



# Lehr- und Handbücher der Politikwissenschaft

Herausgegeben von  
Dr. Arno Mohr

Bisher erschienene Werke:

- Gabriel · Holtmann, Handbuch  
Politisches System der Bundesrepublik Deutschland, 2. Auflage  
Glöckler-Fuchs, Institutionalisierung der europäischen  
Außenpolitik  
Jäger · Welz, Regierungssystem der USA, 2. Auflage  
Lehmkuhl, Theorien Internationaler Politik, 2. Auflage  
Lietzmann · Bleek, Politikwissenschaft – Geschichte und  
Entwicklung  
Mohr (Hrg. mit Claußen, Falter, Prätorius, Schiller, Schmidt,  
Waschkuhn, Winkler, Woyke),  
Grundzüge der Politikwissenschaft, 2. Auflage  
Naßmacher, Politikwissenschaft, 3. Auflage  
Pilz · Ortwein, Das politische System Deutschlands, 2. Auflage  
Schmid, Verbände  
Schumann, Repräsentative Umfrage, 2. Auflage  
Sommer, Institutionelle Verantwortung  
Wagschal, Statistik für Politikwissenschaftler  
Waschkuhn, Demokratietheorien  
Woyke, Europäische Union

# Statistik für Politikwissenschaftler

Von  
Dr. Uwe Wagschal

R. Oldenbourg Verlag München Wien

## **Die Deutsche Bibliothek - CIP-Einheitsaufnahme**

### **Wagschal, Uwe:**

Statistik für Politikwissenschaftler / von Uwe Wagschal. – München ;

Wien : Oldenbourg, 1999

(Lehr- und Handbücher der Politikwissenschaft)

ISBN 3-486-23847-7

© 1999 R. Oldenbourg Verlag

Rosenheimer Straße 145, D-81671 München

Telefon: (089) 45051-0, Internet: <http://www.oldenbourg.de>

Das Werk einschließlich aller Abbildungen ist urheberrechtlich geschützt. Jede Verwertung außerhalb der Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Bearbeitung in elektronischen Systemen.

Gedruckt auf säure- und chlorfreiem Papier

Gesamtherstellung: R. Oldenbourg Graphische Betriebe GmbH, München

ISBN 3-486-23847-7

## Vorwort

Dieses Lehrbuch ist das Resultat verschiedener Lehrveranstaltungen zur statistischen Methodenausbildung, die der Autor am Institut für Politische Wissenschaft der Universität Heidelberg sowie am Fachbereich 8 der Universität Bremen abgehalten hat. Dabei wurden den Teilnehmern Grundkenntnisse der empirischen Politikforschung und der Datenauswertung vermittelt. Das Lehrbuch richtet sich an Anfänger, wobei Vorkenntnisse nicht vorausgesetzt werden. Besonderer Wert wird auf Anschaulichkeit und Verständlichkeit gelegt. Aus diesem Grund wird auch auf einen doppelten Ausweis der weiblichen Form („Student und Studentin“, „StudentInnen“) verzichtet.

Die statistischen Verfahren sind anhand von Beispielen dargestellt, die sich an politikwissenschaftlichen Problemen und Fragestellungen orientieren. Sie sind einfach von Hand nachzurechnen. Für Verbesserungsvorschläge, Anmerkungen und Kritik ist der Autor jederzeit dankbar. Auf Übungsaufgaben und deren Lösungen wurde bewußt verzichtet, da sie den Rahmen des Buches über eine kritische Größe hinaus ausgedehnt hätten und weil jedes Verfahren ausführlich erklärt wird. Um dem Leser trotzdem die Möglichkeit zu geben, das gelernte Wissen zu vertiefen, kann eine Aufgabensammlung (mit Lösungen) beim Autor (Zentrum für Sozialpolitik, Parkallee 39, 28209 Bremen) angefordert werden.

Danken möchte ich Dr. Arno Mohr für die Aufnahme in diese Reihe. Daneben gab es viele tapfere Helfer (Freunde, die es hoffentlich geblieben sind, Studenten und Kollegen), die sich durch dieses Buch oder Teile davon, gearbeitet und es durch ihre Kommentare mit verbessert haben. Herausheben möchte ich Anna Schneider, Petra Riedle und Dr. Herbert Obinger, die sehr viel zum Gelingen mit beigetragen haben. Bernd Brinkhoff hat mir beim Einscannen von Graphiken geholfen. Teile des Buches haben Tobias Ostheim, Petra Heinzelmann, Iris Krimmel, Nico Siegel und Dr. Bernhard Kittel kommentiert. Prof. Johann Bacher gab mir wertvolle Verbesserungsvorschläge bei der Clusteranalyse, während Prof. Bergschlosser, Sven Quenter und Herbert Obinger mich vor Fehlern bei der QCA-Analyse bewahrt haben. Schließlich danke ich Prof. Manfred G. Schmidt für seine Hinweise. Manfred Schmidt hat mich nicht nur zu diesem Buch ermutigt, sondern mir außerdem in Heidelberg und Bremen stets exzellente Arbeitsbedingungen bereitgestellt, deren Finanzierung zum Teil aus dem Leibniz-Preis der Deutschen Forschungsgemeinschaft stammt.



## Inhaltsverzeichnis

1. Einleitung .....	1
2. Geschichte und Arbeitsbereiche der Statistik .....	6
2.1. Geschichte der Statistik .....	6
2.1.1. Die praktische Statistik .....	7
2.1.2. Die Universitätsstatistik .....	7
2.1.3. Politische Arithmetik .....	8
2.1.4. Die Wahrscheinlichkeitsrechnung .....	9
2.2. Definition, Bereiche und Quellen der Statistik .....	11
3. Der sozialwissenschaftliche Forschungsprozeß .....	16
4. Grundbegriffe der Statistik .....	25
4.1 Merkmale, statistische Einheiten und Massen .....	25
4.2. Meßniveaus und Skalierung .....	29
4.3. Das Adäquationsproblem .....	34
4.4. Gütekriterien für Tests und Meßverfahren .....	39
5. Datenerhebung und Datenaufbereitung .....	46
5.1. Primärdatenerhebung .....	46
5.1.1. Die Befragung .....	47
5.1.1.1. Welcher Erhebungsmodus ist zu wählen? .....	53
5.1.1.2. Befragungen im Internet .....	58
5.1.2. Die Beobachtung .....	60
5.1.3. Das Experiment .....	61
5.1.4. Die Inhaltsanalyse .....	64
5.1.5. Das Panel .....	67
5.2. Auswahlverfahren .....	68
5.3. Datenaufbereitung .....	75
6. Eindimensionale Häufigkeitsverteilungen .....	77
6.1. Nominalskalierte Merkmale .....	77
6.2. Ordinalskalierte Merkmale .....	80
6.3. Metrischskalierte Merkmale .....	82
6.4. Weitere Darstellungsmöglichkeiten von Häufigkeitsverteilungen .....	86

---

7. Maßzahlen bei univariaten Häufigkeitsverteilungen .....	93
7.1. Der Begriff der Maßzahl .....	93
7.2. Lageparameter .....	95
7.2.1. Der Modalwert (Modus) .....	95
7.2.2. Der Median .....	97
7.2.3. Das arithmetische Mittel .....	99
7.2.4. Das geometrische Mittel .....	102
7.2.5. Das harmonische Mittel .....	104
7.2.6. Sonstige Mittelwerte .....	105
7.2.7. Lageregeln der Mittelwerte .....	106
7.3. Streuungsparameter .....	107
7.3.1. Die Spannweite R (Range) .....	107
7.3.2. Quantilsabstände .....	109
7.3.3. Die mittlere absolute Abweichung D .....	111
7.3.4. Varianz und Standardabweichung .....	113
7.3.5. Der Variationskoeffizient .....	115
7.3.6. Der Boxplot (Box-and-Whiskers Plot) .....	116
7.3.7. Maße der Gestalt .....	118
7.4. Verhältniszahlen .....	121
7.4.1. Gliederungszahlen .....	121
7.4.2. Beziehungszahlen .....	123
7.4.3. Maßzahlen .....	124
8. Konzentrationsmaße .....	126
8.1. Die Lorenzkurve .....	127
8.2. Der Gini-Koeffizient .....	130
8.3. Weitere relative Konzentrationsmaße .....	135
8.3.1. Der Robin Hood Index .....	135
8.3.2. Verhältnis der Dezile .....	138
8.4. Einkommensungleichheit in Deutschland .....	139
8.5. Absolute Konzentrationsmaße .....	141
8.5.1. Die Konzentrationsrate .....	141
8.5.2. Der Herfindahl-Index .....	143
9. Zweidimensionale Häufigkeitsverteilungen .....	147

---

10. Messen von Zusammenhängen zwischen zwei Merkmalen .....	155
10.1. Die Prozentsatzdifferenz $d\%$ .....	156
10.2. Der Phi-Koeffizient $\phi$ .....	159
10.3. Yules $Q$ .....	164
10.4. Cramers $V$ und Tschuprows $T$ .....	165
10.5. Der Kontingenzkoeffizient $C$ .....	168
10.6. Lambda $\lambda$ .....	169
10.7. Die biseriale Rangkorrelation.....	172
10.8. Die punktbiseriale Korrelation .....	176
10.9. Das Zusammenhangsmaß Eta $\eta$ .....	178
10.10. Der Rangkorrelationskoeffizient $\rho$ nach Spearman .....	181
10.11. Kendalls Rangkorrelationskoeffizient $\tau$ -a .....	184
10.12. Goodman und Kruskals $\gamma$ .....	187
10.13. Kendalls Rangkorrelationskoeffizient $\tau$ -b und $\tau$ -c .....	191
10.14. Der Korrelationskoeffizient $r$ nach Pearson .....	193
10.15. Korrelation und Kausalität .....	203
11. Regressionsanalyse .....	209
11.1. Lineare Einfachregression .....	211
11.1.1. Das Grundprinzip der bivariaten Regression.....	211
11.1.2. Der Determinationskoeffizient $R^2$ .....	218
11.1.3. Annahmen und Probleme .....	222
11.2. Multivariate Regression .....	230
11.2.1. Der Grundgedanke der multivariaten Regression.....	230
11.2.2. Das Problem der Multikollinearität .....	236
11.2.3. Das Problem der Heteroskedastizität.....	239
11.2.4. Die Identifikation von Ausreißern .....	242
12. Clusteranalyse .....	246
12.1. Ziele der Clusteranalyse .....	246
12.2. Ähnlichkeits- und Distanzmaße .....	248
12.2.1. Ähnlichkeitsmaße .....	248
12.2.2. Distanzmaße .....	255
12.3. Verfahren zur Gruppenbildung .....	261
12.3.1. Das „Single-Linkage“ Verfahren .....	263
12.3.2. Das „Complete Linkage“ Verfahren.....	266

