

Masterclass

mc

Markus Janczyk
Roland Pfister

Inferenz- statistik verstehen

Von A wie Signifikanztest
bis Z wie Konfidenzintervall



Springer Spektrum

Masterclass

mc

Markus Janczyk
Roland Pfister

Inferenz- statistik verstehen

Von A wie Signifikanztest
bis Z wie Konfidenzintervall



Springer Spektrum

Inferenzstatistik verstehen

Markus Janczyk · Roland Pfister

Inferenzstatistik verstehen

Von A wie Signifikanztest
bis Z wie Konfidenzintervall

 Springer Spektrum

Markus Janczyk
Roland Pfister
Institut für Psychologie
Lehrstuhl für Psychologie III
Julius-Maximilians-Universität Würzburg
Würzburg, Deutschland

ISBN 978-3-642-34824-2

ISBN 978-3-642-34825-9 (eBook)

DOI 10.1007/978-3-642-34825-9

Mathematics Subject Classification (2010): 62-01, 62F03, 62F10, 62J05, 62J10, 62K10, 62K15, 62P10, 62P15, 62P25

Springer Spektrum

© Springer-Verlag Berlin Heidelberg 2013

Dieses 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.

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier.

Springer Spektrum ist eine Marke von Springer DE. Springer DE ist Teil der Fachverlagsgruppe Springer Science+Business Media
www.springer-spektrum.de

Vorwort

Empirische Daten sind der Grundstein wissenschaftlichen Fortschritts in vielen natur- und sozialwissenschaftlichen Disziplinen. Solche Daten gewinnen, auswerten und die Ergebnisse korrekt interpretieren zu können stellt daher auch eine zentrale Fähigkeit für viele Bereiche der Psychologie, der Soziologie, der Erziehungswissenschaften oder auch der Neurowissenschaften dar und der Erwerb dieser Fähigkeiten ist daher ein zentraler und nicht zu vernachlässigender Teil der universitären Ausbildung in den genannten Fächern. Das vorliegende Buch richtet sich in erster Linie an Studierende dieser Fächer, aber auch an fortgeschrittene Empiriker, die in den genannten Fächern arbeiten und lehren.

Ziel dieses Buches ist es, einen verständlichen Zugang zu häufig benutzten Verfahren der Inferenzstatistik zu bieten und die Leserinnen und Leser dabei zu unterstützen, diese angemessen in ihren eigenen Arbeiten anzuwenden. Besonderes Augenmerk haben wir dabei auf die grundsätzliche Logik des inferenzstatistischen Vorgehens gelegt, mit der Absicht, ein tieferes Verständnis zu ermöglichen und Zusammenhänge zwischen verschiedenen Verfahren zu betonen. Die praktische Anwendung der Verfahren mit SPSS und R sowie Beispiele zur Darstellung der Ergebnisse bilden den Abschluss einzelner Kapitel.

Vorausgesetzt werden einige Grundlagen der deskriptiven Statistik, die im ersten Kapitel kurz zusammengefasst sind. Die drei folgenden Kapitel sind den Grundlagen der Inferenzstatistik gewidmet und liefern das nötige Rüstzeug, um jede Art inferenzstatistischer Tests verstehen zu können. In den verbleibenden Kapiteln werden dann die wichtigsten Verfahren beschrieben, von t -Tests über Varianzanalysen bis hin zu Korrelation und Regression. Hierbei kommen auch wichtige zeitgemäße Themen zur Sprache, wie etwa Konfidenzintervalle, Effektstärken und die Power von Signifikanztests.

