

REVUE DE STATISTIQUE APPLIQUÉE

A.-B. DUFOUR

M. PAGES

Modèles de taille et de forme : que choisir ?

Revue de statistique appliquée, tome 42, n° 1 (1994), p. 5-18

http://www.numdam.org/item?id=RSA_1994__42_1_5_0

© Société française de statistique, 1994, 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/>

MODÈLES DE TAILLE ET DE FORME : QUE CHOISIR ?

A.-B. Dufour, M. Pages

*Laboratoire d'Analyse de Données et Biométrie
Centre de Recherche et d'Innovation en Sport
U.F.R.A.P.S., Université Claude Bernard - Lyon 1
27-29, Boulevard du 11 novembre 1918, 69622 Villeurbanne Cedex*

RÉSUMÉ

L'étude de la taille et de la forme représente une part importante de la recherche en morphologie. Un modèle générique est proposé concernant un profil commun à tous les individus relativement aux différentes variables. Selon celui-ci, les individus sont tous semblables, à un facteur de taille près. La forme met en évidence les écarts entre les individus et le modèle. Elle est étudiée de deux façons différentes : soit sur les données brutes, soit sur les profils. Le choix entre les deux méthodes dépend essentiellement de la problématique : certaines définitions de la forme mettent plutôt en évidence les individus qui s'écartent du profil moyen pour les petites variables, d'autres plutôt l'existence de groupes caractéristiques.

Mots-clés : *Morphologie, Fonction taille, Vecteur forme, A.C.P. non centrée.*

SUMMARY

Size and shape are very important for analyzing morphology. For all individuals, a common profile based on variables is proposed. According to this profile, the individuals looklike each other : they are represented by multivariate proportional vectors. Shape shows differences between the subjects and the model. Two different proceedings are developped : one based on original data and another one based on profiles. Their use is linked with the problem. Some of them are well used for showing differences between the subjects and the common profile for small measures. Others are better for showing the existence of clusters of individuals.

Key-words : *Morphology, Size function, Shape vector, Non centred P.C.A.*

1. Introduction

Les «facteurs» de taille et de forme sont souvent évoqués dans l'interprétation des résultats d'une analyse de données. Or ils ne semblent pas avoir reçu de définition précise et diffèrent selon l'auteur, le domaine d'application, les questions

qui se posent. Le texte qui suit tente de clarifier la situation à travers un exemple anthropométrique mais il est bien évident que ce choix n'est pas exclusif. Ce qui va être dit est transcriptible à d'autres champs d'application comme par exemple celui de la croissance. Quand un enfant grandit, il croît en taille mais il change également de forme de façon plus ou moins désordonnée (Healy et Tanner, 1981).

C. Burt (1946) souligne que «si chaque corps humain conservait exactement les mêmes proportions, les différences entre un corps et un autre consisteraient seulement en différence de taille». Cette situation modèle correspond au cas où tous les individus ont un même profil : ils sont alors identiques à un individu type à un facteur d'agrandissement près (la taille).

Or il existe des divergences par rapport à cet individu modèle qui peuvent être dues tout simplement à une variabilité individuelle ou résulter de l'influence de facteurs écologiques, climatiques, etc... (Collectif, 1986). Elles induisent ainsi différents types physiques (différentes formes). Ce sont ces divergences mêmes qui renseignent sur des questions telles que : sur quels critères regrouper les individus ? quels sont ceux qui ont des aptitudes particulières ? quelles techniques de réadaptation adopter ?

Dans l'orientation d'enfants vers la pratique d'un sport, la morphologie est par exemple un facteur important du choix. Si la stature est naturellement utilisée (Pilardeau et al, 1978), certaines particularités physiques sont parfois plus spécifiques d'une activité que d'une autre (Lot, 1981). La forme est souvent masquée par ce facteur dominant qu'est la taille. Son étude est plus fine si elle est conçue comme un résidu, l'information parasite de la taille enlevée.

1. Définitions et Notations

Les données sont constituées d'un ensemble de valeurs observées réelles x_{ij} strictement positives et homogènes. Elles se présentent sous la forme d'un tableau dont les q colonnes sont associées à des variables quantitatives et les n lignes aux individus.

On définit :

– le total marginal de la i -ème ligne du tableau : $x_{i.} = \sum_{j=1}^q x_{ij}$

– le total marginal de la j -ème colonne du tableau : $x_{.j} = \sum_{i=1}^n x_{ij}$

– le total des x_{ij} noté $x_{..}$

$p_{i.} = \frac{x_{i.}}{x_{..}}$ et $p_{.j} = \frac{x_{.j}}{x_{..}}$ sont les distributions marginales.

On considère le tableau X de terme général $\frac{x_{ij}}{\sqrt{p_{.j}}}$. Soit \mathbb{R}^q l'espace euclidien des individus, d'origine O, muni de la base canonique $(\vec{e}_1 \dots \vec{e}_q)$ et du produit scalaire «classique» (correspondant à la métrique identité).

