

SOME CASE STUDIES OF THE STATISTICAL ANALYSIS OF SEXUAL DIMORPHISM

R. A. Reymont

Paleontological Institute, University of Uppsala

(Submitted for publication September 18, 1968)

Abstract. This paper presents a case study of sexual dimorphism in frogs, martins, turtles, crabs, brine shrimps, and grasshoppers. An analysis of sexual dimorphism in two species of frogs gave the results:

Rana temporaria L. The material comprises samples from Sweden. The variables selected are X_1 =length of body, X_2 =length of tibia, and X_3 =length of praeallux (=heel protruberance). Although small, the value of D^2 is statistically significant. If we regard the scatter clouds of observations represented in terms of the covariance matrices, as ellipsoids (thus assuming the variables to be multivariate normally distributed), in the present case the ellipsoids could be shown to be differently inflated, but to have the same orientation. This suggests the same patterns of growth to occur for males and females but for the females to have a wider range of dispersion.

Rana esculenta L. This material is from various parts of Sweden. The variables are the same as for *R. temporaria*. D^2 is very small and there is little indication of significant morphologic sexual dimorphism, although the average size differences are consistently in favor of the males. There is, however, strong sexual dimorphism in the variances and covariances. The dispersion ellipsoids are significantly differently inflated and their major axes significantly differently oriented. Therefore, males and females of this species show only weak size dimorphism but strong sexual dimorphism in their growth patterns. An analysis of two cranial dimensions also indicated the absence of significant sexual dimorphism in these (length and breadth of cranium for 35 females and 14 males gave a D^2 of 0.197).

Martes americana Turton. The covariance matrices were found to be homogeneous with respect to inflation, but the second and third principal axes differed significantly in orientation. The squared generalized distance is small and does not indicate significant sexual dimorphism to occur.

Chrysemys picta marginata Agassiz. The squared generalized distance is significant and indicates appreciable sexual dimorphism to occur in the carapace of this turtle.

Carcinus maenas L. The material of *Carcinus* was obtained from the Bay of Arcachon, south-western France. The variables chosen for study are: X_1 =maximum breadth of carapace; X_2 =height of carapace; X_3 =distance from posterior of carapace to level of maximum breadth; X_4 =width of posterior of carapace; X_5 =length of right claw;

X_6 =width of right claw; X_7 =length of left claw; X_8 =width of left claw.

The squared generalized distance for heterogeneous covariance matrices is rather large. The scatter ellipsoids are oriented at widely different angles from each other, a result of different differential growth patterns in the two sexes.

The right claw of females is slightly larger and broader than the left claw ($\bar{X}_5=1.23$ and $\bar{X}_6=0.95$ against $\bar{X}_7=1.20$ and $\bar{X}_8=0.88$). The same applies for males ($\bar{X}_5=1.26$ and $\bar{X}_6=0.99$ against $\bar{X}_7=1.24$ and $\bar{X}_8=0.89$).

Artemia salina L. The material analysed derives from eggs obtained from salt pools around Cagliari, Sicily, and reared in brine with a concentration of 140%. The variables are: X_1 =body length; X_2 =length of abdomen; X_3 =length of prosoma. The covariance matrices show very strong heterogeneity, both in the degree of inflation (male dispersions greater than female) of the dispersion ellipsoids as well as in their orientation. The difference mean vector is 0.0368, 0.0438, 0.0286 and the heterogeneous generalized distance was found to be 0.85. Although the dimorphic distance is small, it is nevertheless statistically significant, with $F=19.53$ for 3 and 135 degrees of freedom. Therefore, the morphologic sexual dimorphism in this species is slight, though significant, as regards size differences, but very strong with respect to the patterns of growth.

Omocestus haemorrhoidalis L. This interesting material was obtained by Dr. Dag Gärdefors, Uppsala, from the three isolated distribution areas of the species in Sweden, notably, Kinnekulle, Öland and Gotland.

The variables measured are: X_1 =length of hind femur; X_2 =pronotal length; X_3 =elytron length; X_4 =least width between pronotus ridges.

These results indicate that the sexual dimorphism in *Omocestus* is particularly pronounced and that this is of about the same strength for the populations of Kinnekulle and Öland, but stronger still for the Gotland population. The covariance matrices for all three pairs of samples are heterogeneous, but whereas the Kinnekulle and Öland material show marked differences in both the inflation of the scatter ellipsoids and the relative orientations of all principal axes, the samples from Gotland differ in having almost the same degree of inflation of the ellipsoids, while all sexes are differently oriented.

The relative magnitude of the sexual-dimorphic differences

may be gauged from a consideration of the distances between grasshoppers of the same sex from the various localities. The generalized distances between localities for the same sexes are clearly very much smaller than those between sexes at the same locality. The greatest distance between localities is for females from Öland and Kinnekulle which stands in reasonable agreement with the distance between the males. Noteworthy also is the fact that five of the six comparisons involve heterogeneous covariance matrices. Hence, the females from Gotland and Kinnekulle may have identical growth patterns, whereas the other pairs show assessable differences. This relationship is also clearly shown up by a canonical variate analysis of the samples from each locality. At each geographic location, 6 samples of *Omocestus* were taken, which material was subsequently sexed, thus making a total number of 36 samples, stratified with respect to sex and geographic origin. The plot of the canonical variate means for the first two transformed variables indicates quite clearly that there is a sharp division into two equally sized clusters, the one comprising males, the other females, but there is no clearly discernible geographic clustering. The eigenvectors indicate that the main clustering influence derives from variables X_1 and X_3 and that X_2 is practically without effect for the delineation of morphometric sexual dimorphic differences.

The sensitivity of the test used for homogeneity of variances and covariances to departures from normality is demonstrated by a practical example.

INTRODUCTION

A particularly useful multivariate statistical procedure for the analysis of morphologically manifested sexual dimorphism is by means of the generalized statistical distance, which was originally introduced by the Indian statistician P. C. Mahalanobis for anthropometric studies. Subsequently, the generalized distance has been extended to a wider class of morphometric problems in Biology, including Paleontology.

In applying this procedure in Paleontology one works to a certain extent in the dark, for there is often uncertainty regarding the homogeneity of the samples (i.e., whether the material is chronologically pure and has not been subjected to postmortal transport).

The test cases accounted for in this paper were based on living animals in order to provide a background against which results for comparable groups of fossils may be judged from the point of view of the afore-mentioned sources of heterogeneity. Another problem of some rank concerns the commonness of, and the importance of, the effect of heterogeneity in the variances and covariances of the samples. This has some bearing on the generalized distance as it depends partly on the means of the variables involved

and partly on the variances and covariances of the variables. The details of calculation will not be taken up here and reference is made to Rao (1952) and Reymont (1967). I am greatly indebted to many colleagues for discussions and for material. In particular I should like to mention Professor C. Radakrishna Rao, Dr. Dag Gärdefors, Professor Pierre Jolicoeur, Dr. Barbara Gilchrist, and Mr. Hans-Åke Ramdén. The figures were drawn by Mrs. Inga Thomasson.

STATISTICAL REVIEW

The statistical methods used in our analysis of sexual dimorphism are essentially the well known multivariate statistical procedures of the generalized statistical distance and principal components. The elucidation of one of the problems encountered has required particular treatment; inasmuch as this is primarily of biologic interest it does not appear to have been taken up in the statistical literature.

The generalized statistical distance

Let N_1 and N_2 be the sizes of two samples drawn from two populations, each based on p -variates. The sample means (sample mean vectors) are \bar{x}_{i1} and \bar{x}_{i2} for the first and second samples respectively. The estimated value of the covariance matrix is given by:

$$S = \left(\frac{1}{N_1 + N_2 - 2} \right) \sum_{t=1}^{N_1} (x_{i1t} - \bar{x}_{i1}) (x_{j1t} - \bar{x}_{j1}) + \sum_{t=1}^{N_2} (x_{i2t} - \bar{x}_{i2}) (x_{j2t} - \bar{x}_{j2}).$$

The standard form of the Mahalanobis' generalized statistical distance squared between the two populations, as estimated from the sample for the p characters, is:

$$D_s^2 = \bar{d}' S^{-1} \bar{d}, \quad (1)$$

where S^{-1} is the inverse of matrix S , \bar{d} is the vector of differences between the vectors of means of the two samples, and D_s^2 denotes the usual form of the generalized distance squared.

The test of the hypothesis specifying no difference in the p mean values for the two populations is carried out by means of the statistic

$$\frac{N_1 N_2 (N_1 + N_2 - p - 1)}{p(N_1 + N_2) (N_1 + N_2 - 2)} D_s^2 \quad (2)$$

which may be used as a variance ratio with p and

Table I. Changes in squared generalized distances for increasing divergence in sample sizes for seven variables and for different numbers of variables

Sample 1	Sample 2	D_s^2	T_s^2	D_a^2	D_r^2
N_1	N_2				
32	33	34.53	561.06	34.31	36.03
32	35	35.34	590.74	34.67	36.37
32	37	27.55	472.66	27.04	27.49
32	37	28.94	496.58	28.54	28.78
32	40	31.82	565.68	30.30	31.84
32	40	34.56	614.88	33.08	34.44
32	40	27.60	490.69	27.06	27.25
32	40	34.56	614.88	33.08	34.44
32	47	24.77	471.57	24.80	24.81
32	57	26.62	545.51	26.43	—
32	62	27.76	585.98	27.46	27.46
32	93	27.23	648.24	27.53	27.55
Number of variables	Sample 1	Sample 2	D^2	T^2	
3	32	40	28.41	505.06	
5	32	40	31.87	565.56	

$(N_1 + N_2 - p - 1)$ degrees of freedom. The generalized distance is connected to the Hotelling T^2 by the relationship

$$T^2 = \frac{N_1 N_2}{N_1 + N_2} D_s^2 \tag{3}$$

The Mahalanobis' distance has been widely used without detailed consideration of the effect of inequality of covariance matrices. The robustness of T^2 for unequal covariance matrices obtained from large samples has been studied by Ito and Schull (1964). Other aspects have been treated by Holloway and Dunn (1967) and Hopkins and Clay (1963).

The results of these studies show that significance level and power may be appreciably influenced by inequality in sample sizes. An example of the variation of D^2 with sample size is given in Table I.

