# Statistik

Von
Professor
Henning Mittelbach

R. Oldenbourg Verlag München Wien

**Die Deutsche Bibliothek - CIP-Einheitsaufnahme**

**Mittelbach, Henning:**
Statistik / von Henning Mittelbach. - München ; Wien :
Oldenbourg, 1992
ISBN 3-486-22322-4



Gesamtherstellung: Grafik + Druck, München

ISBN 3-486-22322-4

# Vorwort

Das nachfolgende Skriptum ist im Ausbildungsbereich Ingenieurwissenschaften entstanden, mit Statistik als Nebenfach in ganz unterschiedlichen Fachrichtungen. In allerhöchstens 30 Doppelstunden soll ein ausreichender theoretischer Hintergrund aufgebaut werden; der jeweils fachbezogen praktische Bezug darf ebenfalls nicht zu kurz kommen ... Diese schmale Gratwanderung erfordert Beschränkung auf das Wesentliche:

Die Dreigliederung des Stoffes in Deskriptive Statistik, Einführung in die Wahrscheinlichkeitsrechnung und zuletzt einige Ausführungen zur Testtheorie kann je nach Schwerpunkt vertieft und mit zusätzlichen Aufgaben betont werden. - Vorgegeben ist durch den Text nur das Grundgerüst.

Die Zielgruppe für das vorliegende Buch ist damit umrissen: Es sind all jene Studierenden, die sich mit nur wenig Zeit Basiswissen aneignen möchten und für weitere Fragen dann spezielle Fachliteratur befragen können. Zum Durcharbeiten des Stoffes reichen die Mathematikkenntnisse des Gymnasiums aus; wirklich komplizierte Beweise kommen nicht vor. Umgang mit einfachen Reihen, weiterhin elementare Analysis und Grundkenntnisse zum Integrieren: Mehr wird nicht verlangt ...

Im allgemeinen stößt Statistik auf großes Interesse; Fragen des täglichen Lebens, Nachrichten und Ereignisse verlangen nach Wertung und Kommentierung: Es ist daher eine wichtige Aufgabe jeder Statistikvorlesung, Nachdenklichkeit und Kritikbewußtsein zu fördern. Wacher Verstand und vor allem sichere Interpretationsfähigkeit sind gefordert. Außerdem eignet sich dieses Teilgebiet der Mathematik hervorragend zu historischen Anmerkungen, zu einer Einbettung in die Kulturgeschichte.

Fachliches wie gefühlsmäßiges Vorwissen sind sehr unterschiedlich ausgeprägt. - Statistik muß daher bisherige Erfahrungen korrigieren und präzisieren. Verschiedene Vorurteile und Trugschlüsse sollten durch eine sichere Modellbildung zuverlässig abgebaut bzw. verhindert werden.

Die besondere Bedeutung der Statistik als Anwendungsdisziplin in fast allen Wissenschaften ist unverkennbar; man sagt ihr aber auch nach, daß sie ein hinterhältiges Werkzeug von willfährigen Statistikern im Namen ihrer Auftraggeber sei ... Auch dazu muß eine Vorlesung Stellung beziehen.

Da ein PC mittlerweile schon fast zum Allgemeinbesitz gehört, habe ich im Buch auf einen Tabellenteil verzichtet; die zum Lösen der Aufgaben notwendigen Zahlenwerte lassen sich mit ein paar Programmen maschinell ermitteln: Eine Diskette für DOS-Rechner enthält alle erforderlichen Programme. - Wer mit TURBO Pascal arbeitet, kann die Listings modifizieren und als Grundgerüst für umfangreichere Programme benutzen; für alle anderen Leserinnen und Leser werden die Files zudem praktischerweise compiliert geliefert, so daß man die Programme ganz einfach benutzen kann (siehe Kapitel 28) ...

Besonderer Dank bei der Bewältigung der Schreibarbeit gilt zunächst meinem Freund Hans Härtl (Holzkirchen), der im Entwurf allerhand orthografische und Schreibfehler aufgedeckt hat. Ein im Kapitel 25 als Anwendungsfall vorgestellter (exemplarisch zu verstehender) Intelligenztest wurde mit Unterstützung von Sprachkennern überarbeitet: Frau Fang Hong und weiter Familie Zhao Hong/Yü haben die Beispiele geliefert. - Einer meiner Studenten, Herr Michael Hülskötter (München), hat die Zahlenbeispiele und Aufgaben nachgerechnet und so manche Hinweise auf Fehler und Unklarheiten gegeben. - Schließlich danke ich noch dem Semester IF074W der FHM, das als Testgruppe im SS 92 die Endredaktion des Textes mit immer wieder neuen Entwürfen und Kopien ertragen mußte.

Der Oldenbourg Verlag in München hat die Idee zum Druck sofort aufgegriffen und den Text in seine umfangreiche Fachliteratur zur Statistik aufgenommen. Also wünsche und hoffe ich, daß die oben angedeutete Intention des Buches bei den Lesers ankommt; Kritik ist willkommen ...

Friedberg, an Pfingsten 1992 H. Mittelbach

# Inhaltsverzeichnis

# 1 EINLEITUNG

Der Begriff Statistik leitet sich vom lateinischen Wort *status* für *Zustand* ab; mit dem Aufblühen großer Staatsgebilde schon vor über 2000 Jahren (in den Perserreichen, bei den Ägyptern und Römern, auch in China) entstand bei den Herrschenden das Bedürfnis, Bevölkerungszahlen, Heerscharen, Edelmetallvorräte, Viehbestände und so weiter zahlenmäßig zu erfassen und unter verschiedenen Gesichtspunkten zu katalogisieren. So wird beispielsweise im LUKAS-Evangelium (2, 1-5) eine Volkszählung *) auf Veranlassung von Kaiser AUGUSTUS ausdrücklich erwähnt:

> Es begab sich aber zu der Zeit, daß ein Gebot von dem Kaiser Augustus ausging, daß alle Welt geschätzt würde. Und diese Schätzung war die allererste und geschah zur Zeit, da Cyrenius Landpfleger in Syrien war. Und jedermann ging, daß er sich schätzen ließe, ein jeglicher in seine Stadt. Da machte sich auf auch Joseph aus Galiläa, aus der Stadt Nazareth, in das jüdische Land zur Stadt Davids, die da heißt Bethlehem, darum, daß er von dem Hause und Geschlechte Davids war, auf daß er sich schätzen ließe mit Maria, seinem vertrauten Weibe, die war schwanger. - Und als sie daselbst waren ...

Die hierdurch ausgelöste Wanderungsbewegung zu Zwecken des Zensus hatte einen guten Grund: Doppelzählungen sollten vermieden werden; die Bevölkerung wurde daher am jeweiligen Herkunftsort der Familie registriert. - LUKAS, ein aus heutiger Sicht studierter Mann (Arzt), mag diesen Grund erkannt haben, obwohl seine Hinweise mehr dokumentarischen Charakter im Blick auf die für ihn wichtige Abstammung von Jesus haben.

Die Anfänge der deskriptiven Statistik liegen demnach weit zurück: Schon in früher Zeit wurde eine Grundgesamtheit oder Population oder nur ein Teil - eine Stichprobe - auf ein Merkmal oder gar mehrere hin untersucht; in Listen wird dann markiert, ob der jeweils in Frage stehene Merkmalsträger die interessierende(n) Eigenschaft(en) hat oder nicht. Im allereinfachsten Fall wird nur festgestellt, wieviele Merkmalsträger existieren. - Bis in die Neuzeit hinein wurde ausschließlich auf diese Art Statistik betrieben und noch heute sind diese Verfahren (Volkszählungen!) von grundsätzlicher

---

*) Provinzialzenus in Judäa, ca. 6 v.C. Es fehlt in der Bibel nicht an Warnungen vor Zählungen als Eingriff in die Ordnung Gottes bzw. als Versuch, dessen Absichten zu ergründen. Vgl. 2. SAMUEL, 24,2 oder 1. CHRONIK, 21,2. - Die Folge unerlaubter Zählungen war die Pest, und die fürchtete man im Mittelalter als Strafe Gottes besonders. Jedwelche Zählungen (außer Vermögenserfassungen) waren daher kirchlicherseits verpönt.

