Ökonometrieprüfung SS 2015 - MUSTERLÖSUNG

Fach: Ökonometrie

Prüfer: Prof. Regina T. Riphahn, Ph.D.

Vorbemerkungen:

Bitte beachten Sie:	**Es wird nur der Lösungsbogen eingesammelt und bewertet. Angaben auf dem Aufgabenbogen werden nicht gewertet.**
Anzahl der Aufgaben:	Die Klausur besteht aus 5 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 gesuchten 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: **[15 Punkte]**

1.1 Leiten Sie ausgehend von der Formel $b = (X'X)^{-1}X'y$ für das Modell $y = X\beta + \varepsilon$ die Unverzerrtheit des KQ-Schätzers in Matrizenschreibweise her. Machen Sie kenntlich, an welcher Stelle der Herleitung Sie welche Annahme verwenden. Gehen Sie von einer stochastischen X Matrix aus. [4 Punkte]

i. $E[b] = E[(X'X)^{-1}X'y]$ [Aufgabenstellung]
ii. $E[b] = E[(X'X)^{-1}X'X\beta + (X'X)^{-1}X'\varepsilon]$ [0,5P]
iii. $E[b] = E[\beta + (X'X)^{-1}X'\varepsilon]$ [0,5P]
iv. $E[b] = \beta + (X'X)^{-1}X'E[\varepsilon]$ [0,5P]
v. $E[b] = \beta$ [0,5P]

- Für Schritt iv) wird A2 (Unabhängigkeit von X und ε) benötigt. [1P]
- Für Schritt v) wird A1 ($E[\varepsilon] = 0$) benötigt. [1P]

1.2 Was ändert sich in Ihrer Antwort zu Teilaufgabe 1.1) bei der Herleitung, wenn die Matrix X deterministisch ist? [1 Punkt]

- Bei deterministischen X wird A2 nicht benötigt: $E[b] = \beta + (X'X)^{-1}X'E[\varepsilon] = \beta$ [1P]
- Alternativ: Es wird nur A1 ($E[\varepsilon] = 0$) benötigt. [1P]

1.3 Leiten Sie die Varianz ($Var[b]$) in Matrizenschreibweise her. Gehen Sie von einer stochastischen $(X'X)$ Matrix aus. Welche Annahmen benötigen Sie, um zu zeigen, dass $Var[b] = \sigma^2(X'X)^{-1}$? (10 Punkte)

Herleitung der Varianz (6 Punkte)

i. $Var[b] = Var[(X'X)^{-1}X'y]$ [Aufgabenstellung]
ii. $= Var[\beta + (X'X)^{-1}X'\varepsilon]$ [1P]
iii. $= Var[(X'X)^{-1}X'\varepsilon]$ [1P]
iv. $= (X'X)^{-1}X'Var[\varepsilon]X(X'X)^{-1}$ [1P]
v. $= (X'X)^{-1}X'(\sigma^2 I_N)X(X'X)^{-1}$ [1P]
vi. $= \sigma^2(X'X)^{-1}X'X(X'X)^{-1}$ [1P]
vii. $= \sigma^2(X'X)^{-1}$ [1P]

- Man benötigt die A 2, 3 und 4. Im Einzelnen:
  - Schritt (iv): Um die $(X'X)^{-1}X'$ aus der Varianz ziehen zu können, muss die A 2 (X und ε unabhängig) gelten. [2P]
  - Schritt (v): Um $Var[\varepsilon]$ durch $\sigma^2 I_N$ ersetzen zu können, müssen die A 3 (konstantes σ^2) [1P] sowie die A 4 (die Cov der Störterme ist gleich Null) gelten. [1P]

alternativ:

i. $Var[b] = E[(b-\beta)(b-\beta)']$ [Formelsammlung]
ii. $= E[((X'X)^{-1}X'y - \beta)((X'X)^{-1}X'y - \beta)']$ [Aufgabenstellung]
iii. $= E[((X'X)^{-1}X'\overbrace{(X\beta+\varepsilon)}^{y} - \beta)((X'X)^{-1}X'\overbrace{(X\beta+\varepsilon)}^{y} - \beta)']$ [Aufgabenstellung]

iv. $= E[(\beta + (X'X)^{-1}X'\varepsilon - \beta)(\beta + (X'X)^{-1}X'\varepsilon - \beta)']$ [1P]

v. $= E[(X'X)^{-1}X'\varepsilon\varepsilon'X(X'X)^{-1}]$ [1P]

vi. $= (X'X)^{-1}X'E[\varepsilon\varepsilon']X(X'X)^{-1}$ [1P]

vii. $= (X'X)^{-1}X'(\sigma^2 I_N)X(X'X)^{-1}$ [1P]

viii. $= \sigma^2(X'X)^{-1}X'X(X'X)^{-1}$ [1P]

