
Deskriptive Statistik

Reinhold Kosfeld · Hans Friedrich Eckey
Matthias Türck

Deskriptive Statistik

Grundlagen – Methoden –
Beispiele – Aufgaben

6. Auflage 2016

Reinhold Kosfeld
Universität Kassel
Kassel, Deutschland

Matthias Türk
Siemens Financial Services
München, Deutschland

Hans Friedrich Eckey
Universität Kassel
Kassel, Deutschland

ISBN 978-3-658-13639-0

ISBN 978-3-658-13640-6 (eBook)

DOI 10.1007/978-3-658-13640-6

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 1992, 2000, 2002, 2005, 2008, 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.

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier

Springer Gabler ist Teil von Springer Nature

Die eingetragene Gesellschaft ist Springer Fachmedien Wiesbaden GmbH

Vorwort zur 6. Auflage

Für die 6. Auflage wurde der Text der „Deskriptiven Statistik“ kritisch durchgesehen und überarbeitet. Hierbei wurden Druckfehler korrigiert und Ungenauigkeiten beseitigt. Die didaktisch bewährte Grundkonzeption des Lehrbuchs ist in der Neuauflage fortgeführt worden.

Dem Wunsch vieler Studierender entsprechend sind in allen Kapiteln Übungsaufgaben zu den behandelten statistischen Methoden aufgenommen worden. Die eigenständige Bearbeitung der Aufgaben ist für das Verständnis und die erfolgreiche Absolvierung von Prüfungen unverzichtbar. Zur Kontrolle der Lösungswege und Rechenschritte haben wir Musterlösungen zu den ausgewählten Aufgaben aufgenommen.

Für die Rückmeldungen von Studierenden, die die „Deskriptive Statistik“ bei ihrem Lernprozess und der Prüfungsvorbereitung begleitend verwendet haben, bedanken wir uns herzlich. Hierbei ist insbesondere Frau cand. rer. pol. Sara-Elif Kara hervorzuheben, die auf eine Reihe von „Stolpersteinen“ bei dem Lernprozess aufmerksam gemacht hat. Für Anregungen danken wir auch unseren Kollegen Prof. Dr. Wolfgang Grundmann und Dr. Vahidin Jeleskovic.

Schließlich danken wir Claudia Hasenbalg vom Verlag Springer Gabler für die gute Zusammenarbeit.

Reinhold Kosfeld

Hans-Friedrich Eckey

Matthias Türck

Vorwort zur 4. Auflage

Das Lehrbuch vermittelt anwendungsorientiert den Lehrinhalt der Deskriptiven Statistik, wie er in den Wirtschafts- und Sozialwissenschaften an Universitäten und Fachhochschulen gelehrt wird. Erfahrungen in der Lehre zeigen, dass viele Studierende die Inhalte erst verstehen, wenn sie wenig formal dargestellt sind. Insofern wurde auf manche mathematische Ableitung verzichtet und stattdessen mehr Wert auf Beispiele und die Interpretation gelegt.

Um das Auf- und Nacharbeiten zusätzlich zu vereinfachen, sind verschiedene Darstellungsweisen gewählt worden:

- Normal geschrieben ist der Text, der zum Verständnis der Inhalte unerlässlich ist. Er sollte auf jeden Fall gelesen und verarbeitet werden.
- Besonders wichtige Aussagen sind in einem Kasten dargestellt.
- Grau unterlegt sind weiterführende Erläuterungen, deren Kenntnis zwar wünschenswert, für das Verständnis aber nicht unbedingt erforderlich ist. Hierzu zählen etwa mathematische Ableitungen und Beweise.
- Das Lehrbuch enthält zahlreiche Beispiele. Diese sind durchnummeriert und ihr Ende ist durch das Zeichen "♦" angezeigt. Die Beispiele eignen sich zum selbstständigen Durchrechnen und für die Klausurvorbereitung.

Fortgeschrittene Studierende und Praktiker, beispielsweise aus der Markt- und Meinungsforschung und dem Controlling, die sich über bestimmte Kennzahlen informieren wollen, können hierfür den umfangreichen Index verwenden. Mit Hilfe des Symbolverzeichnisses lassen sich bei Vorkenntnissen auch einzelne Abschnitte im Text ohne Kenntnis der vorangegangenen Kapitel problemlos erschließen.

Das Buch ging aus den Manuskripten zu der Lehrveranstaltung "Statistik I" sowie dem Lehrbuch "Hans-Friedrich Eckey/Reinhold Kosfeld/Christian Dreger 2002: Statistik. Grundlagen – Methoden – Beispiele, 3. Aufl., Wiesbaden" hervor. Erfahrungen in der Lehre sowie neue Möglichkeiten der Computertechnik (bessere grafische Darstellungsmöglichkeiten etc.) haben eine vollständige Neubearbeitung des genannten Lehrbuchs erforderlich gemacht. Insbesondere wurden die Teile "Deskriptive Statistik" und "Induktive Statistik" getrennt, um das Lehrbuch auch für Fachhochschulen und Berufsakademien – bei denen häufig nur die beschreibende Statistik eine Pflichtveranstaltung ist – attraktiv zu gestalten. Die Zweiteilung entspricht zudem der Statistikausbildung an den meisten Universitäten.

