

STICHPROBEN ***&*** ***UMFRAGEN***

Grundlagen der Stichprobenziehung

Marcus Hudec
Christian Neumann



Institut für Statistik der Universität Wien

Vorwort

„Wir leben im Jahrhundert der Statistik!“ lautet der Leitsatz von STAT 4 U, dem Projekt zur Vermittlung statistischer Grundkenntnisse für 10-18jährige. Tatsächlich werden neue Erkenntnisse immer mehr aufgrund von systematischen Auswertungen gesammelter Daten gewonnen. Wichtige Entscheidungen werden in den verschiedensten Bereichen aufgrund dieser Erkenntnisse getroffen.

Das Ziehen von Stichproben stellt ein wichtiges Werkzeug zur Gewinnung von Informationen über eine große Zahl von Merkmalsträgern dar.

Das vorliegende Arbeitsheft zum Thema *Stichproben und Umfragen* gliedert sich in neun Kapitel:

In den ersten beiden Kapiteln werden die Grundbegriffe der Stichprobenziehung vorgestellt und der Sinn von Stichprobenziehungen diskutiert.

Das dritte Kapitel beschäftigt sich mit der Frage, mit welchen Fehlern man bei der Ziehung einer Stichprobe rechnen muss und wie man diesen Fehler möglichst gering halten kann.

Im vierten Kapitel werden verschiedene Arten von Stichproben vorgestellt und deren unterschiedliche Eigenschaften erläutert.

Thema des fünften Kapitels ist der statistische Hintergrund der Stichprobenziehung. Hier wird die zentrale Frage behandelt, wie und mit welcher Sicherheit man Rückschlüsse von Stichprobeninformationen auf die Zusammensetzung der Grundgesamtheit ziehen kann.

Darauf aufbauend widmet sich das sechste Kapitel dem notwendigen Stichprobenumfang, um Ergebnisse mit einer bestimmten Sicherheit und Genauigkeit erzielen zu können.

Im siebenten Kapitel wird die Durchführung einer auf einer Stichprobe basierenden Untersuchung Schritt für Schritt anhand eines Beispiels erläutert, im achten Kapitel wird auf die speziellen Eigenheiten einer Umfrage näher eingegangen.

Das neunte Kapitel ist schließlich der Lösungsteil aller Arbeitsaufgaben.

Im Anhang finden sich einige Kommentare zu Inhalt und Verwendung dieses Arbeitshefts.

Inhalt

1.	Einleitung.....	4
2.	Grundlagen.....	5
2.1.	Was ist eine Stichprobe?.....	5
2.2.	Wofür zieht man Stichproben?.....	7
3.	Stichprobenfehler.....	10
3.1.	Systematischer Fehler.....	10
3.2.	Zufälliger Fehler.....	10
4.	Arten von Stichproben.....	12
4.1.	Willkürliche Auswahl.....	12
4.2.	Auswahl nach Gutdünken.....	12
4.3.	Zufallsauswahl.....	12
4.3.1.	Einfache Zufallsauswahl.....	13
4.3.2.	Geschichtete Zufallsauswahl.....	13
5.	Statistischer Hintergrund.....	15
5.1.	Beispiel.....	15
5.2.	Schätzen von Anteilen.....	15
5.3.	Anwendung.....	22
6.	Stichprobenumfang.....	23
7.	Durchführen einer Stichprobenuntersuchung.....	24
7.1.	Definition des Untersuchungszieles.....	24
7.2.	Definition der Grundgesamtheit.....	24
7.3.	Auswahl der zu erhebenden Merkmale.....	25
7.4.	Gewünschte Genauigkeit.....	25
7.5.	Messmethoden.....	25
7.6.	Probeerhebungen.....	26
7.7.	Auswahl der Stichprobenart.....	26
7.8.	Organisation der Erhebung.....	26
7.9.	Auswertung der Daten.....	27
7.10.	Analysieren der Untersuchung.....	27
8.	Durchführen einer Umfrage.....	28
8.1.	Definition des Untersuchungszieles.....	28
8.2.	Definition der Grundgesamtheit.....	28
8.3.	Auswahl der zu erhebenden Merkmale.....	28
8.4.	Gewünschte Genauigkeit.....	29
8.5.	Messmethoden.....	29
8.5.1.	Art der Fragen.....	29
8.5.2.	Wie soll befragt werden?.....	30
8.6.	Probeerhebungen.....	32
8.7.	Auswahl der Stichprobenart.....	32
8.8.	Organisation der Erhebung.....	32
9.	Lösungsteil.....	33
	Kommentar.....	35

1. Einleitung

Das Ziehen von Stichproben ist im Alltag so weit verbreitet, dass man sich meistens kaum bewusst ist, dass man selbst gerade eine Stichprobe zieht oder Teil einer Stichprobenuntersuchung geworden ist.

Wenn man zum Beispiel beim Kochen „kostet“, ob die Suppe bereits genug gesalzen ist, zieht man eine Stichprobe. Und man ist gut beraten, zwischen dem Salzen und dem Kosten umzurühren, da man sonst leicht zu einer verzerrten Stichprobe und damit zu einem verfälschten (und gar nicht wohlschmeckenden) Resultat gelangen könnte.

Grundsätzlich geht es beim Ziehen von Stichproben darum, Informationen über eine Grundgesamtheit von Personen oder Objekten zu erhalten. In der Praxis ist es nicht immer möglich bzw. wirtschaftlich sinnvoll, für eine Untersuchung alle Personen oder Objekte der Grundgesamtheit heranzuziehen. Vielmehr geht es oft darum, aufgrund von Untersuchungen an einem Teil (einer Stichprobe) auf die Eigenschaften des Ganzen zu schließen (vgl. Abbildung 1). Unter welchen Umständen dieser Schluss möglich ist und was dabei berücksichtigt werden muss, erklärt die Stichprobentheorie, die in diesem Arbeitsheft erläutert wird.

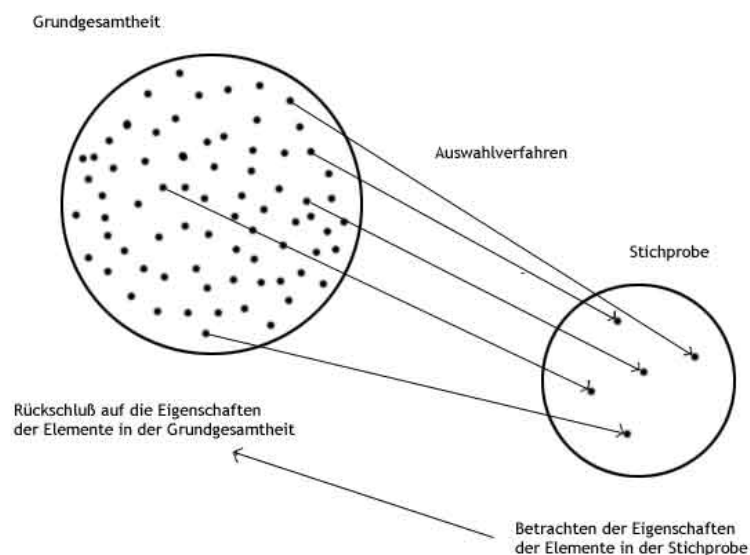


Abbildung 1

Eine der häufigsten Anwendungen der Stichprobenziehung ist die Auswahl der zu befragenden Personen im Rahmen einer Umfrage. Die Ergebnisse einer Umfrage unter einer relativ kleinen Gruppe von Personen sollen Rückschlüsse auf die Meinung einer größeren Personengruppe (etwa alle Wahlberechtigten einer Stadt) geben können. Die Besonderheiten einer Umfrage werden in einem eigenen Kapitel dargestellt. Dabei wird vor allem auf in der Praxis auftauchende Probleme näher eingegangen.

2. Grundlagen

2.1. Was ist eine Stichprobe?

Die sogenannte **Grundgesamtheit** ist jene Menge von Personen oder Objekten, über die man Informationen gewinnen will.

Wird jedes einzelne Element dieser Gruppe auf die gewünschte Information hin untersucht, spricht man von einer **Vollerhebung**.

Meist ist es allerdings aus unterschiedlichen Gründen nicht möglich, die ganze Grundgesamtheit zu untersuchen. In solchen Fällen wird eine **Stichprobe** gezogen. Eine Stichprobe ist eine Teilmenge der Grundgesamtheit. Aufgrund der Eigenschaften der Elemente, die sich in der Stichprobe befinden, soll auf die Eigenschaften aller Elemente der Grundgesamtheit geschlossen werden. Wenn dies möglich ist, bezeichnet man die Stichprobe als repräsentativ für die Grundgesamtheit.

Beispiel: Bei der Auslieferung von Glühbirnen ist ein gewisser Anteil an defekten Glühbirnen zulässig. Da nicht jede einzelne Glühbirne getestet werden kann, wird eine Stichprobe gezogen. Aus der Anzahl an defekten Glühbirnen in der Stichprobe kann auf die Anzahl der defekten Glühbirnen in der gesamten Lieferung geschlossen werden.

Beispiel: In einer Umfrage möchte eine Wochenzeitschrift herausfinden, was „die Österreicher“ zu einer geplanten Steuerreform sagen. Es wäre natürlich viel zu teuer, tatsächlich alle Österreicher in kurzer Zeit zu befragen. So wird nur ein kleiner Teil befragt.





- 1) Liegt in folgenden Beispielen eine Vollerhebung oder eine Stichprobe vor? Benenne die Grundgesamtheit!
- a) Ein Getränkeautomatenhersteller überlegt die Aufstellung eines Automaten in einer Schule. Um den Bedarf herauszufinden, spricht ein Vertreter der Firma mit allen Klassensprechern der Schule und befragt sie über die Trinkgewohnheiten der Schüler.
Vollerhebung oder Stichprobe?
Grundgesamtheit: _____
- b) Eine Automobilzeitschrift möchte herausfinden, ob die Autofahrer ihr Fahrverhalten aufgrund der hohen Benzinpreise geändert haben. An einigen großen Tankstellen werden dazu Umfragen veranstaltet.
Vollerhebung oder Stichprobe?
Grundgesamtheit: _____
- c) Im Zuge der Volkszählung werden alle Österreicher im Abstand von 10 Jahren bezüglich mehrerer Merkmale untersucht.
Vollerhebung oder Stichprobe?
Grundgesamtheit: _____
- d) Ein Redakteur der Zeitschrift „Gourme“t schreibt einen Bericht über die Pizzeria „Dolce Vita“ und probiert eine Tomatensuppe, Lasagne und ein Tiramisu.
Vollerhebung oder Stichprobe?
Grundgesamtheit: _____
- e) Im Zuge der Inventur notieren die Angestellten der Firma „Lager-Löf“ alle vorhandenen Waren mit Stückzahl und Wert.
Vollerhebung oder Stichprobe?
Grundgesamtheit: _____
- f) Karl soll einen Aufsatz mit mindestens 2.000 Wörtern schreiben. Nachdem er 5 Seiten geschrieben hat, zählt er die Anzahl der Wörter auf der ersten Seite. Da die erste Seite bereits 420 Wörter enthält, schätzt er, dass der Aufsatz lang genug sei.
Vollerhebung oder Stichprobe?
Grundgesamtheit: _____

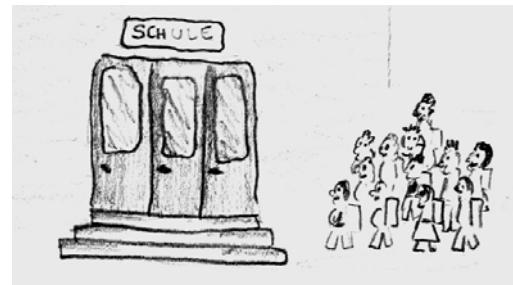
2.2. Wofür zieht man Stichproben?

