Lehrstuhl für Statistik und empirische Wirtschaftsforschung
Prof. Regina T. Riphahn, Ph.D.

Masterprüfung Sommersemester 2022 - MUSTERLÖSUNG

Fach: Ökonometrie

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

Vorbemerkungen:

Anzahl der Aufgaben:	Die Klausur besteht aus 4 Aufgaben, die alle bearbeitet werden müssen. **Es wird nur der Lösungsbogen eingesammelt.**
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: **[20 Punkte]**

Sie sind daran interessiert, die Determinanten der Häufigkeit von jährlichen Schwimmbadbesuchen zu analysieren. Ihnen stehen dazu Daten aus einer Befragung von 3.670 Hobby-Schwimmerinnen und Hobby-Schwimmern aus dem Jahre 2021 mit folgenden Variablen zur Verfügung:

$ln\_svisits_i$	Anzahl der jährlichen Schwimmbadbesuche von Person i im Jahr 2021 (logarithmiert)
age_i	Alter der Person i in Jahren
$female_i$	Dummy-Variable, =1, wenn Person i weiblich, =0 sonst
$status20_i$	Selbsteinschätzung des Leistungszustandes von Person i für das Jahr 2020 (Skala von 1-5)
$status21_i$	Selbsteinschätzung des Leistungszustandes von Person i für das Jahr 2021 (Skala von 1-5)
$eduhigh_i$	Dummy-Variable, =1, wenn Person i eine hohe Bildung hat, =0 sonst

Sie stellen folgendes Regressionsmodell auf und schätzen es anschließend in Stata:

$$ln\_svisits_i = \beta_1 + \beta_2 age_i + \beta_3 female_i + \beta_4 age_i \cdot female_i + \beta_5 status21_i + \beta_6 eduhigh_i + \epsilon_i \quad \text{(Modell I)}$$

```
      Source |       SS       df       MS              Number of obs =    3670
-------------+------------------------------           F(  5,  3664) =  164.96
    Model    |  397.894727     5  79.5789453           Prob > F      =  0.0000
    Residual |  1767.55976  3664  .482412599           R-squared     =  0.1837
-------------+------------------------------           Adj R-squared =  0.1826
       Total |  2165.45449  3669  .590202913           Root MSE      =  .69456

------------------------------------------------------------------------------
 ln_svisits |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0048175    .001016     4.74   0.000     .0028256    .0068094
      female |   .3289255   .0735732     4.47   0.000     .1846769     .473174
  age_female |  -.0042282    .001311    -3.23   0.001    -.0067986   -.0016578
    status21 |   .3233856   .0132505    24.41   0.000     .2974064    .3493647
     eduhigh |    .068207   .0279707     2.44   0.015     .0133674    .1230467
       _cons |   1.104396   .0617708    17.88   0.000      .983287    1.225504
------------------------------------------------------------------------------
```

Runden Sie im Folgenden alle Zahlenangaben auf die dritte Nachkommastelle.

1.1 Interpretieren Sie den geschätzten Koeffizienten b_6 inhaltlich und statistisch. (2 Punkte)

- Die Anzahl an jährlichen Schwimmbadbesuchen im Jahr 2021 ist für Schwimmer mit hoher Bildung c.p. im Mittel um ca. 6,821 % höher als bei nicht hoch gebildeten Schwimmern. [1P]
- Der Koeffizient ist statistisch signifikant auf dem 5% Niveau. [1P]

1.2 Leiten Sie den marginalen Effekt der Variable *age* auf die Anzahl an Schwimmbadbesuche pro Jahr allgemein her. Wie groß ist der marginale Effekt für Frauen? Wie groß ist er für Männer? Interpretieren Sie die Ergebnisse. (4 Punkte)

- Marginaler Effekt allgemein:

$$\frac{\Delta E(ln\_dvisits_i)}{\Delta age_i} = \beta_2 + \beta_4 \cdot female_i \quad \text{[1P]} \tag{1}$$

- Marginaler Effekt für Frauen:

$$b_2 + b_4 \cdot 1 = 0,005 - 0,004 = 0,001\,(0,1\%) \quad [1P]$$

Für Frauen führt ein Anstieg des Alters um ein Jahr c.p. im Mittel zu einem Anstieg der jährlichen Schwimmbadbesuche um ca. 0,1%. [1P]

- Marginaler Effekt für Männer:

$$b_2 + b_4 \cdot 0 = 0,005\,(0,5\%) \quad [0,5P]$$

Für Männer führt ein Anstieg des Alters um ein Jahr c.p. im Mittel zu einem Anstieg der jährlichen Schwimmbadbesuche um ca. 0,5%. [0,5P]

1.3 Berechnen Sie die erwartete Anzahl an Schwimmbadbesuche pro Jahr eines 30-jährigen Mannes mit hoher Bildung, der seinen Leistungszustand 2021 mit 4 bewertet hat. (4 Punkte)

- Vorhersagewert für $ln\_svisits_i$:

$$\widehat{ln\_svisits}_i = b_1 + b_2 \cdot 30 + b_3 \cdot 0 + b_4 \cdot 0 + b_5 \cdot 4 + b_6 \cdot 1$$

$$= 1,104 + 0,005 \cdot 30 + 0,329 \cdot 0 - 0,004 \cdot 0 + 0,323 \cdot 4 + 0,068 \cdot 1 = 2,614 \quad [1P]$$

- Vorhersagewert für $svisits_i$:

$$\widehat{svisits}_i = exp(\widehat{ln\_svisits}_i + 0,5 \cdot \hat{\sigma}^2) = exp(2,614 + 0,5 \cdot 0,695^2) = 17,383 \quad [3P]$$

1.4 Sie vermuten, dass sich gleichzeitig mit dem Effekt von age auch der Effekt von $status21$ zwischen Männern und Frauen unterscheidet und führen einen entsprechenden F-Test auf Geschlechterunterschiede in beiden Effekten auf dem 1%-Signifikanzniveau durch. Sie erhalten folgenden Wert der Teststatistik: $F = 6,10$.

