Aufgabe 1: **[14,5 Punkte]**

Sie interessieren sich für die Determinanten der subjektiven Lebenszufriedenheit. Der Datensatz beinhaltet folgende Variablen:

satlife	Subjektive Lebenszufriedenheit, gemessen auf einer Skala von 1=sehr unzufrieden bis 10=sehr zufrieden
region	=1, wenn Person wohnhaft in Westdeutschland; =0, wenn wohnhaft in Ostdeutschland
educ	Anzahl der Schuljahre
lnincpc	logarithmiertes Haushaltsnettoeinkommen pro Kopf
age	Alter einer Person in Jahren
sqage	Alter einer Person in Jahren zum Quadrat
hhmem05	Anzahl der Haushaltsmitglieder im Alter 0-5 Jahre
hhmem614	Anzahl der Haushaltsmitglieder im Alter 6-14 Jahre
hhmem1518	Anzahl der Haushaltsmitglieder im Alter 15-18 Jahre

Sie schätzen folgendes Modell:

$$satlife_i = \beta_0 + \beta_1 \text{region}_i + \beta_2 \text{educ}_i + \beta_3 \text{lnincpc}_i + \beta_4 \text{age}_i + \beta_5 \text{sqage}_i + \beta_6 \text{hhmem05}_i + \beta_7 \text{hhmem614}_i + \beta_8 \text{hhmem1518}_i + u_i$$

	Koeffizient	Std. fehler	T	Sig.
region	0,3792	0,0768	4,93	0,0000
educ	0,0383	0,0128	3,00	0,0030
lnincpc	0,6373	0,0734	8,68	0,0000
age	-0,0630	0,0225	-2,79	0,0050
sqage	0,0007	0,0003	2,33	0,0200
hhmem05	0,3494	0,1094	3,19	0,0010
hhmem614	0,1239	0,0761	1,63	0,1040
hhmem1518	0,1898	0,0848	2,24	0,0250
Konstante	1,2769	0,7792	1,64	0,1010
SSE	425,91			
SSR	5297,35			
SST	5723,26			
N	2295			

Abhängige Variable: satlife

a) Interpretieren Sie den Effekt des Einkommens (β_3) statistisch und inhaltlich. (2 Punkte)

statistisch signifikant auf 1% Niveau, da p-Wert <0,01

eine Erhöhung des Einkommens um 1% ist im Mittel c.p. korreliert mit einem Zufriedenheitsgewinn von 0,006 Punkten.

b) Eine Politikerin behauptet, dass die subjektive Lebenszufriedenheit um mehr als 0,3 Punkte steigt, wenn ein Kind im Alter bis 5 Jahre im Haushalt lebt. Testen Sie diese Hypothese auf dem 5%-Niveau. Geben Sie Testverfahren, Null- und Alternativhypothese, Teststatistik, Freiheitsgrade, kritischen Wert und Ihre Testentscheidung an. (4,5 Punkte)

einseitiger t-Test

H_0: $\beta_6 \leq 0,3$

H_1: $\beta_6 > 0,3$

$$t_{\hat{\beta}_6} = \frac{\hat{\beta}_6 - \beta_6}{se(\hat{\beta}_6)} = \frac{0,3494 - 0,3}{0,1094} = 0,45$$

$FG: n - k - 1 = 2295 - 8 - 1 = 2286$

$$t_c = t_{\alpha,n-k-1} = t_{0,05,2286} = 1,645$$

Da $t < t_c$ kann die H_0 nicht verworfen werden. Die Behauptung der Politikerin kann nicht bestätigt werden. Es kann somit nicht ausgeschlossen werden, dass der Zufriedenheitsgewinn $\leq 0,3$ ist.

