Bachelorprüfung

Fach: Praxis der empirischen Wirtschaftsforschung
Prüfer: Prof. Regina T. Riphahn, Ph.D.

Name, Vorname	
Matrikelnr.	
E-Mail	
Studiengang	
Semester	
Datum	
Raum, Platz	
Unterschrift	

Vorbemerkungen:

Anzahl der Aufgaben: Die Klausur besteht aus 6 Aufgaben, die alle bearbeitet werden müssen.

Bewertung: Es können maximal 90 Punkte erworben werden. Die maximale Punktzahl ist für jede Aufgabe in Klammern angegeben. Sie entspricht der für die Aufgabe empfohlenen Bearbeitungszeit in Minuten.

Erlaubte Hilfsmittel:
- Formelsammlung (ist der Klausur beigefügt)
- Tabellen der statistischen Verteilungen (sind der Klausur beigefügt)
- Taschenrechner
- Fremdwörterbuch

Wichtige Hinweise:
- Sollte es vorkommen, dass die statistischen Tabellen, die dieser Klausur beiliegen, den exakten Wert der Freiheitsgrade nicht ausweisen, machen Sie dies kenntlich und verwenden Sie den nächstgelegenen Wert.
- Sollte es vorkommen, dass bei einer Berechnung eine erforderliche Information fehlt, machen Sie dies kenntlich und treffen Sie für den fehlenden Wert eine plausible Annahme.

Aufgabe 1: **[14 Punkte]**

Die Bildungsrendite wird auf Basis einer Stichprobe vom Umfang 1045 für männliche Erwerbstätige mit folgenden Regressionsfunktionen empirisch analysiert.

Modell 1: $\log(\text{wage}) = \beta_0 + \beta_1\text{educ} + \beta_2\text{exper} + \beta_3\text{IQ} + u$
Modell 2: $\log(\text{wage}) = \beta_0 + \beta_1\text{educ} + \beta_2\text{exper} + \beta_3\text{IQ} + \beta_4\text{IQfather} + u$

Dabei ist:

wage	Stundenlohn in US-Dollar
educ	Schul- und Berufsausbildung in Jahren
exper	Berufserfahrung in Jahren
IQ	Intelligenzquotient in Punkten
IQfather	Intelligenzquotient des Vaters in Punkten

Der Regressionsoutput von SPSS ist in folgender Tabelle zusammengefasst.

	Modell 1		Modell 2	
	Koeffizient	Standardfehler	Koeffizient	Standardfehler
educ	0.078	0.007	0.078	0.007
exper	0.050	0.004	0.050	0.004
IQ	0.008	0.001	0.008	0.006
IQfather	--	--	-0.003	0.006
Konstante	4.375	0.129	4.374	0.129
SSE	36.02		36.07	
SSR	155.56		155.51	
SST	191.58		191.58	

a) Interpretieren Sie den Effekt eines weiteren Jahres Schul- und Berufsausbildung aus Modell 1 inhaltlich, wenn die anderen Variablen im Modell konstant gehalten werden. (2 Punkte)

Log-level Modell: Ein weiteres Jahr Ausbildung bringt ceteris paribus im Durchschnitt eine Rendite in Höhe von 7,8%.

b) Das angepasste Bestimmtheitsmaß $\bar{R}^2$ für Modell 1 beträgt 0,1857. Berechnen Sie $\bar{R}^2$ für Modell 2 und beurteilen Sie, ob die Aufnahme der Variable IQfather den Erklärungsgehalt des Modells verbessert. (4 Punkte)

Modell 2: Adj. R^2=1-SSR/SST(n-1)/(n-k-1)=1-155,51/191,58*(1045-1)/(1045-4-1)=0,1851*

Der Erklärungsgehalt erhöht sich durch die Aufnahme der zusätzlichen Variable nicht.

c) Der Parameterschätzer für β_4 in Modell 2 weist einen p-Wert von 0,55 auf. Erläutern Sie anhand dieses Wertes die statistische Signifikanz des Effekts. (2 Punkte)

Die H_0: $\beta_4 = 0$ kann verworfen werden, wenn $p < \alpha$. Bei p=0,55 ist der Parameter jedoch auf jedem relevanten Signifikanzniveau α (1%, 5%, 10%) insignifikant.

