

Der Mikrozensus

1991 erstmals Ergebnisse für München

Was ist der Mikrozensus?

Der Mikrozensus ist eine bundesweite, amtliche Repräsentativstatistik, die die Aufgabe hat, Daten über Bevölkerungszahlen und -strukturen, den Arbeitsmarkt und – allgemeiner – die soziale Situation der Bevölkerung, was etwa Fragen der Wohnung oder der Ausbildung angeht, zu erbringen. Offiziell wird der Mikrozensus denn auch als „Stichprobenerhebung über die Bevölkerung und den Arbeitsmarkt“ bezeichnet.

Durchgeführt wird der Mikrozensus jährlich vom Statistischen Bundesamt und den Statistischen Landesämtern als Haushaltsbefragung: Bundesweit werden 1% der Haushalte – effektiv sind das ungefähr 350.000 – nach einem Zufallsverfahren ausgewählt und interviewt. Das Auswahlverfahren ist insofern objektiv, als die Wahrscheinlichkeit, in die Stichprobe aufgenommen zu werden, für jeden Haushalt gleich groß ist.

Die Haushalte der Stichprobe werden von Interviewern besucht und erhalten zwei Fragebögen, wobei die Fragen des ersten beantwortet werden müssen, was in mehreren Gesetzen, die die juristische Grundlage des Mikrozensus bilden (etwa das Mikrozensusgesetz vom 10. Juni 1985, die Mikrozensusverordnung vom 14. Juni 1985) verankert ist; der zweite Bogen hingegen enthält speziellere Fragen zur Person, deren Beantwortung daher freiwillig ist. Persönliche Daten wie Name, Adresse etc. werden jedoch tatsächlich nicht erhoben, sondern dienen nur als technische Hilfsmittel bei der Durchführung.

Zur Notwendigkeit des Mikrozensus

Dem sogenannten „Durchschnittsbürger“ sind derartige statistische Erhebungen natürlich oftmals suspekt, was nicht zuletzt daran liegt, daß Absicht und Abläufe bei Planung, Durchführung und Auswertung meist im Dunkeln bleiben – zum Beispiel stellt man sich die Frage, warum ausgerechnet die Daten des eigenen Haushalts für den Interviewer interessant sein

sollen und nicht die des Nachbarn. Diese Unklarheit erzeugt zuweilen Mißtrauen.

Dazu ist zunächst zu bemerken, daß von Zeit zu Zeit ein einigermaßen verlässliches Bild der aktuellen Lage vonnöten ist, um aufgrund der daraus gewonnenen Erkenntnisse Entscheidungen zu treffen; unmittelbar klar ist das am Beispiel der Wohnungs- oder Verkehrspolitik. Aus Sicht der Exekutive wären hier natürlich Volkszählungen wie etwa die 1987 durchgeführte ideal, doch stehen der immense legitimatorische und organisatorische Aufwand derartiger Großzählungen solchen Wunschvorstellungen entgegen.

Gezwungenermaßen beschränkt man sich daher auf Stichproben, wobei die mathematische Statistik Mittel und Wege bereitstellt, auch hier zu sinnvollen Ergebnissen zu kommen. Ebenso macht sie aber die Grenzen einer solchen Vorgehensweise deutlich, die eben nicht generell dazu geeignet ist, verlässliche Abschätzungen beliebiger Größen zu liefern. Dies zeigt einerseits, daß solchermaßen gewonnene Daten unter Umständen mit Vorbehalt zu interpretieren sind, andererseits, daß Totalerhebungen, also Volkszählungen, notwendig sind und bleiben; sie sind es sogar für den Mikrozensus, wie wir noch sehen werden.

Was nun die technischen Aspekte der Gesamtdurchführung des Mikrozensus in seiner Form ab 1990 angeht, so sollen diese hier etwas transparenter gemacht werden. Leider ist es nicht möglich, z. B. die Stichprobenerzeugung bis ins letzte Detail zu verfolgen, da das entsprechende Regelwerk einmal sehr umfangreich ist und den hier gebotenen Rahmen bei weitem sprengen würde und zum anderen in seiner Gesamtheit vom Statistischen Bundesamt als Internum behandelt wird (und somit dem Autor bedauerlicherweise nicht zur Verfügung stand). Trotzdem können wir eine ungefähre Vorstellung davon vermitteln, welche Stadien eine solche Erhebung von der Konzeption bis zur Präsentation der Resultate durchläuft und welche Probleme in diesem Zusammenhang auftreten.

Die grundsätzliche Vorgehensweise

Das einer Stichprobenerhebung wie dem Mikrozensus zugrundeliegende Schema ist ein relativ einfaches, wenngleich einige Jahre darauf verwendet wurden, das am besten geeignete Verfahren festzulegen.

Gehen wir ganz allgemein von einer demographischen Stichprobenuntersuchung aus, dann können wir zunächst die folgenden Schritte unterscheiden:

(1) *Abgrenzung der Grundgesamtheit*

Dieser Schritt scheint zunächst trivial – man erwartet, daß zumindest die Grundgesamtheit klar sein sollte.

Tatsächlich sind aber schon hier grundsätzliche und wichtige Überlegungen anzustellen: Welche Bevölkerungsgruppen sind zu erfassen? Ist es sinnvoll, nur deutsche Staatsbürger zu zählen, oder sollen Ausländer mitgezählt werden? Wie sind bei der Zählung nach Haushalten die Haushalte zu erfassen? Was sind eigentlich die „Atome“, aus denen die Grundgesamtheit besteht? Können sie in irgendeiner, hinsichtlich des Endergebnisses günstigen Weise definiert werden? Das sind nur einige Fragen, die vorab geklärt werden müssen, weswegen dieser Schritt tatsächlich keineswegs zu vernachlässigen ist.

(2) Stichprobenerzeugung

Die hier auftretenden Probleme sind schon offensichtlicher: Wir brauchen eine zufällige, repräsentative Stichprobe, doch wie gewährleistet man diese Eigenschaften? Wir werden sehen, daß hier einiger Aufwand getrieben werden muß, da dies von entscheidender Bedeutung für die Qualität der Resultate ist.

(3) Erhebung

Dieser Abschnitt eines jeden Stichprobenverfahrens ist bekannt, er stellt die öffentliche Seite des Gesamtvorgangs dar. Ebenso bekannt dürfte sein, daß auch hier Schwierigkeiten zu bewältigen sind, die beim Mikrozensus vom Finden der „abzugrasenden“ Haushalte bis zur (auf-)richtigen Beantwortung der Fragen reichen.

(4) Kompensation

Will man im Sinne der mathematischen Statistik brauchbare Ergebnisse erhalten, also Ergebnisse, bei denen man mit hinreichender Sicherheit davon ausgehen kann, daß sie die Realität treffend charakterisieren, so muß man den in (3) genannten Schwierigkeiten unbedingt die notwendige Aufmerksamkeit schenken! Konnte beispielsweise beim Mikrozensus die Befragung eines Haushalts nicht durchgeführt werden (was durchaus vorkommt), und trifft auf diesen Haushalt ein seltenes Merkmal zu, so kann das Fehlen der Daten dieses Haushalts die Schätzung dieses Merkmals schon beeinflussen. Derartigen Störungen muß in adäquater Weise Rechnung getragen werden – dazu eben dienen Kompensationsverfahren.

(5) Hochrechnung und Fehlerabschätzung

Der letzte Schritt ist, was seinen Bekanntheitsgrad in der Öffentlichkeit betrifft, von gespaltener Natur: Der Begriff Hochrechnung ist in aller Munde, doch verbindet sich häufig mit ihm die Meinung, man müsse die Stichprobenergebnisse nur hochmultiplizieren um ein Bild der Realität zu erhalten. Kaum bekannt aber ist, daß jede Hochrechnung nur zusammen mit einer

gleichzeitigen Fehlerabschätzung sinnvoll ist, mit Hilfe derer sich angeben läßt, ob aufgrund eines hochgerechneten Ergebnisses überhaupt eine verlässliche Aussage zu machen ist. Das kaum vorhandene Bewußtsein für diese Tatsache kann als einer der Hauptgründe für Mißverständnisse im Zusammenhang mit Repräsentativstatistiken angesehen werden. Diese fünf Schritte sollen im folgenden am Beispiel des Mikrozensus näher erläutert werden. Zuvor aber stellen wir einige Überlegungen zur Stichprobenmethodik an.

Schichten und geschichtete Zufallsauswahl

Gehen wir davon aus, daß die Grundgesamtheit einer durchzuführenden Stichprobenerhebung festliegt, so können wir eine Stichprobe ziehen und erhalten nach dem oben beschriebenen Schema diverse Ergebnisse über die Grundgesamtheit. Das scheint schon das gewünschte Resultat zu sein, ist aber unbefriedigend, sobald innerhalb der Grundgesamtheit in gewisser Weise Unterklassen, für die – zusätzlich zum Gesamtergebnis – Einzelergebnisse interessieren, vorgegeben sind.