Bedeutung beim Sammeln von Ausgangsdaten. Ein geschichtlich bedeutendes Beispiel sind die sog. Kapitularien aus der Zeit KARLs des Großen, später dann die Landbücher.

Mit Beginn der Neuzeit entstand unter dem Namen Statistik eine eher empirisch orientierte Staatslehre als beschreibende und bestenfalls aus spekulativen Überlegungen begründete Wissenschaft. Der damals gerade einsetzende weltweite Handel ging Hand in Hand mit einem neuen Staats- und Weltverständnis: Soziomorphe Erklärungsweisen der Natur und die theokratische Staatsidee gerieten immer mehr ins Abseits; der individuelle Erfahrungshorizont weitete sich aus. Kaufleute erzählten von fremden Ländern und die Naturwissenschaften erlebten eine erste Aufbruchsstimmung.

Bis zu jener Zeit gaben vor allem Theologen als die wichtigste gebildete Gruppe kompetente und verbindliche Erklärungen zu allen Erscheinungen des täglichen Lebens ab: Nunmehr zogen sie sich aus diesen Bereichen immer mehr zurück. - Auch heute noch gehört es für viele Gebildete zum guten Ton, von Naturwissenschaften und Technik wenig zu verstehen ...

Europa erlebte in der Mitte des 16. Jahrhunderts den Höhepunkt seiner Expansion. Der Handel mit Waren und Dienstleistungen wurde abstrakter: Man erfand das Brief- und Termingeld; Buchstabenrechnen und der Umgang mit mathematischen Formeln wurden eingeführt. Die *Wissenschaft von den Staatsmerkwürdigkeiten*, wie die Statistik damals hieß, entwickelte zunehmend das Bedürfnis, den politischen und sozialen Raum nicht nur beschreibend, sondern auch prognostisch zu durchdringen.

Grundsätzlich vorhandene und erkennbare Risiken sollten durch kalkulierbare Schätzverfahren tragbar werden. Es ist also kein Zufall, daß sich Christiaan HUYGENS mit einer *Theorie des Zufalls* (1657) beschäftigte, daß in England um 1696 die erste Seeversicherung LLOYD (Name eines Kaffeehausbesitzers) gegründet wurde. An der Universität Helmstedt zog 1660 die Statistik als akademische Disziplin *Notitia rerum publicarum* ein ...

Die Beschäftigung mit Glücksspielen war auslösend für die neue Wissenschaft von der Wahrscheinlichkeit: Abraham de MOIVRE (1667 - 1754) und Pierre LAPLACE (1749 - 1827) seien als Pioniere nur stellvertretend erwähnt. In jener Zeit wurde für die Statistik ein tragendes Fundament gelegt. Sozialwissenschaft, Ökonomie und Naturwissenschaft standen damals noch in enger Beziehung; der bekannte Astronom Edmont HALLEY (1656 - 1742), dem der Nachweis der perodischen Wiederkehr eines aufregenden Himmelsphänomens gelang, das wir heute nach ihm als Kometen benennen, trat auch als Versicherungsmathematiker an die Öffentlichkeit. - Frühe Vertreter der damals sog. *Politischen*

*Arithmetik* hatten das Bedürfnis, stets ihren Bezug zur Wirklichkeit zu betonen: So schreibt ein Mathematiker jener Tage im Vorwort seines Buchs, daß seine Ergebnisse auf Beobachtungen beruhten, ohne die alles Nachdenken vergeblich sei ...

Verweilen wir noch ein wenig in der Geschichte: Mit dem Aufkommen der industriellen Massengesellschaft im 19. Jahrhundert wurde erkannt, daß Grundgesamtheiten nur noch über Stichproben zu erfassen waren. Während bisher die Sozialstatistik fast nur mit vollständig abzählbaren Mengen operierte, stellte sich nun die Aufgabe, aus Stichproben verläßliche Schätzungen für die Population abzuleiten. Die neue Wahrscheinlichkeitsrechnung wurde zum Fundament dieser Entwicklung:

Gestattet eine genaue Kenntnis des Ist-Zustands und der Vergangenheit Prognosen für die Zukunft? Mit welchem Risiko sind solche Aussagen behaftet? - Diese und andere wichtige Fragen konnten nur nach und nach geklärt werden, und teilweise wird daran noch heute gearbeitet. - Sehr auffällig ist, daß diese neuen Verfahren im Rahmen sozialer und ökonomischer Fragestellungen entwickelt worden sind, ihre Bewährungsprobe aber in den Naturwissenschaften bestanden, so vor allem in der Thermodynamik Ludwig BOLTZMANNs (1844 - 1906). Die Anwendung von statistischen Methoden auf die Gesetzmäßigkeiten menschlichen Handelns und Zusammenlebens (mit sehr individuellen und gesellschaftlichen Aspekten) erfolgte hingegen nur zögernd.

Auch heute noch stößt der Versuch, menschliches Miteinander und persönliche Individualität zählend und messend zu durchforschen, auf starke emotionale Ablehnung: Der einzelne versteht sich als unersetzbare und einmalige Person ... doch urplötzlich wird er zum Gegenstand einer Sozialstatistik: Wahrscheinlichkeitstheoretisch begründbare Vorhersagen z.B. über Handlungsweisen, Entscheidungsverhalten, individuelle Nicht-Eignungen werden möglich. Widerstand gegen solches erklärt sich mindestens teilweise dadurch, daß man sich insbesondere in der privaten Sphäre schon vorab im Besitz abschließender Urteile wähnt.

Jedoch wird gerne vergessen, daß die eigenen Verhaltensmuster letzlich aus Erfahrung resultieren, also aus nichts anderem als einer Privatstatistik der Lerninhalte des eigenen Lebens. Die Statistik kann in diesem Zusammenhang als Interpretation oder Versuch verstanden werden, diese eher vorwissenschaftlichen Erfahrungen (eigenes Wissen) zu objektivieren. In diesem Sinn kann Statistik (richtig genutzt) Entscheidungen absichern, ohne jedoch den Entscheidungsträger aus seiner Verantwortung zu entlassen. Der primäre Gewinn liegt vor allem darin, daß die moderne Statistik zugleich ein Maß für das jeweilige Risiko angeben kann.

Im täglichen Leben gebrauchen wir oft Begriffe wie 'selten', 'oft', 'zufällig', 'regelmäßig' usw. Wir interpretieren und kommentieren dann Ereignisse vor dem Hintergrund der eigenen Lebenserfahrung und versuchen dadurch, uns in einer zunehmend undurchschaubaren und auch widersprüchlichen Welt zu orientieren. Unsere Deutungen und Werturteile sind dabei weitgehend durch das jeweilige soziale Umfeld bedingt, mithin also sehr subjektiv. Im Prinzip bleibt die Situation durch Ungewißheit sehr belastet. Als einfacher Ausweg bietet sich der Rückzug auf 'unmittelbar einleuchtende Gewißheiten' und im Grunde unreflektierte und oft ideologisch verhärtete Grundüberzeugungen an, die durch einseitige Selektion entstanden sind. Es ist daher nicht überraschend, daß Statistik gerade in der Anwendung im Bereich aller sozialen Wissenschaften (wie Psychologie, Soziologie, Pädagogik, Politologie, aber auch Medizin) auffällig oft dem widerspricht, was man selber als 'selbstverständlich' und 'natürlich', eben als Alltag erlebt.

