Florian Wickelmaier
Forschungsmethoden und Mathematische Psychologie
Fachbereich Psychologie
Universität Tübingen

Vorlesung Statistik II
SS 2026

Übungsaufgaben

Aufgabe 31. Es soll getestet werden, ob bei einer Klausur (Noten 1 bis 5) die Männer weniger Einser schreiben als die Frauen.

a) Bestimmen Sie per Simulation die Stichprobenumfänge n_1 und n_2 für Frauen und Männer, mit denen ein geeigneter Test folgende minimal relevanten Unterschiede der bedingten Wahrscheinlichkeiten mit ca. 80% Power ($\alpha = 0.05$) findet:

	Note				
	1	2	3	4	5
weiblich	.20	.30	.30	.12	.08
männlich	.10	.30	.30	.20	.10

Gehen Sie dabei von einem Geschlechterverhältnis Frauen:Männer von 4:3 aus.

b) In der Klausur erhielt man dann die in der folgenden Kontingenztabelle dargestellten Ergebnisse:

	Note				
	1	2	3	4	5
weiblich	14	20	20	4	2
männlich	7	20	11	2	5

Schätzen Sie die (bedingten) Wahrscheinlichkeitsverteilungen der Noten in den durch das Geschlecht definierten Subpopulationen.

c) Testen Sie, ob zwischen den Merkmalen Geschlecht und Klausurergebnis ein Zusammenhang besteht bzw. ob die durch das Geschlecht definierten Subpopulationen homogen sind ($\alpha = 0.05$).

Führen Sie alle Berechnungen von Hand durch und überprüfen Sie diese dann durch Verwendung der entsprechenden Funktionen in R.

Aufgabe 32. Bei einer Umfrage sollen Personen ihre Meinung zu einem Verbot von Waffenlieferungen ins Ausland angeben. Dabei sind die Äußerungen "stimme zu", "lehne ab" und "weiß nicht" möglich. Weiterhin sollen die Personen angeben, ob sie bei der nächsten Bundestagswahl SPD, CDU/CSU, die Linke oder eine andere Partei wählen würden.

a) Bestimmen Sie per Simulation den Stichprobenumfang n, mit dem ein geeigneter Test die Unabhängigkeit zwischen der Meinung zum Verbot und dem Wahlverhalten mit ca. 80% Power ($\alpha = 0.01$) ablehnen kann. Gehen Sie dabei von folgenden Verbundwahrscheinlichkeiten aus, die eine Häufung von "stimme zu"-Antworten bei Wählern der Linken repräsentieren:

	Zustimmung	Ablehnung	weiß nicht
SPD	0.09	0.15	0.06
CDU/CSU	0.10	0.22	0.08
Linke	0.05	0.03	0.02
Andere	0.06	0.10	0.04

b) Anschließend seien bei 1000 Personen folgende Daten erhoben worden:

	Zustimmung	Ablehnung	weiß nicht
SPD	100	180	50
CDU/CSU	100	200	100
Linke	60	10	0
Andere	50	100	50

Prüfen Sie für eine Irrtumswahrscheinlichkeit von $\alpha = 0.01$, ob zwischen der Meinung zum Verbot und dem Wahlverhalten ein Zusammenhang besteht.

Führen Sie alle Berechnungen von Hand durch und überprüfen Sie diese dann durch Verwendung der entsprechenden Funktionen in R.

Aufgabe 33. Der Zusammenhang zwischen den Werten eines Intelligenztests (Variable $X \sim N(100, 15^2)$) und eines Gedächtnistests (Variable $Y \sim N(100, 15^2)$) soll geprüft werden. Gehen Sie davon aus, dass Paare aus beiden Variablen unabhängig gemeinsam normalverteilt sind mit unbekannter Korrelation ρ_{XY}.

a) Bestimmen Sie per Simulation den Stichprobenumfang n, um unter obigen Annahmen mit einem geeigneten statistischen Test bei ca. 80% Power ($\alpha = 0.05$) eine Korrelation von $\rho_{XY} = 0.6$ nachweisen zu können, H_0: $\rho_{XY} = 0$.

b) Anschließend wurden bei $n = 10$ Probanden Intelligenz und Gedächtnisleistung ermittelt. Man erhielt die Wertepaare:

X	124	79	118	102	86	89	109	128	114	95
Y	100	94	101	112	76	98	91	73	90	84

Prüfen Sie die Hypothese $\rho_{XY} = 0$ der Unkorreliertheit von X und Y unter Verwendung des oben gewählten statistischen Tests ($\alpha = 0.05$). Interpretieren Sie das Ergebnis mit Blick auf den oben ermittelten Stichprobenumfang.

Führen Sie alle Berechnungen von Hand durch und überprüfen Sie diese dann durch Verwendung der entsprechenden Funktionen in R.

c) Überprüfen Sie die Plausibilität der Normalverteilungssnnahme für jede Variable getrennt mit Hilfe eines Quantil-Normal-Plots.

