

Kohlhammer

Frank Siegmann

Statistik verstehen, nicht rechnen

Band 2: Schließende Statistik

Verlag W. Kohlhammer

Für Marvin, Barbara und Simon

1. Auflage 2017

Alle Rechte vorbehalten

© W. Kohlhammer GmbH, Stuttgart

Gesamtherstellung: W. Kohlhammer GmbH, Stuttgart

Print:

ISBN 978-3-17-031017-9

E-Book-Formate:

pdf: ISBN 978-3-17-031018-6

epub: ISBN 978-3-17-031019-3

mobi: ISBN 978-3-17-031020-9

Für den Inhalt abgedruckter oder verlinkter Websites ist ausschließlich der jeweilige Betreiber verantwortlich. Die W. Kohlhammer GmbH hat keinen Einfluss auf die verknüpften Seiten und übernimmt hierfür keinerlei Haftung.

Inhaltsverzeichnis

Auch dieses Vorwort sollte man lesen	7
0 Statistik ist überall – eine Einführung	11
0.0 Prolog: Binomialkoeffizient und Co.	14
0.1 Warum man nicht einfach nur zählen kann: die Abgrenzung der Schließenden von der Beschreibenden Statistik	16
0.2 Warum man manchmal nicht alle Daten untersuchen kann: das Ziel der Schließenden Statistik	18
0.3 Nur ein Teil vom Ganzen: Warum man Stichproben zieht ...	24
0.4 Was man denn nun wirklich untersucht: Repräsentativität der Daten	28
0.5 Wer mit wem: Kombinatorik und Stichprobenmodelle	33
1 So wird es vermutlich sein – Wahrscheinlichkeiten	41
1.1 Was denn so vorkommen kann: Elementarereignisse und Ereignisse	43
1.2 Objektiv, ausprobieren oder subjektiv: Arten der Wahrscheinlichkeitsrechnung	44
1.3 Unabhängig oder nicht: Rechnen mit Wahrscheinlichkeiten ...	50
1.4 Wahrscheinlichkeiten in der Praxis	57
2 Was es alles so geben kann – denkbare Verteilungsfunktionen	60
2.1 Welche Ergebnisse grundsätzlich möglich sind: die Zufallsvariable	62
2.2 Wie ist denn hier die Mitte und die Streuung: Erwartungswert und Varianz	71
2.3 Tot oder nicht tot: die Binomialverteilung	74
2.4 Es gibt mehr Möglichkeiten als nur gut und böse: die Multinomialverteilung	78
2.5 Tausend Mann im Stadion und ich krieche den Ball vorn Kopp: die Poisson-Verteilung	80
2.6 Die Grundgesamtheit wird kleiner: die Hypergeometrische Verteilung	82
2.7 Letztendlich ist alles wunderbar normal: die stetige Normalverteilung	84

2.8	Es geht noch normaler: die Standardnormalverteilung	94
2.9	Verteilungen in der Praxis	105
3	Wie sieht es in der Grundgesamtheit aus – Schätzverfahren	110
3.1	Wann sie gut ist: Eigenschaften einer Schätzfunktion	113
3.2	Zwischen diesen Grenzen müssen sie liegen: Intervallschätzungen von Parametern der Grundgesamtheit ...	119
3.3	Schätzverfahren in der Praxis	131
4	Die Vermutung stimmt, oder nicht – Hypothesentests	134
4.1	Ob das auch in der Grundgesamtheit so ist: Parametertests...	136
4.2	So sind die Werte verteilt, oder nicht: Anpassungstests	139
4.3	Das hat doch nichts miteinander zu tun: Unabhängigkeitstest	144
4.4	Hypothesentests in der Praxis	151
5	Epilog: Das Ziegenproblem und einige Schlussbemerkungen zur Schließenden Statistik	156
	Anhang 1: Einige Berechnungen zur Kombinatorik	159
	Anhang 2: Einige Berechnungen von Wahrscheinlichkeiten	163
	Anhang 3: Einige vollständige Berechnungen von Verteilungsfunktionen	165
	Anhang 4: Einige vollständige Parameterschätzungen	169
	Anhang 5: Zwei vollständige Hypothesentests	171
	Stichwortverzeichnis	175

Auch dieses Vorwort sollte man lesen

Wer macht sich eigentlich die Mühe, all die unsinnigen Wahrscheinlichkeiten auszurechnen? Die für den Tod durch einen umstürzenden Getränkeautomaten (sie beträgt angeblich 1:112.000). Die, von einem Meteorit erschlagen oder die, vom Blitz getroffen zu werden. Würde es unser Lebensgefühl verbessern, wenn die Wahrscheinlichkeit, im Lotto den Hauptgewinn abzusahnen, nicht 1:13 Millionen, sondern mit 1:6,5 Millionen doppelt so hoch wäre? Beruhigt es irgendjemanden, dass es 1888mal wahrscheinlicher ist im Lotto zu gewinnen statt ein zweites Mal vom Blitz getroffen zu werden oder gar beides (ein Kanadier wurde angeblich zweimal vom Blitz getroffen und gewann im Lotto 1 Million, Wahrscheinlichkeit 1:2,6 Billionen)? Oder dass es 14000mal wahrscheinlicher ist, auf der Toilette zu sterben als durch einen Blitz? Und dass der Tod durch ein Produkt für Rechtshänder dreimal so tödlich für Linkshänder sein kann (das liegt vor allem an den Motorsägen)? Oder wie wäre es mit der vierfachen Wahrscheinlichkeit, ins Krankenhaus zu kommen, weil man Schwierigkeiten beim Anziehen der Hose hatte (beim Hose anziehen zu sterben ist mit 1:30 Millionen auch nicht zu unterschätzen)? Solche extremen Wahrscheinlichkeiten zu berechnen ist vielleicht spaßig, aber letztendlich unsinnig. Alles demmaßen unwahrscheinlich, dass sich eine Berechnung eigentlich verbietet.

