

Prüfung im Fach Mikroökometrie im Wintersemester 2011/12 Aufgaben

Vorbemerkungen:

- Anzahl der Aufgaben:**
- Die Klausur besteht aus 5 Aufgaben.
- Bewertung:**
- Es können maximal 60 Punkte erworben werden. Die Punktzahl ist für jede Aufgabe in Klammern angegeben. Sie entspricht der für die Aufgabe empfohlenen Bearbeitungszeit in Minuten.
- Erlaubte Hilfsmittel:**
- Tabellen der statistischen Verteilungen (sind der Klausur beigelegt)
 - Taschenrechner
 - Fremdwörterbuch
- Wichtige Hinweise:**
- Sollte es vorkommen, dass die statistischen Tabellen, die dieser Klausur beiliegen, den exakten Wert der gesuchten Freiheitsgrade nicht ausweisen, machen Sie dies kenntlich und verwenden Sie den nächstgelegenen Wert.
 - Sollte es vorkommen, dass bei einer Berechnung eine erforderliche Annahme oder Angabe fehlt, machen Sie dies kenntlich und treffen Sie für den fehlenden Wert eine plausible Annahme.

Aufgabe 1 (15 Punkte)

Ihnen stehen folgende Informationen zu 5574 deutschen Rentnern zur Verfügung:

- health = subjektive Einschätzung des Gesundheitszustandes (1 = schlecht, 2 = mittel, 3 = gut)
- age = Alter gemessen in Jahren
- income = Jahreseinkommen gemessen in 1000 Euro
- ndisease = Anzahl chronischer Krankheiten

Die Schätzung eines Ordered Logit Modells mit *health* als abhängiger Variable ergibt folgenden Output:

Ordered logistic regression	Number of obs	=	5574
	LR chi2(3)	=	758.59
	Prob > chi2	=	0.0000
Log likelihood = -4760.7527	Pseudo R2	=	0.0738

health	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
age	-.0299348	.001691	-17.70	0.000	-.033249 -.0266206
income	.089344	.0069124	12.93	0.000	.075796 .102892
ndisease	-.0550935	.0040768	-13.51	0.000	-.0630838 -.0471032
/cut1	-3.175896	.0930344			-3.35824 -2.993552
/cut2	-.8295084	.078223			-.9828226 -.6761941

- 1.1 Interpretieren Sie den Koeffizienten von *ndisease* statistisch und inhaltlich. (2 Punkte)
- 1.2 Erläutern Sie formal den Zusammenhang zwischen dem beobachteten Gesundheitszustand y_i und dem latenten Gesundheitszustand y_i^* . Definieren Sie die verwendeten Größen. (3 Punkte)
- 1.3 Leiten Sie im Rahmen des geordneten Logit Modells den Likelihood-Beitrag für einen Rentner ab, der angibt, bei schlechter Gesundheit zu sein. *Hinweis*: Die logistische Verteilungsfunktion lautet $\Lambda(z) = 1/(1 + \exp(-z))$. (3 Punkte)
- 1.4 Könnten Sie im vorliegenden Fall auch ein multinomiales Logit Modell schätzen? Gehen Sie auf einen Vorteil und einen Nachteil im Vergleich zum Ordered Logit Modell ein. (4 Punkte)
- 1.5 Bei einem Rentner wird eine chronische Krankheit diagnostiziert. Um wie viel müsste das Jahreseinkommen des Rentners steigen, um ihn für die damit verbundene Verschlechterung des Gesundheitszustandes zu kompensieren? (3 Punkte)

Aufgabe 2 (17 Punkte)

Sie interessieren sich für die Determinanten der Studienfachwahl. Ihnen stehen für 295 Abiturienten folgende Informationen zur Verfügung:

- studienfach = gewähltes Studienfach (1 = Medizin, 2 = Musik, 3 = BWL, 4 = VWL)
- abinote = individuelle Abiturnote des Abiturienten
- alquote = studienfachspezifische Arbeitslosenquote

- 2.1 Erklären Sie, warum das Mixed Logit Modell eine geeignete Modellwahl darstellt, wenn *abinote* als abhängige Variable und *alquote* und *studienfach* als erklärende Variablen in der Schätzung verwendet werden. (3 Punkte)
- 2.2 Erläutern Sie die IIA-Annahme und nennen Sie die Konsequenz einer Annahmeverletzung. (2 Punkte)