Wie bereits in den Beispielen weiter oben erläutert, ist es oft aus zeitlichen oder wirtschaftlichen Gründen nicht möglich, eine Vollerhebung durchzuführen.

Eine Grundgesamtheit ist meist so groß, verstreut, schwer zugänglich oder kurzlebig, dass eine Vollerhebung gar nicht möglich ist. Es kann sogar sein, dass eine Vollerhebung ungenauere Ergebnisse liefert als eine sorgfältig geplante Stichprobenerhebung!

Im folgenden werden einzelne Gründe für die Stichprobenziehung anhand von Beispielen näher erläutert.

Wirtschaftlichkeit: Soll zum Beispiel das Telefonierverhalten aller Schüler untersucht werden, wäre es vielleicht sogar theoretisch möglich, alle Schüler einzeln zu befragen. Die Durchführung einer solchen Befragung wäre aber sehr teuer. Demgegenüber kann bei einer gut angelegten Umfrage mit einem kleinen Teil der Schüler ein gutes Ergebnis erzielt werden.



Zeitfaktor: Es kann vorkommen, dass es aus Zeitgründen nicht möglich ist, die gesamte Grundgesamtheit zu untersuchen. Ein etwas drastisches Beispiel wäre etwa eine Seuche, die in einem kleinen Ort ausgebrochen ist und sich schnell ausbreitet. Wenn man erst alle möglicherweise Angesteckten untersuchen will und alle Ergebnisse abwarten muss, hat sich die Seuche höchstwahrscheinlich schon sehr weit verbreitet. In diesem Fall muss die Untersuchung einiger Patienten reichen, um Rückschlüsse auf mögliche Behandlungsmethoden ziehen zu können.



Große Grundgesamtheiten: Ab einer gewissen Größe der Grundgesamtheit wird es nicht nur unwirtschaftlich, sondern unmöglich, alle einzelnen Elemente zu untersuchen. Soll etwa das Telefonieverhalten aller Europäer untersucht werden, können nicht mehr wirklich alle, die ein Telefon besitzen, befragt werden.

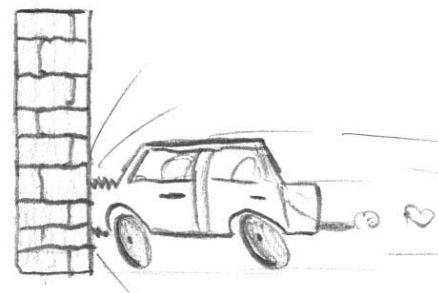


Nicht zugängliche Grundgesamtheiten: Unter Umständen kann es vorkommen, dass dem Forscher gewisse Teile der Grundgesamtheit aus verschiedensten Gründen von vornherein nicht zugänglich sind. Möglicherweise kann man aber aus einer Stichprobe auch auf diesen Teil der Grundgesamtheit schließen.



Zerstörung durch Beobachtung: Bei der Qualitätskontrolle wird bei der Ziehung einer Stichprobe das betreffende Element oft zerstört oder zumindest unbrauchbar gemacht. Als Beispiel dient hier etwa der Crashtest bei PKWs. Diese sollen so gefertigt sein, dass die Insassen bei einem Aufprall mit einer gewissen Geschwindigkeit keine ernsthaften Verletzungen davontragen.

Natürlich wird der einzelne PKW beim Crashtest zerstört. In diesem Fall ist es offensichtlich, dass nur ein kleiner Teil aller Autos einem Crash-Test unterzogen werden kann. Vom Ergebnis dieser Stichprobe wird dann auf die Beschaffenheit aller Fahrzeuge gleicher Bauart geschlossen.





- 2) Würdest du in folgenden Fällen eine Vollerhebung oder eine Stichprobe empfehlen? Begründe deine Entscheidung!
- a) In einer Schulklasse sollen die Schüler über das Ausflugsziel des nächsten Wandertages entscheiden.
 - b) Eine Wochenzeitschrift will herausfinden, wie Nationalratswahlen ausgehen würden, fände sie am nächsten Wochenende statt (sog. „Sonntagsfrage“).
 - c) Verschiedene Marken von Autoreifen sollen einem Vergleichstest unterzogen werden.
- 3) Nenne ein Beispiel, für das eine Vollerhebung aus einem der folgenden Gründe nicht sinnvoll ist!
- a) Wirtschaftlichkeit
 - b) Zeitfaktor
 - c) Große Grundgesamtheiten
 - d) Nicht zugängliche Grundgesamtheiten
 - e) Zerstörung durch Beobachtung
- 4) Nenne ein Beispiel, in dem du eine Vollerhebung gegenüber der Stichprobenziehung bevorzugen würdest!

Aus der Praxis: Jeder Kaufmann muss einmal pro Jahr eine Inventur, d.h. eine Bestandsaufnahme aller seiner im Geschäft befindlichen Waren, durchführen. Bei großen Betrieben bzw. Lagern bedeutet dies eine enorme zeitliche und finanzielle Belastung. Aus diesem Grund hat der Gesetzgeber festgelegt, dass anstelle einer Vollinventur auch eine Inventur auf Stichprobenbasis erfolgen darf.

3. Stichprobenfehler

Jede aus einer Stichprobe gewonnene Information über die Grundgesamtheit ist mit einem Fehler behaftet. Dieser Fehler wird aufgrund der Vorteile einer Stichprobenziehung gegenüber einer Vollerhebung in Kauf genommen. Die Größenordnung des Fehlers kann unter gewissen Umständen, die im fünften Kapitel diskutiert werden, angegeben werden.

Bei den Fehlern im Rahmen von Stichprobenziehungen kann man zwei verschiedene Arten unterscheiden.

3.1. Systematischer Fehler

Dieser Fehler beruht auf einem mangelhaften Stichprobenplan, d.h. die Stichprobe wurde so gezogen, dass sie von vornherein als verzerrt (d.h. nicht repräsentativ) einzustufen ist.

Beispiel: Ein Reiseveranstalter will anhand einer Umfrage herausfinden, wie groß der Anteil der Deutschen ist, die gerne in den Alpen Urlaub machen. Wenn er diese Umfrage unter deutschen Touristen in Tirol macht, wird er wohl einen unverhältnismäßig hohen Anteil an deutschen Urlaubern erhalten, die gern in den Alpen urlauben. Unter den deutschen Urlaubern, die in Tirol Urlaub machen, wird der Anteil derer, die Berge gern haben systematisch größer sein als etwa unter den Deutschen, die nach Mallorca auf Urlaub fahren.

Beispiel: Bei einer Messung kann ein systematischer Fehler durch ungenaue Messgeräte zustande kommen. Wenn eine Waage ohne Belastung bereits 10 kg anzeigt, wird jede Wägung systematisch um 10kg zuviel anzeigen.

In der Praxis beruhen „sensationelle Umfrageergebnisse“ häufig auf solchen systematischen Fehlern. Man kann diesen Fehler nur durch sorgfältiges Vorgehen in allen Phasen der Stichprobenziehung bzw. bei Messungen durch das Verwenden geeichter und kalibrierter Messgeräte gering halten.

3.2. Zufälliger Fehler

Es kann passieren, dass in einer Stichprobe zufälligerweise gerade eine gegenüber der Grundgesamtheit überproportionale Anzahl von Elementen vorhanden ist, die eine bestimmte untersuchte Eigenschaft aufweist.

Beispiel: In einer Wohnsiedlung sind von 2.000 Bewohnern 1.500 (=75%) für den Bau einer Tiefgarage, 500 (=25%) sind dagegen. Für eine Befragung werden 200 Bewohner zufällig ausgewählt.

Welche Anteile an Befürwortern und Gegnern sind in der Stichprobe möglich?

Die 200 Befragten könnten alle aus der Gruppe der Befürworter kommen.

Die 200 Befragten könnten auch alle aus der Gruppe der Gegner stammen.

Es wäre auch möglich, dass 150 Personen (75%) der Befragten in der Stichprobe für den Bau und 50 Personen (25%) gegen den Bau der Tiefgarage sind. In diesem Fall würde der Anteil in der Stichprobe genau dem Anteil in der Grundgesamtheit entsprechen.

Wenn die Stichprobe zufällig gezogen wird, sind nicht alle Zusammensetzungen der Stichprobe gleich wahrscheinlich. Zum Glück ist der letzte Fall viel wahrscheinlicher als die ersten beiden extremen Zusammensetzungen.

Die Gesetze der Wahrscheinlichkeitstheorie erlauben es, die Wahrscheinlichkeiten der verschiedenen Zusammensetzungen zu berechnen. Man kann im konkreten Fall eine Größenordnung angeben, in welchem Intervall sich dieser Fehler höchstwahrscheinlich bewegt. Wie dies genau funktioniert, wird im fünften Kapitel erklärt.



5) *Ein berühmtes Beispiel für eine verzerrte Stichprobe lieferte eine bekannte amerikanische Zeitschrift im Präsidentenwahlkampf 1936 zwischen dem Republikaner Alf Landon und dem Demokraten Franklin Roosevelt. Über 10 Millionen(!) Fragebögen wurden ausgesendet, die Adressen der Befragten wurden aus Telefonbüchern und KFZ-Zulassungsregistern entnommen. Über 2,4 Millionen Personen beantworteten den Fragebogen mit dem Ergebnis, dass Alf Landon einen klaren Sieg erwarten dürfte: 57% der Befragten würden für ihn stimmen. Am Wahltag sah alles aber ganz anders aus: Franklin Roosevelt gewann mit 63% der Stimmen ganz klar vor seinem Konkurrenten.*

- a) *Was war die Grundgesamtheit für diese Umfrage?*
- b) *Was die Stichprobe?*
- c) *Habt ihr eine Idee, warum das Stichprobenergebnis dermaßen vom tatsächlichen Ergebnis abweichen konnte?*

4. Arten von Stichproben

Man kann verschiedene Arten von Stichproben danach unterscheiden, wie die Elemente aus der Grundgesamtheit ausgewählt werden.