Die Antwort auf die Frage, warum man die in Band 1 dargestellte Beschreibende Statistik braucht, fiel mir nicht schwer. Die dort erworbenen Kenntnisse ließen uns viele Aussagen sinnvoll hinterfragen. Ein Beispiel war die Berechnung des Preisniveaus. Eine Kennziffer der Lebenshaltungskosten in Deutschland sagt etwas darüber aus, was unser Geld am Ende des Jahres noch wert ist, wie viel wir also real verlieren. Ein arithmetisches Mittel, das die durchschnittlichen Geburten pro Frau angibt, kann für Prognosen verwendet werden, wie viele von uns es (bei sonst gleichen Rahmenbedingungen, also ohne Krieg, Pest, Pandemien) in 30 Jahren geben wird und ob es für die Rentenzahlungen reicht. Und ob es einen Zusammenhang zwischen dem Austritt von Ozon aus Laserdruckern und Lungenkrebs geben könnte, interessiert vielleicht viele Büroangestellte.

Warum man aber die Schließende Statistik braucht, konnte selbst ich nicht immer ausreichend beantworten. Vielleicht lag es an den abstrusen Rechenbeispielen. Im Ernst: Wer interessiert sich dafür, ob er aus einer Lostrommel eine Primzahl zieht? Wer möchte wissen, wie groß das Intervall ist, das die jährliche Nutzungsdauer im Internet eingrenzt? Reicht es nicht zu wissen, dass sie bei durchschnittlich 34000 Stunden liegt? Wohl nicht. Sonst gäbe es ja nicht so viele Veranstaltungen und so viele Lehrbücher, die sich mit der Schließenden (oder Analytischen oder Induktiven) Statistik beschäftigen würden. Vielleicht liegt es an der Aufbereitung des Stoffes? An den immer gleichen Beispielen, in denen gewürfelt wird, in denen Münzen auf Wappen oder Zahl fallen oder in denen

irgendwelche Warenlieferungen zum Teil fehlerhaft sind. Vielleicht wäre es interessanter, darüber nachzudenken, wie hoch die Überlebenschance ist, wenn Krebs diagnostiziert wurde (30 % von uns, mit steigender Tendenz, werden vermutlich an Krebs sterben). Natürlich hilft diese Wahrscheinlichkeit einem einzelnen Kranken nicht abschließend weiter; was heißt schon 30 %? Aber dass die Mehrheit überlebt, gibt sicher ein bisschen Hoffnung. Läge die Überlebenschance sogar bei 95 %, beruhigt das vielleicht mehr als jedes Beruhigungsmittel es tun könnte. Ein Politiker, der für das Gesundheitswesen zuständig ist, könnte sich durchaus mit der Frage beschäftigen: Ist es besser, die Tabaksteuer drastisch zu erhöhen, so hoch, dass fast keiner mehr raucht? Mit der Folge, dass enorme Defizite im Haushalt entstehen. Oder lassen wir es bei der bisherigen Politik und akzeptieren wir die vielen Todesfälle durchs Rauchen (die Wahrscheinlichkeit, durchs Rauchen zu sterben, beträgt jedes Jahr bis zu 0,5 Prozent; jeder 7. in Deutschland Lebende, im weltweiten Vergleich überdurchschnittlich viele, ist davon betroffen)? Helfen wir mit den Milliarden aus der Tabaksteuer lieber anderen Bedürftigen (an dieser Stelle wollen wir bitte keine Diskussion darüber aufmachen, ob die Einnahmen aus der Tabaksteuer zweckdienlich verwandt werden, das steht auf einem anderen Blatt)?

Schon in der Beschreibenden Statistik gilt, dass alle Erhebungen nur Momentaufnahmen darstellen. Wenn 30 % die Partei XY gewählt haben, dann könnte das Ergebnis am nächsten Tag schon anders aussehen. Wenn auf deutschen Straßen im Jahr z vor allem silberne Autos unterwegs waren, kann das im Jahr $z+1$ anders aussehen. Wenn die Inflationsrate oder die Arbeitslosenquote im März 2,3 % betragen, wird das im April nur selten wieder so sein. In der Beschreibenden Statistik lebten wir also mit Unsicherheiten. Und nicht anders ist es in der hier vorgestellten Schließenden Statistik, auch hier sind die Ergebnisse zwar sehr wahrscheinlich, aber nicht sicher. Deshalb könnte man auch hier generell kritisieren: Was nützen mir Aussagen über Ereignisse, die nur mit einer Wahrscheinlichkeit von X % eintreffen werden? Weil eine wahrscheinliche Aussage besser ist als gar keine! Wenn wir berechnen können, dass die Ausfallwahrscheinlichkeit eines Rechners bei 10 % liegt, dann sollten wir einen zweiten als Backup vorsehen, liegt sie bei 0,1 %, kann man auf einen langen fehlerlosen Betrieb hoffen.

