

Herausforderungen bei der Ziehung repräsentativer Stichproben in der Markt- und Sozialforschung



Prof. Dr. Raimund Wildner
Vorstandsmitglied des NIM e.V. (früher GfK Verein) sowie des BVM e.V., Honorarprofessor für Statistik und BWL, Universität Erlangen-Nürnberg

raimund.wildner@web.de

Für die Ziehung repräsentativer Stichproben fehlt aktuell die wissenschaftliche Fundierung. Für die Quotenauswahl hat eine solche Theorie noch nie existiert und für die Anwendung der Theorie der Zufallsstichprobe ist aufgrund der geringen Teilnahmebereitschaft bei Interviews die Voraussetzung entfallen. Von der Praxis entwickelte Möglichkeiten, dennoch zu qualitativ hochwertigen Stichproben zu kommen, bleiben letztlich unbefriedigend. Der Beitrag schließt mit der Forderung nach einer sachlichen Diskussion zur Entwicklung einer Stichprobentheorie, welche die Notwendigkeiten der Markt- und Sozialforschung berücksichtigt.

Schlagworte:

› Marktforschung › Repräsentativität › Quotenstichprobe › Zufallsstichprobe › Ausschöpfungsquote › Antwortverweigerung

1 Das Ziel: Repräsentative Stichproben

Es ist eine Binsenweisheit, dass selbst große Stichproben irreführend sein können, wenn sie nicht repräsentativ für die Grundgesamtheit sind. Das macht Repräsentativität so wichtig. Doch damit beginnt bereits die erste Schwierigkeit: Was heißt „Repräsentativität“?

Überraschend ist: In den meisten Lehrbüchern der Stichprobentheorie findet man den Begriff nicht. Das gilt für Cochran (1977), Govindarajulu (1999), Chaudhuri (2014) oder Arnab (2017). Für andere ist er ein Synonym für „unverzerrt“, z. B. für Chaudhuri und Stenger (2005, 12) und Thompson (2012, 110).

Schon eher fündig wird man in der Marktforschungsliteratur. In dem Gabler Lexikon Marktforschung definierte der Autor dieses Artikels: „Eine Befragung kann dann als repräsentativ bezeichnet werden, wenn es

möglich ist, unverzerrte Schätzwerte für die Grundgesamtheit zu ermitteln. Ein (häufiger) Sonderfall ist, dass die Stichprobe in allen wesentlichen Merkmalen ein verkleinertes Abbild der Grundgesamtheit ist. Bei Befragungen, bei denen die interessierenden Sachverhalte bei nur vergleichsweise kleinen Bevölkerungsanteilen auftreten ... ist es ratsam, von den relevanten Personen einen höheren Anteil in der Stichprobe zu haben, als ihrem Anteil in der Grundgesamtheit entspricht (disproportionale Stichprobe).“ (Wildner, 2004, 448).

Damit ist auch angesprochen, wie Repräsentativität in der Praxis überprüft wird: Zunächst setzt sie eine Grundgesamtheit voraus, die so definiert wird, dass von jedem existierenden Element eindeutig festgestellt werden kann, ob das Element zur Grundgesamtheit gehört oder nicht. Für diese Grundgesamtheit müssen die Verteilungen wesentlicher Strukturmerkmale bekannt sein, z. B. aus der amtlichen Statistik oder aus sonstigen, allgemein anerkannten Erhebungen wie

z. B. der Media-Analyse. Wesentliche Strukturmerkmale sind dabei solche, von denen vermutet werden kann, dass sie mit den zu erhebenden Daten korrelieren. Am Beispiel einer Erhebung über den Kauf von Nahrungsmitteln wären dies z. B. Haushaltsgröße, soziale Schicht, Region und anderes mehr. Die Stichprobe darf in ihrer Verteilung bezüglich dieser Strukturmerkmale nur unwesentlich von der Grundgesamtheit abweichen. Wird mit disproportionalen Stichproben gearbeitet, so ist dies für jede der Schichten zu fordern, für die unterschiedliche Auswahlwahrscheinlichkeiten angewendet werden. Die Erwartung dabei ist, dass Stichproben, die in den Strukturmerkmalen die Grundgesamtheit abbilden, dies auch für die weiteren erhobenen Merkmale tun und damit diese Grundgesamtheit zutreffend repräsentieren oder eben repräsentativ sind.

