



UNIVERSITY
OF APPLIED SCIENCES

Quantitative Forschungsmethoden

Übungsaufgaben

QAF

Übungsaufgaben zum Modul Quantitative Forschungsmethoden (QAF)

Teil 1: Rechen- und Verständnisaufgaben

Aufgabe 1: Wahr oder Falsch?

1. Der Mittelwert ist ein erwartungstreuer Schätzer.
2. μ ist ein konsistenter und erwartungstreuer Schätzer für den Erwartungswert.
3. Der Fehler 1. Art und der Fehler 2. Art können bei einem statistischen Test gemeinsam auftreten.
4. Der alpha-Fehler (Fehler 1. Art) eines statistischen Tests kann größer als das unterstellte Signifikanzniveau sein.

Aufgabe 2: Statistisches Testen

Es ist ein weit verbreitetes Klischee, dass in Bayern mehr Alkohol konsumiert wird als in anderen Bundesländern. Im Rahmen eines Forschungsprojekts für Suchtprävention soll diese These wissenschaftlich untersucht werden. Sie veranstalten deshalb zwei identische Partys in München und Hamburg und messen am Ende des Abends den Atemalkoholgehalt (in mg/l) bei von je 10 Party-Besucher*innen.

Nehmen Sie im Weiteren an, dass das Merkmal Atemalkoholgehalt normalverteilt ist.

Atemalkoholgehalt in München (M)	0.5	0.7	1.2	0.0	1.5	0.6	0.3	0.2	0.1	0.9
Atemalkoholgehalt in Hamburg (H)	0.4	0.8	1.1	1.2	0.0	0.3	0.4	0.2	0.1	0.1

- Schätzen Sie ein 95%-Konfidenzintervall für den Erwartungswert des Atemalkoholgehalts bei den Münchner Party-Besucher:innen (t-Verteilung im Anhang).
- Testen Sie, ob sich die Erwartungswerte des Atemalkoholgehalts zwischen den Besucher:innen der beiden Partys voneinander unterscheiden ($\alpha = 0.05$). Hinweis: $S_M^2 = 0,24$ und $S_H^2 = 0,18$. Gehen Sie dabei nach folgendem Muster vor:
 - Welcher Test wird gewählt?
 - Welche Annahmen werden dafür getroffen? (Hinweis: Diese können Sie als gegeben annehmen)
 - Signifikanzniveau
 - Statistische Hypothesen
 - Testwert berechnen
 - Ablehnungsbereich definieren
 - Testentscheidung treffen
- Erklären Sie, was alpha- und beta-Fehler im Kontext der Aufgabenstellung inhaltlich bedeuten.

Aufgabe 3: Effektstärken

In einem Unternehmen wird mit dem Personal eine Woche lang ein tägliches Trainingsprogramm zur Steigerung der Fitness eingeführt. Ein Wert von Interesse ist hierbei, wie sich das Training auf die subjektive Einschätzung der Fitness auswirkt. Vor dem Programmstart beurteilten die Mitarbeiterinnen und Mitarbeiter ihre Fitness auf einer Skala von 1 bis 10 (höhere Werte stellen höhere subjektive Fitness dar). Dabei ergab sich ein Mittelwert von 5,4. Nach dem Programm ergab sich ein Mittelwert von 6,8. Die Schätzung für die Streuung der Differenzwerte ergibt $\sqrt{\hat{S}_D^2} = 1,8$.

- a) Berechnen Sie die Effektstärke der Trainingsmaßnahme.
 b) Beurteilen Sie die Größe des Effekts.

Teil 2: SPSS-Output-Interpretation

Aufgabe 1: Deskriptive Statistik

Ihnen liegt folgender SPSS-Output vor:

Häufigkeiten

[DataSet1] /Applications/IBM/SPSS/Statistics/25/Samples/German/survey_sample.sav

Statistiken				
	Familienstand	Anzahl Kinder	Höchster Abschluss	
N	Gültig	2831	2825	2822
	Fehlend	1	7	10

Häufigkeitstabelle

Familienstand				
		Häufigkeit	Prozent	Gültige Prozente
Gültig	Verheiratet	1346	47,5	47,5
	Verwitwet	283	10,0	10,0
	Geschieden	446	15,7	15,8
	Getrennt	93	3,3	3,3
	Nie verheiratet	663	23,4	23,4
	Gesamt	2831	100,0	100,0
Fehlend	KA	1	,0	
Gesamt		2832	100,0	