4.1. Willkürliche Auswahl

In der willkürlichen Auswahl finden sich jene Elemente der Stichprobe wieder, die für den Zieher der Stichprobe am leichtesten zugänglich waren. Diese Art der Auswahl ist ziemlich sicher mit unkontrollierbaren Fehlern behaftet und generell abzulehnen.

Beispiel: Erinnern wir uns an den Reiseveranstalter, der die deutschen Urlauber befragen wollte, ob sie gerne in den Alpen Urlaub machen. Wenn er diese Umfrage in Tirol veranstaltet nur weil er sich selbst gerade dort befindet und ihm das am angenehmsten ist, ist dieses Ergebnis sicher nicht auf alle Deutschen übertragbar.

4.2. Auswahl nach Gutdünken

Diese Art der Auswahl wird in der Markt- und Meinungsforschung oft angewendet, gibt aber auch Anlass zu Auseinandersetzungen. Elemente der Grundgesamtheit werden aufgrund verschiedener Kriterien als „typisch“ für die Grundgesamtheit eingestuft und deshalb genauer untersucht. Es bleibt allerdings die Frage offen, ob die Elemente wirklich in allen untersuchten Eigenschaften als typisch für die Grundgesamtheit eingestuft werden können und ob die Auswahltechnik nicht gerade zu jenen Ergebnissen tendiert, die die Auswählenden erwartet.

Beispiel: Der Reiseveranstalter führt seine Befragung im Ruhrgebiet durch, da dort der „typische“ Deutsche wohnt. Es kann aber leicht sein, dass gerade das Urlaubsverhalten der Einwohner des Ruhrgebiets nicht typisch für ganz Deutschland ist.

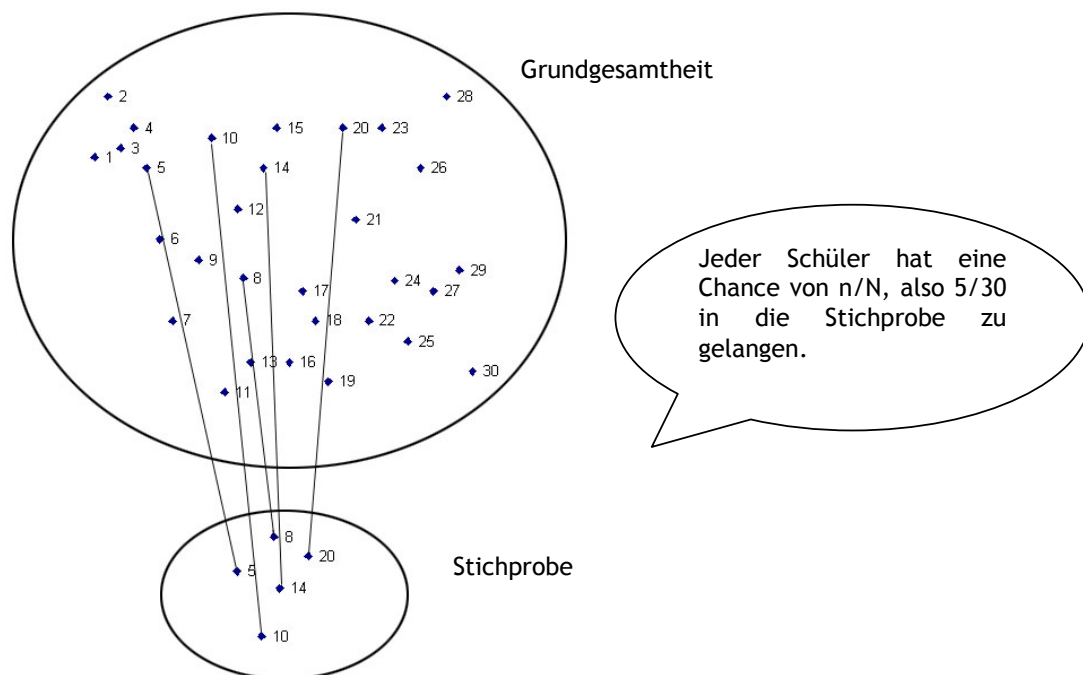
4.3. Zufallsauswahl

Bei einer Zufallsauswahl erhält jedes Element der Grundgesamtheit eine gewisse Wahrscheinlichkeit, ausgewählt zu werden. Die Zufallsauswahl ist eine in der Praxis häufig angewendete Art der Stichprobenziehung, da man in diesem Fall aufbauend auf Erkenntnisse der Wahrscheinlichkeitstheorie Aussagen über die Genauigkeit bzw. Sicherheit der Resultate machen kann. Eine sorgfältig durchgeführte Stichprobenuntersuchung auf der Basis der Zufallsauswahl liefert zudem repräsentative Ergebnisse.

4.3.1. Einfache Zufallsauswahl

Bei der einfachen Zufallsauswahl wird aus allen N Elementen der Grundgesamtheit eine Stichprobe von n Elementen gezogen. Alle n Stichprobenelemente haben dieselbe Wahrscheinlichkeit, ausgewählt zu werden.

Beispiel: Aus einer Klasse von $N = 30$ Schülern sollen $n = 5$ Schüler ins Schülerparlament entsendet werden. Eine einfache Zufallsauswahl könnte folgendermaßen aussehen: Jedem Schüler wird eine Zahl zugewiesen. Diese Zahl wird auf einem Zettel notiert, der Zettel wird gefaltet und alle Zetteln kommen in einen Topf. Nun werden fünf Zetteln blind gezogen. So erhält man eine mögliche Stichprobe von 5 Schülern.

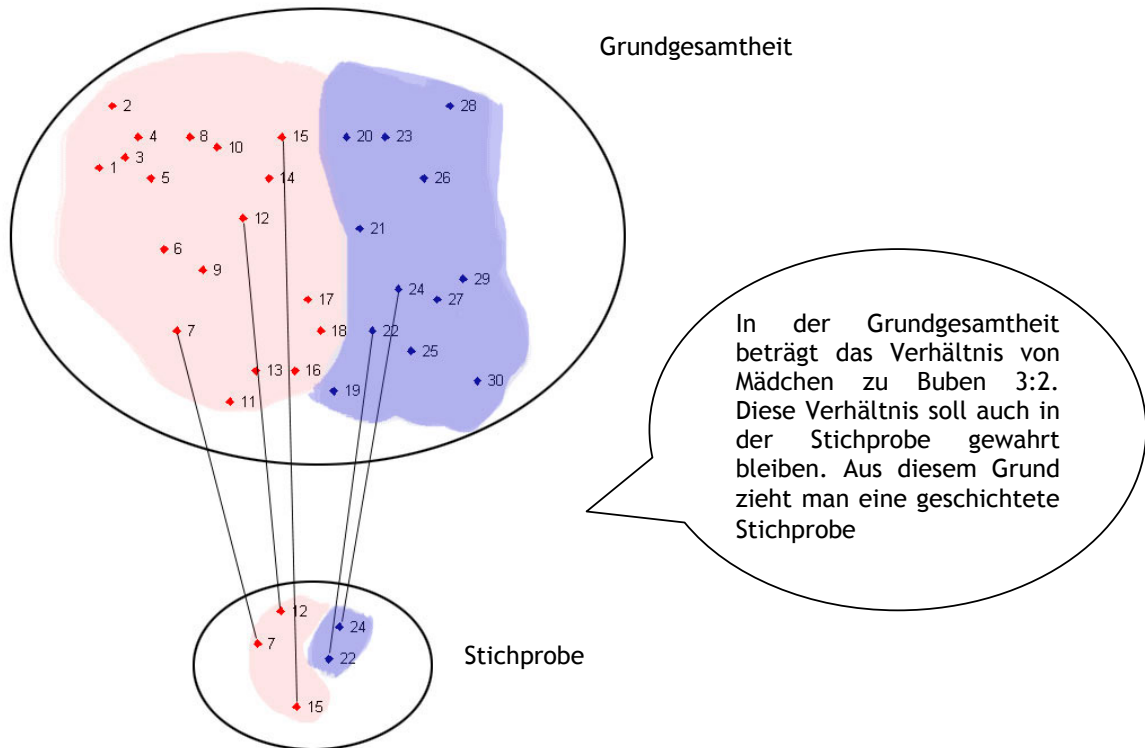


4.3.2. Geschichtete Zufallsauswahl

Bei der geschichteten Zufallsauswahl wird die Grundgesamtheit zunächst in verschiedene Teile (sogenannte Schichten) unterteilt und danach aus jeder Schicht einige Elemente für die Stichprobe entnommen. Diese Vorgangsweise soll dazu dienen, dass in der Stichprobe gewisse Teile der Grundgesamtheit nicht zu stark oder zu schwach vertreten sind.

Beispiel: Die Klasse aus dem vorigen Beispiel wird von 18 Mädchen und 12 Buben besucht. Einerseits soll die Auswahl der Abgesandten ins Parlament zufällig erfolgen, andererseits soll aber auch in der Gesandtschaft das Geschlechterverhältnis der Klasse

gewahrt bleiben. Aus diesem Grund sollen drei Mädchen und zwei Buben ausgewählt werden. Die Grundgesamtheit (alle Schüler der Klasse) wird demnach in zwei Schichten (Mädchen und Buben) geteilt. Innerhalb der Schichten wird dann wieder eine einfache Zufallsauswahl vorgenommen.



- 6) Welcher Stichprobenart würden folgende Auswahlkriterien entsprechen? Werden die Ergebnisse voraussichtlich verzerrt sein? In welche Richtung?
- Ein Interviewer verwendet ein drei Jahre altes Telefonbuch als Auswahlgrundlage für eine Umfrage zur Zufriedenheit in einer Wohnsiedlung.
 - Karl soll für ein Schulprojekt eine Umfrage zur Nutzung des Internets für Kommunikation machen. Er schickt den Fragebogen an alle, die im Adressbuch seines Mailprogramms stehen und bittet, den Fragebogen ausgefüllt zurückzusenden.
 - Eine große Tageszeitung will wissen, wie zufrieden ihre Abonnenten mit dem Service sind. 100 Abonnenten werden aus einer Liste ausgewählt und zu einem Abendessen eingeladen, bei dem sie auch ausführlich über verschiedene Merkmale der Zeitung befragt werden.

5. Statistischer Hintergrund

5.1. Beispiel

In einer Gemeinde leben 5.000 Einwohner. Der Bürgermeister möchte herausfinden, ob diese für oder gegen den Bau einer geplanten Mülldeponie sind. Dazu wird eine Stichprobe von 100 Personen gezogen und diese werden nach ihrer Meinung befragt. 30 Personen (=30%) sind für den Bau der Deponie, 70 Personen (=70%) sind dagegen.

In diesem Beispiel möchte man den Anteil der Befürworter bzw. Gegner der Mülldeponie unter den Bürgern der Gemeinde herausfinden. Aus dem Anteil der Stichprobe soll auf den Anteil in der Grundgesamtheit geschlossen werden.

