

Die Statistik ist uralte. Mit ihr zählte, maß und beschrieb man einst die Mengen von Objekten, die den Herrscher interessierten, also Vorräte, Anzahl der Untertanen, Soldaten usw. Diese „beschreibende Statistik“ existiert heute als offizielle Statistik. Sie sammelt die Fakten möglichst vollzählig und stellt sie dar. In der Antike bemerkten Gelehrte das „Gesetz der großen Zahl“. Sie wiederholten ihre Messungen und berechneten daraus den stets genaueren Mittelwert. Nachdem die Wahrscheinlichkeitsrechnung entstanden ist, entwickelte sich auch die mathematische Statistik, die Stochastik. Diese sammelt die Daten in Stichproben, berücksichtigt ihre Zufälligkeit und schätzt Größen, auch abstrakte. Sie produziert empirisches Wissen.

2.1 Was das Wort „Statistik“ so alles bedeutet

Bis ins späte 19. Jahrhundert verstand man unter dem Begriff „Statistik“ ausschließlich die *beschreibende Statistik*. Sie war ein Bestandteil der „Staatenkunde“. Im Brockhaus-Lexikon von 1819 [2] heißt es zum Begriff Statistik: *Zwei große Kreise bilden den Umfang der geschichtlichen Wissenschaften. Der Kreis der Vergangenheit und der Kreis der Gegenwart. . . . Von jenen beiden Kreisen der Zeit aber wird der Kreis der Vergangenheit durch die Geschichte, der Kreis der Gegenwart durch die Statistik und Geographie dargestellt.* Diese Statistik, die man einst Teil der „Wissenschaft über die Gegenwart“ nannte, bezeichnen wir heute als „beschreibende Statistik“ und verstehen darunter meistens die amtliche Statistik. Nationale und internationale Institutionen sammeln Daten über Bevölkerung, Wirtschaft, Politik, Geographie usw. Diese Daten sollen den gegenwärtigen Zustand der Gesellschaft, des Landes u. ä. möglichst vollständig beschreiben. Die dabei entstehenden Tabellen und

Diagramme selbst werden auch *Statistiken* genannt und beziehen sich immer auf die ihnen zu Grunde liegende reale Gesamtheit. Städte ermitteln z. B. das Alter *aller* Einwohner, meistens unterteilt in Altersgruppen. Deutschland ermittelt z. B. die *gesamte* Stromerzeugung, unterteilt nach Energieträgern. Alles das sind Statistiken. Viele Menschen, einfache Bürger, aber auch Personen, die sich gern öffentlich äußern, verstehen unter dem Begriff „Statistik“ ausschließlich solche Datensammlungen, also Ergebnisse der beschreibenden Statistik. Seit etwa dem Anfang des 20. Jahrhunderts hat sich daraus, durch die Wahrscheinlichkeitsrechnung befruchtet, die *mathematische bzw. schließende Statistik* oder *Stochastik* entwickelt. Sie soll der Gegenstand unserer Betrachtungen sein. Doch zunächst bleiben wir noch etwas bei der beschreibenden Statistik.

2.2 Von prähistorischen Strichlisten und Lotosblumen

Jede Datensammlung ist das Ergebnis von Aktivitäten wie Zählen, Messen oder Wiegen. Man hat Relikte solcher Daten aus der sehr weit zurückliegenden Vergangenheit gefunden. Heute vermutet man, dass das archaische Denken zunächst sehr bildhaft war und hauptsächlich durch die Beziehungen des Einzelnen zur Natur (Nahrung, Sonne, Regen) und zur sozialen Gemeinschaft (Geburt, Tod, Stamm) geprägt worden ist [6]. Irgendwann muss der Mensch begonnen haben zu zählen. Einer der vermutlich ältesten Funde, der das vorzeitliche Zählen belegt, ist ein Wolfsknochen, in den 55 Kerben eingeschnitten sind. Er soll 25 000 bis 30 000 Jahre alt sein. Die ersten 25 Kerben sind in Gruppen zu je 5 angeordnet sind, so, wie mancher Kellner noch heute die Anzahl der Biere auf Bierdeckeln notiert. Dieser Knochen wurde 1937 in Věstonice (Mähren) gefunden [1, 3, 6, 8, 10], siehe Abb. 2.1. Es gibt noch weitere Knochenfunde mit ähnlichen „Kerbungen“. Einige davon sind im Buch von Wußing [10] abgebildet. Später ließ die Kommunikation zwischen den sozialen Gruppen Schriften, und damit auch die entsprechenden Zeichen für Zahlen, entstehen. So findet man z. B. schon in alt ägyptischen Bilderschriften Darstellungen von Mengen. Auf einer Schminkschatulle aus der Zeit von 2850 v. u. Z. sind 6 Lotosblumen dargestellt, was insgesamt die Bedeutung von 6000 Gefangenen haben soll, denn eine Lotosblume bedeutete 1000 Gefangene. Auch das ist eine Statistik.