Geringe Abweichungen bei den Strukturmerkmalen erfordern auch eine ausreichende Stichprobengröße. Dabei lässt sich herleiten, dass die maximale Abweichung eines

Anteils bei einer Sicherheitswahrscheinlichkeit von 95 Prozent näherungsweise gleich 100 dividiert durch die Wurzel des Stichprobenumfangs ist. Bei einem Stichprobenumfang von $n=1.000$ ergibt sich, dass Anteile mit der genannten Sicherheitswahrscheinlichkeit maximal um $100/(1000^{0,5})$ Prozent = 3,2 Prozent abweichen. Dieser Stichprobenumfang wird in der Praxis auch als Untergrenze für eine repräsentative Stichprobe angesehen.

Während die Stichprobentheorie also immer von einem Schätzer für eine Variable ausgeht, ist es das Ziel der Marktforschung ein ganzes Set von Variablen möglichst unverzerrt zu schätzen. Ein Beispiel ist die Media-Analyse, bei der es wichtig ist, dass die Kontaktzahlen für alle abgefragten Medien zutreffend geschätzt werden, um Fehler bei der Buchung von Werbeplätzen und zu hohe bzw. zu niedrige Umsätze mit den Medien zu vermeiden.

Gänzlich andere Fragestellungen ergeben sich bei A/B-Tests, bei denen der Effekt einer unterschiedlichen Behandlung von Test- und Kontrollgruppen untersucht wird. Hier ist weniger die Repräsentativität der Struktur der Stichprobe das Problem, sondern mehr die Repräsentativität des gemessenen Effekts der unterschiedlichen Behandlung, weshalb hier darauf nicht eingegangen wird.

2 Wege zu repräsentativen Stichproben

Für die Ziehung einer repräsentativen Stichprobe gibt es in der Marktforschung zwei grundlegend unterschiedliche Methoden:

Die erste Methode ist die *Quotenstichprobe* (Quota-Verfahren). Quotenstichproben sind Stichproben, bei denen dem Interviewer für bestimmte Merkmale (die so genannten „Quotenmerkmale“) Verteilungen vorgegeben werden, die eingehalten werden müssen. Ansonsten ist der Interviewer bzw. die Interviewerin in der Auswahl der Zielpersonen frei. Insbesondere ist es möglich, beim Nichtantreffen einer Person diese durch eine andere Person mit gleichen Quotenmerkmalen zu ersetzen.

Abstract

Currently, there is no working scientific theory for representative sampling. There was never a theory of quota sampling, and the existing theory of random sampling no longer works because response rates are too small to meet the fundamental preconditions for random sampling. Practitioners have developed ways to get high-quality samples, but they are unsatisfactory. What is needed is a fact-based discussion to develop a sampling theory that takes into account the needs of market and social research.

Keywords: › market research › representativity › quota sampling
› random sampling › response rate › nonresponse rate

In Lehrbüchern über Stichproben wird das Quota-Verfahren meist nicht (z. B. Thompson, 2012; Chaudhuri & Stenger, 2005) oder nur sehr kurz erwähnt (z. B. Cochran, 1977). Der einfache Grund dafür: Es gibt keine Theorie der Quotenstichproben.

