

Repräsentativität

Inhaltsverzeichnis

Repräsentativität	2
Lernhinweise	2
Einführung	2
Theorie (1-26)	3
1. Magische Konzepte?	4
2. Was heisst 'Repräsentativität'?	4
3. Repräsentativität in den Wissenschaften	5
4. Repräsentative Stichprobe als Miniaturmodell	6
5. Repräsentativität als gerechte Vertretung	7
6. Arche-Noah-Prinzip der Repräsentativität	8
7. Typisches	9
8. Nominalistische Definition einer Stichprobe	10
9. Des Kaisers neue Kleider	11
10. Repräsentativität: was das eigentlich ist, wäre zu klären	12
11. Zusammenfassung zum Lernschritt	12
12. Rechnen oder Beobachten?	13
13. Repräsentativität - eine Berner Erfindung?	14
14. Beispiel aus der Geschichte für Kaisers Methode	15
15. Sozialwissenschaftliche Untersuchungen vor der Einführung der repräsentativen Stichproben....	15
16. Widerstand gegenüber dem repräsentativen Sampling	16
17. Verteidigung der Methode	16
18. Vergessen	17
19. Mathematisierung der Stichprobentheorie	17
20. Etablierung der mathematischen Stichprobentheorie	18
21. Zahlenorakel	20
22. Urnenmetapher	22
23. Erreichung von Repräsentativität bei vermuteten Modellverletzungen	23
24. Konsequenzen der wahrscheinlichkeitstheoretischen Stichproben	24
25. Literaturauswahl	25

Repräsentativität

Lernhinweise

Dieses Submodul führt sie ein in die Diskussion um den Begriff der statistischen "Repräsentativität" und der Idee der "repräsentativen" Stichprobe

Mit dieser Diskussion sollen Sie gleichzeitig bekannt werden mit den Fragen

- Nach Stichprobe und Grundgesamtheit.
- Nach der Beurteilung der Güte von statistischen Stichproben.

Dieser Lernschritt soll also eine Vorbereitung darstellen für die nachfolgende Diskussion der Mathematik von Wahrscheinlichkeit und Stichproben.

Lernziele

- Die Voraussetzungen kennenzulernen, aufgrund derer überhaupt statistische Schlussfolgerungen von einer Stichprobe auf die Grundgesamtheit möglich sind.
- Dabei Klärungen hinsichtlich eines alltagsweltlich und sozialwissenschaftlich oft konfus gebrauchte Begrifflichkeit vornehmen, mit der die Güte von Stichproben beurteilt werden: nämlich die Vorstellung von "repräsentativen" Stichproben oder Untersuchungen.
- Ein kleines, aber wichtiges Stück der Wissenschaftsgeschichte der quantitativen Sozialforschung kennenzulernen.

Hinweis zur Bearbeitung

Beim Anklicken des "next" - Buttons werden Sie nach der oben dargestellten Rubriken-Reihenfolge durch die Lektion geführt: (1) Lernhinweis, (2) Einführung, (3) Theorie, (4) Fallbeispiel, (5) Lernkontrolle.

Einführung

Die Sozialforschung will Aussagen über soziale Gruppen vornehmen, oft über ganze Gesellschaften (die Schweiz, die Deutschen, die europäische Bevölkerung).

Es ist unmittelbar einleuchtend, dass nicht alle und alles untersucht werden können, sobald die Gesellschaft eine bestimmte Grösse erreicht hat. In einer kleinen Gemeinschaft stellt sich das Problem der Stichprobe nicht. Es müssen also für die sozialwissenschaftliche Erforschung gegenwärtiger Gesellschaften oft nach bestimmten Kriterien Teile des interessierenden Ganzen ausgewählt werden -- Teile, die wiederum Aussagen über das Ganze zulassen.

Definition: Eine Stichprobe (sample) ist eine definierte Auswahl bestimmter Einheiten eines Ganzen (Grundgesamtheit). Ziel der Stichprobenziehung ist es, aufgrund der ausgewählten Einheiten Aussagen über das nicht erfasste Ganze (die Grundgesamtheit) zu ermöglichen.

- Zwei oder drei Passanten, die ich im Verlaufe des heutigen Tages beobachte und dasselbe absonderliche Verhalten zeigen wie jüngst meine Grossmutter, bilden noch keine Stichprobe. Die Auswahl ist nicht definiert (oder muss erst im Nachhinein definiert werden) und die Grundgesamtheit ist nicht festgelegt (ausser man nehme die Weltbevölkerung).

Im gewissen Sinne soll also die Stichprobe die Grundgesamtheit "repräsentieren". Die Entwicklung von Verfahren, die Aussagen von einer Auswahl von untersuchten Personen oder Einheiten auf das Ganze ermöglichen (engl: sampling), stellt einen der wesentlichen Entwicklungsschritte der quantitativen Sozialforschung dar. Quantitative Sozialforschung in grossen Gesellschaften wäre sonst kaum möglich. Sowohl in der Forschungsliteratur wie bei der Präsentation von sozialwissenschaftlichen Untersuchungen in den Medien, wird von repräsentativen Stichproben gesprochen, um zu unterstreichen, dass die Untersuchung wissenschaftlich sei und Aussage über das Ganze, die Schweiz, die Stadt Zürich, zulasse.

Der Vorwurf, dass die Untersuchung "nicht repräsentativ" sei, entsagt ihr schlicht sämtliche Wissenschaftlichkeit.

Es bietet sich deshalb an, die Problematik der Verbindung von Stichprobe und Grundgesamtheit an dem so häufig verwendeten Begriff der Repräsentativität zu diskutieren, um über die notwendigen Klärungen begreifen zu können, inwiefern Stichproben die Grundgesamtheit "repräsentieren" können.

Theorie (1-26)

Inhaltsübersicht

- 1. Magische Konzepte?
- 2. Was heisst "Repräsentativität"?
- 3. "Repräsentativität" in den Wissenschaften
- 4. Repräsentative Stichprobe als Miniaturmodell
- 5. Repräsentativität als gerechte Vertretung
- 6. Arche-Noah-Prinzip der Repräsentativität
- 7. Typisches
- 8. Nominalistische Definition einer Stichprobe
- 9. Des Kaisers neue Kleider
- 10. Repräsentativität: was das eigentlich ist, wäre zu klären
- 11. Zusammenfassung
- 12. Rechnen oder Beobachten?
- 13. Repräsentativität - eine Berner Erfindung?
- 14. Beispiel aus der Geschichte für Kaisers Methode
- 15. Sozialwissenschaftliche Untersuchungen vor der Einführung der "repräsentativen Stichproben"
- 16. Widerstand gegenüber dem repräsentativen Sampling
- 17. Verteidigung der Methode
- 18. Vergessen
- 19. Mathematisierung der Stichprobentheorie
- 20. Etablierung der mathematischen Stichprobentheorie
- 21. Zahlenorakel
- 22. Urnenmetapher
- 23. Erreichung von "Repräsentativität" bei vermuteten Modellverletzungen
- 24. Konsequenzen der wahrscheinlichkeitstheoretischen Stichproben
- 25. Literaturauswahl

1. Magische Konzepte?

Auf welche Weise die Sozialwissenschaften über repräsentative Stichproben Aussagen über die Gesellschaften vornehmen können, erscheint der Öffentlichkeit zuweilen als Magie, als "Geheimrezept" (Noelle-Neumann). Hier zwei real vorgekommene Beispiele der Verwendung des Begriffs der Repräsentativität und seiner Verwendung:

- Ein junger Soziologe beginnt bei einem Hilfswerk zu arbeiten. Dieses will eine Umfrage durchführen, die der Soziologie als seine erste Tätigkeit leiten soll. Sogleich wird er mit der Frage konfrontiert: wie viele Personen müssen wir befragen, damit die Untersuchung "repräsentativ" ist? Offenbar glaubt der Fragesteller an die Existenz einer geheimen Formel, die das beantworten könnte. Der junge Soziologe ist verzweifelt, und dies zurecht, weil es auf diese Frage, so formuliert, keine Antwort geben kann.
- Ein anderes Beispiel: In einer Webumfrage steht, dass die Untersuchung nicht "repräsentativ" sei. Man kann sich fragen, was ist sie dann? Auch hierauf wird sich schwerlich eine kluge Antwort finden lassen.

Für den Laien schier unmöglich zu verstehen

Doch diese als Magie erscheinenden Praktiken rufen oft auch Skepsis hervor: In einem skeptischen Artikel über Stichprobenverfahren der Wahlforschung schreibt der Spiegel:

- "Durch diese Zufallsauswahl wird - so die Theorie erreicht - erreicht, dass beispielsweise 1000 Befragte ein verkleinertes Abbild der Gesamtheit aller Wahlberechtigten sind. Dann können, was vielen Laien schier unmöglich scheint, die Antworten eines einzelnen Befragten als repräsentativ für 60 000 gelten - für so viele, wie Rüsselsheim oder Weimar Einwohner hat (37, 1998)."

Was hinter der so dargestellten Vorstellung einer "repräsentativen Stichprobe" steht, erscheint freilich nicht nur als unmöglich - es ist auch tatsächlich unmöglich. Ausser es gäbe ein der Wissenschaft bekannte Geheimrezept wirklich, von dem die Meinungsforscherin E. Noelle-Neumann schreibt.