c) Berechnen Sie das angepasste Bestimmtheitsmaß $\bar{R}^2$ für das Modell. (2 Punkte)

$$Adj.\ R^2 = 1 - \frac{SSR/(n-k-1)}{SST/(n-1)} = 1 - \frac{5297,35/2286}{5723,26/2294} = 0,071$$

d) Sie möchten wissen, ob der Zufriedenheitsgewinn aus Einkommen in Ost- und Westdeutschland signifikant verschieden ist. Wählen Sie eine geeignete Modellspezifikation und notieren Sie die Regressionsgleichung, die ihnen das gewünschte Ergebnis liefert. (2 Punkte)

$$satlife_i = \beta_0 + \beta_1 \text{region}_i + \beta_2 \text{educ}_i + \beta_3 \text{lnincpc}_i + \beta_4 \text{age}_i + \beta_5 \text{sqage}_i + \beta_6 \text{hhmem05}_i + \beta_7 \text{hhmem614}_i + \beta_8 \text{hhmem1518}_i + \beta_9 region_i * lnincpc_i + u_i$$

e) Geben Sie das 95% Konfidenzintervall für den geschätzten Parameter $\hat{\beta}_7$ (hhmem614) an und interpretieren sie es. (4 Punkte)

$$95\%\ Konfidenzintervall_{\hat{\beta}_7} = \left[\hat{\beta}_7 \pm c \cdot se(\hat{\beta}_7)\right]$$

$$c = t_{\frac{0,05}{2};\infty} = 1,96$$

$$[0,1239 \pm 1,96 \cdot 0,0761] = [-0,0253; 0,2731]$$

Bei wiederholter Stichprobenziehung liegt in 95% der Fälle der wahre Parameter innerhalb der auf diese Weise bestimmten Konfidenzintervalle.

Aufgabe 2: **[15,5 Punkte]**

Sie interessieren sich für die Determinanten von Mangergehältern. Ihr Datensatz enthält für n=177 Beobachtungen folgende Variablen:

salary	Gehalt gemessen in Tausend $
sales	Firmenumsatz gemessen in Million $
age	Alter der Person in Jahren
male	=1, wenn die Person männlich ist, sonst =0

a) Betrachten Sie das Modell $\log(salary) = \beta_0 + \beta_1 \log(sales) + u$.
Erläutern Sie *kurz* die Auswirkungen auf den KQ-Schätzer $\hat{\beta}_1$ wenn in diesem Modell *salary* in Hundert $ statt in Tausend $ gemessen wird. (2 Punkte)

*Der KQ-*Schätzer *$\hat{\beta}_1$ ändert sich nicht, da $\hat{\beta}_1$ eine Elastizität angibt. Diese ist unabhängig von den Einheiten in denen salary gemessen wird.*

b) Wie verändert sich $\hat{\beta}_1$ wenn *sales* in Tausend \$ statt in 1 Mio. \$ gemessen wird? (1 Punkt)

Auch hier ändert sich der Koeffizient $\hat{\beta}_1$ nicht, Erklärung wie oben.

c) Ein Kollege schlägt Ihnen vor, zu prüfen, ob das Modell aus a) getrennt für Männer und Frauen geschätzt werden sollte. Welchen Test (1% Niveau) können Sie hierzu verwenden? Nennen Sie Testverfahren, Null- und Alternativhypothese, Freiheitsgrade, kritischen Wert und die Testentscheidung. (3,5 Punkte)

Chow-Test:

$H_0: \beta_{0,male=0} = \beta_{0,male=1}$ *und* $\beta_{1,male=0} = \beta_{1,male=1}$ $H_1: H_0$ *gilt nicht*

Freiheitsgrade: Zähler=2 (Anzahl der Restriktionen), Nenner=173 (n–2(1+1))

$c = F_{1\%;2;173} = F_{1\%;2;\infty} = 4{,}61$

Wenn der Wert der Teststatistik größer ist als 4,61 *dann folgt, dass die H_0 verworfen wird und somit die Koeffizienten für Frauen und Männer signifikant verschieden sind.*

d) Sie schätzen in SPSS folgendes Modell:

$$\log(salary) = \beta_0 + \beta_1 \log(sales) + \beta_2 age + \beta_3 male + u$$

Dazu erhalten Sie den unten stehenden Output:

Koeffizienten[a]

