OPEN ACCESS

Übersichtsarbeit

Multinomiale, ordinale und stereotype logistische Regression – eine Einführung in die Regressionsanalyse kategorialer Zielvariablen

Multinomial, ordinal and stereotype logistic regression – an introduction to the regression analysis of categorial outcome variables

Norbert Kersten[1]

1 Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Berlin, Deutschland

Abstract

For about three decades now a number of suitable regression models for categorical outcome variables have been described in the literature. These are also available as analysis tools in statistical software packages. This overview presents some of the models which are suitable for medicine and health sciences. The lesser-known stereotypical model in particular is emphasized since the ordinal and the multinomial models have been in use for some time.
The article is aimed at statistics users: It covers types of categorical data, models and their implementation in statistical software. Estimation and inference theory is omitted. It demonstrates, firstly, model selection depending on the order structure of the outcome variables and, secondly, appropriate interpretation of the model parameters. This is illustrated by one example each for the multinomial and the stereotypical model. Notes to use the software are included.

Keywords: multinomial, ordinal and stereotype regression, logistic regression model, categorial outcome variables

Zusammenfassung

Seit rund drei Jahrzehnten gibt es in der Literatur eine Reihe geeigneter Regressionsmodelle zur Analyse kategorialer Outcome-Variablen. Diese sind inzwischen in den großen Statistik-Softwarepaketen implementiert.
Diese Übersicht soll einige für Medizin und Gesundheitswissenschaften geeignete Modelle darstellen, wobei das Schwergewicht auf den weniger bekannten stereotypen Modellen liegt. Die ordinalen und die multinomialen Modelle sind seit längerem in Gebrauch.
Der Artikel richtet sich an Anwender. Er beschreibt die Typen von kategorialen Daten, Modellen und deren Implementierungen, auf Schätz- und Inferenztheorie wird verzichtet. Es wird gezeigt, wie in Abhängigkeit von der Ordnungsstruktur der Zielvariablen Modelle ausgewählt und wie die Modellparameter interpretiert werden können. Dies wird durch je ein Beispiel zum multinomialen und zum stereotypen Modell illustriert, Erläuterungen zum Umgang mit der Software sind eingeschlossen.

Schlüsselwörter: multinomiale, ordinale und stereotype Regression, logistisches Regressionsmodell, kategoriale Zielvariablen

1 Einleitung

Die Analyse kategorialer Daten ist fester Bestandteil medizinischer und gesundheitswissenschaftlicher Studien. Dies gilt für diverse statistische Prozeduren wie z.B. die Analyse von Kontingenztafeln, aber nur eingeschränkt für Regressionsmodelle mit kategorialen Zielvariablen. Diese Modelle fordern einigen Aufwand bei der Modellwahl und der Interpretation der z.T. zahlreichen Modellparameter. Viele Autoren dichotomisieren aus diesen – oder anderen Gründen – die Zielvariable ihrer Modelle und beschränken sich auf die binäre logistische Regression, andere Autoren sprechen sich gegen ein solches Vorgehen aus, z.B. [1], [2]. Die Reduzierung beobachteter

Kategorien auf Binärdaten ist mit einigem Informationsverlust verbunden, denn für die einzelnen Outcome-Kategorien eines Modells mit nichtreduzierter Zielvariable können durchaus differenzierte und interessante Effekte der Regressoren schätzbar sein. Seit mindestens drei Jahrzehnten gibt es hierfür eine Reihe geeigneter Modellansätze. Diese Arbeit soll eine Übersicht über die vorhandenen Modelle geben, soll zeigen, wie in Abhängigkeit von der Ordnungsstruktur der Zielvariablen Modelle ausgewählt und wie die Modellparameter interpretiert werden können.

In Beobachtungs-, Register- oder experimentellen Studien werden dort, wo keine metrischen Zielvariablen erfasst werden können, Gegebenheiten in Form von Ereignissen, Zuständen, Eigenschaften usw. registriert. Solche Gegebenheiten werden als Mengen von Kategorien A_j (j=1,...,J; J>2) beschrieben, diese Mengen können vollständig geordnet oder auch völlig ungeordnet sein, dazwischen sind auch teilgeordnete Mengen möglich. Je nach Ausgangslage können hier verschiedene Modelle notwendig sein, ihnen gemeinsam ist, dass die Zielvariable einer Multinomialverteilung folgt. Voraussetzung dafür ist, dass eine eindeutige Relation zwischen der Menge der Beobachtungseinheiten und der Menge der Kategorien besteht, d.h. jeder Beobachtungseinheit wird *genau* eine Kategorie zugeordnet.

In Döring/Bortz [3] wird gefordert, die Zuweisung der Beobachtungseinheiten zu Kategorien anhand von Merkmalen bzw. Merkmalskombinationen so vorzunehmen, dass dabei drei Kriterien erfüllt sind:

1. Genauigkeit: Die Merkmalsausprägungen sollen so definiert sein, dass für jede Beobachtungseinheit feststellbar ist, welche Ausprägung vorliegt.
2. Exklusivität: Die Merkmalsausprägungen sollen sich gegenseitig ausschließen, so dass jede Beobachtungseinheit höchsten einer Kategorie zugeordnet wird.
3. Exhaustivität: Die Merkmalsausprägungen sollen ein Merkmal bzw. eine Merkmalskombination erschöpfend beschreiben, so dass jede Beobachtungseinheit mindestens einer Kategorie zugeordnet werden kann.

Die Forderungen in [3] sind in ihrer Gesamtheit äquivalent mit der Forderung einer eindeutigen Relation zwischen der Menge der Beobachtungseinheiten und der Menge der Kategorien. Für den Praktiker wird es nützlicher sein, das Erfülltsein dieser drei Kriterien zu prüfen als die Eindeutigkeit einer Relation nachzuweisen, insbesondere dann, wenn die erhobenen Beobachtungen ohne genaue Vorstellungen über die spätere Datenanalyse in Kategorien überführt werden sollen.

Im gesundheitswissenschaftlichen Kontext sind solche Kategorien z.B. eine Krankheit oder ein Krankheitsstadium mit einer bestimmten Diagnose, der Vitalstatus ohne oder mit einer bestimmten Todesursache. So können in Folge einer Einwirkung von Asbestfaserstäuben sowohl fibrogene Wirkungen im Bereich des Lungenparenchyms (Asbestose), pleurale Läsionen (Plaques und Verkalkungen) als auch Bronchialkarzinom und Mesotheliom auftreten [4]. Diese Krankheiten können einzeln oder gemeinsam vorkommen. Die Kategorienbildung muss folglich Uni- und Multimorbidität bei den Probanden berücksichtigen.

2 Modelle für kategoriale Zielvariablen

Bei N Wiederholungen einer Beobachtung oder eines Experiments werden die Kategorien A_j im Zusammenhang mit weiteren Variablen (Prädiktoren) $X_1,...,X_I$ unter der Zielstellung beobachtet, einen funktionalen Zusammenhang zwischen dem Auftreten der Kategorie A_j und den dabei beobachteten Werten der Variablen $X_1,...,X_I$ herzustellen. Bei den bislang eingesetzten verallgemeinerten linearen Modellen ist dabei von unabhängigen Wiederholungen ausgegangen worden. Inzwischen gestatten die GEE-Modelle auch unterschiedliche Anzahlen von Wiederholungsbeobachtungen innerhalb von Probanden bzw. korrelierte Beobachtungen von Probanden in übergeordneten Einheiten (z.B. Interventionsgruppen) zu analysieren. Von den hier vorgestellten Modellen ist bisher nur das ordinale Modell als GEE-Modell implementiert (Tabelle 1).

Üblicherweise werden dabei die Kategorien A_j mit Zahlenwerten, z.B. 1,...,J, codiert. In einer einzelnen Beobachtung wird nun statt der Kategorien A_j eine Zufallsgröße Y registriert, die Werte aus der Menge {1,...,J} annimmt, beobachtet wird also ein Vektor (Y, $X_1,...,X_I$). Zwischen den bedingten Wahrscheinlichkeiten $p_j=P(Y=j\,|\,X_1,...,X_I)$ und den Prädiktorvariablen X_i (i=1,...,I) ist ein modellmäßiger Zusammenhang herzustellen, dieser Zusammenhang wird üblicherweise in Form eines verallgemeinerten linearen Modells dargestellt. Mit Hilfe einer Verteilungsfunktion F als Linkfunktion und den Modellparametern β_{ij} lassen sich die Wahrscheinlichkeiten p_j als Funktion einer Linearkombinationen der Variablen $X_1,...,X_I$ darstellen:

$$P(Y=j|X_1,\dots,X_I) = F\left(\beta_{0j} + \sum_{i=1}^{I}\beta_{ij}X_i\right) \quad (1)$$

Häufig verwendete Funktionen für F sind die Verteilungsfunktion Φ der Normalverteilung (Probitanalyse) und die Verteilungsfunktion F(t) = 1/(1+exp(-t)) der logistischen Verteilung, wobei in Medizin und Gesundheitswissenschaften vor allem die logistischen Regressionsmodelle Anwendung finden.

Es gibt eine Reihe von Regressionsmodellen, die auf dem logistischen Modellansatz beruhen, aber verschiedene Formen der Parametrisierung aufweisen. Bei Anwendern bekannt sind das multinomiale und das ordinale logistische Modell, weniger bekannt ist das stereotype logistische Modell. Der Einsatz dieser drei Modelle erfolgt je nach gegebener Ordnungstruktur der beobachteten Kategorienmenge. Wesentliche Begriffe für die Auswahl des geeigneten Regressionsmodells sind die *Dimensionalität* und die *Unterscheidbarkeit* [5], [6].