In Fortführung der Gedanken wird verständlich, daß Statistik in allen Bereichen menschlichen Miteinanders vielfach nicht erwünscht ist; mit empirisch gesicherten Ergebnissen wird nämlich das Gefühl gestört, daß man zur Gruppe jener gehört, die 'Bescheid wissen'. Es ist allemal bequemer, sich auf Führungspersönlichkeiten und andere Leitbilder zu verlassen und die eigenen Vor-Urteile bestätigt zu finden, als eben diese Überzeugungen in Frage zu stellen.

Der Mensch des 20. Jahrhunderts befindet sich hier noch ungebrochen in der Tradition des ausklingenden Mittelalters: Jedwelche Versuche, vor allem emotionale Bereiche mit Zähl- und Meßmethoden zu durchdringen, stoßen auf sehr intensive Aversion. – Der geordnete mittelalterliche Kosmos hatte seinen Sinn aus der Projektion göttlicher Ordnungsprinzipien in die Sozialstruktur und in die unbelebte Natur gewonnen; aber die Naturwissenschaften hatten mit diesen Vorstellungen radikal aufgeräumt. Eine Sinndeutung der Welt konnten und wollten sie nicht geben; dafür konnten sie die Natur allfällig erklären. Je mehr sich diese Einsicht verbreitete, desto mehr zogen sich die Allgemeingebildeten von Naturforschung zurück. Auch heute noch ist sie vorwiegend Sache von Spezialisten, die häufig als einseitig verbildet angesehen werden.

Von daher ist es nur konsequent, sich gegen jede quantifizierende Einordnung der Persönlichkeit heftig zu wehren und insbesondere die Vorhersage menschlicher Verhaltensweisen u.ä. abzulehnen. Von den Naturwissenschaften erwarten wir bis in unsere Tage permanenten Fortschritt, auch wenn nunmehr Kritik laut wird und Aufrufe zum Umdenken erfolgen. Diese Entwicklung wurde bisher als ideologiefrei angesehen; erst jetzt beginnt man damit, diese Sichtweise nachdrücklich zu hinterfragen. Für

die Erforschung der Gesetze sozialen Wandels galt diese Ideologiefreiheit schon von Anfang an nicht.

Ein wichtiger psychologischer Aspekt muß noch angesprochen werden:

Im persönlichen Bereich erträgt der Mensch nur ein sehr beschränktes Maß an Ungewißheit. Die Belastungsgrenze liegt dabei umso niedriger, je einschneidender und intimer die Ereignisse und Bedingungen sind, die er selber nicht oder nur wenig bestimmen kann. Das betrifft zuerst alle Fragen der Gesundheit und des Lebensglücks, des sozialen, schulischen, beruflichen Erfolgs oder Mißerfolgs, des Einkommens usw.

Gerade weil hier Ungewißheit besonders abgewehrt werden muß, ist die eigentliche Wahrheit unerwünscht. Eindeutige Gewißheit ist somit am sichersten zu bewahren, wenn man irgendwelche empirischen Bewährungsprüfungen kategorisch ablehnt oder mit allerlei Ausreden relativiert. Vorsorgeuntersuchungen sind eine sinnvolle Sache, aber eben nur für andere! Auch viele Vorurteile gegen Psychologie wie Psychologen haben ihre Ursache letztlich in dieser Grundhaltung. Sie übertragen sich dann zwangsläufig auf die benutzten Methoden, denen heutzutage meist ausgefeilte statistische Modelle zugrunde liegen.

Von Haus aus denkt der Mensch im allgemeinen kausal; Statistik scheint diesem Ansatz zu widersprechen. In der Tat sind simple Schlußfolgerungen nach dem Muster *wenn-dann* immer gefährlich und zeugen von geringer Verantwortung; moderne korrelationsstatistische Verfahren gehen aber nicht von dieser primitiven Denkweise aus: Mit ihnen wird versucht, komplexe Wirkungsgefüge mathematisch zu strukturieren und dann in gegenseitigen Abhängigkeiten zu erklären. - Sog. Intelligenztests sind hier exemplarisch: Intelligenz ist nicht mehr länger etwas, was der einzelne in irgendeiner Quantität (als Intelligenzquotient IQ) hat. Intelligenz wird vielmehr als kompliziertes Gefüge von kooperativen Fähigkeiten und Eigenschaften verstanden, von Faktoren also, die sich an einer einzelnen Person weitgehend wertfrei feststellen, beschreiben und abgrenzen lassen. Auch Begriffe wie 'Begabung', 'Eignung' usw. werden in diesem verbesserten Sinn neu definiert.

Statistische Methoden sollen Wissen nicht nur als mitteilbar und prüfbar aufarbeiten, sie sollen auch den Anwendungs- und Geltungsbereich offenlegen, also positiv zur Verbesserung der Daseinsbedingungen beitragen, beim einzelnen wie in der Gesellschaft. - Devise: Nicht nur munter Material sammeln und sortieren, sondern auch Sicherheitsgrenzen abstecken und den Geltungsbereich benennen!

Dieses neue Methodenbewußtsein trägt viel zur sachlichen Aufklärung der Lebensbedingungen bei. Alle Planungsaufgaben einer modernen Gesellschaft in Verkehr, Wirtschaft, Sozialpolitik, Bildungswesen usw. sind ohne statistische Verfahren nicht mehr denkbar. Analoges gilt für persönliche Entscheidungsprozesse, gleichgültig, ob andere dabei mitwirken oder nicht.

Leider kommen Fehldeutungen und absichtliche Entstellungen immer wieder vor. Teilweise sind dafür Gründe verantwortlich, von denen schon weiter oben die Rede war. So ist die Redensart von der Statistik als einem sog. 'Zahlenfriedhof' oder einer 'geistlosen Zählmaschine' oft Abschirmung vor überraschenden und unbequemen Einsichten oder auch Eingeständnis mangelnden (mathematischen) Verständnisses. Derselben Person oder auch Interessengruppe kann es bei anderer Gelegenheit durchaus willkommen sein, gerade ihre Postulate oder Forderungen mit (sogar manipulierten) Statistiken zu untermauern. – Zum Glück kann die statistische Lüge, im Gegensatz zu den vielfältigen Formen der Unwahrheit im Alltag, im allgemeinen aufgedeckt werden. – Die Formulierung, daß Statistik nur die letzte Steigerungsform der Lüge (mit der Zwischenstufe der Notlüge) sei, ist daher eher ironisch zu verstehen.

Ist es Zufall, daß dieser Satz gerade von jenen besonders gerne zitiert wird, deren eigene Behauptungen und Leitsätze (in der Regel ohne jede Beweiskraft) in Politik, Rechtswesen, Wirtschaft und Schule und so weiter ohne nennenswerten Widerstand hingenommen werden? – Führungspersönlichkeiten wähnen sich oft im Besitz jener Grundwahrheiten, mit denen die Ordnung auf Erden abschließend hergestellt wird. Freilich können Traditionen, seien sie nun schlicht bequem oder auch nur eingewöhnt, oft auch (zugegeben!) sinnvoll, nicht von heute auf morgen über Bord geworfen werden, zumal dann, wenn sie der eigenen Absicherung und vielfältigen Vorteilen dienen.

