

# Willkommen zur Vorlesung Empirische Methoden I

## 11. Sitzung: Stichproben

Prof. Dr. Wolfgang Ludwig-Mayerhofer

Universität Siegen – Philosophische Fakultät, Seminar für Sozialwissenschaften

# Themen der Vorlesung

- Überblick/Einführung
- Auswahlverfahren bei qualitativer Forschung
- Auswahlverfahren bei standardisierter Forschung
- Stichproben in der Umfrageforschung
- Stichproben für andere Erhebungsverfahren
- Stichprobentheorie (5-Minuten-Version)

# Warum Stichproben?

- Totalerhebungen im allgemeinen nicht realisierbar und zu teuer (Ausnahme: Makro-Analysen).
- Der Preis für Stichproben: Unsicherheit.
- Aber: Das Ausmaß der Unsicherheit ist (jedenfalls bei standardisierter Forschung) im Prinzip exakt bestimmbar! Oder sie kann (bei weniger standardisierter Forschung) im Prinzip glaubwürdig reduziert werden.

## Zwei Leitbilder für Stichproben

- Standardisierte Verfahren: Zufallsstichprobe
  - Ausgangspunkt immer: definierte Grundgesamtheit (Population).
  - Zufallsprozesse bei der Auswahl der Stichprobenelemente sollen Genauigkeit des Schlusses auf Grundgesamtheit (Inferenzschluss) berechenbar machen.
- Weniger standardisierte Verfahren: Schrittweise Auswahl bzw. Orientierung an Relevanz
  - Grundgesamtheit kann, muss aber nicht vorher bekannt sein.
  - Auswahl hängt von theoretischen Überlegungen ab.

# Auswahlverfahren bei qualitativer Forschung

- Häufig propagiert: **Theoretisches Sampling**  
Schrittweise Auswahl im Forschungsprozess, je nach bisher erzielten Ergebnissen, bis zur „theoretischen Sättigung“ (es scheinen keine neuen Erkenntnisse mehr erzielbar)
- Weitere Möglichkeiten schrittweiser Auswahl:  
Auswahl extremer, typischer, ‚kritischer‘ (besonders zentraler), politisch wichtiger oder ‚sensibler‘ Fälle
- Faktisches Vorgehen bei Auswahl häufig: Schneeballverfahren (Befragte geben Hinweis auf weitere Fälle)
- Andere Formen (z. B. Vorab-Festlegung) möglich, sollten aber, wenn möglich/erforderlich, ergänzt werden

# Grundprinzipien bei standardisierter Forschung

- Grundgesamtheit **muss** angebar sein (Ausnahme: Experimentelle Forschung).
- Zufallsverfahren sind **regelgeleitete** (also gerade nicht: willkürliche) Verfahren zur Auswahl von Stichprobenelementen.
- Vorgehen in den Grundzügen:
  - Festlegung der Population/Grundgesamtheit (GG)
  - Bestimmung der Auswahlgesamtheit
  - Ziehen der Stichprobe
  - Inferenzschluss (Schluss von Stichprobe auf GG – mit angebarerer Unsicherheit)

# Zufallsstichproben

- Zufallsstichprobe: Jedes Element der GG hat eine angebbare (von null verschiedene) Wahrscheinlichkeit, in die Stichprobe zu gelangen.
- Alle Elemente werden symbolisch verfügbar gemacht (z. B. Adressdatei), oder es werden Verfahren entwickelt, denen zu Folge im Prinzip jedes Element aufgefunden werden kann (z. B. Random Route)
- Welche Elemente konkret in Stichprobe kommen, wird durch Zufallsverfahren (z. B. Zufallszahlen, zufällige Startpunkte) bestimmt.