Wie solche Unterklassen aussehen können und eingeteilt werden, ist am Beispiel des Mikrozensus sofort klar: Offenbar möchte man hier Schätzungen nach Regionen differenzieren können, da verschiedene Regionen völlig verschiedene Strukturen aufweisen können (eine Unterscheidung, die bei Ausmittlung über das gesamte Bundesgebiet verwischt würde) – man denke beispielsweise nur an das bei vielen Merkmalen auftretende Nord-Süd-Gefälle oder, noch aktueller, die zahlreichen Gegensätze zwischen alten und neuen Bundesländern. Tatsächlich sind hier oftmals bundesweite Ergebnisse – eben wegen solcher erheblichen Inhomogenitäten – kaum von Interesse; regionale Ergebnisse stehen eindeutig im Vordergrund.

Schichten sind nun genau diese Unterklassen, in die die Grundgesamtheit aufgeteilt wird. Im Falle des Mikrozensus handelt es sich dabei zunächst um Regionen innerhalb Deutschlands (daher regionale Schichten), darüber hinaus aber auch innerhalb der regionalen Schichten verschiedene, nach der Gebäudegröße zu unterscheidende Arten von Gebäuden (sog. Gebäudeschichten). Dabei stellt München nunmehr eine eigene regionale Schicht dar, so daß der Mikrozensus auch spezielle Aussagen (s. u.) über die Münchener Verhältnisse ermöglicht.

Wie solche Schichten im konkreten Fall zu bilden sind, dafür liefert die mathematische Statistik einige Kriterien: als wichtigstes ist zu nennen, daß jede Schicht für

sich über eine möglichst homogene Struktur verfügen soll, während die Schichten untereinander strukturell möglichst verschieden sein sollen. Dies setzt offenbar schon ein Vorwissen über die interessierenden Merkmalsstrukturen voraus.

Die Auswahl der Stichprobe muß nun offenbar diese Unterteilung in Schichten in angemessener Weise berücksichtigen. Ein naheliegender und beim Mikrozensus tatsächlich zur Anwendung kommender Ansatz ist, je nach Anteil einer Schicht an der Grundgesamtheit einen entsprechenden Anteil der Stichprobe aus der jeweiligen Schicht zu ziehen. Diese Art, die Stichprobe auf die einzelnen Schichten zu verteilen, heißt *proportionale Aufteilung* und liefert im allgemeinen gute Ergebnisse.

Klumpen

Nun wäre es zu einfach, wenn sich die in der mathematischen Statistik zunächst vorausgesetzten Idealbedingungen in der Welt des statistischen Anwenders wiederfänden. Die hier gemeinte „Idealbedingung“ ist die problemlose Gewinnung von Daten, eine Bedingung also, die den Schritt (3) des obigen Schemas betrifft.

Beim Mikrozensus wird dieses Problem deutlich, wenn man den Aufwand abschätzt, der notwendig ist, um Interviewer in 350.000 gleichmäßig verteilte Haushalte zu schicken. Bei einer Fläche des Bundesgebiets von ca. 357.000 km² bedeutet dies, daß ungefähr ein Haushalt pro Quadratkilometer besucht und damit natürlich von jedem Interviewer eine immense Wegstrecke zurückgelegt werden müßte.

Aus diesem Grund faßt man mehrere Haushalte zu einem sog. *Klumpen* zusammen und definiert diese Klumpen gewissermaßen als „Atome“ der Grundgesamtheit. Der gewonnene Vorteil eines geringeren Aufwands muß aber natürlich bezahlt werden, und zwar mit einem – abhängig von Klumpengröße und geschickter Zusammenlegung – mehr oder weniger großen Präzisionsverlust bei den erhobenen Daten, dem sog. *Klumpeneffekt*. Die Theorie legt dabei nahe, bei der Verklumpung darauf zu achten, die Klumpen untereinander möglichst ähnlich zu wählen, innerhalb eines jeden Klumpens aber für Inhomogenität zu sorgen.

Ein Klumpen der Mikrozensus-Auswahlgrundlage umfaßt im Mittel neun Wohnungen, wodurch offensichtlich die von den Interviewern zurückzulegenden Wege schon erheblich verkürzt werden. Genaueres darüber findet sich in den folgenden Abschnitten.

Nach diesen allgemeinen methodischen Überlegungen jetzt aber zur konkreten Vorgehensweise beim Mikrozensus.

Die Auswahlgrundlage

Bei der als Schritt (1) bezeichneten Abgrenzung der Grundgesamtheit darf beim Mikrozensus nach dem in diesem Fall relevanten Volkszählungsgesetz von 1987 nur gebäudeorientiert vorgegangen werden. Als Informationen stehen danach für jedes Gebäude die Zahl der Wohnungen und die Gesamtzahl der wohnberechtigten Personen gegliedert nach Straße und Hausnummer zur Verfügung, wobei diese Daten der Volkszählung von 1987 entnommen werden. Da sich aber offensichtlich seitdem ständig Veränderungen ergeben, muß auch die behördliche Bevölkerungsfortschreibung berücksichtigt werden, mit Hilfe derer notwendige Aktualisierungen durchgeführt werden.

Regionale Schichten

Regionale Schichten sind (bis auf wenige Ausnahmen) geographisch zusammenhängende Flächen, für deren Größe folgende Kriterien zugrundegelegt werden: Als Untergrenze für die Bevölkerungszahl einer regionalen Schicht werden 250.000 Einwohner bzw. bei Großstädten 200.000 Einwohner festgesetzt; landesweit soll aber ein Durchschnitt von 300.000 Einwohnern pro regionaler Schicht erreicht werden. Die Bildung dieser Schichten geschieht im allgemeinen durch Zusammenlegung von Landkreisen, wobei Regierungsbezirksgrenzen nicht überschritten werden.

Dadurch ergeben sich in Bayern 33 regionale Schichten, die im Mittel 330.000 Einwohner (allerdings mit sehr großer Streuung) umfassen. Dies zeigt, daß die Erzeugung regionaler Schichten in der Praxis auf Schwierigkeiten stößt, die wohl vor allem darin begründet sind, daß die Schichten, wie oben schon erwähnt, in ihrer Merkmalsstruktur möglichst homogen sein sollen – ein Umstand, der offensichtlich der Bildung großer Einheiten entgegensteht.

Die regionalen Schichten stellen die Gliederungsgrundlage für alle anderen Untergliederungen und Zusammenfassungen dar.

Regionale Untergruppen

Vor allem zum Zwecke der Ordnung der noch zu definierenden sog. *Auswahleinheiten* untergliedert man die im ersten Schritt erzeugten regionalen Schichten in jeweils eine oder mehrere *Regionale Untergruppen*. Eine solche regionale Untergruppe soll dabei mindestens 100.000 Einwohner umfassen.

In Bayern werden auf diese Art und Weise 50 regionale Untergruppen erzeugt, deren Einwohnerzahl im Mittel bei 156.000 Einwohnern liegt (wiederum mit großer Streuung).

Anpassungsschichten

Die zunächst gebildeten regionalen Schichten sind mit ihren durchschnittlich 330.000 Einwohnern noch zu klein, als daß Hochrechnungen innerhalb dieser Schichten aussagefähige Ergebnisse liefern können. Daher geht man dazu über, eine oder mehrere regionale Schichten zu sog. *Anpassungsschichten* mit mindestens 500.000 Einwohnern zusammenzufassen (Richtwert). Dabei sollen aber Regierungsbezirksgrenzen wiederum nicht überschritten werden.

Aus dieser Zusammenlegung gehen in Bayern 22 Anpassungsschichten mit durchschnittlich 494.000 Einwohnern (bei sehr großen Schwankungen!) hervor. Das Stadtgebiet München bildet dabei eine eigene Anpassungsschicht.

Neben der regionalen Schichtung wird – auch das wurde schon gesagt – eine Gebäudeschichtung eingeführt.

Gebäudeschichten

Da die Erhebungsmethode des Mikrozensus haushaltsspezifisch ist, und da es außerdem naheliegend ist, daß in verschiedenen Gebäuden (etwa einem Einfamilienhaus und einem Miethochhaus) unter vielen Aspekten verschiedenartige Haushalte zu finden sind, bietet es sich an, auch bezüglich Gebäudekriterien eine Schichtendifferenzierung vorzunehmen.

Die vorhandenen Gebäude (wobei als „Grundregister“ die oben genannten Volkszählungsdaten und die behördliche Fortschreibung dienen) werden daher je nach Anzahl ihrer Wohnungen und Bewohner jeweils einer der folgenden vier *Gebäudeschichten* zugeordnet: Die Schicht 1 umfaßt „kleine Gebäude“, das sind Gebäude mit mindestens einer, höchstens vier Wohnungen und einer gewissen Höchstzahl von Personen. In der Schicht 2, der Schicht der „mittleren Gebäude“ finden sich alle Gebäude mit fünf bis zehn Wohnungen und wiederum einer Höchstzahl von Bewohnern.

Die Gebäude der Schicht 3 werden als „große Gebäude“ bezeichnet und haben mindestens elf Wohnungen und eine Bewohnerzahl, die eine gewisse Schranke nicht überschreitet.

Die Schicht 4 schließlich enthält alle Gebäude mit mindestens einer Wohnung, deren Bewohnerzahl aber die Schranken für die ersten drei Schichten sprengt. Außerdem gehören alle Gebäude ohne Wohnungen, aber mit mindestens einem Bewohner zu dieser Schicht. Da in diese Kategorie vor allem Krankenhaus-, Kasernen- oder ähnliche Gebäude fallen, heißt diese Schicht die Schicht der „Anstaltsgebäude“.