L'ensemble des n individus est représenté dans \mathbb{R}^q par un nuage de points X_i dont les coordonnées figurent dans la i -ème ligne du tableau X. On affecte chacun

des individus d'un même poids $\frac{1}{n}$. Le centre de gravité H de ce nuage a pour j -ème coordonnée :

$$\frac{x_{..}}{n} \sqrt{p_{.j}}$$

Remarque : La transformation $\frac{x_{ij}}{\sqrt{p_{.j}}}$ a pour but de donner des importances comparables aux différentes variables; d'autres auraient pu être envisagées ($\frac{x_{ij}}{p_{.j}}$, $\frac{x_{ij}}{\sigma_j}$ où σ_j est l'écart-type de x_j , $\frac{x_{ij}}{\sqrt{\sum_i x_{ij}^2}}$, par exemple). Celle-ci a été retenue car elle correspond (paragraphe 2) à un résultat simple dans le calcul de la projection des points X_i sur \overrightarrow{OH} .

2. Modélisation de l'ensemble des individus. Facteur taille

Les n individus étant décrits relativement aux q variables par les n vecteurs de \mathbb{R}^q $\overrightarrow{X}_1, \overrightarrow{X}_2, \dots, \overrightarrow{X}_n$, on peut faire l'hypothèse modélisatrice qu'ils ne diffèrent entre eux que par un coefficient de proportionnalité c'est-à-dire qu'ils sont représentés par des vecteurs $\overrightarrow{X}_1, \dots, \overrightarrow{X}_2, \dots, \overrightarrow{X}_n$ colinéaires. Cela revient à dire qu'il existe un **profil modèle commun** \overrightarrow{S} , vecteur unitaire, combinaison linéaire des q variables initiales et rassemblant toute l'information de ces q variables pour les individus modélisés. On obtient les points \widehat{X}_i en projetant orthogonalement les points X_i sur l'axe portant \overrightarrow{S} . La forme linéaire de matrice ${}^t s = (s_1 \dots s_j \dots s_q)$ permet d'associer à un vecteur \overrightarrow{X}_i (c'est-à-dire à un individu ou à une ligne du tableau X) la mesure algébrique de sa projection orthogonale $\overrightarrow{\widehat{X}}_i$ sur l'axe portant \overrightarrow{S} . Par définition, nous appelons cette forme linéaire : facteur taille.

La matrice $t = \mathbf{X}s$ a pour éléments les tailles t_i des n individus. Le modèle correspond à l'hypothèse d'une similitude géométrique puisque les vecteurs \overrightarrow{X}_i sont tous colinéaires entre eux et colinéaires à \overrightarrow{S} : $\overrightarrow{X}_i = t_i \overrightarrow{S}$. C'est la propriété requise par Mosimann (1970) pour une «fonction taille» T réelle positive : quels que soient \overrightarrow{X}_i et le nombre réel α positif, $T(\alpha \overrightarrow{X}_i) = \alpha T(\overrightarrow{X}_i)$.

Analytiquement le tableau modélisé de dimensions $(n \times q)$ s'écrit :

$$\widehat{\mathbf{X}} = \mathbf{t} {}^t s = \mathbf{X} s {}^t s.$$

Ses lignes sont les matrices des composantes des vecteurs \overrightarrow{X}_i ; puisqu'elles correspondent à des vecteurs colinéaires, elles sont proportionnelles. Tous les individus modélisés sont des copies les uns des autres à un facteur d'agrandissement près : $\widehat{\mathbf{X}}$ est de rang 1. Le modèle est qualifié de multiplicatif : le terme $i \times j$ du tableau $\widehat{\mathbf{X}}$

est $t_i s_j$, produit d'un coefficient lié à l'individu (t_i , «taille» individuelle), et d'un coefficient lié à la variable (s_j). Cette factorisation de \widehat{X} à l'aide du produit $t^t s$ n'est pas unique puisqu'elle dépend du choix de \vec{S} . Utilisée en vue d'une représentation graphique dans un plan, elle peut privilégier davantage la conservation des relations entre individus ou entre variables (Gabriel, 1971).

Dès les premières applications des méthodes multivariées à la morphométrie, on a su définir le facteur taille comme un effet de la croissance générale de toutes les variables (Cohen, 1940; Pearce et Holland, 1961). On trouve dans la littérature de multiples définitions du facteur taille (Mosiman, 1970; Rao, 1964) dont la plupart découlent du modèle précédemment exposé et varient selon le critère – explicité ou non – utilisé pour choisir \vec{S} . Trois critères sont présentés ici en raison de leur importance dans les applications.

