Klausur zur Vorlesung Methoden Verkehrsökonometrie
für Master-Studierende
WS 2012/13

Lösungsvorschlag

Aufgabe 1 (50 Punkte)

(a)
- Endogene Variable, mittlere Fahrtenhäufigkeit pro Einwohner und Jahr: metrisch bzw. kardinalskaliert
- Exogene Variablen:
  - Preis: metrisch
  - mittlere Entfernung: metrisch
  - mittlere Taktzeit: metrisch
  - Land: Qualitativ bzw. nominalskaliert
  - Nutzungsfreundlichkeit: ordinalskaliert.

(b) Parameterlineares Regressionsmodell für die obigen Faktoren:

$$y(\boldsymbol{x}) = \sum_{j=0}^{6} \beta_j x_j + \epsilon, \quad \epsilon \sim i.i.dN(0, \sigma^2)$$

mit

- $x_0 = 1$: Dummy für den konstanten Term,
- x_1: Preis in Euro,
- x_2: mittlere Entfernung,
- x_3: mittlere Taktzeit,
- x_4: Dummy für Deutschland mit Schweiz als Referenz, also $x_4 = 1$ für D und $x_4 = 0$ für CH (man kann natürlich auch Deutschland als Referenz nehmen, dann wäre $x_4 = 0$ für D und $x_4 = 1$ für CH),
- x_5: Dummy für die beste Nutzerfreundlichkeits-Stufe (die schlechteste ist Referenz),
- x_6: Dummy für die mittlere Nutzerfreundlichkeits-Stufe (die schlechteste ist Referenz).

(c) Interpretationen:

- β_1 ist die Preissensitivität: Je höher der Preis bei ceteris-paribus-Bedingungen, desto geringer die Nutzung, also $\beta_1 < 0$.
- β_2 ist die Zeitsensitivität: Je höher die Zusatzeit bei sonst gleichen bedingungen, desto geringer die Nutzung, also $\beta_2 < 0$.
- β_0 ist die Nutzungshäufigkeit, falls der ÖV nichts kostet, eine Haltestelle vor der Haustür liegt und die richtige Straßenbahn/ der richtige Bus immer genau dann ankommt, wenn man sie/ihn braucht. Also $\beta_0 > 0$.

(d) Arithmetische Mittel:

$$\bar{x}_1 = 2.183, \quad \bar{x}_2 = 11.67, \quad \bar{y} = 177.5.$$

Determinante:

$$\mathrm{Det} S = s_{11}s_{22} - s_{12}^2 = 1.600.$$

Damit

$$\hat{\beta}_1 = \frac{s_{1y}s_{22} - s_{2y}s_{12}}{\det(\mathbf{S})} = -40.4,$$

$$\hat{\beta}_2 = \frac{s_{2y}s_{11} - s_{1y}s_{21}}{\det(\mathbf{S})} = -7.40,$$

$$\hat{\beta}_0 = \bar{y} - \beta_1\bar{x}_1 - \beta_2\bar{x}_2 = 352$$

(e) Impliziter Zeitwert für die Zusatzzeiten: Eine Zusatzzeit Δx_2 muss durch eine Kostenersparnis $-\Delta x_1$ kompensiert werden, so dass sich die Nutzenhäufigkeit nicht ändert: $\Delta y = -\beta_1 \Delta x_1 + \beta_2 \Delta x_2 = 0$, also

$$\text{Zeitwert:} \quad \frac{\Delta x_1}{\Delta x_2} = \frac{\beta_2}{\beta_1} = 0.185\,\text{Euro/Minute}(= 11.1\,\text{Euro/h})$$

(f) (1) Nullhypothese H_0: Das Trivialmodell ist nicht schlechter als das volle unrestringierte Modell

(2) Testvariable:

$$F = \frac{\left(S_{\min}^{(0)} - S_{\min}\right) / \left(J - J^{(0)}\right)}{S_{\min}/(n - J - 1)} \sim F(J - J^{(0)}, n - J - 1)$$

(3) Realisierung: $n = 6$ Datenpunkte, das unrestringierte Modell hat $J = 2$ exogene Variable und das Trivialmodell $J^{(0)} = 0$ exogene Variable, also hat die F-Verteilung $J - J^{(0)} = 2$ Zähler-Freiheitsgrade und $n - J - 1 = 3$ Nenner-Freiheitsgrade. Ferner nach Aufgabenstellung $S_{\min} = 1222$ und $S_{\min}^{(0)} = 4887$. Damit