Dennoch hat die Praxis Qualitätskriterien für das Quoten-Verfahren entwickelt, die sich daraus ergeben, dass die Interviewer bzw. Interviewerinnen die Befragungen in ihrem erweiterten sozialen Netzen durchführen. Um Klumpungen zu vermeiden, müssen daher viele Personen für die Interviews eingesetzt sein, pro Person dürfen nur wenige Interviews durchgeführt werden und es muss sichergestellt sein, dass dieselben Personen nur mit zeitlichem Abstand (ca. sechs Monate) erneut interviewt werden. Auch sind die Quotenmerkmale so zu wählen, dass sie nicht zu schwer und nicht zu leicht zu erfüllen sind und mit dem Erhebungsgegenstand korrelieren (Kreis, Wildner & Kuß, 2021).

Solche qualitativ hochwertigen Quotenstichproben sind leichter und schneller als Zufallsstichproben zu realisieren und haben sich auch bezüglich der Qualität der Ergebnisse bewährt. Man muss Humphrey Taylor Recht geben, wenn er schreibt: „The European quota method has worked reasonable well over many years and has been widely accepted, not only by practitioners and their clients but also by many European academic researchers“. (Taylor, 1998, 978).

Die fehlende Theorie der Quotenstichproben hat zur Folge, dass Formeln für die Varianz von Mittelwerten oder Anteilen nicht existieren. In der Konsequenz können daher streng genommen auch keine Vertrauensintervalle angegeben werden. Da die Praxis auf ein Genauigkeitsmaß aber nicht verzichten kann, wird dennoch häufig ersatzweise das Vertrauensintervall angegeben, das sich bei einer einfachen Zufallsauswahl ergibt. Dies lässt sich auch durch die Ergebnisse von Untersuchungen des Allensbach-Instituts begründen, bei denen die Quotenstichprobe regelmäßig bessere Ergebnisse erzielt hat, als die einfache Zufallsauswahl (Noelle-Neumann & Petersen, 2005). Eine seriöse Ergebnisdarstellung macht jedoch auf die fehlende theoretische Begründung aufmerksam.

Eine weitere – deutlich aufwendigere Möglichkeit – ist das *Bootstrapping* oder *Resampling*. Dabei werden sehr viele (z. B. 1.000) Stichproben mit Zurücklegen aus der vorhandenen Stichprobe gezogen und der interessierende Parameter jeweils geschätzt. „Mit Zurücklegen“ bedeutet, dass jedes Element mehrfach oder auch gar nicht in die Stichprobe gelangen kann, was dazu führt, dass die Parameterschätzungen variieren. Werden diese nun der Größe nach geordnet und die kleinsten und die größten 2,5 Prozent jeweils abgeschnitten, so erhält man eine Schätzung für das 95-Prozent-Vertrauensintervall.

Die zweite Methode ist die Zufallsstichprobe, für die es mehrere Möglichkeiten gibt, wie einfache oder geschichtete Zufallsstichproben, Cluster-Stichproben oder mehrstufige Stichproben, um nur die wichtigsten zu nennen.

Für alle Verfahren der Zufallsauswahl gibt es in den genannten Lehrbüchern vollständig ausgearbeitete Theorien mit Formeln für die Schätzung des Stichprobenmittelwerts und seiner Varianz. Aber diese Theorien und diese Formeln basieren alle auf der Annahme, dass die Wahrscheinlichkeit für jedes Element der Grundgesamtheit, Teil der Stichprobe zu werden, vor der Ziehung der Stichprobe berechnet werden kann und größer als Null ist. Damit das gewährleistet ist, müssen die beide folgenden Voraussetzungen erfüllt sein:

- (1) Alle Elemente der Grundgesamtheit sind in irgendeiner Form aufgelistet.
- (2) Alle (bzw. fast alle) für die Erhebung ausgewählten Elemente der Grundgesamtheit können auch erhoben werden.

