

---

# WiWi klipp & klar

**Reihe herausgegeben von**

Peter Schuster, Schmalkalden, Deutschland

WiWi klipp & klar steht für verständliche Einführungen und prägnante Darstellungen aller wirtschaftswissenschaftlichen Bereiche. Jeder Band ist didaktisch aufbereitet und behandelt ein Teilgebiet der Betriebs- oder Volkswirtschaftslehre, indem alle wichtigen Kenntnisse aufgezeigt werden, die in Studium und Berufspraxis benötigt werden.

Vertiefungsfragen und Verweise auf weiterführende Literatur helfen insbesondere bei der Prüfungsvorbereitung im Studium und zum Anregen und Auffinden weiterer Informationen. Alle Autoren der Reihe sind fundierte und akademisch geschulte Kenner ihres Gebietes und liefern innovative Darstellungen – WiWi klipp & klar.

Weitere Bände in der Reihe <http://www.springer.com/series/15236>

---

Björn Christensen · Sören Christensen ·  
Martin Missong

# Statistik klipp & klar

 Springer Gabler

Björn Christensen  
Fachbereich Wirtschaft  
Fachhochschule Kiel  
Kiel, Deutschland

Sören Christensen  
Mathematisches Seminar  
Christian-Albrechts-Universität zu Kiel  
Kiel, Deutschland

Martin Missong  
Fachbereich Wirtschaftswissenschaft  
Universität Bremen  
Bremen, Deutschland

ISSN 2569-2194                      ISSN 2569-2216 (electronic)  
WiWi klipp & klar  
ISBN 978-3-658-27217-3              ISBN 978-3-658-27218-0 (eBook)  
<https://doi.org/10.1007/978-3-658-27218-0>

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

Springer Gabler

© Springer Fachmedien Wiesbaden GmbH, 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.

Springer Gabler ist ein Imprint der eingetragenen Gesellschaft Springer Fachmedien Wiesbaden GmbH und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

*Dieses Buch ist Gerd Hansen gewidmet,  
dessen charismatischer Lehrstil uns in der  
eigenen Lehre ein steter Ansporn ist.*

## Vorwort

---

Quantitative Datenanalysen sind in einer Wissensgesellschaft eine unverzichtbare Quelle der Informationsgewinnung und Wissensgenerierung. Erschienen Kenntnisse der Anwendung quantitativer Auswertungsverfahren sowie die Interpretation quantitativer Ergebnisse lange als eine vornehmlich für die wissenschaftliche Forschung bedeutsame Qualifikation, so ist heute der Umgang mit statistischen Methoden eine Schlüsselkompetenz auf dem gesamten Arbeitsmarkt über nahezu alle Berufsfelder hinweg. Dazu tragen nicht zuletzt die ubiquitäre Datenverfügbarkeit sowie die Möglichkeiten bei, Daten digital für Analyse- und Prognosezwecke nutzen zu können.

Damit entstehen sowohl zusätzliche Ausbildungsbedarfe als auch die Erfordernis einer Profilierung didaktischer Konzepte in der Vermittlung quantitativer Methodenkompetenzen. Vor dem Hintergrund einer wachsenden Zahl von methodisch Interessierten mit immer heterogeneren analytischen Vorkenntnissen gilt es beispielsweise, bei der Vermittlung statistischer Methoden zunächst soweit wie möglich den (vermeintlich) abstrakten Formalismus zurückzuhalten und den Detaillierungsgrad in der Darstellung statistischer Methoden zu reduzieren. Es erscheint sinnvoll, den anwendungsorientierten Blick auf diese Methoden noch weiter zu betonen und dabei herauszustellen, dass die quantitative Datenanalyse nicht alleine ein Teilgebiet der Angewandten Mathematik darstellt, sondern in weiten Teilen argumentativen Charakter besitzt – angefangen bei der Formulierung der Fragestellung über die Auswahl der Methoden bis hin zum Abwägen, zur Interpretation und zur Diskussion quantitativer Analyseergebnisse.