Weiterführende Informationen zum Buch – Folien, Übungsaufgaben, Errata – sind unter

<http://www.wirtschaft.uni-kassel/Kosfeld/StatistikI.html>

hinterlegt. Angaben zu im Buchhandel erschienenen Aufgabensammlungen sind ebenso wie Nachweise von alternativen Statistik-Lehrbüchern im Literaturverzeichnis zu finden.

Ein besonderer Dank gilt Frau Iris Röttger, die einen großen Teil des Manuskripts in eine druckfertige Form umgesetzt hat. Für Anregungen sei ebenfalls Herrn Dr. Christian Dreger, Frau Dipl.-Mathem. Tatiana Samrowski und Herrn Dipl.-Vw. Günter Schumacher gedankt.

Schließlich danken wir noch Frau Susanne Kramer und Frau Annegret Eckert vom Gabler-Verlag für die gute Zusammenarbeit.

Hans-Friedrich Eckey

Reinhold Kosfeld

Matthias Türck

Inhaltsverzeichnis

| | |
|--|-------------|
| Vorwort | V |
| Abbildungsverzeichnis | XIII |
| Tabellenverzeichnis | XVI |
| Symbolverzeichnis | XVII |
| 1. Gegenstand und Grundbegriffe der Statistik | 1 |
| 1.1 Gegenstand der Statistik..... | 1 |
| 1.2 Statistische Einheiten und Massen..... | 2 |
| 1.3 Merkmale..... | 5 |
| 1.3.1 Merkmale und Merkmalsausprägungen..... | 5 |
| 1.3.2 Skalenniveau..... | 6 |
| 1.3.3 Weitere Charakteristika von Merkmalen..... | 9 |
| Aufgaben..... | 13 |
| Lösungen..... | 14 |
| 2. Datengewinnung | 17 |
| 2.1 Erhebungsarten..... | 17 |
| 2.1.1 Arten der Primärerhebung..... | 21 |
| 2.1.1.1 Befragung..... | 22 |
| 2.1.1.2 Beobachtung..... | 24 |
| 2.1.1.3 Experiment..... | 26 |
| 2.2 Auswahlverfahren..... | 28 |
| 2.2.1 Willkürliche Auswahl..... | 29 |
| 2.2.2 Zufallsauswahl (Stichprobe im engeren Sinne)..... | 30 |
| 2.2.3 Bewusste Auswahl (Beurteilungsstichprobe)..... | 31 |
| Aufgaben..... | 33 |
| Lösungen..... | 35 |
| 3. Univariate Häufigkeitsverteilungen | 37 |
| 3.1 Häufigkeitsverteilungen bei unklassierten Merkmalen..... | 38 |
| 3.1.1 Absolute und relative Häufigkeiten..... | 38 |
| 3.1.2 Darstellung unklassierter Häufigkeitsverteilungen..... | 41 |
| 3.2 Klassierte Häufigkeitsverteilung..... | 46 |
| 3.2.1 Absolute und relative Klassenhäufigkeiten..... | 46 |
| 3.2.2 Darstellung klassierter Häufigkeitsverteilungen..... | 48 |

| | | |
|-----------|--|------------|
| 3.3 | Empirische Verteilungsfunktion | 53 |
| 3.3.1 | Empirische Verteilungsfunktion bei unklassierten Merkmalen | 54 |
| 3.3.2 | Empirische Verteilungsfunktion bei klassierten Merkmalen | 56 |
| | Aufgaben | 62 |
| | Lösungen..... | 63 |
| 4. | Lagemaße..... | 67 |
| 4.1 | Modus (Modalwert) | 68 |
| 4.2 | Median..... | 72 |
| 4.3 | Arithmetisches Mittel..... | 79 |
| 4.4 | Geometrisches Mittel | 89 |
| 4.5 | Harmonisches Mittel | 93 |
| 4.6 | Quantile | 98 |
| | Aufgaben | 103 |
| | Lösungen | 105 |
| 5. | Streuung und Schiefe..... | 109 |
| 5.1 | Streuungsmaße | 110 |
| 5.1.1 | Überblick..... | 110 |
| 5.1.2 | Spannweite | 111 |
| 5.1.3 | Quartilsabstand und Boxplot..... | 113 |
| 5.1.4 | Mittlere absolute Abweichung | 117 |
| 5.1.5 | Varianz und Standardabweichung..... | 120 |
| 5.1.6 | Variationskoeffizient | 130 |
| 5.2 | Schiefe..... | 132 |
| | Aufgaben | 138 |
| | Lösungen | 139 |
| 6. | Konzentration | 145 |
| 6.1 | Konzentrationsbegriff | 145 |
| 6.2 | Konzentrationsraten und Konzentrationskurve..... | 147 |
| 6.3 | Herfindahl-Index und Exponentialindex..... | 151 |
| 6.4 | Rosenbluth-Index | 156 |
| 6.5 | Lorenzkurve und Gini-Koeffizient..... | 160 |
| | Aufgaben | 169 |
| | Lösungen | 170 |
| 7. | Bivariate Häufigkeitsverteilungen und Kontingenz | 175 |
| 7.1 | Gemeinsame Häufigkeiten und Randhäufigkeiten..... | 175 |

