

Comparaison des populations de Kharga et de Saï

André LEGUEBE et Ginette BILLY

1. INTRODUCTION

L'analyse d'un échantillon de 59 hommes de l'île de Saï (Haute Nubie soudanaise) a montré que la population de cette île manifeste une tendance à être morphologiquement plus proche de la population de l'oasis de Kharga (Égypte) que de celles de la vallée du Nil avec lesquelles elle a pu être comparée (Billy, 1987).

Le fait que la littérature ne mentionne, pour l'ensemble des populations introduites dans la comparaison, que les moyennes et les écarts-type des sept variables considérées (stature, longueur de la tête, largeur de la tête, largeur de la face, hauteur morphologique de la face, largeur et hauteur du nez) limite la comparaison à une classification hiérarchique basée sur la matrice des distances C^2_H de Penrose. Disposant des valeurs individuelles des deux populations les

plus proches, Kharga (Twisselmann, 1951) et Saï, il a paru intéressant d'essayer de préciser la nature des différences et des ressemblances observables entre les individus de ces deux populations.

2. DISTRIBUTION DES VARIABLES ET PARAMETRES

Les résultats (moyennes, écarts-type, coefficients de variation) sont donnés dans le tableau 1. La normalité des distributions univariées a été éprouvée au moyen du test de Filliben (Leguebe et Albert, 1989) : les probabilités de r indiquent que les distributions peuvent être considérées comme normales, la probabilité la plus faible, celle de la largeur du nez étant comprise entre 0.01 et 0.02. Si les sujets des deux échantillons sont regroupés, les distributions restent normales (avec $.02 < p < .05$ pour la largeur du nez).

Tableau 1 : Moyennes, écarts-type et coefficients de variation des variables; tests de normalité (méthode de Filliben).

Variables	Kharga (n = 50)					Saï (n = 59)				
	\bar{x}	s	c.v. %	r	prob.	\bar{x}	s	c.v. %	r	prob.
1. Stature (cm)	165.8	6.87	4.2	.992	.50-.75	167.2	6.10	3.6	.991	.50
2. Longueur tête (mm)	189.9	7.80	4.1	.992	.50-.75	190.5	6.17	3.2	.991	.50
3. Largeur tête (mm)	141.2	5.51	3.9	.988	.25-.50	143.3	6.18	3.6	.990	.25-.50
4. Largeur face (mm)	133.6	6.25	4.7	.991	.50-.75	133.3	5.19	3.9	.985	.10-.25
5. Hauteur morph. face (mm)	115.3	7.07	6.1	.979	.05-.10	118.9	6.91	5.8	.993	.50-.75
6. Largeur nez (mm)	37.4	3.22	8.6	.988	.25-.50	39.0	2.56	6.6	.970	.01-.02
7. Hauteur nez (mm)	50.2	4.09	8.2	.988	.25-.50	50.5	3.14	6.2	.987	.10-.25

