



Inputveranstaltung

Methoden der Empirischen Sozialforschung

Thomas Gautschi

Universität Mannheim

27. Januar 2010

Inhalt

- 1 Merkmale und Anliegen der Empirischen Sozialforschung
- 2 Untersuchungsablauf und Forschungsdesigns
- 3 Qualitative Verfahren und Inhaltsanalyse
- 4 Beobachtung und nicht-reaktive Verfahren
- 5 Interview und Fragebogenkonstruktion
- 6 Weitere Aspekte und Formen der Befragung

Gegenstand und Ziele I

- “Empirische Sozialforschung” umfasst die systematische und methodisch geleitete Beschaffung, Verarbeitung und Deutung von Informationen über die individuelle und soziale Wirklichkeit und deren Effekte.
- Im Mittelpunkt stehen dabei bestimmte Ziele, die zu einer Typologie von Untersuchungen in der Sozialforschung führen:
 - ▶ die methodisch geleitete Aufdeckung potentieller Zusammenhänge zwischen Merkmalen oder Variablen (explorative Untersuchungen),
 - ▶ die möglichst genaue Beschreibung von Sachverhalten, Beziehungen zwischen Variablen und Prozessen (deskriptive Untersuchungen),

Gegenstand und Ziele II

- ▶ die möglichst exakte Überprüfung von theoretischen Hypothesen über Merkmalszusammenhänge anhand von empirischen Daten, die möglichst kontrolliert gesammelt oder erhoben werden (theoriegeleitete Untersuchungen),
 - ▶ die systematische und methodisch geleitete Bewertung der Wirksamkeit von praktisch-politischen oder sozialplanerischen Massnahmen (Evaluationsstudien).
- Die Sozialforschung stellt “Werkzeuge” zur Realisierung dieser Zielsetzungen bereit. Ihre Verfahren sind nicht an Disziplinengrenzen gebunden—sie gehören zum “Handwerkszeug”. Zentral sind hierbei
 - ▶ die Sicherstellung der intersubjektiven Nachprüfbarkeit der Vorgehensweise und der gewonnenen Resultate und
 - ▶ die methodisch informierte Kontrolle von “Störfaktoren” und/oder die Entwicklung von solchen Kontrolltechniken.

Methodenvielfalt der Sozialforschung

- Die Wahl von Methoden sollte sich immer an dem jeweiligen Untersuchungsproblem orientieren (und nicht umgekehrt).
- Der Einsatz verschiedener Methoden lässt sich oftmals auch bei der Untersuchung einer einzigen Fragestellung nicht vermeiden (schliesslich ist ein „Hammer“ auch kein Universalwerkzeug!).
- Durch den Einsatz verschiedener Methoden kann man oft die Robustheit von Befunden besser erkennen (z.B. die klassische Studie “Die Arbeitslosen von Marienthal” von Marie Jahoda, Paul F. Lazarsfeld und Hans Zeisel aus den 30er Jahren).
- Welche Techniken bei welchem Erkenntnisziel einzusetzen sind ist Gegenstand der empirischen Sozialforschung.

Warum und was soll erfasst werden?

- **Beschreibung:** Was ist der Fall? Somit deskriptive Aussagen
 - ▶ In der Schweiz beträgt die Arbeitslosenquote im Juni 2008 2.3%
 - ▶ Jede dritte Ehe endet mit einer Scheidung.
- **Erklärung:** Warum ist etwas der Fall? Somit nomologische Aussagen
 - ▶ Individuen mit höherer Ausbildung erzielen auf dem Arbeitsmarkt einen höheren Lohn.
 - ▶ Das Umweltbewusstsein ist in reicheren Ländern immer höher als in ärmeren Ländern.
- **“Soziale Tatbestände”** im weitesten Sinne, z.B.
 - ▶ Verhalten und Einstellungen von Menschen
 - ▶ Eigenschaften von sozialen Systemen
 - ▶ Veränderungen von sozialen Systemen

Was soll geprüft werden?

- Die empirische Sozialforschung beschäftigt sich zu einem grossen Teil mit der Prüfung empirischer Sätze anhand geeigneter Daten.
- Dabei sind die jeweiligen “Wenn, dann” oder “Je, desto” Beziehungen mehr oder weniger präzise durch theoretische Vorüberlegungen begründet.
- Hypothesen sind empirische Sätze über Zusammenhänge von zwei oder mehr Merkmalen (Variablen), die sich nach verschiedenen Aspekten unterscheiden lassen.

Hauptphasen einer Untersuchung

Der Ablauf einer empirischen Untersuchung lässt sich grob in fünf Hauptphasen unterteilen:

- I. Präzise Formulierung der Forschungsfrage
- II. Planung und Vorbereitung der Erhebung
- III. Durchführung der Erhebung von Daten
- IV. Aufbereitung und Auswertung der Daten
- V. Zusammenfassung und Berichterstattung

Vergleiche [Anhang 1](#)

Dimensionen von Forschungsdesigns

- Nach der Formulierung des Untersuchungsproblems ist generell ein geeignetes Forschungsdesign festzulegen.
- Zumindest drei Entscheidungsdimensionen sind wesentlich:
 - Untersuchungsebene: Individual und/oder Kollektivebene.
 - Zeitlicher Erhebungsmodus: Quer- oder Längsschnitterhebung.
 - Varianzkontrolle: Nichtexperimentelles, quasi-experimentelles oder experimentelles Design.
- Diese Entscheide hängen davon ab, welche Ziele und Probleme das Forschungsvorhaben beantworten soll.
- Entscheid quantitative vs. qualitative Verfahren oder für eine Mischung aus beiden Verfahren.

Datenerhebung

- Empirische Prüfungen von Hypothesen erfordern die Beschaffung und/oder Erhebung von Daten.
- Abgesehen von der Übernahme von prozessproduzierten Daten (z.B. amtliche Statistiken, Personalakten) gibt es folgende zentrale Methoden der Datenerhebung:
 - ▶ Befragung
 - ▶ Beobachtung
 - ▶ Inhaltsanalyse
- Dabei ist die Befragung die weitaus am häufigsten angewendete Methode der Datenerhebung in der wissenschaftlichen Sozialforschung.
- Dies gilt, obwohl die Befragung eine ausgesprochen reaktive Erhebungsmethode ist.

Aspekte qualitativer Forschung

Es bestehen Unterschiede und Kontroversen zwischen der quantitativen und der qualitativen Sozialforschung. Eine genauere Diskussion betrifft u.a. die folgenden Punkte:

- Quantitative oder qualitative Forschung
- Ziele der qualitativen Sozialforschung
- Einige Formen der qualitativen Befragung

Quantitative vs. qualitative Forschung

- In der quantitativen Sozialforschung wird eine möglichst hohe Standardisierung von Fragebogen, Interviewerverhalten und Interviewsituation präferiert, weil eine möglichst hohe Objektivität, Reliabilität und Validität bei der Erhebung eine quantitative Datenanalyse begünstigt.
- Qualitative Sozialforschern kritisieren hierbei u.a. die Künstlichkeit der Interviewsituation, die mangelnde Offenheit von strukturierten Befragungen und die Dominanz der Forscherperspektive. Erhebungen im Sinne der qualitativen Forschung betonen deshalb:
 - ▶ “Sinndeutungen” von Befragten (Subjektbezogenheit)
 - ▶ Offenheit bzgl. Fragen, Antworten und Methoden
 - ▶ Untersuchungen in alltäglichen Situationen

Ziele und Formen der qualitativen Sozialforschung

- Qualitative Verfahren werden als alternative Methoden zur Erhebung und Auswertung von Daten angesehen, die nicht nur Exploration, sondern auch Hypothesenprüfung erlauben.
- Zielsetzungen sind dabei:
 - ▶ Deskription empirischer Sachverhalte und sozialer Prozesse
 - ▶ Aufstellung von Klassifikationen oder Typologien
 - ▶ Gewinnung von Hypothesen am empirischen Material
 - ▶ Prüfung von Forschungshypothesen
- In der Praxis der qualitativen Sozialforschung werden u.a. folgende Formen des Interviews verwendet:
 - ▶ Fokussiertes Interview
 - ▶ Narratives Interview
 - ▶ Problemzentriertes Interview

Probleme qualitativer Forschung

- Methodenforschung hat u.a. Probleme von quantitativen Verfahren identifiziert (z.B. soziale Erwünschtheit, Fragekontexteffekte, Interviewereinflüsse).
- Qualitative Methoden sind mit diesen Problemen zumindest genauso konfrontiert, obwohl diese eher selten thematisiert werden.
- Generell stellen sich bei qualitativen Forschungen (v.a. Befragungen) folgende Schwierigkeiten:
 - ▶ Stichprobenproblematik
 - ▶ Güte des Datenmaterials
 - ▶ Auswertungsverfahren
- Prinzipielle Lösungen für diese Probleme werden in der qualitativen Literatur zwar postuliert. Leider wird aber regelmässig auf empirische Nachweise ihrer tatsächlichen Wirkung verzichtet.

Kennzeichen der Inhaltsanalyse

- Die Inhaltsanalyse befasst sich mit der systematischen Erfassung und Auswertung von Texten (dominant!), Bildern und Filmen.
- Gelegentlich wird sie auch als Dokumentenanalyse oder Bedeutungsanalyse bezeichnet. Sie will also nicht nur ein Erhebungs-, sondern auch ein Analyseverfahren sein.
- Im einzelnen lassen sich folgende Punkte vertieft diskutieren:
 - ▶ Definition der sozialwissenschaftlichen Inhaltsanalyse
 - ▶ Anwendungsbereiche der Inhaltsanalyse
 - ▶ Zielsetzung und Kommunikationsprozess
 - ▶ Generelle Vorteile der Inhaltsanalyse

Definition der Inhaltsanalyse

- Die Inhaltsanalyse beschäftigt sich überwiegend mit der Untersuchung von sprachlichen Äußerungen.
- Im Gegensatz zur rein subjektiven Interpretation von Texten ist für die sozialwissenschaftliche Inhaltsanalyse hierbei Systematik und Objektivität kennzeichnend.

Inhaltsanalyse

Die Inhaltsanalyse ist eine empirische Methode zur systematischen, intersubjektiv nachvollziehbaren Beschreibung inhaltlicher und formaler Merkmale von Mitteilungen.

- Somit sollte die Reliabilität der Technik überprüfbar sein: Wer die gleichen explizierten Regeln bezüglich Stichprobenbildung und Auswertung verwendet, sollte analoge Resultate erzielen können!

