*Written test, date*: 9. December 2008 *Course no.* : 02409

*Course name*: Multivariate Statistics “Multivariat Statistik”.

*Aids allowed:* All usual ones

*“Weighting”:* The questions are given equal weight.

This exam is answered by:

(name) (signature) (study no.)

There is a total of 30 questions for the 8 problems. The answers to the 30 questions must be written into the table below.

| **Problem** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Question** | 1.1 | 1.2 | 1.3 | 1.4 | 1.5 | 1.6 | 1.7 | 1.8 | 1.9 | 1.10 |
| **Answer** | | | | | | | | | | |

| **Problem** | 1 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Question** | 1.11 | 2.1 | 2.2 | 2.3 | 2.4 | 2.5 | 3.1 | 3.2 | 3.3 | 3.4 |
| **Answer** | | | | | | | | | | |

| **Problem** | 3 | 3 | 3 | 3 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Question** | 3.5 | 3.6 | 3.7 | 3.8 | 3.9 | 4.1 | 5.1 | 6.1 | 7.1 | 8.1 |
| **Answer** | | | | | | | | | | |

The possible answers for each question are numbered from 1 to 6. If you enter a wrong number, you may correct it by crossing the wrong number in the table and writing the correct answer immediately below. If there is any doubt about the meaning of a correction then the question will be considered not answered.

**Only the front page must be returned**. The front page must be returned even if you do not answer any of the questions or if you leave the exam prematurely. Drafts and/or comments are **not** considered, only the numbers entered above are registered.

A correct answer gives 5 points, a wrong answer gives $-1$ point. Unanswered questions or a 6 (corresponding to “don’t know”) gives 0 points. The total number of points, needed for a satisfactorily answered exam is determined at the final evaluation of the exam.

Remember to write your name, signature and study number on the front page.

*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.*

# About Enclosures A, B, and C.

The data are part of a larger study by Dr. Rick Linthurst while he was doing his PhD at Noth Carolina State University. He considered the aerial biomass of a certain marsh grass *Spartina Alterniflora* in the Cape Fear Estuary of North Carolina.

There are corresponding observations of aerial biomass, and soil characteristics: salinity, acidity, potassium, sodium, zinc. Furthermore, there are recordings of locality and type of vegetation.

The variables are:

| Variable | | Description |
|---|---|---|
| Location | : | OI=Oak Island, SI=Smith Island, SM=Snows Marsh |
| Type | : | DVEG=revegetated "dead" areas, SHRT="short" Spartina areas, TALL="tall" Spartina areas |
| Biomass | : | aerial biomass (g/m$^2$) |
| Salinity | : | salinity (o/oo) |
| pH | : | acidity as measured in water (pH) |
| K | : | potassium (ppm) |
| Na | : | sodium (ppm) |
| Zn | : | zinc (ppm) |

These data are used for different analyses in Enclosures A, B, and C.

# Problem 1.

The marked part of the SAS-program and Enclosure A with the corresponding SAS-output belongs to this problem.

The origin of the data and the different variables are explained on page 2 of this exam.

*The problem continues on the next page*

# Question 1.1.

Consider the correlation matrix where we condition on pH.

The usual test for hypotheses of the type: $\rho_{ij|\mathrm{pH}} = 0$ has the following degrees of freedom:

**1** ☐ 41

**2** ☐ 42

**3** ☐ 43

**4** ☐ 44

**5** ☐ 45

**6** ☐ Don't know.

# Question 1.2.

Consider the correlation matrix where we condition on pH.

The usual test for the hypothesis: $\rho_{\mathrm{Biomass,K|pH}} = 0$:

**1** ☐ is not significant at level 0.0001, but is significant at level 0.001

**2** ☐ is not significant at level 0.001, but is significant at level 0.005

**3** ☐ is not significant at level 0.005, but is significant at level 0.01

**4** ☐ is not significant at level 0.01, but is significant at level 0.025

**5** ☐ is not significant at level 0.025, but is significant at level 0.05

**6** ☐ Don't know.

# Question 1.3.

The partial correlation $\rho_{\mathrm{Salinity,Zn|pH}}$ can be calculated by:

**1** ☐ $-0.42$

**2** ☐ $\frac{-0.42}{\sqrt{(-0.051)^2(-0.72)^2}}$

**3** ☐ $\frac{-0.42-(-0.051)(-0.72)}{\sqrt{(1-(-0.051)^2)(1-(-0.72)^2)}}$

**4** ☐ $\frac{-0.42}{\sqrt{(-0.051)(-0.72)}}$

**5** ☐ Cannot be given using the information provided

**6** ☐ Don't know.

*The problem continues on the next page*

# Question 1.4.

Consider the unconditional correlation matrix used in the factor analysis. Which correlation between two variables will be most significant, when considering the usual test for hypotheses of the type: $\rho_{ij} = 0$

**1** ☐ $\rho_{\text{Salinity,K}}$

**2** ☐ $\rho_{\text{pH,K}}$

**3** ☐ $\rho_{\text{pH,Zn}}$

**4** ☐ $\rho_{\text{pH,Salinity}}$

**5** ☐ $\rho_{\text{K,Na}}$

**6** ☐ Don't know.

# Question 1.5.

The first 3 factors together:

**1** ☐ explain more than or equal to 20%, but less than 50% of the total variation

**2** ☐ explain more than or equal to 50%, but less than 75% of the total variation

**3** ☐ explain more than or equal to 75%, but less than 90% of the total variation

**4** ☐ explain more than or equal to 90%, but less than 95% of the total variation

**5** ☐ explain more than or equal to 95%, but less than 99% of the total variation

**6** ☐ Don't know.

# Question 1.6.

The first normed eigenvector of the correlation-matrix used in the factor analysis is:

**1** ☐ $[-0.21\ -0.61\ \ 0.67\ \ 0.68\ \ 0.75]'$

**2** ☐ $1.88[-0.21\ -0.61\ \ 0.67\ \ 0.68\ \ 0.75]'$

**3** ☐ $\sqrt{1.88}[-0.21\ -0.61\ 0.67\ \ 0.68\ \ 0.75]'$

**4** ☐ $\frac{1}{1.88}[-0.21\ -0.61\ 0.67\ \ 0.68\ \ 0.75]'$

**5** ☐ $\frac{1}{\sqrt{1.88}}[-0.21\ -0.61\ 0.67\ \ 0.68\ \ 0.75]'$

**6** ☐ Don't know.

*The problem continues on the next page*

# Question 1.7.

Consider the correlation-matrix used in the factor analysis. In order to calculate the usual test statistic for the hypothesis: $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4 = \lambda_5$ we (among other things) need the following piece of information:

**1** ☐ 5

**2** ☐ $\frac{0.0428+0.0212}{2}$

**3** ☐ $0.9101 \cdot 0.9356$

**4** ☐ $\sqrt{1.8838}$

**5** ☐ $\frac{0.2138+0.1059}{2}$

**6** ☐ Don't know.

# Question 1.8.

Consider the correlation-matrix used in the factor analysis. The usual test statistic under the hypothesis: $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4 = \lambda_5$ has the following approximate distribution:

**1** ☐ $\chi^2(1)$

**2** ☐ $\chi^2(2)$

**3** ☐ $\chi^2(3)$

**4** ☐ $\chi^2(4)$

**5** ☐ $\chi^2(5)$

**6** ☐ Don't know.

# Question 1.9.

Consider the factor analysis. The variable with the **largest** value of uniqueness is:

**1** ☐ Salinity

**2** ☐ pH

**3** ☐ K

**4** ☐ Na

**5** ☐ Zn

**6** ☐ Don't know.

*The problem continues on the next page*

# Question 1.10.

Consider the factor analysis. The 3 varimax-rotated factors can be interpreted as:

**1** ☐ Varimax-rotated factor 1: mainly describes variation in pH, K, Na, and Zn; Varimax-rotated factor 2: mainly describes variation in K, Na, and Zn; Varimax-rotated factor 3: mainly describes variation in Salinity.

**2** ☐ Varimax-rotated factor 1: mainly describes variation in K and Na; Varimax-rotated factor 2: mainly describes variation in pH and Zn; Varimax-rotated factor 3: mainly describes variation in Salinity.

**3** ☐ Varimax-rotated factor 1: mainly describes variation in Salinity, pH, and Zn; Varimax-rotated factor 2: mainly describes variation in Salinity, K, and Na; Varimax-rotated factor 3: mainly describes variation in pH, K, Na, and Zn.

**4** ☐ Varimax-rotated factor 1: mainly describes variation in Zn; Varimax-rotated factor 2: mainly describes variation in K and Na; Varimax-rotated factor 3: mainly describes variation in Salinity and pH.

**5** ☐ Varimax-rotated factor 1: mainly describes variation in Salinity; Varimax-rotated factor 2: mainly describes variation in pH and K; Varimax-rotated factor 3: mainly describes variation in Na and Zn.

