Regressionsmodelle für ordinale Zielvariablen

Regression models for ordinal response variables

Abstract

Regression models for ordinal response variables are available since the late 1980s, and many statistical software packages have been upgraded to perfom these calculations. However, ordinal regression still seems to be overlooked in health sciences. This might be due to the available literatur, which is critizised to be too technical and hard to understand for endusers. Thus, the aim of this paper is to give an overview of popular ordinal regression models and to outline the advantages of its model results. First, the characteristics and interpretations of the "proportional odds model" are presented. The model is exemplary applied to variables of a dataset. A focus is set to the interpretation of the model results. Secondly, the characteristics and interpretations of the "continuation ratio model" are presented. Again, the model is exemplary applied to variables of a dataset, while the focus is set to the interpretation of the model results. The results of both models are contrasted to results of ordinary logistic regression and discussed with respect of their differences.

Keywords: ordinal regression, proportional odds, continuation ratio

Zusammenfassung

Verfahren zur Auswertung ordinaler Variablen mittels Regressionsanalysen sind seit ca. 25 Jahren in der Literatur beschrieben, und viele Statistik-Softwarepakete wurden in den letzten Jahren um die Funktion dieser speziellen Art der Regression erweitert. Dennoch hält sich der Einsatz der Verfahren bisweilen stark in Grenzen. Dies mag daran liegen, dass die vorhandene Literatur zu spezialisiert und für den Anwender nur schwer verständlich ist. Ziel des Artikels ist es daher, die gängigen Verfahren der ordinalen Regression vorzustellen sowie den Nutzen ihrer Ergebnisse an Beispielen zu verdeutlichen. Der Artikel versteht sich als praktische Einführung in das Thema und beschreibt nacheinander die Methoden der „Proportional Odds Models“ sowie der „Continuation Ratio Models“. Beide Methoden werden beispielhaft angewendet, wobei ein Fokus auf die Interpretation der Modellergebnisse gelegt wird. Die Modellergebnisse beider Verfahren werden den Ergebnissen von alternativen logistischen Regressionsanalysen gegenübergestellt und hinsichtlich ihrer Unterschiede diskutiert.

Schlüsselwörter: ordinale Regression, Proportional Odds, Continuation Ratio

Jörg große Schlarmann[1]
Michael Galatsch[1]

1 Department Pflegewissenschaft, Forschungskolleg FamiLe, Universität Witten/Herdecke, Witten, Deutschland

Einleitung

Skalen mit ordinalem Datenniveau sind fester Bestandteil in gesundheitswissenschaftlichen Studien. Die Abbildung des Gesundheitsstatus von Probanden in den Kategorien *gesund – leicht krank – krank – schwer krank* ist ein Beispiel für eine ordinale Variable. Auch ursprünglich metrische Messgrößen (z.B. *Körpergröße)* können in verschiedene Stufen ordinal operationalisiert werden (z.B. *sehr klein, klein, normal, groß, sehr groß*). Die geordnete Reihenfolge der Stufenausprägungen ist ein wichtiges Merkmal ordinaler Daten. Ein Ziel der Auswertung kann es sein herauszufinden, welche Faktoren das untersuchte Outcome (z.B. *Gesundheitsstatus* oder *Körpergröße)* beeinflussen, und wie stark sich dieser Einfluss darstellt. Zur Beantwortung solcher Frage eignen sich die statistischen Verfahren der Regressionsanalysen [1], ([2], S.407ff). Hierbei entscheidet das Skalenniveau der Ziel-

variablen, welche Art der Regression durchgeführt werden kann (z.B. lineare Regression, Varianzanalyse, logistische Regression).
Obwohl Verfahren zur Durchführung einer Regressionsanalyse bei ordinalen Zielvariablen seit längerer Zeit verfügbar sind [3], [4], [5], [6], [7], hält sich deren Einsatz in Grenzen. Dies kann unter anderem daran liegen, dass die verfügbare Literatur zu spezialisiert und für Anwender nur schwer verständlich ist (vgl. [8], [9]).
Aus diesem Grund soll der vorliegende Artikel einen Einstieg in die gängigen Verfahren der ordinalen Regressionsanalyse geben sowie ihren Nutzen in der Anwendung verdeutlichen. Eine vertiefende Beschreibung inklusive mathematischer Formeln findet sich bei Harrell [6] und Agresti [7].

Ordinale Regression

Eine häufig angewandte Variante zum Umgang mit ordinal-skalierten Outcomes ist, das Zielkriterium zu dichotomisieren. Hierbei werden die ursprünglichen Stufen der ordinalen Variable in zwei Stufen kategorisiert.
Tabelle 1 zeigt die Variable *Gesundheitsstatus* einer fiktiven Untersuchung, in welcher der Gesundheitsstatus von 1.200 Personen erfasst wurde. Die Ausprägung des Gesundheitsstatus liegt in den fünf Stufen *gesund, leicht krank, krank, schwer krank, todkrank* vor. Für eine Analyse dieser Daten könnte es beispielsweise sinnvoll sein, die Gruppe der *gesunden* mit der Gruppe der *erkrankten* zu vergleichen. In Tabelle 2 wurde der Gesundheitsstatus daher in zwei Antwortgruppen dichotomisiert.
Die Zielvariable (*gesund / krank*) könnte auf diesem Weg mittels logistischer Regressionsanalyse ausgewertet werden. Der Nachteil der Dichotomisierung besteht darin, dass Informationen über die geordnete ordinale Reihe verloren gehen.
Eine weitere Variante besteht darin, die Ausprägungen eines ordinalen Zielkriteriums in Zahlen umzuschreiben (z.B. 1,2,3,...,*j*; vgl. Tabelle 3). Dies würde die Möglichkeit eröffnen, die Variable mittels linearer Regression und Varianzanalysen auszuwerten.
Der Nachteil dieser Variante liegt darin, dass den Daten mehr Informationen unterstellt werden, als sie tatsächlich beinhalten: zwar liegen die ordinalen Daten in einer geordneten Reihe vor, jedoch sind die Abstände zwischen den einzelnen ordinalen Stufen *nicht* bekannt. Darüber hinaus handelt es sich um *diskrete* Werte, da jeder ordinalen Stufe eine natürliche ganze Zahl zugeordnet wird. Lineare Modelle setzen *kontinuierliche* Werte voraus, so dass solche Modelle auch Aussagen über (Zwischen-)Stufen treffen können, die in der Zielvariable nicht existieren (z.B. 2,13; 4,97; oder 7). Falls in einer Stufe des ordinalen Ziels viele Fälle vorliegen, ist es wahrscheinlich, dass die Varianz in den einzelnen Stufen heterogen ist (so genannte *Heteroskedaszität*). Hierauf reagieren die gängigen Gütetests der Regression (F-Test und t-Test) sehr sensitiv, so dass die Gültigkeit solcher Modelle abgelehnt würde. Aus den genannten Gründen ist es in keinem Fall zulässig, eine Variable nachträglich durch Zahlenkodierung zu metrisieren.

