

Masterprüfung SS 2017

Fach: Ökonometrie

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

Vorbemerkungen:

Anzahl der Aufgaben: Die Klausur besteht aus 5 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 beigelegt)
- Tabellen der statistischen Verteilungen (sind der Klausur beigelegt)
- 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:**[7 Punkte]**

In Musterstadt gibt es seit einem Monat drei neue Autohändler. Alle haben zur Eröffnung Werbeplakate aufgehängt. Der erste hat ein Plakat aufgehängt und hat seitdem 5 Autos verkauft. Der zweite hat ebenfalls ein Plakat aufgehängt und hat 4 Autos verkauft. Der dritte hat zwei Plakate aufgehängt und hat 8 Autos verkauft. Sie regressieren die Anzahl verkaufter Autos auf die Anzahl der Plakate mit folgendem Modell:

$$\text{Verkaufte_Autos}_i = \beta_1 + \beta_2 \cdot \text{Anzahl_Plakate}_i + \varepsilon_i$$

1.1 Berechnen Sie in Matrixschreibweise mithilfe der KQ-Methode die geschätzten Koeffizienten b_1 und b_2 . [6 Punkte]

1.2 Interpretieren Sie den geschätzten Koeffizienten b_2 inhaltlich. [1 Punkt]

Hinweis: Sollten Sie in Aufgabe 1.1 keine Lösung gefunden haben, verwenden Sie $b_2 = 2$.

Aufgabe 2:**[20 Punkte]**

Eine Krankenkasse beauftragt Sie, die Determinanten der Häufigkeit von Arztbesuchen ihrer Versicherten zu analysieren. Ihnen stehen dazu Daten aus einer Befragung von 3670 Versicherten im Jahre 2016 mit folgenden Variablen zur Verfügung:

- $\ln_dvisits_i$ Anzahl der Arztbesuche von Person i im Jahr 2016 (logarithmiert)
- age_i Alter der Person i in Jahren
- $female_i$ Dummy-Variable, =1, wenn Person i weiblich, =0 sonst
- $hstatus15_i$ Selbsteinschätzung des Gesundheitszustandes von Person i für das Jahr 2015 (Skala von 1-5)
- $hstatus16_i$ Selbsteinschätzung des Gesundheitszustandes von Person i für das Jahr 2016 (Skala von 1-5)
- $eduhigh_i$ Dummy-Variable, =1, wenn Person i eine hohe Bildung (mind. 13 Bildungsjahre) hat, =0 sonst

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

$$\ln_dvisits_i = \beta_1 + \beta_2 age_i + \beta_3 female_i + \beta_4 age_female_i + \beta_5 hstatus16_i + \beta_6 eduhigh_i + \varepsilon_i \quad (\text{Modell I})$$

wobei $age_female_i = age_i \cdot female_i$ gilt.

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

$\ln_dvisits$	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
hstatus16	.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.

2.1 Interpretieren Sie den geschätzten Koeffizienten b_6 inhaltlich und statistisch. [2 Punkte]

2.2 Leiten Sie den marginalen Effekt der Variable age auf die Anzahl an Arztbesuchen allgemein her. Wie groß ist der marginale Effekt für Frauen? Wie groß ist er für Männer? Interpretieren Sie die Ergebnisse. [4 Punkte]