d) Erläutern und begründen Sie formal, inwiefern die Aufnahme der Variable IQfather in das Modell die Präzision der Schätzergebnisse für β_3 verschlechtern könnte. Werden Ihre Überlegungen durch den Regressionsoutput bestätigt? (6 Punkte)

Standardfehler von $\hat{\beta}_3$:

- *Der Parameter wird durch die Berücksichtigung der irrelevanten Variable IQfather mit zu großer Varianz bzw. zu großem Standardfehler geschätzt.*

- *Formale Begründung: Es gilt* $\mathrm{se}(\hat{\beta}_3) = \hat{\sigma}/\sqrt{\mathrm{SST}_3(1-R_3^2)}$. *Dabei ist* R_3^2 *das Bestimmtheitsmaß einer Regression von IQ auf alle anderen erklärenden Variablen. Da IQ und IQfather wohl hoch miteinander korrelieren, dürfte* R_3^2 *in Modell 2 deutlich größer sein als* R_3^2 *in Modell 1. Folglich wird der Standardfehler größer geschätzt.*

- *Der Regressionsoutput bestätigt dies: Der Standardfehler des Parameterschätzers der Variable IQ wird in Modell 2 sechs mal so groß geschätzt.*

Aufgabe 2: **[21 Punkte]**

Sie interessieren sich für die Determinanten des Geburtsgewichts von Säuglingen. Für Ihre Analyse stehen Ihnen folgende Informationen für 1388 Neugeborene zur Verfügung:

bweight	Gewicht eines Kindes bei der Geburt in Pfund
cigs	Anzahl der von der Mutter während der Schwangerschaft durchschnittlich gerauchten Zigaretten pro Tag
feduc	Schulausbildung des Vaters in Jahren
meduc	Schulausbildung der Mutter in Jahren
male	= 1 wenn männlich, = 0 sonst
east	= 1 wenn Eltern aus den neuen Bundesländern stammen, = 0 sonst

Für die Grundgesamtheit wird folgendes Modell unterstellt:

$$bweight_i = \beta_0 + \beta_1 cigs_i + \beta_2 feduc_i + \beta_3 meduc_i + \beta_4 male_i + \beta_5 east_i + u_i$$

Eine Schätzung in SPSS ergibt:

Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,215[a]	,046	,042	1,23195

a. Einflußvariablen : (Konstante), cigs, male, feduc, meduc, east

Koeffizienten[a]

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten		
		Regressionskoeffizi entB	Standardfehler	Beta	T	Sig.
1	(Konstante)	7,013	,227		30,832	,000
	cigs	-,037	,007	-,159	-5,459	,000
	feduc	,032	,017	,070	1,866	,062
	meduc	-,021	,019	-,040	-1,062	,288
	male	,228	,072	,090	3,183	,001
	east	,303	,099	,087	3,061	,002

a. Abhängige Variable: bweight

a) Interpretieren Sie $\hat{\beta}_1$ inhaltlich und bestimmen Sie das zugehörige 99% Konfidenzintervall. (5 Punkte)
Raucht die Mutter durchschnittlich eine Zigarette mehr am Tag während der Schwangerschaft, reduziert sich c.p. das erwartete Geburtsgewicht ihres Kindes um 0,037 Pfund.

$$KI = [\hat{\beta}_1 - t_{\alpha/2,n-k-1} \cdot se(\hat{\beta}_1);\ \hat{\beta}_1 + t_{\alpha/2,n-k-1} \cdot se(\hat{\beta}_1)]$$

$$KI = [-0.037 - 2{,}576 \cdot 0{,}007; -0{,}037 + 2{,}576 \cdot 0{,}007] = [-0{,}055; -0{,}019]$$

b) Was versteht man unter dem sogenannten „Dummy Variable Trap“? Zeigen Sie anhand der Varianzformel für den KQ-Schätzer, welche Folge für die Schätzung der Koeffizienten zu erwarten ist. (4 Punkte)

Entsteht, wenn bei einer abschließenden Anzahl von Kategorien für ein Merkmal für jede Kategorie ein Dummy in das Modell aufgenommen wird und sich zusätzlich eine Konstante im Modell befindet.

$$Var(\widehat{\beta_j}) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2(1 - R_j^2)}$$

R_j^2 in Varianzformel der Dummykoeffizienten wird 1 (perfekte Multikollinearität).

