

Aufgabe 1

[14 Punkte]

Sie möchten untersuchen, wovon die Abwesenheit der Studierenden in den Vorlesungen an einer Universität abhängt. Sie verfügen über einen Datensatz zu 282 Studierenden mit folgenden Variablen:

| | |
|----------------|---|
| <i>skipped</i> | durchschnittliche Anzahl von verpassten Vorlesungen pro Woche |
| <i>male</i> | Dummy-Variable: =1 wenn männlich, sonst =0 |
| <i>job</i> | Dummy-Variable: =1 wenn Student jobbt, sonst =0 |
| <i>alcohol</i> | durchschnittliche Anzahl von Tagen pro Woche an denen Person Alkohol trinkt |
| <i>trans</i> | Genutztes Verkehrsmittel auf dem Weg zur Uni (nur eins der drei möglich): 1-zu Fuß, bzw. öffentliche Verkehrsmittel, 2-mit dem Fahrrad, 3-mit dem Auto |

a) Zunächst betrachten Sie das folgende Modell:

$$skipped = \beta_0 + \beta_1 male + \beta_2 job + \beta_3 alcohol + u$$

Bei einer Schätzung mit SPSS erhalten Sie folgenden Output:

| Modell | Nicht standardisierte Koeffizienten | |
|-------------|-------------------------------------|----------------|
| | Regressions-koeffizient | Standardfehler |
| (Konstante) | 0,411 | 0,155 |
| male | 0,212 | 0,13 |
| job | 0,119 | 0,131 |
| alcohol | 0,255 | 0,05 |

a. Abhängige Variable: skipped

- i) Interpretieren Sie inhaltlich die Größe der geschätzten Parameter $\hat{\beta}_1$ und $\hat{\beta}_3$. (2 Punkte)
 - ii) Testen Sie am 1% Signifikanzniveau, ob der Koeffizient von alcohol signifikant größer als Null ist. Geben Sie an: welchen Test Sie durchführen, die Null- und Alternativhypothese, die Teststatistik, den kritischen Wert, sowie die Testentscheidung. (4 Punkte)
- b) Sie wollen prüfen, b1) ob die genutzten Verkehrsmittel unterschiedliche Einflüsse auf die Vorlesungsabwesenheit haben und b2) ob sich der Effekt der Häufigkeit des Alkoholkonsums auf die Vorlesungsabwesenheit danach unterscheidet, auf welches Verkehrsmittel man auf dem Weg zur Uni angewiesen ist.
- i) Wie sieht die Regressionsgleichung aus, mit der Sie beide Vermutungen in nur einem Modell überprüfen können? Definieren Sie gegebenenfalls neu definierte Variablen. (3 Punkte)
 - ii) Erläutern Sie kurz, wie sie b2) testen können. Geben Sie das Testverfahren und die Nullhypothese an. (2 Punkte)
- c) Anschließend schätzen Sie das folgende Modell:

$$skipped = \beta_0 + \beta_1 male + \beta_2 job + \beta_3 alcohol + \beta_4 male * alcohol + u$$

- i) Bestimmen Sie formal den marginalen Effekt der Häufigkeit des Alkoholkonsums für männliche Studierende. (1 Punkt)
- ii) Bestimmen Sie formal den marginalen Effekt der Häufigkeit des Alkoholkonsums für weibliche Studierende. (1 Punkt)
- iii) Welches Vorzeichen erwarten Sie für den relevanten Parameter, wenn sich die Häufigkeit des Alkoholkonsums auf die Vorlesungsabwesenheit stärker bei weiblichen als bei männlichen Studierenden niederschlägt? (1 Punkt)

Aufgabe 2**[4 Punkte]**

Gegeben sei das wahre Modell $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$.

