

---

# Springer-Lehrbuch

Weitere Informationen zu dieser Reihe finden Sie unter  
<http://www.springer.com/series/1183>

---

Ansgar Steland

# Basiswissen Statistik

Kompaktkurs für Anwender aus Wirtschaft,  
Informatik und Technik

4., überarbeitete Auflage

 Springer Spektrum

Ansgar Steland  
Institut für Statistik und Wirtschaftsmathematik  
RWTH Aachen  
Aachen, Deutschland

ISSN: 0937-7433

Springer-Lehrbuch

ISBN: 978-3-662-49947-4

ISBN: 978-3-662-49948-1 (eBook)

DOI 10.1007/978-3-662-49948-1

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 Berlin Heidelberg 2009, 2013, 2016

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 Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

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.

Planung: Iris Ruland

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier

Springer-Spektrum ist Teil von Springer Nature

Die eingetragene Gesellschaft ist Springer-Verlag GmbH Berlin Heidelberg

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

*Für Heike, Solveig und Adrian*

---

## Vorwort

Modelle und Methoden der angewandten Wahrscheinlichkeitstheorie und Statistik sind aus den modernen Wissenschaften, aber auch aus Industrie und Gesellschaft, nicht mehr wegzudenken. Wirtschaftswissenschaftler, Informatiker und Ingenieure benötigen heutzutage profunde Kenntnisse in diesen Bereichen. Zufallsbehaftete Phänomene sind durch stochastische Ansätze zu modellieren und anfallende Daten durch statistische Methoden zu analysieren. Wahrscheinlichkeitstheorie und Statistik haben sich nicht nur bei klassischen Aufgaben wie der Modellierung und Auswertung von Umfragen, Experimenten oder Beobachtungsstudien bewährt. Sie spielen auch eine entscheidende Rolle für das theoretische Verständnis hochkomplexer Systeme. Dies ist wiederum oftmals die notwendige Grundlage für die Entwicklung moderner Produkte und Dienstleistungen. Beispielhaft seien hier die modernen Finanzmärkte und der Datenverkehr im Internet genannt.

Der in diesem Text behandelte Stoff umfasst hauptsächlich die in der anwendungsorientierten Statistik-Ausbildung für Informatiker, Wirtschaftswissenschaftler und Ingenieure allgemein üblichen Themen. Insbesondere sind die Inhalte der zugehörigen Lehrveranstaltungen an der RWTH Aachen abgedeckt. In diesem Kompaktkurs bin ich sparsam – aber gezielt – mit illustrierenden Beispielen umgegangen. Viele sind so einfach wie möglich gehalten, um das berühmte Aha-Erlebnis zu ermöglichen. Andere wollen motivieren und zeigen daher Anwendungen auf. Ein ausführlicher mathematischer Anhang, *Mathematik – kompakt*, stellt die wichtigsten mathematischen Zusammenhänge, Formeln und Methoden aus Analysis und linearer Algebra zusammen. So ist ein schnelles und zielführendes Nachschlagen möglich.

Das zugrunde liegende didaktische Konzept wurde über viele Jahre an mehreren deutschen Universitäten entwickelt. Studierende tun sich in den ersten Semestern oftmals mit mathematischen Formalismen schwer. Unter dem Motto: „So wenig Formalismus wie möglich, aber so viel wie nötig“ habe ich versucht, diesem Umstand Rechnung zu tragen. Die Erfahrung zeigt, dass hierdurch die eigentlichen mathematischen Inhalte – um die es ja geht – von den Studierenden schneller und leichter erfasst und verstanden werden. So manche Erklärung eines mathematischen Sachverhalts lebt davon, dass der Lehrende seine Worte mit einer kleinen Skizze veranschaulicht oder in Schritten eine Formel entwickelt.

Dies lässt sich in einem Buch nicht umsetzen. Ich habe mich aber bemüht, möglichst viele eingängige verbale Erklärungen aufzunehmen, die sich im Lehralltag bewährt haben.