- 2.3 Berechnen Sie die erwartete Anzahl an Arztbesuchen eines 30-jährigen Mannes mit hoher Bildung, der seinen Gesundheitszustand 2016 mit 4 bewertet hat. [4 Punkte]
- 2.4 Sie vermuten, dass sich gleichzeitig mit dem Effekt von *age* auch der Effekt von *hstatus16* 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]
- 2.5 Ein Kommilitone behauptet, dass Ihre Variable *hstatus16* endogen ist. Nennen Sie einen möglichen Grund für das Vorliegen der Endogenität und erläutern Sie diesen. [1,5 Punkte]
- 2.6 Ihr Kommilitone rät Ihnen, Ihr Endogenitätsproblem durch Instrumentierung von *hstatus16* zu lösen. Er schlägt vor, die Selbsteinschätzung des Gesundheitsstatus aus dem Vorjahr (*hstatus15*) als Instrument in einer two-stage-least-squares (2SLS)-Schätzung Ihres Modells I zu verwenden.
- 2.6.1 Stellen Sie die für die 2SLS-Schätzung benötigten Modellgleichungen auf. Nennen Sie einen Fall, in dem *hstatus15* kein geeignetes Instrument darstellt. [3 Punkte]
- 2.6.2 Wie können Sie testen, ob es sich bei *hstatus15* um ein schwaches Instrument handelt? Geben Sie die Nullhypothese des Tests an. [1,5 Punkte]

Aufgabe 3:

[20 Punkte]

Sie interessieren sich für die Determinanten der am Saisonende erreichten Punktzahl des 1. FC Nürnberg. Dazu liegt ein Datensatz mit Informationen von 1996 bis 2017 vor:

- points_t* Am Saisonende erreichte Punktzahl des 1. FC Nürnberg im Jahr *t*
value_t Marktwert aller Spieler des 1. FC Nürnberg in Millionen € im Jahr *t*
possession_t Durchschnittlicher Ballbesitz in Prozent (1-100) pro Spiel im Jahr *t*
coach_t Dummy-Variable, =1, wenn es im Jahr *t* zu einem Trainerwechsel beim 1. FC Nürnberg kam, =0 sonst
bundesliga_t Dummy-Variable, =1, wenn der 1. FC Nürnberg im Jahr *t* in der 1. Bundesliga gespielt hat, =0 sonst

Es wird folgendes Regressionsmodell aufgestellt und anschließend mit Stata geschätzt:

$$points_t = \beta_1 + \beta_2 value_t + \beta_3 possession_t + \beta_4 coach_t + \beta_5 bundesliga_t + \varepsilon_t$$

Source	SS	df	MS	Number of obs = 22		
Model	688.158228	4	172.039557	F(4, 17)	=	18.85
Residual	465.556058	17	27.3856504	Prob > F	=	0.0000
Total	1153.71429	21	54.9387757	R-squared	=	0.5965
				Adj R-squared	=	0.5648
				Root MSE	=	5.2331

points	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
value	.3790518	.0721279	5.26	0.000	.2376325	.5204711
possession	2.487062	.4710640	5.28	0.000	1.562692	3.411432
coach	-8.744706	2.404095	-3.64	0.000	-13.46227	-4.027140
bundesliga	-20.37705	7.634737	-2.67	0.008	-35.34628	-5.407819
_cons	-74.64235	20.84982	-3.58	0.000	-116.3419	-32.94271

Runden Sie im Folgenden alle Zahlenangaben auf die dritte Nachkommastelle.

- 3.1 Interpretieren Sie den geschätzten Koeffizienten b_3 inhaltlich und statistisch. [2 Punkte]
- 3.2 Begründen Sie, warum der geschätzte Koeffizient b_4 eventuell nicht kausal interpretiert werden kann. Erläutern Sie den Mechanismus anhand eines Beispiels. [2 Punkte]
- 3.3 Der folgende Stata-Output enthält die Ergebnisse eines White-Tests auf Heteroskedastie. Führen Sie den Test auf einem Signifikanzniveau von 10% durch. Geben Sie die Hypothesen und Hilfsregression an. Definieren Sie die abhängige Variable der Hilfsregression. Bestimmen Sie zudem die Teststatistik, Freiheitsgrade, kritischen Wert und Testentscheidung. [7 Punkte]

```
White's test for Ho: ???
against Ha: ???

chi2(?)      =    12,84
Prob > chi2  =    ??????
```

