
Einführung in die statistische Auswertung von Experimenten

Torben Kuhlenkasper · Andreas Handl

Einführung in die statistische Auswertung von Experimenten

Theorie und Praxis mit R

Torben Kuhlenkasper
Hochschule Pforzheim
Pforzheim, Deutschland

Andreas Handl
Bielefeld, Deutschland

ISBN 978-3-662-59053-9 ISBN 978-3-662-59054-6 (eBook)
<https://doi.org/10.1007/978-3-662-59054-6>

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Springer Spektrum

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature 2019

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jedermann benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des jeweiligen Zeicheninhabers sind zu beachten.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag, noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Planung/Lektorat: Iris Ruhmann

Springer Spektrum ist ein Imprint der eingetragenen Gesellschaft Springer-Verlag GmbH, DE und ist ein Teil von Springer Nature

Die Anschrift der Gesellschaft ist: Heidelberger Platz 3, 14197 Berlin, Germany

Vorwort

„Versuchen Sie, einen Sinn in dem zu erkennen, was Sie sehen.... Bleiben Sie neugierig!“ Dieser Wunsch des weltbekannten Physikers Stephen Hawking wurde anlässlich seines Todes von der University of Cambridge in einer Videobotschaft im März 2018 veröffentlicht (s. University of Cambridge 2018). Im gleichen Monat entstand auch die Idee für das vorliegende Buch. Der Wunsch des Ausnahmephysikers war beim Schreiben in den letzten Monaten stets eine große Motivation für mich. Mit dieser Einführung in die statistischen Methoden zur Auswertung von Experimenten möchte ich auch beim Leser sowohl Neugierde entfachen als auch befriedigen: Neugierde, mit welchen Methoden wir einen Sinn in dem erkennen können, was wir sehen und was uns interessiert. Nicht nur in den vielen Disziplinen der Wissenschaft, sondern in nahezu allen Lebensbereichen versuchen wir, Zusammenhänge zu verstehen. Mithilfe von Experimenten ganz unterschiedlicher Art decken wir dabei immer wieder Phänomene auf und ziehen daraus unsere Schlüsse für Wissenschaft und Alltag.

Mit dem vorliegenden Buch möchte ich einen kleinen Beitrag dazu leisten, mit welchen Methoden der Statistik wir unsere Neugierde systematisch befriedigen können. Dazu werden wir eine Vielzahl von statistischen Methoden zur Auswertung von Experimenten kennenlernen und anwenden. Solche Experimente müssen nicht in großen Laboren durchgeführt werden. Die Beispiele in dem Buch sind leicht nachvollziehbar und kommen oft aus ganz alltäglichen Situationen. Genauso hat auch Andreas Handl an der Universität Bielefeld seinen Studierenden die Methoden zur Auswertung von Versuchen vorbildlich erklärt. Dabei hat er auch meine Neugierde geweckt. Auf diesen Aufzeichnungen von Andreas basiert dieses Lehrbuch. Ihm waren dabei sowohl die Beispiele von leicht nachvollziehbaren Experimenten im Vordergrund als auch die mathematischen Methoden zur Auswertung im Hintergrund gleichermaßen wichtig. Ich habe in den letzten Jahren versucht, genau diese Art der Lehre von Andreas aufzugreifen und seine Ideen und sein Ideal weiterzuentwickeln. Ich setze die Aufzeichnungen seit mehreren Jahren in meinen Vorlesungen ein und möchte mich bei allen Kollegen und vor allem bei den vielen Studenten der letzten Jahre für die wertvollen Kommentare bedanken. Für sie ist auch dieses Buch geschrieben!

Das Buch stellt die gängigen Methoden der Varianzanalyse und der nichtparametrischen Alternativen für balancierte Experimente mit festen Faktoren vor. Wenn man jede Untersuchungsanordnung zur Überprüfung von Hypothesen als Experiment auffasst, kann das Buch auch als Einführung in die Analyse von unverbundenen Stichproben verwendet werden. Es ist als Einführungsbuch für Studierende und Wissenschaftler mit Grundkenntnissen der Statistik gedacht. Der Leser des Buches ist bei den kleinen Beispielen zunächst dazu aufgefordert, die Experimente – und somit die Daten – mit Stift und Papier zu analysieren und auszuwerten. Mit diesem Wissen ist es dann sehr einfach, die gleichen Ergebnisse mit der Software R zu erhalten und in Zukunft die eigenen Ideen in R umzusetzen.