**6** ☐ Don't know.

# Question 1.11.

Consider the factor analysis. The factor loading for variable Salinity in varimax-rotated factor 1 can be found by:

**1** ☐ $\begin{bmatrix} 0.70 \\ -0.68 \\ -0.20 \end{bmatrix} \cdot [-0.21\ 0.44\ 0.86]$

**2** ☐ $[-0.21\ 0.44\ 0.86] \cdot \begin{bmatrix} 0.70 \\ -0.68 \\ -0.20 \end{bmatrix}$

**3** ☐ $0.0498^2 + 0.0681^2 0.9877^2$

**4** ☐ $\begin{bmatrix} 0.70 \\ 0.70 \\ -0.13 \end{bmatrix} \cdot [-0.21\ 0.44\ 0.86]$

**5** ☐ $[-0.21\ 0.44\ 0.86] \cdot \begin{bmatrix} 0.70 \\ 0.70 \\ -0.13 \end{bmatrix}$

**6** ☐ Don't know.

# Problem 2.

The marked part of the SAS-program and Enclosure B with the corresponding SAS-output belongs to this problem.

The origin of the data and the different variables are explained on page 2 of this exam.

Option "simple" in the "proc discrim" statement gives various univariate statistics.

## Question 2.1.

Consider the output from the first "proc discrim". The estimate of the variance-covariance matrix used to construct the classification functions is based on:

**1** ☐ 1 degree of freedom

**2** ☐ 14 degrees of freedom

**3** ☐ 28 degrees of freedom

**4** ☐ 29 degrees of freedom

**5** ☐ 30 degrees of freedom

**6** ☐ Don't know.

## Question 2.2.

Consider the output from the first "proc discrim". The value of the usual F-distributed test statistic for the hypothesis $\boldsymbol{\mu}_{\mathrm{OI}} = \boldsymbol{\mu}_{\mathrm{SI}}$ amounts to: 9.0245

In that case the test is

**1** ☐ significant at level 0.25, but not at level 0.1

**2** ☐ significant at level 0.1, but not at level 0.05

**3** ☐ significant at level 0.05, but not at level 0.01

**4** ☐ significant at level 0.01, but not at level 0.005

**5** ☐ significant at level 0.005

**6** ☐ Don't know.

*The problem continues on the next page*

# Question 2.3.

Consider the output from the first "proc discrim". A new observation $(x_{\text{Salinity}}, x_{\text{K}})$ is classified as "OI" if:

**1** □ $-70.47 + 4.02x_{\text{Salinity}} + 0.00846x_{\text{K}} > 0$

**2** □ $-61.17 + 3.85x_{\text{Salinity}} + 0.00326x_{\text{K}} > 0$

**3** □ $-70.47 + 4.02x_{\text{Salinity}} + 0.00846x_{\text{K}} > -61.17 + 3.85x_{\text{Salinity}} + 0.00326x_{\text{K}}$

**4** □ $-70.47 + 4.02x_{\text{Salinity}} + 0.00846x_{\text{K}} < -61.17 + 3.85x_{\text{Salinity}} + 0.00326x_{\text{K}}$

**5** □ $-70.47 + 4.02x_{\text{Salinity}} + 0.00846x_{\text{K}} > 0 \ \wedge\ -61.17 + 3.85x_{\text{Salinity}} + 0.00326x_{\text{K}} > 0$

**6** □ Don't know.

# Question 2.4.

The usual test statistic for the hypothesis: "variable "K" does not contribute to a better discrimination" is found as:

**1** □ $\frac{15+15-2-1}{1}\frac{15\cdot 15(2.4957-0.3718)}{(15+15)(15+15-2)+15\cdot 15\cdot 0.3718}$

**2** □ $\frac{15+15-2-1}{2(15+15-2)}\frac{15\cdot 15}{15+15}2.4957$

**3** □ $\frac{15+15-2-1}{2(15+15-2)}\frac{15\cdot 15}{15+15}0.3718$

**4** □ $\frac{14+14-2-1}{1}\frac{14\cdot 14(2.4957-0.3718)}{(14+14)(14+14-2)+14\cdot 14\cdot 0.3718}$

**5** □ $\frac{14+14-2-1}{2(14+14-2)}\frac{14\cdot 14}{14+14}2.4957$

**6** □ Don't know.

# Question 2.5.

We wish to construct a quadratic discriminant rule with equal priors based on one single variable: "Salinity".

Which one of the following numbers is **not** needed:

**1** □ 33.07

**2** □ 31.33

**3** □ 8.579

**4** □ 7.781

**5** □ 8.381

**6** □ Don't know.

# Problem 3.

The marked part of the SAS-program and Enclosure C with the corresponding SAS-output belongs to this problem. #### indicates that information has been concealed *(Danish: skjult el. fjernet)*.

Note: the first "proc reg" is based on all 45 observations, the second "proc reg" is based on 44 observations (observation 34 has been deleted).

The origin of the data and the different variables are explained on page 2 of this exam.

For several of the following questions we will be referring to the following models:

"Model $M_1$": corresponds to the first "proc reg" and is a regression model based on all 45 observations which might be written as:

$$\begin{bmatrix} 676 \\ 516 \\ \vdots \\ 1560 \end{bmatrix} = \begin{bmatrix} 1 & 5.00 & 35184.5 \\ 1 & 4.75 & 28170.4 \\ & \vdots & \\ 1 & 5.40 & 16892.2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{45} \end{bmatrix}$$

The residuals are assumed identically, independently, and normally distributed.

"Model $M_2$": is a regression model based on all 45 observations which might be written as:

$$\begin{bmatrix} 676 \\ 516 \\ \vdots \\ 1560 \end{bmatrix} = \mu + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_{45} \end{bmatrix}$$

The residuals are assumed identically, independently, and normally distributed.

## Question 3.1.

Using maximum likelihood $\alpha$ in model $M_1$ is estimated at:

**1** ☐ -475.73

**2** ☐ 404.95

**3** ☐ -0.02333

**4** ☐ 0.0893

**5** ☐ 0.0101

**6** ☐ Don't know.

*The problem continues on the next page*

## Question 3.2.

The reduction in variance when going from model $M_2$ to model $M_1$ amounts to the following fraction:

**1** ☐ 1.0

**2** ☐ 0.6584

**3** ☐ 40.48

**4** ☐ 394.85

**5** ☐ 45

**6** ☐ Don't know.

## Question 3.3.

For model $M_1$ the usual (formal) test statistic for the hypothesis: "$R^2 = 0$" has the value:

**1** ☐ 40.48

**2** ☐ -1.74+8.48-2.70

**3** ☐ 0.6584

**4** ☐ 394.85

**5** ☐ 39.45

**6** ☐ Don't know.

## Question 3.4.

Consider model $M_1$. Under the null-hypothesis "$R^2 = 0$" the usual (formal) test statistic is distributed as:

**1** ☐ F(42,44)

**2** ☐ F(3,44)

**3** ☐ F(3,42)

**4** ☐ F(2,44)

**5** ☐ F(2,42)

**6** ☐ Don't know.

*The problem continues on the next page*

# Question 3.5.

Using model $M_1$ the first observation is predicted at:

**1** ☐ 728.30

**2** ☐ -475.72

**3** ☐ 394.85

**4** ☐ 404.94

**5** ☐ -52.30

**6** ☐ Don't know.

# Question 3.6.

Values of Cook's D for model $M_1$ can be compared to a suitable percentile in the following distribution:

**1** ☐ F(1,44)

**2** ☐ F(2,43)

**3** ☐ F(3,42)

**4** ☐ F(4,41)

**5** ☐ F(5,40)

**6** ☐ Don't know.

# Question 3.7.

$\mathrm{RSTUDENT}_{34}$ (corresponding to observation 34) for model $M_1$ is estimated at:

**1** ☐ 1168

**2** ☐ $\frac{1168}{394.85}$

**3** ☐ $\frac{1168}{353.47}$

**4** ☐ $\frac{1168}{394.85\sqrt{1-0.0423}}$

**5** ☐ $\frac{1168}{353.47\sqrt{1-0.0423}}$

**6** ☐ Don't know.

*The problem continues on the next page*

# Question 3.8.

In order to calculate a 95% prediction interval for model $M_1$ we need a suitable percentile in a t-distribution. The value we need is:

**1** ☐ 1.68 (approximately)

**2** ☐ 2.02 (approximately)

**3** ☐ 2.92 (approximately)

**4** ☐ 3.18 (approximately)

**5** ☐ 4.30 (approximately)

**6** ☐ Don't know.

# Question 3.9.

A general linear model "Model $M_3$" based on all 45 observations with separate intercepts for the three localities might be written as:

$$\begin{bmatrix} 676 \\ 516 \\ \vdots \\ 352 \\ \vdots \\ 1560 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 5.00 & 35184.5 \\ 1 & 0 & 0 & 4.75 & 28170.4 \\ & & \vdots & & \\ 0 & 1 & 0 & 3.35 & 12822.0 \\ & & \vdots & & \\ 0 & 0 & 1 & 5.40 & 16892.2 \end{bmatrix} \begin{bmatrix} \alpha_{\mathrm{OI}} \\ \alpha_{\mathrm{SI}} \\ \alpha_{\mathrm{SM}} \\ \gamma_1 \\ \gamma_2 \end{bmatrix} + \begin{bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_{17} \\ \vdots \\ \omega_{45} \end{bmatrix}$$

The residuals are assumed identically, independently, and normally distributed.