Stellen Sie das Modell, das dem F-Test unterliegt, nachvollziehbar dar. Geben Sie anschließend die Null- und Alternativhypothese, den kritischen Wert und das Testergebnis für den durchgeführten F-Test an. (4 Punkte)

- Modell:

$$ln\_svisits_i = \beta_1 + \beta_2 age_i + \beta_3 female_i + \beta_4 age_i \cdot female_i + \beta_5 status21_i$$

$$+ \beta_6 eduhigh_i + \beta_7 status21\_female_i + \epsilon_i \quad [1P]$$

- Hypothesen: $H_0: \beta_4 = \beta_7 = 0$ [0,5P]; H_1: Mind. einer der Koeffizienten ungleich Null. [0,5P]
- Kritischer Wert: $F_{2;3663;1\%} = 4,63$. [1P]
- Testergebnis: Da $F_{empirisch} = 6,10 > 4,63 = F_{kritisch}$ kann die Nullhypothese auf dem 1% Signifikanzniveau verworfen werden. [1P] Die Koeffizienten sind gemeinsam statistisch signifikant.

1.5 Ein Kommilitone behauptet, dass Ihre Variable $status21$ endogen ist. Nennen Sie einen möglichen Grund für das Vorliegen der Endogenität und erläutern Sie diesen. (1,5 Punkte)

- Mögliche Gründe (1 Nennung wird bepunktet):
  1. Umgekehrte Kausalität (reverse causality bias): [0,5P]
     Die Selbsteinschätzung des Leistungszustands für das Jahr 2021 hängt von der Anzahl an Schwimmbadbesuche im Jahr 2021 ab. [1P]
  2. Ausgelassene erklärende Variable (omitted variable bias): [0,5P]
     Faktoren, die mit $status21$ sowie $ln\_svisits$ korrelieren, sind nicht im Modell enthalten. [1P]
  3. Beispielsweise: Wohnortnähe zum Schwimmbad

1.6 Ihr Kommilitone rät Ihnen, Ihr Endogenitätsproblem durch Instrumentierung von $status21$ zu lösen. Er schlägt vor, die Selbsteinschätzung des Leistungsstatus aus dem Vorjahr ($status20$) als Instrument in einer two-stage-least-squares (2SLS)-Schätzung Ihres Modells I zu verwenden.

1.6.1 Stellen Sie die für die 2SLS-Schätzung benötigten Modellgleichungen auf. Nennen Sie einen Fall, in dem $status20$ kein geeignetes Instrument darstellt. (3 Punkte)

1.Stufe:

$$status21_i = \theta_1 + \theta_2 age_i + \theta_3 female_i + \theta_4 age\_female_i + \theta_5 eduhigh_i + \theta_6 status20_i + \upsilon_i \quad [1P]$$

2.Stufe:

$$ln(svisits_i) = \beta_1 + \beta_2 age_i + \beta_3 female_i + \beta_4 age\_female_i + \beta_5 \widehat{status21}_i + \beta_6 eduhigh_i + u_i \quad [1P]$$

- $status20$ korreliert nicht mit $status21$. [1P]
  Alternativ:
- $status20$ korreliert mit ϵ. [1P]

1.6.2 Wie können Sie testen, ob es sich bei $status20$ um ein schwaches Instrument handelt? Geben Sie die Nullhypothese des Tests an. (1,5 Punkte)

- F-Test auf Signifikanz von $status20$ auf der ersten Stufe der 2SLS-Schätzung. [1P]
- Nullhypothese: $H_0 : \theta_6 = 0$ [0,5P]

Aufgabe 2: [20 Punkte]

Sie interessieren sich für die Entwicklung der Mietpreise in Berlin für die Jahre 1995-2020. Es liegen folgende Variablen vor:

ln_price_t	durchschnittlicher Mietpreis pro Quadratmeter im Jahr t, logarithmiert
$year_t$	Kalenderjahr (1995-2020)
pop_t	Nettozuwanderung (in Tausend) in Jahr t

Sie stellen folgendes Regressionsmodell auf und schätzen dieses anschliessend mit Hilfe des Kleinst-Quadrate-Schätzers:

$$ln\_price_t = \beta_1 + \beta_2 year_t + \beta_3 pop_t + \epsilon_t$$

Runden Sie alle Zahlenangaben auf die dritte Nachkommastelle.

2.1 Sie erhalten $b_2 = 0,03491$. Interpretieren Sie den geschätzten Koeffizienten inhaltlich. (1 Punkt)

- Mit jedem zusätzlichen Jahr steigt der durchschnittliche Mietpreis pro Quadratmeter um ca. 3,491 %. [1P]

2.2 Erläutern Sie, was man unter Autokorrelation versteht und welche Konsequenz dies für den KQ Schätzer hat. (2 Punkte)

- Bei Vorliegen von Autokorrelation sind die Störterme zeitlich benachbarter Beobachtungen korreliert. [1P]
- Der KQ Schätzer ist weiterhin unverzerrt, solange $E(\epsilon|x) = 0$ gilt. [0,5P]
- Ohne Korrektur sind die Standardfehler falsch und der KQ Schätzer ist ineffizient. [0,5P]

2.3 Nennen sie zwei ökonometrische Gründe, warum Autokorrelation vorliegen könnte und geben Sie für jeden Grund ein Beispiel an. (2 Punkte)

Jeweils ein Punkt pro Grund

- Dynamische Fehlspezifikation, z.B. verzögerter Preis als erklärende Variable
- Ausgelassene relevante Variablen, z.B. Angebot an Mietwohnungen
- Fehlspezifikation der funktionalen Form, z.B. nicht-linearer Zeittrend

2.4 Beschreiben Sie die Grundidee des Breusch-Godfrey Tests auf Autokorrelation 1. Ordnung. Führen Sie den Test am 1%-Signifikanzniveau für das Modell durch. Geben Sie Hypothesen, Hilfsregression, kritischen Wert und Entscheidungsregel an. Liefert der Test einen Hinweis auf das Vorliegen von Autokorrelation 1. Ordnung, wenn der Wert der Teststatistik 10,423 beträgt. (5 Punkte)