Die erste Bedingung ist für einen Teil der Methoden in der Praxis gegeben. So ermöglicht die vom ADM-Stichprobenclub bereitgestellte Infrastruktur zufällige Stichproben für persönliche Interviews (sogen. „Random-Route-Verfahren“) und für Telefoninterviews (sogen. „Gabler-Häder-Verfahren“) (ADM 2024a). Doch die persönlichen und telefonischen Interviews sind in den letzten Jahren stark zurückgegangen und finden heute nur noch bei einer Minderheit der Stichproben in der Markt- und Sozialforschung Anwendung. Dagegen hat die Bedeutung der online-Interviews stark zugenommen. Laut ADM wurden 2022 64,0 Prozent aller Interviews online durchgeführt (ADM, 2024b).

Bei den Online-Stichproben ist die Tatsache, dass nicht alle Haushalte über das Internet erreichbar sind, noch das kleinere Problem. Denn laut Statistischen Bundesamt (2022) haben immerhin 95,5 Prozent aller Haushalte einen Internetzugang. Zum Vergleich: Die Ausstattungsquote mit Festnetztelefon beträgt im Jahr 2022 nur noch 82,9 Prozent.

Problematisch ist jedoch, dass für Online-Stichproben schon die erste Bedingung zunächst nicht gegeben ist, weil es keine Liste von E-Mail-Adressen aller Personen gibt. In der Praxis hilft man sich so, dass man große Pools von Personen anwirbt, die sich grundsätzlich bereit erklärt haben, sich interviewen zu lassen. Diese Online-Access-Panels waren in 2022 die Basis für 89 Prozent aller Online-Interviews (ADM, 2024b).

Für die Stichprobenqualität ist wichtig, wie die Mitglieder eines Online-Access-Panels gewonnen werden. In der Regel geschieht dies, indem die Besucher bestimmter Websites gefragt werden, ob sie an einem solchen Panel teilnehmen wollen. Hier ist es wichtig, dass eine gute Mischung sehr vieler solcher Websites genutzt wird. Werden z. B. nur Nutzer linksliberaler Medien gefragt, dann werden die Ergebnisse einer Wahlforschung in diese Richtung verzerrt sein. Doch selbst wenn die genannte Bedingung erfüllt ist, wird man die solcherart gewonnenen Teilnehmer nicht als Zufallsauswahl aller in Deutschland lebenden Personen mit Online-Anschluss interpretieren können.

Grundsätzlich ist es auch möglich, die Teilnehmer in persönlichen oder telefonischen Interviews zu rekrutieren, die zufällig ausgewählt wurden. In der Praxis sind solcherart offline rekrutierte Panels aus Kostengründen zwar die Ausnahme. Dennoch steht hier grundsätzlich ein Weg zur Verfügung, um die erste Bedingung auch für Online-Access-Panels erfüllen zu können.

3 Geringe Teilnahmebereitschaft und die Folgen

Im nächsten Schritt ist die Erfüllung der zweiten Bedingung zu prüfen, nämlich dass ein hoher Anteil der ausgewählten Elemente auch tatsächlich in die Stichprobe kommt. An diesem Punkt ist zu definieren, was mit „in die Stichprobe kommen“ gemeint ist. Dazu reicht es nicht aus, zur Teilnahme an einer Umfrage aufgefordert zu werden. Eine Person kommt nur dann in die Stichprobe, wenn ihre Daten auch tatsächlich erhoben werden, das heißt, wenn die Person an der Umfrage auch tatsächlich teilnimmt (*effek-*

tive Stichprobe). Weigert sie sich, dann war die Wahrscheinlichkeit, die für diese Person vorab berechnet wurde, falsch und damit sind auch die aus der Stichprobe ermittelten Schätzwerte falsch. Deshalb ist eine hohe Antwortbereitschaft für die Zufallsauswahl so wichtig.