| | | |
|------------|--|------------|
| 7.2 | Bedingte Häufigkeiten und statistische Unabhängigkeit..... | 181 |
| 7.3 | Bei Unabhängigkeit zu erwartende Häufigkeiten | 186 |
| 7.4 | Phi-Koeffizient und Kontingenzkoeffizient | 190 |
| | Aufgaben..... | 194 |
| | Lösungen..... | 196 |
| 8. | Korrelationsanalyse | 199 |
| 8.1 | Streuungsdiagramm..... | 199 |
| 8.2 | Kovarianz | 202 |
| 8.3 | Korrelationskoeffizient nach Fechner | 208 |
| 8.4 | Korrelationskoeffizient nach Bravais und Pearson | 213 |
| 8.5 | Rangkorrelationskoeffizient nach Spearman..... | 216 |
| 8.6 | Kausalität und Scheinkorrelation | 218 |
| | Aufgaben..... | 220 |
| | Lösungen..... | 222 |
| 9. | Regressionsanalyse | 225 |
| 9.1 | Regressionsfunktion..... | 225 |
| 9.2 | Methode der kleinsten Quadrate und Regressionsgerade..... | 227 |
| 9.3 | Eigenschaften der Kleinst-Quadrate-Regression..... | 236 |
| 9.4 | Güte der Anpassung | 241 |
| | Aufgaben..... | 245 |
| | Lösungen..... | 247 |
| 10. | Verhältnis- und Indexzahlen | 251 |
| 10.1 | Verhältniszahlen..... | 251 |
| 10.2 | Indexzahlen | 254 |
| 10.2.1 | Preisindizes | 255 |
| 10.2.2 | Mengenindizes, Wertindex und Deflationierung..... | 269 |
| 10.2.3 | Umbasieren und Verketteten..... | 274 |
| | Aufgaben | 277 |
| | Lösungen | 278 |
| 11. | Zeitreihenanalyse..... | 281 |
| 11.1 | Aufgaben und Ziele | 281 |
| 11.2 | Zeitreihenzerlegung..... | 282 |
| 11.2.1 | Überblick..... | 282 |
| 11.2.2 | Trend und glatte Komponente..... | 286 |
| 11.2.3 | Gleitende Durchschnitte..... | 293 |

| | |
|--|------------|
| 11.2.4 Saisonkomponente und Saisonbereinigung..... | 297 |
| 11.3 Exponentielle Glättung..... | 306 |
| Aufgaben | 313 |
| Lösungen | 314 |
| Anhang: Das Summenzeichen..... | 321 |
| Literaturverzeichnis..... | 327 |
| Index | 331 |

Abbildungsverzeichnis

| | | |
|-----------------|--|-----|
| Abbildung 1.1: | Bereiche der Statistik..... | 2 |
| Abbildung 1.2: | Arten von statistischen Einheiten | 3 |
| Abbildung 1.3: | Statistische Einheit und Merkmale..... | 5 |
| Abbildung 1.4: | Arten von Messskalen | 6 |
| Abbildung 1.5: | Schema zur Bestimmung des Skalenniveaus..... | 8 |
| Abbildung 1.6: | Skalenniveaus und Rechenoperationen | 9 |
| Abbildung 1.7: | Skalenniveau und Merkmalsart | 10 |
| Abbildung 1.8: | Skalen, Merkmalsarten und weitere Charakteristika | 11 |
| Abbildung 1.9: | Charakterisierung von Merkmalen nach der Anzahl der Merkmalsausprägungen | 12 |
| Abbildung 2.1: | Vorteile der Primär- und Sekundärerhebung..... | 17 |
| Abbildung 2.2: | Externe Quellen für eine Sekundärerhebung..... | 18 |
| Abbildung 2.3: | Arten der Primärerhebung | 21 |
| Abbildung 2.4: | Arten der Befragung | 22 |
| Abbildung 2.5: | Mögliche Beobachtungsfehler | 25 |
| Abbildung 2.6: | Experimenteller Versuchsaufbau..... | 27 |
| Abbildung 2.7: | Auswahlverfahren..... | 28 |
| Abbildung 2.8: | Wichtige Auswahlverfahren | 29 |
| Abbildung 3.1: | Erhebung von Merkmalswerten..... | 37 |
| Abbildung 3.2: | Darstellungsarten für unklassierte Häufigkeitsverteilungen | 41 |
| Abbildung 3.3: | Säulen- und Stabdiagramm..... | 42 |
| Abbildung 3.4: | Kreisdiagramm | 44 |
| Abbildung 3.5: | Abgrenzung von Klassen..... | 47 |
| Abbildung 3.6: | Darstellung einer klassierten Häufigkeitsverteilung..... | 48 |
| Abbildung 3.7: | Histogramm | 50 |
| Abbildung 3.8: | Grafische Darstellung von kumulierten Häufigkeiten | 54 |
| Abbildung 3.9: | Empirische Verteilungsfunktion..... | 55 |
| Abbildung 3.10: | Approximative empirische Verteilungsfunktion | 57 |
| Abbildung 3.11: | Ausschnitt der approx. empirischen Verteilungsfunktion..... | 59 |
| Abbildung 4.1: | Lagemaße, die sich in Abhängigkeit vom Skalenniveau berechnen lassen | 67 |
| Abbildung 4.2: | Uni- und bimodale Verteilung..... | 58 |
| Abbildung 4.3: | Bestimmung des Median durch Interpolation ($p = 3$)..... | 76 |
| Abbildung 4.4: | Anwendungen verschiedener Durchschnittswerte..... | 80 |
| Abbildung 4.5: | p-Quantil..... | 99 |
| Abbildung 4.6: | Bestimmung des p-Quantils durch Interpolation..... | 102 |
| Abbildung 5.1: | Rechts- und linksschiefe Verteilung..... | 110 |

