

Aufgaben zur Vorlesung

Multivariate Verfahren

Übungsblatt IX

- Beantworten Sie folgende Fragen bzw. bearbeiten Sie folgende Arbeitsanweisungen:
 - Was ist der Average Marginal Effect?
 - Nennen Sie drei Erweiterungen des Logit-/Probit-Ansatzes. Wofür werden diese jeweils eingesetzt?
 - Erklären Sie, was es mit der Independence of Irrelevant Alternatives auf sich hat.
 - Was muss bei der Interpretation des marginalen Effekts berücksichtigt werden?
- Nach der Durchführung einer Logit-Regression wollen Sie die Güte Ihres Modells überprüfen. Sie haben einen Cut-Off-Wert von 0.5 gewählt und erhalten daraus die folgende Kreuztabelle über die Häufigkeiten der richtigen und falschen Zuordnungen der Daten:

y_i	\hat{y}_i		
	0	1	
0	20	20	40
1	17	43	60
	37	63	100

- Berechnen Sie die Prognosegenauigkeit, Sensitivität und Spezifität.
 - Skizzieren Sie die zugehörige ROC-Kurve. Können Sie anhand dieser bereits die Güte des Modells abschätzen?
- Ein großer Autohersteller möchte das Kaufverhalten seiner Kunden analysieren. Zu diesem Zweck führt das Unternehmen eine Logit-Regression durch für die Kaufwahrscheinlichkeit eines Autos auf die Regressoren x_1 : Nettoeinkommen des Kunden in EUR und x_2 : Entfernung des Kundenwohnortes vom Werk des Autoherstellers in km. Sei mit $y = 1$ der Kauf eines Autos kodiert. Folgendes Logit-Modell ohne konstanten Term wurde geschätzt:

$$P(y = 1|x') = \Lambda(-0.0006x_1 + 0.025x_2).$$

Nach der Schätzung des Modells kam es leider zu einem Absturz der Computersysteme. Ein Praktikant konnte allerdings noch folgende vier Datenpunkte aus der Stichprobe auf einem Zettel notieren:

Laufende Nummer	x_1	x_2
1	2000	102
2	2500	75
3	1500	60
4	2500	153

- (a) Wie hoch sind die Wahrscheinlichkeiten für einen Autokauf der vier Kunden?
 - (b) Ausgehend vom Datenpunkt 1. Wie hoch ist der marginale Effekt für einen weiteren Kilometer Entfernung des Wohnortes vom Autowerk?
 - (c) Angenommen die vier Datenpunkte seien die komplette Stichprobe. Wie hoch ist der Average Marginal Effect für die beiden Merkmale? Interpretieren Sie die Größe.
4. Für Ihren Arbeitgeber, eine große M&A-Investmentfirma, haben Sie ein polytomes Logit-Modell geschätzt. Mit dessen Hilfe wollen Sie die Wahrscheinlichkeiten für das Fortbestehen ($y = 0$), die Restrukturierung ($y = 1$) und den Konkurs ($y = 2$) eines finanzschwachen Unternehmens beurteilen.
- (a) Zeigen Sie, dass $\frac{\partial P_1}{\partial x_1} = P_1(\beta_{11} - \beta_{11}P_1 - \beta_{21}P_2)$.
 - (b) Die Variable x_1 im Modell ist die Bilanzsumme des Unternehmens in Mio. EUR. Seien nun $P_1 = 0.3$, $P_2 = 0.2$, $\beta_{11} = 0.05$ und $\beta_{21} = 0.15$. Wie hoch ist der marginale Effekt der Variablen x_1 auf P_1 ? Interpretieren Sie das Ergebnis.
 - (c) Ist die Interpretation aus (b) für das Modell allgemeingültig? Was muss zusätzlich berücksichtigt werden?
5. Rechnerübung: Laden Sie den Datensatz *iris* aus der sklearn-Bibliothek herunter (*from sklearn import datasets*). Dieser enthält den Ihnen bereits aus der Clusteranalyse bekannten Schwertlilientendatz. Führen Sie eine polytome Logit-Regression durch.