Schätzer sind nicht mehr zu identifizieren, da Varianz gegen unendlich geht.

c) Wie könnten Sie durch das Einfügen einer zusätzlichen Variable testen, ob Jungs, deren Eltern aus Ostdeutschland stammen signifikant schwerer sind als Jungs, deren Eltern aus Westdeutschland stammen? (2 Punkte)

*Interaktionsterm male*east einfügen und statistische Signifikanz des Koeffizienten überprüfen.*

d) Sind die beiden Variablen *feduc* und *meduc* jeweils am 5% Niveau signifikant? Begründen Sie kurz. Warum ist es sinnvoll, die beiden Variablen in einem F-Test gemeinsam zu testen? (4 Punkte)

Nein, da die p-Werte beider Variablen größer 0,05 sind.
Schulausbildung der Mutter und Schulausbildung des Vaters u. U. korreliert. Dies wird beim t-Test nicht berücksichtigt.

e) Führen Sie für die beiden Variablen *feduc* und *meduc* den F-Test am 5% Signifikanzniveau durch. Verwenden Sie hierfür zusätzlich untenstehenden Output. Es sind Null- und Alternativhypothese, Teststatistik, kritischer Wert, Freiheitsgrade und Testentscheidung anzugeben. (6 Punkte)

Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,209[a]	,044	,042	1,24532

a. Einflußvariablen : (Konstante), east, cigs, male
b. Abhängige Variable: bweight

- *Hypothesen:* $H_0: \beta_2 = \beta_3 = 0$
  $H_1: H_0 \ gilt \ nicht$
- *Teststatistik:* $\frac{(R_u^2 - R_r^2)/q}{(1-R_u^2)/(n-k-1)} = \frac{(0{,}046-0{,}044)/2}{(1-0{,}046)/(1388-5-1)}$ *=1,45*
- *Kritischer Wert:* $F_{\alpha;q,n-k-1} = F_{0{,}05;2{,}1382} = 3{,}00$
- *Vergleich Teststatistik mit kritischem Wert: 1,45<3,00*
  Testentscheidung: Die Nullhypothese kann am 5% Signifikanzniveau nicht verworfen werden. Die Parameter sind auch gemeinsam nicht signifikant von Null verschieden.

Aufgabe 3: **[8 Punkte]**

Das Engelsche Gesetz besagt, dass die Ausgaben für Ernährung vom Haushaltseinkommen abhängen. Sie versuchen, diese ökonomische Hypothese in einem Modell darzustellen. Ihr Datensatz enthält folgende Variablen für 58 Haushalte:

consum	wöchentliche Ausgaben für Ernährung (in €)
wage	wöchentliches Einkommen (in €)
$wage^2$	quadriertes wöchentliches Einkommen

Sie schätzen das Modell $consum_i = \beta_0 + \beta_1 wage_i + \beta_2 wage_i{}^2 + u_i$
und erhalten folgenden SPSS-Output:

Koeffizienten[a]

Modell		Nicht standardisierte Koeffizienten: Regressionskoeffizient B	Nicht standardisierte Koeffizienten: Standardfehler	Standardisierte Koeffizienten: Beta	T	Sig.
1	(Intercept)	50,287	14,812		3,395	,001
	wage	1,393	,096	2,225	14,566	,000
	$wage^2$	-0,001224	,000	-1,298	-8,495	,000

a. Abhängige Variable: consum

a) Bei welchem Einkommen sind die durchschnittlichen Ernährungsausgaben maximal? (3 Punkte)

$$\frac{\partial consum}{\partial wage} = \beta_1 + 2\beta_2 wage$$

$$\beta_1 + 2\beta_2 wage = 0$$

$$wage = \frac{-\beta_1}{2\beta_2} = \frac{-1{,}393}{-2 \cdot 0{,}001224} = 569{,}04$$

Die durchschnittlichen Ernährungsausgaben erreichen ihr Maximum bei einem wöchentlichen Einkommen von 569,04€.

b) Bestimmen Sie die marginalen Effekte für Einkommen von 270€ und 500€ und interpretieren Sie diese. (3 Punkte)

$$\frac{\partial consum}{\partial wage} = \beta_1 + 2\beta_2 wage$$

$$\frac{\partial consum}{\partial wage} = 1{,}393 - 2 \cdot 0{,}001224 \cdot 270 = 0{,}73$$