An ANOVA table for this model is:

| Source of Variation | Sum of squares | Degrees of Freedom |
|---|---|---|
| $\lVert \mathrm{p}_{M_3}(\boldsymbol{Y}) - \mathrm{p}_{M_2}(\boldsymbol{Y}) \rVert^2$ | 13908507 | 4 |
| $\lVert \boldsymbol{Y} - \mathrm{p}_{M_3}(\boldsymbol{Y}) \rVert^2$ | 5262456 | 40 |
| $\sum_{i=1}^{45}(Y_i - \frac{1}{45}\sum_{i=1}^{45} Y_i)^2$ | 19170963 | 44 |

The usual test statistic for testing if there is a difference in intercept between the three locations is found as:

**1** ☐ $\frac{(13908507-5262456)/2}{19170963/44}$

**2** ☐ $\frac{(13908507-6548174)/2}{5262456/42}$

**3** ☐ $\frac{(13908507-12622789)/2}{19170963/44}$

**4** ☐ $\frac{(13908507-12622789)/1}{5262456/40}$

**5** ☐ $\frac{(13908507-12622789)/2}{5262456/40}$

**6** ☐ Don't know.

# Problem 4.

The following data corresponds to a subset of the original data in e.g. Enclosure A

| Location | Type | Salinity | pH |
|---|---|---|---|
| OI | DVEG | 33 | 5.00 |
| OI | SHRT | 33 | 5.05 |
| OI | TALL | 30 | 4.10 |
| SI | DVEG | 30 | 3.25 |
| SI | SHRT | 30 | 3.25 |
| SI | TALL | 29 | 7.10 |
| SM | DVEG | 26 | 4.85 |
| SM | SHRT | 25 | 4.55 |
| SM | TALL | 24 | 5.60 |

## Question 4.1.

The following questions are to be answered: a) Is there a difference in mean content of (Salinity, pH) with respect to "Location"? b) Is there a difference in mean content of (Salinity, pH) with respect to "Type"?

The most sensible way of analysing this is by means of:

**1** ☐ A one-sided multivariate analysis of variance with three observations in each group.

**2** ☐ A two-sided multivariate analysis of variance with one observation per cell.

**3** ☐ Four univariate one-sided analyses of variance, testing for differences in "Location" and "Type" for each of the variables "Salinity" and "Type".

**4** ☐ Two univariate two-sided analyses of variance, testing for differences in "Location" and "Type" for each of the variables "Salinity" and "Type".

**5** ☐ A general linear model, where "Salinity" is the dependent variable and "Location", "Type", and "pH" are independent variables.

**6** ☐ Don't know.

# Problem 5.

Consider the two random vectors: $\boldsymbol{X}$ and $\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{U} \end{bmatrix}$ where $\mathrm{E}(\boldsymbol{Y}) = \boldsymbol{0}$

## Question 5.1.

Then the covariance between $\boldsymbol{X}$ and $\boldsymbol{Y}$, denoted $\mathrm{C}(\boldsymbol{X}, \boldsymbol{Y})$ equals

**1** □ $[\mathrm{D}(\boldsymbol{X}),\ \mathrm{D}(\boldsymbol{U})]$

**2** □ $[\mathrm{C}(\boldsymbol{U}, \boldsymbol{X}),\ \mathrm{D}(\boldsymbol{X})]$

**3** □ $[\mathrm{D}(\boldsymbol{X}),\ \mathrm{C}(\boldsymbol{U}, \boldsymbol{X})]$

**4** □ $[\mathrm{D}(\boldsymbol{X}),\ \mathrm{C}(\boldsymbol{X}, \boldsymbol{U})]$

**5** □ $[\mathrm{C}(\boldsymbol{X}, \boldsymbol{U}),\ \mathrm{D}(\boldsymbol{X})]$

**6** □ Don't know.

# Problem 6.

Consider the random vector: $\boldsymbol{X}$, the constant vector $\boldsymbol{b}$, and the constant matrix $\boldsymbol{A}$

## Question 6.1.

Then $\mathrm{D}(\boldsymbol{A}(\boldsymbol{b} + \boldsymbol{X}))$ equals

**1** □ $\boldsymbol{A}\mathrm{D}(\boldsymbol{X})$

**2** □ $\boldsymbol{b} + \boldsymbol{A}\mathrm{D}(\boldsymbol{X})$

**3** □ $\boldsymbol{b} + \boldsymbol{A}\mathrm{D}(\boldsymbol{X})\boldsymbol{A}'$

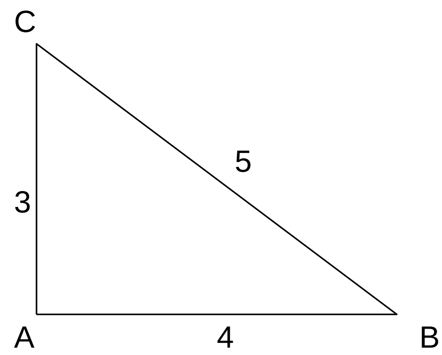
**4** □ $\boldsymbol{A}\mathrm{D}(\boldsymbol{X})\boldsymbol{A}'$

**5** □ $\boldsymbol{A}\boldsymbol{b}\boldsymbol{A}' + \boldsymbol{A}\mathrm{D}(\boldsymbol{X})\boldsymbol{A}'$

**6** □ Don't know.

# Problem 7.

Consider the following spatial layout of data-sampling points $A, B, C$. The distances between the data points is also indicated.




It is assumed that each of the 3 observations have equal variance $\sigma^2$, such that: $\begin{bmatrix} X_A \\ X_B \\ X_C \end{bmatrix} \in \mathrm{N}(\boldsymbol{\mu}, \sigma^2 \boldsymbol{\Sigma})$

Furthermore, the spatial correlation between any pair of data points is assumed to be inversely proportional to the distance between the points.

## Question 7.1.

Then $\Sigma$ equals:

**1** ☐ $\begin{bmatrix} \sigma^2 & \frac{1}{4} & \frac{1}{3} \\ \frac{1}{4} & \sigma^2 & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{5} & \sigma^2 \end{bmatrix}$

**2** ☐ $\begin{bmatrix} 1 & \frac{1}{4} & \frac{1}{3} \\ \frac{1}{4} & 1 & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{5} & 1 \end{bmatrix}$

**3** ☐ $\begin{bmatrix} 1 & \frac{1}{16} & \frac{1}{9} \\ \frac{1}{16} & 1 & \frac{1}{25} \\ \frac{1}{9} & \frac{1}{25} & 1 \end{bmatrix}$

**4** ☐ $\begin{bmatrix} 1 & 4 & 3 \\ 4 & 1 & 5 \\ 3 & 5 & 1 \end{bmatrix}$

**5** ☐ $\begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{5} \end{bmatrix}$

**6** ☐ Don't know.

# Problem 8.

Below is given values for four polynomials for a certain set of numbers.

| $\xi_a$ | $\xi_b$ | $\xi_c$ | $\xi_d$ |
|---|---|---|---|
| 1 | -2 | -1 | -3 |
| -1 | 1 | 3 | -1 |
| -1 | -1 | -3 | 1 |
| 1 | 2 | 1 | 3 |

## Question 8.1.

In that case the following polynomials are pairwise orthogonal, namely:

**1** □ $\xi_a$, $\xi_c$, $\xi_d$

**2** □ $\xi_b$, $\xi_c$, $\xi_d$

**3** □ $\xi_a$, $\xi_b$, $\xi_d$

**4** □ $\xi_a$, $\xi_b$, $\xi_c$

**5** □ $\xi_a$, $\xi_b$, $\xi_c$, $\xi_d$

**6** □ Don't know.

```
/* enclabc.sas   Crtd: 28-11-08 13:09 by BKE. Updt: 02-12-08 23:05 */
/* Purpose: Enclosures for 02409 exam on 9 December 2008 */

*** start of SAS-statements for Enclosure A ***;

title1 'Enclosure A - Linthurst''s Spartina Alterniflora data';

title2 'Print of data';
proc print data=stat2.linthurst;
run;

title2 'Correlations conditioned on pH';
proc corr data=stat2.linthurst;
var Biomass Salinity K Na Zn;
partial pH;
run;

title2 'Factor analysis';
proc factor data=stat2.linthurst corr nfactors=3 rotate=varimax;
var Salinity pH K Na Zn;
run;

*** end of SAS-statements for Enclosure A ***;

*** start of SAS-statements for Enclosure B ***;

title1 'Enclosure B - Linthurst''s Spartina Alterniflora data - locations OI and S

data OIandSI;
set stat2.linthurst;
if Location = 'SM' then delete;

title2 'Print of data';
proc print data=OIandSI;
run;

title2 'Linear discriminant analysis - two variables';
proc discrim data=OIandSI simple wcov wcorr pcov pcorr pool=yes;
class Location;
var Salinity K;
run;

title2 'Linear discriminant analysis - one variable';
proc discrim data=OIandSI pool=yes;
class Location;
var Salinity;
run;

*** end of SAS-statements for Enclosure B ***;

*** start of SAS-statements for Enclosure C ***;

title1 'Enclosure C - Linthurst''s Spartina Alterniflora data';

title2 'Print of data';
proc print data=stat2.linthurst;
run;

title2 'Regression analysis - all 45 observations';
proc reg data=stat2.linthurst;
model Biomass=pH Na / r influence;
run;

data Except34;
set stat2.linthurst;
if _N_=34 then delete;
```



```
title2 'Regression analysis - 44 observations - Obs No 34 deleted';
proc reg data=Except34;
model Biomass=pH Na;
run;

*** end of SAS-statements for Enclosure C ***;
```