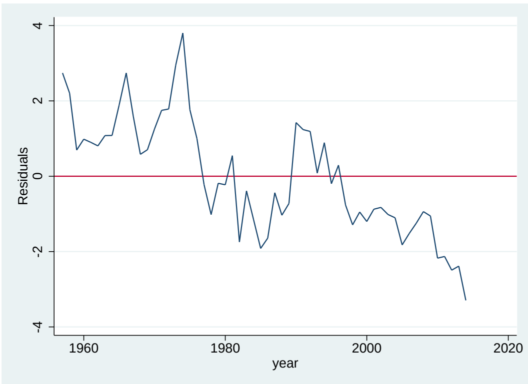
ix. $= \sigma^2(X'X)^{-1}$ [1P]

- Man benötigt die A 2, 3 und 4. Im Einzelnen:
  - Schritt (vi): Um die $(X'X)^{-1}X'$ aus dem Erwartungswertoperator ziehen zu können, muss die A 2 (X und ε unabhängig) gelten. [2P]
  - Schritt (vii): Um $E[\varepsilon\varepsilon']$ durch $\sigma^2 I_N$ ersetzen zu können, müssen die A 3 (konstantes σ^2) (1P) sowie die A 4 (die Cov der Störterme ist gleich Null) gelten. [2P]

Aufgabe 2: **[15 Punkte]**

Sie untersuchen den Zusammenhang zwischen Erträgen aus US-Staatsanleihen (usbond) und Erträgen aus Deutschen Staatsanleihen (gerbond) mit Zeitreihendaten auf Jahresbasis von 1957-2014 (T=58 Beobachtungen).

2.1 Die folgende Abbildung zeigt den Plot der Residuen einer einfachen linearen Regression mit *usbond* als abhängiger und *gerbond* als unabhängiger Variable. Interpretieren Sie das Muster der Residuen mit Hinblick auf Autokorrelation und begründen Sie Ihre Antwort. [2 Punkte]



- Zeitlich aufeinander folgende Residuen e_t und e_{t-1} nehmen häufig gemeinsam negative oder positive Werte an, sodass $Cov(\varepsilon_t, \varepsilon_{t-1}) = 0$ verletzt sein könnte. [1P]
- Der Plot deutet auf positive Autokorrelation hin. [1P]

2.2 Führen Sie unter Verwendung des folgenden Stata-Outputs einen Breusch-Godfrey-Test auf Autokorrelation 1. Ordnung am 5% Signifikanzniveau durch. Geben Sie Null- und Alternativhypothese, Teststatistik, Freiheitsgrade, Entscheidungsregel, Schlusslogik und Testergebnis an. Hinweis: Die Variable L1.e im Stata Output bezeichnet das verzögerte Residuum. [5 Punkte]

- Hypothesen: H0: $\rho = 0$ gegen H1: $\rho \neq 0$ [1P] für $\varepsilon_t = \alpha + \rho\varepsilon_{t-1} + \nu_t$.
- 1 Freiheitsgrad wird berücksichtigt . [1P]

- Entscheidungsregel: H0 verwerfen, falls $LM > \chi^2_{1,\alpha=0.05} = 3,84$. [1P]
- Teststatistik: LM=57*0,714=40,69. [1P]
- Entscheidung: Da $LM = 40,69 > 3,84$, wird die Nullhypothese verworfen [1P]. Es liegt somit Evidenz für Autokorrelation vor.

```
      Source |       SS       df       MS              Number of obs =      57
-------------+------------------------------           F(  2,    54) =   67.42
       Model |  92.1668831     2  46.0834416           Prob > F      =  0.0000
    Residual |  36.9113371    54   .68354328           R-squared     =  0.7140
-------------+------------------------------           Adj R-squared =  0.7034
       Total |   129.07822    56  2.30496822           Root MSE      =  .82677

------------------------------------------------------------------------------
           e |      Coef.   Std. Err.      t    P>|t|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
        L1.e |   .8572559   .0738818    11.60   0.000     .7091317     1.00538
     gerbond |  -.0100386   .0407582    -0.25   0.806    -.0917539    .0716767
       _cons |  -.0350127    .276695    -0.13   0.900    -.5897527    .5197274
------------------------------------------------------------------------------
```

2.3 Nennen Sie zwei inhaltliche Gründe, warum in diesem Fall Autokorrelation vorliegen könnte. Geben Sie für jeden Grund jeweils ein Beispiel an. [2 Punkte]

Jeweils ein halber Punkt pro Beispiel und Grund:

- Auslassen relevanter Variablen (z.B. globales Wirtschaftswachstum), oder
- Dynamische Fehlspezifikation (z. B. verzögerte Anleihen), oder
- Fehlspezifikation der funktionalen Form (z.B. ein nicht-linearer Zusammenhang).