Eine Erhöhung des wöchentlichen Einkommens um einen € auf einem Einkommensniveau von 270€ pro Woche erhöht c.p. die durchschnittlichen wöchentlichen Ausgaben für Ernährung um 0,73€.

$$\frac{\partial consum}{\partial wage} = 1{,}393 - 2 \cdot 0{,}001224 \cdot 500 = 0{,}169$$

Eine Erhöhung des wöchentlichen Einkommens um einen € auf einem Einkommensniveau von 500€ pro Woche erhöht c.p. die durchschnittlichen wöchentlichen Ausgaben für Ernährung um 0,169€.

c) Formulieren Sie das Modell um, sodass Ihnen der geschätzte Lohn-Koeffizient eine Elastizität angibt. (2 Punkte)

$$log(consum) = \beta_0 + \beta_1 \, log(wage) + u$$

Aufgabe 4: **[12 Punkte]**

Die USA importiert eine Chemikalie aus 31 verschiedenen Ländern. Eine Gruppe von 6 Ländern hat einen Zoll auf den Chemikalienexport. Der Datensatz beschreibt die Handelsbeziehung der USA zu jedem der 31 Länder und beinhaltet folgende Variablen:

vol	Importvolumen der Chemikalie aus Land *i* in Tonnen
prod	Volumen der Produktion der Chemikalie im Land *i* in Tonnen
dist	Distanz zwischen den USA und dem Land *i* in km
wk	Wechselkurs (Dollar gegenüber Währung in Land *i*)
zoll	Dummy-Variable; 1 = Land *i* verlangt Zoll, 0 = sonst

Sie schätzen zunächst das Modell

$$\log(vol) = \beta_0 + \beta_1 \log(prod) + \beta_2 \log(dist) + \beta_3 \log(wk) + u$$

und erhalten folgenden Output:

Modell: n=31; SSR=44,274; SST=63,534; R^2=0,304

Koeffizienten[a]

	RegressionkoeffizientB	Standardfehler	T	Sig.
log(prod)	3,107	,474	6,56	0,000
log(dist)	-,214	,897	-0,24	0,812
log(wk)	,993	,37	2,68	0,008
Konstante	-18,201	20,834	-0,87	0,384

a. Abhängige Variable: log(vol)

a) Interpretieren Sie das Schätzergebnis für β_2 statistisch und inhaltlich. (2 Punkte)

1% größere Entfernung zwischen den beiden Ländern ist ceteris paribus korreliert mit einem Importrückgang von 0,214%. Der Parameterschätzer ist statistisch nicht signifikant von null verschieden.

b) Um zu prüfen, ob sich die Handelsbeziehungen für Länder mit und ohne Zoll unterscheiden, führen Sie einen Chow-Test durch. Dazu schätzen Sie nach dem gepoolten Modell zwei weitere Modelle, welche Ihnen folgende Information liefern:

Modell für *zoll=1*: n=6; SSR=0,629; SST=2,250; R^2=0,721
Modell für *zoll=0*: n=25; SSE=19,192; SST=61,284; R^2=0,313

Führen Sie einen Chow-Test durch. Geben Sie Nullhypothese, Teststatistik, Freiheitsgrade, den kritischen Wert und Ihre Testentscheidung an. (7 Punkte)

$$H_0 : \beta_{j,zoll=0} = \beta_{j,zoll=1} \quad j = 0,...,3$$
$$H_1 : H_0 \; falsch$$
$$F = \frac{[SSR_P - (SSR_1 + SSR_2)]/(k+1)}{(SSR_1 + SSR_2)/[n-2(k+1)]}$$

$$Zwischenschritt: SSR_2 = SST_2 - SSE_2 = 61{,}284 - 19.192 = 42{,}092$$

$$F = \frac{[44{,}274-(0{,}629+42{,}092)]/(3+1)}{(0{,}629+42{,}092)/[31-2(3+1)]} = \frac{0{,}3883}{1{,}8574} = 0{,}2090$$

$$F_{4;23;5\%} = 2{,}80$$

H_0 kann nicht verworfen werden, da $F<F_c$.

c) Geben Sie eine alternative Vorgehensweise für den Chow-Test an, um mit nur einem Modell Unterschiede in den Importdeterminanten für Länder mit und ohne Zoll untersuchen und testen zu können. Stellen Sie die unrestringierte Schätzgleichung dar. (3 Punkte)

