

Prüfung im Fach Mikroökonomie SS 2011

Name, Vorname	
Matrikelnr.	
Studiengang	
E-Mail-Adresse	
Unterschrift	

Vorbemerkungen:

- Anzahl der Aufgaben:**
- Die Klausur besteht aus 5 Aufgaben, die alle bearbeitet werden müssen.
- 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 (10.5 Punkte)

1.1 Stellen Sie die Likelihood-Funktion sowie die Log-Likelihood-Funktion für ein binäres Modell in allgemeiner Form auf. (2 Punkte)

1.2 Der Gradienten-Vektor (Score-Vektor) eines Logit-Modells lautet: $\mathbf{s}(\boldsymbol{\beta}) = \sum_{i=1}^N \mathbf{x}_i (y_i - p_i)$.

1.2.1 Geben Sie unter Verwendung dieses Ergebnisses die Bedingung erster Ordnung für den Maximum-Likelihood-Schätzer des Logit-Modells an. (1 Punkt)

1.2.2 Zeigen Sie formal, dass der Mittelwert der vorhergesagten Wahrscheinlichkeiten im Logit-Modell, das mit einer Konstanten geschätzt wird, dem im Sample beobachteten Anteil der Beobachtungen mit $y_i = 1$ entspricht. (3 Punkte)

1.3 Nennen Sie drei Schritte bei der iterativen numerischen Berechnung des ML-Schätzers. Geben Sie drei sinnvolle Abbruchkriterien an. (4.5 Punkte)

Aufgabe 2 (9 Punkte)

Mit einer Regressionsanalyse werden Determinanten der Arbeitszufriedenheit modelliert. In Tabelle 1 sind die Regressionsergebnisse ausgewiesen. Die Variablen sind wie folgt definiert:

- jsat = Zufriedenheit mit der Arbeit (=1, unzufrieden; =2, zufrieden; =3, sehr zufrieden)
- educ = Jahre der Schul- und Berufsausbildung
- age = Alter
- agesq = Alter quadriert
- ln_hours = wöchentliche Arbeitszeit, logarithmiert
- ln_inc = monatliches Bruttoeinkommen, logarithmiert

Tabelle 1: Regressionsergebnisse

Ordered probit regression		Number of obs	=	7559		
		LR chi2(5)	=	80.10		
		Prob > chi2	=	0.0000		
Log likelihood = -7922.0936		Pseudo R2	=	0.0050		

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

	educ	.0135585	.0049779	2.72	0.006	.003802 .0233149
	age	-.0513299	.0084765	-6.06	0.000	-.0679436 -.0347162
	agesq	.0005166	.0000991	5.21	0.000	.0003224 .0007108
	ln_hours	-.1343222	.0408597	-3.29	0.001	-.2144058 -.0542386
	ln_inc	.1091773	.0263342	4.15	0.000	.0575632 .1607915

	/cut1	-1.131448	.193896			-1.511477 -.7514186
	/cut2	.1214197	.1935914			-.2580124 .5008518

2.1 Stellen Sie die Likelihood-Funktion für das in Tabelle 1 geschätzte Modell auf. Geben Sie die Likelihood-Beiträge für die drei Antwortkategorien an. (5 Punkte)

2.2 Beschreiben Sie kurz, was sich an der Likelihood-Funktion bzw. den Likelihood-Beiträgen aus Teilaufgabe 2.1 ändert, wenn das Modell mit Konstante geschätzt wird? (1 Punkt)

2.3 Berechnen und interpretieren Sie die kompensierende Einkommensvariation bei einer Verlängerung der Arbeitszeit. (3 Punkte)

Aufgabe 3 (9 Punkte)

Die Anzahl außerehelicher Affären wird mit Hilfe von Regressionsanalysen geschätzt, wobei Effekte des Alters (*age*), der bisherigen Ehedauer in Jahren (*yrs_married*), der Religiosität (*religious*, = 1, falls religiös; = 0 sonst) und der Selbsteinschätzung der Ehequalität (*sr_marriage*, 1 = sehr unzufrieden, ..., 5 = sehr zufrieden) modelliert werden. 451 der 601 in der Stichprobe enthaltenen verheirateten Personen hatten bislang keine Affäre. Die Schätzergebnisse von drei verschiedenen Regressionsmodellen sind in Tabelle 2 wiedergegeben.