- Erläutern Sie *kurz* die Auswirkungen bezüglich der Unverzerrtheit des KQ-Schätzers $\hat{\beta}_1$, wenn in einer Schätzung die relevante erklärende Variable x_2 nicht berücksichtigt wird. (2 Punkte)
- Welche Auswirkungen ergeben sich für $\hat{\beta}_1$, wenn Sie eine irrelevante Variable x_3 zusätzlich in einer Schätzung aufnehmen? (2 Punkte)

Aufgabe 3**[8,5 Punkte]**

Sie schätzen mit Daten des Sozio-ökonomischen Panels (SOEP) für das Jahr 2008 den Zusammenhang zwischen (logarithmierten) Stundenlöhnen und einer Reihe von erklärenden Größen:

| | |
|---------------------|--|
| <i>ln_hwage</i> | logarithmierter Stundenlohn |
| <i>exper</i> | (Vollzeit) Arbeitsmarkterfahrung (in Jahren) |
| <i>exper2</i> | quadrierte (vollzeit) Arbeitsmarkterfahrung (in Jahren) |
| <i>male</i> | Dummy-Variable: =1 wenn Person männlich, sonst =0 |
| <i>married</i> | Dummy-Variable: =1 wenn Person verheiratet, sonst =0 |
| <i>high_skill</i> | Dummy-Variablen: =1 wenn Person hochqualifiziert ist, sonst =0 |
| <i>medium_skill</i> | - Referenz - |
| <i>low_skill</i> | Dummy-Variablen: =1 wenn Person geringqualifiziert ist, sonst =0 |
| <i>sued</i> | Dummy-Variable: =1 wenn Person in Süddeutschland wohnt, sonst =0 |

Sie schätzen das Modell zunächst ohne den Regionen-Dummy (*sued*) und erhalten folgendes Ergebnis:

Koeffizienten^a

| Modell | Nicht standardisierte Koeffizienten | | Standardisierte Koeffizienten | T | Sig. |
|---------------|-------------------------------------|----------------|-------------------------------|---------|------|
| | Regressionskoeffizient B | Standardfehler | Beta | | |
| 1 (Konstante) | 1,8494 | 0,0163 | | 113,532 | ,000 |
| exper | ,0403 | 0,0021 | ,750 | 19,594 | ,000 |
| exper2 | -,0008 | 0,0001 | -,573 | -15,290 | ,000 |
| male | ,1326 | 0,0136 | ,107 | 9,720 | ,000 |
| married | ,0591 | 0,0138 | ,046 | 4,273 | ,000 |
| high_skill | ,3952 | 0,0137 | ,302 | 28,897 | ,000 |
| low_skill | -,2525 | 0,0211 | -,125 | -11,958 | ,000 |

a. Abhängige Variable: *ln_hwage*

- Berechnen und interpretieren Sie inhaltlich den marginalen Effekt der Arbeitsmarkterfahrung. (2 Punkte)
- Bestimmen Sie den vorhergesagten logarithmierten Stundenlohn einer hochqualifizierten, verheirateten Frau mit 20 Jahren Arbeitsmarkterfahrung. (1,5 Punkte)

Sie schätzen erneut, diesmal unter Verwendung des Regionen-Dummies:

Koeffizienten^a

| Modell | Nicht standardisierte Koeffizienten | | Standardisierte Koeffizienten | T | Sig. |
|---------------|-------------------------------------|----------------|-------------------------------|---------|------|
| | Regressionskoeffizient B | Standardfehler | Beta | | |
| 1 (Konstante) | 1,8320 | 0,0185 | | 98,900 | ,000 |
| exper | ,0404 | ? | ,751 | 19,647 | ,000 |
| exper2 | -,0008 | 0,0001 | -,575 | -15,352 | ,000 |
| male | ,1322 | 0,0136 | ,107 | 9,705 | ,000 |
| married | ? | 0,0138 | ,046 | 4,244 | ,000 |
| high_skill | ,3950 | 0,0137 | ,302 | 28,918 | ,000 |
| low_skill | -,2531 | 0,0211 | -,125 | -11,999 | ,000 |
| sued | ,0438 | 0,0139 | ,035 | ? | ,002 |

a. Abhängige Variable: ln_hwage

- c) Berechnen Sie die 3 fehlenden Werte der Tabelle. (1,5 Punkte)
- d) Sie stellen fest, dass sich die geschätzten *exper*-Koeffizienten und Standardfehler verglichen mit denen aus Teilaufgabe c) nicht ändern. Erläutern Sie *kurz*, warum dies der Fall sein könnte. (2 Punkte)
- e) Ist „*sued*“ eine irrelevante erklärende Variable? Argumentieren Sie statistisch und inhaltlich. (1,5 Punkte)

