

Prüfung im Fach Mikroökonomie im Sommersemester 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 (13,5 Punkte)

Sie analysieren den Effekt von Bildung auf das Alters bei der ersten Hochzeit. Hierfür liegen Ihnen Spelldaten vor, die Beobachtungen im Alter von 16 bis 45 enthalten.

Hinweis: Gehen Sie davon aus, dass keine Person vor dem Alter 16 geheiratet hat.

- 1.1 Erläutern Sie, ob die verwendeten Daten zensiert sind und nennen Sie ggf. die Richtung der Zensierung. Zählen Sie zwei Konsequenzen auf, die durch die Verwendung zensierter Daten in einer KQ-Schätzung entstehen. (4 Punkte)
- 1.2 Was beschreiben die abhängigen Variablen bei Modellen in diskreter und stetiger Zeit? (2 Punkte)
- 1.3 Sie schätzen ein Verweildauermodell in diskreter Zeit mit jährlichen Beobachtungen und unterstellen eine logistische Verteilung ($P(y_i = 1) = \frac{e^{\beta'x_i}}{1+e^{\beta'x_i}}$). Geben Sie den Log-Likelihood-Beitrag für Beobachtung 1 (erste Hochzeit im Alter 16) und Beobachtung 2 (erste Hochzeit im Alter 18) an. (3 Punkte)
- 1.4 Sie analysieren die Dauer bis zur ersten Hochzeit (gemessen in Jahren) mit einer Weibull-Regression. Als erklärende Variable nehmen Sie die Bildungsjahre (*bil_jahre*) auf. Der Koeffizient der Variable *bil_jahre* beträgt $-0,074$. Interpretieren Sie den Koeffizienten der Variable *bil_jahre* inhaltlich. (1,5 Punkte)
- 1.5 Unterstellen Sie, dass sich die duration dependence im Beispiel für Personen mit und ohne Migrationshintergrund unterscheidet. (1) Definieren Sie den Begriff *duration dependence* am Beispiel der ersten Eheschließung. (2) Nennen Sie zwei mögliche Folgen für die Schätzung eines Weibull-Modells, wenn Sie keine Information zum Migrationsstatus haben. (3) Beschreiben Sie außerdem, wie Sie trotz fehlender Information zum Migrationsstatus verhindern, dass der Unterschied in der duration dependence Ihre Schätzergebnisse beeinflusst. (3 Punkte)

Aufgabe 2 (17 Punkte)

Sie untersuchen die Determinanten der erzielten Punkte des 1.FC Nürnberg am Saisonende. Ihnen liegen Informationen zu 30 Spielzeiten vor. Die Variablen in Ihrem Datensatz sind wie folgt kodiert:

points	Anzahl der am Saisonende erzielten Punkte
value	Durchschnittlicher Marktwert aller Spieler des 1. FC Nürnberg in Millionen Euro
first	=1, falls der 1. FC Nürnberg in der 1. Bundesliga spielte, =0 sonst
second	=1, falls der 1. FC Nürnberg in der 2. Bundesliga spielte, =0 sonst
third	=1, falls der 1. FC Nürnberg in der 3. Bundesliga oder niedrigerer Liga spielte, =0 sonst

Tobit regression		Number of obs	=	30
Log likelihood = -307.3193		LR chi2(2)	=	???
		Prob > chi2	=	???
		Pseudo R2	=	0.0186

points	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
value	.97297	.14650	6.64	0.000	.67997 1.26597
first	-38.92351	5.35949	-7.26	0.000	-49.42895 -28.41808
second	-15.06481	4.56820	3.30	0.001	-24.03319 -6.11114
_cons	70.99468	9.73499	7.29	0.000	51.52470 90.46466

Hinweis: Runden Sie alle Zwischenergebnisse auf drei Nachkommastellen.