Nicht ohne Grund hat der Gesetzgeber dafür gesorgt, daß gewisse Tatbestände nur von ihm statistisch durchleuchtet werden dürfen. Er scheut sich dabei nicht, wenig plausible Gründe für diese Restriktionen zur Informationsgewinnung anzuführen. Beispielsweise wird eine durchsichtige Einkommensstatistik mit dem Vorwand verhindert, daß das sog. Steuergeheimnis gewahrt bleiben müsse. Eher mag schon der soziale Friede gestört werden: Denn der Staat (das sind immer auch einzelne Mandatsträger mit persönlicher Vorteilsnahme) ist wenig interessiert, selber krasse Ungereimtheiten aufzudecken. Also behält er diese Informationen für sich und operiert im Beispiel mit irgendwelchen durchschnittlichen Einkommen, die für Außenstehende unanfechtbar sind, aber (wie jeder Statistiker weiß) nur recht geringe Aussagekraft haben. – Auf diese Weise wird vertuscht und beruhigt, aber nicht aufgeklärt.

Informationsquellen und Exekutive sind also in vielen Fällen und zunehmend in einer Hand; darin liegt eine nicht zu unterschätzende Gefahr moderner Datengewinnung. Datenschutzgesetze haben daher durchaus zwei Seiten: Während die Weitergabe eines ziemlich unbedeutenden persönlichen Datums strafwürdig sein kann, werden an anderer Stelle weithin unbemerkt viel umfangreichere und brisantere Datenpakete angesammelt, gegen die sich der einzelne, selbst wenn er davon zufällig Kenntnis erlangt oder doch entsprechendes vermutet, kaum zur Wehr setzen kann.

Das ist ein durchaus aktueller Aspekt, nicht erst seit der noch jungen Wiedervereinigung und der Aufarbeitung der Akten der Firma "VEB Horch und Guck" (so hieß die Stasi im anderen Deutschland). Das Studium statistischer Methoden hat jenseits allen Interesses an der Materie (sei es reine Neugier oder nur der Wunsch nach Anwendung) immer auch das Ziel, das Verantwortungsbewußtsein zu schärfen: Gefordert sind nicht nur Sorgfalt bei der Ermittlung und methodischen Aufbereitung der Daten, sondern auch wacher Verstand und kritische Würdigung bei der Interpretation des Materials ... Und auch die Frage darf gestellt werden, ob denn alles erfaßt werden muß ...

Statistisches Wissen zeichnet sich durch prüfbare Bewährung an der Wirklichkeit aus: Beobachtung und Erfahrung auf der einen Seite, eine gesicherte Theorie auf der Basis mathematischer Methoden auf der anderen wirken sinnvoll zusammen. Eine Sichtweise allein bliebe entweder subjektiv, uneindeutig und von ungewissen Sicherheitsgrad, oder eben nur eine Spielwiese der Wissenschaft. Die Statistik ist das Musterbeispiel einer angewandten Disziplin: Eine Vielzahl einzelner Beobachtungen wird geordnet (klassifiziert), verglichen, ausgewertet und kommentiert. Eine formalisierte Theorie verknüpft die Ergebnisse mit anderen Daten, die auf ähnliche Weise schon gewonnen worden sind. So ergeben sich ganz neue Erkenntnisse, aus denen dann Schlüsse mit sog. Wahrscheinlichkeitscharakter, also mit Angabe eines meßbaren Risikos, gezogen werden können.

Täglich sind wir vor neue Entscheidungen gestellt: Wir können eigentlich nichts besseres tun, als nach solchen wahrscheinlichsten Setzungen zu handeln. Die persönliche Freiheit wird dadurch nicht eingeengt, ja im Gegenteil: Erwartungen und Hoffnungen, Befürchtungen und Ängste werden unter Kontrolle gebracht, rationalisiert. – Die Freiheit der selbstverantwortlichen Entscheidung bleibt, sie wächst sogar ...

Ein paar wichtige Anwendungsbereiche für statistische Methoden seien stellvertretend genannt:

In der Privatwirtschaft:

Produktionsplanung, Kostenanalyse, Absatzstrategie, Werbung und Marktanalyse samt Rentabilität, Lagerhaltung und Vertrieb, Abnahmekontrollen, Versicherungswesen (Prämien). Im Bereich Produktion: Qualitäts- und Normenkontrolle, Terminplanung, Sicherheits- und Unfallforschung, Arbeitsplatzgestaltung, Werkstoffprüfung und Rohanalysen, Exploration und Förderung von Rohstoffen, Recycling, Risikoberechnung und Folgelasten bei Großprojekten.

Im öffentlichen Bereich:

Sozialversicherung und Renten, Wirtschafts- und Steuerpolitik, Verkehrsplanung, öffentliche Dienstleistungen, auch Bildungspolitik, Geldumlauf, Subventionen, Binnen- und Außenhandel, Wettervorhersage, Arbeitsmedizin und anderes.

Alle zivilisierten Länder haben für diese und andere Zwecke Statistische Ämter (in Schweden schon 1756) eingerichtet, Behörden, die im Rahmen gegebener Gesetze diese o.g. und andere Aufgaben unterstützen. Volkszählungen (in Deutschland ab 1742) sind ein Vorgang, wo man sich als Normalbürger dieser Behörden wieder erinnert.

Bei uns ist dies das Statistische Bundesamt in Wiesbaden, eine eher unauffällige, aber doch recht effiziente "Datenfabrik", die neben einer Fülle von speziellen Fachveröffentlichungen der interessierten Öffentlichkeit regelmäßig auch ein sog. Statistisches Jahrbuch anbietet. Nicht vergessen sei der monatliche Lebenshaltungsindex, umgangssprachlich *Warenkorb* genannt, mit dem das Amt auch Signale für die Politik setzt. Unser Bundesamt geht in langer Tradition auf das Kaiserliche Statistische Amt des Deutschen Reiches zurück, das im vorigen Jahrhundert die Vielfalt der damaligen Kleinstaaten "reichsseitig" durch Empfehlungen zur statistischen Basisarbeit koordinierte.

# 2 SKALEN und PARAMETER

Statistik, das ist die Untersuchung und Beschreibung von Gesetzmäßigkeiten bei Massenerscheinungen. Die Masse, bei der solche Erscheinungen auftreten: das kann die Bevölkerung der Bundesrepublik sein, die Menge der Studentinnen und Studenten in einer Vorlesung, die gesamte Produktion von Glühlampen in der Firma Primalux oder auch die Anzahl aller Legehühner im Freistaat Bayern. Diese jeweils deutlich abgrenzbare Gesamtheit von Merkmalsträgern (Objekte) nennt man Grundgesamtheit oder Population. Oft wird dieses Wort auch nur für den tatsächlich zu Beobachtungen verfügbaren Teil der Masse genommen. Alle Individuen (Objekte) – bei fortgeschrittener Methodenkenntnis auch nur eine nach gewissen Regeln repräsentativ ausgewählte Stichprobe – werden auf ein gewisses Merkmal hin befragt, oft auch auf ein ganzes Bündel solcher Merkmale, bei denen später allerhand Querverbindungen von Interesse sein können. – Nahezu jede statistische Untersuchung beginnt mehr oder weniger aufwendig mit einem solchen ersten Schritt, den man Erstellung einer Urliste nennt.