Aus der Digitalisierung von Informationen und der rasanten Entwicklung digitalen Daten- und Informationsaustausches geht jedoch nicht allein ein gesteigerter Bedarf an Kenntnissen von Verfahren der Datenanalyse hervor, sondern an anderer Stelle ein ebenfalls deutlich angestiegenes Angebot medialer Informationsmöglichkeiten über statistische Verfahren. Zu nahezu allen – grundlegenden wie elaborierten – statistischen Verfahren finden sich im weltweiten Datennetz Informationen, die sich in sehr unterschiedlichem Grad einer formalen Darstellung bedienen. Die Form der Darstellung reicht von der textgebundenen Präsentation bis hin zur audio-visuellen Vermittlung in Form von Lehrvideos. Dazu findet man stets auch Anwendungsbeispiele und Übungsaufgaben zum Selbststudium.

Diese beiden Entwicklungen – das Anwachsen sowohl eines interessierten, immer weniger fachgebundenen Interessentenkreises als auch der öffentlich verfügbaren Informationen zu quantitativen Methoden – haben unmittelbare Konsequenzen für die Entwicklung eines modernen statistischen Lehrbuchs auf Grundlagenniveau: Zum ersten muss es den „Blick von oben“ auf das statistische Methodenspektrum entfalten und damit eine Orientierung bieten, welche Verfahren der Datenanalyse in bestimmten Anwendungssituationen sinnvoll und möglich sind. Dabei soll der Fokus auf der Intention dieser Verfahren liegen und sich die Darstellung in formal-mathematischer Hinsicht – mit Verweis auf allgemein verfügbare weitergehende Informationsquellen – auf ein Minimum beschränken. Zum zweiten sollte dieser Blick aus der Vogelperspektive die Verbindungen zwischen den Teilgebieten der Statistik sowie zwischen den einzelnen Verfahren betonen, um ein tiefgehendes Verständnis zu fördern und die statistischen Methodenlehre nicht als loses System einzelner, mehr oder weniger disparater Analysemethoden

den erscheinen zu lassen. Zum dritten sollte das präsentierte Methodenspektrum weit genug gefasst sein, um dem Kreis der Leserinnen und Leser eigene praxistaugliche Anwendungen zu ermöglichen, oder – wo nötig – zu betonen, in welche Richtung der aufgezeigte „Methodenbaukasten“ für die gegebene Fragestellung ergänzt werden müsste, um zu belastbaren Ergebnissen zu gelangen.

Das vorliegende Lehrbuch versucht, diesen drei didaktischen und inhaltlichen Anforderungen ein Stück weit gerecht zu werden. Der Praxisbezug wird hergestellt zum einen durch reale Beispiele und zum anderen durch das Aufgreifen von Medieninhalten, die auf statistische Erhebungen Bezug nehmen und empirische Datenbefunde – mitunter bemerkenswert falsch – interpretieren. An vielen Stellen wird die formale Darstellung durch grafische Illustrationen ergänzt, wenn möglich sogar (nahezu) gänzlich ersetzt. Die grafische Analyse wird auch dazu verwendet, die Beziehung zwischen verschiedenen statistischen Konzepten anschaulich zu illustrieren. Der „Weitwinkel“ der inhaltlichen Perspektive unterliegt in einer Lehrbuchreihe, die als „WiWi klipp & klar“ bewusst auf einen kompakten Buchumfang zielt, klaren Begrenzungen. Den bestehenden Freiraum haben wir dazu genutzt, die Verfahren zur Abhängigkeitsmessung zu betonen, da Fragestellungen zur Interaktion beobachtbarer Größen zumindest die wirtschaftswissenschaftliche Datenanalyse dominieren. Hier entgehen wir der in einführenden statistischen Lehrbüchern mitunter zu beobachtenden Beschränkung auf den (für praktische Anwendung i. d. R. unzulänglichen) zweidimensionalen Zusammenhang zwischen statistischen Merkmalen oder Zufallsvariablen. Mit einer (rudimentären) Darstellung und Diskussion der multiplen linearen Regression und dem multiplen Linearen Modell wird zum einen die Anwendungstauglichkeit verbessert und zum anderen eine Brücke zur Ökonometrie als spezifisch wirtschaftswissenschaftlichem Anwendungsgebiet der quantitativen Methodenlehre geschlagen.