-vollständig interagiertes Modell mit Zoll und F-Test auf gemeinsame Signifikanz der Interaktionsterme.
-Schätzgleichung:

$$\log(vol) = \beta_0 + \beta_1 \log(prod) + \beta_2 \log(dist) + \beta_3 \log(wk) +$$
$$+\beta_4 zoll + \beta_5 zoll \cdot \log(prod) + \beta_6 zoll \cdot \log(dist) + \beta_7 zoll \cdot \log(wk) + u$$

Aufgabe 5: **[21 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	Mittelwerte sind stärker von Ausreißern beeinflusst als Medianwerte.
F	Die Normalverteilung mit einem Erwartungswert von 1 bezeichnet man als Standardnormalverteilung.
W	Der LM-Test kann zum Testen mehrerer linearer Restriktionen genutzt werden.
F	Ist die Annahme MLR6 verletzt, gilt die BLUE-Eigenschaft nicht mehr.
W	Der mit der KQ-Methode geschätzte Parameter ist eine Zufallsvariable.
F	Bei Regressionen durch den Ursprung hat der Steigungsparameter den Wert eins.
W	Die Annahme E(u|x)=0 impliziert Cov(x,u)=0.
F	Die Standardabweichung einer Konstanten ist 1.
F	Ein Schätzer ist konsistent, wenn mit zunehmender Stichprobengröße die Varianz des Schätzers steigt.
W	Eigenschaften des KQ-Schätzers, die nur für eine große Stichprobe ($n \rightarrow \infty$) definiert sind, nennt man asymptotische Eigenschaften.
W	Die χ^2-Verteilung nimmt nur positive Werte an und ist nicht symmetrisch.
W	Zufallsvariablen sind i.i.d., wenn sie statistisch unabhängig und mit der gleichen Dichtefunktion verteilt sind.
W	Die Nullhypothese $H_0\colon \beta_1 + \beta_2 = \theta$ kann mittels t-Test getestet werden.
W	Ein R^2 von eins bedeutet, dass die Residuen alle den Wert null annehmen.
F	Durch Umskalierung einer der erklärenden Variablen können die Effekte der anderen Variablen verzerrt werden.
W	Die Residuen einer Regression berechnen sich als Differenz zwischen tatsächlichen und vorhergesagten Werten für die abhängige Variable.
F	Bei perfekter Multikollinearität geht die Störtermvarianz gegen unendlich.
W	Ändern sich die Koeffizienten einer KQ-Schätzung bei Aufnahme einer zusätzlichen Variablen waren die Schätzer im Ausgangsmodell unter Umständen verzerrt.
F	Die Aufnahme irrelevanter Variablen verzerrt die Koeffizienten der anderen Regressoren.
W	Das Gauss-Markov Theorem besagt, dass der KQ-Schätzer bei Gültigkeit der Annahmen MLR1-MLR5 unter allen linearen, unverzerrten Schätzverfahren die kleinste Varianz hat.
F	Bei einer Regression durch den Ursprung verläuft die Regressionsgerade stets durch den Mittelwert der Daten $(\bar{x}, \bar{y})$.
F	Wenn die Störtermvarianz nicht konstant ist, ergeben sich verzerrte KQ-Schätzer.
W	Messfehler in einer der erklärenden Variablen können dazu führen, dass die Koeffizienten verzerrt sind.
F	Die Verwendung von Polynomen höherer Ordnung einer erklärenden Variablen führt zu perfekter Multikollinearität.
W	Der KQ-Schätzer im einfachen linearen Regressionsmodell entspricht dem Verhältnis der Stichprobenkovarianz zwischen erklärender und abhängiger Variable zur Stichprobenvarianz der erklärenden Variablen.
W	Je kleiner $1-\alpha$, desto enger ist ein Konfidenzintervall.
W	Der p-Wert im F-Test auf Gesamtsignifikanz eines Modells beschreibt die Wahrscheinlichkeit bei einer F-verteilten Zufallsvariable einen größeren Wert als den der errechneten Teststatistik zu finden.
W	Ein Typ 2 Fehler wird begangen, wenn eine falsche H_0 nicht verworfen wird.
W	In einem Test auf Gesamtsignifikanz eines Modells werden so viele Restriktionen getestet wie sich erklärende Variablen im Modell befinden.
F	Liegt der kritische Wert eines zweiseitigen t-Tests genau am 90. Perzentil der t_{n-k-1} Verteilung, wurde ein Signifikanzniveau von 90 Prozent gewählt.
W	Der kritische Wert eines zweiseitigen t-Tests ist bei gleichem Signifikanzniveau stets größer als der eines einseitigen t-Tests.
F	Je höher das R^2, desto eher trifft die Annahme E(u|x) = 0 zu, da die Modellanpassung an die Daten besser ist.
W	Ein Chow-Test kann über getrennte Schätzungen für zwei Gruppen durchgeführt werden.
F	Over-controlling liegt vor, wenn zu viele abhängige Variablen in einer Gleichung berücksichtigt werden.
F	In Regressionsfunktionen können nur quantitative Variablen analysiert werden, da die Mittelwerte von qualitativen Variablen nicht sinnvoll interpretiert werden können.

