

**Hinweise:** Mit \* gekennzeichnete Aufgaben sind mit einer Statistik-Software zu lösen. Die zu verwendenden Daten (SPSS- und Stata-Format), ein kurzes Codebuch, sowie Befehlsreferenzen zu SPSS und Stata findet man unter <http://www.soz.unibe.ch/statistik/>. Bei den Daten handelt es sich um eine Teilstichprobe des „Schweizer Arbeitsmarktsurveys 1998“ mit 1000 Fällen und 20 Variablen (vgl. zum Survey: <http://www.soz.unibe.ch/forschung/wl/daten.asp>).

Alle anderen Aufgaben können von Hand bzw. mit dem Taschenrechner gelöst werden.

Als Grundlage zu den Übungen dient das Lehrbuch: B. Jann (2002), Einführung in die Statistik, München: Oldenbourg.

## Übungen 1: Variablen und Skalenniveaus

- 1-1. Beschreiben Sie die Unterschiede zwischen
- stetigen und diskreten Variablen,
  - quantitativen und qualitativen Variablen.
- 1-2. Bestimmen Sie für die folgenden Variablen jeweils das Skalenniveau. Entscheiden Sie zudem, ob es sich um (quasi-)stetige oder diskrete Merkmale handelt.
- Staatsangehörigkeit
  - Schulnoten
  - Geschwisterzahl
  - Arbeitszufriedenheit (Skala von 1 bis 7)
  - Jahreszahlen
  - Lebensalter
  - Telefonnummer
  - Zivilstand
- 1-3. Für welche Skalenniveaus treffen die folgenden Aussagen jeweils zu?
- Die Addition einer Konstante ist eine zulässige Transformation.
  - Neben der Identität ist nur die „Grösser-Kleiner-Beziehung“ der Messwerte von Bedeutung.
  - Die „Grösser-Kleiner-Beziehung“ der Messwerte ist interpretierbar.
  - Die Verhältnisse von Messwerten sind interpretierbar.
  - Negative Messwerte sind möglich.
  - Die Multiplikation mit einer Konstante ist eine zulässige Transformation.

## Übungen 2: Summenzeichen

- 2-1. Schreiben Sie die folgenden Summen aus.
- $\sum_{i=1}^5 x_i$
  - $\sum_{i=1}^3 x_i^2 y_i$
  - $\sum_{i=1}^2 (x_i + y_i)^2$
  - $[\sum_{i=4}^5 y_i^2]^{1/2}$
  - $\sum_{i=4}^6 \sum_{j=1}^4 x_i y_j$
  - $\sum_{i=1}^2 (x_i^2 + y_i^2)$
  - $\sum_{i=1}^4 (x_i - 5)$

- 2-2. Berechnen Sie mit den nachstehenden Werten die Ergebnisse der in Aufgabe 2-1 aufgeführten Ausdrücke:

$$x_1 = 3; x_2 = 4; x_3 = 4; x_4 = 0; x_5 = 6; x_6 = 2$$

$$y_1 = 2; y_2 = 5; y_3 = 1; y_4 = 3; y_5 = 4; y_6 = 8$$

- 2-3. Gegeben ist die folgende Matrix mit den Werten  $x_{ij}$ :

$i \backslash j$	1	2	3
1	2	9	4
2	3	5	6
3	1	8	5
4	2	7	7

Berechnen Sie die folgenden Summen:

- $\sum_{i=1}^4 x_{i1}$  (Spaltensumme)
- $\sum_{j=1}^3 x_{3j}$  (Zeilensumme)
- $\sum_{i=3}^4 \sum_{j=1}^3 x_{ij}$
- $\sum_{i=1}^4 \sum_{j=1}^3 x_{i2} x_{3j}$

## Übungen 3: Häufigkeitsverteilungen und grafische Darstellung

- 3-1. Gegeben ist die folgende Häufigkeitstabelle:

Zivilstand	$h_j$
ledig	900
verheiratet	1700
geschieden	300
verwitwet	100
	3000

Berechnen Sie die relativen Häufigkeiten der Kategorien und stellen Sie sie in einem Säulendiagramm dar. Berechnen Sie zudem die Winkel der Sektoren, die zur Darstellung der Daten in einem Kreisdiagramm benötigt werden.