Table II. Random sequence of 30 values of D_r^2 for male and female *Omocestus* from Kinnekulle

50.83	66.13	70.24
53.94	63.29	60.75
53.84	55.98	73.14
50.38	68.22	49.89
66.91	54.58	52.98
63.49	55.71	68.54
89.04	60.06	62.65
64.38	53.03	69.70
88.60	54.42	50.63
49.91	68.75	59.74
Average D_r^2	D_s^2	D_h^2
61.66	54.63	54.60

Table III. Values of squared generalized distances

Samples compared	D_s^2	D_h^2	D_a^2
<i>R. esculenta</i> ♀/♂ (raw)	0.037	—	0.037
<i>R. esculenta</i> ♀/♂ (log)	0.190	0.182	0.180
<i>R. temporaria</i> ♀/♂ (log)	0.887	0.928	0.902
<i>A. salina</i> ♀/♂ (log)	0.773	0.854	0.770
<i>C. maenas</i> ♀/♂ (log)	2.88	3.16	3.11
<i>O. haemorrhoidalis</i> ♀/♂ (log) (Kinnekulle)	54.63	54.96	54.83
<i>O. haemorrhoidalis</i> ♀/♂ (raw) (Kinnekulle)	60.54	60.49	60.16
<i>O. haemorrhoidalis</i> ♀/♂ (log) (Kinnekulle)	54.63	54.96	54.83
<i>O. haemorrhoidalis</i> ♀/♂ (log) (Gotland)	77.61	—	79.41
<i>O. haemorrhoidalis</i> ♀/♂ (raw) (Gotland)	—	77.31	—
<i>O. haemorrhoidalis</i> (logs) Kinnekulle ♀/Gotland ♀	1.32	—	1.33
Kinnekulle ♀/Öland ♀	2.55	—	2.53
Öland ♂/Gotland ♂	0.48	0.49	0.49
Kinnekulle ♂/Öland ♂	1.69	1.70	1.70

Anderson and Bahadur (1962) suggested a generalized distance for use in the case where the covariance matrices are not equal. In terms of sample quantities, the distance measure, here denoted D_h , is:

$$D_h = \frac{\max_t 2b'\bar{d}}{(b'S_1 b)^{\frac{1}{2}} + (b'S_2 b)^{\frac{1}{2}}} \tag{4}$$

where $b = (tS_1 + (1-t)S_2)^{-1}\bar{d}$, \bar{d} is the vector of differences of the sample means and S_1 and S_2 are the respective sample covariance matrices. When the population covariance matrices $\Sigma_1 = \Sigma_2$, expression (4) becomes equal to the normal form of the distance. The method of finding D_h is therefore iterative and, without the aid of an electronic computer, must be regarded as prohibitively arduous.

D_h^2 has not yet been related to a test of significance. Therefore if it is desired to carry out an Hotelling's T^2 -test, and it has been ascertained that D_h^2 and D_s^2 are very different, then the significance may be checked by using the Bennet (1951) solution of the generalized Behrens-Fischer problem (cf. Anderson, 1958). The generalized distance squared statistic proposed by the writer (Reyment (1962)), herereferred to as D_r^2 , is defined as

$$D_r^2 = 2\bar{d}' S_r^{-1} \bar{d}, \tag{5}$$

where S_r is the sample covariance matrix of differences obtained from the random pairing of the two samples. Since $N_1 D_r^2 / 2 = T^2$ ($N_1 < N_2$) is an Hotelling

T^2 with p and $(N_1 - p)$ degrees of freedom, the distribution of D_r^2 is known. This squared generalized distance suffers from the disadvantage that in order to obtain a usable value, it is necessary to average the results obtained from a number of permutations. This is naturally only feasible where the calculations are made on a computer. An example of a random sequence of 30 values of D_r^2 is shown in Table II.

An approximate means of producing a squared generalized distance D_a^2 for heterogeneous covariance matrices is by simply averaging the two covariance matrices.

Thus,

$$D_a^2 = \bar{d}' S_a^{-1} \bar{d},$$

where \bar{d} has the same meaning as before,

$$S_a = \frac{1}{2} (S_1 + S_2), \tag{6}$$

and S_1 and S_2 are the sample covariance matrices. The value of D_a^2 is very close to D_h^2 and D_s^2 and for most purposes offers a perfectly adequate solution to the question of a generalized distance for populations with unequal covariance matrices. Comparisons of some squared generalized distances obtained in this study are given in Table III. D_a^2 is clearly always smaller than or sometimes equal to D_h^2 .

The equality of covariance matrices

The generalized test for homogeneity of covariance matrices (Kullback, 1959, p. 317) is:

$$2I(H_1: H_2(*)) = N_1 \log_e (\det S / \det S_1) + N_2 \log_e (\det S / \det S_2) \tag{7}$$

where S_1 and S_2 have the same meaning as in (4), N_1 and N_2 are the corresponding degrees of freedom, "det" denotes the determinant and $NS = N_1 S_1 + N_2 S_2$ ($N = N_1 + N_2$). This is the form of the test used in the computer program employed in this study. It is approximately distributed as χ^2 with $k(k+1)/2$ degrees of freedom, where k is the number of variables. A result indicative of heterogeneity in covariance matrices yielded by this procedure may be sufficient for many statistical purposes. Generally, however, a biological study of the kind considered in this paper requires further information on the nature of the heterogeneity. The first point of interest concerns the range of variability in each sample as represented in the covariance matrices. Conceptually, this situation may be thought of in terms of

inflation of the ellipsoids of scatter and is thus an expression of the degree of multivariate variability in the samples relative to each other. The second point of interest concerns the relative orientations of the ellipsoids of scatter and the connotations hereof in growth interpretation (cf. Jolicoeur and Mosimann 1960).

An approximate test of the hypothesis that a given eigenvector is the i th eigenvector of a very large sample estimate of a covariance matrix may be produced by adopting a procedure suggested by Anderson (1963) (cf. Reymont, 1967, p. 3 and Reymont, 1969).

One computes

$$n(d_i b_i' S_1^{-1} b_i + (1/d_i) b_i' S_1 b_i - 2) \tag{8}$$

where $n+1$ is the sample size of covariance matrix S_1 , d_i is the i th sample eigenvalue of S_1 , and the vector b_i is the i th eigenvector of sample covariance matrix S_2 , based on a very large sample. The criterion is distributed approximately as chisquare with $(p-1)$ degrees of freedom, where p is the number of dimensions involved.

Principal components

This statistical procedure (PCA) is designed for the analysis of the structure of multivariate observations. PCA is useful for investigating the dependence structure occurring in a suite of observations, particularly when no *a priori* patterns of interrelationship are available. In PCA, the observable variates, such as morphologic dimensions, are represented as functions of a smaller number of latent variates (the principal components). The principal components are expressed in terms of linear combinations of the observable variates and the analysis of the dependence structure amounts to the statistical estimation of the coefficients of these linear functions. The first principal component of the observations X is defined as the p -variate linear compound

$$Y_1 = a_{11} X_1 + \dots + a_{p1} X_p = a_1' X$$

with sample variance $s_{Y_1}^2 = a_1' S a_1$.

The coefficients c_1 of the first linear combination of PCA must satisfy the p simultaneous linear equations

$$(S - d_1 I) a_1 = 0,$$

where S is the sample covariance matrix, I is the

Table IV. Squared generalized distance values

Species	Inflation heterogeneity	Orientation heterogeneity	Approximate D_a^2	Heterogeneous D_h^2	Standard D_s^2	$D_h^2 - D_s^2$	Reyment's D_r^2
<i>Carcinus maenas</i>	+	+	3.11	3.16	2.60	0.56 (17.7)*	2.60
<i>Artemia salina</i>	+	+	0.77	0.85	0.77	0.08 (9.4)	0.88
<i>Rana esculenta</i>	+	+	0.180	0.182	0.190	0.08 (44.0)	0.30
<i>Rana temporaria</i>	-	+	0.902	0.928	0.887	0.041 (4.4)	1.12
<i>Omocestus haemorrhoidalis</i>							
Kinneulle	+	+	54.83	54.96	54.63	0.33 (0.6)	59.67
Kinneulle-Gotland ♂♂	+	+	0.49	0.49	0.48	0.01 (2.0)	0.55
Öland-Kinneulle ♂♂	+	+	1.70	1.70	1.69	0.01 (0.6)	1.72
<i>Chrysemys picta marginata</i> (raw data)	+	+	5.56	6.66	5.56	0.10 (1.5)	4.88

* Percentages within parentheses.

identity matrix and d_1 is obtained by solution of the determinantal equation

$$|S - d_1 I| = 0.$$

The vector a_1 is the first eigenvector of S and d_1 its first eigenvalue. In general terms, the j th principal component of S is the linear expression

$$Y_j = a_{1j}X_1 + \dots + a_{pj}X_p,$$

where the coefficients are the elements of the eigenvector corresponding to the j th largest eigenvalue.

The sign and magnitude of a vector element indicate the direction and importance of the contribution of a variable to a particular component. More precisely, the ratio $a_{ij} \sqrt{d_j} / s_i$ yields the correlation of the i th variable and the j th component; here, s_i is the pertinent standard deviation. PCA is in this paper used in the treatment of the *Omocestus* material.

Some empirical results for squared generalized distances

As clearly brought out in the present analysis, significant differences in homogeneity of covariance matrices are of quite common occurrence. Theoretically, the approach by means of the classical D^2 is not permissible and some other way should be sought. However, the differences between the generalized distance obtained by a method acceptable to the theory and that found by ignoring such considerations are mostly so small as to make the more difficult and time-consuming calculations for the former unnecessary. Sets of comparisons between

the three methods of calculation of D^2 outlined in the section on methods are presented in Table IV.

The heterogeneity in covariance matrices for *Carcinus maenas* is exceptionally strong with respect to both the much greater inflation of the scatter ellipsoid of males as well as to the starkly different orientations of the first three significant major axes of the ellipsoids. The difference in D_h^2 and D_s^2 is 17.7%, where the percentage is given by

$$\left[\frac{(D_h^2 - D_s^2) 100}{D_h^2} \right].$$

The heterogeneity in covariance matrices for *Artemia salina* is very strong with respect to inflation as also to orientation of the scatter ellipsoids. These ellipsoids are of a particularly elongated shape, with the first principal axis several hundred times the length of the second principal axis. The difference in D_h^2 and D_s^2 is 9.4%, which, although large, is not

Table V. Analysis of length (x_1) and breadth (x_2) of skull of *Rana esculenta* L. for sexual dimorphism

	Females (N=35)		Males (N=14)	
	x_1	x_2	x_1	x_2
Covariance matrix	17.683	20.290 24.407	18.479	19.095 20.755
Mean vector	22.860	24.397	21.821	22.843
	D^2	T^2	F	
	0.197	1.969	0.96	

difference between D_s^2 and D_h^2 . Here, the lowest estimate of the generalized distance squared is yielded by D_r^2 and the highest estimate by D_h^2 .

THE CASE STUDIES

Rana esculenta and *R. temporaria*

The material forming the basis of this analysis derives from localities in Europe and North Africa. It was extracted from Kauri (1959). The measurements made by Kauri were: total length of body (x_1), length of tibia (x_2) and the length of the praeallux (x_3). The body length was measured from the tip of the snout to the posterior end of the os coccygis, the breadth of the head was measured across the joint of the jaw and the length of the head was taken from the tip of the snout to the occipital border. The value of the praeallux as a diagnostic and useful variable is open to some doubt, as its size is largely determined by the age of the animal (Kauri, 1959, p. 138). For some of the material the length and breadth of the cranium were measured. These variables are only available on a very limited number of individuals and they do not appear in the main analysis. A bivariate treatment is given in Table V. Here, the value of the variance ratio is not significant.