2. Was heisst 'Repräsentativität'?

Bei dem sowohl in den Sozialwissenschaften wie auch in der Öffentlichkeit gebrauchte Begriff "Repräsentativität" handelt es sich allerdings ein typischer Fall, was die Wissenschaftsforschung "Kreolisierung" nennt, die Vermengung der Bedeutungen aus verschiedenen Sprachen: im vorliegenden Fall vermischen sich Wissenschaftssprache und Alltagssprache. Doch wo Rauch ist, ist auch Feuer: denn es lassen sich über die Missverständnisse, die mit dem Begriff einher gehen, auch eine sozialwissenschaftliche Problematik genauer betrachten.

Woher kommt das Wort?

- Der Begriff "Repräsentativität" leitet sich vom Lateinischen her, von *repraesentare*, was sich mit "vergegenwärtigen, vorführen, darstellen" übersetzen liesse. Interessanterweise handelt es sich um eine Kombination von *praesentare*, also darstellen, und dem Präfix *re-*, also "wieder darstellen", was eigentlich schon da ist
- Über das Französische ist der lateinische Begriff in den allgemeinen Sprachgebrauch eingewandert: neben der substantivierten Form "Repräsentant" als "offizieller Vertreter (z.b. eines Volkes oder einer Firma)", auch Abgeordneter. Repräsentativ hiess im Französischen aber auch "stellvertretend, ehrenvoll, würdig".
- Es existieren folglich verschiedene Bedeutungsaspekte von "Repräsentativität".

Repräsentativität im öffentlichen Gebrauch bezieht sich meist nicht auf Statistik

Eine Untersuchung des Gebrauchs des Begriffs der Repräsentativität in der NZZ hat ergeben (Arnold 2001), dass die Begriffe repräsentativ und Repräsentativität keineswegs immer im Zusammenhang mit statistischen Stichproben verwendet werden. Am weitaus am häufigsten werden die Begriffe in den breiteren Bereichen der Politik und der Kultur gebraucht, aber natürlich im Sport.

- In der Politik wird von einer repräsentativen Vertretung des Volkes im Parlament gesprochen, bezüglich der Frage der Quotenregelungen bei der Gleichstellung der Frauen. Im ökonomischen Zusammenhang erscheint "repräsentativ" in Verbindung mit Spitzengehältern von Manager oder in der Frage, inwiefern das Gebaren bestimmter Aktienindizes repräsentativ sei für die wirtschaftliche Entwicklung.
- In der Kunst wird die Frage diskutiert, inwiefern ein Bild "repräsentativ" sei für eine bestimmte Periode eines Malers, oder gar einer bestimmten Epoche, ob die Werkphasen in einem Band über eines Dichters repräsentativ vertreten seien. In der Architektur wird gefragt, ob ein bestimmtes Gebäude denn wirklich repräsentativ sei, im Sinne einen Status aufwies. Insofern können auch Automodelle gleichzeitig "sportlich und repräsentativ" sein, aber auch solide und repräsentativ.
- Selbstverständlich wird der Begriff auch im Sport verwendet, etwa wenn es über ein Skispringen heisst, einer sei "...aus verschiedenen Luken gesprungen und darum nicht repräsentativ".

"Repräsentativ" wird als nicht nur im Sinne der numerischen Abbildung gebraucht, sondern im Sinne eines Statusmarkierens und im Sinne des "Geltens" für etwas. Insofern kann im breiteren Sprachgebrauch, wiederum statistisch gesprochen, ein Einzelnes für ein ganz grosses Ganzes "repräsentativ" sein (ein Werk für eine Epoche). Die Sozialwissenschaften gebrauchen also mit ihrem Begriff, der wie ein Vexierbild ganz verschiedene Bedeutungen aufweisen kann, bei dem aber keine klar definiert ist.

3. Repräsentativität in den Wissenschaften

Doch wer nun denkt, in der Wissenschaft, dem Ort des Definierens und Festlegens, habe nun der Begriff der Repräsentativität eine definierte Bedeutung, die gelehrt werden könne, der irrt. Was denn genau mit einer Repräsentativität gemeint sei, ist erstaunlicherweise nicht genau festgelegt - ungeachtet der häufigen Verwendung des Begriffs in quantitativer sozialwissenschaftlicher Forschung.

William Kruskal und Frederick Mosteller, zwei amerikanische Wissenschaftler, haben die Verwendungsweise des Begriffs "representative sampling" in der Statistik untersucht, seit dem Auftauchen des Konzepts im Jahre 1895 und erkannten eine Reihe gänzlich unterschiedlicher Bedeutungen des Konzepts. Sie haben sie in Karikaturen dargestellt, die hier wiedergegeben werden sollen, weil diese Vorstellungen von statistischer Repräsentativität nach wie vor wirksam sind.

4. Repräsentative Stichprobe als Miniaturmodell



Meaning 3. *Miniature of the population: Model train set.*

Es gibt zunächst die Vorstellung, eine repräsentative Stichprobe verkörpere so etwas wie eine "miniaturisierte" Form der Grundgesamtheit - also beispielsweise eine Mini-Gesellschaft, eine Mini-Schweiz. Alles Merkmale und Wirkungsweisen sind in der Stichprobe gleich wie in der Grundgesamtheit vorhanden, nur geschrumpft halt. Aber die Stichprobe vergegenständlicht eine Gesellschaft im Kleinen. Die Vorstellung der "Miniaturgesellschaft" bildet die weitaus am wichtigste Bedeutung des Begriffs in den statistischen Sozialwissenschaften.

5. Repräsentativität als gerechte Vertretung



Meaning 2. Absence of selective forces: Justice balancing the scales.

Eine andere Vorstellung meint, dass die Untersuchung ausgewogen sei, dass es nicht verzerrende Elemente gibt, welche die Ergebnisse verfremden. Keine Gruppe soll in der Stichprobe über- oder untervertreten sein. Beispielsweise soll eine Studie zur sozialen Ungleichheit darauf achten, dass alle relevanten Bevölkerungsgruppen in gleicher Weise vertreten sind. Hier finden wir eine Schnittstelle zum politischen Feld, gleich wie es in einer (repräsentativen) Demokratie als eine Gerechtigkeitsfrage gilt, wenn bestimmte Gruppen im Parlament über- (ausgebildete Juristen) oder aber untervertreten sind (Frauen beispielsweise), wird es auch als wissenschaftlich inadäquat betrachtet, wenn bestimmte Gruppen in einer Stichprobe unter- oder übervertreten sind.

6. Arche-Noah-Prinzip der Repräsentativität

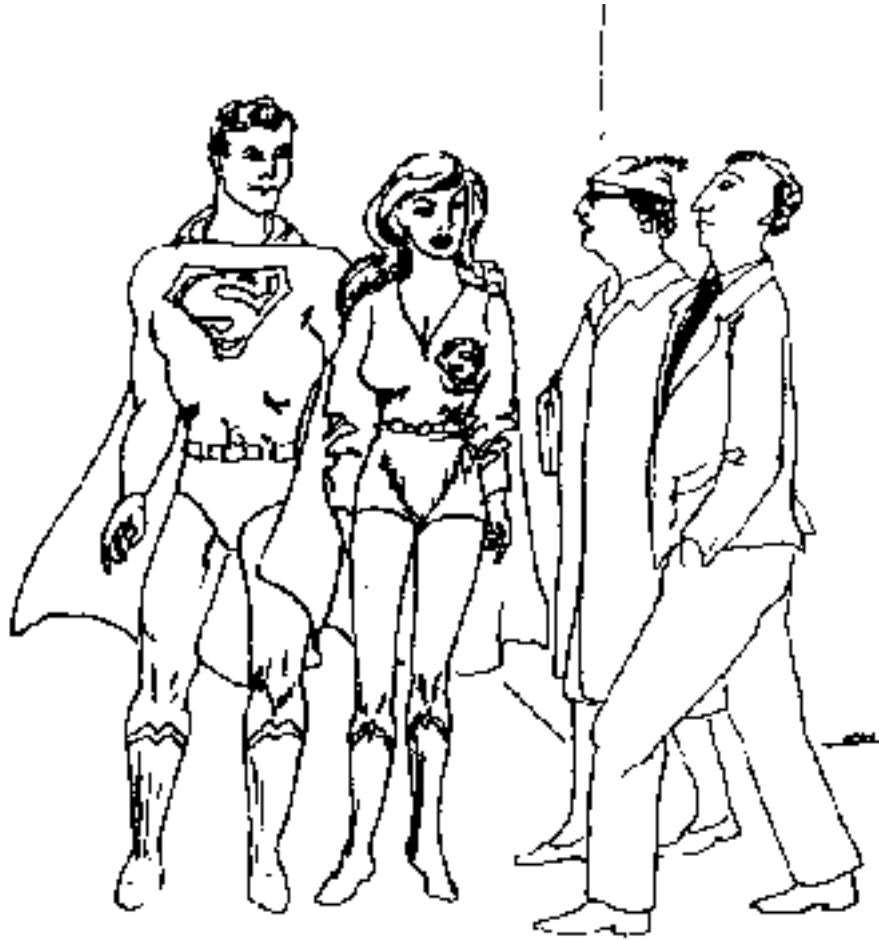


Meaning 5. Coverage of the population: Noah's Ark.

Eine andere Bedeutungshorizont der "repräsentativen Stichprobe" meint, dass in der Stichprobe alle Merkmale einer Grundgesamtheit vertreten sein sollten, mit anderen Worten: wie in der Arche Noah soll jeder "Gattung" in der Arche Noah vertreten sein. Wie in der Arche Noah tritt dabei die numerische Grösse der einzelnen "Gattungen" in den Hintergrund angesichts der Erhaltung der "Arten-Vielfalt".