- Der Breusch-Godfrey Test untersucht, ob verzögerte Störterme einen Einfluss auf den Erklärungsgehalt des Modells haben. [1P]
- Nullyhypothese: $H_0 : \rho = 0$ (Es liegt keine Autokorrelation 1. Ordnung vor.) [0,5P]
- Alternativhypothese: $H_1 : \rho \neq 0$ (Es liegt Autokorrelation vor.) [0,5P]
- Teststatistik: $\chi^2_{empirisch} = (T-1) \cdot R^2$
- Hilfsregression: $\epsilon_t = \alpha_1 + \rho\epsilon_{t-1} + \nu_t$ [1P]
- kritischer Wert: $\chi^2_{kritisch} = \chi^2_1 = 6,635$ [0,5P]
- Entscheidungsregel: Wenn $\chi^2_{empirisch} > \chi^2_{kritisch}$, dann wird die Nullhypothese verworfen. Hier ist $10,423 > 6,635$. Somit liefert der Test einen Hinweis auf das Vorliegen von Autokorrelation 1. Ordnung. [1,5P]

2.5 Erklären Sie, wie der Breusch-Godfrey Test vorgeht, wenn Sie zusätzlich auf Autokorrelation zweiter Ordnung testen. (1 Punkt)

- Es werden zusätzlich die um zwei Perioden verzögerten Residuen als Regressoren in die Hilfsregression aufgenommen. [1P]
- $\epsilon_t = \alpha_1 + \rho_1\epsilon_{t-1} + \rho_2\epsilon_{t-2} + \nu_t$
- Die Freiheitsgrade der Teststatistik erhöhen sich um 1.

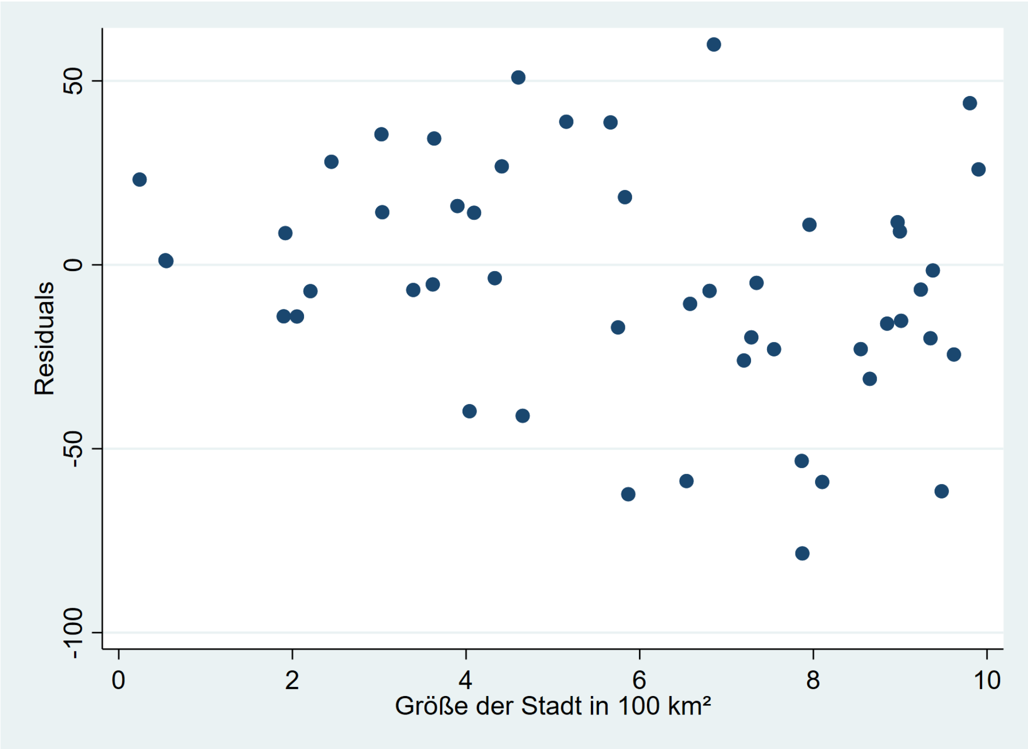
2.6 Ihr Kommilitone beschäftigt sich ebenfalls mit der Entwicklung der Mietpreise in Deutschland. Ihm liegt ein Datensatz aus dem Jahr 2021 vor, der folgende Informationen für verschiedene Städte in Deutschland enthält.

$price_i$	durchschnittliche Mietpreis pro Quadratmeter in Stadt i
$space_i$	Größe der Stadt i in 100 km^2
pop_i	Bevölkerung (in Mio.) in Stadt i

Sie stellen folgendes Regressionsmodell auf und schätzen dieses anschliessend mit Hilfe des Kleinst-Quadrate-Schätzers:

$$price_i = \beta_1 + \beta_2 space_i + \beta_3 pop_i + \epsilon_i$$

Sie stellen die Residuen der Schätzung und die Stadtgröße graphisch dar.



Liegt im vorliegenden Fall Heteroskedastie vor? Erläutern Sie Ihre Antwort. (2 Punkte)

- `Ja, im vorliegenden Fall liegt Heteroskedastie vor. Die Varianz der Störterme ist nicht konstant, sondern variiert mit der Größe der Stadt. [2P]`

2.7 Es sei $Var(\epsilon_i) = \sigma^2 \cdot space_i^4$. Zeigen Sie formal die aus einer GLS-Transformation resultierende Schätzgleichung, welche zu konstanter Störtermvarianz führt. Leiten Sie zudem die Varianz des Störterms im transformierten Modell her. (3 Punkte)

- $\frac{price_i}{space_i^2} = \frac{\beta_1}{space_i^2} + \frac{\beta_2 space_i}{space_i^2} + \frac{\beta_3 pop_i}{space_i^2} + \frac{\epsilon_i}{space_i^2}$ `[1,5P]`
- $Var(\frac{\epsilon_i}{space_i^2}) = \frac{1}{space_i^4} * Var(\epsilon_i) = \frac{1}{space_i^4} \cdot \sigma^2 \cdot space_i^4 = \sigma^2$ `[1,5P]`

2.8 Erklären Sie die Vorgehensweise des White Tests. Gehen Sie auf die Hypothesen, die Teststatistik und die Schlusslogik ein. (4 Punkte)

