

Manuskript zur Lehrveranstaltung Research Methods (Master; M1)

Teil III: Quellen wissenschaftlicher Erkenntnis und Grundlagen empirischer Forschung

Inhalt

- 1 Grundsätzliche Erkenntnisquellen wissenschaftlicher Forschung**
- 2 Empirische Sozialforschung**
 - 2.1 Allgemeine Grundlagen empirischer Forschung
 - 2.1.1 Ziele und Aufgaben der empirischen Sozialforschung
 - 2.1.2 Ablauf empirischer Untersuchungen
 - 2.1.3 Forschungsrichtungen
 - 2.1.4 Begriffe empirischer Forschung
 - 2.2 Angewandte qualitative Sozialforschung
 - 2.2.1 Gütekriterien qualitativer Forschung
 - 2.2.2 Durchführung qualitativer Studien
 - 2.2.3 Ausgewählte qualitative Erhebungsmethoden
 - 2.3 Angewandte quantitative Sozialforschung
 - 2.3.1 Hauptziele quantitativer Untersuchungen
 - 2.3.2 Gütekriterien quantitativer Forschung
 - 2.3.3 Ausgewählte quantitative Erhebungsmethoden
 - 2.3.3 Durchführung quantitativer Studien
- 3 Stichprobenplanung**
 - 3.1 Grundgesamtheit und Stichprobe
 - 3.2 Stichprobenverfahren
 - 3.3 Stichprobengröße
- 4 Statistische Verfahren der Datenauswertung und Hypothesentest**
 - 4.1. Datenauswertung
 - 4.1.1 Univariate Datenauswertung
 - 4.1.2 Bivariate Datenauswertung
 - 4.1.3 Multivariate Datenauswertung
 - 4.2 Hypothesentest
 - 4.2.1 Grundstruktur von Hypothesentests
 - 4.2.2 Chi-Quadrat-Unabhängigkeitstest

Bei dem vorliegenden Skript handelt es sich im Wesentlichen um Text-Auszüge aus:

- Berger, D.: Wissenschaftliches Arbeiten in den Wirtschafts- und Sozialwissenschaften. Hilfreiche Tipps und praktische Beispiele, Wiesbaden 2010
- Kornmeier, M.: Wissenschaftstheorie und wissenschaftliches Arbeiten. Eine Einführung für Wirtschaftswissenschaftler, Heidelberg 2007
- Töpfer, A.: Erfolgreich Forschen. Ein Leitfaden für Bachelor-, Master-Studierende und Doktoranden, 2. Auflage, Heidelberg u.a. 2009/10.
- Einschlägige Werke aus der Statistik

1 Grundsätzliche Erkenntnisquellen wissenschaftlicher Forschung

Wissenschaftliche Arbeiten basieren auf vorhandenem bzw. auf neuem Wissen. Hierbei kann man auf verschiedene Erkenntnisquellen zurückgreifen. Je nach

- Art der **Informationsbeschaffung** (Sekundär- vs. Primärforschung) und
- **Herkunft** der Informationen

ergeben sich vier Möglichkeiten (Abb. 1)

Informationsbeschaffung	Herkunft der Informationen	
	Literatur	Empirie
Sekundärforschung	Literaturstudium (I)	Schreibtischforschung (III) (,Desk research')
Primärforschung	Meta-Analyse (II)	Feldforschung (IV) (,Field research')

Abb. 1: Quellen zur systematischen Gewinnung von Wissen (Kornmeier 2007)

(I) Wer auf bestehendes Wissen zurückgreift, muss dieses ggf. in Frage stellen. Ein Autor hat deshalb zunächst mit Hilfe eines **Literaturstudiums** das „Fundament“ zu schaffen, indem er den Stand der Forschung aufbereitet. Ausgehend von diesem „State of the art“ legt er dann seine eigenen Gedanken und Erkenntnisse dar. Die Qualität und Quantität der so dokumentierten Verarbeitung der relevanten Literatur erlauben dem Leser (und den Gutachtern!) Rückschlüsse, in welchem Maße und in welcher Güte sich ein Autor mit dem Thema auseinandergesetzt hat. „Wissenschaftlich“ Arbeiten heißt in diesem Fall insbesondere (s.a. Teil IV),

- die wichtigsten Literaturquellen zu erschließen,
- diese zusammenfassend bzw. in den für die Thematik relevanten Ausschnitten (d.h. nicht nur nacherzählend) wiederzugeben und
- den Stand der Diskussion in diesem Feld (methoden-)kritisch zu würdigen.

(II) Neben dem klassischen Literaturstudium bietet – als spezieller Forschungsansatz – die so genannte **Meta-Analyse** eine Möglichkeit, Wissen zu bündeln und Erkenntnisse zu generieren. Genau genommen ist dieses Verfahren ein Hybrid aus Primär- und Sekundärforschung:

- Einerseits werden **vorhandene** Daten ausgewertet, was dafür spricht, die Meta-Analyse als **sekundärstatistische** Methode zu bezeichnen;
- Andererseits aber ähnelt der **Ablauf** der Meta-Analyse der Vorgehensweise wie sie für **primärstatistische** Erhebungen üblich ist (Abb. 2 und unten Kap. 2.1.2).

Bei der **Meta-Analyse** ist der Forscher nicht unmittelbar in die empirische Forschung (1. Ordnung) einbezogen, sondern er hat den Status eines **nachträglichen** Analytikers und Systematikers. Meta-Analysen sind dadurch wissenschaftliche Untersuchungen 2.Ordnung, bei denen eine Datenanalyse auf **sekundärstatistischem** Niveau durchgeführt wird. Das Ziel der Meta-Analyse ist, die ermittelten Ergebnisse von verschiedenen vorliegenden wissenschaftlichen Studien zum selben Sachverhalt noch einmal **gemeinsam** zu analysieren und statistisch nach einzelnen Inhalten und Ergebnissen **auszuwerten**. Dieses Vorgehen ist bei der Aufarbeitung des bisherigen Forschungsstandes angebracht und üblich, weil auf diese Weise die **erreichten** Ergebnisse klassifiziert und zusammenfassend quantifiziert werden können. Das Resultat ist dann eine evaluierte Forschungs- und Literaturübersicht.

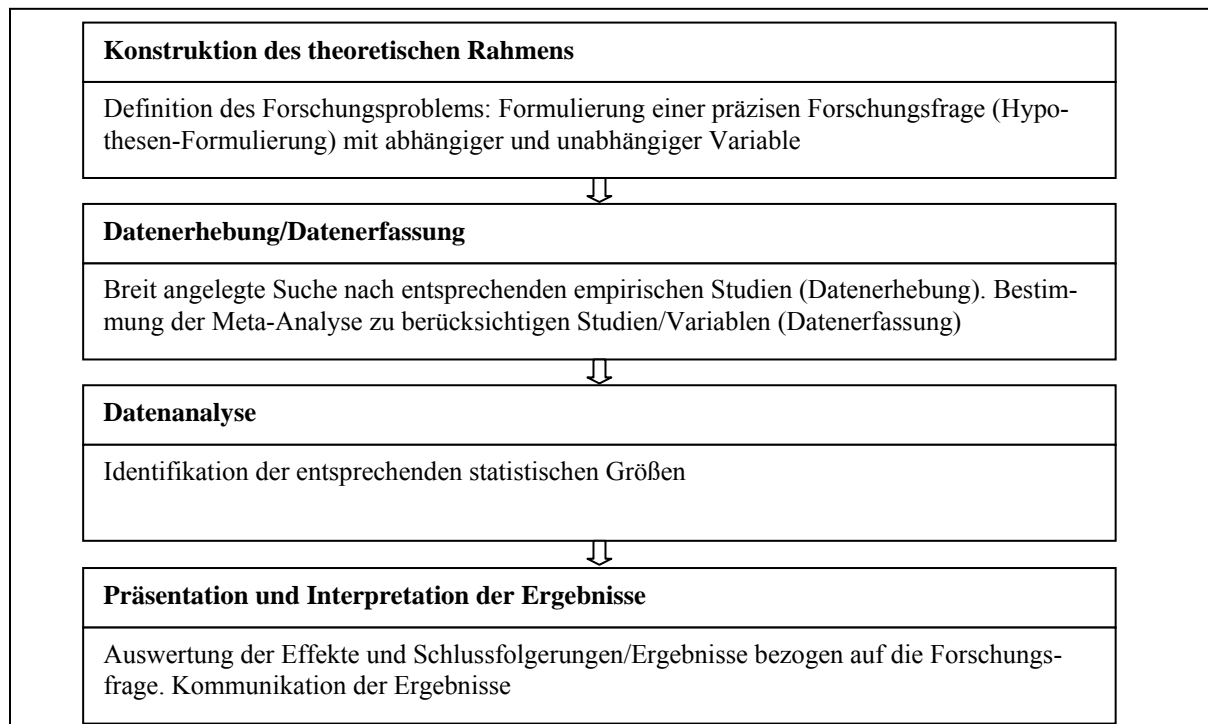


Abb. 2: Ablaufschritte der Meta-Analyse (Kornmeier 2007, Töpfer 2009/10)

(III)/(IV) Neue Erkenntnisse lassen sich auch durch empirische Forschung gewinnen.

- Dabei kann es zweckmäßig sein, **Sekundärforschung** zu betreiben, indem man bereits erhobene Daten re-analysiert bzw. vorliegendes Material (z.B. Informationen des Statistischen Bundesamtes, Geschäftsberichte) für die eigene Fragestellung **aufbereitet** und **auswertet**. Wer sich also für die Sekundäranalyse (Schreibtischforschung) entscheidet, greift auf Datenbestände zurück. Wichtige Quellen der Schreibtischforschung sind u.a. öffentliche Einrichtungen, Ministerien, Behörden, Industrie- und Handwerkskammern, Wirtschaftsverbände, Marktforschungsinstitute, Kreditinstitute, Wirtschaftsdatenbanken, DATEV und Geschäftsberichte (Internet, Messen, Fach- und Branchenzeitschriften).
- Falls die verfügbaren sekundärstatistischen Daten für ein bestimmtes Forschungsvorhaben ungeeignet sind bzw. nicht ausreichen, sind ggf. Daten zu erheben, d.h. es wird **Primärforschung** (Feldforschung) erforderlich. Hierzu bieten sich grundsätzlich verschiedene Möglichkeiten (Erhebungsmethoden) an, von denen die Befragung (z.B. Versorgungszufriedenheit der Verbraucher) die am weitesten verbreitete ist. In Betracht kommen auch Beobachtung und Experiment. Die Wahl der Erhebungsmethode richtet sich nach der Problemstellung, aber beispielsweise auch nach den jeweiligen Vorkenntnissen sowie nach den verfügbaren Ressourcen.

In wissenschaftlichen Arbeiten stützt man sich i.d.R. nicht nur auf eine der o.g. Quellen, sondern man verknüpft theoretische mit empirischen Erkenntnissen. Dabei ist allerdings – unabhängig von der jeweils genutzten Erkenntnisquelle – zu beachten, dass Dritte die Ergebnisse prüfen und bewerten können. Hierzu müssen die einzelnen Schritte einer wissenschaftlichen Arbeit **systematisch** und überdies **so gut dokumentiert** sein, dass jeder Sachverständige die Argumentationslinien sowie die theoretischen bzw. empirischen Ergebnisse **nachvollziehen** kann. In den folgenden Abschnitten steht die empirische (Sozial-)Forschung im Sinne der Primär- bzw. Feldforschung im Vordergrund.

2 Empirische Sozialforschung

Empirische Sozialforschung (empirisch = auf Erfahrung beruhend) bezeichnet die **systematische** Erhebung von Daten über **soziale Tatsachen** durch den Einsatz geeigneter **Methoden**, wie bspw. einer Befragung oder einer Beobachtung, und deren Auswertung (s. Folie).

2.1 Allgemeine Grundlagen empirischer Forschung

In den Wirtschafts- und Sozialwissenschaften existieren eine größere Zahl von Methoden zur Erhebung und Auswertung von Daten. Die Gesamtheit dieser Methoden stellt das **Inventar** der empirischen Sozialforschung dar. Welche Methode sich am besten eignet, hängt vom Untersuchungsziel und der zugrunde liegenden Fragestellung ab. Häufig werden auch verschiedene Methoden **kombiniert**. Erfahrungen und Beobachtungen werden auch im Alltag gemacht. Allerdings bestehen zwischen empirischer Sozialforschung und Alltagswissen erhebliche Unterschiede (Abb. 3).

Empirische Sozialforschung	Alltagswissen
Systematisches Vorgehen	Selektive Beobachtung und Erinnerung
Methodisch begründetes Vorgehen	Ignorieren widersprechender Informationen
Präzise Begrifflichkeit	Implizites Begriffsverständnis
Suche nach Wahrheit	Vorurteile und stereotypes Denken
Intersubjektiv überprüfbares Vorgehen	Aberglaube und Magie

Abb. 3: Wesentliche Unterschiede empirische Sozialforschung und Alltagswissen (Berger 2010)

2.1.1 Ziele und Aufgaben der empirischen Sozialforschung

Das oberste Ziel der empirischen Sozialforschung ist das Gewinnen von Erkenntnissen über Phänomene der sozialen Welt. Die Aufgaben der empirischen Sozialforschung können im Wesentlichen in drei Typen von Untersuchungen zusammengefasst werden:

1. Deskriptive Untersuchungen (Deskription = Beschreibung) richten sich nicht an die Erforschung von Zusammenhänge, sondern auf die genaue Beschreibung des Realitätsausschnittes (z.B. Häufigkeitsverteilungen oder Typen). Hauptziele sind Beschreibungen, Erstellung von Klassifizierungen oder Typologien und der Bericht beobachteter Daten.

2. Explorative Untersuchungen (Exploration = Erkundung) werden durchgeführt, wenn weitgehend unbekannte Phänomene oder Zusammenhänge erforscht werden sollen. Häufig handelt es sich um Voruntersuchungen bzw. Pretests, die einer Hauptstudie vorgeschaltet werden. Es geht in erster Linie darum, **Hypothesen** zu **generieren**, welche in einer Hauptstudie **überprüft** werden. Vorzugsweise kommen qualitative Methoden (z.B.. Fokusgruppen, Tiefeninterviews, unstrukturierte Beobachtung s.u.) zum Einsatz. Die methodische Vorgehensweise ist induktiv oder hermeneutisch.