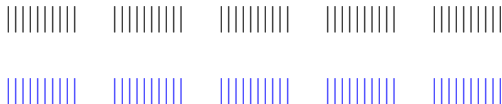
## Einstufige (direkte) Zufallsstichproben

- Einfache Zufallsstichprobe: Auswahlwahrscheinlichkeit für alle Elemente gleich; direkte Zufallsziehung der Elemente für Stichprobe.
- Systematische Stichprobe: Zufällige Bestimmung des „Anfangselements“, dann Ziehung jedes x-ten Elements (problematisch bei Systematik in Anordnung der Elemente).
- Geschichtete Stichprobe: Einteilung der GG in ‚Schichten‘ (nicht im Sinne der Sozialstrukturanalyse!) (z. B. Westdeutsche / Ostdeutsche); (möglichst) einfache Zufallsstichprobe aus den Schichten (bei richtigem Vorgehen größere Genauigkeit als bei einfacher Zufallsstichprobe!).



# Mehrstufige Zufallsstichproben I

- Mehrstufige Auswahl: Zufallsauswahl größerer Einheiten (z. B. Gebiete oder Gemeinden), daraus kleinerer Einheiten (z. B. Personen; oder: zunächst Haushalte und aus diesen Personen).
- Problematisch, wenn größere Einheiten in sich homogen und untereinander heterogen (Beispiel [extrem verkleinert]: 10 Wahlkreise; davon 5 schwarze und 5 blaue):



Konsequenz: Möglichst viele größere Einheiten auswählen!

## Mehrstufige Zufallsstichproben II: Weitere Varianten

- PPS-Design bei unterschiedlich großen Einheiten erster Stufe:: Auswahlwahrscheinlichkeit auf erster Stufe (z. B. Gemeinden) proportional zur Größe; dann zufällige Auswahl der gleichen Anzahl von Personen aus jeder größeren Einheit.
- Klumpenstichprobe: Auswahl größerer Einheiten (z. B. Schulklassen); alle Elemente dieser Einheiten (im Beispiel: Schüler) kommen in die Stichprobe.

## Bewusste Stichproben: Quotenstichprobe

- Quotenstichprobe: Es wird ein Abbild der Bevölkerung festgelegt (z.B.: 1,5 % sind Frauen über 65 mit Abitur, 3,3 % sind Männer mit Hauptschulabschluss zwischen 30 und 40, usw. [Zahlen sind frei erfunden]).
- Interviewer sollen entsprechend viele Personen mit den jeweiligen Merkmalskombinationen suchen (d. h. jeder Interviewer erhält Quotenvorgabe, etwa: Suche 3 Frauen über 65 mit Abitur, 6 Männer mit Hauptschulabschluss zwischen 30 und 40, usw.)
- Problem: Auswahl der Elemente hängt nicht mehr vom Zufall ab, sondern vom (unkontrollierbaren) Verhalten der Interviewer.

# Willkürliche Stichprobe

Bei einer willkürlichen Stichproben werden ‚irgendwelche‘ Personen befragt (oder wählen sich selbst aus für Teilnahme an Befragung).

Beispiele:

- Befragung von Passanten in Fußgängerzone
- Fragebögen, die nach Belieben in Briefkästen geworfen werden
- Befragung von Besuchern einer Webseite

Hier können keinerlei Aussagen über eine irgendwie geartete Grundgesamtheit gemacht werden, weil gar nicht klar ist, welche Grundgesamtheit von der Stichprobe abgebildet wird.

Alle quantitativen Aussagen, die auf der Grundlage willkürlicher Stichproben getroffen werden, und seien diese noch so groß, sind in wissenschaftlicher Hinsicht wertlos und haben keinen Platz in wissenschaftlicher Argumentation.

