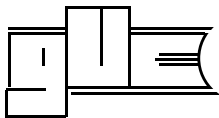


Bodo Sturm

Statistik: Eine Einführung mit R

2., überarbeitete und aktualisierte Auflage



GUC - Verlag der Gesellschaft für
Unternehmensrechnung und Controlling m.b.H.
Chemnitz 2019

Bibliografische Information der Deutschen Nationalbibliothek

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

Sturm, Bodo:

Statistik: Eine Einführung mit R, 2., überarbeitete und aktualisierte Auflage / Bodo Sturm - Chemnitz · Löbnitz: Verlag der GUC, 2019

(Lehrbuchreihe; 12)

ISBN 978-3-86367-060-3

© 2019 by Verlag der GUC - Gesellschaft für Unternehmensrechnung und Controlling m.b.H.

GUC m.b.H. · Chemnitz · Löbnitz

<http://www.guc-verlag.de>

Alle Rechte vorbehalten. Dieses Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist - wenn sie ohne Zustimmung des Verlages erfolgt - unzulässig und strafbar. Dies gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen sowie die Einspeicherung und Verarbeitung in elektronischen Systemen.

Druck und Weiterverarbeitung: Druckerei und Verlag Mike Rockstroh, Aue

Gedruckt auf säurefreiem Papier - alterungsbeständig

Printed in Germany

ISBN 978-3-86367-060-3

Über dieses Buch	9
1 Einführung	11
1.1 Was ist Statistik?	11
1.2 Grundlagen	12
1.3 Aufgaben	17
1.4 Statistik mit R	17
2 Darstellung und Beschreibung qualitativer Daten	25
2.1 Eindimensionale qualitative Daten	25
2.2 Zweidimensionale qualitative Daten	27
2.3 Aufgaben	31
2.4 Anwendung mit R	32
3 Darstellung und Beschreibung quantitativer Daten	37
3.1 Darstellung einer Verteilung	37
3.2 Form und Lagemaße einer Verteilung	44
3.3 Streuung einer Verteilung	48
3.4 Boxplot	55
3.5 Vergleich von Gruppen	56
3.6 Aufgaben	57
3.7 Anwendung mit R	59
4 Assoziation und Korrelation	65
4.1 Korrigierter Kontingenzkoeffizient	65
4.2 Streudiagramm und Kovarianz	69
4.3 Korrelation	71
4.4 Aufgaben	76
4.5 Anwendung mit R	78
4.6 Anhang: Maximale quadratische Kontingenz	84
5 Lineare Regression	87
5.1 Lineares Modell	87
5.2 Regression to the Mean	92
5.3 Modellannahmen	94
5.4 Untersuchung der Fehlerterme	95
5.5 Bestimmtheitsmaß	96
5.6 Nicht-lineare Zusammenhänge	98
5.7 Aufgaben	103
5.8 Anwendung mit R	106
6 Zufall und Wahrscheinlichkeit	111
6.1 Grundbegriffe	111
6.2 Arten von Wahrscheinlichkeiten	112
6.3 Rechenregeln für Wahrscheinlichkeiten	115
6.4 Kontingenztabelle und Wahrscheinlichkeiten	118
6.5 Aufgaben	120
6.6 Anwendung mit R	124

7	Zufallsvariablen und ausgewählte Verteilungen	129
7.1	Zufallsvariablen	129
7.2	Verteilung	130
7.3	Median und Quantile	133
7.4	Erwartungswert und Standardabweichung	133
7.5	Standardisierung	135
7.6	Ausgewählte Verteilungen	136
7.7	Aufgaben	151
7.8	Anwendung mit R	154
7.9	Anhang: Erzeugen von Zufallsvariablen	159
8	Grenzwertsätze	161
8.1	Gesetz der großen Zahl	161
8.2	Bernoullis Gesetz der großen Zahl	165
8.3	Zentraler Grenzwertsatz	168
8.4	Aufgaben	175
8.5	Anwendung mit R	177
9	Schätzung unbekannter Parameter	181
9.1	Punktschätzungen	184
9.2	Stichprobenverteilungen	188
9.3	Konfidenzintervall für den Mittelwert einer Grundgesamtheit	193
9.4	Konfidenzintervall für den Anteilswert einer Grundgesamtheit	199
9.5	Exkurs: Konfidenzintervall für den Anteilswert einer Grundgesamtheit bei wenigen Beobachtungen	204
9.6	Exkurs: Wann liegt eine „große Stichprobe“ vor?	205
9.7	Aufgaben	206
9.8	Anwendung mit R	208
9.9	Anhang: Streuungsmaße	214
10	Hypothesentests für eine Stichprobe	217
10.1	Grundlagen	217
10.2	Hypothesentest für den Mittelwert einer Grundgesamtheit	221
10.3	Exkurs: Der Fehler 2. Art	232
10.4	Hypothesentest für den Anteilswert einer Grundgesamtheit	235
10.5	Aufgaben	238
10.6	Anwendung mit R	239
11	Hypothesentests für zwei Stichproben und Verteilungen qualitativer Daten	245
11.1	Vergleich zweier Lagemaße	245
11.2	Inferenz für Verteilungen qualitativer Daten	259
11.3	Aufgaben	265
11.4	Anwendung mit R	267
11.5	Anhang: Testverfahren im Überblick	277