3. Hypothesenprüfende Untersuchungen beschäftigen sich mit der Prüfung von vermuteten Merkmals- oder Variablenzusammenhängen. Die Hypothesen werden i.d.R. deduktiv, also aus einer allgemeinen Theorie, abgeleitet.

Aus diesen drei Typen von Untersuchungen lassen sich folgende wesentliche Ziele ableiten:

- **Beschreiben:** Es wird versucht, Phänomene und Zusammenhänge zwischen ihnen möglichst präzise zu beschreiben. Diese Beschreibungen dienen als wichtige Basis für wissenschaftliche Theorie- und Hypothesenentwicklung.
- **Erklären:** Kausale Beziehungen zwischen Ereignissen der sozialen Welt sollen erklärt und auf Ursachen zurückgeführt werden.
- **Verstehen:** Man versucht, Ereignisse und Handlungen zu verstehen, indem Sinneszusammenhänge, in die sie eingebettet sind, erfasst werden (Teil I: Hermeneutik). Dies wird in erster Linie mittels explorativer Studien gemacht.

Darüber hinaus sind das **Bewerten**, die Ermittlung der Wirksamkeit oder Unwirksamkeit von Maßnahmen bzgl. einer oder mehrerer Erfolgskriterien, sowie das **Verändern** („praktisches Gestalten“) weitere Zielsetzungen empirischer Forschung. Für die praktische Umsetzung sind Beschreiben, Erklären, Verstehen und Bewerten erforderliche Voraussetzungen.

Der Großteil der empirischen Bachelor- oder Masterarbeiten verfolgt mit ihren Untersuchungen entweder das Ziel, ein Phänomen zu beschreiben (deskriptive Untersuchung) oder kausale Beziehungen zu erklären und auf Ursachen zurückzuführen (hypothesenüberprüfende Untersuchung).

2.1.2 Ablauf empirischer Untersuchungen

Der Ablauf empirischer Studien erfolgt idealerweise in **fünf** aufeinander folgenden **Schritten**:

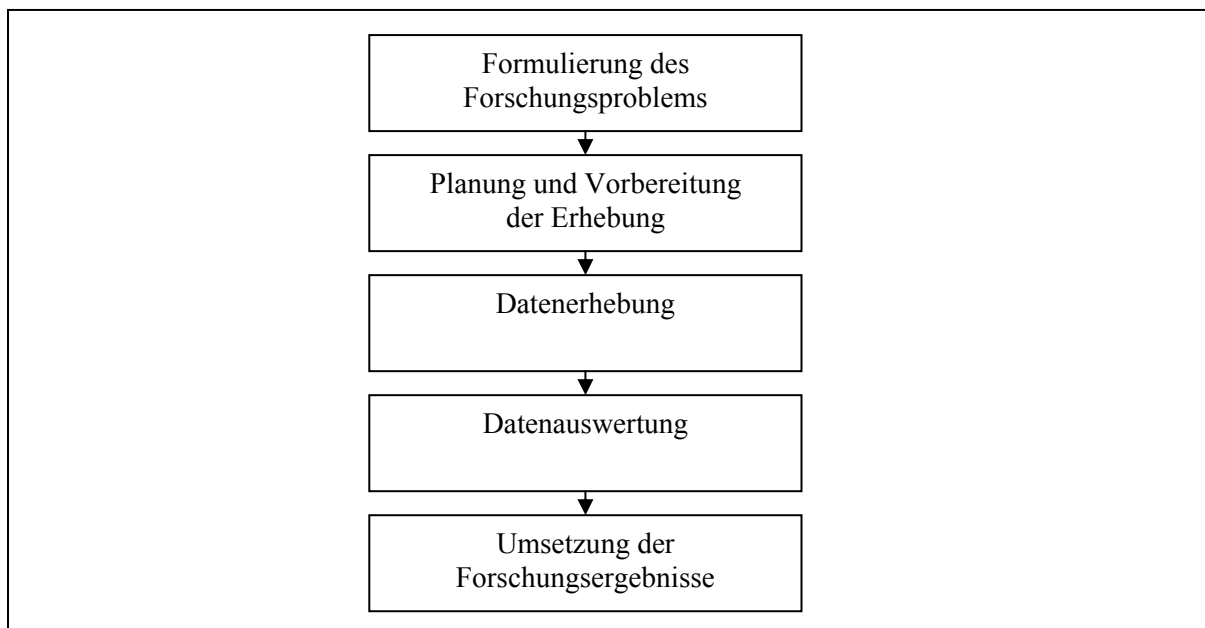


Abb. 4: Ablauf empirischer Untersuchungen

1. **Formulierung des Forschungsproblems:** Diese hängt davon ab, welches Ziel mit dem empirischen Projekt verfolgt wird. Frage: „Was genau möchte ich wissen“? Das Forschungsziel ist eng verbunden mit dem Untersuchungstyp. Handelt es sich bspw. um eine hypothesenprüfende Untersuchung, dann besteht das Forschungsproblem aus der Prüfung von vermuteten Merkmals- oder Variablenzusammenhängen. Das Forschungsproblem kann aber auch von Auftraggeber vorgegeben sein, aus einem praktischen Problem resultieren oder auf wissenschaftliche Literatur zurückzuführen sein, indem Hypothesen generiert wurden, die es zu überprüfen gilt (s. Folie).

2. **Planung und Vorbereitung der Erhebung:** Dieser Schritt kann in vier wesentliche Aufgaben unterteilt werden:
 - (a) **Konstruktion des Erhebungsinstrumentes:** Diese Aufgabe beinhaltet die Definition der Begriffe aus den Forschungshypothesen. Hierzu gehört auch die Konzeptspezifikation für komplexe Begriffe, also die **Zerlegung** der Begriffe in Einzeldimensionen, und die **Messung** der Einzeldimensionen (z.B. Lebensstil hat u.a. die Dimensionen Aktivitäten, Interessen, Meinungen/Einstellungen) und das **Operationalisieren** – d.h., die möglichst genaue Angabe der Vorgehensweise, mit der ein Merkmal erhoben werden soll. Bei einer Befragung versteht man unter Operationalisierung bspw. die genaue Frageformulierung mitsamt den Antwortvorgaben.
 - (b) **Festlegung der Untersuchungsform:** Eine wichtige Entscheidung bei dieser Aufgabe ist es, den zeitlichen Aspekt der Datenerhebung festzulegen. Es kann eine **Querschnittserhebung** (einmalige punktuelle Erhebung) oder eine **Längsschnitterhebung** (=Erhebung zu mehreren Zeitpunkten) gewählt werden. Beim Festlegen der Untersuchungsform ist auch zu überlegen, ob eine Vergleichs- oder Kontrollgruppe herangezogen werden soll. In diesem Fall spricht man von einem Experiment.
 - (c) **Stichprobenverfahren:** Dieser Schritt bezieht sich auf die Bestimmung von **Art und Größe** der Stichprobe. Zunächst ist die Zielpopulation zu definieren; d.h. die Festlegung der Grundgesamtheit (Menge aller potenziellen Untersuchungsobjekte), aus welcher dann die Stichprobe gezogen werden soll. Für welche Art der Stichprobenziehung (z.B. einfache Zufallsauswahl, Quotenauswahl, willkürliche Stichprobe) man sich entscheidet und wie groß diese Stichprobe sein soll, hängt vom Forschungsziel ab.
 - (d) **Pretest:** Jedes neu konstruierte Erhebungsinstrument sollte einem sog. Pretest unterzogen werden. Das bedeutet, dass vorerst nur an einer geringen Personenanzahl aus der Grundgesamtheit getestet wird, um **Informationen** darüber zu erhalten, ob bspw. noch etwas an dem Erhebungsinstrument geändert werden muss. Zum Beispiel könnten bei einem Fragebogen manche Fragen zu komplex sein oder Antwortkategorien fehlen. Wichtig ist, dass der Pretest dieselben Untersuchungskonditionen hat wie die eigentliche Studie, sodass es zu keinen Verzerrungen im Ergebnis kommt.
3. **Datenerhebung:** Die Art der Datenerhebungsmethode wird bereits bei der Konstruktion des Erhebungsinstrumentes bestimmt. Es gilt zu beachten, dass die Erhebungsmethode auch das Messergebnis beeinflussen kann. Wichtige und häufig verwendete Arten sind: Befragung, Beobachtung und Inhaltsanalyse.
4. **Datenauswertung:** Die Datenauswertung besteht aus zwei Aufgabenbereichen, aus dem Aufbau eines **analysefähigen Datenfiles** und aus der **statistischen Datenanalyse**.
 - (a) Erfassung der Daten in einer analysefähigen Form (z.B. mittels Datenfile)
 - (b) Welches Verfahren der statistischen Datenanalyse verwendet wird (univariate, bivariate, multivariate Verfahren, Kap. 4), hängt von der Komplexität der Fragestellung, der Art der Daten und vom statistischen Know how ab.
5. **Umsetzung der Forschungsergebnisse:** Die Umsetzung der Forschungsergebnisse ist ein wesentlicher Aspekt, der die **Anschlussfähigkeit** der empirischen Studie gewährleistet. Die empirischen Erkenntnisse werden in einem **Ergebnisbericht** bzw. im empirischen Teil der wissenschaftlichen Arbeit festgehalten. Auch die methodischen Entscheidungen sollen darin nachvollziehbar dokumentiert werden. Das Erhebungsinstrument sowie

etwaige Datenfiles, meist in anonymisierter Form, werden zwecks Transparenz als Anhang beigelegt. Dieser detaillierte Ergebnisbericht ermöglicht eine kritische Diskussion der Ergebnisse und lässt die Möglichkeit zur Replikation offen.

2.1.3 Forschungsrichtungen

In der empirischen Sozialforschung stehen grundsätzlich zwei verschiedene **Forschungsrichtungen** zur Verfügung, um zu Erkenntnissen zu kommen: die **quantitative** und die **qualitative** Methoden (Froschauer/Lueger 2008, Fink 2008, Atteslander 2008, Kessler 2008a/b). Oft ist auch eine Kombination aus beiden Richtungen sinnvoll, um das Forschungsproblem tiefer zu durchleuchten. Qualitative Methoden werden auch oft als Input für quantitative Methoden verwendet.

(1) Quantitative Sozialforschung

Bei der quantitativen Forschung geht es darum, Verhalten in Form von **Modellen**, Zusammenhängen und zahlenmäßigen Ausprägungen möglichst genau zu beschreiben, zu **erklären** und **vorhersagbar** zu machen. Dabei werden bspw. aus einer Befragung einer möglichst repräsentativen Stichprobe mit Hilfe eines Fragebogens die zahlenmäßigen Ausprägungen bestimmter Merkmale gemessen. Diese Messwerte werden miteinander oder mit anderen Variablen in Beziehung gesetzt und die Ergebnisse auf die Population bezogen (deduktive Vorgehensweise). Um die gleichen Voraussetzungen für die Entstehung der Messwerte innerhalb einer Studie zu gewährleisten, sind die quantitativen Methoden meist **standardisiert** und **strukturiert**, d.h. jeder Befragte bekommt möglichst exakt die gleichen Voraussetzungen bei der Beantwortung der Fragen (gleicher Wortlaut, gleiche Reihenfolge, gleiche Bewertungsskala etc.), um die Aussagen der Befragten untereinander vergleichbar zu machen. Quantitative Methoden eignen sich durch die standardisierte Vorgehensweise und große Stichproben zur **objektiven Messung** von Sachverhalten, zum **Testen** von Hypothesen und zur **Überprüfung statistischer Zusammenhänge** (Metaanalyse), aber auch für **wiederkehrende** Fragestellungen, bei denen Ergebnisse verschiedener Zeitpunkte verglichen werden sollen. Die wichtigsten Methoden sind die Befragung, die Beobachtung, das Panel und das Experiment.

(2) Qualitative Sozialforschung

Die qualitative Forschung ist im Vergleich offener und flexibler. Es geht in erster Linie um das Beschreiben, **Interpretieren** und **Verstehen** von Zusammenhängen, die Aufstellung von Klassifikationen oder Typologien und die **Generierung** von Hypothesen. Die qualitative Befragung zeichnet sich durch eine nicht vorherbestimmte, nicht standardisierte und umfassende Informationen liefernde Herangehensweise aus. Es liegt zwar meist ein grober thematischer Leitfaden zugrunde, die Reihenfolge und Gestaltung der Fragen sind jedoch flexibel und die Antwortmöglichkeiten der Gesprächspartner unbeschränkt. Durch diese Vorgehensweise soll ein tieferer Informationsgehalt der Ergebnisse (Sinnzusammenhang) erreicht werden, ohne allerdings repräsentative und zahlenmäßige Aussagen machen zu können (induktive Vorgehensweise; Hermeneutik). Sie eignet sich deshalb gut, um ausführliche Beschreibungen individueller Meinungen und Eindrücke zu erhalten, wie z.B. für Verbesserungsvorschläge oder zur Erkundung von Ursachen. Die Stichprobenbildung erfolgt nicht, wie bei quantitativen Verfahren, nach Kriterien statistischer Repräsentativität, sondern danach, ob sie geeignet ist, das Wissen über den Untersuchungsgegenstand zu erweitern oder nicht. Daher ist die erforderliche Stichprobengröße im Allgemeinen deutlich geringer als bei quantitativen Verfahren. Ziel der qualitativen Forschung ist es, die Wirklichkeit anhand der subjektiven Sicht der Untersuchungspersonen abzubilden und so mögliche Ursachen für deren Verhalten

nachzuvollziehen und das Verhalten zu verstehen. Wichtige Verfahren sind u.a. das narrative Interview, die Fokusgruppe, die qualitative Inhaltsanalyse und die Beobachtungsanalyse.

Folgende Abbildungen stellen Vor- und Nachteile der qualitativen Methoden (Abb. 5) bzw. der quantitativen Methoden (Abb. 6) gegenüber (s.a. die vergleichende Charakterisierung bei Hienerth u.a. 2008).

Vorteile qualitativer Methoden	Nachteile qualitativer Methoden
Neue, unbekannte Sachverhalte können entdeckt werden	Hoher Zeit- bzw. Kostenaufwand, relativ aufwendige Auswertung
Tieferer Informationsgehalt durch offene Vorgehensweise	Anforderungen an die Qualifikation der durchführenden Person ist relativ hoch
Methode passt sich an den Untersuchungsgegenstand an, flexiblere Vorgehensweise	Größere Subjektivität der Ergebnisse, kaum Generalisierbarkeit
„Wahre“, vollständige Informationen über die subjektive Sicht der Untersuchungspersonen	Qualität der Daten ist von der durchführenden Person abhängig
Möglichkeit, während der Untersuchung nachzufragen	Aus qualitativen Daten können keine zahlenmäßigen Mengenangaben abgeleitet werden.