Für eine soziologische Forschung bedeutet dies beispielsweise, dass eine Untersuchung über Armut in einer grösseren Stadt möglichst versuchen sollte, alle Bevölkerungsgruppen zu erfassen, die von dem Problem betroffen sein könnte - unabhängig von der Tatsache, ob die Personen einen festen Wohnsitz mit eingetragener Telefonnummer besitzen, die sie für ein Meinungsforschungsinstitut erreichbar werden lassen.

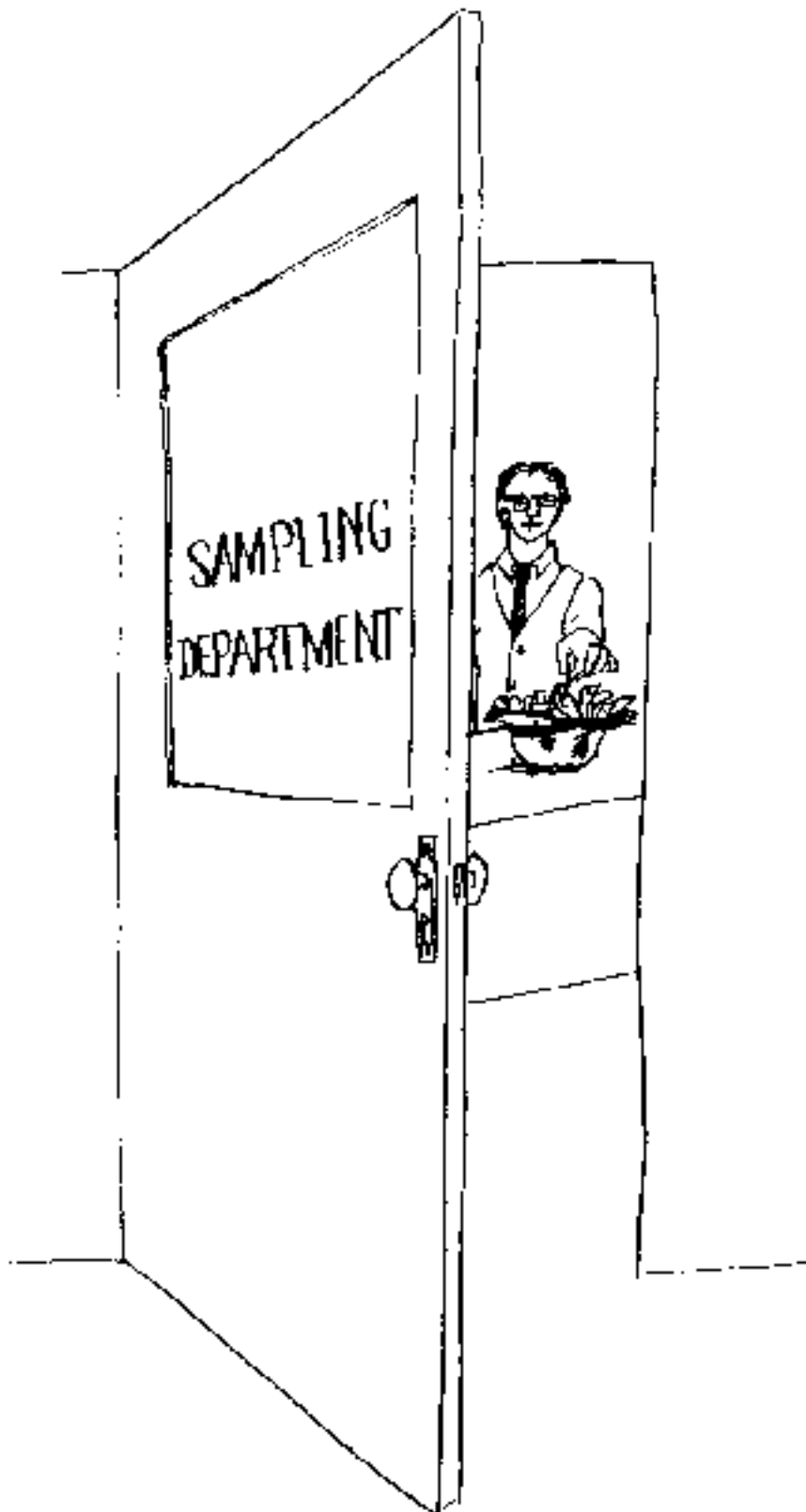
7. Typisches



Meaning 4. Typical or ideal case: Superman and Superwoman and Average man and Average woman.

Eine andere Bedeutungskonnotation von "repräsentativ" meint eigentlich "typisch". Man versucht nicht alle, oder die Durchschnittlichen herauszunehmen. Ein repräsentatives Ergebnis meint hier, dass das Ergebnis einen typischen Bereich, Prozess oder typische Verhaltens- und Einstellungsmuster repräsentiert. In gewissem Sinne kann auch Alphonse Quetelets "homme moyen", des imaginären statistischen Durchschnittsmenschen als Repräsentationskonstruktion in diesem Sinne verstanden werden. Quetelet selbst war gegen jegliche Stichprobenziehungen, nachdem er sie zunächst in Betracht gezogen hatte

8. Nominalistische Definition einer Stichprobe



Meaning 7. Some specific sampling method: The Sam-

Sehr beliebt ist auch eine technische Definition der Repräsentativität: Als repräsentative Stichproben gilt das, was die statistischen Ämter und universitäre Statistik als repräsentative Stichproben definieren - nach ihren eigenen Regeln. Da es sich um eine wissenschaftliche Definition handelt, muss sie auch nicht mit anderen Bedeutungsfeldern kommunizieren.

Dieses Verfahren, Repräsentativität zu festzulegen, ist beliebt in Einführungsbüchern in die Sozialwissenschaftliche Statistik und in Grundstudiumsveranstaltungen zu Statistik.

9. Des Kaisers neue Kleider



Meaning 1. *General, usually unjustified, acclaim for data: The emperor's new clothes.*

Wie als Antwort auf diese nominalistische und in einem gewissen Sinne eigenmächtige Definition von Repräsentativität existiert eine andere, skeptische im wissenschaftlichen Diskussion. Sie will das Konzept entlarven: wie des Kaisers neue Kleider im gleichnamigen Andersen Märchen, sei das Konzept der statistischen Repräsentativität blosser Glauben, das alle gebrauchen (weil sie Interesse daran haben, eine komplexe Problematik auf einen so einfachen Punkt zu bringen), obwohl sie es anders wissen.

10. Repräsentativität: was das eigentlich ist, wäre zu klären



Meaning 6. *Vague term to be made precise: A hatching egg of unknown species.*

Entsprechend gibt es auch seit dem Auftauchen des Begriffs in der Statistik die Vorstellung, dass der Begriff eigentlich zu klären sei - und diese Forderung gibt es auch nach über 100 Jahren, seit er aufgetaucht. In dieser Tradition steht dieser Einführungstext. Nicht in dem Sinne, als dass hier nun die endgültige Lösung angeboten werden soll, sondern in dem Sinne, dass auf die mögliche Klippen und Missverständnisse in der Verwendung des Begriffs hingewiesen und sein bewusster Gebrauch gefördert werden soll.

11. Zusammenfassung zum Lernschritt

- Ein Nachforschen zeigt, dass der Begriff der statistischen Repräsentativität keineswegs in Wissenschaft und Öffentlichkeit keineswegs einheitlich und stimmig gebraucht wird. Es gibt verschiedene Bedeutungskonnotation, die sich ergänzen, überlagern und oft auch widersprechen.
- Die Vorstellung der Repräsentativität als gerechte, der "Wirklichkeit" entsprechende Verteilung in einem Sample widerspricht u.U. beispielsweise dem Arche Noah-Modell der Repräsentativität. Beide wiederum stehen in Konfrontation mit der Vorstellung der Repräsentativität als das Typische, da das Gerechtigkeitsmodell gerade auch nach dem Nichttypischen und das Arche-Noah Modell nach dem möglichst Umfassenden sucht. Diese Vorstellungen haben aber alle, je nach Forschungsfrage ihre Rechtfertigung. Soll der Begriff der "Repräsentativität" deshalb am besten aufgegeben werden?

Rainer Schnell, Paul Hiller und Elke Esser schreiben in Ihrem Methodenlehrbuch:

- "Die Vielfalt und Verschwommenheit der verschiedenen Definitionen bewirkt, dass die bloße Feststellung, bei einer gegebenen Stichprobe handle es sich um eine "repräsentative Stichprobe" oder um eine "Repräsentativbefragung" nichts aussagt; erst recht handelt es sich um kein exakt definiertes Gütekriterium" (287).

Das Kind mit dem Bade ausschütten?

Den Begriff deshalb fallen zulassen, bedeutet aber auch, das Kind mit dem Bade auszuschütten. Denn der Begriff hat sich etabliert als eine Brücke, mit der de facto die Sozialwissenschaften und die breitere Öffentlichkeit miteinander kommunizieren und als real funktionierendes Konzept gilt es ihn ernst zu nehmen. Allerdings empfiehlt es sich genau zu achten, was denn damit kommuniziert zu werden, um als Forscher oder Forscherin nicht missverstanden oder aber berechtigterweise als unseriös zu gelten.

