

Ein Stereotype Regressionsmodell für korrelierte Beobachtungen

Oliver Kuß*; Uwe Hasenbein**;

*Institut für Medizinische Epidemiologie,
Biometrie und Informatik,
Martin-Luther-Universität Halle-Wittenberg, Halle(Saale)

** Institut für Neurologisch-Neurochirurgische
Rehabilitationsforschung (INNRF), Magdeburg

GMDS 2003, Münster, 15.9.2003

Programm

- Motivation
- Probleme
- (Partielle) Lösungen
- Das Modell
- Resultate
- Fazit

Motivation

36 Ärzte wurden um ihre Einschätzung zur optimalen rehabilitativen Versorgung von Patienten mit Schädel-Hirn-Trauma (SHT) gebeten.

Jedem wurden 10 Musterkrankengeschichten vorgelegt und sie sollten das optimale Setting für den jeweiligen Patienten (stationäre, teilstationäre oder ambulante Rehabilitation) auswählen.

Welche Faktoren (sowohl die Ärzte als auch die Patienten betreffend) beeinflussen die Setting-Präferenzen?

Probleme

1. Einschätzungen innerhalb eines Arztes werden korreliert sein.
2. Ist die Zielgröße ordinal oder nominal?
3. Original-Ton U. Hasenbein: „Liegt das teilstationäre Setting näher beim stationären oder beim ambulanten Setting?“

(Teilweise) Lösungen

Lösung Problem 1:

→ Gemischte Modelle

Lösung Probleme 1+2:

→ Gemischte Modelle für ordinale und nominale Zielgrößen

(Tutz/Hennevogl, 1996; Hartzel/Caffo/Agresti, 2001)

Lösung Probleme 2+3

→ Stereotype Regression (Anderson, 1984)

Aber alle drei gleichzeitig???

Das Modell (Lösung Problem 1+2+3)

Fasse Stereotype Modell als nicht-lineares Regressionsmodell auf und erweitere zu nicht-linearem Modell mit gemischten Effekten

Konkret: Gegeben $i = 1, \dots, N$ Cluster (Ärzte), jedes mit T_i Beobachtungen. Y_{ij} sei der Wert der Zielgröße der j -ten Beobachtung in Cluster i , $j = 1, \dots, T_i$ mit Ereigniswahrscheinlichkeit $\pi_{ijr} = P(Y_{ij} = r)$, $r = 1, \dots, R$. x_{ij} ist ein Spaltenvektor von (festen) Kovariablen.

Modellgleichung:

$$\log \left(\frac{\pi_{ijr}}{\pi_{ijR}} \right) = \theta_r + \phi_r \beta x'_{ij} + u_{ir}, \quad r = 1, \dots, R - 1,$$

mit Restriktionen $\theta_R = 0$, $\beta_R = 0$, $\phi_1 = 0$ und $\phi_R = 1$.

Annahme

Die zufälligen Effekte sind multivariat normalverteilt mit unstrukturierter Kovarianzmatrix Σ , d.h. für $u'_i = (u_{i1}, \dots, u_{i,R-1})$ gilt

$$u_i \sim N((0, \dots, 0)', \Sigma) \text{ mit}$$

$$\Sigma = \begin{pmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \cdots & \sigma_{1,R-1}^2 \\ & \sigma_{22}^2 & \cdots & \sigma_{2,R-1}^2 \\ & & \cdots & \vdots \\ & & & \sigma_{R-1,R-1}^2 \end{pmatrix}$$

Eigenschaften

- Herleitung aus dem multinomialen gemischten Modell von Hartzel et al. 2001, durch die Parameter-Restriktion $\beta_r = \phi_r \beta$
- Einführung der zufälligen Effekte in die Modellgleichung modelliert gleichzeitig Heterogenität zwischen und Korrelation innerhalb der Cluster
- Gegeben die zufälligen Effekte sind die Beobachtungen innerhalb eines Clusters unabhängig
- Zu maximierende Likelihoodfunktion ist ein Produkt von N nicht geschlossen lösbaren Integralen

Schätzmethode

Adaptive Gauss-Hermite Quadrature: Numerische Integration, SAS PROC NLMIXED, liefert 'exakte' ML-Schätzer