Abb. 5: Vor- und Nachteile qualitativer Methoden (Berger 2010)

Vorteile quantitativer Methoden	Nachteile quantitativer Methoden
Große Stichproben und repräsentative Ergebnisse	Keine Flexibilität während der Untersuchung
Exakt quantifizierbare Ergebnisse	Kein individuelles Eingehen auf Probanden möglich
Objektivität und Vergleichbarkeit der Ergebnisse	Fragen können vom Probanden falsch interpretiert werden
Ermittlung von statistischen Zusammenhängen möglich	Wird nicht an die subjektive Erfahrungswelt der Probanden angepasst
Meist geringerer Zeitaufwand	Nicht geeignet für Ursachenforschung
Auswertung aufgrund standardisierter Vorgehensweise einfacher	Hohes Vorwissen zur Erstellung des Erhebungsinstrumentes notwendig

Abb. 6: Vor- und Nachteile quantitativer Methoden (Berger 2010)

Zusammengefasst kann man festhalten, dass qualitative Sozialforschung verwendet wird bzw. werden sollen, wenn

- Verstehen forciert und die Wissensbasis erweitert,
- Tatsächliche Umstände aufgeklärt,
- Hypothesen generiert,
- Verhaltensbandbreiten identifiziert,
- Motive, Einstellungen etc. entdeckt, erstmals verstanden,
- Und für zukünftige quantitative Forschung Input geliefert werden soll!

Ansonsten werden (besser) quantitative Methoden verwendet!

2.1.4 Begriffe empirischer Forschung

In diesem Abschnitt werden einige der statistischen Begriffe noch einmal aufgegriffen und detaillierter erläutert, die in den vorhergehenden Abschnitt schon genannt wurden (zu Pkt b. und c. siehe ausführlich Teil II).

(a) Variablen und Daten

Eine Variable ist ein Merkmal, d.h. eine Eigenschaft, die in einer Erhebung untersucht werden soll. In einer Befragung bspw. entspricht ein Merkmal einer gestellten Frage, wie etwa das

Geschlecht, das Alter oder das Ausbildungsniveau. Der Wert einer Variablen ist die Merkmalsausprägung (z.B. weiblich, männlich). Bei Befragungen sind die Merkmalsausprägungen die Antwortmöglichkeiten, die der Befragte angeben kann.

Zusätzlich wird zwischen qualitativen und quantitativen Merkmalen unterschieden. Als **qualitative** Merkmale bezeichnet man Merkmale, bei denen sich die Merkmalsausprägungen eindeutig in Kategorien unterscheiden lassen, die jedoch **keinen mathematischen** Wert annehmen können, wie z.B. das Geschlecht, Meinungen oder Einstellungen. Man spricht auch von **nominalen** oder **ordinalen** Mess- bzw. **Skalenniveau** (siehe c.). Als **quantitative** Merkmale bezeichnet man Merkmale, deren Merkmalsausprägungen **intervallskalierte** Werte annehmen. Typische Beispiele sind Einkommen, Haushaltsausgaben oder der Unternehmenswert. Für diese Merkmale können verschiedene mathematische Rechenoperationen durchgeführt werden, wie die Berechnung des Mittelwertes. Bei den quantitativen Merkmalen unterscheidet man Intervallskalen (ohne natürlichen Nullpunkt) und Ratio- bzw. Verhältnisskalen (mit natürlichem Nullpunkt).

Variablen haben im Zusammenhang mit empirischen Studien verschiedene funktionale Bedeutungen. Es wird unterschieden zwischen **abhängigen** und **unabhängigen** Variablen. Veränderungen der abhängigen Variablen sollen mit dem Einfluss der unabhängigen Variablen erklärt werden, wie z.B. „Wenn jemand ein starker Raucher (unabhängige Variable) ist, dann hat diese Person eine kürzere Lebenserwartung (abhängige Variable)“.

Daten sind Messwerte, die im Rahmen von empirischen Untersuchungen, wie Befragungen, Beobachtungen, Experimente etc. erhoben werden. Daten können entweder **numerisch** oder **kategorial** sein. Numerische Daten besitzen eine mathematische Bedeutung als Messwert, wie Gewicht oder Einkommen. Kategoriale Daten bezeichnen dagegen Merkmale wie Geschlecht, Beruf oder Familienstand. Die Gesamtheit aller numerischen und kategorialen Daten bezeichnet man als **Datensatz** einer Erhebung.

(b) Hypothesen (vgl. Teil II)

Hypothesen sind Aussagen oder Vermutungen über einen bestehenden Sachverhalt, meist über einen Merkmalszusammenhang zwischen mindesten zwei Merkmalen. Hypothesen gehen über den aktuellen Wissensstand hinaus, d.h., mit einer Hypothese wird eine neue Vermutung aufgestellt. Hypothesen sind eindeutig und präzise zu formulieren, damit ihre Aussage empirisch bestätigt (verifiziert) oder widerlegt (falsifiziert) werden kann.

Man unterscheidet als Gegensatzpaar die **Nullhypothese** und die **Alternativhypothese**. Die Nullhypothese sagt aus, dass ein bestimmter Zusammenhang nicht besteht. Diese Annahme soll mittels Hypothesentest verworfen werden, so dass die Alternativhypothese als wahrscheinlich übrig bleibt.

Beispiel

Nullhypothese: Wenn jemand starker Raucher (unabhängige Variable) ist, dann verkürzt er dadurch nicht seine Lebenserwartung (abhängige Variable). D.h. es gibt keinen statistisch signifikanten Zusammenhang zwischen Rauchverhalten und Lebenserwartung.

Alternativhypothese: Wenn jemand starker Raucher (unabhängige Variable) ist, dann verkürzt er dadurch seine Lebenserwartung (abhängige Variable). D.h. es gibt einen statistisch signifikanten Zusammenhang zwischen Rauchverhalten und Lebenserwartung.

Ziel eines statistischen Test ist die Ablehnung (Verwerfung, Falsifikation) der Nullhypothese. Falls diese nicht verworfen werden kann, weil bspw. die Stichprobengröße zu klein ist, besteht aus statistischer Sicht allerdings noch kein Grund, von der Gültigkeit der Nullhypothese auszugehen. Ein statistischer Test kann nur zu einer Annahme der Alternativhypothese, nicht aber zu eine Annahme der Nullhypothese führen. Das heißt, eine Nullhypothese gilt, bis ihr die Fehlerhaftigkeit nachgewiesen werden kann.

Wichtige Arten von Hypothesen wurden in Teil II bereits genannt, es handelt sich im Rahmen empirischer Forschung insbesondere um die Zusammenhangs-, Unterschied- und Veränderungshypothese:

- **Zusammenhangshypothesen:** Hierbei handelt es sich um eine Hypothese, die eine Aussage über Zusammenhänge zwischen den Merkmalsausprägungen zweier (bivariat) oder mehrerer Merkmale (multivariat) trifft. Diese wird meist als „Wenn-dann“ oder „Je-desto“-Hypothese formuliert.
Bsp.: Wenn eine Person am Arbeitsplatz viel Stress (unabhängige Variable) ausgesetzt ist, dann hat diese mehr Krankenstandstage (abhängige Variable). D.h. es besteht ein Zusammenhang zwischen Stress am Arbeitsplatz und Krankenstandtagen.
- **Unterschiedshypothese:** Diese Hypothesenart vergleicht zwei oder mehrere Stichproben, die sich in einer oder mehreren Variablen (=Merkmalen) unterscheiden.
Bsp.: Das Einkaufsverhalten von Frauen in Bezug auf Bekleidung unterscheidet sich eindeutig von jenem der Männer.
- **Veränderungshypothese:** Die Ausprägungen einer Variable (=Merkmal) verändert sich im Laufe der Zeit. Das heißt, es sind mindestens zwei Messzeitpunkte erforderlich.
Bsp.: Je häufiger ein Produkt beworben (unabhängige Variable) wird, desto größer ist sein Bekanntheitsgrad (abhängige Variable).
- **Einzelfallhypothese:** Bei der Einzelfallhypothese bleibt das Untersuchungsfeld auf den Einzelfall beschränkt, man spricht auch von einer Fallstudie. Vermutungen über einen Sachverhalt werden durch systematische Analyse überprüft.
Bsp.: Das Süßwarenprodukt X weist während einer Preisreduktion (unabhängige Variable) eine höhere Absatzmenge (abhängige Variable) auf.

Allgemein ist zu bemerken, dass Hypothesen unterschiedlich konkret formuliert werden. Wünschenswert sind Aussagen, die möglichst präzise die Wirkung einer gegebenen Ursache vorhersagen (**gerichtete Hypothesen**).

Unspezifische (ungerichtete) Hypothesen enthalten keine weitere Angaben, in welcher Form der Zusammenhang bzw. der Unterschied erwartet wird. Spezifische (gerichtete) Hypothesen sagen hingegen die spezifische Richtung und teilweise die Stärke des Zusammenhangs bzw. des Unterschiedes voraus.

(c) Skalenniveaus

In der empirischen Forschung werden je nach Art der erhobenen Daten der Variablen verschiedene Messlatten (Skalen) verwendet. Nicht jede Merkmalsausprägung lässt sich gleich gut in Zahlen darstellen. Während dies beispielsweise für das Jahreseinkommen in Euro sehr einfach ist, ist es für das Geschlecht gar nicht möglich. Das Skalenniveau gibt somit auch vor, welche Rechenoperationen angewendet werden können. Es werden vier Skalenniveaus unterschieden (Abb. 7):

- **Nominalskala:** Die Nominalskala bietet den geringsten statistischen Informationsgehalt. Sie ist eine nicht-metrische (kategoriale) Skala, d.h. ihre Antwortwerte stehen nicht für einen direkt verwendbaren Zahlenwert. Sie wird deshalb auch als qualitativ bezeichnet und lässt nur Aussagen über Gleichheit oder Verschiedenheit zu. Rechenoperationen, die sinnvoll sind, sind Häufigkeiten und die Berechnung des Modus (häufigster Wert).
Bsp: Geschlecht: weiblich, männlich; Wohnort: Wien, Berlin, Frankfurt, Wiesbaden etc.

- **Ordinalskala:** Die in der Hierarchie nachfolgende Skala ist die Ordinalskala. Sie ist eine ebenfalls nicht-metrische (kategoriale) Skala, da ihre Antwortwerte keinen direkt verwendbaren Zahlenwert aufweisen. Die Ordinalskala wird auch Rangskala genannt, da sie die Aufstellung einer Rangordnung mit Hilfe von rangwerten erlaubt. Dies Rangordnung zeigt aber nicht, um wie viel A höher oder besser als B ist, d.h. die Abstände zwischen A und B können nicht abgelesen werden. Zulässige Rechenoperationen sind Häufigkeiten, Modus, die Berechnung des Median (der Wert in der Mitte) und der Quantile (Spannweite von Frequenzverteilungen).
 Bsp.: Ausbildungsniveau: Realschule, Abitur, Hochschule; Fahrzeugklasse: Kleinwagen, untere Mittelklassewagen, obere Mittelklassewagen, Oberklassewagen; Zustimmungsgrad: stimme zu, stimme eher zu, stimme eher nicht zu, lehne ab
- **Intervallskala:** Die Intervallskala ist eine metrische Skala, da ihre Antwortwerte einen direkt verwendbaren Zahlenwert aufweisen und die Abstände zwischen den einzelnen Einheiten gleich groß sind. Daher sind auch verschieden Rechenoperationen sinnvoll und zulässig. Die Intervallskala verfügt zwar über gleich große Skalenabschnitte, weist aber keinen **natürlichen** Nullpunkt auf. Ein typische Beispiel ist die Celsius-Skala zur Temperaturmessung. Zulässige Rechenoperationen sind Häufigkeiten, Modus, Median, Quantile, die Berechnung des Mittelwertes (arithmetisches Mittel, die Varianz und die Standardabweichung; es kann addiert und subtrahiert, nicht aber die Summe gebildet werden.
 Bsp.: Jahreszahlen, Temperaturmessung etc.
- **Ratioskala oder Verhältnisskala:** Diese stellt das höchste Skalenniveau dar. Sie ist ebenfalls eine metrische Skala, die zusätzlich einen natürlichen Nullpunkt aufweist. Dadurch können auch Verhältnisse zwischen Messwerten gebildet werden. Ratioskalierte Daten erlauben somit die Anwendung aller arithmetischen Rechenoperationen sowie die Anwendung aller statistischen Maße.
 Bsp.: Physikalische Maße, wie Flächen-, Größen-, Gewichts-, Längen- und Geschwindigkeitsmaße, Körpergröße, Geldeinheiten etc.

Skalenniveau	Aussagemöglichkeiten	Beispiele	Rechenoperationen
Nominalskala	Gleichheit/Ungleichheit	Geschlecht, Beruf, Marken	Häufigkeit, Modus
Ordinalskala	Gleichheit/Ungleichheit, Rangordnung	Rangplätze, Ausbildungsniveau	Häufigkeit, Modus, Median, Quantile
Intervallskala	Gleichheit/Ungleichheit, Rangordnung, Differenzbildung, ohne natürlichen Nullpunkt	Celsiuskala Jahreszahlen	Häufigkeit, Modus, Median, Quantile, Mittelwert, Varianz, Standardabweichung, Addition, Subtraktion
Ratioskala	Gleichheit/Ungleichheit, Rangordnung, Differenzbildung, mit natürlichen Nullpunkt	Physikalische Maße (wie Flächen, Größen, Gewichte), Körpergröße, Geldeinheiten	Häufigkeit, Modus, Median, Quantile, arithmetisches und geometrisches Mittel, Varianz, Standardabweichung, Addition, Subtraktion, Summe

Abb. 7: Überblick über die vier Skalenniveaus (Berger 2010)

(d) Theorie und Modell (vgl. Teil II)

Eine wissenschaftliche Theorie ist nichts anderes als ein spezifischer Ausschnitt der Realität. Eine Theorie enthält logisch widerspruchsfreie Aussagen über empirisch überprüfbare Zusammenhänge; diese Aussagen können beschreibend (= deskriptiv) oder erklärend (= kausal) sein. Man spricht in diesem Zusammenhang auch von einem integrierten Verbund von Hypothesen. Eine wissenschaftliche Theorie weist i.W. folgende vier Charakteristika auf:

- sie muss mittels empirischer Methoden überprüfbar und wiederholbar sein,

- sie muss objektiv sein,
- sie enthält allgemein gültige, logisch widerspruchsfreie Aussagen und
- diese Aussagen werden systematisch gewonnen.

