

Worum es bei statistischen Verfahren überhaupt geht

Erfolgreich mit Statistik umgehen – im Alltag, im Beruf oder im Studium

Kapitel 1

Statistik kurz und knapp

Die heutige Welt ist so mit Zahlen überfrachtet, dass wir (auch ich!) davon geradezu überwältigt werden. Wenn Sie gegenüber den Statistiken, die Ihnen in den Medien begegnen, eine zynische Grundhaltung entwickelt haben, kann ich Ihnen keinen Vorwurf machen – manchmal geht es mir da auch nicht anders. Aber auch wenn da draußen jede Menge irreführende und falsche Informationen auf Sie warten, gibt es doch jede Menge statistische Arbeiten, die zu kennen sich lohnt. Viele Studien und Verfahren, die mit Daten umgehen, tragen zur Verbesserung unser aller Lebensqualität bei. Ihre Aufgabe besteht nun darin, die verlässlichen von den irreführenden Methoden und Ergebnissen zu unterscheiden – und zwar im Vertrauen auf Ihre eigenen Fähigkeiten. Wenn Sie Statistik und statistische Verfahren wirklich verstehen, verschaffen Sie sich die nötige Sicherheit, die Sie für den gewinnbringenden Umgang mit Zahlen in Alltag, Beruf und Studium brauchen. Genau darum geht es in diesem Buch.

In diesem Kapitel zeige ich auf, welche Rolle die Statistik in unserer heutigen datenlastigen Gesellschaft spielt und was Sie tun können, um in dieser Datenflut nicht nur nicht unterzugehen, sondern sie sogar für sich zu nutzen. Ich nehme Sie mit in die Welt der wissenschaftlichen Anwendung der Statistik und zeige Ihnen, wie man sinnvolle Studien konzipiert, verlässliche Daten erhebt, Informationen analysiert, Ergebnisse interpretiert und daraus eigene Schlüsse zieht. (Und Sie dachten schon, Statistik wäre ein reines Zahlenwerk!)

Statistiken verstehen und nutzen

Unser Alltag wird geradezu überflutet von Statistiken der unterschiedlichsten Art, die wir gar nicht immer als solche wahrnehmen. Das fängt schon an, wenn Sie morgens das Radio einschalten (oder die Zeitung aufschlagen oder sich im Internet informieren) und dort die Wettervorhersage hören (oder lesen) – denn die beruht auf der statistischen Analyse früherer (meteorologischer) Daten und der aktuellen Wetterbedingungen. Am Frühstückstisch lesen Sie vielleicht auf der Müslipackung, welche Anteile von Kohlehydraten, Proteinen und Fett Ihr Morgenmahl hat. Im Büro schlagen Sie Zahlen in Diagrammen

und Tabellen nach, tragen Daten in eine Tabellenkalkulation ein, führen eine Diagnose oder Messungen durch, berechnen Werte, schätzen Kosten, treffen Entscheidungen auf der Grundlage statistischer Werte oder arrangieren Ihr Inventar anhand der Verkaufszahlen des Vormonats.

Wenn Sie Glück haben, nehmen Sie Ihr Mittagessen in einem Restaurant ein, das von 500 befragten Personen zur Nr. 1 erkoren wurde. Was Sie dort für Ihr Essen bezahlen, hängt von Marktdaten ab. Nachmittags haben Sie einen Arzttermin, bei dem Ihr Blutdruck gemessen, Körpergewicht und Körpergröße festgestellt und Blutuntersuchungen durchgeführt werden. Wenn alle Informationen beisammen sind, erhalten Sie einen Bericht, der Ihre Werte mit denen der statistischen Norm vergleicht.

Sie fahren nach Hause in einem Auto, bei dessen Wartung heutzutage in der Regel ein Computer für eine statistische Diagnostik zum Einsatz kam. Zu Hause schalten Sie noch mal die Nachrichten ein, hören die neuesten Statistiken zu Wohnungseinbrüchen und Autodiebstählen, verfolgen die Entwicklung der Börsenkurse und erfahren, wie viele Menschen beim DFB-Pokalspiel im Stadion waren.

Abends putzen Sie sich die Zähne mit einer Zahnpasta, die statistisch erwiesenermaßen gegen Karies wirkt, lesen noch ein paar Seiten in einem *Spiegel*-Bestseller (der aufgrund von statistischen Verkaufszahlen auf der Liste gelandet ist), schlafen ein – und am nächsten Morgen geht alles wieder von vorne los. Aber wie können Sie sicher sein, dass alle diese Statistiken, auf die Sie sich täglich verlassen, auch korrekt sind? In Kapitel 2 befasste ich mich ausführlicher mit einigen Beispielen dafür, welche Rolle Statistik in unserem Alltag und am Arbeitsplatz spielt, welche Auswirkungen sie hat und wie Sie Ihre Wahrnehmung für Statistik schärfen können.



Manche Statistiken sind vage, ungeeignet oder schlicht und einfach falsch. Deshalb müssen Sie Ihre Wahrnehmung für Statistik im Alltag schärfen. Sie können üben, Statistiken bewusster wahrzunehmen und sich die Informationen genauer anzusehen, Fragen zu stellen und eine gesunde Skepsis zu entwickeln. In Kapitel 3 erfahren Sie, wie schlechte Statistiken Sie in die Irre führen können und wie Sie sich die nötigen Fähigkeiten aneignen, um kritisch zu denken und Probleme zu erkennen, anstatt einfach die Ergebnisse für bare Münze zu nehmen.