- 3.4 Warum könnte in dem vorliegenden Modell positive Autokorrelation im Fehlerterm vorliegen? Welche Folgen hat unkorrigierte Autokorrelation für die geschätzten Parameter und deren Standardfehler? [2,5 Punkte]
- 3.5 Führen Sie einen Breusch-Godfrey-Test auf Autokorrelation 1. Ordnung auf einem Signifikanzniveau von 5% durch. Geben Sie die Hypothesen, die Hilfsregression, die Teststatistik, die Freiheitsgrade, den kritischen Wert und die Testentscheidung an. Definieren Sie dabei formal Autokorrelation 1. Ordnung. *Hinweis: Das Bestimmtheitsmaß aus der Hilfsregression lautet: $R^2 = 0,2$.* [6,5 Punkte]

Aufgabe 4:

[13 Punkte]

Sie interessieren sich für die Determinanten des Rauchverhaltens von Amerikanern. Hierfür steht Ihnen ein Datensatz mit 807 Befragten zur Verfügung. Er enthält folgende Variablen:

- $smoker_i$ Dummy-Variable, =1, wenn Person i raucht, =0 sonst
- $educ_i$ Bildung von Person i in Jahren
- age_i Alter von Person i in Jahren
- inc_i Jahreseinkommen von Person i in Dollar

Ihnen sind folgende deskriptive Statistiken zu den Daten gegeben:

Variable	Obs	Mean	Std. Dev.	Min	Max
smoker	807	.3841388	.4866926	0	1
educ	807	12.47088	3.057161	6	18
age	807	41.23792	17.02729	17	88
income	807	19970.58	9498.139	500	32700

Sie führen mit Stata eine Logit-Schätzung des folgenden Modells durch:

$$P(smoker = 1|X) = F(\beta_1 + \beta_2 educ_i + \beta_3 age_i + \beta_4 age_i^2)$$

```
Logistic regression                               Number of obs =      807
                                                  LR chi2(3)      =      47.57
                                                  Prob > chi2     =      0.0000
Log likelihood = -513.71831                    Pseudo R2      =      0.0443
```

smoker	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
educ	-.1350465	.0268896	-5.02	0.000	-.1877492 -.0823438
age	.1060429	.0269834	3.93	0.000	.0531563 .1589295
age^2	-.0013844	.0003073	-4.51	0.000	-.0019867 -.0007822
_cons	-.4479964	.5725251	-0.78	0.434	-1.570125 .6741321

Runden Sie im Folgenden alle Zahlenangaben auf die dritte Nachkommastelle.

- 4.1 Wie hoch ist die durchschnittliche vorhergesagte Wahrscheinlichkeit zu rauchen? [1 Punkt]
- 4.2 Berechnen und interpretieren Sie den marginalen Effekt eines zusätzlichen Bildungsjahres am Mittelwert der Daten. [6 Punkte]
- 4.3 Eine zweite Logit-Schätzung mit dem gleichen Datensatz, bei der zusätzlich zu den bisherigen Variablen noch das Einkommen berücksichtigt wird, erzielt einen Log-Likelihoodwert von $-510,272$. Prüfen Sie, ob das Einkommen einen signifikanten Effekt auf die Wahrscheinlichkeit zu rauchen hat. Führen Sie dazu einen Likelihood-Ratio-Test am 1%-Signifikanzniveau durch. Geben Sie zuerst das restringierte Modell und das unrestringierte Modell an. Benennen Sie dann Hypothesen und kritischen Wert. Treffen Sie eine Testentscheidung. [6 Punkte]

Aufgabe 5 - MC Fragen

[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.	Im Modell $y_t = \beta_1 + \beta_2 x_t + \beta_3 y_{t-1} + \varepsilon_t$ führt Autokorrelation 1. Ordnung in ε
a	zu unobserved heterogeneity (unbeobachtete Heterogenität).
b	zu Endogenität von y_{t-1} .
c	zum attenuation bias.
d	zu einer Verzerrung von β_2 .

