

# Statistik II: Regressions- und Varianz- analyse

Eine Einführung für Studierende der Psychologie

Helge Toutenburg und Christian Heumann  
mit Beiträgen von Michael Schomaker  
überarbeitet von Hans-Georg Sonnenberg

# Inhaltsverzeichnis

<b>1. Vorbemerkungen</b> .....	i
1.1 Vorwort zum Kurs .....	i
1.2 Vorwort zur Kursüberarbeitung .....	iii
1.3 Vorwort zur 2. Kursüberarbeitung .....	1

---

## Teil I. Regressionsanalyse

---

<b>2. Deskriptive lineare Regression</b> .....	5
2.1 Einführung .....	5
2.2 Regression: ein erstes Beispiel .....	7
2.3 Prinzip der kleinsten Quadrate .....	13
2.3.1 Vorsicht bei Ausreißern .....	18
2.3.2 Fazit .....	19
2.4 Güte der Anpassung .....	20
2.4.1 Zerlegung der Varianz der Y-Variablen .....	20
2.4.2 Korrelation und Güte der Anpassung .....	26
2.5 Anwendungsbeispiele zu den Berechnungen .....	27
2.6 Residualanalyse .....	35
2.7 Lineare Transformation der Originaldaten .....	36
2.8 Multiple lineare Regression .....	38
2.9 Lineare Regression mit kategorialen Regressoren ...	40
2.10 Korrelation und Kausalität .....	45
2.11 Aufgaben und Kontrollfragen zur Regression .....	57
2.12 Beispiele für Klausurfragen zum Kapitel .....	68
<b>3. Schließende lineare Regression</b> .....	73
3.1 Einführung .....	73

3.1.1	Modellannahmen	74
3.1.2	Schätzung von $\beta$	75
3.1.3	Schätzung von $\sigma^2$	76
3.2	Klassische Normalregression	76
3.2.1	Maximum-Likelihood-Schätzung	77
3.3	Spezifizieren von linearen Hypothesen	77
3.4	Prüfen der Hypothesen	80
3.4.1	Konfidenzbereiche	87
3.5	Vergleich von Modellen	87
3.5.1	Beispiel zum Modellvergleich	88
3.6	Prädiktorselektionsstrategien	90
3.7	Anwendungsfragen bei multipler Regression	93
3.7.1	Das Problem der Multikollinearität	93
3.7.2	Moderator- und Mediatorvariablen	94
3.8	Anwendungsbeispiel	96
3.9	Aufgaben zur schließenden Regression	100
3.10	Beispiele für Klausurfragen zum Kapitel	103
<b>4.</b>	<b>Logistisches Regressionsmodell</b>	<b>111</b>
4.1	Eine binäre erklärende Variable	121
4.2	Eine stetige erklärende Variable	123
4.3	Ein Beispiel	125
4.4	Odds Ratios vorwärts und rückwärts	128
4.5	Vektor- und Matrixschreibweise	130
4.6	Schätzung der Parameter	132
4.7	Tests	137
4.8	Modellbewertung	140
4.9	Zurück zum Beispiel	144

---

## Teil II. Varianzanalyse

---

<b>5.</b>	<b>Einfaktorielle Experimente</b>	<b>149</b>
5.1	Einführung	149
5.2	Einfaktorielle Varianzanalyse	155
5.2.1	Zerlegung der Fehlerquadratsumme	158
5.3	Effektstärke	165

5.4	Vergleich von einzelnen Mittelwerten	166
5.4.1	Lineare Kontraste	166
5.4.2	Multiple Vergleiche	173
5.5	Einfaktorielle Modelle mit zufälligen Effekten	176
5.6	Rangvarianzanalyse	183
5.6.1	Kruskal–Wallis–Test	183
5.7	Kontrollfragen und Aufgaben	187
5.8	Beispiele für Klausurfragen zum Kapitel	188
<b>6.</b>	<b>Zweifaktorielle Varianzanalyse</b>	<b>191</b>
6.1	Definitionen und Grundprinzipien	191
6.2	Zweifaktorielle Experimente mit Wechselwirkung	196
6.3	Kontrollfragen und Aufgaben	203
6.4	Beispiele für Klausurfragen zum Kapitel	205
<b>7.</b>	<b>Modelle mit Messwiederholungen</b>	<b>207</b>
7.1	Modell und Hypothesentest für eine Population	207
7.2	Kontrollfragen und Aufgaben	217
7.3	Beispiele für Klausurfragen zum Kapitel	218