Im Fließtext des Buches haben wir uns bemüht, auf unnötiges Formelwerk und weiterführende Details zu verzichten, und haben dabei manche mathematische Unschärfe in Kauf genommen. Für besonders interessierte Leserinnen und Leser haben wir jedoch an einigen Stellen grau unterlegte Abschnitte in den Text eingefügt; in diesen finden sich formale Hinweise und Herleitungen, Hintergrundwissen sowie andere wissenswerte Informationen. Ergänzende Textdokumente, sowie Beispiel-

datensätze und kommentierte Auswertungsskripte finden sich zudem ergänzend als Online-Material unter <http://www.springer.com/springer+vs/psychologie/book/978-3-642-34824-2>. Zum Einsatz in Lehrveranstaltungen, stehen dort außerdem die Abbildungen dieses Buches zur Verfügung. Wenngleich das vorliegende Buch in deutscher Sprache verfasst ist, haben wir die in Publikationen übliche, und auch von vielen Statistikprogrammen genutzte, Schreibweise eines Dezimalpunktes durchgängig verwendet (anstelle des deutschen Dezimalkommata). Darüber hinaus haben wir uns, dort wo wir Ergebnisse beispielhaft berichten, an den Richtlinien der American Psychological Association orientiert.

Zum Abschluss möchten wir denjenigen Personen danken, ohne die dieses Buch zwar schneller, aber sicherlich auch weitaus weniger verständlich erschienen wäre. Dies sind vor allem Katharina Schwarz, Thomas Göb und Stefan Friedrich, die mit ihrer sorgfältigen Durchsicht so einige Inkonsistenzen und Probleme aufgedeckt haben, und somit einen wichtigen Anteil an der Endfassung dieses Buches haben, sowie zahlreiche Studentinnen und Studenten, die einzelne Kapitel auf ihre Verständlichkeit hin getestet haben. Gar nicht erschienen wäre dieses Buch vermutlich ohne die vermittelnden Eingriffe von Alexander Heinemann, der dadurch erbitterte Grabenkämpfe der beiden Autoren im Verlauf so mancher Diskussion verhindert hat; so viel Zeit musste stets sein. Wilfried Kunde gewährte uns an seinem Lehrstuhl die nötigen Ressourcen und denkbar besten Arbeitsbedingungen zur Realisierung des Projekts. Unser besonderer Dank gilt ferner den Mitarbeiterinnen und Mitarbeitern des Springer Verlags, die dieses Projekt betreut haben: Alice Blanck, Agnes Herrmann, Clemens Heine und Niels Peter Thomas. Schließlich möchten wir uns bei Dieter Heyer (Halle), Gisela Müller-Plath (Berlin) und Rainer Scheuchenspflug (Würzburg) bedanken, die uns nachhaltig für Statistik und Forschungsmethoden begeistert haben.

Würzburg,
November 2012

Markus Janczyk
Roland Pfister

Inhaltsverzeichnis

1	Einführung und deskriptive Statistik	1
1.1	Wichtige mathematische Schreibweisen	1
1.1.1	Das Summenzeichen	1
1.1.2	Mengentheoretische Schreibweisen	3
1.1.3	Variablentransformationen	4
1.2	Deskriptive Statistik	4
1.2.1	Arithmetisches Mittel	5
1.2.2	Stichprobenvarianz	5
1.3	Vorbemerkungen zu R und SPSS	7
1.3.1	Das Statistikpaket R	7
1.3.2	Das Statistikpaket SPSS	9
2	Grundlagen der Inferenzstatistik 1: Zufallsvariablen	11
2.1	Diskrete Zufallsvariablen	11
2.1.1	Das Konzept der Zufallsvariablen	11
2.1.2	Diskrete Wahrscheinlichkeitsfunktionen	13
2.1.3	Erwartungswert diskreter Zufallsvariablen	15
2.1.4	Varianz diskreter Zufallsvariablen	16
2.2	Stetige Zufallsvariablen	16
2.2.1	Dichtefunktionen stetiger Zufallsvariablen	17
2.2.2	Erwartungswert und Varianz stetiger Zufallsvariablen	18
2.3	Die Normalverteilung	19
3	Grundlagen der Inferenzstatistik 2: Population und Parameterschätzung	23
3.1	Stichprobe vs. Population	23
3.1.1	Das Problem	23
3.1.2	Stichprobenstatistiken und Populationsparameter	25
3.2	Der Populationsparameter μ	25
3.3	Gütekriterien von Parameterschätzern	31
3.4	Der Populationsparameter σ_X^2	32

