

Multinomiale Logistische Regression

Benjamin Schlegel

09. Mai 2016

Die multinomiale logistische Regression wird verwendet, wenn die abhängige Variable nominalskaliert ist. Um eine multinomiale Regression zu schätzen, muss zuerst eine Referenzkategorie gebildet werden. Für alle anderen Fälle wird dann eine logistische Regression vom Modell geschätzt.

Grundsätzlich kann jede Kategorie als Referenzkategorie verwendet werden. Kommen in der Referenzkategorie jedoch nur wenige Fälle vor, wird der Standardfehler grösser. Deshalb ist es ratsam, eine Kategorie zu wählen, die viele Fälle enthält. Ein zweites Kriterium für die Wahl der Referenzkategorie ist die Distanz zu den anderen Parteien. Je grösser die Distanz, desto besser.

Beispiel: Ist die abhängige Variable die Parteiwahl, so sollte man eine rechte oder linke Partei wählen. In der Schweiz ist beispielsweise die SVP geeignet, da sie zum einen extrem ist und zum anderen viele Fälle enthält. Weniger geeignet wäre hingegen die Lega die Ticino, da sie zwar extrem ist, dafür jedoch nur wenige Fälle hat und so die Standardfehler in die Höhe treiben würde.

Nachdem die Referenzkategorie definiert wurde, kann das Modell geschätzt werden. Anders als bei den anderen logistischen Modellen, kann bei den Koeffizienten die Richtung nicht abgelesen werden, d.h. es macht keinen Sinn von positiven oder negativen Effekten zu sprechen. Das hängt damit zusammen, dass die Werte von anderen Werten abhängen. Es ist jedoch möglich, für konkrete Fälle die vorausgesagten Wahrscheinlichkeiten zu berechnen.

Die vorausgesagten Wahrscheinlichkeiten können mit folgender Formel berechnet werden:

$$\pi_k = \frac{e^{X\beta_k}}{1 + e^{X\beta_1} + e^{X\beta_2} + \dots + e^{X\beta_{K-1}}}, k = 1, 2, \dots, K-1 (X\beta = \beta_0 + \beta_1 x_1 + \dots)$$

Die Referenzkategorie kann berechnet werden, indem alle anderen Wahrscheinlichkeitswerte von 1 abgezogen werden. So summieren sich die Möglichkeiten auf 1.

$$\pi_K = 1 - \sum_{k=1}^{K-1} \pi_k$$

weiterführende Literatur

Kühnel, Steffen M. und Dagmar Krebs (2010): Multinomiale und ordinale Regression. In: Christof Wolf und Henning Best (Hrsg.): Handbuch der sozialwissenschaftlichen Datenanalyse. Wiesbaden: VS Verlage. (Deutsch)