Wie alle anderen Fachgebiete hat auch die Statistik eine eigene Fachsprache (manchmal auch etwas lieblos Jargon genannt). In Kapitel 4 erkläre ich kurz einige der am häufigsten gebrauchten statistischen Fachbegriffe. Wenn Sie sich mit dieser Fachsprache vertraut machen, anstatt sich von ihr einschüchtern zu lassen, können Sie Statistik auf einem höheren Niveau verstehen und darüber sprechen. Sie gewinnen an Glaubwürdigkeit, wenn Sie die richtigen Ausdrücke verwenden, um zu beschreiben, warum genau ein bestimmtes statistisches Ergebnis falsch ist. Wenn Sie in Ihren Präsentationen statistische Tabellen, Diagramme und Analysen einsetzen, werden diese durch die richtige Fachsprache effektiver und informativer. (Und wenn Sie das alles nicht überzeugt: Sie müssen die Fachausdrücke schon deshalb lernen, weil ich sie im ganzen Buch verwende, auch wenn ich sie mehrfach erkläre.)

In den nächsten Abschnitten erfahren Sie, in welcher Form Statistik an den einzelnen Phasen der wissenschaftlichen Methodik beteiligt ist.

Aussagekräftige Studien konzipieren

Fragen haben alle – pharmazeutische Unternehmen und Biologen genauso wie Marktforscher, Politiker allgemein oder speziell die Regierung. Zur Beantwortung dieser Fragen brauchen all diese Personengruppen am Ende die Statistik. Insbesondere viele medizinische und psychologische Studien werden durchgeführt, weil bestimmte Menschen auf der Suche nach einer Antwort auf drängende Fragen sind. Beispiele gefällig?

- ✓ Kann dieser Impfstoff den Ausbruch der Grippe verhindern?
- ✓ Wie schätzen die Bundesbürger die aktuelle wirtschaftliche Lage des Landes ein?
- ✓ Führt eine intensive Nutzung sozialer Netze im Internet bei Jugendlichen zu Depressionen?

Wenn eine Forschungsfrage formuliert wurde, besteht der erste Schritt zu ihrer Beantwortung darin, eine aussagekräftige Studie zu planen, um die benötigten Daten zu erheben. Das bedeutet, dass man sich überlegen muss, mit welchen Verfahren man diese Daten gewinnen kann. In diesem Abschnitt gebe ich Ihnen einen kleinen Überblick über die beiden wesentlichen Grundformen von Studien, nämlich Umfragen (auch als Befragungen bezeichnet) und Experimente (auch als Versuche bezeichnet). Außerdem erkläre ich, warum es wichtig ist, die Anlage (das heißt: die Konzeption) einer Studie zu prüfen, bevor man deren Ergebnissen traut.

Umfragen

Eine *Beobachtungsstudie* ist eine Studie, in der Daten von Individuen erhoben werden, ohne Einfluss auf diese zu nehmen. Die häufigste Beobachtungsstudie ist die Umfrage. Bei einer Umfrage (oder Befragung) werden Personen, die aus einer interessierenden Grundgesamtheit ausgewählt wurden, mittels eines Fragebogens befragt. Diese Befragung kann auf vielfältige Weise erfolgen: Sie können gedruckte Fragebögen verwenden, die per Post verschickt werden, Umfragen ins Internet stellen, über das Fernsehen die Zuschauer zur telefonischen Beantwortung einer Frage aufrufen, Personen gezielt telefonisch ansprechen und vieles weitere mehr.



Eine korrekt durchgeführte Umfrage kann ein sehr nützliches Instrument dafür sein, an Informationen zu kommen. Werden Umfragen jedoch nicht ordentlich durchgeführt, kann dabei auch einfach nur Quatsch herauskommen. Mögliche Probleme liegen in den Formulierungen der Fragen, die irreführend sein können, in der Verweigerung der Teilnahme ausgewählter Personen oder im Übersehen ganzer Gruppen der Grundgesamtheit bei der Auswahl der zu Befragenden. Aufgrund solcher möglichen Fehlerquellen muss eine Umfrage immer sehr sorgfältig geplant werden.



Viele Wissenschaftler betreiben einen enormen Aufwand, auch finanzieller Art, um gute Umfragen durchzuführen. Das können Sie merken (anhand von Kriterien, die ich Ihnen noch vorstelle). Da Ihnen aber in den Medien, am Arbeitsplatz und auch im Studium so viele verschiedene Arten von Umfragen begegnen, müssen Sie in der Lage sein, schnell prüfen beziehungsweise beurteilen zu können, wie eine Studie konzipiert und durchgeführt wurde, damit Sie sachkundig auf mögliche Probleme hinweisen können. Die Instrumente, die Sie brauchen, um Befragungen zu beurteilen, finden Sie in Kapitel 16.

Experimente

Bei einem *Experiment* werden die beobachteten Individuen einer oder mehrerer Untersuchungen (»Behandlungen«) unterzogen, die eindeutige Vergleiche zwischen den Personen(gruppen) erlauben. Dabei werden die jeweiligen Ergebnisse aufgezeichnet. Wenn man beispielsweise prüfen möchte, ob ein Medikament sich auf den Blutdruck auswirkt, lässt man vielleicht eine Gruppe 10 mg des Medikaments nehmen und eine weitere Gruppe 20 mg. Häufig wird auch eine Kontrollgruppe eingesetzt, in der die Teilnehmenden ein Scheinmedikament (als Placebo bezeichnet) erhalten oder ein bereits zugelassenes Medikament gegen dieselbe Krankheit.



Gute und glaubwürdige Experimente sind so angelegt, dass mögliche Verzerrungen (auch als systematischer Fehler oder *Bias* bezeichnet) minimiert werden, dass viele verlässliche Daten erhoben und geeignete Vergleiche (von Behandlungs- und Kontrollgruppe) angestellt werden können. Mögliche Probleme bei Experimenten ergeben sich, wenn Wissenschaftler oder Teilnehmende wissen, welche Behandlung angewendet wurde bzw. welche Behandlung sie erhalten haben, wenn Faktoren, die sich auf das Ergebnis auswirken, nicht berücksichtigt werden (beispielsweise das Körpergewicht bei Untersuchungen zur Medikamentendosierung) oder wenn eine Kontrollgruppe fehlt (sodass man die Ergebnisse nicht mit einem Ausgangswert ohne Behandlung vergleichen kann).