12	Hypothesentests für lineare Regression und Korrelation	279
12.1	Inferenz für lineare Regression	279
12.2	Test auf Korrelation	286
12.3	Aufgaben	289
12.4	Anwendung mit R	290
13	Weiterführende Literatur	297
14	Tabellenanhang	299
14.1	Standardnormalverteilung – Verteilungsfunktion	299
14.2	Binomialverteilung – Massenfunktion	300
14.3	t -Verteilung – Verteilungsfunktion	302
14.4	χ^2 -Verteilung – Verteilungsfunktion	303
14.5	Kritische Werte für den MWU-Test	304
14.6	Kritische Werte für den Wilcoxon-Test für gepaarte Stichproben	305
15	Verzeichnisse	307
15.1	Stichwortverzeichnis	307
15.2	Verzeichnis der wichtigsten R-Befehle	309

Über dieses Buch

Dieses Buch bietet – nun in seiner 2., überarbeiteten Auflage – eine Einführung in die Statistik. Abgedeckt werden die Grundlagen der deskriptiven und induktiven Statistik, welche von mir in der Veranstaltung „Wirtschaftsstatistik“ der Bachelor-Studiengänge an der Fakultät Wirtschaftswissenschaft und Wirtschaftsingenieurwesen der HTWK Leipzig gelehrt werden. Im Vordergrund steht dabei die Anwendung statistischer Methoden. Die mathematischen Anforderungen sind auf ein Mindestmaß reduziert. Neben vielen Beispielen zur Anwendung und Verdeutlichung der statistischen Methoden werden zu jedem Kapitel Übungsaufgaben präsentiert. Aufgrund der knappen Zeit, die im Studium für Statistik zur Verfügung steht, ist eine Konzentration auf die essentiellen statistischen Konzepte notwendig. In Kap. 13 werden daher Hinweise für weiterführende Literatur gegeben.

Die Einführung in die Statistik wird in diesem Buch ergänzt mit einer Einführung in R. R ist eine mittlerweile sehr populäre, leistungsfähige und kostenfreie Statistik-Software. Die Kombination der Einführung in die Statistik mit einem Grundkurs in R ermöglicht es, die statistischen Methoden direkt am Rechner anzuwenden und damit besser zu verstehen. Die Einführung in die Statistik-Software R ist an die zuvor betrachteten statistischen Methoden gekoppelt und entspricht im Wesentlichen dem Statistik-Teil der Veranstaltung „Quantitative Methoden“ in den Master-Studiengängen an der Fakultät Wirtschaftswissenschaft und Wirtschaftsingenieurwesen der HTWK Leipzig. Jedes Kapitel im Buch ist daher in drei Teile gegliedert: (i) Statistische Methoden, (ii) Übungsaufgaben und (iii) Anwendung mit R. Dabei sind die Übungsaufgaben auch ohne Software lösbar.

In Kap. 14 (Anhang) sind Tabellen mit Verteilungen ausgewählter Zufallsvariablen zu finden. Ausführliche Lösungen zu den Übungsaufgaben, die im Buch verwendeten Datensätze und den R-Code für Abbildungen und Berechnungen in den Kapiteln gibt es im Internet unter <http://bsturm.htwk-leipzig.de/> oder auf Nachfrage per E-Mail (bodo.sturm@htwk-leipzig.de) beim Autor.

Dieses Buch wäre nicht entstanden ohne die großartige und unermüdliche Unterstützung meines Vaters, Martin Sturm, der die Darstellung inhaltlich und sprachlich deutlich verbessert hat. Vielen Dank für Deine Hilfe! An der Fehlerkorrektur und der Aufbereitung der Lösungen zu den Übungsaufgaben haben weiterhin Stefanie Burkhardt, Andrea Gauselmann, Natascha Götzte, Mathis Kirchner und Philipp Radomski mitgewirkt. Auch ihnen gilt mein Dank. Alle verbleibenden Fehler gehen natürlich auf mein Konto.

Leipzig, im Dezember 2018

Bodo Sturm