- `Nullypothese:` $V(\epsilon_i) = \sigma$ `für alle i (Homoskedastie) [0,5P]`
- `Alternativhypothese:` $H_1 : V(\epsilon_i) \neq V(\epsilon_j)$ `für mind. ein` $i \neq j$ `(Heteroskedastie) [0,5P]`
- `Teststatistik:` $\chi^2_{empirisch} = N \cdot R^2$`, wobei N=Anzahl der Beobachtungen und` R^2 `ist das Bestimmtheitsmaß der Hilfsregression. Die Hilfsregression regressiert die quadrierten Residuen auf alle erklärenden Variablen, deren Quadrate und jeweiligen Interaktionen. [2P]`
- `Schlusslogik: Wenn` $\chi^2_{empirisch} > \chi^2_{kritisch}$`, dann wird die Nullhypothese verworfen. [1P]`

Aufgabe 3: **[20 Punkte]**

Sie interessieren sich für die Determinanten von Immobilienbesitz. Sie nutzen einen Datensatz mit 447 Haushalten mit Wohnsitz in Deutschland. Es liegen folgende Variablen vor:

$immo_i$	Dummy-Variable, =1, wenn Haushalt i eine Immobilie besitzt, =0, sonst
inc_i	Jährliches Einkommen von Familie i, gemessen in 1000 Euro
$educ_i$	Anzahl an Schuljahren des Haushaltsvorstandes
$kids_i$	Anzahl an minderjährigen Kindern in Haushalt i
ost_i	Dummy-Variable, =1, wenn der Wohnsitz von Haushalt i in Ostdeutschland liegt, =0, wenn er in Westdeutschland liegt

Sie stellen folgendes Regressionsmodell auf und schätzen dieses anschliessend mit Hilfe des Kleinst-Quadrate-Schätzers:

$$immo_i = \beta_1 + \beta_2 inc_i + \beta_3 educ_i + \beta_4 kids_i + \beta_5 ost_i + \epsilon_i$$

immo	Coef.	Std. Err.	t
cons	-.24003211	.0543800	4.41
inc	.02643564	.0005736	46.09
educ	.01684755	.0099739	1.69
kids	-.07287933	.0332959	-2.19
ost	-.10557388	.0425521	-2.48

Runden Sie alle Zahlenangaben auf die dritte Nachkommastelle.

3.1 Interpretieren Sie b_3 und b_5 inhaltlich. (2 Punkte)

- `Ein zusätzliches Schuljahr des Haushaltsvorstandes erhöht ceteris Paribus im Mittel die Wahrscheinlichkeit des Immobilienbesitzes um 1,685 Prozentpunkte. [1P]`
- `Für Haushalte in Ostdeutschland ist ceteris paribus im Mittel die Wahrscheinlichkeit des Immobilienbesitzes um 10,6 Prozentpunkte niedriger als für Haushalte in Westdeutschland. [1P]`

3.2 Wie hoch ist die vorhergesagte Wahrscheinlichkeit des Immobilienbesitzes für einen Haushalt in Ostdeutschland mit einem Jahreseinkommen von 30.000 Euro, keinen Kindern, und 12 Jahren Schulbildung des Haushaltsvorstandes? Zeigen Sie Ihren Rechenweg. Beachten Sie die Rundungsvorgaben. (3 Punkte)

- $\hat{immo}_i = -0,240 + 0,026 * 30 + 0,017 * 12 + (-0,073) * 0 + (-0,106) * 1 = 0,638$ [2P]
- Die vorhergesagte Wahrscheinlichkeit des Immobilienbesitzes beträgt 63,8%. [1P]

3.3 Welchen Wert würde b_2 annehmen, wenn inc_i in 100 Euro anstatt in 1.000 Euro gemessen würde? Zeigen Sie Ihren Rechenweg. (3 Punkte)

- $immo_i = \beta_1 + (\beta_2/10)(inc_i * 10) + \beta_3 educ_i + \beta_4 kids_i + \epsilon_i$ [1P]
- $\tilde{\beta_2} = (\beta_2/10)$ [1P]
- b_2 würde den Wert $0,02643564 * 10 = 0,264$ annehmen [1P]

3.4 Ein Kommilitone merkt an, dass Ihr Modell eventuell fehlspezifiziert sein könnte. Mit welchem Test können Sie überprüfen, ob nichtlineare Funktionen des geschätzten Modells dazu beitragen, $immo_i$ zu erklären? Geben Sie den Namen des Tests, die Hilfsregression und die Hypothesen an. (3 Punkte)

- Test: RESET-Test [1P]
- Hilfsregression: $y_i = x_i\prime\beta + \alpha_2 \hat{y_i}^2 + \alpha_3 \hat{y_i}^3 + \alpha_4 \hat{y_i}^4 + \upsilon_i$ [1P]
- es ist möglich auch weniger Polynome aufzunehmen.
- Nullhypothese: $\alpha_2 = \alpha_3 = \alpha_4 = 0$ [0,5P]
- Alternativhypothese: mind. ein $\alpha_j \neq 0$ mit $j = 2, 3, 4$ [0,5P]

3.5 Ihr Kommilitone vermutet, dass sich der Zusammenhang zwischen den erklärenden Variablen und dem Immobilienbesitz zwischen West- und Ostdeutschland unterscheidet und schlägt vor, das obige Regressionsmodell getrennt für Haushalte in West- und Ostdeutschland zu schätzen. Schreiben Sie ein Modell auf, mit dem Sie stattdessen die Hypothese Ihres Kommilitonen mit Hilfe von nur einer Regression testen können. Für welche Parameterwerte ist der Zusammenhang zwischen Einkommen und Immobilienbesitz in Westdeutschland positiv, in Ostdeutschland jedoch negativ? (5 Punkte)

- $immo_i = \beta_1 + \beta_2 inc_i + \beta_3 (inc_i * ost_i) + \beta_4 educ_i + \beta_5 (educ_i * ost_i) + \beta_6 kids_i + \beta_7 (kids_i * ost_i) + \beta_8 ost_i + \epsilon_i$ [3P]
- Für $\beta_2 > 0$ und $\beta_2 + \beta_3 < 0$ ist der Zusammenhang zwischen Einkommen und Immobilienbesitz in Westdeutschland positiv und in Ostdeutschland negativ. [2P]

3.6 Nennen Sie zwei Nachteile eines linearen Wahrscheinlichkeitsmodells. (2 Punkte)

- pro richtige Antwort 1P, max 2P
- Die vorhergesagten Werte für die abhängige Variable können Werte außerhalb [0,1] annehmen.
- Die Varianz der Störterme ist nicht konstant, es liegt Heteroskedastie vor.
- Alternative Antworten möglich.