4	Hypothesentests und Signifikanz	35
4.1	Inhaltliche und statistische Hypothesen	35
4.1.1	Klassifikation statistischer Hypothesen	36
4.1.2	Alternativ- und Nullhypothese	36
4.2	Die Idee des Signifikanztests	37
4.2.1	Eine fiktive Situation.....	37
4.2.2	...und die Logik des Signifikanztests	38
5	Unterschiedshypothesen bei maximal zwei Gruppen: t-Tests	43
5.1	Der t -Test für unabhängige Stichproben	43
5.1.1	Die t -Verteilung	45
5.1.2	Entscheidungen auf Basis kritischer t -Werte	46
5.1.3	Entscheidungen auf Basis von p -Werten	50
5.1.4	Voraussetzungen beim t -Test für zwei Stichproben	52
5.1.5	Testen ungerichteter Hypothesen	53
5.1.6	Eine Beispielrechnung	54
5.2	Der Einstichproben- t -Test	55
5.3	Der t -Test für zwei abhängige Stichproben	56
5.4	Zusammenfassung des Vorgehens	57
5.5	Beispiele und Übungen	58
5.5.1	t -Tests mit R	58
5.5.2	t -Tests mit SPSS	62
6	Konfidenzintervalle	65
6.1	Die allgemeine Form von Konfidenzintervallen	65
6.2	Konfidenzintervalle für Mittelwerte	66
6.2.1	Rechnerische Durchführung	66
6.2.2	Einflussfaktoren auf die Größe von Konfidenzintervallen ...	67
6.2.3	Interpretation von Konfidenzintervallen	67
6.3	Konfidenzintervalle und Hypothesentests	69
6.3.1	Der Einstichprobenfall	69
6.3.2	Konfidenzintervalle für Mittelwertsdifferenzen	70
6.4	Konfidenzintervalle bei abhängigen Stichproben	71
6.5	Gegenüberstellung der Konfidenzintervalle	72
6.6	Konfidenzintervalle mit R und SPSS	73
7	Fehlertypen, Effektstärken und Power	77
7.1	Fehlentscheidungen in der Inferenzstatistik	78
7.2	Effektstärken	79
7.3	Power und Fehler 2. Art	80
7.4	Optimaler Stichprobenumfang	84
7.5	Das Zusammenspiel der Fehler 1. und 2. Art	86
7.6	Beispiele und Übungen	87
7.6.1	Effektstärken mit R	87
7.6.2	Effektstärken mit SPSS	89

8	Einfaktorielle Varianzanalyse	91
8.1	Grundlagen der einfaktoriellen Varianzanalyse	92
8.1.1	Warum Varianzanalysen? α -Inflation und α -Adjustierung	92
8.1.2	Zentrale Begriffe und Hypothesen	93
8.1.3	Die Idee der Varianzanalyse	94
8.2	Rechnerische Durchführung	97
8.2.1	Quadratsummenzerlegung	98
8.2.2	Mittlere Quadratsummen	100
8.2.3	Der F -Bruch und die F -Verteilung	101
8.2.4	Die Entscheidungsregel	103
8.3	Effektstärken und Power	104
8.3.1	Effektstärken in der Population	104
8.3.2	Schätzung der Effektstärke aus den Stichproben	105
8.3.3	Power der Varianzanalyse	106
8.4	Kontraste in der Varianzanalyse	107
8.4.1	Beispiele für Kontraste	108
8.4.2	Die Hypothesen und das Schätzen von Kontrasten	109
8.4.3	Testen von Kontrasthypotesen	110
8.5	Abschließende Bemerkungen	111
8.5.1	Ergebnisdarstellung	111
8.5.2	Konfidenzintervalle	111
8.5.3	Voraussetzungsverletzungen	112
8.5.4	Eine andere Betrachtung der einfaktoriellen Varianzanalyse	113
8.6	Beispiele und Übungen	114
8.6.1	Einfaktorielle Varianzanalysen mit R	114
8.6.2	Einfaktorielle Varianzanalysen mit SPSS	116
9	Mehrfaktorielle Varianzanalyse	119
9.1	Grundlagen der zweifaktoriellen Varianzanalyse	120
9.1.1	Haupteffekte	121
9.1.2	Interaktionseffekte	123
9.1.3	Interpretation und grafische Darstellung	124
9.2	Rechnerische Durchführung	125
9.2.1	Quadratsummenzerlegung	127
9.2.2	Mittlere Quadratsummen	128
9.2.3	Die F -Brüche und die Entscheidungsregeln	129
9.3	Abschließende Bemerkungen	129
9.3.1	Vorteile mehrfaktorieller Varianzanalysen	129
9.3.2	Effektstärken	130
9.3.3	Interpretation und Ergebnisdarstellung	131
9.3.4	Varianzanalysen mit mehr als zwei Faktoren	131
9.4	Beispiele und Übungen	132
9.4.1	Zweifaktorielle Varianzanalysen mit R	132
9.4.2	Zweifaktorielle Varianzanalysen mit SPSS	134