Rana esculenta. The results of the analysis of the pooled Swedish material of *R. esculenta* are given in Table VI. The material comprises 115 females and 70 males, all fully grown; the calculations were made on the logarithmically transformed observations. The tridimensional ellipsoids of scatter are significantly different and the first two significant major axes are differently oriented. In neither case are the differences exceptionally strong in geometric terms, although the formal statistical significance is high. The differences in the means are slight, with males slightly larger than females. The generalized distance is not significant and it may be concluded that there is no statistically important sexual dimorphism in the variables studied.

The elements of the first two eigenvectors indicate that the patterns of growth for males and females are different. There is, however, close agreement in the elements of the third eigenvector, and it is suggested, that this may represent some morphologic factor unaffected by sexual differences of size and shape. The praeallux is without importance in this connexion.

Table VIII. Distance statistics for males and females of *Rana esculenta* and *R. temporaria*

Origin of sample	Total sample size	Number of variables	D^2
<i>Rana esculenta</i>			
Asia Minor	42	3	0.184
Asia Minor	30	6	0.818
Balkans	36	3	0.041
North Africa	29	3	1.401
Vienna	49	3	0.296
<i>Rana temporaria</i>			
Southern Sweden	73	3	1.317
Vienna	47	3	3.023

Rana temporaria. The study material of this species is made up of 91 females and 82 males from northern and southern Sweden. Here, there is an interesting difference from *R. esculenta*; the major axes of the scatter ellipsoids are identically oriented while the ellipsoids are significantly differently inflated. This may be readily appreciated from inspection of the covariance matrices in Table VII; the variance ratios for the variances of x_1 , x_2 and x_3 are all significant. The results of the analysis of *R. temporaria* are shown in Table VII. In contrast to *R. esculenta*, the certainly slight sexual dimorphic differences in the mean vector are significantly different.

Geographic variation in *Rana*

Variation in the strength of dimorphism. Table VIII summarizes the results obtained for various samples of *R. esculenta*, respectively *R. temporaria*. Inasmuch as all sample sizes are small for morphometric work, these values can only be accepted as very approximate. The results of comparisons between samples of the same sex from different geographic locations are shown in Table IX. For purposes of comparison

Table IX. Squared generalized distances between samples of the same sex for *Rana esculenta* and *R. temporaria* for three variables

Localities	Total sample size	D^2
<i>Rana esculenta</i>		
Balkans: Asia ♂♂	57	1.40
North Africa: Spain ♂♂	25	0.63
<i>Rana temporaria</i>		
Southern Sweden: Northern Sweden ♀♀	63	0.068

Table X. Results of calculations for males and females of *Martes americana* Turton (logarithmic data)

x_1	x_2	x_3	x_4	Eigenvalues	Eigenvectors			
					I	II	III	IV
<i>Covariance matrix of males (N=92)</i>								
1.1544	0.9109	1.0330	0.7993	4.5482	0.4121	-0.5208	-0.1484	0.7327
	2.0381	0.7056	1.4083	1.1163	0.5846	0.4025	-0.6809	-0.1806
		1.2100	0.7958	0.6447	0.3894	-0.6411	0.1275	-0.6489
			2.0277	0.1209	0.5803	0.3947	0.7058	0.0971
<i>Covariance matrix of females (N=47)</i>								
0.9617	0.2806	0.9841	0.6775	3.7749	0.3520	-0.5025	-0.2166	0.7594
	1.8475	0.3129	1.2960	1.6047	0.5474	0.6091	-0.5738	-0.0144
		1.2804	0.7923	0.3679	0.4104	-0.5977	-0.2266	-0.6504
			1.7819	0.1240	0.6388	0.1390	0.7567	0.0117
<i>Orientation of ellipsoids</i>								
B^2	β^2	d.f.	D_s^2	T_s^2	$F_{4/134}$	Vector	Chisquare	
15.83	0.3645	10	0.0045	0.14	0.03	1	6.66	
						2	13.16***	
						3	16.67***	
						4	7.48	

with these results, the squared generalized distance between 28 females of *R. temporaria* and 35 females of *R. esculenta* from Vienna was calculated; it is 3.37, which is highly significant.

Martes americana Turton

The data underlying the analysis accounted for here derive from Jolicoeur (1963). The study material was obtained from Montana and consists of adults and subadults. The variables are: total length of humerus = x_1 , maximum epicondylar width of the distal end of the humerus = x_2 , total length of the femur = x_3 ,

maximum width of the distal end of the femur = x_4 . These were measured on the right side of the animals. The results of the calculations are given in Table X.

The covariance matrices are homogeneous as regards degree of inflation, but the second and third principal axes differ significantly in orientation. The value of the squared generalized distance is very small indeed and on the basis of the variables included in this analysis one cannot claim significant sexual dimorphism to occur, as expressed by the squared generalized distance. However, all means of males are slightly greater than those of females

Table XI. Results of calculations for *Chrysemys* (raw data)

x_1	x_2	x_3	Eigenvalues	Eigenvectors			
				I	II	III	
<i>Covariance matrix of females</i>							
138.77	79.15	37.38	195.27	0.8401	0.4881	-0.2365	
	50.04	21.65	3.69	0.4919	-0.8694	-0.0469	
		11.26	1.10	0.2285	0.0770	0.9705	
<i>Covariance matrix of males</i>							
451.39	271.17	168.70	680.41	0.8126	0.5454	-0.2054	
	171.73	103.29	6.50	0.4955	-0.8321	-0.2491	
		66.65	2.86	0.3067	-0.1006	0.9465	
<i>Orientation of ellipsoids</i>							
B^2	β^2	d.f.	D_h^2	T_r^2	$F_{3/23}$	Vector	Chisquare
25.86	0.4239	6	6.66	58.52	17.81***	1	36.77***

Table XII. Results of calculations for *Chrysemys* (logarithmic data)

x_1	x_2	x_3	Eigenvalues ($\times 10^{-4}$)	Eigenvectors			
				I	II	III	
<i>Covariance matrix of females</i> ($\times 10^{-4}$)							
20.88	15.13	15.39	43.953	0.6831	-0.1595	0.7127	
	12.10	11.33	1.128	0.5102	-0.5940	-0.6220	
		12.77	0.679	0.5225	0.7885	-0.3244	
<i>Covariance matrix of males</i> ($\times 10^{-4}$)							
49.78	37.96	47.99	126.746	0.6223	-0.4552	0.6369	
	30.54	37.31	1.415	0.4841	-0.4156	-0.7700	
		48.85	1.000	0.6152	0.7874	-0.0383	
<i>Orientation of ellipsoids</i>							
B^2	β^2	<i>df.</i>	D_s^2	T_s^2	$F_{3/44}$	Vector	Chisquare
14.83	0.4239	6	6.47	77.69	24.77***	1	27.60***
						2	20.61***

(Jolicoeur, 1963, p. 5) and there is no overlap between individual measurements of males and females (Jolicoeur, 1963, Fig. 1).

Hence, the smallest male in the material is larger than the greatest female. Nevertheless, the differences are very slight, even for each variable regarded on its own. The values of *t* are for x_1 , x_2 , x_3 and x_4 , 0.29, 0.30, 0.30, respectively, 0.04. These are very far from significance.

Chrysemys picta marginata Agassiz

The data on the Midland Painted Turtle were obtained from Jolicoeur & Mosimann (1960). They derive from a sample drawn from a single local population. The analysis here accounted for is of particular interest inasmuch as it provides an example of a case where the raw data produce covariance matrices, heterogeneous with respect to each other (Jolicoeur and Mosimann did their principal component analyses on the raw data) while the logarithmically transformed data result in homogeneous covariance matrices.

We shall first consider the logarithmically transformed data. The results of these computations are displayed in Table XII, where the variables are: length of carapace = x_1 , maximum width of carapace = x_2 , maximum height of carapace = x_3 .

For the first set of calculations, the value of B^2 is fairly high, although not significant. The first two principal axes are not parallel, and there is consequently heterogeneity in the orientation of the covariance ellipsoids. The computations for the raw

data are presented in Table XI. In this case, the inflations of the covariance matrices are sufficiently different as to produce a significant value of B^2 (the B^2 of Table XI is almost twice that of Table XII). The effect of the logarithmic transformation of the data is clearly to bring the scatter ellipsoids nearer to each other in size, and in so doing to reduce the extent of inflation heterogeneity. The effect of the logarithmic transformation on the squared generalized distance value is slight; in both cases it is significant.

The foregoing opens up the question of the correct strategy for generalized distance studies of sexual dimorphism. It seems that data of the kind represented by the material of *Artemia*, *Omocestus*, *Rana*, *Martes* and *Carcinus* are best treated via the logarithmic transformation as the variables are not highly compatible. The type of data represented by *Chrysemys*, where the measurements are made on a single structural feature, should not be transformed, as this may tend to obscure the manifestations of dimorphism rather than unveil them.

The reasons for introducing the logarithmic transformation in biologic data are twofold. Firstly, one wishes to bring deviating data as near to multivariate normality as possible and, secondly, variables selected among interlocked growth patterns run a good chance of being bound in some non-linear growth relationships. Although the multivariate calculations may be made on non-linear, non-normal data, the tests of significance are based on the assumption of normality. Where variables may be re-

Table XV. Eigenvalues and eigenvectors for males of *Carcinus maenas* L.

	Eigenvalues	Percentage of total variation	Eigenvectors							
			I	II	III	IV	V	VI	VII	VIII
1	0.1267	67.887	0.3056	0.1717	0.2238	-0.1745	0.0918	-0.1118	-0.0842	-0.8767
2	0.0490	26.283	0.3236	0.1917	0.3151	-0.6168	0.4286	-0.0403	0.2164	0.3828
3	0.0053	2.830	0.2862	0.1651	0.1290	-0.0961	-0.8012	-0.3710	0.2656	0.1221
4	0.0030	1.604	0.2748	0.1465	0.0769	0.6467	0.3803	-0.5505	0.1359	0.1124
5	0.0011	0.574	0.3207	0.2458	-0.1252	-0.0390	-0.0774	-0.0620	-0.8727	0.2193
6	0.0010	0.509	0.3330	0.3144	-0.8122	-0.0481	0.0621	0.1823	0.2949	-0.0651
7	0.0004	0.236	0.3638	0.1876	0.3776	0.3965	-0.1072	0.7127	0.0911	0.0701
8	0.0001	0.076	0.5461	-0.8310	-0.1010	-0.0237	0.0021	0.0084	-0.0167	0.0073

Table XVI. Eigenvalues and eigenvectors for females of *Carcinus maenas* L.

