

# Exam 2009 problem 1

Adapted to omit SAS enclosure, data are given below:

```
data P1;
input Obs species $ number x1 x2 x3;
datalines;
1 conc 1 191 131 53
2 conc 2 185 134 50
3 conc 3 200 137 52
4 conc 4 173 127 50
5 conc 5 171 118 49
6 conc 6 160 118 47
7 conc 7 188 134 54
8 conc 8 186 129 51
9 conc 9 174 131 52
10 conc 10 163 115 47
11 conc 11 190 143 52
12 conc 12 174 131 50
13 conc 13 201 130 51
14 conc 14 190 133 53
15 conc 15 182 130 51
16 conc 16 184 131 51
17 conc 17 177 127 49
18 conc 18 178 126 53
19 conc 19 210 140 54
20 conc 20 182 121 51
21 conc 21 186 136 56
22 Heik 1 186 107 49
23 Heik 2 211 122 49
24 Heik 3 201 114 47
25 Heik 4 242 131 54
26 Heik 5 184 108 43
27 Heik 6 211 118 51
28 Heik 7 217 122 49
29 Heik 8 223 127 51
30 Heik 9 208 125 50
31 Heik 10 199 124 46
32 Heik 11 211 129 49
33 Heik 12 218 126 49
34 Heik 13 203 122 49
35 Heik 14 192 116 49
36 Heik 15 195 123 47
37 Heik 16 211 122 48
38 Heik 17 187 123 47
39 Heik 18 192 109 46
40 Heik 19 223 124 53
41 Heik 20 188 114 48
42 Heik 21 216 120 50
43 Heik 22 185 114 46
44 Heik 23 178 119 47
45 Heik 24 187 111 49
46 Heik 25 187 112 49
47 Heik 26 201 130 54
48 Heik 27 187 120 47
49 Heik 28 210 119 50
50 Heik 29 196 114 51
51 Heik 30 195 110 49
52 Heik 31 187 124 49
;
run;
```

Please note, that there is one and only one correct answer to each question. Furthermore, some of the possible alternative answers may not make sense. When the text refers to SAS-output the values may be rounded to fewer decimal places than in the output itself. Please check that all pages of the exam paper and the enclosure are present.

## Problem 1.

Enclosure A with the SAS-program and the corresponding SAS-output belongs to this problem.

The data are part of a study on two species of male flea-beetles: *Chaetocnema concinna* and *Chaetocnema heikertlinger*.

The variables are:

Variable	Description
species	: "Conc" = <i>Chaetocnema concinna</i> , "Heik" = <i>Chaetocnema heikertlinger</i>
number	: observation number within species
$x_1$	: width of the first joint of the first tarsus in microns (the sum of measurements for both tarsi)
$x_2$	: the same for the second joint
$x_3$	: the maximal width of the aedeagus in the fore-part in microns

We consider a linear discriminant analysis (LDA) with equal priors and losses.

### Question 1.1.

Mahalanobis' distance between the two groups is:

**1**   $\frac{(193.8 - 183.1)^2}{267.2}$

**2**   $| -290.6 - (-245.9) |$

**3**   $1 - 0.0323$

**4**  15.07

**5**  9.52

**6**  Don't know.

The problem continues on the next page

## Question 1.2.

The usual test statistic for testing a difference in mean for the two groups is distributed as  $F(\nu_1, \nu_2)$  where  $\nu_2$  equals:

- 1**  46
- 2**  47
- 3**  48
- 4**  49
- 5**  50
- 6**  Don't know.

## Question 1.3.

The pooled variance estimate for variable  $x_1$  is:

- 1**  267.2
- 2**   $\frac{267.2}{193.8}$
- 3**   $\frac{147.49+222.13}{2}$
- 4**   $\frac{20 \cdot 147.49 + 30 \cdot 222.13}{20+30}$
- 5**   $\frac{21 \cdot 147.49 + 31 \cdot 222.13}{21+31}$
- 6**  Don't know.

## Question 1.4.

A new flea-beetle specimen is collected. It is classified as "Conc" if  $\mathbf{d}' \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} > c$ .

The constant  $c$  is:

- 1**  -44.63
- 2**  44.63
- 3**  15.07
- 4**  9.52
- 5**  10
- 6**  Don't know.

*The problem continues on the next page*

## Question 1.5.

The generalised variance of the pooled covariance matrix is:

- 1**  13618.0
- 2**   $267.2 \cdot 72.1 \cdot 6.6$
- 3**   $e^{\sqrt{15.07}}$
- 4**   $12.1^2 + 7.2^2 + 2.2^2 + 14.9^2 + 6.6^2 + 2.3^2$
- 5**  50
- 6**  Don't know.

## Question 1.6.

OMITTED