

Prüfung im Fach Mikroökonomie im Wintersemester 2017/2018 Aufgaben

Vorbemerkungen:

**Anzahl der
Aufgaben:
Bewertung:**

- Die Klausur besteht aus 4 Aufgaben.
- 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 (17 Punkte)

Sie analysieren die Determinanten der Arbeitszufriedenheit. Hierfür nutzen Sie einen Datensatz, der erwerbstätige Personen über einen Zeitraum von drei Jahren jährlich fragt, wie zufrieden sie mit ihrer derzeitigen Tätigkeit sind. Der Datensatz enthält die folgenden Informationen:

job_satisfaction	=1, wenn Person i im Jahr t sehr zufrieden mit derzeitiger Tätigkeit ist; sonst=0
age	Alter von Person i im Jahr t in Jahren
salary_inc	=1, wenn Gehalt von Person i im Jahr t erhöht wurde; sonst=0
over_educ	=1, wenn Person i im Jahr t für derzeitige Tätigkeit überqualifiziert ist; sonst=0

Tabelle 1: Deskriptive Statistik

	Mittelwert	Std.abw.
job_satisfaction	0,33	0,13
age	36,17	16,48
salary_inc	0,21	0,17
over_educ	0,30	0,16

Tabelle 2: Logit-Schätzung

	Koeffizient	Std.fehler
age	-0,01 ***	0,00
salary_inc	0,55 ***	0,07
over_educ	-0,10 **	0,05
_cons	0.20	0.09

Abhängige Variable: *job_satisfaction*,

Signifikanzniveau: * < 0,1, ** < 0,05, *** < 0,01.

1.1 Stellen Sie die Log-Likelihood Funktion für ein Logit Modell auf und vereinfachen Sie diese (3 Punkte).

Hinweise:

(1) Die allgemeine Log-Likelihood-Funktion für eine binäre Zufallsvariable (y_i) lautet: $\ln(L(\beta)) = \sum_{i=1}^N \{y_i * \ln(\frac{p_i}{1-p_i}) + \ln(1 - p_i)\}$.

(2) Für das Logit-Modell gilt $p_i = \frac{\exp(x_i'\beta)}{1+\exp(x_i'\beta)}$.

1.2 Wie wird der Likelihood-Schätzer aus der Log-Likelihood Funktion abgeleitet? Erläutern Sie verbal. (2 Punkte)

1.3 Interpretieren Sie den Effekt der Variable *salary_inc* inhaltlich am Mittelwert der Daten. (5 Punkte)

Hinweis: Runden Sie alle Zwischenergebnisse auf zwei Nachkommastellen.

1.4 Im weiteren Verlauf Ihrer Analyse schätzen Sie ein Logit Modell mit fixen Effekten auf der Individualebene. Für die Schätzung der individualspezifischen Effekte verwenden Sie den Dummy Variablen Ansatz. Ein Kollege weist Sie darauf hin, dass Sie nun möglicherweise ein *incidental parameter problem* haben. Erläutern Sie, wann dieses Problem auftritt und nennen Sie zwei Folgen für die Schätzung. (3 Punkte)

1.5 Benennen Sie das Schätzverfahren, mit dem sich das *incidental parameter problem* lösen lässt und skizzieren Sie kurz die Idee dieses Modells. (4 Punkte)

Aufgabe 2 (13 Punkte)

Sie untersuchen den Zusammenhang zwischen verschiedenen Charakteristika der Grundschulzeit und der Dauer der Schulbildung. Für Ihre Analyse verwenden Sie einen Datensatz, der folgenden Informationen über 25-jährige Personen enthält:

dropout	=1, wenn Person i die Schule ohne Schulabschluss verlassen hat; sonst=0
educ	Dauer der Schulbildung in Jahren (=fehlender Wert, wenn $dropout=1$)
class_size	Durchschnittliche Zahl der Mitschüler in der Grundschule
male_teacher	=1, wenn Person i einen männlichen Klassenlehrer in der Grundschule hatte; sonst=0

