

# Auswahlverfahren

Verfahren, welche die prinzipiellen  
Regeln zur Konstruktion von  
Stichproben angeben

# Definition der Grundgesamtheit

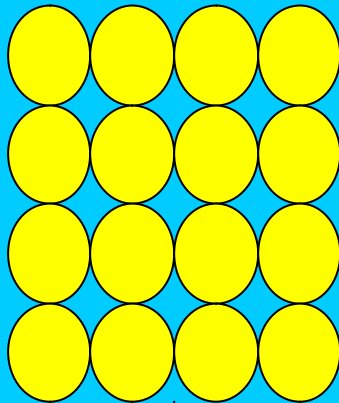
- **Untersuchungseinheit:** Objekt an dem Messungen vorgenommen werden
- **Grundgesamtheit** (Population): Menge von Elementen (Untersuchungseinheiten), über die Aussagen im Rahmen einer Untersuchung gemacht werden sollen
- **Schwierigkeit:** Für beschreibende Studien ist sie relativ einfach festzulegen. Bei Theorie testenden Fragestellungen ist es oft nicht genau klar, auf welche Objekte sich die Hypothesen beziehen

# Vollerhebung und Stichprobe

- **Vollerhebung:** die Daten aller Elemente einer Grundgesamtheit werden erhoben
- **Teilerhebung:** nur eine Teilmenge der Grundgesamtheit wird untersucht
- **Stichprobe:** Die Elemente der Teilerhebung werden durch Regeln bestimmt, die vor der Untersuchung festgelegt werden

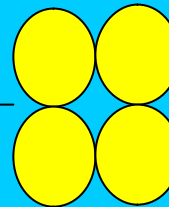
- Die **Auswahlregeln** können auf zufälligen, willkürlichen oder bewussten Auswahlprozessen fußen.
- **Lediglich Zufallsstichproben erlauben die Anwendung inferenzstatistischer Techniken, d. h. nur für sie sind die Fehler, die beim Schluss von der Stichprobe auf die Grundgesamtheit entstehen, berechenbar!**
- **Statistiken:** Maßzahlen, welche sich auf die Stichprobe beziehen
- **Parameter:** Maßzahlen, welche sich auf die Grundgesamtheit beziehen

Grundgesamtheit



n-Ziehungen

Stich-  
probe



Berechnung der  
Stichproben-  
statistiken

Rückschluss auf die  
Grundgesamtheits-  
parameter



# Voll- und Teilerhebung

- **Vor- und Nachteile der Vollerhebung**

- Alle Informationen liegen unverfälscht vor
- Kosten steigen überproportional mit dem Umfang der Population
- Bei destruktiven Tests unsinnig

- **Vor- und Nachteile der Teilerhebung**

- Fehlerhafter Induktionsschluss beim Ableiten der Populationsparameter aus den Stichprobenstatistiken
- Relativ kostengünstig
- Stichprobenergebnisse können genauer sein

# Angestrebte Grundgesamtheit, Auswahlgesamtheit und Inferenzpopulation

- **Angestrebte Grundgesamtheit** umfasst alle Elemente, für die entsprechende Theorie gelten soll
- **Auswahlgesamtheit** umfasst alle Elemente, die eine prinzipielle Chance haben, in eine Stichprobe zu gelangen
- **Inferenzpopulation**: Ist die Grundgesamtheit, über die auf der Basis der vorliegenden Stichprobe tatsächlich Aussagen gemacht werden können bzw. ist die Grundgesamtheit, aus welcher die vorliegende Stichprobe tatsächlich eine Zufallsstichprobe darstellt
- **Angestrebte Grundgesamtheit und Inferenzpopulation sollten möglichst genau übereinstimmen!**

