Klausur zur Vorlesung Verkehrsökonometrie für den Bachelor-Studiengang, SS 2019

Lösungsvorschlag

Insgesamt 120 Punkte

Aufgabe 1 (15 Punkte)

Verkehrserzeugung:

(a) – exogene Variable: personenbezogene Merkmale n_i^g für jede Quelle-Ziel-Gruppe (QZG) g sowie wegezielbezogene Strukturmerkmale S_i^g
   – endogene Variable: Quell- und Zielsummen Q_i^g bzw. Z_i^g für jede QZG und jeden Bezirk
   – Parameter: Spezifische Verkehrsaufkommen σ^g und Erzeugungsraten ϵ^g.

(b) Lineare Gleichungen sind bei QZG vom Typ I die Quellsummen,

$$Q_i^g = \sigma^g n_i^g,$$

und auch die Zielsummen der QZG vom Typ II: Erstere sind beispielsweise linear in den personenbezogenen Merkmalen n_i^g (und den Parametern σ^g)

Nichtlineare Gleichungen sind bei QZG vom Typ I die Zielsummen:

$$Z_i^g = \epsilon^g S_i^g \frac{\sum_j \sigma^g n_j^g}{\sum_j \epsilon^g S_j^g}$$

und auch die Quellsummen der QZG vom Typ II sowie alle Summen der QZG vom Typ III: Sie sind aufgrund der Bedingung der räumlichen Geschlossenheit $\sum_i Q_i = \sum_i Z_i$ und der dadurch entstehenden Quotienten beim gemeinsamen Korrekturfaktor nichtlinear in den exogenen Variablen (und den Parametern)

Hinweis: jeweils eine einzige Gleichung samt Begründung ergibt volle Punktzahl.

(c) Verkettung bedeutet, dass die endogenen Variablen eines Modells an die exogenen des nächsten verkettet werden. Die endogenen Variablen der Erzeugung, die Quell- und Zielsummen, dienen gleichzeitig als exogene Variable der Zielwahl bzw. Verteilung.

(d) Die exogenen Variablen charakterisieren die *konkreten Bezirke* und sind deshalb von Bezirk zu Bezirk und von Untersuchung zu Untersuchung *unterschiedlich*.

Die Modellparameter charakterisieren im wesentlichen *gleichbleibende* Merkmale des Mobilitäts*verhaltens*, wie Zahl der Wege pro Tag und QZG und die Mobilitätskennziffer (Gesamtzahl an Wegen pro Tag).

Generell (nicht verlangt, ggf Extrapunkte): Die Trennung zwischen exogenen Variablen und Parametern ist die Grundlage der Aussage- und Voraussagefähigkeit der Modellierung. Die Parameter eines einmal geschätzten Modells sind fest und können auf eine Vielzahl unterschiedlicher Systeme (hier: Untersuchungsgebiete) angewandt werden.

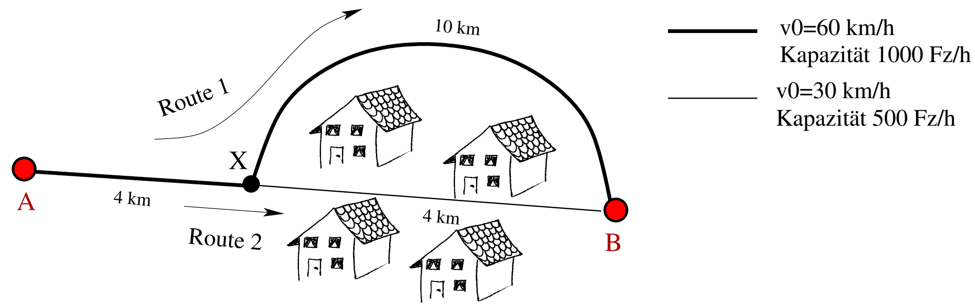
Hinweise:

- Immer mit der konkreten Aufgabenstellung argumentieren, keine Skripttexte kopieren! Beispielsweise ist bei (a) *nicht* verlangt, zu erläutern, was exogene Variable sind, sondern es sollen *Beispiele aus dem Verkehrserzeugungsmodell* angegeben werden

- Auch bei Teil (b): Es ist *nicht* gefragt, was lineare Modelle sind, sondern es sollen lineare Teilmodelle *der Verkehrserzeugung* angegeben werden

Aufgabe 2 (30 Punkte)

Um den Durchgangsverkehr von A nach B durch die Innenstadt (Route 2) zu reduzieren, hat eine Stadt eine Umgehungsstraße in Betrieb genommen (Route 1).



(a) Die Kante von A nach X wird von beiden Routen, also der gesamten Nachfrage benutzt. Damit hängt die Reisezeit auf dieser Kante nicht von der Routenaufteilung ab und fällt damit in der Betrachtung heraus.

(b) Betrachtung der Reisezeiten beider Routen in Minuten nur für die Teilverbindung X-B:

$$\begin{aligned} T_1 &= 10\left(1+\frac{Q_1}{1\,000\,\mathrm{Fz/h}}\right) = 10\left(1+\frac{Q_{\mathrm{AB}}w_1}{1\,000\,\mathrm{Fz/h}}\right) = 10+10qw_1 \\ T_2 &= 8\left(1+\frac{Q_2}{500\,\mathrm{Fz/h}}\right) = 8\left(1+\frac{Q_{\mathrm{AB}}(1-w_1)}{500\,\mathrm{Fz/h}}\right) = 8+16q(1-w_1) \end{aligned}$$

Das Nutzergleichgewicht im Falle einer Nutzung beider Routen, $0 < w_1 < 1$, lautet:

$$\begin{aligned} T_1 &= T_2 \\ 10+10qw_1 &= 8+16q-16qw_1 \\ 26qw_1 &= 16q-2 \end{aligned}$$

Für sehr kleine Nachfragen ($q < 1/8$) wird $w_1 < 0$, also ist die Annahme einer Nutzung beider Routen nicht mehr erfüllt und wir erhalten insgesamt

$$w_1^{\mathrm{UE}} = \begin{cases} \frac{8q-1}{13q} & q > \frac{1}{8} \\ 0 & \text{sonst.} \end{cases} \tag{1}$$

Bei linearen CR-Funktionen gilt für das Systemoptimum

$$w_1^{\mathrm{SO}}(q) = w_1^{\mathrm{UE}}(2q) = \begin{cases} \frac{16q-1}{26q} & q > \frac{1}{16} \\ 0 & \text{sonst.} \end{cases} \tag{2}$$

(c) Damit ein Systemoptimum (SO) bezüglich der Reisezeit stabil ist, muss für die Verkehrsteilnehmer bezüglich der Gesamtkosten (Zeit+Geld) ein Nutzergleichgewicht vorliegen. Der Zeitunterschied zwischen Route 1 und 2 muss daher in Geld "aufgewogen" werden: Mit einem Zeitwert λ =30 €/h=0.5 €/Min ergibt sich für

$Q_{\rm AB} = 1\,000\,\mathrm{Fz/h}$, also $q = 1$ für eine Maut M_2 auf Route 2 folgender Betrag (in €):

$$M_2 = \lambda({T_1}^{\rm SO} - {T_2}^{\rm SO}) = \lambda\,1\,\mathrm{Min} = \ 0.5\,€$$

Hinweis (nicht verlangt): Dieser Reisezeitunterschied und damit diese Maut gilt für alle Verkehrsnachfragen oberhalb von $q/16$ entsprechend $63\,\mathrm{Fz/h}$, also fast immer:

$$\begin{aligned} {T_1}^{\rm SO} - {T_2}^{\rm SO} &= 10 + 10q{w_1}^{\rm SO} - 8 - 16q + 16q{w_1}^{\rm SO} \\ &= 2 - 16q + 26q{w_1}^{\rm SO} \\ &= 2 - 16q + (16q - 1) \\ &= 1 \end{aligned}$$

Auch im SO wird die Innenstadtroute bei $Q_{\rm AB} = 1\,000\,\mathrm{Fz/h}$ noch benutzt. Für eine hundertprozentige Vermeidung muss der Zeitunterschied bei ausschließlicher Nutzung der Route 1 ($w_1 = 1$) in Geld kompensiert werden:

$${M_2}^{100\%} = \lambda(T_1(q = w_1 = 1) - T_2(q = w_1 = 1)) = \lambda(20\,\mathrm{Min} - 8\,\mathrm{Min}) = \ 6\,€$$

Hinweis (nicht verlangt): Im Ggs zur Maut beim SO hängt diese Maut stark von der Verkehrsnachfrage ab:

$$T_1(w_1 = 1) - T_2(w_1 = 1) = 10 + 10q - 8 = 2 + 10q$$

Aufgabe 3 (25 Punkte)

Im Rahmen einer Bachelorarbeit wird eine offene internetbasierte Umfrage zur Verkehrsmittelwahl der erwachsenen Einwohner im Raum Dresden für die Fahrt zur Arbeit bzw. Ausbildung durchgeführt:

Berufsstatus	Altersklasse	zu Fuß	Rad	ÖV	MIV
Student	< 30 Jahre	5	20	24	1
Student	≥ 30 Jahre	1	1	4	4
Nicht-Student	< 30 Jahre	3	3	9	5
Nicht-Student	≥ 30 Jahre	1	1	8	10

(a) Abgrenzungen der Grundgesamtheit (GG):
   - räumliche: Raum Dresden
   - zeitlich: jetzt
   - sachlich: nur erwachsene (erwerbstätige bzw. studierende/in Ausbildung stehende) Personen

(b) Die Grundgesamtheit besteht aus allen erwachsenen Einwohnern, während die Stichprobe aufgrund der sozialen Umgebung (Universität) und der unkontrollierten Auswahl ("offene internetbasierte Umfrage") wohl systematisch mehr Studenten enthält, als es nach dem Anteilswert in der GG zu erwarten ist.

(c) Bis auf die Rad-Verfügbarkeit (Räder müssen nicht angemeldet werden, zumal ist Besitz nicht gleichbedeutend mit Verfügbarkeit) ist alles in Registern abgelegt, also in der GG bekannt. Studenten und jüngere sind tendenziell (!) ökologisch bewusster und haben weniger Geld, was sich auf die Wahl des Kfz als Modus auswirkt. Ebenfalls wirkt sich offensichtlich die Verfügbarkeit oder, als Proxy, der Besitz von Verkehrsmitteln auf die Wahl aus. Beim Geschlecht ist dies nicht a priori klar (sowohl ein begründetes "Ja" als auch "Nein" ergibt volle Punktzahl). Zusammenfassung:

Merkmal	in GG exakt bekannt	für Umfrage relevant
Student-Sein	ja	ja
Alter	ja	ja
Geschlecht	ja	vermutlich
Kfz-Besitz	ja	ja
Rad-Verfügbarkeit	nein	ja

(d) Tabelle der relativen Stichproben-Häufigkeiten:

Berufsstatus	Altersklasse	zu Fuß	Rad	ÖV	MIV	$\sum = f_{\text{St}}$
Student	< 30 Jahre	0.05	0.20	0.24	0.01	0.50
Student	≥ 30 Jahre	0.01	0.01	0.04	0.04	0.10
Nicht-Student	< 30 Jahre	0.03	0.03	0.09	0.05	0.20
Nicht-Student	≥ 30 Jahre	0.01	0.01	0.08	0.10	0.20
$\sum =$ ungewichteter Modal Split		0.10	0.25	0.45	0.20	1.00

Tabelle der relativen Häufigkeiten in der GG und daraus resultierende Entzerrungsfaktoren:

Berufsstatus	Altersklasse	θ	f (vorhergehende Tabelle)	$E = \theta/f$
Student	$<$ 30 Jahre	0.08	0.50	0.16
Student	$\geq$ 30 Jahre	0.02	0.10	0.20
Nicht-Student	$<$ 30 Jahre	0.18	0.20	0.90
Nicht-Student	$\geq$ 30 Jahre	0.72	0.20	3.60
$\sum$		1.00	1.00	

Die entzerrenden Schätzer für den Modal-Split für Modus j erhält man nun durch eine mit den Entzerrungsfaktoren E_i gewichtete Spaltensumme der kombinierten relativen Häufigkeiten f_{ij} dafür, dass ein Befragter der sozioökonomischer Schicht i angehört *und* den Modus j wählt:

$$f_j^{\text{ E}} = \sum_i f_{ij} E_j$$

also

$$\begin{aligned}
f_{\text{Fuss}}^{\text{ E}} &= 0.05 * E_1 + 0.01 * E_2 + 0.03 * E_3 + 0.01 * E_4 = 0.073, \\
f_{\text{Rad}}^{\text{ E}} &= 0.20 * E_1 + 0.01 * E_2 + 0.03 * E_3 + 0.01 * E_4 = 0.097, \\
f_{\text{OEV}}^{\text{ E}} &= 0.24 * E_1 + 0.04 * E_2 + 0.09 * E_3 + 0.08 * E_4 = 0.415, \\
f_{\text{MIV}}^{\text{ E}} &= 0.01 * E_1 + 0.04 * E_2 + 0.05 * E_3 + 0.10 * E_4 = 0.415.
\end{aligned}$$

Hinweis: Dies unterscheidet sich deutlich von den einfachen Summen $\sum_i f_{ij}$ der nicht entzerrten Modal-Split-Schätzer, die direkt in der letzten Zeile der ersten Tabelle angegeben sind.

Aufgabe 4 (50 Punkte)

Ein Verkehrswissenschaftler analysiert die von Nutzern gegebenen Eingaben eines verkehrsmittelübergreifenden Buchungsportales für die Verbindung Dresden-München (one-way). Der Forscher hat Zugriff auf die Anfragen aller Nutzer für diese Verbindung sowie auf die durch das Buchungsportal tatsächlich vermittelten Reisen. Auf der Webseite werden für die drei Verkehrsmittel

$i = 1$: Fernbus $\quad$ $i = 2$: Bahn $\quad$ $i = 3$: Flugzeug

jeweils 5 Verbindungen mit Attributen wie Reisedauer, Kosten, Zahl der Umstiege und Startzeitpunkt (günstig oder nicht) angegeben.

(a) Es ist eine Revealed-Choice-Situation, da tatsächlich realisierte Entscheidungen (gebuchte Reisen) erhoben werden.

(b) Welche Skalierung hat Die Variable "Startzeitpunkt" hat die Werte "günstig" und "nicht günstig". Dies ist ordinalskaliert.

Hinweis: Bei dichotomen (zweiwertigen) Variablen gibt es prinzipiell keinen formalen Unterschied zwischen ordinal- und nominalskaliert und, mittels einer Dummyvariablen, auch keinen formalen Unterschied zu einer metrischen Skalierung. Nominalskaliert bzw. qualitativ wird daher auch als "richtig" bewertet.

(c)
- *Exklusivität:* Ja, man kann nur eine Reise wählen.
- *Vollständigkeit:* Nein, man könnte ja auch gar keine Reise buchen. Dies würde durch eine 16. Option "kein Angebot passt mir" gelöst werden[1]
- *Unabhängigkeit* bzw. *wesentliche Verschiedenheit:* Nein, jede der 5 Angebote innerhalb eines der Verkehrsmodi hat gemeinsame Eigenschaften eben dieser Modi (z.B. langwierige Sicherheitskontrolle vor dem Flug oder keine Bewegungsfreiheit im Bus) und die jeweiligen Alternativen sind daher hochgradig korreliert. Lösung: Jede 5-er-Gruppe auf ein "bestes" Angebot kondensieren.

(d) Bei einer Modellierung mit der diskreten Wahltheorie müssen die Attribute *aller* Alternativen, nicht nur der gewählten, bekannt sein. Dies ist durch die Zusammenfassung in drei Fünfergruppen nicht gegeben, da man bei den beiden nichtgewählten Gruppen ohne zusätzliche Befragung nicht weiß, auf welche Alternative man sich beziehen soll.[2]

(e) Ein *No-Response*-Anteil ist insofern problematisch, da durchaus Grund zur Annahme besteht, dass die eher zeitsensitiven Personen einerseits bevorzugt das Flugzeug wählen und andererseits eher unwillig zur Beantwortung der E-Mail sind. Damit stellt die verbleibende Gruppe eine verzerrte Stichprobe dar.