- 2.1 Erläutern Sie knapp den Unterschied zwischen einer gestutzten und zensierten Verteilung und beschreiben Sie, wie sie mit den beschriebenen Daten diese Verteilungen erhalten. Wie müssten Sie die Daten aufbereiten, um sie für eine Tobit-Schätzung nutzen zu können? (5 Punkte)
- 2.2 Nennen Sie zwei Konsequenzen für die Eigenschaft des Schätzers, wenn eine KQ-Schätzung mit gestutzten Daten durchgeführt wird. (2 Punkte)
- 2.3 Nehmen Sie an, der Datensatz wurde, wie von Ihnen in 2.1 beschrieben, für eine Tobit-Schätzung aufbereitet. Eine Tobit-Schätzung liefert die folgenden Ergebnisse:

Variable	Tobit	
	Koeffizient	Std.fehler
class_size	-0,052 *	(0,028)
male_teacher	0,120	(0,094)
Konstante	8,782 ***	(2,149)
Skalierungsfaktor $\Phi(x'_i\beta/\sigma)$	0,193	
Skalierungsfaktor $(1 - \delta(\alpha))$	0,212	
N	1500	
Log-Likelihood-Wert	-10.903,51	

Signifikanzniveau: * $<0,1$, ** $<0,05$, *** $<0,01$.

Berechnen und interpretieren Sie den (unverzerrten) marginalen Effekt eines zusätzlichen Klassenkameraden auf die Bildungsjahre für alle 25-Jährigen. (2 Punkte)

- 2.4 Führen Sie für die Tobit-Schätzung in 2.3 einen Likelihood-Ratio Test auf Gesamtsignifikanz am 5%-Signifikanzniveau durch. Geben Sie Null- und Alternativhypothese, Freiheitsgrade, kritischen Wert, Teststatistik und Testentscheidung an. (4 Punkte)

Hinweis: Der Wert der Log-Likelihood-Funktion des obigen Modells

(1) ohne Konstante beträgt -11.721,09.

(2) ohne die Variablen *class_size* und *male_teacher* beträgt -11.834,12.

Aufgabe 3 (18,5 Punkte)

Eine Bank bewirbt bei 1000 Kunden ein neues Darlehen. 490 Kunden reagieren auf die Werbemaßnahme und leihen sich den Betrag *loan*. Für eine Schätzung des Zusammenhangs von Darlehensbetrag und Kundencharakteristika stehen folgende Variablen zur Verfügung:

<i>loan</i>	Betrag des Darlehens (=fehlender Wert, wenn kein Darlehen)
<i>female</i>	=1, falls Frau, 0 sonst
<i>age</i>	Alter in Jahren
<i>response</i>	=1, falls Person auf Werbemaßnahme antwortet, 0 sonst
<i>credit</i>	=1, falls Person in der Vergangenheit einen Kredit oder ein Dahrlehen hatte, 0 sonst.

```

Heckman selection model -- two-step estimates      Number of obs      =      1000
(regression model with sample selection)          Censored obs       =           ?
                                                    Uncensored obs     =      490

                                                    Wald chi2(2)       =      19.82
                                                    Prob > chi2        =      0.0002
    
```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
loan					
female	-.0484285	.1737498	-0.28	0.780	-.3889717 .2921148
age	.0816150	.0271189	3.01	0.003	.0284631 .1347671
_cons	2.086416	.8419196	2.48	0.013	.4362843 3.736549
response					
female	.5881135	.0966845	6.08	0.000	.3986154 .7776116
age	.0416800	.0215441	1.93	0.053	-.0005456 .0839057
credit	.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

