

Heinz Holling · Günther Gediga

Statistik – Deskriptive Verfahren

BACHELORSTUDIUM PSYCHOLOGIE



Statistik – Deskriptive Verfahren

Bachelorstudium Psychologie
Statistik – Deskriptive Verfahren
von Prof. Dr. Heinz Holling und PD Dr. Günther Gediga

Herausgeber der Reihe:

Prof. Dr. Eva Bamberg, Prof. Dr. Hans-Werner Bierhoff,
Prof. Dr. Alexander Grob, Prof. Dr. Franz Petermann

Statistik – Deskriptive Verfahren

von
Heinz Holling
und Günther Gediga

HOGREFE  GÖTTINGEN · BERN · WIEN · PARIS · OXFORD · PRAG · TORONTO
CAMBRIDGE, MA · AMSTERDAM · KOPENHAGEN · STOCKHOLM

Prof. Dr. Heinz Holling, geb. 1950. 1969–1976 Studium der Mathematik, Psychologie und Soziologie in Würzburg und Berlin. 1974–1987 Wissenschaftlicher Mitarbeiter an der FU Berlin und der Universität Osnabrück. Promotion 1980 (Dr. phil.) und 1987 (Dr. rer. nat.). 1987 Habilitation. 1987–1993 Vertretungsprofessor an den Universitäten Oldenburg, Münster und Mannheim. Seit 1993 Professor für Statistik und Quantitative Methoden am Psychologischen Institut der Universität Münster.

PD Dr. Günther Gediga, geb. 1953. 1971–1979 Studium der Informatik und Mathematik in Dortmund und Osnabrück. 1986 Promotion in Psychologie. 1994 Habilitation. 1979–2000 wissenschaftlicher Mitarbeiter im Fachgebiet Methodenlehre an der Universität Osnabrück. Seit 2001 außerplanmäßiger Professor am Department of Computer Science an der Brock University in St. Catherines, Kanada, und freiberufliche Tätigkeit als Statistical Consultant. Seit 2008 akademischer Rat an der Universität Münster im Institut für Statistik und Methoden.



Informationen und Zusatzmaterialien zu diesem Buch finden Sie unter www.hogrefe.de/buecher/lehrbuecher/psychlehrbuchplus

Wichtiger Hinweis: Der Verlag hat für die Wiedergabe aller in diesem Buch enthaltenen Informationen (Programme, Verfahren, Mengen, Dosierungen, Applikationen etc.) mit Autoren bzw. Herausgebern große Mühe darauf verwandt, diese Angaben genau entsprechend dem Wissensstand bei Fertigstellung des Werkes abzudrucken. Trotz sorgfältiger Manuskriptherstellung und Korrektur des Satzes können Fehler nicht ganz ausgeschlossen werden. Autoren bzw. Herausgeber und Verlag übernehmen infolgedessen keine Verantwortung und keine daraus folgende oder sonstige Haftung, die auf irgendeine Art aus der Benutzung der in dem Werk enthaltenen Informationen oder Teilen davon entsteht. Geschützte Warennamen (Warenzeichen) werden nicht besonders kenntlich gemacht. Aus dem Fehlen eines solchen Hinweises kann also nicht geschlossen werden, dass es sich um einen freien Warennamen handelt.

© 2011 Hogrefe Verlag GmbH & Co. KG
Göttingen · Bern · Wien · Paris · Oxford · Prag · Toronto
Cambridge, MA · Amsterdam · Kopenhagen · Stockholm
Rohnsweg 25, 37085 Göttingen

<http://www.hogrefe.de>

Aktuelle Informationen · Weitere Titel zum Thema · Ergänzende Materialien

Copyright-Hinweis:

Das E-Book einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlags unzulässig und strafbar.

Der Nutzer verpflichtet sich, die Urheberrechte anzuerkennen und einzuhalten.

Umschlagabbildung: © Paul Maguire – dreamstime.com

Satz: ARThür Grafik-Design & Kunst, Weimar

Format: PDF

ISBN 978-3-8409-2134-6

Nutzungsbedingungen:

Der Erwerber erhält ein einfaches und nicht übertragbares Nutzungsrecht, das ihn zum privaten Gebrauch des E-Books und all der dazugehörigen Dateien berechtigt.

Der Inhalt dieses E-Books darf von dem Kunden vorbehaltlich abweichender zwingender gesetzlicher Regeln weder inhaltlich noch redaktionell verändert werden. Insbesondere darf er Urheberrechtsvermerke, Markenzeichen, digitale Wasserzeichen und andere Rechtsvorbehalte im abgerufenen Inhalt nicht entfernen.

Der Nutzer ist nicht berechtigt, das E-Book – auch nicht auszugsweise – anderen Personen zugänglich zu machen, insbesondere es weiterzuleiten, zu verleihen oder zu vermieten.

Das entgeltliche oder unentgeltliche Einstellen des E-Books ins Internet oder in andere Netzwerke, der Weiterverkauf und/oder jede Art der Nutzung zu kommerziellen Zwecken sind nicht zulässig.

Das Anfertigen von Vervielfältigungen, das Ausdrucken oder Speichern auf anderen Wiedergabegeräten ist nur für den persönlichen Gebrauch gestattet. Dritten darf dadurch kein Zugang ermöglicht werden.

Die Übernahme des gesamten E-Books in eine eigene Print- und/oder Online-Publikation ist nicht gestattet. Die Inhalte des E-Books dürfen nur zu privaten Zwecken und nur auszugsweise kopiert werden.