Die Dimensionalität eines funktionalen Zusammenhanges zwischen einer kategorialen Variablen Y und den Prädiktoren $X_1,\ldots,X_I$ ist durch die Anzahl an linearen Funktionen $\sum_{i=1}^{I}\beta_{ij}X_i$ festgelegt, die benötigt wird, um den funktionalen Zusammenhang zu beschreiben [5]. Wenn nur eine lineare Funktion $\sum_{i=1}^{I}\beta_{ij}X_i$ nötig ist, um zwischen allen Kategorien, auf denen Y basiert, zu unterscheiden, ist die Beziehung eindimensional. Kann die lineare Funktion nur zwischen einer Teilmenge der Kategorien von Y unterscheiden und eine andere lineare Funktion unterscheidet zwischen den Kategorien einer anderen, nicht notwendig zur ersten durchschnittsfreien Teilmenge, ist der funktionale Zusammenhang mindestens zweidimensional. Das ordinale Modell z.B. beschreibt eine 1-dimensionale, das multinomiale Modell eine (J-1)-dimensionale Beziehung. Die Dimensionalität eines funktionalen Zusammenhanges baut auf dem Begriff der Unterscheidbarkeit auf. Ein Paar von Kategorien heißt nichtunterscheidbar hinsichtlich der Variablen X_i, wenn X_i nicht prädiktiv für diese beiden Kategorien ist. Ist z.B. X_i nicht prädiktiv für die Kategorien j_1 und j_2, aber prädiktiv für die beiden Kategorien $\{j_1, j_2\}$ und j_3, dann heißen die Kategorien j_1 und j_2 nicht unterscheidbar für X_i [5].

Im Folgenden werden drei logistische Modelle für kategoriale Zielvariablen dargestellt und es wird gezeigt, wie sich Wirkungszusammenhänge zwischen Outcome- und Prädiktorvariablen mit unterschiedlichen Dimensionen damit modellieren lassen.

2.1 Das multinomiale logistische Regressionsmodell

Die Grundform des multinomialen logistischen Regressionsmodells (MNLM) ist durch

$$P(Y{=}j|X_1,\ldots,X_I) = \frac{\exp\{\beta_{0j}+\sum_{i=1}^{I}\beta_{ij}X_i\}}{\sum_{h=1}^{J}\exp\{\beta_{0h}+\sum_{i=1}^{I}\beta_{ih}X_i\}} \quad \textbf{(2)}$$

gegeben [7]. Der Term im Nenner von (2) dient der Normierung der Wahrscheinlichkeiten, so dass $\sum_{j=1}^{J}P(Y{=}j|X_1,\ldots,X_I)=1$ gilt. Die Parameter β in diesem Modellansatz sind jedoch nicht eindeutig schätzbar, aus diesem Grund müssen Restriktionen an die Parameter gestellt werden. Eine häufig verwendete Restriktion an die Parameter ist, eine der beobachteten Kategorien als Referenz auszuwählen, z.B. das Nichtvorliegen einer Krankheit, und deren Parameter auf Null zu setzen. Zweckmäßigerweise codiert man diese Kategorie mit dem Wert 1 oder J, so dass sich eine einfachere Interpretation der Parameterschätzungen ergibt. Wählt man als Referenz die mit J codierte Kategorie, so ist $\beta_{iJ} = 0$ (i=0,...,I) und (2) nimmt für j=1,...,J-1 die Form

$$P(Y{=}j|X_1,\ldots,X_I) = \frac{\exp\{\beta_{0j}+\sum_{i=1}^{I}\beta_{ij}X_i\}}{1+\sum_{h=1}^{J-1}\exp\{\beta_{0h}+\sum_{i=1}^{I}\beta_{ih}X_i\}} \quad \textbf{(3)}$$

an. Die Wahrscheinlichkeit für die Kategorie A_J lässt sich dann mittels $P(Y{=}J|X_1,\ldots,X_I) = 1-\sum_{j=1}^{J-1}P(Y{=}j|X_1,\ldots,X_I)$ aus den Wahrscheinlichkeiten der übrigen Kategorien rekonstruieren. Eine andere Form der Restriktion wäre es, die Summe der Parameter β_{ij} über j auf Null zu setzen. Ähnlich wie in der Varianzanalyse, wo die Summe der Abweichungen der Zellenmittel vom Gesamtmittel auf Null gesetzt wird, werden dann die Parameter als Abweichungen von einem mittleren Responseniveau interpretiert [7].

Beim MLNM wird – mit Ausnahme der Referenzkategorie – für jede Outcome-Kategorie ein Satz Parameter geschätzt. Die Interpretation dieser Parameter kann aufwendig sein, da nicht nur die Effekte der Prädiktoren in Relation zur Referenzkategorie sondern auch die Effekte hinsichtlich aller anderen Kombinationen je zweier Outcome-Kategorien betrachtet werden sollten. Für die Odds zweier Outcome-Kategorien ($j_1 \neq j_2$) gilt:

$$\frac{P(Y{=}j_1|X_1,\ldots,X_I)}{P(Y{=}j_2|X_1,\ldots,X_I)} = \exp\left\{(\beta_{0j_1}-\beta_{0j_2})+\sum_{i=1}^{I}\left(\beta_{ij_1}-\beta_{ij_2}\right)X_i\right\} \quad \textbf{(4)}$$

Die Relation der Wahrscheinlichkeiten für das Auftreten der Kategorien A_{j_1} oder A_{j_2} unter der Bedingung $(X_1,\ldots,X_J)$ hängt daher von der Differenz der Parameter ab. Eine Outcome-Kategorie wird also häufiger auftreten als eine andere, wenn entsprechende Unterschiede in den Effekten der Prädiktoren und/oder den Achsenabschnittsparametern vorhanden sind.

Ändert sich der Wert eines Prädiktors X_i um einen Einheitsbetrag $\Delta X_i = 1$ (Wechsel zu einer benachbarten Kategorie auf einer kategorialen Skala oder eine metrische Einheit auf einer Intervallskala) folgt:

$$OR_{j_1j_2}(\Delta X_i) = \frac{\dfrac{P(Y{=}j_1|X_1,\ldots,X_i{+}\Delta X_i,\ldots,X_I)}{P(Y{=}j_2|X_1,\ldots,X_i{+}\Delta X_i,\ldots,X_I)}}{\dfrac{P(Y{=}j_1|X_1,\ldots,X_i,\ldots,X_I)}{P(Y{=}j_2|X_1,\ldots,X_i,\ldots,X_I)}} = \exp\left\{\beta_{ij_1}-\beta_{ij_2}\right\} \quad \textbf{(5)}$$

Somit ändert sich das Odds-Verhältnis der Outcome-Kategorie j_1 gegenüber der Outcome-Kategorie j_2 bei Änderung des Prädiktors X_i um einen Einheitsbetrag ΔX_i um den Faktor $e^{\beta_{ij_1}}/e^{\beta_{ij_2}}$. Ist j_2 der Index der Referenzkategorie, dann ist β_{ij_2} gleich 0 und (5) nimmt eine Form an, wie man sie vom binären logistischen Modell kennt. Mit Hilfe der Parameter β_{ij} lassen sich alle Relationen zwischen den Wahrscheinlichkeiten der Outcome-Kategorien und die Effekte der Prädiktoren auf diese Relationen in Form von Odds-Ratios beschreiben.

Das multinomiale Modell liefert die bestmögliche Anpassung an die Daten auf Kosten einer großen Anzahl von Parametern, deren Interpretation schwierig sein kann [8], da Redundanzen unter den Prädiktoreffekten einiger Outcome-Kategorien auftreten können.

2.2. Das ordinale logistische Regressionsmodell

Das ordinale logistische Regressionsmodell (OLM), auch als kumulatives logistisches Modell bekannt, ist eine von mehreren Möglichkeiten [9] zur Analyse ordinaler Outcomes. In seiner formalen Modellstruktur bietet es Vergleichsmöglichkeiten mit den anderen hier vorgestellten Modellen. Im OLM reduziert sich die Zahl der Parameter, indem zwar für jede Outcome-Kategorie ein spezifischer

Achsenabschnittsparameter β_{0j} in das Modell aufgenommen wird, die Effektparameter β_i der Prädiktoren aber als unabhängig von den Outcome-Kategorien angesetzt werden. Praktisch ist es jedoch nicht mit den beiden anderen vorgestellten Modellen vergleichbar, weil die Modellierung nominaler Daten hiermit nicht möglich ist. Voraussetzung ist, dass die zugrunde liegenden Kategorien einer 1-dimensionalen Ordnungsrelation unterliegen und somit in auf- oder absteigender Reihenfolge kumuliert werden können. Das Modell sagt dann die Wahrscheinlichkeit für das Auftreten kumulierter Kategorien bis zu einem vorgegebenen Index j (j<J) voraus:

$$P(Y\leq j|X_1,\ldots,X_I) = \frac{\exp\{\beta_{0j}+\sum_{i=1}^{I}\beta_i X_i\}}{1+\exp\{\beta_{0j}+\sum_{i=1}^{I}\beta_i X_i\}} \quad (6)$$

Im Modellansatz (6) sind die Parameter ohne weitere Restriktionen identifizierbar. Das Verhältnis der Wahrscheinlichkeiten für das Auftreten kumulierter Kategorien $\{A_1,\ldots,A_{j1}\}$ und $\{A_1,\ldots,A_{j2}\}$ unter der Bedingung $(X_1,\ldots,X_I)$ hängt nun nicht mehr von den Prädiktoren ab und reduziert sich auf:

$$P\left(Y\leq j_1|X_1,\ldots,X_I\right)=P\left(Y\leq j_2|X_1,\ldots,X_I\right)\cdot\exp\left\{\beta_{0j_1}-\beta_{0j_2}\right\} \quad (7)$$

D.h. die Odds für die kumulierten Kategorien sind proportional über alle Ausprägungen der Prädiktoren hinweg mit der exponierten Differenz der Achsenabschnittsparameter als Proportionalitätsfaktor. Dies bedeutet, dass dieses Modell nur anwendbar ist, wenn die Prädiktoren in allen Kumulationsstufen gleiche Effekte aufweisen (Proportional-Odds-Bedingung). Man kann dies näherungsweise prüfen, wenn man aus den kumulierten Kategorien $\{A_1,\ldots,A_j\}$ und $\{A_{j+1},\ldots,A_J\}$ (j=1,...,J-1) eine Folge binärer Kategorien bildet und jeweils ein binäres logistisches Regressionsmodell schätzt. Die Regressionskoeffizienten für die Prädiktoren sollten in allen Modellen annähernd gleich sein. Ein gemeinsamer Plot der Log-Odds aller Modelle über einer Prädiktorenachse sollte dann nahezu parallele Graden ergeben.