- Berechnen und interpretieren Sie inhaltlich den fehlenden Wert im Stata-Output ("?"). (1 Punkt)
- Erläutern Sie unter Bezugnahmen auf die Analyse des Darlehensbetrags den Begriff der endogenen Stichprobenselektion. Welche Konsequenz folgt daraus für eine Kleinst-Quadrate-Schätzung? (3 Punkte)
- Erläutern Sie verbal die Vorgehensweise des zweistufigen Heckman-Schätzverfahrens am Beispiel. (4 Punkte)
Hinweis: Formeln müssen nicht angegeben werden.
- Nennen Sie eine zentrale Annahme des Heckman-Schätzverfahrens. Welche Folge hat eine Annahmeverletzung? (2,5 Punkte)
- 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)
- Legen Sie dar, welche Eigenschaften die Variable *credit* aufweisen muss, damit sie als Ausschlussrestriktion verwendet werden kann. Beurteilen Sie, ob die Variable im vorliegenden Fall als Ausschlussrestriktion geeignet ist. (3 Punkte)
- Welche Konsequenz ergibt sich für die Eigenschaften des zweistufigen Heckman-Schätzers, wenn keine Ausschlussrestriktionen gefunden werden. (2 Punkte)

Aufgabe 4 (11,5 Punkte)

Für eine Arztpraxis analysieren Sie die Dauer der individuellen Grippeerkrankungen und ihre Determinanten.

- 4.1 Definieren Sie den Begriff *duration dependence* am Beispiel der Grippendauer. Unterscheiden Sie dabei zwischen negativer bzw. positiver *duration dependence*. (3 Punkte)
- 4.2 Sie analysieren die Dauer von Grippeerkrankungen (gemessen in Tagen) mit einer Weibull-Regression. Als erklärende Variable nehmen Sie das Alter des Patienten in Jahren (*alter*) auf. Interpretieren Sie den Koeffizienten der Variable *alter* inhaltlich und statistisch. Runden Sie auf die dritte Nachkommastelle. (2,5 Punkte)

```
Weibull regression -- log relative-hazard form

No. of subjects =      1445                Number of obs   =      1445
No. of failures =       552                LR chi2(1)      =      165.48
Time at risk    =     80013                Prob > chi2     =       0.0000

Log likelihood = -853.08318

-----+-----
      _t |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
      alter |   .0887867   .0134355     6.61  0.000    .0624535   .1151198
      _cons |  -3.402094   .3010177   -11.30  0.000   -3.992077  -2.81211
-----+-----
      /ln_p |  -.2158398   .0389149    -5.55  0.000   -.2921115  -.1395681
-----+-----
           p |   .8058644   .0313601           .7466852   .8697338
          1/p |   1.240904   .0482896           1.149777   1.339252
-----+-----
```

- 4.3 Erklären Sie den Unterschied zwischen dem Weibull-Modell und dem Cox Proportional-Hazard-Modell. (2 Punkte)
- 4.4 Sie untersuchen anschliessend den Zusammenhang zwischen der Anzahl der Grippeerkrankungen eines Patienten im Jahr (*grippen*) und seinem Alter (*alter*). Sie schätzen ein Poissonmodell für die Daten der gesamten Patientenschaft und erhalten folgenden Output:

```
Iteration 0:  log likelihood = -72190.583
Iteration 1:  log likelihood = -72190.583

Poisson regression                Number of obs   =    55,089
LR chi2(1)                        =     12.50
Prob > chi2                        =     0.0004
Pseudo R2                          =     0.0001

Log likelihood = -72190.583

-----+-----
      grippen |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
      alter |   .0182651   .0051670     3.54  0.000    .0081382   .0283913
      _cons |  -.0915017   .0215336    -4.25  0.000   -.1337068  -.0492966
-----+-----
```

Interpretieren Sie den Koeffizienten des Alters inhaltlich. Runden Sie auf die dritte Nachkommastelle. (1 Punkt)

- 4.5 Erläutern Sie kurz, was unter Unterstreueung zu verstehen ist. Was bedeutet das für die Schätzergebnisse eines Poisson-Modells? (2 Punkte)
- 4.6 Nennen Sie einen Vorteil des Poisson-Modells im Vergleich zum linearen Modell (1 Punkt).