3.7 Sie entscheiden sich nun für die Schätzung eines Probit-Modells anstatt des obigen linearen Wahrscheinlichkeitsmodells. Nennen Sie die in der Probit-Schätzung getroffene Annahme zur Verteilung des Störterms sowie eine alternative Annahme der Verteilung des Störterms. (2 Punkte)

- `Annahme bei Probit: standardnormalverteilter Störterm. [1P]`
- `Alternative: standardlogistische Verteilung des Störterms bei logistischer Regression. [1P]`

Aufgabe 4: **[30 Punkte]**

Bitte geben Sie die zutreffende Antwort **auf Ihrem Multiple-Choice-Lösungsblatt** an. 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 gilt diese als nicht beantwortet. **Angaben auf dem Aufgabenblatt werden nicht gewertet.**

1.	Für ein Modell mit Konstante, 2 Regressoren und 228 Beobachtungen schätzen Sie eine Residuenquadratsumme von 27450. Berechnen Sie die geschätzte Varianz des Störterms.
a	11,04
b	55.30
c	122,00 **X**
d	74,12

2.	Bei Vorliegen von Homoskedastie und Autokorrelation gilt für die Varianz-Kovarianz Matrix des Störterms, dass die Einträge
a	auf der Hauptdiagonalen konstant und Einträge abseits der Diagonalen ungleich 0 sind. **X**
b	auf der Hauptdiagonalen konstant und die Einträge abseits der Diagonalen gleich 0 sind.
c	auf der Hauptdiagonalen unterschiedlich und Einträge abseits der Diagonalen ungleich 0 sind.
d	auf der Hauptdiagonalen unterschiedlich und die Einträge abseits der Diagonalen gleich 0 sind.

3.	Für die Schätzgleichung $\hat{y}_i = 2,5 - 4 \cdot x_{i2} + 2 \cdot x_{i3}$ und die Beobachtung $(y_i, x_{i2}, x_{i3}) = (-1, 1, 1)$ beträgt das Residuum
a	-2,5.
b	-1,5. **X**
c	-0,5.
d	0,5.

4.	Von einem überidentifizierten Modell spricht man, wenn
a	auch die exogenen Regressoren in der ersten Stufe stark mit der endogenen Variable korrelieren.
b	der F-Wert eines Instruments in der ersten Stufe > 10 ist.
c	mehr Instrumente als endogene Regressoren vorliegen. **X**
d	sich OLS und 2SLS Koeffizienten signifikant unterscheiden.

5.	Sie schätzen das Modell $ln\_workhours_i = \beta_1 + \beta_2 male_i + \beta_3 ln\_income_i + \epsilon_i$ mittels einer KQ-Schätzung, wobei *ln_workhours* der Logarithmus der Arbeitszeit in Stunden ist, *male* eine Dummyvariable für Männer ist und *ln_wage* den logarithmierten Stundenlohn darstellt. Der geschätzte Koeffizient b_3 ist $0,12$. Welche Interpretation von b_3 ist richtig?
a	Steigt der Stundenlohn um 12%, so steigt im Durchschnitt c.p. die Arbeitszeit um 1%.
b	Steigt der Stundenlohn um 0,12%, so steigt im Durchschnitt c.p. die Arbeitszeit um 1%.
c	Steigt der Stundenlohn um 1%, so steigt im Durchschnitt c.p. die Arbeitszeit um 12%.
d	Steigt der Stundenlohn um 1%, so steigt im Durchschnitt c.p. die Arbeitszeit um 0,12%. **X**

6.	Ein Typ 2-Fehler tritt auf, wenn man die Nullhypothese
a	verwirft, obwohl diese falsch ist.
b	nicht verwirft, obwohl diese zutrifft.
c	nicht verwirft, obwohl diese falsch ist. **X**
d	verwirft, obwohl diese zutrifft.

7.	Das korrigierte Bestimmtheitsmaß
a	ist immer größer als das Bestimmtheitsmaß R^2.
b	gibt das Verhältnis von Gesamtvariation zu erklärter Variation an.
c	berücksichtigt die zur Schätzung benötigten Freiheitsgrade. **X**
d	zeigt an, ob die Parameter unverzerrt geschätzt werden.

8.	Das Maximum-Likelihood Schätzverfahren
a	minimiert die Summe der quadrierten Residuen.
b	bestimmt die geschätzten Parameter so, dass die Wahrscheinlichkeit für das Auftreten der Stichprobe maximiert wird. **X**
c	bestimmt die Populationsparameter so, dass die Wahrscheinlichkeit für das Auftreten der Stichprobe maximiert wird.
d	bestimmt die Stichprobe so, dass die Wahrscheinlichkeit für die Populationsparameter maximiert wird.

9.	Unabhängigkeit der zwei Zufallsvariablen A und B impliziert,
a	dass die Summe von A und B gleich 0 ist.
b	a und c sind richtig.
c	dass die Kovarianz zwischen A und B ungleich 0 ist.
d	dass die Korrelation zwischen A und B gleich 0 ist. **X**

10.	Das Informationskriterium BIC
a	wird unter Verwendung des R^2 berechnet.
b	bewertet die Unverzerrtheit von Parametern einer Schätzung.
c	kann zum Vergleich nicht-genesteter Modelle verwendet werden. **X**
d	berücksichtigt bei der Berechnung ausschließlich die Anzahl der Regressoren.

11.	Bei Vorliegen von perfekter Multikollinearität ist die Kreuzproduktmatrix $\mathbf{X'X}$
a	invertierbar.
b	**X** singulär.
c	ein Skalar.
d	idempotent.

12.	Bei Vorliegen von Heteroskedastie
a	können die Koeffizienten konsistent geschätzt sein.**X**
b	müssen die Parameter kausal interpretiert werden.
c	gilt das Gauss-Markov-Theorem.
d	sind die Konfidenzintervalle korrekt berechnet.