2.3 Auch die Bibel berichtet von statistischen Erhebungen

In der Bibel findet man mehrere Berichte über statistische Erhebungen. Im Alten Testament (2. Samuel 24, 4) heißt es: . . . „Also zog Joab aus und die Hauptleute des Heeres von dem Könige, dass sie das Volk Israel zählten. . .“. Es folgen die Orte und Völker, die gezählt wurden. Im gleichen Kapitel (2. Samuel 24, 9) findet man dann: . . . „Und Joab gab dem Könige die Summe des Volkes, das gezählet war. Und es waren in Israel acht hundert mal

Abb. 2.1 Wolfsknochen aus den Höhlen des Mährischen Karst, nach [3]. (Dieses Bild stammt ursprünglich aus [1])



tausend starke Männer, die das Schwerdt auszogen; und in Juda fünf hundert mal tausend Mann“. Das ist eine mehr als 3000 Jahre alte Statistik! Im Neuen Testament schildert das Lucas-Evangelium, Kap. 2, die Geburt Christi. Dort heißt es: „*Und es begab sich aber zu der Zeit, dass ein Gebot vom Kaiser Augustus ausging, dass alle Welt geschätzt würde. Und diese Schätzung war die allererste, und geschah zu einer Zeit, da Cyrenius Landpfleger in Syrien war. Und jedermann ging, dass er sich schätzen ließe, ein jeglicher in seiner Stadt“.* Wieder eine statistische Erhebung, nämlich der Bericht über eine Volkszählung, die vor mehr als 2000 Jahren stattgefunden hat!

Tab. 2.1 Die Daten beschreiben einige Eigenschaften Deutschlands 2005 [7]

357 023 km ² Fläche	82 495 000 Einwohner	15 % unter 15 Jahre	16,8 % über 65 Jahre
--------------------------------	----------------------	---------------------	----------------------

2.4 Die offizielle Statistik beschreibt auch noch heute die Welt

Heute fasst die beschreibende Statistik *u. a.* die kollektiven Eigenschaften der von ihr vollständig erfassten Objekte zusammen. Eine solche Statistik enthält z. B. folgende Angaben für die Betrachtungseinheit (das Objekt) „Deutschland“ (im Jahre 2005; Tab. 2.1):

Das sind konkrete Daten, also Zahlenangaben, die einige Eigenschaften Deutschlands beschreiben. Man kann diese Zahlen mit denen aus anderen Jahren oder aus anderen Ländern vergleichen und z. B. untersuchen, wie sich der Prozentsatz der unter 15 jährigen oder der über 65 jährigen im Verlaufe der Zeit entwickelt hat. Diese Angaben sind Tatsachen, an denen man, vorausgesetzt die Daten sind korrekt erhoben worden, nicht zweifeln kann. Anders ist es, wenn abstrakte Größen angegeben werden sollen, die sich der direkten Beobachtung entziehen, wie z. B. Risiken. Dazu braucht man die *Stochastik* bzw. die *mathematische* oder *schließende Statistik*.