Wir hoffen, dass dieser für uns Autoren „erste Schritt“ in Richtung eines auf eine allgemeine Stärkung der „Statistical Literacy“ ausgerichteten Lehrprogramms in der Leserschaft ein weiter aufgeschlossenes Interesse an statistischem Arbeiten und der quantitativen Datenanalyse wecken wird. Diese Hoffnung wird genährt durch den Umstand, dass sich viele der hier präsentierten Beispiele und Illustrationen in unseren Vorlesungsskripten u. a. zur „Analyse von Wirtschaftsdaten“, „Statistik“, „Mathematischen Stochastik“ und zur „Einführung in die Ökonometrie“ didaktisch bewährt haben. Auch unsere Rückgriffe auf die – ebenfalls im Springer-Verlag erschienenen – Sammlungen populärwissenschaftlicher Kolumnen „Achtung: Statistik“ von Björn und Sören Christensen müssten dazu beitragen, dass auch eine heterogene, vielfältig interessierte Leserschaft die Statistik nie als „trockenes“ oder gar „abstraktes“ Wissensgebiet wahrnimmt. Damit es als ergänzende Lektüre zu „klassischen“ und stärker formal orientierten Lehrbüchern sehr gut geeignet sein kann, hat das Buch an vielen Stellen (noch) einen starken Bezug zu etablierten Lehrkonzepten und Inhalten. Schließlich würden wir uns freuen, wenn dieser Band ein tiefergehendes Verständnis quantitativer Methoden, ihrer Systematik und ihrer Zusammenhänge auch bei analytisch weniger aufgeschlossenen Interessenten fördert, den sachgerechten Umgang mit Statistik im Alltag und (evtl. nur vermeintlich) datengestützten medialen Aussagen schult und vielleicht sogar Appetit auf eigene quantitativ-empirische Forschung weckt!

Kiel, im August 2019

Björn Christensen  
Sören Christensen  
Martin Missong

# Inhaltsverzeichnis

---

<b>1</b>	<b>Einleitung: Was ist und was soll Statistik?</b> .....	1
1.1	„Statistiken kann man doch nicht trauen“ .....	2
1.2	„Statistik braucht man überall“ .....	3
1.3	Eine klassische Gliederung .....	4
1.4	Die Gliederung dieses Buches .....	6
	Literaturverzeichnis .....	7
<b>2</b>	<b>Beschreibende Statistik</b> .....	9
2.1	Datenarten .....	10
2.2	Grafische Darstellung der Daten .....	14
2.3	Messen von „Lage“ und „Streuung“ der Daten .....	19
2.4	Abhängigkeitsmessung bei qualitativen Merkmalen .....	28
2.5	Abhängigkeitsmessung bei komparativen Merkmalen .....	35
2.6	Abhängigkeitsmessung bei quantitativen Merkmalen 1: Korrelation .....	39
2.7	Abhängigkeitsmessung bei quantitativen Merkmalen 2: Regression .....	45
2.8	Übungsaufgaben .....	58
	Literaturverzeichnis .....	65
<b>3</b>	<b>Wahrscheinlichkeitsrechnung</b> .....	67
3.1	Was sind Wahrscheinlichkeiten? .....	69
3.2	Bedingte Wahrscheinlichkeiten und Unabhängigkeit .....	75
3.3	Zufallsvariablen und Wahrscheinlichkeitsverteilungen .....	80
3.4	Erwartungswert, Varianz und Co. ....	87
3.5	Urnenexperimente .....	91
3.6	Prominente Verteilungen .....	97
3.7	Hauptsätze der Wahrscheinlichkeitstheorie und Statistik .....	107
3.8	Simulation – Wenn Formeln zu kompliziert werden .....	112
3.9	Fehlschlüsse mit Wahrscheinlichkeiten .....	113
3.10	Übungsaufgaben .....	116
	Literaturverzeichnis .....	124
<b>4</b>	<b>Schließende Statistik</b> .....	125
4.1	Grundfragen der Schließenden Statistik .....	126
4.2	Punktschätzung .....	127
4.3	Konfidenzintervalle .....	132
4.4	Hypothesentests .....	140