Tabelle 2: Regressionsergebnisse

Variable	KQ-Regression		gestutzte Regression		zensierte Regression (Tobit)	
	Koeffizient	Std.fehler	Koeffizient	Std.fehler	Koeffizient	Std.fehler
age	-0.043**	(0.022)	-0.077	(0.118)	-0.160**	(0.078)
yrs_married	0.155***	(0.037)	0.559**	(0.219)	0.539***	(0.134)
religious	-0.487***	(0.111)	-1.531**	(0.617)	-1.723***	(0.405)
sr_marriage	-0.705***	(0.118)	-1.331**	(0.562)	-2.267***	(0.408)
Konstante	5.886***	(0.775)	8.937**	(3.750)	9.083***	(2.659)
Skalierungsfaktor			0.271		0.234	

Signifikanzniveau: * < 0.1, ** < 0.05, *** < 0.01. Im gestutzten Regressionsmodell ist der Skalierungsfaktor $1 - \delta(\alpha)$, wobei $\delta(\alpha) = \lambda(\alpha) \cdot (\lambda(\alpha) - \alpha)$ mit $\lambda(\alpha) = \phi(\alpha)/(1 - \Phi(\alpha))$ und $\alpha = -\mathbf{x}'\beta/\sigma$. Im zensierten Regressionsmodell ist der Skalierungsfaktor $\Phi(\mathbf{x}'\beta/\sigma)$.

- 3.1 Legen Sie dar (z. B. anhand einer Grafik), warum die Ergebnisse der KQ-Schätzung von denen der zensierten Regression abweichen. (2 Punkte)
- 3.2 Nennen Sie drei Arten marginaler Effekte, die auf Basis einer zensierten Regression berechnet werden können. (1.5 Punkte)
- 3.3 Berechnen und interpretieren Sie den (unverzerrten) marginalen Effekt eines weiteren Ehejahres auf die Anzahl der Affären von allen verheirateten Personen (d. h. Personen mit bzw. ohne Affären). (3.5 Punkte)
- 3.4 Das Tobit-Modell wurde unter Verwendung der folgenden Log-Likelihood-Funktion geschätzt. Erklären Sie kurz, was die beiden Summen $\sum_{y_i=0}$ bzw. $\sum_{y_i>0}$ jeweils berechnen. (2 Punkte)

$$\ln L = \sum_{y_i=0} \ln \Phi \left(\frac{0 - \mathbf{x}'_i \beta}{\sigma} \right) + \sum_{y_i>0} \ln \frac{1}{\sigma} \phi \left(\frac{y_i - \mathbf{x}'_i \beta}{\sigma} \right)$$

Aufgabe 4 (11 Punkte)

Die Determinanten des Marktlohns wurden auf Basis einer Stichprobe beschäftigter Arbeitnehmer geschätzt. Zur Berücksichtigung potentiell endogener Stichprobenselektion wird das zweistufige Heckman-Verfahren verwendet. Tabelle 3 gibt die Regressionsergebnisse wieder. Die Variablen sind wie folgt definiert:

- inlf = Indikatorvariable für Erwerbstätigkeit (= 1, falls Person erwerbstätig ist; = 0 sonst)
- lwage = Stundenlohn (logarithmiert)
- educ = Bildung (in Jahren)
- exper = Berufserfahrung (in Jahren)
- expersq = Berufserfahrung quadriert
- nwifeinc = Nichtarbeitseinkommen in Tsd. USD
- age = Alter
- kidslt6 = Anzahl Kinder jünger als 6
- kidsge6 = Anzahl Kinder 6 Jahre und älter

Tabelle 3: Regressionsergebnisse

		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Heckman selection model -- two-step estimates							
(regression model with sample selection)							
	Number of obs				=		753
	Censored obs				=		325
	Uncensored obs				=		428
	Wald chi2(3)				=		51.53
	Prob > chi2				=		0.0000