Ein **Modell** ist eine mathematisch formalisierte Theorie, welche widerspruchsfrei und präzise ist. Das heißt, eine zu untersuchende Realität soll durch bestimmte Erklärungsgrößen im Rahmen einer wissenschaftlichen Theorie abgebildet werden und eine Hypothesenableitung ermöglichen. Da aber nicht alle Aspekte der untersuchten Realität in einem Modell abbildbar sind, erfolgt oftmals eine **Reduktion** oder **Abstraktion**. Anwendbare Modelle stellen **Erklärungen** zur Verfügung und erlauben es, **Prognosen** aufzustellen. Versucht man möglichst alle relevanten Einflussgrößen eines Gesamtzusammenhangs und deren gegenseitige Abhängigkeiten darzustellen, so spricht man von einem **Totalmodell** (z.B. Kaufentscheidungsmodell nach Howard/Sheth 1969; s. Folie). Nimmt man nur einzelne Teilbereiche aus einem Modell, so spricht man von einem **Partialmodell**. D.h. man versucht z.B. aus dem dargestellten Modell (s. Folie) nur die Outputvariablen herauszunehmen und diesen Teilzusammenhang zu beschreiben bzw. zu erklären.

2.2 Angewandte qualitative Sozialforschung

Entschließt man sich im Rahmen der wissenschaftlichen Abschlussarbeit, ein empirisches Forschungsprojekt durchzuführen, so sollte man sich vorab sorgfältig überlegen, mit welcher Forschungsmethode bzw. -richtung der größte Erkenntnisgewinn für die gewählte Problemstellung zu erreichen ist. Zur Untersuchung vieler wirtschafts- und sozialwissenschaftlicher Fragestellungen ist die Erhebung und Auswertung qualitativer Daten notwendig, vor allem dann, wenn es darum geht, **Ursachenforschung** zu betreiben, **Ideen** zu **generieren** oder eine **Grobauswahl** von **Alternativen** (Screening) zu treffen. Die qualitative Forschung bedient sich dieser qualitativen Daten, vor allem verbalisierter oder schriftlicher Daten oder Texte, die in erster Linie Bedeutungen, Strukturen und Veränderungen erfassen. Sie **verzichtet** überwiegend auf **Messungen** und arbeitet mit **Interpretationen** von verbalem oder schriftlichem Material. Oft macht es Sinn, qualitative und quantitative Forschungsmethoden zu kombinieren, um eine Problemstellung auch wirklich tiefgehend analysieren zu können. Man beginnt meist mit einer qualitativen Methode, wie z.B. mit einem qualitativen Interview, um vorerst eine Sachverhalt verstehen zu lernen. Diese verbalen Daten werden verarbeitet und interpretiert. Aus dem Ergebnis können dann zum Beispiel sinnvolle Fragen für eine quantitative Befragung abgeleitet werden.

Die wichtigsten Forschungsziele qualitativer Forschung sind:

- die Beschreibung empirischer Sachverhalte und sozialer Prozesse
- die Gewinnung von Hypothesen aus empirischen Material
- die Prüfung von Hypothesen
- die Aufstellung von Typologien oder Klassifikationen

Bei der **Klassifikationen** wird einer Menge von Elementen bzw. Objekten eine Menge von Merkmalen zugeordnet. Ob ein Element zu einer bestimmten Klasse gehört, ist davon abhängig, ob die vorher festgelegten Merkmale bzw. Merkmalsausprägungen vorliegen oder nicht. Klassifikationen liegen beispielsweise bei der Segmentierung von Kostenarten vor (Merkmal beispielsweise Beschäftigungsreagibilität – fixe und variable Kostenarten). Bei **Typologien** werden Einteilungen zwar auch mit Hilfe von Merkmalen vorgenommen. Allerdings weiß man von ihnen weder, ob sie hinreichend sind, noch, ob man eine vollständige Klassifikation erreichen kann (z. B. Betriebstypen).

Versucht man qualitative Forschungsmethoden zu charakterisieren, so lassen sich folgende typische Aufgaben ableiten, die im Vordergrund stehen:

- Generell wird mit kleineren Stichproben oder Fallstudien gearbeitet.
- Angestrebt wird die Theoriebildung („theorieentwickelnd“; Entdeckungszusammenhang).
- Dies wird dadurch erreicht, dass die eingesetzten Verfahren möglichst offen sind, entweder nicht strukturiert oder halbstrukturiert, dass sie sich an den Untersuchungsgegenstand anpassen und an dessen Relevanz orientieren. Man versucht, Gegenstände immer möglichst in ihrem natürlichen, alltäglichen Umfeld zu erforschen.
- Fokussiert wird bei der Untersuchung die Entstehung und Veränderung sozialer Phänomene. Die subjektive Sicht des Beforschten soll nachvollzogen werden. Man versucht, Regeln, Muster und Strukturen zu erkennen, auch wenn diese dem Beforschten gar nicht bewusst sind.
- Die Ergebnisse repräsentieren idealisierte, von Besonderheiten bereinigte Fälle (Typisierung).

Bei qualitativer Forschung fließen subjektive Wahrnehmungs- und Bewertungsprozesse in die Analyse mit ein, da diese vom Vorverständnis des Forschers geprägt sind. Dieses Vorverständnis muss daher immer offen gelegt und schrittweise am Gegenstand weiterentwickelt werden. Der gesamte Forschungsprozess ist eher zirkulär organisiert, d.h. er besteht aus Forschungszyklen, und nicht linear wie bei quantitativen Studien (Kap.2.2.2). Qualitatives Forschen ist somit der Versuch herauszufinden, wie Menschen einen Sachverhalt sehen, welche individuelle Bedeutung er für sie hat und welche Handlungsmotive in diesem Zusammenhang auftreten. Erleben und Verhalten wird durch **Verstehen** und **Interpretation** bestimmbar. Aus den Ergebnissen werden Theorien konstruiert und Folgerungen für die Praxis gezogen. Repräsentativität, Standardisierung und Strukturierung werden durch Offenheit, Breite, Detaillierung, Nähe und Interdisziplinarität ersetzt.

2.2.1 Gütekriterien qualitativer Forschung

Bei den Gütekriterien handelt es sich um Regeln und Kriterien, die einen Maßstab für die Qualität der Forschung, d.h. einzelner Instrumente und des Forschungsdesign bereitstellen:

- Ein wesentliches Kriterium ist die **intersubjektive Nachvollziehbarkeit**. Im Gegensatz zur quantitativen Forschung geht es nicht um Überprüfbarkeit, da die Replikation aufgrund der begrenzten Standardisierbarkeit nicht möglich ist. Vielmehr geht es darum, den Forschungsprozess mit all seinen Entscheidungen, Problemen und Ergebnissen genauestens zu dokumentieren, damit die Möglichkeit besteht, den Forschungsprozess nachzuvollziehen und die entstandenen Ergebnisse entsprechend beurteilen zu können (Verfahrensdokumentation).
- Die durch Interpretation getroffenen Annahmen müssen vom Forscher argumentiert werden (argumentative **Interpretationsabsicherung**). Zudem ist es sinnvoll, die erhobenen Daten vom Beforschten auf ihre Gültigkeit hin bewerten zu lassen (kommunikative **Validierung**). So kann sichergestellt werden, dass Aussagen korrekt aufgefasst und interpretiert worden sind.
- Ein weiteres wichtiges Gütekriterium ist, dass während des Forschungsprozesses ein Vertrauensverhältnis zum Beforschten hergestellt werden soll (Nähe zum Untersuchungsgegenstand)

2.2.2 Durchführung qualitativer Studien

Die Durchführung qualitativer Studien erfolgt nicht linear in einer Abfolge von Forschungsschritten, wie es bei quantitativen Studien der Fall ist, sondern die Hauptforschungsphase erfolgt in Forschungszyklen.

Zu den Forschungsphasen gehören (s. Folie) die

- Planungsphase
- Orientierungsphase
- Hauptforschungsphase mit Forschungszyklen
(jeder Zyklus besteht aus Erhebungen, Analysen, Prüfungen, Zwischenbilanz der vorläufigen Ergebnisse)
- Forschungsdarstellung

Für welche **Erhebungsmethode** man sich letztendlich entscheidet, hängt in erster Linie vom jeweiligen Erkenntnisinteresse ab. Einige wichtige Erhebungsmethoden werden im folgenden Kapitel (kurz) vorgestellt.

2.2.3 Ausgewählte qualitative Erhebungsmethoden

Methoden der qualitativen Sozialforschung sind bezogen auf den wissenschaftlichen Erkenntnisprozess v.a. einsetzbar auf den ersten drei Ebenen der Erkenntnisgewinnung (Teil I), also Definition, Klassifikation und Deskription. Bei der auf soziale Wirklichkeit ausgerichteten Sozialforschung kommt als wissenschaftlicher Grundsatz anstatt des **deduktiven Erklärens** (vom Allgemeinen/der hypothetisch unterstellten Regelmäßigkeit auf das Besondere/den Einzelfall) vorrangig das **induktive Verstehen** (vom besonderen Einzelfall zur allgemeinen Invarianz) und Vorgehen zum Einsatz. An die Stelle der bei quantitativen Sozialforschungsprojekten notwendigen Vorablegung jedes einzelnen Forschungsschritts tritt das so genannte Prinzip der Offenheit. Die hiermit korrespondierenden Erhebungsmethoden der qualitativen Forschung sind v.a die qualitative Inhaltsanalyse, Beobachtungen und niedrig abstrahierte Befragungen (s. Beispiel in der Vorlesung).

- Die **qualitative Inhaltsanalyse** beschäftigt sich mit der Auswertung von erhobenen Daten in Form von z.B. Texten oder Filmen. Besonders eignet sich die Methode für die systematische, theoriegeleitete Bearbeitung von Textmaterial. Latente Bedeutungen und Sinnstrukturen, unabhängig von der Intention des Textproduzenten, sollen analysiert werden. Beispielsweise können vorliegende Dokumente zu Wertschöpfungsprozessen, Produkten oder Vermarktungskonzepten und generell zu Organisationen erfasst und interpretiert, präzisiert und auf die mit ihrem Einsatz/ihrer Gestaltung beabsichtigte Wirkung (als angestrebte Ziele) untersucht werden. Die Problematik, in einer systematischen Weise die wesentlichen Inhalte vorliegender Materialien/Unterlagen zum jeweiligen Erkenntnisgegenstand herauszuarbeiten, stellt sich hierbei zu Beginn der empirischen Forschungsarbeit. Wichtig ist es daher vorab ein **Kategoriensystem** zu definieren und über das Festlegen eines geeignet aufgegliederten „**Erfassungssystems**“ mit bedeutungsbezogenen Zuordnungsregeln („Kodierregeln“) eine wesentliche Grundlage für das Gewinnen aussagefähiger Ergebnisse zu legen (s. Folie).
- **Niedrig abstrahierte, mündliche Befragungen** kommen ebenfalls häufig als Erhebungsmethode im Rahmen qualitativer Forschung zum Einsatz. Hierzu zählen in erster Linie **narrative Interviews** (geringster Standardisierungsgrad, kaum strukturiert, ohne Leitfaden), bei denen unter Vorgabe eines Rahmenthemas (z.B. Käufererlebnisse,

Arbeitssituationen, Freizeitaktivitäten etc.) zu einem „freien Erzählen“ aufgefordert wird, oder **Leitfadeninterviews** (halbstrukturierte Interviewform, mit Leitfaden; Reihenfolge der Fragestellung ist aber nicht zwingend einzuhalten), deren Schwergewicht auf freien Assoziationen der Beteiligten (offene Fragen) liegt. Zur Datenerhebung werden Leitfäden eingesetzt, die eine Vergleichbarkeit mehrerer Interviews ermöglichen; wichtig ist es hierbei zu beachten, dass genügend Spielraum für spontane Äußerungen vorhanden sein muss. Eine Form des Leitfadeninterviews, die häufig in wissenschaftlichen Arbeiten eingesetzt wird, ist das **Experteninterview**. Bei dieser Form von Interviews steht nicht der Befragte als Person, sondern seine Erfahrungen und Interpretationen im Hinblick auf das interessierende Forschungsthema im Vordergrund. Auch hier wird oft ein halbstandardisierter Fragebogen bzw. Leitfaden eingesetzt, der sicherstellt, dass das relevante Kriterienraster angesprochen wird, aber dennoch genügend Freiraum für individuelle Aussagen lässt. Alternativ oder ergänzend können auch Workshops als themenzentrierte **Gruppendiskussion** (Fokusgruppe) durchgeführt werden.

- **Beobachtungen** stehen vom Prinzip her in engem Zusammenhang zu Experimenten (s. Methoden quantitativer Sozialforschung). Anders als beim inhaltsanalytischen Vorgehen werden im Rahmen von Beobachtungen nicht die materialisierten Ergebnisse menschlicher Handlungen auf ihren Bedeutungsinhalt hin untersucht; jetzt richtet sich das Erkenntnisinteresse direkt auf Handlungs-, Verhaltens- oder Interaktionsprozesse im sozialen Feld. Diese sind in einer unmittelbar sinnesorganischen Wahrnehmung (Sehen, Hören, etc.) zu erfassen und in ihrer „latenten Bedeutung“ vor dem Hintergrund der Bedürfnis- und Motivlagen der beteiligten Akteure einzuordnen. Beobachtung lassen sich nach folgenden Dimensionen differenzieren: **Erkennbarkeit** des Beobachter (offene vs. verdeckte Beobachtungen), **Interagieren** des Beobachters (teilnehmende vs. nichtteilnehmende Beobachtungen), Grad der **Standardisierung** (strukturierte vs. unstrukturierte Beobachtungen). Das generelle Ziel besteht darin, situationspezifisch alle relevanten Einflussfaktoren zu erfassen und damit zu kontrollieren sowie in Bezug zu nachvollziehbarem Handeln und Verhalten zu setzen. In der Forschung ist dies insbesondere in **kaufrelevanten** Situationen im Handel oder in **Führungssituationen** im Unternehmen von Interesse.