Dass ein Interesse an der Schließenden Statistik durchaus Millionen einbringen kann, zeigt das Beispiel »Sell in May and go away«, ein Ausspruch, den alle Börsianer kennen. Aber ist da wirklich etwas dran? In einer Studie kamen Wissenschaftler zu einem Ergebnis, das viele Aktienhändler überrascht hat. Denn danach gibt es in den meisten Jahren seit Einführung des Deutschen Aktienindex (DAX, seit 1988) nur einen unbedeutend kleinen Unterschied zwischen der Aktienentwicklung von November bis April und der zwischen Mai bis Oktober. Im Winter Aktien zu erwerben, um diese dann im Mai zu verkaufen, bringt also nichts. Mit einer Ausnahme: Der Effekt trat (sogar seit 1927) in den Jahren tatsächlich ein, denen die Vereidigung des neuen US-Präsidenten vorausging. Dann hat beispielsweise der Dow Jones im Winterhalbjahr um durchschnittlich 3,82 Prozent zugelegt, im darauffolgenden Sommer aber um 2,58 Prozent nachgegeben. So etwas zu wissen, und ist es nur etwas wahrscheinlicher als andere Entwicklungen, kann einen Gewinn von Millionen ausmachen. Zumal dieser sogenannte Halloween-Effekt mit schöner Regelmäßigkeit alle 4 Jahre auftritt.

Das im ersten Band Gesagte gilt natürlich auch hier: Wir werden mit Statistiken überhäuft; es gibt scheinbar nichts, für das es keine Statistik gibt, und kein Medium, in dem nicht zahlreiche Statistiken zitiert oder interpretiert werden. Und oft haben wir Nutzer den Eindruck, dass eine Statistik nicht zu uns passt oder willkürlich ist. Die Glaubwürdigkeit der Statistik leidet unter dieser Masse an Informationen. Die Zahlen, die uns angeboten werden, erscheinen uns immer weniger glaubhaft. Wer nicht raucht, täglich Sport treibt und sich gesund ernährt, findet sich in einer Statistik nicht wieder, in der die Wahrscheinlichkeit ausgerechnet wurde, in genau diesem Jahr zu sterben. Außerdem interessieren wir uns häufig erst dann intensiver für Statistiken, wenn wir in einer aktuellen Situation Orientierung suchen. Eine Statistik über die Überlebenswahrscheinlichkeit im Falle einer schweren Krankheit beschäftigt uns. Die Wahrscheinlichkeit, das Rentenalter zu erreichen, interessiert uns (und die Rentenkasse) ebenfalls. Zwar ist uns allen bekannt, dass wir etwa 70, 80 Jahre leben werden, also ca. 10 bis 15 Jahre Rente beziehen werden. Aber lohnt sich die private Altersvorsorge, wenn wir unser Leben lang in einem Bergwerk geschuftet haben (wie z. B. unsere Vorgängergeneration im Ruhrgebiet oder im Saarland) und die Wahrscheinlichkeit, früh an einer Staublunge zu sterben, hoch ist? Bei welchem Professor sind die Klausuren einfach und die Wahrscheinlichkeit zu bestehen deshalb hoch? Das sind durchaus Statistiken, die unser Interesse wecken.

Es gibt einen weiteren Grund, sich intensiver mit der Schließenden Statistik auseinander zu setzen. Offensichtlich können viele von uns Wahrscheinlichkeiten kaum korrekt einschätzen. Wir können einfach nicht beurteilen, ob es sinnvoller ist, Maßnahmen zu ergreifen, die unsere Überlebenschance von 98 % auf 100 % steigern oder, ob wir lieber Maßnahmen ergreifen, die unsere Chance von 70 % auf 90 % verbessern. 100 % scheint erstrebenswert, aber eine Steigerung der Wahrscheinlichkeit um 20 % bringt eindeutig mehr.

Deshalb soll auch dieser 2. Band ein einfach gehaltenes Buch zur Statistik sein, das ohne fundierte Kenntnisse der Mathematik lesbar und verständlich bleibt. Eine Einführung zur Schließenden Statistik also, die sich nicht nur an die Studierenden aller Fachrichtungen richtet, sondern auch den interessierten Laien ansprechen möchte, der bisher vergeblich versucht hat, Statistik zu verstehen. Gleichzeitig soll vor allem die oben angedeutete kritische Haltung aufgegriffen werden. Denn da gerade die Schließende Statistik oftmals sehr abstrakt und sogar mit Integralen daherkommt, möchte mancher sich lieber gar nicht mit ihr beschäftigen. Daran ändert auch nichts, dass in den letzten 5 Jahren 40 % mehr Studierende gezählt wurden. Die, die sich gerne für die Mathematik interessieren, sind leider nicht mehr geworden. Oder anders ausgedrückt: Die Wahrscheinlichkeit, dass ein Studierender sich für Mathematik und Statistik interessiert, ist gleichgeblieben (oder eher sogar gesunken). Und natürlich gibt es auch zur Schließenden Statistik passende Aussagen, die uns den Spaß an ihr verderben wollen. »Wenn Sie mal eine Telefonnummer vergessen haben fragen Sie einen Statistiker, er kann Ihnen eine gute Schätzung liefern« ist so ein Ansatz. Davon wollen wir uns nicht leiten, und schon gar nicht beirren lassen.

