

---

## Statistik mit Mathematica

---

Springer-Verlag Berlin Heidelberg GmbH

---

Andreas H. Jäger

Statistik mit  
**M**athematica<sup>®</sup>  
Methoden und ihre  
Anwendungen

Mit zahlreichen Abbildungen

Springer

---

Andreas H. Jäger Dipl.-Phys.  
Parkstraße 7  
97294 Burggrumbach

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

**Statistik mit Mathematica** : Methoden und ihre Anwendungen /  
Andreas Jäger. - Berlin; Heidelberg; New York; Barcelona; Budapest;  
Hongkong; London; Mailand; Paris; Santa Clara; Singapur; Tokio:  
Buch. – 1997 CD-ROM. – 1997

Additional material to this book can be downloaded from <http://extras.springer.com>

ISBN 978-3-642-63894-7 ISBN 978-3-642-59200-3 (eBook)  
DOI 10.1007/978-3-642-59200-3

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, des Vortrags, der Entnahme von Abbildungen und Tabellen, der Funksendung, der Mikroverfilmung oder der Vervielfältigung auf anderen Wegen und der Speicherung in Datenverarbeitungsanlagen, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Eine Vervielfältigung dieses Werkes oder von Teilen dieses Werkes ist auch im Einzelfall nur in den Grenzen der gesetzlichen Bestimmungen des Urheberrechtsgesetzes der Bundesrepublik Deutschland vom 9. September 1965 in der jeweils geltenden Fassung zulässig. Sie ist grundsätzlich vergütungspflichtig. Zuwiderhandlungen unterliegen den Strafbestimmungen des Urheberrechtsgesetzes.

© Springer-Verlag Berlin Heidelberg 1997

Ursprünglich erschienen bei Springer-Verlag Berlin Heidelberg New York 1997  
Softcover reprint of the hardcover 1st edition 1997

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.

Umschlaggestaltung: Künkel & Lopka, Heidelberg

Satz: Mit  $\text{T}_{\text{E}}\text{X}$  erstellte reproduktionsfertige Vorlage vom Autor

SPIN 10535471 33/3142 – 5 4 3 2 1 0 – Gedruckt auf säurefreiem Papier

Birgit gewidmet

---

# Vorwort

Statistik gehört sicherlich nicht zu den typischen Einsatzgebieten von Mathematica. Entwickelt, um zelluläre Automaten zu programmieren, wird Mathematica heute wohl insbesondere in Bereichen der Mathematik und (theoretischen) Physik eingesetzt. Als Programmiersprache in eher empirisch orientierten Wissenschaften – angefangen von der Chemie und Biologie über die Psychologie und Wirtschaftswissenschaften bis hin zur Medizin – trifft man Mathematica eher selten an. Die Bearbeitung statistischer Themen mit Mathematica schließlich fristete bisher ein relativ schattiges Dasein – durchaus zu Unrecht, wie hoffentlich als Tenor aus diesem Buch hervorgehen wird.

Statistik-Programme gibt es in der Tat sehr viele – mit SPSS und SAS sind sicherlich zwei der bekanntesten Vertreter genannt. Ein triftiger Grund, Statistik nun mit Mathematica zu betreiben, kann sicherlich nicht darin liegen, den bisherigen X Programmen eine X+1. Version hinzuzufügen. Tatsächlich liegt die Stärke von Mathematica in der Eigenschaft, (so weit wie möglich) analytisch zu rechnen. Die Gefahr, numerisch bedingten Rundungsfehlern aufzusitzen, ist damit erheblich reduziert (wenngleich keineswegs ausgeschlossen). Ein anderer Vorteil von Mathematica ist, bereits im Vorfeld des Experimentes eingesetzt werden zu können, beispielsweise bei der Entwicklung von Modellen oder mathematischen Gleichungen, bei der Versuchsplanung oder bei der Simulation von Versuchen. Ein weiterer Vorteil von Mathematica gegenüber vielen etablierten Statistik-Programmen ist, daß Mathematica eine Programmiersprache ist und damit sehr flexibel an neue Problemstellungen angepaßt

werden kann. Letztendlich ist Mathematica auch wegen seiner gut ausgebauten graphischen Möglichkeiten bekannt – und davon wird im Rahmen dieses Buches reger Gebrauch gemacht.