2.4 Sie schätzen nun das Modell $usbond_t = \alpha + \beta gerbond_t + \varepsilon_t$ mit dem Cochrane-Orcutt Verfahren und führen danach einen Durbin-Watson-Test auf positive Autokorrelation am 5% Signifikanzniveau durch. Geben Sie Null- und Alternativhypothese, Teststatistik, und eine knappe Interpretation des Testergebnisses an. Hinweis: $\sum_{t=2}^{T}(e_t - e_{t-1})^2 = 59,544$. (6 Punkte)

```
Cochrane-Orcutt AR(1) regression -- iterated estimates

      Source |       SS       df       MS              Number of obs =      57
-------------+------------------------------           F(  1,    55) =    9.74
       Model |  5.76574992     1  5.76574992           Prob > F      =  0.0029
    Residual |  32.5590357    55  .591982467           R-squared     =  0.1504
-------------+------------------------------           Adj R-squared =  0.1350
       Total |  38.3247856    56  .684371171           Root MSE      =   .7694

------------------------------------------------------------------------------
      usbond |      Coef.   Std. Err.      t    P>|t|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
     gerbond |   .2719418    .087137     3.12   0.003     .0973154    .4465682
       _cons |    3.00989   1.589767     1.89   0.064    -.1760739    6.195853
-------------+----------------------------------------------------------------
         rho |   .9319879
------------------------------------------------------------------------------
```

Runden Sie in Ihren Antworten alle Zahlenangaben auf die dritte Nachkommastelle.

- Hypothese: H0: $\rho \leq 0$ gegen H1: $\rho > 0$ [1P]
- Teststatistik: dw=$\frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2} = \frac{59,544}{32,559} = 1,828$ [2P]

- Die kritischen Werte für K=2 Parameter und T=57 Beobachtungen lauten am 5%-Signifikanzniveau: dL = 1.53 und dU = 1.60. [1P]
- Entscheidung: Da dw = 1,828 oberhalb von dU liegt, wird die Nullhypothese nicht verworfen [1P]. Es liegt somit keine Evidenz für positive Autokorrelation in diesem Modell vor. [1P]

Aufgabe 3: **[15 Punkte]**

Sie interessieren sich für die Determinanten des Geburtsgewicht von Neugeborenen. Ihnen steht ein Querschnittsdatensatz für 1387 Neugeborene von Frauen in den USA zur Verfügung, der folgende Informationen enthält:

ln_gew_i	Logarithmiertes Gewicht des Neugeborenen *i* bei der Geburt in Unzen
$ln\_fameink_i$	Logarithmiertes Einkommen der Familie in 1000 US-Dollar
$bildmu_i$	Schulbildung der Mutter des Neugeborenen *i* in Jahren
zig_i	Anzahl der während der Schwangerschaft von der Mutter des Neugeborenen *i* pro Tag gerauchten Zigaretten
$junge_i$	Geschlecht des Neugeborenen *i* (=1, wenn männlich; =0, wenn weiblich)

Sie stellen folgendes lineares Regressionsmodell auf und schätzen dieses anschließend mit Stata:

$$ln\_gew_i = \beta_1 + \beta_2 ln\_fameink_i + \beta_3 bildmu_i + \beta_4 zig_i + \varepsilon_i$$

```
      Source |       SS       df       MS              Number of obs =    1387
-------------+------------------------------           F(  3,  1383) =   12.19
       Model |  1.29867894     3  .432892979           Prob > F      =  0.0000
    Residual |   49.114045  1383  .035512686           R-squared     =  0.0258
-------------+------------------------------           Adj R-squared =  0.0236
       Total |  50.4127239  1386  .036372817           Root MSE      =  .18845

------------------------------------------------------------------------------
      ln_gew |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  ln_fameink |   .0157281    .006081     2.59   0.011     .0037991    .0276571
      bildmu |   .0004566   .0023717     0.19   0.847    -.004196     .0051092
         zig |  -.0040596   .0008704    -4.66   0.000    -.0057671   -.0023521
       _cons |   4.714349   .0298628   157.87   0.000     4.655768     4.77293
------------------------------------------------------------------------------
```

Runden Sie in Ihren Antworten alle Zahlenangaben auf die vierte Nachkommastelle.

3.1 Interpretieren Sie inhaltlich und statistisch den Zusammenhang zwischen Geburtsgewicht und Familieneinkommen. [2 Punkte]

- Inhaltlich: b_2=0,0157: Wenn das Familieneinkommen um 1 Prozent steigt, steigt das Geburtsgewicht c.p. im Mittel um 0,0157 Prozent. [1P]
- Statistisch: $p = 0{,}011 < 0{,}05$: Der Koeffizient ist statistisch signifikant auf dem 5%-Niveau. [1P]

3.2 Erläutern Sie kurz, welche Auswirkungen sich auf die Koeffizienten der Schätzung ergeben würden, wenn das Familieneinkommen nicht in 1000 US-Dollar, sondern in 1000 Euro gemessen werden würde. [2 Punkte]

- Die Umskalierung hätte keinen Effekt auf die Steigungsparameter. [1P]
- Die Konstante würde sich jedoch ändern. [1P]

3.3 Ermitteln Sie die Vorhersage des Geburtsgewichts in Unzen für einen Neugeborenen, dessen Mutter 10 Jahre zur Schule gegangen ist, während der Schwangerschaft pro Tag 10 Zigaretten geraucht und ein Familieneinkommen von 20.000 US-Dollar hat. [5 Punkte]

