



Das log-lineare Bradley-Terry Model

Eine Messung von wechselseitigen Erwartungen zwischen Universitätsbediensteten

Daniela Weber, Maria Weiler



Überblick

- 1 Grundidee
 - Ordinalität und Unabhängige Variablen
- 2 Datenmodellierung
- 3 Paarvergleichsmodell
 - Bradley-Terry Model
 - log-lineare BT Model
 - Model für ordinale Daten
 - Bradley-Terry Model für ordinale Daten
 - Erweiterungen
- 4 Anwendung an Datensatz

Im folgenden wird beschrieben und gezeigt, wie man anhand von Paarvergleichsdaten mit dem LLBT (loglinearen Bradley-Terry Modell) analysieren kann. Die hier verwendeten Daten stammen aus einer Erhebung an der Wirtschaftsuniversität Wien, wobei

- 31 Professoren
- 96 wissenschaftliche Mitarbeiter

befragt wurden, was sie am je anderen am meisten schätzen. Neben dieser Selbsteinschätzung wurden die Versuchspersonen auch gebeten, anzukreuzen, was an ihnen am meisten geschätzt wird.

Daten 2

In der Untersuchung wurden 8 Items vollständig in Vergleich gestellt.

- 1 Fachliche Kompetenz
- 2 Ehrlichkeit/Vertrauenswürdigkeit
- 3 Mitarbeiter: Kommunikationsfähigkeit
Dienstvorgesetzte: Kreativität
- 4 Mitarbeiter: Gerechtigkeit
Dienstvorgesetzte: Genauigkeit/Gewissenhaftigkeit
- 5 Verlässlichkeit
- 6 Wissenschafts- und Forschungsinteresse
- 7 Inspiration/Motivation
- 8 Kooperationsfähigkeit

Ordinalität und subjektspezifische Variablen

Die Originaldaten wurden auf einer ordinalen 5-stufigen Skala beurteilt, die von „viel wichtiger“ und „eher wichtiger“ über „unentschieden“ zu „eher“ und „viel wichtiger“ für das nicht bevorzugte Item reichte.

Als unabhängige Variablen wurden die Institutsgröße durch die Anzahl der Mitarbeiter und das Dienstalter erhoben und in Kategorien eingeteilt.

Variable	Kategorie
Institutsgröße 1	1 bis 6 Mitarbeiter
Institutsgröße 2	über 6 Mitarbeiter
Dienstalter 1	bis 5 Jahre
Dienstalter 2	5 bis 12 Jahre
Dienstalter 3	über 12 Jahre

Itemkontraktion

- zu **Wissensfaktoren** zählen fachliche Kompetenz und Wissenschafts- und Forschungsinteresse
- zu **Charakterfaktoren** zählen Ehrlichkeit/Vertrauenswürdigkeit, Gerechtigkeit, Verlässlichkeit und Genauigkeit/Gewissenhaftigkeit
- zu **Begabungsfaktoren** zählen Inspiration/Motivation, Kooperationsfähigkeit, Kommunikationsfähigkeit und Kreativität

Daten

Der Datensatz data3.csv ist wie folgt aufgebaut:

- Spalten 1:2 subjektspezifische Variablen der wissenschaftlichen Mitarbeiter
- Spalten 3:5 Vergleiche der wissenschaftlichen Mitarbeiter unter Selbsteinschätzung
- Spalten 6:8 Vergleiche der wissenschaftlichen Mitarbeiter unter Fremdeinschätzung
- Spalten 9:10 subjektspezifische Variablen der Dienstvorgesetzten
- Spalten 11:13 Vergleiche der Dienstvorgesetzten unter Selbsteinschätzung
- Spalten 14:16 Vergleiche der Dienstvorgesetzten unter Fremdeinschätzung
- Zeile 1: Spaltenüberschriften

Paarvergleich

Grundidee des Skalierungsverfahrens

- Vergleich von Objekten
- Bestimmung der Position der Objekte auf Präferenzskala
- Untersuchung eventueller Abhängigkeit der Objektbevorzugung weiterer Variablen

Antwortmöglichkeiten bei Paarvergleichen:

- binäre Antworten
- unentschieden Kategorie (ordinale Daten)
- gerankte Daten

Bradley-Terry Model

Im einfachen Fall Entscheidung zwischen zwei Objekten
Basismodel BT Model:

$$\Pi_{(jk)j} = \frac{\pi_j}{\pi_j + \pi_k}$$

$\Pi_{(ik)j}$... Wahrscheinlichkeit, dass Objekt j bevorzugt
 π_j und π_k ... nicht-negative Präferenzparameter (worth
parameter)