Der Effekt eines Prädiktors X_i bei Änderung um einen Einheitsbetrag $\Delta X_i=1$ ergibt sich zu $OR(\Delta X_i)=\exp\{\beta_i\}$. Die Wahrscheinlichkeiten für das Auftreten einzelner Kategorien werden als Differenz der Wahrscheinlichkeiten zweier benachbarter kumulierter Kategorien mit dem Index j und j+1 berechnet. Für den Index J gilt $P(Y=J|X_1,\ldots,X_I) = 1 - P(Y\leq(J-1)|X_1,\ldots,X_I)$.

Das ordinale Regressionsmodell wurde ausführlich und mit Beispielen in [10] dargestellt, ebenso ist dort eine anschauliche Erläuterung der Proportional-Odds-Bedingung zu finden. Einen breiten Überblick zu logistischen Regressionsmodellen mit geordneten Zielvariablen findet man in [11] und [12].

2.3 Das stereotype logistische Regressionsmodell

Das stereotype logistische Regressionsmodell (SLM) wurde von Anderson 1984 unter dem Aspekt entwickelt, die enge proportional-odds-Voraussetzung des ordinalen Modells zu umgehen und andererseits ein sparsameres Modell als das multinomiale logistische Modell zur Verfügung zu haben [5]. Ähnlich zu diesen beiden Modellen ist das stereotype logistische Modell für den Einsatz mit kategorialen abhängigen Variablen gedacht, wobei diese nominal oder ordinal skaliert sein können. Die Outcome-Kategorien können geordnet sein, dies ist aber keine Voraussetzung wie im ordinalen Modell. Insbesondere ist das SLM für teilgeordnete Outcome-Kategorien geeignet, z.B. wenn man zwischen den Kategorien *GESUND* und *VERSTORBEN* verschiedene Krankheiten – oder ihre Stadien – in disjunkte Kategorien fasst, die sich nicht als Gesamtmenge rangskalieren lassen. Es lässt sich auch anstelle des multinomialen logistischen Modells verwenden, wenn einige der Outcome-Kategorien gleiche Effekte der Prädiktoren aufweisen.

Das stereotype logistische Regressionsmodell stellt einen Kompromiss zwischen dem ordinalen und dem multinomialen logistischen Regressionsmodell dar [13]. Es benötigt mehr Parameter als das ordinale Modell, im günstigen Fall aber deutlich weniger als das multinomiale Modell. Wenn sich die Outcome-Kategorien in einer Dimension ordnen lassen, ähnelt das stereotype Modell dem ordinalen Modell, wenn die Outcome-Kategorien keinerlei Ordnungsstruktur aufweisen, dann entspricht das stereotype Modell dem multinomialen Modell mit einer anderen Form der Parametrisierung. Konkret hat das (eindimensionale) stereotype logistische Regressionsmodell die Form:

$$P(Y=j|X_1,\ldots,X_I) = \frac{\exp\{\beta_{0j}+\varphi_j\sum_{i=1}^{I}\beta_i X_i\}}{\sum_{h=1}^{J}\exp\{\beta_{0h}+\varphi_h\sum_{i=1}^{I}\beta_i X_i\}} \quad (8)$$

Analog dem multinomialen logistischen Modell wird die Wahrscheinlichkeit für das Auftreten einer Kategorie A_j unter der Bedingung $(X_1,\ldots,X_J)$ kategorienspezifisch bestimmt, der Effekt der Variablen X_j ist jedoch wie im ordinalen Modell für alle Kategorien gleich. Im Unterschied zum ordinalen Modell werden Parameter φ_j eingeführt, die den summativen Effekt der Variablen $X_1,\ldots,X_J$ gewichten, so dass einerseits die adjustierten Effekte der Prädiktoren erhalten bleiben, andererseits aber der gewichtete Gesamteffekt der Prädiktoren zwischen den Outcome-Kategorien variieren kann. Diese Variation der Effekte „befreit“ das Modell quasi von der proportional-odds-Bedingung. Da auch in diesem Modell die Parameter nicht eindeutig schätzbar sind, werden Restriktionen eingeführt. Im Modellansatz (8) reicht es aus, diese auf die φ-Koeffizienten zu beschränken. Analog zum MNLM wird eine Outcome-Kategorie als Referenzkategorie festgelegt, deren Parameter auf Null gesetzt wird, z.B. $\varphi_1=0$. Für eine weitere Kategorie wird $\varphi_j=1$ festgelegt. Mit diesen beiden Festlegungen sind sowohl die φ- als auch die β-Koeffizienten eindeutig schätzbar.

Hinsichtlich der Interpretation der Parameter ist das stereotype logistische Modell ähnlich aufwendig wie das multinomiale logistische Modell, da analog zu diesem die Effekte aller Kombinationen je zweier Outcome-Kategorien betrachtet werden sollten. Für die Odds zweier Outcome-Kategorien $(j_1\neq j_2)$ gilt:

$$\frac{P(Y{=}j_1|X_1,\ldots,X_I)}{P(Y{=}j_2|X_1,\ldots,X_I)} = \exp\left\{\left(\beta_{0j_1}-\beta_{0j_2}\right)+(\varphi_{j_1}-\varphi_{j_2})\sum_{i=1}^{I}\beta_i X_i\right\} \quad (9)$$

D.h. die Relation der Wahrscheinlichkeiten für das Auftreten der Kategorien A_{j_1} oder A_{j_2} unter gegebenen $(X_1,\ldots,X_I)$ hängt in dem Maße von den Prädiktoren ab, wie sich die Parameter φ1 und φ2 voneinander unterscheiden. Sind beide Parameter gleich, dann ist $e^{\beta_{0j_1}}/e^{\beta_{0j_2}}$ das Verhältnis der Basiswahrscheinlichkeiten der Kategorien A_{j_1} und A_{j_2} in der Referenzgruppe und die Prädiktoren haben keinen Effekt auf die Relation der Wahrscheinlichkeiten. Ändert sich der Wert eines Prädiktors X_i um einen Einheitsbetrag ΔX_i (Wechsel zu einer benachbarten Kategorie auf einer kategorialen Skala oder eine metrische Einheit auf einer Intervallskala) folgt analog zu (5):

$$OR_{j_1j_2}(\Delta X_i) = \exp\left\{(\varphi_{j_1}-\varphi_{j_2})\beta_i\right\} \quad (10)$$

Somit ändert sich das Odds-Verhältnis der Outcome-Kategorie j_1 gegenüber der Outcome-Kategorie j_2 um den Faktor $\left(e^{\varphi_{j_1}}/e^{\varphi_{j_2}}\right)^{\beta_i}$ bei Änderung des Prädiktors X_i um einen Einheitsbetrag ΔX_i. Für die Referenzkategorie ist φ_j=0 zu setzen.

Sind alle φ-Koeffizienten deutlich voneinander verschieden, so lassen sie sich in einer Rangfolge anordnen. Diese Rangfolge bestimmt dann eine Ordnungsrelation der Kategorien A_j, d.h. die Ordinalität der Zielvariablen muss nicht von vorn herein gegeben sein, sondern sie wird über die Effektstärke der Prädiktoren auf die Zielvariable bestimmt. Eine Kategorie A_j, deren Auftretenswahrscheinlichkeit stark durch die Prädiktoren beeinflusst ist, erhält einen adäquat hohen Rang in der betrachteten Kategorienmenge.

Wenn die φ-Koeffizienten zweier Outcome-Kategorien nahezu gleich sind, ergibt sich aus Formel (10) ein Odds-Verhältnis von nahezu 1. Das bedeutet, dass kein Prädiktor X_i einen Beitrag zur Änderung des Odds-Verhältnisses der Kategorien j_1 und j_2 leistet, die Kategorien sind im Modell unter den gegebenen Prädiktoren nicht unterscheidbar. Wenn zwei Outcome-Kategorien nicht unterscheidbar sind, könnten sie zusammengelegt werden. Die Entscheidung dafür kann mit Hilfe eines Likelihood-Quotienten-Testes getroffen werden, indem eine zusätzliche Restriktion an die Parameter φ_i eingeführt wird. Die Testgröße

$$LQ{=}2\cdot\left\{\ln L\left(\varphi_{j_1}\neq\varphi_{j_2}|X_1,\ldots,X_J\right) - \ln L\left(\varphi_{j_1}{=}\varphi_{j_2}|X_1,\ldots,X_J\right)\right\}$$

ist Chi-Quadrat verteilt mit einem Freiheitsgrad, wobei das Modell mit der Parameterrestriktion $\varphi_{j_1}=\varphi_{j_2}$ ein Submodell zum Modell ohne diese Restriktion darstellt.

Der Modellansatz für das SLM kann in verschiedenen Dimensionen ausgeführt werden. Liegen vollständig ranggeordnete Outcome-Kategorien vor, d.h. diese lassen sich auf einer eindimensionalen Achse abbilden, kann ein eindimensionales Modell angepasst werden, wie in (8) dargestellt. Lassen sich die Outcome-Kategorien nicht in einer eindimensionalen Ordnungsrelation abbilden, kann die Dimension des Modells erhöht werden. Ein zweidimensionales Modell hat die Form:

$$P(Y{=}j|X_1,\ldots,X_I) = \frac{\exp\{\beta_{0j}+\varphi_j^{(1)}\sum_{i=1}^{I}\beta_j^{(1)}X_i+\varphi_j^{(2)}\sum_{i=1}^{I}\beta_j^{(2)}X_i\}}{\sum_{j=1}^{J}\exp\{\beta_{0j}+\varphi_j^{(1)}\sum_{i=1}^{I}\beta_j^{(1)}X_i+\varphi_j^{(2)}\sum_{i=1}^{I}\beta_j^{(2)}X_i\}} \quad (11)$$

Die Dimension des Modells entspricht der Anzahl an Parametersätzen, die in das Modell aufgenommen werden. Das Modell bildet aus den Prädiktoren eine entsprechende Anzahl von Linearkombinationen, die dann mit Hilfe der φ-Koeffizienten als Gewichte zu einem Schätzer für die Wahrscheinlichkeiten verrechnet werden. Die maximale Dimension ist gleich dem Minimum aus der um 1 verringerten Anzahl der Outcome-Kategorien und der Anzahl an Prädiktoren: D_{max} = min(J-1,I) [8]. Bei Modellen mit mehr als einer Dimension kann die Rangordnung der φ-Koeffizienten zwischen den Dimensionen variieren, ebenso können die Prädiktorvariablen in den einzelnen Dimensionen unterschiedliche p-Werte aufweisen. Prädiktoren, die in einer Dimension signifikant sind, können in einer anderen Dimension nichtsignifikant sein. Ob die Erhöhung der Anzahl an Dimensionen zu einer besseren Modellanpassung führt, kann ebenfalls mittels Likelihood-Quotiententest geprüft werden, da es sich hier um genestete Modelle handelt. Das Modell mit der niedrigeren Anzahl an Dimensionen stellt ein Submodell zum Modell mit der höheren Anzahl an Dimensionen dar.