2.5 Die Stochastik misst auch abstrakte Größen, wie Risiken, Trends oder Korrelationen

Die gegenwärtige Gesellschaft informiert sich immer häufiger über immer mehr Aspekte ihres Alltags. Es geht nicht mehr nur um das Erfassen konkreter Größen, sondern auch um abstrakte Größen, z. B. um Risiken. Diese lassen sich nicht direkt beobachten, sondern nur schätzen. Man ist bemüht, eine vermutete Gesetzmäßigkeit „statistisch“ nachzuweisen, eine wissenschaftliche Hypothese „auf ihre Signifikanz hin“ zu prüfen, usw. Die moderne Gesellschaft, die sich ziemlich selbstgerecht und anmaßend gern „Wissengesellschaft“ nennt (was wird man wohl in 500 Jahren dazu sagen?), stützt viele ihrer Entscheidungen auf solche Größen, also auf Risiken, Trends oder Korrelationen. Diese sind der direkten Beobachtung nicht zugänglich, denn es sind *kollektive* Eigenschaften, meist von *gedachten* Kollektiven. Eine solche Größe sagt nichts über den Einzelfall aus. Im Straßenverkehr gibt es ein Unfallrisiko. Es wird z. B. in Berlin jährlich auf der Basis der Anzahl der Unfälle im vorangegangenen Jahr geschätzt. Die ermittelte Zahl bezieht sich auf *alle* Teilnehmer am *gesamten* Straßenverkehr. Um die Wahrscheinlichkeit eines Unfalls schätzen zu können, müsste man die ermittelte Zahl von Unfällen auf die gesamte Anzahl der Teilnehmer am Straßenverkehr beziehen. Doch das ist schwer, denn es gibt Fußgänger, Radfahrer, Motorradfahrer und Autofahrer und sie alle sind nur zeitweilig unterwegs. Zudem spielt das Wetter eine Rolle, es gibt Sonnenschein, Regen, Nebel, Schnee und Eis. Angenommen, alles das könnte „im Mittel“ berücksichtigt werden, dann könnte man die Wahrschein-

lichkeit für Unfälle aus den Daten in der zurückliegenden Periode schätzen. Mit diesem Wert kann man jedoch prinzipiell nicht vorhersagen, ob eine bestimmte Person Opfer des Straßenverkehrs werden wird. Denn das „Risiko“ ist eine abstrakte Eigenschaft, die nur für das am Straßenverkehr teilnehmende gesamte *Kollektiv* gilt. Es realisiert sich zwar individuell bei demjenigen, der in einen Unfall verwickelt ist, kann aber nur für das Kollektiv „aller Verkehrsteilnehmer“ ausgedrückt werden. Man kann das Risiko verringern, indem die Straßen in einem guten Zustand gehalten werden und indem für eine vernünftige Verkehrsordnung gesorgt wird. Die Größe des „Risikos“ muss aber bekannt sein, wenn die Stadtverwaltung die Wirksamkeit ihrer Maßnahmen bewerten oder verschiedene Jahre vergleichen soll. Um einen Unfallschwerpunkt zu erkennen, braucht sie die Zahl der Unfälle bezogen auf einzelne Straßenabschnitte, Kreuzungen usw. Dazu muss man die Daten lokal erfassen. Dann lässt sich die Unfallohäufigkeit lokal bewerten und auch die Frage beantworten, ob die ermittelte Zahl im Rahmen der normalen Schwankungen liegt oder ob sie an dieser Kreuzung im Vergleich zu anderen wesentlich erhöht ist. Zur Beantwortung derartiger Fragen brauchen wir die schließende Statistik oder Stochastik. Mit ihrer gedanklichen Grundlage werden wir uns im Folgenden ausführlich befassen.