Eine Möglichkeit, um die „verfilzten“ Bedeutungskonnotationen (Desrosières) von statistischer Repräsentativität begreifen zu können, um mit ihnen umgehen zu können, ist die Vergegenwärtigung der wissenschaftsgeschichtlichen Prozesse, die zu seiner Herausbildung geführt haben, Problemfelder, die auch heute noch wirksam sind und die verstanden werden sollten.

Mit anderen Worten geagt, es soll gefragt werden:

- Welche Fragestellungen und Forschungsprobleme haben dazu geführt, dass Stichproben eingeführt wurden und parallel dazu von Repräsentativität gesprochen wurde?
- Was hat dazu geführt, dass die Vorstellung der statistischen Stichprobe schlussendlich übernommen wird?
- Welche anderen Untersuchungs-Konzepte und Begriffe hat die Praxis und Vorstellung des repräsentativen Samplings ersetzt?

12. Rechnen oder Beobachten?

Die Frage der Stichprobenziehung und die Frage der Repräsentativität, ein Komplex, der wesentlich zur Entwicklung quantitativer Sozialforschung gehören, tauchen in den sozialwissenschaftlichen Debatten gleichzeitig und untrennbar miteinander verbunden auf.

Stichprobenverfahren waren freilich vielerorts schon getestet und anderen Wissenschaften worden, beispielsweise in der Geologie oder in den Agrarwissenschaften, und es lag eine wahrscheinlichkeitstheoretische Begründung von mathematischen Aussagen über die Stichprobenziehung vor. Vielleicht erscheint es aus heutiger Perspektive merkwürdig: die entstehenden Sozialwissenschaften verwendeten weder die Techniken der systematischen Stichprobenziehung noch der Wahrscheinlichkeitsrechnung. Aber nicht, weil kein Wissen darüber vorhanden war, sondern weil frühere statistisch operierende Sozialforscher, wie beispielsweise Alphons Quetelet, strikte dagegen waren.

Die Gründe für die Ablehnung der Anwendung von Stichproben und mathematischen Verfahren, die im Verlaufe der Entwicklung der Sozialwissenschaften verschwanden, dürfen auch heute nicht unterschlagen werden. Vielleicht werden sie vor aktuellen Verhältnissen sogar wieder aktuell.

13. Repräsentativität - eine Berner Erfindung?

Der bedeutendste Ort für die Diskussion wissenschaftliche statistische Fragen stellte in der Wende des 19. zum 20. Jahrhunderts das International Statistical Institute (ISI). Hier tauchte auch das erste Mal der Begriff der Repräsentativität in einem statistischen Zusammenhang auf. In einer Berner Konferenz des ISI im Jahre 1895 präsentierte der Direktor des Norwegischen Zentralbüros für Statistik Anders Nicolai Kiaer ein Papier mit dem Titel *Observations et expériences concernant des dénombrements représentatifs* (Beobachtungen und Erfahrungen mit repräsentativen statistischen Untersuchungen).

Kiaer gab damit den ersten Versuchen im 19. Jahrhundert mit Stichproben, sie hiessen damals "indirekte Methode" oder "Teiluntersuchungen" einen neuen Namen. Doch waren Stichprobenverfahren wenig akzeptiert.

Die Suche nach dem Typischen wird durch die Idee der Miniaturgesellschaft ersetzt

Kiaer ersetzt im gewissen Sinne das Repräsentationsmodell des Typischen, das noch in der Idee des statistischen Durchschnittsmenschen zum Zuge kam, durch Miniaturmodell der Repräsentativität. Es sollten die interessierenden gesellschaftlichen sorgfältig eingeteilt werden (aufgrund von geographischen Kartenmaterial und dann selektive soziale Einheiten (Häuser) ausgewählt werden, diese dann aber sollen eine viel sorgfältigere Untersuchung erfahren, als es über jeglichen Zensus nur möglich wäre.

Beispielsweise sagt er: Die Forscher sollten nicht nur die durchschnittlichen Haushalte (maisons) erheben, sondern jene Haushalte (maisons), welche die unterschiedlichen sozialen und ökonomischen Verhältnisse repräsentieren. Doch diese Auswahl der Untersuchungseinheiten (Häuser, Familien, Individuen) erfolgte systematisch und noch nicht zufällig.

Die Zwei-Fässer-Logik

Kiaer begründet die Stichprobentheorie folgendermassen:

- "Unter einer Repräsentativuntersuchung verstehe ich eine partikuläre Untersuchung aufgrund vieler Beobachtungseinheiten, die über ein Territorium zerstreut sind, um damit eine Miniatur des vollständigen Territoriums herzustellen. Die Beobachtungseinheiten werden nicht willkürlich, sondern nach einer rationalen Gruppierung ausgewählt, die auf Volkszählungsergebnissen beruht, wobei die sich ergebenden Resultate wiederum mit der Volkszählung verglichen werden sollten."

An einem andern Ort sagt er:

- "Man stelle sich zwei Fässer vor, ein sehr grosses und ein kleines, gefüllt mit gemischtem Inhalt. Zeigen nun mehrere Stichproben aus den beiden Fässern, in etwa ähnliche Verteilungen des gemischten Inhalts (also immer etwa ähnlich viele rote und blaue Stücke), so ist die Annahme zulässig, dass der Inhalt der beiden Fässer sich entspricht. Aber dies bedeutet auch: es genügt, das kleinere Fass zu untersuchen, um zu wissen, was im grösseren steckt."

Mit anderen Worten: die Idee des repräsentativen Samplings war ganz stark von einer räumlichen Gesellschaftsvorstellung geprägt und entstand ohne den Einbezug der Wahrscheinlichkeitsrechnung.

Erst in einem zweiten Schritt und als Fortsetzung von Kaiers Gedanken forderte der Statistiker von Bortkiewicz, die Unterschiede zwischen den Verhältnissen in der Stichprobe und den Verhältnissen der Volkszählungen (beispielsweise über den Anteil der Geschlechter oder den Anteil der Arbeiter in der Stichprobe und der Grundgesamtheit) statistisch zu beurteilen.

14. Beispiel aus der Geschichte für Kiaers Methode

Eine der ersten sozialwissenschaftlichen Untersuchungen der Schweiz wurde 1963 im Vorfeld der Expo durchgeführt. Das Projekt hiess "Un jour en Suisse". Neben einer Vielzahl von Interviews sollte auch einem Querschnitt der Schweizer Bevölkerung ein standardisierte Fragekatalog präsentiert werden. In der Expo, in einem Projekt Gulliver, sollten die Besucher ihre Antworten auf bestimmte Fragen, beispielsweise zur Schweizer Armee und der Situation der Frauen, mit den Antworten der "Durchschnittsschweizer" und # "Schweizerinnen" vergleichen können. Die Befragungen wurden in systematisch ausgewählten Gemeinden durchgeführt, nach einem eigenen geografischen Stichprobenplan die Haushalte aufgesucht. Interessierende Merkmale der Interviewten der Stichprobe -- ihr Geschlecht, Alter beispielsweise, -- wurden in Bezug zu den Volkszählungsdaten des Jahres 1960 gestellt (die erwarteten Häufigkeiten). Bei gleichen Anteilen entsprächen sich die entsprechenden Merkmale der Stichprobe ("erhalten") und der Grundgesamtheit ("errechnet"). So konnte also die "Repräsentativität" der Stichprobe überprüft, indem beispielsweise herausgefunden wurde, ob beispielsweise die Katholiken und Katholikinnen untervertreten waren.

	Total	Protestants	Catholiques	Autres	Sans confession	Divers	Non Reponses
Erwartet	1200	640	535	25	0	0	0
Erhalten	1200	638	512	12	24	2	12

Der Bundesrat verbot schlussendlich das Unterfangen. Das neue Verfahren, das unbekannte Wissen, das es an den Tag zu bringen drohte, erschien nicht genehm. Die Untersuchung wurde schlussendlich von dem Pariser Soziologen Luc Boltanski unter dem Titel "Le bonheur Suisse" publiziert. Sie können in diesem Buch nachlesen, weshalb Boltanski zum Schluss kommt, dass das "Bonheur Suisse" letztendlich ein "mal Suisse" sei.

15. Sozialwissenschaftliche Untersuchungen vor der Einführung der repräsentativen Stichproben

Weshalb war Kiaers Idee so neu? Weshalb erregte sie Widerstände? Weshalb versank sie wieder in Vergessen? Um die Kritik an der Methode zu verstehen ist zunächst zu sehen, auf welche Weise sozialwissenschaftliche Untersuchungen vorgenommen wurden, vor der Einführung der Stichproben.

Primat der Vollerhebung

Mit zunehmender Bürokratisierungen, Rationalisierung und Professionalisierung im Verlaufe des 19. Jahrhunderts erhoben die staatlichen Verwaltungen in einer neu entstehenden Zettelwirtschaft vermehrt auch statistisch verwertbare Datenbestände (Statistik leitet sich schon alleine von der Begrifflichkeit von "ragione di state" als Staatsrationalität her).

Diese Datenbestände wurden schon früh für sozialwissenschaftliche Untersuchungen verwendet. Zu den Pionieren zählten beispielsweise Alphons Quetelet aber auch Friedrichs Engels (mit seinen Untersuchungen zur Arbeiterklasse in England). Die frühen statistischen Erhebungen und soziologischen Untersuchungen selbst noch Emile Durkheims beruhten allesamt auf Vollerhebungen, das heisst Daten basierend auf Volkszählungen oder Vollerhebungen in Städten und Gemeinden. Nur einen Teil des Untersuchungsfeldes zu untersuchen, galt als wenig sinnvoll.

