h_da
HOCHSCHULE DARMSTADT
UNIVERSITY OF APPLIED SCIENCES
fbmn
FACHBEREICH MATHEMATIK
UND NATURWISSENSCHAFTEN

Antje Jahn
Multivariate Statistik
Sommersemester 2018

Arbeitsblatt 1

A 1 Erstellen Sie sich eine R-Markdown-Vorlage, die im Dateikopf Namen und Matrikelnummer Ihrer Arbeitsgruppe und „echo=TRUE" als Default-Einstellung für R Code Chunks setzt. Nutzen Sie diese Vorlage dann für alle R-Übungen dieser Vorlesung.

A 2

a) Laden Sie das R-Paket ISLR und schauen Sie sich den darin enthaltenen Datensatz *Default* an. Nutzen Sie *?Default*, um eine Beschreibung des Datensatzes zu erhalten.
b) Definieren Sie eine multivariate Fragestellung, die hier von Interesse sein koennte.
c) Wie werden hinsichtlich dieser Fragestellung die Matrizen X und Y definiert? Geben Sie p und n an. Welche Dimensionen haben X und Y?
d) Schauen Sie sich die Variablendeskriptionen an und beschreiben Sie paarweise Zusammenhaenge zwischen den Variablen mittels Streudiagrammen, Boxplots und Haeufigkeitstabellen: *summary, hist, plot, boxplot, table, prop.table...*
e) Untersuchen Sie paarweise Zusammenhaenge mittels univariater statistischer Tests. Formulieren Sie dazu jeweils H_0 und H_1 und waehlen Sie einen geeigneten statistischen Test (*chisq.test, t.test...*). Interpretieren Sie das Ergebnis. Zu welchen Schlussfolgerungen kommen Sie?

A 3

a) Laden Sie das R-Paket ISLR und schauen Sie sich den darin enthaltenen Datensatz *Hitters* an. Nutzen Sie *?Hitters*, um eine Beschreibung des Datensatzes zu erhalten.
b) Angenommen, Sie interessieren sich dafuer, welche Faktoren das Gehalt eines Spielers beeinflussen. Wie werden hinsichtlich dieser Fragestellung die Matrizen X und Y definiert? Geben Sie p und n an. Welche Dimensionen haben X und Y?
c) Schauen Sie sich die Variablendeskriptionen an, beschreiben Sie paarweise Zusammenhaenge zwischen *salary* und weiteren Variablen mittels Streudiagrammen und Boxplots. Nutzen Sie die Funktion *pairs*, um mehrere paarweise Streudiagramme gleichzeitig darzustellen (Schraenken Sie dazu den Datensatz auf die interessierenden Spalten ein!).
d) Geben Sie das mittlere Einkommen der Spieler mit 95%-Konfidenzintervall unter Normalverteilungsannahme an (Funktion *t.test*). Interpretieren Sie das Ergebnis.

A 4

a) Prüfen Sie nach, dass die Matrix

$$A = \begin{pmatrix} \frac{5}{13} & \frac{12}{13} \\ -\frac{12}{13} & \frac{5}{13} \end{pmatrix}$$

orthogonal ist.

b) Die Matrix

$$A = \begin{pmatrix} 9 & -2 \\ -2 & 6 \end{pmatrix}$$

sei gegeben. Zeigen Sie, dass A positiv definit ist.

c) Die Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$$

sei gegeben.

i) Finden Sie die Eigenwerte λ_1, λ_2 und die normierten Eigenvektoren von A.

ii) Finden Sie $V \in \mathbb{R}^{2\times 2}$ mit $A = V \cdot \operatorname{diag}(\lambda_1, \lambda_2) \cdot V^T$.