Diese Bestimmungen gelten gegebenenfalls auch für zum E-Book gehörende Audiodateien.

Inhaltsverzeichnis

1	Über dieses Buch	11
1.1	Zum Inhalt dieses Buches	13
1.2	Danksagung	15
2	Zur Relevanz der Statistik	17
2.1	Beispiel 1: Die Wahrscheinlichkeit, krank zu sein, bei einer positiven Diagnose eines recht exakten Tests.....	19
2.2	Beispiel 2: Nationale Herkunft und Schulversagen.....	21
2.3	Beispiel 3: Zulassung zum Studium in Abhängigkeit vom Geschlecht	22
2.4	Beispiel 4: Studiendauer und Einstiegsgehalt	24
3	Grundbegriffe und Aufgaben der Statistik	27
3.1	Daten	30
3.2	Stichprobenziehung	31
3.3	Deskriptive Statistik	33
3.4	Inferenzstatistik	35
	Zusammenfassung	36
	Zentrale Begriffe	37
4	Klassifikation von Daten	39
4.1	Das Skalenniveau einer Variablen.....	42
4.1.1	Ratioskalenniveau	42
4.1.2	Intervallskalenniveau	45
4.1.3	Ordinalskalenniveau	49
4.1.4	Nominalskalenniveau.....	51
4.1.5	Beziehungen zwischen den Skalenniveaus	54
4.2	Weitere Klassifikationen von Variablen.....	54
4.2.1	Quantitative und qualitative Variablen.....	54
4.2.2	Diskrete und stetige Variablen	55
4.3	Fehlende Werte.....	56
4.4	Software.....	57
4.4.1	SPSS	57

4.4.2	R.....	57
	Zusammenfassung	59
	Zentrale Begriffe	60
5	Univariate Häufigkeitsverteilungen	61
5.1	Häufigkeitsverteilungen	63
5.2	Tabellarische Darstellung von Häufigkeitsverteilungen	64
5.3	Grafische Darstellung von Häufigkeitsverteilungen	66
5.3.1	Kreis- und Balkendiagramme	67
5.3.2	Stamm-Blatt-Diagramm	69
5.3.3	Histogramm	70
5.4	Formen von Häufigkeitsverteilungen	73
5.4.1	Normalverteilung	73
5.4.2	Gipfel von Verteilungen	75
5.4.3	Schiefe von Verteilungen.....	77
5.5	Darstellung der empirischen Verteilungsfunktion.....	78
5.6	Software.....	82
5.6.1	SPSS.....	82
5.6.2	R.....	83
	Zusammenfassung	84
	Zentrale Begriffe	85
	Notation	86
6	Univariate deskriptive Statistiken.....	87
6.1	Statistiken für die zentrale Tendenz von Variablen	90
6.1.1	Arithmetisches Mittel	90
6.1.2	Median	95
6.1.3	Modus	98
6.2	Quantile	99
6.3	Statistiken für die Streuung von Variablen	102
6.3.1	Die Spannweite einer Variablen	102
6.3.2	Interquartilabstand	104
6.3.3	Abweichungsquadrat, Varianz und Standardabweichung	105
6.4	Statistiken für die Schiefe einer Verteilung.....	116
6.4.1	Lageregeln	117
6.4.2	Schiefekoeffizient für metrische Variablen	118
6.4.3	Quartilskoeffizient der Schiefe	119
6.5	Statistik für die Wölbung einer Verteilung.....	120

6.6	Software.....	123
6.6.1	SPSS.....	123
6.6.2	R.....	123
	Zusammenfassung.....	124
	Zentrale Begriffe.....	126
	Notation.....	127
7	Standardisierung und extreme Werte von Variablen	129
7.1	Standardisierung von Variablen.....	131
7.1.1	Wichtige Standardisierungen.....	134
7.1.2	z -Standardisierung normalverteilter Variablen.....	136
7.1.3	Interpretation von z -standardisierten Variablen.....	139
7.2	Extreme Werte und Ausreißer.....	140
7.3	Boxplots.....	142
7.4	Software.....	146
7.4.1	SPSS.....	146
7.4.2	R.....	146
	Zusammenfassung.....	147
	Zentrale Begriffe.....	148
	Notation.....	149
8	Bivariate deskriptive Statistik	151
8.1	Bivariate Häufigkeitsverteilungen.....	153
8.2	Grafische Darstellung bivariater Häufigkeitsverteilungen.....	158
8.3	Numerische Zusammenhangsmaße für bivariate Verteilungen.....	161
8.4	Zusammenhangsmaße für intervallskalierte Variablen.....	162
8.4.1	Kovarianz.....	162
8.4.2	Korrelation.....	167
8.4.3	Interpretation der Korrelation.....	171
8.5	Zusammenhangsmaße für ordinalskalierte Variablen.....	173
8.5.1	Rangkorrelation nach Spearman.....	176
8.5.2	Zusammenhangsmaße basierend auf Konkordanz und Diskordanz.....	179
8.6	Zusammenhangsmaße für nominalskalierte Variablen.....	187
8.6.1	Fehlerreduktionsmaße.....	193
8.6.2	Zusammenhangsmaße für die Übereinstimmung von Urteilen.....	197
8.6.3	Zusammenhangsmaße für 2×2 -Tabellen.....	199
8.7	Zusammenhangsmaße für Variablen mit unterschiedlichem Skalenniveau.....	204