Anzahl Kinder				
		Häufigkeit	Prozent	Gültige Prozente
Gültig	0	802	28,3	28,4
	1	474	16,7	16,8
	2	743	26,2	26,3
	3	411	14,5	14,5
	4	209	7,4	7,4
	5	86	3,0	3,0
	6	47	1,7	1,7
	7	19	,7	,7
	Acht oder mehr	34	1,2	1,2
	Gesamt	2825	99,8	100,0
Fehlend	KA	7	,2	
Gesamt		2832	100,0	

Höchster Abschluss				
		Häufigkeit	Prozent	Gültige Prozente
Gültig	Niedriger als High School	430	15,2	15,2
	High School	1500	53,0	53,2
	Junior College	209	7,4	7,4
	Bachelor	478	16,9	16,9
	Universitätsabschluss	205	7,2	7,3
	Gesamt	2822	99,6	100,0
Fehlend	KA	10	,4	
Gesamt		2832	100,0	

- a) Geben Sie N an.
 b) Geben Sie an, welche Merkmale erhoben wurden. Benennen Sie auch dessen Merkmalausprägungen und das zugehörige Skalenniveau.
 c) Geben Sie an, wie viele Personen der Stichprobe nie verheiratet waren.
 d) Erklären Sie, was die blau markierte Zahl aussagt.
 e) Wie viele Personen (in %) haben 5 Kinder oder mehr?

Aufgabe 2: Inferenzstatistik

Sie sind bei einem großen Sportverein mit den statistischen Analysen der Trainingsergebnisse betraut. Es soll der Einfluss des Konditionstrainings (leicht, mittel, hart) und des Alters der Spieler (jung, alt) auf die Laufzeit über 1000 m in Minuten untersucht werden. Es ergibt sich nachfolgender SPSS-Output:

Zwischensubjektfaktoren			
		Wertelabel	N
Alter	1	jung	12
	2	alt	12
Konditionstraining	1	leicht	8
	2	mittel	8
	3	hart	8

Deskriptive Statistiken				
Abhängige Variable: Laufzeit in min				
Alter	Konditionstraining	Mittelwert	Standardabweichung	N
jung	leicht	2,400	,2000	4
	mittel	2,375	,2217	4
	hart	2,375	,2217	4
	Gesamt	2,383	,1946	12
alt	leicht	2,500	,2309	4
	mittel	2,750	,3317	4
	hart	2,850	,1732	4
	Gesamt	2,700	,2763	12
Gesamt	leicht	2,450	,2070	8
	mittel	2,563	,3292	8
	hart	2,613	,3137	8
	Gesamt	2,542	,2842	24

Tests der Zwischensubjekteffekte					
Abhängige Variable: Laufzeit in min					
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.
Korrigiertes Modell	,863 ^a	5	,173	3,124	,033
Konstanter Term	155,042	1	155,042	2804,774	,000
Alter	,602	1	,602	10,884	,004
Konditionstraining	,111	2	,055	1,003	,387
Alter * Konditionstraining	,151	2	,075	1,364	,281
Fehler	,995	18	,055		
Gesamt	156,900	24			
Korrigierte Gesamtvariation	1,858	23			

a. R-Quadrat = .465 (korrigiertes R-Quadrat = .316)

- Stellen Sie das angewandte statistische Verfahren im Kontext der Aufgabenstellung dar.
- Schätzen Sie alle Interaktionseffekte des Modells.
Hinweis: Nutzen Sie folgende Formel: $(\alpha\beta)_{jk} = \mu_{jk} + \mu - \alpha_j - \beta_k$
- Testen Sie, ob das Modell (Overall) eine signifikante Erklärungsgüte besitzt ($\alpha = 5\%$).
- Skizzieren Sie ein Profildigramm (Liniendiagramm), welches nur die signifikanten Effekte des Modells berücksichtigt.

Teil 1: Rechen- und Verständnisaufgaben**Aufgabe 1: Wahr oder Falsch?**

1. Wahr.
2. Falsch, weil der Mittelwert ein konsistenter und erwartungstreuer Schätzer für den Erwartungswert ist. μ bezeichnet den Populationsparameter für den Erwartungswert.
3. Falsch. Alpha- und Beta-Fehler können niemals gemeinsam auftreten.
4. Falsch. Der Alpha-Fehler eines statistischen Tests muss immer $\geq \alpha$ sein.