Ich versuche daher, in dieser Einführung zu zeigen, dass für das Verständnis der statistischen Methodenlehre die Beschäftigung mit komplizierten mathematischen

Ansätzen, vielen Formeln und Rechenvarianten nicht zwingend notwendig ist. Wieder fast ohne Formeln und ohne Taschenrechner werden wir uns mit dem Thema auseinandersetzen; es geht uns um das grundsätzliche Verstehen, nicht um das Rechnen. Erst im Anhang werden zusammenfassende Aufgaben besprochen, die (überwiegend) nur mit Hilfe eines Taschenrechners gelöst werden können.

Mein ältester Sohn hat das Ganze wieder Korrektur gelesen und mir zahlreiche Hinweise gegeben, wie man etwas besser formulieren könnte. Meine Frau und mein kleiner Sohn haben mir auch diesmal die ungestörte Ruhe in meinem Arbeitszimmer ermöglicht, die man für so ein Werk braucht. Streit darüber gab es zum Glück keinen. Ein Dank geht auch an Alexander Hau für die Erstellung der zahlreichen Grafiken. Und wieder möchte ich mich besonders bei Herrn Dr. Fliegau vom Verlag Kohlhammer bedanken, der auch diesen zweiten Band intensiv gelesen hat und wieder zahlreiche Anregungen und Verbesserungsvorschläge einbrachte. Schade, dass unsere Zusammenarbeit damit erst einmal zu Ende geht.

Bochum, im Mai 2017

Frank Siegmann

0 Statistik ist überall – eine Einführung

Nach 12-mal Rot muss es gelingen:
»ich setz auf Schwarz und werd' gewinnen,
Roulette ist doch ein Kinderspiel,
man muss nur wissen, wie die Kugel fiel«,
denkt Heinz und setzt nun Haus und Hof.

Tatsächlich ist der Heinz recht doof.

Die Kugel hat nicht mitgezählt,
weshalb sie keine Farbe wählt,
sie fällt auf Schwarz und Rot ganz nach Belieben,
entscheidet neu, wo will sie liegen,
sie weiß nicht, was sie vorher tat,
deshalb an Heinz der gute Rat:

Mit einer Chance eins zu eins,
verlierst erneut nur Du, Du Heinz.

Dieses kleine Gedicht von mir wird Sie vermutlich enttäuschen. Leider ist Statistik so und nicht anders. Sie kann uns nicht helfen, beim Roulette zu gewinnen. Es gibt keine Strategie, und das muss ich in dieser Deutlichkeit vorausschicken, mit der man Glücksspiele überlisten kann. Sätze ich dann hier am Schreibtisch und würde über Statistik schreiben?

Bevor wir diesbezüglich in die Tiefe gehen: Auch eine wiederholte Verdoppelung des Einsatzes scheitert früher oder später am Limit, das Sie setzen dürfen. Ganz abgesehen davon verlangt die Bank Bares, haben Sie so viel davon? Rechnen wir nach:

- 1. Fall: 10 Euro gesetzt, verloren, Einsatz verdoppelt, also 20 Euro gesetzt, dann gewonnen. Sie sind wieder da, wo Sie angefangen haben. Und mehr erreichen Sie mit dieser Strategie nicht.
- 2. Fall: 10 Euro gesetzt, gewonnen, nochmal 10 Euro gesetzt, gewonnen. Mit dieser Strategie gewinnen Sie nach 10 Runden 100 Euro. Nicht vergessen: Wenn Sie sehr, sehr viel Glück hatten.
- 3. Fall: 10 Euro gesetzt, verloren, Einsatz verdoppelt, also 20 Euro gesetzt, wieder verloren, Einsatz erneut verdoppelt, also 40 Euro gesetzt, verloren, verdoppelt, 80 Euro ... Schon nach 10 Runden verlieren Sie 5120 Euro. Zugegeben, es passiert selten, dass dieselbe Farbe 10mal kommt (so selten wie Fall 2). Aber es passiert und Sie sind alle

Einsätze los (also 5120 Euro). Und es kommt ein weiteres Problem hinzu: Sie erreichen vielleicht schon vorher das Tischlimit, dürfen also nicht mehr verdoppeln.

- Natürlich gibt es noch weitere denkbare Gewinn- und Verlustfolgen, aber glauben Sie mir, wirklich besser wird es nicht.

Fazit: Entweder Sie fallen auf Ihren Einsatz zurück oder Sie gewinnen ein bisschen oder Sie scheitern am Limit. Sie gewinnen also im ersten Fall nichts, im zweiten ein wenig, aber im dritten verlieren Sie richtig viel.