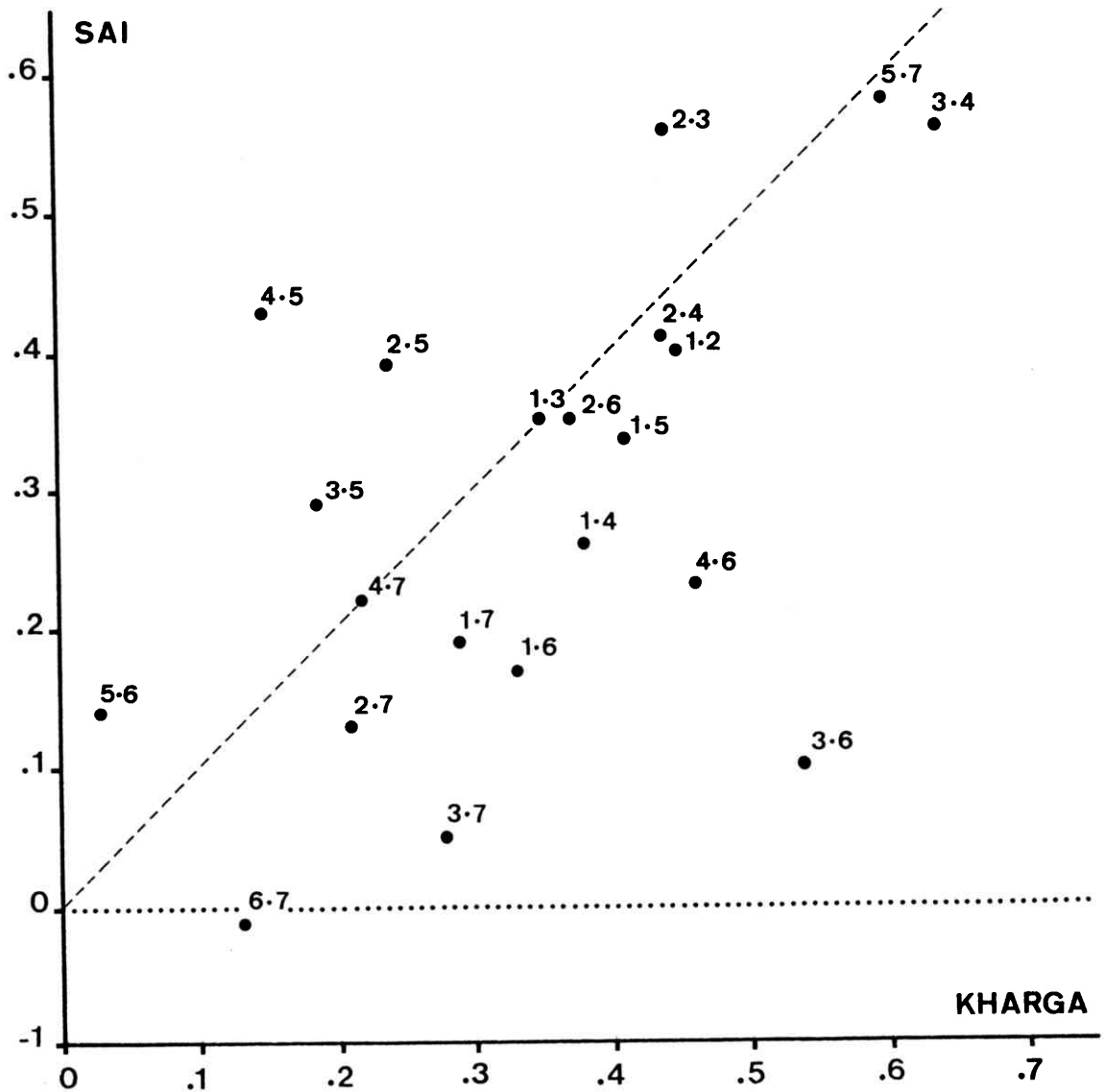


Fig. 1 : Relation entre les coefficients de corrélation entre les différentes variables pour l'oasis de Kharga et l'île de Saï (1 : stature; 2 : longueur de la tête; 3 : largeur de la tête; 4 : largeur de la face; 5 : hauteur morphologique de la face; 6 : largeur du nez; 7 : hauteur du nez).

Tableau 2 : Différences entre les moyennes, tests de normalité pour les deux échantillons groupés et analyses de la variance.

Variables (n=109)	Diff. Kharga-Saï	Normalité r prob.	Anova F prob.
Stature (cm)	-1.37	.992 .25	1.21 .270
Longueur tête (mm)	-0.58	.996 .25-.50	0.20 .660
Largeur tête (mm)	-2.12	.993 .25-.50	4.29 .040*
Largeur face (mm)	0.24	.994 .50	0.05 .830
Hauteur morph. face (mm)	-3.65	.991 .10-.25	7.39 .010*
Largeur nez (mm)	-1.64	.986 .02-.05	8.81 .004*
Hauteur nez (mm)	-0.33	.989 .10	0.23 .640

Bien que les écarts-type et les coefficients de variation de toutes les variables soient plus élevés pour l'échantillon de Kharga que pour celui de Saï, l'hypothèse de l'homogénéité des variances de chacune des variables est acceptée par les tests de Cochran et de Bartlett-Box.

Les dimensions moyennes de Kharga, sont inférieures à celles de Saï, sauf pour la largeur de la face (tableau 2): l'égalité des moyennes est rejetée, par une analyse de la variance, pour la largeur de la tête, pour la hauteur morphologique de la face et pour la largeur du nez.