Die Zuordnung der Gebäude zu den Schichten ist noch weitaus genauer definiert, so daß tatsächlich

jedes Gebäude in eindeutiger Weise klassifiziert werden kann; dies muß uns aber hier im Detail nicht interessieren.

Auswahleinheiten – die „Atome“ der Grundgesamtheit

In dem Abschnitt über Klumpen war schon von der Zusammenfassung von mehreren Haushalten zu größeren Einheiten, eben „Klumpen“, die Rede. Diese Bezeichnung als Klumpen ist der Stichprobentheorie entlehnt und wird im Falle des Mikrozensus durch die spezielle Bezeichnung „Auswahleinheit“ ersetzt. Diese Auswahleinheiten sind genau die zu definierenden Grundeinheiten und bilden die „Atome“ für alle folgenden Schritte, insbesondere die Auswahl der Stichprobe.

Auswahleinheiten sind also ganz einfach Zusammenfassungen von mehreren Haushalten, wobei diese nicht beliebig geschehen können, sondern sich nach der beabsichtigten Schichtung der Stichprobe richten müssen. Dabei ist die regionale Schichtung zu beachten, vor allem aber die Gebäudeschichtung, die auf der Mikroebene der Haushalte offensichtlichere Einschränkungen mit sich bringt. Innerhalb jeder Auswahleinheit müssen also alle Haushalte ein und derselben Regional- und Gebäudeschicht angehören.

Die Erzeugung der Auswahleinheiten

Um zunächst die regionale Schichtung zu gewährleisten, werden alle Gebäude der Grundgesamtheit pro Gemeinde (bzw. pro regionaler Untergruppe, falls eine Gemeinde in regionale Untergruppen aufgeteilt ist) nach Gemeindeteil, Straße und Hausnummer (einschließlich Alpha-Zusatz) sortiert. Ausgehend von dieser Reihenfolge beginnt man dann, innerhalb jeder Straße Haushalte zu Auswahleinheiten zusammenzulegen, wobei wegen der Gebäudeschichtung schichtenweise getrennt vorgegangen wird. Die Auswahleinheiten erhalten hierbei in der Reihenfolge ihrer Bildung laufende Nummern.

In der Schicht 1 besteht eine Auswahleinheit aus mehreren kleinen Gebäuden; man strebt dabei eine Größe von zwölf Haushalten an und läßt im allgemeinen einen Streubereich von 10 bis 13 Haushalten zu. Da das Einteilungsverfahren straßenweise abläuft, können am Ende einer Straße einige Wohnungen übrigbleiben, die zusammen mangels Quantität keine eigene Auswahleinheit mehr bilden können. Diese

Straßenreste faßt man zunächst innerhalb eines Gemeindeteils, dann innerhalb der Gemeinde zu Auswahleinheiten zusammen, wobei man den üblichen Streubereich von 10 bis 13 Wohnungen pro Auswahleinheit beachtet.

Die mittleren Gebäude der Schicht 2 bilden jeweils eine eigene Auswahleinheit. Als Durchschnittsgröße ergeben sich hier sieben Wohnungen pro Auswahleinheit.

Für jedes der großen Gebäude der Schicht 3 berechnet man aus der Zahl seiner Wohnungen sowie der durchschnittlichen Zahl der Personen pro Wohnung die Zahl der zu erzeugenden Auswahleinheiten. Die konkrete Erzeugung der Auswahleinheiten in einem solchen Gebäude erfordert nun aber eine genaue Kenntnis der Gebäudestruktur, also der Anzahl der Geschosse, der Anzahl der Wohnungen in jedem Geschos und weitere Angaben. Diese Informationen stehen jedoch nicht zur Verfügung bzw. können nur per Lokaltermin gewonnen werden (da zur Lokalisierung nur Straße und Hausnummer verwendet werden, was für die Auswahleinheiten der ersten beiden Schichten ja genügt). Daher beschränkt man sich auf die gewissermaßen „abstrakte“ Zerlegung des Gebäudes und numeriert die zu erzeugenden Auswahleinheiten innerhalb des Gebäudes durch. In der Grundgesamtheit werden diese Auswahleinheiten nur mit ihrer Nummer unter der jeweiligen Adresse geführt, also z. B. als vierte Auswahleinheit in der Schloßallee 5. Erst wenn eine dieser Auswahleinheiten in die Stichprobe aufgenommen wird, wird sie konkret innerhalb des Gebäudes abgegrenzt. Bei dieser Abgrenzung orientiert man sich im wesentlichen an Geschossen und strebt eine Richtgröße von sechs Wohnungen an.

Eine Sonderregelung tritt noch in Kraft, wenn sich in einem Gebäude dieser Schicht (der Schicht 3) eine Anstalt befindet, d. h. ein Gebäudeteil mit Wohnungen, die zu nichtselbständigen Haushalten gehören. Ein solcher Anstaltsteil wird (vor der Zerlegung) vom restlichen Gebäude abgetrennt, und die beiden Teile werden separat in die vorher berechnete Teilezahl zerlegt. Dann wird jeweils ein Teil des Anstaltsteils mit einem Teil des restlichen Gebäudes zu einer Auswahleinheit zusammengefaßt. Damit erhält man wieder die gewünschte Anzahl von Auswahleinheiten.

Auch hier gibt es noch wesentlich detailliertere „Spielregeln“, deren genaue Beschreibung aber zu weit führen würde. Der Grund für diese Vorgehensweise ist aber klar: Die nur vor Ort durchführbare Abgrenzung der Auswahleinheiten erfordert, falls sie für alle Gebäude der Schicht 3 erfolgt, einen ungeheuren Aufwand, der vollends unvertretbar wird, wenn man bedenkt, daß sie später nur für ein Prozent der Auswahleinheiten tatsächlich interessiert. Dieses Argument gilt zwar prinzipiell auch für die Auswahleinheiten

der Schichten 1 und 2, doch dort kann die Abgrenzung sofort und ohne großen Aufwand (vielleicht sogar per Computer) geschehen. Die Erzeugung einer Auswahleinheit nach ihrer Aufnahme in die Stichprobe macht aber offenbar eine Genauigkeit und Objektivität der Abgrenzungsregeln nötig, die bewußte oder unbewußte Beeinflussungen so weit wie möglich ausschließen, denn diese würden ja der geforderten Zufälligkeit des Auswahlverfahrens zuwiderlaufen.

In der Schicht 4, der Schicht der Anstaltsgebäude, geht man ähnlich vor. Da hier Wohnungen als Gliederungskriterium weitgehend fehlen, benutzt man hier grundsätzlich die Nachnamen der Bewohner als Ordnungsschlüssel, wozu genauer zu bemerken ist, daß eine Wohnung in einem solchen Gebäude nicht (wie in den anderen Schichten) durch ihre Lage innerhalb des Gebäudes spezifiziert wird (also nach Geschos o. ä.), sondern durch den Namen eines Bewohners, der die jeweilige Wohnung gewissermaßen repräsentiert. Zunächst berechnet man aber aus der Gesamtzahl der Bewohner eines Gebäudes nach einem vorgegebenen Schema die Anzahl der Auswahleinheiten, in die das Gebäude zerlegt werden soll. Wie bei den Gebäuden der Schicht 3 werden dann aber nur die zur Stichprobe gehörigen Auswahleinheiten tatsächlich abgegrenzt. Dabei unterscheidet man allerdings zwischen dem *Anstaltsteil* und dem *Wohnungsteil* eines Gebäudes: dem Anstaltsteil werden die Wohnungen zugerechnet, die durch Personen ohne eigene Haushaltsführung repräsentiert sind, im Wohnungsteil finden sich dagegen die Wohnungen der Personen mit eigener Haushaltsführung. Die Zerlegung des Anstaltsteils des Gebäudes in die vorher berechnete Anzahl von Teilen geschieht einfach durch Einteilung möglichst gleich großer Personengruppen in der alphabetischen Ordnung der Nachnamen der Bewohner. Unabhängig davon wird der Wohnungsteil nach ähnlichen Regeln wie den für die Schicht 3 geltenden in ebenso viele Teile zerlegt und anschließend je ein Teil aus dem Anstaltsteil und ein Teil aus dem Wohnungsteil des Gebäudes zu einer Auswahleinheit zusammengefaßt. Dieses Verfahren ergibt gerade die angestrebte Anzahl von Auswahleinheiten in dem Gebäude.

Sinn der Unterscheidung zwischen Anstalts- und Wohnungsteil bei Gebäuden der Schichten 3 und 4 ist es, auch die nicht in eigenständigen Haushalten lebende Bevölkerung (Senioren in Altenheimen, Soldaten etc.) in der Stichprobe geeignet zu repräsentieren.

An dieser Stelle sollte vielleicht bemerkt werden, daß damit erst der erste Schritt des anfangs gegebenen stufigen Schemas erklärt ist, und dies auch nicht bis ins Detail. Der zunächst so trivial scheinende Schritt der Definition der Grundgesamtheit war also in der Tat ein sehr schwieriger!