Wir können schon jetzt festhalten, dass das, was in der Beschreibenden Statistik die relativen Häufigkeiten waren (wie oft fällt die Kugel auf Rot), in der Schließenden Statistik zu Wahrscheinlichkeiten wird (die Kugel fällt mit einer Wahrscheinlichkeit von 0,5 auf Rot). Das Ergebnis einer Untersuchung lautet also nicht, dass 4 von 52 Karten Asse sind (7,7 %), sondern wir können mit einer Wahrscheinlichkeit von 0,077 erwarten, dass die nächste Karte ein Ass ist.

Und genauso wie relative Häufigkeiten erscheinen auch statistische Wahrscheinlichkeiten oft widersprüchlich: »Die Wahrscheinlichkeit, mit dem Auto zu verunfallen, ist wahrscheinlicher als die, dass es mit dem Flugzeug oder der Bahn passiert«. Die Aussage ist korrekt. Aber ihr fehlt wieder das Wesentliche: Wie ist diese Wahrscheinlichkeit zustande gekommen, was ist die Bezugsgröße? Denn was ist mit Wahrscheinlichkeit gemeint? Die Wahrscheinlichkeit, dass wir bei dem Autounfall sterben werden? Oder zählt dazu auch die Wahrscheinlichkeit, im Rollstuhl zu landen? In Deutschland ereignen sich ca. 2,5 Millionen Unfälle im Straßenverkehr; bei über 43 Millionen zugelassenen Autos liegt die Wahrscheinlichkeit, einen Unfall zu haben, also etwas unter 6 %. Ca. 4000 Getötete (mehr als 10 pro Tag!) ergeben einen Anteil von 0,01 %, 70000 Schwerverletzte 0,16 %. Zum Vergleich: Die Wahrscheinlichkeit bei einem Flugzeugabsturz ums Leben zu kommen, beträgt je nach Quelle etwa 0,00003 % (und das sogar weltweit).

Das, was für die Beschreibende Statistik gilt, gilt also auch für die Schließende, sie ist allgegenwärtig. Schlagen wir dazu einfach eine beliebige medizinische Zeitung auf, Ihnen werden auf Anhieb viele statistische Aussagen auffallen. »Im Schnitt beträgt die Wahrscheinlichkeit sich nach einem Zeckenstich mit Borreliose zu infizieren bei 1,5 bis 6 Prozent. Je länger die Zecke saugt, desto höher ist die Wahrscheinlichkeit einer Ansteckung.« »Studien mit Zwillingen weisen darauf hin, dass es genetische Faktoren gibt, die eine Depression wahrscheinlicher machen.« »Man geht davon aus, dass Dengue-Fieber bald auch bei uns in Deutschland, vor allem in wärmeren Gebieten wie zum Beispiel am Oberrhein, übertragen werden kann.« »Fast alle Fälle von Chorea Huntington entstehen durch Vererbung: Eltern mit entsprechendem Erbmerkmal geben dieses an ihre Nachkommen weiter; dabei reicht es aus, dass nur ein Elternteil die Genveränderung in sich trägt. Chorea Huntington wird über einen sogenannten autosomal-dominanten Erbgang an die Kinder übertragen: die Kinder erben das veränderte Gen somit mit einer Wahrscheinlichkeit von 50 Prozent.« Oder eine naheliegende Frage, die viele junge (und vielleicht auch ältere wohlhabende Herren mit einer jüngeren Partnerin) interessiert: Wie groß ist die Wahrscheinlichkeit, schwanger zu werden? Dazu gibt es massenweise Studien, aus denen ich nur zwei

Ergebnisse zitieren will: »Frauen zwischen 35 bis 39 Jahren haben eine 50 % geringere Chance schwanger zu werden.« »Je Zyklus liegt die Wahrscheinlichkeit bei ca. 25 %, wer genau die Tage zählt und die Körpertemperatur misst, kann diese auf fast 40 % steigern.« Auch die in diesem Zusammenhang eher traurige Frage nach der Wahrscheinlichkeit einer Fehlgeburt wurde berechnet und interessiert sicher vor allem Betroffene: »15 % aller Schwangerschaften enden mit einer Fehlgeburt, das Risiko einer zweiten sinkt auf 5 % und 1 % aller Paare erleben sogar 3 davon.« Oder z. B. in folgender Aussage: »Übergewicht erhöht nicht nur das Diabetesrisiko, sondern auch die Möglichkeit, sich einer Knieersatztherapie unterziehen zu müssen: Bei schwer übergewichtigen Frauen ist das Risiko – in beiden Fällen – zwölf Mal höher. Die Wahrscheinlichkeit für hohen Blutdruck ist fünf Mal so groß wie bei normalgewichtigen Frauen. Männer in einer hohen Gewichtskategorie haben acht Mal so oft Diabetes und sechs Mal so oft Knie-Operationen und hohen Blutdruck.« Wahrscheinlichkeiten über Wahrscheinlichkeiten.