Die zentrale Frage lautet: Wie gut kann man Rückschlüsse vom Anteil in der Stichprobe auf den Anteil in der Grundgesamtheit ziehen?

5.2. Schätzen von Anteilen

Allgemein formuliert steht man bei der Schätzung von Anteilen vor folgendem Problem: In einer Grundgesamtheit mit N Elementen beträgt die Anzahl der Merkmalsträger mit einer bestimmten Merkmalsausprägung eines Merkmals M .

Der Anteil dieser Merkmalsträger an allen Elementen beträgt demnach $p = \frac{M}{N}$.

Beispiel: In unserem Beispiel hat die Grundgesamtheit einen Umfang von $N=5000$. Der Anteil M der Befürworter der Deponie ist unbekannt. Somit ist auch der Anteil der Befürworter $p = \frac{M}{5000}$ unbekannt.

Man zieht eine Stichprobe vom Umfang n und betrachtet in der Stichprobe die Anzahl m mit obiger Merkmalsausprägung des Merkmals.

Beispiel: Unsere Stichprobe hat den Umfang $n=100$. Die Anzahl m der Befürworter in der Stichprobe beträgt $m=30$. Der Anteil der Befürworter in der Stichprobe beträgt also

$$\hat{p} = \frac{m}{n} = \frac{30}{100} = 0,3 \text{ oder } 30\%.$$

Der Stichprobenanteil \hat{p} wird nun als Schätzwert für den Anteil p in der Grundgesamtheit verwendet. Man kann zwar nicht mit Sicherheit sagen, ob \hat{p} ein guter Schätzwert für p ist, aber man kann immerhin ein Intervall angeben, in dem der wahre Anteil p bei gegebenem \hat{p} mit großer Wahrscheinlichkeit liegen wird.

Annahme

Um zu zeigen, wie man die auf der vorigen Seite aufgeworfene Frage beantworten kann, gehen wir von folgender Annahme aus:

Wir nehmen an, dass wir in unserem Beispiel den wahren Anteil von Befürwortern und Gegnern der Mülldeponie in der gesamten Gemeinde kennen.

Von den $N = 5.000$ Bürgern sind $M = 1.500$ (=30%) für den Bau der Deponie und $(N-M) = 3.500$ (=70%) gegen den Bau der Deponie. Anhand einer Stichprobe von $n=100$ Personen soll der Anteil \hat{p} der Befürworter geschätzt werden.

Dieser Fall kommt in der Praxis der Stichprobenziehung natürlich nicht vor. Wenn man den wahren Anteil der Befürworter in der Grundgesamtheit bereits kennt, braucht man keine Stichprobe mehr zu ziehen, um diesen zu schätzen. Wir wollen es hier zur Erläuterung aber trotzdem tun.

In der Stichprobe von $n = 100$ Personen ist theoretisch jede Aufteilung von Befürwortern und Gegnern möglich. Da wir die tatsächliche Aufteilung kennen, können wir die theoretischen Wahrscheinlichkeiten berechnen, mit der jede Aufteilung von Befürwortern und Gegnern in der Stichprobe auftritt. Für die Berechnung dieser Wahrscheinlichkeiten benötigt man die hypergeometrische Verteilung.

Hypergeometrische Verteilung

Die hypergeometrische Verteilung gibt die Wahrscheinlichkeit an, dass in einer Stichprobe vom Umfang n genau m Elemente eine bestimmte Merkmalsausprägung haben, wenn in der Grundgesamtheit N genau M Elemente diese Merkmalsausprägung besitzen.

Die Formel zur Berechnung der Wahrscheinlichkeiten der hypergeometrischen Verteilung lautet:

$$p(X = m) = \frac{\binom{M}{m} \binom{N-M}{n-m}}{\binom{N}{n}}$$

Diese Formel entspricht dem Prinzip „Günstige durch Mögliche“:

Bei einer Grundgesamtheit von N Elementen hat man $\binom{N}{n}$ Möglichkeiten, eine Stichprobe vom Umfang n zu ziehen („Mögliche“).

Die Günstigen entsprechen der Auswahl von m aus insgesamt M Befürwortern sowie $n-m$ aus insgesamt $N-M$ Gegnern.

Wenn man also wissen will, wie groß die Wahrscheinlichkeit ist, dass sich bei gegebener Grundgesamtheit in unserer Stichprobe genau 30 Befürworter der Mülldeponie befinden, setzt man

$$p(X = 30) = \frac{\binom{1500}{30} \binom{3500}{70}}{\binom{5000}{100}} \text{ und erhält eine Wahrscheinlichkeit von etwa 9\%.}$$

Nun kann man diese Wahrscheinlichkeiten natürlich nicht nur für 30 Befürworter, sondern für alle möglichen Anzahlen an Befürwortern von 0-100 berechnen.

m	p(X=m)	m	p(X=m)	m	p(X=m)	m	p(X=m)
0	0,0000						
1	0,0000	26	0,0614	51	0,0000	76	0,0000
2	0,0000	27	0,0723	52	0,0000	77	0,0000
3	0,0000	28	0,0810	53	0,0000	78	0,0000
4	0,0000	29	0,0864	54	0,0000	79	0,0000
5	0,0000	30	0,0877	55	0,0000	80	0,0000
6	0,0000	31	0,0848	56	0,0000		
7	0,0000	32	0,0783	57	0,0000		
8	0,0000	33	0,0690	58	0,0000		
9	0,0000	34	0,0581	59	0,0000		
10	0,0000	35	0,0467	60	0,0000		
11	0,0000	36	0,0360	61	0,0000	86	0,0000
12	0,0000	37	0,0265	62	0,0000	87	0,0000
13	0,0000	38	0,0187	63	0,0000	88	0,0000
14	0,0001	39	0,0126	64	0,0000	89	0,0000
15	0,0002	40	0,0082	65	0,0000	90	0,0000
16	0,0005	41	0,0051	66	0,0000	91	0,0000
17	0,0011	42	0,0030	67	0,0000	92	0,0000
18	0,0022	43	0,0017	68	0,0000	93	0,0000
19	0,0041	44	0,0009	69	0,0000	94	0,0000
20	0,0073	45	0,0005	70	0,0000	95	0,0000
21	0,0120	46	0,0003	71	0,0000	96	0,0000
22	0,0186	47	0,0001	72	0,0000	97	0,0000
23	0,0273	48	0,0001	73	0,0000	98	0,0000
24	0,0377	49	0,0000	74	0,0000	99	0,0000
25	0,0494	50	0,0000	75	0,0000	100	0,0000

Die Wahrscheinlichkeit, dass in der Stichprobe genau 37 Befürworter sind, wenn sich in der Grundgesamtheit 1.500 Befürworter und 3.500 Gegner befinden, beträgt 2,65%.

Tabelle 1

In Abbildung 2 sind die Daten aus der Tabelle in einem Diagramm dargestellt.

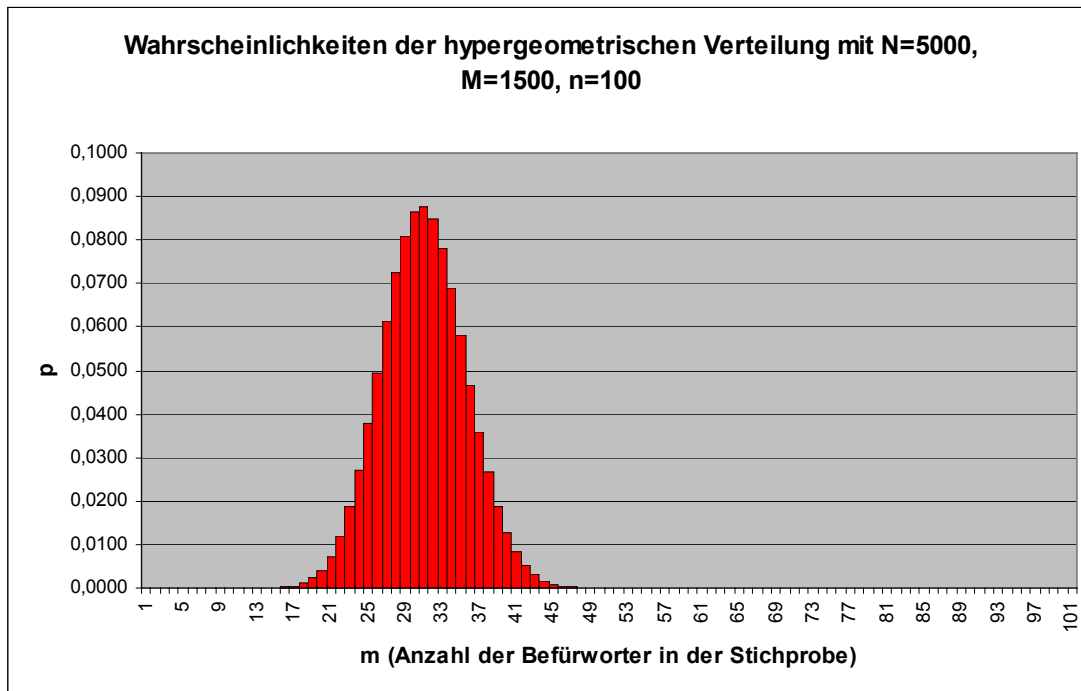


Abbildung 2

Aus der Tabelle und dem Diagramm kann man einige Fragen beantworten:

- Wie groß ist die Wahrscheinlichkeit, in der Stichprobe genau denselben Anteil an Befürwortern (30%) und Gegnern (70%) zu erhalten wie in der Grundgesamtheit?
 $p(X=30) = 0,0877$ also nur etwa 9%.
- Wie groß ist die Wahrscheinlichkeit, in der Stichprobe einen Anteil zu erhalten, der halbwegs in der Nähe des Anteils der Grundgesamtheit liegt, etwa zwischen 20% und 40% Befürwortern?
 Dazu müssen die Wahrscheinlichkeiten für 20, 21, 22, ..., 38, 39, 40 Befürworter addiert werden.
 $p(X=20) + p(X=21) + \dots + p(X=39) + p(X=40) = 0,9799$, also etwa 98%.
- Wie groß ist die Wahrscheinlichkeit, in der Stichprobe mehr Gegner als Befürworter zu haben?
 Dazu müssen die Wahrscheinlichkeiten für 51, 52, 53, ..., 98, 99, 100 Gegner addiert werden.
 $p(X=50) + p(X=51) + \dots + p(X=99) + p(X=100) = 0,0000073$, also etwa 0,00073%.

Wir sehen also: Die Wahrscheinlichkeit, in der Stichprobe genau denselben Anteil wie in der Grundgesamtheit zu haben, ist nicht besonders groß.

Auf der anderen Seite ist es sehr unwahrscheinlich, dass die Aufteilung in der Stichprobe sehr weit von der tatsächlichen Aufteilung in der Grundgesamtheit entfernt liegt.

