

## Masterprüfung SS 2016

**Fach:** Mikroökonomie

**Semester:** Sommersemester 2016

**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 60 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:**

- 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)

- 1.1 Beschreiben Sie die Grundidee des Maximum-Likelihood Schätzers. Ist der ML Schätzer in großen Stichproben unter der Annahme der korrekt spezifizierten Likelihood-Funktion konsistent und/oder effizient? (2 Punkte)
- 1.2 Schreiben Sie die Likelihood-Funktion sowie die Log-Likelihood-Funktion für ein binäres Modell in allgemeiner Form aus. (2 Punkte)
- 1.3 Sie stellen ein Modell mit binärer abhängiger Variable auf und unterstellen eine logistische Verteilung.
- (a) Berechnen Sie das odds-ratio für eine Eintrittswahrscheinlichkeit ( $P(y_i = 1)$ ) von 35%. (1 Punkt)
  - (b) Gehen Sie von  $\beta_k = 0,3$  aus. Wie verändert sich das odds-ratio, wenn der kontinuierliche Regressor  $x_k$  um eine Einheit steigt? (2 Punkte)

### Aufgabe 2 (11 Punkte)

Sie interessieren sich für den Zusammenhang zwischen Aufenthaltsdauer im Gastland und Stundenlohn von Migranten. Ihr Querschnittsdatensatz enthält Informationen aus dem Jahr 2004 zu Migranten, welche zwischen 1990 und 2000 zugewandert und nicht zurück gewandert sind. Sie stellen folgendes Modell auf:

$$wage_i = \beta_0 + \beta_1 YSM_i + \gamma' X_i + u_i \quad (1)$$

wobei  $wage$  den Brutto-Stundenlohn in Euro misst und  $YSM$  die Aufenthaltsdauer in Jahren seit Migration.  $X$  stellt einen Vektor mit Kontrollvariablen dar.

- 2.1 Erläutern Sie den Begriff der endogenen Stichprobenselektion allgemein und anhand des genannten Modells. Welche Konsequenz hat endogene Stichprobenselektion für eine Kleinst-Quadrate Schätzung? (3 Punkte)
- 2.2 Sie nutzen für Ihre Schätzung das zweistufige Heckman-Verfahren.
- (a) Beschreiben Sie verbal die Vorgehensweise des zweistufigen Heckman-Verfahrens für das vorliegende Beispiel. (4 Punkte)
  - (b) Erläutern Sie für das gegebene Beispiel, welche Eigenschaften eine Variable haben muss, damit sie für eine Ausschlussrestriktion in Frage kommt. Nennen Sie eine Konsequenz für das zweistufige Heckman-Verfahren, wenn keine Ausschlussrestriktion vorliegt. (3 Punkte)
- 2.3 Wie können Sie prüfen, ob in der durchgeführten Schätzung Selektionsverzerrung vorliegt? (1 Punkt)

### Aufgabe 3 (10 Punkte)

Sie untersuchen die Determinanten der Anzahl verkaufter Konzerttickets in einem Münchener Konzertsaal mit 2000 Plätzen. Ihnen liegen Informationen zu 489 Konzerten vor.

Die Variablen in Ihrem Datensatz sind wie folgt kodiert:

ticketssold	Zahl der verkauften Tickets für eine Veranstaltung
releases	Anzahl der aufgenommenen Studioalben des Künstlers
rock	1, falls Genre= Rockmusik, 0 sonst
pop	1, falls Genre= Popmusik, 0 sonst
folk	1, falls Genre= Volksmusik, 0 sonst (Referenz)

```
Tobit regression                                Number of obs =          489
                                                LR chi2(4)             =        116.57
                                                Prob > chi2            =         0.0000
Log likelihood = -3079.3193                    Pseudo R2              =         0.0186
```

ticketssold	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
releases	75.37487	15.89734	4.74	0.000	44.19271 106.557
rock	12.64840	7.486000	1.69	0.092	-2.02416 27.32096
pop	-5.68537	2.985340	-1.91	0.057	-11.53663 .1658964
_cons	-150.6481	45.68200	3.30	0.001	61.11138 24033.185

3.1 Erläutern Sie, unter welchen Bedingungen in diesem Fall eine Tobit- anstatt einer KQ-Regression angemessen ist. (3 Punkte)

3.2 Berechnen und interpretieren Sie den (unverzerrten) marginalen Effekt eines weiteren aufgenommenen Studioalbums auf die Anzahl der verkauften Konzerttickets. (2,5 Punkte)

*Hinweis:* Der Skalierungsfaktor  $\Phi(x_i'\beta/\sigma)$  beträgt 0,704.

