

# TEIL 5: AUSWAHLVERFAHREN

## Auswahlverfahren – eine Umschreibung

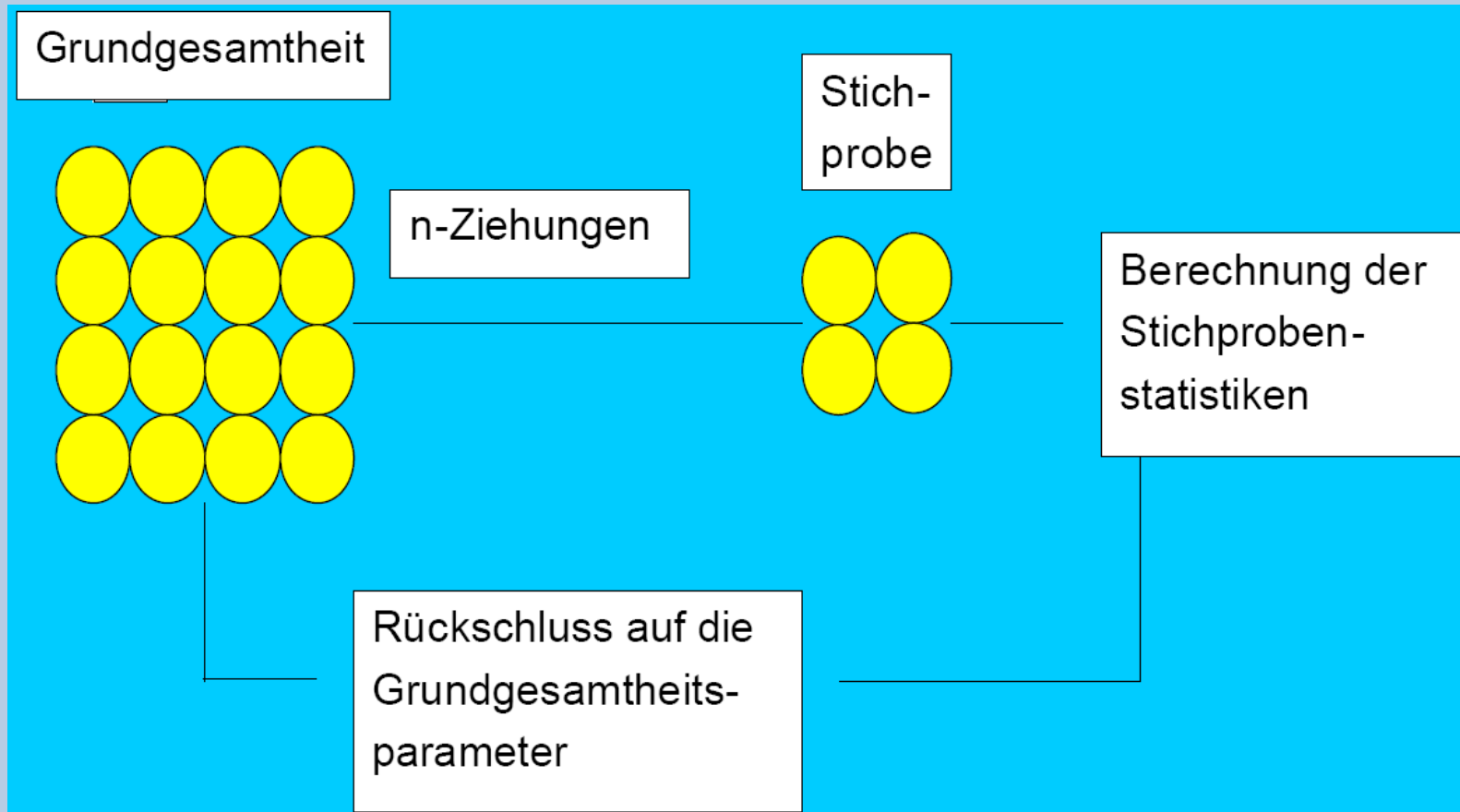
- Beschäftigen sich mit der Problematik, **wie Objekte** (Personen, Gruppen, Organisationen, Staaten etc.) **für eine Untersuchung** (Befragung, Experiment etc.) **ausgewählt werden sollen**
- Stellen also Verfahren dar, welche die **prinzipiellen Regeln zur Konstruktion von Stichproben** angeben