- 3-2. Vervollständigen Sie die Häufigkeitstabelle der Variable „Politisches Interesse“.

Politikinteresse	$h_j$	$f_j$
1. viel	810	...
2. mittel	...	0.45
3. wenig	...	0.28

- 3-3. Bei welchen Arten von Variablen ist die Erstellung eines Histogramms sinnvoll? Welche Bedeutung hat die Fläche einer Säule eines Histogramms?

3-4. Skizzieren Sie die Form (Dichtekurve) einer

- a) bimodalen,
- b) symmetrisch unimodalen,
- c) rechtsschiefen,
- d) linkssteilen Verteilung.

3-5. Gegeben sind die folgenden Häufigkeiten der Variable „Politische Orientierung“ ( $n = 3000$ ):

	links										rechts
Ausprägung	1	2	3	4	5	6	7	8	9	10	
Häufigkeit	75	75	450	600	750	450	300	150	60	90	

Zeichnen Sie die empirische Verteilungsfunktion. Zusatzfrage: Wieso ist dieses Vorgehen genau genommen problematisch?

- \*3-6. Erstellen Sie eine Häufigkeitsauszählung für die Variable „Geschlecht“ (SEX). Wie hoch ist der Anteil Frauen in der Stichprobe?
- \*3-7. Bilden Sie die Variable ALTER mit Hilfe des Geburtsjahres (A5J) und erstellen Sie eine Häufigkeitsauszählung für diese Variable (Befragungsjahr: 1998). Welcher Anteil der Befragten ist älter als 30?
- \*3-8. Erstellen Sie eine Häufigkeitsauszählung für den Ausbildungsabschluss der Frauen und eine Häufigkeitsauszählung für den Ausbildungsabschluss der Männer (Ausbildungsabschluss: G1, Geschlecht: SEX). Zu welchem Schluss gelangen Sie beim Vergleich der Anteile mit Tertiärbildung (höhere Berufsausbildung, Universität)?
- \*3-9. Erstellen Sie eine Häufigkeitsauszählung für die Frage, ob man an den Abstimmungen vom 28.9.1997 teilgenommen hat (I17). Wie hoch war die Abstimmungspartizipation?
- \*3-10. Erstellen Sie eine Häufigkeitsauszählung der Variable aus Aufgabe 3-9 für bis 30-jährige Personen sowie für über 30-jährige Personen. Was folgern Sie aus den Resultaten?
- \*3-11. Erstellen Sie ein Säulendiagramm für das politische Interesse der Befragten (B29). Was lässt sich über die Verteilungsform aussagen?
- \*3-12. Erstellen Sie für die Variable ALTER eine angemessene Grafik. Interpretieren Sie die Unterschiede im Vergleich zu der Verteilung aus Aufgabe 3-11.

## Übungen 4: Univariate Kennzahlen

4-1. Gegeben ist die folgende Datenreihe (Dauer der ersten 16 Interviews des Schweizer Arbeitsmarktsurveys 1998 in Minuten):

34, 44, 48, 37, 35, 58, 35, 37, 39, 40, 31, 35, 45, 31, 56, 60

Berechnen Sie den Modus, den Median, das arithmetische Mittel, das untere und obere Quartil, sowie die Varianz und Standardabweichung.

4-2. Berechnen Sie für die beiden untenstehenden Häufigkeitsverteilungen jeweils den Modus, den Median, das arithmetische Mittel, die Varianz und die Standardabweichung. Vergleichen Sie die Ergebnisse.

Originalwerte		5 Klassen	
$a_j$	$h_j$	$a_j$	$h_j$
80	1	0–199	3
100	1	200–399	11
170	1	400–599	1
200	4	600–799	4
250	1	800–1400	1
300	5		20
360	1		
400	1		
600	2		
700	2		
1400	1		
20			

- 4-3. Welche der folgenden Masszahlen ist bei Vorliegen von Ausreissern besonders wenig störanfällig?
  - a) arithmetischer Mittelwert
  - b) Range
  - c) Median
  - d) Varianz
- 4-4. Eine Variable  $X$  hat den Mittelwert  $\bar{x} = 1.7$ , die Varianz  $s^2 = 0.25$  und die Standardabweichung  $s = 0.5$ . Wie lauten die Koeffizienten einer gemäss

$$Y = 1.3 - 0.7X$$

transformierten Variable?