Les deux populations diffèrent donc par certaines de leurs mesures mais, tenant compte des corrélations qui existent entre les différentes variables, elles diffèrent également pour l'ensemble des mesures considérées : ceci n'est pas nécessairement toujours le cas (Morant, 1926, p. 269; Healy, 1969; Le-guebe, 1991a).

La distance de Mahalanobis entre les centroïdes des deux populations est égale à 1.03 : tout en étant relativement faible, un test de Hotelling montre qu'elle est très significative ($T^2 = 27.81$, $F = 3.75$, 7 et 101 d.d.l., prob = 0.001). L'importance des effectifs est un élément qui favorise l'obtention d'un résultat aussi significatif malgré la faiblesse de la distance.

Tableau 3 : Test d'autocomparaison pour le T^2 de Hotelling (Kharga et Saï) (T^2 obs. = 22.07).

Tirages	$T^2 > T^2$ obs.	\bar{x}	s
1 à 100	0	7.01	3.30
101 à 200	2	7.94	4.10
201 à 300	1	7.49	4.14
301 à 400	0	6.13	3.05
401 à 500	0	7.52	3.74

Ce résultat est confirmé par un test d'autocomparaison (Edgington, 1980; Pontier et Pernin, 1985). Si on calcule la valeur du test d'Hotelling pour 5x100 combinaisons au hasard des 109 individus en deux groupes de 50 et de 59 sujets respectivement, trois

combinaisons seulement sur les 500 donnent une valeur de T^2 plus élevée que pour les échantillons observés : dans le tableau 3, figurent les moyennes et les écarts-type de T^2 pour les cinq séries de 100 tirages.

3. ANALYSES DISCRIMINANTES

En vue de préciser en quoi les deux groupes se différencient, nous avons eu recours à l'analyse discriminante pas à pas (Norusis, 1990) , utilisant comme critère d'introduction des variables la minimisation du lambda de Wilks qui, pour mémoire, correspond au rapport de la variabilité intragroupe à la variabilité totale : λ est égal à l'unité quand les vecteurs moyens des deux groupes sont égaux et il tend vers zéro au fur et à mesure où la variabilité intragroupe diminue par rapport à la variabilité totale. L'ordre d'introduction des variables, les λ de Wilks et le seuil de signification à chaque pas sont donnés dans le tableau 4. La taille n'est pas incluse dans l'analyse parce que la valeur de F est devenue inférieure à 1.

Tableau 4 : Analyse discriminante : ordre d'introduction des variables (Wilks).

	λ	Probabilité
1. Largeur nez	.924	.0037
2. Hauteur face	.877	.0010
3. Largeur face	.846	.0005
4. Largeur tête	.821	.0004
5. Longueur tête	.805	.0004
6. Hauteur nez	.794	.0005
7. Stature		non introduite

C'est évidemment la largeur du nez qui est introduite au premier pas parce que c'est la variable qui présente la valeur la plus significative de F ($F = 8.81$, $p = 0.004$); cette variable est suivie par la hauteur de la face mais au troisième pas déjà, c'est la largeur de la face qui intervient bien que F vaille seulement 0.05 ($p = 0.830$); en fait l'introduction de variables supplémentaires n'améliore pas le classement des individus.

Le caractère significatif du test n'entraîne pas automatiquement que la séparation des sujets des deux populations reproduise de manière précise la répartition initiale : 15 Kharga (30 %) et 16 Saï (27.1 %) sont mal classés, soit 28.4 % des sujets groupés.

L'analyse a été répétée au moyen de l'analyse discriminante logistique plus robuste eu égard aux conditions d'application (Leguebe et Albert, 1989; Albert et Harris, 1987). L'ordre de sélection des variables est identique à celui obtenu précédemment et deux variables seulement sont retenues (tableau 5). On constate que l'addition d'une troisième ou d'une quatrième variable n'améliore pas le classement (tableau 6).

Tableau 5 : Analyse discriminante logistique. Ordre de sélection des variables et probabilités correspondantes.