2.3 Angewandte quantitative Sozialforschung

In zahlreichen wissenschaftlichen Abschlussarbeiten werden quantitative Studien durchgeführt. Als Erhebungsmethode werden dabei gerne Befragungen mittels strukturiertem und standardisiertem Fragebogen genutzt. Diese sind einfach und relativ schnell auszuwerten und lassen wenig Spielraum für freie Interpretationen zu, was für weniger geübte Forscher von Vorteil ist. Im Folgenden wird ein Überblick über die quantitative Forschung gegeben, relevante Methoden vorgestellt und erläutert, wie eine Datenauswertung durchgeführt wird.

Die quantitative Sozialforschung ermöglicht die Messung von Ausschnitten der Realität, die in eine statistische Verarbeitung der Messwerte mündet. Die Quantifizierung, d.h. die Angabe von Zahlenwerten, erfolgt durch Operationalisierung (Messbarmachung) von Merkmalen (s. Folie). Die wichtigsten Unterschiede zu qualitativen Untersuchungen wurden bereits behandelt. An dieser Stelle ist nochmals zu betonen, dass quantitative Verfahren mit **großen Stichproben** arbeiten, da versucht wird, zu **repräsentativen** Ergebnissen zu gelangen. Diese Ergebnisse sind quantifizierbar, d.h. in Zahlen auszudrücken. Durch den hohen Grad an **Objektivität** ist kein individuelles Eingehen auf Probanden möglich. Setzt man sich das Untersuchungsziel Ursachenforschung zu betreiben, dann sollte man also besser auf qualitative Verfahren zurückgreifen.

2.3.1 Hauptziele quantitativer Untersuchungen

Quantitative Untersuchungen können **deskriptiv** (beschreibend) oder **explanativ** (erklärend) angelegt sein:

- Bei **deskriptiven** Untersuchungen geht es in erster Linie darum, Populationen zu beschreiben. Das heißt im Detail, dass Häufigkeiten, Anteile, Mittelwerte etc. von Eigenschaften in einer Population beziffert werden. Dies ist bspw. bei der gesamten amtlichen Statistik der Fall (z.B. Volkszählungen). Da Vollerhebungen aus ressourcens-technischen Gründen in den meisten Fällen nicht möglich sind, werden **Stichprobenerhebungen** durchgeführt. Diese Stichproben sollen **repräsentativ** für die Population (Grundgesamtheit) sein; für ihre Zusammensetzung gibt es verschiedene Möglichkeiten (s.u. Kap. 3). Ziel ist es, mit möglichst genauer **Schätzung** unbekannte Merkmalsausprägungen in der jeweiligen Population zu eruieren.
- Ziel der **explanativen** Untersuchungen ist die Überprüfung und Erklärung von begründeten Hypothesen; explanativ wird hier im Sinn von **hypothesenprüfend** verstanden und steht insofern methodologisch mit der wissenschaftstheoretischen Position des Kritischen Rationalismus in Einklang. Untersucht werden entweder Zusammenhänge, Unterschiede, oder bestimmte Einzelfälle. Darauf begründen (s.o. Kap. 2.1.4 sowie Teil II) sich auch die unterschiedlichen Arten von Hypothesen, die bereits vorgestellt wurden, und hier hinsichtlich ihrer Überprüfungsverfahren nochmals angesprochen werden.
 - Bei **Zusammenhangshypothesen** geht es um die Überprüfung der **Art** und **Intensität** des miteinander Variierens zweier Merkmale (bivariat) oder mehrerer Merkmale (multivariat). Die Überprüfung erfolgt mittels **Chi-Quadrat-Test** oder **Korrelations- und Regressionsanalyse** (Kap. 4)
 - Bei **Unterschiedshypothesen** wird ein **Vergleich** zweier oder mehrerer Stichproben angestellt, die sich in einer oder mehrerer unabhängiger Variablen unterscheiden. Zur Überprüfung werden zum Beispiel **Experimente** durchgeführt, in welchen ein Vergleich mit einer Kontrollgruppe erfolgt. Die Überprüfung kann statisch mittels **t-Test** erfolgen.
 - **Veränderungshypothesen** werden mittels Querschnitt- oder Längsschnittuntersuchungen überprüft.
 - Bei der Überprüfung von **Einzelfallhypothesen** bleibt das Untersuchungsfeld auf einen Einzelfall beschränkt, d.h. es wird eine Fallstudie durchgeführt. Theoriegeleitete Vermutungen werden durch systematische Analyse überprüft.

2.3.2 Gütekriterien quantitativer Forschung

In der quantitativen Sozialforschung werden verschiedene Kriterien verwendet, um die Güte einer Untersuchung beurteilen zu können. Damit die Messergebnisse und die daraus resultierenden Schlussfolgerungen verlässlich sind, müssen zumindest folgende drei Kriterien bzw. Anforderungen bei der Durchführung der quantitativen Untersuchung erfüllt sein:

- **Objektivität** meint hier den Grad, in dem die Ergebnisse einer Messung unabhängig von den Anwendern sind; d.h., dass mehrere Personen, die unabhängig voneinander die Messung durchführen zum gleichen Ergebnis gelangen. Objektivität fordert und besagt demnach, dass keine Verzerrungen durch Forscher oder das Messinstrumentarium stattfinden darf: Der Test gilt dann als vollständig objektiv, wenn Anwender A und B mit dem gleichen Messinstrument zu einem identischen Resultat gelangen.

- Unter **Reliabilität** versteht man das Ausmaß, in dem ein Erhebungsinstrument bei wiederholten Datenerhebungen (auch wenn sie von unterschiedlichen Personen durchgeführt wird) unter gleichen Bedingungen und bei denselben Probanden das gleiche Ergebnis erzielt - die Messung also zuverlässig ist, weil die Messwerte reproduzierbar und über mehrerer Messvorgänge sowie bei mehreren Mess-Personen stabil sind.
- Die **Validität** gibt an, inwieweit ein Erhebungsinstrument tatsächlich die Variable misst, die es zu messen vorgibt. Ein Test ist dann valide, wenn er genau das misst, was gemessen werden soll. Eine Körperwaage beispielsweise gilt demnach als valide, wenn sie genau das misst, was sie messen soll, d.h. das Körpergewicht.

2.3.3 Ausgewählte quantitative Erhebungsmethoden

Mit Blick auf die Wissenschaftstheorie sind die Methoden der quantitativen Sozialforschung im wissenschaftlichen Erkenntnisprozess stärker auf der 4. und 5. Ebene angesiedelt (Teil I). Im Zentrum stehen also die **Erklärung** und die **Prognose** im Rahmen der Theorienbildung und -überprüfung sowie die **Technologie** und dabei mit der Formulierung von Gestaltungs- und Handlungsempfehlungen verbundene Zielsetzungen. Somit werden Methoden der quantitativen Datenerhebung eingesetzt, um **Hypothesen** bzw. kausale Strukturen an der Realität zu **überprüfen**, d.h. zu bestätigen/widerlegen/modifizieren. Für den angestrebten „**hypothetisch-deduktiven**“ Prozess der Erklärung, Prognose und Gestaltung benötigt die quantitative Sozialforschung eine vom Skalenniveau her **gesicherte** Datenbasis. Im Vordergrund stehen deshalb das Messen sowie insbesondere das umfangreiche Quantifizieren sozialer Sachverhalte mittels **standardisierter** Erhebungsinstrumente. Zum Gewinnen repräsentativer Aussagen wird mit **großen** Fallzahlen/Stichproben gearbeitet. Deren hypothesengeleitet Auswertung erfolgt mit **statistischen** Verfahren (Datenauswertung/Hypothesentest). Der Schwerpunkt quantitativer Sozialforschung liegt damit im **Begründungszusammenhang** empirischer Forschungsprojekte zur Untersuchung menschlichen Handelns, Verhaltens und Entscheidens. Hierzu bieten sich vor allem standardisierte, meistens schriftliche, heute aber zunehmend elektronische **Befragungen** an. Zum Einsatz kommen aber auch **Beobachtung** und **Experiment**.

- Quantitative **Befragungen** unterscheiden sich von „offenen, unstrukturierten“ qualitativen Befragungen, bei denen nur minimale Vorgaben und der Befragungsschwerpunkt erst im Gesprächsverlauf festgelegt werden, durch ihren Grad der Strukturierung bzw. Standardisierung. Bei einer „vollständig strukturierten“ quantitativen Befragung werden alle Fragen mit vorgegebenen Antwortkategorien in festgelegter Reihenfolge gestellt. Die quantitative Befragung kennzeichnet eine Vorgehensweise nach standardisierten Grundsätzen. Durch diese Vereinheitlichung von Fragebögen (und Interviews) sollen möglichst vergleichbare Ergebnisse erzielt werden. Quantitative Befragungen können mündlich („face-to-face“), telefonisch, schriftlich oder online durchgeführt werden.
- Unter einer quantitativen **Beobachtung** versteht man das systematische Erfassen und Festhalten von Verhalten zum Geschehenszeitpunkt. Der Forschungsablauf einer quantitativen Beobachtung muss in einem standardisierten Rahmen ablaufen und intersubjektiv überprüfbar sein. Grundsätzlich sind die selben Beobachtungsarten wie bei der qualitativen Forschung möglich.
- In Experimenten werden hypothetisch vermutete Ursachen oder Bedingungen manipuliert, um eine Wirkung oder einen Effekt bei paralleler Kontrolle möglicher Störfaktoren bzw. experimenteller Randbedingungen zu erzielen. Es ist damit eine nach festgelegten Bedingungen ablaufende Überprüfung von bereits vorher theoretisch festgelegten Aussagen. Unterscheiden lassen sich insbesondere Feld- und Laborexperimente.

- bei **Laborexperimenten** wird die Umweltsituation vom Forscher arrangiert: die Bedingungen sind also kontrolliert, künstlich und konstant. Dem Vorteil der Konstanzhaltung und Überwachung der Randbedingungen steht der Nachteil der Realitätsferne gegenüber. Ein Beispiel für ein Laborexperiment ist die Verwendung von Blickaufzeichnungsgeräten in der Werbewirkungsforschung
- Das Feldexperiment wird in einer natürlichen Umgebung durchgeführt (z.B. im Supermarkt), was als Vorteil eine hohe Realitätsnähe, eine geringe Beeinflussung durch den Forscher und somit eine bessere Voraussetzung für eine Verallgemeinerung mit sich führt. Der Nachteil bei einem Feldexperiment liegt in der erschwerten Kontrolle und Aufzeichnung aller Bedingungen und Reaktionen. Ein Beispiel für ein Feldexperiment ist die Untersuchung verschiedener Regalstandplätze in einem Handelsunternehmen in ihrer kaufverhaltensbezogenen Wirkung auf unterschiedliche Ergebnisgrößen (z.B. Absatzzahlen).

2.3.3 Durchführung quantitativer Studien

Der Ablauf einer quantitativen Studie besteht im Wesentlichen aus fünf Schritten:

1. Formulierung und Präzisierung des Forschungsproblems
2. Planung und Vorbereitung der Erhebung
3. Datenerhebung
4. Datenauswertung
5. Berichterstattung

Formulierung und Präzisierung des Forschungsproblems

Ein Forschungsprojekt beginnt mit der **Formulierung** des Forschungsproblems in Form wissenschaftlicher Fragestellungen. Die Entscheidung über das Forschungsproblem, sowie die Auswahl des Forschungszieles hängt vom Typ der Studie ab. So wird entweder eine **überprüfbare** Hypothese oder eine klar definiert **Fragestellung** als Ansatzpunkt herangezogen.

Planung und Vorbereitung der Erhebung

In der zweiten Phase werden zunächst die in den Forschungshypothesen auftretenden Begriffe **definiert** und **operationalisiert**. Durch die Operationalisierung wird versucht, den theoretischen Begriff der Hypothese durch konkrete Indikatoren und Merkmale zu erfassen und dadurch **messbar** zu machen. Darauf folgt die Ausarbeitung des Untersuchungsdesign, die Auswahl an geeigneten Erhebungsmethoden und die Erstellung der Erhebungsinstrumente, wie z.B. Fragebogen oder Beobachtungsprotokoll. Die Wahl der Untersuchungsform hängt von der Fragestellung ab. Zur Planung und Vorbereitung der Erhebung quantitativer Studien muss die zu erhebende **Stichprobe** festgelegt werden (Kap. 3). Die Größe der Stichprobe bestimmt die Schätzgenauigkeit der Untersuchung. Die Stichprobe soll für die Grundgesamtheit repräsentativ sein.

Datenerhebung

Nachdem die Planung sorgfältig durchgeführt und abgeschlossen ist, kann die Erhebung der notwendigen Daten erfolgen. Je nach Fragestellung und Untersuchungsdesign werden für die Erhebung verschiedene Instrumente eingesetzt. Als Arbeitsort der Primärforschung gelten entweder das soziale Feld oder das Labor.

Datenauswertung

In dieser Phase folgen Eingabe, Aufbereitung und Analyse der erhobenen Daten. Zuerst ist die Überprüfung der erhobenen Daten vorzunehmen, um Unvollständigkeiten oder Mängel zu beheben. Das gewonnene Material wird so aufbereitet, dass es der Anwendung von statistischen Analysen dienen kann. Der Erfassung der Daten folgt deren Auswertung und Analyse durch Untersuchung einzelner Variablen (univariate Methoden), oder Auswertungen von Beziehungen zwischen mehreren Variablen (bivariate bzw. multivariate Methoden). Hilfestellung bieten statistische Analyseprogramme, wie SPSS und Office Programme, wie Excel oder Access. Die Ergebnisse von Datenanalysen können in tabellarischer oder grafischer Form dargestellt werden.

Berichterstattung

Am Ende des Forschungsprozesses steht die Schlussfolgerungsphase. Die Schnittstelle zwischen Datenauswertung und Berichterstattung bildet die Interpretation der Ergebnisse, einerseits innerhalb der Untersuchung und andererseits zwischen bestehenden Forschungsergebnissen. Der Forschungsbericht, welcher den empirischen Teil einer wissenschaftlichen Arbeit darstellt, soll die Fragestellung, die einzelnen methodischen Schritte und die Darstellung der Ergebnisse dokumentieren sowie Interpretationen und Schlussfolgerungen (Handlungs- bzw. Gestaltungsempfehlungen) enthalten.