Modell		Nicht standardisierte Koeffizienten: Regressionskoeffizient B	Nicht standardisierte Koeffizienten: Standardfehler	Standardisierte Koeffizienten: Beta	T	Sig.
1	(Konstante)	?	,299	?	16,248	,000
	log(sales)	,204	,028	,481	7,288	,000
	age in years	-,001	?	-,017	-,263	,793
	male	,361	,117	,201	3,095	,002

a. Abhängige Variable: log(salary)

i) Berechnen Sie die beiden fehlenden Werte („?“). (1 Punkt)

$$\hat{\beta}_0 = t_{\hat{\beta}_0} \cdot se(\hat{\beta}_0) = 16{,}248 \cdot 0{,}299 = 4{,}858$$

$$se(\hat{\beta}_2) = \frac{\hat{\beta}_2}{t_{\hat{\beta}_2}} = \frac{-0{,}001}{-0{,}263} = 0{,}0038$$

ii) Interpretieren Sie $\hat{\beta}_1$ und $\hat{\beta}_3$ präzise inhaltlich und statistisch. (4 Punkte)

$\hat{\beta}_1$: *Steigt der Umsatz um 1%, so steigt c.p. das erwartete Vorstandsgehalt um 0,204%. Der Koeffizient ist am 1%-Niveau statistisch signifikant von Null verschieden.*

$\hat{\beta}_3$: *Der erwartete Lohn von Männern liegt c.p. um* $[\exp(0{,}361) - 1] \cdot 100\% = 43{,}5\%$ *über dem von Frauen. Der Koeffizient ist am 1%-Niveau statistisch signifikant von Null verschieden.*

e) Sie wollen untersuchen, ob sich der Effekt des Alters für Männer und Frauen unterscheidet und schätzen daher das Modell:

$$\log(salary) = \beta_0 + \beta_1 \log(sales) + \beta_2 age + \beta_3 male + \beta_4 male * age + u$$

Der dazugehörende Output ergibt:

Model		Nicht standardisierte Koeffizienten: Regressions-koeffizient B	Standardfehler	Standardisierte Koeffizienten: Beta	T	Sig.
1	(Konstante)	5,996	,644		9,311	,000
	log(sales)	,207	,028	,489	7,457	,000
	age	-,022	,012	-,309	-1,925	,056
	male	-1,006	,699	-,560	-1,438	,152
	male_age	,025	,012	,840	1,982	,049

a. Abhängige Variable: log(salary)

Wie hoch ist der mittlere erwartete altersbedingte Lohnanstieg pro Altersjahr für Männer in diesem Modell? Wie hoch ist dieser Effekt für Frauen? (4 Punkte)

$$\left.\frac{\partial \log(\widehat{salary})}{\partial age}\right|_{male=1} = \hat{\beta}_2 + \hat{\beta}_4 \cdot male = (-0{,}022 + 0.025 \cdot 1) \cdot 100\% = 0{,}003 \cdot 100\% = 0{,}3\%$$

$$\left.\frac{\partial \log(\widehat{salary})}{\partial age}\right|_{male=0} = \hat{\beta}_2 \cdot 100\% = -0{,}022 \cdot 100\% = -2{,}2\%$$

Der mittlere erwartete altersbedingte Lohnanstieg beträgt c.p. für Männer 0,3% und für Frauen -2,2%.

Aufgabe 3: **[17,5 Punkte]**

Sie verfügen über einen Datensatz mit folgenden Informationen zu 160 Ländern:

lifeexp	Durchschnittliche Lebenserwartung in Jahren
lgdp_cap	Logarithmiertes Bruttoinlandsprodukt pro Kopf in $
calories	Durchschnittliche tägliche Kalorienaufnahme in Tausend Kalorien

Es kann beobachtet werden, dass die durchschnittliche Lebenserwartung der Bevölkerung mit dem BIP pro Kopf korreliert.

a) Welche Annahme ist nötig, damit eine Korrelation als kausaler Effekt interpretiert werden kann? Erläutern Sie kurz was unter dieser Annahme zu verstehen ist. (2 Punkte)

Nötige Annahme: $E(u|x) = 0$*, mittlere bedingte Unabhängigkeit des Störterms.*

Die Annahme bedeutet z.B., dass x nicht mit u korreliert oder dass der Erwartungswert von u an jeder Stelle von x gleich Null ist.

b) Anhand des Datensatzes wird folgende Gleichung geschätzt. Die Standardfehler der Koeffizienten sind in Klammern angegeben und R^2=0,708.