Variabes	Probabilités
1. Largeur nez	0.003
2. Hauteur morph. face	0.019
3. Largeur face	0.056
4. Largeur tête	0.070
5. Longueur tête	0.167
6. Hauteur nez	0.193
7. Stature	—

Tableau 6 : Analyse discriminante logistique. Classements obtenus pour différents ensembles de variables.

Variables	Larg. nez		+Haut. face		+Larg. face		+Larg. tête	
	classés en 1	classés en 2	classés en 1	classés en 2	classés en 1	classés en 2	classés en 1	classés en 2
1. Kharga	25	25	33	17	29	21	31	19
2. Saï	14	45	16	43	17	42	16	43

Enfin, on a comparé au résultat obtenu pour la paire de variables (largeur nez (6) — hauteur face (5))

conduisant au classement le plus efficace, ceux correspondant aux couples de variables servant à calculer les indices classiques (céphalique, facial et nasal) : les résultats repris dans le tableau 7, indiquent que, de toute évidence, aucune de ces trois combinaisons ne conduit à une discrimination proche de la répartition initiale. Le mauvais classement des individus composant l'échantillon de Saï est une conséquence de la plus grande variabilité observée chez les Kharga.

Tableau 7 : Analyse discriminante logistique. Classements pour les trois paires de variables servant au calcul des indices.

	Long. tête		Haut. face		Haut. nez	
	Larg. tête	Larg. fac	Larg. fac	Larg. nez	Larg. nez	Larg. nez
	classés en 1	classés en 2	classés en 1	classés en 2	classés en 1	classés en 2
1. Kharga	26	24	23	27	23	27
2. Saï	43	16	43	16	45	14

4. ANALYSES EN COMPOSANTES PRINCIPALES

C'est principalement au niveau des coefficients de corrélation entre les variables que l'on observe des différences marquées entre les deux échantillons (tableau 8 et fig. 1) : les coefficients sont donnés avec cinq décimales afin de permettre une utilisation ultérieure de ces échantillons dans d'autres analyses multivariées (Leguebe, 1991b). Pour toutes les paires de variables situées au-dessus de la diagonale de la figure 1, la valeur du coefficient est plus élevée pour les Saï que pour les Kharga à l'inverse des paires situées au-dessous de la diagonale. Ces différences entre les coefficients de corrélation traduisent des différences au niveau de la morphologie. Pour essayer de les préciser, chacun des deux échantillons a été soumis à une analyse en composantes principales.

Tableau 8 : Matrices de corrélation des Kharga (au-dessous de la diagonale) et des Saï (au-dessus de la diagonale).

	1	2	3	4	5	6	7
							SAÏ (n=59)
1. Stature	—	.39740	.35240	.26150	.33860	.17340	.18650
2. Longueur tête	.44631	—	.56300	.41200	.38630	.35100	.12730
3. Largeur tête	.34400	.43532	—	.55700	.29140	.09780	.04900
4. Largeur face	.37751	.44226	.63750	—	.42800	.22560	.21860
5. Hauteur face	.40622	.23854	.19023	.15248	—	.14250	.57810
6. Largeur nez	.32899	.37168	.53519	.45472	.02958	—	.01240
7. Hauteur nez	.29019	.21669	.27986	.21675	.59582	.13375	—
	KHARGA (n=50)						

Les pourcentages de la variance totale expliqués par chacun des vecteurs successifs figurent dans le tableau 9 : seuls les deux premiers vecteurs propres ont une valeur supérieure à l'unité et ils expliquent ensemble quelque 60 % de la variabilité totale. Ceci nous incite à penser qu'une partie importante des variations pourrait ne pas correspondre à des facteurs de structure.

Tableau 9 : Analyse en composantes principales : pourcentages et pourcentages cumulés des valeurs propres des composantes.