12.3.3. Graphische Darstellung der Cluster .....	269
12.3.4. Die Bestimmung der Clusterzahl .....	270
12.3.5. Weitere hierarchisch-agglomerative Fusionierungsverfahren .....	276
12.4. Die drei Welten des Wohlfahrtsstaates .....	278
13. Qualitative Comparative Analysis (QCA) .....	289
14. Statische Testverfahren und Konfidenzintervalle .....	305
14.1. Die Nullhypothese und die Alternativhypothese .....	305
14.2. $\alpha$ -Fehler und $\beta$ -Fehler .....	310
14.3. Konfidenzintervalle und Stichprobengröße .....	313
14.3.1. Der Stichprobenfehler des Mittelwertes .....	313
14.3.2. Die Berechnung von Konfidenzintervallen .....	316
14.3.3. Die Bestimmung des Stichprobenumfanges .....	317
14.4. Der t-Test auf den Unterschied zweier Mittelwerte .....	322
14.5. Verteilungsfreie Testverfahren .....	330
14.5.1. Der Wilcoxon-Test .....	330
14.5.2. Der U-Test nach Mann-Whitney .....	334
15. Die Darstellung quantitativer Informationen .....	338
16. Studienpraktische Tips und Hinweise .....	356
16.1. Hinweise zur Abfassung einer Examensarbeit .....	356
16.2. Statistische Software .....	360
16.3. Hilfreiche Adressen zur Informationsbeschaffung .....	364
16.3.1. Nationale Informationsquellen .....	364
16.3.2. Internationale Informationsquellen .....	365
Anhangstabelle 1: Standardnormalverteilung .....	367
Anhangstabelle 2: t-Verteilung und zweiseitige Signifikanzgrenzen .....	369
Literatur .....	370
Stichwortverzeichnis .....	382

*I would rather be wrong than vague.*

*David Easton*

## 1. Einleitung

Warum sollte man sich im politikwissenschaftlichen Studium mit Statistik beschäftigen? Diese Frage stellen sich viele Studenten im Grundstudium, sich düster an bittere Mathematikstunden in der gymnasialen Oberstufe erinnernd. Die Legitimation eines Themas ergibt sich im wesentlichen aus der Relevanz, die es für andere Bereiche oder Fragestellungen, zum Beispiel für die Forschung oder den Arbeitsmarkt besitzt. Hier schneidet die Statistik im Vergleich zu anderen wissenschaftlichen Hilfsdisziplinen und Mitteln der Erkenntnisgewinnung nicht schlecht ab. Warum ist es also lohnend, sich intensiver mit Statistik zu beschäftigen?

- (1) In den Studienplänen der meisten sozialwissenschaftlichen Fakultäten ist Statistik vorgeschrieben, und bei einem Studienortwechsel werden an der neuen Hochschule oftmals Statistikenntnisse verlangt.
- (2) Ein bedeutender Teil der Politikwissenschaft in Deutschland, aber vor allem in den anglo-amerikanischen Ländern wendet statistische Methoden an. An den meisten politikwissenschaftlichen Instituten in Deutschland wird man daher nicht umhinkönnen, sich mit Statistik zu beschäftigen. Es gibt bereits immer mehr Politikwissenschaftler, die auf Statistikenntnisse Wert legen, so daß diese zum Standard gehören.
- (3) Für das Verständnis eines großen Teils der fachwissenschaftlichen Literatur und Diskussion sind statistische Kenntnisse unentbehrlich.
- (4) Als Qualifikation für den Arbeitsmarkt sind Statistik- und Computerkenntnisse von Vorteil. In nahezu allen akademischen Berufen muß man zumindest Grundkenntnisse in Statistik vorweisen können, mit reiner „Verbalakrobatik“ ist heute kaum noch Karriere zu machen.
- (5) Statistik macht Spaß und ist interessant. Sicher werden viele Studenten das Fach Politikwissenschaft mit der Vorstellung gewählt haben, sich nie mehr mit Mathematik beschäftigen zu müssen, doch dies ist ein Irrtum. Es kann ungemein interessant sein, politikwissenschaftliche Fragen mit mathematischen und statistischen Methoden zu

bearbeiten und zu beantworten. Sinnvoll eingesetzt können sich quantitative und qualitative Politikwissenschaft gegenseitig gut ergänzen.

- (6) Wer bei Sachthemen mitreden will, braucht fundierte Informationen. Bei der Informationsflut in den Medien muß man allerdings aufpassen, keinen Manipulationen zu unterliegen. Der Konsument von Statistiken muß lernen, die wirklich relevanten Informationen zu erkennen. Außerdem dienen Informationen und Statistiken in der Auseinandersetzung zwischen Interessengruppen als die jeweilige „Munition“, das heißt, jede Partei oder Organisation kann sich Zahlen für ihre Zwecke zunutze machen. Daher sollte jeder Politikwissenschaftler in der Lage sein, wesentliche von unwesentlichen Daten zu trennen.
- (7) Statistiken verdichten Informationen zu einer Maßzahl. Deshalb ermöglicht die Statistik eine sachkundige Zusammenfassung und Bündelung von Informationen. Ferner stellt sie die Kennzahlen bereit, mit deren Hilfe man Hypothesen und Theorien überprüfen kann.

In der Politikwissenschaft selbst variiert die Beliebtheit der Statistik je nach Schulenzugehörigkeit des jeweiligen Wissenschaftlers. Die drei Hauptströmungen oder theoretischen Schulen sind (Falter 1987: 295ff.): die normativ-ontologische, die empirisch-analytische und die kritisch-dialektische Schule. Allerdings werden die Bezeichnungen in der Literatur durchaus abgewandelt gebraucht. Diese drei Metatheorien sind zur Orientierung in der Fachwissenschaft und zum generellen Verständnis der Politikwissenschaft unerlässlich.

1. Die normativ-ontologische Politikwissenschaft verbindet politisch-philosophische Fragestellungen mit einer vorwiegend institutionell geprägten Betrachtung des Politischen. Berühmte Vertreter dieser Schule waren E. Fraenkel, A. Bergstraesser, C. J. Friedrich und W. Hennis.

2. Die empirisch-analytische Politikwissenschaft. Die bedeutendste Strömung ist hierbei der Behaviorismus. Die Wurzeln dieser Schule liegen vor allem in den Vereinigten Staaten. Ein herausragender Vertreter dieser Schule war Karl Popper, der den sogenannten Kritischen Rationalismus begründet hat. Ein wichtiger Zweig dieser Schule ist die Wahl- und Einstellungsforschung.

3. Die dritte Schule ist die kritisch-dialektische Politikwissenschaft. Sie ist stark politökonomisch orientiert. Die Vertreter der kritisch-dialektischen Politikwissenschaft sind im wesentlichen gesellschaftskritischem Gedankengut verpflichtet. Prägende Strömungen innerhalb dieser Richtung waren der Marxismus, der Neo-Marxismus und die Frankfurter Schule.

Nicht jede dieser drei Schulen verwendet empirische Methoden zur Analyse von Politik. Vor allem die zweite Schule, die empirisch-analytische Politikwissenschaft und hier besonders der Behaviorismus, propagiert diese Methoden. Mitunter werden sie aber auch von Anhängern des kritisch-dialektischen Ansatzes verwendet.

„Empirie“ ist ein zentraler Begriff in der Charakterisierung der unterschiedlichen Schulen. Deshalb soll dieser Begriff vorab geklärt werden, da er im Verlauf eines sozialwissenschaftlichen Studiums häufig auftreten wird. Seine Verwendung in den Sozialwissenschaften ist nicht eindeutig. Daher wird hier eine häufig verwendete Definition dieses Begriffes angeben.

**Definition: „Empirie bezeichnet die auf Sinneserfahrung, Beobachtung, Messung, Experiment und andere Verfahren gründende wissenschaftliche Erkenntnis.“ (Nohlen 1994: 88).**

Quantitative Methoden stehen im Mittelpunkt der Anwendung von empirischen Verfahren. Leider muß man feststellen, daß der Begriff der „Empirie“ geradezu inflationär benutzt wird und die Verwendungen dieses Begriffs oft nur Lippenbekenntnisse darstellen, ohne daß tatsächlich empirische Ergebnisse zur Stützung eines Arguments, einer Hypothese oder einer Theorie referiert werden.

Ausgangspunkte dieses Buches sind die Geschichte und die Arbeitsbereiche der Statistik (Kapitel 2). Heutzutage wird Statistik eher mit Mathematik assoziiert als mit Politikwissenschaft, obwohl die historischen Wurzeln der Statistik - wie man sehen wird - eng mit politischen Fragestellungen verknüpft sind. Politische und ökonomische und nicht mathematische Interessen ließen den Bedarf nach Statistiken entstehen. Anschließend wird im dritten Kapitel der sozialwissenschaftliche Forschungsprozeß erörtert, der meistens idealtypisch mit Hilfe von Phasenmodellen dargestellt wird. Im vierten Kapitel sollen dann zentrale Grundbegriffe und Probleme der Statistik erörtert und erläutert werden. Um Daten analysieren zu können, muß man sie zunächst erheben. Das weite Feld der Datenerhebung und der Datenaufbereitung soll deshalb im fünften Kapitel behandelt werden. Die anschließenden Kapitel beschäftigen sich mit der Darstellung und der Analyse von Daten.

Den Auftakt hierfür geben eindimensionale Häufigkeitsverteilungen sowie ihre vielfältigen Darstellungsmöglichkeiten. Danach werden die wichtigsten Verteilungsmaßzahlen bei univariaten Häufigkeitsverteilungen (Lage- und Streuungsparameter) vorgestellt. Im darauf folgenden achten Kapitel werden wichtige Konzentrationsmaße vorgestellt. Die Kenntnis von

Konzentrationsmaßen ist deshalb angebracht, da diese in den unterschiedlichsten Teildisziplinen der politikwissenschaftlichen Literatur auftauchen, sei es bei Überlegungen zu Gerechtigkeits- oder Verteilungsfragen, bei Konzentrationserscheinungen auf der Wählerebene oder zur Charakterisierung von Parteiensystemen. Diese Konzentrationsmaße sind vielfältig einsetzbar und erfreuen sich eines regen Gebrauchs.

Das neunte Kapitel behandelt zweidimensionale Häufigkeitsverteilungen und liefert damit die Grundlagen für das Kapitel über die Zusammenhänge. Wenn bis zu diesem Kapitel immer nur eine Variable hinsichtlich ihrer Lage und Streuung betrachtet wurde, werden nun die Zusammenhänge zwischen zwei Variablen untersucht. Dazu werden im zehnten Kapitel die gängigsten Zusammenhänge hinsichtlich ihrer Berechnung und Interpretation erläutert. Das in den Sozialwissenschaften am meisten verwendete statistische Verfahren ist die Regressionsanalyse (Kapitel 11). Die Regressionsanalyse hängt eng mit der im zehnten Kapitel vorgestellten Korrelationsanalyse zusammen. Das Verständnis dieser beiden Methoden öffnet gleichzeitig die Tür für das Verständnis vieler Fachpublikationen. Als weitere Methode wird die Clusteranalyse erörtert, die zu den multivariaten Auswertungsmethoden zählt (Kapitel 12). Eine relativ neue multivariate Untersuchungsmethode, die QCA-Analyse (Qualitative Comparative Analysis) von Charles Ragin wird dann in Kapitel 13 erklärt.

Von Bedeutung sind ferner Schätz- und Testverfahren der Statistik, von denen wichtige Verfahren im Kapitel 14 thematisiert werden. Schließlich sollen im vorletzten Kapitel (Kapitel 15) - ausgehend von den bahnbrechenden Arbeiten Edward Tuftes - Regeln für die graphische Darstellung quantitativer Informationen vorgestellt werden. Wissenschaftliche Argumente können von der Qualität und Erklärungskraft von Schaubildern leben. Abschließend werden im letzten Kapitel studienpraktische Tipps, zum Beispiel zu statistischer Software und zur Abfassung einer wissenschaftlichen Arbeit gegeben.