Aufgabe 2: Statistisches Testen

- a) Schätzen Sie ein 95%-Konfidenzintervall für den Erwartungswert des Atemalkoholgehalts bei den Münchner Party-Besucher:innen. => SB 1; Statistisches Testen, S.46

$$\left[M_X \pm t_{n-1; \frac{\alpha}{2}} \cdot \frac{\hat{S}_X}{\sqrt{n}} \right]$$

- Mittelwert berechnen
- Standardabweichung berechnen
- T-Wert ablesen (siehe Anhang)
- Formel anwenden

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{10} (0,5 + 0,7 + \dots + 0,9) = 0,6$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{10-1} ((0,5 - 0,6)^2 + \dots + (0,9 - 0,6)^2) = 0,24$$

$$t(0,975; 9) = 2,262$$

$$KI = 0,6 \pm 2,262 \frac{\sqrt{0,24}}{\sqrt{10}} \Rightarrow [0,6 - 0,35; 0,6 + 0,35] = \mathbf{[0,25; 0,95]}$$

- b) Testen Sie, ob sich die Erwartungswerte des Atemalkoholgehalts zwischen den Besucher:innen der beiden Partys voneinander unterscheiden ($\alpha = 0.05$). => SB 1; Statistisches Testen, S.27

1. Welcher Test wird gewählt?

t-Test bei unabhängigen Stichproben, Mittelwertvergleich (SB1, S.27), einseitig (gerichteter Test)

2. Annahmen

Normalverteilung, Varianzhomogenität, Skalenniveau mindestens intervallskaliert

3. Signifikanzniveau

Alpha = 5 %

4. Hypothesen

$$H_0 = \mu_M \leq \mu_H$$

$$H_1 = \mu_M > \mu_H$$

5. Testwert berechnen

$$\bar{X}_M = 0,61$$

$$\bar{X}_H = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{10} (0,4 + 0,8 + \dots + 0,1) = 0,5$$

$$t = \frac{M_H - M_M}{\sqrt{\frac{\hat{S}_H^2 - \hat{S}_M^2}{n}}} = \frac{0,60 - 0,46}{\sqrt{\frac{0,18 - 0,24}{10}}} = 0,68$$

6. Ablehnungsbereich

Verteilung der Teststatistik unter H_0 : $t \sim t_{N-2}$ ($N=20$)

Kritischer t-Wert bei $t_{18;0,05} = 1,73$

H_0 ablehnen, falls der errechnete t-Wert größer ist als der kritische t-Wert von 1,73.

7. Testentscheidung

$0,68 < 1,73 \Rightarrow$ Die H_0 wird beibehalten. Auf einem Signifikanzniveau von 5% kann nicht bestätigt werden, dass der mittlere Atemalkoholgehalt sich zwischen den Partybesucher:innen in München und Hamburg unterscheidet.

c) Erklären Sie, was alpha- und beta-Fehler im Kontext der Aufgabenstellung inhaltlich bedeuten.

Alpha-Fehler: Tatsächlich ist der Atemalkoholgehalt der Partybesucher:innen in München größer als der Atemalkoholgehalt der Partybesucher:innen in Hamburg. Trotzdem wird die H_0 abgelehnt.

Beta-Fehler: Tatsächlich ist der Atemalkoholgehalt der Partybesucher:innen in München kleiner oder gleich dem Atemalkoholgehalt der Partybesucher:innen in Hamburg. Trotzdem wird die H_0 nicht abgelehnt.

Aufgabe 3: Effektstärken

c) Berechnen Sie die Effektstärke der Trainingsmaßnahme.

d) Beurteilen Sie die Größe des Effekts.

$$\delta = \frac{6,8 - 5,4}{1,8} = 0,778$$

Nach allgemeiner Konvention handelt es sich hier um einen mittelstarken Effekt.

Teil 2: SPSS-Output-Interpretation**Aufgabe 1: Deskriptive Statistik**

Ihnen liegt folgender SPSS-Output vor:

Häufigkeiten

[DataSet1] /Applications/IBM/SPSS/Statistics/25/Samples/German/survey_sample.sav

Statistiken				
	Familienstand	Anzahl Kinder	Höchster Abschluss	
N	Gültig	2831	2825	2822
	Fehlend	1	7	10

Häufigkeitstabelle

Familienstand					
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Verheiratet	1346	47,5	47,5	47,5
	Verwitwet	283	10,0	10,0	57,5
	Geschieden	446	15,7	15,8	73,3
	Getrennt	93	3,3	3,3	76,6
	Nie verheiratet	663	23,4	23,4	100,0
	Gesamt	2831	100,0	100,0	
Fehlend	KA	1	,0		
Gesamt		2832	100,0		

Anzahl Kinder					
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	802	28,3	28,4	28,4
	1	474	16,7	16,8	45,2
	2	743	26,2	26,3	71,5
	3	411	14,5	14,5	86,0
	4	209	7,4	7,4	93,4
	5	86	3,0	3,0	96,5
	6	47	1,7	1,7	98,1
	7	19	,7	,7	98,8
	Acht oder mehr	34	1,2	1,2	100,0
	Gesamt	2825	99,8	100,0	
Fehlend	KA	7	,2		
Gesamt		2832	100,0		