- Vorhersagewert für ln_gew_i:

$$\widehat{ln\_gew_i} = b_1 + b_2 * ln(20) + b_3 * 10 + b_4 * 10 = 4,7143 + 0,0157 * 2,9957 + 0,0005 * 10 - 0,0041 * 10 = 4,7254 \text{ [2P]}$$

- Vorhersagewert für gew_i:

$$\widehat{gew_i} = exp(\widehat{ln\_gew_i} + 0,5 * \hat{\sigma}^2) = exp(4,7254 + 0,5 * 0,1885^2) = 114,7971 \text{ [2P]}$$

- Das vorhergesagte Geburtsgewicht für einen Neugeborenen, dessen Mutter 10 Jahre zur Schule gegangen ist, während der Schwangerschaft pro Tag 10 Zigaretten geraucht und ein Familieneinkommen von 20.000 US-Dollar hat beträgt 114,7971 Unzen. [1P]

3.4 Sie vermuten, dass sich der Effekt der Bildung der Mutter und der Effekt des Rauchens während der Schwangerschaft auf das Geburtsgewicht für weibliche und männliche Neugeborene unterscheidet und nehmen die entsprechenden Interaktionsterme in ihr Modell auf. Sie führen anschließend einen F-Test auf gemeinsame Signifikanz der neu gebildeten Variablen auf dem 5%-Signifikanzniveau durch und erhalten folgenden Wert der Teststatistik: $F = 1,68$.

Stellen Sie die neu gebildeten Interaktionsterme und das erweiterte Modell dar. Geben Sie anschließend die Null- und Alternativhypothese, den kritischen Wert und das Testergebnis für den durchgeführten F-Test an. [6 Punkte]

- Interaktionsterme: jungebildmu=$junge * bildmu$ und jungezig=$junge * zig$ [1P]
- Modell:

$$ln\_gew_i = \beta_1 + \beta_2 ln\_fameink_i + \beta_3 bildmu_i + \beta_4 zig_i + \beta_5 junge_i + \beta_6 jungebildmu_i + \beta_7 jungezig_i + \varepsilon_i \text{ [2P]}$$

- Hypothesen: $H_0 : b_6 = b_7 = 0$; H_1 : Mind. einer der Koeffizienten ungleich Null. [1P]
- Kritischer Wert: $F_{2,1380;5\%} = 3,00$. [1P]
- Testergebnis: Da $F_{2,1380} = 1,68 < 3,00 = F_{kritisch}$ kann die Nullhypothese auf dem 5% Signifikanzniveau nicht verworfen werden. [1P] Die Koeffizienten sind gemeinsam nicht signifikant.

Aufgabe 4 [15 Punkte]

Sie interessieren sich nun statt für das Geburtsgewicht von Neugeborenen für die Determinanten der Wahrscheinlichkeit, dass eine Mutter während der Schwangerschaft raucht. Sie stellen dafür das folgende lineare Wahrscheinlichkeitsmodell auf:

$$P(rauchen_i = 1|\mathbf{x}_i) = \beta_1 + \beta_2 bildmu_i + \beta_3 bildva_i + \beta_4 gebfolge_i$$

Die folgende Tabelle beschreibt die verwendeten Variablen:

Variable	Mittelwert	Std.abw.	Min	Max	Beschreibung
rauchen	0.1352	0.3421	0	1	=1, falls Mutter des Neugeborenen i während der Schwangerschaft geraucht hat, sonst=0
bildmu	13.1251	2.4174	2	18	=Schulbildung der Mutter des Neugeborenen i in Jahren
bildva	13.1914	2.7413	1	18	=Schulbildung des Vaters des Neugeborenen i in Jahren
gebfolge	1.6138	0.8746	1	6	=Rang des Neugeborenen i in der Geschwisterfolge (=1, wenn Erstgeborener; =2, wenn Zweitgeborenener usw.)

Sie schätzen das Modell mit dem Maximum-Likelihood-Verfahren und erhalten folgenden Output:

```
  Mixed-effects ML regression                    Number of obs      =      1191

                                                 Wald chi2(3)       =     71.64
Log likelihood = -377.00485                      Prob > chi2        =    0.0000

------------------------------------------------------------------------------
     rauchen |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      bildmu |  -.0269401   .0052222    -5.16   0.000    -.0371754   -.0167048
      bildva |  -.0081504   .0045887    -1.78   0.076    -.0171441    .0008433
    gebfolge |   .0017877   .0110598     0.16   0.872    -.0198891    .0234645
       _cons |   .5934029   .0600552     9.88   0.000     .4756968     .711109
------------------------------------------------------------------------------
```

Runden Sie in Ihren Antworten alle Zahlenangaben auf die vierte Nachkommastelle.