Modelle der ordinalen Regression

Die Motivation für eine *ordinale* Regressionsanalyse liegt darin, alle verfügbaren Informationen der geordneten Reihe bei der Beschreibung der Einflussgrößen auf das ordinale Outcome zu verwenden. Hierfür stehen hauptsächlich zwei Verfahren zur Verfügung: das *Proportional Odds Model* (PO) sowie das *Continuation Ratio Model* (CR). Die Modelle schätzen die Wahrscheinlichkeit, eine bestimmte Stufe (oder darüber hinaus) des ordinalen Ziels zu erreichen.
Nachfolgend werden beide Verfahren gesondert vorgestellt.

Proportional Odds Model

Das *Proportional Odds Modell* (PO) ist in der Literatur unter weiteren Synonymen bekannt (vgl. [9], [10]), wie beispielsweise *cumulative odds model* [5], *ordinal logistic model* [11], *cumulative logit model* [12] oder *McCullagh's grouped continuous model* [4], [3]. Es ist eine Erweiterung des binär-logistischen Regressionsmodells für Zielvariablen, deren Ausprägungen mehr als zwei ordinale Stufen beinhalten. Dieses Modell wird vor allem dann verwendet, wenn dem ordinalen Zielkriterium eine latente metrische Größe zu Grunde liegt. Beispielsweise ist die *Körpergröße in Metern* ein metrisches Konstrukt. Wird die *Körpergröße* nicht in Metern, sondern in den Stufen *klein – mittel – groß* erfasst, so erhält man eine ordinale Variable, der ein latentes metrisches Konstrukt zu Grunde liegt. Stetige Zielgrößen sollten jedoch nie ohne Not in Kategorien eingeteilt werden, da dies zu einem unnötigen Informationsverlust führt.
Ein weiteres Einsatzgebiet für PO stellen likert-skalierte Daten dar, da ihnen typischerweise *quasi-metrische* Ausprägungen – bzw. equi-distante Abstände – unterstellt werden ([13], S. 224).
Das PO betrachtet die Wahrscheinlichkeit, eine ranghöhere Stufe des ordinalen Ziels zu erreichen, relativ zur Wahrscheinlichkeit in der aktuellen Stufe zu verbleiben oder in eine niedrigere Stufe zu gelangen. Bei einer ordinalen Variable mit 5 Stufen ergeben sich so 4 Möglichkeiten (Stufenpaare), diese Wahrscheinlichkeiten zu betrachten (vgl. Tabelle 4).
Die Grundannahme des PO (und ebenfalls des CR) besagt, dass die Beziehung zwischen jeden denkbaren zwei Stufenpaaren des ordinalen Ziels statistisch gleich ist ([7], S.46ff; [6], S.333). Dann lässt sich der Einfluss der unabhängigen Variablen jeweils durch *einen* universellen β-Koeffizienten beschrieben, der für jeden Stufenwechsel innerhalb des ordinalen Ziels gültig ist (so genannte „proportional odds" oder „equal slopes"). Aus der Annahme der proportionalen Odds folgt, dass die Abstände zwischen den Stufen des ordinalen Ziels keinen Einfluss auf die Koeffizienten haben. Abbildung 1 veranschaulicht

Tabelle 1: Variable Gesundheitsstatus

	gesund	leicht krank	krank	schwerkrank	todkrank	TOTAL
n	170	230	410	270	120	1.200

Tabelle 2: Dichotomisierte Daten

	gesund	krank	TOTAL
n	170	1.030	1.200

Tabelle 3: Linearisierte Rohdaten

	1	2	3	4	5	TOTAL
n	170	230	410	270	120	1.200

Tabelle 4: Stufenpaare im Proportional Odds Modell

Stufe	gesund	leicht krank	krank	schwerkrank	todkrank	TOTAL
n	170	230	410	270	120	1.200
Paar 1	170	1.030				1.200
Paar 2	400		800			1.200
Paar 3	810			390		1.200
Paar 4	1.080				120	1.200

die „equal slope assumption“: die logistischen Funktionen für die Wahrscheinlichkeit, mindestens Stufe *j* zu erreichen, verläuft für jede Stufe parallel verschoben mit derselben Steigung. Der jeweilige Stufenwechsel (parallele Verschiebung) hat keinen Einfluss auf den Koeffizienten (Steigung der Funktionen).