Einige mit einem Sternchen gekennzeichneten Abschnitte sind etwas anspruchsvoller oder nur für einen Teil der Leserschaft gedacht. Dort werden jedoch auch Themen angesprochen, die einen kleinen Einblick in wichtige Bereiche der modernen angewandten Stochastik und Statistik bieten und vielleicht den einen oder anderen Leser motivieren, in weiterführende Literatur zu schauen.

Mein Dank gilt Barbara Giese, die weite Teile dieses Buchs mit großer Expertise und Sorgfalt getippt und das Layout verbessert hat. Dipl.-Math. Sabine Teller und Dipl.-Math. André Thrun haben das Manuskript sehr gewissenhaft durchgesehen, etliche Tippfehler und Ungenauigkeiten gefunden und Verbesserungsvorschläge gemacht. Frau Lilith Braun vom Springer-Verlag danke ich für die angenehme und vertrauensvolle Zusammenarbeit bei diesem Buchprojekt.

Aachen,  
15. Juli 2007

Ansgar Steland

---

## Vorwort zur zweiten Auflage

Für die zweite Auflage wurden Tippfehler und Ungenauigkeiten korrigiert und an unzähligen Stellen Ergänzungen und Verbesserungen vorgenommen. In Anbetracht der guten Prüfungsergebnisse der Aachener Studierenden, die nach diesem Kompaktkurs lernen, und des überraschenden Verkaufserfolges, wurde das Grundkonzept jedoch beibehalten.

Der Anhang *Mathematik – kompakt* wurde ebenfalls durchgesehen und ergänzt. Die Arbeit mit und die Erstellung von englischen Dokumenten wird immer wichtiger. Zur Unterstützung der Studierenden wurde hierzu ein Glossar mit den wichtigsten Begriffen aus Mathematik, Wahrscheinlichkeitsrechnung und Statistik erstellt. Schließlich wurde ein Anhang mit Tabellen der wichtigsten statistischen Testverteilungen angefügt.

Alle Studierenden, die uns auf Fehler und Verbesserungsmöglichkeiten aufmerksam gemacht haben, gilt mein Dank. Frau Simone Gerwert hat mit großer Sorgfalt und kontinuierlichem Engagement alle Änderungen in das Latex-Dokument eingearbeitet.

Aachen,  
8. September 2009

Ansgar Steland

---

## Vorwort zur dritten Auflage

Die dritte Neuauflage wurde um viele zusätzliche Beispiele ergänzt, um das selbstständige Lernen und Nachbereiten zu erleichtern. Die Einführung von *Meilensteinen* hat sich in den Lehrveranstaltungen sehr bewährt. An Meilensteinen wird in der Praxis – insbesondere bei Projekten – sehr ernsthaft und oftmals bis ins Detail überprüft, inwieweit geplante Aktivitäten erledigt wurden, aufzubauende Fähigkeiten tatsächlich vorhanden sind und gesteckte Ziele erreicht wurden. Bezugnehmend auf die universitäre Lernsituation wurden für die Meilensteine Fragen und Aufgaben konzipiert, die in Form von Lückentexten, einem stärkerem Praxisbezug, offen gestellten Fragen oder Arbeitsaufträgen an die Studierenden als zukünftige Mitarbeiter/innen helfen sollen, den eigenen Wissenstand im Sinne von passivem Verständnis (Nachvollziehen) und aktivem Handlungswissen selbstständig zu überprüfen.

Darüber hinaus wurde der Text gründlich durchgesehen und an vielen Stellen verbessert und ergänzt. Insbesondere wurde der Anhang *Mathematik kompakt* erweitert, auch im Hinblick auf die geänderten Vorkenntnisse der Studierenden aufgrund der verkürzten Abiturzeit. Schließlich wurde das Glossar ausgebaut, um die Arbeit mit englischsprachigen Texten zu erleichtern.