2.6 Das Gesetz der großen Zahl

Die Basis statistischer Schlüsse ist das Gesetz der großen Zahl. Es ist ein universell geltendes Gesetz und wirkt auf zufällige Messfehler und relative Häufigkeiten. Jede Messung ist unvermeidbar mit einem zufälligen Fehler behaftet. Misst man eine Größe mehrfach und bildet den Mittelwert der Messergebnisse, so ist dieser in einer größeren Anzahl von Messungen genauer als in einer geringeren. Das wusste man schon in der Antike. Wußing [10] zitiert den Bericht des Historikers Thudydikes (455–396 v. u. Z.). Dieser schrieb: „... und mochten sie sich irren, so musste doch die Mehrzahl die rechte Summe treffen, zumal sie öfters zählten“. Gibt es ein zufälliges Ereignis, das mit einer bestimmten Wahrscheinlichkeit eintrifft, z. B. die Wahrscheinlichkeit dafür, dass ein Neugeborenes ein Mädchen ist, so lässt sich auch diese Wahrscheinlichkeit auf der Grundlage einer großen Zahl von Beobachtungen genauer schätzen als auf einer kleinen. Es gibt Hinweise darauf, dass die Menschen schon in sehr früher Zeit Erfahrungen mit dem Zufall und der Wahrscheinlichkeit hatten und wussten, dass man mit einer großen Anzahl von Beobachtungen eine gesuchte „undeutliche“ Größe besser erkennen kann als mit einer kleinen. Die mathematische Fassung des Gesetzes der großen Zahlen (es gibt mehrere davon, ein starkes und ein schwaches) geht vermutlich auf Jakob Bernoulli zurück. Er formulierte es 1713 in seiner Schrift *Ars conjectandi* (die Kunst des Vermutens). Dieses Gesetz wurde von vielen Mathematikern weiterentwickelt, bis es im 20. Jahrhundert Chinchine und Kolmogorov (1925) in der heutigen Form formulierten, und zwar mit der erforderlichen mathematischen Strenge. Dass das Gesetz der großen Zahl funktioniert, hat an Wahlabenden anhand der im Fernsehen übertragenen Prognosen für die einzelnen Parteien wohl jeder schon einmal beobachten

können. So lange die Anzahl der ausgezählten Stimmen klein ist, liegen die ermittelten Stimmenanteile der Parteien meistens noch weit weg von den endgültigen Prozentsätzen. Mit der wachsenden Zahl ausgezählter Stimmen wird die Unsicherheit immer kleiner und verschwindet zuletzt, denn sie nähert sich dem endgültigen „wahren“ Wert, der nach der Auszählung aller Stimmen feststeht.

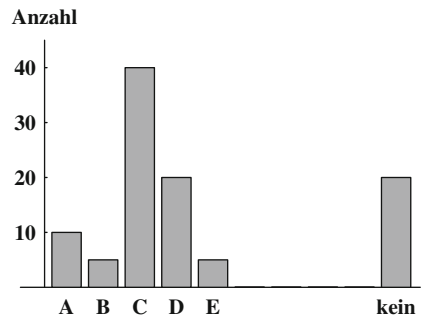
2.7 Stochastik oder die „Kunst des Erratens“

Der Begriff Statistik hat sich im Laufe der Zeit beachtlich erweitert. Im Unterschied zum Brockhaus von 1819 [2] sieht z. B. die 1998 herausgegebene Hutchinson Encyclopedia of Science [5] die Statistik als ein Teilgebiet der Mathematik, das der Sammlung und Interpretation von Daten dient. Auch im Zeit-Lexikon von 2005 [4] wird die Statistik als Teilgebiet der Stochastik (griechisch, als die zum Erraten gehörende Kunst) bezeichnet. Die Stochastik ist ein Komplex von Begriffen und Sätzen aus der Wahrscheinlichkeitstheorie und mathematischen Statistik. Sie ist die Brücke, die von den beobachteten Daten zu einer verallgemeinerten statistischen Aussage führt. Die Brückenpfeiler sind dabei die beobachteten Realisierungen der zufälligen Größe und ein mathematisches Modell. Die Gesetzmäßigkeiten der zufälligen Größen lassen sich durch wiederholtes Messen oder Beobachten herausfiltern, denn auf sie wirkt das Gesetz der großen Zahl. Das mathematische Modell dabei ist ein Wahrscheinlichkeitsmodell, das die zufälligen Größen enthält und dessen Gültigkeit vorausgesetzt werden muss. Die Stochastik als die „zum Erraten gehörende Kunst“ ermöglicht es, aus den Daten jene Antworten zu extrahieren, nach denen gesucht wird. Das geschieht z. B. in vielen der heute sehr beliebten empirischen Studien, deren Ergebnisse so oft in Zweifel gezogen werden. Laut Wikipedia (von 2011) [9] ist die Statistik die Lehre von Methoden zum Umgang mit quantitativen Informationen (den Daten). Sie schafft die Möglichkeit, quantitatives empirisches Wissen zu gewinnen oder theoretische Aussagen abzuleiten, sowie ein vorgeschlagenes Modell durch Beobachtungswerte zu verifizieren.