Aufgabe 4:

[23,5 Punkte]

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

| | |
|--------------------|--|
| <i>satlife</i> | Subjektive Lebenszufriedenheit mit den Ausprägungen 0 (sehr unzufrieden) bis 10 (sehr zufrieden) |
| <i>age</i> | Alter einer Person in Jahren |
| <i>sqage</i> | Alter ² /100 |
| <i>educ</i> | Anzahl der Bildungsjahre |
| <i>sqeduc</i> | Bildungsjahre ² /100 |
| <i>west</i> | Dummy-Variable: =1, wenn Person in den alten Bundesländern lebt, =0 sonst |
| <i>male</i> | Dummy-Variable: =1, wenn Person männlich, =0 sonst |
| <i>unempl</i> | Dummy-Variable: =1, wenn Person arbeitslos, =0 sonst |
| <i>child</i> | Anzahl der Kinder im Haushalt |
| <i>log(incnet)</i> | logarithmiertes Nettomonatseinkommen einer Person in € |

Sie schätzen zunächst folgendes **Modell 1**:

$$satlife_i = \beta_0 + \beta_1 age_i + \beta_2 sqage_i + \beta_3 educ_i + \beta_4 sqeduc_i + \beta_5 west_i + \beta_6 male_i + \beta_7 unempl_i + \beta_8 child_i + u_i$$

| | Modell 1 | | | | Modell 2 | | | |
|-------------|-------------|-------------|--------|-------|-------------|-------------|--------|-------|
| | Koeffizient | Std. fehler | T | Sig. | Koeffizient | Std. fehler | T | Sig. |
| age | -,0327 | ,0018 | -18,03 | 0,000 | -,1153 | ,0040 | -28,69 | 0,000 |
| sqage | ,0282 | ,0018 | 15,71 | 0,000 | ,1282 | ,0047 | 27,18 | 0,000 |
| educ | ,2582 | ,0183 | 14,07 | 0,000 | ,1086 | ,0251 | 4,33 | 0,000 |
| sqeduc | -,6123 | ,0679 | -9,01 | 0,000 | -,1838 | ,0909 | -2,02 | 0,043 |
| west | ,6100 | ,0128 | 47,39 | 0,000 | ,5087 | ,0166 | 30,61 | 0,000 |
| male | -,0297 | ,0108 | -2,75 | 0,006 | -,1625 | ,0154 | -10,53 | 0,000 |
| unempl | -1,2072 | ,0213 | -56,46 | 0,000 | -,7218 | ,0622 | -11,59 | 0,000 |
| child | ,0496 | ,0065 | 7,56 | 0,000 | ,0905 | ,0078 | 11,55 | 0,000 |
| log(incnet) | --- | --- | --- | --- | ,2477 | ,0112 | 22,04 | 0,000 |
| Konstante | 5,1702 | ,1279 | 40,41 | 0,000 | 6,3024 | ,1819 | 34,63 | 0,000 |
| SSE | 27.646 | | | | 8.788 | | | |
| SSR | 306.288 | | | | 137.742 | | | |
| SST | 333.934 | | | | 146.530 | | | |
| N | 102.575 | | | | 54.615 | | | |