Wichtig ist also die sogenannte Ausschöpfung, d. h. der Anteil der ausgewählten Personen, die auch tatsächlich interviewt werden können. Nun kann es verschiedene Gründe geben, dass eine Person nicht erreicht werden kann, wie z. B., dass die Person nicht erreichbar ist, nicht der deutschen Sprache mächtig ist oder dem Interview wegen eines Handicaps nicht folgen kann. Die Erfahrung zeigt, dass der weitaus wichtigste Grund für eine Nichtteilnahme darin besteht, dass eine Person die Teilnahme verweigert, also nicht zur Kooperation bereit ist. In der sehr gut dokumentierten sozialwissenschaftlichen Erhebung Allbus waren 2014 76,5 Prozent aller Ausfälle auf Verweigerung zurückzuführen (Wasmer et al., 2017; eigene Berechnungen). Die Antwortbereitschaft hängt u. a. vom Thema der Befragung, von der Erhebungsmethode und ihrer Dauer ab (Wübbenhorst & Wildner, 2007).

Nun ist Nonresponse kein neues Problem. Es ist so alt ist wie die Markt- und Sozialforschung selbst. Noch vor Jahrzehnten, als Responseraten von 70 Prozent und mehr möglich waren, konnten Forscher davon ausgehen, dass die Annahme der vorab berechenbaren Auswahlwahrscheinlichkeiten näherungsweise erfüllt ist. Statistiker haben eine große Erfahrung mit Annahmen, die näherungsweise gültig sind. Doch solche Responseraten sind nicht mehr möglich. Ein Beispiel dafür ist der von der GESIS beauftragte Allbus, für den ein hoher Aufwand getrieben wird. Hier sanken die Responseraten von 64 Prozent im Jahr 1980 auf 29,5 Prozent im Jahr 2021 (Brückner et al., 1981; Baumann et al., 2024).

Es ist damit zu rechnen, dass die Responseraten weiter zurückgehen werden. So liegen nach Angaben von Praktikern die Ausschöpfungsraten bei Telefoninterviews derzeit in der Regel unter zehn Prozent, bei persönlichen Interviews unter 25 Prozent und bei der Anwerbung von Online-Access-Panelisten

unter fünf Prozent. Damit ist die Annahme der berechenbaren Auswahlwahrscheinlichkeiten nicht einmal näherungsweise gültig, sondern schlichtweg falsch.

Natürlich gibt es eine Theorie über Nonresponse. Es gibt gute Lösungen für Item-Nonresponse, für den Fall also, dass nur einige Fragen nicht beantwortet wurden, andere aber schon. Diese Antworten sowie die Antworten anderer Personen können verwendet werden, um die Lücken durch Imputation zu füllen (Rubin, 1987).

Es gibt jedoch keine wissenschaftlich fundierten und für die Praxis befriedigenden Lösungen für Unit-Nonresponse, für den Fall also, dass eine Person sich weigert, überhaupt an einem Interview teilzunehmen. Wir wissen, dass Unit-Nonresponse zu keinen Verzerrungen führt, wenn die erhobenen Daten unabhängig von der Wahrscheinlichkeit der Teilnahme sind. Wir wissen aber auch, dass diese Annahme im Allgemeinen nicht zutrifft. Es gibt weiter eine Theorie für den Fall, dass eine Zufallsstichprobe der Verweigerer durch eine Erhöhung des Aufwands doch noch erhoben werden kann (z. B. Tripathi & Kare, 1997). Doch auch dies ist nicht realistisch. Schließlich gibt es Versuche, die Nonresponse-Wahrscheinlichkeit zu schätzen und für eine Gewichtung zu nutzen. In der Mehrzahl der Fälle lässt sich dadurch die Verzerrung durch Nonresponse reduzieren, in einigen Fällen erhöhen sich diese jedoch (Blumenstiel & Gummer, 2015), so dass auch diese Lösung nicht befriedigend ist.

