Methoden Verkehrsökonometrie für Master-Studierende

Winter semester 2021/22, Solutions to Tutorial No. 10

Lösungsvorschlag zu Aufgabe 10.1: Relevanz von Einflussfaktoren: Likelihood-Ratio-Test

Allgemein:

Zur Berechnung der anhand der Daten realisierten Werte der LR-Test-Statistiken

$$\lambda = 2\left[\ln L(\hat{\vec{\beta}}) - \ln L^{\mathrm{r}}(\hat{\vec{\beta}}^{\mathrm{r}})\right] \sim \chi^2(J - J_r)$$

bei den verschiedenen Modellvergleichen benötigen wir die Log-Likelihood-Werte der jeweiligen geschätzten Modelle. Anhand der beiden Contourplots erhält man für die Probit- und Logit-Modelle folgende Ergebnisse:

	Probit			Logit		
Modell	$\hat{\beta}_1$	$\hat{\beta}2$	$\ln L(\hat{\vec{\beta}})$	$\hat{\beta}_1$	$\hat{\beta}2$	$\ln L(\hat{\vec{\beta}})$
volles Modell M_1	-0.35	-0.08	-16	-0.3	-0.075	-16
Konstantenmodell M_2	-0.5	–	-19.5	-0.4	–	-19.5
Nur Zeitsensitivität M_3	–	-0.08	-16.5	–	-0.075	-16.5
Trivialmodell M_4	–	–	-20.5	–	–	-20.5

(a) Hier wird das volle Modell mit der Modellvariante "nur Zeitsensitivität verglichen, also $M = M_1$ und $M_r = M_3$. Man erhält

$$\lambda_{\text{Probit}} = 1, \quad \lambda_{\text{Logit}} = 1$$

Im Vergleich mit dem Quantil $\chi^2_{1,0.95} = 3.9$ kann die Nullhypothese H_0: "Die globale Bevorzugung einer Alternative existiert nicht" nicht abgelehnt werden (p-Werte um 30 %).

(b) Hier gilt $M = M_1$ und $M_r = M_2$ und damit

$$\lambda_{\text{Probit}} = 7 > \chi^2_{1,0.95}, \quad \lambda_{\text{Logit}} = 7 > \chi^2_{1,0.95}.$$

Die Zeitsensitivität ist also ein signifikanter Einflussfaktor. Damit ist sowohl unter gaußverteilten (Probit) wie unter gumbelverteilten Zufallsnutzen (Logit-Modell) jeweils das reine Zeitsensitivitätsmodell das unter den vier Modellvarianten zu bevorzugende.

(c) (i) Beim vollen Modell ($M = M_1$) im Vergleich zum Trivialmodell ($M_r = M_4$) haben die beiden zu vergleichenden Modelle eine Parameterzahldifferenz von 2, die Test-Statistik ist also $\chi^2(2)$-verteilt. Es ergibt sich

$$\lambda_{\text{Probit}} = 2 * 4.5 = 9 > \chi^2_{2,0.95} \approx 6, \quad \lambda_{\text{Logit}} = 2 * 4.5 = 9 > \chi^2_{2,0.95}.$$

Die p-Werte sind kleiner als 0.1 %, das volle Modell ist also in der Beschreibung hochsignifikant besser als das Trivialmodell

(ii) Beim Vergleich der Konstanten- mit den Trivialmodellen ($M = M_2$, $M_r = M_4$, Parameterdifferenz 1) ergeben sich LR-Statistiken von etwa 2, also ist das Konstantenmodell im Vergleich mit dem Trivialmodell grenzwertig besser (p-Werte um $1 - F^{(1)}_{\chi^2}(2) = 5\,\%$).

(iii) Der Vergleich des Zeitsensitivitäts- mit dem Trivialmodell ($M = M_3$, $M_r = M_4$, Parameterdifferenz 1) ergibt sich wieder hohe Signifikanz bei den Probit- und Logitmodellen ($p < 1\,\%$), was die Wahl des Zeitsensitivitätsmodells als das "richtige" Modell bestätigt.

Diskussion

Es kann durchaus vorkommen, dass ein echter Beeinflussungsfaktor wie die Reisezeit durch eine globale pauschale Bevorzugung "mehr schlecht als recht" ebenfalls modelliert wird und dieser dann signifikant werden kann. Führt man dann aber den "echten" relevanten Einflussfaktor in das Modell ein, verliert die pauschale Bevorzugung oft ihre Signifikanz.

Lösungsvorschlag zu Aufgabe 10.2: Likelihood-Ratio Test for regression models: $\lambda = T^2$

(a) Nein. Dies ist analog zu den Modellen der diskreten Wahltheorie, bei denen das Trivialmodell durch $V_{ki} = 0$ definiert ist und Konstantenmodellen nur alternativenspezifische Konstanten aufweisen, $V_{ki} = \sum_{j=1}^{K-1} \delta_{jk}$. Keines dieser Modelle enthält auf der rechten Seite den Index i, welcher Eigenschaften der verschiedenen betrachteten statistischen Einheiten und damit exogene Variable anzeigt.