Die Stichprobenziehung

Erstes Kennzeichen eines Stichprobenverfahrens ist der sog. *Auswahlsatz*, also der Stichprobenumfang relativ zum Umfang der Grundgesamtheit. Beim Mikrozensus beträgt dieser Auswahlsatz $\frac{1}{100}$, also 1%, und man spricht folglich von einer 1%-Stichprobe. Gezogen werden aber tatsächlich zwanzig solcher 1%-Stichproben (was nach dem Volkszählungsgesetz von 1987 zulässig ist), wobei nur die erste Stichprobe ausgewertet wird und die restlichen 19 Stichproben lediglich dazu dienen, evtl. zu oft schon in die Stichprobe aufgenommene Auswahleinheiten zu ersetzen. Als Zusatzbedingung fordert man nämlich, daß keine Auswahleinheit mehr als viermal in die Mikrozensusstichprobe einbezogen wird. Wird sie zum fünftenmal gezogen, wird sie durch eine Auswahleinheit aus einer der 19 sog. Ersatzstichproben ersetzt. Dieses Ersetzungsverfahren nennt man auch *Rotation*.

Wir betrachten hier genauer die Erzeugung der Stichprobe, denn an diesem Punkt entscheidet sich offenbar wesentlich die Güte (im Sinne des schwer faßbaren Begriffs „Zufälligkeit“) der gesamten statistischen Erhebung. Die im folgenden beschriebenen Schritte laufen, um die Schichtung der Gesamtheit und die proportionale Aufteilung der Stichprobe zu gewährleisten, innerhalb jeder Regional- und Gebäudeschicht getrennt ab.

Wie oben schon einmal gesagt, werden die Auswahleinheiten innerhalb einer Gemeinde in der Reihenfolge ihrer Erzeugung mit laufenden Nummern versehen. Um sie nun auch gesamtregional zu ordnen, werden sie (wie gesagt: innerhalb jeder Regional- und Gebäudeschicht) nach den Kriterien regionale Untergruppen, Kreis, Gemeindegrößenklasse, Gemeinde und eben dieser laufenden Nummer sortiert. In der so gegebenen Reihenfolge werden jetzt aufsteigend je 100 Auswahleinheiten zu sog. *Zonen* zusammengefaßt. Außerdem bilden je vier aufeinanderfolgende Zonen einen sog. *Block*. Innerhalb jedes Blocks werden die Zonen zufällig mit den Zahlen 1 bis 4 durchnummeriert und auf diese Weise genau einem von vier sog. *Rotationsvierteln* zugeordnet. Dies ist die erste Ordnungsstruktur der Grundgesamtheit.

Um eine zweite, sozusagen zur ersten querliegende Ordnungsstruktur zu schaffen, werden innerhalb jeder Zone die Auswahleinheiten zufällig mit den Zahlen 1 bis 100, den sog. *Stichprobennummern*, durchnummeriert und dementsprechend neu angeordnet. Dadurch sind die Auswahleinheiten einerseits in Zonen, andererseits in *Stichproben* gegliedert.

Aus der Kombination von 100 Stichprobennummern und vier Rotationsviertelnummern ergeben sich 400 verschiedene („schnittfreie“) Klassen von Auswahleinheiten, die wir, um die Bezeichnungsweise zu erleichtern, im folgenden *Rotationsklassen* nennen

(auch wenn das kein Ausdruck der üblichen Mikrozensus-Nomenklatur ist).

Bei der zufälligen Numerierung der Auswahleinheiten innerhalb der Zonen und der Zonen innerhalb der Blöcke müssen natürlich Verfahren verwendet werden, die für „gleichmäßige Zufälligkeit“ sorgen (also dafür, daß jede Numerierungsreihenfolge gleich wahrscheinlich ist), denn erst durch eine derartige Vorgehensweise ist für jede Rotationsklasse die Zuordnung einer bestimmten Auswahleinheit gleich wahrscheinlich (die Wahrscheinlichkeit beträgt also $\frac{1}{400}$ oder 0,25 %). Dies aber ist gleichwertig mit der geforderten, ganz wesentlichen Eigenschaft, daß für jeden Haushalt die Chance, in die Mikrozensus-Stichprobe aufgenommen zu werden, gleich groß ist. Eine weitere wünschenswerte Eigenschaft, nämlich die gleichmäßig dichte Stichprobenauswahl innerhalb der regionalen Schichten (also konkret am Beispiel München: die gleichmäßige Verteilung der Stichprobenauswahleinheiten über das gesamte Stadtgebiet im Gegensatz zur Konzentration auf einen Stadtbezirk, die kaum ein repräsentatives Bild liefern würde) wird außerdem durch das spezielle Stichprobenerzeugungsverfahren garantiert, nämlich zunächst die regional strukturierte Anordnung und dann die beschriebene zufällige Numerierung in zwei Schritten.

Die Stichprobe mit der Nummer 1 ist nun „die“ Mikrozensus-Stichprobe, die darauffolgenden Stichproben 2 bis 20 sind die schon erwähnten *Vorratsstichproben*, die vor allem für die Rotation wichtig sind; denn um zu verhindern, daß einzelne Auswahleinheiten mehr als viermal in die Stichprobe miteinbezogen werden, wird nach einem hier nicht näher zu erläuternden Verfahren jährlich ein Viertel der Stichprobe, nämlich gerade eine Rotationsklasse der Stichprobe, mit einer Rotationsklasse aus den Vorratsstichproben ausgetauscht. Dies motiviert die Bildung der Rotationsviertel.

Damit ist die gewünschte Stichprobe im wesentlichen erzeugt. „Im wesentlichen“ meint: bis auf Zusätze, die im nächsten Abschnitt behandelt werden.

Aktualisierung durch Neubauswahleinheiten

Wie oben schon erläutert wurde, dienen zur Bestimmung der Auswahleinheiten zunächst Volkszählungsdaten, aktuell also Daten aus dem Jahre 1987. Nun dürfte klar sein, daß einige interessierende Daten sich seitdem nicht unwesentlich geändert haben und daher eine Aktualisierung notwendig ist, will man eine einigermaßen aktuelle Situationsbeschreibung erhalten. Also müssen auch Gebäude, die nach dem Stichtag der letzten Volkszählung (dem 25. Mai 1987) errichtet wurden, in der Auswahl berücksichtigt werden, d. h.

die Grundgesamtheit muß noch erweitert werden. Eine Übersicht über die betreffenden Gebäude garantiert dabei die behördliche Registrierung der Neubautätigkeit.

Innerhalb der regionalen Schichten bilden die Neubauten eine eigene zusätzliche Schicht, die wegen zu geringer Fallzahlen aber nicht mehr genauer (nach Gebäudegröße) differenziert wird. Bei der Bildung von Auswahlseinheiten, die auch hier stattfinden muß, werden aber die Gebäudeschichten trotzdem berücksichtigt, das heißt: nur Gebäude derselben Gebäudeschicht werden jeweils zusammengefaßt. Damit können auch die oben beschriebenen bisherigen Regeln für die Zusammenlegung kleiner Gebäude bzw. die Aufteilung größerer Gebäude wiederverwendet werden. Die bei den „normalen“ Auswahlseinheiten angewandte Vorgehensweise nach Straßen ist aber offensichtlich nicht mehr sinnvoll, da im allgemeinen innerhalb einer Straße kaum genügend Neubauten (ein und derselben Gebäudeschicht!) für eine Auswahlseinheit vorhanden sein werden, was offenbar vor allem für Gebäude der Schichten 2, 3 und 4 gilt. Daher erzeugt man die Auswahlseinheiten der Neubauten – kurz: *Neubauauswahlseinheiten* – innerhalb von Gemeinden und, falls auch dieser Rahmen noch zu eng ist, innerhalb von Landkreisen. Wie zuvor schon erhalten die Neubauauswahlseinheiten in der Reihenfolge ihrer Bildung laufende Nummern und werden so geordnet.

Um nun zwanzig 1%-Stichproben der durch die Neubauauswahlseinheiten erweiterten Grundgesamtheit zu erhalten, zieht man gewissermaßen zwanzig 1%-Stichproben aus der Gesamtheit der Neubauauswahlseinheiten und faßt diese mit den schon gezogenen Stichproben der ursprünglichen Grundgesamtheit zusammen. Der Ziehung liegt auch hier ein festgeschriebener Algorithmus zugrunde, der, um regionale Schichtung zu berücksichtigen, für jede regionale Schicht einzeln durchgeführt wird: Durch ein Zufallsverfahren (das wiederum gewissen Objektivitätskriterien genügt) wird eine laufende Nummer bestimmt, ab der 20 Neubauauswahlseinheiten in die schon gezogenen zwanzig 1%-Stichproben übernommen werden. Diese laufende Nummer wird dann um 100 erhöht, woraufhin man ab der sich ergebenden laufenden Nummer wieder 20 Neubauauswahlseinheiten übernimmt. Dieses Verfahren wird fortgesetzt, bis schließlich zwanzig 1%-Stichproben aus den Neubauauswahlseinheiten erzeugt sind.

Damit ist auch der Schritt (2) des anfangs gegebenen Grobschemas abgeschlossen.