4-5. Es gilt:

$$Z = \frac{X - \bar{x}}{s_x}.$$

Wie lauten das arithmetische Mittel und die Standardabweichung der Variable  $Z$ ?

4-6. In einer Stichprobe von 10 Beobachtungen ergaben sich folgende Werte:

90, 92, 95, 96, 98, 102, 106, 107, 109, 118

- a) Berechnen Sie den Median und das arithmetische Mittel.
- b) Nachträglich stellte man fest, dass die Messergebnisse 95, 96 und 106 um jeweils 3 Einheiten zu hoch abgelesen wurden. Wie verändern sich der Median und das arithmetische Mittel nach einer entsprechenden Korrektur der Daten? (Versuchen Sie, die Frage ohne erneute Berechnung zu beantworten.)

- 4-7. Berechnen Sie für die folgende Häufigkeitsverteilung die Spannweite, den Interquartilsabstand, die durchschnittliche Abweichung  $AD$ , die Varianz  $s^2$  und den Variationskoeffizienten  $v$ .

$a_j$	$h_j$	$H_j$
7	7	7
8	14	21
9	20	41
10	12	53
11	5	58
12	2	60
		60

- 4-8. Sie wollen die Mietpreise von Studentenappartements in Bern und Wien vergleichen und haben die folgenden Werte erhoben:

Preise in Bern in CHF	523	414	456	474	341	564	384	404	587	465
Preise in Wien in ÖS	2850	3210	3872	4255	4028	3665	3230	3950	4150	3260

In welcher Datenreihe variieren die Mietpreise stärker?

- 4-9. In einem Wohnviertel mit 10 Haushalten besitzen 5 Haushalte CHF 0 Vermögen und die anderen 5 je CHF 1 Mio.

Zeichnen Sie die Lorenzkurve dieser Vermögensverteilung. Wie hoch ist der Gini-Koeffizient (durch Betrachtung der Grafik bestimmbar; führen Sie zur Kontrolle zusätzlich eine Berechnung des Gini-Koeffizienten durch)?

- \*4-10. Berechnen Sie Masszahlen der zentralen Tendenz und der Streuung für das monatliche Nettoeinkommen (NETINCO) von vollzeiterwerbstätigen Personen ( $C1=1$ ). Wie lauten die Werte für das arithmetische Mittel, den Median und die Standardabweichung? Was lässt sich bei Betrachtung der Masszahlen über die Verteilungsform aussagen?

- \*4-11. Vergleichen Sie die Mittelwerte und Standardabweichungen der monatlichen Nettoeinkommen von vollzeiterwerbstätigen Männern gegenüber vollzeiterwerbstätigen Frauen und erstellen Sie entsprechende Box-Plots. In welcher Gruppe variieren die Einkommen stärker?

## Übungen 5: Bivariate Kennzahlen I

- 5-1. Berechnen Sie mit Hilfe der nachfolgenden Randhäufigkeiten die bei statistischer Unabhängigkeit erwarteten Zellhäufigkeiten.

	$b_1$	$b_2$	
$a_1$	?	?	1610
$a_2$	?	?	1418
	2624	404	3028

- 5-2. Gegeben ist die folgende Kontingenztabelle der Merkmale  $X$  und  $Y$ :

		$Y$	
		0	1
$X$	0		40
	1	60	
		100	

Ergänzen Sie die fehlenden Zahlen und berechnen Sie die Zusammenhangsmasse  $\chi^2$  und  $\phi$ . Was lässt sich über den Zusammenhang zwischen  $X$  und  $Y$  aussagen?

- 5-3. Von den 200 Angestellten eines Betriebes sind 120 männlichen und 80 weiblichen Geschlechts. 50% der Männer und 40% der Frauen sind Raucher. Berechnen Sie einige geeignete Masse, die die Stärke der Abhängigkeit zwischen den Merkmalen „Geschlecht“ und „Rauchen“ ausdrücken.
- 5-4. Gegeben sind folgende Daten bezüglich Gemeindegrösse und Scheidungsrisiko?