Mit zunehmender Zahl der Dimensionen müssen mehr Restriktionen an die φ-Parameter eingeführt werden, um die Identifizierbarkeit der Parameter zu gewährleisten. Für ein zweidimensionales Modell kann z.B. $\varphi_1^{(1)}$=0, $\varphi_2^{(1)}$=0, $\varphi_3^{(1)}$=1 und $\varphi_1^{(2)}$=0, $\varphi_2^{(2)}$=1, $\varphi_3^{(2)}$=0 gesetzt sein. Die übrigen Parameter sind frei variierbar. Wenn das Modell die Dimension J-1 hat, sind alle φ-Parameter auf 0 oder 1 gesetzt. Das stereotype Modell ist dann mit dem multinomialen Modell identisch und stellt nur eine andere Form der Parametrisierung desselben dar.

3 Softwareimplementierungen

Alle großen Statistik-Softwarepakte enthalten Verfahren zur Analyse von kategorialen Daten. Die Darstellung der Implementierungen von Regressionsmodellen für kategoriale Outcome-Variablen ist hier auf die aktuellen Versionen der Programmpakete SPSS, STATA, SAS und SYSTAT eingegrenzt (Tabelle 1), wobei nur einige Spezifika der Implementierungen hervorgehoben werden sollen. Detaillierte Anweisungen zur Nutzung der Programmmodule stehen in den jeweiligen Software-Handbüchern zur Verfügung.

SYSTAT – ein Programmpaket, das vorrangig zur Analyse von Labordaten konzipiert ist – enthält lediglich ein Modul für die multinomiale logistische Regression. In den anderen drei Paketen gibt es z.T. mehrere Module für die multinomiale und die ordinale logistische Regression. Implementierungen für das stereotype logistische Modell beschränken sich dagegen auf SAS und STATA und wurden erst vor einigen Jahren bereitgestellt. Kuss [14] beschreibt ein SAS-Makro zur Schätzung der Parameter ei-

Tabelle 1: Implementierungen von Regressionsmodellen für kategoriale Outcome-Variablen in Statistik-Softwarepaketen

Softwarepaket		MNLM	OLM	SLM
Menü	SPSS	>Regression >Complex Samples	>Regression >GLM >GEE >Complex Samples	-
	STATA	>Categorial Outcomes	>Ordinal Outcomes	>Categorial Outcomes
	SYSTAT	>Regression>Logit	-	-
Syntax	SPSS	NOMREG CSLOGISTIC	PLUM GENLIN CSORDINAL	-
	STATA	MLOGIT	OLOGIT	SLOGIT
	SAS	LOGISTIC	LOGISTIC	NLMIXED, Makro
	SYSTAT	LOGIT	-	-

nes stereotypen Regressionsmodells. Wenige Jahre später bestand die Möglichkeit mithilfe der SAS-Prozedur NLMIXED, einer Prozedur für die Parameterschätzung in nichtlinearen Modellen mit zufälligen Effekten, stereotype Logitanalysen auszuführen [15]. Allerdings war dies mit einigem Aufwand hinsichtlich der Vorbereitung der Datenstruktur verbunden. In einem Paper des SAS Global Forum 2013 wird neben verschiedenen Modellen für ordinale Responsevariablen auch das stereotype logistische Modell dargestellt und die Parameterschätzung mit der Prozedur NLMIXED beschrieben. Die Berechnung von Odds Ratios wird mit Hilfe von Programmierbeispielen demonstriert und die Anpassungsgüte der Modelle getestet [16].

Im gleichen Zeitraum wurde das SLM in STATA eingeführt. Lunt [8] beschreibt das STATA-Modul SOREG zur Schätzung der Modellparameter. Die heute verfügbare STATA-Prozedur SLOGIT wurde 2006 von Long und Freese beschrieben [13]. SLOGIT ist sowohl in der Menü- als auch in der Syntaxform eine komfortable Prozedur mit der stereotype logistische Modelle ohne Aufwand angepasst werden können. Die Spezifikation der Modellvariablen erfolgt analog zu anderen Regressionsanalyseprozeduren. Spezifisch für SLOGIT ist die Möglichkeit, die Dimension eines stereotypen Modells vorzugeben und Modelle unterschiedlicher Dimensionen mittels Likelihood-Quotienten-Test im Postestimation-Paragraph zu testen. Daneben können außer den durch das Modell vorgegebenen weitere Restriktionen an die φ-Parameter gestellt werden, z.B. $\varphi_{j_1}=\varphi_{j_2}$ um für zwei Outcome-Kategorien die Hypothese paralleler Effekte zu prüfen. Solche Restriktionen können sowohl in der Syntax als auch in der Menüsteuerung über die constraint-Anweisung ausgeführt werden. Der entsprechende Likelihood-Quotienten-Test wird ebenfalls über den Postestimation-Paragraph realisiert.

Ein Beispiel zum Umgang mit der Prozedur SLOGIT gibt es in [17]. Der Fokus liegt dort auf dem Einsatz des stereotypen logistischen Modells auf ordinale Responsevariablen bei Verletzung der Proportional-Odds-Bedingung. Es wird explizit auf die unterschiedliche Interpretation der Odds-Ratios verwiesen, die zwangsläufig beim Übergang von einem Modell mit kumulierten Kategorien zu einem mit Verwendung einer Referenzkategorie beachtet werden muss.

Den Implementierungen in den Programmpaketen liegen unterschiedliche Parametrisierungen zu Grunde. Wünschenswert ist eine Interpretation der Parameter, wie man sie vom binären logistischen Modell gewohnt ist. Bei diesem wird die Wahrscheinlichkeit für ein besonderes Ereignis (Krankheit, Unfall usw.) modelliert, Personen ohne ein solches Ereignis bilden die Referenzgruppe. Dieser Ansatz bewirkt, dass sich für einen Prädiktor, der positiv mit der Ereignishäufigkeit assoziiert ist, ein positiver Regressionskoeffizient ergibt. Von den Regressionsmodellen für kategoriale Outcome-Variablen erwartet man eine ähnliche „natürliche" Interpretation der Regressionskoeffizienten. Das multinomiale Regressionsmodell ist in den Softwarepaketen so implementiert, wie in Formel 4 dargestellt. Damit lassen sich alle Parameter ohne Umstände analog zum binären Modell interpretieren.

Beim ordinalen logistischen Regressionsmodell in der Darstellung

$$\ln\left(\frac{P(Y\le j|X_1,\dots,X_I)}{P(Y>j|X_1,\dots,X_I)}\right) = \beta_{0j}+\sum_{i=1}^{I}\beta_i X_i \quad \textbf{(12)}$$

wird die Wahrscheinlichkeit des Auftretens einer Outcome-Kategorie mit dem Index j oder kleiner zur Wahrscheinlichkeit, dass eine Outcome-Kategorie mit einem Index größer als j auftritt, in Relation gesetzt. Beinhaltet die Outcome-Variable z.B. die Schweregrade einer Krankheit in aufsteigender Reihenfolge mit den gesunden Personen als erste (Referenz-)Kategorie, so muss für eine dem binären logistischen Modell entsprechende Interpretation der Koeffizienten die Relation in (12) umgekehrt werden:

$$\ln\left(\frac{P(Y>j|X_1,\dots,X_I)}{P(Y\le j|X_1,\dots,X_I)}\right) = -\ln\left(\frac{P(Y\le j|X_1,\dots,X_I)}{P(Y>j|X_1,\dots,X_I)}\right) \quad \textbf{(13)}$$

Es ergibt sich ein zu (12) analoges Modell mit negativen Vorzeichen bei allen Koeffizienten auf der rechten Seite der Gleichung. In SPSS und STATA ist das ordinale Modell jedoch in der Form

$$\ln\left(\frac{P(Y\le j|X_1,\dots,X_I)}{P(Y>j|X_1,\dots,X_I)}\right) = \beta_{0j}-\sum_{i=1}^{I}\beta_i X_i \quad \textbf{(14)}$$

implementiert. Damit wird durch die Parameter β_i eine positive Assoziation zwischen den Prädiktoren und den

Wahrscheinlichkeiten $P(Y>j|X_1,...X_I)$ für j=2,...,J (sinkende Wahrscheinlichkeit der Outcome-Kategorien mit einem Index $\leq$ j) beschrieben, jedoch müssen die β_{0j} uminterpretiert werden. Diese dienen im Modellansatz (6) der Berechnung der Referenzwahrscheinlichkeiten (Ordinatenabschnitte) für die kumulierten Kategorien A_1, $\{A_1, A_2\}$, $\{A_1, A_2, A_3\}$, ... Für eine Interpretation analog zum binären logistischen Modell benötigt man die β_0-Parameter für die gegenläufig kumulierten Kategorien A_J, $\{A_{J-1}, A_J\}$, $\{A_{J-2}, A_{J-1}, A_J\}$..., wobei der Zusammenhang

$$\beta_0(Y>j) = -\beta_0(Y\leq j)$$

gilt. D.h. bei Umkehrung der Vorzeichen der β_{0j} ist hier außerdem eine Neuzuordnung der Parameter zu den kumulierten Kategorien vorzunehmen.

