MATHÉMATIQUES ET SCIENCES HUMAINES

JEAN-MARC BERNARD

Une formule utile en analyse des comparaisons

Mathématiques et sciences humaines, tome 138 (1997), p. 75-76

<http://www.numdam.org/item?id=MSH_1997__138__75_0>

Math. Inf. Sci. hum., (35^{e} année, n°138, *1997*, pp.75- 76)

UNE FORMULE UTILE EN ANALYSE DES COMPARAISONS

Jean-Marc BERNARD[1]

RÉSUMÉ — *Cette note donne la formule du protocole projeté (orthogonalement)* y^J *d'un protocole de base* x^J *de* R^J*, pondéré par* n_J*, sur le sous-espace de* R^J *dual d'une g-comparaison de* R_J.

SUMMARY — A useful formula for the Analysis of Comparisons
This note gives the formula of the orthogonally-projected protocol y^J *of a basic protocol* x^J *from* R^J*, weighted by* n_J*, on the dual sub-space in* R^J *of a g-comparison from* R_J.

1. INTRODUCTION

Dans l'article de base sur les structures linéaires en jeu en Analyse des Comparaisons (Rouanet, Lépine, 1976), on trouve (p. 15) une formule donnant la somme des carrés associée à une comparaison quelconque sur un protocole pondéré. Par contre, n'y figure pas de formule explicite — qu'on peut pourtant considérer comme logiquement antérieure — du protocole dérivé (par projection orthogonale) correspondant, dont la somme des carrés constitue seulement une statistique particulière. Il faut certainement voir cette absence comme la conséquence de ce qui, à l'époque, était considéré comme prioritaire par ces auteurs, à savoir la recherche d'algorithmes informatisables (donc, à l'époque, peu gourmands en mémoire) et performants pour une refonte profonde des méthodes usuelles d'ANOVA. Dans ce contexte, l'accent avait ainsi surtout été mis sur l'aspect inductif de ces méthodes avec pour but principal le calcul des statistiques usuelles (SC, DL, CM et F), plutôt que, comme plus récemment (*e.g.* Bernard, 1994, ch. II et III), sur leur aspect descriptif qui conduit à donner une place plus importante à la notion de protocole dérivé.

2. LA FORMULE

Nous reprenons les notations de Rouanet & Lépine (1976). On considère l'espace R^J des protocoles sur J muni de la mesure fondamentale n_J et son espace dual R_J des mesures sur J. Soit un protocole x^J de R^J, et une comparaison (sous-espace de contrastes sur J) $\mathcal{C}$ de R_J à P degrés de liberté (ddl) ; on note $\mathcal{C}^*$ le sous-espace dual de $\mathcal{C}$ dans R^J. En

[1] Laboratoire Cognition et Activité Finalisées, CNRS ERX 125, Université Paris 8. 2 rue de la Liberté, 95526 Saint-Denis cedex 02.

termes matriciels, on note $\boldsymbol{X}$ la matrice $[J, 1]$ des x^j, $\boldsymbol{N}$ la matrice diagonale $[J, J]$ des n_j, et $\boldsymbol{C}$ la matrice $[J, P]$ constituée de P vecteurs c_J^p de R_J formant une base de $\mathcal{C}$.

Le protocole y^J, projection orthogonale de x^J sur $\mathcal{C}^*$, est donné par l'expression matricielle suivante ($\boldsymbol{Y}$ est la matrice $[J, 1]$ des y^j) :

$$\boldsymbol{Y} = \boldsymbol{N}^{-1}\boldsymbol{C}(\boldsymbol{C}'\boldsymbol{N}^{-1}\boldsymbol{C})^{-1}\boldsymbol{C}'\boldsymbol{X} \tag{1}$$

La somme des carrés (bruts, *i.e.* non centrés) associée à un protocole quelconque $\boldsymbol{Z}$ de R^J est donnée par $SC_{\boldsymbol{Z}} = \boldsymbol{Z}'\boldsymbol{N}\boldsymbol{Z}$. Si on applique cette formule à $\boldsymbol{Y}$ donné en (1), et en utilisant le fait que les matrices $(\boldsymbol{C}'\boldsymbol{N}^{-1}\boldsymbol{C})^{-1}$ ainsi que $\boldsymbol{N}^{-1}$ sont symétriques, on trouve

$$SC_{\boldsymbol{Y}} = \boldsymbol{X}'\boldsymbol{C}(\boldsymbol{C}'\boldsymbol{N}^{-1}\boldsymbol{C})^{-1}\boldsymbol{C}'\boldsymbol{N}^{-1} \quad \boldsymbol{N} \quad \boldsymbol{N}^{-1}\boldsymbol{C}(\boldsymbol{C}'\boldsymbol{N}^{-1}\boldsymbol{C})^{-1}\boldsymbol{C}'\boldsymbol{X}\,,$$

expression qui, après simplification, redonne la formule générale de la somme des carrés associée à une comparaison (Rouanet, Lépine, 1976, p. 15) :

$$SC_{\boldsymbol{Y}} = \boldsymbol{X}'\boldsymbol{C}(\boldsymbol{C}'\boldsymbol{N}^{-1}\boldsymbol{C})^{-1}\boldsymbol{C}'\boldsymbol{X}\,. \tag{2}$$

Enfin, si on applique (1) au cas particulier d'une comparaison à $P = 1$ ddl, représentée par le contraste c_J, on aboutit à l'expression plus compacte du protocole projeté $\boldsymbol{Y}$, déjà indiquée (avec d'autres notations) dans Bernard (1994, ch. V, eq. (7), p. 65) :

$$y^j = \frac{c_j}{n_j}\left(\frac{\sum c_j x^j}{\sum c_j^2/n_j}\right) \tag{3}$$

3. REMARQUE : GÉNÉRALITÉ DE LA FORMULE (1)

Dans Bernard (1994) nous avons développé l'idée que la construction élaborée par Rouanet et Lépine autour de la notion de *comparaison* était en fait générale et pouvait s'appliquer, non seulement à des mesures de masse totale 0 de R_J (contrastes, puis espaces de contrastes), mais aussi à des mesures de masse totale non nulle. Nous avons ainsi parlé de *g-comparaison* pour désigner un sous-espace quelconque de R_J. Or, aucune des formules précédentes ne présuppose que $\mathcal{C}$ est une comparaison ; ces formules sont donc également valables plus généralement pour une *g-comparaison*[2].

BIBLIOGRAPHIE

BERNARD, J.-M. (1994), "Analyse descriptive des données planifiées. II : Protocoles pondérés, protocoles dérivés, III : Protocoles dérivés pour certaines structures remarquables, et V : Introduction à la notion de comparaison", *Mathématiques, Informatique et Sciences humaines*, 126, 19–26, 27–54 et 61–70.

ROUANET, H., LÉPINE, D. (1976), "Structures linéaires et analyse des comparaisons", *Mathématiques et Sciences humaines*, 56, 5–46.

[2]On vérifiera par exemple aisément qu'en choisissant pour $\mathcal{C}$ la mesure n_J, g-comparaison à 1 ddl qui exprime la "question" de la "moyenne pondérée", on obtient par (1) un protocole projeté y^J constant, avec pour unique valeur la moyenne pondérée de $(x^J,\ n_J)$.