Abhängige Variable: satlife

- Interpretieren Sie den geschätzten Parameter $\hat{\beta}_7$ statistisch unter Verwendung des p-Wertes und inhaltlich. Beurteilen Sie die Stärke des Effektes. (3 Punkte)
- Unterstellen Sie, dass die Variable *sqage* nicht durch 100 geteilt worden wäre. Welche Werte erwarten Sie nun für die Schätzwerte des Regressionskoeffizienten, des Standardfehlers, des t-Wertes und des Signifikanzniveaus des Parameters β_2 ? (3 Punkte)
- Bei welchem Alter ist die durchschnittliche Zufriedenheit minimal? (2,5 Punkte)
- Testen Sie die gemeinsame Signifikanz von *educ* und *sqeduc* auf dem 1%-Niveau. ($SSR_r = 306.396$). Geben Sie Testverfahren, Teststatistik, Nullhypothese, kritischen Wert und Ihre Testentscheidung an. (4,5 Punkte)

Sie schätzen nun folgendes **Modell 2**:

$$satlife_i = \beta_0 + \beta_1 age_i + \beta_2 sqage_i + \beta_3 educ_i + \beta_4 sqeduc_i + \beta_5 west_i + \beta_6 male_i + \beta_7 unempl_i + \beta_8 child_i + \beta_9 \log(incnet) + u_i$$

- Interpretieren Sie auf Basis der Ergebnisse in obiger Tabelle $\hat{\beta}_9$ inhaltlich. (1 Punkt)
- Berechnen Sie, wie sich das Einkommen einer Person ohne Kind ändern müsste, damit das gleiche Zufriedenheitsniveau erreicht wird wie bei einer Person mit einem Kind. (4 Punkte)
- Berechnen Sie das R^2 für Modell 2. Interpretieren Sie den von Ihnen berechneten Wert für R^2 . (3 Punkte)
- Diskutieren Sie unter welchen Umständen die Gauss-Markov-Annahme MLR.4 im Modell 2 verletzt ist. (2,5 Punkte)

Aufgabe 5:

[24 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.