Höchster Abschluss					
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Niedriger als High School	430	15,2	15,2	15,2
	High School	1500	53,0	53,2	68,4
	Junior College	209	7,4	7,4	75,8
	Bachelor	478	16,9	16,9	92,7
	Universitätsabschluss	205	7,2	7,3	100,0
	Gesamt	2822	99,6	100,0	
Fehlend	KA	10	,4		
Gesamt		2832	100,0		

- N = 2831
- Geben Sie an, welche Merkmale erhoben wurden. Benennen Sie auch dessen Merkmalausprägungen und das zugehörige Skalenniveau.
 - Familienstand mit 5 Ausprägungen (verheiratet, verwitwet, geschieden, getrennt, nie verheiratet) => nominal skaliert
 - Anzahl Kinder mit 9 Ausprägungen (0,1,2,3,4,5,6,7, 8 und mehr) => ordinal skaliert
 - Höchster Abschluss mit 5 Ausprägungen (niedriger als High School, High School, Junior College, Bachelor, Uni) => ordinal skaliert
- Geben Sie an, wie viele Personen der Stichprobe nie verheiratet waren.
663
- Erklären Sie, was die blau markierte Zahl aussagt.

75,8% der Personen haben höchstens einen Junior College Abschluss

- Wie viele Personen (in %) haben mindestens 5 Kinder?

$$100\% - 93,4\% = 6,6\%$$

Aufgabe 2: Inferenzstatistik

Sie sind bei einem großen Sportverein mit den statistischen Analysen der Trainingsergebnisse betraut. Es soll der Einfluss des Konditionstrainings (leicht, mittel, hart) und des Alters der Spieler (jung, alt) auf die Laufzeit über 1000 m in Minuten untersucht werden. Es ergibt sich nachfolgender SPSS-Output:

Zwischensubjektfaktoren			
		Wertelabel	N
Alter	1	jung	12
	2	alt	12
Konditionstraining	1	leicht	8
	2	mittel	8
	3	hart	8

Deskriptive Statistiken				
Abhängige Variable: Laufzeit in min				
Alter	Konditionstraining	Mittelwert	Standardabweichung	N
jung	leicht	2,400	,2000	4
	mittel	2,375	,2217	4
	hart	2,375	,2217	4
	Gesamt	2,383	,1946	12
alt	leicht	2,500	,2309	4
	mittel	2,750	,3317	4
	hart	2,850	,1732	4
	Gesamt	2,700	,2763	12
Gesamt	leicht	2,450	,2070	8
	mittel	2,563	,3292	8
	hart	2,613	,3137	8
	Gesamt	2,542	,2842	24

Tests der Zwischensubjekteffekte					
Abhängige Variable: Laufzeit in min					
Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Sig.
Korrigiertes Modell	,863 ^a	5	,173	3,124	,033
Konstanter Term	155,042	1	155,042	2804,774	,000
Alter	,602	1	,602	10,884	,004
Konditionstraining	,111	2	,055	1,003	,387
Alter * Konditionstraining	,151	2	,075	1,364	,281
Fehler	,995	18	,055		
Gesamt	156,900	24			
Korrigierte Gesamtvariation	1,858	23			

a. R-Quadrat = .465 (korrigiertes R-Quadrat = .316)

a) Stellen Sie das angewandte statistische Verfahren im Kontext der Aufgabenstellung dar.

2-Faktorielle Varianzanalyse mit

- Faktor A: Alter (2 Faktorstufen, alt/jung)
- Faktor B: Konditionstraining (3 Faktorstufen, leicht/mittel/hart)

AV: Laufzeit in Minuten

b) Schätzen Sie alle Interaktionseffekte des Modells.

s. SB 2, S. 42

$$(\alpha\beta)_{jk} = \mu_{jk} + \mu - \alpha_j - \beta_k$$

$$(\alpha\beta)_{jung;leicht} = 2,4 + 2,54 - 2,38 - 2,45 = 0,11$$

$$(\alpha\beta)_{jung;mittel} = 2,375 + 2,54 - 2,38 - 2,56 = -0,025$$

Ab hier kann zur Vervollständigung eine Tabelle verwendet werden, da sich alle Effekte zu 0 aufaddieren lassen:

Alter	Training				
	$(\alpha\beta)_{jk}$	1	2	3	
	1	0,11	- 0,025	- 0,085	0
	2	-0,11	0,025	0,085	0
		0	0	0	0

- c) Testen Sie, ob das Modell (Overall) eine signifikante Erklärungsgüte besitzt ($\alpha = 5\%$). Die -Test-Voraussetzungen können Sie als gegeben annehmen.