$$lifeexp = 20{,}941 + 6{,}251 \cdot lgdp\_cap + u$$

(2,535) (0,319)

i) Interpretieren Sie inhaltlich die Größe des Koeffizienten von *lgdp_cap.* (1 Punkt)

Steigt das BIP pro Kopf um 1%, so steigt die durchschnittliche Lebenserwartung im Durchschnitt c.p. um $\frac{6{,}251}{100} = 0{,}06$ *Jahre.*

ii) Geben Sie ein Beispiel für eine Situation, in der β_1 nicht den kausalen Effekt der Variable *lgdp_cap* auf *lifeexp* beschreibt. (2 Punkte)

Eine solche Situation liegt vor, wenn es einen Einflussfaktor gibt, der sowohl lifeexp beeinflusst, als auch mit lgdp_cap korreliert (z.B. durchschnittl. Ausbildungsniveau oder durchschnittl. Gesundheitszustand der Bevölkerung).

c) Sie nehmen die Variable *calories* ins Modell auf und schätzen

$$lifeexp = \beta_0 + \beta_1 lgdp\_cap + \beta_2 calories + \beta_3 calories2 + u$$

mit *calories2* = *calories*^{2}. SPSS liefert folgenden Output:

Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,870[a]	,757	,752	5,40043

a. Einflussvariablen : (Konstante), calories2, lgdp_cap, calories

Koeffizienten[a]

Modell		Nicht standardisierte Koeffizienten	
		Regressions-koeffizient B	Standardfehler
1	(Konstante)	-33,438	12,403
	lgdp_cap	5,049	,450
	calories	43,747	9,507
	calories2	-7,167	1,663

a. Abhängige Variable: lifeexp

i) Berechnen und interpretieren Sie inhaltlich den marginalen Effekt der täglichen Kalorienaufnahme auf die Lebenserwartung am Mittelwert (hier = 2) von calories. (2 Punkte)

Marginaler Effekt von calories: $\frac{\partial lifeexp}{\partial calories} = 43{,}747 - 2 \cdot 7{,}167 \cdot 2 = 15{,}079$.

Steigt die tägliche Kalorienaufnahme ausgehend vom Mittelwert um 1000 Kalorien, so steigt die durchschnittliche Lebenserwartung c.p. im Mittel um ca. 15 Jahre.

ii) Bei welcher Kalorienaufnahme ist die durchschnittliche Lebenserwartung maximal? (2 Punkte)

$43{,}747 - 2 \cdot 7{,}167 \cdot calories = 0$

$2 \cdot 7{,}167 \cdot calories = 43{,}747$

$calories^* = 3{,}05$

Die durchschnittliche Lebenserwartung ist bei einer tägl. Kalorienaufnahme von 3050 Kalorien maximal.

iii) Skizzieren Sie kurz grafisch den Effekt der Kalorienaufnahme auf die Lebenserwartung. Achten Sie dabei auf korrekte Achsenbeschriftungen. (2 Punkte)

Lösung graphisch.

iv) Testen Sie am 1% Signifikanzniveau, ob die im Vergleich zu Aufgabe b) zusätzlich aufgenommenen Parameter β_2 und β_3 gemeinsam signifikant sind. Geben Sie die Null- und Alternativhypothese, die Teststatistik, die Anzahl der Freiheitsgrade, den kritischen Wert und die Testentscheidung an. (4,5 Punkte)

H_0: $\beta_2 = \beta_3 = 0$

H_1: H_0 *trifft nicht zu*

$F = \frac{(R_U^2 - R_R^2)/q}{(1-R_U^2)/(n-k-1)} = \frac{(0{,}757-0{,}708)/2}{(1-0{,}757)/(160-3-1)}$=*15,738*

Freiheitsgrade: Zähler=2 (Anzahl der Restriktionen)

Freiheitsgrade: Nenner= $156\ (n-k-1 = 160-3-1)$

$F_c\ = F_{0{,}01;2;156}\ = 4{,}61$

Da $F > F_c$ *kann* H_0 *am 1% Signifikanzniveau abgelehnt werden. Die beiden Parameter sind gemeinsam signifikant.*

v) Bestimmen Sie die vorhergesagte durchschnittliche Lebenserwartung in einem Land mit einem BIP pro Kopf von 4000$ und durchschnittlicher täglicher Kalorienaufnahme von 3000 Kalorien. (2 Punkte)

$-33{,}438 + 5{,}049 \cdot log(4000) + 43{,}747 \cdot 3 - 7{,}167 \cdot 9 = 75{,}17\, Jahre$

Die vorhergesagte durchschnittliche Lebenserwartung beträgt in diesem Land 75.17 Jahre.