Ohne die Lehre und die Aufzeichnungen von Andreas wäre das Buch nicht möglich gewesen. Ich möchte daher besonders Claudia und Fabian Handl ganz herzlich für das Vertrauen dafür danken, dass ich die Ideen von Andreas aufgreifen und weiterentwickeln kann. Wer mehr über den Bielefelder „Statistiker mit Herz und Verstand“ erfahren möchte, findet unter www.andreashandl.de viele persönliche Informationen über den Initiator dieses Buches.

Auch dem Springer Verlag danke ich für das Vertrauen, sowohl in dieses Einführungsbuch als auch in das für 2020 geplante Buch mit den fortgeschrittenen Methoden zur Auswertung von Experimenten.

Unter www.experimente.kuhlenkasper.de stehen die verwendeten Datensätze des Buches, der Quellcode für die R-Anweisungen sowie die Lösungen für die Übungsaufgaben und auch weitere Informationen für den Leser bereit.

Bad Essen
im Februar 2019

Torben Kuhlenkasper

Inhaltsverzeichnis

- 1 Einführung** 1
- 2 Einführung in R** 7
 - 2.1 R als mächtiger Taschenrechner 7
 - 2.2 Datenstrukturen 10
 - 2.3 Pakete 18
 - 2.4 Einlesen von Daten aus externen Dateien 19
 - 2.5 Selektion unter Bedingungen 24
- 3 Einfaktorielle Experimente** 29
 - 3.1 Grundlagen 29
 - 3.2 Balancierte Experimente mit zwei Faktorstufen 37
 - 3.2.1 Schätzer des Effekts von A 38
 - 3.2.2 *t*-Test 40
 - 3.2.3 Algorithmus von Yates 42
 - 3.3 Einfaktorielle Experimente in R 45
 - 3.4 Übungsaufgaben 47
- 4 Annahmen der Varianzanalyse** 49
 - 4.1 Normalverteilung 50
 - 4.2 Varianzhomogenität 58
 - 4.3 Überprüfung der Annahmen mit R 63
 - 4.4 Übungsaufgaben 65
- 5 Zweifaktorielle Experimente** 67
 - 5.1 Additives Modell 70
 - 5.2 Nichtadditives Modell 80
 - 5.2.1 Der Algorithmus von Yates 91
 - 5.3 Sonderfall $n=1$ 96
 - 5.4 Beispiel eines zweifaktoriellen Experiments 99
 - 5.5 Zweifaktorielle Experimente in R 102
 - 5.6 Übungsaufgaben 107

6	<i>k</i>-faktorielle Experimente	109
6.1	Haupteffekte	110
6.2	Interaktionseffekte zwischen zwei Faktoren	112
6.3	Interaktionseffekte zwischen mehr als zwei Faktoren	113
6.4	Varianzanalyse	114
6.4.1	Algorithmus von Yates	114
6.5	Sonderfall $n=1$	125
6.6	<i>k</i> -faktorielle Varianzanalyse in R	131
6.7	Übungsaufgaben	137
7	Fraktionelle faktorielle Experimente	139
7.1	Grundlagen	139
7.2	2^{4-1} -Experiment	147
7.3	Fraktionelle faktorielle Varianzanalyse in R	152
7.4	Übungsaufgaben	155
8	Alternative Auswertungsmethoden	157
8.1	Welch-Test	157
8.1.1	Welch-Test mit R	161
8.1.2	Übungsaufgabe	162
8.2	Mann-Whitney-Test	163
8.2.1	Bindungen	173
8.2.2	Mann-Whitney-Test mit R	175
8.2.3	Übungsaufgaben	177
8.3	Kruskal-Wallis-Test	177
8.3.1	Bindungen	182
8.3.2	Kruskal-Wallis-Test mit R	185
8.3.3	Übungsaufgaben	186
8.4	Varianzanalyse mit Rängen	186
8.4.1	Varianzanalyse mit Rängen in R	190
8.4.2	Übungsaufgaben	192
Anhang A: R-Funktionen		193
Anhang B: Beweise und Herleitungen		197
Anhang C: Tabellen		211
Literatur		225
Stichwortverzeichnis		227