

2 Binäre Regression

Aufgabe 2

Man betrachte ein Modell mit der dichotomisierten Responsevariable Arbeitslosigkeit ($y \in \{1, 0\}$; 1: Kurzarbeitslosigkeit, 0: Langzeitarbeitslosigkeit) sowie den dummy-kodierten Kovariablen Geschlecht ($x_G \in \{1, 0\}$; 1: männlich, 0: weiblich) und Ausbildung ($x_A \in \{1, 0\}$; 1: Uni-Abschluss, 0: kein Uni-Abschluss). Die Wahrscheinlichkeiten für Kurzarbeitslosigkeit gegeben die Kovariablen seien gegeben durch

Geschlecht	Ausbildung	$P(y = 1 x_A, x_G)$
$x_G = 1$	$x_A = 1$	0.9
	$x_A = 0$	0.7
$x_G = 0$	$x_A = 1$	0.6
	$x_A = 0$	0.4

- Spezifizieren Sie in dieser Situation ein Logit-Modell, das eine (eventuelle) Interaktion zwischen Geschlecht und Ausbildung berücksichtigt. Um welchen Spezialfall handelt es sich bei diesem Modell?
- Bestimmen Sie mit Hilfe von Stift, Papier und ggf. Taschenrechner die Parameter Ihres Modells aus Teilaufgabe (a).
- Bestimmen Sie nun die Parameter Ihres Modells unter Verwendung von R.
- Interpretieren Sie die Parameter Ihres Modells. Berücksichtigen Sie dabei insbesondere die Zusammenhänge zu den Odds bzw. Odds Ratios.
- Es soll nun mit Hilfe eines Likelihood-Quotiententests geprüft werden, ob der Interaktionsterm signifikant ist. Geben Sie dabei Nullhypothese, Teststatistik und Ablehnbereich an. Kann mit den hier zur Verfügung stehenden Informationen eine Testentscheidung getroffen werden?

Aufgabe 3

Gegeben sei ein Zufallsvektor (y, \mathbf{x}) , für den

$$\mathbf{x}|y = i \sim \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}), \quad i \in \{1, 2\}, \quad \pi_i = P(y = i)$$

gilt. Leiten Sie unter Verwendung des Satzes von Bayes die bedingte Verteilung von $y|\mathbf{x}$ her. Welchem Modell entspricht dies?