Die SAS-Prozedur LOGISTIC verwendet nicht den Ansatz (14), sondern den Ansatz (12), so dass für eine Modellinterpretation analog zum binären Modell die Vorzeichen aller Parameter gemäß (13) umzukehren sind.

Beim stereotypen logistischen Modell in STATA wird – ähnlich wie beim ordinalen Modell in SPSS und STATA – ebenfalls eine Vorzeichenumkehr vorgenommen. Desweiteren wird eine Parametertransformation $\theta_{0j} = \varphi_j \cdot \beta_{0j}$ eingeführt [13], so dass (8) die Form

$$P(Y=j|X_1,\ldots,X_I) = \frac{\exp\{\theta_{0j} - \varphi_j \sum_{i=1}^{I} \beta_i X_i\}}{\sum_{j=1}^{J} \exp\{\theta_{0j} - \varphi_j \sum_{i=1}^{I} \beta_i X_i\}} \quad (15)$$

annimmt. Bei Interpretation der Koeffizienten sind die Vorzeichen der β-Parameter umzukehren, damit z.B. Expositionseffekte sachgemäß dargestellt werden. Die θ_{0j} dagegen sind, wie Parallelrechnungen mit dem multinomialen Modell zeigen, in der richtigen Größenordnung und Vorzeichenorientierung und müssen nicht zurückgerechnet werden.

Wer sich bei der Interpretation der Modellparameter beim stereotypen Modell nicht sicher ist, sollte parallel dazu ein multinomiales Modell rechnen und die Odds-Ratios (bezogen auf die Referenzkategorie) aus den verschiedenen Modellen hinsichtlich ihrer Größenordnung und Vorzeichenorientierung miteinander vergleichen.

4 Beispiele

4.1 Beispiel für ein multinomiales logistisches Modell

Für dieses erste Beispiel wurden aus dem Datenbestand einer Studie zum Gesundheitszustand der Erwerbsbevölkerung [18] 1.294 erwerbstätige Frauen selektiert und deren teils recht unvollständige Daten über mehrere Ersetzungen von Fehlwerten gemittelt und gerundet. In dieser Studie sollte untersucht werden, ob neben den körperlichen Arbeitsanforderungen auch psychosoziale Arbeitsbedingungen einen Einfluss auf Gesundheit und Wohlbefinden der Beschäftigten haben.

Als Zielvariable wurde der selbstberichtete Gesundheitszustand über eine Frage aus dem SF36 [19] in den Stufen „ausgezeichnet“, „sehr gut“, „gut“, „mittelmäßig“ und „schlecht“ erfasst. Von der ordinalen Struktur dieser Variablen wird hier abgesehen, stattdessen ist von Interesse wie und mit welcher Effektstärke die Prädiktoren ausgehend von einem mittleren Gesundheitszustand zur Verbesserung oder Verschlechterung des Gesundheitszustandes beitragen.

Von den zahlreichen Variablen zu den Arbeitsbedingungen, den Arbeitsinhalten und dem Arbeitsumfeld sollen hier nur zwei Variablen dienen. Bei den Frauen haben die Arbeitszufriedenheit und der work-family-conflict jeweils einen wesentlichen Zusammenhang mit dem Gesundheitszustand. Die Arbeitszufriedenheit (COPSOQ) bezog sich auf die Arbeitssituation insgesamt und wurde als Zufriedenheitsscore durch Mittelung von sieben Items berechnet. Die Antwortkategorien waren „sehr zufrieden“, „zufrieden“, „unzufrieden“ und „sehr unzufrieden“. Die Codierung der Kategorien erfolgte kongruent zur Codierung des Gesundheitszustandes, um einen positiven Zusammenhang zwischen steigender Arbeitszufriedenheit und gutem Gesundheitszustand zu erhalten. Ein hoher Scorewert beschreibt eine hohe Arbeitszufriedenheit.

Die Variable work-family-conflict erfasst als single-Item den Konflikt zwischen den Anforderungen aus dem Arbeitsverhältnis und den Anforderungen aus dem Privat- und Familienleben. Die Antwortkategorien „stimme voll zu“, „stimme eher zu“, „stimme eher nicht zu“ und „stimme gar nicht zu“ bilden eine kategoriale Variable und wurden entsprechned im Modell verwendet.

Adjustiert wurde auf das Alter im Jahr 2011 (metrisch) und den höchsten Schulabschluss (kategorial) mit den Stufen „Hauptschule oder weniger“, „mittlerer Schulabschluss“ und „Abitur“.

Das multinomiale Modell wurde mit SPSS 21 NOMREG berechnet, der Output ist in Tabelle 2 wiedergegeben. Unter *Deskriptive Statistik* sind beschreibende Parameter zu den beiden metrischen Variablen Alter und Arbeitszufriedenheit aufgeführt, weiter unten unter *Verarbeitete Fälle* die Häufigkeiten für die Kategorien der Variablen selbstberichteter Gesundheitszustand (self-rated-health), work-family-conflict und höchster Schulabschluss. Als Referenzkategorie für die Zielvariable wurde entsprechend der oben formulierten Zielstellung die Kategorie „(3) gut“ gewählt, als Referenzkategorie für den work-family-conflict die Kategorie „(4) es liegt kein Konflikt vor“ und für den höchsten Schulabschluss „(3) Abitur“.

In Tabelle 2 sind unter *Parameterschätzer* in der Spalte B die geschätzten Regressionsparameter für das multinomiale logistische Modell zu finden. Mit deren Hilfe können die Veränderungen in den Vorhersagewahrscheinlichkeiten der Kategorien der Zielvariablen infolge einer Änderung der Ausprägung der Prädiktorvariablen gemäß Formel (5) in Form von Odds-Ratios beschrieben werden. Da hier nur die Abweichungen vom guten Gesundheitszustand betrachtet werden sollen, kann β_{ij_2} für j_2=3 gleich 0 gesetzt und die in der Spalte „Exp(B)“ dargestellten Zahlenwerte als Odds-Ratios relativ zur Referenzkategorie interpretiert werden.

Erwartungsgemäß nimmt mit zunehmenden Alter die Wahrscheinlichkeit für einen ausgezeichneten bzw. sehr

Tabelle 2: SPSS-Output zum multinomialen logistischen Modell

Deskriptive Statistik

	N	Minimum	Maximum	Mittelwert	Standardabweichung
Alter (2011)	1294	21	64	44,76	10,527
Arbeitszufriedenheit	1294	1,00	4,00	2,8241	,52061

Verarbeitete Fälle

		Anzahl	Rand-Prozentsatz
self-rated health	(1) ausgezeichnet	51	3,9%
	(2) sehr gut	241	18,6%
	(3) gut	617	47,7%
	(4) mittelmäßig	342	26,4%
	(5) schlecht	43	3,3%
work-family conflict	(1) stimme voll zu	134	10,4%
	(2) stimme eher zu	352	27,2%
	(3) stimme eher nicht zu	474	36,6%
	(4) stimme gar nicht zu	334	25,8%
höchster Schulabschluss	(1) Hauptschule oder weniger	242	18,7%
	(2) mittlerer Schulabschluss	372	28,7%
	(3) Abitur	680	52,6%
Gesamt		1294	100,0%

Likelihood-Quotienten-Tests

Effekt	Kriterien für die Modellanpassung	Likelihood-Quotienten-Tests		
	-2 Log-Likelihood für reduziertes Modell	Chi-Quadrat	Freiheitsgrade	Signifikanz
Konstanter Term	2691,277[a]	0,000	0	
Alter (2011)	2758,705	67,428	4	,000
Arbeitszufriedenheit	2809,903	118,626	4	,000
work-family conflict	2757,596	66,319	12	,000
höchster Schulabschluss	2745,617	54,340	8	,000

Die Chi-Quadrat-Statistik stellt die Differenz der -2 Log-Likelihoods zwischen dem endgültigen Modell und einem reduziertem Modell dar. Das reduzierte Modell wird berechnet, indem ein Effekt aus dem endgültigen Modell weggelassen wird. Hierbei liegt die Nullhypothese zugrunde, nach der alle Parameter dieses Effekts 0 betragen.

a. Dieses reduzierte Modell ist zum endgültigen Modell äquivalent, da das Weglassen des Effekts die Anzahl der Freiheitsgrade nicht erhöht.

Parameterschätzer

self-rated health[a]		B	Standardfehler	Wald	Freiheitsgrade	Signifikanz	Exp(B)	95% Konfidenzintervall für Exp(B)	
								Untergrenze	Obergrenze
ausgezeichnet	Konstanter Term	-3,242	1,241	6,829	1	,009			
	Alter (2011)	-,044	,015	8,290	1	,004	,957	,928	,986
	Arbeitszufriedenheit	1,218	,338	13,020	1	,000	3,381	1,745	6,553
	[work-family conflict=1]	-1,034	,659	2,460	1	,117	,356	,098	1,294
	[work-family conflict=2]	-1,169	,443	6,952	1	,008	,311	,130	,741
	[work-family conflict=3]	-1,083	,358	9,127	1	,003	,339	,168	,684
	[work-family conflict=4]	0[b]			0				
	[höchster Schulabschluss=1]	-,625	,501	1,556	1	,212	,535	,201	1,429
	[höchster Schulabschluss=2]	-1,021	,399	6,564	1	,010	,360	,165	,787
	[höchster Schulabschluss=3]	0[b]			0				

(Fortsetzung)