1. Welcher Test?

Overall-F-Test, Varianzanalyse

2. Annahmen

Gegeben

3. Signifikanzniveau: $\alpha = 5\%$

4. Hypothesen:

$$H_0 : \mu_{11} = \mu_{12} = \dots = \mu_{23}$$

$$H_1 : \mu_{jk} \neq \mu_{jk^*} \text{ für mindestens ein Paar } (jk \neq jk^*)$$

5. Testwert

$F = 3,124$ (gemäß SPSS Output)

6. Ablehnungsbereich: falls $p\text{-Value} < 0.05$

$p\text{-Value} = 0,033$ (gemäß SPSS Output)

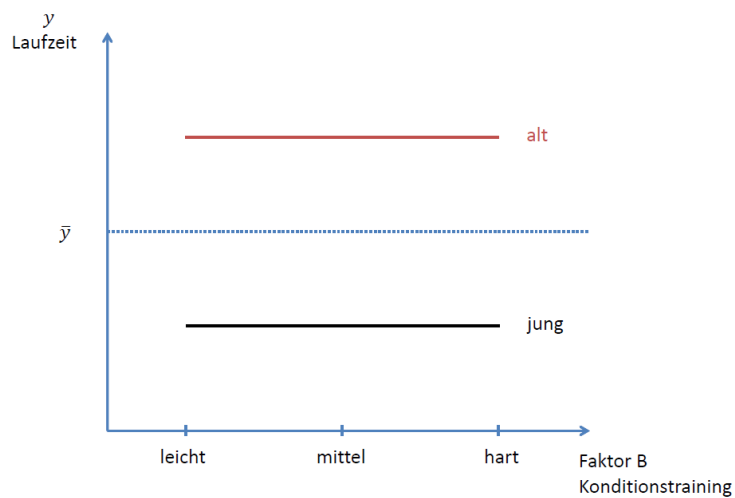
7. Testentscheidung

H_0 ablehnen $\Rightarrow 0,033 < 0,005$

Das Modell liefert einen signifikanten Erklärungsbeitrag für den Erwartungswert der Laufzeit in min ($\alpha = 5\%$).

- d) Skizzieren Sie ein Profildiagramm, welches nur die signifikanten Effekte des Modells berücksichtigt.

Profildiagramm: nur Haupteffekt A (Alter) signifikant



Die Quantile der t-Verteilung:

$n \backslash \alpha$	0,10	0,05	0,025	0,01	0,005	0,001	0,0005	einseitig
	0,20	0,10	0,05	0,02	0,01	0,002	0,001	zweiseitig
1	3,078	6,314	12,71	31,82	63,66	318,3	636,6	
2	1,886	2,920	4,303	6,965	9,925	22,33	31,56	
3	1,638	2,353	3,182	4,541	5,841	10,22	12,92	
4	1,533	2,132	2,776	3,747	4,604	7,173	8,610	
5	1,476	2,015	2,571	3,365	4,032	5,893	6,869	
6	1,440	1,943	2,447	3,143	3,707	5,208	5,959	
7	1,415	1,895	2,365	2,998	3,499	4,785	5,408	
8	1,397	1,860	2,306	2,896	3,355	4,501	5,041	
9	1,383	1,833	2,262	2,821	3,250	4,297	4,781	
10	1,372	1,812	2,228	2,764	3,169	4,144	4,587	
11	1,363	1,796	2,201	2,718	3,106	4,025	4,437	
12	1,356	1,782	2,179	2,681	3,055	3,930	4,318	
13	1,350	1,771	2,160	2,650	3,012	3,852	4,221	
14	1,345	1,761	2,145	2,624	2,977	3,787	4,140	
15	1,341	1,753	2,131	2,602	2,947	3,733	4,073	
16	1,337	1,746	2,120	2,583	2,921	3,686	4,015	
17	1,333	1,740	2,110	2,567	2,898	3,646	3,965	
18	1,330	1,734	2,101	2,552	2,878	3,610	3,922	
19	1,328	1,729	2,093	2,539	2,861	3,579	3,883	
20	1,325	1,725	2,086	2,528	2,845	3,552	3,850	
21	1,323	1,721	2,080	2,518	2,831	3,527	3,819	
22	1,321	1,717	2,074	2,508	2,819	3,505	3,792	
23	1,319	1,714	2,069	2,500	2,807	3,485	3,768	