Quantitative Inhaltsanalyse

Es gibt verschiedene Spielarten der quantitativen Inhaltsanalyse. Ganz grob lassen sich unterscheiden:

Frequenzanalyse Feststellung von Häufigkeiten von Begriffen, Themen, Moden etc. (Z.B.: “Entwicklung von Bartmoden”).

Kontingenzanalyse Aufdeckung von Assoziationsstrukturen (Z.B.: “Gibt es einen positiven Zusammenhang zwischen Freiheit und Gleichheit in sozialistischen Zeitungen?”).

Bewertungsanalyse Messung der Intensität von Bewertungen, die ein Sender zum Ausdruck bringt. (Z.B.: “Werden Politikerinnen in bestimmten Zeitschriften positiver bewertet?”)

Qualitative Inhaltsanalyse

- Die Zielsetzung der qualitativen Inhaltsanalyse ist u.a. die subjektbezogene Auswertung von Antworten auf offene Fragen in unstrukturierten Interviews.
- Ein recht bekanntes Verfahren der qualitativen Inhaltsanalyse zur subjektbezogenen Auswertung irgendeines protokollierbaren kommunikativen Aktes (z.B. sprachliche Äusserung, Bedienungsanleitung) ist die v.a. von Oevermann propagierte “objektive Hermeneutik”.
- Das Verfahren ist aber (v.a. unter quantitativen Sozialforschern) umstritten,
 - ▶ sind Objektivität und Reliabilität kaum gewährleistet,
 - ▶ ist die nötige “Naivität” nicht immer vorhanden (v.a. wenn die Erhebung und Analyse von einer Person durchgeführt wird).

Übung 1: Wertewandel

- Forschungsfrage: Analysieren Sie den Wertewandel einer Gesellschaft im Laufe des 20. Jahrhunderts.
- Vorschlag: Inhaltsanalyse
- Welcher Fragen sind zuerst zu klären?
- Welche Texte würden Sie als analysieren?

Beobachtung als Erhebungsmethode

- Beobachtungsverfahren gehören zu den ältesten Erhebungsmethoden, werden aber inzwischen weit weniger als die Befragung und die Inhaltsanalyse eingesetzt.
- Dennoch sind folgende generellen Aspekte der Beobachtung diskussionswürdig:
 - ▶ Eingrenzung der “Beobachtung”
 - ▶ Vorteile der Beobachtung
 - ▶ Nachteile der Beobachtung
 - ▶ Journalistische Beobachtungen

Eingrenzung der Beobachtung

- In der speziellen Deutung der Sozialforschung ist die Beobachtung eine Erhebungsmethode:

Beobachtung

Es handelt sich dabei um die direkte Beobachtung menschlicher Handlungen, Äusserungen, nonverbaler Reaktionen (Mimik, Gestik) und sozialer Merkmale (Kleidung, Symbole, Gebräuche, Wohnformen), die (1) unter Bezug auf Forschungshypothesen zwar immer beschränkt, aber (2) kontrolliert und systematisch erfolgt.

- Die Betonung von (1) und (2) reflektiert dabei zwei Schwierigkeiten:
 - ▶ Problem der Verzerrung durch selektive Wahrnehmung
 - ▶ Problem der (Fehl-)Interpretation des beobachteten Geschehens

Vorteile der Beobachtung

Generell hat die Beobachtung mehrere Vorzüge:

- Erfassung des tatsächlichen Verhaltens (u.U. auch unbewusstes Verhalten)
- Zugriff auf Informationen über Personen mit eingeschränktem Verbalisierungsvermögen (z.B. Einwegspiegel im Uni-Kindergarten)
- Überschaubarkeit und Verdeutlichung von sozialen Prozessen

Nachteile der Beobachtung I

Generelle Schwachpunkte der Beobachtung sind:

- Zugang zum sozialen Feld ist nicht immer einfach.
- Beeinflussung des sozialen Geschehens (z.B. "Pygmalion-Effekt") und Identifikation mit dem Untersuchungsgegenstand (bis zum "going native").
- Daten über die Verteilung von Merkmalen in einer grossen Population werden nicht erfasst (keine Totalität des Geschehens).
 - ▶ Letzteres ergibt sich insbesondere deshalb, weil nur wenig Einheiten beobachtet werden können. Dies bedeutet, dass bei Beobachtungen oft nur eine beschränkte Generalisierbarkeit besteht!
- Latente Eigenschaften von Situationen und Personen werden nicht erfasst.

Nachteile der Beobachtung II

- Vergangene und unwiederholbare Ereignisse werden von vornherein vernachlässigt.

Was kann man dagegen tun?

- Schulung von Beobachtern
- Einsatz mehrerer Beobachter
- Leitfaden und strukturiertes Beobachtungsschema
- Test auf Reliabilität und Validität

Varianten der Beobachtungstechnik

- Nach bestimmten Kriterien können verschiedene Arten von Beobachtungen unterschieden werden:
 - ▶ Fremdbeobachtung oder Selbstbeobachtung (Introspektion)
 - ▶ Teilnehmende oder nicht teilnehmende Beobachtung
 - ▶ Offene oder verdeckte Beobachtung
 - ▶ Feldbeobachtung oder Laborbeobachtung
 - ▶ Unstrukturierte oder strukturierte Beobachtung
- Diese Formen gehen mit unterschiedlichen positiven und negativen Aspekten einher, die jeweils diskussionswürdig sind.
- Am häufigsten sind die Kombinationen:
 - ▶ Teilnehmende, offene, nicht strukturierte Fremdbeobachtung im Feld
 - ▶ Nichtteilnehmende, verdeckte, strukturierte Fremdbeobachtung im Labor.

Reaktivität und Nicht-Reaktivität

- Erhebungsmethoden heißen reaktiv, wenn die Gefahr besteht, dass der Messvorgang (d.h. das Messinstrument, sein Anwender oder die Messsituation) das Messergebnis beeinflussen und verfälschen kann (aber nicht muss).
- Als reaktiv kann man z.B. die Befragung bezeichnen.
- Dagegen können Formen der Inhaltsanalyse und Beobachtung nicht-reaktiv sein.
- So geht z.B. mit der Analyse von Dokumenten und z.B. der verdeckten Beobachtung sicher keine systematische Verzerrung des Messvorgangs einher.
- Grundidee: Reaktivität wird durch das Wissen der Beteiligten über die Messung erzeugt (kein Wissen, keine Reaktivität).

Nicht-reaktive Verfahren

- Nicht-reaktive Methoden sind keine neuen Klasse von Erhebungsverfahren. Reaktivität oder Nicht-Reaktivität bezeichnet vielmehr eine Eigenschaft unterschiedlicher Methoden.
- Zwei Möglichkeiten: Die Messung erfolgt verdeckt oder die Messung erfolgt ex ante durch Interpretation bereits existierender Daten.
- Die Gewährleistung des “Nichtwissens” ist für unterschiedliche Erhebungsmethoden unterschiedlich gut zu erreichen.
- Arten nicht-reaktiver Verfahren:
 - ▶ Feldexperimente
 - ▶ Unaufdringliche Beobachtung
 - ▶ Verhaltensspuren
 - ▶ Sekundäranalyse prozessproduzierter Daten
 - ▶ Inhaltsanalyse

Verhaltensspuren

- Menschen hinterlassen fast überall Spuren und diese lassen Rückschlüsse auf das Verhalten und ihre Präferenzen zu.
- Den Ausdruck “Verhaltensspuren” kann man als Oberbegriff für die Erfassung und Auswertung von prozessproduzierten Daten und Dokumenten auffassen.
- Beispiele für die Analyse von Verhaltensspuren betreffen u.a.:
 - ▶ Pizzabestellung im Weissen Haus vor dem Golfkrieg
 - ▶ Abnutzung und Abdrücke: Lese-, Seh- und Hörgewohnheiten
 - ▶ Untersuchung von Abfällen zur Konsumfeststellung
 - ▶ Verkaufsziffern von Kondomen als Reaktion auf AIDS
 - ▶ Archivierte Daten (Sozialversicherung, Personalbüro)

Prozessproduzierte Daten und ihre Probleme

- Prozessproduzierte Daten: Daten die im Rahmen sozialer (insbesondere bürokratischer) Prozesse automatisch ohne Eingriff (Planung, Messung etc.) des Forschers erhoben werden.
- Problem: Datenerhebung wird nicht durch den Forscher durchgeführt.
 - ▶ Reliabilität: Zuverlässigkeit der Buchführung oder Personalkarteien?
 - ▶ Validität: Wie aussagekräftig ist z.B. die Steuerklasse für die Familiensituation?
 - ▶ Zugänglichkeit: Prozessproduzierte Daten sind häufig geschützt.
 - ▶ Informationsbeschränkung: Meist enthalten prozessproduzierte Daten nur wenige Informationen

Probleme nicht-reaktiver Methoden I

- Nicht-Reaktivität ist eine wünschenswerte Eigenschaft, dennoch ergeben sich aber auch Probleme.
- Aufgrund der Heterogenität nicht-reaktiver Methoden stellen sie sich für jedes Verfahren etwas anders dar.
- Beispiele für potentielle Schwierigkeiten sind:
 - ▶ Was ist die Stichprobe (z.B. verlorene Briefe, Abfallforschung)?
 - ▶ Sind die Indikatoren gültig (z.B. hat rückgesandter "verlorener" Brief etwas mit Sympathie gegenüber Adressat zu tun)?
 - ▶ Wie reliabel sind die einzelnen Verfahren (z.B. was ist mit "Technik der verlorenen Briefe" bei Regen? Versetzt der Fund einer Münze die Person wirklich in eine bessere Stimmung?)?

Probleme nicht-reaktiver Methoden II

- ▶ Wie valide sind die einzelnen Verfahren (z.B. ist der Autotyp ein guter Indikator für Status)?
- ▶ Welche ethischen Probleme ergeben sich (z.B. bei Täuschungsverfahren)? Entsteht der involvierten Person ein unzumutbarer Schaden?
- Generell empfiehlt sich, wenn möglich, jeweils ein Methodenmix (d.h. Kombination verschiedener Verfahren und ihrer jeweiligen Vorteile).

Vorteile nicht-reaktiver Methoden

- Natürlich: keine Verzerrung durch Reaktivität.
- Hohe interne Validität: gute Kontrollierbarkeit des Stimulus, insbesondere, wenn eine Randomisierung möglich ist (wie im Laborexperiment).
- Sehr hohe externe Validität (logisch: Realitätsnähe).