Tabelle 2: SPSS-Output zum multinomialen logistischen Modell

self-rated health[a]		B	Standard-fehler	Wald	Freiheits-grade	Signi-fikanz	Exp(B)	95% Konfidenzintervall für Exp(B) Unter-grenze	Ober-grenze
sehr gut	Konstanter Term	-1,485	,664	4,999	1	,025			
	Alter (2011)	-,028	,008	12,262	1	,000	,972	,957	,988
	Arbeitszufriedenheit	,878	,183	22,952	1	,000	2,406	1,680	3,445
	[work-family conflict=1]	-1,514	,442	11,718	1	,001	,220	,093	,524
	[work-family conflict=2]	-,421	,220	3,644	1	,056	,656	,426	1,011
	[work-family conflict=3]	-,647	,198	10,672	1	,001	,524	,355	,772
	[work-family conflict=4]	0[b]			0				
	[höchster Schulabschluss=1]	-1,331	,302	19,365	1	,000	,264	,146	,478
	[höchster Schulabschluss=2]	-1,020	,201	25,789	1	,000	,361	,243	,535
	[höchster Schulabschluss=3]	0[b]			0				
mittelmäßig	Konstanter Term	-,021	,583	,001	1	,971			
	Alter (2011)	,039	,008	25,036	1	,000	1,040	1,024	1,056
	Arbeitszufriedenheit	-1,017	,154	43,919	1	,000	,362	,268	,488
	[work-family conflict=1]	,778	,267	8,508	1	,004	2,177	1,291	3,673
	[work-family conflict=2]	,767	,213	12,976	1	,000	2,154	1,419	3,271
	[work-family conflict=3]	,134	,201	,446	1	,504	1,143	,771	1,695
	[work-family conflict=4]	0[b]			0				
	[höchster Schulabschluss=1]	,098	,197	,248	1	,619	1,103	,749	1,624
	[höchster Schulabschluss=2]	,034	,167	,042	1	,837	1,035	,746	1,435
	[höchster Schulabschluss=3]	0[b]			0				
schlecht	Konstanter Term	-2,970	1,425	4,343	1	,037			
	Alter (2011)	,069	,021	10,943	1	,001	1,071	1,028	1,116
	Arbeitszufriedenheit	-1,545	,330	21,919	1	,000	,213	,112	,407
	[work-family conflict=1]	1,370	,590	5,393	1	,020	3,936	1,238	12,512
	[work-family conflict=2]	1,298	,514	6,380	1	,012	3,661	1,337	10,023
	[work-family conflict=3]	-,029	,551	,003	1	,958	,971	,330	2,859
	[work-family conflict=4]	0[b]			0				
	[höchster Schulabschluss=1]	,957	,433	4,880	1	,027	2,604	1,114	6,089
	[höchster Schulabschluss=2]	,552	,418	1,744	1	,187	1,736	,766	3,936
	[höchster Schulabschluss=3]	0[b]			0				

a. Die Referenzkategorie lautet: gut.

b. Dieser Parameter wird auf Null gesetzt, weil er redundant ist.

guten Gesundheitszustand signifikant ab (OR=0.957 bzw. 0.972). Entsprechend nimmt sie für einen mittelmäßigen bzw. schlechten Gesundheitszustand mit dem Alter signifikant zu (OR=1.040 bzw. 1.071). Beim Schulabschluss, hier als Indikator für einen sozialen Gradienten verwendet, sind ähnlich erwartbare Effekte zu beobachten. Probandinnen mit Schulabschlüssen unterhalb des Abiturs weisen einerseits eine verringerte Wahrscheinlichkeit für einen ausgezeichneten bzw. einen sehr guten Gesundheitszustand auf und andererseits eine erhöhte Wahrscheinlichkeit für einen mittelmäßigen bzw. schlechten Gesundheitszustand. Besonders beim schlechten Gesundheitszustand sind mit OR=2.604 für die Kategorie „Hauptschule oder weniger“ und OR=1.736 für den

Tabelle 3: Klassifikationshäufigkeiten der Probandinnen nach maximaler Outcome-Wahrscheinlichkeit

Beobachteter Gesundheitszustand	Vorhergesagter Gesundheitszustand					
	ausgezeichnet	sehr gut	gut	mittelmäßig	schlecht	**Richtig:**
ausgezeichnet	0	12	38	1	0	0,0%
sehr gut	0	52	177	12	0	21,6%
gut	0	41	503	73	0	81,5%
mittelmäßig	0	8	216	118	0	34,5%
schlecht	0	1	17	25	0	0,0%
Insgesamt:	0,0%	8,8%	73,5%	17,7%	0,0%	52,0%

„mittleren Schulabschluss“ auffällige Effekte zu beobachten.

Die beiden vorrangig interessierenden Prädiktoren Arbeitszufriedenheit und work-family-conflict zeigen jeweils spezifische Effekte auf den Gesundheitszustand. Die Arbeitszufriedenheit weist in den Effekten einen starken Gradienten über den Gesundheitskategorien auf. So ist ein Anstieg der Arbeitszufriedenheit um den Betrag 1 signifikant mit einer Erhöhung der Wahrscheinlichkeit für einen ausgezeichneten (OR=3.381) bzw. einen sehr guten Gesundheitszustand (OR=2.406) und einer signifikanten Absenkung der Wahrscheinlichkeit für einen mittelmäßigen bzw. schlechten Gesundheitszustand (OR=0.362 bzw. 0.213) verbunden. Der work-family-conflict zeigt in den Kategorien (1) bis (3) die Tendenz zur Verringerung der Wahrscheinlichkeiten für einen ausgezeichneten und einen sehr guten Gesundheitszustand, in den Kategorien (1) bis (2) ist dagegen eine signifikante Anhebung der Wahrscheinlichkeiten für einen mittelmäßigen (OR=2.177 und 2.154) und einen schlechten Gesundheitszustand (OR=3.936 und 3.661) zu beobachten. Eine höhere Arbeitszufriedenheit (Anhebung des Scorewertes) ist demnach mit einer generell besseren Ausprägung des Gesundheitszustandes verbunden, während der work-family-conflict in den ersten beiden Kategorien eine starke Assoziation mit einer schlechteren Ausprägung des Gesundheitszustandes aufweist.

SPSS NOMREG hat die Option zur Ausgabe einer Klassifikationsmatrix, in der jede Beobachtungseinheit (Stichprobenelement, Proband, etc.) a-posteriori einer Outcome-Kategorie zugeordnet wird. Die Zuordnung erfolgt nach Berechnung der Kategorienwahrscheinlichkeiten gemäß Formel (3) zu der Kategorie mit der maximalen Wahrscheinlichkeit. In Tabelle 3 sind in einer Kreuztabelle die Häufigkeiten nach beobachtetem und klassifiziertem Gesundheitszustand dargestellt. Es fällt zunächst auf, dass keine Zuordnung zu den Kategorien „ausgezeichnet“ und „schlecht“ erfolgt ist. Dafür sind nach der Zuordnung die Kategorien „sehr gut“ und „mittelmäßig“ unterrepräsentiert (8,8% klassifiziert zu 18,6% beobachtet bzw. 17,7% klassifiziert zu 26,4% beobachtet) sowie die Kategorie „gut“ stark überrepräsentiert (81,5% klassifiziert zu 47,7% beobachtet). Richtig klassifiziert wurden 52% der Probandinnen.

Die MNLR neigt im Allgemeinen dazu, für die am häufigsten besetzte Kategorie die besten (genauesten) Vorhersagen zu liefern [20]. Die Ursache für die geringe Klassifikationsgüte liegt in den differenten Skalenniveaus von Beobachtung und Modell-Outcome. Einerseits werden Kategorien beobachtet, andererseits werden Wahrscheinlichkeiten modelliert. Eine Vorhersage im Sinne einer Zuordnung zu einer Kategorie erfolgt demnach als sekundäre Klassifikation der Beobachtungen auf der Basis geschätzter Wahrscheinlichkeiten, d.h. es sind zwei Ebenen des Prozesses zu unterscheiden. Schendera [21] betont dazu, dass Modellgüte und Vorhersagegenauigkeit einander nicht immer entsprechen müssen. „Der Grund ist darin zu sehen, dass die Maße für die Modellgüte einer logistischen Regression auf den Log-Likelihoods (-2LL) basieren, darin aber nicht das modellspezifische Verhältnis zwischen korrekt und inkorrekt klassifizierten Fällen berücksichtigen“ [21]. Im vorliegenden Beispiel konnte eine starke Assoziation von Arbeitszufriedenheit und work-family-conflict mit dem selbstberichteten Gesundheitszustand gezeigt werden. Eine aussagekräftige Risikoanalyse muss also nicht mit einer zufriedenstellenden Klassifikation einhergehen. Für letztere können sich weitere oder andere Prädiktoren als nützlich erweisen.

4.2 Beispiel für ein stereotypes logistisches Modell

Für ein Beispiel zum stereotypen logistischen Modell sollen hier die Daten von 3.752 berufstätigen Männern der Danish Work Environment Cohort Study [22], [23] des Erhebungsjahres 2005 dienen. DWECS ist eine Längsschnittstudie, die per Fragebogen Arbeits- und Gesundheitsbedingungen in der dänischen Arbeitsbevölkerung untersucht. Jankowiak et al. analysierten das Auftreten von Kniegelenksbeschwerden in Abhängigkeit von der Dauer kniender/hockender Tätigkeit [24], wovon hier nur die Schmerzdauer in Tagen pro Jahr herausgegriffen wird. Die Dauer der Schmerzen im Knie wurden mit der Frage „Hatten Sie in den letzten 12 Monaten Schmerzen oder Beschwerden in einem oder beiden Knien?“ erfasst. Die Antwortmöglichkeiten waren mit den Kategorien „0 Tage“, „1–7 Tage“, „8–30 Tage“, „mehr als 30 Tage“ und „jeden Tag“ vorgegeben. Kniende oder hockende Tätigkeit wurde mit der Frage „Erfordert es ihre Tätigkeit, dass Sie knien oder hocken, wenn Sie arbeiten?“ als Zeitanteil eines Arbeitstages erfasst. Unter Nutzung der Zahl der Wochenarbeitsstunden (reguläre Arbeitszeit plus Überstunden und anderer Arbeiten) wurde die Stunden pro Woche mit kniender/hockender Tätigkeit berechnet. Die daraus resultierende Variable „Kniende/hockende Tätigkeit [Stunden/Woche]“ wurde in 5 Kategorien „>12 Stunden“, „12 bis >8 Stunden“, „8 bis >4 Stunden“, „4 bis >0 Stunden“ und „nie“ mit annähernd homogener Verteilung der Wochenstunden gruppiert.

Als wesentliche Kovariaten wurden das Alter (18–29, 30–39, 40–49, 50–59 und ≥60 Jahre) und der Body-Mass-Index (Normalgewicht: BMI 18,5–<25, Übergewicht: BMI 25–<30 und starkes Übergewicht: BMI ≥30) einbezogen. 34 untergewichtige Männer mit einem BMI <18,5 wurden zuvor wegen des geringen Gruppenumfanges aus dem Datensatz entfernt.