Bereits jetzt kann man vielleicht erkennen, wohin die Erklärung führen wird: Selbst wenn man den wahren Anteil nicht kennt, weiß man, dass bei einem Stichprobenanteil von zum Beispiel 80% Befürwortern in einer Stichprobe vom Umfang 100 ein wahrer Anteil von 30% Befürwortern dermaßen unwahrscheinlich ist, dass man dies beruhigt ausschließen kann.

Exkurs: Die folgenden beiden Hinweise stellen Möglichkeiten dar, unter bestimmten Voraussetzungen eine Verteilung durch eine andere Verteilung anzunähern. So gelangt man schließlich zur Normalverteilung, mit der man im allgemeinen leichter rechnen kann als etwa mit der hypergeometrischen Verteilung.

Approximation der hypergeometrischen Verteilung durch die Binomialverteilung:

Die Binomialverteilung ist das Verteilungsmodell zur Darstellung des „Ziehens mit Zurücklegen“. Wenn eine Münze öfters geworfen wird entspricht die Wahrscheinlichkeit p , genau n Erfolge zu erzielen, einer Binomialverteilung mit den Parametern n und p .

Die hypergeometrische Verteilung ist dagegen das Modell zur Darstellung des „Ziehens ohne Zurücklegen“. Auch das obige Beispiel entspricht dieser Vorstellung. Sobald ein Bürger „gezogen“ und befragt wurde, kann er nicht nochmals drankommen.

Bei einer entsprechend großen Grundgesamtheit spielt es nicht mehr eine so große Rolle, ob das Ziehen mit oder ohne Zurücklegen stattfindet. Die Wahrscheinlichkeit, dass ein Element zweimal gezogen wird ist so gering, dass man sie einfach vernachlässigt. Aus diesem Grund kann bei großen Grundgesamtheiten statt der hypergeometrischen Verteilung (mit Parametern N , M , n und m) die Binomialverteilung (mit Parametern n und $p=M/N$) verwendet werden.

Approximation der Binomialverteilung durch die Normalverteilung:

Aber auch die Binomialverteilung kann durch eine noch einfacher zu handhabende Verteilung gut beschrieben werden, nämlich die Normalverteilung. Voraussetzung dafür ist ein genügend großer Stichprobenumfang, als Faustregel gilt: $np(1-p) > 9$.

Variation der Annahme

Wie bereits erwähnt, ist in der Praxis die wahre Aufteilung in der Grundgesamtheit unbekannt.

Trotzdem wollen wir noch ein letztes Mal annehmen, dass wir den wahren Anteil der Grundgesamtheit kennen (25% Befürworter, 75% Gegner).

Nun wollen wir noch vor der Stichprobenuntersuchung wissen, in welchem Bereich der Stichprobenanteil höchstwahrscheinlich liegen wird (vgl. Abbildung 2). Dieses „höchstwahrscheinlich“ setzen wir für diese Untersuchung mit $(1-a) = 95\%$ fest, die Irrtumswahrscheinlichkeit a beträgt also 5%.

Zur Berechnung des Stichprobenanteils verwenden wir nun die Normalverteilungsapproximation (vgl. Exkurs). Die allgemeine Formel lautet

$$P\left(n \cdot \frac{M}{N} - z_{1-\frac{a}{2}} \cdot \sqrt{n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right)} < X < n \cdot \frac{M}{N} + z_{1-\frac{a}{2}} \cdot \sqrt{n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right)}\right) = 1 - a$$

Mit den entsprechenden Werten des Beispiels ergibt sich die Formel zu

$$P\left(100 \cdot 0,3 - 1,96 \cdot \sqrt{100 \cdot 0,3 \cdot 0,7} < X < 100 \cdot 0,3 + 1,96 \cdot \sqrt{100 \cdot 0,3 \cdot 0,7}\right) = 0,95$$

$$P(21,02 < X < 38,98) = 0,95$$

Das heißt, dass bei einer Stichprobe vom Umfang $n=100$ die Anzahl m der Befürworter zu 95% zwischen 21,02 und 38,98 liegen wird.

Nun trennen wir uns von der Annahme, dass wir den wahren Anteil kennen und berechnen das obige Intervall für *alle möglichen Anteile in der Grundgesamtheit* von 0-100%.

Abbildung 3 zeigt alle Intervalle für eine Vielzahl von Anteilen der Grundgesamtheit.

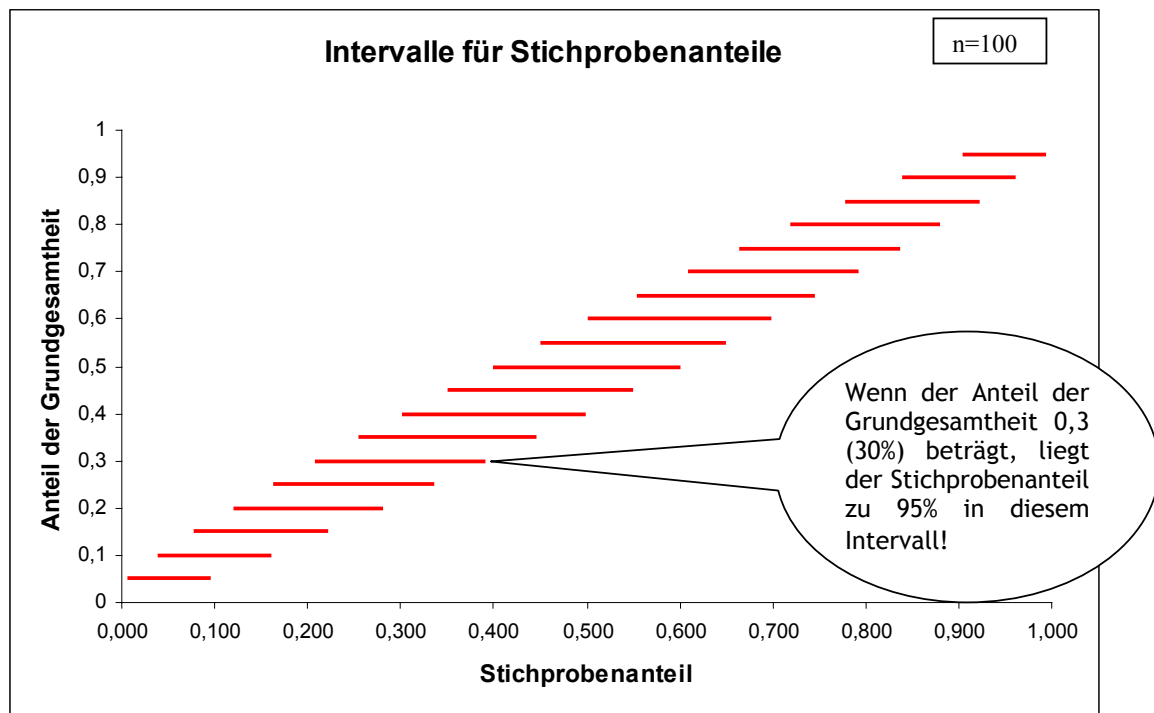


Abbildung 3

Man hat nun festgestellt, dass in der Stichprobe 30 Befürworter sind, der Stichprobenanteil der Befürworter beträgt also 30%. Trägt man eine senkrechte Linie für einen Stichprobenanteil mit diesem Wert ein, kann man ablesen, in welchem Intervall der Anteil in der Grundgesamtheit zu 95% liegt (Abbildung 4).

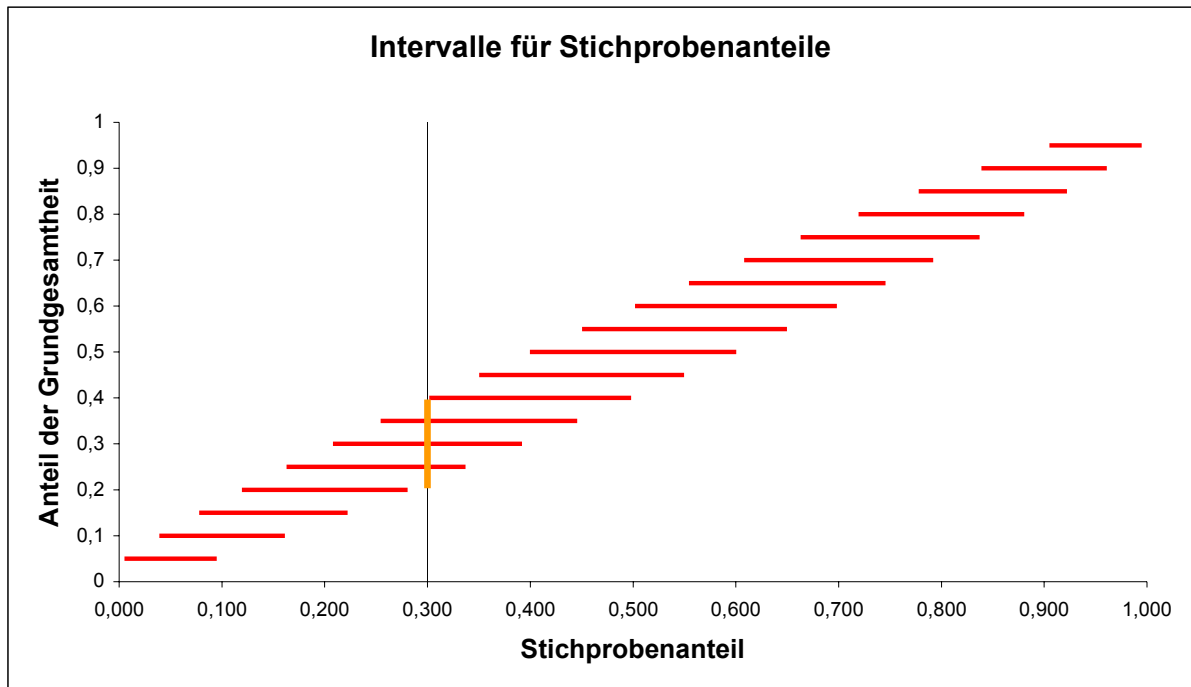


Abbildung 4

Abbildung 4 zeigt die Lösung des Problems: Wenn der Stichprobenanteil einer Stichprobe vom Umfang $n = 100$ gerade 30% beträgt, so liegt der Anteil der Grundgesamtheit mit einer Wahrscheinlichkeit von 95% im Intervall von (21,02% ; 38,98%)

Formal kann man sagen: Liefert eine Stichprobe den Anteilswert \hat{p} , so liegt der Anteilswert p der Grundgesamtheit mit einer Irrtumswahrscheinlichkeit α im Intervall

$$\left(\hat{p} - z \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}; \hat{p} + z \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right).$$

Der Wert z gibt das $1-\alpha/2$ - Quantil der Standardnormalverteilung an.