---

### Teil III. Anhänge

---

<b>8.</b>	<b>Lösungen zu den Klausurfragen</b>	<b>223</b>
8.1	Aufgaben Kapitel 2	223
8.2	Aufgaben Kapitel 3	223
8.3	Aufgabe Kapitel 5	223
8.4	Aufgaben Kapitel 6	223
8.5	Aufgabe Kapitel 7	223
<b>9.</b>	<b>Literatur</b>	<b>225</b>
9.1	Literaturhinweise	225
9.2	Griechische Buchstaben	227
<b>10.</b>	<b>Tabellenanhang</b>	<b>229</b>
10.1	Verteilungsfunktion der Binomialverteilung	229
10.2	Verteilungsfunktion der Standardnormalverteilung	235
10.3	Quantile der $\chi^2$ -Verteilung	239

10.4 Quantile der t-Verteilung .....	241
10.5 Quantile der F-Verteilung .....	243



# 1. Vorbemerkungen

## 1.1 Vorwort zum Kurs

Dieser Kurs ist ein an der Universität München entwickelter Baustein des Moduls "Methoden der empirischen Sozialforschung, Statistik und computergestützte Datenanalyse"(Modul 2) im neuen BSc-Studiengang "Psychologie" der Fakultät Kultur- und Sozialwissenschaften der FernUniversität Hagen.

Der Kurs widmet sich Modellen der Regressions- und Varianzanalyse. Diese gehören zu den linearen Modellen, deren statistische Behandlung meist unter Verwendung der Matrizenrechnung erfolgt, weil so eine sehr kompakte Darstellung der Theorie möglich wird. Da die Leser dieses Kurses vermutlich nicht oder nur bedingt mit Matrixoperationen vertraut sind, wurde neben der Verwendung von Matrizen zusätzlich versucht, eine Kombination von mathematisch exakter und gleichzeitig verständlicher Darstellung zu erreichen. Dies geschieht vor allem durch Einbezug von Beispielen im Text, durch Verknüpfung des Lehrtextes mit SPSS-Präsentationen und durch Übungsaufgaben (für die Regression und für die Varianzanalyse jeweils ein zusätzlicher Aufgabenteil mit Lösungen). Die Leser sollten ihnen zu theoretisch erscheinende Textstellen im ersten Durchgang überblättern.

Ausführliche Beweise, weiterführende Kapitel und detaillierte Literaturangaben (die wegen der besseren Lesbarkeit in diesem Text z.T. weggelassen wurden) findet man in

H. Toutenburg (2003): *Lineare Modelle. Theorie und Anwendungen*. 2. Auflage, Physica Verlag, Heidelberg

H. Toutenburg (1994): *Versuchsplanung und Modellwahl*. Physica Verlag, Heidelberg (zweite Auflage in Vorbereitung)

Viele der im Text verwendeten SPSS-Beispieldateien sind unter folgender Adresse zu finden:

[www.stat.uni-muenchen.de/institut/ag/toutenb/psy/daten.html](http://www.stat.uni-muenchen.de/institut/ag/toutenb/psy/daten.html)<sup>1</sup>

Wir bedanken uns bei den Kollegen Prof. Dr. H.- J. Mittag und Prof. Dr. K.-H. Renner für die Einladung zu diesem Projekt und für die Unterstützung.

München, im November 2008

Helge Toutenburg    Christian Heumann    Michael Schomaker

---

<sup>1</sup> Diese Adresse wird von der Uni München offenbar nicht mehr unterstützt, Sie finden die Dateien inzwischen in der Moodle-Umgebung zu Statistik 2.