Gut aufgebaute Experimente können jedoch dazu beitragen, eine Ursache-Wirkungs-Beziehung zu erkennen, wenn der Unterschied in den Ergebnissen zwischen der Behandlungsgruppe und der Kontrollgruppe statistisch signifikant ist (also mit einer gewissen Wahrscheinlichkeit nicht nur zufällig beobachtet wurde).



Experimente können zur Entwicklung und Prüfung von Arzneimitteln herangezogen oder verwendet werden, um herauszufinden, wie man Nahrungsmittel am besten herstellt und zubereitet, und auch, um festzustellen, ob eine neue Behandlung eine Krankheit heilen oder lindern kann. Zweifellos hat sich unsere Lebensqualität durch gut geplante und durchgeführte Experimente verbessert. Aber nicht alle Experimente sind gut geplant, und es ist wichtig, dass Sie lernen zu erkennen, welche Ergebnisse glaubwürdig sind und welche nicht. Das gilt besonders für Fragen, die Ihnen wichtig sind. Was Sie über Experimente wissen müssen und wie Sie sie beurteilen können, lernen Sie in Kapitel 17.

Erheben von brauchbaren Daten

Nachdem eine Studie, also eine Umfrage oder ein Experiment, geplant wurde, müssen die Teilnehmenden ausgewählt und ein Verfahren entwickelt werden, um die Daten zu erheben. Diese Phase ist ganz entscheidend dafür, ob am Ende glaubwürdige, brauchbare Daten vorliegen. In diesem Abschnitt erkläre ich, worauf es besonders ankommt.

Auswahl einer guten Stichprobe



In der Statistik gilt ganz besonders, was Sie auch aus anderen Bereichen kennen: Ein Ergebnis kann nur so gut sein, wie es die Ausgangsbedingungen erlauben. Wenn Sie schon bei der Auswahl Ihrer Teilnehmenden eine *Verzerrung* einführen, also bestimmte Personen oder Gruppen bevorzugen, dann werden auch Ihre Ergebnisse verzerrt sein. So einfach ist das.

Angenommen, Adrian möchte wissen, was die Menschen in seiner Heimatstadt über die Eröffnung einer Spielbank denken. Er geht ins größte Einkaufszentrum oder Kaufhaus der Stadt, nimmt sein Klemmbrett für Notizen mit und fragt die Leute, die an ihm vorbeikommen, nach ihrer Meinung dazu. Was macht er falsch? Adrian wird auf diese Weise nur die Meinung von Menschen erfahren, die a) in diesem Einkaufszentrum oder Kaufhaus einkaufen, und zwar b) nur an diesem einen Tag und c) nur zu der Zeit, zu der er dort ist, und die d) auch noch Lust haben, ihm zu antworten.

Das sind sehr viele Einschränkungen, die dazu führen, dass die Befragten nicht repräsentativ für alle Menschen in der Stadt sind. Adrian könnte auch eine Webseite ins Netz stellen und dort Menschen auffordern, über die neue Spielbank abzustimmen. Dann stützt sich aber sein Ergebnis nur auf Menschen, die von seiner Webseite wissen, Internetzugang haben und bereit sind zu antworten. Erfahrungsgemäß sind das nur Personen, die ein ausgeprägtes Interesse an der Frage haben. Am Ende hat Adrian dann eine Menge verzerrter Daten von Personen, die nicht stellvertretend für die ganze Stadt sind.



Das Schlüsselwort für die Vermeidung von Verzerrungen ist »zufällig«. Sie müssen Ihre Stichprobe von Personen *zufällig* auswählen – also so, als ob Sie Namen aus einem Hut ziehen (ohne in den Hut hineinzuschauen, am besten mit verbundenen Augen). In der Wissenschaft gibt es verschiedene Methoden, um Personen zufällig auszuwählen, die Sie in Kapitel 16 kennenlernen werden.

Wenn Sie ein Experiment planen, ist die zufällige Auswahl von Teilnehmenden oft ethisch nicht vertretbar, da bei einem Experiment die Teilnehmenden in der Regel einer Behandlung unterzogen werden. Daher geht man anders vor: Man sucht sich Freiwillige für die Teilnahme, die repräsentativ für die zu untersuchende Grundgesamtheit sind, und sorgt dafür, dass die Daten so erhoben werden, dass sich die Ergebnisse auf eine größere Gruppe übertragen lassen. Wie das geht, erfahren Sie in Kapitel 17.

Wenn Sie die Kapitel 16 und 17 durchgearbeitet haben, wissen Sie, wie Sie die von anderen angewandten Methoden zur Auswahl der jeweiligen Stichproben genau prüfen können, und Sie können sogar Ihren eigenen Plan für eine Stichprobenziehung erstellen. Sie wissen dann, ob vorliegende Ergebnisse unbrauchbar sind, weil die Studie einfach nicht gut geplant und durchgeführt wurde.

Verzerrungen in den Daten vermeiden

Als *Verzerrung*, *systematischen Fehler* oder Neudeutsch *Bias* (sprich Bei-ess) bezeichnet man die systematische Bevorzugung bestimmter Personen oder Antworten. Die Verzerrung ist der Erzfeind aller Statistikerinnen und Statistiker, und diese werden alles tun, um sie zu vermeiden. Sie brauchen ein Beispiel? Angenommen, Sie führen eine telefonische Umfrage zur Zufriedenheit am Arbeitsplatz durch. Wenn Sie dafür Menschen zu Hause anrufen, und zwar zwischen 9 und 17 Uhr, bekommen Sie keine Antworten von Personen, die tagsüber arbeiten (oder nur von solchen, die in Heimarbeit tätig sind). Es ist aber denkbar, dass Menschen, die tagsüber auf der Arbeit sind, zufriedener sind als Menschen, die in Nachtschichten arbeiten.