Die Funktion der Statistik innerhalb der Politikwissenschaft ist die einer Hilfswissenschaft. Statistik um der Statistik willen kann nicht das Ziel ihrer Anwendung sein. Im Vordergrund stehen gesellschaftliche, politische und ökonomische Phänomene, die mit Hilfe des statistischen Instrumentariums beschrieben und erklärt werden sollen. Vergleicht man die Politikwissenschaft mit anderen benachbarten Disziplinen, wie der Ökonomie, der Soziologie und der Psychologie, so muß man allerdings konstatieren, daß die Statistik innerhalb der Disziplin Politikwissenschaft - zumindest im deutschsprachigen Raum - schwach ausgeprägt ist.

Aber die Anwendung nimmt ohne Frage zu. Darauf weisen auch verschiedene Umfragen hin, die unter den Fachvertretern der Zunft durchgeführt wurden (Honolka 1986, Klingemann

---

und Falter 1997). Mitte der achtziger Jahre waren Forschungsrichtungen, in denen die Statistik zum Einsatz kam eher marginal, so die Befunde von Honolkas Reputationsstudie unter Deutschlands Politikwissenschaftlern, die 1985 durchgeführt wurde. Elf Jahre später wurden wieder die professionellen Politikwissenschaftler befragt und es zeigt sich, daß Theorieansätze, die quantifizierende Methoden verwenden, einen deutlichen Anstieg erfahren haben (Klingemann und Falter 1997). Ein Befund dieser Umfrage gab zudem Antwort auf die Frage, an welcher Universität oder Forschungsstätte die wichtigsten Arbeiten zu dem Gebiet der Wissenschaftstheorie und der Methoden der Politikwissenschaft geleistet werden. An der Spitze lag die Universität Mannheim, gefolgt von der Universität Mainz sowie dem Wissenschaftszentrum Berlin.

## 2. Geschichte und Arbeitsbereiche der Statistik

### 2.1. Geschichte der Statistik

Statistische Zahlen begegnen jedem Menschen - bewußt oder unbewußt - in großer Zahl. Es beginnt schon mit der Zeitungslektüre am Morgen. Am Frühstückstisch treffen wir auf Wahlergebnisse, die neuesten Arbeitslosenzahlen, die aktuellsten Statistiken zur Inflation und Staatsverschuldung, die Börsen- und Wechselkurse im Wirtschaftsteil und, nicht zu vergessen, die Fußballtabellen und Torjägerlisten im Sportteil. Beispiele für Statistiken und ihre vielfältigen Darstellungsformen - gute wie schlechte - gibt es millionenfach. Statistische Zahlen, Kolonnen und Graphiken drohen einen geradezu zu erschlagen. Beinahe jede Aussage und Meinung kann mit einer Statistik begründet werden - das gleiche gilt oft für die korrespondierende Gegenaussage. Welche Aussage ist nun richtig, fragt man sich. Zur Illustration dient ein Beispiel von Krämer (Krämer 1995: 23):

*„Ein Einzelhändler kauft eine Ware für 100 DM ein und schlägt sie für 200 DM los. Wieviel Prozent macht seine Handelsspanne aus? „Eine Unverschämtheit!“ sagen wir als Kunde. „Ein Aufschlag von glatt 100 Prozent!“*

*„So schlimm ist die Sache nun auch wieder nicht“, sagt der Händler. „50 Prozent Verdienst sind wirklich nicht zuviel.“*

*Offenbar haben hier beide Parteien recht. 100 Mark sind 100 Prozent von 100 Mark und 100 Mark sind 50 Prozent von 200 Mark. Trotzdem ist der Eindruck je nach der Basis durchaus ein anderer.“*

Oft sind Statistiken falsch oder gar bewußt manipuliert. Einige Beispiele hierfür werden wir im Laufe dieses Buches zu Gesicht bekommen. Daneben gibt es eine breite Literatur über „manipulierte Statistiken“ (z.B. Huff 1954; Tufte 1983; Krämer 1995; Ketteler 1997). Bevor wir uns aber mit statistischen Problemen beschäftigen, sollen zuerst jedoch die historischen Wurzeln der heutigen Statistik dargestellt werden.

Theoretische Statistik gibt es seit rund drei Jahrhunderten und die angewandte Statistik seit rund viereinhalb tausend Jahren. Insgesamt kann man vier Quellen der Statistik unterscheiden (Menges 1982; Zwer 1985; Zwer 1986):

- 1.) die praktische Statistik
- 2.) die Universitätsstatistik
- 3.) die Politische Arithmetik
- 4.) die Wahrscheinlichkeitsrechnung

### **2.1.1. Die praktische Statistik**

Der älteste Zweig der Statistik ist die praktische Statistik. Wie der Name sagt, waren praktische Ziele die Ursache für ihre Entstehung. Man interessierte sich für die Zahl der Bewohner und ihre wirtschaftliche Leistungsfähigkeit. Ihre Geschichte beginnt im „Alten Reich“ in Ägypten (2650 - 2190 v. Chr.). Wohl aus fiskalischen Gründen führte man zweijährige Zählungen des Goldes und der landwirtschaftlichen Nutzflächen durch. Weitere Gründe für das Erheben von Statistiken waren der Unterhalt des Militärs sowie große Bauvorhaben (Pyramidenbau). Daneben wurden in anderen Großreichen, wie China (ca. 2300 v. Chr.) und dem persische Großreich (ca. 500 v. Chr.), statistische Zahlen ermittelt.

Im antiken Griechenland hatte die Statistik kaum Bedeutung, was an der relativen Überschaubarkeit der Stadtstaaten lag - ganz im Gegensatz zu den Römern, die mit wachsender Größe ihres Reiches verstärkt statistische Erhebungen durchführten. Bereits 550 v. Chr. soll der König Servius Tullius einen Zensus der römischen Bürger verfügt haben. Bei den Römern wurden erstmals periodische Erhebungen für die gesamte Bevölkerung durchgeführt - insgesamt können 69 Volkszählungen als gesichert gelten. Die berühmteste fand vor rund 2000 Jahren statt und begegnet uns jedes Jahr um die Weihnachtszeit wieder.

Die Germanen kannten keine Statistik, und das Mittelalter war im Hinblick auf die Anwendung von Statistik ebenfalls dürftig. In der Neuzeit änderte sich dies: Es kam, wegen eines gestiegenen Bedarfs an Informationen, zu einer Expansion der Statistik. Zentral war immer noch die Bevölkerungsstatistik. Nach und nach traten dann aber auch andere Wirtschaftsstatistiken hinzu. Dazu kamen die ersten Gründungen statistischer Zentralämter, so 1796 in Schweden, 1797 in Norwegen, 1800 in Frankreich und 1829 in Österreich. 1834 wurde das statistische Zentralbureau des Deutschen Zollvereins und 1871 das Kaiserliche Statistische Reichsamt gegründet.

### **2.1.2. Die Universitätsstatistik**

Von der Datenerhebung aus fiskalischen, militärischen und administrativen Gründen durch die praktische Statistik war es nur ein kleiner Schritt, die vorliegenden quantitativen Informationen zu systematischen Beschreibungen eines Landes beziehungsweise mehrerer Länder zu verwenden. In diesem Sinne trat die Statistik im 16. Jahrhundert als deskriptive

Disziplin in die neuere Wissenschaftsgeschichte ein. Ziel war die systematische Staatenbeschreibung. Der erste dieser „Ländervergleiche“ wurde von dem Italiener Francesco Sansovino (1521-1586) im Jahre 1561 verfaßt, wobei er 22 Staaten verglich. Insofern liegt in der Universitätsstatistik auch eine der historischen Wurzeln der vergleichenden Politikwissenschaft.

Im übrigen hieß die Universitätsstatistik - dieser Name kam erst später auf - „Lehre von den Staatsmerkwürdigkeiten“ - also genau das, was auch von der Politikwissenschaft geleistet wird. Der Name Universitätsstatistik kommt daher, daß die Hauptvertreter dieser Richtung Gelehrte an den entstehenden Universitäten waren, deren Fächerkanon jedoch noch nicht ausdifferenziert war. So umfaßte die „Lehre von den Staatsmerkwürdigkeiten“ die Bereiche Geschichte, Geographie und Staatskunde. Die Hauptvertreter dieser historischen Vorläufer der Statistik kamen aus Deutschland.

Im Jahre 1660 hielt in Helmstedt Hermann Conring (1606-1681) die erste Statistikvorlesung der Geschichte - allerdings ohne sie auch als solche zu bezeichnen. Zum Namensgeber der Statistik wurde Martin Schmeitzel (1679-1747), der an der Universität Halle lehrte. Er hielt eine Vorlesung mit dem Namen „collegium politico-statisticum“, womit der Name geboren war. Im Italienischen bedeutet „statista“ Politiker beziehungsweise Staatsmann. Wie man sieht, kann man die Statistik schon aus Tradition im Fach Politikwissenschaft ansiedeln. Der bedeutendste Vertreter der Universitätsstatistik war aber ein Schüler von Schmeitzel mit dem Namen Gottfried Achenwall (1719-1772). Aber auch diese Schule verschwand wieder. Der Zusammenbruch der Universitätsstatistik hängt mit der Ausdifferenzierung der Wissenschaften und mit der Gründung der nationalen statistischen Ämter zusammen.

### **2.1.3. Politische Arithmetik**

Die Politische Arithmetik war nicht deskriptiv (d.h. beschreibend), sondern analytisch (d.h. erklärend) ausgerichtet. Als erste der historischen Wurzeln suchte sie nach Gesetzmäßigkeiten in Wirtschaft und Gesellschaft. Die Anfänge der Politischen Arithmetik liegen im England des 17. Jahrhunderts. Im Jahre 1662 veröffentlichte der Londoner John Graunt (1620-1674) eine Untersuchung über Bevölkerungsgesetzmäßigkeiten, die auf Geburts- und Totenlisten der Stadt London basierte. Aus ihnen folgerte er Gesetzmäßigkeiten des Bevölkerungswachstums, der Fruchtbarkeit, der Sterblichkeit, des Altersaufbaus usw. Sein besonderes Anliegen war, die Ursachen für die Pestepidemien in London im 17. Jahrhundert zu finden, und

es nicht nur bei einer reinen Auflistung der Toten zu belassen. Dieser neue Ansatz in der Statistik läßt ihn gleichzeitig zu einem der ersten „Policy-Forscher“ werden, denn im Prinzip tat Graunt nichts anderes, als die Gesundheitspolitik seines Landes zu analysieren. Der herausragende Vertreter dieser Schule ist der Nationalökonom William Petty (1623-1687), gleichzeitig ein Freund von Graunt, der ihr auch mit seinem Hauptwerk „Political Arithmetic“ den Namen gab. Bedeutend war auch William Playfair (1759-1823), der als erster Wissenschaftler Graphiken zur Deskription und Analyse seines Forschungsgegenstandes verwendete, auch als Herausgeber von Adam Smith' Werk „Wealth of Nations“.

Die Politische Arithmetik hat sich in Deutschland nie richtig durchsetzen können. Sie blieb vor allem auf England, Holland, Frankreich und Belgien beschränkt. Ein interessanter Vertreter dieser Schule war der Belgier Quéételêt (1796-1874), der mit seinem Werk Generationen von Statistikern beeinflusste. Während bei einigen Vertretern statistische Zahlen als Ausdruck der göttlichen Ordnung bewundert wurden, reduzierte Quéételêt die Erklärung und Analyse der statistischen Zahlen auf das Wirken von Naturgesetzen. Zwar besteht die Aufgabe der Statistik unter anderem darin, von den Individualitäten der Einzelercheinungen abzusehen und mittlere Größen zu berechnen (Mittelwerte), aber Quéételêt wollte die zahlreichen Durchschnitte menschlicher Eigenschaften zu einem „Durchschnittsmenschen“ zusammensetzen. Jede Abweichung von diesem „mittleren Menschen“ deklarierte er als störende Zufälligkeit. Dieses Wesen besitzt mittleres Glück, mittlere Intelligenz, mittlere Schönheit, mittlere Heiratsneigung, mittlere Verbrechenneigung usw... Quéételêt glaubte sich im Besitz einer Formel, nach der die Gesellschaft berechnet werden könnte, was als äußerst fragwürdig zu bewerten ist. Immerhin waren die Politischen Arithmetiker jedoch die ersten, die nach systematischen Regelmäßigkeiten suchten, entweder basierend auf Naturgesetzen oder als Ausdruck einer göttlichen Ordnung.