2.	Ein RESET-Test überprüft
a	mit Hilfe eines t-Tests, ob Fehlspezifikation des Regressionsmodells vorliegt.
b	, ob durch Änderung der Modellspezifikation das angepasste Bestimmtheitsmaß \bar{R}^2 signifikant steigt.
c	auf Basis eines F-Tests, ob ein Problem ausgelassener Variablen vorliegt.
d	ob sich die marginalen Effekte im Modell für Teilgruppen unterscheiden.

3.	Das Modell $Lohn_i = \beta_1 + \beta_2 uni_i + \beta_3 west_i + \beta_4 uni_i \cdot west_i + \varepsilon_i$ wird mittels KQ geschätzt ($uni_i = 1$, falls Uni-Abschluss, $west_i = 1$ falls Westdeutscher). Welche Aussage trifft für die Schätzung des Modells zu?
a	Der marginale Effekt von uni_i für Ostdeutsche lautet: $b_2 + b_4$.
b	Der Lohnunterschied zwischen west- und ostdeutschen Uni-Absolventen lautet: $b_3 + b_4$.
c	Der vorhergesagte Lohn für Westdeutsche mit Uni-Abschluss lautet: $b_1 + b_3 + b_4$.
d	Der Lohnunterschied zwischen westdeutschen Uni-Absolventen und westdeutschen Nicht-Absolventen lautet: b_2 .

4.	Sie schätzen das Modell $y_i = \beta_1^* + \beta_2^* x_i + \varepsilon_i^*$ und Ihr Kommilitone das Modell $y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + \varepsilon_i$. Sei $cov(x, z) < 0$. Dann gilt:
a	$E[b_2^*] < \beta_2$, wenn $\beta_3 < 0$.
b	$E[b_2] < E[b_2^*]$, wenn $\beta_3 < 0$.
c	$E[b_2^*] - \beta_2 > 0$, wenn $\beta_3 > 0$.
d	$E[b_2] \neq E[b_2^*]$, wenn $\beta_3 = 0$.

5.	Eine KQ-Schätzung des Modells $y_i = \beta_1 + \beta_2 x_i + \beta_3 x_i^2 + \varepsilon_i$ ergibt $b_1 = 5$, $b_2 = -0,5$ und $b_3 = 0,025$. Bei welchem Wert von \hat{y} ist der marginale Effekt von x gleich null?
a	5
b	2,5
c	10
d	0,25

6.	Für das Modell $y_i = \beta_1 + \beta_2 x_{1i} + \beta_3 x_{2i} + u_i$ testen Sie $H_0 : \beta_2 = -2$ vs. $H_1 : \beta_2 \neq -2$. Auf einem Signifikanzniveau von 5% führt eine t-Teststatistik von $-2,16$
a	bei $n = 16$ nicht zur Ablehnung von H_0 .
b	bei $n = 15$ zur Ablehnung von H_0 .
c	bei $n = 17$ nicht zur Ablehnung von H_0 .
d	bei $n = 14$ zur Ablehnung von H_0 .

7.	Die Schätzung des Modells $\text{zufriedenheit}_i = \beta_1 + \beta_2 \text{freizeit}_i + \varepsilon_i$ mit KQ ergibt bei 150 Beobachtungen $b_2 = 2,5$, $se(b_2) = 0,5$. Welche Aussage trifft für die Schätzung von β_2 zu?
a	Die untere Grenze des 95 % Konfidenzintervalls für β_2 beträgt 3,48.
b	Der kritische Wert für einen t-Test auf Signifikanz mit $\alpha = 0,01$ beträgt 2,326.
c	Der p-Wert für den beidseitigen t-Test der $H_0: \beta_2 = 1,25$ ist größer als 0,01.
d	Die Nullhypothese $\beta_2 \leq 0$ eines t-Tests wird nicht verworfen, wenn der kritische Wert 1,282 beträgt.