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

lwage							
	educ	.1090655	.015523	7.03	0.000	.0786411	.13949
	exper	.0438873	.0162611	2.70	0.007	.0120163	.0757584
	expersq	-.0008591	.0004389	-1.96	0.050	-.0017194	1.15e-06
	_cons	-.5781032	.3050062	-1.90	0.058	-1.175904	.019698

inlf							
	educ	.1309047	.0252542	5.18	0.000	.0814074	.180402
	exper	.1233476	.0187164	6.59	0.000	.0866641	.1600311
	expersq	-.0018871	.0006	-3.15	0.002	-.003063	-.0007111
	nwifeinc	-.0120237	.0048398	-2.48	0.013	-.0215096	-.0025378
	age	-.0528527	.0084772	-6.23	0.000	-.0694678	-.0362376
	kidslt6	-.8683285	.1185223	-7.33	0.000	-1.100628	-.636029
	kidsge6	.036005	.0434768	0.83	0.408	-.049208	.1212179
	_cons	.2700768	.508593	0.53	0.595	-.7267473	1.266901

mills							
	lambda	.0322619	.1336246	0.24	0.809	-.2296376	.2941613

- 4.1 Erläutern Sie verbal die Vorgehensweise des zweistufigen Heckman-Schätzverfahrens. *Hinweis:* Formeln müssen nicht angegeben werden. (4 Punkte)
- 4.2 Interpretieren Sie ausführlich die in Tabelle 3 vorliegenden Ergebnisse in Bezug auf Selektivität der Stichprobe. (3 Punkte)
- 4.3 Nennen Sie zwei Schwächen des Heckman-Schätzverfahrens und gehen Sie kurz auf ihre Konsequenzen ein. (4 Punkte)

Aufgabe 5 (20.5 Punkte)

Die Wahl der Fischfangmethode von Hobbyanglern soll mit einer Regressionsanalyse erklärt werden. Als erklärende Variablen werden das individuelle Einkommen (*income*, in Tsd. USD gemessen) sowie der Preis der jeweiligen Fischfangmethode (*price*, in USD) verwendet. Es stehen drei Fischfangmethoden (*fishmode*) zur Auswahl: 1 = Strand (*beach*), 2 = Privat-Boot (*private*), 3 = Charter-Boot (*charter*).

- 5.1 Geben Sie für ein Conditional-Logit-Modell, das nur den Preis als erklärende Variable berücksichtigt, die Wahrscheinlichkeit für Alternative j , also $P(y_i = j)$, in einer allgemeinen formalen Notation an. (2 Punkte)
- 5.2 Geben Sie für ein Mixed-Logit-Modell, das Einkommen und Preis als erklärende Variablen berücksichtigt, die Wahrscheinlichkeit für Alternative j , also $P(y_i = j)$, in einer allgemeinen formalen Notation an. (4 Punkte)

5.3 Interpretieren Sie die Steigungsparameter des folgenden Mixed-Logit-Modells inhaltlich. (4.5 Punkte)

```

Alternative-specific conditional logit      Number of obs   =      3012
Case variable: id                        Number of cases  =      1004

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

                                           Wald chi2(3)    =      156.75
Log likelihood = -853.08318              Prob > chi2     =      0.0000
    
```

fish_dv	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
fishmode						
price	-.0213763	.001839	-11.62	0.000	-.0249806	-.017772
beach	(base alternative)					
charter						
income	-.0207736	.0541491	-0.38	0.701	-.1269039	.0853568
_cons	1.673575	.2305556	7.26	0.000	1.221694	2.125455
private						
income	.1078711	.0537996	2.01	0.045	.0024258	.2133165
_cons	.4186247	.2330478	1.80	0.072	-.0381405	.8753899

5.4 Interpretieren Sie die nachfolgend ausgewiesenen marginalen Effekte des Mixed-Logit-Modells inhaltlich. (5 Punkte)

```

Pr(choice = private|1 selected) = .45715058
-----
variable | dp/dx Std. Err. z P>|z| [ 95% C.I. ]
-----|-----
price    |
  beach  | .000456 .000055  8.32 0.000 .000348 .000563
  charter| .004849 .000485 10.00 0.000 .003899 .0058
  private| -.005305 .000463 -11.46 0.000 -.006212 -.004398
-----|-----
casevars |
  income | .031482 .006928  4.54 0.000 .017904 .04506
-----|-----
    
```

5.5 In einer weiteren Schätzung werden die alternativen-spezifischen Fangquoten als zusätzliche erklärende Variable in das Modell aufgenommen. Die Schätzung ergibt einen Log-Likelihood-Wert von -850.61034. Überprüfen Sie, ob sich der Erklärungsgehalt des Modells signifikant verbessert hat. Geben Sie die Teststatistik, Freiheitsgrade und kritischen Wert zum Signifikanzniveau $\alpha = 0.05$ an. Berechnen Sie den empirischen Wert der Teststatistik und treffen Sie eine Testentscheidung. (5 Punkte)