Klausur zur Vorlesung Methoden Verkehrsökonometrie

für Master- und Diplom-Studierende

WS 2011/12

Lösungsvorschlag

Aufgabe 1 (50 Punkte)

(a)
- Skalierung: x_1 und y kardinal bzw metrisch, x_2 nominal bzw. qualitativ.
- keine überflüssigen Var, mögliche fehlende Faktoren z.B. Haltestellendichte
- Linearität OK
- Möglicher Strukturbruch zwischen 1930 und 1960, da es vorher nur eine Geschwindigkeitsklasse gab

(b) i.i.d $\Rightarrow$ 1. unabhängig und 2. identisch verteilt, $N(...)$ $\Rightarrow$ 3. normalverteilt, $N(0,\sigma^2)$ $\Rightarrow$ 4. Erwartungswert $\mu = 0$. Damit enthält $\epsilon_i \sim i.i.d\ N(0,\sigma^2)$ alle vier Spezifikationsbedingungen.

(c) Die Zugklasse und das Jahr sind offensichtlich keine Zufallsvariablen. Ferner kann man die Zugklasse nicht als lineare Funktion des Jahres schreiben, also gibt es keine Multikolinearität bei den exogenen Daten. Schließlich ist der Stichprobenumfang $n = 8$ zwar klein, aber nicht nur gleich sondern sogar größer als der formale Mindestrumfang $J + 2 = 4$.

(d) β_0: erwartete Geschwindigkeit der langsameren Zugklasse im Jahr 1 900, β_1: Anstieg der erwarteten Geschwindigkeit jeder der beiden Zugklassen pro Jahr (anhand der Daten positiv erwartet), β_2: Erwartete Geschwindigkeitsdifferenz zwischen schnellen und langsameren Zügen (aufgrund der Definition ("schneller ist schneller als langsamer") und der Daten wird ein positiver Wert erwartet.

(e) Aus den Daten:

$$\bar{x}_1 = 63.75, \quad \bar{x}_2 = 0.375, \quad \bar{y} = 83$$

Aus den in der Aufgabenstellung gegebenen Varianzen und Kovarianzen, mit einer Formel aus dem Skript:

$$\det \mathbf{S} = s_{11}s_{22} - s_{12}^2 = \underline{\underline{251.4}}$$

und damit

$$\hat{\beta}_1 = \frac{s_{1y}s_{22} - s_{2y}s_{12}}{\det(\mathbf{S})} = \underline{\underline{0.435}},$$

$$\hat{\beta}_2 = \frac{s_{2y}s_{11} - s_{1y}s_{21}}{\det(\mathbf{S})} = \underline{\underline{28.7}},$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}_1 - \hat{\beta}_2\bar{x}_2 = \underline{\underline{44.5}}.$$

(f) Jahr 2020, also $x_1 = 120$: Langsame Züge:

$$\hat{y}(120,0) = \hat{\beta}_0 + 120\hat{\beta}_1 = 92.2\,\text{km/h}$$

Schnelle Züge:

$$\hat{y}(120,1) = \hat{\beta}_0 + 120\hat{\beta}_1 + \hat{\beta}_2 = 121.2\,\text{km/h}$$

(g) Gegeben: Varianzschätzer von $\hat{\beta}_1$:

$$\hat{V}(\hat{\beta}_1) = 0.0078$$

Damit ergibt sich das 5%-Konfidenzintervall zu

$$\beta_1 \in [\hat{\beta}_1 - \Delta\hat{\beta}_1, \hat{\beta}_1 + \Delta\hat{\beta}_1]$$

mit

$$\Delta\hat{\beta}_1 = \sqrt{\hat{V}(\hat{\beta}_1)}t^{(8-3)}_{1-0.05/2} = \sqrt{\hat{V}(\hat{\beta}_1)}t^{(5)}_{0.975} = 0.0883 * 2.57 = 0.228$$

also

$$\mathrm{KI}(\beta_1)^{\alpha=5\%} = [0.183, 0.637].$$

(h)