2.8 Von der Stichprobe zur Häufigkeitsverteilung

Der Gegenstand empirischer Studien ist vielgestaltig. Es muss sich dabei nicht immer um wichtige Fragen handeln. Das zeigen auch die in der Einleitung genannten Umfragen zu Sonnenölen oder Kochsendungen. Die Studenten sollten herausfinden, *in welchem Maße* die verschiedenen Sonnenöle oder Kochsendungen in der Bevölkerung bekannt und beliebt sind. Es wird also nach der relativen Häufigkeit von Menschen gesucht, die das Sonnenöl A oder B oder . . . bevorzugen oder überhaupt keines benutzen. Alle Befragten zusammen bilden eine *Stichprobe*. Entsprechendes gilt für die Kochsendungen, dort wird nach der

Abb. 2.2 Häufigkeitsverteilung (Histogramm) der Sonnenöle



relativen Anzahl von Menschen gesucht, welche die Sendung A, B, ... bevorzugen, oder überhaupt keine. In beiden Fällen bilden alle Befragten zusammen 100 % des Kollektivs, das eine Stichprobe der dort anzutreffenden Menschen ist. Was man eigentlich sucht, ist aber die Verteilung der Merkmale A, B, ... in der gesamten Bevölkerung, der *Grundgesamtheit*. Denn derjenige, der die Untersuchung veranlasst hat, möchte auf der Grundlage dieser Daten entscheiden, ob er für gewisse Sonnenöle intensiver werben muss, ob er bestimmte Marken vom Markt nehmen sollte, usw. bzw. ob bestimmte Kochsendungen gemocht oder nicht gemocht werden. Wie man von einer solchen Stichprobe auf die gesamte Population (die Grundgesamtheit) schließt, ist das Anliegen der Stochastik. Das Resultat der Umfrage ist das Datenmaterial. Daraus ergibt sich eine *Häufigkeitsverteilung*, die nach der Befragung von 100 Personen die in der Abb. 2.2 dargestellte Form haben könnte, die Form eines *Histogramms*.

Bekommen die Auftraggeber dieser Studie, die auf ihrer Grundlage entscheiden wollen, die gewünschte Antwort? Nicht unbedingt. Die Auswahl der Befragten ist zwar zufällig, aber an dieser Stelle Wiens trifft man hauptsächlich Touristen in bevorzugten Altersklassen. Sind 100 Befragte genug? Für die 6 Möglichkeiten (A, B, C, D, E, kein Sonnenöl) sind 100 Befragte keine große Zahl. Um einigermaßen sichere statistische Schlüsse ziehen zu können, müsste sie wohl größer sein. Auch bei der Zusammensetzung der Stichprobe müssen gewisse Regeln eingehalten werden. Die Stichprobe soll sich so zusammensetzen, wie es der zu untersuchenden Zielgruppe entspricht, sie soll „repräsentativ“ sein. Würde eine Umfrage über die in Europa üblichen Sonnenöle in Sydney durchgeführt, wäre das für den europäischen Markt nicht repräsentativ. Die Stichprobe wäre auch nicht repräsentativ, wenn nur Kinder oder Übersechzigjährige befragt würden. Ob die Stichproben der Wiener Studenten repräsentativ waren, das darf bezweifelt werden.

2.9 Messwerte, Fehlerrechnung und Glockenkurve

Nach diesem Beispiel des Ermitteln einer Häufigkeitsverteilung wenden wir uns nun einem weiteren typischen Anwendungsbereich der Statistik zu, nämlich der Reduzierung des Einflusses von Messfehlern durch die Fehlerrechnung. Ungefähr seit dem 17. Jahrhundert