8.8	Übersicht über die behandelten Zusammenhangsmaße	211
8.9	Software	214
8.9.1	SPSS	214
8.9.2	R	216
	Zusammenfassung	217
	Zentrale Begriffe	218
	Notation	222
9	Einfache lineare Regression	223
9.1	Grundlagen der einfachen linearen Regression	225
9.2	Bestimmung der Regressionsgeraden	227
9.3	Standardisierte Regressionskoeffizienten	233
9.4	Zerlegung der Variation in der einfachen linearen Regression	235
9.5	Determinationskoeffizient	239
9.6	Fehlervarianz und Standardschätzfehler	242
9.7	Weitere Statistiken in der einfachen linearen Regression	245
9.8	Voraussetzungen der einfachen linearen Regression	247
9.8.1	Residualplots	248
9.8.2	Homoskedastizität	250
9.8.3	Linearität	251
9.8.4	Ausreißer und einflussreiche Beobachtungen	252
9.9	Software	256
9.9.1	SPSS	256
9.9.2	R	256
	Zusammenfassung	258
	Zentrale Begriffe	258
	Notation	260
10	Das lineare Modell	261
10.1	Grundlagen des linearen Modells	263
10.2	Multiple Regression	265
10.2.1	Beispiel	265
10.2.2	Berechnung der Regressionskoeffizienten	270
10.2.3	Multiple Korrelation und multipler Determinationskoeffizient	271
10.3	Multiple Regression mit unkorrelierten Prädiktoren	273
10.3.1	Korrelierte vs. unkorrelierte Prädiktoren	274
10.3.2	Berechnung und Interpretation der Regressionskoeffizienten	275

10.3.3	Vergleich der Regressionskoeffizienten in der einfachen und multiplen Regression	278
10.3.4	Beitrag einzelner Prädiktoren zur Variationsaufklärung	279
10.4	Multiple Regression mit korrelierten Prädiktoren	280
10.4.1	Beispiel	281
10.4.2	Berechnung und Interpretation der Regressionskoeffizienten	287
10.4.3	Vergleich der Regressionskoeffizienten in der einfachen und multiplen Regression	292
10.4.4	Semipartielle und partielle Korrelation	296
10.4.5	Beitrag einzelner Prädiktoren zur Variationsaufklärung	297
10.4.6	Vergleich der Regressionskoeffizienten im Fall unkorrelierter und korrelierter Prädiktoren	300
10.5	Lineares Modell mit nominalskalierten Prädiktoren	302
10.5.1	Dichotome Prädiktoren	303
10.5.2	Polytome Prädiktoren	306
10.5.3	Variationsaufklärung durch nominalskalierte Prädiktoren	309
10.6	Lineares Modell mit nominalskalierten und intervallskalierten Prädiktoren	310
10.7	Nicht lineare Transformationen	311
10.8	Interaktion der Prädiktoren	315
10.8.1	Zwei binäre Prädiktoren	316
10.8.2	Ein binärer und ein metrischer Prädiktor	321
10.8.3	Zwei metrische Prädiktoren	322
10.8.4	Interaktionen von mehr als zwei Prädiktoren	324
10.9	Überprüfung der Voraussetzungen des linearen Modells	325
10.10	Software	327
10.10.1	SPSS	327
10.10.2	R	328
	Zusammenfassung	329
	Zentrale Begriffe	330
	Notation	331
	Anhang	333
	Literatur	335
	Glossar	337
	Sachregister	349
	Normalverteilungstabelle	353

Kapitel 1

Über dieses Buch

Inhaltsübersicht

1.1	Zum Inhalt dieses Buches	13
1.2	Danksagung	15

1.1 Zum Inhalt dieses Buches

Die Statistik stellt eine wichtige Grundlage für die Beantwortung von Fragestellungen der Natur-, Sozial- oder Erziehungswissenschaften dar. Beispiele für Fragestellungen sind:

- In welchem Ausmaß sind Merkmale, wie Intelligenz oder Motivation, angeboren?
- Welchen Effekt besitzen verschiedene Therapien, wie die Verhaltens- oder Gesprächstherapie, auf die Heilung psychischer Störungen, wie Magersucht, Drogenabhängigkeit oder Depression?
- In welchem Ausmaß hängt die Produktivität in Organisationen von Faktoren wie dem Führungsstil oder der Arbeitszufriedenheit ab?

Die Beantwortung solcher Fragen basiert auf den Daten empirischer Studien. Das Design solcher Studien sowie die Auswertung und Interpretation der Daten aus diesen Studien sind Gegenstand der Statistik.

Der vorliegende Band zur *Deskriptiven Statistik* ist der erste Teil eines insgesamt dreibändigen Lehrbuchs der Statistik. Der gesamte Stoff ist gegliedert in die deskriptive (beschreibende) Statistik (Band 1), Wahrscheinlichkeitstheorie und Schätzverfahren (Band 2) und statistische Testverfahren (Band 3). Mit der zu Beginn erfolgenden relativ ausführlichen Darstellung der deskriptiven Statistik soll der Einstieg in das gesamte Gebiet der Statistik erleichtert werden. Die Inhalte der beschreibenden Statistik sind für viele Studierende einfacher zu verstehen als die abstrakteren Konzepte aus der Wahrscheinlichkeitstheorie und der Inferenzstatistik (Schätzverfahren und statistische Testverfahren), die in den beiden weiteren Bänden folgen.