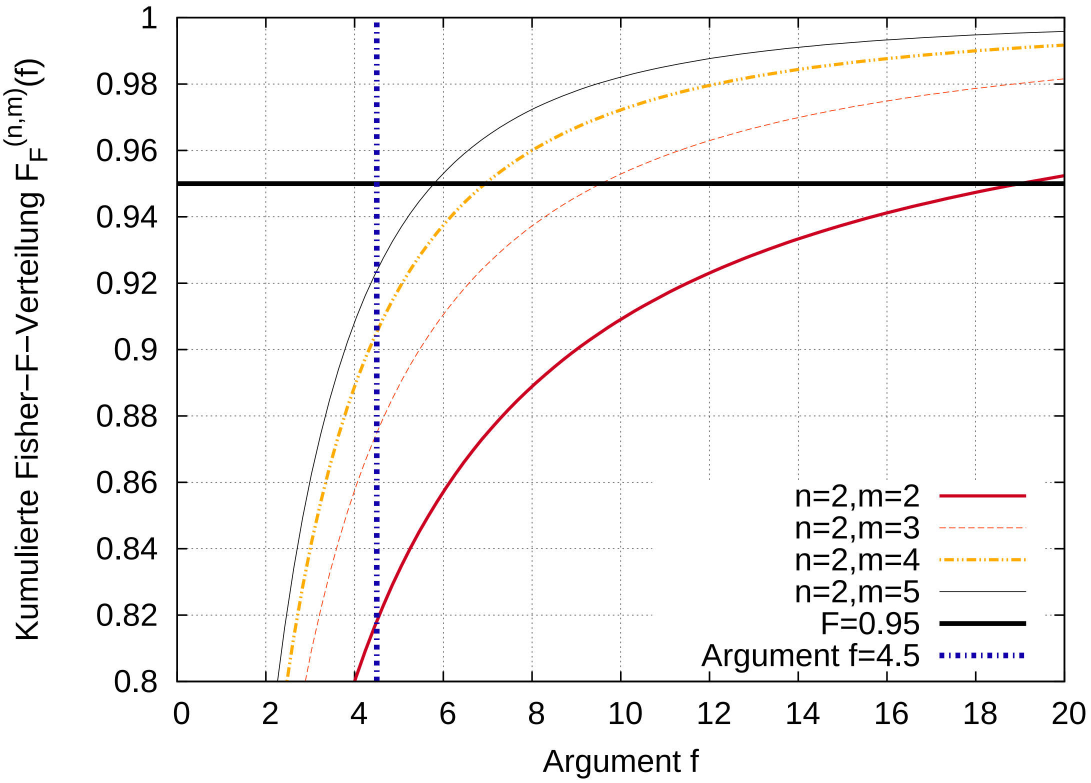
$$f = \frac{(4887 - 1222)/2}{1222/3} = 4.50$$

(4) Testentscheidung: H_0 bei $\alpha = 5\,\%$ ablehnbar, falls $f > f_{0.95}^{(2,3)} \approx 9.5$ (dort, wo die dünne gestrichelte Kurve des Plots auf der letzten Seite de rAufgabenstellung die dicke waagerechte linie bei $F = 0.95$ schneidet). Dies ist nicht der Fall, also ist dass volle Modell *nicht* signifikant besser als das Trivialmodell!

p-Wert: Ablesen der dünne gestrichelte Linie bei $f = 4.50$:

$$1 - p = 0.875 \Rightarrow p = 12.5\,\%$$

(12 % oder 13 % sind auch OK)



Aufgabe 2 (55 Punkte)

(a) Stated-Choice.

(b) Bei den ersten beiden Choice Sets haben die Alternativen 1 und 3 unveränderte Attribute. Ferner sind zwar die Absoluthäufigkeiten dieser Alternativen in Set 2 geringer (da die "dritte" Alternative 2 dann attraktiver ist), die Quotienten sind jedoch unverändert. Also gilt zumindest für diese beiden Choice Sets die IIA-Eigenschaft.

(c) Charakterisierung des Modells:

- Die Referenzalternative ist Alternative 3 (Fuß), da sie von keine alternativenspezifischne Konstante (AC) selektiert wird.
- Das Modell ist parameterlinear
- Charakterisierung der vier exogenen Faktoren:
  * $x_{1ni} = T_{ni}$: Generische Variable: Zeit
  * $x_{2ni} = K_{ni}$: Generische Variable: Kosten
  * $x_{3ni} = \delta_{i1}$: Alternativenspezifische Konstante: Bonus Alternative 1 gegenüber 3
  * $x_{4ni} = \delta_{i1}$: Alternativenspezifische Konstante: Bonus Alternative 2 gegenüber 3.