Composantes	Kharga (n=50)		Saï (n=59)	
	%	% cumul.	%	% cumul.
1	43.9	43.9	40.8	40.8
2	20.1	64.0	18.1	59.9
3	10.6	74.5	13.2	72.1
4	8.1	82.0	10.9	83.1
5	7.4	90.1	7.5	90.6
6	5.2	95.3	5.1	95.7
7	4.7	100.0	4.3	100.0

Examinons, pour chacune des deux populations, la projection des variables dans l'espace défini par les deux premiers axes (fig. 2). On constate que, pour toutes les variables, l'extrémité du vecteur est située au-delà du cercle de distribution équilibrée, cercle qui correspond à la situation pour laquelle toutes les variables interviendraient de manière égale dans la définition des divers axes. L'angle fait par chacun des vecteurs avec un axe déterminé correspond au coefficient de corrélation de la variable avec cet axe. Ainsi, dans le cas des Saï pour l'espace défini par les deux premiers axes, on observe que la taille (1) et la largeur de la face (4) sont fortement corrélées avec l'axe 1 et pratiquement pas avec le deuxième axe, situation qui se différencie de celle observée chez les Kharga. On constatera toutefois, pour l'espace considéré, qu'il existe entre ces deux populations une certaine similitude quant au positionnement des variables; en particulier, la largeur du nez (6) et la hauteur morphologique de la face (5), qui sont introduites aux premiers pas de l'analyse discriminante, sont orthogonales entre elles.

Considérons enfin les valeurs des contributions relatives, exprimées en %, correspondant à la part de la dispersion de chacune des variables expliquée par les axes successifs : une représentation graphique de ces valeurs est donnée dans les figures 3 (Kharga) et 4 (Saï).

Les parties des variables expliquées par les axes successifs, respectivement chez les Kharga et chez les Saï, sont très nettement différentes en particulier pour la hauteur morphologique de la face (5) et la largeur du nez (6); on notera également que le premier axe correspond aussi à une part importante de

la variabilité de la taille (1), de la longueur de la tête (2), de la largeur de la tête (3) et de la largeur de la face (4) notamment.

Un contraste supplémentaire entre la hauteur de la face et la largeur du nez mérite en outre d'être souligné : les deux premiers axes expliquent environ 80 % de la variabilité de la hauteur de la face (5) alors que, pour la largeur du nez (6), ce pourcentage ne s'élève qu'à 60 % pour les Kharga et à 30 % pour les Saï.

Bien que la valeur propre du 3^e axe soit inférieure à l'unité, la partie de la variabilité totale qu'il explique est encore de 10.6 % pour les Kharga et de 13.2 % pour les Saï : ce troisième axe explique chez les Saï une part notable des mesures de largeur(nez (5), tête (3) et face(4)) alors que chez les Kharga, ce sont la stature (1), la longueur de la tête (2) et la hauteur du nez (7) qui sont principalement impliquées.

5. CONCLUSION

L'application de diverses procédures de l'analyse multivariée aux échantillons de l'oasis de Kharga et de l'île de Saï contribue à préciser la nature des différences faibles qui les séparent. Même en ne prenant en considération que deux variables parmi les sept pour lesquelles on dispose des paramètres, la distance entre les centroïdes des deux groupes est significativement différente mais la possibilité de restituer correctement l'appartenance des individus à chacun des deux groupes définis initialement reste faible.

L'ensemble de ces constatations tend à montrer que, pour l'anthropologiste, l'intérêt des analyses multivariées réside plus dans la puissance descriptive qu'elles détiennent, permettant d'appréhender des aspects plus complexes de la morphologie, que dans les résultats des tests auxquels elles aboutissent même si les conditions d'application sont remplies.

Soulignons enfin la possibilité qui s'offre d'approfondir l'analyse de données anciennes limitées aux moyennes et aux écarts-type, données fort nombreuses et qui ont conservé leur intérêt étant donné qu'il n'est, le plus souvent, pas possible de les actualiser. Une première analyse est effectuée sur l'ensemble des données anciennes et des données modernes en utilisant seulement les paramètres de tendance centrale et de dispersion; elle peut être complétée par une ou plusieurs analyses portant sur les échantillons pour lesquels on dispose aussi des matrices d'association entre les variables.

La confrontation des deux séries de résultats est susceptible de projeter un éclairage complémentaire sur l'interprétation des relations pouvant exister entre ces populations.