4.5	Übungsaufgaben .....	160
	Literaturverzeichnis .....	164
<b>5</b>	<b>Anwendung: Lineares Modell und Ökonometrie .....</b>	<b>167</b>
5.1	Das bivariate Lineare Modell .....	169
5.2	Die Annahmen im Linearen Modell .....	172
5.3	Konsequenzen für den KQ-Schätzer .....	174
5.4	Analysemöglichkeiten .....	175
5.5	Übungsaufgaben .....	178
	Literaturverzeichnis .....	179
<b>A</b>	<b>Tabellen .....</b>	<b>181</b>
	<b>Sachverzeichnis .....</b>	<b>185</b>
	<b>Über die Autoren .....</b>	<b>189</b>

# Abbildungsverzeichnis

---

1.1	Ergebnis der „Sonntagsfrage“ am 29. Juni 2018 .....	5
1.2	Repräsentativität im ZDF Politbarometer .....	5
1.3	Genauigkeit im ZDF Politbarometer .....	6
1.4	Detailergebnis im ZDF Politbarometer .....	7
1.5	Teilgebiete der Statistik am Beispiel der „Sonntagsfrage“ .....	8
2.1	Skalenniveaus und Merkmalsarten .....	13
2.2	Häufigkeiten der Bundesländer .....	15
2.3	Kumulierte Häufigkeiten .....	15
2.4	Häufigkeitsdiagramm .....	16
2.5	Unbefriedigendes Häufigkeitsdiagramm .....	16
2.6	Histogramm .....	16
2.7	Alternatives Histogramm .....	16
2.8	Linkssteile bzw. rechtsschiefe Häufigkeitsverteilung .....	17
2.9	Empirische Verteilungsfunktion .....	17
2.10	Approximierende Verteilungsfunktion .....	17
2.11	Heimatland .....	18
2.12	Apothekenschwund in Deutschland .....	18
2.13	Entwicklung des Kindergelds für das erste und zweite Kind .....	19
2.14	Box-Plot .....	24
2.15	Visualisierung der Varianz .....	26
2.16	Bedingte Häufigkeiten $f(c_i b_j)$ .....	29
2.17	Bedingte Häufigkeiten im ZDF-Politbarometer .....	30
2.18	Fiktives Politbarometer-Ergebnis .....	31
2.19	Streudiagramm der Merkmale $X$ und $Y$ im Tourismus-Beispiel .....	40
2.20	Streudiagramm der Merkmale $X$ und $Y$ im Tourismus-Beispiel mit beschreibender Geraden .....	40
2.21	Quadranten der Vorzeichen bei der Korrelation .....	40
2.22	Streudiagramm zum Schokoladenkonsum und Nobelpreisen .....	43
2.23	Anscombe's Quartett .....	44
2.24	Anpassung der Geraden an die Datenpunkte .....	46
2.25	Varianz der empirischen KQ-Residuen $\hat{v}$ .....	48
2.26	Streuungszerlegung .....	49
2.27	Einfaches Bestimmtheitsmaß .....	51
2.28	Regressionsgerade im Tourismusbeispiel .....	51
2.29	Datenpunkte im Mietpreisbeispiel .....	52