	Land	Kleinstadt	Grossstadt
Ehe geschieden	5	10	60
Ehe nicht geschieden	50	60	80

- a) Erstellen Sie eine Tabelle, die die bedingten Häufigkeiten wiedergibt.
- b) Berechnen Sie Cramer's  $V$ , die Kontingenzkoeffizienten  $K$  und  $K^*$ , sowie die PRE-Masse Guttman's  $\lambda$  und Goodman und Kruskal's  $\tau$ . Interpretieren Sie die Resultate.
- c) Fassen Sie die Kategorien „Kleinstadt“ und „Grossstadt“ zusammen, so dass eine  $(2 \times 2)$ -Tabelle entsteht. Wiederholen Sie die Berechnungen und vergleichen Sie.

- 5-5. Berechnen Sie aufgrund der nachfolgenden Kontingenztabelle die Konkordanzmasse Kendall's  $\tau_b$  und Goodman und Kruskal's  $\gamma$ . Wann würde gelten  $\tau_b = \gamma$ ?

$X \backslash Y$	1	2	3	
1	5	2	0	7
2	2	6	7	15
3	1	3	8	12
	8	11	15	34

- 5-6. In einer sozialpsychologischen Studie wird einer Gruppe von 3 Kindern ein Film mit wenigen Gewaltszenen, einer Gruppe von 4 Kindern ein Film mit vielen Gewaltszenen vorgeführt. Anschliessend wird das Spielverhalten beobachtet (Tabelle). Berechnen zwei angemessene Zusammenhangsmasse.

Nr.	Film	Spielverhalten	
1	0	1	Film mit wenig Gewalt = 0
2	1	2	Film mit viel Gewalt = 1
3	1	1	
4	0	2	schwach aggressiv = 1
5	1	2	mittel aggressiv = 2
6	0	1	stark aggressiv = 3
7	1	3	

- \*5-7. Erstellen Sie eine Kontingenztafel mit den Variablen „Mitgliedschaft in einer Gewerkschaft oder einem Angestelltenverband“ (B24) und „Geschlecht“ für erwerbstätige Personen. Wie lautet die Prozentsatzdifferenz? Wie verändert sich die Prozentsatzdifferenz, wenn nur vollzeiterwerbstätige Personen berücksichtigt werden?
- \*5-8. Fassen Sie die Ausprägungen der Variable ALTER (aus Aufgabe 3-7) in 3 Kategorien zusammen: „18–30“, „31–50“ und „über 50“. Dichotomisieren zudem Sie die Variable „Erwerbsstatus“ (1 „erwerbstätig“, 0 „nicht erwerbstätig“). Erstellen Sie mit der dichotomisierten Variable „Erwerbsstatus“ und den Alterskategorien und eine Kreuztafel. Berechnen Sie neben den absoluten Häufigkeiten die Spaltenprozente sowie eine Reihe von Zusammenhangsmassen.

## Übungen 6: Bivariate Kennzahlen II

- 6-1. Berechnen und interpretieren Sie anhand der folgenden Daten die Stärke des linearen Zusammenhangs zwischen Einkommen und Kirchganghäufigkeit. Zeichnen Sie zudem ein Streudiagramm.

Person	A	B	C	D	E	F	G	H	I	J
monatl. Einkommen in Tausend	5	4	2	1	3	2	2	3	3	5
durchschnittl. monatl. Kirchenbesuch	2	1	1	4	0	3	5	2	0	0

- 6-2. Zwei Lehrer I und II bewerten die Leistungen von 14 Schülern (A bis N) rangplatzartig in folgender Weise:

Rang	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Lehrer I	K	D	A	C	E	F	J	G	L	B	N	I	M	H
Lehrer II	A	J	K	E	D	F	L	C	B	H	M	N	G	I

Berechnen Sie die Rangkorrelation. Wie beurteilen Sie die Übereinstimmung der Bewertungen?

- 6-3. Gegeben sind die nachfolgenden Einkommensdaten (monatliches Einkommen in 1000 CHF) für Arbeiter, Angestellte und Selbständige. Welcher Anteil der gesamten Einkommensvariation von 2476.1 wird durch diese Gruppierung erklärt?