### **2.1.4. Die Wahrscheinlichkeitsrechnung**

Derjenige Bereich, der heute mit Statistik am stärksten assoziiert wird, ist die Wahrscheinlichkeitsrechnung. Es ist ein historisches Kuriosum, daß der Name „Statistik“ von einem Extrem, der Universitätsstatistik, über die eine Mittelstellung einnehmende Politische Arithmetik zum anderen Extrem, der Wahrscheinlichkeitsrechnung, übergegangen ist.

Im Mittelpunkt der Wahrscheinlichkeitsrechnung stehen die Gesetze des Zufalls. Für viele ihrer Vertreter ist allein diese Form der Statistik als ein Zweig der angewandten Mathe-

matik anzusehen. Der Fachbegriff für diesen Zweig lautet Stochastik, abgeleitet vom griechischen Verb „vermuten“. Die Wahrscheinlichkeitsrechnung wird als das Mittel gesehen, welches zum Schließen von Zusammenhängen von einer Stichprobe auf eine Grundgesamtheit notwendig ist.

Die wichtigsten Pioniere kommen aus England, Frankreich, und Rußland, aber auch die Schweiz hat mit Leonard Euler (1707-1783) und Deutschland mit Carl-Friedrich Gauss (1777-1855) zwei wichtige Protagonisten beigesteuert. Die Wahrscheinlichkeitsrechnung wurde zuerst durch ein praktisches Interesse von Glücksspielern initiiert. Die französischen Mathematiker Blaise Pascal (1623-1662) und Pierre de Fermat (1601-1665) lösten im 17. Jahrhundert die Fragen, die ihnen ein Glücksspieler vorlegte. Weitere wichtige Namen und Vertreter sind Jakob und Nikolaus Bernoulli aus der Schweiz, Francis Galton, und Karl Pearson aus England, ohne daß diese Liste nur annähernd vollständig wäre (Zwer 1985: 10).

Ein für die Politikwissenschaft relevanter Vertreter der Wahrscheinlichkeitsrechnung ist der Franzose Marquis de Condorcet (1743-1794), ein Mathematiker und Politiker (Präsident der Nationalversammlung 1792), der das nach ihm benannte Wahlparadoxon entdeckt hat, welches - zugegeben - wenig mit der Wahrscheinlichkeitsrechnung zu tun hat. Das Condorcet-Paradoxon illustriert das Problem zyklischer Mehrheiten, das heißt, es können sich widersprüchliche Präferenzordnungen ergeben, wenn alle Gruppenmitglieder beziehungsweise alle Gruppen rational handeln. Folgendes Beispiel soll das Problem veranschaulichen:

Tabelle 2.1: Zyklische Mehrheiten (Condorcet-Paradox)

	Wähler		
	I	II	III
Rangordnung	A	C	B
der	B	A	C
Alternativen	C	B	A

Es gibt 3 Wähler (I, II, III) sowie 3 Alternativen, die zur Entscheidung stehen (A, B, C), wobei die relative Mehrheit das Entscheidungskriterium ist. Wähler I hat die stärksten Präferenzen für Alternative A, geringere für B und die wenigsten für C (vgl. Tabelle 2.1). Im folgenden kommt es zu einer paarweisen Abstimmung der Alternativen. Bei einer ersten Abstimmung der Alternativen B versus C gewinnt die Alternative B mit 2 zu 1 Stimmen, denn die Wähler I und III schätzen B jeweils höher als C. Anschließend kommt es zu einer Abstimmung des Gewinners der ersten Runde, also B, gegen die Alternative A. Hier gewinnt die Alternative A, wiederum mit 2 zu 1 Stimmen (wegen Wähler I und II). Danach wird A gegen

C zur Abstimmung gebracht, mit dem Ergebnis, daß C mit 2 zu 1 gewinnt. Die Präferenzordnung ist folglich nicht transitiv. Transitivität bedeutet: Wenn A mehr als B geschätzt wird und B höher als C, dann muß auch A höher als C geschätzt werden. Es ergibt sich folgende Reihenfolge, wenn immer paarweise über die Alternativen abgestimmt wird („>“ bedeutet „gewinnt gegen“):  $A > B > C > A$ . Je nachdem an welcher Stelle man in diesem Zyklus die paarweisen Abstimmungen beginnt, erhält man ein unterschiedliches Endergebnis. Dieser Sachverhalt wird als zyklische Präferenz (bzw. zyklische Mehrheit) bezeichnet. Dies bedeutet, daß mit Hilfe demokratischer Entscheidungen - hier des Mehrheitsentscheids - Probleme nicht immer widerspruchsfrei gelöst werden können. Darüber hinaus gibt es immer eine überstimmte Mehrheit, da es immer mehr Wähler gibt, die eine andere Alternative vorziehen würden. Da für das tatsächliche Endergebnis entscheidend ist, welche Abstimmungsreihenfolge gewählt wird, fällt dem Versammlungsleiter eine große Bedeutung bei Abstimmungen zu.

## 2.2. Definition, Bereiche und Quellen der Statistik

Wie kann man Statistik definieren? Allgemein bedeutet eine Definition die Feststellung des Gebrauchs eines sprachlichen Ausdrucks, in unserem Fall also der Statistik. Man kann weiter unterteilen in das, was definiert wird (= Definiendum) und in das Definiens, welches derjenige Teil einer Definition ist, der das Definiendum erklärt (= definiert). Der Student sollte sich immer vor Augen halten, daß es viele Möglichkeiten gibt, einen interessierenden Gegenstand zu definieren. Dazu finden sich in der Definitionslehre verschiedene Arten von Definitionen. Eine erste Unterscheidung ist die zwischen festgestellten (oder etymologischen) sowie festgesetzten Definitionen. Etymologische Definitionen beziehen sich auf den Sprachgebrauch, während festgesetzte Definitionen willkürlich - nach subjektiver Meinung des Autors - den Inhalt festlegen. So hat das Wort Statistik einen lateinischen Ursprung im Wort „Status“ (Staat, aber auch Zustand) und - wie bereits erwähnt - im italienischen Wort „Statista“ (Staatsmann) (Kennedy 1985: 67).

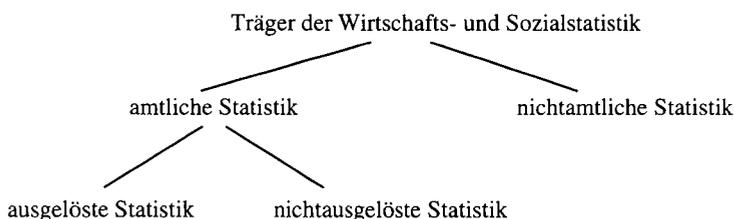
Eine weitere Unterscheidung wird zwischen Nominal- und Realdefinition getroffen. Dabei wird bei einer Nominaldefinition das Definiendum durch einen anderen Begriff ersetzt oder auf diesen zurückgeführt, zum Beispiel: Statistik ist die Anwendung quantitativer Methoden. Realdefinitionen geben die Merkmale des Definiendums an oder beschreiben die Art seines Gebrauchs näher: „Statistik ist Inbegriff theoretisch fundierter, empirischer objektivierter Daten“ (Menges 1982: 19). Die Elemente dieser Definition sind:

- 1.) Die Theoriegebundenheit, die darauf aufmerksam macht, daß willkürliches Datensammeln unwissenschaftlich ist.
- 2.) Die Eigenschaft, daß Daten empirischen Charakter haben müssen, bedeutet, daß sie aus der Wirklichkeit gewonnen werden müssen.
- 3.) Die Daten müssen objektivierbar sein, das heißt, der Erhebungsprozess muß aufgrund von sachlichen und nachvollziehbaren Meßvorschriften gewonnen werden.

Desweiteren kann man noch zwischen deskriptiven (d.h. beschreibenden) sowie präskriptiven Definitionen unterscheiden. Dabei enthalten präskriptive Definitionen Sollens-Aussagen. Sie beinhalten also ethisch-normative Vorstellungen. Anhand des idealtypischen Begriffs der „Globalisierung“, der in jüngster Zeit in der Politikwissenschaft große Aufmerksamkeit erfahren hat, soll die Differenz zwischen beiden Definitionsarten veranschaulicht werden. Holm und Sorensen liefern die beschreibende Definition: „Globalisierung ist die Intensivierung wirtschaftlicher, politischer, sozialer und kultureller Beziehungen über Grenzen hinweg“ (Holm und Sorensen 1995: 4; Übersetzung d.V.). Eine präskriptive Variante wäre: Globalisierung ist der Prozeß hin zu einer friedlichen Weltgesellschaft (in Anlehnung an Zürn 1996: 124).

Nach diesem Exkurs über Definitionen wird danach gefragt, wie Statistik weiter systematisiert werden kann. Statistiken finden sich in nahezu jedem wissenschaftlichen Bereich, zum Beispiel als Wahldaten in der Politikwissenschaft, ökonomische Daten in der Volkswirtschaftslehre, in der Soziologie, Psychologie, Physik, Medizin et cetera. Die statistischen Methoden und Untersuchungsverfahren sind in diesen Disziplinen jedoch immer dieselben. Die Statistik selbst läßt sich in zwei große Bereiche unterscheiden: die Wirtschaftsstatistik und die Statistischen Methoden. Letztere können noch weiter in die deskriptive Statistik sowie in die induktive Statistik (Inferenzstatistik) unterteilt werden. Der Schwerpunkt dieses Buches liegt auf der deskriptiven Statistik, also der beschreibenden Statistik, während der große Bereich der induktiven Statistik („schließende Statistik“), der sich mit Schlüssen von Stichproben auf die Grundgesamtheit beschäftigt, nur am Rande behandelt wird, da er für die praktische politikwissenschaftliche Forschung von etwas geringerer Bedeutung ist. Ein weiteres Unterscheidungsmerkmal der Statistik ist, wer Träger, also Produzent der Statistik ist. Man unterscheidet hierbei nach: 1. amtlicher Statistik und 2. nichtamtlicher Statistik.

Abbildung 2.1.: Träger der Statistik



Quelle: von der Lippe (1990: 4)

Die amtliche Statistik kann weiter untergliedert werden in ausgelöste Statistik und in die nichtausgelöste Statistik (Ressortstatistik). Die ausgelöste Statistik wird von dem Statistischen Bundesamt sowie den Statistischen Landesämtern betrieben. Die Bundesstatistiken werden, soweit es das Bundesstatistikgesetz (BStatG) oder eine sonstige Rechtsvorschrift nichts anderes bestimmt, durch Gesetz angeordnet (§ 5 BStatG). Nach Art. 73 Grundgesetz hat der Bund die ausschließliche Gesetzgebungskompetenz für die Statistik. Die Ressortstatistik betrifft die Ministerien im Bund und in den Ländern sowie deren nachgeordnete Behörden, deren Aufgabe es eben primär nicht ist Statistiken zu erstellen. Neben der amtlichen Statistik steht die nichtamtliche Statistik. Deren Träger (= Produzenten) können die Wirtschaftsverbände, Arbeitgeberverbände, Arbeitnehmerorganisationen, Markt- und Meinungsforschungsinstitute sowie die wirtschaftswissenschaftlichen Forschungsinstitute sein.