1^{er} Critère

L'axe portant \vec{S} vérifie le critère des moindres carrés. Il est porteur de l'inertie projetée maximum c'est-à-dire qu'il minimise $\sum_{i=1}^n \frac{1}{n} \left\| \vec{X}_i - \widehat{\vec{X}}_i \right\|^2$, écart quadratique résiduel moyen. On sait alors que le vecteur \vec{S} est le premier vecteur propre \vec{u} de l'analyse en composantes principales du tableau \mathbf{X} non centré. La matrice des tailles est donnée par la première composante principale : $\mathbf{v} = \mathbf{X}\mathbf{u} = {}^t(v_1, \dots, v_i, \dots, v_n)$. Le tableau modélisé est le tableau des données reconstitué par le premier facteur : $\widehat{\mathbf{X}} = \mathbf{X}\mathbf{u} {}^t\mathbf{u} = \mathbf{v} {}^t\mathbf{u}$. Il s'agit du premier terme de la décomposition en valeurs singulières du tableau \mathbf{X} (Eckart & Young, 1936; Good, 1969), appelée N -technique par Okamoto (1972). Cet auteur la compare à la R -technique obtenue en décomposant en valeurs singulières le tableau \mathbf{X} centré par colonnes, méthode utilisée pour l'étude de la taille par Jolicœur et Mosimann (1960). Mais dans ce cas, la propriété précédemment requise : «deux individus représentés par des vecteurs colinéaires (coefficient α) ont des tailles dans le même rapport α » n'est pas satisfaite. Cela signifie qu'on ne fait pas référence au même modèle : les points X_i sont alors tous modélisés par rapport au point H centre de gravité du nuage (le premier facteur de la décomposition en valeurs singulières de \mathbf{X} centré correspond alors à la taille).

2^{ème} critère

Alors que le premier critère découle d'une démarche mathématique (décomposition de \mathbf{X} en valeurs singulières), l'attitude adoptée pour définir le facteur taille est ici plus pragmatique.

On choisit pour \vec{S} le vecteur unitaire \vec{G} de l'axe joignant l'origine O au centre de gravité H du nuage de points. Le tableau \mathbf{X} ayant pour terme général $\frac{x_{ij}}{\sqrt{p_j}}$, la matrice ligne de ce vecteur est ${}^t\mathbf{g} = (\sqrt{p_1}, \dots, \sqrt{p_j}, \dots, \sqrt{p_q})$. La matrice des tailles individuelles est $\mathbf{X}\mathbf{g} = {}^t(x_1, \dots, x_i, \dots, x_n)$. Le tableau modélisé $\widehat{\mathbf{X}}$ a pour terme général $x_i \cdot \sqrt{p_j}$.

3^{ème} critère

Ce dernier critère correspond à une attitude naturelle dans les recherches anthropométriques : la construction d'indices morphologiques (Vandervael, 1980). On peut être amené à définir le profil commun à l'aide d'un des axes de la base canonique. Supposons que l'une des mesures soit intéressante pour définir la taille. Pour simplifier l'exposé, prenons par exemple la première variable. On est alors amené à choisir pour \vec{S} le premier vecteur de la base : \vec{e}_1 . Il en résulte alors $T(\vec{X}_i) = \frac{x_{i1}}{\sqrt{p.1}}$.

Le tableau modèle $\widehat{\mathbf{X}}$ est dans ce cas le tableau dont tous les éléments sont nuls sauf ceux de la première colonne dont les éléments sont $\frac{x_{i1}}{\sqrt{p.1}}$.

Remarques

En général, les profils communs \vec{u} et \vec{G} choisis dans le premier et le deuxième critère ne sont pas très différents (Gower, 1966). Cela vient de ce que la morphologie des individus obéit à des lois – linéaires en première approximation – c'est-à-dire que les proportions entre les différentes variables sont harmonieusement conservées : le pourcentage d'inertie imputable au premier facteur dans l'A.C.P. de \mathbf{X} non centré est grand. Sans que les individus soient strictement des copies les uns des autres à un facteur près, une grande partie de l'information leur est commune.

Mais s'il existe entre les différentes variables des relations d'allométrie qui induisent un gradient important des «plus petits» vers les «plus grands» individus, les profils \vec{u} communs et \vec{G} peuvent présenter des différences sensibles; ils correspondent à une allométrie appelée isométrie dans la terminologie courante (Jolicœur, 1991). Dans ce cas, certains auteurs utilisent le logarithme des variables; le deuxième critère conduit alors à la fonction taille : $T(\vec{X}_i) = \prod_{j=1}^q x_{ij}^{1/q}$.

3. Vecteur forme. Analyse multivariée des tableaux de résidus.

De façon un peu schématique, nous disons que la forme est «ce qui reste lorsqu'on a enlevé la taille». Deux méthodes sont proposées pour l'étude de la forme.

Equipollence géométrique

On peut proposer pour définition du vecteur forme de l'individu i : $\vec{X}_i - \vec{\widehat{X}}_i$. Les composantes des vecteurs forme des n individus figurent dans le tableau des résidus $\mathbf{X} - \widehat{\mathbf{X}}$, différence entre le tableau initial et le tableau modélisé. Si deux individus i et i' ont même forme, les deux lignes leur correspondant dans le tableau résidu sont les mêmes, les vecteurs forme sont équipollents (figure 1).