	$\bar{x}$	$n$
Arbeiter	3.15	51
Angestellte	4.45	311
Selbständige	4.57	42
Total	4.30	404

- 6-4. Gegeben sind die folgenden Daten zu Geschlecht (0 = „männlich“, 1 = „weiblich“) und politischer Orientierung (1 = „rechts“ bis 10 = „links“):

Person	1	2	3	4	5	6	7	8	9	10
Geschlecht	0	1	1	0	1	1	1	0	0	1
Politische Orientierung	3	2	5	5	4	2	8	1	7	10

Berechnen Sie den punkt-biserialen und den rang-biserialen Korrelationskoeffizienten. Was folgern Sie aus den Resultaten?

- \*6-5. Berechnen Sie einige geeignete Masse für den Zusammenhang zwischen dem Interesse an Politik (B29) und der politischen Orientierung (B30). Was lässt sich über die Stärke des Zusammenhangs aussagen? Könnte ein nicht-monotoner Zusammenhang vorliegen?
- \*6-6. Bilden Sie aufgrund der Variable „Ausbildungsabschluss“ eine Variable „Bildungsjahre“ (Obligatorische Schule: 9 Jahre; Anlehre: 10 J.; Diplommittelschule: 11 J.; Berufslehre/Vollzeitberufsschule: 12 J.; Matura/Berufsmatura: 13 J.; höhere Berufsausbildung: 15 J.; Universität: 18 J.). Wie stark sind für vollzeiterwerbstätige Personen die Korrelationen zwischen den Variablen „monatliches Nettoeinkommen“, „Alter“ und „Bildungsjahre“.

## Übungen 7: Inferenzstatistik

- 7-1. Ein Fabrikant hat nach längerer Beobachtung festgestellt, dass bei der Produktion eines elektronischen Artikels zwei Fehler ( $A$  und  $B$ ) unabhängig voneinander mit den Wahrscheinlichkeiten  $P(A) = 0.1$  und  $P(B) = 0.2$  auftreten. Hat ein Stück nur einen Fehler, so ist es 2. Wahl. Stücke mit beiden Fehlern sind Ausschuss. Wieviel Prozent 1. Wahl, 2. Wahl und Ausschuss sind auf lange Sicht zu erwarten?
- 7-2. Bei einer Fluggesellschaft haben im Schnitt 10 von 100 000 Flügen eine Bombe an Bord (Ereignis  $A$ ). Bei durchschnittlich 400 Flügen geht ein anonymer Hinweis ein (Ereignis  $B$ ). Wenn eine Bombe an Bord war, hat die Gesellschaft in 80% der Fälle einen Hinweis erhalten. Wie gross ist die Wahrscheinlichkeit, dass sich nach einem anonymen Hinweis auch tatsächlich eine Bombe an Bord befindet?
- 7-3. Ermitteln Sie aus geeigneten Tabellen die 0.05- und 0.95-Quantile der  $t(10)$ - und  $t(17)$ -Verteilung, der  $\chi^2(15)$  und  $\chi^2(28)$ -Verteilung, sowie der  $F(3, 7)$ - und  $F(7, 3)$ -Verteilung.
- 7-4. Von den erwachsenen Einwohnern einer Stadt weiss man, dass die Körpergrösse normalverteilt ist mit den Parametern  $\mu = 175$  cm und  $\sigma = 10$  cm. Wie hoch ist die Wahrscheinlichkeit, dass ein zufällig ausgewählter Einwohner (a) höchstens 180 cm, (b) mindestens 160 cm bzw. (c) zwischen 175 cm und 190 cm gross ist?
- 7-5. Von 6 400 zufällig ausgewählten Personen aus der Bevölkerung gaben 2 304 an, Raucher zu sein. Bestimmen Sie das 95%-Konfidenzintervall für den Raucheranteil.