4.1 Interpretieren Sie inhaltlich und statistisch den Zusammenhang zwischen der Rauchwahrscheinlichkeit und dem Rang des Neugeborenen in der Geschwisterfolge. [2 Punkte]

- Inhaltlich: b_4=$0,0018$: Für jedes zusätzliche Geschwisterkind, das in der Geburtenfolge vor dem Neugeborenen i steht, liegt die Rauchwahrscheinlichkeit der Mutter des Neugeborenen i c.p. im Mittel um 0,18 Prozentpunkte höher. [1P]
- Alternativ: Inhaltlich: b_4=$0,0018$: Neugeborene mit einem älteren Geschwisterkind haben c.p. im Mittel eine um 0,18 Prozentpunkte höhere Rauchwahrscheinlichkeit der Mutter als Neugeborene ohne älteres Geschwisterkind, Neugeborene mit zwei älteren Geschwisterkindern haben c.p. im Mittel eine um 0,18 Prozentpunkte höhere Rauchwahrscheinlichkeit der Mutter als Neugeborene mit einem älteren Geschwisterkind usw.. [1P]
- Statistisch: p=$0,872 > 0,1$: Der Koeffizient ist statistisch nicht signifikant. [1P]

4.2 Erläutern Sie kurz allgemein das Vorgehen des Maximum-Likelihood-Verfahrens. Erläutern Sie dabei die Likelihood-Funktion und ihre Bedeutung im Schätzverfahren. [3 Punkte]

- Beim ML-Verfahren wird der Vektor der unbekannten Parameter β so bestimmt, dass die Wahrscheinlichkeit, mit dem Modell genau die vorliegenden Daten zu generieren, maximiert wird. [1P]
- Die Likelihood-Funktion ist dabei die gemeinsame Dichtefunktion für die beobachteten Werte der abhängigen Variable, [1P] die sogenannte Wahrscheinlichkeitsdichtefunktion.
- Der ML-Schätzer wird so gewählt, dass die (Log)-Likelihood-Funktion maximiert wird. [1P]

4.3 Welche Annahme über die Verteilung der Rauchwahrscheinlichkeit wurde bei der vorliegenden Schätzung getroffen? Welches Problem kann hierbei hinsichtlich der vorhergesagten Wahrscheinlichkeiten auftreten? [2 Punkte]

- Es wurde angenommen, dass die Wahrscheinlichkeit $P(rauchen_i = 1|\mathbf{x}_i)$ linear in den Parametern ist. [1P]
- Durch die Annahme der Linearität können die prognostizierten Wahrscheinlichkeiten außerhalb des Intervalls $[0,1]$ liegen. [1P]

4.4 Sie unterstellen nun, dass die Wahrscheinlichkeitsfunktion einer logistischen Verteilung folgt und schätzen das Modell mit einem Logit-Schätzer. Sie erhalten den folgenden Stata-Output:

```
  Logistic regression                               Number of obs   =       1191
                                                    LR chi2(3)      =      68.80
                                                    Prob > chi2     =     0.0000
Log likelihood = -437.37647                         Pseudo R2       =     0.0729

------------------------------------------------------------------------------
     rauchen |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      bildmu | -.2460909     .04779    -5.15   0.000    -.3397576   -.1524243
      bildva | -.0707888   .0395888    -1.79   0.074    -.1483814    .0068039
    gebfolge | -.0117786   .0978076    -0.12   0.904    -.2034779    .1799208
       _cons |    2.1447    .547336     3.92   0.000     1.071942    3.217459
------------------------------------------------------------------------------
```

Runden Sie in Ihren Antworten alle Zahlenangaben auf die vierte Nachkommastelle.

4.4.1 Interpretieren Sie den Koeffizienten für die Schulbildung des Vaters inhaltlich und statistisch. [2 Punkte]

- Inhaltlich: $b_3 = -0,0708$: Die Rauchwahrscheinlichkeit nimmt mit der Schulbildung des Vaters ab. [1P]
- Statistisch: $p = 0,074 < 0,1$: Der Koeffizient ist statistisch signifikant auf dem 10%-Niveau. [1P]

4.4.2 Berechnen und interpretieren Sie den marginalen Effekt der Schulbildung der Mütter für die erstgeborenen Neugeborenen. Nehmen Sie für die übrigen Variablen Stichprobenmittelwerte an. [6 Punkte]

- Marginale Effekte im Logit-Modell allgemein (aus der Formelsammlung):

$$\frac{\partial P(y_i = 1|\bar{\mathbf{x}})}{\partial x_{ik}} = \frac{\exp(\mathbf{x}'_{ik}\beta)}{(1+\exp(\mathbf{x}'_{ik}\beta))^2}\beta_k$$

- Bestimmung des marginalen Effekts im Logit-Modell für $bildmu_i$:

$$\mathbf{x}'_{ik}b\big|_{gebfolge=1,\bar{\mathbf{x}}} = 2,1447 - 0,2461 * 13,1251 - 0,0708 * 13,1914 - 0,0118 * 1 = -2,0311 \quad [3P]$$

$$\frac{\partial P(rauchen_i = 1|gebfolge = 1, \bar{\mathbf{x}})}{\partial bildmu_i} = \frac{\exp(-2,0311)}{(1+\exp(-2,0311))^2}(-0,2461) = -0.0252 \quad [2P]$$