Ein Beispiel eines Pioniers der Sozialstatistik, des Astronomen A. Quetelet (1835), das die Alphabetisierung von Gefängnisinsassen in Frankreich des Jahres 1828 zeigt eine solche Vollerhebung. Heute wird man nicht mehr sämtliche Gefangenen auf ihre Alphabetisierung hin prüfen.

Monographien

Die andere Untersuchungsform der Sozialwissenschaften waren die so genannten Monographien. Üblicherweise werden sie als Einzelfallanalyse dargestellt. Eher sollte man aber von Feldstudien, die u.U. mit statistischen Mitteln operierte, sprechen: eine Region, ein Quartier, eine Stadt wurden hinsichtlich ihrer soziologischen Entwicklung hin möglichst genau untersucht und beschrieben. Pionier dieser Form der Sozialforschung war Frédéric Le Play (1806-1882). Der französische katholische Sozialreformer interessierte sich vor allem für den Zustand der Familie angesichts der beschleunigenden Industrialisierung (*Les ouvriers européenne*, 1877). Le Play ging unabhängig seines nicht-statistischen Verfahrens sehr methodisch vor: Vgl. "Instruction sur la Méthode d'observation dite des monographies de famille, 1862"

16. Widerstand gegenüber dem repräsentativen Sampling

In dieser etablierten Ordnung der Sozialforschung, die zwischen Monographien und Vollerhebung operierte (eine Methodenbuch im heutigen Sinne hätte diese beiden hauptsächlichen Forschungen darzustellen), erschien die Idee der repräsentativen Stichprobe etwas völlig Neues, das Widerstände erwecken musste.

Die Argumentation gegen die Stichprobenziehung kann auf einen Punkt gebracht werden, der auch heute noch von keiner Stichprobentechnik ignoriert werden kann.

Das Argument lautet folgendermassen: Es ist unmöglich die Beobachtung von Tatsachen durch Berechnung zu ersetzen. Eine Stichprobe lässt nur sorgfältige Aussage lediglich über die Stichprobe selbst zu. Alle anderen Aussagen über die Grundgesamtheit sind lediglich Hypothesen. Wissenschaftslogisch müsste man sagen: alle Aussagen über das Ganze aufgrund von Stichproben beruhen auf einem Induktionsschluss (auch wenn damit Hypothesen getestet werden), und wo Induktionsschlüsse gezogen werden, gibt es auch eine Induktionsproblematik. Deshalb vorderten die Verteidiger der Vollerhebung im Wortlaut: "Pas de calcul là où l'observation peut être faite". Keine Rechnerei, wo direkte Beobachtung möglich ist.

Um die Widerstände gegenüber der neuen Methode ungeachtet des gewichtigen Inhaltes zu verstehen, sind auch ein, zwei andere Evidenzen relevant:

- Das ISI war getragen vor allem von Vertretern der entstehenden statistischen Ämtern in den verschiedenen Staaten. Sie bauten unter teils grossen Mühen Institutionen wie die Volkszählungen auf (also Vollerhebungen) und waren um deren Sicherungen respektive Erweiterungen bemüht.
- Im ISI waren vor allem Personen der unmittelbaren Praxis vertreten. Durch die mit der Stichprobentheorie sich abzeichnende Mathematisierung und damit Professionalisierung fühlten sich die mathematischen Laien offenbar bedroht.

17. Verteidigung der Methode

Kiaer und die Vertreter der repräsentativen Stichprobe verteidigten sich in den Debatten über Argumente, die teils auch heute noch Gültigkeit haben:

- Kiaer argumentiert, dass er eine kleine Zahl sorgfältig ausgewählter und sorgfältig unternommener Beobachtungen einer grossen Zahl bloss oberflächlicher und unsorgfältiger Beobachtungen vorziehe. Stichproben können inhaltsreichere Informationen als umfassende und notwendigerweise wenig differenzierte Vollerhebungen. Heute ist dies das Standardargument für die postulierte Überlegenheit von Stichproben gegenüber Vollerhebungen.
- Wissenschaftslogisch nicht ganz korrekt wird darauf hingewiesen, dass der Schluss von der Stichprobe auf die Grundgesamtheit (ein klassischer Induktionsschluss) mit all seinen Problemen dem Schluss entspräche, von einer Vollerhebung auf eine andere, spätere zu schliessen (hierbei handelt sich aber in logischer Hinsicht um einen Analogieschluss).
- Letztendlich aber sind es vor allem Argumente der Effizienz und der Ökonomie die für die Stichprobe sprechen: Vollerhebungen sind teuer, aufwändig und entsprechend schwerfällig. Stichproben entsprächen besser den steigenden Wissensbedarf.

Nichtsdestotrotz forderte Kaier immer eines: das stets sorgfältige Vergleichen der Stichprobenresultate mit der Volkszählung.

18. Vergessen

Es ist nun eine wissenschaftsgeschichtlich interessantes und immer wieder zu beobachtendes Phänomen, dass Auseinandersetzungen in wissenschaftlichen Debatten nicht über Argumente gelöst werden, eine Synthese entsteht, sondern dass sie vergessen werden, oder aber ein Modell schlicht das andere ersetzt. Beides geschah mit der Idee der statistischen Repräsentativität:

Die Diskussion um die repräsentative Methode verschwand wieder aus den wissenschaftlichen Debatten. Den Durchbruch schaffte sie freilich nicht durch neue Argumente oder neue mathematische Verfahren, die waren ja im Prinzip schon lange bereit gelegt, sondern über zwei spektakuläre Ereignisse, welche die Fachwelt ebenso wie die Öffentlichkeit Aufsehen erregen liess und die Stichprobentheorie als forschungstechnisch praktikabel erscheinen liess. Zunächst sollen aber die Mathematisierung der "repräsentativen Stichprobe" und die Definition ihrer Voraussetzungen eingeführt werden.

19. Mathematisierung der Stichprobentheorie

Wie gesehen, die Idee der repräsentativen Stichprobe gründet sich noch in wenig Mathematik. Mehr oder weniger systematisch wird ein gesellschaftlicher Bereich erhoben und nach interessierenden Merkmalen mit der zur Verfügung stehenden Merkmalen der Grundgesamtheit verglichen. Die Mathematisierung des Verfahrens geschah erst in einem zweiten Schritt und weniger spektakulär als die Debatten um die Einführung der Stichprobe und des Konzepts der Repräsentativität. Die Mathematisierung sollte die Aussagemöglichkeiten der Stichproben definieren helfen, sie damit nachvollziehbar und kritisierbar werden lassen.

Arthur Lyon Bowley und die Armutsforschung

Die Mathematisierung der Stichprobenverfahren und damit die Begründung der modernen Stichprobentheorie im heutigen Sinne leistete im wesentlichen der britische, an der London School of Economics lehrenden Mathematiker A. L. Bowley (1869-1957). Der Kontext der Entwicklung einer sozialwissenschaftlichen und mathematischen Stichprobentheorie liegt in Bowleys Engagement für britischen Sozialreformen und, wie er selbst sagt, dem knappen Budget, das für seine Armutsforschungen zur Verfügung stand.

Bowley, fasziniert von Kiaers Idee, erweiterte die Stichprobentheorie, indem er sie mit wahrscheinlichkeitstheoretischen Überlegungen verband. Weshalb aber mathematisierte aber Bowley Kiaers des representative Sampling überhaupt?

Die wahrscheinlichkeitstheoretischen Modelle lassen Aussagen zu über wahrscheinliche Entsprechungsbereiche von Stichprobe und Grundgesamtheit sowie über wahrscheinliche Fehler der Schätzungen von Merkmalen in der Grundgesamtheit, die unbekannt sind nicht von vornherein festgelegt sind – aber dann und nur dann, wenn die Modellbedingungen der Wahrscheinlichkeitsrechnung überhaupt gegeben sind.

Die Frage der Modellbildung

Die Problematik Modellbedingungen hatten Sozialstatistiker wie Quetelet, welche die Wahrscheinlichkeitsrechnung sehr wohl kannten, bislang zurückgeschreckt, sie auf den Bereich der Gesellschaft anzuwenden.

Die Anwendungsbedingungen der Wahrscheinlichkeitsrechnung auf Stichproben sind in einem gängigen Metapher zusammengefasst, nämlich dem sogenannten Urnenmodell.

Ein Universum, oder die Grundgesamtheit, die Population oder das Ganze wird, so Bowley, ist symbolisiert über eine Urne oder eine Gruppe von Urnen, welche eine Anzahl von Kugeln enthalten, die nur hinsichtlich des Kriteriums der Farbe der Kugeln variieren. Aus diesen Urnen werden eine oder mehrere Stichproben gezogen, aufgrund vordefinierter Regeln und danach (mathematisch) gefragt, welches das Entsprechungsverhältnis zwischen den Stichproben und den Kugeln darstellt. Mit einer unterschiedlichen wahrscheinlichkeitstheoretisch festgelegten Sicherheit lässt sich dann sagen, in welcher wahrscheinlichen Spannbreite die effektiven Farbenverhältnisse der Kugeln in der Urne liegen, obwohl nicht der Inhalt der ganzen Urne bekannt ist.