Im Ergebnis zeigt sich, dass die Stichprobenbildung in der Praxis der Markt- und Sozialforschung ein Problem darstellt, für das die Wissenschaft keine Lösungen anbietet. Das trifft weniger für die amtliche Statistik zu, die vom Gesetzgeber teilweise ermächtigt wurde, die Teilnahme mit Bußgeldern zu erzwingen. Es trifft aber für die Marktforschung zu, die solche Möglichkeiten nicht hat und auch nicht haben will. Für die eine Methode, die Quotenstichprobe, gibt es keine Theorie. Und zur anderen Methode, der Zufallsstichprobe, hat der Autor in einem anderen Beitrag die Situation wie folgt beschrieben: „Die in den Lehrbüchern beschriebene Zufallsstichprobe ist so etwas

wie das Einhorn der Marktforschung: Es ist wunderschön, jeder weiß auch, wie es aussieht, aber – leider! – noch niemand hat es bisher gesehen.“ (Wildner, 2020)

4 Was also tun?

Dies bedeutet nun nicht, dass die Zeit der repräsentativen Stichprobenziehungen in der Marktforschung vorbei ist, wie es ein Marketingmanager auf einer ESOMAR-Konferenz schon vor fast 15 Jahren formulierte (persönliche Erinnerung des Autors). Denn wenn für ein Problem geschätzt werden muss, wie viel Geld zur Lösung des Problems benötigt wird, dann braucht man zuverlässige Zahlen darüber, wie viele Menschen von dem Problem betroffen sind. Es gibt also einen deutlichen Bedarf nach repräsentativen Stichproben.

Doch wie kann dieser Bedarf gestillt werden? Denn die Tatsache, dass eine tragfähige Theorie fehlt, macht es viel schwieriger zu repräsentativen Stichproben zu gelangen. Dies erfordert eine große Sorgfalt und viel Erfahrung.

Dazu gibt es neben den Anstrengungen, die Response-Rate zu erhöhen, vor allem die folgenden vier Strategien:

(1) Die erste Strategie ist, so weit wie möglich, Zufallsverfahren anwenden und nicht zufällige Verfahren so spät und wenig wie möglich zum Einsatz bringen. So kann es sinnvoll sein, Telefoninterviews zufällig nach dem Gabler-Häder-Verfahren zu beginnen, gegen Ende der Feldzeit dann aber nur noch Interviews mit Personen durchzuführen, die zu einer ausgewogenen Struktur der Gesamtstichprobe beitragen. Eine ähnliche Strategie kann auch für die Offline-Rekrutierung von Online-Access-Panelteilnehmern sinnvoll sein.

Ein Beispiel für diese Strategie wurde in der Fernsehforschung der GfK angewendet. Für die Teilnahme an der Fernsehforschung ist es erforderlich, dass die Haushalte ein Gerät installieren lassen, das erfasst, welcher Kanal am Fernsehgerät genutzt wird. Naturgemäß ist

die Bereitschaft dazu sehr gering. Die Lösung bestand nun darin, sehr viele kurze Interviews durchzuführen, in dem die für die Panelteilnahme wichtigen Merkmale abgefragt werden. Anschließend wird das System der Fernsehforschung erläutert und die Teilnahmebereitschaft daran abgefragt. Wird nun ein Panelteilnehmer mit einer bestimmten Struktur benötigt, so wird er unter den in Frage kommenden teilnahmebereiten Haushalten mit dieser Struktur zufällig ausgewählt.

- (2) Die zweite Strategie ist im Wesentlichen diejenige, welche bei Quotenstichproben angewendet wird. Man weiß, dass bei Quotenstichproben jeder Interviewer bei der Auswahl seiner Interviewpartner einen Bias hat. Deshalb versucht man durch den Einsatz vieler und unterschiedlicher Interviewer zu erreichen, dass sich die Verzerrungen ausgleichen. Ähnlich können bei der Anwerbung von Online-Panelteilnehmern durch die Nutzung unterschiedlicher Rekrutierungsmöglichkeiten die verfahrensbedingten Verzerrungen ausgeglichen werden (Conley, 2007). Ebenso ist bekannt, dass jede Erhebungsmethode bestimmte Personen anspricht und dadurch zu spezifischen Verzerrungen führt (Wübberhorst & Wildner, 2007). Durch den Einsatz unterschiedlicher Erhebungsverfahren ist es möglich, die Verzerrung der Ergebnisse zu verringern.
- (3) Die dritte Möglichkeit ist der Einsatz von Quotenstichproben, was zu Beginn dieses Beitrags schon ausgeführt wurde.
- (4) Schließlich gibt es noch die vierte Möglichkeit, die Ergebnisse zu gewichten, wodurch die Verzerrung zwar häufig verringert, aber normalerweise nicht beseitigt wird, sondern sogar noch verstärkt werden kann. Besonders gute Möglichkeiten bieten sich hier für Online-Panels, bei denen aufgrund früherer Erhebungen viele Informationen über die Panelteilnehmer vorliegen.