Unter einem Merkmal versteht man dabei eine beobachtbare und abfragbare Eigenschaft, die je nach Merkmalsträger (Proband) in unterschiedlicher Ausprägungsform, Graden vorkommt. Man bezeichnet solche Merkmale mit einem Wort aus der Mathematik als Variable. Eine Variable muß also je nach Fall unterschiedliche Intensität aufweisen, auch wenn es oft schwierig ist, die Ausprägungsgrade gegeneinander abzugrenzen. – Bei der Anwendung von statistischen Methoden geht man dabei von der Annahme aus, daß der Wert im Rahmen der Bandbreite der Variablen am jeweiligen Probanden so oder anders zufällig vorgefunden wird. Begriff und Bedeutung des Wortes Zufall müssen für die Theorie nicht weiter präzisiert werden; es genügt, daß die konkrete Feststellung einer Merkmalsausprägung für den Befrager unvorhersehbar ist, auch wenn sich am Probanden (später) ein komplexes Wirkungsgefüge als irgendwie kausal erkennen läßt:

Ob ein auf der Straße angehaltener Passant Deutscher oder Ausländer ist, ob eine Studentin lieber schlapprige Hosen als enge Röcke trägt, ob eine eben gekaufte Glühlampe funktioniert oder wieviele Eier die bayerische Legehenne Erna wöchentlich legt, das ist (mehr oder weniger) zufällig, auch wenn sich im konkreten Fall (insbesondere nach Befragung) gewisse Gründe fallweise erkennen lassen. Vorsicht bei der Bewertung von Antworten ist immer dann geboten, wenn innere Zusammenhänge eine Rolle spielen (können): Ein Interview etwa unter Passanten zur Meinung über die Finanzbehörden sollte besser nicht in der unmittelbaren Umgebung eines solchen Amts durchgeführt werden; die Ergebnisse wären – da der Befragte vielleicht gerade von dort verärgert herkommt – sicherlich verfälscht ...

Für theoretische Überlegungen und Ableitungen greift man daher gerne auf Modelle zurück, so das Würfelmodell oder das Urnenmodell; beide werden uns noch ausführlich beschäftigen. Hier gilt, daß der einzelne Augenausfall, der einzelne Griff in die Urne vollständig dem Zufall unterliegen, also erst bei vielfacher Wiederholung eine statistische Gesetzmäßigkeit erkannt wird, die für den Einzelversuch Aussagen mit Risikocharakter zuläßt. Entsprechende Versuche haben typische Charakteristika des naturwissenschaftlichen Experiments: Vorabbeschreibung des Verfahrens, Wiederholbarkeit und damit Reproduzierbarkeit der Ergebnisse.

Mit den Methoden der beschreibenden Statistik kann nach der Sichtung des Urmaterials, also dem Sortieren der erfaßen Daten aus der Urliste, mitgeteilt werden, wie das Merkmal in der Population 'verteilt' ist: Im einfachsten Fall ist das eine Aussage darüber, wie oft die einzelnen Ausprägungsgrade der betrachteten Variablen vorkommen. So sind beispielsweise von 56 Studenten eines Semesters insgesamt 22 weiblich. Eine solche Variable nennt man dichotomisch (zweiwertig).

Wären die allermeisten der 56 Studenten Frauen, so wäre ein beliebig herausgegriffener Proband 'aller Voraussicht nach' weiblich ... Eine präzisere Formulierung ist erst mit Methoden der operativen Statistik möglich, auch wenn wir schon zu wissen glauben, was damit gemeint ist.

Die Beschaffung des Urmaterials für eine Statistik kann auf vielerlei Weise geschehen. Sieht man davon ab, daß bereits vorhandene Daten (die sog. "Primärstatistik") neu bearbeitet werden (das Ergebnis heißt dann "Sekundärstatistik"), so erfolgt die Beschaffung der Daten prinzipiell durch Registrieren am Merkmalsträger, durch direktes Befragen (Interview, Zählen, Messen), durch indirektes Befragen (Fragebogen), durch direkte oder indirekte Beobachtung (Tonband, Video, Beobachtung mit dem Teleskop oder Auswerten von Satellitenfotos) oder durch ein Experiment im engeren Sinn des Wortes, also eine Laborsituation. Dies sind die häufigsten und wichtigsten Verfahren. Zählen, Messen und Befragen sind jedermann vertraute Methoden; 'Beobachten' heißt, objektiv wahrnehmbare Eigenschaften, Verhaltensweisen und ähnliches nach festgelegten Kriterien zu erfassen, ohne zunächst irgendwie zu bewerten.

Die wichtige Frage der Objektivität müssen wir in unserem Fall nicht so genau diskutieren; die numerische Datenauswertung im naturwissenschaftlichen und technischen Bereich ist relativ problemlos: Man kann meistens unterstellen, daß die Ergebnisse ohne Beeinflussung des Probanden zustande gekommen sind, auch wenn Fälle denkbar sind, bei denen die Beobachtung selber die Daten verfälscht hat. In erster Linie achtet man auf korrekte (und sinnvolle) Messungen, also Meßfehler und dgl.

Notwendige Voraussetzung für die statistische Bearbeitung von irgendwelchen Daten ist die Möglichkeit, das betrachtete Merkmal, also die Variable, hinsichtlich des Ausprägungsgrades qualitativ oder besser quantitativ klassifizieren zu können. Nach vorgegebenen Regeln wird der Ausprägung durch Zählen eine Ordnungsnummer und damit eine Häufigkeit zugewiesen, oder aber durch Messen eine Maßzahl ermittelt, ein Zahlenwert (Größe) attestiert. So können von 56 Studenten 42 blondes Haar haben, während die Körpergröße X im konkreten Fall des Kommilitonen Niedermayer x = 188 cm beträgt.

Demnach lassen sich Merkmale zunächst grob in zwei Kategorien einteilen, in qualitative und in quantitative. Letztere werden mit einer metrischen Skala verglichen, mit einer Absolutskala oder doch wenigstens Intervallskala.

Alle nicht quantitativen Merkmale sind qualitativ; hier wird eine Unterscheidung hinsichtlich der Qualität getroffen, dies in der Regel im Vergleich mit anderen Probanden. Die Spannbreite der Skalen reicht hier von einer einfachen Nummernskala (wie z.B. Hausnummern als Ordnungsprinzip) über Nominalskalen (Klassifizierung nach Klassen ohne inneren Zusammenhang: Haarfarbe, Einteilung nach Geschlecht) bis hin zu Ordinal- oder Rangskalen, d.h. einer Ordnung nach Intensitätsklassen mit graduellen Unterschieden. Bei Rangskalen verwischt sich der Übergang zu Intervallskalen gelegentlich: Schulnoten sind strenggenommen nur rangskaliert, werden oft aber intervallskaliert interpretiert und in der Folge dann sogar mit arithmetischen Mittelwerten versehen, obwohl dies eigentlich nicht korrekt ist. Das folgende Schema nennt Beispiele.

Die Ergebnisse von nominal skalierten Variablen faßt man zumeist in Häufigkeitstabellen zusammen, wobei zwischen den einzelnen Klassen keine innere Beziehung bestehen muß. Leicht erkennbar ist dann der häufigste Wert, das sog. Dichtemittel (Modalwert), das natürlich von der Klasseneinteilung abhängt: Bei einer Einteilung nach Haarfarben mag die häufigste Farbe z.B. 'blond' sein; das ist das Dichtemittel, nicht etwa die zugehörige Häufigkeit $f_{blond} = 23$. Der Name der Klasse steht für das Dichtemittel! Es kommt oft vor, daß bei der gewählten Klasseneinteilung zwei oder mehr Häufigkeiten mit größtem Wert übereinstimmen. In solchen Fällen spricht man von einer zweigipfeligen (bimodalen) bis mehrgipfeligen Verteilung, und im Grenzfall von einer tafelförmigen Verteilung dann, wenn die Reihenfolge der Klassenaufzählung nicht willkürlich ist, vielmehr bei höherwertiger Skalierung einem inneren Zusammenhang (zwischen den Klassen) folgt.

