

Soit donc les A_q une suite de nombres positifs auxquels on demande que les $\text{Somme}'(q)$ soient proportionnels ; et soit $A_{\text{tot}} = \sum_q A_q$. Il suffira de poser :

$$m_q = A_q / (\text{Card } J_q - 1) ; \text{ d'où } \text{Somme}'(q) = A_q / m_{\text{tot}}.$$

En particulier pour que toutes les questions apportent une même contribution à la trace (cas où tous les A_q valent 1), il suffira de diviser chaque bloc J_q par $(\text{Card } J_q - 1)$; et on aura pour trace' : $A_{\text{tot}} / m_{\text{tot}} = \text{Card } Q / \sum \{1 / (\text{Card } J_q - 1)\}$; formule qu'on rapprochera de celle relative au tableau k en (0,1) : $\text{trace} = (\text{Card } J - \text{Card } Q) / \text{Card } Q$.

(1) Professeur de statistique. Université P. et M. Curie