Zunehmende Wichtigkeit für die Politikwissenschaft, besonders bei der international vergleichenden Staatstätigkeitsforschung, erhalten die internationalen Produzenten von Statistiken, wobei man hier nach supranationalen und internationalen Organisationen unterscheiden kann. Wichtige Datenlieferanten (Adressen siehe Kapitel 16) sind:

- 1.) das Statistische Amt der Europäischen Union mit Sitz in Luxemburg (EUROSTAT),
- 2.) die OECD (Organisation for Economic Co-Operation and Development in Paris),
- 3.) der Internationale Währungsfond (IMF) in Washington,
- 4.) die Weltbank in New York,
- 5.) die ILO (International Labour Organisation) mit Sitz in Genf,
- 6.) die FAO (Food and Agriculture Organisation - Welternährungsbehörde) in Rom,
- 7.) die WTO (Welthandelsorganisation, vormals das GATT) in Genf.

Für die Bundesrepublik Deutschland produziert das Statistische Bundesamt (StBA) die wichtigsten Statistiken. Als selbständige Bundesoberbehörde untersteht es der Dienstaufsicht des Bundesministers des Inneren. Seine vordringlichste Aufgabe ist es, die Statistiken für

Bundeszwecke methodisch und technisch vorzubereiten und weiterzuentwickeln. Um sich über das Arbeitsgebiet und die Arbeitsweise des StBA zu informieren, empfiehlt sich die Publikation „Das Arbeitsgebiet der Bundesstatistik“ (letzte Ausgabe 1988). Daneben gibt es zahlreiche Fachserien mit statistischen Daten (insgesamt 19), die selbst noch einmal tief untergliedert sind, beispielsweise zu Wahlen. Wichtige Informationen enthält auch das „Statistische Jahrbuch für die Bundesrepublik Deutschland“, sowie die monatliche Fachpublikation „Wirtschaft und Statistik“. Aus Sicht der Politikwissenschaft ist interessant, daß der Präsident des StBA gleichzeitig Bundeswahlleiter und somit verantwortlich für die Durchführung der Bundestagswahlen ist. Analoges gilt für die Länderebene und für Großstädte: Die jeweiligen Leiter der statistischen Ämter sind auch für die Durchführung der Wahlen verantwortlich. Besondere Bedeutung genießen ebenso die Publikationen der Deutschen Bundesbank, die der nichtausgelösten Statistik zuzurechnen sind, die neben umfangreichen Statistiken in ihren Monatsberichten noch in fünf Reihen verschiedene Wirtschaftsstatistiken publiziert.

Weitere Möglichkeiten der Datenbeschaffung sind die nationalen und internationalen Datenarchive. Für Deutschland ist hier das Zentralarchiv (ZA) für empirische Sozialforschung in Köln führend. Das ZA archiviert Primärmaterial (Daten, Fragebögen, Kodepläne) und Ergebnisse empirischer Untersuchungen, um sie für wissenschaftliche Sekundäranalysen aufzubereiten und der interessierten Öffentlichkeit zugänglich zu machen. Der Arbeitsbereich des ZA erstreckt sich auf alle Fachgebiete, in denen Verfahren der empirischen und historischen Sozialforschung verwendet werden. Dort kann man jederzeit Datensätze gegen einen geringen Unkostenbeitrag bestellen. Weitere bedeutende internationale Datenarchive sind in Colchester (University of Essex in England) sowie das Archiv des „Inter-university Consortium for Political and Social Research“ (ICPSR), das am „Institute for Social Research at the University of Michigan“ eingerichtet ist. Das ICPSR-Archiv ist das weltgrößte Datenarchiv für sozialwissenschaftliche Daten. Es ist möglich, aus beiden Archiven Datensätze zu beziehen. Schneller, bequemer und oft auch kostengünstiger ist der Weg über das Zentralarchiv in Köln. Weitere Quellen der Datenbeschaffung sind zum Beispiel die Botschaften in Deutschland und die nationalen statistischen Ämter. Durch die zunehmende Technologisierung der Informationsbeschaffung wird diese wesentlich vereinfacht. Die Möglichkeit der Recherche in internationalen Datennetzen, wie dem Internet, macht es mittlerweile jedem Studenten möglich, sich Daten aus verschiedenen Ländern relativ einfach zu beschaffen, da fast alle bedeutenden statistischen Ämter im World Wide Web vertreten sind. Über die Homepage des Statistischen Bun-

desamtes (<http://www.statistik-bund.de/>) kommt man sehr einfach zu diesen Ämtern sowie zu den meisten statistischen Landesämtern in Deutschland. Daneben ist es auch möglich, in allen erwähnten Datenarchiven nach interessanten Datensätzen zu recherchieren.

Wichtige politikwissenschaftliche Datenquellen sind desweiteren noch Datenhandbücher mit politischen und institutionellen Variablen. Hier gibt es eine ganze Reihe bedeutender Publikationen, wie das HIWED (Historische Indikatoren der westeuropäischen Demokratien) Forschungsprojekt von Peter Flora (Flora 1975), das Cross-Polity Handbook von Arthur Banks (Banks 1963), das „Political Data Handbook“ von Lane et al. (Lane, McKay und Newton 1997) sowie das „World Handbook of Political and Social Indicators“ von Taylor und Jodice (Taylor und Jodice 1983), um hier nur einige wenige zu erwähnen.

### 3. Der sozialwissenschaftliche Forschungsprozeß

Es ist allgemein üblich, den sozialwissenschaftlichen Forschungsprozeß in Phasen einzuteilen. Diese Vorgehensweise erscheint plausibel und bietet sich auch intuitiv an, da der Forschungsprozeß ein dynamischer Vorgang ist. Man kann die Phasen des Forschungsprozesses unterschiedlich - mehr oder weniger grob - untergliedern. Im folgenden sollen verschiedene Unterteilungen vorgestellt werden, die sich in der Literatur finden. In Anlehnung an Patzelt (1985: 9 ff.) kann man vier Phasen unterscheiden:

(1) Die Konzeptualisierungsphase:

Der erste Schritt dieser Phase ist die Problemauswahl. In der Konzeptualisierungsphase sollte ferner die Fragestellung klar herausgearbeitet werden, ein Punkt, der zentral für das Gelingen einer empirischen Forschungsarbeit ist. Zudem werden in diesem Forschungsabschnitt die Hypothesen formuliert, die man untersuchen möchte. Voraussetzung dafür ist ein theoretischer Rahmen, in dem man sich bewegt. In dieser Phase stellen sich weiter folgende Fragen: Was für Daten werden erhoben? Individual- oder Aggregatdaten? Führt man Vollerhebungen durch oder zieht man nur eine Stichprobe aus der Grundgesamtheit?

(2) Die Datenerhebungsphase:

Grundlage der Datenerhebungsphase ist der theoriegeleitete Suchprozeß. Willkürliches Datensammeln ist unwissenschaftlich und wird wahrscheinlich zu uninteressanten Ergebnissen führen. Welche Arten von Daten erhoben werden, hängt von den Variablen ab, die sich bei der Hypothesenbildung herauskristallisiert haben. Die erste Phase ist mithin unverzichtbare Voraussetzung für den zweiten Forschungsabschnitt. In dieser Phase werden zudem die Daten in einer Datenmatrix (vgl. Abbildung 3.1.) zusammengefaßt und aufbereitet. In der Datenmatrix werden in der Kopfzeile die Variablen angeordnet (z.B. Arbeitslosenquoten, Antworten zu Einstellungsfragen oder institutionelle Eigenschaften). In der ersten Spalte werden die Untersuchungseinheiten (Merkmalsträger) abgetragen. Dies können Individuen, Länder oder Unternehmen sein. Die Anlage einer Datenmatrix ist für die Analyse der Daten mit Hilfe eines statistischen Softwareprogramms unabdingbar.

(3) Die Auswertungsphase:

In dieser Phase werden die erhobenen Variablen und Daten analysiert. Dazu gibt es mehrere Möglichkeiten, wie die Bildung von Häufigkeitstabellen, die Berechnung von statistischen Kennzahlen, wie Mittelwerte, oder die Berechnung von Zusammenhangsmaßen. Da

man in der Politikwissenschaft primär an Vergleichen interessiert ist, kann man unterschiedliche Vergleichsperspektiven einnehmen. Ein horizontaler Vergleich der Variablen, über eine Untersuchungseinheit gibt Informationen über die „Gestalt“ oder Performanz dieses Merkmalsträgers. Beispielsweise hat eine Person mit ihren Eigenschaften Größe, Geschlecht, Einkommen und Einstellung ein spezifisches Profil. Beim vertikalen Vergleich über alle Merkmalsträger hinweg kann die Verteilung für eine Variable ermittelt werden, z.B. die Einkommensverteilung für alle untersuchten (befragten) Personen.

Weiter kennt man als Vergleichsperspektive den Querschnittvergleich (engl. cross-sectional analysis). Hier werden die Daten über geographische oder sektorale Einheiten, meist Länder oder territoriale Einheiten eines Landes, hinweg zu einem Zeitpunkt  $t$  verglichen. Die Datenanalyse ist statisch, das heißt, Aussagen über den zeitlichen Verlauf sind nicht möglich. Man führt also einen synchronen Vergleich durch.

Beim vertikalen Vergleich der Variablen über die Zeit hinweg läßt sich die Veränderung und Entwicklung der untersuchten Variablen feststellen. Dabei spricht man auch von einer Zeitreihenanalyse (longitudinal analysis = Längsschnittvergleich) oder einem diachronen Vergleich.

Werden sowohl der horizontale als auch der vertikale Vergleich gleichzeitig durchgeführt, spricht man von Panel-Untersuchungen oder Panel-Analysen (Englisch: Pooled Time Series Analysis). Panel-Untersuchungen kommen nicht nur in der Markt-, Wahl- und Einstellungsforschung vor, sondern mittlerweile auch in der vergleichenden Staatstätigkeitsforschung, allerdings hier mit Aggregatdaten. Man analysiert dabei für verschiedene Merkmalsträger über verschiedene Perioden hinweg die selben Variablen.

Schließlich gibt es noch komparativ-statische Vergleiche. Dabei werden zu zwei unterschiedlichen Zeitpunkten die Merkmalsausprägungen der Variablen verglichen. Beispielsweise wäre ein Vergleich der ökonomischen (oder politischen) Situation von 1968 mit der Situation von 1998 ein komparativ-statischer Vergleich, bei dem allerdings jedoch nicht der Übergang und die Veränderung zwischen diesen beiden Zeitpunkten näher untersucht wird. Solche komparativ-statische Vergleiche können auch auf die Querschnittsperspektive ausgedehnt werden.

Bei Stichprobenerhebungen stellt sich außerdem die Frage, ob von der Datenmatrix auf die interessierende Grundgesamtheit zurückgeschlossen werden kann. In diesem Forschungsabschnitt sollen die Daten mit Hilfe des statistischen Instrumentariums analysiert und ausgewertet werden.