Lösung des Beispiels - Zusammenfassung

In einer Gemeinde leben 5.000 Einwohner. Der Bürgermeister möchte herausfinden, ob diese für oder gegen den Bau einer geplanten Mülldeponie wären. Dazu wird eine Stichprobe von 100 Personen gezogen und diese werden nach ihrer Meinung befragt. 30 Personen sind für den Bau der Deponie, 70 sind dagegen.

Wie groß ist der Anteil in der Grundgesamtheit bei einer Irrtumswahrscheinlichkeit von 5%?

$$\left(0,3 - 1,96 \sqrt{\frac{0,3(1-0,3)}{100}}; 0,3 + 1,96 \sqrt{\frac{0,3(1-0,3)}{100}} \right) \Rightarrow (21,02; 38,98)$$

Der Anteil in der Grundgesamtheit liegt zu 95% im Intervall (21,02 ; 38,98)

5.3. Anwendung

Beispiel: Der Bürgermeister einer Gemeinde möchte die Meinung der Einwohner zu einem geplanten Freizeitpark wissen. Aus den 7000 Bürgern der Gemeinde wird eine Stichprobe vom Umfang $n = 300$ gezogen. Die Anzahl der Befürworter in der Stichprobe beträgt 160. In welchem Intervall liegt der Anteil der Grundgesamtheit zu

a) 95% bzw. zu b) 99%?

Der Stichprobenanteil beträgt $160/300=0,533$. Das $1-\alpha/2$ - Quantil der Standardnormalverteilung beträgt bei

a) 1,96 bzw. bei b) 2,58.

Eingesetzt in die obige Formel ergibt dies

$$\text{a) } \left(0,533 - 1,96 \sqrt{\frac{0,533(1-0,533)}{300}}; 0,533 + 1,96 \sqrt{\frac{0,533(1-0,533)}{300}} \right) = (0,477; 0,590)$$

$$\text{b) } \left(0,533 - 2,58 \sqrt{\frac{0,533(1-0,533)}{300}}; 0,533 + 2,58 \sqrt{\frac{0,533(1-0,533)}{300}} \right) = (0,459; 0,608)$$

Zu 95% liegt der Anteil der Grundgesamtheit im Intervall von 47,7% - 59,0%. Wenn man nur eine Irrtumswahrscheinlichkeit von 1% in Kauf nimmt, kann man sagen, dass der Anteil in der Grundgesamtheit zu 99% im Intervall (45,9%; 60,8%) liegt.



7) Ein Tiergartendirektor will in einer Umfrage herausfinden, ob die Mehrheit der Zoo-Besucher trotz einer bevorstehenden Erhöhung der Eintrittspreise den Zoo wieder besuchen wollen. Von 20.000 Besuchern innerhalb einer Woche werden 50 befragt. 30 (=60%) antworten mit „Ja“, 20 (=40%) mit „Nein“.

a) Kann der Direktor mit 5% Irrtumswahrscheinlichkeit sagen, dass die Mehrheit aller Besucher trotzdem wieder kommen wird?

b) Ändert sich das Ergebnis, wenn insgesamt 30.000 Besucher im Zoo waren?

c) Von den 20.000 Besuchern werden 200 befragt. 120 (=60%) antworten mit „Ja“, 80 (=40%) mit „Nein“. Kann der Direktor mit 5% Irrtumswahrscheinlichkeit sagen, dass die Mehrheit aller Besucher trotzdem wieder kommen wird?

6. Stichprobenumfang

Gleichzeitig mit der Bestimmung eines Intervalls hat man nun auch ein Werkzeug in der Hand, mit dem man den notwendigen Stichprobenumfang in einer Untersuchung bestimmen kann.

Dazu muss man zunächst zwei Dinge festlegen: Die Irrtumswahrscheinlichkeit und die gewünschte Genauigkeit.

Je größer die Genauigkeit sein soll und je kleiner die Irrtumswahrscheinlichkeit sein darf, desto größer muss die Stichprobe sein.

Indem man die hergeleitete Formel nach n auflöst, kann man auch die Größe von n angeben.

$$n \geq \frac{4 * z^2_{1-\alpha/2} * p(1-p)}{L^2}$$

In diesem Ausdruck steckt auch p , das vor der Stichprobenziehung normalerweise unbekannt ist. In diesem Fall nimmt man den „worst case“ (schlimmsten möglichen Fall) an, der bei $p=0,5$ liegt. Der Ausdruck ergibt sich dann zu

$$n \geq \frac{z^2_{1-\alpha/2}}{L^2}.$$

Beispiel: Bei einer Umfrage über die Durchführung möchte man ein Ergebnis, das mit einer Irrtumswahrscheinlichkeit von 1% eine Abweichung von maximal +/- 5% vom wahren Wert ergibt. Über die Größe von p ist von vornherein nichts bekannt.

Hier ist $\alpha = 0,01$ und $z_{1-\frac{\alpha}{2}} = 2,58$. L beträgt $0,05 + 0,05 = 0,1$.

Der Umfang der Stichprobe sollte also mindestens $n \geq \frac{2,58^2}{0,1^2} = 666$ betragen, wenn man die gewünschte Genauigkeit bei einer Irrtumswahrscheinlichkeit von 1% erreichen will. Begnügt man sich mit einer Irrtumswahrscheinlichkeit von 5% und einer Genauigkeit von +/- 10%, so reicht ein Stichprobenumfang von $n \geq \frac{1,96^2}{0,2^2} = 96$.



8) Wie viele Personen müssen befragt werden, wenn der Tiergartendirektor ein Ergebnis mit einer Genauigkeit von +/- 2% mit einer Irrtumswahrscheinlichkeit von 5% wünscht?

7. Durchführen einer Stichprobenuntersuchung

Im folgenden werden zehn Punkte aufgelistet, die bei einer Stichprobenuntersuchung beachtet werden sollten. Einige Punkte haben im Rahmen einer Umfrage eine besondere Bedeutung und werden aus diesem Grund in einem eigenen Kapitel besonders behandelt.

- Definition des Untersuchungszieles
- Definition der Grundgesamtheit
- Auswahl der zu erhebenden Merkmale
- Bestimmung der gewünschten Genauigkeit (erforderlichen Fallzahl)
- Messmethoden
- Probeerhebung
- Auswahl der Stichprobenart
- Organisation der Erhebung
- Auswertung der Daten
- Analysieren der Untersuchung

Beispiel: Für die Bewerbung bei olympischen Spielen ist die Akzeptanz der Bevölkerung wichtig. Die Landesregierung möchte herausfinden, wie es um die Zustimmung steht und möchte unter allen Einwohnern ab 14 Jahren eine Umfrage machen.

7.1. Definition des Untersuchungszieles

Die Definition des Untersuchungszieles ist zu Beginn von besonderer Bedeutung. Das Ziel sollte konkret, sinnvoll und mit den gegebenen Mitteln realisierbar sein.

Beispiel: Das Ziel in diesem Fall ist nicht klar definiert, soll aber ein Stimmungsbild in der Bevölkerung zeigen. Noch bevor die bereits relativ kostspielige Bewerbungsvorbereitung begonnen wird, soll herausgefunden werden, ob bei der Bevölkerung überhaupt Interesse an einem solchen Projekt besteht.

7.2. Definition der Grundgesamtheit

Wie bereits beschrieben, besteht die Grundgesamtheit aus der Menge der Elemente, für die man bestimmte Informationen gewinnen will. Aus der Grundgesamtheit wird eine Stichprobe (also eine Teilmenge der Grundgesamtheit) gezogen. Mitunter kann es recht schwierig sein, die Grundgesamtheit genau abzugrenzen, das heißt zu entscheiden, ob

eine Person oder ein Objekt zur Grundgesamtheit gehört oder nicht. Besonders schwierig ist das bei großen Grundgesamtheiten („die Autofahrer“, „die Raucher“, „die Bauern“, „die Senioren“, etc).

Beispiel: Im Beispiel besteht die Grundgesamtheit aus allen Einwohnern des Landes ab 14 Jahren.

7.3. Auswahl der zu erhebenden Merkmale

Dieser Punkt stellt die Umsetzung des Untersuchungszieles in die konkrete Untersuchungssituation dar. Die zu erhebenden Fragen bzw. zu stellenden Fragen sollen alle der Erreichung des Zieles dienen. Aus diesem Grund ist es wichtig, die Fragestellungen auf den Zusammenhang mit dem Ziel zu überprüfen.

Eine besonders wichtige Rolle spielt die Auswahl der Fragen im Zusammenhang mit der Gestaltung von Fragebögen in einer Umfrage. Aus diesem Grund wird auf diesen Punkt in einem eigenen Abschnitt näher eingegangen.

Beispiel: Die Fragen sollten sich tatsächlich nur um die Bewerbung bei den olympischen Spielen drehen und möglichst einfach und klar, aber doch umfassend formuliert werden. Die Bereitschaft zur Beantwortung ist sicher größer, wenn die Fragen einfach gestellt sind, man aber weiß worum es geht und worauf die Fragen hinaus wollen.

7.4. Gewünschte Genauigkeit

Stichprobenuntersuchungen sind stets mit einem gewissen Fehler behaftet. Während man mit genaueren Messmethoden oder größeren Stichproben versuchen kann, den Fehler zu minimieren, muss man auch die Kosten einkalkulieren (Zeit, Geld, geschultes Personal, usw.), die eine exakte Untersuchung mit sich bringt. Eine Abwägung dieser beiden Aspekte führt dazu, dass man jeweils überlegen muss, welche Genauigkeit für das gegebene Untersuchungsziel ausreicht bzw. erzielt werden muss.

Beispiel: Die Vertreter der Landesregierung müssen sich vor Beginn der Befragung überlegen, welche Genauigkeit das Ergebnis haben soll. Da es sich hier nur um ein allgemeines Stimmungsbild handelt und noch nicht zwingend konkrete Entscheidungen davon abhängen, wird vermutlich zunächst eine geringere Genauigkeit zugunsten geringerer Kosten in Kauf genommen.

7.5. Messmethoden

Um Informationen über eine Grundgesamtheit zu gewinnen, kann man eine Stichprobe auf verschiedene Arten untersuchen.

Einerseits kann die Art der Messmethode bestimmt werden. So könnte man zum Beispiel etwas über den Gesundheitszustand eines Menschen erfahren, indem man ihn nach seinem Wohlbefinden befragt (Befragung) oder einen Arzt bestimmte Werte, wie Blutbild oder EKG bestimmen lässt (Messung).

Andererseits stellt sich nach der Entscheidung für eine bestimmte Messmethode die Frage nach der Art der Durchführung. Dieser Punkt wird im folgenden Kapitel im Rahmen der Umfragen erläutert.

Beispiel: Hier steht natürlich die Befragung im Vordergrund.