## Grundbegriffe

- **Untersuchungseinheit:** Objekt an dem Messungen vorgenommen werden
- **Grundgesamtheit** (Population): Menge von Objekten, über die Aussagen im Rahmen einer Untersuchung gemacht werden sollen
- **Vollerhebung:** Es werden Daten zu allen Objekten einer Grundgesamtheit erhoben
- **Teilerhebung:** Nur eine Teilmenge der Grundgesamtheit wird untersucht

- **Stichprobe:** Die Menge der Elemente einer Teilerhebung, welche über festgelegte Auswahlregeln bestimmt wurde
- **Auswahlregeln:** Können auf zufälligen, willkürlichen oder bewussten Auswahlprozessen fußen
- **Statistiken:** Maßzahlen, welche sich auf die Stichprobe beziehen
- **Parameter:** Maßzahlen, welche sich auf die Grundgesamtheit beziehen

## Graphische Darstellung: Grundgesamtheit und Stichprobe



## Vor- und Nachteile: Voll- vs. Teilerhebung

<u>VOLLERHEBUNG</u>	<u>TEILERHEBUNG</u>
Alle Informationen liegen unverfälscht vor	Fehlerhafter Induktionsschluss beim Ableiten der Populationsparameter aus den Stichprobenstatistiken
Kosten steigen überproportional mit dem Umfang der Population	Relativ kostengünstig
Sehr hoher Arbeitsaufwand, somit viele Quellen für (Flüchtigkeits-)Fehler	Stichprobenergebnisse können genauer sein, da der Arbeitsaufwand überschaubar ist

Dunkelblaue Hintergrundfarbe = Vorteil

## Angestrebte Grundgesamtheit, Auswahlgesamtheit und Inferenzpopulation

- **Angestrebte Grundgesamtheit** – umfasst alle Elemente, für die eine vom Forscher zu prüfende Theorie gelten soll
- **Auswahlgesamtheit** – Eine (zumindest virtuelle) Auflistung aller Elemente, die eine echte Chance haben, in eine Stichprobe zu gelangen
- **Inferenzpopulation** – Abstrakte Grundgesamtheit, über die auf der Basis der tatsächlich realisierten Stichprobe Aussagen gemacht werden können<sup>1</sup>