L'étude de la forme se fait à l'aide d'une A.C.P. de ce tableau $\mathbf{X} - \widehat{\mathbf{X}}$ non centré. Le retrait de $\widehat{\mathbf{X}}$ à \mathbf{X} correspond à un centrage multiplicatif (on a vu que le terme général de $\widehat{\mathbf{X}}$ est $t_i s_j$). Les facteurs successifs de cette A.C.P. sont utilisés en vue de la représentation plane des individus et sont qualifiés de facteurs forme. Notons

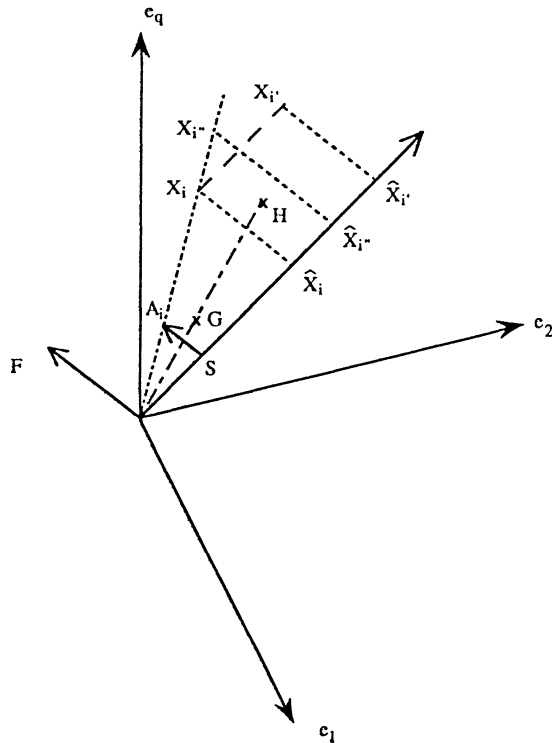


FIGURE 1

Equipollence géométrique : les deux individus i et i' ont même vecteur forme \vec{F}

Similitude géométrique : les deux individus i et i'' ont même vecteur forme $\overline{A_i S}$

toutefois que les facteurs taille et les facteurs forme ne sont pas incorréllés puisqu'ils ne résultent pas de la réalisation d'analyses en composantes principales centrées par variables. Dans tous les cas, l'espace des individus est décomposé en somme directe de deux sous-espaces orthogonaux : l'espace «modèle» de dimension 1 et l'espace «forme» de dimension au plus égale à $(q - 1)$.

Tableau analysé pour l'étude de la forme dans l'équipollence géométrique

critère	terme général	dimensions
1	$\frac{x_{ij}}{\sqrt{p_{.j}}} - u_j v_i$	$n \times q$
2	$\frac{x_{ij}}{\sqrt{p_{.j}}} - x_{i.} \sqrt{p_{.j}}$	$n \times q$
3	$\frac{x_{ij}}{\sqrt{p_{.j}}}$	$n \times (q - 1)$

Remarquons que dans le cas du premier critère, il est inutile de calculer le tableau $\mathbf{X} - \widehat{\mathbf{X}}$: le premier facteur de l'A.C.P. de \mathbf{X} est le facteur taille; les suivants décrivent la forme.

Similitude géométrique

Nous avons vu que les composantes du profil moyen sont celles du vecteur \vec{S} . Soit i un individu dont la taille t_i est obtenue par un des critères du paragraphe 2. Par définition :

- les composantes du vecteur $\frac{\vec{X}_i}{t_i}$ décrivent son **profil**
- son vecteur forme est $\frac{\vec{X}_i}{t_i} - \frac{\vec{X}_i}{t_i} = \frac{1}{t_i} \vec{X}_i - \vec{S}$.

Dans cette optique, si deux individus i et i'' de tailles t_i et t_i'' sont représentés dans \mathbb{R}^q par des vecteurs colinéaires $\vec{X}_i = \alpha \vec{X}_{i''}$, alors ils ont (Mosimann, 1970) :

- le rapport de leur taille égal à α : $t_i = \alpha t_{i''}$
- même vecteur forme.

La dynamique des mesures entre elles est alors prise en compte : Sempé *et al* (1979) parlent de la conservation d'une harmonie dans une croissance différentielle. On est conduit pour étudier la forme à réaliser l'A.C.P. du tableau non centré dont les lignes sont les composantes du vecteur $\frac{1}{t_i} \vec{X}_i - \vec{S}$. Comme dans le cas de l'équipollence, les facteurs successifs de cette A.C.P. sont qualifiés de facteurs forme. Géométriquement, on associe au profil de l'individu i le point A_i intersection du vecteur \vec{X}_i avec l'hyperplan des profils qui passe par l'extrémité du vecteur \vec{S} et lui est orthogonal (figure 1). Remarquons que le point G en général n'appartient pas à cet hyperplan (sauf dans le cas du critère 2) et que le centre de gravité du nuage des points A_i est différent des points S et G .