Die Wahrscheinlichkeitstheorie liefert hier exakte Ausformulierung dessen, was sagbar ist, welche Wahrscheinlichkeitsbereiche der Aussage über das Ganze möglich sind (in die Mathematik der Wahrscheinlichkeitsrechnung wird in Kapitel XXX eingeführt).

Doch was sind die Voraussetzungen überhaupt, dass das Urnenmodell überhaupt auf den Untersuchungsbereich anwendbar ist?

20. Etablierung der mathematischen Stichprobentheorie

Dass überhaupt trotz der Skepsis gegenüber den Verfahren die mathematische Stichprobentheorie sich zu etablieren begann, ist nicht Folge wissenschaftlichen Disputierens in Kongressen und Zeitschriften, sondern ergibt sich, wiederum einem wissenschaftshistorischen Muster folgend, mit Aufsehen erregenden Anwendungen in der Praxis.

Der eine Durchbruch hat dazu geführt, dass Bowley sich in seinem Unterfangen bestätigt führt, die andere, auf der Basis der Idee des mathematisierten representative sampling, bescherte der Öffentlichkeit die Meinungsforschung.

- Ein Erdbeben in Tokyo im Jahre 1923 zerstörte grosse Teile der Volkszählungsdaten. Allerdings lag ein Verzeichnis der Haushalte Tokyos vor. Die Behörden recherchierten jeden 1000 Haushalt, die Haushaltsnummer 500, 1000, 1'5000 usw. des 11 Millionen Haushaltungen umfassenden Registers und publizierten 1924 die Alters- und Geschlechterverteilungen der Bevölkerung Tokyos, die Grösse der Haushalte usw., um damit der Administration wieder die notwendigen Informationen zur Versorgung Tokyos zur Verfügung zu stellen. Die Differenzen zwischen den Ergebnisse der Stichproben und der später vollzogenen Volkszählung waren gering und im Bereich der mathematisch erwartbaren

Fehlerspanne. Dieses Vorgehen beschrieb auch im wesentlichen jenes von Bowley. Die Urne, das war der Stadtplan und das Häuserverzeichnis, aus dem dann jedes n-te Haus gezogen wurde. Sofern keine Gründe vorliegen, dass eine Konstante die Auswahl der n-ten Einheit stört, kommt dieses Vorgehen einer reinen Zufallsauswahl gleich. Von Bowley stammt auch die in der Stichprobentechnik auch heute noch vielfach gebrauchte Masszahl, wie gross denn die Stichprobe sein soll, um Stichprobenfehlerbereiche schätzen zu können: in ungefähr 1000 Einheiten in der Stichprobe, so Bowley, sollte für die meisten Fragestellungen hilfreicher sein. Der Grund für die häufige Verwendung dieser Zahl bei "repräsentativen Stichproben" ist denn heute in wissenschaftlicher Gewohnheit zu finden.

Meinungsforschung und Stichprobe: Ein politisches Spiel

- Spektakulärer ist jedoch der Durchbruch der Stichprobentechnologie in der Meinungsforschung. Sie führte dazu, dass das Ziehen "repräsentativer Stichproben" und Meinungsforschung oft sogar gleichgesetzt werden. Im Französischen wird dafür sogar dasselbe Wort verwendet: Sondages. Diese "Erfolgsgeschichte" des Samplings wird in den Methodenbüchern gerne erzählt.

In den USA zu Beginn des 20. Jahrhunderts war ein politisches Spiel besonders beliebt: aufgrund von Probeabstimmungen sollten die kommenden Präsidentschaftswahlen vorausgesagt werden, zuweilen auch Fragen über andere politische Fragen erhoben werden. Solche straw polls waren an den verschiedensten Veranstaltungen und bei den verschiedensten Gelegenheiten gang und gäbe. Bei allen möglichen und unmöglichen Gelegenheiten wurden Probeabstimmungen durchgeführt und die Stimmen zu einem Gegenstand aufsummiert und miteinander verglichen. Diese Institution hat erstaunlicherweise überlebt: so etwa in sogenannten TED-Umfragen bei Fernsehsendern, bei dem die Fernsehzuschauer per Telefon ihre Meinung abgeben und dafür bezahlen. Aber auch das Internet bietet eine solche Plattform. Fast jedes grössere Webportal bietet heute die Möglichkeit, Abstimmungen online durchzuführen, mit zuweilen etwas bizarren Resultaten und Vorkommnissen.

- Beispiel: Beobachter 18/01
"Schiessobligatorium: Flankenfeuer aus dem Schützengraben. Schweizer Schützenkreise haben bei einer Beobachter-Umfrage kräftig mitgewirkt."
Von Thomas Angeli
«Soll das <Obligatorische> abgeschafft werden, damit in vielen Gemeinden der Schiesslärm reduziert wird?», wollte der Beobachter bei seiner Internetabstimmung im Juli wissen (Nr. 14). Die anfängliche Ja-Mehrheit hatte nicht lange Bestand: Innert weniger Tage kippte das Resultat. Endstand: Satt 76 Prozent waren gegen die Abschaffung des «Obligatorischen». «Wie lange dauert die Umfrage?» 1814 Personen machten bei der Umfrage mit, dreimal mehr als sonst. Des Rätsels Lösung findet sich in der «Schweizerischen Schützenzeitung» vom 19. Juli: «Dank dem Zusammenhalt unter den Schützen änderte sich der Nein-Stimmen-Anteil von Tag zu Tag mehr zu unseren Gunsten», schreibt dort Vreny Zaugg, Aktuarin des Thurgauischen Kantonal-Schützenvereins, über die Nein-Stimmen-Aktion aus Schützenkreisen. Einige Tage zuvor hatte sich Zaugg bei der Beobachter-Redaktion per E-Mail erkundigt, «wie lange die Umfrage betreffend Abschaffen des <Obligatorischen> noch dauert». Bleibt also bloss zu wiederholen, was der Beobachter auch sonst unter die Ergebnisse schreibt: «Die Umfrage ist nicht repräsentativ.»
- Beispiel: CNN vom 11. September 2001.

Created: Tue Sep 11 11:44:17 EDT 2001

Do you know someone who works at the Pentagon or the World Trade Center?		
Yes	 33%	2064 votes
No	 67%	4252 votes
Total: 6316 votes		

This QuickVote is not scientific and reflects the opinions of only those Internet users who have chosen to participate. The results cannot be assumed to represent the opinions of Internet users in general, nor the public as a whole. The QuickVote sponsor is not responsible for content, functionality or the opinions expressed therein.

Related:

- [Get more Quick Vote results](#)

21. Zahlenorakel

Bestimmte Zeitungen widmeten sich in besonderer Weise der Voraussage der Präsidentschaftswahlen über Probeabstimmungen. Sie veröffentlichten Talons in den Zeitungen, die ausgeschnitten und eingesandt werden konnten. Um die Stichprobe über die engeren Klientel der Zeitungen hinaus zu erweitern, wurden beispielsweise auch Adressen von Telefon- und Autoverzeichnissen gesammelt. Die Zeitschrift Literary Digest sammelte die wahrscheinlich umfassendste Adressdatei. Sie umfasste 1895 bereits 350'000 Adressen und 1932 fanden sich gar über 32'000'0000 Anschriften. Für die Wahlen 1928 sollen nicht weniger als 18 Millionen "Wahlzettel" verschickt worden sein. Literary Digest vermochte 1928 wie auch die Jahre zuvor die Wahlergebnisse auch hinreichend präzise vorauszusagen, so etwa prognostizierte sie dem Kandidaten Hoover 63.2% der Stimme und er gewann schlussendlich mit 58.8 % der Stimmen. Literary Digest stand deshalb im Ruf, besonders seriös und zuverlässig zu sein.

Die Blamage

1936 allerdings wandelte sich der bisherige Erfolg in ein Blamage. Zwar waren wiederum über 10 Millionen Stimmzettel verschickt worden und schlussendlich ganze 2.3 Millionen wurden ausgewertet. Literary Digest sagte aber dem falschen Kandidaten den Sieg voraus: es sollte Alfred Landon mit 57 % der Stimmen gewinnen, de facto erhielt er aber nur 38.5%. Das Ergebnis des Siegers Roosevelt wurde gar mit über 20 % überschätzt. Die Blamage vervollständigte die Tatsache, dass eine zuvor verspottete Person mit ihrem American Institute of Public Opinion mit nur wenigen Tausend Stimmen den Sieg Roosevelts richtig voraussagen konnte: George H. Gallup, der als Erfinder der modernen Meinungsforschung gilt, Mehr noch, Gallup, und das musste für die Öffentlichkeit wie ein technisches Wunder klingen, hatte gar das Versagen der Literary Digest, die im übrigen nach diesem Ereignis ihr Erscheinen einstellen musste, ebenfalls vorher gesagt.

Das Projekt des Public Opinoin Polling

George H. Gallup, der Publizistikwissenschaftler (wie seine Tätigkeit wohl heute am ehesten umschrieben werden könnte). Hatte schon seit den früheren dreissiger Jahren mit der bekannt werdenden Stichprobentheorie experimentiert. 1935 präsentierte er sein Verfahren in einem Artikel der Washington Post der Öffentlichkeit. Washington Post stellte auch das Stichprobenverfahren vor: es ist nichts anderes als Kiaers Verfahren des representative Sampling. Die wahrscheinlichkeitstheoretische Fundierung war noch nicht konzeptionell in das Stichprobenverfahren eingelagert, sondern wahrscheinlichkeitstheoretische Berechnungen wurden nur durchgeführt, als ob es sich um eine reine Zufallsstichprobe handelt (es geht das Gerücht, das zuweilen heute noch so verfahren wird). Wiederum kommt die Miniaturvorstellung der repräsentativen Stichprobe zum Zug: "The poll reported today and every poll conducted by the American Institute of Public Opinion is a national election on a small scale."