Die Outcome-Variable „Schmerzdauer in Tagen pro Jahr“ ist an sich eine ordinale Variable, intuitiv ist aber zu erwarten, dass die unterschiedlichen Intervalllängen, auf die sich die Kategorien stützen, der proportional-odds-Annahme entgegenstehen. Und tatsächlich wird in einem OLM die proportional-odds-Hypothese abgewiesen. Die strenge Rangordnung der Outcome-Kategorien lässt nun ein eindimensionales SLM als das geeignete Modell erscheinen, wobei sich eine Schmerzdauer von „0 Tage“ als Referenzkategorie anbietet. Eine Modellschätzung in STATA 13 SLOGIT zeigt jedoch, dass sich dieses Modell nicht signifikant vom Null-Modell unterscheidet:

```
1 dim                           Wald chi2(10) =  14.41
Log likelihood = -4432.9451     Prob > chi2   = 0.1549
```

Ein zweiter Modellansatz mit 2 Dimensionen unterscheidet sich im Likelihood-Quotienten-Test signifikant vom ersten Ansatz und ist deutlich signifikant vom Null-Modell verschieden:

```
2 dim                           Wald chi2(20) =  63.75
Log likelihood = -4421.2141     Prob > chi2   = 0.0000
```

Ein dritter Ansatz mit drei Dimensionen führte zu nichtkonvergentem Verhalten des Schätzalgorithmus. Parallel zu den beiden stereotypen Modellen wurden auch ein OLM und ein MNLM gerechnet und aus den Log-Likelihood-Werten die Werte für Akaike’s Informationskriterium (AIC) bestimmt:

- SLM (dim=2): AIC = 8898,42
- SLM (dim=1): AIC = 8943,54
- OLM: AIC = 8954,44
- MNLM: AIC = 8959,34

Auch hier zeigt sich, dass das zweidimensionale SLM ein deutlich besser angepasstes Modell (vgl. [25]) liefert als das eindimensionale SLM. Das eindimensionale Modell an sich stellt zumindest ein geringfügig besseres Modell als das OLM und das MNLM dar.

Der STATA-Output zum 2-dimensionalen stereotypen logistischen Modell ist in Tabelle 4 dargestellt. Oben im Output sind die Restriktionen an die φ-Parameter aufgelistet, die für das 2-dimensionale Modell gesetzt werden. Aus diesen Restriktionen ergibt sich, dass die Effekte der Prädiktoren auf die 2. Outcome-Kategorie nur durch die β-Schätzer der 1. Dimension bestimmt werden und analog die Effekte der Prädiktoren auf die 3. Outcome-Kategorie nur durch die β-Schätzer der 2. Dimension. Die Effekte der Prädiktoren auf die 4.und 5. Outcome-Kategorie beruhen auf den Parameterschätzern beider Dimensionen. In der Output-Tabelle sind die Modellparameter entsprechend dem implementierten Modellansatz (15) dargestellt, wobei die β- und die φ-Parameter für zwei Dimensionen vorhanden sind. Für eine sachgemäße Modellinterpretation muss bei den β_i-Parametern das Vorzeichen umgekehrt werden.

Zur Interpretation der Modellparameter in Form von Odds-Ratios muss Formel (10) um eine Dimension zu

$$OR_{j_1 j_2}(\Delta X_i)=\exp\left\{(\varphi_{j_1}^{(1)}-\varphi_{j_2}^{(1)})\beta_i^{(1)}+(\varphi_{j_1}^{(2)}-\varphi_{j_2}^{(2)})\beta_i^{(2)}\right\} \quad \textbf{(16)}$$

erweitert werden. Setzt man für j_1 sukzessive alle Kategorien mit Schmerzdauerangaben ab mindestens einem Tag und für j_2 konstant die Kategorie ohne Schmerzen (0 Tage) ein, ergeben sich die in Tabelle 5 aufgeführten Odds-Ratios. Diese lassen sich wie vier simultane binäre logistische Modelle lesen, wobei jede Spalte in Relation zur Referenzkategorie zu sehen ist. Die Basis innerhalb der Prädiktoren (OR=1 gesetzt) bilden die 18–29jährigen, normalgewichtigen Männer ohne kniende/hockende Tätigkeit. In den Spalten von Tabelle 5 sind die Odds-Ratios also mit Blick auf die Basiskategorien der Prädiktoren zu lesen. Bei einer Schmerzdauer von 1–7 Tagen ist bereits ein deutlicher Einfluss der Dauer der knienden oder hockenden Tätigkeit festzustellen, während das Alter und der BMI praktisch keinen Einfluss haben. Ab einer Schmerzdauer von 8 Tagen wird der Einfluss der Arbeit in jeder weiteren Outcome-Kategorie stärker und auch das Alter und das Übergewicht weisen einen zunehmenden Anstieg der Effekte auf. Eine Ausnahme bildet die Altersklasse mit 60 und mehr Jahren, hier ist ein Rückgang des Alterseffektes zu beobachten, der vermutlich auf Selektion beruht. D.h. ältere Arbeitnehmer mit häufigen Schmerzen verringern die Dauer belastender Arbeitsaufgaben, suchen eine andere Tätigkeit oder steigen aus dem Berufsleben aus.

In der untersten Zeile von Tabelle 5 sind die Basis-Odds dargestellt. Aus diesen lassen sich die Prävalenzen für Knieschmerz bei Probanden der Basisgruppe innerhalb jeder Outcome-Kategorie berechnen. Die Prävalenz schmerzfreier Probanden erhält man als Differenzbetrag der Summe dieser vier Prävalenzwerte zu 1.

Tabelle 4: STATA-Output zum 2-dimensionalen stereotypen logistischen Modell

```
Stereotype logistic regression                      Number of obs   =       3752
                                                    Wald chi2(20)   =      63.75
Log likelihood = -4421.2141                         Prob > chi2     =     0.0000

 ( 1)  [phi1_2]_cons = 1
 ( 2)  [phi1_3]_cons = 0
 ( 3)  [phi2_2]_cons = 0
 ( 4)  [phi2_3]_cons = 1
------------------------------------------------------------------------------------------
      Knieschmerz (Dauer)       Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
------------------------------------------------------------------------------------------
dim1
Alter                  60+    .0355068   .2474936     0.14   0.886    -.4495719    .5205854
                     50-59    .0531057   .1494215     0.36   0.722    -.2397551    .3459664
                     40-49   -.0512577   .1395202    -0.37   0.713    -.3247122    .2221969
                     30-39   -.0405772   .1231948    -0.33   0.742    -.2820345    .2008802
BMI
übergewichtig, 25 - < 30     -.0490023   .0811235    -0.60   0.546    -.2080015    .1099968
stark übergewichtig, 30+     -.0924608   .1556052    -0.59   0.552    -.3974414    .2125198

Knien/Hocken (Std./Woche)
                 >12 Std.    -.6803535   .1689842    -4.03   0.000    -1.011556   -.3491506
             >8 - 12 Std.    -.6366934   .1513047    -4.21   0.000    -.9332452   -.3401417
              >4 - 8 Std.    -.1865681   .1539731    -1.21   0.226    -.4883499    .1152137
               bis 4 Std.    -.2957482   .1271158    -2.33   0.020    -.5448907   -.0466058
------------------------------------------------------------------------------------------
dim2
Alter
                       60+   -.2414312   .1996781    -1.21   0.227    -.6327931    .1499306
                     50-59   -.4925973   .1741147    -2.83   0.005    -.8338558   -.1513389
                     40-49   -.4240126   .1345581    -3.15   0.002    -.6877415   -.1602836
                     30-39   -.3676669   .1231286    -2.99   0.003    -.6089945   -.1263392

BMI
übergewichtig, 25 - < 30     -.1547771   .0645487    -2.40   0.016    -.2812902    -.028264
stark übergewichtig, 30+     -.3843202   .1163123    -3.30   0.001    -.6122881   -.1563524

Knien/Hocken (Std./Woche)
                 >12 Std.    -.9850119   .1685972    -5.84   0.000    -1.315456   -.6545675
             >8 - 12 Std.    -.7622529   .1709361    -4.46   0.000    -1.097281   -.4272244
              >4 - 8 Std.    -.2745558    .114008    -2.41   0.016    -.4980073   -.0511042
               bis 4 Std.    -.2685529   .1005285    -2.67   0.008    -.4655851   -.0715207
------------------------------------------------------------------------------------------
                  /phi1_1            0  (base outcome)
                  /phi1_2            1  (constrained)
                  /phi1_3            0  (omitted)
                  /phi1_4     .2999973   .7478273     0.40   0.688    -1.165717    1.765712
                  /phi1_5    -2.077455   2.473508    -0.84   0.401    -6.925442    2.770532
------------------------------------------------------------------------------------------
                  /phi2_1            0  (base outcome)
                  /phi2_2            0  (omitted)
                  /phi2_3            1  (constrained)
                  /phi2_4     1.446282   .4963307     2.91   0.004     .4734916    2.419072
                  /phi2_5     3.349778    1.55246     2.16   0.031     .3070129    6.392543
------------------------------------------------------------------------------------------
                  /theta1            0  (base outcome)
                  /theta2    -1.518573    .120184   -12.64   0.000    -1.754129   -1.283017
                  /theta3    -2.340645   .1478845   -15.83   0.000    -2.630493   -2.050796
                  /theta4    -2.992126   .1868225   -16.02   0.000    -3.358291    -2.62596
                  /theta5      -4.5673   .4461793   -10.24   0.000    -5.441795   -3.692805
------------------------------------------------------------------------------------------
(Knieschmerz (Dauer) = "0 Tage" is the base outcome)
```

Tabelle 5: Aus den Modellparametern errechnete Odds-Ratios für Knieschmerzen bezogen auf schmerzfreie Probanden
(Basis: 18–29jährige, normalgewichtige Männer ohne kniende/hockende Tätigkeit)

	Knieschmerz (Dauer)			
	1–7 Tage	8–30 Tage	> 30 Tage	jeden Tag
Alter				
30–39	1.041	1.444	1.723	3.150
40–49	1.053	1.528	1.875	3.721
50–59	0.948	1.637	2.007	5.815
60+	0.965	1.273	1.403	2.417
BMI				
übergewichtig	1.050	1.167	1.269	1.517
stark übergewichtig	1.097	1.469	1.792	2.990
Knien/Hocken (Std./Woche)				
bis 4 Std.	1.344	1.308	1.611	1.330
>4–8 Std.	1.205	1.316	1.573	1.703
>8–12 Std.	1.890	2.143	3.645	3.424
>12 Std.	1.975	2.678	5.097	6.594
Basis-Odds				
exp(theta)	0.219	0.096	0.050	0.010

5 Diskussion

Ordinale und multinomiale logistische Modelle sind etablierte Verfahren in der statistischen Praxis und in allen großen Statistikprogrammpaketen enthalten. Dem stereotypen Modell dagegen wurde vor rund eineinhalb Jahrzehnten ein kümmerliches Dasein bescheinigt [14], daran scheint sich bis heute wenig geändert zu haben. Modernere Lehrbücher erwähnen das SLM nur kurz, z.B. [7], [26]. Bender und Grouven waren der Auffassung, dass es aufgrund von Berechnungsschwierigkeiten noch eine längere Zeit dauern wird, bis das stereotype Modell in der medizinischen Forschung eingesetzt werden kann [27]. Diese Schwierigkeiten waren aber wenige Jahre danach überwunden [14], [8].