# Zufallsstichproben

- **Einfache Zufallsstichproben:** jede mögliche Stichprobe und auch jedes Element besitzen dieselbe Chance ausgewählt zu werden
- **Geschichtete Zufallsstichproben:** Die Elemente der Grundgesamtheit werden so in Gruppen (Schichten, strata) eingeteilt, dass jedes Element der Grundgesamtheit zu einer – und nur zu einer – Schicht gehört. Danach werden einfache Zufallsstichproben aus jeder Schicht gezogen.
- **Klumpenstichprobe:** eine einfache Zufallsauswahl, bei der die Auswahlregeln nicht auf die Elemente der Grundgesamtheit, sondern auf zusammengefasste Elemente (Klumpen, Cluster) angewendet werden und dann jeweils die Daten *aller* Elemente des ausgewählten Clusters erhoben werden
- **Mehrstufige Auswahlverfahren:** Kombination mehrerer Auswahlverfahren; die Grundgesamtheit wird zuerst in Primäreinheiten eingeteilt, welche die erste Auswahlgrundlage liefern. Aus diesen Primäreinheiten wird dann eine Zufallsstichprobe der Sekundäreinheiten gezogen, die dann bereits die Erhebungseinheiten bilden können. Oft folgen noch weitere Auswahlstufen, welche auf den Sekundäreinheiten (bzw. Tertiäreinheiten usw.) basieren.



# 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 sogar meist angebbar und überprüfbar, dennoch sind inferenzstatistische Techniken nicht anwendbar. Es gibt viele verschiedene Arten bewusster Auswahlen.

# Bewusste Auswahlen

- **Auswahl extremer Fälle:** Selektion derjenigen Fälle (z. B. Leistungssportler), die in Bezug auf ein bestimmtes Merkmal eine extreme Ausprägung besitzen
- **Auswahl typischer Fälle:** „charakteristisch“ für die Grundgesamtheit angesehene Objekte; 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:** beginnt mit einer Person. Diese benennt weitere und zwar aufgrund eines vom Forscher bestimmten Kriteriums, wie z. B. „Wen würden sie in beruflichen Angelegenheiten um Hilfe bitten?“. Diese benannten Personen müssen jetzt ihrerseits angeben wer ihnen aufgrund des Kriteriums in den Sinn kommt.
- **Quotaverfahren**

**Quotaverfahren** basieren auf der Auswahl von Personen in der Art, dass bestimmte Merkmale in der Stichprobe exakt in derselben Häufigkeit vorkommen wie in der Grundgesamtheit; die endgültige Auswahl bleibt dem Interviewer vorbehalten.

Quotaverfahren lassen sich in solche mit unabhängiger oder kombinierter Quote unterscheiden

Quotenvorgaben (in %)

Unabhängige Quote			
	prot.	kath.	Summe
männlich	45	5	50
weiblich	5	45	50
Summe	50	50	100

Kombinierte Quote			
	prot.	kath.	Summe
männlich	25	25	50
weiblich	25	25	50
Summe	50	50	100

# Kritikpunkte am Quotaverfahren

- Quota-Verfahren stellen keine Zufallsauswahlen dar; dadurch ist die Verallgemeinerung erschwert, wenn nicht sogar unmöglich
- Das Verfahren bedingt eine erhöhte Auswahlwahrscheinlichkeit für Personen, die häufig an ihrem Wohnsitz anzutreffen sind
- Verzerrungsgefahr durch die Interviewer, da diese vor allem Personen auswählen werden, die sich kooperativ verhalten. Das Quotaverfahren stellt darum keine Lösung des Problems der Ausfälle dar: Das Problem wird nur verdeckt

# Nonresponse

- **Unit-Nonresponse:** die zu befragende Person antwortet gar nicht
- **Itemnonresponse:** die Person antwortet nur auf einzelne (u. U. heikle) Fragen nicht
- **Problem von Nonresponse:** Unterscheiden sich die Nichtantwortenden in irgendeiner Weise systematisch in den interessierenden Variablen von denen die antworten? Wenn ein Unterschied vorläge, würden die angestrebte und die Inferenzpopulation schon deswegen nicht mehr deckungsgleich sein. Dieses Problem besteht nicht, wenn die Ausfälle stichprobenneutral sind.
- **Ausschöpfungsquoten:** Je stärker die tatsächlich realisierte (die Anzahl der tatsächlich gezogenen Elemente) von der zu realisierende Stichprobe (die Anzahl der Angedachten Elemente) abweicht, desto größer können die Fehler sein. Je größer der Unit-Nonresponse, desto geringer ist die Ausschöpfungsquote.

# Ursachen und Folgen 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 angetroffene Personen. Offensichtlich hängt das nicht erreichbar sein mit bestimmten anderen Merkmalen des Personenkreises zusammen.
- **Verweigerer:** Personen, welche eine Befragung explizit verweigern