Grundformen der Befragung

- Befragungen können nach verschiedenen Unterscheidungsmerkmalen klassifiziert werden. Die wesentlichen Kriterien der Differenzierung sind:
 - ▶ Art der Kommunikation bei der Befragung
 - ▶ Grad der Standardisierung der Befragung
 - ▶ Teilnehmerzahl in der Befragungssituation
- Somit gibt es für Befragungen generell zwei Möglichkeiten:
 - ▶ Quantitativ, mit grossen Fallzahlen und der Möglichkeit der Beschreibung der Population mittels Inferenzstatistik
 - ▶ Qualitativ, explorativ mit kleinen Fallzahlen

Unterscheidung nach Kommunikation

- Nach der Art der Kommunikation können drei Typen von Befragungen unterschieden werden:
 - ▶ Persönliches Interview
 - ▶ Telephonisches Interview
 - ▶ Schriftliche Befragung
- Das persönliche Interview wird oftmals auch als “Face-to-Face” Interview bezeichnet.

Unterscheidung nach Standardisierung

- Befragungen können auch nach dem Grad der Standardisierung differenziert werden.
- Dabei wird ein Kontinuum mit den Polen “vollständig strukturiert” und “vollständig unstrukturiert” (“offen”) zur Unterscheidung verwendet:
 - ▶ Ein vollständig strukturiertes Interview liegt vor, wenn
 - (a) alle Fragen mit
 - (b) vorgegebenen Antwortkategorien in
 - (c) festgelegter Reihenfolge gestellt werden.
 - ▶ Ein vollständig unstrukturiertes Interview liegt vor, wenn nur das Thema der Befragung vorgegeben wird und alles weitere dem Gesprächsverlauf überlassen bleibt.
- Bei stark (wenig) strukturierten Befragungen spricht man auch von quantitativen (qualitativen) Befragungen.

Grad der Strukturierung I

- Es gibt ein Kontinuum zwischen den Extremen. Beispiel:
 - ▶ Wenn am nächsten Sonntag Wahlen wären, würden Sie Partei XY wählen?
 - Ja
 - Nein
 - Weiss nicht
 - ▶ Bitte erzählen Sie mir doch, was Sie von Partei XY und deren Politik halten.
- Der Grad der Strukturierung birgt einige Vor- aber auch Nachteile.
- Je höher die Strukturierung, desto eher sind Objektivität und Reliabilität erfüllt.

Grad der Strukturierung II

- Bei unstrukturierten Interviews hängt das Ergebnis stärker vom Interviewer ab, geringere Durchführungsobjektivität und unter Umständen auch geringere Auswertungsobjektivität.
- Nachteil der Strukturierung: Antwortmöglichkeiten müssen alle bekannt sein (“erschöpfend”), sonst bleiben wichtige Optionen unberücksichtigt.
- Möglichkeit: Man lässt eine offene Kategorie “anderes, nämlich” zu.
 - ▶ Bitte geben Sie an, mit welchem Verkehrsmittel Sie in die letzten Ferien gefahren sind.
 - Auto
 - Bus
 - Zug
 - Flugzeug
 - Anderes, nämlich: _____

Befragung/Interview als soziale Interaktion

- Ein Interview ist eine soziale Interaktion zwischen Befragtem und Interviewer.
- Es gibt Unterschiede zwischen Alltagsgesprächen und Interviews:
 - ▶ Anonymität der Informationen und Personen,
 - ▶ Einmaligkeit der Befragungssituation sowie
 - ▶ fehlende Sanktionierung durch Interviewer
- Wichtig: Der Befragte ist ein Akteur mit eigenen Interessen, Vorstellungen und Handlungsalternativen.
- Ziel des Interviews: Informationen sollen möglichst neutral und ohne Beeinflussung erhoben werden.

Voraussetzung bei der Befragung/Interview

- Die Erhebung unverzerrter Daten durch Interviews erfordert mindestens drei Voraussetzungen.
- Die Kooperationsbereitschaft (z.B. durch Akzeptanz und Interesse an der Untersuchung) seitens der befragten Person muss vorhanden sein.
- Der Befragte muss die Norm, wahrheitsgetreu zu antworten, erfüllen (z.B. Gefahr der sozialen Erwünschtheit).
- Befragter und Interviewer müssen eine "gemeinsame Sprache" sprechen, der Fragebogen muss also dermassen gestaltet sein, dass Fragen und Antwortkategorien allgemein verständlich sind (Kultur).

Klassifikation von Fragen

- Fragen lassen sich nach verschiedenen Kriterien unterscheiden.
- Zentrale Differenzierungen betreffen
 - ▶ Inhaltliche Schwerpunkte von Fragen
 - ▶ Personenbezug von Fragen
 - ▶ Art des Bezugs auf die Zieldimension
 - ▶ Frageformen nach Antwortmöglichkeit
 - ▶ Fragerollen im Befragungskontext
- Diese Unterscheidungsmerkmale führen zu weiteren Differenzierungen und damit einhergehenden Empfehlungen (vgl. die folgenden Folien).

Inhaltliche Fragedifferenzierung

- Fragen beziehen sich auf unterschiedliche Sachverhalte. Die wichtigsten Frageinhalte betreffen
 - Einstellungen
 - Überzeugungen
 - Verhaltensweisen
 - Sozialstatistik
- Es ist klar, dass der inhaltliche Schwerpunkt die jeweilige Fragetechnik vorstrukturiert.
- Ausgehend von der obigen Klassifikation können Fragetypen unterschieden werden!

Differenzierung nach Personenbezug

- Unterscheidet man Fragen danach, ob ein Bezug auf die befragte Person vorliegt oder nicht, so ergibt sich die Klassifikation nach **Selbstauskünften**, z.B. Frage nach eigenem Erwerbseinkommen **Fremdauskünften**, z.B. Frage nach dem Erwerbseinkommen des Ehepartners
- Im Regelfall gilt: Fremdauskünfte sind weniger zuverlässig als Selbstauskünfte!

Unterscheidung nach Zielbezug

- Man kann Fragen danach klassifizieren, ob ein direkter oder indirekter Bezug auf die relevante Zieldimension vorliegt.
- Die Zielsetzung indirekter Fragen ist die Erfassung verborgener Motive (d.h. ohne dass Befragte dies merken).
- Vermeidung zu starker Reaktivität der Befragung.

Unterscheidung nach Frageform

- Differenziert man nach der Form der Fragen, so können die folgenden hauptsächlichsten Fragetypen unterschieden werden:
 - ▶ Geschlossene (Vorgabe der Antwortmöglichkeiten), offene (freie Antwort möglich) oder halboffene Fragen (Hybridfragen);
 - ▶ Geschlossene Fragen nach Art der Antwortkategorien (ja/nein, Rating oder Ranking, Mehrfachantworten).
- Geschlossene Fragen sind eigentlich die Regel.
- Es besteht eine hohe Vergleichbarkeit sowie eine hohe Durchführungs- und Auswertungsobjektivität.

Unterscheidung nach der Fragenrolle

- Klassifiziert man nach der Rolle von Fragen im Befragungskontext, so gibt es neben den “normalen” Fragen insbesondere folgende Typen:
 - “Filterfragen”: leitet nur Betroffene zu entsprechenden Fragen oder Frageblöcken.
 - “Gabelfragen”: leitet nur Betroffene zu entsprechenden Fragen oder Frageblöcken.
 - “Fragetrichter”: v.a. bei heiklen Themen, beginnend mit einfachen/harmlosen Fragen bis zu schwierigen/heiklen Fragen.
 - “Eisbrecherfragen”: Einfache/Harmlose Fragen zu Beginn des Interviews oder eines Frageblocks.
- Diese Fragetypen dienen der Steuerung des Interviews.

Einige Formulieringsregeln I

Folgende Grundregeln sind bei der Frageformulierung hilfreich:

- Formulierung von kurzen, verständlichen und präzisen Fragen.
- Verzicht auf platte Anbietierungen.
- Vermeidung von doppelten Verneinungen.
- Wahl der Antwortkategorien: Bei geschlossenen Fragen sollten sie disjunkt (d.h. nicht überlappend), erschöpfend und präzise sein. Bei Fragen nach der Häufigkeit, Dauer und anderen Zahlenwerten sollten sie offen sein.
- Verzicht auf stark wertbesetzte Begriffe (z.B. Gerechtigkeit, Freiheit)

Einige Formulieringsregeln II

- Keine mehrdimensionalen Fragen (z.B. nicht Zustimmung oder Ablehnung auf “Alkoholgenuss ist gesellschaftsfähig, aber problematisch”, sondern zwei getrennte Fragen).
- Keine Suggestivfragen, welche sozial erwünschte Antworten hervorrufen resp. die Antworten in eine bestimmte Richtung lenken können (z.B. nicht “Meinen Sie auch, dass übermässiger Alkoholgenuss gesellschaftlich problematisch ist”, sondern neutrale Fragestellung).

Weitere Formulieringsregeln I

Zudem sind folgende Grundregeln bei der Frageformulierung zu beachten:

- Vermeidung von nicht erprobten indirekten Fragen (wegen des Problems der uneindeutigen Zurechenbarkeit).
- Polung der Antwortkategorien in unterschiedliche Richtungen bei Fragebatterien (z.B. bei Einstellungsmessung)
- Keine Überforderung von Befragten (z.B. keine Rechenaufgaben)
- Neue Themen sollten mit einer Überleitung eingeleitet werden.

Weitere Formulierungsregeln II

- Sozialdemographische Angaben am Schluss des Fragebogens erfragen.
- Interviews sollten eine vernünftige Länge haben, welche abhängig von der Befragungsform ist:
 - ▶ Persönliche Interviews können bis 90 Minuten dauern.
 - ▶ Telefonische Interviews können bis etwa 40 Minuten dauern.
 - ▶ Schriftliche Fragebögen sollten höchstens 30 Minuten Bearbeitungszeit in Anspruch nehmen.

ABER: Keine Regel ohne Ausnahme!

Entwicklung des Fragebogens

- Nachdem das Untersuchungsziel einer Befragungsstudie bekannt ist, sollte man folgendermassen vorgehen:
 - ▶ Festlegung thematischer Blöcke (Module)
 - ▶ Formulierung von geeigneten Fragen
 - ▶ Anordnung der Fragen nach bestimmten Regeln
 - ▶ Durchführung und Auswertung eines Vortests (Pretest) mit Personen aus der Grundgesamtheit.
- Die Ergebnisse des Pretests führen in der Regel zu Korrekturen und damit zu einer Wiederholung der erwähnten Schritte.
- Erst wenn sich ein Fragebogen in einem Pretest weitgehend bewährt hat, sollte man mit der tatsächlichen Befragung beginnen!