Aufgabe 4: **[25,5 Punkte]**

Wahr oder falsch? Tragen Sie für jede der folgenden Aussagen ein „w" für „wahr" oder ein „f" für „falsch" ein. Für jede richtige Antwort gibt es 0,5 Punkte, für jede falsche Antwort werden 0,5 Punkte abgezogen. Die Gesamtpunktzahl kann nicht negativ werden.

w	Die dummy variable trap entsteht durch perfekte Multikollinearität.
f	Die Aufnahme irrelevanter Variablen führt zu einer Verletzung der Gauss-Markov-Annahme MLR.1.
f	Die F-Verteilung folgt für n→∞ der Normalverteilung.
f	Ein p-Wert von 0,001 zeigt, dass der Parameter höchstens auf dem 10% Niveau signifikant ist.
f	Wenn ein kausaler Zusammenhang zwischen X und Y vorliegt, dann muss corr(X,Y)≠0.
f	Mit zunehmender Störtermvarianz sinkt die Varianz des geschätzten Koeffizienten $\hat{\beta}$.
w	Beim zweiseitigen Hypothesentest bezüglich eines Parameters würden t-Test und F-Test zum gleichen Ergebnis kommen.
f	Eine Nullhypothese, die am 5% Niveau verworfen wurde, muss am 10% Niveau nicht verworfen werden.
f	Der Bevölkerungsparameter β_k ist eine Zufallsvariable.
f	Die Nullhypothese bezieht sich auf geschätzte Werte der betrachteten Parameter.
w	Je größer α, desto kleiner das Konfidenzintervall eines Schätzers.
w	Beim einseitigen Test gegen die Alternativhypothese H_1: β_1<0 befindet sich der Ablehnungsbereich auf der linken Seite der Verteilung.
f	Je größer der Vorhersagefehler, desto kleiner die Varianz des geschätzten Koeffizienten β_k.
f	Bei einer empirischen Analyse sollte das ökonometrische Modell vor dem ökonomischen Modell aufgestellt werden.
f	Bei der ceteris paribus Interpretation eines Koeffizienten wird davon ausgegangen, dass alle anderen Größen gleich Null sind.
f	$E(g(x)) = g(E(x))$
w	Für n → ∞ konvergiert die t-Verteilung gegen die Standardnormalverteilung.
f	$\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2} = 0$, wenn x nicht konstant ist.
f	Ein Schätzer W des Parameters θ heißt unverzerrt, wenn für alle θ gilt: $plim(W) = \theta$.
w	Durch Umskalierung der abhängigen Variable ändert sich auch der Wert der Konstanten.
w	Wenn im Modell $y = \beta_0 + \beta_1 x + \beta_2 x^2 + u$ der Koeffizient $\beta_2 > 0$ ist, dann ergibt sich in der grafischen Darstellung der Abhängigkeit von y in Bezug auf x ein konvexer Verlauf.
w	Bei n → ∞ legitimiert die asymptotische Normalverteilung die Anwendung von t- und F-Tests.
w	Konsistenz und asymptotische Effizienz sind für n → ∞ definiert.
w	Ein geschätzter Parameter $\hat{\beta}$ mit einem Konfidenzintervall [-2; 2] ist statistisch insignifikant.
f	Ein $R^2 = 1$ kann ein Hinweis auf Multikollinearität im einfachen Regressionsmodell sein.
w	Der F-Statistik kann nie negativ werden.
w	Interaktionsterme werden verwendet, um für Teilgruppen unterschiedliche Steigungsparameter abzubilden.
w	Das angepasste Bestimmtheitsmaß $\overline{\mathrm{R}}^2$ kann verwendet werden, um genestete Modelle miteinander zu vergleichen.

