Statistik IV für Nebenfachstudierende

4.3. Multinomiales Logit-Modell

Prof. Dr. Andreas Mayr

Institut für Statistik, LMU München

Sommersemester 2017

Allgemeines Modell

$\underline{x}$

$$\log \frac{\mathbb{P}(Y \; = \; r \; | \underline{x}_i)}{\mathbb{P}(Y \; = \; k \; | \underline{x}_i)} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip} = \underline{x}_i^\top \underline{\beta}_r$$

für $r = 1, ..., q (= k - 1)$

Auflösen nach $\mathbb{P}(Y = r \mid \underline{x}_i)$

$$\log \frac{\mathbb{P}(Y = r \mid \underline{x}_i)}{\mathbb{P}(Y = k \mid \underline{x}_i)} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip} = \underline{x}_i^\top \underline{\beta}_r$$

$$\mathbb{P}(Y = r \mid \underline{x}_i) = \frac{\exp(\underline{x}_i^\top \beta_r)}{1 + \sum_{s=1}^{k-1} \exp(\underline{x}_i^\top \beta_s)}$$

bzw.

$$\mathbb{P}(Y = k \mid \underline{x}_i) = \frac{1}{1 + \sum_{s=1}^{k-1} \exp(\underline{x}_i^\top \beta_s)}$$

Alternative allgemeine Darstellung

$$\mathbb{P}(Y \; = \; r \mid \underline{x}_i) = \frac{\exp(\underline{x}_i^\top \beta_r)}{\sum_{s=1}^k \exp(\underline{x}_i^\top \beta_s)}$$

Problem: Identifizierbarkeit

- Spezialfall nur ein $x_1 = 1$:
  - Parameter $\beta_1, ..., \beta_k$
  - Freie Wahl $\mathbb{P}(Y = 1)$, ..., $\mathbb{P}(y = k - 1)$
  - Zu viele Parameter $\Longrightarrow$ nicht identifizierbar

    Bsp.:

Restriktionen

- Modell braucht Restriktionen zur Eindeutigkeit der Parameter.
- Es gibt grundsätzlich zwei Möglichkeiten:
  1. Wähle Referenzkategorie
  2. Setze systemmatische Nebenbedingung

Referenzkategorie

Setze Referenzkategorie r_0:

Symmetrische Nebenbedingung

Stelle Nebenbedingung an $\underline{\beta}$ (seltener in der Praxis):

Interpretation

Immer gilt:

Interpretation:

Bei Referenzkategorie:

Odds-Ratios

- Ähnlich wie beim binären Logit-Modell ist die Interpretation über Odds-Ratios leichter.
- Diese bezieht sich immer auf die Referenzkategorie.