Die Erhebung der Daten

Über diese Phase bei der Durchführung des Mikrozensus sollen hier keine ausführlicheren Bemerkun-

gen gemacht werden, da vom methodischen Standpunkt aus nur ein Arbeitsschritt, nämlich das Ausfüllen des Fragebogens durch die Stichprobenhaushalte, erledigt werden muß. Nichtsdestoweniger sollte man sich darüber im klaren sein, daß auch dieser Schritt nicht trivial ist – einige Beispiele für mögliche Probleme wurden schon genannt. Da wir uns aber eher für die algorithmische Seite der Erhebung interessieren und sich die hier auftauchenden Probleme nur schwer in ein allgemeines formales Schema pressen lassen, gehen wir zum nächsten Schritt, der Kompensation, über.

Kompensation

Haben wir die systematische Beschreibung der Probleme bei der konkreten Datenerhebung gerade noch übergangen, so müssen wir uns nun mit denselben Schwierigkeiten von einem anderen Standpunkt aus befassen; genauer: da manche dieser Probleme nicht nur dazu geeignet sind, dem Interviewer Kopfschmerzen zu bereiten, sondern darüber hinaus wichtige Ergebnisse erheblich verfälschen können, ist zu überlegen, ob und wie dadurch entstandene Verzerrungen ausgeglichen werden können. Vor allem geht es hier um die Tatsache, daß die beabsichtigte Datenerhebung bei manchen Haushalten nicht durchgeführt werden kann (da diese nicht erreichbar sind oder dies womöglich vortäuschen), daß also manche Merkmale in der Stichprobe öfter auftreten, als dies den ausgefüllten Fragebögen zu entnehmen ist. Solche Ausfälle mögen für einige Merkmale unerheblich sein (z. B. für die Erwerbsquote, allgemeiner für häufige Merkmale), führen aber schnell zu Verzerrungen, sobald sie bei seltenen Merkmalen auftreten (z. B. alleinerziehenden Vätern). Die *Kompensation* dient nun dazu, diese Fehlstände in einigermaßen vernünftiger Weise zu korrigieren, eben zu kompensieren.

Das beim Mikrozensus angewandte Kompensationsverfahren beruht auf zwei Grundideen. Die erste, intuitiv völlig klare Idee ist, aus der Anzahl der gegebenen Antworten und der Anzahl der Ausfälle einen Faktor zu berechnen, mit dem die gewonnenen Fallzahlen eines Merkmals multipliziert werden. Man extrapoliert sozusagen aus den vorhandenen Daten und unterstellt dabei, daß die interessierenden Merkmale in dem Teil der Stichprobe, der für die Ausfälle verantwortlich ist, genauso verteilt sind wie in der restlichen Stichprobe. Dazu sollte bemerkt werden, daß diese Annahme für manche Merkmale sicher mehr als fragwürdig ist – in der Praxis jedoch ist man auf sie angewiesen, will man nicht in reine Spekulation verfallen. Die zweite Idee ist, evtl. vorhandene anderweitige Informationen (z. B. aus den Volkszählungsunterlagen, soweit sie für den Mikrozensus verwendet wer-

den dürfen) dazu zu benutzen, den betroffenen Haushalt einer sog. Kompensationsklasse bzw. einem Kompensationstyp zuzuordnen. Man führt also die Kompensation nicht gleichmäßig über die gesamte Stichprobe durch, was Strukturunterschiede nur wieder verwischen würde, sondern differenziert nach verschiedenen Klassen, deren spezifische Merkmalsstruktur (unter der oben erwähnten Annahme) somit erhalten bleibt.

Zu diesem Zweck unterscheidet man 19 verschiedene *Kompensationsklassen* und unterteilt diese noch feiner in insgesamt 51 verschiedene *Kompensationstypen*. Als Differenzierungskriterien gelten die Haushaltsgröße (eine Person, zwei Personen, mehr als zwei Personen), bei Einpersonenhaushalten das Geschlecht und das Alter (jünger oder älter als 60 Jahre) des Bewohners, außerdem die Staatsangehörigkeit (Deutscher oder Ausländer) der den Haushalt repräsentierenden Person und bei Deutschen die Frage, ob die betreffende Wohnung als Haupt- oder Nebenwohnsitz gilt; zusätzlich sieht man noch eine eigene Kompensationsklasse für ausgefallene Personen in Gemeinschaftsunterkünften (d. h. vor allem Gebäuden der Schicht 4) vor. Insbesondere sind die Kompensationstypen also haushaltsbezogen. Die Kompensationstypen lassen auch fehlende Informationen zu, erfassen also auch Haushalte, bei denen nur über einen Teil (oder im Extremfall über gar keines) der Klassifikationsmerkmale etwas bekannt ist. Das ist auch notwendig, da die Informationen, aufgrund derer über die Zuordnung eines ausgefallenen Haushalts zu einem der Kompensationstypen entschieden wird, nicht aus behördlichen Quellen kommen können (da sie dort nicht vorhanden sind) und damit nicht als „vollständig“ garantiert werden können. Vielmehr ist es hier notwendig, auf Informationen aus früheren Mikrozensus-Befragungen, an denen der betreffende Haushalt eventuell schon beteiligt war, zurückzugreifen bzw. es der Initiative des Interviewers vor Ort zu überlassen, möglichst viele der Zuordnungsmerkmale des Haushalts herauszufinden. Weniger schwierig ist es schließlich, auch regionale Differenzen nicht einzuebnen: Man unterscheidet dazu zusätzlich noch zwischen regionalen Untergruppen und führt also insgesamt die Kompensation in Klassen durch, die sowohl durch soziale als auch durch regionale Kriterien definiert sind.

Für jede dieser Klassen errechnet man einen (haushaltsbezogenen) *Kompensationsfaktor* k gemäß

$$k = \frac{b+a}{b} \geq 1,$$

wobei b die Anzahl der tatsächlich befragten Haushalte und a die Anzahl der ausgefallenen Haushalte bezeichnet. Offenbar ist k nur geringfügig größer als eins. Die so gewonnenen Kompensationsfaktoren faßt man innerhalb der Kompensationsklassen (getrennt

nach regionalen Untergruppen) zu sog. *verdichteten Kompensationsfaktoren* zusammen, und zwar genauer folgendermaßen: Sind

$$k_1, \dots, k_r$$

die Kompensationsfaktoren der zu einer bestimmten Kompensationsklasse (und einer bestimmten regionalen Untergruppe) gehörenden Kompensationstypen, so ist der verdichtete Kompensationsfaktor K dieser Kompensationsklasse in dieser regionalen Untergruppe definiert als

$$K = k_1 \dots k_r.$$

Also hat man 19 verdichtete Kompensationsfaktoren pro regionaler Untergruppe (entsprechend den 19 Kompensationsklassen). Die Kompensationsklassen und -faktoren werden uns im nächsten Abschnitt, der die Vorgehensweise bei der Hochrechnung der gewonnenen Ergebnisse beschreibt, wieder begegnen.

Hochrechnung und Fehlerabschätzung

Offensichtlicher Grundgedanke eines jeden Stichprobenverfahrens ist es, daß die in der Grundgesamtheit vorhandenen relativen Anteile eines jeden interessierenden Merkmals sich in der Stichprobe (annähernd) wiederfinden, daß also Proportionen (weitgehend) erhalten bleiben: das ist gemeint, wenn von einer repräsentativen Stichprobe die Rede ist. Zur Erzeugung einer derartigen Stichprobe im Spezialfall des Mikrozensus wurde in den vorherigen Abschnitten schon einiges berichtet, aber es sollte klar sein, daß trotz aller regionalen und anderweitigen Differenzierungen ein exaktes, proportional verkleinertes Abbild der Grundgesamtheit praktisch nicht erreichbar ist. Der Grund dafür ist offensichtlich die Zufälligkeit der Auswahl, die zwar – wie sich mit Mitteln der Mathematik noch präzisieren läßt – mit großer Wahrscheinlichkeit dem beschriebenen Ideal nahekommt, jedoch auch Verformungen zuläßt. Die Größe dieser Verformungen, in der Statistik auch *Fehler* genannt, ist dabei abhängig von der Qualität der Stichprobe und kann dementsprechend kontrolliert werden. Doch ist hier zu betonen, daß statistische Fehler, anschaulich gesprochen Abbildungsunschärfen, nie ganz zum Verschwinden gebracht werden können.