Enclosure A - Linthurst's Spartina Alterniflora data 1
Print of data

| Obs | Location | Type | Biomass | Salinity | pH | K | Na | Zn |
|---|---|---|---|---|---|---|---|---|
| 1 | OI | DVEG | 676 | 33 | 5.00 | 1441.67 | 35184.5 | 16.4524 |
| 2 | OI | DVEG | 516 | 35 | 4.75 | 1299.19 | 28170.4 | 13.9852 |
| 3 | OI | DVEG | 1052 | 32 | 4.20 | 1154.27 | 26455.0 | 15.3276 |
| 4 | OI | DVEG | 868 | 30 | 4.40 | 1045.15 | 25072.9 | 17.3128 |
| 5 | OI | DVEG | 1008 | 33 | 5.55 | 521.62 | 31664.2 | 22.3312 |
| 6 | OI | SHRT | 436 | 33 | 5.05 | 1273.02 | 25491.7 | 12.2778 |
| 7 | OI | SHRT | 544 | 36 | 4.25 | 1346.35 | 20877.3 | 17.8225 |
| 8 | OI | SHRT | 680 | 30 | 4.45 | 1253.88 | 25621.3 | 14.3516 |
| 9 | OI | SHRT | 640 | 38 | 4.75 | 1242.65 | 27587.3 | 13.6826 |
| 10 | OI | SHRT | 492 | 30 | 4.60 | 1281.95 | 26511.7 | 11.7566 |
| 11 | OI | TALL | 984 | 30 | 4.10 | 553.69 | 7886.5 | 9.8820 |
| 12 | OI | TALL | 1400 | 37 | 3.45 | 494.74 | 14596.0 | 16.6752 |
| 13 | OI | TALL | 1276 | 33 | 3.45 | 525.97 | 9826.8 | 12.3730 |
| 14 | OI | TALL | 1736 | 36 | 4.10 | 571.14 | 11978.4 | 9.4058 |
| 15 | OI | TALL | 1004 | 30 | 3.50 | 408.64 | 10368.6 | 14.9302 |
| 16 | SI | DVEG | 396 | 30 | 3.25 | 646.65 | 17307.4 | 31.2865 |
| 17 | SI | DVEG | 352 | 27 | 3.35 | 514.03 | 12822.0 | 30.1652 |
| 18 | SI | DVEG | 328 | 29 | 3.20 | 350.73 | 8582.6 | 28.5901 |
| 19 | SI | DVEG | 392 | 34 | 3.35 | 496.29 | 12369.5 | 19.8795 |
| 20 | SI | DVEG | 236 | 36 | 3.30 | 580.92 | 14731.9 | 18.5056 |
| 21 | SI | SHRT | 392 | 30 | 3.25 | 535.82 | 15060.6 | 22.1344 |
| 22 | SI | SHRT | 268 | 28 | 3.25 | 490.34 | 11056.3 | 28.6101 |
| 23 | SI | SHRT | 252 | 31 | 3.20 | 552.39 | 8118.9 | 23.1908 |
| 24 | SI | SHRT | 236 | 31 | 3.20 | 661.32 | 13009.5 | 24.6917 |
| 25 | SI | SHRT | 340 | 35 | 3.35 | 672.15 | 15003.7 | 22.6758 |
| 26 | SI | TALL | 2436 | 29 | 7.10 | 528.65 | 10225.0 | 0.3729 |
| 27 | SI | TALL | 2216 | 35 | 7.35 | 563.13 | 8024.2 | 0.2703 |
| 28 | SI | TALL | 2096 | 35 | 7.45 | 497.96 | 10393.0 | 0.3205 |
| 29 | SI | TALL | 1660 | 30 | 7.45 | 458.38 | 8711.6 | 0.2648 |
| 30 | SI | TALL | 2272 | 30 | 7.40 | 498.25 | 10239.6 | 0.2105 |
| 31 | SM | DVEG | 824 | 26 | 4.85 | 936.26 | 20436.0 | 18.9875 |
| 32 | SM | DVEG | 1196 | 29 | 4.60 | 894.79 | 12519.9 | 20.9687 |
| 33 | SM | DVEG | 1960 | 25 | 5.20 | 941.36 | 18979.0 | 23.9841 |
| 34 | SM | DVEG | 2080 | 26 | 4.75 | 1038.79 | 22986.1 | 19.9727 |
| 35 | SM | DVEG | 1764 | 26 | 5.20 | 898.05 | 11704.5 | 21.3864 |
| 36 | SM | SHRT | 412 | 25 | 4.55 | 989.87 | 17721.0 | 23.7063 |
| 37 | SM | SHRT | 416 | 26 | 3.95 | 951.28 | 16485.2 | 30.5589 |
| 38 | SM | SHRT | 504 | 26 | 3.70 | 939.83 | 17101.3 | 26.8415 |
| 39 | SM | SHRT | 492 | 27 | 3.75 | 925.42 | 17849.0 | 27.7292 |
| 40 | SM | SHRT | 636 | 27 | 4.15 | 954.11 | 16949.6 | 21.5699 |
| 41 | SM | TALL | 1756 | 24 | 5.60 | 720.72 | 11344.6 | 19.6531 |
| 42 | SM | TALL | 1232 | 27 | 5.35 | 782.09 | 14752.4 | 20.3295 |
| 43 | SM | TALL | 1400 | 26 | 5.50 | 773.30 | 13649.8 | 19.5880 |
| 44 | SM | TALL | 1620 | 28 | 5.50 | 829.26 | 14533.0 | 20.1328 |
| 45 | SM | TALL | 1560 | 28 | 5.40 | 856.96 | 16892.2 | 19.2420 |



Enclosure A - Linthurst's Spartina Alterniflora data 2
Correlations conditioned on pH

The CORR Procedure

1 Partial Variables: pH
5 Variables: Biomass Salinity K Na Zn

Simple Statistics

| Variable | N | Mean | Std Dev | Sum | Minimum | Maximum |
|---|---|---|---|---|---|---|
| pH | 45 | 4.60222 | 1.24699 | 207.10000 | 3.20000 | 7.45000 |
| Biomass | 45 | 1001 | 660.07856 | 45036 | 236.00000 | 2436 |
| Salinity | 45 | 30.26667 | 3.71973 | 1362 | 24.00000 | 38.00000 |
| K | 45 | 797.62289 | 297.60234 | 35893 | 350.73000 | 1442 |
| Na | 45 | 16597 | 6882 | 746852 | 7887 | 35185 |
| Zn | 45 | 17.87524 | 8.27981 | 804.38580 | 0.21050 | 31.28650 |

Simple Statistics

| Variable | Partial Variance | Partial Std Dev |
|---|---|---|
| pH | | |
| Biomass | 178618 | 422.63225 |
| Salinity | 14.12083 | 3.75777 |
| K | 90593 | 300.98729 |
| Na | 48400365 | 6957 |
| Zn | 33.56475 | 5.79351 |

Pearson Partial Correlation Coefficients, N = 45
Prob > |r| under H0: Partial Rho=0

| | Biomass | Salinity | K | Na | Zn |
|---|---|---|---|---|---|
| Biomass | 1.00000 | -0.10034 | -0.34690 | -0.38398 | -0.14918 |
| | | 0.5170 | 0.0211 | 0.0101 | 0.3338 |
| Salinity | -0.10034 | 1.00000 | -0.01968 | 0.16066 | -0.66285 |
| | 0.5170 | | 0.8991 | 0.2975 | <.0001 |
| K | -0.34690 | -0.01968 | 1.00000 | 0.79353 | 0.12651 |
| | 0.0211 | 0.8991 | | <.0001 | 0.4132 |
| Na | -0.38398 | 0.16066 | 0.79353 | 1.00000 | 0.12992 |
| | 0.0101 | 0.2975 | <.0001 | | 0.4006 |
| Zn | -0.14918 | -0.66285 | 0.12651 | 0.12992 | 1.00000 |
| | 0.3338 | <.0001 | 0.4132 | 0.4006 | |

Enclosure A – Linthurst's Spartina Alterniflora data 3
Factor analysis

The FACTOR Procedure

Correlations

| | Salinity | pH | K | Na | Zn |
|---|---|---|---|---|---|
| Salinity | 1.00000 | -0.05133 | -0.02063 | 0.16227 | -0.42083 |
| pH | -0.05133 | 1.00000 | 0.01923 | -0.03772 | -0.72217 |
| K | -0.02063 | 0.01923 | 1.00000 | 0.79210 | 0.07361 |
| Na | 0.16227 | -0.03772 | 0.79210 | 1.00000 | 0.11705 |
| Zn | -0.42083 | -0.72217 | 0.07361 | 0.11705 | 1.00000 |

/><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><>

Enclosure A – Linthurst's Spartina Alterniflora data 4
Factor analysis

The FACTOR Procedure
Initial Factor Method: Principal Components

Prior Communality Estimates: ONE

Eigenvalues of the Correlation Matrix: Total = 5 Average = 1

| | Eigenvalue | Difference | Proportion | Cumulative |
|---|---|---|---|---|
| 1 | 1.88382449 | 0.14149384 | 0.3768 | 0.3768 |
| 2 | 1.74233065 | 0.68816912 | 0.3485 | 0.7252 |
| 3 | 1.05416153 | 0.84036735 | 0.2108 | 0.9361 |
| 4 | 0.21379417 | 0.10790501 | 0.0428 | 0.9788 |
| 5 | 0.10588916 | | 0.0212 | 1.0000 |

3 factors will be retained by the NFACTOR criterion.