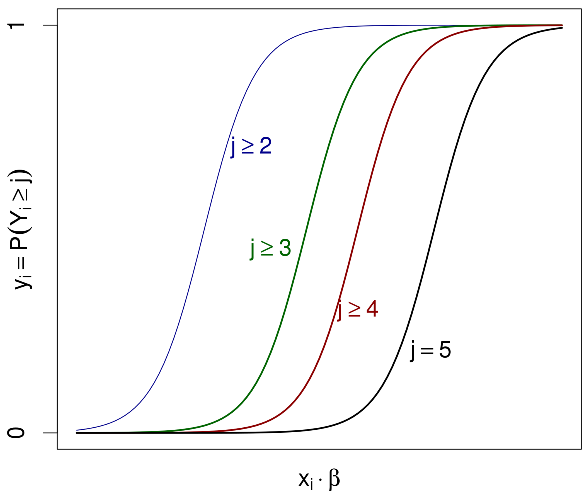

Abbildung 1: Illustration der „proportional odds assumption“ bei 5 Antwortstufen. Der jeweilige Stufenwechsel (parallele Verschiebung) hat keinen Einfluss auf den Koeffizienten (Steigung der Funktionen).

Das klingt zunächst verwirrend, da es ein Merkmal von ordinalen Daten ist, dass die Abstände zwischen den einzelnen Stufen nicht equi-distant sind. Als Beispiel seien hier Schulnoten erwähnt: die Ausprägung der Note *ausreichend* (z.B. die Fehlerspanne, die sich ein Schüler leisten kann) ist eine andere als die der Note *sehr gut*. Jedoch wird das Erreichen einer bestimmten Note von mehreren Faktoren beeinflusst. Während beispielsweise die reine *Lernzeit* für den Wechsel von der Note *ausreichend* auf die Note *befriedigend* wahrscheinlich anders ist als beim Wechsel von *gut* nach *sehr gut*, könnten *finanzielle Ressourcen* (z.B. für Lehrmaterial oder Nachhilfe) einen equidistanten Einfluss auf die erreichte Note haben. Daher muss die Gültigkeit der „equal slopes assumption“ (bzw. der „proportional odds“) zunächst für jeden Prädiktor überprüft werden, bevor das errechnete Modell als gültig angesehen werden kann. Eine Methode hierfür ist die Berechnung eines Regressionsmodells *mit* und eines *ohne* der Annahme von „equal slopes“. Wenn „proportional odds“ vorliegen, ist die Devianz (und somit auch die log-Likelihood) beider Modelle annähernd gleich. Dies kann mittels χ^2-Tests überprüft werden ([7], S.68), [10]. Die Nullhypothese besagt, dass proprotional odds vorliegen. Allerdings muss beachtet werden, dass die Tests zum Modellvergleich keine große Power haben, so dass es trotz Unterschied zu einem p-Wert größer 0.05 kommen kann (beta-Fehler).

Kann die Annahme lediglich für *einige* Prädiktoren im Modell bestätigt werden, stehen alternativ die so genannten *Partial Proportional Odds*-Modelle (PPO) zur Verfügung [14]. Diese liefern für Prädiktoren, welche die „equal slopes assumption“ nicht erfüllen, gesonderte Koeffizienten pro Stufenwechsel.

Proportional Odds Modelle haben folgende Vorteile (vgl. [9]):

- der universelle Koeffizient ändert sich nicht, falls Stufen zusammengefasst oder deren Definitionen verändert werden.
- bei einer Umkehrung des Stufenwechsels (z.B. von *hoch* nach *niedrig* statt von *niedrig* nach *hoch*) ändert sich lediglich das Vorzeichen des Koeffizienten.
- das Modell kann auch bei wenig besetzten Kategorien sinnvoll angewendet werden.
- der Koeffizient ist leicht interpretierbar, da er den Effekt für jeden Stufenwechsel innerhalb der ordinalen Zielvariable zusammenfasst.
- es erfolgt keine α-Fehler-Kumulation wie beim multiplen Testen.

Daher ist das PO das derzeit meist-genutzte Modell zur Berechnung von ordinalen Regressionen [9], [15].

Eine Mindest-Fallzahl zur Durchführung eines PO ist in der Literatur nicht beschrieben, so dass auf klassische Methoden der Fallzahlkalkulation zurückgegriffen werden muss.

Beispiel NEXT

Um die Anwendung des PO zu veranschaulichen, werden im folgenden Beispiel Daten von deutschen Altenpflegenden aus der NEXT-Studie zum Erhebungszeitpunkt t_0 (2002–2003) (vgl. [16]) verwendet. Die Auswertungen dienen ausschließlich als Beispiel und haben keinen Anspruch auf Vollständigkeit. Eine detailierte Auswertung der NEXT-Daten finden sich bei Hasselhorn et al. 2005 [16]. Als Zielvariable dient die ordinale Version des *Work Ability Index* (*WAI*) [17]. Mit diesem Instrument ist es möglich, die Übereinstimmung der eigenen Fähigkeiten mit den Anforderungen des Arbeitsplatzes zu erfassen. Der Index reicht von 7 bis 49 Punkte, wobei 7–27 Punkte eine schlechte, 28–36 Punkte eine mäßige, 37–43 eine gute und 44 bis 49 Punkte eine sehr gute Arbeitsfähigkeit bedeuten ([18], S. 19). Als mögliche Prädiktoren wurden die Variablen *Arbeitszufriedenheit* (*JobSatisf*) und der *Arbeit/Familie Konflikt* (*ConfWorkFam*) in das Regressionsmodel eingefügt. Die Arbeitszufriedenheit wurde mit vier Fragen des Copenhagen Psychosocial Questionnaire (COPSOQ) [19] erhoben. Hohe Werte stehen dabei für eine hohe Arbeitszufriedenheit. Mit der Variablen *ConfWorkFam* wurde operationalisiert, inwieweit die Arbeit der Befragten mit dem Familienleben konfligiert [20]. Die Skala besteht aus 5 Items und ein hoher Wert zeigt an, dass die Arbeit das Familienleben stark beeinträchtigt.