	Eigenvalues	Percentage of total variation	Eigenvectors							
			I	II	III	IV	V	VI	VII	VIII
1	0.0848	84.127	0.3582	-0.0017	-0.0191	0.1057	0.1819	-0.5316	0.3269	0.6615
2	0.0101	10.056	0.3409	0.0032	-0.0151	0.1860	0.1957	-0.4266	0.2708	-0.7452
3	0.0041	4.084	0.2641	0.0369	0.9234	-0.2657	-0.0266	0.0301	-0.0630	-0.0111
4	0.0008	0.773	0.3103	0.0305	0.1291	0.8334	0.0623	0.3964	-0.1550	0.0806
5	0.0005	0.521	0.3930	-0.0800	-0.1416	-0.1436	-0.5976	0.3614	0.5573	-0.0148
6	0.0003	0.251	0.3994	-0.7813	-0.1641	-0.1998	0.1336	0.0523	-0.3776	0.0001
7	0.0002	0.160	0.3529	0.4033	-0.1734	-0.0432	-0.4981	-0.3070	-0.5816	-0.0105
8	0.0000	0.028	0.3885	0.4671	-0.2302	-0.3558	0.5486	0.3870	-0.0195	0.0110

Table XVII. Orientation, homogeneity and distance statistics for males and females of *Carcinus maenas* L. from Arcachon

B^2	β^2	d.f.	Orientation of ellipsoids (χ^2 value)			D_h^2	T_r^2	$F_{8/20}$
			vector 1	vector 2	vector 3			
171.02	5.537	36	118.26***	279.18***	120.85***	3.16	27.30	2.22

Table XVIII. Correlations of principal components with original variables for *Carcinus maenas* L.

	Components for males			Components for females		
	I	II	III	I	II	III
x_1	0.9271	0.3241	0.1386	0.9941	-0.0016	-0.0117
x_2	0.8841	0.3259	0.1758	0.9939	0.0032	-0.0097
x_3	0.9046	0.3247	0.0833	0.7892	0.0382	0.6081
x_4	0.8800	0.2918	0.0503	0.9616	0.0327	0.0882
x_5	0.8905	0.4246	-0.0710	0.9838	-0.0692	-0.0781
x_6	0.7908	0.4646	-0.3938	0.8247	-0.5577	-0.0746
x_7	0.9107	0.2922	0.1930	0.9176	0.3625	-0.0995
x_8	0.7259	-0.6873	-0.0274	0.9080	0.3775	-0.1185

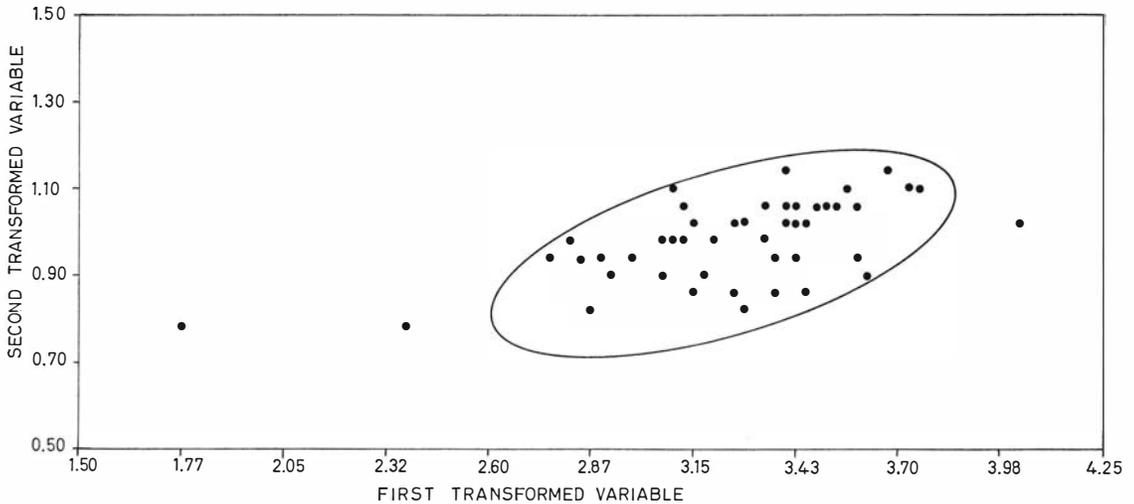


Fig. 1. Plot of the first two PCA transformed variables for males of *Carcinus*.

different from zero apart from that between variables 6 and 8 for males. The high standard deviation for the maximum breadth of left claws of males is surprising.

Rather interestingly, the sample difference mean vector $\{0.0039, 0.0108, -0.0112, 0.0220, -0.0314, -0.0404, -0.0324, -0.0099\}$, shows the females are slightly larger than males, on the average, with respect to three variables, notably, x_1 , x_2 and x_4 . This indicates, that the carapace dimensions maximum breadth, maximum height and posterior width of carapace tend to be greater in females, whereas the line of the maximum width is higher up in males. The cheliped dimensions are greater in males.

Analysis of Dimorphism. The eigenvalues and eigenvectors for the covariance matrices of males and females are given in Tables XV and XVI. The plots of the first transformed variables are shown in Fig. 1 (males) and Fig. 2 (females). The plot for females comprises too few observations for conclusions concerning the shape of the scatter cloud; that for the males is clearly approximately ellipsoidal in form. This is what obtains for multivariate normally distributed variables.

The values of the various homogeneity and distance statistics are given in Table XVII. Although the generalized distance is large, it does not attain the level of significance of the variance ratio owing, mainly, to the small sample sizes. Three major axes are sufficiently long to be of importance. Their strongly different orientations indicate pronouncedly

different patterns of growth to occur in males and females. The ellipsoid of scatter for males is several times greater than that of the females, thus indicating a possible wider range of variation for males.

The correlations of the 8 variables with the first three principal components are given in Table XVIII. For both males and females, all variables are strongly to very strongly correlated with the first component. The second component for males is strongly negatively correlated with x_8 and weakly with several others. For females, the relationships with the second component are different and there are no significant correlations with the third component. For males, the third component is significantly negatively correlated with x_6 .

Artemia salina L.

The brine shrimp, *Artemia salina*, has long been known to be particularly strongly affected morphologically by the concentration of salts occurring in the water in which it is living. Gilchrist (1960) summarized several years of work on the reaction of the morphology of *Artemia* to salinities of 35‰ and 140‰ under standard conditions of temperature and feeding. She found that the extent to which growth is influenced by the salinity of the medium varies with the sex of the animal and the stock from which it is derived. Changes in body proportions influenced by salinity were analysed and it was concluded, that the intrinsic factors of sex and genetic constitution modify the effects of external salinity. Moreover, brine shrimps from different localities

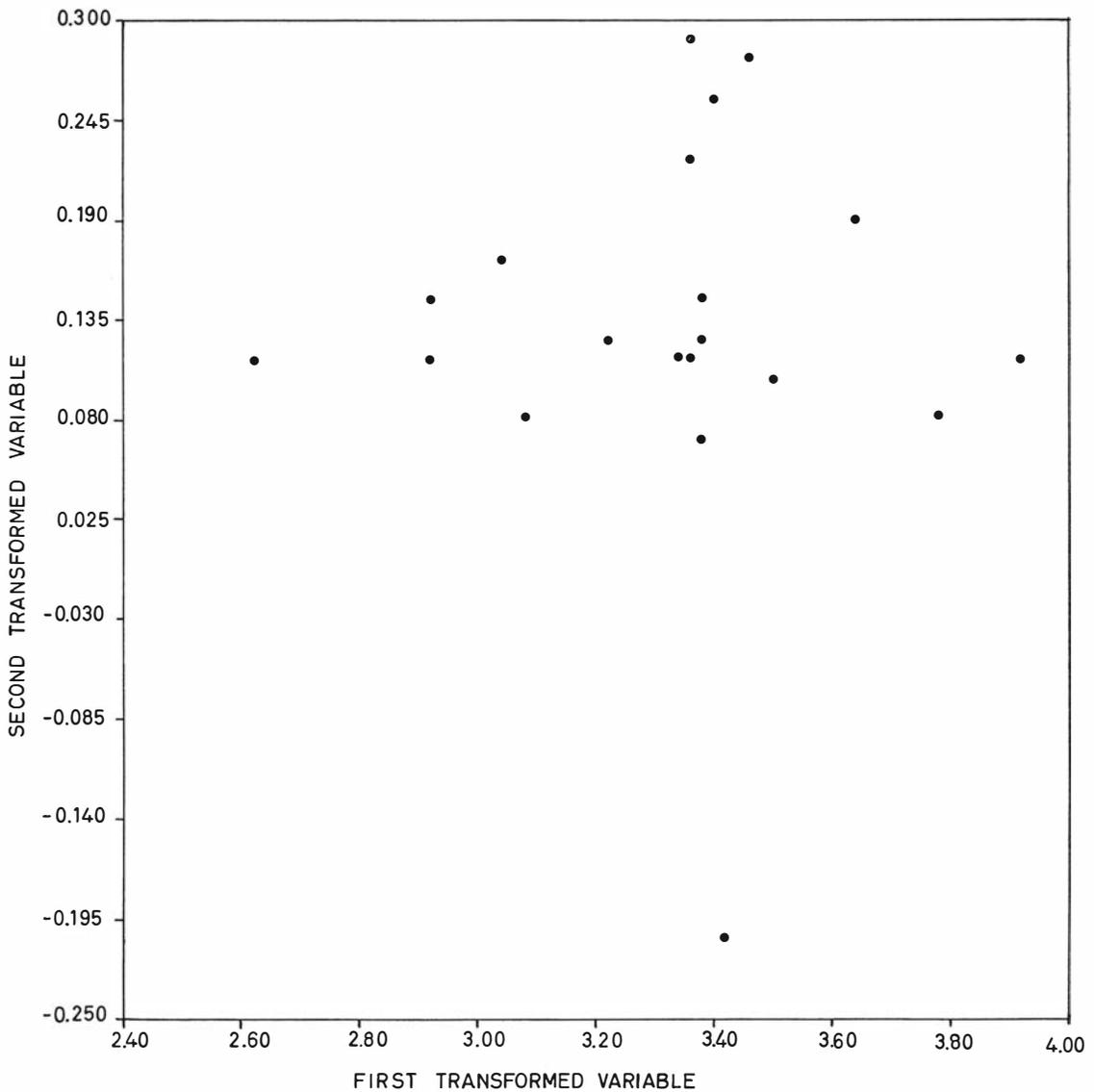


Fig. 2. Plot of the first two PCA transformed variables for females of *Carcinus*.

are morphologically distinct. This was verified by a canonical variate analysis by the writer (Reyment, 1965). The variables chosen here for analysis are length of body (x_1), length of abdomen (x_2), and length of prosoma (x_3).