| | | |
|-----------------|---|-----|
| Abbildung 5.2: | Verschiedene Streuungsmaße | 110 |
| Abbildung 5.3: | Spannweite bei Häufigkeitsverteilungen..... | 111 |
| Abbildung 5.4: | Quartilsabstand..... | 113 |
| Abbildung 5.5: | Boxplot..... | 115 |
| Abbildung 5.6: | Varianzberechnung bei klassierten Merkmalen | 126 |
| Abbildung 5.7: | Lineare Transformation ($b = 1$)..... | 129 |
| Abbildung 5.8: | Absolute und relative Streuungsmaße..... | 131 |
| Abbildung 5.9: | Symmetrische Verteilungen | 132 |
| Abbildung 5.10: | Asymmetrische Verteilungen..... | 133 |
| Abbildung 5.11: | Quartilsbereiche | 135 |
| Abbildung 5.12: | Quartilsbereiche bei einer rechtsschiefen Ver- teilung..... | 136 |
| Abbildung 6.1: | Relative und absolute Konzentration | 145 |
| Abbildung 6.2: | Verschiedene Konzentrationsmaße | 146 |
| Abbildung 6.3: | Konzentrationskurve | 148 |
| Abbildung 6.4: | Konzentrationskurve bei minimaler und maxi- maler Konzentration..... | 148 |
| Abbildung 6.5: | Konzentrationskurve und Dekonzentrationsfläche | 156 |
| Abbildung 6.6: | Dekonzentrationsfläche bei minimaler Konzentration..... | 157 |
| Abbildung 6.7: | Dekonzentrationsfläche bei maximaler Konzentration..... | 157 |
| Abbildung 6.8: | Dekonzentrationsfläche A..... | 158 |
| Abbildung 6.9: | Lorenzkurve bei drei Klassen..... | 162 |
| Abbildung 6.10: | Lorenzkurve bei drei Klassen..... | 167 |
| Abbildung 7.1: | Übersicht verschiedener Zusammenhangsmaße | 190 |
| Abbildung 8.1: | Streuungsdiagramm bei einem positiven oder negativen Zusammenhang..... | 199 |
| Abbildung 8.2: | Streuungsdiagramm, wenn kein Zusammenhang besteht | 200 |
| Abbildung 8.3: | Streuungsdiagramm mit verschobenen Koordi- natensystem | 202 |
| Abbildung 8.4: | Grafische Veranschaulichung der Kreuzprodukte | 204 |
| Abbildung 8.5: | Streuungsdiagramm bei perfekter positiver Korrelation | 212 |
| Abbildung 8.6: | Parabolische Beziehung | 213 |
| Abbildung 8.7: | Mögliche Kausalitäten | 218 |
| Abbildung 8.8: | Scheinkorrelation | 219 |
| Abbildung 9.1: | Unterstellte Beziehung in der Regression | 225 |
| Abbildung 9.2: | Funktionswerte..... | 226 |
| Abbildung 9.3: | Fehlergrößen | 227 |
| Abbildung 9.4: | Quadrierte Fehlergrößen | 228 |
| Abbildung 9.5: | Regressionsgerade mit und ohne Ausreißer | 238 |
| Abbildung 9.6: | Abweichungsquadrate | 241 |
| Abbildung 10.1: | Verschiedene Verhältniszahlen..... | 251 |

| | | |
|------------------|---|-----|
| Abbildung 10.2: | Verschiedene Indizes | 254 |
| Abbildung 10.3: | Übersicht über die behandelten Preisindizes | 288 |
| Abbildung 10.4: | Unterstellter Warenkorb beim Laspeyres-Preis- index | 265 |
| Abbildung 10.5: | Unterstellte Warenkörbe beim Paasche-Preis- index | 265 |
| Abbildung 10.6: | Preisindizes bei substitutionellem Verbraucher- verhalten | 227 |
| Abbildung 11.1: | Unterschied zwischen Zeitreihenanalyse und Ökonometrie | 281 |
| Abbildung 11.2: | Aufgaben der Zeitreihenanalyse | 282 |
| Abbildung 11.3: | Zeitreihendiagramm der systematischen Kompo- nenten | 283 |
| Abbildung 11.4: | Komponenten einer Zeitreihe | 284 |
| Abbildung 11.5: | Einfache Modelle der Zeitreihenzerlegung | 284 |
| Abbildung 11.6: | Gleitende Durchschnittsbildung bei ungerader Ordnung ($p = 3$) | 293 |
| Abbildung 11.7: | Berechnung der p -gliedrigen gleitenden Durch- schnitte | 290 |
| Abbildung 11.8: | Saisonausschläge mit konstanter Amplitude | 298 |
| Abbildung 11.9: | Saisonausschläge mit zunehmender Amplitude | 302 |
| Abbildung 11.10: | Gewichtsfaktor α und Gewichtungsfunktion | 311 |
| Abbildung 11.12: | Reaktionen der Vorhersage bei verschiedenen α | 309 |