Das vorliegende Buch versteht sich als Gebrauchsanweisung für Mathematica im Bereich der statistischen Auswertung und graphischen Darstellung experimenteller Resultate. Was **nicht** versucht wurde, ist, ein Lehrbuch zur Statistik zu verfassen – hierzu gibt es eine Reihe guter Publikationen, die man teilweise in Kap. 4 aufgelistet findet. Ausgeklammert wurde auch der gesamte Bereich der statistischen Versuchsplanung – Literatur hierzu findet man im Publikationsverzeichnis auf der beiliegenden CD-ROM. Im ersten Kapitel wird kurz das Mathematica-System und seine Handhabung vorgestellt und erläutert, wie die beiliegenden Programme und Programm-Bibliotheken installiert werden. Im zweiten Kapitel werden einige statistische Anwendungen vorgestellt, deren Befehle in Kap. 3 näher erläutert werden. In Kap. 4 schließlich befindet sich eine unkommentierte Zusammenstellung der Statistik-Befehle, ein Verzeichnis zur Mathematica- und Statistik-Lehrbüchern, das Stichwort-Register und weitere Informationen.

Unvermeidbarerweise wird der eine oder andere Anwender eine bestimmte statistische Methode vermissen oder ein implementiertes Verfahren als etwas „exotisch“ empfinden. In der statistischen Auswertung tritt der Fall durchaus nicht selten auf, zwischen mehreren Methoden entscheiden zu müssen. Das Buch versucht, in solchen Situationen Kriterien an die Hand zu geben, die geeignete Methode ausfindig zu machen und richtig einzusetzen. Prinzipiell gilt jedoch die Faustregel, wonach statistische Verfahren zur Auswertung beispielsweise in Publikationen **immer** namentlich genannt werden müssen. Der Einsatz eines eventuell weniger etablierten Verfahrens ist dann durchaus nicht von vorne herein „illegal“, sondern u.U. lediglich Anlaß zur Diskussion. Unzulässig hingegen ist die Anwendung mehrerer Verfahren in der Hoffnung, eines werde beispielsweise eine opportun erscheinende Hypothese schon bestätigen (*fishing for the right hypothesis*).

Abschließend will ich mich bei allen Personen bedanken, die dieses Buchprojekt unterstützt und z.T. erst ermöglicht haben. An erster Stelle ist hier Herr Hermann Engesser vom Springer-Verlag Heidelberg zu nennen, ferner Frau Dorothea Glaunsinger und Frau Ulrike Drechsler, beide ebenfalls vom Springer-Verlag. Weiterhin bedanke ich mich bei Frau Silke Garotti vom Bibliographischen Institut, Mannheim, die das Projekt in der Anfangsphase mit betreut hat. Mein ganz besonderer Dank schließlich gilt meiner Frau, Birgit Pfeufer, die eine Engelsgeduld aufbrachte und mich sehr motiviert hat.

Andreas H. Jäger, März 1997



---

# Inhalt

<b>Einführung</b> .....	<b>1</b>
1.1	Wie dieses Buch zu handhaben ist ..... 2
1.2	Zum Aufbau dieses Buches ..... 3
1.3	Installation und Betrieb von Mathematica ..... 4
1.4	Starten von Mathematica-Programmen ..... 5
1.5	Mathematica-Notebooks und Zellen ..... 7
1.6	Zur Anwendung von Mathematica-packages ..... 9
1.7	Einlesen von Daten-Dateien ..... 10
1.8	Mut zur Liste: Zur Datenstruktur der Anwender-Programme .... 11
1.9	Hinweise zur Installation der Programme und der Programm-Bibliotheken ..... 14
1.10	Mathematica-Befehle und Optionen ..... 16
<b>Anwendungen</b> .....	<b>17</b>
2.1	Bevölkerungsstatistik – Beispiel für eine ausführliche deskriptive Statistik ..... 19
2.1.1	Das Notebook (nach einer Auswertung) ..... 20
2.2	Widerstandsmessung mit der Wheatstone’schen Brücke – Beispiel für deskriptive Statistik und Fehlerfortpflanzung ..... 27
2.2.1	Experimentelle Grundlagen ..... 27
2.2.2	Das Notebook (nach einer Auswertung) ..... 28
2.2.3	Einsatz des Notebooks in anderen Bereichen ..... 32
2.3	Dosis-Wirkungs-Analyse ..... 33