Aachen,  
5. Februar 2013

Ansgar Steland



---

## Vorwort zur vierten Auflage

Erneut wurde der Text kritisch durchgesehen, didaktisch verbessert und um eine Reihe von Beispielen und Illustrationen ergänzt. Die *Meilensteine* haben sich zusammen mit zusätzlichen Online-Materialien, wie einer Smartphone-tauglichen Formelsammlung, als wertvolle Hilfestellung für die Studierenden bewährt. Sie sind in Blöcke strukturiert, die sich am Lehr- und Lernverlauf orientieren. Herr M. Sc. Andreas Sommer hat die dritte Auflage sorgfältig durchgesehen. Bei der Überarbeitung hat Frau Stefanie Truong an einigen Stellen mitgeholfen. Diese wurden von Herrn Sommer und Frau M. Sc. Katharina Bosch mit sehr großer Sorgfalt durchgesehen. Dem Springer-Verlag danke ich für die vertrauensvolle Zusammenarbeit.

Aachen,  
7. Januar 2016

Ansgar Steland

---

# Inhaltsverzeichnis

<b>1</b>	<b>Deskriptive und explorative Statistik</b>	<b>1</b>
1.1	Motivation und Beispiele	1
1.2	Grundgesamtheit und Stichproben	3
1.3	Merkmale und ihre Klassifikation	5
1.4	Studiendesigns	7
1.4.1	Experimente und Beobachtungsstudien	8
1.4.2	Zeitreihen	8
1.4.3	Querschnittsstudie versus Longitudinalstudie	9
1.5	Aufbereitung von univariaten Daten	9
1.5.1	Datenmatrix	10
1.5.2	Nominale und ordinale Daten	11
1.5.3	Metrische Daten	14
1.6	Lagemaße	22
1.7	Streuungsmaße	31
1.7.1	Nominale und ordinale Merkmale: Die Entropie	31
1.7.2	Metrische Merkmale	33
1.8	Schiefe versus Symmetrie	36
1.9	Quantile und abgeleitete Kennzahlen	37
1.9.1	Empirische Quantile und Quartilsabstand	37
1.9.2	Fünf-Punkte-Zusammenfassung und Boxplot	39
1.9.3	QQ-Plot (Quantildiagramm)	41
1.10	Konzentrationsmessung*	42
1.10.1	Lorenzkurve	42
1.10.2	Gini-Koeffizient	44
1.10.3	Herfindahl-Index	46
1.11	Deskriptive Korrelationsanalyse	47
1.11.1	Nominale Merkmale	47
1.11.2	Metrische Merkmale	53
1.11.3	Ordinale Merkmale	59
1.11.4	Grenzen der Korrelationsrechnung	60

1.12	Deskriptive Regressionsrechnung .....	61
1.12.1	Die Ausgleichsgerade .....	61
1.12.2	Anpassungsgüte .....	64
1.12.3	Grenzen der Regressionsrechnung .....	66
1.13	Deskriptive Zeitreihenanalyse* .....	66
1.13.1	Indexzahlen .....	67
1.13.2	Zerlegung von Zeitreihen .....	70
1.13.3	Bestimmung und Bereinigung der Trendkomponente .....	70
1.13.4	Bestimmung einer periodischen Komponente .....	71
1.14	Meilenstein .....	73
<b>2</b>	<b>Wahrscheinlichkeitsrechnung .....</b>	<b>75</b>
2.1	Grundbegriffe .....	76
2.1.1	Zufallsexperimente und Wahrscheinlichkeit .....	77
2.1.2	Chancen (Odds)* .....	84
2.1.3	Siebformel* .....	85
2.1.4	Ereignis-Algebra* .....	86
2.2	Bedingte Wahrscheinlichkeiten .....	88
2.2.1	Begriff der bedingten Wahrscheinlichkeit .....	88
2.2.2	Satz von der totalen Wahrscheinlichkeit .....	90
2.2.3	Satz von Bayes .....	92
2.3	Mehrstufige Wahrscheinlichkeitsmodelle .....	93
2.4	Unabhängige Ereignisse .....	96
2.5	Zufallsvariablen und ihre Verteilung .....	99
2.5.1	Die Verteilung einer Zufallsvariable .....	101
2.5.2	Die Verteilungsfunktion .....	102
2.5.3	Quantilfunktion und $p$ -Quantile .....	103
2.5.4	Diskrete Zufallsvariablen .....	104
2.5.5	Stetige Zufallsvariablen .....	106
2.5.6	Unabhängigkeit von Zufallsvariablen und Zufallsstichproben ..	108
2.5.7	Verteilung der Summe: Die Faltung .....	111
2.6	Erwartungswert, Varianz und Momente .....	112
2.6.1	Erwartungswert .....	112
2.6.2	Varianz .....	115
2.6.3	Momente und Transformationen von Zufallsvariablen .....	117
2.6.4	Entropie* .....	118
2.7	Diskrete Verteilungsmodelle .....	119
2.7.1	Bernoulli-Verteilung .....	119
2.7.2	Binomialverteilung .....	120
2.7.3	Hypergeometrische Verteilung .....	125
2.7.4	Geometrische Verteilung und negative Binomialverteilung .....	126
2.7.5	Poisson-Verteilung .....	127