w	Die Varianz des vorhergesagten Wertes $\hat{y}$ ist am kleinsten, wenn alle erklärenden Variablen an ihrem Stichprobenmittelwert betrachtet werden.
f	Dummyvariablen können nicht verwendet werden, um ordinale Informationen zu berücksichtigen.
f	Der Chow-Test ist ein Spezialfall des t-Tests.
f	Ein Vorteil des linearen Wahrscheinlichkeitsmodells ist, dass die vorhergesagten Werte außerhalb des [0;1]-Intervalls liegen können.
w	Die Gruppe, für die die Dummy-Variable mit Null kodiert ist, nennt man Referenzgruppe.
w	Die Wahrscheinlichkeitsdichtefunktion f beschreibt für jedes x die Wahrscheinlichkeit, mit der die diskrete Zufallsvariable X den Wert x annimmt.
w	Eine Zufallsvariable X mit Mittelwert σ und Varianz μ^2 lässt sich standardisieren als $Z = \frac{X-\sigma}{\mu}$.
w	Die Eigenschaft der Unverzerrtheit gilt unabhängig von der Stichprobengröße.
f	Konsistenz erfordert mittlere bedingte Unabhängigkeit $E(u\|x) = 0$.
f	Die BLUE-Eigenschaft gilt nicht, wenn die Residuen nicht normalverteilt sind.
w	Wenn die abhängige Variable nicht normalverteilt ist, sind t- und F-Tests nur asymptotisch gültig.
f	Wenn ein Regressor x_j mit dem Störterm korreliert, ist der Schätzer $\hat{\beta}_j$ verzerrt aber konsistent.
w	Ein positives Residuum $\hat{u}_i$ nach einer KQ-Schätzung bedeutet, dass y_i unterschätzt wird.
w	Im Modell $y = \beta_0 + \beta_1 x + u$ ist $\hat{\beta}_1 > 0$ wenn Cov(x,y)>0.
w	Bei $R^2 = 1$ gilt $\hat{u}_i = 0$ für jedes i.
w	Der Standardfehler des KQ-Schätzers ist eine Zufallsvariable.
w	Das Modell $y = \beta_0 + \frac{\beta_1}{x} + u$ ist linear in Parametern.
f	Ist die Varianz des Fehlerterms u für alle Beobachtungen identisch, so bezeichnet man ihn als heteroskedastisch.
f	Im einfachen Modell kann der KQ-Schätzer berechnet werden, solange Var(x)<0.
f	Je höher das R^2, desto niedriger ist die Wahrscheinlichkeit beim F-Test auf Gesamtsignifikanz die H_0 zu verwerfen.
f	Je mehr Restriktionen sie gemeinsam beim F-Test testen, desto höher ist das Signifikanzniveau.
w	Der KQ-Schätzer ist unverzerrt, wenn sein Erwartungswert gleich dem wahren Parameter ist.
f	Bei zwei unverzerrten Schätzern $\hat{\beta}_1$ und $\tilde{\beta}_1$ für einen unbekannten Bevölkerungsparameter β_1 ist $\tilde{\beta}_1$ effizient, wenn $E(\tilde{\beta}_1) < E(\hat{\beta}_1)$.

Aufgabe 5: **[17 Punkte]**

Welche Antwort ist richtig? Kreuzen Sie nur **eine Antwort** pro Aufgabe an. Falls mehrere Aussagen korrekt sind, kreuzen Sie **nur** die entsprechende **Antwortkombination** an. Für jede richtige Antwort gibt es 1 Punkt. Für falsche Antworten werden keine Punkte abgezogen.

1.	Die Annahme MLR.4 ist verletzt, wenn	
a	☑	die funktionale Form falsch spezifiziert wurde.
b	☐	die erklärenden Variablen multikollinear sind.
c	☐	irrelevante Variablen eingefügt wurden.
d	☐	Messfehler in der abhängigen Variable vorliegen.
e	☐	a und b
f	☐	a und d