7.6. Probeerhebungen

Es hat sich in der Praxis als sehr sinnvoll erwiesen, die geplante Messmethode zunächst in kleinem Rahmen auszuprobieren, bevor die gesamte Stichprobe untersucht wird. Verständnisprobleme sowohl auf Seiten der Erhebenden als auch potentieller Befragter können so bereits im Vorfeld ausgeräumt werden, sodass die tatsächliche Untersuchung bessere Ergebnisse liefert.

Beispiel: Der an sich fertige Fragebogen wird an 30 potentielle Befragte verteilt, die diesen ausfüllen und etwaige Unklarheiten oder Probleme melden. Aufgrund dieser Erkenntnisse können die Fragen noch umformuliert oder näher erläutert werden.

7.7. Auswahl der Stichprobenart

Je nach Fragestellung sollte die passende Stichprobenart eingesetzt werden, wobei die Faktoren Zeit, Kosten und erzielbare Genauigkeit bei gegebenem Stichprobenumfang berücksichtigt werden müssen.

Beispiel: Im Beispiel wird die einfache Zufallsstichprobe der beste Weg sein.

7.8. Organisation der Erhebung

Wird eine Untersuchung von mehreren Personen gleichzeitig durchgeführt, sollte die Arbeit von einem Projektleiter gesteuert werden. Die Erhebenden sollten alle mit der Art der Messung vertraut sein. Das bedeutet, dass sie bei einer Messung alle die Messgeräte gleich bedienen können oder bei einer Umfrage die gleichen Fragen mehr oder weniger auf dieselbe Art stellen.

Beispiel: Eine Vereinheitlichung kann erzielt werden, indem etwa die Interviewer gemeinsam eingeschult werden und auf die wichtigen Merkmale der Befragung hingewiesen werden.

7.9. Auswertung der Daten

Eine besonders wichtige Aufgabe ist die Auswertung der gesammelten Daten. In diesem Zusammenhang sollte man sich mit Fragen der Plausibilität, fehlenden Antworten, Fehlergrenzen und Gestaltung der Auswertung beschäftigen.

Beispiel: In diesem Zusammenhang ergeben sich viele interessante Fragestellungen. Unter der Plausibilitätsprüfung versteht man die Kontrolle, ob die Antworten im logischen Kontext überhaupt möglich sind. Antwortet jemand etwa auf die Frage „Betreiben Sie selbst Sport?“ mit „Nein“ und auf die darauffolgende Frage „Welche Sportarten betreiben sie?“ mit „Tennis, Radfahren und Golf“, liegt hier ein logischer Widerspruch vor. Solche widersprüchlichen Antworten werden zumeist wie fehlende Antworten behandelt.

Es wird öfters vorkommen, dass Befragte eine Frage nicht beantworten können oder wollen. Ein Problem kann dies dann darstellen, wenn gerade eine bestimmte Teilmenge nicht bereit ist, auf die Fragen zu antworten. Möglicherweise ist der Großteil der Bevölkerung eher für die Bewerbung für die olympischen Spiele, aber nicht bereit, extra einen Fragebogen auszufüllen und einzusenden. Die Gegner hingegen möchten ihren Unmut zur Geltung bringen und füllen den Fragebogen aus. So entsteht eine Verzerrung, die unter dem Fachbegriff „non-response-bias“ bekannt ist.

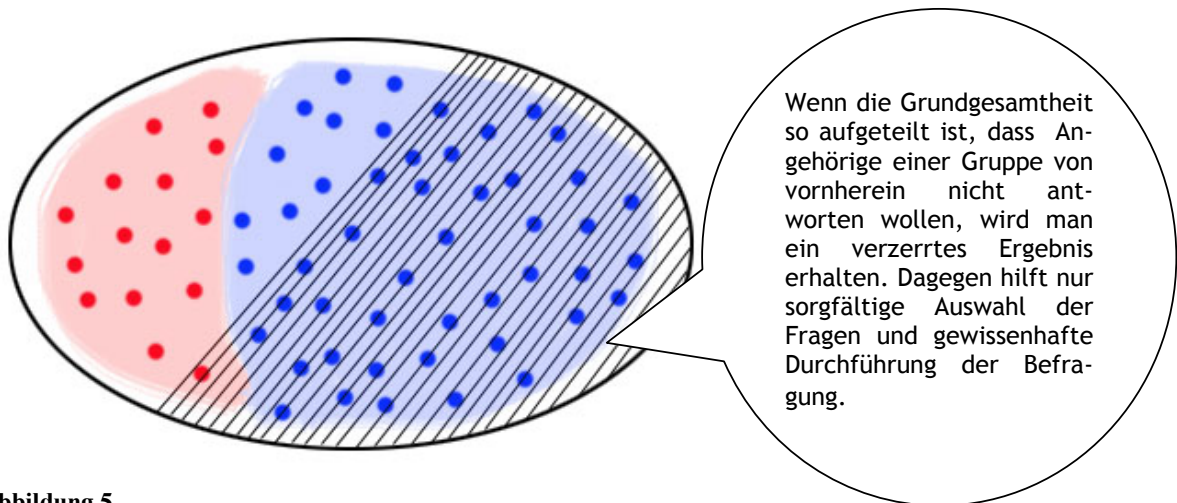


Abbildung 5

7.10. Analysieren der Untersuchung

Eine gemeinsame Nachbetrachtung aller Teilnehmer an der Untersuchung dient zur Evaluierung der vergangenen Arbeit und zur Verbesserung zukünftiger Projekte.

Beispiel: Eine gemeinsame Nachbesprechung aller Interviewer dient dem Erfahrungsaustausch, aus dem sich Verbesserungsmöglichkeiten für die Zukunft ergeben.

8. Durchführen einer Umfrage

Eine Umfrage ist in vieler Hinsicht eine besondere Anwendung einer Stichprobenuntersuchung. Aufgrund von Befragungen soll die Meinung von Personen zu unterschiedlichsten Themen herausgefunden werden.

Selbstverständlich gelten auch hier alle Punkte des vorigen Kapitels, es gibt aber noch einige Besonderheiten, die bei der Durchführung einer Umfrage berücksichtigt werden sollten.

8.1. Definition des Untersuchungszieles

Im Zusammenhang mit der Gestaltung einer Umfrage ist die Definition eines realistischen Untersuchungszieles von besonders großer Wichtigkeit. Ein strukturierter Plan sollte sich mit folgenden Fragen beschäftigen:

- Inhalt der Fragen
- Art der Fragen
- Wer soll befragt werden?
- Wie soll befragt werden?

8.2. Definition der Grundgesamtheit

Die Definition der Grundgesamtheit steht natürlich in engem Zusammenhang mit dem Untersuchungsziel. Dabei muss man sich die Frage stellen, ob es im Rahmen der gegebenen Mittel (Zeit, Geld, Personal) überhaupt möglich sein kann, sinnvolle Aussagen über die angestrebte Grundgesamtheit erlangen zu können oder ob etwa eine eingeschränktere Fragestellung aussagekräftigere Ergebnisse liefern könnte.

8.3. Auswahl der zu erhebenden Merkmale

Bei der Auswahl der zu erhebenden Merkmale muss einerseits das Untersuchungsziel im Auge behalten werden, andererseits muss man sich aber auch überlegen, wie der Inhalt der Fragen gestaltet werden muss, um ein möglichst aussagekräftiges Ergebnis zu erzielen.

Bei diesem Punkt sollte man berücksichtigen, dass die Befragten die Fragen üblicherweise nicht beantworten müssen und für die Beantwortung der Fragen freiwillig ihre Zeit opfern. Aus diesem Grund sollten die Fragen jedenfalls so aussehen, dass die Befragten leicht an der Umfrage teilnehmen können. Wenn der Befragte das Gefühl hat, den Sinn der Umfrage oder die Fragen nicht zu verstehen bzw. die Art der Befragung ablehnt, wird seine Bereitschaft zur Teilnahme an der Umfrage nicht gegeben sein.

Aus diesem Grund sollten die Fragen möglichst einfach gestellt werden. Komplizierte Satzkonstruktionen oder Fremdwörter sollten nach Möglichkeit vermieden werden.

Auch sollte es möglichst unterlassen werden, Fragen zu stellen, die die Befragten voraussichtlich nicht beantworten wollen oder können; etwa Fragen nach gesellschaftlich abgelehntem oder strafbarem Verhalten.

Beispiel: Auf die Fragen: „Bist du schüchtern?“ oder „Verkaufst du Drogen?“ wird man wohl kaum von allen eine ehrliche Antwort bekommen.

8.4. Gewünschte Genauigkeit

Hier sind vor allem zwei Fragen zu berücksichtigen:

- Wie viele Personen werden befragt?
- Wie genau werden die Personen befragt?

8.5. Messmethoden

Der Frage der Auswahl von Messmethoden wird im Rahmen von Befragungen besonderer Raum gewidmet.

- Art der Fragen
- Art der Befragung

8.5.1. Art der Fragen

Bei der Art der Fragen unterscheidet man grundsätzlich zwei Arten:

Geschlossene Fragen

Bei geschlossenen Fragen sind die Antwortmöglichkeiten bereits vorgegeben. Hier ist es wichtig, dass die verschiedenen Alternativen überschaubar, gut gegeneinander abgegrenzt sind und vor allem den gesamten Bereich möglicher Antworten abdecken. Unter den geschlossenen Fragen können noch mehrere Arten unterschieden werden.

Ja/Nein-Fragen: Hier gibt es nur eine Antwortmöglichkeit, ohne Abstufungen dazwischen; diese Fragenart ist meist einfach zu verstehen und leicht auszuwerten, schränkt aber den Spielraum der Befragung stark ein.

Beispiel: Auf die Frage: Bist du Raucher? kann man nur mit „Ja“ oder „Nein“ antworten. Diese Unterscheidung vernachlässigt den Umfang des Zigarettenkonsums und den Umstand, wie lange jemand raucht oder nicht raucht.

Mehrere Auswahlmöglichkeiten: Hier ist wieder nur eine Antwort zulässig, es gibt aber bereits mehrere Möglichkeiten für den Befragten. Wichtig ist hierbei, die Fragestellung so zu formulieren, dass Mehrfachantworten nicht vorkommen können.

Beispiel: Wieviele Zigaretten rauchst du üblicherweise pro Tag? Die Antwortmöglichkeiten erstrecken sich von 0 über 1-5, 6-10, 11-20, 21-40, mehr und erlauben so bereits eine viel bessere Unterscheidung als eine Ja/Nein-Frage.

Mehrere Antwortmöglichkeiten: Hierbei stehen mehrere Antworten zur Auswahl und es dürfen auch mehrere Antworten gegeben werden. Die Darstellung der Ergebnisse solcher Fragen muss bereits anders erfolgen, als wenn nur eine Antwortmöglichkeit gegeben wäre.

Beispiel: Wohin würdest du nächstes Jahr gerne auf Urlaub fahren?

Offene Fragen

Bei offenen Fragen dürfen die Antworten frei formuliert werden. Die Antworten sind am schwierigsten auszuwerten, da sie meist im Nachhinein in Kategorien eingeteilt werden müssen.