2.8	Stetige Verteilungsmodelle .....	129
2.8.1	Stetige Gleichverteilung .....	130
2.8.2	Exponentialverteilung .....	130
2.8.3	Normalverteilung .....	131
2.8.4	Betaverteilung* .....	132
2.8.5	Gammaverteilung* .....	132
2.9	Erzeugung von Zufallszahlen* .....	134
2.10	Zufallsvektoren und ihre Verteilung.....	134
2.10.1	Verteilungsfunktion und Produktverteilung .....	135
2.10.2	Diskrete Zufallsvektoren .....	138
2.10.3	Stetige Zufallsvektoren .....	140
2.10.4	Bedingte Verteilung und Unabhängigkeit .....	143
2.10.5	Bedingte Erwartung .....	145
2.10.6	Erwartungswertvektor und Kovarianzmatrix .....	146
2.11	Grenzwertsätze und Konvergenzbegriffe .....	150
2.11.1	Das Gesetz der großen Zahlen .....	150
2.11.2	Der Hauptsatz der Statistik .....	153
2.11.3	Der zentrale Grenzwertsatz.....	153
2.11.4	Konvergenzbegriffe* .....	158
2.12	Verteilungsmodelle für Zufallsvektoren.....	159
2.12.1	Multinomialverteilung.....	159
2.12.2	Die zweidimensionale Normalverteilung .....	161
2.12.3	Multivariate Normalverteilung .....	164
2.13	Erzeugende Funktionen, Laplace-Transformierte* .....	166
2.14	Markov-Ketten* .....	169
2.14.1	Modell und Chapman-Kolmogorov-Gleichung .....	169
2.14.2	Stationäre Verteilung und Ergodensatz .....	172
2.15	Meilensteine .....	173
2.15.1	Lern- und Testfragen Block A .....	173
2.15.2	Lern- und Testfragen Block B.....	173
2.15.3	Lern- und Testfragen Block C.....	175
<b>3</b>	<b>Schließende Statistik</b> .....	<b>177</b>
3.1	Grundbegriffe .....	177
3.2	Schätzprinzipien.....	179
3.2.1	Nichtparametrische Schätzung .....	179
3.2.2	Likelihood-Schätzung .....	181
3.3	Gütekriterien für statistische Schätzer .....	190
3.3.1	Erwartungstreue.....	190
3.3.2	Konsistenz.....	193
3.3.3	Effizienz .....	194
3.3.4	Mittlerer quadratischer Fehler.....	195

