

## Prüfung im Fach Mikroökonomie im Sommersemester 2017

### Aufgaben

#### Vorbemerkungen:

**Anzahl der  
Aufgaben:  
Bewertung:**

- Die Klausur besteht aus 5 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 (17,5 Punkte)

Sie analysieren die Determinanten der Häufigkeit des Schwimmbadbesuchs mit einem geordnetem Probit Modell. Hierzu verwenden Sie folgende Variablen:

pool Schwimmbadbesuch (=1, nie; =2, manchmal; =3, oft)  
 income Jahreseinkommen (in 1.000 Euro)  
 age Alter in Jahren  
 sea =1, falls Wohnort am Meer (Distanz zwischen Wohnort und Meer < 100km); sonst =0

Ordered probit regression	Number of obs	=	8793
	LR chi2(?)	=	?
	Prob > chi2	=	?
Log likelihood = -8353.129	Pseudo R2	=	?

  

pool	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
income	.0253862	.0143191	1.773	0.076	-0.0027 .0534534
age	-.015490	.0003601	-43.5	0.000	-0.0162 -0.01481
sea	-.048421	.0054711	-8.85	0.000	-0.0591 -0.03769
/cut1	.4685231	.1095631			.253780 .6832666
/cut2	3.567936	.1123528			3.34772 3.788148

- 1.1 Interpretieren Sie den Koeffizienten von *income* statistisch und inhaltlich. (2 Punkte)
- 1.2 Berechnen Sie die kompensierende Altersvariation für eine Person, die in der Nähe des Meeres lebt. Interpretieren Sie das Ergebnis. (2 Punkte)
- 1.3 Stellen Sie die Likelihood-Funktion für das geschätzte Modell auf. Geben Sie die Likelihood-Beiträge für die drei Antwortkategorien an. (5 Punkte)
- 1.4 Was ändert sich an der Likelihood-Funktion bzw. den Likelihood-Beiträgen aus Teilaufgabe 1.3, wenn das Modell mit Konstante geschätzt wird? Beschreiben Sie kurz. (1 Punkt)
- 1.5 Definieren und erläutern Sie das McFadden  $R^2$ . Berechnen Sie das Maß für die vorliegende Schätzung. *Hinweis:* Wert der Log-Likelihood-Funktion des obigen Modells (3,5 Punkte)
  - (1) ohne die Variable *income* beträgt -8428,267.
  - (2) ohne die Variablen *age* und *sea* beträgt -8579,137.
  - (3) nur mit Konstante beträgt -8619,105.
- 1.6 Alternativ zum geordneten Probit Modell könnten Sie auch ein multinomiales Logit Modell schätzen. Geben Sie jeweils einen Vor- und Nachteil des multinomialen Logit Modells in diesem Zusammenhang an. (4 Punkte)

## Aufgabe 2 (16 Punkte)

Im Auftrag eines Marktforschungsinstituts analysieren Sie die Dauer des Samstagseinkaufs. Dazu erheben Sie Daten zu den eintreffenden Personen am Eingang des Parkplatzes eines Supermarkts zwischen 11 und 12 Uhr.

- 2.1 Ist Ihre Stichprobe ein Stock Sample oder ein Flow Sample? Erläutern Sie kurz. (1 Punkt)
- 2.2 Definieren Sie den Begriff *duration dependence* am Beispiel des Samstagseinkaufs. Unterscheiden Sie zusätzlich zwischen negativer bzw. positiver *duration dependence*. (3 Punkte)
- 2.3 Sie analysieren die Dauer des Samstagseinkaufs (gemessen in Minuten) in stetiger Zeit mit einer Weibull-Regression. Als erklärende Variable nehmen Sie das Alter und das Geschlecht einer Person auf (wobei  $female = 1$ , wenn Frau;  $female = 0$ , wenn Mann). Interpretieren Sie den Koeffizienten der Variable  $female$  inhaltlich und statistisch. (3 Punkte)

```

Weibull regression -- log relative-hazard form

Weibull regression -- accelerated failure-time form

No. of subjects =          100          Number of obs   =          100
No. of failures =           80
Time at risk    =          1136

LR chi2(2)      =          52.05
Prob > chi2     =          0.0000

-----+-----
      _t |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
    female |   .0907665   .013616     6.67   0.000   - .1174534   -.0640796
      age |  -1.049168   .1889778    -5.55   0.000   -1.419558   -.6787786
     _cons |    6.14794   .5107206    12.04   0.000    5.146946    7.148934
-----+-----
    /ln_p |   .1750802   .0860646     2.03   0.042    .0063967    .3437637
-----+-----
         p |   1.191342   .1025323          1.006417   1.410245
        1/p |   .8393897   .0722417          .7090965   .9936237
-----+-----

```

- 2.4 Erläutern Sie anhand des Beispiels, ob Ihre Daten rechts- oder linkszensiert sind. Entnehmen Sie dem Stata-Output, wie viele Personen davon betroffen sind. (2 Punkte)
- 2.5 In einer weiteren Weibull-Schätzung wird das Alter der Person ( $age$ ) mittels eines Polynoms dritten Grades abgebildet. Die Schätzung ergibt einen Log-Likelihood-Wert von  $-128,50229$ . Überprüfen Sie, ob sich der Erklärungsgehalt des Modells signifikant verbessert hat. Geben Sie die Teststatistik, Freiheitsgrade und kritischen Wert zum Signifikanzniveau  $\alpha = 0,01$  an. Berechnen Sie den empirischen Wert der Teststatistik und treffen Sie eine Testentscheidung. (5 Punkte)
- 2.6 Erklären Sie zwei Unterschiede zwischen dem Weibull-Modell und dem Cox-Proportional-Hazard-Modell. (2 Punkte)