10	Varianzanalyse mit Messwiederholung	137
10.1	Ein einfacher Zugang zur Messwiederholungsanalyse	137
10.2	Der Umgang mit interindividuellen Unterschieden	139
10.3	Abhängige vs. unabhängige Stichproben	140
10.4	Abschließende Bemerkungen	141
10.4.1	Voraussetzungsverletzungen	141
10.4.2	Effektstärken	142
10.4.3	Konfidenzintervalle	142
10.4.4	Mehrfaktorielle Varianzanalyse mit Messwiederholung	143
10.4.5	Gemischte Varianzanalyse	143
10.5	Beispiele und Übungen	144
10.5.1	Varianzanalysen mit Messwiederholung mit R	144
10.5.2	Varianzanalysen mit Messwiederholung mit SPSS	145
11	Zusammenhangshypothesen: Korrelation und Regression	149
11.1	Zusammenhang und Abhängigkeit von Variablen	149
11.2	Die Produkt-Moment-Korrelation	150
11.2.1	Rechnerische Bestimmung	151
11.2.2	Inferenzstatistik der Korrelation	154
11.3	Einfache lineare Regression	155
11.3.1	Bestimmung einer Regressionsgeraden	156
11.3.2	Inferenzstatistik der einfachen linearen Regression	158
11.4	Abschließende Bemerkungen	159
11.4.1	Multiple lineare Regression	160
11.4.2	Partialkorrelation	160
11.5	Beispiele und Übungen	162
11.5.1	Korrelation und Regression mit R	162
11.5.2	Korrelation und Regression mit SPSS	163
12	Anmerkungen zum Schluss	167
	Literaturverzeichnis	169
	Sachregister	173

Kapitel 1

Einführung und deskriptive Statistik

Wissenschaftlicher Fortschritt beruht in vielen Disziplinen auf kreativen Ideen und Fragestellungen, die nur auf Basis empirischer Daten beantwortet werden können. Nach der Datenerhebung steht man daher i.d.R. vor einem großen Datensatz, den sog. Rohdaten; um sich einen Überblick über sie zu verschaffen und sie übersichtlich darzustellen, bedient man sich der Methoden der **deskriptiven Statistik** („beschreibende Statistik“). Mit diesen Methoden können die wesentlichen Aspekte der Daten einfach und anschaulich ausgedrückt werden. Allerdings lassen sich mit ihnen lediglich Aussagen über den erhobenen Datensatz machen, z.B. über den Mittelwert einer Variablen in der untersuchten Stichprobe. Für Aussagen die über die Stichprobe hinausgehen sind hingegen Methoden der Inferenzstatistik („schließende Statistik“) gefragt.

Im diesem Kapitel behandeln wir Grundlagen der deskriptiven Statistik, die für das Verständnis der weiteren Kapitel unerlässlich sind. Nach einer Klärung relevanter Begriffe und Schreibweisen werden wir die wichtigsten und gebräuchlichsten deskriptiven Maße – Mittelwert und Varianz – und ihre Berechnung vorstellen.