W	Mit Interaktionseffekten können für unterschiedliche Gruppen verschiedene Steigungsparameter bestimmt werden.
W	Ein lineares Wahrscheinlichkeitsmodell kann vorhergesagte Werte außerhalb des [0;1] Intervalls generieren.
W	Eine Dummy-Variable wird in Bezug auf eine Referenzgruppe interpretiert.
W	Ausreißer-Beobachtungen können durch Logarithmierung der betroffenen Variable an Einfluss verlieren.
F	Variablen mit Ausprägungen ≤ 0 werden in ökonometrischen Modellen häufig logarithmiert.
F	Die Effekte von standardisierten Variablen können nicht direkt miteinander verglichen werden.
F	Interaktionsterme werden stets in zweiter Potenz ins Schätzmodell integriert.

Aufgabe 6: **[14 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.	Ein Intervallschätzer	
a	❑	ist bei sonst gleichen Bedingungen für ein Signifikanzniveau von 0,05 breiter als für ein Signifikanzniveau von 0,1.
b	❑	wird unter Verwendung der geschätzten Varianz des Störterms berechnet.
c	❑	liefert nur für Querschnittsdaten ein interpretierbares Ergebnis.
d	❑	wird enger, wenn der zugehörige Punktschätzer sinkt.
e	☒	a und b.
f	❑	c und d.

2.	Der Effekt eines Regressors x_1 im Modell $\hat{y} = \hat{\beta}_o + \hat{\beta}_1 x_1$ ist bei negativer Korrelation mit einer ausgelassenen erklärenden Variable	
a	❑	unterschätzt, wenn die ausgelassene Variable einen positiven Effekt auf y hat.
b	❑	überschätzt, wenn die ausgelassene Variable einen negativen Effekt auf y hat.
c	❑	unterschätzt, wenn die ausgelassene Variable einen negativen Effekt auf y hat.
d	☒	a und b.
e	❑	a und c.
f	❑	keine der Antworten.

3.	Bei welchen der folgenden Modelle ist aufgrund ihrer Spezifikation mindestens eine der Gauss-Markov Annahmen zum linearen Regressionsmodell verletzt?	
a	☒	$\log(y) = \beta_0 + \beta_1 \log(x) + \beta_2 \log(x^2) + u$
b	❑	$\log(y) = \beta_0 + \beta_1 \log(x) + \beta_2 [\log(x)]^2 + u$
c	❑	$y = \beta_0 + \beta_1 \sqrt{x} + u$
d	❑	$y = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + u$
e	❑	a und c.
f	❑	a und d.

4.	Die Varianz des KQ-Schätzers β_1 ist ceteris paribus umso kleiner, je	
a	❑	größer die Varianz der Störterme ist.
b	☒	größer der Stichprobenumfang n ist.
c	❑	kleiner die Streuung der erklärenden x Variablen.
d	❑	größer die Streuung der abhängigen y Variable.
e	❑	a und c.
f	❑	keine der genannten Antworten.

5.	Der Zusammenhang zwischen zwei Variablen	
a	❑	wird mit dem Standardfehler gemessen.
b	☒	wird mit dem Korrelationskoeffizienten als linearer Zusammenhang gemessen.
c	❑	wird nur mit experimentellen Daten verlässlich gemessen.
d	❑	wird mit der Standardabweichung gemessen.
e	❑	a und b.
f	❑	a und d.