Wir können deshalb festhalten: Auch das Themengebiet der Schließenden Statistik ist heute aus vielerlei Gründen und nicht zu Unrecht in allen Wissenschaftsdisziplinen fest verankert. Denn überall werden schließende statistische Methoden eingesetzt: in der Medizin (wenn es darum geht, die Wahrscheinlichkeit einer zweiten Fehlgeburt zu berechnen), in der Psychologie (wenn man die Heilungschancen beim Einsatz verschiedener Psychotherapien vergleicht), in der Wahlnacht (wenn man die Prognosen zu ersten Hochrechnungen verdichtet), bei Banken (Stresstests zur Vorhersage von Reaktionen auf Marktturbulenzen), in der Qualitätskontrolle (Stichprobenverfahren in der Produktion), in der Wirtschaftspolitik (Wahrscheinlichkeit von Steuervermeidungsstrategien), im Gesundheitswesen (bei der Abschätzung von Risiken und Nebenwirkungen, Stichwort seltene und sehr seltene Nebenwirkungen), in den Rechtswissenschaften (Wahrscheinlichkeit eines Rückfalls bei Straftätern).

Grundsätzlich gilt: Wäre alles auf der Welt vorhersehbar, also determiniert, bräuchte man die Schließende Statistik nicht mehr. Ob dieser Zustand durch fortschreitenden Erkenntnisstand jemals erreicht wird, darf wohl bezweifelt werden. Denn letztendlich lassen sich alle Ereignisse unseres Lebens und unserer Umwelt vereinfacht in zwei Kategorien unterteilen, über die sich schon zahlreiche Philosophen den Kopf zerbrochen haben: determinierte und stochastische Ereignisse. Die Entfernung zwischen Erde und Mond ist ein streng determinierter Prozess, man weiß immer sehr genau, wann welcher Abstand gemessen werden kann. Das Gleiche gilt für viele andere physikalische Gesetzmäßigkeiten: den Zeitpunkt des Sonnenaufgangs, die nächsten Gezeiten, die Beschleunigung beim freien Fall, der Verbrauch eines Motors bei konstanter Last. Aber Wahrscheinlichkeiten sind nicht wie Gleichungen, die Zahl ist nicht exakt. Sicheres Wissen gibt es in der Wirtschaft und Soziologie nicht (dort nennt man so etwas auch *ceteris paribus*). Wir müssen mit der Unsicherheit umgehen und diese richtig kalkulieren. Wenn eine amerikanische Radiostation den Wetterbericht verliert, gibt sie Prozentzahlen für bestimmte Arten von Niederschlägen an (z. B. »twenty percent chance of rain«). Dass es im Winter kälter wird, ist vorhersehbar (manchmal scheint selbst das nicht mehr »wahrscheinlich«); wie kalt es tatsächlich wird, entzieht sich aber einer sicheren Vorhersage. Zugvögel fliegen im Herbst in den Süden, ein genaues Datum ist

aber auch hier nicht vorhersehbar. Gleichwohl ist es »besser« zu wissen, dass es dann so ungefähr geschieht.

Schließen wir die einführenden Bemerkungen mit einem netten Spruch, der uns Mut machen soll: »Da lernt man also Mittelwerte und berechnet Wahrscheinlichkeiten und dann steht man trotzdem grübelnd vor dem Backofen und fragt sich, welche der vier Schienen mit der größten Wahrscheinlichkeit die mittlere ist.«

0.0 Prolog: Binomialkoeffizient und Co.

Im Rahmen der Schließenden Statistik gibt es wie in der Beschreibenden Statistik nichts Mathematisches, vor dem man wirklich Angst haben müsste. Auch hier sind »Addition«, »Subtraktion«, »Multiplikation« und »Division« die einzigen arithmetischen Operationen, die man kennen muss. Nur vier Dinge könnten einen Leser dieses Statistikbuches beunruhigen:

$$N^n, N! \text{ bzw. } n! \text{ und } \binom{N}{n}.$$

N soll die Anzahl der Elemente in einer Grundgesamtheit und n die Menge in einer Stichprobe sein (wir werden den Unterschied zwischen Grundgesamtheit und Stichprobe später genauer auflösen). Nehmen wir nun an, dass wir 4 verschiedene Cocktailzutaten im Kühlschrank haben (N) und daraus einfach 2 beliebige auswählen möchten. N^n ist damit $4^2 = 16$; es gibt also 16 Möglichkeiten, 2 von 4 Cocktailzutaten aus dem Kühlschrank zu holen. Aus den 4 Zutaten ABCD also AA, AB, AC, AD, BA, BB, BC, BD, CA, CB, CC, CD, DA, DB, DC, DD.

$N!$ (sprich N Fakultät) bzw. $n!$ bedeutet, dass man das Produkt aller Zahlen von N bzw. n bis 1 bildet. Bei $N = 4$ ist das also:

$$4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24,$$

bei $n = 2$ ist das:

$$2! = 2 \cdot 1 = 2.$$

Der sogenannte Binomialkoeffizient (man spricht » N über n «)

$$\binom{N}{n}$$

ist eine verkürzte Form von

$$\frac{N!}{n! \cdot (N - n)!}$$

bzw.

$$\frac{N \cdot (N - 1) \cdot (N - 2) \cdot (N - 3) \cdot \dots \cdot (N - n + 1)}{n!}.$$

Die verschiedenen Möglichkeiten, 2 verschiedene (!) Zutaten zufällig aus dem Kühlschrank zu holen, sind:

$$\begin{aligned} \binom{N}{n} &= \binom{4}{2} = \frac{N!}{n! \cdot (N - n)!} = \frac{4!}{2! \cdot (4 - 2)!} \\ &= \frac{4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1 \cdot 2 \cdot 1} \end{aligned}$$

bzw.

$$= \frac{4 \cdot 3}{2 \cdot 1} = 6.$$

Zur Verdeutlichung: Es sind die Zutaten AB, AC, AD, BC, BD, CD, da im Gegensatz zu oben nicht zweimal die gleiche Zutat gewählt werden darf.