- 2.1 Erläutern Sie, warum in diesem Fall eine Tobit- anstatt einer KQ-Regression angemessen ist. (1 Punkt)

2.2 Bestimmen und interpretieren Sie den (unverzerrten) marginalen Effekt einer Erhöhung des durchschnittlichen Marktwertes der Spieler auf die latente Anzahl der erreichten Punkte. (2 Punkte)

Hinweise:

Skalierungsfaktor $\Phi(x'_i\beta/\sigma) = 0,003$

Skalierungsfaktor $(1 - \delta(\alpha)) = 0,005$

2.3 Überprüfen Sie mit einem Wald-Test am 1%-Signifikanzniveau die Hypothese, dass sich die Höhe der erreichten Punkte am Saisonende in mindestens der 1. oder 2. Bundesliga von der in der 3. Bundesliga unterscheidet. Geben Sie Null- und Alternativhypothese, Freiheitsgrade, kritischen Wert, Teststatistik und Testentscheidung an. (5,5 Punkte)

Hinweise:

(1) Die Teststatistik des Wald-Tests lautet: $W = \hat{\beta}' \widehat{Var}(\hat{\beta})^{-1} \hat{\beta} \sim \chi_k^2$

(2) Verwenden Sie folgende geschätzte inverse Kovarianzmatrix des Schätzvektors $\hat{\beta} = (\hat{\beta}_{ek1} \hat{\beta}_{ek2})'$:

$$\widehat{Var}(\hat{\beta})^{-1} = \begin{pmatrix} 0,30 & 0 \\ 0 & -0,20 \end{pmatrix}$$

2.4 Führen Sie einen Likelihood-Ratio-Test auf Gesamtsignifikanz auf dem 10%-Signifikanzniveau durch. Geben Sie Null- und Alternativhypothese, Freiheitsgrade, kritischen Wert, Teststatistik und Testentscheidung an. (4 Punkte)

Hinweis: Der Log-Likelihood-Wert eines Modells, welches nur eine Konstante enthält, beträgt -334,135.

2.5 Zusätzlich analysieren Sie die Determinanten der Abstiegswahrscheinlichkeit des 1.FC Nürnberg. Hierfür nutzen Sie den gleichen Datensatz wie zuvor, mit folgenden zusätzlichen Informationen:

relegated = 1, wenn der 1.FC Nürnberg am Saisonende abgestiegen ist, = 0 sonst

coach = 1, wenn der Trainer am Saisonanfang gewechselt wurde, = 0 sonst

Deskriptive Statistik		
	Mittelwert	Std.abw.
relegated	0.21	0.17
coach	0.25	0.20
first	0.60	0.26
second	0.30	0.16

Logit-Schätzung		
	Koeffizient	Std.fehler
coach	-0.20***	0.05
first	0.32***	0.10
second	0.09	0.16
_cons	0.05	0.09

Abhängige Variable: relegated

Signifikanzniveau: *** < 0,01

Berechnen und interpretieren Sie den Effekt der Variable *coach* inhaltlich am Mittelwert der Daten. (4,5 Punkte)

Aufgabe 3 (16,5 Punkte)

Für eine Krankenkasse untersuchen Sie die Determinanten des subjektiven Gesundheitsempfindens. Für die Analyse nutzen Sie einen Datensatz mit folgenden Informationen:

<i>gesundheit</i>	Subjektives Gesundheitsempfinden (1=gut, 2=durchschnittlich, 3=schlecht)
<i>raucher</i>	=1, wenn Raucher; =0 sonst
<i>mann</i>	=1, wenn Mann; =0 sonst
<i>koerpergewicht</i>	Körpergewicht (in kg)

Die Schätzung eines geordneten Probit-Modells liefert folgende Ergebnisse:

```

Iteration 0:  log likelihood = -15512.456
Iteration 1:  log likelihood = -15013.235
Iteration 2:  log likelihood = -15013.008
Iteration 3:  log likelihood = -15013.008

Ordered probit regression              Number of obs   =    15,237
                                      LR chi2(3)      =    998.90
                                      Prob > chi2     =    0.0000
Log likelihood = -15013.008           Pseudo R2      =    ???