2.30	Datenpunkte, geschätzte Regressionsgerade und Regressionsgleichung im Mietpreisbeispiel .....	52
2.31	Datenpunkte Gartenfläche und Mietpreis .....	56
2.32	„Chart of Doom“ .....	61
2.33	„Chart of Doom“ mit Folgezeitraum .....	65
3.1	Mengenoperationen und Verknüpfung von Ereignissen .....	72
3.2	Wahrscheinlichkeitsbaum zum Scharlach-Beispiel .....	77
3.3	Verteilungsfunktion Würfelwurf .....	82
3.4	Verteilungsfunktion der $Exp(2)$ -Verteilung .....	83
3.5	Verteilungsfunktion des Sofortaufstehers .....	86
3.6	Verteilungsfunktion des Fast-Sofortaufstehers .....	86
3.7	Verteilungsfunktion des Umdrehers .....	86
3.8	Verteilungsfunktion des Variablen .....	87
3.9	Urne mit $N$ Kugeln, wir ziehen $n$ mal .....	92
3.10	Wahrscheinlichkeitsfunktionen der Laplace-Verteilung beim Würfelwurf .....	97
3.11	Wahrscheinlichkeitsfunktionen der Bernoulli-Verteilung mit $p = 0,3$ (oben) und $p = 0,8$ (unten) .....	98
3.12	Wahrscheinlichkeitsfunktionen der Binomial-Verteilung mit $p = 0,3$ und $n = 5$ (oben) sowie $n = 10$ (unten) .....	98
3.13	Wahrscheinlichkeitsfunktionen der Geometrischen Verteilung mit $p = 0,5$ (oben) sowie $p = 0,7$ (unten) .....	100
3.14	Wahrscheinlichkeitsfunktionen der Poisson-Verteilung mit $\lambda = 0,4$ (oben) sowie $\lambda = 5$ (unten) .....	100
3.15	Dichte der Rechteckverteilung .....	102
3.16	Dichten der Exponentialverteilung zu verschiedenen Parametern $\lambda$ .....	103
3.17	Dichten der Normalverteilung zu verschiedenen Parametern .....	104
3.18	Dichte der $t$ -Verteilung für verschiedene Werte von $k$ .....	105
3.19	Dichte der $\chi^2$ -Verteilung für verschiedene Werte von $k$ .....	105
3.20	Zwei Realisierungen der mittlere Augenzahl beim Werfen eines fairen Würfels .....	108
3.21	Wahrscheinlichkeitsverteilung der mittleren Augenzahl beim Werfen mit $n = 1, 2, 4, 6$ Würfeln .....	109
3.22	Dichten der mittleren Brenndauer von $n$ Glühbirnen ( $Exp(1)$ -Verteilung) für unterschiedliche Anzahlen $n$ .....	109
3.23	log>Returns des DAX: empirische Dichte (durchgezogen) und Normalverteilungsdichte (gepunktet) mit geschätzten Parametern .....	115
3.24	Wahrscheinlichkeitsbaum zur Aufgabe 5 (Terroristen) .....	121
4.1	Wahre Umfrageergebnisse und simulierte zu einer Umfrage mit 1600 Teilnehmern .....	133
4.2	Kritischer Wert $k$ und kritischer Bereich $K$ .....	149
4.3	Fehlerwahrscheinlichkeit erster Art bei vorgegebenem $k$ .....	149

**Abbildungsverzeichnis**

4.4	Kritischer Wert bei vorgegebenem $\alpha$ .....	150
4.5	Irrtumswahrscheinlichkeiten $\alpha$ und $\beta$ .....	150
4.6	Unsinnige Ausrichtung des kritischen Bereichs .....	151
4.7	Testniveau und p-Wert .....	154
4.8	Ergebnis des Unabhängigkeitstests .....	160
5.1	Lineares Modell und KQ-Schätzung .....	170
5.2	Verschiedene Stichproben desselben Linearen Modells .....	171
5.3	Empirische Verteilung des KQ-Steigungskoeffizienten .....	172
5.4	Erwartungswert der Störgrößen .....	173
5.5	Varianz der Störgrößen .....	173
5.6	Verteilung von $\hat{\beta}_1$ , 250 Stichproben mit $n = 120$ .....	175