Grundlegendes Ziel aller drei Bände ist es, die Statistik ausführlich und verständlich darzustellen. Es wird sehr viel Wert darauf gelegt, den jeweiligen Sinn der statistischen Konzepte aufzuzeigen und die Bedeutung, die die Statistik für die Gewinnung neuer Erkenntnisse hat, vor Augen zu führen. Dazu werden einerseits einfache, übersichtliche Beispieldaten aus dem vertrauten Alltagsgeschehen, wie der Schule, verwendet und andererseits Daten „echter“ empirischer Studien. Da eine wesentliche Zielgruppe dieses Buches Studierende der Psychologie bilden, haben wir vorwiegend Beispiele aus der psychologischen Forschung gewählt. Nichtsdestoweniger ist dieser Band auch für Studierende anderer Studiengänge aus den Natur-, Sozial- und Erziehungswissenschaften als grundlegende Einführung geeignet.

Gegenstand dieses Buches sind die grundlegenden Verfahren der uni- und bivariaten deskriptiven Statistik sowie die einfache Regression. Weiterhin erfolgt eine ausführliche Behandlung des linearen Modells, das in der statistischen Praxis sehr häufig benutzt wird. Bei der Darstellung der einzelnen Verfahren wird der Interpretation der Ergebnisse eine besondere Aufmerksamkeit gewidmet.

Um die Darstellung der statistischen Konzepte nicht zu überfrachten, wird auf mathematische Ableitungen und Beweise, die für das grundlegende Verständnis nicht unbedingt erforderlich sind, verzichtet. Für besonders interessierte Leserinnen und Leser werden diese Beweise auf der folgenden Website zur Verfügung gestellt:

<http://www.hogrefe.de/buecher/lehrbuecher/psychlehrbuchplus>



Im Text wird jeweils auf diese Website hingewiesen, zusätzlich werden die entsprechenden Stellen durch das „Internetsymbol“ am Seitenrand gekennzeichnet.

Um die Inhalte dieses Lehrbuchs übersichtlich darzustellen, verwenden wir im Text verschiedene Arten von Kästen. Diese jeweils unterschiedlich gekennzeichneten Kästen enthalten:

- Beispiele
- Anwendungen von Software (SPSS und R)
- Zusammenfassungen

Wichtige Formeln werden ebenfalls durch einen Rahmen hervorgehoben, damit sie unmittelbar als solche zu erkennen sind. Bei der Einführung neuer Begriffe werden diese Begriffe im Text kursiv hervorgehoben und zudem am Seitenrand aufgeführt. Am Seitenrand befindet sich des Weiteren das bereits erwähnte Internetsymbol.

Die Vermittlung von Statistik ist heutzutage ohne die Verwendung von EDV-Programmen undenkbar und es gibt mittlerweile eine Vielzahl sehr leistungsfähiger Statistikprogramm Pakete. Im vorliegenden Lehrbuch nutzen wir die Softwarepakete SPSS und R.

Das Programmpaket IBM SPSS Statistics, kurz SPSS genannt, ist eine seit jeher sehr weit verbreitete Statistiksoftware. In vielen Organisationen, in denen Natur-, Sozial- und Erziehungswissenschaftler beschäftigt sind, wird dieses Programmpaket verwendet. SPSS hat allerdings den Nachteil, dass es kostenpflichtig ist.

Das Open-Source-Programm R ist ebenfalls weit verbreitet, hingegen kostenlos verfügbar. R ist allerdings nicht ganz so benutzerfreundlich wie SPSS. Jedoch bietet auch R eine Menüsteuerung mittels des sogenannten R-Commanders an, der für Anfänger eine einfache Benutzung der wesentlichen Funktionen dieses Programms erlaubt. Ein großer Vorteil von R besteht darin, neueste statistische Verfahren aus zuverlässigen Quellen einzubinden. Ebenso ist es mit R im Vergleich zu vielen anderen Statistikprogrammen relativ einfach möglich, eigene statistische Programme zu entwickeln.

Im vorliegenden Buch behandeln wir am Ende der Kapitel jeweils die notwendigen Schritte zur Datenanalyse mittels SPSS und R. Für SPSS sowie für R stellen wir die jeweiligen Menübefehle dar. Da der Funktionsumfang des R-Commanders begrenzt ist, werden zusätzlich die wesentlichen Befehle der R-Syntax aufgeführt. Für Studierende, die diese Programmsysteme nicht kennen, stellen wir auf der oben genannten Website eine kurze Einführung für R mit zahlreichen Literaturhinweisen zur Verfügung. Für die Einführung in die Arbeit mit SPSS möchten wir auf das Buch von Leonhart (2010) verweisen.

Auf der Website befinden sich neben den oben bereits angesprochenen Inhalten viele weitere Ressourcen zu diesem Lehrbuch. Dabei handelt es sich um die in diesem Buch benutzten Datensätze und die entsprechenden Kommandos für die Programme SPSS und R. Somit können die in diesem Buch dargestellten statistischen Analysen eigenständig nachvollzogen werden. Des Weiteren enthält die Website ergänzende Inhalte und Literaturhinweise.