# Stichproben in der Umfrageforschung I

- Mündliche (face-to-face) Befragung
  - Auswahl: Einwohnermeldeamt (Personen) vs. Random Route (Haushalte)
  - Auswahl der Personen im Haushalt muss ebenfalls zufällig geschehen (Schweden-Schlüssel [Zufallszahl] oder „Last-Birthday“-Methode)
  - Problem der Unterrepräsentanz von Personen aus großen Haushalten
- Postalische Befragung
  - Im Allgemeinen nur sinnvoll auf Grundlage von Adresslisten

# Stichproben in der Umfrageforschung II

- Online-Befragung
  - Sollte auf (aktuellen) Listen beruhen.
- Telephonische Befragung
  - Telephonlisten vs. RDD (Random Digit Dialing)
  - In Deutschland wegen Möglichkeit, Eintrag in Telefonbuch zu unterbinden, Listen nicht mehr so gut wie früher.
  - Besonderheiten des RDD für Deutschland: siehe nächste Seite!
  - Für echte Zufallsauswahl muss wieder in jedem Haushalt zufällig eine Person ausgewählt werden!

## RDD in Deutschland (nach GESIS)

- Telekom benutzt „Hunderterblöcke“, z.B. 31 74 00 bis 31 74 99
- Innerhalb jedes Blocks sind nicht alle, aber einige bis viele Nummern vergeben.
- Viele Blöcke sind ganz leer, d. h. nach dem obigen (fiktiven) Block könnte es z. B. weitergehen mit 32 18 00 bis 32 18 99 (Nummern 31 75 00 bis 32 17 99 sind nicht vergeben)
- Um nicht zu viele „Leernummern“ (ohne Anschluss) zu erzeugen, geht man von existierenden Nummern aus und erzeugt zufällig weitere Nummern aus dem entsprechenden Block (z. B. 31 74 13 → weitere Nummern 31 74 08, 31 74 21, 31 74 54, 31 74 59, 31 74 73).

# Stichproben bei anderen Erhebungsverfahren

- Die grundlegenden Ideen der Zufallsstichprobe und ihrer Varianten gelten auch hier, aber die Praxis der Stichprobenziehung kann deutlich abweichen.
- Sehr gute Voraussetzungen für Zufallsstichproben weisen z. B. prozessproduzierte Daten auf (fortlaufende Nummerierung von Akten bzw. symbolische Abbildung aller ‚Akten‘ in Datei); aber auch inhaltsanalytische Verfahren, sofern das Ausgangsmaterial (Grundgesamtheit) gut repräsentiert ist.
- Bei Beobachtungsverfahren *kann* dagegen die Stichprobenziehung schwierig sein: Was sind die grundlegenden Einheiten (z. B. Situationen, Zeiträume, Orte, bzw. Personen in bestimmten Situationen, Zeiträumen . . . )?



# Bedeutung der Qualität von Stichproben I

- Das mit großem Abstand wichtigste Kriterium für die Beurteilung von Stichproben ist deren Qualität.
- Die Größe einer Stichprobe hat nur dann irgendeine Bedeutung, wenn es sich um Stichprobe hoher Qualität (korrekt gezogene Zufallsstichprobe) handelt. Eine sehr große, aber schlechte Stichprobe hat nicht mehr Aussagekraft als eine kleine schlechte Stichprobe, nämlich überhaupt keine.
- Bei einer guten Stichprobe ist die Stichprobengröße von Bedeutung für die Genauigkeit des Schlusses von der Stichprobe auf die Grundgesamtheit (mehr dazu in der kommenden Vorlesung).

## Bedeutung der Qualität von Stichproben II

- Wichtig im vorgenannten Sinn ist einzig die Größe der Stichprobe, nicht der Auswahlsatz (also der Anteil der Grundgesamtheit, der letztlich in die Stichprobe gelangt). Eine Stichprobe von 1200 Wahlberechtigten der Bundesrepublik entspricht einem Auswahlsatz von  $2 / 100\,000$  – und ist für normale Ansprüche an die Genauigkeit völlig ausreichend!
- Bei (echten) Experimenten steht die Qualität der Stichprobe weniger im Vordergrund; hier entscheidet das Gelingen der Randomisierung. Die Frage der Gewinnung der Stichproben kann allerdings wichtig sein für die Frage der Generalisierung der Experimente auf andere Bevölkerungsgruppen (90 Prozent der Experimente in der Psychologie werden an Studierenden durchgeführt!).