

---

# Ein SAS<sup>®</sup>-Makro zur Schätzung des Stereotype Regressionsmodells

**Oliver Kuß**

Abt. Klinische Sozialmedizin, Universitätsklinikum  
Heidelberg,  
Bergheimer Str. 58, 69115 Heidelberg  
[Oliver\\_Kuss@med.uni-heidelberg.de](mailto:Oliver_Kuss@med.uni-heidelberg.de)

**John Hendrickx**

Institut für Soziologie, Universität Nijmegen,  
P.O. Box 9108, 6500 HK Nijmegen, Niederlande  
[J.Hendrickx@mailbox.kun.nl](mailto:J.Hendrickx@mailbox.kun.nl)

4. Konferenz für SAS<sup>®</sup>-Anwender  
in Forschung und Entwicklung (KSFE)  
9. - 10. März 2000  
Justus-Liebig-Universität Gießen

---

# **Programm**

- 1. Ausgangspunkt (Beispieldatensatz)**
- 2. Potentielle Modelle zur Auswertung**
- 3. Stereotype Modell**
- 4. Schätzung des Stereotype Modells mit Hilfe eines SAS<sup>®</sup>-Makros**
- 5. Ergebnisse Beispieldatensatz**
- 6. Fazit**
- 7. Literatur**

---

# 1. Ausgangspunkt (Beispieldatensatz)

- Prospektive Kohortenstudie
- Alle Auszubildenden der AUDI AG der Einstellungsjahrgänge 1990-94 (Ingolstadt) bzw. 1990-93 (Neckarsulm)
- N=1910

## **Ziel**

Bestimmung von Risikofaktoren für die Bildung von berufsbedingten Handekzemen

## **Zielgröße**

Schweregrad des Handekzems (kein, leicht, schwer)

---

## 2. Potentielle Modelle zur Auswertung

**Gegeben:** Diskrete Zielgröße  $Y$  mit  $J$  Ausprägungen,  
 $k$  Kovariablen  $x$

**binäre** Zielgröße ( $J=2$ )

$$\log\left(\frac{p}{1-p}\right) = \alpha + \beta'x$$

mit  $p = P(Y=1 | x)$

- Logistisches Regressionsmodell
- $1+k$  Parameter

**ordinale** Zielgröße

$$\log\left(\frac{p_j}{1-p_j}\right) = \alpha_j + \beta'x$$

mit  $p_j = P(Y \leq j | x)$ ,  $j = 1, \dots, J-1$

- Proportional Odds-Modell
- $(J-1) + k$  Parameter
- Hergeleitet für zugrundeliegende stetige Zielgröße, die nur in Klassen beobachtet werden kann
- Geeignet ???

---

## nominale Zielgröße

$$\log\left(\frac{p_j}{p_1}\right) = \alpha_j + \beta_j'x$$

mit  $p_j = P(Y=j|x)$ ,  $j = 2, \dots, J$  und  $\alpha_1=0$ ,  $\beta_1=0$

- Multinomiales logistisches Modell
- Interpretation durch Vergleich mit Referenzkategorie
- $(J-1) + (J-1)*k$  Parameter  
→ Geeignet ???

---

### 3. Stereotype Modell

Reduktion der Parameterzahl durch Einführung von neuen Parametern  $\theta_j$  mittels

$$\beta_j = \theta_j \beta$$

**Modellgleichung:**

$$\log\left(\frac{p_j}{p_1}\right) = \alpha_j + \theta_j \beta'x$$

mit

$p_j = P(Y=j|x)$ ,  $j = 2, \dots, J$  und

$\theta_1 = 0$ ,  $\theta_J = 1$

**Eigenschaften:**

- Interpretation der  $\theta_j$  :

$$\log\left(\frac{p_j}{p_{j'}}\right) = \alpha_j - \alpha_{j'} + (\theta_j - \theta_{j'})\beta'x$$

Durchgang über die Kategorien der Zielgröße liefert bei Ordinalität eine Anordnung der  $\theta_j$

$(\theta_1 \leq \dots \leq \theta_J)$

- Tests auf Ordnung bzw. Ununterscheidbarkeit der  $\theta_j$  sind möglich
- Ordinalität wird bezüglich der Kovariablen abgebildet
- $2(J-1)+k$  Parameter