The basic statistics are shown in Table XIX. The material underlying the analysis here accounted for was reared in a salinity of 140‰. The eigenvalues and eigenvectors of the covariance matrices are shown in Table XX. The differences in growth patterns, as represented by the elements of the eigenvectors, are slight but real.

The results for the squared distance and homogeneity computations are given in Table XXI. Although the generalized distance squared is small, this indicating that there is little size difference between males and females, it is highly significant.

Omocestus haemorrhoidalis L.

Introduction. The study material of *Omocestus haemorrhoidalis* was made available by fil.lic. Dag Gärdefors, Uppsala. The particular value of this material lies partly in the fundamental issues in-

Table XIX. Basic statistics for males and females of *Artemia salina* L. (logarithmic values)

	Females (N=139)			Males (N=136)		
	x_1	x_2	x_3	x_1	x_2	x_3
Mean vectors	0.8030	0.5233	0.4789	0.7662	0.4795	0.4503
Covariance matrices	0.00284	0.00323 0.00368	0.00241 0.00275 0.00206	0.00372	0.00463 0.00577	0.00272 0.00336 0.00200

Table XX. Eigenvalues and eigenvectors of the covariance matrices for males and females of *Artemia salina* L.

Females					Males				
Eigenvalues	Percentage of total variation	Eigenvectors			Eigenvalues	Percentage of total variation	Eigenvectors		
		I	II	III			I	II	III
0.008571	99.95	0.5751	-0.0608	-0.8160	0.011459	99.755	0.5700	-0.0725	-0.8185
0.000004	0.05	0.6554	0.6311	0.4150	0.000028	0.244	0.7090	0.5470	0.4450
+0.000000	0.00	0.4896	-0.7733	-0.4028	+0.000000	0.001	0.4155	-0.8340	0.3631

volved and partly in the careful statistically based collecting procedures employed by Gärdefors.

In Sweden *O. haemorrhoidalis* is only known to occur in three areas, notably, the hill of Kinnekulle, Västergötland, and the islands of Öland and Gotland. The reason for this profound degree of isolation has not yet been satisfactorily explained (in more southerly Europe the distribution is largely continuous). A seemingly plausible case might be argued for the calcareous nature of much of the bedrock of these regions, which, combined with latitude, might constitute a decisive distributional contollant. However, despite numerous searches by Dr. Gärdefors, *O. haemorrhoidalis* has not been found outside the immediate vicinity of Kinnekulle even although the same strata occur in several other places in Västergötland. Each area was divided into six sampling

localities, each of which comprised a collecting zone made up of a circle with a diameter of 40 m.

Canonical variate analysis. The 36 samples of males and females from the three areas were subjected to canonical variate analysis. A preliminary univariate analysis of variance by Dr. Gärdefors had shown that important differences were to be expected and also that differences within areas may exceed differences between areas. The basic statistics for the 36 samples are given in Table XXII. It should be noted that the theoretical requirements for the applicability of canonical variates are not exactly met in the material. As will become apparent further on in this section, the material lacks complete homogeneity in the variances and covariances; however, this is not of such an order as to constitute a troublesome source of inadequacy in the analysis. The variables on which our analyses are based are: length of hind femur (x_1), pronotal length (x_2), elytron length (x_3) and the least width between the pronotus ridges (x_4). *Omocestus* does not produce a swarming phase in Sweden.

The detailed analyses through the generalized distance may be given an approximate diagrammatic interpretation by means of the canonical variates, based on the 36 samples. The plot of the first two

Table XXI. Generalized distances and homogeneity values for *Artemia salina*

B^2	β^2	d.f.	Orientation of ellipsoids (χ^2)	D_h^2	T_r^2	$F_{3/153}$
<i>Vector I</i>						
582.97	0.0714	6	484.84	0.85	59.47	19.53***

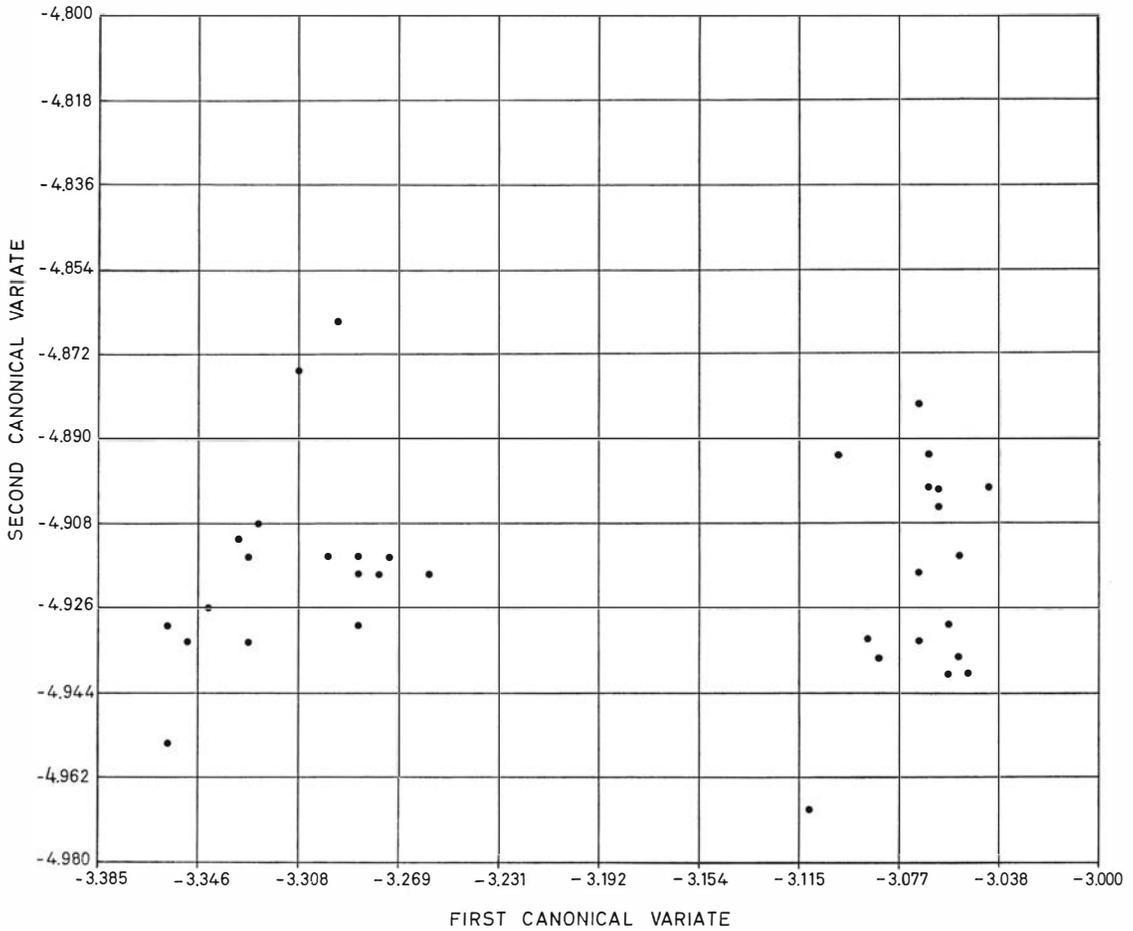


Fig. 3. Plot of first two canonical variate means for *Omocestus*.

canonical variate means is shown in Fig. 3. In this figure, the clear subdivision into male and female groups is clearly manifested. This analysis is an approximate appraisal of the interrelationships between the samples owing to the existence of heterogeneity in the variances and covariances.

Generalized distance analysis. The generalized distance analysis was designed to examine the strength of sexual dimorphism in the three populations and the differences in multivariate means between the populations. The latter aim was achieved by studying the distances between males, respectively females, from the three areas. The analyses are displayed in Tables XXII, XXIII, XXIV.

The generalized distances between males and females of the samples from the three populations are, in summary:

	Kinneulle	Öland	Gotland
	54.96	85.92	77.61

These are exceptionally high values and confirm the particularly strong sexual dimorphism in this species of grasshoppers. The sample sizes are large and these distances therefore reasonably stable. There would therefore appear to be populational differences in the strength of sexual dimorphism.

In addition to the dimorphism of size there is also the dimorphism of growth differences. This is strong for the population of Kinneulle, where significant differences in orientation occur in all the principal axes of the scatter ellipsoids and these ellipsoids are significantly differently inflated. The growth differences for the population of Öland are much less than those of the Kinneulle population and are confined to significant differences in the orientation

Table XXII. Generalized distance analysis for males ($N = 110$) and females ($N = 101$) of *Omocestus haemorrhoidalis* L. from Kinnekulle

x_1 x_2 x_3 x_4				Eigenvalues ($\times 10^{-4}$)	Eigenvectors			
Covariance ($\times 10^{-4}$) and correlation matrices for ♂♂					Males			
					I	II	III	IV
2.28	2.31	2.14	1.69	13.836	0.2892	-0.0753	-0.2571	0.9190
0.5099	9.01	2.91	1.91	7.515	0.6245	-0.6117	0.4718	-0.1147
0.6318	0.4319	5.04	2.07	3.522	0.4171	-0.1358	-0.8183	-0.3714
0.3621	0.2059	0.2985	9.55	1.006	0.5936	0.7757	0.2039	-0.0662
Covariance ($\times 10^{-4}$) and correlation matrices for ♀♀					Females			
					I	II	III	IV
4.13	3.04	3.74	1.86	17.635	0.2877	-0.4055	0.2499	0.8308
0.6862	4.76	3.37	0.95	10.003	0.2421	-0.4752	0.6694	-0.5172
0.6721	0.5641	7.49	3.11	2.519	0.4369	-0.5344	-0.6944	-0.2032
0.2359	0.1118	0.2933	15.04	1.252	0.8171	0.5693	0.0850	-0.0306
Difference mean vector								
-0.00735	0.10852	0.10004	0.07796	Orientation of ellipsoids				
B^2	β^2	<i>d.f.</i>	D_h^2	T_r^2	$F_{100/4}$	Vector		Chisquare
47.87***	0.2063	10	54.96	3013.27	730.72***	1	78.08***	
						2	62.61***	
						3	58.55***	
						4	24.26***	