Tabelle 5 beschreibt den Datensatz hinsichtlich der Fälle für jede Stufe des *WAI* sowie der entsprechenden Mittelwerte und Standardabweichungen der Variablen *ConfWorkFam* und *JobSatisf*. Es ist erkennbar, dass eine niedrige *WAI*-Stufe mit hohen *ConfWorkFam*-Werten und niedrigen *JobSatisf*-Werten einhergeht. Entsprechend sind bei hoher *WAI*-Stufe niedrige *ConfWorkFam*-Werte und hohe *JobSatisf*-Werte abzulesen.

Unter Annahme der „equal slopes assumption“ wird im Folgenden ein PO-Modell erstellt, welches den Einfluss von *ConfWorkFam* und *JobSatisf* auf den *WAI* darstellen soll. Die Berechnungen wurden mit Hilfe der freien Software R [21] und dessen Paket *VGAM* [10] erstellt. Die Ergebnisse sind in Tabelle 6 dargestellt.

Die Tabelle beginnt mit allgemeinen Werten zur Modellbeschreibung. *logLike* ist die geschätze Log-Likelihood des Modells. Die *Devianz* entspricht der Summe der Quadrierten Devianz-Abweichungen und berechnet sich aus –(2· *logLike*). Durch das Regressionsmodell wird die Log-Likelihood Schätzung maximiert und somit die Devianz minimiert (was einer Verbesserung entspricht). Beide Werte können zur Prüfung der „equal slopes assumption“ verwendet werden. *AIC* entspricht dem Informationskriterium nach Hirotugu Akaike. Es dient zum Vergleich verschiedener Modelle, wobei ein kleinerer Wert für ein *besseres* Modell steht ([7], S. 69). Der Wert *Nagelkerke* R^2 ist vergleichbar mit dem R^2 der linearen Regression, welcher angibt, wieviel Prozent der Streuung durch das Modell erklärt werden kann (der hier niedrige Wert von 29,3% spricht dafür, dass *WAI* durch weitere Prädiktoren beeinflusst wird, die im vorliegenden Modell nicht berücksichtigt wurden). Es folgen p-Werte der Devianz- und Likelihood-Ratio-χ^2-Tests zur Überprüfung der „equal slope assumption“. Beide Tests sind im Falle der NEXT-Daten nicht signifikant (p=0,11 und p=0,11), so dass die Nullhypothese der „proportional odds“ beibehalten werden kann.

Als primäre Ergebnisse folgen die Schätzwerte der beiden Prädiktoren *ConfWorkFam* und *JobSatisf* mit Konfidenzintervall und Standardabweichung. Mit der *Wald*-Statistik (und den dazugehörigen p-Werten) wird geprüft, ob die Variable einen signifikanten Beitrag zum Regressionsmodell liefert. Der Einfluss beider Variablen ist signifikant. Die Koeffizienten sind in Form von logit-Wahrscheinlichkeiten angegeben, was sie schlecht interpretierbar macht. Man sieht allerdings an der Variablen *ConfWorkFam*, dass der Schätzwert negativ ist. Das bedeutet, dass ein Anstieg von *ConfWorkFam* eine Abnahme des *WAI* beeinflusst. Zur besseren Interpretation können die Schätzwerte durch Exponenzierung in *Odds Ratios* (OR) umgerechnet werden. Odds Ratios geben an, wie hoch die Chance ist, dass ein bestimmtes Ereignis (hier: Stufenwechsel im *WAI*) eintrifft ([2], S.442ff).

So erhält man für *ConfWorkFam* $OR = e^{\beta} = 0{,}562$ und für *JobSatisf* OR=3,913. Die Chance, in eine höhere Stufe des *WAI* zu gelangen, ändert sich um den Faktor 3,913, sobald der Wert von *JobSatisf* sich um 1 erhöht. Probanden haben also für jeden Anstieg der *JobSatisf* um 1 eine 4fach höhere Chance in eine höhere *WAI*-Stufe zu gelangen. Es bedeutet aber auch, dass bei einer *Abnahme* der Variablen *JobSatisf* um den Wert 1 die Chance zur Erreichung einer höheren *WAI*-Stufe 4fach *geringer* ist. Ebenso ist dann die Chance zu Erreichung einer *niedrigeren* *WAI*-Stufe 4fach höher.

Eine OR kleiner als 1 entspricht einem negativen Zusammenhang. Dies ist bei der Variabel *ConfWorkFam* der Fall: für jeden Anstieg in *ConfWorkFam* ändert sich die Chance

Tabelle 5: Deskriptive Auswertung der NEXT-Daten [mean (SD)]

WAI	schlecht	mäßig	gut	sehr gut	Gesamt
n	52	151	156	56	415
ConfWorkFam	3,65 (1,17)	3,16 (1,01)	2,54 (0,98)	2,15 (0,86)	2,85 (1,10)
JobSatisf	1,96 (0,44)	2,41 (0,50)	2,69 (0,48)	2,80 (0,44)	2,51 (0,54)

Tabelle 6: Proportional Odds Modell der NEXT-Daten

logLike	Deviance	AIC	Nagelkerke R^2	χ^2 (Dev)	χ^2 (likR)
–460,27	920,5	930,55	0,293	p=0,11	p=0,11
–456,46	912,9				

Coefficient	Estimate (Conf.Int.)	SD	Wald	p-Wert	Sig.
ConfWorkFam	–0,576 (–0,766 – –0,386)	0,097	–5,940	2,86e-09	***
JobSatisf	1,364 (0,968 – 1,760)	0,202	6,750	1,47e-11	***
schlecht | mäßig	0,621	0,648	0,958		
mäßig | gut	–1,770	0,654	–2,707		
gut | sehr gut	–4,058	0,676	–6,004		