Factor Pattern

| | Factor1 | Factor2 | Factor3 |
|---|---|---|---|
| Salinity | -0.20953 | 0.43826 | 0.86414 |
| pH | -0.61158 | 0.55173 | -0.51738 |
| K | 0.66595 | 0.65111 | -0.19569 |
| Na | 0.68318 | 0.66544 | 0.02439 |
| Zn | 0.74543 | -0.61571 | -0.02911 |

Variance Explained by Each Factor

| Factor1 | Factor2 | Factor3 |
|---|---|---|
| 1.8838245 | 1.7423306 | 1.0541615 |

Final Communality Estimates: Total = 4.680317

| Salinity | pH | K | Na | Zn |
|---|---|---|---|---|
| 0.98271746 | 0.94611993 | 0.90572598 | 0.91014691 | 0.93560638 |



Enclosure A – Linthurst's Spartina Alterniflora data 5
Factor analysis

The FACTOR Procedure
Rotation Method: Varimax

Orthogonal Transformation Matrix

| | 1 | 2 | 3 |
|---|---|---|---|
| 1 | 0.70085 | -0.68462 | -0.20026 |
| 2 | 0.70168 | 0.61120 | 0.36617 |
| 3 | -0.12828 | -0.39715 | 0.90874 |

Rotated Factor Pattern

| | Factor1 | Factor2 | Factor3 |
|---|---|---|---|
| Salinity | 0.04981 | 0.06812 | 0.98772 |
| pH | 0.02488 | 0.96140 | -0.14566 |
| K | 0.94870 | 0.01976 | -0.07278 |
| Na | 0.94261 | -0.07068 | 0.12902 |
| Zn | 0.09414 | -0.87510 | -0.40118 |

Variance Explained by Each Factor

| Factor1 | Factor2 | Factor3 |
|---|---|---|
| 1.8005062 | 1.7001063 | 1.1797041 |

Final Communality Estimates: Total = 4.680317

| Salinity | pH | K | Na | Zn |
|---|---|---|---|---|
| 0.98271746 | 0.94611993 | 0.90572598 | 0.91014691 | 0.93560638 |

```
Enclosure B - Linthurst's Spartina Alterniflora data - locations OI and SI onl 6
                              Print of data

Obs   Location    Type   Biomass   Salinity    pH        K          Na         Zn

  1      OI       DVEG      676        33     5.00   1441.67    35184.5    16.4524
  2      OI       DVEG      516        35     4.75   1299.19    28170.4    13.9852
  3      OI       DVEG     1052        32     4.20   1154.27    26455.0    15.3276
  4      OI       DVEG      868        30     4.40   1045.15    25072.9    17.3128
  5      OI       DVEG     1008        33     5.55    521.62    31664.2    22.3312
  6      OI       SHRT      436        33     5.05   1273.02    25491.7    12.2778
  7      OI       SHRT      544        36     4.25   1346.35    20877.3    17.8225
  8      OI       SHRT      680        30     4.45   1253.88    25621.3    14.3516
  9      OI       SHRT      640        38     4.75   1242.65    27587.3    13.6826
 10      OI       SHRT      492        30     4.60   1281.95    26511.7    11.7566
 11      OI       TALL      984        30     4.10    553.69     7886.5     9.8820
 12      OI       TALL     1400        37     3.45    494.74    14596.0    16.6752
 13      OI       TALL     1276        33     3.45    525.97     9826.8    12.3730
 14      OI       TALL     1736        36     4.10    571.14    11978.4     9.4058
 15      OI       TALL     1004        30     3.50    408.64    10368.6    14.9302
 16      SI       DVEG      396        30     3.25    646.65    17307.4    31.2865
 17      SI       DVEG      352        27     3.35    514.03    12822.0    30.1652
 18      SI       DVEG      328        29     3.20    350.73     8582.6    28.5901
 19      SI       DVEG      392        34     3.35    496.29    12369.5    19.8795
 20      SI       DVEG      236        36     3.30    580.92    14731.9    18.5056
 21      SI       SHRT      392        30     3.25    535.82    15060.6    22.1344
 22      SI       SHRT      268        28     3.25    490.34    11056.3    28.6101
 23      SI       SHRT      252        31     3.20    552.39     8118.9    23.1908
 24      SI       SHRT      236        31     3.20    661.32    13009.5    24.6917
 25      SI       SHRT      340        35     3.35    672.15    15003.7    22.6758
 26      SI       TALL     2436        29     7.10    528.65    10225.0     0.3729
 27      SI       TALL     2216        35     7.35    563.13     8024.2     0.2703
 28      SI       TALL     2096        35     7.45    497.96    10393.0     0.3205
 29      SI       TALL     1660        30     7.45    458.38     8711.6     0.2648
 30      SI       TALL     2272        30     7.40    498.25    10239.6     0.2105

/><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>>
Enclosure B - Linthurst's Spartina Alterniflora data - locations OI and SI onl 7
                 Linear discriminant analysis - two variables

                            The DISCRIM Procedure

            Observations      30          DF Total               29
            Variables          2          DF Within Classes      28
            Classes            2          DF Between Classes      1


                           Class Level Information

             Variable                                                  Prior
  Location   Name         Frequency       Weight     Proportion    Probability

  OI         OI                  15      15.0000       0.500000       0.500000
  SI         SI                  15      15.0000       0.500000       0.500000
```



```
Enclosure B - Linthurst's Spartina Alterniflora data - locations OI and SI onl 8
                 Linear discriminant analysis - two variables

                            The DISCRIM Procedure
                       Within-Class Covariance Matrices

                          Location = OI,      DF = 14

                  Variable           Salinity                  K

                  Salinity             7.7810            56.9951
                  K                   56.9951        151893.8632

--------------------------------------------------------------------------------


                          Location = SI,      DF = 14

                  Variable           Salinity                  K

                  Salinity           8.380952          87.078095
                  K                 87.078095        6923.210264

--------------------------------------------------------------------------------

/><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>>
Enclosure B - Linthurst's Spartina Alterniflora data - locations OI and SI onl 9
                 Linear discriminant analysis - two variables

                            The DISCRIM Procedure

               Pooled Within-Class Covariance Matrix,      DF = 28

                  Variable           Salinity                  K

                  Salinity            8.08095           72.03660
                  K                  72.03660        79408.53672
```

```
Enclosure B - Linthurst's Spartina Alterniflora data - locations OI and SI on 10
                  Linear discriminant analysis - two variables

                           The DISCRIM Procedure
            Within-Class Correlation Coefficients  /  Pr > |r|

                              Location = OI

                   Variable       Salinity            K

                   Salinity        1.00000      0.05243
                                                 0.8528

                   K               0.05243      1.00000
                                    0.8528

--------------------------------------------------------------------------------

                              Location = SI

                   Variable       Salinity            K

                   Salinity        1.00000      0.36150
                                                 0.1855

                   K               0.36150      1.00000
                                    0.1855

--------------------------------------------------------------------------------
/><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>>
Enclosure B - Linthurst's Spartina Alterniflora data - locations OI and SI on 11
                  Linear discriminant analysis - two variables

                           The DISCRIM Procedure

         Pooled Within-Class Correlation Coefficients  /  Pr > |r|

                   Variable       Salinity            K

                   Salinity        1.00000      0.08993
                                                 0.6427

                   K               0.08993      1.00000
                                    0.6427
```



```
Enclosure B - Linthurst's Spartina Alterniflora data - locations OI and SI on 12
                  Linear discriminant analysis - two variables

                           The DISCRIM Procedure
                             Simple Statistics

                               Total-Sample

                                                                       Standard
   Variable          N          Sum          Mean      Variance       Deviation

   Salinity         30    966.00000      32.20000       8.57931          2.9290
   K                30        22461     748.69800        123265        351.0916
--------------------------------------------------------------------------------


                              Location = OI

                                                                       Standard
   Variable          N          Sum          Mean      Variance       Deviation

   Salinity         15    496.00000      33.06667       7.78095          2.7894
   K                15        14414     960.92867        151894        389.7356
--------------------------------------------------------------------------------


                              Location = SI

                                                                       Standard
   Variable          N          Sum          Mean      Variance       Deviation

   Salinity         15    470.00000      31.33333       8.38095          2.8950
   K                15         8047     536.46733          6923         83.2058
--------------------------------------------------------------------------------


                   Pooled Covariance Matrix Information

                                      Natural Log of the
                       Covariance     Determinant of the
                      Matrix Rank      Covariance Matrix

                                2               13.36375
```