$$\begin{aligned} y(\boldsymbol{x}) &= \beta_0 + \beta_1 x_1 \begin{cases} 1 & x_2 = 0 \\ 0 & \text{sonst.} \end{cases} + \beta_2 x_1 \begin{cases} 0 & x_2 = 0 \\ 1 & \text{sonst.} \end{cases} + \beta_3 x_2 + \epsilon \\ &= \beta_0 + \beta_1 x_1(1 - x_2) + \beta_2 x_1 x_2 + \beta_3 x_2 + \epsilon \end{aligned}$$

mit

- β_1: Anstiegsrate der erwarteten Geschwindigkeit der langsameren Zugklasse pro Jahr
- β_2: Anstiegsrate der erwarteten Geschwindigkeit der schnelleren Zugklasse pro Jahr
- β_3: Differenz der erwarteten Geschwindigkeiten der beiden Zugklasse im Jahr $x_1 = 0$, also 1900.
- β_0 unverändert (erwartete Geschwindigkeit der langsameren Zugklasse im Jahr 1900)

Man könnte das Modell völlig äquivalent sogar einfacher formulieren:

$$y(\boldsymbol{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_1 x_2 + \beta_3 x_2 + \epsilon$$

Die Parameter β_0, β_1 und β_3 behalten ihre Bedeutung während β_2 nun die *Differenz* der Geschwindigkeitsanstiegsraten der schnelleren bezüglich der langsameren Zugklasse bedeutet.

Aufgabe 2 (20 Punkte)

(a) Matrix des vollen Aufwandes $\mathbf{B} = (\mathbf{1} - \mathbf{A})^{-1}$ bei 2×2 Matrizen:

$$\begin{aligned}\mathbf{B} &= \begin{pmatrix} 1-A_{11} & -A_{12} \\ -A_{21} & 1-A_{22} \end{pmatrix}^{-1} = \frac{1}{(1-A_{11})(1-A_{22})} \begin{pmatrix} 1-A_{22} & A_{12} \\ A_{21} & 1-A_{11} \end{pmatrix} \\ &= \begin{pmatrix} 1.13 & 0.032 \\ 0.64 & 1.45 \end{pmatrix}\end{aligned}$$

(b) – einen Euro ÖPNV-Nachfrage: *Nachfrage*vektor (in €)

$$\boldsymbol{y} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Insgesamt muss von der Wirtschaft dafür der Erzeugungsvektor mit den Komponenten

$$\begin{aligned} x_1 &= B_{11}y_1 + B_{12}y_2 = 1.125, \\ x_2 &= B_{21}y_1 + B_{22}y_2 = 0.643 \end{aligned}$$

erzeugt werden. Da nur CO_2 betrachtet wird, reduziert sich die Emissionsfaktorenmatrix $\mathbf{C}$ der direkten Emissionen auf einen Emissionsfaktorenvektor $\boldsymbol{c}$ der beiden Sektoren. Dieser ist nach Aufgabenstellung durch

$$\boldsymbol{c}^{\mathrm{T}} = (0.1\,\mathrm{kgCO_2}, 0.5\,\mathrm{kgCO_2})$$

gegeben. Damit ist der CO_2 Ausstoß pro Euro Nachfrage im Verkehrssektor gegeben durch

$$e_{\mathrm{CO2}} = \boldsymbol{c}^{\mathrm{T}} \cdot \boldsymbol{x} = \boldsymbol{c}^{\mathrm{T}} \cdot \mathbf{B}\,\boldsymbol{y} = 0.1x_1 + 0.5x_2 = 0.43(\text{ kg CO}_2/\text{Euro})$$