Übung 2: Fragebogenkonstruktion

- Forschungsfrage: Analysieren Sie das Umweltbewusstsein, das Umweltwissen und das Umweltverhalten der Schweizer Bevölkerung.
- Welche thematischen Blöcke schlagen Sie vor?
- Wie ist die Abfolge der Blöcke im Fragebogen?
- Formulieren Sie pro Block ein bis zwei geeignete Fragen mit entsprechenden Antwortkategorien
- Welche Angaben zu Soziodemographie sind unabdingbar?
- Obwohl nicht besprochen, wie sieht die Grundgesamtheit und Ihr Stichprobendesign aus?

Durchführung von Befragungen

- Es gibt eine Reihe von Fragen, die für die Durchführung einer Befragungsstudie zu beantworten sind:
 - ▶ Wie werden die Befragungen durchgeführt (z.B. telefonische oder persönliche Interviews mit standardisiertem oder teilstandardisiertem Fragebogen)?
 - ▶ Wie werden die Befragten ausgewählt (z.B. Zufallstichprobe vs. Quotenauswahl, Personenwahl per Schwedenschlüssel oder mit Geburtagsmethode)?
 - ▶ Werden die Zielpersonen vorher angeschrieben und, falls ja, wie wird das Befragungsanliegen im Anschreiben formuliert?
 - ▶ Wird die Umfrage eigens für einen Auftraggeber durchgeführt oder in eine Mehrthemenbefragung (Omnibus-Befragung) eingeschoben?

Weitere Durchführungsfragen

- Wie lange ist das Interview insgesamt?
- Wie und wo werden die Interviewer geschult?
- Wie viele Interviews führt ein Interviewer durch (Ermüdung, Fälschung)?
 - ▶ Face-to-Face: max. 10–20 Interviews
- Wie ist die Höhe und Art der Honorierung der Interviewer?
- Wie werden die Interviews kontrolliert (Minimalstandard: Telephonischer Rückruf bei 20 bis 40% der Befragten)?
- Wie viele Kontaktversuche sind vorgeschrieben, um eine Zielperson zu erreichen (3 bis 5 Kontaktversuche bei persönlichen Interviews sind üblich)?
- Welche bisherigen Umfragen zur Thematik wurden wie und mit welchen Resultaten durchgeführt?

Fehlerquellen im Interview

- Abgesehen von Fälschungen und etwaigen Auswahlfehlern lassen sich die folgenden Fehlerquellen bei Interviews unterscheiden:
 - ▶ Merkmale der Befragten (Soziale Erwünschtheit, Response-Set, Meinungslosigkeit)
 - ▶ Merkmale der Frage (Effekte der Formulierung, Positionierung und Antwortkategorien)
 - ▶ Merkmale des Interviewers und der Befragungssituation (Interviewereffekte, Anwesenheit Dritter, weitere Situationseinflüsse, Sponsorship-Effekt)
- Ausgangspunkt vieler Fehlerquellen ist die Vorstellung, dass eine befragte Person einen “wahren” Wert der jeweils interessierenden Variable hat, ihre Antwort jedoch nicht nur von diesem davon abhängt.

Telephonische Befragung I

- In den letzten Jahrzehnten hat sich die telephonische Befragung, nach anfänglicher Skepsis der Methodenforscher, in den Industrieländern immer mehr ausgebreitet.
- Zwei Drittel aller Umfragen werden mittlerweile telefonisch durchgeführt.
- Die telephonische Befragung bietet eine Reihe von Vorteilen, z.B.:
 - ▶ Leichte Stichprobenziehung (z.B. RDD bei unvollständigen Telefonregistern)
 - ▶ Geringe Interviewkosten dank Computerassistiertem Telephoninterview (CATI)
 - ▶ Hohe Kontrolle der Interviewer
 - ▶ Hohe Anonymität und damit gering(ere) soziale Erwünschtheit
 - ▶ Hohe Standardisierung

Telephonische Befragung II

- Weitere Vorteile einer telefonischen Befragung sind:
 - ▶ Hohe Ausschöpfungsquoten von bis zu 70% u.a. wegen grosser Anzahl Kontaktversuche
 - ▶ Optimale Befragungszeit von rund 30 Minuten erlaubt etwa 100 Fragen
 - ▶ Methodische Studien zeigen keine wesentlichen Qualitätsunterschiede zwischen face-to-face und telefonischer Befragung.
 - ▶ Filterfragen sind in telefonischen Interviews dank CATI besonders einfach zu bewältigen.
 - ▶ CATI erlaubt Konsistenztests während des Interviews

Schriftliche Befragung I

- In gewissen Fällen sind schriftliche Befragungen aber besser geeignet als telefonische Umfragen.
- Schriftliche Befragungen haben u.U. gewisse Vorteile:
 - ▶ Befragte können Fragen besser durchdenken
 - ▶ Interviewereffekte treten nicht auf, höhere Anonymität und damit geringere soziale Erwünschtheit
 - ▶ Aufwand und Kosten sind geringer
- Die Nachteile von schriftlichen Befragungen wiegen aber dessen Vorteile vielfach nicht auf:
 - ▶ Interviewer kann bei Verständnisproblemen nicht helfen, keine Rückfragen möglich.
 - ▶ Fragebogen muss relativ einfach und selbsterklärend gestaltet sein.

Schriftliche Befragung II

- ▶ Filterfragen stellen Befragte oft vor Probleme.
- ▶ Es ist unsicher, wer den Fragebogen ausgefüllt hat.
- ▶ Aufwendige Stichprobenziehung (z.B. Random Route) ist oft unumgänglich
- ▶ Rücklaufquote (Ausschöpfungsquote) muss durch zusätzliche Massnahmen (Nachfassaktionen) erhöht werden, wobei hier einige methodische Erfahrungen (z.B. Versand eines Vorbrieffs nur bei einziger Nachfassaktion, bei mehreren Erinnerungen ist Vorbrief entbehrlich; kleinere Belohnungen erhöhen Rücklauf) vorliegen.

Face-to-face Befragung

- Vorteile von face-to-face Befragungen:
 - ▶ Interviewer kann nachfragen, Anwesenheit Dritter kontrollieren und falls notwendig eigene Einschätzungen über die Situation vornehmen.
 - ▶ Verwendung von Bildmaterial o.ä. ist möglich.
 - ▶ Komplexe Fragen können gestellt und erläutert werden.
 - ▶ Es sind Anonymisierungsverfahren (z.B. RRT, Antwortkuverts) einsetzbar.
- Nachteile von face-to-face Befragungen:
 - ▶ Interviews sind teuer
 - ▶ Feldkontrolle ist sehr aufwendig
 - ▶ Anfälliger für Effekte der sozialen Erwünschtheit

Online Befragung I

- Formen der Onlinebefragung sind die Befragung per eMail oder die Online Befragung im Internet.
- Vorteile von Online Befragungen:
 - ▶ Geringe Kosten da sowohl Druck als auch Versandkosten entfallen.
 - ▶ Arbeitsschritte der nachträglichen Kodierung und Dateneingabe entfallen.
 - ▶ Programmierte Filterführung und Konsistenzprüfung sind bei einer Internetbefragung (wie bei CATI) möglich.
 - ▶ Flexibilität der graphischen Gestaltung.
- Nachteile von Online Befragungen:
 - ▶ Geringe Verbreitung (60% der Haushalte) und damit selten oder nur bei "verzerrter" Population einsetzbar.

Online Befragung II

- ▶ Bei besonderen Stichproben aber (z.B. Unternehmensstudie, Hochschulen) durchaus verwendbar.
- ▶ Datensicherheit (Übertragungssicherheit) wird von den Befragten kritisch beurteilt.
- ▶ Geringe Ausschöpfung da z.B. weniger verpflichtenden Charakter.
- ▶ Gewisse Fragentypen (z.B. Wissensfragen) können weniger gut gestellt werden.
- ▶ Das grösste Problem ist jedoch häufig, dass die Grundgesamt unbekannt ist.

Qualitative Befragung I

- Häufige Zielsetzungen qualitativer Befragungen
 - ▶ explorativ beziehungsweise theoriebildend
 - ▶ Subjektperspektive / Sinndeutung
- Anforderung an die qualitative Datenerhebung
 - ▶ Subjektbezogenheit
 - ▶ Offenheit bezüglich der Fragen, Antworten und Methoden
 - ▶ Untersuchung in alltäglichen Situationen
- Die qualitative Befragung kennt drei Interviewtypen
 - ▶ Das fokussierte Interview: Bezug auf eine konkrete Situation, die alle Befragten gemeinsam haben, Herausarbeitung der wesentlichen Merkmale.

Qualitative Befragung II

- ▶ Das narrative Interview: Weitgehend unstrukturiertes Interview, offene Erzählung auf einen Stimulus (Themensetzung), am Ende evtl. Nachfragen durch den Interviewer.
- ▶ Problemzentriertes Interview: Einsatz eines Leitfadens (→ stärkere Strukturierung), Steuerung durch den Interviewer.
- Wie jede Befragung, hat auch die qualitative Befragung Nachteile
 - ▶ Auswahl der Stichprobe: meist bewusst, quasi-experimentell (z.B. 2x2 Design). Verzerrungen, Verallgemeinerbarkeit?
 - ▶ Reliabilität und Validität: Abschätzung wird meist nicht vorgenommen.
 - ▶ Datenauswertung: Beruht meist auf Interpretationsverfahren der Texte. Durchführungsobjektivität?
- Aber, häufig ist der kombinierte Einsatz beider Methodentypen sinnvoll.
 - ▶ Die Arbeitslosen von Marienthal (Jahoda et al. 1960)

Nicht- und vorexperimentelle Designs I

- Ziel vieler Analysen ist die Messung von Unterschieden (Varianz) und der Zuordnung der Unterschiede zu kausalen Einflussfaktoren (“causal inference”).
- Eine solche Varianzkontrolle ist kann am besten über ein experimentelles Design erreicht werden. Dabei unterscheiden wir verschiedene “Arten von Experimenten”.
- **Nichtexperimentelle Studien** sind z.B. für die Umfrageforschung charakteristisch.
- Im Gegensatz zur experimentellen Forschung werden bei diesen ex-post-facto-Designs üblicherweise Vergleichs- und Kontrollgruppen nicht explizit vor der Erhebung gebildet.