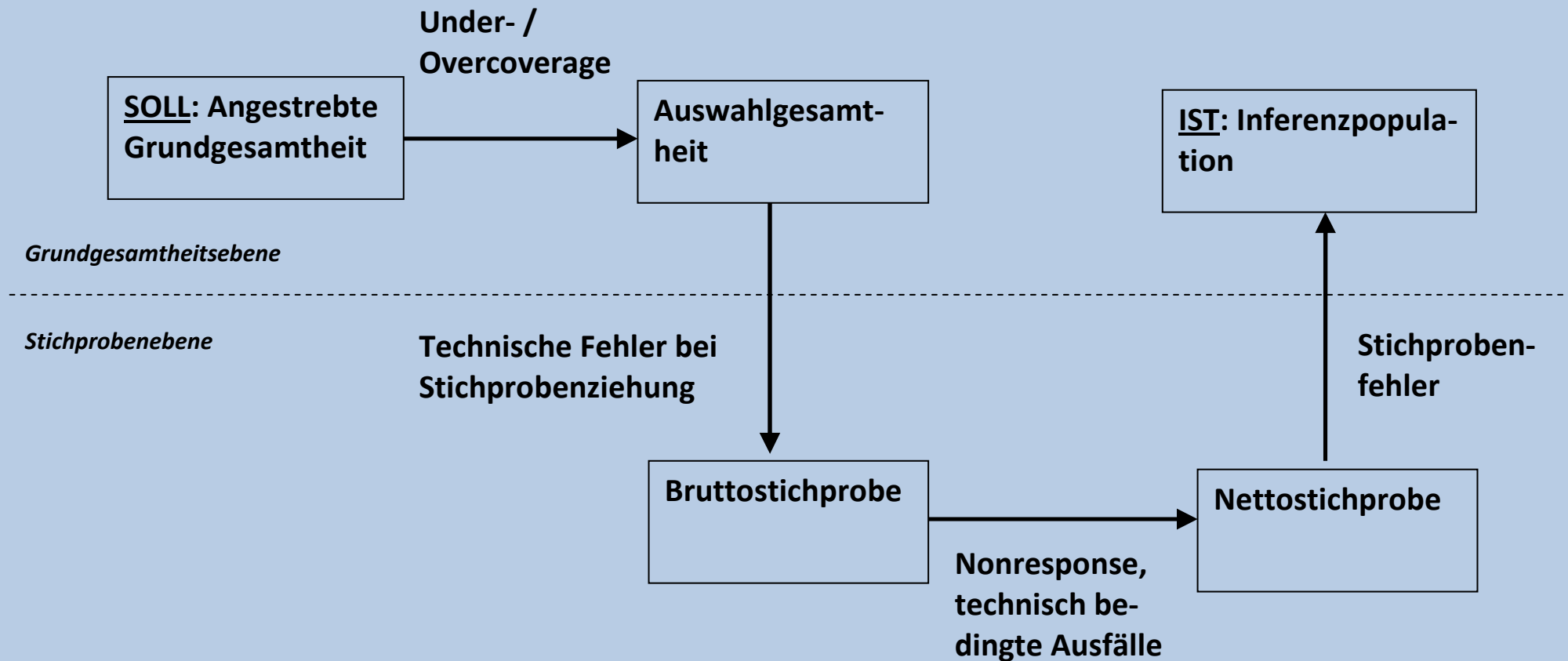
---

<sup>1</sup> m.a.W. die Grundgesamtheit, aus welcher die vorliegende Stichprobe eine reine unverzerrte Zufallsstichprobe darstellt

**Ziel: Angestrebte Grundgesamtheit und Inferenzpopulation sollten  
möglichst genau übereinstimmen!**



## Verhältnis zwischen Populationsbegriffen / Fehlerquellen:



## ***Begriffserläuterungen:***

**Under- / Over-Coverage:** In der Auswahlgesamtheit fehlen Objekte, die zur Grundgesamtheit gehören (under); in der Auswahlgesamtheit sind Objekte aufgeführt, die gar nicht zur Grundgesamtheit gehören (over).

***Technische Fehler bei Stichprobenziehung:*** Die Stichprobenziehung ist eine eigenständige Prozedur, die in der Regel mit Softwareprogrammen vorgenommen wird. Mögliche Probleme: Lückenhaft eingelesene Daten aus der Auswahlgesamtheit; Software kann nicht alle Daten lesen; das Ergebnis der Stichprobenziehung, also die eigentliche Stichprobe, wird lückenhaft oder fehlerbehaftet an die Feldkoordinatoren weitergeleitet.

**Bruttostichprobe:** Die auf Basis der Auswahlgesamtheit gezogene Stichprobe

**Nettostichprobe:** Die Menge der auf Basis der Bruttostichprobe im eigentlichen Datenerhebungsprozess untersuchten Objekte

**Nonresponse:** s.u.

**Technisch bedingte Ausfälle:** z.B. durch Interviewer / Feldkoordinatoren versehentlich vernichtete oder mit einem falschen Status versehene Fälle

**Stichprobenfehler:** Der allgemeine Rückschlussfehler, wenn von einer bestimmten Menge der Stichprobenobjekte auf eine größere Menge der Objekte in der Grundgesamtheit geschlossen wird (s.o.)

## Zufallsstichproben

- Auswahlen, bei denen ein **Zufallsprozess** über die **Aufnahme eines Elementes in die Stichprobe** entscheidet
- Jedes Element aus der Auswahlgesamtheit hat eine **mathematisch bestimmbare Chance** (Inklusionswahrscheinlichkeit), in die Stichprobe zu gelangen
- Lediglich Zufallsstichproben erlauben die **Anwendung inferenzstatistischer Techniken**, d. h. nur für sie sind die **Fehler**, die beim Schluss von der Stichprobe auf die Grundgesamtheit entstehen, **berechenbar** (Stichprobenfehler)

## ***Inklusionswahrscheinlichkeit:***

Die Wahrscheinlichkeit für ein Element, in die Stichprobe zu gelangen

## ***Unterscheidung:***

- Jedes Element besitzt **dieselbe** Inklusionswahrscheinlichkeit
- Elemente haben, in Abhängigkeit gewisser Merkmale, **unterschiedliche Inklusionswahrscheinlichkeiten**

## Arten von Zufallsstichproben

### Einfache Zufallsstichprobe:

- Einstufiger Ziehungsprozess, in dem jedes Element gleich behandelt wird (*Merkmale* der Elemente aus der Auswahlgesamtheit spielen keine Rolle)
- Jede mögliche Stichprobe und auch jedes Element besitzen **die-selbe Chance, ausgewählt zu werden** (gleiche Inklusionswahrscheinlichkeiten)