Enclosure B - Linthurst's Spartina Alterniflora data - locations OI and SI on 13
Linear discriminant analysis - two variables

The DISCRIM Procedure

Pairwise Generalized Squared Distances Between Groups

$$D^2(i|j) = (\bar{X}_i - \bar{X}_j)' \, COV^{-1} (\bar{X}_i - \bar{X}_j)$$

Generalized Squared Distance to Location

| From Location | OI | SI |
|---|---|---|
| OI | 0 | 2.49566 |
| SI | 2.49566 | 0 |

Linear Discriminant Function

$$\text{Constant} = -.5 \, \bar{X}_j' \, COV^{-1} \bar{X}_j \qquad \text{Coefficient Vector} = COV^{-1} \bar{X}_j$$

Linear Discriminant Function for Location

| Variable | OI | SI |
|---|---|---|
| Constant | -70.47019 | -61.16618 |
| Salinity | 4.01653 | 3.84833 |
| K | 0.00846 | 0.00326 |



Enclosure B - Linthurst's Spartina Alterniflora data - locations OI and SI on 14
Linear discriminant analysis - two variables

The DISCRIM Procedure
Classification Summary for Calibration Data: WORK.OIANDSI
Resubstitution Summary using Linear Discriminant Function

Generalized Squared Distance Function

$$D_j^2(X) = (X-\bar{X}_j)' \, COV^{-1} (X-\bar{X}_j)$$

Posterior Probability of Membership in Each Location

$$Pr(j|X) = \exp(-.5 \, D_j^2(X)) \, / \, SUM_k \exp(-.5 \, D_k^2(X))$$

Number of Observations and Percent Classified into Location

| From Location | OI | SI | Total |
|---|---|---|---|
| OI | 9 | 6 | 15 |
| | 60.00 | 40.00 | 100.00 |
| SI | 1 | 14 | 15 |
| | 6.67 | 93.33 | 100.00 |
| Total | 10 | 20 | 30 |
| | 33.33 | 66.67 | 100.00 |
| Priors | 0.5 | 0.5 | |

Error Count Estimates for Location

| | OI | SI | Total |
|---|---|---|---|
| Rate | 0.4000 | 0.0667 | 0.2333 |
| Priors | 0.5000 | 0.5000 | |

Enclosure B – Linthurst's Spartina Alterniflora data – locations OI and SI on 15
Linear discriminant analysis – one variable

The DISCRIM Procedure

| | | | |
|---|---|---|---|
| Observations | 30 | DF Total | 29 |
| Variables | 1 | DF Within Classes | 28 |
| Classes | 2 | DF Between Classes | 1 |

Class Level Information

| Location | Variable Name | Frequency | Weight | Proportion | Prior Probability |
|---|---|---|---|---|---|
| OI | OI | 15 | 15.0000 | 0.500000 | 0.500000 |
| SI | SI | 15 | 15.0000 | 0.500000 | 0.500000 |

Pooled Covariance Matrix Information

| Covariance Matrix Rank | Natural Log of the Determinant of the Covariance Matrix |
|---|---|
| 1 | 2.08951 |

/><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>>

Enclosure B – Linthurst's Spartina Alterniflora data – locations OI and SI on 16
Linear discriminant analysis – one variable

The DISCRIM Procedure

Pairwise Generalized Squared Distances Between Groups

$$D^2(i|j) = (\bar{X}_i - \bar{X}_j)' \, COV^{-1} \, (\bar{X}_i - \bar{X}_j)$$

Generalized Squared Distance to Location

| From Location | OI | SI |
|---|---|---|
| OI | 0 | 0.37179 |
| SI | 0.37179 | 0 |

Linear Discriminant Function

$$\text{Constant} = -.5 \, \bar{X}_j' \, COV^{-1} \, \bar{X}_j \qquad \text{Coefficient Vector} = COV^{-1} \, \bar{X}_j$$

Linear Discriminant Function for Location

| Variable | OI | SI |
|---|---|---|
| Constant | -67.65319 | -60.74642 |
| Salinity | 4.09193 | 3.87743 |



Enclosure B – Linthurst's Spartina Alterniflora data – locations OI and SI on 17
Linear discriminant analysis – one variable

The DISCRIM Procedure
Classification Summary for Calibration Data: WORK.OIANDSI
Resubstitution Summary using Linear Discriminant Function

Generalized Squared Distance Function

$$D_j^2(X) = (X-\bar{X}_j)' \, COV^{-1} \, (X-\bar{X}_j)$$

Posterior Probability of Membership in Each Location

$$Pr(j|X) = \exp(-.5 \, D_j^2(X)) \, / \, \text{SUM}_k \exp(-.5 \, D_k^2(X))$$

Number of Observations and Percent Classified into Location

| From Location | OI | SI | Total |
|---|---|---|---|
| OI | 9<br>60.00 | 6<br>40.00 | 15<br>100.00 |
| SI | 5<br>33.33 | 10<br>66.67 | 15<br>100.00 |
| Total | 14<br>46.67 | 16<br>53.33 | 30<br>100.00 |
| Priors | 0.5 | 0.5 | |

Error Count Estimates for Location

| | OI | SI | Total |
|---|---|---|---|
| Rate | 0.4000 | 0.3333 | 0.3667 |
| Priors | 0.5000 | 0.5000 | |

Enclosure C – Linthurst's Spartina Alterniflora data 18
Print of data

| Obs | Location | Type | Biomass | Salinity | pH | K | Na | Zn |
|---|---|---|---|---|---|---|---|---|
| 1 | OI | DVEG | 676 | 33 | 5.00 | 1441.67 | 35184.5 | 16.4524 |
| 2 | OI | DVEG | 516 | 35 | 4.75 | 1299.19 | 28170.4 | 13.9852 |
| 3 | OI | DVEG | 1052 | 32 | 4.20 | 1154.27 | 26455.0 | 15.3276 |
| 4 | OI | DVEG | 868 | 30 | 4.40 | 1045.15 | 25072.9 | 17.3128 |
| 5 | OI | DVEG | 1008 | 33 | 5.55 | 521.62 | 31664.2 | 22.3312 |
| 6 | OI | SHRT | 436 | 33 | 5.05 | 1273.02 | 25491.7 | 12.2778 |
| 7 | OI | SHRT | 544 | 36 | 4.25 | 1346.35 | 20877.3 | 17.8225 |
| 8 | OI | SHRT | 680 | 30 | 4.45 | 1253.88 | 25621.3 | 14.3516 |
| 9 | OI | SHRT | 640 | 38 | 4.75 | 1242.65 | 27587.3 | 13.6826 |
| 10 | OI | SHRT | 492 | 30 | 4.60 | 1281.95 | 26511.7 | 11.7566 |
| 11 | OI | TALL | 984 | 30 | 4.10 | 553.69 | 7886.5 | 9.8820 |
| 12 | OI | TALL | 1400 | 37 | 3.45 | 494.74 | 14596.0 | 16.6752 |
| 13 | OI | TALL | 1276 | 33 | 3.45 | 525.97 | 9826.8 | 12.3730 |
| 14 | OI | TALL | 1736 | 36 | 4.10 | 571.14 | 11978.4 | 9.4058 |
| 15 | OI | TALL | 1004 | 30 | 3.50 | 408.64 | 10368.6 | 14.9302 |
| 16 | SI | DVEG | 396 | 30 | 3.25 | 646.65 | 17307.4 | 31.2865 |
| 17 | SI | DVEG | 352 | 27 | 3.35 | 514.03 | 12822.0 | 30.1652 |
| 18 | SI | DVEG | 328 | 29 | 3.20 | 350.73 | 8582.6 | 28.5901 |
| 19 | SI | DVEG | 392 | 34 | 3.35 | 496.29 | 12369.5 | 19.8795 |
| 20 | SI | DVEG | 236 | 36 | 3.30 | 580.92 | 14731.9 | 18.5056 |
| 21 | SI | SHRT | 392 | 30 | 3.25 | 535.82 | 15060.6 | 22.1344 |
| 22 | SI | SHRT | 268 | 28 | 3.25 | 490.34 | 11056.3 | 28.6101 |
| 23 | SI | SHRT | 252 | 31 | 3.20 | 552.39 | 8118.9 | 23.1908 |
| 24 | SI | SHRT | 236 | 31 | 3.20 | 661.32 | 13009.5 | 24.6917 |
| 25 | SI | SHRT | 340 | 35 | 3.35 | 672.15 | 15003.7 | 22.6758 |
| 26 | SI | TALL | 2436 | 29 | 7.10 | 528.65 | 10225.0 | 0.3729 |
| 27 | SI | TALL | 2216 | 35 | 7.35 | 563.13 | 8024.2 | 0.2703 |
| 28 | SI | TALL | 2096 | 35 | 7.45 | 497.96 | 10393.0 | 0.3205 |
| 29 | SI | TALL | 1660 | 30 | 7.45 | 458.38 | 8711.6 | 0.2648 |
| 30 | SI | TALL | 2272 | 30 | 7.40 | 498.25 | 10239.6 | 0.2105 |
| 31 | SM | DVEG | 824 | 26 | 4.85 | 936.26 | 20436.0 | 18.9875 |
| 32 | SM | DVEG | 1196 | 29 | 4.60 | 894.79 | 12519.9 | 20.9687 |
| 33 | SM | DVEG | 1960 | 25 | 5.20 | 941.36 | 18979.0 | 23.9841 |
| 34 | SM | DVEG | 2080 | 26 | 4.75 | 1038.79 | 22986.1 | 19.9727 |
| 35 | SM | DVEG | 1764 | 26 | 5.20 | 898.05 | 11704.5 | 21.3864 |
| 36 | SM | SHRT | 412 | 25 | 4.55 | 989.87 | 17721.0 | 23.7063 |
| 37 | SM | SHRT | 416 | 26 | 3.95 | 951.28 | 16485.2 | 30.5589 |
| 38 | SM | SHRT | 504 | 26 | 3.70 | 939.83 | 17101.3 | 26.8415 |
| 39 | SM | SHRT | 492 | 27 | 3.75 | 925.42 | 17849.0 | 27.7292 |
| 40 | SM | SHRT | 636 | 27 | 4.15 | 954.11 | 16949.6 | 21.5699 |
| 41 | SM | TALL | 1756 | 24 | 5.60 | 720.72 | 11344.6 | 19.6531 |
| 42 | SM | TALL | 1232 | 27 | 5.35 | 782.09 | 14752.4 | 20.3295 |
| 43 | SM | TALL | 1400 | 26 | 5.50 | 773.30 | 13649.8 | 19.5880 |
| 44 | SM | TALL | 1620 | 28 | 5.50 | 829.26 | 14533.0 | 20.1328 |
| 45 | SM | TALL | 1560 | 28 | 5.40 | 856.96 | 16892.2 | 19.2420 |