8.	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 $b_2 < 0$, $b_3 > 0$, $b_4 < 0$ und $b_5 < 0$ ist
a	der marginale Effekt von x_2 für große x_2 -Werte negativ und betragsmäßig zunehmend.
b	der marginale Effekt von x_1 für kleine x_1 -Werte negativ und betragsmäßig zunehmend.
c	der marginale Effekt von x_2 für kleine x_2 -Werte negativ und betragsmäßig abnehmend.
d	der marginale Effekt von x_1 für große x_1 -Werte positiv und betragsmäßig abnehmend.

9.	Welche Aussage bezüglich des Durbin-Wu-Hausman Test ist richtig?
a	Die Nullhypothese besagt, dass die instrumentierte Variable endogen ist.
b	Bei Ablehnung der Nullhypothese liegt Endogenität vor.
c	In der Hilfsregression wird die Instrumentenvariable auf die geschätzten Residuen aus der 1. Stufe der 2SLS-Schätzung regressiert.
d	Die Anzahl der Zählerfreiheitsgrade des kritischen Wertes entspricht der Anzahl der Instrumente.

10.	Welche Aussage bezüglich des Bestimmtheitsmaßes R^2 im multiplen Regressionsmodell bei N Beobachtungen ist richtig?
a	Im Modell ohne Konstante gilt $R^2 = 0$.
b	Wenn $\sum_{i=1}^n e_i = 0$ gilt, gilt $R^2 = 1$.
c	Wenn $R^2 = 0,49$ gilt, gilt für das angepasste Bestimmtheitsmaß $\bar{R}^2 = 1 - 0,51 \cdot \frac{N-1}{N-K}$.
d	Im Modell mit Konstante und $K > 1$ gilt $0 \leq R^2 < 0,5$.

11.	Ein Konfidenzintervall
a	wird mit c.p. steigender Präzision schmaler.
b	wird mit c.p. steigender Stichprobengröße schmaler.
c	wird mit c.p. sinkendem Signifikanzniveau α schmaler.
d	hat stets negative Intervallgrenzen, wenn $b_j < 0$.

12.	Sei $V(\varepsilon) = \sigma^2 \Psi$ die Varianz-Kovarianz-Matrix der Störterme. Liegt Heteroskedastie und Autokorrelation vor, ist die Matrix Ψ
a	eine untere Dreiecksmatrix.
b	eine obere Dreiecksmatrix.
c	eine symmetrische Matrix.
d	eine Diagonalmatrix.

13.	Der Breusch-Pagan Test
a	kann die falsche H_0 mit höherer Wahrscheinlichkeit verwerfen als der White-Test.
b	ist allgemeiner als der White-Test.
c	prüft in der Hilfsregression, ob e^2 durch die ursprünglichen Regressoren und deren Quadrate erklärt werden kann.
d	hat eine höhere Anzahl an Freiheitsgraden J als der White-Test.

14.	Unkorrigierte Heteroskedastie im linearen Regressionsmodell führt zu
a	Verzerrung des KQ-Schätzers.
b	Inkonsistenz des KQ-Schätzers.
c	Effizienz des KQ-Schätzers.
d	fehlerhaft geschätzten Konfidenzintervallen.

15.	Es sei $V(\epsilon_i) = \sigma^2 \cdot \text{Einkommen}_i^4$. Zu konstanter Störtermvarianz führt die Transformation
a	$\epsilon_i / \text{Einkommen}_i^{16}$.
b	$\epsilon_i / \text{Einkommen}_i^8$.
c	$\epsilon_i \cdot \text{Einkommen}_i^4$.
d	$\epsilon_i / \text{Einkommen}_i^2$.