Auch die Wahrscheinlichkeit, im Lotto 6 Richtige zu erzielen, lässt sich über den Binomialkoeffizienten leicht berechnen; exakt ist sie:

$$\binom{N}{n} = \binom{49}{6}$$

$$= \frac{49 \cdot 48 \cdot 47 \cdot 46 \cdot \dots \cdot 3 \cdot 2 \cdot 1}{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 \cdot 48 \cdot 47 \cdot 46 \cdot 45 \cdot \dots \cdot 3 \cdot 2 \cdot 1}$$

bzw.

$$= \frac{49 \cdot (49 - 1) \cdot (49 - 2) \cdot (49 - 3) \cdot \dots \cdot (49 - 6 + 1)}{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}$$

$$= \frac{49 \cdot 48 \cdot 47 \cdot 46 \cdot 45 \cdot 44}{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 13.983.816.$$

Verlieren wir noch ein Wort zu einer der unbeliebtesten Rechenmethoden, die Schüler jemals gequält hat, dem Integral. Vereinfacht gesagt: Was in der Beschreibenden Statistik das Summenzeichen ist, ist in der Schließenden Statistik (manchmal) das Integral. Manchmal deshalb, weil es nur dann verwendet werden muss, wenn wir stetige Variablen betrachten (wir erinnern uns: stetige Variable können einen Tatbestand beliebig genau messen, während diskrete Merkmale nur bestimmte Werte annehmen können).

Summiert man die Werte $x_1 = 110$, $x_2 = 111$ und $x_3 = 112$ auf, so erhält man im diskreten Fall $\sum x_i = 333$; beim Integral wäre das Ergebnis kein anderes. Addiert man die Körpergrößen von 3 Jugendfußballern, die 110, 111 und 112 cm groß sind, so erhält man also:

$$\sum_{110}^{112} x_i = 333.$$

Allerdings ist dieser Wert auf ganze cm gerundet; in Wirklichkeit ist das erste Kind genau 110,25 oder noch genauer 110,2538 oder noch genauer ...cm groß. Es gibt also keinen exakten Wert, sondern immer nur einen, den man noch genauer fassen könnte. Deshalb arbeitet man mit Wertebereichen, z. B. den Körpergrößen zwischen 110,00000 und 111,00000..., also Bereichen, die sehr viele einzelne Werte enthalten könnten. Und diese Bereiche schließt ein Integral ein. Wenn wir dann wissen wollen, wie groß die Wahrscheinlichkeit dafür ist, dass unter den Kindern nur diejenigen berücksichtigt wurden, die zwischen 110 und 112 groß sind, dann berechnen wir dazu ein Integral, in das alle diese Kinder fallen, egal, wie exakt gemessen wird. Wir berechnen

$$W(x) = \int_{110}^{112} f(x_i) dx.$$

Folgerichtig ist ein Kind, das in einem Jahr (z. B. zwischen seinem 5. und 6. Lebensjahr) von 101,0000.... auf 113,0000.... Zentimeter gewachsen ist, 12,0000.... Zentimeter größer geworden. Deshalb schreibt man dann

$$\int_{101}^{113} f(x_i) dx.$$

Ein Fußballspiel dauert in der Regel 90 Minuten, der Zeitraum zwischen der 50. und 60. Minute lässt sich auch über das entsprechende Integral berechnen; ein Tor, das in diesem Zeitraum fällt, fällt genau genommen in der 56,0523...-ten Minute.

Genauer brauchen wir uns hier über Integrale nicht den Kopf zu zerbrechen, denn zum Glück können wir immer dann, wenn integriert werden muss, auf Tabellen zugreifen, aus denen wir den Wert nur ablesen müssen. Perfekt.

0.1 Warum man nicht einfach nur zählen kann: die Abgrenzung der Schließenden von der Beschreibenden Statistik

Statt von der Abgrenzung der Deskriptiven von der Induktiven, Analytischen Statistik zu sprechen, wollen wir von der Abgrenzung der Beschreibenden von der Schließenden Statistik reden, was mir persönlich besser gefällt, weil es den Unterschied deutlich macht, ohne dass man sich in einer doch nicht so ganz geläufigen (früheren) Fremdsprache zurechtfinden muss.