| | |
|--|---|
| | Die F-Verteilung ist eine symmetrische Verteilungsfunktion. |
| | $R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$. |
| | Eine Annahme des einfachen linearen Modells ist $E(u x) = 1$. |
| | Große p-Werte sind Evidenz gegen H_0 . |
| | Der geschätzte Koeffizient $\hat{\beta}$ lässt sich aus dem Produkt aus t-Wert und Standardfehler von $\hat{\beta}$ berechnen. |
| | Das angepasste R^2 kann nicht negativ werden. |
| | Im einfachen linearen Regressionsmodell liegt der Punkt (\bar{x}, \bar{y}) auf der Regressionsgeraden. |
| | Wenn es gelingt, alle relevanten unabhängigen Variablen konstant zu halten, lässt sich der ceteris paribus Effekt als kausale Wirkung interpretieren. |
| | Der Parameter β_1 im Modell $\log(y) = \beta_0 + \beta_1 \log(x_1) + u$ wird als Semielastizität interpretiert. |
| | $Var(X - 2Y) = Var(X) + 4Var(Y) - 4Cov(X, Y)$. |
| | Multikollinearität führt zu verzerrten Schätzern. |
| | Das Gauss-Markov-Theorem besagt, dass unter bestimmten Bedingungen das KQ-Verfahren das effizienteste der linearen unverzerrten Schätzverfahren ist. |
| | Für die Unverzerrtheit des KQ-Schätzers spielt $E\{u x\} = 0$ keine Rolle. |
| | Man kann Dummy-Variablen aus stetigen Variablen bilden. |
| | Im multiplen linearen Modell ist das R^2 das geeignete Maß, um über die Aufnahme zusätzlicher erklärender Größen zu entscheiden. |
| | Eine standardisierte stetige Variable hat den Erwartungswert von 0 und eine Standardabweichung von 1. |
| | Bei unverzerrten Schätzern ist jeder Schätzwert mit dem wahren Wert identisch. |
| | Ein Schätzer ist unverzerrt, wenn mit größer werdender Stichprobe die Varianz sinkt. |
| | Der F-Wert ist nie negativ. |
| | Für große Stichproben gelten t-Tests asymptotisch auch ohne normalverteilte Störterme. |
| | Die Varianz einer Konstanten ist gleich dem Wert dieser Konstante. |
| | Durch Verwendung von Polynomen der erklärenden Variablen können in einer Regression nicht-lineare Effekte dieser Variablen abgebildet werden. |
| | Das lineare Regressionsmodell heißt linear, weil alle erklärenden Größen nur in erster Potenz vorkommen dürfen. |
| | Der Lagrange-Multiplier Test verwendet im 2. Schritt die Residuen des restringierten Modells als abhängige Variable. |
| | Von zwei unverzerrten Schätzern $\hat{\beta}_1$ und $\tilde{\beta}_1$ für einen unbekanntem Bevölkerungsparameter β_1 ist $\tilde{\beta}_1$ effizienter, wenn $Var(\tilde{\beta}_1) < Var(\hat{\beta}_1)$. |
| | Im einfachen linearen Modell $y_i = \beta_0 + \beta_1 x_i + u_i$ bezeichnet β_1 den Achsenabschnittsparameter. |
| | Benutzt man 15 Bundesland-Dummies (einer pro Land), um für Unterschiede zwischen den 16 Bundesländern zu kontrollieren, so kann über das ausgelassene Bundesland keine Aussage getroffen werden. |
| | Over-controlling führt zu hohen Standardfehlern und zu verzerrten Parametern. |
| | Selbst wenn es im einfachen linearen Modell keine Variation in der erklärenden Variable gibt, kann der Effekt der Variable mit KQ geschätzt werden. |
| | Genestete Modelle können mittels R^2 verglichen werden. |
| | Eine Modellspezifikation mit Polynomen 3. Ordnung ist flexibler als eine quadratische Modellspezifikation. |
| | Schätzt man im multiplen linearen Modell ohne Achsenabschnitt, dann geht man davon aus, dass $\beta_0 = 1$. |
| | Die Störtermvarianz ändert sich bei Umskalierung der abhängigen Variable. |

| | |
|--|---|
| | Bei einer normalverteilten Zufallsvariablen ist in der Grundgesamtheit der Medianwert gleich dem arithmetischen Mittel. |
| | Das multiple Regressionsmodell erlaubt die Berechnung von Partialeffekten. |
| | Durch die Aufnahme mehrerer Kontrollvariablen im multiplen Regressionsmodell können kausale Interpretationen plausibler werden. |
| | Folgt $\log(Y)$ einer Normalverteilung, dann ist Y lognormal verteilt. |
| | Hat der Fehlerterm u für alle Beobachtungen die gleiche Varianz, so bezeichnet man ihn als homoskedastisch. |
| | Auch wenn gilt $E\{u x\} \neq 0$ kann im Rahmen einer linearen Schätzung durch eine binäre Variable der Effekt einer Politikmaßnahme evaluiert werden. |
| | Bei der „dummy variable trap“ sind die Regressionskoeffizienten eindeutig bestimmbar aber nicht effizient |
| | Mehrere Dummy-Variablen in einer Schätzung führen zu „dummy variable trap“. |
| | Werden irrelevante Größen in ein Schätzmodell aufgenommen so führt dies zu „omitted variable bias“. |
| | Wird im Modell $y = \beta_0 + \delta_0 \text{Geschlecht}$ die Variable <i>Geschlecht</i> mit den Werten 1-männlich, 2-weiblich kodiert, so ist δ_0 verzerrt. |
| | Die Zählerfreiheitsgrade beim F-Test geben die Anzahl der getesteten Restriktionen an. |
| | Im multiplen linearen Regressionsmodell mit k erklärenden Größen ergeben sich bei der Herleitung des KQ-Schätzers genau k Bedingungen 1. Ordnung. |
| | Im Modell $y = \beta_0 + \beta_1 \log(x_1) + u$ gibt β_1 an, um wie viel Prozent sich die abhängige Variable ändert, wenn x_1 um eine Einheit steigt. |
| | Die Varianz des Fehlerterms kann im einfachen linearen Modell mittels $\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^n \hat{u}_i^2$ unverzerrt geschätzt werden. |
| | Die SSR im restringierten Modell ist immer mindestens genauso groß ist wie die SSR im unrestringierten Modell. |