## 1.2 Vorwort zur Kursüberarbeitung

Nachdem dieser Kurs zwei Semester im BSc-Studiengang „Psychologie“ der Fakultät Kultur- und Sozialwissenschaften der FernUniversität Hagen eingesetzt worden ist, erhielt ich Gelegenheit zu einer ersten Überarbeitung. Diese konnte auf zahlreichen Beiträgen von Studenten und Mentoren in der Elearning-Umgebung Moodle zu diesem Kurs aufbauen sowie auf der Erfahrung aus einer ganzen Reihe von Durchführungen der von Prof. Dr. Renner entwickelten Präsenzveranstaltungen. Sie steht im Rahmen der kontinuierlichen Weiterentwicklung des gesamten Lehrmaterials im Bereich des Moduls 2.

Die Überarbeitung hatte das Ziel, die mathematische Prägnanz der Darstellung zu erhalten, auch wenn dies für Studienanfänger im Fach Psychologie eine Hürde darstellen kann und ein bestimmtes Maß an Einarbeitung verlangt. Gleichwohl wurden doch einige spezielle Gebiete und Fallunterscheidungen der 1. Version ausgelassen und die mathematische Beweisführung teilweise in Exkurse ausgelagert. Die Erwartung ist, dass Studierende auch im weiteren Verlauf ihres Studiums, zum Beispiel beim Empirisch-Experimentellen Praktikum oder bei der Erstellung der Bachelorarbeit, Nutzen daraus ziehen, wenn sie auf das Kursmaterial zurückgreifen können.

Ich selbst machte Bekanntschaft mit diesem Stoffgebiet im Rahmen meines Psychologiestudiums in Vorlesungen von Prof. Dr. Wottawa und Dr. Rochel an der Universität Bochum mit dem Konzept des „Allgemeinen Linearen Modells“ und wurde später durch den Umgang mit Statistik-Software geprägt. Im Sommersemester 2009 hatte ich dann Gelegenheit, die Aufbereitung des Themas durch Herrn Prof. Dr. Holger Dette von der Fakultät für Mathematik<sup>2</sup> für die Bachelorstudenten des 2. Semesters der Psychologischen Fakultät der Universität Bochum zu begleiten. Wichtig waren für mich außerdem wertvolle Hinweise von Dr. Raphael Diepgen, der seit vielen Jahren diese Themen vermittelt. Über das

---

<sup>2</sup> <http://www.ruhr-uni-bochum.de/mathematik3/lehre/ss09/methodenlehre2.html>

dabei akkumulierte „kulturelle Kapital“ informiert beispielsweise die Sammlung der Bochumer Klausuren<sup>3</sup> aus einem Zeitraum von fast 20 Jahren, die über den angegebenen Link einzusehen ist.

Dieser Kurs ist aus meiner Sicht, nicht zuletzt aufgrund der Schwierigkeit und Breite des Stoffes und der Heterogenität der Vorkenntnisse der Adressaten, eine herausfordernde „Baustelle“, für die ich weiterhin gern Feedback und Meinungen entgegennehme. Studierhinweise, Beispieldaten sowie Informationen zur Gewichtung bezüglich der Klausuraufgaben werden im Laufe des Semsters in Moodle bereitgestellt.

Hagen, im Oktober 2009

Hans-Georg Sonnenberg

---

<sup>3</sup> [http://www.ruhr-uni-bochum.de/psy-methoden/methodenlehre/klausur\\_archiv.html](http://www.ruhr-uni-bochum.de/psy-methoden/methodenlehre/klausur_archiv.html)

## 1.3 Vorwort zur 2. Kursüberarbeitung

Es ist meine traurige Pflicht Sie zu informieren, dass der Erstautor dieses Kurses, Herr Prof. Dr. Dr. Helge Toutenburg, im November 2009 plötzlich und unerwartet verstorben ist.

Die jüngste Kursüberarbeitung konnte sich neben der Neufassung des Kapitels zur „Logistischen Regression“ auf die Korrektur von Fehlern konzentrieren, die bei einem Text dieser Komplexität nicht zuletzt dank der aufmerksamen Lektüre vieler Studierender noch entdeckt wurden. Als neue Leser seien Sie daran erinnert, dass Sie im Umgang mit dem Lehrtext Fragen aller Art in der Moodle-Lernumgebung stellen können. Ausserdem steht in der „Virtuellen Universität“ die Videoaufzeichnung eines Repetitoriums zu dieser Kurseinheit mit Aufgaben und SPSS-Datenbeispielen für Sie bereit.

Hagen, im März 2010

Hans-Georg Sonnenberg