Tableau analysé pour l'étude de la forme dans la similitude géométrique

critère	terme général	dimensions
1	$\frac{x_{ij}}{v_i \sqrt{p \cdot j}} - u_j$	$n \times q$
2	$\frac{x_{ij}}{x_i \cdot \sqrt{p \cdot j}} - \sqrt{p \cdot j}$	$n \times q$
3	$\frac{\sqrt{p \cdot 1}}{x_{i1}} \frac{x_{ij}}{\sqrt{p \cdot j}}$	$n \times (q - 1)$

Remarque

L'A.C.P. réalisée dans le cadre du critère 2 correspond à l'analyse factorielle des correspondances du tableau des x_{ij} à condition que les pondérations des individus soient p_i au lieu de $\frac{1}{n}$, ce qui est discutable dans le cadre des données que nous nous proposons d'analyser. En effet, les tableaux traités par l'A.F.C. présentent en général une symétrie ligne-colonne car ils correspondent au croisement de deux tableaux (à n individus lignes) d'indicatrices de variables qualitatives; c'est de ce croisement que résulte le choix des pondérations p_i et p_j , les individus ayant eux, la pondération $\frac{1}{n}$. Dans le cadre de l'A.F.C., les points S et G sont confondus et coïncident avec le centre de gravité des points A_i puisqu'ils sont munis de la pondération p_i . En effet,

$$\sum_i p_i \frac{x_{ij}}{x_i \sqrt{p_j}} = \sqrt{p_j}$$
4. Exemple

La pratique d'un sport nécessite des qualités physiques spécifiques. Celles-ci sont d'autant plus marquées que le niveau dans lequel évolue l'athlète est élevé. Le rôle particulier de certaines mesures anthropométriques dans une discipline donnée est connu des spécialistes (Thomas, 1975). A la demande de la Fédération Française de Hand-Ball, nous nous sommes intéressés à la définition morphologique des joueurs (Dufour et al, 1987). Il s'agit de l'étude de 8 caractères anthropométriques mesurés sur 64 sportifs appartenant au plus haut niveau de compétition (française et internationale). Les variables sont la stature (1), la taille assis (2), le diamètre biacromial (3), l'envergure (4), la hauteur utilisable (5), la longueur des membres inférieurs (6), la longueur des membres supérieurs (7) et l'empan de la main porteuse de la balle (8). L'empan est une mesure de la main : distance maximum (doigts écartés) entre l'extrémité du pouce et celle de l'auriculaire, main posée bien à plat. Les mesures, toutes exprimées ici en centimètres, ont été publiées (en millimètres) dans «Le Modèle Euclidien en Analyse de Données» (Pontier *et al*, 1990).

L'intérêt de séparer la taille et la forme est double. La stature est un critère de sélection dans la plupart des sports (Lot, 1981) mais peut-elle être identifiée au facteur taille? Bien décrire ce dernier, c'est se préparer à une bonne étude de la forme c'est-à-dire la description de la spécificité morphologique des cinq postes de jeu : ailier, pivot, arrière centre, arrière latéral et gardien.

Nous allons exposer les principaux résultats obtenus dans les différentes analyses proposées.

1. Les facteurs taille

Comme nous l'avons déjà souligné (remarque du paragraphe 2), les profils communs \vec{w} (critère 1) et \vec{G} (critère 2) sont pratiquement identiques et fournissent le même classement pour la taille, la stature \vec{e}_1 (critère 3) étant, elle, à part (tableau 1).

Si nous analysons les tailles des individus, que ce soit les valeurs individuelles v_i de la première composante principale (critère 1) ou que ce soit les totaux lignes

TABLEAU 1
Présentation des trois vecteurs permettant de définir le facteur taille.

variables	profil commun	\vec{u}	\vec{G}	\vec{e}_1
stature		0,4461	0,4461	1
taille assis		0,3240	0,3241	0
diamètre biacromial		0,2001	0,2002	0
envergure		0,4518	0,4518	0
hauteur utilisable		0,5010	0,5009	0
longueur des membres supérieurs		0,2864	0,2864	0
longueur des membres inférieurs		0,3067	0,3066	0
empan		0,1574	0,1574	0

x_i . du tableau initial (critère 2), les résultats sont semblables. Nous présentons à titre d'exemple, les valeurs individuelles de quatre handballeurs (tableau 2).