## Geschichtete Zufallsstichprobe:

- Die Elemente der Grundgesamtheit werden vor der Ziehung anhand von *Merkmalen* in Gruppen (Schichten) eingeteilt
- **Jedes Element der Grundgesamtheit** muss hierbei **exakt zu einer Schicht** gehören
- Dann wird **für jede Schicht** gesondert eine einstufige Stichprobe mit gleichen Inklusionswahrscheinlichkeiten innerhalb der Schicht gezogen
- Unterscheidung proportionale vs. disproportionale Schichtung

## Klumpenstichprobe:

- Eine einstufige Zufallsauswahl von **zusammengefassten Elementen** (Klumpen, Cluster) mit gleichen Inklusionswahrscheinlichkeiten
- Innerhalb der gezogenen Clusters werden Daten zu **allen Ziel-elementen** erhoben

## Mehrstufige Auswahlverfahren:

- Die Grundgesamtheit wird in **hierarchisch zueinander stehende Stufen** geteilt (z.B. räumlich: Stadt, Stadtteil, Stimmbezirk, Straße, Haushalt, Person; oder: Schule, Schulklasse, Schulkind)
- Auf jeder dieser Stufen findet ein **eigenständiger Auswahlprozess** statt



## Arten von Zufallsstichproben und Inklusionswahrscheinlichkeiten der Zielelemente (Elemente, zu denen Daten erhoben werden)

Art Stichprobe	Gleiche Inklusionswahrscheinlichkeiten der Zielelemente	Differierende Inklusionswahrscheinlichkeiten der Zielelemente
Einfache Zufallsstichprobe	Ja	Nein
Geschichtete Zufallsstichprobe	Proportionale Schichtung	Disproportionale Schichtung
Klumpenstichprobe	(Ja)*	(Nein)*
Mehrstufiges Verfahren	Nur unter ganz speziellen Bedingungen**	i.d.R. Ja

\*Gleiche Inklusionswahrscheinlichkeiten wären nur dann nicht gegeben, wenn die Klumpenelemente nach dem Prinzip der disproportionalen Schichtung gezogen werden (unwahrscheinlich).

\*\* Das ist nur dann der Fall, wenn innerhalb einer Stufe, die weiter differenziert wird, jedes Element aus gleich vielen Unterelementen besteht und die Elemente auf jeder Stufe mit gleichen Inklusionswahrscheinlichkeiten gezogen werden. Beispiel:

Haushalte stellen die erste Stufe dar. Jeder Haushalt besitzt dieselbe Inklusionswahrscheinlichkeit. Dann wird innerhalb der gezogenen Haushalte eine Person (zweite Stufe, Zielelement) gezogen, wobei **innerhalb des Haushaltes** jede Person ebenfalls die gleiche Inklusionswahrscheinlichkeit besitzt. Damit ist aber noch lange nicht gesichert, dass die **haushaltsübergreifende** Inklusionswahrscheinlichkeit für jede Person gleich ist (was gleiche Inklusionswahrscheinlichkeiten der Zielelemente bedeuten würde). Denn das würde nur dann zutreffen, wenn jeder Haushalt **aus derselben Anzahl an Personen** bestünde. Das ist ja bekanntlich hier, und auch in den meisten anderen Beispielen aus der Praxis, nicht der Fall. So hat in dem Beispiel eine Person aus einem großen Haushalt eine kleinere Inklusionswahrscheinlichkeit inne, als eine Person aus einem kleinen Haushalt.

## Willkürliche und bewusste Auswahlen

- **Willkürliche Auswahlen** (Auswahlen aufs „Geratewohl“): Die Entscheidung über die Aufnahme eines Elementes der Grundgesamtheit in die Stichprobe erfolgt **unkontrolliert und liegt nur im Ermessen des Auswählenden**
- **Bewusste Auswahlen** (Auswahlen nach „Gutdünken“): Erfolgen zwar **nach einem Auswahlplan**, d.h. die diesem Plan zugrunde liegenden Kriterien sind meist angebbbar
  - dennoch sind inferenzstatistische Techniken nicht anwendbar, da **keine Zufallsziehung** stattfindet

## Arten von bewussten Auswahlen

- **Auswahl extremer Fälle:** Selektion derjenigen Fälle (z. B. Leistungssportler), die in Bezug auf ein bestimmtes Merkmal eine extreme Ausprägung besitzen
- **Auswahl typischer Fälle:** Auswahl von als „charakteristisch“ für die Grundgesamtheit angesehenen Objekten; das Problem liegt in der Angabe der Kriterien, was ein typischer Fall ist