Ordinalskalen lassen bereits Vergleiche wie *besser/schlechter* oder auch *mehr/weniger* zu; neben dem Modalwert läßt sich hier zusätzlich noch ein Zentralwert definieren: Zur Bestimmung

werden alle Daten (Meßwerte) z.B. in aufsteigender Reihenfolge angeordnet, um dann jenen Wert zu suchen, der diese Reihe nach Position halbiert. Sind z.B. die 17 Klausurnoten

1 1 2 2 2 3 3 3 3 4 4 4 4 4 4 5 5

aufsteigend sortiert, so ist der häufigste Wert offenbar die Note Vier (Name der Klasse), während der Zentralwert noch Drei auf Position 9 = (17+1)/2 ist. - Bei 18 Noten kann dessen Bestimmung auf Probleme stoßen, da wir uns z.B. bei einer zusätzlich weiteren Fünf auf Drei oder Vier einigen müßten. In solchen Fällen wählt man, sofern möglich, dann das arithmetische Mittel aus den beiden Nachbarklassen, was hier - da Drei und Vier ja Namen von Klassen und nicht Werte sind, allerdings fragwürdig wäre. Dies gilt erst recht für das gerne berechnete arithmetische Mittel bei Noten! Betont sei, daß bei Ordinalskalen die einzelnen Klassen nicht im folgenden Sinn am Beispiel der Noten verglichen werden können: *Mit der Note Zwei ist jemand doppelt so gut wie mit einer Vier.* Der Zweier ist lediglich besser, sonst nichts ...

| Skalentyp | Mögliche Lagemaße |
|---|---|
| a) nicht-metrische Skalen | |
| Nummernskala<br>(Hausnummern, KFZ-Nummern, ...) | keine |
| Nominalskala<br>Geschlecht, Augenfarbe, Qualität, Familienstand ... | häufigster Wert: Mode |
| Ordinalskala<br>Rangfolgen, Noten, Dienstgrade, Brinell-Härte bei Mineralien ... | Zentralwert: Median |
| b) metrische Skalen | Arithmet. Mittel |
| Intervallskala<br>Lautstärke in dB, Temperatur in Grad Celsius, Kalenderzeit | |
| Absolutskala<br>Größe, Gewicht, Kelvinskala, Kinder je Familie ... | |

Übersicht : Skalierungen

Der arithmetische Mittelwert ist erst bei metrischen Skalen zulässig. Die allgemein bekannte Berechnung sei einstweilen

übergangen. Die beiden unter b) genannten Skalentypen unterscheiden sich nur dadurch, daß der Nullpunkt einer Absolutskala aus einem inneren Grunde zwingend festgelegt werden muß, während Nullpunkte bei Intervallskalen mehr oder weniger willkürlich (operational) definiert werden: Am Beispiel etwa der Temperatur in Grad Kelvin bzw. Celsius (bei 0 Grad Celsius gefriert definitionsgemäß Wasser) wird der Unterschied besonders deutlich.

Bei metrischen Skalen können Differenzen zwischen Werten aussagekräftig verwendet werden, d.h. gleiche Differenzen sind an allen Stellen der Skala gleichwertig. Jedoch ist eine Aussage *Bei 20 Grad Celsius ist es doppelt so warm wie bei 10 Grad.* falsch, während eine entsprechende Formulierung mit Kelvingraden richtig wäre!

Die Übersicht der vorigen Seite bringt von oben nach unten steigende statistische Qualität; je weiter unten ein Variablentyp plaziert werden kann, desto ausgefeiltere Methoden existieren zu seiner Untersuchung. Generell gilt dabei, daß eine für "schlechtere" Merkmalstypen geeignete Methode auch bei "besseren" noch zulässig ist, aber nicht umgekehrt, wie das Beispiel des arithmetischen Mittels bei Noten zeigt.

Insbesondere metrische Skalen bzw. deren Variable können noch nach einem anderen Gesichtspunkt unterteilt werden: Wird der Wert durch Abzählen (u.U. bis ins Unendliche) gefunden, so nennt man die Variable diskret. Andernfalls – meist erfolgt dann die Bestimmung des Wertes mit einer im Prinzip beliebig fein geteilten Meßlatte – heißt die Variable stetig. In diesem Sinne sind Länge und Gewicht stetige Variable; hingegen ist die Größe 'Pulsschläge pro Minute' diskret. Aber alle drei sind metrisch, ja sogar absolut (auch rational genannt). Hingegen sind Variable mit nur endlich vielen Ausprägungsformen stets diskret, also alle unter a) aufgeführten.

Zum Begriff des Messens ist noch eine Anmerkung erforderlich: Das zu messende Merkmal (die empirische Gegebenheit) kann u.U. auf ganz verschiedene Weisen numerisch abgebildet, skaliert werden. Das hängt von der Meßvorschrift (der Zuordnung zum numerischen Relativ) ab. Eine Mindestforderung ist aber, daß Beziehungen zwischen den Objekten wie z.B. *kleiner/größer, doppelt so viel* und dgl. stets durch die Meßzahlen in gleicher Weise wiedergegeben werden müssen, diese also eine Repräsentation der ursprünglichen Beziehungen darstellen. Die Meßvorschrift muß im mathematischen Sinn eine affine Abbildung sein. Die Entdeckung und Anwendung eines physikalischen Gesetzes sollte z.B. nicht davon abhängen, ob man die Temperatur in Grad Celsius oder Reaumur angibt; Ausdehnungskoeffizienten können fallweise eben umgerechnet, verschiedene Skalen durch lineare (!) Transformationen angepaßt werden ...

Die bisher genannten Mittelwerte sind sog. statistische Maßzahlen. Diese auch Parameter genannten Kenngrößen für Verteilungen lassen sich entsprechend unterschiedlichsten Bedürfnissen wie folgt einteilen:

**Maßzahlen der Lage (Mittelwerte, Schwerpunktsmaße)**
**Maßzahlen der Streuung (um den Schwerpunkt)**
**Maßzahlen der Form (Schiefe, weiter noch Exzeß).**

Mittelwerte: das sind Lagemaße der zentralen Tendenz, der sog. Schwerpunktslage. Wie oben gesagt, gibt es bereits auf dem niedrigen Nominalniveau den Modalwert oder Mode, d.i. der Name der am stärksten besetzten Klasse in einer u.U. recht willkürlichen Klasseneinteilung zur Merkmalsausprägung.

Bei Rangskalen ist der zusätzlich definierbare Zentralwert meist von höherer, weitergehender Aussagekraft. Kommt im Falle einer metrischen Skala noch das arithmetische Mittel hinzu, so stehen alle drei Mittelwerte in einer auffälligen inneren Beziehung zueinander, die etwas mit der "Form" der jeweiligen Verteilung zu tun hat. Betrachten wir dazu das Beispiel einer fiktiven Einkommensverteilung, einer Variablen, die an sich (unabhängig von der Erfassung aus mehr praktischen Gründen) als metrisch (zudem stetig, absolut) betrachtet werden kann:

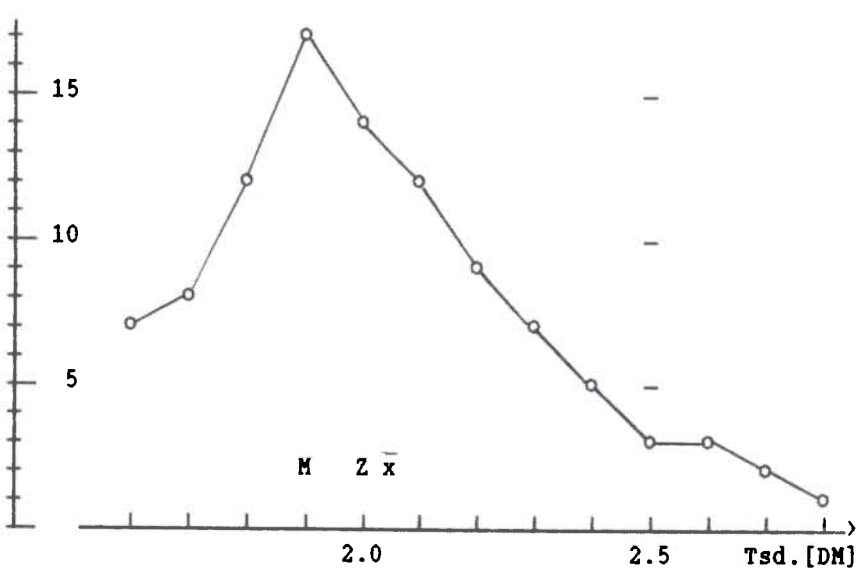


**Abb.: Einkommensverteilung von n = 100 Probanden**

Im Bild wird angenommen, daß insgesamt 100 ermittelte Monatseinkommen in Klassen der Breite 100 DM (gerundet) eingeteilt

worden sind. Es läßt sich zeigen, daß das arithmetische Mittel $\bar{x}$ statt an den Urwerten noch mit recht guter Näherung über die Klassen bestimmt werden kann; im obigen Beispiel findet man

$$\bar{x} = (7*1600 + 8*1700 + \ldots + 1*2800) / 100 \approx 2030 \text{ [DM]}.$$