- Die Wahrscheinlichkeit, dass die Mutter während der Schwangerschaft raucht, sinkt für die Mütter von Erstgeborenen c.p. im Mittel um 2,52 Prozentpunkte mit jedem zusätzlichen Bildungsjahr. [1P]

Aufgabe 5 - MC Fragen **[30 Punkte]**

Bitte geben Sie die zutreffende Antwort **auf Ihrem Multiple-Choice-Lösungsblatt** an. **Angaben auf dem Aufgabenblatt werden <u>nicht</u> gewertet.** Zu jeder Frage gibt es genau eine richtige Antwort. Für jede korrekt beantwortete Frage erhalten Sie einen Punkt. Falsche Antworten führen nicht zu Punktabzug. Bei mehr oder weniger als einer markierten Antwort auf eine Frage auf dem Lösungsblatt gilt diese als nicht beantwortet.

1.	Wenn für die Störterme eines linearen Regressionsmodells gilt, dass $\varepsilon \sim i.i.d.(0, \sigma^2)$, dann
a	ist der KQ Schätzer in kleinen Stichproben normalverteilt.
b	müssen die Standardfehler in Hinsicht auf Autokorrelation korrigiert werden.
c	sind t-Tests asymptotisch gültig. **X**
d	sollte die abhängige Variable so transformiert werden, dass sie normalverteilt ist.

2.	Ein konsistenter Schätzer
a	ist auch ein unverzerrter Schätzer.
b	gehört zur Klasse der Best Unbiased Linear Estimators (BLUE).
c	minimiert die Varianz des Störterms.
d	konvergiert asymptotisch zu dem unbekannten Populationsparameter. **X**

3.	Bei Vorliegen von Heteroskedastie und Autokorrelation gilt für die Varianz-Kovarianz Matrix des Störterms, dass die Einträge
a	auf der Hauptdiagonalen identisch und die Einträge abseits der Diagonalen nicht alle gleich 0 sind.
b	auf der Hauptdiagonalen identisch und die Einträge abseits der Diagonalen alle gleich 0 sind.
c	auf der Hauptdiagonalen unterschiedlich und die Einträge abseits der Diagonalen nicht alle gleich 0 sind. **X**
d	auf der Hauptdiagonalen unterschiedlich und die Einträge abseits der Diagonalen alle gleich 0 sind.

4.	Der KQ-Schätzer $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ lässt sich nur berechnen, wenn $\mathbf{X}'\mathbf{X}$
a	ein Skalar ist.
b	nicht symmetrisch ist.
c	invertierbar ist. **X**
d	linear abhängige Spalten aufweist.

5.	Das Kleinstquadrateverfahren basiert auf der Minimierung
a	der quadrierten Likelihoodfunktion.
b	der Summe der Residuen.
c	der Summe der quadrierten Residuen. **X**
d	der quadrierten Summe der Störterme.

6.	Eine χ^2-verteilte Variable entsteht durch
a	die Quadrierung einer standardnormalverteilten Variable.
b	die Quadrierung einer F-verteilten Variable.
c	die Summierung mehrerer quadrierter standardnormalverteilter Variablen. **X**
d	die Summierung mehrerer quadrierter F-verteilter Variablen.

7.	Ein Typ II Fehler liegt vor, wenn man
a	eine wahre Nullhypothese ablehnt.
b	eine falsche Nullhypothese ablehnt.
c	eine wahre Nullhypothese nicht ablehnt.
d	eine falsche Nullhypothese nicht ablehnt. **X**

8.	Wenn für zwei Zufallsvariablen X und Y gilt, dass $E(Y\|X) = E(Y) = 0$, dann
a	$cov(X,Y) = 0$. **X**
b	sind die Variablen X und Y statistisch unabhängig.
c	$Var(Y\|X) = Var(Y)$.
d	$E(X\|Y) = E(X) = 0$.

9.	Was ergibt sich aus dem Produkt der Matrizen $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 0 \end{pmatrix}$ und $B = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}$?
a	$AB = \begin{pmatrix} 7 \\ 19 \end{pmatrix}$. **X**
b	$AB = \begin{pmatrix} 10 & 5 \\ 7 & 19 \end{pmatrix}$.
c	$AB = \begin{pmatrix} 7 & 19 \end{pmatrix}$.
d	$AB = \begin{pmatrix} 10 & 7 \\ 19 & 5 \end{pmatrix}$.

10.	Wenn $\Psi = \mathbf{I}$ gilt, dann impliziert $Var(\varepsilon) = \sigma^2\Psi$,
a	dass $cov(\varepsilon_i, \varepsilon_j) = 0$ für alle $i \neq j$. **X**
b	dass Heteroskedastie vorliegt.
c	dass $cov(\varepsilon_i, \varepsilon_j) \neq 0$ für alle $i \neq j$.
d	dass $Var(\varepsilon) = 1$.