3.3 Überprüfen Sie mit einem Wald-Test am 5% Signifikanzniveau die Hypothese, dass sich die Zahl der verkauften Konzertkarten in mindestens einem Genre von der Zahl der verkauften Tickets von volksmusikalischen Künstlern unterscheidet. Geben Sie die Null- und Alternativhypothese, Freiheitsgrade, den kritischen Wert und die Testentscheidung an. (4,5 Punkte)

*Hinweis:* Die Teststatistik des Wald-Tests lautet:  $W = \hat{\beta}'\widehat{Var}(\hat{\beta})^{-1}\hat{\beta} \sim \chi_k^2$ . Verwenden Sie folgende geschätzte inverse Kovarianzmatrix des Schätzvektors:  $\hat{\beta} = (\hat{\beta}_{rock}, \hat{\beta}_{pop})'$ :  $\begin{pmatrix} 0,958 & 0 \\ 0 & 2,371 \end{pmatrix}$

### Aufgabe 4 (13 Punkte)

Sie besuchen ein Nürnberger Fitnessstudio und analysieren die Dauer des Trainings (gemessen in Minuten) von 1500 Besuchern in Abhängigkeit vom Alter.

*Hinweis:* Runden Sie alle Zwischenergebnisse auf die dritte Nachkommastelle.

4.1 Definieren Sie den Begriff *duration dependence* am Beispiel des Trainings. Unterscheiden Sie zusätzlich zwischen negativer und positiver *duration dependence*. (3 Punkte)

4.2 Sie analysieren die Dauer des Trainings in stetiger Zeit (in Minuten) mit einer Weibull-Regression. Als erklärende Variable nehmen Sie das Alter der Personen in Jahren (*age*) auf. Interpretieren Sie den Koeffizienten der Variable *age* statistisch und inhaltlich präzise. (3 Punkte)

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

No. of subjects =	1500	Number of obs =	1500
No. of failures =	632		
Time at risk =	70152		
Log likelihood =	-593.02118	LR chi2(1) =	165.48
		Prob > chi2 =	0.0000

  

_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
age	.0597775	???	???	???	-.1740935 .2936486
_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 In einer weiteren Weibull-Schätzung wird zusätzlich der höchste Schulabschluss der Person mittels Indikatorvariablen (0/1) in das Modell aufgenommen, wobei zwischen kein Schulabschluss, Hauptschulabschluss, Realschulabschluss und Abitur unterschieden wird.

Die Schätzung ergibt einen Log-Likelihood-Wert von -578,403. Überprüfen Sie, ob sich der Erklärungsgehalt des Modells signifikant verbessert hat. Geben Sie die Teststatistik, Freiheitsgrade und kritischen Wert zum Signifikanzniveau  $\alpha = 0,1$  an. Berechnen Sie den empirischen Wert der Teststatistik und treffen Sie eine Testentscheidung. (5 Punkte)

4.4 Erklären Sie den Unterschied zwischen dem Weibull-Modell und dem Cox-Proportional-Hazard-Modell. (2 Punkte)

### Aufgabe 5 (19 Punkte)

Mit Daten von 8265 Individuen möchten Sie die Determinanten von Freizeitaktivitäten analysieren: 1=*Kino*, 2=*Theater* und 3=*Oper*. Als unabhängige Variablen kodieren Sie das Monatseinkommen in 500 Euro (*Einkommen*) und das Geschlecht *Frau* (=1 wenn Beobachtung weiblich, =0 wenn Beobachtung männlich). Ein multinomiales Logit-Modell ergibt folgende Schätzergebnisse:

```

Iteration 0: log likelihood = -10828.9
Iteration 1: log likelihood = -10724.39
Iteration 2: log likelihood = -10723.622
Iteration 3: log likelihood = -10723.622

