

Seminar: Statistische Herausforderungen im Umgang mit fehlenden bzw. fehlerbehafteten Daten

im Wintersemester 2014/15

Prof. Dr. Thomas Augustin,
Eva Endres

Institut für Statistik, LMU

26. September 2014

① Allgemeine Aspekte und Anforderungen

② Inhalt und Themen

③ Themenvergabe und Termine

- 1 Allgemeine Aspekte und Anforderungen
- 2 Inhalt und Themen
- 3 Themenvergabe und Termine

Allgemeine Ziele eines Seminars

- Ziel: Einführung in/Vertiefung von wissenschaftlichen Arbeitstechniken
- Kollektives Erarbeiten eines komplexeren Themengebiets
 - + Individuelles Expertentum in einem Aspekt (dabei auch: Recherche, Reflexion, Aufbereitung, Präsentation mündlich und schriftlich, Vernetzung von Theorie/Praxis ...)
 - + Überblick über das Gebiet und Einordnen in Gesamtzusammenhang (Themenbereich, Studium, individuelle Profilbildung)
 - + Wissenschaftliche Kooperation und Kommunikation (Austausch, Diskussion, Kritik üben und aushalten, Anregungen für eigene Arbeit aufnehmen und erteilen, ...)

Allgemeine Aspekte und Anforderungen

- Im Kontakt mit Hauptbetreuer(in) als Feedback-Unterstützung (reaktive Betreuung!):
- **Aufbereitung** und **Vertiefung** des Basistextes in einer selbst gewählten Richtung
- Engagierte Teilnahme an der **Diskussion**. Vorbereitung anhand des bereitgestellten **Vorbereitungsmaterials**
 - Einreichung des Vorbereitungsmaterial (erste Fassung der Seminararbeit oder notfalls gut kommentierte Vortragsfolien) spätestens eine Woche vor dem Vortrag
- Vortragsfolien und die Hausarbeit sollen dem Standard unserer Disziplin entsprechend mit dem Textsatzprogramm **L^AT_EX** erstellt werden.
 - Zusatzangebot: Tutorien am 24.10.14 und 31.10.14
- **Fachliche Voraussetzungen** einer erfolgreichen Seminarteilnahme: Lineare Regressionsmodelle

Allgemeine Aspekte und Anforderungen

6 ECTS (Bachelor) / 9 ECTS (Master):

1. Referat (ca. 40 min),
2. Hausarbeit,
3. regelmäßige Teilnahme.

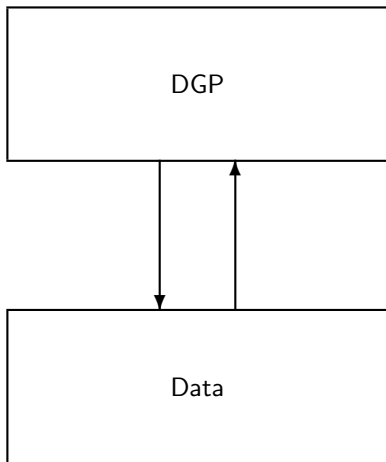
3 ECTS:

1. regelmäßige Teilnahme,
2. aktive Diskussionsbeteiligung,
3. mündliche Prüfung (ca. 15 Minuten).

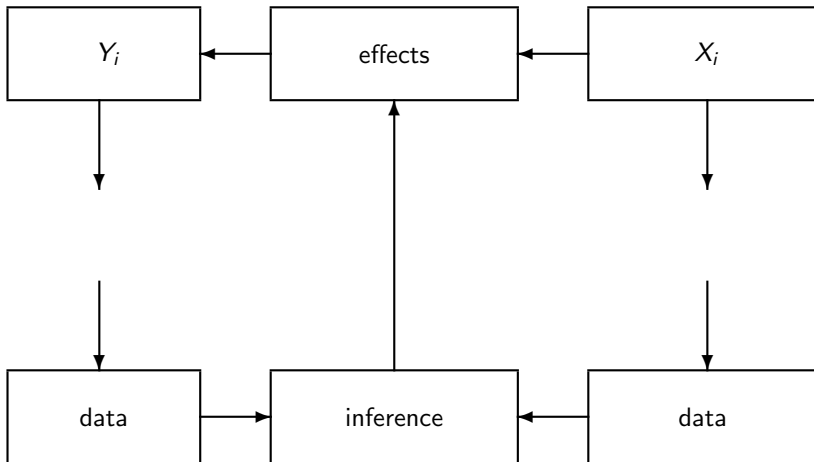
Alternativ zur mündlichen Prüfung: Kürzerer Seminarvortrag mit zusätzlichen kommentierten Folien

- ① Allgemeine Aspekte und Anforderungen
- ② Inhalt und Themen
- ③ Themenvergabe und Termine

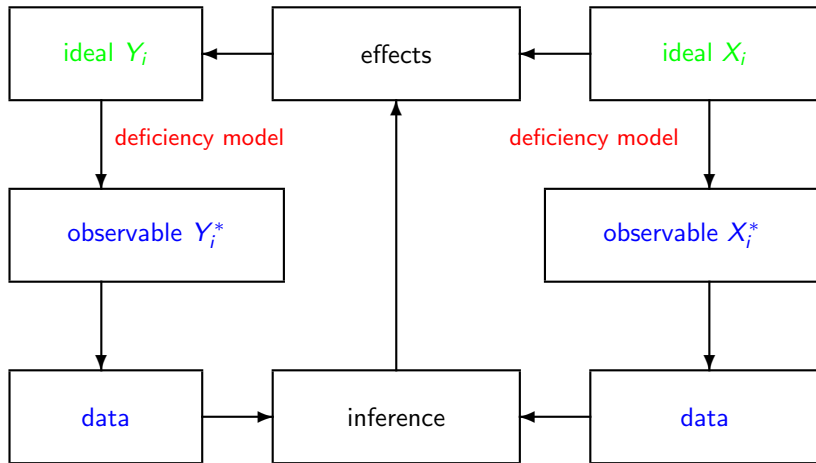
Statistische Modellierung



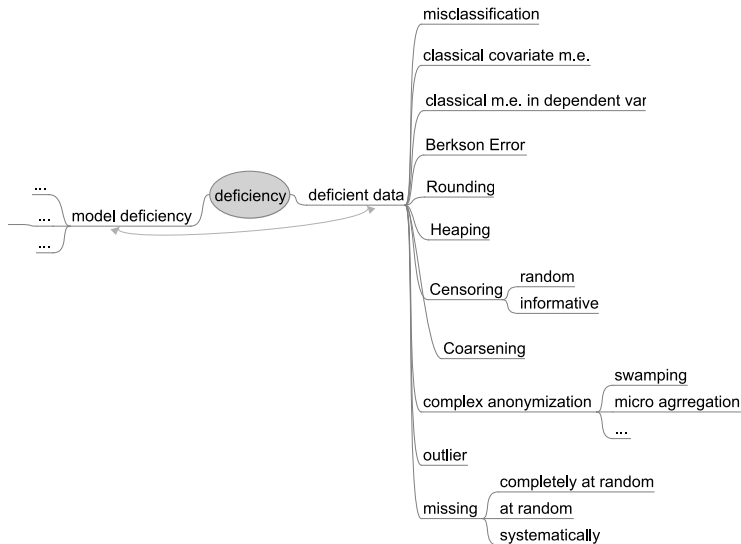
Statistische Modellierung



Die Zwei-Ebenen (Sprachen) Sicht



Spinney of Deficiencies



Themen

- **Fehler-in-den-Variablen**
- **Fehlklassifikation**
- **Fehlende Daten**

Themen

- **Fehler-in-den-Variablen**

Überblick über Messfehler und ihre Auswirkungen in der linearen Regression (B)

Literatur:

Carroll, R., Ruppert, D., Stefanski, L. and Crainiceanu, C. (2006).
Measurement Error in Nonlinear Models: A Modern Perspective, 2nd edn,
Chapman & Hall, Boca Raton, USA, Kap. 1–3

- **Fehlklassifikation**
- **Fehlende Daten**

Themen

- **Fehler-in-den-Variablen**

Regressionskalibrierung (B/M)

Literatur:

Carroll, R., Ruppert, D., Stefanski, L. and Crainiceanu, C. (2006).
Measurement Error in Nonlinear Models: A Modern Perspective, 2nd edn,
Chapman & Hall, Boca Raton, USA, Kap. 4

- **Fehlklassifikation**
- **Fehlende Daten**

Themen

- **Fehler-in-den-Variablen**

Simulation Extrapolation (SIMEX) (B/M)

Literatur:

Carroll, R., Ruppert, D., Stefanski, L. and Crainiceanu, C. (2006).
Measurement Error in Nonlinear Models: A Modern Perspective, 2nd edn,
Chapman & Hall, Boca Raton, USA, Kap. 5

- **Fehlklassifikation**

- **Fehlende Daten**

Themen

- **Fehler-in-den-Variablen**

Korrigierte Scorefunktionen (M)

Literatur:

Carroll, R., Ruppert, D., Stefanski, L. and Crainiceanu, C. (2006).
Measurement Error in Nonlinear Models: A Modern Perspective, 2nd edn,
Chapman & Hall, Boca Raton, USA, Kap. 7

- **Fehlklassifikation**
- **Fehlende Daten**

Themen

- **Fehler-in-den-Variablen**

Likelihood/Quasi-Likelihood für Messfehler (M)

Literatur:

Carroll, R., Ruppert, D., Stefanski, L. and Crainiceanu, C. (2006).
Measurement Error in Nonlinear Models: A Modern Perspective, 2nd edn,
Chapman & Hall, Boca Raton, USA, Kap. 8

- **Fehlklassifikation**
- **Fehlende Daten**

Themen

- **Fehler-in-den-Variablen**

Bayesianische Methoden für Messfehler (M)

Literatur:

Carroll, R., Ruppert, D., Stefanski, L. and Crainiceanu, C. (2006).
Measurement Error in Nonlinear Models: A Modern Perspective, 2nd edn,
Chapman & Hall, Boca Raton, USA, Kap. 9

- **Fehlklassifikation**

- **Fehlende Daten**

Themen

- **Fehler-in-den-Variablen**

Semiparametrische Regression (M)

Literatur:

Carroll, R., Ruppert, D., Stefanski, L. and Crainiceanu, C. (2006).
Measurement Error in Nonlinear Models: A Modern Perspective, 2nd edn,
Chapman & Hall, Boca Raton, USA, Kap. 13

- **Fehlklassifikation**
- **Fehlende Daten**

Themen

- **Fehler-in-den-Variablen**

Nonparametrische Schätzung (Dekonvolution) (M)

Literatur:

Carroll, R., Ruppert, D., Stefanski, L. and Crainiceanu, C. (2006).
Measurement Error in Nonlinear Models: A Modern Perspective, 2nd edn,
Chapman & Hall, Boca Raton, USA, Kap. 12

- **Fehlklassifikation**
- **Fehlende Daten**

Themen

- **Fehler-in-den-Variablen**

Messfehler in Überlebenszeit-Modellen (M)

Literatur:

Carroll, R., Ruppert, D., Stefanski, L. and Crainiceanu, C. (2006).
Measurement Error in Nonlinear Models: A Modern Perspective, 2nd edn,
Chapman & Hall, Boca Raton, USA, Kap. 14

- **Fehlklassifikation**
- **Fehlende Daten**

Themen

- **Fehler-in-den-Variablen**

Instrumentenvariablen (M)

Literatur:

Carroll, R., Ruppert, D., Stefanski, L. and Crainiceanu, C. (2006).
Measurement Error in Nonlinear Models: A Modern Perspective, 2nd edn,
Chapman & Hall, Boca Raton, USA, Kap. 6

- **Fehlklassifikation**
- **Fehlende Daten**

Themen

- **Fehler-in-den-Variablen**

Fehler in der abhängigen Variable (B/M)

Literatur:

Carroll, R., Ruppert, D., Stefanski, L. and Crainiceanu, C. (2006).
Measurement Error in Nonlinear Models: A Modern Perspective, 2nd edn,
Chapman & Hall, Boca Raton, USA, Kap. 15

- **Fehlklassifikation**
- **Fehlende Daten**

Themen

- Fehler-in-den-Variablen
- Fehlklassifikation

Fehlklassifikation; MCSIMEX (B/M)

Literatur:

Lederer, W. and Küchenhoff, H. (2006).

A short introduction to the SIMEX and MCSIMEX, *The Newsletter of the R Project* **6**: 26–31.

Carroll, R., Ruppert, D., Stefanski, L. and Crainiceanu, C. (2006).

Measurement Error in Nonlinear Models: A Modern Perspective, 2nd edn, Chapman & Hall, Boca Raton, USA, Kap. 15.3

Gustafson, P. (2003).

Measurement Error and Misclassification in Statistics and Epidemiology: Impacts and Bayesian Adjustments, Chapman & Hall, Boca Raton, USA, Kap. 3

- Fehlende Daten

Themen

- Fehler-in-den-Variablen
- Fehlklassifikation

Weiteres zu Fehlklassifikation (Likelihood/Bayes) (B/M)

Literatur:

Carroll, R., Ruppert, D., Stefanski, L. and Crainiceanu, C. (2006).
Measurement Error in Nonlinear Models: A Modern Perspective, 2nd edn,
Chapman & Hall, Boca Raton, USA, Kap. 15.3

Gustafson, P. (2003).
*Measurement Error and Misclassification in Statistics and Epidemiology:
Impacts and Bayesian Adjustments*, Chapman & Hall, Boca Raton, USA,
Kap. 5

- Fehlende Daten

Themen

- Fehler-in-den-Variablen
- Fehlklassifikation
- Fehlende Daten

Übersicht zu fehlenden Daten (B)

Literatur:

Little, R. and Rubin, D. (2002).

Statistical Analysis with Missing Data, Wiley & Sons, Inc., Hoboken, USA.

Themen

- Fehler-in-den-Variablen
- Fehlklassifikation
- Fehlende Daten

Imputationsverfahren (B/M)

Literatur:

Rässler, S., Rubin, D. and Zell, E. (2013).

Imputation, *Wiley Interdisciplinary Reviews: Computational Statistics*

5: 20–29.

Themen

- Fehler-in-den-Variablen
- Fehlklassifikation
- Fehlende Daten

EM-Algorithmus (B)

Literatur:

McLachlan, G. and Krishnan, T. (2004).

The em-algorithm, *Technical Report 24*, Humboldt-Universität Berlin, Center for Applied Statistics and Economics.

Themen

- Fehler-in-den-Variablen
- Fehlklassifikation
- Fehlende Daten

Statistical Matching/Data Fusion (B/M)

Literatur:

D'Orazio, M., Di Zio, M. and Scanu, M. (2006).

Statistical Matching: Theory and Practice, Wiley & Sons, Ltd., Sussex, England, Kap. 1 & 2

Themen

- Fehler-in-den-Variablen
- Fehlklassifikation
- Fehlende Daten

Treatment-Evaluationsproblematik (B/M)

Literatur:

Caliendo, M. and Hujer, R. (2006).

The microeconomic estimation of treatment effects - an overview, in O. Hübler and J. Frohn (eds), *Modern Econometric Analysis - Surveys on Recent Developments*, Springer, Heidelberg, Germany.

Themen

- Fehler-in-den-Variablen
- Fehlklassifikation
- Fehlende Daten

Vergrößerte Daten (M)

Literatur:

Heitjan, D. and Rubin, D. (1991).

Ignorability and coarse data, *The Annals of Statistics* **19**.

Themen

- Fehler-in-den-Variablen
- Fehlklassifikation
- Fehlende Daten

Partielle Identifikation (M)

Literatur:

Manski, C. (2003).

Partial Identification of Probability Distributions, Springer, New York, USA.

Themen

- Fehler-in-den-Variablen
- Fehlklassifikation
- Fehlende Daten

Anonymisierungstechniken (B/M)

Literatur:

Anonymisierungsansätze und Anonymisierungsverfahren (2005).
in G. Ronning, R. Sturm, J. Höhne, R. Lenz, M. Rosemann, M. Scheffler and
D. Vorgrimler (eds), *Statistik und Wissenschaft: Handbuch zur*
Anonymisierung wirtschaftsstatistischer Mikrodaten, Statistisches Bundesamt,
Wiesbaden, Germany.

- ① Allgemeine Aspekte und Anforderungen
- ② Inhalt und Themen
- ③ Themenvergabe und Termine**

Themenvergabe und Termine

Kalender