Auch bei der eigentlichen Erhebung von Daten mittels Umfragen müssen Sie auf Verzerrungen achten. Ein Beispiel: Manche Fragebögen sind einfach zu lang. Was passiert, wenn jemand die Befragung nach der Hälfte abbricht? Oder wie gehen Sie damit um, wenn Ihnen jemand falsche Auskünfte erteilt, also beispielsweise angibt, 5.000 Euro im Monat zu verdienen, obwohl es in Wirklichkeit nur 2.500 Euro sind? Und was machen Sie, wenn Sie eine Antwort bekommen, die in Ihrer Liste möglicher Antworten gar nicht enthalten ist? Bei der Erhebung von Daten mittels Befragungen können jede Menge Probleme auftreten, die Sie erkennen müssen.



Hinsichtlich möglicher Verzerrungen und der Datenerhebung sind Experimente manchmal noch schwieriger durchzuführen als Befragungen. Angenommen, Sie möchten den Blutdruck untersuchen. Was tun Sie, wenn mitten in Ihrer Studie Ihr Blutdruckmessgerät kaputt geht? Was passiert, wenn jemand die Messung vorzeitig abbrechen möchte? Wie gehen Sie damit um, wenn während des Experiments die Teilnehmenden oder die Untersuchenden abgelenkt werden? Oder wenn eine Blutuntersuchung genau eine Stunde nach der Einnahme eines Medikaments erfolgen muss, aber keine Vene für die Blutentnahme zugänglich ist? Das sind nur ein paar Beispiele dafür, was alles schiefgehen kann, wenn Sie Daten für Experimente erheben. Auf so etwas müssen Sie sich einstellen, und wenn solche Probleme auftreten, müssen Sie sie erkennen.

Nachdem Sie Kapitel 16 (zu Stichproben und Umfragen) und Kapitel 17 (zu Experimenten) durchgearbeitet haben, sollten Sie in der Lage sein, unverzerrte Stichproben auszuwählen und Daten ohne Verzerrung zu erheben, sowie ein Gespür dafür entwickeln, wie sich vermeintliche Kleinigkeiten nachhaltig auf Ergebnisse auswirken können. Außerdem können Sie dann die Glaubwürdigkeit statistischer Ergebnisse beurteilen und sich überzeugend dazu äußern, da Sie wissen, wovon Sie reden.

Sinnvolle Zusammenfassungen erstellen

Nachdem verlässliche Daten erhoben wurden, besteht der nächste Schritt darin, diese sinnvoll zusammenzufassen, um sich einen Überblick zu verschaffen. In der Statistik gibt es dazu zwei wichtige Verfahren: Zahlen und Bilder (also Diagramme und Schaubilder).

Beschreibende Statistik



Als *beschreibende Statistik* bezeichnet man in erster Linie Zahlen, die einen Datensatz oder eine Menge von Daten mit deren wichtigen Merkmalen beschreiben:

- ✓ Wenn es sich um kategoriale Daten handelt (für die Personen in Gruppen eingeteilt werden, beispielsweise nach Geschlecht oder Parteizugehörigkeit), werden diese üblicherweise mit der Anzahl der Individuen pro Gruppe (als *Häufigkeit* bezeichnet) oder dem prozentualen Anteil der Individuen pro Gruppe (als *relative Häufigkeit* bezeichnet) zusammengefasst.
- ✓ *Numerische Daten* stehen für Messungen oder Zählwerte, bei denen die Zahlen eine numerische Bedeutung haben (wie Körpergröße und Körpergewicht). Bei numerischen Daten kann mehr als die Anzahl oder der prozentuale Anteil zusammengefasst werden. Dazu gehören unter anderem:
 - Maße für die Mitte (die angeben, wo sich der »Mittelpunkt« der Daten befindet)
 - Maße für die Streuung (die angeben, wie unterschiedlich die Daten sind oder wie nah sie am Mittelwert liegen)
 - gegebenenfalls auch Zahlen, die die Beziehung zwischen zwei Variablen angeben (wie Körpergröße und Körpergewicht)



Manche Größen der beschreibenden Statistik sind in bestimmten Situationen geeigneter als andere. Häufig ist das arithmetische Mittel nicht das beste Maß für die Mitte eines Datensatzes, sondern eher der Median. Auch ist die Standardabweichung nicht das einzige Maß für Variabilität, der Interquartilsabstand hat ebenfalls seine guten Seiten. Sie müssen die verschiedenen Kenngrößen der beschreibenden Statistik kennen sowie sie interpretieren und bewerten können, die Ihnen täglich begegnen, und auch wissen, in welchen Fällen (beziehungsweise bei welchen Fragestellungen) eine andere als die angegebene Größe besser wäre.

Die häufigsten Größen der beschreibenden Statistik werden in Kapitel 5 berechnet, interpretiert, verglichen und bewertet. Dazu zählen für kategoriale Daten absolute Häufigkeiten und relative Häufigkeiten (Anzahlen und Prozente) sowie Durchschnittswert (arithmetisches Mittel), Median, Standardabweichung, Perzentile und deren Kombinationen für numerische Daten.

Diagramme und Schaubilder

In Diagrammen und Schaubildern werden Daten grafisch zusammengefasst. Sie sind so organisiert, dass sie Ihnen entweder blitzschnell einen groben Überblick über die Daten vermitteln oder ein bestimmtes Ergebnis genauer darstellen. In unserer Welt mit ihren schnellen Infos und superkurzen O-Tönen sind Diagramme und Schaubilder sehr beliebt. Die meisten davon stellen deutlich, fair und sinnvoll dar, was ausgesagt werden soll. Allerdings gibt es auch Beispiele, in denen etwas zu viel künstlerische Freiheit am Werk ist, was dann zu irreführenden oder einfach falschen Diagrammen und Schaubildern führt.