Weitere: (teilweise in Arbeit) NPMLE, MCMC, Posterior Mean Estimation, Marginale Modelle

Resultate

Zielgröße: Optimales Reha-Setting (stationär [Referenz, 61%], teilstationär [TS, 19%], ambulant [AB, 20%])

Kovariablen: alle binär

- Ist der Arzt ein Neurologe? (NEURO, 63%)
- Ist der Arzt eine Facharzt? (FACHARZT, 49%)
- Liegt das SHT mehr als 3 Monate zurück? (ZEIT, 51%)
- Ist der Patient schwer behindert? (SCHWERE, 83%)

	RE Multinomiales Modell		RE Stereotype Modell
	Feste Effekte		
	TS	AB	
$\hat{\beta}_{NEURO}$	-0.56 (0.74)	-0.36 (0.90)	1.36 (0.93)
$\hat{\beta}_{FACHARZT}$	-0.63 (0.78)	-0.04 (0.94)	0.40 (0.88)
$\hat{\beta}_{ZEIT}$	2.51 (0.41)	3.43 (0.50)	4.23 (0.58)
$\hat{\beta}_{SCHWERE}$	-1.94 (0.43)	-3.29 (0.47)	-2.60 (0.53)
$\hat{\phi}_2$	–	–	0.55 (0.09)
	Zufällige Effekte		
$\hat{\sigma}_1^2$	1.47 (0.35)	–	2.01 (0.96)
$\hat{\sigma}_2^2$	–	1.87 (0.43)	2.62 (1.20)
$\hat{\sigma}_{12}^2$	2.54 (1.11)		1.94 (0.93)
	Modellwahlkriterien		
AIC	499.3		487.5
BIC	516.7		503.3

SAS PROC NLMIXED

```
proc nlmixed data=sht;
  parms u2=-1.5 u3=-3
        phi2=0.5
        b_neurologe=1.5 ...
        sd2=2 sd3=2 cov23=2;

  eta = b_neurologe*neurologe + b_facharzt*facharzt +
        b_zeit*zeit + b_schwere*schwere;

  p1=          1 / (1+exp(u2 + phi2*eta)+exp(u3 + eta));
  p2= exp(u2 + phi2*eta) / (1+exp(u2 + phi2*eta)+exp(u3 + eta));
  p3= exp(u3 +          eta) / (1+exp(u2 + phi2*eta)+exp(u3 + eta));

  if resp1 then loglike=log(p1);  else
  if resp2 then loglike=log(p2);  else
  loglike=log(p3);
  model setting ~ general(loglike);
  random u2 u3 ~ normal([a2,a3],[sd2,cov23,sd3])
          subject=arzt;

run;
```

Fazit

- Ein Stereotype Regressionsmodell für korrelierte Beobachtungen konnte hergeleitet werden.
- Faktoren, die ärztliche Setting-Präferenzen für die Rehabilitation von Patienten nach SHT beeinflussen, konnten identifiziert werden. Darüberhinaus gewinnt man zusätzliche Information über den Abstand der Kategorien
- Schätzung mit Standard-Software ist möglich
- Noch zu tun: Simulationsuntersuchungen, Andere Schätzverfahren

Literatur

Anderson JA, (1984): Regression and Ordered Categorical Variables. Journal of the Royal Statistical Society, Series B, 46, 1-30.

Hasenbein U, Kuss O, Bäumer M, Schert C, Schneider H, Wallesch CW, (2003): Physicians' preferences and expectations in traumatic brain injury rehabilitation - results of a case-based questionnaire survey. Disability and Rehabilitation, 25, 136-142.

Hartzel J, Agresti A, Caffo B, (2001): Multinomial Logit Random Effects Models. Statistical Modelling, 1, 81-102.

Kuss O, (2003): Modelling Physicians' Recommendations for Optimal Medical Care by Random Effects Stereotype Regression. In: Verbeke G, Molenberghs G, Aerts M, and Fieuws S [Eds.]: Proceedings of the 18th International Workshop on Statistical Modelling. Leuven: Katholieke Universiteit Leuven, 245-249.

Tutz G, Hennevogl W, (1996): Random effects in ordinal regression models. Computational Statistics & Data Analysis, 22, 537-557.