### Aufgabe 3 (12,5 Punkte)

In der Befragung für das Marktforschungsinstitut (vgl. Aufgabe 2) haben Sie auch Informationen zu den gekauften Produkten gesammelt. Nun wollen Sie die Determinanten der Höhe der Ausgaben für Bio-Lebensmittel mit einem Tobit Modell untersuchen. Hierzu nutzen Sie folgende Variablen:

<i>bio</i>	Ausgaben für Bio-Lebensmittel (in Euro zwischen 0 und 500)
<i>age</i>	Alter in Jahren
<i>hh_small</i>	=1, falls 1 bis 2 Personen im Haushalt wohnen; sonst =0
<i>hh_medium</i>	=1, falls 3 bis 5 Personen im Haushalt wohnen; sonst =0
<i>hh_large</i>	=1, falls mehr als 5 Personen im Haushalt wohnen; sonst =0

Tobit regression		Number of obs	=	146
		LR chi2(?)	=	???
		Prob > chi2	=	???
Log likelihood = -682.816		Pseudo R2	=	0.6582

  

bio	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
age	3.241859	1.546680	2.101	0.038	0.210366 6.273352
hh_medium	-0.64176	0.348210	-1.84	0.067	-1.32425 0.040732
hh_large	-2.54361	0.589431	-0.43	0.667	-1.40964 0.900920
_cons	1.295467	0.091819	14.11	0.000	1.115499 1.475434
/sigma	0.971583	0.095347			0.784705 1.158461

- Erläutern Sie knapp den Unterschied zwischen Stützung und Zensierung auch in Bezug auf die betrachtete Stichprobe. Welcher Fall liegt in der Aufgabenstellung vor? Begründen Sie Ihre Antwort. (2 Punkte)
- Welchen marginalen Effekt beschreibt der Koeffizient der Variable *age*? Interpretieren Sie diesen Koeffizienten inhaltlich. Warum ist die Interpretation dieses Koeffizienten in der Regel wenig aussagekräftig? (3 Punkte)
- 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)

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

- Überprüfen Sie mit einem Wald-Test am 10%-Signifikanzniveau die Hypothese, dass die Ausgaben für Bio-Lebensmittel nicht von der Zahl der Haushaltsmitglieder abhängt. Geben Sie Null- und Alternativhypothesen, Freiheitsgrade, den kritischen Wert und die Testentscheidung an. (5,5 Punkte)

*Hinweise:*

- Die Teststatistik des Wald-Tests lautet:  $W = \hat{\boldsymbol{\beta}}' \hat{Var}(\hat{\boldsymbol{\beta}})^{-1} \hat{\boldsymbol{\beta}} \sim \chi^2_k$ .
- Verwenden Sie folgende geschätzte inverse Kovarianzmatrix des Schätzvektors  $\hat{\boldsymbol{\beta}} = (\hat{\beta}_{hh\_medium}, \hat{\beta}_{hh\_large})'$ :

$$\hat{Var}(\hat{\boldsymbol{\beta}})^{-1} = \begin{pmatrix} 10 & 0 \\ 0 & 5 \end{pmatrix}$$

- Runden Sie alle Zwischenschritte auf die dritte Nachkommastelle.

#### Aufgabe 4 (8 Punkte)

4.1 Beschreiben Sie die Grundidee des Maximum-Likelihood Schätzers. (1 Punkt)

4.2 Erklären Sie kurz die Invarianzeigenschaft eines ML-Schätzers  $\hat{\theta}_{ML}$  gegenüber Parametertransformationen (1 Punkt).

4.3 Gegeben sei die Log-Likelihood-Funktion für eine normalverteilte Zufallsvariable  $y_i \sim N(\mathbf{x}'_i \boldsymbol{\beta}, \sigma^2)$ :

$$\ln L(\boldsymbol{\beta}, \sigma^2) = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^N (y_i - \mathbf{x}'_i \boldsymbol{\beta})^2.$$

Zeigen Sie anhand der Transformation  $\theta^2 = 1/\sigma^2$  die Invarianzeigenschaft für  $\hat{\sigma}_{ML}^2$ . (6 Punkte)

#### Aufgabe 5 (6 Punkte)

Sie untersuchen den Zusammenhang zwischen der Kinderzahl einer Frau und ihrem Alter in Monaten (*age*). Ihnen liegen Informationen zu 200 Frauen vor. Sie schätzen ein Poissonmodell und erhalten folgenden Output:

```
Iteration 0:  log pseudolikelihood = -182.75759
Iteration 1:  log pseudolikelihood = -182.75225
Iteration 2:  log pseudolikelihood = -182.75225

Poisson regression                               Number of obs   =       200
                                                Wald chi2(1)    =       80.15
                                                Prob > chi2     =       0.0000
Log pseudolikelihood = -182.75225              Pseudo R2      =       0.2118

-----
      num_awards |          Coef.   Robust Std. Err.   z    P>|z|    [95% Conf. Interval]
-----+-----
           age |   .0070152     .0010461    6.71   0.000    .0049648    .0090656
           _cons |   5.247124     .6476195    8.10   0.000    3.977814    6.516435
-----
```

5.1 Interpretieren Sie den Koeffizienten des Alters inhaltlich und statistisch. Runden Sie auf die dritte Nachkommastelle. (2 Punkte)

5.2 Erläutern Sie kurz, was unter Überstreuung zu verstehen ist. Was bedeutet das für die Schätzergebnisse? (2 Punkte)

5.3 Erläutern Sie den Unterschied zwischen einem Poisson und einem Negbin Modell. (2 Punkte)