Im Kapitel 4 stelle ich die wichtigsten Arten von Diagrammen und Schaubildern für die Zusammenfassung von kategorialen oder numerischen Daten vor. Sie erfahren dort, wie man diese Grafiken erstellt, welchem Zweck sie dienen und wie man sie interpretiert. Außerdem zeige ich Ihnen, in welcher vielfältiger Weise Diagramme und Schaubilder irreführend sein können und wie Sie mögliche Probleme schnell erkennen. Sie sollten in diesen Problemfällen stutzig werden und auch benennen können, was Ihr Misstrauen erregt hat. Ein paar Beispiele:

- ✓ Einige der einfacheren Darstellungen für kategoriale Daten sind Tortendiagramme und Balkendiagramme, die eine Variable in ihre Bestandteile aufteilen, beispielsweise das Geschlecht oder die Apps, die Jugendliche auf ihren Smartphones haben. So kann ein Balkendiagramm beispielsweise Meinungen in Gestalt von fünf Balken abbilden, die von »stimme überhaupt nicht zu« bis »stimme vollkommen zu« reichen. In Kapitel 6 finden Sie alle wichtigen Infos dazu, wie Sie diese Diagramme erstellen, interpretieren und, noch wichtiger, prüfen, ob sie der dargestellten Sache gerecht werden. Sie werden sich wahrscheinlich wundern, was man bei einem einfachen Balkendiagramm alles falsch machen kann.
- ✓ Für numerische Daten wie Größe, Gewicht, Zeit oder Mengen braucht man andere Arten von Diagrammen. Zu ihrer Zusammenfassung werden Histogramme und Box-Plots verwendet, die sehr informativ sein und Datenmengen wunderbar zutreffend abbilden können. Aber natürlich können auch diese Darstellungen irreführend sein – aus Unkenntnis oder sogar mit Absicht. (Die Hintergründe dazu finden Sie in Kapitel 7.)



Diagramme und Schaubilder begegnen Ihnen tagtäglich – schlagen Sie einfach eine Zeitung auf, und Sie werden mühelos mehrere davon finden. Schauen Sie mit Ihren Statistikkenntnissen im Hinterkopf genau hin, um irreführende Darstellungen zu erkennen, bevor Sie sich zu vorschnell zu Meinungen oder Handlungen hinreißen lassen.

Verteilungen bestimmen

Eine *Variable* ist ein Merkmal, das gezählt, gemessen oder kategorisiert wird. Beispiele für Variable sind Geschlecht, Alter, Körpergröße, Körpergewicht und die Zahl der Haustiere, die jemand sich hält. Eine *Verteilung* ist eine Auflistung der möglichen Werte einer Variablen

(oder Intervalle von Werten) zusammen mit der Angabe, wie häufig (oder mit welcher Dichte) diese auftreten. Das Geschlecht unter Neugeborenen in Deutschland verteilte sich beispielsweise 2018 auf 51,3 Prozent männliche und 48,7 Prozent weibliche Kinder.



Für verschiedene Arten von Variablen gibt es auch verschiedene Verteilungen. In Einführungskursen in die Statistik werden die folgenden drei am häufigsten behandelt, die auch im wahren Leben viele Anwendungen haben:

- ✓ Eine Variable, die die Anzahl der Erfolge bei einer bestimmten Zahl von Versuchen misst (beispielsweise die Anzahl der Menschen, die nach Einnahme eines bestimmten Medikaments gesund wurden), hat eine *Binomialverteilung*.
- ✓ Eine Variable, die Werte annimmt, die sich anhand einer sogenannten Glockenkurve verteilen, wie die Noten in Klassenarbeiten, hat eine *Normalverteilung*.
- ✓ Eine Variable, die auf den Mittelwerten von Stichproben beruht und für die Sie nur wenige Daten haben, wie beispielsweise zehn Teilnehmer in einer Studie zu einem Gewichtsabnahme-Programm, sollten Sie sinnvollerweise mit einer *t*-Verteilung beschreiben.

In Bezug auf Verteilungen müssen Sie herausfinden, welche Verteilung eine bestimmte Variable hat, wie Sie – über die Häufigkeitsfragestellung hinausgehend – Wahrscheinlichkeiten bestimmen und wie Sie feststellen, was der (arithmetische) Mittelwert und die Standardabweichung der Ergebnisse über einen langen Zeitraum wären. (Wenn im Übrigen im Buch verkürzend nur von »Mittelwert« die Rede ist, ist damit das arithmetische Mittel, also der Durchschnittswert, gemeint.) Damit Sie sich diese Kenntnisse erarbeiten können, habe ich gleich drei Kapitel für Sie geschrieben, nämlich eines pro genannter Verteilung: Kapitel 8 widmet sich der Binomialverteilung, Kapitel 9 der Normalverteilung und Kapitel 10 der *t*-Verteilung.



Falls Sie gerade einen Einführungskurs in Statistik besuchen (oder überhaupt einen Statistikkurs), wissen Sie schon, dass zu den schwierigsten Inhalten die Stichprobenverteilung und der zentrale Grenzwertsatz gehören (zwischen denen ein enger Zusammenhang besteht). Kapitel 11 führt Sie Schritt für Schritt durch diese Themen. Sie verstehen dann, was eine Stichprobenverteilung ist, wofür man sie braucht und wie sie die Grundlage für Datenanalysen in Gestalt von Hypothesentests und Konfidenzintervallen bildet (mit der Datenanalyse befasst sich der nächste Abschnitt noch genauer). Wenn Sie den zentralen Grenzwertsatz verstehen, können Sie auch schwierige Aufgaben wesentlich einfacher lösen. Die Grundlagen für dieses Verständnis liefert Ihnen Kapitel 11.

Fachgerechte Analysen durchführen

Wenn Sie jetzt also Ihre Daten erhoben und mit Zahlen und Bildern beschrieben haben, fängt der eigentliche Spaß an: Sie begeben sich dann in die Blackbox, die man als *statistische Analyse* bezeichnet. Wenn Ihre Studie fachgerecht angelegt wurde, lassen sich die

ursprünglichen Fragen mit der geeigneten Analyse beantworten. Dem Wort »geeignet« kommt hier besondere Bedeutung zu.