Die am stärksten besetzte Klasse von DM 1900 stellt den Modalwert M dar mit folgender Bedeutung: Die meisten der Befragten haben monatlich eben diesen Betrag zur Verfügung.

Der Zentralwert Z wird durch die Klasse 2000 DM bestimmt. Er markiert die "Mitte" der Verteilung in folgendem Sinn: Die Hälfte der Befragten hat monatlich höchstens diesen Betrag zur Verfügung, der Rest (das sind ebenfalls n/2 = 50 Probanden) mindestens. Man findet Z durch Auszählen und Suchen jener Klasse, in welcher nach Rangfolge des Einkommens geordnet der Proband $n \approx 50$ fällt: $7 + 8 + 12 + 17 < 50$; mit dem weiteren Summanden 14 wird diese Ordnungsnummer übertroffen. – Im vorliegenden Fall gilt offenbar

$$M < Z < \bar{x}.$$

Man nennt eine solche Verteilung augenscheinlich links-steil, oder nicht so einsichtig auch rechts-schief. Sie ist jedenfalls keineswegs symmetrisch, wie Verteilungen "der schönen Dinge" meistens. Bei einigermaßen symmetrischen Verteilungen von metrisch skalierbaren Variablen gilt angenähert $M \approx Z \approx \bar{x}$. Analog wird der Begriff rechts-steil (links-schief) definiert. Beispielsweise ist die Verteilung der Sterberate nach Altersklassen stark links-schief; wer alt ist, stirbt eher. Unser obiges Beispiel ist sehr geeignet, über den vagen Begriff "mittleres Einkommen" aus der amtlichen Publizistik nachzudenken ... Man überlege sich vor allem den Grenzfall vieler Probanden mit wenig Einkommen, denen sich ein Millionär zur "Schönung" von $\bar{x}$ hinzugesellt.

Im allgemeinen hat eine Verteilung nur einen Gipfel M. Nur dann sollte man die anderen Mittelwerte (insb. $\bar{x}$, soweit möglich) angeben. Mehrgipfelige Verteilungen weisen meistens auf ein verdecktes Kriterium hin, nach dem man zusätzlich hätte klassifizieren sollen. Beispiel:

Eine nicht geschlechtsspezifische Stichprobe zur Körpergröße wird vermutlich zwei Gipfel aufweisen, einen für Männer und einen für Frauen. Es hat wenig Sinn, von einer durchschnittlichen Körpergröße aller Bundesbürger zu reden: Normalerweise will man wissen, wie groß (im Mittel) Frauen und Männer für sich sind. Immerhin ist denkbar (und anderswo in der Biologie auch offensichtlich), daß diese beiden Werte stark voneinander abweichen.

Der Form nach gleiche Verteilungen können offenbar sehr unterschiedliche Lagemaße haben. Sind $x_i$ Meßwerte mit dem Mittelwert $\bar{x}$, so hat die Verteilung X mit den Meßwerten $x_i + c$ den Mittelwert $\bar{x} + c$, ist also bei gleicher Form um c verschoben, was oft zur vereinfachenden Berechnung von Mittelwerten ausgenutzt wird. Dies gilt auch für Z und M.

Umgekehrt können Verteilungen mit gleichem Mittelwert (den man unpräzise auch oft "Durchschnitt" nennt) ganz unterschiedliche Form aufweisen, breit oder schmal sein, symmetrisch oder nicht usw. Neben Maßzahlen der Lage dienen daher weiter Maßzahlen der Streuung (Dispersionsmaße oder Variabilitätsmaße) einer differenzierteren Betrachtung:

Die Spannweite (Variationsbreite, Range)

$$R = x_{max} - x_{min}$$

kann schon bei Rangskalierungen eingesetzt werden. Allerdings ist deren Aussagekraft meist gering, sofern die Probanden nicht nahe beieinander liegen: Noten, Weiten im Springen auf einer Olympiade als Grenzfälle. Im ersten Fall kommen meistens alle Noten vor, während auf einer Meisterschaft nur die Besten gegeneinander kämpfen.

Die mittlere Abweichung

$$D = \frac{1}{n} \left( \sum_{i=1}^{n} |x_i - \bar{x}| \right)$$

ist ebenfalls leicht zu berechnen; sie wird meistens auf das arithmetische Mittel $\bar{x}$ bezogen, aber auch auf andere Mittelwerte. Verwendet man z.B. den Modalwert M anstelle von $\bar{x}$, wie es in Warenzeitschriften (z.B. Test) meist gemacht wird, so ergibt sich jene Abweichung vom häufigsten Wert, mit welcher der Käufer in einem beliebigen Geschäft rechnen kann.

Bei metrischen Skalen, so vor allem im technisch-wirtschaftlichen Bereich, ist die sog. empirische Standardabweichung s das gebräuchlichste Maß:

$$s := \sqrt{\frac{1}{n} \left( \sum_{i=1}^{n} (x_i - \bar{x})^2 \right)} \quad .$$

s wird stets auf das arithmetische Mittel $\bar{x}$ bezogen:

$$\bar{x} := \frac{1}{n} \sum_{i=1}^{n} x_i \quad .$$

In der Formel für s erkennt man, daß die Standardabweichung vor allem durch die von $\bar{x}$ weit entfernten Werte beeinflußt wird: s wird groß, wenn wenn die $x_i$ stark streuen.

$s^2$ heißt Varianz; $V := 100 \cdot s/\bar{x}$ wird Variationskoeffizient (in Prozent) genannt; bei kleinem $\bar{x}$, das durch negative $x_i$ erzeugt worden sein kann, ist bei V Vorsicht geboten: Große Werte von V (gar über 50 %) weisen auf schlechte Klassifizierung bei der Auswertung hin. - Generell sollten $\bar{x}$ wie s nur bei größerem n berechnet werden.

Der Vollständigkeit halber seien noch zwei andere Mittelwerte genannt:

$$GM := \sqrt[n]{x_1 * \ldots * x_n} \qquad (\text{alle } x_i > 0)$$

heißt geometrisches Mittel. Es wird bei multiplikativ verknüpften Merkmalsdaten (jährlichen Wachstumsraten, Zinssätze u. dgl.) benutzt, um dort sinnvolle Mittelwerte anzugeben.

Für sehr spezielle Sachzusammenhänge gibt es ein harmonisches Mittel:

$$HM := \frac{n}{1/x_1 + 1/x_2 + \ldots + 1/x_n} \qquad (\text{alle } x_i > 0) \ .$$

Mit dieser Formel (oder durch Gewichtung beim arithmetischen Mittel!) kann z.B. folgende Aufgabe gelöst werden:

Jemand kauft an einem Tag für 10 DM Orangen, das Stück zu 50 Pfennig, am anderen Tag ebenfalls für 10 DM, das Stück zu 40 Pfennigen. Was kostete eine Orange durchschnittlich? (Antwort: DM 0.44).