1.2 Danksagung

Zur Entstehung dieses Buches haben nicht nur die Autoren beigetragen, sondern viele weitere Personen. Zunächst einmal möchten wir uns an dieser Stelle ganz herzlich bei den ca. 250 Studierenden von zwei Jahrgängen des Bachelor-Studiengangs Psychologie an der Westfälischen Wilhelms-Universität Münster bedanken. Die konstruktive Kritik dieser Leser¹ war für die Überarbeitung der vorläufigen Manuskripte sehr hilfreich. Ein besonderer Dank gebührt Frau Kathrin Gediga, M. A. Sie hatte nicht nur die Ideen für das Design vieler Grafiken, sondern hat auch ihre Umsetzung kompetent übernommen.

¹Natürlich gab es auch kritische Leserinnen. Wir werden jedoch im Folgenden immer dann, wenn beide Geschlechter gemeint sind, die überkommene männliche Form benutzen. Damit wird u. E. die Darstellung vereinfacht. Wenn wir explizit die weibliche Form nutzen, sind nur Personen weiblichen Geschlechts gemeint.

Frau cand. psych. Evelyn Alex, Herr Dipl.-Psych. Jonas Bertling, Frau Dipl.-Math. Anna Doeblér, Herr Dr. Philipp Doeblér, Frau cand. psych. Linda Fromm, Frau stud. psych. Ruth Helmich, Frau stud. psych. Inken Kirschbaum, Frau stud. psych. Miriam Miesen, Herr Dipl.-Psych. Philipp Nagels, Frau cand. psych. Marlene Pacharra, Frau stud. psych. Jana Scharfen, Frau Margret Unger und Frau Dipl.-Psych. Nina Zeuch haben eine vorläufige Endversion sorgfältig gelesen. Ihnen gebührt für die vielen sehr wertvollen Kommentare ein sehr großer Dank.

Weiterhin haben Herr Prof. Dr. Hans-Werner Bierhoff und Herr Prof. Dr. Franz Petermann ein vorläufiges Manuskript dieses Bandes aufmerksam gelesen und uns wertvolles Feedback gegeben. Dafür sei ihnen hier ganz herzlich gedankt. Schließlich gilt unser Dank Frau Dipl.-Psych. Susanne Weidinger, Herrn Dr. Michael Vogtmeier und Herrn Stefan Reins M. A. vom Hogrefe Verlag, die uns in gewohnter Weise kompetent betreut haben.

Kapitel 2

Zur Relevanz der Statistik

Inhaltsübersicht

2.1	Beispiel 1: Die Wahrscheinlichkeit, krank zu sein, bei einer positiven Diagnose eines recht exakten Tests	19
2.2	Beispiel 2: Nationale Herkunft und Schulversagen.....	21
2.3	Beispiel 3: Zulassung zum Studium in Abhängigkeit vom Geschlecht	22
2.4	Beispiel 4: Studiendauer und Einstiegsgehalt	24

Im vorliegenden Kapitel möchten wir anhand von vier Beispielen aufzeigen, warum der Einsatz statistischer Verfahren wichtig ist, d. h. umgekehrt, wie man ohne die Anwendung statistischer Kenntnisse schnell zu fehlerhaften Schlüssen kommen kann.

2.1 Beispiel 1: Die Wahrscheinlichkeit, krank zu sein, bei einer positiven Diagnose eines recht exakten Tests

Das erste, recht einfache, häufig aber verblüffende Beispiel behandelt die Interpretation von Diagnosen anhand von Testverfahren. Wir betrachten einen Test für eine Krankheit, von der in der Bevölkerung 25 von 10 000 Personen betroffen sind. Der Test hat folgende – für moderne Testverfahren nicht unübliche – Eigenschaften. Wenn eine Krankheit vorliegt, besitzt der Test eine Genauigkeit von 100 %, d. h., dieser Test liefert bei Personen, die erkrankt sind, mit einer Wahrscheinlichkeit von 100 % ein richtiges Ergebnis und mit einer Wahrscheinlichkeit von 0 % ein falsches Ergebnis. Liegt keine Krankheit vor, liefert der Test in 95.2 % aller Fälle ein korrektes Ergebnis und damit in 4.8 % aller Fälle ein falsches Resultat.

Ausgehend von den oben dargestellten Daten dürften viele Personen, die ein positives Testresultat, d. h. die Diagnose *krank*, erhalten, eine Erkrankung befürchten und dementsprechend sehr besorgt sein. Es ist auch durch viele Experimente belegt, dass die meisten Personen davon ausgehen, mit einer sehr hohen Wahrscheinlichkeit erkrankt zu sein, wenn bei solchen Testverfahren eine positive Diagnose erfolgt (Hell, Fiedler & Gigerenzer, 1993).

Aber wie groß ist tatsächlich die Wahrscheinlichkeit, bei einem positiven Testresultat erkrankt zu sein? Wie wir gleich zeigen werden, beträgt diese Wahrscheinlichkeit lediglich 5%! Viele Personen machen sich also völlig unnötig Sorgen, erkrankt zu sein, obwohl die Diagnose *positiv* vorliegt.

Bevor wir das vorliegende Beispiel eingehend analysieren, stellt sich die Frage, warum viele Personen in der vorliegenden Situation spontan dazu tendieren, bei einer positiven Diagnose mit einer viel zu hohen Wahrscheinlichkeit von einer Erkrankung auszugehen. Diese Fehleinschätzung hängt mit den ursprünglich zur Verfügung gestellten Informationen zusammen. Die beiden angegebenen Fehlerquoten des Tests mit 0 % bzw. 4.8 % sind in der Tat nicht schlecht und verleiten zu der Auffassung, dass im Fall einer positiven Diagnose maximal 4.8 % Fehler auftreten und damit 95.2 % der Personen auch erkrankt sind.