Analysen gibt es in vielen Formen und Ausführungen. Es ist unglaublich wichtig, die jeweils zur Situation passende Analyse zu wählen, die Ergebnisse richtig zu interpretieren, die Grenzen des benutzten Verfahrens zu kennen und auch beurteilen zu können, ob andere Personen jeweils die richtige Analyse gewählt und eine zulässige Schlussfolgerung daraus gezogen haben.

In diesem Buch finden Sie alle Informationen und Werkzeuge, die Sie brauchen, um Daten mit den Methoden zu analysieren, die am häufigsten in Einführungsveranstaltungen zur Statistik unterrichtet werden: Konfidenzintervalle, Hypothesentests, Korrelation und Regression und die Analyse von Kontingenztafeln (auch Zweiwegtabellen genannt). Dieser Abschnitt gibt Ihnen schon mal einen kurzen Überblick darüber.

Fehlergrenze und Konfidenzintervalle

Häufig begegnen Ihnen Statistiken, mit deren Hilfe versucht wird, Werte zu schätzen, die sich auf eine sehr große Grundgesamtheit beziehen – Umfrageergebnisse sind ein typisches, tägliches Beispiel dafür. Aus den Medien erfahren Sie den durchschnittlichen Benzinpreis in Deutschland, was die Bevölkerung von der Kanzlerin hält oder wie viele Stunden Jugendliche pro Tag oder Woche im Internet verbringen.

Aber niemand kann für diese Angaben eine einzige Zahl nennen und behaupten, dass diese einen zutreffenden Wert für die gesamte Grundgesamtheit darstellt, sofern er (oder sie) nicht jedes einzelne Mitglied dieser Grundgesamtheit befragt oder vermessen hat. So lesen Sie beispielsweise in der Zeitung, dass 60 Prozent der Deutschen mit Angela Merks Haltung in der (immer noch schwelenden) Eurokrise einverstanden sind, aber Sie hat niemand gefragt, wie also könnten alle Deutschen befragt worden sein? Und weil eben nicht alle Deutschen gefragt wurden, wissen Sie, dass eine einzige Zahlenangabe nicht ausreicht.

Zu entsprechenden statistischen Aussagen gelangen Meinungsforscher in Wirklichkeit, indem sie eine Stichprobe aus der Grundgesamtheit befragen (beispielsweise ruft das Institut für Demoskopie Allensbach 1.000 zufällig ausgewählte Personen an), dann die Ergebnisse der Stichprobe analysieren und daraus Rückschlüsse auf die gesamte Grundgesamtheit (beispielsweise alle Deutschen) ziehen.



Daran ist wichtig zu verstehen, dass sich die Ergebnisse von Stichprobe zu Stichprobe unterscheiden – wenn Sie die Befragung mit derselben Anzahl von Personen wiederholen, aber andere Menschen anrufen, erhalten Sie etwas andere Werte. Diese Variabilität muss mit dem Ergebnis angegeben werden (was aber häufig nicht passiert). Die statistische Größe, die angibt, wie genau ein Stichprobenergebnis ist, bezeichnet man als *Fehlergrenze*. In diesem Zusammenhang bedeutet *Fehler* nicht, dass tatsächlich ein Fehler gemacht wurde. Vielmehr verweist die Fehlergrenze darauf, dass eben nicht die ganze Grundgesamtheit untersucht wurde und dass es daher eine gewisse Abweichung zwischen dem Stichprobenergebnis und dem wahren Wert geben wird, den Sie für die Grundgesamtheit schätzen wollen.

Angenommen, jemand stellt fest, dass 60 Prozent der 1.000 befragten Personen mit der Haltung der Kanzlerin in der Eurokrise einverstanden sind und dass die Fehlergrenze plus oder minus 2 Prozentpunkte beträgt. Dieses Endergebnis, das man dann als eine Spannweite wahrscheinlicher Werte angeben würde, die sich von 58 Prozent bis 62 Prozent erstreckt, bezeichnet man als *Konfidenzintervall*.



Wir alle haben es mit Ergebnissen zu tun, für die es Fehlergrenzen und Konfidenzintervalle gibt, und angesichts der heutigen Datenflut müssen viele Menschen auch am Arbeitsplatz damit umgehen. Sie sollten daher wissen, welche Faktoren die Fehlergrenze beeinflussen (beispielsweise der Stichprobenumfang) und was man für ein gutes Konfidenzintervall braucht. Bei Bedarf sollten Sie auch selbst in der Lage sein, ein Konfidenzintervall zu berechnen.

In Kapitel 12 erfahren Sie alles, was Sie über die Fehlergrenze wissen müssen: alle ihre Bestandteile, was damit gemessen wird und was nicht und wie man sie in verschiedenen Situationen berechnet. Kapitel 13 führt Sie Schritt für Schritt durch Formeln, Berechnungen und Interpretationen von Konfidenzintervallen für den Mittelwert einer Grundgesamtheit, einen Anteil an einer Grundgesamtheit und die Differenz zwischen den Mittelwerten oder Anteilen zweier Grundgesamtheiten.

Hypothesentests

Ein wesentlicher Bestandteil von Forschungsstudien sind Hypothesentests. Ein *Hypothesentest* ist ein Verfahren, bei dem Sie Daten verwenden, um eine Behauptung über eine Grundgesamtheit zu bestätigen oder zu widerlegen. So könnte ein Politiker beispielsweise behaupten, dass 80 Prozent der Menschen in seinem Bundesland mit ihm einer Meinung sind – aber stimmt das auch? Oder ein Pizzadienst gibt an, alle Pizzen in höchstens 30 Minuten zu liefern – aber ist das wirklich so? In der Medizinforschung werden ständig Hypothesentests durchgeführt, um herauszufinden, ob ein bestimmtes Medikament wirkt, ob ein neues Medikament weniger Nebenwirkungen hat als ein altes oder welche Diät am wirksamsten zur Gewichtsabnahme verhilft.