Aufgabe 6:

[16 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. | Wenn der p-Wert für den Signifikanztest eines Parameters den Wert 0,06 annimmt, dann |
| a | <input type="checkbox"/> wird H_0 am 5% Niveau verworfen. |
| b | <input type="checkbox"/> wird H_0 am 10% Niveau verworfen. |
| c | <input type="checkbox"/> beträgt der Typ I Fehler bei Verwerfen von H_0 6%. |
| d | <input type="checkbox"/> beträgt der Typ II Fehler bei Verwerfen von H_0 6%. |
| e | <input type="checkbox"/> a und c |
| f | <input type="checkbox"/> b und c |

| | |
|----|---|
| 2. | Beim Testen gegen die Alternative $H_1: \beta_j < 0$, wird H_0 verworfen, wenn |
| a | <input type="checkbox"/> $ t < -c$ (c =kritischer Wert) |
| b | <input type="checkbox"/> $ t < c$ |
| c | <input type="checkbox"/> $t < -c$ |
| d | <input type="checkbox"/> $t > -c$ |
| e | <input type="checkbox"/> b und c |
| f | <input type="checkbox"/> b und d |

| | | |
|----|--|---|
| 3. | Ein 99% Konfidenzintervall für einen Parameter β_j | |
| a | <input type="checkbox"/> | kann nicht mit dem Konfidenzintervall für einen Parameter β_k überlappen. |
| b | <input type="checkbox"/> | schließt bei signifikantem Parameterschätzer den Wert Null ein. |
| c | <input type="checkbox"/> | ist abhängig von der Stichprobengröße. |
| d | <input type="checkbox"/> | enthält mit Wahrscheinlichkeit 99% den wahren Wert. |
| e | <input type="checkbox"/> | a, b und c |
| f | <input type="checkbox"/> | b und d |

| | | |
|----|--|----------------------------|
| 4. | In einer quadratischen Spezifikation $y = \beta_1 x + \beta_2 x^2 + u$ ergibt sich ein umgekehrt u-förmiger Verlauf der Regressionsgeraden, wenn | |
| a | <input type="checkbox"/> | $\beta_1 < 0, \beta_2 < 0$ |
| b | <input type="checkbox"/> | $\beta_1 > 0, \beta_2 < 0$ |
| c | <input type="checkbox"/> | $\beta_1 > 0, \beta_2 > 0$ |
| d | <input type="checkbox"/> | $\beta_1 = 0, \beta_2 < 0$ |
| e | <input type="checkbox"/> | $\beta_1 < 0, \beta_2 > 0$ |
| f | <input type="checkbox"/> | keine der Antworten |

| | | |
|----|---|--|
| 5. | Welche Aussage beschreibt keine Eigenschaft des KQ-Schätzers? | |
| a | <input type="checkbox"/> | Jeder vorhergesagte Wert \hat{y} liegt auf der Regressionsgeraden. |
| b | <input type="checkbox"/> | Die Stichprobenkovarianz zwischen x_i und u_i ist 1. |
| c | <input type="checkbox"/> | y_i lässt sich über \hat{y}_i und \hat{u}_i berechnen. |
| d | <input type="checkbox"/> | $\sum_{i=1}^n \hat{u}_i^2 = 0$. |
| e | <input type="checkbox"/> | a und c |
| f | <input type="checkbox"/> | b und d |