Gallup verfuhr so, dass er die amerikanische Gesellschaft a priori nach demografischen und ökonomischen Merkmalen in sogenannte cross sections einteilte (soziologisch könnte man von einer Art in die statistische Ordnung übersetzte Milieutheorie sprechen), darauf achtete, dass in seiner Stichprobe eine genügende Anzahl von jeder Cross section erfasst sind. Auf diese Weise, in Verbindung mit Volkszählungsdaten, lässt sich der Meinungs-Anteil in den Stichproben auf die Gesamtbevölkerung hochrechnen.

Der Irrtum der Literary Digest

Aufgrund dieses Stichprobenverfahren erkannte Gallup, dass Literary Digest hinsichtlich der Wahlprognosen 1936 fehl liegen musste. Die Stichprobe von Literary Digest umfasste die polemisch so genannte tel-auto-Öffentlichkeit, Zeitungsabonnenten und InhaberInnen von Automobilen und Telefonanschlüssen. All dies vergegenständlichte dazumals klare Schichtmerkmale, weshalb das Sample von Literary Digest eindeutig auf die Mittleren und oberen Schichten fokussierte. Dies war für die Prognose solange kein Problem, als Schichtmerkmale nicht entscheidend waren für die Wahlpräferenzen, sondern andere Kategorien (beispielsweise der amerikanische Osten versus die ehemaligen Südstaaten). Doch im Zuge der 30er Jahre verschärfte sich die Weltwirtschaftskrise und Roosevelt trat mit einem einmaligen Reformprogramm in den Wahlkampf, weshalb er von den ökonomischen defavorisierten Schichten klar bevorzugt wurde. So sollen Arbeitslose mit bis zu 80% Roosevelt gewählt haben, während die obersten Einkommenskategorien lediglich mit ungefähr 30 Prozent für Roosevelts Partei stimmten.

Dieser Kampf zwischen "haves" und "have-nots" ist Literary Digest rein strukturell entgangen, konnte gar nicht erkannt werden. Ohne geplante Stichprobe konnte Literary Digest nur zählen und die Qualität seiner Stichprobe nicht beurteilen (und genau dies gilt auch für die heutigen Formen der straw polls).

Ein wissenschaftlicher Durchbruch?

Freilich, dieses Ereignis als wissenschaftlicher Durchbruch der quantitativen Sozialwissenschaften zu feiern, ist riskant. Zunächst gilt zu bemerken, dass Literary Digest über mehrere Jahrzehnte hinaus durchaus präzise Aussagen zu liefern vermochte, also "repräsentative Samples" zu erstellen vermochte. Des weiteren ist zu bemerken, dass auch Gallups Methode mehr auf politischer Erfahrung denn auf statistischer Logik beruhte. Seine Miniaturgesellschaft beruhte auf der Annahme, dass, wenn cross sections korrekterweise so gewählt werden, die hauptsächlichen Berufs- und Einkommensgruppen repräsentiert sind, alle anderen Eigenschaften und Meinungen in der amerikanischen Bevölkerung ebenfalls richtig "repräsentiert" sind - was eine heikle, aber in der Stichprobentheorie in ähnlicher Form immer noch vorherrschende Vorannahme bildet. Gallup musste jedenfalls erfahren, dass diese Annahme, dass eine adäquate Abbildung bestimmter Merkmalsverteilungen in der Stichprobe (Berufsgruppen, Geschlecht, Alter) keineswegs bedeutet, dass andere (wie Meinungen) ebenfalls richtig abgebildet werden, als sein Verfahren in den Präsidentschaftswahlen 1948 eine ähnliche Schlappe

erlebte wie Literary Digest, ohne dass bis heute bekannt wäre, weshalb. Die Meinungsforschung jedenfalls hat überlebt -- freilich hin und wieder auch ihre Fehleinschätzungen (vgl. den einleitenden Spiegel-Artikel), denn die Wahlforschung ist eine der wenigen sozialwissenschaftlichen Stichprobenverfahren, deren Adäquatheit nach den Wahlen tatsächlich zu Tage treten vermag.

22. Urnenmetapher

Obwohl heute ein ganzes Feld von statistischen Stichprobenverfahren existieren, die unter dem Stichwort der repräsentativen Stichprobe zusammengefasst sind, ist es notwendig, die basalen Erfordernissen der Übertragbarkeit des Urnenmodells auf soziale Phänomene bewusst zu bleiben. Einige Aspekte dieser nicht immer einfachen Übertragung sollen hier aufgelistet werden. Die Anwendung der Stichprobentheorie fordert, dass der Untersuchungsbereich die wesentliche Eigenschaften einer Urne teilt. Dies zieht die Frage nach sich nach der:

- **Symbolisierbarkeit des Universums**

Um überhaupt "Kugeln" für die Stichproben greifbar zu machen, müssen sie als Katalog, heute zumeist in elektronischer Form vorliegen. Denn nur dann lassen sich, entweder mit Bleistift und Zahlrahmen oder mit Computer die Stichproben auswählen. Es muss sichergestellt werden, dass alle interessierenden Einheiten der Grundgesamtheit überhaupt erfassen lassen. Sie müssen in einem Katalog, mit einem Index versehen, vollständig symbolisch präsent sein. Dies stellt unter Umständen, wie unschwer erkennbar ist, grosse technische wie auch datenschützerische Probleme. Sei es, weil ein Verzeichnis gar nicht vorliegt, sei es weil es nicht zugänglich ist, sei es, weil es unvollständig ist.

-> Nehmen wir an, es sollen die working poor in der Schweiz untersucht werden. Neben dem hier nicht zu diskutierenden Problem, auf welche Weise "working poor" definiert wird, stellt sich die Frage, was heisst eigentlich "in der Schweiz". Gemeint sei wohl die Wohnbevölkerung. Doch gehören hier beispielsweise die Saisonniers und die GrenzgängerInnen dazu, die sehr von dem Phänomen betroffen werden können? Und was geschieht mit den Personen ohne festen Wohnsitz? Selbst das Bundesamt für Statistik kennt eine ganze Reihe von Bevölkerungsdefinitionen -> Dokument

- **Definition der Einheiten**

Zunächst ist die Frage, was die Kugeln, respektive die Einheiten des Sozialen, vergegenständlichen. Es ist eine Einsicht der Soziologie, dass nicht alle soziale Phänomene auf Individuen, respektive Summen von Individuen rückbezogen werden können. Vielleicht interessieren Familien als soziale Einheiten, Betriebe, Organisationen, Gemeinden oder gar Staaten. Wie kann aber heute eine Familie definiert werden? Die Definition dieser "Kugeln" ist mitunter nicht sehr einfach.

- **Zugänglichkeit der Einheiten**

Damit verbunden ist ein weiteres Problem: damit die wahrscheinlichkeitstheoretische Aussagen überhaupt zulässig sind, muss jede Einheit der Urne (der Grundgesamtheit) Universums die gleiche Chance haben, sich in der Stichprobe zu finden. Das heisst, die Kugeln in der Urne müssen gut durchmischt sein und alle für die stichprobengreifende Hand in gleicher Weise zugänglich. Per definitionem ist aber die Urne undurchsichtig. Wer sagt denn, dass in der Urne alle Kugeln gut durchmischt sind, wenn ich nur Stichproben daraus nehmen kann?

Beispiel: Stichprobe, Katalog und Grundgesamtheit bei Telefoninterviews

Anhand der Stichprobenbasis der Telefoninterviews soll diese Praktik kurz diskutiert werden.

- Der weitaus grösste Teil der sogenannten repräsentativen Stichprobe wird mit Telefoninterviews gemacht. Teils ausgeklügte Verfahren sollen verhindern, dass bestimmte Haushaltsmitglieder die grösseren Chance haben, mit dem Befragungsinstitut in Kontakt zu treten. Das Problem bleibt: welche Grundgesamtheit wird hier im Zeitalter von Handy, ISDN, liberalisierten Telekommärkten ohne Eintragspflicht für Telefonnummern erfasst?
- Nicht zuletzt bestehen auch datenschutzrechtliche Probleme bei der Erstellung eines perfekten Katalogs der Bevölkerung. Die Begehren der Statistiker stossen hier auch auf Widerstände in der Bevölkerung selbst
- Es stellt sich damit die Frage, welche Gesellschaft eine wahrscheinlichkeitstheoretische fundierte "repräsentative" Stichprobe repräsentiert, die mit Telefoninterviews gemacht wurde. Genau genommen ist es die folgende: "In Haushalten mit einem eingetragenen Telefonanschluss eines Ortsnetzes der Swisscom lebende Personen, die während des Untersuchungszeitraumes hätten erreicht werden können und sich zu einer Teilnahme an der Befragung entschlossen hätten." (Schnell, S. 259)

Konsequenzen aus den Forderungen der Urnenhypothese

- Was sich hier zeigt ist: die oben genannten Bedingungen sind alles Voraussetzungen, die erfüllt sein müssen, damit die Stichprobe mit Hilfe der Wahrscheinlichkeitsrechnung "repräsentativ" gemacht werden könnte. Denn alleine diese Voraussetzungen lassen die sinnvolle Anwendung der Wahrscheinlichkeitsrechnung zu (angewandt kann sie freilich immer werden), um damit Aussagen über den wahrscheinlichen Streubereich unbekannter Merkmale zu formulieren.
- Deshalb besteht die Auffassung, dass Zufallsstichproben als einzige als einzige Stichprobenart als "repräsentativ" bezeichnet werden können. Es wird gar gefordert, die Begriffe der Wahrscheinlichkeitsstichprobe und der Repräsentativität synonym zu verwenden. Freilich erwies sich dann zumindest der Begriff der statistischen Repräsentativität als überflüssig. Ungeachtet dessen weist er auf Fragen hin, die seit Kiaers Ausführungen nicht alleine über die Anwendung statistisch-mathematischer Theorie lösbar ist. Denn dass die wahrscheinlichkeitstheoretischen Bedingungen vorliegen für die Stichprobenziehung, bleibt im Verlaufe der Stichprobenziehung selbst nicht überprüfbare Hypothese.
- Oder mit anderen Worten: dass die Bedingungen gegeben sind, lässt sich nicht mit den Mitteln Stichprobe selbst überprüfen. Oder mit anderen Worten: die Stichprobe ist gerade dadurch definiert, dass nur einen Teil des Gesamten darstellt, dessen Geltung gerade über die Rahmenbedingungen definiert sind, die selbst die Stichprobe bestimmen.