2.4	Digitale Blotanalyse .....	38
2.4.1	Das Notebook (nach einer Auswertung).....	40
2.5	Bestimmung der elektrischen Leitfähigkeit von Hefezellwänden mittels Elektrorotation – Beispiel für lineare Regression .....	46
2.5.1	Zur Elektrorotation .....	47
2.5.2	Das Notebook (nach einer Auswertung).....	48
2.6	Vergleich der elektrischen Leitfähigkeit dreier Kalium-Halogenid-Lösungen .....	55
2.6.1	Das Notebook (nach einer Auswertung).....	56
2.7	Arterieller Blutdruck – Beispiel für eine Zeitreihen-Analyse .....	58
2.7.1	Das Notebook (nach einer Auswertung).....	59
<b>Statistische Grundlagen .....</b>		<b>65</b>
3.0	Gemeinsame Routinen .....	67
3.1	Univariate Deskriptive Statistik .....	73
3.1.1	Schema zur deskriptiven Statistik .....	74
3.1.2	Graphische Darstellung von Daten .....	75
3.1.3	Test auf Normalität .....	85
3.1.4	Ausreißer-Behandlung .....	92
3.1.5	Daten-Reduktion .....	97
3.1.6	Fehlerfortpflanzung .....	101
3.2	Bivariate deskriptive Statistik .....	103
3.2.1	Darstellung bivariater Datensätze bei einer unabhängigen und einer abhängigen Variablen ..	104
3.2.2	Darstellung bivariater Datensätze bei zwei unabhängigen Variablen .....	113
3.3	Korrelations-Analyse .....	119
3.3.1	Einfache lineare Korrelation .....	120
3.3.2	Partielle lineare Korrelation .....	125
3.3.3	Multiple lineare Korrelation .....	126
3.4	Lineare Regression .....	128
3.4.1	Befehle zur linearen Regression .....	129
3.4.2	Approximations-Methoden .....	135
3.4.3	Das Problem überangepaßter Funktionen .....	136
3.4.4	Zur Extrapolation approximierter Gleichungen ..	139
3.4.5	Hinweis zu nichtlinearen Gleichungen .....	140

3.5	Nichtlineare Regression .....	141
3.5.1	Der Befehl .....	141
3.5.2	Levenberg-Marquardt-Verfahren .....	143
3.5.3	Konfidenz-Bereiche der geschätzten Parameter und Reparametrisierung .....	145
3.6	Zeitreihen-Analyse .....	147
3.6.1	Darstellung des Datenmaterials .....	149
3.6.2	Zeitreihen-Glättung .....	154
3.6.3	Ermittlung des Trends .....	158
3.6.4	Ermittlung der Periodizität .....	159
3.7	Hypothesen-Tests von Verteilungen .....	161
3.7.1	Schema zu den Hypothesen-Tests .....	161
3.7.2	Vergleiche eines Schätzers mit einem theoretischen Erwartungswert .....	162
3.7.3	Parametrischer Vergleich zweier Verteilungen ...	163
3.7.4	Parameterfreier Vergleich zweier Verteilungen ...	165
3.7.5	Chi <sup>2</sup> - und Kolmogoroff-Smirnow- Anpassungstest .....	168
3.8	Kontingenz-Analyse .....	171
3.8.1	Vierfelder-Test (2·2-Felder-Test) .....	172
3.8.2	2·k-Felder-Test .....	174
3.8.3	m·2-Felder-Test .....	175
3.8.4	m·k-Felder-Test .....	175
3.9	Varianzanalyse .....	177
3.9.1	Schema zur Varianzanalyse .....	178
3.9.2	Hinweise zur Datenstruktur .....	179
3.9.3	Abstandsquadrat-Summen .....	180
3.9.4	Test auf Varianzhomogenität .....	181
3.9.5	Einfache Varianzanalyse .....	183
3.9.6	H-Test nach Kruskal und Wallis .....	188
<b>Anlagen .....</b>		<b>191</b>
4.1	Schnellübersicht der Befehle .....	192
4.2	Liste der package-Namen unter MS-DOS .....	205

---

4.3	Literatur zu Mathematica .....	206
4.3.1	Lehrbücher zu Mathematica .....	206
4.3.2	Zusammenstellung weiterer Publikationen zu Anwendungen von Mathematica .....	212
4.3.3	MathSource .....	214
4.4	Zusammenstellung einiger Statistik-Lehrbücher .	215
4.5	Übersicht zum Inhalt der CD-ROM .....	219
4.6	Quellen .....	221
4.7	Index .....	229