Aber all diese Möglichkeiten wurden von der Praxis entwickelt und haben keine theoretische Grundlage. Sie sind letztlich Kunst-

handwerk. Und natürlich gilt für die so gewonnenen Stichproben das Gleiche wie für die Quotenstichprobe, nämlich dass die Formeln der Stichprobentheorie für die Schätzparameter und ihre Varianzen nicht mehr anwendbar sind. Damit sind sie letztlich unbefriedigend.

Notwendig ist daher eine offene Diskussion zwischen Praktikern und Wissenschaftlern über zufällige und nicht zufällige Möglichkeiten der Stichprobenziehung. Eine solche Diskussion fordert auch Hedewig-Mohr (2020). Sie berichtet von wissenschaftlichen Tagungen, die aber ohne Ergebnisse blieben. Und sie hält fest: „In dieser Diskussion um Methoden wird weiter aneinander vorbei argumentiert und nicht offen miteinander diskutiert. Es fehlt Transparenz und bei manchen Beteiligten auch der Wille dazu.“ (Hedewig-Mohr, 2020, 10)

Dies gilt es zu ändern und zu praktikablen und dabei von der Wissenschaft begründeten Vorgehensweisen zu kommen. Der aktuelle Zustand ist jedenfalls für eine Branche, die sich die Wissenschaftlichkeit ihres Vorgehens auf die Fahnen geschrieben hat, unbefriedigend.

Management-Takeaway

Sogenannte Zufallsstichproben in der Markt- und Sozialforschung erfüllen aufgrund der aktuell geringen Teilnahmebereitschaft an Interviews nicht die Anforderung an Zufallsstichproben. Notwendig ist daher eine genaue Dokumentation, wie die Stichprobe zustande gekommen ist, sowie eine vorsichtige Interpretation der Signifikanzen und der Vertrauensintervalle.