Aufgabe 34. In einem Wahrnehmungsexperiment wurden $n = 10$ Versuchspersonen aufgefordert zu jeweils einem Reiz, der in unterschiedlichem physikalischen Abstand (Merkmal X) vom Betrachter platziert war, eine Schätzung von dessen Entfernung (Merkmal Y) abzugeben. Für die Versuchspersonen $i = 1, ..., 10$ ergaben sich folgende Wertepaare (in m):

x_i	20	20	30	40	40	60	70	90	90	90
y_i	16	26	25	35	55	62	95	70	110	118

Verwenden Sie bei allen Tests eine Irrtumswahrscheinlichkeit von 5%.

a) Zeichnen Sie die Wertepaare (x_i, y_i) in ein Streudiagramm ein.

b) Schätzen Sie die Parameter α und β des Regressionsmodells

$$Y_i = \alpha + \beta\, X_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2) \text{ i.i.d.}, \quad i = 1, ..., 10$$

und tragen Sie die resultierende Stichproben-Regressionsgerade in das Streudiagramm ein.

c) Bestimmen Sie für den verwendeten Stichprobenumfang, den verwendeten physikalischen Abstand und $\sigma = 16\,\text{m}$ per Simulation die Power eines Tests der Hypothese H_0: $\beta = 0$, falls stattdessen $\beta = 1$ gelten sollte.

d) Berechnen Sie den Determinationskoeffizienten und die F-Statistik des Overall-Tests.

   (a) Liefert dieser Test ein statistisch signifikantes Ergebnis?

   (b) Was bedeutet dieses Ergebnis inhaltlich?

e) Welche inhaltliche Bedeutung hat die Annahme, dass $\alpha = 0$ und $\beta = 1$ ist, also für alle $i = 1, ..., 10$ das Modell

$$Y_i = X_i + \varepsilon_i$$

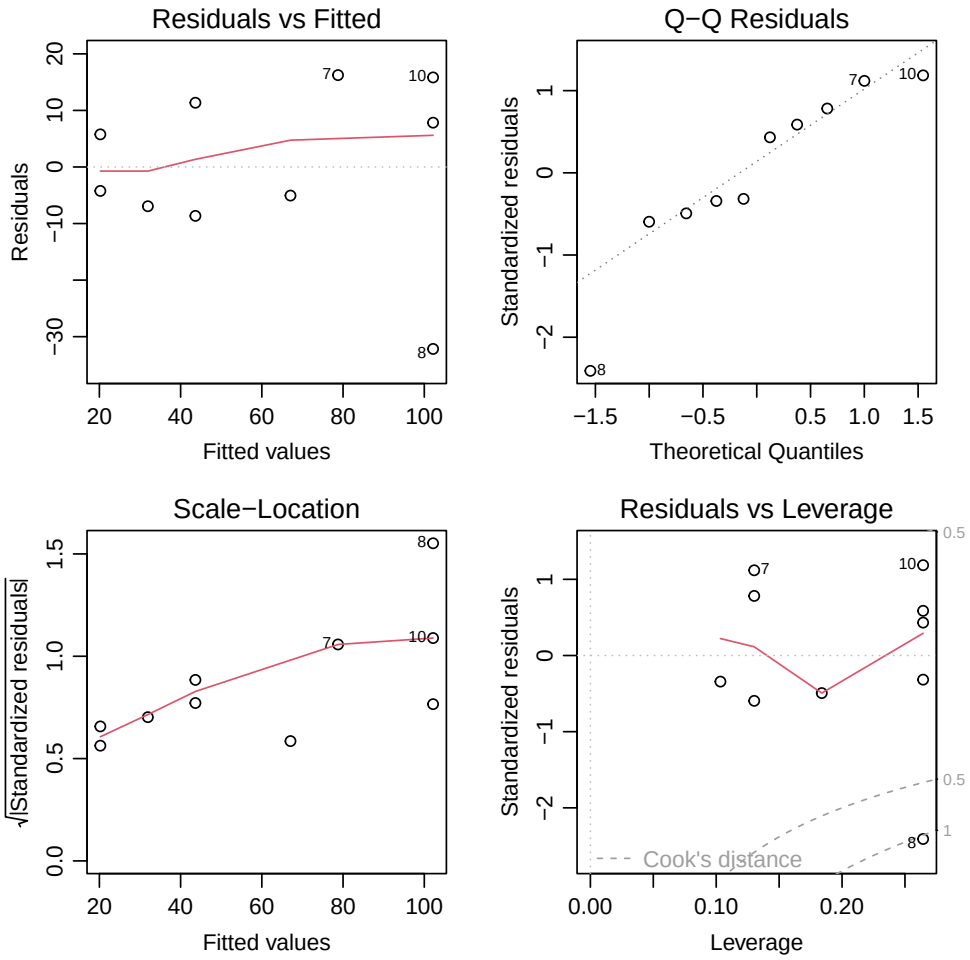
gilt mit $E(\varepsilon_i) = 0$?

f) Überprüfen Sie die beiden Hypothesen $\alpha = 0$ und $\beta = 1$ durch geeignete statistische Tests.

g) Welche Werte für die geschätzte Entfernung werden durch die Regressionsgerade vorhergesagt?

h) Geben Sie Prognoseintervalle für die vorhergesagten Werte bei $x_0 = 20$ und $x_0 = 90$ an.

i) Eine Analyse der Daten mit R liefert folgende Grafiken zur Modelldiagnose



Welche Aussagen lassen sich hieraus zu folgenden Fragen ableiten?

(a) Sind die Residuen approximativ normalverteilt?

(b) Kann die Annahme der Homoskedastizität als erfüllt angesehen werden?

(c) Gibt es Datenpunkte, die einen dominierenden Einfluss auf die Ergebnisse der Regressionsanalyse haben?

Führen Sie alle Berechnungen von Hand durch und überprüfen Sie diese durch Verwendung der entsprechenden Funktionen in R.