Tabellenverzeichnis

| | | |
|---------------|--|-----|
| Tabelle 1.1: | Fragen zur Abgrenzung statistischer Massen..... | 3 |
| Tabelle 2.1: | Vor- und Nachteile der verschiedenen Befragungsformen..... | 23 |
| Tabelle 3.1: | Häufigkeitstabelle bei unklassierten Merkmalen | 41 |
| Tabelle 3.2: | Häufigkeitstabelle bei klassierten Merkmalen | 49 |
| Tabelle 4.1: | Spezielle Quantile | 99 |
| Tabelle 6.1: | Marktbeherrschende Stellung nach § 19 GWB..... | 147 |
| Tabelle 6.2: | US-Fusionsrichtlinien | 152 |
| Tabelle 7.1: | Zweidimensionale Häufigkeitsverteilung für absolute Häufigkeiten | 177 |
| Tabelle 7.2: | Zweidimensionale Häufigkeitsverteilung für relative Häufigkeiten..... | 179 |
| Tabelle 7.3: | Bedingte Häufigkeitsverteilungen von Y gegeben x_j | 182 |
| Tabelle 7.4: | Bedingte Häufigkeitsverteilungen von X gegeben y_k | 182 |
| Tabelle 7.5: | Vierfeldertafel | 163 |
| Tabelle 9.1: | Verschiedene synonyme Begriffe für die Variablen..... | 225 |
| Tabelle 10.1: | Typisierung von Verhältniszahlen | 251 |
| Tabelle 11.1: | Bedeutung des Gewichtungsfaktors α | 311 |

Symbolverzeichnis

- Lateinische Buchstaben

| | |
|-----------|---|
| a | Regressionskoeffizient, der den Ordinatenabschnitt angibt; autonomes Glied der Regressionsgleichung $\hat{y}_i = a + b \cdot x_i$ Konstanter Betrag, um den alle Merkmalswerte bei der Lineartransformation erhöht werden: $y_i = a + b \cdot x_i$ Absolute Häufigkeit bei einer Vierfeldertafel, die die Anzahl statistischer Einheiten mit einer Eins bei beiden Merkmalen angibt |
| \hat{a} | Kleinst-Quadrate-Schätzer für den Regressionskoeffizienten a |
| a_i | i-ter Einzelwert des Merkmals A reelle Zahl bei der Schwarzschen Ungleichung |
| A_t | Abgangsmasse in der Periode von t-1 bis t |
| av_o | Unterer "adjacent value" beim Boxplot |
| av_u | Oberer "adjacent value" beim Boxplot |
| b | Regressionskoeffizient, der das Steigungsmaß angibt Faktor, um den alle Merkmalswerte bei der Lineartransformation erhöht werden: $y_i = a + b \cdot x_i$ Absolute Häufigkeit bei einer Vierfeldertafel, die die Anzahl statistischer Einheiten mit einer Eins beim Merkmal X und einer Null beim Merkmal Y angibt |
| \hat{b} | Kleinst-Quadrate-Schätzer für den Regressionskoeffizienten b |
| b_i | i-ter Einzelwert des Merkmals B reelle Zahl bei der Schwarzschen Ungleichung |
| b_k | Klassenbreite der k-ten Klasse |
| B_t | Bestand zum Zeitpunkt t |
| B_{t-1} | Bestand zum Zeitpunkt t-1 |
| c | Anzahl der Ausprägungen bzw. Klassen eines Merkmals Y bei zweidimensionalen Häufigkeiten (c steht für columns) Absolute Häufigkeit bei einer Vierfeldertafel, die die Anzahl statistischer Einheiten mit einer Null beim Merkmal X und einer Eins beim Merkmal Y angibt |
| c_i | Merkmalanteil des i-ten Merkmalsträgers an der Merkmalssumme |

| | |
|-----------------|--|
| C_i | i-te Konzentrationsrate |
| c_t | Konjunkturkomponente |
| d | Absolute Häufigkeit bei einer Vierfeldertafel, die die Anzahl statistischer Einheiten mit einer Null bei beiden Merkmalen angibt |
| D | Modus (Modalwert) |
| \bar{d} | Durchschnittliche jährliche Abweichung der unnormierten Saisonkomponente s_j^* von null |
| d_i | Differenz zwischen dem Rang von x_i und dem Rang von y_i |
| d_{ij} | Trendbereinigte Zeitreihe mit dem Jahr i und dem Jahresteil j |
| d_k | Häufigkeitsdichte der k -ten Klasse |
| d_t | Trendbereinigte Zeitreihenwerte |
| $d_{\bar{x}}$ | Durchschnittliche Abweichung vom arithmetischen Mittel |
| $d_{\tilde{x}}$ | Durchschnittliche Abweichung vom Median |
| e | Eulersche Zahl |
| E | Exponentialindex |
| e_t | Prognosefehler bei der exponentiellen Glättung |
| F | Flächeninhalt des Kreises Fläche zwischen der Gleichverteilungsgeraden G und der Lorenzkurve L |
| F' | Fläche unterhalb der Lorenzkurve L |
| f_i | Wachstumsfaktor der i -ten Zeiteinheit |
| F_j | Flächeninhalt des j -ten Kreissektors |
| F'_k | Fläche unterhalb der Lorenzkurve von der k -ten Klasse |
| g | Schiefekoeffizient (Momentschiefe) Ganzzahliger Teil des Produktes $n \cdot p$ bei der Quantilsbestimmung |
| G | Gleichverteilungsgerade |
| g_i | i -tes Gewicht bei der Berechnung des Herfindahl-Index |
| g_j | j -tes Gewicht zur Berechnung eines Preisindex |
| g_{j0}^L | j -tes Gewicht zur Berechnung eines Preisindex nach Laspeyres; Ausgabenanteil des j -ten Gutes in der Basisperiode 0 |
| $g_{j,0t}^P$ | j -tes Gewicht zur Berechnung eines Preisindex nach Paasche in der arithmetischen Mittelwertform; fiktiver Ausgabenanteil |