Literatur

- ADM (2024a). Das Referenzsystem für bevölkerungsrepräsentative Studien in Deutschland. Abruf von <https://www.adm-ev.de/leistungen/arbeitsgemeinschaft-adm-stichproben/>.
- ADM (2024b). Die Marktforschung in Zahlen. Quantitative Interviews nach Befragungsart. Abruf von <https://www.adm-ev.de/die-branche/mafo-zahlen/>.
- Arnab, R. (2017). *Survey Sampling. Theory and Applications*. London: Elsevier Ltd.
- Blumenstiel, J. E., & Gummer, T. (2015). Prävention, Korrektur oder beides? Drei Wege zur Reduzierung von Nonresponse Bias mit Propensity Scores. In J. Schupp, & W. Wolf (Hrsg.), *Nonresponse Bias* (pp. 13-44). Wiesbaden: Springer VS.
- Baumann, H., Schulz, S., & Thiesen, S. (2024). *ALLBUS 2021 – Variable Report, Studien-Nr. 5280, GESIS Datenarchiv für Sozialwissenschaften*, 2. Auflage. Mannheim.
- Brückner, E., Kirschner, H.-P., Porst, R., Prüfer, P., & Schmidt, P. (1982). Methodenbericht Allgemeine Bevölkerungsumfrage der Sozialwissenschaften – ALLBUS 1980 (Nationaler Sozialer Survey), ZUMA-Arbeitsbericht Nr. 1981/07. Abruf von https://www.researchgate.net/publication/47801005_Methodenbericht_Allgemeine_Bevölkerungsumfrage_der_Sozialwissenschaften_-_ALLBUS_1980_Nationaler_Sozialer_Survey.
- Chaudhuri, A. (2014). *Modern Survey Sampling*. London: Taylor & Francis Group.
- Chaudhuri, A., & Stenger, H. (2005). *Stichprobenziehung bei Umfragen. Theorie und Methoden*. 2. Auflage. Stuttgart: Chapman & Hall.
- Cochran, W. G. (1977). *Sampling Techniques*. 3. Auflage. Hoboken, NJ: Wiley.
- Conley, P. (2007). Online Market Research. In M. van Hamersfeld, M., & C. de Bont (Hrsg.), *Market Research Handbook* (pp. 401-420). 5. Auflage. Hoboken, NJ: Wiley.
- Govindarajulu, Z. (1999). *Elements of Sampling Theory and Methods*. Upper Saddle River, NJ: Prentice Hall.
- Hedewig-Mohr, S. (2020). Zufall, Zufall über alles – Sind Zufallsstichproben heute noch machbar und was können die Alternativen sein? *planung & analyse* (4/2020), 8-10.
- Kreis, H., Wildner, R., Kuß, A. (2021). *Marktforschung – Datenerhebung und Datenanalyse*. Wiesbaden: Springer Gabler.
- Noelle-Neumann, E., & Petersen, T. (2005). *Alle, nicht jeder – Einführung in die Methoden der Demoskopie*. Wiesbaden: Springer Gabler
- Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. Hoboken, NJ: Wiley.
- Statistisches Bundesamt (Hrsg.) (2022). Fachserie 15 Reihe 2: Wirtschaftsrechnungen – Laufende Wirtschaftsrechnungen – Ausstattung privater Haushalte mit ausgewählten Gebrauchsgütern. Abruf von https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/Ausstattung-Gebrauchsgueter/Publikationen/Downloads-Ausstattung/ausstattung-privater-haushalte-2150200227004.pdf?__blob=publicationFile.
- Taylor, H. (1997). Opinion Polling. In ESOMAR (Hrsg.), *ESOMAR Handbook of Market and Opinion Research*. Amsterdam.
- Thompson, S. K. (2012). *Sampling*. 3. Auflage. Hoboken, NJ: Wiley.
- Tripathi, T. P., & Kare, B. B. (2007). Estimation of Mean Vector in Presence of Non-Response. *Communication in Statistics – Theory and Methods*, 26(9), 2255-2269. Doi.org/10.1080/03610929708832045.
- Wasmer, M., Blohm, M., Walter, J., Jutz, R., & Scholz, E. (2017). Konzeption und Durchführung der "Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften" (ALLBUS) 2014. (GESIS Papers, 2017/20). Köln: GESIS – Leibniz-Institut für Sozialwissenschaften. Doi.org/10.21421/ssoar.53370.
- Wildner, R. (2004). Stichwort Stichprobe. In D. K. Tschulin, & B. Helmig (Hrsg.), *Gabler Lexikon Marktforschung* (S. 448). Wiesbaden: Gabler.
- Wildner, R. (11.11.2020). Das Einhorn der Marktforschung - Herausforderungen bei der Stichprobenziehung in der Markt- und Sozialforschung. Abruf von <https://www.horizont.net/planung-analyse/nachrichten/-herausforderungen-bei-der-stichprobenziehung-in-der-markt--und-sozialforschung-186468>.
- Wübberhorst, K., & Wildner, R. (2007). Festnetzlos, nicht erreichbar und genervt: Gehen der Umfrageforschung die Interviewpartner aus? *Jahrbuch der Absatz- und Verbrauchsforschung* (4/2007), 340-360.