(b) Likelihood-Funktion bei bekannter Varianz σ_ϵ^2:

$$L(\beta_0) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_\epsilon^2}} e^{-\frac{(y_i-\beta_0)^2}{2\sigma_\epsilon^2}}.$$

Log-Likelihood:

$$l(\beta_0) = \ln L(\beta_0) = \sum_{i=1}^{n} \left[\frac{-1}{2} \ln(2\pi\sigma_\epsilon^2) - \frac{(y_i - \beta_0)^2}{2\sigma_\epsilon^2} \right]$$

Maximieren:

$$\frac{\partial l(\beta_0)}{\partial \beta_0} = \sum_{i=1}^{n} \left[\frac{y_i - \beta_0}{\sigma_\epsilon^2} \right] \stackrel{!}{0} \quad \Rightarrow \quad \hat{\beta}_0 = \bar{y}.$$

(c) Die Log-Likelihood des restringierten Modells (Trivialmodell) lautet schlicht

$$\ln L^{\mathrm{r}} = \sum_{i=1}^{n} \left[\frac{-1}{2} \ln(2\pi\sigma_\epsilon^2) - \frac{(y_i - \mu_0)^2}{2\sigma_\epsilon^2} \right]$$

Bei Formulierung der LR-Teststatisik λ fallen die konstanten Summanden $-1/2 \ln(2\pi\sigma_\epsilon^2)$ der beiden Loglikelihoods weg und es ergibt sich

$$\begin{aligned} \lambda &= 2 \left[\ln L(\hat{\beta}_0) - \ln L^{\mathrm{r}} \right] \\ &= 2 \sum_{i=1}^{n} \left[-\frac{(y_i - \bar{y})^2}{2\sigma_\epsilon^2} + \frac{(y_i - \mu_0)^2}{2\sigma_\epsilon^2} \right] \end{aligned}$$

Dies lässt sich folgendermaßen umformen:

$$\begin{aligned} \lambda\sigma_\epsilon^2 &= \sum_{i=1}^{n} \left[-(y_i - \bar{y})^2 + (y_i - \mu_0)^2 \right] \\ &= \sum_{i=1}^{n} \left[-y_i^2 + 2y_i\bar{y} - \bar{y}^2 + y_i^2 - 2y_i\mu_0 + \mu_0^2 \right] \\ &= \sum_{i=1}^{n} \left[2y_i(\bar{y} - \mu_0) + \mu_0^2 - \bar{y}^2 \right] \\ &= n \left[\bar{y}^2 + \mu_0^2 - 2\bar{y}\mu_0 \right] \\ &= n \left(\bar{y} - \mu_0 \right)^2 \end{aligned}$$

Nun ist aber bei Zutreffen der Nullhypothese wegen der i.i.d gaußverteilten Zufallsvariablen

$$\bar{y} \sim N \left(\mu_0, \frac{\sigma_\epsilon^2}{n} \right)$$

also gehorcht

$$\sqrt{\lambda} = \sqrt{n} \frac{\bar{y} - \mu_0}{\sigma_\epsilon} \sim N(0,1) \tag{1}$$

einer Standardnormalverteilung. Da die $\chi^2(m)$-Verteilungen als Summe von m Quadraten unabhängiger standardnormalverteilter Größen definiert sind, gilt für λ selbst

$$\lambda \sim \chi^2(1).$$

(d) Bei bekannter Varianz wird der Student-t-Test mit der $T(n-1)$-verteilten Testvariablen $T = \sqrt{n}(\bar{y} - \mu_0)/s$ zum Z-Test mit der standardnormalverteilten Testvariablen $Z = \sqrt{n}(\bar{y} - \mu_0)/\sigma_\epsilon$. Vergleich mit (1) zeigt, dass

$$\lambda = Z^2 \tag{2}$$

Eine einfache graphische Veranschaulichung zeigt, dass Z^2 genau dann größer als das $1-\alpha$ Quantil $z_{1-\alpha}^2 = \chi_{1,1-\alpha}^2$ ist, wenn der *Betrag* von Z das $1-\alpha/2$-Quantil überschreitet, also wegen (2)

$$|Z| > z_{1-\alpha/2} \Leftrightarrow \lambda > \chi_{1,1-\alpha}^2$$

Ergebnis und Diskussion

Der LR-Test zeigt *genau dann* eine Signifikanz des Parameters β_0 und damit ein gegenüber dem Trivialmodell signifikant "besseres" Konstantenmodell an, wenn der *symmetrische* Z-Test (also der T-Test bei bekannter Varianz) der Nullhypothese "Wahrer Parameterwert=Wert im restringiertem Modell" anschlägt.

Schätzte man allerdings σ_ϵ^2 ebenfalls durch die ML-Methode, erhielte man bei kleinen Datensätzen bedeutsame verzerrte Schätzer. In diesem Sinne ist die ML-Methode, einschließlich des Likelihood-Ratio-Tests, nur *asymptotisch*, also bei großem Datenumfang $n \gg J$ genau. (J gibt die Gesamtzahl der zu schätzenden Parameter an). Insbesonder gilt nur asymptotisch, dass die LR-Statistik (bei Gültigkeit der Nullhypothese) χ^2-verteilt ist.