# Ma 322: Biostatistics
# Homework Assignment 11

Prof. Wickerhauser

Due Friday, April 20th, 2012

Read Chapter 16, "Working with Multivariate Data," pages 288–318 of our text.

1. The following 40 ordered pairs $x = (x_1, x_2)$ are samples from a bivariate normal population:

| $x_1$ | $x_2$ | $x_1$ | $x_2$ | $x_1$ | $x_2$ | $x_1$ | $x_2$ |
|---|---|---|---|---|---|---|---|
| 2.864810 | −0.087901 | 2.388924 | −0.112396 | 2.404386 | 1.536228 | 0.980159 | −1.113963 |
| 0.579622 | 2.072845 | −1.170284 | 0.211460 | −1.153178 | 0.435754 | 0.739514 | 2.413948 |
| 1.384192 | 4.185621 | 2.157917 | 3.993882 | −2.040037 | −0.076255 | 1.189135 | −0.800904 |
| 3.015638 | 2.956750 | 1.360922 | 1.483508 | −0.156409 | 0.444964 | 1.827972 | 0.590482 |
| 0.852800 | 0.633167 | 0.258943 | 1.706435 | −0.467125 | −0.712590 | 2.863697 | −1.876853 |
| 1.744421 | 2.453734 | 1.788729 | −1.266549 | 2.108316 | −2.300278 | 1.364329 | 1.972333 |
| 4.754022 | 1.574119 | 2.610398 | −0.411356 | 1.432215 | 1.049123 | 1.041985 | 0.760463 |
| 0.469449 | 1.740265 | 0.090927 | 2.289402 | 1.998294 | 3.047970 | 2.124222 | 0.543565 |
| 1.226427 | 1.741965 | 2.167013 | 1.948388 | −0.963964 | −1.826650 | 1.367142 | 1.569296 |
| 1.122402 | −1.337069 | 1.074869 | 2.284006 | −0.124088 | 0.895195 | 1.873769 | 1.341474 |

---

(a) Estimate the population mean $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$ and the variance matrix $\Sigma^2$ in the bivariate normal density $\frac{1}{\sqrt{2\pi}\det A}\, e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-2}(x-\mu)}$ of this population.

(b) Compute the eigenvalues of the estimated matrix.

2. The following data gives the hypothetical concentrations of three amino acids in centipede hæmolymph (mg/100ml) labeled by gender:

| Male | | | Female | | |
|---|---|---|---|---|---|
| Alanine | Aspartic Acid | Tyrosine | Alanine | Aspartic Acid | Tyrosine |
| 7.0 | 17.0 | 19.7 | 7.3 | 17.4 | 22.5 |
| 7.3 | 17.2 | 20.3 | 7.7 | 19.8 | 24.9 |
| 8.0 | 19.3 | 22.6 | 8.2 | 20.2 | 26.1 |
| 8.1 | 19.8 | 23.7 | 8.3 | 22.6 | 27.5 |
| 7.9 | 18.4 | 22.0 | 6.4 | 23.4 | 28.1 |
| 6.4 | 15.1 | 18.1 | 7.1 | 21.3 | 25.8 |
| 6.6 | 15.9 | 18.7 | 6.4 | 22.1 | 26.9 |
| 8.0 | 18.2 | 21.5 | 8.6 | 18.8 | 25.5 |

(a) Perform three analyses of variance on the three amino acid concentrations individually to test whether their concentrations are the same in males and females.

(b) Using multivariate analysis of variance, analyze the three amino acid concentrations together to determine whether their concentrations are the same in males and females.

3. The following data is from a hypothetical experiment involving 10 male and 10 female birds. Half the birds of each sex were given a hormone treatment and half were not. Two measurements were then made on each bird: plasma calcium concentration (in mg/100 ml) and rate of evaporative water loss (in mg/min). Perform a two-factor bivariate Model I MANOVA on the data.

| Hormone Treatment | | | | No Hormone Treatment | | | |
|---|---|---|---|---|---|---|---|
| Female | | Male | | Female | | Male | |
| Plasma Calcium | Water Loss | Plasma Calcium | Water Loss | Plasma Calcium | Water Loss | Plasma Calcium | Water Loss |
| 16.5 | 76 | 14.5 | 80 | 39.1 | 71 | 32.0 | 65 |
| 18.4 | 71 | 11.0 | 72 | 26.2 | 70 | 23.8 | 69 |
| 12.7 | 64 | 10.8 | 77 | 21.3 | 63 | 28.8 | 67 |
| 14.0 | 66 | 14.3 | 69 | 35.8 | 59 | 25.0 | 56 |
| 12.8 | 69 | 10.0 | 74 | 40.2 | 60 | 29.3 | 52 |

4. For this problem, use the amino acid concentration data in Problem 2.

(a) Plot all pairs of amino acid concentrations on a $3 \times 3$ grid of graphs using the `R` command `pairs()`. Identify the plotted points by sex using "x" for males and "o" for females.

(b) Plot the 3-d scatterplot amino acid concentrations.

(Hint: `install.packages("scatterplot3d")`.)

(c) Find the principal components of the amino acid data and scree plot their importance. (Hint: `screeplot()` and `princomp()` are included in the standard `R` installation.)

(d) A centipede has the following amino acid concentrations in its hæmolymph:

| Amino Acid | Concentration (mg/100ml) |
|---|---|
| Alanine | 7.5 |
| Aspartic Acid | 18.1 |
| Tyrosine | 22.1 |

Use linear discriminant analysis to judge whether it is likelier to be male or female.

(Hint: `install.packages("MASS")` for function `lda()`.)

(e) Use cross-validation on the given data to estimate the probabilities of correctly classifying male and female centipedes from the concentrations of the three amino acids in their hæmolymph.