```

gesundheit	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
raucher	.4262583	.0239555	17.79	0.000	.3793063	.4732103
mann	-.2421375	.0200439	-12.08	0.000	-.2814228	-.2028522
koerpergewicht	.0136727	.0005243	26.08	0.000	.0126451	.0147003
/cut1	.9584058	.0405198			.8789886	1.037823
/cut2	2.069846	.0423859			1.986771	2.152921

- 3.1 Erläutern Sie formal und in Worten den Zusammenhang zwischen dem beobachteten Gesundheitsstatus y_i und der latenten abhängigen Variable y_i^* . Gehen Sie von einem Modell ohne Konstante aus. (3,5 Punkte)
- 3.2 Stellen Sie unter Berücksichtigung der Standardnormalverteilung die Likelihood-Funktion für das geordnete Probit-Modell auf. (3 Punkte)
- 3.3 Interpretieren Sie den Koeffizienten für *raucher* inhaltlich und statistisch. (3 Punkte)
- 3.4 Berechnen Sie die kompensierende Variation im Körpergewicht für Frauen. Interpretieren Sie das Ergebnis. (3 Punkte)
- 3.5 Der Datensatz enthält eine weitere Variable zum subjektiven Gesundheitsempfinden (*gesundheit_dtl*), die das Gesundheitsempfinden in 10 Kategorien unterteilt. Ein Kollege schlägt vor, ein Zähldatenmodell mit *gesundheit_dtl* als abhängige Variable zu schätzen. Halten Sie diesen Vorschlag für sinnvoll? Begründen Sie Ihre Antwort mit zwei Argumenten. (2 Punkte)
- 3.6 Nennen Sie eine Verwendung des McFadden R^2 und berechnen Sie das McFadden R^2 für die oben abgebildete Schätzung. (2 Punkte)

Aufgabe 4 (13 Punkte)

Sie analysieren die Determinanten der Stimmabgabe für die politische Partei TUR bei der türkischen Parlamentswahl. Für Ihre Analyse verwenden Sie einen Datensatz, der 500 türkischstämmige Wahlberechtigte in Deutschland beinhaltet. Ihnen stehen folgende Variablen zur Verfügung:

vote	=1, falls Person für die Partei TUR gestimmt hat, 0 sonst
female	=1, falls Frau, 0 sonst
age	Alter in Jahren
participation	=1, falls Person bei der Parlamentswahl abgestimmt hat, 0 sonst
tv	=1, falls Person regelmäßig türkische Fernsehsender verfolgt, 0 sonst.

```

Heckman selection model -- two-step estimates   Number of obs   =       500
(regression model with sample selection)      Censored obs    =       300
                                              Uncensored obs  =       200

                                              Wald chi2(2)    =       21.82
                                              Prob > chi2     =       0.0000
  
```

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

vote						
female	-.0816150	.0271189	-3.01	0.003	-.1347671	-.0284631
age	.0055080	.0012175	4.52	0.000	.0031215	.0078946
_cons	.4304413	.3210011	1.34	0.180	-.1987092	1.059592

participation						
female	-.0484285	.1737498	-0.28	0.780	-.3889717	.2921148
age	.0090438	.0003700	24.42	0.000	.0083177	.0097698
tv	.1431383	.0377298	3.79	0.000	.0691819	.2170947
_cons	.1376843	.2114778	0.65	0.515	-.2768454	.5522140

mills						
lambda	.5611673	.1115717	5.03	0.000	.3424908	.7798438

- 4.1 Erläutern Sie unter Bezugnahmen auf die Analyse des Wahlverhaltens den Begriff der endogenen Stichprobenselektion. Welche Konsequenz folgt daraus für eine Kleinste-Quadrate-Schätzung? (3 Punkte)
- 4.2 Erläutern Sie verbal die Vorgehensweise des zweistufigen Heckman-Schätzverfahrens am vorliegenden Beispiel. (4 Punkte)
Hinweis: Formeln müssen nicht angegeben werden.
- 4.3 Anhand welcher Statistik im oben stehenden Stata-Output kann die Selektivität der Stichprobe beurteilt werden? Was misst die Statistik allgemein? Erläutern Sie, ob im Beispiel Selektivität vorliegt. (3 Punkte)
- 4.4 Legen Sie dar, welche Eigenschaften die Variable *tv* aufweisen muss, damit sie als Ausschlussrestriktion verwendet werden kann. Erläutern Sie, ob die Variable im vorliegenden Fall als Ausschlussrestriktion geeignet ist. (3 Punkte)