Am häufigsten werden die folgenden Aspekte einer Grundgesamtheit geprüft:

- ✓ Der (arithmetische) Mittelwert der Grundgesamtheit (Ist die durchschnittliche Lieferzeit von 30 Minuten realistisch?)
- ✓ Ein Anteil an der Grundgesamtheit (Stimmen wirklich 80 Prozent der Bevölkerung der Meinung eines bestimmten Politikers zu oder doch weniger?)
- ✓ Die Differenz zwischen den (arithmetischen) Mittelwerten oder Anteilen an zwei Grundgesamtheiten (Stimmt es, dass man mit dieser neuen Diät im Schnitt fünf Kilo mehr abnimmt als mit einer beliebten älteren Methode? Stimmt es, dass dieses neue Medikament den Blutdruck um 10 Prozent stärker senkt als das bisherige?)



Hypothesentests werden in allen möglichen Bereichen eingesetzt, die sich auf Ihr Leben auswirken: medizinische Studien, Werbung, Wählerbefragungen und praktisch überall, wo Vergleiche anhand von Mittelwerten oder Anteilen gezogen werden. Im Berufsleben werden Hypothesentests beispielsweise im Marketing häufig eingesetzt, wo man herausfinden möchte, ob eine bestimmte Art von Anzeige wirksamer ist als eine andere oder ob eine bestimmte Personengruppe mehr oder weniger von einem Produkt kauft als im Vorjahr.

Dabei erfahren Sie oft nur die Schlussfolgerung aus einem Hypothesentest (beispielsweise, dass das neue Medikament wirksamer ist und weniger Nebenwirkungen hat als das, das Sie aktuell nehmen), aber nichts über die Methoden, mit denen jemand zu diesen Schlussfolgerungen gelangt ist. Kapitel 14 befasst sich mit allen Einzelheiten und Voraussetzungen für Hypothesentests, damit Sie diese selbst durchführen und entsprechende Aussagen besser beurteilen können. Kapitel 15 geht dann nahtlos zu Schritt-für-Schritt-Anleitungen dafür über, wie Sie in bestimmten Situationen einen Hypothesentest anlegen und durchführen (für den Mittelwert einer Grundgesamtheit, für einen Anteil an einer Grundgesamtheit, für die Differenz zwischen den Mittelwerten zweier Grundgesamtheiten und so weiter).

Nachdem Sie die Kapitel 14 und 15 durchgearbeitet haben, werden Sie wesentlich besser für Situationen gerüstet sein, in denen Sie wissen möchten, an welche Zielgruppe sich Ihr Marketing richten sollte, welches Autoreifenmodell länger hält, ob eine bestimmte Diät wirksam ist, oder auch, wenn Sie vor einer wichtigen Entscheidung stehen, beispielsweise, welches Operationsverfahren Sie wählen sollten.

Korrelation, Regression und Kontingenztafeln

Ein wichtiges Forschungsziel besteht häufig darin, Verbindungen zwischen Variablen zu erkennen. Beispiele gefällig?

- ✓ Welche Art der Lebensführung erhöht oder senkt das Krebsrisiko?
- ✓ Welche Nebenwirkungen hat ein bestimmtes neues Medikament?
- ✓ Kann ich meinen Cholesterinspiegel mit diesem neuen pflanzlichen Präparat senken?
- ✓ Führt langes Surfen im Internet zu einer Gewichtszunahme?

Das Erkennen von Verbindungen zwischen Variablen erlaubt es, bessere Medikamente und Behandlungen zu entwickeln, lässt Marketingfachleute ihre Zielgruppen effektiver ansprechen und gibt Politikern die Informationen, auf die sie ihre Argumente für oder gegen bestimmte politische Maßnahmen stützen.



Die Suche nach Beziehungen zwischen Variablen ist inzwischen ein Riesengeschäft, und so gibt es dafür unglaubliche Mengen an statistischen Ergebnissen. Aber können Sie auch erkennen, welche davon richtig und welche falsch sind? Viele Entscheidungen werden anhand solcher Studien getroffen, und es ist wichtig zu wissen, welche Standards erfüllt sein müssen, damit ein Ergebnis als glaubwürdig gelten kann. Das gilt ganz besonders für Aussagen über Ursachen und Wirkungen.

Kapitel 18 beschäftigt sich ausführlich mit diesem Thema. Von den Details und Feinheiten der grafischen Darstellung zweier numerischer Variablen (wie Dosierung und Blutdruck) über das Finden und Interpretieren von *Korrelationen* (der Stärke und Richtung einer linearen Beziehung zwischen x und y) und das Ermitteln der Geradengleichung, die am besten an die Daten angepasst ist (und wann dies zulässig ist), bis hin zu Erklärungen, wie Sie diese Ergebnisse nutzen, um Vorhersagen über eine Variable aufgrund der anderen Variablen zu machen (was man *Regression* nennt). Sie bekommen auch das nötige Werkzeug an die Hand, um zu untersuchen, wann eine Linie Daten gut wiedergibt und wann nicht und welche Schlüsse Sie ziehen dürfen (und welche nicht), wenn die Linie gut angepasst ist.

Methoden zur Aufdeckung und Beschreibung der Zusammenhänge zwischen zwei kategorialen Variablen (wie die Zahl der pro Tag genommenen Tabletten und das Auftreten oder Ausbleiben von Schwindel) behandle ich ausführlich in Kapitel 19. Dort finden Sie auch Informationen dazu, wie man Daten in *Kontingenztafeln* (*Zweiwegtabellen*) anordnet (in denen die möglichen Werte der einen Variable in den Zeilen und die Werte der anderen Variablen in den Spalten stehen), die Ergebnisse interpretiert, die Daten aus solchen Tabellen nutzt, um Beziehungen zu erkennen, und die Daten auf Unabhängigkeit prüft. Und wie überall in diesem Buch beschreibe ich auch dort Strategien, mit denen Sie die Glaubwürdigkeit solcher Analysen kritisch prüfen können.