Nicht- und vorexperimentelle Designs II

- Eine Gruppenbildung erfolgt bei ihnen regelmässig erst im Rahmen der Datenanalyse auf nicht-zufällige Weise.
- Von den nichtexperimentellen Designs sind die **vorexperimentellen Designs** zu unterscheiden.
- Zu ihrer Kennzeichnung ist die **experimentelle Standardnotation** einzuführen:
 - ▶ Üblicherweise wird die experimentelle Manipulation (Stimulus) mit X und
 - ▶ die Messung oder Beobachtung (Observation) mit O bezeichnet,
 - ▶ wobei die Sequenz der Symbole ihre zeitliche Abfolge ausdrückt.

Nicht- und vorexperimentelle Designs III

- Vorexperimentelle Designs sind dann die Anordnungen
 - ▶ $X O$ (Nachher-Messung)
 - ▶ $X O_{\text{const}}$ (Design falscher Vergleichswerte) und
 - ▶ $O_1 X O_2$ (Vorher-Nachher-Messung).
- Insgesamt lassen sich zwei wesentliche Probleme solcher Designs identifizieren:
 - 1 Keine Beobachtungen und daher fehlende Varianz in der unabhängigen oder der abhängigen Variablen (“Design falscher Vergleichswerte”).
 - 2 Keine Möglichkeit einer Identifikation der (v.a. in den Human- und Sozialwissenschaften wesentlichen) “Reifung” zwischen den Messungen.

Experimentelle und quasi-experimentelle Designs

Von einem experimentellen Design wird gesprochen, wenn

- mindestens zwei Gruppen vor der Erhebung gebildet werden,
- die Versuchspersonen den Gruppen nach einem Zufallsverfahren (Randomisierung) zugewiesen werden und
- die unabhängige Variable durch den Forscher manipuliert wird (Stimulus).
- Quasi-experimentelle Designs sind Versuchsanordnungen, die dem Vorbild des Experiments nahekommen und der experimentellen Logik folgen, jedoch nicht alle Anforderungen (idR keine Randomisierung) eines experimentellen Designs erfüllen.
- **Vergleiche** Anhang 2

Vor- und Nachteile von Experimenten I

- Der experimentelle Stimulus wird im Experiment produziert und geht der vermuteten Wirkung zeitlich voraus.
- Durch die Randomisierung werden verzerrende Effekte durch Drittvariablen neutralisiert (Gewährleistung der “internen Validität”).
- Insbesondere wenn es sich um Laborexperimente handelt, stellt sich die Frage,
 - ▶ ob die experimentelle Befunde generalisiert werden können (Problem der “externen Validität”).
 - ▶ wie gut die Isolierung des Stimulus funktioniert (Problem der “internen Validität”).
- Es ist stets zu bedenken, dass Experimente durch Reaktivität verzerrt sein können.

Vor- und Nachteile von Experimenten II

- Es kann sehr aufwendig sein, komplexe Zusammenhänge experimentell zu prüfen (adäquate Anreize, genügend Lernzeit für Probanden).
- Bei der Untersuchung sozialer Zusammenhänge ist die Randomisierung aus ethischen oder praktischen Gründen oft unmöglich.
- Bei kleinen Fallzahlen kann eine ungleiche Verteilung der Drittvariablen trotz Randomisierung auftreten.

Zum Begriff der "Operationalisierung" I

- Empirische Prüfungen von Hypothesen erfordern die Zuordnung von beobachtbaren Merkmalen (Indikatoren) zu den theoretischen Begriffen.
- Dies ist oft unproblematisch (z.B. Geschlecht), v.a. bei einer 1:1 Beziehung zwischen Indikator und theoretischem Begriff.
- In vielen Fällen ist es aber nicht immer ganz einfach, v.a. bei der Prüfung von Hypothesen
 - ▶ mit bestenfalls indirekt beobachtbaren oder
 - ▶ unscharfen vieldeutigen Konstrukten.

Zum Begriff der “Operationalisierung” II

- **Operationalisierung** ist definiert als hinreichend genaue Anweisungen, wie die in einem theoretischen Begriff spezifizierten Eigenschaften beobachtbaren Sachverhalten zugeordnet werden sollen.
- Messung, Skalierung und Indexbildung sind somit Spezialfälle der Operationalisierung.
- **Vergleiche** Anhang 3

Objektivität, Reliabilität und Validität I

- Messungen sollten objektiv (wiederholbar), reliabel (zuverlässig) und valide (gültig) sein.
 - ▶ Der Grad der Objektivität eines Messinstrumentes drückt aus, in welchem Ausmass die Messergebnisse unabhängig sind von der jeweiligen Person, die das Messinstrument anwendet.
 - ▶ Die Reliabilität eines Messinstruments drückt aus, inwieweit Messergebnisse reproduzierbar sind.
 - ▶ Ein Messinstrument ist in dem Ausmass valide, in dem es das zu messende Merkmal tatsächlich misst.
- Diese Gütekriterien der Messung können auf Erfüllung geprüft werden.
- Dies geschieht sehr oft auf Basis von Korrelationsanalysen.
- Objektivität, Reliabilität und Validität beziehen sich allesamt auf das Messinstrument und seinen Einsatz.

Objektivität, Reliabilität und Validität II

- Die drei Kriterien stehen in einem hierarchischen Verhältnis:
 - ▶ Objektivität ist eine notwendige, aber keine hinreichende Bedingung für Reliabilität.
 - ▶ Reliabilität ist eine notwendige, aber keine hinreichende Bedingung für Validität.
- Objektivität und Reliabilität sind somit notwendige Minimalanforderungen an ein Messinstrument und dessen Einsatz, das Hauptziel einer Messung ist jedoch deren Validität!

Messung und Skalentypen von Variablen I

- Es besteht eine Hierarchie der Skalenniveaus von Variablen in dem Sinn, dass mit einem höhergestelltem Messniveau zwar immer auf ein tiefergestelltes Messniveau, aber prinzipiell nicht auf ein noch höheres Messniveau überführt werden kann.
- Die kumulative Hierarchie der Skalentypen (vom untersten zum obersten Messniveau) für Variablen lautet:
 - ▶ Nominalskala
 - ▶ Ordinalskala
 - ▶ Intervallskala
 - ▶ Ratioskala
 - ▶ Absolutskala

Messung und Skalentypen von Variablen II

- Die untersten beiden Skalenniveaus nennt man **kategorial**, die obersten drei **metrisch**.
- Je höher das Skalenniveau, desto informativer sind die Messwerte, aber desto weniger umfangreich ist die Klasse der erlaubten Transformationen für die Variable.
- Die Messung erlaubt Aussagen über die Relation der empirischen Relationen von Objekten (Unterscheidbarkeit, Rangordnung, spezifizierbar konstante Abstände, natürlicher Nullpunkt, natürliche Intervalle und Nullpunkt).
- Skalenniveaus können über diese Relationen definiert werden.
- **Vergleiche** Anhang 4

Skalenniveau und Messfehler

- Die numerische Relation (Messniveau) muss nicht zwangsweise die empirische Relation auf dem gleichen Niveau abbilden.
- Aber, die tatsächliche empirische Relation sollte mit der Messung bestmöglichst erfasst werden.
- Das Problem dabei ist, dass
 - ▶ einen Konflikt mit dem Erhebungsaufwand geben kann,
 - ▶ sich (unsystematische) Fehler in die Messung einschleichen können.

Numerische Relation und Skalenniveau

- Beispiel: Die Erhebung des persönlichen Bruttomonatseinkommens?
- Erhebung in offener Frage oder in Kategorien?
- Tatsächliche empirische Relation ist das Bruttomonatseinkommen auf Franken und Rappen genau!
- Was erhoben wird, also die numerische Relation, ist abhängig von der Fragestellung und bestimmt damit auch das Skalenniveau.

Numerische Relation und Messfehler

- Fast keine Messung (v.a. bei latenten Konstrukten) ist fehlerlos.
- Messfehler sind dann ein Problem, wenn sie systematisch auftreten.
- Messfehler, welche unsystematisch auftreten sind unter bestimmten Bedingungen unproblematisch (für die Statistik).
- Latentes Konstrukt (unbeobachtbar) \longrightarrow Messung \longleftarrow Zufallsmessfehler
- Unter welchen Bedingungen beeinflussen Messfehler das Ergebnis nicht?

Stichproben: Worum geht es?

- Das Ziel (quantitativ) empirischer Forschung ist, Aussagen über eine gegebene Population, z.B. der Schweizer Bevölkerung zu machen.
 - ▶ Deskriptiv z.B. Mittelwert des Einkommens oder Altersverteilung einer Gesellschaft
 - ▶ Erklärend, z.B. trifft eine bestimmte Hypothese für diese Population zu?
- Das Problem ergibt sich aus dem Umstand, dass Gesamterhebungen (der Population) in der Regel nicht möglich sind.
- Die Lösung ist die Verwendung eines “verkleinerten” Abbildes (Stichprobe)

Auswahlproblematik und Grundbegriffe

- Allein aus Kostengründen empfehlen sich oftmals Teilerhebungen und damit Ziehungen von Stichproben.
 - ▶ Wie muss diese Stichprobe zusammengesetzt sein (Wahrscheinlichkeitsauswahl, andere Alternativen)?
 - ▶ Wie gross muss die Stichprobe sein (Stichprobentheorie)?
- Teilerhebungen haben verschiedene Aspekte, die sich in der folgenden Reihenfolge diskutieren lassen:
 - ▶ Selektionsbias und Beispiele
 - ▶ Grundbegriffe der Teilerhebung
 - ▶ Arten von Stichprobenverfahren

Selektionsbias und Beispiele

- Grössere Stichproben produzieren im Regelfall genauere Schätzungen der realen Gegebenheiten als kleinere Stichproben.
- Diese Regel trifft nicht zu, wenn eine grosse Stichprobe stark verzerrt ist—kleinere, aber unverzerrte Stichproben werden dann genauere Schätzungen liefern!
- Verzerrungen ergeben sich z.B. bei Passantenbefragungen in der Fussgängerzonen und den im Fernsehen üblichen Umfragen.

Grundbegriffe der Teilerhebung I

Grundgesamtheit := die für die jeweilige Untersuchung relevante Population (z.B. alle in der Schweiz stimmberechtigten Personen).