Kontingenztafel

Output darstellbar durch zweidimensionale Kontingenztafel
wobei

$y_{(jk)k}$... Anzahl der Präferenzen für k (Poisson-Verteilung)

n_{jk} ... Anzahl der Vergleiche

Vergleich	Entscheidung			Anzahl an Vergleichen
	Objekt 2	Objekt 2	Objekt 3	
(12)	$y_{(12)1}$	$y_{(12)2}$	–	n_{12}
(13)	$y_{(13)1}$	–	$y_{(13)3}$	n_{13}
(23)	–	$y_{(23)3}$	$y_{(23)3}$	n_{23}

log-lineare BT Model (LLBT)

Basis Model als log-lineares Model

$$\ln m_{(jk)k} = \mu_{(jk)k} + \lambda_j^O - \lambda_k^O$$

$\mu_{(jk)k}$... Störparameter (zB Wechselwirkungen)

λ_j^O ... Objektparameter

Zusammenhang zwischen Objektparameter und Präferenzparameter:

$$\Pi_{(jk)j} = \frac{\sqrt{\frac{\pi_j}{\pi_k}}}{\sqrt{\frac{\pi_k}{\pi_j}} + \sqrt{\frac{\pi_j}{\pi_k}}}$$

$$\frac{1}{2} \ln \pi_j = \lambda_j$$

Entscheidungsstruktur bei drei Objekten

Für drei Objekte erhält man folgende Designmatrix:

Vergleich	Entscheidung	Anzahl	μ	O_i	O_j	O_k
(ij)	O_i	$y_{(ij)i}$	1	1	-1	0
(ij)	O_j	$y_{(ij)j}$	1	-1	1	0
(ik)	O_i	$y_{(ik)i}$	2	1	0	-1
(ik)	O_k	$y_{(ik)k}$	2	-1	0	1
(jk)	O_j	$y_{(jk)j}$	3	0	1	-1
(jk)	O_k	$y_{(jk)k}$	3	0	-1	1

ordinale Daten

Paarvergleichsmodel sehr aufwendig, daher Möglichkeit mehr Information zu sammeln

Nicht nur binäre Antworten, sondern ordinale Antwortskala

- Unentschiedenkategorie möglich
- Grad an Präferenz wird angegeben

Bradley-Terry Model für ordinale Daten

allgemeine Bradley-Terry Model für ordinale Daten:

$$\Pi_{(ij)h} = a_{ij} c_h \left(\frac{\pi_j}{\pi_i + \pi_j} \right)^h \left(\frac{\pi_i}{\pi_i + \pi_j} \right)^{H-h}, h = 0, 1, \dots, H$$

c_h ... Kategorieparameter (Verzerrung)

a_{ij} ... Normalisierungskonstante

log-lineare BT für ordinale Daten (1)

log-lineare Bradley-Terry Model für ordinale Daten

$$\Pi_{(ij)h} = a_{ij}^* c_h \left(\frac{\sqrt{\pi_i}}{\sqrt{\pi_j}} \right)^{H-2h}$$

$$a_{ij}^* = a_{ij} / [\sqrt{\pi_i/\pi_j} + \sqrt{\pi_j/\pi_i}]^H$$

des weiteren sei $y_{(ij)h}$... Anzahl der Antworten der Kategorie h

y_{ij+} ... Anzahl aller Vergleiche der Objekte i und j

$y_{(ij)h}$ sind multinominal verteilt

log-lineare BT für ordinale Daten (2)

als log-lineares BT für ordinale Daten erhalten wir analog:

$$m_{(ij)h} = y_{ij} a_{ij}^* c_h \left(\frac{\sqrt{\pi_i}}{\sqrt{\pi_j}} \right)^{H-2h}$$

$$\ln m_{(ij)h} = \mu_{ij} + \gamma_h + (H - 2h)\gamma_i - (H - 2h)\gamma_j$$

wobei $(1/2) \ln \pi_i = \gamma_i$

$\gamma_h = \ln c_h$

$\mu_{ij} = \ln y_{ij} + \ln a_{ij}^*$

Erweiterungen

Es besteht Möglichkeit Bradley-Terry Model zu erweitern durch:

- Reihungseffekte
- Indifferenzentscheidungen
- Subjektspezifische Kovariate
- Objektspezifische Kovariate

subjektspezifische Kovariate

$$\ln m_{ijhl} = \mu_{ijl} + \lambda_l^S + \gamma_h + (H - 2h)(\gamma_i^O + \gamma_{il}^{OS}) - (H - 2h)(\gamma_j^O + \gamma_{jl}^{OS})$$

where m_{ijhl} ... erwartete Anzahl an Antworten der Kategorie h ,
wobei Probanden der Kovariatenklasse j

λ_l^S ... Haupteffekte der Subjekte

γ_{il}^{OS} ... Wechselwirkungsterm zwischen Objekt j und
subjektspezifischen Kovariaten

da log-lineares Model fitten als GLM möglich

Datensatz

siehe Kalkulationen in R