13.	Mithilfe eines linearen Regressionsmodells lassen sich
a	keine Koeffizienten schätzen.
b	keine Semielastizitäten schätzen.
c	keine Elastizitäten schätzen.
d	Keine der Aussagen ist richtig. **X**

14.	Der marginale Effekt in einem Logit-Modell
a	entspricht dem Wert des Koeffizienten.
b	hat nicht das gleiche Vorzeichen wie der geschätzte Koeffizient.
c	wird immer als Elastizität interpretiert.
d	hängt von den Ausprägungen der erklärenden Variablen ab.**X**

15.	Ein Strukturbruchtest überprüft,
a	ob ein Problem der Modellspezifikation vorliegt. **X**
b	ob zusätzliche erklärende Variablen die Varianz des Störterms negativ werden lässt.
c	ob die Kreuzproduktmatrix $\mathbf{X'X}$ invertierbar ist.
d	ob bei Zeitreihendaten Autokorrelation vorliegt.

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

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

18.	Die FGLS-Schätzung bei heteroskedastischen Störtermen beruht darauf, dass
a	nur die abhängige Variable so transformiert wird, dass Homoskedastie vorliegt.
b	Beobachtungen mit großer Störtermvarianz ein kleineres Gewicht erhalten als Beobachtungen mit kleiner Störtermvarianz.**X**
c	Beobachtungen mit großer Störtermvarianz ein größerers Gewicht erhalten als Beobachtungen mit kleiner Störtermvarianz.
d	die Form der Heteroskedastie bekannt ist.

19.	Die Alternativhypothese im Breusch-Pagan Test besagt,
a	dass sowohl Heteroskedastie als auch Autokorrelation vorliegen.
b	dass weder Heteroskedastie noch Autokorrelation vorliegen.
c	dass Heteroskedastie vorliegt.**X**
d	dass Autokorrelation vorliegt.

20.	Welche Folge hat die Aufnahme einer irrelevanten Variable in ein empirisches Modell, das mithilfe der OLS-Methode geschätzt werden soll?
a	Die Varianz der Schätzer steigt.**X**
b	Die Schätzer sind inkonsistent.
c	Die Schätzer sind verzerrt.
d	Der Betrag des Schätzers steigt.

21.	Mit welchem Test kann unter Annahme der Gültigkeit der Instrumente die Endogenität einer Variable überprüft werden?
a	Hausman-Test.**X**
b	Breusch-Godfrey Test.
c	Strukturbruchtest.
d	Wald-Test.

22.	Der Durbin-Watson Test ist nicht verlässlich, wenn
a	die Störterme normalverteilt sind.
b	die Regressoren nicht-stochastisch sind.
c	das Modell eine Regressionskonstante enthält.
d	Autokorrelation 2. Ordnung vorliegt. **X**

23.	Liegt Heteroskedastie vor, so
a	sind die Standardfehler falsch berechnet. **X**
b	ist die Schätzung effizient.
c	ist der KQ-Schätzer BLUE.
d	ist die Varianz der Störterme konstant.

24.	Je stärker die Korrelation zwischen zwei erklärenden Variablen,
a	desto schmaler das Konfidenzintervall.
b	desto geringer die Varianz.
c	desto kleiner die t-Werte. **X**
d	desto geringere Stichprobengrößen sind vorteilhaft.

25.	Welche Aussage ist richtig?
a	Die Teststatistik des Wald-Tests folgt asymptotisch der t-Verteilung.
b	In kleinen Stichproben sind Likelihood-Ratio- und Lagrange-Multiplier-Test aquivalent.
c	Der Likelihood-Ratio Test ist nur bei nicht-genesteten Modellen anwendbar.
d	Keine der Aussagen ist richtig. **X**

26.	Der geschätzte Koeffizient b_2 des Modells $y_i = \beta_1 + \beta_2 x_i + \epsilon_i$ ist überschätzt, falls eine ausgelassene Variable z_i
a	positiv mit x_i korreliert ist und y_i positiv beeinflusst. **X**
b	negativ mit x_i korreliert ist und y_i positiv beeinflusst.
c	mit x_i unkorreliert ist und y_i negativ beeinflusst.
d	positiv mit x_i korreliert ist und y_i negativ beeinflusst.

27.	Im log-linearen Modell $ln(y_i) = \beta_1 + \beta_2 x_i + \epsilon_i$
a	werden prozentuale Änderungen von x_i betrachtet.
b	ist β_2 eine Elastizität.
c	ist die abhängige Variable eine lineare Funktion von x_i. **X**
d	ist der KQ-Schätzer nicht berechenbar.

28.	Der marginale Effekt in einem Modell mit binärer abhängiger Variable
a	kann nicht mit KQ geschätzt werden.
b	hat das selbe Vorzeichen wie der geschätzte Koeffizient. **X**
c	unterscheidet sich nicht zwischen dem Logit- und Probit-Schätzer.
d	beschreibt die Änderung der unabhängigen Variable bei Änderung der abhängigen Variable um eine Einheit.

29.	Folgendes lineare Modell $lohn = \beta_1 + \beta_2 \cdot educ + \beta_3 \cdot educ^2 + \beta_4 \cdot frau + \beta_5 \cdot frau \cdot educ + \epsilon$ wird mit dem KQ Verfahren geschätzt. Dabei ist $lohn$ der Stundenlohn, $educ$ die Anzahl der Bildungsjahre und $frau$ eine Dummy-Variable für Frauen. Wann ist der marginale Effekt der Bildung von Frauen größer als der marginale Effekt für Männer?
a	Wenn $\beta_1 < 0$.
b	Wenn $\beta_4 > 0$.
c	Wenn $\beta_5 > 0$. **X**
d	keine Aussage möglich.

