



## Statistik III: Multivariate Stat. Verfahren

### Übungsblatt 3

**Bearbeitung:** Do. 27.5.2004, 16.00 Uhr.

Alle Aufgaben können im Mac-Cip Pool bearbeitet werden.

- Erstellen Sie einen Contour-Plot einer 2-dim. Normalverteilung mit den Parametern  $\mu = (1, 2.5)$  und  $\Sigma = \begin{pmatrix} 2 & -3 \\ -3 & 5 \end{pmatrix}$  für Wahrscheinlichkeiten  $p_{xy} = 0.5, 0.8, 0.95, 0.99$ .
  - Schätzen Sie für die beiden Variablen *palmitoleic* und *oleic* aus dem **Olive Oils** Datensatz (siehe Blatt 2, Aufgabe 5) den Mittelwertvektor  $\mu$  und die Varianz-Kovarianzmatrix  $\Sigma$  mittels  $\bar{X}$  und  $S^2$ , und vergleichen Sie die daraus resultierenden Isolinien mit denen des 2-dim. Kerndichteschätzers.
- Erstellen Sie eine Scatterplot Matrix für den **Transpirations** Datensatz aus der Vorlesung. Tragen Sie dort ein:
  - $\bar{X}$ ,  $\mu_0$ , und die Konfidenzellipse
  - die simultanen Konfidenz Intervalle als Projektion der Konfidenzellipse
  - die Bonferroni korrigierten Konfidenz Intervalle.
- Beschreiben Sie formal die Abbildung von Seite 38 unten im Vorlesungsskript, und erstellen Sie sie in R.
  - Zeigen Sie dass gilt:
$$\sigma_{11} + \dots + \sigma_{pp} = \lambda_1 + \dots + \lambda_p$$
wobei  $\lambda_i$  die Eigenwerte der Varianz- Kovarianz Matrix von  $X$  und  $\sigma_{ii}$  deren Varianzen sind.
- Der Datensatz **Studenten** enthält Informationen über Tests an 82 Studenten aus drei Fachbereichen: Technologie (=1), Architektur (=2) und Medizintechnik (=3). Gemessen wurden jeweils Leistung in (1) Eignung, (2) Mathe, (3) Sprachen und (4) Allgemeinbildung.
  - Stellen Sie die Ergebnisse möglichst kompakt in einem Trellis Plot dar, so dass die Leistungen einfach verglichen werden können.
  - Stellen Sie die Testergebnisse in einer Scatterplot Matrix dar, und unterscheiden Sie dabei die Gruppen mittels drei verschiedener Farben.
  - Testen Sie mit einer MANOVA, ob die Unterschiede zu einem Niveau von  $\alpha = 0.01$  signifikant sind.
- Berechnen Sie für die **Olive Oils** Daten die Hauptkomponenten mittels einer Eigenwertzerlegung der Korrelationsmatrix. Wie groß sind die einzelnen Varianzkomponenten? Stellen Sie die drei ersten Hauptkomponenten in Scatterplots dar.