Stichprobe (Sample) := eine Auswahl von Elementen der Grundgesamtheit (z.B. eine wie auch immer gebildete Teilmenge der Wahlberechtigten).

Erhebungseinheiten := Elemente der Population, auf die sich die Auswahl bezieht und die überhaupt die Chance haben, in die Stichprobe aufgenommen zu werden (z.B. alle in der Schweiz stimmberechtigten Personen, welche über einen Telefonanschluss verfügen).

Grundbegriffe der Teilerhebung II

Untersuchungseinheiten := Elemente, auf die sich die Untersuchung bezieht und die nicht unbedingt mit den Erhebungseinheiten übereinstimmen müssen (z.B. wahlberechtigte Angehörige, über deren politische Präferenzen befragte Erhebungseinheiten Auskunft geben).

Stichprobenverfahren := eine explizite Vorschrift/Regel, welche die Auswahl der Elemente der Grundgesamtheit festlegt (z.B. zufällige Auswahl der zu befragenden Wahlberechtigten).

Stichprobenumfang := die Anzahl, n , der ausgewählten Elemente der Population, N .

Arten von Stichprobenverfahren

- Es lassen sich drei Hauptgruppen von Stichprobenverfahren unterscheiden, die allesamt in der Praxis verwendet werden:
 - ▶ Willkürliche Auswahlen
 - ▶ Wahrscheinlichkeitswahlen (Zufallsauswahlen)
 - ▶ Bewusste Auswahlen (Quotenauswahlen)
- Das erste und letzte Auswahlverfahren führen zu Stichproben, die insbesondere in den Medien oft als “repräsentativ” bezeichnet werden.
- Diese Deutung ist jedoch extrem fragwürdig!

Willkürliche Auswahl

- Der Vorgang der Stichprobenziehung wird hier nicht kontrolliert, sondern ist beliebig.
 - ▶ Z.B. Befragung willkürlich ausgewählter Personen
- Psychologische und ökonomische Experimente beruhen oftmals auf willkürlichen Auswahlen (Teilnahme nach freiwilliger Meldung).
- Sofern nur Zusammenhangshypothesen geprüft werden sollen und generell Aussagen über die Grundgesamtheit nicht interessieren, ist diese Vorgehensweise legitim und oft sinnvoll!

Wahrscheinlichkeitsauswahl

- Das Resultat einer Wahrscheinlichkeitsauswahl wird als Zufallsstichprobe bezeichnet.
- Bei diesem Auswahlverfahren hat jedes Element der Grundgesamtheit eine positiv von Null verschiedene und angebbare Wahrscheinlichkeit der Berücksichtigung in der Stichprobe.
- Zufallsstichproben sind die Voraussetzung für die Anwendbarkeit der statistischen Theorie des Schliessens von der Stichprobe auf die Population (Inferenzstatistik).

Bewusste Auswahl (Quotenauswahl) I

- Die Stichprobe wird bewusst nach vorgegebenen Regeln gezogen, wobei sich diese Regeln üblicherweise auf Quoten beziehen (Quotenauswahlverfahren, so dass die Quoten der Verteilung in der Grundgesamtheit entsprechen).
- Die Quotenanweisung legt fest, wie viele Personen mit welchen Merkmalen oder Merkmalskombinationen zu befragen sind.
- Aber, eine willkürliche Auswahl innerhalb der durch die Quoten festgelegten Gruppen.
- Zielsetzung der bewussten Auswahl ist die bewusste Konstruktion einer Stichprobe, welche ein "verkleinertes Abbild" der Grundgesamtheit ist.

Bewusste Auswahl (Quotenauswahl) II

- Diese Auswahlssystematik kann nur funktionieren, wenn man die Merkmalsverteilungen in der Population vorab kennt.
- Entgegen der Praxis der kommerziellen Wahl- und Meinungsforschung lassen sich Quotenverfahren nicht für statistische Schlüsse von der Stichprobe auf die Gesamtheit verwenden!
- “Repräsentativität” herrscht lediglich bezüglich den quotierten Merkmalen und nie für alle anderen (unbeobachteten) Merkmale.

Argumente der Quotenbefürworter (und ihre “Widerlegung”)

- Quotenmerkmale sind mit anderen Merkmalen korreliert und mithin auch für diese repräsentativ.
 - ▶ Ob dies zutrifft ist eine empirische Frage, die nicht a priori gelten kann!
- Innerhalb der Quote würde eine zufällige Auswahl getroffen, mithin handle es sich um Annäherung an geschichtete Zufallsstichproben.
 - ▶ Zufallsauswahl innerhalb der Quote dürfte kaum zutreffen, in der Regel werden diejenigen Person interviewt, die auch auskunftbereit sind.
- Bei detaillierter Quote und homogener Verteilung innerhalb der Quote seien die Stichproben auch ohne Zufallsauswahl unverzerrt.
 - ▶ Annahme der Homogenität nicht-beobachteter Merkmale ist eine empirische Frage und nicht a priori entscheidbar.

Arten der Wahrscheinlichkeitsauswahl

- Zufallsstichproben (Random Samples) erhält man als Ergebnis einer Wahrscheinlichkeitsauswahl.
- Die Inferenzstatistik (Schluss von Stichprobe auf Grundgesamtheit) setzt generell Zufallsstichproben voraus.
- Man unterscheidet:
 - ▶ Einfache Zufallsstichproben (EPSEM, SRS)
 - ▶ Gebietsverfahren und Personenwahl
 - ▶ Mehrstufige Zufallsstichproben und PPS
 - ▶ Klumpenstichprobe (Cluster Sample)
 - ▶ Stichprobenziehung nach Schichtung
- **Vergleiche** Anhang 5

Problem selektiver Stichproben

- Probleme der Stichprobenziehung bei Umfragen, z.B.:
 - ▶ Auswahl aus Telefonregister in Regionen mit niedriger Telefondichte
 - ▶ Viele Haushalte ohne Festnetzanschluss (nur Handy)
 - ▶ Internetumfragen (Selektion der Haushalte mit PC)
 - ▶ Leserbefragung, Befragungen von Fernsehzuschauern oder unter Hörern bestimmter Sender (Selektivität der Stichprobe)
- Probleme der Stichprobenziehung bei Dokumenten, z.B.:
 - ▶ Dokumente stellen keine Zufallsauswahl aller Dokumente dar (z.B. Patentmeldungen, nicht alle Erfindungen werden zum Patent angemeldet oder Kriminalitätsstatistik, nicht alle Delikte finden den gleichen Niederschlag in der Polizeistatistik).

Wichtige Aspekte bei Zufallsauswahl I

- Ist man an der Verallgemeinerungsfähigkeit der gewonnenen Aussagen (v.a. Bezug nicht nur auf die Stichprobe, sondern auf die Grundgesamtheit) interessiert, so hat man eine Zufallsstichprobe zu ziehen.
- Nur unter diesen Umständen kann das Verfahren der Inferenzstatistik (d.h. der von einer Stichprobe auf die Grundgesamtheit schliessenden Statistik) eingesetzt werden.
- Bereits vor der Erhebung der Stichprobendaten ist hierfür zu bedenken, dass der Vertrauensbereich von statistischen Schätzungen bei Zufallsstichproben im Regelfall umgekehrt proportional zur positiven Quadratwurzel aus dem Stichprobenumfang ist (das \sqrt{n} -Gesetz).

Wichtige Aspekte bei Zufallsauswahl II

- Der Stichprobenumfang n ist also zu vervierfachen, wenn man die Präzision einer Schätzung verdoppeln will, die auf der Grundlage einer Zufallsauswahl gewonnen werden soll.
- Daneben ist zu bedenken, dass bei allgemeinen Bevölkerungsumfragen mit Wahrscheinlichkeitsauswahl oftmals eine erhebliche Diskrepanz zwischen Ziel- und Umfragepopulation besteht.
- In solchen Umfragen werden z.B. keine Obdachlosen, Spitzenpolitiker, Krankenhaus- und Heiminsassen, Häftlinge, Kasernenbewohner, Ordensleute, Hausmeister öffentlicher Gebäude etc. erfasst.
- Es besteht ein “Undercoverage”, weil die Befragungen regelmässig nur in Privathaushalten durchgeführt werden!

Fehlerquellen bei Zufallsauswahl

- Stichprobenfehler (sampling variability)
- Systematische Fehler aufgrund der Stichprobenziehung:
 - ▶ Missachtung einzelner Schritte z.B. durch Interviewer bei Haushaltsauswahl oder Personenauswahl im Haushalt.
- Verzerrung, die nicht durch das Auswahlverfahren verursacht wurden (non-sampling bias):
 - ▶ Non-Response (Verweigerung)
 - ▶ Item Non-Response
 - ▶ Messfehler
 - ▶ Fehlerquellen im Interview (soziale Erwünschtheit)
 - ▶ Undercoverage

Sampling Biases

- Konzentriert man sich auf diejenigen Fehler, die mit dem Auswahlverfahren zu tun haben, so ergeben sich folgende Fehler:
 - ▶ **Zufallsfehler der Stichprobe:** Diese “Sampling Variability” kann grösser oder kleiner sein. Der jeweilige Zufallsfehler der Stichprobe kann mit statistischen Methoden abgeschätzt werden (siehe Methodenliteratur).
 - ▶ **Systematische Stichprobenfehler:** Sie treten auf, wenn Elemente der Population mit unterschiedlicher Wahrscheinlichkeit in die Stichprobe eingehen. Bei bekannten Auswahlwahrscheinlichkeiten lässt sich diese Verzerrung durch eine nachträgliche Gewichtung korrigieren. Schwerlich korrigierbar sind dagegen systematische Auswahlfehler, die sich aufgrund der Nichteinhaltung von Auswahlregeln (z.B. Random Route, Schwedenschlüssel) ergeben.

Elementare Zusammenhangsanalyse

- Ein Schwerpunkt empirischer Sozialforschung ist die Prüfung von Zusammenhangshypothesen.
- Beschränkt man sich auf Hypothesen über Zusammenhänge zwischen lediglich zwei Variablen, so ist eine bivariate Analyse durchzuführen.
- Grundlegend sind dabei folgende Verfahren:
 - ▶ Korrelation und Regression (metrische Variablen)
 - ▶ Mittelwertsvergleich (metrisches Y und dichotomes X)
 - ▶ Kreuztabellenanalyse (kategoriale Variablen)
 - ▶ Prozentsatzdifferenz (dichotome Variablen)
- Ein generelles Problem bivariater Analysen: Korrelation, Kausalität und Scheinkorrelationen.