30.	Die Aufnahme einer irrelevanten Variable z in die KQ-Schätzung des Modells $y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + \epsilon_i$
a	führt zu inkonsistenten Koeffizientenschätzern, wenn die Stichprobe klein ist.
b	erhöht das adjustierte Bestimmtheitsmaß.
c	führt zu einer Verzerrung von β_2, wenn $cov(x, z) = 0$.
d	hat einen Einfluss auf die Präzision der Schätzung. **X**

Tabelle 2: Perzentile der *t*-Verteilung

Zelleneintrag: x, sodass Prob$\left[t_n \le x\right] = P$, mit n Freiheitsgraden

n \ P	0.75	0.9	0.95	0.975	0.99	0.995
1	1.000	3.078	6.314	12.706	31.821	63.657
2	0.817	1.886	2.920	4.303	6.965	9.925
3	0.765	1.638	2.353	3.182	4.541	5.841
4	0.741	1.533	2.132	2.776	3.747	4.604
5	0.727	1.476	2.015	2.571	3.365	4.032
6	0.718	1.440	1.943	2.447	3.143	3.707
7	0.711	1.415	1.895	2.365	2.998	3.500
8	0.706	1.397	1.860	2.306	2.896	3.355
9	0.703	1.383	1.833	2.262	2.821	3.250
10	0.700	1.372	1.813	2.228	2.764	3.169
11	0.697	1.363	1.796	2.201	2.718	3.106
12	0.696	1.356	1.782	2.179	2.681	3.055
13	0.694	1.350	1.771	2.160	2.650	3.012
14	0.692	1.345	1.761	2.145	2.624	2.977
15	0.691	1.341	1.753	2.131	2.602	2.947
16	0.690	1.337	1.746	2.120	2.583	2.921
17	0.689	1.333	1.740	2.110	2.567	2.898
18	0.688	1.330	1.734	2.101	2.552	2.878
19	0.688	1.328	1.729	2.093	2.539	2.861
20	0.687	1.325	1.725	2.086	2.528	2.845
21	0.686	1.323	1.721	2.080	2.518	2.831
22	0.686	1.321	1.717	2.074	2.508	2.819
23	0.685	1.320	1.714	2.069	2.500	2.807
24	0.685	1.318	1.711	2.064	2.492	2.797
25	0.684	1.316	1.708	2.060	2.485	2.787
26	0.684	1.315	1.706	2.056	2.479	2.779
27	0.684	1.314	1.703	2.052	2.473	2.771
28	0.683	1.313	1.701	2.048	2.467	2.763
29	0.683	1.311	1.699	2.045	2.462	2.756
30	0.683	1.310	1.697	2.042	2.457	2.750
35	0.682	1.306	1.690	2.030	2.438	2.724
40	0.681	1.303	1.684	2.021	2.423	2.705
45	0.680	1.301	1.679	2.014	2.412	2.690
50	0.679	1.299	1.676	2.009	2.403	2.678
60	0.679	1.296	1.671	2.000	2.390	2.660
70	0.678	1.294	1.667	1.994	2.381	2.648
80	0.678	1.292	1.664	1.990	2.374	2.639
90	0.677	1.291	1.662	1.987	2.368	2.632
100	0.677	1.290	1.660	1.984	2.364	2.626
∞	0.674	1.282	1.645	1.960	2.326	2.576

Quelle: In R generiert

Tabelle 4b: 99% Perzentile der *F*-Verteilung

Zelleneintrag: f, sodass Prob $\left[F_{n1,n2} \leq f\right] = 0.99$

n2	n1 = Freiheitsgrade des Zählers 1	2	3	4	5	6	7	8	9
1	4052.18	4999.50	5403.35	5624.58	5763.65	5858.99	5928.36	5981.07	6022.47
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
50	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.79
70	7.01	4.92	4.07	3.60	3.29	3.07	2.91	2.78	2.67
100	6.90	4.82	3.98	3.51	3.21	2.99	2.82	2.69	2.59
∞	6.66	4.63	3.80	3.34	3.04	2.82	2.66	2.53	2.43

n2	n1 = Freiheitsgrade des Zählers 10	12	15	20	30	40	50	60	∞
1	6055.85	6106.32	6157.28	6208.73	6260.65	6286.78	6302.52	6313.03	6362.68
2	99.40	99.42	99.43	99.45	99.47	99.47	99.48	99.48	99.50
3	27.23	27.05	26.87	26.69	26.50	26.41	26.35	26.32	26.14
4	14.55	14.37	14.20	14.02	13.84	13.75	13.69	13.65	13.47
5	10.05	9.89	9.72	9.55	9.38	9.29	9.24	9.20	9.03
6	7.87	7.72	7.56	7.40	7.23	7.14	7.09	7.06	6.89
7	6.62	6.47	6.31	6.16	5.99	5.91	5.86	5.82	5.66
8	5.81	5.67	5.52	5.36	5.20	5.12	5.07	5.03	4.87
9	5.26	5.11	4.96	4.81	4.65	4.57	4.52	4.48	4.32
10	4.85	4.71	4.56	4.41	4.25	4.17	4.12	4.08	3.92
15	3.80	3.67	3.52	3.37	3.21	3.13	3.08	3.05	2.88
20	3.37	3.23	3.09	2.94	2.78	2.69	2.64	2.61	2.43
25	3.13	2.99	2.85	2.70	2.54	2.45	2.40	2.36	2.18
30	2.98	2.84	2.70	2.55	2.39	2.30	2.25	2.21	2.02
40	2.80	2.66	2.52	2.37	2.20	2.11	2.06	2.02	1.82
50	2.70	2.56	2.42	2.27	2.10	2.01	1.95	1.91	1.70
70	2.59	2.45	2.31	2.15	1.98	1.89	1.83	1.78	1.56
100	2.50	2.37	2.22	2.07	1.89	1.80	1.74	1.69	1.45
∞	2.34	2.20	2.06	1.90	1.72	1.61	1.54	1.50	1.16

Quelle: In R generiert

Formeln Ökonometrie

I. Mathematische Grundlagen

i. Algebra

$$(AB)' = B'A'$$

$$(A')' = A$$

$$AA^{-1} = I \text{ und } A^{-1}A = I$$

$$\left(A^{-1}\right)' = \left(A'\right)^{-1}$$

$$(AB)^{-1} = B^{-1}A^{-1}$$

Ableitung von Matrizen

Für die Matrix A und die Vektoren x und c gilt bei passender Ordnung:

$$\frac{\partial c'x}{\partial x} = c$$

$$\frac{\partial Ax}{\partial x} = A'$$

Wenn A symmetrisch ist: $\frac{\partial x'Ax}{\partial x} = 2Ax$

ii. Varianz, Kovarianz und Korrelationskoeffizient

Varianz:

$$\sigma^2 = V\{Y\} = E\left\{\left(Y - E(Y)\right)^2\right\}$$
$$= E\{Y^2\} - E\{Y\}^2$$