1.1 Wichtige mathematische Schreibweisen

1.1.1 Das Summenzeichen

Das Summenzeichen spielt in diesem Buch (und auch allgemein in der Statistik) eine wichtige Rolle und wird als abkürzende Schreibweise für eine Summe verwendet. Als Beispiel betrachten wir die Daten von fünf Untersuchungseinheiten – z.B. Personen – auf einer Variablen X . Variablen werden im Folgenden als Großbuchstaben geschrieben; einzelne, konkrete Werte dieser Variablen als Kleinbuchstaben:

$$x_1 = 3, \quad x_2 = 5, \quad x_3 = 1, \quad x_4 = 0, \quad x_5 = -1 .$$

Die Bestandteile des **Summenzeichens** sind in Abb. 1.1 dargestellt. Die Summe $x_1 + x_2 + x_3 + x_4 + x_5 = 8$ lässt sich damit auch kurz schreiben als:

$$\sum_{i=1}^5 x_i = 8 .$$

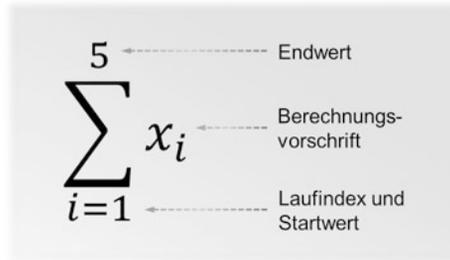


Abb. 1.1 Die Bestandteile des Summenzeichens. Dabei bezeichnet i den *Laufindex*; grundsätzlich kann jedoch jeder Buchstabe für den Laufindex verwendet werden. Die Zahl 1 ist der *Startwert*, 5 der *Endwert* und der Ausdruck nach dem Summenzeichen (in diesem Fall x_i) stellt die *Berechnungsvorschrift* dar

Relevanter wird der Gebrauch des Summenzeichens dann, wenn nicht nur fünf, sondern unendlich viele Werte betrachtet werden oder zumindest ihre genaue Anzahl vorher nicht bekannt ist; dies kommt in der Statistik häufig vor. In Formel 1.1 betrachten wir n Werte, wobei die exakte Anzahl n nicht weiter spezifiziert ist. Dies ist z.B. dann der Fall, wenn der Stichprobenumfang (noch) nicht bekannt ist, und Formeln daher allgemein gehalten werden sollen:

$$x_1 + x_2 + \dots + x_{n-1} + x_n = \sum_{i=1}^n x_i . \quad (1.1)$$

Einige wichtige *Rechenregeln mit dem Summenzeichen* sind:

- Sei a eine konstante reelle Zahl (Kurzschreibweise: $a \in \mathbb{R}$), dann gilt:

$$\sum_{i=1}^n ax_i = a \sum_{i=1}^n x_i .$$

- Sei $a \in \mathbb{R}$, dann gilt:

$$\sum_{i=1}^n a = \underbrace{a + a + a + \dots + a}_{n\text{-mal}} = na .$$

- Seien X und Y zwei Variablen, dann gilt:

$$\sum_{i=1}^n (x_i + y_i) = \sum_{i=1}^n x_i + \sum_{i=1}^n y_i .$$

- Ähnliches gilt für die Multiplikation aber i.A. *nicht*:

$$\sum_{i=1}^n (x_i \cdot y_i) \neq \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i .$$

- Leicht zu übersehen, und daher eine häufige Fehlerquelle, ist die Position des Exponenten:

$$\sum_{i=1}^n x_i^2 \neq \left(\sum_{i=1}^n x_i \right)^2 .$$

Beispielsweise ergeben sich für die fünf Werte des Eingangsbeispiels

$$\sum_{i=1}^5 (x_i^2) = 36 \quad \text{und} \quad \left(\sum_{i=1}^5 x_i \right)^2 = 64 .$$

1.1.2 Mengentheoretische Schreibweisen

In diesem Abschnitt geben wir einen kurzen Überblick über Schreibweisen aus der Mengenlehre, die wir im Folgenden verwenden werden.