6.	Ein unter den Gauss-Markov Annahmen unverzerrter Schätzer für die Varianz der Störterme u_i im Modell $y_i = \beta_0 + \beta_1 x_i + u_i$ ist	
a	❑	$\sum_i u_i^2$
b	❑	$\sum_i (u_i - \bar{u})^2$
c	❑	$E(u_i - \bar{u})^2$
d	❑	$\frac{1}{n}\sum_i u_i^2$
e	❑	$\frac{1}{n-1}\sum_i u_i^2$
f	☒	keine der genannten Antworten.

7.	Die Annahme $x \sim i.i.d.(\mu, \sigma^2)$ impliziert unter anderem, dass	
a	❑	$E(\mu) = 0$
b	❑	$Var(\mu) = \sigma^2$
c	☒	$Var(x) = \sigma^2$
d	❑	$E(\mu) = x$
e	❑	a und b.
f	❑	c und d.

8.	Beim Umskalieren der abhängigen Variable im Modell $y = \beta_0 + \beta_1 x + u$ ändern sich	
a	❑	die geschätzten Koeffizienten nicht.
b	❑	die Standardfehler.
c	❑	die t-Statistiken nicht.
d	❑	die F-Statistik.
e	❑	a und d.
f	☒	b und c.

9.	Der p-Wert in einem einseitigen t-Test mit H_1: $\beta_1 > 0$	
a	❑	ist das Signifikanzniveau des Tests, in dem der empirische t-Wert gleich dem kritischen Wert der t-Verteilung ist.
b	❑	kann nicht größer 1 sein.
c	❑	gibt die Wahrscheinlichkeit an, dass die t-Verteilung einen Wert annimmt, der größer ist als der berechnete t-Wert.
d	❑	a und b.
e	☒	a, b und c.
f	❑	keine der Antworten.

10.	Das angepasste Bestimmtheitsmaß $\bar{R}^2$	
a	❑	kann nicht sinken, wenn zusätzliche Variablen in das Modell aufgenommen werden.
b	☒	steigt, sobald der Betrag der t-Statistik einer zusätzlich aufgenommenen Variable ≥ 1 ist.
c	❑	kann nicht negativ werden.
d	❑	hat die gleiche Interpretation wie das Bestimmtheitsmaß R^2.
e	❑	a und c.
f	❑	b und d.

11.	Dummy-Variablen	
a	☐	nehmen nur Werte zwischen 0 und 1 an.
b	☒	nehmen typischerweise die Ausprägungen 0 und 1 an.
c	☐	können mehr als zwei Ausprägungen haben.
d	☐	sind für qualitative Informationen nicht definiert.
e	☐	a, b, c und d.
f	☐	keine der Antworten.

12	Der Typ 2 Fehler	
a	☒	wird begangen, wenn eine falsche H_0 nicht verworfen wird.
b	☐	wird begangen, wenn eine wahre H_0 verworfen wird.
c	☐	entspricht der Wahrscheinlichkeit 1-α.
d	☐	entspricht dem Signifikanzniveau α.
e	☐	Antworten a und c sind zutreffend.
f	☐	Antworten a, b, und d sind zutreffend.

13.	Die H_0: $\beta_1 = \beta_2 = \beta_3 = 0$ in einem F-Test auf dem 5% Signifikanzniveau für ein Modell mit Konstante β_0 (19 Beobachtungen)	
a	☐	wird verworfen, wenn die Anzahl der Regressoren sieben und F = 3,58 beträgt.
b	☐	wird verworfen, wenn die Anzahl der Regressoren sieben und F = 7,0 beträgt.
c	☐	wird verworfen, wenn die Beobachtungszahl auf 100 steigt, die Anzahl der Regressoren sieben und F = 2,72 beträgt.
d	☐	a und c.
e	☒	b und c.
f	☐	keine der Antworten.

14.	Der KQ-Schätzer	
a	☐	minimiert die Summe der quadrierten vertikalen Abweichungen von der Regressionsgeraden.
b	☐	minimiert die Residuenquadratsumme.
c	☐	minimiert den unerklärten Teil der Streuung in der abhängigen Variable.
d	☐	maximiert R^2.
e	☐	a und b.
f	☒	Alle Antworten sind zutreffend.