Dieser Umstand relativiert natürlich die Ergebnisse eines jeden Stichprobenverfahrens, denn er macht es notwendig, jedes Resultat im Kontext seiner Unschärfe zu sehen. Als Konsequenz dieser Sichtweise müssen manche Ergebnisse, die mit einer zu großen Unschärfe behaftet sind, sogar als unbrauchbar bewertet werden: das Abbild der Grundgesamtheit ist sozusagen an diesen Stellen zu verschwommen. Was aber ist genau Unschärfe, und wann ist sie zu groß? Die mathematische Statistik liefert hierfür einige Begriffe und Verfahren, anhand derer der etwas

schwammige Unschärfebegriff präzisierbar und vor allem quantifizierbar ist. Stichwort ist hier die *Standardabweichung* einer Stichprobe. Die Standardabweichung ist eine Zahl, die nur aus den Stichprobenwerten und dem Stichprobenumfang errechnet wird und ein Maß dafür liefert, wie stark die beobachteten Werte streuen. Offenbar deuten sehr wenig streuende Beobachtungswerte darauf hin, daß der „wahre“ Wert (also eine Zahl, die in vorgegebener Weise die tatsächliche Ausprägung eines Merkmals beschreibt) in dem von ihnen abgedeckten Bereich liegt; sehr stark streuende Beobachtungswerte legen hingegen die Vermutung nahe, daß das interessierende Merkmal in der Stichprobe nur mit sehr großen Unsicherheiten erfaßt ist, und daß das Stichprobenergebnis nur wenig aussagekräftig, im Extremfall also völlig unbrauchbar ist. Um nun nicht jedesmal zwei Zahlen – das Stichprobenergebnis und die zugehörige Standardabweichung – vergleichen zu müssen, bevorzugt man den relativen Standardfehler, der als die relative Größe der Standardabweichung im Vergleich zum Stichprobenergebnis definiert ist. Üblicherweise wird der relative Standardfehler in Prozent angegeben. Grob gesprochen bedeutet beispielsweise ein relativer Standardfehler von 13 % bei einem beobachteten Wert von 100, daß davon ausgegangen werden muß, daß der „wahre“ Wert genauso gut irgend eine andere Zahl zwischen 87 und 113 (also 13% ober- und unterhalb von 100) sein kann, wobei nicht einmal ganz ausgeschlossen ist, daß er noch weiter vom beobachteten Wert entfernt liegt. Diese ungefähre Beschreibung kann mathematisch wesentlich präziser gefaßt werden, doch soll uns dies hier nicht weiter interessieren. Die Erfahrung zeigt nun, daß Stichprobenergebnisse, die mit einem relativen Standardfehler von 15 % oder mehr behaftet sind, nicht mehr dazu geeignet sind, einigermaßen stichhaltige Erkenntnisse über die Grundgesamtheit zu erhalten. Die Wahl von 15% als maximal zulässigen relativen Standardfehler ist dabei mehr oder weniger willkürlich; ebenso kann man auch noch größere Genauigkeitsanforderungen stellen und beispielsweise nur Stichprobenergebnisse mit relativen Standardfehlern von höchstens 5% (oder noch weniger) für brauchbar erklären – offenbar muß dies im Einzelfall entschieden werden.

Die Hochrechnung des Mikrozensus erfolgt allgemein getrennt nach sogenannten *Anpassungsklassen*. Zur Definition der Anpassungsklassen verwendet man drei Abgrenzungskriterien, und zwar erstens Anpassungsschichten (bayernweit gibt es 22 davon), zweitens die Staatsangehörigkeit der Personen der Stichprobe (Deutsche/Ausländer) und drittens das Geschlecht. In Bayern ergeben sich daraus also 132 Anpassungsklassen. Innerhalb dieser Klassen ist im folgenden offensichtlich eine verlässliche Information über die Größe der Grundgesamtheit (hier also: Bevölke-

rungszahlen, genauer Zahlen der wohnberechtigten Bevölkerung) notwendig, um exakte Hochrechnungsfaktoren zu bestimmen: es ist einleuchtend, daß bei einer 1%-Stichprobe die Grundgesamtheit etwa hundertmal größer ist als die Stichprobe, doch sei daran erinnert, daß die Stichprobenziehung des Mikrozensus haushaltsbezogen erfolgt, während die Hochrechnung – wie an der Definition der Anpassungsklassen erkennbar ist – zumindest teilweise personenbezogen abläuft, so daß die zu verwendenden Hochrechnungsfaktoren keineswegs klar sind! Die somit unerläßlichen Bezugsgrößen liefert die behördliche Bevölkerungsfortschreibung, wobei als Bezugszeitraum („Stichtag“) in der Regel die letzte feiertagsfreie Woche im April des jeweiligen Jahres gilt (die „Feiertagsfreiheit“ fordert man, da in den Mikrozensus-Fragebögen nach der in der vergangenen Woche geleisteten Zahl von Arbeitsstunden gefragt wird). Allerdings werden damit die in der Behördenfortschreibung auftretenden Fehler auf den Mikrozensus übertragen. Diese Unschönheit muß jedoch in Kauf genommen werden, da keine besseren Quellen zur Verfügung stehen.

Verschiedene Kompensationsfaktoren für die Hochrechnung

Bezeichnet x die Anzahl der Stichprobenpersonen einer Kompensationsklasse, die durch die Mikrozensus-Stichprobe ermittelt wurden, und ist K der verdichtete Kompensationsfaktor dieser Kompensationsklasse, so errechnet sich die kompensierte Stichprobenbevölkerungszahl z als

$$z = K \cdot x.$$

Damit sieht man Stichprobenausfälle als kompensiert an und betrachtet die Zahl z als die tatsächlich im Mikrozensus ermittelte Stichprobenbevölkerungszahl der betreffenden Kompensationsklasse. Indem man die so erhaltenen Bevölkerungszahlen aller zu einer bestimmten Anpassungsklasse gehörenden Kompensationsklassen aufaddiert, erhält man die (kompensierte) Stichprobenbevölkerungszahl der jeweiligen Anpassungsklasse (differenziert nach Geschlecht und Staatsangehörigkeit). Mit der aus der Bevölkerungsfortschreibung bekannten Gesamtbevölkerungszahl B (die genauer die Zahl der wohnberechtigten Bevölkerung bezeichnet) der Anpassungsklasse ergibt sich der *Hochrechnungs-* bzw. *Anpassungsfaktor* A in offensichtlicher Weise als

$$A = \frac{B}{z}.$$

Offenbar ist dieser Anpassungsfaktor personenbezogen (da ja z. B. nach dem Geschlecht unterschieden wird), d. h. innerhalb eines Haushalts können ver-

schiedene Anpassungsfaktoren gelten. Um nun einerseits genügend feine Anpassungen von personenspezifischen Merkmalen durchführen zu können, aber andererseits auch in der Lage zu sein, ohne größeren Aufwand haushaltsspezifische Merkmalsdaten hochzurechnen, definiert man sowohl personen- als auch haushaltsspezifische Faktoren. Zunächst berechnet man zu jeder Person eines Haushalts aus dem für den Haushalt gültigen Kompensationsfaktor k und dem für die betreffende Person gültigen Anpassungsfaktor A einen sog. *Personenfaktor* p gemäß

$$p = A \cdot k,$$

der personenbezogen ist und daher für Anpassungen personenbezogener Merkmale, wie etwa Bevölkerungs(teil-)zahlen, Erwerbsquoten etc. benutzt wird. Einen für jeden Haushalt einheitlichen Faktor h , den sog. *Haushaltsfaktor*, berechnet man daraufhin aus den Personenfaktoren der zu dem Haushalt gehörenden Personen. Sind also

$$p_1, \dots, p_r$$

die entsprechenden Personenfaktoren, so ist der Haushaltsfaktor h definiert durch

$$h = \frac{1}{r} \sum_{i=1}^r p_i = \frac{p_1 + \dots + p_r}{r},$$

also sinnvollerweise unabhängig davon, welche Person den Haushalt repräsentiert, da alle Personen mit gleicher Gewichtung berücksichtigt sind. Dieser Haushaltsfaktor wird bei Anpassungen haushalts- oder familienspezifischer Merkmale verwendet.

Alle diese Vorbereitungen dienen dazu – das sei nochmals bemerkt –, eine Hochrechnung zu gewährleisten, die vorhandene regionale und soziale Unterschiede bei den interessierenden Merkmalen so wenig wie möglich verwischt, letztlich also genau das vermeidet, was die Statistik fälschlicherweise oft zum Schreckbild werden läßt, nämlich rücksichtslose Nivellierung.

Die tatsächliche Anpassung läuft innerhalb der schon erwähnten Anpassungsklassen ab, wobei natürlich alle in den obigen Absätzen getroffenen Unterscheidungen berücksichtigt und die jeweils zu verwendenden Haushalts- bzw. Personenfaktoren einbezogen werden (je nachdem, ob man gerade haushalts- oder personenbezogene Daten behandelt). Interessiert man sich also beispielsweise für die geschätzte, d. h. hochgerechnete Anzahl N von Haushalten innerhalb einer Anpassungsklasse, auf die ein bestimmtes (haushaltsbezogenes) Merkmal zutrifft, so geht man folgendermaßen vor: Man zählt innerhalb jedes Kompensationstyps alle Haushalte, die das interessierende Merkmal tragen, zusammen, wobei aber jeder dieser Haushalte nicht als ein Haushalt („ein“ im Sinne von „1“) gezählt wird, sondern die Haushaltsfaktoren (die geringfügig größer als 1 sind) der betreffendem Haushalte addiert werden – sozusagen eine verallgemei-

nete Form des Zählens. Dies ergibt eine etwas höhere Anzahl von Fällen als in der Stichprobe ausgewiesen, doch genau das ist ja der gewünschte Kompensationseffekt! Führt man dieses Zählverfahren für alle Kompensationstypen einer Anpassungsklasse durch, so erhält man als Summe der ermittelten Fallzahlen innerhalb der Kompensationstypen die Gesamtfallzahl y des interessierenden Merkmals für die betrachtete Anpassungsklasse. Aus dieser Zahl y und dem für die Anpassungsklasse gültigen Anpassungsfaktor A erhält man die hochgerechnete Fallzahl N dieses Merkmals, auch als Besetzungszahl des Merkmals bezeichnet, durch

$$N = A \cdot y.$$

Auch wenn hier einige Details vernachlässigt wurden, ist dies das Verfahren, mit dem sämtliche Mikrozensusergebnisse ermittelt werden und dessen Grundgedanken auch dem statistischen Laien zugänglich sein sollten. Etwas schwieriger wird es bei der Fehlerrechnung, da hier einige Argumente und Formeln nicht mehr ohne weiteres motiviert werden können.