3 Stichprobenplanung

Eine der wesentlichen Aufgaben bei der Fragebogenerhebung ist die **Stichprobe**. Folgende Fragen sollten in diesem Zusammenhang beantwortet werden:

- Welche Personen sollen befragt werden (Definition der Grundgesamtheit)?
- Wie wähle ich die Personen aus (Totalerhebung oder Stichprobenerhebung)?
- Wie viele Personen müssen befragt werden (Stichprobenumfang)

3.1 Grundgesamtheit und Stichprobe

Möchte man eine Befragung der Einwohner von Deutschland durchführen, so wäre eine mögliche Definition der angestrebten Grundgesamtheit „alle Einwohner von Deutschland über 18 Jahren“. Entschließt man sich beispielsweise zur Durchführung einer telefonischen Befragung, so entspricht die Auswahlgesamtheit der Definition „alle Einwohner von Deutschland über 18 Jahren, die einen Telefonanschluss haben“. Die Stichprobe (Sample) wiederum ist eine Auswahl von Elementen der Grundgesamtheit. Sie soll so gewählt werden, dass sie möglichst repräsentativ (stellvertretend) für die zugrunde liegende Grundgesamtheit (Population) ist, so dass vom Sample auf die Grundgesamtheit geschlossen werden kann. Die Wahrung der Repräsentanz fordert somit, dass aus der Grundgesamtheit eine Teilmenge gezogen wird, die die gleichen charakteristischen Eigenschaften wie die Grundmenge aufweist.

Die für eine statistische Untersuchung **relevanten** Merkmalsträger fasst man zur **Grundgesamtheit** zusammen. Wird bei einer statistischen Erhebung die Grundgesamtheit vollständig erfasst, so spricht man von einer **Vollerhebung** oder **Totalerhebung**, anderenfalls von einer **Teilerhebung** oder **Stichprobenerhebung**.

- Eine **Vollerhebung** empfiehlt sich, wenn die Zahl der Personen in der Grundgesamtheit „überschaubar“ ist (z.B. BWL-Professoren an deutschen Fachhochschulen).

- Zeit- und Kostengründe zwingen indessen häufig dazu, sich bei der Datenerhebung auf einen Teil der Grundgesamtheit zu beschränken (Teilerhebung).

Entsprechende Stichproben lassen sich mit verschiedenen Verfahren ziehen. Im Allgemeinen unterscheidet man (als Gruppen von Auswahlverfahren) die **Zufallsauswahl**, die **nicht-zufallsgesteuerte Auswahl** und **komplexe Verfahren** (s. Folie):

3.2 Stichprobenverfahren

(1) Zufallsauswahl

Wer eines der folgenden Verfahren anwendet, gibt jedem Element der Grundgesamtheit dieselbe **berechenbare** Chance, in die Stichprobe zu gelangen. Voraussetzung ist, dass für die Ziehung ein vollständiges Verzeichnis der Grundgesamtheit vorliegt (z.B. Kundendatei; Verzeichnis aller BWL-Professoren an deutschen Fachhochschulen).

- Die **systematische Auswahl** mit Zufallsstart (= sog. Herausgreifen des n-ten Falles) ist einfach zu handhaben. Angenommen man verfügt über ein vollständiges Verzeichnis aller Elemente in der Grundgesamtheit, z.B. eine alphabetisch geordnete Kundendatei mit 10.000 Adressen; um eine Stichprobe von 1.000 Personen zu ziehen, wählt man jede zehnte Adresse aus (d.h. Herausgreifen des n-ten, also hier: zehnten Falles)
- Bei der **Lotteriewahl** werden Zettel o.ä. aus einem Behälter gezogen, weshalb sie für große Stichproben zu aufwendig ist.
- Wer **Zufallszahlen** verwendet, ordnet jedem Element der Grundgesamtheit eine Zahl zu, wobei diese Elemente in einem zweiten Schritt mit Hilfe eines Zufallszahlengenerators ausgewählt werden. Auch dieses Verfahren ist in der Regel zu aufwendig und damit wenig praktikabel.

(2) Nicht-zufallsgesteuerte Verfahren

- Bei der **willkürlichen Auswahl** benötigt man keinen Erhebungsplan, sondern geht aufs Geratewohl zu. Ein Beispiel hierfür ist die Befragung von Kunden, die zu einer bestimmten Stunde ein Geschäft betreten. Je nach Tageszeit wird man dann in der Mehrzahl Berufstätige, Rentner, Schüler oder Studierende antreffen, deren Angaben aber keinesfalls die Meinung sämtlicher Kunden widerspiegeln. Folglich ist der Wert solcher Befunde stark eingeschränkt.
- Wer **bewusst** auswählt, muss die Struktur der Grundgesamtheit kennen:
 - Beim **‚Cut off-Verfahren‘**, das sich besonders bei Studien im Investitionsgüter- und im Handelssektor bewährt hat, berücksichtigt man nur die wichtigsten Elemente, z.B. Handelsunternehmen ab einem bestimmten Umsatzvolumen.
 - Beim **Quota-Verfahren** werden Merkmalsquoten vorgegeben, die bei der Auswahl der Probanden zu berücksichtigen sind. Hierbei geht man davon aus, dass innerhalb der Grundgesamtheit die zu untersuchenden Merkmale stark unterschiedlich verteilt sind. Man teilt daher (z.B. auf Basis statistischer Unterlagen) die Gesamtheit in verschiedene Quoten auf, z.B. nach Alter, Bildungsgrad, Geschlecht etc. Die beabsichtigte Stichprobe wird nun, je nach ermittelten Prozentanteilen, auf die einzelnen Quoten verteilt. Das Quota-Verfahren ist einfach zu handhaben sowie eine zeitsparende und flexible Methode zur Datenerhebung.

Beispiel Quota-Verfahren

Angenommen es sollen 60 Personen in vier Altersklassen (unter 18 Jahren; 18 bis 40 Jahre; 41 bis 64 Jahre; über 64 Jahre) befragt werden. Um die Grundgesamtheit näherungsweise abzubilden, sollen in der Gruppe der Befragten – analog zur Grundgesamtheit – lediglich 20% der Probanden unter 18 bzw. über 64 Jahre alt sein, während die beiden anderen Altersgruppen zu 10% (= 18 bis 40 Jahre) bzw. zu 50% (=41 bis 64 Jahre) in der Stichprobe vertreten sein sollen. Außerdem soll (entsprechend der **Grundgesamtheit**) darauf geachtet werden, dass ebenso viele Frauen wie Männer (= 50%/50%). Auf Basis dieser Vorgaben muss der Interviewer bspw. darauf achten, dass lediglich 3 männliche und 3 weibliche Personen zwischen 18 und 40 Jahren befragt werden. Für die Altersgruppe 41 bis 64 Jahre benötigt man hingegen jeweils 15 männliche und 15 weibliche Probanden.

Altersklasse	Geschlecht		Summe
	Männlich (= 50%)	Weiblich (= 50%)	
Unter 18 Jahre (= 20%)	6	6	12
18 bis 40 Jahre (= 10%)	3	3	6
41 bis 64 Jahre (= 50%)	15	15	30
Über 64 Jahre (= 20%)	6	6	12
Summe	30	30	60

Abb.8: Beispielhafter Quotenplan als Grundlage des Quota-Verfahrens (Kornmeier 2007)

(3) Komplexe Formen der Stichprobenziehung

- Beim **Klumpenverfahren** teilt man die Grundgesamtheit in Untergruppen auf, um daraus dann – zufällig oder systematisch – einzelne Gruppen auszuwählen. Diese wiederum werden vollständig in die Stichprobe einbezogen. Voraussetzung ist, dass dieses Sample „typisch“ (und damit in gewissem Sinne „repräsentativ“) für die Grundgesamtheit ist. Anstatt bspw. alle Mitarbeiter eines nationalen Kreditinstitutes zu befragen, könnte man sich darauf beschränken, lediglich in allen Filialen **einer Stadt** eine Vollerhebung durchzuführen. Mit Blick auf Kosten- und Zeitersparnis wäre diese Vorgehensweise durchaus vorteilhaft. Sie liefert allerdings schlechte Ergebnisse, wenn die Stichprobe „nicht typisch“ ist, die Filialmitarbeiter in Stadt A also andere Vorstellungen haben als jene in Stadt B oder in der ländlichen Region C („**Klumpeneffekt**“).
- Hierarchisch strukturierte Grundgesamtheiten (z.B. regionale Niederlassungen eines nationalen Kreditinstituts → Filialen innerhalb der regionalen Niederlassungen → Mitarbeiter der einzelnen Filialen) kann man sich bei der **mehrstufigen Auswahl** zunutze machen. So wäre es möglich, zunächst unter den **Niederlassungen** eine Stichprobe zu ziehen, um daraus wiederum **Filialen** auszuwählen. Abschließend zieht man unter den **Mitarbeitern** der jeweils ausgewählten Filialen erneut eine Stichprobe.
- Bei der **geschichteten Auswahl** wird eine heterogene Grundgesamtheit (z.B. Handelsunternehmen) anhand bestimmter Kriterien in möglichst homogene Schichten aufgeteilt, etwa in A-, B- und C-Handelsunternehmen (bspw. Abhängig vom Umsatzvolumen pro Jahr). In diesem Zusammenhang ist auch zu klären, ob die fraglichen Segmente in der Stichprobe gleich stark vertreten sein sollen (= proportionale Stichprobe) oder ob man manchen Schichten – z.B. nach Maßgabe ihrer Relevanz (etwa Handelsunternehmen im A-Segment – einen größeren Anteil zubilligt (= disproportionale Stichprobe).

3.3 Stichprobengröße

(statistische Überlegungen hierzu in der Vorlesung)

Bei der Planung einer empirischen Untersuchung muss man sich überlegen, welche Stichprobengröße notwendig ist, um eine gewünschte Schätzwertgenauigkeit zu erreichen.

Das Ausmaß des verantwortbaren Schätzfehlers hängt von der notwendigen **Genauigkeit** der Untersuchungsergebnisse und damit von der Fragestellung der Untersuchung ab.

Geht es z.B. darum, ein Wahlergebnis zu prognostizieren, vor allem wenn ein knappes Resultat zu erwarten ist, so muss man mit möglichst großer Genauigkeit arbeiten. In einer Pilot-Studie, etwa zur Untersuchung von Freizeitbeschäftigungen bei Jugendlichen, dürfte hingegen eine kleinere Stichprobe und damit ein Verzicht auf sehr hohe Genauigkeit, vertretbar sein.

Sofern die Größe der Grundgesamtheit N nicht sehr klein ist ($N > 100.000$), spielt sie bei der Bestimmung der Stichprobe eine zu vernachlässigende Rolle. Damit bleibt der erforderliche Stichprobenumfang gleich, ob man nun die Bevölkerung eines ganzen Landes oder nur einer Stadt erfassen will. D.h. der Umfang einer Stichprobe hängt nicht von der Größe der Grundgesamtheit N, sondern vom **Signifikanzniveau** t (= Sicherheitsgrad) und dem **Konfidenzintervall** e (= Genauigkeit bzw. Schwankungsbreite des Messwertes). Der Umfang einer Stichprobe hängt demnach davon ab,

- wie **genau** das Stichprobenergebnis sein soll und
- mit welcher **Sicherheit** die Aussagen zutreffen sollen, wobei sich
 - eine Sicherheit von mindestens 95,5% mit
 - einer Genauigkeit von 5% als empfehlenswert erwiesen hat.

Um die Stichprobengröße bestimmen zu können, müssen t und e vorab festgelegt und in folgende Formel(n) eingesetzt werden:

$$n = \frac{t^2 \cdot p \cdot q}{e^2} \text{ für } N > 100.000 \quad \text{bzw.} \quad n = \frac{t^2 \cdot p \cdot q \cdot N}{t^2 \cdot p \cdot q + e^2 \cdot (N - 1)} \text{ für } N < 100.000$$

Dabei gilt:

n	=	Stichprobenumfang
t	=	Sicherheitsgrad: für t = 1 gilt: 68,3 % Sicherheit für t = 2 gilt: 95,5% Sicherheit für t = 3 gilt: 99,7% Sicherheit (ablesbar aus tabellierten Normalverteilungen in Statistikbüchern)
e	=	Schwankungsbreite des Messwertes, Genauigkeit
p	=	Anteil der Elemente in der Stichprobe, welche die Merkmalsausprägung aufweisen
q	=	Anteil der Elemente in der Stichprobe, welche die Merkmalsausprägung nicht aufweisen (wenn p und q im Voraus nicht bekannt sind, geht man von einer Gleichverteilung aus, also jeweils 50%: p = 0,5; q = 0,5)
N	=	Grundgesamtheit

Bei zahlreichen Erhebungen übersteigt die Grundgesamtheit die Mindestgröße (um ein Vielfaches), so dass die Berechnung des Stichprobenumfangs nur davon abhängt, wie genau das Stichprobenergebnis sein soll (Schwankungsbreite) und mit welcher Sicherheit die Aussage getroffen werden soll (Wahrscheinlichkeit). D.h. wenn man bspw. repräsentative Umfrage in Deutschland und eine in Österreich durchführen möchte, muss man deshalb nicht in Deutschland (ca.) 10-mal mehr Personen befragen.

Folgende Beispiele verdeutlichen die Anwendung oben angeführter Formeln für eine große und eine kleine Stichprobe:

Bsp. 1: Bei einer Grundgesamtheit von 5.000 Personen ergibt sich ein Stichprobenumfang von 370 Probanden, wenn man eine Sicherheit von 95,5% und eine Genauigkeit von 5% anstrebt. Berücksichtigt man überdies, dass bei **schriftlichen** Befragungen lediglich ein Teil der Angeschriebenen antwortet und legt man deshalb eine (geschätzte) Rücklaufquote von 20% zugrunde, dann müssten demnach 1.850 Fragebögen versandt werden.