```

```

Multinomial logistic regression      Number of obs =      8265
LR chi2(6) =      210.56
Prob > chi2 =      0.0000
Pseudo R2 =      0.0097
Log likelihood = -10723.622

```

Freizeit	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Kino	(base outcome)					
Theater						
Einkommen	.096491	.0388303	2.48	0.013	.0203851	.172597
Frau	-.180340	.0409951	-4.40	0.000	-.2606892	-.099991
_cons	.3855938	.1234583	3.12	0.002	.14362	.6275675
Oper						
Einkommen	.416778	.0335549	12.42	0.000	.3510116	.4825444
Frau	.1555685	.0495409	3.14	0.002	.0584701	.2526668
_cons	1.989291	.1031735	19.28	0.000	1.787075	2.191508

5.1 Erläutern Sie knapp die *independence of irrelevant alternatives* (IIA) Annahme im multinomialen Logit-Modell. (1 Punkt)

5.2 Sie möchten einen Hausman Test der IIA-Annahme durchführen. Hierzu schätzen Sie zusätzlich zu dem oben ausgewiesenen Output ein Modell, welches die Alternative *Theater* ignoriert. Die Teststatistik lautet:  $W = (b - B)' \cdot (V_b - V_B)^{-1} \cdot (b - B)$ . Erläutern Sie kurz die vier Komponenten der Teststatistik  $b$ ,  $B$ ,  $V_b$  und  $V_B$  sowie die Dimensionen der vier Matrizen. (4 Punkte)

5.3 Sie führen den in 5.2 beschriebenen Hausman Test auf dem 1% Signifikanzniveau durch und erhalten eine Teststatistik von 9,49. Stellen Sie Null- und Alternativhypothese auf und erläutern Sie die Eigenschaften der Schätzer unter den beiden Hypothesen. Geben Sie Freiheitsgrade, Entscheidungsregel mit kritischem Wert und Testentscheidung an. (4,5 Punkte)

5.4 Berechnen und interpretieren Sie die marginalen Effekte der Variable *Einkommen* auf die drei Freizeitaktivitäten. (3,5 Punkte)

*Hinweis:* Der marginale Effekt für die Basiskategorie *Kino* muss aus den Angaben berechnet werden. Runden Sie alle Zwischenergebnisse auf die dritte Nachkommastelle.

```

Average marginal effects      Number of obs =      8265
Model VCE      : OIM

Expression      : Pr(Freizeit==Theater), predict(outcome(2))
dy/dx w.r.t.    : Einkommen

```

	dy/dx	Delta-method Std. Err.	z	P> z	[95% Conf. Interval]	
Einkommen	???	???	???	???	.0155461	.0311461

```

Average marginal effects          Number of obs   =      8265
Model VCE      : OIM

Expression   : Pr(Freizeit==Oper), predict(outcome(3))
dy/dx w.r.t. : Einkommen

-----
            |          Delta-method
            |          dy/dx   Std. Err.      z    P>|z|      [95% Conf. Interval]
-----+-----
Einkommen  |          ???       ???          ???    ???       .0524799   .0748959
-----

```

5.5 Bestimmen Sie mittels der oben angegebenen Schätzergebnisse das Vorzeichen des marginalen Effekts der Variable *Frau* auf die Wahrscheinlichkeiten der Wahl von Theater und Oper. Zeigen Sie Ihren Rechenweg. (6 Punkte)

*Hinweis:* Mit  $j = 1$  als Basiskategorie gilt:

$$\frac{\partial P(y_i = j)}{\partial x_{ik}} = P(y_i = j)[\beta_{jk} - \beta_{2k}P(y_i = 2) - \beta_{3k}P(y_i = 3)] \quad \forall j = 1, 2, 3.$$

Runden Sie alle Zwischenergebnisse auf die dritte Nachkommastelle.