A5 Discriminant Analysis

Problem

For two groups, show that the discriminant function coefficient vector $\mathbf{a}$, that maximises the discriminant criterion $\frac{\left(\bar{z}_1 - \bar{z}_2\right)^2}{s_z^2} = \frac{\left[\mathbf{a}'\left(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2\right)\right]^2}{\mathbf{a}'\mathbf{S}_{\mathbf{pl}}\mathbf{a}}$, is given by $\mathbf{a} = \mathbf{S}_{\mathbf{pl}}^{-1}\left(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2\right)$.

Problem

For $z = \mathbf{a}'\mathbf{y}$ with $\mathbf{a} = \mathbf{S}_{\mathbf{pl}}^{-1}\left(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2\right)$, show that

$$\frac{\left(\bar{z}_1 - \bar{z}_2\right)^2}{s_z^2} = \left(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2\right)' \mathbf{S}_{\mathbf{pl}}^{-1}\left(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2\right)$$

Problem

For several groups, show that the third discriminant function $z_3 = \mathbf{a}_3'\mathbf{y}$ has the largest discriminant criterion value λ_3 among all linear combination of the $y's$ that are uncorrelated with both z_1 and z_2.

Problem

Four psychological tests were given to 32 men and 32 women. The data are given in the table below. The variables are

y_1 = pictorial inconsistencies

y_2 = paper form board

y_3 = tool recognition

y_4 = vocabular

Males				Females			
y_1	y_2	y_3	y_4	y_1	y_2	y_3	y_4
15	17	24	14	13	14	12	21
17	15	32	26	14	12	14	26
15	14	29	23	12	19	21	21
13	12	10	16	12	13	10	16
20	17	26	28	11	20	16	16
15	21	26	21	12	9	14	18
15	13	26	22	10	13	18	24
13	5	22	22	10	8	13	23
14	7	30	17	12	20	19	23
17	15	30	27	11	10	11	27
17	17	26	20	12	18	25	25
17	20	28	24	14	18	13	26
15	15	29	24	14	10	25	28
18	19	32	28	13	16	8	14
18	18	31	27	14	8	13	25
15	14	26	21	13	16	23	28
18	17	33	26	16	21	26	26
10	14	19	17	14	17	14	14
18	21	30	29	16	16	15	23
18	21	34	26	13	16	23	24
13	17	30	24	2	6	16	21
16	16	16	16	14	16	22	26
11	15	25	23	17	17	22	28
16	13	26	16	16	13	16	14
16	13	23	21	15	14	20	26
18	18	34	24	12	10	12	9
16	15	28	27	14	17	24	23
15	16	29	24	13	15	18	20
18	19	32	23	11	16	18	28
18	16	33	23	7	7	19	18
17	20	21	21	12	15	7	28
19	19	30	28	6	5	6	13

For this psychological data,

(a) Determine the discriminant function coefficient vector

(b) Find the standardised coefficients

(c) Calculate t-tests for individual variables

(d) Compare the results of (b) and (c) as to the contribution of the variables to separation of the two groups.

(e) Find the partial F for each variable, and compare with the standardised coefficients?

Problem

The data in the table below are judges' scores on fish prepared by three methods. Twelve fish were cooked by each method, and several judges tasted fish samples and rated each on four variables: y_1 = aroma, y_2 = flavour, y_3 = texture, and y_4 = moisture. Each entry in the table is an average score for the judges on that fish. Using this data, do the following:

(a) Find the eigenvectors of $\mathbf{E}^{-1}\mathbf{H}$.

(b) Carry out tests of significance for the discriminant functions and find the relative importance of each $\left(\frac{\lambda_i}{\sum_{j=1}^{s}\lambda_j}\right)$. Do these two procedures agree as to the number of important discriminant functions?

(c) Find the standardised coefficients and comment on the contribution of the variables to separation of groups.

(d) Find the partial F for each variable. Do they rank the variables in the same order as the standardised coefficients for the first discriminant function?

(e) Plot the first two discriminant functions for each observation and for the mean vectors.

Method 1				Method 2				Method 3			
y_1	y_2	y_3	y_4	y_1	y_2	y_3	y_4	y_1	y_2	y_3	y_4
5.4	6.0	6.3	6.7	5.0	5.3	5.3	6.5	4.8	5.0	6.5	7.0
5.2	6.2	6.0	5.8	4.8	4.9	4.2	5.6	5.4	5.0	6.0	6.4
6.1	5.9	6.0	7.0	3.9	4.0	4.4	5.0	4.9	5.1	5.9	6.5
4.8	5.0	4.9	5.0	4.0	5.1	4.8	5.8	5.7	5.2	6.4	6.4
5.0	5.7	5.0	6.5	5.6	5.4	5.1	6.2	4.2	4.6	5.3	6.3
5.7	6.1	6.0	6.6	6.0	5.5	5.7	6.0	6.0	5.3	5.8	6.4
6.0	6.0	5.8	6.0	5.2	4.8	5.4	6.0	5.1	5.2	6.2	6.5
4.0	5.0	4.0	5.0	5.3	5.1	5.8	6.4	4.8	4.6	5.7	5.7
5.7	5.4	4.9	5.0	5.9	6.1	5.7	6.0	5.3	5.4	6.8	6.6
5.6	5.2	5.4	5.8	6.1	6.0	6.1	6.2	4.6	4.4	5.7	5.6
5.8	6.1	5.2	6.4	6.2	5.7	5.9	6.0	4.5	4.0	5.0	5.9
5.3	5.9	5.8	6.0	5.1	4.9	5.3	4.8	4.4	4.2	5.6	5.5