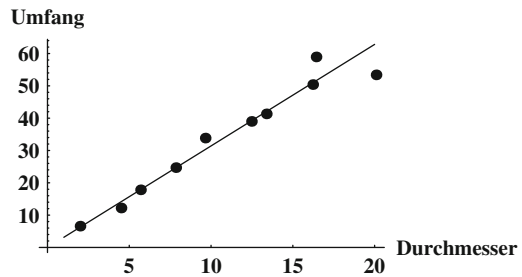
Tab. 2.2 Häufigkeitsverteilung von Messwerten

Messwerte im Bereich (cm)	Strichliste	Anzahl
97,5–97,99	I	1
98,0–98,49	IIII I	6
98,5–98,99	IIII	4
99,0–99,49	IIII III	8
99,5–99,99	IIII I	6
100,0–100,49	IIII IIIII I	11
100,5–100,99	IIII	5
101,0–101,49	IIII	4
101,5–101,99	IIII	4
102,0–102,49		0
102,5–102,99	I	1

benutzt die Naturwissenschaft in wachsendem Maße Messgeräte. Ihre Genauigkeit ließ oft zu wünschen übrig. Man weiß schon sehr lange, dass man durch Wiederholungen von Messungen und ihre Mittelung die Genauigkeit verbessern kann. Carl Friedrich Gauß hat seit 1794 im Zusammenhang mit der Beobachtung des Planetoiden Pallas die Methode der kleinsten Quadrate entwickelt, verwendet und 1809 publiziert. Es geht um Folgendes: Ein unbekanntes Maß soll bestimmt werden, etwa eine Entfernung. Dazu stellt man sich vor, dass die wahre Entfernung der gesuchte Messwert ist, zu diesem ist aber stets ein zufälliger Fehler addiert. In einer langen Messreihe wird der Mittelwert aller zufälligen Fehler fast Null, weil sich die Fehler ausgleichen. Der gemittelte Messwert nähert sich dadurch dem wahren Wert. Die Häufigkeitsverteilung aller Messfehler folgt (unter bestimmten Bedingungen) einer Gaußschen Normalverteilung. Sie ist allgemein als Glockenkurve bekannt. Stellen wir uns vor, wir hätten eine Kante von 100 cm in 50 Wiederholungen gemessen. Das Ergebnis könnte die folgende Strichliste sein, siehe Tab. 2.2.

Hätten wir diese Kante viel öfter gemessen und hätte unser Messgerät keinen systematischen Fehler, so würde sich aus so einer Strichliste eine symmetrische Häufigkeitsverteilung in Glockenform entwickeln, deren Mittelwert eine Schätzung der wahren Kantenlänge ergibt. Die Breite der Häufigkeitsverteilung würde etwas über die Genauigkeit unseres Messgerätes aussagen. Im Unterschied zu den vorherigen Beispielen müssen wir hier die Messwerte in Zahlenbereiche einsortieren, denn unser Maß ist eine kontinuierliche Größe. Im Sonnenöl-Beispiel ging es um die möglichst genaue Ermittlung relativer Häufigkeiten, hier interessiert uns der mittlere Messwert. Um einen Messwert auf diese Weise genau genug zu bestimmen, muss die Stichprobe groß genug sein und das Messgerät darf keinen systematischen Fehler haben. Wie groß die Stichprobe sein muss, das hängt von der Genauigkeit der einzelnen Messungen ab (der Streuung der Messwerte) und natürlich auch von der gewünschten Genauigkeit des Mittelwertes.