In der Beschreibenden Statistik untersuchen wir eine statistische Masse vollständig so wie sie ist. Wir berechnen, wie viele Arbeitslose es zu einem bestimmten Stichtag in Deutschland gibt (beispielsweise am 31.3. des Jahres); von der örtlichen Arbeitsagentur bis hin zur Bundesagentur für Arbeit werden dazu an verschiedenen Erhebungsstellen die Zahlen im Einzelnen erhoben und zusammengezählt. Man spricht dann von einer Vollerhebung, weil alle, in diesem Fall alle Arbeitslosen, voll erhoben wurden. Auch die Durchführung einer Bundestagswahl erfordert Methoden der Beschreibenden Statistik. Die Stimmen, die einzelnen Parteien gegeben wurden (Häufigkeiten), werden von den Wahlbüros an die städtischen Wahlämter und von dort an die Landeswahlleiter und letztendlich an den Bundeswahlleiter weitergegeben. Auch das ist eine Vollerhebung, denn wirklich alle Stimmen aller Wahlberechtigten (ob sie nun wählen oder nicht, also auch die der Nichtwähler) werden erfasst und finden zumindest in der Wahlbeteiligung Berücksichtigung.

Gleichzeitig kennen wir im Zusammenhang mit Wahlen aber auch eine andere Methode, eine aus der Schließenden Statistik. Vor der Wahl nämlich und auch gleich nach Schließung der Wahllokale werden Prognosen und Hochrechnungen veröffentlicht. Prognosen darüber, wie viele Stimmen denn welche Partei vermutlich insgesamt erzielen wird. Oft sind diese Zahlen recht nah am Endergebnis. Sie sind nicht auf den Punkt genau, aber doch so gut, dass sie für einen ersten Jubel oder erste Enttäuschungen sorgen. Aber natürlich sind sie auf einem anderen Weg zustande gekommen als die Zahlen einer Vollerhebung. Solche Teilerhebungen sind Stichproben, die hochgerechnet werden. Wenn also in Deutschland 60 Millionen Bürger wahlberechtigt sind und man etwa 20.000 (typische) Bürger sozusagen als Stellvertreter befragt, dann rechnet man das

Ergebnis entsprechend hoch. 20.000 von 60.000.000 sind 0,03 %. Wählen von diesen 20.000 befragten 2.500 die Partei A, dann sind das hochgerechnet $250/0,0003 = 8.333.333$ Wähler insgesamt (= 13,89 % der Stimmen).

Halten wir also fest: In der Schließenden Statistik (manche sagen auch Analytische Statistik, weil man die Daten halt nicht nur beschreibt, sondern auch analysiert) geht es immer um Stichproben aus einer sogenannten Grundgesamtheit. Dabei verwechseln wir bitte nicht die Begriffe Grundgesamtheit und statistische Masse. Wenn man in der Beschreibenden Statistik alle statistischen Einheiten, also alle Schüler einer Schule oder eines Bundeslandes untersucht, dann ist das die statistische Masse. Wenn ich nur einige Klassen auswähle und der Meinung bin, das Ergebnis dieser ausgewählten Klassen stehe stellvertretend für alle Klassen (also die statistische Masse), dann handelt es sich um eine Stichprobe. Dennoch kann die statistische Masse auch Stichprobe sein, wenn man beispielsweise alle statistischen Einheiten (alle Schüler des Bundeslandes) untersucht hat (was in dieser Untersuchung die Zielgröße, die statistische Masse, war) und dieses Ergebnis als Ausgangspunkt einer Hochrechnung auf alle Schüler in Deutschland nutzt. Dann ist die vorherige statistische Masse nun Stichprobe für eine noch größere Grundgesamtheit.

Unterzieht man alle deutschen Banken einem Stresstest, handelt es sich um eine Vollerhebung im Rahmen der Beschreibenden Statistik, die einzelnen Banken sind die statistischen Einheiten, alle zusammen die statistische Masse. Untersucht man nur die Banken und nicht die Sparkassen, erhält man eine (sicher nicht repräsentative) Stichprobe, von der aus man auf die Grundgesamtheit aller Banken und Sparkassen (fehlerhaft) schließen könnte.

Untersucht das Organisationsteam im beschaulichen Südtiroler Ort Sexten die Brenndauer von Wunderkerzen (die man an Sylvester an Luftballone bindet und zu heimeliger Musik zwecks stillem Feuerwerk in die Luft steigen lässt) hat man sicherlich nur einige wenige untersucht (um herauszufinden, ob sie wirklich lange genug für den gewünschten Effekt brennen). Man hat also eine Stichprobe gezogen, eine Frage der Schließenden Statistik. Und man hat eben nicht wie im Rahmen der Beschreibenden Statistik die statistische Masse insgesamt untersucht (alle Wunderkerzen). Hört sich komplizierter an als es ist.

Letztendlich bestimmt die Fragestellung, was statistische Masse und was Grundgesamtheit ist. Ordnen wir die Begriffe einfach den verschiedenen statistischen Methoden zu. Bewegen wir uns im Rahmen der Beschreibenden Statistik, reden wir von statistischen Einheiten und der statistischen Masse, bewegen wir uns auf dem Gebiet der Schließenden Statistik, untersuchen wir Stichproben und Grundgesamtheiten. Eine Stichprobe kann auch die statistische Masse sein, über die man im Rahmen der Beschreibenden Statistik abschließend urteilen möchte. Aber eine statistische Einheit ist immer Teil der statistischen Masse, Stichprobe immer ein Teil der Grundgesamtheit.

Stellen wir uns zum besseren Verständnis noch ein paar Fragen und ordnen sie der Beschreibenden oder Schließenden Statistik zu: