

Alexandra Grand
Regina Dittrich
Reinhold Hatzinger

Präferenzmodelle in der Praxis

Analyse von Paarvergleichen,
Likert Items und Rankings mit
R-prefmod

Alexandra Grand, Regina Dittrich, Reinhold Hatzinger

Präferenzmodelle in der Praxis

Analyse von Paarvergleichen, Likert Items
und Rankings mit **R**-prefmod

Bibliografische Information Der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Alle Angaben in diesem Buch sowie auf der Webseite zu diesem Buch erfolgen trotz sorgfältiger Bearbeitung ohne Gewähr, eine Haftung der Autoren oder des Verlages ist ausgeschlossen.
Die gewerbliche Nutzung der in diesem Buch angeführten Datensätze ist nicht zulässig.

© 2017 Facultas Verlags- und Buchhandels AG
facultas, 1050 Wien, Österreich

Alle Rechte, insbesondere das Recht der Vervielfältigung und der Verbreitung sowie der Übersetzung, sind vorbehalten.

Satz: Facultas Verlags- und Buchhandels AG
Einbandgestaltung: Atelier Reichert, Stuttgart
Druck und Bindung: CPI – Ebner & Spiegel, Ulm
Printed in Germany

utb-Titel-Nr.: 3785
ISBN: 978-3-8252-3785-1
ISBN: 978-3-8385-3785-6 (Online-Leserecht)

Vorwort

Dieses Buch entstand aus der Idee, eine kompakte und einfache Einführung in die Analyse von Präferenzdaten verschiedenster Art mit Hilfe der Methode der Paarvergleiche zu geben. In diesem Buch wird gezeigt, wie unterschiedliche Präferenz- bzw. Paarvergleichsmodelle in der statistischen Programmumgebung R – dank des von Reinhold Hatzinger programmierten Pakets `prefmod` – relativ rasch und unkompliziert berechnet werden können und wie deren Ergebnisse anhand von Präferenzskalen visualisiert und interpretiert werden können.

An Paarvergleichsmodellen wird bereits seit vielen Jahren geforscht. Bislang blieben die in internationalen Fachzeitschriften publizierten Arbeiten zu Paarvergleichsmodellen jedoch weitgehend Fachinteressierten vorbehalten. An einer breiten Anwendung mangelt es bis dato. Es war uns deshalb ein großes Anliegen die Methode der Paarvergleiche einer breiteren Leserschaft bekannt zu machen bzw. näher zu bringen, Interesse zu wecken und zum Ausprobieren in R zu animieren. Dieses Buch soll vor allem eine einfache Anleitung geben, wie erhobene Daten mit Hilfe von Paarvergleichsmodellen analysiert und interpretiert werden können.

Die ursprüngliche Idee eines kompakten, „kleinen“ Taschenbuchs wurde angesichts des großen Spektrums an Paarvergleichsmodellen und deren zahlreichen Erweiterungsmöglichkeiten rasch verworfen. Stattdessen wurde ein Buch gestaltet, das neben den Paarvergleichs-Basismodellen auch einige wichtige Erweiterungsmöglichkeiten umfasst. Gleichzeitig wurde darauf geachtet, dass das Buch für Studierende bzw. für all jene die an der Erstellung relativer Rangordnungen interessiert sind, verständlich bleibt und nachvollziehbar ist (auch ohne besondere statistische Kenntnisse oder Vorkenntnisse im Umgang mit der Statistiksoftware R). Ziel dieses Buches war es, neben einer kurzen theoretischen Einführung in die Grundlagen unterschiedlicher Paarvergleichsmodelle, einen anwendungsorientierten Schwerpunkt mit Erläuterung aller zur Berechnung notwendiger Schritte in R und der Interpretation der Ergebnisse zu setzen.

Inhaltsverzeichnis

Vorwort	i
1. Aufbau des Buchs	1
1.1. Typologische und symbolische Gestaltungselemente	3
1.2. Programmversion von R	5
2. Einleitung	7
3. Paarvergleiche	11
3.1. Arten von Paarvergleichen	13
3.2. Paarvergleichs-/Präferenzmodelle	14
4. Das Bradley-Terry Modell (BT)	15
4.1. Das log-lineare BT Modell (LLBT Modell)	18
I. Log-lineare Bradley-Terry Modelle (LLBT)	19

Echte Paarvergleiche

5. Das log-lineare Bradley-Terry Modell (LLBT)	21
5.1. Theorie	21
5.1.1. Designstruktur des LLBT Modells	23
5.1.2. Berechnung des LLBT Modells	24
5.1.3. Interpretation der Objektparameter	26
5.1.4. Berechnung der Werteparameter π	27
5.1.5. Prüfung der Modellgüte (Goodness-of-fit)	28
5.1.6. Erweiterungsmöglichkeiten des LLBT Modells	29
5.2. Anwendung in R – Beispiel 1: <i>Schokolade</i>	30
5.2.1. Beschreibung des Datensatzes, Datenvorbereitung	31
5.2.2. Deskriptive Statistik	34
5.2.3. Modellschätzung mit dem R-Paket <i>prefmod</i>	35

5.2.4.	Interpretation	39
5.2.5.	Modellgüte, Konfidenzintervalle	44
6.	Das LLBT Modell mit drei Antwortkategorien (<i>ties</i>)	49
6.1.	Theorie	49
6.1.1.	Designstruktur des LLBT Modells mit <i>ties</i>	51
6.2.	Anwendung in R – Beispiel 2: <i>CEMS</i>	52
6.2.1.	Beschreibung des Datensatzes, Datenvorbereitung . . .	52
6.2.2.	Deskriptive Statistik	55
6.2.3.	Modellschätzung mit dem R-Paket <i>prefmod</i>	56
6.2.4.	Modellgüte	58
6.2.5.	Interpretation	58
7.	Das LLBT Modell mit einer kategorialen Subjektkovariate	61
7.1.	Theorie	61
7.1.1.	Interpretation	63
7.1.2.	Modellselektion	64
7.2.	Anwendung in R – Beispiel 2: <i>CEMS</i>	66
7.2.1.	Beschreibung des Datensatzes, Datenvorbereitung . . .	66
7.2.2.	Deskriptive Statistik	67
7.2.3.	Modellschätzung mit dem R-Paket <i>prefmod</i>	70
7.2.4.	Modellselektion, Modellgüte	73
7.2.5.	Interpretation	75
8.	Das LLBT Modell mit zwei kategorialen Subjektkovariaten	81
8.1.	Theorie	81
8.1.1.	Berechnung der Objektparameter	82
8.1.2.	Modellselektion	83
8.2.	Anwendung in R – Beispiel 2: <i>CEMS</i>	85
8.2.1.	Modellschätzung	85
8.2.2.	Modellselektion	86
8.2.3.	Berechnung und Interpretation der Objekt- und Werte- parameter	90
9.	Eine metrische Subjektkovariate (ohne <i>ties</i>)	95
9.1.	Theorie	95
9.2.	Anwendung in R – Beispiel 3: <i>Lernmethoden</i>	96
9.2.1.	Beschreibung des Datensatzes, Datenvorbereitung . . .	96

9.2.2. Deskriptive Statistik	98
9.2.3. Modellschätzung mit dem R-Paket <i>prefmod</i>	101
9.2.4. Berechnung der Objektparameter	104
9.2.5. Berechnung und Darstellung der Werteparameter	106
10. Das LLBT Modell mit einer objekt-spezifischen Kovariate	111
10.1. Theorie	111
10.2. Anwendung in R – Beispiel 2: <i>CEMS</i>	114
10.2.1. Modellschätzung mit dem R-Paket <i>prefmod</i>	114
10.2.2. Objekt- und Werteparameter	119
10.2.3. Modellselektion	121
II. Log-lineare Paarvergleichs Pattern Modelle	123
11. Paarvergleichs Pattern Modelle	125
11.1. Theorie	125
Echte Paarvergleiche	
12. Das Pattern Modell bei unabhängigen Entscheidungen	133
12.1. Theorie	133
12.1.1. Designstruktur des Pattern Modells	133
12.1.2. Interpretation der Objektparameter	134
12.1.3. Berechnung der Werteparameter	135
12.1.4. Der Unterschied zwischen LLBT und Pattern Modell	135
12.2. Anwendung in R – Beispiel 1: <i>Schokolade</i>	136
12.2.1. Modellschätzung mit dem R-Paket <i>prefmod</i>	136
12.2.2. Berechnung der Objekt- und Werteparameter	140
12.2.3. Grafische Darstellung der Objekt- und Werteparameter	141
13. Das Pattern Modell mit Abhängigkeiten in den Entscheidungen	143
13.1. Theorie	143
13.1.1. Designstruktur des Pattern Modells mit Abhängigkeiten in den Entscheidungen	146
13.1.2. Interpretation der Abhängigkeits- und Objektparameter	147

14. Das Pattern Modell mit Abhängigkeiten und einer kategorialen Subjektkovariate	149
14.1. Theorie	149
14.2. Anwendung in R – Beispiel 4: <i>Interview mit Sportlerinnen</i> . . .	150
14.2.1. Beschreibung des Datensatzes	150
14.2.2. Modellschätzung mit dem R-Paket <i>prefmod</i>	151
14.2.3. Objektparameter	154
14.2.4. Werteparameter	156
15. Das Pattern Modell mit Abhängigkeiten und ties	159
15.1. Theorie	159
15.2. Anwendung in R – Beispiel 2: <i>CEMS</i>	161
15.2.1. Beschreibung des Datensatzes, Datenvorbereitung . . .	161
15.2.2. Modellschätzung mit dem R-Paket <i>prefmod</i>	162
15.2.3. Grafische Darstellung der Objekt- und Werteparameter	166
Hergeleitete Paarvergleiche	
16. Das Pattern Modell für Rangdaten	169
16.1. Theorie	169
16.1.1. Designstruktur des Pattern Modells für Rangdaten . . .	172
16.2. Anwendung in R – Beispiel 5: <i>Restaurant</i>	173
16.2.1. Beschreibung des Datensatzes, Datenvorbereitung . . .	173
16.2.2. Modellschätzung mit dem R-Paket <i>prefmod</i>	175
16.2.3. Modellgüte	179
16.2.4. Erweiterungsmöglichkeiten	180
17. Das Pattern Modell für Ratingdaten	183
17.1. Theorie	183
17.1.1. Designstruktur des Pattern Modells für Ratingdaten mit Berücksichtigung von <i>ties</i>	187
17.2. Anwendung in R – Beispiel 6: <i>Umweltgefahr</i>	190
17.2.1. Beschreibung des Datensatzes, Datenvorbereitung . . .	190
17.2.2. Modellschätzung mit dem R-Paket <i>prefmod</i>	191
17.2.3. Ausgabe, Visualisierung der Objekt- und Werteparameter	195

Appendix	197
A. R-Grundlagen	199
A.1. Einlesen externer Dateien in R	199
A.2. Installation der benötigten R-Pakete	200
B. Zusammenfassung wichtiger Symbole	203
Literaturverzeichnis	205
Index	209

1. Aufbau des Buchs

Dieses Buch ermöglicht Ihnen einen Streifzug durch Paarvergleichsmodelle und zeigt deren Anwendung mit Hilfe der frei verfügbaren Software R-prefmod. Wir haben das Buch nach der Devise „*learning by doing*“ aufgebaut, mit dem Ziel Sie zum „Mitmachen“ zu animieren.

Das Buch gewährt zu Beginn einen kurzen Einblick in Erhebungsmöglichkeiten von Präferenzen. Kapitel 3 setzt gleich anschließend mit einer einführenden Beschreibung der Methode der Paarvergleiche, anhand derer Präferenzdaten verschiedenster Art analysiert werden können, fort. Die Analyse von Paarvergleichsdaten erfolgt auf Basis des von Bradley & Terry (1952) definierten Paarvergleichmodells (abgekürzt mit BT Modell), das in Kapitel 4 vorgestellt wird.

Der Kern dieses Buches besteht aus zwei Teilen. Im Teil I dieses Buches beschäftigen wir uns mit log-linearen Bradley-Terry Modellen (abgekürzt mit LLBT) und im Teil II liegt der Schwerpunkt auf log-linearen Paarvergleichs-Pattern Modellen (in Kurzform auch Pattern Modelle genannt). Jeder der beiden Teilbereiche beginnt mit einem Kapitel, in dem das jeweilige Basismodell (also das LLBT Basismodell und das Pattern Basismodell) zunächst theoretisch beschrieben und dann angewandt wird. Die jeweils nachfolgenden Kapitel beziehen sich auf Erweiterungsmöglichkeiten der Basismodelle. Die Gliederung der Kapitel erfolgt nach einem einheitlichen Schema: Zuerst werden theoretische Grundlagen erläutert und das jeweilige Modell formal definiert. Wir haben darauf geachtet, dass sich der Theorieteil auf das zum Verständnis Wesentliche beschränkt. Anschließend folgt ein (reales) Anwendungsbeispiel. Die Berechnung log-linearer BT Modelle kann auf verschiedene Weise in R erfolgen. Zwei Möglichkeiten werden in diesem Buch angeführt – eine „benutzerfreundliche“ Variante, die mit weniger R-Code (Befehle die in R eingegeben werden) auskommt und eine etwas „aufwändigere“ Variante, die allerdings mehr individuelle Gestaltungsmöglichkeiten zulässt. An dieser Stelle sollen Sie animiert werden die in diesem Buch beschriebenen Auswertungsschritte nicht nur nachzuvollziehen, sondern auch direkt in R „auszuprobieren“. Die Software R kann direkt von <https://cran.r-project.org/> heruntergela-

1. Aufbau des Buchs

den und für verschiedene Betriebssysteme installiert werden. Der im Buch angeführte R-Code ist in farbig serifizierter **Schrift** geschrieben und kann 1 : 1 in R eingegeben werden. Folgen Sie einfach den einzelnen Schritten und Sie werden feststellen, dass die Auswertung relativ rasch und einfach durchgeführt werden kann. Die berechneten Modellparameter werden angeführt, passende Grafiken erstellt und die Ergebnisse sowohl statistisch als auch inhaltlich interpretiert.

Die kapitelweise Trennung von Theorie und Anwendung ermöglicht dem/der Leser/in bei Bedarf einfach und rasch entweder im Theorie- oder Anwendungsteil des jeweiligen Modells nachzuschlagen.

Im Anhang befinden sich zusätzliche Erläuterungen zur Installation der benötigten R-Pakete und zum Einlesen von Datensätzen in R sowie ein Verzeichnis der wichtigsten Symbole, die in diesem Buch verwendet wurden. Auf der Webseite zu diesem Buch werden neben den Datensätzen und den R-Programmcodes auch Zusatzinformationen für Interessierte zu einzelnen Abschnitten bereitgestellt.

Welche Zielgruppe möchten wir ansprechen?

Das Buch richtet sich in erster Linie an Studierende, beispielsweise der Sozial- und Wirtschaftswissenschaften oder der Psychologie und an all jene die an quantitativer empirischer Forschung interessiert sind.

Unser Ziel war es das Buch so zu gestalten, dass alle zum Verständnis notwendigen theoretischen Grundlagen, Schritte und R-Befehle möglichst einfach und nachvollziehbar dargestellt werden. Über die Grundlagen hinausgehende statistische, mathematische Kenntnisse oder softwaretechnische Kenntnisse in R, sind daher nicht unbedingt Voraussetzung, aber wenn vorhanden natürlich von Vorteil. Das Buch bietet eine Einführung für Interessierte und zeigt Schritt für Schritt praxisorientiert Abläufe von Präferenzanalysen – beginnend bei einem realen Beispieldatensatz auf den ein Paarvergleichsmodell angewandt und mit R-prefmod ausgewertet wird bis hin zur grafischen Darstellung und Interpretation der Ergebnisse. Es eignet sich auch als Nachschlagewerk und könnte zum Selbststudium und als Inspiration für die Realisierung eigener Forschungsprojekte dienen.

1.1. Typologische und symbolische Gestaltungselemente

Den Text dieses Buches haben wir mit verschiedenen Schrifttypen, Hervorhebungen und Symbolen versehen, um diesen übersichtlich und leicht lesbar zu gestalten.

Verwendete Schrifttypen

1. Serifenlose Schrift:

Als Grundschrift wurde eine serifenlose Schrift gewählt.

2. Serifierte Schrift (Schreibmaschinenschrift):

Dieser Schrifttyp wird eingesetzt für

- Variablennamen (z.B. Geschlecht),
- Funktionen, die mit Klammern geschrieben werden (z.B. `plot()`),
- Optionsnamen bzw. Argumente von Funktionen (z.B. `nitems=3`),
- Dateinamen (z.B. `paarvergleiche.RData`).

Funktionen (die zum ersten Mal vorkommen) werden mit ihren wichtigsten Argumenten in schwarzer serifizierter Schrift in folgender Weise gekennzeichnet:

► **Funktion:**

```
llbtPC.fit(data, nitems = ...)
```

Hervorhebungen

- Schriftstil: *kursiv* um Begriffe hervorzuheben
- Schriftfarbe: **farbig** für wichtige Begriffe und besondere Hervorhebungen

Weitere Gestaltungsweisen und Symbole

- Beispiele werden in folgender Form dargestellt:

BEISPIEL _____
Wie lautet die Berechnung für ...

2. Einleitung

Das Konzept der Präferenzen beruht darauf, Wahlmöglichkeiten zwischen verschiedenen Alternativen zu haben und diese hinsichtlich eines oder mehrerer Kriterien zu beurteilen (Hatzinger, Dittrich & Salzberger, 2009). Präferenzen sind subjektiv und nicht direkt beobachtbar (d.h. latent). Doch wie können Präferenzen erfasst und schließlich messbar gemacht werden?

BEISPIEL

Nehmen wir als Beispiel Musikpräferenzen her. Eine einfache Möglichkeit Musikpräferenzen zu erheben wäre es Personen direkt zu fragen welcher von beispielsweise drei vorgespielten Songs ihnen am besten gefällt. Mit Hilfe der Antworten (die beobachtbar sind) können dann Rückschlüsse auf die Musikpräferenzen gezogen werden.

Zur Erhebung von Präferenzen gibt es verschiedene Ansätze (Kotler & Bliemel, 2001):

- Präferenzen können anhand von **Rangordnungen** (Rankings) erhoben werden. Beispielsweise werden Befragte gebeten Fruchtjogurts verschiedener Hersteller nach ihren Präferenzen zu ordnen. Der Aussagegehalt solcher Präferenzordnungen ist begrenzt, denn es kann nicht unmittelbar abgelesen werden, wie groß die Präferenzunterschiede zwischen den Fruchtjogurts sind. Dies liegt daran, dass Rangbeurteilungen z.B. das beste Fruchtjogurt, das zweitbeste Jogurt nichts über den Abstand zwischen den beiden Fruchtjogurts aussagen (siehe auch Kapitel 16.1).
- Eine weitere Möglichkeit wäre eine Präferenzenerhebung anhand **direkter Bewertungen** einzelner Objekte, indem zum Beispiel mit Hilfe eines vierstufigen Ratingformats, das von *sehr wichtig* bis *nicht wichtig* reicht, verschiedene Serviceleistungen eines Wellnesshotels bewertet werden. Bei dieser Methode besteht ein Problem darin, dass den Bewertungen Zahlen zugeordnet werden (z.B. *sehr wichtig* = 4, ..., *nicht wichtig* = 1).

2. Einleitung

Oft werden diese Zahlen dann für statistische Analysen als (direkte) Messungen verwendet und erhalten Bedeutungen die sie nicht haben. Zum Beispiel werden die Zahlen als metrisch interpretiert, d.h. dass die Differenz zwischen 1 und 2 dasselbe bedeutet wie zwischen 3 und 4 usw. (siehe auch Kapitel 17.1 oder Dittrich et al., 2007).

- Die [Methode der Paarvergleiche](#) (Exkurs siehe Kasten auf Seite 12) ist eine weitere Methode zur Erhebung von Präferenzen um, auf Basis subjektiver Präferenzen, Rangordnungen erstellen und darüberhinaus Abstände zwischen Objekten quantifizieren (d.h. in numerischer Form messbar machen) zu können.

Anhand von Paarvergleichen können relative Vergleiche jeweils zweier Objekte, Items usw. durchgeführt werden. Paarweise Vergleiche erleichtern bzw. vereinfachen das Urteil der befragten Personen, besonders in Situationen in denen es schwierig ist Bewertungen vorzunehmen. Die Befragten müssen lediglich entscheiden welches von beiden Objekten präferiert (bevorzugt) wird oder einen Aspekt oder eine Eigenschaft stärker widerspiegelt (oder ob sie unentschieden sind).

Ein Beispiel wäre die Beurteilung von Urlaubsdestinationen hinsichtlich des subjektiv empfundenen Sicherheitsrisikos. Das relative Sicherheitsrisiko von beispielsweise drei Destinationen (bezeichnet mit A, B und C) könnte einfach erhoben werden, indem Destinationspaare gebildet werden, wie zum Beispiel Destination A vs. B, C vs. A und B vs. C. Die Befragten sollen dann in jedem der insgesamt drei Vergleiche entscheiden, welche der beiden präsentierten Destinationen sie als *sicherer* empfinden (im Sinne von beispielsweise Destination B ist *sicherer* als A, C ist *sicherer* als A und C ist *sicherer* als B).

Die Methode der Paarvergleiche hat einen breiten Anwendungsbereich. Beispielsweise wird diese Methode in der Ästhetikforschung (z.B. McManus, 1980), bei Produkt-/Dienstleistungsvergleichen (z.B. Hatzinger & Mazanec, 2007) oder Produktverkostungen in denen es um Geschmackspräferenzen geht (z.B. Gridgeman, 1955), bei der Wahlforschung zur Erhebung von Parteipräferenzen (z.B. Zeileis et al., 2012), in der Medizin zur Messung von Schmerzen (z.B. Matthews & Morris, 1995), der Bewertung von wissenschaftlichen Journalen (z.B. Vana et al., 2016) oder auch zur Erforschung von Geräuschen (z.B. Ellermeier et al., 2004)

herangezogen.

Anwendungen für diverse wissenschaftliche Fragestellungen aus dem Bereich Marketing, Behavioural Finance und Human Resource Management können im Buch „Präferenzanalyse mit R“ (Hatzinger et al., 2009) nachgelesen werden.