| | |
|-------------------------|--|
| | des j-ten Gutes |
| g_{jt}^P | j-tes Gewicht zur Berechnung eines Preisindex nach Paasche in der harmonischen Mittelwertform; Ausgabenanteil des j-ten Gutes in der Berichtsperiode t |
| g_P | Schiefemaß von Pearson |
| g_Q | Quartilskoeffizient |
| g_t | Glatte Komponente (= Trend und Konjunkturkomponente) |
| g_Y | Schiefemaß von Yule |
| H | Harmonisches Mittel |
| h_j | Relative Häufigkeit der j-ten Merkmalsausprägung von X |
| H_j | Kumulierte relative Häufigkeit bis zur j-ten Merkmalsausprägung |
| $h_{j\bullet}$ | Relative Randhäufigkeit der j-ten Merkmalsausprägung von X |
| h_{jk} | Gemeinsame relative Häufigkeit der Merkmalskombination (x_j, y_k) |
| \tilde{h}_{jk} | Bei Unabhängigkeit von X und Y zu <i>erwartende</i> gemeinsame relative Häufigkeit |
| h_k | Relative Häufigkeit der k-ten Klasse |
| H_k | Kumulierte relative Häufigkeit bis zur k-ten Klasse |
| $h_{\bullet k}$ | Relative Randhäufigkeit der k-ten Merkmalsausprägung von Y |
| $H(x)$ | Empirische Verteilungsfunktion bzw. approximative empirische Verteilungsfunktion |
| $h(x_j^* y_k^*)$ | Relative Häufigkeit von x_j^* , die sich für ein gegebenes y_k^* ergibt bzw. relative Häufigkeit von x_j^* unter der Bedingung y_k^* |
| $h(y_k^* x_j^*)$ | Relative Häufigkeit von y_k^* , die sich für ein gegebenes x_j^* ergibt bzw. relative Häufigkeit von y_k^* unter der Bedingung x_j^* |
| $\text{int}(n \cdot p)$ | Integer-Funktion gibt den ganzzahligen Teil von $n \cdot p$ aus |
| j^* | j^* = Laufindex, der mit der maximalen Häufigkeit korrespondiert: $h_{j^*} = \max\{h_j\}$ bzw. $n_{j^*} = \max\{n_j\}$ Merkmalsausprägung, bei der die relativen kumulierten Häufigkeiten 0,5 erreichen oder durchlaufen Merkmalsausprägung, bei der die relativen kumulierten Häufigkeiten p erreichen oder durchlaufen |

| | |
|----------------|--|
| k bzw. k_j | Anzahl der verfügbaren trendbereinigten Zeitreihenwerte eines Jahresteils j |
| K | Kontingenzkoeffizient |
| k^* | Modalklasse Medianklasse Quantilsklasse |
| K^* | Normierter Kontingenzkoeffizient |
| K_α | Verallgemeinerung des Herfindahl-Index |
| K_H | Herfindahl-Index |
| K'_H | Approximativer Herfindahl-Index |
| lg | Dekadischer Logarithmus |
| ln | Natürlicher Logarithmus |
| \log | Allgemeiner Logarithmus |
| L | Lorenzkurve |
| m | Anzahl der realisierten Ausprägungen des Merkmals X Anzahl der Güter bei der Indexberechnung Anzahl der Teilintervalle |
| M | Median Durchschnittswert einer Verhältniszahl |
| m_k | Klassenmitte der k -ten Klasse |
| m_{k^*} | Klassenmitte der Modalklasse |
| $m_k(\bar{x})$ | k -tes zentrales Moment; Verallgemeinerung der Varianz |
| m_t | Trend |
| $m_{t-1,t}$ | Meßzahl mit gleitender Basis |
| m_{0t} | Meßzahl der Berichtszeit t zur Basiszeit 0 |
| m_{0t}^j | Meßzahl der Berichtszeit t zur Basiszeit 0 für das j -te Gut |
| $\max\{h_j\}$ | Maximaler Wert von h_j |
| $\min\{h_j\}$ | Minimaler Wert von h_j |
| MSE | Mittlerer quadratischer Fehler (mean square error) |
| n | Anzahl der Merkmalsträger |
| n^+ | Die Anzahl der Beobachtungen, die auf einen positiven Zusammenhang hinweisen, also im ersten und dritten Quadranten liegen |