Table XXIII. Generalized distance analysis for males ($N = 120$) and females ($N = 117$) of *Omocestus haemorrhoidalis* L. from Öland

x_1 x_2 x_3 x_4				Eigenvalues ($\times 10^{-4}$)	Eigenvectors			
Covariance ($\times 10^{-4}$) and correlation matrices for ♂♂					Males			
					I	II	III	IV
5.04	3.78	3.84	2.13	17.222	0.3907	-0.4390	-0.1963	0.7849
0.7292	5.34	3.40	3.06	8.030	0.4269	-0.3271	-0.6351	-0.5544
0.6886	0.5934	6.17	2.78	2.366	0.4398	-0.4213	0.7465	-0.2679
0.2700	0.3772	0.3192	12.33	1.245	0.6867	0.7230	0.0284	0.0697
Covariance ($\times 10^{-4}$) and correlation matrices for ♀♀					Females			
					I	II	III	IV
8.03	6.35	7.23	4.00	24.957	0.5185	-0.2783	-0.1138	0.8005
0.8300	7.28	5.96	3.61	7.953	0.4655	-0.2491	-0.6957	-0.4870
0.8310	0.7197	9.42	4.39	2.320	0.5500	-0.2697	0.7091	-0.3492
0.4156	0.3938	0.4205	11.56	1.059	0.4604	0.8875	-0.0155	0.0082
Difference mean vector								
0.00923	0.01981	0.12203	0.09839	Orientation of ellipsoids				
B^2	β^2	<i>d.f.</i>	D_h^2	T_r^2	$F_{232/4}$	Vector		Chisquare
11.83	0.1830	10	85.92	5090.05	1256.27***	1	19.33***	
						2	14.46**	
						3	1.87	
						4	5.60	

Table XXIV. Generalized distance analysis for males (N = 128) and females (N = 95) of *Omocestus haemorrhoidalis* L. from Gotland

x_1 x_2 x_3 x_4				Eigenvectors				
Covariance ($\times 10^{-4}$) and correlation matrices for ♂♂				Eigenvalues ($\times 10^{-4}$)	Males			
					I	II	III	IV
3.29	2.74	2.26	1.88	15.912	0.2432	0.3926	0.3825	0.8002
0.6844	4.88	3.03	1.69	8.222	0.2759	0.5789	0.4811	-0.5978
0.5535	0.6108	5.05	1.99	2.100	0.2893	0.5418	-0.7888	0.0234
0.2746	0.2033	0.2343	14.22	1.210	0.8838	-0.4661	0.0028	-0.0413
Covariance ($\times 10^{-4}$) and correlation matrices for ♀♀				Eigenvalues ($\times 10^{-4}$)	Females			
					I	II	III	IV
5.79	5.02	4.68	2.33	19.050	0.4501	0.3578	0.0567	0.8162
0.7887	7.01	3.99	2.30	9.657	0.4611	0.3932	0.6409	-0.4712
0.7334	0.5674	7.04	3.16	3.005	0.4843	-0.2729	-0.7613	-0.3338
0.2700	0.2418	0.3311	12.90	1.032	0.5918	-0.8018	0.0805	0.0195
Difference mean vector				Orientation of ellipsoids				
-0.00163	0.11308	0.11614	0.09176		$F_{218/4}$	Vector	Chisquare	
B^2	β^2	d.f.	D_h^2	T_r^2				
22.08	0.2005	10	77.61	4232.27	1043.70***	1	9.51**	
						2	23.25***	
						3	16.31***	
						4	29.20***	

Table XXV. Generalized distance analysis for pairs of the same sex from different areas

Pair	B^2	β^2	D_h^2	D_s^2	Variance ratio	Orientation of ellipsoids			
						$\chi^2_{(1)}$	$\chi^2_{(2)}$	$\chi^2_{(3)}$	$\chi^2_{(4)}$
Öland-Gotland ♀♀	8.04	0.2078		0.62	8.04*** (4/207)	3.60	2.58	2.31	1.31
Kinneulle-Gotland ♀♀	10.76	0.2219		1.32	15.92*** (4/191)	9.22	11.86	10.88	5.46
Öland-Kinneulle ♀♀	21.62	0.2005		2.55	34.08*** (4/213)	17.23***	10.88	5.57	8.03
Öland-Gotland ♂♂	16.65	0.1750		0.59	9.09*** (4/243)	8.42	26.06***	3.16	26.25***
Kinneulle-Gotland ♂♂	47.69	0.1836	0.49		7.40*** (4/109)	9.82*	25.99***	72.92***	129.13***
Öland-Kinneulle ♂♂	57.53	0.1890	1.70		23.00*** (4/109)	18.17***	54.53***	40.10***	141.73***

Table XXVI. Eigenvalues of *Omocestus* covariance matrices, expressed as percentages

Locality	Sex	Eigenvalues as percentages			
		First	Second	Third	Fourth
Kinneulle	♀♀	56.147	31.849	8.019	3.985
Öland	♀♀	68.773	21.915	6.394	2.918
Gotland	♀♀	58.179	29.492	9.177	3.152
Kinneulle	♂♂	53.463	29.039	13.610	3.888
Öland	♂♂	59.667	27.821	8.199	4.313
Gotland	♂♂	57.980	29.959	7.652	4.409

represented by significant differences in orientation in all principal axes of the scatter ellipsoids.

The distances squared between pairs of the same sex are very much smaller than the sexual dimorphic differences. The largest values are those between males, respectively, females in the Öland-Kinneulle comparison. It seems therefore, that the three populations differ in the dimensions analysed.

Within sample analysis. The percentages of the total variation of each of the four eigenvalues for the six samples are given in Table XXVI. These percentages are of the same order of magnitude and reflect the general homogeneity in the material.

of the first two principal axes. In the case of the Gotland population, growth-dimorphic unlikenesses are

Table XXVII. Confidence bounds ($\times 10^{-4}$) for eigenvalues of *Omocestus*

	First	Second	Third	Fourth
<i>Kinnekulle</i>				
Males	10.944–18.807	5.944–10.215	2.786–4.788	0.796–1.368
Females	13.822–24.351	7.841–13.813	1.974–3.478	0.981–1.728
<i>Öland</i>				
Males	13.744–23.055	6.408–10.750	1.889–3.168	0.993–1.667
Females	19.866–33.556	6.330–10.693	1.847–3.120	0.843–1.424
<i>Gotland</i>				
Males	12.780–21.075	6.604–10.890	1.687–2.782	0.972–1.603
Females	14.832–26.620	7.518–13.494	2.340–4.199	0.803–1.442

Table XXVIII. Correlation of first two principal components with original variables

	x_1	x_2	x_3	x_4
<i>Kinnekulle</i>				
Males	0.7120	0.7741	0.6909	0.7146
	0.1367	0.5588	0.1657	-0.6882
Females	0.5948	0.4662	0.6705	0.8849
	0.6314	0.6892	0.6176	-0.4644
<i>Öland</i>				
Males	0.7226	0.7670	0.7351	0.8117
	0.5544	0.4013	0.4808	-0.5836
Females	0.9138	0.8621	0.8510	0.6766
	0.2769	0.2605	0.2478	-0.7363
<i>Gotland</i>				
Males	0.5348	0.4979	0.5134	0.9350
	0.6206	0.7510	0.6911	-0.3545
Females	0.8163	0.7603	0.7965	0.7191
	0.4619	0.4616	0.3195	-0.6938

The confidence bounds for eigenvalue percentages shown in Table XXVII vary mostly within fairly narrow bounds and the only sample deviating to any extent is that for females from Öland.

The first principal component for all samples is interpretable as a size-variational factor. The elements of the eigenvectors show considerable differences from each other. Males of the samples from Gotland

and Öland display a similar pattern for the first principal component in that x_1 , x_2 and x_3 are equally weighted. This is also so for the females and there is therefore a chance that the Kinnekulle population is, in this respect, less like the other two populations than they resemble each other. The second principal components for males from the Öland and Gotland populations are more like each other than they are the population from Kinnekulle. However, the females for both samples resemble each other with respect to the second principal component.

The correlations of the first two principal components with the original variables are shown in Table XXVIII. The point of interest here lies in the same overall patterns of correlation. The graphs of the first two transformed observations and in one case the first three transformed observations are given in Figs. 4–10. The shapes are mostly ellipsoidal and indicate reasonable agreement with the multivariate normal distribution. The third principal components for males and females of the Gotland and Öland samples show considerable agreement, but all differ from the patterns for males and females from Kinnekulle. Thus, the principal component analysis points towards a closer relationship between the populations of Öland and Gotland than between either of these with Kinnekulle.

Table XXIX. Some results for untransformed data for pairs of samples of males and females of *Omocestus*

Locality	B^2	D_h^2	Orientations of ellipsoid axes (χ^2)			
			1	2	3	4
Kinnekulle	48.00	60.49	23.87***	25.80***	25.49***	11.81*
Gotland	51.08	77.31	13.90**	11.78**	34.45***	36.21***
Öland	45.07	78.72	31.05***	7.39	8.02	39.21***

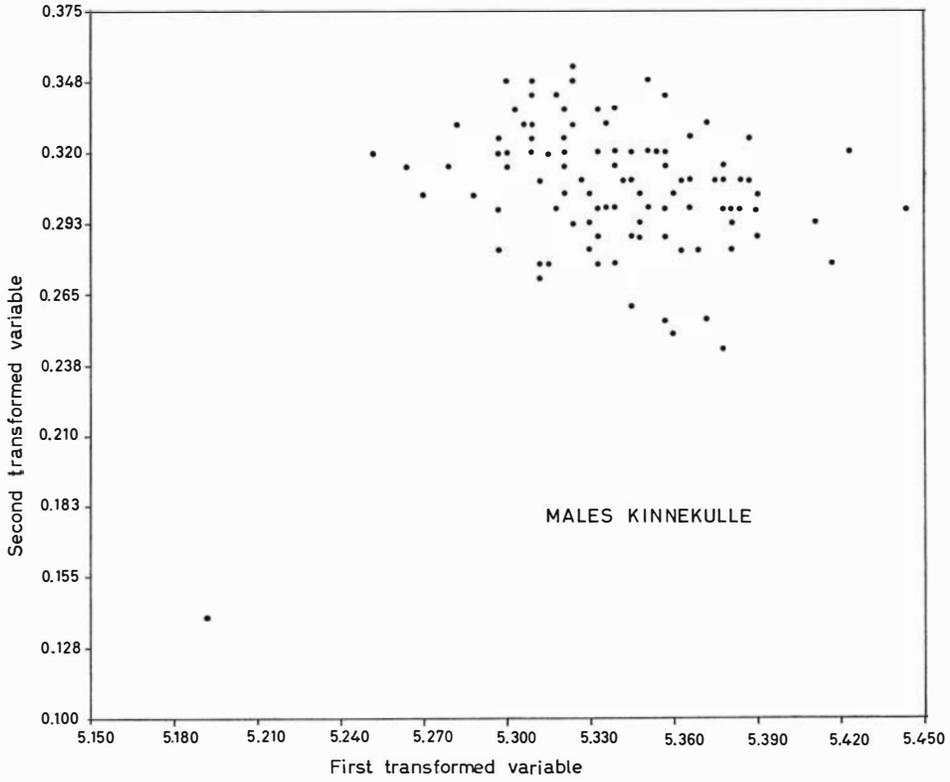


Fig. 4. Graph of first two transformed variables of PCA for male *Omocestus* from Kinnekulle.

