Multivariates Schätzen und Testen

Aufgabe 1: Kovarianzmatrix

In einer Stichprobe der Größe $n = 3$ erheben Sie zwei Merkmale X_1 und X_2 und erhalten folgende Datenmatrix:

$$\mathbf{X} = \begin{pmatrix} 1 & -4 \\ 0 & 2 \\ 2 & -1 \end{pmatrix}.$$

Schätzen Sie die Kovarianzmatrix von $\boldsymbol{X} = (X_1, X_2)$.

Aufgabe 2: Multivariate Tests I

In einer Studie (Seal (1964), S. 106) wurde die Länge (x_1) und Breite (x_2) des Schädels bei 35 erwachsenen weiblichen Fröschen untersucht. Es ergaben sich folgende empirische Größen:

$$\bar{\boldsymbol{x}}_w = \begin{bmatrix} 22.860 \\ 24.397 \end{bmatrix}, \quad \boldsymbol{\Sigma}_w = \begin{bmatrix} 17.683 & 20.290 \\ & 24.407 \end{bmatrix}.$$

a) Testen Sie, unter der Annahme $(x_1, x_2)^T \sim N_2(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ mit

$$\boldsymbol{\mu}_0 = \begin{bmatrix} 20 \\ 25 \end{bmatrix} \quad \boldsymbol{\Sigma}_0 = \begin{bmatrix} 20 & 0 \\ 0 & 25 \end{bmatrix}$$

die Hypothese $H_0: \ \boldsymbol{\mu}_w = \boldsymbol{\mu}_0$ auf einem Signifikanzniveau von $\alpha = 0.05$.

b) Testen Sie, ob $\boldsymbol{\mu}_w = \boldsymbol{\mu}_0$ bei unbekannter Kovarianzmatrix ($\alpha = 0.05$).

c) Äquivalente Messungen wurden bei 14 erwachsenen männlichen Fröschen durchgeführt, mit folgenden Ergebnissen:

$$\bar{\boldsymbol{x}}_m = \begin{bmatrix} 21.821 \\ 22.843 \end{bmatrix}, \quad \boldsymbol{\Sigma}_m = \begin{bmatrix} 18.479 & 19.095 \\ & 19.273 \end{bmatrix}.$$

Überprüfen Sie unter der Annahme identischer Kovarianzmatrizen in beiden Subpopulationen, ob $\boldsymbol{\mu}_w = \boldsymbol{\mu}_m$ gilt ($\alpha = 0.05$).

Aufgabe 3: Multivariate Tests II

Für die Fußball-Bundesliga liegen für 16 verschiedene Vereine die beiden Merkmale ‘Anzahl an Toren in der 1. Halbzeit’ und ‘Anzahl an Toren in der 2. Halbzeit’ vor, jeweils für die Saisons 2013/14 ($\boldsymbol{X}^{(1)}$) und 2014/15 ($\boldsymbol{X}^{(2)}$). Mittels eines geeigneten multivariaten Tests soll überprüft werden, ob sich die beiden Merkmale zwischen den betrachteten Saisons unterscheiden.

	$\boldsymbol{x}_i^{(1)}$		$\boldsymbol{x}_i^{(2)}$		$\boldsymbol{d}_i$	
	1. HZ	2. HZ	1. HZ	2. HZ	1. HZ	2. HZ
Bayern München	37	57	29	51		
VfL Wolfsburg	26	37	34	38		
Bor. M’Gladbach	27	32	24	29		
Bor. Dortmund	32	48	22	25	10	23
Eintr. Frankfurt	18	22	23	33	-5	-11
Bayer Leverkusen	32	28	18	44	14	-16
Werder Bremen	20	22	24	26	-4	-4
Schalke 04	30	33	19	23	11	10
Mainz 05	19	33	17	28	2	5
1899 Hoffenheim	34	38	25	24	9	14
Hamburger SV	20	31	11	14	9	17
Hertha BSC	19	21	16	20	3	1
Hannover 96	22	24	15	25	7	-1
SC Freiburg	12	31	13	23	-1	8
VfB Stuttgart	26	23	16	26	10	-3
FC Augsburg	26	21	13	30	13	-9

a) Vervollständigen Sie die fehlenden Werte der Differenzenvektoren $\boldsymbol{d}_i = \boldsymbol{x}_i^{(1)} - \boldsymbol{x}_i^{(2)}$ in obiger Tabelle! Zeigen Sie außerdem, dass gilt

$$\bar{\boldsymbol{d}} = \begin{pmatrix} 5.1 \\ 2.6 \end{pmatrix}$$

b) Die geschätzte Kovarianzmatrix des obigen Differenzenvektors ist gegeben durch

$$\boldsymbol{S} = \begin{pmatrix} 45.3 & 11.3 \\ 11.3 & 108.2 \end{pmatrix}$$

Testen Sie mit einem geeigneten Test zum Signifkanzniveau von 5% und unter der Annahme von normalverteilten Variablen für beide Saisons, ob sich die erwarteten Anzahlen an Toren für Halbzeit 1 und Halbzeit 2 über die Saisons hinweg signifkant verändert haben.

c) Lesen Sie die Rohdaten `buligoals.rda` von der Übungshompage in `R` ein und vollziehen Sie den gesamten Test nach.