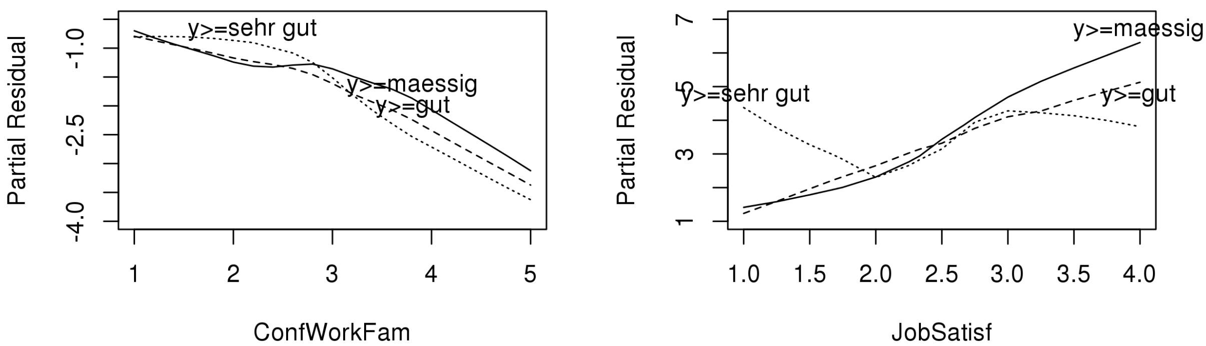

Abbildung 2: Partial Residual Plots des PO

zur Erreichung einer höheren *WAI*-Stufe um den Faktor 0,562 (und ist somit niedriger). OR kleiner als 1 sind schwer interpretierbar. Zur Erleichterung können die Werte der Variablen *ConfWorkFam* bei der Erstellung des Modells umgekehrt werden. Hierdurch ändert sich das Vorzeichen des Koeffizienten (aus –0,576 wird 0,576). Dies liefert eine *OR*=1,78. Bei der Interpretation der OR beschreibt man wegen der Umkehrung nicht mehr die Chance bei *Erhöhung* der Variablen *ConfWorkFam* um den Wert 1, sondern die bei *Abnahme* um den Wert 1. Das bedeutet, dass bei einer Abnahme des *ConfWorkFam* um den Wert 1 sich die Chance in eine höhere *WAI*-Stufe zu gelangen um den Faktor 1,78 erhöht. Entsprechend steigt die Chance eine *niedrigere WAI*-Stufe zu erreichen um den Faktor 1,78, wenn der Wert von *ConfWorkFam* sich um 1 erhöht.

Tabelle 6 schließt mit den Schwellenwerten der einzelnen Stufen des *WAI*. Diese Werte werden normalerweise nicht zur Interpretation herangezogen.

Als Alternative zu den χ^2-Tests kann mittels *Partialer Residual-Plots* überprüft werden, ob die Abweichungen der beobachteten Werte von der Modellvorhersage pro *WAI*-Stufe annähernd parallel verlaufen [9]. Abbildung 2 zeigt diese Plots für die Variablen *ConfWorkFam* und *JobSatisf*.

Das Plot der Variablen *ConfWorkFam* zeigt tendenziell gleiche Verläufe für jede Stufe von *WAI*. Dies ist ein Hinweis darauf, dass die Annahme eines universellen β-Koeffizienten für *ConfWorkFam* zutreffend ist.

Das Plot der Variablen *JobSatisf* zeigt annähernd gleiche Verläufe in den ersten 3 Stufen des *WAI*. Für die Stufe *„sehr gut“* zeigen sich Abweichungen bei (vor allem) niedrigen und hohen *JobSatisf*-Werten, was ein Hinweis auf Modellverzerrung sein kann.

Zum Vergleich wurden für jedes Stufenpaar des *WAI* logistische Regressionen gerechnet. Tabelle 7 zeigt die Ergebnisse im Vergleich zum PO. Es ist erkennbar, dass im PO kleinere Standardfehler sowie schärfere Konfidenzgrenzen erreicht werden.

Die Schätzwerte des PO liegen innerhalb der Konfidenzgrenzen der logistischen Regressionen.

Tabelle 7: Logistische Regressionen und PO der NEXT-Daten

logReg 1	–0,459	–0,805 – –0,130	0,171	2,077	1,392 – 2,820	0,363
logReg 2	–0,587	–0,814 – –0,368	0,113	1,344	0,876 – 1,835	0,244
logReg 3	–0,631	–0,974 – –0,310	0,169	0,835	0,189 – 1,520	0,338
PO	–0,576	–0,770 – –0,386	0,098	1,364	0,969 – 1,770	0,204

Continuation Ratio Model

Als zweite Gruppe stehen die *Continuation Ratio Modelle* (CR) zur Verfügung. Auch hier sind zahlreiche synonyme Bezeichnungen in der Literatur vertreten (vgl. [8], [9], [10]), wie z.B. *logistic continuation ratio model* [22], *proportional logit hazard model* [23] oder *sequentielle Mod*elle [7]. Diese Variante wird vor allem dann angewendet, wenn das Outcome *Stadien* repräsentiert, die ein Proband durchlaufen muss ([6], S. 338). Als Beispiel seien hier Universitätsabschlüsse genannt: Studenten beginnen *ohne Abschluss* und können über *Bachelor*-, *Master*- und *PhD*-Abschlüsse bis zum *Professor* aufsteigen. Bei Krankheiten beginnt ein Individuum mit dem Status *gesund* und durchläuft die Stadien *leicht krank, krank, schwer krank* bis *todkrank*.

Das Modell trifft Aussagen über die Wahrscheinlichkeit, in das nächste Stadium oder darüber hinaus zu gelangen. Das bedeutet, dass im Gegensatz zu PO-Modellen eine eindeutige Richtung vorgegeben ist, welche von der Fragestellung bestimmt wird. Nehmen wir als Beispiel die Daten des Gesundheitsstatus aus Tabelle 1. Wenn es Ziel der Untersuchung ist, Risikofaktoren herauszustellen, die mit einer Verschlechterung des Gesundheitsstatus einhergehen, wäre ein Vorwärts-Modell angezeigt. Tabelle 8 zeigt die einzelnen Stufenpaare dieses Verfahrens. Im Gegensatz zu PO-Modellen werden bei der Stufenpaarbildung die Daten von niedrigeren („gesünderen") Stadien verworfen.

Wenn es Ziel der Untersuchung ist, die Wirksamkeit verschiedener Therapiemethoden zu bestimmen, die mit einer Verbesserung des Gesundheitsstatus einhergehen, wäre ein Rückwärts-Modell angezeigt. Tabelle 9 zeigt die einzelnen Stufenpaare dieses Verfahrens. Hier werden entsprechend die Daten von höheren („kränkeren") Stadien bei der Paarbildung verworfen.

Die errechneten Koeffizienten von Vorwärts- und Rückwärts-Methode sind **nicht** equivalent, so dass sorgfältig geprüft werden muss, welche Richtung zur Beantwortung der Fragestellung sinnvoll ist.

CR-Modelle gehen ebenfalls von der „equal slope assumption" aus, d.h. auch hier wird angenommen, dass der Einfluss der Prädiktoren in den einzelnen Stadien derselbe ist. Diese Annahme muss auf Gültigkeit getestet werden, bevor das errechnete Modell angewendet werden kann. Hält die Annahme bei einigen Prädiktoren nicht stand, können alternativ *Partial Continuation Ratio*-Modelle (PCR) gerechnet werden [24]. Wie schon beim PO erhält man für solche Prädiktoren gesonderte Koeffizienten für jedes Stufenpaar.

CR-Modelle können (im Gegensatz zum PO) mit jeder Software erstellt werden, die logistische Regressionen berechnen kann. Armstrong und Sloan [5] geben eine Anleitung zur Restrukturierung des Datensatzes, die zuvor durchgeführt werden muss: für jedes Stufenpaar (vgl. Tabelle 9) wird ein neuer Dummy-Datensatz erstellt. Diese Datensätze werden um die Variablen *Ziel* und *Stufe* erweitert. Die Variable *Stufe* erhält bei allen verwendeten Fällen den Wert des aktuellen Stufenpaars. Beim Beispiel aus Tabelle 9 wäre dieser Wert beim ersten Stufenpaar =1. Der Wert der Variablen *Ziel* wird bei allen Fällen der „dunkelgrauen" Gruppe auf =1 gesetzt, und bei Fällen der „hellgrauen" Gruppe auf =0. Dieser Vorgang wird für jedes Stufenpaar wiederholt, so dass im Beispiel von Tabelle 9 vier Dummy-Datensätze vorliegen. Diese werden nun zu einem neuen Datensatz zusammengeführt. Anschließend fittet man mit dem neuen Datensatz eine logistische Regression auf die Variable *Ziel*. Die Ergebnisse entsprechen dem CR. Möchte man ein *Partiales CR* rechnen, muss bei der Modellierung die Variable *Stufe* als *Interaktion* zu den betroffenen Variablen hinzugefügt werden [9]. Dies liefert gesonderte Koeffizienten für jedes Stufenpaar. Eine solche Umkodierung ist für PO nicht möglich, da mit diesem Verfahren so genannte *singuläre Matrizen* erzeugt werden, die eine eindeutige Auflösung der Modellgleichungen mathematisch unmöglich machen.

Beispiel NEXT

Am Beispiel des NEXT-Datensatz wird nachfolgend ein CR berechnet. Als Methode wurde die *Rückwärts*-Variante gewählt, da in diesem Falle die *Verschlechterung* des *WAI* von Interesse sein soll. Tabelle 10 zeigt die Modellergebnisse.

Die χ^2-Tests zur Überprüfung der „equal slopes assumption" sind nicht signifikant (p=0,23). Das bedeutet, dass die Annahme der universellen Koeffizienten im vorliegenden CR beibehalten werden kann. Da in diesem Beispiel die *Rückwärts*-Variante berechnet wurde, sind die Vorzeichen der Prädiktoren im Vergleich zum PO (vgl. Tabelle 6) umgekehrt. Somit geht die Erhöhung der Variablen *ConfWorkFam* mit einer Verschlechterung des *WAI* einher. Durch Exponenzierung der Koeffizienten erhält man die Odds Ratio der Variablen. Die OR für *ConfWorkFam* beträgt 1,572. Somit steigt mit jeder Erhöhung der Variablen *ConfWorkFam* die Chance in eine niedrigere *WAI*-Stufe zu gelangen um den Faktor 1,572. Der Koeffizient der Variablen *JobSatisf* hat ein negatives Vorzeichen, was ein umgekehrt-proportionales Verhältnis anzeigt. Die OR für *JobSatisf* beträgt 0,285. Somit ändert sich mit jedem Anstieg der Variablen *JobSatisf* die Chance in eine niedrigere *WAI*-Stufe zu gelangen um den Faktor 0,285 (und

Tabelle 8: Continuation Ratio Modell (Vorwärts)

Stufe	**gesund**	**leicht krank**	**krank**	**schwerkrank**	**todkrank**	**TOTAL**
n	170	230	410	270	120	1.200
Paar 1	170	1030				1.200
Paar 2	–	230	800			1.030
Paar 3	–	–	410	390		800
Paar 4	–	–	–	270	120	390

Tabelle 9: Continuation Ratio Modell (Rückwärts)

Stufe	**gesund**	**leicht krank**	**krank**	**schwerkrank**	**todkrank**	**TOTAL**
n	170	230	410	270	120	1.200
Paar 1	1.080				120	1.200
Paar 2	810			270	–	1.080
Paar 3	400		410	–	–	810
Paar 4	170	230	–	–	–	400

Tabelle 10: Continuation Ratio Modell der NEXT-Daten

logLike	**Deviance**	**AIC**	**Nagelkerke R^2**	**χ^2 (Dev)**	**χ^2 (likR)**
–460,23	920,5	930,46	0,293	p=0,23	p=0,23
–457,40	914,8				