23. Erreichung von Repräsentativität bei vermuteten Modellverletzungen

Es empfiehlt es sich deshalb, in der Forschungspraxis nicht bloss von der Gegenüberstellung von Grundgesamtheit, die erforscht werden soll, und Stichprobe auszugehen sondern zumindest von einem dreistufigen Modell:

- Der anvisierten Grundgesamtheit (der target population).
- Der zum Katalog erhobenen Stichprobenbasis (frame population, beispielsweise das Telefonverzeichnis).
- Und letztendlich die Stichprobe, deren Basis nicht die Grundgesamtheit ist, sondern notwendigerweise der indizierte Katalog.

Damit ergeben sich ganz verschiedenartige "Repräsentationsprobleme".

Wie ersichtlich ist, besteht zwischen Target- und Framepopulation ein wesentlicher analytischer und technischer Unterschied. Die Diskussion der Qualität der Stichprobe sollte alle drei Ebenen berücksichtigen. Es ist evident, dass ein Rückbezug der Stichprobe, die Darstellung ihrer "Repräsentativität" und der Ausfallquoten auf bloss die Frame Population (beispielsweise wie üblich die telefonisch erreichbaren Haushalte) viele systematische Fehlerquellen ignoriert. Es handelt sich im übrigen um einen gängigen Trick, die Qualität der Stichproben zu schönen.

Strategien um registrierte Verzerrungen zu vermeiden

Um (mutmassliche) Verzerrungen zu vermeiden, kennen die Stichprobentechniker zwei hauptsächliche Verfahren. Beide beruhen darauf, dass Stichprobenmerkmale hinsichtlich ihrer Verteilung mit der Verteilung bekannter Merkmale in der Frame-, seltener in der Targetpopulation in Bezug gesetzt werden. Findet sich, dass ein Merkmal oder eine Merkmalskombination zuwenig vertreten ist, dann bieten sich zwei Strategien an:

- *Poststratification:*
Einheiten mit untervertretenen Merkmalen werden zusätzlich erhoben. Damit werden aber die wahrscheinlichkeitstheoretischen Modellannahmen verletzt und die Errechnung von Stichprobenkennwerte wird problematisch und die Beeinflussung anderer Merkmale ist nicht abzusehen.
- *Gewichtung:*
Die verzerrten Merkmale werden mit einer Gewichtungsfaktor versehen, die mathematisch den untervertretenen Merkmalen mehr Gewicht verleiht und den übervertretenen Merkmalen weniger. Der Einsatz dieser mathematischen Strategie ist umstritten: die Gefahr ist vorhanden, dass der Teufel mit dem Beelzebube ausgetrieben wird. Denn das korrekte Entsprechungsverhältnis von bekannten Merkmalen kann durch eine künstlich hergestellte Verzerrung auf einem unbekannten anderen zur Folge haben.

24. Konsequenzen der wahrscheinlichkeits-theoretischen Stichproben

Aus der Problematik heraus, dass die Voraussetzungen der statistischen Stichproben nicht mit der statistischen Stichprobenziehung selbst geprüft werden können, ergeben sich auch Probleme, so etwas wie eine Repräsentativität, oder "globale Repräsentativität" definiert werden kann. Oder selbst wenn strikte nach wahrscheinlichkeitstheoretischen Regeln vorgegangen wird, bedeutet dies noch nicht, dass die Voraussetzung der Anwendung von wahrscheinlichkeitstheoretischen Sätzen erfüllt sind. Welche Konsequenzen ergeben sich daraus?

Der Begriff der statistischen Repräsentativität sollte vermieden werden. In wissenschaftlicher Hinsicht handelt sich um eine Leerformel. In der renommierten International Encyclopedia of the Social Sciences ist entsprechend zu lesen:

- "There is no such thing as a "representative", "unbiased", "fair", or otherwise "acceptable" sample: such adjectives are strictly applicable to the sampling process that produces the sample, and their application to the sample (...). A sample can be judged only in relation to the process that produced it. The central concepts of selection bias and sampling error have no meaning except in this context."

Man könnte es auch anders formulieren: alle Stichproben sind "repräsentativ" aber es fragt sich für welche Population (Grundgesamtheit).

Sofern der Begriff "Repräsentativität" sich in der Kommunikation mit der wissenschaftlichen Umwelt als unerlässlich erweist, empfiehlt es sich, ihn zu erläutern und dies geschieht idealerweise aufgrund einer allgemeinen Erläuterung des Stichprobenverfahrens mit dem Hinweis auf die Problematik des Begriffs der Repräsentativität. Dies liesse sich beispielsweise so formulieren: die über Telefoninterviews gewonnen Stichprobe entspricht nach dem Merkmalen "Geschlecht" und "Alter" relativ genau der ständigen Wohnbevölkerung der Schweiz. Über weitere Entsprechungsverhältnisse lassen sich keine präzisen Angaben machen oder: weitere Entsprechungsverhältnisse wurden nicht geprüft. Wissenschaftsoptimisten und –optimistinnen können hinzufügen: Es besteht die Vermutung, dass auch die Verteilung anderer Merkmale. Eine solche Aussage ist präziser, ehrlicher und deshalb auch "wissenschaftlicher".

Hinsichtlich einer fachspezifischen Diskussion wäre noch weitere Informationen vonnöten. Beispielsweise:

- Der Definition der Zielpopulation (target population) wie der katalogisierten respektive katalogisierten Gesamtheit (Frame population).
- Der Diskussion der Entsprechungsproblemen von target und frame population.
- Der Diskussion der Ausfälle.

Letztendlich gilt es aber auch zu prüfen, ob die verwendete Gesellschaftsvorstellung überhaupt dem Urnenmodell entspricht. Das heisst, ob überhaupt die untersuchte Gesellschaft als eine Summe isolierter Einheiten verstanden werden kann. Der Soziologe Herbert Blumer war ein hartnäckiger Kritiker der Stichprobentheorie. Paraphrasiert lautet sein Argument wie folgt: Die Gesellschaft funktioniert doch wie ein Organismus. So soll mir doch jemand erklären, wie ich eine repräsentative Stichprobe aus einer Katze ziehen soll.

25. Literaturauswahl

- Bowley, Albert L., 1936: The application of sampling to economic and sociological problems, Journal of the American Statistical Association, 34, S. 474-456.
- Desrosières, Alain, 1993: La politique des grands nombres. Histoire de la raison statistique. Paris, Éditions La Découverte.
- Desrosières, Alain, 2002: Three Studies on the History of Sampling Surveys: Norway, Russia-USSR-, United States, Science in Context, 15, S. 377-383.
- Didier, Emmanuel, 2002: Sampling and Democracy: Representativeness in the First United States Surveys, Science in Context, 15, S. 427-445.
- Gabler, Siegfried, Jürgen H. P. Hoffmeyer-Zlotnik und Dagmar Krebs, 1994: Gewichtung in der Umfragepraxis. Opladen, Westdeutscher Verlag.
- Gallup, George, 1944: A guide to public opinion polls. Princeton, Princeton University Press.
- Keller, Felix, 2001: Archäologie der Meinungsforschung. Mathematik und die Erzählbarkeit des Politischen. Konstanz, UVK.
- Kruskal, William und Frederick Mosteller, 1980: Representative Sampling, IV: the History of the Concept in Statistics, 1895-1939, International Statistical Review, 48, S. 169-195.
- Martin, John L., 1984: The Genealogy of Public Opinion Polling, Annals, 472, S. 12-23.
- Quételet, Lambert Adolphe, 1995: Sur l'homme et le développement de ses facultés ou Essai de physique sociale. Tome seconde (1835). Paris, Bachelier.
- Schnell, Rainer, Paul B. Hill und Elke Esser, 1995: Zum Begriff der Repräsentativität. In: Schnell, Rainer, Paul B. Hill und Elke Esser, Methoden der empirischen Sozialforschung. München, Wien, Oldenbourg, S. 286-295.

- Stuart, Alan, 1968: Nonprobability Sampling. In: David L. Sills (Hrsg.), International Encyclopedia of the Social Sciences, Vol. 13. New York, MacMillan, S. 612-616.