Mit Blick auf die Abb. von Seite 20 wird vorerst als einfaches Schiefemaß

$$sk := \frac{\bar{x} - M}{s}$$

definiert; man erkennt direkt, daß sk > 0 für linkssteile Verteilungen zutrifft. Bei leidlich symmetrischen Verteilungen ($sk \approx 0$) kann man auf Formparameter verzichten; dann reichen zur Beschreibung Maße der Lage und Streuung aus. Diese und andere Maßzahlen verdichten die Informationen der Urliste im statistischen Sinne, machen (neben grafischen Möglichkeiten) Mitteilungen prägnanter und einfacher ...

Die vorne definierte Spannweite R ist meistens ohne große Aussagekraft; wenn nur ein einziger Proband nach oben oder unten stark abweicht, ändert sich R spürbar, ohne daß die Verteilung wesentlich anders aussieht.

Zum Median Z paßt als Streuungsmaß besser der sog. mittlere Quartilabstand MQA. Als erstes Quartil bezeichnet man jenen Beobachtungswert, unterhalb welchem 25 Prozent aller Werte liegen. Das zweite Quartil ist der Median selbst. Die 75-Prozent-Grenze heißt analog drittes Quartil. Als MQA wird nunmehr die Hälfte der Differenz zwischen dem ersten und dem dritten Quartil erklärt.

Dieses Streuungsmaß ist gegen einzelne starke Abweichungen ziemlich "stabil", wie man sich leicht an Beispielen deutlich machen kann: Ist bei einem Sportwettbewerb der Sieger extrem gut, der Verlierer dazu auffällig schlecht, so kann das übrige Feld gleichwohl dicht beieinander liegen, also die Streuung im Sinne des MQA klein sein.

Über den Begriff Quartil hinaus kann bei Intervallskalen noch der sog. Prozentwert (Perzentil) eingeführt werden. Er beschreibt die Position eines Meßwerts im Vergleich zu den restlichen in folgender Weise: Unter dem Perzentil P versteht man denjenigen Punkt auf der Skala, unterhalb dem p Prozent der in der Verteilung vorkommenden Meßwerte liegen.

Das Perzentil 25 (im Konkreten dann durch einen Meßwert auf der Skala definiert, z.B. 170 cm), ist also das erste Quartil. Zur Bestimmung beliebiger Perzentile zieht man die (geordnete) Häufigkeitstabelle der Verteilung mit Aufwärtskumulation heran (nächstes Kapitel).

Interessant ist auch manchmal die umgekehrte Fragestellung: Gegeben ist irgendein Skalenwert und wir wollen wissen, wieviel Prozent der Meßwerte unterhalb dieses Werts liegen. Diese Zahlenangabe nennt man Prozentrang des Meßwerts. Prozentränge sind schon auf Ordinalniveau der Variablen erklärbar.

Die näherungsweise Bestimmung aus Häufigkeitstabellen ist nur Zähl- und Rechenaufwand. Bei Bedarf findet man in der Literatur exakte Formeln zur schnellen Bestimmung.

# 3 ELEMENTARE METHODEN

Am Anfang jeder statistischen Untersuchung steht eine Auflistung aller Daten in jener Reihenfolge, wie sie angefallen sind, die sog. Urliste:

$x_1, x_2, x_3, \ldots, x_n.$

Sind diese Daten in einer ersten Bearbeitung der Größe nach geordnet, so spricht man von der primären Verteilungstafel (Beispiel S. 18 oben).

Solche Urlisten können natürlich auch mit Paaren, Tripeln usw. von Ausgangsdaten auftreten, etwa bei einer Untersuchung des Zusammenhangs von Körpergröße X und Gewicht M

$(x_1, m_1), (x_2, m_2), \ldots$

In diesem Fall von Wertepaaren je Proband betrachtet man eine bivariable Verteilung.

Bleiben wir aber zunächst bei einer monovariablen Verteilung. Gewisse Daten können mehrfach vorkommen, was man durch Häufigkeiten $f_i$ in der Tabelle andeutet, die anstelle der Urliste sogleich als Strichliste entstanden sein könnte, insbesondere bei Nominaldaten:

| Werte X | | $f_i$ |
|---|---|---|
| $a_1$ | | 0 |
| $a_2$ | //// | 4 |
| $a_3$ | ///// | 5 |
| $a_4$ | /// | 3 |
| ... | ... | ... |
| $a_n$ | | 0 |

**Abb.: Häufigkeitstabelle**

Unterstellen wir aber z.B. metrische Werte $x_i$, so bedeuten die $a_i$ bereits eine Einteilung nach Klassen, und sie sind daher von den tatsächlich auftretenden $x_i$ genau zu unterscheiden! Ein konkreter Wert $x_k$ fällt dann in jene Klasse $a_i$, für die

$$a_i - b/2 < x_k \leq a_i + b/2.$$

gilt. Die Differenz $a_{i+1} - a_i$ heißt Klassenbreite b; der Klassenbezeichner $a_i$ ist die sog. Klassenmitte; vorteilhaft ist dabei (wenn möglich) die Wahl einer ganzen Zahl. Außerdem sollten alle Klassen gleich breit sein.

Zu vermeiden sind nach Möglichkeit offene Klassen, die am Rand der Verteilung auftreten können, wenn z.B. alle Daten $x \leq a_0$ zu einer einzigen Klasse zusammengefaßt werden sollen. - Für offene Klassen sind nämlich Klassenmitte sowie Klassenbreite offenbar nicht definiert.

"In der Mitte" einer Verteilung sollten keine leeren Klassen vorkommen; dies kann u.U. von der Kategorisierung des Merkmals abhängen.

Gibt es k verschiedene Klassen mit den Häufigkeiten $f_i$, so gilt für das arithmetische Mittel $\bar{x}$ angenähert

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{k} a_i * f_i \qquad \text{mit } n = \sum_{i=1}^{k} f_i \ ,$$

wobei insgesamt n Probanden erfaßt worden sind. - Der exakte Wert entsprechend der Formel von Seite 22 unten wird sich bei guter Klasseneinteilung nur wenig von diesem Näherungswert unterscheiden.

Es gilt die Regel, daß eine Zusammenfassung nach etwa sieben bis 10 Klassen günstig ist, wenn nicht besondere Gründe dagegen sprechen. Bei einer Untersuchung zur Körpergröße wird man beispielsweise die Klassenbreite 1 cm zugrunde legen und damit zwangsläufig mehr Klassen in Kauf nehmen.

Diese Ausführungen gelten sinngemäß für multivariable Verteilungen; am Beispiel von Körpergröße/Gewicht würde man die Daten tabellarisch in einer Kontingenztafel zusammenstellen:

| Größe [cm] | Gewicht [kg] ... | 50 | 51 | 52 | 53 | ... | |
|---|---|---|---|---|---|---|---|
| ... | | | | | | | ... |
| 160 | | 1 | | | | | ... |
| 161 | | | 2 | | | | ... |
| 162 | | 1 | | 1 | 1 | | ... |
| ... | | | | | | | ... |
| | ... | ... | ... | ... | ... | ... | n |

**Abb: Verteilungstafel einer bivariablen Verteilung**

Die Summe aller Spaltensummen oder Zeilensummen ist dann die Anzahl n aller erfaßten Datenpaare. Berechnet man $\bar{x}$ und $\bar{y}$ , so wird das Paar $(\bar{x}, \bar{y})$ der Schwerpunkt der Verteilung (X,Y). Im Beispiel wäre dies eine fiktive Person mit Durchschnittsgröße und Durchschnittsgewicht der erfaßten Population.