

Name:	Vorname:	Matrikel-Nr.:
-------	----------	---------------

## Klausur zur Vorlesung Methoden Verkehrsökometrie für Diplom-Studierende WS 2011/12

### Aufgabe 1 (50 Punkte)

Um zu analysieren ob und wieviel die Eisenbahn in den letzten Jahrzehnten schneller geworden ist, werden die gemessenen Durchschnittsgeschwindigkeiten für die langsameren Zugtypen (S-Bahnen, Regional-, Vorortzüge etc.) und die schnelleren Züge (IC, EC, ICE etc.) analysiert:

Jahr $x_1$ (seit 1900)	0	30	60	60	80	80	100	100
Zugkategorie $x_2$	0	0	0	1	0	1	0	1
Geschw (km/h) $y$	45	60	65	90	86	105	83	130

Die Geschwindigkeitsentwicklung soll mit dem linearen Regressionsmodell

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

beschrieben werden. Dabei kann  $x_2$  die Werte 0 und 1 (langsame bzw. schnelle Züge) annehmen. Vor 1960 gibt es diese Unterteilung allerdings nicht und die Züge wurden der langsameren Klasse zugeordnet.

- (a) Betrachten Sie zunächst die funktionale Spezifikation: Wie sind die exogenen und endogenen Variablen skaliert? Enthält das Modell überflüssige exogene Variable, wenn ja welche? Fehlen exogene Faktoren? Wenn ja, geben Sie mindestens einen davon an. Kann Linearität angenommen werden oder muss man ggf. transformieren? Kann Homogenität in den Daten angenommen werden oder gibt es Gründe für einen Strukturbruch?
- (b) Sie können annehmen, dass die unbestimmten Anteile der acht Datensätze  $i$  durch  $\epsilon_i \sim i.i.d N(0, \sigma^2)$  charakterisiert sind. Erläutern Sie diese Charakterisierung und zeigen Sie, dass sie alle vier Bedingungen einer korrekten statistischen Spezifikation enthält.
- (c) Zeigen Sie anhand des Sachverhalts und der Daten der obigen Tabelle, dass das Modell bezüglich der Daten korrekt spezifiziert ist, also die exogenen Variablen keine Zufallsvariablen sind, keine Multi-Kolinearität vorliegt und der Stichprobenumfang ausreichend ist.
- (d) Geben Sie anschauliche Bedeutungen für die drei Parameter an. Welche Vorzeichen erwarten Sie für die geschätzten Werte von  $\beta_1$  und  $\beta_2$ ? (Begründung!)
- (e) Bestimmen Sie den LSE-Schätzer des Parametervektors unter Verwendung der bereits ausgerechneten Varianzen und Kovarianzen  $s_{11} = 1250$ ,  $s_{12} = 6.41$ ,  $s_{22} = 0.234$  sowie  $s_{1y} = 728$  und  $s_{2y} = 9.50$ .

Name:	Vorname:	Matrikel-Nr.:
-------	----------	---------------

- (f) Nehmen Sie in dieser und der folgenden Teilaufgabe folgenden Schätzvektor und die zugehörige Varianz-Kovarianzmatrix an:

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} 43 \\ 0.41 \\ 29 \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} 40 & -0.44 & -1.5 \\ -0.44 & 0.0078 & -0.21 \\ -1.5 & -0.21 & 42 \end{pmatrix}.$$

Prognostizieren Sie (ohne Fehlerabschätzung) die erwarteten mittleren Geschwindigkeiten der langsameren und schnelleren Zugklassen im Jahr 2020.

- (g) Berechnen Sie zu einer Fehlerwahrscheinlichkeit von 5% das Konfidenzintervall von  $\beta_1$ . *Hinweis:* Sie müssen keine Residualvarianz ausrechnen, verwenden Sie die obige Varianz-Kovarianzmatrix!
- (h) Das Modell berücksichtigt nicht, dass sich die Geschwindigkeiten der langsameren und der schnelleren Züge mit unterschiedlichen Raten über die Zeit ändern können. Erweitern Sie das Modell (also die funktionale Modellspezifikation) so, dass Unterschiede erfasst werden können.

### Aufgabe 2 (20 Punkte)

Betrachtet werden der Sektor des ÖPNV und der Sektor 2, welcher die sonstige Volkswirtschaft umfasst. Die Koeffizienten der Matrix des direkten Aufwandes sind

$$A_{11} = 0.1, \quad A_{12} = 0.02, \quad A_{21} = 0.4, \quad A_{22} = 0.3.$$

- (a) Berechnen Sie die Matrix des vollen Aufwandes.
- (b) Da die meisten ÖPNV-Verkehrsmittel elektrifiziert sind, fallen für die Erbringung des Wertes von einem Euro an Transportdienstleistungen vergleichsweise geringe Direktmissionen von 0.1 kg CO<sub>2</sub> an. Allerdings benötigt der ÖPNV große Mengen an Ressourcen von der sonstigen Wirtschaft, bei der 0.5 kg CO<sub>2</sub> pro Euro Waren- oder Dienstleistungswert anfallen. Wie hoch sind ÖPNV-Emissionen an CO<sub>2</sub> pro Euro an Endnachfrage einschließlich Verflechtungseffekte? Wie hoch sind die Emissionen der sonstigen Wirtschaft pro Euro Nachfrage der Endverbraucher einschließlich Verflechtungseffekte?

Name:	Vorname:	Matrikel-Nr.:
-------	----------	---------------

### Aufgabe 3 (50 Punkte)

Ein Autohausbesitzer verwendet viel Aufwand an Beratung und kostenlosen Probefahrten, obwohl nicht jeder Kunde auch ein Auto kauft. Bei manchen hat er eher den Eindruck, dass sie nur die neuen Modelle probefahren möchten. Um sich besser auf die "ernsthafte" Kundschaft zu fokussieren, fragt er nun unverfänglich jeden Interessenten, wie alt das aktuelle Auto sei und ob er/sie dieses als Neuwagen gekauft habe. Außerdem bietet er unterschiedliche Rabatte an und notiert, ob der Kunde nun ein Auto bei ihm kauft:

Alter des aktuellen Autos (Jahre)	1	3	5	7	10	5	8	10	12	14
Rabatt (1 000 €)	1	2	3	1	0	3	2	2	1	3
Bisher Neuwagen (0=nein,1=ja)	0	0	1	0	1	1	1	1	0	0
Auto gekauft (0=nein,1=ja)	0	0	0	0	0	1	1	0	0	1

- (a) Handelte es sich um eine Revealed-Choice oder eine Stated-Choice - Befragung? Begründen Sie ihre Antwort in einem Satz.
- (b) Die Daten sollen mit dem Binomial-Logit-Modell mit folgenden deterministischen Nutzenfunktionen analysiert werden:

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

(Alternative  $i = 1$ : Autokauf, Alternative  $i = 2$ : kein Auto gekauft,  $n$ : Personenindex;  $T_n$ : Alter des aktuellen Autos in Jahren,  $R_n$ : Rabatt in 1 000 €). Unterscheiden Sie die Einflussfaktoren nach generischen und sozioökonomischen Variablen sowie alternativenspezifischen Konstanten.

- (c) Muss man das Fahrzeugalter wie oben alternativenspezifisch definieren, also die Nutzenänderung durch diesen Faktor wie oben durch den Summand  $\Delta V_{ni} = T_n \delta_{n1}$  spezifizieren, oder wäre auch  $\Delta V_{ni} = T_n$  möglich? Begründen Sie ihre Antwort!
- (d) Das Logitmodell soll nun mit der Maximum-Likelihood-Methode geschätzt werden. Dabei müssen vier Merkmalssummen in den Daten und im Modell gleich sein. Formulieren und interpretieren Sie die Merkmalssummen. Geben Sie die realisierten Werte und die vom Modell für  $\beta = \mathbf{0}$  vorausgesagten Werte an.
- (e) Die Modellkalibrierung ergab den geschätzten Parametervektor

$$\hat{\beta} = (-9.2, 0.35, 2.2, 1.3)^T$$

Warum ist  $\hat{\beta}_1$  so negativ?

- (f) Ein neuer Kunde hat zur Zeit ein 5 Jahre altes, damals neu gekauftes Auto. Der Händler bietet 2 000 € Rabatt an. Mit welcher Wahrscheinlichkeit greift der Kunde zu?

Name:	Vorname:	Matrikel-Nr.:
-------	----------	---------------

- (g) Um zu überprüfen, ob die Art des bisherigen Autos (Neu- oder Gebrauchtwagen) oder der Rabatt die Kaufentscheidung signifikant beeinflusst, werden die einfacheren Modelle

$$\text{M1: } V_{ni} = [\beta_1 + \beta_2 T_n + \beta_3 R_n] \delta_{n1}$$

$$\text{M2: } V_{ni} = [\beta_1 + \beta_2 T_n] \delta_{n1}$$

geschätzt. Die maximierte Log-Likelihood beträgt bei den restringierten Modellen  $\ln L_{\max}^{\text{M1}} = -3.52$  und  $\ln L_{\max}^{\text{M2}} = -5.78$ , während die Log-Likelihood des geschätzten vollen Modells  $\ln L_{\max} = -3.43$  beträgt. Führen Sie für die beiden restringierten Modelle Likelihood-Ratio-Tests gegenüber dem vollen Modell bei einer Fehlerwahrscheinlichkeit von 10% durch. Interpretieren Sie das Ergebnis.

- (h) Neben Kunden mit bisherigen Neu- oder Gebrauchtwagen könnte es auch Kunden ohne bisheriges Auto geben. Wie muss man das Modell erweitern, um dies zu berücksichtigen?

### Quantile $t_{\alpha}^{(n)}$ der Studentischen $t$ -Verteilung mit $n$ Freiheitsgraden

$n$	$\alpha = 0.60$	0.70	0.80	0.90	0.95	0.975	0.990	0.995	0.999	0.9995
1	0.325	0.727	1.376	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	0.289	0.617	1.061	1.886	2.920	4.303	6.965	9.925	22.327	31.598
3	0.277	0.584	0.978	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.271	0.569	0.941	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.267	0.559	0.920	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.265	0.553	0.906	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.263	0.549	0.896	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.262	0.546	0.889	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.261	0.543	0.883	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.260	0.542	0.879	1.372	1.812	2.228	2.764	3.169	4.144	4.587
$\infty$	0.253	0.524	0.842	1.282	1.645	1.960	2.326	2.576	3.090	3.291

### Quantile $q_{\alpha}^{(n)}$ der $\chi^2$ -Verteilung mit $n$ Freiheitsgraden

$n$	$\alpha = 0.9900$	0.9750	0.9500	0.9000	0.8000	0.5000	0.2000	0.1000	0.05000
1	6.635	5.034	3.821	2.706	1.656	0.4589	0.06540	0.01638	0.004230
2	9.210	7.378	5.991	4.605	3.219	1.386	0.4463	0.2107	0.1026
3	11.34	9.348	7.815	6.251	4.642	2.366	1.005	0.5843	0.3518
4	13.28	11.14	9.488	7.779	5.989	3.357	1.649	1.064	0.7106
5	15.09	12.83	11.07	9.236	7.289	4.351	2.343	1.610	1.145
6	16.81	14.45	12.59	10.64	8.558	5.348	3.070	2.204	1.635
7	18.48	16.01	14.07	12.02	9.803	6.346	3.822	2.833	2.167
8	20.10	17.54	15.51	13.36	11.03	7.344	4.594	3.490	2.733
9	21.67	19.03	16.92	14.68	12.24	8.343	5.380	4.168	3.325
10	23.22	20.49	18.31	15.99	13.44	9.342	6.179	4.865	3.940