16.	Der Cochrane-Orcutt Schätzer
a	korrigiert nicht für negative Autokorrelation im Störterm.
b	korrigiert für Autokorrelation höherer als 1. Ordnung.
c	nutzt die Beobachtung für $t = 1$ nicht.
d	liefert immer die gleichen Standardfehler wie der Prais-Winston Schätzer.

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

18.	Die obere und untere Grenze für den kritischen Wert eines Durbin-Watson Tests auf positive Autokorrelation für ein lineares Modell mit 100 beobachteten Perioden, einer Konstante und 2 unabhängigen Variablen beträgt auf einem Signifikanzniveau von 5%
a	$d_L = 1,63$ und $d_W = 1,72$.
b	$d_L = 1,65$ und $d_W = 1,69$.
c	$d_L = 1,62$ und $d_W = 1,71$.
d	$d_L = 1,61$ und $d_W = 1,74$.

19.	Wenn keine Autokorrelation 1. Ordnung in den Residuen vorliegt, ist der Wert der Durbin-Watson Teststatistik nahe
a	-2.
b	0
c	2.
d	4.

20.	Wenn unter den erklärenden Variablen eines linearen Regressionsmodells perfekte Multikollinearität vorliegt, gilt für die Matrix $X'X$, dass sie
a	nicht den vollen Rang hat.
b	regulär ist.
c	invertierbar ist.
d	eine Determinante ungleich Null hat.

21.	Wird ein Modell einmal mit Probit und einmal mit Logit geschätzt
a	sind die marginalen Effekte identisch.
b	sind die geschätzten Koeffizienten identisch.
c	sind beide Schätzungen effizient.
d	ist mindestens eine der beiden Schätzungen nicht konsistent.

22.	Wenn eine KQ-Schätzung nicht konsistent ist, dann
a	sind die Koeffizienten immer verzerrt.
b	enthalten die Koeffizienten keine validen Informationen über Korrelationen zwischen den Variablen.
c	ist der Erwartungswert des Residuums trotzdem 0.
d	ist das angepasste R^2 kleiner als 0.

23.	Die Matrix $X'X$
a	ist immer invertierbar.
b	ist immer quadratisch.
c	hat immer einen niedrigeren Rang als die Matrix X .
d	hat immer einen höheren Rang als die Matrix X .

24.	Ein Unterschied zwischen Wald-Test und Likelihood Ratio-Test ist, dass der Wald-Test
a	weniger Freiheitsgrade hat.
b	mehr Restriktionen auf einmal testen kann.
c	präziser ist.
d	nur eine Schätzung erfordert.

25.	Der marginale Effekt in einem Probit-Modell mit nur einer unabhängigen Variable
a	ist am größten am Mittelwert der unabhängigen Variable.
b	ist am größten bei Beobachtungen mit $x = 0$.
c	kann immer als Semi-Elastizität interpretiert werden.
d	ist betragsmäßig immer kleiner als der Koeffizient.

26.	Wenn das Pseudo- R^2 einer Logit-Schätzung 0,8 beträgt, die Log-Likelihood des unrestringierten Modells -100 und die des restringierten Modells -120, wie hoch ist dann die Anzahl der Beobachtungen?
a	5
b	10
c	40
d	100

27.	Bei unabhängigen Beobachtungen entspricht die Likelihoodfunktion
a	dem Produkt der individuellen Wahrscheinlichkeitsdichten der beobachteten abhängigen Variablen.
b	der Summe der quadrierten individuellen Wahrscheinlichkeitsdichten der beobachteten abhängigen Variablen.
c	der Summe der individuellen Wahrscheinlichkeitsdichten der beobachteten abhängigen Variablen.
d	dem Logarithmus der individuellen Wahrscheinlichkeitsdichten der beobachteten abhängigen Variablen.

28.	Ein Messfehler in x führt bei einer KQ-Schätzung des Modells $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$
a	zum Problem des attenuation bias.
b	zur Überschätzung von β_2 .
c	zu endogenen Störtermen.
d	zu einer Unterschätzung der Störtermvarianz.