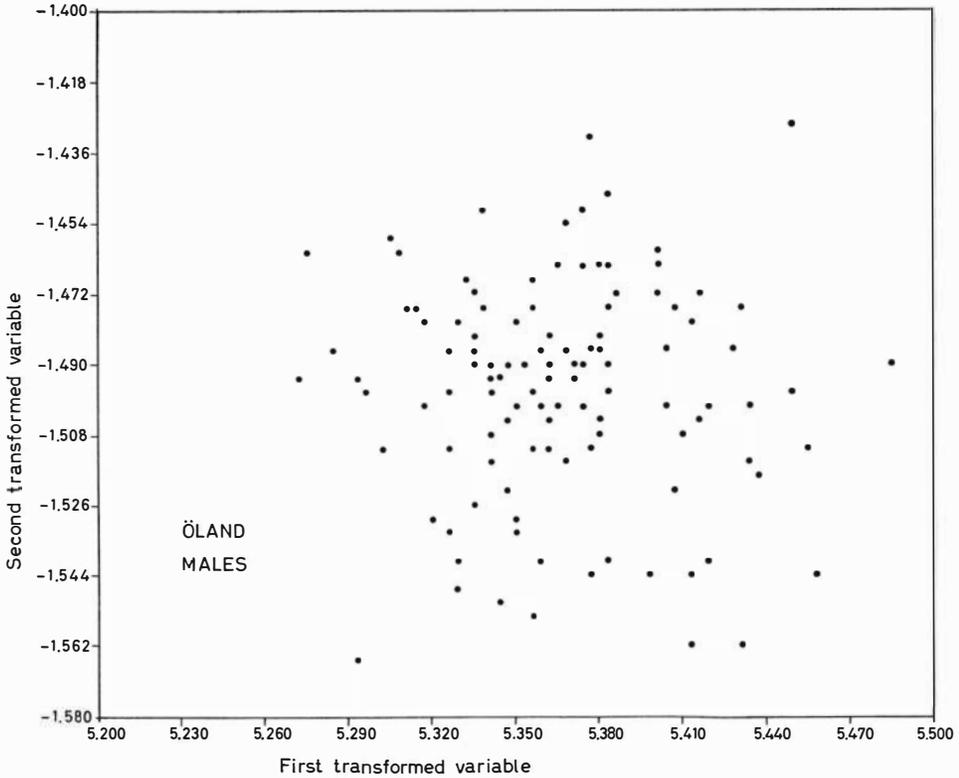


Fig. 5. Graph of first two transformed variables of PCA for male *Omocestus* from Öland.

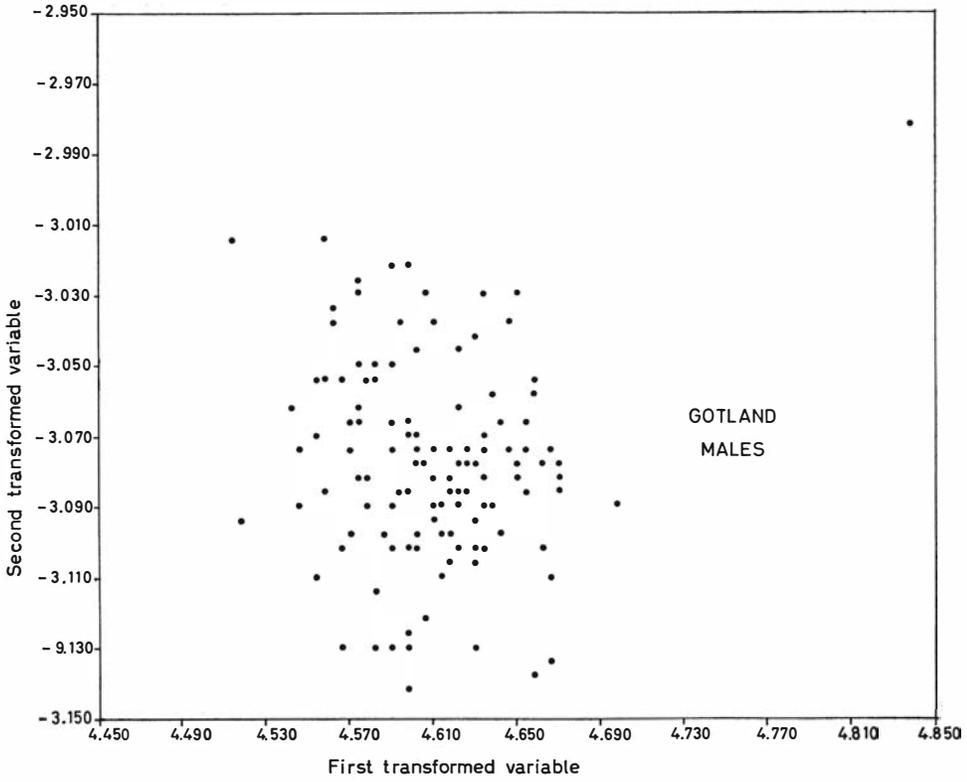


Fig. 6. Graph of first two transformed variables of PCA for male *Omocestus* from Gotland

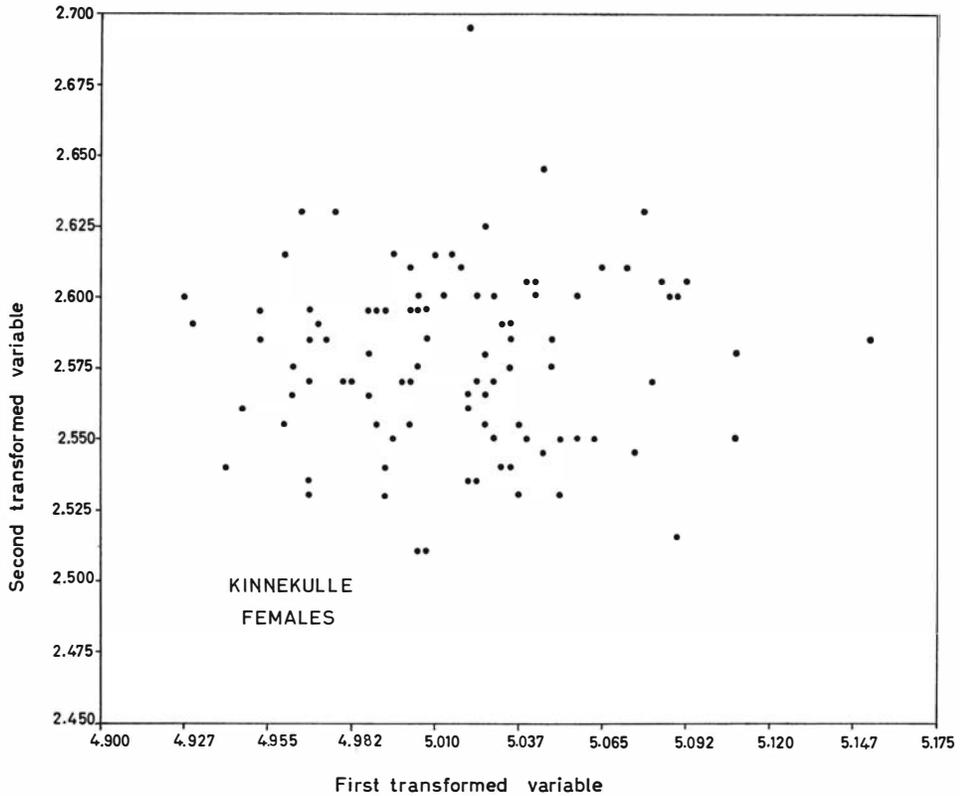


Fig. 7. Graph of first two transformed variables of PCA for female *Omocestus* from Kinnekulle.

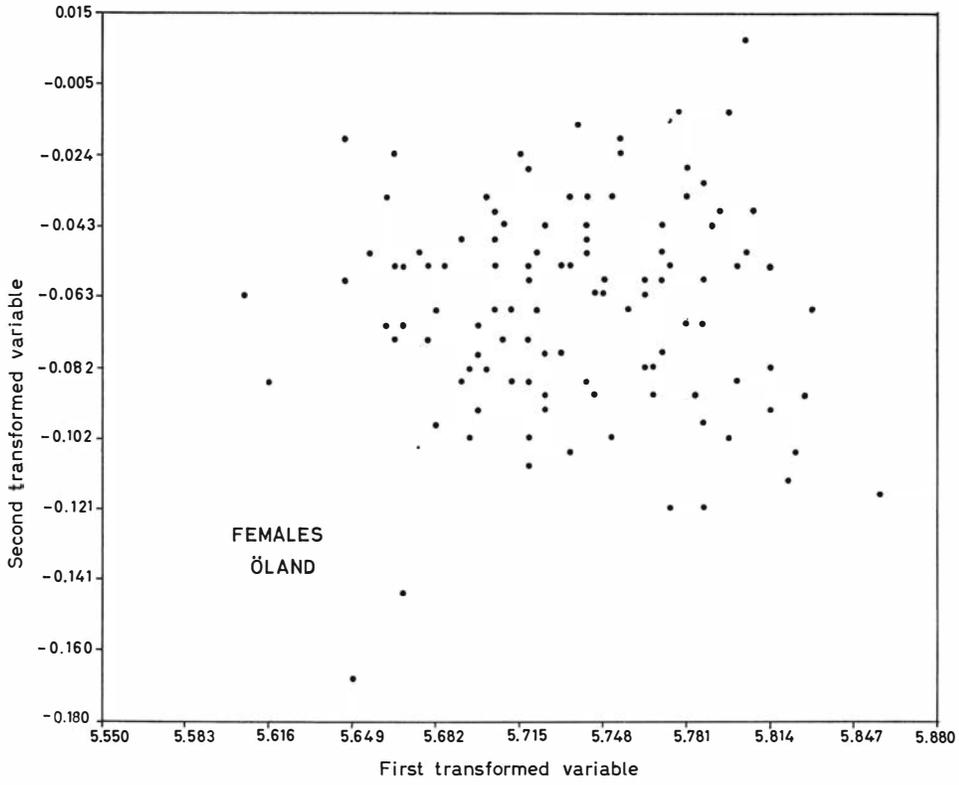


Fig. 8. Graph of first two transformed variables of PCA for female *Omocestus* from Öland.