Anderson bezeichnete das stereotype Modell als „important for assessed or judged ordered response variables“ [5]. Solche Variablen beruhen nicht auf einer Messung, deren Werte anschließend kategorisiert wurden, sie beruhen auf Bewertungen anhand einer oder mehrerer Eigenschaften eines Untersuchungsgegenstandes und eines anschließenden „Urteils“ über die Kategorienzugehörigkeit. Ein solches Urteil kann durch einen Untersucher oder durch einen Probanden erstellt werden. In letzterem Fall ist in Betracht zu ziehen, dass die Zuordnung zu einer Kategorie auf der Basis individueller Bewertungsmaßstäbe erfolgen kann und nicht auf der Basis eines standardisierten und trainierten Verfahrens, wie es ein Studienmitarbeiter ausführen würde. Entsprechend können für unterschiedliche Probandengruppen Zusammenhänge zu jeweils anderen Prädiktorengruppen bestehen. Die Analyse einer solchen Zielvariablen in einem Regressionsmodell kann diese Zusammenhänge mehr oder weniger gut erfassen, je nachdem ob die Dimension des Problems entsprechend angesetzt wurde. Das stereotype Modell bietet im Gegensatz zum ordinalen und multinomialen Modell den Vorteil, dass man sich nicht a priori auf eine bestimmte Dimension des Zusammenhanges festlegen muss. Im Prozess der Datenanalyse kann die erforderliche Dimension des Models bei 1 beginnend ermittelt werden. Das SLM ist somit ein geeignetes Modell für die Analyse von Fragebogendaten.

Vermutlich erfüllen nur wenige Datensätze aus der Medizin oder den Gesundheitswissenschaften die proportional-odds-Bedingung, andererseits wird auch kaum ein Datensatz den Einsatz eines Modells mit der Dimension J-1 erfordern. Das stereotype Modell sollte vielen Daten aus diesen Disziplinen gerecht werden und dem Untersucher ein Werkzeug in die Hand geben, mit dem er sich der tatsächlichen Dimension seines Problems Schritt für Schritt nähern kann. Mit der Implementierung in STATA ist das SLM seit einigen Jahren einem großen Nutzerkreis zugänglich, einem vielfältigen Einsatz zur Datenanalyse steht also nichts mehr entgegen. Die vorliegende Arbeit ist insbesondere ein Versuch, den Nutzen des SLM für die medizinische und gesundheitswissenschaftliche Forschung herauszustellen.

Anmerkung

Danksagung

Der Autor dankt Frau Dipl.-Math. Anne Pohrt für die kritische Durchsicht des Manuskriptes.

Interessenkonflikte

Der Autor erklärt, dass er keine Interessenkonflikte in Zusammenhang mit diesem Artikel hat.

Literatur

1. Subramanian SV, Huijts T, Avendano M. Self-reported health assessments in the 2002 World Health Survey: how do they correlate with education? Bull World Health Organ. 2010 Feb;88(2):131-8. DOI: 10.2471/BLT.09.067058
2. Manor O, Matthews S, Power C. Dichotomous or categorical response? Analysing self-rated health and lifetime social class. Int J Epidemiol. 2000 Feb;29(1):149-57. DOI: 10.1093/ije/29.1.149
3. Döring N, Bortz J. Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften. 5. Aufl. Berlin: Springer; 2016. DOI: 10.1007/978-3-642-41089-5
4. Drechsel-Schlund C, Butz M, Haupt B, Plinske W, Francks HP. Asbestverursachte Berufskrankheiten in Deutschland – Entstehung und Prognose. Sankt Augustin: HVBG; 2003.
5. Anderson JA. Regression and ordered categorical variables (with discussion). J R Statist Soc B. 1984;46(1):1-30. Available from: www.jstor.org/stable/2345457
6. Lunt M. Prediction of ordinal outcomes when the association between predictors and outcome differs between outcome levels. Stat Med. 2005 May;24(9):1357-69. DOI: 10.1002/sim.2009

7. Tutz G. Die Analyse kategorialer Daten: Anwendungsorientierte Einführung in Logit-Modellierung und kategoriale Regression. München, Wien: R. Oldenbourg; 2000.
8. Lunt M. Stereotype ordinal regression. STATA Technical Bulletin. 2001;61(sg163):12-8.
9. Liu I, Agresti A. The analysis of ordered categorical data: An overview and a survey of recent developments. Test. 2005;14(1):1-73. DOI: 10.1007/BF02595397
10. Große Schlarmann J, Galatsch M. Regressionsmodelle für ordinale Zielvariablen. GMS Med Inform Biom Epidemiol. 2014;10(1):Doc05. DOI: 10.3205/mibe000154
11. Abreu MN, Siqueira AL, Caiaffa WT. Regressão logística ordinal em estudos epidemiológicos [Ordinal logistic regression in epidemiological studies]. Rev Saude Publica. 2009 Feb;43(1):183-94. DOI: 10.1590/S0034-89102009000100025
12. Fullerton AS. A conceptual framework for ordered logistic regression models. Sociol Methods Res. 2009;38(2):306-47. DOI: 10.1177/0049124109346162
13. Long JS, Freese J. Regression models for categorical dependent variables using Stata. 2nd ed. Texas: Stata Press; 2006.
14. Kuss O, Hendricks J. Ein SAS-Makro zur Schätzung des Stereotype Regressionsmodells. 4. Konferenz für SAS-Anwender in Forschung und Entwicklung (KSFE); 9.-10. März 2000; Justus-Liebig-Universität Gießen.
15. Kuss O. On the estimation of the stereotype regression model. Comput Stat Data Anal. 2006;50:1877-90. DOI: 10.1016/j.csda.2005.02.013
16. High R. Models for ordinal response data. SAS Global Forum; 2013 Apr 28 - May 1; San Francisco. Statistics and Data Analysis, Paper 445-2013.
17. Liu X. Fitting stereotype logistic regression models for ordinal response variables in educational research (Stata). J Modern Appl Stat Met. 2014;13(2):528-45.
18. Formazin M. JCQ 2.0: Testentwicklung und -validierung. Arbeit und Gesundheit - Befragung in Bochum. BauA Aktuell. 2013;1:5.
19. Bullinger M. Erfassung der gesundheitsbezogenen Lebensqualität mit dem SF-36-Health Survey. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2000;43:190–7. DOI: 10.1007/s001030050034
20. Hosmer DW, Lemeshow S. Applied Logistic Regression. New York: John Wiley & Sons; 2000. DOI: 10.1002/0471722146
21. Schendera CFG. Regressionsanalyse mit SPSS. München: Oldenbourg; 2008. DOI: 10.1524/9783486710625
22. Burr H, Bjorner JB, Kristensen TS, Tüchsen F, Bach E. Trends in the Danish work environment in 1990-2000 and their associations with labor-force changes. Scand J Work Environ Health. 2003 Aug;29(4):270-9. DOI: 10.5271/sjweh.731
23. Feveile H, Olsen O, Burr H, Bach E. Danish Work Environment Cohort Study 2005: from idea to sampling design. Stat Trans. 2007;8(3):441-58.
24. Jankowiak S, Kersten N, Burr H, Liebers F, Latza U. Berufliche Kniebelastung und späteres Auftreten von Knieschmerzen in der Danish Work Environment Cohort Study 1990–2005. In: 8. Jahrestagung der Deutschen Gesellschaft für Epidemiologie und 1. Internationales LIFE Symposium Abstractband: Was hält uns gesund? Was macht uns krank? Mit Epidemiologie den Zivilisationskrankheiten auf der Spur; 24. - 27. September 2013; Leipzig. ID: 241.
25. Burnham KP, Anderson DR. Model selection and multi-model inference: a practical information-theoretic approach. 2nd ed. New York: Springer; 2002.
26. Hilbe JM. Logistic Regression Models. Boca Raton: CRC; 2011.
27. Bender R, Grouven U. Ordinal logistic regression in medical research. J R Coll Physicians Lond. 1997 Sep-Oct;31(5):546-51.

Korrespondenzadresse:

Dr. Norbert Kersten
Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Fachbereich Arbeit und Gesundheit, Nöldnerstrasse 40–42, 10317 Berlin, Germany
kersten.norbert@baua.bund.de

Bitte zitieren als
Kersten N. Multinomiale, ordinale und stereotype logistische Regression – eine Einführung in die Regressionsanalyse kategorialer Zielvariablen. GMS Med Inform Biom Epidemiol. 2016;12(1):Doc01. DOI: 10.3205/mibe000163, URN: urn:nbn:de:0183-mibe0001635

Artikel online frei zugänglich unter
http://www.egms.de/en/journals/mibe/2016-12/mibe000163.shtml

Veröffentlicht: *18.02.2016*