- 7-6. In einer statistischen Qualitätskontrolle bei der Herstellung von Massenprodukten entnimmt man im Allgemeinen in kurzen Abständen kleine Stichproben, um den Produktionsprozess laufend unter Kontrolle zu halten. Bei welcher der drei folgenden Stichproben muss korrigierend in die Produktion eingegriffen werden, wenn für ein Werkstück im Mittel ein Durchmesser von  $\mu_0 = 10$  cm eingehalten werden soll (bei einem Signifikanzniveau  $\alpha = 0.01$ ; die Standardabweichung der Produktion beträgt  $\sigma = 0.1$  cm)?

Stichprobe 1:	9.91	10.00	10.05	10.10	10.01
Stichprobe 2:	10.18	9.97	10.00	10.05	10.06
Stichprobe 3:	9.80	9.80	9.92	10.03	9.80

- 7-7. Testen Sie anhand der folgenden Angaben, ob Schweizer im Durchschnitt ein höheres Bildungsniveau (gemessen in Bildungsjahren) aufweisen als Ausländer ( $\alpha = 0.05$ ).

Schweizer:	$\bar{x}_1 = 13.34$	$s_1 = 3.11$	$n_1 = 2502$
Ausländer:	$\bar{x}_2 = 12.73$	$s_2 = 3.84$	$n_2 = 361$

- 7-8. In einem Bewerbungsexperiment wurden auf zufällig ausgewählte Stelleninserate jeweils zwei fiktive, bis auf das Geschlecht des Bewerbers inhaltlich äquivalente Bewerbungen eingereicht. In 17 von insgesamt 55 Versuchen wurden beide Bewerber zu einem Bewerbungsgespräch eingeladen. In 6 Versuchen wurde nur der Mann eingeladen und in 14 Versuchen nur die Frau. Hat das Geschlecht einen signifikanten Einfluss auf die Bewerbungschancen (bei  $\alpha = 0.05$ )?

- 7-9. Gegeben ist die folgende Kontingenztafel:

Kirchengangshäufigkeit	Konfession			Total
	katholisch	reformiert	andere	
mehrmals im Monat	8	1	6	15
mehrmals im Jahr	22	13	2	37
seltener	21	30	21	72
Total	51	44	29	124

Kann die Annahme, dass die beiden Merkmale unabhängig sind, verworfen werden?

- \*7-10. Besteht eine signifikante Mittelwertsdifferenz zwischen den monatlichen Einkommen von vollzeiterwerbstätigen Männern und vollzeiterwerbstätigen Frauen?
- \*7-11. Existiert ein signifikanter Zusammenhang zwischen dem Interesse an Politik (B29) und der abgeschlossenen Ausbildung (G1)?
- \*7-12. Testen Sie, ob sich die Anteile erwerbstätige Frauen und Männer, die Mitglied in einer Gewerkschaft oder einem Angestelltenverband sind, signifikant unterscheiden. Zu welchem Schluss gelangen Sie, wenn Sie nur *vollzeiterwerbstätige* Personen berücksichtigen?
- \*7-13. Testen Sie den Zusammenhang zwischen dem Interesse an Politik (B29) und dem Geschlecht? Ändert sich die Interpretation, wenn nur Personen mit Tertiärausbildung berücksichtigt werden?

## Übungen 8: Lineare Regression

- 8-1. Gegeben sind die folgenden Daten:

Person	A	B	C	D	E	F	G	H	I	J
monatl. Einkommen in Tausend	5	4	2	1	3	2	2	3	3	5
durchschnittl. monatl. Kirchenbesuch	2	1	1	4	0	3	5	2	0	0