Eine grundsätzliche Überlegung zu Schätzfehlern

Unabhängig von bestimmten Merkmalsarten, also etwa haushalts-, familien- oder personenbezogenen Merkmalen, liefert die mathematische Statistik Formeln, um bei Stichprobenverfahren mit geschichteter Zufallsauswahl den Fehler, mit dem ein Ergebnis behaftet ist, abzuschätzen. Die genaue Angabe dieser Formeln ist hier allerdings wenig sinnvoll, da sie für den Laien nur schlecht motivierbar sind und ihr Verständnis einiges an mathematischem Vorwissen voraussetzt. Trotzdem können auch hier einige plausible Überlegungen angestellt werden.

Gehen wir zunächst davon aus, daß wir uns für ein Merkmal interessieren, von dem wir erwarten, daß es sehr oft auftreten wird, z.B. das Merkmal „Person ist weiblich“, das in über der Hälfte aller Fälle anzutreffen ist. Wir stellen uns vor, daß dieser Anteil sich in der Stichprobe entsprechend wiederfinden wird, daß also auch in der Stichprobe etwa zur Hälfte (oder etwas mehr) Frauen enthalten sein werden. Nun wurde schon angesprochen, daß die Vorstellung, die Stichprobe stelle ein proportional verkleinertes Bild der Grundgesamtheit dar, insofern eingeschränkt werden muß, als diese „Verkleinerung“ als eine Abbildung mit stochastischen Unschärfen zu verstehen ist. Halten wir uns aber an das Merkmal „Person ist weiblich“, das bundesweit in ca. 40 Millionen Fällen zutrifft, in der Stichprobe also ungefähr 400.000 mal vorkommen

wird, so bedeutet diese stochastische Unschärfe, daß eben nicht genau jede hundertste Frau in die Stichprobe ausgewählt wird, sondern nur mit zufälligen Streuungen in etwa jede hundertste. Ob jedoch bei einem ca. 400.000 mal in der Stichprobe vorhandenen Merkmal ein Fall mehr oder weniger ausgewählt wird, fällt offensichtlich kaum ins Gewicht. Das bedeutet jedoch, daß die Schätzung des weiblichen Bevölkerungsanteils relativ „scharf“ durchgeführt werden kann, daß also hier keine allzu gravierenden statistischen Fehler zu erwarten sind.

Ganz anders sieht es aber aus, wenn ein seltenes Merkmal wie z. B. „Person ist alleinerziehender Vater“ betrachtet wird, noch dazu, wenn man sich auf Teile der Grundgesamtheit wie z. B. eine Anpassungsschicht zurückzieht. Nimmt man etwa an, daß das genannte Merkmal in einer bestimmten Anpassungsschicht 2000 mal vorkommt, so sind für die Stichprobe in dieser Anpassungsschicht ungefähr 20 Fälle dieses Merkmals zu erwarten. Ob nun aber durch die Zufallsauswahl ein oder zwei Personen dieses Merkmals mehr oder weniger in die Stichprobe aufgenommen werden, bewirkt bei der Hochrechnung offenbar sehr unterschiedliche Ergebnisse. Diese Überlegung macht klar, daß das Auftreten seltener Merkmale innerhalb kleiner Teile der Grundgesamtheit nur sehr unsicher geschätzt werden kann und verdeutlicht, an welche Grenzen man etwa bei der Regionalisierung von Mikrozensusergebnissen stößt.

Als wichtige Tatsache, die auch an den entsprechenden mathematischen Fehlerformeln abgelesen werden kann, ist also erstens festzuhalten, daß nicht alle interessierenden Merkmale bei Stichprobenverfahren wie dem Mikrozensus gleichmäßig genau geschätzt werden können (Faustregel: Je häufiger ein Merkmal, desto besser seine Schätzung), und daß darüber hinaus nicht etwa der relative Anteil eines Merkmals (also die Größe, die man zunächst mit den Attributen „selten“ und „häufig“ in Verbindung bringt) an der betrachteten Grundgesamtheit entscheidend für seine Schätzbarkeit ist, sondern (noch schlimmer!) die absolute Fallzahl des Merkmals in der Grundgesamtheit. Dies bedeutet speziell für den Mikrozensus, daß mit abnehmender Größe der betrachteten regionalen oder sozialen Einheiten immer weniger Merkmale sinnvoll geschätzt werden können, für sehr kleine Einheiten wie regionale Schichten also nur noch Eckdaten wie Bevölkerungszahlen, Erwerbsquoten etc.

Übrigens motivieren dieselben mathematischen Formeln die im Abschnitt über Schichten aufgestellte Forderung, Schichten so zu wählen, daß jede Schicht intern möglichst gleichmäßig strukturiert ist. Anders ausgedrückt: Die Güte des „Schichten-Designs“ geht direkt in die Fehlerrechnung ein und ist somit entscheidend für die Brauchbarkeit von Stichprobenergebnissen.

Alternative Abschätzung des Fehlers

Neben der exakten, „klassischen“ Formel zur Abschätzung des relativen Standardfehlers, in die alle Einzelergebnisse eines Merkmals aus den verschiedenen Schichten eingehen, gibt es noch andere Verfahren, den Fehler bei der Schätzung eines Merkmals zumindest ungefähr einzugrenzen. Ein Verfahren, das hier angewendet werden soll, benötigt nur die absoluten Besetzungszahlen eines Merkmals in der betrachteten Grundgesamtheit, die allerdings noch in Relation zum Umfang der Grundgesamtheit gesetzt wird. Die von uns später betrachtete Grundgesamtheit wird das Stadtgebiet München sein, das, wie schon gesagt, seit 1991 eine eigene Anpassungsschicht bildet, so daß tatsächlich münchenspezifische Mikrozensus-Ergebnisse vorliegen.

Ohne nun wieder auf den genaueren mathematischen Hintergrund einzugehen, wenden wir nur die resultierenden Regeln zur Fehlerschätzung an. Dazu müssen wir noch einige Begriffe einführen. Bei der eingehenderen Untersuchung der Hochrechnungsfehler beim Mikrozensus zeigt es sich, daß man drei *Merkmalskategorien* einteilen kann, die jeweils Merkmale mit näherungsweise gleichem Fehlerverhalten enthalten und zusammen alle Merkmale erfassen. Die erste Merkmalskategorie enthält personenbezogene Merkmale, die Bevölkerungsgruppen oder Gruppen von Erwerbstätigen auszeichnen, aber nicht speziell Gruppen von Ausländern oder Gruppen von Erwerbstätigen in Land- und Forstwirtschaft. Diese Kategorie wird *B/E* genannt (Bevölkerung/Erwerbstätige). Passend dazu bilden die Merkmale, die Gruppen von Ausländern oder Erwerbstätigen in Land- und Forstwirtschaft abgrenzen, eine eigene Merkmalskategorie *A/L* (Ausländer/Landwirtschaft). Damit sind alle personenbezogenen Merkmale klassifiziert, und so erfaßt die letzte Kategorie *H* (Haushalte) alle haushaltsbezogenen Merkmale.

Der Grund für die Sonderbehandlung von Ausländern und Erwerbstätigen in Land- und Forstwirtschaft ist, daß man bei diesen Gruppen oftmals andere Merkmalsstrukturen erwartet als beim Rest der Bevölkerung. Da aber beide Gruppen Minderheiten bilden, können hier offenbar nur relativ niedrige Fallzahlen auftreten, was ein schlechteres Fehlerverhalten impliziert.

Mit geeigneten mathematischen Methoden lassen sich für jede der drei Merkmalskategorien und jede gewünschte Ergebnisgenauigkeit (also eine Obergrenze für den relativen Standardfehler) Mindestfallzahlen angeben, mit denen entschieden werden kann, ob über ein bestimmtes Merkmal sinnvolle Angaben gemacht werden können. Bei bundesweit

betrachteten Ergebnissen zeigt sich beispielsweise, daß in allen drei Merkmalskategorien mindestens 12.000 Fälle (hochgerechnet) eines bestimmten Merkmals notwendig sind, um dieses Merkmal mit einem relativen Standardfehler von nicht mehr als 10% schätzen zu können. Legt man dagegen 2% als Schranke für den relativen Standardfehler fest, so sind für Merkmale der Kategorien B/E und H mindestens 300.000 Fälle notwendig, während in der Kategorie A/L erst 400.000 Fälle für ein entsprechend genaues Ergebnis sorgen. Fallzahlen von 400.000 und mehr werden allerdings nur sehr wenige (oder gar keine) Merkmale dieser Kategorie erreichen, so daß unter diesen Genauigkeitsanforderungen über viele Merkmale schlicht und einfach keine Aussage gemacht werden kann.