Tabelle 2.1 enthält die Daten, die der oben dargestellten Situation zugrunde liegen. Hier ist zu sehen, dass nur 25 der 500 Personen (5%), die als *krank* diagnostiziert wurden, tatsächlich auch erkrankt sind.

Tabelle 2.1: Anzahl der gesunden und kranken Personen mit der Diagnose *krank* bzw. *gesund*

Diagnose	Personen		Gesamt
	krank	gesund	
krank	25	475	500
gesund	0	9500	9500
Gesamt	25	9975	10000

Sensitivität eines Tests
Spezifität eines Tests

Die Angaben zur Genauigkeit eines Testverfahrens werden *Sensitivität* und *Spezifität* genannt.

Tabelle 2.2: Prozentualer Anteil der kranken bzw. gesunden Personen für die Gesamtheit der als *krank* diagnostizierten Personen (obere Zeile) und *gesund* diagnostizierten Personen (untere Zeile)

Diagnose	Personen		Gesamt
	krank	gesund	
krank	5%	95%	100%
gesund	0%	100%	100%

Die Sensitivität eines Tests gibt den Anteil aller erkrankten Personen an, die die richtige Diagnose, also die Diagnose *krank*, erhalten, und ist im vorliegenden Beispiel mit 100% nicht zu übertreffen (vgl. Tab. 2.3). Die Spezifität eines Tests gibt den Anteil aller richtig diagnostizierten Personen unter den gesunden Personen an. Im vorliegenden Fall ist die Spezifität mit 95.2% ebenfalls recht hoch (vgl. Tab. 2.3).

Tabelle 2.3: Prozentualer Anteil der als *krank* bzw. *gesund* diagnostizierten Personen für die Gesamtheit der kranken Personen (linke Spalte) und die Gesamtheit der gesunden Personen (rechte Spalte)

Diagnose	Personen	
	krank	gesund
krank	100%	4.8%
gesund	0%	95.2%
Gesamt	100%	100%

Die beiden Kennwerte, Sensitivität und Spezifität, stellen sogenannte bedingte Wahrscheinlichkeiten dar, das heißt zum einen die Wahrscheinlichkeit, die Diagnose *krank* zu erhalten unter der Bedingung *erkrankt zu sein*, und zum anderen die Wahrscheinlichkeit, die Diagnose *gesund* zu erhalten unter der Bedingung *gesund zu sein*. Bei solchen bedingten Wahrscheinlichkeiten werden, wie Tabelle 2.3 verdeutlicht, nicht alle Personen berücksichtigt, sondern jeweils nur Teilmengen: Für die Sensitivität wird die Teilmenge der erkrankten Personen betrachtet und für die Spezifität die Teilmenge der gesunden Personen.

Die gesuchte Wahrscheinlichkeit, *krank zu sein* unter der Bedingung *Diagnose krank*, ist ebenfalls eine bedingte Wahrscheinlichkeit. Sie beruht jedoch auf einer anderen Teilmenge der Personen. Hier werden nur solche Personen in Betracht gezogen, die die Diagnose *krank* erhalten haben (vgl. Tab. 2.2). Diese bedingte Wahrscheinlichkeit hängt jedoch nicht unmittelbar von der Sensitivität oder Spezifität eines Testverfahrens ab. Im Rahmen der Wahrscheinlichkeitstheorie (Band 2) wird gezeigt, welche Beziehungen zwischen den Ausgangsdaten, wie beispielsweise in Tabelle 2.1 dargestellt, und den verschiedenen bedingten Wahrscheinlichkeiten bestehen. Dann wird auch gezeigt, dass eine solch geringe Wahrscheinlichkeit für eine tatsächliche Erkrankung bei der erhaltenen Diagnose *krank* insbesondere dann auftreten kann, wenn der prozentuale Anteil aller erkrankten Personen relativ gering ist. (Jedoch gibt es zahlreiche eher selten auftretende Krankheiten.)

Im vorliegenden Beispiel ist der Fehlschluss, mit hoher Wahrscheinlichkeit erkrankt zu sein, schon durch die genauere Inspektion einer tabellarischen Darstellung der Ausgangsdaten zu erkennen und benötigt keine Anwendung der Wahrscheinlichkeitsrechnung. Wie wir aber später sehen werden, ist für komplexere Situationen eine wahrscheinlichkeitstheoretische Analyse unerlässlich.

2.2 Beispiel 2: Nationale Herkunft und Schulversagen

In einem zweiten Beispiel möchten wir nun zeigen, dass Fehlinterpretationen statistischer Daten häufig darauf beruhen, dass diese Daten aggregiert, das heißt auf einer bestimmten Ebene, z. B. Länder-, Schul- oder Klassenebene, zusammengefasst sind. Die hier gezeigte Fehlinterpretation statistischer Daten wird auch *ökologischer Fehlschluss* genannt und geht auf eine Studie von Robertson (1950) zurück.