– einen Euro sonstige Nachfrage: *Nachfrage*vektor (in €)

$$\boldsymbol{y} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Insgesamt muss von der Wirtschaft dafür der Erzeugungsvektor mit den Komponenten

$$\begin{aligned} x_1 &= B_{11}y_1 + B_{12}y_2 = 0.032, \\ x_2 &= B_{21}y_1 + B_{22}y_2 = 1.45 \end{aligned}$$

erzeugt werden. Damit ist der CO_2 Ausstoß pro Euro Nachfrage im sonstigen Sektur gegeben durch

$$e_{\mathrm{CO2}} = \boldsymbol{c}^{\mathrm{T}} \cdot \boldsymbol{x} = \boldsymbol{c}^{\mathrm{T}} \cdot \mathbf{B}\,\boldsymbol{y} = 0.1x_1 + 0.5x_2 = 0.73(\text{ kg CO}_2/\text{Euro}).$$

Der ÖPNV erzeugt also insgesamt weniger Emissionen pro Euro, aber der Unterschied ist deutlich geringer als bei den Direktemissionen.

Aufgabe 3 (50 Punkte für Diplom, 20 Punkte für Master)

[*Hinweis:* Aufgabenteile (a)-(d) und (h) waren nur für Diplomstudierende, die restlichen drei Fragen für alle]

(a) Es ist eine Revealed-Choice-Befragung, da bereits durchgeführte Entscheidungen erfragt wurden.

(b)
- Alternativenspezifische Konstante: δ_{n1}
- Sozioökonomisch: Alter des bisherigen Kfz und ob er als Neu- oder als Gebrauchtwagen gekauft wurde
- Generisch: Eigentlich der Rabatt. Doch da dieser vom Händler wieder individuell an die Person gekoppelt wurde und er für die Alternative 2 (Nichtkauf) nicht relevant ist, hat man doch wieder die Struktur einer sozioökonomischen Variablen.

(c) Ja, das Fahrzeugalter muss alternativenspezifisch formuliert werden, da es eine sozioökonomische Variable ist bzw. sich nicht zwischen den Alternativen unterscheidet.

(d)
- Merkmalssumme X_1: Gesamte Zahl der gekauften Neuwagen
- Merkmalssumme X_2: Summe der Fahrzeugalter der Kunden, welche ein neues Auto kaufen (bei den Daten sind dies 3 Kunden)
- Merkmalssumme X_3: Summe des Rabatts, welcher den Käufern gewährt wurde (bei den Daten sind dies 3 Käufer)
- Merkmalssumme X_4: Wieviel der Käufer auch bisher einen Neuwagen fuhren

also

$$\begin{aligned}
\text{Daten:} X_1^{\text{data}} &= 3, \quad \text{Modell } \boldsymbol{\beta} = 0 : X_1^{\text{mod}} = 10/2 = 5 \\
\text{Daten:} X_2^{\text{data}} &= 27, \quad \text{Modell } \boldsymbol{\beta} = 0 : X_2^{\text{mod}} = 75/2 = 37.5 \\
\text{Daten:} X_3^{\text{data}} &= 8, \quad \text{Modell } \boldsymbol{\beta} = 0 : X_3^{\text{mod}} = 18/2 = 9 \\
\text{Daten:} X_4^{\text{data}} &= 2, \quad \text{Modell } \boldsymbol{\beta} = 0 : X_4^{\text{mod}} = 5/2 = 2.5
\end{aligned}$$

(e) [=(a) für Master] Weil es unattraktiv ist, ein neues Auto zu kaufen, wenn man bereits einen 0 Jahre alten Wagen besitzt und vom Händler keinen Rabatt bekommt.

(f) [=(b) für Master] Einsetzen der Wahrscheinlichkeit für das Binomial-Logitmodell für $T_n = 5$ (5 Jahre alt), $R_n = 2$ (2 000 € Rabatt, Neuwagen-Dummy=1 ("der Alte war damals ein Neuer") unter Verwendung des Parameterschätzers $\hat{\boldsymbol{\beta}} = (-9.2, 0.35, 2.2, 1.3)^{\mathrm{T}}$ der Aufgabenstellung:

$$\begin{aligned}
V_1 &= \hat{\beta}_1 + 5\hat{\beta}_2 + 2\hat{\beta}_3 + \hat{\beta}_4 = -1.75, \\
N &= e^{V_1} + e^{V_2} = e^{-1.75} + 1 = 1.174, \\
P_1 &= \frac{e^{V_1}}{N} = \underline{\underline{0.148}}
\end{aligned}$$