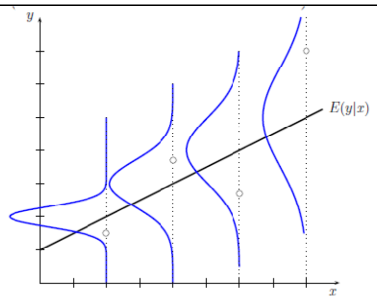
2.	Im einfachen linearen Regressionsmodell $y_i = \beta_0 + \beta_1 x_i + u_i$ mit $\beta_0 = 0$, ist	
a	☑	die Varianz des geschätzten Koeffizienten β_1 kleiner, wenn ohne Konstante geschätzt wird.
b	☐	die Varianz des geschätzten Koeffizienten β_1 größer, wenn ohne Konstante geschätzt wird.
c	☐	der geschätzte Koeffizient β_1 beim Schätzen ohne Konstante immer verzerrt.
d	☐	der geschätzte Koeffizient β_1 beim Schätzen ohne Konstante immer unverzerrt.
e	☐	a und d
f	☐	b und c

3.		Ein sehr hoher Wert der t-Statistik
a	☐	ergibt sich nur bei sehr hoher Varianz des geschätzten Koeffizienten.
b	☑	deutet auf statistische Signifikanz des Koeffizienten hin.
c	☐	kann auf eine sehr ungenaue Schätzung hindeuten.
d	☐	ergibt sich nur bei einem betraglich kleinen Koeffizienten.
e	☐	a und b
f	☐	a und c

4.		Perfekte Multikollinearität
a	☐	führt zu einer ungenauen Schätzung der Koeffizienten.
b	☐	liegt vor, wenn eine Variable linear und quadratisch berücksichtigt wird.
c	☐	führt zu kollinearen Parametern, welche sich mittels F-Test auf gemeinsame Signifikanz testen lassen.
d	☐	führt zu verzerrter Schätzung der Koeffizienten.
e	☐	c und d
f	☑	keine der Antworten

5.		Das korrekt spezifizierte Modell sei $\hat{y} = \widehat{\beta_0} + \widehat{\beta_1}x_1 + \widehat{\beta_2}x_2$, Sie schätzen aber $\tilde{y} = \widetilde{\beta_0} + \widetilde{\beta_1}x_1$.
a	☐	$\widetilde{\beta_1}$ ist nach oben verzerrt, wenn $cov(x_1 x_2) < 0$ und $\beta_2 < 0$
b	☐	$\widetilde{\beta_1}$ ist unverzerrt, wenn $cov(x_1 x_2) = 0$ und $\beta_2 < 0$
c	☐	$\widetilde{\beta_1}$ ist unverzerrt, wenn $cov(x_1 x_2) < 0$ und $\beta_2 = 0$
d	☐	$\widetilde{\beta_1}$ ist nach unten verzerrt, wenn $cov(x_1 x_2) < 0$ und $\beta_2 > 0$
e	☐	a und d
f	☑	alle genannten Antworten

6.



Was ist in dieser Abbildung dargestellt?

a	☐	Heteroskedastizität mit konstanter Varianz von u
b	☐	Heteroskedastizität mit abnehmender Varianz von u bei steigendem x.
c	☐	Homoskedastizität mit zunehmender Varianz von u bei steigendem x.
d	☑	Heteroskedastizität mit zunehmender Varianz von u bei steigendem x.
e	☐	Homoskedastizität mit abnehmender Varianz von u bei steigendem x.
f	☐	keine der genannten Antworten

7.		Die Eigenschaft der Unverzerrtheit
a	☐	ist auch für kleine Stichproben definiert.
b	☐	ist nur für große Stichproben definiert.
c	☐	ist im linearen Regressionsmodell eine asymptotische Eigenschaft.
d	☐	gilt unter MLR.1–MLR.4.
e	☐	b und c
f	☑	a und d

8.		Im Modell $\log(y) = \beta_0 + \beta_1 \log(x) + u$ mit $\beta_1 > 1$
a	☐	ist β_1 der marginale Effekt von x auf y.
b	☐	ist β_1 der marginale Effekt von log(x) auf y.
c	☐	resultiert eine 1% Änderung in y aus einer 1% Änderung in x.
d	☑	wird β_1 als Elastizität interpretiert.
e	☐	a und d
f	☐	a, c und d