2.3 Warum könnte die IIA-Annahme im vorliegenden Fall verletzt sein? Nennen Sie eine Lösungsmöglichkeit. (3 Punkte)

2.4 Sie fassen die Studienfächer VWL und BWL zur Kategorie Wirtschaftswissenschaften (Wiwi) zusammen. Interpretieren Sie die Koeffizienten von *alquote* und *abinote* im folgenden Mixed Logit Modell inhaltlich. (4 Punkte)

```

Alternative-specific conditional logit      Number of obs   =      885
Case variable: id                        Number of cases  =      295

Alternative variable: studienfach        Alts per case:  min =       3
                                           avg =       3.0
                                           max =       3

Log likelihood = -255.5512                Wald chi2(3)    =       7.19
                                           Prob > chi2     =      0.0661

```

choice	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

studienfach						
alquote	-.0436522	.0329333	-1.33	0.185	-.1082003 .020896	

Medizin	(base alternative)					

Musik						
abinote	-.6390918	.3105483	-2.06	0.040	-1.247755 -.0304283	
_cons	-.1604799	.4730256	-0.34	0.734	-1.087593 .7666333	

Wiwi						
abinote	.4242441	.4496724	0.94	0.345	-.4570977 1.305586	
_cons	-1.255797	.6455486	-1.95	0.052	-2.521049 .0094554	

2.5 Erläutern Sie die Vorgehensweise bei einem Lagrange Multiplier Tests. *Hinweis*: Die Formel für die Teststatistik muss nicht angegeben werden. (5 Punkte)

Aufgabe 3 (11.5 Punkte)

Die Determinanten von Bonuszahlungen wurden mit einem Regressionsmodell geschätzt. 41 der 616 in der Stichprobe enthaltenen Individuen haben keine Bonuszahlung erhalten (d. h. *bonus*= 0). Die Variablen sind wie folgt definiert:

- bonus* = Höhe der Bonuszahlung pro Arbeitsstunde in US-Dollar
- educ* = Schulausbildung in Jahren
- exper* = Berufserfahrung in Jahren
- expersq* = Berufserfahrung in Jahren (quadriert)
- ind1* = 1, falls Arbeitsplatz im primären Sektor; 0 sonst (Referenz: tertiärer Sektor)
- ind2* = 1, falls Arbeitsplatz im sekundären Sektor; 0 sonst (Referenz: tertiärer Sektor)

Ein Tobit-Modell liefert folgende Ergebnisse:

[Output siehe nächste Seite]

```

. tobit bonus educ exper expersq ind1 ind2, ll(0)

Tobit regression                               Number of obs   =       616
                                                LR chi2(5)      =         ?
                                                Prob > chi2     =       0.0000
Log likelihood = -595.28643                    Pseudo R2      =       0.1002

```

bonus	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
educ	.0664911	.0097017	6.85	0.000	.0474383 .0855439
exper	.062486	.0075171	8.31	0.000	.0477235 .0772486
expersq	-.0010172	.0001621	-6.27	0.000	-.0013356 -.0006987
ind1	.3803796	.4503483	0.84	0.399	-.5040387 1.264798
ind2	-.0471082	.0912512	-0.52	0.606	-.2263123 .1320958
_cons	-.5989004	.1429254	-4.19	0.000	-.879585 -.3182157
/sigma	.6322063	.0188922			.5951047 .6693079

Obs. summary: 41 left-censored observations at bonus<=0
 575 uncensored observations
 0 right-censored observations

- 3.1 Erläutern Sie den Unterschied zwischen Stützung und Zensierung. Welcher Fall liegt in der Aufgabenstellung vor? (3 Punkte)
- 3.2 Überprüfen Sie mit einem Wald-Test am 5%-Signifikanzniveau die Hypothese, dass die Höhe der Bonuszahlungen in mindestens einem Wirtschaftssektor (d.h. im primären und/oder sekundären Sektor) von der im tertiären Sektor verschieden ist. Geben Sie Null- und Alternativhypothesen, Freiheitsgrade, den kritischen Wert und die Testentscheidung an. *Hinweise:* (1) Die Teststatistik des Wald-Tests lautet: $W = \hat{\beta}' \hat{Var}(\hat{\beta})^{-1} \hat{\beta} \sim \chi_k^2$. (2) Verwenden Sie folgende geschätzte inverse Kovarianzmatrix des Schätzvektors $\hat{\beta} = (\hat{\beta}_{ind1}, \hat{\beta}_{ind2})'$. (4.5 Punkte)

$$\hat{Var}(\hat{\beta})^{-1} = \begin{pmatrix} 4.93 & 0 \\ 0 & 120.10 \end{pmatrix}$$