| | | |
|----|---|--|
| 6. | Wenn die Kovarianz von zwei Zufallsvariablen Null ist (d.h. $Cov(X, Y) = 0$), dann | |
| a | <input type="checkbox"/> | sind die zwei Variablen unabhängig. |
| b | <input type="checkbox"/> | sind die zwei Variablen unkorreliert. |
| c | <input type="checkbox"/> | besteht kein Zusammenhang zwischen den Variablen. |
| d | <input type="checkbox"/> | kann Skalieren der Variablen die Kovarianz ändern. |
| e | <input type="checkbox"/> | a, b und c |
| f | <input type="checkbox"/> | b, c und d |

| | | |
|----|--|--|
| 7. | Folgende Annahmen müssen zutreffen, damit mit dem KQ-Verfahren im multiplen Modell ohne Konstante mit k erklärenden Größen unverzerrt geschätzt werden kann: | |
| a | <input type="checkbox"/> | Eine betrachtete Zufallsstichprobe folgt dem Bevölkerungsmodell. |
| b | <input type="checkbox"/> | Keine unabhängige Variable ist konstant. |
| c | <input type="checkbox"/> | $E(u x_1 \dots x_k) = 0$. |
| d | <input type="checkbox"/> | Der Fehlerterm u ist homoskedastisch. |
| e | <input type="checkbox"/> | a, b und c |
| f | <input type="checkbox"/> | alle Antworten |

| | | |
|----|--|---|
| 8. | Im einfachen linearen Modell $y_i = \beta_0 + \beta_1 x_{1i} + u_i$ erhalten wir den KQ-Schätzer $\hat{\beta}_1$ durch die Minimierung folgender Zielfunktion S: | |
| a | <input type="checkbox"/> | $\sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i})^2$ |
| b | <input type="checkbox"/> | $\sum_{i=1}^n (y_i - \bar{y})^2$ |
| c | <input type="checkbox"/> | $\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ |
| d | <input type="checkbox"/> | $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ |
| e | <input type="checkbox"/> | a und b |
| f | <input type="checkbox"/> | a und d |

| | | |
|----|--------------------------|--|
| 9. | Ein Chow-Test | |
| a | <input type="checkbox"/> | kann nur durch Verwendung von Interaktionstermen durchgeführt werden. |
| b | <input type="checkbox"/> | basiert auf einer Chi-Quadrat verteilten Teststatistik. |
| c | <input type="checkbox"/> | prüft, ob Regressionskoeffizienten nach Gruppen unterschiedlich ausfallen. |
| d | <input type="checkbox"/> | kann durch Schätzung getrennter Regressionen für zwei Gruppen durchgeführt werden. |
| e | <input type="checkbox"/> | b und c |
| f | <input type="checkbox"/> | c und d |

| | | |
|-----|---|---|
| 10. | Bei einer logarithmierten abhängigen Variable wird der marginale Effekt einer erklärenden Dummy-Variable: | |
| a | <input type="checkbox"/> | als $(\exp(\beta) - 1) \cdot 100$ berechnet. |
| b | <input type="checkbox"/> | approximativ als $100 \cdot \beta$ berechnet. |
| c | <input type="checkbox"/> | als Prozentgröße interpretiert. |
| d | <input type="checkbox"/> | relativ zur Referenzgruppe interpretiert. |
| e | <input type="checkbox"/> | a, c und d |
| f | <input type="checkbox"/> | alle Antworten |

| | | |
|-----|---|---|
| 11. | Ein Schätzer eines unbekanntem Parameters | |
| a | <input type="checkbox"/> | ist unverzerrt, wenn sein Erwartungswert mit dem wahren Wert identisch ist. |
| b | <input type="checkbox"/> | ist verzerrt wenn sein Wert von Stichprobe zu Stichprobe variiert. |
| c | <input type="checkbox"/> | ist effizient nur wenn er gleichzeitig symmetrisch ist. |
| d | <input type="checkbox"/> | ist eine Zufallsvariable. |
| e | <input type="checkbox"/> | a und d |
| f | <input type="checkbox"/> | a, b und d |