Coefficient	**Estimate (Conf.Int.)**	**SD**	**Wald**	**p-Wert**	**Sig.**
ConfWorkFam	0,452 (0,286 – 0,619)	0,085	5,329	9,88e-08	***
JobSatisf	–1,256 (–1,611 – –0,901)	0,181	–6,938	3,98e-12	***
sehr gut | schlechter	4,042	0,606	6,667		
gut | schlechter	2.088	0,581	3,593		
mäßig | schlechter	0,128	0,573	0,224		

Tabelle 11: Logistische Regressionen und CR der NEXT-Daten

	ConfWorkFam			JobSatisf		
	Est.	CI	SD	Est.	CI	SD
logReg 1	0,631	0,310 – 0,974	0,169	–0,835	–1,520 – –0,189	0,338
logReg 2	0,486	0,256 – 0,724	0,119	–1,260	–1,778 – –0,768	0,257
logReg 3	0,147	–0,209 – 0,512	0,183	–1,801	–2,636 – –1,036	0,406
CR	0,452	0,286 – 0,619	0,084	–1,256	–1,611 – –0,901	0,181

ist somit immer niedriger). Zur besseren Interpretation können die Werte der Variablen *JobSatisf* bei der Modellerstellung umgekehrt werden. Hierdurch ändert sich das Vorzeichen des Koeffizienten. So erhält man *OR*=3,512. Bei der Interpretation der OR beschreibt man wegen dieser Umkehrung nicht mehr die Chance bei *Erhöhung* der Variablen *JobSatisf* um den Wert 1, sondern die bei *Abnahme* um den Wert 1. Für Probanden bedeutet demnach eine Abnahme von *JobSatisf* um den Wert 1, dass sich ihre Chance zur Erreichung einer niedrigeren *WAI*-Stufe um das annähernd dreieinhalbfache erhöht.

Da beim CR die Richtung des Stufenwechsels vorgegeben ist, können (im Gegensatz zum PO) keine Aussagen über das Erreichen einer *höheren WAI*-Stufe getroffen werden. Zum Vergleich wurden für jedes Stufenpaar (vgl. Tabelle 9) des *WAI* logistische Regressionen gerechnet. Tabelle 11 zeigt die Ergebnisse im Vergleich zum CR. Es ist erkennbar, dass im CR kleinere Standardfehler sowie schärfere Konfidenzgrenzen erreicht werden.

Die Schätzwerte des CR liegen innerhalb der Konfidenzgrenzen der logistischen Regressionen.

Für das Stufenpaar *logReg* 3 umfasst das Konfidenzintervall der Variablen *ConfWorkFam* den Wert 0. Der Koeffizient ist somit in diesem Modell nicht gültig, da der wahre Wert sowohl positiv als auch negativ sein könnte.

Diskussion

Der vorliegende Artikel beschreibt die Verwendung ordinaler Regressionsmodelle bei der Auswertung ordinaler Zielvariablen. Am Beispiel des NEXT-Datensatzes konnten die Vorteile dieser Verfahren aufgezeigt werden. Zwar liefern auch die hier erstellten logistischen Regressionsmodelle akzeptable Ergebnisse, jedoch haben die Modellwerte der ordinalen Regression kleinere Standardfehler und schärfere Konfidenzgrenzen. Einschränkend muss gesagt werden, dass diese Werte besonders beim CR durch die Erzeugung der Dummy-Variablen (und der damit einhergehenden Vergrößerung des Samples) besser ausfallen als bei logistischen Regressionen.

Am Beispiel aus Tabelle 11 lässt sich zeigen, dass ordinale Modelle auch dann gültige Ergebnisse erzeugen können, wenn logistische Regressionen keine signifikanten Werte liefern.

Beim Einsatz ordinaler Regressionen müssen die Modellannahmen sorgfältig geprüft werden. Kann die „equal slopes assumption“ der Prädiktoren nicht verworfen werden, liefern PO und CR leicht interpretierbare Koeffizienten, die den Effekt der Prädiktoren für alle Stufen des ordinalen Ziels zusammenfassen. Da die Einflüsse in *universellen* Koeffizienten beschrieben werden, treten keine Effekte des multiplen Testens auf (Kumulation der α-Fehler).

Die Tests zur Prüfung der Modellannahmen gehen in der Nullhypothese von „equal slopes“ aus. Das Problem hierbei ist, dass statistische Tests grundsätzlich nie die Nullhypothese *bestätigen* können. Fallen die Modelltests nicht-signifikant aus, bedeutet dies lediglich, dass aus den *vorliegenden Daten* sowie mit der vorhanden Test-Power keine Verletzung der Annahme ersichtlich ist. Dennoch kann eine solche Verletzung theoretisch vorliegen.

Forscher müssen sich im Vorfeld darüber im Klaren sein, welches Modell geeignet ist, die Forschungsfrage zu beantworten. Das PO eignet sich, um allgemeine Aussagen über die Beeinflussung der Zielvariable zu treffen. Ist eine bestimmte *Richtung* von Interesse, oder repräsentiert die Zielvariable *Stadien*, die durchlaufen werden, eignet sich das CR zur Beschreibung gerichteter Effekte. Jedoch sind nur im PO die Koeffizienten (durch Vorzeichenwechsel) für den Wechsel in sowohl höhere als auch niedrigere Stufen gültig. Forscher müssen sich über diese Einschränkung im Klaren sein, wenn sie CR-Modelle interpretieren. Kann die „equal slopes assumption“ nur für einige Prädiktoren bestätigt werden, stehen erweiterte Modelle wie PPO und PCR zur Verfügung. Da diese jedoch separate Koeffizienten pro Stufenwechsel liefern, tritt – ebenso wie bei logistischen Regressionen pro Stufenwechsel – eine Kumulation der α-Fehler auf.