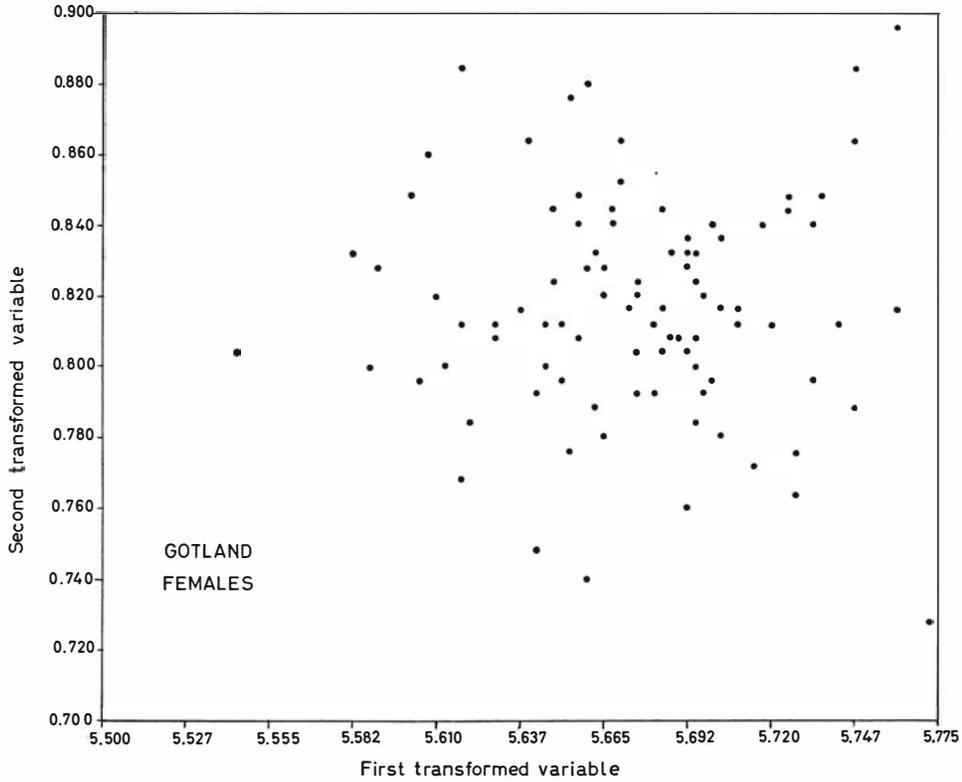


Fig. 9. Graph of first two transformed variables of PCA for female *Omocestus* from Gotland.

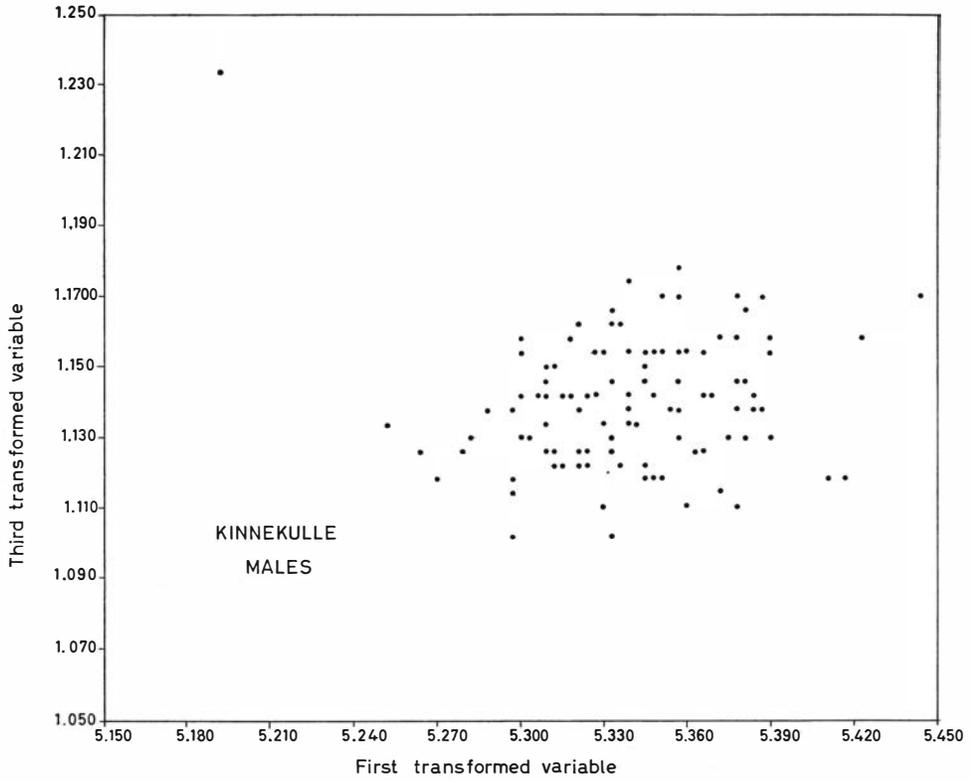


Fig. 10. Graph of first and third transformed variables for males of *Omocestus* from Kinnekulle.

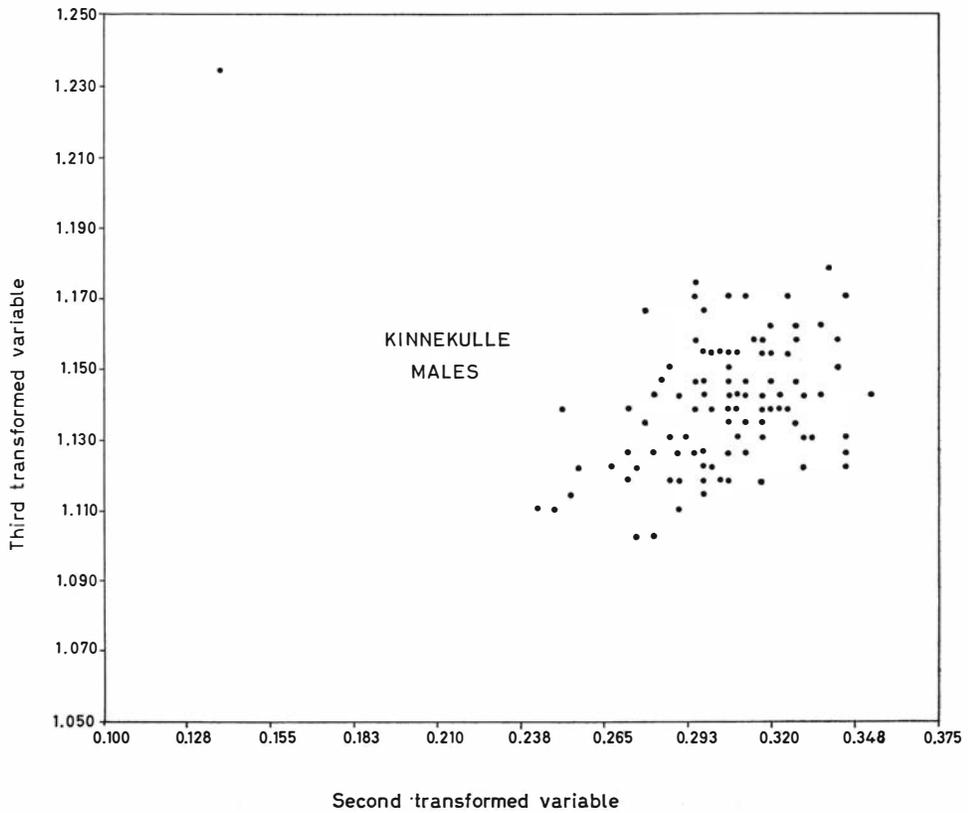


Fig. 11. Graph of second and third transformed variables of PCA for male *Omocestus* from Kinnekulle.

It is instructive to compare the results obtained using the raw data with those based on the logarithmically transformed data. These are given in Table XXIX. For Kinnekulle, the values of B^2 are close, but D_h^2 is slightly greater for the untransformed data. The main difference lies in the orientations of the ellipsoid axes. Although in all cases, the axes are differently oriented, the significances are higher for the untransformed data. For Gotland, the value of B^2 for untransformed data is more than twice that for transformed data although the squared generalized distances are about the same. The differences in orientations of the ellipsoid axes are about the same, with all of them showing significant non-collinearity. The Öland samples differ greatly in B^2 and the squared generalized distance for raw data is appreciably less than that of the transformed data. There are also differences in the results for the tests of collinearity. The explanation of these divergencies would appear to be mainly in the normalizing effect of the logarithmic transformation and the sensitivity to departures from normality of the test used for homogeneity of covariance matrices.

The calculations were made with Uppsala University computer grant 104104 and the research supported by Swedish Natural Science Research Council Grants 2320-17 and 2320-24.

REFERENCES

- Anderson, T. W. 1958. *Introduction to multivariate statistical analysis*. Wiley & Sons, New York.
- 1963. Asymptotic theory for principal component analysis. *Ann. math. Statist.* 34, 122-148.
- Anderson, T. W. and Bahadur, R. R. 1962. Classification into two multivariate normal distributions with different covariance matrices. *Ann. math. Statist.* 33, 420-431.
- Bauchou A. 1966. La vie des Crabes. *Encyclopédie biol.* 46, 138 pp.
- Bennet, B. M. 1951. Note on the solution of the generalized Behrens-Fisher problem. *Ann. Inst. Statist. Math.* 2, 87-90.
- Gilchrist, B. M. 1960. Growth and form in the brine shrimp *Artemia salina* (L.) *Proc. zool. Soc. Lond.* 13, 221-235.
- Holloway, J. N. and Dunn, O. J. 1967. The robustness of Hotelling's T^2 . *J. Amer. statist. Ass.* 62, 124-136.
- Hopkins, J. W. and Clay, P. P. F. 1963. Some empirical distributions of bivariate T^2 and homoscedasticity criterion M under unequal variance and leptokurtosis. *J. Amer. statist. Ass.* 58, 1048-1053.
- Huxley, J. 1932. *Problems of relative growth*. Methuen, London.
- Ito, K. and Schull, W. J. 1964. On the robustness of the T_0^2 test in multivariate analysis of variance when variance-covariance matrices are not equal. *Biometrika* 51, 71-82.

- Jolicoeur, P. 1963. The degree of generality of robustness in *Martes americana*. *Growth* 27, 1-27.
- Jolicoeur, P. and Mosimann, J. E. 1960. Size and shape variation in the painted turtle. A principal component analysis. *Growth* 24, 339-354.
- Kauri, H. 1959. *Die Rassenbildung bei europäischen Rana-Arten und die Gültigkeit der Klimaregeln*, 172 pp. Gleerupska Universitetsbokhandels Förlag AB Lund.
- Kinne, O. 1964. The effects of temperature and salinity on marine and brackish water animals. II. Salinity and temperature salinity combinations. *Ann. Rev. Oceanogr. Mar. Biol.* 2, 281-339.
- Kullback, S. 1959. *Information theory and statistics*. Wiley and Sons, New York.
- Rao, C. R. 1952. *Advanced statistical methods in biometric research*. Wiley & Sons, New York, London.
- Reyment, R. A. 1962. Observations on homogeneity of covariance matrices in paleontologic biometry. *Biometrics* 18, 1-11.
- 1965. Studies on Nigerian Upper Cretaceous and Lower Tertiary Ostracoda Part 3: Stratigraphical, paleoecological and biometrical conclusions. *Stockh. Contr. Geol.* 14, 151 pp.
- 1967. A multivariate paleontological growth problem. *36th Sess. Inst. int. Statist. Sydney* 2/81 (11), 9 pp.
- 1969. A multivariate paleontological growth problem. *Biometrics*, 25, 1-8.