Enclosure C – Linthurst's Spartina Alterniflora data 19
Regression analysis – all 45 observations

The REG Procedure
Model: MODEL1
Dependent Variable: Biomass

| | |
|---|---|
| Number of Observations Read | 45 |
| Number of Observations Used | 45 |

Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 2 | 12622789 | 6311394 | 40.48 | <.0001 |
| Error | 42 | 6548174 | 155909 | | |
| Corrected Total | 44 | 19170963 | | | |

| | | | |
|---|---|---|---|
| Root MSE | 394.85302 | R-Square | 0.6584 |
| Dependent Mean | 1000.80000 | Adj R-Sq | 0.6422 |
| Coeff Var | 39.45374 | | |

Parameter Estimates

| Variable | DF | Parameter Estimate | Standard Error | t Value | Pr > \|t\| |
|---|---|---|---|---|---|
| Intercept | 1 | -475.72558 | 273.52269 | -1.74 | 0.0893 |
| pH | 1 | 404.94819 | 47.76984 | 8.48 | <.0001 |
| Na | 1 | -0.02333 | 0.00866 | -2.70 | 0.0101 |

Enclosure C – Linthurst's Spartina Alterniflora data 20
Regression analysis – all 45 observations

The REG Procedure
Model: MODEL1
Dependent Variable: Biomass

Output Statistics

| Obs | Dependent Variable | Predicted Value | Std Error Mean Predict | Residual | Std Error Residual | Student Residual |
|---|---|---|---|---|---|---|
| 1 | 676.0000 | 728.2994 | 173.0289 | -52.2994 | 354.9 | -0.147 |
| 2 | 516.0000 | 790.6737 | 116.6292 | -274.6737 | 377.2 | -0.728 |
| 3 | 1052 | 607.9657 | 104.8359 | 444.0343 | 380.7 | 1.166 |
| 4 | 868.0000 | 721.1943 | 94.2690 | 146.8057 | 383.4 | 0.383 |
| 5 | 1008 | 1033 | 151.5494 | -25.1357 | 364.6 | -0.0689 |
| 6 | 436.0000 | 974.6417 | 99.8677 | -538.6417 | 382.0 | -1.410 |
| 7 | 544.0000 | 758.3190 | 71.2276 | -214.3190 | 388.4 | -0.552 |
| 8 | 680.0000 | 728.6497 | 97.8559 | -48.6497 | 382.5 | -0.127 |
| 9 | 640.0000 | 804.2751 | 112.3122 | -164.2751 | 378.5 | -0.434 |
| 10 | 492.0000 | 768.6224 | 104.0595 | -276.6224 | 380.9 | -0.726 |
| 11 | 984.0000 | 1001 | 99.2980 | -16.6010 | 382.2 | -0.0434 |
| 12 | 1400 | 580.8784 | 82.8612 | 819.1216 | 386.1 | 2.122 |
| 13 | 1276 | 692.1251 | 100.8508 | 583.8749 | 381.8 | 1.529 |
| 14 | 1736 | 905.1530 | 75.5669 | 830.8470 | 387.6 | 2.144 |
| 15 | 1004 | 699.7344 | 96.7311 | 304.2656 | 382.8 | 0.795 |
| 16 | 396.0000 | 436.6425 | 87.4361 | -40.6425 | 385.1 | -0.106 |
| 17 | 352.0000 | 581.7640 | 90.8719 | -229.7640 | 384.3 | -0.598 |
| 18 | 328.0000 | 619.9103 | 114.5134 | -291.9103 | 377.9 | -0.772 |
| 19 | 392.0000 | 592.3191 | 92.4480 | -200.3191 | 383.9 | -0.522 |
| 20 | 236.0000 | 516.9662 | 87.5820 | -280.9662 | 385.0 | -0.730 |
| 21 | 392.0000 | 489.0515 | 88.7624 | -97.0515 | 384.7 | -0.252 |
| 22 | 268.0000 | 582.4561 | 100.8485 | -314.4561 | 381.8 | -0.824 |
| 23 | 252.0000 | 630.7266 | 117.0747 | -378.7266 | 377.1 | -1.004 |
| 24 | 236.0000 | 516.6482 | 95.2490 | -280.6482 | 383.2 | -0.732 |
| 25 | 340.0000 | 530.8736 | 85.4120 | -190.8736 | 385.5 | -0.495 |
| 26 | 2436 | 2161 | 142.2901 | 275.1025 | 368.3 | 0.747 |
| 27 | 2216 | 2313 | 159.5760 | -97.4706 | 361.2 | -0.270 |
| 28 | 2096 | 2299 | 155.8941 | -202.7106 | 362.8 | -0.559 |
| 29 | 1660 | 2338 | 161.0225 | -677.9311 | 360.5 | -1.880 |
| 30 | 2272 | 2282 | 154.2708 | -10.0414 | 363.5 | -0.0276 |
| 31 | 824.0000 | 1012 | 68.8377 | -187.5817 | 388.8 | -0.482 |
| 32 | 1196 | 1095 | 68.6295 | 101.0039 | 388.8 | 0.260 |
| 33 | 1960 | 1187 | 68.9177 | 772.7004 | 388.8 | 1.987 |
| 34 | 2080 | 911.6030 | 81.2538 | 1168 | 386.4 | 3.024 |
| 35 | 1764 | 1357 | 77.3420 | 407.0149 | 387.2 | 1.051 |
| 36 | 412.0000 | 953.4275 | 59.6969 | -541.4275 | 390.3 | -1.387 |
| 37 | 416.0000 | 739.2849 | 66.6226 | -323.2849 | 389.2 | -0.831 |
| 38 | 504.0000 | 623.6767 | 72.9866 | -119.6767 | 388.0 | -0.308 |
| 39 | 492.0000 | 626.4832 | 72.1539 | -134.4832 | 388.2 | -0.346 |
| 40 | 636.0000 | 809.4419 | 62.7349 | -173.4419 | 389.8 | -0.445 |
| 41 | 1756 | 1527 | 87.4039 | 228.6406 | 385.1 | 0.594 |
| 42 | 1232 | 1347 | 70.3736 | -114.6318 | 388.5 | -0.295 |
| 43 | 1400 | 1433 | 76.6287 | -33.0933 | 387.3 | -0.0854 |
| 44 | 1620 | 1412 | 74.6000 | 207.5083 | 387.7 | 0.535 |
| 45 | 1560 | 1317 | 70.2203 | 243.0339 | 388.6 | 0.625 |