(g) [=(c) für Master] Likelihood-Ratio Test für das Modell M1 gegenüber dem Ausgangsmodell M in den bekannten 4 Stufen:

(1) H_0: Modell M ist nicht "besser" als M1

(2) Die Test-Statistik ist das Likelihood-Verhältnis ("Likelihood-ratio")

$$\lambda = 2\left[\ln L_{\max} - \ln L_{\max}^{\mathrm{M1}}\right] \sim \chi^2(1)$$

(3) Die Realisierung ist nach den Angaben der Aufgabenstellung,

$$\ln L_{\max} = -3.43, \quad \ln L_{\max}^{\mathrm{M1}} = -3.52, \quad \ln L_{\max}^{\mathrm{M2}} = -5.78,$$

gegeben durch

$$\lambda_{\mathrm{data}}^{\mathrm{M1}} = 2 * (-3.43 + 3.52) = 0.18$$

(4) Entscheidung: H_0 bei einer Fehlerwahrsch. von 10 % ablehnbar, falls $\lambda_{\mathrm{data}}^{\mathrm{M1}} > \chi^2_{0.9,1} = 2.706$. Dies ist keine wahre Aussage, also ist H_0: "M1 ist genau so aussagekräftig wie M" nicht ablehnbar.

Der Vergleich von M mit M2 geht analog, nur dass hier wegen des Unterschieds von 2 Modellparametern (M hat vier, M2 hat zwei) zwei Freiheitsgrade bei der χ^2-Verteilung anzusetzen sind:

(1) H_0: Modell M ist nicht "besser" als M2

(2) $\lambda = 2\left[\ln L_{\max} - \ln L_{\max}^{\mathrm{M2}}\right] \sim \chi^2(2)$

(3) Realisierung $\lambda_{\mathrm{data}}^{\mathrm{M2}} = 2 * (-3.43 + 5.78) = 4.70$

(4) Entscheidung: H_0 bei $\alpha = 10\,\%$ ablehnbar, falls $\lambda_{\mathrm{data}}^{\mathrm{M2}} > \chi^2_{0.9,2} = 4.605$. Dies ist erfüllt, H_0 ist also be 10 % Fehlerwahrscheinlichkeit gerade noch ablehnbar. [Bei 5 % allerdings nicht mehr, $p = 9.56\,\%$]

(h) Die nominalskalierte exogene Variable "bisher Neuwagen" wird zur dreiwertigen nominalskalierten Variable "Kauf des letzten Kfz" mit den Ausprägungen 0: "Gebrauchtwagen", 1: "Neuwagen" und 2: "Bisher kein Kfz". Wie immer (Modelle der diskreten Wahltheorie, aber auch Regressionsmodelle!) bildet man bei m-wertigen nominalskalierten Variablen $m-1$ quasilineare Einflussfaktoren, also hier 3-1=2 Faktoren. Die neuen deterministischen Nutzenfunktionen lauten also nun

$$V_{ni} = \beta_1\delta_{n1} + \beta_2 T_n\delta_{n1} + \beta_3 R_n\delta_{n1} + \beta_4\delta_{n1}\begin{cases} 1 & \text{Neuwagen} \\ 0 & \text{sonst.}\end{cases} + \beta_5\delta_{n1}\begin{cases} 1 & \text{Gebrauchtwagen} \\ 0 & \text{sonst.}\end{cases}$$

Hierbei gibt β_4 die Attraktivität eines Neukaufs für bisherige Neuwagenbesitzer gegenüber Kfz-Nichtbesitzern und β_5 die Attraktivität eines Neukaufs für bisherige Gebrauchtwagenbesitzer gegenüber den Kfz-Nichtbesitzern unter *ceteris-paribus-Bedingungen* an.