Kovarianz:

$$\sigma_{YX} = \text{cov}\{Y, X\} = E\left\{\left(Y - E(Y)\right)\left(X - E(X)\right)\right\}$$
$$= E\{XY\} - E\{X\}E\{Y\}$$

Korrelationskoeffizient:

$$\rho_{YX} = \frac{\text{cov}\{Y, X\}}{\sqrt{V\{X\} \cdot V\{Y\}}} = \frac{\sigma_{YX}}{\sigma_Y \sigma_X}, \quad -1 \leq \rho_{YX} \leq 1$$

X,Y sind *unkorreliert*, wenn cov{Y,X} = 0

***Rechenregeln*:**

Wenn a, b, c, d Skalare und X, Y Zufallsvariablen sind:

$$V\{aY + b\} = a^2 \, V\{Y\}$$

$$V\{aY + bX\} = a^2 V\{Y\} + b^2 V\{X\} + 2ab\,\text{cov}\{Y, X\}$$

II. Annahmen im linearen Modell

A 1 $E\{\varepsilon_i\} = 0$ i = 1,2,..., N

A 2 $\{x_1, \ldots, x_N\}$ und $\{\varepsilon_1, \ldots, \varepsilon_N\}$ sind unabhängig

A 3 $V\{\varepsilon_i\} = \sigma^2$ i = 1,2,…, N

A 4 $\text{cov}\{\varepsilon_i, \varepsilon_j\} = 0$ i,j = 1,…, N, $i \neq j$

A 5 $\varepsilon \sim N(0, \sigma^2 I_N)$

A 5' $\varepsilon_i \sim NID(0, \sigma^2)$

A 6 $\frac{1}{N}\sum_{i=1}^{N} x_i x_i'$ konvergiert gegen eine positiv definite nichtsinguläre Matrix Σ_{xx}.

A 7 $E\{x_i \varepsilon_i\} = 0$

A 8 x_t und ε_t sind für gegebenes t statistisch unabhängig

A 9 $V\{\varepsilon \mid X\} = \sigma^2 \, \text{Diag}\{h_i^2\} = \sigma^2 \Psi$

A 10 $E\{\varepsilon \mid X\} = 0$

A 11 $\varepsilon_t \sim IID(0, \sigma^2)$

A 12 ε_t ist über die Zeit unkorreliert, mit Erwartungswert 0.

III. Das Lineare Regressionsmodell

Lösung für β:

$$b = \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \sum_{i=1}^{N} x_i y_i \quad \text{wenn } \sum_{i=1}^{N} x_i x_i' \text{ invertierbar}$$

ist bzw.

$b = (X'X)^{-1} X'y$ wenn $X'X$ invertierbar ist

Lösung für b_2 wenn $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$:

$$b_2 = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

Varianz des KQ Schätzers:

$$V\{b \mid X\} = \sigma^2 (X'X)^{-1} = \sigma^2 \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1}$$

Unverzerrter Schätzer für σ^2:

$$s^2 = \frac{1}{N-K}\sum_{i=1}^{N} e_i^2$$

IV. Maximum Likelihood

Likelihood Funktion im Modell mit einer binären abhängigen Variable:

$$L(\beta) = \prod_{i=1}^{N} P\{y_i = 1 | x_i; \beta\}^{y_i} \; P\{y_i = 0 | x_i; \beta\}^{1-y_i}$$

Log-Likelihood Funktion im Modell mit einer binären abhängigen Variable:

$$\log L(\beta) = \sum_{i=1}^{N} y_i \log F(x_i'\beta) + \sum_{i=1}^{N} (1-y_i)\log(1-F(x_i'\beta))$$

Marginale Effekte im Probit und Logit Modell:

Probit:

$$\frac{\partial \Phi(x_i'\beta)}{\partial x_{ik}} = \phi(x_i'\beta)\cdot\beta_k$$

Logit:

$$\frac{\partial \Lambda(x_i'\beta)}{\partial x_{ik}} = \frac{\exp(x_i'\beta)}{(1+\exp(x_i'\beta))^2}\cdot\beta_k$$

V. Gütemaße

$$R^2 = \frac{\sum_{i=1}^{N}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^{N} e_i^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}$$

Angepasstes R^2

$$\bar{R}^2 = 1 - \frac{\sum_{i=1}^{N} e_i^2 \,/(N-K)}{\sum_{i=1}^{N}(y_i - \bar{y})^2 \,/(N-1)}$$

AIC

$$AIC = \log\frac{1}{N}\sum_{i=1}^{N} e_i^2 + \frac{2K}{N}$$

BIC

$$BIC = \log\frac{1}{N}\sum_{i=1}^{N} e_i^2 + \frac{K}{N}\log N$$

pseudo R^2

$$\text{pseudo}R^2 = 1 - \frac{1}{1+2(\log L_1 - \log L_0)/N}$$

Mc Fadden R^2

$$\text{McFadden } R^2 = 1 - (\log L_1 / \log L_0)$$

VI. Tests

Kritischer Wert bei einem einseitigen Test

$$P\{t_k > t_{N-K;\alpha}\} = \alpha$$

Konfidenzintervall zum Niveau 1 - α

$$b_k - t_{N-K;\frac{\alpha}{2}}\cdot se(b_k) < \beta_k < b_k + t_{N-K;\frac{\alpha}{2}}\cdot se(b_k)$$

Teststatistik für einen F-Test auf gemeinsame Signifikanz

$$F = \frac{(S_0 - S_1)/J}{S_1/(N-K)} \sim F_{J,N-K}$$

$$F = \frac{(R_1^2 - R_0^2)/J}{(1-R_1^2)/(N-K)}$$

Teststatistik Goldfeld-Quandt-Test

$$\lambda = \frac{s_A^2}{s_B^2} \sim F_{N_A-K,N_B-K}$$

Teststatistik Durbin-Watson-Test

$$dw = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}$$

Teststatistik Wald-Test

$$\xi_W = (R\hat{\theta} - q)'[R\hat{V}R']^{-1}(R\hat{\theta} - q) \sim \chi_J^2$$

Teststatistik Likelihood-Ratio-Test

$$\xi_{LR} = -2\left[\log L(\tilde{\theta}) - \log L(\hat{\theta})\right] \sim \chi_J^2$$