Bsp.2: Es soll eine Studie über das Rauchverhalten der Einwohner eines Landes durchgeführt werden. Aus vorherigen Untersuchungen sei bekannt, dass die Verteilung von Rauchern zu Nichtrauchern in der Grundgesamtheit (alle Einwohner älter als 16 Jahre) dem Verhältnis 40 zu 60 entspricht. Somit hat p den Wert 0,4 und q den Wert 0,6 (sollte diese Verteilung nicht bekannt sein, so geht man von einer Gleichverteilung mit $p = 0,5$ und $q = 0,5$ aus). Es wird eine Irrtumswahrscheinlichkeit von 5% angesetzt; dies entspricht einem Signifikanzniveau von 95%; wofür ein $t = 1,96$ gilt (siehe Tabellenwerk der Statistik; bei einem Signifikanzniveau von bspw. 99% gilt $t = 2,58$). Geht man ferner davon aus, dass die Schwankungsbreite des Messwertes $\pm 3\%$ ($e = 0,03$) entsprechen soll, ergibt sich durch Einsetzen Formel eine Stichprobengröße von 1.025 Personen ($n = 1,96^2 * 0,4 * 0,6 / 0,03^2 = 1.024,43$).

4 Statistische Verfahren der Datenauswertung und Hypothesentest

(hier nur ansatzweise behandelt, ein Rück-Blick in die Statistik-Vorlesung bzw. ein zweiter Blick in Statistikbücher wird dringend empfohlen; z.B. Eckey u.a. 2005, Bamberg u.a. 2006, Ziegler 2004)

4.1 Datenauswertung

Mit statistischen Verfahren lässt sich erhobenes Datenmaterial auswerten, darstellen und auf vermutete Regelmäßigkeiten (zusammenhänge zwischen Variablen/Unterschiede zwischen Merkmalsträgern) untersuchen. Die Unterscheidung der statistischen Verfahren erfolgt i.d.R. danach, ob es sich um univariate, bivariate oder multivariate Methoden (bzw. Auswertungen) handelt. Einige wichtige Verfahren werden im Folgenden angesprochen.

4.1.1 Univariate Datenauswertung

(Verfahren zur Charakterisierung der Verteilung einzelner Merkmale)

Bei der univariaten Auswertung werden statistische Verfahren verwendet, bei denen nur eine Variable analysiert wird, d.h. es werden keine Beziehungen zwischen Variablen geprüft. Man versucht hiermit, sich einen ersten Einblick in die Struktur der Daten zu verschaffen, indem man die Verteilungseigenschaften (s. Verteilungshypothese) der einzelnen Variablen untersucht. Die Untersuchung univariater Verteilungen gilt daher nur als erster Schritt einer Analyse. Im Wesentlichen handelt es sich dabei um die Auszählung von absoluten Häufigkeiten, die Berechnung von relativen Häufigkeiten und Mittel-, Modal- oder Medianwerten mit ihren statistischen Zusatzwerten.

1. Häufigkeiten

Die absolute Häufigkeit ist die Anzahl der Werte einer Ausprägung, bei der relativen Häufigkeit erfolgt diese Angabe in Prozent. Die Ermittlung von Häufigkeiten erstreckt sich v.a. auf jene Variablen, die nominal ausgeprägt sind.

2. Arithmetisches und geometrisches Mittel

Das arithmetische Mittel (für Variablen mit zumindest Intervallcharakter) wird als Quotient aus der Summe aller beobachteten Ausprägungen und der Anzahl der Werte/Fälle n berechnet. Das geometrische Mittel (für Verhältniszahlen) ergibt sich als n -te Wurzel aus dem Produkt der einzelnen Merkmalswerte.

3. Modus

Der Modus (Modalwert) ist der häufigste Wert in einer Häufigkeitsverteilung. Für die Bestimmung genügt nominales Skalenniveau. Bsp.: {1,1,1,2,2,2,3,4,4,4,4,4, 5,5} → der Modus = 4.

4. Median

Der Median ist der „Wert in der Mitte“. In einer Beobachtungsreihe ist der Median dadurch charakterisiert, dass (mindestens) 50% der Beobachtungen größer oder gleich diesem Wert und (mindestens) 50% der Beobachtungen kleiner oder gleich diesem Wert sind. Sortiert man die Beobachtungswerte der Größe nach, so ist der Median bei einer ungeraden Anzahl von Beobachtungen der Wert der in der Mitte dieser Folge liegenden Beobachtung. Bei einer geraden Anzahl von Beobachtungen gibt es nicht ein einziges mittleres Element, sondern zwei. Hier sind die Werte der beiden mittleren Beobachtungen ein Median der Stichprobe. Voraussetzung für die Berechnung ist zumindest Ordinalität. Bsp. 1: Messwerte 1,2,4,5,13,17,19 → Median = 5. Bsp. 2: Messwerte 1,1,2,3,4,14 → Median $0,5 * (2+3) = 2,5$ (d.h. 50% der Messwerte liegen unter 2,5 und 50% darüber).

5. Streuungswerte

a) Spannweite

Die Spannweite ist definiert als die Differenz zwischen dem größten und kleinsten Messwert einer Verteilung, d.h. Maximum minus Minimum. Der Vorteil liegt in der einfachen Ermittlung, der Nachteil darin, dass die Spannweite von Extremwerten („Ausreißern“ beeinflusst wird.

b) Quartilsabstand

Die Quartile trennen die unteren und oberen 25 Prozent der Fälle einer Verteilung von den mittleren 50 Prozent der Fälle. Das (zweite) 0,5- Quartil (Q₂, Zentralwert s. Folie) halbiert die Verteilung und ist mit dem Median identisch, das 0,25-Quartil heißt unteres Quartil (Q₁), das 0,75-Quartil oberes Quartil (Q₃). Der Quartilsabstand (Strecke Q₁Q₃) ist demnach die Länge des Intervalls, das die mittleren 50 Prozent der Fälle einer Beobachtungsreihe umfasst. Dieser Wert ist stabiler als die Spannweite, weil er nicht von Extremwerten der Verteilung beeinflusst wird.

c) Standardabweichung

Die Standardabweichung ist das gebräuchlichste Streuungsmaß, da es Informationen sämtlicher Messwerte ausschöpft. Sie ist ein Maß für die Streuung der Werte einer Zufallsvariable um ihren Mittelwert. Vereinfacht gesagt, ist die Standardabweichung die durchschnittliche Entfernung aller gemessenen Ausprägungen eines Merkmals vom Durchschnitt. Die Standardabweichung für eine Stichprobe wird mit „s“ angegeben, für die Grundgesamtheit mit „σ“. Eine kleinere Standardabweichung gibt in der Regel an, dass die gemessenen Ausprägungen eines Merkmals eher eng um den Mittelwert liegen, eine größere Standardabweichung gibt eine stärkere Streuung an. Wichtig ist zu beachten, dass die Standardabweichung immer die Einheit des Merkmals trägt. Für normalverteilte Merkmale gilt, dass innerhalb der Entfernung einer Standardabweichung nach oben und unten vom Mittelwert rund 68 Prozent aller Antwortwerte liegen, im Umkreis von zwei Standardabweichungen sind es rund 95 Prozent aller Werte (Kap. 3.3).

Bsp.: Gefragt wurden 600 Studenten, wie viel Geld sie im Schnitt beim Mittagessen ausgeben. Der Mittelwert liegt bei 4,50 Euro, die Standardabweichung bei $s = 0,5$. Das heißt, dass die durchschnittliche Entfernung aller Antworten zum Mittelwert 0,50 Euro beträgt. Das Merkmal weist eine glockenartige Normalverteilung auf. Daher lässt sich ableiten, dass rund 68 Prozent aller Befragten der Stichprobe mittags zwischen 4,00 Euro und

5,00 Euro ausgeben (4,50 +/- 0,50 Euro). Rund 95 Prozent der Befragten geben zwischen 3,50 Euro und 5,50 Euro aus (4,50 +/- 2 mal 0,50 Euro).

4.1.2 Bivariate Datenauswertung

(Verfahren zur Beurteilung des Verhaltens zweier Merkmale)

Die Analyse von Beziehungen zwischen zwei Variablen, sowie die Erforschung gemeinsamer Häufigkeitsverteilungen wird als bivariate Auswertung bezeichnet. War die univariate Analyse noch rein deskriptiver Natur, die sich auf Verteilungshypothesen bezieht, so wendet sich die bivariate Analyse mehr der Überprüfung von (Zusammenhangs- bzw. Wirkungs-)Hypothesen zu. Die am meisten verwendeten bivariaten Analysemethoden sind die Kontingenzanalyse, die Korrelations- und die Regressionsanalyse.

1. Kontingenzanalyse

Die Kontingenzanalyse ist eine Analysemethode für die Abhängigkeit zwischen nominal skalierten Variablen; sie kann also feststellen, ob ein Zusammenhang zwischen den zu untersuchenden Variablen besteht oder nicht. Mit Hilfe einer Kontingenztabelle, in denen die Häufigkeitsverteilungen der Ausprägungen zweier Variablen gegeneinander aufgetragen sind, lassen sich die Größe Chi-Quadrat (quadrierte Kontingenz) oder der Kontingenzkoeffizient als Indikator für den Grad der Interdependenz (Maßzahl für die Stärke der Abhängigkeit) zwischen den Variablen berechnen. Eine wichtige Anwendung findet die Kontingenzanalyse in der schließenden Statistik bei der Überprüfung von Zusammenhangshypothesen (z.B. zur Verteilung erworbener Produkte und Geschlecht, etwa dass Frauen mehr Bio- bzw. Ökoprodukte kaufen als Männer) im sog. Kontingenz- bzw. Chi-Quadrat-Unabhängigkeitstest (Kap. 4.2.2). Beispiel siehe Vorlesung!

Kontingenzanalyse: Wie stark ist der Zusammenhang zwischen zwei nominal skalierten Variablen?

2. Varianzanalyse

Die Varianzanalyse ist ein Verfahren, das die Beeinflussung einer abhängigen metrischen Variablen (z.B. Ausgaben beim Mittagessen in Euro) durch eine (oder mehrere) unabhängige nominale Variable (z.B. Geschlecht) untersucht. Wenn auch die unabhängige Variable metrisch skaliert ist, geht die Varianzanalyse in eine Regressionsanalyse über. Beispiel s. Vorlesung!

Varianzanalyse: Wie stark unterscheiden sich gemessene Ausprägungen bestimmter Variablen und wie stark ist die Streuung von gemessenen metrischen Merkmalswerten innerhalb und zwischen Gruppen?

3. Korrelationsanalyse

Die Korrelationsanalyse überprüft die Stärke eines linearen Zusammenhangs zweier metrisch skalierten Variablen. Das Maß für den Grad der Beeinflussung ist der sog. Korrelationskoeffizient. Dieser variiert von - 1.0 bis + 1.0, womit nicht nur die Stärke des Zusammenhangs, sondern auch die Richtung angezeigt wird.

Interpretation: 0 = es besteht kein Zusammenhang (unkorreliert); bis 0,2 = sehr geringe Korrelation; bis 0,5 = geringe Korrelation; bis 0,7 = mittlere Korrelation; bis 0,9 = hohe Korrelation; über 0,9 = sehr hohe Korrelation

Häufig wird anstelle des Korrelationskoeffizienten r das Bestimmtheitsmaß r^2 angegeben (je näher r^2 an 1 liegt, desto höher ist die Wahrscheinlichkeit des linearen Zusammenhangs. Ist $r^2 = 0$, liegt kein Zusammenhang vor. Das Bestimmtheitsmaß stellt also eine Maßzahl für die Güte der Anpassung dar. Bei der Anwendung der Korrelationsanalyse ist zu berücksichtigen, dass der Wert des Koeffizienten keine Aussage über Kausalzusammenhänge (Kausalität) macht. Hierfür ist die Regressionsanalyse besser geeignet. Ein Beispiel für den Einsatz der Korrelationsanalyse im Marketing ist die Frage nach der Stärke des Zusammenhangs zwischen dem Haushaltseinkommen und den Ausgaben für Lebensmittel. Beispiel s. Vorlesung!

Korrelationsanalyse: Wie stark ist der Zusammenhang zwischen zwei ordinal oder metrisch skalierten Variablen

4. Regressionsanalyse

Die Regressionsanalyse unterstellt im Gegensatz zur Korrelationsanalyse eine eindeutige Richtung des Zusammenhangs mit einer klaren „je-desto-Beziehung“. D.h., man schaut sich z.B. an, wie sich der Absatz eines Produktes ändert, wenn die Werbung verstärkt wird. **Untersucht** wird der Zusammenhang zwischen einer abhängigen und einer (oder mehrere) unabhängigen Variablen. Voraussetzung ist eine metrische Skalierung. Zusätzlich können **Prognosen** der abhängigen Variablen gestellt werden.

- Untersuchung des Zusammenhangs: Mit Hilfe des Bestimmtheitsmaßes kann z.B. die Stärke des Zusammenhangs von Werbung auf den Absatz des Produktes angegeben werden. Durch einen Vergleich der standardisierten Regressionskoeffizienten kann man auf die relative Bedeutung der unabhängigen Variablen schließen. Es lässt sich somit feststellen, dass bspw. das Einkommen einen doppelt so starken Einfluss wie das Alter besitzt.
- Prognose der abhängigen Variablen: Man kann z.B. berechnen, um wie viel Stück der Absatz des Produktes höchstwahrscheinlich steigen wird, wenn sich bspw. die Werbeausgaben um 100 GE erhöhen. Andererseits lässt sich für jedes Unternehmen, von dem die Werbeausgaben bekannt sind, der Absatz des Produktes schätzen. In der Literatur finden man daher häufig die Anwendung der Regressionsanalysen im Zusammenhang mit Absatzprognosen im weitesten Sinn.

Die Berechnung der **Regressionsgeraden** folgt dem Schema: $y = a + bx$

y = abhängige Variable; x = unabhängige Variable; a = konstanter Wert, Residualwert (zwischen 0 und 1); b = Steigung der Geraden

Die Werte der abhängigen Variablen y und der unabhängigen Variablen x sind beobachtbar, während die Parameter der Regressionsfunktion a und b zu schätzen sind. Das Ziel der Regressionsanalyse besteht darin, die Parameter der Regressionsfunktion so zu schätzen, dass die Summe der quadrierten Residualgrößen minimiert wird. Dies wird auch als Methode der kleinsten Quadrate bezeichnet. Die Betrachtung der quadrierten Abweichungen bietet den Vorteil, dass große Abweichungen eine höhere Bedeutung zugemessen wird als kleinere Abweichungen und die Lösung der Zielfunktion algorithmisch einfach zu ermitteln ist. Beispiel s. Vorlesung!

Regressionsanalyse: Wie stark ist der funktionale Zusammenhang ($y = f(x)$) zwischen zwei metrisch skalierten Variablen, der i.d.R. als Ursache-Wirkungs- (=Kausal)-Zusammenhang interpretiert wird.