TABLEAU 2
Présentation des tailles et des classements par la taille de quatre individus.

individus	13	59	8	61
classement sur 64	13-ème	64-ème	27-ème	62-ème
valeur individuelle v_i (critère 1)	950.9971	851.9922	929.4989	866.3977
valeur individuelle x_i . (critère 2)	951,0	852,0	929,5	866,4
stature	187,0	170,0	182,0	175,0
taille assis	103,8	91,5	96,3	95,0
diamètre biacromial	37,0	39,5	38,0	31,5
envergure	194,0	178,0	193,0	176,0
hauteur utilisable	240,0	203,0	232,0	213,0
longueur des membres supérieurs	78,5	69,2	77,5	72,2
longueur des membres inférieurs	83,2	78,5	85,7	80,0
empan	27,5	22,3	25,0	23,7

2. Les facteurs forme

Les facteurs taille, critères 1 et 2 étant identiques, les tableaux de résidus (paragraphe 3) sont, par voie de conséquence, identiques. Une seule des deux études de forme est donc réalisée (critère 2). Les résultats présentés concernent l'A.C.P. du tableau de terme général $\frac{x_{ij}}{\sqrt{p \cdot j}} - x_i \cdot \sqrt{p \cdot j}$ pour l'équipollence, et l'A.C.P. du tableau de terme général $\frac{x_{ij}}{x_i \cdot \sqrt{p \cdot j}}$ pour la similitude géométrique.

Les variables

Il y a identité des représentations des variables pour l'équipollence géométrique (figure 2a) et pour la similitude géométrique (figure 3a), une fois le facteur taille enlevé. L'axe horizontal oppose deux groupes de caractères : la hauteur utilisable (5) et la longueur des membres supérieurs (7) d'une part, l'empan (8), la taille assis (2) et le diamètre biacromial (3) d'autre part. L'axe vertical oppose l'envergure (4) et la longueur des membres inférieurs (6) au diamètre biacromial (3) et à la hauteur utilisable (5).

Dans le cas du critère 3, on observe que le premier axe est un classement des plus petites aux plus grandes variables pour l'équipollence (figure 4) comme pour la similitude géométrique. Il faut donc étudier d'autres axes pour mieux décrire la forme. C'est ainsi que les résultats de l'étude de la forme : critère 3, axes 2-3 (figure 4) et critère 2, axes 1-2, sont semblables, quel que soit le modèle choisi. Une conséquence pratique des analyses suivant le critère 3 est que, la sélection des sportifs de haut niveau par la seule stature ne suffit pas, non seulement pour la définition de la taille mais également pour la compréhension de la forme.

Les individus

Compte tenu de ce qui précède, une seule étude des individus est présentée : celle issue du critère 2. L'examen de la représentation des individus dans le premier plan factoriel de l'équipollence géométrique (figure 2b) permet de mieux comprendre cette étude de la forme. Les individus situés à l'extrême droite du plan (individus 18,23,21,17,44) se caractérisent par la longueur des membres supérieurs et la hauteur utilisable. Ces athlètes appartiennent aux postes de gardien et d'arrières latéraux. Ils ont un rôle de défense : empêcher l'équipe adverse de marquer. A gauche et en haut du graphique (individus 59,13,50,22,9) se situent les arrières centre et les pivots. Ils ont un rôle d'interception et de distribution de la balle. Ils sont confrontés directement aux membres de l'équipe adverse. Un grand empan permet de bien tenir la balle, un diamètre biacromial (largeur des épaules) élevé de s'imposer; une taille assis élevée signifie un centre de gravité bas donc une bonne stabilité du joueur. Enfin, le bas du graphique (individus 62,39,56) est constitué par les ailiers, les joueurs les plus rapides. Ils se caractérisent par l'envergure et la longueur des membres inférieurs.

Il existe une légère différence dans la représentation des individus entre l'équipollence et la similitude géométrique (figure 3b). Prenons par exemple les individus 13 et 59. La représentation issue de l'équipollence géométrique les place côte à côte tandis que la représentation de la similitude les dissocie plus sur l'axe horizontal. Il en est de même pour l'individu 8 qui se confond totalement avec l'individu 61 dans l'équipollence et qui s'exprime dans la similitude géométrique. Si nous examinons les tailles liées à ces quatre hand-balleurs (tableau 2), nous constatons que la similitude géométrique attire les grands individus au centre du nuage de points et met plus en évidence les petits individus.

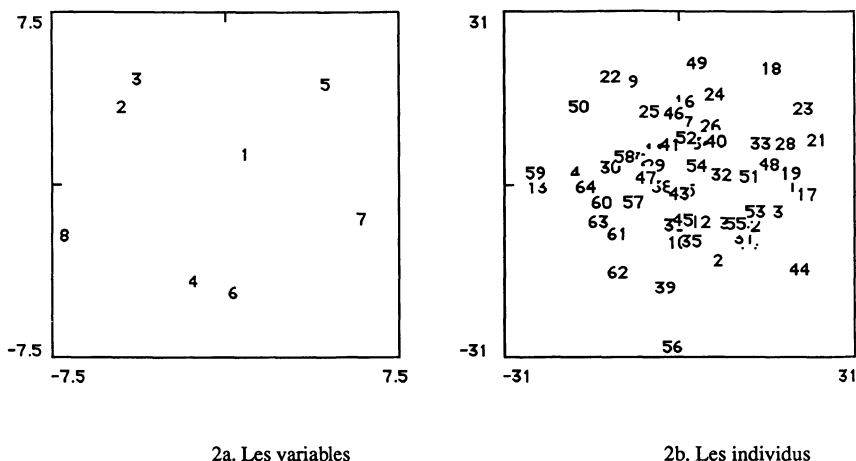


FIGURE 2
 Représentation de la forme sur les axes 1 et 2
 Méthode de l'équipollence géométrique
 Critère de taille 2