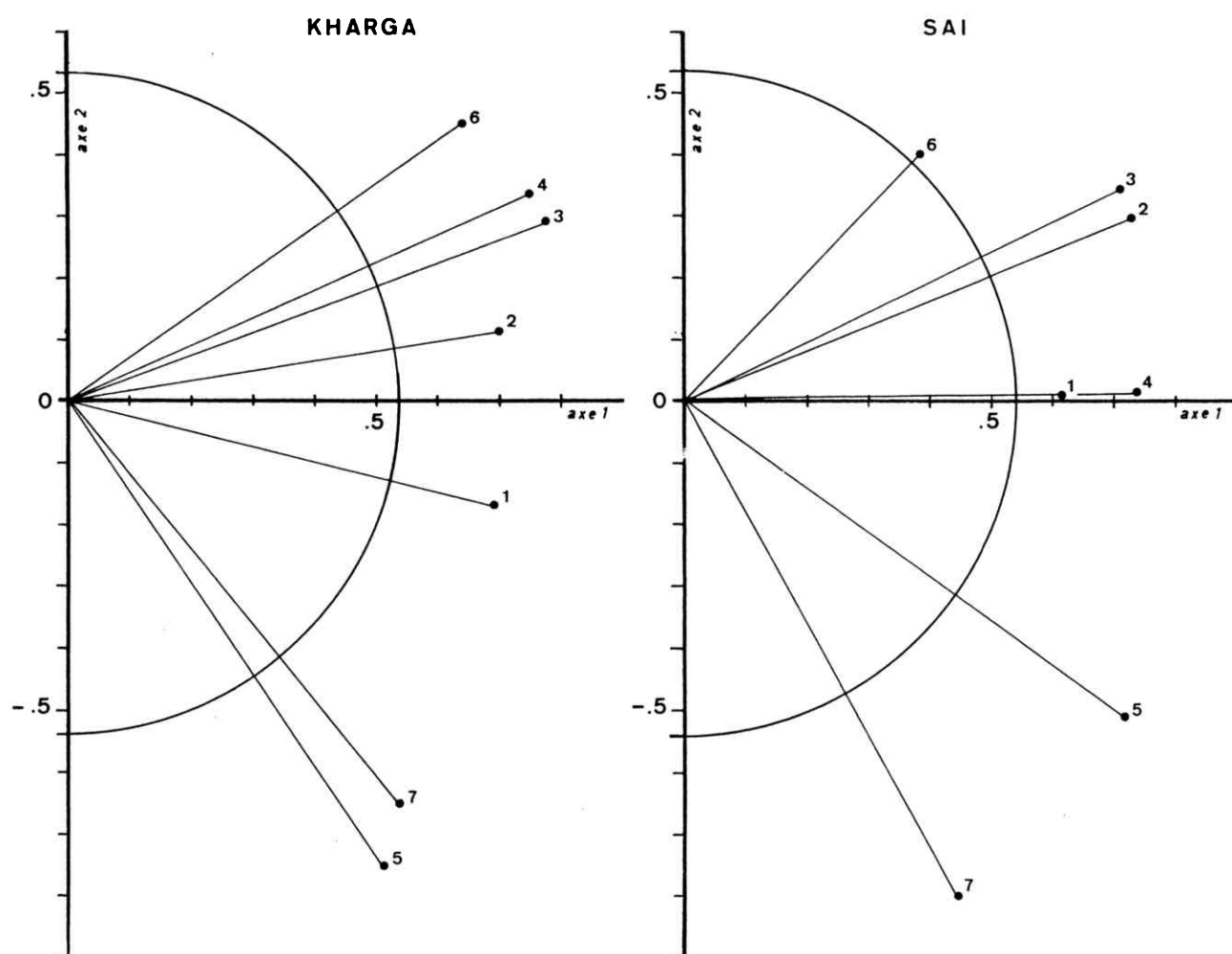


Fig. 2 : Projection des variables dans l'espace des deux premiers axes pour les échantillons de Kharga et de Saï (1 : stature; 2 : longueur de la tête; 3 : largeur de la tête; 4 : largeur de la face; 5 : hauteur morphologique de la face; 6 : largeur du nez; 7 : hauteur du nez).