(d) Kalibrierungsbedingungen:

- Merkmalssumme 1: $X_1^{\text{data}} = \sum_{n=1}^{4} \sum_{i=1}^{3} y_{ni} T_{ni} = 10\,150$ [genau ausrechnen!]: Realisierte Gesamtzeit in Minuten
- Merkmalssumme 2: $X_2^{\text{data}} = \sum_{n=1}^{4} \sum_{i=1}^{3} y_{ni} K_{ni} = 40$: Realisierte Gesamtkosten in €
- Merkmalssumme 3: $X_3^{\text{data}} = \sum_{n=1}^{4} y_{n1} = 130$: Realisierte Gesamtzahl an ÖV-Entscheidungen
- Merkmalssumme 4: $X_4^{\text{data}} = \sum_{n=1}^{4} y_{n2} = 215$: Realisierte Gesamtzahl an Entscheidungen für das Rad.

(e) Diskussion der Schätzergebnisse:

- Die Vorzeichen von $\hat{\beta}_1$ und $\hat{\beta}_2$ entsprechen den Erwartungen, da beide negativ sein sollen (Zusatzkosten und Zusatzzeiten mindern den Nutzen). Die alternativenspezifischen Konstanten β_3 und β_4 können beide Vorzeichen haben. Hier wird der ÖV gegenüber Schusters Rappen positiv und das Rad negativ bewertet, was nicht *a priori* unplausibel ist.
- Impliziter Zeitwert:

$$\hat{\beta}_1/\hat{\beta}_2 = 0.119\,\text{Euro/Minute} = 7.1\,\text{Euro/h}.$$

- Impliziter Zeitwert durch direkten Vergleich der Sets 3 und 4: 2 Euro entsprechen 20 Minuten, also 0.1 Euro/Minute bzw. 6 Euro/h.

(f) Testvariable “Schätzer in Einheiten der Standardabweichung”:

$$z = \frac{-0.78}{0.37} = -2.1$$

Die Standardnormalverteilung geht aus den Student-t-Verteilungen im Limes unendlich vieler Freiheitsgrade hervor, so dass die beigefügte Tabelle der Standardnormalverteilung verwendet werden kann.[1] Der Schwellwert bei einer Fehlerwahrscheinlichkeit von 5 % beträgt $z_{1-\alpha/2} = z_{0.975} = 1.96$. Die Nullhypothese H_0: “β_4 ist nicht signifikant von null verschieden” kann (bei einer Fehlerwahrscheinlichkeit von 5 %) abgelehnt werden, falls $|z| > 1.96$. Dies ist der Fall, also $\hat{\beta}_4$ signifikant von null verschieden (der p-Wert liegt allerdings bei 4 %, die Signifikanz ist also nicht besonders hoch).

Aufgabe 3 (15 Punkte)

(a) Versuchsplan des *full factorial design*:

$$\begin{aligned}
&\text{Choice Set 1:}(-20\,\text{min}, -1\ €),\\
&\text{Choice Set 2:}(-20\,\text{min}, 1\ €),\\
&\text{Choice Set 3:}(0\,\text{min}, -1\ €),\\
&\text{Choice Set 4:}(0\,\text{min}, 1\ €),\\
&\text{Choice Set 5:}(20\,\text{min}, -1\ €),\\
&\text{Choice Set 6:}(20\,\text{min}, 1\ €).
\end{aligned}$$

(b) Die Zeit- und Kostendifferenzen sind bei den drei Choice Sets $(-20\,\text{min}, -1\ €)$, $(20\,\text{min}, -1\ €)$, und $(20\,\text{min}, 1\ €)$ positiv korreliert. Also ist dieser Versuchsplan *nicht* nach dem orthogonalem Design aufgebaut.

[1] In der Literatur wird dieser Test im Zusammenhang mit der Maximum-Likelihood-Methode meist *t-Test* genannt, obwohl er in der Regel in der asymptotischen Näherung unendlich vieler Freiheitsgrade ausgewertet wird, so dass man eigentlich von einem *z-Test* (d.h. die Testvariable ist standardnormalverteil) sprechen müsste.