| | | |
|-----|---|--|
| 12. | Im Modell $y = \beta_0 + \beta_1 x_1 + u$ kommt es aufgrund einer ausgelassenen Variablen x_2 zu einer positiver Verzerrung von $\hat{\beta}_1$, wenn: (Hilfsregression: $x_1 = \delta_0 + \delta_1 x_2 + \varepsilon$) | |
| a | <input type="checkbox"/> | $\beta_2 = 0$ und $\tilde{\delta}_1 > 0$ |
| b | <input type="checkbox"/> | $\beta_2 < 0$ und $\tilde{\delta}_1 < 0$ |
| c | <input type="checkbox"/> | $\beta_2 > 0$ und $\tilde{\delta}_1 = 0$ |
| d | <input type="checkbox"/> | $\beta_2 = 0$ und $\tilde{\delta}_1 = 0$ |
| e | <input type="checkbox"/> | a und c |
| f | <input type="checkbox"/> | keine der Antworten |

| | | |
|-----|---|---|
| 13. | In einer geschätzten Lohnregression $\log(wage) = \hat{\beta}_0 + \hat{\beta}_1 educ + \hat{\beta}_2 exper + \hat{\beta}_3 exper^2$ | |
| a | <input type="checkbox"/> | ist $\hat{\beta}_1$ als Semi-Elastizität interpretierbar. |
| b | <input type="checkbox"/> | würde eine Steigerung in educ um eine Einheit eine approximative Änderung in wage in Höhe von $100 \cdot \hat{\beta}_1$ Prozentpunkten verursachen. |
| c | <input type="checkbox"/> | wird der marginale Effekt von exper als partielle Ableitung von $\log(wage)$ nach exper berechnet. |
| d | <input type="checkbox"/> | ist die Lohnwirkung einer zusätzlichen Einheit exper sowohl von educ als auch vom Ausgangsniveau in exper abhängig. |
| e | <input type="checkbox"/> | b, c und d |
| f | <input type="checkbox"/> | a und c |

| | | |
|-----|---------------------------------------|---|
| 14. | Von Heteroskedastie spricht man, wenn | |
| a | <input type="checkbox"/> | irrelevante erklärende Variablen in einem Regressionsmodell aufgenommen werden. |
| b | <input type="checkbox"/> | eine erklärende Variable x_j mit dem Fehlerterm u korreliert ist. |
| c | <input type="checkbox"/> | Parameter in nicht-linearer Weise in das Regressionsmodell aufgenommen werden. |
| d | <input type="checkbox"/> | die Varianz des Fehlerterms von der Ausprägung der erklärenden Größen abhängt. |
| e | <input type="checkbox"/> | a und d |
| f | <input type="checkbox"/> | b und d |

| | | |
|-----|---|---|
| 15. | Wird im Modell $y = \beta_0 + \beta_1 x_1 + u$ die unabhängige Variable mit 100 multipliziert, so | |
| a | <input type="checkbox"/> | wird β_1 durch 100 geteilt. |
| b | <input type="checkbox"/> | werden β_0 und β_1 durch 100 geteilt. |
| c | <input type="checkbox"/> | muss y durch 100 geteilt werden. |
| d | <input type="checkbox"/> | muss y mit 100 multipliziert werden. |
| e | <input type="checkbox"/> | a und d |
| f | <input type="checkbox"/> | b und c |

| | | |
|-----|------------------------------|---|
| 16. | Der Lagrange Multiplier Test | |
| a | <input type="checkbox"/> | verwendet eine χ^2 -verteilte Teststatistik. |
| b | <input type="checkbox"/> | wird u.a. zum Testen mehrerer linearer Restriktionen verwendet. |
| c | <input type="checkbox"/> | ist mit dem Chow-Test identisch wenn $q=2$. |
| d | <input type="checkbox"/> | schätzt zwei Regressionsmodelle. |
| e | <input type="checkbox"/> | a, b und d |
| f | <input type="checkbox"/> | a und b |