ökologischer Fehlschluss

Basierend auf den Daten aus der Volkszählung des Jahres 1930 in den USA berechnete dieser Autor für 48 US-Bundesstaaten den Zusammenhang zwischen der Quote (prozentualer Anteil) an Immigranten und der

Quote an Analphabeten. Den Ergebnissen zufolge wiesen Bundesstaaten mit einem höheren Anteil an Immigranten im Mittel auch eine höhere Analphabetisierungsquote auf. Aus diesem Befund kann man nun nicht den Schluss ziehen, dass unter den Immigranten die Quote an Analphabeten höher wäre als unter den in den USA geborenen Personen bzw. dass die Wahrscheinlichkeit, ein Analphabet zu sein, für einen Immigranten höher wäre als für eine in den USA geborene Person. Tatsächlich stellte sich heraus, dass die Quote an Analphabeten unter den Immigranten geringer war als unter den in den USA geborenen Personen.

Solche Fehlinterpretationen – wie die einer höheren Wahrscheinlichkeit für Analphabetismus unter Immigranten – entstehen immer dann recht schnell, wenn Zusammenhänge von einer Aggregationsebene unzulässigerweise auf eine andere Aggregationsebene übertragen werden. Das folgende fiktive Beispiel verdeutlicht diese Situation noch plakativer.

In drei Schulen seien die in Tabelle 2.4 dargestellten Quoten der Kinder von ausländischen Mitbürgern und die Quoten der Schulabbrecher gegeben.

Tabelle 2.4: Anteile von Schülern ausländischer Herkunft und Schulabbrechern an allen Schülern bei drei Schulen

Schule	Kinder ausländischer Mitbürger	Schulabbrecher
A	40%	40%
B	30%	30%
C	20%	20%

Tatsächlich kann in allen drei Schulen die Situation vorliegen, dass kein einziges Kind ausländischer Mitbürger ein Schulabbrecher ist. Somit ist anhand dieser Darstellung ein Zusammenhang zwischen den beiden Merkmalen nicht aufrechtzuerhalten. Eine analoge Argumentation gilt auch für viele weitere Behauptungen, wie z. B. für den Zusammenhang des Anteils ausländischer Mitbürger mit der Auftretenshäufigkeit bestimmter Straftaten in verschiedenen Städten.

2.3 Beispiel 3: Zulassung zum Studium in Abhängigkeit vom Geschlecht

Auch in unserem dritten Beispiel geht es um die Aggregation von Daten. Es wird gezeigt, dass es für bestimmte Fragestellungen nötig ist, die ursprünglichen Daten in einer bestimmten Weise aufzuteilen. Es handelt

sich hier um das sogenannte *Simpson-Paradox*, das von Edward Simpson (1951) eingehend untersucht, jedoch von Statistikern wie Karl Pearson schon um 1900 dargestellt wurde.

Die im Simpson-Paradox beschriebene Situation ist mehrfach Gegenstand politischer Diskussionen und juristischer Auseinandersetzungen gewesen. Das Simpson-Paradox wurde insbesondere durch den folgenden Fall bekannt. Die University of California, Berkeley, wurde verklagt, da im Herbst 1973 die Zulassungsquote für das Graduiertenstudium für weibliche Bewerber geringer ausfiel als für männliche Bewerber. Betrachtete man jedoch die Zulassungsquoten gesondert für die einzelnen Fakultäten, ließen sich die Vorwürfe einer Diskriminierung nicht aufrechterhalten. Bei der Mehrzahl der Fakultäten hatten die weiblichen Bewerber sogar bessere Chancen als die männlichen Bewerber.

Anhand der folgenden vereinfachten, fiktiven Situation sei dieses Phänomen verdeutlicht. Betrachten wir eine Universität mit den zwei Fachbereichen *Psychologie* und *Ingenieurwissenschaften*. Insgesamt bewerben sich an dieser Universität jeweils 1000 Frauen und 1000 Männer. Für beide Fächer besteht ein Numerus clausus und die Auswahl für beide Fächer erfolgt anhand der mittleren Abiturnote. Dabei sind die mittleren Abiturnoten der weiblichen Bewerber sowohl für das Fach *Psychologie* als auch für das Fach *Ingenieurwissenschaften* höher als die entsprechenden mittleren Abiturnoten der männlichen Bewerber.

Tabelle 2.5 zeigt die fiktiven Aufnahmequoten dieser Universität ohne Aufschlüsselung für die beiden Fachbereiche. Genau die Hälfte (50 %) aller Bewerbungen wird akzeptiert. Dennoch werden lediglich 34 % der weiblichen Bewerber im Vergleich zu 66 % der männlichen Bewerber zugelassen.

Tabelle 2.5: Fiktive Aufnahmequoten einer Universität

	Bewerber		Quote
	angenommen	gesamt	
Frauen	340	1000	34 %
Männer	660	1000	66 %
Gesamt	1000	2000	50 %

Wird jedoch jeder Fachbereich für sich betrachtet (vgl. Tab. 2.6), so ändert sich das Bild. Die Aufnahmequote für die Bewerberinnen ist in beiden Fachbereichen erheblich größer.

Tabelle 2.6: Fiktive Aufnahmequoten einer Universität aufgeschlüsselt nach Fachbereichen

	männliche Bewerber		Quote
	angenommen	gesamt	
Psychologie	20	200	10 %
Ingenieurwissenschaften	640	800	80 %
Gesamt	660	1 000	66 %
	weibliche Bewerber		Quote
	angenommen	gesamt	
Psychologie	160	800	20 %
Ingenieurwissenschaften	180	200	90 %
Gesamt	340	1 000	34 %

Dass sich der Einfluss des Geschlechts auf die Gesamtzulassungsquote anders verhält als auf die einzelnen Zulassungsquoten in beiden Fachbereichen liegt darin begründet, dass sich erheblich mehr Männer als Frauen für die Ingenieurwissenschaften bewerben und hier die Zulassungsquote bedeutend höher ist als für das Fach Psychologie.