Korrelation oder doch nicht?

- Eine medizinische Untersuchung in einem Betrieb mit $N = 400$ Mitarbeitern ergab folgendes Bild:

Bluthochdruck	Geschlecht		Total
	Mann	Frau	
ja	73	48	121
	0.380	0.231	
nein	119	160	279
	0.620	0.769	
Total	192	208	400

- $\chi^2(1) = 10.5673 > 3.8414588$ und Cramér's $V = 0.1625$
- Also ein signifikanter Zusammenhang zwischen Geschlecht und Bluthochdruck. Stimmt das?

Korrelation oder doch nicht?

- Eine mögliche Erklärung liefert die Tatsache, ob jemand raucht oder nicht?

Bluthochdruck	Raucher		Total
	Geschlecht		
	Mann	Frau	
ja	58	45	103
	0.659	0.413	
nein	30	64	94
	0.341	0.587	
Total	88	109	197

Bluthochdruck	Nichtraucher		Total
	Geschlecht		
	Mann	Frau	
ja	15	9	24
	0.144	0.091	
nein	89	90	179
	0.856	0.909	
Total	104	99	203

$$\chi^2(1) = 11.8346, \text{ Cramér's } V = 0.2451$$

$$\chi^2(1) = 1.3833, \text{ Cramér's } V = 0.0825$$

Korrelation oder Kausalität

- Nicht immer, wenn ein statistischer Zusammenhang (Korrelation) zwischen zwei Variablen besteht, liegt auch eine direkte Ursache–Wirkungsbeziehung vor (Kausalität).
- Um Anhaltspunkte zu haben, ob wir von einer Korrelation auf eine Kausalbeziehung schliessen können, müssen mindestens folgende Bedingungen erfüllt sein:
 - 1 X und Y müssen korreliert sein
 - 2 X geht Y zeitlich voraus
 - 3 Die Korrelation bleibt erhalten, auch wenn man für Drittvariablen kontrolliert
- Aber, was Ursache und was Wirkung ist, ist vielfach unklar und kann nie statistisch erklärt werden. Dazu benötigen wir eine Theorie!

Korrelation und Regression

- Die lineare Assoziation zwischen zwei oder mehr zumindest intervallskalierten Variablen y und x kann erfasst werden durch:
 - ▶ den Produkt–Moment oder Bravais–Pearson Korrelationskoeffizienten r_{xy} ,
 - ▶ die Schätzung einer linearen Regressionsgleichung.
- Für nicht–lineare Zusammenhänge sowie nicht–metrische abhängige Variablen gibt es entsprechende Regressionsverfahren.

Inferenzstatistische Grundlagen I

- Die angesprochenen Konzepte und Verfahren sind allesamt beschreibender Natur (deskriptive Statistik) und beziehen sich jeweils nur auf das vorliegende Datenmaterial.
- Häufig ist man aber an der Verallgemeinerungsfähigkeit der gewonnenen Aussagen (v.a. Bezug nicht nur auf die Stichprobe, sondern auf die Grundgesamtheit bei der Stichprobenziehung) interessiert.
- Sind die Daten im Rahmen einer Zufallsstichprobe erhoben worden, so kann man zur Gewinnung derartiger Aussagen die Verfahren der Inferenzstatistik (d.h. der von einer Stichprobe auf die Grundgesamtheit schliessenden Statistik) einsetzen.

Inferenzstatistische Grundlagen II

- Wichtige inferenzstatistische Verfahren sind Signifikanztests.
- Grob gesprochen wird in ihrem Rahmen systematisch geprüft, ob die anhand der Stichprobendaten gewinnbaren Einsichten “überzufällig” (d.h. nicht nur rein zufallsbedingt) auftreten.
- Bei Signifikanztests handelt es sich um statistische Prüfungen von Hypothesen anhand gegebener Stichprobendaten.
- Gefragt wird, ob eine gegebene Beobachtung oder Schätzung mit einer explizit formulierten Hypothese für die Population hinreichend vereinbar ist.
- Getestet wird stets die so genannte Nullhypothese H_0 gegen ihr Gegenteil, die sogenannte Alternativhypothese H_1 .

Inferenzstatistische Grundlagen III

- Das Testen von Hypothesen beruht dabei auf Regeln für die Entscheidung, ob die H_0 verworfen werden kann oder nicht.
- Das statistische Testen von Hypothesen beantwortet somit immer folgende Frage:

Statistisches Testen

Ist eine gegebene Beobachtung oder Schätzung (z.B. der Punktschätzer $\hat{\theta}$) mit einer formulierten Hypothese "vereinbar" (d.h. liegt sie hinreichend nahe bei der Hypothese) oder nicht?

Wichtige Aspekte bei Signifikanztests I

- Durch statistische Hypothesen- oder Signifikanztests wird niemals mit letzter Sicherheit die Gültigkeit einer Hypothese nachgewiesen.
- Die Entscheidung “Beibehaltung von H_0 ” (d.h. Ablehnung der H_1) ist nicht unter statischer Kontrolle (für diese Situation wird ja keine maximale Irrtumswahrscheinlichkeit (Fehler der 2. Art) festgelegt).
- Die Nichtsignifikanz eines Zusammenhangs ist v.a. bei kleinen Stichproben mithin kein sicherer Nachweis, dass tatsächlich kein Zusammenhang existiert.
- Bei sehr grossen Stichproben werden aus rein mathematischen Gründen bereits recht kleine Abweichungen oder Zusammenhänge statistisch signifikant (wie z.B. in der Astrologie–Studie von Gunther Sachs).

Wichtige Aspekte bei Signifikanztests II

- Insbesondere bei grossen Stichproben ist daher, wenn immer möglich, die Grösse des postulierten Effektes anzugeben.
- Es empfiehlt sich also, die Bedeutsamkeit oder Relevanz signifikanter Befunde abzuschätzen.

Phase I: Forschungsproblem

- Die Formulierung des Forschungsproblems kann mehrere Gründe haben.
- Ein Forschungszusammenhang ergibt sich z.B.
 - ▶ aus der wissenschaftlichen Literatur
 - ▶ aus neuen Ideen des Forschers
 - ▶ aus praktischen Problem
 - ▶ durch einen Auftraggeber
 - ▶ durch Anregung aus explorativen Studien (Einzelbeobachtung)
 - ▶ aus der Evaluation (politischer) Massnahmen

Phase II: Planung und Vorbereitung

- Die Planungsphase unterteilt sich in vier Teile:
 - ▶ Konstruktion des Erhebungsinstruments, d.h. Definition der Begriffe und Operationalisierung.
 - ▶ Festlegen der Untersuchungsform.
 - ▶ Stichprobenverfahren, d.h. Definition der Population (Grundgesamtheit) sowie Art und Umfang der Stichprobe.
 - ▶ Pretest (Test des Erhebungsinstrumentes auf Logik des Instruments, Verständnis und Schwere der Fragen und Varianz auf dem Antwortverhalten).
- Die einzelnen Komponenten können nicht zeitlich nacheinander abgearbeitet werden, da die Entscheidungen in den einzelnen Komponenten z.T. voneinander abhängig sind.

Phase III: Durchführung

- Die Erhebungsform wird meist schon in früheren Planungsphasen festgelegt.
- Mögliche Erhebungsformen sind
 - ▶ Befragung (face-to-face, telefonisch, schriftlich)
 - ▶ Beobachtung (diverse Formen)
 - ▶ Inhaltsanalyse
 - ▶ verschiedene nicht-reaktive Verfahren
- Vor der Durchführung der Erhebung sind z.B. die Interviewerschulung und -kontrolle zu planen.

Phase IV: Datenaufbereitung und -auswertung

- Die Datenauswertung erfordert den “Aufbau” eines Datenfiles.
- Das heisst, sofern die Daten noch nicht in elektronischer Form vorliegen, müssen diese erfasst werden.
- Danach erfolgt die Fehlerkontrolle und Logikprüfung der Daten.
- Die statistische Datenanalyse erfordert u.a. die Bildung von Indizes und die Konstruktion von Skalen.
- Auswertungen der Daten mittels univariater Deskription sowie bi- und multivariaten Analysen.

Phase V: Forschungsbericht

- Forschungsergebnisse werden, je nach “Grund” der Forschung, in
 - ▶ Fachzeitschriften veröffentlicht,
 - ▶ an Konferenzen präsentiert und/oder
 - ▶ in Forschungsberichten zusammengefasst.
- Vielfach werden die Daten nach Abschluss eines Projektes anderen Forschern zugänglich gemacht.

Zurück

Klassisches experimentelles Design

- Bezeichnen wir mit R die Zufallsaufteilung auf Versuchs- und Kontrollgruppe (Randomisierung).
- Dann sieht das **klassische experimentelle Design** folgendermassen aus:

$$\begin{array}{c|cc} R & X & O_1 \quad (\text{Versuchsgruppe}) \\ R & & O_2 \quad (\text{Kontrollgruppe}) \end{array}$$

- Die Randomisierung sorgt für die Kontrolle (Gleichverteilung) aller unbeobachteten Merkmale auf die Versuchsgruppen.
- Unter diesen Umständen lassen sich Unterschiede in den Gruppen auf den Stimulus zurückführen (und hängen nicht mit den unbeobachteten Merkmalen zusammen).

Varianten experimenteller Designs I

- Es gibt eine Vielzahl von Erweiterungen und Modifikationen des Standarddesigns (siehe Methodenliteratur).
- Ein Beispiel ist die Erweiterung auf mehrere verschiedene Stimuli, bei der jede Gruppe sowohl Kontroll- als auch Versuchsgruppe ist:

R	X_1	O_1	(Versuchsgruppe 1)
R	X_2	O_2	(Versuchsgruppe 2)
R	X_3	O_3	(Versuchsgruppe 3)
	\vdots	\vdots	
R	X_k	O_k	(Versuchsgruppe k)

Varianten experimenteller Designs II

- Ein weiteres Beispiel ist die Vorher–Nachher–Messung bei Versuchs- und Kontrollgruppe:

$$\begin{array}{l|l} R & O_1 \quad X \quad O_2 \quad (\text{Versuchsgruppe}) \\ R & O_3 \quad \quad O_4 \quad (\text{Kontrollgruppe}) \end{array}$$

- Der Vorteil dieses Designs liegt in der Kontrolle der “Ausgangssituation”.
- Aber, die Vorhermessung kann die Nachhermessung beeinflussen.