(f) Die Alternative $i = 3$ (Flug) hat keine AC und ist daher die Referenzalternative. Die Parameter β_1 und β_2 geben den subjektiven "Anfangsbonus" für eine Bus- bzw. Bahnfahrt gegenüber einem Flug dar, wenn alle weiteren Attribute insgesamt für alle Alternativen eine gleiche Bewertung geben würden.

---

[1] Die durchaus empfohlene Möglichkeit, direkt beim Verkehrsanbieter zu buchen, wird hier außen vor gelassen.

[2] Eleganter geht dies mit dem *Nested-Logit-Modell*, was derartige verschachtelte Entscheidungen berücksichtigt, siehe Master-Vorlesung "Methoden Verkehrsökonometrie".

(g) Die Zeitsensitivitäten unterscheiden sich für die drei Modi, sind also alternativenspezifisch modelliert. Hingegen ist die Geldbewertung immer gleich, was einer generischen Spezifikation entspricht. Umstiege sind generell nachteilig, also wird $\beta_U < 0$ erwartet. [3] Der Dummy W soll =1 sein, wenn die Wunschzeit angeboten wird und 0 sonst, also $W > 0$.

(h) Für die spontane Fahrt nach München ergeben sich folgende deterministische Nutzenfunktionen:

$$\begin{aligned} V_1 &= -0.5 + 6.5\beta_{T_1} + 25\beta_K + \beta_W = -8.75, \\ V_2 &= 1.5 + 5.83\beta_{T_2} + 84\beta_K + 2\beta_{U_2} + \beta_W = -7.78, \\ V_3 &= 3.83\beta_{T_3} + 80\beta_K + \beta_{U_3} = -10.64 \end{aligned}$$

und damit der Nenner

$$N = \sum_{i=1}^{3} e^{V_i} = 0.000600$$

und die Auswahlwahrscheinlichkeiten

$$P_{\text{ Bus}} = \frac{e_1^V}{N} = 26.4\,\%, \quad P_{\text{ Bahn}} = \frac{e_2^V}{N} = 69.6\,\%, \quad P_{\text{ Flug}} = \frac{e_3^V}{N} = 4.0\,\%.$$

(i) Es soll mit kontrollierbarer Fehlerwahrscheinlichkeit $\alpha = 5\,\%$ gezeigt werden, dass das Umsteigen als Nachteil wahrgenommen wird. Also muss man auf *positives* $\beta_{U_3} > 0$ testen:

1. $H_0 : \beta_{U_3} > 0$
2. Testvariable $T = \frac{\hat{\beta}_{U_3}}{\sqrt{V(\hat{\beta}_{U_3})}} \sim N(0,1)$
3. Realisierung aus Stichprobe: $t_{\text{ data}} = -2/\sqrt{0.64} = -2.5$
4. Entscheidung: H_0 abgelehnt bei $\alpha = 5\,\%$, falls $t_{\text{ data}} < t_\alpha = -t_{1-\alpha} = -1.645$. Dies ist der Fall

p-Wert:

$$p = P(T < t_{\text{ data}} | \beta_{U_3} = 0) = \Phi(t_{\text{ data}}) = \Phi(-2.5) = 1 - \Phi(2.5) = 1 - 0.9938 = 0.0062$$

Nun zweiter Test:

1. $H_0 : \beta_{U_3} > -1$
2. Testvariable $T = \frac{\hat{\beta}_{U_3}+1}{\sqrt{V(\hat{\beta}_{U_3})}} \sim N(0,1)$
3. Realisierung aus Stichprobe: $t_{\text{ data}} = -1/\sqrt{0.64} = -1.25$
4. Entscheidung: H_0 abgelehnt bei $\alpha = 5\,\%$, falls $t_{\text{ data}} < t_\alpha = -t_{1-\alpha} = -1.645$. Dies ist *nicht* der Fall

[3] Ein Umstieg bei einem Flug ist "nerviger" als bei Bahn/Bus, die Umstiege werden deshalb alternativenspezifisch modelliert (nicht verlangt).