Abbildung 3.1.: Die Datenmatrix

	Variable 1	Variable 2	.....	Variable k
Untersuchungseinheit 1				
Untersuchungseinheit 2				
:				
:				
Untersuchungseinheit n				

## (4) Die Interpretationsphase:

Die als Endprodukt angestrebten wissenschaftlichen Analysen entstehen in dieser Periode. Die Fülle der statistischen Daten und Ergebnisse sind zu interpretieren und miteinander in Beziehung zu setzen. Dabei sollten auch die Methoden kritisch reflektiert werden. Die Forschungsfragen und die aufgestellten Hypothesen sind zu beantworten, wobei eine deutliche und widerspruchsfreie Sprache anzustreben ist.

Bortz (1989: 3 ff.) unterscheidet insgesamt 6 Phasen der empirischen Forschung:

- (1) die Erkundungsphase,
- (2) die theoretische Phase,
- (3) die Planungsphase,
- (4) die Untersuchungsphase,
- (5) die Auswertungsphase,
- (6) die Entscheidungsphase.

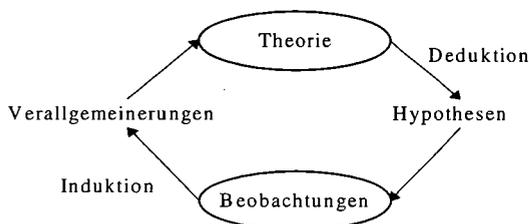
Dabei können die ersten drei Phasen analog zur ersten Phase bei Patzelt gesehen werden. Die Untersuchungsphase bei Bortz entspricht der Datenerhebungsphase, ebenso wie die Auswertungs- und Entscheidungsphase den korrespondierenden Phasen bei Patzelt. Die stärkere Gewichtung der ersten Forschungsphasen macht durchaus Sinn. Insbesondere die Betonung einer theorie- beziehungsweise hypothesengeleiteten Forschung ist für die Sozialwissenschaften zentral. Sie macht auch darauf aufmerksam, daß Hypothesen am Beginn der empirischen Untersuchung formuliert werden sollten und nicht erst am Ende. Will sagen, eine nachträgliche Anpassung der Hypothesen ist problematisch, da dann den Ergebnissen eine gewisse Willkür und Zufälligkeit nicht abgesprochen werden kann. Je fundierter und durchdachter die Vorarbeiten einer empirischen Forschungsarbeit sind, desto eher wird das Resultat positiv sein. Deswegen sollten Studierende oder Wissenschaftler diesen Komplex nicht auf die leichte Schulter nehmen. Eine wohlformulierte Fragestellung, die sowohl theoretische als auch prak-

tische Relevanz besitzt, und ein kluges Forschungskonzept mit einem detaillierten Zeitplan sind oftmals schon mehr als die halbe Miete für eine gelungene Seminar-, Abschluß- oder Forschungsarbeit.

Was ist eigentlich der Unterschied zwischen Theorie und Hypothese, nachdem diese Begriffe nun bereits mehrfach erwähnt wurden? Im umgangssprachlichen Gebrauch werden beide Begriffe oftmals synonym verwendet, obwohl sie sich wesentlich unterscheiden. Vom etymologischen Ursprung her bedeutet *theoria* im griechischen „Schauen“ und „Betrachten“. Nach Seiffert (1994: 368) lassen sich drei Verwendungen dieses Begriffes unterscheiden. 1) Theorie als Gegensatz zur Praxis. Hier ist der Theoriebegriff sehr weitgefaßt und bedeutet alles, was nicht „praktisches Tun“ ist. 2) Theorie als wissenschaftliches Lehrgebäude. In diesem Sinn ist Theorie vor allem ein Fachausdruck für eine wissenschaftstheoretische Grundposition, wie die bereits angesprochenen Metatheorien (z.B. die kritisch-dialektische Theorie oder die Systemtheorie). 3) Theorie ist im Sinne des logischen Empirismus zu verstehen. Dieses Wortungetüm meint die Verknüpfung von empirischen (d.h. von der Erfahrung bestimmten) und normativen (d.h. logischen, vom „Denken“ bestimmten) Aussagen. Ein Beispiel hierfür wäre in der Politikwissenschaft die Parteienherrschaftstheorie (Hibbs 1977), die davon ausgeht, daß die Staatstätigkeit eine Funktion der regierenden Parteien ist. Parteien unterschiedlicher politischer Couleur verursachen einen unterschiedlichen Output der Staatstätigkeit, z.B. bei der Bekämpfung der Arbeitslosigkeit oder der Staatsverschuldung.

Unter einer Hypothese (aus dem griechischen Wort *hypothesis* = Grundlage, Voraussetzung abgeleitet) versteht man „eine begründete (bereits geprüfte oder noch zu prüfende) Vermutung über die Beschaffenheit eines Sachverhalts oder die Art, Stärke und Richtung des Zusammenhangs zwischen zwei oder mehreren Variablen“ (Schmidt 1994: 168). Hypothesen haben im Forschungsprozeß einen vorläufigen Charakter. Sie sind stets Teilmenge einer Theorie. Mehrere logisch widerspruchsfreie Hypothesen, die in einem inhaltlichen Zusammenhang stehen, können eine Theorie bilden (Galtung 1970: 451). Mit Hilfe der Deduktion, d.h. der logischen Ableitung aus einer Theorie, werden Hypothesen formuliert. Eine Arbeitshypothese aus der bereits angesprochenen Parteienherrschaftstheorie wäre: „Linke Regierungen werden - unter sonst gleichen Umständen - für niedrigere Arbeitslosenquoten sorgen als bürgerliche Regierungen.“ Das Problem der Aufstellung und des Testens von Hypothesen ist zentral für das empirische Arbeiten. Ihm ist deshalb ein eigenes Kapitel (Kapitel 14) gewidmet, in dem näher auf weitere Besonderheiten eingegangen wird.

Abbildung 3.2.: Theoriebildung



Die Diskussion des Theoriebegriffs wirft weitere Fragen auf, insbesondere nach der Art der Theoriebildung und Theorieüberprüfung. Hier wird gemeinhin zwischen deduktiver und induktiver Methode unterschieden. Es handelt sich hierbei um einen entscheidenden Komplex jeder empirischen Forschung, mitunter sogar Streitpunkt zwischen konkurrierenden Schulen. Kennedy (1985) bezeichnet diesen Streit in seiner lesenswerten „Einladung zur Statistik“ als „die große Debatte“ in den Sozialwissenschaften. Wenn wir Theorien aus der Erfahrung, Beobachtung, Befragung und anderen empirischen Methoden bilden, gewinnen wir sie über die Induktion. Leiten wir Hypothesen mittels logischer und rationaler Überlegungen ab, so gehen wir deduktiv vor (siehe Abbildung 3.2.). Deduktion und Induktion sind zunächst Gegensatzpaare. Während die Deduktion die Ableitung des Besonderen aus dem Allgemeinen beinhaltet, ist die Induktion das Gegenteil, also die Ableitung des Allgemeinen (d.h. der Theorie) aus dem Besonderen.

Der in Abbildung 3.2. abgebildete Kreislauf der Wissenschaft weist allerdings auch auf eine Beziehung zwischen Deduktion und Induktion hin. Ausgehend von der Formulierung einer allgemeinen Theorie und ihrer logischen Überprüfung werden im Forschungsprozeß mittels Deduktion spezielle Hypothesen abgeleitet. Im Zuge der Forschung werden dann die interessierenden Variablen operationalisiert, und anschließend die relevanten Daten erhoben. Sofern die Daten für die Untersuchung nützlich sind, wird die Theorie dann anhand der „harten“ Daten der Wirklichkeit überprüft. Eine solche Forschungsstrategie wird auch als deduktiv-nomologischer Erklärungsansatz bezeichnet (von Alemann und Tönnemann 1995: 40ff., Schnell, Hill und Esser 1995: 55ff.). Folgendes Beispiel soll die Struktur einer deduktiv-nomologischen Erklärung veranschaulichen. In der modernen Demokratietheorie lautet eine Theorie, daß demokratisch verfaßte Staaten nicht untereinander Krieg führen (Schmidt 1997: 316, Rousseau et al. 1996: 512). Dies ist das allgemeine Gesetz, es wird auch als Explanans bezeichnet. Die zweite Komponente einer deduktiv-nomologischen Erklärung ist die An-

fangsbedingung (Antecedenz- oder auch Randbedingung). In unserem Beispiel ist die Randbedingung, daß Deutschland und Frankreich demokratische Staaten sind. Das Explanandum, dies kann synonym als spezielles Ergebnis oder als ein Sachverhalt zu einer bestimmten Zeit und zu einem bestimmten Ort bezeichnet werden, wird auf logischem Weg aus der Theorie und der zweiten Prämisse, d.h. der Randbedingung, abgeleitet. Das Ergebnis - nach logischer Ableitung - ist: Deutschland und Frankreich führen untereinander keinen Krieg. Das postulierte Gesetz - „Demokratien führen untereinander keinen Krieg“ - muß solange als richtig gelten, bis es falsifiziert wird. Was bedeutet Falsifikation? Im Wissenschaftsprogramm des Kritischen Rationalismus gibt es kein Wahrheitskriterium, d.h. Theorien sind nicht verifizierbar. Dagegen gilt, daß Theorien dann empirisch-wissenschaftlich sind, wenn sie falsifizierbar, d.h. empirisch widerlegbar sind. Dies bedeutet, daß man Gegenbeispiele finden muß, um die Theorie zu widerlegen. In unserem Beispiel würde das Ereignis, daß zwei demokratisch verfaßte Staaten untereinander Krieg führen, zu einer Falsifizierung der Theorie führen. Wie man leicht sieht, stellt dies eine recht rigide Forderung dar. Dies gilt umso mehr, wenn wir an einer strikt deterministischen Auslegung unseres Gesetzes festhalten, das heißt keine Abweichungen zulassen. In den Sozialwissenschaften ist man allerdings davon abgekommen, sich auf solche strikt deterministischen Aussagen zu stützen. Man formuliert die Theorien nicht mehr deterministisch, sondern probabilistisch, d.h. als wahrscheinlichkeitstheoretische oder statistische Aussage. In unserem Fall könnte die Theorie so lauten: „Die Wahrscheinlichkeit, daß demokratisch verfaßte Staaten untereinander Krieg führen, ist sehr gering.“ Neben den allgemeinen Problemen, wie man das Niveau der Demokratie in den einzelnen Staaten mißt sowie einen „Krieg“ operationalisiert, kommt es zu dem Problem der Abschätzung der „sehr geringen Wahrscheinlichkeit“, die man letztlich über eine Auszählung der Kriegstätigkeit approximiert.

Andererseits können wir auf dem Wege der Induktion auch Theorien bilden. Während die Theorie „Demokratien führen keine Kriege untereinander“, tatsächlich auf deduktivem Wege erstmals von Immanuel Kant in seiner Schrift „Vom ewigen Frieden“ postuliert wurde, wurde folgende Theorie auf induktiven Wege gewonnen. Sie lautet: „Demokratien haben dieselbe Wahrscheinlichkeit für kriegerische Auseinandersetzungen mit Nichtdemokratien, wie die Nichtdemokratien untereinander“. Diese „Systemnullhypothese“ bedeutet, daß Demokratien sich nicht signifikant in ihrem Konfliktverhalten von Nichtdemokratien unterscheiden, wenn ihre potentiellen Gegner Nichtdemokratien sind. Allerdings ist diese, lange Zeit als gesichert angesehene Auffassung inzwischen in die Kritik geraten (Rousseau et al. 1995). Die

Formulierung dieser Theorie beruht nicht auf Deduktion, sondern auf empirischen Beobachtungen, die dann zu einem Gesetz verallgemeinert wurden. Bei dieser Form der Induktion schließt man von einer Teilmenge (Summe der Staaten in einer Zeitperiode) auf die Gesamtmenge (Summe aller Staaten zu allen Zeiten). Die wissenschaftliche Leistungsfähigkeit der Induktion ist allerdings umstritten. Insbesondere Popper (1969) kritisierte diese Methode: Noch so viele Beobachtungen eines Sachverhaltes können nicht ausschließen, daß es Gegenbeispiele gibt, die die aufgestellte Theorie widerlegen können. Ziel ist deshalb die ständige Überprüfung, das heißt der Versuch der Falsifizierung der Theorie. Damit zusammenhängend richtet sich Poppers Kritik gegen eine vermeintlich mögliche Verifizierung von Theorien, was er der induktiven Vorgehensweise unterstellt. Dieses ist nicht möglich (Popper 1969: 15): „Ein empirisch-wissenschaftliches System muß an der Erfahrung scheitern können“.