9.		Gilt $E(u|x) = 0$ dann
a	❑	kann der KQ-Schätzer als kausal interpretiert werden.
b	❑	ist der Störterm gleich Null, unabhängig von x.
c	❑	$Cov(x, u) = 0$.
d	❑	ist die ceteris paribus Interpretation legitim.
e	☑	a, c und d
f	❑	alle Antworten

10.		Unter den Gauss-Markov Annahmen (MLR.1–MLR.5) ist der KQ-Schätzer
a	❑	unverzerrt.
b	❑	konsistent.
c	❑	effizient.
d	❑	asymptotisch effizient.
e	❑	a und c
f	☑	alle Antworten

11.		Wenn das Modell $y = \beta_1 \mathrm{x}_1 + \beta_2 \mathrm{x}_2 + u$ mit der KQ-Methode geschätzt wird,
a	❑	wird $y - \hat{\beta}_1 \mathrm{x}_1 - \hat{\beta}_2\, \mathrm{x}_2$ minimiert.
b	❑	kann man das R^2 als Maß der Schätzgüte verwenden.
c	❑	ist $\hat{y} = \hat{\beta}_1 + \hat{\beta}_2 + \mathrm{u}$ der vorhergesagte Wert von y an den Stellen $\mathrm{x}_1 = 1$ und $\mathrm{x}_2 = 1$.
d	❑	ist der Mittelwert der Residuen gleich Null.
e	❑	c und d
f	☑	keine der Antworten

12.		Im Modell $\log(y) = \beta_0 + \beta_1 x_1 + \beta_2 \log(x_2) + u$
a	☑	kann β_1 als Semielastizität interpretiert werden.
b	❑	kann β_2 als Semielastizität interpretiert werden.
c	❑	ändert sich bei einer einprozentigen Änderung von x_2 y um β_2 Prozentpunkte.
d	❑	ändert sich bei einer einprozentigen Änderung von x_1 y um β_1 Prozent.
e	❑	a und c
f	❑	c und d

13.		Bei konsistenten Schätzverfahren
a	❑	ist der Erwartungswert des Schätzers mit dem wahren Wert identisch.
b	❑	liegt der Wahrscheinlichkeitsgrenzwert des Schätzers umso näher am wahren Wert, je größer die Stichprobe.
c	❑	ist der Schätzer unverzerrt.
d	❑	sinkt die Varianz des Schätzers, je größer die Stichprobe.
e	☑	b und d
f	❑	alle Antworten

14.		Sei *X* eine Zufallsvariable mit $E(X) = \mu$. Dann gilt, dass
a	❑	$var(X) \equiv E[(X - \mu)^2] = \sigma^2$.
b	❑	$\sigma^2 = E(X^2) - \mu^2$.
c	❑	$sd(aX) = a^2 sd(X)$.
d	❑	die Varianz von *X* zwischen 0 und 1 liegt.
e	☑	a und b
f	❑	a und d

15.		Im Modell $y = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \beta_3 x_2 + u$ mit $\beta_1 > 0$
a	❑	ist der marginale Effekt von x_1 auf *y* konstant wenn $\beta_2 \neq 0$.
b	☑	ergibt sich ein mit x_1 steigender marginaler Effekt von x_1 auf *y* wenn $\beta_2 > 0$.
c	❑	ist β_1 inkonsistent geschätzt wenn $\beta_2 > 0$.
d	❑	hängt der marginale Effekt von x_1 auf *y* von der Ausprägung von x_2 ab.
e	❑	b und c
f	❑	keine der Antworten

16.	Wenn in einem linearen Modell die abhängige Variable y eine Dummyvariable ist, dann	
a	❑	entsteht das Problem der dummy variable trap.
b	❑	gilt $\Delta P(x = 0\|y) = \beta_j \Delta x_j$.
c	☑	variiert die Varianz des Störterms über die Beobachtungseinheiten.
d	❑	können keine Dummyvariablen als erklärende Variablen verwendet werden.
e	❑	a und c
f	❑	b und c

17.	Interaktionsterme in Modellen werden verwendet, um	
a	❑	die Stichprobengröße zu erhöhen.
b	❑	Homoskedastie aufzudecken.
c	❑	das R^2 zu bestimmen.
d	❑	Multikollinearität zu beheben.
e	❑	a und b
f	☑	keine der Antworten