29.	Was ist eine Eigenschaft des KQ-Schätzers in einem linearen Modell mit binärer abhängiger Variable?
a	Der Fehlerterm ist homoskedastisch.
b	Bei großer Stichprobengröße ist er effizienter als Probit.
c	Bei gewöhnlich berechneten Standardfehlern sind t-Tests auf Signifikanz der Kovariate ungültig.
d	Die Summe der errechneten Wahrscheinlichkeiten ist gleich $\frac{N}{2} \cdot \bar{y}$.

30.	Die Verletzung welcher Annahme schließt die kausale Interpretation einer Schätzung aus? (s. Formelsammlung)
a	A2
b	A3
c	A7
d	A11

Tabelle 2: Perzentile der t-Verteilung

Zelleneintrag: x , sodass $\text{Prob}[t_n \leq x] = P$, mit n Freiheitsgraden

P n	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 3: Perzentile der χ^2 -Verteilung

Zelleneintrag: c, sodass $\text{Prob}[\chi_n^2 \leq c] = P$, mit n Freiheitsgraden

n \ P	0.005	0.01	0.025	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.975	0.99	0.995
1	0.00004	0.0002	0.001	0.004	0.016	0.102	0.455	1.323	2.706	3.842	5.024	6.635	7.879
2	0.01	0.02	0.05	0.10	0.21	0.58	1.39	2.77	4.61	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	0.58	1.21	2.37	4.11	6.25	7.81	9.35	11.34	12.84
4	0.21	0.30	0.48	0.71	1.06	1.92	3.36	5.39	7.78	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	1.61	2.67	4.35	6.63	9.24	11.07	12.83	15.09	16.75
6	0.68	0.87	1.24	1.64	2.20	3.45	5.35	7.84	10.64	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.22	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.39	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.55	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	7.58	10.34	13.70	17.28	19.68	21.92	24.73	26.76
12	3.07	3.57	4.40	5.23	6.30	8.44	11.34	14.85	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	9.30	12.34	15.98	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	10.17	13.34	17.12	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	11.04	14.34	18.25	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	11.91	15.34	19.37	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	12.79	16.34	20.49	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.86	13.68	17.34	21.60	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	14.56	18.34	22.72	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	15.45	19.34	23.83	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	16.34	20.34	24.93	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	17.24	21.34	26.04	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	18.14	22.34	27.14	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	19.04	23.34	28.24	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	19.94	24.34	29.34	34.38	37.65	40.65	44.31	46.93
30	13.79	14.95	16.79	18.49	20.60	24.48	29.34	34.80	40.26	43.77	46.98	50.89	53.67
35	17.19	18.51	20.57	22.47	24.80	29.05	34.34	40.22	46.06	49.80	53.20	57.34	60.27
40	20.71	22.16	24.43	26.51	29.05	33.66	39.34	45.62	51.81	55.76	59.34	63.69	66.77
45	24.31	25.90	28.37	30.61	33.35	38.29	44.34	50.98	57.51	61.66	65.41	69.96	73.17
50	27.99	29.71	32.36	34.76	37.69	42.94	49.33	56.33	63.17	67.50	71.42	76.15	79.49

Quelle: In R generiert

Tabelle 4b: 99% Perzentile der F-Verteilung

Zelleneintrag: f, sodass $\text{Prob}[F_{n_1, n_2} \leq f] = 0.99$

n2 \ n1 = Freiheitsgrade des Zählers	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	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