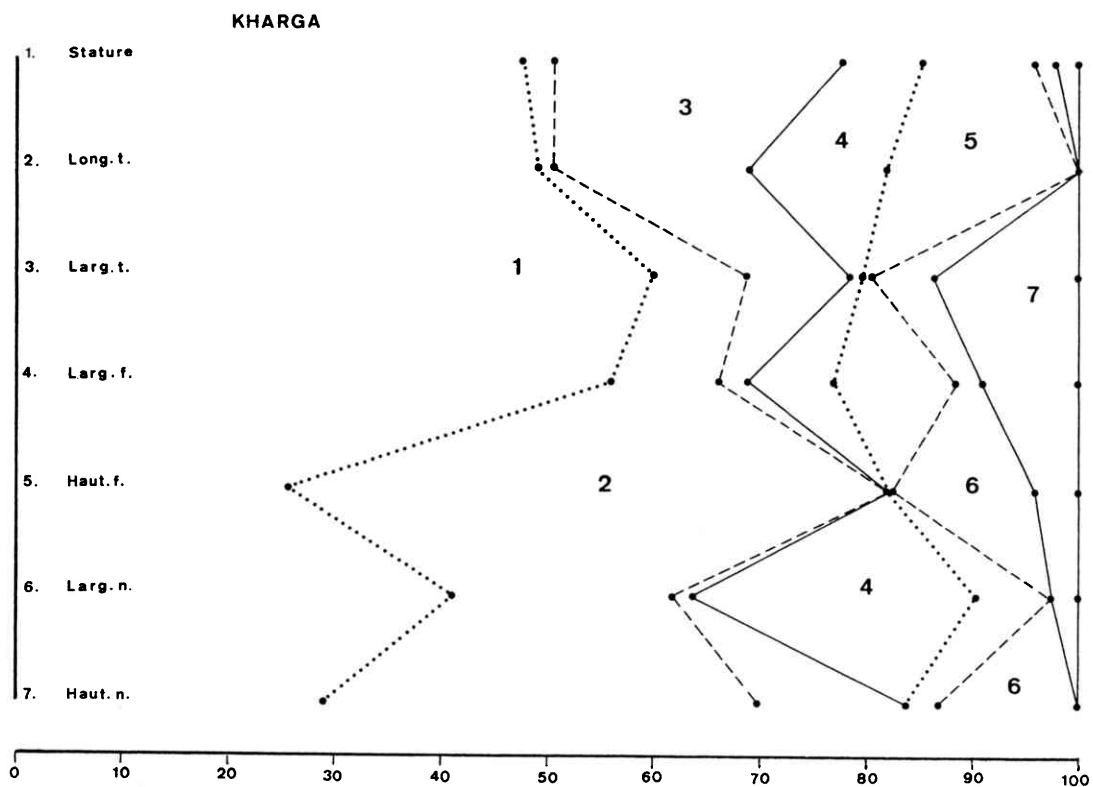


Fig. 3 : Répartition des contributions relatives des diverses variables selon les axes pour l'échantillon de Kharga (les chiffres 1 à 7 dans le corps du graphique correspondent aux vecteurs).