Beispiel für einen eher problematischen Gebrauch der Induktion, wäre die These (oder „Theorie“) Günter Oggers, daß Deutschlands Manager „Nieten in Nadelstreifen“ (Ogger 1992) seien. Ogger präsentiert Einzelbeispiele von Fehlverhalten und generalisiert dann über die Gruppe der Unternehmer und Manager hinweg. Seine „theoretische“ Aussage - verkürzt zusammengefaßt - lautet: Die deutschen Unternehmer sind Versager, weil sie nur an ihrem kurzfristigen, privaten Interesse orientiert sind. Aus der Teilmenge erfolgloser, teilweise auch krimineller Unternehmer wird hier, um eine provokante These zu lancieren (in seinem Terminus „zu beweisen“ - was, wie wir bereits wissen, unmöglich ist), auf sämtliche Unternehmer geschlossen. Wie man sich leicht vorstellen kann, ist es einfach, diese These zu falsifizieren, denn in der Grundgesamtheit aller Unternehmer finden sich auch solche, die ihr Unternehmen sowohl zum Wohl der Anteilseigner als auch zum Wohl der Arbeitnehmer führen. Generell stellt diese Art der Theoriebildung einen Kardinalfehler in den Sozialwissenschaften dar. Von einer kleinen Stichprobe - im Extremfall einer Einerstichprobe - wird hier auf die Grundgesamtheit zurückgeschlossen. Die Erfahrungsgrundlage, die „empirische Basis“ also, ist eindeutig unzureichend.

Der Forschungsprozeß wird von Müller und Schmidt (1979), in dem dritten hier vorgestellten Phasenmodell, noch weiter als bei Patzelt und Bortz untergliedert. Sie stellen acht Phasen fest, die einer idealtypischen Untersuchung in der etablierten empirischen Politikforschung zu Grunde liegen (Müller und Schmidt 1979: 8f.):

- (1) Jede Untersuchung nimmt ihren Ausgangspunkt von einer wissenschaftlichen und/oder gesellschaftlich bedingten Problemstellung. Dieses Problem ist zu beschreiben und/oder zu erklären.

- (2) Auf der Basis von vorhandenen Erkenntnissen und Überlegungen werden Hypothesen über Strukturmerkmale und Bedingungsfaktoren aufgestellt.
- (3) Übersetzung der in den Hypothesen enthaltenen theoretischen Konzepte in beobachtbare und meßbare Variablen.
- (4) Die Anordnung der Variablen und Untersuchungseinheiten. Welche Art von Untersuchung soll durchgeführt werden? Welcher Untersuchungszeitraum soll analysiert werden?
- (5) Entscheidung über die Art und Weise der Informationsbeschaffung.
- (6) Wahl der Methoden der Datenanalyse und Festlegung von Kriterien, nach denen die Hypothesen als bestätigt oder nicht bestätigt angesehen werden.
- (7) Der tatsächliche Forschungsprozeß.
- (8) Erstellung eines Forschungsberichtes.

Über die angesprochenen Modelle hinaus lassen sich in der Fachliteratur weitere Phasenmodelle und Hinweise zum Ablauf eines Forschungsprozesses finden (Schnell, Hill, Esser 1995; Roth 1993; von Alemann und Tönnemann 1995). Es bleibt festzustellen, daß die Einteilungen in verschiedene Phasen durchaus nach den angewandten Untersuchungsmethoden variieren können, letztendlich sind alle diese Modelle Heuristiken, das heißt eine methodische Anleitung um Probleme zu lösen.

Ohne die Werturteilsdiskussion (Weber 1904) in den Sozialwissenschaften aufzurollen, kann man doch Kriterien angeben, wie eine „gute“ empirische Forschung aussehen sollte und welche Kriterien sie zu erfüllen hat. Diese Kriterien orientieren sich an dem Forschungsprogramm des Kritischen Rationalismus. Anhänger einer normativ oder einer kritisch-dialektisch orientierten Politikwissenschaft würden sicherlich einige Punkte dieses Programms kritisieren. Jedoch können diese Punkte - als Anhaltspunkte - eine Hilfe für eine gelungene empirische Arbeit darstellen.

- (1) Jede Forschung sollte theoretisch fundiert sein.
- (2) Der Untersuchungsgegenstand sollte entweder für die Wissenschaft und/oder für die Praxis relevant (das heißt bedeutend) sein.
- (3) Die Forschung soll nachprüfbar sein, der Leser sollte in der Lage sein - unter vertretbarem Arbeitsaufwand - die Ergebnisse zu überprüfen. Das heißt, wichtige Daten sollten publiziert werden und die Variablen sollten genau spezifiziert werden.
- (4) Die Forschung sollte empirisch sein, das heißt, man soll sich auf Phänomene beschränken, die beobachtet und mit Daten quantifiziert werden können.

- (5) Die Forschung sollte vergleichend sein.
- (6) Die politikwissenschaftliche Forschung sollte interdisziplinär sein.
- (7) Die theoretischen Konzepte sollten mit mehreren Indikatoren überprüft werden.
- (8) Die aufgestellten Hypothesen sollten, falls möglich, mit verschiedenen statistischen Methoden und Maßzahlen überprüft werden.

## 4. Grundbegriffe der Statistik

### 4.1 Merkmale, statistische Einheiten und Massen

Bei der Durchführung einer Datenerhebung muß vorher (a priori) festgelegt werden, welche Tatbestände erhoben werden und wer als Träger der Informationen in Frage kommt. Träger von Informationen oder interessierenden Eigenschaften heißen statistische Einheiten. Diese können zum Beispiel natürliche Personen (bei Geburten und Eheschließungen), Gegenstände (Gebäude bei der Gebäudezählung) oder auch einzelne Unternehmen (d.h. juristische Personen bei Gewinnermittlungen) sein. Politikwissenschaftliche Beispiele sind einzelne Parteien oder Regierungen.

Von der statistischen Einheit ist die statistische Masse zu unterscheiden. Die statistischen Einheiten bilden zusammen die statistische Masse. Die statistische Masse ist mithin die nach sachlichen, räumlichen und zeitlichen Kriterien gebildete Gesamtheit der statistischen Einheiten. Beispiel: Die Bevölkerung ist die statistische Masse bei einer Volkszählung.

Es ist wichtig, daß sich die Einheiten der statistischen Masse voneinander unterscheiden. Insbesondere müssen die statistischen Einheiten gegenseitig nach sachlichen, örtlichen und zeitlichen Identifikationsmerkmalen genau abgegrenzt sein. Bei einem Merkmalsträger werden meist verschiedene Merkmale erhoben (z.B. Alter, Geschlecht, Einkommen, parteipolitische Präferenz etc.). Als Merkmalsausprägungen bezeichnet man die verschiedenen Werte oder Kategorien, die ein Merkmal annehmen kann.

Jede statistische Einheit besitzt im allgemeinen eine Reihe von Eigenschaften. Es sind dabei vier Begriffe zu unterscheiden:

1. Merkmal
2. Merkmalsträger
3. Merkmalsausprägungen
4. Merkmalswert.

Eine Eigenschaft, die bei einer statistischen Untersuchung interessiert, heißt Merkmal (z.B. Alter, Einkommen, politische Einstellung). An dem Merkmalsträger (= statistische Einheit) wird das Merkmal erhoben. Die möglichen Realisationen werden Merkmalsausprägungen genannt. Ein Merkmal besitzt somit eine Menge von Merkmalsausprägungen. Als Merkmalswert wird die an einem Merkmalsträger ermittelte Merkmalsausprägung bezeichnet. Tabelle 4.1. soll dies für unterschiedliche Merkmale verdeutlichen. Während für die ersten drei

Beispiele die Merkmalsträger jeweils natürliche Personen sind, handelt es sich bei den beiden anderen Fällen um Institutionen (Haushalte beziehungsweise Regierungen). Das Merkmal Alter ist aber nicht nur an natürliche Personen gebunden, sondern es kann auch für Institutionen, zum Beispiel für Demokratien erhoben werden. Hier würde es die Dauer der demokratischen Herrschaft messen.

Tabelle 4.1.: Beispiele für Merkmale, Merkmalsausprägungen und Merkmalswerte

Merkmal	Merkmalsträger	Merkmalsausprägung (mögliche Realisationen)	Merkmalswert (Realisation am Merkmalsträger)
Alter	natürliche Person	0, 1, 2, ....., Jahre	28
Geschlecht	natürliche Person	männlich/weiblich	männlich
Politische Selbsteinstufung (auf einer Skala von 1 ... 10)	natürliche Person	1, 2, 3, 4, ....., 8, 9, 10 links.....rechts	5
Haushaltsgröße	Haushalt	1, 2, .... Personen	1
Parteilpolitische Zusammen- setzung <sup>1</sup> der Regierung	Regierung	1, 2, 3, 4, 5	1

Anmerkung: 1. Dieser Indikator beruht auf einer Berechnung der Kabinettsitzanteile der Parteien. Dabei wird den untersuchten Regierungen für jedes Jahr ein Wert zugewiesen, der aus fünf Klassifikationstypen abgeleitet wird (Schmidt 1982).

1 = Hegemonie bürgerlicher Parteien,

2 = Dominanz bürgerlicher Parteien,

3 = Patt zwischen bürgerlichen und Links-Parteien,

4 = Dominanz einer sozialdemokratischen Partei,

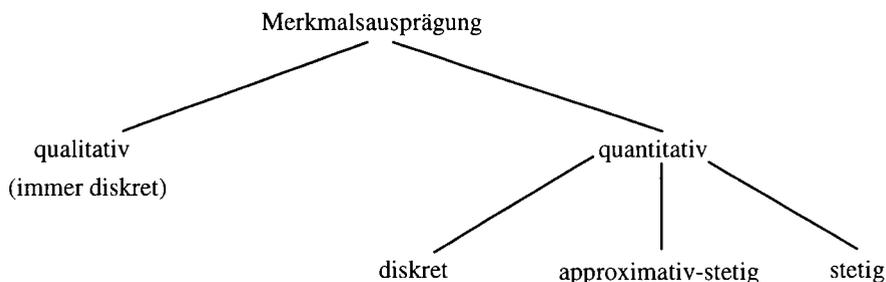
5 = Hegemonie einer sozialdemokratischen Partei.

Von Hegemonie spricht man, wenn der Kabinettsitzanteil 100 % beträgt, von Dominanz, wenn der Kabinettsitzanteil zwischen 66,6 % und 100 % liegt und von einem Patt, wenn die Kabinettsitzanteile der beiden Lager jeweils zwischen 33,3 % und 66,6 % liegen.