3.4	Testverteilungen .....	195
3.4.1	$t$ -Verteilung .....	196
3.4.2	$\chi^2$ -Verteilung .....	196
3.4.3	$F$ -Verteilung .....	197
3.5	Konfidenzintervalle .....	197
3.5.1	Konfidenzintervall für $\mu$ .....	198
3.5.2	Konfidenzintervalle für $\sigma^2$ .....	199
3.5.3	Konfidenzintervall für $p$ .....	200
3.5.4	Konfidenzintervall für $\lambda$ (Poisson-Verteilung) .....	202
3.6	Einführung in die statistische Testtheorie .....	202
3.7	1-Stichproben-Tests .....	206
3.7.1	Motivation .....	207
3.7.2	Stichproben-Modell .....	207
3.7.3	Gauß- und $t$ -Test .....	207
3.7.4	Vorzeichentest und Binomialtest .....	218
3.8	2-Stichproben-Tests .....	220
3.8.1	Verbundene Stichproben .....	220
3.8.2	Unverbundene Stichproben .....	222
3.8.3	Wilcoxon-Test .....	226
3.8.4	2-Stichproben Binomialtest .....	228
3.9	Korrelationstests .....	229
3.9.1	Test auf Korrelation .....	230
3.9.2	Rangkorrelationstest .....	231
3.10	Lineares Regressionsmodell .....	231
3.10.1	Modell .....	231
3.10.2	Statistische Eigenschaften der KQ-Schätzer .....	233
3.10.3	Konfidenzintervalle .....	234
3.11	Multiple lineare Regression (Lineares Modell)* .....	237
3.11.1	Modell .....	237
3.11.2	KQ-Schätzung .....	238
3.11.3	Verteilungseigenschaften .....	240
3.11.4	Anwendung: Funktionsapproximation .....	241
3.12	Analyse von Kontingenztafeln .....	241
3.12.1	Vergleich diskreter Verteilungen .....	242
3.12.2	Chiquadrat-Unabhängigkeitstest .....	243
3.13	Elemente der Bayes-Statistik* .....	244
3.13.1	Grundbegriffe .....	244
3.13.2	Minimax-Prinzip .....	245
3.13.3	Bayes-Prinzip .....	246
3.14	Meilensteine .....	250
3.14.1	Lern- und Testfragen Block A .....	250
3.14.2	Lern- und Testfragen Block B .....	251

<b>Anhang A</b>	<b>Mathematik – kompakt</b> .....	253
A.1	Notationen .....	253
A.1.1	Griechische Buchstaben (Auswahl) .....	253
A.1.2	Mengen und Zahlen .....	253
A.2	Platzhalter, Variablen und Termumformungen .....	253
A.3	Punktfolgen und Konvergenz .....	255
A.3.1	Konvergenz von Folgen .....	256
A.3.2	Summen und Reihen .....	257
A.4	Ungleichungen .....	259
A.5	Funktionen .....	260
A.5.1	Spezielle Funktionen .....	261
A.5.2	Grenzwert von Funktionen .....	264
A.5.3	Stetigkeit .....	264
A.5.4	Potenzreihen* .....	265
A.6	Differenzialrechnung .....	265
A.6.1	Ableitung .....	265
A.6.2	Elastizität .....	267
A.6.3	Höhere Ableitungen .....	268
A.7	Taylorpolynom und Taylorentwicklung .....	268
A.8	Optimierung von Funktionen .....	269
A.9	Integration .....	271
A.9.1	Stammfunktion .....	272
A.9.2	Integrationsregeln .....	274
A.9.3	Uneigentliches Integral .....	274
A.10	Vektoren .....	275
A.10.1	Lineare Unabhängigkeit .....	276
A.10.2	Skalarprodukt und Norm .....	277
A.11	Matrizen .....	278
A.12	Lösung linearer Gleichungssysteme .....	282
A.12.1	Gauß-Verfahren .....	283
A.12.2	Determinanten .....	286
A.13	Funktionen mehrerer Veränderlicher .....	287
A.13.1	Partielle Differenzierbarkeit und Kettenregel .....	288
A.13.2	Lineare und quadratische Approximation, Hessematrix .....	291
A.13.3	Optimierung von Funktionen .....	292
A.13.4	Optimierung unter Nebenbedingungen .....	293
A.14	Mehrdimensionale Integration .....	295
<b>Anhang B</b>	<b>Glossar</b> .....	297
B.1	Deutsch – Englisch .....	297
B.2	Englisch – Deutsch .....	300

<b>Anhang C Tabellen</b> .....	303
C.1 Normalverteilung .....	303
C.2 <i>t</i> -Verteilung .....	305
C.3 $\chi^2$ -Verteilung .....	307
C.4 <i>F</i> -Verteilung .....	309
<b>Literatur</b> .....	319
<b>Sachverzeichnis</b> .....	321