1 = stature, 2 = taille assis, 3 = diamètre biacromial, 4 = envergure, 5 = hauteur utilisable, 6 = longueur des membres inférieurs, 7 = longueur des membres supérieurs, 8 = empan

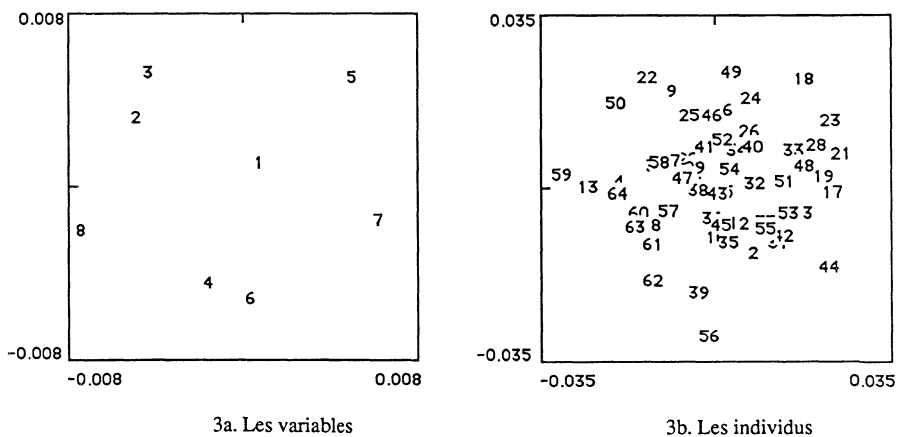


FIGURE 3
 Représentation de la forme sur les axes 1 et 2
 Méthode de la similitude géométrique
 Critère de taille 2

1 = stature, 2 = taille assis, 3 = diamètre biacromial, 4 = envergure, 5 = hauteur utilisable, 6 = longueur des membres inférieurs, 7 = longueur des membres supérieurs, 8 = empan

Remarque

Notons que dans cet exemple, les résultats de l'analyse des correspondances sont également semblables à ceux de l'étude de la forme – critère 2 (similitude géométrique). Ceci est dû à une relative égalité des pondérations sur les individus. En effet, en A.C.P., les n individus ont tous même importance : $\frac{1}{n} = 0,015625 (n = 64)$ alors qu'en A.F.C., la pondération utilisée est la distribution marginale des lignes $p_{i.} = \frac{x_{i.}}{x_{..}}$ (cf remarque, paragraphe 3). Le plus grand handballeur de notre échantillon (n°14) a un poids $p_{14.} = 0,016865 (x_{14.} = 993,6; x_{..} = 58915,7)$ et le plus petit (n°59) un poids $p_{59.} = 0,014461$ (tableau 2); les pondérations des autres joueurs variant entre ces deux valeurs proches de la pondération uniforme.

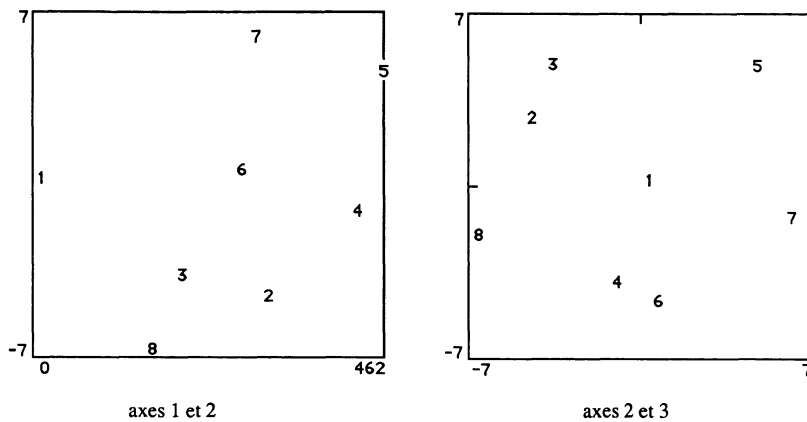


FIGURE 4
Représentation de la forme
Méthode de l'équipollence géométrique
Critère de taille 3

1 = stature, 2 = taille assis, 3 = diamètre biacromial, 4 = envergure, 5 = hauteur utilisable, 6 = longueur des membres inférieurs, 7 = longueur des membres supérieurs, 8 = empan

Conclusion

Les données contiennent toujours une structure inhérente au matériel étudié. Elle peut être banale comme dans l'étude de la croissance : les enfants d'un groupe d'âge ont une taille inférieure à ceux d'un autre groupe, plus âgés. Elle peut être moins évidente comme par exemple la comparaison de la taille de crânes de groupes ethniques différents sur la base de la longueur antéropostérieure maximum, de la largeur transverse maximum et de la hauteur du crâne.