Varianten experimenteller Designs III

- Ein drittes Beispiel ist **Solomons Vier-Gruppen Versuchsplan**:

R	O_1	X	O_2
R	O_3		O_4
R		X	O_5
R			O_6

- Wirkung des Stimulus wird zweifach überprüft und der (mögliche Effekt) des Messinstruments wird isoliert.
- Bei kausalem Zusammenhang erwarten wir $O_2 > O_1$, $O_2 > O_4$, $O_5 > O_6$ sowie $O_5 > O_3$.

Quasi-Experimente und ihre Probleme

- Quasi-Experimente sind in der Hauptsache Experimente ohne Randomisierung (gebräuchlich v.a. in der Evaluationsforschung).
- Es gibt verschiedene Anordnungen (siehe Methodenliteratur), wovon hier insbesondere die Versuchsanordnung mit nicht gleichartiger Kontrollgruppe besprochen wird.
- Letztere entspricht einem Design mit Vorher-Nachher Messung ohne Zufallsaufteilung:

O_1 X O_2 (Versuchsgruppe)

O_3 O_4 (Kontrollgruppe)

- Drittvariablen können das Ausmass des Effektes von X verzerren!
- Weitere Probleme sind die Selbstselektion (ins Experiment) und ein systematischer Ausfall (während dem Experiment) von Probanden.

Regressionseffekt und Zeitreihen

- Eine wesentliche Fehlerquelle beim quasi-experimentellen Design mit nicht-gleichartiger Kontrollgruppe ist der sogenannte “Regressions-effekt” (statistische Tendenz zur Mitte und nicht Effekt des Stimulus).
- Der Effekt tritt bei starken Abweichungen von lediglich einer Vorher- und Nachhermessung wahrscheinlich auf.
- Regressionseffekte können u.a. mit einem Zeitreihen-Design kontrolliert werden:

$O_1 O_2 O_3 O_4 X O_5 O_6 O_7 O_8$

- Mit längerfristigen Zeitreihen sind Regressionseffekt und Reifung identifizierbar, nicht aber zwischenzeitliche Ereignisse.

Zurück

Messung und Operationalisierung

Messung (klassische Definition)

Messung ist die Zuordnung von Zahlen zu Objekten nach bestimmten Regeln.

Messung (modernen Definition)

Die Messung unterstellt die Existenz einer zumindest eindeutigen Abbildung zwischen einem empirischen und einem numerischen Relativ (im Sinne der axiomatischen Messtheorie).

Die moderne Definition ist präziser, die klassische Definition ist allgemeiner!

Direkte und indirekte Messung I

- **Direkte Messung:** Direkt beobachtbarer Merkmale wie z.B. Gewicht, Körpergröße oder Anzahl geraucher Zigaretten pro Tag.
- **Indirekte Messung:** Nicht direkt beobachtbarer Größen, d.h. theoretische Konstrukte wie z.B. Intelligenz, Leistungsbereitschaft, Umweltbewusstsein oder Lebensqualität in einer Stadt.
- Die Messung von direkt beobachtbaren Merkmalen ist relativ unproblematisch.

Direkte und indirekte Messung II

- Die Messung nicht unmittelbar beobachtbarer Merkmale oder unscharfen vieldeutigen Konstrukten erfordert die Zuordnung von beobachtbaren Merkmalen (**Indikatoren**) zu den theoretischen Begriffen.
 - ▶ Die Messung des Werts einer Währung (z.B. Big-Mac-Index)
 - ▶ Die Messung des Wohlstands von Nationen (z.B. BIP, HDI)
 - ▶ Die Messung des Body-Mass Index
- Es bedarf also einer genauen Definition der Begriffe, die in einer Theorie auftauchen und gemessen werden sollen (Konzeptspezifizierung).
- Konzeptspezifizierung → Operationalisierung → Messung.

Zurück

Skalenniveaus und Beispiele I

- **Nominalskala:** Objekte oder Merkmale können lediglich klassifiziert bzw. unterschieden werden, eine Rangfolge ist nicht möglich:

$$a = b \text{ sowie } a \neq b$$

Beispiel: Geschlecht, Hautfarbe, Nationalität.

- **Ordinalskala:** Objekte oder Merkmale lassen sich in eine Rangfolge bringen:

$$a < b, a = b \text{ sowie } a > b$$

Beispiel: Mit einer Guttman-Skala erhobene Merkmale, Resultate von Paarvergleichen (z.B. ordinaler Nutzen), Schulnoten.

Skalenniveaus und Beispiele II

- **Intervallskala:** Objekte oder Merkmale können in eine Rangfolge gebracht werden und die Interpretation der Abstände ist möglich:

$$a < b < c < d \text{ und damit } b - a = d - c$$

Beispiel: Kalenderzeit, Temperatur in Grad Celsius oder Fahrenheit, kardinaler Nutzen, mit Likert-Skalen erhobene Merkmale.

- **Ratioskala:** Objekte oder Merkmale verfügen zusätzlich den Eigenschaften der Intervallskala über einen absoluten Nullpunkt:

$$\frac{a}{b} = \frac{c}{d}$$

Beispiel: Ehedauer, Einkommen, Preise.

Skalenniveaus und Beispiele III

- **Absolutskala**: Objekte oder Merkmale verfügen über einen absoluten Nullpunkt und genau festgelegte Skalenabstände:

$$a = b$$

Beispiel: Demonstrantenzahlen, Wahrscheinlichkeiten.

Zulässige Transformationen I

In Abhängigkeit von ihrem Messniveau kann die Variable y in eine andere y' überführt (transformiert) werden.

Nominalskala: $y' = f(y)$ mit der beliebigen Funktion f (eindeutige Transformation).

Ordinalskala: $y' = f(y)$ mit der Funktion f , die eine positive erste Ableitung ($df/dy > 0$) besitzt (positive monotone, d.h. ordnungserhaltende Transformation).

Intervallskala: $y' = a + by$ mit dem beliebigem Nullpunkt a und der Konstante $b > 0$ (positive lineare Transformation).

Ratioskala: $y' = by$ mit Konstante $b > 0$ (positive affine, d.h. proportionale Transformation).

Absolutskala: $y' = y$ (identitätsbewahrende oder keine Transformation).

Zurück

Einfache Zufallsstichproben (EPSEM, SRS)

- Eine einfache Zufallsstichprobe liegt vor, wenn
 - ① die Auswahl direkt in einem einstufigen Auswahlvorgang erfolgt und
 - ② die Auswahlwahrscheinlichkeit für alle Elemente der Grundgesamtheit identisch grösser als Null ist.
- Sind beide Kriterien erfüllt, so liegt ein “Simple Random Sampling” (SRS) vor.
- Ist dagegen nur (2) gegeben, dann spricht man von der “Equal Probability Selection Method” (EPSEM).
- Wie kann man (1) und (2) erfüllen?
 - ▶ Listenauswahl (Verzeichnis sämtlicher Untersuchungseinheiten, z.B. Telefonbuch, Einwohnermeldeliste)
 - ▶ Random-Digit-Dialing (RDD-Methode)
 - ▶ Lotteriewahl

Gebietsverfahren und Personenwahl

- Um eine SRS aus einem Gebiet zu erhalten, arbeitet man mit dem “Random Route” Verfahren (und eventuell mit der Verfeinerung “Adressrandom”).
- Zur Zufallsauswahl innerhalb von Mehr–Personen–Haushalten verwendet man folgende Verfahren:
 - ▶ Bei persönlichen Interviews: “Schwedenschlüssel”; dieser informiert über eine (für die jeweilige Haushaltgrösse relevante) Zufallszahl, die je nach der Sequenz des Alters der Haushaltsmitglieder die zu interviewende Person bestimmt.
 - ▶ Bei telephonischen Interviews: Befragt wird die zur Grundgesamtheit gehörende Person im Haushalt, die zuletzt Geburtstag hatte. Oder Schwedenschlüssel.

Mehrstufige Zufallsstichproben und PPS

- Erstreckt sich eine Zufallsauswahl über mehrere Ebenen, so liegt eine mehrstufige Zufallsstichprobe vor.
- Die typische Zufallsstichprobe einer nationalen Umfrage in der Schweiz ist dreistufig:
 - 1 Auswahl von Gemeinden mit Zufallszahlen
 - 2 Auswahl von Haushalten per Random Route
 - 3 Auswahl von zu befragenden Haushaltsmitgliedern per Schweden-schlüssel
- Ein Spezialfall zweistufiger Zufallsstichproben sind PPS-Samples.
- Hier ist die Auswahlwahrscheinlichkeit auf der ersten Stufe proportional zur Grösse der Auswahleinheiten (“Probability Proportional to Size”, PPS).
- Unterschiedliche Auswahlwahrscheinlichkeiten innerhalb des Haushaltes können durch Gewichtung korrigiert werden.

Klumpenstichprobe (Cluster Sample)

- Eine Klumpenauswahl ist eigentlich ein Spezialfall einer mehrstufigen Auswahl.
- Bei ihr werden zunächst einige Klumpen (Cluster) zufällig aus vielen ausgewählt.
- Dann werden alle Elemente jedes Klumpens in die Stichprobe aufgenommen (d.h. Auswahlwahrscheinlichkeit von 1).
- Beispiel: Zufällig Auswahl von einigen Schulklassen eines Schulhauses, dann Befragung aller Schüler in den jeweiligen Schulklassen.
- Die Schätzung ist umso genauer, je heterogener die Elemente in den Klumpen und je homogener die Elemente zwischen den Klumpen sind.

Stichprobenziehung nach Schichtung

- Sind Informationen über die Merkmalsverteilungen in der Grundgesamtheit vorhanden und ...
- ... ist die Heterogenität zumindest eines Merkmals sehr hoch, so kann man das Fehlerintervall der Schätzung reduzieren.
- Dazu wird die Population zunächst in Schichten (z.B. Einkommens-, Berufsgruppen, Geschlecht) zerlegt und dann aus jeder Schicht separat eine Zufallstichprobe gezogen.
- Die Schätzung wird dann umso präziser sein, je homogener die Elemente in einer Schicht sind.
- Voraussetzung: Schichtanteil muss bekannt sein und Zugehörigkeit der Untersuchungseinheiten muss erkennbar sein.
- Vorteil: genauere Schätzungen bei "seltenen Merkmalsträgern" (z.B. Scheidungsstudien).

[Zurück](#)