11.	Autokorrelation im Störterm kann behoben werden durch
a	die Aufnahme von ausgelassenen relevanten erklärenden Variablen. **X**
b	eine Durbin-Watson Transformation.
c	das Weglassen statistischer Ausreißer.
d	eine Schätzung mit robusten Standardfehlern.

12.	Bei Autokorrelation in Form von moving average Störprozessen
a	ist es möglich, dass einige Fehlerterme nicht miteinander korreliert sind.**X**
b	sind alle Fehlerterme unkorreliert.
c	hängt die Varianz eines Störterms von der Vorperiode ab.
d	ist der Prais-Winsten Schätzer zur Behebung der Autokorrelation geeignet.

13.	Die Annahme $\varepsilon \sim i.i.d.(0, \sigma^2)$
a	schließt Heteroskedastie und Autokorrelation aus. **X**
b	schließt Heteroskedastie, nicht aber Autokorrelation aus.
c	schließt Homoskedatie und Autokorrelation aus.
d	schließt Homoskedatie, nicht aber Autokorrelation aus.

14.	Der Durbin-Watson Test
a	eignet sich zum Testen auf Autokorrelation höherer Ordnung.
b	eignet sich zum Testen auf ausgelassene Variablen.
c	ist nur bei Schätzung mit Konstante gültig. **X**
d	ist nur anwendbar, wenn das Modell mindestens zwei Steigungsparameter enthält.

15.	Wenn autokorrelierte Störterme vorliegen, dann
a	ist die Varianz-Kovarianz Matrix des Störterms keine diagonale Matrix. **X**
b	ist der Feasible GLS-Schätzer BLUE.
c	ist der KQ Schätzer verzerrt.
d	sind die mit KQ-Schätzern ausgewiesenen Standardfehler gültig.

16.	Der marginale Effekt in einem linearen Wahrscheinlichkeitsmodell
a	ist bei Verwendung des KQ-Verfahrens nicht konstant.
b	entspricht immer dem Wert des Koeffizienten.
c	kann vom Niveau der erklärenden Variablen abhängen. **X**
d	hat nicht das gleiche Vorzeichen wie der geschätzte Koeffizient.

17.	Welche Aussage ist richtig?
a	Beim Wald-Test ist die Teststatistik asymptotisch F-verteilt.
b	Der geschätzte Anteil der abhängigen Variablen mit Ausprägung 0 entspricht im Logit-Modell (mit Konstante) dem beobachteten Anteil in der Stichprobe. **X**
c	Der marginale Effekt am Mittelwert der Daten ist beim Probit-Modell immer größer als der mittlere marginale Effekt.
d	Beim Likelihood-Ratio Test wird die Nullhypothese verworfen, wenn gilt: Log-Likelihood-Wert im unrestringierten Modell > Log-Likelihood-Wert im restringierten Modell.

18.	Beim ML-Verfahren führen normalverteilte Störterme ($\varepsilon \sim n.i.d.(0, \sigma^2)$)
a	zu asymptotisch unverzerrten ML-Schätzern.
b	zu einer eindeutigen analytischen Lösung für den ML-Schätzer.
c	im binären Fall zu Wahrscheinlichkeiten im Intervall [0,1].
d	im linearen Wahrscheinlichkeitsmodell zur Übereinstimmung von ML- und KQ-Schätzern. **X**

19.	Bei $\{Y_1, Y_2, ..., Y_N\} \sim i.i.d.$ entspricht die logarithmierte Likelihoodfunktion
a	der Summe der quadrierten logarithmierten Dichten der einzelnen Beobachtungen der abhängigen Variable.
b	dem Logarithmus der Summe der Dichten der einzelnen Beobachtungen der abhängigen Variable.
c	der Summe der logarithmierten Dichten der einzelnen Beobachtungen der abhängigen Variable. **X**
d	dem Produkt der logarithmierten Dichten der einzelnen Beobachtungen der abhängigen Variable.

20.	Ein RESET-Test überprüft,
a	ob sich die marginalen Effekte im Modell für Teilgruppen unterscheiden.
b	ob die Konstante signifikant von Null verschieden ist.
c	ob durch die Änderung der Modellspezifikation das R^2 signifikant steigt. **X**
d	ob statt abnehmenden marginale Effekten zunehmende vorliegen.

21.	Die Aufnahme einer irrelevanten Variable x_3 in ein lineares Regressionsmodell $y_i = \beta_1 + \beta_2 x_{i2} + \varepsilon_i$
a	erhöht die Varianz für den Parameterschätzer b_2, wenn $cov(x_2, x_3) \neq 0$. **X**
b	führt bei kleinen Stichproben zu inkonsistenten Schätzern.
c	erhöht die Varianz des Störterms.
d	führt bei Multikollinearität zu verzerrten Schätzern.