Tabelle 5: Durbin-Watson Teststatistik

d_L und d_U am 5% Signifikanzniveau

n	k=2		k=3		k=4		k=5		k=6		k=11		k=16	
	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21				
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15	0.16	3.30		
17	1.13	1.38	1.02	1.53	0.90	1.71	0.78	1.90	0.67	2.21	0.20	3.18		
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.15	0.24	3.07		
19	1.18	1.40	1.08	1.54	0.97	1.68	0.86	1.85	0.75	2.10	0.29	2.97		
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	2.06	0.34	2.89	0.06	3.68
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	2.02	0.38	2.81	0.09	3.58
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.99	0.42	2.73	0.12	3.55
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.96	0.47	2.67	0.15	3.41
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.64	0.51	2.61	0.19	3.33
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.92	0.54	2.57	0.22	3.25
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.90	0.58	2.51	0.26	3.18
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.89	0.62	2.47	0.29	3.11
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.88	0.65	2.43	0.33	3.05
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.86	0.68	2.40	0.36	2.99
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.85	0.71	2.36	0.39	2.94
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.84	0.74	2.33	0.43	2.99
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.83	0.77	2.31	0.46	2.84
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.83	0.80	2.28	0.49	2.80
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.82	0.82	2.26	0.52	2.75
35	1.40	1.52	1.34	1.53	1.28	1.65	1.22	1.73	1.16	1.81	0.85	2.24	0.55	2.72
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.81	0.87	2.22	0.58	2.68
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80	0.89	2.20	0.60	2.65
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.80	0.91	2.18	0.63	2.61
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.80	0.93	2.16	0.65	2.59
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79	0.95	2.15	0.68	2.56
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.79	1.04	2.09	0.79	2.44
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.79	1.11	2.04	0.88	2.35
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.78	1.17	2.01	0.96	2.28
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77	1.22	1.98	1.03	2.23
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77	1.27	1.96	1.09	2.18
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77	1.30	1.95	1.14	2.15
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77	1.34	1.94	1.18	2.12
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77	1.37	1.93	1.22	2.09
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77	1.40	1.92	1.26	2.07
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78	1.42	1.91	1.29	2.06
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78	1.44	1.90	1.32	2.04
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78	1.46	1.90	1.35	2.03
150	1.72	1.75	1.72	1.76	1.71	1.76	1.69	1.77	1.68	1.79	1.61	1.86	1.54	1.52
200	1.76	1.78	1.76	1.78	1.75	1.79	1.74	1.80	1.73	1.81	1.68	1.86	1.62	1.61

Anmerkung: k ist die Anzahl der Regressoren mit Konstante.

Quelle: Johnston, J. and J. DiNardo, 1997, *Econometric Methods*, 4th ed., New York: McGraw-Hill

Formeln Ökonometrie

I. Mathematische Grundlagen

i. Algebra

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

$$(A')' = A$$

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

$$(A^{-1})' = (A')^{-1}$$

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

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:

$$\begin{aligned} \sigma^2 &= V\{Y\} = E\{(Y - E(Y))^2\} \\ &= E\{Y^2\} - E\{Y\}^2 \end{aligned}$$

Kovarianz:

$$\begin{aligned} \sigma_{YX} &= \text{cov}\{Y, X\} = E\{(Y - E(Y))(X - E(X))\} \\ &= E\{XY\} - E\{X\}E\{Y\} \end{aligned}$$

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 $\text{cov}\{Y, X\} = 0$

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 $E\{\varepsilon_i\} = 0 \quad i = 1, 2, \dots, N$

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

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

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

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

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

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 $V\{\varepsilon | X\} = \sigma^2 \text{Diag}\{h_i^2\} = \sigma^2 \Psi$

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

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

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.

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

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:

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

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 $1 - \alpha$

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

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

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

Teststatistik Wald-Test

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

Teststatistik Likelihood-Ratio-Test

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

VII. IV Schätzer

$$\hat{\beta}_{IV} = \left(\sum_{i=1}^N z_i x_i' \right)^{-1} \sum_{i=1}^N z_i y_i$$

$$\sqrt{N}(\hat{\beta}_{IV} - \beta) \rightarrow N(0, \sigma^2 (\sum_{xz} \sum_{zz}^{-1} \sum_{zx})^{-1})$$