| | |
|------------------|---|
| n^- | Die Anzahl der Beobachtungen, die auf einen negativen Zusammenhang hinweisen, also im zweiten und vierten Quadranten liegen |
| n_j | Absolute Häufigkeit der j-ten Merkmalsausprägung des Merkmals X |
| $n_{j\bullet}$ | Absolute Randhäufigkeit der j-ten Merkmalsausprägung von X |
| n_{jk} | Gemeinsame absolute Häufigkeiten der Merkmalskombination (x_j, y_k) |
| \tilde{n}_{jk} | Bei Unabhängigkeit von X und Y zu <i>erwartende</i> absolute gemeinsame Häufigkeit |
| n_k | Absolute Klassenhäufigkeit der k-ten Klasse |
| $n_{\bullet k}$ | Absolute Randhäufigkeit der k-ten Merkmalsausprägung von Y |
| N_j | Kumulierte absolute Häufigkeit bis zur j-ten Merkmalsausprägung |
| N_k | Kumulierte absolute Häufigkeit bis zur k-ten Klasse |
| N_t | Nominale Größe in der Periode t |
| p | Anzahl der Klassen |
| | Prozentpunkt bei Quantilen |
| | Preis eines Gutes |
| | Ordnung des gleitenden Durchschnitts |
| | Anzahl der Jahresabschnitte (z. B. bei Monatsdaten ist $p = 12$) |
| P | Preisindex |
| P_{jt} | Preis des j-ten Gutes in der Berichtsperiode t |
| P_{j0} | Preis des j-ten Gutes in der Basisperiode 0 |
| \bar{P}_t | Durchschnittspreis aller m Güter in der Berichtsperiode t |
| \bar{P}_0 | Durchschnittspreis aller m Güter in der Basisperiode 0 |
| P_{kt}^* | Umbasierte Indexreihe der Berichtsperiode t zur Basisperiode k |
| P_{0t}^* | Verkettete Indexreihe der Berichtsperiode t zur Basisperiode 0 |
| P_{0t}^C | Preisindex nach Carli |
| P_{0t}^D | Preisindex nach Dutot |
| P_{0t}^{impl} | Impliziter Preisindex |

| | |
|---------------------|--|
| P_{0t}^L | Preisindex nach Laspeyres |
| P_{0t}^P | Preisindex nach Paasche |
| Q | Quartilsabstand Mengenindex |
| Q_E | Durch den Regressionsansatz erklärte Abweichungsquadratsumme |
| q_{jt} | Menge des j-ten Gutes in der Berichtsperiode t |
| q_{j0} | Menge des j-ten Gutes in der Basisperiode 0 |
| q_k | Merkmalsanteil der k-ten Klasse |
| Q_k | Kumulierter Merkmalsanteil bis zur k-ten Klasse |
| Q_R | Nicht erklärte Abweichungsquadratsumme (Residualabweichungsquadratsumme) |
| Q_T | Gesamte (totale) Abweichungsquadratsumme |
| Q_{0t}^L | Mengenindex nach Laspeyres |
| Q_{0t}^P | Mengenindex nach Paasche |
| $Q(a, b)$ | Minimierungsfunktion in der Regression |
| r | Anzahl der Ausprägungen bzw. Klassen eines Merkmals X bei zweidimensionalen Häufigkeiten (r steht für rows) Korrelationskoeffizient nach Bravais und Pearson Radius (halber Durchmesser) eines Kreises |
| R | Spannweite Gini-Koeffizient |
| R^* | Normierter Gini-Koeffizient |
| R^2 | Determinationskoeffizient (Bestimmtheitsmaß) |
| r_F | Korrelationskoeffizient nach Fechner |
| R_{\max} | Maximaler Wert des Gini-Koeffizienten |
| r_s | Rangkorrelationskoeffizient nach Spearman |
| R_t | Reale Größe in der Periode t |
| $r(x_i)$ | Rang von x_i |
| $R(x)$ | Resthäufigkeitsfunktion |
| $r(y_i)$ | Rang von y_i |
| s | Standardabweichung; Quadratwurzel aus der Varianz s^2 |
| $s = \min \{r, c\}$ | Minimale Anzahl der Zeilen und Spalten einer Kreuztabelle |