- **Konzentrationsprinzip:** Auswahl derjenigen Fälle, bei denen ein interessierendes Merkmal so stark ausgeprägt ist, dass diese Fälle **nahezu die gesamte Verteilung** in der Grundgesamtheit bestimmen
- **Schneeball-Verfahren:**
  - Eine untersuchte Person **benennt weitere Personen** nach einem vom Forscher bestimmten **Kriterium**
  - Diese benannten Personen werden untersucht und wiederum gebeten, ihrerseits **weitere Personen anzugeben** usw.
    - ⇒ sinnvoll bei Merkmalen, welche **sehr selten** auftauchen oder zur **Erfassung von Netzwerken**

- **Quotaverfahren:**

- Auswahl derart, dass der **Anteil bestimmter Merkmalsausprägungen** in der Stichprobe exakt dem **Anteil in der Grundgesamtheit entspricht**
- Die **endgültige Auswahl** der einzelnen Objekte bleibt der **Willkür der durchführenden Person** vorbehalten → keine Zufallsauswahl!
- Erhöhte Auswahlwahrscheinlichkeit für Personen, die **gut erreichbar** sind, oder dem Auswähler „**gut gefallen**“
  - ⇒ das Quotaverfahren stellt somit keine Lösung des Problems der Ausfälle dar – **Das Problem wird nur verdeckt!**

## Zufallsstichprobenziehungsverfahren in der wissenschaftlichen Umfragepraxis

- Stichprobenziehung aus Registern (Telefonbuch, Einwohnermeldeamt, Register der Agentur für Arbeit, Ausländerzentralregister etc.)
- Telefonische Stichproben: „Gabler Häder“-Verfahren, Handynummer-Stichproben, „last-birthday-Methode“
- Persönlich-mündliche Stichproben: Random-Walk-Verfahren
- Online-Stichproben: Online vs. Offline-Rekrutierungsstrategien

## Nonresponse

- **Unit-Nonresponse:** Die zu befragende Person antwortet gar nicht / verweigert die Teilnahme an der Befragung
- **Item-Nonresponse:** Die Person antwortet nur auf einzelne (u. U. heikle) Fragen nicht

## Problem niedriger „Ausschöpfungsquoten“:

- **Ausschöpfungsquote:** Verhältnis der **tatsächlich realisierten Untersuchungen** (= Netto) zu der **Anzahl der Elemente in der ursprünglich gezogenen Stichprobe** (= Brutto)
- Je größer Unit-Nonresponse, desto geringer ist die Ausschöpfungsquote
- Je geringer die Ausschöpfungsquote, desto größer können Schlussfolgerungsfehler (Stichprobenfehler) sein



## Problem von systematischen Nonresponse-Fällen:

- **Systematisches Nonresponse:** Nicht-Teilnehmer, die sich **systematisch** in den interessierenden Variablen von den Teilnehmern unterscheiden
  - ➔ angestrebte und die Inferenzpopulation sind **nicht mehr deckungsgleich**

## Mögliche Ursachen von Unit-Nonresponse:

- **Nicht-Befragbare** – Personen, welche aufgrund körperlicher Gebrechen oder schwerer akuter Erkrankung aus der Stichprobe ausscheiden
- **Schwer-Erreichbare** – Trotz mehrfacher Kontaktversuche nicht an ihrem Wohnsitz angetroffenen Personen; die Nicht-Erreichbarkeit kann mit bestimmten anderen Merkmalen des Personenkreises zusammenhängen (z.B. stressiger Beruf)
- **Verweigerer** – Personen, welche die Teilnahme an einer Befragung (aus diversen Gründen) explizit verweigern

## Typische (geäußerte) Verweigerungsgründe:

- Zeitmangel (auch keine Terminfindung möglich)
- Keine Lust, kein Interesse, "habe Wichtigeres zu tun"
- Bedenken bzgl. Datenschutz / Anonymität / Stichprobenziehung
- Angst vor Verkäufen
- Angst vor heiklen Fragen
- Befürchtung, zu dem Thema nichts zu sagen zu haben
- „Immer diese ständigen Befragungen“
- Diese Befragung bringt nichts, Zweck dieser Befragung unklar
- Generelle Resignation: Umfragen ändern nichts
- Prinzipiell keine Teilnahme
- Person hat "oft genug" an Befragungen teilgenommen
- Person fühlt sich zu alt
- Weiterleitung auf das Mobiltelefon
- Anderer Modus (z.B. online oder postalisch) gewünscht
- Wortlose Verweigerung