Abb. 2.3 Kreisdurchmesser und Kreisumfänge zur Schätzung von π



2.10 Die empirische Bestimmung von Konstanten

Die auf Beobachtungs- oder Messwerten beruhenden statistischen Methoden können auch zur empirischen Bestimmung unbekannter Konstanten verwendet werden. Angenommen, wir wüssten nicht, dass das Verhältnis des Kreisumfangs zum Durchmesser konstant und gleich π ist. (Der Wert von π ist mit 4 Stellen nach dem Komma gleich 3,1416, π ist eine irrationale Zahl und heute bis auf mehr als eine Billion Stellen bekannt.) Schon Archimedes (287–212 v. u. Z.) hat die Konstante π angenähert bestimmt, indem er eine obere und eine untere Schranke für π durch ein in den Kreis eingeschriebenes und ein den Kreis umschreibendes 96-Eck berechnete. Wir sind also ziemlich ungebildet, wissen nichts über die Konstante π und wollen sie empirisch mit Hilfe der Statistik bestimmen. Dazu könnten wir die Durchmesser und Umfänge mehrerer Kreise messen und die Ergebnisse in einem Koordinatensystem grafisch darstellen. Das mögliche Ergebnis zeigt Abb. 2.3. In ihr sind über den 10 gemessenen Durchmessern verschiedener Kreise ihre Umfänge dargestellt. Der Anstieg der Ausgleichsgeraden ist linear und ungefähr gleich π . Wollten wir dabei eine bestimmte Genauigkeit der Schätzung von π erreichen, so wäre es notwendig, auch die Anzahl der Messungen und ihre Genauigkeit richtig zu wählen. Im Gegensatz zu der im absoluten Sinne geltenden Genauigkeitsgrenze von Archimedes kann der nach unserer Methode berechnete Wert mit einer bestimmten Irrtumswahrscheinlichkeit falsch sein. Das ist die wesentliche Eigenschaft statistisch bestimmter Größen. Doch es wird wohl kaum jemand auf die Idee kommen, die überaus gut bekannte Konstante π auf unsere Weise zu bestimmen! In der Praxis existieren jedoch viele unbekannte Abhängigkeiten, die man nur statistisch bestimmen kann. Auch dazu braucht man die Stochastik.

Statistische Untersuchungen beruhen immer auf Beobachtungswerten. Wenn es Daten aus einer Stichprobe oder Messwerte sind, gibt es immer den Einfluss des Zufalls. Durch das Gesetz der großen Zahl wissen wir, dass die Auswirkung des Zufalls mit wachsender Anzahl von Beobachtungswerten abnimmt. Während es in der beschreibenden Statistik hauptsächlich darum geht, das Charakteristische der Daten dem jeweiligen Zweck der Untersuchung angemessen als Tatsache auszudrücken, geht die Stochastik weiter: Sie sucht nach allgemeineren Aussagen und muss deshalb den Zufallseinfluss berücksichtigen, dem die Daten unterliegen.

Wichtige Begriffe

Beschreibende Statistik	Sammelt Daten, berechnet aus ihnen charakteristische Größen und stellt sie dar.
Schließende Statistik, mathematische Statistik	Sammelt Daten in Stichproben, sieht sie als Realisierungen zufälliger Größen an und schätzt daraus charakteristische Größen.
Stochastik	Griechisch: Kunst des Erratens. Synonym für mathematische Statistik.
Häufigkeitsverteilung	Absolute oder relative Anzahl aller beobachteten Werte.
Histogramm	Grafische Darstellung der Häufigkeitsverteilung.

Literatur

1. Absolon, K.: Dokumente und Beweise der Fähigkeiten des fossilen Menschen zu zählen im Mährischen Paläolithikum. *Artibus Asiae*. **20**(2–3), 123–150 (1957)
2. Brockhaus Lexikon, Leipzig.: (1819)
3. Detlefsen, M.: Kerbknöchen und Kerbhölzer. *Wissenschaft und Fortschritt*. **27**, 396 (1977)
4. Die Zeit.: Das Lexikon in 20 Bänden. Zeitverlag Gerd Bucerius & Co., Hamburg (2005)
5. The Hutchinson Encyclopedia of Science. Helicon Publ. Group Ltd, Oxford (1998)
6. Klix, F.: *Erwachendes Denken*, S. 204. Deutscher Verlag der Wissenschaften, Berlin (1985)
7. Redaktion Weltalmanach (Hrsg.): *Der Fischer Weltalmanach 2005*. Fischer Taschenbuch, Frankfurt a. M. (2006)
8. Struik, D. J.: *Abriss der Geschichte der Mathematik*. Deutscher Verlag der Wissenschaften, Berlin (1961)
9. Wikipedia (2011)
10. Wufing, H.: *6000 Jahre Mathematik*, Bd. 1, S. 75. Springer, Berlin Heidelberg (2008)

Statistisch gesichert und trotzdem falsch?

Vom (Un-)Wesen statistischer Schlüsse

Härtler, G.

2014, XI, 143 S., Softcover

ISBN: 978-3-662-43356-0