| | |
|--------------------|---|
| s, s_1, s_2 | Teilstrecken |
| S | Gesamtstrecke |
| | Merkmalssumme |
| s_j | normierte Saisonkomponente |
| s_j^* | unnormierte Saisonkomponente |
| S_k | Merkmalssumme der k-ten Klasse |
| s_t | Saisonkomponente |
| s^2 | Varianz |
| s_{ext}^2 | Externe Varianz (Varianz zwischen den Klassen) |
| s_{int}^2 | Interne Varianz (Varianz innerhalb den Klassen) |
| s_k^2 | Varianz der k-ten Klasse |
| s_m^2 | Approximative Varianz bei klassierten Daten |
| $s_{\hat{u}}^2$ | Residualvarianz |
| $s_{\hat{u}x}$ | Kovarianz zwischen X und den Residuen |
| s_x^2 | Varianz von X |
| s_{xy} | Kovarianz zwischen X und Y |
| $s_{x'y'}$ | Kovarianz zwischen den lineartransformierten Merkmalen X und Y |
| s_y^2 | Varianz von Y |
| | Varianz der lineartransformierten x-Werte: $y_i = a + b \cdot x_i$ |
| $s_{\hat{y}}^2$ | Varianz der Regressionswerte |
| $\text{sgn}(x_i)$ | Signum-Funktion gibt bei einer negativen Zahl eine -1 und bei einer positiven Zahl eine 1 zurück; Beispiele: $\text{sgn}(-4) = -1$ und $\text{sgn}(3) = 1$ |
| t, t_1, t_2 | Benötigte Zeit für Teilstrecken |
| T | Benötigte Zeit insgesamt |
| $\tan \alpha$ | Tangens vom Winkel α |
| u_i | i-te Fehlergröße in der Regression: $u_i = y_i - \hat{y}_i$ |
| \hat{u}_i | i-te geschätzte Fehlergröße (Residuum) in der Regression |
| $\bar{\hat{u}}_i$ | Arithmetisches Mittel der Residuen \hat{u}_i |
| u_t | Restkomponente bei einer Zeitreihe |
| v | Variationskoeffizient |
| v, v_1, v_2 | Geschwindigkeit für Teilstrecken |

| | |
|-----------------|--|
| V | Durchschnittsgeschwindigkeit |
| v_{jk} | Abweichungen der beobachteten von den bei Unabhängigkeit erwarteten Häufigkeiten; prozentual interpretierbar |
| w | Geometrisches Mittel |
| w_i | Wachstumsrate der i -ten Zeiteinheit |
| w_o | Oberer "whisker" beim Boxplott |
| w_u | Unterer "whisker" beim Boxplott |
| W_{0t} | Wertindex |
| x_i | Beobachtungswert (Einzelwert) ($i = 1, \dots, n$) |
| $x^{(i)}$ | i -ter Wert von in <i>aufsteigender</i> Reihenfolge geordneten Beobachtungswerten |
| $x^{(i)}$ | i -ter Wert von in <i>absteigender</i> Reihenfolge geordneten Beobachtungswerten |
| x_j^* | j -te Ausprägung eines Merkmals X |
| x_{k-1}' | Untergrenze der k -ten Klasse |
| x_k' | Obergrenze der k -ten Klasse |
| \tilde{x} | Median, Zentralwert |
| \tilde{x}_p | p -Quantil |
| \bar{x} | arithmetisches Mittel |
| \bar{x}_k | Klassenmittelwert; arithmetisches Mittel der k -ten Klasse |
| \bar{x}_m | Approximatives arithmetisches Mittel bei klassierten Daten |
| y_i | Lineartransformierte x -Werte: $y_i = a + b \cdot x_i$ i -te Ausprägung des Merkmals Y |
| y_k^* | k -te Ausprägung eines Merkmals Y bei zweidimensionalen Häufigkeitsverteilungen |
| y_t | t -ter Zeitreihenwert |
| (y_t) | Zeitreihe ($t = 1, \dots, n$) |
| y_{ij}^* | Saisonbereinigter Zeitreihenwert |
| \hat{y}_i | Regressionswert |
| \hat{y}_t | Prognosewert für die Periode t |
| \bar{y} | Arithmetisches Mittel der lineartransformierten x -Werte Arithmetisches Mittel des Merkmals Y |
| $\bar{\hat{y}}$ | Arithmetisches Mittel der Regressionswerte \hat{y}_i |

| | |
|---------------|---|
| \bar{y}_t^p | p-gliedriger Durchschnitt |
| Z | Im Hintergrund stehende Größe (Faktor) |
| Z_t | Zugangsmasse in der Periode von t-1 bis t |

- griechische Buchstaben

| | |
|------------|--|
| α | Winkel Gewichtungsfaktor bei der exponentiellen Glättung |
| α_j | Winkel des j-ten Kreissektors beim Kreisdiagramm |
| ∂ | Operator beim Differenzieren |
| ϕ | Phi-Koeffizient |
| π | Konstante zur Berechnung des Flächeninhaltes von Kreisen ($\pi = 3,14$) |
| \prod | Produktoperator; Beispiel: $\prod_{i=1}^n x_i = x_1 \cdot x_2 \cdot \dots \cdot x_n$ |
| \sum | Summenzeichen (siehe Anhang) |
| χ^2 | quadratische Kontingenz (Chi-Quadrat) |

- Sonstige Symbole

| | |
|--------------------|---|
| $[x'_{k-1}; x'_k)$ | Intervall zwischen x'_{k-1} und x'_k , wobei x'_{k-1} zum Intervall gehört (eckige Klammer) und x'_k nicht (runde Klammer). |
| $(x'_{k-1}; x'_k]$ | Intervall zwischen x'_{k-1} und x'_k , wobei x'_k zum Intervall gehört (eckige Klammer) und x'_{k-1} nicht (runde Klammer). |
| \neq | ungleich |
| \approx | näherungsweise |
| $ b $ | Absolutwert von b (Minus- werden durch Pluszeichen ersetzt; Beispiele: $ -3 = 3$, $ 3 = 3$) |
| ∞ | Unendlichkeitszeichen |