Dans tous les cas, cette structure (facteur taille) est éliminée en étudiant les résidus (facteur forme) entre les valeurs observées et les valeurs modélisées. L'étude

de la forme peut être appréhendée de deux façons différentes : soit à partir des données brutes, soit à partir des profils.

L'exemple traité issu d'un problème anthropométrique n'a pas révélé de différences très significatives entre les méthodes que ce soit pour les critères de taille ou pour le vecteur forme utilisé. Cela est dû pour une grande part, à l'homogénéité de l'échantillon : un groupe de 64 handballeurs français de même niveau de pratique sportive. Le choix entre les analyses proposées reste lié à la problématique de départ. Si une répartition en agrégats est soupçonnée, il est préférable de prendre comme facteur de taille, la première composante principale de l'A.C.P. non centrée (critère 1). S'il existe une tendance non linéaire, comme par exemple une relation d'allométrie, on choisira plutôt le profil moyen (critère 2). Enfin, l'expérimentateur peut souhaiter privilégier une variable (empirisme, routine, résultats préalables...); l'analyse – critère 3 est alors la plus appropriée.

Dans le cas de grands écarts entre les distributions marginales des lignes, la similitude géométrique gomme ces différences, et par là même, doit être préférée à l'équipollence pour l'étude de la forme.

Références Bibliographiques

- BURT C. (1946). L'étude factorielle des types physiques, *Biotypologie*, t.VIII, 1-2, 42-55.
- COHEN J.I. (1940). Are differences in size between parts of body due to general or specifics, *Proc. Nat. Acad. Scien. Washington*, 26, 524-526.
- Collectif (1986). L'homme, son évolution, sa diversité. Manuel d'Anthropologie Physique, sous la direction de D. Ferembach, C.Suzanne, M.C. Chamla, ed C.N.R.S., Paris, 572p.
- DUFOUR A.B., ROUARD A., PONTIER J., MAURIN L. (1987). Profil morphologique des hand-balleurs français de haut niveau. *Science et Motricité*, 2 : 3-9.
- ECKART C., YOUNG G., (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, vol. 1, 3 : 211-218.
- GABRIEL K.R. (1971). The biplot graphic display of matrices with application to principal component analysis. *Biometrika*, 58, 3 : 453-467.
- GOOD I.J. (1969). Some applications of the singular decomposition of a matrix. *Technometrics*, 11 : 823-831.
- GOWER J.C. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 53, 3-4 : 325-338.
- HEALY M.J.R., TANNER J.M. (1981). Size and shape in relation to growth and form, *Symp. Zool. Soc. Lond.*, 46, 19-35.
- JOLICOEUR P., (1991). Introduction à la biométrie, Décarie-Masson, 1e édition, Montréal, 300p.
- JOLICOEUR P., MOSIMANN J.E. (1960), Size and shape variation in the painted turtle. A principal component analysis, *Growth*, 24, 339-354.

- LOT J.F. (1981). Profils anthropométriques et sport, Thèse de médecine n°162, Bordeaux II, 93p.
- MOSIMANN J.E. (1970). Size Allometry : size and shape variables with characterizations of the Lognormal and generalized Gamma distributions, *J.A.S.A.*, 65, 930-945.
- OKAMOTO M. (1972). Four techniques of principal component analysis, *Journal of Japan Stat. Soc.*, 2, 63-69.
- PILARDEAU P., JOUBLIN M., VALERI L., ENGEL J.C. (1978). Choix du sport et taille. *Médecine du Sport*, t. 52, 5, 285-287.
- PEARCE S.C., HOLLAND D.A. (1961). Analyse des composantes, outil en recherche biométrique, *Biométrie-Praximétrie*, t. II, 3-4, 159-177.
- PONTIER J., DUFOUR A.B., NORMAND M. (1990). Le modèle euclidien en analyse des données, éditions de l'Université de Bruxelles – Ellipses – collection SMA, 428p.
- RAO C.R. (1964). The use and interpretation of principal components analysis in applied research, *Sankhya*, 26, 329-357.
- SEMPE M., ROY-PERNOT M.P., PEDRON G. (1979). Auxologie, Méthode et Séquences, Théraplix, Paris, 205p.
- THOMAS R. (1975). La morphologie (chapitre VI), in *La Réussite Sportive*, P.U.F., Paris, 121-143.
- VANDERVAEL F. (1980). Biométrie Humaine, ed Masson, Paris, 165p.