Output Statistics

```
                            Cook's                Hat Diag      Cov
Obs   -2-1 0 1 2               D    RStudent          H      Ratio     DFFITS

  1  |      |      |       0.002    -0.1456     0.1920    1.3284    -0.0710
  2  |     *|      |       0.017    -0.7240     0.0872    1.1337    -0.2238
```



Enclosure C – Linthurst's Spartina Alterniflora data 21
Regression analysis – all 45 observations

The REG Procedure
Model: MODEL1
Dependent Variable: Biomass

Output Statistics

```
                             Cook's                Hat Diag      Cov
Obs   -2-1 0 1 2                D    RStudent          H      Ratio     DFFITS

  3  |      |**    |        0.034     1.1716     0.0705    1.0477     0.3226
  4  |      |      |        0.003     0.3789     0.0570    1.1281     0.0932
  5  |      |      |        0.000    -0.0681     0.1473    1.2603    -0.0283
  6  |    **|      |        0.045    -1.4273     0.0640    0.9929    -0.3731
  7  |     *|      |        0.003    -0.5472     0.0325    1.0871    -0.1004
  8  |      |      |        0.000    -0.1257     0.0614    1.1440    -0.0321
  9  |      |      |        0.006    -0.4297     0.0809    1.1539    -0.1275
 10  |     *|      |        0.013    -0.7221     0.0695    1.1122    -0.1973
 11  |      |      |        0.000    -0.0429     0.0632    1.1474    -0.0112
 12  |      |****  |        0.069     2.2186     0.0440    0.8003     0.4762
 13  |      |***   |        0.054     1.5551     0.0652    0.9683     0.4108
 14  |      |****  |        0.058     2.2445     0.0366    0.7882     0.4376
 15  |      |*     |        0.013     0.7913     0.0600    1.0928     0.1999
 16  |      |      |        0.000    -0.1043     0.0490    1.1295    -0.0237
 17  |     *|      |        0.007    -0.5933     0.0530    1.1063    -0.1403
 18  |     *|      |        0.018    -0.7687     0.0841    1.1244    -0.2330
 19  |     *|      |        0.005    -0.5173     0.0548    1.1153    -0.1246
 20  |     *|      |        0.009    -0.7256     0.0492    1.0881    -0.1651
 21  |      |      |        0.001    -0.2494     0.0505    1.1270    -0.0575
 22  |     *|      |        0.016    -0.8205     0.0652    1.0951    -0.2167
 23  |    **|      |        0.032    -1.0044     0.0879    1.0957    -0.3118
 24  |     *|      |        0.011    -0.7283     0.0582    1.0982    -0.1810
 25  |      |      |        0.004    -0.4906     0.0468    1.1081    -0.1087
 26  |      |*     |        0.028     0.7429     0.1299    1.1868     0.2870
 27  |      |      |        0.005    -0.2669     0.1633    1.2781    -0.1179
 28  |     *|      |        0.019    -0.5541     0.1559    1.2453    -0.2381
 29  |   ***|      |        0.235    -1.9414     0.1663    0.9904    -0.8671
 30  |      |      |        0.000    -0.0273     0.1526    1.2686    -0.0116
 31  |      |      |        0.002    -0.4780     0.0304    1.0903    -0.0846
 32  |      |      |        0.001     0.2569     0.0302    1.1031     0.0453
 33  |      |***   |        0.041     2.0630     0.0305    0.8244     0.3657
 34  |      |******|        0.135     ######     0.0423    0.5374     0.7103
 35  |      |**    |        0.015     1.0525     0.0384    1.0319     0.2102
 36  |    **|      |        0.015    -1.4031     0.0229    0.9557    -0.2146
 37  |     *|      |        0.007    -0.8275     0.0285    1.0528    -0.1417
 38  |      |      |        0.001    -0.3051     0.0342    1.1055    -0.0574
 39  |      |      |        0.001    -0.3428     0.0334    1.1026    -0.0637
 40  |      |      |        0.002    -0.4406     0.0252    1.0873    -0.0709
 41  |      |*     |        0.006     0.5891     0.0490    1.1021     0.1337
 42  |      |      |        0.001    -0.2918     0.0318    1.1033    -0.0529
 43  |      |      |        0.000    -0.0844     0.0377    1.1165    -0.0167
 44  |      |*     |        0.004     0.5306     0.0357    1.0921     0.1021
 45  |      |*     |        0.004     0.6209     0.0316    1.0793     0.1122
```

Output Statistics

| Obs | DFBETAS Intercept | pH | Na |
|---|---|---|---|
| 1 | 0.0379 | -0.0103 | -0.0663 |
| 2 | 0.0936 | -0.0208 | -0.1928 |
| 3 | -0.0582 | -0.0492 | 0.2604 |
| 4 | -0.0199 | -0.0068 | 0.0721 |

Enclosure C – Linthurst's Spartina Alterniflora data 22
Regression analysis – all 45 observations

The REG Procedure
Model: MODEL1
Dependent Variable: Biomass

Output Statistics

| Obs | DFBETAS Intercept | DFBETAS pH | DFBETAS Na |
|---|---|---|---|
| 5 | 0.0181 | -0.0094 | -0.0247 |
| 6 | 0.1783 | -0.0908 | -0.2907 |
| 7 | -0.0084 | 0.0217 | -0.0513 |
| 8 | 0.0081 | 0.0014 | -0.0256 |
| 9 | 0.0522 | -0.0121 | -0.1083 |
| 10 | 0.0662 | -0.0059 | -0.1627 |
| 11 | -0.0083 | 0.0030 | 0.0086 |
| 12 | 0.3886 | -0.3201 | -0.1114 |
| 13 | 0.3688 | -0.2332 | -0.2471 |
| 14 | 0.3164 | -0.1477 | -0.2367 |
| 15 | 0.1777 | -0.1130 | -0.1155 |
| 16 | -0.0169 | 0.0174 | -0.0010 |
| 17 | -0.1236 | 0.0943 | 0.0539 |
| 18 | -0.2164 | 0.1416 | 0.1462 |
| 19 | -0.1108 | 0.0825 | 0.0523 |
| 20 | -0.1373 | 0.1184 | 0.0348 |
| 21 | -0.0475 | 0.0422 | 0.0102 |
| 22 | -0.1988 | 0.1427 | 0.1083 |
| 23 | -0.2892 | 0.1858 | 0.2022 |
| 24 | -0.1617 | 0.1295 | 0.0638 |
| 25 | -0.0886 | 0.0768 | 0.0204 |
| 26 | -0.1109 | 0.2365 | -0.1022 |
| 27 | 0.0401 | -0.0949 | 0.0512 |
| 28 | 0.1062 | -0.2047 | 0.0742 |
| 29 | 0.3309 | -0.7187 | 0.3399 |
| 30 | 0.0050 | -0.0099 | 0.0038 |
| 31 | 0.0191 | -0.0161 | -0.0414 |
| 32 | 0.0214 | -0.0009 | -0.0233 |
| 33 | -0.1184 | 0.1557 | 0.1151 |
| 34 | -0.2086 | 0.0799 | 0.4858 |
| 35 | 0.0344 | 0.0733 | -0.1122 |
| 36 | -0.0335 | 0.0076 | -0.0346 |
| 37 | -0.0826 | 0.0663 | 0.0046 |
| 38 | -0.0360 | 0.0338 | -0.0022 |
| 39 | -0.0355 | 0.0356 | -0.0082 |
| 40 | -0.0325 | 0.0243 | -0.0025 |
| 41 | -0.0020 | 0.0703 | -0.0668 |
| 42 | 0.0059 | -0.0264 | 0.0110 |
| 43 | 0.0018 | -0.0091 | 0.0052 |
| 44 | -0.0174 | 0.0578 | -0.0222 |
| 45 | -0.0322 | 0.0611 | 0.0064 |

| | |
|---|---|
| Sum of Residuals | 0 |
| Sum of Squared Residuals | 6548174 |
| Predicted Residual SS (PRESS) | 7394089 |



Enclosure C – Linthurst's Spartina Alterniflora data 23
Regression analysis – 44 observations – Obs No 34 deleted

The REG Procedure
Model: MODEL1
Dependent Variable: Biomass

| | |
|---|---|
| Number of Observations Read | 44 |
| Number of Observations Used | 44 |

Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 2 | 12857163 | 6428582 | 51.45 | <.0001 |
| Error | 41 | 5122657 | 124943 | | |
| Corrected Total | 43 | 17979821 | | | |

| | | | |
|---|---|---|---|
| Root MSE | 353.47258 | R-Square | 0.7151 |
| Dependent Mean | 976.27273 | Adj R-Sq | 0.7012 |
| Coeff Var | 36.20634 | | |

Parameter Estimates

| Variable | DF | Parameter Estimate | Standard Error | t Value | Pr > \|t\| |
|---|---|---|---|---|---|
| Intercept | 1 | -424.64004 | 245.32425 | -1.73 | 0.0910 |
| pH | 1 | 401.52948 | 42.77556 | 9.39 | <.0001 |
| Na | 1 | -0.02709 | 0.00783 | -3.46 | 0.0013 |