8.5.2. Wie soll befragt werden?

Hier unterscheidet man grundsätzlich drei Typen:

- Face to face Befragung in Form eines Interviews
- Telefonbefragung
- Briefbefragung (heute auch mittels Internet möglich)

Was sind die entscheidenden Charakteristika der unterschiedlichen Befragungstypen?

Interview

In einem persönlichen Interview treten einander ein Interviewer und ein Befragter gegenüber. Dies führt zu einer Reihe von Besonderheiten eines face-to-face Interviews (Vor- und Nachteilen), die hier kurz erläutert werden:

Der direkte Kontakt zwischen Interviewer und Befragtem bedeutet, dass der Befragte jedenfalls nicht mehr anonym ist. Auch wenn ihm die vertrauliche Behandlung der Antworten zugesichert wird, muss er sie doch gegenüber einer Person aussprechen. Das kann die Bereitschaft, wahrheitsgemäße Antworten zu geben, beeinflussen.

Die Antworten des Befragten sind von der Person des Interviewers abhängig. Die Art und Weise, wie Fragen gestellt werden, wie die verschiedenen Antwortmöglichkeiten

präsentiert werden und die Art des Gesprächsklimas zwischen Interviewer und Befragtem sind Faktoren, die das Befragungsergebnis maßgeblich beeinflussen können.

Stehen sich Interviewer und Befragter direkt gegenüber, wird der Befragte einem gewissen zeitlichen Druck ausgesetzt, da er die Fragen nicht in beliebiger Zeit beantworten kann.

Ein Vorteil der direkten Befragung besteht allerdings darin, dass dem Befragten unklar oder unverständlich formulierte Fragen nochmals mit anderen Worten erläutert werden können.

Briefbefragung

In einer brieflichen Befragung treten zunächst einmal die Aspekte des Interviews mit umgekehrten Vorzeichen auf: der Befragte bleibt anonym, er kann die Fragen in beliebiger Zeit durchlesen und beantworten, ohne irgendeiner Beeinflussung ausgesetzt zu sein. Auf der anderen Seite sind zusätzliche Informationen nicht möglich, sodass die Fragen einfach gehalten werden müssen.

Briefliche Umfragen bringen aber noch andere Nachteile mit sich:

Die Antworten sind nicht spontan, Rückfragen sind aber ausgeschlossen.

Die Fragen müssen nicht in einer bestimmten Reihenfolge beantwortet werden. Aufbauende Fragen können so nicht gestellt werden.

Die Rücklaufquote ist nicht besonders hoch.

Telefonbefragung

Die Befragung via Telefon ist wiederum ein ganz eigener Typ, der aber eher dem Interview ähnlich ist.

Aspekte der Anonymität und der Abhängigkeit vom Interviewer sind sehr ähnlich, bei der Telefonbefragung spielt meist auch der Zeitdruck eine gesteigerte Rolle. Möglichst wenige und einfach gestellte Fragen sind hier von zentraler Bedeutung. Der Vorteil der Telefonbefragung besteht darin, dass in kurzer Zeit sehr viele Personen befragt werden können, selbst wenn die Teilnehmerquote nicht sehr hoch ist.

Befragung via Internet

Eine moderne und sehr beliebte neue Befragungsart ist die Befragung via Internet. Auch hier gibt es zwei Unterscheidungen: Einerseits ist dies die Web-Poll, bei der die Betrachter auf bestimmten Seiten ihre Meinung zu einem Thema kundgeben dürfen. In diesem Fall kann der Befrager allerdings überhaupt nicht kontrollieren, wer die Fragen

beantwortet, eine Einschränkung auf einmalige Beantwortung pro Tag oder IP-Adresse ist bestenfalls möglich.

Auf der anderen Seite kann das Web natürlich genutzt werden, um Briefumfragen statt per Post via Internet durchzuführen. Das vereint die Vorteile der Briefbefragung mit zusätzlichen Vorteilen der schnellen Datenübertragung. So sind etwa Rückfragen durchaus vorstellbar, die Spontaneität und eine vorgegebene Reihenfolge der Beantwortung der Fragen bleibt aber weiterhin auf der Strecke.

8.6. Probeerhebungen

Gerade bei der Planung einer Umfrage sollte der Fragebogen an einigen Personen getestet werden. Unklar formulierte Fragen oder unerwartete Antworten können eine ordentliche Auswertung stark beeinflussen und sollten von vornherein vermieden werden.

8.7. Auswahl der Stichprobenart

Für Umfragen eignet sich eine Zufallsauswahl aus der gegebenen Grundgesamtheit normalerweise am besten. Bei bestimmten Zielsetzungen kann es sein, dass eine geschichtete Stichprobenziehung notwendig ist.

Beispiel: In einer Umfrage sollen die Einkaufsgewohnheiten von Männern und Frauen herausgefunden werden. Hier ist klar, dass man die Grundgesamtheit zunächst in Männer und Frauen aufteilen muss.

8.8. Organisation der Erhebung

Die Organisation der Erhebung wird von praktischen Gesichtspunkten geleitet. Soll etwa eine Telefonbefragung durchgeführt werden, sollten die Interviewer gut telefonieren können, etc.

9. Lösungsteil

1)

- a) *Stichprobe; Grundgesamtheit: alle Schüler der Schule*
- b) *Stichprobe; Grundgesamtheit: alle Autofahrer*
- c) *Vollerhebung; Grundgesamtheit: alle Österreicher*
- d) *Stichprobe; Grundgesamtheit: alle Speisen der Pizzeria*
- e) *Vollerhebung; Grundgesamtheit: Warenbestand der Firma*
- f) *Stichprobe; Grundgesamtheit: alle Wörter des Aufsatzes*

2)

- a) *Vollerhebung: Keiner der Gründe für eine Stichprobenziehung trifft zu*
- b) *Stichprobe: v.a. große Grundgesamtheit, Wirtschaftlichkeit*
- c) *Stichprobe: Durch den Test werden die Reifen unbrauchbar*

3)

- a) *Befragung aller Konsumenten zur Meinung zu einem neuen Produkt*
- b) *Befragung aller Fernsehzuschauer zu ihrem Wunschprogramm am Abend desselben Tages*
- c) *Qualitätskontrolle bei jeder Milchpackung*
- d) *Umfrage unter allen Einsiedlern zum Thema „Wohnen am Land“*
- e) *Belastungstest für einen neuen Werkstoff*

4) *Wenn keiner der Gründe für eine Stichprobenziehung zutrifft*

5)

- a) *Grundgesamtheit: alle Wahlberechtigten in der Präsidentenwahl*
- b) *2,4 Millionen Beantworter der Fragebögen*
- c) *1936 waren in den USA Autos sowie Telefone eher den reicheren Bevölkerungsschichten vorbehalten. Die republikanischen Wähler gehörten ebenfalls eher zu den reicheren Schichten.*

6)

- a) *Die Angabe ist nicht ganz klar: Falls der Interviewer etwa die ersten 50 Bewohner der Wohnsiedlung, die er im Telefonbuch findet, befragen will, liegt eine willkürliche Auswahl vor. Wenn er sich alle notiert und dann zufällig auswählt, liegt eben eine Zufallsauswahl vor. Auf alle Fälle wird das Ergebnis eine gewisse Verzerrung aufweisen, da ja alle in den*

letzten drei Jahren eingezogenen Personen nicht berücksichtigt werden. Je nachdem, ob diese Anzahl eher groß oder gering ist, kann es zu Abweichungen kommen.

- b) Diese Auswahlmethode ist völlig willkürlich und kommt nur dadurch zustande, dass es sich Karl besonders einfach machen will. Das Ergebnis wird verzerrt sein, da Karl ja nur gerade die Personen befragt, die das Internet bereits zur Kommunikation benutzen.
- c) Wenn die Personen zufällig aus der Liste ausgewählt wurden, liegt eine Zufallsauswahl vor. Das Ergebnis könnte unter Umständen insofern verzerrt sein, als dass die Eingeladenen wahrscheinlich eben aufgrund der Einladung eine besonders gute Meinung vom Service der Zeitung haben.

7)

- a) Nein, das Intervall aller Besucher, die wiederkommen wollen, liegt bei einer Irrtumswahrscheinlichkeit von 5% im Intervall (46,4% ; 73,6%).
- b) Nein.
- c) Ja, das Intervall liegt nun komplett jenseits der 50% - Marke. (53,2% ; 66,8%)

8) 1225

Kommentar

Zu 1. Einleitung

Dem handlungsorientierten Grundsatz von STAT4U entsprechend, soll natürlich die selbständige Durchführung von Stichprobenuntersuchungen im Vordergrund stehen. Das Arbeitsheft selbst enthält keine längeren Beispiele, sondern ist als Einführung bzw. Begleitung während einer Untersuchung gedacht. Einige Anregungen zu Stichprobenuntersuchungen befinden sich in den fachspezifischen Beispielen zu STAT4U.

Zu Aufgabe 2)

In der Praxis ist die Abgrenzung der Grundgesamtheit oft nicht trivial und erfordert einiges an Überlegung. Bei den Übungsbeispielen wurde versucht, möglichst eindeutige Beispiele zu wählen, wobei sich auch hier bei einigen Beispielen Ansatzpunkte für Diskussionen ergeben könnten.

Zu a)

Streng genommen besteht die Grundgesamtheit hier aus allen möglichen potentiellen Kunden der aufzustellenden Getränkeautomaten. Dass sich der Vertreter der Automatenfirma nur für die Trinkgewohnheiten der Schüler interessiert, liegt wahrscheinlich daran, dass er den potentiellen Getränkekonsum von Lehrern oder externen Besuchern als vernachlässigbar gering einschätzt. Der Vertreter sieht also als interessierende Grundgesamtheit nur die Menge aller Schüler.

Die Befragung der Klassensprecher stellt grundsätzlich eine Stichprobenziehung dar, da die Klassensprecher nur einen Teil der Grundgesamtheit darstellen (wie auch immer diese exakt definiert ist).

Denkbar ist natürlich, dass alle Klassensprecher vor ihrem Gespräch mit dem Vertreter alle ihre Klassenkameraden über deren Trinkgewohnheiten befragt haben. In diesem Fall könnte man von einer Vollerhebung sprechen.

Zu b)

Aus der Grundgesamtheit „die Autofahrer“ ergeben sich natürlich auch Abgrenzungsprobleme, die aber in diesem Fall keine Rolle spielen.

Zu Aufgabe 5)

Dies ist ein gutes Beispiel dafür, dass bei Vorliegen eines systematischen Fehlers in der Stichprobenziehung auch eine noch so große Stichprobe keine Garantie für ein gutes Ergebnis bietet. Eine viel kleinere, aber nach verschiedenen Merkmalen geschichtete Stichprobe wäre in diesem Fall sicherlich aussagekräftiger gewesen.