4.1.3 Multivariate Datenauswertung

(Verfahren zur Entdeckung von Strukturen)

Unter **multivariater** Auswertung versteht man die Analyse von Beziehungen zwischen drei und mehr metrischen Variablen und die Erforschung gemeinsamer Häufigkeitsverteilungen dieser Variablen. Zu den multivariaten Analyseverfahren zählen u.a. die **Diskriminanzanalyse**, die **Clusteranalyse** und die **Faktorenanalyse**. Eine Besprechung dieser Verfahren würde das vorgesehene Zeitkontingent dieser Lehrveranstaltung sprengen. Sofern eine wissenschaftliche Arbeit den Einsatz derartiger Analyseverfahren erfordern, sollten entsprechende Lehrbücher der Statistik bzw. einschlägige Publikationen (z.B. „Multivariate Analysemethoden“ von Backhaus u.a.) zu Rate gezogen werden.

4.2 Hypothesentests

4.2.1 Grundstruktur von Hypothesentests

Testverfahren gehen von einer vorliegenden Stichprobe aus. Ihre zentrale Problemstellung lautet: wie kann auf der Basis einer Stichprobe entschieden werden, ob bestimmte Hypothesen über die zugrunde liegende Grundgesamtheit richtig oder falsch sind. Ausgangspunkt eines Hypothesentests ist also zunächst die Formulierung einer statistischen Hypothese. Besondere Bedeutung in der praktischen Forschung haben sog. **Signifikanztests**.

Null- und Alternativhypothese

Ausgehend von den Stichprobenergebnissen soll durch statistische Tests, eine sog. **Nullhypothese** H_0 über bestimmte Eigenschaften in der Grundgesamtheit **überprüft** werden. Diese Eigenschaften können sich auf bestimmte Parameter (z.B. Erwartungswert), die Verteilung eines Merkmals oder die Unabhängigkeit von Merkmalen in der Grundgesamtheit beziehen. Die Nullhypothese beinhaltet den „Status quo“, d.h. eine allgemein vertretende (und bisher als richtig unterstellte) Meinung, eine Herstellerangabe, eine wissenschaftliche Arbeitshypothese etc., die richtig oder falsch sein kann. Sie soll jedenfalls geglaubt werden, bis genügend Beweise für das Gegenteil erbracht sind. Einer Nullhypothese steht daher die Alternativhypothese H_1 gegenüber, deren Inhalt genau der komplementäre Sachverhalt zu H_0 ist. Immer dann, wenn die Nullhypothese H_0 nicht zutrifft, gilt automatisch die Alternativhypothese H_1 . Nachgewiesen soll also eigentlich die Alternativhypothese, während die Nullhypothese einen Sachverhalt beinhaltet, der **bezweifelt** wird. In diesem Sinn kann man auch eine Analogie zu einem Strafprozess ziehen: Die Nullhypothese umfasst dann den Status quo, also die Unschuldsvermutung des Angeklagten. Erst wenn genügend Beweise gefunden sind und man mit großer Sicherheit die Unschuldsvermutung ablehnen kann, wird der Angeklagte verurteilt.

Bsp.: Ein Automobilwerk behauptet, dass ein von ihm gefertigter PKW einen durchschnittlichen Spritverbrauch von höchstens 7,6 l/100 km. Die Überprüfung einer Testzeitschrift anhand von 40 zufällig ausgewählten Autos ergab einen Durchschnittsverbrauch von 7,8 l/100 Km. Die Nullhypothese bezieht sich auf den Sachverhalt, der in Frage gestellt wird. Die Testzeitschrift möchte nachweisen, dass die Herstellerangabe falsch ist. Damit stehen sich zwei Hypothesen gegenüber, nämlich

- die Nullhypothese (Herstellerangabe des Automobilwerks) $H_0: \mu \leq 7,6$, die angezweifelt und überprüft werden soll, und
- die Alternativhypothese (Testzeitschrift) $H_1: \mu > 7,6$.

Im Beispiel ist hinsichtlich des Spritverbrauchs für die Testzeitschrift nur ein höherer Wert als die Herstellerangabe relevant. Kein Autofahrer würde sich über einen zu geringen

Benzinverbrauch beklagen. Damit liegt hier ein **einseitiger** Test mit $H_0: \mu \leq 7,6$. Sind hingegen für die Entscheidung über Annahme oder Ablehnung der Nullhypothese Abweichungen nicht nur nach einer, sondern zwei Richtungen relevant, spricht man von **zweiseitigen** Tests.

Anwendung in wissenschaftlichen Arbeiten: Der in der durchgeführten Forschungsstudie behauptete Zusammenhang wird als **Alternativhypothese** formuliert. Hier wird also z.B. der vermutete Zusammenhang, die unterstellte Wirkungsbeziehung oder der vermutete Unterschied zwischen den 2 Variablen(-gruppen) ausgedrückt. Die Nullhypothese hat die hierzu gegenteilige Aussage; sie besagt grundsätzlich, dass der mit der Alternativhypothese formulierte Zusammenhang oder Unterschied nicht besteht. Mit Blick auf die Methodenlehre des **Kritischen Rationalismus** wird die Nullhypothese also gebildet, um falsifiziert und damit verworfen zu werden. Ist dies der Fall, dann wird die Alternativhypothese als vorläufig bestätigt angesehen.

Da Aussagen über die Grundgesamtheit auf der Basis von Stichprobendaten mit einer Unsicherheit behaftet sind, kann nur mit einer bestimmten Wahrscheinlichkeit festgestellt werden, ob die Nullhypothese zutrifft oder nicht. Die Testentscheidung ist daher grundsätzlich immer mit einer gewissen Unsicherheit verbunden, d.h. es können Fehlentscheidungen auftreten. Zwischen der Testentscheidung und dem Wahrheitsgehalt der Nullhypothese sind folgende Kombinationen möglich (Abb. 9):

Wirklichkeit Entscheidung	Nullhypothese ist richtig	Nullhypothese ist falsch
Nullhypothese wird nicht abgelehnt	Richtige Entscheidung	Falsche Entscheidung (Fehler 2. Art)
Nullhypothese wird abgelehnt	Falsche Entscheidung (Fehler 1. Art)	Richtige Entscheidung

Abb. 9: Testentscheidung und Realität

Bei den meisten (Signifikanz-)Tests wird insbesondere darauf Wert gelegt, den sog. Fehler 1. Art, der ein fälschliches Verwerfen der Nullhypothese („Ablehnung von H_0 , wenn H_0 richtig ist,“) beinhaltet, zu **kontrollieren**. Dies geschieht durch die Forderung, dass eine derartige Fehlentscheidung lediglich mit einer als zulässig vorgegebenen kleinen **Irrtumswahrscheinlichkeit** α , dem sog. **Signifikanzniveau**, vorkommen darf. Bei einer Ablehnung der Nullhypothese H_0 ist man sich also **relativ sicher**, mit der Annahme der Alternativhypothese H_1 die richtige Entscheidung getroffen zu haben. Wie hoch die Irrtumswahrscheinlichkeit α konkret zu wählen ist, ist von der vorliegenden Fragestellung ab. Ist ein Fehler 1. Art mit hohen Konsequenzen verbunden, beispielsweise wenn es um die Sicherheit von Atomkraftwerken geht, dann wird man einen sehr geringen Wert, z.B. 0,1%, vorgeben. Bei ökonomischen Fragestellungen setzt man das Signifikanzniveau α in der Regel auf 5%.

Ist man beispielsweise bereit, in 5 von 100 Fällen einen Fehler 1. Art zu begehen, d.h. eine richtige Hypothese abzulehnen, dann beträgt das Signifikanzniveau $\alpha = 0,05$ (5%). Aussagen, denen also eine Wahrscheinlichkeit von 95% zugrunde liegt, nennt man in der Prüfstatistik **signifikant**. Erhöht man das Signifikanzniveau auf $\alpha = 0,01$ (1%), wird von **sehr signifikant** gesprochen; bei $\alpha = 0,01$ (0,1%) von **hoch signifikant**.

Abbildung 10 zeigt den Prozess der Hypothesenprüfung als Ablaufdiagramm. Hierzu noch einmal folgende Erläuterung: Nachdem die Alternativhypothese H_1 , z.B. bezogen auf einen Zusammenhang, eine Wirkung oder einen Unterschied zwischen 2 Variablen(-gruppen), eingangs aufgestellt wurde, besteht das Ziel, die Existenz dieser vermuteten Beziehung empirisch zu prüfen. Hierzu wird die komplementäre Nullhypothese H_0 herangezogen bzw. formuliert, dass als keine vermutete Beziehung existiert. Nach dem Einsatz statistischer Testverfahren zur Überprüfung der Daten auf Vereinbarkeit mit der Nullhypothese H_0 lässt sich das empirische Signifikanzniveau ermitteln. Wenn – im rechten Fall – die

Wahrscheinlichkeit dafür, dass die Daten der Nullhypothese H_0 entsprechen, größer als 0,05 (α) ist, wird die Nullhypothese H_0 angenommen bzw. beibehalten und damit die Alternativhypothese H_1 falsifiziert. Da diese Alternativhypothese H_1 dann offensichtlich keinen Erkenntniszugewinn liefert, ist sie – entsprechend dem Kritischen Rationalismus – zu modifizieren bzw. neu zu formulieren. Wenn – im linken Fall – die Irrtumswahrscheinlichkeit für die Ablehnung der Nullhypothese H_0 kleiner als 0,05 ist, dann wird die Nullhypothese H_0 abgelehnt und von einer vorläufigen Bestätigung der Alternativhypothese H_1 ausgegangen.

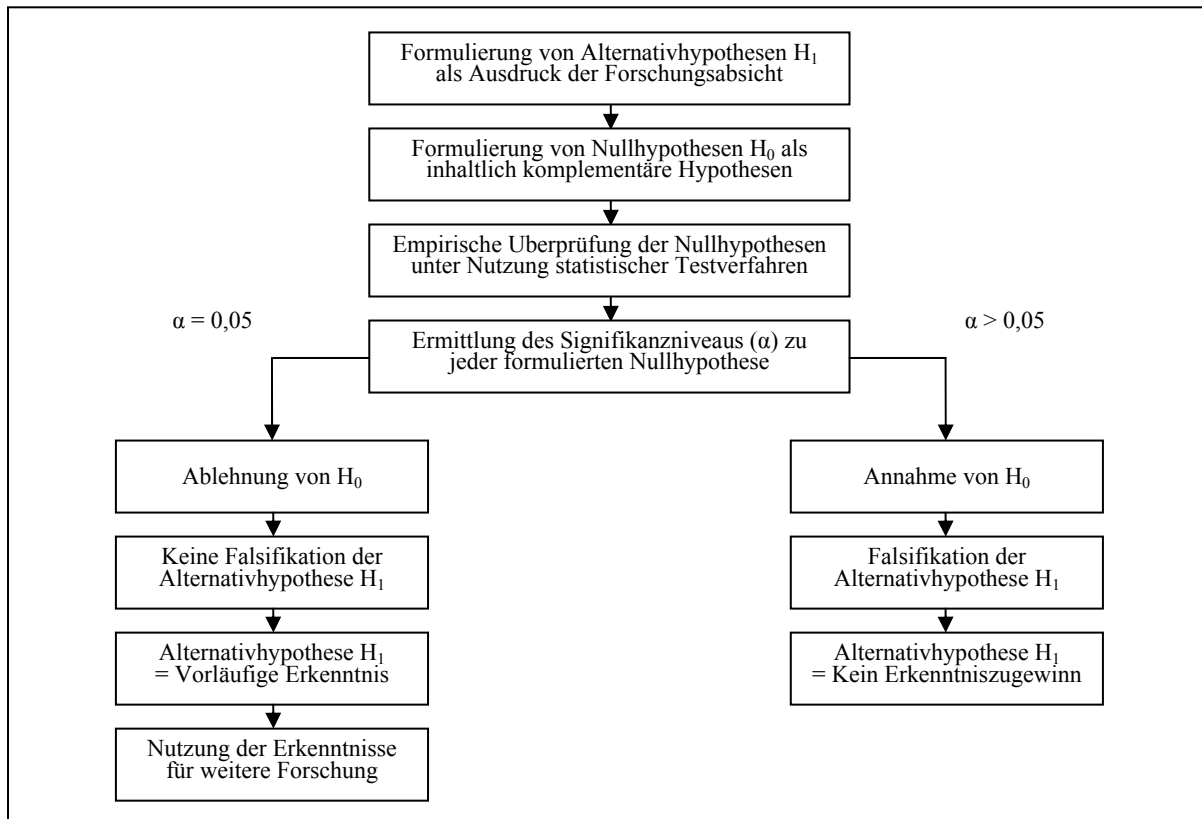


Abb. 10: Prozess der Hypothesenprüfung (Töpfer 2009/10)

Zu unterscheiden sind die statistischen Testverfahren ferner danach, ob Verteilungsannahmen für die Grundgesamtheit vorausgesetzt sind, oder nicht. Im ersten Fall spricht man von **parametrischen** Tests, im zweiten Fall von **nichtparametrischen** Tests. Einen guten Einblick in die einzelnen Arten parametrischer Test mit jeweiliger Vorgehensweise und zahlreichen Beispielen erhält man z.B. bei Eckey/Kosfeld/Türck (Wahrscheinlichkeitsrechnung und Induktive Statistik) und Bamberg/Baur (Statistik). Im folgenden Kapitel wird exemplarisch der **Chi-Quadrat-Unabhängigkeitstest** als ein Beispiel für einen nichtparametrischen Test skizziert.

4.2.2 Chi-Quadrat-Unabhängigkeitstest

Nichtparametrische Tests sind insbesondere dann relevant, wenn aufgrund kleiner Stichprobenumfänge Verteilungsannahmen nicht erfüllt sind. Ein in den Sozialwissenschaften weit verbreitetes Testverfahren ist der sog. **Chi-Quadrat-Unabhängigkeitstest** (χ^2 -Test). Mit dem Chi-Quadrat-Unabhängigkeitstest kann untersucht werden, ob zwischen zwei Merkmalen ein Zusammenhang besteht, ob also zwei Zufallsvariablen X und Y unabhängig sind oder nicht. Hierzu reicht es bereits aus, wenn nominalskalierte Merkmale vorliegen. Die Vorgehensweise bei diesem Test wird in der Vorlesung anhand eines Beispiels illustriert.