Wie wir gesehen haben, können die Merkmalsausprägungen unterschiedliche Formen annehmen: Zum einen als Zahlen und zum anderen als Wörter. Wir bezeichnen nun diejenigen Merkmalsausprägungen als quantitative Merkmale, die durch Zahlenwerte realisiert werden. Alle anderen Merkmale werden als qualitativ bezeichnet. Die quantitativen Merkmalsausprägungen lassen sich weiter danach unterteilen, ob sie diskret oder stetig sind. Mathematisch spricht man dann von einem diskreten Wert, wenn in einem geschlossenen Intervall endlich abzählbare Merkmalswerte vorliegen. Liegen in einem geschlossenen Intervall unendliche (potentielle) Merkmalsausprägungen vor, dann heißt dies stetig. Diskrete Variablen werden manchmal wie kontinuierliche (= stetige) Variablen behandelt (z.B. bei einer durchschnittlichen Haushaltsgröße von 2,43 Personen). Oft ist es möglich, verschiedene Merkmale sowohl diskret als auch approximativ-stetig zu erfassen. Beispiele sind das Alter, das normalerweise in Jahren angegeben wird, sich aber auch bis in Sekunden erfassen läßt sowie das Einkommen (Einkommensklassen versus Erfassung bis auf Pfennigbeträge). Eine solch tief-

gehende Erfassung wird als approximativ-stetig bezeichnet. Zur Veranschaulichung dient Abbildung 4.1.

Abbildung 4.1.: Typologie der Merkmalsausprägungen



In der Politikwissenschaft haben wir es überwiegend mit diskreten Merkmalen beziehungsweise Variablen zu tun. Üblicherweise wird der Begriff Variable zumeist bei quantitativen Merkmalen verwendet. Was bedeutet nun dieser Begriff „Variable“ überhaupt?

**Definition: Eine Variable ist ein veränderliches Merkmal, bei dem der Merkmalsträger mindestens zwei Merkmalsausprägungen annehmen kann.**

Im Gegensatz dazu ist eine Konstante eine unveränderliche Größe, ohne Variation des Merkmals. Die Variable Geschlecht kann beispielsweise zwei Merkmalsausprägungen annehmen, während die Variable „Einkommen“ dagegen eine potentiell große Anzahl an Merkmalsrealisationen aufweist. Zu erwähnen ist, daß selbst qualitative Merkmale für statistische Analysen verwendet werden können, dazu aber in quantitative Merkmale „umgeformt“ (d.h. umkodiert) werden müssen. Beim Merkmal Geschlecht können wir der Merkmalsausprägung männlich den Wert 1 und der Merkmalsausprägung weiblich den Wert 2 zuweisen. Weitere Beispiele für qualitative Merkmale sind Familienstand und Beruf. Eine Kodierung des Familienstandes (analog zu den Allbus-Umfragen) könnte wie folgt aussehen:

Verheiratet und lebt mit Ehepartner zusammen	= 1,
Verheiratet und getrennt lebend	= 2,
Verwitwet	= 3,
Geschieden	= 4,
Ledig	= 5.

Nach der Art der Messung unterscheidet man schließlich noch manifeste und latente Variablen. Manifeste Variablen sind direkt beobachtbar, während latente Variablen nur indirekt

gemessen werden können. Die latenten Merkmale sind in den Sozialwissenschaften von großer Bedeutung. Wenn man diese latenten Merkmale erfassen will, muß man sich daher überlegen, wie sie am sinnvollsten gemessen werden können. Oft versucht man von manifesten Meinungsäußerungen auf latente Einstellungen zurückzuschließen. Beispiel: Die politische Einstellung in der Bevölkerung soll gemessen und in verschiedene Kategorien klassifiziert werden, z.B. in links-extrem, gemäßigt-links, liberal, gemäßigt-rechts, rechts-extrem. Man kann nun die Einstellungen zu gewissen Themen abfragen wie: Ausländer, Gleichberechtigung, Arbeitslosigkeit, politische Aktivität sowie die Selbsteinstufung auf einer Links-Rechts-Skala (der Fragenkatalog läßt sich natürlich erweitern). Von den Ergebnissen der Umfragen versucht man dann auf die Einstellung in der Bevölkerung zu schließen.

Je nach Vollständigkeit der Daten spricht man von einer Grundgesamtheit oder einer Teilgesamtheit. Zur Grundgesamtheit gelangt man über eine Vollerhebung (Totalerhebung) aller statistischen Einheiten. Bei einer Teilgesamtheit spricht man analog von einer Teil- oder Stichprobenerhebung.

Wie bereits erörtert, müssen die statistischen Massen in sachlicher, räumlicher und zeitlicher (mitunter auch aus „institutioneller“) Sicht klar abgegrenzt sein. Bezüglich der zeitlichen Abgrenzung können die statistischen Massen entweder für einen bestimmten Zeitpunkt oder für einen Zeitraum definiert sein. Dementsprechend unterscheidet man Bestandsmassen (engl. stocks), die zeitpunktbezogen und Bewegungsmassen (= Ereignismassen, englisch = flows), die zeitraumbezogen sind.

Beispiele für Bestandsmassen sind:

1. Wohnbevölkerung zum 1. Januar eines Jahres.
2. Der Bestand aller Personenkraftwagen in Deutschland zum Ende eines Jahres.
3. Der Schuldenstand eines Landes am Ende eines Jahres.
4. Alle Bundestagsabgeordneten zu Beginn der Wahlperiode.

Beispiele für Bewegungsmassen sind:

1. Geburten eines Jahres.
2. Verkehrsunfälle innerhalb eines Jahres.
3. Das jährliche Haushaltsdefizit.
4. Die Abgeordneten, die neu in den Bundestag einziehen.

Über die Bestandsfortschreibung sind Bestands- und Bewegungsmassen miteinander verknüpft. Sie heißen dann korrespondierende Massen. Ein Beispiel für korrespondierende Massen sind die Bestandsmasse Bevölkerung und die Bewegungsmassen Geburten, Sterbe-

fälle sowie Zu- und Abwanderungen. Die Fortschreibung erfolgt über folgende Gleichung: Endbestand (B) = Anfangsbestand (AB) + Zugänge (Z) - Abgänge (A). Für das Beispiel der Staatsverschuldung lautet die Fortschreibungsformel: Schuldenstand (am 31.12.1998) = Schuldenstand (31.12.1997) + Bruttokreditaufnahme (in 1998) - Tilgungen (in 1998). Ein weiteres Beispiel ist die Lagerbestandsfortschreibung, in der - auf Basis von einem gegebenen Anfangsbestand - die jeweiligen Zu- und Abgänge erfaßt werden.

## 4.2. Meßniveaus und Skalierung

Die Unterscheidung in quantitative und qualitative Merkmale hängt eng mit der Frage nach dem Meßniveau und dem Skalentypus zusammen. Beide Ausdrücke werden synonym verwandt. Skalenniveaus sind in zweifacher Hinsicht von Bedeutung: Zum einen lassen sich die sozialwissenschaftlich relevanten Variablen sinnvoll klassifizieren, und zum anderen kann man aufgrund des vorliegenden Skalenniveaus entscheiden, welche statistischen Kennzahlen und Verfahren zur Deskription und Analyse zulässig sind. Dabei gilt: Je höher das Skalenniveau, desto mehr statistische Verfahren können angewandt werden.

Folgendes Beispiel soll auf die Problematik aufmerksam machen. Wir haben das qualitative Merkmal Religion erhoben mit den folgenden Ausprägungen:

1 = römisch-katholisch,

2 = evangelisch

3 = sonstige Religionszugehörigkeit.

Bei einer Umfrage erhalten wir das Resultat, daß 12 Personen katholisch und 8 Personen evangelisch sind. 12 Personen geben „sonstiges“ an. Es ist sinnlos, hier eine Mittelwertbildung mit Hilfe des arithmetischen Mittels<sup>1</sup> vorzunehmen - mit einem „Durchschnittswert“ von 2 - und von einer mittleren Konfession, die evangelisch sei, zu sprechen. Die Zuordnung der Zahlen zu den Merkmalsausprägungen kann hier, bei sogenannten nominalskalierten Daten, „frei“ erfolgen. Es spielt keine Rolle, ob „katholisch“ oder „evangelisch“ den Wert 1 zugewiesen bekommt. Eine Vertauschung würde gleichfalls das Ergebnis unserer Berechnung verändern und somit die Sinnlosigkeit dieses Tuns verdeutlichen. Dieses kleine Beispiel zeigt, wie wichtig es ist, sich zu vergegenwärtigen, welche mathematischen Operationen für welches

<sup>1</sup> Das arithmetische Mittel wird in Kapitel 7.2.3. ausführlich behandelt. Damit der Leser die (unzulässige) „Berechnung“ nachvollziehen kann: Dieser Mittelwert berechnet sich aus der Summe der Merkmalswerte der einzelnen Merkmalsträger, wobei diese dann durch die Zahl der Merkmalsträger dividiert wird.

Meßniveau zulässig sind.

Bevor eine statistische Untersuchung durchgeführt werden kann, muß das interessierende Phänomen zunächst gemessen werden. Die Zuordnung bestimmter Merkmalswerte zu den Merkmalsträgern heißt messen. Eine sehr ähnliche Definition ist diejenige von Friedrichs (1973). Danach besteht der Meßvorgang an sich aus drei Komponenten: dem empirischen Relativ, dem numerischen Relativ und der Abbildungsvorschrift. In unserem Religionsbeispiel ist das empirische Relativ die einzelne Religion der Merkmalsträger, die über die Abbildungsvorschrift („ordne Personen mit katholischer Religionszugehörigkeit den Wert 1 zu“) in das numerische Relativ überführt wird. Ein weiteres Beispiel für das Messen ist: Wir können einem Merkmalsträger den Merkmalswert „männlich“ oder „weiblich“ zuordnen, je nachdem welches Geschlecht er/sie aufweist.

**Definition: „Messen ist die systematische Zuordnung einer Menge von Zahlen oder Symbolen zu den Ausprägungen einer Variablen, mithin auch zu den Objekten“ (Friedrichs 1973: 97).**

Die Qualität der Messung hängt generell davon ab, wie gut die Eigenschaften durch Zahlen ausgedrückt werden können. Je nachdem, was gemessen wird, sind vier Meßniveaus (= Skalenniveaus) zu unterscheiden:

1. die Nominalskala,
2. die Ordinalskala,
3. die Intervallskala,
4. die Ratio- oder auch Verhältnisskala.

Nominal- und Ordinalskala bezeichnet man auch als topologische oder als nichtmetrische Skalen, Intervall- und Verhältnisskala dagegen als metrische Skalen.

Die Nominalskala (oder auch Kategorialeskala) stellt die einfachste Form des Messens dar. Sie liegt vor, wenn die Ausprägungen des untersuchten Merkmals durch die zugeordneten Zahlen lediglich unterschieden werden sollen. Die Zahlen erlauben keinerlei rechnerische Transformationen, sie dienen lediglich der Kodierung und Verschlüsselung der Daten sowie zur Unterscheidung der einzelnen Merkmalsträger, zum Beispiel das Merkmal Geschlecht mit männlich = 1 und weiblich = 2, oder das Merkmal Religion oder auch Parteilabels auf einem Wahlzettel. Substantiell spielt es keine Rolle, wie die Abbildungsvorschrift hier lautet. Wichtig ist lediglich die Unterscheidbarkeit der Merkmalsträger. Ein Beispiel aus politikwissenschaftlicher Perspektive ist die Unterscheidung von föderativen und nicht föderativen Ländern