---

## 4. Schätzung des Stereotyp Modells

**In der Vergangenheit:** Einzellösungen oder informelle Beurteilungen von multinomialen Modellen

**Aber:** Es existierte kein routinemäßig anwendbares Programm/Makro in einem Standard-Software-Paket

**Jetzt:** SAS-Makros %MCL

(J. Hendrickx, <http://baserv.uci.kun.nl/~johnh/mcl/>)

### Prinzip:

- Betrachte multinomiales Modell als konditionales logistisches Modell → PROC PHREG, Umordnung des Datensatzes mit %MCLGEN
- Iterative, alternierende Schätzung von  $\beta$  und  $\theta$  durch %MCLEST, Übergabe der Parameter und wiederholter Aufruf von PROC PHREG mit PROC IML

**Problem:** Berechnete Standardfehler der Parameter nicht gültig →  $\alpha$ -Adjustierung oder Bootstrap-Korrektur

---

## 5. Ergebnisse Beispieldatensatz

### Beobachtet wurden:

167 schwere HE

103 leichte HE

1640 ohne HE

### 3 Kovariablen:

Feuchtarbeit ( $\geq 3$ h)

Anamnesticus Handekzem

Anamnesticus Beugeneckzem

	Multinomiales Modell		Stereotype Modell
	OR (95%-CI) kein/leicht	OR (95%-CI) kein/schwer	OR (95%*-CI)
Feuchtarbeit	2.42 (1.47, 3.98)	2.31 (1.54, 3.47)	2.33 (1.67, 3.24)
Anamn. HE	6.06 (2.78, 13.23)	5.41(2.75, 10.65)	5.55 (3.11, 9.90)
Anamn. BE	3.25 (1.48, 7.11)	3.42 (1.79, 6.54)	3.33 (1.90, 5.74)
$\theta_2$	.	.	1.03 (0.75,1.24)

---

## 6. Fazit

Das Stereotype Modell stellt im medizinischen Bereich eine nützliche Alternative zu den bekannten Regressionsmodellen mit mehrstufigen Zielgrößen dar

### Gründe:

- Modellierung der Zielgröße als multidimensionales Konstrukt
- Prozess des Durchlaufens von mehreren Stadien einer Krankheit kann in natürlicher Weise modelliert werden (Greenland, 1994)
- Allgemein zugängliche Software ist verfügbar

Methodische Probleme müssen ernstgenommen werden, sollten aber Anstoß für zukünftige Arbeit sein

---

## 7. Ausblick/Update

Kritik von Holtbrügge/Schumacher:

Verzerrte Schätzer, unterschätzte Standardfehler der Parameterschätzer im Vergleich zum Proportional Odds

**Aber:** Verzerrung minimal, verschwindet ab  $N > 200$ .

**Desweiteren:**

Schätzmethode !!!

Stereotype Modell kann auch als multivariates nicht-lineares Regressionsmodell aufgefasst werden,

Parameterschätzung durch LS oder ML (PROC MODEL), asymptotisch gute Eigenschaften

→ Simulationsstudie???

---

## 8. Literatur

- Ananth CV, Kleinbaum DG. Regression Models for Ordinal Responses: A Review of Methods and Applications. *Int J Epidemiol*, 26, 1323-1333, 1997.
- Anderson, JA. Regression and Ordered Categorical Variables. *J R Statist Soc B*, 46, 1-30, 1984.
- DiPrete T. Adding Covariates to Loglinear Models for the study of social mobility. *American Sociological Review*, 55, 757-773, 1990.
- Greenland S. Alternative Models for Ordinal Logistic Regression. *Stat Med*, 13, 1665-1677, 1994.
- Greenwood C, Farewell V. A Comparison of regression models for Ordinal Data in an Analysis of Transplanted-Kidney Function. *Can J Stat*, 16, 325-335, 1988.
- Hendrickx J, Ganzeboom HBG. Occupational Status Attainment in the Netherlands, 1920-1990. A Multinomial Logistic Analysis. *European Sociological Review*, 14, 387-403, 1998.
- Holtbrügge W, Schumacher M. A Comparison of Regression Models for the Analysis of Ordered Categorical Data. *Applied Statistics*, 40, 249-259, 1991.

Makros unter:

<http://baserv.uci.kun.nl/~johnh/mcl/>