Prinzipiell genauso kann man für eine einzelne Anpassungsschicht vorgehen, in unserem Fall die Anpassungsschicht 3 in Bayern, nämlich das Stadtgebiet München. Bei der jetzigen Form des Mikrozensus ist man jedoch gezwungen, bundesdurchschnittliche Strukturen zugrunde zu legen, so daß keine Fehlerschätzung im exakten Sinne, immerhin aber eine größenordnungsmäßige Eingrenzung des Fehlers durchgeführt werden kann. Geht man also beispielsweise für die Mikrozensusergebnisse 1991 von einer Bevölkerungszahl von 1,234 Millionen Personen (nur Hauptwohnsitzbevölkerung) und einer Gesamtzahl von 717.000 Haushalten in München aus, so ergeben sich zu verschiedenen Fehlerschranken die folgenden Mindestbesetzungszahlen (hochgerechnet):

Fehlerschranken für Mikrozensusergebnisse in München

Fehlerschranke	Mindestbesetzungszahl (in 1000)	
	Merkmalskategorie B/E	Merkmalskategorie H
10%	14	13
9%	17	16
8%	21	20
7%	28	27
6%	38	37
5%	56	54
4%	91	86
3%	176	162
2%	469	355
1%	1026	616

Zu der Tabelle ist zu bemerken, daß die angegebenen Besetzungszahlen für die Kategorie B/E als Personenzahlen zu verstehen sind, während in der Kategorie H von Haushaltzahlen die Rede ist. Außerdem sollten die Schranken „cum grano salis“ verstanden werden, da die ihnen zugrundeliegenden Berechnungen (wie bereits gesagt) nur näherungsweise Abschätzungen des relativen Standardfehlers erlauben. Bemerkenswert ist, daß für Merkmale der Kategorie A/L, also Merkmale, die (zumindest in München) vor allem die ausländische Bevölkerung betreffen, innerhalb einer Anpassungsschicht wie München die für die beiden anderen Merkmalskategorien verwendete Fehlerabschätzung keine sinnvollen Ergebnisse liefert, so daß Ergebnisse zu Merkmalen der Kategorie A/L im allgemeinen als unbrauchbar gelten müssen, da sie entweder mit zu großen Fehlern behaftet sind bzw. über die Größenordnung dieses Fehlers nichts ausgesagt werden kann.

Eine Auswahl der Mikrozensus-Ergebnisse für München 1991

Bevor man sich daran macht, Mikrozensusergebnisse für München zu verwenden, sollte man sich darüber Gedanken machen, welche Konsequenzen aus der obigen Fehlerabschätzung zu ziehen sind, d. h. also zum Beispiel, welche Fehlerschranken man anlegen will, um brauchbare von unbrauchbaren Ergebnissen zu unterscheiden. Wir haben schon gesehen, daß sehr große Besetzungszahlen vonnöten sind, um in einer regionalen Schicht wie München einen einigermaßen zuverlässigen Eindruck von der Häufigkeit eines Merkmals zu erhalten, und diese Einschränkung wird um so gravierender, je strengere Genauigkeitsforderungen man stellt. Speziell für die Zwecke der Städtestatistik, die sich in den Bereichen, die auch vom

Mikrozensus abgedeckt werden und die daher für kritische Zahlenvergleiche, oder wo möglich, gegenseitige Ergänzung prädestiniert sind, weitestgehend auf behördliche Fortschreibungen oder ähnliche Register stützt, die eine ziemlich genaue Darstellung der „Wirklichkeit“ versprechen, sollten diese Genauigkeitsforderungen aus Konsistenzgründen aber besonders rigoros ausfallen. Denn es ist völlig sinnlos, hinreichend sicheren Daten (wie beispielsweise Bevölkerungsdaten aus dem Meldewesen) Daten mit hohen relativen Standardfehlern gegenüberzustellen, da ja, wenn eine Gegenüberstellung solch verschiedener Zahlen durchgeführt wird, die Beziehung zwischen den Zahlen interessiert, also z.B. die Differenz oder

das Verhältnis der beiden Zahlen. Diese Größen erben aber grob gesprochen mindestens die statistische Unsicherheit der mit dem größten Fehler behafteten Ausgangsgröße, werden also im allgemeinen noch unschärfer.

Eine Betrachtung der Besetzungszahlen in den Mikrozensus-Tabellen für München zusammen mit den groben Fehlerabschätzungen für jede dieser Zahlen zeigt nun aber, daß nur sehr wenige von ihnen einer Fehlerschranke von 3% bzw. sogar 2% standhalten. Dies sei einmal an folgender Tabelle aus dem Mikrozensus 1991 illustriert, in der die Münchener Privathaushalte nach dem Alter der jeweiligen Bezugsperson aufgeschlüsselt sind:

Privathaushalte nach Alter der Bezugsperson

Alter der Bezugsperson		Männlich	Weiblich	Insgesamt
		(in 1000 / RSF)	(in 1000 / RSF)	(in 1000 / RSF)
		(RSF = Relativer Standardfehler)		
jünger als 25 J.	insgesamt	25 / 8%	26 / 8%	52 / 6%
	dar. 1-Personen-HH	20 / 8%	22 / 8%	42 / 6%
von 25 b.u. 45 J.	insgesamt	189 / 3%	96 / 4%	285 / 3%
	dar. 1-Personen-HH	86 / 4%	67 / 5%	153 / 4%
von 45 b.u. 65 J.	insgesamt	169 / 3%	66 / 5%	235 / 3%
	dar. 1-Personen-HH	51 / 6%	45 / 6%	96 / 4%
65 J. und älter	insgesamt	66 / 5%	80 / 5%	146 / 4%
	dar. 1-Personen-HH	13 / 10%	74 / 5%	88 / 4%
Haushalte insgesamt		449 / 2%	268 / 3%	717 / 0

Nur sechs der insgesamt 27 aufgeführten Merkmale (in der Tabelle schattiert) sind hier mit einem relativen Standardfehler von 3% oder weniger schätzbar, die 2%-Fehlerschranke unterschreitet sogar nur ein einziges Merkmal, nämlich die Gesamtheit der Münchener Haushalte mit männlicher Bezugsperson (am Hauptwohnsitz; denn alle dem Autor vorliegenden Daten beziehen sich nur auf die Bevölkerung am Hauptwohnsitz). Obwohl also die im Mikrozensus erfragten Merkmale ein weites Spektrum abdecken (von Bevölkerungs- und Erwerbstätigenzahlen über die finanzielle und soziale Situation der Haushalte bis hin zu speziellen Merkmalen wie den Versicherungsarten der Arbeitnehmer oder der Bildungsstruktur der Erwerbstätigen etc.), sind auf der Ebene einer – bundesweit gesehen – so kleinen Grundgesamtheit wie München kaum brauchbare Ergebnisse zu erhoffen –

ein aus Sicht der Städtestatistik eher ernüchterndes Ergebnis, das jedoch andererseits zeigt, daß bundesweite Stichprobenerhebungen (also alle Erhebungen außer Volkszählungen) die kleinräumliche Statistik, wie sie von den statistischen Ämtern der Großstädte betrieben wird, nicht ersetzen können. Nichtsdestotrotz hat der Mikrozensus seine Berechtigung als statistische Erhebung, die dazu dient, im Zeitraum zwischen zwei Volkszählungen über demographische und andere großräumige und langfristige Entwicklungen auf dem laufenden zu bleiben.

Um den geduldigen Leser angesichts der doch sehr stark eingeschränkten Verwendbarkeit der Mikrozensus-Ergebnisse für München jedoch nicht vollends zu enttäuschen, sei zum Abschluß die „bunte Liste“ der Merkmale aufgeführt, die nach der erwähnten Fehlerabschätzung als hinreichend sicher gelten können,

wobei „hinreichend sicher“ hier bedeutet, daß die betreffenden Merkmale mit einem relativen Standardfehler von 2% oder weniger behaftet sind. In den Besetzungszahlen der personenspezifischen Merkmale ist dabei nur die Hauptwohnsitzbevölkerung

berücksichtigt, in den Besetzungszahlen der haushaltsspezifischen Merkmale sind ausschließlich Haushalte gezählt, die als Hauptwohnsitz ihrer Bezugsperson gelten.

Ausgewählte Mikrozensusergebnisse für München

Merkmal	Besetzungszahl (in 1000)
Bevölkerung insgesamt	1234
Männliche Bevölkerung	593
Weibliche Bevölkerung	641
Bevölkerung jünger als 45 Jahre	699
Bevölkerung im Alter von 15 bis unter 45 Jahren	551
Bevölkerung im Alter von 15 bis unter 65 Jahren	892
Männliche Bevölkerung im Alter von 15 bis unter 65 Jahren	452
Ledige/Verwitwete/Geschiedene jünger als 45 Jahre	476
Bevölkerung mit Lehre als berufsbildendem Abschluß	482
Bevölkerung mit überwiegendem Lebensunterhalt durch Erwerbstätigkeit	628
Erwerbspersonen	696
Erwerbspersonen im Alter von 15 bis unter 65 Jahren	688
Erwerbstätige	667
Nichterwerbspersonen	538
Nichterwerbstätige	566
Privathaushalte insgesamt	717
1-Personen-Privathaushalte	378
Privathaushalte mit männlicher Bezugsperson	449
Privathaushalte mit erwerbstätiger Bezugsperson	481
Privathaushalte ohne Kinder	553

Thomas Klein (stud. math.)