- 3.3 Führen Sie einen Likelihood-Ratio-Test auf Gesamtsignifikanz am 1%-Signifikanzniveau durch. Geben Sie Null- und Alternativhypothese, Teststatistik, Freiheitsgrade, Schlusslogik und Testentscheidung an. *Hinweis:* Der Log-Likelihood-Wert eines Modells, das nur eine Konstante enthält, beträgt -661.5941. (4 Punkte)

Aufgabe 4 (12.5 Punkte)

Eine Bank bewirbt bei 925 Kunden ein neues Finanzmarktprodukt. 470 Kunden reagieren auf die Werbemaßnahme und investieren den Betrag *invest*. Für eine Schätzung des Zusammenhangs von Investitionsbetrag und Kundencharakteristika stehen folgende Variablen zur Verfügung:

- female = 1, falls Frau, 0 sonst
- age = Alter in Jahren
- agesq = Alter in Jahren (quadriert)
- response = 1, falls Person auf Werbemaßnahme antwortet, 0 sonst
- activity = 1, falls Person in der Vergangenheit in ein Finanzprodukt investiert hat, 0 sonst.

```

Heckman selection model -- two-step estimates      Number of obs   =      925
(regression model with sample selection)         Censored obs    =      455
                                                Uncensored obs  =      470

                                                Wald chi2(3)    =      19.82
                                                Prob > chi2     =      0.0002

```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

invest						
female	-.0484285	.1737498	-0.28	0.780	-.3889717	.2921148
age	.081615	.0271189	3.01	0.003	.028463	.134767
agesq	-.0675975	.0263961	-2.56	0.010	-.1193329	-.015862
_cons	2.086416	.8419196	2.48	0.013	.4362843	3.736549

response						
female	.5881135	.0966845	6.08	0.000	.3986154	.7776116
age	.04168	.0215441	1.93	0.053	-.0005456	.0839057
agesq	-.0409823	.0206072	-1.99	0.047	-.0813717	-.0005928
activity	.5611673	.1115717	5.03	0.000	.3424908	.7798438
_cons	-1.497584	.5368219	-2.79	0.005	-2.549735	-.4454319

mills						
lambda	.4304413	.3210011	1.34	0.180	-.1987092	1.059592

- 4.1 Erläutern Sie unter Bezugnahmen auf die Analyse des Investitionsbetrags den Begriff der endogenen Stichprobenselektion. Welche Konsequenz folgt daraus für eine Kleinst-Quadrate-Schätzung? (3 Punkte)
- 4.2 Nennen Sie eine zentrale Annahme des Heckman-Schätzverfahrens. Welche Folge hat eine Annahmeverletzung? (2.5 Punkte)
- 4.3 Anhand welcher Statistik im oben stehenden Stata-Output kann die Selektivität der Stichprobe beurteilt werden? Was misst die Statistik? Interpretieren Sie den gegebenen Schätzwert. (3 Punkte)
- 4.4 Legen Sie dar, welche Eigenschaften die Variable *activity* aufweisen muss, damit sie als Ausschlussrestriktion verwendet werden kann. Beurteilen Sie, ob die Variable im vorliegenden Fall als Ausschlussrestriktion geeignet ist. (2 Punkte)
- 4.5 Welche Konsequenz ergibt sich für das zweistufige Heckman-Schätzverfahren, wenn keine Ausschlussrestriktionen gefunden werden. (2 Punkte)

Aufgabe 5 (4 Punkte)

Definieren Sie knapp folgende Konzepte und Begriffe, die bei Verweildauermodellen verwendet werden.

- (a) Spell
- (b) linkszensierte Daten
- (c) duration dependence
- (d) Stock Sample