Dieses Beispiel zeigt sehr deutlich, wie eine „einfache“ Analyse von Daten zu Fehlschlüssen führen kann. Die Analyse der Daten auf der Ebene der gesamten Universität ohne Differenzierung in die einzelnen Fachbereiche führt zu gravierenden Fehlinterpretationen. Generell gilt, dass bei statistischen Analysen der Einfluss von möglichen Drittvariablen, hier die Zugehörigkeit zu einem Fachbereich, zu bedenken und gegebenenfalls mit zu analysieren ist, wie auch das folgende Beispiel deutlich macht.

2.4 Beispiel 4: Studiendauer und Einstiegsgehalt

Das letzte hier dargestellte Beispiel stammt aus dem Buch von Krämer (2009, S. 165 f.). Dieser Autor zitiert eine Studie aus dem *Handelsblatt* mit dem Titel „Methusalems machen Kasse“, in der statistisch „nachgewiesen“ wurde, dass ein längeres Studium zu einem höheren Einstiegsgehalt führt. In dieser Untersuchung wurden die Daten von Studierenden dreier Studienfächer, BWL (an der Fachhochschule, FH), Physik und Chemie, erhoben. Abbildung 2.1 zeigt die von uns diesbezüglich nachgestellten Daten.

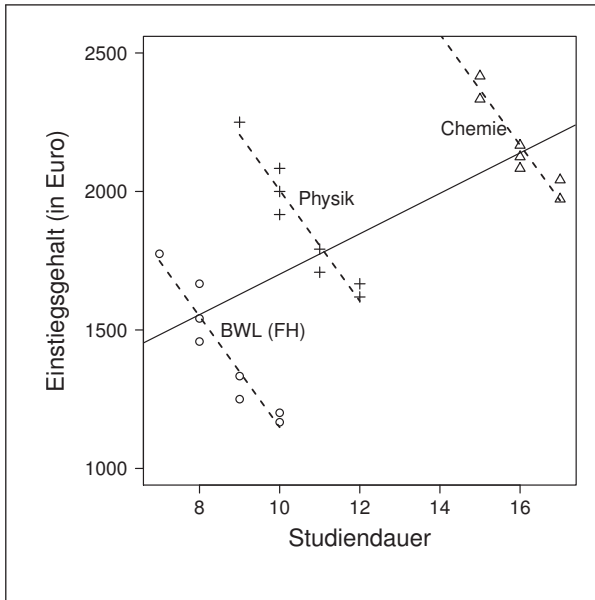


Abbildung 2.1: Studiendauer und Einstiegsgehalt für drei Studienfächer

Wie Abbildung 2.1 zeigt, steigt – wenn man die Fachzugehörigkeit außer Acht lässt – das Einstiegsgehalt im Durchschnitt mit zunehmender Studiendauer. Damit liegt ein positiver Zusammenhang zwischen diesen beiden Merkmalen vor, der allerdings wenig plausibel ist. Betrachtet man hingegen den Zusammenhang zwischen Studiendauer und Einstiegsgehalt für jedes Studienfach getrennt, resultiert ein negativer Zusammenhang, der die Realität angemessener widerspiegelt.

Der positive Zusammenhang zwischen Einstiegsgehalt und Studiendauer über die drei Fächer hinweg kommt dadurch zustande, dass das Einstiegsgehalt auch (sehr stark) von dem Studienfach abhängt und die Studiendauer mit dem Studienfach zusammenhängt. Ausgehend von Durchschnittswerten führt das kürzeste Studium der BWL an der Fachhochschule zu dem geringsten Einstiegsgehalt, das länger dauernde Studium der Physik an einer Universität zu einem höheren Einstiegsgehalt. Das Studium der Chemie, das häufig mit der Promotion abgeschlossen wird, dauert am längsten und erbringt das höchste Einstiegsgehalt.

Das zuletzt aufgeführte Beispiel zeigt besonders anschaulich Fehlinterpretationen auf, die entstehen, wenn der Zusammenhang von zwei Merkmalen betrachtet wird, ohne dass der Einfluss weiterer Merkmale berücksichtigt wird. Studienfach und Studiendauer sind hier – statistisch gesprochen – korreliert. Im Falle einer solchen *Korrelation* von Merkmalen wird nicht mehr der „reine“ Effekt eines Merkmals (*Studiendauer*

Korrelation

er) auf ein Kriterium (*Einstiegsgehalt*) erhoben, sondern der gemeinsame Einfluss zweier Merkmale (*Studiendauer* und *Studienfach*).

Im letzten Beispiel ging es lediglich um eine Drittvariable – um die Variable *Studienfach* –, deren Einfluss in einer grafischen Darstellung erkannt werden mag. In vielen Fällen liegt aber eine komplexere Situation mit mehreren Drittvariablen vor, die eine differenzierte statistische Analyse unerlässlich macht. Hierzu dient insbesondere das lineare Modell, das wir im letzten Kapitel dieses Buches darstellen.

Kapitel 3

Grundbegriffe und Aufgaben der Statistik

Inhaltsübersicht

3.1	Daten	30
3.2	Stichprobenziehung	31
3.3	Deskriptive Statistik	33
3.4	Inferenzstatistik	35
	Zusammenfassung	36
	Zentrale Begriffe	37