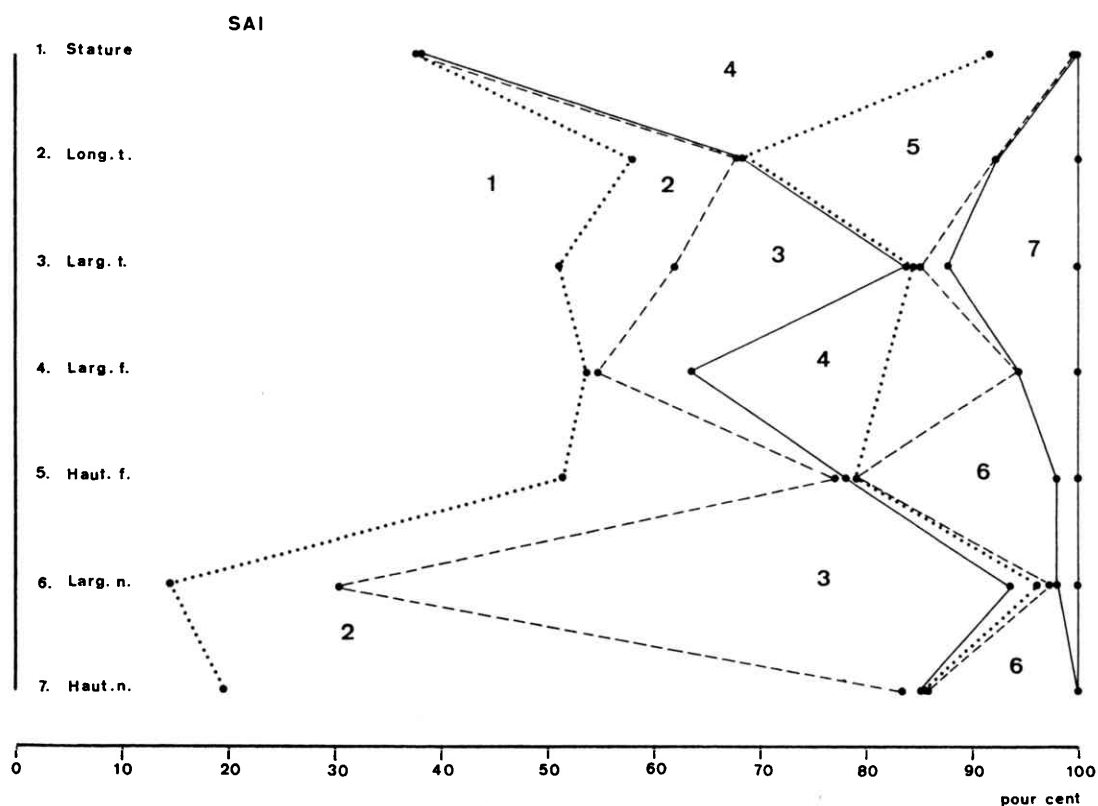


Fig. 4 : Répartition des contributions relatives des diverses variables selon les axes pour l'échantillon de Saï (les chiffres 1 à 7 dans le corps du graphique correspondent aux vecteurs).

Bibliographie

- ALBERT, A. et HARRIS, E.K., 1987. *Multivariate interpretation of clinical laboratory data*. New York and Basel, Marcel Dekker, 312 p.
- BILLY, G., 1987. La population de l'île de Saï en Haute Nubie soudanaise. *Bull. Mém. Soc. Anthropol. Paris*, série 14, 4 (4) : 249-260.
- EDGINGTON, E.S., 1980. *Randomization tests*. New York, Marcel Dekker Inc., 12+287p.
- HEALY, M.J.R., 1969. Rao's paradox concerning multivariate tests of significance. *Biometrics*, 25 : 411-413.
- LEGUEBE, A., 1991a. Apparentes contradictions entre les résultats d'analyses uni- et multivariées. *Homo*, Stuttgart, 42(2) : 171-178.
- LEGUEBE, A. 1991b. Publishing results of multivariate analysis. *Homo*, Stuttgart (sous presse)
- LEGUEBE, A. et ALBERT, A., 1981. Détermination du sexe de crânes au moyen de l'analyse discriminante logistique. *Z. Morph. Anthropol.*, 72 (2) : 171-179.
- LEGUEBE, A. et ALBERT, A., 1989. Test de normalité graphique en paléontologie humaine. *Z. Morph. Anthropol.*, 77 (3) : 259-271.
- MORANT, G.M., 1926. Studies on Palaeolithic man. I. The Chancelade skull and its relationship to the modern Eskimo skull. *Annals of Eugenics*, 1(3/4) : 257-276.
- NORUSIS, M.J., 1990. *SPSS/PC+. Advanced Statistics 4.0*. Chicago.
- PONTIER, J. et PERNIN, M.O., 1985. Principes et pratique des tests d'autocomparaison. *Biométrie-Praximétrie*, 25 : 67-90.
- TWIESSELMANN, F., 1951. Expédition anthropologique du Dr D.J.H. Nyéssen. I.- L'oasis de Kharga. *Bull. Inst. roy. Sci. nat. Belgique*, 27(14) : 36 p.

Adresse des auteurs : A. LEGUEBE
 Institut royal des Sciences naturelles de Belgique,
 Section d'Anthropologie et de Préhistoire,
 rue Vautier, 29
 B-1040 Bruxelles (Belgique)

G. BILLY
 Faculté des Sciences, Laboratoire d'Anthropologie,
 avenue Albert Thomas, 123
 F-87060 Limoges (France)