Berechnen Sie

- die Parameter  $\alpha$  und  $\beta$  der Regressionsgeraden (und tragen Sie die Gerade in ein Streudiagramm ein),
  - die geschätzte Kirchgangshäufigkeit einer Person mit einem Monatseinkommen von CHF 3 000.– und
  - den Anteil nicht-erklärter Varianz, wobei das Einkommen die unabhängige Variable sei.
  - Berechnen Sie zudem die lineare Korrelation zwischen den Variablen und testen Sie diese auf Signifikanz ( $\alpha = 0.05$ ).
- 8-2. Folgende These soll überprüft werden: „Je grösser das Interesse einer Person an Mathematik und Naturwissenschaften, desto positiver ist ihre Einstellung zur Statistikvorlesung.“ In einer Stichprobe von 89 Studenten wurde für das Interesse an Mathematik ein Mittelwert von  $\bar{x} = 3.730$  und für die Einstellung zum Statistikkurs ein Mittelwert von  $\bar{y} = 4.169$  ermittelt. Für die Varianzen und die Kovarianz ergaben sich folgende Werte:  $s_X^2 = 3.449$ ;  $s_Y^2 = 1.824$ ;  $s_{XY} = 0.353$ .
- Wie lautet die Regressionsgerade zur Vorhersage der Einstellung zur Statistik?
  - Wie lautet die standardisierte Regressionsgleichung und wie ist sie zu interpretieren?
  - Wie hoch ist der Anteil erklärter Varianz?
  - Liefert das Modell einen signifikanten Erklärungsbeitrag (für  $\alpha = 0.05$ )?
- 8-3. Gegeben sind die folgenden Regressionskoeffizienten und Standardfehler:

	$\hat{\beta}$	$\hat{\sigma}$
Konstante	1.631	.440
Umweltbetroffenheit (0–10)	$-2.51 \cdot 10^{-2}$	.016
Umweltwissen (0–10)	.115	.023
Umweltbewusstsein (0–10)	.191	.028
Geschlecht (1 'weiblich')	.586	.096
Alter in Jahren	$1.03 \cdot 10^{-2}$	.003

Abhängige Variable: Umweltverhalten (0–10),  $n = 1287$

- Berechnen Sie die  $T$ -Werte der Koeffizienten. Welche Koeffizienten sind auf einem Niveau von  $\alpha = 0.05$  signifikant von null verschieden?
- Schätzen Sie den Wert des Umweltverhaltens für einen 35-jährigen Mann mit Maximalwerten auf den drei erklärenden Umweltvariablen.

- \*8-4. Schätzen Sie für vollzeitbeschäftigte Personen die folgenden Regressionsgleichungen (Alter aus Aufgabe 3-7; Bildungsjahre aus Aufgabe 6-6):

- a)  $\text{Einkommen} = \alpha + \beta \cdot \text{Alter}$
- b)  $\text{Einkommen} = \alpha + \beta \cdot \text{Bildungsjahre}$
- c)  $\text{Einkommen} = \alpha + \beta \cdot \text{Geschlecht}$
- d)  $\text{Einkommen} = \beta_0 + \beta_1 \cdot \text{Alter} + \beta_2 \cdot \text{Bildungsjahre} + \beta_3 \cdot \text{Geschlecht}$

Interpretieren Sie die Koeffizienten in den Modellen a) bis c). Was lässt sich über die Veränderung der Koeffizienten aussagen, wenn die Einflüsse der drei Variablen gleichzeitig geschätzt werden (letztes Modell). Welcher Anteil der Einkommensvarianz wird im letzten Modell erklärt?

- \*8-5. Erstellen Sie Streudiagramme mit eingezeichneter Regressionsgerade für die Modelle aus 8-4a und 8-4b.
- \*8-6. Schätzen Sie getrennt für Frauen und Männer die folgende, aus der Humankapitaltheorie abgeleitete Einkommensregression:

$$\ln(Y) = \beta_0 + \beta_1 \cdot EDUC + \beta_2 \cdot EXP + \beta_3 \cdot EXP^2 + \beta_4 \cdot \ln(H)$$

( $Y$  entspricht dem monatlichen Nettoeinkommen;  $EDUC$  den Bildungsjahren;  $EXP$  der Berufserfahrung, die berechnet werden soll als  $\text{Alter} - EDUC - 6.5$ ;  $H$  entspricht der wöchentlichen Arbeitszeit). Wie ist der Effekt eines zusätzlichen Bildungsjahres zu interpretieren? Wie ist der Effekt eines zusätzlichen Jahres Berufserfahrung zu interpretieren?

- \*8-7. Bilden Sie einen einfachen additiven Index des Umweltbewusstseins aus den Variablen S17A bis S17H (Standardisieren Sie den Index auf 0 „tiefes UB“ bis 10 „hohes UB“; Hinweis: einige Items sollten umgepolt werden). Welchen Einfluss hat die politische Links-Rechts-Orientierung (B30) auf das Umweltbewusstsein?