22.	Sie schätzen das Modell $ln(y_i) = \beta_1 + \beta_2 ln(x_{i2}) + \beta_3 x_{i3} + \varepsilon_i$ mit KQ und erhalten $b_2 = 0,5$ und $b_3 = 2$. Ceteris paribus und im Mittel gilt:
a	steigt x_{i3} um 1 Einheit steigt, so steigt y um 20%.
b	sinkt x_{i3} um eine Einheit, so sinkt y um 0,02%.
c	steigt x_{i2} um 1%, so steigt y um 0,5 Prozentpunkte.
d	steigt x_{i2} um 1%, so steigt y um 0,5%. **X**

23.	Das korrigierte Bestimmtheitsmaß im multiplen Regressionsmodell und bei endlicher Stichprobe
a	steigt c.p. mit Zunahme der Residuenquadratsumme.
b	ist immer kleiner als das Bestimmtheitsmaß R^2. **X**
c	eignet sich nicht zum Vergleich von Modellen, bei denen das eine Modell ein Spezialfall des anderen ist.
d	entspricht dem Anteil der erklärten Variation an der gesamten Variation der abhängigen Variablen.

24.	Die Schätzung des Modells $einkommen_i = \beta_1 + \beta_2 berufserfahrung_i + \varepsilon_i$ mit KQ ergibt bei 188 Beobachtungen $b_2 = 2,4$, $se(b_2) = 1,2$. Welche Aussage trifft für die Schätzung vonβ_2 zu?
a	b_2 ist signifikant auf dem 1%-Niveau, wenn $p = 0,02$.
b	Der kritische Wert für einen rechtsseitigen t-Test mit $\alpha = 0,05$ beträgt 1,95.
c	Die Teststatistik für den linksseitigen t-Test mit $H_0 : \beta_2 \geq 0$ lautet -2.
d	Die Nullhypothese beim zweiseitigen t-Test auf Signifikanz wird verworfen, wenn $\lvert$kritischer Wert$\rvert < 2$. **X**

25.	Für ein Modell ohne Konstante, 3 Regressoren und 204 Beobachtungen schätzen Sie eine Residuenquadratsumme von 80400. Welchen Wert hat die geschätzte Standardabweichung des Störterms?
a	20 **X**
b	400
c	80,4
d	201

26.	Sie schätzen das Modell $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$ und ihr Kommilitone das Modell $y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + \varepsilon_i$. In ihrem Modell wurde β_2 positiv verzerrt geschätzt, wenn
a	$cov(x,z) < 0$ und $\beta_2 < 0$.
b	$cov(x,z) = 0$ und $\beta_3 < 0$.
c	$cov(x,z) > 0$, $\beta_2 > 0$ und $\beta_3 < 0$.
d	$cov(x,z) > 0$ und $\beta_3 > 0$. **X**

27.	Im linearen Modell $y_i = \beta_1 + \beta_2 x_{i1} + \beta_3 x_{i1}^2 + \beta_4 x_{i2} + \beta_5 x_{i2}^2 + \varepsilon_i$ mit $\beta_2 > 0$, $\beta_3 > 0$, $\beta_4 < 0$ und $\beta_5 > 0$ ist
a	der marginale Effekt von x_1 für alle Werte von x_2 konstant. **X**
b	der marginale Effekt von x_2 für kleine x_2-Werte positiv und abnehmend.
c	der marginale Effekt von x_1 für kleine x_1-Werte positiv und abnehmend.
d	der marginale Effekt von x_2 konstant abnehmend.

28.	Das Informationskriterium AIC$= log \frac{1}{N} \sum_{i=1}^{N} e_i^2 + \frac{2K}{N}$
a	steigt bei Hinzunahme relevanter Variablen quadratisch an.
b	dient zum Vergleich nicht-genesteter Modelle. **X**
c	bleibt bei Hinzunahme irrelevanter Variablen konstant.
d	ist ein ansteigendes Maß für die Schätzgüte eines Modells.

29.	Welche Aussage trifft für die Schätzung des Modells zu, wobei $alter_i$ in Jahren, $abitur_i$ $0/1$ und $frau_i$ $0/1$: $Lohn_i = \beta_1 + \beta_2 abitur_i + \beta_3 frau_i + \beta_4 alter_i + \beta_5 abitur_i \cdot frau_i + \beta_6 alter_i \cdot frau_i + \varepsilon_i$?
a	Der marginale Effekt von $abitur_i$ für Männer lautet: $\beta_2 + \beta_5$.
b	Die Nullhypothese beim Chow-Test auf Geschlechterunterschiede lautet: $\beta_5 = \beta_6 = 0$.
c	Der vorhergesagte Lohn für 20-jährige Frauen ohne Abitur lautet: $\beta_1 + \beta_3 + \beta_4 * 20 + \beta_6 * 20$. **X**
d	β_3 misst den Lohnunterschied von Männern und Frauen ohne Abitur.

30.	Für das Modell $zufriedenheit_i = \beta_1 + \beta_2 freizeit_i + \beta_3 freizeit_i^2 + \varepsilon_i$ ergibt eine KQ-Schätzung b_2 = 10,4 und b_3=-2,08. Bei welchem Freizeitkonsum wird die Zufriedenheit maximiert?
a	5
b	8,32
c	2,5 **X**
d	12,48