Glaubwürdige Schlussfolgerungen ziehen



Für die statistische Analyse verwendet man in der Wissenschaft Statistiksoftware, die auf Formeln beruht. Aber die Formeln wissen nicht, ob sie gerade sachgerecht angewendet werden, und sie warnen Sie auch nicht, wenn die Ergebnisse inhaltlich nicht passen. Ein Computer kann Ihnen am Ende eben nicht sagen, was ein Ergebnis bedeutet – das müssen Sie schon selbst herausfinden. Überall in diesem Buch zeige ich auf, welche Arten von Schlussfolgerungen Sie nach einer Analyse ziehen dürfen und welche nicht. Die folgenden Abschnitte geben Ihnen eine Einführung in das Ziehen angemessener Schlussfolgerungen.

Überbordende Ergebnisse

Zu den häufigsten Fehlern bei Schlussfolgerungen zählt es, die Ergebnisse zu übertreiben oder auf eine größere Gruppe zu verallgemeinern, als eigentlich in der Studie untersucht wurde. So möchte eine Professorin beispielsweise wissen, welche Werbespots im Vorabendprogramm am beliebtesten sind. Sie befragt also 100 Studierende in ihrer Vorlesung, welcher

Spot ihnen am Abend zuvor am besten gefallen hat. Aus den Antworten bildet sie eine Liste der fünf beliebtesten Spots und schließt daraus, dass diese in Deutschland am beliebtesten sind. In Wirklichkeit erfährt sie so jedoch nur, welche Werbespots *ihre Studierenden* bevorzugen, denn diese sind nicht repräsentativ für das gesamte Fernsehpublikum.

Behauptungen zu Ursache und Wirkung hinterfragen

Eine Situation, in der Schlussfolgerungen oft unzulässig sind, tritt ein, wenn eine Regressionsanalyse ergibt, dass eine Beziehung zwischen zwei Variablen besteht und man daraus schließt, dass es sich um eine Ursache-Wirkung-Beziehung handelt.

Angenommen, eine Forscherin führt eine Gesundheitsumfrage durch und stellt fest, dass Personen, die täglich Vitamin C einnehmen, weniger oft angeben, erkältet zu sein, als Personen, die nicht täglich Vitamin C einnehmen. Aufgrund dieser Ergebnisse gibt sie eine Pressemeldung heraus, die besagt, dass Vitamin C Erkältungen vorbeugt.

Zwar ist es durchaus möglich, dass Vitamin C diese vorbeugende Wirkung hat, aber anhand ihrer Daten allein darf die Forscherin das nicht behaupten. Bei ihrer Studie handelte es sich um eine Beobachtungsstudie. Sie hat also keine Kontrollen bezüglich anderer Faktoren durchgeführt, die sowohl mit Vitamin C als auch mit Erkältungen zusammenhängen. Beispielsweise könnte es sein, dass Menschen, die täglich Vitamin C nehmen, allgemein sehr gesundheitsbewusst sind, sich gesund ernähren, sich öfter die Hände waschen und mehr Sport treiben, was sich ebenfalls günstig auf die Vorbeugung von Erkältungen auswirken könnte.



Eine Aussage über Ursache und Wirkung können Sie, strenggenommen, immer erst treffen, wenn Sie ein kontrolliertes Experiment durchführen und Schlüsse anhand beobachteter Beziehungen ziehen. (Über Experimente habe ich mich weiter oben in diesem Kapitel schon ausführlicher geäußert.)

Detektivischen Spürsinn statt bloßer Skepsis entwickeln

In der Statistik geht es um viel mehr als um Zahlen. Wenn Sie Statistik wirklich verstehen wollen, müssen Sie lernen, wie man zulässige Schlussfolgerungen aus Daten und Datenanalysen zieht. Sie müssen clever genug sein, um nicht alles zu glauben, was Sie hören oder lesen. Zuerst sollten Sie immer fragen, woher die Informationen stammen, was mit den Daten gemacht wurde und wie bestimmte Schlüsse gezogen wurden. Diese Punkte erwähne ich zwar auch im übrigen Buch immer wieder, aber in Kapitel 20 konzentriere ich mich dann noch einmal ganz auf diesen wichtigen Punkt. Dort zeige ich Ihnen zehn Wege auf, wie Sie detektivischen Spürsinn mit Köpfchen entwickeln und Fehler erkennen, die die Medien, aber auch Wissenschaftlerinnen und Wissenschaftler häufig machen.



Statistik mit reiner Skepsis oder sogar Zynismus zu begegnen, ist einfach, besonders, wenn man mal hinter die Kulissen geschaut hat. Aber das wollen Sie ja nicht. Sie können in Statistiken jede Menge sinnvoller Informationen finden und zu Ihrem Vorteil nutzen. Mit einer ordentlichen Portion gesunder Skepsis können Sie sich Folgendes vornehmen:

1. Werden Sie zur aufgeklärten Nutzerin oder zum aufgeklärten Nutzer der statistischen Informationen, die Ihnen täglich begegnen.
2. Werden Sie zur ersten Ansprechpartnerin oder zum ersten Ansprechpartner für Statistikfragen in Ihrem Umfeld. Wenn eine Fragestellung Sie überfordert, können Sie immer noch auf Statistikfachleute verweisen, aber Statistikgrundkenntnisse sind sehr gefragt und viel zu wenig verbreitet. Machen Sie sich unentbehrlich!

Wenn Sie dieses Buch lesen und die Informationen darin aktiv nutzen, werden Sie statistische Ergebnisse schon bald mit Selbstvertrauen bewerten können. Sie können dann eigene statistische Untersuchungen auf überzeugende Weise durchführen. Und Sie werden bereit sein, ein anstehendes Projekt im Büro anzupacken, eine nervige Werbung kritisch zu hinterfragen oder die nächste Statistik-Klausur mit Bravour zu meistern!