Mengen bestehen aus einer „Ansammlung von Elementen“. Dabei werden die **Mengen** an sich üblicherweise mit Großbuchstaben und ihre Elemente mit Kleinbuchstaben bezeichnet. Ist z.B. a ein **Element der Menge** A , wird dies ausgedrückt durch $a \in A$ („ a ist Element von A “). Die Menge der natürlichen Zahlen wird mit \mathbb{N} , und die der reellen Zahlen mit \mathbb{R} bezeichnet. Werden bestimmte Mengen explizit eingeführt, so werden ihre Elemente in geschweiften Klammern geschrieben. Für die Menge A der Zahlen 1, 2, 3 und 4 kann man schreiben:

$$A = \{1, 2, 3, 4\} = \{1, \dots, 4\} = \{x | x \in \mathbb{N} \text{ und } 1 \leq x \leq 4\} .$$

Alle drei Varianten bezeichnen die gleiche Menge; die Lesart der dritten Variante ist „ A ist die Menge aller Zahlen x , für die gilt: x ist Element der natürlichen Zahlen und liegt zwischen 1 und 4 (jeweils einschließlich)“.

Gelegentlich werden wir eine Mengenschreibweise in Zusammenhang mit dem Summenzeichen verwenden. Sollen alle Elemente der Menge A aufsummiert werden, kann dies geschrieben werden als:

$$\sum_{a \in A} a = 10 .$$

Darüber hinaus werden wir gelegentlich eine bestimmte Berechnung für verschiedene Gruppen bzw. Bedingungen durchführen. Dies lässt sich durch den sog. „All(o)quantor“ \forall ausdrücken. Das Zeichen \forall wird dabei gelesen als „für alle“. Als Beispiel betrachten wir drei Werte x_1 , x_2 und x_3 , zu denen wir jeweils 10 addieren wollen. Dies lässt sich schreiben als:

$$x'_i = x_i + 10 \quad \forall i \in \{1, 2, 3\}.$$

1.1.3 Variablentransformationen

Manchmal bildet man aus einer Variablen eine neue Variable, indem die Werte mit einem bestimmten Faktor multipliziert werden und/oder bestimmte Werte zu ihnen addiert werden. Würde man also bspw. aus der Variablen X eine neue Variable $aX + b$ bilden, so wäre dies so zu verstehen: Wir nehmen jeden einzelnen Wert x_i , multiplizieren ihn mit a und addieren dann b dazu. Dies nennt man eine **lineare Transformation**.

Ganz ähnlich kann man natürlich auch neue Variablen generieren, indem zwei bestehende Variablen miteinander kombiniert werden, z.B. durch Addition oder Multiplikation. Gibt es bspw. zwei Variablen X und Y , dann bezeichnet die neue Variable $Z = X + Y$ die Addition der jeweiligen Messwertpaare: $z_i = x_i + y_i$.

1.2 Deskriptive Statistik

Der erste Schritt einer jeden Datenauswertung ist die Beschreibung der Daten; dies ist die Aufgabe der deskriptiven Statistik. Zwei besonders wichtige Maße sind hierbei das *arithmetische Mittel* und die (*Stichproben-*)*Varianz* und beide werden in den folgenden Kapiteln sehr häufig zur Anwendung kommen.

Zur Veranschaulichung sind in Tab. 1.1 zwei Beispieldatensätze gegeben, auf die sich die nachfolgenden Berechnungen beziehen.

Tabelle 1.1 Zwei Beispieldatensätze mit je $n = 12$ Versuchspersonen. Zu jeder Versuchsperson ist ein dazugehöriger (arbiträrer) Messwert auf einer Variablen X_1 und X_2 angegeben

	Versuchsperson											
	1	2	3	4	5	6	7	8	9	10	11	12
Beispiel 1 (X_1)	4	4	4	4	5	5	5	5	6	6	6	6
Beispiel 2 (X_2)	5	5	5	5	6	6	6	4	4	4	3	7