Schlussfolgerung

Die vorgestellten ordinalen Regressionsmodelle zeigen erweiterte, aber bisher nicht oft verwendete Möglichkeiten zur Auswertung ordinaler Daten. Seit einigen Jahren existieren für viele Statistikprogramme Funktionen zur Durchführung ordinaler Regressionen. Der Informationsgewinn, der sich durch ordinale Regressionsmodelle ergibt, sollte in zukünftigen Forschungsprojekten mehr berücksichtigt werden.

Anmerkungen

Danksagung

Wir bedanken uns bei Prof. Dr. Herbert Mayer und Dr. Jessica Hirsch für methodologische Rückmeldungen.

Förderung

Diese Publikation ist im Rahmen des vom BMBF geförderten Forschungskollegs „FamiLe“ entstanden (FKZ 01KX1113A).

Interessenkonflikte

Die Autoren erklären, dass sie keine Interessenkonflikte in Zusammenhang mit diesem Artikel haben.

Literatur

1. Sheather SJ. A Modern Approach to Regression with R. New York: Springer Science; 2009.
2. Hartung J. Statistik. Lehr- und Handbuch der angewandten Statistik. 15th ed. München: Oldenbourg Verlag; 2009.
3. McCullagh P. Regression models for ordinal data (with discussion). J R Stat Soc Series B Stat Methodol. 1980;42(2):109-42.
4. McCullagh P, Nelder JA. Generalized Linear Models. New York: Chapman and Hall; 1989.
5. Armstrong BG, Sloan M. Ordinal regression models for epidemiologic data. Am J Epidemiol. 1989 Jan;129(1):191-204.
6. Harrell FE. Regression Modeling Strategies, With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York: Springer; 2001.
7. Agresti A. Analysis of Ordinal Categorical Data. 2nd ed. New York: Wiley & Sons; 2010. DOI: 10.1002/9780470594001
8. Bender R. Introduction to the use of regression models in epidemiology. Methods Mol Biol. 2009;471:179-95. DOI: 10.1007/978-1-59745-416-2_9
9. Bender R, Benner A. Calculating Ordinal Regression Models in SAS and S-Plus. Biom J. 2000;42(6):677-99. DOI: 10.1002/1521-4036(200010)42:6<677::AID-BIMJ677>3.0.CO;2-O
10. Yee TW. The VGAM package for Categorical Data Analysis. J Stat Softw. 2010;32(10). Available from: http://www.jstatsoft.org/v32/i10

11. Scott SC, Goldberg MS, Mayo NE. Statistical assessment of ordinal outcomes in comparative studies. J Clin Epidemiol. 1997 Jan;50(1):45-55. DOI: 10.1016/S0895-4356(96)00312-5

12. Ananth CV, Kleinbaum DG. Regression models for ordinal responses: a review of methods and applications. Int J Epidemiol. 1997 Dec;26(6):1323-33. DOI: 10.1093/ije/26.6.1323

13. Bortz J, Döring N. Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. 4th ed. Heidelberg: Springer; 2006. DOI: 10.1007/978-3-540-33306-7

14. Peterson B, Harrell FE Jr. Partial proportional odds model for ordinal response variables. Appl Stat. 1990;39(2):205-17. DOI: 10.2307/2347760

15. Fox J, Hong J. Effect Displays in R for Multinominal and Proportional-Odds Logit Models: Extensions to the effects Package. J Stat Softw. 2009;32(1). Available from: http://www.jstatsoft.org/v32/i01

16. Hasselhorn HM, Müller BH, Tackenberg P, Kümmerling A, Simon M. Berufsausstieg bei Pflegepersonal. Berlin: Verlag für neue Wissenschaft; 2005.

17. Tuomi K, Ilmarinen J, Jahkola A, Katajarinne L, Tulkki A. Work Ability Index. Helsinki: Finnish Institute of Occupational Health; 1998.

18. Hasselhorn HM, Freude G. Der Work Ability Index – ein Leitfaden. Dresden: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin; 2007. (Schriftenreihe der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin: Sonderschrift; S 87). Available from: http://www.sopol.at/get_file.php?id=1248

19. Kristensen TS, Borg V, Hannerz H. Copsoq, Socioeconomic status and psychosocial work environment. Results from a national Danish study. Copenhagen: National Institute of Occupational Health; 2001.

20. Netemeyer R, Boles J, McMurrian R. Development and validation of workfamily conflict and familiy work conflict scales. J Appl Psychol. 1996;81(4):400-10. DOI: 10.1037/0021-9010.81.4.400

21. R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing; 2011.

22. Greenland S. Alternative models for ordinal logistic regression. Stat Med. 1994 Aug;13(16):1665-77. DOI: 10.1002/sim.4780131607

23. Cox C. Multinomial regression models based on continuation ratios. Stat Med. 1997;7(3):435-41. DOI: 10.1002/sim.4780070309

24. Harrell FEJ, Margolis PA, Gove S, Mason KE, Mulholland EK, Lehmann D, Muhe L, Gatchalian S, Eichenwald HF. Tutorial in biostatistics: Development of a clinical prediction model for an ordinal outcome: The World Health Organization Multicentre Study of Clinical Signs and Etiological Agents of Pneumonia, Sepsis and Meningitis in Young Infants. Stat Med. 1998;17(8):909-44. DOI: 10.1002/(SICI)1097-0258(19980430)17:8<909::AID-SIM753>3.0.CO;2-O

Korrespondenzadresse:

Jörg große Schlarmann
Department Pflegewissenschaft, Universität Witten/Herdecke, Stockumer Straße 12, 58453 Witten, Deutschland
schlarmann@uni-wh.de

Bitte zitieren als
Schlarmann JG, Galatsch M. Regressionsmodelle für ordinale Zielvariablen. GMS Med Inform Biom Epidemiol. 2014;10(1):Doc05. DOI: 10.3205/mibe000154, URN: urn:nbn:de:0183-mibe0001542

Artikel online frei zugänglich unter
http://www.egms.de/en/journals/mibe/2014-10/mibe000154.shtml

Veröffentlicht: *19.03.2014*